2012 IEEE 53rd Annual Symposium on Foundations of Computer Science

New Limits to Classical and Quantum Instance Compression

Andrew Drucker
CSAIL, MIT
Cambridge, MA
andy.drucker@gmail.com

***Abstract*—Given an instance of a hard decision problem, a limited goal is to *compress* that instance into a smaller, equivalent instance of a second problem. As one example, consider the problem where, given Boolean formulas $\psi^1, \ldots, \psi^t$, we must determine if at least one ψ^j is satisfiable. An OR-*compression scheme* for SAT is a polynomial-time reduction that maps $(\psi^1, \ldots, \psi^t)$ to a string z, such that z lies in some "target" language L' if and only if $\bigvee_j [\psi^j \in \mathrm{SAT}]$ holds. (Here, L' can be arbitrarily complex.) AND-compression schemes are defined similarly. A compression scheme is *strong* if $|z|$ is polynomially bounded in $n = \max_j |\psi^j|$, independent of t.**

Strong compression for SAT seems unlikely. Work of Harnik and Naor (FOCS '06/SICOMP '10) and Bodlaender, Downey, Fellows, and Hermelin (ICALP '08/JCSS '09) showed that the infeasibility of strong OR-compression for SAT would show limits to instance compression for a large number of natural problems. Bodlaender et al. also showed that the infeasibility of strong AND-compression for SAT would have consequences for a different list of problems. Motivated by this, Fortnow and Santhanam (STOC '08/JCSS '11) showed that if SAT is strongly OR-compressible, then $\mathsf{NP} \subseteq \mathsf{coNP/poly}$. Finding similar evidence against AND-compression was left as an open question.

We provide such evidence: we show that strong AND- *or* OR-compression for SAT would imply *non-uniform, statistical zero-knowledge proofs* for SAT—an even stronger and more unlikely consequence than $\mathsf{NP} \subseteq \mathsf{coNP/poly}$. Our method applies against *probabilistic* compression schemes of sufficient "quality" with respect to the reliability and compression amount (allowing for tradeoff). This greatly strengthens the evidence given by Fortnow and Santhanam against probabilistic OR-compression for SAT. We also give variants of these results for the analogous task of *quantum instance compression*, in which a polynomial-time quantum reduction must output a quantum state that, in an appropriate sense, "preserves the answer" to the input instance.

The central idea in our proofs is to exploit the information bottleneck in an AND-compression scheme for a language L in order to fool a cheating prover in a proof system for $\overline{L}$. Our key technical tool is a new method to "disguise" information being fed into a compressive mapping; we believe this method may find other applications.

***Keywords*-instance compression; kernelization; polynomial hierarchy; quantum compression**

I. Introduction

Given an instance of a hard decision problem, we may hope to *compress* that instance into a smaller, equivalent instance, either of the same or of a different decision problem. Here we do not ask to be able to recover the original instance from the smaller instance; we only require that the new instance have the same (yes/no) *answer* as the original. Such *instance compression* may be the first step towards obtaining a solution; this has been a central technique in the theory of fixed-parameter-tractable algorithms [1], [2]. Strong compression schemes for certain problems would also have important implications for cryptography [3]. Finally, compressing an instance of a difficult problem may also be a worthwhile goal in its own right, since it can make the instance easier to store and communicate [3].

A natural goal is to design an efficient reduction that achieves compression on instances that are particularly "simple" in some respect. Toward this end, the versatile framework of *parametrized problems* [1] has been extensively used to study instance compression. A parametrized problem is a decision problem in which every instance has an associated *parameter value* k, giving some measure of the complexity of a problem instance.[1] As an example, one can parametrize a Boolean formula ψ by the number of distinct variables appearing in ψ. An ambitious goal for a parametrized problem P is to compress an arbitrary instance x of the decision problem for P into an equivalent instance x' of a second, "target" decision problem, where the output length $|x'|$ is bounded by a *polynomial* in $k = k(x)$. If P has such a reduction running in time $\mathrm{poly}(|x| + k)$, we say P is *strongly compressible*; we say P is *strongly self-compressible* if the target problem of the reduction is P itself. (A strong self-compression reduction is usually referred to as a *polynomial kernelization*.)

A. Previous work: results and motivation

Let VAR-SAT denote the Satisfiability problem for Boolean formulas, parametrized by the number of distinct variables in the formula. In their study of instance compression for NP-hard problems, Harnik and Naor [3] asked whether VAR-SAT is strongly compressible.[2] They showed that a *deterministic* strong compression reduction for VAR-SAT (with any target problem) would yield a construction of collision-resistant hash functions based on any one-way function—a long-sought goal.

[1] The parameter k is explicitly given as part of the input to the algorithm.
[2] Strictly speaking, they asked a slightly different question whose equivalence to this one was pointed out in [4].

