Solutions to quantum weak coin flipping*

Atul Singh Arora†1, Jérémie Roland‡2, Chrysoula Vlachou§3, and Stephan Weis¶4

[1]Institute for Quantum Information and Matter and Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, California, USA
[2]Université libre de Bruxelles, Brussels, Belgium
[3] Instituto de Telecomunicações Lisbon and Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Lisbon Portugal
[4] Wald-Gymnasium, Berlin, Germany

August 25, 2022

Abstract

Weak coin flipping is an important cryptographic primitive, as it is the strongest known secure two-party computation primitive, that classically becomes secure only when certain assumptions are made (e.g. computational hardness), while quantumly there exist protocols that achieve arbitrarily close to perfect security. This breakthrough result was established by C. Mochon in 2007 [arXiv:0711.4114], however, his proof of existence was partially non-constructive, thus, setting back the proposal of explicit protocols. In this work, we report three different solutions to the quantum weak coin flipping problem. In particular, we propose different methods that result—either analytically or numerically—in the operators needed to construct weak coin flipping protocols with different levels of security, including nearly perfect security. In order to develop these methods, we study the quantum weak coin flipping problem from both an algebraic and a geometric perspective. We also analytically construct illustrative examples of weak coin flipping protocols achieving different levels of security.

*Parts of the original work presented in this paper have been presented in the STOC19 (https://dl.acm.org/doi/10.1145/3313276.3316306) and SODA21 (https://epubs.siam.org/doi/10.1137/1.9781611976465.58) conferences.

†atul.singh.arora@gmail.com, asarora@caltech.edu

‡Jeremie.Roland@ulb.be

§chrysoula.vlachou@lx.it.pt

¶weis@waldgymnasium.de

1 Introduction

The problem we study in this paper is rather easy to state. Suppose there are two parties, conventionally called Alice and Bob, who are placed in physically remote locations and can communicate with each other using a communication channel. They wish to exchange messages over this channel in order to agree on a random bit, while having *a priori* known opposite preferred outcomes. This is easy to do—Alice flips a coin and sends a message with the outcome to Bob. However, this requires Bob to trust Alice. Can Bob modify the scheme to be sure that Alice did not cheat? More generally, can we construct a protocol, which involves an exchange of messages over a communication channel, to decide on a random bit while ensuring that an honest party, i.e. one that follows the protocol, can not be deceived? It turns out that if one communicates over a classical communication channel, then a cheating party can always force their desired outcome on the honest party; unless we make further assumptions, such as computational hardness. On the other hand, if Alice and Bob use a quantum communication channel, then protocols solving this problem up to vanishing errors have been shown to *exist* [Moc07]. This was a seminal result from 2007, however there is a non-constructive part in this analysis, which implies that we know there exists a solution without knowing the solution itself. In this paper, we build upon the previous pioneering works to develop explicit protocols for *quantum weak coin flipping*, as this problem is referred to in the literature.

The coin flipping problem, since 1983 when it was introduced by M. Blum [Blu83], occupies an interesting place in the overall landscape of cryptography. In 1994 it was shown that the—even today—widely used public key cryptosystem RSA can be broken using a quantum computer [Sho94]. A decade earlier, a method for key distribution using quantum channels was proposed [BB84] whose security, in principle, relied only on the validity of the laws of physics. It was thought that quantum mechanics could also revolutionize *secure two-party computation*. This is another branch of cryptography comprising protocols in which two distrustful parties wish to jointly compute a function on their inputs without having to reveal these inputs to each other. Success here, was marred by a cascade of impossibility results. In a central result of classical cryptography, it was shown that a primitive called *oblivious transfer* is universal for secure two-party computation [Kil88], but there exists no protocol that offers perfect security without relying on further assumptions, such as computational hardness; classical secure two-party computation with perfect security is thus impossible [Col07]. Oblivious transfer deprived quantum mechanics of being the panacea for cryptography, as it was shown that even if the communication is quantum, oblivious transfer can not be implemented with perfect security [Lo97; CKS13]. *Bit commitment*, a secure two-party computation primitive weaker than oblivious transfer was targeted, but it turned out to be also impossible—in the same sense—even in the quantum setting [CK11]. Thus, *coin flipping* was considered, an even weaker secure two-party computation primitive, which has two variants: *strong* and *weak* coin flipping (WCF). In a coin flipping protocol the two distrustful parties need to establish a shared random bit; for strong coin flipping the preferences of the parties are unknown to each other, in contrast to WCF where the parties have a priori known opposite preferences. And while strong coin flipping suffered the same fate as that of oblivious transfer and bit commitment [CK09], WCF was poised for fame; it is the strongest known primitive in the two-party setting which admits no secure classical protocol, but can be implemented over a quantum channel with near perfect security [Moc07].

In particular, in a quantum strong coin flipping protocol a dishonest party can successfully cheat with probability at least $\frac{1}{\sqrt{2}}$ [Kit03], and the best known explicit protocol[1] has a cheating probability of $\frac{1}{2}+\frac{1}{4}$ [Amb04]. As for WCF, the existence of protocols with arbitrarily perfect security was proved non-constructively, by elaborate successive reductions of the problem based on the formalism introduced earlier by A. Y. Kitaev for the study of strong coin flipping [Kit03]. Consequently, the structure of the protocols

[1] A strong coin flipping protocol with the minimum cheating probability is known [CK09], but relies on the use of near perfect WCF as a black box.

whose existence is proved was lost. A systematic verification led to a simplified proof of existence by D. Aharonov et al. [Aha+14b], but over a decade later an explicit, nearly perfectly secure WCF protocol was missing, despite various approaches ranging from the distillation of a protocol using the proof of existence to numerical search [NST14; NST15]. While an explicit WCF protocol has remained evasive, several connections have been discovered. In particular, nearly perfect WCF provides, via black-box reductions, optimal protocols for strong coin flipping [CK09], bit commitment [CK11] and a variant of oblivious transfer [CGS13]. It is also used to implement other cryptographic tasks such as leader election [Gan09] and dice rolling [AS10].

The most significant advance in the study of WCF was the invention of the so-called point games, attributed to A. Y. Kitaev by C. Mochon [Moc07]. In this context, there are three equivalent formalisms which can be used to describe WCF protocols and their security properties: explicit protocols given by pairs of dual semi-definite programs (SDPs), Time Dependent Point Games (TDPGs) and Time Independent Point Games (TIPGs). The existence of quantum WCF protocols with almost perfect security was established using the TIPG formalism [Moc07], however the proposal of explicit protocols was hindered by the fact that no constructive method was given for obtaining a protocol from a TDPG, while they were shown to be equivalent. In this work, we start by constructing a new framework that allows us to convert point games into protocols granted that we can find unitaries satisfying certain constraints; we further use perturbative methods to obtain a protocol with cheating probability $\frac{1}{2}+\frac{1}{10}$, improving the former best known protocol with cheating probability $\frac{1}{2}+\frac{1}{6}$ [Moc05].[2] We then introduce a more systematic method for converting the point games used by C. Mochon (including the ones approaching perfect security) into explicit unitaries, which in turn can be readily converted into explicit WCF protocols. This approach also circumvents part of the previous formalism, thus leading to a simplification of the overall context. However, it is tailored for the aforementioned point games and it is not expected a priori to work in general. To address this, we develop a numerical algorithm that allows us to provably perform the non-constructive step in the aforementioned conversion of a TDPG into an explicit protocol. This, in effect, permits us to numerically construct WCF protocols corresponding to *any* TIPG. Finally, we give another analytic solution to the point games employed by C. Mochon, which is inspired by the techniques used in the numerical solution and it is exact; it is not affected by the algorithm's numerical accuracy.

Below, we briefly introduce the various formalisms[3] in order to be able to informally describe and summarize our contributions in Section 1.1. In Section 2, we present these formalisms in more detail, as we need to build on these results afterwards.

Let us start with two features of WCF. First, without loss of generality, we can say that, if the outcome value of the bit is 0 it means that Alice won, while Bob wins on outcome 1; since in a WCF protocol the parties have opposite known preferences this is just a matter of labeling. Second, there are four situations which can arise in a WCF scenario, of which three are of interest. Let us denote by HH the situation where both Alice and Bob are honest, i.e. they follow the protocol. We want the protocol to be such that both Alice and Bob (a) win with equal probability and (b) are in agreement with each other. In the situation HC where Alice is honest and Bob is cheating, the protocol must protect Alice from a cheating Bob, who tries to convince her that he has won. His probability of succeeding by using his best cheating strategy is denoted by P_B^*, where the subscript denotes the cheating party. The CH situation where Bob is honest and Alice is cheating naturally points us to the corresponding definition of P_A^*. The situation CC where both players are cheating is not of interest to us as nothing can be said with respect to the protocol; neither party is actually following it.

The trivial example of a WCF protocol is where Alice flips a coin and reveals the outcome to Bob over the telephone. A cheating Alice can simply lie and always win against an honest Bob; that means $P_A^* = 1$.

[2] Strictly speaking, these are families of protocols whose cheating probability approaches the said value asymptotically.

[3] We have suppressed the technical details.

On the other hand, a cheating Bob can not do anything to convince Alice that he has won, unless it happens by random chance on the coin flip. This corresponds to $P_B^* = \frac{1}{2}$. We say that a protocol has *bias* ϵ if neither party can force their preferred outcome with probability greater than $1/2 + \epsilon$, for $\epsilon \geq 0$. For the above naive protocol the bias is $\epsilon = \max[P_A^*, P_B^*] - \frac{1}{2}$, which amounts to $\epsilon = \frac{1}{2}$; the worst possible. Constructing protocols where one of the parties is protected is nearly trivial; constructing protocols where neither party is able to cheat against an honest party is the real challenge.

Given a WCF protocol it is not a priori clear how the maximum success probability of a cheating party, $P_{A/B}^*$, should be computed, as their strategy space can be dauntingly large. It turns out that all quantum WCF protocols can be defined using the exchange of a message register interleaved with the parties applying the unitaries U_i locally (see Figure 1) until a final measurement— say Π_A denoting Alice won and Π_B denoting Bob won—is made in the end. Computing P_A^* in this case reduces to a semi-definite program (SDP) in ρ, the corresponding quantum state: maximize $P_A^* = \text{tr}(\Pi_A \rho)$ given the constraint that the honest party follows the protocol. Similarly for computing P_B^* we can define another SDP. Using SDP duality one can turn this maximization problem over cheating strategies into a minimization problem over dual variables $Z_{A/B}$. Any dual feasible assignment then provides an upper bound on the cheating probabilities $P_{A/B}^*$. Handling SDPs is, in general, straightforward, but in this case, there are two SDPs, and we must optimize both simultaneously. Note that we assume that the protocol is known and we are trying to bound P_A^* and P_B^*. However, our goal is to find good protocols. Therefore, we would like a formalism which allows us to do both, construct protocols *and* find the associated P_A^* and P_B^*. A. Y. Kitaev gave us such a formalism.

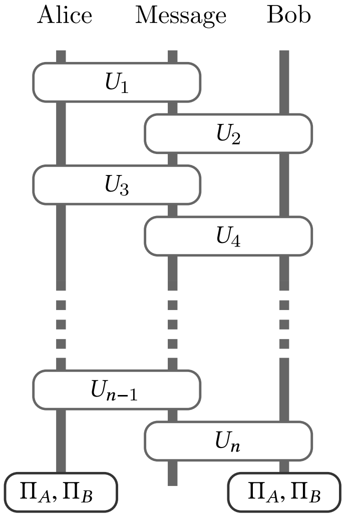


Figure 1: General structure of a WCF protocol.

He converted this problem about matrices (Z, ρ and U) into a problem about points on a plane, and C. Mochon called it "Kitaev's Time Dependent Point Game formalism" (TDPG). Therein, we are concerned with a sequence of frames, also referred to as configurations. Each frame is a finite collection of points in the positive quadrant of the xy-plane with probability weights assigned to them. This sequence must start with a fixed frame and end with a frame that has only one point. The fixed starting frame consists of two points at $(0, 1)$ and $(1, 0)$ with equal weights $1/2$. The end frame must be a single point, say at (β, α), with weight 1. The objective of the protocol designer is to get this end point as close to the point $(\frac{1}{2}, \frac{1}{2})$ as possible by transitioning through intermediate frames (see Figure 2) following certain rules. The magic

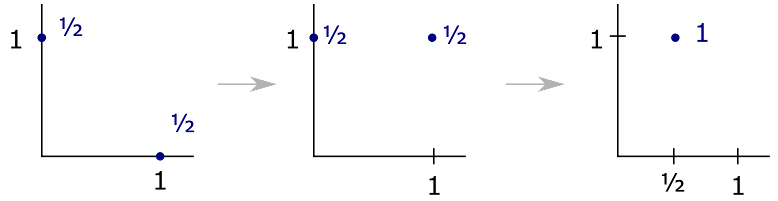


Figure 2: Point game.

of this formalism, roughly stated, is that if one abides by these rules, then corresponding to every such sequence of frames, there exists a WCF protocol with $P_A^* = \alpha$, $P_B^* = \beta$.

Let us now describe these rules. Consider a given frame and focus on a set of points that fall along a vertical (or horizontal) line. Let the y (or x) coordinate of the ith point be given by z_{g_i} and its weight by p_{g_i}, and let z_{h_i} and p_{h_i} denote the corresponding quantities for the points in the subsequent frame. Then, the following conditions must hold:

1. the probabilities are conserved, viz. $\sum_i p_{g_i} = \sum_i p_{h_i}$
2. for all $\lambda > 0$

$$\sum_i \frac{\lambda z_{g_i}}{\lambda + z_{g_i}} p_{g_i} \leq \sum_i \frac{\lambda z_{h_i}}{\lambda + z_{h_i}} p_{h_i}. \tag{1}$$

From one frame to the next, we can either make a horizontal or a vertical transition. By combining these sequentially we can obtain the desired form of the final frame, i.e. a single point. The points in the frames and the rules of the transitions arise from the variables $Z_{A/B}$ of the dual SDP and their constraints, respectively. Just as the state ρ evolves through the protocol, so do the dual variables $Z_{A/B}$. The points and their weights in the TDPG are exactly the eigenvalue pairs of $Z_{A/B}$ with the probability weight assigned to them by the honest state $|\psi\rangle$ at a given point in the protocol. Given an explicit WCF protocol and a feasible assignment for the dual variables witnessing a given bias, it is straightforward to construct the TDPG. However, going backwards, constructing the WCF dual from a TDPG is non-trivial and no general construction is known.

Before proceeding, it is useful to encode the points on a line and their weights into a function from the interval $[0, \infty)$ to itself. Let

$$\llbracket a \rrbracket (z) = \delta_{a,z}, \tag{2}$$

i.e. $\llbracket a \rrbracket (z)$ is zero when $z \neq a$ and one when $z = a$. The *transition* from a given frame to the next is written as $\sum_i p_{g_i} \llbracket z_{g_i} \rrbracket \to \sum_i p_{h_i} \llbracket z_{h_i} \rrbracket$. The corresponding *function* is written as $t = \sum_i p_{h_i} \llbracket z_{h_i} \rrbracket - \sum_i p_{g_i} \llbracket z_{g_i} \rrbracket$. If the transition/function satisfies the conditions 1. and 2. above, it is termed as a *valid transition/function*.

If we restrict ourselves to transitions involving only one initial and one final point, i.e. $\llbracket z_g \rrbracket \to \llbracket z_h \rrbracket$, the second condition reduces to $z_g \leq z_h$. This is called a *raise*, and it means that we can always increase the coordinate of a single point. What about going from one initial point to many final points, i.e. $\llbracket z_g \rrbracket \to \sum_i p_{h_i} \llbracket z_{h_i} \rrbracket$? Note that the points before and after must lie along either a horizontal or a vertical line. The second condition in this case becomes $1/z_g \geq \langle 1/z_h \rangle$, which means that the harmonic mean of the final points must be greater than or equal to that of the initial point, where $\langle f(z_h) \rangle := \left(\sum_i f(z_{h_i}) p_{h_i} \right) / \left(\sum_j p_{h_j} \right)$. This is called a *split*. Finally, we can ask what happens upon merging many points into a single point, i.e. $\sum_i p_{g_i} \llbracket z_{g_i} \rrbracket \to \llbracket z_h \rrbracket$. The second condition becomes $\langle z_g \rangle \leq z_h$, which means that the final position must not be smaller than the average initial position. This is called a *merge*. While these three valid transitions do not exhaust the set of possible valid moves, they are enough to construct games approaching bias 1/6.

Let us consider a simple game as an example (see Figure 2). We start with the initial frame and raise the point $(1, 0)$ vertically to $(1, 1)$; this is a raise, an allowed move. Next we merge the points $(0, 1)$ and

$(1, 1)$ using a horizontal merge. The x-coordinate of the resulting point can at best be $\frac{1}{2}.0 + \frac{1}{2}.1 = \frac{1}{2}$ where we used the fact that both points have weight $1/2$. Thus, we end up with a single point having all the weight at $(\frac{1}{2}, 1)$. This formalism tells us that there must exist a protocol which yields $P_A^* = 1$ while $P_B^* = \frac{1}{2}$, which is exactly the telephone protocol that we presented earlier. It is a neat consistency check but it yields the worst possible bias. This is because we did not use the split move. If we use a split once, we can, by essentially matching the weights, already obtain a game with $P_A^* = P_B^* = \frac{1}{\sqrt{2}}$. Various protocols corresponding to this bias were found [SR02; NS03; KN04] before the point game formalism was known. In fact, this bias, $\epsilon = \frac{1}{\sqrt{2}} - \frac{1}{2}$, is exactly the lower bound for the bias of *strong* coin flipping protocols. It was an exciting time —we imagine—as the technique used to bound strong coin flipping failed for WCF. The matter was not resolved, and this protocol remained the best known WCF protocol for some time; until C. Mochon showed that using multiple splits at the beginning followed by a raise, and thereafter simply using merges, we can obtain a game with bias almost $1/6$ [Moc05]. Obtaining lower biases, however, is not a straightforward extension of the above, and we need other moves which can not be decomposed into the three basic ones: splits, merges and raises.

1.1 Contributions

1.1.1 TDPG-to-Explicit-protocol Framework (TEF) and a protocol approaching bias 1/10

In Section 3 we provide a framework for converting a TDPG into an explicit WCF protocol. We start by defining a "canonical form" for any given frame of a TDPG, which allows us to write the WCF dual variables, Zs, and the honest state $|\psi\rangle$ associated with each frame of the TDPG. We then define a sequence of quantum operations, unitaries and projections, which describe how Alice and Bob transition from the initial to the final frame. It turns out that there is only one non-trivial quantum operation, U, in the sequence. Using the SDP formalism we write the constraints at each step of the sequence on the Zs and show that they are indeed satisfied. The aforementioned constraints can be summarized as in Theorem 1 below. In Section 3 one can find the full version, Theorem 31, together with its proof and a detailed description of the framework.

Theorem 1 (TEF constraint (simplified))**.** *If a unitary matrix U acting on the space span$\{|g_1\rangle, |g_2\rangle \dots, |h_1\rangle, |h_2\rangle \dots\}$ satisfying the constraints*[4]

$$U|v\rangle = |w\rangle,$$
$$\sum_i x_{h_i} |h_i\rangle\langle h_i| - \sum_i x_{g_i} E_h U |g_i\rangle\langle g_i| U^\dagger E_h \geq 0 \tag{3}$$

can be found for every transition (see Definition 14 and Definition 15) of a TDPG, then an explicit protocol with the corresponding bias can be obtained using the TDPG–to–Explicit–protocol Framework (TEF). Here, $\{|g_i\rangle\}$, $\{|h_i\rangle\}$ are orthonormal vectors. If the transition is horizontal, then

- *the initial points have x_{g_i} as their x-coordinate and p_{g_i} as their corresponding probability weight,*
- *the final points have x_{h_i} as their x-coordinate and p_{h_i} as their corresponding probability weight,*
- *E_h is a projection onto the span $\{|h_i\rangle\}$ space,*
- *$|v\rangle = \sum_i \sqrt{p_{g_i}} |g_i\rangle / \sqrt{\sum p_{g_i}}$, $|w\rangle = \sum_i \sqrt{p_{h_i}} |h_i\rangle / \sqrt{\sum p_{h_i}}$.*

If the transition is vertical, the x_{g_i} and x_{h_i} become the y-coordinates y_{g_i} and y_{h_i} with everything else unchanged.

[4] We use $A \geq B$ to mean that $A - B$ has non-negative eigenvalues; we implicitly assume that A and B are Hermitian.

The TDPG already specifies the coordinates x_{h_i}, x_{g_i} and the probabilities p_{h_i}, p_{g_i} satisfying the scalar condition Equation (1), therefore our task reduces to finding the correct U which satisfies the matrix constraints Equation (3). Given such a unitary U we show in detail how we can progressively build the sequence of unitaries corresponding to the complete WCF protocol. In fact, we need to reverse the order of the operations in the sequence we get in order to obtain the final protocol. Note that the message register is initially decoupled, then it gets entangled, and finally emerges decoupled again. This simplifies the analysis, and also entails that we don't need to keep it coherent for the whole duration of the protocol. Keeping the message register coherent for each round individually is sufficient. We continue by introducing what we call the *blinkered unitary*, that satisfies the required constraints for split and merge moves. In particular, any valid transition from m initial to n final points that can be implemented by means of the blinkered unitary, can be seen as a combination of an $m \to 1$ merge and an $1 \to n$ split (see Section 3.1.1 and Appendix C). With these the former best known explicit protocol with bias $1/6$ [Moc05] can already be derived from its TDPG. We finally study the family of TDPGs with bias $1/10$ and isolate the precise moves required to implement it. These can not be produced by a combination of merges and splits, therefore, we need to look beyond the blinkered unitary. We give analytic expressions for the required unitaries and show that they satisfy the corresponding constraints. This allows us, in effect, to convert the family of games with bias $1/10$, proposed by C. Mochon into explicit protocols, thus breaking the bias $1/6$ barrier. However, we essentially guessed the form that the blinkered unitary and the unitaries of the $1/10$ game should have in these cases, and then showed that they indeed satisfy the required constraints. Games achieving lower biases, though, correspond to larger unitary matrices, therefore this approach becomes untenable. We overcome this issue in Section 4, where we find a way to systematically construct the unitaries for the whole family of C. Mochon's games achieving bias $\epsilon(k) = 1/(4k+2)$ for arbitrary integers $k > 0$.

1.1.2 Exact Unitaries for C. Mochon's assignments—an algebraic solution

As we saw, TEF allows us to convert any TDPG into an explicit protocol, granted that the unitaries satisfying Equation (3) can be found corresponding to each valid transition used in the game (see Theorem 1). Using A. Y. Kitaev's and C. Mochon's formalism [Moc07], we have that the following—an even weaker requirement—is enough (see Section 4.1): Suppose that a valid function, t, can be written as a sum of valid functions. Then, in order to obtain the *effective solution* for t (see Definition 34), it suffices to find unitaries corresponding to the valid functions appearing in the sum.

We consider the class of valid functions that C. Mochon uses in his family of point games approaching bias $\epsilon(k) = \frac{1}{4k+2}$ for an arbitrary integer $k > 0$. These are of the form (see Definition 32)

$$t = \sum_{i=1}^{n} \frac{-f(x_i)}{\prod_{j \neq i}(x_j - x_i)} [\![x_i]\!] ,$$

where $0 \leq x_1 < x_2 \cdots < x_n \in \mathbb{R}$, $f(x)$ is a polynomial[5], and the notation follows Equation (2). We refer to these as *f-assignments* and in particular, when f is a monomial, we call them *monomial assignments*. We observe that the f-assignments can be expressed as a sum of monomial assignments, and we give formulas for the unitaries corresponding to these monomial assignments. There are four types of monomial assignments—which we call balanced or unbalanced (depending on whether the number of points with negative weights in the point game is equal to the number of points with positive weight or not) and aligned or misaligned (depending on whether the power of the polynomial $f(x)$ is even or odd). The formulas for their *solutions* (see Definition 34) and their proofs of correctness comprise most of Section 4 whose central result is summarized in the following theorem.

[5] with some restrictions which we suppress for brevity

Theorem 2 (informal[6]). *Let t be an f-assignment (see Definition 32). Then, t can be expressed as $t = \sum_i \alpha_i t'_i$ where $\alpha_i > 0$ and t'_i are monomial assignments (see Definition 32). Each t'_i admits a solution (see Definition 34) given in either Proposition 39, Proposition 40, Proposition 41 or Proposition 42, depending on the form of t'_i.*

Here, we present the case of a balanced and aligned monomial assignment given by $t = \sum_{i=1}^{n} x_{h_i}^{m} p_{h_i} [\![x_{h_i}]\!] - \sum_{i=1}^{n} x_{g_i}^{m} p_{g_i} [\![x_{g_i}]\!]$ with $b = m/2$ being an integer. Given an orthonormal basis $\{|h_1\rangle, |h_2\rangle \dots |h_n\rangle, |g_1\rangle, |g_2\rangle \dots |g_n\rangle\}$, the *solution* (see Definition 34) is

$$O = \sum_{i=-b}^{n-b-1} \left(\frac{\Pi_{h_i}^{\perp} (X_h)^i |w'\rangle \langle v'| (X_g)^i \Pi_{g_i}^{\perp}}{\sqrt{c_{h_i} c_{g_i}}} + \text{h.c.} \right)$$

where $X_h := \sum_{i=1}^{n} x_{h_i} |h_i\rangle \langle h_i|$, $|w\rangle := \sum_{i=1}^{n} \sqrt{p_{h_i}} |h_i\rangle$, $|w'\rangle := (X_h)^b |w\rangle$, $c_{h_i} := \langle w'| (X_h)^i \Pi_{h_i}^{\perp} (X_h)^i |w'\rangle$,

$$\Pi_{h_i}^{\perp} := \begin{cases} \text{projector orthogonal to span}\{(X_h)^{-|i|+1} |w'\rangle, (X_h)^{-|i|+2} |w'\rangle \dots, |w'\rangle\} & i < 0 \\ \text{projector orthogonal to span}\{(X_h)^{-b} |w'\rangle, (X_h)^{-b+1} |w'\rangle, \dots (X_h)^{i-1} |w'\rangle\} & i > 0 \\ \mathbb{I} & i = 0, \end{cases}$$

and X_g, $|v\rangle$, $|v'\rangle$, c_{g_i}, $\Pi_{g_i}^{\perp}$ are defined analogously.

In Section 4.5 we illustrate as an example the construction of a WCF protocol with bias 1/14 from the corresponding point game by means of the TEF and the analytical solutions to the monomial assignments.

Having found these unitaries, we have effectively solved our problem, since the TEF allows the conversion of point games—including the ones with arbitrarily small bias—into WCF protocols with the respective bias. We should also note that this approach bypasses one of C. Mochon's reductions, thus providing not only a solution but also a simplification of the formalism. However, we considered a specific family of point games (the one proposed by C. Mochon). Our next contribution provides a solution which is applicable beyond these point games.

1.1.3 Elliptic Monotone Align (EMA) algorithm

We now introduce the Elliptic Monotone Align (EMA) algorithm (Section 5), which allows us to numerically find the unitary corresponding to *any* strictly valid function (see Definition 127), thus being applicable beyond the family of point games used above. If we remove the projector in Equation (3), we can express the inequality as $X_h \geq U X_g U^\dagger$ where X_h, X_g are diagonal matrices with positive entries (see Section 5.1). It is possible to show that, without loss of generality, we can restrict ourselves to orthogonal matrices (see Appendix E). Once we restrict to real numbers, the set of vectors $\mathcal{E}_{X_h} := \{|u\rangle \mid \langle u| X_h |u\rangle = 1\}$ describes the boundary of an ellipsoid, since $\sum_i u_i^2/(x_{h_i}^{-1}) = 1$ for x_{h_i} fixed and u_i variable. Similarly, $\mathcal{E}_{O X_g O^T}$ represents a rotated ellipsoid where O is orthogonal (see Figure 3). The larger the x_{h_i} (or x_{g_i}) is, the higher is the curvature of the ellipsoid along the associated direction. Geometrically, the aforesaid inequality can be seen as the containment of the $\mathcal{E}_{X_h}$ ellipsoid inside the $\mathcal{E}_{O X_g O^T}$ ellipsoid, as we describe in Section 5.2. The orthogonal matrix we are looking for also has the property $O|v\rangle = |w\rangle$ from Equation (3). Imagine that in addition, we have $\langle w| X_h |w\rangle = \langle v| X_g |v\rangle$ which in terms of the point game means that the average is preserved; such is the case for the merge move. In terms of the ellipsoids, this means that the ellipsoids touch along the $|w\rangle$ direction. More precisely, the point $|c\rangle := |w\rangle / \sqrt{\langle w| X_h |w\rangle}$ belongs to both $\mathcal{E}_{X_h}$ and $\mathcal{E}_{O X_g O^T}$. Since the inequality tells us that the smaller $\mathcal{E}_{X_h}$ is contained inside the larger $\mathcal{E}_{O X_g O^T}$, and we now know that they touch at $|c\rangle$, we conclude that their normal vectors evaluated at $|c\rangle$ must be equal. Furthermore, the inner ellipsoid must be more curved than the outer one at the point of contact. Mark the point $|c\rangle$ on the $\mathcal{E}_{O X_g O^T}$ ellipsoid, and imagine rotating the $\mathcal{E}_{X_g}$ ellipsoid to the $\mathcal{E}_{O X_g O^T}$ ellipsoid. The normal vector at $|c\rangle$ must be mapped to the normal vector of $\mathcal{E}_{X_h}$ at this point. It turns out that to evaluate

[6] We suppressed some constraints on f for brevity.

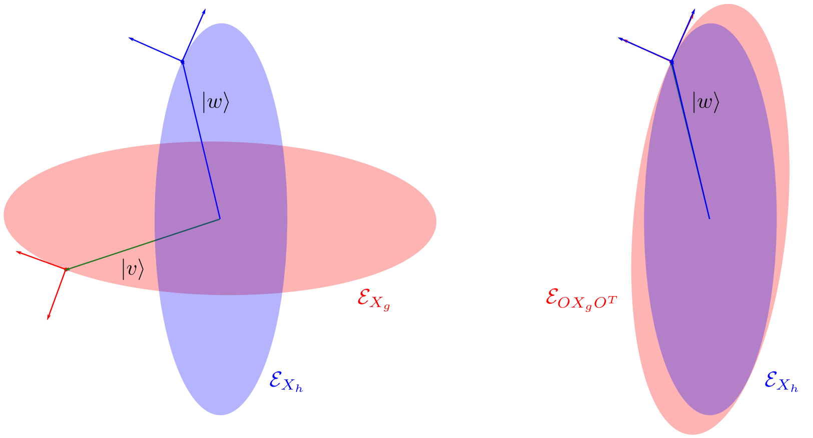


Figure 3: On the left the ellipsoids correspond to the diagonal matrices X_g and X_h, and the vectors $|w\rangle$ and $|v\rangle$ indicate the direction. On the right, the larger ellipsoid is now rotated to OX_gO^T, and the point of contact is along the vector $|w\rangle = O|v\rangle$.

the normal vectors $|n_h\rangle$ on $\mathcal{E}_{X_h}$ and $|n_g\rangle$ on $\mathcal{E}_{X_g}$ at the marked point, we only need to know $X_h, X_g, |v\rangle$ and $|w\rangle$. Complete knowledge of O is not required and yet we can be sure that $O|n_g\rangle = |n_h\rangle$ which means O must have a term $|n_h\rangle\langle n_g|$. In fact, we can even evaluate the curvature from the aforesaid quantities. It turns out that when this condition is expressed precisely, it becomes an instance of the same problem we started with one less dimension, allowing us to iteratively find O, which so far we only assumed to exist.

This, however, only works under our assumption that $\langle w|X_h|w\rangle = \langle v|X_g|v\rangle$. Whenever this is not the case, we resort to the following method. Recall that a monotone function μ is defined to be a function which has the property "$x \geq y \implies \mu(x) \geq \mu(y)$". An operator monotone function is a generalization of the aforesaid property to matrices, which in our notation can be expressed as "$X_h \geq OX_gO^T \implies \mu(X_h) \geq O\mu(X_g)O^T$". It is known that for a certain class of operator monotone functions, μ, the inverse μ^{-1} is also an operator monotone. Using these results in conjunction with results from [Aha+14b] we can show that, after an appropriate scaling of the ellipsoids, there is always an operator monotone μ such that $\langle w|\mu(X_h)|w\rangle = \langle v|\mu(X_g)|v\rangle$. This result also admits a simple geometric interpretation. It means that to establish that $\mathcal{E}_{X_h}$ is inside $\mathcal{E}_{OX_gO^T}$, we can—instead of looking at all different directions and make sure that $\mathcal{E}_{X_h}$ is indeed inside $\mathcal{E}_{OX_gO^T}$— look along a single direction $|w\rangle$ and make sure that all the different ellipsoids $\mathcal{E}_{\mu(X_h)}$ are inside the corresponding $\mathcal{E}_{O\mu(X_g)O^T}$ ellipsoids along just this direction, for every operator monotone μ in the class indicated earlier. Since the orthogonal matrix which solves the initial problem also solves the one mapped by μ, we can use our technique on the latter to proceed. It is essentially a combination of these steps that constitutes our EMA algorithm, which is informally summarized below.

Definition 3 (EMA algorithm (informal))**.** Given a valid transition, the algorithm proceeds in three phases.

1. INITIALIZATION
   - Bring the final points close to zero until the corresponding ellipsoids start to touch (tightening procedure).
   - Find the spectrum of the matrices which represent the ellipsoid. Evaluate the smallest matrix size n needed to represent the problem using ellipsoids.
   - Using the aforesaid, define $\left(X_h^{(n)}, X_g^{(n)}, \left|w^{(n)}\right\rangle, \left|v^{(n)}\right\rangle\right) := \underline{X}^{(n)}$ where the superscript denotes the size of the matrix and vectors.

2. ITERATION

   Input: $\underline{\mathrm{X}}^{(k)}$

   Output: $\underline{\mathrm{X}}^{(k-1)}$, the vector $\left|u_h^{(k)}\right\rangle$ and the orthogonal matrices $\bar{O}_g^{(k)}, \bar{O}_h^{(k)}$

   Procedure:

   - Shrink the outer ellipsoid until it touches the inner ellipsoid (tightening procedure).
   - Use operator monotone functions to make the ellipsoids touch along the $|w\rangle$ direction.
   - Evaluate the curvatures and the normal vector along the $|w\rangle$ direction.
   - Use the curvatures to specify $\underline{\mathrm{X}}^{(k-1)}$ and find the orthogonal matrices $\bar{O}_g^{(k)}, \bar{O}_h^{(k)}$.

3. RECONSTRUCTION

   Evaluate $O^{(n)}$ recursively using $O^{(k)} = \bar{O}_g^{(k)} \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)} \right) \bar{O}_h^{(k)}$.

Theorem 4 (Correctness of the EMA algorithm (informal))**.** *Given a transition of a TDPG, the EMA Algorithm always finds a U such that the constraints in Theorem 1 are satisfied.*

In Section 5.4 one can find the complete algorithm and the proof of its correctness; in particular Definition 79 and Theorem 80 are the corresponding formal statements. The results obtained from a numerical implementation of the EMA algorithm are discussed in Section 5.5.

Despite the apparent simplicity of the main argument there were many difficulties we had to address in the course of this approach. First, we had to extend the results about operator monotone functions in order to use them for the tightening procedure and be certain that the solution stays unchanged under these transformations. We also extended some results related to different representations of the aforesaid transitions, as these situations arise in the tightening procedure (see Section 5.3.1). Finding an easy method for evaluating the curvatures—by means of the Weingarten map—(see Sections 5.2 and 5.4 and Appendix F) is also a key ingredient of our analysis. The trickiest part of the algorithm is to handle the cases where one of the tangent directions of an ellipsoid has infinite curvature. In these cases, the analysis presented so far breaks down, as the normal vector is no longer well-defined; this is for instance the case for the split move. To address this issue we had to introduce what we call the wiggle-v method (see Section 5.3.3 and Section 5.4).

By employing the EMA algorithm we can numerically convert any point game—including those with arbitrarily small bias— into a WCF protocol. However, such a numerical approach has downsides, as it relies on assumptions concerning the accuracy of the algorithm. For instance, we assume that the algorithm can find the roots of polynomials and diagonalize matrices with arbitrary precision; this is not exactly the case, though, as errors arise related to these processes. Therefore, we also propose an analytical solution based on this geometric approach and inspired by the EMA algorithm, as described below in Section 1.1.4.

1.1.4 Exact unitaries for C. Mochon's assignments—a geometric solution

Our last contribution is yet another analytic solution for the f-assignments of C. Mochon's point games, based this time on the ellipsoid picture. The main result of this geometric approach is summarized in a restatement of Theorem 2 as follows:

Theorem 5 (informal[7])**.** *Let t be an f-assignment (see Definition 32). Then, t can be expressed as $t = \sum_i \alpha_i t_i'$ where $\alpha_i > 0$ and t_i' are monomial assignments (see Definition 32), and each t_i' admits a solution (see Definition 34) of the form given in either Proposition 102 or Proposition 104.*

[7] Some constraints on f have been suppressed for brevity.

We prove this result in Section 6. We first show how to construct the solution for the simplest f-assignment; the one for which $f(x) = x^0 = 1$. We call it the f_0-assignment because of the zeroth degree of the polynomial. This construction has all the basic ingredients needed for finding the solution corresponding to higher degree monomial assignments, and can be explained intuitively.

Consider an f_0-assignment to which we are applying the aforementioned ellipsoid approach, assuming that the contact condition, i.e. $\langle w| X_h |w\rangle = \langle v| X_g |v\rangle$ holds, until the point where we have obtained a problem of the same form with one less dimension. It turns out that for an f_0-assignment, this contact condition holds, and, in fact, it continues to hold for all sub-instances of the problem analogously. More precisely, these conditions correspond to $\langle w| (X_h)^k |w\rangle = \langle v| (X_g)^k |v\rangle$, for successively larger integers $k > 0$. Furthermore, since the assignment is a valid function, we know that an O satisfying the necessary conditions exists (see Corollary 144 and Lemma 146). This allows us to iteratively find the solution, O, for the f_0-assignment in Section 6.1. This procedure, however, breaks down for monomial assignments as the contact condition ceases to hold after a number of iterations. The key idea then, is to invert the inequality and use $X_h^{-1} \leq OX_g^{-1}O^T$ instead of $X_h \geq OX_gO^T$. Intuitively, for a monomial assignment of degree m, the k that appears above in the contact condition starts from m; using the inverse leads to successively smaller k and this allows us to iteratively find the solution, O, to monomial assignments in Section 6.2. Having the solutions for monomial assignments effectively gives us the solution to any f-assignment, thus permitting us to construct WCF protocols for the whole family of these point games, including the ones achieving bias arbitrarily close to zero.

2 Existence of almost perfect quantum WCF protocols

The contents of this section are based on two works: the first is by C. Mochon [Moc07]—part of which is attributed to A.Y. Kitaev—and the second is by D. Aharonov, A. Chailloux, M. Ganz, I. Kerenidis and L. Magnin [Aha+14b], who simplified and verified the former. We present the results needed to understand and build our analysis upon, but for their proofs we refer to the original works [Moc07; Aha+14b].

2.1 WCF protocol as an SDP and its dual

Any WCF protocol can be expressed in the following general form (see [Amb04] and page 9 of [Moc07]):

Definition 6 (WCF protocol with bias ϵ)**.** For n even, an n-message WCF protocol between two parties, Alice and Bob, is described by

- three Hilbert spaces: A and B corresponding to Alice's and Bob's private work-spaces (Bob does not have any access to A and, similarly, Alice for B) and a message space M;
- an initial product state $|\psi_0\rangle = |\psi_{A,0}\rangle \otimes |\psi_{M,0}\rangle \otimes |\psi_{B,0}\rangle \in A \otimes M \otimes B$;
- a set of n unitaries $\{U_1, \dots U_n\}$ acting on $A \otimes M \otimes B$ with $U_i = U_{A,i} \otimes \mathbb{I}_B$ for i odd and $U_i = \mathbb{I}_A \otimes U_{B,i}$ for i even;
- a set of honest states $\{|\psi_i\rangle : i \in [n]\}$ defined as $|\psi_i\rangle = U_i U_{i-1} \dots U_1 |\psi_0\rangle$;
- a set of n projectors $\{E_1, \dots E_n\}$ acting on $A \otimes M \otimes B$ with $E_i = E_{A,i} \otimes \mathbb{I}_B$ for i odd, and $E_i = \mathbb{I}_A \otimes E_{B,i}$ for i even, such that $E_i |\psi_i\rangle = |\psi_i\rangle$;
- two positive operator valued measures (POVMs) $\{\Pi_A^{(0)}, \Pi_A^{(1)}\}$ acting on A and $\{\Pi_B^{(0)}, \Pi_B^{(1)}\}$ acting on B.

The WCF protocol proceeds as follows:

- In the beginning, Alice holds $|\psi_{A,0}\rangle |\psi_{M,0}\rangle$ and Bob $|\psi_{B,0}\rangle$.
- For $i = 1$ to n:
  - If i is odd, Alice applies U_i and measures the resulting state with the POVM $\{E_i, \mathbb{I} - E_i\}$. On the first outcome, she sends the message qubits to Bob; on the second outcome, she ends the protocol by outputting "0", i.e, she declares herself to be the winner.
  - If i is even, Bob applies U_i and measures the resulting state with the POVM $\{E_i, \mathbb{I} - E_i\}$. On the first outcome, he sends the message qubits to Alice; on the second outcome, he ends the protocol by outputting "1", i.e., he declares himself to be the winner.
  - Alice and Bob measure their part of the state with the final POVM and output the outcome of their measurements. Alice wins on outcome "0" and Bob on outcome "1".

The WCF protocol has the following properties:

- Correctness: When both parties are honest, their outcomes are always the same: $\Pi_A^{(0)} \otimes \mathbb{I}_M \otimes \Pi_B^{(1)} |\psi_n\rangle = \Pi_A^{(1)} \otimes \mathbb{I}_M \otimes \Pi_B^{(0)} |\psi_n\rangle = 0$.
- Balanced: When both parties are honest, they win with probability $1/2$:
  $P_A = \left|\Pi_A^{(0)} \otimes \mathbb{I}_M \otimes \Pi_B^{(0)} |\psi_n\rangle\right|^2 = \frac{1}{2}$ and $P_B = \left|\Pi_A^{(1)} \otimes \mathbb{I}_M \otimes \Pi_B^{(1)} |\psi_n\rangle\right|^2 = \frac{1}{2}$.

- ϵ-biased: When Alice is honest, the probability that both parties agree on Bob winning is $P_B^* \leq \frac{1}{2} + \epsilon$. Conversely, when Bob is honest, the probability that both parties agree on Alice winning is $P_A^* \leq \frac{1}{2} + \epsilon$.

For a depiction of the protocol see Figure 4.

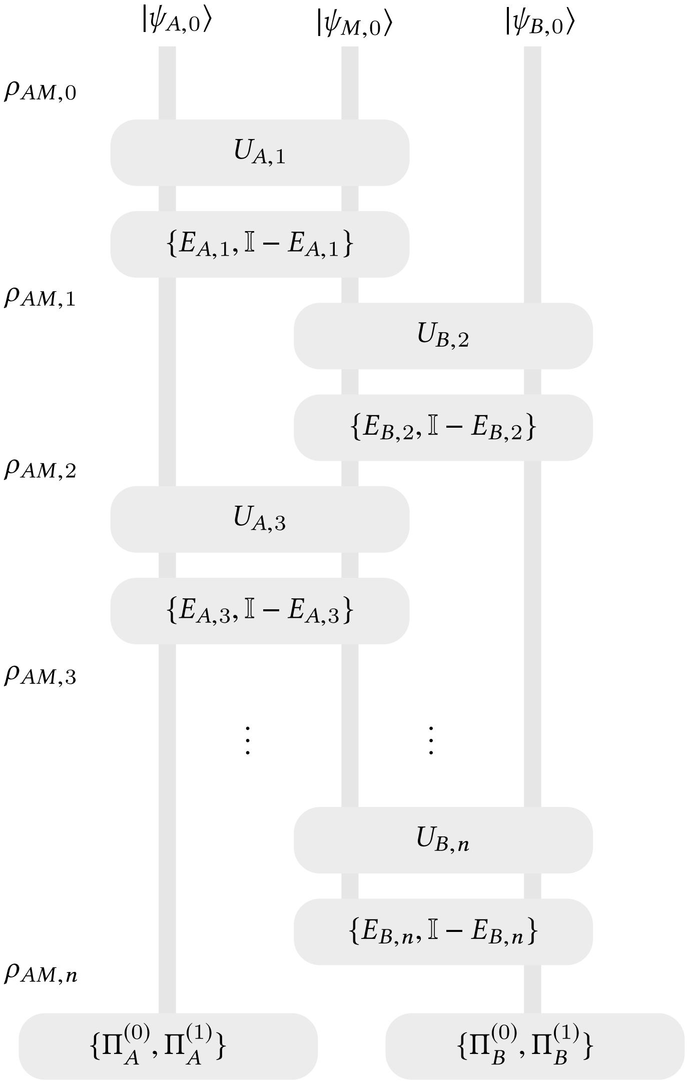

$|\psi_{A,0}\rangle$
$|\psi_{M,0}\rangle$
$|\psi_{B,0}\rangle$
$\rho_{AM,0}$
$U_{A,1}$
$\{E_{A,1}, \mathbb{I} - E_{A,1}\}$
$\rho_{AM,1}$
$U_{B,2}$
$\{E_{B,2}, \mathbb{I} - E_{B,2}\}$
$\rho_{AM,2}$
$U_{A,3}$
$\{E_{A,3}, \mathbb{I} - E_{A,3}\}$
$\rho_{AM,3}$
$\vdots$
$\vdots$
$U_{B,n}$
$\{E_{B,n}, \mathbb{I} - E_{B,n}\}$
$\rho_{AM,n}$
$\{\Pi_A^{(0)}, \Pi_A^{(1)}\}$
$\{\Pi_B^{(0)}, \Pi_B^{(1)}\}$

Figure 4: Every quantum WCF protocol can be cast into this general form.

To define the bias of the protocol, we need to know P_A^* and P_B^* corresponding to the best possible cheating strategy of the opponent. This is formalized by the following (primal) semi-definite program:

Theorem 7 (Primal)**.**
$P_B^* = \max Tr((\Pi_A^{(1)} \otimes \mathbb{I}_M)\rho_{AM,n})$ *over all* $\rho_{AM,i}$ *satisfying the constraints*

- $Tr_M(\rho_{AM,0}) = Tr_{MB}(|\psi_0\rangle\langle\psi_0|) = |\psi_{A,0}\rangle\langle\psi_{A,0}|$;
- *for i odd,* $Tr_M(\rho_{AM,i}) = Tr_M(E_i U_i \rho_{AM,i-1} U_i^\dagger E_i)$;
- *for i even,* $Tr_M(\rho_{AM,i}) = Tr_M(\rho_{AM,i-1})$.

$P_A^* = \max Tr((\mathbb{I}_M \otimes \Pi_B^{(0)})\rho_{MB,n})$ *over all* $\rho_{BM,i}$ *satisfying the constraints*

- $Tr_M(\rho_{MB,0}) = Tr_{AM}(|\psi_0\rangle\langle\psi_0|) = |\psi_{B,0}\rangle\langle\psi_{B,0}|$;
- *for i even,* $Tr_M(\rho_{MB,i}) = Tr_M(E_iU_i\rho_{MB,i-1}U_i^\dagger E_i)$;
- *for i odd,* $Tr_M(\rho_{MB,i}) = Tr_M(\rho_{MB,i-1})$.

Remark 8. In fact, one can restrict to unitaries without loss of generality (see page 9 of [Moc07]) by simulating the projections as coherent measurements and absorbing them into the final measurement. Generality is not lost because (a) the projections can only improve the bias and (b) a protocol with projections can be converted into one without projections. The use of projectors, though, can simplify the proofs, as we will see later. One could have, in addition to the measurement $\{E_i, \mathbb{I} - E_i\}$, introduced a similar measurement, say $\{F_i, \mathbb{I} - F_i\}$, before the unitary. This would yield $\mathrm{tr}_M(\rho_{AM,i}) = \mathrm{tr}_M(E_iU_iF_i\rho_{AM,i-1}F_iU_i^\dagger E_i)$ for the SDP of P_B^*.

Notice that P_B^* depends on Alice's actions specified in the protocol —as we optimize over all possible actions of Bob—and thus involves variables such as $\rho_{AM,i}$ and $U_{A,i}$. Analogously, P_A^* depends on Bob's actions.

A feasible solution to an optimization problem satisfies the constraints but is not necessarily optimal. For the primal problems, a feasible solution gives a lower bound on P_A^* and P_B^* (for details and the proof see [Aha+14b; Moc07]). Instead, we can consider the corresponding dual problems, a feasible solution to which gives an upper bound on P_A^* and P_B^* (for details and the proof see [Aha+14b; Moc07]). We can further prove that in this case strong duality holds [Moc07; Aha+14b]; this means that the optimal value of the dual program yields P_A^* and P_B^* exactly and not just a bound. In terms of the protocol, it means that there exist cheating strategies corresponding to the optimal values of the dual.

Theorem 9 (Dual)**.**
$P_B^* = \min Tr(Z_{A,0}|\psi_{A,0}\rangle\langle\psi_{A,0}|)$ *over all* $Z_{A,i}$ *under the constraints*

1. $\forall i, Z_{A,i} \geq 0$;
2. *for i odd,* $Z_{A,i-1} \otimes \mathbb{I}_M \geq U_{A,i}^\dagger E_{A,i}(Z_{A,i} \otimes \mathbb{I}_M)E_{A,i}U_{A,i}$;
3. *for i even,* $Z_{A,i-1} = Z_{A,i}$;
4. $Z_{A,n} = \Pi_A^{(1)}$.

$P_A^* = \min Tr(Z_{B,0}|\psi_{B,0}\rangle\langle\psi_{B,0}|)$ *over all* $Z_{B,i}$ *under the constraints*

1. $\forall i, Z_{B,i} \geq 0$;
2. *for i even,* $\mathbb{I}_M \otimes Z_{B,i-1} \geq U_{B,i}^\dagger E_{B,i}(\mathbb{I}_M \otimes Z_{B,i})E_{B,i}U_{B,i}$;
3. *for i odd,* $Z_{B,i-1} = Z_{B,i}$;
4. $Z_{B,n} = \prod_B^{(0)}$.

We add one more constraint to the above dual SDPs.

5. $|\psi_{A,0}\rangle$ *is an eigenvector of* $Z_{A,0}$ *with eigenvalue* $\beta > 0$ *and* $|\psi_{B,0}\rangle$ *is an eigenvector of* $Z_{B,0}$ *with eigenvalue* $\alpha > 0$.

Remark 10. As in Remark 8, the dual SDP for P_B^* would have yielded the constraint

$$Z_{A,i-1} \otimes \mathbb{I}_M \geq F_{A,i}U_{A,i}^\dagger E_{A,i}\left(Z_{A,i} \otimes \mathbb{I}_M\right)E_{A,i}U_{A,i}F_{A,i} \qquad \text{for } i \text{ odd.}$$

In the next subsection we will see why the fifth constraint is useful; before that we define the *dual feasible points* to be those that satisfy this constraint:

Definition 11 (dual feasible points)**.** We call *dual feasible points* any two sets of matrices $\{Z_{A,0}, \dots, Z_{A,n}\}$ and $\{Z_{B,0}, \dots, Z_{B,n}\}$ that satisfy the conditions 1 to 5 as listed in Theorem 9.

Related to the dual feasible points, the following proposition also holds (see [Aha+14b; Moc07] for the proof):

Proposition 12. *$P_A^* = \inf \alpha$ and $P_B^* = \inf \beta$ where the infimum is over all dual feasible points and β, α are defined in constraint 5 of the definition of the dual feasible points.*

2.2 TDPGs with EBM transitions/functions

We would like now to remove all inessential information from the two aforesaid dual problems; that is the basis information. A. Y. Kitaev achieved this by considering, at a given step, the dual variables Z_A, Z_B as observables with $|\psi\rangle$ governing the probability. This combines the evolution of the certificates on cheating probabilities with the evolution of the honest state—the state obtained when none of the parties is cheating.[8] Let us start with the definition of the function 'Prob', which essentially permits us to remove the basis dependence of the dual SDP.

Definition 13 (Prob)**.** Consider $Z \geq 0$ and let $\Pi^{[z]}$ represent the projector on the eigenspace of eigenvalue $z \in \text{spectrum}(Z)$. We have $Z = \sum_z z\Pi^{[z]}$. Let $|\psi\rangle$ be a vector, not necessarily normalized. We define the function $\text{Prob}[Z, \psi] : [0, \infty) \to [0, \infty)$ as

$$\text{Prob}[Z, \psi](z) = \begin{cases} \langle\psi| \Pi^{[z]} |\psi\rangle & \text{if } z \in \text{sp}(Z) \\ 0 & \text{else.} \end{cases}$$

If $Z = Z_A \otimes \mathbb{I}_M \otimes Z_B$, using the same notation, we define the 2−variate function $\text{Prob}[Z_A, Z_B, \psi] : [0, \infty) \times [0, \infty) \to [0, \infty)$, with finite support, as

$$\text{Prob}[Z_A, Z_B, \psi](z_A, z_B) = \begin{cases} \langle\psi| \Pi^{[z_A]} \otimes \mathbb{I}_M \otimes \Pi^{[z_B]} |\psi\rangle & \text{if } (z_A, z_B) \in \text{sp}(Z_A) \times \text{sp}(Z_B), \\ 0 & \text{else.} \end{cases}$$

We would like the point game framework to be protocol-independent. The following definitions of *Expressible by Matrices* (EBM) transitions facilitate such a description.

Definition 14 (Line Transition)**.** A line transition is an ordered pair of finitely supported functions $g, h : [0, \infty) \to [0, \infty)$, which we conveniently denote as $g \to h$.

Definition 15 (EBM line transition)**.** Let $g, h : [0, \infty) \to [0, \infty)$ be two functions with finite supports. The line transition $g \to h$ is EBM if there exist two matrices $0 \leq G \leq H$ and a vector $|\psi\rangle$, not necessarily normalized, such that $g = \text{Prob}\,[G, |\psi\rangle]$ and $h = \text{Prob}\,[H, |\psi\rangle]$.

Definition 16 (EBM transition)**.** Let $g, h : [0, \infty) \times [0, \infty) \to [0, \infty)$ be two functions with finite supports. The transition $g \to h$ is an

- EBM horizontal transition if for all $y \in [0, \infty)$, $g(., y) \to h(., y)$ is an EBM line transition, and

[8] Originally, using a similar maneuver, A. Y. Kitaev settled the solvability of the quantum strong coin flipping problem by giving a lower bound on its bias [Kit03].

- EBM vertical transition if for all $x \in [0, \infty)$, $g(x, .) \longrightarrow h(x, .)$ is an EBM line transition.

Remark 17. When clear from the context, we refer to an EBM line transition also as an EBM transition.

When we wrote the dual SDP, the order of the constraints got inverted, i.e. the condition associated with the final measurements and states appeared first and the condition associated with the initial state appeared in the end. We expect the final state to be an EPR-like state to which two points of the point game can be associated in the basis-independent description of the dual (2−variate function from Definition 13), while the initial state of the protocol should be unentangled and correspond to a single point in the basis-independent description of the dual. The rules for moving these points must be related to the dual constraints and they are already formalized into EBM transitions. The notation

$$\llbracket x_g, y_g \rrbracket (x, y) = \begin{cases} 1 & x_g = x \text{ and } y_g = y \\ 0 & \text{else} \end{cases}$$

is useful for the description of the EBM point games that follows.

Definition 18 (EBM point game)**.** An EBM point game is a sequence of functions $\{g_0, g_1, \ldots, g_n\}$ with finite support such that

- $g_0 = 1/2 \llbracket 0, 1 \rrbracket + 1/2 \llbracket 1, 0 \rrbracket$;
- for all even i, $g_i \longrightarrow g_{i+1}$ is an EBM vertical transition;
- for all odd i, $g_i \longrightarrow g_{i+1}$ is an EBM horizontal transition;
- $g_n = 1 \llbracket \beta, \alpha \rrbracket$ for some $\alpha, \beta \in [0, 1]$. We call $\llbracket \beta, \alpha \rrbracket$ the final point of the EBM point game.

Since we started with a WCF protocol, considered its dual and re-expressed it as a TDPG (which is just a basis-independent representation), the following proposition (for the proof see [Aha+14b]) should not come as a surprise.

Proposition 19 (WCF $\implies$ EBM point game)**.** *Given a WCF protocol with cheating probabilities P_A^* and P_B^*, along with a positive real number $\delta > 0$, there exists an EBM point game with final point $\llbracket P_B^* + \delta, P_A^* + \delta \rrbracket$.*

The converse statement—given an EBM TDPG the corresponding WCF protocol can be constructed—is not as easy to see, but indeed it holds. By using only "allowed moves" one can be sure that there exists a corresponding sequence of unitaries U_i, measurements $\Pi_{A/B}$ and an initial state $|\psi_0\rangle$ complemented by the dual variables $Z_{A,i}$ and $Z_{B,i}$ which certify the bias corresponding to the coordinates of the final point in the point game.

Theorem 20 (EBM to protocol)**.** *Given an EBM point game with final point $\llbracket \beta, \alpha \rrbracket$, there exists a WCF protocol with $P_A^* \leq \alpha$ and $P_B^* \leq \beta$.*

One can find a proof in [Moc07; Aha+14b]. We also sketch an alternative proof later in Section 3 after Theorem 31. This establishes the equivalence between EBM TDPGs and WCF protocols.

2.3 TDPGs with valid functions

To check whether a given transition is EBM is not an easy task; A. Y. Kitaev and C. Mochon [Moc07] introduced the following alternative characterization of EBM line transitions in order to simplify the analysis. We use the notation from Equation (2).

Proposition 21. *Let $g \longrightarrow h$ where $g = \sum_{i=1}^{n_g} p_{g_i} [\![x_{g_i}]\!]$ and $h = \sum_{i=1}^{n_h} p_{h_i} [\![x_{h_i}]\!]$ with all x_{g_i}, x_{h_i} being non-negative and distinct ($x_{g_i} \neq x_{g_j}$ and $x_{h_i} \neq x_{h_j}$ for every $i \neq j$), and $p_{g_i}, p_{h_i} > 0$. Then, the transition is EBM if it is* strictly valid, *i.e. the following equality holds and the inequalities are* strictly *satisfied:*

$$\sum_{i=1}^{n_h} p_{h_i} = \sum_{i=1}^{n_g} p_{g_i}$$
$$\sum_{i=1}^{n_h} p_{h_i} \frac{\lambda x_{h_i}}{\lambda + x_{h_i}} \geq \sum_{i=1}^{n_g} p_{g_i} \frac{\lambda x_{g_i}}{\lambda + x_{g_i}} \quad \forall \lambda > 0, \quad \textit{and} \quad \sum_{i=1}^{n_h} x_{h_i} p_{h_i} \geq \sum_{i=1}^{n_g} x_{g_i} p_{g_i}.$$

Conversely, a transition is valid, *i.e. satisfies these inequalities (see also Definition 120), if the transition $g \longrightarrow h$ is EBM.*

Whenever g and h have disjoint support, we can equivalently consider the function $t = h - g$ and rewrite the above relationships as

$$\sum_{x \in \operatorname{supp}(t)} t(x) = 0$$
$$\sum_{x \in \operatorname{supp}(x)} t(x) f_\lambda(x) \geq 0 \quad \forall \lambda > 0, \quad \text{and} \quad \sum_{x \in \operatorname{supp}(x)} t(x) x \geq 0,$$
$$\text{where} \quad f_\lambda(x) = \frac{\lambda x}{(\lambda + x)}. \tag{4}$$

We may therefore speak of valid and EBM *functions* instead of transitions (see Definition 120, Definition 106 and Corollary 123), and rephrase Definition 18 in terms of EBM or valid functions instead of transitions (see Definition 109). We can also extend the definitions of EBM line/horizontal/vertical transitions to EBM or valid line/horizontal/vertical functions (see Definition 108). The proof of the above Proposition 21, as presented in [Moc07; Aha+14b], uses an interesting connection with operator monotone functions through conic duality arguments, and in Appendix A we include the conic duality analysis leading to the proof. In [Aha+14b; Moc07] the authors start by noticing that the set of EBM functions is a convex cone K and its dual, K^*, is the set of operator monotone functions. Then, they show that the bi-dual cone, $K^{**} = \operatorname{cl}(K)$, is the set of valid functions. This way, they prove that the set of EBM functions and the set of valid functions are the same up to closures. We also briefly sketch this proof in Appendix B in the course of showing that the aforementioned sets are also equal to the set of functions satisfying the constraints of our framework introduced in Section 3.

This alternative characterization of EBM transitions significantly simplifies the analysis, as it removes entirely the matrices and trades them for scalar conditions, albeit infinitely many of them, one for each $\lambda > 0$. This simplification, though, comes with a catch. The two cones—the cone of EBM functions and the dual of the cone of operator monotone functions—are the same, but given an element in the second we do not have a recipe for finding the matrices certifying that it is an EBM function; only their existence is guaranteed and this is where C. Mochon's approach becomes non-constructive. Without these matrices, we can not find the protocol.

Below, we present examples of valid transitions describing allowed moves in a TDPG. In the first, we see what can be done with a single point; we can increase its coordinates to raise it in the frame. The second example is that of merging two (or more) points into one, and the third is about splitting a single point into two (or more). The proofs for the validity of these transitions can be found in [Moc07].

Example 22 (Point raise)**.** $p [\![x_g]\!] \longrightarrow p [\![x_h]\!]$ with $x_h \geq x_g$ is a valid transition.

Example 23 (Point merge)**.** $p_{g_1} [\![x_{g_1}]\!] + p_{g_2} [\![x_{g_2}]\!] \to (p_{g_1} + p_{g_2}) [\![x_h]\!]$ with $x_h \geq \frac{p_{g_1}x_{g_1}+p_{g_2}x_{g_2}}{p_{g_1}+p_{g_2}}$ is a valid transition, or generally $\sum_i p_{g_i} [\![x_{g_i}]\!] \to (\sum_i p_{g_i}) [\![x_h]\!]$ with $x_h \geq \langle x_g \rangle$ is a valid transition.

Example 24 (Point split)**.** $p_g [\![x_g]\!] \to p_{h_1} [\![x_{h_1}]\!] + p_{h_2} [\![x_{h_2}]\!]$ with $p_g = p_{h_1} + p_{h_2}$ and $\frac{p_g}{x_g} \geq \frac{p_{h_1}}{x_{h_1}} + \frac{p_{h_2}}{x_{h_2}}$ is a valid transition, or generally $(\sum_i p_{h_i}) [\![x_g]\!] \to \sum_i p_{h_i} [\![x_{h_i}]\!]$ with $\frac{1}{x_g} \geq \left\langle \frac{1}{x_h} \right\rangle$ is a valid transition.

2.4 Time-Independent Point Games (TIPGs)

The protocol with bias 1/6 [Moc05] can be expressed as a TDPG which uses only the moves described in Example 22, Example 23, Example 24. However, the point game formalism can be further simplified, and it is in this simplified formalism that C. Mochon constructed his family of point games achieving arbitrarily small bias. Instead of considering the entire sequence of horizontal and vertical transitions, he focused on just two functions (hence the name time-independent), as described below:

Definition 25 (TIPG)**.** A *time-independent point game (TIPG)* is a valid horizontal function, denoted by a, and a valid vertical function, denoted by b, such that

$$a + b = 1 [\![\beta, \alpha]\!] - \frac{1}{2} [\![0, 1]\!] - \frac{1}{2} [\![1, 0]\!]$$

for some $\alpha, \beta > 1/2$. Further

- we call the point $[\![\beta, \alpha]\!]$ the final point of the game, and
- we call the set $\mathcal{S} = ((\text{supp}(a) \cup \text{supp}(b)) \setminus \text{supp}(a + b)$, the set of intermediate points.

Remark 26. When clear from the context, we may use the word TIPG even when $a + b$ is not necessarily $[\![\beta, \alpha]\!] - \frac{1}{2} ([\![0, 1]\!] + [\![1, 0]\!])$ but some other function, c, with finite support in $[0, \infty) \times [0, \infty)$ satisfying $\sum_{x \in \text{supp}(c)} c(x) = 0$.

Theorem 27 (TIPG to valid point games)**.** *Given a TIPG with a valid horizontal function a and a valid vertical function b such that* $a+b = 1 [\![\beta, \alpha]\!] - \frac{1}{2} [\![0, 1]\!] - \frac{1}{2} [\![1, 0]\!]$*, we can construct, for all* $\epsilon > 0$*, a valid point game with its final point being* $[\![\beta + \epsilon, \alpha + \epsilon]\!]$*, where the number of transitions depends on* ϵ*.*

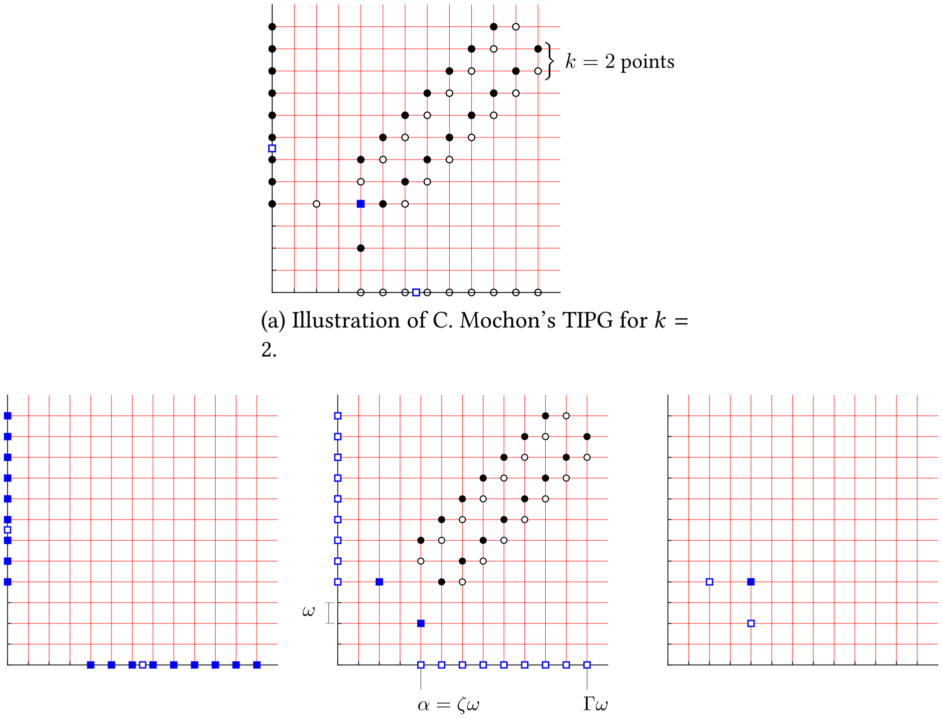
If we have a valid point game we can combine the horizontal and vertical functions to obtain a and b. In particular, if the valid point game with final point $[\![\beta, \alpha]\!]$ is specified by $a_1, a_2 \ldots a_n$ valid horizontal and $b_1, b_2 \ldots b_n$ valid vertical functions, then the corresponding TIPG is specified by $a = \sum_{i=1}^n a_i$ and $b = \sum_{i=1}^n b_i$, which are horizontally and vertically valid, respectively, and satisfy $a + b = [\![\beta, \alpha]\!] - \frac{1}{2} [\![0, 1]\!] - \frac{1}{2} [\![1, 0]\!]$.

It is a little counter-intuitive that a TIPG can be converted to a valid TDPG with an arbitrarily small cost on the bias, as it is not clear how a time ordered sequence can be extracted. We might run into causal loops; we expect a point to be present to create another point which in turn is required to produce the first point. The trick is to employ the method of the *catalyst state* [Moc07; Aha+14b]: we deposit a little bit of weight wherever there is negative weight for a, for instance, and then we can implement a scaled down round of a and b. The scaling is proportional to the weight that is placed in the beginning. Repeating this procedure multiple times yields the required final state along with the catalyst state which stays unchanged. Absorbing the catalyst state leads to a small increase in the bias. The number of rounds increases with how small we want this increase in the bias. In particular, we have the following corollary, whose proof along with the proof of Theorem 27 can be found in [Moc07; Aha+14b].

Corollary 28. *Consider a TIPG with a valid horizontal function* $a = a^+ - a^-$ *and a valid vertical function* $b = b^+ - b^-$ *such that* $a + b = [\![\beta, \alpha]\!] - \frac{1}{2} [\![0, 1]\!] - \frac{1}{2} [\![1, 0]\!]$*. Let* Γ *be the largest coordinate of all the points that appear in the TIPG. Then, for all* $\epsilon > 0$*, we can construct a point game with* $O\left(\frac{\|b\|\Gamma^2}{\epsilon^2}\right)$ *valid transitions and final point* $[\![\beta + \epsilon, \alpha + \epsilon]\!]$*.*

2.5 C. Mochon's TIPG achieving bias $\epsilon(k) = 1/(4k+2)$

The existence of quantum WCF protocols with arbitrarily small bias was established by C. Mochon as follows. He constructed a family of TIPGs, parametrized by an integer $k > 0$, such that the final point is $[\![\frac{1}{2} + \epsilon(k), \frac{1}{2} + \epsilon(k)]\!]$, where $\epsilon(k) = 1/(4k+2)$. To achieve bias $\epsilon(k)$ we must have $k = O(\frac{1}{\epsilon})$, therefore for large k we achieve almost zero bias [Moc07] (see Figure 5a). Let us briefly describe the general structure of these games. Apart from their initial points, $[\![0,1]\!]$ and $[\![1,0]\!]$, all the other points involved are placed on a regular lattice, i.e. at locations of the form $[\![a\omega, b\omega]\!]$ where $a, b \in \mathbb{N}$ and $\omega \in (0,\infty)$. The final point of the games is at $[\![\alpha,\alpha]\!]$ for $\alpha = \zeta\omega = \frac{1}{2} + O\left(\frac{1}{k}\right)$ where $\zeta \in \mathbb{N}$, and in general, they have the following three stages (see Figure 5b):



(a) Illustration of C. Mochon's TIPG for $k = 2$.

(b) C. Mochon's TIPG may be understood in three stages, the initial *splits*, the *ladder*, and the *raises*.

Figure 5: Mochon's TIPG

1. *Split.* The point $[\![0,1]\!]$ is vertically split into many points along the y-axis. The resulting points lie between $\zeta\omega$ and $\Gamma\omega$ with $\zeta, \Gamma \in \mathbb{N}$. Analogously, the point $[\![1,0]\!]$ is horizontally split into many points along the x-axis.

2. *Ladder.* This is the main non-trivial move of the games parametrized by an integer $k > 0$, and it consists of points along the diagonal and along the axes (see the second image in Figure 5b). The points on the axis are transformed by the ladder into the final points $[\![\alpha - k\omega, \alpha]\!]$ and $[\![\alpha, \alpha - k\omega]\!]$.

3. *Raise.* The two points $[\![\alpha - k\omega, \alpha]\!]$ and $[\![\alpha, \alpha - k\omega]\!]$ are raised to the final point $[\![\alpha,\alpha]\!]$.

For each integer $k > 0$ there exist parameters $\omega, \Gamma \in (0,\infty)$ such that the two initial splits are valid, the *ladder* corresponds to a horizontally and vertically valid function, and $\alpha = \frac{1}{2} + O\left(\frac{1}{k}\right)$ (see [Moc07; Aha+14b]).

The key technical tool that C. Mochon introduced is the following: given a set of point coordinates, he constructed a way of assigning non-trivial weights to them such that this assignment is valid while still retaining considerable freedom. This weight assignment is parametrized by a polynomial and works for essentially all polynomials up to a certain degree. In other words, he simplified the validity condition by restricting to a class of functions which are easy to manipulate and valid by construction.

Lemma 29. *Let*

- $x_1, x_2 \dots x_n$ *be non-negative and distinct real numbers,*
- f *be a polynomial of degree at most* $n-1$ *satisfying* $f(-\lambda) \geq 0$ *for all* $\lambda \geq 0$.

Then, $a = \sum_{i=1}^{n} \frac{-f(x)}{\prod_{j \neq i}(x_j - x_i)} [\![x_i]\!]$ *is a valid function.*

This function a, i.e. the function that C. Mochon uses to assign the probability weights to the points of his TIPGs, is what we call in our analysis an f-assignment; in Section 4 we provide for this Definition 32, which is tailored to our purpose of constructing an analytical solution.

3 TDPG-to-Explicit-protocol Framework (TEF) and bias 1/10 game and protocol

In this section, we give a framework for converting TDPGs into explicit protocols granted that an EBM-like condition (see Definition 15) holds, and we then use it to construct the unitaries that specify WCF protocols approaching bias 1/10.

Our goal is to construct a protocol (see Definition 6) such that its dual (see Theorem 9) matches a given TDPG. The main difference in our construction, compared to [Aha+14b] and [Moc07], is that the message register decouples after each round by suitably placing the cheat-detection projectors. Consequently, the non-trivial constraint that the dual matrices must satisfy turns out to be similar to the EBM condition. With Definition 13 in mind, intuitively, the most natural way of constructing Z and $|\psi\rangle$, given an arbitrary frame, is to construct an entangled state that encodes the weight, and define Z to contain the coordinates corresponding to this weight. The so-called *Canonical Form* makes this precise.

Definition 30 (Canonical Form)**.** The tuple $(|\psi\rangle, Z^A, Z^B)$ is said to be in the Canonical Form with respect to a set of points in a frame of a TDPG[9] if $|\psi\rangle = \sum_i \sqrt{P_i}\, |ii\rangle_{AB} \otimes |\varphi\rangle_M$, $Z^A = \sum x_i\, |i\rangle\langle i|_A$ and $Z^B = \sum y_i\, |i\rangle\langle i|_B$ where $|\varphi\rangle_M$ represents the state of extra uncoupled registers which might be present.

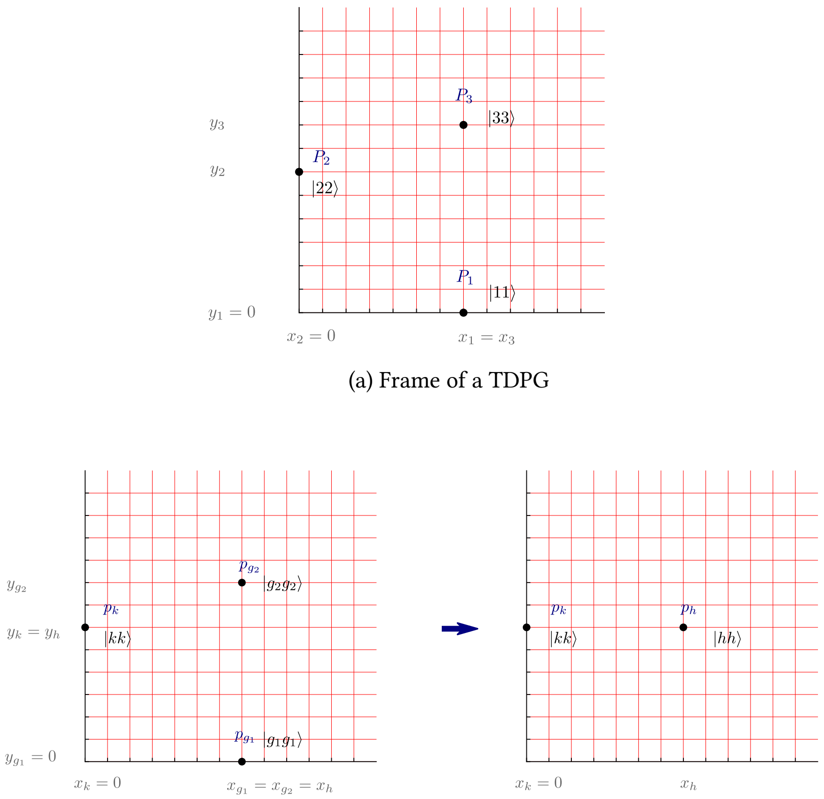
The label $|ii\rangle$ corresponds to a point with coordinates x_i, y_i and weight P_i in the frame (see also Figure 6a). It is tempting to imagine that we systematically construct, from each frame of a TDPG, a canonical form of $|\psi\rangle$ *s* and Zs, and the unitaries can be deduced from the evolution of $|\psi\rangle$. However, this approach has two problems: first, the unitaries are not necessarily decomposable into moves by Alice and Bob who communicate only through the message register and, second, the constraints imposed on consecutive Zs, of the form $Z_{n-1} \otimes \mathbb{I} \geq U_n^\dagger (Z_n \otimes \mathbb{I}) U_n$, are not satisfied in general. Our approach solves these issues. The outputs of our framework are variables indexed as $|\psi_{(i)}\rangle$, $Z_{(i)}$, $U_{(i)}$ (see Definition 18 and Proposition 19) and they are produced in the reverse time convention with respect to the protocol. This means that the variables at the ith step of the protocol (which follows the forward time convention) would be given by $|\psi_i\rangle = |\psi_{(N-i)}\rangle$, $Z_i = Z_{(N-i)}$ and $U_i = U^\dagger_{(N-i)}$. Furthermore, our results extend naturally to the case where U_i may not be unitary and contains projections, e.g. $U_i E_i = E_{(N-i)} U^\dagger_{(N-i)}$. After presenting the framework, we construct the unitaries that implement the three basic moves of a TDPG given in Example 22, Example 24 and Example 23. When generalized to n points these moves are enough to construct the protocol with bias 1/6 from its TDPG [Moc05; Moc07], however they do not exhaust the set of possible moves and we need to consider more advanced ones to go below bias 1/6.

3.1 The framework

Let us start with an informal outline of our framework. Assume that a canonical description is given. Let the labels on the points we want to transform be $\{g_i\}$, and let us also assume that we wish to apply a horizontal-transition, i.e. Alice performs the non-trivial step. Let the labels of the points that will be left unchanged be $\{k_i\}$ (see Figure 6b). We can write the state as

$$|\psi_{(1)}\rangle = \left(\sum_i \sqrt{p_{g_i}}\, |g_i g_i\rangle_{AB} + \sum_i \sqrt{p_{k_i}}\, |k_i k_i\rangle_{AB} \right) \otimes |m\rangle_M .$$

[9] One could define the canonical form for any frame but we only use it for those arising from TDPGs.



(a) Frame of a TDPG

(b) The points which are unchanged from one frame to another are labeled by $\{k_i\}$. Among the points that change, the initial ones are labeled by $\{g_i\}$ and the final ones by $\{h_i\}$.

Figure 6: Illustrations for the Canonical Form

We[10] want Bob to send his part of $|g_i\rangle$ states to Alice through the message register. One way is that to conditionally swap to obtain

$$\left|\psi_{(2)}\right\rangle = \sum_i \sqrt{p_{g_i}}\, |g_i g_i\rangle_{AM} \otimes |m\rangle_B + \sum_i \sqrt{p_{k_i}}\, |k_i k_i\rangle_{AB} \otimes |m\rangle_M .$$

This way, all the points align along the $y-$axis, while the respective $x-$coordinates remain the same due to the fact that it is a horizontal transition. Let $\{h_i\}$ be the labels of the new points after the transformation. We assume that h_i, g_i and k_i index orthonormal vectors. Alice can update the probabilities and labels by locally performing a unitary to obtain

$$\left|\psi_{(3)}\right\rangle = \sum_i \sqrt{p_{h_i}}\, |h_i h_i\rangle_{AM} \otimes |m\rangle_B + \sum_i \sqrt{p_{k_i}}\, |k_i k_i\rangle_{AB} \otimes |m\rangle_M .$$

It is precisely this step which yields the non-trivial constraint. Bob must now accept this by ‘unswapping’ to get

$$\left|\psi_{(4)}\right\rangle = \left(\sum_i \sqrt{p_{h_i}}\, |h_i h_i\rangle_{AB} + \sum_i \sqrt{p_{k_i}}\, |k_i k_i\rangle_{AB} \right) \otimes |m\rangle_M .$$

[10] To be explicit, for $\mathcal{X} \in \{\mathcal{A}, \mathcal{M}, \mathcal{B}\}$, the Hilbert space $\mathcal{X}$ is the span of the orthonormal vectors $\{\{|g_i\rangle_X\}_i, \{|k_i\rangle_X\}_i, \{|h_i\rangle_X\}_i, |m\rangle\}$

In the actual protocol the sequence is in the reverse time convention. Note also that we add a few extra frames to the final TDPG to go from a given frame to the next of the original TDPG. This is irrelevant, when resource usage is not of interest, as the bias does not change.

