Practical Order-Revealing Encryption with Limited Leakage

Nathan Chenette[1(✉)], Kevin Lewi[2], Stephen A. Weis[3], and David J. Wu[2]

[1] Rose-Hulman Institute of Technology, Terre Haute, USA
chenett1@rose-hulman.edu
[2] Stanford University, Stanford, USA
[3] Facebook, Inc., Menlo Park, USA

Abstract. In an order-preserving encryption scheme, the encryption algorithm produces ciphertexts that preserve the order of their plaintexts. Order-preserving encryption schemes have been studied intensely in the last decade, and yet not much is known about the security of these schemes. Very recently, Boneh et al. (Eurocrypt 2015) introduced a generalization of order-preserving encryption, called order-revealing encryption, and presented a construction which achieves this notion with best-possible security. Because their construction relies on multilinear maps, it is too impractical for most applications and therefore remains a theoretical result.

In this work, we build efficiently implementable order-revealing encryption from pseudorandom functions. We present the first efficient order-revealing encryption scheme which achieves a simulation-based security notion with respect to a leakage function that precisely quantifies what is leaked by the scheme. In fact, ciphertexts in our scheme are only about 1.6 times longer than their plaintexts. Moreover, we show how composing our construction with existing order-preserving encryption schemes results in order-revealing encryption that is strictly more secure than all preceding order-preserving encryption schemes.

1 Introduction

A symmetric encryption scheme is order-preserving if the ciphertexts preserve the numeric ordering of their underlying plaintexts. The notion of order-preserving encryption (OPE) was introduced by Agrawal et al. [1] who showed how it could be used to efficiently answer range queries over encrypted data, as well as sorting queries, searching queries, and more. Indeed, existing OPE solutions have been implemented in practice [43,46] for these exact purposes. Since the introduction of OPE, there has been a plethora of work on analyzing the security of various OPE schemes, found both in the cryptography community and the database community. However, it is troubling that in spite of the numerous practical applications of OPE, the security of the best candidate OPE schemes is still not well understood.

Prior Work. The first OPE construction by Agrawal et al. [1] relied on heuristics and lacked a formal security analysis. Subsequently, Boldyreva et al. [7] gave

T. Peyrin (Ed.): FSE 2016, LNCS 9783, pp. 474–493, 2016.
DOI: 10.1007/978-3-662-52993-5_24

the first formal security definitions for OPE schemes. Boldyreva et al. introduced two primary notions for security of an OPE scheme. The first notion of security for an OPE scheme is called indistinguishability under an ordered chosen plaintext attack (IND-OCPA). The IND-OCPA definition can be viewed as a generalization of semantic security [31], and effectively says that encryptions of a sequence of messages should reveal nothing about the underlying messages other than their ordering. However, in the same work, Boldyreva et al. showed that no efficient order-preserving encryption scheme can be IND-OCPA secure, even in settings where the size of the ciphertext space is exponentially larger than the size of the plaintext space.

In light of this lower bound for OPE schemes that satisfy IND-OCPA security, Boldyreva et al. introduced a weaker notion of security (POPF-CCA security) where the encryption function for the OPE scheme is compared to a random order-preserving function—that is, the encryption algorithm for an OPE scheme behaves like a truly random order-preserving function. Under this definition, an OPE scheme inherits the properties of a random order-preserving function.[1] In the same work, Boldyreva et al. gave an explicit construction of an OPE scheme that satisfies POPF-CCA security. However, the POPF-CCA security definition does not precisely specify the information that is leaked by an OPE scheme that achieves this definition. In fact, a scheme that achieves this notion of security does not even satisfy semantic security for a single encryption, and indeed, in subsequent work, Boldyreva et al. [8] showed that ciphertexts in their OPE scheme leak approximately the first half of the bits of the underlying plaintexts. In addition, they introduce several new security definitions in order to better quantify the information leakage of OPE schemes that are POPF-CCA secure.

Recently, Boneh et al. [9] proposed a generalization of OPE called order-revealing encryption (ORE). In an OPE scheme, the ciphertexts are numeric-valued, and the ordering of the underlying plaintexts is determined by numerically comparing the ciphertexts. In contrast, in an ORE scheme, the ciphertexts are not constrained to any particular form, and instead, there is a publicly computable comparison function which takes two ciphertexts and outputs the numeric ordering of the underlying plaintexts[2]. Although this generalization may at first seem subtle, Boneh et al. constructed an ORE scheme from multilinear maps that achieves the "best-possible" notion of security, which is equivalent to the IND-OCPA security notion for order-preserving encryption.

The main drawback of the Boneh et al. ORE construction is that it relies on complicated tools and strong assumptions on these tools, and as such, is currently impractical to implement.

[1] This definition is inspired by the similar definition for PRF security [28], which compares the output of a keyed function to that of a truly random function.

[2] This application was also observed and independently achieved by Goldwasser et al. [29] using indistinguishability obfuscation.

1.1 Our Contributions

We now summarize the main contributions of this work, which include a new simulation-based security notion for ORE, along with a practical construction of an ORE scheme which achieves this security notion. We also show how our new construction can be used to achieve a strictly stronger notion of security compared to other stateless and efficiently implementable (e.g., constructions that do not rely on powerful primitives such as multilinear maps and indistinguishability obfuscation) OPE and ORE encryption schemes.

Security Model. In our work, we take the general approach of Boneh et al. in constructing an ORE scheme, except we take a more efficient route. Our first contribution is a new security definition for order-revealing encryption schemes that both allows for and explicitly models the leakage in the scheme. Our design goals for introducing this new security model are twofold: first, the security model should enable constructions that are efficiently implementable, and second, it should provide a precise quantification of any information leaked by the scheme. The two primary notions of security, IND-OCPA and POPF-CCA, introduced by Boldyreva et al. [7] each satisfy one of these two properties. In particular, all non-interactive, stateless[3] ORE schemes that achieve IND-OCPA security require strong cryptographic primitives such as multilinear maps or indistinguishability obfuscation [9,29], and thus, are not efficiently implementable today. At the other end of the spectrum, it is difficult to precisely quantify the leakage of schemes that satisfy POPF-CCA security. The work by Boldyreva et al. [8] provides some concrete lower and upper bounds for the leakage under the strong assumption that the plaintexts are drawn from a uniform distribution. For more general distributions, the leakage remains unclear.

In our work, we give a simulation-based definition of security for ORE with respect to a leakage function $\mathcal{L}$. In other words, our definition states that whatever an adversary is able to deduce from seeing encryptions of messages $m_1, \ldots, m_t$, it could also deduce given only the leakage $\mathcal{L}(m_1, \ldots, m_t)$. The "best-possible" security for ORE would correspond to the case where the leakage function simply outputs whether $m_i < m_j$ for all pairs of messages m_i and m_j. By allowing for the possibility of additional leakage, it becomes possible to construct practical ORE schemes from standard assumptions. Thus, our constructions provide a concrete trade-off between security and efficiency. Our security definitions are similar to the simulation-based definitions that have been considered previously in the searchable symmetric encryption literature [14,22].

Constructions. In our main construction, we show how to construct an ORE scheme from one-way functions (more precisely, from pseudorandom functions (PRFs) [28]). This particular ORE scheme reveals slightly more information than just the ordering of the underlying messages. Specifically, two ciphertexts encrypting messages m_1 and m_2 also reveal the index of the first bit in m_1 and

[3] There are "mutable" order-preserving encryption schemes [35,36,42] that do satisfy IND-OCPA, but they require stateful encryption, and oftentimes, an interactive protocol to "update" ciphertexts. We survey some of these constructions in Sect. 1.2.

m_2 that differ. In other words, our ORE scheme leaks some information about the relative distance between the underlying messages.

We give a brief overview of our PRF-based construction. The secret key in our scheme consists of a PRF key k. The output space of the PRF is the set $\{0, 1, 2\}$. Each ciphertext consists of the bits of the message blinded by the outputs of the PRF evaluated on the prefixes of the message. More precisely, to encrypt an n-bit message $m = m_1 m_2 \cdots m_n$, the encryption algorithm effectively computes the following for each $i \in [n]$:

$$u_i = F(k, m_1 m_2 \cdots m_{i-1}) + m_i \pmod 3.$$

Note that to support variable-length PRF inputs, we simply pad the input. We describe our construction in greater detail in Sect. 3. The ciphertext is then the tuple $\mathsf{ct} = (u_1, \ldots, u_n)$ of blinded values.

