Randomness Bounds for Private Simultaneous Messages and Conditional Disclosure of Secrets

Akinori Kawachi[1] and Maki Yoshida[2]

[1] Mie University
kawachi@info.mie-u.ac.jp
[2] National Institute of Information and Communications Technology
maki-yos@nict.go.jp

Abstract. In cryptography, the private simultaneous messages (PSM) and conditional disclosure of secrets (CDS) are closely related fundamental primitives. We consider k-party PSM and CDS protocols for a function f with a common random string, where each party P_i generates a message and sends it to a referee P_0.

We consider bounds for the optimal length ρ of the common random string among k parties (or, *randomness complexity*) in PSM and CDS protocols with perfect and statistical privacy through combinatorial and entropic arguments. (i) We provide general connections from the optimal total length λ of the messages (or, *communication complexity*) to the randomness complexity ρ. (ii) We also prove randomness lower bounds in PSM and CDS protocols for general functions. (iii) We further prove randomness lower bounds for several important explicit functions.

They contain the following results: For PSM protocols with perfect privacy, we prove $\lambda - 1 \leq \rho$ as the general connection. From the general lower bound, we prove $\rho \geq 2^{(k-1)n} - 1$ for a general function $f : (\{0,1\}^n)^k \to \{0,1\}$ under universal reconstruction, in which P_0 is independent of f. This implies that the Feige-Killian-Naor PSM protocol for a general function [Proc. STOC '94, pp.554–563] is optimal with respect to randomness complexity. We also provide a randomness lower bound $\rho > kn - 2$ for a generalized inner product function. This implies the optimality of the 2-party PSM protocol for the inner-product function of Liu, Vaikuntanathan, and Wee [Proc. CRYPTO 2017, pp.758–790].

For CDS protocols with perfect privacy, we show $\rho \geq \lambda - \sigma$ as the general connection by similar argument to those for PSM protocols, where σ is the length of secrets. We also obtain randomness lower bounds $\rho \geq (k-1)\sigma$ for XOR, AND, and generalized inner product functions. These imply the optimality of Applebaum and Arkis's k-party CDS protocol for a general function [Proc. TCC 2018, pp.317–344] up to a constant factor in a large k.

1 Introduction

1.1 Background and Related Work

The private simultaneous messages (PSM) and conditional disclosure of secrets (CDS) are closely related primitives in cryptography. (See Figure 1 for models of PSM and CDS.) These primitives have been studied broadly for the last two decades from viewpoints of information-theoretic security since these are regarded as natural and fundamental primitives in the information-theoretic cryptography. In particular, the investigation of efficiency, or *complexity* (for instance, necessary and sufficient amount of communication overhead, common random resources, etc.) for these primitives is quite important towards deep understanding for strong notions such as information-theoretic secrecy and privacy.

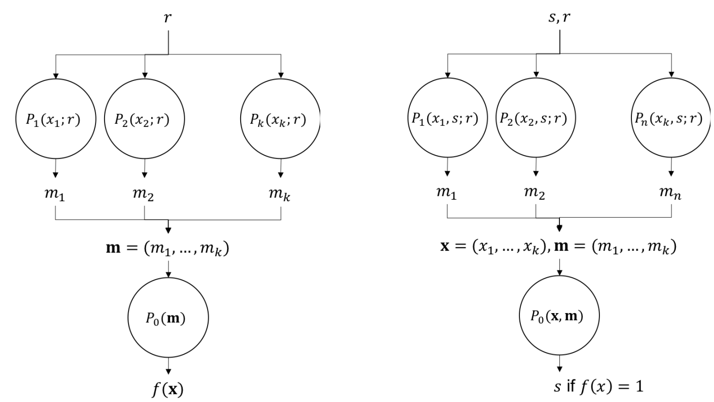
PSM is a model of multi-party secure computation in a minimal scenario. The 2-party PSM protocol was originally introduced by Feige, Kilian, and Naor [21], and the notion of PSM was explicitly defined and extended to k-party protocols by Ishai and Kushilevitz [27]. PSM is an important primitive to provide several cryptographic constructions such as generalized oblivious transfers based on standard oblivious transfers [27], distributed multiparty secure computation over some general network topology [35], and more. In particular, it is known that PSM is equivalent with the (decomposable) randomized encodings, which also has a number of cryptographic applications. (See survey papers, e.g., [2,26].)

CDS is originally introduced by Gertner, Ishai, Kushilevitz, and Malkin for efficient realization of symmetrically private information retrieval schemes [24]. As well as PSM, CDS provides several cryptographic applications such as attribute-based encryption schemes [9, 37], priced oblivious transfer [1] and secret sharing schemes for uniform access structures [3, 12, 13] other than the symmetrically private information retrieval.

In a PSM protocol, k parties $P_1, \dots, P_k$ individually have private inputs $x_1 \in \{0,1\}^{n_1}, \dots, x_k \in \{0,1\}^{n_k}$ and a common random string r. P_i first computes a message $m_i \in \{0,1\}^{\lambda_i}$ on $x_i \in \{0,1\}^{n_i}$ and $r \in \{0,1\}^{\rho}$, and sends m_i to the referee P_0. Then, P_0 outputs $f(x_1, \dots, x_k)$ from $m_1, \dots, m_k$ for a function $f : \prod_{i\in[k]}\{0,1\}^{n_i} \to \{0,1\}^v$ without learning anything except the output value $f(x_1, \dots, x_k)$. In other words, there are distributions D_i ($i \in [2^v]$) such that the joint distribution of $(P_j(x_j))_{j\in[k]}$ is identical (or statistically close) to $D_{f(x_1,\dots,x_k)}$.

In a CDS protocol, $P_1, \dots, P_k$ send messages $m_1 \in \{0,1\}^{\lambda_1}, \dots, m_k \in \{0,1\}^{\lambda_k}$ on individual inputs $x_1, \dots, x_k$, common random string $r \in \{0,1\}^{\rho}$, and secret $s \in \{0,1\}^{\sigma}$ to the referee P_0, respectively. If a predicate f is satisfied with $x_1, \dots, x_k$, P_0 succeeds recovering the secret from $x_1, \dots, x_k$ and $m_1, \dots, m_k$. Otherwise, P_0 learns nothing, i.e., there is a distribution D_0 such that the joint distribution $(P_i(x_i, s; r))_i$ is identical (or statistically close) to D_0 for every $x_1, \dots, x_k$ for which $f(x_1, \dots, x_k) = 0$.

In this paper, we are particularly interested in the minimum length of the messages, or *communication complexity*, and the common random string, or *randomness complexity*, in order to achieve the information-theoretic privacy as a



(a) Private Simultaneous Messages (PSM)

(b) Conditional Disclosure of Secrets (CDS)

Fig. 1. Multi-party computation with the minimal communication pattern

measure of the efficiency of the protocols. It is very important to understand limitations for efficiency of PSM and CDS protocols by demonstrating the lower bounds of those complexity measures towards the optimal protocol constructions.

There are many studies of upper bounds for the communication complexity from explicit constructions of PSM and CDS protocols and lower bounds of the communication complexity. On the contrary, lower bounds for the randomness complexity are much less known in studies of PSM and CDS protocols so far, while several results are known for randomness complexity in more general models of the secure computation [22,25,28–30]. (See Section A and Table 1 for more details on the related results of PSM and CDS.)

The randomness complexity is important cryptographic resources as so is the communication complexity. For instance, the common random string in CDS protocols corresponds to a random string shared among databases which is hidden from users querying to databases in the application to symmetrically private information retrieval schemes [24] and the public key used for every encryption in the application to attribute-based encryptions [9]. Thus, it is directly connected to the resource for secret and public information in the applications.

To the best knowledge of the authors, there is only one known study by Pillai, Prabhakaran, Prabhakaran, and Sridhar [36] for randomness lower bounds in PSM and CDS protocols. They focused on the 2-bit input functions, and proved the optimality of the 2-party PSM protocol for the 2-bit input AND function given by [21] with respect to the length of a common random string.

Table 1. Complexity bounds for CDS and PSM protocols

Scheme	Ref.	Bound	Functions	Comm.	Rand.
2-party PSM	FKN94 [21]	Upper†	General	2^n+n+1	2^n+n
	BIKK14, LVW17 [12,32]	Upper	General	$O(2^{n/2})$	$O(2^{n/2})$
	LVW17 [32] (Corollary 3)	Upper	f_{IP} [1]	$2n+2$	$2n$
	PPPS19, FKN94 [21,36]	Matching	1-bit input f_{and} [3]	$2\log_2 3$	$\log_2 6$
	folklore	Matching	1-bit input f_{xor} [2]	2	1
	BIKK14 [12]	Lower†	Non-explicit	2^n	
	AHMS20 [6]	Lower	Non-explicit	$3n-O(\log n)$	
	AHMS20 [6]	Lower	Explicit**	$3n-O(\log n)$	
k-party PSM	BKN18 [13] ($k>3$)	Upper	General	$O(k^3 2^{kn/2})$	$O(k^3 2^{kn/2})$
	Theorem 2	Lower	General	λ	$\lambda-1$
	Theorem 3	Lower†	General	$2^{(k-1)n}$	$2^{(k-1)n}-1$
	Theorem 4	Lower	f_{IP} [1]	kn	$kn-2$
2-party CDS	GKW15 [23]	Upper	General	$\sigma 2^{(n/2)+1}$	$\sigma 2^{(n/2)+1}$
	LVW17 [32]	Upper	General	$\sigma 2^{o(n)}$	$\sigma 2^{o(n)}$
	GKW15 [23] (Appendix C)	Upper	f_{IP} [1]	$(n+2)\sigma$	$(n+2)\sigma$
	GKW15 [23] (Appendix C)	Upper	1-bit input f_{and} [3]	3σ	3σ
	GKW15 [23]	Lower	General	$\Omega(\log n)$	
	GKW15 [23] (Section 4)	Lower	General	2σ	
	GKW15 [23] (Section 5)	Lower*	f_{IP} [1]	$\Omega(n)$	
	AARV17 [4]	Lower	Non-explicit	$\Omega(n)$	
	AV21 [8]	Lower§	f_{IP} [1]	$\Omega(n)$	
k-party CDS	AA18 [3] ($\sigma \geq 2^{kn-1}$)	Upper	General	$(4k-2)\sigma$	$(4k-4)\sigma$
	BP18 [14]	Upper*	General	$O(\sigma 2^{(k-1)n/2})$	
	LVW18 [33]	Upper	General	$2^{O(\sqrt{kn}\log(kn))}$	
	BFMP17 [11] (See also [14])	Lower*	Non-explicit	$\Omega(k^{-1}2^{(k-1)n/2})$	
	Lemma 1	Lower	General	λ	$\lambda-\sigma$
	Theorem 5 ($2^\sigma \leq 2\lvert\mathcal{C}\rvert$)	Lower†‡	General	$\Omega\left(\frac{\sigma 2^{kn}}{\log\lvert\mathcal{C}\rvert}\right)$	$\Omega\left(\frac{\sigma(2^{kn}-\log\lvert\mathcal{C}\rvert)}{\log\lvert\mathcal{C}\rvert}\right)$
	Theorem 6	Lower	f_{IP} [1]	$k\sigma$	$(k-1)\sigma$
	Theorem 6, Theorem 15	Matching	f_{and} [3]	$k\sigma$	$(k-1)\sigma$
	Theorem 6	Lower	f_{xor} [2]	$k\sigma$	$(k-1)\sigma$
	Theorem 7	Lower	θ-nontrivial	$\theta\sigma$	$(\theta-1)\sigma$

$*$: The protocol has linear reconstruction. $**$: The assumption that some hitting-set generator exists is required. $\dagger$: The protocol has universal reconstruction. $\ddagger$: $\mathcal{C}$ is a set of possible referees. §: The protocol admits either a small constant privacy error or a small constant correctness error. [1]: $f_{\mathrm{IP}}(x_1,\ldots,x_k)=\langle x_1,x_2\rangle \oplus \cdots \oplus \langle x_{k-1},x_k\rangle$. [2]: $f_{\mathrm{xor}}(\mathbf{x})=\bigoplus_{i=1}^k x_i$. [3]: $f_{\mathrm{and}}(\mathbf{x})=\bigwedge_{i=1}^k x_i$.

1.2 Our Results

In this paper, we focus on the randomness complexity in PSM and CDS protocols with perfect and statistical privacy through combinatorial and entropic arguments (Table 2). Particularly, (i) we provide general connections from the communication complexity to randomness complexity. (ii) We prove randomness lower bounds in PSM and CDS protocols for general functions. (iii) We further prove randomness lower bounds for several important explicit functions.

Table 2. Summary of results on randomness complexity ρ and communication complexity λ and used techniques

k-party PSM	Connection between ρ and λ in PSM	Theorem 1	Entropic
		Theorem 2	Combinatorial
	ρ and λ for general functions	Theorem 3	Combinatorial
	ρ and λ for an explicit function	Theorem 4	Entropic
k-party CDS	Connection between ρ and λ in CDS	Lemma 1	Entropic
	ρ and λ for general functions	Theorem 5	Combinatorial
	ρ and λ for explicit functions	Theorems 6 & 7	Entropic

We describe more details of our results below. In the following, we denote by $\bar{n} = n_1 + \cdots + n_k$ the total input bit length of k parties. Let σ denote the bit length of secrets. See Table 1 for summary of previous results and ours. In the table, we suppose the case of the perfect privacy for CDS and PSM protocols and the secret length is equal to one if the symbol σ does not appear in the complexity.

We can simply prove the following connection from the communication complexity to randomness complexity for k-party PSM protocols by an entropic argument.

Theorem 1. *Let λ be the communication complexity for a function $f : \{0,1\}^{\bar{n}} \to \{0,1\}$ of k-party PSM protocols. Then, the randomness complexity ρ for f of a k-party PSM protocol is larger than $\lambda - \bar{n} - 1$.*

On the other hand, we can prove a stronger connection by employing a combinatorial argument.

Theorem 2. *Consider k-party PSM protocols with the perfect correctness for a function $f : \{0,1\}^{\bar{n}} \to \{0,1\}$. Then, the randomness complexity ρ of those with the perfect privacy is at least $\lambda - 1$ for the communication complexity λ. The randomness complexity ρ of those with the δ-statistical privacy is at least $\lambda - 1 - \log(1 - 2\delta)^{-1}$.*

The key observation for the proof of Theorem 2 is a similarity between the privacy of PSM protocols and secrecy of symmetric-key encryption schemes. As a fundamental fact shown by Shannon (e.g., see Section 9.1 in [20]), in order to achieve the perfect secrecy, the length of secret keys must be at least that of the plaintexts to be encrypted. In some sense, the privacy of the inputs of $P_1, \cdots, P_k$ is "encrypted" by the common random string shared among them. Thus, a similar combinatorial argument presented by Shannon appears to work well for the randomness lower bound for PSM protocols. However, there are technical differences between encryption schemes and PSM protocols. For example, the original proof for the secrecy of encryption schemes explicitly used the injective property of the encryption function; however, we cannot assume such a property for $P_1, \ldots, P_k$ in PSM protocols. Furthermore, the referee needs to learn the output value of f from given messages that depend on f. Namely, the message distributions that may depend on f for some output value should be distinct from those for the other output values. In order to resolve the challenges, we exploit the communication lower bounds of PSM protocols.

