Communication with Imperfectly Shared Randomness

Clément L. Canonne*
Columbia University
New York, NY 10027
ccanonne@cs.columbia.edu

Venkatesan Guruswami†
Carnegie Mellon University
Pittsburgh, PA 15213
guruswami@cmu.edu

Raghu Meka
Microsoft Research
1288 Pear Avenue
Mountain View, CA 94043
meka@microsoft.com

Madhu Sudan
Microsoft Research
1 Memorial Drive
Cambridge, MA 02142
madhu@mit.edu

ABSTRACT

The communication complexity of many fundamental problems reduces greatly when the communicating parties share randomness that is independent of the inputs to the communication task. Natural communication processes (say between humans) however often involve large amounts of shared correlations among the communicating players, but rarely allow for perfect sharing of randomness. Can the communication complexity benefit from shared correlations as well as it does from shared randomness? This question was considered mainly in the context of simultaneous communication by Bavarian et al. [1]. In this work we study this problem in the standard interactive setting and give some general results. In particular, we show that every problem with communication complexity of k bits with perfectly shared randomness has a protocol using imperfectly shared randomness with complexity $2^{\Omega(k)}$ bits. We also show that this is best possible by exhibiting a promise problem with complexity k bits with perfectly shared randomness which requires $2^{\Omega(k)}$ bits when the randomness is imperfectly shared. Along the way we also highlight some other basic problems such as compression, and agreement distillation, where shared randomness plays a central role and analyze the complexity of these problems in the imperfectly shared randomness model.

The technical highlight of this work is the lower bound that goes into the result showing the tightness of our general connection. This result builds on the intuition that communication with imperfectly shared randomness needs to be less sensitive to its random inputs than communication with perfectly shared randomness. The formal proof invokes results about the small-set expansion of the noisy hypercube and an invariance principle to convert this intuition to a proof, thus giving a new application domain for these fundamental results.

Categories and Subject Descriptors

E.4 [**Coding and Information Theory**]: Formal models of communication; H.1.1 [**Models and Principles**]: Systems and Information Theory—*Information theory*

General Terms

Theory

1. INTRODUCTION

The availability of shared randomness can lead to enormous savings in communication complexity when computing some basic functions whose inputs are spread out over different communicating players. A basic example of this is Equality Testing, where two players Alice and Bob have inputs $x \in \{0,1\}^n$ and $y \in \{0,1\}^n$ and need to determine if $x = y$. Deterministically this takes n bits of communication. This reduces to $\Theta(\log n)$ bits if Alice and Bob can toss coins and they are allowed some error. But if they share some randomness $r \in \{0,1\}^*$ independent of x and y then the communication cost drops to $O(1)$. (See, for instance, [11]).

A more prevalent example of a communication problem is compression with uncertain priors. Here Alice has a distribution P on a universe $[N] = \{1, \ldots, N\}$, and a message $m \in [N]$ chosen according to the distribution P. Alice is allowed to send some bits to Bob and Bob should output m and the goal is to minimize the expected number of bits that Alice sends Bob (over the random choice of m). If Bob knows the distribution P exactly then this is the classical compression problem, solved for example by Huffman coding. In most forms of natural communication (e.g., think about the next email you are about to send), Alice and Bob are not perfectly aware of the underlying context to their exchange, but have reasonably good ideas about each other. One way to model this is to say that Bob has a distribution Q that is

*Research supported in part by NSF CCF-1115703 and NSF CCF-1319788. Some of this work was done when the author was an intern at Microsoft Research New England.

†Some of this work was done when the author was visiting Microsoft Research New England. Research supported in part by NSF CCF-0963975.

ITCS'15, January 11–13, 2015, Rehovot, Israel.

ACM 978-1-4503-3333-7/15/01 ...$15.00.
http://dx.doi.org/10.1145/2688073.2688099.

close to the distribution P that Alice is working with, but is not identical to P. Compressing information down to its entropy in the presence of such uncertainty (i.e., $P \neq Q$) turns out to be possible if Alice and Bob share randomness that is independent of (P, Q, m) as shown by Juba et al. [9]. However it remains open as to whether such compression can be effected deterministically, without the shared randomness — the best known schemes can only achieve a compression length of roughly $O(H(P) + \log\log N)$, where $H(P) = \sum_{i \in [N]} P(i) \log 1/P(i)$ denotes the entropy of P.[1]

In both examples above it is natural to ask the question: can the (presumed) savings in communication be achieved in the absence of perfect sharing of randomness? The question especially makes sense in the latter context where the essential motivation is that Alice and Bob are not in perfect synchrony with each other: If Alice and Bob are not perfectly aware of the distributions P and Q, why should their randomness be identical?

