Additive Conditional Disclosure of Secrets and Applications

***** Draft. 30.05.2005 *****

Sven Laur[1] and Helger Lipmaa[2]

[1] Helsinki University of Technology, Finland
[2] Cybernetica AS and University of Tartu, Estonia

Abstract. The main goal of this paper is to study the power of additively homomorphic public-key cryptosystems, and in particular, to study what can be achieved under the sole assumption that such a cryptosystem is IND-CPA secure. We extend the conditional disclosure of secrets transformation to work over additively homomorphic public-key cryptosystems, and using this, construct several novel protocols that are all one-round, computationally receiver-private (assuming that the underlying cryptosystem is IND-CPA secure and sufficiently rough) and statistically sender-private, in the complexity-theoretic model against a malicious adversary. In particular, we propose an oblivious transfer protocol with log-squared communication, a millionaire's protocol with logarithmic communication, and a few simpler protocol for tasks from linear algebra and privacy-preserving data mining. We hope that by presenting our results and in particular, our real applications, we can help to popularise the notion of conditional disclosure of secrets that has, unfortunately, until now been studied only in an handful of papers.
Keywords. Conditional disclosure of secrets, homomorphic encryption, malicious model, millionaire's problem, oblivious transfer, two-party computation.

1 Introduction

A relatively large family of well-known cryptographic protocols can be jointly described as follows: Receiver generates a secret key and public key pair for an IND-CPA secure additively homomorphic public-key cryptosystem. He sends the public key to Sender, and possibly proves its correctness. Later, during the protocol, he transfers some encrypted values to Sender. Sender performs a number of protocol-dependent operations on the ciphertexts and returns the resulting ciphertexts to Receiver. Some noteworthy tasks that can be efficiently solved by using such *additively homomorphic one-round protocols* include computationally-private information retrieval [AIR01,Ste98,Lip04], millionaire's problem [BK04], and various tasks of linear algebra (e.g., private matrix multiplication) and privacy-preserving data mining (e.g., private scalar product and private set intersection cardinality).

Additively homomorphic one-round protocols can usually be shown to be computationally receiver-private in the complexity-theoretic model, given the IND-CPA security of the underlying public-key cryptosystem PKC. Statistical sender-privacy and correctness (e.g., that the parties know their inputs) are usually guaranteed by using zero-knowledge proofs and proofs of knowledge. However, in practice, correctness is often left out from the security requirements; such a "relaxed" security definition that only considers privacy is *de facto* standard in the case of computationally-private information retrieval, oblivious transfer and oblivious keyword search protocols [NP01,AIR01,Lip04,FIPR05]. Such protocols are statistically sender-private in the complexity-theoretic model without using zero-knowledge proofs (of knowledge).

As the first contribution of this paper, we propose precise privacy definitions for additively homomorphic one-round protocols. As done in the mentioned papers on oblivious transfer, we are only interested in the privacy (more precisely, in computational receiver-privacy and statistical sender-privacy). Moreover, we require that an additively homomorphic one-round protocol remains private even under a concurrent execution with (a restricted/polynomial number of) other additively homomorphic one-round protocols, even if the same key pair of Receiver is used in all such protocols. This is a strictly stronger requirement than the one given say in [Gol04], but in our setting, it holds almost straightforwardly in the case of one-round protocols. Proven privacy in such a multi-user setting enables one to perform the possibly costly key generation and key correctness verification phase only once, in the initialisation phase; the same key can then used many times in many different protocols.

According to our definitions, computational receiver-privacy of an additively homomorphic one-round protocol in the malicious model follows directly from the IND-CPA security of the underlying public-key cryptosystem. On the other hand, well-known additively homomorphic one-round protocols are often statistically sender-private only in the semi-honest model, that is, when Receiver encrypts correct values. Often, there is no guarantee of

sender-privacy whatsoever in the case of malicious Receiver and in some protocols, a malicious Receiver can retrieve Sender's input during a single protocol run.

To make additively homomorphic one-round protocols private in the malicious model, one must guarantee that Receiver obtains only the "required amount of information" even if he encrypts invalid inputs. To the best of our knowledge, the only transformation that transforms one-round protocols, secure in the semi-honest model, into one-round protocols, secure in the malicious model, and works in the complexity-theoretic model is the *conditional disclosure of secrets* (CDS) transformation of Gertner, Ishai, Kushilevitz and Malkin [GIKM00]. In the case of a single server (that we are interested in this paper), the first CDS transformation was proposed in [AIR01]. According to [AIR01], a *conditional disclosure of secrets* ($\mathsf{CDS}_d^{\mathcal{S}}$) protocol for a set $\mathcal{S}$ and for some d is a two-party protocol, at the end of what Receiver obtains a secret $t \in \mathbb{Z}_d$ specified by Sender only if Receiver's private input ϱ belongs to $\mathcal{S}$, and "non-relevant" information, otherwise. As briefly mentioned in [AIR01], one can use a suitable additively homomorphic one-round protocol for $\mathsf{CDS}_d^{\mathcal{S}}$ to transform a large class of additively homomorphic one-round protocols, sender-private in the semi-honest model, to additively homomorphic one-round protocols, sender-private against malicious adversaries in the complexity-theoretic model. The resulting CDS transformation is especially efficient with protocols where the correctness of Receiver's inputs can be verified publicly and in particular without the knowledge of Receiver's secret key. We call such protocols *conventional*; almost all additively homomorphic one-round protocols in the literature are conventional.

The CDS transformation of Aiello, Ishai and Reingold works only in conjunction with an IND-CPA secure homomorphic cryptosystem PKC that has plaintext space of prime order and where Sender can verify whether a public key is correct without any interaction with Receiver. The only widely known such cryptosystem, ElGamal [El 84], is multiplicatively homomorphic modulo a prime n, while many real-life protocols require PKC to be additively homomorphic modulo some integer n. In the case of all widely-known additively homomorphic cryptosystems, n is a large composite integer with large prime factors.

We modify the CDS transformation so that it can be used in conjunction with a PKC that satisfies substantially weaker properties; in particularly, it has to be IND-CPA secure, and the smallest prime factor $\Phi(n)$ of n should not be too small. Let us call such PKC *CDS-friendly*; as an example, the cryptosystems from [Pai99,DJ01] are CDS-friendly. Our construction consists of three steps. First, we design an additively homomorphic one-round 1-out-of-N-oblivious transfer protocol for ℓ-bit strings, with $\ell := \lfloor \log_2 \Phi(n) - \log_2 N - \lambda + 1 \rfloor$, where N is the number of Sender's inputs and $2^{-\lambda}$ is the desired security level for Sender, that is computationally receiver-private and statistically sender-private assuming that the underlying additively homomorphic public-key cryptosystem is CDS-friendly. In this protocol, honest Receiver's first message is just an homomorphic encryption of the database index. This protocol is the first oblivious transfer protocol with such properties and therefore interesting in its own right. We then show how to transform Lipmaa's recent 1-out-of-N-computationally-private information retrieval protocol [Lip04], with communication $\Theta(k \cdot \log^2 N + \ell \cdot \log N)$, into an additively homomorphic one-round statistically sender-private oblivious transfer protocol for ℓ-bit strings with a minimal increase in the communication, given that PKC is a CDS-friendly length-flexible additively homomorphic public-key cryptosystem; the communication increases to $\Theta(\sqrt{\log N} \cdot 2^{\sqrt{\log N}} \cdot \ell \cdot k)$ if PKC is not length-flexible [Ste98]. This is the first such transformation that works under the sole complexity-theoretic assumption that PKC is IND-CPA secure. Here, k is the security parameter (a constant or a polylogarithmic value in $\ell \cdot N$, depending on the security model).

Second, we propose an additively homomorphic one-round protocol for $\mathsf{CDS}_{2^\ell}^{\mathcal{S}}$ that works with any CDS-friendly additively-homomorphic public-key cryptosystem. For this, we use the new oblivious transfer protocol. This results in the polylogarithmic communication and linear computation in $\sharp\mathcal{S}$ for any set $\mathcal{S}$, given that PKC is a CDS-friendly length-flexible additively homomorphic public-key cryptosystem; again, the communication increases to quasipolynomial if PKC is not length-flexible. We use arithmetic circuit evaluation to reduce the communication and computational complexity of the $\mathsf{CDS}_{2^\ell}^{\mathcal{S}}$ protocol for some specific interesting sets $\mathcal{S}$.

Third, we propose the *conditional disclosure of secrets (CDS) transformation* that transforms an arbitrary conventional additively homomorphic one-round protocol Π (where the valid input set Valid has an efficient $\mathsf{CDS}_d^{\mathsf{Valid}}$ protocol), secure in the semi-honest model, to a protocol that is secure in the malicious model. The basic idea is as in [AIR01]: in parallel with the original protocol, Receiver and Sender execute a conditional disclosure of secrets protocol for Receiver's input. Sender masks the output with secrets, corresponding to all different Receiver's inputs. As a result, Receiver will obtain any of the outputs only if his all inputs belong to the valid input sets. The resulting protocol is secure in the malicious model, has one round and is—in many cases, interesting in practice—surprisingly computation and communication efficient. The only security assumption is that the underlying public-key cryptosystem is IND-CPA secure. The CDS transformation is efficient whenever the valid input set $\mathcal{S}$ has an efficient conditional disclosure of secrets protocol, and the number of outputs L is reasonably small. If we require only computational sender-privacy then the communication can be reduced almost L times. We de-

fine rigorously all the needed primitives and propose corresponding security proofs in the language of concrete security.

Until now, the CDS transformation has been overlooked until now by most of the researchers, with only a couple of published papers that do more than mention it and with most of the contemporary papers using zero-knowledge proofs in a situation where the CDS transformation would provide a simpler solution. This unfortunate situation might be partially due to the relatively small number of applications proposed for this transformation. To remedy this situation and to popularise the CDS transformation, throughout this paper, we propose several interesting applications to demonstrate the power of the new tools. We hope that they will motivate further studies in this area.

First, based on Lipmaa's recent log-squared computationally-private information retrieval protocol [Lip04], we construct an oblivious transfer protocol and a conditional disclosure of secrets protocol with log-squared communication, for a length-flexible PKC, and with quasi-polylogarithmic communication, for any PKC. (See Cor. 2.) Second, we construct a private millionaire's protocol with logarithmic communication. (See Cor. 4.) Third, we construct efficient private protocols for a few other tasks like scalar product and set intersection cardinality. (See Sect. 6.) All constructed protocols are one-round, computationally receiver-private and statistically sender-private solely under the assumption that the underlying additively homomorphic public-key cryptosystem is CDS-friendly. Note also that the first two protocols do not directly use the generic CDS transformation but related techniques from this paper.

