Electronic Colloquium on Computational Complexity, Report No. 325 (2026)

On the Complexity of Symmetric Boolean Functions in Multiparty Private Simultaneous Messages

Marshall Ball[1], Peter Crawford-Kahrl[1], Vladimir Podolskii[2], and Morgan Prior[2]

[1] *New York University (marshall@cs.nyu.edu, peter.crawford.kahrl@gmail.com)*
[2] *Tufts University (podolskii.vv@gmail.com, morgan.prior@tufts.edu)*

Abstract

Multiparty Private Simultaneous Messages (PSM), also known as Decomposable Randomized Encoding (DRE) or information-theoretic garbling, is a model for non-interactive secure computation. Each party encodes its input using shared randomness and sends a message to an evaluator who does not know that randomness. The evaluator must recover the function output without learning anything else about the inputs. Perfect privacy requires identical transcript distributions on inputs with the same output. We also study the weaker notion of *support privacy*, which requires only that these distributions have the same support.

We study the communication complexity of symmetric Boolean functions when each of the n parties holds one input bit.

Lower bounds. We prove a $(1 - o(1))n \log n$ lower bound for support-private PSM protocols computing OR_n, and as a consequence for MAJ_n and EMAJ_n as well. We also prove an $\Omega(n \log q)$ lower bound for MOD_q when $2 \leq q \leq n$, with a constant independent of q. An alternative proof technique recovers most of these bounds as well as the previously known approach based on Nechiporuk's method.

Upper bounds. We construct perfectly private PSM protocols of total communication $\widetilde{O}(n^{3/2})$ for every symmetric Boolean function, improving on the known quadratic bound. For MAJ_n, known branching-program constructions yield an upper bound of $O(n \log^3 n \log \log n)$. For EMAJ_n the lower bound is tight. For MOD_q with prime q the lower bound is tight up to a constant factor. The matching upper bound for OR_n was previously known. Under support privacy, we obtain $\widetilde{O}(n)$ communication for every symmetric Boolean function.

Earlier lower-bound techniques from [BHI+20] yield only trivial linear bounds for symmetric Boolean functions. Concurrent and independent work by Holmgren and Wu [HW26] proves a $\omega(n)$ lower bound for OR_n using the same approach as our alternative technique. Both that work and the present work used substantial AI assistance.

ISSN 1433-8092

1 Introduction

Feige, Kilian, and Naor [FKN94] introduced a minimal model for secure computation, later named Private Simultaneous Messages (PSM) by Ishai and Kushilevitz [IK97]. The model is closely related to Decomposable Randomized Encodings (DRE) [IK02,AIK06,Ish13] and information-theoretic garbling [Yao86]. We refer to [BHR12,Ish13,App17] for surveys of these objects and their cryptographic applications.

In a PSM protocol for a function f, each of n parties holds an input x_i and locally computes a message m_i using a shared random string r. The parties send their messages to an evaluator, who has access to neither the inputs nor the shared randomness and must compute $f(x_1, \ldots, x_n)$. Privacy requires that the evaluator learn nothing else about the inputs. We study two information-theoretic notions of privacy, both defined through the *transcript*: the tuple of messages sent by all parties. Perfect privacy requires identical transcript distributions whenever $f(x) = f(x')$. Support privacy requires only that these distributions have the same support.

A central goal in this setting is to understand *the cost of privacy*. We study this cost through communication complexity. For a Boolean function f, we distinguish the individual message complexity $C_{\text{indiv}}(f)$ from the total communication complexity $C(f)$. The former measures the logarithm of the largest local message alphabet, minimized over protocols for f. The latter measures the length of the longest transcript, again minimized over protocols. We give the precise definitions in Section 2.2.

Without privacy, each party can send its input directly to the evaluator. When each $x_i \in \{0,1\}$, this costs n bits in total. Only a few functions, such as XOR_n, are known to admit private protocols with just n bits of communication.

For the general setting, the paper [BKN18] showed that any function on n bits can be computed with communication complexity $O(n^3 2^{n/2})$. This was later slightly improved by Assouline and Liu [AL21] when the input of each player is large. On the other hand, the authors in [BHI+20] showed that almost all functions require communication complexity $\Omega(n^2/\log n)$, and exhibited explicit functions with such cost (e.g. the distinct elements function). Closing the gap between upper and lower bounds for the "worst" functions remains a major open question.

We study this cost for *symmetric Boolean functions*: $f : \{0,1\}^n \to \{0,1\}$ such that

$$f(x_1, \ldots, x_n) = f(x_{\pi(1)}, \ldots, x_{\pi(n)}) \text{ for any permutation } \pi : [n] \xrightarrow{\sim} [n].$$

An equivalent way to express this condition is the following: there exists $g : \{0, \ldots, n\} \to \{0,1\}$ such that for all $x \in \{0,1\}^n$,

$$f(x) = g(\|x\|) \text{ where } \|x\| = \ \#\text{ of 1's in } x.$$

We focus on the following symmetric Boolean functions:

- Disjunction/OR: $\mathsf{OR}(x) = x_1 \vee \cdots \vee x_n$ (and, similarly, Conjunction/AND).
- Majority: $\mathsf{MAJ}(x) = 1$ if and only if $\|x\| \geq n/2$.
- Exact Majority: $\mathsf{EMAJ}(x) = 1$ if and only if $\|x\| = \lfloor n/2 \rfloor$.
- Modulo q: $\mathsf{MOD}_q(x) = 1$ if and only if $\|x\| \equiv 0 \pmod{q}$.

1.1 Our Results

Lower Bounds. The authors in [BHI+20] adapted Nechiporuk's method to prove lower bounds for multiparty PSMs with one input bit per party. Their method relies on the number of distinct subfunctions obtained by restricting input variables. For symmetric Boolean functions, it yields only the trivial lower bound: n.

In this paper we provide a simple combinatorial technique to prove lower bound for OR function. More specifically, we prove the following result.

Theorem 1 (Same as Corollary 15)**.** *For any perfectly correct, support-private PSM protocol computing* OR*, the communication complexity* $C(\mathsf{OR}) \geq (1 - o(1))n \log n$*. (The same bound holds for* AND*.)*

The matching upper bound of $O(n \log n)$ is known for this function [Ish13].

As a direct consequence we obtain lower bounds for other standard symmetric functions. The bound for EMAJ is tight, and the bound for MOD_q is tight up to a constant factor for prime q.

Theorem 2 (Same as Corollary 16)**.** *For any perfectly correct, support-private PSM protocol computing* MAJ *or* EMAJ*, the communication complexity is at least* $(1 - o(1))n \log n$*. For* MOD_q *with* $2 \leq q \leq n$*, the communication complexity of such protocols satisfies* $C(\mathsf{MOD}_q) = \Omega(n \log q)$*, where the constant in* Ω *does not depend on* q*.*

The drawback of the technique we use to prove Theorems 1 and 2 is that it cannot give lower bounds better than $\Omega(n \log n)$. This might be fine for symmetric functions, for which we know no function of higher cost, but it is not satisfactory for general functions. We introduce an alternative combinatorial technique to obtain lower bounds that does not suffer from this restriction. This technique is employed using the following framework:

1. Use our combinatorial technique to prove a lower bound on the *individual* message complexity $C_{\text{indiv}}(f)$.
2. Lift the lower bound on $C_{\text{indiv}}(f)$ to a bound on the *total* communication complexity $C(f)$ (see Lemma 7 below).

Step 1 uses only support privacy.[1] Section 4 proves the "agreement lemmas" that supply the combinatorial bound. Section 5 translates these lemmas into individual message lower bounds, and the lifting lemma in Section 2.2 converts them into total communication lower bounds.

This technique proves a superlinear lower bound for the OR function and recovers the same lower bounds as above (up to a constant factor) for MAJ, EMAJ and MOD_q (the last one only for q bounded away from n).

However, the advantage of this technique over the one used to prove Theorems 1 and 2 is that it readily generalizes to arbitrary Boolean functions and is capable of proving nearly quadratic lower bounds. In particular, we show that our technique subsumes the method based on Nechiporuk's bound used in [BHI+20], which implies that this method works in the support-private setting. Previously, Beimel et al. [BIKL25] showed that the method of [BHI+20] works for weak privacy.

[1] Perfect privacy implies support privacy, so a lower bound under support privacy also applies to perfectly private protocols.

Upper Bounds. For perfectly private PSM protocols, we improve the known $O(n^2)$ communication bound for symmetric Boolean functions [SESN23] to $\widetilde{O}(n^{3/2})$. Section 6 proves:

Theorem 3 (same as Theorem 36)**.** *For any symmetric Boolean function $f\colon \{0,1\}^n \to \{0,1\}$ there is a perfectly correct, perfectly private PSM of total complexity $\widetilde{O}(n^{3/2})$.*

For MAJ, a combination of known results matches our lower bound up to polylogarithmic factors:

Corollary 4 (same as Corollary 40)**.** *There is a perfectly correct, perfectly private simultaneous messages protocol for MAJ with total communication $O(n \log^3 n \log\log n)$.*

For EMAJ and for MOD_q with prime q, the upper bounds are tight up to constant factors in the stated regimes; in other words, they match the lower bounds we establish in Section 5.2. The upper bounds use variants of the protocol in [Ish13, Claim 4.5], which we recall in Theorem 12.

The universal support-private construction of [BIKL25] gives $\widetilde{O}(n^2)$ communication. For the special case of symmetric Boolean functions, we obtain a quasilinear bound:

Theorem 5 (same as Theorem 37)**.** *For any symmetric Boolean function $f\colon \{0,1\}^n \to \{0,1\}$ there is a perfectly correct, support-private PSM of total complexity $\widetilde{O}(n)$.*

This upper bound also limits lower-bound techniques that use only support privacy: such techniques cannot yield more than quasilinear bounds for symmetric Boolean functions. The constructions and proofs of these upper bounds appear in Section 6.

1.2 Overview of our Techniques

We provide a new general combinatorial framework that reduces proving lower bounds in the PSM model for arbitrary functions to studying a certain combinatorial object: f-cube systems (see Section 2.4). The connection is simple: fixing the shared randomness of an n-party PSM assigns a transcript to each vertex of the Boolean input cube $\{0,1\}^n$. Vertices correspond to parties' inputs, and the transcript at that vertex consists of the messages sent with those inputs and that randomness. Iterating through choices of randomness, generates a system of such cubes. We say that v is the S-corner of a cube if it is the transcript produced when exactly the parties in S hold a 1. For a bit $b \in \{0,1\}$, we say that an S-corner is a b-corner if the value of the function f on the corresponding input is b.

Perfect correctness and support privacy impose constraints on such cube systems. Perfect correctness forbids any transcript from appearing at both a 0-corner and a 1-corner. Support privacy, on the other hand, tells us that the same transcript must correspond to many b-corners across the cube system. Formally, if v occurs at a b-corner, it means that for every input S on which f outputs b, there is some cube whose S-corner is v.

$\Omega(n \log n)$ lower bounds. Let's consider the case of OR. The bound is elementary enough that we can effectively sketch the whole proof here.

Let Z be the set of 0-corners and let U be the set of all transcripts that occur. Write k_i for the size of party i's message alphabet and $g = (\prod_i k_i)^{1/n}$ for the geometric mean of the alphabet sizes. We will bound $|Z|/|U|$ from both sides by showing

$$\frac{1}{g} \le \frac{|Z|}{|U|} \le e^{-n/g}.$$

Taking logarithms gives $g \ln g \geq n$. For $n \geq 3$, if $g \leq n/\ln n$, then $g < n$ and $g \ln g < n$, a contradiction. Thus $g > n/\ln n$, and the sum of the logarithms of the local alphabet sizes satisfies

$$\sum_i \log_2 k_i = n \log_2 g \geq n \log_2(n/\ln n) = (1 - o(1))n \log_2 n.$$

For the upper bound on $|Z|/|U|$, select one cube for each transcript in Z. The indicator function χ_Q of a cube Q is 1 at its vertices and 0 elsewhere. Each selected cube contains just the point of Z for which it was chosen, so the indicators are linearly independent over $\mathbb{F}_2$. Since all the indicators vanish outside U, we can view them as vectors with one entry per transcript in U. This gives the dimension bound $|Z| \leq |U|$. If we hide some transcripts and the restricted indicators remain independent, the same argument bounds $|Z|$ by the smaller number of visible transcripts in U.

Choose a symbol a_i in each coordinate and hide every point that uses any of these symbols in its corresponding coordinate. The visible vertices still determine each cube Q. If only one of its two symbols in coordinate i is visible, the other must be a_i. For the dimension argument, reconstruction must also work for linear combinations of indicators. A coordinate line fixes all but one coordinate and lets the remaining coordinate range over its alphabet. Each line meets a Boolean cube in zero or two vertices, so every indicator and every linear combination of indicators has even parity on each line. If a point uses a_i in coordinate i, parity determines its value from the values at the other points on its line in direction i. Each of those points uses one fewer chosen symbol, so we can recover all hidden entries by induction. A linear combination that vanishes on the visible points therefore vanishes everywhere, which preserves independence.

Choose the a_i greedily. In coordinate i, choose the most frequent symbol among the transcripts of U still visible. This hides at least a $1/k_i$ fraction of those transcripts. After all n choices, at most a fraction $\prod_i (1 - 1/k_i)$ of U remains visible. Applying the arithmetic mean-geometric mean inequality on the reciprocals $1/k_i$, we get

$$\frac{|Z|}{|U|} \leq \prod_i (1 - 1/k_i) \leq e^{-\sum_i 1/k_i} \leq e^{-n/g}.$$

For the lower bound, support privacy lets us realize any transcript with output 1 when only party i holds a 1. Flipping that bit with the same shared randomness gives a zero-output transcript on the same line. Thus every line in direction i that meets U contains a point of Z. These lines are disjoint and each contains at most k_i points, so $|Z|/|U| \geq 1/k_i$ for every i. At least one k_i is at most g, and hence,

$$\frac{|Z|}{|U|} \geq \frac{1}{g}$$

This completes the proof.

Alternative lower bound approach. Our alternative lower bound argument exploits the fact that when each party's alphabet is small, some transcript must occupy corners for which f outputs different values.

We now sketch how to prove such lower bounds on the cube systems for symmetric functions by forcing a transcript collision. Suppose a cube has v as its S-corner. Then each party's message at that corner is fixed. The message that each party would've sent had it held the opposite bit, however, still remains unfixed.