By Theorem 2 with the communication lower bounds of k-party PSM protocols with universal reconstruction for a general function, we can obtain explicit randomness lower bounds as follows.

Theorem 3. *Suppose that $n = n_1 = \cdots = n_k$. Consider k-party PSM protocols that have universal reconstruction with the perfect correctness for a function $f : \{0,1\}^{\bar{n}} \to \{0,1\}$. Then, it holds that $\lambda \geq 2^{(k-1)n}$. Furthermore, if they have the perfect privacy (the δ-statistical privacy, respectively), it holds that $\rho \geq 2^{(k-1)n} - 1$ ($\rho \geq 2^{(k-1)n} - 1 - \log(1-2\delta)^{-1}$, respectively).*

Note that the communication lower bound of $2^{(k-1)n}$ is a simple generalization of the one given in [12] for the case when $k = 2$. This bound shows the 2-party Feige-Kilian-Naor PSM protocol with universal reconstruction for a general function [21] is optimal (up to additive factors) with respect to not only the communication complexity but also the randomness complexity.

We also provide a direct proof of communication and randomness lower bounds for a generalized inner product function by an entropic argument, which shows the randomness optimality (up to an additive constant factor) of the 2-party PSM protocol for the inner product function given by Liu et al. [32].

Theorem 4. *Suppose that $n = n_1 = \cdots = n_k$. Consider k-party PSM protocols with perfect privacy for a generalized inner product function $f_{\mathrm{IP}} : (x_1, \ldots, x_k) \mapsto \langle x_1, x_2 \rangle \oplus \cdots \oplus \langle x_{k-1}, x_k \rangle$, where $\langle \cdot, \cdot \rangle$ is the inner product modulo 2, i.e., $\langle x, y \rangle = \oplus_{j \in [n]} x[j] \cdot y[j]$ for $x = (x[1], \ldots, x[n]), y = (y[1], \ldots, y[n]) \in \{0,1\}^n$. Then, it holds that $\lambda \geq kn$. Furthermore, if they have the perfect privacy (the δ-statistical privacy, respectively), it holds that $\rho > kn - 2$ ($\rho + 2\delta\lambda > (1-2\delta)kn - 2(1-2\delta) + 2\delta \log \delta$, respectively).*

The connection from the communication complexity is also quite useful for the randomness complexity of CDS protocols. We obtain another randomness lower bound characterized by λ and the secret length σ, which can be directly derived from the data processing inequality.

Lemma 1. *For every function f, we have $\rho > \lambda - \sigma - 1$, where ρ (λ, respectively) is the randomness complexity (communication complexity, respectively) of k-party CDS protocols for f, and σ is the length of secrets.*

Note that this lemma holds whether the privacy is perfect or statistical. From Lemma 1, if we want to show randomness lower bounds of CDS protocols, it suffices to prove communication lower bounds of the protocols.

The best known lower bound of the communication complexity is linear in input length [6] for 2-party CDS protocols with a single-bit secret. From the above lemma, we also obtain linear randomness lower bounds. In order to show higher communication lower bounds (and consequently, randomness lower bounds), we examine two sorts of CDS protocols in this paper.

The first one is a k-party CDS protocol for a general function under restriction of reconstruction procedures. While exponential communication upper bounds of CDS protocols for general functions were known such as a k-party upper bound of $2^{O(\sqrt{kn}\log(kn))}$ by [33], much less is known for lower bounds for short secrets [4,6,23], and super-linear communication lower bounds remain open for short secrets. (See Table 1 for more results.) As an attempt to breaking the barrier, we provide communication lower bounds, and consequently, randomness lower bounds from Lemma 1, of CDS protocols for a general function under the restriction of possible referees.

Theorem 5. *Suppose $n_1 = \cdots = n_k$. Let $\mathcal{F} := \{f : \{0,1\}^{\bar{n}} \to \{0,1\}\}$ and $\mathcal{C}$ be a set of possible referees in CDS protocols. Suppose that it holds that $2^\sigma \leq 2|\mathcal{C}|$ for the secret length σ. If a k-party CDS protocol for $\mathcal{F}$ with universal reconstruction has the perfect correctness and perfect privacy, it holds that*

$$\lambda = \Omega(2^{\bar{n}} \cdot \sigma / \log|\mathcal{C}|),$$

and if it has δ-statistical privacy, it holds that

$$\lambda = \Omega\left(\frac{2^{\bar{n}} \cdot \log\left\{(2^{-\sigma} + \delta)^{-1}\right\}}{\log|\mathcal{C}|}\right).$$

For example, if we consider a single-bit secret ($\sigma = 1$) and a non-trivially small number of possible referees, say $|\mathcal{C}| \leq 2^{2^{0.1\lambda}}$, the communication lower bound must be exponential in $\bar{n}$. From Lemma 1, we also obtain the randomness lower bounds, e.g., $\rho \geq \lambda - \sigma = \Omega((2^{\bar{n}} - \log|\mathcal{C}|) \cdot \sigma / \log|\mathcal{C}|)$ in the case of the perfect privacy.

The proof is based on Larsen and Simkin's argument for lower bounds of share size in secret sharing schemes [31]. Similar to the communication complexity for CDS protocols, determining the tight bound of share complexity (i.e., minimum share size) for k-party secret sharing schemes for a general access structure is one of the longstanding open problems in cryptography. While the best known upper bound is $2^{0.892k}$ shown by Applebaum, Beimel, Farràs, Nir, and Peter [5], the lower bound for an explicit access structure is $k/\log k$, which was presented by Csirmaz [16, 17].

Larsen and Simkin proved the share lower bounds of $\Omega(2^k/(\sqrt{k}\log|\mathcal{C}|))$ for a non-explicit access structure by specifying a set $\mathcal{C}$ of possible reconstruction functions (or, the referees). Their result implies that an exponential share complexity can be obtained if a reconstruction function is implemented by polynomial-size circuits. As mentioned above, the best known upper and lower bounds for communication complexity of 2-party CDS protocols are $2^{o(n)}$ [32] and $\Omega(n)$ [4], respectively: therefore, it is evident that there remains a large gap between upper and lower bounds of the communication complexity. We arrange the argument of Larsen and Simkin against differences between definitions of secret sharing schemes and CDS protocols to obtain the communication lower bounds of k-party CDS protocols for a non-explicit function.

Note that this result provides nontrivial lower bounds only for short secrets from the condition $2^\sigma \leq 2|\mathcal{C}|$. For example, it can provide only trivial constant lower bounds in the case of exponentially long secrets such as the setting of [3].

The second one is k-party CDS protocols for some class of explicit functions with potentially long secrets. Unlike previous communication lower bounds for CDS protocols from Theorem 5 and derived by [4, 6, 23] with respect to the input length $\bar{n}$, we also show other communication lower bounds for explicit functions with respect to the secret length σ, and hence, we can obtain high communication lower bounds (and randomness lower bounds) for long secrets. Let $f_{\mathrm{xor}}(\mathbf{x}) = \bigoplus_{i=1}^k x_i$ and let $f_{\mathrm{and}}(\mathbf{x}) = \bigwedge_{i=1}^k x_i$ where $n_1 = n_2 = \cdots = n_k = 1$. We prove the following communication and randomness lower bounds of k-party CDS protocols for these explicit functions.

Theorem 6. *If a k-party CDS protocol with a σ-bit secret for $f \in \{f_{\mathrm{xor}}, f_{\mathrm{and}}, f_{\mathrm{IP}}\}$ has perfect correctness and perfect privacy, then $\lambda \geq k\sigma$ and $\rho \geq (k-1)\sigma$. If the protocol has δ-statistical privacy, we have*

$$\lambda \geq \frac{1-2\delta}{1+2\delta k} k\sigma + \frac{2\delta k}{1+2\delta k}\log\delta, \quad \textit{and} \quad \rho \geq (k-1)\sigma + 2\delta k(\lambda + \sigma - \log\delta).$$

An interesting point is that the overhead factor of k is inherent even in the CDS setting where a referee does not collude with any parties (minimal resiliency). This result provides tight lower bounds (up to a constant factor) for a large k with respect to both of the communication and randomness lower bounds for k-party CDS protocols with long secrets. Applebaum and Arkis's construction of a k-party CDS protocol for a general function with exponentially long secrets actually gave the communication upper bounds of $(4k-2)\sigma$ and the randomness upper bounds of $(4k-4)\sigma$ [3].

In [8], Applebaum and Vasudevan also provided a linear lower bound for f_{IP} via the communication complexity games in the cases of 2-party CDS protocols with imperfect privacy and those with imperfect correctness. While it is not clear whether our proof technique can apply to the case of imperfect correctness, our technique provides a more direct proof and generalization to k-party cases.

The key techniques for these lower bounds are based on information-theoretic arguments developed for communication lower bounds in multi-party secure computation protocols [18,19]. Data, Prabhakaran and Prabhakaran [19] proved the first *generic* communication lower bounds in a variety of 3-party secure computation protocols *without* common random strings. In particular, they showed an explicit function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^{n-1}$ such that a referee must interact with two parties at least $3n-1$ bits. Damgård, Larsen, and Nielsen [18] considered k-party secure computation protocols with and without preprocessing. Under both conditions, they obtained communication lower bounds for *explicit Boolean* functions combined with Fano's inequality for statistical security. However, their communication lower bounds decrease as the resiliency is weakened. Our techniques are tailored for CDS protocols so that smooth application of the data processing inequality to derive the lower bounds is made possible. In particular, we use *non-trivial* input pairs for conditioning and applying the chain rule.

Furthermore, we can prove communication and randomness lower bounds for a wide class of functions, we name θ-nontrivial functions, by extending the above entropic arguments. (For the details of θ-nontriviality, see Definition 5 in Section 4.)

Theorem 7. *If a k-party CDS protocol with a σ-bit secret for a θ-nontrivial function f has perfect correctness and perfect privacy, then $\lambda \geq \theta\sigma$ and $\rho \geq (\theta-1)\sigma$. If the protocol has δ-statistical privacy, we have*

$$\lambda \geq \frac{1-2\delta}{1+2\delta\theta}\theta\sigma + \frac{2\delta\theta}{1+2\delta\theta}\log\delta, \quad \textit{and} \quad \rho \geq (\theta-1)\sigma - 2\delta\theta(\lambda+\sigma-\log\delta).$$

The class of θ-nontrivial functions contains not only f_{xor} and f_{IP} but also natural functions appeared in the area of secret sharing schemes such as θ-threshold access functions, θ-uniform access functions, etc.

1.3 Organization

The remaining part of this paper is organized as follows. In Sect. 2, we define the models of PSM and CDS more formally, and then, we provide basic notions and a useful corollary in the information theory. In Sects. 3 and 4, we demonstrate results on PSM and CDS protocols with the perfect privacy, respectively. We put the details of our results on PSM and CDS protocols with the statistical privacy at Sects. B and C in the supplementary material, respectively. Each of Sects. 3, 4, B, and C has the same organization: We first demonstrate the general connections from the communication complexity to randomness complexity. Then, we provide lower bounds for general functions, and subsequently, those for explicit functions.

2 Preliminaries

In this section, we introduce the formal definitions of PSM and CDS protocols. In addition, we review the fundamental notions and entropic tools that we use for proofs of lower bounds.

2.1 Private Simultaneous Messages

Let $\mathcal{X}_i$ and $\mathcal{M}_i$ denote finite domains of inputs and messages of the i-th party ($i \in [k]$). Also, let $\mathcal{X}, \mathcal{M}$ and $\mathcal{R}$ denote finite domains of the total inputs, total messages, and common random strings, respectively. Let $\mathcal{Y}$ be a finite domain of an output and let $f : \mathcal{X} \to \mathcal{Y}$ be a function.

In a PSM protocol, there are $k+1$ parties $P_0, P_1, \ldots, P_k$. For $i \in [k]$, P_i is connected only to P_0 by secure point-to-point channels. P_0 is a special party called a *referee*. Each party P_i with $i \in [k]$ sends the referee P_0 a message m_i that is computed on an input $x_i \in \mathcal{X}_i$ and a common random string $r \xleftarrow{\$} \mathcal{R}$, where $r \xleftarrow{\$} \mathcal{R}$ means that r is sampled uniformly at random from $\mathcal{R}$. Note that the common random string r is not given to the referee P_0. Eventually, P_0 gives an output $y \in \mathcal{Y}$ from messages $m_1, \ldots, m_k$.

A PSM protocol is required to satisfy correctness and privacy. The correctness means that the output is correct, and the privacy means that some procedure can generate a distribution that is identical to (or statistically close to) the view of the referee only from the output.

More formally, a PSM protocol for a function $f : \mathcal{X} \to \mathcal{Y}$ is modeled as follows. For every $i \in [k]$, let P_i be an algorithm that gives a message $m_i \in \mathcal{M}_i$ on an input $x_i \in \mathcal{X}_i$ and a common random string $r \xleftarrow{\$} \mathcal{R}$ at the start of the protocol. Therefore, we can write $m_i = P_i(x_i; r)$. Let P_0 be an algorithm that receives a sequence of messages $\mathbf{m} = (m_1, \ldots, m_k) = (P_1(x_1; r), \ldots, P_k(x_k; r))$ from $P_1, \ldots, P_k$. For simplicity, let $\mathbf{P}(\mathbf{x}; r) = (P_1(x_1; r), \ldots, P_k(x_k; r))$. We call $\mathbf{P}$ a message function of the PSM protocol. Also, let $\mathbf{P}(\mathbf{x})$ denote a run of $\mathbf{P}(\mathbf{x}; r)$ with uniformly random $r \in \mathcal{R}$. Eventually, P_0 gives an output $f(\mathbf{x}) = y \in \mathcal{Y}$ for every input $\mathbf{x} = (x_1, \ldots, x_k)$ on a sequence of messages $\mathbf{m}$ for f.

The correctness and privacy are defined as follows.

Definition 1 (Perfect correctness). *We say that a PSM protocol for f has perfect correctness if for every input $\mathbf{x} = (x_1, \ldots, x_k) \in \mathcal{X}$ and every common random string $r \in \mathcal{R}$, it holds that $P_0(\mathbf{P}(\mathbf{x}; r)) = f(\mathbf{x})$.*

Definition 2 (Perfect privacy and δ-statistical privacy). *We say that a PSM protocol for f has δ-statistical privacy if there exists a simulator* sim *for every input $\mathbf{x} = (x_1, \ldots, x_k) \in \mathcal{X}$ it holds that $\Delta(\mathrm{sim}(f(\mathbf{x})), \mathbf{P}(\mathbf{x})) \leq \delta$. In particular, it is said to have perfect privacy if it has 0-statistical privacy.*

Let $\mathcal{F}$ be a function family. If a PSM protocol works for every function $f \in \mathcal{F}$, we say that it is a PSM protocol for $\mathcal{F}$. In particular, if its referee is independent of $f \in \mathcal{F}$, we say that it has universal reconstruction.