Road-map. In Section 2, we give preliminaries. In Section 3, we define additively homomorphic one-round protocols and their security. In Section 4, we propose a new additively homomorphic one-round protocol for oblivious transfer and prove its security. In Section 5, we propose an additively homomorphic one-round protocol for conditional disclosure of secrets and show how to implement it efficiently for many interesting sets. In Section 6, we present our generic CDS transform and prove its security. In Appendix D, we discuss about the optimality of our padding method.

2 Preliminaries

For an integer s, let $[s] := \{1, 2, \ldots, s\}$. For an integer n, let $\Phi(n)$ be the smallest prime divisor of n. We say that n is *p-rough* if $\Phi(n) \geq p$. The statistical difference of two distributions X and Y over a discrete support Z is defined as $\mathsf{Dist}\,(X\|Y) := \max_{S\subseteq Z} |\Pr[X \in S] - \Pr[Y \in S]|$. For an arbitrary set Z, let $U(Z)$ denote the uniform distribution over it; we sometimes identify Z with $U(Z)$. A set Z with a binary operation $\circ : Z^2 \to Z$ is a quasigroup iff $\forall a, b \in Z$ there exist unique $x, y \in Z$ such that $ax = b$ and $ya = b$. In particular, for every $a \in Z$, $a \circ U(Z) = U(Z)$.

Public-key cryptosystem is a triple $\mathsf{PKC} = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$, where Gen is a key generation algorithm that returns a secret and public key pair $(\mathsf{sk}, \mathsf{pk})$, Enc is a randomized encryption algorithm and Dec is a decryption algorithm with the usual syntax. For fixed PKC and for a fixed public key pk, let $\mathcal{R}$ be the randomness space, let $\mathcal{M}$ be the plaintext space and let $\mathcal{C}$ be the ciphertext space.

For an adversary A, define $\mathsf{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathsf{PKC}}(A) := 2 \cdot |\Pr[(\mathsf{sk}, \mathsf{pk}) \leftarrow \mathsf{Gen}, \quad (m_0, m_1) \leftarrow A(\mathsf{pk}), b \leftarrow U(\mathbb{Z}_2) : A(\mathsf{pk}, m_0, m_1, \mathsf{Enc}_{\mathsf{pk}}(m_b; U(\mathcal{R}))) = b] - \frac{1}{2}|$, where the probability is taken over the coin tosses of Gen and A, and over the choice of random variables. We say that PKC is (ε, τ)-*IND-CPA-secure* if $\mathsf{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathsf{PKC}}(A) \leq \varepsilon$ for any probabilistic algorithm A that works in time τ.

A public-key cryptosystem PKC is *homomorphic*, if for any key pair $(\mathsf{sk}, \mathsf{pk})$, any $x_1, x_2 \in \mathcal{M}$ and any $r_1, r_2 \in \mathcal{R}$, $\mathsf{Enc}_{\mathsf{pk}}(x_1; r_1) \cdot \mathsf{Enc}_{\mathsf{pk}}(x_2; r_2) = \mathsf{Enc}_{\mathsf{pk}}(x_1 + x_2; r_1 \circ r_2)$, where $+$ is a group operation in $\mathcal{M}$ and $\circ$ is a groupoid operation in $\mathcal{R}$. We say that PKC is *additively homomorphic* if $\mathcal{M} = \mathbb{Z}_n$ for some n, and *multiplicatively homomorphic*, if $\mathcal{M} = \mathbb{Z}^*_n$ for some n. Several homomorphic cryptosystems [El 84,OU98,NS98,Pai99,DJ01,DJ03] are IND-CPA secure under reasonable complexity assumptions; from these, the ElGamal cryptosystem [El 84] is multiplicatively homomorphic (and the only one where $\mathcal{M}$ has a prime order), while other cryptosystems (e.g., [Pai99,DJ01]) are additively homomorphic with a usually rough composite n.

The Paillier cryptosystem [Pai99] (as modified in [DJ01]) is one of the most efficient known IND-CPA secure public-key cryptosystems. Here, $\mathcal{M} = \mathbb{Z}_n$, $\mathcal{R} = \mathbb{Z}^*_n$ and $\mathcal{C} = \mathbb{Z}^*_{n^2}$ for an RSA modulus n. Thus, n is $\sqrt{n}/2$-rough. The Paillier cryptosystem is IND-CPA secure, assuming that the Decisional Composite Residuosity Problem (DCRP) is hard [Pai99], and additively homomorphic.

Later, we need an efficient zero-knowledge correctness proof KProof for pk. KProof can be omitted if the correctness of pk can be verified interactively (e.g., if pk is valid iff it is an element of a residue class ring). We require that if $(\mathsf{sk}, \mathsf{pk}) \in \mathsf{Gen}$ then in the case of an honest prover and verifier, the verifier accepts. We say that PKC

is ε_{zk}-*sound* if for any invalid public key, the probability that KProof accepts is less than ε_{zk}. Within this paper, we will assume that KProof is computationally zero-knowledge with perfect hiding; this considerably simplifies the presentation. For example, if KProof is imperfectly hiding, we will have to talk about colliding Senders who pool together their advantages they get by observing their runs of KProof. Our results however also work in the general case but then the proofs become more involved. (See, e.g., App. B.)

3 Additively Homomorphic One-Round Protocols and Their Security

Let ϱ denote the private input of Receiver and σ denote the private input of Sender. In this paper, we consider one-round (i.e., two-message) protocols between Receiver and Sender that implement the following functionality for some public function f: an unbounded Receiver learns $f(\varrho, \sigma)$ and nothing more, and a computationally bounded Sender learns no new information. More precisely, we are interested in *additively homomorphic one-round protocols* that consist of the next phases. In the initialisation phase (usually shared by different protocols), Receiver generates his key pair $(\mathsf{sk}, \mathsf{pk})$ for PKC by executing Gen. After that, Receiver (as a prover) and an arbitrary interested Sender (as a verifier) invoke the KProof protocol. Sender halts when the proof is incorrect, and Receiver can halt when Sender behaves maliciously. Sender either accepts or rejects the key proof. After running Gen and KProof, the same key pair $(\mathsf{sk}, \mathsf{pk})$ can be reused—possibly, in parallel—in many (although a restricted number of) instantiations of possibly different protocols with possibly different Senders. Since KProof is executed rarely, it can be relatively complex and thus can be perfectly hiding.

A concrete additively homomorphic one-round protocol Π is specified by a triple of efficient algorithms $(\mathsf{Query}, \mathsf{Transfer}, \mathsf{Recover})$, and by three efficiently samplable distributions $\mathcal{R}_Q$, $\mathcal{R}_T$ and $\mathcal{R}_R$. In the first message of a protocol instantiation, Receiver sends a randomised message $\mathsf{msgq} \leftarrow \mathsf{Query}_{\mathsf{pk}}(\varrho; U(\mathcal{R}_Q))$, for some $\varrho \in \mathcal{M}_Q$, to Sender. We additionally assume that one can efficiently verify, given only pk, that $\mathsf{msgq} \in \mathsf{Query}_{\mathsf{pk}}(\mathcal{M}_Q; \mathcal{R}_Q)$. In our setting, msgq is a tuple of ciphertexts, all encrypted under the same key, and thus this verification can be done efficiently given that membership in $\mathcal{C}$ can be tested efficiently. In the second message, Sender replies with $\mathsf{msgt} \leftarrow \mathsf{Transfer}_{\mathsf{pk}}(\sigma, \mathsf{msgq}; U(\mathcal{R}_T))$. In our case, this means that Sender applies a number of randomized operations to the received ciphertexts, and returns the resulting ciphertexts. We assume that $\mathsf{msgt} = \perp$ if Sender does not have the public key, Sender halts or msgq is malformed. Finally, Receiver recovers the answer by computing $\mathsf{Recover}_{\mathsf{sk}}(\varrho, \mathsf{msgt}; U(\mathcal{R}_R))$. In our setting, this means that he decrypts the received ciphertexts, and applies some local algorithm to the resulting plaintexts. We call such a protocol an *additively homomorphic one-round protocol for* f. The *communication* of an additively homomorphic one-round protocol is equal to $|\mathsf{msgq}| + |\mathsf{msgt}|$.

We say that Π is (ε, τ)-*receiver-private* in the malicious model, if for any adversary A with the working time τ,

$$\mathsf{Adv}^{\mathsf{RECPRI}}_{\Pi}(A) := 2 \cdot \max \left| \Pr \left[\begin{array}{l} (\mathsf{sk}, \mathsf{pk}) \leftarrow \mathsf{Gen}, \mathsf{R} \text{ and } A \text{ run } \mathsf{KProof}, \\ b \leftarrow U(\mathbb{Z}_2), \mathsf{msgq} \leftarrow \mathsf{Query}_{\mathsf{pk}}(\varrho_b; U(\mathcal{R}_Q)) : \\ A(\mathsf{pk}, \varrho_0, \varrho_1, \mathsf{msgq}) = b \end{array} \right] - \frac{1}{2} \right| \leq \varepsilon \ ,$$

Here, probability is taken over the coin tosses of Gen, KProof, Query, A, and over the choice of random variables. The maximum is taken over all possible inputs (ϱ_0, ϱ_1). If KProof is perfectly hiding then it can be omitted from the definition. Privacy in the semi-honest model is defined as usually.

For defining sender-privacy, we note that in our case, there exists a function $\mathsf{Extract}$, such that $\mathsf{Extract}_{\mathsf{sk}}(\mathsf{Query}_{\mathsf{pk}}(\varrho; r)) = \varrho$ for every $\varrho \in \mathcal{M}_Q$ and $r \in \mathcal{R}_Q$; $\mathsf{Extract}$ just decrypts all ciphertexts in $\mathsf{Query}_{\mathsf{pk}}(\cdot; \cdot)$. We require that the only piece of new information that a potentially malicious Receiver obtains by running the protocol is $f(\varrho^*, \sigma)$, where $\varrho^* = \mathsf{Extract}_{\mathsf{sk}}(\mathsf{msgq})$ is his submitted input. The existence of $\mathsf{Extract}$ makes it easy to mix inputs from different protocols since then merging queries from different protocols yields a consistent output even if Receiver is malicious. More formally, we define the *sender-privacy* of Π by requiring that there exists an unbounded simulator Sim that for every unbounded receiver A, on access to A's first message msgq in the protocol, to A's random tape r_A, and to $f(\mathsf{Extract}_{\mathsf{sk}}(\mathsf{msgq}), \sigma)$, generates a output that is statistically indistinguishable from transcript of a protocol run between A and the honest Sender, given the same r_A. That is, we define

$$\mathsf{Adv}^{\mathsf{SENPRI}}_{\Pi, \mathsf{Sim}}(A) := \max_{\substack{(\mathsf{sk}, \mathsf{pk}) \in \mathsf{Gen} \\ (\varrho^*, \sigma)}} \mathsf{Dist}\left(\mathsf{Sim}_{\mathsf{pk}}(\varrho^*, f(\varrho^*, \sigma), r_A) \| (\mathsf{msgq}, \mathsf{Transfer}_{\mathsf{pk}}(\sigma, \mathsf{msgq}; U(\mathcal{R}_T)))\right) \ ,$$

where $\mathsf{msgq} \leftarrow A(\mathsf{pk}, \varrho^*; r_A)$, σ is the private input of Sender, and $\varrho^* \leftarrow \mathsf{Extract}_{\mathsf{sk}}(\mathsf{msgq})$. The probability is taken over the coin tosses of $\mathsf{Transfer}$ and A. We say that Π is *statistically ε'-sender-private* if for some unbounded simulator Sim and for every unbounded algorithm A, $\mathsf{Adv}^{\mathsf{SENPRI}}_{\Pi,\mathsf{Sim}}(A) \leq \varepsilon'$. Sender-privacy is said to be *perfect* if for some unbounded simulator Sim and for every unbounded algorithm A, $\mathsf{Adv}^{\mathsf{SENPRI}}_{\Pi,\mathsf{Sim}}(A) = 0$. We say that Π is $(\varepsilon, \tau; \varepsilon')$-*private* if it is (ε, τ)-receiver-private and ε'-sender-private.