To compare encryptions $\mathsf{ct} = (u_1, \ldots, u_n)$ and $\mathsf{ct}' = (u'_1, \ldots, u'_n)$ of messages m and m', the evaluator first finds the first index i for which $u_i \neq u'_i$. Since u_i and u'_i are functions of just the first i bits of m and m', respectively, the first index i for which $u_i \neq u'_i$ is the first bit of m and m' that differ. After identifying the i^{th} bit that differs, the evaluator uses u_i and u'_i to determine which message has 0 as the i^{th} bit and which message has 1[4]. Conversely, if $u_i = u'_i$ for all i, then $\mathsf{ct}_i = \mathsf{ct}'_i$, and so $m = m'$. Security of this construction follows from the security of the PRF (Theorem 3.2).

Ciphertexts in our candidate scheme are $\lceil n \cdot \log_2 3 \rceil \approx \lceil 1.6\, n \rceil$ bits, where n is the bit-length of the message. As a point of comparison, ciphertexts in the OPE scheme of Boldyreva et al. [7] are only $n + 1$ bits long. While the ciphertexts in our scheme are longer (by a multiplicative factor $\log_2 3$), the authors of [8] note that even if the size of the ciphertext space is increased beyond $n + 1$ bits in the Boldyreva et al. scheme, the security of their construction does not improve by any noticeable amount.

We then explain in Sect. 3.2 how to convert our ORE scheme into an OPE scheme, at the expense of longer ciphertexts. This is useful for applications where it is more convenient to have a numeric ciphertext space and for order relations to be computable without a "custom" comparison function. The transformation we describe is natural and does not reduce the security of the original ORE scheme. In particular, we note that the resulting OPE scheme does *not* behave like a random order-preserving function (the ideal object from the POPF-CCA security notion). Thus, the scheme is able to achieve stronger security than the Boldyreva et al. OPE scheme.

Comparison with Existing Schemes. First, we note in Sect. 2.3 that the security of any OPE scheme can be "augmented" by applying ORE encryption on top of OPE encryption. The resulting scheme is at least as secure as the underlying OPE scheme, and moreover, inherits the security properties of the ORE scheme. Hence, by composing our ORE construction with existing OPE constructions, we obtain ORE schemes that are at least as secure.

[4] Either $u_i + 1 = u'_i \pmod 3$, in which case $m < m'$, or $u_i - 1 = u'_i \pmod 3$, in which case $m > m'$.

While composing an OPE scheme with an ORE scheme yields a scheme that is at least as secure as the underlying OPE scheme, we show that even without this composition, our basic ORE scheme still achieves stronger security guarantees according to the one-wayness metrics introduced by Boldyreva et al. [8] for analyzing the leakage of random order-preserving functions (and by extension, any OPE scheme that is POPF-CCA secure). In our work, we introduce two generalized one-wayness notions and show that under a uniform plaintext distribution,[5] our basic ORE scheme achieves strictly stronger security compared to OPE schemes that are POPF-CCA secure. Specifically, Boldyreva et al. [8] show that a random order-preserving function leaks half of the most-significant bits of the messages with probability close to 1. In contrast, under the same settings, we can show that our basic ORE scheme will not leak *any constant* fraction of the message bits with *overwhelming* probability.

1.2 Related Work

In recent years, there have been numerous works on order-preserving encryption and related notions [1,7,8,35,36,38,41,42,44,47]. In this section, we survey some of these works.

Security Definitions. Though the POPF-CCA security definition introduced by Boldyreva et al. [7] is similar in flavor to PRF security, it is not immediately evident what kind of information the output of a random order-preserving function leaks about its input. In a follow-up work [8], Boldyreva et al. introduce several notions (based on definitions of one-wayness [27] for one-way functions) to capture the information leakage in schemes that are POPF-CCA secure. They show that a random order-preserving function leaks at least half of the bits in each message.

Teranishi et al. [47] also introduce a stronger indistinguishability-based notion (stronger than the one-wayness definitions from [8], but weaker than IND-OCPA) for OPE schemes, as well as a construction that achieves these stronger notions. Notably, their definition ensures that under a uniform message distribution, any fraction of the low-order bits of the messages being encrypted are hidden.

Recently, Naveed et al. [40] analyzed the information leaked by order-preserving encryption used in practical scenarios.

Modular OPE. Boldyreva et al. also introduced the notion of modular OPE as a possible extension of standard OPE [8]. In modular OPE, a modular shift is applied to each plaintext before applying OPE—so the scheme is not order-preserving, but naturally supports "wrap-around" range queries. Their modular OPE scheme adds an extra layer of security to vanilla OPE, but it is worth noting that leakage of a small amount of information (say, a single plaintext-ciphertext pair) reveals the shift value and nullifies this added security. Subse-

[5] This is the only distribution for which we have concrete analysis of the leakage in any POPF-CCA secure scheme.

quently, Mavroforakis et al. [38] designed several protocols to avoid leaking the shift value while using modular OPE schemes in practice.

Mutable OPE. Popa et al. [42] introduced a related notion of a mutable order-preserving encoding scheme which can be viewed as a two-party protocol that allows a user to insert and store encrypted values in a database such that the database is able to perform comparisons and range queries on the encrypted values without learning anything more about the values. Their construction is interactive and leverages stateful encryption. By working in this setting, the authors are able to circumvent the Boldyreva et al. [7] lower bound for order-preserving encryption and show that their scheme is IND-OCPA secure.

In subsequent work, Kerschbaum and Schröpfer [36] improved on the communication complexity of the Popa et al. construction at the expense of increasing the amount of client-side state. Specifically, in their construction, the amount of persistent state the client has to maintain increases linearly in the number of elements inserted into the database. More recently, Kerschbaum [35] introduced a new notion of frequency-hiding OPE that introduces additional randomness to hide whether multiple ciphertexts encrypt the same value. Their notions provide a strictly stronger guarantee than IND-OCPA.

Very recently, Roche et al. [44] introduced the notion of partial order-preserving encodings, which optimizes for the setting where there are a huge number of insertion queries but only a moderate number of range queries. Their protocol improves upon the round-complexity for insertions compared to the Popa et al. protocol [42], and requires the client to maintain less state than the Kerschbaum-Schröpfer construction [36]. All of the schemes described here require stateful encryption and employ an interactive encryption procedure.

ORE. Order-revealing encryption schemes, as introduced by Boneh et al. [9] provide another method of circumventing the Boldyreva et al. lower bound [7]. In an ORE scheme, the public comparison operation is not required to correspond to numerically comparing the ciphertexts, and in fact, the ciphertexts themselves need not be elements of a numeric, well-ordered set. This type of relaxation was previously considered by Pandey and Rouselakis [41] in the context of property-preserving encryption. In a property-preserving encryption scheme, there is a publicly computable function that can be evaluated on ciphertexts to determine the value of some property on the underlying plaintexts. Order-revealing encryption can thus be viewed as a property-preserving encryption scheme for the comparison operation. Pandey and Rouselakis introduce and explore several indistinguishability-based notions of security for property-preserving encryption; however, they do not construct an order-revealing encryption scheme.

To the best of our knowledge, all existing ORE schemes that provide IND-OCPA security either rely on very strong (and currently impractical) cryptographic primitives such as indistinguishability obfuscation [29] and cryptographic multilinear maps [9], or only achieve a weaker notion of security [3,12] when instantiated with simple cryptographic primitives such as public key cryptography. For the constructions based on indistinguishability obfuscation or multilinear maps [9,29], security of the ORE scheme is conditional on the conjectured

security of cryptographic multilinear maps [2,10,20,21,23,26,37][6]. However, in the last few months, numerous attacks [11,16–19,33,39] on these multilinear maps have emerged, raising some doubts about the security of constructions that leverage them.

To avoid multilinear maps in favor of more well-studied number-theoretic or lattice-based assumptions, one can apply arity-amplification techniques [3,12] to a single-input functional encryption scheme based on simpler assumptions such as learning with errors [30] or semantically-secure public-key encryption [32,45]. However, due to limitations of the underlying functional encryption schemes, the resulting ORE scheme only provides "bounded-message" security—that is, security only holds if there is an *a priori* (polynomial) bound on the maximum number of messages that will be encrypted. Moreover, the length of the ciphertexts in this scheme grows *polynomially* in the bound on the number of messages that will be encrypted. These constraints severely limit the practicality of the resulting ORE scheme. To obtain full semantic security, it would be necessary to apply the arity-amplification transformation to a more powerful functional encryption scheme, but to date, the only known candidates of such schemes rely again on indistinguishability obfuscation [24] or multilinear maps [25].

Recently, Bun and Zhandry [13] investigated the connection between order-revealing encryption and problems in learning theory.

