The Power of Shared Randomness in Uncertain Communication*

Badih Ghazi[1] and Madhu Sudan[2]

1 **MIT, CSAIL, Cambridge, MA, USA**
badih@mit.edu
2 **Harvard John A. Paulson School of Engineering and Applied Sciences, Cambridge, MA, USA**
madhu@cs.harvard.edu

Abstract

In a recent work (Ghazi et al., SODA 2016), the authors with Komargodski and Kothari initiated the study of *communication with contextual uncertainty*, a setup aiming to understand how efficient communication is possible when the communicating parties imperfectly share a huge context. In this setting, Alice is given a function f and an input string x, and Bob is given a function g and an input string y. The pair (x, y) comes from a known distribution μ and f and g are guaranteed to be close under this distribution. Alice and Bob wish to compute $g(x, y)$ with high probability. The lack of agreement between Alice and Bob on the function that is being computed captures the uncertainty in the context. The previous work showed that any problem with one-way communication complexity k in the standard model (i.e., without uncertainty[1]) has *public-coin* communication at most $O(k(1 + I))$ bits in the uncertain case, where I is the mutual information between x and y. Moreover, a lower bound of $\Omega(\sqrt{I})$ bits on the public-coin uncertain communication was also shown.

However, an important question that was left open is related to the power that public randomness brings to uncertain communication. Can Alice and Bob achieve efficient communication amid uncertainty without using public randomness? And how powerful are public-coin protocols in overcoming uncertainty? Motivated by these two questions:

- We prove the first separation between private-coin uncertain communication and public-coin uncertain communication. Namely, we exhibit a function class for which the communication in the standard model and the public-coin uncertain communication are $O(1)$ while the private-coin uncertain communication is a growing function of n (the length of the inputs). This lower bound (proved with respect to the uniform distribution) is in sharp contrast with the case of public-coin uncertain communication which was shown by the previous work to be within a constant factor from the certain communication. This lower bound also implies the first separation between public-coin uncertain communication and deterministic uncertain communication. Interestingly, we also show that if Alice and Bob *imperfectly share* a sequence of random bits (a setup weaker than public randomness), then achieving a constant blow-up in communication is still possible.
- We improve the lower-bound of the previous work on public-coin uncertain communication. Namely, we exhibit a function class and a distribution (with mutual information $I \approx n$) for which the one-way certain communication is k bits but the one-way public-coin uncertain communication is at least $\Omega(\sqrt{k} \cdot \sqrt{I})$ bits.

Our proofs introduce new problems in the standard communication complexity model and prove lower bounds for these problems. Both the problems and the lower bound techniques may be of general interest.

* A full version of the paper is available at https://arxiv.org/abs/1705.01082 [7].
[1] In other words, under the promise that $f = g$.

44th International Colloquium on Automata, Languages, and Programming (ICALP 2017).
Editors: Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl;
Article No. 49; pp. 49:1–49:14


Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1998 ACM Subject Classification E.4 Coding and Information Theory

Keywords and phrases randomness, uncertainty, communication, imperfectly shared randomness, lower bounds

Digital Object Identifier 10.4230/LIPIcs.ICALP.2017.49

1 Introduction

In many forms of communication (e.g., human, computer-to-computer), the communicating parties share some context (e.g, knowledge of a language, operating system, communication protocol, encoding/decoding mechanisms.). This context is usually a) *huge* and b) *imperfectly shared* among the parties. Nevertheless, in human communication, very efficient communication is usually possible. Can we come up with a mathematical analogue of this phenomenon of efficient communication based on a huge but imperfectly shared context? Motivated by this general question, the study of "communication amid uncertainty" has been the subject of a series of recent work starting with Goldreich, Juba and Sudan [13, 8] followed by [12, 14, 15, 11, 4]. While early works were very abstract and general, later works (starting with Juba, Kalai, Khanna and Sudan [12]) tried to explore the ramifications of uncertainty in Yao's standard communication complexity model [26]. In particular, the more recent works relax the different pieces of context that were assumed to be *perfectly shared* in Yao's model, such as shared randomness [4], and in a recent work of the authors with Komargodski and Kothari the *function being computed* [6].

Specifically, [6] study the following functional notion of uncertainty in communication. Their setup builds on – and generalizes – Yao's classical model of (distributional) communication complexity, where Alice has an input x and Bob has an input y, with (x, y) being sampled from a distribution μ. Their goal is to communicate minimally so as to compute some function $g(x, y)$ (with high probability over the choice of (x, y)). The understated emphasis of the model is that for many functions g, the communication required is much less than the lengths of x or y, the entropy of x or y or even the conditional entropy of x given y.

The question studied by [6] is: How much of this gain in communication is preserved when the communicating parties do not exactly agree on the function being computed? (We further discuss the importance of this question in Section 1.2.) This variation of the problem is modelled as follows: Alice is given a Boolean function f and an input string x, and Bob is given a Boolean function g and an input string y where (x, y) is sampled from a known distribution μ as before, and (f, g) is chosen (adversarially) from a known class $\mathcal{F}$ of pairs of functions that are close in terms of the Hamming distance Δ_μ (weighted according to μ). Alice and Bob wish to compute $g(x, y)$. Alice's knowledge of the function f (which is close but not necessarily equal to g) captures the uncertainty in the knowledge of the context.

We define the *public-coin uncertain communication complexity* $\mathsf{PubCCU}^{\mu}_{\epsilon}(\mathcal{F})$ as the minimum length of a two-way public-coin protocol whose output is correct with probability at least $1-\epsilon$ over its internal randomness and that of (x, y). We similarly define the *private-coin uncertain communication complexity* $\mathsf{PrivCCU}^{\mu}_{\epsilon}(\mathcal{F})$ by restricting to private-coin protocols. Clearly, $\mathsf{PubCCU}^{\mu}_{\epsilon}(\mathcal{F}) \leq \mathsf{PrivCCU}^{\mu}_{\epsilon}(\mathcal{F})$. The quantities $\mathsf{owPubCCU}^{\mu}_{\epsilon}(\mathcal{F})$ and $\mathsf{owPrivCCU}^{\mu}_{\epsilon}(\mathcal{F})$ are similarly defined by restricting to one-way protocols.[2]

[2] Note that the uncertain model is clearly a generalization of Yao's model which corresponds to the particular case where $\mathcal{F} = \{(f, f)\}$ for some fixed function f. On the other hand, the uncertain model

The previous work ([6]) gave an upper bound on $\mathsf{owPubCCU}^{\mu}_{\epsilon}(\mathcal{F})$ whenever $\mathcal{F}$ consists of functions g whose one-way distributional complexity is small. More precisely, denote by $\mathsf{owCC}^{\mu}_{\epsilon}(g)$ the one-way communication complexity of g in the standard distributional model.[3] Namely, $\mathsf{owCC}^{\mu}_{\epsilon}(g)$ is the minimum length of a one-way deterministic protocol computing g with probability at least $1-\epsilon$ over the randomness of (x,y). Then, [6] showed that if $\mathcal{F}$ consists of pairs (f,g) of functions that are at distance δ, and if $\mathsf{owCC}^{\mu}_{\epsilon}(f), \mathsf{owCC}^{\mu}_{\epsilon}(g) \leq k$, then for every positive θ, $\mathsf{owPubCCU}^{\mu}_{\epsilon+2\delta+\theta}(\mathcal{F}) \leq O_{\theta}(k \cdot (1+I(x;y)))$, where $I(x;y)$ denotes the mutual information between x and y.[4] Note that if μ is a product distribution and if we let the parameter θ be a small constant, then the blow-up in communication is only a constant factor. However, the protocol of [6] crucially uses public randomness, and one of the main motivations behind this work is to understand how large the blow-up would be in the case where Alice and Bob have access to weaker types of randomness (or no randomness at all).