We now fill in the details and prove the correctness of the above sequence of steps.

1. **First frame.**

$$\begin{aligned}
\left|\psi_{(1)}\right\rangle &= \left(\sum_i \sqrt{p_{g_i}}\, |g_i g_i\rangle_{AB} + \sum_i \sqrt{p_{k_i}}\, |k_i k_i\rangle_{AB}\right) \otimes |m\rangle_M \\
Z^A_{(1)} &= \sum_i x_{g_i} |g_i\rangle \langle g_i|_A + \sum_i x_{k_i} |k_i\rangle \langle k_i|_A \\
Z^B_{(1)} &= \sum_i y_{g_i} |g_i\rangle \langle g_i|_B + \sum_i y_{k_i} |k_i\rangle \langle k_i|_B \,.
\end{aligned}$$

Proof. Follows from the assumption of starting with a Canonical Form. □

2. **Bob sends to Alice.** With $y \geq \max\{y_{g_i}\}$ the following choice

$$\begin{aligned}
\left|\psi_{(2)}\right\rangle &= \sum_i \sqrt{p_{g_i}}\, |g_i g_i\rangle_{AM} \otimes |m\rangle_B + \sum_i \sqrt{p_{k_i}}\, |k_i k_i\rangle_{AB} \otimes |m\rangle_M \\
U_{(1)} &= U_{BM}^{\mathrm{SWP}\{\vec{g},m\}} \\
Z^A_{(2)} &= Z^A_{(1)} \quad \text{and} \quad Z^B_{(2)} = y \mathbb{I}_B^{\{\vec{g},m\}} + \sum_i y_{k_i} |k_i\rangle \langle k_i|_B \,,
\end{aligned}$$

is a viable choice, in the sense that it satisfies the properties (1) $\left|\psi_{(2)}\right\rangle = U_{(1)} \left|\psi_{(1)}\right\rangle$, and (2) $U^\dagger_{(1)} \left(Z^B_{(2)} \otimes \mathbb{I}_M\right) U_{(1)} \geq \left(Z^B_{(1)} \otimes \mathbb{I}_M\right)$.

Proof. We have to prove that the above properties (1) and (2) are satisfied. (1) It follows trivially from the defining action of $U_{(1)}$.
(2) For ease of notation, let $U = U_{(1)}$ and note that $U^\dagger = U$, so that we can write

$$\begin{aligned}
& U \left(Z^B_{(2)} \otimes \mathbb{I}_M\right) U \\
&= y \left(U \left(\mathbb{I}_B^{\{\vec{g},m\}} \otimes \mathbb{I}_M^{\{\vec{g},m\}}\right) U + U \underbrace{\left(\mathbb{I}_B^{\{\vec{g},m\}} \otimes \mathbb{I}_M^{\{\vec{k},\vec{h}\}}\right)}_{\text{outside } U\text{'s action space}} U \right) \\
&+ U \underbrace{\left(\sum y_{k_i} |k_i\rangle \langle k_i| \otimes \mathbb{I}\right)}_{\text{outside } U\text{'s action space}} U = Z_{(2)} \otimes \mathbb{I}_M \geq Z_{(1)} \otimes \mathbb{I}_M
\end{aligned}$$

so long[11] as $y \geq y_{g_i}$ which is guaranteed by the choice of y. □

[11] By the action space of U we mean the space where U acts non-trivially.

3. **Alice's non-trivial step.** Consider the following choice

$$\begin{aligned}\left|\psi_{(3)}\right\rangle &= \sum_i \sqrt{p_{h_i}}\left|h_i h_i\right\rangle_{AM} \otimes |m\rangle_B + \sum_i \sqrt{p_{k_i}}\left|k_i k_i\right\rangle_{AB} \otimes |m\rangle_M \\ E_{(2)}U_{(2)} &= E_{(2)}\left(|w\rangle\langle v| + \text{other terms acting on span}\{\,|h_i h_i\rangle, |g_i g_i\rangle\}\right)_{AM} \\ Z^A_{(3)} &= \sum_i x_{h_i}|h_i\rangle\langle h_i| + \sum_i x_{k_i}|k_i\rangle\langle k_i| \quad \text{and} \quad Z^B_{(3)} = Z^B_{(2)}\end{aligned}$$

where

$$|v\rangle = \frac{\sum_i \sqrt{p_{g_i}}\,|g_i g_i\rangle}{\sqrt{\sum_i p_{g_i}}},\ |w\rangle = \frac{\sum_i \sqrt{p_{h_i}}\,|h_i h_i\rangle}{\sqrt{\sum_i p_{h_i}}}, E_{(2)} = \left(\sum |h_i\rangle\langle h_i|_A + \sum |k_i\rangle\langle k_i|_A\right) \otimes \mathbb{I}_M$$

subject to the condition

$$\sum x_{h_i}|h_i h_i\rangle\langle h_i h_i| \geq \sum x_{g_i} E_{(2)} U_{(2)} |g_i g_i\rangle\langle g_i g_i| U^\dagger_{(2)} E_{(2)} \tag{5}$$

and the conservation of probability, viz. $\sum p_{g_i} = \sum p_{h_i}$. We claim that this choice is viable, in the sense that it satisfies the conditions (1) $E_{(2)}\left|\psi_{(3)}\right\rangle = U_{(2)}\left|\psi_{(2)}\right\rangle$, and (2) $Z^A_{(3)} \otimes \mathbb{I}_M \geq E_{(2)} U_{(2)}\left(Z^A_{(2)} \otimes \mathbb{I}_M\right) U^\dagger_{(2)} E_{(2)}$.

Proof. We must show that (1) and (2) as above hold. For (1) we observe that $E_{(2)}\left|\psi_{(3)}\right\rangle = \left|\psi_{(3)}\right\rangle$ and the statement holds by construction of $U_{(2)}$.
(2) Consider the space $\mathcal{H} = \text{span}\{|g_1 g_1\rangle, |g_2 g_2\rangle \ldots, |h_1 h_1\rangle, |h_2, h_2\rangle \ldots\}$ which is a subspace of $\mathcal{A} \otimes \mathcal{M}$ (space of Alice and the message register). One can write $\mathcal{A} \otimes \mathcal{M} = \mathcal{H} \oplus \mathcal{H}^\perp$. We separate all expressions which act on the $\mathcal{H}$ space from the rest. We start with the RHS, excluding the $U_{(2)}$'s,

$$Z^A_{(2)} \otimes \mathbb{I}_M = \underbrace{\sum x_{g_i}|g_i g_i\rangle\langle g_i g_i|}_{\text{I}} + \sum x_{g_i}|g_i\rangle\langle g_i| \otimes (\mathbb{I} - |g_i\rangle\langle g_i|) + \sum x_{k_i}|k_i\rangle\langle k_i| \otimes \mathbb{I}.$$

Note that $Z^A_{(2)} \otimes \mathbb{I}_M$ is block diagonal with respect to $\mathcal{H} \oplus \mathcal{H}^\perp$, with term I making the first block (corresponding to $\mathcal{H}$), and the rest constituting the second block. Next consider the LHS,

$$Z^A_{(3)} \otimes \mathbb{I}_M = \underbrace{\sum x_{h_i}|h_i h_i\rangle\langle h_i h_i|}_{\text{I}} + \sum x_{h_i}|h_i\rangle\langle h_i| \otimes (\mathbb{I} - |h_i\rangle\langle h_i|) + \sum x_{k_i}|k_i\rangle\langle k_i| \otimes \mathbb{I},$$

which is also block diagonal with respect to $\mathcal{H} \oplus \mathcal{H}^\perp$ and has only term I in the first block. Consequently, only on these will $U_{(2)}$ have a non-trivial action (as $U_{(2)}$ is of the form $\begin{bmatrix} U & 0 \\ 0 & \mathbb{I}_{\mathcal{H}^\perp}\end{bmatrix}$ wrt $\mathcal{H} \oplus \mathcal{H}^\perp$). Let us first evaluate the non-$\mathcal{H}$ part where we only need to apply the projector. The result after separating equations where possible is

$$\sum x_{h_i}|h_i\rangle\langle h_i| \otimes (\mathbb{I} - |h_i\rangle\langle h_i|) \geq 0, \text{ and } \sum (x_{k_i} - x_{k_i})|k_i\rangle\langle k_i| \otimes \mathbb{I} \geq 0,$$

which imply $x_{h_i} \geq 0$. The non-trivial part yields

$$\sum x_{h_i}|h_i h_i\rangle\langle h_i h_i| \geq \sum x_{g_i} E_{(2)} U_{(2)} |g_i g_i\rangle\langle g_i g_i| U^\dagger_{(2)} E_{(2)}$$

completing the proof. □

4. **Bob accepts Alice's change.** The following holds:

$$\left|\psi_{(4)}\right\rangle = \left(\sum_i \sqrt{p_{h_i}} \left|h_i h_i\right\rangle_{AB} + \sum_i \sqrt{p_{k_i}} \left|k_i k_i\right\rangle_{AB}\right) \otimes |m\rangle_M$$

$$E_{(3)} U_{(3)} = E_{(3)} U_{BM}^{\mathrm{SWP}\{\vec{h},m\}}$$

$$Z_{(4)}^A = Z_{(3)}^A \quad \text{and} \quad Z_{(4)}^B = y \sum_i |h_i\rangle \langle h_i| + \sum_i y_{k_i} |k_i\rangle \langle k_i|_B ,$$

where $E_{(3)} = \left(\sum |h_i\rangle \langle h_i| + \sum |k_i\rangle \langle k_i|\right)_B \otimes \mathbb{I}_M$.

Proof. We have to prove: (1) $E_{(3)} \left|\psi_{(4)}\right\rangle = U_{(3)} \left|\psi_{(3)}\right\rangle$ and (2) $Z_{(4)}^B \otimes \mathbb{I}_M \geq E_{(3)} U_{(3)} \left(Z_{(3)}^B \otimes \mathbb{I}_M\right) U_{(3)}^\dagger E_{(3)}$. The first equality (1) can be shown by a direct application of $U^\dagger E$ on $\left|\psi_{(4)}\right\rangle$, where E, U denote $E_{(3)}$ and $U_{(3)}$, respectively, in this proof for ease of notation.

(2) Note that

$$EU\left(\mathbb{I}_B^{\{\vec{g},m\}} \otimes \mathbb{I}_M^{\{\vec{h},\vec{g},\vec{k},m\}}\right) U^\dagger E = EU\left(\mathbb{I}_B^{\{m\}} \otimes \mathbb{I}_M^{\{\vec{h},\vec{g},\vec{k},m\}}\right) U^\dagger E + E\left(\mathbb{I}_B^{\{\vec{g}\}} \otimes \mathbb{I}_M^{\{\vec{h},\vec{g},\vec{k},m\}}\right) E$$

$$= EU\left(\mathbb{I}_B^{\{m\}} \otimes \mathbb{I}_M^{\{\vec{h},m\}}\right) U^\dagger E = \sum |h_i\rangle \langle h_i| \otimes \mathbb{I}_M^{\{m\}}.$$

Since the other term in $Z_{(3)}^B \otimes \mathbb{I}$ is not in the action space of U it follows that

$$EU(Z_{(3)}^B \otimes \mathbb{I})U^\dagger E = y \sum |h_i\rangle \langle h_i| \otimes \mathbb{I}_M^{\{m\}} + \sum y_{k_i} |k_i\rangle \langle k_i| \otimes \mathbb{I}_M.$$

It only remains to show that $Z_{(4)}^B \otimes \mathbb{I}_M \geq EU\left(Z_{(3)}^B \otimes \mathbb{I}_M\right) U^\dagger E$ which holds as $y \sum |h_i\rangle \langle h_i| \otimes \mathbb{I}_M \geq y \sum |h_i\rangle \langle h_i| \otimes \mathbb{I}_M^{\{m\}}$ and the y_{k_i} term is common. □

The above analysis can be distilled in the following Theorem 31, which includes the so-called *TEF constraints*, i.e. the conditions of our framework, which—when satisfied by some unitary—permit the transformation of the TDPG into an explicit WCF protocol by means of this unitary. The proof of the theorem is a straightforward consequence of the above.

Theorem 31. *For an x-transition (where Alice performs the non-trivial step)*

$$\sum_{i=1}^{n_k} p_{k_i} [\![x_{k_i}]\!] + \sum_{i=1}^{n_g} p_{g_i} [\![x_{g_i}]\!] \to \sum_{i=1}^{n_h} p_{h_i} [\![x_{h_i}]\!] + \sum_{i=1}^{n_k} p_{k_i} [\![x_{k_i}]\!]$$

to be implementable under the TDPG-to-Explicit-protocol Framework (TEF) it suffices to find a $U_{(2)}$ that satisfies the inequality

$$\sum_{i=1}^{n_h} x_{h_i} |h_i h_i\rangle \langle h_i h_i|_{AM} \geq \sum_{i=1}^{n_g} x_{g_i} E_{(2)}^h U_{(2)} |g_i g_i\rangle \langle g_i g_i|_{AM} U_{(2)}^\dagger E_{(2)}^h \tag{6}$$

and the honest action constraint $U_{(2)} |v\rangle = |w\rangle$, where $|h_i\rangle$ and $|g_i\rangle$ are orthonormal basis vectors,

$$|v\rangle = \mathcal{N}\left(\sum \sqrt{p_{g_i}} |g_i g_i\rangle_{AM}\right) \quad \textit{and} \quad |w\rangle = \mathcal{N}\left(\sum \sqrt{p_{h_i}} |h_i h_i\rangle_{AM}\right)$$

for $\mathcal{N}(|\psi\rangle) = |\psi\rangle / \sqrt{\langle\psi|\psi\rangle}$, $E^h = \left(\sum_{i=1}^{n_h} |h_i\rangle \langle h_i|_A + \sum |k_i\rangle \langle k_i|_A\right) \otimes \mathbb{I}_M$ with $U_{(2)}$'s non-trivial action restricted to span $\left\{\{|g_i g_i\rangle_{AM}\}, \{|h_i h_i\rangle_{AM}\}\right\}$, and $|k_i\rangle$ correspond to the points that are left unchanged in the transition.

Theorem 31 also leads to Theorem 20 by showing that the EBM condition implies that the inequality appearing in Theorem 31 can be satisfied. There are two difficulties, the first is that in Equation (6) there is a projector and the second is that the matrices have a certain dimension; neither of the two holds for EBM. We address these issues in Section 5, by arguing that the projector can be seen as the limiting case of one of the matrices having diverging eigenvalues (see Section 5.1). We also show in Appendix E that it is sufficient to restrict to orthogonal matrices, and—as C. Mochon proved—in particular to matrices of size $n_h + n_g - 1$ (see Lemma 146). Together, these establish Theorem 20.

The set of functions satisfying the TEF constraints is the closure of the set of EBM functions which, in turn, is the set of valid functions (see Appendix B). Thus, using the TEF, we can directly associate valid games with WCF protocols, granted that the non-trivial unitary, $U_{(2)}$, can be found. This allows us to skip the notion of strictly valid functions (see Definition 127) and makes our approach simpler, in some sense, compared to the previous one which relied on strictly valid functions.

3.1.1 Special case: the blinkered unitary

So far we have not specified the non-trivial $U_{(2)}$—which we call U from now, and, similarly, $E^h_{(2)} = E$—beyond requiring it to have a certain action on the honest state. We now define the *blinkered unitary*, an important class of such unitaries, as

$$U = |w\rangle\langle v| + |v\rangle\langle w| + \sum_i |v_i\rangle\langle v_i| + \sum_i |w_i\rangle\langle w_i| + \mathbb{I}^{\text{outside }\mathcal{H}},$$

where $\mathcal{H} = \text{span}\{|g_1g_1\rangle, |g_2g_2\rangle \ldots, |h_1h_1\rangle, |h_2, h_2\rangle \ldots\}$. We can ignore the last term and restrict our analysis to the $\mathcal{H}$-operator space, where $|v\rangle, \{|v_i\rangle\}$ form a complete orthonormal basis with respect to $\text{span}\{|g_ig_i\rangle\}$, and so do $|w\rangle, \{|w_i\rangle\}$ for $\text{span}\{|h_ih_i\rangle\}$. The blinkered unitary can be used to implement the two non-trivial operations of the set of basic moves, namely the merge (see Example 23) and the split (see Example 24).

- Merge: $g_1, g_2 \longrightarrow h_1$
  From the very definitions we construct

  $$|v\rangle = \frac{\sqrt{p_{g_1}}|g_1g_1\rangle + \sqrt{p_{g_2}}|g_2g_2\rangle}{N}, \; |v_1\rangle = \frac{\sqrt{p_{g_2}}|g_1g_1\rangle - \sqrt{p_{g_1}}|g_2g_2\rangle}{N}, \; |w\rangle = |h_1h_1\rangle$$

  with $N = \sqrt{p_{g_1} + p_{g_2}}$ and $U = |w\rangle\langle v| + |v\rangle\langle w| + |v_1\rangle\langle v_1| = U^\dagger$. We need

  $$EU|g_1g_1\rangle = \frac{\sqrt{p_{g_1}}|w\rangle}{N} \text{ and } EU|g_2g_2\rangle = \frac{\sqrt{p_{g_2}}|w\rangle}{N},$$

  since the constraint is $x_h|h_1h_1\rangle\langle h_1h_1| \geq \sum x_{g_i}EU|g_ig_i\rangle\langle g_ig_i|U^\dagger E$; it becomes $x_h \geq \frac{p_{g_1}x_{g_1}+p_{g_2}x_{g_2}}{N^2}$, which is precisely the merge condition (see Example 23).

- Split: $g_1 \longrightarrow h_1, h_2$
  Here, we construct

  $$|v\rangle = |g_1g_1\rangle, \; |w\rangle = \frac{\sqrt{p_{h_1}}|h_1h_1\rangle + \sqrt{p_{h_2}}|h_2h_2\rangle}{N}, \; |w_1\rangle = \frac{\sqrt{p_{h_2}}|h_1h_1\rangle - \sqrt{p_{h_1}}|h_2h_2\rangle}{N}$$

  with $N = \sqrt{p_{h_1} + p_{h_2}}$ and $U = |v\rangle\langle w| + |w\rangle\langle v| + |w_1\rangle\langle w_1| = U^\dagger$. We evaluate $EU|g_1g_1\rangle = |w\rangle$ which we substitute into the constraint to obtain

  $$x_{h_1}|h_1h_1\rangle\langle h_1h_1| + x_{h_2}|h_2h_2\rangle\langle h_2h_2| - x_{g_1}|w\rangle\langle w| \geq 0.$$

This yields the matrix equation

$$\begin{aligned}
&\begin{bmatrix} x_{h_1} & \\ & x_{h_2} \end{bmatrix} - \frac{x_{g_1}}{N^2}\begin{bmatrix} p_{h_1} & \sqrt{p_{h_1}p_{h_2}} \\ \sqrt{p_{h_1}p_{h_2}} & p_{h_2} \end{bmatrix} \geq 0 \\
&\mathbb{I} \geq \frac{x_{g_1}}{N^2}\begin{bmatrix} \frac{p_{h_1}}{x_{h_1}} & \sqrt{\frac{p_{h_1}}{x_{h_1}}\frac{p_{h_2}}{x_{h_2}}} \\ \sqrt{\frac{p_{h_1}}{x_{h_1}}\frac{p_{h_2}}{x_{h_2}}} & \frac{p_{h_2}}{x_{h_2}} \end{bmatrix} \\
&\frac{x_{g_1}}{N^2}\left(\frac{p_{h_1}}{x_{h_1}} + \frac{p_{h_2}}{x_{h_2}}\right) \leq 1,
\end{aligned}$$

where in the first step we used the fact that for $F > 0$, $F - M \geq 0 \Rightarrow \mathbb{I} - \sqrt{F}^{-1}M\sqrt{F}^{-1} \geq 0$, and the last equation is obtained by writing the matrix as $\langle\psi|\,|\psi\rangle$, and then demanding $1 \geq \langle\psi|\psi\rangle$. This last equation is exactly the split condition (see Example 24).

The above two conditions can be readily generalized for an $m \longrightarrow 1$ point merge and a $1 \longrightarrow n$ points split, respectively, by simply constructing the appropriate vectors (see Appendix C). Furthermore, for a

- general $m \longrightarrow n$: $g_1, g_2 \dots g_m \longrightarrow h_1, h_2 \dots h_n$ transition,
  we can show that one obtains the constraint $\frac{1}{\sum_{i=1}^m p_{g_i} x_{g_i}} \geq \sum_{i=1}^n p_{h_i}\frac{1}{x_{h_i}}$, by using the appropriate blinkered unitary (see Appendix C).

This class of unitaries is enough to convert the $1/6$ game into an explicit protocol, but falls short for point games going beyond this bias; the general $m \longrightarrow n$ blinkered transition effectively behaves like an $m \longrightarrow 1$ merge followed by a $1 \longrightarrow n$ split, which are insufficient to break the $1/6$ limit. Next, we show how to construct the unitaries for a WCF protocol approaching bias $1/10$.

3.2 Bias $1/10$ game and protocol

In Section 2.5 we presented the family of TIPGs achieving bias $\epsilon(k) = 1/(4k+2)$, where k is the number of points involved in the non-trivial step, which was proposed by C. Mochon. Here, we consider the game with $k = 2$ and by means of the TEF we construct the corresponding WCF protocol with bias $1/10$.

We assume an equally spaced n-point lattice given by $x_j = x_0 + j\delta x$ where $\delta x = \delta y$ is small and x_0 will be determined through the constraints[12]; similarly $y_j = y_0 + j\delta y$ and we also define $\Gamma_{k+1} = y_{n-k} = x_{n-k}$. Let $P(x_j)$ be the probability weight associated with the point $[x_j, 0]$ which is such that

$$\sum_{j=1}^{n} P(x_j) = \frac{1}{2} \text{ and } \sum_{j=1}^{n} \frac{P(x_j)}{x_j} = \frac{1}{2}.$$

Similarly with the point $(0, y_j)$ we associate $P(y_j)$ where $y_j = x_j$ as we also assume that $x_0 = y_0$. These choices explicitly impose symmetry between Alice and Bob which in turn means that we only have to do the analysis for one of them.

With respect to Figure 7 we use the assignment that C. Mochon employed (see Definition 32) with $f(y_i) = (y_{-2} - y_i)(\Gamma_1 - y_i)(\Gamma_2 - y_i)$ as $\frac{f(y_j)c(x_j)}{\prod_{k\neq j}(y_k - y_j)}$. The probabilities become

$$\begin{aligned}
P_2(y_{j+2}) &= \frac{-f(y_{j+2})c(x_j)}{4\cdot 3(\delta y)^2 y_{j+2}}, \; P_1(y_{j+1}) = \frac{-f(y_{j+1})c(x_j)}{3\cdot 2(\delta y)^2 y_{j+1}}, \\
P_1(x_j) &= \frac{-f(y_{j-1})c(x_j)}{3\cdot 2(\delta y)^2 y_{j-1}}, \; P_2(x_j) = \frac{-f(y_{j-2})c(x_j)}{4\cdot 3(\delta y)^2 y_{j-2}}, \; P(x_j) = \frac{f(0)c(x_j)\delta y}{y_{j+2}y_{j+1}y_{j-1}y_{j-2}}
\end{aligned}$$

[12] Essentially, x_0 provides a bound on P_B^*.

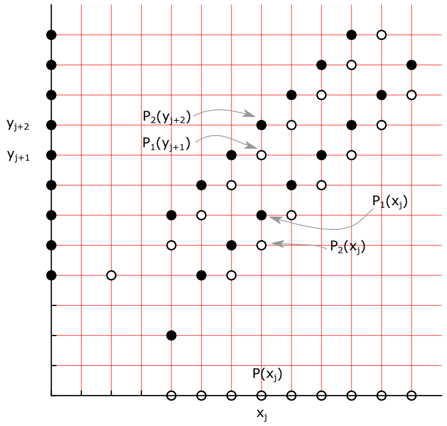


Figure 7: 1/10-bias TIPG: The $3 \longrightarrow 2$ move

where we added the minus sign to account for the fact that f is negative for coordinates between y_{-2} and Γ_1. Imposing the symmetry constraint $P_1(y_j) = P_1(x_j)$ we get $c(x_j) = \frac{c_0 f(x_j)}{x_j}$, where c_0 is a constant. Similarly, the symmetry constraint for P_2 entails $P_2(y_j) = P_2(x_j)$. Finally, we can evaluate $P(x_j) = \frac{c_0 x_0 (x_0 - x_j)}{x_j^5}\delta x + O(\delta x^2)$ which, in the limit $\delta x \longrightarrow 0$, means that

$$\sum P(x_j) = \frac{1}{2} = \sum \frac{P(x_j)}{x_j} \longrightarrow \int_{x_0}^{\Gamma} \frac{(x_0 - x)dx}{x^5} = \int_{x_0}^{\Gamma} \frac{(x_0 - x)dx}{x^6}.$$

This evaluates to

$$x_0 \int_{x_0}^{\Gamma} \left(\frac{1}{x^5} - \frac{1}{x^6}\right) dx = \int_{x_0}^{\Gamma} \left(\frac{1}{x^4} - \frac{1}{x^5}\right) dx \Rightarrow x_0 = \frac{3}{5} \implies \epsilon = \frac{3}{5} - \frac{1}{2} = \frac{1}{10}.$$

Below we consider the moves involved in this game in order to construct the unitaries of the corresponding protocol and we prove that they are valid transitions. For ease of notation, henceforth, we use $|g_1\rangle$ instead of $|g_1 g_1\rangle$, and similarly $|h_1\rangle$ instead of $|h_1 h_1\rangle$.

3.2.1 The $3 \longrightarrow 2$ move and its validity

Here, we consider the $3 \longrightarrow 2$ move, i.e., a transition from 3 initial to 2 final points.

Recall that

$$|v\rangle = \frac{\sqrt{p_{g_1}}\,|g_1\rangle + \sqrt{p_{g_2}}\,|g_2\rangle + \sqrt{p_{g_3}}\,|g_3\rangle}{N_g}$$

and let

$$|v_1\rangle = \frac{\sqrt{p_{g_3}}\,|g_2\rangle - \sqrt{p_{g_2}}\,|g_3\rangle}{N_{v_1}}, \quad |v_2\rangle = \frac{-\frac{(p_{g_2}+p_{g_3})}{\sqrt{p_{g_1}}}\,|g_1\rangle + \sqrt{p_{g_2}}\,|g_2\rangle + \sqrt{p_{g_3}}\,|g_3\rangle}{N_{v_2}}$$

where $N_{v_1}^2 = p_{g_3} + p_{g_2}$ and $N_{v_2}^2 = \frac{(p_{g_2}+p_{g_3})^2}{p_{g_1}} + p_{g_2} + p_{g_3}$. Also,

$$|w\rangle = \frac{\sqrt{p_{h_1}}\,|h_1\rangle + \sqrt{p_{h_2}}\,|h_2\rangle}{N_h} \text{ and } |w_1\rangle = \frac{\sqrt{p_{h_2}}\,|h_1\rangle - \sqrt{p_{h_1}}\,|h_2\rangle}{N_h}.$$

Now we define

$$\left|v_1'\right\rangle = \cos\theta\left|v_1\right\rangle + \sin\theta\left|v_2\right\rangle \text{ and } \left|v_2'\right\rangle = \sin\theta\left|v_1\right\rangle - \cos\theta\left|v_2\right\rangle,$$

where $\cos\theta \approx 1$, and the full unitary as

$$U = |w\rangle\langle v| + \left(\alpha\left|v_1'\right\rangle + \beta\left|w_1\right\rangle\right)\left\langle v_1'\right| + \left|v_2'\right\rangle\left\langle v_2'\right| + \left(\beta\left|v_1'\right\rangle - \alpha\left|w_1\right\rangle\right)\left\langle w_1\right| + |v\rangle\langle w|,$$

where $|\alpha|^2 + |\beta|^2 = 1$ for $\alpha, \beta \in \mathbb{C}$[13]. We need terms of the form $EU|g_i\rangle$ with $E = \mathbb{I}^{\{h_i\}}$. This entails that EU acts on the $\{|g_i\rangle\}$ space as

$$EUE_g = |w\rangle\langle v| + \beta\left|w_1\right\rangle\left\langle v_1'\right| = |w\rangle\langle v| + \beta\left|w_1\right\rangle\left(\cos\theta\left\langle v_1\right| + \sin\theta\left\langle v_2\right|\right),$$

where E_g is the projector on the $\{|g_i\rangle\}$ space. Consequently we have

$$\begin{aligned}
EU|g_1\rangle &= \frac{\sqrt{p_{g_1}}}{N_g}|w\rangle + \left[\cos\theta\cdot 0 - \sin\theta\frac{p_{g_2}+p_{g_3}}{\sqrt{p_{g_1}}N_{v_2}}\right]\beta\left|w_1\right\rangle\\
EU|g_2\rangle &= \frac{\sqrt{p_{g_2}}}{N_g}|w\rangle + \left[\cos\theta\frac{\sqrt{p_{g_3}}}{N_{v_1}} + \sin\theta\frac{\sqrt{p_{g_2}}}{N_{v_2}}\right]\beta\left|w_1\right\rangle\\
EU|g_3\rangle &= \frac{\sqrt{p_{g_3}}}{N_g}|w\rangle + \left[-\cos\theta\frac{\sqrt{p_{g_2}}}{N_{v_1}} + \sin\theta\frac{\sqrt{p_{g_3}}}{N_{v_2}}\right]\beta\left|w_1\right\rangle.
\end{aligned}$$

Recall that the constraint equation was

$$\sum x_{h_i}|h_i\rangle\langle h_i| - \sum x_{g_i}EU|g_i\rangle\langle g_i|U^\dagger E \geq 0$$

where the first sum becomes

$$\begin{bmatrix} \langle x_h\rangle & \frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1}-x_{h_2}) \\ \text{h.c.} & \frac{p_{h_2}x_{h_1}+p_{h_1}x_{h_2}}{N_h^2} \end{bmatrix}$$

in the $|w\rangle, |w_1\rangle$ basis. Since we plan to use the 3 → 2 move with one point on the axis, we take $x_{g_1} = 0$. Consequently we only need to evaluate

$$\begin{aligned}
x_{g_2}EU|g_2\rangle\langle g_2|U^\dagger E &\doteq x_{g_2}\begin{bmatrix} \frac{p_{g_2}}{N_g^2} & \beta\left(\cos\theta\frac{\sqrt{p_{g_3}p_{g_2}}}{N_gN_{v_1}} + \sin\theta\frac{p_{g_2}}{N_gN_{v_2}}\right) \\ \text{h.c.} & \left(\cos\frac{\sqrt{p_{g_3}}}{N_{v_1}} + \sin\theta\frac{\sqrt{p_{g_2}}}{N_{v_2}}\right)^2|\beta|^2 \end{bmatrix}\\
x_{g_3}EU|g_3\rangle\langle g_3|U^\dagger E &\doteq x_{g_3}\begin{bmatrix} \frac{p_{g_3}}{N_g^2} & \beta\left(-\cos\theta\frac{\sqrt{p_{g_2}p_{g_3}}}{N_gN_{v_1}} + \sin\theta\frac{p_{g_3}}{N_gN_{v_2}}\right) \\ \text{h.c.} & \left(-\cos\frac{\sqrt{p_{g_2}}}{N_{v_1}} + \sin\frac{\sqrt{p_{g_3}}}{N_{v_2}}\right)^2|\beta|^2 \end{bmatrix}
\end{aligned}$$

which means that the constraint equation becomes

$$\begin{bmatrix} \langle x_h\rangle - \left\langle x_g\right\rangle & \frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1}-x_{h_2}) - \beta\cos\theta\frac{\sqrt{p_{g_2}p_{g_3}}}{N_gN_{v_1}}(x_{g_2}-x_{g_3}) - \beta\sin\theta\left\langle x_g\right\rangle\frac{N_g}{N_{v_2}} \\ \text{h.c.} & \frac{p_{h_2}x_{h_1}+p_{h_1}x_{h_2}}{N_h^2} - |\beta|^2\left[\frac{\cos^2\theta}{N_{v_1}^2}(p_{g_3}x_{g_2}+p_{g_2}x_{g_3}) + \frac{\sin^2\theta}{\left(N_{v_2}^2/N_g^2\right)}\left\langle x_g\right\rangle + \frac{2\cos\theta\sin\theta\sqrt{p_{g_3}p_{g_2}}}{N_{v_1}N_{v_2}}(x_{g_2}-x_{g_3})\right] \end{bmatrix} \geq 0.$$

Since this transition is average non-decreasing viz. $\langle x_h\rangle - \left\langle x_g\right\rangle \geq 0$ (see Lemma 135 and Lemma 33), we set the off-diagonal elements of the matrix above to zero and show that the second diagonal element is

[13] There is some freedom in choosing U in the sense that $\alpha|v\rangle + \beta|w_1\rangle$ would also work instead of $\alpha\left|v_1'\right\rangle + \beta|w_1\rangle$(in that case $|v\rangle\langle w|$ should be replaced by $|v_1\rangle\langle w|$), as these do not influence the constraint equation.

positive. Setting the off-diagonal to zero one can obtain θ by solving the quadratic equation in terms of β although the expression is not particularly pretty. To establish existence and positivity we need to simplify our expressions.

So far, everything was exact. To proceed, we write $\theta\frac{N_g}{N_{v_2}} = O(\delta y)$ at most (where $\delta y = \delta x$ is the lattice spacing) and we take δy to be small. Thus, to first order in $\theta\frac{N_g}{N_{v_2}}$, the constraints become

$$\frac{\frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1} - x_{h_2}) - \beta\frac{\sqrt{p_{g_2}p_{g_3}}}{N_g N_{v_1}}(x_{g_2} - x_{g_3})}{\beta\left\langle x_g\right\rangle} = \theta\frac{N_g}{N_{v_2}} + O(\delta y^2)$$

and
$$\frac{p_{h_2}x_{h_1} + p_{h_1}x_{h_2}}{N_h^2} - |\beta|^2\left[\frac{p_{g_3}x_{g_2} + p_{g_2}x_{g_3}}{N_{v_1}^2} + 2\theta\frac{N_g}{N_{v_2}}\frac{\sqrt{p_{g_3}p_{g_2}}}{N_g N_{v_1}}(x_{g_2} - x_{g_3})\right] + O(\delta y^2) \geq 0.$$

If our claim is wrong when we evaluate $\theta\frac{N_g}{N_{v_2}}$, we will get zero order terms but as we show later, indeed, $\theta\frac{N_g}{N_{v_2}} = O(\delta y^2)$. With respect to Figure 7 we have

$$P_2(y_{j+2}) = p_{h_2} = \frac{-f(y_{j+2})}{4\cdot 3\delta y^2 y_{j+2}},\; P_1(y_{j+1}) = p_{g_3} = \frac{-f(y_{j+1})}{3\cdot 2\delta y^2 y_{j+1}}$$
$$P_1(x_j) = p_{h_1} = \frac{-f(y_{j-1})}{3\cdot 2\delta y^2 y_{j-1}},\; P_2(x_j) = p_{g_2} = \frac{-f(y_{j-2})}{4\cdot 3\delta y^2 y_{j-2}},\; P(x_j) = p_{g_1} = \frac{f(0)\delta y}{y_{j+2}y_{j+1}y_{j-1}y_{j-2}},$$

where we assumed $f(0) > 0$ and $f(y) < 0$ for $y > y_0'$, $y_0' = y_0 + \delta y$, and we scaled by δy. We now convert all expressions to first order in δy:

$$f(y_{j+m}) = f(y_j) + \frac{\partial f}{\partial y}m\delta y + O(\delta y^2) \Rightarrow \frac{1}{y_{j+m}} = \frac{1}{y_j} - m\frac{\delta y}{y_j^2} + O(\delta y^2),$$

where $\frac{\partial f}{\partial y}$ is $\frac{\partial f(y)}{\partial y}|_{y_j}$. We define and evaluate

$$P_k^m = \frac{-f(y_{j+m})}{k\delta y^2 y_{j+m}} = \frac{1}{ky_j\delta y^2}\left[-f - m\delta y\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right) + O(\delta y^2)\right],$$

where f means $f(y_j)$. In this notation

$$p_{h_2} = P_{12}^2,\; p_{h_1} = P_6^{-1} \quad\text{and}\quad p_{g_2} = P_{12}^{-2},\; p_{g_3} = P_6^1.$$

With an eye at the off-diagonal condition we evaluate

$$P_{k_1}^{m_1}P_{k_2}^{m_2} = \frac{1}{k_1k_2}\left(\frac{1}{y_j\delta y^2}\right)^2\left[f^2 + f\delta y\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right)(m_1 + m_2) + O(\delta y^2)\right]$$

and
$$P_{k_1}^{m_1} + P_{k_2}^{m_2} = \frac{1}{y_j\delta y^2}\left[-\left(\frac{1}{k_1} + \frac{1}{k_2}\right)f - \left(\frac{m_1}{k_1} + \frac{m_2}{k_2}\right)\delta y\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right) + O(\delta y^2)\right].$$

Moreover, we have

$$\sqrt{p_{h_1}p_{h_2}} = \sqrt{P_{12}^2P_6^{-1}} = \frac{1}{y_j\delta y^2}\sqrt{\frac{1}{12\cdot 6}\left[f^2 + f\delta y\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right) + O(\delta y^2)\right]}$$
$$N_h^2 = P_{12}^2 + P_6^{-1} = \frac{1}{4y_j\delta y^2}\left[-f + O(\delta y^2)\right],$$

and similarly

$$\sqrt{p_{g_2}p_{g_3}} = \sqrt{P_{12}^{-2}P_6^1} = \frac{1}{y_j\delta y^2}\sqrt{\frac{1}{12\cdot 6}\left[f^2 - f\delta y\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right) + O(\delta y^2)\right]}$$

$$N_g^2 = P_{12}^{-2} + P_6^1 + p_{g_1} = \frac{1}{4y_j\delta y^2}\left[-f + O(\delta y^2)\right] \text{ and } N_{v_1}^2 = \frac{1}{4y_j\delta y^2}\left[-f + O(\delta y^2)\right],$$

where we already neglected the terms that contribute to the ratio $\frac{N_g}{N_{v_2}}$ in higher than first order. Actually, for $\beta = 1$

$$\theta\frac{N_g}{N_{v_2}} = \frac{4\sqrt{\frac{1}{12\cdot 6}}(-3\delta y)\left[f(\cancel{1} + \frac{\delta y}{2f}\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right)) - f(\cancel{1} - \frac{\delta y}{2f}\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right)) + O(\delta y^2)\right]}{\langle x_g\rangle} = O(\delta y^2).$$

This shows that to first order the off-diagonal term is zero for $\theta = 0$. Now, we show that the second diagonal element is positive to first order in δy. Using the fact that $\theta\frac{N_g}{N_{v_2}} = O(\delta y^2)$, the positivity condition reads

$$\frac{p_{h_2}x_{h_1} + p_{h_1}x_{h_2}}{N_h^2} - \frac{p_{g_3}x_{g_2} + p_{g_2}x_{g_3}}{N_{v_1}^2} + O(\delta y^2) \geq 0,$$

which, in turn, becomes

$$\frac{P_{12}^2 y_{j-1} + P_6^{-1}y_{j+2}}{N_h^2} - \frac{P_6^1 y_{j-2} + P_{12}^{-2}y_{j+1}}{N_{v_1}^2} + O(\delta y^2) = 2\delta y + O(\delta y^2) \geq 0.$$

This establishes the validity of the 3 → 2 transition for a closely spaced lattice. Note that only the proof of validity was done perturbatively to first order in δy. The unitary itself is known exactly, as θ can be obtained by solving the quadratic. Using $f(y) = (y_0' - y)(\Gamma_1 - y)(\Gamma_2 - y)$ we can implement the last two moves in Figure 7 as they constitute a 3 → 1 and a 2 → 1 merge. The only remaining task is to implement the 2 → 2 move of the last step, because previously we assumed $\sqrt{p_{g_2}} \neq 0$.

3.2.2 The 2 → 2 move and its validity

We claim that the 2 → 2 move can be implemented using

$$U = |w\rangle\langle v| + (\alpha|v\rangle + \beta|w_1\rangle)\langle v_1| + |v\rangle\langle w| + (\beta|v\rangle - \alpha|w_1\rangle)\langle w_1|$$

where as before $|\alpha|^2 + |\beta|^2 = 1$,

$$|v\rangle = \frac{1}{N_g}\left(\sqrt{p_{g_1}}|g_1\rangle + \sqrt{p_{g_2}}|g_2\rangle\right), |w\rangle = \frac{1}{N_h}\left(\sqrt{p_{h_1}}|h_1\rangle + \sqrt{p_{h_2}}|h_2\rangle\right),$$

$$|v_1\rangle = \frac{1}{N_g}\left(\sqrt{p_{g_2}}|g_1\rangle - \sqrt{p_{g_1}}|g_2\rangle\right) \text{ and } |w_1\rangle = \frac{1}{N_h}\left(\sqrt{p_{h_2}}|h_1\rangle - \sqrt{p_{h_1}}|h_2\rangle\right).$$

We evaluate the constraint equation using

$$EU|g_1\rangle = \frac{\sqrt{p_{g_1}}|w\rangle + \beta e^{-i\phi_g}e^{i\phi_h}\sqrt{p_{g_2}}|w_1\rangle}{N_g}, \ EU|g_2\rangle = \frac{\sqrt{p_{g_2}}|w\rangle - \beta e^{-i\phi_g}e^{i\phi_h}\sqrt{p_{g_1}}|w_1\rangle}{N_g},$$

and

$$EU|g_1\rangle\langle g_1|U^\dagger E = \frac{1}{N_g^2}\begin{array}{c|cc} & \langle w| & \langle w_1| \\ \hline |w\rangle & p_{g_1} & \beta e^{i(\phi_h - \phi_g)}\sqrt{p_{g_2}p_{g_1}} \\ |w_1\rangle & \text{h.c.} & |\beta|^2 p_{g_2} \end{array}$$

as

$$\begin{bmatrix} \langle x_h \rangle - \langle x_g \rangle & \frac{1}{N_g^2}\left[\sqrt{p_{h_1}p_{h_2}}(x_{h_1}-x_{h_2}) - \beta\sqrt{p_{g_1}p_{g_2}}(x_{g_1}-x_{g_2})\right] \\ \text{h.c.} & \frac{1}{N_g^2}\left[p_{h_2}x_{h_1}+p_{h_1}x_{h_2} - |\beta|^2\left(p_{g_2}x_{g_1}+p_{g_1}x_{g_2}\right)\right] \end{bmatrix} \geq 0,$$

where we absorbed the phase freedom in β, a free parameter, which will be fixed shortly. We use the same strategy as above and take the first diagonal element to be zero. We must show that

$$\sqrt{\frac{p_{h_1}p_{h_2}}{p_{g_1}p_{g_2}}}\frac{(x_{h_1}-x_{h_2})}{(x_{g_1}-x_{g_2})} = \beta \leq 1, \text{ and } \frac{1}{N_g^2}\left[p_{h_2}x_{h_1}+p_{h_1}x_{h_2} - |\beta|^2\left(p_{g_2}x_{g_1}+p_{g_1}x_{g_2}\right)\right] \geq 0.$$

For this transition $f(y_{j-2}) = 0$, which we use to write

$$f(y_{j+k}) = \left.\frac{\partial f}{\partial y}\right|_{y_{j-2}}(k+2)\delta y = -(k+2)\alpha\delta y, \text{ with } \alpha = -\left.\frac{\partial f}{\partial y}\right|_{y_{j-2}} = (\Gamma_1 - y_{j-2})(\Gamma_2 - y_{j-2}).$$

From Figure 8 we have

$$p_{h_1} = P_1(x_j) = \frac{-f(y_{j-1})}{3\cdot 2\delta y^2 y_{j-1}} = \frac{\alpha + O(\delta y)}{6\delta y y_j}, \; p_{h_2} = P_2(y_{j+2}) = \frac{-f(y_{j+2})}{4\cdot 3\delta y^2 y_{j+2}} = \frac{\alpha + O(\delta y)}{3\delta y y_j}$$

$$x_{h_1} = y_{j-1}, \; x_{h_2} = y_{j+2}$$

while

$$p_{g_1} = P(x_j) = \frac{f(0)\delta y}{y_{j+2}y_{j+1}y_{j-1}y_{j-2}} = \frac{f(0)\delta y + O(\delta y^2)}{y_j^4}, \; p_{g_2} = P_1(y_{j+1}) = \frac{-f(y_{j+1})}{3\cdot 2\delta y^2 y_{j+1}} = \frac{\alpha + O(\delta y)}{2\delta y y_j}$$

$$x_{g_1} = 0, \; x_{g_2} = y_{j+1}.$$

This entails

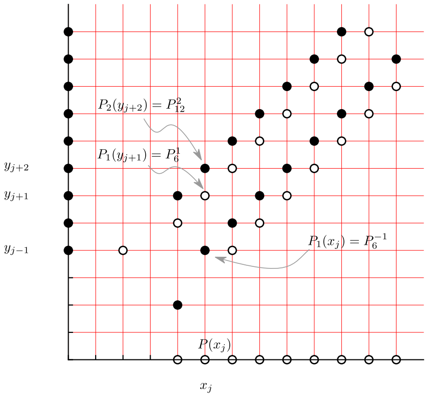


Figure 8: The first 2 ⟶ 2 transition

$$\beta = \sqrt{\frac{p_{h_1}p_{h_2}}{p_{g_1}p_{g_2}}}\frac{(x_{h_1}-x_{h_2})}{(x_{g_1}-x_{g_2})} = \sqrt{\frac{y_0'\alpha + O(\delta y)}{f(0)}} = \sqrt{\frac{(\Gamma_1 - y_{j-2})(\Gamma_2 - y_{j-2}) + O(\delta y)}{\Gamma_1\Gamma_2}} \leq 1,$$

where we used $f(0) = y_0'\Gamma_1\Gamma_2$ and the fact that δy is small compared to Γs. Analogously, for the second condition we have

$$\frac{1}{N_g^2}\left[p_{h_2}x_{h_1}+p_{h_1}x_{h_2} - |\beta|^2\left(p_{g_2}x_{g_1}+p_{g_1}x_{g_2}\right)\right] \geq \frac{1}{N_g^2}\left[p_{h_2}x_{h_1}+p_{h_1}x_{h_2} - p_{g_2}x_{g_1}\right]$$

$$= \frac{1}{2\delta y N_g^2}\left[\alpha + O(\delta y)\right] = \frac{1}{2\delta y N_g^2}\left[(\Gamma_1 - y_{j-2})(\Gamma_2 - y_{j-2}) + O(\delta y)\right] \geq 0,$$

where the last step holds for δy small enough. The $2 \longrightarrow 2$ move corresponding to the leftmost (see Figure 9) and bottom-most set of points can be shown to be implementable similarly.

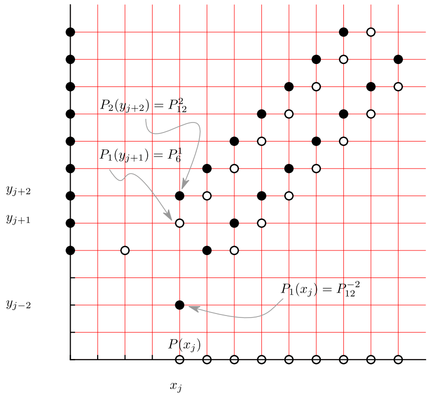


Figure 9: The final $2 \longrightarrow 2$ transition.

4 Approaching bias $\epsilon(k) = 1/(4k+2)$: an algebraic solution

While we succeeded at constructing the unitaries involved in the bias $1/10$ protocol, we did not follow a systematic method. Here, we construct the unitaries corresponding to the valid functions that characterize C. Mochon's point games (see Lemma 29). These, together with the TEF, allow us to construct explicit WCF protocols with bias approaching $\epsilon(k) = 1/(4k+2)$ for arbitrary integers $k > 0$. In this section, the unitaries we construct are additionally real (i.e. orthogonal matrices). As we shall see in Section 5, since this restriction anyway does not lead to a loss of generality,[14] all vector spaces considered here are over the field of real numbers.

Notation:

- For a Hermitian matrix A with spectral decomposition (including zero eigenvalues) $A = \sum_i a_i \left|i\right\rangle \left\langle i\right|$, we define the pseudo-inverse or the generalized inverse of A as $A^{\dashv} := \sum_{i:|a_i|>0} a_i^{-1} \left|i\right\rangle \left\langle i\right|$.

- We write functions t with finite support in the following two ways (unless otherwise stated): (1) as $t = \sum_{i=1}^{n} p_i \llbracket x_i \rrbracket$ where we assume $p_i > 0$ for all $i \in \{1, 2 \ldots n\}$ and that $x_i \neq x_j$ for $i \neq j$ and (2) as $t = \sum_{i=1}^{n_h} p_{h_i} \llbracket x_{h_i} \rrbracket - \sum_{i=1}^{n_g} p_{g_i} \llbracket x_{g_i} \rrbracket$ where p_{h_i} and p_{g_i} are strictly positive and x_{h_i} and x_{g_i} are all distinct.

4.1 The f−assignments

We start with the definition of f-assignments tailored to the purpose of the analysis that follows.

Definition 32 (f-assignments)**.** Given a set of real numbers $0 \leq x_1 < x_2 \cdots < x_n$ and a polynomial of degree at most $n-2$ satisfying $f(-\lambda) \geq 0$ for all $\lambda \geq 0$, an *f-assignment* is given by the function

$$t = \sum_{i=1}^{n} \underbrace{\frac{-f(x_i)}{\prod_{j\neq i}(x_j - x_i)}}_{:=p_i} \llbracket x_i \rrbracket = h - g,$$

(up to a positive multiplicative factor) where h contains the positive part of t and g the negative part (without any common support), viz. $h = \sum_{i:p_i>0} p_i \llbracket x_i \rrbracket$ and $g = \sum_{i:p_i<0} (-p_i) \llbracket x_i \rrbracket$.

- When f is a monomial, viz. has the form $f(x) = cx^q$, where $c > 0$ and $q \geq 0$ we call the assignment a *monomial assignment*. For $q = 0$ we call the assignment an *f_0-assignment*.

- We say that an assignment is *balanced* if the number of points with negative weights, $p_i < 0$, equals the number of points with positive weights, $p_i > 0$. We say an assignment is *unbalanced* if it is not balanced.

- We say that a monomial assignment is *aligned* if the degree of the monomial is an even number ($q = 2(b-1), b \in \mathbb{N}$). We say that a monomial assignment is *misaligned* if it is not aligned.

An f_0-assignment starts with a point that has a negative weight regardless of the total number of points and thereafter, the sign alternates. With this as the base structure, working out the signs of the weights for monomial assignments is facilitated. The only mathematical property which is needed to find an analytic solution, turns out to be the following:

[14] Briefly, we introduce the so-called Expressible-By-Real-Matrices (EBRM) transitions and functions, which are the analogue of EBM transitions and functions with the further restriction that the matrices and vectors involved are real; we then show that we can reduce the problem from EBM to EBRM transitions (see Appendix E).

Lemma 33. *Fix integers $m \leq n-2$ and $n \geq 2$. Consider an f-assignment of the form $t = \sum_i \frac{-(-x_i)^m}{\prod_{j \neq i}(x_j - x_i)} [\![x_i]\!]$ for n points $0 \leq x_1 < \cdots < x_n$ and use it to implicitly define p_{h_i} and p_{g_i} as follows: $t = \sum_i (x_{h_i})^m p_{h_i} [\![x_{h_i}]\!] - \sum_i (x_{g_i})^m p_{g_i} [\![x_{g_i}]\!]$. Let $\langle x^l \rangle := \sum_i (x_{h_i})^l p_{h_i} - \sum_i (x_{g_i})^l p_{g_i}$. Then, $\langle x^l \rangle = 0$ for $0 \leq l \leq n-2$. Further, $\langle x^{n-1} \rangle := \sum_i (x_{h_i})^{n-1} p_{h_i} - \sum_i (x_{g_i})^{n-1} p_{g_i} = (-1)^{m+n}$ which is strictly positive when $n+m$ is even (i.e. when t is unbalanced misaligned and balanced aligned (see Definition 32)).*

For the proof we refer to Appendix D.1.

Suppose that the f-assignment[15] can be decomposed into a sum of valid functions, and let us call these valid functions in the decomposition, *constituents*. In Section 1.1.2 we claimed that in order to implement the valid function corresponding to an f-assignment it suffices to implement the constituent functions. Let us briefly justify this claim. The difficulty is that the constituent functions might be negative at various locations, where there are no points present. A similar difficulty was encountered while transforming a TIPG into a TDPG, and it was handled using the technique of the catalyst state (following [Moc07; Aha+14b]), as we described in Section 2.4 after Theorem 27. This technique also applies here; for the f-assignment of the TIPG, we can again use a catalyst state, scale the constituent functions accordingly, and proceed thereafter as in the original proof [Aha+14b], to obtain the corresponding TDPG. The orthogonal matrices for the constituent functions are, thus, sufficient to get a TDPG with the same bias as for the $f-$assignment. This motivates Definition 34 below. We can then apply the TEF from Section 3 to the TDPG and obtain a WCF protocol approaching the same bias as the TIPG that we started with, in the limit of infinite rounds of communication. We emphasize that while we used the term *valid functions*, it was only for convenience and not a necessity.[16]

Definition 34 (Solving an assignment)**.** Given a finitely supported function $t = \sum_{i=1}^{n_h} p_{h_i} [\![x_{h_i}]\!] - \sum_{i=1}^{n_g} p_{g_i} [\![x_{g_i}]\!]$ and $\{|g_1\rangle, |g_2\rangle \dots |g_{n_g}\rangle, |h_1\rangle, |h_2\rangle \dots |h_{n_h}\rangle\}$ an orthonormal basis, we say that an orthogonal matrix O *solves* t if O satisfies the following: $O|v\rangle = |w\rangle$ and $X_h \geq E_h O X_g O^T E_h$ where $|v\rangle = \sum_{i=1}^{n_g} \sqrt{p_{g_i}} |g_i\rangle$, $|w\rangle = \sum_{i=1}^{n_h} \sqrt{p_{h_i}} |h_i\rangle$, $X_h = \sum_{i=1}^{n_h} x_{h_i} |h_i\rangle\langle h_i|$, $X_g = \sum_{i=1}^{n_g} x_{g_i} |g_i\rangle\langle g_i|$ and the projector $E_h = \sum_{i=1}^{n_h} |h_i\rangle\langle h_i|$. Moreover, we say that t has an *effective solution* if $t = \sum_{i \in I} t'_i$ and t'_i has a solution for all $i \in I$, where I is a finite set.

We first give a decomposition of an f-assignment into monomials which happens to be quite general. Another decomposition and its applications are given in Appendix D.2.

Lemma 35 (f-assignment as a sum of monomials)**.** *Consider a set of real coordinates satisfying $0 \leq x_1 < x_2 \cdots < x_n$ and let $f(x) = (r_1 - x)(r_2 - x)\dots(r_k - x)$ where $k \leq n-2$. Let $t = \sum_{i=1}^n p_i [\![x_i]\!]$ be the corresponding f-assignment. Then*

$$t = \sum_{l=0}^{k} \alpha_l \left(\sum_{i=1}^{n} \frac{-(-x_i)^l}{\prod_{j \neq i}(x_j - x_i)} [\![x_i]\!] \right),$$

where $\alpha_l \geq 0$.

[15] While an f-assignment is a valid function for all polynomials f satisfying the conditions in Definition 32, in what follows, we restrict to polynomials f with real roots. In fact, to be consistent with Definition 32, the roots must additionally be non-negative.

[16] We already argued along similar lines in Section 3 after Theorem 31 but for readability, we adapt the reasoning to the present discussion. Recall from Section 2.3 that (strictly) valid transitions were introduced as an alternative characterization of EBM transitions (see Proposition 21) to simplify the formalism—to show that a transition is EBM we need to verify constraints involving matrices, while to check its validity we need to verify constraints on scalars. In Appendix B we establish that the set of TEF functions, i.e. the set of functions which satisfy the constraints of Theorem 31, is equal to the set of valid functions (see Definition 126), which, in turn, due to conic duality, is equal to the closure of the set of EBM functions (see Definition 106). Here, we find the matrices themselves, therefore, we may work with TEF functions directly and circumvent the part of the formalism which uses conic duality (see Appendix A).

In the course of our analysis we have to use matrix inverses, therefore having a coordinate equal to zero breaks our argument. However, we can use the following lemma that tells us that the solution to the f-assignment is invariant under a shift of the origin.

Lemma 36. *Consider a set of real coordinates satisfying* $0 \leq x_1 < x_2 \cdots < x_n$ *and let* $f(x) = (a_1 - x)(a_2 - x)\dots(a_k - x)$ *where* $k \leq n-2$ *and the roots* $\{a_i\}_{i=1}^k$ *of* f *are non-negative. Let* $t = \sum_{i=1}^n p_i \llbracket x_i \rrbracket$ *be the corresponding* f*-assignment. Consider a set of real coordinates satisfying* $0 < x_1 + c < x_2 + c \cdots < x_n + c$ *where* $c > 0$ *and let* $f'(x) = (a_1 + c - x)(a_2 + c - x)\dots(a_k + c - x)$*. Let* $t' = \sum_{i=1}^n p'_i \llbracket x'_i \rrbracket$ *be the corresponding* f*-assignment with* $x'_i := x_i + c$*. The solution to* t *and to* t' *are the same.*

Proof sketch. We write $t = \sum_{i=1}^{n_h} p_{h_i} \llbracket x_{h_i} \rrbracket - \sum_{i=1}^{n_g} p_{g_i} \llbracket x_{g_i} \rrbracket$ and define $X_h := \sum_{i=1}^{n_h} x_{h_i} |h_i\rangle, X_g := \sum_{i=1}^{n_g} x_{g_i} |g_i\rangle$. If t is solved by O then we must have $X_h \geq E_h O X_g O^T E_h$. We then show that $X_h + c\mathbb{I}_h \geq E_h O(X_g + c\mathbb{I}_g)O^T E_h$, where $\mathbb{I}_h := \sum_{i=1}^{n_h} |h_i\rangle\langle h_i|$ and $\mathbb{I}_g := \sum_{i=1}^{n_g} |g_i\rangle\langle g_i|$. Together with the observation that $p'_i = p_i$ as the c's cancel, this establishes that O also solves t'. Since c is an arbitrary real number, it follows that O solves t if and only if it solves t'.

We now establish $X_h \geq E_h O X_g O^T E_h \iff X_h + c\mathbb{I}_h \geq E_h O(X_g + c\mathbb{I}_g)O^T E_h$. Observe that

$$\begin{aligned} X_h \geq E_h O X_g O^T E_h &\iff E_h(X_h - O X_g O^T)E_h \geq 0 && \because X_h = E_h X_h E_h \\ \iff E_h(X_h + c\mathbb{I}_{hg} - O(X_g - c\mathbb{I}_{hg})O^T)E_h \geq 0 &\iff X_h + c\mathbb{I}_h \geq E_h O(X_g + c\mathbb{I}_{hg})O^T E_h, \end{aligned}$$

where $\mathbb{I}_{hg} := \mathbb{I}$. Further,

$$X_g + c\mathbb{I}_{hg} \geq X_g + c\mathbb{I}_g \implies E_h O(X_g + c\mathbb{I}_{hg})O^T E_h \geq E_h O(X_g + c\mathbb{I}_g)O^T E_h$$

which together yield

$$X_h \geq E_h O X_g O^T E_h \iff X_h + c\mathbb{I}_h \geq E_h O(X_g + c\mathbb{I}_g)O^T E_h.$$

□

Having decomposed the f-assignment into a sum of monomial assignments, we now give the solution for each one of the different types of monomial assignments.

4.2 Solution to the f_0-assignment

We begin with the solution of the f_0-assignment. We first look at the balanced case, where the number of points involved, $2n$, is even. This corresponds to an $n \to n$ transition, i.e. a transition from n initial points to n final points.

4.2.1 The balanced case

Proposition 37 (Solution to balanced f_0-assignments). *Let*

- $t = \sum_{i=1}^n p_{h_i} \llbracket x_{h_i} \rrbracket - \sum_{i=1}^n p_{g_i} \llbracket x_{g_i} \rrbracket$ *be an* f_0*-assignment over* $\{x_1, x_2 \dots x_{2n}\}$
- $\{|h_1\rangle, |h_2\rangle \dots |h_n\rangle, |g_1\rangle, |g_2\rangle \dots |g_n\rangle\}$ *be an orthonormal basis, and*
- *finally*

$$X_h := \sum_{i=1}^n x_{h_i} |h_i\rangle\langle h_i| \doteq diag(x_{h_1}, \dots x_{h_n}, \underbrace{0, \dots 0}_{n\text{-}zeros}), X_g := \sum_{i=1}^n x_{g_i} |g_i\rangle\langle g_i| \doteq diag(\underbrace{0, \dots 0}_{n\text{-}zeros}, x_{g_1}, \dots x_{g_n}),$$

$$|w\rangle := \sum_{i=1}^n \sqrt{p_{h_i}} |h_i\rangle \doteq (\sqrt{p_{h_1}}, \dots \sqrt{p_{h_n}}, \underbrace{0, \dots 0}_{n\text{-}zeros})^T, |v\rangle := \sum_{i=1}^n \sqrt{p_{g_i}} |g_i\rangle \doteq (\underbrace{0, \dots 0}_{n\text{-}zeros}, \sqrt{p_{g_1}}, \dots \sqrt{p_{g_n}})^T.$$

Then,

$$O := \sum_{i=0}^{n-1}\left(\frac{\Pi^{\perp}_{h_{i-1}}(X_h)^i|w\rangle\langle v|(X_g)^i\Pi^{\perp}_{g_{i-1}}}{\sqrt{c_{h_i}c_{g_i}}} + h.c.\right)$$

satisfies $X_h \geq E_h O X_g O^T E_h$ *and* $O|v\rangle = |w\rangle$*, where* $E_h := \sum_{i=1}^n |h_i\rangle\langle h_i|$, $\Pi^{\perp}_{h_{-1}} = \Pi^{\perp}_{g_{-1}} = \mathbb{I}$,

$$\Pi^{\perp}_{h_i} := \textit{projector orthogonal to span}\{(X_h)^i|w\rangle, (X_h)^{i-1}|w\rangle, \dots|w\rangle\}, c_{h_i} := \langle w|(X_h)^i\Pi^{\perp}_{h_{i-1}}(X_h)^i|w\rangle,$$

and analogously

$$\Pi^{\perp}_{g_i} := \textit{projector orthogonal to span}\{(X_g)^i|v\rangle, (X_g)^{i-1}|v\rangle, \dots|v\rangle\}, c_{g_i} := \langle v|(X_g)^i\Pi^{\perp}_{g_{i-1}}(X_g)^i|v\rangle.$$

Proof. Using Lemma 33 for $2n$ points, we get

$$\left\langle x^k\right\rangle = 0 \quad \text{for} \quad k \in \{0, 1, 2\dots, 2n-2\}, \tag{7}$$

and

$$\left\langle x^{2n-1}\right\rangle > 0. \tag{8}$$

We define the basis of interest here, essentially using the Gram-Schmidt method. Let

$$\begin{aligned}|w_0\rangle &:= |w\rangle\\ |w_1\rangle &:= \frac{(\mathbb{I}-|w_0\rangle\langle w_0|)(X_h)|w\rangle}{\sqrt{c_{h_1}}}\\ &\vdots\\ |w_k\rangle &:= \frac{\left(\mathbb{I}-\sum_{i=0}^{k-1}|w_i\rangle\langle w_i|\right)(X_h)^k|w\rangle}{\sqrt{c_{h_k}}}.\end{aligned} \tag{9}$$

We indicate the term with the highest power of X_h appearing in $|w_k\rangle$ by

$$\mathcal{M}(|w_k\rangle) = \left\langle x_h^{2k}\right\rangle \cdot (X_h)^k|w\rangle$$

where the scalar factor represents the dependence on the highest power of x_h (appearing as $\langle x_h^l\rangle$) in $|w_k\rangle$. For instance, here the $\langle x_h^{2k}\rangle$ factor comes from $\sqrt{c_{h_k}}$. Note that the projectors can be expressed in terms of these vectors more concisely,

$$\Pi_{h_i} := \mathbb{I} - \Pi^{\perp}_{h_i} = \sum_{j=0}^{i}|w_j\rangle\langle w_j|.$$

It also follows that O can be re-written as $O = \sum_{j=0}^{n-1}\left(|w_j\rangle\langle v_j| + |v_j\rangle\langle w_j|\right)$, where $|v_j\rangle$ is analogously defined. It is evident that $O|v\rangle = |w\rangle$. Let $D = X_h - E_h O X_g O^T E_h$ and note that $\langle v_j|D|v_i\rangle = 0$ (because $X_h|v_i\rangle = 0$ and $E_h|v_i\rangle = 0$[17]). We assert that it has the following rank-1 form

$$D = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \langle w_{n-1}|D|w_{n-1}\rangle \end{bmatrix}$$

[17] The conclusion holds even without the projector as O maps span$(|v_1\rangle, |v_2\rangle, \dots|v_n\rangle)$ to span$(|w_1\rangle, |w_2\rangle\dots|w_n\rangle)$ on which X_g has no support.

in the $(|w_0\rangle, |w_1\rangle, \dots |w_{n-1}\rangle)$ basis, together with $\langle w_{n-1}| D |w_{n-1}\rangle > 0$. To see this, we simply compute

$$\langle w_i| D \left|w_j\right\rangle = \langle w_i| X_h \left|w_j\right\rangle - \langle w_i| O X_g O^T \left|w_j\right\rangle = \langle w_i| X_h \left|w_j\right\rangle - \langle v_i| X_g \left|v_j\right\rangle .$$

For (i, j) for any $0 \le i, j \le n-1$ except for the case where both $i = j = n-1$, the two terms are the same. This is because the term with the highest possible power l (of $\left\langle x^l \right\rangle$) in $\langle w_i| X_h \left|w_j\right\rangle$ can be deduced by observing

$$\mathcal{M}(\langle w_i|) X_h \mathcal{M}(\left|w_j\right\rangle) = \left\langle x_h^{2i} \right\rangle \cdot \left\langle x_h^{2j} \right\rangle \cdot \left\langle x_h^{i+j+1} \right\rangle . \tag{10}$$

For the analogous expression with gs to be the same, we must have $2i, 2j$ and $i + j + 1 \le 2n - 2$, using Equation (7). The first two conditions are always satisfied (for $0 \le i, j \le n-1$). The last can only be violated when $i = j = n - 1$. This establishes that the matrix has the asserted form. To prove the positivity of $\langle w_{n-1}| D |w_{n-1}\rangle$, consider $\langle w_{n-1}| X_h |w_{n-1}\rangle$ and $\langle v_{n-1}| X_g |v_{n-1}\rangle$. When these terms are expanded in powers of $\left\langle x_h^k \right\rangle$ and $\left\langle x_g^k \right\rangle$ respectively, only terms with $k > 2n - 2$ would remain; the others would get canceled due to Equation (7). From Equation (9) it follows that

$$\langle w_{n-1}| D |w_{n-1}\rangle = \frac{1}{c_{h_{n-1}}} \langle w| (X_h)^{2n-2+1} |w\rangle - \frac{1}{c_{g_{n-1}}} \langle v| (X_g)^{2n-2+1} |v\rangle$$

and it is not hard to see that $c_{h_{n-1}} = c_{h_{n-1}}(\left\langle x_h^{2n-2} \right\rangle, \left\langle x_h^{2n-3} \right\rangle, \dots, \left\langle x_h^{1} \right\rangle)$ does not depend on $\left\langle x_h^{2n-1} \right\rangle$ (and analogously for $c_{g_{n-1}}$). Also, $c_{h_{n-1}} = c_{g_{n-1}} =: c_{n-1}$. We thus have

$$\langle w_{n-1}| D |w_{n-1}\rangle = \frac{\left\langle x_h^{2n-1} \right\rangle}{c_{n-1}} > 0$$

using Equation (8). Hence, $X_h - E_h O X_g O^T E_h \ge 0$.
In the above, we assumed span$\{|w\rangle, X_h |w\rangle, X_h^2 |w\rangle, \dots, X_h^n |w\rangle\}$ equals span$\{|h_1\rangle, |h_2\rangle \dots |h_n\rangle\}$ which is justified by Lemma 134. □

4.2.2 The unbalanced case

We now consider unbalanced f_0-assignments, and we start by reviewing the result we just proved from a slightly different perspective to see where it fails in this case. We write $D_{ij} = \langle w_i| D \left|w_j\right\rangle$, and note that the maximum power, l, which appears as $\left\langle x^l_{g/h} \right\rangle$ is given by $\max\{2i, 2j, i + j + 1\}$. This yields a matrix with each term depending on the power as

$$D = \begin{bmatrix} D_{00}(\langle x \rangle) & & & \\ D_{10}(\left\langle x^2 \right\rangle, \dots) & D_{11}(\left\langle x^3 \right\rangle, \dots) & & \text{h.c.} \\ D_{20}(\left\langle x^4 \right\rangle, \dots) & D_{21}(\left\langle x^4 \right\rangle, \dots) & D_{22}(\left\langle x^5 \right\rangle, \dots) & \\ & & & \ddots \end{bmatrix} .$$

We represent this dependence as

$$\mathcal{M}(D) = \begin{bmatrix} \langle x \rangle & & & \\ \left\langle x^2 \right\rangle & \left\langle x^3 \right\rangle & & \\ \left\langle x^4 \right\rangle & \left\langle x^4 \right\rangle & \left\langle x^5 \right\rangle & \\ & & & \ddots \end{bmatrix} .$$

For concreteness, consider the balanced f_0-case over $\{x_1, x_2, x_3, x_4\}$, where $\langle x \rangle = \left\langle x^2 \right\rangle = 0$ and $\left\langle x^3 \right\rangle > 0$. For this two-dimensional case, we have

$$\mathcal{M}(D) = \begin{bmatrix} 0 & 0 \\ 0 & \left\langle x^3 \right\rangle \end{bmatrix} \geq 0.$$

Using the same method for an f_0-assignment over $\{x_1, x_2 \dots x_5\}$, we have $\langle x \rangle = \left\langle x^2 \right\rangle = \left\langle x^3 \right\rangle = 0$ and $\left\langle x^4 \right\rangle > 0$, and trying to solve in three dimensions, we would obtain

$$\mathcal{M}(D) = \begin{bmatrix} 0 & 0 & \left\langle x^4 \right\rangle \\ 0 & 0 & \left\langle x^4 \right\rangle \\ \left\langle x^4 \right\rangle & \left\langle x^4 \right\rangle & \left\langle x^5 \right\rangle \end{bmatrix} \tag{11}$$

which does not seem to work directly. It turns out that the projector appearing in the TEF constraint, removes the troublesome part and yields a zero matrix. This unbalanced assignment takes three points to two points. We define $X_h := \mathrm{diag}(x_{h_1}, x_{h_2}, 0, 0, 0)$, $|w\rangle = (\sqrt{p_{h_1}}, \sqrt{p_{h_2}}, 0, 0, 0)$ along with $|w_0\rangle := |w\rangle$ and $|w_1\rangle := (\mathbb{I} - |w_0\rangle \langle w_0|) X_h |w_0\rangle$. We can write $E_h = \sum_{i=0}^{1} |w_i\rangle \langle w_i|$ and have the same unitary as before, except that now $|v_2\rangle$ is left unchanged, i.e. $O = \sum_{i=0}^{1} |w_i\rangle \langle v_i| + |v_2\rangle \langle v_2|$. We can show that $D' = X_h - E_h O X_g O^T E_h \geq 0$ because every vector in $|\psi\rangle \in \mathrm{span}\{|v_0\rangle, |v_1\rangle, |v_2\rangle\}$ satisfies $D' |\psi\rangle = 0$ (as $X_h |\psi\rangle = 0$ and $E_h |\psi\rangle = 0$). This entails that it suffices to restrict to a 2×2 matrix in $\mathrm{span}\{|w_0\rangle, |w_1\rangle\}$. From Equation (11) this is zero, hence $D' = 0$. By generalizing this example, we can obtain the solution for an unbalanced f_0-assignment, as presented in the following Proposition:

Proposition 38 (Solution to unbalanced f_0-assignments)**.** *Let*

- $t = \sum_{i=1}^{n-1} p_{h_i} [\![x_{h_i}]\!] - \sum_{i=1}^{n} p_{g_i} [\![x_{g_i}]\!]$, *be an f_0-assignment over* $0 < x_1 < x_2 \dots < x_{2n-1}$
- $\{|h_1\rangle, |h_2\rangle \dots |h_{n-1}\rangle, |g_1\rangle, |g_2\rangle \dots |g_n\rangle\}$ *be an orthonormal basis, and*
- *finally*

$$X_h := \sum_{i=1}^{n-1} x_{h_i} |h_i\rangle \langle h_i| \doteq \mathit{diag}(x_{h_1}, \dots x_{h_{n-1}}, \underbrace{0, \dots 0}_{n\ zeros}), X_g := \sum_{i=1}^{n} x_{g_i} |g_i\rangle \langle g_i| \doteq \mathit{diag}(\underbrace{0, \dots 0}_{n-1\ zeros}, x_{g_1}, \dots, x_{g_n}),$$

$$|w\rangle := \sum_{i=1}^{n-1} \sqrt{p_{h_i}} |h_i\rangle \doteq (\sqrt{p_{h_1}}, \dots \sqrt{p_{h_{n-1}}}, \underbrace{0, \dots 0}_{n\ zeros})^T, |v\rangle := \sum_{i=1}^{n} \sqrt{p_{g_i}} |g_i\rangle \doteq (\underbrace{0, \dots 0}_{n-1\ zeros}, \sqrt{p_{g_1}}, \dots \sqrt{p_{g_n}})^T$$

- *and* $E_h := \sum_{i=1}^{n-1} |h_i\rangle \langle h_i|$.

Then,

$$O := \left(\sum_{i=0}^{n-2} \frac{\Pi^{\perp}_{h_{i-1}} (X_h)^i |w\rangle \langle v| (X_g)^i \Pi^{\perp}_{g_{i-1}}}{\sqrt{c_{h_i} c_{g_i}}} + h.c. \right) + \frac{\Pi^{\perp}_{g_{n-2}} (X_g)^{n-1} |v\rangle \langle v| (X_g)^{n-1} \Pi^{\perp}_{g_{n-2}}}{c_{g_i}}$$

satisfies $X_h \geq E_h O X_g O^T E_h$ *and* $E_h O |v\rangle = |w\rangle$, *where* $\Pi^{\perp}_{h_{-1}} = \Pi^{\perp}_{g_{-1}} = \mathbb{I}$,

$$\Pi^{\perp}_{h_i} := \textit{projector orthogonal to span}\{(X_h)^i |w\rangle, (X_h)^{i-1} |w\rangle, \dots |w\rangle\}, c_{h_i} := \langle w| (X_h)^i \Pi^{\perp}_{h_{i-1}} (X_h)^i |w\rangle,$$

and analogously

$$\Pi^{\perp}_{g_i} := \textit{projector orthogonal to span}\{(X_g)^i |v\rangle, (X_g)^{i-1} |v\rangle, \dots |v\rangle\}, c_{g_i} := \langle v| (X_g)^i \Pi^{\perp}_{g_{i-1}} (X_g)^i |v\rangle.$$

Proof. In this case, we use Lemma 33 for $2n-1$ points. We have

$$\left\langle x^k \right\rangle = 0 \tag{12}$$

but this time, $k \in \{0, 1, \ldots 2n-3\}$ and $\left\langle x^{2n-2} \right\rangle > 0$. We define the basis similarly by setting $|w_0\rangle := |w\rangle$ and for all $k \in \mathbb{Z}$ satisfying $0 \leq k \leq n-2$ we have

$$|w_k\rangle := \frac{\Pi^{\perp}_{h_{k-1}} (X_h)^k |w\rangle}{\sqrt{c_{h_k}}} = \frac{\left(\mathbb{I} - \sum_{i=0}^{k-1} |w_i\rangle \langle w_i|\right) (X_h)^k |w\rangle}{\sqrt{c_{h_k}}}.$$

We also define $|v_0\rangle := |v\rangle$ and for all $k \in \mathbb{Z}$ satisfying $0 \leq k \leq n-1$ we have

$$|v_k\rangle := \frac{\Pi^{\perp}_{g_{k-1}} (X_g)^k |v\rangle}{\sqrt{c_{g_k}}} = \frac{\left(\mathbb{I} - \sum_{i=0}^{k-1} |v_i\rangle \langle v_i|\right) (X_g)^k |v\rangle}{\sqrt{c_{h_k}}}.$$

This means that $O = \sum_{i=0}^{n-2} (|w_i\rangle \langle v_i| + |v_i\rangle \langle w_i|) + |v_n\rangle \langle v_n|$ and so $E_h O |v\rangle = |w\rangle$ follows directly. To establish $D := X_h - E_h O X_g O^T E_h \geq 0$, it suffices to show $\langle w_i| D \left|w_j\right\rangle \geq 0$ for $i, j \in \mathbb{Z}$ satisfying $0 \leq i, j \leq n-2$. Just as in the balanced case, this is because $D |v_i\rangle = 0$, as $X_h |v_i\rangle = 0$ and $E_h |v_i\rangle = 0$. As before, we denote the highest-power term of X_h appearing in $|w_k\rangle$, for k in $\{0, 1 \ldots n-2\}$, by

$$\mathcal{M}(|w_k\rangle) = \left\langle x_h^{2k} \right\rangle \cdot (X_h)^k |w\rangle$$

and analogously, the highest power of X_g appearing in $|v_k\rangle$ for k in $\{0, 1, \ldots n-2\}$, by

$$\mathcal{M}(|v_k\rangle) = \left\langle x_g^{2k} \right\rangle \cdot (X_g)^k |v\rangle .$$

Again, the highest power l of $\left\langle x^l \right\rangle$ in $\langle w_i| D \left|w_j\right\rangle$ is $\max\{2j, 2i, i+j+1\}$ which can be deduced by evaluating

$$\mathcal{M}(\langle w_i|) X_h \mathcal{M}(\left|w_j\right\rangle) = \left\langle x_h^{2j} \right\rangle \cdot \left\langle x_h^{2i} \right\rangle \cdot \left\langle x_h^{i+j+1} \right\rangle, \text{ and similarly}$$

$$\mathcal{M}(\langle v_i|) E_h O X_g O E_h \mathcal{M}(|v_i\rangle) = \left\langle x_g^{2j} \right\rangle \cdot \left\langle x_g^{2i} \right\rangle \cdot \left\langle x_g^{i+j+1} \right\rangle .$$

The highest possible power is attained for $i = j = n-2$. This yields $2n-3$ and thus, using Equation (12), we conclude that $\langle w_i| D \left|w_j\right\rangle = 0$ for all $0 \leq i, j \leq n-2$. □

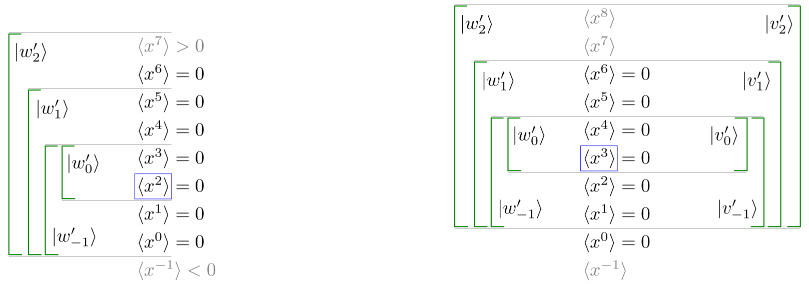
4.3 Solution to monomial assignments

There are four different types of monomial assignments depending on whether they are balanced or unbalanced and aligned or misaligned. One could find a single expression for all of them but it does not seem to aid clarity, therefore we present the solutions to all four different types separately. To go beyond the previous solutions to the f_0−assignments, the additional technique we introduce here is the use of the pseudo-inverses X_h^{-1} and X_g^{-1}, but the idea is essentially unchanged.

4.3.1 The balanced case

Even monomials (as opposed to odd monomials) seem to align well at the bottom; see Figure 10a, therefore justifying our choice to call them aligned (as opposed to misaligned).