Other schemes. Numerous ad hoc or heuristic order-preserving encryption schemes [6,34,48] have been proposed in the literature, but most lack formal security analysis.

2 Order-Revealing Encryption

In this section, we establish and review some conventions that we use in this work, and also formally define our security notions for our encryption schemes.

Preliminaries. For $n \in \mathbb{N}$, we write $[n]$ to denote the set of integers $\{1, \ldots, n\}$, and $\mathbb{Z}_n$ to denote the additive group of integers modulo n. If $\mathcal{P}(x)$ is a predicate on x, we write $\mathbf{1}(\mathcal{P}(x))$ to denote the indicator function for $\mathcal{P}$: that is, $\mathbf{1}(\mathcal{P}(x)) = 1$ if and only if $\mathcal{P}(x) = 1$, and 0 otherwise. If $x, y \in \{0,1\}^*$ are bit-strings, we write $x \| y$ to denote the concatenation of x and y. For a finite set S, we write $\mathsf{Unif}(S)$ to denote the uniform distribution on S. We say a function $f(\lambda)$ is negligible in a security parameter λ if $f = o(1/\lambda^c)$ for all $c \in \mathbb{N}$. We write $\mathrm{negl}(\lambda)$ to denote a negligible function in λ and $\mathrm{poly}(\lambda)$ to denote a polynomial in λ. We say an event occurs with negligible probability if the probability of the event is $\mathrm{negl}(\lambda)$, and it occurs with overwhelming probability if the complement of the event occurs with negligible probability. Finally, we review the definition of a pseudorandom function (PRF) [28]. Let $\mathsf{Funs}[\mathcal{D}, \mathcal{R}]$ denote the set of all functions from a domain $\mathcal{D}$ to a range $\mathcal{R}$. In this paper, we specialize the domain of our PRFs to $\{0,1\}^n$.

[6] To date, the only concrete instantiations of indistinguishability obfuscation [4,5,24,49] leverage multilinear maps.

Definition 2.1 (Pseudorandom Function [28]). Fix a security parameter λ. A PRF $F : \mathcal{K} \times \{0,1\}^n \to \mathcal{R}$ with key space $\mathcal{K}$, domain $\{0,1\}^n$, and range $\mathcal{R}$ is secure if for all efficient adversaries $\mathcal{A}$,

$$\left| \Pr\left[k \xleftarrow{R} \mathcal{K} : \mathcal{A}^{F(k,\cdot)}(1^\lambda) = 1\right] - \Pr\left[f \xleftarrow{R} \mathsf{Funs}[\{0,1\}^n, \mathcal{R}] : \mathcal{A}^{f(\cdot)}(1^\lambda) = 1\right] \right| = \mathrm{negl}(\lambda).$$

2.1 Order-Revealing Encryption

An order-revealing encryption (ORE) scheme is a tuple of algorithms $\Pi = (\mathsf{ORE.Setup}, \mathsf{ORE.Encrypt}, \mathsf{ORE.Compare})$ defined over a well-ordered domain $\mathcal{D}$ with the following properties:

- $\mathsf{ORE.Setup}(1^\lambda) \to \mathsf{sk}$. On input a security parameter λ, the setup algorithm ORE.Setup outputs a secret key sk.
- $\mathsf{ORE.Encrypt}(\mathsf{sk}, m) \to \mathsf{ct}$. On input the secret key sk and a message $m \in \mathcal{D}$, the encrypt algorithm ORE.Encrypt outputs a ciphertext ct.
- $\mathsf{ORE.Compare}(\mathsf{ct}_1, \mathsf{ct}_2) \to b$. On input two ciphertexts ct_1, ct_2, the compare algorithm ORE.Compare outputs a bit $b \in \{0,1\}$.

Remark 2.2 (Public Parameters). In general, the setup algorithm of an ORE scheme can also output public parameters pp which are then passed as an additional input to the comparison algorithm, as is done in Boneh et al. [9]. However, none of our constructions require these public parameters, so we omit them in this work for simplicity.

Remark 2.3 (Support for Decryption). As described, our definition of an order-revealing encryption scheme does not include a "decryption" function. However, this omission is without loss of generality. To decrypt a message, the holder of the secret key can use the secret key to encrypt messages of her choosing, apply the comparison algorithm, and perform binary search to recover the message. An alternative method that avoids the need for binary search is to augment each ORE encryption of a message m with an encryption of m under a CPA-secure symmetric encryption scheme. The secret key of the ORE scheme would also include the key for the symmetric encryption scheme. As long as the underlying encryption scheme is CPA-secure, including this additional ciphertext does not compromise security. For the remainder of this work, we use the schema described above that does not explicitly specify a decryption function.

Correctness. Fix a security parameter λ. An ORE scheme $\Pi = (\mathsf{ORE.Setup}, \mathsf{ORE.Encrypt}, \mathsf{ORE.Compare})$ over a well-ordered domain $\mathcal{D}$ is correct if for $\mathsf{sk} \leftarrow \mathsf{ORE.Setup}(1^\lambda)$, and all messages $m_1, m_2 \in \mathcal{D}$,

$$\Pr[\mathsf{ORE.Compare}(\mathsf{ct}_1, \mathsf{ct}_2) = \mathbf{1}(m_1 < m_2)] = 1 - \mathrm{negl}(\lambda),$$

where $\mathsf{ct}_1 \leftarrow \mathsf{ORE.Encrypt}(\mathsf{sk}, m_1)$ and $\mathsf{ct}_2 \leftarrow \mathsf{ORE.Encrypt}(\mathsf{sk}, m_2)$, and the probability is taken over the random coins in ORE.Setup and ORE.Encrypt.

Security. We now give our simulation-based notion of security for an ORE scheme. As described in Sect. 1.1, our security definition is parameterized by a leakage function $\mathcal{L}$, which exactly specifies what is leaked by an ORE scheme.

Definition 2.4 (Security of ORE with Leakage). *Fix a security parameter $\lambda \in \mathbb{N}$. Let $\Pi_{\mathsf{ore}} = (\mathsf{ORE.Setup}, \mathsf{ORE.Encrypt}, \mathsf{ORE.Compare})$ be an ORE scheme. Let $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_q)$ be an adversary for some $q \in \mathbb{N}$. Let $\mathcal{S} = (\mathcal{S}_0, \mathcal{S}_1, \ldots, \mathcal{S}_q)$ be a simulator, and let $\mathcal{L}(\cdot)$ be a leakage function. We define the experiments $\mathsf{REAL}^{\mathrm{ORE}}_{\mathcal{A}}(\lambda)$ and $\mathsf{SIM}^{\mathrm{ORE}}_{\mathcal{A},\mathcal{S},\mathcal{L}}(\lambda)$ as follows:*

$\mathsf{REAL}^{\mathrm{ORE}}_{\mathcal{A}}(\lambda)$:

1. $\mathsf{sk} \leftarrow \mathsf{ORE.Setup}(1^\lambda)$
2. $(m_1, \mathsf{st}_{\mathcal{A}}) \leftarrow \mathcal{A}_1(1^\lambda)$
3. $c_1 \leftarrow \mathsf{ORE.Encrypt}(\mathsf{sk}, m_1)$
4. for $2 \le i \le q$:
   (a) $(m_i, \mathsf{st}_{\mathcal{A}}) \leftarrow \mathcal{A}_i(\mathsf{st}_{\mathcal{A}}, c_1, \ldots, c_{i-1})$
   (b) $c_i \leftarrow \mathsf{ORE.Encrypt}(\mathsf{sk}, m_i)$
5. output $(c_1, \ldots, c_q)$ and $\mathsf{st}_{\mathcal{A}}$

$\mathsf{SIM}^{\mathrm{ORE}}_{\mathcal{A},\mathcal{S},\mathcal{L}}(\lambda)$:

1. $\mathsf{st}_{\mathcal{S}} \leftarrow \mathcal{S}_0(1^\lambda)$
2. $(m_1, \mathsf{st}_{\mathcal{A}}) \leftarrow \mathcal{A}_1(1^\lambda)$
3. $(c_1, \mathsf{st}_{\mathcal{S}}) \leftarrow \mathcal{S}_1(\mathsf{st}_{\mathcal{S}}, \mathcal{L}(m_1))$
4. for $2 \le i \le q$:
   (a) $(m_i, \mathsf{st}_{\mathcal{A}}) \leftarrow \mathcal{A}_i(\mathsf{st}_{\mathcal{A}}, c_1, \ldots, c_{i-1})$
   (b) $(c_i, \mathsf{st}_{\mathcal{S}}) \leftarrow \mathcal{S}_i(\mathsf{st}_{\mathcal{S}}, \mathcal{L}(m_1, \ldots, m_i))$
5. output $(c_1, \ldots, c_q)$ and $\mathsf{st}_{\mathcal{A}}$

We say that Π_{ore} is a secure ORE scheme with leakage function $\mathcal{L}(\cdot)$ if for all polynomial-size adversaries $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_q)$ where $q = \mathrm{poly}(\lambda)$, there exists a polynomial-size simulator $\mathcal{S} = (\mathcal{S}_0, \mathcal{S}_1, \ldots, \mathcal{S}_q)$ such that the outputs of the two distributions $\mathsf{REAL}^{\mathrm{ORE}}_{\mathcal{A}}(\lambda)$ and $\mathsf{SIM}^{\mathrm{ORE}}_{\mathcal{A},\mathcal{S},\mathcal{L}}(\lambda)$ are computationally indistinguishable.

