Csirmaz's Duality Conjecture and Threshold Secret Sharing

Andrej Bogdanov
University of Ottawa, Canada

Abstract

We conjecture that the smallest possible share size for binary secrets for the t-out-of-n and $(n-t+1)$-out-of-n access structures is the same for all $1 \le t \le n$. This is a strenghtening of a recent conjecture by Csirmaz (*J. Math. Cryptol.*, 2020). We prove the conjecture for $t = 2$ and all n. Our proof gives a new $(n-1)$-out-of-n secret sharing scheme for binary secrets with share alphabet size n.

2012 ACM Subject Classification Theory of computation → Randomness, geometry and discrete structures; Theory of computation → Cryptographic primitives; Mathematics of computing → Information theory; Security and privacy → Mathematical foundations of cryptography

Keywords and phrases Threshold secret sharing, Fourier analysis

Digital Object Identifier 10.4230/LIPIcs.ITC.2023.3

Funding *Andrej Bogdanov*: This work was supported by RGC GRF grant CUHK 14301519 and NSERC grant RGPIN-2023-05006.

Acknowledgements Part of the research was carried out while the author was with the Chinese University of Hong Kong. I thank the anonymous ITC 2023 reviewers for helpful suggestions.

An *access structure* $\mathcal{A}$ over n parties is a nonempty monotone set system over ground set $\{1, \dots, n\}$. A *secret sharing scheme* [7, 1] for $\mathcal{A}$ with secret alphabet Σ is a collection of joint distributions $(X_1(\sigma), \dots, X_n(\sigma))$ with $\sigma \in \Sigma$ taking values in Γ^n such that

Secrecy: If $S \notin \mathcal{A}$ then $(X_i(\sigma) : i \in S)$ are identically distributed for all $\sigma \in \Sigma$.
Reconstruction: If $R \in \mathcal{A}$ then $(X_i(\sigma) : i \in R)$ determine σ with probability 1.

The *information rate* of the scheme is the ratio $\log|\Sigma| / \log|\Gamma|$ of the secret size and the share size. The *dual* of $\mathcal{A}$ is the access structure $\mathcal{A}^* = \{\overline{S} \colon S \notin \mathcal{A}\}$. Csirmaz [4] asks whether the following duality conjecture holds:

▶ **Conjecture 1.** *If $\mathcal{A}$ has a secret sharing scheme of information rate ρ for some secret alphabet size $|\Sigma|$, then $\mathcal{A}^*$ has a secret sharing scheme of information rate at least ρ for some secret alphabet size $|\Sigma'|$.*

As supporting evidence, Csirmaz shows that duality holds for the polymatroid relaxation of $\mathcal{A}$. This is a relaxation whose variables are the joint entropies of subsets of shares and whose constraints consist of a (in general incomplete) set of linear inequalities. On the other hand, he proves that duality fails for a relaxed asymptotic notion of secrecy. It is natural to consider the following even stronger conjecture:

▶ **Conjecture 2.** *For every Σ, if $\mathcal{A}$ has a secret sharing scheme of information rate ρ for secret alphabet Σ, then so does $\mathcal{A}^*$.*

In the case when Σ is the order of a finite field and the scheme is restricted to be linear, Conjecture 2 is known to hold (see Lemma 7.2 in [5]).

My motivation for Conjecture 2 is that it can be tested on threshold schemes. Such schemes have asymptotic information rate 1 as $|\Sigma|$ grows with the number of parties so Conjecture 1 does not say anything new about them. In contrast, when $|\Sigma| < n$, Conjecture 2 appears to be open for threshold schemes.

Here I study Conjecture 2 for threshold schemes and binary secrets, i.e., $|\Sigma| = 2$. This specialization is formulated as Conjecture 3. The *t-out-of-n* access structure consists of all t-element subsets of $\{1, \dots, n\}$.