We point out that understanding the type of randomness that is needed in order to cope with uncertainty is a core question in the setup of communication with contextual uncertainty: *If Alice and Bob do not (perfectly) agree on the function being computed, why can we assume that they (perfectly) agree on the shared randomness?*

1.1 Our Contributions

We prove several results about the power of shared randomness in uncertain communication.

Private and Imperfectly Shared Randomness

Our first result (Theorem 1) shows that private-coin protocols are much weaker than public-coin protocols in the setup of communication with contextual uncertainty. Far from obtaining a constant factor blow-up in communication, private-coin protocols incur an increase that is a growing function of n when dealing with uncertainty.

Let $\mathcal{U} \triangleq \mathcal{U}_{2n}$ be the uniform distribution on $\{0,1\}^{2\cdot n}$. For positive integers t and n, we define $\log^{(t)}(n)$ by setting $\log^{(1)}(n) = \log n$, and $\log^{(i)}(n) = \max(\log\log^{(i-1)}(n), 1)$ for all $i \in \{2,\ldots,t\}$.

▶ **Theorem 1** (Lower-bound on private-coin uncertain protocols). *For every sufficiently small $\delta > 0$, there exist a positive integer $\ell \triangleq \ell(\delta)$ and a function class $\mathcal{F} \triangleq \mathcal{F}_{\delta}$ such that*

- **(i)** *For each $(f,g) \in \mathcal{F}$, we have that $\Delta_{\mathcal{U}}(f,g) \leq \delta$.*
- **(ii)** *For each $(f,g) \in \mathcal{F}$, we have that $\mathsf{owCC}^{\mathcal{U}}_{0}(f), \mathsf{owCC}^{\mathcal{U}}_{0}(g) \leq \ell$.*
- **(iii)** *For every $\eta > 0$ and $\epsilon \in (4\delta, 0.5]$, we have that $\mathsf{PrivCCU}^{\mathcal{U}}_{\epsilon/2-2\delta-\eta}(\mathcal{F}) = \Omega(\eta^2 \cdot \log^{(t)}(n))$ for some positive integer $t = \Theta((\epsilon/\delta)^2)$.*

In Theorem 1, the inputs x and y are binary strings of length n and $\mathcal{F}$ is a family of pairs of functions, which each function mapping $\{0,1\}^n \times \{0,1\}^n$ to $\{0,1\}$. Also, the parameter η

can also be viewed as a particular case of Yao's model via an exponential blow-up in the input size. For more on this view (which turns out to be ineffective in our setup), we refer the reader to Note 9 at the end of this section.

[3] By the "easy direction" of Yao's min-max principle, we can without loss of generality consider deterministic (instead of public-coin) protocols when defining $\mathsf{owCC}^{\mu}_{\epsilon}(g)$. We point out that this is not true in the uncertain case.

[4] One interpretation of the dependence of the communication in the uncertain setup on the mutual information $I(x;y)$ is that the players are better able to use the correlation of their inputs in the standard case than in the uncertain case.

can possibly depend on n. We point out that Theorem 1 also implies the first separation between *deterministic* uncertain protocols and public-coin uncertain protocols[5].

▶ Note 2. We point out that the relative power of private-coin and public-coin protocols in the uncertain model is both conceptually and technically different from the standard model. Specifically, the randomness is potentially used in the standard model in order to fool an adversary selecting the input pair (x, y), whereas in the uncertain model, it is used to fool an adversary selecting the pair (f, g) of functions that are promised to be close. This promise makes the task of proving lower bounds against private-coin protocols in the uncertain model (e.g., Theorem 1) significantly more challenging than in the standard model.[6] Moreover, a well-known theorem due to Newman [21] shows that in the standard model, any public-coin protocol can be simulated by a private-coin protocol while increasing the communication by an additive $O(\log n)$ bits. By contrast, there is no known analogue of Newman's theorem in the uncertain case!

▶ Note 3. The construction that we use to prove Theorem 1 cannot give a separation larger than $\Theta(\log \log n)$. Thus, showing a separation of $\omega(\log \log n)$ between private-coin and public-coin protocols in the uncertain case would require a new construction. For more details, see Note 13.

In light of Theorem 1, it is necessary for Alice and Bob to share some form of randomness in order to only incur a constant blow-up in communication for product distributions. Fortunately, it turns out that it is not necessary for Alice and Bob to *perfectly* share a sequence of random coins. If Alice is given a uniform-random string r of bits and Bob is given a string r' obtained by independently flipping each coordinate of r with probability 0.49, then efficient communication is still possible!

More formally, for $\rho \in [0, 1]$, define $\mathsf{owIsrCCU}^{\mu}_{\epsilon,\rho}(\mathcal{F})$ in the same way that we defined $\mathsf{owPubCCU}^{\mu}_{\epsilon}(\mathcal{F})$ except that instead of Alice and Bob having access to public randomness, Alice will have access to a sequence r of independent uniformly-random bits, and Bob will have access to a sequence r' of bits obtained by independently flipping each coordinate of r with probability $(1 - \rho)/2$. Note that this setup of imperfectly shared randomness interpolates between the public randomness and private randomness setups, i.e., $\mathsf{owIsrCCU}^{\mu}_{\epsilon,1}(\mathcal{F}) = \mathsf{owPubCCU}^{\mu}_{\epsilon}(\mathcal{F})$ and $\mathsf{owIsrCCU}^{\mu}_{\epsilon,0}(\mathcal{F}) = \mathsf{owPrivCCU}^{\mu}_{\epsilon}(\mathcal{F})$.