Remark 2.5 (IND-OCPA Security). We briefly note how the IND-OCPA definition of security is captured by this definition. Let $\mathcal{L}$ be the following leakage function:

$$\mathcal{L}(m_1, \ldots, m_t) = \{\mathbf{1}(m_i < m_j) : 1 \le i < j \le t\}.$$

If an ORE scheme is secure with leakage $\mathcal{L}$, then it is IND-OCPA secure.

2.2 Order-Preserving Encryption (OPE)

An OPE scheme [1,7] is a special case of an ORE scheme, where the ciphertext space is required to be a well-ordered range $\mathcal{R}$ and moreover, for two ciphertexts $\mathsf{ct}_1, \mathsf{ct}_2 \in \mathcal{R}$, the comparison algorithm outputs 1 if $\mathsf{ct}_1 < \mathsf{ct}_2$. For simplicity, we can write an OPE scheme as a tuple of algorithms $\Pi = (\mathsf{OPE.Setup}, \mathsf{OPE.Encrypt})$ defined over a well-ordered domain $\mathcal{D}$ and well-ordered range $\mathcal{R}$ with the following properties:

- $\mathsf{ORE.Setup}(1^\lambda) \to \mathsf{sk}$. On input a security parameter λ, the setup algorithm ORE.Setup outputs a secret key sk.
- $\mathsf{ORE.Encrypt}(\mathsf{sk}, m) \to \mathsf{ct}$. On input the secret key sk and a message $m \in \mathcal{D}$, the encrypt algorithm OPE.Encrypt outputs a ciphertext $\mathsf{ct} \in \mathcal{R}$.

Correctness. An OPE scheme $\Pi = (\mathsf{OPE.Setup}, \mathsf{OPE.Encrypt})$ over a well-ordered domain $\mathcal{D}$ and well-ordered range $\mathcal{R}$ is correct if $\mathsf{sk} \leftarrow \mathsf{OPE.Setup}(1^\lambda)$, and all messages $m_1, m_2 \in \mathcal{D}$,

$$m_1 < m_2 \quad \iff \quad \mathsf{OPE.Encrypt}(\mathsf{sk}, m_1) < \mathsf{OPE.Encrypt}(\mathsf{sk}, m_2)$$

with overwhelming probability.

2.3 Composing OPE with ORE

By composing an ORE scheme with an OPE scheme, we obtain an ORE scheme whose security is at least as strong as the security of the underlying OPE scheme. Let $\Pi_{\mathsf{ope}} = (\mathsf{OPE.Setup}, \mathsf{OPE.Encrypt})$ be an OPE scheme and $\Pi^{\mathsf{in}}_{\mathsf{ore}} = (\mathsf{ORE^{in}.Setup}, \mathsf{ORE^{in}.Encrypt}, \mathsf{ORE^{in}.Compare})$ be an ORE scheme. Consider the following composed construction $\Pi_{\mathsf{ore}} = (\mathsf{ORE.Setup}, \mathsf{ORE.Encrypt}, \mathsf{ORE.Compare})$ of an ORE scheme with an OPE scheme:

- $\mathsf{ORE.Setup}(1^\lambda)$. The setup algorithm runs $\mathsf{sk}_1 \leftarrow \mathsf{OPE.Setup}(1^\lambda)$ and $\mathsf{sk}_2 \leftarrow \mathsf{ORE^{in}.Setup}(1^\lambda)$. The secret key is $\mathsf{sk} = (\mathsf{sk}_1, \mathsf{sk}_2)$.
- $\mathsf{ORE.Encrypt}(\mathsf{sk}, \mathsf{m})$. The encryption algorithm outputs $\mathsf{ORE^{in}.Encrypt}(\mathsf{sk}_2, \mathsf{OPE.Encrypt}(\mathsf{sk}_1, \mathsf{m}))$.
- $\mathsf{ORE.Compare}(\mathsf{ct}_1, \mathsf{ct}_2)$. The compare algorithm computes and outputs the value $\mathsf{ORE^{in}.Compare}(\mathsf{ct}_1, \mathsf{ct}_2)$.

Correctness of Π_{ore} follows immediately from the correctness of $\Pi^{\mathsf{in}}_{\mathsf{ore}}$ and Π_{ope}. Furthermore, we note that under our simulation-based definition of security, the composed scheme Π_{ore} is at least as secure as Π_{ope}. This intuition is formalized in the following remark, whose proof follows immediately by construction.

Remark 2.6 (Security of Composed Scheme). For any leakage function $\mathcal{L}(\cdot)$, if the OPE scheme Π_{ope} is secure with leakage function $\mathcal{L}(\cdot)$, then the ORE scheme Π_{ore} is also secure with leakage function $\mathcal{L}(\cdot)$.

3 Main Construction

In this section, we give a construction of an ORE scheme for the set of n-bit positive integers with the following leakage function:

$$\mathcal{L}_{\mathsf{f}}(m_1, \ldots, m_t) := \{(\mathsf{ind}_{\mathsf{diff}}(m_i, m_j), \mathbf{1}(m_i < m_j)) : 1 \le i < j \le t\}, \qquad (3.1)$$

where $\mathsf{ind}_{\mathsf{diff}}(x, y)$ gives the index of the first bit where x and y differ. If $x = y$, we set $\mathsf{ind}_{\mathsf{diff}}(x, y) = n + 1$. In other words, for $x \neq y$, if $x = x_1 \cdots x_n$ and $y = y_1 \cdots y_n$, then $\mathsf{ind}_{\mathsf{diff}}(x, y)$ is the smallest index $\ell \in [n]$ for which $x_\ell \neq y_\ell$.

Construction. Fix a security parameter $\lambda \in \mathbb{N}$, and take an integer $M \ge 3$. Let $F : \mathcal{K} \times ([n] \times \{0,1\}^{n-1}) \to \mathbb{Z}_M$ be a secure PRF. We define our ORE scheme $\Pi_{\mathsf{ore}} = (\mathsf{ORE.Setup}, \mathsf{ORE.Encrypt}, \mathsf{ORE.Compare})$ as follows:

- $\mathsf{ORE.Setup}(1^\lambda)$. The setup algorithm chooses a uniformly random PRF key k for F. The secret key is $\mathsf{sk} = k$.
- $\mathsf{ORE.Encrypt}(\mathsf{sk}, m)$. Let $b_1 \cdots b_n$ be the binary representation of m and let $\mathsf{sk} = k$. For each $i \in [n]$, the encryption algorithm computes

$$u_i = F(k, (i, b_1 b_2 \cdots b_{i-1} \| 0^{n-i})) + b_i \pmod{M},$$

  and outputs the tuple $(u_1, u_2 \ldots, u_n)$.
- $\mathsf{ORE.Compare}(\mathsf{ct}_1, \mathsf{ct}_2)$. The compare algorithm first parses

$$\mathsf{ct}_1 = (u_1, u_2, \ldots, u_n)$$
$$\mathsf{ct}_2 = (u'_1, u'_2, \ldots, u_n),$$

  where $u_1, \ldots, u_n, u'_1, \ldots, u'_n \in \mathbb{Z}_M$. Let i be the smallest index where $u_i \neq u'_i$. If no such index exists, output 0. If such an index exists, output 1 if $u'_i = u_i + 1 \pmod{M}$, and 0 otherwise.

3.1 Correctness and Security

We now show that the above ORE scheme Π_{ore} is correct and secure against the leakage function $\mathcal{L}_{\mathsf{f}}$ from Eq. (3.1). We give the proof of the following theorem in the full version of this paper [15].

Theorem 3.1 *The ORE scheme Π_{ore} is correct.*

Next, we state and prove the security theorem for Π_{ore}.

