Verifiable Homomorphic Oblivious Transfer and Private Equality Test

Helger Lipmaa

Laboratory for Theoretical CS, Department of CS&E
Helsinki University of Technology, P.O.Box 5400, FIN-02015 HUT, Espoo, Finland
helger@tcs.hut.fi

Abstract. We describe slightly modified version (that we call the HOT protocol) of the Aiello-Ishai-Reingold oblivious transfer protocol from Eurocrypt 2001. In particular, the HOT protocol will be what we call weakly secure when coupled with many different homomorphic semantically secure public-key cryptosystems. Based on the HOT protocol, we construct an efficient verifiable oblivious transfer protocol and an efficient verifiable private equality test. As a concrete application of our results, we propose a novel protocol called proxy verifiable private equality test, and apply it to a cryptographic auction scheme to improve its security.
Keywords: cryptographic auctions, homomorphic encryption, verifiable oblivious transfer, verifiable private equality test.

1 Introduction

In a two-party $\binom{n}{1}$-*oblivious transfer (OT)* protocol the chooser receives a chosen single input from the database of n items, without the sender getting to know which element was retrieved. We first present a concise proof that a slightly modified version (that we call the *homomorphic oblivious transfer* or the HOT protocol) of the $\binom{n}{1}$-OT protocol of [AIR01] is perfectly sender-private iff for all possible private keys x of the used homomorphic semantically secure public-key cryptosystem, the corresponding plaintext space is a cyclic group of prime order M. Additionally, we show that the HOT protocol is computationally sender-private when M is composite but hard to factor by the chooser. This makes it possible to use the recent Damgård-Jurik cryptosystem [DJ03] in this context.

We then also introduce another security notion for oblivious transfer protocols, *weak sender-privacy*, that is sufficient whenever the oblivious transfer protocol does not have to be chooser-verifiable. Intuitively, a protocol is weakly sender-private if the chooser will never obtain information about more than one item from the database; however, the Chooser can still obtain information about a single item of the database even if his input to the protocol is out of the bounds. We show that the $\binom{n}{1}$-HOT protocol is weakly sender-private whenever $\mathcal{M}_{\Pi}(x)$ is a residue class ring with $\Phi(M) > n$, where $\Phi(M)$ is the smallest prime divisor of M. A weakly sender-private $\binom{n}{1}$-HOT protocol can be made sender-private by accompanying it with a zero-knowledge argument that chooser's input was in the correct range. In this case, some suitable homomorphic cryptosystems are [El 84,Pai99,DJ01,DJ03], and possibly [NS98,OU98]. Therefore, the $\binom{n}{1}$-HOT protocol can be based on different hardness assumptions (like the DCRA assumption of Paillier [Pai99]), made to work efficiently with long strings (in the case of Damgård-Jurik cryptosystems [DJ01,DJ03]), and efficiently thresholded (in the case of [El 84,DJ03]).

In a verifiable (also known as "committed" [CvdGT95,CD97,CC00]) oblivious transfer protocol, the chooser obtains sender's commitment to every database element and can later verify if these elements were equal to some other elements, used in other parts of the higher-level protocol. In the new verifiable homomorphic oblivious transfer protocol (Protocol 2), the chooser and the sender execute the HOT protocol so that the chooser obtains the random number that was used by the sender to commit to the chosen database element. Security of the verifiable HOT protocol depends additionally on the security of the employed homomorphic commitment scheme Γ, and on a simple relation between the sizes of plaintext spaces of Π and Γ. In particular, the verifiable HOT protocol based on the ElGamal cryptosystem and on the CGHN commitment scheme [CGHN01] is perfectly sender-private (unlike the recent slightly less efficient verifiable oblivious transfer protocol of [AJL03] that offers only statistical sender-privacy), and allows efficient reconstruction of the transmitted data item (unlike, again, [AJL03]).

After that, we show how to use the ideas, developed while constructing the HOT and the verifiable HOT protocols, in another context. *Private equality test (PET)* [FNW96,NP99,BST01] (let the Chooser to know whether the private inputs W_{Cho} and W_{Sen} of the Chooser and the Sender are equal without leaking any other information) is yet another widely used cryptographic protocol. We propose a new two-round homomorphic PET (HPET) protocol that is very similar to the $\binom{n}{1}$-HOT protocol. Previously known PET protocols [FNW96,NP99,BST01] were significantly less efficient. The HPET protocol is perfectly sender-private, when based on a homomorphic semantically secure public-key cryptosystem with a prime M like the ElGamal [El 84]. Computational privacy is achieved when the decrypter cannot factor M [DJ03]. As with the HOT protocol, we show how to make the HPET protocol verifiable, although the concrete technique for this will be different.

Finally, we propose a novel application for the new verifiable HPET protocol. Namely, we show that it can be generalised to the proxy verifiable HPET protocol and then use the latter to increase the security of the probably most efficient currently known ($(b+1)$st-price sealed-bid) cryptographic auction scheme without threshold trust by Lipmaa, Asokan and Niemi [LAN02]. More precisely, we show how to make the payment enforcement phase of [LAN02] more secure by not revealing the contract price either to the bidders or to the seller, before all the bidders have shown by using the proxy verifiable HPET protocol whether their bid was equal to the (yet unknown to them) value of the highest bid. We hope to see more applications of the proxy verifiable HPET protocol in the future, especially since to the best of our knowledge, no efficient proxy PET protocols were known previously at all.

All the proofs in this paper are slightly simplified due to the lack of space.

Road-map. We start the paper by describing cryptographic building blocks (Section 2). Section 3 defines some properties of the public-key cryptosystems that we need later. Our main contribution starts with Section 4, where we propose the new oblivious transfer protocols and prove their security. In Section 5, we describe a new private equality test protocol, together with some extensions. Finally, in Section 6 we propose some applications of the new protocols. In particular, we demonstrate how to use the proxy verifiable PET protocol in auctions.

2 Preliminaries and Cryptographic Building Blocks

Throughout this paper, let k be the security parameter. We assume that the reader knows standard complexity-theoretic notions like negligibility and probabilistic polynomial time (PPT); we take the latter to be equivalent to "efficiently computable". For a positive integer x, let $\varPhi(x)$ denote the smallest prime divisor of x. Let $\varphi(x)$ be the Euler's totient function of x. Recall that if $x = \prod_i p_i^{c_i}$ for different primes p_i then $\varphi(x) = x \cdot \prod_i (1 - 1/p_i)$.

For a distribution (random variable) X, let $x \leftarrow X$ denote the assignment of x according to X. We often identify sets with the uniform distributions on them, and algorithms with their output distributions, assuming that the algorithm that outputs this distribution is clear from the context or just straightforward to construct. The statistical difference of two distributions X and Y over the discrete support U is defined as $\Delta(X||Y) := \max_{S \subseteq U} |\Pr[X \in S] - \Pr[Y \in S]|$.

Homomorphic Semantically-Secure Cryptosystems. Let $\varPi = (\mathcal{G}_\varPi, E, D)$ be a public-key cryptosystem, where $\mathcal{G}_\varPi$ is the key generation algorithm $\mathcal{G}_\varPi : 1^k \mapsto (x, K)$, E is the encryption algorithm $E_K : (m; r) \mapsto E_K(m; r)$ and D is the decryption algorithm $D_K : c \mapsto D_K(c)$. Assume that for every possible private key x, the corresponding message space $\mathcal{M}_\varPi(x)$ is an Abelian group with the group operation $+$, and that the corresponding ciphertext space $\mathcal{C}_\varPi(x)$ is a Abelian group with the group operation $\cdot$. We denote the space of random coins by $\mathcal{R}_\varPi(x)$. (In particular, this notation indicates that $\mathcal{M}_\varPi(x)$, $\mathcal{R}_\varPi(x)$ and $\mathcal{C}_\varPi(x)$ might be unknown to the encrypter, although this is usually not the case.)