▶ **Conjecture 3.** *If there exists a t-out-of-n scheme for binary secrets and share alphabet size γ then there also exists a $(n-t+1)$-out-of-n scheme for binary secrets and share alphabet size γ.*

The conjecture is true for every $n \geq 2$ when $t \in \{1, n\}$. Let $\gamma_2(\mathcal{A})$ denote the smallest possible share alphabet size for binary secrets and access structure $\mathcal{A}$. When $t = 1$ and $t = n$ one-bit secrets are possible and clearly optimal, so $\gamma_2(1\text{-out-of-}n) = \gamma_2(n\text{-out-of-}n) = 2$. A more interesting case is $t \in \{2, n-1\}$.

▶ **Proposition 4.** *For all $n \geq 2$, $\gamma_2(2\text{-out-of-}n) = n$.*

The lower bound $\gamma_2(2\text{-out-of-}n) \geq n$ was proved by Kilian and Nisan (see [2]). When n is a power of a prime (i.e., a finite field order) the upper bound can be obtained from Shamir's secret sharing with "infinity" as one of the evaluation points (see e.g. [3]). An alternative construction, which was communicated to me by Ilan Komargodski around 2016, works for all n. A variant of it is shown in the proof of Proposition 4 below.

If duality were to hold the same bound should be expected for the $(n-1)$-out-of-n access structure. The required lower bound was shown by Bogdanov, Guo, and Komargodski [2]. When n is a power of a prime the upper bound can also be derived from Shamir's scheme. The main result here is that this bound can be matched for non-prime powers n:

▶ **Theorem 5.** *For all $n \geq 2$, $\gamma_2((n-1)\text{-out-of-}n) = n$.*

The smallest example for which Theorem 5 is new is $n = 6$. This is a good example to keep in mind for the rest of the discussion.

Perspective: Lower bounds on alphabet size

There are two methods for lower bounding $\gamma_2(t\text{-out-of-}n)$ that give incomparable results. The analysis of Kilian and Nisan (KN) shows $\gamma_2(t\text{-out-of-}n) \geq n - t + 2$ for all $t \geq 2$. The analysis of Bogdanov, Guo, and Komargodski (BGK) shows the same lower bound for $\gamma_2((n-t+1)\text{-out-of-}n)$. Among the two, KN is more intuitive. They reduce their statement to the special case $t = 2$. When $t = 2$ let X_i and Y_i denote the i-th party's share of zero and one, respectively. Assuming the shares of zero and one are sampled independently, the KN bound follows from the two inequalities

$$1 = \mathbb{E}[1] \geq \mathbb{E}\big[|\{i\colon X_i = Y_i\}|\big] = \sum_{i=1}^{n} \Pr[X_i = Y_i] \geq \sum_{i=1}^{n} \frac{1}{|\Gamma|} = \frac{n}{|\Gamma|}.$$

The first inequality is by correctness of reconstruction (if $X_i = Y_i$ and $X_j = Y_j$ is possible the corresponding values would reconstruct to both zero and one) and the second one is by secrecy (X_i and Y_i are identically distributed, so $\Pr[X_i = Y_i]$ is a collision probability). The middle equality is linearity of expectation.

In contrast, BGK work directly with the probability mass functions p_0, p_1 of the shares of zero and one. They derive two types of constraints on the Fourier transform $\hat{f}$ of the real-valued function $f = p_1 - p_0$ over Γ^n. The first type is a reformulation of secrecy in the Fourier domain:

$$|\hat{f}(\chi)|^2 = 0 \qquad \text{for every } \chi \text{ such that } \operatorname{Supp}\chi \notin \mathcal{A}, \tag{BGK1}$$

where $\operatorname{Supp}\chi = \{i\colon \chi_i \neq 0\}$ is the support of the character χ viewed as an element of $\mathbb{Z}_q^n$ where $q = |\Gamma|$. The second type of constraint is the following (somewhat mysterious) relaxation of reconstruction:

$$\sum_A \Big(\sum_{\chi\colon \operatorname{Supp}\chi = A} |\hat{f}(\chi)|^2 \Big) \Big(-\frac{1}{q-1} \Big)^{|A \setminus B|} \geq 0 \qquad \text{for all } B \in \mathcal{A}. \tag{BGK2}$$

This system of constraints is a linear program in the variables $|\hat{f}(\chi)|^2, \chi \in \mathbb{Z}_q^n$. The BGK lower bound follows from its infeasibility when $q < n$ and $\mathcal{A}$ is the $(n-1)$-out-of-n access structure.