Theorem 3.2 *The order-revealing encryption scheme Π_{ore} is secure with respect to leakage function $\mathcal{L}_{\mathsf{f}}$ (Definition 2.4) under the PRF security of F.*

Proof. Fix a security parameter λ and let $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_q)$ where $q = \mathrm{poly}(\lambda)$ be an efficient adversary for the ORE security game (Definition 2.4). To prove security, we give an efficient simulator $\mathcal{S} = (\mathcal{S}_0, \ldots, \mathcal{S}_q)$ for which the outputs of the distributions $\mathsf{REAL}^{\mathrm{ORE}}_{\mathcal{A}}(\lambda)$ and $\mathsf{SIM}^{\mathrm{ORE}}_{\mathcal{A},\mathcal{S},\mathcal{L}_{\mathsf{f}}}(\lambda)$ are computationally indistinguishable.

We use a hybrid argument. We begin by defining the hybrid experiments:

- **Hybrid** H_0: This is the real experiment $\mathsf{REAL}^{\mathrm{ORE}}_{\mathcal{A}}(\lambda)$.
- **Hybrid** H_1: Same as H_0, except during $\mathsf{ORE.Setup}$, a random function $f \xleftarrow{\mathrm{R}} \mathsf{Funs}[([n] \times \{0,1\}^{n-1}), \mathbb{Z}_M]$ is chosen. In all invocations of $\mathsf{ORE.Encrypt}$, the function $F(k, \cdot)$ is replaced by $f(\cdot)$.

Hybrids H_0 and H_1 are computationally indistinguishable under the PRF security of F. Thus, it suffices to show that there exists a simulator $\mathcal{S}$ such that the distribution of outputs in H_1 is computationally indistinguishable from $\mathsf{SIM}^{\mathrm{ORE}}_{\mathcal{A},\mathcal{S},\mathcal{L}_{\mathsf{f}}}(\lambda)$.

Description of the Simulator. We now describe the simulator $\mathcal{S} = (\mathcal{S}_0, \ldots, \mathcal{S}_q)$. First, $\mathcal{S}_0$ initializes an empty lookup tables $\mathsf{L} : [q] \times [n] \to \mathbb{Z}_M$.

It then outputs $\mathsf{st}_{\mathcal{S}} = \mathsf{L}$. Then, for each $t \in [q]$, after the adversary outputs a query m_t, the simulation algorithm $\mathcal{S}_t$ is invoked on input $\mathsf{st}_{\mathcal{S}} = \mathsf{L}$ and $\mathcal{L}_{\mathsf{f}}(m_1, \ldots, m_t)$. In particular, $\mathcal{L}_{\mathsf{f}}(m_1, \ldots, m_t)$ contains the values $\mathbf{1}(m_j < m_t)$ and $\mathsf{ind}_{\mathsf{diff}}(m_j, m_t)$ for all $j \in [t-1]$, where $\mathsf{ind}_{\mathsf{diff}}(m_j, m_t)$ is the index of the first bit in m_j and m_t that differ. For each $s \in [n]$, there are three cases to consider:

- **Case 1:** There exists a $j \in [t-1]$ such that $\mathsf{ind}_{\mathsf{diff}}(m_j, m_t) > s$. If there are multiple j for which $\mathsf{ind}_{\mathsf{diff}}(m_j, m_t) > s$, let j be the smallest one. Then, the simulator sets $\overline{u}_s = \mathsf{L}(j, s)$.
- **Case 2:** For each $\ell \in [t-1]$, $\mathsf{ind}_{\mathsf{diff}}(m_\ell, m_t) \leq s$, and there exists a $j \in [t-1]$ for which $\mathsf{ind}_{\mathsf{diff}}(m_j, m_t) = s$. If there are multiple j for which $\mathsf{ind}_{\mathsf{diff}}(m_j, m_t) = s$, let j be the smallest one. Then, the simulator sets $\overline{u}_s = \mathsf{L}(j, s) - (1 - 2 \cdot \mathbf{1}(m_j < m_t)) \pmod M$.
- **Case 3:** For each $\ell \in [t-1]$, $\mathsf{ind}_{\mathsf{diff}}(m_\ell, m_t) < s$. In this case, the simulator samples $y \xleftarrow{\mathrm{R}} \mathbb{Z}_M$ and sets $\overline{u}_s = y$.

For each $s \in [n]$, the simulator adds the mapping $(t, s) \mapsto \overline{u}_s$ to L. Finally, the simulator $\mathcal{S}_t$ outputs the ciphertext $\overline{\mathsf{ct}}_t = (\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_n)$ and the updated state $\mathsf{st}_{\mathcal{S}} = \mathsf{L}$. This completes the description of the simulator $\mathcal{S}$.

Correctness of the Simulation. We show that the simulator $\mathcal{S} = (\mathcal{S}_0, \ldots, \mathcal{S}_q)$ perfectly simulates the distribution in hybrid H_2. Let $(\mathsf{ct}_1, \ldots, \mathsf{ct}_q)$ be the joint distribution of the ciphertexts output in hybrid H_2, and let $(\overline{\mathsf{ct}}_1, \ldots, \overline{\mathsf{ct}}_q)$ be the joint distribution of the ciphertexts output by the simulator. We proceed inductively in the number of queries q. The base case ($q = 0$) follows trivially.

Suppose now that $(\mathsf{ct}_1, \ldots, \mathsf{ct}_{t-1}) \equiv (\overline{\mathsf{ct}}_1, \ldots, \overline{\mathsf{ct}}_{t-1})$ for some $t \in [q]$. We show that the statement holds for $t+1$. Consider the distributions of ct_t and $\overline{\mathsf{ct}}_t$. First, for any $j \in [t]$, write ciphertext ct_j as $(u_{j,1}, u_{j,2}, \ldots, u_{j,n})$ and $\overline{\mathsf{ct}}_j$ as $(\overline{u}_{j,1}, \overline{u}_{j,2}, \ldots, \overline{u}_{j,n})$. In addition, for $j \in [t]$, we write $b_{j,s}$ to denote the s^{th} bit of m_j. For each $s \in [n]$, we consider three cases:

- **Case 1:** There exists a $j \in [t-1]$ such that $\mathsf{ind}_{\mathsf{diff}}(m_j, m_t) > s$. If there are multiple j for which $\mathsf{ind}_{\mathsf{diff}}(m_j, m_t) > s$, let j be the smallest one. This means that m_j and m_t share a prefix of length at least s. Let $p \in \{0,1\}^{s-1}$ be the first $s-1$ bits of this common prefix. Then, in hybrid H_1, we have

$$u_{t,s} = f(s, p \| 0^{n-s}) + b_{t,s} = u_{j,s}.$$

  In the simulation, $\overline{u}_{t,s} = \mathsf{L}(j, s) = \overline{u}_{j,s}$. Since $j < t$, we conclude from the induction hypothesis that $u_{t,s}$ and $\overline{u}_{t,s}$ are identically distributed.
- **Case 2:** For each $\ell \in [t-1]$, $\mathsf{ind}_{\mathsf{diff}}(m_\ell, m_t) \leq s$, and there exists a $j \in [t-1]$ such that $\mathsf{ind}_{\mathsf{diff}}(m_j, m_t) = s$. If there are multiple j for which $\mathsf{ind}_{\mathsf{diff}}(m_j, m_t) = s$, let j be the smallest one. This means that m_j and m_t share a prefix $p \in \{0,1\}^{s-1}$ of length $s-1$. Then, in hybrid H_1, we have

$$u_{t,s} = f(s, p \| 0^{n-s}) + b_{t,s} \pmod M.$$

  In the simulation,

$$\overline{u}_{t,s} = \mathsf{L}(j, s) - (1 - 2 \cdot \mathbf{1}(m_j < m_t)) = \overline{u}_{j,s} - (1 - 2 \cdot \mathbf{1}(m_j < m_t)) \pmod M.$$

In hybrid H_2, $u_{j,s} = f(s, p\|0^{n-s}) + b_{j,s}$. By assumption, $b_{j,s} \neq b_{t,s}$, so we can write $b_{t,s} = b_{j,s} - (1 - 2 \cdot \mathbf{1}(m_j < m_t))$. Thus, in hybrid H_2, we have

$$u_{t,s} = f(s, p\|0^{n-s+1}) + b_{t,s} = u_{j,s} - (1 - 2 \cdot \mathbf{1}(m_j < m_t)) \pmod{M}.$$

By the inductive hypothesis, $u_{j,s}$ and $\overline{u}_{j,s}$ are identically distributed, so we conclude that $u_{t,s}$ and $\overline{u}_{t,s}$ are identically distributed.