Above, we omitted the security parameter k by assuming that the adversary works in time that is less than some fixed public constant τ, then also k is a constant. Sometimes, one needs security against adversaries that work in time, polynomial in the input size ν of the protocol Π. Then, k will depend on ν. More precisely, assume that the underlying computationally hard problem, with input n of size $\nu := \log_2 n$, can be broken in time $L_n[a, b] := \exp(a(\ln n)^b \cdot (\ln \ln n)^{1-b})$ for some $0 < b \leq 1$. To guarantee security against such *polynomial adversaries*, it is necessary that $L_n[a, b] = \omega(\nu^c)$ for every constant c, or that $k^b \cdot \ln^{1-b} k = \omega(\ln \nu)$. Omitting the logarithmic factor, we get that $k = \Omega(\ln^{1/b} \nu)$. For example, when basing a protocol on the Decisional Composite Residuosity Assumption with $b = 1/3$, we must assume that $k = \Omega(\log^{3-o(1)} \nu)$.

Computationally-private information retrieval and oblivious transfer. During a 1-*out-of-*N *computationally-private information retrieval* (CPIR^N_d) protocol for elements from $\mathbb{Z}_d$, Receiver fetches $\sigma[\varrho]$ from the database $\sigma = (\sigma[1], \ldots, \sigma[N])$, $\sigma[i] \in \mathbb{Z}_d$, so that a computationally bounded Sender does not know which entry Receiver is learning. In the following, we will also need the case where $\sigma = (\sigma[h])_{h \in \mathcal{S}}$ for an arbitrary public set $\mathcal{S} \subseteq \mathbb{Z}_n$, in this case we call the resulting protocol a $\mathsf{CPIR}^{\mathcal{S}}_d$ protocol. We assume that N (or $\mathcal{S}$) is public. With a few exceptions (for example, [Cha04]), all one-round computationally-private information retrieval protocols (e.g., [Ste98,AIR01,Lip04]) are additively homomorphic one-round protocols. The most efficient known CPIR^N_d protocol by Lipmaa [Lip04] is also an additively homomorphic one-round protocol. When based on the Damgård-Jurik length-flexible additively homomorphic public-key cryptosystem [DJ01], it has communication $(\log^2 N + (s + \frac{3}{2}) \cdot \log N + 1)k$, where $k := \log_2 n$ is the security parameter, and $s := \lceil \frac{1}{k} \cdot \log_2 d \rceil$. A CPIR^N_d protocol is a *computationally chooser-private and statistically sender-private* 1-*out-of-*N *oblivious transfer protocol for elements from* $\mathbb{Z}_d$ (an OT^N_d protocol) if also Sender's privacy is guaranteed; $\mathsf{OT}^{\mathcal{S}}_d$ protocols are defined analogously.

The next private additively homomorphic one-round protocol for 1-out-of-N-oblivious transfer was defined by Aiello, Ishai and Reingold [AIR01]. Assume that PKC is an IND-CPA secure homomorphic cryptosystem such that for all possible secret keys sk, $\mathcal{M}$ is a cyclic group with public prime order. To obtain the element $\sigma[\varrho]$, Receiver sends to Sender a random encryption $c = \mathsf{Enc}_{\mathsf{pk}}(\varrho; U(\mathcal{R}))$. For all $i \in [N]$, Sender replies with $c_i \leftarrow (c \cdot \mathsf{Enc}_{\mathsf{pk}}(-i; 0))^{U(\mathcal{M})} \cdot \mathsf{Enc}_{\mathsf{pk}}(\sigma[i]; U(\mathcal{R}))$. Receiver obtains $\sigma[\varrho] \leftarrow \mathsf{Dec}_{\mathsf{sk}}(c_\varrho)$. Unfortunately, the only well-known homomorphic public-key cryptosystem that works over message spaces of prime order, ElGamal, is multiplicative.

For our applications, we need a one-round oblivious transfer protocol that works on groups of composite order. In [Lip03,Cha04], the authors have tried to generalise the Aiello-Ishai-Reingold protocol correspondingly. Lipmaa [Lip03] claimed that the Aiello-Ishai-Reingold protocol is a "weakly" sender-private 1-out-of-N-oblivious transfer protocol, under the weakened assumption that n is N-rough; weak security meaning that a malicious Receiver will, even in the case of incorrect inputs, obtain information about exactly one database element. Lipmaa's proof is however faulty, because in the case of a malicious Receiver, the Aiello-Ishai-Reingold protocol leaks information about L database elements, where L is the number of different prime factors of n. Namely, assume that $n = \prod p_i^{\alpha_i}$ for different primes $p_1 < p_2 < \cdots < p_L$. If Receiver's input ϱ' is such that $\varrho' = \varrho_i \mod p_i$ for some mutually different values $\varrho_i \in [N]$, then Receiver can straightforwardly compute the value $\sigma[\varrho_i] \mod p_i$ even if $\varrho \notin [N]$. A receiver who knows how to factor n can therefore easily, by using the Chinese Remaindering Theorem, compute the required ϱ'. The same observation underlies, in a constructive way, Chang's 2-out-of-N-oblivious transfer protocol; [Cha04] actually proved that if n is a product of two safe primes then no more information than $\sigma[\varrho_i] \mod p_i$, $i \in [2]$, is revealed. Based on this observation, Chang [Cha04] proposed also a 1-out-of-N-oblivious transfer protocol; however, since there a honest Receiver has to encrypt values that depend on the secret key then according to our definition it is not an additively homomorphic protocol. In particular, it will be unusable in the additive CDS transformation.

Millionaire's problem. Yao's millionaire problem is as follows: given Receiver's private input ϱ and Sender's private input σ from some set $\mathbb{Z}_d$, decide whether $\varrho > \sigma$. Though there have been proposed numerous protocols for this problem (see, for example, [Fis01,BK04,ST04]), none of the proposals is completely satisfactory. For example, one of the most elegant previous millionaire's protocols, a one-round additively homomorphic one-round protocol proposed by Blake and Kolesnikov [BK04], is sender-private only in the semi-honest model, while a somewhat different protocol by [ST04] uses zero-knowledge proofs to achieve privacy in the malicious model.

PRIVATE INPUT: Receiver has $\varrho \in \{1,\ldots,N\}$ and Sender has a tuple $\sigma[1],\ldots,\sigma[N] \in \mathbb{Z}_{2^\ell}$.
PRIVATE OUTPUT: Receiver obtains $\sigma[\varrho]$.

$\mathsf{Query}_{\mathsf{pk}}(\varrho;\cdot)$: Set $\mathsf{Query}_{\mathsf{pk}}(\varrho;r) \leftarrow \mathsf{Enc}_{\mathsf{pk}}(\varrho;r)$ for $r \leftarrow U(\mathcal{R})$.
$\mathsf{Transfer}_{\mathsf{pk}}(\sigma,\mathsf{msgq};\cdot)$:
  If $\mathsf{msgq} \notin \mathcal{C}$ then return $\mathsf{Transfer}_{\mathsf{pk}}(\mathsf{msgq},\sigma;\cdot) \leftarrow \perp$.
  For $j \in [N]$
    Set $\mathsf{padded}(\sigma[j]) \leftarrow \sigma[j] + 2^\ell \cdot z_j$, where $z_j \leftarrow U(\mathbb{Z}_{T_j})$ for $T_j \leftarrow \lfloor (n-\sigma[j]-1)\cdot 2^{-\ell} \rfloor$.
    Set $c_j \leftarrow (\mathsf{msgq}\cdot \mathsf{Enc}_{\mathsf{pk}}(-j;0))^{s_j} \cdot \mathsf{Enc}_{\mathsf{pk}}(\mathsf{padded}(\sigma[j]);U(\mathcal{R}))$, where $s_j \leftarrow U(\mathcal{M})$.
  Return $\mathsf{Transfer}_{\mathsf{pk}}(\mathsf{msgq},\sigma;t,s,r) \leftarrow (c_1,\ldots,c_N)$.
$\mathsf{Recover}_{\mathsf{sk}}(\varrho,\mathsf{msgt};\cdot)$: Return $\mathsf{Dec}_{\mathsf{sk}}(c_\varrho) \mod 2^\ell$.

Protocol 1: An additively homomorphic one-round protocol for $\mathsf{OT}^N_{2^\ell}$

Composability of additively homomorphic one-round protocols. We next show that privacy of additively homomorphic one-round protocols is preserved under reasonable concurrent (parallel and sequential) executions. For an honest sender, we assume that KProof has been successful (otherwise Sender just does not participate in any protocol).

Lemma 1. *Assume that* KProof *is ε_{zk}-sound. Assume that additively homomorphic one-round protocols Π_i are ε_i-sender-private on the promise that the public key* pk *is valid. Then a concurrent composition Π of protocols Π_i, that all share the same key pair* $(\mathsf{sk},\mathsf{pk})$*, is an ε'-sender-private protocol, where $\varepsilon' = \max\{\varepsilon_{zh}, \varepsilon_1 + \ldots + \varepsilon_s\}$.*

The proof is given in App. A. Composing additively homomorphic one-round protocols preserves the receiver-privacy; the corresponding result—that we will not need in this paper but which is needed to ascertain the validity of the multi-user setting—is given in App. B. As a result, we have established that any protocol based on additively homomorphic one-round protocols that either share or do not share the same key pair preserves privacy, provided that the KProof protocol runs are executed in isolation. The latter is quite easy to achieve in practice by introducing timing limits for key proof protocols.

4 Additively Homomorphic One-Round Protocol for Oblivious Transfer

Next, we propose a new 1-out-of-N-oblivious transfer protocol (see Protocol 1) that achieves sender-privacy under the assumption that n is sufficiently rough. We first need the following technical lemma about the difference of two related distributions.