The question of communication with imperfectly shared randomness was considered recently in the work of Bavarian et al. [1]. They consider the setting where Alice and Bob have randomness r and s respectively, with some known correlation between r and s, and study the implications of correlated randomness in the simultaneous message communication model (where a referee gets messages from Alice and Bob and computes some joint function of their inputs). Their technical focus is on the different kinds of correlations possible between r and s, but among basic results they show that equality testing has a $O(1)$ communication complexity protocol with correlated shared randomness.

In this work we are concerned with the setting of general communication protocols, where Alice and Bob interact to determine the value of some function. From some perspectives, this setting does not seem to offer a major difference between "private randomness" and "perfectly shared randomness" — Newman [14] shows that the communication complexity in the former setting can be larger by at most an additive $\log n$ term, where n is the input size. "Imperfectly shared randomness" being in between the two models cannot therefore be too far from them either. However, problems like compression above highlight a different perspective. There N is the size of the universe of all possible messages, and compression to $\log N$ bits of communication is trivial and uninteresting. Even a solution with $\log\log N$ bits of communication is not completely satisfactory. The real target is $O(H(P))$ bits of communication, which may be a constant independent of the universe size N (and for natural communication, the set of possible messages could be thought of as an infinitely large set). Thus the gap between the communication complexity with perfectly shared randomness and imperfectly shared randomness remains a very interesting question, which we explore in this paper.

[1]We stress that the setting of uncertain compression is completely different from that of compression with the "wrong distribution", a well-studied question in information theory. In the "wrong distribution problem" (see, for instance, [4, Theorem 5.4.3]) the sender and receiver agree on the distribution, say P, but both have it wrong and the distribution the message comes from is R. This leads to a compression length of $\mathbb{E}_{m \sim R}[\log(1/P(m))] \approx H(R) + D(R\|P)$. The important aspect here is that while the compression is not as good, there is no confusion between sender and receiver; and the latter is the focus of our problem.

We provide a formal description of our models and results in the following section, and here give an informal preview. We consider communication complexity in a simplified setting of imperfectly shared randomness: Alice has a uniform binary string $r \in \{0,1\}^m$ and Bob has a string s obtained by flipping each bit of r independently with some tiny probability. (While this setting is not the most general possible, it seems to capture the most interesting aspects of the "lack of prior agreement" between Alice and Bob.) Our main contributions in this work are the introduction of some new problems of interest in the context of communication complexity, and a comparison of their communication complexity with/without perfect sharing of randomness.

The first problem we study is the complexity of *compression with uncertain priors*. We show that any distribution P can be compressed to $O(H(P))$ bits even when the randomness is not perfectly shared. As in the analogous result of Juba et al. [9] this protocol sheds some light on natural communication processes, and introduces an error-correcting element that was not previously explained.

The next problem we mention is that of *agreement distillation*. Here Alice and Bob try to agree on a small random string using little communication. This is a natural problem to study in the context of communication complexity with imperfect randomness, since an efficient solution for this problem would allow Alice and Bob to convert any protocol using perfectly shared randomness into one that relies only on imperfectly shared randomness. It turns out that the zero-communication version of this question, where Alice and Bob are not allowed to communicate at all with each other, was studied by Bogdanov and Mossel [2]. They give a very strong negative result for this problem, showing that the probability that Alice and Bob can agree on a k-bit string is exponentially small in k. By a simple reduction we show that this implies that $o(k)$ bits of communication are insufficient to get agreement on k bits. Conversely, we also show that Alice and Bob can get a constant factor advantage — so they can communicate αk bits for some $\alpha < 1$. Such a result seems implicit in [2].

Returning to our work, we next attempt to get a general conversion of communication protocols from the perfectly-shared setting to the imperfectly-shared setting. We introduce a complete promise problem GapInnerProduct which captures two-way communication, and use it to show that any problem with k bits of communication with perfectly shared randomness also has a $\min\{\exp(k), k + \log n\}$ bit (one-way) protocol with imperfectly shared randomness. While the protocol is simple, we feel its existence is somewhat surprising; and indeed it yields a very different protocol for equality testing when compared with Bavarian et al. [1].