We say that $\varPi$ is *homomorphic*, if $E_K(m_1; r_1) \cdot E_K(m_2; r_2) = E_K(m_1 + m_2; r_1 \circ r_2)$ for some deterministic binary operation $\circ : \mathcal{R}_\varPi(x)^2 \to \mathcal{R}_\varPi(x)$. Then $E_K(m; r)^s = E_K(m^s; \mathsf{rf}_e(r, s))$ for another deterministic mapping rf_e. Given that $\mathsf{rf}_e(r, s+1) = \mathsf{rf}_e(r, s) \circ r$, we will denote $\mathsf{rf}_e(r, s)$ by r^s.

For an algorithm A, define $\mathsf{Adv}^{\mathsf{sem}}_{\varPi,k}(A) := \big|\Pr[(x, K) \leftarrow \mathcal{G}_\varPi(1^k), (m_0, m_1) \leftarrow A(1^k, K), r \leftarrow \mathcal{R}_\varPi(x), b \leftarrow [0, 1], c \leftarrow E_K(m_b; r) : A(1^k, K, m_0, m_1, c) = b] - \frac{1}{2}\big|$ to be the advantage that A has over random guessing when trying to distinguish random encryption of two elements, chosen by herself. We say that Π is *semantically secure* if for all PPT algorithms A, $\mathsf{Adv}^{\mathsf{sem}}_{\Pi,k}(A)$ is negligible in k. This definition is polynomially equivalent to other common definitions of semantical security.

A classical example of an homomorphic semantically secure public-key cryptosystem is the ElGamal public-key cryptosystem [El 84] with $E_K(m;r) = (mh^r; g^r)$; it works over any family of multiplicative groups where the Decisional Diffie-Hellman Assumption is true. In particular, $\mathcal{M}_\Pi(x)$ may be a subgroup of $\mathbb{Z}_p^*$, generated by an element of order q, where p and q are primes such that $q \mid (p-1)$. In another important case, $\mathcal{M}_\Pi(x)$ is a prime-order subgroup of a suitable elliptic curve group. Another example of an homomorphic semantically secure public-key cryptosystem is the Paillier public-key cryptosystem [Pai99], where as modified by [CGHN01,DJ01], $E_K(m;r) = (1+mN)r^N \mod N^2$ for $N = pq$, $\mathcal{M}_\Pi(x) = \mathbb{Z}_N$ and $\mathcal{R}_\Pi(x) = \mathbb{Z}_N^*$. Here, $E_K(m_1;r_1) \cdot E_K(m_2;r_2) = E_K(m_1+m_2; r_1r_2)$.

Homomorphic commitment schemes. In a commitment scheme $\Gamma = (\mathcal{G}_\Gamma, C)$, the committer sends an element $m \leftarrow \mathcal{M}_\Gamma(x)$ of the plaintext space to the receiver in a committed form, $c \leftarrow C_K(m;r)$, where (x,K) is generated by $\mathcal{G}_\Gamma(1^k)$ and $r \leftarrow \mathcal{R}_\Gamma(x)$. We denote the commitment space of Γ by $\mathcal{C}_\Gamma(x)$. In the context of our paper, all commitment schemes are required to be perfectly (or at least statistically) hiding and computationally binding. More precisely, for an algorithm A, define $\mathsf{Adv}^{\mathsf{hide}}_{\Pi,k}(A) := \bigl|\Pr[(x,K) \leftarrow \mathcal{G}_\Gamma(1^k), (m_0,m_1) \leftarrow A(1^k,K), r \leftarrow \mathcal{R}_\Gamma(x), b \leftarrow [0,1], c \leftarrow C_K(m_b;r) : A(1^k,K,m_0,m_1,c) = b] - \frac{1}{2}\bigr|$ to be the advantage that A has over random guessing when trying to distinguish random commitments of two elements, chosen by herself. We say that Γ is *statistically hiding* if for all (not necessarily PPT) algorithms A, $\mathsf{Adv}^{\mathsf{hide}}_{\Gamma,k}(A)$ is negligible in k. We allow Γ to be a trapdoor commitment scheme. That is, if A has access to the secret key x, she can break the binding property. Γ is *homomorphic* if for any (m_1,m_2,r_1,r_2), $C_K(m_1;r_1)C_K(m_2;r_2) = C_K(m_1+m_2; r_1 \circ r_2)$ for some binary operator $\circ$. We will sometimes assume that $\mathcal{R}_\Gamma(x)$ has a unit element 1.

In the Pedersen commitment scheme [Ped91], the setting is the same as in the ElGamal public-key cryptosystem, and $C_K(m;r) := g^m h^r$ for $r \in \mathcal{R}_\Gamma(x)$. In the CGHN [CGHN01] trapdoor commitment scheme, $N = pq$, $C_K(m;r,s) = (1+mN)r^N h^s \mod N^2$, where $h = \alpha^N(1+\beta N) \mod N^2$ for random $\alpha \leftarrow \mathbb{Z}_N^*$ and $\beta \leftarrow \mathbb{Z}_N \setminus \{0\}$, $r \leftarrow \mathbb{Z}_N^*$ and $s \leftarrow \mathbb{Z}_N$. Then $C_K(m_1;r_1,s_1)C_K(m_2;r_2,s_2) = C_K(m_1+m_2; r_1r_2, s_1+s_2)$.

$\binom{n}{1}$-Oblivious Transfer. During an $\binom{n}{1}$-oblivious transfer protocol, the chooser receives precisely one, chosen by himself, item from the database $\mu = (\mu_1, \dots, \mu_n)$ of n items, maintained by the sender. The sender does not get to know which item was transferred. In the general case, the index i in μ_i does not have to be an integer (indeed, we will not require it in the following), it is sufficient that different elements of μ are indexed by different elements of some set $\mathcal{I} = (\mathcal{I}_1, \dots, \mathcal{I}_n)$. However, for the sake of simplicity we will denote the ith element of the database by μ_i (and not by $\mu_{\mathcal{I}_i}$).

Importantly, most of the cryptography can be based on the oblivious transfer [Kil88]. Additionally, efficient oblivious transfer is necessary, since oblivious transfer is often the most expensive part of cryptographic protocols. An example is Yao's two-party computation model, where the proxy oblivious transfer [NPS99] is the only sub-protocol that requires public-key operations.

The security of an (information-theoretically sender-private) $\binom{n}{1}$-oblivious transfer protocol is usually defined in two parts. We will follow the definitions of [NP01, Section 2.1.1]. (It is possible to switch the security requirements so as to require information-theoretical chooser-privacy and computational sender-privacy, but corresponding protocols will be out of the scope of this paper. See, e.g., [Tze02].)

Denote a run of interactive protocol between A who has private input a and random tape r_a and between B who has private input b and a random tape r_b as $(A, B)[a, r_a; b, r_b]$. As usually, define Cho's view $\mathsf{view}_{\mathsf{Cho}}[\sigma, r_{\mathsf{Cho}}; \mu, r_{\mathsf{Sen}}]$ in the oblivious transfer protocol $(\mathsf{Cho}, \mathsf{Sen})[\sigma, r_{\mathsf{Cho}}; \mu, r_{\mathsf{Sen}}]$ as the concatenation of its private input σ, random tape r_{Cho}, the protocol transcript, and its private output μ_σ. The view of Sen is defined dually.

Computational Chooser-Privacy: For an algorithm A executing the sender's part in the oblivious transfer protocol $(\mathsf{Cho}, A)[\sigma, r_{\mathsf{Cho}}; \mu, r_A]$, define $\mathsf{Adv}^{\mathsf{otcho}}_{\mathsf{Cho},k}(A) := \Pr[(\sigma_0, \sigma_1, \mu') \leftarrow A(1^k, \mu, r_A), b \leftarrow [0, 1] : A(1^k, \mu, r_A, \mathsf{view}_A[\sigma_b, r_{\mathsf{Cho}}; \mu', r_A]) = b']$ to be the probability that after observing an execution of the protocol $(\mathsf{Cho}, A)[\sigma_b, r_{\mathsf{Cho}}; \mu, r_{\mathsf{Sen}}]$, A can predict which of the two possible choices σ_0 and σ_1 was used by the chooser. We call an oblivious transfer protocol *(computationally) chooser-private* if $\mathsf{Adv}^{\mathsf{otcho}}_{\mathsf{Cho},k}(A)$ is negligible for any PPT algorithm A.