Now we consider two cubes, one whose S-corner is v and one whose T-corner is v, where $T \subsetneq S$. Suppose that for each party i in some set $X \subseteq T$, the message i would've sent if its input bit flipped

is the same in both of the two cubes. If the parties in X switch from holding 1 to holding 0 in both cubes, the two cubes will produce the same transcript, as the $(S \setminus X)$-corner of the first cube equals the $(T \setminus X)$-corner of the second. If f takes a different value on inputs of Hamming weight $|T| - |X|$ than it does on inputs of Hamming weight $|S| - |X|$, we have reached a contradiction: two inputs with different f-values that produce the same transcript.

Our agreement lemmas guarantee us the existence of such a pair of cubes, and they are specialized based on f. For $f = \mathsf{OR}$, we take $X = T$, which means that the $(S \setminus T)$-corner of the first cube, where at least one party holds a 1, corresponds to the $\emptyset$-corner of the second, where no party holds a 1. The proof that such an S and T exist relies heavily on results from extremal combinatorics; in particular, it combines Higman's lemma with Ramsey's theorem.

For $f \in \{\mathsf{MAJ}, \mathsf{EMAJ}, \mathsf{MOD}_q\}$, it is enough to choose X so that exactly p parties of T lie outside X, where p is a Hamming weight threshold value for f. In this case, we argue that if such S and T did not exist, we would have error-correcting codes whose parameters beat Plotkin's bound.

We also show that our framework subsumes the Nechiporuk method for PSM protocols used in [BHI+20]. In particular, we show that given a sufficiently large lower bound on the number of distinct subfunctions of f, one can use a simple counting argument instead of specialized agreement lemma to guarantee the existence of a contradictory transcript. This result applies to arbitrary Boolean functions.

The agreement lemmas allow us to bound the size of the largest alphabet for any one party, so we use a lifting lemma (Lemma 7) to translate this individual bound into a bound on the total communication. The idea behind the lifting lemma is as follows: if a protocol has small total communication complexity, it cannot have too many long messages. Given a transcript with many short messages, we can consider the same problem restricted to the short message coordinates. Then, given a bound on the individual communication complexity of the restricted function, we can show that there must exist a transcript with even fewer short messages, which, after some iterations, yields a contradiction.

Upper bounds. Our general upper bounds compose an outer universal protocol with an inner protocol. The outer protocol evaluates a symmetric function from residues of the Hamming weight modulo small primes. Each inner protocol privately computes one outer message from the parties' input bits, without revealing the corresponding residue. We use the perfectly private universal protocol of [BKN18] for the $\widetilde{O}(n^{3/2})$ bound and the support-private construction of [BIKL25] for the $\widetilde{O}(n)$ bound. The upper bound for MAJ instead follows by converting a known permutation branching program into a PSM protocol.

1.3 Related Work

Concurrent lower bounds. Independently of this work, Holmgren and Wu [HW26] proved a $\omega(n)$ lower bound on the complexity of OR_n in the perfectly correct, perfectly private setting using the techniques that is very similar to our alternative approach. Their OR argument and ours share a common combinatorial core: privacy and correctness produce a family of locally colored subsets, while Ramsey's theorem and Higman's lemma force an incompatible agreement. Their work also characterizes symmetric functions admitting constant-size local alphabets in terms of constant periodicity, extends this characterization to sufficiently small statistical error, and proves quadratic lower bounds for nonexplicit Boolean functions. Both our paper and [HW26] obtained this result with substantial help from AI. Under the additional assumption that the support of the all-zero encoding is invariant under coordinate permutations, they obtain an $\Omega(n \log n)$ lower bound for OR.

Earlier lower-bound methods. Ball, Holmgren, Ishai, Liu, and Malkin [BHI+20] adapted Nechiporuk's method to DREs, obtaining $\Omega(n^2/\log n)$ lower bounds for explicit Boolean functions. For symmetric Boolean functions, restrictions remain symmetric, and the subfunction-counting method gives only a trivial linear bound. Ball and Randolph [BR22] extended this approach to many-party PSMs with longer local inputs. Shinagawa and Nuida [SN24] developed a different combinatorial representation using abstract simplicial complexes, yielding explicit, nonasymptotic lower bounds for concrete functions.

Weak privacy. The qualitative requirement that an adversary cannot rule out a possible secret was studied in secret sharing by Beimel and Franklin [BF07]. Beimel, Ishai, Kushilevitz, and Li [BIKL25] developed a quantitative notion of p-weak privacy that bounds likelihood ratios and includes the equal-support notion considered here when no quantitative lower bound on p is imposed. They also showed that the DRE lower-bound method of [BHI+20] extends to weak privacy, and constructed universal weakly private encodings. We use their construction as a black box for our support-private upper bound. Our term *support privacy* emphasizes equality of transcript supports, without requiring that the probabilities of those transcripts be close.

Upper bounds. Universal PSM constructions for arbitrary functionalities include those of Beimel, Kushilevitz, and Nissim [BKN18] and Assouline and Liu [AL21]. For symmetric functions with one input bit per party, Shinagawa, Eriguchi, Satake, and Nuida [SESN23] gave an $O(n^2)$ protocol based on quadratic residues, generalizing earlier quadratic-residue protocols of Feige, Kilian, and Naor and of Ishai. Eriguchi and Shinagawa [ES25] later developed constructions using universal reconstruction, and Nuida [Nui26] improved the histogram encodings underlying this approach. Eriguchi [Eri26] subsequently related the communication complexities of symmetric and general functionalities in both PSM and CDS, including improved PSM upper bounds when the number of parties exceeds the local alphabet size. These works also study related models such as ad hoc and robust PSM. We focus on the regime of a growing number of parties with one input bit per party, and study both quantitative lower bounds and the effect of relaxing perfect privacy to support privacy.

2 Preliminaries

Throughout, we use $[n] := \{1, \ldots, n\}$. We write $S \in \binom{[n]}{\geq k}$ when $S \subseteq [n]$ and $|S| \geq k$. For $S \subseteq [n]$, we take $\chi_S \in \{0,1\}^n$ to denote the characteristic vector of S: $(\chi_S)_i = 1 \iff i \in S$. Unless otherwise specified, $\log n$ denotes the base 2 logarithm of n. For a vector $x \in \{0,1\}^n$, we use $\|x\|$ to denote the Hamming weight of x. We use $\mathbb{1}[X]$ to denote the indicator function for the event X.

Our bounds concern the following explicit functions. Here $x = (x_1, x_2, \ldots, x_n)$ where each $x_i \in \{0,1\}$.

- The OR function, $\mathsf{OR}_n(x) := \bigvee_{i=1}^n x_i = 1 - (1-x_1)(1-x_2)\ldots(1-x_n)$.
- The Majority function, $\mathsf{MAJ}_n(x) = \mathbb{1}[\|x\| \geq n/2]$.
- The Modulo q function, $\mathsf{MOD}_{q,n}(x) = \mathbb{1}[\|x\| \equiv 0 \pmod q]$.
- The Exact Majority function, $\mathsf{EMAJ}_n(x) = \mathbb{1}[\|x\| = \lfloor n/2 \rfloor]$.

When the size of the input is clear from context, we omit the subscript n in our function names.

2.1 Private Simultaneous Messages

Definition (PSM for f)**.** Let $n \in \mathbb{N}$. Let $\mathcal{R}$ be a nonempty finite set from which the shared random string is drawn uniformly. For each $i \in [n]$, let $\mathcal{M}_i \subseteq \{0,1\}^*$ be a set of possible messages (not necessarily of equal length, and possibly containing the empty message of length 0), and let $\mathsf{M}_i \colon \{0,1\} \times \mathcal{R} \to \mathcal{M}_i$ be the function which returns the message that party i sends. Let $\mathsf{M} = (\mathsf{M}_1, \ldots, \mathsf{M}_n)$. Let $\mathsf{Rec} \colon \mathcal{M}_1 \times \cdots \times \mathcal{M}_n \to \{0,1\}$ be some recovery function. To reduce notation, we occasionally will write $\mathsf{M}(x,r) := (\mathsf{M}_1(x_1,r), \ldots, \mathsf{M}_n(x_n,r))$.

The pair $(\mathsf{M}, \mathsf{Rec})$ is a *PSM (n,K)-protocol for f* if $\max_i |\mathcal{M}_i| \leq K$ and it is perfectly correct, meaning that $\mathsf{Rec}(\mathsf{M}(x,r)) = f(x)$ for every $x \in \{0,1\}^n$ and $r \in \mathcal{R}$.

The protocol is a *perfect PSM* if it also satisfies *perfect privacy*. For all $x, x' \in \{0,1\}^n$ such that $f(x) = f(x')$, when r is drawn uniformly from $\mathcal{R}$ we have

$$(\mathsf{M}_1(x_1,r), \ldots, \mathsf{M}_n(x_n,r)) \stackrel{d}{=} (\mathsf{M}_1(x'_1,r), \ldots, \mathsf{M}_n(x'_n,r)) \ . \tag{1}$$

The authors in [BIKL25] introduced a weaker notion inspired by differential privacy. A protocol $(\mathsf{M}, \mathsf{Rec})$ is *p-weakly private* if

$$\Pr[\mathsf{M}(x,r) = y] \geq p \cdot \Pr[\mathsf{M}(x',r) = y]$$

for all x, x' such that $f(x) = f(x')$ and all y. Observe that perfect privacy is p-weak privacy in the case that $p = 1$.

A protocol is *weakly private* if it is p-weakly private for some $p > 0$. For the finite protocols considered here, this is equivalent to equality of transcript supports on inputs with the same output. We call this property *support privacy* and use the following equivalent formulation in our lower bounds:

Definition (Support Privacy)**.** Let $(\mathsf{M}, \mathsf{Rec})$ be a *PSM (n,K)-protocol for f*. Let $\mathsf{M}(x, \mathcal{R})$ denote $\{\mathsf{M}(x,r) \mid r \in \mathcal{R}\}$. Then we say that $(\mathsf{M}, \mathsf{Rec})$ satisfies *support privacy* if the following holds:

- **(Support Privacy)** For all $x, x' \in \{0,1\}^n$ such that $f(x) = f(x')$, $\mathsf{M}(x, \mathcal{R}) = \mathsf{M}(x', \mathcal{R})$, or, equivalently, for any x and any x' such that $f(x') = f(x)$ and any $r \in \mathcal{R}$, there exists $r' \in \mathcal{R}$ such that $\mathsf{M}(x,r) = \mathsf{M}(x',r')$.

2.2 Communication Complexity Metrics

We discuss a few worst-case metrics for communication complexity in the PSM/DRE setting. Fix a privacy requirement, either support privacy or perfect privacy. In each of the three definitions below, the minimum ranges over perfectly correct protocols satisfying the chosen privacy requirement. We suppress this requirement in the notation. Relations between measures discussed below also apply to both privacy settings. All lower bounds in the paper hold for support privacy, and thus imply lower bounds for perfect privacy as well. Upper bounds are stated directly as protocols, together with the privacy notion they achieve.

We first measure the largest local message alphabet. In an (n,K)-PSM protocol, every party's alphabet has size at most K. The *individual message complexity* of a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is the logarithm of the smallest such K:

$$C_{\text{indiv}}(f) = \min \{\log K \mid \text{ there exists an } (n,K)\text{-PSM for } f\} \, .$$

A slightly stronger metric, studied in [BHI+20, BR22], considers communication complexity to be the sum of individual message complexities of all parties:

$$C_{\text{indiv-sum}}(f) = \min\left\{ \sum_{i=1}^{n} \log K_i \mid \text{ there is a PSM for } f \text{ where } \mathsf{M} : \{0,1\}^n \times \mathcal{R} \to [K_1] \times \cdots \times [K_n] \right\}.$$

Note that these complexity measures are closed under taking a subfunction.

For a function $f : \{0,1\}^n \to \{0,1\}$, a set $I \subseteq [n]$, and assignment $z \in \{0,1\}^n$, denote by $f|_{\bar{I}=z} : \{0,1\}^{|I|} \to \{0,1\}$ the restriction obtained by fixing the coordinates outside I according to z:

$$f|_{\bar{I}=z}(x) = f(u) \text{ where } u_i = \begin{cases} x_i & \text{if } i \in I, \\ z_i & \text{if } i \notin I. \end{cases}$$

Lemma 6. *For every function $f : \{0,1\}^n \to \{0,1\}$, subset $I \subseteq [n]$, and assignment $z \in \{0,1\}^n$, we have*

$$C_{\text{indiv-sum}}(f) \geq C_{\text{indiv-sum}}(f|_{\bar{I}=z}) \qquad \text{and} \qquad C_{\text{indiv}}(f) \geq C_{\text{indiv}}(f|_{\bar{I}=z}).$$

Both inequalities hold for perfect privacy and for support privacy.

Proof. Fix a PSM protocol $(\mathsf{M}, \mathsf{Rec})$ for f with shared randomness R. Let $T = (M_j(z_j, R))_{j \in \bar{I}}$, and fix any t in the support of T. To compute $f|_{\bar{I}=z}$, the players in I draw their shared randomness from the distribution of R conditioned on $T = t$. Each player $i \in I$ then sends $M_i(x_i, R)$. The referee inserts t in the coordinates $\bar{I}$ and applies Rec. The resulting full transcript is a transcript of the original protocol on the extension u of x by z, so the referee outputs $f(u) = f|_{\bar{I}=z}(x)$.

For privacy, take x, x' with $f|_{\bar{I}=z}(x) = f|_{\bar{I}=z}(x')$. Their extensions u, u' have the same f-value, so their original transcript distributions are equal under perfect privacy and have equal supports under support privacy. Conditioning both on $T = t$ and projecting onto I preserves the respective equality. Finally, each player in I uses a subset of its original message alphabet. Thus both the sum and the maximum of the logarithms of the alphabet sizes can only decrease. Minimizing over protocols for f proves the claim. □

We also consider the worst-case total bit length of a transcript. The referee receives an ordered tuple of (clearly delineated) messages, and we only charge for bits in each message. Thus f has *communication complexity* at least k if every PSM protocol $(\mathsf{M}, \mathsf{Rec})$ for f has an input x and randomness r for which $|\mathsf{M}(x, r)| \geq k$.

$$C(f) := \min_{(\mathsf{M},\mathsf{Rec})\text{-PSM for } f} \max_{x,r} \sum_{i=1}^{n} |\mathsf{M}_i(x_i, r)|.$$

Note that for both privacy definitions and any f, we have $C_{\text{indiv-sum}}(f) \geq C(f)$. Indeed, consider a protocol with individual party alphabet sizes $K_1, \ldots, K_n$ attaining $C_{\text{indiv-sum}}(f)$. Injectively encode each $\mathcal{M}_i$ into strings of length at most $\lfloor \log K_i \rfloor$. Injectiveness preserves correctness and privacy, and the resulting transcript has length at most $\sum_i \lfloor \log K_i \rfloor \leq \sum_i \log K_i$.