0272-5428/12 $26.00 © 2012 IEEE
DOI 10.1109/FOCS.2012.71

In fact, Harnik and Naor showed that for their applications, it would suffice to achieve strong compression for a simpler parametrized problem, the "OR(SAT) *problem:"* this is the Satisfiability problem for Boolean formulas expressed as disjunctions $\psi = \bigvee_{j=1}^{t} \psi_j$, where the parameter is now defined as the maximum bit-length of any subformula ψ_j. Strong compression for VAR-SAT easily implies strong compression for OR(SAT). Harnik and Naor defined a hierarchy of decision problems called the "VC hierarchy," which can be modeled as a class of parametrized problems (see [4]). They showed that a strong compression reduction for any of the problems "above" OR(SAT) in this hierarchy would also imply strong compression for OR(SAT); this includes parametrized versions of natural problems like the Clique and Dominating Set problems. While Harnik and Naor's primary motivation was to *find* a strong compression scheme for OR(SAT) to use in their cryptographic applications, their work also provides a basis for showing *negative* results: in view of the reductions in [3], any evidence against strong compression for OR(SAT) is also evidence against strong compression for a variety of other parametrized problems.

In subsequent, independent work, Bodlaender, Downey, Fellows, and Hermelin [5] also studied the compressibility of OR(SAT) and of related problems; these authors' motivations came from the theory of *fixed-parameter tractable (FPT)* algorithms [1]. A strong self-compression reduction for P provides the basis for an FPT algorithm for P: on input x, first compress x, then solve the equivalent, compressed instance. This is one of the most widely-used schemas for developing FPT algorithms.

Strong self-compression reductions are known for parametrized versions of many natural NP-complete problems, such as the Vertex Cover problem; see, e.g., the survey [2]. However, for many other such parametrized problems, including numerous problems known to admit FPT algorithms (such as OR(SAT)), no strong compression reduction is known. Bodlaender et al. [5] conjectured that no strong self-compression reduction exists for OR(SAT). They made a similar conjecture for the closely-related "AND(SAT) *problem,"* in which one is given Boolean formulas $\psi_1, \ldots, \psi_t$ and asked to decide whether $\bigwedge_{j=1}^{t}[\psi_j \in \mathrm{SAT}]$ holds—that is, whether every ψ_j is individually satisfiable. As with OR(SAT), we parametrize AND(SAT) by the maximum bit-length of any ψ_j.

Bodlaender et al. showed that these conjectures (sometimes referred to as the "OR-" and "AND-conjectures") would have considerable explanatory power. First, they showed [5, Theorem 1] that the nonexistence of strong self-compression reductions for OR(SAT) would rule out strong self-compression for a large number of other natural parametrized problems; these belong to a class we call "OR-expressive problems."[3] Under the assumption that AND(SAT) does not have strong self-compression, Bodlaender et al. ruled out strong self-compression reductions for a second substantial list of problems [5, Theorem 2], belonging to a class we call "AND-expressive." Despite the apparent similarity of OR(SAT) and AND(SAT), no equivalence between the compression tasks for these two problems is known.

In light of their results, Bodlaender et al. asked for complexity-theoretic evidence against strong self-compression for OR(SAT) and AND(SAT). Fortnow and Santhanam [4] provided the first such evidence: they showed that if OR(SAT) has a strong compression reduction (to any target problem), then $\mathsf{NP} \subseteq \mathsf{coNP/poly}$ and the Polynomial Hierarchy collapses to its third level.

The techniques of [5], [4] were refined and extended by many researchers to give further evidence against efficient compression for parametrized problems, e.g., in [7], [8], [9], [6], [10], [11], [12], [13], [14], [15]. (See [14] for further discussion and references.) As one notable development that is relevant to our work, Dell and Van Melkebeek [8] combined the techniques of [5], [4] with new ideas to provide tight compression-size lower bounds for certain problems that *do* admit polynomial kernelizations. Researchers also used ideas from [5], [4] in other areas of complexity, giving new evidence of lower bounds for the length of PCPs [4], [8] and for the density of NP-hard sets [16].

Finding evidence against strong compression for AND(SAT) was left as an open question by these works, however. The limits of *probabilistic* compression schemes for OR(SAT) and for OR-expressive problems (including VAR-SAT) also remained unclear. The results and techniques of [4] give evidence only against some restrictive sub-classes of probabilistic compression schemes for OR(SAT): schemes avoiding false negatives, or schemes whose error probability or randomness use is severely restricted.

B. Our results

1) Results on classical compression: We complement the results of [4] by providing evidence against strong compression for AND(SAT): we prove that such a compression scheme, to any target problem, would also imply $\mathsf{NP} \subseteq \mathsf{coNP/poly}$. Our techniques extend naturally (and in a strong fashion) to the *probabilistic* setting with two-sided error, in which we expect the compression reduction to obey some success-probability guarantee on every input. We show that any sufficiently "high-quality" compression scheme for AND(SAT) would imply $\mathsf{NP} \subseteq \mathsf{coNP/poly}$. Here, "quality" is defined by a certain relationship between the reliability

[3]The class of OR-expressive problems, defined in the full version of our paper, is not identical to the class described in [5], but it is closely related and contains their class, as well as other classes of problems identified in [3], [6].

and the compression amount of the reduction, and allows for tradeoff.