Proposition 39 (Solution to balanced aligned monomial assignments)**.** *Let*



(a) $2n = 8$, $m = 2b = 2$. Balanced aligned monomial assignment

(b) $2n = 8$, $m = 2b - 1 = 3$. Balanced misaligned monomial assignment

Figure 10: Balanced monomial assignments

- *$m = 2b$ be an even non-negative integer*
- *$t = \sum_{i=1}^{n} x_{h_i}^m p_{h_i} [\![x_{h_i}]\!] - \sum_{i=1}^{n} x_{g_i}^m p_{g_i} [\![x_{g_i}]\!]$, be a monomial assignment over $0 < x_1 < x_2 \cdots < x_{2n}$*
- *$\{|h_1\rangle, |h_2\rangle \dots |h_n\rangle, |g_1\rangle, |g_2\rangle \dots |g_n\rangle\}$ be an orthonormal basis, and*
- *finally*

$$X_h := \sum_{i=1}^{n} x_{h_i} |h_i\rangle \langle h_i| \doteq diag(x_{h_1}, \dots x_{h_n}, \underbrace{0, \dots 0}_{n\ zeros}), X_g := \sum_{i=1}^{n} x_{g_i} |g_i\rangle \langle g_i| \doteq diag(\underbrace{0, \dots 0}_{n\ zeros}, x_{g_1}, \dots x_{g_n}),$$

$$|w\rangle := \sum_{i=1}^{n} \sqrt{p_{h_i}} |h_i\rangle \doteq (\sqrt{p_{h_1}}, \dots \sqrt{p_{h_n}}, \underbrace{0, \dots 0}_{n\ zeros})^T \text{ and } |w'\rangle := (X_h)^b |w\rangle,$$

$$|v\rangle := \sum_{i=1}^{n} \sqrt{p_{g_i}} |g_i\rangle \doteq (\underbrace{0, \dots 0}_{n\ zeros}, \sqrt{p_{g_1}}, \dots \sqrt{p_{g_n}})^T \text{ and } |v'\rangle := (X_g)^b |v\rangle.$$

Then,

$$O := \sum_{i=-b}^{n-b-1} \left(\frac{\Pi_{h_i}^{\perp} (X_h)^i |w'\rangle \langle v'| (X_g)^i \Pi_{g_i}^{\perp}}{\sqrt{c_{h_i} c_{g_i}}} + h.c. \right)$$

satisfies $X_h \geq E_h O X_g O^T E_h$ and $E_h O |v'\rangle = |w'\rangle$, where we write $(X_{h/g})^{-k}$ instead of $(X_{h/g}^{-1})^k$ (for $k > 0$), $E_h := \sum_{i=1}^{n} |h_i\rangle \langle h_i|, c_{h_i} := \langle w'| (X_h)^i \Pi_{h_i}^{\perp} (X_h)^i |w'\rangle$

$$\Pi_{h_i}^{\perp} := \begin{cases} \text{projector orthogonal to span}\{(X_h)^{-|i|+1} |w'\rangle, (X_h)^{-|i|+2} |w'\rangle \dots, |w'\rangle\} & i < 0 \\ \text{projector orthogonal to span}\{(X_h)^{-b} |w'\rangle, (X_h)^{-b+1} |w'\rangle, \dots (X_h)^{i-1} |w'\rangle\} & i > 0 \\ \mathbb{I} & i = 0, \end{cases}$$

and analogously $c_{g_i} := \langle v'| (X_g)^i \Pi_{g_i}^{\perp} (X_g)^i |v'\rangle$ and

$$\Pi_{g_i}^{\perp} := \begin{cases} \text{projector orthogonal to span}\{(X_g)^{-|i|+1} |v'\rangle, (X_g)^{-|i|+2} |v'\rangle \dots, |v'\rangle\} & i < 0 \\ \text{projector orthogonal to span}\{(X_g)^{-b} |v'\rangle, (X_g)^{-b+1} |v'\rangle, \dots (X_g)^{i-1} |v'\rangle\} & i > 0 \\ \mathbb{I} & i = 0. \end{cases}$$

Proof. The orthonormal basis of interest here is

$$\left|w'_i\right\rangle := \frac{\Pi^{\perp}_{h_i}(X_h)^i \left|w'\right\rangle}{\sqrt{c_{h_i}}}, \text{ which entails} \tag{13}$$

$$\Pi^{\perp}_{h_i} = \begin{cases} \mathbb{I}_h & i = 0 \\ \mathbb{I}_h - \sum_{j=i+1}^{0} \left|w'_j\right\rangle \left\langle w'_j\right| & i < 0 \\ \mathbb{I}_h - \sum_{j=-b}^{i-1} \left|w'_j\right\rangle \left\langle w'_j\right| & i > 0 \end{cases} \tag{14}$$

where $\mathbb{I}_h := E_h$. We define $\left|v'_i\right\rangle$ and $\Pi^{\perp}_{g_i}$ analogously. Here, we keep track of both the highest and lowest power, l in $\langle w'| X_h^l |w'\rangle$ and $\langle v'| X_g^l |v'\rangle$, which appear in the matrix elements $\left\langle w'_i\right| D \left|w'_j\right\rangle$. To this end, we use $\left\langle x_h^l\right\rangle' := \langle w'| X_h^l |w'\rangle = \langle w| X_h^{l+2b} |w\rangle$ and $\left\langle x_g^l\right\rangle' := \langle v'| X_g^l |v'\rangle = \langle v| X_g^{l+2b} |v\rangle$. We denote the minimum and maximum powers, l, by

$$\mathcal{M}(\left|w'_i\right\rangle) = \begin{cases} \left(\left\langle x_h^0\right\rangle' |w'\rangle, \left\langle x_h^0\right\rangle' |w'\rangle\right) & i = 0 \\ \left(\left\langle x_h^{-2|i|}\right\rangle' (X_h)^{-|i|} |w'\rangle, \left\langle x_h^0\right\rangle' |w'\rangle\right) & i < 0 \\ \left(\left\langle x_h^{-2b}\right\rangle' (X_h)^{-b} |w'\rangle, \left\langle x_h^{2i}\right\rangle' (X_h)^i |w'\rangle\right) & i > 0, \end{cases}$$

and we define $D := X_h - E_h O X_g O^T E_h \doteq \left\langle w'_i\right| (X_h - E_h O X_g O^T E_h) \left|w'_j\right\rangle$, as usual. It suffices to restrict to the span of the $\{\left|w'_i\right\rangle\}$ basis because $X_h \left|v'_i\right\rangle = 0$ and $E_h \left|v'_i\right\rangle = 0$. The lowest power, l, appearing in D is attained for $i = j = -b$ (as $-b \leq i, j \leq n - b - 1$). This can be evaluated to be $-2b$ by observing that

$$\mathcal{M}(\left\langle w'_{-b}\right|) X_h \mathcal{M}(\left|w'_{-b}\right\rangle) = \left(\left\langle x_h^{-2b}\right\rangle' \left\langle x_h^{-2b}\right\rangle' \left\langle x_h^{-2b+1}\right\rangle', \left\langle x_h^0\right\rangle' \left\langle x_h^0\right\rangle' \langle x_h\rangle'\right),$$

where we multiplied component-wise. To find the highest power, l, in the matrix D, note that for $i, j > 0$ we have

$$\mathcal{M}(\left\langle w'_i\right|) X_h \mathcal{M}(\left|w'_j\right\rangle) = \left(\left\langle x_h^{-2b}\right\rangle' \left\langle x_h^{-2b+1}\right\rangle' \left\langle x_h^{-2b}\right\rangle', \left\langle x_h^{2i}\right\rangle' \left\langle x_h^{2j}\right\rangle' \left\langle x_h^{i+j+1}\right\rangle'\right)$$

so $l = \max\{2i, 2j, i + j + 1\}$. As argued for the f_0-assignment, $l = 2n - 2b - 1$ for $i = j = n - b - 1$, otherwise $l < 2n - 2b - 1$. Thus, only the $D_{n-b-1,n-b-1}$ term in D, depends on $\left\langle x_h^{2n-2b-1}\right\rangle'$. All other terms, at most, depend on $\left\langle x_h^{-2b}\right\rangle', \left\langle x_h^{-2b+1}\right\rangle', \ldots \left\langle x_h^{2n-2b-2}\right\rangle'$, i.e. $\left\langle x_h^0\right\rangle, \left\langle x_h^1\right\rangle, \ldots \left\langle x_h^{2n-2}\right\rangle$. The analogous argument for $\left\langle v'_i\right| X_g \left|v'_j\right\rangle$, the observation that $\left\langle w'_i\right| D \left|w'_j\right\rangle = \left\langle w'_i\right| X_h \left|w'_j\right\rangle - \left\langle v'_i\right| X_g \left|v'_j\right\rangle$, and the fact that $\left\langle x^0\right\rangle = \left\langle x^1\right\rangle = \cdots = \left\langle x^{2n-2}\right\rangle = 0$ entail that these terms vanish. It remains to show that $D_{n-b-1,n-b-1} \geq 0$. Noting that in $\left\langle w'_{n-b-1}\right| D \left|w'_{n-b-1}\right\rangle$, the only term which would not get cancelled due to the aforesaid reasoning, must come from the part of $\left|w'_{n-b-1}\right\rangle$ containing $X_h^{n-b-1} |w'\rangle$. It suffices to show that the coefficient of this term is positive because we know that $\left\langle x^{2n-2b-1}\right\rangle' = \left\langle x^{2n-1}\right\rangle > 0$. We know this coefficient to be exactly $1/c_{h_{n-b-1}}$ (see Equation (14) and Equation (13)) establishing that $D \geq 0$. □

To proceed further, it is helpful to have a more concise way of viewing the proof. Let us consider a concrete example of a balanced aligned monomial assignment with $2n = 8$ and $m = 2b = 2$ (see Figure 10a). We represent the range of dependence of $\left\langle w'_0\right| X_h \left|w'_0\right\rangle$ on $\left\langle x_h^l\right\rangle$ diagrammatically by enclosing in a left bracket, the terms $\left\langle x^3\right\rangle = \langle x\rangle'$ and $\left\langle x^2\right\rangle = \left\langle x^0\right\rangle'$ (replacing $|w\rangle$ with $\left|w'_0\right\rangle$) and writing $\left|w'_0\right\rangle$ next to it. Similarly, for $\left|w'_{-1}\right\rangle, \left|w'_1\right\rangle$ and $\left|w'_2\right\rangle$ we enclose in a left bracket, the terms

$$\{\left\langle x^0\right\rangle, \left\langle x^1\right\rangle, \left\langle x^2\right\rangle, \left\langle x^3\right\rangle\} = \{\left\langle x^{-2}\right\rangle', \left\langle x^{-1}\right\rangle', \ldots \langle x\rangle'\},$$

$$\left\{\left\langle x^0\right\rangle, \left\langle x^1\right\rangle, \dots, \left\langle x^5\right\rangle\right\} = \left\{\left\langle x^{-2}\right\rangle', \left\langle x^{-1}\right\rangle', \dots \left\langle x^3\right\rangle'\right\},$$
$$\text{and } \left\{\left\langle x^0\right\rangle, \left\langle x^1\right\rangle, \dots \left\langle x^7\right\rangle\right\} = \left\{\left\langle x^{-2}\right\rangle', \left\langle x^{-1}\right\rangle', \dots \left\langle x^5\right\rangle'\right\},$$

respectively. The highest power l of $\left\langle x_h^l\right\rangle$ that appears in $\left\langle w_i'\right| X_h \left|w_j'\right\rangle$ is $l = 7$ when (and only when) $i = j = 2$. Thus, the matrix D, restricted to the subspace spanned by the $\{\left|w_i'\right\rangle\}$ basis (again, we can safely ignore the subspace span$\{\left|v_i'\right\rangle\}$ because $D\left|v_i'\right\rangle = 0$), has only one non-zero entry which we saw was positive as $\left\langle x^7\right\rangle > 0$.

A direct extension of this analysis to the balanced misaligned monomial assignment fails, as we can see concretely in the case with $2n = 8$ and $m = 2b - 1 = 3$ (see Figure 10b). From hindsight, we write both the $\left|v_i'\right\rangle$s and the $\left|w_i'\right\rangle$s. We start with $\left|w_0'\right\rangle = X_h^{3/2}\left|w\right\rangle$ and $\left|v_0'\right\rangle = X_g^{3/2}\left|v_0\right\rangle$, and as before, enclose the terms $\left\{\left\langle x^0\right\rangle' = \left\langle x^3\right\rangle, \left\langle x^1\right\rangle' = \left\langle x^4\right\rangle\right\}$ in a left bracket. We then multiply $\left|w_0'\right\rangle$ with X_h^{-1} (and $\left|v_0'\right\rangle$ with X_g^{-1} respectively) and project out the components along the previous vectors. We represent these by $\left|w_{-1}'\right\rangle$ and $\left|v_{-1}'\right\rangle$, and in the figure we enclose the terms $\left\{\left\langle x\right\rangle = \left\langle x^{-2}\right\rangle', \left\langle x^2\right\rangle = \left\langle x^{-1}\right\rangle' \dots \left\langle x^4\right\rangle = \left\langle x\right\rangle'\right\}$ in the left and right brackets. We do not go lower, because then we pickup a dependence on $\left\langle x^{-1}\right\rangle$ which persists for subsequent vectors. In general, we stop after taking b steps down (here $b = 1$). We go up by multiplying $\left|w_0'\right\rangle$ with X_h (and $\left|v_0'\right\rangle$ with X_g resp.) and projecting out the components along the previous vectors. We represent these by $\left|w_1'\right\rangle$ and $\left|v_1'\right\rangle$, and in the figure we enclose the terms $\left\{\left\langle x\right\rangle = \left\langle x^{-2}\right\rangle', \left\langle x^2\right\rangle = \left\langle x^{-1}\right\rangle' \dots \left\langle x^6\right\rangle = \left\langle x^3\right\rangle'\right\}$ in the brackets. Finally, we construct $\left|w_2'\right\rangle$ and $\left|v_2'\right\rangle$ by taking a step up using X_h and X_g, respectively. These are essentially fixed to be the vectors orthogonal to the previous ones, once we restrict to span$\{\left|h_1\right\rangle, \left|h_2,\right\rangle \dots \left|h_n\right\rangle\}$ and span$\{\left|g_1\right\rangle, \left|g_2,\right\rangle \dots \left|g_n\right\rangle\}$. Taking a step down using X_h^{-1} and X_g^{-1} we could have constructed $\left|w_{-2}'\right\rangle$ and $\left|v_{-2}'\right\rangle$, but these are the same as $\left|w_2'\right\rangle$ and $\left|v_2'\right\rangle$, as we have a 3-dimensional space. If we were to use $O = \sum_{i=-1}^{2}\left(\left|w_i'\right\rangle\left\langle v_i'\right| + \text{h.c.}\right)$ then we would have obtained dependence on $\left\langle x^7\right\rangle$ in the row corresponding to $\left|w_2'\right\rangle$ and a dependence on $\left\langle x^8\right\rangle$ for the term $\left\langle w_2'\right| D \left|w_2'\right\rangle$. This already hints that the matrix is negative because it has the form $\begin{bmatrix} 0 & b \\ b & c \end{bmatrix}$ with $b \neq 0$; thus this choice cannot work. We therefore define $O := \left(\sum_{i=-1}^{1}\left|w_i'\right\rangle\left\langle v_i'\right| + \text{h.c.}\right) + \left|w_2'\right\rangle\left\langle w_2'\right| + \left|v_2'\right\rangle\left\langle v_2'\right|$. Further, instead of using

$$X_h \geq E_h O X_g O^T E_h \tag{15}$$

for establishing positivity, we equivalently use

$$E_h \geq \left(X_h^{\dashv}\right)^{1/2} O X_g O^T \left(X_h^{\dashv}\right)^{1/2}, \tag{16}$$

which is easily obtained by multiplying by $(X_h^{\dashv})^{1/2}$ on both sides. The reason is that to establish positivity, we must include $\left|w_2'\right\rangle$ in the basis (we can neglect the null vectors of E_h), and even though the RHS of Equation (15) would not contribute, the LHS would get non-trivial contributions along the rows. Using the inverses allows us to remove this dependence. To see this, note that span$\{\left|w_{-1}'\right\rangle, \left|w_0'\right\rangle \dots \left|w_2'\right\rangle\}$ equals the h-space, i.e. span$\{\left|h_1\right\rangle, \left|h_2\right\rangle \dots \left|h_n\right\rangle\}$. Further, span$\{X_h^{1/2}\left|w_i'\right\rangle\}_{i=-1}^{2}$ also equals the h-space (but the vectors are not, in general, orthonormal any more). Finally, observe that $X_h^{1/2}\left|w_2'\right\rangle$ is a null vector of the RHS of Equation (16). Therefore, to prove the positivity it suffices to restrict to span$\{X_h^{1/2}\left|w_i'\right\rangle\}_{i=-1}^{1}$. An arbitrary normalized vector in this space can be written as

$$\left|\psi\right\rangle = \frac{\sum_{i=-1}^{1}\alpha_i X_h^{1/2}\left|w_i'\right\rangle}{\sqrt{\sum_{i,j=-1}^{1}\alpha_i\alpha_j\left\langle w_i'\right| X_h \left|w_j'\right\rangle}} \implies X_g^{1/2} O^T (X_h^{\dashv})^{1/2}\left|\psi\right\rangle = \frac{\sum_{i=-1}^{1}\alpha_i X_g^{1/2}\left|v_i'\right\rangle}{\sqrt{\sum_{i,j=-1}^{1}\alpha_i\alpha_j\left\langle w_i'\right| X_h \left|w_j'\right\rangle}}$$

$$\implies \left\langle\psi\right| (X_h^{\dashv})^{1/2} O X_g O^T (X_h^{\dashv})^{1/2}\left|\psi\right\rangle = \frac{\sum_{i,j=-1}^{1}\alpha_i\alpha_j\left\langle v_i'\right| X_g \left|v_j'\right\rangle}{\sum_{i,j=-1}^{1}\alpha_i\alpha_j\left\langle w_i'\right| X_h \left|w_j'\right\rangle} = 1,$$

where we get equality by noting that $\left\langle v'_i \middle| X_g \middle| v'_j \right\rangle$s depend on (at most) $\left\{\left\langle x_g \right\rangle, \left\langle x_g^2 \right\rangle \dots \left\langle x_g^6 \right\rangle\right\}$ and analogously $\left\langle w'_i \middle| X_h \middle| w'_j \right\rangle$ depend on (at most) $\left\{\left\langle x_h \right\rangle, \left\langle x_h^2 \right\rangle \dots \left\langle x_h^6 \right\rangle\right\}$, which are the same as $\left\langle x^i \right\rangle = 0$ for $i \in \{0, 1, \dots 6\}$. Since we proved the RHS of Equation (16) equals 1 for all normalized $|\psi\rangle$s, we conclude that we have the correct unitary.

Proposition 40 (Solution to balanced misaligned monomial assignments)**.** *Let*

- $m = 2b - 1$ *be an odd non-negative integer (i.e.* $b \geq 1$*)*
- $t = \sum_{i=1}^{n} x_{h_i}^m p_{h_i} [\![x_{h_i}]\!] - \sum_{i=1}^{n} x_{g_i}^m p_{g_i} [\![x_{g_i}]\!]$, *be a monomial assignment over* $\{x_1, x_2 \dots x_{2n}\}$
- $(|h_1\rangle, |h_2\rangle \dots |h_n\rangle, |g_1\rangle, |g_2\rangle \dots |g_n\rangle)$ *be an orthonormal basis*
- *finally*

$$X_h := \sum_{i=1}^{n} x_{h_i} |h_i\rangle \langle h_i| \doteq diag(x_{h_1}, \dots x_{h_n}, \underbrace{0, \dots 0}_{n\ zeros}), X_g := \sum_{i=1}^{n} x_{g_i} |g_i\rangle \langle g_i| \doteq diag(\underbrace{0, \dots 0}_{n\ zeros}, x_{g_1}, \dots x_{g_n}),$$

$$|w\rangle := (\sqrt{p_{h_1}}, \dots \sqrt{p_{h_n}}, \underbrace{0, \dots 0}_{n\ zeros})\ and\ |w'\rangle := (X_h)^{b-\frac{1}{2}} |w\rangle$$

$$|v\rangle := (\underbrace{0, \dots 0}_{n\ zeros}, \sqrt{p_{g_1}}, \dots \sqrt{p_{g_n}})\ and\ |v'\rangle := (X_g)^{b-\frac{1}{2}} |v\rangle .$$

Then,

$$O := \sum_{i=-b+1}^{n-b-1} \left(\frac{\Pi_{h_i}^{\perp} (X_h)^i |w'\rangle \langle v'| (X_g)^i \Pi_{g_i}^{\perp}}{\sqrt{c_{h_i} c_{g_i}}} + h.c. \right) \quad + \frac{\Pi_{g_{n-b}}^{\perp} (X_g)^{n-b} |v'\rangle \langle v'| (X_g)^{n-b} \Pi_{g_{n-b}}^{\perp}}{c_{g_{n-b+1}}}$$
$$+ \frac{\Pi_{h_{n-b}}^{\perp} (X_h)^{n-b} |w'\rangle \langle w'| (X_h)^{n-b} \Pi_{h_{n-b}}^{\perp}}{c_{h_{n-b}}}$$

satisfies $X_h \geq E_h O X_g O^T E_h$ *and* $E_h O |v'\rangle = |w'\rangle$*, where we write* $X_{h/g}^{-k}$ *instead of* $(X_{h/g}^{\dashv})^k$ *for* $k > 0$, $c_{h_i} := \langle w'| (X_h)^i \Pi_{h_i}^{\perp} (X_h)^i |w'\rangle$,

$$\Pi_{h_i}^{\perp} := \begin{cases} projector\ orthogonal\ to\ span\{(X_h^{\dashv})^{|i|-1} |w'\rangle, (X_h^{\dashv})^{|i|-2} |w'\rangle \dots, |w'\rangle\} & i < 0 \\ projector\ orthogonal\ to\ span\{(X_h^{\dashv})^{b-1} |w'\rangle, (X_h^{\dashv})^{b-2} |w'\rangle, \dots, |w'\rangle, X_h |w'\rangle, \dots (X_h)^{i-1} |w'\rangle\} & i > 0 \\ \mathbb{I} & i = 0, \end{cases}$$

and analogously $c_{g_i} := \langle v'| (X_g)^i \Pi_{g_i}^{\perp} (X_g)^i |v'\rangle$,

$$\Pi_{g_i}^{\perp} := \begin{cases} projector\ orthogonal\ to\ span\{(X_g^{\dashv})^{|i|-1} |v'\rangle, (X_g^{\dashv})^{|i|-2} |v'\rangle \dots, |v'\rangle\} & i < 0 \\ projector\ orthogonal\ to\ span\{(X_g^{\dashv})^{b-1} |v'\rangle, (X_g^{\dashv})^{b-2} |v'\rangle, \dots |v'\rangle, X_g |v'\rangle, \dots (X_g)^{i-1} |v'\rangle\} & i > 0 \\ \mathbb{I} & i = 0. \end{cases}$$

Proof. The proof is very similar to that of Proposition 39. The orthonormal basis of interest here is

$$\left| w'_i \right\rangle := \frac{\Pi_{h_i}^{\perp} (X_h)^i |w'\rangle}{\sqrt{c_{h_i}}}$$

which entails

$$\Pi_{h_i}^{\perp} = \begin{cases} \mathbb{I}_h & i = 0 \\ \mathbb{I}_h - \sum_{j=i-1}^{0} \left| w'_j \right\rangle \left\langle w'_j \right| & i < 0 \\ \mathbb{I}_h - \sum_{j=-b+1}^{i} \left| w'_j \right\rangle \left\langle w'_j \right| & i > 0 \end{cases}$$

where $\mathbb{I}_h := E_h$. We define $\left|v_i'\right\rangle$ and $\Pi_{g_i}^{\perp}$ analogously. Our strategy is to keep track of the highest and lowest powers, l, in $\langle w'|\, X_h^l \,|w'\rangle$ and $\langle v'|\, X_g^l \,|v'\rangle$, which appear in the matrix elements $\left\langle w_i'\right| X_h \left|w_j'\right\rangle$ and $\left\langle v_i'\right| X_g \left|v_j'\right\rangle$. For brevity we write $\left\langle x_h^l\right\rangle' := \langle w'|\, X_h^l \,|w'\rangle$ and $\left\langle x_g^l\right\rangle' := \langle v'|\, X_g^l \,|v'\rangle$. The minimum and maximum powers, l, are denoted by

$$\mathcal{M}(\left|w_i'\right\rangle) = \begin{cases} \left(\left\langle x_h^0\right\rangle' |w'\rangle, \left\langle x_h^0\right\rangle' |w'\rangle\right) & i = 0 \\ \left(\left\langle x_h^{-2|i|}\right\rangle' (X_h)^{-|i|} |w'\rangle, \left\langle x_h^0\right\rangle' |w'\rangle\right) & i < 0 \\ \left(\left\langle x_h^{-2(b-1)}\right\rangle' (X_h)^{-(b-1)} |w'\rangle, \left\langle x_h^{2i}\right\rangle' (X_h)^{i} |w'\rangle\right) & i > 0. \end{cases}$$

Establishing $X_h \geq E_h O X_g O^T E_h$ is equivalent to establishing

$$E_h \geq X_h^{-1/2} O X_g O^T X_h^{-1/2}. \tag{17}$$

It is easy to see that $X_h^{1/2}\left|w_{n-b}'\right\rangle$ is a vector with zero eigenvalue for the RHS as $X_g O^T \left|w_{n-b}'\right\rangle = 0$. Any vector $|\psi\rangle \in \text{span}\{|g_1\rangle, |g_2\rangle \ldots |g_n\rangle\}$ is a vector with zero eigenvalue for both the LHS and the RHS. Thus, for the positivity we can restrict to $\text{span}\{|h_1\rangle, |h_2\rangle, \ldots |h_n\rangle\}\backslash\text{span}\{X_h^{1/2}\left|w_{n-b}'\right\rangle\}$, i.e. to vectors in the h-space orthogonal to $X_h^{1/2}\left|w_{n-b}'\right\rangle$. It turns out to be easier to test for positivity on a larger space. It is clear that $\text{span}\left\{X_h^{1/2}\left|w_i'\right\rangle\right\}_{i=-b+1}^{n-b} = \text{span}\{|h_1\rangle, |h_2\rangle \ldots |h_n\rangle\} = \text{span}\{\left|w_i'\right\rangle\}_{i=-b+1}^{n-b}$, (due to Lemma 134). As neglecting vectors with components along $X_h^{1/2}\left|w_{n-b}'\right\rangle$ suffices to satisfy Equation (17), we can restrict to $\text{span}\{X_h^{1/2}\left|w_i'\right\rangle\}_{i=-b+1}^{n-b-1}$ (which might still contain vectors with components along $X_h^{1/2}\left|w_{n-b}'\right\rangle$ as the basis vectors are not orthogonal but it only means that we check for positivity over a larger set of vectors). These ensure that the troublesome vectors $\left|w_{n-b}'\right\rangle$ and $\left|v_{n-b}'\right\rangle$ do not appear in the remaining analysis. Let $|\psi\rangle = \left(\sum_{i=-b+1}^{n-b-1} \alpha_i X_h^{1/2}\left|w_i'\right\rangle\right)/c$ where $c = \sqrt{\langle\psi|\psi\rangle}$. To establish Equation (17), it is enough to show that for all choices of α_is,

$$1 \geq \langle\psi|\, X_h^{-1/2} O X_g O^T X_h^{-1/2} \,|\psi\rangle = \frac{\sum_{i,j=-b+1}^{n-b-1} \alpha_i\alpha_j \left\langle v_i'\right| X_g \left|v_j'\right\rangle}{\sum_{i,j=-b+1}^{n-b-1} \alpha_i\alpha_j \left\langle w_i'\right| X_h \left|w_j'\right\rangle} = 1 \tag{18}$$

where the second step follows from $X_g^{1/2} O^T X_h^{-1/2} |\psi\rangle = \sum_{i=-b+1}^{n-b-1} \alpha_i X_g^{1/2} \left|v_i'\right\rangle$ and the last step follows from the counting argument below. Start by noting that

$$\left\langle x_h^i\right\rangle' = \left\langle x_h^{i+2b-1}\right\rangle \text{ and } \left\langle x^0\right\rangle = \langle x\rangle = \cdots = \left\langle x^{2n-2}\right\rangle = 0. \tag{19}$$

To determine the highest power of l in $\langle w'|\, X_h^l \,|w'\rangle$ which appears in the matrix elements $\left\langle w_i'\right| X_h \left|w_j'\right\rangle$ (for $-b+1 \leq i, j \leq n-b-1$) it suffices to consider the expectation values $\left\langle w_{n-b-1}'\right| X_h \left|w_{n-b-1}'\right\rangle$. To this end, we evaluate

$$\begin{aligned} &\mathcal{M}(\left\langle w_{n-b-1}'\right|) X_h \mathcal{M}(\left|w_{n-b-1}'\right\rangle) \\ &= \left(\left\langle x_h^{-2(b-1)}\right\rangle' \left\langle x_h^{-2(b-1)}\right\rangle' \left\langle x_h^{-2(b-1)+1}\right\rangle', \left\langle x_h^{2(n-b-1)}\right\rangle' \left\langle x_h^{2(n-b-1)}\right\rangle' \left\langle x_h^{2(n-b-1)+1}\right\rangle'\right) \\ &= \left(\langle x_h\rangle \langle x_h\rangle \left\langle x_h^2\right\rangle, \left\langle x_h^{2n-3}\right\rangle \left\langle x_h^{2n-3}\right\rangle \left\langle x_h^{2n-2}\right\rangle\right). \end{aligned}$$

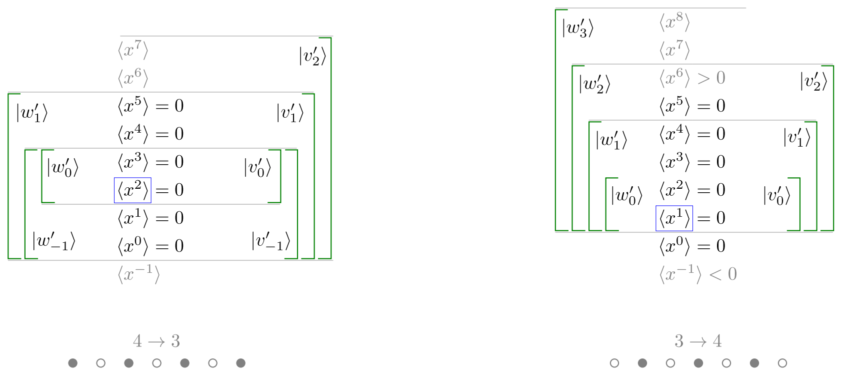
The highest power is, manifestly, $l = 2n-2$. To find the lowest power l in $\langle w'|\, X_h^l \,|w'\rangle$ appearing in $\left\langle w_i'\right| X_h \left|w_j'\right\rangle$ (for $-b+1 \leq i, j \leq n-b-1$) it suffices to consider $\left\langle w_{-b+1}'\right| X_h \left|w_{-b+1}'\right\rangle$. To this end, we evaluate

$$\begin{aligned}\mathcal{M}(\langle w'_{-b+1}|)X_h\mathcal{M}(|w'_{-b+1}\rangle) &= \left(\left\langle x_h^{-2(b-1)}\right\rangle' \left\langle x_h^{-2(b-1)}\right\rangle' \left\langle x_h^{-2(b-1)+1}\right\rangle', \langle x_h^0\rangle' \langle x_h^0\rangle' \langle x_h\rangle'\right) \\ &= \left(\langle x_h\rangle \langle x_h\rangle \langle x_h^2\rangle, \left\langle x_h^{2b-1}\right\rangle \left\langle x_h^{2b-1}\right\rangle \left\langle x_h^{2b}\right\rangle\right).\end{aligned}$$

The lowest power is, manifestly, $l = 1$. We thus conclude that the numerator of Equation (18) is a function of $\langle x_h\rangle, \langle x_h^2\rangle, \dots \langle x_h^{2n-2}\rangle$ and, an analogous argument entails that the denominator is a function of $\langle x_g\rangle, \left\langle x_g^2\right\rangle, \dots \left\langle x_g^{2n-2}\right\rangle$ with the same form. Using Equation (19), we conclude that the numerator and the denominator are the same. □

4.3.2 The unbalanced case

The techniques we have used so far also work when the number of points in a monomial assignment are odd, i.e. for unbalanced monomial assignments, both aligned and misaligned. We illustrate how the solution is constructed by considering a concrete example of an unbalanced aligned monomial assignment. We start with $2n - 1 = 7$ points and $m = 2b = 2$ (see Figure 11a). We use the diagrammatic representation introduced previously. In this case, we have 4 initial and 3 final points; the standard basis is $\{|g_1\rangle, |g_2\rangle, \dots |g_4\rangle, |h_1\rangle, |h_2\rangle, |h_3\rangle\}$. The basis of interest is again constructed by starting at $|w'\rangle$ and using X_h^{-1} until we reach $\langle x^0\rangle$, and then by using X_h until the space is spanned (analogously for $|v'\rangle$ with X_g^{-1} and X_g). It is $\{|v'_{-1}\rangle, |v'_0\rangle, |v'_1\rangle, |v'_2\rangle\}$ and $\{|w'_{-1}\rangle, |w'_0\rangle, |w'_1\rangle\}$. In the same vein as the earlier solutions, we define $O := \sum_{i=-1}^{1} (|w'_i\rangle \langle v'_i| + \text{h.c.}) + |v'_2\rangle \langle v'_2|$. In $X_h \geq E_h O X_g O^T E_h$, the $|v'_2\rangle$ term is removed by the projector, $E_h := \sum_{i=1}^{3} |h_i\rangle \langle h_i|$. Using $\langle x^0\rangle = \langle x\rangle = \dots = \langle x^5\rangle = 0$ and the counting arguments from before, it follows that $D = X_h - E_h O X_g O^T E_h = 0$.



(a) $2n - 1 = 7$; $m = 2b = 2$. Unbalanced aligned monomial assignment.

(b) $2n - 1 = 7$; $m = 2b - 1 = 1$. Unbalanced misaligned monomial assignment.

Figure 11: Visualizing unbalanced monomial assignment with simple examples.

For an unbalanced misaligned monomial assignment let us consider the example with $2n - 1 = 7$ and $m = 2b - 1 = 1$. We have 3 initial and 4 final points; the standard basis is $\{|g_1\rangle, |g_2\rangle, |g_3\rangle, |h_1\rangle, |h_2\rangle, \dots |h_4\rangle\}$. We construct the basis of interest by starting at $|w'\rangle$ and using X_h until the space is spanned (analogously for $|v'\rangle$ with X_g). More generally, we first go down for $b - 2$ steps (which is zero in this case), until $\langle x\rangle$ is reached in the diagram. The bases are $\{|v'_0\rangle, |v'_1\rangle, |v'_2\rangle\}$ and $\{|w'_0\rangle, |w'_1\rangle, |w'_2\rangle, |w'_3\rangle\}$. As before, we define $O := \sum_{i=0}^{2} (|w'_i\rangle \langle v'_i| + \text{h.c.}) + |w'_3\rangle \langle w'_3|$. This time we use $E_h \geq X_h^{-1/2} O X_g O^T X_h^{-1/2}$ which is equivalent to $X_h \geq E_h O X_g O^T E_h$ for $E_h := \sum_{i=1}^{4} |h_i\rangle \langle h_i|$. Using an argument similar to the balanced misaligned case, we can reduce the positivity condition to

$$1 \geq \frac{\sum_{i,j=0}^{2} \alpha_i \alpha_j \left\langle v_i' \middle| X_g \middle| v_j' \right\rangle}{\sum_{i,j=0}^{2} \alpha_i \alpha_j \left\langle w_i' \middle| X_h \middle| w_j' \right\rangle}$$

but the counting argument doesn't make the fraction 1. This is because we now have an $\left\langle x_h^6 \right\rangle$ dependence in the denominator and an $\left\langle x_g^6 \right\rangle$ dependence in the numerator. However, we also know that this term only appears in $\left\langle w_2' \middle| X_h \middle| w_2' \right\rangle$ that too with a positive coefficient (as we saw in the unbalanced f_0-assignment). Further, we know $\left\langle x_h^6 \right\rangle > \left\langle x_g^6 \right\rangle$ and therefore we can conclude that the numerator is smaller than the denominator ensuring the inequality is always satisfied. We state the general solution for both these cases and prove their correctness below.

Proposition 41 (Solution to unbalanced aligned monomial assignments)**.** *Let*

- *$m = 2b$ be an even non-negative integer*
- *$t = \sum_{i=1}^{n-1} x_{h_i}^m p_{h_i} [\![x_{h_i}]\!] - \sum_{i=1}^{n} x_{g_i}^m p_{g_i} [\![x_{g_i}]\!]$, be a monomial assignment over $\{x_1, x_2 \dots x_{2n-1}\}$*
- *$(|h_1\rangle, |h_2\rangle \dots |h_{n-1}\rangle, |g_1\rangle, |g_2\rangle \dots |g_n\rangle)$ be an orthonormal basis*
- *finally*

$$X_h := \sum_{i=1}^{n-1} x_{h_i} |h_i\rangle \langle h_i| \doteq diag(x_{h_1}, \dots x_{h_{n-1}}, \underbrace{0, \dots 0}_{n \text{ zeros}}), X_g := \sum_{i=1}^{n} x_{g_i} |g_i\rangle \langle g_i| \doteq diag(\underbrace{0, \dots 0}_{n-1 \text{ zeros}}, x_{g_1}, \dots x_{g_n}),$$

$$|w\rangle := (\sqrt{p_{h_1}}, \dots \sqrt{p_{h_{n-1}}}, \underbrace{0 \dots 0}_{n \text{ zeros}}) \text{ and } |w'\rangle := (X_h)^b |w\rangle,$$

$$|v\rangle := (\underbrace{0, 0, \dots 0}_{n-1 \text{ zeros}}, \sqrt{p_{g_1}}, \sqrt{p_{g_2}} \dots \sqrt{p_{g_n}}) \text{ and } |v'\rangle := (X_g)^b |v\rangle.$$

Then

$$O := \sum_{i=-b}^{n-b-2} \left(\frac{\Pi_{h_i}^{\perp} (X_h)^i |w'\rangle \langle v'| (X_g)^i \Pi_{g_i}^{\perp}}{\sqrt{c_{h_i} c_{g_i}}} + h.c. \right) + \frac{\Pi_{g_{n-b-1}}^{\perp} (X_g)^{n-b-1} |v'\rangle \langle v'| (X_g)^{n-b-1} \Pi_{g_{n-b-1}}^{\perp}}{c_{g_{n-b-1}}}$$

satisfies $X_h \geq E_h O X_g O^T E_h$ and $E_h O |v'\rangle = |w'\rangle$, where by $X_{h/g}^{-k}$ we mean $(X_{h/g}^{\dashv})^k$ for $k > 0$, and all $c_{h_i}, c_{g_i}, \Pi_{h_i}^{\perp}, \Pi_{g_i}^{\perp}$ are as defined in Proposition 39.

Proof. Many observations from the proof of Proposition 39 carry over to this case. We import the definitions of $\{|w_i'\rangle\}_{i=-b}^{n-b-2}$ and $\{|v_i'\rangle\}_{i=-b}^{n-b-1}$, together with the observations that $\mathcal{M}(\langle w_{-b}'|) X_h \mathcal{M}(|w_{-b}'\rangle)$ has no dependence on a term $\left\langle x_h^l \right\rangle'$ with $l < -2b$ and that $\mathcal{M}(\langle w_{n-b-2}'|) X_h \mathcal{M}(|w_{n-b-2}'\rangle)$ has no dependence on a term $\left\langle x_h^l \right\rangle'$ with $l > 2n-2b-4+1 = 2n-3-2b$. We can restrict to $\text{span}\{|w_{-b}'\rangle, |w_{-b+1}'\rangle \dots |w_{n-b-2}'\rangle\}$ to establish the positivity of $D := X_h - E_h O X_g O^T E_h$. Using the analogous observation for $\mathcal{M}(\langle v_{-b}'|) X_g \mathcal{M}(|v_{-b}'\rangle)$ and $\mathcal{M}(\langle v_{n-b-2}'|) X_g \mathcal{M}(|v_{n-b-2}'\rangle)$, along with the fact that $\left\langle x^l \right\rangle' = \left\langle x^{l+2b} \right\rangle$ and $\left\langle x^0 \right\rangle = \left\langle x^1 \right\rangle = \dots = \left\langle x^{2n-3} \right\rangle = 0$, it follows that $D = 0$. □

Proposition 42 (Solution to unbalanced misaligned monomial assignments)**.** *Let*

- *$m = 2b - 1$ be an odd non-negative integer*
- *$t = \sum_{i=1}^{n} x_{h_i}^{m} p_{h_i} [\![x_{h_i}]\!] - \sum_{i=1}^{n-1} x_{g_i}^{m} p_{g_i} [\![x_{g_i}]\!]$ be a monomial assignment over $\{x_1, x_2 \ldots x_{2n-1}\}$*
- *$(|h_1\rangle, |h_2\rangle \ldots |h_n\rangle, |g_1\rangle, |g_2\rangle \ldots |g_{n-1}\rangle)$ be an orthonormal basis*
- *finally*

$$X_h := \sum_{i=1}^{n} x_{h_i} |h_i\rangle \langle h_i| \doteq diag(x_{h_1}, \ldots x_{h_n}, \underbrace{0, \ldots 0}_{n-1\ zeros})X_g := \sum_{i=1}^{n-1} x_{g_i} |g_i\rangle \langle g_i| \doteq diag(\underbrace{0, \ldots 0,}_{n\ zeros} x_{g_1}, \ldots x_{g_{n-1}}),$$

$$|w\rangle := (\sqrt{p_{h_1}}, \ldots \sqrt{p_{h_n}}, \underbrace{0, \ldots 0}_{n-1\ zeros}) \ and\ |w'\rangle := (X_h)^{b-\frac{1}{2}} |w\rangle,$$

$$|v\rangle := (\underbrace{0, \ldots 0,}_{n\ zeros} \sqrt{p_{g_1}}, \ldots \sqrt{p_{g_{n-1}}}) \ and\ |v'\rangle := (X_g)^{b-\frac{1}{2}} |v\rangle .$$

Then

$$O := \sum_{i=-b+1}^{n-b-1} \left(\frac{\Pi_{h_i}^{\perp} (X_h)^i |w'\rangle \langle v'| (X_g)^i \Pi_{g_i}^{\perp}}{\sqrt{c_{h_i} c_{g_i}}} + h.c. \right) + \frac{\Pi_{h_{n-b}}^{\perp} (X_h)^{n-b} |w'\rangle \langle w'| (X_h)^{n-b} \Pi_{h_{n-b}}^{\perp}}{c_{h_{n-b}}},$$

satisfies $X_h \geq E_h O X_g O^T E_h$ and $E_h O |v'\rangle = |w'\rangle$, where by $X_{h/g}^{-k}$ we mean $(X_{h/g}^{\dashv})^k$ for $k > 0$, and all $c_{h_i}, c_{g_i}, \Pi_{h_i}^{\perp}, \Pi_{g_i}^{\perp}$ are as defined in Proposition 40.

Proof. For this proof, we can use the definitions and observations from the proof of Proposition 40. We import the definitions of $\{|w_i'\rangle\}_{i=-b+1}^{n-b}$ and $\{|v_i'\rangle\}_{i=-b+1}^{n-b-1}$ along with the observation that

$$\mathcal{M}(\langle w_{-b+1}'|) X_h \mathcal{M}(|w_{-b+1}'\rangle)$$

has no dependence on a term $\langle x_h^l \rangle'$ with $l < -2b + 2$ and

$$\mathcal{M}(\langle w_{n-b-1}'|) X_h \mathcal{M}(|w_{n-b-1}'\rangle)$$

has no dependence on a term $\langle x^l \rangle$ with $l > 2n-2b-1$.Also from the previous proof we have that establishing $X_h \geq E_h O X_g O^T E_h$ is equivalent to establishing

$$1 \geq \frac{\sum_{i,j=-b+1}^{n-b-1} \alpha_i \alpha_j \langle v_i' | X_g | v_j' \rangle}{\sum_{i,j=-b+1}^{n-b-1} \alpha_i \alpha_j \langle w_i' | X_h | w_j' \rangle}$$

for all real $\{\alpha_i\}_{i=-b+1}^{n-b-1}$. We know that $\langle x \rangle = \langle x^2 \rangle = \cdots = \langle x^{2n-3} \rangle = 0$. As we have the dependence on $\langle x_h^{2n-2} \rangle$, we can't conclude that the fraction is one. However, as we saw in the proof of Proposition 39, dependence on $\langle x_h^{2n-2} \rangle$ in the denominator only appears in the $\langle w_{n-b-1}' | X_h | w_{n-b-1}' \rangle$ term, that too with the positive coefficient, $1/c_{h_{n-b-1}}$. The analogous statement holds for the numerator. This, using $\langle x^{2n-2} \rangle > 0$, entails that the denominator is larger than or equal to the numerator, concluding the proof. □

4.4 Main result

Our observations so far can be combined to prove Theorem 2, which we formally state here.

Theorem 43. *Let t be an f-assignment (see Definition 32) on strictly positive coordinates (without loss of generality; see Lemma 36). Suppose f has real and strictly positive roots. Then, in order to obtain its effective solution (see Definition 34), it suffices to write it as $t = \sum_i \alpha_i t'_i$ where α_i are positive and t'_i are monomial assignments (see Definition 32 and Lemma 35). Furthermore, each t'_i admits a solution given by either Proposition 39, Proposition 40, Proposition 41, or Proposition 42.*

Proof. In Section 4.1 we established that it suffices to express an f-assignment as a sum of monomial assignments and find the solution for each one of them, in order to find the solution to the f-assignment. A monomial assignment now, can be balanced or unbalanced and aligned or misaligned (see Definition 32). The solution in each case is given by either Proposition 39, Proposition 40, Proposition 41, or Proposition 42. □

4.5 Example: a bias-$1/14$ protocol

The f-assignment for the TIPG approaching bias $\epsilon(3) = 1/14$ ($k = 3$ for $\epsilon(k) = \frac{1}{4k+2}$) has the following form. Let

$$x'_0 = 0 < r'_1 < r'_2 < x'_1 < x'_2 < x'_3 < x'_4 < x'_5 < x'_6 < r'_3 < r'_4 < r'_5.$$

This is an f-assignment (see Figure 12) on $\{x'_0, x'_1 \dots x'_6\}$ with $f'(x) = (r'_1 - x)(r'_2 - x)(r'_3 - x)(r'_4 - x)(r'_5 - x)$ viz.

$$t' = \sum_{i=0}^{6} \frac{-f'(x'_i)}{\prod_{j\neq i}(x'_j - x'_i)} \llbracket x'_i \rrbracket .$$

For a positive number Δ, we can consider an f-assignment on $\{x_0, x_1 \dots x_6\}$ where $x_i = x'_i + \Delta$, with $f(x) = (r_1 - x)(r_2 - x)\dots(r_5 - x)$ where $r_i = r'_i + \Delta$ viz.

$$t = \sum_{i=0}^{6} \frac{-f(x_i)}{\prod_{j\neq i}(x_j - x_i)} \llbracket x_i \rrbracket .$$

Lemma 36 guarantees that the solution to t and t' are the same. We decompose t into a sum of monomial assignments, i.e.

$$t = \underbrace{\sum_{i=0}^{6} \frac{-r_1r_2r_3r_4r_5}{\prod_{j\neq i}(x_j - x_i)} \llbracket x_i \rrbracket}_{\text{I}} + \underbrace{\sum_{i=0}^{6} \frac{-\overbrace{(r_2r_3r_4r_5 + r_1r_3r_4r_5 + r_1r_2r_3r_5 + r_1r_2r_3r_4)}^{:=\alpha_1}(-x_i)}{\prod_{j\neq i}(x_j - x_i)} \llbracket x_i \rrbracket}_{\text{II}}$$

$$+ \underbrace{\sum_{i=0}^{6} \frac{-\alpha_2(-x_i)^2}{\prod_{j\neq i}(x_j - x_i)} \llbracket x_i \rrbracket}_{\text{III}} + \underbrace{\sum_{i=0}^{6} \frac{-\alpha_3(-x_i)^3}{\prod_{j\neq i}(x_j - x_i)} \llbracket x_i \rrbracket}_{\text{IV}} + \underbrace{\sum_{i=0}^{6} \frac{-\alpha_4(-x_i)^4}{\prod_{j\neq i}(x_j - x_i)} \llbracket x_i \rrbracket}_{\text{V}} + \underbrace{\sum_{i=0}^{6} \frac{-\alpha_5(-x_i)^5}{\prod_{j\neq i}(x_j - x_i)} \llbracket x_i \rrbracket}_{\text{VI}},$$

where α_l is the coefficient of $(-x)^l$ in $f(x)$. Since the total number of points in each assignment are 7, they are unbalanced monomial assignments. Terms I, III and V each have an even powered monomial therefore they correspond to the aligned case. Their solutions, thus, are given in Proposition 41. Analogously, the remaining terms II, IV and VI each have an odd powered monomial therefore they correspond to the misaligned case. Their solutions, thus, given in Proposition 42.

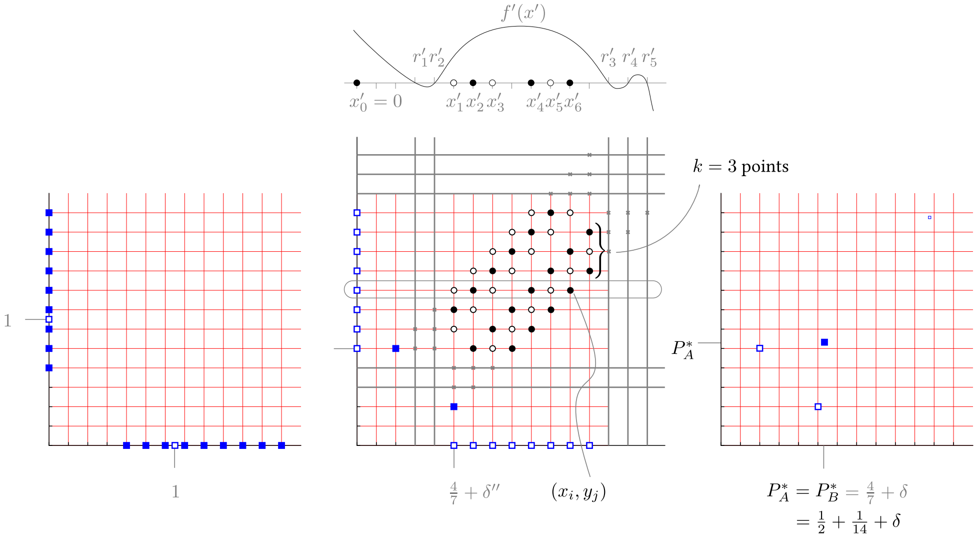


Figure 12: The TDPG (or equivalently, the reversed protocol) approaching bias $\epsilon(k = 3) = 1/14$ may be seen as proceeding in three stages, as illustrated by the three images (left to right). *First*, the initial points (indicated by unfilled squares) are split along the axes (indicated by the filled squares). *Second*, the points on the axes (unfilled squares) are transferred, by means of the ladder described in Section 2.5 (indicated by the circles), into two final points (filled squares). *Third*, the two points from the previous step (unfilled squares) and the catalyst state (indicated, after being raised into one point by the little unfilled box) are merged into the final point (filled box). The second stage is illustrated by the TIPG,—or more precisely, by its main move, the ladder—approaching bias $1/14$. The weight of these points is given (up to a constant) by the f−assignment shown above. The roots of the polynomial correspond to the locations of the vertical lines and the location of the points in the graph is representative of the general construction.

Let us now see how all these pieces fit together to give the full protocol. We describe the procedure in the language of TDPGs each step of which can be thought of as a short-hand to denote an exchange and manipulation of qubits between Alice and Bob, granted that the associated unitaries are known. As we have already done all the hard work in finding these unitaries[18], we can now proceed at this level of description. Concretely, the bias $1/14$ game (see Figure 12) goes as follows:

1. The first frame. This simply corresponds to the function $\frac{1}{2}\left(\llbracket 0,1 \rrbracket + \llbracket 1,0 \rrbracket\right)$.

2. The split. Deposit weights along the axis as specified by the TIPG; more precisely, split the point $\llbracket 0,1 \rrbracket$ into a set of points along the y−axis and analogously, split the point $\llbracket 1,0 \rrbracket$ into a set of points along the x−axis, to match the distribution of points along the axis by the bias $1/14$ game.

3. The Catalyst State. Deposit a small amount of weight, δ_{catalyst}, at all the points that appear in the TIPG. This can be done by raising the points which are along the y−axis, i.e. if the points along the axes are denoted as $\sum_i p_{\text{split},i} \llbracket 0, y_i \rrbracket$, then raise them to obtain $\sum_i (p_{\text{split},i} - \delta_{\text{split},i}) \llbracket 0, y_i \rrbracket + \sum_{i,j} \delta_{\text{catalyst}} \llbracket x_i, y_j \rrbracket$, where $\delta_{\text{catalyst}} > 0$ can be chosen to be arbitrarily small and the second sum is over points (x_i, y_j) which appear in the TIPG (excluding the axes[19]).

4. The Ladder.

[18] In this section we found the unitaries for f-assignments and in Section 3 we found those corresponding to splits and merges.

[19] One needs to use the analogous procedure, i.e. use $\sum_i p_{\text{split},i} \llbracket x_i, 0 \rrbracket$ as well for the one point of the TIPG which has a y−coordinate smaller than that of the points along the y−axis.

   (a) Denote the monomial decomposition of the valid functions by constituent valid functions. Globally scale these constituent valid functions sufficiently so that no negative weight appears when they are applied.
   (b) Apply all the scaled down constituent horizontal valid functions.
   (c) Apply all the scaled down constituent vertical valid functions.
   (d) Repeat these two steps until all the weight has been transferred from the axes into the two final points of the ladder[20].

   The unitaries corresponding to these constituent valid functions correspond to the solutions of the monomial assignments.

5. Raise and merge. Raise and merge the last two points into the point $(1-\delta')\,[\![\frac{4}{7}+\delta'', \frac{4}{7}+\delta'']\!]$ where δ' represents the total weight used by the catalyst, while δ'' comes from the truncation of the ladder. Then, using the method developed in the proof of Theorem 27 in [Aha+14b; Moc07], the catalyst state can be absorbed to obtain a single point $[\![\frac{4}{7}+\delta, \frac{4}{7}+\delta]\!]$. Thus, $P_A^* = P_B^* = \frac{1}{2}+\frac{1}{14}+\delta$, where δ can be made arbitrarily small by making the catalyst state smaller and the ladder longer.

The protocol is the reverse: it starts with a single point corresponding to uncorrelated states and whose coordinates encode the cheating probabilities, and ends with two points along the axis with equal weights, corresponding to the state $\frac{|AA\rangle+|BB\rangle}{\sqrt{2}}$.

[20] It would automatically become impossible to apply the moves once the weights on the axes becomes sufficiently small.

5 Elliptic Monotone Align (EMA) algorithm

So far we have exclusively studied C. Mochon's point games. In the following, we construct a numerical algorithm that generates the required unitary for any given Λ-valid function (see Definition 120). Note that corresponding to any WCF protocol with valid functions, one can find a WCF protocol with strictly valid functions (see Lemma 132), which in turn are Λ-valid for some finite Λ (see Lemma 129, Corollary 123); thus we do not lose generality by restricting to Λ-valid functions.

5.1 The Canonical Projective Form (CPF) and the Canonical Orthogonal Form (COF)

One can formalize the TEF constraint Equation (5) by associating to each transition, what we call a *Canonical Projective Form* (CPF). This essentially compiles the coordinates and weights which define a transition, into vectors and matrices.

Definition 44 (Canonical Projective Form (CPF) for a given transition)**.** For a given transition (see Definition 14) the *Canonical Projective Form (CPF)* is given by the set of $m \times m$ matrices X_h, X_g, O, D and m-dimensional vectors $|v\rangle$, $|w\rangle$ given by

$$X_h := \sum_{i=1}^{n_h} x_{h_i} |h_i\rangle \langle h_i| \, , \, X_g := \sum_{i=1}^{n_g} x_{g_i} |g_i\rangle \langle g_i| \, ,$$

$$|w\rangle := \sum_{i=1}^{n_h} \sqrt{p_{h_i}} \, |h_i\rangle \, , \, |v\rangle := \sum_{i=1}^{n_g} \sqrt{p_{g_i}} \, |g_i\rangle \, ,$$

$$D := X_h - E_h O X_g O^\dagger E_h$$

and O is a unitary which satisfies $O|v\rangle = |w\rangle$,
where $E_h = \sum |h_i\rangle \langle h_i|$, $\left\{ |g_1\rangle , |g_2\rangle \dots |g_{n_g}\rangle , |h_1\rangle , |h_2\rangle \dots |h_{n_h}\rangle \right\}$ are orthonormal basis vectors and $m = n_g + n_h$.

A transition satisfying the TEF constraint Equation (5) corresponds to $D \geq 0$ for the associated CPF, motivating the following definition.

Definition 45 (legal CPF)**.** A CPF is *legal* if $D \geq 0$.

So far, we have only recompiled the TEF constraint, into a supposedly more usable form, i.e. into a CPF. However, removing the projector from the CPF, allows one to interpret the *legality* condition above, geometrically and this, as we mentioned in the introduction, is at the heart of the EMA algorithm. To this end, we define a *Canonical Orthogonal Form* for each transition.

Definition 46 ((n,ξ)-COF for a transition, ξ-COF for a transition)**.** For a given transition (see Definition 14) and two numbers $n \geq \max(n_h, n_g)$ and $\xi \geq \max(x_{h_1}, x_{h_2} \dots x_{h_{n_h}})$, an (n,ξ)*-COF* is given by the set of $n \times n$ matrices X_h, X_g, O, D and vectors $|v\rangle$, $|w\rangle$ where

$$X_h := \text{diag}(x_{h_1}, x_{h_2} \dots , x_{h_{n_h}}, \xi, \xi \dots) \text{ and } X_g := \text{diag}(x_{g_1}, x_{g_2} \dots , x_{g_{n_g}}, 0, 0 \dots),$$

$$|v\rangle := \sum_{i=1}^{n_g} \sqrt{p_{g_i}} \, |i\rangle \text{ and } |w\rangle := \sum_{i=1}^{n_h} \sqrt{p_{h_i}} \, |i\rangle \, ,$$

$$D := X_h - O X_g O^T$$

and O is an orthogonal matrix which satisfies $O|v\rangle = |w\rangle$.

A ξ*-COF* is an (n,ξ)-COF with $n = n_h + n_g - 1$.

The legality condition this time, is defined in the limit of $\xi \to \infty$. We explain this shortly.

Definition 47 (n-legal COF, legal COF)**.** An (n,ξ)-COF is an *n-legal COF* if $D \geq 0$ in the limit $\xi \to \infty$. A *legal COF* is a ξ-COF such that $D \geq 0$ in the limit $\xi \to \infty$.

The EMA algorithm would produce a legal COF for a given Λ-valid transition. However, to use TEF, we need a legal CPF for the transition. We therefore first observe that a legal COF also yields a legal CPF.

Imagine one finds a legal COF corresponding to some transition. Then, they can sandwich D between a positive matrix, say E, as EDE to get

$$\left[\begin{array}{c|ccc} X_h & & & \\ \hline & 1 & & \\ & & \ddots & \\ & & & 1 \end{array}\right] - \underbrace{\left[\begin{array}{ccc|ccc} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ \hline & & & \xi^{-1/2} & & \\ & & & & \ddots & \\ & & & & & \xi^{-1/2} \end{array}\right]}_{:=E} UX_gU^\dagger \left[\begin{array}{ccc|ccc} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ \hline & & & \xi^{-1/2} & & \\ & & & & \ddots & \\ & & & & & \xi^{-1/2} \end{array}\right].$$

Note that $D \geq 0 \iff EDE \geq 0$, because E is diagonal. Since the COF is legal, $D \geq 0$ for $\xi \to \infty$ and in this limit E becomes a projector. After some relabeling (and appropriately expanding the space to $m = n_g + n_h$ dimensions) the inequality reduces to a CPF. This observation readily extends to the n-legal case where $n \leq n_g + n_h$. These arguments establish the following statement:

Proposition 48. *Consider a transition (see Definition 14). If there exists an n-legal COF corresponding to it, then there also exists a legal CPF for this transition.*

One can also show the reverse, i.e. that given a legal CPF one can find the corresponding $m-$legal COF. In particular, we are given

$$\left[\begin{array}{c|ccc} X_h & & & \\ \hline & 0 & & \\ & & \ddots & \\ & & & 0 \end{array}\right] - \underbrace{\left[\begin{array}{ccc|ccc} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ \hline & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{array}\right]}_{:=E_h} U \left[\begin{array}{c|c} 0 & \\ \hline & X_g \end{array}\right] U^\dagger \left[\begin{array}{ccc|ccc} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ \hline & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{array}\right] \geq 0.$$

Replacing the appended diagonal zeros in the first matrix (the one containing X_h) with 1s yields an equivalent inequality. Next, we can conjugate by a permutation matrix to get

$$\left[\begin{array}{c|c} 0 & \\ \hline & X_g \end{array}\right] = \tilde{U} \left[\begin{array}{c|c} X_g & \\ \hline & 0 \end{array}\right] \tilde{U}.$$

Finally, we write the diagonal zeros in E_h as $1/\xi^{1/2}$ and use the reverse of the trick above to recover an $m-$legal COF with $m := n_g + n_h$. This sketches the proof of the following proposition. The full proof requires some care, and since we do not use it further, we omit it.

Proposition 49. *Consider a transition (see Definition 14). If there exists a legal CPF corresponding to it, then there also exists an m-legal COF for this transition with $m := n_g + n_h$.*

Two aspects of the definition of COFs merit further discussion. To this end, recall that to convert an arbitrary point game into a protocol (with essentially the same bias), it suffices to consider Λ-valid transitions. Further, a Λ-valid transition is also an EBM transition (see Corollary 123). Thus, it suffices to show that EBM transitions admit a COF (for details, see Appendix E). With this in mind, the first aspect of the definition of COFs which is noteworthy, is that we used orthogonal matrices. This was because one can show that EBM transitions are the same as EBRM transitions, i.e. EBM but with the additional constraint that the matrices involved are real (see Appendix E.1). The second noteworthy aspect is that the dimensions of the matrices involved is taken to be $n_g + n_h$ while for EBM/EBRM, we did not have any such constraint. It turns out one can always generate matrices of size $n_g + n_h - 1$ from arbitrary sized matrices corresponding to an EBM/EBRM transition (see Appendix E.2). In the following, we discuss the geometric interpretation of the COF.

5.2 The inequality as containment of ellipsoids and Convex Geometry tools

In this section, we show that the matrix inequality of the form $0 \leq G \leq H$ can be geometrically viewed as the containment of an ellipsoid inside another. Consider an unnormalized vector $|u\rangle = \sum_j u_j \left|h_j\right\rangle$ with $u_j \in \mathbb{R}$. The set $\{|u\rangle \mid \langle u| X_h |u\rangle = 1\}$, where $X_h = \operatorname{diag}(x_{h_1}, x_{h_2} \dots)$, describes the surface of an ellipsoid. This is easy to see as the constraint corresponds to

$$x_{h_1} u_1^2 + x_{h_2} u_2^2 + \cdots = 1 \text{ which is of the form } \frac{u_1^2}{a_1^2} + \frac{u_2^2}{a_2^2} + \cdots = 1.$$

This is the equation of an ellipsoid in the variables $\{u_i\}$ with axes $a_1 = 1/\sqrt{x_{h_1}}, a_2 = 1/\sqrt{x_{h_2}} \dots$. An inequality would correspond to points inside or outside the ellipsoid. If we start with some arbitrary (possibly unnormalized) vector $|u\rangle$, then the point on the ellipse along this direction is given by $\mathcal{E}_h(|u\rangle) = \frac{|u\rangle}{\sqrt{\langle u|X_h|u\rangle}}$. Finally, the set $\left\{|u\rangle \mid \langle u| OX_gO^\dagger |u\rangle = 1\right\}$ also corresponds to the equation of an ellipsoid with axes $\left\{1/\sqrt{x_{g_i}}\right\}$ except that it is rotated by O. We can define a similar map from a vector $|u\rangle$ to a point on the rotated ellipsoid as $\mathcal{E}_g(|u\rangle) = \frac{|u\rangle}{\sqrt{\langle u|UX_gU^\dagger|u\rangle}}$. Accordingly, our inequality can be seen as

$$\begin{aligned} X_h - OX_gO^\dagger \geq 0 &\iff \langle u| X_h |u\rangle - \langle u| OX_gO^\dagger |u\rangle \geq 0 \qquad \forall\, |u\rangle \\ &\iff \langle u| OX_gO^\dagger |u\rangle \leq 1 \qquad \forall \{|u\rangle \mid \langle u| X_h |u\rangle = 1\}, \end{aligned}$$

which means that every point denoted by $|u\rangle$ on the h-ellipsoid must be on or inside the g-ellipsoid. If $\langle x_h\rangle - \langle x_g\rangle = 0$, then for $|u\rangle = |w\rangle$ the inequality saturates. This, in turn, means that also for $\mathcal{E}_h(|w\rangle)$ the inequality is saturated as it is the same vector up to a scaling, with the difference being that $\mathcal{E}_h(|w\rangle)$ is a point on the h-ellipsoid. Since the inequality is saturated it means that the ellipsoids must touch at this point, i.e., $\mathcal{E}_g(|w\rangle) = \mathcal{E}_h(|w\rangle)$; moreover, the curvature of the smaller ellipsoid should be larger than the curvature of the larger ellipsoid at the point of contact. Therefore, the curvature of an ellipsoid and the normal vectors along the direction of the contact are important in our analysis.

To formalize our approach we apply tools from Convex Geometry in the case of ellipsoids. One can find more details in Appendix **??** and a presentation of these results on general convex bodies in Section 2.5 of the book by R. Schneider [Sch09]. The central tool of our analysis is the so-called *Weingarten map*, which— defined intuitively—is the differential of the normal vector at a given point on the ellipsoid (or any manifold in general). Employing the Weingarten map permits us to evaluate curvatures and related quantities at the point of contact, as briefly described below. Let the point of contact be $|c\rangle = \sum c_i |i\rangle$. Then, the normal vector at this point is $|u(c)\rangle = \mathcal{N}\left(\sum x_i c_i |i\rangle\right)$, where we denote $\mathcal{N}(|\psi\rangle) := |\psi\rangle / \sqrt{\langle\psi|\psi\rangle}$. For a normalized direction vector $|u\rangle$ the support function corresponding to an ellipsoid X is given by

$$h(u) = \sqrt{\langle u| X^{-1} |u\rangle} = \sqrt{\sum x_i^{-1} u_i^2}. \tag{20}$$

The derivative of the support function yields the point on the ellipsoid where the tangent plane corresponding to the direction $|u\rangle$ touches the said ellipsoid, and it is given as $\partial_i h(u) = \frac{x_i^{-1} u_i}{h(u)}$. Furthermore, the second derivative of the support function

$$\partial_j \partial_i h(u) = \frac{1}{h(u)} \left(-\frac{x_j^{-1} x_i^{-1} u_i u_j}{h^2(u)} + x_i^{-1} \delta_{ij} \right) \tag{21}$$

contains as eigenvalues the radii of the curvature of the ellipsoid at the aforesaid point and as eigenvectors the directions of the principle curvature. Hereafter, we omit the $1/h$ factor as it cancels out in the equations of interest.

5.3 Definitions and lemmas for the EMA algorithm

Solving the WCF problem has been reduced to finding explicit matrices for a given EBM transition $g = \sum_{i=1}^{n_g} p_{g_i} [\![x_{g_i}]\!] \longrightarrow h = \sum_{i=1}^{n_h} p_{h_i} [\![x_{h_i}]\!]$ where g and h have disjoint support or, equivalently, for a given EBM function $a = h - g = \sum_{i=1}^{n_h} p_{h_i} [\![x_{h_i}]\!] - \sum_{i=1}^{n_g} p_{g_i} [\![x_{g_i}]\!]$. Here we describe our EMA algorithm, which runs by converting iteratively the above problem into the same problem of one less dimension until it is solved. We start by setting the notation involved.

At step k of the iteration, the transition $g \longrightarrow h$ and the associated function $a = h - g$ are given by $g^{(k)} \longrightarrow h^{(k)}$ and $a^{(k)}$, respectively, and they remain fixed for each step. Therefore, we do not write an explicit dependence on it in the following definitions. By $[x_{\min}, x_{\max}]$ we denote the smallest interval containing $\operatorname{supp}(a)$, and we call it the *support domain* for a. Similarly, we refer to the smallest interval containing $\operatorname{supp}(g) \cup \operatorname{supp}(h)$ as the *transition support domain* for the transition $g \longrightarrow h$. We use the variables $\chi, \xi \in \mathbb{R}$ to denote an interval $[\chi, \xi] \supseteq [x_{\min}, x_{\max}]$. In Section 1.1.3 we motivated the use of operator monotone functions to make the ellipsoids touch along a certain direction; the following definitions are tailored to this purpose.

Definition 50 (f_λ on (α, β))**.** $f_\lambda : (\alpha, \beta) \longrightarrow \mathbb{R}$ is defined for $\lambda \in \mathbb{R} \backslash [-\beta, -\alpha]$ as $f_\lambda(x) := \frac{-1}{\lambda + x}$.

Definition 51 (f_λ on $[\alpha, \beta]$)**.** $f_\lambda : [\alpha, \beta] \longrightarrow \mathbb{R} \cup \{\infty, -\infty\}$[21] is defined for $\lambda \in \mathbb{R} \backslash (-\beta, -\alpha)$. For $\lambda \in \mathbb{R} \backslash [-\beta, -\alpha]$ we define $f_\lambda(x) := \frac{-1}{\lambda + x}$.

For $\lambda = -\beta$ and $-\alpha$ we keep the same defn except for $x = \beta, \alpha$ in which case we define $f_{-\beta}(\beta) := \infty$ and $f_{-\alpha}(\alpha) := -\infty$.

As we also described in Section 1.1.3, we have to expand the smaller ellipsoid until it touches the larger one. An ellipsoid corresponding to a positive diagonal matrix, X_h (as in Section 5.2), is smaller than another ellipsoid corresponding to γX_h for $0 < \gamma < 1$. If the X_h matrix corresponds to a function h, what would be the corresponding function for γX_h? The following definition of h_γ formalizes the answer. We also introduce l_γ which helps us check the validity condition for a transition.

Definition 52 ($l_\gamma, l_\gamma^1, a_\gamma$)**.** Consider the transition $g \longrightarrow h$ and let $a = h - g$. For $\gamma \in (0, 1]$ we define the finitely supported functions $h_\gamma : \mathbb{R} \longrightarrow [0, \infty)$ and $a_\gamma(x) : \mathbb{R} \longrightarrow \mathbb{R}$ as

$$h_\gamma(x) := h(x/\gamma) \quad \text{and} \quad a_\gamma(x) := h_\gamma(x) - g(x).$$

Let $S_\gamma = [x_{\min}(\gamma), x_{\max}(\gamma)]$ be the support domain of a_γ.
We define $l_\gamma : \mathbb{R} \backslash [-x_{\max}(\gamma), -x_{\min}(\gamma)] \longrightarrow \mathbb{R}$ as

$$l_\gamma(\lambda) := \sum_{x \in \operatorname{supp}(a_\gamma)} a_\gamma(x) f_\lambda(x),$$

where f_λ is defined on S_γ. We also define $l_\gamma^1 := \sum_{x \in \operatorname{supp}(a_\gamma)} a_\gamma(x) x$.

[21] This is the extended real line with $1/\infty = -1/\infty := 0$.

We now define a function, m, which, given the transition $g \longrightarrow h$, indicates if the transition corresponding to the scaled ellipsoid $g \longrightarrow h_\gamma$ is valid.

Definition 53 ($m(\gamma, \chi, \xi)$)**.** We define $m : ((0,1], \mathbb{R}, \mathbb{R}) \longrightarrow \{0,1\}$ to be

$$m(\gamma, \chi, \xi) := \begin{cases} 0 & \text{if any of the following root conditions hold} \\ 1 & \text{else,} \end{cases}$$

where the first root condition is satisfied if there exists a $\lambda \in \mathbb{R}\backslash(-\xi, -\chi)$ such that $l_\gamma(\lambda) = 0$, and the second root condition is satisfied if $l_\gamma^1 = 0$.

As we are dealing with different representations of the same object, it is useful to define a relation between the matrix instance of the problem—involving matrices—and its function instance that involves transitions and functions.

Definition 54 (Matrix Instance, $\underline{\mathrm{X}}$ $\longrightarrow$ Function Instance, $\underline{\mathrm{x}}$)**.** For a *Matrix Instance* defined as the tuple $\underline{\mathrm{X}} := (X_h, X_g, |w\rangle, |v\rangle)$, where X_h, X_g are diagonal matrices and $|w\rangle, |v\rangle$ are vectors on $\mathbb{R}^n$ for some n with equal norm, i.e. $\langle w|w\rangle = \langle v|v\rangle$, we define the *Function Instance* to be the tuple $\underline{\mathrm{x}} : (g, h, a)$, where $h = \mathrm{Prob}[X_h, |w\rangle]$, $g = \mathrm{Prob}[X_g, |v\rangle]$ and $a = h - g$.

Definition 55 (Attributes of the Function Instance, $\underline{\mathrm{x}}$)**.** For a given tuple $\underline{\mathrm{x}} := (g, h, a)$ as in Definition 54 we define the attributes $n_h, n_g, \{p_{g_i}\}, \{p_{h_i}\}, \{x_{g_i}\}, \{x_{h_i}\}$ as they appear by declaring $g \longrightarrow h$ to be a transition, i.e.,

- n_h as the number of times h is non-zero,
- n_g as the number of times g is non-zero,
- $\{p_{h_i}\}, \{x_{h_i}\}, \{p_{g_i}\}, \{x_{g_i}\}$ implicitly as $h = \sum_{i=1}^{n_h} p_{h_i} [\![x_{h_i}]\!]$, $g = \sum_{i=1}^{n_g} p_{g_i} [\![x_{g_i}]\!]$ for $p_{h_i}, p_{g_i} > 0$.

The *support domain* for a is denoted by $[x_{\min}, x_{\max}]$, i.e., the attributes $x_{\min}, x_{\max}$ are defined to be such that $[x_{\min}, x_{\max}]$ is the smallest interval containing $\mathrm{supp}(a)$.

Definition 56 (Attributes of the Matrix Instance, $\underline{\mathrm{X}}$)**.** For a tuple $\underline{\mathrm{X}}$ as defined in Definition 54 we have the following:

- *Spectral domain*: The *spectral domain* of $\underline{\mathrm{X}}$ is denoted by $[\chi, \xi]$ where the attributes χ, ξ are such that $[\chi, \xi]$ is the smallest interval containing $\mathrm{spec}\{X_g \oplus X_h\}$.
- *Solution*: We say O is a *solution* to the matrix instance $\underline{\mathrm{X}} = \big(X_h, X_g, |w\rangle, |v\rangle\big)$ if $X_h \geq O X_g O^T$ and $O|v\rangle = |w\rangle$.
- *Notation*: With respect to a standard orthonormal basis $\{|i\rangle\}$, we use the *notation* $X_h := \sum_{i=1}^k y_{h_i} |i\rangle\langle i|$, $X_g := \sum_{i=1}^k y_{g_i} |i\rangle\langle i|$, $|w\rangle := \sum_{i=1}^k \sqrt{q_{h_i}} |i\rangle$, and $|v\rangle := \sum_{i=1}^k \sqrt{q_{g_i}} |i\rangle$.

With the notation in place, we can now state and prove some results that we need in our algorithm. First, we generalize the results obtained in [Aha+14b] about operator monotones and their relation with EBM functions. Second, we prove some results that formalize the intuitive notions of tightening—stretching the smaller ellipsoid until it touches the larger ellipsoid. Finally, we generalize them in the case where the curvature of the smaller ellipsoid becomes infinite.

5.3.1 Generalizations

Our approach for achieving the aforementioned generalization is schematically represented in Figure 13. Our starting point would be the relations between EBM, EBRM and COF which we already outlined (see the discussion above Section 5.2) while deferring the details to Appendix E. For readers familiar with those details, our main objective here is twofold. First, we generalize the relation between EBM and EBRM functions (see Corollary 144) from matrices with their spectrum in $[0, \Lambda]$ to matrices with their spectrum in $[\chi, \xi]$. Second, we extend the result from valid functions to valid transitions, including the case of overlapping support. To establish the first, our strategy is to find a relation between $[0, \Lambda]$-valid functions (see Definition 120) and $[\chi, \xi]$-valid functions (which we define shortly) and then a relation between $[0, \Lambda]$-EBRM functions (see Definition 140) and $[\chi, \xi]$-EBRM functions (which again, we define shortly). Then we use the link (see Corollary 144) between $[0, \Lambda]$-valid and $[0, \Lambda]$-EBRM functions to establish the equivalence of $[\chi, \xi]$-valid functions and $[\chi, \xi]$-EBRM functions. Along the way we sharpen our understanding of operator monotone functions which should make the definitions of f_λ, l and m (see Definition 52 and Definition 53) obvious. The second objective can be met with by a single, albeit, slightly long argument.

$$\begin{tikzcd}
\begin{array}{c} a(x) \text{ is} \\ \Lambda - \text{valid} \end{array} \arrow[r, Leftrightarrow, "\text{Corollary Corollary 144}"] \arrow[d, Leftrightarrow, "\text{Cor. Corollary 65}"] & \begin{array}{c} a(x) \text{ is} \\ \Lambda - \text{EBRM} \end{array} \arrow[d, Leftrightarrow, "\text{Cor. Corollary 67}"] & \\
\begin{array}{c} a(x' - \chi) \text{ is} \\ [\chi, \xi] - \text{valid} \end{array} \arrow[r, Leftrightarrow, "\text{Lemma 68}"] & \begin{array}{c} a(x' - \chi) \text{ is} \\ [\chi, \xi] - \text{EBRM} \end{array} \arrow[d, Leftrightarrow, "\text{Lemma 69}"] & \\
 & \begin{array}{c} h \longrightarrow g \text{ is} \\ [\chi', \xi'] \text{ EBRM} \\ \text{if } a = h - g \text{ and} \\ [\chi, \xi] \cup \text{spec}(g + h) \subset [\chi', \xi'] \end{array} \arrow[r, Rightarrow, shift left, "\text{Lemma 146}"] \arrow[r, Leftarrow, shift right, "\text{trivial}"'] & {[\chi', \xi'] - \text{COF}}
\end{tikzcd}$$

Figure 13: Generalization schematised.

Let us start with extending our definition of the COF to accommodate matrices with their spectrum in $[\chi, \xi]$.

Definition 57 (COF with spectrum in $[\chi, \xi]$)**.** For a given transition $g \longrightarrow h$ let $[\chi, \xi]$ be such that it contains $\text{supp}(g)$ and $\text{supp}(h)$. We define the COF with its spectrum in $[\chi, \xi]$ by the set of $n \times n$ matrices X_h, X_g, O, D and vectors $|v\rangle$, $|w\rangle$ where

$$X_h := \text{diag}\{x_{h_1}, x_{h_2} \dots, x_{h_{n_h}}, \xi, \xi \dots\} \text{ and } X_g := \text{diag}\{x_{g_1}, x_{g_2} \dots, x_{g_{n_g}}, \chi, \chi \dots\},$$

$$|v\rangle := \sum_{i=1}^{n_g} \sqrt{p_{g_i}} \, |i\rangle \text{ and } |w\rangle := \sum_{i=1}^{n_h} \sqrt{p_{h_i}} \, |i\rangle \, ,$$

$$D := X_h - O X_g O^\dagger, \qquad n = n_g + n_h - 1$$

and O is an orthogonal matrix which satisfies $|v\rangle = O\,|w\rangle$.

Definition 58 (Legal COF with spectrum in $[\chi, \xi]$)**.** A COF with spectrum in $[\chi, \xi]$ is legal if $D \geq 0$.

Definition 59 (Operator monotone functions on $[\chi, \xi]$)**.** A function $f : [\chi, \xi] \to \mathbb{R}$ is operator monotone on $[\chi, \xi]$ if for all real symmetric matrices H, G with $\mathrm{spec}(H \oplus G) \in [\chi, \xi]$ and $H \geq G$ we have $f(H) \geq f(G)$.