If Conjecture 3 were true, BGK would be a direct consequence of it and KN. Thus a natural first step towards Conjecture 3 would be to seek an alternative proof of BGK. The Conjecture itself suggests a route for such a proof: Assume that a $(n-t+1)$-out-of-n scheme with impossibly good share alphabet size γ_2 exists. Use this scheme to construct a t-out-of-n scheme with the same parameters. BGK offers a possible clue about this transformation: A feasible solution to the linear program (BGK1-BGK2) for access structure $\mathcal{A}$ should correspond to a secret sharing scheme for $\mathcal{A}^*$.

I do not know how to construct this transformation. For the purposes of investigating this potential "duality" between a secret sharing scheme and its Fourier transform it should be instructive to compare known secret sharing schemes for $\mathcal{A}$ and $\mathcal{A}^*$ and their Fourier transforms. I discovered the proof of Theorem 5 by working backwards from this correspondence. In the case of $(n-1)$-out-of-n schemes, the constraints (BGK1-BGK2) provide substantial information about what a scheme for this access structure should look like, if one exists at all. The scheme itself was obtained by reverse engineering f (and the distributions p_0 and p_1) from its Fourier transform. It would be interesting if the same result can be obtained by direct construction.

Concrete challenges

Figure 1 shows the best currently known lower and upper bounds on $\gamma_2(t\text{-out-of-}n)$ for small values of t and n. Except for the entries in bold, the upper bounds follow from Shamir's scheme, while the lower bounds are from KN or BGK. The upper bound for γ_2(3-out-of-5) can be obtained from a $(6, 4, 3)$ MDS code over $\mathbb{F}_4$ (see e.g. [6, Chapter 11]). The upper bound for γ_2(2-out-of-6) is from Proposition 4. The upper bound for γ_2(5-out-of-6) is from Theorem 5. The obvious next challenges are to calculate γ_2(3-out-of-6), and γ_2(4-out-of-6), or for those who prefer prime n, γ_2(3-out-of-7) and γ_2(5-out-of-7).

Constructions

Proof of the upper bound in Proposition 4. Shares of zero are n random symbols in $\Gamma = \{0, \ldots, n-1\}$ all equal to one another, while shares of one are a random cyclic permutation of the sequence $(0, 1, \ldots, n-1)$. Reconstruct to zero if the shares are equal and to one if they are different. The scheme is secret because the marginal distribution of every share is uniform (and therefore identical) in both cases. ◀

n \ t	1	2	3	4	5	6	7
2	2	2					
3	2	3	2				
4	2	4	4	2			
5	2	5	**4**	5	2		
6	2	**6**	5-7	5-7	**6**	2	
7	2	7	6-7	5-7	6-7	7	2

Figure 1 Upper and lower bounds on $\gamma_2(t\text{-out-of-}n)$.

The proof of Theorem 5 uses Fourier analysis of functions $f\colon \mathbb{Z}_q^n \to \mathbb{C}$. The q^n character functions

$$\chi(x) = \chi(x_0, \dots, x_{n-1}) = \exp\Big(\frac{2\pi i}{n} \cdot (x_0\chi_0 + \dots + x_{n-1}\chi_{n-1})\Big)$$

with $(\chi_0, \dots, \chi_{n-1}) \in \mathbb{Z}_q^n$ (also denoted by χ) form an orthonormal basis of the linear space of such functions with respect to the inner product $\langle f, g\rangle = \mathbb{E}[f(x)\overline{g(x)}]$ for x chosen uniformly at random from $\mathbb{Z}_q^n$. The Fourier transform of f is the unique function $\hat{f}\colon \mathbb{Z}_q^n \to \mathbb{C}$ for which $f = \sum_{\chi\in\mathbb{Z}_q^n} \hat{f}(\chi)\cdot\chi$. The Fourier coefficients $\hat{f}(\chi)$ are given by $\langle f, \chi\rangle$. Parseval's identity states that $\langle f, g\rangle = \sum_{\chi\in\mathbb{Z}_q^n} \hat{f}(\chi)\cdot\overline{\hat{g}(\chi)}$.