Lastly, our *main technical result* is a matching lower bound giving a parameterized family of promise problems, SparseGapInnerProduct, where the k'th problem can be solved with k bits of communication with perfect randomness, but requires $\exp(\Omega(k))$ bits with imperfect sharing. This result builds a new connection between influence of variables and communication complexity, which may be of independent interest. Finally we conclude with a variety of open questions.

2. MODEL, FORMAL DESCRIPTION OF RESULTS AND MAIN IDEAS

Throughout the paper, we denote by $\mathbb{Z}^+$ the set of positive integers, and by $[n]$ the set $\{1,\ldots,n\}$. Unless specified otherwise, all logarithms are in base 2. We also recall, for $x \in [0,1]$, the definition of the binary entropy function $h(x) = -x\log x - (1-x)\log(1-x)$; furthermore, for any $p \in [0,1]$, we will write $\mathrm{Bern}(p)$ for the Bernoulli distribution on $\{0,1\}$ with parameter p, and $\mathrm{Bern}^n(p)$ for the product distribution on $\{0,1\}^n$ of n independent Bernoulli random variables. For a distribution P over a domain Ω, we write $H(P) = \sum_{x\in\Omega} P(x)\log(1/P(x))$ for its entropy, and $x \sim P$ to indicate that x is drawn from P. $\mathcal{U}_\Omega$ denotes the uniform distribution over Ω.

Finally, for two elements $x, y \in \{+1,-1\}^n$, their *Hamming distance* $\mathrm{dist}(x,y)$ is defined as the number of coordinates in which they differ (and similarly for $x, y \in \{0,1\}^n$).

2.1 Model

We use the familiar model of communication complexity, augmented by the notion of correlated shared randomness. Recall that in the standard model, two players, Alice and Bob, have access to inputs x and y respectively. A protocol Π specifies the interaction between Alice and Bob (who speaks when and what), and concludes with Alice and Bob producing outputs w_A and w_B respectively. A communication problem P is (informally) specified by conditions on the inputs and outputs (x, y, w_A, w_B). In usual (promise) problems this is simply a relationship on the 4-tuple. In sampling problems, this may be given by requirements on the distribution of this output given x and y. For functional problems, $P = (f_A, f_B)$ and the conditions require that $w_A = f_A(x,y)$ and $w_B = f_B(x,y)$. A randomized protocol is said to solve a functional problem P if the outputs are correct with probability at least $2/3$. The (worst-case) complexity of a protocol Π, denoted $\mathrm{cc}(\Pi)$ is the maximum over all x, y of the expected number of bits communicated by Π. This is the main complexity measure of interest to us, although distributional complexity will also be considered, as also any mix. (For instance, the most natural measure in compression is a max-average measure.)

We will be considering the setting where Alice and Bob have access to an arbitrarily long sequence of correlated random bits. For this definition it will be convenient to let a random bit be an element of $\{+1,-1\}$. For $\rho \in [-1,+1]$, we say a pair of bits (a,b) are *ρ-correlated (uniform) bits* if $\mathbb{E}[a] = \mathbb{E}[b] = 0$ and $\mathbb{E}[ab] = \rho$. We will consider the performance of protocols when given access to sequences (r, r') where each coordinate pair (r_i, r'_i) are ρ-correlated uniform bits chosen independently for each i. We shall write $r \sim_\rho r'$ for such ρ-correlated pairs.

The *communication complexity of a problem* P with access to ρ-correlated bits, denoted[2] $\text{isr-cc}_\rho(P)$ is the minimum over all protocols Π that solve P with access to ρ-correlated bits of $\mathrm{cc}(\Pi)$. For integer k, we let $\mathsf{ISR\text{-}CC}_\rho(k)$ denote the collections of problems P with $\text{isr-cc}_\rho(P) \le k$. The one-way communication complexity and simultaneous message complexities are defined similarly (by restricting to appropriate protocols) and denoted $\text{isr-cc}^{\mathrm{ow}}_\rho(P)$ and $\text{isr-cc}^{\mathrm{sm}}_\rho(P)$ respectively. The corresponding complexity classes are denoted similarly by $\mathsf{ISR\text{-}CC}^{\mathrm{ow}}_\rho(k)$ and $\mathsf{ISR\text{-}CC}^{\mathrm{sm}}_\rho(k)$.