Statistical Sender-Privacy: We make the comparison to the ideal implementation, using a trusted third party that receives μ from the sender, receives σ from the chooser, and tells the chooser μ_σ. We assume that μ_σ is garbage (i.e., a random value from some μ-independent set $\mathcal{T}$) if $\sigma \notin \mathcal{I}$.

We define the security by showing that for every algorithm A, one can define a simulator S that, given only private input σ, random tape r_A, and private output μ_σ of A, generates output that is statistically indistinguishable from the view of A that reacts with the honest sender Sen. More precisely, for a sender Sen and an algorithm S, define $\mathsf{Adv}^{\mathsf{otsen}}_{\mathsf{Sen},k}(A, S) := \Delta\left(S(1^k, \sigma, r_A, \mu_\sigma) || \mathsf{view}_A[\sigma, r_A; \mu, r_{\mathsf{Sen}}]\right)$. We say that the oblivious transfer protocol is *statistically sender-private* if for every (not necessarily PPT) A there exists a (not necessarily PPT) S, such that $\mathsf{Adv}^{\mathsf{otsen}}_{\mathsf{Sen},k}(A, S)$ is negligible in k. As usually, sender-privacy is perfect when $\mathsf{Adv}^{\mathsf{otsen}}_{\mathsf{Sen},k}(A, S) = 0$.

As argued, e.g., in [NP01, Section 2.1.2], an oblivious transfer protocol does not have to guarantee the correctness (even if Cho is honest but Sen is not, Cho will still receive Sen's input μ_σ). Following this convention, also we will leave it up to the application protocols to provide security in this sense.

The next $\binom{n}{1}$-oblivious transfer (OT) protocol by Aiello, Ishai and Reingold [AIR01] provides perfect sender-privacy and computational chooser-privacy. Assume that $\Pi = (\mathcal{G}_\Pi, E, D)$ is an homomorphic semantically secure public-key cryptosystem that works over a plaintext space $\mathbb{Z}_M$ of prime order $M = |\mathcal{M}_\Pi(x)|$. The sender Sen has a vector $\mu = (\mu_1, \dots, \mu_n) \in \mathbb{Z}_M^n$. The chooser Cho has made a choice σ. The AIR protocol works as follows: (a) Cho generates a secret/public key pair $(x, K) \leftarrow \mathcal{G}_\Pi(1^k)$. Cho generates a random coin $r \leftarrow \mathcal{R}_\Pi(x)$ and computes $c \leftarrow E_K(\sigma; r)$. He sends (K, c) to Sen. (b) Sen performs the following, for $i \in [1, n]$: Generate random $(r_i, s_i) \leftarrow \mathcal{R}_\Pi(x) \times \mathcal{M}_\Pi(x)$. Compute $c_i \leftarrow E_K(\mu_i; 0) \cdot (c \cdot E_K(-i; 0))^{s_i} \cdot E_K(0; r_i)$. Send c_i to Cho. (c) Cho obtains $\mu_\sigma \leftarrow D_K(c_\sigma)$. As a consequence, the AIR protocol requires n online encryptions by the sender. A similar but slightly less efficient $\binom{n}{1}$-OT protocol was independently proposed by Naor and Pinkas [NP01, Section 4.1].

Often one needs a $\binom{n}{1}$-oblivious transfer protocol to be *sender-verifiable* (also known as "committed") in the next sense [CvdGT95,CD97,AJL03]: after the oblivious transfer protocol, the chooser obtains sender's commitment c_i to every database element that can be later used in various zero-knowledge proofs or arguments. Recently, Ambainis, Jakobsson and Lipmaa proposed probably the first two-round verifiable oblivious transfer protocol [AJL03]; their protocol was based on decoupling the Naor-Pinkas oblivious transfer protocol and the Pedersen commitment scheme. Briefly, the Naor-Pinkas protocol uses a sub-protocol to recover a key that was used to encrypt the database element. The Ambainis-Jakobsson-Lipmaa (AJL) protocol uses the same sub-protocol to recover a nonce that was used to commit to the database element.

Private Equality Test. At the end of the private equality test (PET, also known as "comparing information without leaking it" or "socialist millionaires's problem") protocol, the Chooser Cho gets to know whether Sender's input W_{Sen} equals to that of the Chooser, W_{Cho}. Cho will not get to know anything else about W_{Sen}, while Sen should not have any private output at all. Exactly as in the case of oblivious transfer, the security is divided into statistical sender-privacy and computational chooser-privacy. The security definitions are standard and we omit them due to the space constraints.

Previously proposed PET protocols [FNW96,NP99,BST01] had an extra emphasis on developing fair protocols where both the Chooser and the Sender get to know the result of comparison. None of these protocols is however really efficient even when simplified so as not to have the fairness property. For example, the PET protocol from [BST01] requires multiple rounds and zero-knowledge proofs of knowledge. One application, considered at the end of our paper actually relies on the asymmetric nature of our PET protocols.

3 Affine Public-Key Cryptosystems

Next we describe a new property of homomorphic semantically secure public-key cryptosystems that will be necessary in the later described protocols. First, recall that a finite cyclic Abelian group is isomorphic to some residue class group $\mathbb{Z}_N$. Now, let $\mathcal{D}$ and $\mathcal{D}' \neq 0$ be two distributions of elements of $\mathbb{Z}$. We say that $\mathcal{D}'$ *affinely ε-approximates* $\mathcal{D}$ on additive group G if for every $g, g' \in G$, $g \neq 0$, $\Delta\left(\mathcal{D}' \cdot g + g' || \mathcal{D}\right) \leq \varepsilon$. We call G *ε-affine* if such distributions $\mathcal{D}$ and $\mathcal{D}'$ exist. We say that G is computationally ε-affine if it is ε-affine under the condition that g and g' must be generated by a PPT algorithm. We say that G is (computationally) non-affine if it is not (computationally) 1/2-affine.

Assume that the order of G is public. First, if G is a cyclic group of prime order, one can define $\mathcal{D}' := |G|$ and $\mathcal{D} := G$. Then G is 0-affine. If G is a cyclic group of composite order, $G \cong \mathbb{Z}_M$, then for any generator g of G, all elements ag for $\gcd(a, |G|) = 1$ are generators, while for a with $\gcd(a, |G|) \neq 1$, $|\langle ag \rangle| \leq |G|/2$. Therefore, G is non-affine. On the other hand, if one assumes that it is hard to factor $|G|$ then G will be computationally 0-affine. If G is an acyclic group, then every element $g \in G$ generates a nontrivial subgroup $\langle g \rangle$ of G of order $\leq G/2$. In this case, any choice of $\mathcal{D}' \neq 0$ leads to non-affinity even in the computational sense.

Let $\varepsilon = (\varepsilon_k)$ be a family of probabilities. We say that $\Pi = (\mathcal{G}_\Pi, E, D; \mathcal{S}, \mathcal{T})$ is an *ε-affine public-key cryptosystem*, if $\Pi' = (\mathcal{G}_\Pi, E, D)$ is a homomorphic semantically secure public-key cryptosystem, $\mathcal{S}$ and $\mathcal{T}$ are PPT algorithms, with $\mathcal{S}(1^k, K) \subset \mathbb{Z}$, $\mathcal{T}(1^k, K) \subseteq \mathcal{M}_\Pi(x)$ with $|\mathcal{T}(1^k, K)| > 1$, and for every security parameter k, key pair $(x, K) \in \mathcal{G}_\Pi(1^k)$, $\mathsf{Adv}^{\mathsf{affine}}_{\Pi,x} := \max_{a \in \mathcal{M}_\Pi(x) \setminus \{0\}, b \in \mathcal{M}_\Pi(x)} \Delta\left(\mathcal{S}(1^k, K)a + b || \mathcal{T}(1^k, K)\right) \leq \varepsilon_k$. Therefore, Π is *perfectly affine* if for every x, $\mathcal{M}_\Pi(x)$ is a cyclic group with known prime order. We say that Π is *computationally affine* if for every x, $\mathcal{M}_\Pi(x)$ is a cyclic group with known composite order under the assumption that it is hard even for the decrypter to factor M. (If M is not known, perfect affinity may change to statistical affinity.)