The lifting lemma below aims to convert individual message lower bounds for restrictions of f into a total communication lower bound for f.

Lemma 7. *Let $f : \{0,1\}^n \to \{0,1\}$ be such that there are integers $1 \leq s \leq n$ and $B \geq 0$ where for every*

$$I \subseteq [n] \text{ such that } |I| \geq s$$

there exists $z \in \{0,1\}^n$ such that

$$C_{indiv}(f|_{\bar{I}=z}) \geq B + 1.$$

Then,

$$C(f) \geq (B+1)(n-s+1).$$

Remark 8. Note that for any $I \subseteq [n]$ of size s,

- $\mathsf{OR}_n|_{\bar{I}=0} \equiv \mathsf{OR}_s$,
- $\mathsf{MAJ}_n|_{\bar{I}=z} \equiv \mathsf{MAJ}_s$ where z has $\lceil n/2 \rceil - \lceil s/2 \rceil$ ones outside I,
- $\mathsf{EMAJ}_n|_{\bar{I}=z} \equiv \mathsf{EMAJ}_s$ where z has $\lfloor n/2 \rfloor - \lfloor s/2 \rfloor$ ones outside I,
- $\mathsf{MOD}_{q,n}|_{\bar{I}=0} \equiv \mathsf{MOD}_{q,s}$

These restrictions explain how individual alphabet bounds can supply the hypotheses of the lifting lemma.

In general for a symmetric function $f(x) = g(\|x\|)$, if $\boldsymbol{a}$ is a string with a ones outside of I then $f|_{\bar{I}=\boldsymbol{a}}(u) = g(a + \|u\|)$. In other words, in order to lower bound the communication complexity of a symmetric function, it suffices to find a (not too large) slice on which you can prove an alphabet size bound.

Remark 9. In the Lemma above, the bound on f will follow if one only knows that f has the property that there are integers s, B as above such that for every $J \subseteq [n]$ with $|J| \geq s$ there is $I \subseteq J$ and z such that $C_{\text{indiv}}(f|_{\bar{I}=z}) \geq B+1$.

The idea of the proof is that no short transcript can have too many long messages. Thus, given a transcript with many short messages, we can induce a PSM with the same privacy property on the function restricted to the short message coordinates. Then, given an individual message bound on PSMs for the restricted function, we can show that there must be a transcript with fewer short messages, which (after iterating) yields a contradiction.

Proof of Lemma 7. The same argument applies to perfect privacy and support privacy, using the same privacy requirement in both complexity measures.

Suppose for the sake of contradiction that

$$C(f) < (B+1)(n-s+1).$$

Let $(\mathsf{M}, \mathsf{Rec})$ be a protocol that realizes this inequality.

Say coordinate i is B-*long* in a transcript $\boldsymbol{v} = (v_1, \ldots, v_n)$ if $|v_i| > B$. Otherwise say that i is B-short for $\boldsymbol{v}$.

Let $\boldsymbol{v}$ be a longest transcript, with length $|\boldsymbol{v}| = k < (B+1)(n-s+1)$. Suppose $\boldsymbol{v}$ has ℓ B-long coordinates. Observe that $\ell \leq n-s$, so we have $n-\ell \geq s$ B-short coordinates. Let I ($|I| = n-\ell$) denote the set of B-short coordinates. To simplify the notation assume $I = [n-\ell]$.

By the condition on f, we know there exists a restriction z such that

$$C_{\text{indiv}}(f|_{\bar{I}=z}) \geq B+1.$$

Note that this condition implies that $\text{Im}(f|_{\bar{I}=z}) = \{0,1\}$. Hence for either $b \in \{0,1\}$, there exists y_b such that $f|_{\bar{I}=z}(y_b) = b$. For $y \in \{0,1\}^{n-\ell}$ define $u(y) \in \{0,1\}^n$ such that $u(y)_i = (y)_i$ if $i \in I$ and z_i otherwise. By perfect privacy, for *any* y such that $f|_{\bar{I}=z}(y) = b$,

$$\mathsf{M}(u(y), r) \stackrel{d}{=} \mathsf{M}(u(y_b), r).$$

Under support privacy, these transcript supports are equal instead. Now, we can define a new PSM with the same privacy property for $f|_{\bar{I}=z}$, $(\mathsf{M}', \mathsf{Rec}')$. We use M_i's as before, but instead of using R directly, we condition on hitting $\boldsymbol{v}_{\bar{I}}$ on inputs $z_{\bar{I}}$. Namely, let

$$R' = (R \mid \mathsf{M}_i(z_i, r) = v_i, \forall i \notin I).$$

The conditioning event has positive probability: choose $b = \mathsf{Rec}(\boldsymbol{v})$ and the corresponding y_b above; support privacy ensures that $\boldsymbol{v}$ occurs on $u(y_b)$. Then, to encode we use the following for $i \in I$:

$$\mathsf{M}'_i(x_i, r') = \mathsf{M}_i(x_i, r').$$

To reconstruct from $\boldsymbol{v}'$ we simply embed $\boldsymbol{v}'$ into $\boldsymbol{v}_{\bar{I}}$:

$$\mathsf{Rec}'(\boldsymbol{v}') = \mathsf{Rec}(\hat{u}(\boldsymbol{v}')) \text{ where } \hat{u}(\boldsymbol{v}')_i = \begin{cases} v'_i & \text{if } i \in I, \\ v_i & \text{if } i \notin I. \end{cases}$$

Perfect correctness and the chosen privacy property are inherited from $(\mathsf{M}, \mathsf{Rec})$.

To verify privacy, fix y, y' with $f|_{\bar{I}=z}(y) = f|_{\bar{I}=z}(y')$ and fix a transcript $\boldsymbol{v}'$. Perfect privacy gives

$$\begin{aligned}
\Pr[\mathsf{M}'(y) = \boldsymbol{v}'] \Pr[\mathsf{M}_{\bar{I}}(z_{\bar{I}}) = \boldsymbol{v}_{\bar{I}}] &= \Pr[\mathsf{M}_I(y) = \boldsymbol{v}' \mid \mathsf{M}_{\bar{I}}(z_{\bar{I}}) = \boldsymbol{v}_{\bar{I}}] \Pr[\mathsf{M}_{\bar{I}}(z_{\bar{I}}) = \boldsymbol{v}_{\bar{I}}] \\
&= \Pr[\mathsf{M}(u(y)) = \hat{u}(\boldsymbol{v}')] \\
&= \Pr[\mathsf{M}(u(y')) = \hat{u}(\boldsymbol{v}')] \\
&= \Pr[\mathsf{M}_I(y') = \boldsymbol{v}' \mid \mathsf{M}_{\bar{I}}(z_{\bar{I}}) = \boldsymbol{v}_{\bar{I}}] \Pr[\mathsf{M}_{\bar{I}}(z_{\bar{I}}) = \boldsymbol{v}_{\bar{I}}] \\
&= \Pr[\mathsf{M}'(y') = \boldsymbol{v}'] \Pr[\mathsf{M}_{\bar{I}}(z_{\bar{I}}) = \boldsymbol{v}_{\bar{I}}].
\end{aligned}$$

Dividing by $\Pr[\mathsf{M}_{\bar{I}}(z_{\bar{I}}) = \boldsymbol{v}_{\bar{I}}]$ yields $\Pr[\mathsf{M}'(y) = \boldsymbol{v}'] = \Pr[\mathsf{M}'(y') = \boldsymbol{v}']$. For support privacy, replace equality by equivalence of positivity in this display; the conditional supports are therefore equal.

However, because $(\mathsf{M}', \mathsf{Rec}')$ is a PSM protocol for $f|_{\bar{I}=z}$, it is subject to $C_{\mathrm{indiv}}(f|_{\bar{I}=z}) \geq B + 1$. If all its messages had length at most B, each local alphabet would have at most $2^{B+1} - 1$ symbols, a contradiction. Hence there must be some input y^* and $r^* \in \mathrm{Supp}(R') = \mathrm{Supp}(R \mid \mathsf{M}_i(z_i, r) = v_i, \forall i \notin I)$ such that $\mathsf{M}'(y^*, r^*)$ has at least one B-long coordinate. Hence if we consider

$$\boldsymbol{v}^* = \mathsf{M}(u(y^*), r^*)$$

we note that because the original B-long coordinates are unchanged, we now have more ($> \ell$) B-long coordinates in total. Let ℓ^* be the number of B-long coordinates in $\boldsymbol{v}^*$.

If $\ell^* > n - s$, then we have $|\boldsymbol{v}^*| \geq (B+1)(n-s+1)$.

Otherwise, $\ell^* \leq n - s$ and hence $\boldsymbol{v}^*$ has at least s B-short coordinates. In this case, we can apply the argument above again iteratively. As the number of B-long coordinates strictly increases each time, eventually we will produce a transcript with $> n - s$ B-long coordinates and reach a contradiction. □

2.3 Some Known Protocols

We use the following protocols as black boxes.

Theorem 10 (Beimel–Kushilevitz–Nissim [BKN18])**.** *Let $k, N \geq 2$ be integers. Every function*

$$f : [N]^k \longrightarrow \{0, 1\}$$

admits a k-party private simultaneous messages protocol with perfect correctness and perfect privacy, such that

$$C_{\textit{indiv-sum}}(f) = O\left(k^3 N^{k/2}\right).$$

Theorem 11 ([BIKL25]). *Let $f : \{0,1\}^n \to \{0,1\}$ be a function. Then there is a support-private, perfectly correct PSM for f with complexity $\widetilde{O}(n^2)$.*

The actual result from [BIKL25] proves p-weak privacy for some small p, but support privacy suffices for our results.

Theorem 12 (Ishai [Ish13]). *There exists a PSM protocol for OR with total communication complexity $C(\mathsf{OR}) \leq (1 + o(1))n \log n$.*

We recall the protocol for completeness.

Proof. Let $p > n$ be the smallest prime greater than n. From Bertrand's postulate we have $p < 2n$. All arithmetic that follows is modulo p.

1. The shared randomness will consist of $r_1, \ldots, r_n$ uniformly random elements of $\mathbb{Z}_p$ such that $\sum_i r_i = 0$ and a uniformly random $s \in \mathbb{Z}_p^*$.
2. The ith party, holding $x_i \in \{0,1\}$ then computes message $m_i = r_i + x_i \cdot s$.
3. The evaluator then computes $y = \sum_i m_i$ and outputs 1 if and only if $y \neq 0$.

When $\mathsf{OR}_n(x) = 0$, the transcript is uniform over message tuples whose sum is zero. When $\mathsf{OR}_n(x) = 1$, the sum is uniform over the nonzero field elements, and the transcript is uniform conditional on that sum. In either case, its distribution depends only on the output. In this protocol, each party sends a single element from $\mathbb{Z}_p$, yielding an overall communication cost of $n\lceil \log p \rceil \leq n(\log n + 2)$. □

2.4 f-Cube Systems

We represent transcript supports by a combinatorial object called an *f-cube system*. This representation isolates the correctness and privacy properties used by our lower bounds.

Definition (f-Cube System). Fix alphabets $\mathcal{M}_1, \ldots, \mathcal{M}_n$ and take

$$\Omega = \mathcal{M}_1 \times \cdots \times \mathcal{M}_n.$$

A *Boolean Cube System* (with respect to Ω) consists of a pair of disjoint sets

$$Z_0, Z_1 \subseteq \Omega$$

and a nonempty collection of possibly degenerate n-dimensional Boolean cubes $\mathcal{Q}$. A cube $Q \in \mathcal{Q}$ is specified by two symbols at each coordinate $i \in [n]$

$$\mathsf{M}_{i,0}(Q), \mathsf{M}_{i,1}(Q) \in \mathcal{M}_i.$$

The two symbols may coincide on an irrelevant coordinate; perfect correctness forces them to be distinct whenever f depends on that coordinate. We denote the *S-corner of a cube $Q \in \mathcal{Q}$* as

$$Q(S)_i = \begin{cases} \mathsf{M}_{i,1}(Q) & i \in S, \\ \mathsf{M}_{i,0}(Q) & i \notin S. \end{cases}$$

We say $\mathcal{Q}$ is an *f-Cube System* if

1. For all $Q \in \mathcal{Q}$, $b \in \{0,1\}$, and $S \subseteq [n]$, $f(\chi_S) = b \iff Q(S) \in Z_b$.

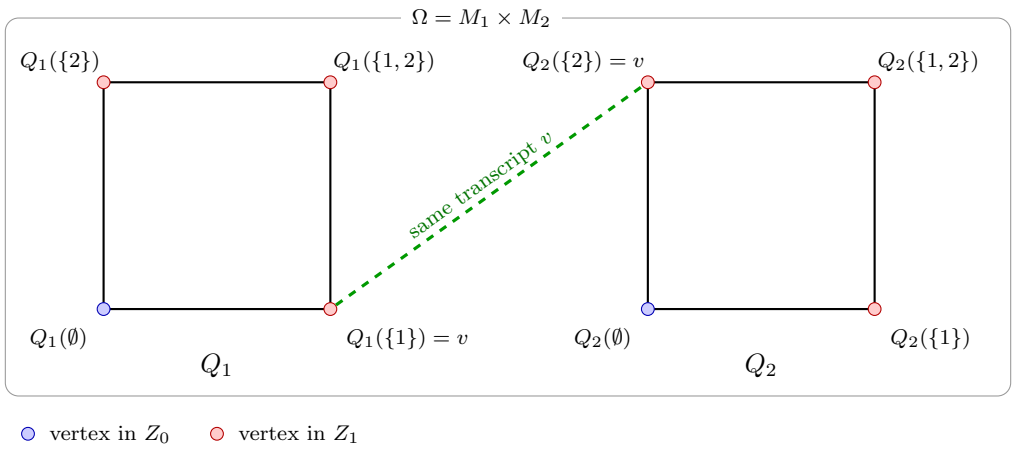


Figure 1: Illustration of an f-cube system for $f = \mathsf{OR}_2$. Vertices in Z_b are transcripts decoded as b. Since $Q_1(\{1\}) = v$ is a 1-vertex, support privacy requires a cube Q_2 with $Q_2(\{2\}) = v$.