– **Case 3:** For each $\ell \in [t-1]$, $\mathsf{ind}_{\mathsf{diff}}(m_\ell, m_t) < s$. Let $p \in \{0,1\}^{s-1}$ be the first $s-1$ bits of m_t. In hybrid H_1, we have

$$u_{t,s} = f(s, p\|0^{n-s}) + b_{t,s} \pmod{M},$$

while in the simulation $\overline{u}_{t,s}$ is a uniformly random string. By assumption, none of the messages $m_1, \ldots, m_{t-1}$ begin with the prefix p. Since f is a truly random function, the value of $f(s, p\|0^{n-s})$ is uniform in $\mathbb{Z}_M$ and independent of all other ciphertexts. Thus, $u_{t,s}$ and $\overline{u}_{t,s}$ are identically distributed.

We conclude that for all $s \in [n]$, $u_{t,s} \equiv \overline{u}_{t,s}$. Since the components of each ciphertext are constructed independently in both hybrid H_1 and in the simulation, this suffices to show that ct_t and $\overline{\mathsf{ct}}_t$ are identically distributed. The claim then follows by induction on t. □

Space usage. The order-revealing encryption scheme Π_{ore} on n-bit inputs produces encryptions of size $\lceil n \cdot \log_2 M \rceil$. By setting $M = 3$, an encryption of an n-bit message under Π_{ore} consists of only $\lceil n \cdot \log_2 3 \rceil \approx 1.59\, n$ bits. In the full version, we describe a "d-ary" generalization of Π_{ore} that further reduces the size of the ciphertexts in the ORE scheme, but with a slight loss in security. Specifically, we construct an ORE scheme where an encryption of an n-bit message has length approximately $n \cdot \log_d(2d-1)$ for any integer $d \geq 2$. Since $\log_d(2d-1)$ is a monotonically decreasing function in d, larger values of d yield shorter ciphertexts, but increased leakage.

3.2 Conversion to OPE

In this section, we explain how to convert Π_{ore}, an ORE scheme, into an OPE scheme. This means that ciphertexts of the resulting OPE scheme can be compared using the normal comparison function on numbers. To do this, we apply a simple transformation of any ciphertext ct of Π_{ore} into a number c that lies in the range $[0, M^n - 1]$ for which direct numeric comparisons of two numbers c_1 and c_2 reveal the order relation of the underlying plaintexts.

Recall that in Π_{ore}, ciphertexts are of the form $\mathsf{ct} = (u_1, u_2 \ldots, u_n)$, where for each $i \in [n]$, u_i lies in the range $\mathbb{Z}_M$. The ciphertext in the resulting OPE scheme is taken to be the $\lceil n \cdot \log_2 M \rceil$-bit number

$$c = \sum_{i=1}^{n} u_i \cdot M^{n-i} \qquad \in [0, M^n - 1]. \tag{3.2}$$

Intuitively, we view $u_1 u_2 \cdots u_n$ as a base-M representation of the OPE ciphertext. Correctness follows similarly to Π_{ore}, except here, there is a non-zero probability of error (as opposed to Π_{ore} where correctness held with probability 1). We claim that for any two messages $m_1, m_2 \in [0, 2^n - 1]$,

$$m_1 < m_2 \qquad \Longleftrightarrow \qquad c_1 < c_2,$$

with probability $1 - 1/M$, where $c_1, c_2 \in [0, M^n - 1]$ are the ciphertexts obtained by first invoking $\mathsf{ORE.Encrypt}$ on m_1, m_2, respectively, and then applying the transformation in Eq. (3.2). To see this, let $i \in [n]$ be the first bit position on which m_1 and m_2 differ. Observe that the numeric comparison of the OPE ciphertexts behaves identically as the ORE comparison procedure, except when the output of the PRF on the first $i - 1$ bits of the messages is the value $M - 1$[7]. However, by PRF security, this event happens with probability $1/M$, and thus, correctness holds with probability $1 - 1/M$. For instance, if $M = 2^\lambda$ (that is, λ bits), correctness holds with overwhelming probability. For practical scenarios, it may be suitable to only take $M \approx 2^{40}$ (the failure probability in this case is 2^{-40}).

Security of the resulting OPE scheme follows identically from security of Π_{ore}, as the transformation from ciphertexts ct to numbers c is bijective. We note that while this scheme is order-preserving, it does not behave like a random order-preserving function, and thus, does not inherit the security limitations associated with such OPE schemes [8]. In fact, our simulation-based security model and associated security theorem (Theorem 3.1) enables us to precisely specify the information leakage in this order-preserving encryption scheme.

In the full version, we describe a "d-ary" generalization of Π_{ore}. While this generalization does not reduce the size of the resulting ciphertexts in the ORE scheme, it does yield shorter ciphertexts in the OPE instantiation (by approximately a $\log_2 d$ multiplicative factor), with a slight loss in security. Correctness in this generalized scheme holds with probability $1 - d/M$.

4 Comparison to Existing OPE Schemes

We now compare the leakage of our order-revealing encryption scheme to that of existing order-preserving encryption schemes by Boldyreva et al. [7,8]. As explained in Sect. 2.3, composing any existing OPE scheme with an ORE scheme results in a new ORE scheme which is at least as secure as the underlying OPE scheme[8]. In this section, we show that even *without* the composition, our construction still achieves stronger security according to the metrics proposed by Boldyreva et al.

[7] If no reduction modulo M occurs in the $\mathsf{ORE.Encrypt}$ encryption, then numerically comparing the transformed ciphertexts is identical to evaluating the $\mathsf{ORE.Compare}$ procedure (since all relations hold over the integers).

[8] In most cases, the security of the composed scheme is strictly greater than that of the base OPE scheme since our ORE construction provides semantic security for a single ciphertext, whereas existing OPE schemes generally do not.

The security definition achieved by an order-preserving encryption scheme is that the encryption function behaves like a random order-preserving function (ROPF) from the plaintext space to the ciphertext space. While this definition has the same flavor as that for PRFs, the behavior of a truly random function is very different from that of a random order-preserving function. In particular, the output of an order-preserving function is not independent of its input, and thus, reveals some information about the input. It turns out that quantifying the exact information leakage is a non-trivial task in general. However, under certain assumptions (for example, if the messages are drawn from a uniform distribution), it is possible to obtain concrete upper bounds on the information leakage [8]. In particular, Boldyreva et al. propose two security notions, window one-wayness and window distance one-wayness, to analyze the security of an OPE scheme. In our setting, the nature of our security definition allows us to analyze the construction under a more generalized set of definitions compared to [8]. We present our analysis for window one-wayness here, and defer the analysis of window distance one-wayness to the full version.

4.1 One-Wayness

One of the most basic requirements of an encryption scheme is that it is one-way. Given a ciphertext, an adversary that does not have the secret key should not be able to recover the underlying message. In the standard definition of one-wayness [27], the adversary is given the encryption of a random message, and its goal is to guess the message. This is a very weak notion of security, and even if an encryption is one-way, the adversary might still be able to deduce nontrivial information about the message given only the ciphertext. To address this, Boldyreva et al. [7] introduce a more general notion of one-wayness where the adversary is allowed to guess a contiguous interval (a window) in the one-wayness challenge. The adversary succeeds if the message is contained within the interval. Moreover, the adversary is given multiple encryptions (of random messages) and succeeds if it outputs an interval that contains at least one of the messages.

The notion of window one-wayness is useful for arguing that an adversary does not learn many of the *most significant* bits of the message, but if all bits of the message are equally sensitive, then this definition is less useful. In our work, we present a more general definition of one-wayness, where instead of outputting an interval, the adversary is allowed to specify a set of guesses. To allow the adversary to specify a super-polynomially-sized set of guesses, we instead require the adversary to submit a circuit C that encodes its set ($C(x) = 1$ if and only if x is in the set). By requiring that the circuit encodes a contiguous interval, we recover the window one-wayness definition by Boldyreva et al. [8]. We now give our generalized definition.