4 Homomorphic Oblivious Transfer Protocols

Simplified notation. To simplify the notation, from now on we will omit the arguments $(1^k, K)$ of $\mathcal{S}$ and $\mathcal{T}$, the argument x of $\mathcal{M}_\Pi$ and $\mathcal{R}_\Pi$, and the argument $\tilde{x}$ of $\mathcal{M}_\Gamma$ and $\mathcal{R}_\Gamma$.

4.1 Simpler Protocol without Sender-Verifiability

Protocol 1 depicts the new homomorphic oblivious transfer protocol. A very similar protocol was proposed in [AIR01]; we will provide comparisons later in this section.

Theorem 1. *Let k be the security parameter. Let $\Pi = (\mathcal{G}_\Pi, E, D; \mathcal{S}, \mathcal{T})$ be a (statistically or computationally) ε-affine homomorphic semantically secure public-key cryptosystem for some $\varepsilon = (\varepsilon_k)_k$. Let the database size n be polynomial in k. The HOT protocol depicted by Protocol 1 is a secure oblivious transfer protocol between the chooser* Cho *and the sender* Sen *in the next sense. When*

Protocol 1 The homomorphic oblivious transfer protocol

Private input: Cho has an index $\sigma \in \mathcal{I}$, Sen has $\mu = (\mu_1, \dots, \mu_n)$.
Private output: Cho has μ_σ.

1. The chooser generates a new key pair $(x, K) \leftarrow \mathcal{G}_\Pi(1^k)$, a random coin $r \leftarrow \mathcal{R}_\Pi$, and sets $c \leftarrow E_K(\mathcal{I}_\sigma; r)$. He sends (K, c) to the sender.
2. For $i \in [1, n]$, the sender chooses random $s_i \leftarrow \mathcal{S}$ and $r_i \leftarrow \mathcal{R}_\Pi$, computes $c_i \leftarrow E_K(\mu_i; 0) \cdot (c \cdot E_K(-\mathcal{I}_i; 0))^{s_i} \cdot E_K(0; r_i)$, and sends c_i to the chooser.
3. The chooser outputs $\mu_\sigma \leftarrow D_K(c_\sigma)$.

Π is semantically secure, then the HOT is computationally chooser-private. Let $M = |\mathcal{M}_\Pi|$. Sender's privacy is (a) perfect when $\varepsilon_k = 0$, (b) computational, with the best adversary having success $n\varepsilon_k$ when ε_k is negligible in k and Π is computationally ε-affine.

Proof. Correctness: If both players are honest then $c_i = E_K(\mu_i + s_i(\sigma - i)); r^{s_i} \circ r')$ and $D_K(c_\sigma) = \mu_\sigma$, and thus this protocol is correct.

Chooser-privacy: If the sender can distinguish the views $\{E_K(\sigma; \mathcal{R}_\Pi)\}$ and $\{E_K(\sigma'; \mathcal{R}_\Pi)\}$ then Π is not semantically secure. (More precisely, if one can violate the chooser-privacy in time t with probability δ, then one can violate the semantical security of Π in time $t + \mathit{const}$ and with probability δ.)

Statistical sender-privacy: We construct the next unbounded simulator S of A: S executes A instruction-by-instruction, except that when A sends a message c to the sender Sen, S interrupts and answers to c with $(c_1, \dots, c_n)$, where c_i is computed as follows: if $i := D_K(c) \in \mathcal{I}$ then $c_i \leftarrow c^{s_i} \cdot E_K(\mu_i - s_i D_K(c); \mathcal{R}_\Pi)$ for random $s_i \leftarrow \mathcal{S}$, otherwise $c_i \leftarrow E_K(\mathcal{T}; \mathcal{R}_\Pi)$.

Now, if $\sigma := D_K(c) \notin \mathcal{I}$ (the opposite case $D_K(c) \in \mathcal{I}$ is analogous), the output distribution of the simulator (for fixed random tape ρ of S, and for fixed c) is $(\rho; c; \dots, E_K(\mathcal{T}; \mathcal{R}_\Pi), \dots; \mu_\sigma)$, while the output distribution of A is $(\rho; c; \dots, c^{s_i} \cdot E_K(\mu_i - s_i \mathcal{I}_\sigma; \mathcal{R}_\Pi), \dots; \mu_\sigma)$ for random $s_i \leftarrow \mathcal{S}$. For a fixed c, the difference between these two distributions is $\mathsf{Adv}^{\mathsf{otsen}}_{\mathsf{Sen},k}(A, S) \leq n \cdot \max_{a \neq 0} \Delta\left(E_K(\mathcal{S}a + \mu_i; \mathcal{R}_\Pi) || E_K(\mathcal{T}; \mathcal{R}_\Pi))\right) \leq n \cdot \max_{a \neq 0, b} \Delta\left(E_K(\mathcal{S}a + b; \mathcal{R}_\Pi) || E_K(\mathcal{T}; \mathcal{R}_\Pi)\right) = n \cdot \max_{a \neq 0, b} \Delta\left(\mathcal{S}a + b || \mathcal{T}\right) = n \cdot \mathsf{Adv}^{\mathsf{affine}}_{\Pi, x}$. Both claims follow straightforwardly. □

Weak Server-Privacy. Only a few homomorphic semantically secure public-key cryptosystems are affine, as seen from Table 1. Fortunately, it comes out that the HOT protocol is sender-private under much broader settings when we slightly weaken the security definitions.

We say that the oblivious transfer protocol provides *weak sender-privacy* if the chooser will retrieve more than an ideal amount of information about at most one value μ_i, where $i = \sigma$ when the Chooser has private input $\sigma \in \mathcal{I}$. Weak sender-privacy is sufficient in almost all cases when the oblivious transfer protocol is not required to be chooser-verifiable. (Chooser-verifiability can be defined as the requirement that the chooser must be able to prove that the database element she received was indexed by her choice.) An example application where weak sender-privacy is sufficient is the paid database queries setting, where the database maintainer is only interested in the number of the items that the client will obtain, and not that the indices of the obtained items satisfy any requirements.

Often (as in the case of the oblivious transfer protocol, proposed in Sect. 4), a weakly sender-private oblivious transfer protocol can be transfered to a statistically sender-private one by accompanying it with a suitable zero-knowledge proof (or argument) that $\sigma \in \mathcal{I}$. Importantly, as we will see from the next theorem, there exist settings where the new oblivious transfer protocol is weakly sender-private but not statistically sender-private.

Theorem 2. *Assume the same setting as in Theorem 1. Additionally, assume that $\mathcal{M}_\Pi$ is a cyclic group with a generator g, $\mathcal{I}_i = ig$ and that $\Phi(M) > n$. Then the HOT protocol is weakly sender-private. Moreover, a statistically weakly sender-private HOT protocol can be made statistically sender-private if before the second step of Protocol 1, the chooser argues in statistical zero-knowledge that c is an encryption of σ for some $\sigma \in \mathcal{I}$.*

Proof (Sketch). As in Theorem 1, the advantage $\mathsf{Adv}^{\mathsf{otsen}}_{\mathsf{Sen},k}(A, S)$ is bound by $n \max_{a \neq 0, b} \Delta\left(\mathcal{S}ag + b || \mathcal{T}\right)$. Define $\mathcal{S} := \mathbb{Z}_M$ and $\mathcal{T} := \mathcal{M}_\Pi$. When $a = (\pm\sigma)g$ for $\gcd(\sigma, M) = 1$ then $\mathcal{S}a + b = \mathcal{T}$ for any b. If $a = \sigma g$ for $\gcd(\sigma, M) \neq 1$ then the chooser will see $n-1$ random encryptions that are distributed as $E_K(\mathcal{T}; \mathcal{R}_\Pi)$, and one encryption of a value $E_K(\mu_i + \mathcal{S}(\sigma - i)g; \mathcal{R}_\Pi)$, this is since $\Phi(M) \geq n$. From the latter she might be able to derive some information about μ_i but this is allowed by the security definition.

