Exponential Lower Bounds for Secret Sharing

Kasper Green Larsen[*]
Aarhus University

Mark Simkin[†]
Aarhus University

Abstract

A secret sharing scheme allows a dealer to distribute shares of a secret among a set of n parties $P = \{p_1, \dots, p_n\}$ such that any authorized subset of parties can reconstruct the secret, yet any unauthorized subset learns nothing about it. The family $\mathcal{A} \subseteq 2^P$ of all authorized subsets is called the access structure. Classic results show that if $\mathcal{A}$ contains precisely all subsets of cardinality at least t, then there exists a secret sharing scheme where the length of the shares is proportional to $\lg n$ bits plus the length of the secret. However, for general access structures, the best known upper bounds have shares of length exponential in n, whereas the strongest lower bound shows that the shares must have length at least $n/\lg n$. Beimel conjectured that the exponential upper bound is tight, but proving it has so far resisted all attempts. In this paper, we almost prove Beimel's conjecture by showing that there exists an access structure $\mathcal{A}$, such that any secret sharing scheme for $\mathcal{A}$ must have either exponential share length, or the function used for reconstructing the secret by authorized parties must have an exponentially long description. As an example corollary, we conclude that if one insists that authorized parties can reconstruct the secret via a constant fan-in boolean circuit of size polynomial in the share length, then there exists an access structure that requires a share length that is exponential in n.

1 Introduction

A secret sharing scheme allows a dealer to distribute shares of a secret among a set of parties $P = \{p_1, \dots, p_n\}$ such that any authorized subset $A \subseteq P$ can reconstruct the secret, yet any unauthorized subset learns nothing about it. The family $\mathcal{A} \subseteq 2^{\{p_1,\dots,p_n\}}$ of all authorized subsets is called the access structure. Secret sharing was introduced independently by Shamir [Sha79] and Blakley [Bla79], who presented constructions for threshold access structures that contains all subsets with a cardinality larger than some threshold t. The first construction for general (monotone) access structures was presented by Ito, Saito, and Nishizeki [ISN89].

The main measure of efficiency for secret sharing schemes is the share size. For threshold access structures it is known that Shamir's secret sharing, which has a share size of $\Theta(\lg n)$, is optimal up to additive constants [BGK16]. This stands in stark contrast to the smallest share sizes we can achieve for general monotone access structures. The construction of Ito, Saito, and Nishizeki has a share size of $\mathcal{O}(2^n/\sqrt{n})$ and 29 years later the best known upper bound on the share size, due to Liu and Vaikuntanathan [LV18], is still $2^{0.994n}$. A widely believed conjecture suggests that these upper bounds are, up to constants, the best ones one can hope for. More concretely, Beimel conjectured:

Conjecture 1 ([Bei96, Bei11])**.** *There exists an $\epsilon > 0$ such that for every integer n there exists an access structure with n parties for which every secret sharing scheme distributes shares of length exponential in the number of parties, that is, $2^{\epsilon n}$.*

Proving this conjecture is a major open problem in the research area of secret sharing schemes. Karnin, Greene, and Hellman [KGH83] initiated a line of works [CDGV92, BDGV93, Csi95, Csi96] that proved different lower bounds on the share size using tools from information theory. The best of those lower

[*] Email: larsen@cs.au.dk. Supported by a Villum Young Investigator grant and an AUFF starting grant.

[†] Email: simkin@cs.au.dk. Supported by the European Unions's Horizon 2020 research and innovation program under grant agreement No 669255 (MPCPRO) and No 731583 (SODA).

bounds is due to Csirmaz [Csi95, Csi96], who uses Shannon information inequalities to prove that there exists an explicit access structure that requires shares of size $\Omega(n/\lg n)$. Csirmaz himself and subsequent works [BO09, MPY13] indicate that it is unlikely that one can prove a super-polynomial lower bound on the share size using such information inequalities.

A different line of works focuses on linear secret sharing schemes, where the shared secret is a linear combination of the shares. Many of the existing schemes, e.g. [Sha79], are linear and applications like multiparty computation [BOGW88, CCD88, RBO89] crucially rely on this property. Karchmer and Wigderson [KW93] introduce monotone span programs and show that these are closely related to linear secret sharing schemes. Through the lens of monotone span programs, a series of works obtained increasingly stronger lower bounds. Karchmer and Wigderson prove the first super-linear lower bound on the share size. Babai, Gál, and Wigderson [BGW99] prove the first super-polynomial lower bound. Finally, Robere et al. [RPRC16] prove a exponential lower bound, thus closing the gap between upper and lower bound for the case of linear secret sharing schemes.

Several works consider different flavors of the original secret sharing notion. Beimel and Franklin [BF07] consider a relaxed security notion of weak privacy, which only requires that any unauthorized subset can not exclude any secret value with certainty. The unauthorized subset can, however, conclude that some secret is more probable than another one. The authors show that this notion is strictly weaker than the original notion of secret sharing by constructing schemes with share sizes that are impossible for secret sharing schemes with perfect privacy. Among other results, the authors construct a weakly-private secret sharing scheme for the threshold access structures, where the share size is independent of n. The authors conclude that any sensible lower bound proof has to make use of the privacy requirement of secret sharing schemes. Applebaum et al. [AARV17, AA18] consider the efficiency of secret sharing schemes for large secrets. The authors show that, for a certain class of access structures, one can construct secret sharing schemes, where the share size does not grow with an increasing number n of parties. Their approach requires the secrets to be exponentially large in n.