We also show that beyond a second, somewhat more demanding quality threshold, probabilistic compression reductions either for AND(SAT) *or* for OR(SAT) would imply the existence of *non-uniform, statistical zero-knowledge proofs* for NP languages—a stronger (and even more unlikely) consequence than NP $\subseteq$ coNP/poly. The more-demanding quality threshold in this second set of results is still rather modest, and allows us to prove the following result as a special case:

Theorem I.1 (Informal)**.** *Suppose that either of* AND(SAT) *or* OR(SAT) *is strongly compressible, with success probability* $\geq .5 + 1/\operatorname{poly}(n)$ *for an* AND *or* OR *of length-*n *formulas. Then there are non-uniform, statistical zero-knowledge proofs for all languages in* NP *(which implies* NP $\subseteq$ coNP/poly*).*

At the other extreme, where we consider compression schemes with more modest compression amounts, but with greater reliability, our techniques yield the following result:

Theorem I.2 (Informal)**.** *Let* $t(n) : \mathbb{N}^+ \to \mathbb{N}^+$ *be any polynomially bounded function. Suppose there is a compression scheme compressing an* AND *of* $t(n)$ *length-*n *SAT instances into an instance* z *of a second decision problem* L'*, where* $|z| \leq C \cdot t(n) \log t(n)$ *for some* $C > 0$*. If the scheme's error probability on such inputs is bounded by a sufficiently small inverse-polynomial in* n *(depending on* $t(n)$ *and* C*), then there are non-uniform, statistical zero-knowledge proofs for all languages in* NP*. The corresponding result also holds for* OR*-compression.*

Our results give the first strong evidence of hardness for compression of AND(SAT). They also greatly strengthen the evidence given by Fortnow and Santhanam against *probabilistic* compression for OR(SAT), and provide the first strong evidence against probabilistic compression for the potentially-harder problem VAR-SAT. Using our results on the infeasibility of compression for AND(SAT) and OR(SAT), and building on [3], [5], [4], we give new complexity-theoretic evidence against strong compressibility for a list of interesting parametrized problems with FPT algorithms. This is the first strong evidence against strong compressibility for any of the ten "AND-expressive" problems identified in [5]. For the numerous "OR-expressive" problems identified in [3], [5] and other works, this strengthens the negative evidence given by [4].

Our methods also extend the known results on limits to compression for parametrized problems that do possess polynomial kernelizations: we can partially extend the results of Dell and Van Melkebeek [8] to the case of probabilistic algorithms with two-sided error. For example, for $d > 1$ and any $\varepsilon > 0$, Dell and Van Melkebeek proved that if the Satisfiability problem for N-variable d-CNFs has a polynomial-time compression reduction with output-size bound $O(N^{d-\varepsilon})$, then NP $\subseteq$ coNP/poly. Their result applies to co-nondeterministic reductions, and to probabilistic reductions without false negatives; we prove that the result also holds for probabilistic reductions with two-sided error, as long as the success probability of the reduction is at least $.5 + N^{-\beta}$ for some $\beta = \beta(d, \varepsilon) > 0$. Using reductions described in [8], we also obtain quantitatively-sharp limits to probabilistic compression for several other natural NP-complete problems, including the Vertex Cover and Clique problems on graphs and hypergraphs. (However, the limits we establish do not give lower bounds on the cost of *oracle communication protocols*; these protocols are a generalization of compression reductions, studied in [8], to which that work's results do apply.)

Our results about AND(SAT) and OR(SAT) follow from more general results about arbitrary languages. For any language L, we follow previous authors and consider the "OR(L) *problem,*" where one is given a collection $x^1, \ldots, x^t$ of strings, and asked to determine whether at least one of them is a member of L. We show that if a sufficiently "high-quality" probabilistic poly-time compression reduction exists for the OR(L) problem, then $L \in$ NP/poly. We also show that a poly-time compression scheme for OR(L) meeting a more demanding standard of quality implies that L possesses non-uniform statistical zero-knowledge proof systems, and lies in NP/poly $\cap$ coNP/poly. (For deterministic compression, the conclusion $L \in$ coNP/poly was established earlier in [4].) Applying these results to $L := \overline{\text{SAT}}$ gives our hardness-of-compression results for AND(SAT); applying the second set of results to $L := \text{SAT}$ gives our improved negative results for OR(SAT).

2) Results on quantum compression: Up to this point we've discussed compression reductions in which the input and output are both classical bit-strings. But from the perspective of quantum computing and quantum information [17], it's natural to ask about the power of reductions that output a *quantum state* over some number of quantum bits, or "qubits." If quantum computers become a practical reality, quantum instance compression schemes could help to store and transmit hard computational problems; compressing an instance might also be a first step towards its solution by a quantum algorithm.

We propose a quantum generalization of classical instance compression: a *quantum compression reduction* for a language L is a quantum algorithm that, on input x, outputs a quantum state ρ on some number q of qubits—hopefully with $q \ll |x|$, to achieve significant compression. Our correctness requirement is that there should exist *some* quantum measurement $\mathcal{M}_q$, depending only on q, such that for every x compressing to q qubits, $\mathcal{M}_q(\rho) = L(x)$ holds with high probability over the inherent randomness in the measurement $\mathcal{M}_q(\rho)$. We do not require that $\mathcal{M}_q$ be an efficiently-performable measurement; this is by analogy to

the general version of the classical compression task, in which the target language of the reduction may be arbitrary.