2.2 Conditional Disclosure of Secrets

Let $\mathcal{X}_i$ and $\mathcal{M}_i$ denote finite domains of inputs and messages of the i-th party ($i \in [k]$). Also, let $\mathcal{X}, \mathcal{M}$ and $\mathcal{R}$ denote finite domains of the total inputs, total messages, and common random strings, respectively. Let $\mathcal{S}$ be a finite domain of a secret and let $f : \mathcal{X} \to \{0, 1\}$ be a predicate.

There are $k + 1$ parties $P_0, P_1, \ldots, P_k$ in a CDS protocol as well as a PSM protocol, where P_0 is the referee. For $i \in [k]$, P_i is connected only with P_0 by secure point-to-point channels. Each party P_i with $i \in [k]$ sends the referee P_0 a message $m_i \in \mathcal{M}_i$ that is computed on an input $x_i \in \mathcal{X}_i$, secret $s \in \mathcal{S}$ and common random string $r \in \mathcal{R}$. Eventually, P_0 learns a secret $s \in \mathcal{S}$ from inputs $x_1, \ldots, x_k$ and messages $m_1, \ldots, m_k$ without a common random string r if $f(x_1, \ldots, x_k) = 1$ for a given function $f : \mathcal{X} \to \{0, 1\}$; otherwise P_0 learns nothing for $s \in \mathcal{S}$.

A CDS protocol is required to satisfy correctness and privacy. The correctness means that the referee outputs a correct secret $s \in \mathcal{S}$ if $f(\mathbf{x}) = 1$, and the privacy means that some procedure can generate a distribution that is identical to (or statistically close to) the view of the referee only from the inputs $\mathbf{x}$, for which $f(\mathbf{x}) = 0$.

More formally, a CDS protocol for a predicate f is modeled as follows. For every $i \in [k]$, let P_i be an algorithm that gives a message $m_i = P_i(x_i, s; r) \in \mathcal{M}_i$ on an input $x_i \in \mathcal{X}_i$, secret $s \in \mathcal{S}$, and common random string $r \overset{\$}{\leftarrow} \mathcal{R}$ at the start of the protocol, where the secret s is common among all parties $P_1, \ldots, P_k$. We also define a message function $\mathbf{P}$ of the CDS protocol as $\mathbf{P}(\mathbf{x}, s; r) = (P_1(x_1, s; r), \ldots, P_k(x_k, s; r))$. Let P_0 be an algorithm that receives a sequence of messages $\mathbf{P}(\mathbf{x}, s; r) = (P_1(x_1, s; r), \ldots, P_k(x_k, s; r))$ from $P_1, \ldots, P_k$. Eventually, P_0 outputs $s' \in \mathcal{S}$ on sequences of inputs $\mathbf{x}$ and messages $\mathbf{P}(\mathbf{x}, s; r)$.

The correctness and privacy are defined as follows.

Definition 3 (Perfect correctness). *We say that a CDS protocol for f has perfect correctness if for every input $\mathbf{x} = (x_1, \ldots, x_k) \in \mathcal{X}$, every secret $s \in \mathcal{S}$, and every common random string $r \in \mathcal{R}$, if $f(\mathbf{x}) = 1$, it holds that*

$$P_0(\mathbf{x}, P_1(x_1, s; r), \ldots, P_k(x_k, s; r)) = s.$$

Definition 4 (Perfect privacy and δ-statistical privacy). *We say that a CDS protocol for f has δ-statistical privacy if there exists a simulator* sim *for every input $\mathbf{x} = (x_1, \ldots, x_k) \in \mathcal{X}$ for which $f(\mathbf{x}) = 0$, and every secret $s \in \mathcal{S}$*

$$\Delta(\mathrm{sim}(\mathbf{x}), \mathbf{P}(\mathbf{x}, s)) \leq \delta,$$

where $\mathbf{P}(\mathbf{x}, s)$ denotes a random run of $\mathbf{P}(\mathbf{x}, s)$ with uniformly random $r \in \mathcal{R}$. In particular, it is said to have perfect privacy if it has 0-statistical privacy.

Let $\mathcal{F}$ be a predicate family. If a CDS protocol works for every predicate $f \in \mathcal{F}$, we say that it is a CDS protocol for $\mathcal{F}$. In particular, if its referee is independent of $f \in \mathcal{F}$, we say that it has universal reconstruction.

2.3 Information Theory

For a random variable X, let $H(X)$ be the Shannon entropy. That is, $H(X) = -\sum_{x \in \mathcal{X}} \Pr(X = x) \log \Pr(X = x)$, where $\mathcal{X}$ is the sample space of X. For random variables X_1 and X_2, let $H(X_1X_2)$, $H(X_1|X_2)$, and $I(X_1;X_2)$ be the joint entropy, the conditional entropy, and the mutual information, respectively.

From Fano's inequality (Theorem 2.10.1 in [15]), the following corollary follows.

Corollary 1 (Corollary 1 in [18]). *Assume $\Delta((X,Y),(X',Y')) \leq \delta$, and let $\mathcal{X}\mathcal{Y}$ be the support set of X, Y. Then, we have $|H(X|Y) - H(X'|Y')| \leq 2\delta(\lg(|\mathcal{X}||\mathcal{Y}|) - \lg\delta)$.*

3 Randomness Bounds for PSM with Perfect Privacy

In this section, we show randomness lower bounds of k-party PSM protocols with the perfect privacy. In order to show the randomness lower bounds, we first provide general connections from communication lower bound to randomness lower bound by entropic and combinatorial arguments. Then, we obtain a randomness lower bound of PSM protocols for general functions from the connections with a communication lower bound. Furthermore, we prove randomness lower bounds for a generalized inner product function by an entropic argument.

For k-party PSM protocols, a party P_i is defined as a map $P_i : \mathcal{X}_i \times \mathcal{R} \to \mathcal{M}_i$ for every $i \in [k]$, and a referee P_0 is defined as $P_0 : \prod_{i=1}^{k} \mathcal{M}_i \to \{0,1\}$ in this section.

3.1 Connections from Communication to Randomness in PSM

We first show a simple connection from an entropic viewpoint. Let $\mathbf{X} = (X_1, \ldots, X_k)$ denote the random variables describing the inputs $\mathbf{x} = (x_1, \ldots, x_k)$. Let R be the random variables describing the common random string r. Let $\mathbf{M} = (M_1, \ldots, M_k)$ be the random variables describing $\mathbf{m} = (m_1, \ldots, m_k)$ after a run on $\mathbf{X}$ and R.

Theorem 8. *For every function f and every PSM protocol for f, it holds that $H(R) \geq H(\mathbf{M}) - H(\mathbf{X})$.*

Proof. Recall $M_i(x_i) = P_i(x_i; R)$. From the data processing inequality, we have $H(\mathbf{X}R) \geq H(\mathbf{M})$. Because $\mathbf{X}$ and R are mutually independent, it holds $H(R) \geq H(\mathbf{M}) - H(\mathbf{X})$. Note that this inequality depends on neither the correctness requirement nor privacy requirement. Thus, the statement of this theorem holds for both perfect and statistical cases. □

From the above theorem, we immediately obtain the following corollary from the source coding theorem ($\rho \geq H(R)$ and $\lambda \geq H(\mathbf{M}) > \lambda - 1$), which provides a weak connection between lower bounds of communication and randomness in PSM protocols.

Corollary 2 (Theorem 1 in Sect. 1). *Let λ be the communication complexity for a function $f : (\{0,1\}^n)^k \to \{0,1\}$ of a k-party PSM protocol. Then, the randomness complexity ρ for f of the k-party PSM protocol is larger than $\lambda - nk - 1$.*

We next provide a better lower bound of the randomness domain size from the message one by a combinatorial argument. Let $\mathcal{X}_{[b]} := \{\mathbf{x} \in \mathcal{X} : f(\mathbf{x}) = b\}$ and let $\mathcal{M}_{[b]} := \{\mathbf{P}(\mathbf{x}; r) \in \mathcal{M} : \mathbf{x} \in \mathcal{X}_{[b]}, r \in \mathcal{R}\}$ for $b \in \{0,1\}$. We also define $\mathcal{M}_{[b]}(r) := \{\mathbf{P}(\mathbf{x}; r) : \mathbf{x} \in \mathcal{X}_{[b]}\} \subseteq \mathcal{M}_{[b]}$.

Theorem 9. *If a k-party PSM protocol has perfect correctness and perfect privacy, it holds that $|\mathcal{R}| \geq \max_{b\in\{0,1\}}|\mathcal{M}_{[b]}|$.*

Proof. Fix any $b \in \{0,1\}$. If a PSM protocol has perfect correctness and perfect privacy, there exists a distribution D_b such that for every $\mathbf{x} \in \mathcal{X}_{[b]}$ we have $\mathbf{P}(\mathbf{x}) \equiv D_b$ and $\mathrm{Supp}(D_b) = \mathcal{M}_{[b]}$. (Note that we have $\mathrm{Supp}(D_b) = \mathrm{Supp}(\mathbf{P}(\mathbf{x})) = \cup_{\mathbf{x}' \in \mathcal{X}_{[b]}} \mathrm{Supp}(\mathbf{P}(\mathbf{x}')) = \mathcal{M}_{[b]}$ for every $\mathbf{x} \in \mathcal{X}_{[b]}$ from the perfect privacy.) It obviously holds that $|\mathrm{Supp}(\mathbf{P}(\mathbf{x}))| \leq |\mathcal{R}|$. Therefore, we obtain $|\mathcal{R}| \geq |\mathrm{Supp}(\mathbf{P}(\mathbf{x}))| = |\mathrm{Supp}(D_b)| = |\mathcal{M}_{[b]}|$. This lower bound holds for any $b \in \{0,1\}$, and thus, we obtain $|\mathcal{R}| \geq \max_{b\in\{0,1\}}|\mathcal{M}_{[b]}|$. □

From the above theorem, we immediately obtain the following corollary, which provides the connection from communication lower bounds to randomness lower bounds in PSM protocols.

Corollary 3 (Perfect-privacy part of Theorem 2 in Sect. 1). *Let λ be the communication complexity for a function $f : [N]^k \to \{0,1\}$ of a k-party PSM protocol that has the perfect correctness and perfect privacy. Then, the randomness complexity ρ for f of the k-party PSM protocol is at least $\lambda - 1$.*

3.2 Lower Bounds for General Functions in PSM

We next apply Theorem 9 to the randomness optimality of the PSM protocol provided by [21] with universal reconstruction for a general function. We consider k-party PSM protocols with universal reconstruction for a family of functions $\mathcal{F} = \{f : \prod_{i=1}^{k} \mathcal{X}_i \to \{0,1\}\}$. Note that P_0 must be independent of $f \in \mathcal{F}$ for the universal reconstruction, but $P_1, \ldots, P_k$ may depend on f.

For simplicity, we assume $\mathcal{X}_i = [N]$ for every i. For a party P_i, let $P_i(x)$ denote a random run of P_i on any fixed $x \in [N]$. Similarly, let $\mathbf{P}(\mathbf{x})$ denote a random run of $P_1, \ldots, P_k$ on $\mathbf{x} \in [N]^k$. Let $\mathrm{Supp}(\mathbf{P}(\mathbf{x})) := \left\{\mathbf{m} \in \prod_{i=1}^{k} \mathcal{M}_i : \Pr[\,\mathbf{P}(\mathbf{x}) = \mathbf{m}\,] > 0\right\}$. We define $\mathcal{M}_i(r) := \mathrm{Im}(P_i(\,\cdot\,; r))$ and $\mathcal{M}(r) := \prod_{i=1}^{k} \mathcal{M}_i(r)$.

In order to prove the randomness lower bound from Theorem 9, we prove communication lower bounds of k-party PSM protocols with universal reconstruction for general functions, which is a simple generalization of the communication lower bounds for the 2-party version in [12].

Lemma 2. *If a k-party PSM protocol with universal reconstruction for $\mathcal{F}$ has perfect correctness, it holds that $|\mathcal{M}(r)| \geq 2^{N^{k-1}}$ for every $r \in \mathcal{R}$.*

Proof. Fix any $r \in \mathcal{R}$. In the execution of the PSM for $f \in \mathcal{F}$ with fixed r, the k parties individually generate $P_1^f(x_1; r) \in \mathcal{M}_1(r), \ldots, P_k^f(x_k; r) \in \mathcal{M}_k(r)$ on input $\mathbf{x} = (x_1, \ldots, x_k)$, where the superscript f indicates that every P_i may depend on $f \in \mathcal{F}$. For each $f \in \mathcal{F}$, the tuple $\mathcal{S}^f(r) = (P_1^f(x_1; r))_{x_1 \in [N]} \circ \cdots \circ (P_k^f(x_k; r))_{x_k \in [N]}$ determines all values $(f(\mathbf{x}))_{\mathbf{x} \in [N]^k}$, i.e., $f \in \mathcal{F}$, from the perfect correctness and universal reconstruction, where the operator $\circ$ denotes concatenation of tuples. Then, it must hold that $\prod_{i=1}^{k} |\mathcal{M}_i(r)|^N = |\mathcal{M}(r)|^N \geq 2^{N^k}$, and thus, $|\mathcal{M}(r)| \geq 2^{N^{k-1}}$. □

By combining Theorem 9 with Lemma 2, we can obtain the randomness lower bounds for the PSM protocol with universal reconstruction.

Theorem 10 (Perfect-privacy part of Theorem 3 in Sect. 1). *If a k-party PSM protocol $\Pi^{\mathcal{F}}$ with universal reconstruction has perfect correctness and perfect privacy, it holds that $|\mathcal{R}| \geq 2^{N^{k-1}}/2$. Furthermore, if its message function $\mathbf{P}$ is surjective, that is, $\mathcal{M} = \cup_{f \in \mathcal{F}, x \in \mathcal{X}, r \in \mathcal{R}} \mathbf{P}^f(x; r)$, it holds that $|\mathcal{R}| \geq |\mathcal{M}|/2$.*

Proof. We can identify the PSM protocol with universal reconstruction as the one whose message function $\mathbf{P}$ takes the truth table of f and the private inputs $(x_1, \ldots, x_k)$ of k parties as an input. From Theorem 9, it holds $|\mathcal{R}| \geq |\mathcal{M}_{[b]}|$ for every $b \in \{0, 1\}$. Since $\max_{b \in \{0,1\}} |\mathcal{M}_b(r)| \geq |\mathcal{M}(r)|/2$ and $|\mathcal{M}(r)| \geq 2^{N^{k-1}}$ by Lemma 2, we obtain $|\mathcal{R}| \geq 2^{N^{k-1}}/2$.

