Statistically Sender-Private OT From LPN and Derandomization*

Nir Bitansky† Sapir Freizeit‡

Tel Aviv University

February 2022

Abstract

We construct a two-message oblivious transfer protocol with statistical sender privacy (SSP OT) based on the Learning Parity with Noise (LPN) Assumption and a standard Nisan-Wigderson style derandomization assumption. Beyond being of interest on their own, SSP OT protocols have proven to be a powerful tool toward minimizing the round complexity in a wide array of cryptographic applications from proofs systems, through secure computation protocols, to hard problems in statistical zero knowledge (SZK).

The protocol is plausibly post-quantum secure. The only other construction with plausible post quantum security is that of Brakerski and Dötling (TCC '18) based on the Learning with Errors (LWE) Assumption. Lacking the geometric structure of LWE, our construction and analysis rely on a different set of techniques.

Technically, we first construct an SSP OT protocol in the common random string model from LPN alone, and then derandomize the common random string. Most of the technical difficulty lies in the first step. Here we prove a robustness property of the inner product randomness extractor to a certain type of linear splitting attacks. A caveat of our construction is that it relies on the so called *low noise regime* of LPN. This aligns with our current complexity-theoretic understanding of LPN, which only in the low noise regime is known to imply hardness in SZK.

1 Introduction

Learning Parity with Noise [BFKL93, BKW03] is a prominent hardness assumption in cryptography. The search version of the problem LPN_ε postulates that given access to polynomially many samples $(\mathbf{a}_i, \mathbf{a}_i^t\mathbf{s} + e_i)$ where $\mathbf{s} \leftarrow \mathbb{F}_2^n$ is a uniformly random secret, each $\mathbf{a}_i \leftarrow \mathbb{F}_2^n$ is a uniformly random vector, and each $e_i \leftarrow \mathrm{Bern}(\varepsilon)$ is a random Bernouli noise bit, it is hard to find the secret $\mathbf{s}$. In the decision version, which is equivalently hard [KSS10], the samples are indistinguishable from completely random samples, where $e_i \leftarrow \mathbb{F}_2$ is uniformly random.

Much of the appeal of the LPN assumption stems from its direct relation to the long-studied problem of decoding random linear codes, as well as its plausible resilience to quantum attacks. Furthermore, in terms of applications, LPN has led to simple and efficient constructions, for a both symmetric-key and asymmetric-key primitives (c.f. [Ale03, KPC+11, Pie12, DGH+19]). Yet, our understanding of LPN, both in terms of hardness and in terms of applications, still seems to be lacking, and in particular to be far behind our understanding of its cousin, the Learning with Errors (LWE) assumption [Reg05]. In LWE, instead of

*Supported by ISF grants 18/484 and 19/2137, by Len Blavatnik and the Blavatnik Family Foundation, by the Blavatnik Interdisciplinary Cyber Research Center at Tel Aviv University, and by the European Union Horizon 2020 Research and Innovation Program via ERC Project REACT (Grant 756482).

†nbitansky@gmail.com. Member of the checkpoint institute of information security.

‡sapirfreizeit@gmail.com

$\mathbb{F}_2$, we consider $\mathbb{F}_q$ for a large (at least polynomial) modulus q, and instead of Bernouli noise e_i, we consider (discrete) Gaussian noise e_i of norm $\ll q$.

Although the two have a similar flavour, the geometric structure endowed by the two mentioned differences has made LWE substantially more versatile than LPN. While LWE has led to a wide array of applications, including ground-breaking ones such as fully-homomorphic encryption [Gen09, BV11], the set of applications known from LPN is far more restricted (see also Section 1.3). At the same time, there is no formal indication that LPN is less powerful than LWE, and the effort to expand its reach continues.

Statistically Sender-Private OT. One powerful primitive that has been constructed from LWE [BD18] and has yet to be achieved under LPN is *two-message statistically sender-private oblivious transfer* (SSP OT in short) [NP01, AIR01]. Recall that in an OT protocol [Rab05, EGL85], the sender S holds two messages $(m_0\ m_1)$ and the receiver R holds a choice bit $c \in \{0,1\}$. The goal is for R to learn the message m_c of its choice, without learning anything on the other message m_{1-c}, and without having S learn anything about the choice c. SSP OT requires that this is done in minimal round complexity with a single message from the receiver R and a single message returned from the sender S. Security is also taken to the extreme, requiring that sender privacy, namely the hiding of m_{1-c}, is statistical (statistical receiver-privacy is impossible in this setting, as it would enable a non-uniform malicious receiver to learn both sender messages).

As for the formal security notion, the gold-standard simulation guarantee against malicious parties is known to be unobtainable (even with computational sender privacy), without reliance on some form of setup. In contrast, in the common random string model, Döttling, Garg, Hajiabadi, and Wichs [DGH$^+$19] construct a simulatable protocol with computational security (for both the receiver and sender) from $\mathrm{LPN}_{n^{\frac{1}{2}-\varepsilon}}$. The standard security notion in this setting, introduced in [NP01, AIR01], relaxes the simulation requirement in a meaningful manner. On the receiver side, receiver messages corresponding to different choice bits should be computationally indistinguishable. On the sender side, any receiver message information-theoretically fixes a choice $c^* \in \{0,1\}$, so that sender messages corresponding to different m_{1-c^*} are statistically indistinguishable.

Such SSP OT protocols have turned out to be highly useful, in particular toward obtaining protocols with low round complexity. They have been used to achieve two-message (statistically) witness indistinguishable protocols [BGI$^+$17, KKS18] and weak zero-knowledge protocols [JKKR17, BKP19], multi-party computation protocols with minimal round complexity [AJ17, BGJ$^+$17, BGJ$^+$18], improved round complexity for non-malleable commitments [KS17, Khu17], malicious circuit privacy for fully-homomorphic encryption [OPP14], and correctness amplification for indistinguishability obfuscators [BV16].

Up until recently, SSP OT protocols were only known based on number-theoretic assumptions such as DDH [AIR01, NP01] and QR and DCR [HK12], which are not resilient to quantum attacks. Brakerski and Dötling then constructed SSP OT based on LWE, which is the first construction to be plausibly post-quantum secure [BD18]. Their construction strongly relies on the tight connection between LWE and lattices and in particular, the transference principle [Ban93]. Aiming to construct SSP OT from LPN, we are once again faced with the fact that LPN lacks the geometrical structure of LWE; in particular, an analogous transference principle is not known.

1.1 Our Results

We construct SSP OT assuming $\mathrm{LPN}_{\frac{\log^2 n}{n}}$ and a standard Nisan-Wigderson style derandomization assumption, namely that there exists functions with (uniform) time complexity $2^{O(n)}$ and non-deterministic circuit complexity $2^{\Omega(n)}$. Toward this, we first construct SSP OT in the common random string model, which is already meaningful on its own, and where most of the technical difficulty lies. We then show how to derandomize the common random string.

In more detail, we prove the following three results:

1. Assuming $\mathrm{LPN}_{\frac{\log^2 n}{n}}$, there exists SSP OT in the *common random string model*.

2. Any SSP OT in the common random string model, can be converted to one in a relaxed model, where the receiver need not trust the common string. We refer to this as the *sender random string model,* as the sender can generate the common string.

3. Under the aforementioned derandomization assumption, any SSP OT in the sender random string model can be transformed into one in the plain model, provided that it has a certain *bad-crs certification* property. We prove that the construction from the first result (in the common random string model) satisfies this property, and that it is preserved by the transformation given by the second result.

On Low-Noise LPN. Our construction relies on LPN in the so called low-noise regime, where we could expect at most quasi-polynomial hardness [BKW03]. This indeed makes it mostly of theoretical interest. Improving the noise rate is an intriguing problem that may very well require a significant leap in our understanding of the complexity of LPN. Indeed, a folklore fact is that SSP OT (even in the common random string model) implies lossy public-key encryption, and thus a construction of SSP OT from LPN_ε would imply that $\mathrm{LPN}_\varepsilon \in \mathrm{BPP}^{\mathrm{SZK}}$. However, so far it is only known that $\mathrm{LPN}_\varepsilon \in \mathrm{BPP}^{\mathrm{SZK}}$ for $\varepsilon = O(\log^2 n/n)$ [BLVW19], and there is no indication that this is also true for larger ε. We also note that there are in fact much more basic primitives than SSP OT, such as collision resistant hashing (which is not even broken in SZK), that to date are only known in the low noise regime.

On the Derandomization Assumption. Starting from the work of Barak, Ong, and Vadhan [BOV07], the use of Nisan-Wigderson syle derandomization has become quite commonly used in cryptographic applications (c.f. [BV17, HNY17, BDK+19]). The corresponding assumption is a worst-case assumption that is considered to be a natural generalization of the assumption that $\mathbf{EXP} \not\subset \mathbf{NP}$. We also note that there is a universal candidate for the assumption, by instantiating the hard function with any $\mathbf{E}$-complete language under linear reductions. In the body, we actually use an even weaker uniform variant of the derandomization assumption. (See further discussion in [BOV07].)

1.2 Technical Overview

We now provide a technical overview of our constructions and proofs. Most of the overview is dedicated to our protocol in the common random string model (CRS), where most of the technical challenge lies. We then explain the second step in which the CRS is derandomized.

A Basic Protocol. We start by describing the basic protocol in the CRS model.

- The CRS $(\mathbf{A}, \mathbf{v})$ will consist of a random matrix $\mathbf{A} \leftarrow \mathbb{F}_2^{\ell \times n}$ and a random vector $\mathbf{v} \leftarrow \mathbb{F}_2^{\ell}$, for a parameter $\ell = \mathrm{poly}(n)$.
- The receiver, with choice $c \in \{0,1\}$, samples a secret $\mathbf{s} \leftarrow \mathbb{F}_2^n$ and a noise vector $\mathbf{e} \leftarrow \mathrm{Bern}(\varepsilon)^\ell$, and sends $\mathbf{v}_0 := \mathbf{A}\mathbf{s} + \mathbf{e} + c\mathbf{v}$.
- The sender, with messages $m_0, m_1 \in \{0,1\}$, samples a vector $\mathbf{x} \leftarrow \mathcal{X}$ from some low hamming-weight distribution $\mathcal{X}$ on $\mathbb{F}_2^\ell$, and sends back $(\mathbf{x}^t\mathbf{A}, \mathbf{x}^t\mathbf{v}_0 + m_0, \mathbf{x}^t\mathbf{v}_1 + m_1)$, where $\mathbf{v}_1 = \mathbf{v}_0 + \mathbf{v}$.
- The receiver, now uses $\mathbf{s}$ to compute $\mathbf{x}^t\mathbf{v}_c + m_c - \mathbf{x}^t\mathbf{A}\mathbf{s} = m_c + \mathbf{x}^t\mathbf{e}$.

Correctness and SSP Against Semi-Honest Receivers. In the above basic protocol, the computational privacy of the sender's choice follows directly from LPN_ε. The essential tradeoff is between correctness and statistical sender privacy (SSP). On one hand, to ensure correctness we aim that $\mathbf{x}^t\mathbf{e} = 0$ with high enough probability, and thus want $\mathbf{x}$ to be as sparse as possible. On the other hand, given that $\mathbf{x}^t\mathbf{A}$ already leaks n bits of information about $\mathbf{x}$, it should have min-entropy greater than n, and thus cannot be too sparse.

To understand how to balance this tradeoff, let us first restrict attention to a simple case of semi-honest receivers that follow the protocol as prescribed (here in fact the receiver may also send the CRS). A simple intuition for SSP in this setting is the fact that for the negative choice bit $1-c$, the receiver obtains

$$\mathbf{x}^t\mathbf{v}_{1-c} + m_{1-c} = \mathbf{x}^t(\mathbf{A}\mathbf{s}+\mathbf{e}) + \mathbf{x}^t\mathbf{v} + m_{1-c} \ .$$

Here the inner product $\mathbf{x}^t\mathbf{v}$ acts as a strong randomness extractor, so as long as the min-entropy remaining in $\mathbf{x}$ is large enough $\mathbf{H}_\infty(\mathbf{x} \mid \mathbf{x}^t\mathbf{A}) \gg n$, the message m_{1-c} will remain statistically hidden. A back of the envelop calculation shows that this already restricts our possible choice of parameters. It requires that we choose $\varepsilon = O(\log^2 n/n)$, which lets us choose $\mathcal{X}$ so to guarantee that $\mathbf{x}^t\mathbf{e} = 0$ with probability at least $\frac{1}{2} + \frac{1}{\mathrm{poly}(n)}$ (the exact choice of $\mathcal{X}$ does not matter at this point, e.g. it can be $\mathrm{Bern}(\delta)^\ell$ for an appropriate δ). Once we have this correctness guarantee, we can amplify it, using standard parallel repetition.

Malicious Receivers. The main technical challenge and the bulk of our work is proving that the above protocol is in fact also SSP against malicious receivers (for an appropriate choice of distribution $\mathcal{X}$). The challenge lies in the fact that a malicious receiver may now choose $\mathbf{v}_0$ arbitrarily and adaptively depending on the seed $\mathbf{v}$ and the matrix $\mathbf{A}$. Still, we need to ensure that any $\mathbf{v}_0 \in \mathbb{F}_2^\ell$, now chosen as a function of $\mathbf{v}, \mathbf{A}$, fixes some $c^* \in \{0,1\}$ such that $\mathbf{x}^t\mathbf{v}_{1-c^*}$ is statistically close to uniform for $\mathbf{x} \leftarrow \mathcal{X}$, even given $\mathbf{x}^t\mathbf{A}$. (To be more precise, as described so far, the sender's message includes an extra bit of leakage on $\mathbf{x}$, since it includes both $\mathbf{x}^t\mathbf{v}_0 + m_0$ and $\mathbf{x}^t\mathbf{v}_1 + m_1$. In the actual scheme, we use two independent samples $\mathbf{x}_0$ and $\mathbf{x}_1$ for these two parts, so this is not an issue.)

One could hope that the inner product extractor is generally resilient to such *linear splitting attacks.* That is, given the seed $\mathbf{v}$ the attacker may split it adaptively to $\mathbf{v}_0, \mathbf{v}_1$ that sum to $\mathbf{v}$, we can still hope that one of the seeds still functions as a good extractor. However, it turns out that this is not generally true. As an example, consider the distribution $\mathcal{X}'$, where $\mathbf{x}$ is sampled by first choosing a uniformly random $\mathbf{x} \leftarrow \mathbb{F}_2^\ell$, then flipping a random bit $b \leftarrow \{0,1\}$, and then zeroing out the first or second half of $\mathbf{x}$ according to b. Then an attacker, given a seed $\mathbf{v} = (v_1, \ldots, v_\ell)$ could split $\mathbf{v}$ into its two halves $\mathbf{v}_0 = (v_1, \ldots, v_{\frac{\ell}{2}}, 0, \ldots, 0)$ and $\mathbf{v}_1 = (0, \ldots, 0, v_{\frac{\ell}{2}+1}, \ldots, v_\ell)$. Then neither $\mathbf{v}_0$ nor $\mathbf{v}_1$ is a good extractor: if we leak b, then although $\mathbf{x}|b$ has high entropy, either bit will be predictable with probability $3/4$. Indeed, this counter example strongly relies on the fact that $\mathbf{v}_0$ is chosen adaptively depending on $\mathbf{v}$.