Lemma 2. *Fix integers n, ℓ and κ, such that $2^\ell \le \kappa \le n/2$ and $\gcd(n,\kappa) = 1$. For $x \in \{0,1\}^\ell$, denote $\mathcal{Z}_x := \left\{x + \kappa t : 0 \le t < \frac{n-x}{\kappa}\right\}$. Let $\mathcal{Z}_x^y$ denote the distribution that one gets by first uniformly choosing a random element of $\mathcal{Z}_x$ and then reducing it modulo y. Then, for any $x \in \{0,1\}^\ell$ and for any non-trivial factor p of n, $\mathsf{Dist}\left(\mathcal{Z}_x^p \| U(\mathbb{Z}_p)\right) < \frac{\kappa}{2\Phi(n)}$.*

The proof of this technical lemma is given in App. C.

Theorem 1. *Let $\mathsf{PKC} = (\mathsf{Gen},\mathsf{Enc},\mathsf{Dec})$ be an additively homomorphic public-key cryptosystem that satisfies the next requirements: (a) it is $(\varepsilon_{\mathsf{PKC}},\tau)$-IND-CPA secure and ε_{zk}-sound, (b) n is $(2^{\lambda-1}\cdot 2^\ell N)$-rough for some $\lambda \in \mathbb{Z}^+$, (c) $\mathcal{R}$ is a quasigroup, (d) $\mathcal{M}$ and $\mathcal{R}$ are efficiently samplable, and (e) membership in $\mathcal{C}$ can be efficiently verified. Assume that Sender has verified the corrected of* pk *by using* KProof *and halted if the verification failed. Fix $\ell \leftarrow \lfloor \log_2 \Phi(n) - \log_2 N - \lambda + 1 \rfloor$. Then Protocol 1 is an $(\varepsilon_{\mathsf{PKC}},\tau - \mathcal{O}(1);\varepsilon_{zk}+\varepsilon')$-private additively homomorphic one-round protocol for $\mathsf{OT}^N_{2^\ell}$, where $\varepsilon' = 2^\ell N/\Phi(n) \le 2^{-\lambda}$.*

Proof. Correctness is straightforward, since $\mathsf{Dec}_{\mathsf{sk}}(c_\varrho) = (\varrho-\underline{\varrho})\cdot s_\varrho + \mathsf{padded}(\sigma[\underline{\varrho}]) = \sigma[\underline{\varrho}] + 2^\ell \cdot z_j < n$. Computational receiver-privacy follows directly from the IND-CPA security of PKC, since Sender sees only a random encryption of ϱ. Statistical sender-privacy: Assume that Sender is honest and that pk is valid. Then, for any j, c_j is an encryption of $v_j := (\varrho - j)s_j + \mathsf{padded}(\sigma[j])$ where v_j is distributed as $V_j := (\varrho-j)U(\mathbb{Z}_n) + \sigma[j] + 2^\ell \cdot U(\mathbb{Z}_{T_j})$. Since $\mathcal{R}$ is a quasigroup, c_j is a random encryption of v_j. Let $q := \gcd(\varrho - j, n) \neq 0$. Clearly, $\Pr[v_j \equiv y \pmod n] = \frac{q}{n}\cdot \Pr[U(\mathcal{Z}_{\sigma[j]}) \equiv y \mod q]$. Now, Lemma 2 assures that $\mathsf{Dist}\left(U(\mathbb{Z}_n)\|V_j\right) = \mathsf{Dist}\left(U(\mathbb{Z}_q)\|V_j \mod q\right) \le \frac{2^\ell}{2\Phi(n)}$. As all N distributions V_j are independent,

$\mathsf{Dist}\left((V_1,\ldots,V_N)\|\mathbb{Z}_n^N\right) \leq \frac{N2^\ell}{2\Phi(n)} \leq \frac{N}{2\Phi(n)} \cdot \frac{\Phi(n)}{N} \cdot 2^{-\lambda+1} = 2^{-\lambda}$. Here, the second inequality follows from the assumption (b). Thus, we simulate Receiver's view as follows: Compute sk corresponding to pk, halt if it does not exist. Set $\varrho \leftarrow \mathsf{Dec}_{\mathsf{sk}}(\mathsf{msgq})$, set $\mathsf{msgt} = \perp$ if msgq is not a valid ciphertext. For all $j \in [N]$ do: If $j \neq \varrho$, set $c_j \leftarrow \mathsf{Enc}_{\mathsf{pk}}(U(\mathcal{M}); U(\mathcal{R}))$. If $j = \varrho$, set $c_j \leftarrow \mathsf{Enc}_{\mathsf{pk}}(\mathsf{padded}(\sigma[\varrho]); U(\mathcal{R}))$. Output $c = (c_1,\ldots,c_N)$. If adversary is semi-honest then simulation is perfect, otherwise we get a statistical difference that is smaller than $2^{-\lambda}$. □

All well-known homomorphic public-key cryptosystems [El 84,OU98,NS98,Pai99,DJ01,DJ03] have the required properties. In all practical situations, we can assume that $\lambda = 80$ and $N \leq 2^{80}$, then a 2^{160}-rough n is sufficient. In this case, $\ell \leq 81 - \log_2 N$, this is sufficient if inputs are Boolean. If PKC is Paillier's cryptosystem then n is $\sqrt{n}/2$-rough, and consequently, one can take $\ell = \left\lfloor \frac{1}{2}\log_2 n - \log_2 N - \lambda \right\rfloor$. For $\log_2 n = 1024$ and $\lambda = 80$, we get $\ell = \lfloor 433 - \log_2 N \rfloor$. Finally, Protocol 1 can be straightforwardly modified to transfer $\ell' > \ell$ bits by repeating the second message of the proposed oblivious transfer protocol $\lceil \ell'/\ell \rceil$ times.

Corollary 1. *Let $\varepsilon_{\mathsf{PKC}}$, ε_{zk}, ε', τ and ℓ be as in Thm. 1, and let $\mathcal{S} \subseteq \mathcal{M}$ be an arbitrary index set. There exists an additively homomorphic one-round protocol for $\mathsf{OT}_{2^\ell}^{\mathcal{S}}$ that is $(\varepsilon_{\mathsf{PKC}}, \tau - \mathcal{O}(1); \varepsilon_{zk} + \varepsilon')$-private.*

Proof. As in Prot. 1, but compute $c_j \leftarrow (c \cdot \mathsf{Enc}_{\mathsf{pk}}(-h_j; 0))^{s_j} \cdot \mathsf{Enc}_{\mathsf{pk}}(\mathsf{padded}(\sigma[j]); U(\mathcal{R}))$ for $h_j \in \mathcal{S}$, and set $\mathsf{msgt} \leftarrow \{c_j : j \in \mathcal{S}\}$. □

Given an arbitrary $\mathsf{CPIR}_{d'}^N$ protocol with $d' \geq \mathcal{C}$, one can construct an almost as efficient $\mathsf{OT}_{d'}^N$ as follows: as in Protocol 1, Receiver sends $\mathsf{Enc}_{\mathsf{pk}}(\varrho; r)$ to Sender, who computes the values c_i. After that, Receiver uses the $\mathsf{CPIR}_{d'}^N$ protocol to retrieve c_ϱ. In particular, [Lip04] gives us the next result.

Corollary 2. *Let ε' and ℓ be as in Thm. 1. Let PKC be a length-flexible additively homomorphic cryptosystem [DJ01] that satisfies the same properties as required in Thm. 1. There exists a one-round $(2\varepsilon \cdot \log_2 N, \tau - \mathrm{polylog}(N); \varepsilon')$-private $\mathsf{OT}_{2^\ell}^N$ protocol with communication $\Theta(k \cdot \log^2 N + \ell \cdot \log N)$, where k is a possibly non-constant security parameter.*

Proof. Correctness is obvious. Receiver-privacy is the same as in Lipmaa's computationally-private information retrieval protocol. Sender-privacy follows from Thm. 1. □

For a non-length-flexible PKC, [Ste98] gives an $\mathsf{OT}_{2^\ell}^N$ protocol with communication $\Theta(\sqrt{\log N} \cdot 2^{\sqrt{\log N}} \cdot \ell \cdot k)$. Discussion on the optimality of the used padding scheme has been moved to App. D.

5 Additively Homomorphic One-Round Protocol for Conditional Disclosure of Secrets

For the purposes of the current paper, a conditional disclosure of secrets protocol $\mathsf{CDS}_d^{\mathcal{S}}$ for a public set $\mathcal{S}$ and inputs from $\mathbb{Z}_d$ is an additively homomorphic one-round protocol for the next functionality: $f(\varrho, \sigma) = \sigma$ if $\varrho \in \mathcal{S}$ and $f(\varrho, \sigma)$ is a random σ-independent element otherwise. Based on the previous section, we can now prove the next result (the communication will again increase if PKC is not length-flexible):

Corollary 3. *Let ε' and ℓ be as in Thm. 1. Let PKC be a length-flexible additively homomorphic cryptosystem [DJ01] that satisfies the same properties as required in Thm. 1. There exists a one-round $(2\varepsilon \cdot \log_2 N, \tau - \mathrm{polylog}(N); \varepsilon')$-private one-round additively homomorphic one-round protocol for $\mathsf{CDS}_{2^\ell}^N$ with communication $\Theta(k \cdot \log^2 N + \ell \cdot \log N)$, where k is a possibly non-constant security parameter and $N = \sharp\mathcal{S}$.*

Proof. Follows from Cor. 1 and Cor. 2 by executing $\mathsf{OT}_{2^\ell}^{\mathcal{S}}$ with database σ', where $\sigma'[i] = \sigma$ for all $i \in \mathcal{S}$. □

Next, we will show how to use explicit circuit evaluation to construct, given an arbitrary set $\mathcal{S}$, an additively homomorphic one-round protocol for $\mathsf{CDS}_d^{\mathcal{S}}$ that can be much more efficient than the generic construction implied by Cor. 3. Namely, we specify the predicate $[\varrho \in \mathcal{S}]$ by a set of constraints on Receiver's input ϱ, where for the sake of efficiency, ϱ might be broken down to several smaller inputs from which one can reconstruct the original inputs by using affine operators. The constraints can be written down as a suitable *monotone Boolean formula with affine zero tests $\Psi_{\mathcal{S}}$*, where we allow Boolean operations $\wedge$ and $\vee$ and affine zero tests $[\sum \alpha_i \varrho_i \stackrel{?}{=} \beta]$. We assume that the $\vee$ gates have an arbitrary fan-in. We will motivate the next discussion with a real problem, the millionaire's, with $\mathcal{S}$ equal to $\mathcal{GT}_M(y) := \{x \in \{0,1\}^M : x > y\}$, for M-bit strings that are split into M one-bit inputs. Writing $x = (x_{M-1},\ldots,x_0)$, we get that $\Psi_{\mathcal{GT}_M(y)} := ([x_{M-1} = 1] \wedge [y_{M-1} = 0]) \vee ([x_{M-1} = y_{M-1}] \wedge [x_{M-2} =$

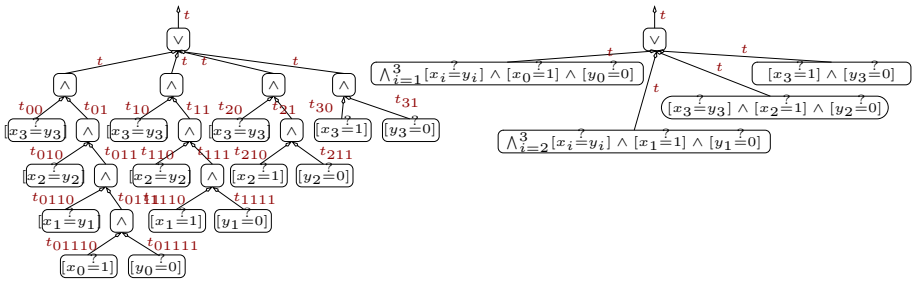

Fig. 1. Circuit for $\mathcal{GT}_4(y)$: unoptimised and optimised versions.