Note that when $\rho = 1$ we get the standard model of communication with shared randomness. We denote this measure by $\text{psr-cc}(P) = \text{isr-cc}_1(P)$, and write $\mathsf{PSR\text{-}CC}(k)$ for the corresponding complexity class. Similarly, when $\rho = 0$ we get communication complexity with private randomness $\text{private-cc}(P) = \text{isr-cc}_0(P)$. We note that $\text{isr-cc}_\rho(P)$ is non-increasing in ρ. Combined with Newman's Theorem [14], we obtain:

PROPOSITION 1. *For every problem P with inputs $x, y \in \{0,1\}^n$ and $0 \le \rho \le \rho' \le 1$ we have*

$$\begin{aligned}\text{psr-cc}(P) \le \text{isr-cc}_{\rho'}(P) &\le \text{isr-cc}_\rho(P)\\ &\le \text{private-cc}(P) \le \text{psr-cc}(P) + O(\log n).\end{aligned}$$

The proposition also holds for one-way communication, and (except for the last inequality) simultaneous messages.

2.2 Problems, Results and Techniques

We now define some of the new problems we consider in this work and describe our main results.

2.2.1 Compression

Definition 1. For $\delta > 0$, $\Delta \ge 0$ and integers ℓ, n, the *uncertain compression problem* $\textsc{Compress}^{\ell,n}_{\Delta,\delta}$ is a promise problem with Alice getting as input the pair (P, m), where $P = (P_1,\ldots,P_n)$ is a probability distribution on $[n]$ and $m \in [n]$. Bob gets a probability distribution Q on $[n]$. The promises are that $H(P) \le \ell$ and for every $i \in [n]$, $|\log(P_i/Q_i)| \le \Delta$. The goal is for Bob to output m, i.e., $w_B = m$ with probability at least $1-\delta$. The measure of interest here is the maximum, over (P,Q) satisfying the promise, of the expected one-way communication complexity when m is sampled according to P.

When $\Delta = 0$, this is the classical compression problem and Huffman coding achieves a compression length of at most $\ell + 1$; and this is optimal for "prefix-free" compressions. For larger values of Δ, the work of [9] gives an upper bound of $\ell + 2\Delta + O(1)$ in the setting of perfectly shared randomness (to get constant error probability). In the setting of deterministic communication or private randomness, it is open if this communication complexity can be bounded by a function of ℓ and Δ alone (without dependence on n). (The work of [6] studies the deterministic setting.) Our first result shows that the bound of [9] can be extended naturally to the setting of imperfectly shared randomness.

THEOREM 1 (). *For every $\epsilon, \delta > 0$ and $0 < \rho \le 1$ there exists $c = c_{\epsilon,\delta,\rho}$ such that for every ℓ, n, we have* $\text{isr-cc}^{\mathrm{ow}}_\rho\left(\textsc{Compress}^{\ell,n}_{\Delta,\delta}\right) \le \frac{1+\epsilon}{1-h((1-\rho)/2)}(H(P) + 2\Delta + c)$.

We stress that the notation $\text{isr-cc}^{\mathrm{ow}}_\rho\left(\textsc{Compress}^{\ell,n}_{\Delta,\delta}\right)$ describes the *worst-case* complexity over P with entropy $H(P) \le \ell$ of the *expected* compression length when $m \leftarrow P$. The protocol that achieves this bound is a simple modification of the protocol of [9]. Roughly, Alice and Bob use their correlated randomness to define a "redundant and ambiguous dictionary" with words of every length for every message. Alice communicates using a word of appropriate length given the

[2] All throughout "isr" stands for *imperfect shared randomness*, while *psr* refers to *perfect* shared randomness.

distribution P, and Bob decodes using maximum likelihood decoding given Q. The main difference in our case is that Alice and Bob work knowing their dictionaries do not match exactly (as if they spelled the same words differently) and so use even longer words during encoding and decoding with some error-correction to allow for spelling errors. Details can be found in the full version [3].

2.2.2 Agreement distillation

Next we turn to a very natural problem in the context of imperfect sharing of randomness. Can Alice and Bob communicate to distill a few random bits from their large collection r and r' (of correlated random bits), bits on which they can agree perfectly?