Proof of the upper bound in Theorem 5. Let $f\colon \mathbb{Z}_n^n \to \mathbb{C}$ be the function whose Fourier transform is

$$\hat{f}(\chi) = \begin{cases} 1, & \text{if } \chi \text{ is a cyclic shift of } (0, 1, \dots, n-1) \text{ or } (n-1, n-2, \dots, 0), \\ 0, & \text{if not.} \end{cases}$$

As will be shown shortly (or derived from symmetry of $\hat{f}$ under negation) f is real-valued. Shares of zero and one are sampled from the disjoint distributions p_0 and p_1 obtained by writing $f = C(p_0 - p_1)$ for a suitable normalizing constant $C > 0$. In more detail, let

$$p_0(x) = \begin{cases} C\cdot f(x), & \text{if } f(x) \ge 0 \\ 0, & \text{otherwise,} \end{cases} \quad \text{and} \quad p_1(x) = \begin{cases} -C\cdot f(x), & \text{if } f(x) \le 0 \\ 0, & \text{otherwise,} \end{cases}$$

where C is the factor that scales p_0 and p_1 to probability mass functions. The scaling factor is the same because $\hat{f}(0) = 0$.

Security follows from the fact that $\hat{f}(\chi)$ vanishes on all characters χ of Hamming weight at most $n-2$. In more detail, the advantage of any distinguisher D is

$$C\sum_{x\in\mathbb{Z}_n^n} D(x)f(x) = \frac{C}{n^n}\,\mathbb{E}\big[D(x)\overline{f(x)}\big] = \frac{C}{n^n}\sum_{\chi\in\mathbb{Z}_n^n} \hat{D}(\chi)\overline{\hat{f}(\chi)}$$

by Parseval's identity. If D depends on at most $n-2$ variables then $\hat{D}(\chi) = 0$ unless $|\chi| \le n-2$. As $\hat{f}(\chi) = 0$ for all χ of size at most $n-2$ the advantage of D must be zero.

To show reconstruction, f is calculated using the inverse Fourier formula. Letting $x = (x_0, \ldots, x_{n-1})$,

$$\begin{aligned}
f(x) &= \sum_{\chi \in \mathbb{Z}_n^n} \hat{f}(\chi)\chi(x) \\
&= \sum_{t \in \mathbb{Z}_n} \exp\Big(\frac{2\pi i}{n} \cdot \sum_{k=0}^{n-1} (k+t)x_k\Big) + \sum_{t \in \mathbb{Z}_n} \exp\Big(\frac{2\pi i}{n} \cdot \sum_{k=0}^{n-1} (-k+t)x_k\Big) \\
&= \sum_{t \in \mathbb{Z}_n} \exp\Big(\frac{2\pi i t}{n} \cdot \sum_{k=0}^{n-1} x_k\Big)\Big(\exp\Big(\frac{2\pi i}{n} \cdot \sum_{k=0}^{n-1} kx_k\Big) + \exp\Big(-\frac{2\pi i}{n} \cdot \sum_{k=0}^{n-1} kx_k\Big)\Big) \\
&= \Big(\sum_{t \in \mathbb{Z}_n} \exp\Big(\frac{2\pi i t}{n} \sum_{k=0}^{n-1} x_k\Big)\Big) \cdot 2\cos\Big(\frac{2\pi}{n} \sum_{k=0}^{n-1} kx_k\Big) \\
&= n \cdot \mathbf{1}\big(x_0 + \cdots + x_{n-1} = 0\big) \cdot 2\cos\Big(\frac{2\pi}{n} \cdot (x_1 + 2x_2 + \cdots + (n-1)x_{n-1})\Big).
\end{aligned}$$

Any $n-1$ of the n values $x_0, \ldots, x_{n-1}$ determine the remaining one on the set of inputs where f does not vanish. These values will satisfy the constraint $x_0 + \cdots + x_{n-1} = 0$ from which the missing x_i can be determined. This in turn determines the value of f and therefore the secret, which equals $\operatorname{sign} f(x)$ up to a change in representation. ◀