Despite all progress that was made, a lower bound on the share size of secret sharing schemes for general access structures remained out of reach.

1.1 Our Contribution

In this work we make a significant step towards proving Beimel's conjecture. Informally, we show that either the total share size or the computational effort for reconstructing the secret has to be exponential in n. A bit more formally, let us consider a secret sharing scheme Σ for some access structure $\mathcal{A}$ that takes a 1-bit secret as input and outputs n shares, which are at most k bits long in total. Let $\mathcal{F}$ be some family of reconstruction functions. We require that for any authorized subset of parties $A \subseteq \mathcal{A}$, there exists at least one function in $\mathcal{F}$ that these parties can use to reconstruct the correct secret with probability at least $3/4$. For any $A \notin \mathcal{A}$, we require that all functions in $\mathcal{F}$ reconstruct the correct secret with probability at most $1/4$. These correctness and privacy requirements are very weak. Neither do we require perfect correctness, nor do we require privacy against an unauthorized set of parties that may use some function outside of $\mathcal{F}$ to reconstruct the secret. Proving a lower bound for such a secret sharing scheme makes our result only stronger, since any lower bound we can prove here also applies to any secret sharing scheme with better correctness and privacy guarantees. In this work we prove:

Theorem 1 (Informal). *There exists an access structure $\mathcal{A}$ such that any secret sharing scheme Σ for $\mathcal{A}$ with domain of secrets $\{0,1\}$ and total share length k satisfies*

$$\lg(|\mathcal{F}|) \cdot k = \Omega(2^n/\sqrt{n}).$$

Our result does not fully prove Beimel's conjecture, but it tells us that any secret sharing scheme for 1-bit secrets for general access structures, which has a reconstruction function whose description is sub-expontially large in n, must have a share size that is exponential in n.

To get a better feeling of what $\mathcal{F}$ is, one can, for example, imagine it to be the set of all functions from $\{0,1\}^k \to \{0,1\}$ that are computable by a constant fan-in boolean circuit of some size $t(k) \geq k$. Any one circuit can compute exactly one function, there are a constant amount of different gates types, and for any gate with constant fan-in, there are $t(k)^{\mathcal{O}(1)}$ choices for the input wires. It follows that there are at most $t(k)^{O(t(k))}$ different reconstruction functions in $\mathcal{F}$. Now, if for example $t(k) \leq k^c$ for a constant $c \geq 1$ (decoding by a circuit of size polynomial in the secret share length), then our theorem says that there exists an access structure $\mathcal{A}$ for which the share length k must be exponential in n. On the other hand, if k is for example polynomial in n, then our theorem tells us that there exists some access structure $\mathcal{A}$ which requires an exponentially large reconstruction circuit.

We prove Theorem 1 via a counting argument, meaning that we do not explicitly provide an access structure $\mathcal{A}$ that is affected by the lower bound. The high-level idea of our proof is as follows. Assume that there exists some secret sharing scheme $\Sigma_{\mathcal{A}}$ for every access structure $\mathcal{A}$ with the desired correctness and privacy properties and a total share size of $k^{\mathcal{A}} \leq k$. In the first step, we construct a family D that contains all access structures $\mathcal{A}$ of a certain type and we show that the size of this family is $2^{\Omega(2^n/\sqrt{n})}$. By the pingeon-hole principle, we know that the description of any $\mathcal{A} \in \mathcal{D}$ is at least $\lg |\mathcal{D}| = \Omega(2^n/\sqrt{n})$ bits long. On the other hand, we show that for any $\mathcal{A} \in \mathcal{D}$ one can use $\Sigma_{\mathcal{A}}$ to construct a $\mathcal{O}(\lg |\mathcal{F}| \cdot k)$-bit long *lossless* encoding from which $\mathcal{A}$ can be uniquely recovered. Combining the two observations directly yields the theorem stated above. The main challenge in realizing this proof idea lies in the construction of an appropriate encoding (and decoding) algorithm with the desired efficiency. Our encoding algorithm proceeds in two steps. First, we exploit the correctness and privacy properties of our secret sharing scheme to construct a randomized lossless encoding algorithm that works well for 99% of the sets A in any given $\mathcal{A}$ and encodes them into $\mathcal{O}(\lg |\mathcal{F}| \cdot k)$ bits. A careful analysis reveals that we can simply write out the remaining 1% of $A \in \mathcal{A}$ as part of the encoding and still obtain a lower bound on $\lg |\mathcal{F}| \cdot k$.

Proving lower bounds via such encoding arguments has been done quite extensively in the area of data structure lower bounds, see e.g. [PD06, PV10, Lar12a, Lar12b, VZ13, CKL18] and was also used recently to prove optimality of the Johnson-Lindenstrauss lemma in dimensionality reduction [LN17] and to prove optimality of ORAMs without balls-in-bins assumptions [LN18].

2 Formal Model and Result

In this section, we formally define secret sharing schemes and the precise conditions under which our lower bounds holds. Except for the security requirements, we define a secret sharing scheme precisely as in [Bei11].