Definition 2. In the AGREEMENT-DISTILLATION$^k_\gamma$ problem, Alice and Bob have no inputs. Their goal is to output w_A and w_B satisfying the following properties:

(i) $\Pr[\, w_A = w_B \,] \geq \gamma$;

(ii) $H_\infty(w_A) \geq k$; and

(iii) $H_\infty(w_B) \geq k$

where $H_\infty(X) = \min_x \log \frac{1}{\Pr[\, X=x \,]}$.

The (slightly) special case of this problem where Alice and Bob are not allowed to communicate at all was considered by Bogdanov and Mossel [2]. The setting where some communication is allowed is closely related but we describe the results in our language anyway.

A trivial way to distill randomness would be for Alice to toss random coins and send their outcome to Bob. This would achieve $\gamma = 1$ and communication complexity of k for k bits of entropy. Our first proposition says that with non-trivial correlation, some savings can always be achieved over this naive protocol.

PROPOSITION 2. *For every $\rho > 0$, it is the case that* isr-cc$^{\text{ow}}_\rho$(AGREEMENT-DISTILLATION$^k_\gamma$) $= (h(\frac{1-\rho}{2}) + o_k(1)) \cdot k$ *with $\gamma = 1 - o_k(1)$. In particular, for every $\rho > 0$ there exists $\alpha < 1$ such that for every sufficiently large k, we have* isr-cc$^{\text{ow}}_\rho$(AGREEMENT-DISTILLATION$^k_{1/2}$) $\leq \alpha k$.

We prove this proposition in the full version [3]. We note that this proposition is similar in spirit to Theorem 4 of [2] with the difference that they use their "better strategy" to improve γ with zero-communication, but γ remains exponentially small. We use communication to increase γ close to 1.

Our next theorem says that these linear savings are the best possible: one cannot get away with $o(k)$ communication unless $\rho = 1$. This theorem follows immediately from Theorem 1 of [2] and a simple reduction that converts protocols with communication to zero-communication protocols with a loss in γ.

THEOREM 2. *For every $\rho > 0$ there exists $\epsilon > 0$ such that* isr-cc$_\rho$(AGREEMENT-DISTILLATION$^k_\gamma$) $\geq \epsilon k - \log \frac{1}{\gamma}$.

The full version [3] contains details of this proof.

2.2.3 General relationships between perfect and imperfect sharing

Our final target in this work is to get some general relationships for communication complexity in the settings of perfect and imperfectly shared randomness. Our upper bounds for communication complexity are obtained by considering a natural promise problem, that we call GAPINNERPRODUCT, which is a "hard problem" for communication complexity. We use a variant, SPARSEGAPINNERPRODUCT, for our lower bounds. We define both problems below.

Definition 3. The GAPINNERPRODUCT$^n_{c,s}$ problem has parameters $n \in \mathbb{Z}^+$ (dimension), and $c > s \in [0,1]$ (completeness and soundness). Both yes- and no-instances of this problem have inputs $x, y \in \{0,1\}^n$. An instance (x, y) is a yes-instance if $\langle x, y \rangle \geq cn$, and a no-instance if $\langle x, y \rangle < sn$. The SPARSEGAPINNERPRODUCT$^n_{q,c,s}$ is a restriction of GAPINNERPRODUCT$^n_{c,s}$ where both the yes- and the no-instances are sparse, i.e., $\|x\|_2^2 \leq n/q$.

In the full version [3] we show that GAPINNERPRODUCT$^n_{c,s}$ is "hard" for PSR-CC(k) with $c = (2/3)2^{-k}$ and $s = (1/3)2^{-k}$. Then we show that this problem is in ISR-CC$^{\text{ow}}_\rho$(poly$(1/(c-s))$). Putting the two results together we get the following theorem giving a general upper bound on isr-cc$^{\text{ow}}_\rho(P)$ in terms of psr-cc(P) for any promise problem P.

THEOREM 3. *$\forall \rho > 0$, $\exists c < \infty$ such that $\forall k$, we have* PSR-CC$(k) \subseteq$ ISR-CC$^{\text{ow}}_\rho(c^k)$.

We prove this theorem in the full version [3].