The second claim of the theorem (about the zero-knowledge argument) is straightforward. Moreover, if $\mathcal{I}_i$ is encoded as g^i for some group element $g \in \mathcal{M}_\Pi$ then one can show efficiently that $j \in \mathcal{I}$ by using protocols from [DJ01,LAN02]; for $\mathcal{I}_i = i$ the corresponding proofs can be found from [Bou00,Lip03]. (See [Lip03] for some other possible encodings.) □

Comparison with [AIR01]. The HOT protocol is a generalisation of the protocol of Aiello, Ishai and Reingold [AIR01, Section 5] to a wider selection of plaintext spaces. (Namely, [AIR01] considered only the case when M is a prime.) Careful specification of parameters and the definition of affine cryptosystems allowed us to prove that the protocol is "almost" as secure in cases, not considered in [AIR01]. In particular, as argued earlier, weak sender-privacy is sufficient always when one does not require chooser-verifiability. In most of the real-life scenarios, one does not require chooser-verifiability; in almost all such cases, one can use weakly sender-private variants of the HOT protocol that were not considered in [AIR01]. However, when chooser-verifiability is needed, one will also usually need sender-verifiability, a property not provided by HOT protocol and thus also not by the AIR protocol from [AIR01]. (See Section 4.2 for a new sender-verifiable oblivious transfer protocol.)

Table 1. Some homomorphic semantically secure public-key cryptosystems Π that make the HOT protocol at least weakly sender-private. The middle column shows whether the corresponding PET protocol from Section 5 is secure

Π	Sender-privacy	Weak sender-privacy
Sender-private HOT		
[El 84]	Yes (perfect)	Yes (perfect)
[DJ03]	Yes (computational)	Yes (perfect)
Weakly sender-private HOT		
[Pai99]	No	Yes (perfect)
DJ01 [DJ01]	No	Yes (perfect)
[NS98]	No	If $\Phi(M)$ is large (perfect)
[OU98]	No	If $\Phi(p-1)$ is large (statistical)

Discussion. Importantly, one has quite a flexible choice between possible underlying homomorphic semantically secure public-key cryptosystem Π when one only goes for the weak sender-security. Table 1 shows that the HOT is weakly sender-private based on most of the widely known homomorphic semantically secure public-key cryptosystems, and statistically sender-private when based on two known homomorphic semantically secure public-key cryptosystems. From the mentioned homomorphic semantically secure public-key cryptosystems, [NS98] offers a flexible choice of the value $\Phi(M)$ in the range $[3, 2^{11}]$, and for other public-key cryptosystems, $\Phi(M)$ is anyways required to be large for the public-key cryptosystem to be semantically secure. (However, it is not known whether the Naccache-Stern cryptosystem is semantically secure if M is known to Sen.) The Okamoto-Uchiyama public-key cryptosystem [OU98] is a notable exception since there M is not public, and $\Phi(M)$ is not required to be large. Still, even in this case one gets statistical weak sender-privacy by choosing $\mathcal{S} = \mathbb{Z}_{2^{k+\ell/2}}$, where ℓ is the key length.

If combined with the Damgård-Jurik cryptosystem from [DJ03], it becomes possible to use extremely large message spaces. If combined with the ElGamal cryptosystem, one can easily distribute the role of the sender. From the strictly efficiency point of view, the best underlying homomorphic semantically secure public-key cryptosystem would be the ElGamal based on (say) elliptic curves and $\mathcal{I}_i$ is defined as g^i for some generator g. Then $c \leftarrow (g^\sigma h^r, g^r)$ and $c_i \leftarrow (\mu_i g^{(\sigma-i)s_i} h^{rs_i+r_i}; g^{rs_i+r_i})$.

4.2 Verifiable HOT Protocol

Protocol 1 by itself is not (sender-)verifiable but it can be made verifiable by borrowing some ideas from the recent AJL verifiable oblivious transfer protocol by Ambainis, Jakobsson and Lipmaa [AJL03]. More precise, we use the HOT protocol so that the chooser obtains a random nonce m_σ that is used also when the sender commits to μ_σ. The chooser will thus only be able to recover the value of μ_σ. On the other hand, for every i, the sender commits to μ_i, using a

Protocol 2 The verifiable HOT protocol

PRIVATE INPUTS: Cho has σ, Sen has μ.
PRIVATE OUTPUTS: Cho obtains μ_σ.

1. Cho creates a key pair $(\tilde{x}, \tilde{K}) \leftarrow \mathcal{G}_\Gamma(1^k)$ and a key pair $(x, K) \leftarrow \mathcal{G}_\Pi(1^k)$. Cho creates a random $r \leftarrow \mathcal{R}$ and computes $c \leftarrow E_K(\mathcal{I}_\sigma; r)$. He sends $(K, \tilde{K}, c)$ to Sen.
2. For all i, Sen creates random $r_i \leftarrow \mathcal{R}$ and $(m_i, s_i) \leftarrow \mathcal{T} \times \mathcal{S}$, computes $v_i \leftarrow (c \cdot E_K(-\mathcal{I}_i; 0))^{s_i} \cdot E_K(m_i; r_i)$ and $c_i \leftarrow C_{\tilde{K}}(\mu_i; \mathsf{tr}(m_i))$. She sends (v_i, c_i) to Cho.
3. Cho outputs $\tilde{\mu}_\sigma \leftarrow \mathsf{retrieve}(c_\sigma \cdot C_{\tilde{K}}(0; \mathsf{tr}(D_K(v_\sigma))^{-1}))$.

random value $\mathsf{tr}(m_i)$ that is known to her. This means that she can use standard zero-knowledge techniques to prove properties of μ_i even for $i \neq \sigma$.

Theorem 3. *Let k be the security parameter. Assume that $\Pi = (\mathcal{G}_\Pi, E, D; \mathcal{S}, \mathcal{T})$ is an ε-affine homomorphic semantically secure public-key cryptosystem and that Γ is a homomorphic perfectly hiding commitment scheme. For fixed $(x, K) \leftarrow \mathcal{G}_\Pi(1^k)$ and $(\tilde{x}, \tilde{K}) \leftarrow \mathcal{G}_\Gamma(1^k)$, assume the existence of two deterministic PPT functions $\mathsf{tr} : \mathcal{M}_\Pi \to \mathcal{R}_\Gamma$ and $\mathsf{retrieve} : C_{\tilde{K}}(m; 1) \mapsto m$. Then Protocol 2 is (a) perfectly sender-private if Γ is perfectly hiding, tr is an injection, $|\mathcal{M}_\Pi| = |\mathcal{R}_\Gamma|$ is a prime and $\mathcal{T}$ and $\mathcal{S}$ are defined as usually; (b) statistically sender-private if Γ is statistically hiding, $(|\mathcal{M}_\Pi| - |\mathcal{R}_\Gamma|)/|\mathcal{R}_\Gamma|$ is negligible and tr is a suitable mapping.*

Proof. CORRECTNESS: If parties are honest then $v_i = E_K(s_i(\mathcal{I}_\sigma - \mathcal{I}_i) + m_i; s_i r + r_i)$, $c_i = C_{\tilde{K}}(\mu_i; \mathsf{tr}(m_i))$ and thus $D_K(v_\sigma) = m_\sigma$, $\tilde{\mu}_\sigma = \mathsf{retrieve}(c_\sigma \cdot C_{\tilde{K}}(0; \mathsf{tr}(m_\sigma) \cdot \mathsf{tr}(m_\sigma)^{-1}) = \mathsf{retrieve}(C_{\tilde{K}}(\mu_\sigma; 1)) = \mu_\sigma$.

CHOOSER-PRIVACY: straightforward, given that Π is semantically secure.

SENDER-PRIVACY: Assume $D_K(c) = \mathcal{I}_\sigma$ for $\sigma \in [1, n]$ (the opposite case is analogous). Denote the distribution $C_{\tilde{K}}(\mathcal{M}_\Gamma; \mathcal{R}_\Gamma)$ by Z and the distribution $((E_K(\tilde{m} + \mathcal{S}(\mathcal{I}_\sigma - \mathcal{I}_i); \mathcal{R}_\Pi), C_{\tilde{K}}(\mu_i; \mathsf{tr}(\tilde{m})))$, where $\tilde{m} \leftarrow \mathcal{T}$, by Y_i. We construct the next unbounded simulator S for A: S executes A step-by-step, except that when A makes a query c to the sender Sen, S interrupts and answers it with $(v_1, c_1, \ldots, v_n, c_n)$, where (v_i, c_i) is computed as follows: $(v_i, c_i) \leftarrow (E_K(\mathcal{T}; \mathcal{R}_\Pi), Z)$ when $i \neq \sigma$, and $(v_i, c_i) \leftarrow Y_\sigma$ when $i = \sigma$.