Definition 1. *Let $\{p_1, \ldots, p_n\}$ be a set of parties. A collection $\mathcal{A} \subseteq 2^{\{p_1,\ldots,p_n\}}$ is monotone if $B \in \mathcal{A}$ and $B \subseteq C$ imply $C \in \mathcal{A}$. An* access structure *is a monotone collection $\mathcal{A} \subseteq 2^{\{p_1,\ldots,p_n\}}$ of non-empty subsets of $\{p_1, \ldots, p_n\}$. Sets in $\mathcal{A}$ are called* authorized*, and sets not in $\mathcal{A}$ are called unauthorized.*

Definition 2. *Let $\{p_1, \ldots, p_n\}$ be a set of parties. A distribution scheme $\Sigma = (\Pi, \mu)$ with domain of secrets $\{0, 1\}$ is a pair, where μ is a probability distribution on some finite set R called the set of random strings and Π is a mapping from $\{0, 1\} \times R$ to a set of n-tuples $\{0, 1\}^{k_1} \times \cdots \times \{0, 1\}^{k_n}$, where $\{0, 1\}^{k_j}$ is called the* domain of shares *of p_j. A dealer distributes a secret $b \in \{0, 1\}$ according to Σ by first sampling a random string $r \in R$ according to μ, computing a vector of shares $\Pi(b, r) = (s_1, \ldots, s_n)$ and privately communicating each share s_j to party p_j. For a set $A \subseteq \{p_1, \ldots, p_n\}$, we denote $\Pi(b, r)_A$ as the restriction of $\Pi(b, r)$ to its A-entries.*

When designing secret sharing schemes, one would typically consider larger domains of secrets than just a single bit as in Definition 2. In this paper we are proving a lower bound, so focusing on the simplest possible setting of a secret consisting of a single bit only makes our lower bound stronger and the proof simpler. The lower bound we prove in this paper holds for secret sharing schemes that are computationally more efficient when authorized parties reconstruct the secret than when unauthorized parties attempt to. We define this formally in the following:

Definition 3. *Let $\{p_1, \ldots, p_n\}$ be a set of parties, let $\mathcal{A} \subseteq 2^{\{p_1,\ldots,p_n\}}$ be an access structure and $\Sigma = (\Pi, \mu)$ a distribution scheme with domain of secrets $\{0, 1\}$ and domain of shares $\{0, 1\}^{k_1} \times \cdots \times \{0, 1\}^{k_n}$. Let $\mathcal{F}$ be* *a family of functions from* $\cup_{i=1}^{\infty} \left(\{0,1\}^i \to \{0,1\}\right)$ *and let* $\mathcal{U}$ *be the uniform distribution on* $\{0,1\}$*. We say that* $(\mathcal{F}, \mathcal{A}, \Sigma)$ *is an* efficient secret sharing scheme *if it satisfies the following two conditions:*

- *For any* $A \in \mathcal{A}$*, there exists a function* $f_A \in \left(\mathcal{F} \cap \left(\{0,1\}^{\sum_{j \in A} k_j} \to \{0,1\}\right)\right)$ *such that*

$$\left| \Pr_{b \sim \mathcal{U}, r \sim \mu}[f_A(\Pi(b,r)_A) = b] - \Pr_{b \sim \mathcal{U}, r \sim \mu}[f_A(\Pi(b,r)_A) \neq b] \right| \geq 3/4.$$

- *For any* $A \notin \mathcal{A}$*, it holds for all functions* $f \in \left(\mathcal{F} \cap \left(\{0,1\}^{\sum_{j \in A} k_j} \to \{0,1\}\right)\right)$ *that*

$$\left| \Pr_{b \sim \mathcal{U}, r \sim \mu}[f(\Pi(b,r)_A) = b] - \Pr_{b \sim \mathcal{U}, r \sim \mu}[f(\Pi(b,r)_A) \neq b] \right| \leq 1/4.$$

For intuition on Definition 3, consider as an example instantiating $\mathcal{F}$ to be the set that contains for each i, the set of all functions from $\{0,1\}^i \to \{0,1\}$ that are computable by a constant fan-in boolean circuit of size $t(i) \leq i^c$ for a constant $c > 1$, i.e. $\mathcal{F}$ contains functions computable by polynomially sized circuits. With this choice of $\mathcal{F}$, consider an access structure $\mathcal{A}$. A distribution scheme Σ gives an efficient secret sharing scheme $(\mathcal{F}, \mathcal{A}, \Sigma)$ precisely if any authorized set of parties $A \in \mathcal{A}$ can recover the secret using *some* constant fan-in boolean circuit with size polynomial in the share length, whereas no unauthorized set of parties can recover the secret using *any* constant fan-in boolean circuit with size polynomial in the share length. We can thus think of $\mathcal{F}$ as defining the computational resources with which authorized parties can recover the secret, but unauthorized parties cannot.

Discussion 1. When designing secret sharing schemes, one would typically insist that authorized parties can reconstruct the secret with probability $1 - \text{negl}(n)$. Similarly, one would insist that unauthorized parties cannot reconstruct the secret except with probability $\text{negl}(n)$. Since we are proving a lower bound, using the constants 3/4 and 1/4 in Definition 3 only makes our results stronger.

Discussion 2. One could consider allowing randomization in the algorithms used for reconstructing the secret, both for the authorized and unauthorized parties. That is, a natural extension of Definition 3 would say that *there exists a distribution* γ_A *over functions in* $\left(\mathcal{F} \cap \left(\{0,1\}^{\sum_{j \in A} k_j} \to \{0,1\}\right)\right)$ *such that* $\Pr_{b \sim \mathcal{U}, r \sim \mu, f_A \sim \gamma_A}[\cdots$. We remark that the definition would be equivalent to Definition 3 since one can always fix the randomness in f_A to achieve the same guarantees (equivalent to one direction of Yao's minimax principle).