Our results for quantum compression are analogous to our results in the classical case. First, we show that for any language L, if a sufficiently "high-quality" quantum polynomial-time compression reduction exists for the $\mathrm{OR}(L)$ problem, then L possesses a *non-uniform, 2-message quantum interactive proof system* (with a single prover). Second, we show that a sufficiently higher-quality quantum polynomial-time compression reduction for $\mathrm{OR}(L)$ implies that L possesses a *non-uniform quantum statistical zero-knowledge proof system*. Remarkably, the two "quality thresholds" in our quantum results are essentially *the same* as in the corresponding results for the classical case.[4] It follows that, unless there exist surprisingly powerful quantum proofs of unsatisfiability for Boolean formulas, the limits we establish for probabilistic compression of $\mathrm{AND}(\mathrm{SAT})$ and $\mathrm{OR}(\mathrm{SAT})$ hold just as strongly for quantum compression.[5]

C. Our techniques

1) The overall approach: We first describe our techniques for the classical case; these form the basis for the quantum case as well. Our two general results, giving complexity upper bounds on any language L for which $\mathrm{OR}(L)$ has a sufficiently high-quality compression reduction, are both based on a single reduction that we describe next. This reduction applies to compression reductions mapping some number $t(n) \leq \mathrm{poly}(n)$ of inputs of length n to an output string z of length $|z| = O(t(n) \log t(n))$.

Fix any language L such that $\mathrm{OR}(L)$ has a possibly-probabilistic compression reduction

$$R(x^1, \ldots, x^t) : \ \{0,1\}^{t \times n} \ \longrightarrow \ \{0,1\}^{\leq t'} \ ,$$

with some target language L', along with parameters t', t satisfying $t' \leq O(t \log t) \leq \mathrm{poly}(n)$. We will use R to derive upper bounds on the complexity of L.

A simple, motivating observation is that if we take a string $y \in L$ and "insert" it into a tuple $\overline{x} = (x^1, \ldots, x^t)$ of elements of $\overline{L}$, replacing some x^j to yield a modified tuple $\overline{x}'$, then the values $R(\overline{x}), R(\overline{x}')$ are *different* with high probability—for, by the "OR-respecting" property of R, we will (with high probability) have $R(\overline{x}) \in \overline{L'}$, $R(\overline{x}') \in L'$. More generally, for any *distribution* $\mathcal{D}$ over t-tuples of inputs from $\overline{L}$, let $\mathcal{D}[y, j]$ denote the distribution obtained by sampling $\overline{x} \sim \mathcal{D}$ and replacing x^j with y; then the two output distributions $R(\mathcal{D})$, $R(\mathcal{D}[y, j])$ are *far apart* in statistical distance. (Of course, the strength of the statistical-distance lower bound we get will depend on the reliability of our compression scheme.)

We want this property to serve as the basis for an interactive proof system by which a computationally powerful Prover can convince a skeptical polynomial-time (but non-uniform) Verifier that a string y lies in L. The idea for our initial, randomized protocol (which we will later derandomize) is that Prover will make his case by showing his ability to *distinguish* between the two R-output distributions described above, when Verifier privately chooses one of the two distributions, samples from it, and sends the sample to Prover.[6] But to make our proof system meaningful, Verifier also needs to *fool* a cheating Prover in the case $y \notin L$. Thus, we want to choose $\mathcal{D}, j$ in such a way that the distributions $R(\mathcal{D}), R(\mathcal{D}[y, j])$ are as close as possible whenever $y \notin L$.

We may not be able to achieve this for an index j that is poorly-chosen; to avoid a bad choice, we choose j *uniformly at random*. The compression scheme R doesn't have room in its output string to copy its entire input, so there is reason for hope. This invites us to search for a distribution $\mathcal{D}^*$ over $\left(\overline{L}_n\right)^t$ with the following properties:

(i) For every $y \in \overline{L}_n$, for uniform j the value $\mathbb{E}_j \left[||R(\mathcal{D}^*) - R(\mathcal{D}^*[y, j])||_{\mathrm{stat}} \right]$ is "not too large;"[7]

(ii) $\mathcal{D}^*$ is efficiently sampleable, given non-uniform advice of length $\mathrm{poly}(n)$.

Condition (i) is quite demanding: we need a single distribution $\mathcal{D}^*$ rendering R insensitive to the insertion of *any* string $y \in \overline{L}_n$. Condition (ii) is also demanding: $\overline{L}_n$ may be a complicated set, and in general we can only hope to sample from distributions over $\left(\overline{L}_n\right)^t$ in which t-tuples are formed out of a fixed "stockpile" of $\mathrm{poly}(n)$ elements of $\overline{L}_n$, hard-coded into the non-uniform advice.

Remarkably, it turns out that such a distribution $\mathcal{D}^*$ can always be found. In item (i), we can force the two distributions to be non-neglibly close (with expected statistical distance $\leq 1 - \frac{1}{\mathrm{poly}(n)}$) whenever the output-size bound t' obeyed by R is $O(t \log t)$; the distributions will be much closer when $t' \ll t$. We call our key result, guaranteeing the existence of such a $\mathcal{D}^*$, the *"Disguising-Distribution Lemma."*

Assuming this lemma for the moment, we use $\mathcal{D}^*$ as above to reduce any membership claim for L to a distinguishing task for a Prover-Verifier protocol. Given any input y, we've constructed two distributions $\mathfrak{R} = R(\mathcal{D}^*)$ and $\mathfrak{R}' = R(\mathcal{D}^*[y, \mathbf{j}])$ (with $\mathbf{j}$ uniform), where each distribution is sampleable in non-uniform polynomial time. Our analysis guarantees some lower bound $D = D(n)$ on $||\mathfrak{R} - \mathfrak{R}'||_{\mathrm{stat}}$ in the case $y \in L$, and some upper bound $d = d(N)$ on this

[4] We do place a minor additional restriction on quantum compression reductions for $\mathrm{OR}(L)$: we require that the reduction, on input $(x^1, \ldots, x^t)$, outputs a quantum state of size determined by $(\max_j |x^j|)$ and t.