Then the advantage of A is

$$\begin{aligned}
\mathsf{Adv}^{\mathsf{otsen}}_{\mathsf{Sen}}(k)(A, S) \leq & \sum_{i \neq \sigma} \Delta\left(Y_i || (E_K(\mathcal{T}; \mathcal{R}_\Pi), Z)\right) \\
\leq & \sum_{i \neq \sigma} \max_{a \neq 0, b} \Delta\left((\mathcal{S}a + b, C_{\tilde{K}}(\mu_i; \mathsf{tr}(\mathcal{T}))) || (\mathcal{T}, Z)\right) \\
\leq & \sum_{i \neq \sigma} \max_{a \neq 0, b} \Delta\left(\mathcal{S}a + b || \mathcal{T}\right) + \sum_{i \neq \sigma} \Delta\left(C_{\tilde{K}}(\mu_i; \mathsf{tr}(\mathcal{T})) || Z\right) \\
\leq & n \cdot \left(\mathsf{Adv}^{\mathsf{affine}}_{\Pi, x} + \Delta\left(\mathsf{tr}(\mathcal{T}) || \mathcal{R}_\Gamma\right) + \mathsf{Adv}^{\mathsf{hide}}_{\Gamma, k}(A)\right) \enspace .
\end{aligned}$$

The claim follows. □

Table 2. Comparison of some verifiable oblivious transfer protocols, with specified homomorphic semantically secure public-key cryptosystem Π and homomorphic commitment scheme Γ. Here we have always $\mathcal{T} = \mathcal{S} = \mathbb{Z}_{|\mathcal{R}_\Gamma|}$ and thus $\mathsf{tr}(m) = m$.

Π	Γ	Sender's priv.	$\mathsf{retrieve}(c)$	Verifiable	Online work (exp/enc/comm)
Naor-Pinkas [NP01]					
ElGamal	(Pedersen)	Perfect	Easy (decryption)	No	$4n/n/-$
AIR [AIR01] and HOT (this paper)					
ElGamal	—	Perfect	Easy (decryption)	No	$-/n/-$
Ambainis-Jakobsson-Lipmaa [AJL03]					
ElGamal	Pedersen	Statistical	Hard (DL)	Yes	$4n/n/n$
Verifiable HOT (this paper)					
ElGamal	Pedersen	Perfect	Hard (DL)	Yes	$-/n/n$
ElGamal	CGHN	Statistical	$(c-1)/N \mod N^2$	Yes	$-/2n/n$

Straightforwardly, for the weak sender-privacy it suffices to replace the requirement that $\mathsf{Adv}^{\mathsf{affine}}_{\Pi,x}$ is negligible in k by the requirement that $\Phi(\mathcal{M}_\Pi) > n$.

Comparison with previous work. Recall that the up to now most efficient (and the only two-round) verifiable oblivious transfer protocol by Ambainis-Jakobsson-Lipmaa protocol [AJL03] was statistically private, and at the end of the AJL protocol, the chooser had to compute discrete logarithm to recover the value of μ_σ. The verifiable HOT protocol from Protocol 2 solves either—but not both—of these problems, when based on suitable Π and Γ. See Table 2 for a comparison of the verifiable HOT protocol (with the ElGamal cryptosystem but different Γ) with some previous work.

When Γ is the Pedersen commitment scheme with $x = \tilde{x}$ and $K = \tilde{K}$, and $\mathcal{I}_i = g^i$ for some generator g, the resulting scheme will be somewhat similar to [AJL03] with $v_i = (g^{s_i(\sigma-i)} m_i h^{s_i r + r_i}, g^{s_i r + r_i})$. Then $\mathcal{R}_\Gamma = \mathcal{M}_\Pi = \mathbb{Z}_q$, tr is the identity function, $\mathcal{S} = \mathcal{T} = \mathbb{Z}_q$, and the resulting protocol will be both computationally chooser-private and perfectly sender-private under the DDH assumption. (Recall that the AJL protocol from [AJL03] was only statistically sender-private.) Similarly to [AJL03], the drawback of this protocol is that the chooser obtains $C_{\tilde{K}}(\mu_\sigma; 0) = g^{\mu_\sigma}$, from which he has to recover μ_σ by computing a discrete logarithm.

The use of the CGHN [CGHN01] trapdoor commitment scheme as Γ enables one to get rid of the latter drawback with the cost of making the protocol only statistically sender-private. Recall that in the CGHN commitment scheme the chooser recovers $\tilde{c}_\sigma = C_{\tilde{K}}(\mu_\sigma; 1) = (1 + \mu_\sigma N) \mod N^2$, from which he can efficiently compute $\mu_\sigma = (\tilde{c}_\sigma - 1)/N \mod N^2$. However, in this case $|\mathcal{R}_\Gamma| \approx |\mathcal{M}_\Pi|^2$, assuming that the public keys of Π and Γ have the same length. There are at least three different methods for overcoming this obstacle: (a) Choosing twice longer keys for the public-key cryptosystem, so that $|\mathcal{M}_\Pi| \geq |\mathcal{R}_\Gamma| \approx N^2$; this

Protocol 3 New PET protocol, where $\Pi = (\mathcal{G}_\Pi, E, D; \mathcal{S}, \mathcal{T})$ is an affine homomorphic semantically secure public-key cryptosystem

Private inputs: Chooser has W_{Cho}, Sender has W_{Sen}.
Private outputs: Chooser has 0 if $W_{\mathsf{Cho}} = W_{\mathsf{Sen}}$ or garbage, otherwise.

1. Chooser generates a new key pair $(x, K) \leftarrow \mathcal{G}_\Pi(1^k)$, a random $r \leftarrow \mathcal{R}_\Pi$, and sets $c \leftarrow E_K(W_{\mathsf{Cho}}g; r)$. He sends (K, c) to Sender.
2. Sender generates random $s \leftarrow \mathcal{S}$ and $r' \leftarrow \mathcal{R}_\Pi$. She sends $c' \leftarrow (c \cdot E_K(-W_{\mathsf{Sen}}g; 0))^s \cdot E_K(0; r')$ to the Chooser.
3. Chooser accepts that $W_{\mathsf{Cho}} = W_{\mathsf{Sen}}$ iff $D_K(c') = 0$.

might however be impractical; (b) Setting tr to be a pseudorandom number generator; this results in a mere computational privacy; (c) Letting Sen to generate two different random numbers m_i and m_i', and to use the HOT protocol twice so that the Chooser obtains both m_i and m_i', and then use both to commit to μ_i. In all three cases, $\mathsf{Adv}^{\mathsf{otsen}}_{\mathsf{Sen}}(k)(A, S) \le 2n\Delta\left(\tau(\mathcal{T})||\mathcal{R}_\Gamma\right)$ is negligible. We suggest, even if this results in a slightly less efficient protocol, to use the third recommendation.

5 Private Equality Test and Enhancements

The Homomorphic Private Equality Test Protocol. Assume that a possible wealths W is encoded as Wg for a generator g of the cyclic group $\mathcal{M}_\Pi$. (Other encodings might also work) The new homomorphic private equality test (HPET) protocol (Protocol 3) is in a sense just a—although not a straightforward—simplification of the HOT protocol. Namely, it corresponds to the conditional disclosure of a single element $\mu_{W_{\mathsf{Sen}}} = 0$, where instead of $i = W_{\mathsf{Sen}}$, the sender uses $i = W_{\mathsf{Cho}}$. Thus, $\mu_{W_{\mathsf{Sen}}} = 0$ will be revealed only when $W_{\mathsf{Sen}} = W_{\mathsf{Cho}}$; otherwise the chooser will obtain a random element of $\mathcal{M}_\Pi$. Therefore, unsurprisingly, the PET protocol is sender-private exactly when based on a Π that also makes the HOT protocol sender-private.