$1] \wedge [y_{M-2} = 0]) \vee ([x_{M-1} = y_{M-1}] \wedge [x_{M-2} = y_{M-2}] \wedge [x_{M-3} = 1] \wedge [y_{M-3} = 0]) \vee \cdots \vee ([x_{M-1} = y_{M-1}] \wedge [x_{M-2} = y_{M-2}] \cdots \wedge [x_1 = y_1] \wedge [x_0 = 1 \wedge y_0 = 0])$.

The general idea of the circuit evaluation process is as follows (see Fig. 1, left): Construct a circuit where every internal node implements a Boolean operation and every leaf implements an affine zero test. Process the circuit recursively from top to bottom. Assign a random secret $t \leftarrow U(\mathbb{Z}_{2^\ell})$ to the output wire of the circuit. For every $\wedge$ gate ψ with secret t' assigned to its output wire, pick $t'_1 \leftarrow U(\mathbb{Z}_{2^\ell})$ and $t'_2 \leftarrow t' - t'_1 \mod 2^\ell$, and assign these values to the two input wires of ψ. For every $\vee$ gate, just push the output secret downwards. In the query phase of the resulting $\mathsf{CDS}_d^{\mathcal{S}}$ protocol, Receiver transfers $P_i \leftarrow \mathsf{Enc}_{\mathsf{pk}}(\varrho_i; U(\mathcal{R}))$ for every $i \in [M]$. In the transfer phase of the protocol, for every leaf ψ with the corresponding affine zero test $[\sum_{i=1}^M \alpha_i \varrho_i \stackrel{?}{=} \beta]$ and output secret t_ψ, Sender replies with $c_\psi \leftarrow (\prod_{i=1}^M P_i^{\alpha_i} \cdot \mathsf{Enc}_{\mathsf{pk}}(-\beta; 0))^{U(\mathcal{M})} \cdot \mathsf{Enc}_{\mathsf{pk}}(\mathsf{padded}(t_\psi); U(\mathcal{R}))$, where $\mathsf{padded}(t_\psi)$ is defined as in Protocol 1. In the recovery phase, Receiver decrypts ciphertexts that correspond to the correct branch in the circuit, and recovers all the secrets (modulo 2^ℓ). Therefore, Receiver transfers M ciphertexts and Sender transfers $L(\Psi_{\mathcal{S}})$ ciphertexts, where $L(\Psi_{\mathcal{S}})$ is the number of affine zero tests. Clearly, this protocol is correct. Following our motivating example, $\Psi_{\mathcal{GT}_M(y)}$ (see Fig. 1, left), we get a circuit with $L(\Psi_{\mathcal{GT}_M(y)}) = M(M+1)/2+M$. Therefore, applying the previous construction results in a $\mathsf{CDS}_{2^\ell}^{\mathcal{GT}_M(y)}$ protocol with the communication of $M(M+1)/2+2M$ ciphertexts.

We can do better by allowing leaf gates that perform a conjunction of several zero tests. In such a circuit, a uniformly random secret t is propagated to bottom as previously. Also, the query phase remains unchanged. Now, let ψ be an arbitrary leaf gate that corresponds to a conjunction of v_ψ different zero tests, $\bigwedge_{j=1}^{v_\psi} [\sum_{i=1}^M \alpha_{ij} \varrho_i \stackrel{?}{=} \beta_j]$; let t_ψ be the output secret of ψ. In the transfer phase, Sender first computes $w_{\psi,j} \leftarrow \prod_{i=1}^M P_i^{\alpha_{ij}} \cdot \mathsf{Enc}_{\mathsf{pk}}(-\beta_j; 0)$, for $j \in [v_\psi]$, and then replies with $c_\psi \leftarrow \prod_{j=1}^{v_\psi} w_{\psi,j}^{U(\mathcal{M})} \cdot \mathsf{Enc}_{\mathsf{pk}}(\mathsf{padded}(t_\psi); U(\mathcal{R}))$. Since $\mathsf{Dec}_{\mathsf{sk}}(c_\psi) = \sum_{i=1}^{v_\psi} \mathsf{Dec}_{\mathsf{sk}}(w_{\psi,i}) \cdot U(\mathbb{Z}_n) + \mathsf{padded}(t_\psi)$, Receiver learns nothing unless all the affine zero tests of ψ are satisfied; in particular, Receiver does not even learn which zero test fails. We call this protocol $\mathsf{CircuitCDS}_{2^\ell}^{\mathcal{S}}$. More precisely, let $L_2(\Psi_{\mathcal{S}})$ be the number of leaves (that is, of conjunctive affine zero tests) in the latter circuit. Then we get the following result.

Theorem 2. *Let $\Psi_{\mathcal{S}} : \{0,1\}^M \to \{0,1\}$ be a public monotone Boolean formula with conjunctive affine zero tests. Let* PKC *be an additively homomorphic public-key cryptosystem that satisfies the same requirements as required in Thm. 1, and* padded *be the corresponding padding with $\ell \leftarrow \lfloor \log_2 \Phi(n) - \log_2 L_2(\Psi_{\mathcal{S}}) - \lambda + 1 \rfloor$. Then the* $\mathsf{CircuitCDS}_{2^\ell}^{\mathcal{S}}$ *protocol is $(M \cdot \varepsilon, \tau - \mathcal{O}(1); \varepsilon_{zk} + \varepsilon')$-private with $\varepsilon' \leq 2^{-\lambda}$. The communication of* $\mathsf{CircuitCDS}_{2^\ell}^{\mathcal{S}}$ *is $M + L_2(\Psi_{\mathcal{S}})$ ciphertexts.*

Proof. Correctness is clear. Receiver-privacy is also straightforward, since Sender sees only M encryptions of 0's and 1's. Sender-privacy: the simulator Sim works as follows. First, it computes the secret key sk corresponding to pk; it halts in the case of failure. Then, it decrypts all inputs P_i and obtains the corresponding input ϱ. If $\Psi_{\mathcal{S}}(\varrho) = 0$ then Sim replies with $L_2(\Psi_{\mathcal{S}})$ random encryptions. Otherwise Sim propagates $t = f(\varrho, \sigma)$ down to the leaf level and computes the corresponding Transfer messages. If Receiver is honest, the simulation is perfect. Otherwise, the statistical difference between the replies is at most ε' as in the proof of Thm. 1. □

Going back to the motivating example, $L_2(\Psi_{\mathcal{GT}_M(y)}) = M$. (See Fig. 1, right. Here, Sender just transfers the same secret M times.) Therefore, under the same assumptions as in Thm. 2, there exists an $(M \cdot \varepsilon, \tau - \mathcal{O}(1); \varepsilon_{zk} + \varepsilon')$-private additively homomorphic one-round protocol for $\mathsf{CDS}_{2^\ell}^{\mathcal{GT}_M(y)}$, with the communication of $2M$ ciphertexts.

One-round millionaire's protocol with logarithmic communication. Before describing the general CDS transformation, we will show how to generalise our methodology to (some) private sets $\mathcal{S}$ that depend on σ. More precisely, assume that Receiver has a private input ϱ and that Server has a private input (σ, t) with $t \in \mathbb{Z}_{2^\ell}$. Note that the $\mathsf{CircuitCDS}^{\mathcal{S}}_{2^\ell}$ protocol is most efficient if $\Psi_{\mathcal{S}}$ is written down in a disjunctive normal form over affine zero tests, $\Psi_{\mathcal{S}} = \bigvee_{i=1}^{L} \bigwedge_{j=1}^{n_i} [\sum_{i=1}^{M} \alpha_i \varrho_i \stackrel{?}{=} \beta]$. Now, modify the $\mathsf{CircuitCDS}^{\mathcal{S}}_{2^\ell}$ protocol as follows. Fix a *public* value t (for example, $t = 0$) and push it down the circuit. For every leaf gate, let Sender to compute c_ψ as previously, but return the values c_ψ in a random order. Be careful to do that so that the number of accepting leaf gates is always either 0 (if $\Psi_{\mathcal{S}} = 0$) or some non-zero constant (if $\Psi_{\mathcal{S}} = 1$); this can be done efficiently for many interesting sets $\mathcal{S}$. Therefore, by testing that at least one of the ciphertexts c_ψ encrypts 0, Receiver gets to know whether $\Psi_{\mathcal{S}}(\varrho, \sigma)$ is true or not.

Since $L_2(\Psi_{\mathcal{GT}_M(y)}) = M$, we get a new one-round protocol for millionaire's problem of M-bit strings that is secure against malicious adversaries, with communication of $2M$ ciphertexts, assuming only that the underlying additively homomorphic public-key cryptosystem is IND-CPA secure.

Corollary 4. *The just described protocol is an $(\varepsilon, \tau; \varepsilon')$-private additively homomorphic one-round protocol for the millionaire's problem. The error probability is $M \cdot 2^{-\ell}$, where $\mathbb{Z}_{2^\ell}$ is the secret space.*

Proof. The security claims are obvious. If Receiver and Sender are semi-honest and $x \leq y$, all M replies are random encryptions. □

Conditional oblivious transfer. A *conditional oblivious transfer* ($\mathsf{COT}^{\mathcal{S}}_d$) protocol [DOR99] is a protocol, at the end of which Receiver obtains t only if $\Psi_{\mathcal{S}}(\varrho, \sigma) = 1$ for some public set $\mathcal{S}$ of valid Receiver's and Sender's input pairs, and no information, otherwise. To implement $\mathsf{COT}^{\mathcal{S}}_d$, we use the same idea as in the case of the millionaire's problem with only one modification: the secret to push down the circuit is $t' = 0^\kappa || t$, where say $\kappa = 80$. This approach works for sets $\mathcal{S}$ that have an efficient implementation for formula $\Psi_{\mathcal{S}}$.