2. For $b \in \{0,1\}$, $\forall \boldsymbol{v} \in Z_b$ and $\forall S \subseteq [n]$ such that $f(\chi_S) = b$, there is a cube $Q \in \mathcal{Q}$ with S-corner $\boldsymbol{v}$: $Q(S) = \boldsymbol{v}$.

Figure 1 shows two cubes in an OR_2-cube system. The next lemma states that an f-cube system follows from a PSM protocol for f.

Lemma 13. *A perfectly correct support-private PSM protocol for f with message space $\Omega = \mathcal{M}_1 \times \cdots \times \mathcal{M}_n$ implies an f-cube system with respect to Ω.*

Proof. Let $\mathcal{R}$ be the set of all possible random strings. Consider the set of cubes $\mathcal{Q} = \{\mathtt{Cube}(r) : r \in \mathcal{R}\}$, where

$$\mathtt{Cube}(r) := \{(\mathsf{M}_1(x_1, r), \ldots, \mathsf{M}_n(x_n, r)) \mid x \in \{0,1\}^n\}.$$

Intuitively, the vertices of $\mathtt{Cube}(r)$ correspond to the possible message patterns sent by the parties when the randomness is fixed to r and the input varies over all n-bit strings. Here each cube retains its ordered local endpoints $(\mathsf{M}_i(0,r), \mathsf{M}_i(1,r))$, even when two cubes have the same vertex set.

Let $\mathtt{Cube}_S(r) = \mathsf{M}(\chi_S, r)$ denote the S-corner.

For $b \in \{0,1\}$, define $Z_b \subseteq \Omega$ to be

$$Z_b := \{(\mathsf{M}_1(x_1, r), \ldots, \mathsf{M}_n(x_n, r)) \mid f(x) = b, r \in \mathcal{R}\} = \{\mathtt{Cube}_S(r) \mid f(\chi_S) = b, r \in \mathcal{R}\}.$$

In other words, Z_b is the set of vertices in $\mathtt{Cube}(r)$ corresponding to a message pattern on which Rec outputs b. We claim the result is an f-cube system with respect to Ω.

Note that property 1 follows by definition: for any $S \subseteq [n]$, we have $f(\chi_S) = b \implies \forall r \in \mathcal{R}, \mathtt{Cube}_S(r) \in Z_b$. Perfect correctness implies that Z_0 and Z_1 are indeed disjoint, as required.

Finally, note that for any χ_S and $\boldsymbol{v} = \mathsf{M}(\chi_S, r) \in Z_b$ we have $f(\chi_S) = b$. By support privacy, for any $S' \subseteq [n]$ such that it is also the case that $f(\chi_{S'}) = b$, there exists some choice of randomness r' such that $\mathsf{M}(\chi_{S'}, r') = \boldsymbol{v}$. Since each cube corresponds to some choice of randomness, this is exactly property 2.

Hence, $\mathcal{Q}, Z_0, Z_1$ is indeed an f-cube system with respect to Ω. □

3 Lower Bounds for Symmetric Functions

In this section we prove the following theorem.

Theorem 14. $C_{\textit{indiv-sum}}(\mathsf{OR}_n) \geq (1 - o(1))\, n \log_2 n$.

From this theorem and from Lemma 7 we immediately get the following corollary.

Corollary 15. $C(\mathsf{OR}_n) \geq (1 - o(1)) n \log n$.

Proof of Corollary 15. Recall that C_{indiv} is the logarithm of the largest alphabet size. It is at least the average of the logarithms of all n alphabet sizes, so $C_{\text{indiv}}(\mathsf{OR}_n) \geq C_{\text{indiv-sum}}(\mathsf{OR}_n)/n$, and from Theorem 14 we have that $C_{\text{indiv}}(\mathsf{OR}_n) \geq (1-o(1)) \log_2 n$. Now we apply Lemma 7 with $s = n/\log n$. Fixing all variables outside of a set I to 0 turns OR_n into $\mathsf{OR}_{|I|}$, and for $|I| \geq s$ we have

$$C_{\text{indiv}}(\mathsf{OR}_{|I|}) \geq (1 - o(1)) \log_2 s \geq (1 - o(1)) \log_2 n.$$

So the lemma gives $C(\mathsf{OR}_n) \geq (n - s + 1) \cdot (1 - o(1)) \log_2 n$, and the corollary follows. □

It is not difficult to see that the same technique gives an $(1-o(1))n \log n$ lower bound for AND_n and NOR_n as well.

Similarly, we can prove lower bounds for other symmetric functions.

Corollary 16. *We have the following lower bounds:*

$$C(\mathsf{MAJ}_n) \geq (1 - o(1))n \log n, \quad C(\mathsf{EMAJ}_n) \geq (1 - o(1))n \log n.$$

For $2 \leq q \leq n$ we have

$$C(\mathsf{MOD}_q) \geq \Omega(n \log q).$$

Proof of Corollary 16. Similarly to the previous corollary, from Theorem 14 we have that $C_{\text{indiv}}(\mathsf{OR}_n) \geq (1 - o(1)) \log_2 n$.

Note that MAJ_n contains $\mathsf{OR}_{n/2}$ as a subfunction, EMAJ_n contains $\mathsf{AND}_{n/2}$ as a subfunction, and MOD_q contains NOR_{q-1} as a subfunction. From these facts and Lemma 6, we have

$$\begin{aligned} C_{\text{indiv}}(\mathsf{MAJ}_n) &\geq (1 - o(1)) \log_2 n, \\ C_{\text{indiv}}(\mathsf{EMAJ}_n) &\geq (1 - o(1)) \log_2 n, \\ C_{\text{indiv}}(\mathsf{MOD}_q) &\geq \Omega(\log_2 q). \end{aligned}$$

Similarly to the previous corollary, using Lemma 7, we obtain the desired inequalities. □

The rest of the section is devoted to the proof of Theorem 14.

Consider a perfectly correct, support private PSM for OR_n in which player i has an alphabet $\mathcal{M}_i$ of size k_i. Denote its individual-sum complexity by C, that is,

$$C = \sum_i \log_2 k_i.$$

It will be convenient to denote by $q = \left(\prod_i k_i\right)^{1/n}$ the geometric mean of the alphabet sizes. In particular, $C = n \log_2 q$, so we need a lower bound on q.

Consider the discrete cube $K = \mathcal{M}_1 \times \cdots \times \mathcal{M}_n$ and the Boolean cube structure on it induced by the protocol. Recall that each value of the shared randomness gives a Boolean cube $Q \subseteq K$: once the randomness is fixed, player i sends one of two messages depending on the input bit, and Q consists of the tuples of messages on all 2^n inputs. The two messages of each player are different (otherwise the all-zeros input and an input with a single one would give the same tuple, which contradicts perfect correctness), so Q is a product of n two-element sets. We call these cubes the

cubes of the protocol. For $S \subseteq [n]$, the S-corner of Q is the tuple of messages on the input that has ones exactly in the positions from S.

Denote by Z the support of all 0-inputs in K and by O the support of all 1-inputs. Denote by $U = Z \cup O$ the total support of the protocol in K. The only 0-input is the all-zeros input, so Z is the set of the $\emptyset$-corners of the cubes of the protocol, O is the set of all their other corners, and U is the union of these cubes. By perfect correctness, O and Z are disjoint. By support privacy, all 1-inputs have the same support, so for every $v \in O$ and every nonempty $S \subseteq [n]$ there is a cube of the protocol whose S-corner is v.

The plan of the proof is simple: we bound the fraction of U occupied by Z from above (this is the main part) and from below, and then compare the two bounds.

We start with the upper bound. Among all Boolean cubes of the protocol pick, for each $z \in Z$, one cube Q_z that contains z, and denote by $\mathcal{F}$ the family of the picked cubes. Each cube of the protocol contains exactly one point of Z, namely its $\emptyset$-corner (the other corners are in O), so the cubes Q_z are different for different z. In particular, $|\mathcal{F}| = |Z|$. For each cube Q consider its characteristic function $\chi_Q \colon K \to \mathbb{F}_2$ that for any point of K outputs 1 iff the point is in Q (sometimes it is better to think of χ_Q as a $|K|$-dimensional vector over $\mathbb{F}_2$). Note that the functions χ_Q for $Q \in \mathcal{F}$ are linearly independent. Indeed, their restrictions to Z are the indicators of single points: χ_{Q_z} is equal to 1 on z and to 0 on the rest of Z.

All these functions are zero outside of U, so for now we only get the trivial bound $|Z| = |\mathcal{F}| \le |U|$. The next lemma strengthens this observation and is the main (and maybe the only) ingredient of the proof: the functions remain independent even if we forget their values on a large part of K.

Lemma 17. *For any choice of $a_i \in \mathcal{M}_i$ for each i denote $a = (a_1, \dots, a_n)$ and*

$$K_a = (\mathcal{M}_1 \setminus \{a_1\}) \times \cdots \times (\mathcal{M}_n \setminus \{a_n\}),$$

that is, K_a consists of the points of K that differ from a in all coordinates. Then for every a we have $|\mathcal{F}| = |Z| \le |U \cap K_a|$. As a consequence,

$$|Z| \le |U| \prod_i \left(1 - \frac{1}{k_i}\right). \tag{2}$$

Proof. Consider the discrete cube K. A *line* in K is a set of points in which all coordinates but one are fixed and the unfixed coordinate ranges over all possible values. If the unfixed coordinate is the i-th one, we say that this is a line in direction i.

Note that for any Boolean cube Q the function χ_Q has an even number of ones on each line (more precisely, either 0 or 2). Indeed, Q is a product of two-element sets, so a line in direction i that intersects Q contains exactly two points of Q, one for each of the two possible values of the i-th coordinate. Thus the sum of χ_Q over each line is 0.

The observation behind the lemma is that the values of χ_Q outside of K_a can be reconstructed from its values on K_a. More precisely, the value of χ_Q at any point of K is a linear combination (over $\mathbb{F}_2$) of its values at some points of K_a, and this combination is the same for all cubes Q. To see this, say that the i-th coordinate of a point v is *bad* if $v_i = a_i$, so that K_a consists of the points with no bad coordinates. We use induction on the number of bad coordinates of a point v. If there are none, there is nothing to prove. If the i-th coordinate of v is bad, consider the line through v in direction i. The value of χ_Q at v is equal to the sum of its values at all other points of this line, and each of these points has one bad coordinate fewer than v.

From the observation we can conclude that the functions χ_Q for $Q \in \mathcal{F}$ are independent even when restricted to K_a.[2] Indeed, if some linear combination of these functions is zero on K_a, then by the observation it is zero on the whole K, and this is possible only for the trivial combination. All cubes of the protocol are inside U, so the restricted functions can be nonzero only on $U \cap K_a$. This means that we have $|\mathcal{F}|$ independent vectors in a space of dimension $|U \cap K_a|$, and thus $|Z| = |\mathcal{F}| \le |U \cap K_a|$.

To prove the second inequality we need to find a for which $U \cap K_a$ is small. We can just pick $a_1, \dots, a_n$ one by one in a greedy way, each time removing the maximal possible number of the remaining points from the support. Here the remaining points are the points of U that differ from $a_1, \dots, a_{i-1}$ in the first $i-1$ coordinates, and we pick as a_i the most popular value of the i-th coordinate among them. This removes at least a $1/k_i$ fraction of the remaining points. In the end the remaining points are exactly the points of $U \cap K_a$, and there are at most $|U| \prod_i (1 - 1/k_i)$ of them. Alternatively, we can pick a uniformly at random and use the expectation argument: each point of U gets into K_a with probability $\prod_i (1 - 1/k_i)$, so for some a the size of $U \cap K_a$ is at most its expectation $|U| \prod_i (1 - 1/k_i)$. □

For the lower bound, we need one more simple observation. Consider the whole cube system and consider any line in direction i in K. If it contains a point of U, it must contain a point of Z too. This follows from support privacy: if the line contains a point $v \in O$, then there is a cube Q of the protocol whose $\{i\}$-corner is v. The $\emptyset$-corner of the same cube Q differs from v only in the i-th coordinate, so it is on the same line, and it is in Z.

Now we can count the number of lines. The lines in direction i are disjoint and each of them consists of k_i points. Each line that contains a point of U also contains a point of Z, so there are at most $|Z|$ such lines, and together they have at most $k_i|Z|$ points. Thus $|Z| \ge |U|/k_i$ for every i. Denoting $k = \min_i k_i$ we get

$$|U| \le k|Z|. \tag{3}$$

Combining (2) and (3) we get

$$|U| \le k|Z| \le k|U| \prod_i \left(1 - \frac{1}{k_i}\right).$$

Canceling out $|U|$ and using $1 + x \le e^x$ for $x = -1/k_i$ we get

$$1 \le k \prod_i \left(1 - \frac{1}{k_i}\right) \le \exp\left(\ln k - \sum_i \frac{1}{k_i}\right).$$

Taking logarithms of both sides and using AM-GM on $\frac{1}{k_1}, \dots, \frac{1}{k_n}$ we get

$$\ln k \ge \sum_i \frac{1}{k_i} \ge n \left(\prod_i \frac{1}{k_i}\right)^{1/n} = \frac{n}{q},$$

where recall that $q = \left(\prod_i k_i\right)^{1/n}$. Observing that $k \le q$ (the first one is the minimum of the k_i's, the second one is their geometric mean) we get

$$n \le q \ln q. \tag{4}$$

[2] If this is not intuitive, think of the functions χ_Q as column vectors organized in a matrix; the rows of the matrix correspond to the points of K. We noted above that the columns of the matrix are independent, and thus the matrix rank is equal to the number of columns. The observation gave us that each row corresponding to a point outside of K_a can be expressed as a linear combination of the rows corresponding to points of K_a. Removing these rows does not change the rank, so the columns are still independent.