Discussion 3. Our definition may seem superficially similar to the definition of weakly-private secret sharing schemes by Beimel and Franklin [BF07]. Their definition states that any unauthorized set cannot exclude any potential secret with probability 1. It does, however, allow the adversary to guess the secret correctly with a probability that is arbitrarily close to 1. In contrast to their definition, ours is strictly stronger, since it requires a sharp upper bound on the probability that an unqualified set of parties guesses the correct secret.

We are ready to present our main theorem in its full generality:

Theorem 2. *Let* $\{p_1, \ldots, p_n\}$ *be a set of parties and let* $\mathcal{F}$ *be a family of functions from* $\cup_{i=1}^{\infty} \left(\{0,1\}^i \to \{0,1\}\right)$*. There exists an access structure* $\mathcal{A} \subseteq 2^{\{p_1,\ldots,p_n\}}$ *such that any efficient secret sharing scheme* $(\mathcal{F}, \mathcal{A}, \Sigma)$ *with domain of secrets* $\{0,1\}$ *and domain of shares* $\{0,1\}^{k_1} \times \cdots \times \{0,1\}^{k_n}$ *with* $k = \sum_j k_j$*, satisfies*

$$\lg(|\mathcal{F} \cap \left(\cup_{i=1}^{k} \left(\{0,1\}^i \to \{0,1\}\right)\right)|) \cdot k = \Omega(2^n/\sqrt{n}).$$

To appreciate Theorem 2, consider instantiating $\mathcal{F}$ to be the set that contains for each i, the set of all functions from $\{0,1\}^i \to \{0,1\}$ that are computable by a constant fan-in boolean circuit of size $t(i)$ (with $t(i) \geq i$). A simple counting argument shows that $|\mathcal{F} \cap \left(\cup_{i=1}^{k} \left(\{0,1\}^i \to \{0,1\}\right)\right)| \leq t(k)^{O(t(k))}$ (A circuit computes only one function and there are $t(k)^{O(1)}$ choices for the input wires to each gate, there are $O(1)$ choices for the function computed by each gate, and there are $t(k)$ gates). Theorem 2 thus gives us that there must exist an access structure $\mathcal{A}$ such that any efficient secret sharing scheme $(\mathcal{F}, \mathcal{A}, \Sigma)$ with domain of shares $\{0,1\}^{k_1} \times \cdots \{0,1\}^{k_n}$ with $k = \sum_j k_j$ must satisfy $t(k)\lg(t(k))k = \Omega(2^n/n^{1/2})$. If we plug in polynomially sized constant fan-in boolean circuits, i.e. $t(i) \leq i^c$ for a constant $c \geq 1$, this gives us that $k^{c+2} = \Omega(2^n/n^{1/2}) \Rightarrow k = 2^{\Omega(n)}$, i.e. any secret sharing scheme for $\mathcal{A}$ must have shares with exponential length if decoding can be done by constant fan-in boolean circuits with size polynomial in the share length. Moreover, the lower bound holds *even if* we *only require* that unauthorized parties cannot reconstruct the secret using a polynomially sized constant fan-in boolean circuit (polynomial in the length of the shares). Notice that since this is a lower bound, it only makes the result stronger than if we e.g. required that a computationally unbounded set of parties cannot reconstruct the secret. We can also deduce from Theorem 2 that the size of the decoding circuit must be exponential in n, regardless of the share length.

Another interesting instantiation of Theorem 2 is to let $\mathcal{F}$ consist of all functions computable by a Turing machine with at most 10^6 states and alphabet $\{0,1\}$ (or some other constant number of states). Then $|\mathcal{F} \cap (\{0,1\}^i \to \{0,1\})| = O(1)$ and the lower bound says that there exists an access structure $\mathcal{A}$ for which any efficient secret sharing scheme $(\mathcal{F}, \mathcal{A}, \Sigma)$ must satisfy $k^2 = \Omega(2^n/\sqrt{n}) \Rightarrow k = 2^{\Omega(n)}$, i.e. shares must have exponential length if the secret can be reconstructed by authorized parties using a Turing machine with at most 10^6 states and binary alphabet. The lower bound holds as long as we require that unauthorized parties cannot recover the secret using a Turing machine with at most 10^6 states and alphabet $\{0,1\}$.

An even more exotic instantiation of Theorem 2 follows by letting $\mathcal{F}$ contain, for every i, the set of functions from $\{0,1\}^i \to \{0,1\}$ that are computable by a C-program with up to t ASCII characters. A counting argument shows that $|\mathcal{F} \cap \left(\cup_{i=1}^{k} \left(\{0,1\}^i \to \{0,1\}\right)\right)| \leq k2^{O(t)}$ (there are $2^{O(t)}$ sequences of t ASCII characters, and any program computes at most one function from $\{0,1\}^i \to \{0,1\}$) and we conclude that it must be the case that there exists an access structure $\mathcal{A}$ such that any efficient secret sharing $(\mathcal{F}, \mathcal{A}, \Sigma)$ must have $(t + \lg k) \cdot k = \Omega(2^n/\sqrt{n})$. This means that either the length of the C-program has to grow exponentially with the number of parties n, or the length of the shares has to grow exponentially with n. Thus if we insist on short shares, then the C-programs for reconstructing the secret have to be extremely non-uniform, and if we insist on reconstructing secrets using C-programs of any constant length t independent of n, then the shares must have exponential length. This lower bound holds as long as we require that unauthorized parties cannot recover the secret via a C-program of length t or less.