6 CDS Transformation for Conventional Protocols

In this section, we present a generic transformation from private in the semi-honest model protocols to private in the malicious model protocols. It can be called as a *compiler* since this transformation can be constructed in a relatively automatic manner. More precisely, fix an additively homomorphic one-round protocol Π. Denote Receiver's input by $\varrho = (\varrho_1, \ldots, \varrho_M)$, Sender's input by $\sigma = (\sigma_1, \ldots, \sigma_N)$ and Receiver's output by $\delta = (\delta_1, \ldots, \delta_L)$. Here, w.l.o.g., we assume that ϱ_i, σ_i and δ_i belong to $\mathbb{Z}_{2^\ell}$, where ℓ is as in Thm. 1. Larger inputs and outputs can be obtained straightforwardly. The query phase consists of sending the elements $\mathsf{Enc}_{\mathsf{pk}}(\varrho_i; r_i)$ and the transfer phase consists of sending the elements $\mathsf{Enc}_{\mathsf{pk}}(\delta_j; r''_j)$ for some r_i and r''_j. We assume that ϱ, σ and δ have already been modified to facilitate efficient circuit evaluation. For example, in the case of $\mathcal{GT}_M(y)$, every ϱ_i is a Boolean value.

We say that an additively homomorphic one-round protocol Π is *conventional* if (a) the input ϱ of an honest Receiver belongs to some publicly known set Valid that in particular does not depend on the value of sk, and (b) sender-privacy is guaranteed if $\varrho \in \mathsf{Valid}$. Most of the known additively homomorphic one-round protocols are indeed conventional, Chang's oblivious transfer protocol [Cha04] being one of the few exceptions. Importantly, $\mathsf{CircuitCDS}$ is a conventional additively homomorphic one-round protocol. Our next transformation works only for conventional protocols since in an unconventional protocol, Sender does not know the set Valid and thus cannot execute the CDS protocol. To simplify the implementations, we assume that if $\varrho \notin \mathsf{Valid}$ then for any input value of an honest Sender, $f(\varrho, \sigma)$ is defined to be a uniformly random value from some fixed set.

Let Π be a conventional additively homomorphic one-round protocol for functionality f with L outputs from $\mathbb{Z}_{2^\ell}$, and let Π^{cds} be a conventional additively homomorphic one-round protocol for $\mathsf{CDS}^{\mathcal{S}}_{2^\ell}$, where ℓ is as defined in Thm. 1. The idea is to compose an instantiation of Π with an instantiation Π^{cds}, on the same inputs ϱ and σ, as follows. Assume that the query phase of both Π and Π^{cds} is the same; this is possible since in the previously constructed additively homomorphic one-round protocol for $\mathsf{CDS}^{\mathsf{Valid}}_{2^\ell}$, Receiver learns a secret $t \in \mathbb{Z}^L_{2^\ell}$ iff $\varrho \in \mathsf{Valid}$, and Sender learns the corresponding ciphertexts $P_i = \mathsf{Enc}_{\mathsf{pk}}(\varrho_i; U(\mathcal{R}))$. Therefore, in the transfer phase, Sender can use the ciphertexts P_i as input to an additively homomorphic one-round protocol Π that evaluates f. Finally, Sender masks the outputs $(\Delta_1, \ldots, \Delta_L)$ of Π with sub-secrets t_i and sends corresponding encryptions $\Delta_i \cdot \mathsf{Enc}_{\mathsf{pk}}(t_i; r_i)$ for $r_i \leftarrow U(\mathcal{R})$ to Receiver. Therefore, Receiver can peel off the masks t_i iff her inputs are in the correct range.

PRIVATE INPUT: Receiver has inputs $\varrho = (\varrho_1, \ldots, \varrho_M)$, Sender has inputs $\sigma = (\sigma_1, \ldots, \sigma_N)$.
PRIVATE OUTPUT: Receiver obtains $(\delta_1, \ldots, \delta_L) = f(\varrho; \sigma)$ where $\delta_j \in \mathbb{Z}_{2^\ell}$.

$\mathsf{Query}_{\mathsf{pk}}(\varrho; \cdot)$:
For $i \in [M]$: Set $P_i \leftarrow \mathsf{Query}_{\mathsf{pk}}(\varrho_i; r_i)$, $r_i \leftarrow U(\mathcal{R}_Q)$. Send $(P_1, \ldots, P_M)$ to Sender.
$\mathsf{Transfer}_{\mathsf{pk}}(\sigma, \mathsf{msgq}; \cdot)$:
For $j \in [L]$:
Compute $\widehat{t}_j \leftarrow U(\mathbb{Z}_{n-2^\ell})$ and a set of ciphertexts $\{c_{ij}\}$ from the output secret $t_j \leftarrow \widehat{t}_j \mod 2^\ell$ as in Π^{cds}.
Compute Δ_j as in the original protocol. Set $\mathsf{mask}_j \leftarrow \mathsf{Enc}_{\mathsf{pk}}(t_j; 0)$ and $\Delta'_j \leftarrow \Delta_j \cdot \mathsf{mask}_j$.
Send $(\Delta'_1, \ldots, \Delta'_L; \{c_{i1}\}, \ldots, \{c_{iL}\})$ to Receiver.
$\mathsf{Recover}_{\mathsf{sk}}(\varrho, \mathsf{msgt}; \cdot)$:
For $j \in [L]$: Set $t_j \leftarrow \mathsf{Recover}^{\mathsf{cds}}_{\mathsf{sk}}(\varrho, \{c_{ij}\})$ as in Π^{cds}. Set $\delta'_j \leftarrow \mathsf{Recover}_{\mathsf{sk}}(\varrho, \Delta'_j)$ and $\delta_j \leftarrow \delta'_j - t_j \mod 2^\ell$.
Return $(\delta_1, \ldots, \delta_L)$.

Protocol 2: Private computation of a function f in malicious model by using additive CDS transformation

Theorem 3. *Fix an additively homomorphic public-key cryptosystem* PKC. *Let* Gen *and* KProof *be as usually, and fix a concrete secret and public key pair* $(\mathsf{sk}, \mathsf{pk})$. *Let* $\Pi^{\mathsf{cds}} = (\mathsf{Query}, \mathsf{Transfer}^{\mathsf{cds}}, \mathsf{Recover}^{\mathsf{cds}})$ *be an* $(\varepsilon, \tau; \varepsilon'_1)$*-private additively homomorphic one-round protocol for* $\mathsf{CDS}^{\mathsf{Valid}}_{2^\ell}$. *Let* $\Pi = (\mathsf{Query}, \mathsf{Transfer}, \mathsf{Recover})$ *be an* $(\varepsilon, \tau; \varepsilon'_2)$*-private conventional additively homomorphic one-round protocol for computing* f *in the semi-honest model.* Π'*, depicted by Prot. 2, is an* $(\varepsilon, \tau - \mathcal{O}(1); \varepsilon'_1 + \varepsilon'_2 + \varepsilon'_3)$*-private additively homomorphic one-round protocol for computing* f *in the malicious mode, where* $\varepsilon'_3 = \frac{2^\ell}{n} L$.

Proof. Correctness: If $\varrho \in \mathsf{Valid}$ and both parties follow the protocol then recovery phase of the $\mathsf{CDS}^{\mathsf{Valid}}_{2^\ell}$ protocol is successful, Receiver obtains $\widehat{t}_j \mod 2^\ell$ and consequently the correct end-result, as there are no modular wrappings. Receiver-privacy: Consider an adversary A that obtains advantage ε against Prot. 2; A can used against the Π^{cds} protocol, since the query phase is exactly the same. For the same reason, Π^{cds} cannot be more secure than Π. Sender-privacy: Clearly, Π' is a parallel execution of two additively homomorphic one-round protocols. Therefore, it is $(\varepsilon'_1 + \varepsilon'_2)$-sender-private implementation of the functionalities $\widehat{f}_j(\varrho, \sigma) = (\delta_j + t_j, t_j)$, if $\varrho \in \mathsf{Valid}$, and $\widehat{f}_j(\varrho, \sigma) = (\delta_j + t_j, \perp)$, if $\varrho \notin \mathsf{Valid}$. The claim follows as t_j are almost random plaintexts and ε'_3/L is the statistical difference between $U(\mathbb{Z}_{n-2^\ell})$ and $U(\mathbb{Z}_n)$. □

With a slight modification (setting $\widehat{t}_j \leftarrow U(\mathbb{Z}_n)$ and using the CDS on a $2^{\ell+1}$ bit secret where one bit indicates the modular wrap), one can remove the addend ε'_3. Note that this theorem does not require PKC to be an additively homomorphic public-key cryptosystem, and thus the same proof goes through also with a multiplicatively homomorphic public-key cryptosystem.

Comparison with related work. A well-known alternative to the additive CDS transformation is to let Receiver to prove in zero-knowledge that $(P_1, \ldots, P_M)$ encrypts a value from Valid; this means that either the resulting protocol takes at least three messages or that the protocol is only secure in the common reference string (or random oracle) model. As we have shown, one can use a mixture of arithmetic and Boolean formulas to construct an efficient additively homomorphic one-round protocol for $\mathsf{CDS}^{\mathcal{S}}_d$. A similar efficiency can be achieved by using non-interactive zero-knowledge proofs that work over additively homomorphic public-key cryptosystems, but compared to them, the additive CDS transformation uses simpler basic components. The difference in efficiency comes in the use of a oblivious transfer instead of a zero-knowledge disjunctive proof [CDS94]: the first can be done in the complexity-theoretic model very efficiently, while non-interactive zero-knowledge proofs are not possible in the complexity-theoretic model, and are somewhat more complex to implement in the random oracle model

Compared to the CDS transformation from [AIR01], the additive CDS transformation is applicable in a wider setting since there exist efficient protocols that crucially rely on additively homomorphic public-key cryptosystems. Using the transformation from [AIR01] in these cases is either impossible or requires one to rely on the Decisional Diffie-Hellman assumption in addition to the assumption that PKC is IND-CPA secure. Note that the Aiello-Ishai-Reingold transformation works also only for conventional protocols.

Optimisations. The communication overhead of the additive CDS transformation is linear in the number of outputs. Therefore, it is not advantageous to use the transformation for functions with many outputs (e.g., private matrix operations). However, if computational sender-privacy is sufficient, one can use an arbitrary pseudo-random generator prg to stretch the transformation's secret to implement the functionalities $\widehat{f}_j(\varrho, \sigma) = (\delta_j + \mathsf{prg}(t), t)$

if $\varrho \in \mathsf{Valid}$ and $\widehat{f}_j(\varrho, \sigma) = (\delta_j + \mathsf{prg}(t), \perp)$ if $\varrho \notin \mathsf{Valid}$, for a single random t. Such a protocol remains computationally sender-private as long as prg is cryptographically strong.