Now, we suppose that $\mathbf{P}$ is surjective. Then, $\{\mathcal{M}_{[b]}\}_{b \in \{0,1\}}$ is a partition of $\mathcal{M}$ and it holds $|\mathcal{R}| \geq \max_{b \in \{0,1\}} |\mathcal{M}_{[b]}| \geq |\mathcal{M}|/2$ by Theorem 9. Thus, we obtain $|\mathcal{R}| \geq |\mathcal{M}|/2$. □

Theorem 10 shows the optimality of the famous FKN protocol [21] for general functions with respect to the randomness complexity, which is a 2-party PSM protocol for $\mathcal{F}$ with universal reconstruction that has perfect correctness and perfect privacy. Its message function is defined as follows. Fix $N := 2^n$ and $\mathcal{F} := \{f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}\}$. Let $(r, s) \in \{0,1\}^N \times \{0,1\}^n$ be common random string between two parties P_1 and P_2. For their private inputs $x_1, x_2 \in \{0,1\}^n$ and a function $f \in \mathcal{F}$, let $P_1^f(x_1; r, s) := (f(x_1, i + s))_{i \in [N]}$, and $P_2^f(x_2; r, s) := (x_2 - s, r_{x_2})$, where the operators $+$ and $-$ are under modulo N. The message space $\mathcal{M}$ of this protocol is $\{0,1\}^N \times \{0,1\}^n \times \{0,1\}$ and randomness space $\mathcal{R}$ is $\{0,1\}^N \times \{0,1\}^n$. It is easy to see that the message function $\mathbf{P}^f(\mathbf{x}; r, s) = (P_1^f(x_1; r, s), P_2^f(x_2; r, s))$ is surjective on the message space $\mathcal{M}$.

By Theorem 10, any PSM protocol with universal reconstruction that has perfect correctness and perfect privacy requires the randomness space of size at least $2^N \cdot 2^n$ if the range of its message function is $\{0,1\}^N \times \{0,1\}^n \times \{0,1\}$, which is achieved by the FKN protocol. So, the randomness complexity of their protocol is optimal for the message space $\{0,1\}^N \times \{0,1\}^n \times \{0,1\}$. There would be a room for improvement of the communication complexity and randomness complexity since the best lower bound of the communication complexity given in [12] (see also Lemma 2) is 2^N. Even if the upper bound of the communication complexity could be improved to 2^N, the randomness complexity is at least $2^N/2$, as shown in Theorem 10.

3.3 Lower Bounds for Explicit Functions in PSM

Next, we focus on randomness and communication lower bounds of PSM protocols for a generalized inner product function. For simplicity, we also assume $\mathcal{X}_i = [N]$ for every i and $N = 2^n$ for some n. Let $\oplus$ denote the XOR operation. For even number k, define f_{IP} by $f_{\mathrm{IP}}(\mathbf{x}) = \langle x_1, x_2 \rangle \oplus \cdots \oplus \langle x_{k-1}, x_k \rangle$, where $\langle \cdot, \cdot \rangle$ is the inner product modulo 2, i.e., $\langle x, y \rangle = \oplus_{j \in [n]} x[j] \cdot y[j]$ for $x = (x[1], \ldots, x[n]), y = (y[1], \ldots, y[n]) \in \{0,1\}^n$. For odd number k, define f_{IP} by $f_{\mathrm{IP}}(\mathbf{x}) = \langle x_1, x_2 \rangle \oplus \cdots \oplus \langle x_{k-2}, x_{k-1} \rangle \oplus \langle x_k, x_1 \rangle$, i.e., the inner product is applied to x_1 twice.

We now prove lower bounds for k-party PSM protocols for f_{IP} as follows:

Theorem 11 (Perfect-privacy part of Theorem 4 in Sect. 1). *If a PSM protocol for f_{IP} has perfect correctness and perfect privacy, it holds that $H(R) \geq H(\mathbf{X} \mid f_{\mathrm{IP}}(\mathbf{X})))$ and $H(\mathbf{M}) \geq H(\mathbf{X})$. For uniformly random inputs, it holds that $H(R) > kn - 2$ and $H(M) \geq kn$.*

Proof. We prove some basic facts about PSM from the perfect correctness.

Claim. If a PSM protocol for f_{IP} has perfect correctness, then there exists a function f_i for every $i \in [k]$ such that $x_i = f_i(m_i, r)$ for every $x_i \in \mathcal{X}_i = \{0,1\}^n$ and every $r \in \mathcal{R}$. Then, $\mathbf{x} = (f_1(m_1, r), \ldots, f_k(m_k, r))$.

Proof. For simplicity, consider the case of $i = 1$. f_1 can be implemented by computing $x_1 = (x_1[1], \ldots, x_1[n])$ for given m_1 and r as follows. For $j \in [n]$, set $x_2[j] = 1$, $x_2[j'] = 0$ for $j' \neq j$, and $x_i = (0, \ldots, 0)$ for $i \in [k] \setminus \{1, 2\}$, and then execute $P_0(m_1, P_2(x_2; r_2), \ldots, P_k(x_k; r_k))$. From perfect correctness, the output is $x_1[j]$. For a general $i \in [k]$, x_2 is replaced by its paired input value and the other inputs are set to $(0, \ldots, 0)$. In particular, for f_k with an odd number k, $x_k = (x_k[1], \ldots, x_k[n])$ is obtained for given m_k and r as follows. For $j \in [n]$, set $x_1[j] = 1$, $x_1[j'] = 0$ for $j' \neq j$, and $x_2 = \cdots = x_{k-1} = (0, \ldots, 0)$, and then execute $P_0(P_1(x_1; r_1), \ldots, P_{k-1}(x_{k-1}; r_{k-1}), m_k)$. □

Then, we have

$$
\begin{aligned}
H(\mathbf{M}) &\geq H(\mathbf{M} \mid R) \\
&= H(\mathbf{M}, \mathbf{X} \mid R) - H(\mathbf{X} \mid \mathbf{M}, R) \\
&\geq H(\mathbf{X} \mid R) - H(\mathbf{X} \mid \mathbf{M}, R) && \text{(Independence of } \mathbf{X} \text{ and R)} \\
&= H(\mathbf{X}). && \text{(Claim 3.3)}
\end{aligned}
$$

In a similar manner, we also have

$$
\begin{aligned}
H(R) &\geq H(R \mid \mathbf{M}) \\
&= H(R\mathbf{X} \mid \mathbf{M}) - H(\mathbf{X} \mid R\mathbf{M}) && \text{(Chain rule)} \\
&\geq H(\mathbf{X} \mid \mathbf{M}) - H(\mathbf{X} \mid R\mathbf{M}) && \text{(Independence of } \mathbf{X} \text{ and } R) \\
&= H(\mathbf{X} \mid \mathbf{M}) && \text{(Claim 3.3)} \\
&= H(\mathbf{X} \mid \mathrm{sim}(f_{\mathrm{IP}}(\mathbf{X}))) && \text{(Perfect privacy)} \\
&\geq H(\mathbf{X} \mid f_{\mathrm{IP}}(\mathbf{X})). && \text{(Data processing inequality)}
\end{aligned}
$$

There are 2^{kn} possible $\mathbf{x}$'s. We know that if just one of every paired value of $\mathbf{x}$ is $(0,\ldots,0)$, then $f_{\mathrm{IP}}(\mathbf{x}) = 0$: otherwise, $f_{\mathrm{IP}}(\mathbf{x}) = 0$ and $f_{\mathrm{IP}}(\mathbf{x}) = 1$ are equally likely. Thus, the number of $\mathbf{x}$ with $f_{\mathrm{IP}}(\mathbf{x}) = 1$ is slightly smaller than that with $f_{\mathrm{IP}}(\mathbf{x}) = 0$. In other words, $f_{\mathrm{IP}}(\mathbf{x}) = 1$ leaks slightly more information on $\mathbf{x}$. Let y and Y be the output $f_{\mathrm{IP}}(\mathbf{x})$ and the random variable describing $f_{\mathrm{IP}}(\mathbf{x})$, respectively. Let $n' = kn/2$. For uniformly random inputs $\mathbf{x}$, it follows that

$$\begin{aligned}
H(\mathbf{X} \mid f_{\mathrm{IP}}(\mathbf{X})) &\geq - \sum_{\mathbf{x}\in\mathcal{X}, y\in\{0,1\}} \Pr(\mathbf{x}, y) \log \Pr(\mathbf{x} \mid y) \\
&= - \sum_{\mathbf{x}: f_{\mathrm{IP}}(\mathbf{x})=0} \Pr(\mathbf{x}, 0) \log \Pr(\mathbf{x} \mid 0) \sum_{\mathbf{x}: f_{\mathrm{IP}}(\mathbf{x})=1} \Pr(\mathbf{x}, 1) \log \Pr(\mathbf{x} \mid 1) \\
&= 2^{-n'-1} \cdot (2^{n'} + 1) \log 2^{n'-1} \cdot (2^{n'} + 1) \\
&> n' - 1 + \log(2^{n'} - 1) > kn - 2.
\end{aligned}$$

Thus, the statement of the theorem holds. □

We can extend the above randomness lower bounds for a PSM protocol with the perfect privacy to one with the statistical privacy. See Theorem 17, Corollary 4, and Theorem 19 in Sect. B.

4 Randomness Bounds for CDS with Perfect Privacy

In this section, we consider CDS protocols in which a party $P_i : \mathcal{X}_i \times \mathcal{S} \times \mathcal{R} \to \mathcal{M}_i$ for every $i \in [k]$ and a referee $P_0 : \mathcal{X} \times \mathcal{M} \to \mathcal{S}$, where $\mathcal{X} = \mathcal{X}_1 \times \cdots \mathcal{X}_k$ and $\mathcal{M} \subseteq \mathcal{M}_1 \times \cdots \times \mathcal{M}_k$.

Let S denote the random variable describing the secret $s \in \mathcal{S}$. Let R be the random variable describing the common random string $r \in \mathcal{R}$. For $x_i \in \mathcal{X}_i$, let $M_i(x_i)$ be the random variable describing m_i after a run on x_i, S, and R. That is, $M_i(x_i) = P_i(x_i, S; R)$. For $\mathbf{x} = (x_1, \ldots, x_k)$, let $\mathbf{M}(\mathbf{x})$ be the random variable describing $\mathbf{m}$ after a run on $\mathbf{x}$, S, and R.

4.1 Connections from Communication to Randomness in CDS

By using the following lemma, we can show a randomness lower bound from the communication complexity and secret length by an entropic argument, and thus, it is useful to prove randomness lower bounds directly from the communication complexity.

Lemma 3 (Lemma 1 in Sect. 1). *For every function f and every CDS protocol for f, it holds that $H(R) \geq H(\mathbf{M}(\mathbf{x})) - H(S)$.*

Proof. Recall $M_i(x_i) = P_i(x_i, S; R)$. From the data processing inequality, we have $H(SR) \geq H(\mathbf{M}(\mathbf{x}))$. Because S and R are mutually independent, the statement of the lemma holds. □

Recall that we can provide a better connection from communication to randomness for randomness lower bounds in PSM protocols by a combinatorial argument. (See Theorem 3 and Corollary 3.) In contrast to PSM protocols, the following CDS protocol with no common random string shows that combinatorial arguments cannot generally improve the connection obtained from the entropic argument of Lemma 3. Let $\vee$ denote the bit-wise OR operation. Let $f_{\mathrm{or}}(\mathbf{x}) = \bigvee_{i=1}^{k} \bigvee_{j=1}^{n_i} x_i[j]$ where $x_i \in \{0,1\}^{n_i}$ for $i \in [k]$.

Theorem 12. *An randomness-optimal CDS protocol for* f_{or} *with perfect correctness and perfect privacy satisfies* $\mathcal{M} = \mathcal{S}^k$ *and* $\mathcal{R} = \emptyset$.

Proof. The optimal protocol is given by as follows:

- $P_i(x_i, s; r)$: If $x_1[1] \vee \cdots \vee x_i[n_i] = 1$, then set $m_i = s$, and otherwise set m_1 to the null string.
- $P_0(m_1, \ldots, m_k)$: If some m_i is not null, then output m_i, and otherwise output a special symbol $\perp$.

The perfect correctness is obvious. Thus, we show the perfect privacy. If $f_{\mathrm{or}}(\mathbf{x}) = 0$, then all $x_i = (0, \ldots, 0)$. Then, all messages are null strings. Thus, $\mathrm{sim}(\mathbf{x})$ is defined by outputting null strings. $\square$

4.2 Lower Bounds for General Functions in CDS

We next consider a k-party CDS protocol with universal reconstruction for a general function. In the following theorem, we obtain a communication lower bound under restriction of possible reconstruction functions, and thus, we also obtain the corresponding randomness lower bound from Lemma 3.

Theorem 13 (Perfect-privacy part of Theorem 5 in Sect. 1). *Let* $\mathcal{F} := \{f : \mathcal{X} \to \{0,1\}\}$ *and let* $\mathcal{C}$ *be a set of possible referees in CDS protocols. Suppose* $|\mathcal{C}| \geq |\mathcal{S}|/2$. *If a* k*-party CDS protocol with universal reconstruction for* $\mathcal{F}$ *has perfect correctness and perfect privacy, it holds that* $\lambda \geq |\mathcal{X}| \cdot \sigma / \log|\mathcal{C}|$.

Proof. We consider an encoding procedure from any given $f \in \mathcal{F}$ by using the CDS protocol. Given $f \in \mathcal{F}$, sample $s_i \xleftarrow{\$} \mathcal{S}$ and $r_i \xleftarrow{\$} \mathcal{R}$ for each $i \in [T]$, where T is specified later. From the perfect correctness, it holds that for every $\mathbf{x}$ for which $f(\mathbf{x}) = 1$, we have $\Pr_{\{s_i, r_i\}_{i\in[T]}}[\, P_0(\mathbf{x}, (\mathbf{P}(\mathbf{x}, s_i; r_i))) = s_i \,] = 1.$ for every $i \in [T]$. From the perfect privacy and the following claim, it holds that for every P_0^* and every $\mathbf{x}$ for which $f(\mathbf{x}) = 0$, we have $\Pr_{s_i, r_i}[\, P_0^*(\mathbf{x}, (\mathbf{P}(\mathbf{x}, s_i; r_i))) = s_i \,] = 1/|\mathcal{S}|$ for every $i \in [T]$.

Claim. If a CDS protocol has perfect privacy, for every referee P_0^* and every $\mathbf{x} \in \mathcal{X}$ for which $f(\mathbf{x}) = 0$ we have $\sum_{s\in\mathcal{S}} \Pr_r[\, P_0^*(\mathbf{x}, \mathbf{P}(\mathbf{x}, s; r)) = s \,] = 1$.