▶ **Theorem 4** (Uncertain protocol using imperfectly shared randomness)**.** *Let $\rho \in (0, 1]$ and μ be a* product *distribution. Let $\mathcal{F}$ consist of pairs (f, g) of functions with $\Delta_\mu(f, g) \leq \delta$, and $\mathsf{owCC}^{\mu}_{\epsilon}(f), \mathsf{owCC}^{\mu}_{\epsilon}(g) \leq k$. Then, for every positive θ, $\mathsf{owIsrCCU}^{\mu}_{\epsilon+2\delta+\theta,\rho}(\mathcal{F}) \leq O_\theta(k/\rho^2)$.*

The *imperfectly shared randomness* model in Theorem 4 was recently independently introduced (in the setup of communication complexity) by Bavarian, Gavinsky and Ito [2] and by Canonne, Guruswami, Meka and Sudan [4] (and it was further studied in [5]). Moreover, our proof of Theorem 4 is based on combining the uncertain protocol of [6] and the locality-sensitive-hashing based protocol of [4].

We point out that Theorem 4 also holds for more general i.i.d. sources of correlated randomness than the one described above. More precisely, for i.i.d. (not necessarily binary)

[5] This uses the fact that private-coin communication complexity is no larger than deterministic communication complexity, both in the certain and uncertain setups.

[6] In particular, the diagonilization-based arguments that imply a separation between the public-coin and the private-coin communication complexities of the Equality function in the standard model completely fail when we impose such a promise.

sources of (imperfectly) shared randomness with *maximal correlation*[7] ρ, the work of Witsenhausen [24] along with the protocols of [4] and [6] imply an uncertain protocol with $O_\theta(k/\rho^2)$ bits of communication.

Public Randomness

We now turn to our next result where we consider the dependence of the upper bound of [6] on the mutual information $I \triangleq I(X;Y)$ in the case of public-coin protocols. The previous work [6] proved a lower bound of $\Omega(\sqrt{I})$ on this dependence, but their lower-bound does not grow with k. We improve this lower bound to $\Omega(\sqrt{k} \cdot \sqrt{I})$.

▶ **Theorem 5** (Improved lower-bound on public-coin uncertain protocols)**.** *For every sufficiently small $\delta > 0$ and every positive integers k, n such that $k = o(\exp(\sqrt{n}))$, there exist an input distribution μ on input pairs $(X, Y) \in \{0,1\}^{k \cdot n} \times \{0,1\}^{k \cdot n}$ with mutual information $I \approx k \cdot n$ and a function class $\mathcal{F} \triangleq \mathcal{F}_{\delta,k,n}$ such that*[8]

- **(i)** *For each $(f, g) \in \mathcal{F}$, we have that $\Delta_\mu(f, g) \leq \delta$.*
- **(ii)** *For each $(f, g) \in \mathcal{F}$, we have that $\mathsf{owCC}_0^\mu(f), \mathsf{owCC}_0^\mu(g) \leq k$.*
- **(iii)** $\mathsf{owPubCCU}_\epsilon^\mu(\mathcal{F}) = \Omega(\sqrt{k} \cdot \sqrt{I})$ *for some absolute constant $\epsilon > 0$ independent of δ.*

As will be explained in detail in Section 2.2, the proof of Theorem 5 is based on an extension of the lower bound construction of [6], which is then analyzed using different techniques.

▶ Note 6. The construction that we use to prove Theorem 5 cannot give a lower bound larger than $\tilde{\Theta}(\sqrt{k} \cdot \sqrt{I})$. Thus, improving on the lower bound in Theorem 5 by more than logarithmic factors in k and I would require a new construction.

New Communication Problems

Our lower bounds in Theorems 1 and 5 are derived by defining new problems in *standard* communication complexity (i.e., without uncertainty) and proving lower bounds for these problems. We describe these problems and our results on these next.

The construction that we use to prove Theorem 1 requires us to understand the following "subset-majority with side information" setup. Alice is given a subset $S \subseteq [n]$ and a string $x \in \{\pm 1\}^n$, and Bob is given a subset $T \subseteq [n]$ and a string $y \in \{\pm 1\}^n$. The subsets S and T are adversarially chosen but are promised to satisfy $S \subseteq T$, $|T| = \ell$ and $|T \setminus S| \leq \delta \cdot \ell$ for some fixed parameters ℓ and δ. The strings x and y are chosen independently and uniformly at random. Alice and Bob wish to compute the function $\mathsf{SubsetMaj}((S,x),(T,y)) \triangleq \mathsf{Sign}(\sum_{i \in T} x_i y_i)$. In words, $\mathsf{SubsetMaj}((S,x),(T,y))$ is equal to 0 if x and y differ on a majority of the coordinates in subset T, and 1 otherwise. Note that S does not directly appear in the definition of the function $\mathsf{SubsetMaj}$ but it can serve as useful side-information for Alice.[9] What is the private-coin communication complexity of computing $\mathsf{SubsetMaj}$ on every (S,T)-pair satisfying the above promise and with high probability over the random choice of (x,y) and over the private randomness? We prove the following (informally stated) lower bound.

---

[7] The *maximal correlation* of a pair (X,Y) of random variables (with support $\mathcal{X} \times \mathcal{Y}$) is defined as $\rho(X,Y) \triangleq \sup \mathbb{E}[F(X)G(Y)]$ where the supremum is over all functions $F : \mathcal{X} \to \mathbb{R}$ and $G : \mathcal{Y} \to \mathbb{R}$ with $\mathbb{E}[F(X)] = \mathbb{E}[G(Y)] = 0$ and $\mathsf{Var}[F(X)] = \mathsf{Var}[G(Y)] = 1$. It is not hard to show that the binary source of imperfectly shared randomness defined in the paragraph preceding Theorem 4 has maximal correlation ρ.

[8] We note that $I \approx k \cdot n$ means that $I/(k \cdot n) \to 0$ as $n \to \infty$.

[9] Note that we could have alternatively defined $\mathsf{SubsetMaj}$ in terms of S, and let T serve as the potentially useful side-information. Our lower bound would also apply to this setup.