From this, by the standard argument, we can deduce that $q > n/\ln n$. Indeed, the function $x \ln x$ is increasing for $x \geq 1$, so if $q \leq n/\ln n$, then $q \ln q \leq (n/\ln n) \cdot \ln(n/\ln n) < n$ for $n \geq 3$, which contradicts (4). Thus

$$C = n \log_2 q \geq n \log_2 \frac{n}{\ln n} = n \log_2 n - n \log_2 \ln n = (1 - o(1))\, n \log_2 n,$$

and Theorem 14 is proved.

3.1 Matching Upper Bounds

For OR the matching upper bound was shown by Ishai [Ish13] and is provided in Theorem 12 above.

For EMAJ and for MOD_q with prime q, matching upper bounds follow from variants of the OR protocol in Theorem 12. We include the constructions for completeness.

Theorem 18. *If q is prime, there exists a perfectly correct, perfectly private PSM protocol for $\mathsf{MOD}_{q,n}$ of cost $n\lceil \log q \rceil$.*

Proof. All arithmetic that follows is modulo q. Consider the following protocol.

1. The shared randomness will consist of $r_1, \ldots, r_n$ uniformly random elements of $\mathbb{Z}_q$ such that $\sum_i r_i = 0$ and a uniformly random $s \in \mathbb{Z}_q^*$.
2. The ith party, holding $x_i \in \{0, 1\}$ then computes message $m_i = r_i + x_i \cdot s$.
3. The evaluator then computes $y = \sum_i m_i$ and outputs 1 if and only if $y = 0$.

It is not difficult to see that if $\sum_{i=1}^n x_i \equiv 0 \pmod q$, then the evaluator sees uniformly random messages that sum to 0. On the other hand, if $\sum_{i=1}^n x_i \not\equiv 0 \pmod q$, then the evaluator sees uniformly random messages summing to a uniformly random non-zero element. In this protocol, each party sends a single element from $\mathbb{Z}_q$, yielding an overall communication cost of $n\lceil \log q \rceil$. □

Theorem 19. *There exists a perfectly correct, perfectly private PSM protocol for EMAJ_n of cost $(1 + o(1))n \log n$.*

Proof. Let $p > n$ be the smallest prime greater than n. All arithmetic that follows is modulo p. Consider the following protocol.

1. Sample s uniformly from $\mathbb{Z}_p^*$. Then sample $(r_1, \ldots, r_n)$ uniformly from the tuples in $\mathbb{Z}_p^n$ satisfying $\sum_i r_i = -s\lfloor n/2 \rfloor$.
2. The ith party, holding $x_i \in \{0, 1\}$ then computes message $m_i = r_i + x_i \cdot s$.
3. The evaluator then computes $y = \sum_i m_i$ and outputs 1 if and only if $y = 0$.

The message sum is $s(\|x\| - \lfloor n/2 \rfloor)$. Since $p > n$ and $s \neq 0$, this sum is zero if and only if $\mathsf{EMAJ}_n(x) = 1$. In that case, the transcript is uniform over zero-sum tuples. Otherwise, its sum is uniform over $\mathbb{Z}_p^*$, and the transcript is uniform conditional on that sum. Its distribution therefore depends only on the output. Each party sends one element of $\mathbb{Z}_p$. By Bertrand's postulate $p < 2n$, so the total communication is $(1 + o(1))n \log n$. □

4 Agreement Lemmas

In this section we provide combinatorial framework needed for our alternative approach to lower bounds. The agreement lemmas below show that small local alphabets force compatible colorings on nested sets. Section 5 uses these agreements to produce transcripts with conflicting outputs thus resulting in the lower bounds on communication complexity in PSM setting.

Although the bounds we prove for symmetric functions are not stronger than the bounds obtained by the other approach (and for OR they are weaker), we still provide them, since the technique might be of independent interest.

4.1 Agreement Lemma for OR

The combinatorial lemma we need for OR says that for any finite set of colors, Σ, there is an integer N such that if for each $S \subseteq [N]$, we define a coloring of its members using Σ, then there must exist non-empty $A, B \subseteq [N]$ such that A is a strict subset of B and the B-coloring agrees with the A-coloring on all members of A.

Lemma 20 (Agreement Lemma for OR)**.** *For every finite alphabet Σ there is an integer $N(\Sigma)$, depending only on $|\Sigma|$, such that the following holds for every $n \geq N(\Sigma)$. For every family of functions $\{\alpha_S : S \to \Sigma\}$ indexed by the nonempty sets $S \subseteq [n]$, there exist nonempty sets $T \subsetneq S \subseteq [n]$ with $\alpha_T = \alpha_S|_T$.*

The proof of this lemma in turn relies on variants of two classic results. We present derivations of these variations in Appendix A for completeness.

Lemma 21 (Simplified version of Higman's Lemma)**.** *Let Σ be a finite alphabet. Then there is an integer $m = m(\Sigma)$ such that for every sequence of strings over Σ,*

$$w_1, \ldots, w_m \text{ with } w_i \in \Sigma^i \text{ for each } i \in [m]$$

there must be some $i < j$ such that w_i is a subsequence of w_j.

Theorem 22 (Iterated Finite version of Ramsey's Theorem)**.** *Let Σ be a finite alphabet and $m \geq 1$. Then there is an integer $N(\Sigma, m)$ such that the following holds:*

If for each integer $k \in [1, m]$, every k-element subset of an $N(\Sigma, m)$-element set is colored by a string in Σ^k, then there is an m-element subset H such that for each $k \in [1, m]$, all k-element subsets of H have the same Σ^k-color.

Proof of Lemma 20. Let $m = m(\Sigma)$ be given by Lemma 21, and set $N(\Sigma) = N(\Sigma, m)$ using Theorem 22.

For any k-set $A = \{a_1 < \cdots < a_k\}$ encode its local coloring by the ordered string

$$\chi_k(A) = (\alpha_A(a_1), \ldots, \alpha_A(a_k)) \in \Sigma^k.$$

Now by Theorem 22 there is an m-element set $B = \{b_1 < \cdots < b_m\}$ on which for all $k \in [1, m]$, χ_k is constant; in other words, for each k, all k-subsets of B received the same Σ^k-color. For each k, define $v_k \in \Sigma^k$ to be the common color of χ_k. This gives a sequence of strings

$$v_1, v_2, \ldots, v_m \text{ such that } \forall i, v_i \in \Sigma^i.$$

In particular, $v_i = (\alpha_{B_i}(b_1), \ldots, \alpha_{B_i}(b_i))$ for $B_i = \{b_1, \ldots, b_i\}$.

By Lemma 21, there are $s, t \in [1, m]$ such that $t < s$ and v_t is a subsequence of v_s. Let $1 \leq j_1 < \cdots < j_t \leq s$ be the indices of v_s that realize the subsequence relation $(\forall i \in [t], (v_t)_i = (v_s)_{j_i})$. Consider the two sets

$$\begin{aligned} S &:= \{b_1, \ldots, b_s\} && \text{(the first } s \text{ elements of } B\text{)}, \\ T &:= \{b_{j_1}, \ldots, b_{j_t}\} && \text{(the corresponding subseq. from } S\text{)}. \end{aligned}$$

By homogeneity, the color of T is v_t and the color of S is v_s. By the subsequence relation, these colorings agree on the elements of T, or, in other words,

$$\alpha_T = \alpha_S|_T.$$

□

4.2 Agreement Lemma for MAJ and Friends

For majority, we only need approximate agreement. The combinatorial lemma we need for the majority function says that, for any n, if the number of colors is at most $c_0\sqrt{n}$ for a sufficiently small absolute constant $c_0 > 0$ and n is sufficiently large, then for any family of colorings $\alpha_S : S \to \Sigma$ with $|\Sigma| = k$ and $S \in \binom{[n]}{\geq \lceil n/2 \rceil}$, there must exist nonempty sets $A, B \in \binom{[n]}{\geq \lceil n/2 \rceil}$ such that A is a strict subset of B and the B-coloring disagrees with the A-coloring on at most $p = \lceil n/2 \rceil - 1$ members of A.

Lemma 23 (Agreement Lemma for Majority)**.** *Consider sufficiently large n and denote $p = \lceil \frac{n}{2} \rceil - 1$. Suppose for any alphabet Σ with $|\Sigma| \leq c_0\sqrt{n}$ for a sufficiently small absolute constant $c_0 > 0$ we are given a family of functions*

$$\alpha_S : S \to \Sigma,$$

where $S \in \binom{[n]}{\geq p+1}$. Then there exist non-empty sets $T \in \binom{[n]}{\geq p+1}$ and $S \in \binom{[n]}{\geq p+1}$ such that

1. *$T \subsetneq S$,*
2. *the functions α_T and $\alpha_S|_T$ disagree on at most p points.*

The proof of the lemma relies on Plotkin's bound for error-correcting codes. In it, we argue that if such sets T and S did not exist, we could use the α_S's to construct an error-correcting code whose efficiency violates Plotkin's bound when the alphabet size is small.

Lemma 24 (Plotkin's bound (for k-ary codes) [Qui03])**.** *For a code C with alphabet size k, codeword length n, and distance $d > \left(1 - \frac{1}{k}\right) n$, we have that $|C| \leq \frac{kd}{kd-(k-1)n}$.*

Proof of Lemma 23. Assume that such a pair of sets S and T does not exist and let $k := |\Sigma|$. Let $\ell = \lfloor \frac{p}{2k} \rfloor$ and for all $i = 1, \ldots, \ell$ let $S_i = [p + i]$. Consider the functions $\alpha_{S_i}|_{[p]}$ for all i as vectors $v_i = (\alpha_{S_i}(1), \ldots, \alpha_{S_i}(p)) \in \Sigma^p$. By our assumption, any two of the functions α_{S_i} differ on at least $p+1$ coordinates. Since the compared functions have at most ℓ coordinates outside $[p]$, the vectors v_i differ in at least $p + 1 - \ell$ coordinates. Thus, $\{v_i\}$ forms a code of size ℓ, codeword length p, and distance $d \geq p + 1 - \ell \geq (1 - \frac{1}{2k})p$.

Then, by Plotkin's bound and the fact that distance d cannot exceed codeword length p, we have

$$\ell \leq \frac{kd}{kd - (k-1)p} \leq \frac{kd}{k(1 - \frac{1}{2k})p - (k-1)p} = \frac{kd}{\frac{p}{2}} \leq 2k.$$

Thus, by definition of ℓ we can write

$$\frac{p}{2k} - 1 < \ell \leq 2k$$

and conclude that

$$k = \Omega(\sqrt{n}),$$

contradicting $k \leq c_0\sqrt{n}$ when c_0 is sufficiently small. □

The same code argument gives the following generalization.

Lemma 25 (Generalized Agreement Lemma)**.** *Suppose for integers n and $1 \leq p \leq n$ and an alphabet Σ we are given a family of functions*

$$\alpha_S : S \to \Sigma,$$

where $S \subseteq [n]$. Consider a sequence of sets

$$S_1 \subsetneq S_2 \subsetneq \ldots \subsetneq S_\ell,$$

with $\ell \geq a\sqrt{p}$ for a fixed constant $a > 0$, such that $|S_i| = p + i$. Suppose that for all $i < j$ the functions α_{S_i} and $\alpha_{S_j}|_{S_i}$ disagree on more than p points. Then $|\Sigma| = \Omega_a(\sqrt{p})$.

Proof. Write $k = |\Sigma|$ and suppose it is a sufficiently small constant times $\sqrt{p}$. Then, for sufficiently large p, we have

$$t = \min\{l, \lfloor p/(2k) \rfloor\} > 2k.$$

Apply the code argument from Lemma 23 to the first t sets of the sequence $S_1, \ldots, S_l$, restricting the colorings to a common p-element set. It yields $t \leq 2k$, a contradiction, and hence the claimed alphabet lower bound. □

We use this generalized agreement lemma for EMAJ and MOD_q.

5 Alternative Approach to Lower Bounds

We use the cube-system representation from Section 2.4. For a fixed shared random string $r \in \mathcal{R}$, the cube $\mathtt{Cube}(r)$ has vertex set

$$\{\mathsf{M}(x, r) \mid x \in \{0,1\}^n\} \subseteq \mathcal{M}_1 \times \cdots \times \mathcal{M}_n.$$

For $S \subseteq [n]$, its S-corner is $\mathtt{Cube}_S(r) = \mathsf{M}(\chi_S, r)$. Throughout this section, *cubes* refer to the members of $\{\mathtt{Cube}(r) \mid r \in \mathcal{R}\}$.

5.1 A Superlinear Lower Bound for OR

We devote this section to proving the following theorem, which gives a weaker version of Corollary 15.

Theorem 26. *For any constant K there exists N such that for all $n \geq N$, no (n, K)-protocol $(\mathsf{M}, \mathsf{Rec})$ satisfies perfect correctness and support privacy for OR.*

The proof of Theorem 26 uses the cube-system representation of perfectly correct, support-private PSMs developed in Section 2.4. Perfect correctness implies that in any given cube $\mathtt{Cube}(r)$ in a cube system $\mathcal{Q}$ for OR, exactly one vertex $v \in \mathtt{Cube}(r)$ will be such that $\mathsf{Rec}(v) = 0$. We refer to the unique vertex in a cube that causes the recovery function Rec to output 0 as the *root* of that cube, and our argument exploits constraints on which positions roots may occupy in cubes in $\mathcal{Q}$. As an initial observation, we note that if some $v \in \mathtt{Cube}(r)$ is a root and v is also a vertex in some other cube $\mathtt{Cube}(r')$, then v must also be a root in $\mathtt{Cube}(r')$. Indeed, if it were not, $\mathsf{Rec}(v)$ would output a different value depending on whether the randomness is r or r', which violates perfect correctness. By the same reasoning, if $v \in \mathtt{Cube}(r)$ is a non-root, then v is a non-root in all other cubes in which it appears.

Support privacy implies that for any $b, b' \in \{0,1\}^n \setminus \{0^n\}$ and any $r \in \mathcal{R}$, there exists $r' \in \mathcal{R}$ such that $\mathsf{M}(b, r) = \mathsf{M}(b', r')$. In the language of cubes, each non-root v occurs as an S-corner for every nonempty $S \subseteq [n]$.

For example, $\{\mathtt{Cube}_\emptyset(r) \mid r \in \mathcal{R}\}$ is the set of all roots in the cube system $\mathcal{Q}$.

Proof of Theorem 26. Correctness is clearly infeasible for $K = 1$, so fix any $K \geq 2$. Let $N(K)$ be as promised in Lemma 20.