Back to Our Case. While we cannot simply rely on the inner product being a strong extractor, in our case the leakage on $\mathbf{x}$ has a specific form $\mathbf{x}^t\mathbf{A}$, and we also have the liberty of choosing the distribution $\mathcal{X}$ (provided that the previous correctness guarantees still hold). Indeed, we manage to prove that for an appropriate choice of $\mathcal{X}$ SSP does hold. We now proceed to describe our choice of distribution $\mathcal{X}$, and the main steps in the proof, which is quite intricate.

Inspired by the LPN smoothing reduction of Brakerski et al. [BLVW19], we choose a distribution $\mathcal{X}$ that behaves somewhat nicely in terms of Fourier analysis. Specifically, we use the *sampling with replacement* distribution

$$\mathcal{X}_{\ell,k} := \sum_{i=1}^{k} U\{\mathbf{e}_1, \ldots, \mathbf{e}_l\} \ ,$$

which is the sum of k independent random unit vectors over $\mathbb{F}_2^\ell$. By choosing the hamming weight $k \approx n/\log n$, we guarantee that $\mathbf{x}^t\mathbf{e} = 0$ with probability noticeably greater than half as required. Considering the attacker's choice of $\mathbf{v}_0$ and the corresponding $\mathbf{v}_1 = \mathbf{v}_0 + \mathbf{v}$, we aim to show that for some $\mathbf{w} \in \{\mathbf{v}_0, \mathbf{v}_1\}$ we are guaranteed that $\mathbf{x}^t\mathbf{w}$ is close to uniform, even given $\mathbf{x}^t\mathbf{A}$.

Relating Statistical Unpredictability to Coset Balance. We first observe that if $\mathbf{w}$ is *too close* to the code generated by $\mathbf{A}$, namely $\mathbf{w} = \mathbf{A}\mathbf{s}^* + \mathbf{e}^*$ for some low hamming weight $\mathbf{e}^*$, then $\mathbf{x}^t\mathbf{w} = \mathbf{x}^t\mathbf{A}\mathbf{s}^* + \mathbf{x}^t\mathbf{e}^*$ becomes predictable. Indeed, $\mathbf{x}^t\mathbf{A}\mathbf{s}^*$ is determined by $\mathbf{x}^t\mathbf{A}$ and $\mathbf{x}^t\mathbf{e}^*$ is likely to be zero (this is exactly what enables correctness). This is in fact also the case if $\mathbf{w}$ is *too far* from some codeword, namely $\mathbf{w} = \mathbf{A}\mathbf{s}^* + \mathbf{1} + \mathbf{e}^*$,

where $\mathbf{1}$ is the all one vector. Predicting $\mathbf{x}^t\mathbf{w}$ is similar to the previous case, except that we need to also predict $\mathbf{x}^t\mathbf{1}$, but this will be exactly $k \mod 2$.

In conclusion, to guarantee statistical unpredictability, it is necessary that all the vectors $\mathbf{w} + \mathbf{As}$ in the coset $\mathbf{w} + \mathbf{A} = \{\mathbf{w} + \mathbf{As} : \mathbf{s} \in \mathbb{F}_2^n\}$ will be rather *balanced*, namely they should have hamming weight $\|\mathbf{w} + \mathbf{As}\|_0 \approx \ell/2$. Using Fourier analysis, we show that to some extent this is also sufficient. That is, we characterize the unpredictability of $\mathbf{x}^t\mathbf{w}$ in terms of the *balance parameter* $\beta_{\mathbf{t}} := 1 - (2/\ell)\|\mathbf{t}\|_0$ of any coset member $\mathbf{t} \in \mathbf{w} + \mathbf{A}$. Specifically, we prove:

$$SD\left(\left(\mathbf{x}^t\mathbf{A}, \mathbf{x}^t\mathbf{w}\right), \left(\mathbf{x}^t\mathbf{A}, u\right)\right) \leq \frac{1}{2}\sum_{\mathbf{s}\in\mathbb{F}_2^n} \left|\beta_{\mathbf{As}+\mathbf{w}}\right|^k \ ,$$

for any matrix $\mathbf{A} \in \mathbb{F}_2^{\ell\times n}$, vector $\mathbf{w} \in \mathbb{F}_2^\ell$, $\mathbf{x} \leftarrow \mathcal{X}_{\ell,k}$, and $u \leftarrow \text{Bern}(1/2)$. The proof can be found in Section 3.2.1.

Our goal is thus to show that for at least one $\mathbf{w} \in \{\mathbf{v}_0, \mathbf{v}_1 = \mathbf{v}_0 + \mathbf{v}\}$, the total balance $\sum_{\mathbf{s}\in\mathbb{F}_2^n} \left|\beta_{\mathbf{As}+\mathbf{w}}\right|^k$ is negligible. To show this, we prove that the following two coset balance properties hold with overwhelming probability over the choice of the CRS $(\mathbf{v}, \mathbf{A})$:

1. **Property 1: v is A-balanced for sums.** This property means that for any decomposition $\mathbf{v}_0 + \mathbf{v}_1 = \mathbf{v}$, for at least one $\mathbf{w} \in \{\mathbf{v}_0, \mathbf{v}_1\}$, the coset $\mathbf{w} + \mathbf{A}$ is *somewhat balanced*. Specifically, for every $\mathbf{s}$, $|\beta_{\mathbf{As}+\mathbf{w}}| \leq 3/5$.
2. **Property 2: A is affinely balanced.** This property means that in any coset $\mathbf{w}+\mathbf{A}$ most members are *well balanced*. Specifically, except for a set E of at most $2^{o(k)}$ vectors $\mathbf{s}$, it holds that $|\beta_{\mathbf{As}+\mathbf{w}}| \leq 2^{-\omega(n/k)}$.

Combining these two properties, we can guarantee that for the $\mathbf{w} \in \{\mathbf{v}_0, \mathbf{v}_1\}$ such that Property 1 holds:

$$\sum_{\mathbf{s}} \left|\beta_{\mathbf{As}+\mathbf{w}}\right|^k \leq \sum_{\mathbf{s}\in E} (3/5)^k + \sum_{\mathbf{s}\notin E} 2^{-\omega(n)} \ ,$$

which is negligible for our choice of $k \approx n/\log n$. We refer the reader to Section 3 for more details regarding the proof, and in particular the proof that the above two properties hold.

From the CRS Model to the Sender Random String Model. We now explain how to compile an SSP OT protocol (S, R) in the CRS model to a protocol $(\mathsf{S}', \mathsf{R}')$ in the sender random string model (SRS), where receiver privacy is guaranteed even if the common string is chosen by a malicious sender. The transformation is based on the idea of reverse randomization from Dwork and Naor's *NIZK to ZAP transformation* [DN00].

In the new protocol, the sender random string consists of many random strings $\mathsf{scrs}_1, \ldots, \mathsf{scrs}_k$. The receiver R', given the sender random string, will sample a single random string rcrs of its own, and will generate k corresponding common strings $\mathsf{crs}_i = \mathsf{scrs}_i \oplus \mathsf{rcrs}$ for the underlying protocol (S, R). It will then run the underlying S in k parallel copies using crs_i and his choice bit c. The sender S' will secret share each of its two messages m_0 and m_1 into $m_0^1, \ldots, m_0^k$ and $m_1^1, \ldots, m_1^k$, and respond in each copy i by running the underlying S with messages m_0^i, m_1^i.

The computational receiver privacy is shown via a standard hybrid argument. For SSP, k is chosen to be large enough to guarantee that for any receiver choice of rcrs, at least one crs_i will ensure SSP, this is sufficient due to the use of secret sharing.

From the SRS Model to the Plain Model Using Derandomization. We now explain how to derandomize the SRS to get a protocol in the plain model. Here we again draw inspiration from the case of ZAPs. Barak, Ong, and Vadhan [BOV07] observe that in ZAPs a *bad CRS*, namely one relative to which there exist false proofs, can be identified non-deterministically in *fixed* polynomial time (for a given false statement, the certificate for badness is an accepting ZAP). This allows them to derandomize the CRS using *hitting set generators* (HSG) against co-nondeterministic circuits, which in turn can be constructed from the aforementioned worst-case assumption [GST03]. Such a generator G deterministically computes in polynomial time

a set $S = \{\mathsf{crs}_i\}$ of strings. G guarantees that if a random string crs is *not* bad with high probability, then the set S will include *at least one* string crs_i that is not bad. This is sufficient for derandomizing ZAPs, by running parallel ZAP instances with each crs_i.

In our setting a bad SRS is one for which SSP does not hold. If such badness is certifiable then we can rely on a similar transformation. As in ZAPs, we can run the SSP OT protocol with each crs_i in the generated set S, and like the transformation from the previous paragraph, use secret sharing on the sender's end to guarantee SSP. However, unlike the case of ZAPs, for a general SSP OT the badness of a given crs might not be certifiable. Hence we need to require this explicitly from the underlying SSP OT. This means that we have to guarantee that our SSP OT has this additional *bad CRS certification* property.

Guaranteeing Bad CRS Certification. This boils down to showing that our protocol in the CRS model has the property; indeed, it is not hard to show that the transformation to the SRS model would preserve bad CRS certification. Recall that in our construction, we proved that if a CRS $(\mathbf{A}, \mathbf{v})$ possesses Properties 1 and 2, then it is not bad. It is not hard to see that if Property 1 — $\mathbf{v}$ is $\mathbf{A}$-balanced for sums — is not satisfied then this can be certified. The witness is a decomposition $\mathbf{v}_0, \mathbf{v}_1$ such that $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_1$ along with $\mathbf{s}_0, \mathbf{s}_1$, such that both $\mathbf{A}\mathbf{s}_i + \mathbf{v}_i$ are not somewhat balanced.

In contrast, it is not clear how to certify Property 2 — $\mathbf{A}$ is affinely balanced. For this purpose we identify an alternative algebraic property that is both certifiable and implies affine balance; surprisingly we call it *strong affine balance*. The property states that for any $\mathbf{w}$, in any set of $d \approx n/\log^2 n$ linearly independent vectors $\mathbf{s}_1, \ldots, \mathbf{s}_d \in \mathbb{F}_2^n$ at least one coset member $\mathbf{A}\mathbf{s}_i + \mathbf{w}$ is well balanced. If this property does not hold, then this can be certified; the witness is $\mathbf{w}, \mathbf{s}_1, \ldots, \mathbf{s}_d$ that do not satisfy the property. Furthermore, we show that strong affine balance implies affine balance. We refer the reader to Section 3 for more details.

1.3 More Related Work

Other Applications of LPN. The works of Brakerski et al. and Yu et al. build collision-resistant hash function based on $LPN_{\log^2 n/n}$ [BLVW19, YZW+17] (the latter also shows certain tradeoffs between hardness and shrinkage) . [BLSV18] construct anonymous identity-based encryption assuming the hardness of $LPN_{\log^2 n/n}$. Brakerski, Mour, and Koppula [BKM20] construct non-interactive zero-knowledge arguments based on $\mathrm{LPN}_{n^{-(1/2+\epsilon)}}$ and the existence of trapdoor-hash-functions (which can be constructed from DDH). Bartusek et al. construct maliciously-secure, two-round reusable multiparty computation in the CRS model based on $\mathrm{LPN}_{1/n^{1-\varepsilon}}$ [BGSZ21].

The Hardness of LPN. The gap between LWE and LPN is also expressed in hardness results. While the hardness of LWE can be based on the worst-case hardness of long-studied lattice problems (c.f. [Reg05, Pei09, BLP+13]), worst-case to average case reductions for LPN have only been recently discovered and are still very limited (they essentially show that solving the relatively "easy case" of $\mathrm{LPN}_{\frac{\log^2 n}{n}}$ in the worst case can be reduced to solving a very "hard case" of $\mathrm{LPN}_{\frac{1}{2}-\frac{1}{\mathrm{poly}(n)}}$ in the average case) [BLVW19, YZ21].

2 Preliminaries

We rely on the following standard notation.

- Throughout, we identify $\{0,1\}^\ell$ with $\mathbb{F}_2^\ell$ in the natural way, addition and multiplication of elements in $\mathbb{F}_2$ refers to the corresponding field operations.
- We denote vectors and matrices in bold, whereas scalars are not bold.
- For a binary vector $\mathbf{x}$, we denote by $\|\mathbf{x}\|_0$ the hamming weight of $\mathbf{x}$.
- We denote by $\mathbb{E}[X]$ the expected value of random variable X.

- For a distribution D, $x \leftarrow D$ denotes sampling x from D. For a set S, $x \leftarrow S$ denotes uniformly sampling from S.

We rely on the standard notions of Turing machines and Boolean circuits.

- We say that a Turing machine is PPT if it is probabilistic and runs in polynomial time.
- For a PPT algorithm M, we denote by $M(x;r)$ the output of M on input $\mathbf{x}$ and random coins r. For such an algorithm, and any input x, we may write $m \in M(x)$ to denote the fact that m is in the support of $M(x;\cdot)$.
- A polynomial-size circuit family $\mathcal{C}$ is a sequence of circuits $\mathcal{C} = \{C_n\}_{n\in\mathbb{N}}$, such that each circuit C_n is of polynomial size $n^{O(1)}$ and has $n^{O(1)}$ input and output bits. We also consider probabilistic circuits that may toss random coins.
- We follow the standard convention of modeling any efficient adversary as a family of polynomial-size circuits. For an adversary A corresponding to a family of polynomial-size circuits $\{\mathsf{A}_n\}_{n\in\mathbb{N}}$, we sometimes omit the subscript n, when it is clear from the context.
- A function $f : \mathbb{N} \to [0,1]$ is negligible if $f(n) = n^{-\omega(1)}$ and is noticeable if $f(n) = n^{-O(1)}$.
- Two ensembles of random variables $\mathcal{X} = \{X_i\}_{n\in\mathbb{N},i\in I_n}$, $\mathcal{Y} = \{Y_i\}_{n\in\mathbb{N},i\in I_n}$ over the same set of indices $I = \cup_{n\in\mathbb{N}} I_n$ are said to be *computationally indistinguishable* (respectively, *statistically indistinguishable*), denoted by $\mathcal{X} \approx_c \mathcal{Y}$, if for every polynomial-size (respectively, unbounded) distinguisher $\mathsf{A} = \{\mathsf{A}_n\}_{n\in\mathbb{N}}$ there exists a negligible function μ such that for all $n \in \mathbb{N}, i \in I_n$,

$$\left|\Pr\left[\mathsf{A}(X_i) = 1\right] - \Pr\left[\mathsf{A}(Y_i) = 1\right]\right| \leq \mu(n) \ .$$

Definition 2.1 (Distribution $\mathcal{X}_{\ell,k}$: sampling with replacement)**.** *Let $\ell, k \in \mathbb{N}$. We denote by $\mathcal{X}_{\ell,k}$ the distribution over $\mathbb{F}_2^\ell$, where $\mathbf{x} \leftarrow \mathcal{X}_{\ell,k}$ is the sum of k uniformly random standard basis vectors, sampled independently with repetitions:*

$$\mathbf{x} := \sum_{i=1}^{k} \mathbf{x}_i, \ \textit{where } \forall i \in [k], \ \mathbf{x}_i \leftarrow \{\mathbf{e}_1, \ldots, \mathbf{e}_\ell\} \ ,$$

where $\mathbf{e}_j$ is the j-th standard basis vector.