▶ **Lemma 7.** *Any private-coin protocol computing* SubsetMaj *on every* (S,T)*-pair satisfying the promise and with high probability over the random choice of* (x,y) *and over the private randomness should communicate at least* $\log^{(t)}(n)$ *bits for some positive integer* t *that depends on* δ *and the error probability.*

The proof of Theorem 5 is based on a construction that leads to the question described next regarding the communication complexity of a particular block-composed function. Namely, consider the following "majority composed with subset-parity with side information" setup. Alice is given a sequence of subsets $S \triangleq (S^{(i)} \subseteq [n])_{i\in[k]}$ and a sequence of strings $x \triangleq (x^{(i)} \in \{0,1\}^n)_{i\in[k]}$, and Bob is given a sequence of subsets $T \triangleq (T^{(i)} \subseteq [n])_{i\in[k]}$ and a sequence of strings $y \triangleq (y^{(i)} \in \{0,1\}^n)_{i\in[k]}$. We consider the following distribution μ on $((S,x),(T,y))$. Independently for each $i \in [k]$, we sample $((S^{(i)},x^{(i)}),(T^{(i)},y^{(i)}))$ as follows: $S^{(i)}$ is a uniform-random subset and $T^{(i)}$ is an ϵ-noisy[10] copy of $S^{(i)}$, and independently $x^{(i)}$ is a uniform-random string and $y^{(i)}$ is an ϵ-noisy copy of x. Here, ϵ is a positive parameter that can depend on n and k. Alice and Bob wish to compute the function $\mathsf{Maj} \circ \mathsf{SubsetParity}((S,x),(T,y)) \triangleq \mathsf{Sign}\big(\sum_{i=1}^{k}(-1)^{\langle T^{(i)},x^{(i)}\oplus y^{(i)}\rangle}\big)$ where $T^{(i)}$ denotes both the subset and its 0/1 indicator vector, the inner product is over $\mathbb{F}_2$, and $x^{(i)} \oplus y^{(i)}$ is the coordinate-wise XOR of $x^{(i)}$ and $y^{(i)}$. What is the communication complexity of computing $\mathsf{Maj} \circ \mathsf{SubsetParity}$ with high probability over the distribution μ? We prove the following lower bound.

▶ **Lemma 8.** *Any* 1*-way protocol computing* $\mathsf{Maj} \circ \mathsf{SubsetParity}$ *with high probability over the distribution* μ *should communicate* $\Omega(k \cdot \epsilon \cdot n)$ *bits.*

In Section 2.1, we outline the proof of Theorem 1 and explain how it leads to the setup of Lemma 7 and how we prove Lemma 7. In Section 2.2, we outline the proof of Theorem 5 and explain how it leads to the setup of Lemma 8 and how we prove Lemma 8.

Before doing so, we discuss some conceptual implications of our results.

1.2 Implications

Functional uncertainty models much of the day-to-day interactions among humans, where a person is somewhat aware of the objectives of the other person she is interacting with, but do not know them precisely. Neither person typically knows exactly what aspects of their own knowledge may be relevant to the interaction, yet they do manage to have a short conversation. This is certainly a striking phenomenon, mostly unexplained in mathematical terms. This line of works aims to explore such phenomena. It is important to understand what mechanisms may come into play here, and what features play a role. Is the ability to make random choices important? Is shared information crucial? Is there a particular measure of distance between functions that makes efficient communication feasible? In order to understand such questions, one first needs to have a ground-level understanding of communication with functional uncertainty. This work tackles several basic questions that remain unexplored.

An ideal model for communication would only assume a constant amount of perfectly shared context between the sender and receiver, such as the knowledge of an encoding/decoding algorithm, one universal Turing machine, etc. Solutions to most interesting communication problems seem to assume a shared information which grows with the length of the inputs.

[10] This means that the indicator vector of $T^{(i)}$ is obtained by independently flipping each coordinate of the indicator vector of $S^{(i)}$ with probability ϵ.

Recent work showed that in many of these scenarios some assumptions about the shared context can be relaxed to an imperfect sharing, but these results are often brittle and break when two or more contextual elements are simultaneously assumed to be imperfectly shared. Our work raises the question of whether *imperfectly shared randomness* would be sufficient to overcome *functional uncertainty*. We show that this is indeed the case for product distributions, but the loss for non-product distributions might be much larger (for this and other open questions, we refer the reader to our conclusion Section 6). Such results highlight the delicate nature of the role of shared context in communication. They beg for a more systematic study of communication which at the very least should be able to mimic the aims, objectives and phenomena encountered in human communication.

▶ Note 9. As mentioned in Footnote 2, the uncertain model is clearly a generalization of Yao's model. Strictly speaking, the uncertain model can also be viewed as a particular case of Yao's model by regarding the function(s) that is being computed as part of the inputs of Alice and Bob, which results in an exponential blow-up in the input-size. This latter view turns out to be fruitless for our purposes. Indeed, from this perspective, all the different well-studied communication functions (such as Equality, Set Disjointness, Pointer Jumping, etc.) are regarded as special cases of one "universal function"! More importantly, this view completely blurs the distinction between the *goal* of the communication (i.e., the function to compute) and the inputs of the parties. On a technical level, it does not simplify the task of proving the lower bounds in Theorems 1 and 5 in any way since it does not capture the promise that the two functions (given to Alice and Bob) are close in Hamming distance. Thus, in the rest of this paper, we stick to the former view and use the expressions "uncertain model" and "standard model" to refer to the setups with and without uncertainty, respectively.

2 Overview of Proofs

2.1 Overview of Proof of Theorem 1

Reduction to Lemma 7

In order to prove Theorem 1, we need to devise a function class for which circumventing the uncertainty is much easier using public randomness than using private randomness. One general setup in which Bob can leverage public randomness to resolve some uncertainty regarding Alice's knowledge is the following "small-set intersection" problem. Assume that Alice is given a subset $S \subseteq [n]$, and Bob is given a subset $T \subseteq [n]$ such that T contains S and $|T| = \ell$, where we think of ℓ as being a large constant. Here, Bob knows that Alice has a subset of his own T but he is uncertain which subset Alice has. Using public randomness, a standard 1-way hashing protocol communicating $\tilde{O}(\ell)$ bits allows Bob to determine S with high probability. On the other hand, using only private randomness, the communication complexity of this task is $\Theta(\log \log n)$ bits.

With the above general setup in mind, we consider functions f_S indexed by small subsets S of coordinates on which they depend. Since we want the functions f_S and f_T to be close in Hamming distance, we enforce $|T \setminus S|$ to be small for every pair (f_S, f_T) of functions in our class, and we let each function f_S be "noise-stable". Since we want our function f_S to genuinely depend on all coordinates in S, the majority function $f_S(x, y) = \mathsf{Sign}(\sum_{i \in S} x_i y_i)$ for $x, y \in \{\pm 1\}^n$ arises as a natural choice. We also let x and y be independent uniform-random strings. In this case, it can be seen that if $|T \setminus S|$ is a small constant fraction of $|T|$, then the quadratic polynomials $\sum_{i \in S} x_i y_i$ and $\sum_{i \in T} x_i y_i$ behave like standard Gaussians with correlation close to 1, and the quadratic threshold functions $f_S(x, y)$ and $f_T(x, y)$ are thus close in Hamming distance.

Note that in the certain case, i.e., when both Alice and Bob agree on S, they can easily compute $f_S(x, y)$ by having Alice send to Bob the ℓ bits $(x_i)_{i \in S}$. Moreover, if Alice and Bob are given access to public randomness in the uncertain case, Bob can figure out S via the hashing protocol mentioned above using $\tilde{O}(\ell)$ bits of communication, which would reduce the problem to the certain case[11]. The bulk of the proof will be to lower-bound the private-coin uncertain communication. Note that by the choice of our function class and distribution, this is equivalent to proving Lemma 7.