Assume for the sake of contradiction there exists a perfectly correct support-private PSM protocol for OR_n for $n \geq N(K)$ with $|\mathcal{M}_i| \leq K$.

Fix some $r^* \in \mathcal{R}$ and consider $\mathtt{Cube}(r^*)$. Let v be the particular non-root

$$v := \mathtt{Cube}_{[n]}(r^*).$$

For every nonempty $S \subseteq [n]$, support privacy implies that there exists r_S (and an associated cube $\mathtt{Cube}(r_S)$) such that

$$\mathtt{Cube}_S(r_S) = v.$$

Take $\gamma^S := \mathtt{Cube}_\emptyset(r_S)$ to be the root of this cube. Note that we have

$$\gamma_i^S = v_i \iff i \notin S.$$

Now for each $i \in [n]$, fix any injective (partial) coloring

$$\lambda_i : \mathcal{M}_i \setminus \{v_i\} \to [K-1].$$

Finally, define a coloring of each nonempty S as follows:

$$\text{for } i \in S, \qquad \alpha_S(i) = \lambda_i(\gamma_i^S).$$

This is well defined because of our observation that $\gamma_i^S = v_i \iff i \notin S$.

Applying Lemma 20, we get that there is a particular S and nonempty T such that $T \subsetneq S$ and

$$\alpha_T = \alpha_S|_T.$$

By the injectivity of λ_i's, this implies

$$\forall i \in T, \quad \gamma_i^T = \gamma_i^S.$$

Now consider the $S \setminus T$ corner of $\mathtt{Cube}(r_S)$, namely $\mathtt{Cube}_{S\setminus T}(r_S)$. Because $S \setminus T$ is non-empty, this is a non-root; that is, $\mathsf{Rec}(\mathtt{Cube}_{S\setminus T}(r_S)) = 1$. On the other hand, we will show that $\gamma^T = \mathtt{Cube}_{S\setminus T}(r_S)$, hence $\mathtt{Cube}_{S\setminus T}(r_S)$ is also a root, which contradicts perfect correctness.

To see this final claim, we give a case analysis:

- If $i \in T$, then $i \notin S \setminus T$ and $\mathtt{Cube}_{S\setminus T}(r_S)_i = \gamma_i^S$ but also for $i \in T$ we have $\gamma_i^S = \gamma_i^T$. Hence, $\mathtt{Cube}_{S\setminus T}(r_S)_i = \gamma_i^T$.
- If $i \in S \setminus T$, then $\mathtt{Cube}_{S\setminus T}(r_S)_i = v_i$, but because $i \notin T$ we also have $\gamma_i^T = v_i$.
- Finally, if $i \notin S$, then $i \notin T$ and $\mathtt{Cube}_{S\setminus T}(r_S)_i = \gamma_i^S = v_i = \gamma_i^T$.

□

Observe that by Remark 8 and 9, the hypothesis of Lemma 7 is satisfied. Applying Lemma 7 with $s = \lceil n/2 \rceil$ and any fixed integer B, we lift the lower bound on the individual message complexity $C_{\text{indiv}}(\mathsf{OR})$ given by Theorem 26 to obtain the following lower bound on the PSM communication complexity $C(\mathsf{OR})$:

Corollary 27. *The PSM communication complexity* $C(\mathsf{OR}) = \omega(n)$.

Remark 28. The agreement-lemma proof gives only $\omega(n)$ lower bound for OR_n. We remark that this approach cannot give a substantially better bound. To see why, let $m := m(\Sigma)$ denote the smallest integer such that the condition of Higman's lemma holds; that is, m is the smallest integer such that for every sequence of strings over Σ,

$$w_1, \ldots, w_m \text{ with } w_i \in \Sigma^i \text{ for each } i \in [m]$$

there must be some $i < j$ such that w_i is a subsequence of w_j. It is known that $m(\Sigma)$ grows faster than any primitive recursive function of $|\Sigma|$ [Fri01].

As a result, it is natural to ask if we could prove a substantially sharper version of Lemma 20 directly, without relying on Higman's Lemma. It is easy to see that this is not the case. Call a sequence $w_1, \ldots, w_m$ with $w_i \in \Sigma^i$ bad if it does not have a pair w_i, w_j such that w_i is a subsequence of w_j. Any bad sequence induces a coloring function violating Lemma 20. Indeed, given a bad sequence $w_1, \ldots, w_m$, if $S = \{s_1 < \ldots < s_k\} \subseteq [m]$, then we can define $\alpha_S(s_j) = (w_k)_j$. Then any $T \subsetneq S$ such that $\alpha_T = \alpha_S$ on T would contradict that $w_1, \ldots, w_m$ was bad. Hence, the threshold $N(\Sigma)$ in Lemma 20 also necessarily grows rapidly with $|\Sigma|$.

5.2 Lower Bounds for MAJ and Friends

In this section, we prove the following theorem, which provides an alternative proof of the lower bound for MAJ function.

Theorem 29. *Any* (n, K)*-protocol* $(\mathsf{M}, \mathsf{Rec})$ *that satisfies perfect correctness and support privacy for* MAJ *must have* $K = \Omega(\sqrt{n})$.

For simplicity of notation we only consider the case of odd $n = 2p + 1$. The proof readily generalizes to arbitrary n. Indeed, fixing one input of MAJ_{2p} to zero gives MAJ_{2p-1}; conditioning on an occurring message of the fixed party preserves support privacy and the local alphabets.

For a cube $\mathtt{Cube}(r)$ and an input x with $\|x\| \leq p$, consider the vertex $v = \mathsf{M}(x, r)$. By perfect correctness we have $\mathsf{Rec}(v) = 0$. In this case we call v a *0-vertex* of a cube. Similarly, for x with $\|x\| \geq p + 1$ and $v = \mathsf{M}(x, r)$ we have $\mathsf{Rec}(v) = 1$. These vertices are called 1-*vertices* of the cube. It is still convenient to denote by the *root* of $\mathtt{Cube}(r)$ the vertex $v = \mathsf{M}(0^n, r)$.

Support privacy implies that for any $b, b' \in \{0,1\}^n_{\geq p+1}$ and any $r \in \mathcal{R}$, there exists $r' \in \mathcal{R}$ such that $\mathsf{M}(b, r) = \mathsf{M}(b', r')$. Phrased in the language of cubes, if $v \in \mathtt{Cube}(r)$ is some 1-vertex of cube $\mathtt{Cube}(r)$, for every other possible position of a 1-vertex in a cube, there must exist some other $\mathtt{Cube}(r')$ so that v has that position within the cube $\mathtt{Cube}(r')$.

Proof of Theorem 29. Consider a perfectly correct support-private PSM for MAJ_n with $|\mathcal{M}_i| \le K$.

Fix some $r^* \in \mathcal{R}$ and consider $\mathtt{Cube}(r^*)$. Let v be the vertex that is opposite to the root:

$$v := \mathtt{Cube}_{[n]}(r^*).$$

For every $S \in \binom{[n]}{\ge p+1}$, support privacy implies there exists r_S (and an associated cube $\mathtt{Cube}(r_S)$) such that

$$\mathtt{Cube}_S(r_S) = v.$$

Take $\gamma^S := \mathtt{Cube}_\emptyset(r_S)$ to be the root of this cube. Note that we have

$$\gamma_i^S = v_i \iff i \notin S.$$

Now for each $i \in [n]$, fix any injective (partial) coloring

$$\lambda_i : \mathcal{M}_i \setminus \{v_i\} \to [K-1].$$

Finally, define a coloring of each $S \in \binom{[n]}{\ge p+1}$ as follows:

$$\text{for } i \in S, \qquad \alpha_S(i) = \lambda_i(\gamma_i^S).$$

(This is well defined because of our observation that $\gamma_i^S = v_i \iff i \notin S$.)

Assuming $K \le C\sqrt{n}$ for small enough constant C and applying Lemma 23 we get that there are sets $S, T \subseteq [n]$, each of size at least $p+1$, such that $T \subsetneq S$ and functions α_T and $\alpha_S|_T$ disagree on at most p inputs. By the injectivity of the λ_i's, this implies that

$$\gamma_i^T \ne \gamma_i^S$$

on at most p coordinates $i \in T$. Equivalently, the roots γ^T and γ^S disagree on at most p coordinates of T.

Consider a subset $X \subseteq T$ such that $|T \setminus X| = p$ and $\gamma_i^T = \gamma_i^S$ for all $i \in X$. Note that $\mathtt{Cube}_{T\setminus X}(r_T)$ and $\mathtt{Cube}_{S\setminus X}(r_S)$ are the same point in the message space. At the same time $|T \setminus X| = p$ and $|S \setminus X| \ge p+1$, which contradicts perfect correctness. □

Using Lemma 25, this argument readily generalizes to EMAJ and MOD_q functions. This gives alternative proofs of lower bounds for these functions (in the case of MOD_q with restricted range of parameters).

Theorem 30. *Any (n, K)-protocol $(\mathsf{M}, \mathsf{Rec})$ that satisfies perfect correctness and support privacy for EMAJ must have $K = \Omega(\sqrt{n})$.*

Theorem 31. *Any (n, K)-protocol $(\mathsf{M}, \mathsf{Rec})$ that satisfies perfect correctness and support privacy for MOD_q for $2 \le q \le (1-c)n$ for some constant $c > 0$ must have $K = \Omega(\sqrt{q})$.*

We claim that the proof for MAJ generalizes to symmetric functions f satisfying the following condition: there exist $1 \le p$ and $\ell = \Omega(\sqrt{p})$ with $p + \ell \le n$ such that $f(x) \ne f(y)$ when $\|x\| = p$ and $\|y\| = p + i$ for any $i \in [\ell]$. Use the same root colorings on $S_i = [p+i]$, $i \in [\ell]$, but make the reference transcript $v := \mathtt{Cube}_{S_1}(r^*)$ instead of $\mathtt{Cube}_{[n]}(r^*)$. An agreement except at p coordinates gives coincident corners at weights p and $p + |S \setminus T|$ where $1 \le |S \setminus T| \le \ell$, as in the majority proof. For EMAJ_n, take $p = \lfloor n/2 \rfloor$ and $\ell = n - p$; for $\mathsf{MOD}_{q,n}$, take $p = q$ and $\ell = \min\{q-1, n-q\} = \Omega_c(q)$. Bounded parameters follow from the function being non-constant.

Observe that by Remark 8 and 9, the hypothesis of Lemma 7 is satisfied. Using Lemma 7, we lift the lower bounds on the individual message complexities $C_{\text{indiv}}(\mathsf{MAJ})$, $C_{\text{indiv}}(\mathsf{EMAJ})$, and $C_{\text{indiv}}(\mathsf{MOD}_q)$ given by Theorems 29, 30, and 31 to obtain the following lower bounds on the PSM communication complexity:

Corollary 32. *The functions* MAJ_n *and* EMAJ_n *have communication complexity* $\Omega(n \log n)$*. The function* MOD_q *has communication complexity* $C(\mathsf{MOD}_q) = \Omega(n \log q)$ *when* $2 \leq q \leq (1-c)n$ *for any fixed constant* $c > 0$*.*

For MAJ and EMAJ, use $s = \lceil n/2 \rceil$ and an integer $B = \Theta(\log n)$ below the alphabet lower bound minus one in Lemma 7. For MOD_q with growing q, use $s = \lceil (1-c/2)n \rceil$ and a sufficiently small $B = \Theta(\log q)$: every zero restriction to $m \geq s$ variables satisfies $2 \leq q \leq (1-c')m$ for $c' = c/(2-c) > 0$, and $n - s + 1 = \Omega_c(n)$. Bounded $q \geq 2$ follows from the $\Omega(n)$ correctness bound, since at fixed randomness every relevant party can choose a nonempty message.

5.3 General Lower Bound

In this section we provide the general form of our lower bound approach.

Definition (Consistent Coloring Family)**.** Consider a function $f\colon \{0,1\}^n \to \{0,1\}$, a positive integer K, and a bit $b \in \{0,1\}$. For any $x \in \{0,1\}^n$ consider a coloring α_x of $[n]$ in colors $[K]$. We say that the family of colorings $\{\alpha_x \mid x \in \{0,1\}^n\}$ is a *b-consistent K-coloring family* for f if the following condition holds.

For any two inputs $x, y \in \{0,1\}^n$ such that $f(x) = f(y) = b$ consider the agreement set $A_{xy} = \{i \in [n] \mid \alpha_x(i) = \alpha_y(i)\}$. Then we require that for any z such that $z_i = 1$ implies $i \in A_{xy}$ we have $f(x \oplus z) = f(y \oplus z)$.

Theorem 33. *Consider a function* $f\colon \{0,1\}^n \to \{0,1\}$*. Suppose that for some* K *and* $b \in \{0,1\}$*, there is no* b*-consistent* K*-coloring family for* f*. Then there is no* (n, K)*-protocol for* f *satisfying perfect correctness and support privacy.*

Proof. Suppose there is an (n, K)-protocol for f satisfying perfect correctness and support privacy.

Consider the message space $\mathcal{M}_1 \times \cdots \times \mathcal{M}_n$ with the hypercube structure defined above. Relabel each $\mathcal{M}_i$ injectively in $[K]$. The set of inputs with output b is nonempty, since otherwise every coloring family is b-consistent.

Consider $b \in \{0,1\}$ from the statement of the theorem and consider a b-vertex v in the message space, that is $v \in Z_b$. By support privacy, for any S with $f(\chi_S) = b$ there is a cube $\mathtt{Cube}(r_S)$ such that $\mathtt{Cube}_S(r_S) = v$. Define a coloring $\alpha_{\chi_S}(i) = \mathtt{Cube}_{\overline{S}}(r_S)_i$, and extend the family arbitrarily to inputs with $f(x) \neq b$.

By the hypothesis in the theorem this coloring cannot be b-consistent. That is, there are inputs $x, y \in \{0,1\}^n$ with $f(x) = f(y) = b$ and there is $S \subseteq A_{xy}$ such that $f(x \oplus \chi_S) \neq f(y \oplus \chi_S)$. Consider sets S_x and S_y such that $\chi_{S_x} = x$ and $\chi_{S_y} = y$. In cubes $\mathtt{Cube}(r_{S_x})$ and $\mathtt{Cube}(r_{S_y})$ we have

$$\mathtt{Cube}_{S_x}(r_{S_x}) = \mathtt{Cube}_{S_y}(r_{S_y}) = v.$$

Then by the coloring agreement we also have

$$\mathtt{Cube}_{S_x \triangle S}(r_{S_x}) = \mathtt{Cube}_{S_y \triangle S}(r_{S_y}),$$

which is a contradiction with perfect correctness since

$$f(\chi_{S_x \triangle S}) = f(x \oplus \chi_S) \neq f(y \oplus \chi_S) = f(\chi_{S_y \triangle S}).$$

□

We next show that the technique outlined in Sections 4 and 5 subsumes the Nechiporuk bound of [BHI+20].