Theorem 4. *Let k be the security parameter. Assume that $\Pi = (\mathcal{G}_\Pi, E, D; \mathcal{S}, \mathcal{T})$ is an ε-affine homomorphic semantically secure public-key cryptosystem, such that it is computationally hard for the decrypter to factor $M \leftarrow |\mathcal{M}_\Pi|$ for any $x \leftarrow \mathcal{G}_\Pi(1^k)$.*

Let $W_{\mathsf{Sen}} \in \mathcal{M}_\Pi$ and $W_{\mathsf{Cho}} \in \mathcal{M}_\Pi$ be Sender's and Chooser's inputs. Let $\mathcal{M}_\Pi$ be a cyclic group with generator g. Then Protocol 3, denoted as HPET, is chooser-private. Moreover, (a) if $\mathcal{M}_\Pi$ is a cyclic group of public prime order, then the HPET protocol is perfectly correct and sender-private, and (b) if $\mathcal{M}_\Pi$ is a cyclic group of public composite order, where it is hard for the chooser and the sender to factor $|\mathcal{M}_\Pi|$, then this protocol is computationally correct and sender-private.

Proof. CORRECTNESS: When both parties are honest then $c' = E_K(s(W_{\mathsf{Cho}} - W_{\mathsf{Sen}})g; r^s \circ r')$. Thus, $m = 0$ iff (a) $W_{\mathsf{Sen}} = W_{\mathsf{Cho}}$ or (b) $M \mid s(W_{\mathsf{Cho}} - W_{\mathsf{Sen}})g$. The latter can only happen when $\gcd(s(W_{\mathsf{Cho}} - W_{\mathsf{Sen}}), M) \neq 1$, that is, when M is composite, and either the chooser or the sender can find factors of M. (As previously, we will not care about correctness in the case when Sender is dishonest, leaving it up to an higher level protocol to deal with that.) CHOOSER-PRIVACY: follows straightforwardly from the semantical security.

STATISTICAL SENDER-PRIVACY (Sketch): In this case, the simulator S knows an answer to the question $W_{\mathsf{Sen}} \overset{?}{=} W_{\mathsf{Cho}}$ and nothing more about the Sender's wealth. He answers the query c with c', distributed as $E_K(\mathcal{T}; \mathcal{R}_\Pi)$, if $D_K(c) \neq W_{\mathsf{Sen}}$, and as $E_K(0; \mathcal{R}_\Pi)$ if $D_K(c) = W_{\mathsf{Sen}}$. Clearly, the difference between S's output and the real view is $\leq \Delta\left(E_K(\mathcal{S}(W_{\mathsf{Cho}} - W_{\mathsf{Sen}})g; \mathcal{R}_\Pi) || E_K(\mathcal{T}; \mathcal{R}_\Pi)\right) \leq \mathsf{Adv}^{\mathsf{affine}}_{\Pi,x}$. □

The HPET protocol is severely more efficient than the BST (Boudot-Schoenmakers-Traoré) protocol [BST01] or the protocol from [NP99]. However, the later can be modified (with significant cost in efficiency) so as to provide fairness, i.e., to guarantee that the Sender will only get to know whether $W_{\mathsf{Sen}} = W_{\mathsf{Cho}}$ if also the Chooser will get to know that. It is unclear yet if our protocol can be modified to become fair, but this is also not our intention.

Unfortunately, the number of currently known homomorphic cryptosystems where the decryption can be performed without knowing the factorisation of $|\mathcal{M}_\Pi|$ is small: the only known examples are [El 84,DJ03]. (See the second column of Tbl. 1.)

Verifiable PET. (Sketch.) Here, we use the same notation as in previous theorems. In a verifiable PET protocol, the Chooser sends $c \leftarrow E_K(W_{\mathsf{Cho}}; r)$ to the Sender, who replies with (v, c'), where $v \leftarrow E_K(s(W_{\mathsf{Cho}} - W_{\mathsf{Sen}})g + m; r^s \circ r')$ and $c' \leftarrow C_{\tilde{K}}(W_{\mathsf{Sen}} \cdot \tilde{g}; \mathsf{tr}(m))$, for $m \leftarrow \mathcal{T}$. Here, $\mathsf{tr} : \mathcal{M}_\Pi \rightarrow \mathcal{R}_\Gamma$ and $\tilde{g}$ is an element of $\mathcal{M}_\Gamma$ of order at least $\mathcal{M}_\Pi$. Clearly, this protocol is correct and secure under reasonable assumptions. The security proof is similar to that, presented in Theorem 3.

Proxy verifiable HPET. In the $\binom{n}{1}$*-proxy private equality test* there is one Alice, n different "Bobs" $B_1, \dots, B_n$, and a new party called Peggy the Proxy. At the end of the proxy PET protocol, Peggy will get to know whether Alice is as wealthy as B_i, Bob the ith, for *all* $i \in [1, n]$, while neither Alice nor any of $B_1, \dots, B_n$ will obtain any new information information. Next, we propose a proxy verifiable homomorphic private equality test protocol (see Protocol 4) that bases on a ε-affine homomorphic semantically secure public-key cryptosystem $\Pi = (\mathcal{G}_\Pi, E, D; \mathcal{S}, \mathcal{T})$ that satisfies the same requirements as Π in Thm. 4. (We omit the security proofs.)

This protocol is basically a modification of the HPET protocol with a proxy Peggy who transmits Alice's and B_i's messages to their partners. As a drawback, Protocol 4 reveals W_A to Peggy on step 5, but importantly, this only happens

Protocol 4 The proxy verifiable HPET protocol

PRIVATE INPUTS: Alice has W_A, B_i has W_{B_i}. PRIVATE OUTPUTS: For all i, Peggy has 0 if $W_A = W_{B_i}$ or garbage, otherwise.

1. Alice generates new private key pairs $(x, K) \leftarrow \mathcal{G}_\Pi(1^k)$ and $(\tilde{x}, \tilde{K}) \leftarrow \mathcal{G}_\Gamma(1^k)$, a random $r \leftarrow \mathcal{R}_\Pi$, and sets $c \leftarrow E_K(W_A; r)$. She sends $(K, \tilde{K}, c)$ to Peggy.
2. Peggy forwards $(K, \tilde{K}, c)$ to players $B_1, \dots, B_{\mathcal{B}}$.
3. For every i, B_i creates a random $m_i \leftarrow \mathcal{T}$, computes $v_i = E_K(m_i + s_i(W_A - W_B); r^{s_i} \circ r'_i)$ for random $s_i \leftarrow \mathcal{S}$ and $r'_i \leftarrow \mathcal{R}_\Pi$, and sets $c_i \leftarrow C_{\tilde{K}}(W_{B_i}; \mathsf{tr}(m_i))$. He sends (v_i, c_i) together with his signature over $(K, \tilde{K}, c, c, v)$ to Peggy.
4. Peggy collects all values $\{v_i, c_i\}$, and signs (at an a priori fixed time) their joint commitment. He sends the signed commitment χ to Alice.
5. Alice sends W_A, x and her signature on (W_A, χ, x) to Peggy.
6. For every i, Peggy decrypts v_i by using the key x, and obtains a message $\tilde{m}_i \in \mathcal{M}_\Pi$. She decides that $W_A = W_{B_i}$ iff $c_i = C_{\tilde{K}}(W_A; \mathsf{tr}(\tilde{m}_i))$.

after Peggy has committed to B_i-s' answers: if Peggy would get to know x before forwarding $(K, \tilde{K}, c)$ on step 2, she might be able, in collaboration with some B_i, to stop the protocol before sending the commitment χ to Alice if the outcome is not suitable for Peggy. This attack is relevant in, e.g., the auction scenario (see Sect. 6), and is one of the reasons why x is sent to Peggy only at the end of the protocol. As we will also see in Sect. 6, in some applications revealing W_A at the end of the protocol is actually desirable.

Second, more secure, proxy verifiable HPET protocol. (Sketch.) In an alternative protocol to Protocol 4, instead of sending x to Peggy, Alice receives (v, c) from Peggy, obtains all messages $\tilde{m}_i$, and then proves in zero-knowledge whether v_i commits to W_A for all $i \in [1, n]$. This protocol is obviously more secure than the first protocol (since x and thus also W_A will not be revealed to Peggy), but requires at least one additional round and more communication.