Definition 4.1 (Generalized One-Wayness). *Fix a plaintext space $\mathcal{D}$ and let $\Pi = (\mathsf{ORE.Setup}, \mathsf{ORE.Encrypt}, \mathsf{ORE.Compare})$ be an ORE over $\mathcal{D}$. The (r, z)-*generalized one-wayness advantage *of an adversary $\mathcal{A}$ against Π is given by*

$$\mathsf{Adv}^{\mathsf{gow}}_{r,z,\Pi}(\mathcal{A}) \stackrel{\mathsf{def}}{=} \Pr[\mathsf{Expt}^{\mathsf{gow}}_{r,z,\Pi,\mathcal{A}}(1^\lambda) = 1],$$

where the (r, z)-generalized one-wayness experiment $\mathsf{Expt}^{\mathsf{gow}}_{r,z,\Pi,\mathcal{A}}(1^\lambda)$ is defined as follows:

Experiment $\mathsf{Expt}^{\mathsf{gow}}_{r,z,\Pi,\mathcal{A}}(1^\lambda)$**:**

1. $\mathsf{sk} \leftarrow \mathsf{ORE.Setup}(1^\lambda)$
2. sample $m_1, \ldots, m_z$ uniformly from $\mathcal{D}$ without replacement
3. for $i \in [z]$, $\mathsf{ct}_i \leftarrow \mathsf{ORE.Encrypt}(\mathsf{sk}, m_i)$
4. $C \leftarrow \mathcal{A}(\mathsf{ct}_1, \ldots, \mathsf{ct}_z)$, where $C : \mathcal{D} \to \{0,1\}$ is a circuit of size $\mathrm{poly}(\lambda)$
5. output 1 if $C(m_i) = 1$ for some $i \in [z]$ and $|\{x \in \mathcal{D} : C(x) = 1\}| \leq r$; otherwise, output 0

Remark 4.2 (Comparison with Existing One-Wayness Notions). By restricting the parameters (r, z) and the classes of circuits the adversary is allowed to output, Definition 4.1 captures many existing notions of one-wayness. For example, when $r = z = 1$, we recover the usual notion of one-wayness [27]. When the underlying plaintext space is the ring $\mathbb{Z}_M$ for some integer M and we require that the circuit output by the adversary encodes a contiguous interval of length at most r in $\mathbb{Z}_M$, our definition corresponds to the notion of window one-wayness introduced by Boldyreva et al. [8].

We now state our security theorem, but defer the proof to the full version.

Theorem 4.3 *Fix a security parameter λ and a plaintext space $\{0,1\}^n$ where $n = \omega(\log \lambda)$. Let Π_{ore} be the ORE scheme given at the beginning of Sect. 3. Then, for any constant $\varepsilon \in (0, 1]$, any $z = \mathrm{poly}(\lambda)$, and all efficient adversaries $\mathcal{A}$,*

$$\mathsf{Adv}^{\mathsf{gow}}_{r,z,\Pi_{\mathsf{ore}},\mathcal{A}}(1^\lambda) = \mathrm{negl}(\lambda),$$

where $r = 2^{n(1-\varepsilon)}$.

Comparison to existing schemes. When discussing the notion of one-wayness, we will always assume that the message-space is super-polynomial in the security parameter. Otherwise, the trivial adversary that just guesses a random point in the message space will succeed with non-negligible probability.

In [8], Boldyreva et al. give an upper bound on the one-wayness advantage of any (possibly computationally unbounded) adversary $\mathcal{A}$ against a random order-preserving function ROPF. This corresponds to setting $r = 1$ in our definition. They show [8, Theorem 4.1] that for $z = \mathrm{poly}(\lambda)$, $\mathsf{Adv}^{\mathsf{gow}}_{1,z,\mathsf{ROPF},\mathcal{A}} = \mathrm{negl}(\lambda)$.

The same statement holds for our ORE construction assuming a computationally bounded adversary: simply instantiate Theorem 4.3 with $\varepsilon = 1$.

In addition to giving an upper bound on an adversary's ability to guess the plaintext from the ciphertext, Boldyreva et al. also give a lower bound on the advantage for the case when r is large. In particular, they exhibit an efficient adversary $\mathcal{A}$ against an ROPF such that $\mathsf{Adv}^{\mathsf{gow}}_{r,z,\mathsf{ROPF},\mathcal{A}}(1^\lambda) = 1 - 2e^{-b^2/2}$ for a constant b when $r = O(\sqrt{2^n})$ and for any z [8, Theorem 4.2][9]. In other words, the authors describe a concrete adversary that is able to break the generalized one-wayness of any POPF-CCA-secure scheme (with probability close to 1) if the adversary is allowed to specify a set with $r = O(\sqrt{2^n})$ elements, even when $z = 1$. An intuitive way to understand this result is that given the output of an ROPF, an adversary can deduce roughly half of the bits of the associated input. In contrast, in our ORE scheme, if the adversary only sees a polynomial number of ciphertexts ($z = \text{poly}(\lambda)$), then invoking Theorem 4.3 with $\varepsilon = 1/2$, we have that for all efficient adversaries $\mathcal{A}$, $\mathsf{Adv}^{\mathsf{gow}}_{r,z,\Pi_{\mathsf{ore}},\mathcal{A}}(1^\lambda) = \text{negl}(\lambda)$ where $r = \sqrt{2^n}$. In fact, as Theorem 4.3 demonstrates, the adversary's advantage remains negligible even if we further increase the size of the sets the adversary is allowed to submit.

Intuitively, our results show that if the adversary only sees a polynomial number of ciphertexts, then it does not learn any constant fraction ε of the bits in the underlying plaintext from each ciphertext. In contrast, with an ROPF, and correspondingly, any OPE scheme that realizes a ROPF, each ciphertext alone leaks *half* of the most-significant bits of the underlying plaintext.

Similarly, while the OPE scheme by Teranishi et al. [47] can be shown to hide any constant fraction of the least significant bits of the plaintext, no such guarantee exists for the other bits of the plaintext. Note though that the security notion proposed in [47] is indistinguishability-based and hence, stronger than the one-wayness security notions. In fact, our basic ORE construction (by itself) does not achieve their indistinguishability-based definition. However, by composing our ORE construction with their OPE construction, we obtain a resulting ORE scheme which is strictly more secure, since it inherits the security properties of the underlying OPE scheme as well as semantic security for a single ciphertext (Sect. 2.3, Remark 2.6).

5 Conclusions

In this work, we introduced a new notion of security for order-preserving, and more generally, order-revealing encryption. Our simulation-based security notion is defined with respect to a leakage function which precisely characterizes what the ciphertexts in the scheme leak about the underlying messages. We then give a practical order-revealing encryption scheme which achieves this security notion for a specific leakage function. By composing our ORE construction with

[9] Strictly speaking, the adversary they describe is for the window one-wayness experiment, but any adversary that succeeds in the window one-wayness experiment also succeeds in the generalized one-wayness experiment (Definition 4.1).

existing OPE schemes, we obtain an ORE scheme with increased security. It is our hope that having a concrete leakage model will enable practitioners to make better-informed decisions on whether an ORE scheme is appropriate for their particular application. We conclude with several open problems:

1. Can we construct a practical ORE scheme with stronger security guarantees?
2. Can we reduce the ciphertext length of our ORE scheme while still maintaining a similar level of security?
3. Is it possible to build a practical ORE scheme with best-possible security from standard assumptions?

Acknowledgments. We would like to thank Sam Kim for helpful discussions about ORE, and Adam O'Neill for useful insights in shrinking the ciphertext size of our main construction. We also thank the anonymous reviewers for their helpful comments. This work was partially supported by an NSF Graduate Research Fellowship. Opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of Facebook.

References

1. Agrawal, R., Kiernan, J., Srikant, R., Xu, Y.: Order-preserving encryption for numeric data. In: SIGMOD, pp. 563–574 (2004)
2. Albrecht, M.R., Farshim, P., Hofheinz, D., Larraia, E., Paterson, K.G.: Multilinear maps from obfuscation. In: TCC (2016)
3. Ananth, P., Jain, A.: Indistinguishability obfuscation from compact functional encryption. In: CRYPTO, pp. 308–326 (2015)
4. Applebaum, B., Brakerski, Z.: Obfuscating circuits via composite-order graded encoding. In: Dodis, Y., Nielsen, J.B. (eds.) TCC 2015, Part II. LNCS, vol. 9015, pp. 528–556. Springer, Heidelberg (2015)
5. Barak, B., Garg, S., Kalai, Y.T., Paneth, O., Sahai, A.: Protecting obfuscation against algebraic attacks. In: Nguyen, P.Q., Oswald, E. (eds.) EUROCRYPT 2014. LNCS, vol. 8441, pp. 221–238. Springer, Heidelberg (2014)
6. Binnig, C., Hildenbrand, S., Färber, F.: Dictionary-based order-preserving string compression for main memory column stores. In: ACM SIGMOD, pp. 283–296 (2009)
7. Boldyreva, A., Chenette, N., Lee, Y., O'Neill, A.: Order-preserving symmetric encryption. In: Joux, A. (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 224–241. Springer, Heidelberg (2009)
8. Boldyreva, A., Chenette, N., O'Neill, A.: Order-preserving encryption revisited: improved security analysis and alternative solutions. In: Rogaway, P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 578–595. Springer, Heidelberg (2011)
9. Boneh, D., Lewi, K., Raykova, M., Sahai, A., Zhandry, M., Zimmerman, J.: Semantically secure order-revealing encryption: multi-input functional encryption without obfuscation. In: Oswald, E., Fischlin, M. (eds.) EUROCRYPT 2015. LNCS, vol. 9057, pp. 563–594. Springer, Heidelberg (2015)
10. Boneh, D., Silverberg, A.: Applications of multilinear forms to cryptography. Contemp. Math. **324**(1), 71–90 (2003)
11. Boneh, D., Wu, D.J., Zimmerman, J.: Immunizing multilinear maps against zeroizing attacks. In: IACR Cryptology ePrint Archive 2014/930 (2014)