We rely on the following basic lemmas.

Lemma 2.2 (Piling-Up Lemma [Mat93])**.** *Let $v_1, .., v_k \in \mathbb{F}_2$ i.i.d random variables such that $\mathbb{E}[v_i] = \varepsilon$, then:*

$$\Pr\left[\sum_{i=1}^{k} v_i = 1\right] = \frac{1}{2} - \frac{1}{2}(1-2\varepsilon)^k \ .$$

Lemma 2.3 (Random Vectors Are Balanced)**.** *Let $\ell \in \mathbb{N}$ and $\beta \geq \sqrt{1/\ell}$, then:*

$$\Pr_{\mathbf{w}\leftarrow\mathbb{F}_2^\ell}\left[\left|\|\mathbf{w}\|_0 - \frac{\ell}{2}\right| \geq \beta\ell\right] \leq 2^{-\beta^2\ell} \ .$$

The latter lemma follows directly from a Chernoff-Hoeffding bound.

2.1 Learning Parity with Noise

We recall the Learning Parity with Noise (LPN) assumption.

Definition 2.4 (LPN Assumption)**.** *For noise rate $\varepsilon(n) \in [0, 1/2]$, the LPN_ε assumption is that for any $m(n) = n^{O(1)}$,*

$$\{\mathbf{A}, \mathbf{As} + \mathbf{e}\}_{n\in\mathbb{N}} \approx_c \{\mathbf{A}, \mathbf{u}\}_{n\in\mathbb{N}} \ ,$$

where $\mathbf{A} \leftarrow \mathbb{F}_2^{m\times n}, \mathbf{s} \leftarrow \mathbb{F}_2^n, \mathbf{e} \leftarrow \mathrm{Bern}(\varepsilon)^m$, and $\mathbf{u} \leftarrow \mathbb{F}_2^m$.

2.2 Derandomization: Hitting Set Generators

We next define hitting set generators (HSGs) and state relevant results from the literature. We address both HSGs against non-uniform circuits as well against uniform algorithms. The non-uniform version is somewhat more common in the literature and simpler to state. However, the (weaker) uniform version will suffice for our purpose.

Definition 2.5 (Co-nondeterministic Circuits and Algorithms)**.** *A co-nondeterministic boolean circuit $C(x, w)$ (respectively, uniform algorithm $A(x, w)$) takes x as a primary input and w as a witness. We define $C(x) := 0$ (respectively, $A(x) := 0$) if and only if there exists w such that $C(x, w) = 0$ (respectively, $A(x, w) = 0$).*

Definition 2.6 (Hitting Set Generators)**.** *A deterministic polynomial-time algorithm $H(1^m, 1^s)$ that outputs a set of strings of length m, is a hitting set generator against co-nondeterministic circuits, if for every $m, s \in \mathbb{N}$, and every co-non-deterministic circuit $C : \{0, 1\}^m \to \{0, 1\}$ of size at most s:*

$$\Pr_{x\leftarrow\{0,1\}^m} [C(x) = 1] > 1/2 \Longrightarrow \exists y \in H(1^m, 1^s) : C(y) = 1 \ .$$

Definition 2.7 (Uniform Hitting Set Generators)**.** *A deterministic polynomial-time algorithm $H(1^m, 1^{s(m)})$ that outputs a set of strings of length m, is a hitting set generator against co-non-deterministic **uniform algorithms**, if for every co-nondeterministic uniform algorithm $A : \{0, 1\}^* \to \{0, 1\}$ of running time at most $s(m)$, and for sufficiently large m:*

$$\Pr_{x\leftarrow\{0,1\}^m} [A(x) = 1] > 1/2 \Longrightarrow \exists y \in H(1^m, 1^{s(m)}) : A(y) = 1 \ .$$

In the literature, a more general notion of ε-HSGs is often defined, where the bound $1/2$ is replaced by ε. In terms of computational assumptions, this difference is inconsequential due to general amplification results for HSGs [GST03].

Theorem 2.8 ([MV99])**.** *Assume there exists a function f in $\mathbf{E} = \boldsymbol{Dtime}(2^{O(n)})$ with non-deterministic circuit complexity $2^{\Omega(n)}$. Then, there exists an efficient HSG against co-nondeterministic circuits.*

Gutfreund, Shaltiel and Ta-Shma [GST03] show that HSGs against co-nondeterministic *uniform* algorithm can be obtained from a relaxed (uniform) hardness assumption.

Definition 2.9 (AM)**.** *A probabilistic nondeterministic algorithm $A(x, r, y)$ takes in addition to its regular input x a randomness input r as well a nondeterministic input y. We say that A computes a function $f : \{0, 1\}^* \to \{0, 1\}$ if for any x:*

- $f(x) = 1 \Longrightarrow \Pr_r [\exists y : A(x, r, y) = 1] = 1$,
- $f(x) = 0 \Longrightarrow \Pr_r [\exists y : A(x, r, y) = 1] \leq \frac{1}{2}$.

***AM** is the class of all languages decidable by probabilistic nondeterministic algorithms running in time $\mathrm{poly}(n)$ where $n = |x|$. Similarly, **AMTIME**$(t(n))$ is the class of all languages decidable by probabilistic nondeterministic algorithms running in time at most $t(n)$. Finally, $[i.o. - \boldsymbol{AMTIME}](t(n))$ denotes the class of all languages which have a probabilistic nondeterministic $t(n)$-running-time algorithm deciding them for infinitely-many input lengths.*

Theorem 2.10 ([GST03]). *Assume* $\boldsymbol{E} \not\subseteq [i.o. - \boldsymbol{AMTIME}](2^{\delta n})$ *for some* $\delta > 0$. *Then, there exists an efficient HSG against co-nondeterministic uniform algorithms.*

Note that the uniform assumption (as in Theorem 2.10) is indeed a relaxation of the non-uniform one (as in Theorem 2.8), since non-uniformity can simulate randomness. We also note that both assumptions are worst-case assumptions, and that similar (or stronger) assumptions have by now become quite common in cryptographic applications (c.f. [BOV07, BV17, HNY17]).

2.3 Statistical Sender-Private Oblivious Transfer

Oblivious Transfer (OT) is a protocol between two parties: a sender S and a receiver R. The sender input consists of two secret messages m_0, m_1, and the receiver input is a secret choice bit c. The protocol allows the receiver to learn m_c, and guarantees that the receiver gains no information regarding m_{1-c}, whereas the sender gains no information regarding the receiver choice bit c. We focus on statistical sender privacy (SSP); namely, sender privacy holds even against unbounded malicious receivers. Receiver privacy is computational. Furthermore, we restrict attention to protocols with two messages (one from each party).

We consider three models of trusted setup:

- **The common random string model:** Here a common random string crs is generated once and for all. The string is trusted by both the receiver and sender.
- **The sender random string model:** This model is similar to the common random string model, except that the receiver need not trust the string crs; namely, receiver privacy holds for any choice of crs (even if adversarially made by the sender).
- **The plain model:** Here there is no trusted setup at all. Equivalently, the setup procedure generating crs is deterministic.

We next define the notion in the common random string model, and then extend it to the sender random string model and the plain model. In all definitions, n indicates the security parameter.

Definition 2.11 (Two-message Statistically Sender-Private OT in CRS model). *A two-message Statistically sender-private OT in the common-random-string model consists of PPT algorithms* $\mathsf{R} = (\mathsf{R.Enc}, \mathsf{R.Dec})$ *and* S*, and an associated polynomial* ρ*, with the following syntax:*

1. $\mathsf{R.Enc}(\mathsf{crs}, c)$*: Gets* $\mathsf{crs} \in \{0,1\}^{\rho(n)}$ *and choice bit* $c \in \mathbb{F}_2$ *and outputs a message* rm *and secret key* sk.
2. $\mathsf{S}(\mathsf{crs}, m_0, m_1, \mathsf{rm})$*: Gets* $\mathsf{crs} \in \{0,1\}^{\rho(n)}$*, two bits* $m_0, m_1 \in \mathbb{F}_2$*, and* rm*, and outputs a message* sm.
3. $\mathsf{R.Dec}(\mathsf{crs}, \mathsf{sk}, \mathsf{sm})$*: Gets* $\mathsf{crs} \in \{0,1\}^{\rho(n)}$*, secret key* sk*, and message* sm*, and outputs a message bit.*

We require the following:

- ***Correctness:*** *For every* c, m_0, m_1,

$$\Pr\left[\mathsf{R.Dec}(\mathsf{crs}, \mathsf{sk}, \mathsf{sm}) = m_c \;\middle|\; \begin{array}{l} \mathsf{crs} \leftarrow \{0,1\}^{\rho(n)} \\ (\mathsf{rm}, \mathsf{sk}) \leftarrow \mathsf{R.Enc}(\mathsf{crs}, c) \\ \mathsf{sm} \leftarrow \mathsf{S}(\mathsf{crs}, m_0, m_1, \mathsf{rm}) \end{array}\right] \geq 1 - n^{-\omega(1)} \ .$$

- ***Receiver Privacy:***

$$\left\{\mathsf{crs}, \mathsf{rm} \;\middle|\; \begin{array}{l} \mathsf{crs} \leftarrow \{0,1\}^{\rho(n)} \\ (\mathsf{rm}, \mathsf{sk}) \leftarrow \mathsf{R.Enc}(\mathsf{crs}, 0) \end{array}\right\}_{n\in\mathbb{N}} \approx_c \left\{\mathsf{crs}, \mathsf{rm} \;\middle|\; \begin{array}{l} \mathsf{crs} \leftarrow \{0,1\}^{\rho(n)} \\ (\mathsf{rm}, \mathsf{sk}) \leftarrow \mathsf{R.Enc}(\mathsf{crs}, 1) \end{array}\right\}_{n\in\mathbb{N}} \ .$$

- ***Statistical Sender Privacy:*** *There exists an (unbounded)* OTExt*, such that for any (unbounded)* R^**:*

$$\left\{ \mathsf{crs}, \mathsf{sm} \;\middle|\; \begin{array}{l} \mathsf{crs} \leftarrow \{0,1\}^{\rho(n)} \\ \mathsf{rm} \leftarrow \mathsf{R}^*(\mathsf{crs}) \\ \mathsf{sm} \leftarrow \mathsf{S}(\mathsf{crs}, m_0, m_1, \mathsf{rm}) \end{array} \right\}_{n,m_0,m_1} \approx_s \left\{ \mathsf{crs}, \mathsf{sm} \;\middle|\; \begin{array}{l} \mathsf{crs} \leftarrow \{0,1\}^{\rho(n)} \\ \mathsf{rm} \leftarrow \mathsf{R}^*(\mathsf{crs}) \\ b \leftarrow \mathsf{OTExt}(\mathsf{crs}, \mathsf{rm}) \\ \mathsf{sm} \leftarrow \mathsf{S}(\mathsf{crs}, m_b, m_b, \mathsf{rm}) \end{array} \right\}_{n,m_0,m_1},$$

  where $n \in \mathbb{N}, m_0, m_1 \in \{0,1\}$*.*

We now derive the definitions in the sender-random-string model and in the plain model.

Definition 2.12 (Two-message Statistically Sender-Private OT in SRS model)**.** *A two-message Statistically sender-private OT in the sender-random-string model is defined similarly to Definition 2.11, except that receiver privacy holds for any choice of* crs*:*

$$\left\{ \mathsf{rm} \;\middle|\; (\mathsf{rm}, \mathsf{sk}) \leftarrow \mathsf{R.Enc}(\mathsf{crs}, 0) \right\}_{n,\mathsf{crs}} \approx_c \left\{ \mathsf{rm} \;\middle|\; (\mathsf{rm}, \mathsf{sk}) \leftarrow \mathsf{R.Enc}(\mathsf{crs}, 1) \right\}_{n,\mathsf{crs}},$$

where $n \in \mathbb{N}, \mathsf{crs} \in \{0,1\}^{\rho(n)}$*.*

Definition 2.13 (Two-message Statistically Sender-Private OT in plain model)**.** *A two-message Statistically sender-private OT in the plain model is defined similarly to Definition 2.11, except that* crs *is ignored by all algorithms.*

2.3.1 Enhancements

We define two natural enhancements to the definition of SSP-OT protocols in the CRS/SRS model. Relying on these enhancements, we will show transformations between the three models (CRS, SRS, and plain). Furthermore, our core protocol (presented in Section 3) will satisfy these enhancements.

Bad CRS Certification. The first enhancement is for sender privacy, roughly saying that there is an NP witness for a CRS being "bad for sender privacy". The exact definition follows.

Definition 2.14 (Bad CRS Certification)**.** *A two-message SSP OT protocol in the CRS/SRS model has* bad CRS certification *if there exists a set* B *such that:*

- ***Statistical Sender Privacy Outside*** B***:*** *There exists an (unbounded)* OTExt*, such that for any (unbounded)* R^**:*

$$\left\{ \mathsf{sm} \;\middle|\; \begin{array}{l} \mathsf{rm} \leftarrow \mathsf{R}^*(\mathsf{crs}) \\ \mathsf{sm} \leftarrow \mathsf{S}(\mathsf{crs}, m_0, m_1, \mathsf{rm}) \end{array} \right\}_{n,\mathsf{crs},m_0,m_1} \approx_s \left\{ \mathsf{sm} \;\middle|\; \begin{array}{l} \mathsf{rm} \leftarrow \mathsf{R}^*(\mathsf{crs}) \\ b \leftarrow \mathsf{OTExt}(\mathsf{crs}, \mathsf{rm}) \\ \mathsf{sm} \leftarrow \mathsf{S}(\mathsf{crs}, m_b, m_b, \mathsf{rm}) \end{array} \right\}_{n,\mathsf{crs},m_0,m_1},$$

  where $n \in \mathbb{N}, \mathsf{crs} \in \{0,1\}^{\rho(n)} \setminus \mathsf{B}, m_0, m_1 \in \{0,1\}$*.*

- ***Negligible Density:*** $\Pr\left[\mathsf{crs} \in \mathsf{B} \;\middle|\; \mathsf{crs} \leftarrow \{0,1\}^{\rho(n)}\right] \leq n^{-\omega(1)}$*.*

- ***Certification:*** $\mathsf{B} \in \boldsymbol{NP}$*.*

Remark 2.1 (Relation to Plain SSP). We note that the plain SSP definition is in fact equivalent to the first two above conditions. That is, if SSP holds, then there exists a set B satisfying the first two conditions, and vice versa. The fact that the first two conditions imply SSP follows directly from the definition. The other direction follows by an averaging argument.