Theorem 3 is obviously tight already because of known gaps between one-way and two-way communication complexity. For instance, it is well known that the "indexing" problem (where Alice gets a vector $x \in \{0,1\}^n$ and Bob an index $i \in [n]$ and they wish to compute x_i) has one-way communication complexity of $\Omega(n)$ with perfectly shared randomness, while its deterministic two-way communication complexity is at most $\log n + 2$. However one could hope for tighter results capturing promise problems P with low psr-cc$^{\text{ow}}(P)$, or to give better upper bounds on isr-cc(P) for P with low psr-cc(P). Our next theorem rules out any further improvements to Theorem 3 when n is sufficiently large (compared to k). We do so by focusing on the problem SPARSEGAPINNERPRODUCT. In the full version [3] we show that psr-cc$^{\text{ow}}$(SPARSEGAPINNERPRODUCT$^n_{q,c,s}$) $= O(\text{poly}(\frac{1}{q(c-s)}) \log q)$ for every q, n and $c > s$. In particular if say $c = 1/(2q)$ and $s = 1/(4q)$ the one-way communication complexity with perfectly shared randomness reduces to $O(\log q)$, in contrast to the poly(q) upper bound on the one-way communication complexity with imperfectly shared randomness.

Our main technical theorem shows that this gap is necessary for every $\rho < 1$. Specifically in the full version we show that

$$\text{isr-cc}_\rho(\text{SPARSEGAPINNERPRODUCT}^n_{q,c=.9/q,s=.6/q}) = \Omega(\sqrt{q}) \,.$$

Putting the two together we get a strong converse to Theorem 3, stated below.

THEOREM 4. *For every k, there exists a promise problem $P = (P_n)_{n \in \mathbb{Z}^+}$ such that* psr-cc$^{\text{ow}}(P) \leq k$, *but for every $\rho < 1$ it is the case that* isr-cc$_\rho(P) = 2^{\Omega_\rho(k)}$.

Remarks on the proofs.

Theorem 3 and Theorem 4 are the technical highlights of this paper and we describe some of the ideas behind them here.

Theorem 3 gives an upper bound for isr-cc$^{\text{ow}}_{\rho}$ for problems with low psr-cc. As such this ought to be somewhat surprising in that for known problems with low probabilistic communication complexity (notably, equality testing), the known solutions are very sensitive to perturbations of the randomness. But the formulation in terms of GapInner-Product suggests that any such problem reduces to an approximate inner product calculation; and the theory of metric embeddings, and examples such as locality sensitive hashing, suggest that one can reduce the dimensionality of the problems here significantly and this may lead to some reduced complexity protocols that are also robust to the noise of the ρ-correlated vectors. This leads us to the following idea: To estimate $\langle x, y\rangle$, where $x, y \in \{0,1\}^n$, Alice can compute $a = \langle g_1, x\rangle$ where g_1 is a random n-dimensional spherical Gaussian and send a (or the most significant bits of a) to Bob. Bob can compute $b = \langle g_2, y\rangle$ and $a \cdot b$ is an unbiased estimator (up to normalization) of $\langle x, y\rangle$ if $g_1 = g_2$. This protocol can be easily shown to be robust in that if g_2 is only ρ-correlated with g_1, $a \cdot b$ is still a good estimator, with higher variance. And it is easy to convert a collection of ρ-correlated bits to ρ-correlated Gaussians, so it is possible for Alice and Bob to generate the g_1 and g_2 as desired from their imperfectly shared randomness. A careful analysis (of a variant of this protocol) shows that to estimate $\langle x, y\rangle$ to within an additive error $\epsilon\|x\|_2\|y\|_2$, it suffices for Alice to send about $1/\epsilon^2$ bits to Bob, and this leads to a proof of Theorem 3.