Lemma 60. *$f(x)$ is an operator monotone function on $[\chi, \xi]$ if and only if $f'(x') = f(x' - x_0)$ is an operator monotone function on $[\chi + x_0, \xi + x_0]$.*

Proof. Consider real symmetric matrices $H \geq G$ with $\mathrm{spec}(H \oplus G) \in [\chi, \xi]$ and let $f(x)$ be operator monotone on $[\chi, \xi]$. We must consider $f'(x') = f(x = x' - x_0)$ which is the same as $f'(x + x_0) = f(x)$. We show that f' is an operator monotone on $[\chi + x_0, \xi + x_0]$. Note that $H' := H + x_0\mathbb{I}$ and $G' := G + x_0\mathbb{I}$ are such that $H' \geq G'$ and $\mathrm{spec}(H' \oplus G') \in [\chi + x_0, \xi + x_0]$. Note that $f'(H') = f(H)$ and $f'(G') = f(G)$ because

$$f'(H') = f'(H + x_0\mathbb{I}) = O_h f'(H_d + x_0\mathbb{I}) O_h^T = O_h f(H_d) O_h^T = f(H)$$

and similarly for G where $H = O_h H_d O_h^T$ for O_h orthogonal and H_d diagonal. Since f is operator monotone on $[\chi, \xi]$ we have $f(H) \geq f(G)$ which entails $f'(H') \geq f'(G')$. Since this holds for all H', G' with their $\mathrm{spec}(H' \oplus G') \in [\chi + x_0, \xi + x_0]$ we can conclude that f' is an operator monotone on $[\chi + x_0, \xi + x_0]$. The other way follows by setting $\chi + x_0$ to χ, $\xi + x_0$ to ξ, x_0 to $-x_0$ but since all these were arbitrary to start with, the reasoning goes through unchanged. □

Corollary 61 (Characterisation of operator monotone functions on $[0, \Lambda]$)**.** *Any operator monotone function $f : [0, \Lambda] \to \mathbb{R}$ can be written as*

$$f(x) = c_0 + c_1 x - \int \frac{1}{\lambda + x} d\tilde{\omega}(\lambda)$$

with the integral ranging over $\lambda \in (-\infty, -\Lambda) \cup (0, \infty)$ satisfying $\int \frac{1}{\lambda(1+\lambda)} d\tilde{\omega}(\lambda) < \infty$.

Proof. Consider the characterisation given in Lemma 118 according to which we had $f(x) = c_0' + c_1 x + \int \frac{\lambda x}{\lambda + x} d\omega(\lambda)$ with $\int \frac{\lambda}{1+\lambda} d\omega(\lambda) < \infty$. We can write

$$f(x) = c_0' + c_1 x + \int \left(\lambda - \frac{\lambda^2}{\lambda + x}\right) d\omega(\lambda) = c_0 + c_1 x - \int \frac{\lambda^2 d\omega(\lambda)}{\lambda + x}$$

where with $d\tilde{\omega} = \lambda^2 d\omega(\lambda)$ we obtain the claimed form. Note that the finiteness of $\int \frac{\lambda}{1+\lambda} d\omega$ is necessary to conclude that $c_0 = c_0' + \int \frac{\lambda}{1+\lambda} d\omega$ is also finite. □

Corollary 62 (Characterisation of operator monotone functions on $[\chi, \xi]$)**.** *Any operator monotone function $f' : [\chi, \xi] \to \mathbb{R}$ can be written as*

$$f'(x') = c_0' + c_1' x' - \int \frac{1}{\lambda' + x'} d\tilde{\omega}'(\lambda')$$

with the integral over $\lambda' \in (-\infty, -\xi) \cup (-\chi, \infty)$ satisfying $\int \frac{1}{(\lambda' + \chi)(1 + \lambda' + \chi)} d\tilde{\omega}'(\lambda') < \infty$.

Proof. We follow the convention that $x' \in [\chi, \xi]$ while $x \in [0, \xi - \chi]$. From Lemma 60 we know that $f(x)$ is operator monotone on $[0, \xi - \chi]$ if and only if $f'(x') = f(x' - \chi)$ is operator monotone on $[\chi, \xi]$ where

$x' = x + \chi$. Since we already have a characterisation for $f(x)$ we can characterise $f'(x')$ as $f(x' - \chi)$. From Corollary 61 we have

$$f'(x') = c_0 + c_1(x' - \chi) - \int \frac{d\tilde{\omega}(\lambda)}{\lambda + x' - \chi} = c_0' + c_1 x' - \int \frac{d\tilde{\omega}'(\lambda')}{\lambda' + x'}$$

where $\lambda' = \lambda - \chi$. Since we had $\lambda \in (-\infty, -(\xi - \chi)) \cup (0, \infty)$ it entails $\lambda' \in (-\infty, -\xi) \cup (-\chi, \infty)$. The condition on the integral $\int \frac{d\tilde{\omega}(\lambda)}{\lambda(\lambda+x)} < \infty$ can be expressed in terms of λ' as $\int \frac{d\tilde{\omega}'(\lambda')}{(\lambda'+\chi)(1+\lambda'+\chi)} < \infty$ with $d\tilde{\omega}'(\lambda') = d\tilde{\omega}(\lambda' + \chi)$. With $c_1 = c_1'$ and $c_0' = c_0 - c_1\chi$ we obtain the claimed form. □

We now generalize the definition of Λ-valid functions to $[\chi, \xi]$−valid functions.

Definition 63 ($[\chi, \xi]$−valid function)**.** A finitely supported function $a : \mathbb{R} \to \mathbb{R}$ with $\mathrm{supp}(a) \in [\chi, \xi]$ is $[\chi, \xi]$−valid if for every operator monotone function f on $[\chi, \xi]$ we have $\sum_{x \in \mathrm{supp}(a)} a(x) f(x) \geq 0$.

Remark 64. Since in Corollary 62 $d\tilde{\omega}'$ is a measure, to establish $[\chi, \xi]$ validity of functions, it would suffice to restrict our attention to operator monotones $f'(x') = x'$, $f'(x') = -\frac{1}{\lambda'+x'}$ with $x' \in [\chi, \xi]$, $\lambda' \in (-\infty, -\xi) \cup (-\chi, \infty)$.

By shifting the characterisation of operator monotone functions we can shift valid functions as well.

Corollary 65 ($a(x)$ is $[\chi, \xi]$−valid $\iff$ $a(x' - x_0)$ is $[\chi + x_0, \xi + x_0]$−valid)**.** *A finitely supported function $a : \mathbb{R} \to \mathbb{R}$ with supp$(a) \in [\chi, \xi]$ is $[\chi, \xi]$−valid if and only if the function $a'(x') := a(x' - x_0) : [0, \infty) \to \mathbb{R}$ is $[\chi - x_0, \xi - x_0]$−valid.*

Proof. a is $[\chi, \xi]$ valid entails $\sum_{x \in \mathrm{supp}(a)} a(x) f(x) \geq 0$ for all f operator monotone on $[\chi, \xi]$. We can write the sum as $\sum a(x' - x_0) f(x' - x_0) \geq 0$. Using Lemma 60 we note that $f'(x') = f(x' - x_0)$ is operator monotone on $[\chi + x_0, \xi + x_0]$. For $a'(x') = a(x' - x_0)$ we thus have $\sum a'(x') f'(x') \geq 0$ which means $a'(x')$ is a $[\chi + x_0, \xi + x_0]$−valid function. The other way follows similarly. □

We have, thus, established a relation between $[0, \Lambda]$-valid functions and $[\chi, \xi]$- valid functions, and we proceed to establish its analogue for EBRM functions.

Definition 66 (EBRM on $[\chi, \xi]$)**.** A finitely supported function $a : \mathbb{R} \to \mathbb{R}$ is EBRM on $[\chi, \xi]$ if there exist real symmetric matrices $H \geq G$ with their spectrum in $[\chi, \xi]$ and a vector $|w\rangle$ such that $a = \mathrm{Prob}[H, |w\rangle] - \mathrm{Prob}[G, |w\rangle]$.

Corollary 67 ($a(x)$ is EBRM on $[\chi, \xi]$ $\iff$ $a(x + \chi)$ is EBRM on $[0, \xi - \chi]$)**.** *A finitely supported function $a : \mathbb{R} \to \mathbb{R}$ with supp$(a) \in [\chi, \xi]$ is EBRM on $[\chi, \xi]$ if and only if the function $a'(x) := (x + \chi) : [0, \infty) \to \mathbb{R}$ is EBRM on $[0, \xi - \chi]$.*

Proof. If a is EBRM on $[\chi, \xi]$ it follows that there exist real symmetric matrices with $H \geq G$ and a vector $|w\rangle$ such that $\mathrm{spec}[H \oplus G] \in [\chi, \xi]$ and $a = \mathrm{Prob}[H, |w\rangle] - \mathrm{Prob}[G, |w\rangle]$. Clearly, $H' := H - \chi\mathbb{I} \geq G - \chi\mathbb{I} =: G'$ and $a'(x) = \mathrm{Prob}[H', |w\rangle] - \mathrm{Prob}[G', |w\rangle] = a(x + \chi)$ with $\mathrm{spec}[H' \oplus G'] \in [0, \xi - \chi]$. This means a' is EBRM on $[0, \xi - \chi]$. The other way follows similarly. □

We combine the above to prove the equivalence between $[\chi, \xi]$-valid and $[\chi, \xi]$-EBRM function:

Lemma 68 ($a(x)$ is $[\chi, \xi]$−valid function $\iff$ $a(x)$ is EBRM on $[\chi, \xi]$)**.** *A finitely supported function $a : \mathbb{R} \to \mathbb{R}$ with supp$(a) \in [\chi, \xi]$ being $[\chi, \xi]$-valid is equivalent to it being $[\chi, \xi]$-EBRM.*

Proof. From Corollary 65 we know that $a(x)$ being $[\chi, \xi]$-valid is equivalent to $a(x + \chi)$ being Λ-valid with $\Lambda = \xi - \chi$. From Corollary 144 we know that $a(x + \chi)$ is equivalently EBRM on $[0, \xi - \chi]$. Finally using Corollary 67 we know that $a(x + \chi)$ being EBRM on $[0, \xi - \chi]$ is equivalent to $a(x)$ being EBRM on $[\chi, \xi]$. □

It only remains to extend the result from valid functions to valid transitions:

Lemma 69 (EBRM function $\iff$ EBRM transition even with common support)**.** *If we write an EBRM function a with spectrum in $[\chi', \xi']$ as $a = h - g$ with $h, g : [0, \infty) \to [0, \infty)$ which may have common support, then $g \to h$ is an EBRM transition with spectrum in $[\chi, \xi]$ and with the smallest matrix size being at most $n_g + n_h - 1$, where $[\chi, \xi]$ is the smallest interval containing $[\chi', \xi']$ and supp$(h) \cup$ supp(g).*

Conversely, if $g \to h$ is an EBRM transition with spectrum in $[\chi, \xi]$ and with the smallest matrix size being at most $n_g + n_h - 1$, with $h, g : [0, \infty) \to [0, \infty)$ which may have common support, then $a = h - g$ is an EBRM function with its spectrum in $[\chi, \xi]$.

Proof. To prove the first statement we write $a = a^+ - a^-$ with $a^+ = \sum_{i=1}^{n'_h} p'_{h_i} [\![x_{h_i}]\!]$ and $a^- = \sum_{i=1}^{n'_g} p'_{g_i} [\![x_{g_i}]\!]$, for $a^+, a^- : [0, \infty) \to [0, \infty)$, represent the positive and the negative parts of a. Note that a^+ and a^- by virtue of this definition can't have any common support. Consider $\Delta = \sum_{i=1}^{n_\Delta} c_i [\![x_i]\!] : [0, \infty) \to [0, \infty)$ to be such that $h = a^+ + \Delta$ and $g = a^- + \Delta$. This is always the case because $h - g = a$. Consider the case where $\text{supp}(\Delta) \cap \text{supp}(a) = \emptyset$. In this case $n_g = n'_g + n_\Delta$ and $n_h = n'_h + n_\Delta$. Since a is an EBRM function we have a legal COF, viz $O' X'_g O'^T \leq X'_h$ and $|w'\rangle = O' |v'\rangle$, of dimension $n' = n'_g + n'_h - 1$ from Lemma 146. To obtain the matrices corresponding to $g \to h$ we expand the space to $n = n_g + n_h - 1$ dimensions and define $X_g = X'_g \oplus X$, $X_h = X'_h \oplus X$, $O = O' \oplus \mathbb{I}$, $|v\rangle = |v'\rangle + \sum_{i=n'}^{n} \sqrt{c_{i+1-n'}} |i\rangle$ where $X = \text{diag}\{x_1, x_2 \ldots x_{n_\Delta}\}$. This is just an elaborate way of adding the points in Δ to the matrices and the vectors in such a way that the part corresponding to Δ remains unchanged. The other cases can be similarly demonstrated with the only difference being in the relation between n_g, n'_g and n_h, n'_h. Suppose Δ is non-zero only at one point. If Δ adds a point where a^- had a point then it does not contribute to increasing the number of points in g (that is $n_g = n'_g$), but it does increase the number in h (that is $n_h = n'_h + 1$). This means that we have one extra dimension to find the matrices certifying $g \to h$ is EBRM which is precisely what is needed to append that extra idle point as described above. Similarly one can reason for adding a point where a^+ had a point and finally extend it to the most general case of $\text{supp}(\Delta) \cap \text{supp}(a) \neq \emptyset$ which may involve multiple points.

For the converse, since $g \to h$ is an EBRM transition from Lemma 146 we know that it admits a legal COF, that is $O X_g O^T \leq X_h$ and $O |v\rangle = |w\rangle$ with dimension $n_g + n_h - 1$. To show that $a = h - g = a^+ - a^-$ is an EBRM function it suffices to show that a is a valid function. This follows directly from the COF and operator monotones as $O f(X_g) O^T \leq f(X_h)$ implies $\langle v| f(X_g) |v\rangle \leq \langle w| f(X_h) |w\rangle$ which in turn is $\sum h(x) f(x) - \sum g(x) f(x) \geq 0$ and that is the same as $\sum a(x) f(x) \geq 0$ for all f operator monotone on the spectrum of $X_h \oplus X_g$, viz. a is valid. From Lemma 68 we conclude that a is also EBRM with size at most $n_g + n_h - 1$. □

5.3.2 For the finite part

In this subsection we formalize the notions of tightening and stretching of ellipsoids and present some results that are relevant to perform the tightening/stretching in the course of the EMA algorithm.

Fact 70 (Weyl's Monotonicity Theorem (see [Bha13]))**.** *If H is positive semi-definite and A is Hermitian then $\lambda_j^\downarrow(A + H) \geq \lambda_j^\downarrow(A)$ for all j where $\lambda_j^\downarrow(M)$ represents the j^{th} largest eigenvalue of the Hermitian matrix M.*

Corollary 71. *If $H \geq G$ then $\lambda_j^\downarrow(H) \geq \lambda_j^\downarrow(G)$ for all j.*

We now state a continuity condition which we subsequently use to establish that when we stretch the h ellipsoid, there would always exist the perfect amount of stretching that makes the h ellipsoid just touch the g ellipsoid. The non-triviality here is that we have to conclude this without fully knowing the ellipsoids.

Claim 72 (Continuity of l)**.** *Let $[x_{\min}, x_{\max}]$ be the smallest interval containing supp(a). $l(\lambda)$ is continuous in the intervals $\lambda \in (-x_{\min}, \infty]$ and $\lambda \in [-\infty, -x_{\max})$ (see Definition 52).*

Proof. Since $l(\lambda)$ is just a rational function of λ it suffices to show that the denominator doesn't become zero in the said range. The roots of the denominator are of the form $\lambda + x = 0$ for $x \in \{\{x_{g_i}\}, \{x_{h_i}\}\}$. Hence the largest root is $\lambda = -x_{\text{min}}$ and the smallest $\lambda = -x_{\text{max}}$. Neither of the intervals defined in the statement contain any roots and therefore we can conclude that $l(\lambda)$ is continuous therein. Note that the function f_λ on $[x_{\text{min}}, x_{\text{max}}]$ is not even defined for λ in $(-x_{\text{max}}, -x_{\text{min}})$. □

Lemma 73 (Tightening with the matrix spectrum unknown)**.** *Consider a finitely supported valid function a. Let $[x_{\text{min}}(\gamma), x_{\text{max}}(\gamma)]$ be the smallest interval containing supp(a_γ) (see Definition 52). Consider $m(\gamma, x_{\text{min}}(\gamma), x_{\text{max}}(\gamma))$ as a function of γ (see Definition 53). Then, m has at least one root in the interval* $(0, 1]$.

Proof. To prove the claim it suffices to show that $l_\gamma(\lambda)$ has a root in the range $(0, \infty)$ for some $\gamma \in (0, 1]$. Note that we are given a valid function a which means supp$(a) \in [0, \infty)$. We assume that $l_{\gamma=1}(\lambda) > 0$ for all $\lambda \in (0, \infty)$ because if this was not the case then we trivially have $\gamma = 1$ as a root, i.e. $m(1, x_{\text{min}}(1), x_{\text{max}}(1)) = 0$. Since $\sum h(x) = \sum g(x)$, we have

$$\lambda l(\lambda) = \sum h(x)(\lambda f_\lambda(x) + 1) - \sum g(x)(\lambda f_\lambda(x) + 1) = \sum h(x)\frac{x}{\lambda + x} - \sum g(x)\frac{x}{\lambda + x}.$$

Therefore for the remainder of this proof we redefine $f_\lambda = \frac{1}{\lambda}\frac{x}{\lambda+x}$ without changing the value of l or by extension l_γ (the $1/\lambda$ factor is partly why we restricted λ to $(0, \infty)$ instead of the more general $(-x_{\text{min}}, \infty)$). Note that $\lim_{\gamma\to 0^+} l_\gamma(\lambda) < 0$ for all $\lambda \in (0, \infty)$ because $h_\gamma(x) = h(x/\gamma)$ which means $\lim_{\gamma\to 0} \sum h_\gamma(x) f_\lambda(x) = \lim_{\gamma\to 0} \sum h(x) f_\lambda(\gamma x) = 0$ since $\lim_{x\to 0} f_\lambda(x) = 0$. This, in turn, means that $\lim_{\gamma\to 0^+} l_\gamma(\lambda) = -\sum g(x) f_\lambda(x) < 0$. Each term constituting $l_\gamma(\lambda)$ is finite for $\lambda \in (0, \infty)$ since for $\lambda > 0$ the denominators are of the form $\lambda + x$ which are always positive. Hence $l_\gamma(\lambda)$ as a function of $\lambda \in [0, \infty)$ and $\gamma \in (0, 1]$ is continuous. By continuity then between $\gamma = 0^+$ and $\gamma = 1$ there should be a root.

It remains to justify why we extended the range of λ from $(0, \infty)$ to $(-\infty, -x_{\text{max}}) \cup (-x_{\text{min}}, \infty)$ in the definition of m (see Definition 53) as it appears in the statement of the lemma. This is due to the fact that $l_\gamma(\lambda)$ is continuous for λ in the stated range (see Lemma 72) and so there might be a root which appears in the extended range. □

Once we are guaranteed that there is at least one perfect stretching amount, we want to know the spectrum of the matrices. We state a slightly more general result which is a direct consequence of the previous results.

Lemma 74 (Matrix spectrum from a valid function)**.** *Consider a valid function a, i.e. an a such that $l(\lambda) \geq 0$ and $l^1 \geq 0$ for all $\lambda \in [0, \infty)$ (see Definition 52) and let $[\chi, \xi]$ be such that for all $\lambda \in [-\infty, -\xi) \cup (-\chi, \infty]$ we have $l(\lambda) \geq 0$. Then, there exists a legal COF, corresponding to the function a, with its spectrum contained in* $[\chi, \xi]$.

Proof. Since $l(\lambda) \geq 0$ for $\lambda \in (-\infty, -\xi) \cup (-\chi, \infty)$ and $l^1 \geq 0$ we know from Corollary 62 that a is $[\chi, \xi]$-valid. From Lemma 68 we know that a is EBRM on $[\chi, \xi]$. Finally, from Lemma 146 we know that there exists a legal COF with spectrum in $[\chi, \xi]$. □

Recall that our analysis involves operator monotone functions f with the property that f^{-1} is also an operator monotone (see Section 1.1.3). We now establish that f_λs (see Definition 51) also have this property.

Lemma 75 ($H \geq G \iff f_\lambda(H) \geq f_\lambda(G)$)**.** *Let H, G be real symmetric matrices, $[\chi, \xi]$ be the smallest interval containing spec$[H \oplus G]$ and f_λ be on (χ, ξ) (see Definition 50; f_λ is defined for $\lambda \in \mathbb{R}\backslash[-\xi, -\chi]$). Then, $H \geq G$ if and only if $f_\lambda(H) \geq f_\lambda(G)$.*

Proof. $H \geq G \implies f_\lambda(H) \geq f_\lambda(G)$ because f_λ is an operator monotone function for matrices with spectrum in $[\chi, \xi]$. We prove the converse. We find the inverse function of f_λ and show that it is also an operator monotone. Start with recalling that for $x \in [\chi, \xi]$ we have

$$y = f_\lambda(x) = \frac{-1}{\lambda + x} \implies x = -\frac{1}{y} - \lambda$$

where $\lambda \in \mathbb{R} \backslash [\chi, \xi]$. Thus $f_\lambda^{-1}(y) = -\frac{1}{y} - \lambda$. For a given λ either $f_\lambda(\chi)$ and $f_\lambda(\xi)$ are both greater than zero or both less than zero. Hence the operator monotones $f'_{\lambda'}(y)$ on $[f_\lambda(\chi), f_\lambda(\xi)]$ permit $\lambda' = 0$. Consequently $f'_{\lambda'=0}(y) = \frac{-1}{y}$ is an operator monotone on $[f_\lambda(\chi), f_\lambda(\xi)]$. A constant is also an operator monotone. Thus we conclude $f_\lambda^{-1}(y)$ is an operator monotone on the required interval establishing the converse. □

5.3.3 For the infinite part; wiggle-v

The results of the previous section permit us to tighten the ellipsoids as needed, and after that we need to find the operator monotone f_λ for which the ellipsoids touch along a certain direction. Under the action of this operator monotone it is possible that the curvature along some direction becomes infinite, i.e., the corresponding matrix has a divergence (see the case of an ellipse getting mapped to a line). Our algorithm fails in this situation because the associated normal vector is ill-defined. To remedy this problem, we show that tightness is preserved under the action of f_λ. This means that if for some λ' we consider the ellipsoids obtained by applying $f_{\lambda'}$ and we find that they touch along a certain direction, then for some other $\lambda'' \neq \lambda'$ they continue to touch but along a different direction. This allows us to consider the sequence leading to the divergence, and we use it in the analysis of the algorithm. We start by showing this result in the case where everything is well-defined, and then extend it to the divergent case.

Lemma 76 (Strict inequality under f_λ). *$H > G$ if and only if $f_\lambda(H) > f_\lambda(G)$ where f_λ is on $(\chi, \xi) \supset$ spec$[H \oplus G]$.*

Proof. Note that $H > G \iff H' := H + \lambda\mathbb{I} > G + \lambda\mathbb{I} =: G'$, where $\lambda \in (\mathbb{R} \cup \{\infty, -\infty\}) \backslash [-\xi, -\chi]$ by Definition 50. There can be two cases, either both matrices are strictly positive or both are strictly negative. Let us consider the former, and the latter follows similarly. We have

$$H' > G' > 0 \iff \mathbb{I} > H'^{-1/2} G' H'^{-1/2} \iff \mathbb{I} < H'^{1/2} G'^{-1} H'^{1/2} \iff H'^{-1} < G'^{-1},$$

where the first and third inequalities follow from the fact that multiplication by a positive matrix doesn't affect the inequality and the second follows from matrix diagonalization. The last inequality is the same as $f_\lambda(H) > f_\lambda(G)$ completing our proof. □

Corollary 77 (Tightness preservation under f_λ). *Let $H \geq G$ and f_λ be on $(\chi, \xi) \supset$ spec$[H \oplus G]$. There exists a $|w\rangle$ such that $\langle w| (H - G) |w\rangle = 0$ if and only if there exists a $|w_\lambda\rangle$ such that $\langle w_\lambda| (f_\lambda(H) - f_\lambda(G)) |w_\lambda\rangle = 0$.*

Proof. The contrapositive of the aforesaid condition is that $f_\lambda(H) > f_\lambda(G)$ if and only if $H > G$ which holds from Lemma 76. □

Lemma 78 (Extending tightness preservation under f_λ to apparently divergent situations). *Let X_h, X_g be diagonal matrices with spec$[X_h] \in (\chi, \xi]$, spec$[X_g] \in [\chi, \xi)$ and let f_λ be on $[-\xi, -\chi]$. Let, further, O be an orthogonal matrix such that $X_h \geq O X_g O^T$.*

There exists a vector $|w\rangle$ such that $\langle w| \left(f_{-\xi}(X_h) - O f_{-\xi}(X_g) O^T\right) |w\rangle = 0$ if and only if there exists a $|w_\lambda\rangle$ such that $\langle w_\lambda| \left(f_\lambda(X_h) - O f_\lambda(X_g) O^T\right) |w_\lambda\rangle = 0$ for a $\lambda \in \mathbb{R} \backslash (-\xi, -\chi)$.

Similarly, there exists a vector $|w\rangle$ such that $\langle w| \left(f_{-\chi}(X_h) - O f_{-\chi}(X_g) O^T\right) |w\rangle = 0$ if and only if there exists a $|w_\lambda\rangle$ such that $\langle w_\lambda| \left(f_\lambda(X_h) - O f_\lambda(X_g) O^T\right) |w_\lambda\rangle = 0$ for a $\lambda \in \mathbb{R} \backslash (-\xi, -\chi)$.

Proof. In this tightness statement the problem is that X_h has an eigenvalue ξ which means that $f_{-\xi}(X_h)$ is not well-defined. We assume that X_h can be expressed as

$$X_h = \begin{bmatrix} X'_h & \\ & \xi\mathbb{I}'' \end{bmatrix}$$

where X'_h has no eigenvalue equal to ξ and $\mathbb{I}''$ is the identity matrix in the subspace. We can write

$$\begin{aligned} X_h > OX_gO^T &\iff \begin{bmatrix} f_\lambda(X'_h) & \\ & f_\lambda(\xi\mathbb{I}'') \end{bmatrix} > Of_\lambda(X_g)O^T \text{ for } \lambda \in \mathbb{R}\backslash[-\xi,-\chi] \\ &\iff \begin{bmatrix} f_\lambda(X'_h) & \\ & \mathbb{I}'' \end{bmatrix} > \begin{bmatrix} \mathbb{I}' & \\ & f_\lambda(\xi\mathbb{I}'')^{-1/2} \end{bmatrix} Of_\lambda(X_g)O^T \begin{bmatrix} \mathbb{I}' & \\ & f_\lambda(\xi\mathbb{I}'')^{-1/2} \end{bmatrix} \lambda \in \mathbb{R}\backslash[-\xi,-\chi] \end{aligned}$$

where the last expression has a well-defined limit for $\lambda = -\xi$. This establishes the contrapositive of the statement we wanted to prove once we note the following: If $\langle w| \left(f_{-\xi}(X_h) - Of_{-\xi}(X_g)O^T\right) |w\rangle = 0$, then $\begin{bmatrix} 0 & \\ & \mathbb{I}'' \end{bmatrix} |w\rangle = 0$, otherwise due to the spectrum constraint of X_g the aforesaid expression would be ∞. This entails

$$\langle w| \left(\begin{bmatrix} f_{-\xi}(X'_h) & \\ & \mathbb{I}'' \end{bmatrix} - \begin{bmatrix} \mathbb{I}' & \\ & f_{-\xi}(\xi\mathbb{I}'')^{-1/2} \end{bmatrix} Of_{-\xi}(X_g)O^T \begin{bmatrix} \mathbb{I}' & \\ & f_{-\xi}(\xi\mathbb{I}'')^{-1/2} \end{bmatrix} \right) |w\rangle = 0.$$

The proof for the case of $f_{-\chi}(X_g)$ follows similarly. □

5.4 The algorithm

We first present our algorithm and then we motivate and prove the correctness of each step.

Definition 79 (EMA Algorithm)**.** Given a finitely supported function a (we assume it is Λ-valid (see Definition 120)) proceed in the following three phases:

- **PHASE 1: INITIALIZATION**
  - **Tightening procedure**: Let $[x_{\min}(\gamma'), x_{\max}(\gamma')]$ be the support domain for $a_{\gamma'}$ (see Definition 52) where $\gamma' \in (0,1]$ is a variable. Let $\gamma \in (0,1]$ be the largest root of $m(\gamma', x_{\min}(\gamma'), x_{\max}(\gamma'))$, and let $x_{\max} := x_{\max}(\gamma)$ and $x_{\min} := x_{\min}(\gamma)$.
  - **Spectral domain for the representation**: Find the smallest interval $[\chi, \xi]$ such that $l_\gamma(\lambda) \geq 0$ for $\lambda \in (\mathbb{R} \cup \{\infty, -\infty\})\backslash[\chi, \xi]$. If $\text{supp}(g), \text{supp}(h)$ is not contained in $[\chi, \xi]$ then from all expansions of $[\chi, \xi]$ that contain the aforesaid sets, pick the smallest. Relabel this interval to $[\chi, \xi]$.
  - **Shift**: Transform $a(x) \longrightarrow a'(x') := a(x' + \chi - 1)$, where instead of 1 any positive constant would do (see Corollary 67). Similarly transform

    $$g(x) \longrightarrow g'(x') := g(x' + \chi - 1) \qquad \text{and} \qquad h(x) \longrightarrow h'(x') := h(x' + \chi - 1).$$

    Relabel a' to a, g' to g and h' to h.
  - **The matrices**: For $n := n_g + n_h - 1$ define $n \times n$ matrices with spectrum in $[\chi, \xi]$ and n-dimensional vectors as

    $$X_g^{(n)} = \text{diag}[\chi, \chi, \dots x_{g_1}, x_{g_2} \dots, x_{g_{n_g}}] \text{ and } X_{h_\gamma}^{(n)} = \text{diag}[\gamma x_{h_1}, \gamma x_{h_2}, \dots, \gamma x_{h_{n_h}}, \xi, \xi, \dots],$$
    $$\left|v^{(n)}\right\rangle \doteq \left[0, 0 \dots, \sqrt{p_{g_1}}, \sqrt{p_{g_2}}, \dots, \sqrt{p_{g_{n_g}}}\right], \left|w^{(n)}\right\rangle \doteq \left[\sqrt{p_{h_1}}, \sqrt{p_{h_2}}, \dots, \sqrt{p_{h_{n_h}}}, 0, 0 \dots\right],$$

    where $g = \sum_{i=1}^{n_g} p_{g_i} [\![x_{g_i}]\!]$ and $h = \sum_{i=1}^{n_h} p_{h_i} [\![x_{h_i}]\!]$.

  - **Bootstrapping the iteration**:
    - Basis: $\left\{\left|t_{h_i}^{(n+1)}\right\rangle\right\}$ where $\left|t_{h_i}^{(n+1)}\right\rangle := |i\rangle$ for $i = 1, 2 \dots n$ with $|i\rangle$ referring to the standard basis in which the matrices and the vectors were originally expressed.
    - Matrix Instance: $\underline{\mathrm{X}}^{(n)} = \{X_h^{(n)}, X_g^{(n)}, \left|w^{(n)}\right\rangle, \left|v^{(n)}\right\rangle\}$.

- **PHASE 2: ITERATION**
  - Objective: Find the objects $\left|u_h^{(k)}\right\rangle, \bar{O}_g^{(k)}, \bar{O}_h^{(k)}$ and $s^{(k)}$.
  - Input: Assume we are given
    - Basis: $\left\{\left|t_{h_i}^{(k+1)}\right\rangle\right\}$
    - Matrix Instance: $\underline{\mathrm{X}}^{(k)} = \left(X_h^{(k)}, X_g^{(k)}, \left|w^{(k)}\right\rangle, \left|v^{(k)}\right\rangle\right)$ with attribute $\chi^{(k)} > 0$
    - Function Instance: $\underline{\mathrm{X}}^{(k)} \to \underline{\mathrm{x}}^{(k)} = \left(h^{(k)}, g^{(k)}, a^{(k)}\right)$
  - Output:
    - Basis: $\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_i}^{(k)}\right\rangle\right\}$
    - Matrix Instance: $\underline{\mathrm{X}}^{(k-1)} = \left(X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle\right)$ with attribute $\chi^{(k-1)} > 0$
    - Function Instance: $\underline{\mathrm{X}}^{(k-1)} \to \underline{\mathrm{x}}^{(k-1)} = \left(h^{(k-1)}, g^{(k-1)}, a^{(k-1)}\right)$
    - Unitary Constructors: Either $\bar{O}_g^{(k)}$ and $\bar{O}_h^{(k)}$ are returned or $\bar{O}^{(k)}$ is returned. If $\bar{O}^{(k)}$ is returned, set $\bar{O}_g^{(k)} := \bar{O}^{(k)}$ and $\bar{O}_h^{(k)} = \mathbb{I}$.
    - Relation: If $s^{(k)}$ is not specified, define $s^{(k)} := 1$.
      If $s^{(k)} = 1$ then use
      $$O^{(k)} := \bar{O}_h^{(k)} \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}_g^{(k)}$$
      else use
      $$O^{(k)} := \left[\bar{O}_h^{(k)} \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}_g^{(k)}\right]^T.$$
  - Algorithm:
    - **Boundary condition: If** $n_g = 0$ and $n_h = 0$ **then** set $k_0 = k$ and **jump to PHASE 3**.
    - **Tighten**: Define $X_{h_{\gamma'}}^{(k)} := \gamma' X^{(k)}$ where $\gamma' \in (0, 1]$ is a variable. Let γ be the largest root of $m(\gamma', \chi_{\gamma'}^{(k)}, \xi_{\gamma'}^{(k)})$ for $a^{(k)}$ where $\chi_{\gamma'}^{(k)}, \xi_{\gamma'}^{(k)}$ are such that $[\chi_{\gamma'}^{(k)}, \xi_{\gamma'}^{(k)}]$ is the smallest interval containing $\mathrm{spec}[X_{h_{\gamma'}}^{(k)} \oplus X_g^{(k)}]$. Relabel $X_{h_\gamma}^{(k)}$ to $X_h^{(k)}$, $\chi_\gamma^{(k)}$ to $\chi^{(k)}$ and $\xi_\gamma^{(k)}$ to $\xi^{(k)}$, and $a_\gamma^{(k)}$ to $a^{(k)}$, $h_\gamma^{(k)}$ to $h^{(k)}$, $l_\gamma^{(k)}$ to $l^{(k)}$. Update $x_{\min}$ and $x_{\max}$ to be such that $\mathrm{supp}(a^{(k)}) \in [x_{\min}^{(k)}, x_{\max}^{(k)}]$ is the smallest such interval. Define $s^{(k)} := 1$.
    - **Honest align**: **If** $l^{1(k)} = 0$ **then** define $\eta = -\chi^{(k)} + 1$
      $$X_h'^{(k)} := X_h^{(k)} + \eta, \quad X_g'^{(k)} := X_g + \eta.$$
      Else: Pick a root λ of the function $l^{(k)}(\lambda')$ in the domain $\mathbb{R} \backslash (-\xi^{(k)}, -\chi^{(k)})$. In the following two cases we consider the function f_λ on $[\chi^{(k)}, \xi^{(k)}]$.
      - If $\lambda \neq -\chi^{(k)}$ then: Let $\eta = -f_\lambda(\chi^{(k)}) + 1$ where any positive constant could be chosen instead of 1. Define
        $$X_h'^{(k)} := f_\lambda(X_h^{(k)}) + \eta, \quad X_g'^{(k)} := f_\lambda(X_g^{(k)}) + \eta.$$

- If $\lambda = -\chi^{(k)}$ then: Update $s^{(k)} = -1$. Let $\eta = -f_\lambda(\xi^{(k)}) - 1$ where any positive constant could be chosen instead of 1. Define

  $$X_h'^{(k)} := X_g''^{(k)}, \quad X_g'^{(k)} := X_h''^{(k)},$$

  where $X_h''^{(k)} := -f_\lambda(X_h^{(k)}) - \eta, X_g''^{(k)} := -f_\lambda(X_g^{(k)}) - \eta$, and make the replacements

  $$\left|v^{(k)}\right\rangle \to \left|w^{(k)}\right\rangle \text{ and } \left|w^{(k)}\right\rangle \to \left|v^{(k)}\right\rangle.$$

* **Remove spectral collision: If** $\lambda = -\chi^{(k)}$ or $\lambda = -\xi^{(k)}$ **then**
  1. **Idle point: If** for some j', j, we have $q_{g_{j'}}^{(k)} = q_{h_j}^{(k)}$ and $y_{g_{j'}}^{(k)} = y_{h_j}^{(k)}$ **then** the solution is given by Definition 81
     Jump to **End**.
  2. **Final Extra**: **If** for some j, j' we have $q_{g_{j'}}^{(k)} > q_{h_j}^{(k)}$ and $y_{g_{j'}}^{(k)} = y_{h_j}^{(k)}$ **then** the solution is given by Definition 82
     Jump to **End**.
  3. **Initial Extra**: **If** for some j, j' we have $q_{g_{j'}}^{(k)} < q_{h_j}^{(k)}$ and $y_{g_{j'}}^{(k)} = y_{h_j}^{(k)}$ **then** the solution is given by Definition 83
     Jump to **End**.
* **Evaluate the Reverse Weingarten Map**:
  1. Consider the point $\left|w^{(k)}\right\rangle / \sqrt{\left\langle w^{(k)}\right| X_h'^{(k)} \left|w^{(k)}\right\rangle}$ on the ellipsoid $X_h'^{(k)}$. Evaluate the normal at this point as $\left|u_h^{(k)}\right\rangle = \mathcal{N}\left(\sum_{i=1}^{n_h^{(k)}} \sqrt{p_{h_i}^{(k)}} x_{h_i}'^{(k)} \left|t_{h_i}^{(k+1)}\right\rangle\right)$. Similarly evaluate $\left|u_g^{(k)}\right\rangle$, the normal at the point $\left|v^{(k)}\right\rangle / \sqrt{\left\langle w^{(k)}\right| X_g'^{(k)} \left|w^{(k)}\right\rangle}$ on the ellipsoid $X_g'^{(k)}$.
  2. Evaluate the Reverse Weingarten maps $W_h'^{(k)}$ and $W_g'^{(k)}$ along $\left|u_h^{(k)}\right\rangle$ and $\left|u_g^{(k)}\right\rangle$, respectively. For a given diagonal matrix $X = \sum_i y_i \left|i\right\rangle \left\langle i\right| > 0$ and normal vector $\left|u\right\rangle = \sum_i u_i \left|i\right\rangle$ the Reverse Weingarten map is given by $W_{ij} = \left(-\frac{y_j^{-1} y_i^{-1} u_i u_j}{r^2} + y_i^{-1}\delta_{ij}\right)$ where $r = \sqrt{\sum y_i^{-1} u_i^2}$.
  3. Find the eigenvectors and eigenvalues of the Reverse Weingarten maps. The eigenvectors of W_h' form the h tangent and normal vectors $\left\{\left\{\left|t_{h_i}^{(k)}\right\rangle\right\}, \left|u_h^{(k)}\right\rangle\right\}$. The corresponding radii of curvature are obtained from the eigenvalues $\left\{\{r_{h_i}^{(k)}\}, 0\right\} = \left\{\{c_{h_i}^{(k)-1}\}, 0\right\}$ which are the inverses of the curvature values. The tangents are labeled in decreasing order of radii of curvature (which is increasing order of curvature). Similarly for the g tangent and normal vectors. Fix the sign freedom in the eigenvectors by requiring $\left\langle t_{h_i}^{(k)} | w^{(k)}\right\rangle \geq 0$ and $\left\langle t_{g_i}^{(k)} | v^{(k)}\right\rangle \geq 0$.
* **Finite Method: If** $\lambda \neq -\xi^{(k)}$ and $\lambda \neq -\chi^{(k)}$, (finite case) **then**
  1. $\bar{O}^{(k)} := \left|u_h^{(k)}\right\rangle \left\langle u_g^{(k)}\right| + \sum_{i=1}^{k-1} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{g_i}^{(k)}\right|$
  2. $\left|v^{(k-1)}\right\rangle := \bar{O}^{(k)} \left|v^{(k)}\right\rangle - \left\langle u_h^{(k)}\right| \bar{O}^{(k)} \left|v^{(k)}\right\rangle \left|u_h^{(k)}\right\rangle$
     and $\left|w^{(k-1)}\right\rangle := \left|w^{(k)}\right\rangle - \left\langle u_h^{(k)} | w^{(k)}\right\rangle \left|u_h^{(k)}\right\rangle$.
  3. Define $X_h^{(k-1)} := \mathrm{diag}\{c_{h_1}^{(k)}, c_{h_2}^{(k)} \ldots, c_{h_{k-1}}^{(k)}\}, X_g^{(k-1)} := \mathrm{diag}\{c_{g_1}^{(k)}, c_{g_2}^{(k)} \ldots c_{g_{k-1}}^{(k)}\}$.

    4. **Jump** to **End**.

  - **Wiggle-v Method: If** $\lambda = -\xi^{(k)}$ or $\lambda = -\chi^{(k)}$ (infinite case) **then**

    1. $\left|u_h^{(k)}\right\rangle$ renamed to $\left|\bar{u}_h^{(k)}\right\rangle$, $\left|u_g^{(k)}\right\rangle$ remains the same.
    2. Let $\tau = \cos\theta := \left\langle u_g^{(k)} | v^{(k)} \right\rangle / \left\langle \bar{u}_h^{(k)} | w^{(k)} \right\rangle$. Let $\left|\bar{t}_h^{(k)}\right\rangle$ be an eigenvector of $X_h^{\prime(k)-1}$ with zero eigenvalue. Redefine

       $$\left|u_h^{(k)}\right\rangle := \cos\theta \left|\bar{u}_h^{(k)}\right\rangle + \sin\theta \left|\bar{t}_h^{(k)}\right\rangle,$$

       $$\left|t_{h_k}^{(k)}\right\rangle = s\left(-\sin\theta \left|\bar{u}_h^{(k)}\right\rangle + \cos\theta \left|\bar{t}_h^{(k)}\right\rangle\right)$$

       where the sign $s \in \{1, -1\}$ is fixed by demanding $\left\langle t_{h_k}^{(k)} | w^{(k)} \right\rangle \geq 0$.
    3. $\bar{O}^{(k)}$ and $\left|v^{(k-1)}\right\rangle, \left|w^{(k-1)}\right\rangle$ are evaluated as step 1. and 2. of the finite case above.
    4. Define

       $$X_h^{\prime(k-1)} := \mathrm{diag}\{c_{h_1}^{(k)}, \ldots, c_{h_{k-1}}^{(k)}\}, X_g^{\prime(k-1)} := \mathrm{diag}\{c_{g_1}^{(k)}, \ldots, c_{g_{k-1}}^{(k)}\}.$$

       Let $[\chi'^{(k-1)}, \xi'^{(k-1)}]$ denote the smallest interval containing $\mathrm{spec}[X_h^{\prime(k-1)} \oplus X_g^{\prime(k-1)}]$. Let $\lambda' = -\chi'^{(k-1)} + 1$ where instead of 1 any positive number would also work. Consider $f_{\lambda''}$ on $[\chi'^{(k-1)}, \xi'^{(k-1)}]$, and let $\eta = -f_{\lambda'}(\chi'^{(k-1)}) + 1$. Define

       $$X_h^{(k-1)} := f_{\lambda'}(X_h^{\prime(k-1)}) + \eta, \qquad X_g^{(k-1)} := f_{\lambda'}(X_g^{\prime(k-1)}) + \eta.$$

    5. **Jump** to **End**.

  - **End**: Restart **PHASE 2** with the newly obtained $(k-1)$-sized objects.

- **PHASE 3: RECONSTRUCTION**

  Let k_0 be the iteration at which the algorithm stops. Using the relation

  $$O^{(k)} = \bar{O}_g^{(k)} \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)} \right) \bar{O}_h^{(k)}$$

  (or its transpose if $s^{(k)} = -1$), evaluate $O^{(k_1)}$ from $O^{(k_0)} := \mathbb{I}_{k_0}$, then $O^{(k_2)}$ from $O^{(k_1)}$, then $O^{(k_3)}$ from $O^{(k_2)}$ and so on until $O^{(n)}$ is obtained. $O^{(n)}$ solves the matrix instance $\underline{X}^{(n)}$ we started with. In terms of EBRM matrices, the solution is given by $H = X_h^{(n)}$, $G = O^{(n)} X_g O^{(n)T}$, and $|w\rangle = \left|w^{(n)}\right\rangle$.

Theorem 80 (Correctness of the EMA Algorithm)**.** *Given a* Λ*-valid function, a, the EMA algorithm (see Definition 79) always finds an orthogonal matrix O of size at most* $n \times n$ *where* $n = n_g + n_h$*, such that the constraints on O stated in Theorem 1 corresponding to the function a, are satisfied.*

Definition 81 (Spectral Collision: Case Idle Point)**.**

$$\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle, \left|t_{h_2}^{(k)}\right\rangle, \ldots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\} \overset{\text{component-wise}}{:=}$$
$$\left\{\left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \ldots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \ldots \left|t_{h_k}^{(k+1)}\right\rangle\right\},$$

$$\bar{O}^{(k)} := \sum_{i=1}^{k} |a_i\rangle \left\langle t_{h_i}^{(k+1)}\right|,$$

$$\text{where} \qquad \{|a_1\rangle, |a_2\rangle \dots |a_k\rangle\} \overset{\text{component-wise}}{:=}$$

$$\begin{cases} \left\{\left|t^{(k+1)}_{h_1}\right\rangle, \left|t^{(k+1)}_{h_2}\right\rangle, \dots \left|t^{(k+1)}_{h_{j-1}}\right\rangle, \left|t^{(k+1)}_{h_{j'}}\right\rangle, \left|t^{(k+1)}_{h_j}\right\rangle, \left|t^{(k+1)}_{h_{j+1}}\right\rangle, \dots, \left|t^{(k+1)}_{h_{j'-1}}\right\rangle, \left|t^{(k+1)}_{h_{j'+1}}\right\rangle \dots \left|t^{(k+1)}_{h_k}\right\rangle\right\} & j < j' \\ \left\{\left|t^{(k+1)}_{h_1}\right\rangle, \left|t^{(k+1)}_{h_2}\right\rangle, \dots \left|t^{(k+1)}_{h_{j'-1}}\right\rangle, \left|t^{(k+1)}_{h_{j'+1}}\right\rangle \dots \left|t^{(k+1)}_{h_{j-1}}\right\rangle, \left|t^{(k+1)}_{h_{j'}}\right\rangle, \left|t^{(k+1)}_{h_j}\right\rangle, \left|t^{(k+1)}_{h_{j+1}}\right\rangle \dots \left|t^{(k+1)}_{h_k}\right\rangle\right\} & j > j' \\ \left\{\left|t^{(k+1)}_{h_1}\right\rangle, \left|t^{(k+1)}_{h_2}\right\rangle, \dots \left|t^{(k+1)}_{h_k}\right\rangle\right\} & j = j', \end{cases}$$

and

$$X^{(k-1)}_h := \sum_{i \neq j} y^{(k)}_{h_i} \left|t^{(k+1)}_{h_i}\right\rangle \left\langle t^{(k+1)}_{h_i}\right| \text{ and } X^{(k-1)}_g := \bar{O}^{(k)} X^{(k)}_g \bar{O}^{(k)T} - y_{h_j} \left|t^{(k+1)}_{h_j}\right\rangle \left\langle t^{(k+1)}_{h_j}\right|,$$

$$\left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\left|w^{(k)}\right\rangle - \sqrt{p_{h_j}} \left|t^{(k+1)}_{h_j}\right\rangle\right] \text{ and } \left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\bar{O}^{(k)} \left|v^{(k)}\right\rangle - \sqrt{p_{h_j}} \left|t^{(k+1)}_{h_j}\right\rangle\right].$$

This specifies the matrix instance $\underline{\mathrm{X}}^{(k-1)} := \{X^{(k-1)}_h, X^{(k-1)}_g, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle\}$.

Definition 82 (Spectral Collision: Case Final Extra)**.** $\underline{\mathrm{X}}^{(k-1)} := (X^{(k-1)}_h, X^{(k-1)}_g, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle)$, where $X^{(k-1)}_h = \sum_{i=1}^{k-1} y^{(k-1)}_{h_i} \left|t^{(k)}_{h_i}\right\rangle \left\langle t^{(k)}_{h_i}\right|$, $X^{(k-1)}_g = \sum_{i=1}^{k-1} y^{(k-1)}_{g_i} \left|t^{(k)}_{h_i}\right\rangle \left\langle t^{(k)}_{h_i}\right|$, $\left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1} \sqrt{q^{(k-1)}_{g_i}} \left|t^{(k)}_{h_i}\right\rangle\right]$, $\left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1} \sqrt{q^{(k-1)}_{h_i}} \left|t^{(k)}_{h_i}\right\rangle\right]$ and the coordinates and weights are given by

$$\begin{aligned} \left\{q^{(k-1)}_{h_1}, \dots q^{(k-1)}_{h_{k-1}}\right\} &\overset{\text{component-wise}}{=} \left\{q^{(k)}_{h_1}, q^{(k)}_{h_2} \dots, q^{(k)}_{h_{j-1}}, q^{(k)}_{h_{j+1}}, \dots q^{(k)}_{h_k}\right\} \\ \left\{q^{(k-1)}_{g_1}, \dots q^{(k-1)}_{g_{k-1}}\right\} &\overset{\text{component-wise}}{=} \left\{q^{(k)}_{g_2} \dots, q^{(k)}_{g_{j'-1}}, q^{(k)}_{g_{j'}} - q^{(k)}_{h_j}, q^{(k)}_{g_{j'+1}}, q^{(k)}_{g_{j'+2}} \dots q^{(k)}_{g_k}\right\} \\ \left\{y^{(k-1)}_{g_1}, \dots y^{(k-1)}_{g_{k-1}}\right\} &\overset{\text{component-wise}}{=} \left\{y^{(k)}_{g_2}, \dots y^{(k)}_{g_k}\right\} \\ \left\{y^{(k-1)}_{h_1}, \dots y^{(k-1)}_{h_{k-1}}\right\} &\overset{\text{component-wise}}{=} \left\{y^{(k)}_{h_1}, \dots y^{(k)}_{h_{j-1}}, y^{(k)}_{h_{j+1}} \dots, y^{(k)}_{h_k}\right\}, \end{aligned}$$

the basis is given by

$$\left\{\left|u^{(k)}_h\right\rangle, \left|t^{(k)}_{h_1}\right\rangle \dots \left|t^{(k)}_{h_{k-1}}\right\rangle\right\} \overset{\text{component-wise}}{=}$$
$$\left\{\left|t^{(k+1)}_{h_j}\right\rangle, \left|t^{(k+1)}_{h_1}\right\rangle, \left|t^{(k+1)}_{h_2}\right\rangle, \dots \left|t^{(k+1)}_{h_{j-1}}\right\rangle, \left|t^{(k+1)}_{h_{j+1}}\right\rangle, \left|t^{(k+1)}_{h_{j+2}}\right\rangle \dots \left|t^{(k+1)}_{h_k}\right\rangle\right\}.$$

The orthogonal matrices are given by $\bar{O}^{(k)}_h := \sum \left|t^{(k+1)}_{h_i}\right\rangle \langle a_i|$ where

$$\{|a_1\rangle, \dots |a_k\rangle\} \to \left\{\left|u^{(k)}_h\right\rangle, \left|t^{(k)}_{h_1}\right\rangle \dots \left|t^{(k)}_{h_{k-1}}\right\rangle\right\} \text{ and } \bar{O}^{(k)}_g := \tilde{O}^{(k)} \bar{O}^{(k)}_h,$$

$$\begin{aligned} \text{with } \tilde{O}^{(k)} :=& \mathcal{N}\left[\sqrt{q^{(k)}_{h_j}} \left|u^{(k)}_h\right\rangle + \sqrt{q^{(k)}_{g_{j'}} - q^{(k)}_{h_j}} \left|t^{(k)}_{h_{j'}}\right\rangle\right] \mathcal{N}\left[\sqrt{q^{(k)}_{g_1}} \left\langle u^{(k)}_h\right| + \sqrt{q^{(k)}_{g_{j'}}} \left\langle t^{(k)}_{h_{j'}}\right|\right] \\ &+ \mathcal{N}\left[\sqrt{q^{(k)}_{g_{j'}} - q^{(k)}_{h_j}} \left|u^{(k)}_h\right\rangle - \sqrt{q^{(k)}_{h_j}} \left|t^{(k)}_{h_{j'}}\right\rangle\right] \mathcal{N}\left[\sqrt{q^{(k)}_{g_{j'}}} \left\langle u^{(k)}_h\right| - \sqrt{q^{(k)}_{g_1}} \left\langle t^{(k)}_{h_{j'}}\right|\right] \\ &+ \sum_{i \in \{1, \dots k\} \setminus j'} \left|t^{(k)}_{h_i}\right\rangle \left\langle t^{(k)}_{h_i}\right|. \end{aligned}$$

Definition 83 (Spectral Collision: Case Initial Extra). $\underline{\mathrm{X}}^{(k-1)} := (X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle)$ where $X_h^{(k-1)} = \sum_{i=1}^{k-1} y_{h_i}^{(k-1)} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{h_i}^{(k)}\right|$, $X_g^{(k-1)} = \sum_{i=1}^{k-1} y_{g_i}^{(k-1)} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{h_i}^{(k)}\right|$, $\left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1} \sqrt{q_{g_i}^{(k-1)}} \left|t_{h_i}^{(k)}\right\rangle\right]$, $\left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1} \sqrt{q_{h_i}^{(k-1)}} \left|t_{h_i}^{(k)}\right\rangle\right]$ and the coordinates and weights are given by

$$\begin{aligned}
\left\{q_{h_1}^{(k-1)}, \dots q_{h_{k-1}}^{(k-1)}\right\} &\overset{\text{component-wise}}{=} \left\{q_{h_1}^{(k)} \dots, q_{h_{j-1}}^{(k)}, q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}, q_{h_{j+1}}^{(k)}, q_{h_{j+2}}^{(k)} \dots q_{h_{k-1}}^{(k)}\right\} \\
\left\{q_{g_1}^{(k-1)}, \dots q_{g_{k-1}}^{(k-1)}\right\} &\overset{\text{component-wise}}{=} \left\{q_{g_1}^{(k)}, q_{g_2}^{(k)} \dots, q_{g_{j'-1}}^{(k)}, q_{g_{j'+1}}^{(k)}, \dots q_{g_k}^{(k)}\right\} \\
\left\{y_{g_1}^{(k-1)}, \dots y_{g_{k-1}}^{(k-1)}\right\} &\overset{\text{component-wise}}{=} \left\{y_{g_1}^{(k)}, \dots y_{g_{j'-1}}^{(k)}, y_{g_{j'+1}}^{(k)} \dots, y_{g_k}^{(k)}\right\} \\
\left\{y_{h_1}^{(k-1)}, \dots y_{h_{k-1}}^{(k-1)}\right\} &\overset{\text{component-wise}}{=} \left\{y_{h_1}^{(k)}, \dots y_{h_{k-1}}^{(k)}\right\},
\end{aligned}$$

the basis is given by

$$\begin{aligned}
&\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle \dots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\} \overset{\text{component-wise}}{=} \\
&\left\{\left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \left|t_{h_{j+2}}^{(k+1)}\right\rangle \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\}.
\end{aligned}$$

The orthogonal matrices are given by $\bar{O}_h^{(k)} := \tilde{O}^{(k)} \sum |a_i\rangle \left\langle t_{h_i}^{(k+1)}\right|$ where

$$\{|a_1\rangle, \dots |a_k\rangle\} \overset{\text{component-wise}}{=} \left\{\left|t_{h_1}^{(k)}\right\rangle, \left|t_{h_2}^{(k)}\right\rangle \dots \left|t_{h_{k-1}}^{(k)}\right\rangle, \left|u_h^{(k)}\right\rangle\right\},$$

$$\begin{aligned}
\tilde{O}^{(k)} :=& \mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}} \left|u_h^{(k)}\right\rangle + \sqrt{q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}} \left|t_{h_j}^{(k)}\right\rangle\right] \mathcal{N}\left[\sqrt{q_{h_k}^{(k)}} \left\langle u_h^{(k)}\right| + \sqrt{q_{g_j}^{(k)}} \left\langle t_{h_j}^{(k)}\right|\right] \\
&+ \mathcal{N}\left[\sqrt{q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}} \left|u_h^{(k)}\right\rangle - \sqrt{q_{g_{j'}}^{(k)}} \left|t_{h_j}^{(k)}\right\rangle\right] \mathcal{N}\left[\sqrt{q_{g_j}^{(k)}} \left\langle u_h^{(k)}\right| - \sqrt{q_{h_k}^{(k)}} \left\langle t_{h_j}^{(k)}\right|\right] \\
&+ \sum_{i \in \{1, \dots k\} \setminus j} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{h_i}^{(k)}\right|,
\end{aligned}$$

and $\bar{O}_h^{(k)}$ is given by the basis change $\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\} \rightarrow \left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle \dots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\}$.

Below, we describe in detail the exact steps of the algorithm and then provide a proof or justification for the claims made in each step.

5.4.1 PHASE 1: INITIALIZATION

We are given a Λ-valid transition $g \longrightarrow h$ and the EBRM function $a = h - g$. Since the function is EBRM we know that there exist matrices $H \geq G$ and a vector $|\psi\rangle$ such that $a = \mathrm{Prob}[H, |\psi\rangle] - \mathrm{Prob}[G, |\psi\rangle]$. We also know that the maximum matrix size we need to consider is $n_g + n_h - 1$. We want to know the spectrum of the matrices involved to proceed. In terms of the ellipsoids $H \geq G$ means that the H ellipsoid is inside the G-ellipsoid. We try to expand the H-ellipsoid (which means scaling down the matrix H) until it touches the G-ellipsoid. If we already knew H and G, we could find the spectrum; however what we know is the function $a = h - g$. We use the equivalence between EBRM and valid functions to perform the tightening procedure even without knowing the matrices. We use $a_\gamma = h_\gamma - g$, where $h_\gamma(x) = h(x/\gamma)$ and check if a_γ stays valid as we shrink γ from one to zero. We stop when we see any signature of tightness. Using this a_γ we determine the spectrum of the matrices certifying the EBRM claim.

We start with the tightening procedure until we find some operator monotone labeled by λ for which $l_{\gamma'}(\lambda)$ disappears. This corresponds to the ellipsoids touching, since after applying this operator monotone the ellipsoids must touch along the $|w\rangle$ direction.

Tightening procedure Let $[x_{\min}(\gamma'), x_{\max}(\gamma')]$ be the support domain for $a_{\gamma'}$ with $\gamma' \in (0,1]$ a variable, and $\gamma \in (0,1]$ the largest root of $m(\gamma', x_{\min}(\gamma'), x_{\max}(\gamma'))$. Let $x_{\max} := x_{\max}(\gamma)$ and $x_{\min} := x_{\min}(\gamma)$.

First we must show that there would indeed be a root of m as a function of γ' in the range $(0,1]$. This is a direct consequence of Lemma 73. Second we must show that if we can find the matrices certifying a_γ is EBRM, then we can find the matrices certifying a is EBRM. This follows from the observation that $\gamma X_h \geq O X_g O^T \Rightarrow X_h \geq \gamma X_h \geq O X_g O^T$.

We found a signature of tightness, and we can proceed to find the spectrum of the matrices involved.

Spectral domain for the representation Find the smallest interval $[\chi, \xi]$ such that $l_\gamma(\lambda) \geq 0$ for $\lambda \in (\mathbb{R} \cup \{\infty, -\infty\}) \backslash [\chi, \xi]$. If $\mathrm{supp}(g), \mathrm{supp}(h)$ are not contained in $[\chi, \xi]$ then from all expansions of $[\chi, \xi]$ that contain the aforesaid sets, pick the smallest. Relabel this interval to $[\chi, \xi]$.

The interval so obtained contains the spectrum of the matrices that certify a_γ is EBRM. This is justified by Lemma 74 using the fact that $l^1_\gamma \geq 0$ from the previous step.

We need our matrices to be positive to be able to use the elliptic picture. We therefore shift the spectrum so that the smallest eigenvalue required is one (or any positive number).

Shift Transform

$$a(x) \longrightarrow a'(x') := a(x' + \chi - 1)$$

where instead of 1 any positive constant would do (see Corollary 67). Similarly transform

$$g(x) \longrightarrow g'(x') := g(x' + \chi - 1) \text{ and } h(x) \longrightarrow h'(x') := h(x' + \chi - 1).$$

Relabel a' to be a, g' to be g and h' to be h. We do not deduce h and g from a as its positive and negative part because they might now have common support due to the tightening procedure.

We use Corollary 67 to deduce that if $a(x)$ is EBRM with spectrum in $[\chi, \xi]$ then $a'(x') = a(x' + \chi - 1)$ is EBRM with spectrum in $[1, \xi - \chi + 1]$. We must also show that if we can find the matrices certifying a' is EBRM then we can find the matrices certifying a is EBRM. This is a direct consequence of the fact that $X'_h \geq O X'_g O^T \iff X_h - (\chi - 1)\mathbb{I} \geq O(X_g - (\chi - 1)\mathbb{I})O^T$. The orthogonal matrix, O, remains unchanged.

With the spectrum determined and adjusted to the elliptic picture, we fix everything except the orthogonal matrix by using the COF (up to a permutation).

The matrices For $n := n_g + n_h - 1$ we define $n \times n$ matrices with spectrum in $[\chi, \xi]$ and n-dimensional vectors as

$$X_g^{(n)} = \mathrm{diag}[\chi, \chi, \dots x_{g_1}, x_{g_2} \dots, x_{g_{n_g}}] \text{ and } X_{h_\gamma}^{(n)} = \mathrm{diag}[\gamma x_{h_1}, \gamma x_{h_2}, \dots, \gamma x_{h_{n_h}}, \xi, \xi, \dots],$$
$$\left|v^{(n)}\right\rangle \doteq \left[0, 0 \dots, \sqrt{p_{g_1}}, \sqrt{p_{g_2}}, \dots, \sqrt{p_{g_{n_g}}}\right], \left|w^{(n)}\right\rangle \doteq \left[\sqrt{p_{h_1}}, \sqrt{p_{h_2}}, \dots, \sqrt{p_{h_{n_h}}}, 0, 0 \dots\right],$$

where $g = \sum_{i=1}^{n_g} p_{g_i} [\![x_{g_i}]\!]$ and $h = \sum_{i=1}^{n_h} p_{h_i} [\![x_{h_i}]\!]$. Note that n_g and n_h may be different. We use Lemma 69 to deduce that $g \longrightarrow h$ is a valid transition from the validity of a. Then, we use Lemma 146 to write the diagonal matrices as described above given the valid transition $g \longrightarrow h$, up to a permutation. Our objective is to find a matrix $O^{(n)}$ such that $O^{(n)} \left|v^{(n)}\right\rangle = \left|w^{(n)}\right\rangle$, while satisfying the inequality $X_h^{(n)} \geq O^{(n)} X_g^{(n)} O^{(n)T}$. Finally we employ the description containing the basis and the matrix instance, which can be iteratively reduced to a simpler form.

Bootstrapping the iteration

- Basis: $\left\{\left|t_{h_i}^{(n+1)}\right\rangle\right\}$ where $\left|t_{h_i}^{(n+1)}\right\rangle := |i\rangle$ for $i = 1, 2 \dots n$ where $|i\rangle$ refers to the standard basis in which the matrices and the vectors were originally written.
- Matrix Instance: $\underline{X}^{(n)} = \{X_h^{(n)}, X_g^{(n)}, \left|w^{(n)}\right\rangle, \left|v^{(n)}\right\rangle\}$.

5.4.2 PHASE 2: ITERATION

We take as input the matrices X_g, X_h together with the vectors $|w\rangle, |v\rangle$ and produce the same objects with one less dimension. We also output objects that, once we have iteratively reduced the problem to triviality, can be put together to find the orthogonal matrix O. See Figure 14 for a schematic reference.

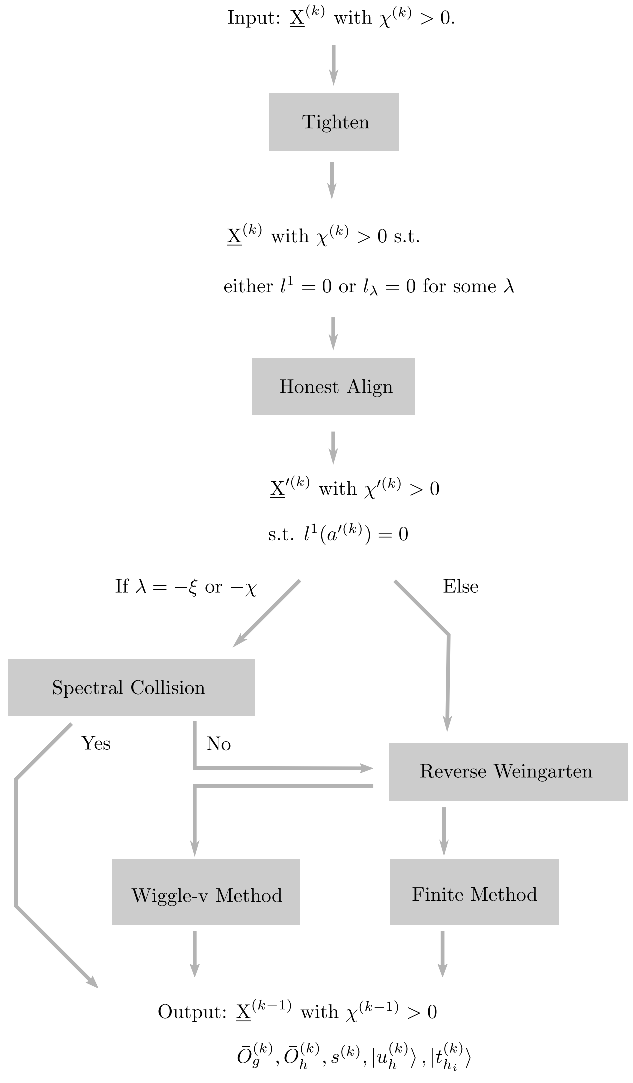

Input: $\underline{X}^{(k)}$ with $\chi^{(k)} > 0$.
Tighten
$\underline{X}^{(k)}$ with $\chi^{(k)} > 0$ s.t.
either $l^1 = 0$ or $l_\lambda = 0$ for some λ
Honest Align
$\underline{X}'^{(k)}$ with $\chi'^{(k)} > 0$
s.t. $l^1(a'^{(k)}) = 0$
If $\lambda = -\xi$ or $-\chi$
Else
Spectral Collision
Yes
No
Reverse Weingarten
Wiggle-v Method
Finite Method
Output: $\underline{X}^{(k-1)}$ with $\chi^{(k-1)} > 0$
$\bar{O}_g^{(k)}, \bar{O}_h^{(k)}, s^{(k)}, |u_h^{(k)}\rangle, |t_{h_i}^{(k)}\rangle$

Figure 14: Overview of the main step, the iteration, of the algorithm (excluding the boundary condition).

- Objective: Find the objects $\left|u_h^{(k)}\right\rangle, \bar{O}_g^{(k)}, \bar{O}_h^{(k)}$ and $s^{(k)}$. These objects together relate $O^{(k)}$ to $O^{(k-1)}$ where $O^{(k)}$ and $O^{(k-1)}$ solve the matrix instances $\underline{X}^{(k)}$ and $\underline{X}^{(k-1)}$, respectively.

- Input: We assume we are given
  - Basis: $\left\{\left|t_{h_i}^{(k+1)}\right\rangle\right\}$
  - Matrix Instance: $\underline{\mathrm{X}}^{(k)} = \left(X_h^{(k)}, X_g^{(k)}, \left|w^{(k)}\right\rangle, \left|v^{(k)}\right\rangle\right)$, with attribute $\chi^{(k)} > 0$
  - Function Instance: $\underline{\mathrm{X}}^{(k)} \longrightarrow \underline{\mathrm{x}}^{(k)} = \left(h^{(k)}, g^{(k)}, a^{(k)}\right)$
- Output:
  - Basis: $\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_i}^{(k)}\right\rangle\right\}$
  - Matrix Instance: $\underline{\mathrm{X}}^{(k-1)} = \left(X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle\right)$ with attribute $\chi^{(k-1)} > 0$
  - Function Instance: $\underline{\mathrm{X}}^{(k-1)} \longrightarrow \underline{\mathrm{x}}^{(k-1)} = \left(h^{(k-1)}, g^{(k-1)}, a^{(k-1)}\right)$
  - Unitary Constructors: Either $\bar{O}_g^{(k)}$ and $\bar{O}_h^{(k)}$ are returned or $\bar{O}^{(k)}$ is returned. If $\bar{O}^{(k)}$ is returned, set $\bar{O}_g^{(k)} := \bar{O}^{(k)}$ and $\bar{O}_h^{(k)} = \mathbb{I}$.
  - Relation: If $s^{(k)}$ is not specified, define $s^{(k)} := 1$.
    If $s^{(k)} = 1$ then use
    $$O^{(k)} := \bar{O}_h^{(k)} \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}_g^{(k)}$$
    else use
    $$O^{(k)} := \left[\bar{O}_h^{(k)} \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}_g^{(k)}\right]^T.$$

    Our task is to solve the matrix instance $\underline{\mathrm{X}}^{(k)}$, i.e. find an orthogonal matrix $O^{(k)}$ such that $X_h^{(k)} \geq O^{(k)} X_g^{(k)} O^{(k)T}$ and $O^{(k)}\left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle$. We assume that the solution exists and show that the solution to the smaller instance, denoted by $\underline{\mathrm{X}}^{(k-1)}$, must also exist. More precisely, we show that $O^{(k)}$ must have the form $O^{(k)} = \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}^{(k)}$ which satisfies the aforesaid constraints, granted that we can find $O^{(k-1)}$ acting on a $k-1$-dimensional Hilbert space orthogonal to $\left|u_h^{(k)}\right\rangle$ and satisfies constraints of the same form in the smaller dimension, viz. $X_h^{(k-1)} \geq O^{(k-1)} X_g^{(k-1)} O^{(k-1)T}$ and $O^{(k-1)}\left|v^{(k-1)}\right\rangle = \left|w^{(k-1)}\right\rangle$. Hence the assumption that $O^{(k)}$ is a solution allows us to deduce that $O^{(k-1)}$ must also be a solution. This allow us to iteratively solve the problem. In certain trivial cases, where a point appears both before and after a transition i.e., $X_g^{(k)}$ and $X_h^{(k)}$ have a common eigenvalue, the solution is of the form
    $$O^{(k)} = \bar{O}_h^{(k)} \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}_g^{(k)}.$$

    Finally, in one of the "infinite" cases denoted by the "wiggle-v method" the solution has the form
    $$O^{(k)} = \left[\left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}^{(k)}\right]^T.$$
- Algorithm:
  If we reach a stage where the vector constraints have disappeared then we can simply stop:
  - **Boundary condition: If** $n_g = 0$ and $n_h = 0$ **then** set $k_0 = k$ and **jump to PHASE 3**.
    We assumed that an $O^{(k)}$ satisfying the necessary constraints exists, which means that there exists an $O^{(k)}$ such that $X_h^{(k)} \geq O^{(k)} X_g^{(k)} O^{(k)T}$ as there is no vector $\left|v^{(k)}\right\rangle, \left|w^{(k)}\right\rangle$ to impose

further constraints. Using Corollary 71 with $H = X_h^{(k)}$ and $G = O^{(k)} X_g^{(k)} O^{(k)T}$ we conclude that $O^{(k)}$ need only be a permutation matrix. Note that this step can never be entered right after the $\underline{\mathrm{X}}^{(n)}$ instance as we start with assuming $n_g, n_h > 0$. Further, since the protocol by construction always returns X_h and X_g in the ascending order the permutation matrix is $\mathbb{I}$.

Finally, we deal with the case, where we need to expand the inner H-ellipsoid (which corresponds to shrinking the H matrix) until it touches the outer G-ellipsoid working with the function a.

- **Tighten**: Define $X_{h_{\gamma'}}^{(k)} := \gamma' X^{(k)}$ where $\gamma' \in (0, 1]$ is a variable. Let γ be the largest root of $m(\gamma', \chi_{\gamma'}^{(k)}, \xi_{\gamma'}^{(k)})$ for $a^{(k)}$ where $\chi_{\gamma'}^{(k)}, \xi_{\gamma'}^{(k)}$ are such that $[\chi_{\gamma'}^{(k)}, \xi_{\gamma'}^{(k)}]$ is the smallest interval containing $\mathrm{spec}[X_{h_{\gamma'}}^{(k)} \oplus X_g^{(k)}]$. Relabel $X_{h_\gamma}^{(k)}$ to $X_h^{(k)}$, $\chi_\gamma^{(k)}$ to $\chi^{(k)}$ and $\xi_\gamma^{(k)}$ to $\xi^{(k)}$ and $a_\gamma^{(k)}$ to $a^{(k)}$, $h_\gamma^{(k)}$ to $h^{(k)}$, $l_\gamma^{(k)}$ to $l^{(k)}$ (for ease of notation). Update $x_{\min}$ and $x_{\max}$ to be such that $\mathrm{supp}(a^{(k)}) \in [x_{\min}^{(k)}, x_{\max}^{(k)}]$ is the smallest such interval. Define $s^{(k)} := 1$.
  Our task is to show that m as a function of γ' has a root. Unlike the first tightening procedure this time we know the spectrum of the matrices involved. Since we know that the matrix instance has a solution we know that $l_{\gamma'=1}(\lambda) \geq 0$ and $l_{\gamma'=1}^1 \geq 0$ for $\lambda \in (\mathbb{R} \cup \{\infty, -\infty\}) \backslash [\chi_{\gamma'=1}^{(k)}, \xi_{\gamma'=1}^{(k)}]$ using Lemma 68. We also know that $\chi_{\gamma'}^{(k)} > 0$ which means that $a^{(k)}$ is a valid function (as deduced by the function instance of $\underline{\mathrm{X}}^{(k)}$). Thus we can show that $m(\gamma')$ has a root in the required range by using the same reasoning as in Lemma 73.
  The tightening procedure guarantees that we find a λ corresponding to an operator monotone. After applying this function, the ellipsoids—which we do not even know completely yet—must touch along the $|w\rangle$ direction. This is the key to reducing the problem to a smaller instance of itself. Recall the picture with the H ellipsoid contained inside the G ellipsoid. If, in addition, we know that they touch at a given point then it must be so that the inner ellipsoid is more curved than the outer ellipsoid. When expressed algebraically, this condition essentially becomes the requirement that an ellipsoid $H^{(k-1)}$, encoding the curvature of the ellipsoid $H^{(k)}$ at the point of contact, must be contained inside the corresponding $G^{(k-1)}$ ellipsoid which encodes the curvature of the $G^{(k)}$ ellipsoid. The vector condition also reduces similarly. Subtleties arise when λ happens to have boundary values in its allowed range as this yields infinities.
- **Honest align**: **If** $l^{1(k)} = 0$ **then** define $\eta = -\chi^{(k)} + 1$

  $$X_h'^{(k)} := X_h^{(k)} + \eta \text{ and } X_g'^{(k)} := X_g + \eta.$$

  Else: Pick a root λ of $l^{(k)}(\lambda')$ in the domain $\mathbb{R} \backslash (-\xi^{(k)}, -\chi^{(k)})$. In the following two cases we consider the function f_λ on $[\chi^{(k)}, \xi^{(k)}]$.
  - If $\lambda \neq -\chi^{(k)}$, then let $\eta = -f_\lambda(\chi^{(k)}) + 1$ where any positive constant can be chosen instead of 1. Define

    $$X_h'^{(k)} := f_\lambda(X_h^{(k)}) + \eta \mathrm{X}_g'^{(k)} := f_\lambda(X_g^{(k)}) + \eta.$$

  - If $\lambda = -\chi^{(k)}$, then update $s^{(k)} = -1$. Let $\eta = -f_\lambda(\xi^{(k)}) - 1$ where any positive constant could be chosen instead of 1. Define

    $$X_h'^{(k)} := X_g''^{(k)} \text{ and } X_g'^{(k)} := X_h''^{(k)},$$

    $$\text{where } X_h''^{(k)} := -f_\lambda(X_h^{(k)}) - \eta \text{ and } X_g''^{(k)} := -f_\lambda(X_g^{(k)}) - \eta,$$

    and make the replacements

    $$\left|v^{(k)}\right\rangle \to \left|w^{(k)}\right\rangle \text{ and } \left|w^{(k)}\right\rangle \to \left|v^{(k)}\right\rangle.$$

If $\lambda = -\chi^{(k)}$ or $-\xi^{(k)}$ it means that at least one of the matrices (among $X_g^{(k)}$ and $X_h^{(k)}$ under f_λ) diverges. We must remove eigenvalues common to both matrices as isolating the divergence makes it easier to handle.

- **Remove spectral collision: If** $\lambda = -\chi^{(k)}$ or $\lambda = -\xi^{(k)}$ **then**

  If it so happens that the coordinate and its corresponding probability are the same we must leave the associated vector unchanged (up to a relabeling). The following simply formalizes this procedure and encodes the remaining non-trivial part into a problem of one less dimension.

  1. **Idle point: If** for some j', j, we have $q_{g_{j'}}^{(k)} = q_{h_j}^{(k)}$ and $y_{g_{j'}}^{(k)} = y_{h_j}^{(k)}$ **then** the solution is given by

$$\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle, \left|t_{h_2}^{(k)}\right\rangle, \dots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\} \overset{\text{component-wise}}{:=}$$
$$\left\{\left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\},$$

$$\bar{O}^{(k)} := \sum_{i=1}^{k} \left|a_i\right\rangle \left\langle t_{h_i}^{(k+1)}\right|,$$

$$\text{where } \{\left|a_1\right\rangle, \left|a_2\right\rangle \dots \left|a_k\right\rangle\} \overset{\text{component-wise}}{:=}$$

$$\begin{cases} \left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'}}^{(k+1)}\right\rangle, \left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \dots, \left|t_{h_{j'-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'+1}}^{(k+1)}\right\rangle \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\} & j < j' \\ \left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_{j'-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'+1}}^{(k+1)}\right\rangle \dots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'}}^{(k+1)}\right\rangle, \left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\} & j > j' \\ \left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\} & j = j', \end{cases}$$

$$X_h^{(k-1)} := \sum_{i \neq j} y_{h_i}^{(k)} \left|t_{h_i}^{(k+1)}\right\rangle \left\langle t_{h_i}^{(k+1)}\right|, X_g^{(k-1)} := \bar{O}^{(k)} X_g^{(k)} \bar{O}^{(k)T} - y_{h_j} \left|t_{h_j}^{(k+1)}\right\rangle \left\langle t_{h_j}^{(k+1)}\right|,$$

$$\left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\left|w^{(k)}\right\rangle - \sqrt{p_{h_j}} \left|t_{h_j}^{(k+1)}\right\rangle\right], \left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\bar{O}^{(k)} \left|v^{(k)}\right\rangle - \sqrt{p_{h_j}} \left|t_{h_j}^{(k+1)}\right\rangle\right].$$

  This specifies $\underline{\mathrm{X}}^{(k-1)} := \{X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle\}$.

  Jump to **End**.

  We want to find an $O^{(k)}$ such that $X_h^{(k)} \geq O^{(k)} X_g^{(k)} O^{(k)T}$ and $O^{(k)}\left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle$. We do this in two stages. First, we re-arrange the entries of $X_g^{(k)}$ as $X_g'^{(k)} := O_p^{(k)} X_g^{(k)} O_p^{(k)T}$ and define $\left|v_p^{(k)}\right\rangle := O_p^{(k)}\left|v\right\rangle$ for an $O_p^{(k)}$ to be specified later. The re-arrangement is such that $x_{g_{j'}}$ sits at the j,j location while the rest of the elements of $X_g'^{(k)}$ are arranged in increasing order. Second, we solve our initial problem under the assumption that $j = j'$. The non-trivial part here is showing that we can consider $O^{(k)}$ to be of the form $\left(\left|j\right\rangle\left\langle j\right| + O^{(k-1)}\right)\bar{O}^{(k)}$ without loss of generality.