[5] We remark that *3-message* quantum interactive proofs are known to be fully as powerful as quantum interactive proofs in which polynomially many messages are exchanged [18], and that these proof systems are equal in power to PSPACE in the uniform setting [19]. However, 2-message quantum proof systems seem much weaker, and are not known to contain coNP.

[6] Such distinguishing tasks have seen many uses in theoretical computer science, and we rely upon known protocols of this kind in our work.

[7] For our purposes, it actually suffices to bound $||R(\mathcal{D}^*) - R(\mathcal{D}^*[y, \mathbf{j}])||_{\mathrm{stat}}$, where $\mathbf{j}$ is a uniform value sampled "internally" as part of the distribution. However, our techniques will yield the stronger property in condition (i) above.

distance when $y \notin L$. (These parameters depend on the reliability and compression guarantees of R.) If $D(n) - d(n) \geq \frac{1}{\mathrm{poly}(n)}$, we can give non-uniform distinguishing protocols for L, which can converted to public-coin protocols and then non-uniformly derandomized to show that $L \in \mathsf{NP/poly}$. If, more strongly, $D(n)^2 - d(n) \geq \frac{1}{\mathrm{poly}(n)}$ then, using a powerful result due to Sahai and Vadhan [20], we can derive a non-uniform, statistical zero-knowledge proof system for L. This also implies $L \in \mathsf{NP/poly} \cap \mathsf{coNP/poly}$.

2) The Disguising-Distribution Lemma: The Disguising-Distribution Lemma is a "generic" result about the behavior of compressive mappings; it uses no properties of R other than R's output-size bound. In view of its generality and interest, we are hopeful that the lemma will find other applications.

Our proof of this lemma uses two central ideas. First, we interpret the search for the "disguising distribution" $\mathcal{D}^*$ as a two-player game between a "disguising player" (choosing $\mathcal{D}^*$) and an opponent who chooses y; we can then apply simple yet powerful principles of game theory. Second, to build a winning strategy for the disguising player, we will exploit an information bottleneck in R stemming from its compressive property.[8]

To describe the proof, it is helpful to first understand how one may obtain the distribution $\mathcal{D}^*$ if we drop the efficient-sampleability requirement on $\mathcal{D}^*$, and focus on the "disguising" requirement (condition (i)). To build $\mathcal{D}^*$ in this relaxed setting, we will appeal to the *minimax theorem* for two-player, zero-sum games; applied here, it tells us that to guarantee the existence of a $\mathcal{D}^*$ that succeeds in disguising all strings $y \in \overline{L}_n$, it is enough to show how to build a $\mathcal{D}^*_Y$ that succeeds in expectation, when y is sampled from some fixed (but arbitrary) distribution Y over $\overline{L}_n$.

Here, a natural idea springs to mind: let $\mathcal{D}^*_Y$ just be a product distribution over t copies of Y! In this case, inserting $y \sim Y$ into $\mathcal{D}^*_Y$ at a random location is equivalent to conditioning on the outcome of a randomly-chosen coordinate of a sample from $\mathcal{D}^*_Y$. The intuition here is that, due to the output-size bound on R, the distribution $R(\mathcal{D}^*_Y)$ shouldn't have enough "degrees of freedom" to be affected much by this conditioning.

We show that for any distribution $\mathcal{D}$ over $\{0,1\}^n$, if $\overline{x} = (x^1, \ldots, x^t) \sim \mathcal{D}^{\otimes t}$ then conditioning on the value of x^j for a uniformly-chosen index $j \in [t]$ has bounded expected effect on the output distribution $R(\overline{x})$. That is, the *expected statistical distance* between the pre- and post-conditioned distributions is bounded non-negligibly away from 1 (provided that $t' \leq O(t \log t)$). We refer to this important property of R as *"distributional stability."*

In our original proof that our compressive mapping R is distributionally stable, we gave a simple (non-constructive) way to use R as a one-shot *encoding method* for independent, unbiased bits $b_1, \ldots, b_t$. The encoding Enc has a desirable property: for each component $j \in [t]$ whose expected "influence" on the output distribution of R is noticeable (when we fix a single value $x^j \sim \mathcal{D}$), our encoding transmits b_j with noticeable advantage over a random guess. We can then deduce strong upper bounds on the influence of a typical component j, using the output-size bound on R and elementary information-theoretic bounds on the reliability of compressive encodings. This analysis succeeds when $t' \leq t - 2$. In our original draft, we used more elaborate techniques (which involved modifying the mapping R itself) to analyze the case when $t \leq t' \leq O(t \log t)$.