Proof. From the perfect privacy, there exists a simulator sim such that for every $s \in \mathcal{S}$ and every $\mathbf{x}$ for which $f(\mathbf{x}) = 0$ we have $\Pr_{\mathrm{sim}}[\, P_0^*(\mathbf{x}, \mathrm{sim}(\mathbf{x})) = s \,] =$

$\Pr_r[P_0^*(\mathbf{x}, \mathbf{P}(\mathbf{x}, s; r)) = s]$. Therefore, it holds that $\sum_{s\in\mathcal{S}} \Pr_r[P_0^*(\mathbf{x}, \mathbf{P}(\mathbf{x}, s; r)) = s] = \sum_{s\in\mathcal{S}} \Pr_{\mathrm{sim}}[P_0^*(\mathbf{x}, \mathrm{sim}(\mathbf{x})) = s] = 1$. □

Therefore, we can see that

$$\Pr_{\{s_i,r_i\}_{i\in[T]}}[\forall i \in [T],\ P_0(\mathbf{x}, (\mathbf{P}(\mathbf{x}, s_i; r_i))) = s_i] = \begin{cases} 1 & \text{if } f(\mathbf{x}) = 1; \\ |\mathcal{S}|^{-T} & \text{otherwise.} \end{cases} \tag{1}$$

Let $T := \lceil \frac{\log|\mathcal{C}|+1}{\log|\mathcal{S}|} \rceil$. Note that $T \geq 1$ since $|\mathcal{S}| \leq 2|\mathcal{C}|$. By the union bound, for every $\mathbf{x}$ for which $f(\mathbf{x}) = 0$, from Eq. (1), it holds that

$$\begin{aligned} &\Pr_{\{s_i,r_i\}_{i\in[T]}}[\exists P_0^*, \forall i \in [T],\ P_0^*(\mathbf{x}, (\mathbf{P}(\mathbf{x}, s_i; r_i))) = s_i] \\ &\leq \sum_{P_0^*} \Pr_{\{s_i,r_i\}_{i\in[T]}}[\forall i \in [T],\ P_0^*(\mathbf{x}, (\mathbf{P}(\mathbf{x}, s_i; r_i))) = s_i] \leq |\mathcal{C}| \cdot |\mathcal{S}|^{-T} \leq 1/2. \end{aligned}$$

Thus, there exists a non-empty set $\mathcal{T} \subset (\mathcal{S} \times \mathcal{R})^T$ (precisely, of size at least $(1/2)|(\mathcal{S} \times \mathcal{R})^T|$) such that for every $(s_i, r_i)_{i\in[T]} \in \mathcal{T}$ it holds that

$$\forall i \in [T],\ P_0(\mathbf{x}, \mathbf{P}(\mathbf{x}, s_i; r_i)) = s_i, \quad \text{and} \quad \forall P_0^*, \exists i \in [T],\ P_0^*(\mathbf{x}, \mathbf{P}(\mathbf{x}, s_i; r_i)) \neq s_i. \tag{2}$$

Therefore, there exists a $(\hat{s}_i, \hat{r}_i)_{i\in[T]}$ that satisfies Eq. (2). Then, a sequence $(\mathbf{P}(\mathbf{x}, \hat{s}_i; \hat{r}_i), \hat{s}_i)_{i\in[T]}$ provides an encoding of f since we can determine if $f(\mathbf{x}) = 1$ on any given $\mathbf{x} \in \mathcal{X}$ by checking $P_0(\mathbf{x}, \mathbf{P}(\mathbf{x}, \hat{s}_i; \hat{r}_i)) = \hat{s}_i$ for every $i \in [T]$. The description length $|f|$ of this encoding of f is at most $O(T\cdot\lambda + T\cdot\log|\mathcal{S}|)$, and $|f|$ should be at least $\log|\mathcal{F}| = |\mathcal{X}|$. Therefore, we have $\lambda = \Omega(|\mathcal{X}| \cdot \log|\mathcal{S}|/\log|\mathcal{C}|)$. □

By extending the above argument, we can obtain a communication lower bound for a δ-statistical privacy version. See Theorem 21 in Section C.

4.3 Lower Bounds for Explicit Functions in CDS

In this section, we further discuss communication and randomness lower bounds of CDS protocols with perfect privacy for several explicit functions.

Let $\oplus$ and $\wedge$ denote the bit-wise XOR and AND operations, respectively. Let $f_{\mathrm{xor}}(\mathbf{x}) = \bigoplus_{i=1}^k \bigoplus_{j=1}^{n_i} x_i[j]$. Let $f_{\mathrm{and}}(\mathbf{x}) = \bigwedge_{i=1}^k \bigwedge_{j=1}^{n_i} x_i[j]$. For $i \in [k]$, we define $\mathcal{X}_{<i} := \mathcal{X}_1 \times \cdots \times \mathcal{X}_{i-1}$, $\mathcal{X}_{>i} := \mathcal{X}_{i+1} \times \cdots \times \mathcal{X}_k$, and $\mathcal{X}_{-i} := \mathcal{X}_{<i} \times \mathcal{X}_{>i}$. We then consider $\mathbf{x}_{<i}$, $\mathbf{x}_{>i}$, and $\mathbf{x}_{-i}$ as elements in $\mathcal{X}_{<i}$, $\mathcal{X}_{>i}$, and $\mathcal{X}_{-i}$, respectively. We say that f is *nontrivial* if there exists an input pair $(\mathbf{x}^{(0)}, \mathbf{x}^{(1)})$ and index i such that $f(\mathbf{x}^{(0)}) = 0, f(\mathbf{x}^{(1)}) = 1$ and $\mathbf{x}_{-i}^{(0)} = \mathbf{x}_{-i}^{(1)}$. Then, we call $(\mathbf{x}_0, \mathbf{x}_1)$ a *nontrivial* input pair on index i for f.

We first prove communication and randomness lower bounds of CDS protocols for three functions, $f_{\mathrm{xor}}, f_{\mathrm{and}}$, and f_{IP} by entropic arguments. (See Sect. 3.3 for the definition of f_{IP}.)

Theorem 14 (Perfect-privacy part of Theorem 6 in Sect. 1). *If a CDS protocol for* $f : (\{0,1\}^n)^k \to \{0,1\} \in \{f_{\mathrm{xor}}, f_{\mathrm{and}}, f_{\mathrm{IP}}\}$ *has perfect correctness and perfect privacy, then it holds that* $\lambda \geq k \cdot H(S)$, *and* $\rho \geq (k-1) \cdot H(S)$. *In particular, for uniformly random secrets, it holds that* $\lambda \geq k \cdot \log|\mathcal{S}|$ *and* $\rho \geq (k-1) \cdot \log|\mathcal{S}|$.

Proof. To derive lower bounds, we prove the following claim.

Claim. Let $f \in \{f_{\mathrm{xor}}, f_{\mathrm{and}}, f_{\mathrm{IP}}\}$. There is $\mathbf{x} \in (\{0,1\}^n)^k$ such that $f(\mathbf{x}) = 1$ and for every $i \in [k]$, $\mathbf{x}$ is one element of a nontrivial input pair on i for f.

Proof. For $f \in \{f_{\mathrm{xor}}, f_{\mathrm{and}}\}$, choose an arbitrarily $\mathbf{x}$ with $f(\mathbf{x}) = 1$. For $f = f_{\mathrm{IP}}$, set $\mathbf{x} = ((1,0,\ldots,0),\ldots,(1,0\ldots,0))$ if $\lceil k/2 \rceil$ is odd, and otherwise, $\mathbf{x} = ((1,0,\ldots,0),\ldots,(1,0,\ldots,0),(0,0\ldots,0))$. We define $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(0)}$ by $\mathbf{x}^{(1)} = \mathbf{x}$, $\mathbf{x}^{(0)}_{-i} = \mathbf{x}_{-i}$, $x^{(0)}_i = x_i \oplus (1,0,\ldots,0)$. Then, $f(\mathbf{x}^{(1)}) = f(\mathbf{x}) = 1$, $f_{\mathrm{xor}}(\mathbf{x}^{(0)}) = f_{\mathrm{xor}}(\mathbf{x}) \oplus 1 = 0$, $f_{\mathrm{and}}(\mathbf{x}^{(0)}) = f_{\mathrm{and}}(\mathbf{x}) \wedge (1 \oplus 1) = 0$, and $f_{\mathrm{IP}}(\mathbf{x}^{(0)}) = f_{\mathrm{IP}}(\mathbf{x}) \oplus 1 \cdot 1 = 0$. Then, the claim holds. □

We use the above $\mathbf{x}^{(1)} = \mathbf{x}$ and $\mathbf{x}^{(0)}$. Let $M^{(1)}_i$ and $M^{(0)}_i$ denote $M_i(x^{(1)}_i) = P_i(x^{(1)}_i; R)$ and $M_i(x^{(0)}_i) = P_i(x^{(0)}_i, R)$, respectively. Let $\mathbf{M}^{(1)}_{i:j}$ and $\mathbf{M}^{(0)}_{i:j}$ with $i \leq j$ denote $(M^{(1)}_i,\ldots,M^{(1)}_j)$ and $(M^{(0)}_i,\ldots,M^{(0)}_j)$, respectively. Then, from the chain rule, for any $1 \leq i \leq k$, we obtain

$$H(\mathbf{M}^{(1)}_{1:i}) = H(M^{(1)}_i | \mathbf{M}^{(1)}_{1:i-1}) + H(\mathbf{M}^{(1)}_{1:i-1}).$$

We have

$$\begin{aligned}
H(M^{(1)}_i \mid \mathbf{M}^{(1)}_{1:i-1}) &\geq H(M^{(1)}_i \mid \mathbf{M}^{(1)}_{1:i-1}\mathbf{M}^{(1)}_{i+1:k}) \\
&= H(M^{(1)}_i S \mid \mathbf{M}^{(1)}_{1:i-1}\mathbf{M}^{(1)}_{i+1:k}) \\
&\quad - H(S \mid M^{(1)}_i \mathbf{M}^{(1)}_{1:i-1}\mathbf{M}^{(1)}_{i+1:k}) && \text{(Chain rule)} \\
&\geq H(S \mid \mathbf{M}^{(1)}_{1:i-1}\mathbf{M}^{(1)}_{i+1:k}) - H(S \mid \mathbf{M}^{(1)}_{1:k}) \\
&= H(S \mid \mathbf{M}^{(1)}_{1:i-1}\mathbf{M}^{(1)}_{i+1:k}) && \text{(Perfect correctness)} \\
&\geq H(S \mid M^{(0)}_i \mathbf{M}^{(1)}_{1:i-1}\mathbf{M}^{(1)}_{i+1:k}) \\
&\geq H(S \mid \mathbf{M}(\mathbf{x}^{(0)})) && \text{(Relation of } (\mathbf{x}^{(1)}, \mathbf{x}^{(0)})) \\
&\geq H(S \mid \mathrm{sim}(\mathbf{x}^{(0)}))) && \text{(Perfect privacy)} \\
&\geq H(S). && \text{(Independence of } S)
\end{aligned}$$

By summing up for $1 \leq i \leq k$, we have $H(\mathbf{M}(\mathbf{x})) \geq k \cdot H(S)$. Here, it holds that $\lambda \geq H(\mathbf{M}(\mathbf{x})) \geq k \cdot H(S)$. From Lemma 3 and $\rho \geq H(R)$, it follows that $\rho \geq H(R) \geq (k-1) \cdot H(S)$.Thus, the theorem holds. □

For f_{and}, we can obtain the matching communication and randomness bounds from an optimal construction of a CDS protocol for f_{and}.

Theorem 15. *Let* $\mathcal{S} = \{0,1\}^\sigma$. *The communication complexity and randomness complexity of* k*-party CDS protocols for* f_{and} *with perfect correctness and perfect privacy is at most* $k\sigma$ *and* $(k-1)\sigma$, *respectively.*

Proof. To show the upper bounds, we construct a CDS protocol as follows:

- Setup: For $s \in \{0,1\}^\sigma$, choose $r_1, \ldots, r_{k-1} \in \{0,1\}^\sigma$ and set $r_k = \bigoplus_{1 \le i \le k-1} r_i$ where $\oplus$ is the element-wise XOR.
- $P_i(x_i, s; r)$: If $x_i[1] \wedge \cdots \wedge x_i[n_i] = 1$, then for $i = 1$, set $m_i = s \oplus r_i$ and for $i \neq i$, set $m_i = r_i$. Otherwise set m_i to the null string.
- $P_0(m_1, \ldots, m_k)$: If some m_i is the null string, then output a special symbol $\perp$, and otherwise output $m_1 \oplus \cdots \oplus m_k$.

The perfect correctness is obvious. Thus, we show the perfect privacy. If $f(\mathbf{x}) = 0$, some message is the null string and thus the remaining messages are mutually independent and random. $\mathrm{sim}(\mathbf{x})$ is defined by outputting random m_i for i with $x_i[1] \wedge \cdots x_i[n_i] = 1$ and the null string for i with $x_i[1] \wedge \cdots x_i[n_i] = 0$. $\square$

Next, we prove lower bounds of CDS protocols for some family of functions, named θ-nontrivial functions, which are defined below.

Definition 5. *A function f is θ-nontrivial if there exist input pairs $(\mathbf{x}^{(1):j}, \mathbf{x}^{(0):j})_{j \in [\theta]}$ and indices $(i_j)_{j \in [\theta]}$ such that for every $j \in [\theta]$, $f(\mathbf{x}^{(1):j}) = 1$, $f(\mathbf{x}^{(0):j}) = 0$, $\mathbf{x}^{(1):j}_{-i_j} = \mathbf{x}^{(0):j}_{-i_j}$, and $\mathbf{x}^{(1):j}_{I_{j-1}} = \mathbf{x}^{(1):j-1}_{I_{j-1}}$ with $I_{j-1} = \{i_1, \ldots, i_{j-1}\}$. We call $(\mathbf{x}^{(1):j}, \mathbf{x}^{(0):j})_{j \in [\theta]}$ nontrivial input pairs on $(i_j)_{j \in [\theta]}$ for f. We also call $(i_j)_{j \in [\theta]}$ a nontrivial index sequence for f.*

The famous three types of access functions appeared in secret sharing schemes are nontrivial functions.