Specifically, given that plain SSP holds, consider the malicious receiver R^* that given crs, chooses the message rm that maximizes the statistical distance between $\mathsf{S}(\mathsf{crs}, m_0, m_1, \mathsf{rm})$ and $\mathsf{S}(\mathsf{crs}, m_b, m_b, \mathsf{rm})$ where b is the extracted bit. Letting ν be the statistical distance for a random crs, it holds that for all but a $\sqrt{\nu}$ fraction of crs, the maximal statistical distance between $\mathsf{S}(\mathsf{crs}, m_0, m_1, \mathsf{rm})$ and $\mathsf{S}(\mathsf{crs}, m_b, m_b, \mathsf{rm})$, over any choice of rm, is at most $\sqrt{\nu}$. The corresponding set B consists of this $\sqrt{\nu}$ fraction.

CRS-Free Correctness. The second enhancement is for the correctness property, saying that correctness holds for *any* choice of CRS.

Definition 2.15. *A two-message SSP OT protocol in the CRS/SRS model has* CRS-free correctness *if:*

$$\min_{\substack{\mathsf{crs}\in\{0,1\}^{\rho(n)} \\ m_0,m_1,c\in\{0,1\}}} \Pr\left[\mathsf{R.Dec}(\mathsf{crs},\mathsf{sk},\mathsf{sm}) = m_c \;\middle|\; \begin{array}{c}(\mathsf{rm},\mathsf{sk}) \leftarrow \mathsf{R.Enc}(\mathsf{crs},c) \\ \mathsf{sm} \leftarrow \mathsf{S}(\mathsf{crs},m_0,m_1,\mathsf{rm})\end{array}\right] \geq 1 - n^{-\omega(1)} \ ,$$

where the probability is over the coins of the sender S *and receiver* R*.*

In Section 4.1, we show that this property can always be obtained for free with no additional assumptions.

3 Two-Message SSP OT in the CRS Model

In this section, we present our two-message, statistically sender-private oblivious transfer in the common random string model. We prove the following theorem.

Theorem 3.1. *Under the* $LPN_{\frac{\log^2(n)}{n}}$ *assumption, there exists a two-message statically-sender-private OT protocol in the CRS model. Moreover, the protocol has CRS-free correctness and bad-CRS certification.*

We describe the protocol in Figure 1 and then proceed to analyze it. We describe the protocol in its interactive form. The receiver algorithms $\mathsf{R.Enc}$ and $\mathsf{R.Dec}$ correspond to the generation of the receiver message, and the decryption of the sender message, respectively.

Protocol 1

$\mathbf{A} \leftarrow \mathbb{F}_2^{\ell\times n}, \mathbf{v} \leftarrow \mathbb{F}_2^{\ell}$,
$\mathsf{crs} := (\mathbf{A}, \mathbf{v})$

$\mathsf{R}(c \in \mathbb{F}_2)$		$\mathsf{S}(m_0, m_1 \in \mathbb{F}_2)$
$\mathbf{s} \leftarrow \mathbb{F}_2^n$, $\mathbf{e} \leftarrow \mathrm{Bern}(\varepsilon)^\ell$,		
$\mathbf{v}_0 := \mathbf{A}\mathbf{s} + \mathbf{e} + c \cdot \mathbf{v}$.	$\xrightarrow{\mathsf{rm} := \mathbf{v}_0}$	
		$\mathbf{v}_1 := \mathbf{v}_0 + \mathbf{v}$.
		$\forall b \in \{0,1\}, i \in [r]$:
		$\mathbf{x}_{b,i} \leftarrow \mathcal{X}_{\ell,k}$,
		$\mathbf{y}_i^b := \mathbf{x}_{b,i}^t \mathbf{A}$,
		$z_i^b := \mathbf{x}_{b,i}^t \mathbf{v}_b + m_i$,
	$\mathsf{sm} := \begin{pmatrix} \mathbf{y}_1^0, z_1^0, ..., \mathbf{y}_r^0, z_r^0 \\ \mathbf{y}_1^1, z_1^1, ..., \mathbf{y}_r^1, z_r^1 \end{pmatrix} \xleftarrow{}$	
$\forall i \in [r]$ $w_i := \mathbf{y}_i^c \mathbf{s} + z_i^c$,		
Output $\mathrm{Maj}(w_1, ..., w_r)$.		

Figure 1: Two-message statistically-sender-private OT in the CRS model

Parameters: $n \in \mathbb{N}$ is the security parameter, $\delta > 1$ is a constant, $\ell = n^{1+\frac{1}{\delta}}$, $\varepsilon = \frac{\log^2(n)}{n}$, $k = 4\delta \cdot \frac{n}{\log(n)}$, and $r = n^{64\delta+1}$.

3.1 Correctness and Receiver Privacy

We first prove correctness.

Proposition 3.2. *The protocol is correct.*

Proof. In fact, we show CRS-free correctness. Fix any $\mathbf{A} \in \mathbb{F}_2^{\ell \times n}$, $\mathbf{v} \in \mathbb{F}_2^{\ell}$, $c, m_0, m_1 \in \mathbb{F}_2$. Recall that, R samples $\mathbf{s} \leftarrow \mathbb{F}_2^n$, $\mathbf{e} \leftarrow \mathrm{Bern}(\varepsilon)^\ell$, sends $\mathbf{v}_0 := \mathbf{A}\mathbf{s} + \mathbf{e} + c \cdot \mathbf{v}$, and then S sets $\mathbf{v}_1 := \mathbf{v}_0 + \mathbf{v}$. It follows that $\mathbf{v}_c = \mathbf{A}\mathbf{s} + \mathbf{e}$ and for every $i \in [r]$,

$$\begin{aligned} z_i^c &= \mathbf{x}_{c,i}^t \mathbf{A}\mathbf{s} + \mathbf{x}_{c,i}^t \mathbf{e} + m_c \\ w_i &= \mathbf{x}_{c,i}^t \mathbf{e} + m_c \qquad\qquad \text{where } \mathbf{x}_{c,i} \leftarrow \mathcal{X}_{\ell,k}. \end{aligned}$$

Thus, it suffices to show that the majority of $(\mathbf{x}_{c,1}^t \mathbf{e}, ..., \mathbf{x}_{c,r}^t \mathbf{e})$ equals 1 with negligible probability.

Claim 3.3. *Let $\mathbf{e} \leftarrow \mathrm{Bern}(\varepsilon)^\ell$ with $\varepsilon \le \frac{1}{8}$, and $\mathbf{x}_1, ..., \mathbf{x}_r \leftarrow \mathcal{X}_{\ell,k}$ be independent random variables. Then:*

$$\Pr_{\mathbf{e},\mathbf{x}_1,...,\mathbf{x}_r}\left[\mathrm{Maj}(\mathbf{x}_1^t\mathbf{e}, ..., \mathbf{x}_r^t\mathbf{e}) = 1\right] \le exp\left(-\varepsilon\ell/3\right) + exp\left(-r \cdot 2^{-16k\varepsilon}/4\right) \le n^{-\omega(1)}.$$

Proof. First, by *Lemma* 2.2, for any $\mathbf{w} \in \mathbb{F}_2^\ell$ with $\|\mathbf{w}\|_0 = \eta \le 2\varepsilon\ell$,

$$\Pr_{\mathbf{x} \leftarrow \mathcal{X}_{\ell,k}}\left[\mathbf{x}^t\mathbf{w} = 1\right] = \Pr\left[\sum_{i=1}^{k} \mathrm{Bern}\left(\frac{\eta}{\ell}\right) = 1\right] = \frac{1}{2} - \frac{1}{2}\left(1 - 2\frac{\eta}{\ell}\right)^k \le \frac{1}{2} - \frac{1}{2}\left(1 - 4\varepsilon\right)^k \le \frac{1}{2} - \frac{1}{2} \cdot 2^{-8\varepsilon k} \ ,$$

where we used the fact that $\forall y \in [0, 1/2]$, $2^{-2y} \le 1 - y$.
For any such $\mathbf{w}$, it holds by Chernoff-Hoeffding that

$$\Pr_{\mathbf{x}_1,...,\mathbf{x}_r}\left[\mathrm{Maj}(\mathbf{x}_1^t\mathbf{w}, ..., \mathbf{x}_r^t\mathbf{w}) = 1\right] \le \exp(-r \cdot 2^{-16\varepsilon k}/4) \ .$$

Also by multiplicative Chernoff,

$$\Pr_{\mathbf{e} \leftarrow \mathrm{Bern}(\varepsilon)^\ell}\left[\|\mathbf{e}\|_0 > 2\varepsilon\ell\right] \le \exp(-\varepsilon\ell/3) \ .$$

Overall,

$$\begin{aligned} &\Pr_{\mathbf{e},\mathbf{x}_1,...,\mathbf{x}_r}\left[\mathrm{Maj}(\mathbf{x}_1^t\mathbf{e}, ..., \mathbf{x}_r^t\mathbf{e}) = 1\right] \\ &\qquad \le \Pr_{\mathbf{e}}\left[\|\mathbf{e}\|_0 > 2\varepsilon\ell\right] + \max_{\mathbf{e}: \atop \|\mathbf{e}\|_0 \le 2\varepsilon\ell} \Pr_{\mathbf{x}_1,...,\mathbf{x}_r}\left[\mathrm{Maj}(\mathbf{x}_1^t\mathbf{e}, ..., \mathbf{x}_r^t\mathbf{e}) = 1\right] \\ &\qquad \le \exp(-\varepsilon\ell/3) + \exp(-r \cdot 2^{-16\varepsilon k}/4) \\ &\qquad \le \exp(-n^{\frac{1}{\delta}}(\log^2 n)/3) + \exp(-n^{64\delta+1} \cdot n^{-64\delta}/4) \le n^{-\omega(1)} \ , \end{aligned}$$

where the first to last inequality is by our setting of the parameters. □

This concludes the proof of CRS-free correctness. □

Receiver privacy follows directly from the LPN_ε assumption:

Proposition 3.4. *Under the LPN_ε assumption, the protocol satisfies receiver privacy.*

Proof. Under LPN_ε, the receiver message $\mathbf{v}_0$ is pseudorandom, regardless of its choice bit c. □

3.2 Statistical Sender Privacy Analysis

In this section we analyze the statistical sender privacy of the protocol. First, in Section 3.2.1, we relate the statistical sender privacy to a certain measure of balance on code cosets. Then in Section 3.2.2, we analyze the required balance conditions, and deduce sufficient conditions for them to hold. Finally, in Section 3.2.3, we tie the two together to deduce statistical sender privacy with bad-CRS certification.

3.2.1 Statistical Distance and Balanced Cosets

To prove statistical sender privacy, we aim to characterize which matrices $\mathbf{A} \in \mathbb{F}_2^{\ell \times n}$ and vectors $\mathbf{w} \in \mathbb{F}_2^{\ell}$ are such that $\mathbf{x}^t\mathbf{w}$ is statistically close to uniform even given the leakage $\mathbf{x}^t\mathbf{A}$, when $\mathbf{x} \leftarrow \mathcal{X}_{\ell,k}$. We prove the following proposition, which relates the relevant statistical distance to how balanced are vectors in the coset $\mathbf{w} + \mathbf{A}$ of the linear code given by $\mathbf{A}$.

Lemma 3.5. *Let $\mathbf{A} \in \mathbb{F}_2^{\ell \times n}$, $\mathbf{w} \in \mathbb{F}_2^{\ell}$. Also, for $\mathbf{t} \in \mathbb{F}_2^{\ell}$, let $\beta_{\mathbf{t}} := 1 - \frac{2}{\ell}\left\|\mathbf{t}\right\|_0$. Then,*

$$SD\left(\left(\mathbf{x}^t\mathbf{A}, \mathbf{x}^t\mathbf{w}\right), \left(\mathbf{x}^t\mathbf{A}, u\right)\right) \leq \frac{1}{2}\sum_{\mathbf{s}\in\mathbb{F}_2^n} \left|\beta_{\mathbf{A}\mathbf{s}+\mathbf{w}}\right|^k \quad ,$$

where $\mathbf{x} \leftarrow \mathcal{X}_{\ell,k}$, and $u \leftarrow \mathrm{Bern}\left(\frac{1}{2}\right)$.

We prove the lemma using Fourier analysis on the Boolean cube. We start by recalling the definition of the Hadamard matrix (corresponding to the Boolean Fourier transform), and then state and prove two lemmas needed to prove lemma 3.5.

Definition 3.6 (Hadamard matrix)**.** *The Hadamard matrices $\{\mathbf{H}^{\otimes n} \in \{\pm 1\}^{2^n \times 2^n}\}_{n\in\mathbb{N}}$ are defined inductively:*

$$\mathbf{H}^{\otimes 0} = \begin{pmatrix} 1 \end{pmatrix} \quad ,$$
$$\mathbf{H}^{\otimes n} = \begin{pmatrix} \mathbf{H}^{\otimes(n-1)} & \mathbf{H}^{\otimes(n-1)} \\ \mathbf{H}^{\otimes(n-1)} & -\mathbf{H}^{\otimes(n-1)} \end{pmatrix} \quad .$$

Note that for every $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$: $\mathbf{H}^{\otimes n}_{\mathbf{x},\mathbf{y}} = (-1)^{\langle \mathbf{x},\mathbf{y}\rangle}$, where we identify strings in $\mathbb{F}_2^n$ with indices in $[2^n]$ in the natural way.

Lemma 3.7. *Let $\ell \in \mathbb{N}$, $n \in \mathbb{N}$, $\mathbf{A} \in \mathbb{F}_2^{\ell \times n}$, $\mathbf{w} \in \mathbb{F}_2^{\ell}$, and let D be a distribution over $\mathbb{F}_2^{\ell}$, then:*

$$\begin{pmatrix} \Pr_{\mathbf{x}\leftarrow D}\left[\mathbf{x}^t\mathbf{A} = \vec{\mathbf{0}}, \mathbf{x}^t\mathbf{w} = 0\right] - \Pr_{\mathbf{x}\leftarrow D}\left[\mathbf{x}^t\mathbf{A} = \vec{\mathbf{0}}, \mathbf{x}^t\mathbf{w} = 1\right] \\ \vdots \\ \Pr_{\mathbf{x}\leftarrow D}\left[\mathbf{x}^t\mathbf{A} = \vec{\mathbf{1}}, \mathbf{x}^t\mathbf{w} = 0\right] - \Pr_{\mathbf{x}\leftarrow D}\left[\mathbf{x}^t\mathbf{A} = \vec{\mathbf{1}}, \mathbf{x}^t\mathbf{w} = 1\right] \end{pmatrix} = \frac{1}{2^n}\mathbf{H}^{\otimes n} \begin{pmatrix} \mathbb{E}_{\mathbf{x}\leftarrow D}\left[(-1)^{\mathbf{x}^t(\mathbf{A}\vec{\mathbf{0}}+\mathbf{w})}\right] \\ \vdots \\ \mathbb{E}_{\mathbf{x}\leftarrow D}\left[(-1)^{\mathbf{x}^t(\mathbf{A}\vec{\mathbf{1}}+\mathbf{w})}\right] \end{pmatrix} \quad ,$$

where above we consider all 2^n strings $\vec{\mathbf{0}}, \ldots, \vec{\mathbf{1}} \in \{0,1\}^n$ according to lexicographic order.