Private scalar product protocol. Assume that Receiver has a Boolean vector ϱ of dimension M and Sender has Boolean vector σ of the same dimension. In a *private scalar product protocol* protocol, Receiver's private output is δ, such that $\delta = \sum_{i=1}^{M} \sigma_i \varrho_i$, and Sender has no private output. It is simple to compute this functionality in the semi-honest model. Assume that c_i is a random encryption of $\varrho[i]$. Then, Sender sends $d_i = \sum_{i=1}^{M} c_i^{\sigma[i]} \cdot \mathsf{Enc}_{\mathsf{pk}}(0; r_i)$, for random r_i, to Receiver. Receiver decrypts d_i. It is straightforward to apply the additive CDS transformation to get a protocol that is sender-private in the malicious model. This protocol also computes the private set intersection. Similar ideas can be used to construct private protocols for many other related problems.

Acknowledgements. We would like to thank XXX for useful comments. The work was partially supported by the Finnish Academy of Sciences and by the Estonian Science Foundation.

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*, volume 2045 of *Lecture Notes in Computer Science*, pages 119–135, Innsbruck, Austria, 6–10 May 2001. Springer-Verlag.

[BK04] Ian F. Blake and Vladimir Kolesnikov. Strong Conditional Oblivious Transfer and Computing on Intervals. In Lee [Lee04], pages 515–529.

[CDS94] Ronald Cramer, Ivan Damgård, and Berry Schoenmakers. Proofs of Partial Knowledge and Simplified Design of Witness Hiding Protocols. In Yvo G. Desmedt, editor, *Advances in Cryptology—CRYPTO '94*, volume 839 of *Lecture Notes in Computer Science*, pages 174–187, Santa Barbara, USA, August 21–25 1994. Springer-Verlag.

[Cha04] Yan-Cheng Chang. Single Database Private Information Retrieval with Logarithmic Communication. In Josef Pieprzyk and Huaxiong Wang, editors, *The 9th Australasian Conference on Information Security and Privacy (ACISP 2004)*, volume 3108 of *Lecture Notes in Computer Science*, pages 50–61, Sydney, Australia, 13–15 July 2004. Springer-Verlag.

[DJ01] Ivan Damgård and Mads Jurik. A Generalisation, a Simplification and Some Applications of Paillier's Probabilistic Public-Key System. In Kwangjo Kim, editor, *Public Key Cryptography 2001*, volume 1992 of *Lecture Notes in Computer Science*, pages 119–136, Cheju Island, Korea, 13–15 February 2001. Springer-Verlag.

[DJ03] Ivan Damgård and Mads Jurik. A Length-Flexible Threshold Cryptosystem with Applications. In Rei Safavi-Naini, editor, *The 8th Australasian Conference on Information Security and Privacy*, volume 2727 of *Lecture Notes in Computer Science*, pages 350–364, Wollongong, Australia, July 9-11 2003. Springer-Verlag.

[DOR99] Giovanni Di Crescenzo, Rafail Ostrovsky, and Sivaramakrishnan Rajagopalan. Conditional Oblivious Transfer and Timed-Release Encryption. In Stern [Ste99], pages 74–89.

[El 84] Taher El Gamal. A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms. In G. R. Blakley and David Chaum, editors, *Advances in Cryptology: Proceedings of CRYPTO 84*, volume 196 of *Lecture Notes in Computer Science*, pages 10–18, Santa Barbara, California, USA, 19–22 August 1984. Springer-Verlag, 1985.

[FIPR05] Michael J. Freedman, Yuval Ishai, Benny Pinkas, and Omer Reingold. Keyword Search and Oblivious Pseudorandom Functions. In Joe Kilian, editor, *The Second Theory of Cryptography Conference, TCC 2005*, volume 3378 of *Lecture Notes in Computer Science*, pages 303–324, Cambridge, MA, USA, February 10–12, 2005. Springer Verlag.

[Fis01] Marc Fischlin. A Cost-Effective Pay-Per-Multiplication Comparison Method for Millionaires. In David Naccache, editor, *Topics in Cryptology - CT-RSA 2001, The Cryptographer's Track at RSA Conference 2001*, volume 2020 of *Lecture Notes in Computer Science*, pages 457–472, San Francisco, CA, USA, 8–12 April 2001. Springer-Verlag. ISBN 3-540-41898-9.

[GIKM00] Yael Gertner, Yuval Ishai, Eyal Kushilevitz, and Tal Malkin. Protecting Data Privacy in Private Information Retrieval Schemes. *Journal of Computer and System Sciences*, 60(3):592–629, June 2000.

[Gol04] Oded Goldreich. *Foundations of Cryptography: Basic Applications*. Cambridge University Press, 2004.

[Lee04] Pil Joong Lee, editor. *Advances on Cryptology — ASIACRYPT 2004*, volume 3329 of *Lecture Notes in Computer Science*, Jeju Island, Korea, December 5-9 2004. Springer-Verlag.

[Lip03] Helger Lipmaa. Verifiable Homomorphic Oblivious Transfer and Private Equality Test. In Chi Sung Laih, editor, *Advances on Cryptology — ASIACRYPT 2003*, volume 2894 of *Lecture Notes in Computer Science*, pages 416–433, Taipei, Taiwan, November 30–December 4, 2003. Springer-Verlag.

[Lip04] Helger Lipmaa. An Oblivious Transfer Protocol with Log-Squared Total Communication. Technical Report 2004/063, International Association for Cryptologic Research, February 25 2004.

[NP01] Moni Naor and Benny Pinkas. Efficient Oblivious Transfer Protocols. In *Proceedings of the Twelfth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 448–457, Washington, DC, USA, January 7–9 2001. ACM Press.

[NS98] David Naccache and Jacques Stern. A New Public Key Cryptosystem Based on Higher Residues. In *5th ACM Conference on Computer and Communications Security*, pages 59–66, San Francisco, CA, USA, 3–5 November 1998. ACM Press.

[OU98] Tatsuaki Okamoto and Shigenori Uchiyama. A New Public-Key Cryptosystem as Secure as Factoring. In Kaisa Nyberg, editor, *Advances in Cryptology — EUROCRYPT '98*, volume 1403 of *Lecture Notes in Computer Science*, pages 308–318, Helsinki, Finland, May 31 – June 4 1998. Springer-Verlag.

[Pai99] Pascal Paillier. Public-Key Cryptosystems Based on Composite Degree Residuosity Classes. In Stern [Ste99], pages 223–238.

[ST04] Berry Schoenmakers and Pim Tuyls. Practical Two-Party Computation Based on the Conditional Gate. In Lee [Lee04], pages 119–136.

[Ste98] Julien P. Stern. A New and Efficient All or Nothing Disclosure of Secrets Protocol. In Kazuo Ohta and Dingyi Pei, editors, *Advances on Cryptology — ASIACRYPT '98*, volume 1514 of *Lecture Notes in Computer Science*, pages 357–371, Beijing, China, 18–22 October 1998. Springer-Verlag.

[Ste99] Jacques Stern, editor. *Advances in Cryptology — EUROCRYPT '99*, volume 1592 of *Lecture Notes in Computer Science*, Prague, Czech Republic, 2–6 May 1999. Springer-Verlag.

A Proof of Lemma 1

We prove slightly more than required by the definition. This is because the joint protocol that is based on additively homomorphic one-round sub-protocols might have more than two rounds, as some of Receiver's inputs might depend on Sender's answers. Thus, in the proof we construct a universal simulator that outputs the protocol messages in chronological order.

Proof. Follows from a standard hybrid argument. First, consider the case when pk is valid and KProof is not executed at all. Let A be the unbounded malicious receiver. Then the universal black-box Sim for the entire protocol Π just forwards any query msgq_i to the appropriate simulator Sim_i of Π_i and returns answers to A. Finally, it reconstructs the protocol transcript and outputs it. By induction, it is straightforward to show that the statistical difference between the original protocol transcript and the simulated one is at most $\varepsilon_1+\cdots+\varepsilon_s$: If $s = 1$ then the claim is trivial. Assume that the claim holds for all $s < s_0$ and consider protocol with s_0 sub-protocols. Let msgq_i be the first message. Consider a hybrid protocol run $\mathcal{H}$ where the first message is sent according to protocol Π_i and all other replies are simulated. Obviously, the statistical distance between the total simulation and hybrid run $\mathcal{H}$ is at most ε_i. Similarly, the induction assumption guarantees that the statistical difference between the hybrid run $\mathcal{H}$ and the complete protocol run is at most $\varepsilon_1 + \cdots + \varepsilon_{i-1} + \varepsilon_{i+1} + \cdots + \varepsilon_s$. Thus, the claim follows from the triangle inequality. For the general case, note that KProof messages are independent from Sender's inputs. Therefore, there is a difference with the previous case when KProof succeeds for an invalid key. Hence, the statistical difference between the transcripts is at most $\max\{\varepsilon_{zk}, \varepsilon_1 + \cdots + \varepsilon_s\}$. □

Though Lemma 3 suggests that the general protocol structure should be static, Lemma 1 shows that even adaptive choice of protocol order cannot reveal more information about Sender's inputs than specified by the protocol outputs.

B Composing Preserves Receiver-Privacy

We prove a slightly stronger result. Namely, we require only that the KProof is executed at any time in an non-concurrent manner. Note that the definition of receiver-privacy is adequate even for multi-round protocols. For the lemma we also need the next definition. We say that KProof is $(\varepsilon_h, \tau_1, \delta, \tau_2)$-*hiding*, if for any malicious verifier A with an auxiliary input s and working in time τ_1, there exist a black-box simulator Sim_A that works in time $\tau_1 + \delta$, such that for any distinguisher D working in time τ_2,

$$\mathsf{Adv}^{\mathsf{ZKHID}}_{\mathsf{ZK}}(A) := \max_{D,s} |\Pr[D^{A(s)} = 1] - \Pr[D^{\mathsf{Sim}_A(s)} = 1]| \leq \varepsilon_h \ ,$$

where the distinguisher is allowed to inspect the output of the oracle, the maximum is taken over all distinguishers working in time τ_2, and the probability is taken over the coin tosses of A, Sim_A and D. We omit τ_2 from the definition if we consider only a trivial distinguisher that just forwards the outputs.

Lemma 3. *Assume that* KProof *is* $(\varepsilon_h, \tau, \delta)$*-hiding for any key pair* $(\mathsf{pk}, \mathsf{sk})$*. Assume that no messages of other protocols are sent by Receiver during* KProof*. Then a concurrent composition* Π *of* $(\varepsilon_1, \tau), \ldots, (\varepsilon_s, \tau)$*-receiver-private additively homomorphic one-round protocols* Π_i*, that all share the same key pair* $(\mathsf{sk}, \mathsf{pk})$*, is an* $(\varepsilon_h + \varepsilon_1 +$

$\ldots + \varepsilon_s, \tau')$*-receiver-private protocol where* $\tau' = \tau - \delta - \Delta$ *and* $\Delta = O(s)$ *is the time to form the first message of the protocol* Π.