12. Brakerski, Z., Komargodski, I., Segev, G.: From single-input to multi-input functional encryption in the private-key setting. In: IACR Cryptology ePrint Archive 2015/158 (2015)
13. Bun, M., Zhandry, M.: Order-revealing encryption and the hardness of private learning. In: IACR Cryptology ePrint Archive 2015/417 (2015)
14. Chang, Y.-C., Mitzenmacher, M.: Privacy preserving keyword searches on remote encrypted data. In: Ioannidis, J., Keromytis, A.D., Yung, M. (eds.) ACNS 2005. LNCS, vol. 3531, pp. 442–455. Springer, Heidelberg (2005)
15. Chenette, N., Lewi, K., Weis, S.A., Wu, D.J.: Practical order-revealing encryption with limited leakage. In: IACR Cryptology ePrint Archive 2015/1125 (2015)
16. Cheon, J.H., Han, K., Lee, C., Ryu, H., Stehlé, D.: Cryptanalysis of the multilinear map over the integers. In: Oswald, E., Fischlin, M. (eds.) EUROCRYPT 2015. LNCS, vol. 9056, pp. 3–12. Springer, Heidelberg (2015)
17. Cheon, J.H., Lee, C., Ryu, H.: Cryptanalysis of the new CLT multilinear maps. In: IACR Cryptology ePrint Archive (2011) Observation of strains: 934 (2015)
18. Coron, J.-S.: Cryptanalysis of GGH15 multilinear maps (2015)
19. Coron, J.-S., Gentry, C., Halevi, S., Lepoint, T., Maji, H.K., Miles, E., Raykova, M., Sahai, A., Tibouchi, M.: Zeroizing without low-level zeroes: new MMAP attacks and their limitations. In: CRYPTO, pp. 247–266 (2015)
20. Coron, J.-S., Lepoint, T., Tibouchi, M.: Practical multilinear maps over the integers. In: Canetti, R., Garay, J.A. (eds.) CRYPTO 2013, Part I. LNCS, vol. 8042, pp. 476–493. Springer, Heidelberg (2013)
21. Coron, J.-S., de Lepoint, T., Tibouchi, M.: New multilinear maps over the integers. In: CRYPTO, pp. 267–286 (2015)
22. Curtmola, R., Garay, J.A., Kamara, S., Ostrovsky, R.: Searchable symmetric encryption: improved definitions and efficient constructions. In: ACM CCS, pp. 79–88 (2006)
23. Garg, S., Gentry, C., Halevi, S.: Candidate multilinear maps from ideal lattices. In: Johansson, T., Nguyen, P.Q. (eds.) EUROCRYPT 2013. LNCS, vol. 7881, pp. 1–17. Springer, Heidelberg (2013)
24. Garg, S., Gentry, C., Halevi, S., Raykova, M., Sahai, A., Waters, B.: Candidate indistinguishability obfuscation and functional encryption for all circuits. In: FOCS, pp. 40–49 (2013)
25. Garg, S., Gentry, C., Halevi, S., Zhandry, M.: Fully secure functional encryption without obfuscation. In: IACR Cryptology ePrint Archive 2014/666 (2014)
26. Gentry, C., Gorbunov, S., Halevi, S.: Graph-induced multilinear maps from lattices. In: Dodis, Y., Nielsen, J.B. (eds.) TCC 2015, Part II. LNCS, vol. 9015, pp. 498–527. Springer, Heidelberg (2015)
27. Goldreich, O.: The Foundations of Cryptography - Volume 1, Basic Techniques. Cambridge University Press, Cambridge (2001)
28. Goldreich, O., Goldwasser, S., Micali, S.: How to construct random functions (extended abstract). In: FOCS, pp. 464–479 (1984)
29. Goldwasser, S., et al.: Multi-input functional encryption. In: Nguyen, P.Q., Oswald, E. (eds.) EUROCRYPT 2014. LNCS, vol. 8441, pp. 578–602. Springer, Heidelberg (2014)
30. Goldwasser, S., Kalai, Y.T., Popa, R.A., Vaikuntanathan, V., Zeldovich, N.: Reusable garbled circuits and succinct functional encryption. In: STOC, pp. 555–564 (2013)
31. Goldwasser, S., Micali, S.: Probabilistic encryption. J. Comput. Syst. Sci. **28**(2), 270–299 (1984)

32. Gorbunov, S., Vaikuntanathan, V., Wee, H.: Functional encryption with bounded collusions via multi-party computation. In: Safavi-Naini, R., Canetti, R. (eds.) CRYPTO 2012. LNCS, vol. 7417, pp. 162–179. Springer, Heidelberg (2012)
33. Hu, Y., Huiwen, J.: Cryptanalysis of GGH map. In: IACR Cryptology ePrint Archive 2015/301 (2015)
34. Kadhem, H., Amagasa, T., Kitagawa, H.: A secure and efficient order preserving encryption scheme for relational databases. In: KMIS, pp. 25–35 (2010)
35. Kerschbaum, F.: Frequency-hiding order-preserving encryption. In: ACM CCS, pp. 656–667 (2015)
36. Kerschbaum, F., Schröpfer, A.: Optimal average-complexity ideal-security order-preserving encryption. In: ACM CCS, pp. 275–286 (2014)
37. Langlois, A., Stehlé, D., Steinfeld, R.: GGHLite: more efficient multilinear maps from ideal lattices. In: Nguyen, P.Q., Oswald, E. (eds.) EUROCRYPT 2014. LNCS, vol. 8441, pp. 239–256. Springer, Heidelberg (2014)
38. Mavroforakis, C., Chenette, N., O'Neill, A., Kollios, G., Canetti, R.: Modular order-preserving encryption, revisited. In: ACM SIGMOD, pp. 763–777 (2015)
39. Minaud, B., Fouque, P.-A.: Cryptanalysis of the new multilinear map over the integers. In: IACR Cryptology ePrint Archive 2015/941 (2015)
40. Naveed, M., Kamara, S., Wright, C.V.: Inference attacks on property-preserving encrypted databases. In: CCS (2015)
41. Pandey, O., Rouselakis, Y.: Property preserving symmetric encryption. In: Pointcheval, D., Johansson, T. (eds.) EUROCRYPT 2012. LNCS, vol. 7237, pp. 375–391. Springer, Heidelberg (2012)
42. Popa, R.A., Li, F.H., Zeldovich, N.: An ideal-security protocol for order-preserving encoding. In: IEEE Symposium on Security and Privacy, pp. 463–477 (2013)
43. Popa, R.A., Redfield, C.M.S., Zeldovich, N., Balakrishnan, H.: Cryptdb: protecting confidentiality with encrypted query processing. In: SOSP, pp. 85–100 (2011)
44. Roche, D., Apon, D., Choi, S.G., Yerukhimovich, A.: POPE: Partial order-preserving encoding. In: Cryptology ePrint Archive, Report 2015/1106 (2015)
45. Sahai, A., Seyalioglu, H.: Worry-free encryption: functional encryption with public keys. In: ACM CCS, pp. 463–472 (2010)
46. Skyhigh Networks Inc. https://www.skyhighnetworks.com/. Accessed 11 Dec 2015
47. Teranishi, I., Yung, M., Malkin, T.: Order-preserving encryption secure beyond one-wayness. In: Sarkar, P., Iwata, T. (eds.) ASIACRYPT 2014, Part II. LNCS, vol. 8874, pp. 42–61. Springer, Heidelberg (2014)
48. Xiao, L., Yen, I-L., Huynh, D.T.: Extending order preserving encryption for multi-user systems. In: IACR Cryptology ePrint Archive, (2011) Observation of strains: 192 (2012)
49. Zimmerman, J.: How to obfuscate programs directly. In: Oswald, E., Fischlin, M. (eds.) EUROCRYPT 2015. LNCS, vol. 9057, pp. 439–467. Springer, Heidelberg (2015)