Proof of Lemma 7

To prove Lemma 7, the high-level intuition is that a protocol solving the uncertain problem should be essentially revealing to Bob the subset S that Alice holds. Formalizing this intuition turns out to be challenging, especially that a private-coin protocol solving the uncertain problem is only required to output a *single bit* which is supposed to equal the Boolean function $f_T(x, y)$ with high probability over (x, y) and over the private randomness. In fact, this high-level intuition can be shown not to hold in certain regimes[12]. Moreover, the standard proofs that lower bound the communication of small-set intersection do not extend to lower-bound the communication complexity of f_T.

To lower-bound the private-coin communication of solving the uncertain task by a growing function of n, we consider the following *shift communication game*. Bob is given a sorted tuple $\sigma = (\sigma_1, \ldots, \sigma_t)$ of integers with $1 \leq \sigma_1 < \cdots < \sigma_t \leq n$, and Alice is either given the prefix $(\sigma_1, \ldots, \sigma_{t-1})$ of length $t-1$ of σ or the suffix $(\sigma_2, \ldots, \sigma_t)$ of length $t-1$ of σ. Bob needs to determine the input of Alice. We show that a celebrated lower bound of Linial [19] on the *chromatic number* of certain related graphs implies a lower bound of $\log^{(t+1)}(n)$ on the private-coin communication of the shift communication game. We then show that any private-coin protocol solving the uncertain task can be turned into a private-coin protocol solving the shift-communication game with a constant (i.e., independent of n) blow-up in the communication.

2.2 Overview of Proof of Theorem 5

Reduction to Lemma 8

The proof of Theorem 5 builds on the lower-bound construction of [6] which we recall next. Let μ be the distribution over pairs $(x, y) \in \{0,1\}^{2n}$ where x is uniform-random and y is an ϵ-noisy copy of x with $\epsilon = \sqrt{\delta/n}$. Then, the mutual information between x and y satisfies $I \approx n$. For each $S \subseteq [n]$, consider the function $f_S(x, y) \triangleq \langle S, x \oplus y \rangle$ where the inner product is over $\mathbb{F}_2$, $x \oplus y$ denotes the coordinate-wise XOR of x and y, and S is used to denote both the subset and its 0/1 indicator vector. Moreover, consider the class $\mathcal{F}$ of all pairs of functions (f_S, f_T) where $|S \triangle T| \leq \sqrt{\delta n}$. It can be seen that for such S and T, the distance between f_S and f_T under μ is at most δ. If Alice and Bob both know S, then Alice can send the single bit $\langle S, x \rangle$ to Bob who can then output the correct answer $\langle S, x \oplus y \rangle = \langle S, x \rangle \oplus \langle S, y \rangle$. This means that the certain communication is 1 bit. Using the well-known discrepancy method, [6] showed a lower bound of $\Omega(\sqrt{n})$ bits on the communication of the associated uncertain problem. Since in this case $I \approx n$, this in fact lower-bounds the uncertain communication by

[11] Alternatively, Alice and Bob can run the protocol of [6] which would communicate $O(\ell)$ bits.

[12] For example, for constant error probabilities, the 1-way randomized communication complexity of small-set intersection is known to be $\Theta(\ell \cdot \log(\ell))$ bits (see, e.g., [3]) whereas the public-coin protocol of [6] can compute f_T with $O(\ell)$ bits of communication.

$\Omega(\sqrt{I})$ bits. For this construction, this lower bound turns out to be tight up to a logarithmic factor.

To improve the lower-bound from $\sqrt{I}$ to $\sqrt{k} \cdot \sqrt{I}$, we consider the following "block-composed" framework. Let $\{f_{S^{(i)}}(x^{(i)}, y^{(i)}) : i \in [k]\}$ be k independent copies of the above base problem of [6] and consider computing the composed function $g\big(f_{S^{(1)}}(x^{(1)}, y^{(1)}), \ldots, f_{S^{(k)}}(x^{(k)}, y^{(k)})\big)$ for some outer function $g : \{0,1\}^k \to \{0,1\}$. For any choice of g, the certain communication of the composed function would be at most k bits. When choosing the outer function g to use in our lower bound, we thus have two objectives to satisfy. First, g has to be sufficiently *hard* in the sense that its average-case decision tree complexity with respect to the uniform distribution on $\{0,1\}^k$ should be $\Omega(k)$; otherwise, it will not be the case that the uncertain communication of computing g on k copies of the base problem is at least k times the uncertain communication of the base problem. Second, g has to be *noise stable* in order to be able to upper bound the distance between $g\big(f_{S^{(1)}}(\cdot), \ldots, f_{S^{(k)}}(\cdot)\big)$ and $g\big(f_{T^{(1)}}(\cdot), \ldots, f_{T^{(k)}}(\cdot)\big)$.

Note that setting g to be a dictator function would satisfy the noise-stability property, but it clearly would not satisfy the hardness property, as the composed function would be equal to the base function and would thus have uncertain communication $\tilde{O}(\sqrt{n})$ bits. Another potential choice of g is to set it to the parity function on k bits. This function would satisfy the hardness property, but it would strongly violate the noise stability property that is crucial to us. This leads us to setting g to the majority function on k bits, which is well-known to be noise stable, and has average-case decision-tree complexity $\Omega(k)$ with respect to the uniform distribution on $\{0,1\}^k$. In fact, the noise stability of the majority function readily implies an upper bound of $O(\sqrt{\delta})$ on the distance between any pair of composed functions that are specified by tuples of subsets $(S^{(1)}, \ldots, S^{(k)})$ and $(T^{(1)}, \ldots, T^{(k)})$ with $|S^{(i)} \triangle T^{(i)}| \leq \sqrt{\delta n}$ for each $i \in [k]$. The crux of the proof will be to lower-bound the uncertain communication of the majority-composed function by $\Omega(k\sqrt{n})$, which amounts to proving Lemma 8. Since in this block-composed framework the mutual information satisfies $I \approx kn$, this would imply the lower bound of $\Omega(\sqrt{k}\sqrt{I})$ on the uncertain communication in Part (iii) of Theorem 5.

Proof of Lemma 8

We first point out that the average-case *quantum* decision tree complexity of Maj_k with respect to the uniform distribution is $\tilde{O}(\sqrt{k})$ [1]. This implies that any communication complexity lower-bound method that extends to the quantum model cannot prove a lower bound larger than $\tilde{O}(\sqrt{k} \cdot \sqrt{n})$ on our uncertain communication[13]. In particular, we cannot solely rely on the discrepancy bound (as done in [6]), since this bound is known to lower-bound quantum communication. Similarly, the techniques of [22, 23, 18] rely on the generalized discrepancy bound (originally due to [16]) which also lower-bounds quantum communication. Moreover, the recent results of [20] only apply to product distributions (i.e., where Alice's input is independent of Bob's input) in contrast to our case where the inputs of Alice and Bob are very highly-correlated. Finally, the recent works of [9, 10] do not imply lower bounds on the average-case complexity with respect to the distribution that arises in our setup.