  Let us start with the first step. We denote the orthogonal matrix $O = \sum_i \left|b_i\right\rangle\left\langle a_i\right|$ by $\{\left|a_1\right\rangle, \left|a_2\right\rangle, \dots \left|a_k\right\rangle\} \to \{\left|b_1\right\rangle, \left|b_2\right\rangle, \dots \left|b_k\right\rangle\}$ where $\{\left|b_i\right\rangle\}$ and $\{\left|a_i\right\rangle\}$ are two orthonormal basis. With this notation and for $j < j'$, we define $O_p^{(k)}$ as

$$\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\} \to$$
$$\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'}}^{(k+1)}\right\rangle, \left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \dots, \left|t_{h_{j'-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'+1}}^{(k+1)}\right\rangle \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\},$$

for $j' < j$ we define it as

$$\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\} \to$$
$$\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_{j'-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'+1}}^{(k+1)}\right\rangle \dots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'}}^{(k+1)}\right\rangle, \left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\}$$

and for $j' = j$ we set $O_p^{(k)} = \mathbb{I}^{(k)}$. For the second step, we solve under the assumption that $j' = j$. We have $X_g^{'(k)} = \text{diag}\{x'_{g_1}, x'_{g_2} \dots x'_{g_k}\}$ and $X_h^{(k)} = \text{diag}\{x_{h_1}, x_{h_2} \dots x_{h_k}\}$ which are such that $x_{h_j} = x'_{g_j}$; $\left|v'^{(k)}\right\rangle \doteq (\sqrt{q'_{g_1}}, \sqrt{q'_{g_2}}, \dots \sqrt{q'_{g_k}})^T$, $\left|w^{(k)}\right\rangle \doteq (\sqrt{q_{h_1}}, \sqrt{q_{h_2}}, \dots \sqrt{q_{h_k}})^T$ are such that $q_{h_j} = q'_{g_j}$. Let us define the matrix instance to be $\underline{X}'^{(k)} = \{X_h^{(k)}, X_g'^{(k)}, \left|v'^{(k)}\right\rangle, \left|w^{(k)}\right\rangle\}$. We have to find an $O'^{(k)}$ such that $X_h^{(k)} \geq O'^{(k)} X_g'^{(k)} O'^{(k)T}$ and $O'^{(k)} \left|v'^{(k)}\right\rangle = |w\rangle$.
Let $\underline{X}'^{(k-1)} = \left\{X_h^{(k-1)}, X_g'^{(k-1)}, \left|v'^{(k-1)}\right\rangle, \left|w^{(k-1)}\right\rangle\right\}$ be the matrix instance obtained after removing the j^{th} entry from the vectors, i.e., $\left|v'^{(k-1)}\right\rangle := \sum_{i\neq j} \sqrt{q'_{g_i}} \left|t_{h_i}^{(k+1)}\right\rangle$, $\left|w^{(k-1)}\right\rangle := \sum_{i\neq j} \sqrt{q_{h_i}} \left|t_{h_i}^{(k+1)}\right\rangle$ and similarly define
$X_g'^{(k-1)} = \text{diag}\{x'_{g_1}, x'_{g_2} \dots x'_{g_{j-1}}, x'_{g_{j+1}}, \dots x_{g_k}\}$ and
$X_h^{(k-1)} = \text{diag}\left\{x_{h_1}, x_{h_2} \dots x_{h_{j-1}}, x_{h_{j+1}}, \dots x_{h_k}\right\}$. Note that $a^{(k)} = a^{(k-1)}$ as the j^{th} point gets canceled. This means that if there is an $O'^{(k)}$ satisfying the aforementioned constraints $a^{(k)}$ is EBRM on the spectral domain of $\underline{X}^{(k)}$. Since $a^{(k)} = a^{(k-1)}$ we know that $a^{(k-1)}$ is also EBRM on the same domain. From Lemma 69 (we justify that k is large enough separately) we conclude that there must also exist an $O'^{(k-1)}$ which satisfies $X_h^{(k-1)} \geq O'^{(k-1)} X_g'^{(k-1)} O'^{(k-1)T}$ and $O'^{(k-1)} \left|v'^{(k-1)}\right\rangle = \left|w^{(k)}\right\rangle$.
With all this in place we can claim that without loss of generality we can write $O'^{(k)} = \left|t_{h_j}\right\rangle \left\langle t_{h_j}\right| + O'^{(k-1)}$ because if we can find some other $\tilde{O}'^{(k)}$ which satisfies the required constraints then there exists an $O'^{(k-1)}$ which satisfies the corresponding constraints in the smaller dimension and that means we can show $O'^{(k)}$ also satisfies the required constraints

$$X_h^{(k)} = x_{h_j} \left|t_{h_j}^{(k+1)}\right\rangle \left\langle t_{h_j}^{(k+1)}\right| + X_h^{(k-1)} \geq x_{g_j} \left|t_{h_j}^{(k+1)}\right\rangle \left\langle t_{h_j}^{(k+1)}\right| + O'^{(k-1)} X_g'^{(k-1)} O'^{(k-1)}$$
$$0 \left(\left|t_{h_j}^{(k+1)}\right\rangle \left\langle t_{h_j}^{(k+1)}\right| + O'^{(k-1)}\right) X_g'^{(k)} \left(\left|t_{h_j}^{(k+1)}\right\rangle \left\langle t_{h_j}^{(k+1)}\right| + O'^{(k-1)}\right)^T = O'^{(k)} X_g'^{(k)} O'^{(k)T},$$

along with

$$O'^{(k)} \left|v'^{(k)}\right\rangle = \sqrt{q'_{g_j}} \left|t_{h_j}^{(k+1)}\right\rangle + O'^{(k-1)} \left|v'^{(k-1)}\right\rangle = \sqrt{q'_{g_j}} \left|t_{h_j}^{(k+1)}\right\rangle + \left|w^{(k-1)}\right\rangle = \left|w^{(k-1)}\right\rangle.$$

It remains to combine the two steps to produce the matrix $\bar{O}^{(k)}$, the vectors $\left\{\left|n_h^{(k)}\right\rangle, \left\{\left|t_{h_i}^{(k)}\right\rangle\right\}\right\}$, along with $\underline{X}^{(k-1)}$. We use $X_g'^{(k)} = O_p^{(k)} X_g^{(k)} O_p^{(k)T}$ from the first step and substitute it in the inequality which we showed would hold, i.e.

$$X_h^{(k)} \geq O'^{(k)} X_g'^{(k)} O'^{(k)T} = O'^{(k)} O_p^{(k)} X_g O_p^{(k)T} O'^{(k)T}$$

and using $O_p^{(k)} \left|v^{(k)}\right\rangle = \left|v'^{(k)}\right\rangle$ we have

$$O'^{(k)} \left|v'^{(k)}\right\rangle = O'^{(k)} O_p^{(k)} \left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle.$$

Comparing to the form $X_h^{(k)} \geq O^{(k)} X_g^{(k)} O^{(k)T}$,
$O^{(k)} \left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle$ for $O^{(k)} = \left(\left|n_h^{(k)}\right\rangle \left\langle n_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}^{(k)}$, we get $\bar{O}^{(k)} = O_p^{(k)}$, $\left|n_h^{(k)}\right\rangle =$

$\left|t^{(k+1)}_{h_j}\right\rangle$ and $O^{(k-1)} = O'^{(k-1)}$. Note that this $O^{(k)}$ is consistent with comparing the equality with the form $O^{(k)}\left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle$. The basis for the $(k-1)$-dimensional problem, is the same as before except for the fact that we removed $\left|t^{(k+1)}_{h_j}\right\rangle$. We define
$\left\{\left|t^{(k)}_{h_1}\right\rangle, \left|t^{(k)}_{h_2}\right\rangle \dots \left|t^{(k)}_{h_{k-1}}\right\rangle\right\} = \left\{t^{(k+1)}_{h_1}, t^{(k+1)}_{h_2} \dots t^{(k+1)}_{h_{j-1}}, t^{(k+1)}_{h_{j+1}}, \dots t^{(k+1)}_{h_k}\right\}$.
Identifying $\underline{\mathrm{X}}^{(k-1)} = \left\{X^{(k-1)}_h, X^{(k-1)}_g, \left|v^{(k-1)}\right\rangle, \left|w^{(k-1)}\right\rangle\right\}$
with $\underline{\mathrm{X}}'^{(k-1)} = \left\{X^{(k-1)}_h, X'^{(k-1)}_g, \left|v'^{(k-1)}\right\rangle, \left|w^{(k-1)}\right\rangle\right\}$ we complete the argument since $O^{(k-1)}$ was already identified with $O'^{(k-1)}$.

2. **Final Extra**: **If** for some j, j' we have $q^{(k)}_{g_{j'}} > q^{(k)}_{h_j}$ and $y^{(k)}_{g_{j'}} = y^{(k)}_{h_j}$ **then** the solution is $\underline{\mathrm{X}}^{(k-1)} := (X^{(k-1)}_h, X^{(k-1)}_g, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle)$ where $X^{(k-1)}_h = \sum_{i=1}^{k-1} y^{(k-1)}_{h_i}\left|t^{(k)}_{h_i}\right\rangle\left\langle t^{(k)}_{h_i}\right|$,
$X^{(k-1)}_g = \sum_{i=1}^{k-1} y^{(k-1)}_{g_i}\left|t^{(k)}_{h_i}\right\rangle\left\langle t^{(k)}_{h_i}\right|$,
$\left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1}\sqrt{q^{(k-1)}_{g_i}}\left|t^{(k)}_{h_i}\right\rangle\right]$, and
$\left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1}\sqrt{q^{(k-1)}_{h_i}}\left|t^{(k)}_{h_i}\right\rangle\right]$, where the coordinates and weights are given by

$$\begin{aligned}
\left\{q^{(k-1)}_{h_1}, \dots q^{(k-1)}_{h_{k-1}}\right\} &\overset{\text{component-wise}}{=} \left\{q^{(k)}_{h_1}, q^{(k)}_{h_2} \dots, q^{(k)}_{h_{j-1}}, q^{(k)}_{h_{j+1}}, \dots q^{(k)}_{h_k}\right\}\\
\left\{q^{(k-1)}_{g_1}, \dots q^{(k-1)}_{g_{k-1}}\right\} &\overset{\text{component-wise}}{=} \left\{q^{(k)}_{g_2} \dots, q^{(k)}_{g_{j'-1}}, q^{(k)}_{g_{j'}} - q^{(k)}_{h_j}, q^{(k)}_{g_{j'+1}}, q^{(k)}_{g_{j'+2}} \dots q^{(k)}_{g_k}\right\}\\
\left\{y^{(k-1)}_{g_1}, \dots y^{(k-1)}_{g_{k-1}}\right\} &\overset{\text{component-wise}}{=} \left\{y^{(k)}_{g_2}, \dots y^{(k)}_{g_k}\right\}\\
\left\{y^{(k-1)}_{h_1}, \dots y^{(k-1)}_{h_{k-1}}\right\} &\overset{\text{component-wise}}{=} \left\{y^{(k)}_{h_1}, \dots y^{(k)}_{h_{j-1}}, y^{(k)}_{h_{j+1}} \dots, y^{(k)}_{h_k}\right\},
\end{aligned}$$

the basis is given by

$$\begin{aligned}
&\left\{\left|u^{(k)}_h\right\rangle, \left|t^{(k)}_{h_1}\right\rangle \dots \left|t^{(k)}_{h_{k-1}}\right\rangle\right\} \overset{\text{component-wise}}{=}\\
&\left\{\left|t^{(k+1)}_{h_j}\right\rangle, \left|t^{(k+1)}_{h_1}\right\rangle, \left|t^{(k+1)}_{h_2}\right\rangle, \dots \left|t^{(k+1)}_{h_{j-1}}\right\rangle, \left|t^{(k+1)}_{h_{j+1}}\right\rangle, \left|t^{(k+1)}_{h_{j+2}}\right\rangle \dots \left|t^{(k+1)}_{h_k}\right\rangle\right\}.
\end{aligned}$$

The orthogonal matrices are given by $\bar{O}^{(k)}_h := \sum \left|t^{(k+1)}_{h_i}\right\rangle\left\langle a_i\right|$ where

$$\{|a_1\rangle, \dots |a_k\rangle\} \to \left\{\left|u^{(k)}_h\right\rangle, \left|t^{(k)}_{h_1}\right\rangle \dots \left|t^{(k)}_{h_{k-1}}\right\rangle\right\},$$

and $\bar{O}^{(k)}_g := \tilde{O}^{(k)}\bar{O}^{(k)}_h$, where

$$\begin{aligned}
\tilde{O}^{(k)} :=& \mathcal{N}\left[\sqrt{q^{(k)}_{h_j}}\left|u^{(k)}_h\right\rangle + \sqrt{q^{(k)}_{g_{j'}} - q^{(k)}_{h_j}}\left|t^{(k)}_{h_{j'}}\right\rangle\right]\mathcal{N}\left[\sqrt{q^{(k)}_{g_1}}\left\langle u^{(k)}_h\right| + \sqrt{q^{(k)}_{g_{j'}}}\left\langle t^{(k)}_{h_{j'}}\right|\right]\\
&+ \mathcal{N}\left[\sqrt{q^{(k)}_{g_{j'}} - q^{(k)}_{h_j}}\left|u^{(k)}_h\right\rangle - \sqrt{q^{(k)}_{h_j}}\left|t^{(k)}_{h_{j'}}\right\rangle\right]\mathcal{N}\left[\sqrt{q^{(k)}_{g_{j'}}}\left\langle u^{(k)}_h\right| - \sqrt{q^{(k)}_{g_1}}\left\langle t^{(k)}_{h_{j'}}\right|\right]\\
&+ \sum_{i\in\{1,\dots k\}\setminus j'}\left|t^{(k)}_{h_i}\right\rangle\left\langle t^{(k)}_{h_i}\right|.
\end{aligned}$$

Jump to **End**.
We are given $\underline{\mathrm{X}}^{(k)} = (X^{(k)}_h, X^{(k)}_g, \left|w^{(k)}\right\rangle, \left|v^{(k)}\right\rangle)$,
where $X^{(k)}_h = \sum_{i=1}^{k} y^{(k)}_{h_i}\left|t^{(k+1)}_{h_i}\right\rangle\left\langle t^{(k+1)}_{h_i}\right|$,
$X^{(k)}_g = \sum_{i=1}^{k} y^{(k)}_{g_i}\left|t^{(k+1)}_{h_i}\right\rangle\left\langle t^{(k+1)}_{h_i}\right|$, $\left|v^{(k)}\right\rangle = \sum_{i=1}^{k} q^{(k)}_{g_i}\left|t^{(k+1)}_{h_i}\right\rangle$, and $\left|w^{(k)}\right\rangle = \sum_{i=1}^{k} q^{(k)}_{h_i}\left|t^{(k+1)}_{h_i}\right\rangle$

with the corresponding function instance being $\underline{\mathrm{x}}^{(k)} = (h^{(k)}, g^{(k)}, a^{(k)})$, where $a^{(k)} = \sum_{i\in\{1,\dots k\}\setminus j} q_{h_i}^{(k)}[y_{h_i}] - \sum_{i\in\{1,\dots k\}\setminus j'} q_{g_i}^{(k)}[y_{g_i}] - (q_{g_{j'}}^{(k)} - q_{h_j}^{(k)})[y_{h_j}]$. Since we assume that $\underline{\mathrm{X}}^{(k)}$ has a solution it follows that $a^{(k)}$ is $[\chi,\xi]$-valid. Thus the transition $g^{(k-1)} := a_-^{(k)} \to a_+^{(k)} =: h^{(k-1)}$ is also $[\chi,\xi]$-valid where $g^{(k-1)}$ comprises $n_g^{(k-1)} = n_g^{(k)}$ points and $h^{(k-1)}$ comprises $n_h^{(k-1)} = n_h^{(k)} - 1$ points (using the attributes corresponding to the function instance $(h^{(k-1)}, g^{(k-1)}, h^{(k-1)} - g^{(k-1)})$. We denote this by $g = \sum_{i=1}^{n_g} p_{g_i}[x_{g_i}]$ and $h = \sum_{i=1}^{n_h} p_{h_i}[x_{h_i}]$). Since $k = n_g^{(k)} + n_h^{(k)} - 1$ the aforesaid relation yields $k-1 = n_g^{(k-1)} + n_h^{(k-1)} - 1$. We conclude that the matrix instance
$\underline{\mathrm{X}}^{(k-1)} := (X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle)$,
where $X_h^{(k-1)} = \sum_{i=1}^{k-1} y_{h_i}^{(k-1)} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{h_i}^{(k)}\right|$,
$X_g^{(k-1)} = \sum_{i=1}^{k-1} y_{g_i}^{(k-1)} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{h_i}^{(k)}\right|$,
$\left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1} \sqrt{q_{g_i}^{(k-1)}} \left|t_{h_i}^{(k)}\right\rangle\right]$
and $\left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1} \sqrt{q_{h_i}^{(k-1)}} \left|t_{h_i}^{(k)}\right\rangle\right]$, has a solution for

$$\left\{q_{h_1}^{(k-1)}, \dots q_{h_{k-1}}^{(k-1)}\right\} \overset{\text{component-wise}}{=} \left\{q_{h_1}^{(k)}, q_{h_2}^{(k)} \dots, q_{h_{j-1}}^{(k)}, q_{h_{j+1}}^{(k)}, \dots q_{h_k}^{(k)}\right\}$$
$$\left\{q_{g_1}^{(k-1)}, \dots q_{g_{k-1}}^{(k-1)}\right\} \overset{\text{component-wise}}{=} \left\{q_{g_2}^{(k)} \dots, q_{g_{j'-1}}^{(k)}, q_{g_{j'}}^{(k)} - q_{h_j}^{(k)}, q_{g_{j'+1}}^{(k)}, q_{g_{j'+2}}^{(k)} \dots q_{g_k}^{(k)}\right\}$$
$$\left\{y_{g_1}^{(k-1)}, \dots y_{g_{k-1}}^{(k-1)}\right\} \overset{\text{component-wise}}{=} \left\{y_{g_2}^{(k)}, \dots y_{g_k}^{(k)}\right\}$$
$$\left\{y_{h_1}^{(k-1)}, \dots y_{h_{k-1}}^{(k-1)}\right\} \overset{\text{component-wise}}{=} \left\{y_{h_1}^{(k)}, \dots y_{h_{j-1}}^{(k)}, y_{h_{j+1}}^{(k)} \dots, y_{h_k}^{(k)}\right\},$$

as the corresponding function instance $\underline{\mathrm{x}}^{(k-1)}$ is indeed given by $(h^{(k-1)}, g^{(k-1)}, a^{(k-1)} = a^{(k)})$. Here $\{\left|t_{h_i}^{(k)}\right\rangle\}$ constitute an orthonormal basis which we will soon relate to $\left|t_{h_i}^{(k+1)}\right\rangle$. We used $q_{g_1}^{(k)} = 0$ as $y_{g_1}^{(k)} = \chi$. To verify this note that $k-1 > n_g^{(k-1)}$, which means that many q_{g_i} are zero; by convention we write the smallest eigenvalue, χ first to increase the matrix size so the first $i = 1, 2 \dots \left(k - 1 - n_g^{(k-1)}\right)$ q_is are zero. This means that there must exist an $O^{(k-1)}$ which solves $\underline{\mathrm{X}}^{(k-1)}$.

Let us see carefully the following basis change. Note that $X'_h \geq O' X'_g O'^T$ with $O' \left|v'\right\rangle = \left|w'\right\rangle$ is equivalent to $X_h \geq O X_g O^T$ with $O\left|v\right\rangle = \left|w\right\rangle$ where $O = \bar{O}_h^T O' \bar{O}_g$, $\bar{O}_g\left|v\right\rangle = \left|v'\right\rangle$, $\bar{O}_h\left|w\right\rangle = \left|w'\right\rangle$, $\bar{O}_h X_h \bar{O}_h^T = X'_h$, $\bar{O} X_g \bar{O}_g^T = X'_g$ which is easy to see by a simple substitution. We first expand the matrix $\underline{\mathrm{X}}^{(k-1)}$ to k dimensions as follows. We had $X_h^{(k-1)} \geq O^{(k-1)} X_g^{(k-1)} O^{(k-1)T}$ with $O^{(k-1)}\left|v^{(k-1)}\right\rangle = \left|w^{(k-1)}\right\rangle$ which we expand as

$$\underbrace{y_{h_j}^{(k)} \left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + X_h^{(k-1)}}_{:=X_h'^{(k)}} \geq$$
$$\underbrace{\left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right)}_{:=O'^{(k)}} \underbrace{\left(y_{h_j}^{(k)} \left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + X_g^{(k-1)}\right)}_{:=X_g'^{(k)}} \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right)^T$$

with $\left|v'^{(k)}\right\rangle = \mathcal{N}\left[\sqrt{q_{h_j}^{(k)}} \left|u_h^{(k)}\right\rangle + \left|v^{(k-1)}\right\rangle\right]$
and $\left|w'^{(k)}\right\rangle = \mathcal{N}\left[\sqrt{q_{h_j}^{(k)}} \left|u_h^{(k)}\right\rangle + \left|w^{(k-1)}\right\rangle\right]$.

The matrix instance $\underline{\mathrm{X}}'^{(k)} := (X_h'^{(k)}, X_g'^{(k)}, \left|v'^{(k)}\right\rangle, \left|w'^{(k)}\right\rangle)$ yields $\underline{\mathrm{x}}'^{(k)} = \underline{\mathrm{x}}^{(k)}$. We can now use the equivalence we pointed out above to establish a relation between $X_h^{(k)} \geq$

$O^{(k)}X_g^{(k)}O^{(k)T}$ and $X_h'^{(k)} \geq O'^{(k)}X_g'^{(k)}O'^{(k)T}$ by finding $\bar{O}_g$ and $\bar{O}_h$. We define, somewhat arbitrarily,

$$\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle \dots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\} \overset{\text{component-wise}}{=}$$
$$\left\{\left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \left|t_{h_{j+2}}^{(k+1)}\right\rangle \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\}.$$

We require $\bar{O}_h^{(k)} \left|w^{(k)}\right\rangle = \left|w'^{(k)}\right\rangle$. This is a permutation matrix given by $\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\} \to \left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle \dots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\}$, and this yields $\bar{O}_h^{(k)T}X_h'^{(k)}\bar{O}_h^{(k)} = X_h^{(k)}$. It remains to find $\bar{O}_g^{(k)}$ which we require to satisfy $\bar{O}_g^{(k)} \left|v^{(k)}\right\rangle = \left|v'^{(k)}\right\rangle$. First, observe that $\bar{O}_h^{(k)} \left|v^{(k)}\right\rangle = \sqrt{q_{g_1}^{(k)}} \left|u_h^{(k)}\right\rangle + \sum_{i=2}^k \sqrt{q_{g_i}^{(k)}} \left|t_{h_{i-1}}^{(k)}\right\rangle$. We now apply

$$\begin{aligned}\tilde{O}^{(k)} := \mathcal{N}\left[\sqrt{q_{h_j}^{(k)}}\left|u_h^{(k)}\right\rangle + \sqrt{q_{g_{j'}}^{(k)} - q_{h_j}^{(k)}}\left|t_{h_{j'}}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{g_1}^{(k)}}\left\langle u_h^{(k)}\right| + \sqrt{q_{g_{j'}}^{(k)}}\left\langle t_{h_{j'}}^{(k)}\right|\right] \\ + \mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)} - q_{h_j}^{(k)}}\left|u_h^{(k)}\right\rangle - \sqrt{q_{h_j}^{(k)}}\left|t_{h_{j'}}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}}\left\langle u_h^{(k)}\right| - \sqrt{q_{g_1}^{(k)}}\left\langle t_{h_{j'}}^{(k)}\right|\right] \\ + \sum_{i\in\{1,\dots k\}\setminus j'}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|\end{aligned}$$

to get $\bar{O}_g^{(k)} \left|v^{(k)}\right\rangle = \left|v'^{(k)}\right\rangle$, where we defined $\bar{O}_g^{(k)} := \tilde{O}^{(k)}\bar{O}_h^{(k)}$. Using $y_{h_j}^{(k)} = y_{g_{j'}}^{(k)}$ we can also see that $\bar{O}_g^{(k)T}X_g'^{(k)}\bar{O}_g^{(k)}$ is essentially $X_g^{(k)}$ with $\chi^{(k)}$ at $\left|t_{h_1}^{(k+1)}\right\rangle$ replaced by $y_{g_{j'}} (= y_{h_j})$. One can conclude therefore that $X_g'^{(k)} \geq \bar{O}_g^{(k)}X_g^{(k)}\bar{O}_g^{(k)T}$. Substituting we get

$$\begin{aligned} X_h'^{(k)} &\geq O'^{(k)}X_g'^{(k)}O'^{(k)T} \geq O'^{(k)}\bar{O}_g^{(k)}X_g^{(k)}\bar{O}_g^{(k)T}O'^{(k)T} \\ \iff \bar{O}_h^{(k)T}X_h'^{(k)}\bar{O}_h^{(k)} &\geq \underbrace{\bar{O}_h^{(k)T}O'^{(k)}\bar{O}_g^{(k)}}_{:=O^{(k)}}X_g^{(k)}\bar{O}_g^{(k)T}O'^{(k)T}\bar{O}_h^{(k)} \\ \iff X_h^{(k)} &\geq O^{(k)}X_g^{(k)}O^{(k)T}\end{aligned}$$

and similarly

$$\begin{aligned} O'^{(k)}\left|v'^{(k)}\right\rangle = \left|w'^{(k)}\right\rangle &\iff O'^{(k)}\bar{O}_g^{(k)}\left|v^{(k)}\right\rangle = \bar{O}_h^{(k)}\left|w^{(k)}\right\rangle \\ &\iff O^{(k)}\left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle,\end{aligned}$$

concluding the proof.

3. **Initial Extra**: **If** for some j, j' we have $q_{g_{j'}}^{(k)} < q_{h_j}^{(k)}$ and $y_{g_{j'}}^{(k)} = y_{h_j}^{(k)}$ **then** the solution is $\underline{X}^{(k-1)} := (X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle)$, where $X_h^{(k-1)} = \sum_{i=1}^{k-1} y_{h_i}^{(k-1)}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|$,
$X_g^{(k-1)} = \sum_{i=1}^{k-1} y_{g_i}^{(k-1)}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|$,
$\left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1}\sqrt{q_{g_i}^{(k-1)}}\left|t_{h_i}^{(k)}\right\rangle\right]$,

and $\left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1}\sqrt{q_{h_i}^{(k-1)}}\left|t_{h_i}^{(k)}\right\rangle\right]$, with the coordinates and weights given by

$$\begin{aligned}
&\left\{q_{h_1}^{(k-1)},\dots q_{h_{k-1}}^{(k-1)}\right\} \overset{\text{component-wise}}{=} \left\{q_{h_1}^{(k)}\dots, q_{h_{j-1}}^{(k)}, q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}, q_{h_{j+1}}^{(k)}, q_{h_{j+2}}^{(k)}\dots q_{h_{k-1}}^{(k)}\right\}\\
&\left\{q_{g_1}^{(k-1)},\dots q_{g_{k-1}}^{(k-1)}\right\} \overset{\text{component-wise}}{=} \left\{q_{g_1}^{(k)}, q_{g_2}^{(k)}\dots, q_{g_{j'-1}}^{(k)}, q_{g_{j'+1}}^{(k)},\dots q_{g_k}^{(k)}\right\}\\
&\left\{y_{g_1}^{(k-1)},\dots y_{g_{k-1}}^{(k-1)}\right\} \overset{\text{component-wise}}{=} \left\{y_{g_1}^{(k)},\dots y_{g_{j'-1}}^{(k)}, y_{g_{j'+1}}^{(k)}\dots, y_{g_k}^{(k)}\right\}\\
&\left\{y_{h_1}^{(k-1)},\dots y_{h_{k-1}}^{(k-1)}\right\} \overset{\text{component-wise}}{=} \left\{y_{h_1}^{(k)},\dots y_{h_{k-1}}^{(k)}\right\},
\end{aligned}$$

and the basis is given by

$$\begin{aligned}
&\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle\dots\left|t_{h_{k-1}}^{(k)}\right\rangle\right\} \overset{\text{component-wise}}{=}\\
&\left\{\left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle,\dots\left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \left|t_{h_{j+2}}^{(k+1)}\right\rangle\dots\left|t_{h_k}^{(k+1)}\right\rangle\right\}.
\end{aligned}$$

The orthogonal matrices are given by $\bar{O}_h^{(k)} := \tilde{O}^{(k)}\sum|a_i\rangle\left\langle t_{h_i}^{(k+1)}\right|$ where

$$\{|a_1\rangle,\dots|a_k\rangle\} \overset{\text{component-wise}}{=} \left\{\left|t_{h_1}^{(k)}\right\rangle, \left|t_{h_2}^{(k)}\right\rangle\dots\left|t_{h_{k-1}}^{(k)}\right\rangle, \left|u_h^{(k)}\right\rangle\right\},$$

$$\begin{aligned}
\tilde{O}^{(k)} := {}& \mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}}\left|u_h^{(k)}\right\rangle + \sqrt{q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}}\left|t_{h_j}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{h_k}^{(k)}}\left\langle u_h^{(k)}\right| + \sqrt{q_{g_j}^{(k)}}\left\langle t_{h_j}^{(k)}\right|\right]\\
&+\mathcal{N}\left[\sqrt{q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}}\left|u_h^{(k)}\right\rangle - \sqrt{q_{g_{j'}}^{(k)}}\left|t_{h_j}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{g_j}^{(k)}}\left\langle u_h^{(k)}\right| - \sqrt{q_{h_k}^{(k)}}\left\langle t_{h_j}^{(k)}\right|\right]\\
&+\sum_{i\in\{1,\dots k\}\backslash j}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|
\end{aligned}$$

and $\bar{O}_h^{(k)}$ is given by the basis change $\left\{\left|t_{h_1}^{(k+1)}\right\rangle,\dots\left|t_{h_k}^{(k+1)}\right\rangle\right\} \to \left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle\dots\left|t_{h_{k-1}}^{(k)}\right\rangle\right\}$.
Jump to **End**.

This proof is very similar to the previous.
We are given $\underline{X}^{(k)} = (X_h^{(k)}, X_g^{(k)}, \left|w^{(k)}\right\rangle, \left|v^{(k)}\right\rangle)$,
where $X_h^{(k)} = \sum_{i=1}^k y_{h_i}^{(k)}\left|t_{h_i}^{(k+1)}\right\rangle\left\langle t_{h_i}^{(k+1)}\right|$,
$X_g^{(k)} = \sum_{i=1}^k y_{g_i}^{(k)}\left|t_{h_i}^{(k+1)}\right\rangle\left\langle t_{h_i}^{(k+1)}\right|$, $\left|v^{(k)}\right\rangle = \sum_{i=1}^k q_{g_i}^{(k)}\left|t_{h_i}^{(k+1)}\right\rangle$, $\left|w^{(k)}\right\rangle = \sum_{i=1}^k q_{h_i}^{(k)}\left|t_{h_i}^{(k+1)}\right\rangle$ with the corresponding function instance being $\underline{x}^{(k)} = (h^{(k)}, g^{(k)}, a^{(k)})$ where

$$a^{(k)} = \sum_{i\in\{1,\dots k\}\backslash j} q_{h_i}^{(k)}[y_{h_i}] + (q_{h_j}^{(k)} - q_{g_{j'}}^{(k)})[y_{h_j}] - \sum_{i\in\{1,\dots k\}\backslash j'} q_{g_i}^{(k)}[y_{g_i}].$$

Since we assume that $\underline{X}^{(k)}$ has a solution it follows that $a^{(k)}$ is $[\chi,\xi]$-valid. Thus, the transition $g^{(k-1)} := a_-^{(k)} \to a_+^{(k)} =: h^{(k-1)}$ is also $[\chi,\xi]$-valid where $g^{(k-1)}$ comprises $n_g^{(k-1)} = n_g^{(k)} - 1$ points and $h^{(k-1)}$ comprises $n_h^{(k-1)} = n_h^{(k)}$ points (using the attributes corresponding to the function instance $(h^{(k-1)}, g^{(k-1)}, h^{(k-1)} - g^{(k-1)})$) We use the notation $g = \sum_{i=1}^{n_g} p_{g_i}[x_{g_i}]$ and $h = \sum_{i=1}^{n_h} p_{h_i}[x_{h_i}]$). Since $k = n_g^{(k)} + n_h^{(k)} - 1$, the aforesaid relation yields $n_g^{(k-1)} + n_h^{(k-1)} - 1 = k-1$. We conclude that the matrix instance $\underline{X}^{(k-1)} := (X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle)$ where $X_h^{(k-1)} = \sum_{i=1}^{k-1} y_{h_i}^{(k-1)}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|$, $X_g^{(k-1)} = \sum_{i=1}^{k-1} y_{g_i}^{(k-1)}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|$, $\left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1}\sqrt{q_{g_i}^{(k-1)}}\left|t_{h_i}^{(k)}\right\rangle\right]$,

and
$\left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1}\sqrt{q_{h_i}^{(k-1)}}\left|t_{h_i}^{(k)}\right\rangle\right]$ has a solution for

$$\begin{aligned}
\left\{q_{h_1}^{(k-1)}, \dots q_{h_{k-1}}^{(k-1)}\right\} &\overset{\text{component-wise}}{=} \left\{q_{h_1}^{(k)} \dots, q_{h_{j-1}}^{(k)}, q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}, q_{h_{j+1}}^{(k)}, q_{h_{j+2}}^{(k)} \dots q_{h_{k-1}}^{(k)}\right\} \\
\left\{q_{g_1}^{(k-1)}, \dots q_{g_{k-1}}^{(k-1)}\right\} &\overset{\text{component-wise}}{=} \left\{q_{g_1}^{(k)}, q_{g_2}^{(k)} \dots, q_{g_{j'-1}}^{(k)}, q_{g_{j'+1}}^{(k)}, \dots q_{g_k}^{(k)}\right\} \\
\left\{y_{g_1}^{(k-1)}, \dots y_{g_{k-1}}^{(k-1)}\right\} &\overset{\text{component-wise}}{=} \left\{y_{g_1}^{(k)}, \dots y_{g_{j'-1}}^{(k)}, y_{g_{j'+1}}^{(k)} \dots, y_{g_k}^{(k)}\right\} \\
\left\{y_{h_1}^{(k-1)}, \dots y_{h_{k-1}}^{(k-1)}\right\} &\overset{\text{component-wise}}{=} \left\{y_{h_1}^{(k)}, \dots y_{h_{k-1}}^{(k)}\right\},
\end{aligned}$$

as the corresponding function instance $\underline{x}^{(k-1)}$ is indeed given by
$(h^{(k-1)}, g^{(k-1)}, a^{(k-1)} = a^{(k)})$. Here $\{\left|t_{h_i}^{(k)}\right\rangle\}$ constitute an orthonormal basis which we will soon relate to $\left|t_{h_i}^{(k+1)}\right\rangle$. We used the fact that $q_{h_k}^{(k)} = 0$ as $y_{h_k}^{(k)} = \xi$. To verify this note that $k-1 > n_h^{(k-1)}$ which means that many q_{h_i} are zero; by convention we write the smallest eigenvalue, x_{h_1} first all the way until $x_{h_{n_h}}$ and then to increase the matrix size we append zeros so the $i = n_h, n_h+1 \dots k$ yield $q_{h_i} = 0$. This means that there must exist an $O^{(k-1)}$ which solves $\underline{X}^{(k-1)}$.
As far as the basis change is concerned, we have that $X_h' \geq O'X_g'O'^T$ with $O'\left|v'\right\rangle = \left|w'\right\rangle$ is equivalent to $X_h \geq OX_gO^T$ with $O\left|v\right\rangle = \left|w\right\rangle$, where $O = \bar{O}_h^T O' \bar{O}_g$, $\bar{O}_g\left|v\right\rangle = \left|v'\right\rangle$, $\bar{O}_h\left|w\right\rangle = \left|w'\right\rangle$, $\bar{O}_h X_h \bar{O}_h^T = X_h'$, $\bar{O}X_g\bar{O}_g^T = X_g'$.
We first expand the matrix $\underline{X}^{(k-1)}$ to k dimensions as follows. We already had $X_h^{(k-1)} \geq O^{(k-1)}X_g^{(k-1)}O^{(k-1)T}$ with $O^{(k-1)}\left|v^{(k-1)}\right\rangle = \left|w^{(k-1)}\right\rangle$ and we expand it as

$$\underbrace{y_{h_j}^{(k)}\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right| + X_h^{(k-1)}}_{:=X_h'^{(k)}} \geq$$
$$\underbrace{\left(\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right| + O^{(k-1)}\right)}_{:=O'^{(k)}}\underbrace{\left(y_{h_j}^{(k)}\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right| + X_g^{(k-1)}\right)}_{:=X_g'^{(k)}}\left(\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right| + O^{(k-1)}\right)^T,$$

with $\left|v'^{(k)}\right\rangle = \mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}}\left|u_h^{(k)}\right\rangle + \left|v^{(k-1)}\right\rangle\right]$
and $\left|w'^{(k)}\right\rangle = \mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}}\left|u_h^{(k)}\right\rangle + \left|w^{(k-1)}\right\rangle\right]$. The matrix instance
$\underline{X}'^{(k)} := (X_h'^{(k)}, X_g'^{(k)}, \left|v'^{(k)}\right\rangle, \left|w'^{(k)}\right\rangle)$ yields $\underline{x}'^{(k)} = \underline{x}^{(k)}$. We can now use the equivalence we pointed out above to establish a relation between $X_h^{(k)} \geq O^{(k)}X_g^{(k)}O^{(k)T}$ and $X_h'^{(k)} \geq O'^{(k)}X_g'^{(k)}O'^{(k)T}$ by finding $\bar{O}_g$ and $\bar{O}_h$. We define, somewhat arbitrarily,

$$\begin{aligned}
&\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle \dots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\} \overset{\text{component-wise}}{=} \\
&\left\{\left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \left|t_{h_{j+2}}^{(k+1)}\right\rangle \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\}.
\end{aligned}$$

We require $\bar{O}_g^{(k)}\left|v^{(k)}\right\rangle = \left|v'^{(k)}\right\rangle$. This is a permutation matrix given by $\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\} \rightarrow \left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle \dots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\}$.
We have $\bar{O}_g^{(k)T}X_g'^{(k)}\bar{O}_g^{(k)} = X_g^{(k)}$ as $y_{h_j}^{(k)} = y_{g_{j'}}^{(k)}$. It remains to find $\bar{O}_h^{(k)}$ which we require to

satisfy $\bar{O}_h^{(k)}\left|w^{(k)}\right\rangle = \left|w'^{(k)}\right\rangle$. Let us define $\bar{O}_h^{(k)} = \tilde{O}^{(k)}\left(\sum_{i=1}^k |a_i\rangle\left\langle t_{h_i}^{(k+1)}\right|\right)$, and observe that for $\tilde{O}^{(k)} = \mathbb{I}$ we have $\bar{O}_h^{(k)}\left|w^{(k)}\right\rangle = q_{h_k}^{(k)}\left|u_h^{(k)}\right\rangle + \sum_{i=1}^{k-1} q_{h_i}^{(k)}\left|t_{h_i}^{(k)}\right\rangle$ where

$$\{|a_1\rangle, \dots |a_k\rangle\} \overset{\text{component-wise}}{=} \left\{\left|t_{h_1}^{(k)}\right\rangle, \left|t_{h_2}^{(k)}\right\rangle \dots \left|t_{h_{k-1}}^{(k)}\right\rangle, \left|u_h^{(k)}\right\rangle\right\}.$$

If we define

$$\begin{aligned}\tilde{O}^{(k)} :=& \mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}}\left|u_h^{(k)}\right\rangle + \sqrt{q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}}\left|t_{h_j}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{h_k}^{(k)}}\left\langle u_h^{(k)}\right| + \sqrt{q_{g_j}^{(k)}}\left\langle t_{h_j}^{(k)}\right|\right] \\ &+ \mathcal{N}\left[\sqrt{q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}}\left|u_h^{(k)}\right\rangle - \sqrt{q_{g_{j'}}^{(k)}}\left|t_{h_j}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{g_j}^{(k)}}\left\langle u_h^{(k)}\right| - \sqrt{q_{h_k}^{(k)}}\left\langle t_{h_j}^{(k)}\right|\right] \\ &+ \sum_{i\in\{1,\dots k\}\backslash j}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|,\end{aligned}$$

we get $\bar{O}_h^{(k)}\left|w^{(k)}\right\rangle = \left|w'^{(k)}\right\rangle$. We can also see that $\bar{O}_h^{(k)T} X_g'^{(k)}\bar{O}_h^{(k)}$ is essentially X_g with $\xi^{(k)}$ at $\left|t_{h_k}^{(k+1)}\right\rangle$ replaced by y_{h_j}. We therefore conclude that $X_g'^{(k)} \geq \bar{O}_g^{(k)} X_g^{(k)}\bar{O}_g'^{(k)}$. Substituting we obtain

$$\begin{aligned} X_h'^{(k)} &\geq O'^{(k)} X_g'^{(k)} O'^{(k)T} \geq O'^{(k)}\bar{O}_g^{(k)} X_g^{(k)}\bar{O}_g^{(k)T} O'^{(k)T} \\ \iff \bar{O}_h^{(k)T} X_h'^{(k)}\bar{O}_h^{(k)} &\geq \underbrace{\bar{O}_h^{(k)T} O'^{(k)}\bar{O}_g^{(k)}}_{:=O^{(k)}} X_g^{(k)}\bar{O}_g^{(k)T} O'^{(k)T}\bar{O}_h^{(k)} \\ \iff X_h^{(k)} &\geq O^{(k)} X_g^{(k)} O^{(k)T},\end{aligned}$$

and similarly

$$\begin{aligned} O'^{(k)}\left|v'^{(k)}\right\rangle = \left|w'^{(k)}\right\rangle &\iff O'^{(k)}\bar{O}_g^{(k)}\left|v^{(k)}\right\rangle = \bar{O}_h^{(k)}\left|w^{(k)}\right\rangle \\ &\iff O^{(k)}\left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle,\end{aligned}$$

which completes the proof.

- **Evaluate the Reverse Weingarten Map**:
  1. Consider the point $\left|w^{(k)}\right\rangle / \sqrt{\left\langle w^{(k)}\right| X_h'^{(k)}\left|w^{(k)}\right\rangle}$ on the ellipsoid $X_h'^{(k)}$. Evaluate the normal at this point as $\left|u_h^{(k)}\right\rangle = \mathcal{N}\left(\sum_{i=1}^{n_h^{(k)}}\sqrt{p_{h_i}^{(k)}} x_{h_i}'^{(k)}\left|t_{h_i}^{(k+1)}\right\rangle\right)$. Similarly evaluate $\left|u_g^{(k)}\right\rangle$, the normal at the point $\left|v^{(k)}\right\rangle / \sqrt{\left\langle w^{(k)}\right| X_g'^{(k)}\left|w^{(k)}\right\rangle}$ on the ellipsoid $X_g'^{(k)}$.
  2. Evaluate the Reverse Weingarten maps $W_h'^{(k)}$ and $W_g'^{(k)}$ along $\left|u_h^{(k)}\right\rangle$ and $\left|u_g^{(k)}\right\rangle$, respectively. For a given diagonal matrix $X = \sum_i y_i |i\rangle\langle i| > 0$ and normal vector $|u\rangle = \sum_i u_i |i\rangle$ the Reverse Weingarten map is given by $W_{ij} = \left(-\frac{y_j^{-1} y_i^{-1} u_i u_j}{r^2} + y_i^{-1}\delta_{ij}\right)$ where $r = \sqrt{\sum y_i^{-1} u_i^2}$.
  3. Find the eigenvectors and eigenvalues of the Reverse Weingarten maps. The eigenvectors of W_h' form the h tangent and normal vectors $\left\{\left\{\left|t_{h_i}^{(k)}\right\rangle\right\}, \left|u_h^{(k)}\right\rangle\right\}$. The corresponding radii of

curvature are obtained from the eigenvalues $\left\{\{r_{h_i}^{(k)}\},0\right\} = \left\{\{c_{h_i}^{(k)-1}\},0\right\}$ which are the inverses of the curvature values. The tangents are labeled in decreasing order of radii of curvature (i.e.,increasing order of curvature). Similarly for the g tangent and normal vectors. Fix the sign freedom in the eigenvectors by requiring $\left\langle t_{h_i}^{(k)}|w^{(k)}\right\rangle \geq 0$ and $\left\langle t_{g_i}^{(k)}|v^{(k)}\right\rangle \geq 0$.

– **Finite Method: If** $\lambda \neq -\xi^{(k)}$ and $\lambda \neq -\chi^{(k)}$, i.e. if it is the finite case **then**

1. $\bar{O}^{(k)} := \left|u_h^{(k)}\right\rangle\left\langle u_g^{(k)}\right| + \sum_{i=1}^{k-1}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{g_i}^{(k)}\right|$
2. $\left|v^{(k-1)}\right\rangle := \bar{O}^{(k)}\left|v^{(k)}\right\rangle - \left\langle u_h^{(k)}\right|\bar{O}^{(k)}\left|v^{(k)}\right\rangle\left|u_h^{(k)}\right\rangle$ and $\left|w^{(k-1)}\right\rangle := \left|w^{(k)}\right\rangle - \left\langle u_h^{(k)}|w^{(k)}\right\rangle\left|u_h^{(k)}\right\rangle$.
3. Define $X_h^{(k-1)} := \mathrm{diag}\{c_{h_1}^{(k)}, c_{h_2}^{(k)} \dots, c_{h_{k-1}}^{(k)}\}$,
   $X_g^{(k-1)} := \mathrm{diag}\{c_{g_1}^{(k)}, c_{g_2}^{(k)} \dots c_{g_{k-1}}^{(k)}\}$.
4. **Jump** to **End**.

First, we need to prove that $O^{(k)}$ must have the form
$\left(\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right| + O^{(k-1)}\right)\bar{O}^{(k)}$
for $\bar{O}^{(k)} := \left|u_h^{(k)}\right\rangle\left\langle u_g^{(k)}\right| + \sum_{i=1}^{k-1}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{g_i}^{(k)}\right|$ if $O^{(k)}$ is to be a solution of the matrix instance $\underline{\mathrm{X}}^{(k)}$. This is best explained by imagining that Arthur is trying to find the orthogonal matrix and Merlin already knows the orthogonal matrix but has still been following the steps performed so far. Recall that we are now at a point where

$$\sum a'(x)x = \langle w|X_h'|w\rangle - \langle v|X_g'|v\rangle = \langle w|X_h'|w\rangle - \langle w|OX_g'O^T|w\rangle = 0.$$

From Merlin's point of view along the $|w\rangle$ direction the ellipsoids X_h' and $OX_g'O^T$ touch. Suppose he started with the ellipsoids X_h', X_g' and only subsequently rotated the second one. He can mark the point along the direction $|v\rangle$ on the X_g' ellipsoid as the point that would after rotation touch the X_h' ellipsoid because as $X_g' \longrightarrow OX_g'O^T$ the point along the $|v\rangle$ direction would get mapped to the point along the direction $O|v\rangle = |w\rangle$. Now, since the ellipsoids touch it must be so, Merlin deduces, that the normal of the ellipsoid X_g' at the point $|v\rangle/\sqrt{\langle v|X_g'|v\rangle}$ is mapped to the normal of the ellipsoid X_h' at the point $|w\rangle/\sqrt{\langle w|X_h'|w\rangle}$ when X_g' is rotated to $OX_g'O^T$, i.e. $O\left|u_g\right\rangle = |u_h\rangle$.

From Arthur's point of view, who has been following Merlin's reasoning, in addition to knowing that O must satisfy $O|v\rangle = |w\rangle$ he now knows that it must also satisfy $O\left|u_g\right\rangle = |u_h\rangle$.

Merlin further concludes that the curvature of the X_g' ellipsoid at the point $|v\rangle/\sqrt{\langle v|X_g'|v\rangle}$ must be larger than the curvature of the X_h' ellipsoid at the point $|w\rangle/\sqrt{\langle w|X_h'|w\rangle}$. To be precise, he needs to find a method for evaluating this curvature. He knows that the brute-force way of doing this is to find a coordinate system with its origin on the said point and then imagining the manifold, locally, as a function from $n-1$ coordinates to one coordinate, call it $x_n(x_1, x_2 \dots x_{n-1})$. The curvature of this object is a generalization of the second derivative which forms a matrix with its elements given by $\partial_{x_i}\partial_{x_j}x_n$. Since this matrix is symmetric he knows it can be diagonalized. The directions of the eigenvectors of this matrix he calls the principle directions of curvature and the curvature values are the corresponding eigenvalues. He recalls that there is a simpler way of evaluating these principle directions and curvatures by using the Weingarten map. The eigenvectors of the Reverse Weingarten map W_h', evaluated for X_h' at $|w\rangle$, yield the normal and tangent vectors and the corresponding eigenvalues are the radii of curvature (curvature is the inverse of the radius of curvature). Similarly for the Reverse Weingarten map W_g' evaluated for X_g' at

$|v\rangle$. With this knowledge Merlin can write, for some $\tilde{O}_{ij} \in \mathbb{R}$ such that $\sum_j \tilde{O}_{ij}\tilde{O}_{jk} = \delta_{ik}$,

$$\begin{aligned} O^{(k)} &= |u_h\rangle \langle u_g| + \sum_{i,j} \tilde{O}_{ij} \left|t_{h_i}\right\rangle \left\langle t_{g_j}\right| \\ &= \left(|u_h\rangle \langle u_h| + \underbrace{\sum_{i,j} \tilde{O}_{ij} \left|t_{h_i}\right\rangle \left\langle t_{h_j}\right|}_{=O^{(k-1)}} \right) \left(\underbrace{|u_h\rangle \langle u_g| + \sum_{i} \left|t_{h_i}\right\rangle \left\langle t_{g_i}\right|}_{=\bar{O}^{(k)}} \right). \end{aligned}$$

He then turns to his intuition about the curvature of the smaller ellipsoid being more than that of the larger ellipsoid. He observes that equivalently, the radius of curvature of the smaller ellipsoid must be smaller than that of the larger ellipsoid. To make this precise he notes that the Weingarten map W_g' gets transformed to $OW_g'O^T$ when X_g' is rotated as $OX_g'O^T$. He considers the point $|w\rangle / \sqrt{\langle w| X_h' |w\rangle}$, which is shared by both the X_h' and the $OX_g'O^T$ ellipsoid. It must be so that along all directions in the tangent plane, the smaller X_h' ellipsoid must have a smaller radius of curvature than the $OX_g'O^T$ ellipsoid, i.e. for all $|t\rangle \in \text{span}\{\left|t_{h_i}\right\rangle\}$, $\langle t| W_h' |t\rangle \leq \langle t| OW_g'O^T |t\rangle$. Restricting his attention to the tangent space he deduces the statement is equivalent to $W_h' \leq OW_g'O^T$. He writes this explicitly as $\sum c_{h_i}^{-1} \left|t_{h_i}\right\rangle \left\langle t_{h_i}\right| \leq \sum c_{g_i}^{-1} O \left|t_{g_i}\right\rangle \left\langle t_{g_i}\right| O^T$. Now he uses the form of O he had deduced to obtain $\sum c_{h_i}^{-1} \left|t_{h_i}\right\rangle \left\langle t_{h_i}\right| \leq \sum c_{g_i}^{-1} O^{(k-1)} \left|t_{h_i}\right\rangle \left\langle t_{h_i}\right| O^{(k-1)T}$. From this he concludes that the inequality $X_h^{(k-1)} \geq O^{(k-1)} X_g^{(k-1)} O^{(k-1)T}$ must hold.
Arthur summarizes that Merlin's reasoning entails that $O^{(k)}$ must always have the form

$$O^{(k)} = \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)} \right) \bar{O}^{(k)},$$

and that $O^{(k-1)}$ must satisfy the constraint
$X_h^{(k-1)} \geq O^{(k-1)} X_g^{(k-1)} O^{(k-1)T}$. Merlin, surprised by the similarity of the constraint he obtained with the one he started with, extends his reasoning to the vector itself. He knows that $O^{(k)} \left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle$ but now he substitutes for $O^{(k)}$ to obtain
$\left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)} \right) \bar{O}^{(k)} \left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle$.
He observes that $O^{(k-1)}$ can not influence the $\left|u_h^{(k)}\right\rangle$ component of the vector $\bar{O}^{(k)} \left|v^{(k)}\right\rangle$.
He thus projects out the $\left|u_h^{(k)}\right\rangle$ component to obtain

$$O^{(k-1)} \underbrace{\left(\bar{O}^{(k)} \left|v^{(k)}\right\rangle - \langle u_h| \bar{O}^{(k)} \left|v^{(k)}\right\rangle |u_h\rangle \right)}_{=\left|v^{(k-1)}\right\rangle} = \underbrace{\left|w^{(k)}\right\rangle - \left\langle u_h^{(k)} | w^{(k)} \right\rangle \left|u_h^{(k)}\right\rangle}_{=\left|w^{(k-1)}\right\rangle}.$$

With this, Arthur realizes, he can reduce his problem involving a k-dimensional orthogonal matrix into a smaller problem in $k-1$ dimensions with exactly the same form. Since Merlin's orthogonal matrix was an arbitrary solution and the constraints involved do not depend explicitly on it but only on the initial problem, Arthur concludes that this reduction must hold for all solutions.

- **Wiggle-v Method: If** $\lambda = -\xi^{(k)}$ or $\lambda = -\chi^{(k)}$ **then**
The above method of matching the normals works well as long as the appropriate operator monotone—the one that gives X_h' and X_g' for which $|w\rangle / \sqrt{\langle w| X_h' |w\rangle}$ is a point on both X_h'

and $OX_g'O^T$)—doesn't yield infinities. If infinities arise, it means that one of the directions involved has infinite curvature which, in turn, means that the component of the normal along this direction can be arbitrary. In other words, imagine a line contained inside an ellipsoid and centered at its origin that touches its boundaries. The line can be seen as an ellipse with infinite curvature along one of the directions. The normal of the line at its tip is arbitrary and therefore we can't require the usual condition that the normals of the two curves must coincide. The solution is to consider the sequence leading to the aforesaid situation.

1. $\left|u_h^{(k)}\right\rangle$ is renamed to $\left|\bar{u}_h^{(k)}\right\rangle$, $\left|u_g^{(k)}\right\rangle$ and remains the same.
2. Let $\tau = \cos\theta := \left\langle u_g^{(k)} | v^{(k)} \right\rangle / \left\langle \bar{u}_h^{(k)} | w^{(k)} \right\rangle$. Let $\left|\bar{t}_h^{(k)}\right\rangle$ be an eigenvector of $X_h'^{(k)-1}$ with zero eigenvalue. Redefine
$$\left|u_h^{(k)}\right\rangle := \cos\theta \left|\bar{u}_h^{(k)}\right\rangle + \sin\theta \left|\bar{t}_h^{(k)}\right\rangle$$
$$\left|t_{h_k}^{(k)}\right\rangle = s\left(-\sin\theta \left|\bar{u}_h^{(k)}\right\rangle + \cos\theta \left|\bar{t}_h^{(k)}\right\rangle\right),$$
where the sign $s \in \{1, -1\}$ is fixed by requiring $\left\langle t_{h_k}^{(k)} | w^{(k)} \right\rangle \geq 0$.
3. $\bar{O}^{(k)}$ and $\left|v^{(k-1)}\right\rangle, \left|w^{(k-1)}\right\rangle$ are evaluated as step 1 and 2 of the finite case.
4. Define
$$X_h'^{(k-1)} := \text{diag}\{c_{h_1}^{(k)}, c_{h_2}^{(k)}, \dots, c_{h_{k-1}}^{(k)}\}$$
$$X_g'^{(k-1)} := \text{diag}\{c_{g_1}^{(k)}, c_{g_2}^{(k)}, \dots, c_{g_{k-1}}^{(k)}\}.$$
Let $[\chi'^{(k-1)}, \xi'^{(k-1)}]$ denote the smallest interval containing spec$[X_h'^{(k-1)} \oplus X_g'^{(k-1)}]$. Let $\lambda' = -\chi'^{(k-1)} + 1$ where instead of 1 any positive number would also work. Consider $f_{\lambda''}$ on $[\chi'^{(k-1)}, \xi'^{(k-1)}]$, and let $\eta = -f_{\lambda'}(\chi'^{(k-1)}) + 1$. Define
$$X_h^{(k-1)} := f_{\lambda'}(X_h'^{(k-1)}) + \eta \text{ and } X_g^{(k-1)} := f_{\lambda'}(X_g'^{(k-1)}) + \eta.$$
5. **Jump** to **End**.
We start with the case $\lambda = -\xi^{(k)}$. The case with $\lambda = -\chi^{(k)}$ follows analogously. For the moment we assume $\eta = 0$ for simplicity; for $\eta \neq 0$ the argument goes through essentially unchanged. Since $\langle w| f_{-\xi}(X_h) |w\rangle - \langle v| f_{-\xi}(X_g) |v\rangle = 0$, we conclude that $y_{h_i}^{(k)} = \xi$ implies $q_{h_i} = 0$. After the application of the map $f_{-\xi}$ these $y_{h_i}^{(k)}$s and $y_{g_i}^{(k)}$s would become infinities but $\left\langle t_{h_i}^{(k+1)} |w\right\rangle$ and $\left\langle t_{g_i}^{(k+1)} |v\right\rangle$ would be zero (we suppressed the superscripts for $\left|v^{(k)}\right\rangle$ and $\left|w^{(k)}\right\rangle$). Since the eigenvalues are arranged in ascending order in $X_h^{(k)}$ we have $y_{h_k}^{(k)} = \xi$ and the corresponding vector is $\left|t_{h_k}^{(k+1)}\right\rangle =: |\bar{t}_h\rangle$. It is useful to define $\left|\tilde{t}_{h_i}\right\rangle = \left|t_{h_i}\right\rangle$ for $i = 1, 2, \dots j-1$ and $\left|\bar{t}_{h_l}\right\rangle = \left|t_{h_i}\right\rangle$ for $i = j, j+1, \dots k$, $l = (i-j)+1$ where j is the smallest i for which $x_{h_i} = \xi$ (their existence is a straightforward consequence of dimension counting, $k \geq n_g + n_h - 1$). We focus on the two-dimensional plane spanned by $|w\rangle$ and $|\bar{t}_h\rangle$.
From Merlin's point of view, since he has a solution $O^{(k)}$ to the matrix instance
$$\underline{\mathrm{X}}^{(k)} = \{X_h^{(k)}, X_g^{(k)}, \left|w^{(k)}\right\rangle, \left|v^{(k)}\right\rangle\},$$
his solution is also a solution to the matrix instance
$$\underline{\mathrm{X}}^{(k)}(\lambda) := \left\{f_\lambda(X_h^{(k)}), f_\lambda(X_g^{(k)}), \left|w^{(k)}\right\rangle, \left|v^{(k)}\right\rangle\right\}$$

for $\lambda \leq -\xi$ but close enough to $-\xi$ such that $f_\lambda(X_h), f_\lambda(X_g) > 0$. This is a consequence of f_λ being operator monotone. Using Corollary 77 and Lemma 78 we know that since the ellipsoids corresponding to the matrix instance $\underline{\mathrm{X}}(-\xi)$ touch along $|w\rangle$ – as we are given that $\langle w| f_{-\xi}(X_h) |w\rangle - \langle w| O f_{-\xi}(X_g) O^T |w\rangle = \langle w| f_{-\xi}(X_h) |w\rangle - \langle v| f_{-\xi}(X_g) |v\rangle = 0$— there must also exist some vector $|c(\lambda)\rangle$ such that
$\langle c(\lambda)| f_\lambda(X_h) |c(\lambda)\rangle - \langle c(\lambda)| O f_\lambda(X_g) O^T |c(\lambda)\rangle = 0$; that is the ellipsoids corresponding to the matrix instance $\underline{\mathrm{X}}(\lambda)$ touch along the said direction. To meet the other conditions of the lemma it suffices to assume that X_h and X_g do not have a common eigenvalue which in turn is guaranteed by the "remove spectral collision" part of the algorithm.
It is easy to convince oneself that $\lim_{\lambda\to -\xi} |c(\lambda)\rangle = |w\rangle$[22]. We can write

$$|w\rangle = \sum_{i=1}^{j-1} q_{h_i} \left|\tilde{t}_{h_i}\right\rangle \text{ because } \left\langle \bar{t}_{h_i} | w \right\rangle = 0.$$

There is no such restriction on $|c(\lambda)\rangle$ which can have the more general form $|c(\lambda)\rangle = \sum_{i=1}^{j-1} c(\lambda)_i \left|\tilde{t}_{h_i}\right\rangle + \sum_{i=j}^{k} c(\lambda)_i \left|\bar{t}_{h_l}\right\rangle$ where $l = (i-j)+1$. Restating one of the limit conditions for $i = j, j+1 \ldots k$, we have $\lim_{\lambda \to -\xi} |c(\lambda)_i\rangle = 0$. If O is a solution it entails that

$$\acute{O}(\lambda) := \left(\sum_{i=1}^{j-1} \left|\tilde{t}_{h_i}\right\rangle \left\langle \tilde{t}_{h_i}\right| + \sum_{i,m=1}^{k-j+1} Q(\lambda)_{im} \left|\bar{t}_{h_i}\right\rangle \left\langle \bar{t}_{h_m}\right| \right) O$$

is also a solution, where $Q(\lambda)$ is an orthogonal matrix in the space spanned by $\left\{\left|\bar{t}_{h_i}\right\rangle\right\}$. This is a consequence of the fact that $\{\left|\bar{t}_{h_i}\right\rangle\}$ spans an eigenspace—with the same eigenvalue $f_\lambda(\xi)$ – of $f_\lambda(X_h)$. We can use this freedom to ensure that the point of contact always has the form

$$|c(\lambda)\rangle = \sum_{i=1}^{j-1} c(\lambda)_i \left|\tilde{t}_{h_i}\right\rangle + \bar{c}(\lambda) \left|\bar{t}_h\right\rangle$$

where $\bar{c}(\lambda) = \sqrt{\sum_{i=j}^{k} c(\lambda)_i^2}$ which must vanish in the limit $\lambda \to -\xi$ as its constituents disappear in the said limit. Similarly,
$\lim_{\lambda\to-\xi} c(\lambda)_i = q_{h_i}$.
Next we evaluate the normals $|u_h(\lambda)\rangle$ at $|c(\lambda)\rangle$ for the ellipsoid represented by $f_\lambda(X_h)$ and the normal $|\bar{u}_h\rangle$ at $|w\rangle$ for the ellipsoid represented by $f_{-\xi}(X_h)$ to show that $\lim_{\lambda\to-\xi} |u_h(\lambda)\rangle \neq |\bar{u}_h\rangle$ (see Figure 15). The right-most term in
$|u_h(\lambda)\rangle = \mathcal{N}\left[\sum_{i=1}^{j-1} f_\lambda(y_{h_i}) c(\lambda)_i \left|\tilde{t}_{h_i}\right\rangle + f_\lambda(\xi)\bar{c}(\lambda) \left|\bar{t}_h\right\rangle\right]$ has $f_\lambda(\xi)$ approaching infinity and $\bar{c}(\lambda)$ approaching zero as λ tends to $-\xi$. This is why it can have a finite component along $|\bar{t}_h\rangle$. On the other hand, $|\bar{u}_h\rangle = \mathcal{N}\left[\sum_{i=1}^{j-1} f_{-\xi}(y_{h_i}) q_{h_i} \left|\tilde{t}_{h_i}\right\rangle\right]$ which has no component along $|\bar{t}_h\rangle$. Since $\lim_{\lambda\to-\xi} f_\lambda(y_{h_i}) = f_{-\xi}(y_{h_i})$ and $\lim_{\lambda\to-\xi} c(\lambda)_i = q_{h_i}$ for $i \in \{1, 2 \ldots j-1\}$, we can write

$$\lim_{\lambda\to-\xi} |u_h(\lambda)\rangle = \cos\theta \, |\bar{u}_h\rangle + \sin\theta \, |\bar{t}_h\rangle := |u_h\rangle \, .$$

We must use $|u_h\rangle$ instead of $|\bar{u}_h\rangle$ to be able to use the reasoning of the finite method. However, we do not know $\cos\theta$ yet. We proceed as in the finite method with the assumption that $|c(\lambda)\rangle$ is known and then use a consistency condition to find $\cos\theta$ in terms of known quantities. At this point we re-introduce the superscripts as we reduce the

[22] Since $f_\lambda(X_h)$ is very close to $f_{-\xi}(X_h)$, the vectors satisfying the condition should also be very close.

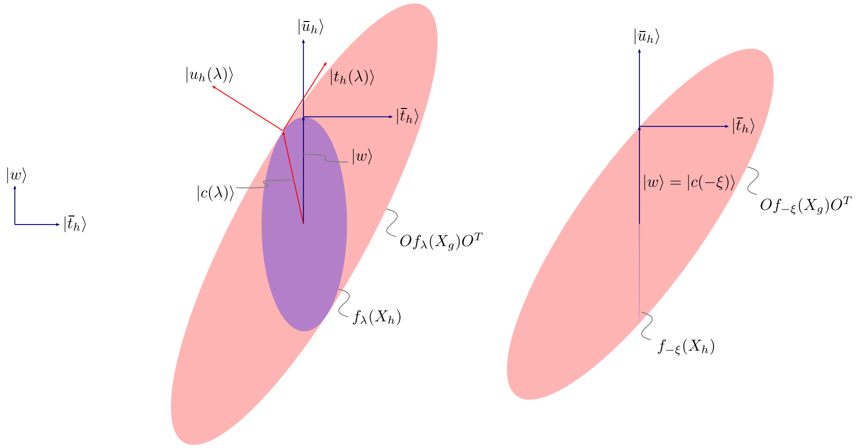


Figure 15: A sequence leading to infinite curvature.

dimension of the problem. Let the normal and tangent vectors at $O^T\,|c(\lambda)\rangle$ for $f_\lambda(X_g)$ be given by $\left\{\left|u_g^{(k)}(\lambda)\right\rangle, \left\{t_{gi}^{(k)}(\lambda)\right\}\right\}$. Similarly at $|c(\lambda)\rangle$ for $f_\lambda(X_h)$ the normal and tangent vectors are $\left\{\left|u_h^{(k)}(\lambda)\right\rangle, \left\{t_{h_i}^{(k)}(\lambda)\right\}\right\}$. From the finite method we know that $O^{(k)}(\lambda) := \left(|u_h(\lambda)\rangle\,\langle u_h(\lambda)| + O^{(k-1)}\right)\bar{O}^{(k)}$ where $\bar{O}^{(k)} = \left|u_h^{(k)}(\lambda)\right\rangle\left\langle u_g^{(k)}(\lambda)\right| + \sum_i \left|t_{h_i}^{(k)}\right\rangle\left\langle t_{gi}^{(k)}\right|$ can be used to reduce the problem into a smaller instance of itself. In particular, we must have $\left\langle u_h^{(k)}(\lambda)|w\right\rangle = \left\langle u_h^{(k)}(\lambda)\right| O^{(k)}(\lambda)\,|v\rangle = \left\langle u_g^{(k)}(\lambda)|v\right\rangle$ because $O^{(k-1)}$ can influence only the subspace spanned by $\left\{\left|t_{h_i}^{(k)}\right\rangle\right\}$, and the component of the vectors $|w\rangle$ and $O^{(k)}\,|v\rangle$ along $\left|u_h^{(k)}(\lambda)\right\rangle$ must match for consistency.

We can determine $\cos\theta$ by taking the limit of the aforesaid condition as $\langle u_h|w\rangle = \langle u_g|v\rangle$ (we suppressed the superscripts again). Substituting $|u_h\rangle = \cos\theta\,|\bar{u}_h\rangle + \sin\theta\,|\bar{t}_h\rangle$ we obtain $\cos\theta = \frac{\langle u_g|v\rangle}{\langle\bar{u}_h|w\rangle}$.

We proceed to find the limit of the reverse Weingarten maps. The reverse Weingarten map for $f_\lambda(X_g)$ along the normal $|u_g(\lambda)\rangle$ has a well-defined limit as $\lambda \longrightarrow -\xi$. We consider the case for $f_\lambda(X_h)$ along the normal $|u_h(\lambda)\rangle$. The support function as defined in Equation (20) is finite in the limit $\lambda \longrightarrow -\xi$[23]. Let us denote it by $h(\lambda)$. The reverse Weingarten map as defined in Equation (21) is given by

$$(W_h(\lambda))_{im} = -\frac{1}{h(\lambda)^2}\frac{u_{h_i}(\lambda)u_{h_m}(\lambda)}{f_\lambda(y_{h_i})f_\lambda(y_{h_m})} + \frac{\delta_{im}}{f_\lambda(x_{h_i})}.$$

Since $\lim_{\lambda\to-\xi}|u(\lambda)\rangle$ is well-defined, $\lim_{\lambda\to-\xi}h(\lambda)$ is finite, and we only need to show that $\lim_{\lambda\to-\xi}1/f_\lambda(y_{h_i})$ is well-defined. We assumed η is zero so $f_{-\xi}(y_{h_i}) \neq 0$. If η is not zero we must consider $f_{-\xi}(y_{h_i}) + \eta$ everywhere but that changes no argument. For $i = 1, 2\ldots j-1$, $f_{-\xi}(y_{h_i})$ is finite but for $i = j, j+1\ldots k$, $f_{-\xi}(y_{h_i})$ it is not well-defined. However $1/f_{-\xi}(y_{h_i}) = 0$, and we therefore conclude that

$$\lim_{\lambda\to-\xi}(W_h(\lambda))_{im} = \begin{cases} -\frac{1}{h^2}\frac{u_{h_i}u_{h_m}}{f_{-\xi}(y_{h_i})f_{-\xi}(y_{h_m})} + \frac{\delta_{im}}{f_{-\xi}(y_{h_i})} & i, m \in \{1, 2\ldots j-1\} \\ 0 & i, m \in \{j, j+1\ldots k\} \end{cases},$$

[23] Use the definition of the normal to get $\sum x_i^{-1}u_i^2 = \sum x_i^{-1}x_i^2c_i^2 = \sum x_ic_i^2 = \langle c|\,X\,|c\rangle$, plug in $|c\rangle = |w\rangle$, $X = f_{-\xi}(X_h)$ and then use $\langle w|\,f_{-\xi}(X_h)\,|w\rangle - \langle v|\,f_{-\xi}(X_g)\,|v\rangle = 0$ which means both must be finite; not that we already dealt with the troublesome case of $\infty - \infty$ in the "remove spectral collision" part of the algorithm.

which is simply the reverse Weingarten map evaluated for $f_{-\xi}(X_h)$ along $|u_h\rangle = \cos\theta\,|\bar{u}_h\rangle + \sin\theta\,|\bar{t}_h\rangle$ and $\cos\theta = \langle u_g|v\rangle / \langle \bar{u}_h|w\rangle$. It remains to relate W_h with the reverse Weingarten map, $\bar{W}_h$, evaluated for $f_{-\xi}(X_h)$ along $|\bar{u}_h\rangle$. It is easy to see that $W_h = \bar{W}_h$ because only the $\cos\theta\,|\bar{u}_h\rangle$ part contributes to the non-zero portion of W_h and the $\cos\theta$ factor gets canceled due to the h^2 term. Moreover, the normal vector is an eigenvector of the reverse Weingarten map evaluated along it, with eigenvalue zero. This tells us that if there are tangent vectors with zero radius of curvature then the normal is not uniquely defined. Since both $|\bar{u}_h\rangle$, $|\bar{t}_h\rangle$ have zero eigenvalues for $\bar{W}_h (= W_h)$ and $|u\rangle = \cos\theta\,|\bar{u}_h\rangle + \sin\theta\,|\bar{t}_h\rangle$ we define $|t_h\rangle := s\,(\sin\theta\,|\bar{u}_h\rangle - \cos\theta\,|\bar{t}_h\rangle)$ to span the same space so that $|u\rangle$ is the correct normal vector and $|t_h\rangle$ is the correct tangent vector corresponding to the point $|w\rangle$ of $f_{-\xi}(X_h)$.
The final step is to convert the condition on the reverse Weingarten map into a condition on the Weingarten map itself. After extracting the tangent vectors appropriately, one simply needs to add a constant before inverting to obtain the Weingarten map condition. This is done in the last step and completes the proof of the wiggle-v method for $\lambda = -\xi$.
To see how the same reasoning applies to the $\lambda = -\chi$ case first note that for $\lambda \geq -\chi$ we have $f_\lambda(X_h), f_\lambda(X_g) < 0$ (assuming $\eta = 0$ as before). The condition $f_\lambda(X_h) \geq O f_\lambda(X_g) O^T$ can then be expressed as $-f_\lambda(X_g) \geq -O^T f_\lambda(X_h) O$ with $O^T\,|w\rangle = |v\rangle$ which can now be reasoned analogously to the above analysis.

- **End**: Restart **PHASE 2** with the newly obtained $(k-1)$ sized objects.

The dimension after every iteration is $k-1 \geq n_g^{(k-1)} + n_h^{(k-1)} - 1$ starting with the assumption $k \geq n_g^{(k)} + n_h^{(k)} - 1$. The reason is that either $n_g^{(k-1)} = n_g^{(k)} - 1$ or $= n_g^{(k)}$. Similarly, either $n_h^{(k-1)} = n_h^{(k)} - 1$ or $= n_h^{(k)}$. Justification of this is simply that we remove at least one component from the two vectors (from the $n_g^{(k)}$ for the wiggle-v). To see this, note that in the finite case we remove one from both as we express the vector in a new basis. This new basis is the space where the vector has finite support. We then remove one of the components in the sub-problem. In the infinite case, it is possible that we remove one and add one for $n_h^{(k-1)}$, assuming it is the usual wiggle-v, but we necessarily reduce $n_g^{(k-1)}$ as this is similar to the finite case. For the other wiggle-w, g and h get swapped but the counting stays the same.

5.4.3 PHASE 3: RECONSTRUCTION

Let k_0 be the iteration at which the algorithm stops. Using the relation

$$O^{(k)} = \bar{O}_g^{(k)} \left(\left| u_h^{(k)} \right\rangle \left\langle u_h^{(k)} \right| + O^{(k-1)} \right) \bar{O}_h^{(k)}$$

(or its transpose if $s^{(k)} = -1$), evaluate $O^{(k_1)}$ from $O^{(k_0)} := \mathbb{I}_{k_0}$, then $O^{(k_2)}$ from $O^{(k_1)}$, then $O^{(k_3)}$ from $O^{(k_2)}$ and so on until $O^{(n)}$ is obtained. $O^{(n)}$ solves the matrix instance $\underline{X}^{(n)}$ that we started with. In terms of EBRM matrices, the solution is given by $H = X_h^{(n)}$, $G = O^{(n)} X_g O^{(n)T}$ and $|w\rangle = \left|w^{(n)}\right\rangle$.

5.5 Preliminary implementation

A preliminary implementation of the algorithm on python [ARW18], which is usable but not automated enough for an end-user, yielded the following results.

1. *f_0-assignments need neither padding nor operator monotones.* We have $\langle x_h\rangle = \langle x_g\rangle$ which means that for the first iteration we do not need to use any operator monotone function. Surprisingly, though,

we saw that even for subsequent iterations, we do not need operator monotones; this also explained why[24] we did not need padding, i.e. the solution had size $n \times n$ for $n = n_g = n_h$. In Section 6 we prove it analytically, and follow this geometric approach to construct a more general solution covering these assignments as well.

2. *Moves in the bias* 1/18 *protocol do not need padding (no wiggle-v).* We already know analytically that there are specific cases where padding is required. However, when we tried to numerically implement the moves involved in protocols going as low as $\epsilon = 1/18$, as proposed by C. Mochon, we found that in no case was padding necessary, which means the wiggle-v method was never invoked.

3. *Trick to improve the precision of the EMA algorithm.* The algorithm tries to find a λ such that $\langle w| f_\lambda(X_h) |w\rangle - \langle v| f_\lambda(X_g) |v\rangle = 0$. In the finite case, for consistency, we must also have $\langle w|n_h(\lambda)\rangle = \left\langle v|n_g(\lambda)\right\rangle$. This is because in subsequent steps, the orthogonal space is affected, therefore, if the component of the honest states along the normals is not mapped correctly, it would not get fixed later; this would mean there is no solution as we are only imposing necessary conditions. We observed that, numerically, we get a better precision if we use the latter condition for fine-tuning the result—after applying the former for obtaining a more course-grained solution. While analytically, the first condition implies the latter exactly, this ceases to be the case numerically due to the finiteness of precision. We understand this improvement as a consequence of the honest state being explicitly mapped correctly (up to the computer's precision) if we use the method involving normals, while in the latter this should happen implicitly.

We also pinpointed the following limitations of this implementation:

1. *Limited wiggle-v.* We have not fully implemented the wiggle-v method which means that it would be cumbersome to apply it to the general merge and split, for instance. However, for them we already give the explicit Blinkered Unitaries. For the rest, as we already saw, it does not even seem necessary.

2. *Other issues.* Sometimes due to noise, arising from finiteness of the precision, our global minimizer gets trapped into local minima and has to be guided manually by looking at the graph. This means that a refined algorithm should be able to solve this problem. Further, we did not implement the systematic method defined by the EMA algorithm for finding the spectrum of the matrices, but it appears that almost any guess works for the assignments used by C. Mochon.

[24] To see this, note that the only time we spill over to the extra dimensions, is when we use the wiggle-v method. Otherwise, we stay inside the first $\max(n_g, n_h)$ dimensions.

6 Approaching bias $\epsilon(k) = 1/(4k+2)$ — a geometric solution

It is natural to ask how the analytic solution from Section 4, which was algebraic in nature, and the numerical solution from Section 5, which was geometric in nature, are related. The goal of this section is to shed some light on this connection. Here, we again construct analytic solutions to monomial assignments (see Section 4), but this time, using a geometric approach. To this end, we combine and extend ideas from both Section 4 and Section 5.[25] We hope that having multiple ways of solving the monomial assignment aids the construction of a general analytic solution which works for all valid functions.

We begin our discussion by contrasting our approach here with the EMA algorithm (see Section 5) and the algebraic solution (see Section 4). Recall that the EMA algorithm resorted to numerical algorithms for two purposes: (1) diagonalizing matrices and (2) solving polynomial equations. Since we seek an analytic solution here, we must somehow address these issues. Issue (1) is handled using three techniques. *First*, we recast the problem using isometries instead of unitaries. Consequently, unlike the EMA algorithm, where in order to consider sub-instances of the problem one had to determine a basis for the tangent space of the associated initial ellipsoids, here we always consider matrices of the same dimension, but each subproblem is described by matrices of one *rank* less than its parent problem. That helps, as it allows one to reduce the rank, using only one vector which, in turn, admits an analytic description. *Second*, we derive and use analytic expressions for the various geometric properties. In the EMA algorithm, their computation relied on the aforementioned basis of the tangent space. *Finally*, we restrict ourselves to f-assignments (see Definition 32). The reason for the restriction is essentially the same as that for the algebraic solution—f-assignments are a sum of monomial assignments which are easier to analyze. This is also related to issue (2) which arises in the EMA algorithm because, recall, that ellipsoids need to be stretched and aligned so that the contact point is along the desired direction. This was crucial for reducing the dimension of the problem, which is what ultimately led to the solution. Monomial assignments, we show, have the special property that they are automatically always aligned. This may be seen as the geometric manifestation of the properties of monomial assignments which were used to construct the algebraic solution.

We introduce some notation which partially overlaps with that of Section 4, but diverges as it is built further. Suppose S is a 4-tuple (an ordered list with 4 elements) and we wish to refer to the third element of S. We write this as

$$(*, *, p, *) := S. \tag{22}$$

We represent the concatenation of two tuples as $(a, b, c) \oplus (d, e) = (a, b, c, d, e)$. A matrix of rank at most k is denoted by $M^{\bar{k}}$. We always use a bar in the superscript to distinguish it from powers. For instance, $(M^{\bar{k}})^2$ refers to the square of the rank k matrix $M^{\bar{k}}$. Given a projector Π, we denote the set $\{\Pi |v\rangle \mid |v\rangle \in \mathbb{R}^n\}$ by $\Pi\mathbb{R}^n$. Recall that in Section 4 we introduced the use of the symbol, $\dashv$, to represent the inverse of a matrix $G \geq 0$ on its non-zero eigenspace, and we called it the pseudo-inverse of G.