Proof. Follows from a standard hybrid argument. First, consider the receiver-privacy when KProof is not executed at all, i.e., when we have an ideal implementation of KProof. Let A be an adversary running in time τ' such that $\mathsf{Adv}^{\mathsf{RECPRI}}_{\Pi}(A) > \varepsilon_1 + \cdots + \varepsilon_s$ and let (ϱ_0, ϱ_1) with $\varrho_i = (\varrho_{i1}, \ldots, \varrho_{is})$ be the corresponding challenge pair. Now consider the distributions $D_i = (\mathsf{msgq}_{01}, \ldots, \mathsf{msgq}_{0i}, \mathsf{msgq}_{1,i+1}, \ldots, \mathsf{msgq}_{1n})$ where $\mathsf{msgq}_{bj} = \mathsf{Query}_{\mathsf{pk}}(\varrho_{bj}, U(\mathcal{R}_{Q_j}))$ is the first message of Π_j. Since A distinguishes D_0 and D_s with probability $\mathsf{Adv}^{\mathsf{RECPRI}}_{\Pi}(A)$, there exists an i such that A distinguishes D_{i-1} and D_i with probability larger than ε_i. Given pk, ϱ_0, ϱ_1 and msgq_{bi}, one can generate an element of D_{i-b} within time Δ. Thus, there exist an adversary A' running in time $\tau' + \Delta$ that has $\mathsf{Adv}^{\mathsf{RECPRI}}_{\Pi_i}(A') > \varepsilon_i$.

For the general case, we have to consider the effect of KProof. W.l.o.g. we can assume that KProof is executed after all messages msgq_i are sent, since messages can be delayed with constant increase in running-time. Now we can treat messages msgq_i as a part of the auxiliary input of A. Since KProof is ε_h-hiding, there exists a simulator Sim_A that without access to sk can fool any time-bounded distinguisher, and in particular, the distinguisher that just forwards the output. Hence $|\Pr[A = 1] - \Pr[\mathsf{Sim}_A = 1]| \leq \varepsilon_h$ and consequently Sim_A has advantage $\mathsf{Adv}^{\mathsf{RECPRI}}_{\Pi}(\mathsf{Sim}_A) > \varepsilon_1 + \cdots + \varepsilon_s$ whenever $\mathsf{Adv}^{\mathsf{RECPRI}}_{\Pi}(A) > \varepsilon_1 + \cdots + \varepsilon_s + \varepsilon_h$. As the construction of A is independent of sk so is the construction of Sim_A. From the first part of the proof we get that there exists an adversary Sim'_A such that $\mathsf{Adv}^{\mathsf{RECPRI}}_{\Pi_i}(\mathsf{Sim}'_A) > \varepsilon_i$ for some i. The latter is impossible unless the working time of Sim'_A is larger than τ. The claim follows. □

It might seem that malicious Sender can mount more powerful attack, as malformed replies can change input values of Receiver's inputs. However, Lemma 3 assures that Sender cannot distinguish protocol runs even if Receiver's inputs are known. To summarise, receiver-privacy is guaranteed unless protocol description depends on Sender's replies, but latter is unavoidable.

Note that this is the only place in the paper we explicitly deal with the imperfectly hiding KProof. As seen from the proof, dealing with it introduces some technicalities that are relatively straightforward to also include to other results. Finally, in the multi-user setting, we have to consider the case of colluding Senders. Dealing with this case is not difficult, we would only have to change ε_h with $X\varepsilon_h$, where X is the number of colluding Senders.

C Proof of Lemma 2

Proof. Let p be a non-trivial factor of n. Fix $x \in \{0,1\}^\ell$. Denote $s := \sharp\mathcal{Z}_x = ap + r$ for $r \in [0, p-1]$. As $\gcd(\kappa, p) = 1$, then κ is a generator of $\mathbb{Z}_p$. Therefore, we can partition $\mathbb{Z}_p$ into two sets $\mathcal{T}_0 = \{z : \Pr[\mathcal{Z}_x^p = z] = a/s\}$ and $\mathcal{T}_1 = \{z : \Pr[\mathcal{Z}_x^p = z] = (a+1)/s\}$. Thus, $\mathsf{Dist}\,(\mathcal{Z}_x^p \| U(\mathbb{Z}_p)) = \max\{\sharp\mathcal{T}_0 \cdot (\frac{1}{p} - \frac{a}{s}), \sharp\mathcal{T}_1 \cdot (\frac{a+1}{s} - \frac{1}{p})\}$. Since $r = s - ap$, $\sharp\mathcal{T}_0 = p - r$ and $\sharp\mathcal{T}_1 = r$, we can conclude that $\mathsf{Dist}\,(\mathcal{Z}_x^p \| U(\mathbb{Z}_p)) = \max\{\frac{(p-r)r}{sp}, \frac{r(p-r)}{sp}\} = \frac{r(p-r)}{sp} \leq \frac{p}{4s} \leq \frac{n}{4\Phi(n)s}$. As $s \geq \frac{n-\kappa}{\kappa}$, it follows that $\mathsf{Dist}\,(\mathcal{Z}_x^p \| U(\mathbb{Z}_p)) \leq \frac{n}{4\Phi(n)} \cdot \frac{\kappa}{n-\kappa} = \frac{n}{n-\kappa} \cdot \frac{\kappa}{4\Phi(n)} \leq \frac{\kappa}{2\Phi(n)}$. □

D Optimal padding scheme

The choice of a good padding scheme in Prot. 1 is crucial, since it must be guarantee both security and relatively high value of ℓ. One may wonder whether there exist more efficient padding schemes for Protocol 1. Indeed, the padding used in Protocol 1 is suboptimal, but it is quite close to optimal bounds. More formally, assume that padded is an efficiently computable function that maps an input σ to $\mathsf{padded}(\sigma)$, independently from msgq. If the padding depends on msgq, then the scheme either uses generic homomorphic operations or is specially tailored for the concrete cryptosystem. The former corresponds to a new generic design of homomorphic oblivious transfer and the latter is not a general solution.

Secondly, we have to consider security. We say that a padding scheme padded is ε'*-secure* if for any two inputs σ_0 and σ_1, the statistical difference between $r \cdot U(n) + \mathsf{padded}(\sigma_0)$ and $r \cdot U(n) + \mathsf{padded}(\sigma_1)$ is less than ε' for any $r \neq 0$. It is easy to verify that Protocol 1 is sender-private if and only if the padding is ε'-secure. Under this restriction, we can state upper bounds for throughputs of secure padding schemes. More formally, let η denote the ratio between the transfer bandwidth (how many bits can be transferred by a single plaintext) and the theoretical limit (how many bits are in a single plaintext). For our padding, $\eta = \frac{\ell}{\log_2 n}$. For the sake simplicity, we will give the upper bound to η only in the case the plaintext order n is a multiple of two distinct primes, this result can be straightforwardly generalised.

Lemma 4. *Let n be a product of two primes $p < q$. If Protocol 1 is ε'-sender-private, then the throughput is bounded from above by*

$$\eta_* = \frac{\log_2 p - \log_2(1-2\varepsilon')}{\log_2 p + \log_2 q} \ .$$

Proof. W.l.o.g., assume that $N < p$ and that the input range of the padding scheme is $[\kappa] = \{1, \ldots, \kappa\}$. Let $\mathcal{Z}_x$ denote the output range of $\mathsf{padded}(x)$ for $x \in [\kappa]$, and let $\mathcal{Z}_x^p$ (resp., $\mathcal{Z}_x^q$) be the output range of $\mathsf{padded}(x)$ reduced modulo p (resp., q).

A malicious Receiver can choose ϱ so that $\mathsf{Dec}_{\mathsf{sk}}(c_i) \equiv qs_i + \mathsf{padded}(\sigma[i]) \pmod n$ and $\mathsf{Dec}_{\mathsf{sk}}(c_j) \equiv ps_j + \mathsf{padded}(\sigma[j]) \pmod n$ for $i \neq j$. But then $\mathsf{Extract}_{\mathsf{sk}}(\mathsf{msgq}) \notin [N]$ and the simulator gets no information about the output. Therefore, for any input pair $x, y \in [\kappa]$

$$\mathsf{Dist}\left(\mathcal{Z}_x^p \| \mathcal{Z}_y^p\right) \leq 2\varepsilon \quad \text{and} \qquad \mathsf{Dist}\left(\mathcal{Z}_x^q \| \mathcal{Z}_y^q\right) \leq 2\varepsilon \ ,$$

as all $\mathcal{Z}_x^p$, $\mathcal{Z}_y^p$, $\mathcal{Z}_x^q$ and $\mathcal{Z}_y^q$ are ε-close the output distribution of the simulator (modulo p or q). We prove upper bounds to the number κ of padding sets provided that $\mathsf{Dist}\left(\mathcal{Z}_x^p \| \mathcal{Z}_y^p\right) \leq \varepsilon$. The analysis for the other factor is symmetrical.

For $x \in [\kappa]$, let $\mathcal{R}_x = \mathbb{Z}_p \setminus \mathcal{Z}_x^p$. As $\mathcal{Z}_x^p$ and $\mathcal{Z}_1^p$ are ε-close, we get that $\Pr[z \leftarrow \mathcal{Z}_1^p : z \in \mathcal{R}_x] \leq \varepsilon$. Let $\mathcal{R}$ be the multi-set containing all elements of $\mathcal{R}_1, \ldots, \mathcal{R}_q$. Then the total probability mass of $\mathcal{R}$ with respect to the distribution over $\mathcal{Z}_1^p$ is less than $q\varepsilon$. Since the probability distributions $\mathcal{Z}_x^p$ and $\mathcal{Z}_1^p$ are ε-close, we can conclude that $\Pr[z \leftarrow \mathcal{Z}_1^p : z \in \mathcal{Z}_x^p] \geq 1 - \varepsilon$. The latter implies that there can be at most $\frac{q\varepsilon}{1-\varepsilon}$ additional sets $\mathcal{Z}_x,\ x \geq q$, because by the Chinese Remainder Theorem there exists one-to-one correspondence between the elements of $\mathcal{R}$ and the elements of $\mathbb{Z}_{pq}$ that are not covered by the sets $\mathcal{Z}_1, \ldots, \mathcal{Z}_q$. To summarise, if $\mathsf{Dist}\left(\mathcal{Z}_x^p \| \mathcal{Z}_y^p\right) \leq \varepsilon$ for all $x, y \in \{1, \ldots, \kappa\}$, then $\kappa \leq q + \frac{\varepsilon q}{1-\varepsilon} = \frac{q}{1-\varepsilon}$. Since ε-sender-privacy enforces that $\mathsf{Dist}\left(\mathcal{Z}_x^p \| \mathcal{Z}_y^p\right) \leq 2\varepsilon$, and analogously, the bound must also hold for the second prime factor, then the claim follows. □

Lemma 4 shows that the padding scheme of Protocol 1 is close to optimal if n is a product of two primes. For the standard parameters of Paillier' cryptosystem, the difference between η and η_* is roughly 20%. As encoding and decoding consist from bit-shifts and addition the padding is extremely efficient.