Theorem 34 (Nechiporuk Bound)**.** *Consider a function $f\colon \{0,1\}^n \to \{0,1\}$ and $S \subseteq [n]$. Denote by $N_S(f)$ the number of subfunctions of f on S. Then every perfectly correct, support-private protocol for f with local alphabet sizes K_i satisfies $\sum_{i\in S} \log K_i \geq \log N_S(f) - 1$.*

Before proving Theorem 34 we observe that the following corollary is immediate.

Corollary 35. *Consider a function $f\colon \{0,1\}^n \to \{0,1\}$ and let $V = (V_1, \ldots, V_m)$ be a partition of $[n]$, that is $[n] = \bigsqcup_{i=1}^m V_i$. Then*

$$C_{\textit{indiv-sum}}(f) \geq \sum_{i=1}^{m} \log N_{V_i}(f) - m.$$

To address the communication metric in Theorem 34, we analyze the communication cubes directly rather than apply Theorem 33.

Proof of Theorem 34. Denote the size of S by k and for notational convenience assume that $S = [k]$.

Without loss of generality assume that at least half of the subfunctions of f on S output 0 on 0-input. Denote the set of these subfunctions by $\mathcal{F}_S$. In particular, $|\mathcal{F}_S| \geq N_S(f)/2$. For each function $g \in \mathcal{F}_S$ fix an input $z \in \{0,1\}^{n-k}$ on $\overline{S}$ such that $g = f\,|_{\overline{S}=z}$. Denote the set of all of these inputs by $Z \subseteq \{0,1\}^{n-k}$. In particular, $|Z| = |\mathcal{F}_S| \geq N_S(f)/2$.

In the communication cube consider an arbitrary supported 0-transcript $v \in Z_0$. By support privacy for any 0-input $x \in \{0,1\}^n$ of f there is a cube that has v as an x-corner. For each $z \in Z$ consider input $(\vec{0}, z)$ to f (note that it is a 0-input) and fix such a cube C_z.

Now for each $z \in Z$ consider the input $(\vec{1}, z)$. For the sake of contradiction suppose $\prod_{i\in S} K_i < N_S(f)/2$. Then there are fewer than $N_S(f)/2$ possible tuples of messages of these players. Thus, for two inputs in Z these transcripts are the same. Denote these inputs by z_1 and z_2 and their corresponding functions in $\mathcal{F}_S$ by g_1 and g_2.

There is an input $y \in \{0,1\}^k$ such that $g_1(y) \neq g_2(y)$. As a result,

$$f(y, z_1) = g_1(y) \neq g_2(y) = f(y, z_2).$$

At the same time, by the choice of z_1 and z_2 the corner (y, z_1) in C_{z_1} and the corner (y, z_2) in C_{z_2} are the same vertex in the communication cube. This is a contradiction with the perfect correctness. □

6 Upper Bounds

6.1 Perfect Privacy for any Symmetric Function in $\widetilde{O}(n^{3/2})$ bits

In this section, we prove the following theorem.

Theorem 36. *For any symmetric Boolean function $f\colon \{0,1\}^n \to \{0,1\}$ there is a perfectly correct, perfectly private PSM of total complexity $\widetilde{O}(n^{3/2})$.*

Proof. We compose an outer protocol, which evaluates the function from residues of the Hamming weight, with inner protocols that privately compute the outer messages. This follows Eriguchi [Eri26], who composes an outer PSM on histogram coordinates with a private selector for each outer message. We use small CRT residues of the Hamming weight as the outer inputs, which reduces the size of the selector tables.

Write $f(x) = g(\|x\|)$ for $g : \{0, \ldots, n\} \to \{0,1\}$.

The outer protocol uses Theorem 10 to evaluate g on the weight reconstructed from its residues modulo small primes. We assume n is sufficiently large. (For the finitely many smaller n, we can use any universal perfectly private PSM.)

Let $p_1, \ldots, p_k$ be distinct prime numbers and let $Q = \max_i p_i$. We want the following properties of this set of primes:

1. $p_1 \ldots p_k > n$,
2. $Q^k \leq n\mathsf{polylog}(n)$.

It is not hard to see that these properties are achievable with $k, Q = \mathsf{polylog}(n)$ (see Lemma 44 in Appendix A.3).

Using Chinese Remainder Theorem we can define a function

$$\mathsf{CRT}\colon \mathbb{Z}_{p_1} \times \ldots \times \mathbb{Z}_{p_k} \to \{0, 1, \ldots, n\},$$

such that $z = \mathsf{CRT}(y_1, \ldots, y_k)$ whenever $z \in \{0, \ldots, n\}$ satisfies $z = y_i \pmod{p_i}$ for all i; set the value to 0 if no such z exists.

The outer protocol will be computing the function

$$h = g \circ \mathsf{CRT}\colon \mathbb{Z}_{p_1} \times \ldots \times \mathbb{Z}_{p_k} \to \{0, 1\}.$$

By Theorem 10, extending this function arbitrarily to $[Q]^k$, there is a protocol for this function, in which k players with inputs $y_1, \ldots, y_k$ use shared randomness R and send messages $m_1(y_1, R), \ldots, m_k(y_k, R)$ to the referee, who is then able to compute $h(y_1, \ldots, y_k)$ without learning any other information about $y_1, \ldots, y_k$. Encode each outer message at a fixed length by rounding up the logarithm of its alphabet size. If we denote the length of the message m_i by l_i, then for this protocol we have

$$l_1 + \ldots + l_k = O(k^3 Q^{k/2}) = \widetilde{O}(\sqrt{n}).$$

However, in our setting we have n players that have inputs $x_1, \ldots, x_n \in \{0, 1\}$. The obstacle is that no individual player knows the residues $y_i = \sum_j x_j \pmod{p_i}$.

Each inner protocol lets the referee recover one outer message $m_i(y_i, R)$ without revealing the residue y_i. Each invocation uses fresh inner randomness, independent of R and the other inner invocations. The referee then applies the outer decoder to these messages.

For the inner protocol we arrive at the following more general problem. We have n players holding private bits $x_1, \ldots, x_n$, and also sharing a secret function $T\colon \mathbb{Z}_p \to \{0, 1\}^l$ (in the outer protocol, $T = m_i(\cdot, R)$, where R is unknown to the referee). The goal of the players is to reveal $T\left(\sum_i x_i \pmod{p}\right)$ to the referee without revealing any other information.

We now describe the inner protocol. The players sample independent uniform elements $s_1, \ldots, s_n \in \mathbb{Z}_p$ to mask the input bits. Denote $s = \sum_i s_i \pmod{p}$. The players also generate the family of uniformly random functions

$$R_i\colon \mathbb{Z}_p \to \{0, 1\}^l$$

for $i = 1, \ldots, n-1$.

The first player sends $a_1 = x_1 + s_1$ and $R_1(a_1)$.

The i-th player for $i = 2, \ldots, n-1$ sends $a_i = x_i + s_i$ as well as the whole table of the function $Q_i(y) = R_{i-1}(y - a_i) \oplus R_i(y)$.

The last player sends $a_n = x_n + s_n$ and the whole table of the function $Q_n(y) = R_{n-1}(y - a_n) \oplus T(y - s)$.

Each message depends only on the sender's input bit, the shared function T, and the shared randomness, so the parties can compute their messages simultaneously.

The referee computes

$$R_1(a_1) \oplus Q_2(a_1 + a_2) \oplus Q_3(a_1 + a_2 + a_3) \oplus \ldots \oplus Q_n(a_1 + \ldots + a_n),$$

which after expansion simplifies to

$$T(a_1 + \ldots + a_n - s) = T(x_1 + \ldots + x_n).$$

To verify perfect privacy, fix the input bits and the table T. Conditional on the masked inputs $a_1, \ldots, a_n$, the map from $(R_1, \ldots, R_{n-1})$ to $(Q_2, \ldots, Q_n)$ is a bijection: recover R_{n-1} from Q_n and T, then recover the remaining R_i backwards. Hence the Q_i are independent uniform tables. Given only the output $t = T(\sum_i x_i)$, a simulator samples the masked inputs and tables uniformly and sets the first party's remaining message to

$$b = t \oplus \bigoplus_{i=2}^{n} Q_i(a_1 + \cdots + a_i).$$

The telescoping identity above gives $b = R_1(a_1)$ in the real protocol, so this simulation has exactly the real transcript distribution. Composing it with the outer simulator proves perfect privacy of the full protocol.

For communication, sending all masked inputs a_i requires at most $n\lceil \log p \rceil$ bits of communication and sending the function tables requires npl bits of communication. In total, to implement the protocol for one function T they send at most $n\lceil \log p \rceil + npl$ bits.

Applying the inner protocol to all messages $m_1, \ldots, m_k$ of the outer protocol gives total communication at most $nk\lceil \log Q \rceil + nQ \sum_i l_i = \widetilde{O}(n^{3/2})$ bits. □

6.2 Support-Private Protocols for Symmetric Functions

In this section, we prove the following theorem.

Theorem 37. *For any symmetric Boolean function $f\colon \{0,1\}^n \to \{0,1\}$ there is a perfectly correct, support-private PSM of total complexity $\widetilde{O}(n)$.*

We use the same composition as in Section 6.1, replacing the outer protocol with the support-private protocol of Theorem 11.

Proof. As in the perfectly private construction, it suffices to handle large enough n. For these large enough n, write $f(x) = g(\|x\|)$ and choose the distinct primes from Lemma 44, so that the following properties hold.

1. $p_1 \ldots p_k > n$,
2. $Q^k \leq n\mathsf{polylog}(n)$

where $Q = \max_i p_i$. As in that construction, we use the Chinese Remainder Theorem to define a function

$$\mathsf{CRT}\colon \mathbb{Z}_{p_1} \times \ldots \times \mathbb{Z}_{p_k} \to \{0, 1, \ldots, n\},$$

that returns the unique representative in $\{0, \ldots, n\}$ when one exists, and 0 otherwise. Let CRT' denote a version of this function with binarized inputs, assigning 0 to invalid residue encodings. In other words, define

$$\mathsf{CRT}' \colon \{0,1\}^{\lceil \log p_1 \rceil} \times \ldots \times \{0,1\}^{\lceil \log p_k \rceil} \to \{0, 1, \ldots, n\}.$$

The outer protocol will be computing the function

$$h = g \circ \mathsf{CRT}' \colon \{0,1\}^{\lceil \log p_1 \rceil} \times \ldots \times \{0,1\}^{\lceil \log p_k \rceil} \to \{0,1\}.$$

Now since $Q^k \leq n\mathsf{polylog}(n)$, we have that $L = \sum_{i=1}^{k} \lceil \log p_i \rceil = O(\log n)$. By Theorem 11, there is a protocol for h with communication $\widetilde{O}(L^2) = \widetilde{O}(\log^2 n)$. The protocol provided by Theorem 11 will contain L messages, since it assumes each party holds only one bit. First give each virtual message a fixed-length encoding, with a length prefix if needed; the possible factor L is polylogarithmic. However, we will concatenate components associated with the same prime residue to obtain k messages $m_1(y_1, R), \ldots, m_k(y_k, R)$, where $m_i(y_i, R)$ is a concatenation of the messages for all the players whose bit encodes the value of the residue y_i.

As in the proof of Theorem 36, we cannot just use the protocol above because the players do not individually know the inputs $y_i = \sum_j x_j \pmod{p_i}$. Instead of implementing the outer protocol above explicitly, our players will use an inner protocol to compute messages m_i for all i. More specifically, they will send the information to the referee to compute each message m_i (by inner protocol), and after that the referee will be able to use the outer protocol to compute $h(y_1, \ldots, y_k)$. Letting ℓ_i denote the length of $m_i(y_i, R)$, we have from the upper bound on the outer protocol that $\sum_i \ell_i = \mathsf{polylog}(n)$. The inner protocol from Theorem 36 suffices to send the information for all messages $m_1, \ldots, m_k$, so we omit the details here. Use independent inner randomness for these invocations. For an outer transcript m, let $K(m)$ denote the distribution of its inner simulations. The full transcript support is the union of $\operatorname{Supp} K(m)$ over possible outer transcripts m; equal outer supports therefore imply equal full supports. This gives us an upper bound of $nk\lceil \log Q \rceil + nQ \sum_i \ell_i = \widetilde{O}(n)$ bits of communication. □

6.3 Upper Bound for Majority

Combining a known permutation branching program for MAJ_n with a standard PSM construction gives communication $O(n \log^3 n \log \log n)$.

Definition (strict permutation branching program)**.** A *strict permutation branching program* of width w and length ℓ consists of the following:

1. A sequence of variable indices $i_1, \ldots, i_\ell$.
2. For each layer j and each bit b, a permutation $P_j(b)$ of $\{1, \ldots, w\}$.
3. A start state s_0.
4. An accepting set $A \subseteq \{1, \ldots, w\}$.

On input x it outputs 1 if and only if the image of s_0 under permutation $P_\ell(x_{i_\ell}) \ldots P_1(x_{i_1})$ is in A.

The following lemma is implicit in Klauck and Podder [KP14a] (the proof can be found in Appendix 10.5 of the full version of the paper [KP14b]).

Lemma 38 (Klauck and Podder [KP14a])**.** *The function MAJ_n can be computed by a strict permutation branching program of width $O(\log n)$ and length $O(n \log^2 n)$.*

The next lemma summarizes the standard conversion of permutation branching programs into perfectly private PSM protocols. The construction goes back to Kilian [Kil88]; see also [Ish13, Section 4.1.3] for its formulation over general groups. Papakonstantinou and Yang [PY14] give the permutation version for width 5; the same construction generalizes to arbitrary width. We give the details in Appendix A.4.

Lemma 39 (Kilian [Kil88])**.** *If f is computable by a strict permutation branching program of width w and length ℓ, with a fixed initial state and a fixed final accepting set, then there is a perfectly private simultaneous messages protocol for f with total communication $O(w\ell \log w)$.*

These two facts readily give us the following corollary.