Proof. For any $k \in \mathbb{N}$ and $\mathbf{B} \in \mathbb{F}_2^{\ell \times k}$, consider the distribution $(\mathbf{x}^t\mathbf{B}, \mathbf{x}^t\mathbf{w})_{\mathbf{x}\leftarrow D}$. Note that for any $\mathbf{b} \in \mathbb{F}_2^k$:

$$\Pr_{\mathbf{x}\leftarrow D}\left[\mathbf{x}^t\mathbf{B} = \mathbf{b}\right] = \mathbb{E}_{\mathbf{x}\leftarrow D}\left[\mathbb{1}_{\mathbf{x}^t\mathbf{B}=\mathbf{b}}\right] = \mathbb{E}_{\mathbf{x}\leftarrow D}\left[\frac{1}{2^k}\sum_{\mathbf{s}\in\mathbb{F}_2^k}(-1)^{\langle \mathbf{x}^t\mathbf{B}+\mathbf{b},\mathbf{s}\rangle}\right] = \frac{1}{2^k}\sum_{\mathbf{s}\in\mathbb{F}_2^k}(-1)^{\langle \mathbf{b},\mathbf{s}\rangle} \cdot \mathbb{E}_{\mathbf{x}\leftarrow D}\left[(-1)^{\mathbf{x}^t\mathbf{B}\mathbf{s}}\right] \quad ,$$

which means

$$\begin{pmatrix} \Pr_{\mathbf{x}\leftarrow D}\left[\mathbf{x}^t\mathbf{B} = \vec{\mathbf{0}}\right] \\ \vdots \\ \Pr_{\mathbf{x}\leftarrow D}\left[\mathbf{x}^t\mathbf{B} = \vec{\mathbf{1}}\right] \end{pmatrix} = \frac{1}{2^k}\mathbf{H}^{\otimes k} \begin{pmatrix} \mathbb{E}_{\mathbf{x}\leftarrow D}\left[(-1)^{\mathbf{x}^t\mathbf{B}\vec{\mathbf{0}}}\right] \\ \vdots \\ \mathbb{E}_{\mathbf{x}\leftarrow D}\left[(-1)^{\mathbf{x}^t\mathbf{B}\vec{\mathbf{1}}}\right] \end{pmatrix} \quad .$$

Now, Because $\mathbf{H}^{\otimes(n+1)} = \begin{pmatrix} \mathbf{H}^{\otimes n} & \mathbf{H}^{\otimes n} \\ \mathbf{H}^{\otimes n} & -\mathbf{H}^{\otimes n} \end{pmatrix}$, the lemma follows when taking $\mathbf{B} = (\mathbf{w}|\mathbf{A}) \in \mathbb{F}_2^{\ell \times (n+1)}$. □

Lemma 3.8. *Let* $\mathbf{A} \in \mathbb{F}_2^{\ell \times n}$, $\mathbf{w} \in \mathbb{F}_2^{\ell}$. *Also, for* $\mathbf{t} \in \mathbb{F}_2^{\ell}$ *let* $\beta_{\mathbf{t}} := 1 - \frac{2}{\ell}\left\|\mathbf{t}\right\|_0$. *Then for* $\mathbf{x} \leftarrow \mathcal{X}_{\ell,k}$,

$$\begin{pmatrix} \Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \vec{\mathbf{0}}, \mathbf{x}^t\mathbf{w} = 0\right] - \Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \vec{\mathbf{0}}, \mathbf{x}^t\mathbf{w} = 1\right] \\ \vdots \\ \Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \vec{\mathbf{1}}, \mathbf{x}^t\mathbf{w} = 0\right] - \Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \vec{\mathbf{1}}, \mathbf{x}^t\mathbf{w} = 1\right] \end{pmatrix} = \frac{1}{2^n}\mathbf{H}^{\otimes n}\begin{pmatrix} \beta^k_{A\vec{\mathbf{0}}+\mathbf{w}} \\ \vdots \\ \beta^k_{A\vec{\mathbf{1}}+\mathbf{w}} \end{pmatrix},$$

where above we consider all 2^n *strings* $\vec{\mathbf{0}}, \ldots, \vec{\mathbf{1}} \in \{0,1\}^n$ *according to lexicographic order.*

Proof. The lemma follows directly from Lemma 3.7, and the observation that for any $\mathbf{t} \in \mathbb{F}_2^{\ell}$,

$$\mathbb{E}_{\mathbf{x}\leftarrow\mathcal{X}_{\ell,k}}\left[(-1)^{\langle \mathbf{x},\mathbf{t}\rangle}\right] = \mathbb{E}_{\mathbf{x}_1,\ldots,\mathbf{x}_k\leftarrow\{\mathbf{e}_1,..,\mathbf{e}_\ell\}}\left[\prod_{i=1}^{k}(-1)^{\langle \mathbf{x}_i,\mathbf{t}\rangle}\right] = \prod_{i=1}^{k}\mathbb{E}_{\mathbf{x}_1\leftarrow\{\mathbf{e}_1,..,\mathbf{e}_\ell\}}\left[(-1)^{\langle \mathbf{x}_1,\mathbf{t}\rangle}\right] = \left(1 - \frac{2}{\ell}\left\|\mathbf{t}\right\|_0\right)^k.$$

□

We are now ready to prove Lemma 3.5.

Proof of Lemma 3.5.

$$\begin{aligned}
&SD_{\mathbf{x}\leftarrow\mathcal{X}_{\ell,k},u\leftarrow\mathrm{Bern}(\frac{1}{2})}\left(\left(\mathbf{x}^t\mathbf{A}, \mathbf{x}^t\mathbf{w}\right), \left(\mathbf{x}^t\mathbf{A}, u\right)\right) \\
&= \frac{1}{2}\sum_{\mathbf{a}\in\mathbb{F}_2^n}\left(\left|\Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \mathbf{a}, \mathbf{x}^t\mathbf{w} = 0\right] - \frac{1}{2}\Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \mathbf{a}\right]\right|\right. \\
&\qquad\left. + \left|\Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \mathbf{a}, \mathbf{x}^t\mathbf{w} = 1\right] - \frac{1}{2}\Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \mathbf{a}\right]\right|\right) \\
&= \frac{1}{2}\sum_{\mathbf{a}\in\mathbb{F}_2^n}\left(\frac{1}{2}\left|\Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \mathbf{a}, \mathbf{x}^t\mathbf{w} = 0\right] - \Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \mathbf{a}, \mathbf{x}^t\mathbf{w} = 1\right]\right|\right. \\
&\qquad\left. + \frac{1}{2}\left|\Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \mathbf{a}, \mathbf{x}^t\mathbf{w} = 1\right] - \Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \mathbf{a}, \mathbf{x}^t\mathbf{w} = 0\right]\right|\right) \\
&= \frac{1}{2}\sum_{\mathbf{a}\in\mathbb{F}_2^n}\left|\Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \mathbf{a}, \mathbf{x}^t\mathbf{w} = 0\right] - \Pr_{\mathbf{x}}\left[\mathbf{x}^t\mathbf{A} = \mathbf{a}, \mathbf{x}^t\mathbf{w} = 1\right]\right| \\
&= \frac{1}{2}\left\|\begin{pmatrix} \Pr_{\mathbf{x}\leftarrow\mathcal{X}_{\ell,k}}\left[\mathbf{x}^t\mathbf{A} = \vec{\mathbf{0}}, \mathbf{x}^t\mathbf{w} = 0\right] - \Pr_{\mathbf{x}\leftarrow\mathcal{X}_{\ell,k}}\left[\mathbf{x}^t\mathbf{A} = \vec{\mathbf{0}}, \mathbf{x}^t\mathbf{w} = 1\right] \\ \vdots \\ \Pr_{\mathbf{x}\leftarrow\mathcal{X}_{\ell,k}}\left[\mathbf{x}^t\mathbf{A} = \vec{\mathbf{1}}, \mathbf{x}^t\mathbf{w} = 0\right] - \Pr_{\mathbf{x}\leftarrow\mathcal{X}_{\ell,k}}\left[\mathbf{x}^t\mathbf{A} = \vec{\mathbf{1}}, \mathbf{x}^t\mathbf{w} = 1\right] \end{pmatrix}\right\|_1 \\
&= \frac{1}{2}\left\|\frac{1}{2^n}\mathbf{H}^{\otimes n}\begin{pmatrix} \beta^k_{\mathbf{A}\vec{\mathbf{0}}+\mathbf{w}} \\ \vdots \\ \beta^k_{\mathbf{A}\vec{\mathbf{1}}+\mathbf{w}} \end{pmatrix}\right\|_1 \qquad \text{(by Lemma 3.8)} \\
&\le \frac{1}{2}\left\|\frac{1}{2^{n/2}}\mathbf{H}^{\otimes n}\begin{pmatrix} \beta^k_{\mathbf{A}\vec{\mathbf{0}}+\mathbf{w}} \\ \vdots \\ \beta^k_{\mathbf{A}\vec{\mathbf{1}}+\mathbf{w}} \end{pmatrix}\right\|_2 \qquad \text{(Cauchy-Schwartz)} \\
&= \frac{1}{2}\left\|\begin{pmatrix} \beta^k_{\mathbf{A}\vec{\mathbf{0}}+\mathbf{w}} \\ \vdots \\ \beta^k_{\mathbf{A}\vec{\mathbf{1}}+\mathbf{w}} \end{pmatrix}\right\|_2 \qquad (2^{-n/2}\mathbf{H}^{\otimes n} \text{ is orthonormal})
\end{aligned}$$

$$\leq \frac{1}{2}\left\| \begin{pmatrix} \beta^k_{\mathbf{A}\vec{0}+\mathbf{w}} \\ \vdots \\ \beta^k_{\mathbf{A}\vec{1}+\mathbf{w}} \end{pmatrix} \right\|_1 .$$

This concludes the proof. □

3.2.2 Balance of Code Cosets

Following Lemma 3.5 from the previous section, in this section we analyze the balance properties of code cosets. Concretely, our goal is to find sufficient conditions to guarantee that no matter how an adversarial receiver decomposes $\mathbf{v}$ into $\mathbf{v}_0 + \mathbf{v}_1 = \mathbf{v}$, it must be that one of the cosets $\mathbf{v}_i + \mathbf{A}$ will be balanced (in which case we can invoke Lemma 3.5).

Step I: In Any Decomposition, One Coset is Somewhat Balanced. Our first step is to show that when $\mathbf{v}$ as chosen at random (as in the CRS), then any sum decomposition $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_1$ will induce at least one coset $\mathbf{v}_i + \mathbf{A}$ in which all members are *somewhat balanced*. Jumping ahead, this balance alone will not suffice, and our second step will deal with the additional balance properties required.

Definition 3.9. *For all $\mathbf{A} \in \mathbb{F}_2^{\ell \times n}, \mathbf{v} \in \mathbb{F}_2^\ell$, we use the (abuse of) notion $\|\mathbf{A} + \mathbf{v}\|_0$ to denote the minimal distance between $\mathbf{v}$ and the image of $\mathbf{s} \mapsto \mathbf{A}\mathbf{s}$, formally: $\|\mathbf{A} + \mathbf{v}\|_0 := \min_{\mathbf{s} \in \mathbb{F}_2^n} \|\mathbf{A}\mathbf{s} + \mathbf{v}\|_0$.*

Note that $\|\mathbf{A} + \mathbf{v}\|_0$ satisfies the triangle inequality:

$$\|\mathbf{A} + (\mathbf{v}_0 + \mathbf{v}_1)\|_0 \leq \|\mathbf{A} + \mathbf{v}_0\|_0 + \|\mathbf{A} + \mathbf{v}_1\|_0 .$$

Definition 3.10. *Let $\mathbf{A} \in \mathbb{F}_2^{\ell \times n}$, $\mathbf{v} \in \mathbb{F}_2^\ell$. We say that $\mathbf{v}$ is $\mathbf{A}$-balanced for sums if for all $\mathbf{v}_0, \mathbf{v}_1$ such that $\mathbf{v}_0 + \mathbf{v}_1 = \mathbf{v}$, there exists $i \in \{0, 1\}$ such that for all $\mathbf{s} \in \mathbb{F}_2^n$:*

$$\frac{\ell}{5} \leq \|\mathbf{A}\mathbf{s} + \mathbf{v}_i\|_0 \leq \frac{4\ell}{5} .$$

Proposition 3.11 (A random $\mathbf{v}$ is $\mathbf{A}$-balanced for sums)**.** *For any $\mathbf{A} \in \mathbb{F}_2^{\ell \times n}$:*

$$\Pr_{\mathbf{v} \leftarrow \mathbb{F}_2^\ell}[\mathbf{v} \text{ is } \mathbf{A}\text{-balanced for sums}] \geq 1 - \frac{2^{n+1}}{2^{\Omega(\ell)}} .$$

Proof. Define $\mathbf{A}' := (\mathbf{A}|\vec{\mathbf{1}}) \in \mathbb{F}_2^{\ell \times (n+1)}$. Observe that:

$$\Pr_{\mathbf{v} \leftarrow \mathbb{F}_2^\ell}\left[\|\mathbf{A}' + \mathbf{v}\|_0 \geq \frac{2}{5}\ell\right] \geq \Pr_{\mathbf{v} \leftarrow \mathbb{F}_2^\ell}\left[\forall \mathbf{s} \in \mathbb{F}_2^{n+1} : \left|\|\mathbf{A}'\mathbf{s} + \mathbf{v}\|_0 - \frac{\ell}{2}\right| \leq \frac{1}{10}\ell\right] \geq 1 - \frac{2^{n+1}}{2^{\Omega(\ell)}} ,$$

where the above follows from Lemma 2.3 and the fact that for any $\mathbf{s}$, $\mathbf{A}'\mathbf{s} + \mathbf{v}$ is uniformly random over $\mathbb{F}_2^\ell$, as well as a union bound.

Now, for any $\mathbf{v}_0, \mathbf{v}_1 \in \mathbb{F}_2^\ell$ such that $\mathbf{v}_0 + \mathbf{v}_1 = \mathbf{v}$, by the triangle inequality,

$$\|\mathbf{A}' + \mathbf{v}\|_0 \geq \frac{2}{5}\ell \Longrightarrow \exists i \in \{0, 1\} : \|\mathbf{A}' + \mathbf{v}_i\|_0 \geq \frac{1}{5}\ell .$$

Finally, observe that for every $\mathbf{w} \in \mathbb{F}_2^\ell$ and $\gamma \leq 1/2$, if $\|\mathbf{A}' + \mathbf{w}\|_0 \geq \gamma\ell$ then $\forall \mathbf{s} \in \mathbb{F}_2^n : \left|\|\mathbf{A}\mathbf{s} + \mathbf{w}\|_0 - \frac{1}{2}\ell\right| \leq (\frac{1}{2} - \gamma)\ell$. Indeed, for all $\mathbf{s} \in \mathbb{F}_2^\ell$:

$$\|\mathbf{A}\mathbf{s} + \mathbf{w}\|_0 \geq \gamma\ell \Longrightarrow \gamma\ell - \frac{1}{2}\ell \leq \|\mathbf{A}\mathbf{s} + \mathbf{w}\|_0 - \frac{1}{2}\ell ,$$

$$\left\|\mathbf{A}\mathbf{s} + \mathbf{w} + \vec{\mathbf{1}}\right\|_0 \geq \gamma\ell \Longrightarrow \ell - \|\mathbf{A}\mathbf{s} + \mathbf{w}\|_0 \geq \gamma\ell \Longrightarrow \|\mathbf{A}\mathbf{s} + \mathbf{w}\|_0 - \frac{1}{2}\ell \leq \frac{1}{2}\ell - \gamma\ell .$$

The lemma now follows when setting $\gamma = \frac{1}{5}$. □

Step II: Almost All Coset Members Are Well Balanced. The balance property defined above is still not sufficient for meaningfully invoking the statistical distance bound given by Lemma 3.5. Indeed, directly using the bound $\sum_{\mathbf{s}\in\mathbb{F}_2^n} |\beta_{\mathbf{As}+\mathbf{w}}|^k$ given by the lemma would require that the maximum bias β is such that $\beta \ll 2^{-n/k}$. However, in our case, to guarantee correctness $k \approx n/\log n$ and bounding β by a constant is insufficient. Using a more careful analysis, we will prove that in fact, for a random matrix $\mathbf{A}$, it is the case that in all cosets $\mathbf{w}+\mathbf{A}$, almost all members are well (rather than somewhat) balanced, and in particular have maximal bias $\beta \ll 2^{-n/k}$. This will allow using the relatively weak balance property from Step I on a sufficiently small set. We proceed with the relevant definitions and analysis.