We briefly revisit the ellipsoid picture introduced in Section 5, this time adapting the notation to accommodate low rank matrices.

Definition 84 (Ellipsoid and Map)**.** Given an $n \times n$ matrix $G \geq 0$, let Π be a projector onto the non-zero eigenvalue eigenspace of G. The *ellipsoid* associated with G is given by $S_G := \{|s\rangle \in \Pi\mathbb{R}^n | \langle s| G |s\rangle = 1\}$. The *ellipsoid map*, $\mathcal{E}_G : \Pi\mathbb{R}^n \longrightarrow \Pi\mathbb{R}^n$, is defined as $\mathcal{E}_G(|v\rangle) = |v\rangle / \sqrt{\langle v| G |v\rangle}$.

Notice that for $G = \sum_i g_i |i\rangle \langle i|$ and $|s\rangle = \sum_i s_i |i\rangle$, the equation $\langle s| G |s\rangle = 1$ can be written as $\sum_i g_i s_i^2 = 1$, which clearly describes an ellipsoid. As motivated in Section 1.1.3, and then extensively used in Section 5, our interest in the geometry of ellipsoids stems from its connection with matrix inequalities which appear

[25] We stumbled upon this solution first, and constructed the algebraic solution later. However, in this presentation, we chose to flip the order for clarity.

in EBRM transitions (see Corollary 144). Let $H \geq 0$ and $G \geq 0$. One can rewrite a matrix inequality as follows:

$$\begin{aligned} H - OGO^T \geq 0 &\iff \langle s|\, H\, |s\rangle - \langle s|\, OGO^T\, |s\rangle \geq 0 \qquad \forall\ |s\rangle \\ &\iff \langle s|\, OGO^T\, |s\rangle \leq 1 \qquad \forall\ \{|s\rangle \,|\, \langle s|\, H\, |s\rangle = 1\}\,. \end{aligned}$$

From Definition 84 one can interpret the last step as stating that along all directions $|s\rangle$, the ellipsoid corresponding to H is inside the ellipsoid corresponding to OGO^T. If H and G are fixed, then finding the orthogonal matrix O can be seen as rotating the G ellipsoid in such a way that the H ellipsoid always remains inside.

The curvature of the ellipsoid at a given point may be given by the Weingarten Map, as we saw in Section 5 and Appendix F. In practice, it is easier to first evaluate the Reverse Weingarten Map, which we denote by W, and then take its pseudo-inverse, W^{-1}, to obtain the Weingarten Map itself. Suppose the ellipsoid under consideration is associated with $G \geq 0$. If G and G^{-1} are known then one can find analytic expressions for W and W^{-1} (see Appendix F), which are summarized in the following definition.

Definition 85 (Normal Function, Weingarten Map, Reverse Weingarten Map, Orthogonal Component)**.** Given a matrix $G \geq 0$, its pseudo-inverse G^{-1} and a vector $|v\rangle$, such that $G\,|v\rangle \neq 0$, we define the following functions. We use $\left\langle G^j \right\rangle = \left\langle v|G^j|v \right\rangle$ and $\mathcal{N}(|v\rangle) = |v\rangle / \langle v|v\rangle$ to denote normalization.

- The Normal Function from $G, |v\rangle$ to a vector $|u\rangle$ is defined as

$$|u(G, |v\rangle)\rangle := \frac{G\,|v\rangle}{\langle G^2\rangle}.$$

- The Weingarten Map from $G, |v\rangle$ to a matrix W^{-1} is defined as

$$W^{-1}(G, |v\rangle) := \sqrt{\frac{\langle G\rangle}{\langle G^2\rangle}} \left(G + \frac{\langle G^3\rangle}{\langle G^2\rangle^2} G\,|v\rangle\,\langle v|\,G - \frac{1}{\langle G^2\rangle}\left(G\,|v\rangle\,\langle v|\,G^2 + G^2\,|v\rangle\,\langle v|\,G\right)\right).$$

- The Reverse Weingarten Map from $G, G^{-1}, |v\rangle$ to W is defined as

$$W(G, G^{-1}, |v\rangle) := \sqrt{\frac{\langle G^2\rangle}{\langle G\rangle}} \left(G^{-1} - \frac{|v\rangle\,\langle v|}{\langle G\rangle} \right).$$

- The Orthogonal Component from $G, |v\rangle$ to $|e\rangle$ is defined as

$$|e(G, |v\rangle)\rangle := \mathcal{N}\left[|v\rangle - \langle u|v\rangle\,|u\rangle\right],$$

  where $|u\rangle = |u(G, |v\rangle)\rangle$.

- The Orthogonal Component from $|v'\rangle, |v\rangle$ to $|e\rangle$ is defined as

$$|e(|v'\rangle, |v\rangle)\rangle := \mathcal{N}[|v\rangle - \langle v'|v\rangle\,|v'\rangle].$$

The following lemma and remark provide properties of the Weingarten map that are relevant for our analysis. Lemma 86 states that evaluating the Weingarten map at a given point of a rotated ellipsoid is the same as evaluating it for the non-rotated ellipsoid and then rotating it, and Remark 87 shows that W necessarily has one less rank compared to G.

Lemma 86. *Let $G \geq 0$ be an $n \times n$ rank k matrix and Q be an isometry from the non-trivial k-dimensional subspace of G to an arbitrary k-dimensional subspace. Then*

$$W^{-1}(QGQ^T, Q\left|v\right\rangle) = QW^{-1}(G, \left|v\right\rangle)Q^T.$$

Remark 87. With reference to Definition 85, let $W = W(G, G^{-1}, \left|v\right\rangle)$, $W^{-1} = W^{-1}(G, \left|v\right\rangle)$ and $\left|u\right\rangle = \left|u(G, \left|v\right\rangle)\right\rangle$. Then $WG\left|v\right\rangle = W\left|u\right\rangle = 0$ and $W^{-1}G\left|v\right\rangle = W^{-1}\left|u\right\rangle = 0$. This may be seen by a direct computation or by inspection of the proofs of Lemma 156 and Lemma 154.

6.1 Solution to the f_0-assignment

Recall that a valid function is the same as an EBRM function (see Corollary 144). Given a valid function $t = \sum_i p_{h_i} [\![x_{h_i}]\!] - \sum_i p_{g_i} [\![x_{g_i}]\!]$, it is easy to re-write the matrices that appear in the EBRM description into a form which satisfies[26] $H \geq OGO^T$, $O\left|v\right\rangle = \left|w\right\rangle$, where $\left|v\right\rangle \doteq (\sqrt{p_{g_1}}, \sqrt{p_{g_2}} \dots)$ and $\left|w\right\rangle \doteq (\sqrt{p_{h_1}}, \sqrt{p_{h_2}} \dots)$ while $H = \operatorname{diag}(x_{h_1}, x_{h_2} \dots)$ and $G = \operatorname{diag}(x_{g_1}, x_{g_2} \dots)$. As we saw in Section 4, it suffices to restrict to monomial assignments (see Definition 32), i.e. assignments of the form

$$t = \sum_{i=1}^{n} \frac{-(-x_i)^k}{\prod_{j \neq i}(x_j - x_i)}$$

for $0 \leq x_1 < x_2 \dots < x_n$ with $0 \leq k \leq n-2$, to convert Mochon's games into explicit protocols.

Recall that for f_0-assignments, $\left\langle x^k \right\rangle = 0$ for all $0 \leq k \leq n-2$ and $\left\langle x^{n-1} \right\rangle \neq 0$ (see Lemma 33). The ellipsoids H and OGO^T touch along the vector $\left|w\right\rangle$ if $\left\langle w\right| H \left|w\right\rangle = \left\langle w\right| OGO^T \left|w\right\rangle = \left\langle v\right| G \left|v\right\rangle$. This is the case here, since $\left\langle w\right| H \left|w\right\rangle - \left\langle v\right| G \left|v\right\rangle = \left\langle x \right\rangle = 0$. This, in turn, means that the normal along $\left|v\right\rangle$ of the G ellipsoid, $\left|u(G, \left|v\right\rangle)\right\rangle$ (see Definition 85) must be mapped to the normal along $\left|w\right\rangle$ of the H ellipsoid, $\left|u(H, \left|w\right\rangle)\right\rangle$, i.e. O must have the form $O = \left|u(H, \left|w\right\rangle)\right\rangle \left\langle u(G, \left|v\right\rangle)\right| + Q$ where Q represents the action of O from the space orthogonal to $\left|u(G, \left|v\right\rangle)\right\rangle$ onto the space orthogonal to $\left|u(H, \left|w\right\rangle)\right\rangle$. Furthermore, since $H \geq OGO^T$ we must have (see Definition 85)

$$W(H, H^{-1}, \left|w\right\rangle) \geq QW(G, G^{-1}, \left|v\right\rangle)Q^T,$$

i.e., the curvature of the H ellipsoid at $\left|w\right\rangle$—which is given by $W(H, H^{-1}, \left|w\right\rangle)$—must be greater than that of the OGO^T ellipsoid along $\left|v\right\rangle$[27]—which is given by $QW(G, G^{-1}, \left|v\right\rangle)Q^T$. The component of $\left|v\right\rangle$ along $\left|u(G, \left|v\right\rangle)\right\rangle$ is mapped to $\left|u(H, \left|w\right\rangle)\right\rangle$ under the action of O, which, so far, has only been partially specified. The remaining component is $\left|e(G, \left|v\right\rangle)\right\rangle$ and analogously for $\left|w\right\rangle$, the remaining component is $\left|e(H, \left|w\right\rangle)\right\rangle$. Using these Ws and $\left|e\right\rangle$s as the new matrices and vectors, it turns out that one can apply this argument repeatedly (when the number of points, n, is even) to completely specify O. Clearly, though, this notation rapidly becomes complicated, therefore we introduce the so-called *Matrix Instance* and the *Weingarten Iteration Map*. The former is similar to the one introduced in Section 5 albeit with the difference that here we use isometries and slightly different nomenclature.

Definition 88 (Matrix Instance and its properties)**.** Let

- $n \geq k$ be positive integers,
- $\mathcal{H}^{\bar{k}}$ and $\mathcal{G}^{\bar{k}}$ be two k dimensional Hilbert spaces,

[26] See the discussion after Theorem 31; we suppressed the details about the dimensions and the spectra of matrices.

[27] We used the fact that $W(G, G^{-1}, \left|v\right\rangle)\left|u(G, \left|v\right\rangle)\right\rangle = 0$

- $H \geq 0$, $G \geq 0$ be $n \times n$ non-zero matrices of rank at most k, such that H has support only on $\mathcal{H}^{\bar{k}}$ and analogously G has support only on $\mathcal{G}^{\bar{k}}$,
- $|w\rangle \in \mathcal{H}^{\bar{k}}$ and $|v\rangle \in \mathcal{G}^{\bar{k}}$ be vectors of equal norm, $|u_h\rangle \in \mathcal{H}^{\bar{k}}$ and $|u_g\rangle \in \mathcal{G}^{\bar{k}}$ be vectors with unit norm,

A *matrix instance* is defined to be the tuple $\underline{\mathrm{X}}^{\bar{k}} := (H, G, |w\rangle, |v\rangle)$ and the set of all matrix instances (of $n \times n$ dimensions) is denoted by $\mathbb{X}^n$.

We define the following properties of a matrix instance.

- Let $Q : \mathcal{G}^{\bar{k}} \to \mathcal{H}^{\bar{k}}$ be an isometry, i.e. $Q^T Q = \mathbb{I}_g$ and $QQ^T = \mathbb{I}_h$ where $\mathbb{I}_h$ is the identity in $\mathcal{H}^{\bar{k}}$ and similarly $\mathbb{I}_g$ is the identity in $\mathcal{G}^{\bar{k}}$. We say that Q *solves* the *matrix instance* $\underline{\mathrm{X}}^{\bar{k}}$ if and only if
$$H \geq QGQ^T \qquad \text{and} \qquad Q|v\rangle = |w\rangle .$$
- We say that $\underline{\mathrm{X}}^{\bar{k}}$ satisfies the *contact condition* if and only if $\langle w| H |w\rangle = \langle v| G |v\rangle$.
- We say that $\underline{\mathrm{X}}^{\bar{k}}$ satisfies the *component condition* if and only if $\langle w| H^2 |w\rangle = \langle v| G^2 |v\rangle$.

Definition 89 (Weingarten Iteration Map)**.** Consider a matrix instance $\underline{\mathrm{X}}^{\bar{k}} =: \left(H^{\bar{k}}, G^{\bar{k}}, \left|w^{\bar{k}}\right\rangle, \left|v^{\bar{k}}\right\rangle\right)$ and let (see Definition 85)
$$\left|v^{\overline{k-1}}\right\rangle := \left|e\left(G^{\bar{k}}, \left|v^{\bar{k}}\right\rangle\right)\right\rangle, \qquad \left|w^{\overline{k-1}}\right\rangle := \left|e\left(H^{\bar{k}}, \left|w^{\bar{k}}\right\rangle\right)\right\rangle,$$
$$G^{\overline{k-1}} := W^{\dashv}\left(G^{\bar{k}}, \left|v^{\bar{k}}\right\rangle\right), \qquad H^{\overline{k-1}} := W^{\dashv}\left(H^{\bar{k}}, \left|w^{\bar{k}}\right\rangle\right).$$
Then we define the *Weingarten Iteration Map* $\mathcal{W} : \mathbb{X}^n \to \mathbb{X}^n$ by its action
$$\underline{\mathrm{X}}^{\bar{k}} \mapsto \left(H^{\overline{k-1}}, G^{\overline{k-1}}, \left|w^{\overline{k-1}}\right\rangle, \left|v^{\overline{k-1}}\right\rangle\right) =: \underline{\mathrm{X}}^{\overline{k-1}}.$$

So far, we only relied on the properties of the f_0-assignment for establishing that the *contact condition* holds, i.e. $\langle w| H |w\rangle = \langle v| G |v\rangle$. The rest of the argument about was actually quite general. We state it in terms of matrix instances and prove it below.

Lemma 90. *Consider a matrix instance $\underline{\mathrm{X}}^{\bar{k}} := (H, G, |w\rangle, |v\rangle)$ which satisfies both the contact and the component condition. Let $\left|u_h^{\bar{k}}\right\rangle := |u(H, |w\rangle)\rangle$, $\left|u_g^{\bar{k}}\right\rangle := |u(G, |v\rangle)\rangle$ and $\underline{\mathrm{X}}^{\overline{k-1}} := \mathcal{W}(\underline{\mathrm{X}}^{\bar{k}})$ (see Definition 95). If $Q^{\bar{k}}$ solves the matrix instance $\underline{\mathrm{X}}^{\bar{k}}$ then*
$$Q^{\bar{k}} = \left|u_h^{\bar{k}}\right\rangle \left\langle u_g^{\bar{k}}\right| + Q^{\overline{k-1}}, \tag{23}$$
where $Q^{\overline{k-1}}$ solves the matrix instance $\underline{\mathrm{X}}^{\overline{k-1}}$.

Proof. Let $\left(H^{\bar{k}}, G^{\bar{k}}, \left|w^{\bar{k}}\right\rangle, \left|v^{\bar{k}}\right\rangle\right) := \underline{\mathrm{X}}^{\bar{k}}$ and $\left(H^{\overline{k-1}}, G^{\overline{k-1}}, \left|w^{\overline{k-1}}\right\rangle, \left|v^{\overline{k-1}}\right\rangle\right) := \underline{\mathrm{X}}^{\overline{k-1}}$. The matrix inequality $H^{\bar{k}} \geq Q^{\bar{k}} G^{\bar{k}} \left(Q^{\bar{k}}\right)^T$ describes the containment of the ellipsoid corresponding to $H^{\bar{k}}$ inside the ellipsoid corresponding to $Q^{\bar{k}} G^{\bar{k}} \left(Q^{\bar{k}}\right)^T$. The two ellipsoids touch along the $\left|w^{\bar{k}}\right\rangle$ direction if and only if
$$\left\langle w^{\bar{k}}\right| H^{\bar{k}} \left|w^{\bar{k}}\right\rangle = \left\langle w^{\bar{k}}\right| Q^{\bar{k}} G^{\bar{k}} \left(Q^{\bar{k}}\right)^T \left|w^{\bar{k}}\right\rangle = \left\langle v^{\bar{k}}\right| G^{\bar{k}} \left|v^{\bar{k}}\right\rangle,$$

where the last step follows from noting $Q^{\bar{k}}\left|v^{\bar{k}}\right\rangle = \left|w^{\bar{k}}\right\rangle$ and the fact that $Q^{\bar{k}}$ is an isometry. This is precisely the contact condition. The component condition ensures that the components of the probability vectors along their respective normals are the same, viz. $\left\langle w^{\bar{k}}|u_h^{\bar{k}}\right\rangle = \left\langle v^{\bar{k}}|u_g^{\bar{k}}\right\rangle$ (see Lemma 153). From this we can deduce the following three necessary conditions.

First, that Equation (23) holds. Indeed, the normal along $\left|w^{\bar{k}}\right\rangle$ (see Lemma 153) of the ellipsoid $H^{\bar{k}}$ and that of the ellipsoid $Q^{\bar{k}}G^{\bar{k}}Q^{\bar{k}T}$ must be the same. This in turn means that $Q^{\bar{k}}$ must map the normal $\left|u_g^{\bar{k}}\right\rangle$ along $\left|v^{\bar{k}}\right\rangle$ of the ellipsoid $G^{\bar{k}}$ to the normal $\left|u_h^{\bar{k}}\right\rangle$ along $|w\rangle$ of the ellipsoid $H^{\bar{k}}$, viz. $\left|u_g^{\bar{k}}\right\rangle := \left|u\left(G^{\bar{k}}, \left|v^{\bar{k}}\right\rangle\right)\right\rangle \mapsto \left|u_h^{\bar{k}}\right\rangle := \left|u\left(H^{\bar{k}}, \left|w^{\bar{k}}\right\rangle\right)\right\rangle$ (see Definition 85). Consequently,

$$Q^{\bar{k}} = \left|u_h^{\bar{k}}\right\rangle\left\langle u_g^{\bar{k}}\right| + Q^{\overline{k-1}}, \tag{24}$$

where $Q^{\overline{k-1}} : \mathcal{G}^{\overline{k-1}} \to \mathcal{H}^{\overline{k-1}}$ is an isometry as the action on the normals is completely determined.

Second, the curvature along $\left|w^{\bar{k}}\right\rangle$ of the ellipsoid $H^{\bar{k}}$ must be greater than that of the ellipsoid $Q^{\bar{k}}G^{\bar{k}}\left(Q^{\bar{k}}\right)^T$ along the same direction, viz.

$$\begin{aligned}
H^{\overline{k-1}} = W^{\dashv}\left(H^{\bar{k}}, \left|w^{\bar{k}}\right\rangle\right) \geq W^{\dashv}&\left(Q^{\bar{k}}G^{\bar{k}}\left(Q^{\bar{k}}\right)^T, Q^{\bar{k}}\left|v^{\bar{k}}\right\rangle\right) \\
&= Q^{\bar{k}}W^{\dashv}\left(G^{\bar{k}}, \left|v^{\bar{k}}\right\rangle\right)\left(Q^{\bar{k}}\right)^T \\
&= Q^{\overline{k-1}}\underbrace{W^{\dashv}\left(G^{\bar{k}}, \left|v^{\bar{k}}\right\rangle\right)}_{=G^{\overline{k-1}}}\left(Q^{\overline{k-1}}\right)^T \qquad \because \quad W^{\dashv}\left(G^{\bar{k}}, \left|v^{\bar{k}}\right\rangle\right)\left|u_g^{\bar{k}}\right\rangle = 0; \\
&= Q^{\overline{k-1}}G^{\overline{k-1}}\left(Q^{\overline{k-1}}\right)^T. \qquad \text{see } \textit{Remark } 87
\end{aligned}$$

Finally, since $Q^{\bar{k}}\left|v^{\bar{k}}\right\rangle = \left|w^{\bar{k}}\right\rangle$ by acting with a projector on both sides, we obtain $\left(\mathbb{I}_h^{\bar{k}} - \left|u_h^{\bar{k}}\right\rangle\left\langle u_h^{\bar{k}}\right|\right)Q^{\bar{k}}\left|v^{\bar{k}}\right\rangle = \left(\mathbb{I}_h^{\bar{k}} - \left|u_h^{\bar{k}}\right\rangle\left\langle u_h^{\bar{k}}\right|\right)\left|w^{\bar{k}}\right\rangle$. Using $\left(\mathbb{I}_h^{\bar{k}} - \left|u_h^{\bar{k}}\right\rangle\left\langle u_h^{\bar{k}}\right|\right)Q^{\bar{k}} = \left(\mathbb{I}_h^{\bar{k}} - \left|u_h^{\bar{k}}\right\rangle\left\langle u_h^{\bar{k}}\right|\right)Q^{\bar{k}}\left(\mathbb{I}_g^{\bar{k}} - \left|u_g^{\bar{k}}\right\rangle\left\langle u_g^{\bar{k}}\right|\right)$ in the LHS (follows from Equation (24)) and Definition 85 for $|e(.,.)\rangle$, one obtains the equation $Q^{\overline{k-1}}\left|v^{\overline{k-1}}\right\rangle = \left|w^{\overline{k-1}}\right\rangle$. These show that $Q^{\overline{k-1}}$ indeed solves $\underline{X}^{\overline{k-1}}$. □

6.1.1 The balanced case

Proposition 91 (The balanced f_0-solution). *Let $t = h - g = \sum_{i=1}^{2n} p_i [\![x_i]\!]$ be an f_0-assignment over the real coordinates $0 \leq x_1 < x_2 \cdots < x_{2n}$. Let $h = \sum_{i=1}^{n} p_{h_i}[\![x_{h_i}]\!]$, $g = \sum_{i=1}^{n} p_{g_i}[\![x_{g_i}]\!]$ where $p_{h_i}, p_{g_i} > 0$, and $\{x_{h_i}\}$ and $\{x_{g_i}\}$ are all distinct. Consider the matrix instance $\underline{X}$:=$(X_h, X_g, |w\rangle, |v\rangle)$
where $X_h \doteq diag(x_{h_1}, x_{h_2}\ldots x_{h_n})$, $X_g \doteq diag(x_{g_1}, x_{g_2}\ldots x_{g_n})$,
$|w\rangle \doteq (\sqrt{p_{h_1}}, \sqrt{p_{h_2}}\ldots\sqrt{p_{h_n}})^T$, $|v\rangle \doteq \left(\sqrt{p_{g_1}}, \sqrt{p_{g_2}}\ldots\sqrt{p_{g_n}}\right)^T$. The orthogonal matrix*

$$O = \sum_{k=1}^{n}\left|u_h^{\bar{k}}\right\rangle\left\langle u_g^{\bar{k}}\right|$$

solves $\underline{X} =: \underline{X}^{\bar{n}}$ (see Definition 88) where the Weingarten Iteration Map (see Definition 95) is used to evaluate $\underline{X}^{\overline{k-1}} = \mathcal{W}(\underline{X}^{\bar{k}})$. This, in turn, is used to obtain $\left|u_h^{\bar{k}}\right\rangle = \left|u(H^{\bar{k}}, \left|w^{\bar{k}}\right\rangle)\right\rangle$ and $\left|u_g^{\bar{k}}\right\rangle = \left|u(G^{\bar{k}}, \left|v^{\bar{k}}\right\rangle)\right\rangle$ for all k starting from $k = n$, with $\left(H^{\bar{k}}, G^{\bar{k}}, \left|w^{\bar{k}}\right\rangle, \left|v^{\bar{k}}\right\rangle\right) := \underline{X}^{\bar{k}}$.

To prove Proposition 91, we use the following lemma which follows from Lemma 162 and Lemma 166 proved in Appendix I.

Lemma 92 (Up Contact/Component Lemma)**.** *Consider the matrix instance* $\underline{X}^{\bar{n}} := (H^{\bar{n}}, G^{\bar{n}}, \left|w^{\bar{n}}\right\rangle, \left|v^{\bar{n}}\right\rangle)$*. Suppose the Weingarten Iteration Map (see Definition 95) is applied l times to obtain*

$$\underline{X}^{\overline{n-l}} := \left(H^{\overline{n-l}}, G^{\overline{n-l}}, \left|w^{\overline{n-l}}\right\rangle, \left|v^{\overline{n-l}}\right\rangle\right).$$

Then,

$$\left\langle v^{\overline{n-l}}\right| \left(G^{\overline{n-l}}\right)^m \left|v^{\overline{n-l}}\right\rangle = r\left(\left\langle (G^{\bar{n}})^{m-1}\right\rangle, \left\langle (G^{\bar{n}})^{m}\right\rangle \dots, \left\langle (G^{\bar{n}})^{2l+m}\right\rangle\right),$$

where $m \geq 1$ and r is a multi-variate function which does not have an implicit dependence on $\left\langle (G^{\bar{n}})^i\right\rangle := \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^i \left|v^{\bar{n}}\right\rangle$ for any i. Analogously, for H and $|w\rangle$ we have

$$\left\langle w^{\overline{n-l}}\right| \left(H^{\overline{n-l}}\right)^m \left|w^{\overline{n-l}}\right\rangle = r\left(\left\langle (H^{\bar{n}})^{m-1}\right\rangle, \left\langle (H^{\bar{n}})^{m}\right\rangle \dots, \left\langle (H^{\bar{n}})^{2l+m}\right\rangle\right).$$

This lemma relates the contact condition of the l-th matrix instance, i.e. the one obtained after applying the Weingarten Iteration Map l times, to the expectation values associated with the first matrix instance. These expectation values, for the f_0-solution, are $\left\langle x^k\right\rangle = \left\langle (H^{\bar{n}})^k\right\rangle - \left\langle (G^{\bar{n}})^k\right\rangle = 0$ for all $0 \leq k \leq n-2$, which means that the contact condition also holds for the l-th matrix instance, thereby allowing one to repeatedly use Lemma 90 to determine the solution, O.

Proof of Proposition 91. We have already done most of the work by proving Lemma 90 and Lemma 92. To use the Weingarten iteration once for the matrix instance $\underline{X} =: \underline{X}^{\bar{n}} =: (H^{\bar{n}}, G^{\bar{n}}, \left|w^{\bar{n}}\right\rangle, \left|v^{\bar{n}}\right\rangle)$, we must show that $\underline{X}^{\bar{n}}$ satisfies the contact condition (see Definition 93 and Lemma 90), viz.

$$\begin{aligned} &\left\langle w^{\bar{n}}\right| H^{\bar{n}} \left|w^{\bar{n}}\right\rangle - \left\langle v^{\bar{n}}\right| G^{\bar{n}} \left|v^{\bar{n}}\right\rangle = \left\langle H^{\bar{n}}\right\rangle - \left\langle G^{\bar{n}}\right\rangle \\ &= \sum_{i=1}^{n} p_{h_i} x_{h_i} - \sum_{i=1}^{n} p_{g_i} x_{g_i} = \sum_{i=1}^{n} p_i x_i = \langle x\rangle = 0, \end{aligned}$$

which holds due to Lemma 33. After iterating for l steps, suppose the matrix instance one obtains is $\underline{X}^{\overline{n-l}}$. To check if another Weingarten iteration is possible, we must check if the contact condition holds, i.e. if

$$\begin{aligned} &\left\langle w^{\overline{n-l}}\right| H^{\overline{n-l}} \left|w^{\overline{n-l}}\right\rangle - \left\langle v^{\overline{n-l}}\right| G^{\overline{n-l}} \left|v^{\overline{n-l}}\right\rangle = \\ &r\left(\left\langle (H^{\bar{n}})^1\right\rangle, \left\langle (H^{\bar{n}})^2\right\rangle \dots, \left\langle (H^{\bar{n}})^{2l+1}\right\rangle\right) - r\left(\left\langle (G^{\bar{n}})^1\right\rangle, \left\langle (G^{\bar{n}})^2\right\rangle \dots, \left\langle (G^{\bar{n}})^{2l+1}\right\rangle\right) \end{aligned}$$

vanishes. We used Lemma 92 with $m = 1$ to obtain the RHS. Note that

$$\left\langle (H^{\bar{n}})^k\right\rangle - \left\langle (G^{\bar{n}})^k\right\rangle = \left\langle x^k\right\rangle. \tag{25}$$

If $2l+1 \leq 2n-2$ then from Lemma 33 it follows that both terms become identical and hence the difference indeed vanishes.[28] A similar argument can be used to obtain the condition $2l+2 \leq 2n-2$ which corresponds to the component condition (see Definition 93). Assuming $O =: O^{\bar{n}}$ solves $\underline{X}^{\bar{n}}$, until $l = n-2$, one can iterate—using the Weingarten Iteration Map, $\mathcal{W}$, and the Normal Function (see Definition 85)— to obtain $\left|u_h^{\bar{n}}\right\rangle, \left|u_h^{\overline{n-1}}\right\rangle, \dots, \left|u_h^{\overline{n-l}}\right\rangle, \dots, \left|u_h^{\bar{1}}\right\rangle$ and $\left|u_g^{\bar{n}}\right\rangle, \left|u_g^{\overline{n-1}}\right\rangle, \dots, \left|u_g^{\overline{n-l}}\right\rangle, \dots, \left|u_g^{\bar{1}}\right\rangle$ which completely determine $O^{\bar{n}}$.

It only remains to prove that there exists an O which solves the matrix instance $\underline{X}^{\bar{n}}$. We outline this proof in Appendix H. □

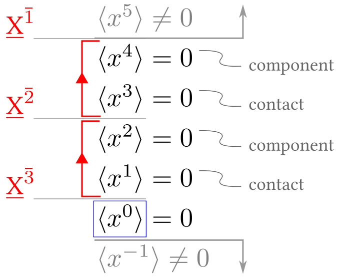


Figure 16: Power diagram for a balanced f_0-assignment with $2n = 6$ points. Starting upwards from $\langle x^0\rangle$, two iterations are completed before encountering the instance where the contact condition does not hold and the normals do not match.

It helps to represent the main argument succinctly using Figure 16. We start right above $\langle x^0\rangle$ with the matrix instance $\underline{X}^{\bar{n}}$. Set $n = 3$ for concreteness. The contact condition at this step corresponds to $\langle x^1\rangle = 0$, which is true as the power is less than or equal to $2n - 2$ (here $2n - 2 = 4$; see Lemma 33). We can thus apply the Weingarten iteration (see Definition 95) which is indicated by the arrow[29] from $\langle x^1\rangle$ to $\langle x^2\rangle$. This yields $\underline{X}^{\overline{n-1}}$ and we can proceed with checking if $\langle x^3\rangle = 0$, which is true as the power is ≤ 4, and therefore we can again iterate to obtain $\underline{X}^{\overline{n-2}}$, which in this illustration is $\underline{X}^{\bar{1}}$. At this point, we have solved the problem as we can evaluate $\left|u_h^{\bar{3}}\right\rangle, \left|u_h^{\bar{2}}\right\rangle, \left|u_h^{\bar{1}}\right\rangle$ and $\left|u_g^{\bar{3}}\right\rangle, \left|u_g^{\bar{2}}\right\rangle, \left|u_g^{\bar{1}}\right\rangle$ form $\underline{X}^{\bar{3}}, \underline{X}^{\bar{2}}, \underline{X}^{\bar{1}}$ respectively to write $O = \sum_{k=1}^{3} \left|u_h^{\bar{k}}\right\rangle \left\langle u_g^{\bar{k}}\right|$. Note that having an even number of total points, $x_1 < x_2 \cdots < x_{2n}$, ensures that there is a proper *alignment* in the diagram in the sense that both the contact condition for $\underline{X}^{\bar{2}}$, $\langle x^3\rangle = 0$, and the component condition, $\langle x^4\rangle = 0$, hold. As we saw in the proof, the contact condition essentially requires that the component of $\left|w^{\bar{k}}\right\rangle$ along $\left|u_h^{\bar{k}}\right\rangle$ is the same as the component of $\left|v^{\bar{k}}\right\rangle$ along $\left|u_g^{\bar{k}}\right\rangle$. If this does not hold, then we do not have $O\left|v\right\rangle = \left|w\right\rangle$, which not only means that we don't have a solution, but also that our approach, which was based on that assumption, fails.

6.1.2 The unbalanced case

In the unbalanced case—where the total number of points is odd—the component condition ceases to hold at the last step, while the contact condition still holds. This means that we can no longer apply the Weingarten Iteration Map as the premise for Lemma 90 is not true. We have already encountered this situation in Section 5 and the wiggle-v method we used there also works here. Let us recall that argument using our present notation. So far, we reasoned that if one ellipsoid is contained inside another, $H \geq QGQ^T$, and they touch along a vector, $\left|w\right\rangle$, then the normal $\left|u_g\right\rangle$ of the G ellipsoid along $\left|v\right\rangle = Q^T\left|w\right\rangle$ must be mapped to the normal $\left|u_h\right\rangle$ of the H ellipsoid along $\left|w\right\rangle$, by the isometry Q. This analysis requires that the normal is well-defined which is true if the matrices have finite spectra. However, as we pointed out in Appendix A some valid functions can not be expressed by matrices (EBM) having a finite spectrum and the merge move was an example. To visualize this, think of the QGQ^T ellipsoid as a circle, the H ellipsoid as a line and the $\left|w\right\rangle$ vector pointing along this line (see Figure 17; image on the right). The normal to the H ellipsoid along the point of contact can have an arbitrary component along the vector perpendicular to $\left|w\right\rangle$. If the line is seen as an approximation to a squeezed circle, then it is clear that a very small *wiggle* in $\left|w\right\rangle$ can significantly affect the normal. As we have already seen this more precisely in Section 5, we content ourselves with

[28] The number of points here is $2n$; in the Lemma they are denoted by n.

[29] It, strictly speaking, goes from below $\langle x^1\rangle$ to above $\langle x^2\rangle$; the idea was just to indicate the inclusion of the two terms for the matrix instance $\underline{X}^{\bar{3}}$.

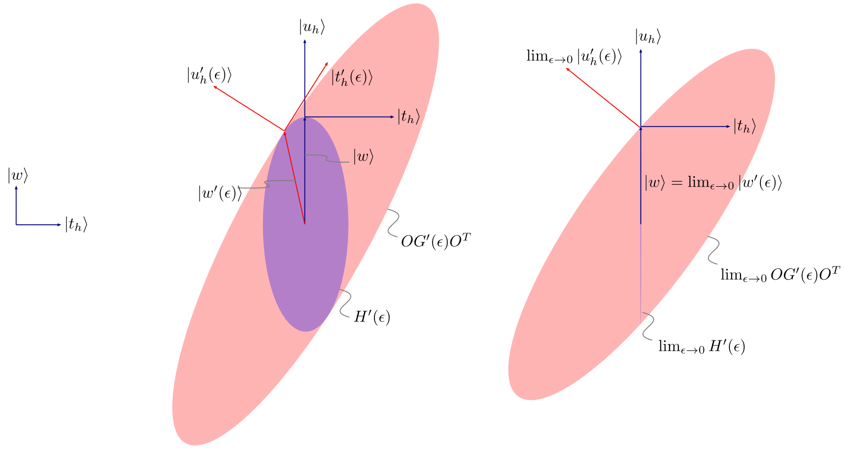


Figure 17: The infinite curvature case, where the wiggle-v method is applied.

the observation that there is a freedom in the choice of the normal. We can fix this freedom by requiring that the component condition is satisfied. Denote the direction of infinite curvature (in our "circle-line" example it was the vector perpendicular to $|w\rangle$), by $|t_h\rangle$. The freedom in correcting the normal can be expressed by parametrizing it as

$$\left|u'_h\right\rangle := \cos\theta\,|u_h\rangle + \sin\theta\,|t_h\rangle \tag{26}$$

where $|u_h\rangle := |u(H, |w\rangle)\rangle$. Enforcing the component condition, $\langle w|u_h\rangle = \left\langle v|u_g\right\rangle$, fixes θ, the parameter which completely specifies the corrected normal, $\left|u'_h\right\rangle$. One can now apply Lemma 90 with $\left|u'_h\right\rangle$ instead of $|u_h\rangle$. To formalize this procedure, we define an object, *Extended Matrix Instance*, which is designed to hold certain additional quantities derived from the initial matrix instance, i.e. normals and inverse of the matrices.

Definition 93 (Extended Matrix Instance and its properties)**.** Let

- $n \geq k$ be positive integers,
- $\mathcal{H}^{\bar{k}}$ and $\mathcal{G}^{\bar{k}}$ be two k dimensional Hilbert spaces,
- S_h be the set of $n \times n$ non-zero matrices of rank at most k with support only on $\mathcal{H}^{\bar{k}}$, i.e.

  $$S_h := \{n \times n \text{ matrices } M : M \geq 0 \text{ has rank at most } k, \text{ and support only on } \mathcal{H}^{\bar{k}}\}$$

  and analogously,

  $$S_g := \{n \times n \text{ matrices } M : M \geq 0 \text{ has rank at most } k, \text{ and support only on } \mathcal{G}^{\bar{k}}\}$$

- $H \in S_h$, $G \in S_g$, $H_{\text{inv}} \in S_h \cup \{[.]\}$, $G_{\text{inv}} \in S_g \cup \{[.]\}$
- $|w\rangle \in \mathcal{H}^{\bar{k}}$ and $|v\rangle \in \mathcal{G}^{\bar{k}}$ be vectors of equal norm,
- $|u_h\rangle \in \{|u\rangle \in \mathcal{H}^{\bar{k}} : \langle u|u\rangle = 1\} \cup \{|.\rangle\}$ and $\left|u_g\right\rangle \in \{|u\rangle \in \mathcal{G}^{\bar{k}} : \langle u|u\rangle = 1\} \cup \{|.\rangle\}$.

A *matrix instance* is defined as the tuple $\underline{\mathrm{X}}^{\bar{k}} = (H, G, |w\rangle, |v\rangle)$. An *extended matrix instance* is defined as the tuple[30] $\underline{\mathrm{M}}^{\bar{k}} := \underline{\mathrm{X}}^{\bar{k}} \oplus \left(H_{\text{inv}}, G_{\text{inv}}, |u_h\rangle, \left|u_g\right\rangle\right)$ for the $\underline{\mathrm{X}}^{\bar{k}}$ matrix instance.

[30] H_{inv} is supposed to explicitly hold the expression for H^{-1} in terms H and its powers. Analogously for G_{inv}.

The extended matrix instance is *partially specified* if H_{inv} or G_{inv} equal $[.]$ or if $|u_h\rangle$ or $|u_g\rangle$ equal $|.\rangle$. We say that an extended matrix instance is *completely specified* if it is not partially specified.

The set of all matrix instances of $n \times n$ dimensions is denoted by $\mathbb{X}^n$ and the set of all extended matrix instances is denoted by $\mathbb{M}^n$. We now define some of their properties.

- Let $Q : \mathcal{G}^{\bar{k}} \to \mathcal{H}^{\bar{k}}$ be an isometry, i.e. $Q^T Q = \mathbb{I}_g$ and $QQ^T = \mathbb{I}_h$ where $\mathbb{I}_h$ is the identity in $\mathcal{H}^{\bar{k}}$ and similarly $\mathbb{I}_g$ is the identity in $\mathcal{G}^{\bar{k}}$. We say that Q *solves* the *matrix instance* $\underline{\mathrm{X}}^{\bar{k}}$ if and only if
$$H \geq QGQ^T \qquad \text{and} \qquad Q\,|v\rangle = |w\rangle\,.$$
Similarly we say that Q *resolves* (reverse solves) the matrix instance if and only if
$$H \leq QGQ^T \qquad \text{and} \qquad Q\,|v\rangle = |w\rangle\,.$$

- We say that $\underline{\mathrm{X}}^{\bar{k}}$ satisfies the *contact condition* if and only if $\langle w|\,H\,|w\rangle = \langle v|\,G\,|v\rangle$. Similarly for $\underline{\mathrm{M}}^{\bar{k}}$.

- We say that $\underline{\mathrm{X}}^{\bar{k}}$ satisfies the *component condition* if and only if $\langle w|\,H^2\,|w\rangle = \langle v|\,G^2\,|v\rangle$. Similarly for $\underline{\mathrm{M}}^{\bar{k}}$.

- We say that $\underline{\mathrm{X}}^{\bar{k}}$ has *wiggle-w room* (ϵ) *along* $|t_h\rangle$ if and only if H has an eigenvector $|t_h\rangle$ with eigenvalue $1/\epsilon$ which has no overlap with $|w\rangle$, viz. $H\,|t_h\rangle = \epsilon^{-1}\,|t_h\rangle$ and $\langle w|t_h\rangle = 0$. Similarly, we say that $\underline{\mathrm{X}}^{\bar{k}}$ has *wiggle-v room* (ϵ) along $|t_g\rangle$ if and only if G has an eigenvector $|t_g\rangle$ with eigenvalue $1/\epsilon$ which has no overlap with $|v\rangle$, viz. $G\,|t_g\rangle = \epsilon^{-1}\,|t_g\rangle$ and $\langle v|t_g\rangle = 0$. For brevity, we say $\underline{\mathrm{X}}^{\bar{k}}$ has *wiggle-w/v room.*

Below we define the *Normal Initialization Map* which formalizes the evaluation of the normals to initialize a partially specified extended matrix instance.We also revisit the *Weingarten Iteration Map* by extending Definition 89 to include extended matrix instances as well. This map now takes a rank k extended matrix instance and constructs a rank $k-1$ extended matrix instance, which however is only partially specified, as the normal vectors are left unspecified. In order to completely specify this $k-1$ rank extended matrix instance we need to use the above two maps together.

Definition 94 (Normal Initialization Map)**.** Given a matrix instance $\underline{\mathrm{X}}^{\bar{k}} =: (H, G, |w\rangle, |v\rangle)$, $H^{\dashv}$, and $G^{\dashv}$ the *normal initialization map* $\mathcal{U} : \mathbb{X}^n \to \mathbb{M}^n$ (see Definition 93) is defined by its action

$$\underline{\mathrm{X}}^{\bar{k}} \mapsto \underline{\mathrm{X}}^{\bar{k}} \oplus (H^{\dashv}, G^{\dashv}, |u(H, |w\rangle)\rangle\,, |u(G, |v\rangle)\rangle).$$

Given an extended matrix instance $\underline{\mathrm{M}}^{\bar{k}}$, let $(*, \cdots *, |u_h\rangle\,, |u_g\rangle) := \underline{\mathrm{M}}^{\bar{k}}$ (see Equation (22)). The *normal initialization map* $\mathcal{U} : \mathbb{M}^n \to \mathbb{M}^n$ leaves all components of $\underline{\mathrm{M}}^{\bar{k}}$ unchanged, except for $|u_h\rangle$ and $|u_g\rangle$ which are mapped as (see Definition 85)

$$|u_h\rangle \mapsto |u(H, |w\rangle)\rangle \qquad \text{and} \qquad |u_g\rangle \mapsto |u(G, |v\rangle)\rangle\,.$$

Definition 95 (Weingarten Iteration Map)**.** Consider a matrix instance $\underline{\mathrm{X}}^{\bar{k}} =: \left(H^{\bar{k}}, G^{\bar{k}}, \left|w^{\bar{k}}\right\rangle, \left|v^{\bar{k}}\right\rangle\right)$ and let (see Definition 85)

$$\begin{aligned} \left|v^{\overline{k-1}}\right\rangle &:= \left|e\left(G^{\bar{k}}, \left|v^{\bar{k}}\right\rangle\right)\right\rangle, & \left|w^{\overline{k-1}}\right\rangle &:= \left|e\left(H^{\bar{k}}, \left|w^{\bar{k}}\right\rangle\right)\right\rangle, \\ G^{\overline{k-1}} &:= W^{\dashv}\left(G^{\bar{k}}, \left|v^{\bar{k}}\right\rangle\right), & H^{\overline{k-1}} &:= W^{\dashv}\left(H^{\bar{k}}, \left|w^{\bar{k}}\right\rangle\right). \end{aligned}$$

We define the *Weingarten Iteration Map* $\mathcal{W} : \mathbb{X}^n \longrightarrow \mathbb{X}^n$ by its action

$$\underline{\mathrm{X}}^{\bar{k}} \mapsto \left(H^{\overline{k-1}}, G^{\overline{k-1}}, \left|w^{\overline{k-1}}\right\rangle, \left|v^{\overline{k-1}}\right\rangle\right) =: \underline{\mathrm{X}}^{\overline{k-1}}.$$

Consider an extended matrix instance $\underline{\mathrm{M}}^{\bar{k}} =: \underline{\mathrm{X}}^{\bar{k}} \oplus S$ and let $\left((H^{\bar{k}})^{\dashv}, (G^{\bar{k}})^{\dashv}, *, *\right) := S$ (see Equation (22)). Let (see Definition 85)

$$(G^{\overline{k-1}})^{\dashv} := W\left(G^{\bar{k}}, (G^{\bar{k}})^{\dashv}, \left|v^{\bar{k}}\right\rangle\right) \qquad \text{and} \qquad (H^{\overline{k-1}})^{\dashv} := W\left(H^{\bar{k}}, (H^{\bar{k}})^{\dashv}, \left|w^{\bar{k}}\right\rangle\right).$$

We define the *Weingarten Iteration Map* $\mathcal{W} : \mathbb{M}^n \longrightarrow \mathbb{M}^n$ by its action

$$\underline{\mathrm{M}}^{\bar{k}} \mapsto \underline{\mathrm{X}}^{\overline{k-1}} \oplus \left((H^{\overline{k-1}})^{\dashv}, (G^{\overline{k-1}})^{\dashv}, |.\rangle, |.\rangle\right) =: \underline{\mathrm{M}}^{\overline{k-1}}.$$

The Weingarten Iteration and the Normal initialization maps fail when applied to cases involving infinite curvature, e.g. unbalanced f_0-assignment. To remedy this, we formalize the wiggle-w/v part of our approach. We start with the *Wiggle-w/v Normal Initialization Map* and continue with the *Wiggle-w/v Iteration Map*.

Definition 96 (Wiggle-w/v Normal Initialization Map)**.** Consider a matrix instance $\underline{\mathrm{X}}^{\bar{k}}$, let $(H, G, |w\rangle, |v\rangle) := \underline{\mathrm{X}}^{\bar{k}}$ with wiggle-w room along $|t_h\rangle$ (see Definition 93). The *Wiggle-w Normal Initialization Map* $\mathcal{U}_w : \mathbb{X}^n \longrightarrow \mathbb{M}^n$ is defined by its action

$$\underline{\mathrm{X}}^{\bar{k}} \mapsto \underline{\mathrm{X}}^{\bar{k}} \oplus \left([.], [.], \cos\theta\, |u\,(H, |w\rangle)\rangle + \sin\theta\, |t_h\rangle, |u\,(G, |v\rangle)\rangle\right)$$

where $\cos\theta := \langle v|u\,(G, |v\rangle)\rangle\, / \langle w|u\,(H, |w\rangle)\rangle$ (see Definition 94).

Given an extended matrix instance $\underline{\mathrm{M}}^{\bar{k}}$, let $\left(*, \cdots *, |u_h\rangle, \left|u_g\right\rangle\right) := \underline{\mathrm{M}}^{\bar{k}}$ (see Equation (22)), the *Wiggle-w Normal Initialization Map* $\mathcal{U}_w : \mathbb{M}^n \longrightarrow \mathbb{M}^n$ is defined by its action on $|u_h\rangle$ and $\left|u_g\right\rangle$ (see Definition 94) as

$$|u_h\rangle \mapsto \cos\theta\, |u\,(H, |w\rangle)\rangle + \sin\theta\, |t_h\rangle \qquad \text{and} \qquad \left|u_g\right\rangle \mapsto |u(G, |v\rangle)\rangle\,.$$

Similarly, for a matrix instance $(H, G, |w\rangle, |v\rangle) := \underline{\mathrm{X}}^{\bar{k}}$ with wiggle-v room along $\left|t_g\right\rangle$ (see Definition 93). The *Wiggle-v Normal Initialization Map* $\mathcal{U}_v : \mathbb{X}^n \longrightarrow \mathbb{M}^n$ is defined by its action

$$\underline{\mathrm{X}}^{\bar{k}} \mapsto \underline{\mathrm{X}}^{\bar{k}} \oplus \left([.], [.], |u\,(H, |w\rangle)\rangle, \cos\theta\, |u\,(G, |w\rangle)\rangle + \sin\theta \left|t_g\right\rangle\right)$$

where $\cos\theta := \langle w|u\,(H, |w\rangle)\rangle\, / \langle v|u\,(G, |v\rangle)\rangle$ (see Definition 94).

Given an extended matrix instance $\underline{\mathrm{M}}^{\bar{k}}$, let $\left(*, \cdots *, |u_h\rangle, \left|u_g\right\rangle\right) := \underline{\mathrm{M}}^{\bar{k}}$ (see Equation (22)), the *Wiggle-v Normal Initialization Map* $\mathcal{U}_v : \mathbb{M}^n \longrightarrow \mathbb{M}^n$ is defined by its action on $|u_h\rangle$ and $\left|u_g\right\rangle$ (see Definition 94) as

$$|u_h\rangle \mapsto |u\,(H, |w\rangle)\rangle \qquad \text{and} \qquad \left|u_g\right\rangle \mapsto \cos\theta\, |u\,(G, |v\rangle)\rangle + \sin\theta \left|t_g\right\rangle.$$

Definition 97 (Wiggle-w/v Iteration Map)**.** Consider an extended matrix instance $\underline{\mathrm{M}}^{\bar{k}}$ and let

$$\left(H^{\bar{k}}, G^{\bar{k}}, \left|w^{\bar{k}}\right\rangle, \left|v^{\bar{k}}\right\rangle, (H^{\bar{k}})^{\dashv}, (G^{\bar{k}})^{\dashv}, \left|u_h^{\bar{k}}\right\rangle, \left|u_g^{\bar{k}}\right\rangle\right) := \underline{\mathrm{M}}^{\bar{k}}.$$

Further, let[31] (see Definition 94)

$$\begin{aligned} \left|v^{\overline{k-1}}\right\rangle &= \left|e\left(G^{\bar{k}}, \left|v^{\bar{k}}\right\rangle\right)\right\rangle, & \left|w^{\overline{k-1}}\right\rangle &= \left|e\left(\left|u_h^{\bar{k}}\right\rangle, \left|w^{\bar{k}}\right\rangle\right)\right\rangle, \\ G^{\overline{k-1}} &= W^{\dashv}\left(G^{\bar{k}}, \left|v^{\bar{k}}\right\rangle\right), & H^{\overline{k-1}} &= W^{\dashv}\left(H^{\bar{k}}, \mathcal{N}\left((H^{\bar{k}})^{\dashv}\left|u_h^{\bar{k}}\right\rangle\right)\right), \\ (G^{\overline{k-1}})^{\dashv} &= W\left(G^{\bar{k}}, (G^{\bar{k}})^{\dashv}, \left|v^{\bar{k}}\right\rangle\right), & (H^{\overline{k-1}})^{\dashv} &= W\left(H^{\bar{k}}, (H^{\bar{k}})^{\dashv}, \mathcal{N}\left((H^{\bar{k}})^{\dashv}\left|u_h^{\bar{k}}\right\rangle\right)\right). \end{aligned}$$

The *Wiggle-w Iteration Map* $\mathcal{W}_w : \mathbb{M}^n \to \mathbb{M}^n$ is defined by its action

$$\underline{M}^{\bar{k}} \mapsto \left(H^{\overline{k-1}}, G^{\overline{k-1}}, \left|w^{\overline{k-1}}\right\rangle, \left|v^{\overline{k-1}}\right\rangle, (H^{\overline{k-1}})^{\dashv}, (G^{\overline{k-1}})^{\dashv}, |.\rangle, |.\rangle\right) =: \mathrm{M}^{\overline{k-1}}.$$

Similarly, consider an extended matrix instance $\underline{M}^{\bar{k}}$ and let

$$\left(H^{\bar{k}}, G^{\bar{k}}, \left|w^{\bar{k}}\right\rangle, \left|v^{\bar{k}}\right\rangle, (H^{\bar{k}})^{\dashv}, (G^{\bar{k}})^{\dashv}, \left|u_h^{\bar{k}}\right\rangle, \left|u_g^{\bar{k}}\right\rangle\right) := \underline{M}^{\bar{k}}.$$

Further, let (see Definition 94)

$$\begin{aligned} \left|v^{\overline{k-1}}\right\rangle &= \left|e\left(\left|u_g^{\bar{k}}\right\rangle, \left|v^{\bar{k}}\right\rangle\right)\right\rangle, & \left|w^{\overline{k-1}}\right\rangle &= \left|e\left(H^{\bar{k}}, \left|w^{\bar{k}}\right\rangle\right)\right\rangle, \\ G^{\overline{k-1}} &= W^{\dashv}\left(G^{\bar{k}}, \mathcal{N}\left((G^{\bar{k}})^{\dashv}\left|u_g^{\bar{k}}\right\rangle\right)\right), & H^{\overline{k-1}} &= W^{\dashv}\left(H^{\bar{k}}, \left|w^{\bar{k}}\right\rangle\right), \\ (G^{\overline{k-1}})^{\dashv} &= W\left(G^{\bar{k}}, (G^{\bar{k}})^{\dashv}, \mathcal{N}\left((G^{\bar{k}})^{\dashv}\left|u_g^{\bar{k}}\right\rangle\right)\right), & (H^{\overline{k-1}})^{\dashv} &= W\left(H^{\bar{k}}, (H^{\bar{k}})^{\dashv}, \left|w^{\bar{k}}\right\rangle\right). \end{aligned}$$

The *Wiggle-v Iteration Map* $\mathcal{W}_v : \mathbb{M}^n \to \mathbb{M}^n$ is defined by its action

$$\underline{M}^{\bar{k}} \mapsto \left(H^{\overline{k-1}}, G^{\overline{k-1}}, \left|w^{\overline{k-1}}\right\rangle, \left|v^{\overline{k-1}}\right\rangle, (H^{\overline{k-1}})^{\dashv}, (G^{\overline{k-1}})^{\dashv}, |.\rangle, |.\rangle\right) =: \mathrm{M}^{\overline{k-1}}.$$

Finally, we can state the analogue of Lemma 90 in the case where infinite curvatures arise and the wiggle-w/v method is employed.[32]

Lemma 98. *Consider an extended matrix instance $\underline{M}^{\bar{k}}$ with wiggle-w room ϵ along $\left|t_h^{\bar{k}}\right\rangle$ (see Definition 93). Assume it is completely specified (Definition 93), and it satisfies both $\mathcal{U}_w(\underline{M}^{\bar{k}}) = \underline{M}^{\bar{k}}$ (see Definition 96) and the contact condition (see Definition 93). Let $\left(*, \cdots *, \left|u_h^{\bar{k}}\right\rangle, \left|u_g^{\bar{k}}\right\rangle\right) := \underline{M}^{\bar{k}}$ and $\underline{M}^{\overline{k-1}} := \mathcal{W}_w(\underline{M}^{\bar{k}})$ (see Definition 97). We assert that if $Q^{\bar{k}}$ solves $\underline{M}^{\bar{k}}$ in the limit of $\epsilon \to 0$ then*

$$Q^{\bar{k}} = \left|u_h^{\bar{k}}\right\rangle\left\langle u_g^{\bar{k}}\right| + Q^{\overline{k-1}}, \tag{27}$$

where $Q^{\overline{k-1}}$ solves $\underline{M}^{\overline{k-1}}$.

Similarly, consider an extended matrix instance $\underline{M}^{\bar{k}}$ with wiggle-v room ϵ along $\left|t_h^{\bar{k}}\right\rangle$ (see Definition 93). Assume it is completely specified (see Definition 93), and it satisfies both $\mathcal{U}_v(\underline{M}^{\bar{k}}) = \underline{M}^{\bar{k}}$ and the contact

[31] Recall from the unbalanced f_0-solution that a very small change in the position vector could lead to a significant change in the normal and therefore also in the calculation of the curvature. Thus, here we infer the correct position as $(\left|H^{\dashv}\right\rangle|u_h\rangle)$ given the corrected normal $|u_h\rangle$.

[32] Even though the solution works in the limit of $\epsilon \to 0$, this is not non-physical, as it corresponds to allowing projections in the description of the protocol; see Section 5.

condition. Let $\left(*,\cdots *,\left|u_h^{\bar{k}}\right\rangle,\left|u_g^{\bar{k}}\right\rangle\right) := \underline{M}^{\bar{k}}$ *and* $\underline{M}^{\overline{k-1}} := \mathcal{W}_v(M^{\bar{k}})$. *We assert that if* $Q^{\bar{k}}$ *resolves* $\underline{M}^{\bar{k}}$ *in the limit of* $\epsilon \to 0$ *then*

$$Q^{\bar{k}} = \left|u_h^{\bar{k}}\right\rangle \left\langle u_g^{\bar{k}}\right| + Q^{\overline{k-1}},$$

where $Q^{\overline{k-1}}$ *resolves* $\underline{M}^{\overline{k-1}}$.

Proof. The basic idea of the proof is that the component of the normal along the $\left|t_h^{\bar{k}}\right\rangle$ direction can be taken to be arbitrary in the limit of $\epsilon \to 0$ (see Section 5.4.2). Let us consider a slightly different sequence of matrix instances, parametrized by ϵ, $\underline{X}'^{\bar{k}}(\epsilon) =: \left(H'^{\bar{k}}(\epsilon), G'^{\bar{k}}(\epsilon), \left|w'^{\bar{k}}(\epsilon)\right\rangle, \left|v'^{\bar{k}}(\epsilon)\right\rangle\right)$, which as $\epsilon \to 0$ converges to $\lim_{\epsilon\to 0} \underline{X}^{\bar{k}}(\epsilon) =: \left(H^{\bar{k}}, G^{\bar{k}}, \left|w^{\bar{k}}\right\rangle, \left|v^{\bar{k}}\right\rangle\right)$ (see Figure 17). One can use operator monotones to construct such a sequence explicitly and show that the solution of all these instances is the same as a function of ϵ. While the parameters specifying the matrix instances converge, the normal itself does not converge accordingly i.e.,

$$\lim_{\epsilon\to 0}\left|u_h^{\bar{k}}(\epsilon)\right\rangle \neq \left|u\left(H'^{\bar{k}}, \left|w'^{\bar{k}}\right\rangle\right)\right\rangle.$$

This is because a small wiggle in $\left|w^{\bar{k}}\right\rangle$ can significantly affect the normal as the curvature along one of the directions diverges. Hence, given $H^{\bar{k}}$ evaluating the normal along $\lim_{\epsilon\to 0}\left|w^{\bar{k}}(\epsilon)\right\rangle$ is not the same as evaluating the normal along $\left|w^{\bar{k}}\right\rangle$.

We can iterate $\underline{X}'^{\bar{k}}(\epsilon)$ using Definition 95 and Lemma 90, and because the complete solution doesn't depend on ϵ, we can use it to iterate $\underline{X}^{\bar{k}}$. Since it is along $\left|t_h^{\bar{k}}\right\rangle$ where the curvature diverges as $\epsilon \to 0$, the component of the normal along this direction gets ill-defined. Therefore[33]

$$\lim_{\epsilon\to 0}\left|u_h'^{\bar{k}}(\epsilon)\right\rangle = \cos\theta\left|u\left(H^{\bar{k}}, \left|w^{\bar{k}}\right\rangle\right)\right\rangle + \sin\theta\left|t_h^{\bar{k}}\right\rangle,$$

where $\cos\theta$ remains to be determined. The contact condition, $\left\langle u_h'^{\bar{k}}(\epsilon)|w'^{\bar{k}}(\epsilon)\right\rangle = \left\langle u_g'^{\bar{k}}(\epsilon)|v'^{\bar{k}}(\epsilon)\right\rangle$, in the limit $\epsilon \to 0$ becomes $\cos\theta\left\langle u\left(H^{\bar{k}}, \left|w^{\bar{k}}\right\rangle\right)|w^{\bar{k}}\right\rangle = \left\langle u_g^{\bar{k}}|v^{\bar{k}}\right\rangle$, since $\left\langle w^{\bar{k}}|t_h^{\bar{k}}\right\rangle = 0$, thus fixing $\cos\theta$. We define $\left|u_h^{\bar{k}}\right\rangle := \lim_{\epsilon\to 0}\left|u_h'^{\bar{k}}(\epsilon)\right\rangle$, and using $\mathcal{W}(\underline{X}'^{\bar{k}}(\epsilon)) =: \underline{X}'^{\overline{k-1}}(\epsilon) =: \left(H'^{\overline{k-1}}(\epsilon), G'^{\overline{k-1}}(\epsilon), \left|w'^{\overline{k-1}}(\epsilon)\right\rangle, \left|v'^{\overline{k-1}}(\epsilon)\right\rangle\right)$, in the limit $\epsilon \to 0$, we define

$$\underline{X}^{\overline{k-1}} =: \left(H^{\overline{k-1}}, G^{\overline{k-1}}, \left|w^{\overline{k-1}}\right\rangle, \left|v^{\overline{k-1}}\right\rangle\right).$$

Since the diverging term is in $H'^{\bar{k}}(\epsilon)$ and not in $G'^{\bar{k}}(\epsilon)$ it follows that $G^{\overline{k-1}}$ and $\left|v^{\overline{k-1}}\right\rangle$ can be evaluated using the usual rule specified by the Weingarten Iteration Map, $\mathcal{W}$ on $\underline{X}^{\bar{k}}$. The relatively non-trivial part is to show that $H^{\overline{k-1}}$ and $\left|w^{\overline{k-1}}\right\rangle$ can be analogously defined using the correct normal, $\left|u_h^{\bar{k}}\right\rangle$. Given a direction of contact $|w\rangle$, the normal vector of the ellipsoid represented by H is along $H|w\rangle$, or—by running the same argument backwards— given a normal vector $|u\rangle$, one can obtain the direction of the point of contact as $H^{-1}|u\rangle$. Since $\left|w^{\bar{k}}\right\rangle$ can not be reliably used to derive quantities we use $\left|u_h^{\bar{k}}\right\rangle$ to evaluate the Weingarten

[33] Subtleties about degeneracies in $\left|t_h^{\bar{k}}\right\rangle$ are not hard to handle; see Section 5.4.2.

map[34] as in Definition 97. If Q solves $\underline{\mathrm{X}}'^{\bar{k}}(\epsilon)$ then from Lemma 90 we have $Q^{\bar{k}} = \left|u_h'^{\bar{k}}(\epsilon)\right\rangle \left\langle u_g'^{\bar{k}}(\epsilon)\right| + Q^{\overline{k-1}}(\epsilon)$, where $Q^{\bar{k}}$ solves $\underline{\mathrm{X}}'^{\bar{k}}(\epsilon)$ but doesn't depend on ϵ. Taking the limit and using the correct normals we obtain Equation (27). □

We have introduced everything we need in order to present the solution to the unbalanced f_0-assignment:

Proposition 99 (The unbalanced f_0-solution). *Let $t = h - g = \sum_{i=1}^{2n-1} p_i \left[\!\left[x_i\right]\!\right]$ be an f_0-assignment over the real coordinates $0 \le x_1 < x_2 \cdots < x_{2n-1}$. Let $h = \sum_{i=1}^{n-1} p_{h_i} \left[\!\left[x_{h_i}\right]\!\right]$, $g = \sum_{i=1}^{n} p_{g_i} \left[\!\left[x_{g_i}\right]\!\right]$ where $p_{h_i}, p_{g_i} > 0$, and $\{x_{h_i}\}$ and $\{x_{g_i}\}$ are all distinct. Consider the matrix instance $\underline{X}$:=$(X_h, X_g, |w\rangle, |v\rangle)$, where $X_h \doteq diag(x_{h_1}, x_{h_2} \dots x_{h_{n-1}}, 1/\epsilon)$, $X_g \doteq diag(x_{g_1}, x_{g_2} \dots x_{g_{n-1}}, x_{g_n})$, $|w\rangle \doteq (\sqrt{p_{h_1}}, \sqrt{p_{h_2}} \dots \sqrt{p_{h_{n-1}}}, 0)^T$, $|v\rangle \doteq \left(\sqrt{p_{g_1}}, \sqrt{p_{g_2}} \dots \sqrt{p_{g_{n-1}}}, \sqrt{p_{g_n}}\right)^T$. In the limit of $\epsilon \longrightarrow 0$, the orthogonal matrix*

$$O = \sum_{k=1}^{n} \left|u_h^{\bar{k}}\right\rangle \left\langle u_g^{\bar{k}}\right|$$

solves $\underline{X} =: \underline{X}^{\bar{n}}$ (see Definition 93) where the Weingarten Iteration Map (see Definition 95) is used to evaluate $\underline{X}^{\overline{k-1}} = \mathcal{W}(\underline{X}^{\bar{k}})$ until $k = 2$, starting from $k = n$. The Normal Initialization Map (see Definition 94) is used until $k = 3$ to obtain $\left|u_h^{\bar{k}}\right\rangle$ and $\left|u_g^{\bar{k}}\right\rangle$, viz. $\mathcal{U}(\underline{X}^{\bar{k}}) =: \left(*, \cdots *, \left|u_h^{\bar{k}}\right\rangle, \left|u_g^{\bar{k}}\right\rangle\right)$. The Wiggle-w Normal Initialization Map (see Definition 94) is used to evaluate $\left|u_h^{\bar{2}}\right\rangle$ and $\left|u_g^{\bar{2}}\right\rangle$, viz. $\mathcal{U}_w(\underline{X}^{\bar{2}}) =: \left(*, *, \left|w^{\bar{2}}\right\rangle, \left|v^{\bar{2}}\right\rangle\right) \oplus \left(*, *, \left|u_h^{\bar{2}}\right\rangle, \left|u_g^{\bar{2}}\right\rangle\right)$. Finally, $\left|u_h^{\bar{1}}\right\rangle := \left|e\left(\left|u_h^{\bar{2}}\right\rangle, \left|w^{\bar{2}}\right\rangle\right)\right\rangle$ and $\left|u_g^{\bar{1}}\right\rangle := \left|e\left(\left|u_g^{\bar{2}}\right\rangle, \left|v^{\bar{2}}\right\rangle\right)\right\rangle$.

Proof. The proof is essentially the same as that for the balanced case until the very last step. After iterating for l steps, suppose the matrix instance one obtains is $\underline{\mathrm{X}}^{\overline{n-l}}$. To check if another Weingarten iteration is possible, we must check if

$$\begin{aligned} &\left\langle w^{\overline{n-l}}\right| \left(H^{\overline{n-l}}\right)^m \left|w^{\overline{n-l}}\right\rangle - \left\langle v^{\overline{n-l}}\right| \left(G^{\overline{n-l}}\right)^m \left|v^{\overline{n-l}}\right\rangle = \\ &r\left(\left\langle (H^{\bar{n}})^m\right\rangle, \left\langle (H^{\bar{n}})^{m+1}\right\rangle \dots, \left\langle (H^{\bar{n}})^{2l+m}\right\rangle\right) - r\left(\left\langle (G^{\bar{n}})^m\right\rangle, \left\langle (G^{\bar{n}})^{m+1}\right\rangle \dots, \left\langle (G^{\bar{n}})^{2l+m}\right\rangle\right) \end{aligned} \tag{28}$$

vanishes for both $m = 1$ and $m = 2$, i.e.,

$$\left\langle x^{2l+1}\right\rangle = 0 \text{ and } \left\langle x^{2l+2}\right\rangle = 0 \tag{29}$$

and for their lower power analogues (see Equation (25)). The $m = 1$ case is the contact condition and $m = 2$ is the component condition (see Definition 93). If $2l + 2 \le 2n - 3$ then from Lemma 33 (we use $2n - 1$ instead of n in the lemma) it follows that both terms in Equation (28) become identical and hence the difference indeed vanishes. Consequently, until $l = n - 3$, one can iterate to obtain $\underline{\mathrm{X}}^{\bar{n}}, \underline{\mathrm{X}}^{\overline{n-1}}, \dots \underline{\mathrm{X}}^{\bar{3}}, \underline{\mathrm{X}}^{\bar{2}}$ which in turn can be used to determine $\left|u_h^{\bar{n}}\right\rangle, \left|u_h^{\overline{n-1}}\right\rangle, \dots, \left|u_h^{\bar{3}}\right\rangle$ and $\left|u_g^{\bar{n}}\right\rangle, \left|u_g^{\overline{n-1}}\right\rangle, \dots, \left|u_g^{\bar{3}}\right\rangle$ (see Definition 94). Since $\left\langle x^{2n-3=2(n-2)+1}\right\rangle = 0$ but $\left\langle x^{2n-2=2(n-2)+2}\right\rangle \neq 0$, we can use Definition 96 on $\underline{\mathrm{X}}^{\overline{2=n-(n-2)}}$ to determine $\left|u_h^{\bar{2}}\right\rangle$ and $\left|u_g^{\bar{2}}\right\rangle$. The vectors $\left|w^{\bar{1}}\right\rangle$ and $\left|v^{\bar{1}}\right\rangle$ are fixed by the requirement that O is orthogonal and $O\,|v\rangle = |w\rangle$. In Appendix H we show that there exists O solving the matrix instance $\underline{\mathrm{X}}^{\bar{n}}$, therefore using Lemma 90 and Lemma 98 we completely determine $O = \sum_{k=1}^{n} \left|u_h^{\bar{k}}\right\rangle \left\langle u_g^{\bar{k}}\right|$. □

In Figure 18 we show an example of an f_0-assignment with 5 points.

[34] It is not hard to see why $H^{\overline{k-1}}$ does not diverge as ϵ goes to zero (granted there was only one diverging eigenvalue in $H^{\bar{k}}$ to start with). The idea is simply to use the reverse Weingarten map; this suppresses the divergence into zero, then one projects out a rank-one subspace. If there was only one zero eigenvalue and if the subspace includes this eigenspace (spanned by a single eigenvector), then the resulting matrix would not have any zero eigenvalues. This can then be inverted to obtain the Weingarten map which is now finite and well-defined.

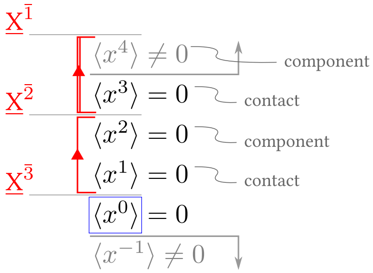


Figure 18: Power diagram representative of an unbalanced f_0-assignment with 5 points (again $n = 3$). Starting upwards from $\langle x^0 \rangle$, one iteration is completed before encountering the instance where the contact condition still holds but the normals do not match, thus the wiggle-w method (double-line arrows) is employed.

6.2 Solution to monomial assignments

For the f_0-assignments every iteration led to an increase in the power of x in the expectation value $\langle x^k \rangle$—we were moving upwards in the power diagram, see Figure 16 and Figure 18. For monomial assignments, though, this is not exactly the case, as there are iterations that lead to a decrease in the power of x—we also need to move downwards in the power diagram, see Figure 19 and Figure 20. This decrease corresponds to inverting the coordinates in the f−assignment, and in Appendix G we show that this transformation leads to the transposition of the solution, i.e. if O solves an f−assignment, t, then O^T solves the f−assignment resulting from the inversion of the coordinates in t. In the same vein, a monomial with the highest permissible degree can be seen as an f_0−assignment, when it comes to its solution (see also Example 159 for an illustration of an alternative solution to this assignment). In this context, to obtain the solutions to general monomial assignments we need to combine iterations using the matrices and their inverses. The following *Flip Map* switches between these two kinds of iterations.

Definition 100 (Flip Map)**.** Consider an extended matrix instance $\underline{M}^{\bar{n}} =: \left(H, G, |w\rangle, |v\rangle, H^{-1}, G^{-1}, |u_h\rangle, |u_g\rangle\right)$. We define the *Flip Map* $\mathcal{F} : \mathbb{M}^n \to \mathbb{M}^n$ as $\underline{M}^{\bar{n}} \mapsto \left(H^{-1}, G^{-1}, |w\rangle, |v\rangle, H, G, |u_h\rangle, |u_g\rangle\right) =: \mathcal{F}(M^{\bar{n}})$.

We also need a way to keep track of the powers in the contact and component conditions of the matrix instances after a certain number of iterations in both directions. To this end, we state the following lemma which can be proven by combining Lemma 164, Lemma 165 and Lemma 166 in Appendix I.

Lemma 101 (Up-then-Down Contact/Component Lemma)**.** *Consider the extended matrix instance*

$$\underline{M}'^{\bar{n}} := \mathcal{U}(H'^{\bar{n}}, G'^{\bar{n}}, \left|w'^{\bar{n}}\right\rangle, \left|v'^{\bar{n}}\right\rangle, (H'^{\bar{n}})^{-1}, (G'^{\bar{n}})^{-1}, |.\rangle, |.\rangle).$$

Suppose the Normal Initialization Map and the Weingarten Iteration Map (see Definition 94 and Definition 95) are applied k times to obtain $\underline{M}'^{\overline{n-k}}$. Let $n - k = d$ and consider $\underline{\tilde{M}}^{\bar{d}} = \mathcal{U}(\mathcal{F}(\underline{M}'^{\bar{d}}))$. Suppose the Normal Initialization Map and the Weingarten Iteration map are applied l more times to obtain $\underline{\tilde{M}}^{\overline{d-l}} =: \left(\tilde{H}^{\overline{d-l}}, \tilde{G}^{\overline{d-l}}, \left|\tilde{w}^{\overline{d-l}}\right\rangle, \left|\tilde{v}^{\overline{d-l}}\right\rangle, *, \cdots *\right)$. *Then,*

$$\left\langle \tilde{v}^{\overline{n-k-l}} \right| \left(\tilde{G}^{\overline{n-k-l}}\right)^{\mu} \left|\tilde{v}^{\overline{n-k-l}}\right\rangle = r\left(\left\langle (G'^{\bar{n}})^{-(2l+\mu)}\right\rangle, \ldots, \left\langle (G'^{\bar{n}})^{2k-1+\mu}\right\rangle, \left\langle (G'^{\bar{n}})^{2k+\mu}\right\rangle\right)$$

where $\mu \geq 1$ and r is a multi-variate function which does not have an implicit dependence on $\left\langle (G'^{\bar{n}})^i \right\rangle := \left\langle v'^{\bar{n}}\right| (G'^{\bar{n}})^i \left|v'^{\bar{n}}\right\rangle$ for any i. The corresponding statement for H and $|w\rangle$ also holds.

From Definition 32 we know that a monomial problem can either be balanced or unbalanced. We find the solution in these two cases separately and in both cases we differentiate between aligned and misaligned assignments (see Definition 32).

Proposition 102 (Solving the Balanced Monomial Problem). *Let*

$$t = \sum_{i=1}^{2n} -\frac{(-x_i)^m}{\prod_{j\neq i}(x_j - x_i)} \llbracket x_i \rrbracket = \sum_{i=1}^{n} x_{h_i}^m p_{h_i} \llbracket x_{h_i} \rrbracket - \sum_{i=1}^{n} x_{g_i}^m p_{g_i} \llbracket x_{g_i} \rrbracket$$

be a balanced monomial assignment over the real coordinates $0 < x_1 < x_2 \cdots < x_{2n-1} < x_{2n}$ *(see Definition 32) where* $p_{h_i}, p_{g_i} > 0$ *and* $\{x_{h_i}\}$ *and* $\{x_{g_i}\}$ *are all distinct. For* $m = 0$ *and* $m = 2n-2$ *the problem reduces to the* f_0*-assignment (see Proposition 99) using Corollary 158 in the latter case. For the remaining cases consider the corresponding matrix instance* $\underline{X}^{\bar{\eta}} := (X_h^{\bar{\eta}}, X_g^{\bar{\eta}}, (X_h^{\bar{\eta}})^b \left|w\right\rangle, (X_g^{\bar{\eta}})^b \left|v\right\rangle)$ *where*

- *if* $b = m/2$ *is an integer (the* aligned case*) then* $\eta = n$, $j' = j = 1$,

$$X_h^{\bar{n}} \doteq diag(x_{h_1}, x_{h_2} \dots x_{h_n}), \qquad X_g^{\bar{n}} \doteq diag(x_{g_1}, x_{g_2} \dots x_{g_n}),$$
$$\left|w^{\bar{n}}\right\rangle \doteq (\sqrt{p_{h_1}}, \sqrt{p_{h_2}} \dots \sqrt{p_{h_n}}), \qquad \left|v^{\bar{n}}\right\rangle \doteq (\sqrt{p_{g_1}}, \sqrt{p_{g_2}} \dots \sqrt{p_{g_n}}).$$

- *else if* $b = m/2$ *is not an integer (the* misaligned case*) then* $\eta = n+1$, $j' = 3$, $j = 4$,

$$X_h^{\overline{n+1}} \doteq diag(x_{h_1}, x_{h_2} \dots x_{h_n}, 1/\epsilon), \qquad X_g^{\overline{n+1}} \doteq diag(x_{g_1}, x_{g_2} \dots x_{g_n}, \epsilon),$$
$$\left|w^{\overline{n+1}}\right\rangle \doteq (\sqrt{p_{h_1}}, \sqrt{p_{h_2}} \dots \sqrt{p_{h_n}}, 0), \qquad \left|v^{\overline{n+1}}\right\rangle \doteq (\sqrt{p_{g_1}}, \sqrt{p_{g_2}} \dots \sqrt{p_{g_n}}, 0).$$

Let $k = \left\lfloor \frac{2n-2-m}{2} \right\rfloor$. *In the limit of* $\epsilon \longrightarrow 0$, *the matrix instance is solved by*

$$O = \sum_{i=\eta}^{\eta-k+1} \left|u_h'^{\bar{i}}\right\rangle \left\langle u_g'^{\bar{i}}\right| + \sum_{i=\eta-k}^{j} \left|\tilde{u}_h^{\bar{i}}\right\rangle \left\langle \tilde{u}_g^{\bar{i}}\right| + (1 - \delta_{j,j'}) \sum_{i=j'}^{1} \left|u_h'^{\bar{i}}\right\rangle \left\langle u_g'^{\bar{i}}\right|,$$

where the terms of the first sum are evaluated in the same way regardless of the alignment. We start with $\underline{M}'^{\bar{\eta}} := \mathcal{U}\left(\underline{X}^{\bar{\eta}} \oplus \left((X_h^{\bar{\eta}})^{-1}, (X_g^{\bar{\eta}})^{-1}, \left|.\right\rangle, \left|.\right\rangle\right)\right)$ *(see Definition 93, Definition 94 and Definition 95) and define*

$$\underline{M}'^{\bar{l}} =: \left(*, \cdots *, \left|u_h'^{\bar{l}}\right\rangle, \left|u_h'^{\bar{l}}\right\rangle\right) \qquad for \quad \eta - k + 1 \leq l \leq \eta$$

using the relations

$$\underline{M}'^{\overline{l-1}} := \mathcal{U}(\mathcal{W}(\underline{M}'^{\bar{l}})) \qquad for \quad \eta - k + 1 \leq l - 1 \leq \eta - 1.$$

The terms of the second sum are also the same in both cases. We start with $\underline{\tilde{M}}^{\overline{\eta-k}} := \mathcal{U}(\mathcal{F}(\underline{M}'^{\overline{\eta-k}}))$ *and using the relations*

$$\underline{\tilde{M}}^{\overline{l-1}} := \mathcal{U}(\mathcal{W}(\underline{\tilde{M}}^{\bar{l}})) \qquad for \quad j' \leq l - 1 \leq \eta - k - 1$$

we define

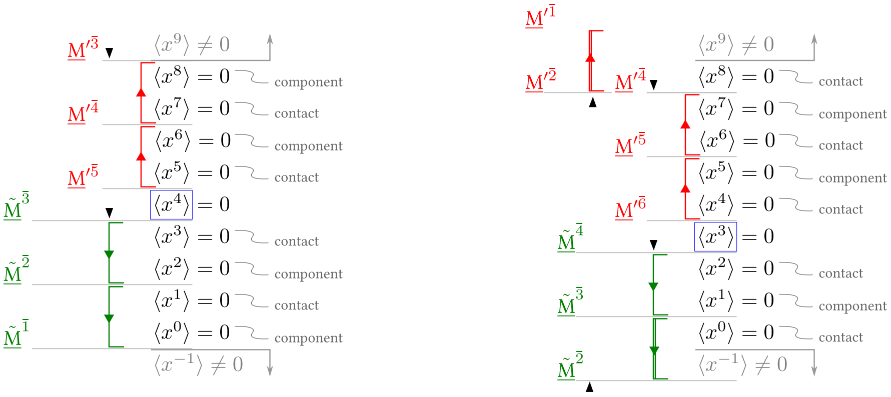
$$\left(*, \cdots *, \left|\tilde{u}_h^{\bar{l}}\right\rangle, \left|\tilde{u}_g^{\bar{l}}\right\rangle\right) := \underline{\tilde{M}}^{\bar{l}} \qquad for \quad j \leq l \leq \eta - k.$$

At this point, the aligned problem is solved, and we use the following relations to specify the terms of the third sum, which solves the misaligned problem:

$$\underline{\tilde{M}}^{\bar{3}} := \mathcal{U}_v(\mathcal{W}(\underline{\tilde{M}}^{\bar{4}}))$$
$$\underline{M}'^{\bar{2}} := \mathcal{U}_w(\mathcal{F}(\mathcal{W}_v(\underline{\tilde{M}}^{\bar{3}}))) =: \left(*, *, \left|w'^{\bar{2}}\right\rangle, \left|v'^{\bar{2}}\right\rangle, *, *, \left|u'^{\bar{2}}_h\right\rangle, \left|u'^{\bar{2}}_g\right\rangle\right)$$
$$\left|u'^{\bar{1}}_h\right\rangle := \left|e\left(\left|u'^{\bar{2}}_h\right\rangle, \left|w'^{\bar{2}}\right\rangle\right)\right\rangle \quad \textit{and} \quad \left|u'^{\bar{1}}_g\right\rangle := \left|e\left(\left|u'^{\bar{2}}_g\right\rangle, \left|v'^{\bar{2}}\right\rangle\right)\right\rangle,$$

where we used Definition 100, Definition 96 and Definition 97.