Next we turn to the proof of Theorem 4, which shows a roughly matching lower bound to Theorem 3 above. The insight to this proof comes from examining the "Gaussian protocol" above carefully and contrasting it with the protocol used in the perfect randomness setting. In the latter case Alice uses the randomness to pick one (or few) coordinates of x and sends some function of these bits to Bob achieving a communication complexity of roughly $\log(1/\epsilon)$, using the fact that only $O(\epsilon n)$ bits of x are non-zero. In the Gaussian protocol Alice sends a very "non-junta"-like function of x to Bob; this seems robust to the perturbations of the randomness, but leads to $1/\epsilon^2$ bits of communication. This difference in behavior suggests that perhaps functions where variables have low "influence" cannot be good strategies in the setting of perfect randomness, and indeed we manage to prove such a statement in the full version of this paper (see Theorem 6.8 in [3]). The proof of this theorem uses a variant of the invariance principle that we prove (see Theorem 7.1 in [3]), which shows that if a communication protocol with low-influences works in a "product-distributional" setting, it will also work with inputs being Gaussian and with the same moments. This turns out to be a very useful reduction. The reason that SparseGapInnerProduct has nice psr-cc$^{\text{ow}}$ protocols is the asymmetry between the inputs of Alice and the inputs of Bob — inputs of Alice are sparse! But with the Gaussian variables there is no notion of sparsity and indeed Alice and Bob have symmetric inputs and so one can now reduce the "disjointness" problem from communication complexity (where now Alice and Bob hold sets $A, B \subseteq [1/\epsilon]$, and would like to distinguish $|A \cap B| = 0$ from $|A \cap B| = 1$) to the Gaussian inner product problem. Using the well-known lower bound on disjointness, we conclude that $\Omega(1/\epsilon)$ bits of communication are necessary and this proves Theorem 6.8 in [3].

Of course, all this rules out only one part of the solution space for the communication complexity problem, one where Alice and Bob use functions of low-influence. To turn this into a general lower bound we note that if Alice and Bob use functions with some very influential variables, then they should agree on which variable to use (given their randomness r and r'). Such agreement on the other hand cannot happen with too high a probability by our lower bound on Agreement-Distillation (from Theorem 2). Putting all these ingredients together gives us a proof of Theorem 4.

The full version of this paper [3] contains proofs of all the theorems mentioned above as also the description of an invariance principle suitable for communication complexity.

3. CONCLUSIONS

In this paper we carried out an investigation of the power of imperfectly shared randomness in the context of communication complexity. There are two important aspects to the perspective that motivated our work: First, the notion that in many forms of natural communication, the communicating parties understand each other (or "know" things about each other) fairly well, but never perfectly. This imperfection in knowledge/understanding creates an obstacle to many of the known solutions and new solutions have to be devised, or new techniques need to be developed to understand whether the obstacles are barriers. Indeed for the positive results described in this paper, classical solutions do not work and the solutions that ended up working are even "provably" different from classical solutions. (In particular they work hard to preserve "low influence").

However, we also wish to stress a second aspect that makes the problems here interesting in our view, which is an aspect of scale. Often in communication complexity our main motivation is to compute functions with sublinear communication, or prove linear lower bounds. Our work, and natural communication in general, stresses the setting where inputs are enormous, and the communication complexity one is considering is tiny. This models many aspects of natural communication where there is a huge context to any conversation which is implicit. If this context were known exactly to sender and receiver, then it would play no significant mathematical role. However in natural communication this context is not exactly known, and resolving this imperfection of knowledge before communicating the relevant message would be impossibly hard. Such a setting naturally motivates the need to study problems of input length n, but where any dependence on n in the communication complexity would be impractical.

We note that we are not at the end of the road regarding questions of this form: Indeed a natural extension to communication complexity might be where Alice wishes to compute $f_A(x, y)$ and Bob wishes to compute $f_B(x, y)$ but Alice does not know f_B and Bob does not know f_A (or have only approximate knowledge of these functions). If x and y are n-bits strings, f_A and f_B might require 2^n bits to describe and this might be the real input size. There is still a trivial upper bound of $2n$ bits for solving any such communication problem, but it would be interesting to study when, and what form of, approximate knowledge of f_A and f_B helps improve over this trivial bound.

Turning to the specific questions studied in this paper a fair number of natural questions arise that we have not been able to address in this work. For instance, we stuck to a specific and simple form of correlation in the randomness shared by Alice and Bob. One could ask what general forms of randomness (r, r') are equally powerful. In particular if the distribution of (r, r') is known to both Alice and Bob, can they convert their randomness to some form of correlation in the sense used in this paper (in product form with marginals being uniform)?

In the Agreement-Distillation problem the goal was for Alice and Bob to agree perfectly on some random string. What if their goal is only to generate more correlated bits than they start with? What is possible here and what are the limits?