Definition 3.12 ((β, D)-Affine-Balance)**.** *Let $\beta \in [0,1]$, $D \in \mathbb{N}$ and $\mathbf{A} \in \mathbb{F}_2^{\ell\times n}$. We say that $\mathbf{A}$ is (β, D)-affinely-balanced if for all $\mathbf{w} \in \mathbb{F}_2^\ell$ there exists a set $E_\mathbf{w} \subseteq \mathbb{F}_2^n$ such that $|E_\mathbf{w}| < D$, and:*

$$\forall \mathbf{s} \in \mathbb{F}_2^n \setminus E_\mathbf{w} : (1-\beta)\frac{\ell}{2} \leq \|\mathbf{As}+\mathbf{w}\|_0 \leq (1+\beta)\frac{\ell}{2}$$

We also define (and achieve) a stronger balance property that will be useful for showing bad-CRS certification.

Definition 3.13 ((β, d)-Strong-Balance)**.** *Let $\beta \in [0,1]$, $d \in \mathbb{N}$ and $\mathbf{A} \in \mathbb{F}_2^{\ell\times n}$. We say that $\mathbf{A}$ is (β, d)-strongly-balanced if for all $\mathbf{w} \in \mathbb{F}_2^\ell$, and any set of d **linearly independent** vectors $\mathbf{s}_1, ..., \mathbf{s}_d \in \mathbb{F}_2^n$, there exists some $i \in [d]$ such that:*

$$(1-\beta)\frac{\ell}{2} \leq \|\mathbf{As}_i+\mathbf{w}\|_0 \leq (1+\beta)\frac{\ell}{2} \ .$$

Proposition 3.14 (From strong to affine balance)**.** *Any $\mathbf{A} \in \mathbb{F}_2^{\ell\times n}$ which is (β, d)-strongly-balanced, is also $(\beta, 2^d)$-affinely balanced.*

Proof. The proposition follows from the fact that any set of 2^d vectors over $\mathbb{F}_2^n$ contains a set of d linearly independent vectors. □

Proposition 3.15. *For $\beta \geq \ell^{-1/2}$, a random $\mathbf{A} \xleftarrow{\$} \mathbb{F}_2^{\ell\times n}$ is (β, d)-strongly-balanced with probability at least:*

$$\Pr_{\mathbf{A}\leftarrow\mathbb{F}_2^{\ell\times n}}[\mathbf{A} \text{ is } (\beta,d)\text{-strongly-balanced}] \geq 1 - 2^{\ell+n\cdot d+2d-\frac{1}{4}d\beta^2\ell} \ .$$

Proof. Assume $\mathbf{A} \xleftarrow{\$} \mathbb{F}_2^{\ell\times n}$. We will bound the probability that $\mathbf{A}$ is not (β, d)-strongly-balanced. This happens when there exists $\mathbf{w} \in \{0,1\}^\ell$ and d linearly independent vectors $\mathbf{s}_1, ..., \mathbf{s}_d$ such that $\forall i \in [d]$: $\left|\|\mathbf{As}_i+\mathbf{w}\|_0 - \frac{\ell}{2}\right| > \frac{1}{2}\beta\ell$. We bound the probability that such $\mathbf{w}, \mathbf{s}_1, ..., \mathbf{s}_d$ exist.
First, for any fixed $\mathbf{w} \in \mathbb{F}_2^\ell$ and $0 \neq \mathbf{s} \in \mathbb{F}_2^n$ it holds that $\mathbf{As}+\mathbf{w}$ is uniformly distributed over $\{0,1\}^\ell$, and therefore from *Lemma* 2.3 we get:

$$\Pr_{\mathbf{A}}\left[\left|\|\mathbf{As}+\mathbf{w}\|_0 - \frac{\ell}{2}\right| > \frac{1}{2}\beta\ell\right] \leq 2^{-\frac{1}{4}\beta^2\ell} \ .$$

Similarly, for any fixed $\mathbf{w} \in \mathbb{F}_2^\ell$, and any set of d linearly independent vectors $\{\mathbf{s}_1, ..., \mathbf{s}_d\} \subseteq \mathbb{F}_2^n$ it holds that $(\mathbf{As}_1+\mathbf{w}, ..., \mathbf{As}_d+\mathbf{w})$ is uniformly distributed over $\{0,1\}^{\ell\times d}$, and independence implies:

$$\Pr_{\mathbf{A}}\left[\forall i \in [d] : \left|\|\mathbf{As}_i+\mathbf{w}\|_0 - \frac{\ell}{2}\right| > \frac{1}{2}\beta\ell\right] \leq \left(2^{-\frac{1}{4}\beta^2\ell}\right)^d$$

Finally, by the union bound, and the fact that $\binom{m}{k} \leq \left(\frac{m\cdot e}{k}\right)^k$,

$$\Pr_{\mathbf{A}}\left[\exists \mathbf{w} \in \{0,1\}^\ell, \text{linearly independent } \mathbf{s}_1, ..., \mathbf{s}_d : \forall i \in [d], \left|\|\mathbf{As}_i+\mathbf{w}\|_0 - \frac{\ell}{2}\right| > \frac{1}{2}\beta\ell\right] \leq 2^\ell \binom{2^n}{d}\left(2^{-\frac{1}{4}\beta^2\ell}\right)^d$$
$$\leq 2^{\ell+n\cdot d+2d-\frac{1}{4}d\beta^2\ell} \ .$$

□

3.2.3 Putting Things Together: Statistical Sender Privacy

We are now ready to prove that the protocol is statistically-sender-private. In fact we will prove the stronger property of bad-CRS certifiability.

Proposition 3.16. *The protocol is statistically-sender-private. Moreover, it is bad-CRS certifiable.*

Proof. In what follows, $\delta = 1+\Theta(1)$, $\ell = n^{1+\frac{1}{\delta}}$ and $k = 4\delta n/\log(n)$ are as previously set in our construction, and $d(n) = n/\log^2(n)$, $\beta(n) = 4\sqrt{n/\ell} = 4n^{-1/(2\delta)}$.

We first define the set of bad CRSs:

$$\mathsf{B} := \bigcup_n \left\{ (\mathbf{A}, \mathbf{v}) \in \mathbb{F}_2^{\ell \times n} \times \mathbb{F}_2^{\ell} \;\middle|\; \mathbf{A} \text{ is } \textbf{not } (\beta, d)\text{-strongly-balanced } \textbf{OR } \mathbf{v} \text{ is } \textbf{not } \mathbf{A}\text{-balanced for sums} \right\} .$$

We next establish each of the three requirements of sender-statistical-privacy with bad CRS certification.

Claim 3.17. *Sender statistical privacy outside of* B *is satisfied.*

Proof. Fix $(\mathbf{A}, \mathbf{v}) \notin \mathsf{B}$ and any decomposition $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_1$. Since $\mathbf{v}$ is $\mathbf{A}$-balanced for sums, there exists $i \in \{0,1\}$ such that for all $\mathbf{s} \in \mathbb{F}_2^n$:

$$\frac{\ell}{5} \leq \|\mathbf{A}\mathbf{s} + \mathbf{v}_i\|_0 \leq \frac{4\ell}{5} \ . \tag{1}$$

Let $i \in \{0,1\}$ be (the minimal) such that the above holds.

$$\text{The extractor } \mathsf{OTExt}(\mathbf{A}, \mathbf{v}, \mathbf{v}_0) \text{ outputs } 1 - i.$$

To conclude the proof, we bound the statistical distance $SD\left((\mathbf{x}^t\mathbf{A}, \mathbf{x}^t\mathbf{v}_i), (\mathbf{x}^t\mathbf{A}, u)\right)$ for $\mathbf{x} \leftarrow \mathcal{X}_{\ell,k}$, $u \leftarrow \text{Bern}\left(\frac{1}{2}\right)$. In what follows, for $\mathbf{t} \in \mathbb{F}_2^{\ell}$ let $\beta_{\mathbf{t}} := 1 - \frac{2}{\ell}\|\mathbf{t}\|_0$. Also, let $E_{\mathbf{v}_i} \subseteq \mathbb{F}_2^n$ be the set given by Definition 3.12, where its existence is guaranteed by Proposition 3.14 and the fact that $\mathbf{A}$ is (β, d)-strongly-balanced.

$$\begin{aligned}
&SD\left(\left(\mathbf{x}^t\mathbf{A}, \mathbf{x}^t\mathbf{v}_i\right), \left(\mathbf{x}^t\mathbf{A}, u\right)\right) \\
&\qquad \leq \sum_{\mathbf{s} \in \mathbb{F}_2^n} |\beta_{\mathbf{A}\mathbf{s}+\mathbf{v}_i}|^k && \text{(By Lemma 3.5)} \\
&\qquad = \sum_{\mathbf{s} \in E_{\mathbf{v}_i}} |\beta_{\mathbf{A}\mathbf{s}+\mathbf{v}_i}|^k + \sum_{\mathbf{s} \notin E_{\mathbf{v}_i}} |\beta_{\mathbf{A}\mathbf{s}+\mathbf{v}_i}|^k \\
&\qquad \leq |E_{\mathbf{v}_i}| \cdot \max_{\mathbf{s} \in \mathbb{F}_2^n} |\beta_{\mathbf{A}\mathbf{s}+\mathbf{v}_i}|^k + 2^n \cdot \max_{s \notin E_{\mathbf{v}_i}} |\beta_{\mathbf{A}\mathbf{s}+\mathbf{v}_i}|^k \\
&\qquad \leq 2^d \cdot \max_{\mathbf{s} \in \mathbb{F}_2^n} |\beta_{\mathbf{A}\mathbf{s}+\mathbf{v}_i}|^k + 2^n \cdot \beta^k && \text{(By Definition 3.12)} \\
&\qquad \leq 2^d \cdot (3/5)^k + 2^n \cdot \beta^k && \text{(By Equation (1))} \\
&\qquad = 2^{\frac{n}{\log^2 n}} \cdot (3/5)^{\frac{4\delta n}{\log n}} + 2^n \cdot (4n^{-1/(2\delta)})^{\frac{4\delta n}{\log n}} && \text{(By our parameter setting)} \\
&\qquad = 2^{-\Omega(n/\log n)} + 2^n \cdot 2^{\frac{8\delta n}{\log n} - 2n} \\
&\qquad = 2^{-\Omega(n/\log n)} \ .
\end{aligned}$$

□

Claim 3.18. B *is certifiable.*

Proof. $(\mathbf{A}, \mathbf{v}) \in \mathsf{B}$ if and only if either one of the following holds:

- **$\mathbf{A}$ is not (β, d)-strongly-balanced:** there exist $\mathbf{w} \in \mathbb{F}_2^\ell$ and d linearly independent vectors $\mathbf{s}_1, .., \mathbf{s}_d \in \mathbb{F}_2^n$ such that $\forall i \in [d]$:
$$\left| \|\mathbf{A}\mathbf{s}_i + \mathbf{w}\|_0 - \frac{\ell}{2} \right| > \beta \frac{\ell}{2} \ .$$
- **$\mathbf{v}$ is not $\mathbf{A}$ balanced for sums:** there exist $\mathbf{v}_0, \mathbf{v}_1 \in \mathbb{F}_2^\ell$, $\mathbf{s}_0, \mathbf{s}_1 \in \mathbb{F}_2^n$ such that $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_1$ and for both $i \in \{0, 1\}$:
$$\left| \|\mathbf{A}\mathbf{s}_i + \mathbf{v}_i\|_0 - \frac{\ell}{2} \right| > \frac{3}{5} \cdot \frac{\ell}{2} \ .$$

Given $(\mathbf{w}, \mathbf{s}_1, .., \mathbf{s}_d)$, respectively $(\mathbf{v}_0, \mathbf{v}_1, \mathbf{s}_0, \mathbf{s}_1)$, the first, respectively the second, condition can be efficiently checked. Hence $\mathsf{B} \in \mathbf{NP}$. □

Claim 3.19. B *has negligible density*

Proof. By *Proposition* 3.15,

$$\begin{aligned} \Pr_{\mathbf{A} \leftarrow \mathbb{F}_2^{\ell \times n}} [\mathbf{A} \text{ is not } (\beta, d)\text{-strongly-balanced}] & \\ &\leq 2^{\ell + n \cdot d + 2d - \frac{1}{4} d \beta^2 \ell} \\ &= 2^{n^{1+1/\delta} + \frac{n(n+2)}{\log^2 n} - 4\frac{n^2}{\log^2 n}} \\ &= 2^{-\Omega(n^2/\log^2 n)} \ . \end{aligned}$$

By *Proposition* 3.11, for every $\mathbf{A} \in \mathbb{F}_2^{\ell \times n}$:

$$\Pr_{\mathbf{v} \leftarrow \mathbb{F}_2^\ell} [\mathbf{v} \text{ is not } \mathbf{A}\text{-balanced for sums}] \leq \frac{2^{n+1}}{2^{\Omega(\ell)}} \leq 2^{-\Omega(n^{1+1/\delta})} \ .$$

Overall, by the union bound,

$$\Pr_{\mathbf{A}, \mathbf{v}} [(\mathbf{A}, \mathbf{v}) \in \mathsf{B}] \leq 2^{-\Omega(n^{1+1/\delta})} \ .$$

□

This concludes the proof of Proposition 3.16.

□

4 From the CRS Model to the SRS and Plain Models

In this section, we show transformations from the CRS model to the SRS model, and then to the plain model.

4.1 From the CRS Model to the SRS Model

In this section, we show how to transform any two-message SSP OT in the common random string model into one in the sender random string model. Recall that this model is similar to the common random string model, except that receiver privacy holds even for an adversarial (rather than random) choice of the common string. The transformation is based on the idea of reverse randomization from [DN00] (tracing back to [Lau83]).