To circumvent the above obstacles, we use a new approach that is tailored to our setup and that is outlined next. Let Π be a 1-way protocol solving the uncertain task with high probability. We consider the information that Π reveals about the *inputs to the outer*

[13] Thus, since $I \approx k \cdot n$ in our block-composed framework, such methods cannot be used to improve the lower-bound of $\Omega(\sqrt{I})$ of [6] by more than logarithmic factors.

function, i.e., about the length-k binary string $\left(f_{S^{(1)}}(x^{(1)}, y^{(1)}), \ldots, f_{S^{(k)}}(x^{(k)}, y^{(k)})\right)$. We call this quantity the *intermediate information cost* of Π, and we argue that it is at least $\Omega(k)$ bits. To do so, we recall the Hamming distance function HD_k defined by $\mathsf{HD}_k(u, v) = 1$ if the Hamming distance between u and v is at least $k/2$ and $\mathsf{HD}_k(u, v) = 0$ otherwise. We upper bound the information complexity of computing HD_k over the uniform distribution on $\{0, 1\}^{2k}$ by the intermediate information cost of Π. We do so by giving an information-cost preserving procedure where Alice and Bob are given independent uniformly distributed u and v (respectively) and use their private and public coins in order to simulate the input distribution (X, Y) of our uncertain problem. The known 1-way lower bound of [25] on HD_k under the uniform distribution then implies that Π reveals $\Omega(k)$ bits of information to Bob about the tuple $\left(f_{S^{(1)}}(x^{(1)}, y^{(1)}), \ldots, f_{S^{(k)}}(x^{(k)}, y^{(k)})\right)$. This allows Bob to guess this tuple with probability 0.51^k. We then apply the strong direct product theorem for discrepancy of [17] which, along with the discrepancy-based lower bound on the communication of the base uncertain problem of [6], implies that Π should be communicating at least $\Omega(k\sqrt{n})$ bits.

Organization of the rest of the paper

In Section 3, we give some preliminaries. In Sections 4 and 5, we describe the proofs of Theorems 1 and 5 respectively. All the missing proofs as well as the proof of Theorem 4 appear in the full version. In Section 6, we conclude with some interesting open questions.

3 Preliminaries

For a real number x, we define $\mathsf{Sign}(x)$ to be 1 if $x \geq 0$ and 0 if $x < 0$. For a set S, we write $X \in_R S$ to indicate that X is a random variable that is uniformly distributed on S. For a positive integer n, we let $[n] \triangleq \{1, \ldots, n\}$. For a real number x, we denote $\exp(x)$ a quantity of the form $2^{\Theta(x)}$. For any two subsets $S, T \subseteq [n]$, we let $S \setminus T$ be the set of all elements of S that are not in T. We let $S \triangle T$ be the symmetric difference of S and T, i.e., the union of $S \setminus T$ and $T \setminus S$. For functions $f, g : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$ and any distribution μ on $\mathcal{X} \times \mathcal{Y}$, we define the distance $\Delta_\mu(f, g) \triangleq \Pr_{(x,y)\sim\mu}[f(x, y) \neq g(x, y)]$ as the Hamming distance between the values of f and g, weighted with respect to μ. If μ is the uniform distribution on $\mathcal{X} \times \mathcal{Y}$, we drop the subscript μ and denote Δ_μ by Δ. We next recall the standard communication complexity model of Yao [26]. For any function $f\colon \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$, we denote by $\mathsf{CC}(f)$, $\mathsf{owCC}(f)$, $\mathsf{PrivCC}_\epsilon(f)$ and $\mathsf{owPrivCC}_\epsilon(f)$ its two-way deterministic, one-way deterministic, two-way private-coin and one-way private-coin communication complexity respectively. For any distribution μ over $\mathcal{X} \times \mathcal{Y}$, we denote by $\mathsf{CC}^\mu_\epsilon(f)$ and $\mathsf{owCC}^\mu_\epsilon(f)$ the two-way distributional and one-way distributional communication complexity of over μ with error ϵ, respectively.

We next recall the model of communication with contextual uncertainty. For more details on this model, we refer the reader to [6]. In this setup, Alice knows a function f and is given an input x, and Bob knows a function g and is given an input y. Let $\mathcal{F} \subseteq \{f\colon \mathcal{X} \times \mathcal{Y} \to \{0, 1\}\}^2$ be a family of pairs of Boolean functions with domain $\mathcal{X} \times \mathcal{Y}$, and μ be a distribution on $\mathcal{X} \times \mathcal{Y}$. We say that a public-coin (resp. private-coin) protocol Π ϵ-computes $\mathcal{F}$ over μ if for every $(f, g) \in \mathcal{F}$, we have that Π outputs the value $g(x, y)$ with probability at least $1 - \epsilon$ over the randomness of $(x, y) \sim \mu$ and over the public (resp. private) randomness of Π.

▶ **Definition 10** (Contextually Uncertain Communication Complexity)**.** Let μ be a distribution on $\mathcal{X} \times \mathcal{Y}$ and $\mathcal{F} \subseteq \{f\colon \mathcal{X} \times \mathcal{Y} \to \{0, 1\}\}^2$. The *two-way (resp. one-way) public-coin communication complexity of $\mathcal{F}$ under contextual uncertainty*, denoted $\mathsf{PubCCU}^\mu_\epsilon(\mathcal{F})$ (resp. $\mathsf{owPubCCU}^\mu_\epsilon(\mathcal{F})$), is the minimum over all two-way (resp. one-way) public-coin protocols Π

that ϵ-compute $\mathcal{F}$ over μ, of the maximum communication complexity of Π over $(f,g) \in \mathcal{F}$, (x,y) from the support of μ and settings of the public coins.

Similarly, the *two-way (resp. one-way) private-coin communication complexity of $\mathcal{F}$ under contextual uncertainty* $\mathsf{PrivCCU}^{\mu}_{\epsilon}(\mathcal{F})$ (resp. $\mathsf{owPrivCCU}^{\mu}_{\epsilon}(\mathcal{F})$) is defined by restricting to two-way (resp. one-way) private-coin protocols.