Corollary 40. *There is a perfectly private simultaneous messages protocol for MAJ_n with total communication $O(n \log^3 n \log\log n)$.*

References

[AIK06] Benny Applebaum, Yuval Ishai, and Eyal Kushilevitz. Cryptography in NC^0. *SIAM Journal on Computing*, 36(4):845–888, 2006.

[AL21] Léonard Assouline and Tianren Liu. Multi-party PSM, revisited: Improved communication and unbalanced communication. In *Theory of Cryptography: 19th International Conference, TCC 2021, Raleigh, NC, USA, November 8–11, 2021, Proceedings, Part II 19*, pages 194–223. Springer, 2021.

[App17] Benny Applebaum. Garbled circuits as randomized encodings of functions: a primer. In *Tutorials on the Foundations of Cryptography: Dedicated to Oded Goldreich*, pages 1–44. Springer, 2017.

[BF07] Amos Beimel and Matthew K. Franklin. Weakly-private secret sharing schemes. In *Theory of Cryptography – TCC 2007*, volume 4392 of *Lecture Notes in Computer Science*, pages 253–272. Springer, 2007.

[BHI+20] Marshall Ball, Justin Holmgren, Yuval Ishai, Tianren Liu, and Tal Malkin. On the complexity of decomposable randomized encodings, or: How friendly can a garbling-friendly prf be? In *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*. Schloss-Dagstuhl-Leibniz Zentrum für Informatik, 2020.

[BHR12] Mihir Bellare, Viet Tung Hoang, and Phillip Rogaway. Foundations of garbled circuits. In *Proceedings of the 2012 ACM conference on Computer and communications security*, pages 784–796, 2012.

[BIKL25] Amos Beimel, Yuval Ishai, Eyal Kushilevitz, and Hanjun Li. Cryptography with weak privacy. Cryptology ePrint Archive, Paper 2025/1978, 2025.

[BKN18] Amos Beimel, Eyal Kushilevitz, and Pnina Nissim. The complexity of multiparty PSM protocols and related models. In Jesper Buus Nielsen and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part II*, volume 10821 of *Lecture Notes in Computer Science*, pages 287–318. Springer, 2018.

[BR22] Marshall Ball and Tim Randolph. A note on the complexity of private simultaneous messages with many parties. In Dana Dachman-Soled, editor, *3rd Conference on Information-Theoretic Cryptography, ITC 2022, Cambridge, MA, USA, July 5-7, 2022*, volume 230 of *LIPIcs*, pages 7:1–7:12. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[CK90] Christian Choffrut and Juhani Karhumäki. Words, automata, grammars. In *Handbook of Theoretical Computer Science, Volume B: Formal Models and Semantics*, pages 343–456. Elsevier and MIT Press, 1990.

[Die17] Reinhard Diestel. *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer Publishing Company, 5th edition, 2017.

[Dus10] Pierre Dusart. Estimates of some functions over primes without rh, 2010.

[Eri26] Reo Eriguchi. On the communication complexity of PSM and CDS for symmetric functions. In *Advances in Cryptology – EUROCRYPT 2026*, volume 16543 of *Lecture Notes in Computer Science*, pages 155–184. Springer, 2026.

[ES25] Reo Eriguchi and Kazumasa Shinagawa. Efficient multiparty private simultaneous messages for symmetric functions. In *Advances in Cryptology – EUROCRYPT 2025*, volume 15605 of *Lecture Notes in Computer Science*, pages 240–269. Springer, 2025.

[FKN94] Uriel Feige, Joe Kilian, and Moni Naor. A minimal model for secure computation (extended abstract). In Frank Thomson Leighton and Michael T. Goodrich, editors, *Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing, 23-25 May 1994, Montréal, Québec, Canada*, pages 554–563. ACM, 1994.

[Fri01] Harvey M. Friedman. Long finite sequences. *Journal of Combinatorial Theory, Series A*, 95(1):102–144, 2001.

[GRS15] Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer. *Ramsey Theory*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Hoboken, NJ, 2nd edition, 2015.

[Hig52] Graham Higman. Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society*, 3(1):326–336, 1952.

[HW26] Justin Holmgren and Kewen Wu. Improved lower bounds for decomposable randomized encoding. *CoRR*, abs/2609.18020, 2026.

[IK97] Yuval Ishai and Eyal Kushilevitz. Private simultaneous messages protocols with applications. In *Proceedings of the Fifth Israeli Symposium on Theory of Computing and Systems*, pages 174–183. IEEE, 1997.

[IK02] Yuval Ishai and Eyal Kushilevitz. Perfect constant-round secure computation via perfect randomizing polynomials. In *Automata, Languages and Programming: 29th International Colloquium, ICALP 2002 Málaga, Spain, July 8–13, 2002 Proceedings 29*, pages 244–256. Springer, 2002.

[Ish13] Yuval Ishai. Randomization techniques for secure computation. In *Secure Multi-Party Computation*, pages 222–248. IOS press, 2013.

[Kil88] Joe Kilian. Founding cryptography on oblivious transfer. *Proceedings of the twentieth annual ACM symposium on Theory of computing*, 1988.

[Kőn27] Dénes Kőnig. Über eine Schlussweise aus dem Endlichen ins Unendliche. *Acta Scientiarum Mathematicarum (Szeged)*, 3(2–3):121–130, 1927.

[KP14a] Hartmut Klauck and Supartha Podder. New bounds for the garden-hose model. In Venkatesh Raman and S. P. Suresh, editors, *34th International Conference on Foundation of Software Technology and Theoretical Computer Science, FSTTCS 2014, New Delhi, India, December 15-17, 2014*, volume 29 of *LIPIcs*, pages 481–492. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2014.

[KP14b] Hartmut Klauck and Supartha Podder. New bounds for the garden-hose model. *CoRR*, abs/1412.4904, 2014.

[Nui26] Koji Nuida. Improved private simultaneous messages protocols for symmetric functions with universal reconstruction. In *Advances in Information and Computer Security – IWSEC 2025*, volume 16208 of *Lecture Notes in Computer Science*, pages 229–246. Springer, 2026.

[PY14] Periklis A. Papakonstantinou and Guang Yang. Cryptography with streaming algorithms. In Juan A. Garay and Rosario Gennaro, editors, *Advances in Cryptology - CRYPTO 2014 - 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part II*, volume 8617 of *Lecture Notes in Computer Science*, pages 55–70. Springer, 2014.

[Qui03] Jörn Quistorff. Some remarks on the plotkin bound. *Electron. J. Comb.*, 10, 2003.

[Ram87] Frank P Ramsey. On a problem of formal logic. In *Classic Papers in Combinatorics*, pages 1–24. Springer, 1987.

[SESN23] Kazumasa Shinagawa, Reo Eriguchi, Shohei Satake, and Koji Nuida. Private simultaneous messages based on quadratic residues. *Designs, Codes and Cryptography*, 91(12):3915–3932, 2023.

[SN24] Kazumasa Shinagawa and Koji Nuida. Explicit lower bounds for communication complexity of PSM for concrete functions. In *Progress in Cryptology – INDOCRYPT 2023*, volume 14460 of *Lecture Notes in Computer Science*, pages 45–61. Springer, 2024.

[Yao86] Andrew Chi-Chih Yao. How to generate and exchange secrets. In *27th annual symposium on foundations of computer science (Sfcs 1986)*, pages 162–167. IEEE, 1986.

A Derivation of Combinatorial Results

A.1 Simple Higman's Lemma Proof

For strings u, v, write $u \preceq v$ when u is a subsequence of v.

We first recall Higman's Lemma (for finite alphabets and ordered by equality).

Lemma 41 (Higman's Lemma [Hig52,CK90])**.** *If Σ is a finite alphabet, then every infinite sequence of finite strings*

$$w_1, w_2, w_3, \ldots$$

contains $i < j$ such that $w_i \preceq w_j$.

We use König's lemma to derive a finite bound from Higman's infinite-sequence theorem.

Lemma 42 (König's Lemma [Kőn27]; see also Lemma 8.1.2 in [Die17])**.** *Let $\mathcal{T}$ be an infinite rooted tree in which every vertex has finitely many children. Then $\mathcal{T}$ contains an infinite path starting at the root.*

The finite variant follows by applying König's lemma to the tree of finite bad sequences.

Proof of (Simple) Higman's Lemma, Lemma 21. Call a sequence $w_1, \ldots, w_\ell$ such that $|w_i| = i$ "bad" if for all $i < j$, $w_i \not\preceq w_j$.

Suppose the lemma is false, so a bad sequence of length ℓ exists for every ℓ.

Note that if a sequence is "bad," any prefix of that sequence is also "bad."

Thus we can build an arbitrarily deep tree of "bad" sequences by letting vertices at depth ℓ be "bad" sequences of length ℓ, $w_1, \ldots, w_\ell$, with parent being the $(\ell-1)$-length prefix, $w_1, \ldots, w_{\ell-1}$ (which is also "bad"). Note that this tree is finitely branching because a node at depth ℓ can only have $|\Sigma^{\ell+1}|$ many children (because we required $|w_i| = i$). Hence by Lemma 42, $\mathcal{T}$ contains an infinite path from which we can derive an infinite "bad" sequence, a sequence of strings

$$w_1, w_2, w_3, \ldots \qquad (\text{with } |w_i| = i, \forall i)$$

such that there does not exist $i < j$ such that $w_i \preceq w_j$, contradicting Lemma 41. □

A.2 Iterated Finite Ramsey

We first recall the finite Ramsey theorem.

Theorem 43 (Finite Ramsey's Theorem [Ram87, GRS15])**.** *For all positive integers k, m, r there is an integer $R_k(m; r)$ such that whenever the k-subsets of a set X of size at least $R_k(m; r)$ are colored with r colors, then there is a subset*

$$Y \subseteq X, \qquad |Y| = m,$$

such that all k-element subsets of Y have the same color.

The desired form follows by induction.

Proof of Theorem 22. Let $r_k = |\Sigma|^k$ be the number of possible colors for k-sets. Define

$$N_m := m, N_{k-1} := R_k(N_k; r_k) \text{ for } k = m, m-1, \ldots, 1.$$

Take $N(\Sigma, m) = N_0$. Now, we iteratively invoke Theorem 43. Start with X_0 a set of size N_0. Then inductively, at step k we have a set X_{k-1} of size $N_{k-1} = R_k(N_k; r_k)$ so by Theorem 43 we have that there must be some

$$X_k \subseteq X_{k-1}, \qquad |X_k| = N_k \text{ where all } k\text{-sets have the same color } u_k \in \Sigma^k.$$

Moreover, by the inductive hypothesis, for ℓ such that $1 \leq \ell \leq k-1$, ℓ-subsets of X_k have the same color because X_k is a subset of X_{k-1}. Hence ℓ-subsets of X_k are the same color for any $1 \leq \ell \leq k$.

Thus, we have constructed a tower of sets

$$X_0 \supseteq X_1 \supseteq \cdots \supseteq X_m, \qquad |X_k| = N_k,$$

where $H = X_m$ has size $N_m = m$ and for any $k \leq m$, all k-subsets have the same color. □

A.3 Choice of Small Primes

Lemma 44. *For all sufficiently large n, there are distinct primes $p_1, \dots, p_k$, with $Q = \max_i p_i$, such that $p_1 \cdots p_k > n$, $Q^k = O(n \log^3 n)$, and $k, Q = \mathsf{polylog}(n)$.*

Proof. Put $t = \lceil \log n \rceil$ and $X = t^3$. Standard prime-counting estimates, for example [Dus10, Theorem 6.9], give $\pi(2X) - \pi(X) = \Omega(X/\log X)$. Partition $(X, 2X]$ into t intervals of length X/t. One interval contains $\Omega(X/(t \log X))$ primes, more than $\lfloor \log n / \log X \rfloor + 1$ for sufficiently large n. Select primes in this interval until their product P first exceeds n. Then

$$k \le \lfloor \log n / \log X \rfloor + 1, \qquad Q \le 2X, \qquad P \le nQ.$$

Since all selected primes lie in one such interval, $Q/p_i \le 1 + 1/t$, and hence

$$Q^k = P \prod_i (Q/p_i) \le nQ(1 + 1/t)^k = O(nX).$$

□

A.4 Permutation Branching Program Conversion

Proof of Lemma 39. Write the transitions as permutations $P_j(x_{i_j})$ of $[w]$, with initial state s_0 and accepting set A. Sample independent uniform permutations $\rho_0, \dots, \rho_\ell$. The owner of x_{i_j} sends $E_j = \rho_j P_j(x_{i_j}) \rho_{j-1}^{-1}$; one party also sends $\alpha = \rho_0(s_0)$ and $A' = \rho_\ell(A)$. The evaluator tests whether $E_\ell \cdots E_1(\alpha) \in A'$, giving perfect correctness.

Put $P = P_\ell \cdots P_1$ and $E = E_\ell \cdots E_1$. The map from the masks to $(\rho_0, E_1, \dots, E_\ell)$ is bijective, so the E_j and ρ_0 are independent uniform permutations. Conditional on the E_j, the pair

$$(\alpha, E^{-1}A') = (\rho_0(s_0), \rho_0(P^{-1}A))$$

is a uniformly relabeled point and set, whose law depends only on whether $P(s_0) \in A$. A simulator therefore samples the E_j and α uniformly, then a uniform $|A|$-element set A' containing $E(\alpha)$ exactly when the output is one. Communication is

$$\ell \lceil \log(w!) \rceil + \lceil \log w \rceil + w = O(w\ell \log w).$$

□

AI usage disclosure. Work on this project began in spring 2023. The authors made initial progress on the complexity of OR_n for small constant n and suspected that extremal combinatorics might provide the key to stronger bounds. On April 29, 2026, ChatGPT-5.5 Pro was used to find a viable superlinear lower bound for the OR function (Theorem 26), indeed using classic results by Ramsey and Higman. The authors subsequently generalized and extended the lower bound approach without AI. Upper bounds were obtained with partial AI assistance, exploring branching program results (Section 6.3), and parameter verification in composition constructions (i.e., Section 6.1 and Section 6.2). Finally, ChatGPT-6.2 Astra was used in November 2026 to obtain the tight lower bound for the OR function, presented in Section 3.

AI tools were used for proof verification and editorial assistance. The authors take full responsibility for the contents.

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il