In what follows, we denote the original protocol by (S, R) and its CRS length by ρ and construct a new protocol $(\mathsf{S}', \mathsf{R}')$ with SRS length ρ^2. The transformation is presented in Figure 2.

We prove:

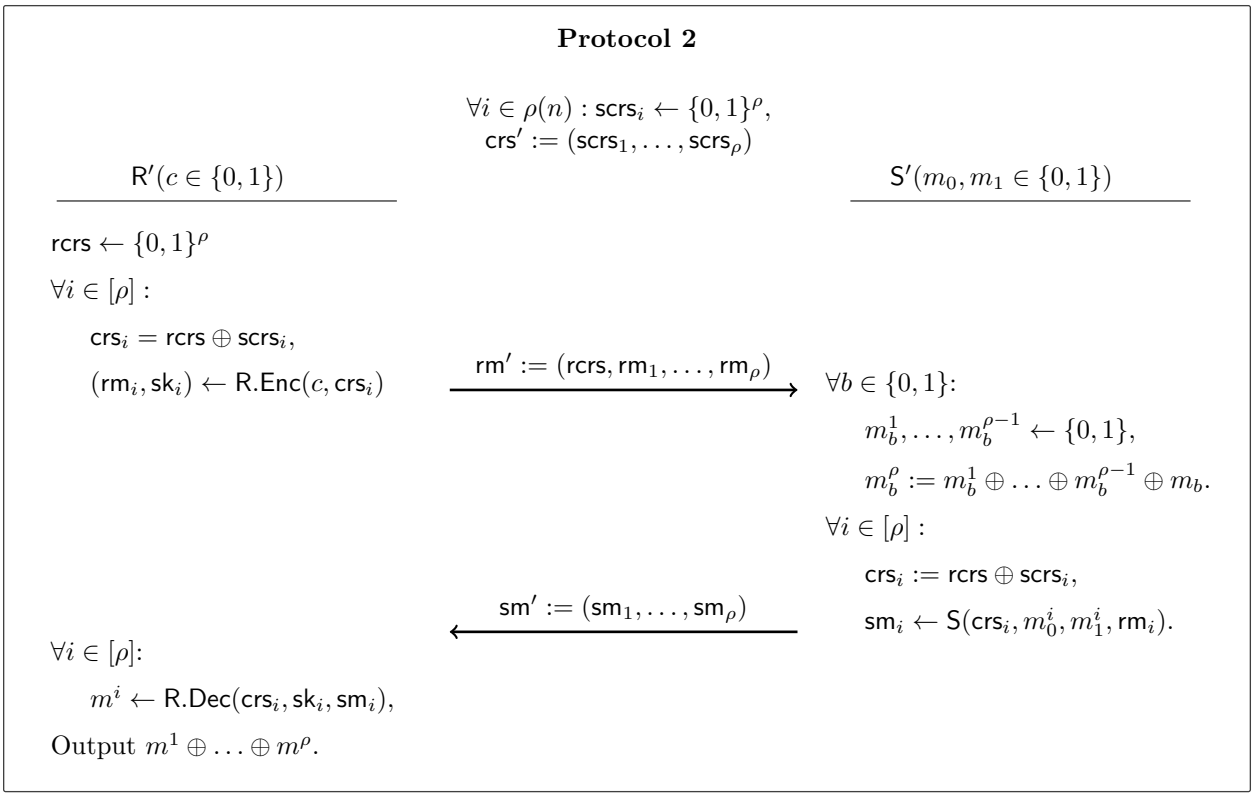

Figure 2: Two-message statistically-sender-private OT in the SRS model

Theorem 4.1. *Assuming* (S, R) *is a two-message statistically sender-private OT in CRS model, then* $(\mathsf{S}', \mathsf{R}')$ *is a two-message statistically sender-private OT in SRS model. Moreover,* $(\mathsf{S}', \mathsf{R}')$ *has CRS-free correctness (even if* (S, R) *does not), and if* (S, R) *has bad-CRS certification so does* $(\mathsf{S}', \mathsf{R}')$.

Corollary 4.2. *Under the* $LPN_{\frac{\log^2(n)}{n}}$ *assumption, there exists a two-message statically-sender-private OT protocol in the SRS model. Moreover, the protocol has CRS-free correctness and bad-CRS certification.*

To prove Theorem 4.1, we prove each of the required properties. We first prove CRS-free correctness.

Proposition 4.3. *Protocol* $(\mathsf{S}', \mathsf{R}')$ *has CRS-free correctness.*

Proof. Fix any CRS $\mathsf{crs}' = (\mathsf{scrs}_1, \ldots, \mathsf{scrs}_\rho)$. Then, note that the marginal of each $\mathsf{crs}_i = \mathsf{rcrs} \oplus \mathsf{scrs}_i$ is uniform since $\mathsf{rcrs} \leftarrow \{0,1\}^\rho$. It follows that except with negligible probability $n^{-\omega(1)}$ over the coins of $(\mathsf{S}', \mathsf{R}')$, the receiver learns m_c^i. By the union bound, the receiver learns $m_c = \oplus_i m_c^i$, except with negligible probability $\rho n^{-\omega(1)}$. □

Proposition 4.4. *Protocol* $(\mathsf{S}', \mathsf{R}')$ *satisfies receiver privacy.*

Proof. Receiver privacy for an adversarial choice of $\mathsf{crs}' = (\mathsf{scrs}_1, \ldots, \mathsf{scrs}_\rho)$ follows by a standard hybrid argument. We consider $\rho + 1$ hybrids, where in hybrid i, in the first i copies of the underlying protocol (S, R), the receiver uses the choice bit 0, and in the last $\rho - i$ copies, it uses the choice bit 1. Note that the difference between hybrid $i-1$ and hybrid i are only in the i-th copy, and any distinguisher D' between the hybrids directly implies a distinguisher D against the receiver privacy of the underlying protocol (S, R). The distinguisher D given rm and crs, will set $\mathsf{rcrs} = \mathsf{crs} \oplus \mathsf{scrs}_i$ and will run the distinguisher D', simulating

$$\mathsf{crs}_1, \mathsf{rm}_1, \ldots, \mathsf{crs}_{i-1}, \mathsf{rm}_{i-1}, \mathsf{crs}_{i+1}, \mathsf{rm}_{i+1}, \ldots, \mathsf{crs}_\rho, \mathsf{rm}_\rho$$

on its own, and planting $\mathsf{crs}, \mathsf{rm}$ as $\mathsf{crs}_i, \mathsf{rm}_i$. □

Proposition 4.5. *Protocol* $(\mathsf{S}', \mathsf{R}')$ *is statistically-sender-private. Moreover, if* (S, R) *is bad-CRS certifiable so is* $(\mathsf{S}', \mathsf{R}')$.

Proof. By the statistical sender privacy of (S, R) (and its equivalent formulation in Remark 2.1) there exists a set B such that:

- **Statistical Sender Privacy Outside B:** There exists an (unbounded) OTExt, such that for any (unbounded) R^*:

$$\left\{ \mathsf{sm} \;\middle|\; \begin{array}{l} \mathsf{rm} \leftarrow \mathsf{R}^*(\mathsf{crs}) \\ \mathsf{sm} \leftarrow \mathsf{S}(\mathsf{crs}, m_0, m_1, \mathsf{rm}) \end{array} \right\}_{n,\mathsf{crs},m_0,m_1} \approx_s \left\{ \mathsf{sm} \;\middle|\; \begin{array}{l} \mathsf{rm} \leftarrow \mathsf{R}^*(\mathsf{crs}) \\ b \leftarrow \mathsf{OTExt}(\mathsf{crs}, \mathsf{rm}) \\ \mathsf{sm} \leftarrow \mathsf{S}(\mathsf{crs}, m_b, m_b, \mathsf{rm}) \end{array} \right\}_{n,\mathsf{crs},m_0,m_1},$$

  where $n \in \mathbb{N}, \mathsf{crs} \in \{0,1\}^{\rho(n)} \setminus \mathsf{B}, m_0, m_1 \in \{0,1\}$.

- **Negligible Density:** $\Pr\left[\mathsf{crs} \in \mathsf{B} \;\middle|\; \mathsf{crs} \leftarrow \{0,1\}^{\rho(n)}\right] \leq n^{-\omega(1)}$.

Furthermore, if (S, R) is bad-CRS certifiable then $\mathsf{B} \in \mathbf{NP}$.

We define the extractor $\mathsf{OTExt}'(\mathsf{crs}', \mathsf{rm}')$ as follows:

- Parse
$$\mathsf{crs}' = (\mathsf{scrs}_1, \ldots, \mathsf{scrs}_\rho), \qquad \mathsf{rm}' = (\mathsf{rcrs}, \mathsf{rm}_1, \ldots, \mathsf{rm}_\rho) \ .$$
- For the first i such that $\mathsf{crs}_i = \mathsf{rcrs} \oplus \mathsf{scrs} \notin \mathsf{B}$, output $\mathsf{OTExt}(\mathsf{crs}_i, \mathsf{rm}_i)$, or $\perp$ if not such i exists.

We now consider a new set of bad CRSs B', which includes all $\mathsf{crs}' = (\mathsf{scrs}_1, \ldots, \mathsf{scrs}_\rho)$ such that there exists rcrs so that for every i $\mathsf{scrs}_i \oplus \mathsf{rcrs} \in \mathsf{B}$. First, we note that B' has negligible density:

$$\begin{aligned} \Pr_{\mathsf{scrs}_1,\ldots,\mathsf{scrs}_\rho}\left[\exists \mathsf{rcrs} \in \{0,1\}^{\rho(n)} \text{ s.t. } \forall i \in [\rho] : \mathsf{scrs}_i \oplus \mathsf{rcrs} \in B\right] & \\ \leq \sum_{\mathsf{rcrs} \in \{0,1\}^{\rho(n)}} \Pr_{\mathsf{scrs}_1,\ldots,\mathsf{scrs}_\rho}\left[\forall i \in [\rho] : \mathsf{scrs}_i \oplus \mathsf{rcrs} \in B\right] & \\ \leq 2^\rho \cdot n^{-\omega(1)\cdot\rho} & \\ \leq n^{-\omega(1)} \ . & \end{aligned}$$

Given any $\mathsf{crs}' = (\mathsf{scrs}_1, \ldots, \mathsf{scrs}_\rho)$ where there exists i such that $\mathsf{crs}_i = \mathsf{scrs}_i \oplus \mathsf{rcrs} \notin \mathsf{B}$, we consider a hybrid experiment where in the i-th copy, $\mathsf{S}(\mathsf{crs}_i, m_0^i, m_1^i, \mathsf{rm}_i)$ is replaced with $\mathsf{S}(\mathsf{crs}_i, m_b^i, m_b^i, \mathsf{rm}_i)$ where $b = \mathsf{OTExt}(\mathsf{crs}_i, \mathsf{rm}_i)$. This hybrid is statistically close to the actual experiment, but is independent of the share m_{1-b}^i, and thus also of the message m_{1-b}. It follows that for any R^*,

$$\left\{ \mathsf{sm}' \;\middle|\; \begin{array}{l} \mathsf{rm}' \leftarrow \mathsf{R}^*(\mathsf{crs}') \\ \mathsf{sm}' \leftarrow \mathsf{S}'(\mathsf{crs}', m_0, m_1, \mathsf{rm}') \end{array} \right\}_{n,\mathsf{crs}',m_0,m_1} \approx_s \left\{ \mathsf{sm}' \;\middle|\; \begin{array}{l} \mathsf{rm}' \leftarrow \mathsf{R}^*(\mathsf{crs}') \\ b \leftarrow \mathsf{OTExt}'(\mathsf{crs}', \mathsf{rm}') \\ \mathsf{sm}' \leftarrow \mathsf{S}'(\mathsf{crs}', m_b, m_b, \mathsf{rm}') \end{array} \right\}_{n,\mathsf{crs}',m_0,m_1},$$

where $n \in \mathbb{N}, \mathsf{crs}' \in \{0,1\}^{\rho^2} \setminus \mathsf{B}', m_0, m_1 \in \{0,1\}$. This proves sender statistical privacy.

It is left to note that if $\mathsf{B} \in \mathbf{NP}$ then also $\mathsf{B}' \in \mathbf{NP}$. Indeed a witness for membership of $(\mathsf{scrs}_1, \ldots, \mathsf{scrs}_\rho)$ in B' is a string $\mathsf{rcrs} \in \{0,1\}^\rho$, and witnesses $w_1, \ldots, w_\rho$ such that each w_i is a witness for the fact that $\mathsf{scrs}_i \oplus \mathsf{rcrs} \in \mathsf{B}$. □

4.2 From the SRS Model to the Plain Model

In this section, we show how to transform any two-message SSP OT in the sender random string model that has CRS-free correctness and bad-CRS certification into one in the plain model. We do this assuming the existence of hitting set generators (HSGs) against co-non-deterministic uniform algorithms, which are in turn known from worst-case uniform assumption commonly used for derandomizing **AM**. Similar (or even stronger) assumptions have become rather common in in the cryptographic literature. (See more details in Section 2.2.)

In what follows, we denote the original protocol by (S, R), its CRS length by ρ, and corresponding bad CRS set by B. Let D_{B} be the co-non-deterministic decider that outputs 0 on every $x \in \mathsf{B}$ and 1 on $x \notin \mathsf{B}$, and let $t(m) = m^{O(1)}$ be its running time. Also let H be hitting-set generator against co-non-deterministic uniform algorithms. We construct a new protocol $(\mathsf{S}', \mathsf{R}')$. The transformation is presented in Figure 3.

Protocol 3

$\mathsf{R}'(c \in \{0,1\})$		$\mathsf{S}'(m_0, m_1 \in \{0,1\})$
$(\mathsf{crs}_1, \ldots, \mathsf{crs}_k) \leftarrow H(1^\rho, 1^{t(\rho)})$		
$\forall i \in [k] : (\mathsf{rm}_i, \mathsf{sk}_i) \leftarrow \mathsf{R.Enc}(c, \mathsf{crs}_i).$		
	$\xrightarrow{\mathsf{rm}' := (\mathsf{rm}_1, \ldots, \mathsf{rm}_k)}$	$\forall b \in \{0,1\}:$
		$m_b^1, \ldots, m_b^{k-1} \leftarrow \{0,1\},$
		$m_b^k := m_b^1 \oplus \ldots \oplus m_b^{k-1} \oplus m_b.$
		$(\mathsf{crs}_1, \ldots, \mathsf{crs}_k) \leftarrow H(1^\rho, 1^{t(\rho)}),$
		$\forall i \in [k] :$
	$\xleftarrow{\mathsf{sm}' := (\mathsf{sm}_1, \ldots, \mathsf{sm}_k)}$	$\mathsf{sm}_i \leftarrow \mathsf{S}(\mathsf{crs}_i, m_0^i, m_1^i, \mathsf{rm}_i).$
$\forall i \in [k]:$		
$m^i \leftarrow \mathsf{R.Dec}(\mathsf{crs}_i, \mathsf{sk}_i, \mathsf{sm}_i),$		
Output $m^1 \oplus \ldots \oplus m^k$.		