In the study of perfectly shared randomness, Newman's Theorem [14] is a simple but powerful tool, showing that $O(\log n)$ bits of randomness suffice to deal with problems on n bit inputs. When randomness is shared imperfectly, such a randomness reduction is not obvious. Indeed for the problem of equality testing, the protocol of [1] uses 2^n bits of randomness, and our Gaussian protocol (which can solve this with one-way communication) uses $\text{poly}(n)$ bits. Do $O(\log n)$ bits of imperfectly shared randomness suffice for this problem? How about for general problems?

Finally almost all protocols we give for imperfectly shared randomness lead to two-sided error. This appears to be an inherent limitation (with some philosophical implications) but we do not have a proof. It would be nice to show that one-sided error with imperfectly shared randomness cannot lead to any benefits beyond that offered by private randomness.

Acknowledgments

We thank Brendan Juba for his helpful notes [8] on the invariance principle. We thank the anonymous referees for their valuable comments and pointers.

References

[1] Mohammad Bavarian, Dmitry Gavinsky, and Tsuyoshi Ito. On the role of shared randomness in simultaneous communication. In Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias, editors, *ICALP (1)*, volume 8572 of *Lecture Notes in Computer Science*, pages 150–162. Springer, 2014. ISBN 978-3-662-43947-0.

[2] Andrej Bogdanov and Elchanan Mossel. On extracting common random bits from correlated sources. *CoRR*, abs/1007.2315, 2010. URL `http://arxiv.org/abs/1007.2315`.

[3] Clément L. Canonne, Venkatesan Guruswami, Raghu Meka, and Madhu Sudan. Communication with imperfectly shared randomness. *CoRR*, abs/1411.3603, 2014. URL `http://arxiv.org/abs/1411.3603`.

[4] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley Publishing, New York, 1991.

[5] Venkatesan Guruswami, Johan Håstad, Rajsekar Manokaran, Prasad Raghavendra, and Moses Charikar. Beating the random ordering is hard: Every ordering CSP is approximation resistant. *SIAM J. Comput.*, 40 (3):878–914, 2011.

[6] Elad Haramaty and Madhu Sudan. Deterministic compression with uncertain priors. In *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science*, ITCS '14, pages 377–386, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2698-8. doi: 10.1145/2554797.2554832. URL `http://doi.acm.org/10.1145/2554797.2554832`.

[7] Marcus Isaksson and Elchanan Mossel. Maximally stable Gaussian partitions with discrete applications. *Israel Journal of Mathematics*, 189:347–396, June 2012.

[8] Brendan Juba. 18.177 course project: Invariance principles, 2009. URL `http://people.seas.harvard.edu/~bjuba/papers/18177-report.pdf`.

[9] Brendan Juba, Adam Tauman Kalai, Sanjeev Khanna, and Madhu Sudan. Compression without a common prior: an information-theoretic justification for ambiguity in language. In Bernard Chazelle, editor, *ICS*, pages 79–86. Tsinghua University Press, 2011. ISBN 978-7-302-24517-9.

[10] Bala Kalyanasundaram and Georg Schnitger. The probabilistic communication complexity of set intersection. *SIAM J. Discrete Math.*, 5(4):545–557, 1992. doi: 10.1137/0405044. URL `http://dx.doi.org/10.1137/0405044`.

[11] Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, New York, NY, USA, 2006. ISBN 9780521029834. URL `http://books.google.com/books?id=dHH7rdhKwzsC`.

[12] Michel Ledoux and Michel Talagrand. *Probability in Banach spaces: Isoperimetry and processes*. Springer, Berlin, 1991. ISBN 3540520139.

[13] Elchanan Mossel. Gaussian bounds for noise correlation of functions. *Geometric and Functional Analysis*, 19(6):1713–1756, 2010. ISSN 1016-443X. doi: 10.1007/s00039-010-0047-x. URL `http://dx.doi.org/10.1007/s00039-010-0047-x`.

[14] Ilan Newman. Private vs. common random bits in communication complexity. *Inf. Process. Lett.*, 39(2): 67–71, July 1991. ISSN 0020-0190. doi: 10.1016/0020-0190(91)90157-D. URL `http://dx.doi.org/10.1016/0020-0190(91)90157-D`.

[15] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014. ISBN 9781107038325. URL `http://books.google.com/books?id=5xlvAwAAQBAJ`.