Example 1 (θ-threshold, θ-uniform, and monotone access functions). The famous three types of access functions appeared in secret sharing schemes are nontrivial functions. A θ-threshold access function $f_{\mathrm{thr}} : \{0,1\}^k \to \{0,1\}$ is defined so that $f_{\mathrm{thr}}(\mathbf{x}) = 1$ if and only if the weight of $\mathbf{x}$ is not smaller than θ. Then, f_{thr} is θ-nontrivial. For simplicity, consider the 3-threshold access function for $k = 4$. Let

$$\begin{aligned}
&\mathbf{x}^{(1):1} = \mathbf{x}^{(1):2} = \mathbf{x}^{(1):3} = (1,1,1,0),\\
&\mathbf{x}^{(0):1} = (0,1,1,0), \mathbf{x}^{(1):2} = (1,0,1,0), \mathbf{x}^{(1):3} = (1,1,0,0),
\end{aligned}$$

where a nontrivial index sequence is $(i_j)_{j \in [3]} = (1,2,3)$. Then, for $j = 1$, it holds that

$$\begin{aligned}
f_{\mathrm{thm}}(\mathbf{x}^{(1):1}) &= f_{\mathrm{thm}}(1,1,1,0) = 1,\\
f_{\mathrm{thm}}(\mathbf{x}^{(0):1}) &= f_{\mathrm{thm}}(0,1,1,0) = 0,\\
\mathbf{x}^{(1):1}_{-1} &= \mathbf{x}^{(0):1}_{-1} = (1,1,0),
\end{aligned}$$

and $I_0 = \emptyset$. Similarly, for $j = 2$,

$$\begin{aligned}
&f_{\mathrm{thm}}(\mathbf{x}^{(1):2}) = 1,\\
&f_{\mathrm{thm}}(\mathbf{x}^{(0):2}) = f_{\mathrm{thm}}(1,0,1,0) = 0,\\
&\mathbf{x}^{(1):2}_{-2} = \mathbf{x}^{(0):2}_{-2} = (1,1,0),\ \mathbf{x}^{(1):2}_{I_1} = \mathbf{x}^{(1):1}_{I_1} = 1,
\end{aligned}$$

where $I_1 = \{1\}$. For $j = 3$,

$$\begin{aligned}
&f_{\text{thm}}(\mathbf{x}^{(1):3}) = 1,\\
&f_{\text{thm}}(\mathbf{x}^{(0):3}) = f_{\text{thm}}(1,1,0,0) = 0,\\
&\mathbf{x}_{-3}^{(1):3} = \mathbf{x}_{-3}^{(0):3} = (1,1,0),\ \mathbf{x}_{I_2}^{(1):3} = \mathbf{x}_{I_2}^{(1):2} = (1,1),
\end{aligned}$$

where $I_2 = \{1,2\}$. Thus, the 3-threshold access function is a 3-nontrivial function. For any $\theta \geq 1$ and any $k \geq \theta$, we can show a similar example of nontrivial input pairs. A θ-uniform access function $f_{\text{uni}} : \{0,1\}^k \to \{0,1\}$ is defined so that $f_{\text{uni}}(\mathbf{x}) = 1$ (resp. $f_{\text{uni}}(\mathbf{x}) = 0$) if the weight of $\mathbf{x}$ is larger than θ (resp. smaller than θ). Then, f_{uni} is at least θ-nontrivial because some $\mathbf{x}$ of weight θ with $f_{\text{uni}}(\mathbf{x}) = 1$ can be used as $\mathbf{x}^{(1):1} = \mathbf{x}^{(1):2} = \cdots = \mathbf{x}^{(1):\theta}$. In such a way, for a monotone access function f, a minimal weight input $\mathbf{x}$ with $f(\mathbf{x}) = 1$ can be used as one of the nontrivial input pairs. Thus, letting θ be the maximal weight of such inputs, the monotone access function is θ-nontrivial.

The lower bounds for nontrivial functions are given by the following theorem. The complexity depends on the parameter θ of nontriviality.

Theorem 16 (Perfect-privacy part of Theorem 7 in Sect. 1). *If a CDS protocol for a θ-nontrivial function f has perfect correctness and perfect privacy, then it holds that $\lambda \geq \theta H(S)$ and $\rho \geq (\theta - 1)H(S)$. In particular, for uniformly random secrets, it holds that $\lambda \geq \theta \log|\mathcal{S}|$ and $\rho \geq (\theta - 1)\log|\mathcal{S}|$.*

Proof. For a θ-nontrivial function f, we use nontrivial input pairs $(\mathbf{x}^{(1):j}, \mathbf{x}^{(0):j})_{j\in[\theta]}$ on $(i_j)_{j\in[\theta]}$. For each $j \in [\theta]$, Let $M_i^{(1):j}$ and $M_i^{(0):j}$ denote $M_i(x_i^{(1):j}) = P_i(x_i^{(1):j}, S; R)$ and $M_i(x_i^{(0):j}) = P_i(x_i^{(0):j}, S; R)$, respectively. For $I_j = \{i_1, \ldots, i_j\}$, let $\bar{I}_j = [k] \setminus I_j$ and $I_0 = \emptyset$.

Then, from the chain rule, for any $1 \leq j \leq \theta$, we obtain

$$H(\mathbf{M}_{\bar{I}_{j-1}}^{(1):j}) = H(M_{i_j}^{(1):j} | \mathbf{M}_{\bar{I}_j}^{(1):j}) + H(\mathbf{M}_{\bar{I}_j}^{(1):j}).$$

We have

$$\begin{aligned}
H(M_{i_j}^{(1):j} | \mathbf{M}_{\bar{I}_j}^{(1):j}) &\geq H(M_{i_j}^{(1):j} \mid \mathbf{M}_{\bar{I}_j}^{(1):j} \mathbf{M}_{I_{j-1}}^{(1):j}) \\
&= H(M_{i_j}^{(1):j} S \mid \mathbf{M}_{\bar{I}_j}^{(1):j} \mathbf{M}_{I_{j-1}}^{(1):j}) \\
&\quad - H(S \mid M_{i_j}^{(1):j} \mathbf{M}_{\bar{I}_j}^{(1):j} \mathbf{M}_{I_{j-1}}^{(1):j}) && \text{(Chain rule)} \\
&\geq H(S \mid \mathbf{M}_{\bar{I}_j}^{(1):j} \mathbf{M}_{I_{j-1}}^{(1):j}) \\
&\quad - H(S \mid \mathbf{M}^{(1):j}) \\
&= H(S \mid \mathbf{M}_{\bar{I}_j}^{(1):j} \mathbf{M}_{I_{j-1}}^{(1):j}) && \text{(Perfect correctness)} \\
&\geq H(S \mid M_{i_j}^{(0):j} \mathbf{M}_{\bar{I}_j}^{(1):j} \mathbf{M}_{I_{j-1}}^{(1):j}) \\
&\geq H(S \mid \mathbf{M}(\mathbf{x}^{(0):j})) && \text{(Relation between } \mathbf{x}^{(1):j} \text{ and } \mathbf{x}^{(0):j}) \\
&\geq H(S \mid \text{sim}(\mathbf{x}^{(0):j}))) && \text{(Perfect privacy)} \\
&\geq H(S). && \text{(Independence of } S)
\end{aligned}$$

By summing up for $1 \leq j \leq \theta$, we have $H(\mathbf{M}(\mathbf{x})) \geq \theta H(S)$. Here, it holds that $\lambda \geq H(\mathbf{M}(\mathbf{x})) \geq H(S)$. From Lemma 3 and $\rho \geq H(R)$, it follows that $\rho \geq H(R) \geq (\theta - 1)H(S)$. Thus, the theorem holds. $\square$

The statistical-privacy version of the above theorem will be provided in Theorem 23 of Sect. C.2.

References

1. Aiello, B., Ishai, Y., Reingold, O.: Priced oblivious transfer: how to sell digital goods. In Proc. EUROCRYPT 2001. pp. 118–134 (2001)
2. Applebaum, B.: Garbled circuits as randomized encodings of functions: a primer. Tutorials on the Foundations of Cryptography, pp.1–44. (2017)
3. Applebaum, B., Arkis, B.: On the power of amortization in secret sharing: d-uniform secret sharing and CDS with constant information rate. In Proc. TCC 2018. pp.317–344 (2018)
4. Applebaum, B., Arkis, B., Raykov, P., Vasudevan, P.N.: Conditional disclosure of secrets: amplification, closure, amortization, lower-bounds, and separations. In Proc. CRYPTO 2017. pp. 727–757 (2017)
5. Applebaum, B., Beimel, A., Farràs, O., Nir, O., Peter, N.: Secret sharing schemes for general and uniform access structures. In Proc. EUROCRYPT 2019 Part III. pp. 441–471 (2019)
6. Applebaum, B., Holenstein, T., Mishra, M., Shayevitz, O.: The communication complexity of private simultaneous messages, revisited. J. Crypto. **33**, 916–953 (2020)
7. Assouline, L., Liu, T.: Multi-party PSM, revisited. To appear in TCC 2021. Cryptology ePrint 2019/657 (2019)
8. Applebaum, B., Vasudevan, P.N.: Placing conditional disclosure of secrets in the communication complexity universe. J. Crypto. **34**(2), 11 (2021)
9. Attrapadung, N.: Dual system encryption via doubly selective security: framework, fully secure functional encryption for regular languages, and more. In Proc. EUROCRYPT 2014. pp. 557–577 (2014)
10. Ball, M., Holmgren, J., Ishai, Y., Liu, T., Malkin, T.: On the complexity of decomposable randomized encodings, or: how friendly can a garbling-friendly PRF be? In Proc. ITCS 2020. pp. 86:1–86:22 (2020)
11. Beimel, A., Farràs, O., Mintz, Y., Peter, N.: Linear secret-sharing schemes for forbidden graph access structures. In Proc. TCC 2017. pp. 394–423 (2017)
12. Beimel, A., Ishai, Y., Kumaresan, R., Kushilevitz, E.: On the cryptographic complexity of the worst functions. In Proc. TCC 2014. pp. 317–342 (2014)
13. Beimel, A., Kushilevitz, E., Nissim, P.: The complexity of multiparty PSM protocols and related models. In Proc. EUROCRYPT 2018 Part II. pp. 287–318 (2018)
14. Beimel, A., Peter, N.: Optimal linear multiparty conditional disclosure of secrets protocols. In Proc. ASIACRYPT 2018 Part III. pp. 332–362 (2018)
15. Cover, T.M., Thomas, J.A.: Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing). Wiley-Interscience, (2006)
16. Csirmaz, L.: The size of a share must be large. In Proc. EUROCRYPT '94. pp. 13–22 (1995)
17. Csirmaz, L.: The dealer's random bits in perfect secret sharing schemes. Studia Scientiarum Mathematicarum Hungarica **32**(3), 429–438 (1996)

18. Damgård, I., Larsen, K.G., Nielsen, J.B.: Communication lower bounds for statistically secure MPC, with or without preprocessing. In Proc. CRYPTO '19. pp. 61–84 (2019).
19. Data, D., Prabhakaran, V.M., Prabhakaran, M.M.: Communication and randomness lower bounds for secure computation. IEEE Transactions on Information Theory **62**(7), 3901–3929 (2016)
20. Delfs, H., Knebl, H.: Introduction to Cryptography: Principles and Applications. Springer, third edn. (2015)
21. Feige, U., Kilian, J., Naor, M.: A minimal model for secure computation (extended abstract). In Proc. STOC '94. pp. 554–563 (1994)
22. Gál, A., Rosén, A.: $\omega(\log n)$ lower bounds on the amount of randomness in 2-private computation. SIAM J. Comput. **34**(4), 946–959 (2005)
23. Gay, R., Kerenidis, I., Wee, H.: Communication complexity of conditional disclosure of secrets and attribute-based encryption. In Proc. CRYPTO 2015 Part II. pp.485–502 (2015)
24. Gertner, Y., Ishai, Y., Kushilevitz, E., Malkin, T.: Protecting data privacy in private information retrieval schemes. JCSS **60**(3), 592–629 (2000)
25. Goyal, V., Ishai, Y., Song, Y.: Tight bounds on the randomness complexity of secure multiparty computation. In Proc. CRYPTO 2022. (2022) Cryptology ePrint 2022/799
26. Ishai, Y.: Randomization techniques for secure computation. In: Secure Multi-Party Computation, Cryptology and Information Security Series, vol. 10, pp. 222–248. IOS Press (2013)
27. Ishai, Y., Kushilevitz, E.: Private simultaneous messages protocol with applications. In Proc. Israel Symposium on Theory of Computing and Systems. pp. 174–183 (1997)
28. Kushilevitz, E., Ostrovsky, R., Rosén, A.: Amortizing randomness in private multiparty computations. SIAM J. Discrete Math. **16**(1), 533–544 (2003)
29. Kushilevitz, E., Ostrovsky, Prouff, E., R., Rosén, A., Thillard, A., Vergnaud, D.: Lower and upper bounds on the randomness complexity of private compuations of AND. SIAM J. Discrete Math. **35**(1), 465–484 (2021)
30. Kushilevitz, E., Rosén, A.: A randomness-rounds tradeoff in private computation. SIAM J. Discrete Math. **11**(1), 61–80 (1998)
31. Larsen, K.G., Simkin, M.: Secret sharing lower bound: either reconstruction is hard or shares are long. In Proc. SCN 2020. pp.566–578 (2020).
32. Liu, T., Vaikuntanathan, V., Wee, H.: Conditional disclosure of secrets via non-linear reconstruction. In Proc. CRYPTO 2017. pp.758–790 (2017)
33. Liu, T., Vaikuntanathan, V., Wee, H.: Towards breaking the exponential barrier in general secret sharing. In Proc. EUROCRYPT 2018. pp.567–596 (2018)
34. Newman, I.: Private vs. common random bits in communication complexity. Inf. Process. Lett., 39(2):67–71 (1991).
35. Parter, M., Yogev, E.: Distributed algorithms made secure: a graph theoretic approach. In Proc. SODA 2019. pp.1693–1710 (2019)
36. Pillai, S.R.B., Prabhakaran, M., Prabhakaran, V.M., Sridhar, S.: Optimality of a protocol by Feige-Kilian-Naor for three-party secure computation. In Proc. INDOCRYPT 2019. pp. 216–226 (2019)
37. Wee, H.: Dual system encryption via predicate encodings. In Proc. TCC 2014. pp. 616–637 (2014)

A Survey on Related Results

Feige et al. constructed 2-party PSM protocols of polynomial communication complexity for functions in the class NL, which was later generalized to k parties and extended to the classes mod_pL and #L by Ishai et al. [27]. Furthermore, Feige et al. provided another 2-party PSM protocol of exponential communication complexity for an arbitrary function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ with universal reconstruction, in which the referee is independent of f. The latter protocol showed upper bounds $2^n + n + 1$ of the communication complexity and $2^n + n$ of the randomness complexity in the case of general functions with universal reconstruction.

Beimel, Ishai, Kumaresan, and Kushilevitz constructed a 2-party PSM protocol for an arbitrary function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ in which the referee can depend on f [12], which differs from the universal reconstruction. Their protocol demonstrated that the dependency of f for the referee actually provides a quadratic improvement over [21]; the upper bounds for complexity of the communication and randomness are $O(\sqrt{2^n})$. Later, Liu et al. provided a more direct construction of the same complexity [32].

The upper bound of [12,32] was extended to multi-party versions of PSM protocols by Beimel, Kushilevitz, and Nissim [13]. They constructed k-party PSM protocols for $f : (\{0,1\}^n)^k \to \{0,1\}$ of communication complexity $O_k(\sqrt{2^{kn}})$ for $k \geq 6$, and better bounds for $3 \leq k \leq 5$, which improves the previous upper bounds $O_k(2^{(k-1)n})$ [21,27], where O_k hides a constant factor in k. Most recently, Assouline and Liu [7] improved the upper bound presented by Beimel et al. [13]. For example, they showed the upper bounds of $O_k(2^{(k-1)n/2})$ for infinitely many k. Liu et al. [32] also improve the upper bound to a polynomial in the input length by specifying the class to homogeneous polynomials including the inner product [32].