Figure 3: Two-message statistically-sender-private OT in the plain model

We prove:

Theorem 4.6. *Assuming (S, R) is a two-message statistically sender-private OT in SRS model with CRS-free correctness and bad-CRS certification, then $(\mathsf{S}', \mathsf{R}')$ is a two-message statistically sender-private OT in plain model.*

Corollary 4.7. *Under the $LPN_{\frac{\log^2(n)}{n}}$ assumption, and the existence of hitting-set generators against co-nondeterministic algorithms, there exists a two-message statically-sender-private OT protocol in the plain model.*

To prove Theorem 4.6, we prove each of the required properties.

Proposition 4.8. *Protocol $(\mathsf{S}', \mathsf{R}')$ is correct.*

Proof. By the CRS-free correctness of the underlying protocol, correctness holds for each crs_i output by the hitting set generator H. It follows that except with negligible probability $n^{-\omega(1)}$ over the coins of $(\mathsf{S}', \mathsf{R}')$, the receiver learns all m_c^i and $m_c = \oplus_i m_c^i$. □

Proposition 4.9. *Protocol* $(\mathsf{S}', \mathsf{R}')$ *satisfies receiver privacy.*

Proof. Recall that the underlying protocol is secure in the SRS model, implying that receiver privacy holds for any choice of CRS. In particular, it holds with respect to each crs_i output by H. Receiver privacy follows by a straightforward hybrid argument. □

Proposition 4.10. *Protocol* $(\mathsf{S}', \mathsf{R}')$ *is statistically-sender-private.*

Proof. Since B has negligible density, and in particular density at most $1/2$, the hitting set generator H outputs at least one string $\mathsf{crs}_i \notin \mathsf{B}$. The rest of the proof is identical to that of proposition 4.5. □

References

[AIR01] William Aiello, Yuval Ishai, and Omer Reingold. Priced oblivious transfer: How to sell digital goods. In Birgit Pfitzmann, editor, *Advances in Cryptology - EUROCRYPT 2001, International Conference on the Theory and Application of Cryptographic Techniques, Innsbruck, Austria, May 6-10, 2001, Proceeding*, volume 2045 of *Lecture Notes in Computer Science*, pages 119–135. Springer, 2001.

[AJ17] Prabhanjan Ananth and Abhishek Jain. On secure two-party computation in three rounds. In Yael Kalai and Leonid Reyzin, editors, *Theory of Cryptography - 15th International Conference, TCC 2017, Baltimore, MD, USA, November 12-15, 2017, Proceedings, Part I*, volume 10677 of *Lecture Notes in Computer Science*, pages 612–644. Springer, 2017.

[Ale03] Michael Alekhnovich. More on average case vs approximation complexity. In *44th Symposium on Foundations of Computer Science (FOCS 2003), 11-14 October 2003, Cambridge, MA, USA, Proceedings*, pages 298–307. IEEE Computer Society, 2003.

[Ban93] W. Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(4):625–636, 1993.

[BD18] Zvika Brakerski and Nico Döttling. Two-message statistically sender-private OT from LWE. In Amos Beimel and Stefan Dziembowski, editors, *Theory of Cryptography - 16th International Conference, TCC 2018, Panaji, India, November 11-14, 2018, Proceedings, Part II*, volume 11240 of *Lecture Notes in Computer Science*, pages 370–390. Springer, 2018.

[BDK+19] Marshall Ball, Dana Dachman-Soled, Mukul Kulkarni, Huijia Lin, and Tal Malkin. Non-malleable codes against bounded polynomial time tampering. In Yuval Ishai and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2019 - 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Darmstadt, Germany, May 19-23, 2019, Proceedings, Part I*, volume 11476 of *Lecture Notes in Computer Science*, pages 501–530. Springer, 2019.

[BFKL93] Avrim Blum, Merrick L. Furst, Michael J. Kearns, and Richard J. Lipton. Cryptographic primitives based on hard learning problems. In Douglas R. Stinson, editor, *Advances in Cryptology - CRYPTO '93, 13th Annual International Cryptology Conference, Santa Barbara, California, USA, August 22-26, 1993, Proceedings*, volume 773 of *Lecture Notes in Computer Science*, pages 278–291. Springer, 1993.

[BGI+17] Saikrishna Badrinarayanan, Sanjam Garg, Yuval Ishai, Amit Sahai, and Akshay Wadia. Two-message witness indistinguishability and secure computation in the plain model from new assumptions. In Tsuyoshi Takagi and Thomas Peyrin, editors, *Advances in Cryptology - ASIACRYPT 2017 - 23rd International Conference on the Theory and Applications of Cryptology and Information Security, Hong Kong, China, December 3-7, 2017, Proceedings, Part III*, volume 10626 of *Lecture Notes in Computer Science*, pages 275–303. Springer, 2017.

[BGJ+17] Saikrishna Badrinarayanan, Vipul Goyal, Abhishek Jain, Dakshita Khurana, and Amit Sahai. Round optimal concurrent MPC via strong simulation. In Yael Kalai and Leonid Reyzin, editors, *Theory of Cryptography - 15th International Conference, TCC 2017, Baltimore, MD, USA, November 12-15, 2017, Proceedings, Part I*, volume 10677 of *Lecture Notes in Computer Science*, pages 743–775. Springer, 2017.

[BGJ+18] Saikrishna Badrinarayanan, Vipul Goyal, Abhishek Jain, Yael Tauman Kalai, Dakshita Khurana, and Amit Sahai. Promise zero knowledge and its applications to round optimal MPC. In Hovav Shacham and Alexandra Boldyreva, editors, *Advances in Cryptology - CRYPTO 2018 - 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2018, Proceedings, Part II*, volume 10992 of *Lecture Notes in Computer Science*, pages 459–487. Springer, 2018.

[BGSZ21] James Bartusek, Sanjam Garg, Akshayaram Srinivasan, and Yinuo Zhang. Reusable two-round MPC from LPN. *IACR Cryptol. ePrint Arch.*, page 316, 2021.

[BKM20] Zvika Brakerski, Venkata Koppula, and Tamer Mour. NIZK from LPN and trapdoor hash via correlation intractability for approximable relations. In Daniele Micciancio and Thomas Ristenpart, editors, *Advances in Cryptology - CRYPTO 2020 - 40th Annual International Cryptology Conference, CRYPTO 2020, Santa Barbara, CA, USA, August 17-21, 2020, Proceedings, Part III*, volume 12172 of *Lecture Notes in Computer Science*, pages 738–767. Springer, 2020.

[BKP19] Nir Bitansky, Dakshita Khurana, and Omer Paneth. Weak zero-knowledge beyond the black-box barrier. In Moses Charikar and Edith Cohen, editors, *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 1091–1102. ACM, 2019.

[BKW03] Avrim Blum, Adam Kalai, and Hal Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. *J. ACM*, 50(4):506–519, 2003.

[BLP+13] Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, pages 575–584. ACM, 2013.

[BLSV18] Zvika Brakerski, Alex Lombardi, Gil Segev, and Vinod Vaikuntanathan. Anonymous ibe, leakage resilience and circular security from new assumptions. In Jesper Buus Nielsen and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part I*, volume 10820 of *Lecture Notes in Computer Science*, pages 535–564. Springer, 2018.

[BLVW19] Zvika Brakerski, Vadim Lyubashevsky, Vinod Vaikuntanathan, and Daniel Wichs. Worst-case hardness for LPN and cryptographic hashing via code smoothing. In Yuval Ishai and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2019 - 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Darmstadt, Germany, May 19-23, 2019, Proceedings, Part III*, volume 11478 of *Lecture Notes in Computer Science*, pages 619–635. Springer, 2019.

[BOV07] Boaz Barak, Shien Jin Ong, and Salil P. Vadhan. Derandomization in cryptography. *SIAM J. Comput.*, 37(2):380–400, 2007.

[BV11] Zvika Brakerski and Vinod Vaikuntanathan. Efficient fully homomorphic encryption from (standard) LWE. In Rafail Ostrovsky, editor, *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011*, pages 97–106. IEEE Computer Society, 2011.

[BV16] Nir Bitansky and Vinod Vaikuntanathan. Indistinguishability obfuscation: From approximate to exact. In Eyal Kushilevitz and Tal Malkin, editors, *Theory of Cryptography - 13th International Conference, TCC 2016-A, Tel Aviv, Israel, January 10-13, 2016, Proceedings, Part I*, volume 9562 of *Lecture Notes in Computer Science*, pages 67–95. Springer, 2016.

[BV17] Nir Bitansky and Vinod Vaikuntanathan. A note on perfect correctness by derandomization. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *Advances in Cryptology - EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Paris, France, April 30 - May 4, 2017, Proceedings, Part II*, volume 10211 of *Lecture Notes in Computer Science*, pages 592–606, 2017.

[DGH+19] Nico Döttling, Sanjam Garg, Mohammad Hajiabadi, Daniel Masny, and Daniel Wichs. Two-round oblivious transfer from CDH or LPN. *IACR Cryptol. ePrint Arch.*, page 414, 2019.

[DN00] Cynthia Dwork and Moni Naor. Zaps and their applications. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*, pages 283–293. IEEE Computer Society, 2000.

[EGL85] Shimon Even, Oded Goldreich, and Abraham Lempel. A randomized protocol for signing contracts. *Commun. ACM*, 28(6):637–647, 1985.

[Gen09] Craig Gentry. Fully homomorphic encryption using ideal lattices. In Michael Mitzenmacher, editor, *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 169–178. ACM, 2009.

[GST03] Dan Gutfreund, Ronen Shaltiel, and Amnon Ta-Shma. Uniform hardness vs. randomness tradeoffs for arthur-merlin games. In *18th Annual IEEE Conference on Computational Complexity (Complexity 2003), 7-10 July 2003, Aarhus, Denmark*, pages 33–47. IEEE Computer Society, 2003.

[HK12] Shai Halevi and Yael Tauman Kalai. Smooth projective hashing and two-message oblivious transfer. *J. Cryptol.*, 25(1):158–193, 2012.

[HNY17] Pavel Hubácek, Moni Naor, and Eylon Yogev. The journey from NP to TFNP hardness. In Christos H. Papadimitriou, editor, *8th Innovations in Theoretical Computer Science Conference, ITCS 2017, January 9-11, 2017, Berkeley, CA, USA*, volume 67 of *LIPIcs*, pages 60:1–60:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.

[JKKR17] Abhishek Jain, Yael Tauman Kalai, Dakshita Khurana, and Ron Rothblum. Distinguisher-dependent simulation in two rounds and its applications. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part II*, volume 10402 of *Lecture Notes in Computer Science*, pages 158–189. Springer, 2017.

[Khu17] Dakshita Khurana. Round optimal concurrent non-malleability from polynomial hardness. In Yael Kalai and Leonid Reyzin, editors, *Theory of Cryptography - 15th International Conference, TCC 2017, Baltimore, MD, USA, November 12-15, 2017, Proceedings, Part II*, volume 10678 of *Lecture Notes in Computer Science*, pages 139–171. Springer, 2017.

[KKS18] Yael Tauman Kalai, Dakshita Khurana, and Amit Sahai. Statistical witness indistinguishability (and more) in two messages. In Jesper Buus Nielsen and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part III*, volume 10822 of *Lecture Notes in Computer Science*, pages 34–65. Springer, 2018.

[KPC+11] Eike Kiltz, Krzysztof Pietrzak, David Cash, Abhishek Jain, and Daniele Venturi. Efficient authentication from hard learning problems. In Kenneth G. Paterson, editor, *Advances in Cryptology - EUROCRYPT 2011 - 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tallinn, Estonia, May 15-19, 2011. Proceedings*, volume 6632 of *Lecture Notes in Computer Science*, pages 7–26. Springer, 2011.

[KS17] Dakshita Khurana and Amit Sahai. How to achieve non-malleability in one or two rounds. In Chris Umans, editor, *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 564–575. IEEE Computer Society, 2017.

[KSS10] Jonathan Katz, Ji Sun Shin, and Adam D. Smith. Parallel and concurrent security of the HB and hb$^+$ protocols. *J. Cryptol.*, 23(3):402–421, 2010.

[Lau83] Clemens Lautemann. BPP and the polynomial hierarchy. *Inf. Process. Lett.*, 17(4):215–217, 1983.

[Mat93] Mitsuru Matsui. Linear cryptanalysis method for DES cipher. In Tor Helleseth, editor, *Advances in Cryptology - EUROCRYPT '93, Workshop on the Theory and Application of of Cryptographic Techniques, Lofthus, Norway, May 23-27, 1993, Proceedings*, volume 765 of *Lecture Notes in Computer Science*, pages 386–397. Springer, 1993.

[MV99] Peter Bro Miltersen and N. V. Vinodchandran. Derandomizing arthur-merlin games using hitting sets. In *40th Annual Symposium on Foundations of Computer Science, FOCS '99, 17-18 October, 1999, New York, NY, USA*, pages 71–80. IEEE Computer Society, 1999.

[NP01] Moni Naor and Benny Pinkas. Efficient oblivious transfer protocols. In S. Rao Kosaraju, editor, *Proceedings of the Twelfth Annual Symposium on Discrete Algorithms, January 7-9, 2001, Washington, DC, USA*, pages 448–457. ACM/SIAM, 2001.

[OPP14] Rafail Ostrovsky, Anat Paskin-Cherniavsky, and Beni Paskin-Cherniavsky. Maliciously circuit-private FHE. In Juan A. Garay and Rosario Gennaro, editors, *Advances in Cryptology - CRYPTO 2014 - 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part I*, volume 8616 of *Lecture Notes in Computer Science*, pages 536–553. Springer, 2014.

[Pei09] Chris Peikert. Public-key cryptosystems from the worst-case shortest vector problem: extended abstract. In Michael Mitzenmacher, editor, *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 333–342. ACM, 2009.

[Pie12] Krzysztof Pietrzak. Cryptography from learning parity with noise. In Mária Bieliková, Gerhard Friedrich, Georg Gottlob, Stefan Katzenbeisser, and György Turán, editors, *SOFSEM 2012: Theory and Practice of Computer Science - 38th Conference on Current Trends in Theory and Practice of Computer Science, Špindlerův Mlýn, Czech Republic, January 21-27, 2012. Proceedings*, volume 7147 of *Lecture Notes in Computer Science*, pages 99–114. Springer, 2012.

[Rab05] Michael O. Rabin. How to exchange secrets with oblivious transfer. *IACR Cryptol. ePrint Arch.*, page 187, 2005.

[Reg05] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In Harold N. Gabow and Ronald Fagin, editors, *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, pages 84–93. ACM, 2005.

[YZ21] Yu Yu and Jiang Zhang. Smoothing out binary linear codes and worst-case sub-exponential hardness for LPN. In Tal Malkin and Chris Peikert, editors, *Advances in Cryptology - CRYPTO 2021 - 41st Annual International Cryptology Conference, CRYPTO 2021, Virtual Event, August 16-20, 2021, Proceedings, Part III*, volume 12827 of *Lecture Notes in Computer Science*, pages 473–501. Springer, 2021.

[YZW+17] Yu Yu, Jiang Zhang, Jian Weng, Chun Guo, and Xiangxue Li. Collision resistant hashing from learning parity with noise. *IACR Cryptol. ePrint Arch.*, page 1260, 2017.