In more detail, the reconstruction procedure is this: Given shares $x_0, \ldots, x_{n-1}$ except for x_i, first compute $x_i = -\sum_{j \neq i} x_j \bmod n$, then output the sign of $\cos(2\pi(\sum kx_k)/n)$. (The cosine will never evaluate to zero because p_0 and p_1 assign zero probability to those shares.) Two alternative descriptions of $\operatorname{sign}\cos(2\pi(\sum kx_k)/n)$ are

- the parity of $\lfloor(\sum kx_k)/n\rceil$, where $\lfloor\cdot\rceil$ is the closest integer,
- the indicator of $|\lfloor\sum kx_k\rceil_n| < n/4$, where $\lfloor\cdot\rceil_n$ is the unique integer in the set $(-n/2, n/2]$ congruent modulo n.

The reconstruction procedure is clearly efficient. Its running time is quasilinear in n. How about sharing? Perfect sampling of the shares is not even possible in a model where the random seed is uniform over some finite domain! The reason is that some of the probabilities are irrational numbers. The scheme has perfect secrecy and reconstruction, but any realistic implementation of it must be imperfect.

It is possible to deduce from general considerations that if there exists a bit secret sharing scheme, then there exists one over the same share alphabet in which all probabilities are rational. The reason is that once the sign-pattern of f is fixed (i.e., once it is determined which shares reconstruct to zero and which reconstruct to one), finding the share probabilities that satisfy the secrecy constraints amounts to solving a linear program with rational coefficients. If this linear program is feasible then a rational solution must exist.

Nevertheless, even if imperfections in sampling are allowed, it is unclear how efficient a $(n-1)$-out-of-n scheme with share alphabet size n can be. Is it possible to sample an ϵ-approximation to the shares in time polynomial in n and $1/\epsilon$ for all n and ϵ?

To summarize, the crucial property of f is that its *weak sign* can be determined from any subset of shares that allow reconstruction. By weak sign I mean that one of the non-exclusive conclusions $f(x) \leq 0$ or $f(x) \geq 0$ can be reached only from knowledge of those coordinates of x that fall inside the reconstruction set. If an f with this property can be constructed under the constraints (BGK1) then reconstruction is possible. In the proof of Theorem 5 the cyclic structure of the nonvanishing Fourier coefficient plays a useful role. If, for example, $\hat{f}(\chi)$ was chosen to equal 1 on all characters of weight $n-1$ it appears that reconstruction wouldn't be possible.

Finally, notice the symmetry between the secret sharing scheme in the proof of Proposition 4 and the construction of $\hat{f}$ in the proof of Theorem 5. Is this a coincidence or an instance of duality?

References

1 G.R. Blakley. Safeguarding cryptographic keys. In *Proceedings of the 1979 AFIPS National Computer Conference*, pages 313–317, Monval, NJ, USA, 1979. AFIPS Press.

2 Andrej Bogdanov, Siyao Guo, and Ilan Komargodski. Threshold secret sharing requires a linear-size alphabet. *Theory of Computing*, 16(2):1–18, 2020. doi:10.4086/toc.2020.v016a002.

3 Ronald Cramer, Ivan Damgård, and Jesper Buus Nielsen. *Secure Multiparty Computation and Secret Sharing*. Cambridge University Press, 2015. URL: http://www.cambridge.org/de/academic/subjects/computer-science/cryptography-cryptology-and-coding/secure-multiparty-computation-and-secret-sharing?format=HB&isbn=9781107043053.

4 László Csirmaz. Secret sharing and duality. *J. Math. Cryptol.*, 15(1):157–173, 2020. doi:10.1515/jmc-2019-0045.

5 Satyanarayana V. Lokam. *Complexity Lower Bounds Using Linear Algebra*. Now Publishers Inc., Hanover, MA, USA, 2009.

6 Ron Roth. *Introduction to Coding Theory*. Cambridge University Press, 2006. doi:10.1017/CBO9780511808968.

7 Adi Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, November 1979. doi:10.1145/359168.359176.