Several researchers pointed out to us that the distributional stability property can be established in a different way, using Kullback-Leibler divergence and an inequality due to Pinsker. This approach allows us to analyze the case when $t' \leq (1 + \varepsilon)t$, for a modest $\varepsilon > 0$. As this author noted later, the divergence-based approach can be combined with an alternative to Pinsker's inequality—a bound due to Vajda (see [22], [23])—to show that the mapping R has a non-negligible amount of distributional stability as long as $t' \leq O(t \log t)$. Thus we feel that the divergence-based approach is ultimately the most convenient one to work with in general, and we will use it here. Colleagues also helped me to understand that the distributional stability property for mappings with $t' \leq (1 + \varepsilon)t$ can also be established using other similar, known results that follow from the same divergence/Pinsker-based techniques: a lemma of Raz [24], and the "Average Encoding Theorem" of Klauck et al. [25]. The latter was used in [25] to identify a stability property for trace and Hellinger distance metrics, for the inputs to a quantum communication problem. Their proof is for inputs drawn from the uniform distribution, but extends readily to general distributions and can be used to derive the kind of lemma we need. We feel all of these approaches to proving distributional stability are worth understanding.[9]

Using the distributional-stability property of compressive mappings under product input-distributions, we then establish a certain "sparsified variant" of this property, which allows us to replace each copy of $\mathcal{D}$ with a small set sampled from $\mathcal{D}$; this is an important tool in addressing the efficient-sampleability requirement on our desired $\mathcal{D}^*$. Armed with this variant, we apply the minimax theorem to show that there exists a *distribution* $\mathfrak{D}$ over product input-distributions to R—with each product distribution defined over small subsets of S—such that, in expectation, $\mathfrak{D}$ disguises the random insertion of any string $y \in S$ at a uniformly-

[8]This is hardly the first paper in which such a bottleneck plays a crucial and somewhat unexpected role. For example, an interesting and slightly similar application of information-theoretic tools to the study of *metric embeddings* was found recently by Regev [21].

[9]R. Impagliazzo suggested the use of Raz's lemma; S. Vadhan also helped me to understand the connection. A. Nayak and S. Vadhan suggested direct proofs of distributional stability based on divergence and Pinsker's inequality, which we now use as our main approach. D. van Melkebeek also suggested the relevance of Pinsker's inequality.

chosen position j. Finally, we obtain our desired "disguising distribution" $\mathcal{D}^*$ as a sparsified version of $\mathfrak{D}$, using a result due to Lipton and Young [26] and, independently, to Althöfer [27], that guarantees the existence of sparsely-supported, nearly-optimal strategies in 2-player, zero-sum games.

3) Extension to the quantum case: Our techniques for studying quantum compression, presented in the full version, are closely analogous to the classical case. The main technical difference is that the output $R(\mathcal{D})$ of our compression reduction, on any input distribution $\mathcal{D}$, is now a (mixed) *quantum state*. In this setting, to carry out an analogue of the argument sketched in Sections I-C1 and I-C2, we need a "disguising distribution" for R that meets a modified version of condition (i) from Section I-C1:

(i') For every $y \in \overline{L}_n$, if we select $j \in [t]$ uniformly then, *for any quantum measurement* $\mathcal{M}$, the expected statistical distance $\mathbb{E}_j\left[||\mathcal{M}(R(\mathcal{D}^*)) - \mathcal{M}(R(\mathcal{D}^*[y,j]))||_{\mathrm{stat}}\right]$ is not too large.

A basic measure of distance between quantum states, the *trace distance*, is relevant here: if two states ρ, ρ' are at trace distance $||\rho - \rho'||_{\mathrm{tr}} \leq \delta$, then for any measurement $\mathcal{M}$, the statistical distance $||\mathcal{M}(\rho) - \mathcal{M}(\rho')||_{\mathrm{stat}}$ is at most δ. (In fact, this property *characterizes* the trace distance.) Thus to satisfy condition (i'), it will be enough to construct $\mathcal{D}^*$ so as to upper-bound $\mathbb{E}_j[||R(\mathcal{D}^*) - R(\mathcal{D}^*[y,j])||_{\mathrm{tr}}]$, for uniformly-chosen j. We do this by essentially the same techniques as in the classical case. The one significant difference is that here, we need to establish a "stability property" for trace distance, analogous to the classical stability property for statistical distance. This can be obtained using the same basic divergence-based techniques as in the classical case, and following [25].[10]

4) A more "elementary" proof?: Researchers have asked whether a shorter proof of our most "basic" result, on the hardness of strong compression for AND(SAT), is possible. Motivated by this, we have found an alternative, quantitatively-weaker proof that resembles our original proof, but avoids using any information-theoretic tools. It also makes no recourse to the minimax theorem, instead taking a more incremental approach to defining the needed non-uniform advice. The resulting proof bears some resemblance to the work of Fortnow and Santhanam [4] on the hardness of compression for OR(SAT). See the full version.

II. Preliminaries

A. Probability and information theory background

All distributions in this paper will take finitely many values; let $\mathrm{supp}(\mathcal{D})$ be the set of values assumed by $\mathcal{D}$ and let $\mathcal{D}(u) := \Pr[\mathcal{D} = u]$. Let $\mathcal{D}^{\otimes t}$ denote a t-tuple of independent samples from $\mathcal{D}$. We let $\mathcal{U}_K$ denote the uniform distribution over a multiset K. The *statistical distance* of two distributions $\mathcal{D}, \mathcal{D}'$ over a shared universe of outcomes is defined as $||\mathcal{D} - \mathcal{D}'||_{\mathrm{stat}} := \frac{1}{2}\sum_{u \in \mathrm{supp}(\mathcal{D}) \cup \mathrm{supp}(\mathcal{D}')} |\mathcal{D}(u) - \mathcal{D}'(u)|$. We will use the following familiar "distinguishability interpretation" of the statistical distance. Suppose a value $b \in \{0,1\}$ is selected uniformly, unknown to us, and a sample $u \in U$ is drawn from $\mathcal{D}$ if $b = 0$, or from $\mathcal{D}'$ if $b = 1$. We observe u, and our goal is to correctly guess b. It is a basic fact that, for *any* $\mathcal{D}, \mathcal{D}'$, our maximum achievable success probability in this "distinguishing" experiment is precisely $\frac{1}{2}(1 + ||\mathcal{D} - \mathcal{D}'||_{\mathrm{stat}})$.