Finally, if one insist that authorized parties can *efficiently* reconstruct the secret via a C-program of length at most t ASCII characters, then the previous lower bound is strengthened. That is, we can now let $\mathcal{F}$ contain, for every i, the set of functions from $\{0,1\}^i \to \{0,1\}$ that are computable by a C-program with up to t ASCII characters that terminates in at most h steps. If we insist that authorized parties can reconstruct the secret by running such a C-program, then the lower bound $(t + \lg k) \cdot k = \Omega(2^n/\sqrt{n})$ holds even if we only require that unauthorized parties cannot reconstruct the secret via a C-program of length t and running time at most h steps.

3 Lower Bound Proof

To prove Theorem 2, let $\{p_1, \ldots, p_n\}$ be a set of parties and $\mathcal{F}$ a family of functions from $\cup_{i=1}^{\infty} \left(\{0,1\}^i \to \{0,1\}\right)$. Assume that there is a parameter k such that it holds for all access structures $\mathcal{A} \subseteq 2^{\{p_1,\ldots,p_n\}}$, that there exists an efficient secret sharing scheme $(\mathcal{F}, \mathcal{A}, (\Pi_{\mathcal{A}}, \mu_{\mathcal{A}}))$ with domain of secrets $\{0,1\}$ and domain of shares $\{0,1\}^{k_1^{\mathcal{A}}} \times \cdots \times \{0,1\}^{k_n^{\mathcal{A}}}$ with $\sum_j k_j^{\mathcal{A}} = k^{\mathcal{A}} \leq k$.

We will prove a lower bound on $\lg(|\mathcal{F} \cap \left(\cup_{i=1}^{k} \left(\{0,1\}^i \to \{0,1\}\right)\right)|) \cdot k$ via a counting argument. The high level intuition is that two distinct access structures $\mathcal{A}_1$ and $\mathcal{A}_2$ must be different either in terms of the shares they use, or in terms of the procedures used for reconstructing the secrets. Since there are overwhelmingly many distinct access structures, this gives a lower bound on either the share length (a lower bound on k), or on the descriptional size of the procedures used for reconstructing secrets (a lower bound on $\lg(|\mathcal{F} \cap \left(\cup_{i=1}^{k} \left(\{0,1\}^i \to \{0,1\}\right)\right)|)$).

More formally, let $\mathcal{D}$ be the family containing all access structures $\mathcal{A} \subseteq 2^{\{p_1,\dots,p_n\}}$ such that $\mathcal{A}$ contains no sets A of cardinality less than $\lfloor n/2 \rfloor$ and $\mathcal{A}$ contains all sets A of cardinality more than $\lfloor n/2 \rfloor$. We claim that $|\mathcal{D}| = 2^{\binom{n}{\lfloor n/2 \rfloor}} = 2^{\Omega(2^n/\sqrt{n})}$. To see this, observe that $\mathcal{A}$ is monotone for any choice of subsets with cardinality $\lfloor n/2 \rfloor$ that we might include in it. Since there are $\binom{n}{\lfloor n/2 \rfloor}$ subsets of cardinality $\lfloor n/2 \rfloor$, we conclude that there are $2^{\binom{n}{\lfloor n/2 \rfloor}}$ ways of choosing which subsets to include in $\mathcal{A}$.

We will show that we can encode any $\mathcal{A} \in \mathcal{D}$ into

$$\lambda = O(\lg(|\mathcal{F} \cap \left(\cup_{i=1}^{k} \left(\{0,1\}^i \to \{0,1\}\right)\right)|) \cdot k) + 0.1 \cdot \binom{n}{\lfloor n/2 \rfloor}$$

bits and still uniquely recover $\mathcal{A}$ from the encoding alone. The encoding procedure thus defines an injective mapping from $\mathcal{D}$ to $\{0,1\}^\lambda$. By the pigeon-hole principle, this implies that

$$\begin{aligned} \lambda &\geq \lg|\mathcal{D}| \Rightarrow \\ O(\lg(|\mathcal{F} \cap \left(\cup_{i=1}^{k} \left(\{0,1\}^i \to \{0,1\}\right)\right)|) \cdot k) &\geq 0.9 \cdot \binom{n}{\lfloor n/2 \rfloor} \Rightarrow \\ \lg(|\mathcal{F} \cap \left(\cup_{i=1}^{k} \left(\{0,1\}^i \to \{0,1\}\right)\right)|) \cdot k &= \Omega(2^n/\sqrt{n}). \end{aligned}$$

We are ready to describe our encoding and decoding procedures. For ease of notation, define $\mathcal{F}_{\leq k}$ as

$$\mathcal{F}_{\leq k} := \mathcal{F} \cap \left(\cup_{i=1}^{k} \left(\{0,1\}^i \to \{0,1\}\right)\right)$$

and define

$$\mathcal{F}_{=k} := \mathcal{F} \cap \left(\{0,1\}^k \to \{0,1\}\right).$$

Encoding. Let $\mathcal{A} \in \mathcal{D}$. Our procedure for uniquely encoding $\mathcal{A}$ is as follows:

1. For $i = 1,\dots,T$ for a parameter T to be fixed, consider sampling $b_i \sim \mathcal{U}$ as a uniform random bit, and sample $r_i \sim \mu_{\mathcal{A}}$. Let $A \subseteq \{p_1,\dots,p_n\}$ be an arbitrary set of cardinality $\lfloor n/2 \rfloor$ and define $k_A^{\mathcal{A}} = \sum_{j \in A} k_j^{\mathcal{A}}$. By Definition 3, it holds that:

   - If $A \in \mathcal{A}$, then there exists a function $f_A \in \mathcal{F}_{=k_A^{\mathcal{A}}}$ such that

   $$\left| \Pr_{b_i,r_i}[f_A(\Pi_{\mathcal{A}}(b_i,r_i)_A) = b_i] - \Pr_{b_i,r_i}[f_A(\Pi_{\mathcal{A}}(b_i,r_i)_A) \neq b_i] \right| \geq 3/4.$$

   - If $A \notin \mathcal{A}$, then for all functions $f \in \mathcal{F}_{=k_A^{\mathcal{A}}}$, it holds that

   $$\left| \Pr_{b_i,r_i}[f(\Pi_{\mathcal{A}}(b_i,r_i)_A) = b_i] - \Pr_{b_i,r_i}[f(\Pi_{\mathcal{A}}(b_i,r_i)_A) \neq b_i] \right| \leq 1/4.$$

   We use this observation as follows: We set $T = c \lg|\mathcal{F}_{\leq k}|$ for a sufficiently large constant $c > 1$. If $A \notin \mathcal{A}$, then since $|\mathcal{F}_{=k_A^{\mathcal{A}}}| \leq |\mathcal{F}_{\leq k}|$, we can use a Chernoff bound and a union bound over all $f \in \mathcal{F}_{=k_A^{\mathcal{A}}}$ to conclude that with probability at least 99/100, it holds simultanously for all $f \in \mathcal{F}_{=k_A^{\mathcal{A}}}$ that

   $$||\{i : f(\Pi_{\mathcal{A}}(b_i,r_i)_A) = b_i\}| - |\{i : f(\Pi_{\mathcal{A}}(b_i,r_i)_A) \neq b_i\}|| < T/3.$$

   At the same time, if $A \in \mathcal{A}$ and we have $T = c \lg|\mathcal{F}_{\leq k}|$, then with overwhelming probability, we will have that there exists at least one function $f \in \mathcal{F}_{=k_A^{\mathcal{A}}}$ such that

   $$||\{i : f(\Pi_{\mathcal{A}}(b_i,r_i)_A) = b_i\}| - |\{i : f(\Pi_{\mathcal{A}}(b_i,r_i)_A) \neq b_i\}|| > T/3.$$

Thus intuitively, the variables $b_1, \dots, b_T$ and $r_1, \dots, r_T$ reveal whether A is in $\mathcal{A}$ or not, i.e. they carry information about $\mathcal{A}$. We exploit this as follows: Let χ_A be the random variable taking the value 1 if the test

$$\exists f \in \mathcal{F}_{=k^{\mathcal{A}}_A} : ||\{i : f(\Pi_{\mathcal{A}}(b_i, r_i)_A) = b_i\}| - |\{i : f(\Pi_{\mathcal{A}}(b_i, r_i)_A) \neq b_i\}|| > T/3?$$

correctly predicts whether $A \in \mathcal{A}$. Then $\Pr[\chi_A = 1] \geq 99/100$. Let $\mathcal{S}$ be the family of all subsets of $\{p_1, \dots, p_n\}$ that have cardinality $\lfloor n/2 \rfloor$. It follows by linearity of expectation that $\mathbb{E}[\sum_{A \in \mathcal{S}} \chi_A] \geq 99|\mathcal{S}|/100$. This means that there must exist a choice values $\hat{b}_1, \dots, \hat{b}_T$ and $\hat{r}_1, \dots, \hat{r}_T$ such that the test $\exists f \in \mathcal{F}_{=k^{\mathcal{A}}_A} : ||\{i : f(\Pi_{\mathcal{A}}(\hat{b}_i, \hat{r}_i)_A) = \hat{b}_i\}| - |\{i : f(\Pi_{\mathcal{A}}(\hat{b}_i, \hat{r}_i)_A) \neq \hat{b}_i\}|| > T/3?$ correctly predicts whether $A \in \mathcal{A}$ for at least $99|\mathcal{S}|/100$ sets $A \in \mathcal{S}$. Fix such values.

2. Write down $\lg k$ bits specifying $k^{\mathcal{A}}$, followed by k bits specifying $k^{\mathcal{A}}_1, \dots, k^{\mathcal{A}}_n$ (this can be done by writing a length k bit string, where positions $\sum_{i=1}^{j} k^{\mathcal{A}}_i$ are set to 1 for all $j = 1, \dots, n$). Then write down the bits $\hat{b}_1, \cdots, \hat{b}_T$ and $\Pi_{\mathcal{A}}(\hat{b}_1, \hat{r}_1)), \cdots, \Pi_{\mathcal{A}}(\hat{b}_T, \hat{r}_T))$ for a total of at most $\lg k + k + T(1+k)$ bits.

3. Let $\bar{\mathcal{S}}$ be the subset of sets from $\mathcal{S}$ where the prediction is incorrect. Encode $\bar{\mathcal{S}}$ as a subset of $\mathcal{S}$ using $\lg \binom{n}{\lfloor n/2 \rfloor} \leq n$ bits to specify $|\bar{\mathcal{S}}|$ and $\lg \binom{|\mathcal{S}|}{|\bar{\mathcal{S}}|} \leq |\bar{\mathcal{S}}| \lg(e|\mathcal{S}|/|\bar{\mathcal{S}}|) \leq (|\mathcal{S}|/100) \lg(100e) < 0.1 \cdot \binom{n}{\lfloor n/2 \rfloor}$ bits to specify the subset.