6 Applications

Applications of the verifiable oblivious transfer protocol. In [AJL03], Ambainis, Jakobsson and Lipmaa proposed several protocols for the cryptographic randomised response technique. Their first protocol—that bases on their own verifiable oblivious transfer protocol—can be made more efficient (and also perfectly private for the respondent) by using the verifiable HOT protocol instead. Note that at least in their application a weakly sender-private oblivious transfer protocol with a trapdoor commitment scheme will be sufficient. See, e.g., [CvdGT95,CD97,CC00] for more applications for the verifiable HOT protocol.

Auctions. The LAN auction scheme [LAN02] is (probably) the most efficient secure cryptographic $(b + 1)$st auction scheme without threshold trust; in largescale auctions with many participants it requires 10–100 times less communication than the Naor-Pinkas-Sumner scheme [NPS99]. On the other hand, the LAN scheme has two principal drawbacks. First, the involved trusted auction authority A will get to know the bid statistics. As argued in [LAN02], this is not a weakness from the economic viewpoint when relying on the assumption that the occasional seller and the well-established business authority A do not collaborate.

Second, the LAN scheme has only an optimistic payment enforcement procedure. Namely, after the seller has received the value of the bth highest bid X_b from A, reliable winner determination is only possible when all the bidders (or at least b highest bidders) will complete a zero-knowledge proof that shows whether they bid more than X_b or not. Clearly, it may be difficult to force the bidders to collaborate at this time—especially after they know the value of X_b—, and it may be hard to distinguish between the malicious bidders (who want to disrupt the auctions, lose their interest in participation since they are not winning, or are not willing to pay as much), shills and bidders that have some genuine problems with their software or hardware. Moreover, some bidders might object to such enforcement even if they have no desire to cheat, by whatever moral or psychological reasons.

By using the proxy verifiable HPET protocol (Protocol 4), one can eliminate the second problem of the LAN scheme for $b \leq 1$ with a moderate increase in the communication complexity. The basic idea of our solution is that after the third party A has computed the bth highest bid X_b, he will not send X_b to the seller P, as it was done in the original protocol of [LAN02]. Instead, the seller will act as a proxy in $(b-1)$ parallel proxy verifiable HPET protocols with the inputs $X_1, \dots, X_{b-1}$ from A and the input b_i (B_i's bid) from the bidder B_i. After the 3rd step of the proxy verifiable PET protocol, neither the seller nor any of the bidders knows X_j for any j. Thus, none of the bidders (including the shills who cooperate with the auctioneer) has the motivation to discontinue participation in the auction. In particular, the seller has no better strategy than to be honest in step 4 of Protocol 4. Moreover, he will receive $X_1, \dots, X_{b-1}$ only on step 5 of the proxy verifiable HPET protocol, after his commitment and thus his actions are accountable. The drawback of this solution is that the seller will get to know $X_1, \dots, X_{b-1}$.

Alternatively, the participants can use the alternative proxy verifiable HPET protocol that was sketched before; in this case, no X_j will be leaked to the seller, but the communication complexity of the whole scheme increases somewhat, since the authority must provide $b-1$ zero-knowledge arguments of plaintext equality. One can most probably apply the proxy verifiable HPET protocol also to other protocols in an analogous manner.

Acknowledgements

This work was partially supported by the Finnish Defense Forces Research Institute of Technology. We would like to thank Yuval Ishai, Markus Jakobsson and Benny Pinkas for useful comments during various stages of writing this paper.

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.

[AJL03] Andris Ambainis, Markus Jakobsson, and Helger Lipmaa. Cryptographic Randomized Response Techniques. Technical Report 2003/027, International Association for Cryptologic Research, February 10 2003.

[Bou00] Fabrice Boudot. Efficient Proofs that a Committed Number Lies in an Interval. In Bart Preneel, editor, *Advances in Cryptology — EUROCRYPT 2000*, volume 1807 of *Lecture Notes in Computer Science*, pages 431–444, Bruges, Belgium, May 14–18 2000. Springer-Verlag.

[BST01] Fabrice Boudot, Berry Schoenmakers, and Jacques Traoré. A Fair and Efficient Solution to the Socialist Millionaires' Problem. *Discrete Applied Mathematics*, 111(1–2):23–36, 2001.

[CC00] Christian Cachin and Jan Camenisch. Optimistic Fair Secure Computation. In Mihir Bellare, editor, *Advances in Cryptology — CRYPTO 2000*, volume 1880 of *Lecture Notes in Computer Science*, pages 93–111, Santa Barbara, USA, 20–24 August 2000. International Association for Cryptologic Research, Springer-Verlag.

[CD97] Ronald Cramer and Ivan Damgård. Linear zero-knowledge – a note on efficient zero-knowledge proofs and arguments. In *Proceedings of the Twenty-Nineth Annual ACM Symposium on the Theory of Computing*, pages 436–445, 1997.

[CGHN01] Dario Catalano, Rosario Gennaro, Nick Howgrave-Graham, and Phong Q. Ngueyn. Paillier's Cryptosystem Revisited. In *8th ACM Conference on Computer and Communications Security*, pages 206–214, Philadelphia, Pennsylvania, USA, 6–8 November 2001. ACM Press.

[CvdGT95] Claude Crépeau, Jeroen van de Graaf, and Alain Tapp. Committed Oblivious Transfer and Private Multi-Party Computation. In Don Coppersmith, editor, *Advances in Cryptology — CRYPTO '95*, volume 963 of *Lecture Notes in Computer Science*, pages 110–123, Santa Barbara, USA, 27–31 August 1995. 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*, Lecture Notes in Computer Science, Wollongong, Australia, July 9-11 2003. Springer-Verlag. To appear.

[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.

[FNW96] Ron Fagin, Moni Naor, and Peter Wrinkler. Comparing Information Without Leaking It. *Communications of the ACM*, 39:77–85, May 1996.

[Kil88] Joe Kilian. Founding Cryptography on Oblivious Transfer. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, pages 20–31, Chicago, Illinois, USA, 2–4 May 1988. ACM Press.

[LAN02] Helger Lipmaa, N. Asokan, and Valtteri Niemi. Secure Vickrey Auctions without Threshold Trust. In Matt Blaze, editor, *Financial Cryptography — Sixth International Conference*, volume 2357 of *Lecture Notes in Computer Science*, pages 87–101, Southhampton Beach, Bermuda, March 11–14 2002. Springer-Verlag.

[Lip03] Helger Lipmaa. On Diophantine Complexity and Statistical Zero-Knowledge Arguments. In Chi Sung Laih, editor, *Advances on Cryptology — ASIACRYPT 2003*, Lecture Notes in Computer Science, Taipei, Taiwan, November 30–December 4 2003. Springer-Verlag. This volume.

[NP99] Moni Naor and Benny Pinkas. Oblivious Transfer and Polynomial Evaluation. In *Proceedings of the Thirty-First Annual ACM Symposium on the Theory of Computing*, pages 245–254, Atlanta, Georgia, USA, 1–4 May 1999. ACM Press.

[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.

[NPS99] Moni Naor, Benny Pinkas, and Reuben Sumner. Privacy Preserving Auctions and Mechanism Design. In *The 1st ACM Conference on Electronic Commerce*, Denver, Colorado, November 1999.

[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 Jacques Stern, editor, *Advances in Cryptology — EUROCRYPT '99*, volume 1592 of *Lecture Notes in Computer Science*, pages 223–238, Prague, Czech Republic, 2–6 May 1999. Springer-Verlag.

[Ped91] Torben P. Pedersen. Non-Interactive and Information-Theoretic Secure Verifiable Secret Sharing. In J. Feigenbaum, editor, *Advances in Cryptology—CRYPTO '91*, volume 576 of *Lecture Notes in Computer Science*, pages 129–140, Santa Barbara, California, USA, August 11–15 1991. Springer-Verlag, 1992.

[Tze02] Wen-Guey Tzeng. Efficient 1-Out-n Oblivious Transfer Schemes. In David Naccache and Pascal Paillier, editors, *Public Key Cryptography '2002*, volume 2274 of *Lecture Notes in Computer Science*, pages 159–171, Paris, France, February12–14 2002. Springer-Verlag.