The *entropy* of a random variable X is defined as $H(X) := -\sum_{x \in \mathrm{supp}(X)} \Pr[X = x] \log_2(\Pr[X = x])$. The *mutual information* $I(X;Y)$ between r.v.'s X, Y is defined as $I(X;Y) := H(X) + H(Y) - H((X,Y))$. The following standard fact is proved in the full version:

Lemma II.1. *If $X^1, \ldots, X^t$ are independent, then*

$$I(Y; (X^1, \ldots, X^t)) \;\geq\; \sum_{j \in [t]} I(Y; X^j) \;.$$

The (binary) *Kullback-Leibler divergence*, or *KL divergence*, is a useful, non-symmetric measure of difference between random variables, defined as

$$D_{\mathrm{KL}}(X||Y) := \sum_{x \in \mathrm{supp}(X)} \Pr[X = x] \cdot \log_2\left(\Pr[X = x] / \Pr[Y = x]\right) \;.$$

Note that D_{KL} may be infinite. We have the following basic equivalence (see [29, Chapter 2]):

Fact II.2. *Let X, Y be any random variables; let X' be distributed as X and independent of Y. Then $I(X;Y) = D_{\mathrm{KL}}((X,Y)||(X',Y))$.*

A proof of the following important result can be found in [29] (see Lemma 11.6.1, p. 370).

Theorem II.3 (Pinsker's inequality)**.** *For any r.v.'s Z, Z', $D(Z||Z') \;\geq\; \frac{2}{\ln 2} \cdot ||Z - Z'||^2_{\mathrm{stat}}$.*

When $||Z - Z'||_{\mathrm{stat}} \approx 1$, the following bound, a slight weakening of Vajda's inequality (see [22], [23]), gives better information on the divergence:

Theorem II.4 (Vajda's inequality)**.** *For any r.v.'s Z, Z', $D(Z||Z') \;\geq\; \frac{1}{\ln 2}\left(\ln\left(\frac{1}{1 - ||Z - Z'||_{\mathrm{stat}}}\right) - 1\right)$.*

B. Statistical distance promise problems

We assume familiarity with promise problems, and with promise versions of standard complexity classes, including pr-AM, the promise version of the "Arthur-Merlin" complexity class AM. It is well-known (and follows from [30]) that pr-AM $\subseteq$ pr-NP/poly.

For a Boolean circuit $\mathcal{C}$ with one or more outputs, let $\mathcal{D}_C$ be the induced output distribution when $\mathcal{C}$ is fed a uniformly

[10] Our original approach to proving distributional stability also admits a quantum version, using Nayak's bound [28] on the reliability of quantum random access codes.

random input. For parameters $0 \leq d \leq D \leq 1$, define the promise problem $\mathrm{SD}^{\geq D}_{\leq d} = (\Pi_Y, \Pi_N)$ as follows:

$$\Pi_Y \ := \ \{\langle C, C' \rangle : ||\mathcal{D}_C - \mathcal{D}_{C'}||_{\mathrm{stat}} \geq D\} \ ,$$

$$\Pi_N \ := \ \{\langle C, C' \rangle : ||\mathcal{D}_C - \mathcal{D}_{C'}||_{\mathrm{stat}} \leq d\} \ .$$

In this definition, both $d = d(n)$ and $D = D(n)$ may be parameters depending on the input length $n = |\langle C, C' \rangle|$.

We will use the class pr-SZK of promise problems having *statistical zero-knowledge proofs*. It follows from a powerful result of Sahai and Vadhan [20] that this class can be defined as the set of promise problems having deterministic, polynomial-time many-to-one reductions to $\mathrm{SD}^{\geq 2/3}_{\leq 1/3}$.[11] pr-$\mathsf{SZK}$ is known to be contained in pr-$\mathsf{AM} \cap$ pr-$\mathsf{coAM} \subseteq$ pr-$\mathsf{NP/poly} \cap$ pr-$\mathsf{coNP/poly}$ [31], [32], and to be closed under complement [33]. We also have:

Theorem II.5. *Let $0 \leq d = d(n) < D = D(n) \leq 1$ be (not necessarily computable) parameters.*

1) *If $D > d + \frac{1}{\mathrm{poly}(n)}$, then $\mathrm{SD}^{\geq D}_{\leq d} \in$ pr-$\mathsf{NP/poly}$.*
2) *If we have the stronger gap $D^2 > d + \frac{1}{\mathrm{poly}(n)}$, then $\mathrm{SD}^{\geq D}_{\leq d}$ is many-to-one reducible to $\mathrm{SD}^{\geq 2/3}_{\leq 1/3} \in$ pr-SZK, in non-uniform polynomial time.*

Item 1 uses a standard distinguishing protocol and non-uniform derandomization; item 2 follows from [20] (and is described as Theorem 1 in [34]).