4 Construction for Private-Coin Uncertain Protocols

We now describe the construction that is used to prove Theorem 1. Each function in our universe is specified by a subset $S \subseteq [n]$ and is of the form $f_S : \{\pm 1\}^n \times \{\pm 1\}^n \to \{0,1\}$ with $f_S(X,Y) \triangleq \mathsf{Sign}(\sum_{i \in S} X_i Y_i)$ for all $X, Y \in \{\pm 1\}^n$. The function class is then defined by

$$\mathcal{F}_\delta \triangleq \{(f_S, f_T) : S \subseteq T,\ |T| = \ell \text{ and } |T \setminus S| \leq \delta' \cdot \ell\},$$

where $\delta' = \alpha \cdot \delta^2$ for some sufficiently small positive absolute constant α, and $\ell = \ell(\delta)$ is a sufficiently large function of δ. The input pair (X,Y) is drawn from the uniform distribution on $\{\pm 1\}^{2n}$. The proof of Part (i) of Theorem 1 follows from the fact that the the polynomials $\sum_{i \in S} X_i Y_i$ and $\sum_{i \in T} X_i Y_i$ behave like zero-mean Gaussians with unit-variance and correlation $\sqrt{1-\delta'}$. The proof of Part (ii) of Theorem 1 is immediate and is given in the full version for completeness. To prove Part (iii) of Theorem 1, the next definition – which is based on the graphs studied by Linial [19]– will be crucial to us.

▶ **Definition 11** (Shift Communication Game $\mathcal{G}_{m,t}$)**.** Let m and t be positive integers with $t \leq m$. In the communication problem $\mathcal{G}_{m,t}$, Bob is given a sorted tuple $\sigma = (\sigma_1, \ldots, \sigma_t)$ of distinct integers with $1 \leq \sigma_1 < \cdots < \sigma_t \leq m$. In the YES case, Alice is given the prefix $(\sigma_1, \ldots, \sigma_{t-1})$ of length $t-1$ of σ. In the NO case, Alice is given the suffix $(\sigma_2, \ldots, \sigma_t)$ of length $t-1$ of σ. Alice and Bob need to determine which of the YES and NO cases occurs.

Lemma 12 lower-bounds the private-coin communication complexity of $\mathcal{G}_{m,t}$. Its proof uses Linial's lower bound on the chromatic number of related graphs.

▶ **Lemma 12.** *There is an absolute constant c such that for every sufficiently small $\epsilon > 0$, we have that* $\mathsf{PrivCC}_{\epsilon}(\mathcal{G}_{m,t}) \geq c \cdot \log^{(t+2)}(m)$.

The proof of Part (iii) of Theorem 1 – which is the main part in the proof of Theorem 1 – is deferred to the full version.

▶ Note 13. As mentioned in Note 3, the above construction cannot give a separation larger than $\Theta(\log \log n)$. This is because using private randomness, Bob can learn the set S using $O(\log \log n)$ bits of communication (see, e.g., [3]). Additionally, Alice can send the coordinates of X indexed by the elements of S to Bob who can then compute $f_S(X,Y)$.

5 Construction for Public-Coin Uncertain Protocols

In this section, we describe the construction that is used to prove Theorem 5. We set

$$\delta' \triangleq c \cdot \delta^2 \quad \text{and} \quad \epsilon \triangleq \sqrt{\delta'/n}, \tag{$\dagger$}$$

where δ is the parameter from the statement of Theorem 5, and $c > 0$ is a small-enough absolute constant. To define our input distribution, we first define a slightly more general distribution μ_η. The support of μ_η is $\{0,1\}^{kn} \times \{0,1\}^{kn}$ and we will view the coordinates of a sample $(x,y) \sim \mu_\eta$ as $x = (x^{(i)})_{i \in [k]}$ and $y = (y^{(i)})_{i \in [k]}$ with $x^{(i)}, y^{(i)} \in \{0,1\}^n$ for all

$i \in [k]$. A sample $(x, y) \sim \mu_\eta$ is generated by letting $x \in_R \{0,1\}^{kn}$ and for all $i \in [k]$ and $j \in [n]$, independently setting $y_j^{(i)}$ to be an η-noisy copy of $x_j^{(i)}$. In other words, we set $y_j^{(i)} = x_j^{(i)}$ w.p. $1-\eta$ and $y_j^{(i)} = 1 - x_j^{(i)}$ w.p. η. Our input distribution is then $\mu_{2\epsilon-2\epsilon^2}$.

We now define our function class $\mathcal{F}_\epsilon$. Each function in our universe is specified by a sequence of subsets $S \triangleq (S^{(i)} \subseteq [n])_{i\in[k]}$ and it is of the form $f_S : \{0,1\}^{2\cdot k\cdot n} \to \{0,1\}$ with[14]

$$f_S(x,y) \triangleq \mathsf{Sign}\Big(\sum_{i\in[k]} (-1)^{\langle S^{(i)}, x^{(i)} \oplus y^{(i)}\rangle}\Big) \tag{1}$$

for all $x, y \in \{0,1\}^{k\cdot n}$, where in Eq. (1) the inner product is over $\mathbb{F}_2$, the sum is over $\mathbb{R}$ and $x^{(i)} \oplus y^{(i)}$ denotes the coordinate-wise XOR of the two length-n binary strings $x^{(i)}$ and $y^{(i)}$. The function class is then defined by $\mathcal{F}_\epsilon \triangleq \{(f_S, f_T) : |S^{(i)} \triangle T^{(i)}| \le \epsilon \cdot n \text{ for all } i \in [k]\}$.

The proof of Part (i) of Theorem 5 follows from known bounds on the noise stability of the majority function. The proof of Part (ii) is immediate. To prove Part (iii), we first define (as in [6]) a communication problem in the standard distributional model that reduces to solving the contextually-uncertain problem specified by the function class $\mathcal{F}_\epsilon$ and the distribution $\mu_{2\epsilon-2\epsilon^2}$. For distributions ϕ and ψ, we denote by $\phi \otimes \psi$ the joint distribution of a sample from ϕ and an independent sample from ψ. The new problem is defined as follows.

- **Inputs:** Alice's input is a pair (S, x) where $S \triangleq (S^{(i)} \subseteq [n])_{i\in[k]}$ and $x \in \{0,1\}^{k\cdot n}$. Bob's input is a pair (T, y) where $T \triangleq (T^{(i)} \subseteq [n])_{i\in[k]}$ and $y \in \{0,1\}^{k\cdot n}$.
- **Distribution:** Let $\mathcal{D}_q$ be the distribution on the pair (S, T) of sequences of k subsets of $[n]$, which is defined by independently setting, for each $i \in [k]$, $S^{(i)}$ to be a uniformly-random subset of $[n]$, and $T^{(i)}$ to be a q-noisy copy of $S^{(i)}$. The distribution on the inputs of Alice and Bob is then given by $\nu_\epsilon \triangleq \mathcal{D}_\epsilon \otimes \mu_{2\epsilon-2\epsilon^2}$ with $\epsilon = \sqrt{\delta'/n}$.
- **Function:** The goal is to compute the function $F : \{0,1\}^{2kn} \times \{0,1\}^{2kn} \to \{0,1\}$ defined by $F((S,x),(T,y)) = f_T(x,y) = \mathsf{Sign}\big(\sum_{i\in[k]} (-1)^{\langle T^{(i)}, x^{(i)} \oplus y^{(i)}\rangle}\big)$.