Next we argue how to recover $\mathcal{A}$ from the above encoding:

Decoding.

1. Read the first $\lg k + k$ bits to recover $k^{\mathcal{A}}$ and $k^{\mathcal{A}}_1, \dots, k^{\mathcal{A}}_n$. Then use the following $T(k+1)$ bits to recover $\hat{b}_1, \dots, \hat{b}_T$ and $\Pi_{\mathcal{A}}(\hat{b}_1, \hat{r}_1), \dots, \Pi_{\mathcal{A}}(\hat{b}_T, \hat{r}_T)$.

2. For each $A \in \mathcal{S}$, iterate over all $f \in \mathcal{F}_{=k^{\mathcal{A}}_A}$ and compute the value

$$\Delta_f := \Big| |\{i : f(\Pi_{\mathcal{A}}(\hat{b}_i, \hat{r}_i)_A) = \hat{b}_i\}| - |\{i : f(\Pi_{\mathcal{A}}(\hat{b}_i, \hat{r}_i)_A) \neq \hat{b}_i\}| \Big|.$$

   Observe that the decoder can extract $\Pi_{\mathcal{A}}(\hat{b}_i, \hat{r}_i)_A$ from $\Pi_{\mathcal{A}}(\hat{b}_i, \hat{r}_i)$ since the decoder knows $k^{\mathcal{A}}_1, \dots, k^{\mathcal{A}}_n$. Thus the decoder can indeed compute Δ_f. If there is at least one f with $\Delta_f \geq T/3$, we initially predict that $A \in \mathcal{A}$ and otherwise, we predict that $A \notin \mathcal{A}$. These predictions are correct, except for $A \in \bar{\mathcal{S}}$.

3. Finally we read the last part of the encoding to determine which sets A that were predicted incorrectly in step 2. Together with the correct predictions from step 2., this recovers $\mathcal{A}$.

Analysis. Finally we derive the lower bound. We have just argued that we can give a unique encoding of each $\mathcal{A} \in \mathcal{D}$, hence the length of the encoding must be at least $\lg |\mathcal{D}| = \binom{n}{\lfloor n/2 \rfloor}$ bits. But the above encoding uses at most:

$$\lg k + k + T(1+k) + n + 0.1 \cdot \binom{n}{\lfloor n/2 \rfloor}$$

bits. Thus we must have

$$\begin{aligned} \lg k + k + T(1+k) + n + 0.1 \cdot \binom{n}{\lfloor n/2 \rfloor} &\geq \binom{n}{\lfloor n/2 \rfloor} \Rightarrow \\ Tk = \Omega\left(\binom{n}{\lfloor n/2 \rfloor}\right) &= \Omega(2^n/\sqrt{n}). \end{aligned}$$

But $T = c \lg |\mathcal{F}_{\leq k}|$ and we conclude:

$$\lg |\mathcal{F}_{\leq k}| \cdot k = \Omega(2^n/\sqrt{n}).$$

References

[AA18] Benny Applebaum and Barak Arkis. On the power of amortization in secret sharing: d-uniform secret sharing and CDS with constant information rate. In Amos Beimel and Stefan Dziembowski, editors, *TCC 2018: 16th Theory of Cryptography Conference, Part I*, volume 11239 of *Lecture Notes in Computer Science*, pages 317–344, Panaji, India, November 11–14, 2018. Springer, Heidelberg, Germany.

[AARV17] Benny Applebaum, Barak Arkis, Pavel Raykov, and Prashant Nalini Vasudevan. Conditional disclosure of secrets: Amplification, closure, amortization, lower-bounds, and separations. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology – CRYPTO 2017, Part I*, volume 10401 of *Lecture Notes in Computer Science*, pages 727–757, Santa Barbara, CA, USA, August 20–24, 2017. Springer, Heidelberg, Germany.

[BDGV93] Carlo Blundo, Alfredo De Santis, Luisa Gargano, and Ugo Vaccaro. On the information rate of secret sharing schemes (extended abstract). In Ernest F. Brickell, editor, *Advances in Cryptology – CRYPTO'92*, volume 740 of *Lecture Notes in Computer Science*, pages 148–167, Santa Barbara, CA, USA, August 16–20, 1993. Springer, Heidelberg, Germany.

[Bei96] Amos Beimel. *Secure schemes for secret sharing and key distribution.* Technion-Israel Institute of technology, Faculty of computer science, 1996.

[Bei11] Amos Beimel. Secret-sharing schemes: a survey. In *International Conference on Coding and Cryptology*, pages 11–46. Springer, 2011.

[BF07] Amos Beimel and Matthew K. Franklin. Weakly-private secret sharing schemes. In Salil P. Vadhan, editor, *TCC 2007: 4th Theory of Cryptography Conference*, volume 4392 of *Lecture Notes in Computer Science*, pages 253–272, Amsterdam, The Netherlands, February 21–24, 2007. Springer, Heidelberg, Germany.

[BGK16] Andrej Bogdanov, Siyao Guo, and Ilan Komargodski. Threshold secret sharing requires a linear size alphabet. In Martin Hirt and Adam D. Smith, editors, *TCC 2016-B: 14th Theory of Cryptography Conference, Part II*, volume 9986 of *Lecture Notes in Computer Science*, pages 471–484, Beijing, China, October 31 – November 3, 2016. Springer, Heidelberg, Germany.

[BGW99] László Babai, Anna Gál, and Avi Wigderson. Superpolynomial lower bounds for monotone span programs. *Combinatorica*, 19(3):301–319, 1999.