Proof. We first prove that O solves $\underline{X}^{\bar{n}}$ in the *aligned case*, i.e. when $b = m/2$ is an integer (see Figure 19 and note that $\eta = n$ in this case). We denote the components of $\underline{M}'^{\bar{l}}$ by



$\underline{M}'^{\bar{3}}$
$\underline{M}'^{\bar{4}}$
$\underline{M}'^{\bar{5}}$
$\underline{\tilde{M}}^{\bar{3}}$
$\underline{\tilde{M}}^{\bar{2}}$
$\underline{\tilde{M}}^{\bar{1}}$
$\langle x^9\rangle \neq 0$
$\langle x^8\rangle = 0$ component
$\langle x^7\rangle = 0$ contact
$\langle x^6\rangle = 0$ component
$\langle x^5\rangle = 0$ contact
$\langle x^4\rangle = 0$
$\langle x^3\rangle = 0$ contact
$\langle x^2\rangle = 0$ component
$\langle x^1\rangle = 0$ contact
$\langle x^0\rangle = 0$ component
$\langle x^{-1}\rangle \neq 0$
$\underline{M}'^{\bar{1}}$
$\underline{M}'^{\bar{2}}$
$\underline{M}'^{\bar{4}}$
$\underline{M}'^{\bar{5}}$
$\underline{M}'^{\bar{6}}$
$\underline{\tilde{M}}^{\bar{4}}$
$\underline{\tilde{M}}^{\bar{3}}$
$\underline{\tilde{M}}^{\bar{2}}$
$\langle x^9\rangle \neq 0$
$\langle x^8\rangle = 0$ contact
$\langle x^7\rangle = 0$ component
$\langle x^6\rangle = 0$ contact
$\langle x^5\rangle = 0$ component
$\langle x^4\rangle = 0$ contact
$\langle x^3\rangle = 0$
$\langle x^2\rangle = 0$ contact
$\langle x^1\rangle = 0$ component
$\langle x^0\rangle = 0$ contact
$\langle x^{-1}\rangle \neq 0$

Figure 19: Power diagram representative of the aligned (left) and misaligned (right) balanced monomial assignment for $2n = 10$ with $m = 4$ (left) and $m = 3$ (right).

$$\left(H'^{\bar{l}}, G'^{\bar{l}}, \left|w'^{\bar{l}}\right\rangle, \left|v'^{\bar{l}}\right\rangle, * \dots, *\right) := \underline{M}'^{\bar{l}}.$$

We start by checking if $\underline{M}'^{\bar{n}}$ satisfies the contact condition, that is $\left\langle w'^{\bar{n}}\right| H'^{\bar{n}} \left|w'^{\bar{n}}\right\rangle = \left\langle v'^{\bar{n}}\right| G'^{\bar{n}} \left|v'^{\bar{n}}\right\rangle$. The LHS is $\left\langle w^{\bar{n}}\right| (X^{\bar{n}}_h)^{2b+1} \left|w^{\bar{n}}\right\rangle = \left\langle (X^{\bar{n}}_h)^{m+1}\right\rangle$ and similarly the RHS is $\left\langle (X^{\bar{n}}_g)^{m+1}\right\rangle$. The contact condition can then be expressed as $\left\langle x^{m+1}\right\rangle = 0$, and similarly the component condition as $\left\langle x^{m+2}\right\rangle = 0$. From Lemma 33, we know that they hold for $m + 2 \leq 2n - 2$, i.e. $m \leq 2n - 4$ (see Figure 19 with $2n = 10$ which means that m can be at most 6 for the contact/component conditions to hold). Assuming $m \leq 2n - 4$ we can apply the Weingarten Iteration Map (Definition 95) and use Lemma 90 along with the Normal Initialization Map (see Definition 94) to construct part of the solution, viz. use $\underline{M}'^{\overline{l-1}} := \mathcal{U}(\mathcal{W}(\underline{M}'^{\bar{l}}))$. Suppose we iterate κ times to obtain $\underline{M}'^{\overline{n-\kappa}}$. The contact condition now corresponds to

$$\left\langle w'^{\overline{n-\kappa}}\right| H'^{\overline{n-\kappa}} \left|w'^{\overline{n-\kappa}}\right\rangle = \left\langle v'^{\overline{n-\kappa}}\right| G'^{\overline{n-\kappa}} \left|v'^{\overline{n-\kappa}}\right\rangle.$$

The RHS can be written as

$$r\left(\left\langle w'^{\bar{n}}\right| (H'^{\bar{n}})^1 \left|w'^{\bar{n}}\right\rangle, \left\langle w'^{\bar{n}}\right| (H'^{\bar{n}})^2 \left|w'^{\bar{n}}\right\rangle \dots \left\langle w'^{\bar{n}}\right| (H'^{\bar{n}})^{2\kappa+1} \left|w'^{\bar{n}}\right\rangle\right)$$

using Lemma 92. Similarly for the LHS. The contact condition can then be expressed as $\left\langle x^{2\kappa+1+m}\right\rangle = 0$ Similarly, the component condition can be expressed as $\left\langle x^{2\kappa+2+m}\right\rangle = 0$. From Lemma 33, we know that these conditions hold if $2\kappa + 2 + m \leq 2n - 2$ which yields $\kappa \leq n - b - 2 = k - 1$. Therefore, if O solves the matrix instance then it must have the form $O = \sum_{l=1}^{n-k+1} \left|u'^{\bar{l}}_{h}\right\rangle \left\langle u'^{\bar{l}}_{g}\right| + Q^{\overline{n-k}}$, where $Q^{\overline{n-k}}$ is an isometry acting on the orthogonal space which remains to be determined. To proceed, we can apply the Weingarten Iteration Map to $\underline{\mathrm{M}}'^{\overline{n-k+1}}$ and obtain $\mathcal{W}(\underline{\mathrm{M}}'^{\overline{n-k+1}}) =: \underline{\mathrm{M}}'^{\overline{n-k}}$, but this instance satisfies neither the contact nor the component condition (corresponds to $\underline{\mathrm{M}}'^{\bar{3}}$ in Figure 19).

Let $(H, G, |w\rangle, |v\rangle, H^{\dashv}, G^{\dashv}, *, *) := \underline{\mathrm{M}}'^{\overline{n-k}}$. Solving $\underline{\mathrm{M}}'^{\overline{n-k}}$ corresponds to finding a Q such that $Q|v\rangle = |w\rangle$ and $H \geq QGQ^{T}$. The matrix inequality can equivalently be written as $H^{\dashv} \leq QG^{\dashv}Q^{T}$. Using $H^{\dashv}$ and $G^{\dashv}$ decreases the powers and thereby allows us to proceed. We evaluate

$$\underline{\tilde{\mathrm{M}}}^{\overline{n-k}} = \mathcal{U}(\mathcal{F}(\mathcal{W}(\underline{\mathrm{M}}'^{\overline{n-k+1}})),$$

and let $\underline{\tilde{\mathrm{M}}}^{\bar{l}} =: \left(\tilde{H}^{\bar{l}}, \tilde{G}^{\bar{l}}, \left|\tilde{w}^{\bar{l}}\right\rangle, \left|\tilde{v}^{\bar{l}}\right\rangle\right)$ (this step is indicated by the small triangles next to $\underline{\mathrm{M}}'^{\bar{3}}$ and $\underline{\tilde{\mathrm{M}}}^{\bar{3}}$ in Figure 19). Let the matrix instance one obtains after iterating l times using $\underline{\tilde{\mathrm{M}}}^{\overline{l-1}} := \mathcal{U}(\mathcal{W}(\underline{\tilde{\mathrm{M}}}^{\bar{l}}))$ starting with $\underline{\tilde{\mathrm{M}}}^{\overline{n-k}}$ be $\underline{\tilde{\mathrm{M}}}^{\overline{n-k-l}}$. The contact condition for $\underline{\tilde{\mathrm{M}}}^{\overline{n-k-l}}$ is

$$\left\langle \tilde{w}^{\overline{n-k-l}} \right| \tilde{H}^{\overline{n-k-l}} \left| \tilde{w}^{\overline{n-k-l}} \right\rangle = \left\langle \tilde{v}^{\overline{n-k-l}} \right| \tilde{G}^{\overline{n-k-l}} \left| \tilde{v}^{\overline{n-k-l}} \right\rangle,$$

which effectively becomes $\left\langle x^{-(2l+1)+m} \right\rangle = 0$ using Lemma 101, noting that the lowest power is relevant here, and that $\left|w'^{\bar{n}}\right\rangle = (X^{\bar{n}}_{h})^{m/2} \left|w^{\bar{n}}\right\rangle$. Similarly for $\left|v'^{\bar{n}}\right\rangle$. Analogously, the component condition yields $\left\langle x^{-(2l+2)+m} \right\rangle = 0$. From Lemma 33, we know that these conditions hold if $0 \leq -(2l + 2) + m$ which yields $l \leq b - 1$. This means that the rank, i.e. $n - k - l$, until which the contact/component condition holds is $\cancel{n - n} + 1 \cancel{+b - b} + 1 = 2$ (included), where we used $k = n - b - 1$. Hence, if $Q^{\overline{n-k}}$ resolves $\underline{\tilde{\mathrm{M}}}^{\overline{n-k}}$ then it must have the form $Q^{\bar{k}} = \sum_{l=n-k}^{1} \left|\tilde{u}^{\bar{l}}_{h}\right\rangle \left\langle \tilde{u}^{\bar{l}}_{g}\right|$, due to Lemma 90, which completely specifies $Q^{\bar{k}}$ and, together with the previous argument, proves that O solves $\underline{\mathrm{M}}^{\bar{n}}$.

Let us move to the misaligned case (i.e. when $m/2$ is not an integer; see Figure 19). We can proceed as in the aligned case until the contact/component condition is violated. After κ steps the said condition is $\left\langle x^{2\kappa+2+m} \right\rangle = 0$ which holds until $2\kappa + 2 + m \leq 2n - 2$ (using Lemma 33). This corresponds to $\kappa \leq \frac{2n-2-m}{2} - 1$, which yields $\kappa \leq k - 1$. Hence $\underline{\mathrm{M}}'^{\overline{\eta-k+1}}$ is the last instance satisfying the required contact/component conditions (this corresponds to $\underline{\mathrm{M}}'^{\bar{5}}$ in Figure 19; use $(n + 1) - (k - 1)$ with $n = 5$, $k = 2$). Supposing O solves $\underline{\mathrm{X}}^{\overline{n+1}}$ we deduce (using Lemma 90 and the arguments from the previous case) that it must have the form $O = \sum_{l=\eta}^{\eta-k+1} \left|u'^{\bar{l}}_{h}\right\rangle \left\langle u'^{\bar{l}}_{g}\right| + Q^{\overline{\eta-k}}$. At the instance $\underline{\mathrm{M}}'^{\overline{\eta-k}} = \mathcal{W}(\underline{\mathrm{M}}'^{\eta-k+1})$ we flip as before to obtain $\underline{\tilde{\mathrm{M}}}^{\overline{\eta-k}} = \mathcal{U}(\mathcal{F}(\underline{\mathrm{M}}'^{\overline{n-k}}))$ (these are indicated by the triangles next to $\underline{\mathrm{M}}'^{\bar{4}}$ and $\underline{\tilde{\mathrm{M}}}^{\bar{4}}$ in Figure 19). We write the contact/component condition after l iterations as $\left\langle x^{-(2l+2)+m} \right\rangle = 0$ which from Lemma 33 holds if $0 \leq -(2l + 2) + m$. This in turn yields $l \leq m/2 - 1$ entailing that the rank, i.e. $\eta - k - l$, until which the contact/component condition holds is $\cancel{\eta} + 1 - (\cancel{\eta} - 1 + \lfloor -m/2 \rfloor) - (\lfloor m/2 \rfloor - 1) = 4$ (this corresponds to $\underline{\tilde{\mathrm{M}}}^{\bar{4}}$ in Figure 19). Continuing with the argument for the form of O, we can deduce (again, using Lemma 90 and the previous reasoning) that $Q^{\overline{\eta-k}} = \sum_{l=\eta-k}^{4} \left|\tilde{u}^{\bar{l}}_{h}\right\rangle \left\langle \tilde{u}^{\bar{l}}_{g}\right| + Q^{\bar{3}}$. Since $\underline{\tilde{\mathrm{M}}}^{\bar{4}}$ satisfies the required contact/component conditions, we can iterate once more. However, at this point, only the contact condition holds but the component condition does not (see Figure 19). Consider $\underline{\tilde{\mathrm{M}}}^{\bar{3}} = \mathcal{U}_{v}(\mathcal{W}(\underline{\tilde{\mathrm{M}}}^{\bar{4}}))$ and let $(\tilde{H}^{\bar{3}}, \tilde{G}^{\bar{3}}, *, \cdots *) := \underline{\tilde{\mathrm{M}}}^{\bar{3}}$. We can not apply Lemma 90 on $\underline{\tilde{\mathrm{M}}}^{\bar{3}}$ but we can apply Lemma 98 as $\underline{\tilde{\mathrm{M}}}^{\bar{3}}$ has wiggle-v room ϵ along $|n + 1\rangle$ (see Definition 93). To see this, note that the probability vectors had no component along $|n + 1\rangle$ and that we inverted the matrices using the flip map. This yields $Q^{\bar{3}} = \left|\tilde{u}^{\bar{3}}_{h}\right\rangle \left\langle \tilde{u}^{\bar{3}}_{g}\right| + Q^{\bar{2}}$. The lemma also lets us proceed by the application of the Wiggle-v Iteration map (see Definition 97) $\underline{\tilde{\mathrm{M}}}^{\bar{2}} = \mathcal{W}_v(\underline{\tilde{\mathrm{M}}}^{\bar{3}})$. Since at this point even the contact condition does not hold, we again apply the flip map and the Wiggle-w Initialization Map to obtain $\underline{\mathrm{M}}'^{\bar{2}} = \mathcal{U}_w(\mathcal{F}(\underline{\tilde{\mathrm{M}}}^{\bar{2}}))$. Instead of decreasing the power of x, the contact condition of this instance corresponds to increasing the power of x, i.e. the contact condition for $\underline{\mathrm{M}}'^{\bar{2}}$ corresponds to $\left\langle x^{2(k-1)+2+m+1}\right\rangle = 0$ which in turn holds if $2k+m+1 \le 2n-2$. Indeed, $0 = \cancel{2n-2}+2\lfloor -m/2\rfloor+m+1 \le \cancel{2n-2} = 0$ (substituting for $n = 5, k = 2, m = 3$ we get $8 = 2\cdot 2+3+1 \le 2\cdot 5-2 = 8$). Since $\underline{\mathrm{M}}'^{\bar{2}}$ has wiggle-w room ϵ along $|n+1\rangle$, we were justified at applying the Wiggle-w Initialization Map (see Lemma 98). This and the orthogonality of O, determine the form of $Q^{\bar{2}} = \left|u_h'^{\bar{2}}\right\rangle\left\langle u_g'^{\bar{2}}\right| + \left|u_h'^{\bar{1}}\right\rangle\left\langle u_g'^{\bar{1}}\right|$, which in turn completely determines the solution, O. □

For unbalanced monomial assignments, either there is a misalignment at the top or at the bottom. If the misalignment is at the top, it is easier to start by going downwards. To facilitate the tracking of powers in this case we need the following lemma—similar to Lemma 101— which can be proved by combining Lemma 164, Lemma 165 and Lemma 166 in Appendix I.

Lemma 103 (Down-then-Up Contact/Component Lemma). *Consider the matrix instance*

$$\underline{\tilde{M}}^{\bar{n}} := \mathcal{U}\left((H'^{\bar{n}})^{\dashv}, (G'^{\bar{n}})^{\dashv}, \left|w'^{\bar{n}}\right\rangle, \left|v'^{\bar{n}}\right\rangle, H'^{\bar{n}}, G'^{\bar{n}}, |.\rangle, |.\rangle\right).$$

Suppose the Normal Initialization Map and the Weingarten Iteration Map (see Definition 94 and Definition 95) are applied k times to obtain $\underline{\tilde{M}}^{\overline{n-k}}$. Let $n-k = d$ and consider $\underline{M}'^{\bar{d}} = \mathcal{U}(\mathcal{F}(\underline{\tilde{M}}^{\bar{d}}))$. Suppose the Normal Initialization Map and the Weingarten Iteration map are applied l more times to obtain $\underline{M}'^{\overline{d-l}} =: \left(H'^{\overline{d-l}}, G'^{\overline{d-l}}, \left|w'^{\overline{d-l}}\right\rangle, \left|v'^{\overline{d-l}}\right\rangle, *, \cdots *\right)$. *Then,*

$$\left\langle v'^{\overline{n-k-l}}\right|\left(G'^{\overline{n-k-l}}\right)^{\mu}\left|v'^{\overline{n-k-l}}\right\rangle = r\left(\left\langle (G'^{\bar{n}})^{-(2k+\mu)}\right\rangle, \dots, \left\langle (G'^{\bar{n}})^{2l+\mu-1}\right\rangle, \left\langle (G'^{\bar{n}})^{2l+\mu}\right\rangle\right),$$

where $\mu \ge 1$ and r is a multi-variate function which does not have an implicit dependence on $\left\langle (G'^{\bar{n}})^{i}\right\rangle := \left\langle v'^{\bar{n}}\right|(G'^{\bar{n}})^{i}\left|v'^{\bar{n}}\right\rangle$ for any i. The corresponding statement for H and $|w\rangle$ also holds.

The solution to the unbalanced monomial problem is as follows.

Proposition 104 (Solving the Unbalanced Monomial Problem). *Let*

$$t = \sum_{i=1}^{2n-1} -\frac{(-x_i)^m}{\prod_{j\neq i}(x_j - x_i)}\llbracket x_i\rrbracket = \sum_{i=1}^{n_h} x_{h_i}^m p_{h_i}\llbracket x_{h_i}\rrbracket - \sum_{i=1}^{n_g} x_{g_i}^m p_{g_i}\llbracket x_{g_i}\rrbracket$$

be an unbalanced monomial assignment over the real coordinates $0 < x_1 < x_2 \cdots < x_{2n-1}$ (see Definition 32) where $p_{h_i}, p_{g_i} > 0$, and $\{x_{h_i}\}$ and $\{x_{g_i}\}$ are all distinct. For $m = 0$ and $m = 2n-3$ the problem reduces to the f_0-assignment (see Proposition 99) using Corollary 158 in the latter case. For the remaining cases, consider the matrix instance $\underline{X}^{\bar{n}} := (X_h^{\bar{n}'}, X_g^{\bar{n}'}, (X_h^{\bar{n}'})^b|w\rangle, (X_g^{\bar{n}'})^b|v\rangle)$ where

- *if $n_h = n$ (the wiggle-v case; corresponds to odd m)*

$$X_h^{\bar{n}} \doteq diag(x_{h_1}, x_{h_2}\dots x_{h_{n-1}}, x_{h_n}), \qquad X_g^{\bar{n}} \doteq diag(x_{g_1}, x_{g_2}\dots x_{g_{n-1}}, \epsilon),$$
$$\left|w^{\bar{n}}\right\rangle \doteq (\sqrt{p_{h_1}}, \sqrt{p_{h_2}}\dots\sqrt{p_{h_{n-1}}}, \sqrt{p_{h_n}}), \qquad \left|v^{\bar{n}}\right\rangle \doteq (\sqrt{p_{g_1}}, \sqrt{p_{g_2}}\dots\sqrt{p_{g_{n-1}}}, 0),$$

- *else if $n_g = n$ (the wiggle-w case; corresponds to even m)*

$$X_h^{\bar{n}} \doteq diag(x_{h_1}, x_{h_2}\dots x_{h_{n-1}}, 1/\epsilon), \qquad X_g^{\bar{n}} \doteq diag(x_{g_1}, x_{g_2}\dots x_{g_{n-1}}, x_{g_n}),$$
$$\left|w^{\bar{n}}\right\rangle \doteq (\sqrt{p_{h_1}}, \sqrt{p_{h_2}}\dots\sqrt{p_{h_{n-1}}}, 0), \qquad \left|v^{\bar{n}}\right\rangle \doteq (\sqrt{p_{g_1}}, \sqrt{p_{g_2}}\dots\sqrt{p_{g_{n-1}}}, \sqrt{p_{g_n}}).$$

Consider the wiggle-v case. Let $k = \frac{2n-3-m}{2}$*. In the limit of* $\epsilon \longrightarrow 0$,

$$O = \sum_{i=n}^{n-k+1} \left| u_h'^{\bar{i}} \right\rangle \left\langle u_g'^{\bar{i}} \right| + \sum_{i=n-k}^{1} \left| \tilde{u}_h^{\bar{i}} \right\rangle \left\langle \tilde{u}_g^{\bar{i}} \right|$$

solves the matrix instance $\underline{X}^{\bar{n}}$ *where the terms in the sum are defined as follows. We start with* $\underline{M}'^{\bar{n}} := \mathcal{U}(\underline{X}^{\bar{n}} \oplus \left((X_h^{\bar{n}})^{-1}, (X_g^{\bar{n}})^{-1}, |.\rangle, |.\rangle\right)$ *(see Definition 94 and Definition 95) and using the relation*

$$\underline{M}'^{\overline{l-1}} := \mathcal{U}(\mathcal{W}(\underline{M}'^{\bar{l}})) \qquad \textit{for} \quad n-k+1 \leq l-1 \leq n-1,$$

we define

$$\left(*, \cdots *, \left| u_h'^{\bar{l}} \right\rangle, \left| u_g'^{\bar{l}} \right\rangle\right) := \underline{M}'^{\bar{l}} \qquad \textit{for} \quad n-k+1 \leq l \leq n.$$

These define the terms of the first sum. For the terms of the second sum we start with $\tilde{\underline{M}}^{\overline{n-k}} := \mathcal{U}(\mathcal{F}(\underline{M}'^{\overline{n-k}}))$ *and using the relation*

$$\tilde{\underline{M}}^{\overline{l-1}} := \mathcal{U}(\mathcal{W}(\tilde{\underline{M}}^{\bar{l}})) \qquad \textit{for} \quad 3 \leq l-1 \leq n-k-1,$$

we define

$$\left(*, \cdots *, \left| \tilde{u}_h^{\bar{l}} \right\rangle, \left| \tilde{u}_g^{\bar{l}} \right\rangle\right) := \tilde{\underline{M}}^{\bar{l}} \qquad \textit{for} \quad 2 \leq l \leq n-k.$$

Finally, we define (see Definition 96)

$$\tilde{\underline{M}}^{\bar{2}} := \mathcal{U}_v(\mathcal{W}(\tilde{\underline{M}}^{\bar{3}})) =: \left(*, *, \left| \tilde{w}^{\bar{2}} \right\rangle, \left| \tilde{v}^{\bar{2}} \right\rangle, * \cdots *\right),$$

$\left| \tilde{u}_h^{\bar{1}} \right\rangle := \left| e\left(\left| \tilde{u}_h^{\bar{2}} \right\rangle, \left| \tilde{w}^{\bar{2}} \right\rangle\right) \right\rangle$ *and* $\left| \tilde{u}_g^{\bar{1}} \right\rangle := \left| e\left(\left| \tilde{u}_g^{\bar{2}} \right\rangle, \left| \tilde{v}^{\bar{2}} \right\rangle\right) \right\rangle$.

Consider the wiggle-w case. Let $k = \frac{m}{2}$*. In the limit of* $\epsilon \longrightarrow 0$,

$$O = \sum_{i=n}^{n-k+1} \left| \tilde{u}_h^{\bar{i}} \right\rangle \left\langle \tilde{u}_g^{\bar{i}} \right| + \sum_{i=n-k}^{1} \left| u_h'^{\bar{i}} \right\rangle \left\langle u_g'^{\bar{i}} \right|$$

solves the matrix instance $\underline{X}^{\bar{n}}$ *where the terms in the sum are defined as follows. We start with*

$$\tilde{M}^{\bar{n}} := \mathcal{U}\left(\mathcal{F}\left(\underline{X}^{\bar{n}} \oplus \left((X_h^{\bar{n}})^{-1}, (X_g^{\bar{n}})^{-1}, |.\rangle, |.\rangle\right)\right)\right)$$

(see Definition 94, Definition 100 and Definition 95) and using the relation

$$\tilde{\underline{M}}^{\overline{l-1}} := \mathcal{U}(\mathcal{W}(\tilde{\underline{M}}^{\bar{l}}) \qquad \textit{for} \quad n-k+1 \leq l-1 \leq n-1,$$

we define

$$\left(*, \cdots *, \left| u_h'^{\bar{l}} \right\rangle, \left| u_g'^{\bar{l}} \right\rangle\right) := \tilde{\underline{M}}^{\bar{l}} \qquad \textit{for} \quad n-k+1 \leq l \leq n.$$

These determine the terms of the first sum. For the terms of the second sum we start with $\underline{M}'^{\overline{n-k}} := \mathcal{U}(\mathcal{F}(\underline{\tilde{M}}^{\overline{n-k}}))$ *and using*

$$\underline{M}'^{\overline{l-1}} := \mathcal{U}(\mathcal{W}(\underline{M}'^{\bar{l}})) \qquad \text{for} \quad 3 \leq l-1 \leq n-k-1,$$

we define

$$\left(*, \cdots *, \left|u_h'^{\bar{l}}\right\rangle, \left|u_g'^{\bar{l}}\right\rangle\right) := \underline{M}'^{\bar{l}} \qquad \text{for} \quad 2 \leq l \leq n-k.$$

Finally, we define (see Definition 96)

$$\underline{M}'^{\bar{2}} := \mathcal{U}_w(\mathcal{W}(\underline{M}'^{\bar{3}})) =: (*, *, \left|w'^{\bar{2}}\right\rangle, \left|v'^{\bar{2}}\right\rangle, * \cdots *),$$

$\left|u_h'^{\bar{1}}\right\rangle := \left|e\left(\left|u_h'^{\bar{2}}\right\rangle, \left|w'^{\bar{2}}\right\rangle\right)\right\rangle$ *and* $\left|u_g'^{\bar{1}}\right\rangle := \left|e\left(\left|u_g'^{\bar{2}}\right\rangle, \left|v'^{\bar{2}}\right\rangle\right)\right\rangle.$

Proof. From Figure 20 it is clear that the wiggle-v case is essentially the same as the balanced misaligned monomial until the second to last step (the wiggle-w step after wiggle-v is not needed). Furthermore, the wiggle-w case is essentially the same as the wiggle-v case except that we must start by going downwards, i.e. using $\underline{\tilde{M}}^{\bar{n}}$ and then flip to $\underline{M}'^{\bar{k}}$ to go upwards and end with a wiggle-w iteration. The arguments for the contact/component conditions go through unchanged using Lemma 103. □

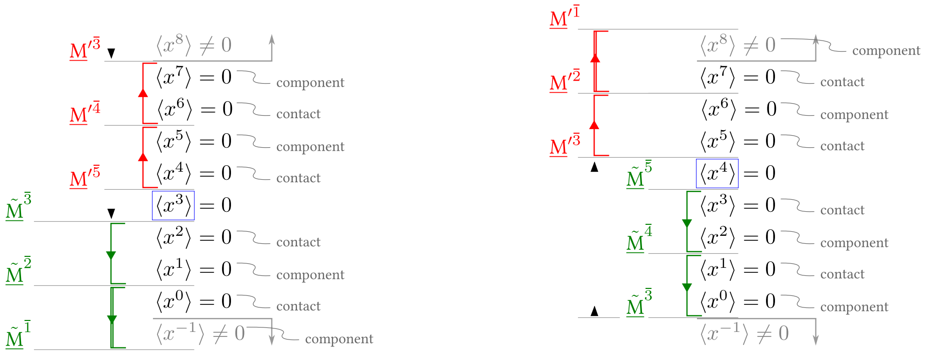


Figure 20: Power diagram representative of the unbalanced monomial assignment for $n = 4$ $(2n - 1 = 7)$ with $m = 3$ (left; wiggle-v case) and $m = 4$ (right; wiggle-w case).

Combining the results together, we can now formally state and prove the following theorem.

Theorem 105. *Let t be an f-assignment (see Definition 32) over strictly positive coordinates (without loss of generality; see Lemma 36). Suppose f has real and strictly positive roots. Decompose the f-assignment as $t = \sum_i \alpha_i t_i'$, where α_i are positive and t_i' are monomial assignments (see Definition 32 and Lemma 35). Then, each t_i' admits a solution (see Definition 34) given by Proposition 102 or Proposition 104 depending on its type; this is the effective solution (see Definition 34) to the f-assignment t.*

Proof. In Section 4.1 we established that in order to determine the solution to an f-assignment, it is sufficient to express it as a sum of monomial assignments and find the solution for each one of them. A monomial assignment is either balanced—in which case its solution is given by Proposition 102—or it is unbalanced—in which case its solution is given by Proposition 104. □

7 Future work

As we now have a construction for quantum WCF protocols with arbitrarily small biases, one can focus on the following aspects of the problem.

Optimality Various questions about the optimality of WCF protocols are unanswered.

- *C. Mochon's Games.* In Section 4, in order to find the solution to the f-assignment, we expressed it as a sum of monomial assignments; this yields an increase in dimensions, which in turn corresponds to an increase in the number of qubits required.[35] One approach towards reducing this, could be to understand the connection between the perturbatively defined unitary from Section 3 and the exact one in Section 4, corresponding to the 1/10-bias protocols. Another approach could be to try reducing the dimension using a standard technical lemma from [Moc07], which is stated as Lemma 146 here. In [Moc07], for converting a TIPG into a TDPG, the catalyst state is used. For the point game with $\epsilon(1) = 1/6$ bias we can, by inspection, obtain a TDPG which only requires both parties to hold qutrits locally and exchange one qubit at each round. However, if one uses the catalyst state approach, then the dimension of the space scales with the number rounds, which in turn diverges as the bias $\epsilon(1) = 1/6$ is approached. Can we convert C. Mochon's TIPG into a TDPG by harnessing this game structure? There is a notion of time-ordering at play—any TIPG in which points can be moved about without involving causal loops, can be easily converted into a TDPG. The challenge is to formalize this procedure and explore whether it can be used to lower bound the bias. The techniques recently introduced by C. Miller [Mil20] for proving a lower bound on the number of rounds needed in WCF protocols may be helpful in this respect.

- *E. Pelchat-P. Hø yer games.* E. Pelchat and P. Høyer [HP13] proposed another family of TIPGs which achieve arbitrarily low bias as well. It might be interesting to see if an explicit WCF protocol can be obtained corresponding to these games; hopefully in fewer dimensions. The construction is based on considering a combination of valid and invalid basic moves which together form valid non-trivial moves. The challenge is to find a decomposition such that each term stays valid and is still, only slightly non-trivial. The corresponding unitaries could perhaps be determined perturbatively, if they involve a constant number of points unlike the terms in the decomposition of the f-assignments.

- *Framework.* Constructing general tools to optimize and test the optimality of a TIPG for the number of points and rounds in the associated TDPG would be very useful to both constructing better protocols as well as benchmarking the existing ones. One way of doing this is related to the two general methods (see [Moc07] and Section 3), which are known for converting a TDPG into an explicit WCF protocol, granted that certain unitaries are known. Understanding how they compare and if they are optimal under an appropriately defined notion of optimality is useful. We already know that, in terms of the time duration for which the message register must be kept coherent, the recently introduced method of Section 3 is better. Nevertheless, in some cases, it is sub-optimal, as it fails to produce a 1/6-bias protocol from its TDPG which matches the resource usage of the 1/6 protocol given by C. Mochon. An obvious starting point could be to adapt the framework to work better in this particular case.

Relaxing assumptions The assumptions we made to obtain the protocols are not realistic.

- *System size.* The size of the incoming system containing the message is assumed to be known, however, this is hard to enforce physically. One way of dropping this assumption could be to adapt the

[35] The dimension of the Hilbert space is expected to scale exponentially with the number of points involved in the f-assignment.

following technique introduced in Ref [Him+17]—imposition of constraints on average energy for prepare-and-measure-like scenarios. The main challenge in our setting is that the parties do not trust each other. Nevertheless, the tools developed in [Him+17] should prove to be useful.

- *Noise.* Adding noise in a WCF protocol can cause a disagreement even when both parties are honest. It has been shown that in the absence of noise but in the present of losses, WCF can still be performed with a certain bias [Ber+09]. An interesting question is whether there exist lower bounds to the lossy but noiseless setting. One way of proceeding could be to generalize the A. Y. Kitaev/C. Mochon frameworks in order to handle an additional outcome corresponding to aborting the protocol, and to constraint the unitaries in such a way that the cheating player can control the losses. Even a preliminary understanding of this procedure should allow us to construct protocols with improved bias. One hurdle is the number of rounds which, in the loss tolerant protocol, varies depending on the strategy of the malicious player, while the A. Y. Kitaev/C. Mochon framework is designed for protocols with a constant number of rounds. Here one should be able to extend the notion of the "catalyst state". Quantum computation is realistic due to error correction. This, however, does not necessarily mean that WCF can be performed in such a setting, as it is not obvious how we can correct errors in this adversarial scenario without compromising the security. Consider the simplest error correcting code and the simplest WCF protocol. The honest case should work, but in the case where a malicious party is involved, the evaluation of the bias involves the communication of the syndrome, the error decoding and finally its correction by means of a unitary. These steps can be directly adapted into the A. Y. Kitaev/C. Mochon framework with the seemingly minor alteration that a malicious party can influence the unitary of the honest player in a way which is consistent with the noise model. The challenge here would be to make the security claim independent of the noise model. An appropriate relaxation of the constraints in the dual problem might be the key to this conundrum. Recently, generic techniques have been proposed to study adversarial cryptographic settings [GRS18] which might prove to be the right language for describing such a relaxation. One way of approaching the problem could be to further generalize the technique so that it facilitates the handling of noise without any error correction. This step itself should be of independent value as its results would serve as benchmarks against which error correction based schemes must be compared. A simultaneous but complementary approach could be to construct protocols which are robust against specific models of noise, such as those appearing in quantum optics. The insights from the two approaches should quicken the advance towards the final construction.

- *Device Dependence.* Device-independent WCF protocols have been suggested and involve the exchange of quantum boxes [Aha+14a]. Their bias, however, is abysmal and to date, no improvement has been reported and no lower bound on the bias is known. The first step could be to redefine the protocol in a generalizable way; perhaps construct successively worse protocols—by, for instance, using fewer boxes—and subsequently, consider them as belonging to the same family. One could try to use PR-boxes or non-signaling boxes to understand the behavior better. A complementary approach could be to construct the analogue of the A. Y. Kitaev/C. Mochon framework where instead of qubits and unitaries, one studies more abstract objects which simulate the exchange of boxes and are only constrained by their statistics. Recently, WCF protocols were also considered in the context of general probabilistic theories [SS19], that are used to extend the impossibility results theories beyond quantum. They used conic duality which is the key point of A. Y. Kitaev/C. Mochon frameworks and hence, this approach could be a starting point.

A fundamental connection It is known that nearly perfect WCF implies optimal strong coin flipping [CK09]. Does this work the other way around? This question may be more general than quantum, since the construction in [CK09] is purely classical. One way of proceeding could be to try and construct optimal strong coin flipping protocols directly by adapting the A. Y. Kitaev/C. Mochon technique and using known, simpler protocols as a starting point. The insight might not only help answer this question but also yield another construction for nearly perfect WCF.

Acknowledgements

This research was supported by the Belgian Fonds de la Recherche Scientifique - FNRS, under grant no R.50.05.18.F (QuantAlgo). The QuantAlgo project has received funding from the QuantERA European Research Area Network (ERA-NET) Cofund in Quantum Technologies implemented within the European Union's Horizon 2020 program. The most substantial part of the work was done while ASA was at the Université libre de Bruxelles, Belgium. He acknowledges support from the Belgian Fonds pour la Formation á la Recherche dans l'Industrie et dans l'Agriculture - FRIA, under grant no 1.E.081.17F. ASA also acknowledges funding provided by the Institute for Quantum Information and Matter. CV also acknowledges support from the SQIG-Security and Quantum Information Group. This work is funded by the FCT – Fundação para a Ciência e a Tecnologia through national funds FCT I.P. and, when eligible, by COMPETE 2020 FEDER funds, under the Scientific Employment Stimulus - Individual Call (CEEC Individual) 2020.03274.CEECIND/CP1621/CT0003 and Award UIDB/50008/2020. We are thankful to Nicolas Cerf, Tom Van Himbeeck, Kishor Bharti, Stefano Pironio and Ognyan Oreshkov for various insightful discussions.

References

[Aha+14a] Nati Aharon et al. "Weak Coin Flipping in a Device-Independent Setting." In: *Revised Selected Papers of the 6th Conference on Theory of Quantum Computation, Communication, and Cryptography - Volume 6745*. TQC 2011. Madrid, Spain: Springer-Verlag New York, Inc., 2014, pp. 1–12. ISBN: 978-3-642-54428-6. DOI: 10.1007/978-3-642-54429-3_1. URL: http://dx.doi.org/10.1007/978-3-642-54429-3_1.

[Aha+14b] Dorit Aharonov et al. "A simpler proof of existence of quantum weak coin flipping with arbitrarily small bias." In: *SIAM Journal on Computing* 45.3 (2014), pp. 633–679. DOI: 10.1137/14096387x. arXiv: 1402.7166.

[Amb04] Andris Ambainis. "A new protocol and lower bounds for quantum coin flipping." In: *Journal of Computer and System Sciences* 68.2 (2004), pp. 398–416. DOI: 10.1016/j.jcss.2003.07.010. arXiv: 0204022 [quant-ph].

[ARW18] Atul Singh Arora, Jérémie Roland, and Stephan Weis. *Weak Coin Flipping*. 2018. URL: https://atulsingharora.github.io/WCF.

[AS10] Nati Aharon and Jonathan Silman. "Quantum dice rolling: a multi-outcome generalization of quantum coin flipping." In: *New Journal of Physics* 12.3 (2010), p. 033027. DOI: 10.1088/1367-2630/12/3/033027.

[BB84] Charles H. Bennett and Gilles Brassard. "Public-Key Distribution and Coin Tossing." In: *Int. Conf. on Computers, Systems and Signal Processing*. 1984, pp. 175–179.

[Ber+09] Guido Berlín et al. "Fair loss-tolerant quantum coin flipping." In: *Physical Review A* 80.6 (2009). DOI: 10.1103/physreva.80.062321.

[Bha13] Rajendra Bhatia. *Matrix Analysis*. Springer New York, 2013. URL: https://www.ebook.de/de/product/25252147/rajendra_bhatia_matrix_analysis.html.

[Blu83] Manuel Blum. "Coin Flipping by Telephone a Protocol for Solving Impossible Problems." In: *SIGACT News* 15.1 (1983), pp. 23–27. ISSN: 0163-5700. DOI: 10.1145/1008908.1008911. URL: http://doi.acm.org/10.1145/1008908.1008911.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. DOI: 10.1017/cbo9780511804441.

[CGS13] André Chailloux, Gus Gutoski, and Jamie Sikora. "Optimal bounds for semi-honest quantum oblivious transfer." In: *Chicago Journal of Theoretical Computer Science, 2016* (2013). arXiv: 1310.3262v2. URL: http://arxiv.org/abs/1310.3262v2.

[CK09] André Chailloux and Iordanis Kerenidis. "Optimal Quantum Strong Coin Flipping." In: *50th FOCS*. 2009, pp. 527–533. DOI: 10.1109/FOCS.2009.71. arXiv: 0904.1511.

[CK11] André Chailloux and Iordanis Kerenidis. "Optimal Bounds for Quantum Bit Commitment." In: *52nd FOCS*. 2011, pp. 354–362. DOI: 10.1109/FOCS.2011.42. arXiv: 1102.1678.

[CKS13] André Chailloux, Iordanis Kerenidis, and Jamie Sikora. "Lower bounds for Quantum Oblivious Transfer." In: *Quantum Information & Computation* 13.1-2 (2013), pp. 158–177. arXiv: 1007.1875.

[Col07] Roger Colbeck. "Impossibility of secure two-party classical computation." In: *Phys. Rev. A* 76 (6 2007), p. 062308. DOI: 10.1103/PhysRevA.76.062308. URL: https://link.aps.org/doi/10.1103/PhysRevA.76.062308.

[Fri18] Tobias Fritz. *Does the set of operator monotone functions become larger if we restrict ourselves to real symmetric matrices?* 2018. URL: https://mathoverflow.net/questions/298359/does-the-set-of-operator-monotone-functions-become-larger-if-we-restrict-ourselv.

[Gan09] Maor Ganz. "Quantum Leader Election." In: (2009). arXiv: 0910.4952v2. URL: https://arxiv.org/abs/0910.4952v2.

[GRS18] Gus Gutoski, Ansis Rosmanis, and Jamie Sikora. "Fidelity of quantum strategies with applications to cryptography." In: *Quantum* 2 (2018), p. 89. DOI: 10.22331/q-2018-09-03-89.

[Hag89] William W. Hager. "Updating the Inverse of a Matrix." In: *SIAM Review* 31.2 (1989), pp. 221–239. DOI: 10.1137/1031049.

[Him+17] Thomas Van Himbeeck et al. "Semi-device-independent framework based on natural physical assumptions." In: *Quantum* 1 (2017), p. 33. DOI: 10.22331/q-2017-11-18-33.

[HP13] Peter Høyer and Edouard Pelchat. "Point Games in Quantum Weak Coin Flipping Protocols." MA thesis. University of Calgary, 2013. URL: http://hdl.handle.net/11023/873.

[Kil88] Joe Kilian. "Founding Crytpography on Oblivious Transfer." In: *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*. STOC '88. Chicago, Illinois, USA: Association for Computing Machinery, 1988, pp. 20–31. ISBN: 0897912640. DOI: 10.1145/62212.62215. URL: https://doi.org/10.1145/62212.62215.

[Kit03] Alexei Kitaev. "Quantum coin flipping." Talk at the 6th workshop on Quantum Information Processing. 2003.

[KN04] Iordanis Kerenidis and Ashwin Nayak. "Weak coin flipping with small bias." In: *Information Processing Letters* 89.3 (2004), pp. 131–135. DOI: 10.1016/j.ipl.2003.07.007.

[Lo97] Hoi-Kwong Lo. "Insecurity of quantum secure computations." In: *Phys. Rev. A* 56 (2 1997), pp. 1154–1162. DOI: 10.1103/PhysRevA.56.1154. URL: https://link.aps.org/doi/10.1103/PhysRevA.56.1154.

[Mil20] Carl A. Miller. "The Impossibility of Efficient Quantum Weak Coin Flipping." In: *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*. New York, NY, USA: Association for Computing Machinery, 2020, pp. 916–929. ISBN: 9781450369794. URL: https://doi.org/10.1145/3357713.3384276.

[Moc05] Carlos Mochon. “Large family of quantum weak coin-flipping protocols.” In: *Phys. Rev. A* 72 (2005), p. 022341. DOI: 10.1103/PhysRevA.72.022341. arXiv: 0502068 [quant-ph].

[Moc07] Carlos Mochon. “Quantum weak coin flipping with arbitrarily small bias.” In: *arXiv:0711.4114* (2007). arXiv: 0711.4114.

[NS03] Ashwin Nayak and Peter Shor. “Bit-commitment-based quantum coin flipping.” In: *Phys. Rev. A* 67 (1 2003), p. 012304. DOI: 10.1103/PhysRevA.67.012304. URL: https://link.aps.org/doi/10.1103/PhysRevA.67.012304.

[NST14] Ashwin Nayak, Jamie Sikora, and Levent Tunçel. “A search for quantum coin-flipping protocols using optimization techniques.” In: *Mathematical Programming* 156.1-2 (2014), pp. 581–613. DOI: 10.1007/s10107-015-0909-y. arXiv: 1403.0505.

[NST15] Ashwin Nayak, Jamie Sikora, and Levent Tunçel. “Quantum and classical coin-flipping protocols based on bit-commitment and their point games.” In: (2015). arXiv: 1504.04217v1. URL: http://arxiv.org/abs/1504.04217v1.

[Sch09] Rolf Schneider. *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge University Press, 2009. DOI: 10.1017/cbo9781139003858.

[Sho94] Peter W. Shor. “Algorithms for quantum computation: discrete logarithms and factoring.” In: *Proceedings 35th Annual Symposium on Foundations of Computer Science*. IEEE Comput. Soc. Press, 1994. DOI: 10.1109/sfcs.1994.365700.

[SM50] Jack Sherman and Winifred J. Morrison. “Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix.” In: *The Annals of Mathematical Statistics* 21.1 (1950), pp. 124–127. DOI: 10.1214/aoms/1177729893.

[SR02] Robert W. Spekkens and Terry Rudolph. “Quantum Protocol for Cheat-Sensitive Weak Coin Flipping.” In: *Phys. Rev. Lett. vol 89, 227901 (2002)* 89.22 (Feb. 21, 2002). DOI: 10.1103/PhysRevLett.89.227901. arXiv: quant-ph/0202118v2 [quant-ph].

[SS19] Jamie Sikora and John H. Selby. “On the impossibility of coin-flipping in generalized probabilistic theories via discretizations of semi-infinite programs.” In: (2019). arXiv: 1901.04876 [quant-ph].

A Connection with conic duality

In Section 2, we showed the existence of quantum WCF protocols with arbitrarily small biases, however, we took the characterization of EBM transitions on faith (see Proposition 21). Here, we present the analysis leading to this alternative characterization of EBM transitions. We see that the set of EBM functions is a convex cone, and the dual of this cone happens to be the set of operator monotone functions which have a surprisingly elegant and simple characterization. To harness this, we use the known fact that for a closed convex cone, the dual of the dual is the original cone itself (also called a bi-dual). So, the bi-dual of the cone of EBM functions equals the cone of EBM functions (up to closures). The dual of operator monotone functions has an easy description because operator monotone functions have an easy description. Combining these, we obtain a more useful characterization of EBM functions. This result was first presented by C. Mochon and A. Y. Kitaev, but it was proved using matrix perturbation theory [Moc07]. The argument we just sketched, however, was also outlined therein. In this section, we present the approach of D. Aharonov, A. Chailloux, M. Ganz, I. Kerenidis and L. Magnin [Aha+14b] who worked out a simpler proof, along the lines alluded to by C. Mochon and A. Y. Kitaev. The proofs of all the results we present here (unless referred otherwise) can be found in [Aha+14b].

A.1 Formalizing the equivalence between transitions and functions

Working with functions instead of transitions is rather useful as it will become evident in the course of this analysis.

Definition 106 (K, EBM functions)**.** A function $a : [0,\infty) \to \mathbb{R}$ with finite support is an *EBM function* if the line transition $a^- \to a^+$ is EBM (see Definition 15), where $a^+ : [0,\infty) \to [0,\infty)$ and $a^- : [0,\infty) \to [0,\infty)$ denote, respectively, the positive and the negative part of a (i.e. $a = a^+ - a^-$ with $\text{supp}(a^+) \cap \text{supp}(a^-) = \phi$ and $a^\pm \geq 0$).

We denote by K the set of EBM functions.

Definition 107 (K_Λ, EBM functions on $[0,\Lambda]$)**.** For any finite Λ, a function $a : [0,\Lambda) \to \mathbb{R}$ with finite support is an *EBM function with support on* $[0,\Lambda]$ if the line transition $a^- \to a^+$ is EBM with its spectrum in $[0,\Lambda]$, where $a^- : [0,\Lambda)[0,\infty)$ and $a^+ : [0,\Lambda) \to [0,\infty)$ denote, respectively, the positive and the negative part of a.

We denote the set of EBM functions with support on $[0,\Lambda]$ by K_Λ.

If the functions g, h denoting the transition $g \to h$ have no common support, then the function description uniquely captures the said transition. In this section we restrict to such transitions and therefore use them (i.e. functions and transitions) interchangeably.

It is useful to abstract the different characterizations of EBM functions into a property $\mathcal{P}$ that the function must satisfy, and we can define games that use these $\mathcal{P}$-functions. This facilitates the handling of subtleties which arise in proving that the set of EBM functions is the same as the set of $\mathcal{P}$-functions for specific $\mathcal{P}$s.

Definition 108 (Horizontal and vertical $\mathcal{P}$-functions)**.** A $\mathcal{P}$-function $a : [0,\infty) \to \mathbb{R}$ is a function with finite support that has the property $\mathcal{P}$. A function $t : [0,\infty) \times [0,\infty) \to \mathbb{R}$ is a

- *horizontal $\mathcal{P}$-function* if for all $y \geq 0$, $t(.,y)$ is a $\mathcal{P}$-function;
- *vertical $\mathcal{P}$-function* if for all $x \geq 0$, $t(x,.)$ is a $\mathcal{P}$-function.

Suppose $\mathcal{P}$-functions are EBM functions. Consider a TDPG given by the following sequence of valid transitions

$$t_0 = p_0 = \frac{1}{2}\left(\llbracket 0,1 \rrbracket + \llbracket 1,0 \rrbracket\right) \to p_1 \to p_2 \cdots \to p_n = \llbracket \beta,\alpha \rrbracket$$

and define $t_1 = p_1 - p_0,\ t_2 = p_2 - p_1,\ \dots t_i = p_i - p_{i-1}$. Clearly, $p_1 = t_1 + t_0,\ p_2 = t_2 + \underbrace{t_1 + t_0}_{p_1},\ \dots p_j = \sum_{i=1}^{j} t_i$.

This effectively shows how one can construct a TDPG consisting of valid functions, $\{t_i\}$, instead of valid transitions, and motivates the following definition:

Definition 109 (point games with $\mathcal{P}$-functions)**.** A point game with $\mathcal{P}$-functions is a set $\{t_1, \dots, t_n\}$ of n $\mathcal{P}$-functions alternatively horizontal and vertical such that

- $\frac{1}{2}\llbracket 0, 1\rrbracket + \frac{1}{2}\llbracket 1, 0\rrbracket + \sum_{i=1}^{n} t_i = \llbracket \beta, \alpha\rrbracket$;
- $\forall j \in \{1, \dots, n\}$, $\frac{1}{2}\llbracket 0, 1\rrbracket + \frac{1}{2}\llbracket 1, 0\rrbracket + \sum_{i=1}^{j} t_i \geq 0$.

We call $\llbracket \beta, \alpha\rrbracket$ the final point of the game.

The first condition simply encodes the initial and final frame configurations while the second condition ensures that the ‘‘p_is’’ are non-negative, i.e. each intermediate frame configuration is sensible.

Lemma 110 ((time-dependent) point game with EBM functions $\implies$ (time-dependent) point game with EBM transitions)**.** *Given a TDPG with n EBM functions and final point $\llbracket \beta, \alpha\rrbracket$ we can construct a TDPG with n EBM transitions and final point $\llbracket \beta, \alpha\rrbracket$.*

A.2 Operator monotone functions and valid functions

Let us start with the definition of a convex cone.

Definition 111 (convex cone)**.** A set C in a vector space V is a cone if for all $x \in C$ and for all $\lambda > 0$, $\lambda x \in C$. It is convex if for all $x, y \in C$, $x + y \in C$.

Noting that the state $|\psi\rangle$ in the definition of an EBM function is unnormalized, the set of EBM functions is easily seen to form a cone. By taking a direct sum we can establish convexity as well.

Let V be a set of vectors where the vectors are functions from $[0, \infty) \rightarrow [0, \infty)$ with finite support.

- V is an infinite dimensional vector space spanned by $\{\llbracket x\rrbracket\}_{x\in[0,\infty)}$, since we can express each element of V as $v = \sum_x v(x)\llbracket x\rrbracket$ where the sum is over the finite support[36] of v.
- The norm is given as $\|v\| := \|v\|_1 = \sum_x |v(x)|$.

Lemma 112. *K is a convex cone. Also, for any $\Lambda \in (0, \infty)$, K_Λ is a convex cone.*

Definition 113 (dual cone)**.** Let C be a cone in a normed vector space V. We denote by V' the space of continuous linear functionals from V to $\mathbb{R}$. The dual cone of a set $C \subseteq V$ is

$$C^* = \{\Phi \in V' | \forall a \in C, \Phi(a) \geq 0\}\,.$$

For our purpose, linear functionals can be thought of simply as functions which map objects in the cone to a non-negative real number with the added property of being linear in its argument. We can now formally define operator monotone functions.

Definition 114 (operator monotone functions)**.** A function $f : [0, \infty) \rightarrow \mathbb{R}$ is operator monotone if for all $0 \leq X \leq Y$ we have $f(X) \leq f(Y)$.

[36] If instead of a sum, we had used an integral, we would have had to use a Dirac delta function/distribution. However, restricting to finitely supported functions suffices for our purpose.

Definition 115 (operator monotone functions on $[0, \Lambda]$)**.** A function $f : [0, \Lambda] \to \mathbb{R}$ is operator monotone on $[0, \Lambda]$ if for all $0 \le X \le Y$ with spectrum in $[0, \Lambda]$ we have $f(X) \le f(Y)$.

The pivotal result here is the equivalence between the cone of operator monotone functions and the dual cone of EBM functions. To state this formally, we consider the following isomorphism. There is a bijective mapping between $\Phi \in V'$ (the space of linear functionals; see Definition 113) and f_Φ which is defined as $f_\Phi(x) = \Phi(\llbracket x \rrbracket)$. By linearity, for any $h = \sum_x h(x) \llbracket x \rrbracket$ we have $\Phi(\sum_x h(x) \llbracket x \rrbracket) = \sum_x h(x) f_\Phi(x)$. We can therefore see elements of the dual cone as functions on real numbers.

Lemma 116. *$\Phi \in K^*$, the dual to the set of EBM functions, if and only if f_Φ is operator monotone in $[0, \infty]$. Also, for any $\Lambda \in (0, \infty)$, $\Phi \in K^*_\Lambda$ if and only if f_Φ is operator monotone on $[0, \Lambda]$.*

Lemma 117 (Characterization of operator monotone functions [Bha13])**.** *Any operator monotone function $f : [0, \infty) \to \mathbb{R}$ can be written as*

$$f(x) = c_0 + c_1 x + \int_0^\infty \frac{\lambda x}{\lambda + x} d\omega(\lambda),$$

for a measure ω satisfying $\int_0^\infty \frac{\lambda}{1+\lambda} d\omega(\lambda) < \infty$.

Lemma 118 (characterization of operator monotone functions on $[0, \Lambda]$ [Bha13])**.** *Any operator monotone function $f : [0, \Lambda] \to \mathbb{R}$ can be written as*

$$f(x) = c_0 + c_1 x + \int \frac{\lambda x}{\lambda + x} d\omega(\lambda)$$

with the integral ranging over $\lambda \in (-\infty, -\Lambda) \cup (0, \infty)$ satisfying $\int \frac{\lambda}{1+\lambda} d\omega(\lambda) < \infty$ where ω is a measure.

As will become clear when we discuss the dual of the cone of operator monotones, it suffices to consider the extremal rays of the cone, i.e. operator monotones of the form $\lambda x/(\lambda + x)$ (together with 1 and x).

It is known that the bi-dual of a cone is the closure of the cone itself (see [BV04] for details and the proof):

Fact 119. *Let $C \subseteq V$ be a convex cone, then $C^{**} = cl(C)$ where C^* is the dual cone of C.*

Essentially, we define from hindsight the bi-dual of EBM functions to be the cone of valid functions. Since the dual of EBM functions has an easy characterization, the bi-dual also has an easy characterization which is why we are interested in it.

Definition 120 (Λ-valid functions)**.** A function $a : [0, \Lambda] \to \mathbb{R}$ with finite support on $[0, \Lambda]$ is Λ-valid if $a \in K^{**}_\Lambda$.

To be able to use the aforementioned fact we note that the cone of interest, the cone of EBM functions, is closed when the matrices involved have a bounded spectrum. In this case, it means that the cone of valid functions is the same as the cone of EBM functions. We state this precisely below.

Lemma 121. *For $\Lambda \in (0, \infty)$, K_Λ is closed (which implies $K^{**}_\Lambda = K_\Lambda$).*

Corollary 122. *For $\Lambda \in (0, \infty)$, $K_\Lambda = \{a \in V | \forall \Phi \in K^*_\Lambda, \Phi(a) \ge 0\}$. Further, $a \in K_\Lambda$ if and only if $\sum_x a(x) = 0$, $\sum_x x a(x) \ge 0$ and $\forall \lambda \in (-\infty, -\Lambda] \cup (0, \infty)$, $\sum_x \frac{\lambda x}{\lambda + x} a(x) \ge 0$.*

A seemingly cumbersome but useful restatement of this result is the following:

Corollary 123 (EBM on $[0, \Lambda]$ is equivalent to Λ valid)**.** *A function $a : [0, \Lambda] \to \mathbb{R}$ with finite support on $[0, \Lambda]$ is EBM on $[0, \Lambda]$ if and only if a is Λ-valid, i.e., it satisfies $\sum_x a(x) = 0$, $\sum_x x a(x) \ge 0$ and $\forall \lambda \in (-\infty, -\Lambda] \cup (0, \infty)$, $\sum_x \frac{\lambda x}{\lambda + x} a(x) \ge 0$.*

All the statements made here assume that the matrices used in EBM functions have a finite spectrum. Our EMA algorithm (see Section 5) heavily relies on this part of the analysis of [Aha+14b].

A.3 Strictly valid functions are EBM functions

To be able to simplify the conditions one needs to check, it is useful to remove the condition on the spectrum of the positive semi-definite matrices involved. This is evident from the range of λ one needs to use in the characterization of operator monotone functions (compare Lemma 118 and Lemma 117).

It is easy to describe the interior of the dual of a cone. It is also possible to relate the interior with the closure of the cone in finite dimensions[37](see [BV04] for details and proofs).

Fact 124. *Let C be a convex set, then $int(C) = int(cl(C))$.*

Fact 125. *Let C be a cone in the finite-dimensional vector space V.*
Then $int(C^*) = \{\Phi \in V' | \forall a \in C \backslash \{0\}, \Phi(a) > 0\}$.

It turns out that K is not closed[38], and recall that K^* is the set of operator monotone functions on $[0, \infty)$. Recall also that $K^{**} = \mathrm{cl}(K)$. Using Fact 124 we conclude that $\mathrm{int}(K^{**}) = \mathrm{int}(K)$. In finite dimensions, restricting to the interior of K^{**} is easy, as stated in Fact 125. We could then simply consider points in the interior of K^* which in turn would guarantee membership in K. For infinite dimensions this result continues to hold. To see this, we define valid and strictly valid functions.

Definition 126 (valid function)**.** A function $a : [0, \infty) \to \mathbb{R}$ with finite support is valid if for every operator monotone function $f : [0, \infty) \to \mathbb{R}$ we have $\sum_{x \in \mathrm{supp}(h)} f(x)a(x) \geq 0$.

Definition 127 (strictly valid function)**.** A function $a : [0, \infty) \to \mathbb{R}$ with finite support is strictly valid if for every non-constant operator monotone function $f : [0, \infty) \to \mathbb{R}$ we have $\sum_{x \in \mathrm{supp}(a)} f(x)a(x) > 0$.

One can use the characterization of operator monotone functions to explicitly characterize the set of valid and strictly valid functions (just as we did for Λ-valid functions; see Corollary 122).

Lemma 128. *Let $a : [0, \infty) \to \mathbb{R}$ be a function with finite support such that $\sum_x a(x) = 0$. The function a is a strictly valid function if and only if for all $\lambda > 0$, $\sum_x \frac{-a(x)}{\lambda + x} > 0$ and $\sum_x x.a(x) > 0$.*

The function a is valid if and only if for all $\lambda > 0$, $\sum_x \frac{-a(x)}{\lambda + x} \geq 0$ and $\sum_x x.a(x) \geq 0$.

The set of strictly valid functions can be also shown to be Λ-valid for some finite Λ. This means that it would also be EBM on $[0, \Lambda]$ which in turn means it would be an EBM function. We hence have the following.

Lemma 129. *Any strictly valid function is an EBM function.*

Similarly we define valid and strictly valid *transitions*.

Definition 130 (Valid and strictly valid line transitions)**.** Let $g, h : [0, \infty) \to \mathbb{R}$ be two functions with finite support. The transition $g \to h$ is valid (resp., strictly valid) if the function $h - g$ is valid (resp., strictly valid).

[37] This reasoning fails for infinite dimensions.

[38] One example is the merge function. Let $p_{g_1} + p_{g_2} = 1$, $x_{g_2} > x_{g_1} > 0$. Consider the sequence $t_1, t_2 \dots t_k$ where $t_k := \llbracket \langle x_g \rangle + \frac{1}{k} \rrbracket - p_{g_1} \llbracket x_{g_1} \rrbracket - p_{g_2} \llbracket x_{g_2} \rrbracket$. This sequence, in the limit $k \to \infty$ is just a merge. One can show that for any finite k, t_k can be shown to be EBM using matrices with a finite spectrum (this is because for a $2 \to 1$ transition, it suffices to restrict to 2×2 matrices (see Lemma 146) and then one can consider the most general unitary to reach the conclusion). However, as $k \to \infty$, the spectra of the matrices involved diverges. Thus, while the elements of the sequence $t_1, t_2 \dots t_k \dots$ are contained in K, its limit is not. This argument does not apply to K_Λ (confirming the fact that K_Λ is closed) because after some finite k, t_k ceases to be in K_Λ.

A.4 From point games with valid functions to point games with EBM functions

If we construct a point game with valid functions we can convert it into a game with EBM functions with an arbitrarily small overhead on the bias. The trick is to raise the coordinates of all the final points (ones with positive weight) a little at each step, to convert a valid function into a strictly valid function.

Theorem 131 (valid to EBM)**.** *Given a point game with $2m$ valid functions and final point $[\![\beta, \alpha]\!]$ and any $\epsilon > 0$, we can construct a point game with $2m$ EBM functions and final point $[\![\beta + \epsilon, \alpha + \epsilon]\!]$.*

Lemma 132. *Fix $\epsilon > 0$. Given a point game with $2m$ valid functions and final point $[\![\beta, \alpha]\!]$ we can construct a point game with $2m$ strictly valid functions and final point $[\![\beta + \epsilon, \alpha + \epsilon]\!]$.*

B TEF functions = valid functions = closure of EBM functions

Let the set of TEF functions be the set of finitely supported functions $t = h - g$, such that for the associated transition $h \longrightarrow g$, the conditions in Theorem 31 can be satisfied for some unitary U. An equivalent definition would be to require the transition to admit a legal CPF (see Definition 44 in Section 5). We assume that h and g are non-negative functions without common support. Then, the following lemma holds:

Lemma 133 (TEF = closure of EBM = valid)**.** *The set of the TEF functions, the set of valid functions (see Definition 126) and the closure of the set of the EBM functions (see Definition 106) are the same.*

Proof sketch. We start by recalling that the set of EBM functions is an open set. From Definition 106 we can see that the matrix H may have eigenvectors which have no support on $|\psi\rangle$. Consequently, one can consider a sequence of EBM functions t_i such that the $\lim_{i\to\infty} t_i = t$ is well-defined, while the associated matrix $\lim_{i\to\infty} H_i$ has a diverging eigenvalue. Such a case arises, for instance, when we have a merge move in the point game. For concreteness, let x_{g_1}, x_{g_2} be the coordinates of two points that are going to be merged into a single point with coordinate $x_h = p_{g_1}x_{g_1} + p_{g_2}x_{g_2}$, and let p_{g_1}, p_{g_2} be their respective probability weights, with $p_{g_1} + p_{g_2} = 1$. Furthermore, let $t_i = [\![x_h + 1/i]\!] - p_{g_1}[\![x_{g_1}]\!] - p_{g_2}[\![x_{g_2}]\!]$. One can verify that for all finite values of i, t_i is EBM, but its limit $t = [\![x_h]\!] - p_{g_1}[\![x_{g_1}]\!] - p_{g_2}[\![x_{g_2}]\!]$ is not EBM (we omit the details for the sake of brevity), thus concluding that the set of EBM functions is open.

To show that the closure of this set is the same as the set of the TEF functions, we need to establish that the limit of any such sequence belongs to the set of TEF functions. This requires a combination of certain results from Section 5. In particular, the relationship between the COF and the CPF permits one to trade the divergence of such a matrix H for appropriate projectors. This is exactly the origin of the projectors E_h that appear in our analysis. The matrices $H \geq G$ and the vector $|\psi\rangle$ corresponding to an EBM transition, can be expressed in the COF,[39] $X_h \geq OX_gO^T$. Essentially, the same orthogonal matrix O also satisfies the TEF inequality Equation (3)[40]. The TEF inequality may, in fact, be seen as the limit where H's eigenvalues diverge to infinity. Thus, the limit t of the sequence t_i indeed belongs to the set of TEF functions and this argument readily extends to all relevant sequences.

Finally, in Appendix A we described how the authors of [Aha+14b] prove that the set of valid functions is the same as the closure of the set of EBM functions. In particular, they start by observing that the set of EBM functions is a convex cone K, and its dual cone K^* is the set of operator monotone functions. The bi-dual K^{**} is the set of valid functions, and the fact that $K^{**} = \text{cl}(K)$ completes the proof. Since we just showed that the closure of the set of EBM functions is the same as the set of TEF functions, we can also conclude that the set of valid functions is the same as the set of TEF functions. □

[39] Recall that X_h and X_g are diagonal matrices containing the eigenvalues of H and G, respectively (in addition to X_h possibly having a large eigenvalue with multiplicities and X_h possibly having zero eigenvalues)

[40] Observe that the TEF inequality is closely related to the CPF.

C Blink $m \to n$ transition

C.1 Completing an orthonormal basis

Consider an orthonormal complete set of basis vectors $\{|g_i\rangle\}$ and a vector $|v\rangle = \frac{\sum_i \sqrt{p_i}|g_i\rangle}{\sqrt{\sum_i p_i}}$. We describe a scheme for constructing vectors $|v_i\rangle$ such that $\{|v\rangle, \{|v_i\rangle\}\}$ is a complete orthonormal set of basis vectors. We can do it inductively, but here instead we choose to do it by examples, as we believe it helps gain some intuition and demonstrates the generalizable argument right away. We define the first vector to be

$$|v_1\rangle = \frac{\sqrt{p_1}\,|g_1\rangle - \frac{p_1}{\sqrt{p_2}}\,|g_2\rangle}{\sqrt{p_1 + \frac{p_1^2}{p_2}}} = \frac{\sqrt{p_1}\,|g_1\rangle - \sqrt{p_2}\,|g_2\rangle}{\sqrt{p_1 + p_2}},$$

which is normalized and orthogonal to $|v\rangle$. The next vector is

$$|v_2\rangle = \frac{\sqrt{p_1}\,|g_1\rangle + \sqrt{p_2}\,|g_2\rangle - \frac{(p_1+p_2)}{\sqrt{p_3}}\,|g_3\rangle}{\sqrt{p_1 + p_2 + \frac{(p_1+p_2)^2}{p_3}}}$$

which is again normalized and orthogonal to $|v_1\rangle$.

Similarly we can construct the $(k+1)^{\text{th}}$ basis vector as

$$|v_k\rangle = \frac{\sum_{i=1}^{k} \sqrt{p_k}\,|g_k\rangle - \frac{\sum_{i=1}^{k} p_k}{\sqrt{p_{k+1}}}\,|g_{k+1}\rangle}{N_k},$$

where $N_k = \sqrt{\sum_{i=1}^{k} p_k + \frac{(\sum_{i=1}^{k} p_k)^2}{p_{k+1}}}$ and, thus, obtain the full set.

C.2 Analysis of the $3 \to 2$ transition

Recall that the constraint equation is

$$\underbrace{\sum x_{h_i}\,|h_{ii}\rangle\,\langle h_{ii}|}_{\text{I}} + \underbrace{x\mathbb{I}^{\{g_{ii}\}}}_{\text{II}} \geq \underbrace{\sum x_{g_i} U\,|g_{ii}\rangle\,\langle g_{ii}|\,U^{\dagger}}_{\text{III}},$$

where we have introduced the notation $|h_{ii}\rangle = |h_i h_i\rangle$. The $g_1, g_2, g_3 \to h_1, h_2$ transition requires us to know

$$U = |v\rangle\,\langle w| + |w\rangle\,\langle v| + |v_1\rangle\,\langle v_1| + |v_2\rangle\,\langle v_2| + |w_1\rangle\,\langle w_1|\,.$$

Using the procedure above we can evaluate the vectors of interest as

$$|v\rangle = \frac{\sqrt{p_{g_1}}\,|g_{11}\rangle + \sqrt{p_{g_2}}\,|g_{22}\rangle + \sqrt{p_{g_3}}\,|g_{33}\rangle}{N_g}, \quad |v_1\rangle = \frac{\sqrt{p_{g_1}}\,|g_{11}\rangle - \frac{p_{g_1}}{\sqrt{p_{g_2}}}\,|g_{22}\rangle}{N_{g_1}},$$

$$|v_2\rangle = \frac{\sqrt{p_{g_1}}\,|g_{11}\rangle + \sqrt{p_{g_2}}\,|g_{22}\rangle - \frac{(p_{g_1}+p_{g_2})}{\sqrt{p_{g_3}}}\,|g_{33}\rangle}{N_{g_2}},$$

$$|w\rangle = \frac{\sqrt{p_{h_1}}\,|h_{11}\rangle + \sqrt{p_{h_2}}\,|h_{22}\rangle}{N_h} \quad \text{and} \quad |w_1\rangle = \frac{\sqrt{p_{h_2}}\,|h_{11}\rangle - \sqrt{p_{h_1}}\,|h_{22}\rangle}{N_h},$$

where N_g, N_{g_1}, N_{g_2}, N_h are normalization factors. In fact we want to express the constraints in this basis, and to evaluate the first term of the LHS in the constraint equation we use the above to find

$$|h_{11}\rangle = \frac{\sqrt{p_{h_1}}\,|w\rangle + \sqrt{p_{h_2}}\,|w_1\rangle}{N_h} \quad \text{and} \quad |h_{22}\rangle = \frac{\sqrt{p_{h_2}}\,|w\rangle - \sqrt{p_{h_1}}\,|w_1\rangle}{N_h},$$

which leads to

$$\begin{aligned} \mathrm{I} &= x_{h_1}\,|h_{11}\rangle\langle h_{11}| + x_{h_2}\,|h_{22}\rangle\langle h_{22}| \\ &= \frac{1}{N_h^2}\left[\begin{array}{c|cc} & \langle w| & \langle w_1| \\ \hline |w\rangle & p_{h_1}x_{h_1} + p_{h_2}x_{h_2} & \sqrt{p_{h_1}p_{h_2}}(x_{h_1} - x_{h_2}) \\ |w_1\rangle & \sqrt{p_{h_1}p_{h_2}}(x_{h_1} - x_{h_2}) & p_{h_2}x_{h_1} + p_{h_1}x_{h_2} \end{array}\right]. \end{aligned}$$

Evaluation of II is nearly trivial after expressing the identity in this basis

$$\mathrm{II} = x(|v\rangle\langle v| + |v_1\rangle\langle v_1| + |v_2\rangle\langle v_2|) = \left[\begin{array}{c|ccc} & \langle v| & \langle v_1| & \langle v_2| \\ \hline |v\rangle & x & & \\ |v_1\rangle & & x & \\ |v_2\rangle & & & x \end{array}\right].$$

For the last term $\mathrm{III} = \underbrace{x_{g_1}U\,|g_{11}\rangle\langle g_{11}|\,U^\dagger}_{\text{(i)}} + \underbrace{x_{g_2}U\,|g_{22}\rangle\langle g_{22}|\,U^\dagger}_{\text{(ii)}} + \underbrace{x_{g_3}U\,|g_{33}\rangle\langle g_{33}|\,U^\dagger}_{\text{(iii)}}$, we evaluate

$$\begin{aligned} U\,|g_{11}\rangle &= \frac{\sqrt{p_{g_1}}}{N_g}\,|w\rangle + \frac{\sqrt{p_{g_1}}}{N_{g_1}}\,|v_1\rangle + \frac{\sqrt{p_{g_1}}}{N_{g_2}}\,|v_2\rangle, \\ U\,|g_{22}\rangle &= \frac{\sqrt{p_{g_2}}}{N_g}\,|w\rangle + \frac{\left(-\frac{p_{g_1}}{\sqrt{p_{g_2}}}\right)}{N_{g_1}}\,|v_1\rangle + \frac{\sqrt{p_{g_2}}}{N_{g_2}}\,|v_2\rangle \ \text{and} \\ U\,|g_{33}\rangle &= \frac{\sqrt{p_{g_3}}}{N_g}\,|w\rangle + 0\,|v_1\rangle + \frac{\left(-\frac{p_{g_1}+g_{g_2}}{\sqrt{p_{g_3}}}\right)}{N_{g_2}}\,|v_2\rangle. \end{aligned}$$

For the first term we have $\text{(i)} = x_{g_1}p_{g_1}\left[\begin{array}{c|ccc} & \langle v_1| & \langle v_2| & \langle w| \\ \hline |v_1\rangle & \frac{1}{N_{g_1}^2} & \frac{1}{N_{g_1}N_{g_2}} & \frac{1}{N_{g_1}N_g} \\ |v_2\rangle & \frac{1}{N_{g_2}N_{g_1}} & \frac{1}{N_{g_2}^2} & \frac{1}{N_{g_2}N_g} \\ |w\rangle & \frac{1}{N_gN_{g_1}} & \frac{1}{N_gN_{g_2}} & \frac{1}{N_g^2} \end{array}\right]$.

For the second term, we re-write $U\,|g_{22}\rangle = \sqrt{p_{g_2}}\left(\frac{1}{N_g}\,|w\rangle - \frac{1}{N'_{g_1}}\,|v_1\rangle + \frac{1}{N_{g_2}}\,|v_2\rangle\right)$ with $N'_{g_1} = \frac{p_{g_2}}{p_{g_1}}N_{g_1}$,

to obtain $\text{(ii)} = x_{g_2}p_{g_2}\left[\begin{array}{c|ccc} & \langle v_1| & \langle v_2| & \langle w| \\ \hline |v_1\rangle & \frac{1}{N'^2_{g_1}} & -\frac{1}{N'_{g_1}N_{g_2}} & -\frac{1}{N'_{g_1}N_g} \\ |v_2\rangle & -\frac{1}{N_{g_2}N'_{g_1}} & \frac{1}{N_{g_2}^2} & \frac{1}{N_{g_2}N_g} \\ |w\rangle & -\frac{1}{N_gN'_{g_1}} & \frac{1}{N_gN_{g_2}} & \frac{1}{N_g^2} \end{array}\right]$,

and finally $U\,|g_{33}\rangle = \sqrt{p_{g_3}}\left(\frac{1}{N_g}\,|w\rangle + 0\,|v_1\rangle - \frac{1}{N'_{g_2}}\,|v_2\rangle\right)$ with $N'_{g_2} = \frac{p_{g_3}}{p_{g_1}+p_{g_2}}$,

to get $\text{(iii)} = x_{g_3}p_{g_3}\left[\begin{array}{c|ccc} & \langle v_1| & \langle v_2| & \langle w| \\ \hline |v_1\rangle & & & \\ |v_2\rangle & & \frac{1}{N'^2_{g_2}} & -\frac{1}{N'_{g_2}N_g} \\ |w\rangle & & -\frac{1}{N_gN'_{g_2}} & \frac{1}{N_g^2} \end{array}\right]$.

Now we can combine all of these into a single matrix and try to obtain some simpler constraints.

$$M \overset{\text{def}}{=} \left[\begin{array}{c|ccccc} & \langle v| & \langle v_1| & \langle v_2| & \langle w| & \langle w_1| \\ \hline |v\rangle & x & & & & \\ |v_1\rangle & & x-\frac{x_{g_1}p_{g_1}}{N_{8_1}^2}-\frac{x_{g_2}p_{g_2}}{N_{g_1}^{\prime 2}} & -\frac{x_{g_1}p_{g_1}}{N_{g_1}N_{g_2}}+\frac{x_{g_2}p_{g_2}}{N'_{g_1}N_{g_2}} & -\frac{x_{g_1}p_{g_1}}{N_{g_1}N_g}+\frac{x_{g_2}p_{g_2}}{N'_{g_1}N_g} & \\ |v_2\rangle & & -\frac{x_{g_1}p_{g_1}}{N_{g_2}N_{g_1}}+\frac{x_{g_2}p_{g_2}}{N_{g_2}N'_{g_1}} & x-\frac{x_{g_1}p_{g_1}}{N_{g_2}^2}-\frac{x_{g_2}p_{g_2}}{N_{g_2}^2}-\frac{x_{g_3}p_{g_3}}{N_{g_2}^{\prime 2}} & -\frac{x_{g_1}p_{g_1}}{N_{g_2}N_g}-\frac{x_{g_2}p_{g_2}}{N_{g_2}N_g}+\frac{x_{g_3}p_{g_3}}{N'_{g_2}N_g} & \\ |w\rangle & & -\frac{x_{g_1}p_{g_1}}{N_gN_{g_1}}+\frac{x_{g_2}p_{g_2}}{N_gN'_{g_1}} & -\frac{x_{g_1}p_{g_1}}{N_gN_{g_2}}-\frac{x_{g_2}p_{g_2}}{N_gN_{g_2}}+\frac{x_{g_3}p_{g_3}}{N_gN'_{g_2}} & \frac{p_{h_1}x_{h_1}+p_{h_2}x_{h_2}}{N_h^2}-\frac{1}{N_g^2}\sum_i x_{g_i}p_{g_i} & \frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1}-x_{h_2}) \\ |w_1\rangle & & & & \frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1}-x_{h_2}) & \frac{p_{h_2}x_{h_1}+p_{h_1}x_{h_2}}{N_h^2} \end{array}\right] \geq 0.$$

Despite this appearing to be a complicated expression, we can conclude that it is always so that the larger x is the looser is the constraint. To show this and simplify the calculation, note that M can be split into a scalar condition, $x \geq 0$ – from the $|v\rangle\langle v|$ part – and a sub-matrix which we choose to write as

$$\begin{array}{c|c|c} & \langle v_1| \quad \langle v_2| & \langle w| \quad \langle w_1| \\ \hline \begin{array}{c}|v_1\rangle \\ |v_2\rangle\end{array} & C & B^T \\ \hline \begin{array}{c}|w\rangle \\ |w_1\rangle\end{array} & B & A \end{array} \geq 0.$$

We $\begin{bmatrix} C & B^T \\ B & A \end{bmatrix} \geq 0 \iff \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \geq 0 \iff C \geq 0,\ A - BC^{-1}B^T \geq 0,\ (\mathbb{I} - CC^{-1})B^T = 0$, using Shur's Complement condition for positivity where C^{-1} is the generalized inverse. We can take x to be sufficiently large so that $C > 0$ and thereby make sure that $\mathbb{I} - CC^{-1} = 0$. Then, the only condition of interest is

$$A - BC^{-1}B^T \geq 0.$$

Actually, we can do even better than this. Note that if $C > 0$ then $C^{-1} > 0$ and that the second term is of the form

$$\underbrace{\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}}_{B}\underbrace{\begin{bmatrix} \alpha & \gamma \\ \gamma & \beta \end{bmatrix}}_{C^{-1}}\underbrace{\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}}_{B^T} = \begin{bmatrix} \begin{bmatrix} a & b \end{bmatrix}\begin{bmatrix} \alpha & \gamma \\ \gamma & \beta \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} & 0 \\ 0 & 0 \end{bmatrix} \geq 0,$$

because $C^{-1} > 0$. We can therefore write the constraint equation as $A \geq BC^{-1}B^T \geq 0$ and note that $A \geq 0$ is a necessary condition. This also becomes a sufficient condition in the limit that $x \longrightarrow \infty$ because $C^{-1} \longrightarrow 0$ in that case. Thus, we have reduced the analysis to simply checking if

$$\begin{bmatrix} \frac{p_{h_1}x_{h_1}+p_{h_2}x_{h_2}}{N_h^2} - \frac{1}{N_g^2}\sum_i x_{g_i}p_{g_i} & \frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1}-x_{h_2}) \\ \frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1}-x_{h_2}) & \frac{p_{h_2}x_{h_1}+p_{h_1}x_{h_2}}{N_h^2} \end{bmatrix} \geq 0.$$

This is a 2×2 matrix and can be checked for positivity using the trace and determinant method or we can use again Schur's Complement conditions. Here, however, we intend to use a more general technique. Let us introduce

$$\langle x_g \rangle \overset{\text{def}}{=} \frac{1}{N_g^2}\sum_i x_{g_i}p_{g_i},\ \left\langle \frac{1}{x_h} \right\rangle \overset{\text{def}}{=} \frac{1}{N_h^2}\sum_i \frac{p_{h_i}}{x_{h_i}}.$$

Term (I) and one element from term (III) constitute a matrix A which can be written as

$$A = x_{h_1}|h_{11}\rangle\langle h_{11}| + x_{h_2}|h_{22}\rangle\langle h_{22}| - \langle x_g \rangle |w\rangle\langle w| = \begin{array}{c|cc} & \langle h_{11}| & \langle h_{22}| \\ \hline |h_{11}\rangle & x_{h_1} & \\ |h_{22}\rangle & & x_{h_2} \end{array} - \langle x_g \rangle |w\rangle\langle w|.$$

We use $F - M \geq 0 \iff \mathbb{I} - \sqrt{F}^{-1} M \sqrt{F}^{-1} \geq 0$ for $F > 0$, to obtain $\mathbb{I} \geq \left\langle x_g \right\rangle |w''\rangle \langle w''|$, where $|w''\rangle = \frac{\sqrt{\frac{p_{h_1}}{x_{h_1}}}|h_{11}\rangle + \sqrt{\frac{p_{h_2}}{x_{h_2}}}|h_{22}\rangle}{N_h}$. Normalizing this we get $|w'\rangle = \frac{|w''\rangle}{\sqrt{\left\langle \frac{1}{x_h} \right\rangle}}$ which entails $\mathbb{I} \geq \left\langle x_g \right\rangle \left\langle \frac{1}{x_h} \right\rangle |w'\rangle \langle w'|$ and that leads us to the final condition $\frac{1}{\langle x_g \rangle} \geq \left\langle \frac{1}{x_h} \right\rangle$.