As well as upper bounds, the lower bounds of communication complexity for PSM protocols were also investigated. Beimel et al. provided the lower bound 2^n of communication complexity for 2-party PSM protocols with universal reconstruction [12], which approximately matches the upper bound $2^n + n + 1$ by [21]. Applebaum, Holenstein, Mishra, and Shayevitz proved the lower bound $3n - O(\log n)$ of the communication complexity of PSM protocols for a non-explicit function by the random function argument [6]. They also constructed explicit functions, for which the communication complexity of PSM protocols is at least $3n - O(\log n)$ under the assumption that some hitting-set generator exists by partially derandomizing the proof for the non-explicit function. Most recently, Ball, Holmgren, Ishai, Liu, and Malkin proved nearly quadratic lower bounds in k for codeword length of k-decomposable randomized encoding for the element distinctness function [10], which equivalently corresponds to a k-party PSM protocol with 1-bit private inputs.

Gay, Kerenidis, and Wee showed the upper bounds $\sigma 2^{(n/2)+1}$ of communication complexity for a two-party CDS protocol with a σ-bit secret for a general function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ with linear reconstruction, in which the referee can be represented as a linear function in messages, and the first non-trivial lower bounds $\Omega(\log n)$ for a general function of 2-party general CDS protocols by revealing relationships of the communication complexity between CDS and 2-party one-way communication models [23]. Furthermore, they provided tight upper and lower bounds of 2-party CDS protocols with linear reconstruction for several explicit functions and parameter settings, and also provided applications to attribution-based encryption.

The lower bounds for a general function were later improved exponentially by Applebaum, Arkis, Raykov, and Vasudevan [4]. They showed the lower bound $\Omega(n)$ of communication complexity for 2-party general CDS protocols. Applebaum, Holenstein, Mishra, and Shayevitz also showed lower bounds for 2-party CDS protocols that satisfy some combinatorial properties from lower bounds of a weak variant of PSM protocols [6].

The upper bounds of communication complexity for k-party CDS protocols were studied by Beimel and Peter [14]. They constructed a k-party CDS protocol with linear reconstruction for $f : \{0,1\}^{n'} \times (\{0,1\}^n)^{k-1} \to \{0,1\}$ of communication complexity $O(2^{(k-1)n/2})$, which is independent of n'. This corresponds to the lower bounds provided by Beimel, Farràs, Mintz, and Peter up to a polynomial factor in k in linear reconstruction settings [11].

Liu, Vaikuntanathan, and Wee improved the upper bounds presented in [23] via non-linear reconstruction. They showed the subexponential upper bound $2^{o(n)}$ of communication complexity by constructing 2-party CDS protocols with non-linear reconstruction [32]. They also extended the result to constructions of PSM protocols for some explicit functions.

For a long secret, Applebaum and Arkis constructed k-party CDS protocols for a general function whose communication complexity of each party is bounded by a constant information rate with respect to the secret. More specifically, if the length σ of the secrets is exponential in the input length of the function, the communication complexity is at most 4σ for each party, and thus, the total communication complexity is at most $4k\sigma$.

Most recently, Applebaum and Vasudevan succeeded to prove new lower bounds for 2-party CDS protocols in [8] by relating them with communication complexity games such as $\mathrm{coNP^{cc}}, \mathrm{PP^{cc}}, \mathrm{AM^{cc}}$, etc. In particular, they showed a linear lower bound of 2-party CDS protocols with imperfect privacy (with imperfect correctness, respectively) for the inner product function by reducing it to a lower bound in $\mathrm{coNP^{cc}}$ (in $\mathrm{PP^{cc}}$, respectively).

B Randomness Bounds for PSM with Statistical Privacy

In this section, we provide statistical-privacy versions of the results in Sect. 3.

B.1 Connections from Communication to Randomness in PSM with Statistical Privacy

We will provide connections from communication complexity to randomness complexity in PSM protocols with statistical privacy. We first prove a statistical-privacy version of Theorem 9, namely, a randomness lower bound from communication complexity of PSM protocols with statistical privacy.

Theorem 17. *If a k-party PSM protocol has the perfect correctness and δ-statistical privacy, it holds that $|\mathcal{R}| > (1-2\delta)\min_{r\in\mathcal{R}}|\mathcal{M}_{[b]}(r)|$ for every $b \in \{0,1\}$.*

Proof. Fix any $b \in \{0,1\}$. From the δ-statistical privacy, there exists a distribution D_b such that we have $\Delta(\mathbf{P}(\mathbf{x}), D_b) \leq \delta$ for every $\mathbf{x} \in \mathcal{X}_{[b]}$. For every $\mathbf{x}, \mathbf{x}' \in \mathcal{X}_{[b]}$, we have $\Delta(\mathbf{P}(\mathbf{x}), \mathbf{P}(\mathbf{x}')) \leq \Delta(\mathbf{P}(\mathbf{x}), D_b) + \Delta(D_b, \mathbf{P}(\mathbf{x}')) \leq 2\delta$ from the triangle inequality. We now fix $\mathbf{x}_0 \in \mathcal{X}_{[b]}$ arbitrarily. It holds $\Delta(\mathbf{P}(\mathbf{x}_0), \mathbf{P}(\mathbf{x})) \leq 2\delta$ for every $\mathbf{x} \in \mathcal{X}_{[b]}$.

For contradiction, we assume that $|\mathcal{R}| \leq (1-2\delta)\min_{r\in\mathcal{R}}|\mathcal{M}_{[b]}(r)|$. We define a binary random variable A_r as follows. Choose $\mathbf{x}$ from $\mathcal{X}_{[b]}$ uniformly at random. Set $A_r = 1$ if $\mathbf{P}(\mathbf{x}; r) \in \mathrm{Supp}(\mathbf{P}(\mathbf{x}_0))$; otherwise, 0. From the perfect correctness, we have

$$\Pr_{\mathbf{x}}[\, A_r = 1\,] \leq \frac{|\mathrm{Supp}(\mathbf{P}(\mathbf{x}_0))|}{|\mathcal{M}_{[b]}(r)|} \leq \frac{|\mathcal{R}|}{\min_{r\in\mathcal{R}}|\mathcal{M}_{[b]}(r)|} \leq 1 - 2\delta.$$

Therefore, we have $E[A_r] \leq 1-2\delta$. Let $A := \sum_{r\in\mathcal{R}} A_r$. From the linearity of expectation, $E[A] = \sum_{r\in\mathcal{R}} E[A_r] \leq (1-2\delta)|\mathcal{R}|$. This implies that $\Pr_{\mathbf{x}}[\, A \leq (1-2\delta)|\mathcal{R}|\,] > 0$, i.e., there exists $\mathbf{x}^* \in \mathcal{X}_{[b]}$ such that $|\{r : \mathbf{P}(\mathbf{x}^*; r) \in \mathrm{Supp}(\mathbf{P}(\mathbf{x}_0))\}| \leq (1-2\delta)|\mathcal{R}|$. Thus, we have $\Delta(\mathbf{P}(\mathbf{x}^*), \mathbf{P}(\mathbf{x}_0)) > 2\delta$. Contradiction. □

Corollary 4 (Statistical-privacy part of Theorem 2 in Sect. 1). *Let λ be the communication complexity for a function $f : [N]^k \to \{0,1\}$ of a k-party PSM protocol that has the perfect correctness and δ-statistical privacy. Then, the randomness complexity ρ for f of a k-party PSM protocol is at least $\lambda + \log(1-2\delta) - 1$.*

Next, we prove a randomness upper bound from the communication complexity. In the setting of statistical privacy, we follow the same derandomization approach as those of Newman [34] and Applebaum and Vasudevan [8]. For the derandomization, we use a variant of pseudorandom generator called non-Boolean pseudorandom generator (nbPRG), which is defined below.

Definition 6. *We say that a function $G : \mathcal{L} \to \mathcal{R}$ ϵ-fools a function $D : \mathcal{R} \to \mathcal{M}$ if it holds $\Delta(D(G(s)), D(r)) \leq \epsilon$ for $s \xleftarrow{\$} \mathcal{L}$ and $r \xleftarrow{\$} \mathcal{R}$. We say that G is a $(\mathcal{D}, \epsilon)$-nbPRG if G ϵ-fools every $D \in \mathcal{D}$.*

By a probabilistic argument, we can show the existence of nb-PRGs with good parameters. (See Claim 2 in [8].)

Lemma 4. *For every finite sets $\mathcal{R}$, $\mathcal{M}$, family $\mathcal{D} \subseteq \{D : \mathcal{R} \to \mathcal{M}\}$, and $\epsilon > 0$, there exists a function $G : \mathcal{L} \to \mathcal{R}$ where $|\mathcal{L}| = |\mathcal{M}|^2 \cdot \log|\mathcal{M}| \cdot \log|\mathcal{D}| \cdot \epsilon^{-2}$ such that G is a $(\mathcal{D}, \epsilon)$-nbPRG.*

By applying the good nbPRG obtained in Lemma 4 to PSM protocols with statistical privacy, we can sparsify the randomness domain with an additional privacy error, as shown in the following theorem.

Theorem 18. *Suppose that we have a k-party PSM protocol for $f : \mathcal{X} \to \{0,1\}$ of message domain size $|\mathcal{M}|$ and randomness domain size $|\mathcal{R}|$ with perfect correctness and δ-statistical privacy. Then, there exists a k-party PSM protocol for $f : \mathcal{X} \to \{0,1\}$ of message domain size $|\mathcal{M}|$ and randomness domain size $|\mathcal{M}|^2 \cdot \log|\mathcal{M}| \cdot \log|\mathcal{X}| \cdot \delta^{-2}$ with perfect correctness and 2δ-statistical privacy.*

Proof. From Lemma 4, there exists a $(\{\mathbf{P}(\mathbf{x};\cdot)\}_{\mathbf{x}\in\mathcal{X}}, \delta)$-nbPRG $G : \mathcal{L} \to \mathcal{R}$, where $|\mathcal{L}| \leq |\mathcal{M}|^2 \cdot \log|\mathcal{M}| \cdot \log|\mathcal{X}| \cdot \delta^{-2}$. Let $P_i : \mathcal{X}_i \times \mathcal{R} \to \mathcal{M}_i$ be the original message function of the i-th party in the k-party PSM protocol for f. For every i, we define a corresponding new message function $P'_i : \mathcal{X}_i \times \mathcal{L} \to \mathcal{M}_i$ as $P'_i(x_i, G(s))$ for a common random string $s \xleftarrow{\$} \mathcal{L}$. Then, the new message function $\mathbf{P'} = (P'_i)_{i\in[k]}$ and the original referee R provide the desired PSM protocol for f, as shown below.

The new PSM protocol has the perfect correctness since for every $r \in \{G(s) : s \in \mathcal{L}\} \subseteq \mathcal{R}$ we have $R(\mathbf{P}(\mathbf{x}; r)) = f(\mathbf{x})$ from the perfect correctness of the original PSM protocol.

Since the original PSM protocol has δ-privacy, it holds $\Delta(\mathrm{sim}(f(\mathbf{x})), \mathbf{P}(\mathbf{x})) \leq \delta$. Then, we have from the triangular inequality

$$\Delta(\mathrm{sim}(f(\mathbf{x})), \mathbf{P'}(\mathbf{x})) \leq \Delta(\mathrm{sim}(f(\mathbf{x})), \mathbf{P}(\mathbf{x})) + \Delta(\mathbf{P}(\mathbf{x}), \mathbf{P'}(\mathbf{x})) \leq 2\delta.$$

Thus, the new PSM protocol has 2δ-privacy. □

From the above theorem, we directly obtain the following connection.

Corollary 5. *Let λ be the communication complexity for a function $f : [N]^k \to \{0,1\}$ of a k-party PSM protocol that has perfect correctness and δ-statistical privacy. Then, the randomness complexity ρ for f of a k-party PSM protocol is at most $2\lambda + \log\lambda + \log(k\log N) + 2\log\delta^{-1}$.*

Therefore, we can see that a $(2\lambda + \log\lambda + O(1))$-bit random string is sufficient for k-party PSM protocols of communication complexity λ for $f : [N]^k \to \{0,1\}$ with perfect correctness and $O(1)$-privacy from Corollary 5.

B.2 Lower Bounds for General Functions in PSM with Statistical Privacy

We provide a statistical-privacy version of Theorem 10, that is, communication and randomness lower bounds of k-party PSM protocols for general functions with universal reconstruction that has perfect correctness and δ-statistical privacy.

Theorem 19 (Statistical-privacy part of Theorem 3 in Sect. 1). *If a PSM protocol with universal reconstruction has perfect correctness and δ-statistical privacy, it holds that $|\mathcal{R}| \geq (1-2\delta)2^{N^{k-1}}/2$.*

Proof. The proof is similar to that of Theorem 10. As before, we identify $f \in \mathcal{F}$ as a part of inputs to the message function $\mathbf{P}$. From Theorem 17, it holds $|\mathcal{R}| \geq (1-2\delta)\min_{r\in\mathcal{R}}\max_{b\in\{0,1\}}|\mathcal{M}_{[b]}(r)|$. From Lemma 2, it holds $\max_{b\in\{0,1\}}|\mathcal{M}_{[b]}(r)| \geq 2^{N^{k-1}}/2$ for every $r \in \mathcal{R}$. Therefore, we obtain $|\mathcal{R}| \geq (1-2\delta)2^{N^{k-1}}/2$. □

B.3 Lower Bounds for Explicit Functions in PSM with Statistical Privacy

We provide a proof of the randomness lower bound for k-party PSM protocols for f_{IP} with δ-statistical privacy as follows:

Theorem 20 (Statistical-privacy part of Theorem 4 in Sect. 1). *If a PSM protocol for f_{IP} has perfect correctness and δ-statistical privacy, it holds that*

$$H(R) \geq H(\mathbf{X} \mid f(\mathbf{X})) - 2\delta(\lambda + kn - \log\delta).$$

For uniformly random inputs, it holds that $H(R) > (1-2\delta)kn - 2 - 2\delta(\lambda - \log\delta)$.