[Bla79] G.R. Blakley. Safeguarding cryptographic keys. pages 313–317. AFIPS Press, 1979.

[BO09] Amos Beimel and Ilan Orlov. Secret sharing and non-shannon information inequalities. In Omer Reingold, editor, *TCC 2009: 6th Theory of Cryptography Conference*, volume 5444 of *Lecture Notes in Computer Science*, pages 539–557. Springer, Heidelberg, Germany, March 15–17, 2009.

[BOGW88] Michael Ben-Or, Shafi Goldwasser, and Avi Wigderson. Completeness theorems for non-cryptographic fault-tolerant distributed computation (extended abstract). In *20th Annual ACM Symposium on Theory of Computing*, pages 1–10, Chicago, IL, USA, May 2–4, 1988. ACM Press.

[CCD88] David Chaum, Claude Crépeau, and Ivan Damgård. Multiparty unconditionally secure protocols (extended abstract). In *20th Annual ACM Symposium on Theory of Computing*, pages 11–19, Chicago, IL, USA, May 2–4, 1988. ACM Press.

[CDGV92] Renato M. Capocelli, Alfredo De Santis, Luisa Gargano, and Ugo Vaccaro. On the size of shares for secret sharing schemes. In Joan Feigenbaum, editor, *Advances in Cryptology – CRYPTO'91*, volume 576 of *Lecture Notes in Computer Science*, pages 101–113, Santa Barbara, CA, USA, August 11–15, 1992. Springer, Heidelberg, Germany.

[CKL18] Diptarka Chakraborty, Lior Kamma, and Kasper Green Larsen. Tight cell probe bounds for succinct boolean matrix-vector multiplication. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 1297–1306, 2018.

[Csi95] László Csirmaz. The size of a share must be large. In Alfredo De Santis, editor, *Advances in Cryptology – EUROCRYPT'94*, volume 950 of *Lecture Notes in Computer Science*, pages 13–22, Perugia, Italy, May 9–12, 1995. Springer, Heidelberg, Germany.

[Csi96] László Csirmaz. The dealer's random bits in perfect secret sharing schemes. *Studia Scientiarum Mathematicarum Hungarica*, 32(3):429–438, 1996.

[ISN89] Mitsuru Ito, Akira Saito, and Takao Nishizeki. Secret sharing scheme realizing general access structure. *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)*, 72(9):56–64, 1989.

[KGH83] Ehud Karnin, Jonathan Greene, and Martin Hellman. On secret sharing systems. *IEEE Transactions on Information Theory*, 29(1):35–41, 1983.

[KW93] Mauricio Karchmer and Avi Wigderson. On span programs. In *Structure in Complexity Theory Conference, 1993., Proceedings of the Eighth Annual*, pages 102–111. IEEE, 1993.

[Lar12a] Kasper Green Larsen. The cell probe complexity of dynamic range counting. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 85–94, 2012.

[Lar12b] Kasper Green Larsen. Higher cell probe lower bounds for evaluating polynomials. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*, pages 293–301, 2012.

[LN17] Kasper Green Larsen and Jelani Nelson. Optimality of the johnson-lindenstrauss lemma. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 633–638, 2017.

[LN18] Kasper Green Larsen and Jesper Buus Nielsen. Yes, there is an oblivious RAM lower bound! In *Advances in Cryptology - CRYPTO 2018 - 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2018, Proceedings, Part II*, pages 523–542, 2018.

[LV18] Tianren Liu and Vinod Vaikuntanathan. Breaking the circuit-size barrier in secret sharing. In Ilias Diakonikolas, David Kempe, and Monika Henzinger, editors, *50th Annual ACM Symposium on Theory of Computing*, pages 699–708, Los Angeles, CA, USA, June 25–29, 2018. ACM Press.

[MPY13] Sebastià Martín Mollevı́, Carles Padró, and An Yang. Secret sharing, rank inequalities and information inequalities. In Ran Canetti and Juan A. Garay, editors, *Advances in Cryptology – CRYPTO 2013, Part II*, volume 8043 of *Lecture Notes in Computer Science*, pages 277–288, Santa Barbara, CA, USA, August 18–22, 2013. Springer, Heidelberg, Germany.

[PD06] Mihai Pătraşcu and Erik D. Demaine. Logarithmic lower bounds in the cell-probe model. *SIAM Journal on Computing*, 35(4):932–963, 2006.

[PV10] Mihai Pătraşcu and Emanuele Viola. Cell-probe lower bounds for succinct partial sums. In *Proc. 21st ACM/SIAM Symposium on Discrete Algorithms (SODA)*, pages 117–122, 2010.

[RBO89] Tal Rabin and Michael Ben-Or. Verifiable secret sharing and multiparty protocols with honest majority (extended abstract). In *21st Annual ACM Symposium on Theory of Computing*, pages 73–85, Seattle, WA, USA, May 15–17, 1989. ACM Press.

[RPRC16] Robert Robere, Toniann Pitassi, Benjamin Rossman, and Stephen A. Cook. Exponential lower bounds for monotone span programs. In Irit Dinur, editor, *57th Annual Symposium on Foundations of Computer Science*, pages 406–415, New Brunswick, NJ, USA, October 9–11, 2016. IEEE Computer Society Press.

[Sha79] Adi Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, 1979.

[VZ13] Elad Verbin and Qin Zhang. The limits of buffering: A tight lower bound for dynamic membership in the external memory model. *SIAM J. Comput.*, 42(1):212–229, 2013.