The next proposition follows from a simple application of the Chernoff bound.

▶ **Proposition 14.** *For any $\theta > 0$, $\mathsf{owPubCCU}_\theta^{\mu_{2\epsilon-2-\epsilon^2}}(\mathcal{F}_\epsilon) \ge \mathsf{owCC}_{\theta+\theta'}^{\nu_\epsilon}(F)$ with $\theta' = 2^{-\Theta(\epsilon\cdot n)}$.*

In the full version, we prove the following lower bound on $\mathsf{owCC}_\theta^{\nu_\epsilon}(F)$, which along with Proposition 14 and the settings of ϵ and δ' in (†), implies Part (iii) of Theorem 5.

▶ **Lemma 15.** *For every sufficiently small positive constant θ, $\mathsf{owCC}_\theta^{\nu_\epsilon}(F) = \Omega(k \cdot \epsilon \cdot n)$.*

6 Open Questions

As mentioned in Notes 3 and 6 in Section 1, significantly improving the bounds from Theorems 1 and 5 seems to require fundamentally new constructions, and is a very important question. Moreover, is there an analogue of the protocol in Theorem 4 for *non-product* distributions?

Another very important and intriguing open question is whether efficient communication under contextual uncertainty is possible in the multi-round setup. Namely, if k is the r-round certain communication, can we upper bound the r-round uncertain communication by some function of k, I and possibly r? Even for $r = 2$ and when the uncertain protocol is allowed

[14] We will use the symbol $S^{(i)}$ to denote both the subset of $[n]$ and its corresponding 0/1 indicator vector.

to use public randomness, no non-trivial protocols are known in this setting. On the other hand, no separations are known for this case (beyond those known for $r = 1$) even if the protocols are restricted to be deterministic.

Acknowledgements. The authors would like to thank Ilan Komargodski, Pravesh Kothari and Mohsen Ghaffari and the anonymous reviewers for very helfpul discussions and pointers.

References

1. Andris Ambainis and Ronald De Wolf. Average-case quantum query complexity. *Journal of Physics A: Mathematical and General*, 34(35):6741, 2001.
2. Mohammad Bavarian, Dmitry Gavinsky, and Tsuyoshi Ito. On the role of shared randomness in simultaneous communication. In *Automata, Languages, and Programming*, pages 150–162. Springer, 2014.
3. Joshua Brody, Amit Chakrabarti, Ranganath Kondapally, David P. Woodruff, and Grigory Yaroslavtsev. Beyond set disjointness: the communication complexity of finding the intersection. In *Proceedings of the 2014 ACM symposium on Principles of distributed computing*, pages 106–113. ACM, 2014.
4. Clément Louis Canonne, Venkatesan Guruswami, Raghu Meka, and Madhu Sudan. Communication with imperfectly shared randomness. In *Innovations in Theoretical Computer Science, ITCS*, pages 257–262, 2015.
5. Badih Ghazi, Pritish Kamath, and Madhu Sudan. Communication complexity of permutation-invariant functions. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 1902–1921, 2016. doi:10.1137/1.9781611974331.ch134.
6. Badih Ghazi, Ilan Komargodski, Pravesh Kothari, and Madhu Sudan. Communication with contextual uncertainty. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 2072–2085, 2016. doi:10.1137/1.9781611974331.ch144.
7. Badih Ghazi and Madhu Sudan. The power of shared randomness in uncertain communication. *arXiv preprint arXiv:1705.01082*, 2017. URL: https://arxiv.org/abs/1705.01082.
8. Oded Goldreich, Brendan Juba, and Madhu Sudan. A theory of goal-oriented communication. *J. ACM*, 59(2):8, 2012.
9. Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles are nonnegative juntas. In *Proceedings of the 47th Symposium on Theory of Computing (STOC). ACM*, 2015.
10. Mika Göös, Toniann Pitassi, and Thomas Watson. The landscape of communication complexity classes. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 22, page 49, 2015.
11. Elad Haramaty and Madhu Sudan. Deterministic compression with uncertain priors. In *Innovations in Theoretical Computer Science, ITCS*, pages 377–386, 2014.
12. Brendan Juba, Adam Tauman Kalai, Sanjeev Khanna, and Madhu Sudan. Compression without a common prior: an information-theoretic justification for ambiguity in language. In *Innovations in Computer Science, ICS*, pages 79–86, 2011.
13. Brendan Juba and Madhu Sudan. Universal semantic communication I. In *40th Annual ACM Symposium on Theory of Computing*, pages 123–132, 2008.
14. Brendan Juba and Madhu Sudan. Efficient semantic communication via compatible beliefs. In *Innovations in Computer Science, ICS*, pages 22–31, 2011.
15. Brendan Juba and Ryan Williams. Massive online teaching to bounded learners. In *Innovations in Theoretical Computer Science, ITCS*, pages 1–10, 2013.

16 Hartmut Klauck. Lower bounds for quantum communication complexity. In *Foundations of Computer Science, 2001. Proceedings. 42nd IEEE Symposium on*, pages 288–297. IEEE, 2001.

17 Troy Lee, Adi Shraibman, and Robert Spalek. A direct product theorem for discrepancy. In *Computational Complexity, 2008. CCC'08. 23rd Annual IEEE Conference on*, pages 71–80. IEEE, 2008.

18 Troy Lee and Shengyu Zhang. Composition theorems in communication complexity. In *Automata, Languages and Programming*, pages 475–489. Springer, 2010.

19 Nathan Linial. Locality in distributed graph algorithms. *SIAM Journal on Computing*, 21(1):193–201, 1992.

20 Marco Molinaro, David P. Woodruff, and Grigory Yaroslavtsev. Amplification of one-way information complexity via codes and noise sensitivity. In *Automata, Languages, and Programming*, pages 960–972. Springer, 2015.

21 Ilan Newman. Private vs. common random bits in communication complexity. *Information processing letters*, 39(2):67–71, 1991.

22 Alexander A. Sherstov. The pattern matrix method for lower bounds on quantum communication. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 85–94. ACM, 2008.

23 Yaoyun Shi and Yufan Zhu. Quantum communication complexity of block-composed functions. *arXiv preprint arXiv:0710.0095*, 2007.

24 Hans S. Witsenhausen. On sequences of pairs of dependent random variables. *SIAM Journal on Applied Mathematics*, 28(1):100–113, 1975.

25 David P. Woodruff. *Efficient and private distance approximation in the communication and streaming models*. PhD thesis, Citeseer, 2007.

26 Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *11h Annual ACM Symposium on Theory of Computing*, pages 209–213, 1979.