In fact all the techniques used in reaching this result can be extended to the $m \to n$ transition case as well and so the aforesaid result holds in general.

D Approaching bias $\epsilon(k) = 1/(4k+2)$

Lemma 134. *Consider an n-dimensional vector space. Given a diagonal matrix $X = diag(x_1, x_2 \ldots x_n)$ and a vector $|c\rangle = (c_1, c_2 \ldots, c_n)$ where all the x_is are distinct and all the c_i are non-zero, the vectors $|c\rangle, X|c\rangle, \ldots X^{n-1}|c\rangle$ span the vector space.*

Proof. We write the vectors as

$$|\tilde{w}_i\rangle = X^{i-1}|c\rangle = \begin{bmatrix} x_1^{i-1}c_1 \\ x_2^{i-1}c_2 \\ \vdots \\ x_n^{i-1}c_n \end{bmatrix}.$$

We show that the set of vectors are linearly independent, which is equivalent to showing that the determinant of the matrix containing the vectors as rows (or equivalently as columns) is non-zero, i.e.

$$\det\left(\underbrace{\begin{bmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & & x_n \\ x_1^2 & x_2^2 & & x_n^2 \\ \vdots & & \ddots & \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{bmatrix}}_{:=\tilde{X}} \begin{bmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_n \end{bmatrix} \right) = c_1 \cdot c_2 \cdot \ldots c_n \cdot \det \tilde{X}$$

is non-zero. Notice that $\tilde{X}$ is the so-called Vandermonde matrix (restricted to being a square matrix) and its determinant, known as the Vandermonde determinant, is $\det(\tilde{X}) = \prod_{1 \leq i \leq j \leq n}(x_j - x_i) \neq 0$ as x_is are distinct. As c_is are all non-negative our proof is complete. □

D.1 Proof of Lemma 33

In our proof we will need the following Lemma 135, which gives a property of the $f-$assignments.

Lemma 135. $\sum_{i=1}^{n} \frac{f(x_i)}{\prod_{j \neq i}(x_j - x_i)} = 0$ *where $f(x_i)$ is a polynomial of order $k \leq n-2$ where $x_i \in \mathbb{R}$ are distinct.*

The proof can be found in [Moc07; Aha+14b].

Proof of Lemma 33. The equality $\left\langle x^k \right\rangle = 0$ for $k \leq n-2$ is a direct consequence of Lemma 135, and we proceed to prove the inequality $\left\langle x^{n-1} \right\rangle > 0$. Suppose for now that (we prove it in the end)

$$\sum_{i=1}^{n} \frac{x_i^{n-1}}{\prod_{j \neq i}(x_j - x_i)} = (-1)^{n-1}. \tag{30}$$

Define $p(x_i) = \frac{-(-x_i)^m}{\prod_{j\neq i}(x_j - x_i)}$ so that $t = \sum_i p(x_i) \llbracket x_i \rrbracket$. Observe that

$$\begin{aligned}
\left\langle x^{n-1} \right\rangle &= \sum_i x_i^{n-m-1} p(x_i) \\
&= \sum_i (-1)^m x_i^{n-1} \frac{-1}{\prod_{j\neq i}(x_j - x_i)} \\
&= (-1)^m(-1) \sum_i \frac{x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)} \\
&= (-1)^m(-1)(-1)^{n-1} = (-1)^{m+n}
\end{aligned}$$

where we used Equation Equation (30).

It remains to prove Equation Equation (30). We show that $d(n) = \sum_{i=1}^n \frac{x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)} = (-1)^{n-1}$ by induction. The base of the induction gives us $d(2) = \frac{x_1}{x_2 - x_1} + \frac{x_2}{x_1 - x_2} = -1$. We continue by assuming that it holds for $d(n)$ and take

$$\begin{aligned}
d(n+1) &= \sum_{i=1}^{n+1} \frac{x_i^n}{\prod_{j\neq i}(x_j - x_i)} = \sum_{i=1}^{n+1} \frac{-(x_{n+1} - x_i)x_i^{n-1} + x_{n+1}x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)} \\
&= -\sum_{i=1}^{n+1} (x_{n+1} - x_i) \frac{x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)} + x_{n+1} \underbrace{\sum_{i=1}^{n+1} \frac{x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)}}_{= 0, \text{ from Lemma } 135} \\
&= -\sum_{i=1}^{n} \frac{x_{n+1} - x_i}{x_{n+1} - x_i} \frac{x_i^{n-1}}{\prod_{j\neq i,n+1}(x_j - x_i)} + (x_{n+1} - x_{n+1}) \frac{x_{n+1}^{n-1}}{\prod_{j\neq n+1}(x_j - x_{n+1})} = -d(n).
\end{aligned}$$

This completes the proof. □

D.2 Restricted decomposition into f_0-assignments

The monomial decomposition we presented in Section 4.1 is not unique. Here, we give another useful decomposition that, however, only works in a restricted case; that is when the roots of f are right roots, as described below.

Lemma 136 (f with right roots to f_0)**.** *Consider a set of real coordinates satisfying $0 < x_1 < x_2 \cdots < x_n$ and let $f(x) = (r_1 - x)(r_2 - x)\dots(r_k - x)$ where $k \leq n-2$ and the roots $\{r_i\}_{i=1}^k$ of f are right roots, i.e. they are such that for every root r_i there exists a distinct coordinate $x_j < r_i$. Let $t = \sum_{i=1}^n p_i \llbracket x_i \rrbracket$ be the corresponding f-assignment. Then, there exist f_0-assignments, $\{t_{0;j}\}$, on a subset of $(x_1, x_2 \dots x_n)$, such that $t = \sum_{i=1}^m \alpha_i t_{0;i}$ where $\alpha_i > 0$ is a real number and $m > 0$ is an integer.*

Proof. For simplicity, assume that $x_i < r_i$, $\forall i$, but the argument works in general. We can, then, write

$$\begin{aligned}
t &= \sum_{i=1}^n \frac{-f(x_i)}{\prod_{j\neq i}(x_j - x_i)} \llbracket x_i \rrbracket \\
&= \sum_{i=1}^n \left(\frac{-(r_1 - x_1)(r_2 - x_i)\dots(r_k - x_i)}{\prod_{j\neq i}(x_j - x_i)} + \frac{-(x_1 - x_i)(r_2 - x_i)\dots(r_k - x_i)}{\prod_{j\neq i}(x_j - x_i)} \right) \llbracket x_i \rrbracket \\
&= (r_1 - x_1) \sum_{i=1}^n \frac{-(r_2 - x_i)\dots(r_k - x_i)}{\prod_{j\neq i}(x_j - x_i)} \llbracket x_i \rrbracket + \sum_{i=2}^n \frac{-(r_2 - x_i)\dots(r_k - x_i)}{\prod_{j\neq i,1}(x_j - x_i)} \llbracket x_i \rrbracket,
\end{aligned}$$

where the first term has the same form that we started with (except for a positive constant which is irrelevant for the validity condition; see Definition 126) but with the polynomial having one less degree. The second term also has the same form, except that the number of points involved has been reduced. Note how this process relies crucially on the fact that $r_1 - x_1 > 0$; otherwise the term on the left would, by itself, not correspond to a valid move. This process can be repeated until we obtain a sum of f_0-assignments on various subsets of $(x_1, x_2 \dots x_n)$. □

The advantage of this decomposition is that we can immediately apply it to the f-assignment of the bias-1/10 game. This is relevant because constructing solutions to f_0-assignments is relatively easy and so they, together with this result, allow us to derive the 1/10 bias protocol circumventing the perturbative approach that we used in Section 3.

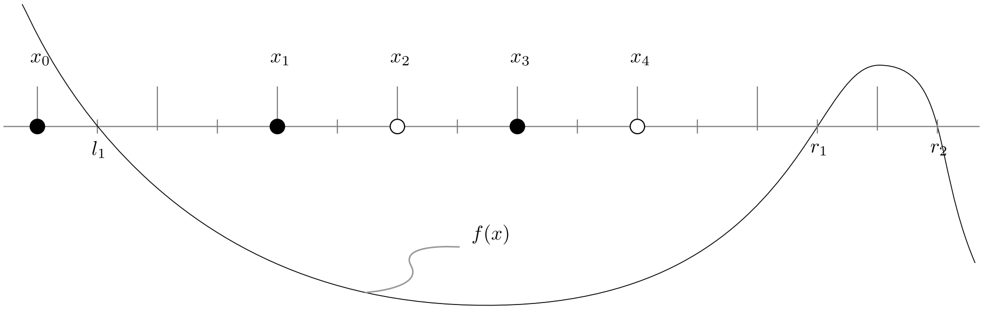


Figure 21: The main 1/10 move involves $n = 5$ points. f has $k = 3$ roots, all of which are right roots.

Example 137 (The main 1/10 move.)**.** The key move in the 1/10-bias point game has its coordinates given by x_0, x_1, x_2, x_3, x_4 and roots given by l_1, r_1, r_2 which satisfy $x_0 < l_1 < x_1 < x_2 < x_3 < x_4 < r_1 < r_2$. Each root is a right root here because $x_0 < l_1$, $x_3 < r_1$, $x_4 < r_2$. Hence, from Lemma 136, this assignment can be expressed as a combination of f_0-assignments defined over subsets of the initial set of coordinates and each f_0-assignment admits a simple solution given by Proposition 37 and Proposition 38 .

Another simple example is the class of f-assignments describing merge moves (see Example 23). We place the roots of f in such a way that all points, except one, have negative weights.

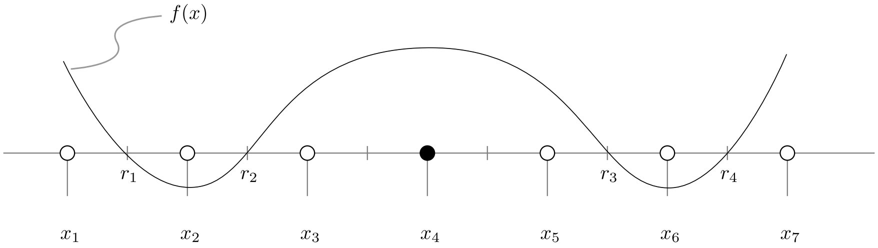


Figure 22: Merge involving $n = 7$ points. f has in total $k = n - 3 = 4$ right roots.

Example 138 (Merge)**.** For merges (see Figure 22) we only get right-roots and hence, we can write them as sums of f_0-assignments and obtain the solution using Proposition 37 and Proposition 38. For n points, the polynomial has degree $n - 3$ and so $\langle x \rangle = 0$, just as expected for a merge.

This scheme fails for moves corresponding to lower bias games. For instance, the main move of the bias $1/14$ game has its coordinates given by $x_0, x_1, x_2, x_3, x_4, x_5, x_6$ and the roots of f are l_1, l_2, r_1, r_2, r_3 satisfying $x_0 < l_1 < l_2 < x_1 < x_2 \cdots < x_6 < r_1 < r_2 < r_3$. Here, we can either consider l_1 to be a right root, in which case l_2 is a left root (i.e. a root which is not a right root). Or we can consider l_2 to be a right root in which case l_1 becomes a left root. Thus for games with bias $1/14$ and less, we must revert to Lemma 35, which means we can not – at least by this scheme – avoid finding the solution to all the monomial assignments.

Since we mentioned the merge move, for completeness let us consider also the split move (see Example 24). The situation (see Figure 23) is similar to that of merge but with one key distinction: the polynomial has degree $n-2$; it has $n-3$ right roots and one left root. Thus, it can not be expressed as a sum of f_0-assignments using Lemma 136. Of course, merges and splits by themselves are not of much interest in this discussion because we already know that the Blinkered Unitary solves them both (see Section 3.1.1).

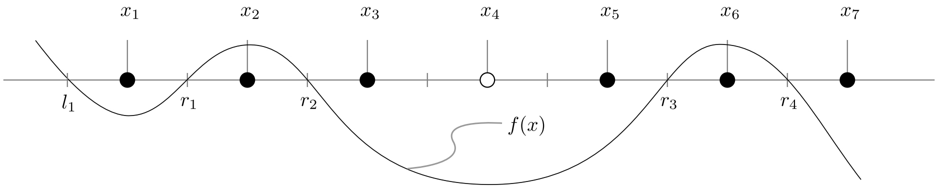


Figure 23: Split involving 7 points. f has $k = n - 2 = 5$ roots; 4 right and one left.

Notice that using the method in Appendix G we can convert left to right roots and combine the different monomial assignment solutions to obtain the full solution.

E From EBM to EBRM to COF

In Appendix A we saw how D. Aharonov et al. prove that valid functions are equivalent to EBM functions following their work [Aha+14b]. Here, we show that we can, in fact, do even better. Instead of EBM functions, we consider Expressible-By-Real-Matrices (EBRM) functions, where the matrices are additionally restricted to be real. Let this set be given by K'. It turns out that its dual K'^* is also the set of operator monotone functions [Fri18] viz. $K'^* = K^*$. The proof for $K = K^{**}$ from [Aha+14b], can be applied to the real case as it is to get $K' = K^{**}$. Since C. Mochon's point games use valid functions, the aforesaid simplification justifies why it suffices to restrict to real matrices.

E.1 Equivalence of EBM and EBRM

First we define EBRM transitions and functions similar to their EBM analogues except with the further restriction that the matrices and vectors involved are real.

Definition 139 (EBRM transitions)**.** Let $g, h : [0, \infty) \to [0, \infty)$ be two functions with finite supports. The transition $g \to h$ is EBRM if there exist two real matrices $0 \le G \le H$ and a vector $|\psi\rangle$, not necessarily normalized, such that $g = \text{prob}[G, \psi]$ and $h = \text{prob}[H, \psi]$.

Definition 140 (K', EBRM functions; K'_Λ, EBRM functions on $[0, \Lambda]$)**.** A function $a : [0, \infty) \to \mathbb{R}$ with finite support is an EBRM function if the transition $a^- \to a^+$ is EBRM, where $a^+ : [0, \infty) \to [0, \infty)$ and $a^- : [0, \infty) \to [0, \infty)$ denote, respectively, the positive and the negative part of a, i.e. $a = a^+ - a^-$. We denote by K' the set of EBRM functions.

For any finite $\Lambda \in (0,\infty)$, a function $a : [0,\Lambda) \to \mathbb{R}$ with finite support is an EBRM function with support on $[0,\Lambda]$ if the transition $a^- \to a^+$ is EBRM with its spectrum in $[0,\Lambda]$, where a^+ and a^- denote, respectively, the positive and the negative part of a. We denote by K'_Λ the set of EBRM functions with support on $[0,\Lambda]$.

Definition 141 (Real operator monotone functions)**.** A function $f : (0,\infty) \to \mathbb{R}$ is a real operator monotone if for all real matrices $0 \le A \le B$ we have $f(A) \le f(B)$.

A function $f : (0,\Lambda) \to \mathbb{R}$ is a real operator monotone on $[0,\Lambda]$ if for all real matrices $0 \le A \le B$ with spectrum in $[0,\Lambda]$ we have $f(A) \le f(B)$.

Lemma 142. *K'^*_Λ is the set of real operator monotones on $[0,\Lambda]$.*

Proof sketch. In [Aha+14b] the authors showed that the dual of the set of EBM functions on $[0,\Lambda]$, denoted as K^*_Λ, is the set of operator monotone functions on $[0,\Lambda]$. Their proof can be adapted here by restricting to real matrices to show that K'^*_Λ is the set of real operator monotone functions on $[0,\Lambda]$. □

Lemma 143. *$K^*_\Lambda = K'^*_\Lambda$ and $K^* = K'^*$, i.e. the set of operator monotones on $[0,\Lambda]$ equals the set of real operator monotones on $[0,\Lambda]$ and the set of operator monotones equals the set of real operator monotones.*

Corollary 144. *$K'_\Lambda = K'^{**}_\Lambda = K^{**}_\Lambda = K_\Lambda$, i.e. the set of EBRM functions on $[0,\Lambda]$ = the set of Λ-valid functions as the dual of EBRM functions = the set of Λ-valid functions as the dual of EBM functions = the set of EBM functions on $[0,\Lambda]$.*

Corollary 145. *Any strictly valid function is EBRM.*

Let us sketch the proof of Lemma 143. The set of real operator monotones must contain the set of operator monotones, since the latter are – by definition – required to work in the real case as well. The set of real operator monotones might be bigger, but that is not the case, as we can encode an $n \times n$ Hermitian matrix into a $2n \times 2n$ real symmetric matrix. This is achieved by replacing each complex number $\alpha + i\beta$ with the matrix

$$\alpha \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} + \beta \begin{bmatrix} & -1 \\ 1 & \end{bmatrix}.$$

This corresponds to writing a complex matrix $W = W_{\mathfrak{R}} + iW_{\mathfrak{I}}$ as a real symmetric matrix

$$W' = \begin{bmatrix} W_{\mathfrak{R}} & -W_{\mathfrak{I}} \\ W_{\mathfrak{I}} & W_{\mathfrak{R}} \end{bmatrix},$$

where $W_{\mathfrak{R}}$ and $W_{\mathfrak{I}}$ are real. For a Hermitian $W^\dagger = W$ we must have $W_{\mathfrak{R}} = W_{\mathfrak{R}}^T$ and $W_{\mathfrak{I}} = -W_{\mathfrak{I}}^T$ which makes $W' = W'^T$ a symmetric matrix. The spectra of W and W' are the same. This is established by looking at the diagonal decomposition, $W = USU^\dagger$, which would get transformed to $W' = U'S'U'^\dagger$. Since S is real S' is also real with doubly degenerate eigenvalues (except for the degeneracy already present in S). Thus if we have $0 \le A \le B$ then we would also have $0 \le A' \le B'$ as $A - B$ and $A' - B'$ would have the same spectrum; we used A' and B' to represent real symmetric analogues of the Hermitian matrices A and B. The other direction is trivial. Hence we have an equivalence which means that requiring a function to be operator monotone on complex matrices is the same as requiring it to be operator monotone on real symmetric matrices of at most twice the size. This establishes that the set of real operator monotones is the same as the set of operator monotones.

E.2 EBRM to COF

Having reduced our problem from the set of EBM transitions to the set of EBRM transitions, we now show that each EBRM transition can be written in the COF, see Section 5.1, thus showing that $\Lambda-$valid functions admit matrices of the COF. The result is actually due to C. Mochon and A. Y. Kitaev [Moc07], but we present the proof here, as there was a minor mistake in one of the arguments that we corrected. The interesting part is showing that we can always restrict to matrices of a certain size that depends on the number of points in the transition.

Lemma 146. *For every EBRM transition $g \longrightarrow h$ with spectrum in $[a, b]$ there exists an orthogonal matrix O, diagonal matrices X_h, X_g (with no multiplicities except possibly those of a and b) of size at most $n_g + n_h - 1$ such that*

$$O \underbrace{\begin{bmatrix} x_{g_1} & & & & \\ & \ddots & & & \\ & & x_{g_{n_g}} & & \\ & & & a & \\ & & & & \ddots \end{bmatrix}}_{:=X_g} O^T \leq \begin{bmatrix} x_{h_1} & & & & \\ & \ddots & & & \\ & & x_{h_{n_h}} & & \\ & & & b & \\ & & & & \ddots \end{bmatrix} = X_h,$$

and the vector $|\psi\rangle := \sum_{i=1}^{n_h} \sqrt{p_{h_i}} |i\rangle = \sum_{i=1}^{n_g} \sqrt{p_{g_i}} O |i\rangle$.

Proof. An EBRM entails that we are given $G \leq H$ with their spectrum in $[a, b]$ and a $|\psi\rangle$ such that

$$g = \text{Prob}[G, |\psi\rangle] = \sum_{i=1}^{n_g} p_{g_i}[x_{g_i}] \quad \text{and} \quad h = \text{Prob}[H, |\psi\rangle] = \sum_{i=1}^{n_h} p_{h_i}[x_{h_i}],$$

with $p_{g_i}, p_{h_i} > 0$ and $x_{g_i} \neq x_{g_j}$, $x_{h_i} \neq x_{h_j}$ for $i \neq j$, but the dimension and multiplicities can be arbitrary. First we show that one can always choose the eigenvectors $|g_i\rangle$ of G with eigenvalue x_{g_i} such that $|\psi\rangle = \sum_{i=1}^{n_g} \sqrt{p_{g_i}} |g_i\rangle$. Consider P_{g_i} to be the projector on the eigenspace with eigenvalue x_{g_i}. Note that $|g_i\rangle := \frac{P_{g_i}|\psi\rangle}{\sqrt{\langle\psi|P_{g_i}|\psi\rangle}}$ fits the bill. Similarly, we choose $|h_i\rangle$ such that $|\psi\rangle = \sum_{i=1}^{n_h} \sqrt{p_{h_i}} |h_i\rangle$. Consider now the projector onto the $\{|g_i\rangle\}$ space $\Pi_g = \sum_{i=1}^{n_g} |g_i\rangle \langle g_i|$. Note that this does not have all eigenvectors with eigenvalues $\in \{x_{g_i}\}$. Similarly, we define $\Pi_h = \sum_{i=1}^{n_h} |h_i\rangle \langle h_i|$. We further define $G' := \Pi_g G \Pi_g + a(\mathbb{I} - \Pi_g)$ and $H' := \Pi_h H \Pi_h + b(\mathbb{I} - \Pi_h)$. These definitions are useful as we can show $G' \leq H'$. From $G = \Pi_g G \Pi_g + (\mathbb{I} - \Pi_g) G (\mathbb{I} - \Pi_g)$ we can conclude that $\Pi_g G \Pi_g + a(\mathbb{I} - \Pi_g) \leq G$. This entails $G' \leq G$. Using a similar argument one can also establish that $H \leq H'$. Combining these we get $G' \leq H'$.

Consider the projector $\Pi :=$ projector on span$\{\{|g_i\rangle\}_{i=1}^{n_g}, \{|h_i\rangle\}_{i=1}^{n_h}\}$, and note that it has dimension at most $n_g + n_h - 1$, because $|\psi\rangle$ is in the span of $\{|g_i\rangle\}$ and in the span of $\{|h_i\rangle\}$, therefore at least one of the basis vectors is not independent. We have $G'' := \Pi G' \Pi \leq \Pi H' \Pi =: H''$, since we can always conjugate an inequality by a positive semi-definite matrix on both sides. Note also that $\Pi |\psi\rangle = |\psi\rangle$ which means that the matrices and the vectors have the claimed dimension. We now establish that $\text{Prob}[H'', |\psi\rangle] = h$ and $\text{Prob}[G'', |\psi\rangle] = g$. We first write the projector tailored to the g basis as $\Pi = \Pi_g + \Pi_{g_\perp}$, where $\Pi_{g_\perp}$ is meant to enlarge the space to the span$\{h_i\}_{i=1}^{n_h}$. With this we evaluate

$$G'' = \left(\Pi_g + \Pi_{g_\perp}\right) \left[\Pi_g G \Pi_g + a(\mathbb{I} - \Pi_g)\right] \left(\Pi_g + \Pi_{g_\perp}\right) = \Pi_g G \Pi_g + a\Pi_{g_\perp}.$$

Then $\text{Prob}[G'', |\psi\rangle] = g$. By a similar argument one can establish the same claim for h. Since G'' and H'' have no multiplicities except possibly in a and b, respectively, we conclude that we can always restrict to the claimed dimension and form. □

Corollary 147. *For every EBRM transition the corresponding COF is legal.*

F The Weingarten map

Consider a curve in the plane specified by a function f. Its curvature is related to the rate of change of the tangents of f, i.e. the second derivative of f. For a surface in arbitrary dimensions specified by f, the corresponding quantity becomes a matrix $\partial_i \partial_j f$. The eigenvalues of this matrix give us the curvature along the corresponding eigenvector, however in practice it is a rather cumbersome calculation. We can use a more general method to easily obtain an analytic solution to this problem for ellipsoids; this method is based on the so-called *Weingarten map.* Intuitively, it is defined as the differential of the normal at a given point on the manifold, and it turns out to be effectively the same as finding the aforementioned matrix of second derivatives.

Definition 148 (Weingarten Map (Informal[41], from [Sch09][42])**.** Let K be a manifold specified by the heads of vectors in $\mathbb{R}^n$. Denote the tangent space of K at $|x\rangle \in K$ by $T_{|x\rangle}K$. Let $|u_K(|x\rangle)\rangle$ be the outer unit normal vector of K at $|x\rangle$. The map $|u_k(|x\rangle)\rangle : K \longrightarrow \mathbb{S}^{n-1} \subset \mathbb{R}^n$ as defined is called the *spherical image map*, or the *Gauss map*, of the interior of the manifold K. Its differential at $|x\rangle$, $d(|u_K\rangle)_k =: W_x$ maps $T_{|x\rangle}K$ to itself. The linear map $W_x : T_{|x\rangle}K \longrightarrow T_{|x\rangle}K$ is called the *Weingarten map.*

The so-called *Reverse Weingarten map* is easier to calculate, and useful due to the following result.

Theorem 149 (Informal)**.** *[Sch09] The inverse of the Weingarten map equals the reverse Weingarten map, for well-behaved surfaces.*

We omit the exact statement of the theorem and the definition of the Reverse Weingarten map as they are not directly relevant for us. We simply work with a formula for the Weingarten map as follows.

Definition 150 (Support Function [Sch09])**.** Given a manifold specified by a set S of vectors and a normalized vector $|u\rangle$, the support function is defined as

$$h_S(|u\rangle) := \sup_{|s\rangle \in S} \langle s|u\rangle .$$

Theorem 151 (Formula for evaluating the Reverse Weingarten Map (Informal[43] from [Sch09]))**.** *Consider a convex surface specified by a set S of vectors. Given a normalized vector $|u\rangle$, the Reverse Weingarten map, W, evaluated along the normal specified by $|u\rangle$ is given by*

$$(W)_{ij} = \left.\frac{\partial^2 h_S(|u'\rangle)}{\partial u'_i \partial u'_j}\right|_u , \text{ where } h_S(|u'\rangle) \text{ is the support function.}$$

Assuming that we can invert a matrix, using Theorem 151 and Theorem 149 we can obtain the Weingarten map. For ellipsoids we have:

Lemma 152. *Given an $n \times n$ matrix $G \geq 0$, the support function corresponding to the ellipsoid S_G along a normal $|u\rangle$ of the manifold is given by*

$$h_{S_G}(|u\rangle) = \sqrt{\langle u| G^{-1} |u\rangle}.$$

[41] There are qualifying conditions on K which we suppressed.

[42] The convention therein for T and K is slightly different.

[43] The qualifying conditions on the surface and certain technicalities have been omitted.

In our analysis, we typically know the point at which we wish to evaluate the curvature. The calculation of the support function requires the normal at that point, which can be evaluated as follows:

Lemma 153 (Normal)**.** *Given an $n \times n$ matrix $G \geq 0$, consider the manifold S_G associated with it. Let $|v\rangle \in \Pi\mathbb{R}^n$ be a vector such that $\mathcal{E}_G(|v\rangle)$ is well-defined ($\langle v| G |v\rangle \neq 0$) where Π is as defined in Definition 84. The normal at $\mathcal{E}_G(|v\rangle)$ – which we also refer to as the normal along $|v\rangle$ – is given by $|u\rangle = G|v\rangle / \sqrt{\langle v| G^2 |v\rangle}$.*

Proof. Consider $G = \operatorname{diag}(x_{g_1}, x_{g_2} \dots x_{g_n})$ and let $|v\rangle = (v_1, v_2 \dots v_n)$. The surface S_G is determined by the constraint $\langle v| G |v\rangle = 1$ which is equivalent to $\sum_{i=1}^n x_{g_i} v_i^2 = 1$. Changing the constant 1 can be thought of as scaling the surface. Treating $\sum_{i=1}^n x_{g_i} v_i^2$ as a scalar function, its gradient points along the outward normal: $|u\rangle \propto \sum_{j=1}^n \frac{\partial}{\partial v_j} \sum_{i=1}^n x_{g_i} v_i^2 |j\rangle \propto \sum_{j=1}^n x_{g_j} v_j |j\rangle \propto G|v\rangle$. □

With these ingredients we can now evaluate the Reverse Weingarten Map.

Lemma 154 (Reverse Weingarten Map)**.** *Given an $n \times n$ matrix $G \geq 0$, and a vector $|v\rangle \in \Pi\mathbb{R}^n$ where Π is as defined in Definition 84, the Reverse Weingarten Map associated with the surface S_G, evaluated at the point $\mathcal{E}_G(|v\rangle)$ is given by*

$$W_G := \sqrt{\frac{\langle G^2\rangle}{\langle G\rangle}} \left(G^{\dashv} - \frac{|v\rangle\langle v|}{\langle G\rangle} \right), \text{ where } \left\langle G^j \right\rangle := \langle v| G^j |v\rangle .$$

Proof. We prove this for $G > 0$ (for $G \geq 0$ but $G \not> 0$, it follows analogously by restricting to the non-zero eigenspace). Let the spectral decomposition of G be given by $G = \sum_{i=1}^n x_{g_i} |g_i\rangle\langle g_i|$, and let $|v\rangle = \sum_{i=1}^n c_i |g_i\rangle$. From Lemma 153 the normal along $|v\rangle$ is given by $|u\rangle = G|v\rangle / \sqrt{\langle v| G^2 |v\rangle}$. Writing $|u\rangle = \sum_{i=1}^n u_i |g_i\rangle$, the u_is are fixed. Then, the support function evaluated along the normal $|u\rangle$ is given by (we denote $h_{S_G}(|u\rangle)$ by h)

$$\begin{aligned} h &= \sqrt{\langle u| G^{\dashv} |u\rangle} = \sqrt{\textstyle\sum_{i=1}^n x_{g_i}^{-1} u_i^2} && \text{from Lemma 152} \\ \implies (W_G)_{ij} = \frac{\partial^2 h}{\partial u_i \partial u_j} &= -\tfrac{1}{h^3} x_{g_j}^{-1} x_{g_i}^{-1} u_j u_i + \tfrac{x_{g_i}^{-1}}{h} \delta_{ij} && \text{from Theorem 151} \\ \implies W_G &= -\tfrac{1}{h^3} G^{\dashv} |u\rangle\langle u| G^{\dashv} + \tfrac{G^{\dashv}}{h}, \end{aligned}$$

where we used the more general notation $G^{\dashv} = G^{-1}$. Substituting $|u\rangle$ in the expression for h and W_G we obtain $h = \sqrt{\frac{\langle G\rangle}{\langle G^2\rangle}}$ and

$$W_G = \frac{1}{h} G^{\dashv} - \frac{1}{h^3} \frac{|v\rangle\langle v|}{\langle G^2\rangle} = \sqrt{\frac{\langle G^2\rangle}{\langle G\rangle}} \left(G^{\dashv} - \frac{|v\rangle\langle v|}{\langle G\rangle} \right).$$

When $G \geq 0$ and has zero eigenvalues, the spectral decomposition has m elements with $m < n$, i.e. $G = \sum_{i=1}^m x_{g_i} |g_i\rangle\langle g_i|$. The sum $\sum_{i=1}^m x_{g_i}^{-1} u_i^2$ would then correspond to $\langle u| G^{\dashv} |u\rangle$. Similar replacements can be made to generalize the proof for $G \geq 0$. □

Inverting the Reverse Weingarten Map is not too hard due to the following result.

Theorem 155 (Sherman-Morrison formula [SM50; Hag89])**.** *Let A be an $n \times n$ invertible matrix and let $|a\rangle$, $|b\rangle$ be n-dimensional vectors. Then, $A + |a\rangle\langle b|$ is invertible if and only if $1 + \langle b| A^{-1} |a\rangle \neq 0$. Furthermore, if this is the case, then*

$$(A + |a\rangle\langle b|)^{-1} = A^{-1} - \frac{A^{-1}|a\rangle\langle b| A^{-1}}{1 + \langle b| A^{-1} |a\rangle}.$$

Combining the above we can also evaluate the Weingarten map.

Lemma 156 (Weingarten Map)**.** *Given an $n \times n$ matrix $G \geq 0$, the Weingarten Map associated with the surface S_G, evaluated at the point $\mathcal{E}_G\left(|v\rangle\right)$ is given by*

$$W_G^{\dashv} = \sqrt{\frac{\langle G\rangle}{\langle G^2\rangle}}\left(G + \frac{\langle G^3\rangle}{\langle G^2\rangle^2} G\,|v\rangle\,\langle v|\,G - \frac{1}{\langle G^2\rangle}\left(G\,|v\rangle\,\langle v|\,G^2 + G^2\,|v\rangle\,\langle v|\,G\right)\right)$$

where $\langle G\rangle := \langle v|\,G\,|v\rangle$.

Proof. We prove for $G > 0$ and the proof for $G \geq 0$ follows analogously. By a direct computation we have $W_G G\,|v\rangle = 0$ (see Lemma 154), and applying Theorem 155 we obtain

$$W^{-1} = \sqrt{\frac{\langle G\rangle}{\langle G^2\rangle}}\left(G + \frac{G\,|v\rangle\,\langle v|\,G}{\langle G\rangle \cdot 0}\right), \tag{31}$$

where we set $A = G^{\dashv} = G^{-1}$ and $|a\rangle = |b\rangle = G\,|v\rangle\,/\sqrt{\langle G\rangle}$ (after pulling out the $1/\sqrt{\langle G^2\rangle\,/\,\langle G\rangle}$ factor). Using appropriate interpolations (for instance $|a\rangle = -\,|b\rangle = (1-\epsilon)G\,|v\rangle\,/\sqrt{\langle G\rangle}$ instead of $G\,|v\rangle\,/\sqrt{\langle G\rangle}$), we can make the second term well-defined and have it diverge only as some parameter ϵ vanishes. The quantity we are interested in is $W^{\dashv} = \Pi_u^{\perp} W \Pi_u^{\perp}$, where $\Pi_u^{\perp} = \mathbb{I} - |u\rangle\,\langle u|$ and $|u\rangle = G\,|v\rangle\,/\sqrt{\langle G\rangle}$. If the positive inverse is to be well-defined, the second term in Equation (31) should disappear after the projection, i.e. $\Pi_u^{\perp} G\,|v\rangle\,\langle v|\,G\Pi_u^{\perp} = 0$. Indeed, it does because $G\,|v\rangle \propto |u\rangle$. The non-vanishing contribution must then come from the first term in Equation (31), $\Pi_u^{\perp} G \Pi_u^{\perp} = (\mathbb{I} - |u\rangle\,\langle u|)\,G\,(\mathbb{I} - |u\rangle\,\langle u|)$, which entails

$$W^{\dashv} = \sqrt{\frac{\langle G\rangle}{\langle G^2\rangle}}\Pi_u^{\perp} G \Pi_u^{\perp} = \sqrt{\frac{\langle G\rangle}{\langle G^2\rangle}}\left(G - \frac{G^2\,|v\rangle\,\langle v|\,G}{\langle G^2\rangle} - \frac{G\,|v\rangle\,\langle v|\,G^2}{\langle G^2\rangle} + \frac{\langle G^3\rangle}{\langle G^2\rangle}\frac{G\,|v\rangle\,\langle v|\,G}{\langle G^2\rangle}\right).$$

The case for $G \geq 0$ where G has zero eigenvalues carries through. This can be seen by viewing the Sherman Morrison formula as a "correction" to an inverse when one entry of the matrix is changed. The inverse of G we are interested in is the positive inverse $G^{\dashv}$. The entry of the matrix that we change is in this positive subspace. Restricting the analysis to this subspace, the matrix G can be viewed as positive, yielding the required generalization. □

G The $-1/x$ transformation

In Section 6.2 we claimed that the $f-$assignments transform in a useful way under $x_i \mapsto 1/x_i$, and if O is the solution to an $f-$assignment, t, then O^T is the solution to the assignment that is derived from t under the aforementioned transformation. Here, we prove this claim. Recall Lemma 69 which tells us that a function is EBRM in $[\chi, \xi]$ if and only if it is $[\chi, \xi]-$valid. For the operator monotone $f_\lambda(x) = \frac{-1}{\lambda + x}$, this corresponds to requiring $\sum_i p_i f_\lambda(x_i) \geq 0$ for all $\lambda \in (-\infty, \infty)\backslash[-\xi, -\chi]$, permitting us to replace $[\![x_i]\!]$ with $[\![1/x_i]\!]$ at the cost of a minus sign. Notice that we can also use this transformation to convert left to right roots (see Appendix D.2).

Lemma 157. *Let $\chi, \xi > 0$. A function $t = \sum_i p_i [\![x_i]\!]$ is EBRM in $[\chi, \xi]$ if and only if the function $t' = \sum_i -p_i [\![1/x_i]\!]$ is EBRM in $[1/\xi, 1/\chi]$. Further, if O solves the matrix instance corresponding to t with spectrum in $[\chi, \xi]$, then O^T solves the matrix instance corresponding to t' with spectrum in $[1/\xi, 1/\chi]$.*

Proof. We start with the only if part ($\Rightarrow$). We are given H, G with spectrum in $[\chi, \xi]$ and a vector $|w\rangle$ such that $t = \text{Prob}[H, |w\rangle] - \text{Prob}[G, |w\rangle]$ and $H \geq G$. Further, $H \geq G \Leftrightarrow H^{-1} \leq G^{-1}$. Using the spectral decomposition we have $t' = \text{Prob}[G^{-1}, |w\rangle] - \text{Prob}[H^{-1}, |w\rangle]$. Defining $H' := G^{-1}, G' := H^{-1}$ and

$|w'\rangle = |w\rangle$, we have $t' = \text{Prob}[H', |w'\rangle] - \text{Prob}[G', |w'\rangle]$ and $H' \geq G'$, where H' and G' have their spectra in $[1/\xi, 1/\chi]$. The other direction ($\Leftarrow$) follows similarly by using a basis in which $H = X_h$ is diagonal, writing $G = OX_g O^T$ and noting that $O^{-1} = O^T$. □

Corollary 158. *Let $0 < x_1 < x_2 < \cdots < x_n$. Then, O^T solves a matrix instance corresponding to*

$$t = \sum_{i=1}^{n} \frac{\left(-\frac{1}{x_i}\right)^k}{\prod_{i \neq j}\left(\frac{1}{x_j} - \frac{1}{x_i}\right)} [\![x_i]\!],$$

if and only if O solves the corresponding matrix instance associated with the monomial assignment

$$t' = \sum_{i=1}^{n} \frac{-\left(-\frac{1}{x_i}\right)^k}{\prod_{i \neq j}\left(\frac{1}{x_j} - \frac{1}{x_i}\right)} [\![1/x_i]\!] = \sum_{i=1}^{n} \frac{-(-\omega_i)^k}{\prod_{i \neq j}(\omega_j - \omega_i)} [\![\omega_i]\!],$$

where $\omega_i = 1/x_i$.

Example 159 (Solving the simplest monomial assignment)**.** Suppose the assignment we wish to solve is

$$t = \sum_{i=1}^{2n} -\frac{(-x_i)^{2n-2}}{\prod_{j \neq i}(x_j - x_i)} [\![x_i]\!] = \sum_{i=1}^{2n} \tilde{p}_i [\![x_i]\!],$$

where $0 < x_1 < x_2 \cdots < x_n$. This can be solved using the f_0−solution (see Proposition 91) by writing $t = \sum_{i=1}^{2n} \frac{1}{\prod_{j \neq i}(\omega_j - \omega_i)} [\![x_i]\!]$, where $\omega_i = 1/x_i$, which is in turn equivalent to solving $t' = \sum_{i=1}^{2n} -\frac{1}{\prod_{j \neq i}(\omega_j - \omega_i)} [\![\omega_i]\!]$ (see Corollary 158 with $k = 0$). Instead, we solve this problem using another method; we use X^{-1} instead of X as in the usual f_0−solution and the fact that $\sum_i \tilde{p}_i x_i^{-k} = 0$ for $k \leq 2n - 2$ (see Lemma 33). Let us write t as

$$t = \sum_{i=1}^{n} \tilde{p}_{h_i} [\![x_{h_i}]\!] - \sum_{i=1}^{n} \tilde{p}_{g_i} [\![x_{g_i}]\!] = \sum_{i=1}^{n} x_{h_i}^{2n-2} p_{h_i} [\![x_{h_i}]\!] - \sum_{i=1}^{n} x_{g_i}^{2n-2} p_{g_i} [\![x_{g_i}]\!].$$

Let the matrix instance corresponding to t be given by $\underline{\mathrm{X}}^{\bar{n}} := \left(X_h^{\bar{n}}, X_g^{\bar{n}}, (X_h^{\bar{n}})^{n-1} \left|w^{\bar{n}}\right\rangle, (X_g^{\bar{n}})^{n-1} \left|v^{\bar{n}}\right\rangle\right)$, where

$$X_h^{\bar{n}} \doteq \text{diag}(x_{h_1}, x_{h_2} \ldots x_{h_n}), \qquad X_g^{\bar{n}} \doteq \text{diag}(x_{g_1}, x_{g_2} \ldots x_{g_n}),$$
$$\left|w^{\bar{n}}\right\rangle \doteq (\sqrt{p_{h_1}}, \sqrt{p_{h_2}} \ldots \sqrt{p_{h_n}}), \qquad \left|v^{\bar{n}}\right\rangle \doteq (\sqrt{p_{g_1}}, \sqrt{p_{g_2}} \ldots \sqrt{p_{g_n}}).$$

Solving the matrix instance $\underline{\mathrm{X}}^{\bar{n}}$ requires us to find an orthogonal matrix O such that $X_h^{\bar{n}} \geq OX_g^{\bar{n}}O^T$ and $O(X_g^{\bar{n}})^{n-1} \left|v^{\bar{n}}\right\rangle = (X_h^{\bar{n}})^{n-1} \left|w^{\bar{n}}\right\rangle$. The matrix inequality can be equivalently written as $\tilde{X}_h^{\bar{n}} \leq O\tilde{X}_g^{\bar{n}}O^T$, where $\tilde{X}_h^{\bar{n}} = (X_h^{\bar{n}})^{-1}$ and $\tilde{X}_g^{\bar{n}} = (X_g^{\bar{n}})^{-1}$. Note that under a change of the direction of the matrix inequality the arguments used in the proof of Lemma 90 go through unchanged. We can therefore consider the matrix instance $\underline{\tilde{\mathrm{X}}}^{\bar{n}} := \left(\tilde{X}_h^{\bar{n}}, \tilde{X}_g^{\bar{n}}, \left|\tilde{w}^{\bar{n}}\right\rangle, \left|\tilde{v}^{\bar{n}}\right\rangle\right)$, where $\left|\tilde{w}^{\bar{n}}\right\rangle := (X_h^{\bar{n}})^{n-1} \left|w^{\bar{n}}\right\rangle$ and $\left|\tilde{v}^{\bar{n}}\right\rangle := (X_g^{\bar{n}})^{n-1} \left|v^{\bar{n}}\right\rangle$. After iterating for l steps, suppose the matrix instance one obtains is $\underline{\tilde{X}}^{\overline{n-l}}$. To check if another isometric iteration is possible, we must check if the contact condition (see Definition 93) holds, i.e. if

$$\begin{aligned}
&\left\langle \tilde{w}^{\overline{n-l}} \middle| \tilde{H}^{\overline{n-l}} \middle| \tilde{w}^{\overline{n-l}} \right\rangle - \left\langle \tilde{v}^{\overline{n-l}} \middle| \tilde{G}^{\overline{n-l}} \middle| \tilde{v}^{\overline{n-l}} \right\rangle \\
&= r\left(\left\langle \tilde{w}^{\bar{n}} \middle| (\tilde{X}_h^{\bar{n}})^1 \middle| \tilde{w}^{\bar{n}} \right\rangle, \left\langle \tilde{w}^{\bar{n}} \middle| (\tilde{X}_h^{\bar{n}})^2 \middle| \tilde{w}^{\bar{n}} \right\rangle \dots, \left\langle \tilde{w}^{\bar{n}} \middle| (\tilde{X}_h^{\bar{n}})^{2l+1} \middle| \tilde{w}^{\bar{n}} \right\rangle\right) \\
&\qquad - r\left(\left\langle \tilde{v}^{\bar{n}} \middle| (\tilde{X}_g^{\bar{n}})^1 \middle| \tilde{v}^{\bar{n}} \right\rangle, \left\langle \tilde{v}^{\bar{n}} \middle| (\tilde{X}_g^{\bar{n}})^2 \middle| \tilde{v}^{\bar{n}} \right\rangle \dots, \left\langle \tilde{v}^{\bar{n}} \middle| (\tilde{X}_g^{\bar{n}})^{2l+1} \middle| \tilde{v}^{\bar{n}} \right\rangle\right) \\
&= r\left(\left\langle (X_h^{\bar{n}})^{2n-3} \right\rangle, \left\langle (X_h^{\bar{n}})^{2n-4} \right\rangle \dots, \left\langle (X_h^{\bar{n}})^{2n-2l-3} \right\rangle\right) \\
&\qquad - r\left(\left\langle (X_g^{\bar{n}})^{2n-3} \right\rangle, \left\langle (X_g^{\bar{n}})^{2n-4} \right\rangle \dots, \left\langle (X_g^{\bar{n}})^{2n-2l-3} \right\rangle\right)
\end{aligned}$$

vanishes. We used Lemma 92 (with $m = 1$) to obtain the RHS, and we continue using the convention that $\left\langle (X_h^{\bar{n}})^k \right\rangle = \langle w^{\bar{n}} | (X_h^{\bar{n}})^k | w^{\bar{n}} \rangle$ and similarly $\left\langle (X_g^{\bar{n}})^k \right\rangle = \langle v^{\bar{n}} | (X_g^{\bar{n}})^k | v^{\bar{n}} \rangle$. Recall that from Equation (25)

$$\left\langle (H^{\bar{n}})^k \right\rangle - \left\langle (G^{\bar{n}})^k \right\rangle = \left\langle x^k \right\rangle. \tag{32}$$

If $0 \le 2n - 2l - 3 \le 2n - 2$, then from Lemma 33 it follows that both terms become identical and hence the difference indeed vanishes (one can similarly verify the component condition). Therefore, until $l = n - 2$ (included), one can apply the Weingarten Iteration to obtain $\left|\tilde{u}_h^{\bar{n}}\right\rangle, \left|\tilde{u}_h^{\overline{n-1}}\right\rangle, \dots, \left|\tilde{u}_h^{\overline{n-l}}\right\rangle, \dots, \left|\tilde{u}_h^{\bar{1}}\right\rangle$ and $\left|\tilde{u}_g^{\bar{n}}\right\rangle, \left|\tilde{u}_g^{\overline{n-1}}\right\rangle, \dots, \left|\tilde{u}_g^{\overline{n-l}}\right\rangle, \dots, \left|\tilde{u}_g^{\bar{1}}\right\rangle$, which completely determine $O = \sum_{i=1}^n \left|\tilde{u}_h^i\right\rangle \left\langle\tilde{u}_g^i\right|$. The argument can, as before, be concisely represented using a diagram (see Figure 24).

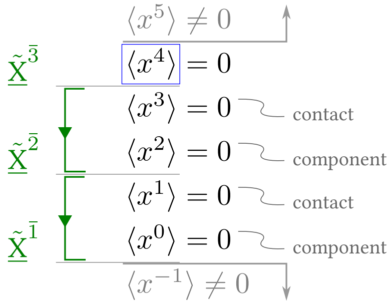

$\langle x^5 \rangle \neq 0$
$\underline{\tilde{\mathrm{X}}}^{\bar{3}}$
$\boxed{\langle x^4 \rangle} = 0$
$\langle x^3 \rangle = 0$
contact
$\underline{\tilde{\mathrm{X}}}^{\bar{2}}$
$\langle x^2 \rangle = 0$
component
$\langle x^1 \rangle = 0$
contact
$\underline{\tilde{\mathrm{X}}}^{\bar{1}}$
$\langle x^0 \rangle = 0$
component
$\langle x^{-1} \rangle \neq 0$

Figure 24: Power diagram representative of the simplest monomial assignment for $2n = 6$ points.

H Existence of solutions to matrix instances and their dimensions

Our goal here is to show that certain matrix instances can be solved with low-dimensional matrices.

From Lemma 160 and Lemma 161 below, we know that a solution to a matrix instance corresponding to a $[\chi, \xi]$-valid function always exists, granted that we pad the matrices with χ and ξ to have their size equal to $n \times n$ with $n = n_g + n_h - 1$. We can, however, do even better. Consider the matrix instance $\underline{\mathrm{X}}^{\bar{k}}$ in the notation introduced in Lemma 161. The eigenspace of H on which $|w\rangle$ has a component is of size n_h (similarly for G, $|v\rangle$ and n_g). Every time we iterate using the Weingarten map, we remove one component from both H and $|w\rangle$ from within this eigenspace (similarly for G and $|v\rangle$). Consequently, in the subsequent step, the eigenspace of $H^{\overline{k-1}}$ on which $\left|w^{\overline{k-1}}\right\rangle$ has a component, is of size $n_h - 1$ (similarly for $G^{\overline{k-1}}$, $\left|v^{\overline{k-1}}\right\rangle$ the size becomes $n_g - 1$) where the matrix instance after the Weingarten Iteration map was taken to be $\underline{\mathrm{X}}^{\overline{k-1}} =: (H^{\overline{k-1}}, G^{\overline{k-1}}, \left|w^{\overline{k-1}}\right\rangle, \left|v^{\overline{k-1}}\right\rangle)$. For the balanced f_0-assignment, we end up with a matrix instance

$\underline{\mathrm{X}}^{l}$ =: $(H^{\bar{l}}, G^{\bar{l}}, 0, 0)$ where the vectors disappear. The matrices $H^{\bar{l}}$ and $G^{\bar{l}}$ only have ξ and χ, respectively, as their eigenvalues and then, we trivially have $H^{\bar{l}} > G^{\bar{l}}$. In fact, this part of the matrix plays no role and can be removed. This justifies why we could assume that even without padding with χs and ξs, the matrix instance corresponding to the f_0-assignment had a solution.

The padding becomes important, however, when we use the wiggle-w (or wiggle-v) map to iterate. Consider again the matrix instance $\underline{X}^{\bar{k}}$ in the notation introduced in Lemma 161 and note that $\xi \to \infty$ in these cases. The eigenspace of H on which $|w\rangle$ has a component is of size n_h. Whenever we iterate using the wiggle-w map, we do not remove any component from H and $|w\rangle$ from within this eigenspace. This is because we introduce an extra dimension, and then project out one dimension, leaving the overall dimension of the space unchanged. The dimension for the G and $|v\rangle$ case, however, drops as before. Again, when we reach a matrix instance $\underline{\mathrm{X}}^{\bar{l}}$, where the vectors disappear we can use the reasoning above to justify that matrices with fewer padded dimensions also have a solution.

Lemma 160. *(Restatement of Lemma 146) Let $t = h - g = \sum_{i=1}^{m} p_i [\![x_i]\!]$ be a $[\chi, \xi]$-valid function where $h =: \sum_{i=1}^{n_h} p_{h_i} [\![x_{h_i}]\!]$ and $g =: \sum_{i=1}^{n_g} p_{g_i} [\![x_{g_i}]\!]$ have disjoint support and $p_{h_i}, p_{g_i} > 0$ for $i \in \{1, 2 \dots n_h\}$ and $\{1, 2 \dots n_g\}$, respectively. Let X_h and X_g be $n \times n$ diagonal matrices, where $n = n_h + n_g - 1$, given by*

$$X_h = diag(x_{h_1}, x_{h_2}, \dots x_{h_{n_h}}, \xi, \xi \dots \xi) \text{ and } X_g = diag(x_{g_1}, x_{g_2}, \dots x_{g_{n_g}}, \chi, \chi \dots \chi).$$

Then, there exists an orthogonal matrix O which solves the matrix instance $\underline{X}^{\bar{n}} := (X_h, X_g, |w\rangle, |v\rangle)$.

Lemma 161. *Let k, n_h and n_g be strictly positive integers such that $k \geq n_h$ and $k \geq n_g$. Consider a matrix instance $\underline{X}^{\bar{k}} =: (H, G, |w\rangle, |v\rangle)$ where*

$$H = \sum_{i=1}^{n_h} x_{h_i} |h_i\rangle \langle h_i| + \sum_{i=n_h+1}^{k} \xi |h_i\rangle \langle h_i|, \qquad |w\rangle = \sum_{i=1}^{n_h} \sqrt{p_{h_i}} |h_i\rangle$$

$$\text{and } \; G = \sum_{i=1}^{n_g} x_{g_i} |g_i\rangle \langle g_i| + \sum_{i=n_g+1}^{k} \chi |g_i\rangle \langle g_i|, \qquad |v\rangle = \sum_{i=1}^{n_g} \sqrt{p_{g_i}} |g_i\rangle$$

such that $x_{h_i} \neq x_{g_j}, p_{h_i}, p_{g_j} > 0$ hold for all $i \in \{1, 2 \dots n_h\}$, $j \in \{1, 2 \dots n_g\}$, and $\mathcal{H}^{\bar{k}} = span\{|h_i\rangle\}$, $\mathcal{G}^{\bar{k}} = span\{|g_i\rangle\}$ (see Definition 93). If the isometry $Q : \mathcal{H}^{\bar{k}} \to \mathcal{G}^{\bar{k}}$ solves the matrix instance $\underline{X}^{\bar{k}}$, then the function

$$t = \sum_{i=1}^{n_h} p_{h_i} [\![x_{h_i}]\!] - \sum_{i=1}^{n_g} p_{g_i} [\![x_{g_i}]\!]$$

is $[\chi, \xi]$-valid; which is equivalent to being $[\chi, \xi]$-EBRM.

I Lemmas for the contact and component conditions

Lemma 162. *Consider the matrix instance $\underline{X}^{\bar{n}} := (H^{\bar{n}}, G^{\bar{n}}, |w^{\bar{n}}\rangle, |v^{\bar{n}}\rangle)$. Suppose that the Weingarten Iteration Map (see Definition 95) is applied l times to obtain $\underline{X}^{\overline{n-l}} := \left(H^{\overline{n-l}}, G^{\overline{n-l}}, \left|w^{\overline{n-l}}\right\rangle, \left|v^{\overline{n-l}}\right\rangle\right)$. Then, for any l, the expectation value $\left\langle v^{\overline{n-l}}\right| G^{\overline{n-l}} \left|v^{\overline{n-l}}\right\rangle$ is a function of the expectation values $\left\langle v^{\bar{n}}\right| (G^{\bar{n}})^p \left|w^{\bar{n}}\right\rangle = \langle (G^{\bar{n}})^p \rangle$, where the powers p range from 0 to $2l + 1$ at most. The corresponding statement for H and $|w\rangle$ also holds.*

Proof. Using once the Weingarten Iteration Map we obtain:

$$\left|v^{\overline{n-1}}\right\rangle = \left|v^{\bar{n}}\right\rangle - \frac{\langle G^{\bar{n}}\rangle}{\langle (G^{\bar{n}})^2\rangle} G^{\bar{n}} \left|v^{\bar{n}}\right\rangle \tag{33}$$

$$G^{\overline{n-1}} = G^{\bar{n}} + \frac{\langle (G^{\bar{n}})^3\rangle}{\langle (G^{\bar{n}})^2\rangle^2} G^{\bar{n}} \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| G^{\bar{n}} - \frac{1}{\langle (G^{\bar{n}})^2\rangle} \left(G^{\bar{n}} \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^2 + (G^{\bar{n}})^2 \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| G^{\bar{n}} \right). \tag{34}$$

If we continue to iterate accordingly and express everything in terms of $\left|v^{\bar{n}}\right\rangle$ and $G^{\bar{n}}$, which are known, after l steps we obtain:

$$\left|v^{\overline{n-l}}\right\rangle = \sum_{i=0}^{l} \alpha_i (G^{\bar{n}})^i \left|v^{\bar{n}}\right\rangle \tag{35}$$

$$G^{\overline{n-l}} = G^{\bar{n}} + \sum_{i,j=0}^{l+1} \alpha_{i,j} (G^{\bar{n}})^i \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^j, \tag{36}$$

where the multiplicative factors a_i and $a_{i,j}$ also contain terms of the form $\langle (G^{\bar{n}})^p\rangle$, in which p ranges between the minimum and maximum powers appearing in the sum; see Remark 163.

Indeed, we can use induction to prove that Equation (35) and Equation (36) hold for all l.
The base of the induction $l = 1$ immediately gives us Equation (33) and Equation (34), , and for $l + 1$ the Weingarten Iteration Map gives

$$\left|v^{\overline{n-l-1}}\right\rangle = \left|v^{\overline{n-l}}\right\rangle - \frac{\langle G^{\overline{n-l}}\rangle}{\langle (G^{\overline{n-l}})^2\rangle} (G^{\overline{n-l}}) \left|v^{\overline{n-l}}\right\rangle$$

$$G^{\overline{n-l-1}} = G^{\bar{n}} + \frac{\langle (G^{\overline{n-l}})^3\rangle}{\langle (G^{\overline{n-l}})^2\rangle^2} G^{\overline{n-l}} \left|v^{\overline{n-l}}\right\rangle \left\langle v^{\overline{n-l}}\right| G^{\overline{n-l}} - \frac{1}{\langle (G^{\overline{n-l}})^2\rangle} \left(G^{\overline{n-l}} \left|v^{\overline{n-l}}\right\rangle \left\langle v^{\overline{n-l}}\right| (G^{\overline{n-l}})^2 + (G^{\overline{n-l}})^2 \left|v^{\overline{n-l}}\right\rangle \left\langle v^{\overline{n-l}}\right| G^{\overline{n-l}} \right).$$

Replacing $\left|v^{\overline{n-l}}\right\rangle$ and $G^{\overline{n-l}}$ from Equation (35) and Equation (36) we get

$$\left|v^{\overline{n-l-1}}\right\rangle = \sum_{i=0}^{l+1} \alpha_i (G^{\bar{n}})^i \left|v^{\bar{n}}\right\rangle \tag{37}$$

$$G^{\overline{n-l-1}} = G^{\bar{n}} + \sum_{i,j=0}^{l+2} \alpha_{i,j} (G^{\bar{n}})^i \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^j, \tag{38}$$

which proves that Equation (35) and Equation (36) are valid for all l.
We can now complete our proof by expressing $\left\langle v^{\overline{n-l}}\right| G^{\overline{n-l}} \left|v^{\overline{n-l}}\right\rangle$ in terms of $\langle G^{\bar{n}}\rangle$. Substituting from Equation (35) and Equation (36), we get:

$$\begin{aligned} \left\langle v^{\overline{n-l}}\right| G^{\overline{n-l}} \left|v^{\overline{n-l}}\right\rangle &= \sum_{i=0}^{l} \alpha_i \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^{i+1} \sum_{j=0}^{l} \alpha_j (G^{\bar{n}})^j \left|v^{\bar{n}}\right\rangle \\ &+ \sum_{i=0}^{l} \alpha_i \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^i \sum_{i',j'=0}^{l+1} \alpha_{i',j'} (G^{\bar{n}})^{i'} \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^{j'} \sum_{j=0}^{l} \alpha_j (G^{\bar{n}})^j \left|v^{\bar{n}}\right\rangle. \end{aligned} \tag{39}$$

In Equation (39), we see that the minimum expectation value is $\langle (G^{\bar{n}})^0\rangle$, while the maximum is $\langle (G^{\bar{n}})^{2l+1}\rangle$, which concludes the proof. □

Remark 163. Notice that we left a_i and $a_{i,j}$ undetermined and we even used the same notation for them; obviously a_i and $a_{i,j}$ are different in Equation (35), Equation (36), Equation (37),Equation (38) and Equation (39). For our proof their specific form is not relevant, but what is rather important are the minimum and maximum powers, p, in $\langle (G^{\bar{n}})^p\rangle$ that might appear in $\left\langle v^{\overline{n-l}}\right| G^{\overline{n-l}} \left|v^{\overline{n-l}}\right\rangle$. To estimate them, it suffices to observe that the minimum power in $\left|v^{\overline{n-l}}\right\rangle$ comes from the first term $\left|v^{\bar{n}}\right\rangle$ and is 0, while the maximum power that appears in $\left|v^{\overline{n-l}}\right\rangle$ comes from $\langle (G^{\overline{n-l+1}})^2\rangle$ (see Definition 95) and is equal to $2l$. In $G^{\overline{n-l}}$, however, we can find an even higher power appearing in the $a_{i,j}$'s coming from $\langle (G^{\overline{n-l+1}})^3\rangle$ (see Definition 95) and is equal to $2l+1$. In total these powers are always between the minimum and maximum powers on Equation (39), thus the factors a_i and $a_{i,j}$ do not need to be specified.

Lemma 164. *Consider the extended matrix instance* $\underline{M}^{\bar{n}} := \mathscr{U}(H^{\bar{n}}, G^{\bar{n}}, \left|w^{\bar{n}}\right\rangle, \left|v^{\bar{n}}\right\rangle, (H^{\bar{n}})^{\dashv}, (G^{\bar{n}})^{\dashv}, |.\rangle, |.\rangle)$*. Suppose the Normal Initialization Map and the Weingarten Iteration Map (see Definition 94 and Definition 95) are applied* l *times to obtain* $\underline{M}^{\overline{n-l}}$*, viz. applying* $\underline{M}^{\overline{i-1}} = \mathscr{U}(\mathscr{W}(\underline{M}^{\bar{i}}))$ l *times. Then, for any* l*, the expectation value* $\left\langle v^{\overline{n-l}}\right| (G^{\overline{n-l}})^{\dashv} \left|v^{\overline{n-l}}\right\rangle$ *is a function of the expectation values* $\left\langle v^{\bar{n}}\right| (G^{\bar{n}})^p \left|w^{\bar{n}}\right\rangle = \langle (G^{\bar{n}})^p\rangle$*, where the powers* p *range from* 0 *to* $2l+1$ *at most. The corresponding statement for* H *and* $|w\rangle$ *also holds.*

Proof. First, we need to specify the form of $(G^{\overline{n-l}})^{\dashv}$ as a function of $G^{\bar{n}}$ and $\left|v^{\bar{n}}\right\rangle$. The first iteration gives

$$\left|v^{\overline{n-1}}\right\rangle = \left|v^{\bar{n}}\right\rangle - \frac{\langle G^{\bar{n}}\rangle}{\langle (G^{\bar{n}})^2\rangle} G^{\bar{n}} \left|v^{\bar{n}}\right\rangle \tag{40}$$

$$(G^{\overline{n-1}})^{\dashv} = (G^{\bar{n}})^{\dashv} - \frac{\left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right|}{\langle G^{\bar{n}}\rangle}. \tag{41}$$

Continuing the iterations to l and using Lemma 162 we obtain:

$$\left|v^{\overline{n-l}}\right\rangle = \sum_{i=0}^{l} \alpha_i (G^{\bar{n}})^i \left|v^{\bar{n}}\right\rangle \tag{42}$$

$$(G^{\overline{n-l}})^{\dashv} = (G^{\bar{n}})^{\dashv} + \sum_{i,j=0}^{l-1} \alpha_{i,j} (G^{\bar{n}})^i \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^j. \tag{43}$$

Indeed, by induction we can prove that Equation (42) and Equation (43) hold for all l.
The base of the induction $l = 1$ immediately gives us Equation (40) and Equation (41), which hold.
For the $l+1$ instance, the Weingarten Iteration Map gives us:

$$\left|v^{\overline{n-l-1}}\right\rangle = \left|v^{\overline{n-l}}\right\rangle - \frac{\langle G^{\overline{n-l}}\rangle}{\langle (G^{\overline{n-l}})^2\rangle} (G^{\overline{n-l}}) \left|v^{\overline{n-l}}\right\rangle \tag{44}$$

$$(G^{\overline{n-l-1}})^{\dashv} = (G^{\overline{n-l}})^{\dashv} - \frac{\left|v^{\overline{n-l}}\right\rangle \left\langle v^{\overline{n-l}}\right|}{\langle G^{\overline{n-l}}\rangle} \tag{45}$$

Replacing $\left|v^{\overline{n-l}}\right\rangle$ and $(G^{\overline{n-l}})^{\dashv}$ from Equation (42) and Equation (43), we get

$$\left|v^{\overline{n-l-1}}\right\rangle = \sum_{i=0}^{l+1} \alpha_i (G^{\bar{n}})^i \left|v^{\bar{n}}\right\rangle \tag{46}$$

$$(G^{\overline{n-l-1}})^{\dashv} = (G^{\bar{n}}) + \sum_{i,j=0}^{l} \alpha_{i,j} (G^{\bar{n}}) \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^j, \tag{47}$$

which concludes our inductive proof.

Now that we proved that Equation (42) and Equation (43) hold for any l, we can proceed to the calculation of the corresponding expectation value:

$$\begin{aligned}\langle (G^{\overline{n-l}})^{\dashv}\rangle &= \left\langle v^{\overline{n-l}}\right| (G^{\overline{n-l}}) \left| v^{\overline{n-l}}\right\rangle = \sum_{i,j=0}^{l} \alpha_i\alpha_j \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^{i+j-1} \left| v^{\bar{n}}\right\rangle \\ &+ \sum_{i=0}^{l} \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^{i} \sum_{i',j'=0}^{l-1} \alpha_{i',j'} (G^{\bar{n}})^{i'} \left| v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^{j'} \sum_{j=0}^{l} \alpha_j (G^{\bar{n}})^{j} \left| v^{\bar{n}}\right\rangle, \end{aligned} \tag{48}$$

where we have used $(G^{\bar{n}})^{\dashv} = (G^{\bar{n}})^{-1}$, since $G^{\bar{n}}$ is full rank.
Observe that the minimum power in the expectation value is $\langle G^{\bar{n}}\rangle$, while the maximum is $\langle (G^{\bar{n}})^{2l-1}\rangle$. Recall though that in the multiplicative factors α_i and $\alpha_{i,j}$ there are higher powers in the expectation values $\langle (G^{\bar{n}})^{2l+1}\rangle$, which from now on are the highest. Since we are iterating with respect to $G^{\dashv}$ the powers are not growing any more, but they rather decrease and we are interested on the minimum powers that are reduced with each iteration. □

Lemma 165. *Consider the extended matrix instance*
$\underline{\tilde{M}}^{\bar{n}} := \mathscr{U}((H^{\bar{n}})^{\dashv}, (G^{\bar{n}})^{\dashv}, \left|\tilde{w}^{\bar{n}}\right\rangle, \left|\tilde{v}^{\bar{n}}\right\rangle, H^{\bar{n}}, G^{\bar{n}}, |.\rangle, |.\rangle)$. *Suppose the Normal Initialization Map and the Weingarten Iteration Map (see Definition 94 and Definition 95) are applied k times to obtain $\underline{\tilde{M}}^{\overline{n-k}}$, viz. applying $\underline{\tilde{M}}^{\overline{i-1}} = \mathscr{U}(\mathscr{W}(\underline{\tilde{M}}^{\bar{i}}))$ k times. Then, for any k, the expectation value $\left\langle \tilde{v}^{\overline{d-k}}\right| \tilde{G}^{\overline{d-k}} \left|\tilde{v}^{\overline{d-k}}\right\rangle$ is a function of the expectation values $\left\langle v^{\bar{n}}\right| (G^{\bar{n}})^{p} \left|w^{\bar{n}}\right\rangle = \langle (G^{\bar{n}})^{p}\rangle$, where the minimum power p that might appear is $-(2k+1)$. The corresponding statement for H and $|w\rangle$ also holds.*

Proof. The first iteration gives:

$$\left|\tilde{v}^{\overline{d-1}}\right\rangle = \left|\tilde{v}^{\bar{d}}\right\rangle - \frac{\langle \tilde{G}^{\bar{d}}\rangle}{\langle (\tilde{G}^{\bar{d}})^2\rangle} \tilde{G}^{\bar{d}} \left|\tilde{v}^{d}\right\rangle \tag{49}$$

$$\tilde{G}^{\overline{d-1}} = \tilde{G}^{\bar{d}} + \frac{\langle (\tilde{G}^{\bar{d}})^3\rangle}{\langle (\tilde{G}^{\bar{d}})^2\rangle^2} \tilde{G}^{\bar{d}} \left|\tilde{v}^{\bar{d}}\right\rangle \left\langle \tilde{v}^{\bar{d}}\right| \tilde{G}^{\bar{d}} - \frac{1}{\langle (\tilde{G}^{\bar{d}})^2\rangle} \left(\tilde{G}^{\bar{d}} \left|\tilde{v}^{\bar{d}}\right\rangle \left\langle \tilde{v}^{\bar{d}}\right| (\tilde{G}^{\bar{d}})^2 + (\tilde{G}^{\bar{d}})^2 \left|\tilde{v}^{\bar{d}}\right\rangle \left\langle \tilde{v}^{\bar{d}}\right| \tilde{G}^{\bar{d}} \right). \tag{50}$$

Continuing for k iterations, we can prove by induction that:

$$\left|\tilde{v}^{\overline{d-k}}\right\rangle = \sum_{i=0}^{k} \alpha_i (G^{\bar{n}})^{i-k} \left|v^{\bar{n}}\right\rangle \tag{51}$$

$$\tilde{G}^{\overline{d-k}} = (G^{\bar{n}})^{\dashv} + \sum_{i,j=0}^{k} \alpha_{i,j} (G^{\bar{n}})^{i-(k+1)} \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^{j-(k+1)}. \tag{52}$$

The base of the induction $k = 1$ gives us Equation (49) and Equation (50), which hold, while for $k + 1$, we obtain:

$$\begin{aligned}\left|\tilde{v}^{\overline{d-k-1}}\right\rangle &= \left|\tilde{v}^{\overline{d-k}}\right\rangle - \frac{\langle \tilde{G}^{\overline{d-k}}\rangle}{\langle (\tilde{G}^{\overline{d-k}})^2\rangle} \tilde{G}^{\overline{d-k}} \left|\tilde{v}^{\overline{d-k}}\right\rangle \\ \tilde{G}^{\overline{d-k-1}} &= \tilde{G}^{\overline{d-k}} + \frac{\langle (\tilde{G}^{\overline{d-k}})^3\rangle}{\langle (\tilde{G}^{\overline{d-k}})^2\rangle^2} \tilde{G}^{\overline{d-k}} \left|\tilde{v}^{\overline{d-k}}\right\rangle \left\langle \tilde{v}^{\overline{d-k}}\right| \tilde{G}^{\overline{d-k}} \\ &- \frac{1}{\langle (\tilde{G}^{\overline{d-k}})^2\rangle} \left(\tilde{G}^{\overline{d-k}} \left|\tilde{v}^{\overline{d-k}}\right\rangle \left\langle \tilde{v}^{\overline{d-k}}\right| (\tilde{G}^{\overline{d-k}})^2 + (\tilde{G}^{\overline{d-k}})^2 \left|\tilde{v}^{\overline{d-k}}\right\rangle \left\langle \tilde{v}^{\overline{d-k}}\right| \tilde{G}^{\overline{d-k}} \right). \end{aligned}$$

Substituting $\left|\tilde{v}^{\overline{d-k}}\right\rangle$ and $\tilde{G}^{\overline{d-k}}$ from Equation (51) and Equation (52), we get:

$$\begin{aligned}\left|\tilde{v}^{\overline{d-k-1}}\right\rangle &= \sum_{i=0}^{k+1} \alpha_i (G^{\bar{n}})^{i-k-1} \left|v^{\bar{n}}\right\rangle \\ \tilde{G}^{\overline{d-k-1}} &= (G^{\bar{n}})^{\dashv} + \sum_{i,j=0}^{k+1} \alpha_{i,j} (G^{\bar{n}})^{i-k-2} \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^{j-k-2},\end{aligned}$$

confirming that Equation (51) and Equation (52) hold for all k. Thus, for any k the corresponding expectation value can be written as:

$$\begin{aligned}\left\langle \tilde{v}^{\overline{d-k}}\right| \tilde{G}^{\overline{d-k}} \left|\tilde{v}^{\overline{d-k}}\right\rangle &= \sum_{i=0}^{k} \alpha_i \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^{i-k} (G^{\bar{n}})^{-1} \sum_{j=0}^{k} \alpha_j (G^{\bar{n}})^{j-k} \left|v^{\bar{n}}\right\rangle \\ &+ \sum_{i=0}^{l+k} \alpha_i \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^{i-k} \sum_{i',j'=0}^{k} \alpha_{i',j'} (G^{\bar{n}})^{i'-(k+1)} \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^{j'-(k+1)} \sum_{j=0}^{k} \alpha_j (G^{\bar{n}})^{j-k} \left|v^{\bar{n}}\right\rangle .\end{aligned}$$

We observe that the minimum power that can appear in the expectation values is $-(2k+1), \forall k$. The factors α_i and $\alpha_{i,j}$, also contain terms of the form $\langle (G^{\bar{n}})^p \rangle$, which behave as explained previously. □

Lemma 166. *Consider the matrix instance $\underline{X}^{\bar{n}} := (H^{\bar{n}}, G^{\bar{n}}, \left|w^{\bar{n}}\right\rangle, \left|v^{\bar{n}}\right\rangle)$. Using the Weingarten Iteration Map (see Definition 95) once, we obtain:*

$$\left|v^{\overline{n-1}}\right\rangle = \left|v^{\bar{n}}\right\rangle - \frac{\langle G^{\bar{n}} \rangle}{\langle (G^{\bar{n}})^2 \rangle} G^{\bar{n}} \left|v^{\bar{n}}\right\rangle \tag{53}$$

$$\begin{aligned}G^{\overline{n-1}} &= G^{\bar{n}} + \frac{\langle (G^{\bar{n}})^3 \rangle}{\langle (G^{\bar{n}})^2 \rangle^2} G^{\bar{n}} \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| G^{\bar{n}} \\ &- \frac{1}{\langle (G^{\bar{n}})^2 \rangle} \left(G^{\bar{n}} \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^2 + (G^{\bar{n}})^2 \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| G^{\bar{n}} \right).\end{aligned} \tag{54}$$

Then, for any power m, the expectation value $\left\langle v^{\overline{n-1}}\right| (G^{\overline{n-1}})^m \left|v^{\overline{n-1}}\right\rangle$ can be expressed in terms of the expectation values $\left\langle v^{\bar{n}}\right| (G^{\bar{n}})^p \left|v^{\bar{n}}\right\rangle = \langle (G^{\bar{n}})^p \rangle$ with p being at most $m+2$. The corresponding statement for H and $|w\rangle$ also holds.

Proof. The first step is to prove that for any power m:

$$(G^{\overline{n-1}})^m = (G^{\bar{n}})^m + \sum_{i,j=0}^{m+1} \alpha_{i,j} (G^{\bar{n}})^i \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^j \tag{55}$$

Note that some of the $\alpha_{i,j}$ can be zero. Indeed, we can use induction to prove Equation (55).
The base of the induction $m = 1$ gives us Equation (53) and Equation (54), which hold, while for $m+1$ we have

$$(G^{\overline{n-1}})^{m+1} = (G^{\overline{n-1}})^m \cdot G^{\overline{n-1}}, \tag{56}$$

and substituting from Equation (53), Equation (54) and Equation (55), we get

$$\begin{aligned}(G^{\overline{n-1}})^{m+1} &= \left[(G^{\bar{n}})^m + \sum_{i,j=0}^{m+1} \alpha_{i,j} (G^{\bar{n}})^i \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^j \right] \\ &\cdot \left[G^{\bar{n}} + \frac{\langle (G^{\bar{n}})^3 \rangle}{\langle (G^{\bar{n}})^2 \rangle^2} G^{\bar{n}} \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| G^{\bar{n}} - \frac{1}{\langle (G^{\bar{n}})^2 \rangle} \left(G^{\bar{n}} \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^2 + (G^{\bar{n}})^2 \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| G^{\bar{n}} \right) \right] \\ &= (G^{\bar{n}})^{m+1} + \sum_{i,j=0}^{m+2} \alpha_{i,j} (G^{\bar{n}n})^i \left|v^{\bar{n}}\right\rangle \left\langle v^{\bar{n}}\right| (G^{\bar{n}})^j,\end{aligned}$$

proving that Equation (55) holds for all m. With this in place, we can proceed to prove our main claim about the corresponding expectation value:

$$\begin{aligned}
&\left\langle v^{\overline{n-1}}\right| (G^{\overline{n-1}})^m \left|v^{\overline{n-1}}\right\rangle \\
&= \left(\left\langle v^{\bar{n}}\right| - \frac{\langle G^{\bar{n}}\rangle}{\langle (G^{\bar{n}n})^2\rangle}\left\langle v^{\bar{n}}\right| G^{\bar{n}}\right)\left((G^{\bar{n}})^m + \sum_{i,j=0}^{m+1}\alpha_{i,j}(G^{\bar{n}})^i\left|v^{\bar{n}}\right\rangle\left\langle v^{\bar{n}}\right|(G^{\bar{n}})^j\right)\left(\left|v^{\bar{n}}\right\rangle - \frac{\langle G^{\bar{n}}\rangle}{\langle (G^{\bar{n}})^2\rangle}G^{\bar{n}}\left|v^{\bar{n}}\right\rangle\right) \\
&= \langle (G^{\bar{n}})^m\rangle + a\langle (G^{\bar{n}})^{m+1}\rangle + b\langle (G^{\bar{n}})^{m+2}\rangle + \sum_{i,j=0}^{m+2}\alpha_{i,j}\langle (G^{\bar{n}})^i\rangle\langle (G^{\bar{n}})^j\rangle \\
&= \sum_{i,j=0}^{m+2}\alpha'_{i,j}\langle (G^{\bar{n}})^i\rangle\langle (G^{\bar{n}})^j\rangle,
\end{aligned}$$

which completes our proof that the highest power is $m + 2$ for any m. Notice that we did not fully specified the scalar factors $a, b, \alpha_{i,j}, \alpha'_{i,j}$, as it is easy to verify, as previously, that they do not contain any higher powers. □