Proof. From the perfect correctness, Claim 3.3 in Theorem 11 also holds. Then, we have

$$\begin{aligned}
H(R) &\geq H(R \mid \mathbf{M}) \\
&= H(R\mathbf{X} \mid \mathbf{M}) - H(\mathbf{X} \mid R\mathbf{M}) && \text{(Chain rule)} \\
&\geq H(\mathbf{X} \mid \mathbf{M}) - H(\mathbf{X} \mid R\mathbf{M}) && \text{(Independence of } \mathbf{X} \text{ and } R\text{)} \\
&= H(\mathbf{X} \mid \mathbf{M}) && \text{(Claim 3.3)} \\
&= H(\mathbf{X} \mid \mathrm{sim}(f_{\mathrm{IP}}(\mathbf{X}))) - 2\delta(\log(|\mathcal{X}||\mathcal{M}|) - \log\delta) && \text{(Statistical privacy, Cor. 1)} \\
&\geq H(\mathbf{X} \mid f_{\mathrm{IP}}(\mathbf{X})) - 2\delta(\lambda + kn - \log\delta). && \text{(Data processing inequality)}
\end{aligned}$$

Let y and Y be the output $f_{\mathrm{IP}}(\mathbf{x})$ and the random variable describing $f_{\mathrm{IP}}(\mathbf{x})$, respectively. Let $n' = kn/2$. For uniformly random inputs $\mathbf{x}$, it follows that

$$\begin{aligned}
H(\mathbf{X} \mid f_{\mathrm{IP}}(\mathbf{X})) &\geq -\sum_{\mathbf{x}\in\mathcal{X}, y\in\{0,1\}} \Pr(\mathbf{x}, y)\log\Pr(\mathbf{x} \mid y) \\
&= -\sum_{\mathbf{x}: f_{\mathrm{IP}}(\mathbf{x})=0} \Pr(\mathbf{x}, 0)\log\Pr(\mathbf{x} \mid 0) \sum_{\mathbf{x}: f_{\mathrm{IP}}(\mathbf{x})=1} \Pr(\mathbf{x}, 1)\log\Pr(\mathbf{x} \mid 1) \\
&= 2^{-2n'} \cdot 2^{n'-1} \cdot (2^{n'}+1)\log 2^{n'-1} \cdot (2^{n'}+1) \\
&\quad + 2^{-2n'} \cdot 2^{n'-1} \cdot (2^{n'}-1)\log 2^{n'-1} \cdot (2^{n'}+1) \\
&= 2^{-n'-1} \cdot (2^{n'}+1)\log 2^{n'-1} \cdot (2^{n'}+1) \\
&> n' - 1 + \log(2^{n'}-1) \\
&> kn - 2.
\end{aligned}$$

Thus, the statement of the theorem holds. □

C Randomness Bounds for CDS with Statistical Privacy

C.1 Lower Bounds for General Functions in CDS with Statistical Privacy

We provide a statistical-privacy version of Theorem 13, that is, communication lower bounds of k-party CDS protocols with universal reconstruction for a general function that has statistical privacy. In the following theorem, we By Lemma 3, we also obtain the corresponding randomness lower bound from Lemma 3 as in the case of Theorem 13.

Theorem 21 (Statistical-privacy part of Theorem 5 in Sect. 1). *Let $\mathcal{F} := \{f : \mathcal{X} \to \{0,1\}\}$ and let $\mathcal{C}$ be a set of possible referees in CDS protocols. Suppose $|\mathcal{C}| \geq (|\mathcal{S}|^{-1} + \delta)^{-1}/2$. If a CDS protocol for $\mathcal{F}$ has perfect correctness and δ-statistical privacy, it holds that*

$$\lambda = \Omega\left(\frac{|\mathcal{X}| \cdot \log\left\{(|\mathcal{S}|^{-1} + \delta)^{-1}\right\}}{\log|\mathcal{C}|}\right).$$

Proof. The proof of Theorem 21 follows almost the same direction with that of Theorem 13. The major difference is the counterpart of Claim 4.2:

Claim. If a CDS protocol has δ-statistical privacy, for every referee P_0^*, every $s \in \mathcal{S}$, and every $\mathbf{x} \in \mathcal{X}$ for which $f(\mathbf{x}) = 0$,

$$\sum_{s \in \mathcal{S}} \Pr_r\left[\, P_0^*(\mathbf{x}, \mathbf{P}(\mathbf{x}, s; r)) = s \,\right] \leq 1 + \delta|\mathcal{S}|$$

holds.

Proof. From the δ-statistical privacy and triangle inequality, there exists a simulator sim such that for every $s \in \mathcal{S}$ and every $\mathbf{x}$ for which $f(\mathbf{x}) = 0$ we have

$$\Delta(\mathbf{P}(\mathbf{x}, s; \cdot), \mathrm{sim}(\mathbf{x}; \cdot)) \leq \delta.$$

Then, from the definition of Δ, it holds that

$$\left|\Pr_r\left[\, P_0^*(\mathbf{x}, \mathbf{P}(\mathbf{x}, s; r)) = s \,\right] - \Pr_{\mathrm{sim}}\left[\, P_0^*(\mathbf{x}, \mathrm{sim}(\mathbf{x})) = s \,\right]\right| \leq \delta.$$

By the triangle inequality,

$$\left|\sum_{s \in \mathcal{S}} \Pr_r\left[\, P_0^*(\mathbf{x}, \mathbf{P}(\mathbf{x}, s; r)) = s \,\right] - \sum_{s \in \mathcal{S}} \Pr_{\mathrm{sim}}\left[\, P_0^*(\mathbf{x}, \mathrm{sim}(\mathbf{x})) = s \,\right]\right| \leq \delta|\mathcal{S}|.$$

Thus, the statement follows directly from the above inequality. □

Rather than Eq. (1), we obtain the following in the case of statistical privacy.

$$\Pr_{\{s_i, r_i\}_{i \in [T]}}\left[\, \forall i \in [T],\, P_0(\mathbf{x}, (\mathbf{P}(\mathbf{x}, s_i; r_i))) = s_i \,\right] \begin{cases} = 1 & \text{if } f(\mathbf{x}) = 1; \\ \leq (|\mathcal{S}|^{-1} + \delta)^T & \text{otherwise.} \end{cases} \tag{3}$$

By Eq. (3) and the union bound, setting $T := \lceil \frac{\log|\mathcal{C}|+8}{\log\{(|\mathcal{S}|^{-1}+\delta)^{-1}\}} \rceil$, for every $\mathbf{x}$ for which $f(\mathbf{x}) = 0$,

$$\Pr_{\{s_i,r_i\}_{i\in[T]}} [\, \exists P_0^*,\, \forall i \in [T],\, P_0^*(\mathbf{x}, (\mathbf{P}(\mathbf{x}, s_i; r_i))) = s_i \,] \leq |\mathcal{C}| \cdot (|\mathcal{S}|^{-1} + \delta)^T \leq 1/2.$$

(Note that $T \geq 1$ since $|\mathcal{C}| \geq (|\mathcal{S}|^{-1} + \delta)^{-1}/2$.) The remaining part of the proof proceeds similar to that of Theorem 13, and we can obtain the lower bound. □

C.2 Lower Bounds for Explicit Functions in CDS with Statistical Privacy

In this section, we provide communication and randomness lower bounds of CDS protocols for explicit functions with statistical privacy as generalized results of Theorems 14 and 16

We first prove the lower bounds for $f_{\mathrm{xor}}, f_{\mathrm{and}}$, and f_{IP}.

Theorem 22 (Statistical-privacy part of Theorem 6). *If a CDS protocol for $f : (\{0,1\}^n)^k \to \{0,1\} \in \{f_{\mathrm{xor}}, f_{\mathrm{and}}, f_{\mathrm{IP}}\}$ has perfect correctness and δ-statistical privacy, then it holds that*

$$\begin{aligned}
\lambda &\geq \frac{k}{1+2\delta k} H(S) - \frac{2\delta k}{1+2\delta k} (\log|\mathcal{S}| - \log \delta), \\
\rho &\geq (k-1)H(S) - 2\delta k(\lambda + H(S) - \log \delta).
\end{aligned}$$

Proof. To derive lower bounds, we also use Claim 4.3 in Theorem 14 and the values of $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(0)}$ defined in the proof of the claim. Let $M_i^{(1)}$ and $M_i^{(0)}$ denote $M_i(x_i^{(1)}) = P_i(x_i^{(1)}; R)$ and $M_i(x_i^{(0)}) = P_i(x_i^{(0)}, R)$, respectively. Let $\mathbf{M}_{i:j}^{(1)}$ and $\mathbf{M}_{i:j}^{(0)}$ with $i \leq j$ denote $(M_i^{(1)}, \ldots, M_j^{(1)})$ and $(M_i^{(0)}, \ldots, M_j^{(0)})$, respectively. Then, from the chain rule, for any $1 \leq i \leq k$, we obtain

$$H(\mathbf{M}_{1:i}^{(1)}) = H(M_i^{(1)} | \mathbf{M}_{1:i-1}^{(1)}) + H(\mathbf{M}_{1:i-1}^{(1)}).$$

We have

$$\begin{aligned}
H(M_i^{(1)} \mid \mathbf{M}_{1:i-1}^{(1)}) &\geq H(M_i^{(1)} \mid \mathbf{M}_{1:i-1}^{(1)} \mathbf{M}_{i+1:k}^{(1)}) \\
&= H(M_i^{(1)} S \mid \mathbf{M}_{1:i-1}^{(1)} \mathbf{M}_{i+1:k}^{(1)}) \\
&\quad -H(S \mid M_i^{(1)} \mathbf{M}_{1:i-1}^{(1)} \mathbf{M}_{i+1:k}^{(1)}) && \text{(Chain rule)} \\
&\geq H(S \mid \mathbf{M}_{1:i-1}^{(1)} \mathbf{M}_{i+1:k}^{(1)}) \\
&\quad -H(S \mid \mathbf{M}_{1:k}^{(1)}) \\
&= H(S \mid \mathbf{M}_{1:i-1}^{(1)} \mathbf{M}_{i+1:k}^{(1)}) && \text{(Perfect correctness)} \\
&\geq H(S \mid M_i^{(0)} \mathbf{M}_{1:i-1}^{(1)} \mathbf{M}_{i+1:k}^{(1)}) \\
&\geq H(S \mid \mathbf{M}(\mathbf{x}^{(0)})) && \text{(Relation between } \mathbf{x}^{(1)} \text{ and } \mathbf{x}^{(0)}) \\
&\geq H(S \mid \mathrm{sim}(\mathbf{x}^{(0)}))) \\
&\quad -2\delta(\log(|\mathcal{M}||\mathcal{S}|) - \log \delta) && \text{(Statistical privacy, Cor. 1)} \\
&\geq H(S) - 2\delta(\lambda + \log|\mathcal{S}| - \log \delta). && \text{(Independence of } S)
\end{aligned}$$

By summing up for $1 \le i \le k$, we have

$$H(\mathbf{M}(\mathbf{x})) + 2\delta k\lambda \ge kH(S) - 2\delta k(\log|\mathcal{S}| - \log \delta).$$

Here, it holds that $\lambda \ge H(\mathbf{M}(\mathbf{x}))$. Thus, we have

$$\lambda + 2\delta k\lambda \ge kH(S) - 2\delta k(\log|\mathcal{S}| - \log \delta),$$

and then

$$\lambda \ge \frac{k}{1+2\delta k}H(S) - \frac{2\delta k}{1+2\delta k}(\log|\mathcal{S}| - \log \delta).$$

From Lemma 3 and $\rho \ge H(R)$, it follows that

$$\rho \ge H(R) \ge (k-1)H(S) - 2\delta k(\lambda + \log|\mathcal{S}| - \log \delta).$$

Thus, the statement of the theorem holds. □

We next prove the lower bounds for nontrivial functions.

Theorem 23 (Statistical-privacy part of Theorem 7). *Let f be a θ-nontrivial function. If a CDS protocol for f has perfect correctness and δ-statistical privacy, then it holds that*

$$\begin{aligned}\lambda &\ge \frac{\theta}{1+2\delta\theta}H(S) - \frac{2\delta\theta}{1+2\delta\theta}(\log|\mathcal{S}| - \log \delta),\\ \rho &\ge (\theta-1)H(S) - 2\delta\theta(\lambda + \log|\mathcal{S}| - \log \delta).\end{aligned}$$

Proof. For a θ-nontrivial function f, we use nontrivial input pairs $(\mathbf{x}^{(1):j}, \mathbf{x}^{(0):j})_{j\in[\theta]}$ on $(i_j)_{j\in[\theta]}$. For each $j \in [\theta]$, Let $M_i^{(1):j}$ and $M_i^{(0):j}$ denote $M_i(x_i^{(1):j}) = P_i(x_i^{(1):j}, S; R)$ and $M_i(x_i^{(0):j}) = P_i(x_i^{(0):j}, S; R)$, respectively. For $I_j = \{i_1, \ldots, i_j\}$, let $\bar{I}_j = [k] \setminus I_j$ and $I_0 = \emptyset$.

Then, from the chain rule, for any $1 \le j \le \theta$, we obtain

$$H(\mathbf{M}^{(1):j}_{\bar{I}_{j-1}}) = H(M^{(1):j}_{i_j}|\mathbf{M}^{(1):j}_{\bar{I}_j}) + H(\mathbf{M}^{(1):j}_{\bar{I}_j}).$$

We have

$$\begin{aligned}
H(M^{(1):j}_{i_j}|\mathbf{M}^{(1):j}_{\bar{I}_j}) &\ge H(M^{(1):j}_{i_j} \mid \mathbf{M}^{(1):j}_{\bar{I}_j}\mathbf{M}^{(1):j}_{I_{j-1}})\\
&= H(M^{(1):j}_{i_j}S \mid \mathbf{M}^{(1):j}_{\bar{I}_j}\mathbf{M}^{(1):j}_{I_{j-1}})\\
&\quad -H(S \mid M^{(1):j}_{i_j}\mathbf{M}^{(1):j}_{\bar{I}_j}\mathbf{M}^{(1):j}_{I_{j-1}}) && \text{(Chain rule)}\\
&\ge H(S \mid \mathbf{M}^{(1):j}_{\bar{I}_j}\mathbf{M}^{(1):j}_{I_{j-1}})\\
&\quad -H(S \mid \mathbf{M}^{(1):j})\\
&= H(S \mid \mathbf{M}^{(1):j}_{\bar{I}_j}\mathbf{M}^{(1):j}_{I_{j-1}}) && \text{(Perfect correctness)}\\
&\ge H(S \mid M^{(0):j}_{i_j}\mathbf{M}^{(1):j}_{\bar{I}_j}\mathbf{M}^{(1):j}_{I_{j-1}})\\
&\ge H(S \mid \mathbf{M}(\mathbf{x}^{(0):j})) && \text{(Relation between } \mathbf{x}^{(1):j} \text{ and } \mathbf{x}^{(0):j})\\
&\ge H(S \mid \mathrm{sim}(\mathbf{x}^{(0):j})))\\
&\quad -2\delta(\log(|\mathcal{M}||\mathcal{S}|) - \log \delta) && \text{(Statistical privacy, Cor. 1)}\\
&\ge H(S) - 2\delta(\lambda + \log|\mathcal{S}| - \log \delta). && \text{(Independence of } S)
\end{aligned}$$

By summing up for $1 \leq j \leq \theta$, we have

$$H(\mathbf{M}(\mathbf{x})) + 2\delta\theta\lambda \geq \theta H(S) - 2\delta\theta(\log|\mathcal{S}| - \log\delta).$$

Here, it holds that $\lambda \geq H(\mathbf{M}(\mathbf{x}))$. Thus, we have

$$\lambda + 2\delta\theta\lambda \geq \theta H(S) - 2\delta\theta(\log|\mathcal{S}| - \log\delta),$$

and then

$$\lambda \geq \frac{\theta}{1 + 2\delta\theta} H(S) - \frac{2\delta\theta}{1 + 2\delta\theta}(\log|\mathcal{S}| - \log\delta).$$

From Lemma 3 and $\rho \geq H(R)$, it follows that

$$\rho \geq H(R) \geq (\theta - 1)H(S) - 2\delta\theta(\lambda + \log|\mathcal{S}| - \log\delta).$$

Thus, the statement of the theorem holds. $\square$