C. f-compression reductions

Our next definition is modeled on definitions in [5], [4].

Definition II.6 (Probabilistic f-compression reductions)**.** *Let L, L' be two languages, and let $f : \{0,1\}^* \to \{0,1\}$ be a Boolean function. Let $t_1(n), t_2(n) : \mathbb{N}^+ \to \mathbb{N}^+$ and $\xi(n) : \mathbb{N}^+ \to [0,1]$ be given.*

A probabilistic f-compression reduction for *L, with parameters $(t_1(n), t_2(n), \xi(n))$ and target language L', is a randomized mapping $R(x^1, \ldots, x^m)$ outputting a string z, such that for all $(x^1, \ldots, x^{t_1(n)}) \in \{0,1\}^{t_1(n) \times n}$,*

1) $\Pr_R[L'(z) = f\left(L(x^1), \ldots, L(x^{t_1(n)})\right)] \geq 1 - \xi(n)$;
2) $|z| \leq t_2(n)$.

If some reduction R as above is computable in probabilistic polynomial time, we say that L is PPT-f-compressible *with parameters $(t_1(n), t_2(n), \xi(n))$. (This does not require that $(t_1(n), t_2(n), \xi(n))$ themselves be computable.)*

III. Technical lemmas

In this section we present our main technical lemmas. Our final goal in this section will be the "Disguising-Distribution Lemma," our key technical tool for our main results.

[11] We don't use the fact that this problem is complete for pr-SZK, but it affords a convenient working definition.

A. Distributional stability

Here we define the notion of "distributional stability" described in Section I-C2.

Definition III.1. *Let U be some finite universe, and let $t, n \geq 1$ be integers. Given a possibly-randomized mapping $F(x^1, \ldots, x^t) : \{0,1\}^{t \times n} \to U$, and a distribution $\mathcal{D}$ over $\{0,1\}^n$, for $j \in [t]$ let*

$$\gamma_j \ := \ \mathop{\mathbb{E}}_{y \sim \mathcal{D}} \left[\left|\left| F\left(\mathcal{D}^{\otimes(j-1)}, y, \mathcal{D}^{\otimes(t-j)}\right) - F\left(\mathcal{D}^{\otimes t}\right) \right|\right|_{\mathrm{stat}} \right] .$$

*For $\delta \in [0,1]$, say that F is δ-*distributionally stable (or δ-DS) with respect to $\mathcal{D}$ *if $\frac{1}{T}\sum_{j=1}^{T} \gamma_j \leq \delta$.*

Lemma III.2. *Let $R(x^1, \ldots, x^t) : \{0,1\}^{t \times n} \to \{0,1\}^{\leq t'}$ be any possibly-randomized mapping, for any $n, t, t' \in \mathbb{N}^+$. R is δ-distributionally stable with respect to any input distribution $\mathcal{D}$ over $\{0,1\}^n$, where we may take either of the following two bounds:*

1) $\delta := \sqrt{\frac{\ln 2}{2} \cdot \frac{t'+1}{t}}$;
2) $\delta := 1 - 2^{-\frac{t'}{t} - 3}$.

When $t'/t = 1 - \Omega(1)$, the bound given in item 1 above is within constant factors of the bound from our original distributional stability lemma. When $t'/t = 1 - \alpha \approx 1$, the bound in Lemma III.2, item 1 is smaller by a $\Theta\left(\log \frac{1}{\alpha}\right)$ factor. We don't know how to prove a version of item 2 above with our original methods; this alternative bound is important for our work. In an earlier draft we used a more complicated workaround to prove the results we will obtain using Lemma III.2, item 2.

Proof of Lemma III.2: Define independent random variables $X^j \sim \mathcal{D}$ over $\{0,1\}^n$, for $j \in [t]$. Let $\mathbf{R} := R(X^1, \ldots, X^t)$.

We have $H(\mathbf{R}) \leq \log_2(|\{0,1\}^{\leq t'}|) < t' + 1$. Thus, the mutual information $I((X^1, \ldots, X^t); \mathbf{R})$ is less than $t' + 1$. By the independence of the X^js, Lemma II.1 gives

$$\sum_{j \in [t]} I(X^j; \mathbf{R}) \ < \ t' + 1 \ . \tag{1}$$

By Fact II.2,

$$I(X^j; \mathbf{R}) \ = D_{\mathrm{KL}}\left((X^j, \mathbf{R}) \,||\, (Y^j, \mathbf{R})\right), \tag{2}$$

where $Y^j \sim \mathcal{D}$ is independent of $\mathbf{R}$. By Theorem II.3,

$$D_{\mathrm{KL}}\left((X^j, \mathbf{R}) \,||\, (Y^j, \mathbf{R})\right) \ \geq \ \frac{2}{\ln 2} \cdot ||(X^j, \mathbf{R}) - (Y^j, \mathbf{R})||^2_{\mathrm{stat}}$$

$$= \frac{2}{\ln 2} \cdot \mathbb{E}_{x^j \sim \mathcal{D}} \left[\left|\left| R\left(\mathcal{D}^{\otimes(j-1)}, x^j, \mathcal{D}^{\otimes(t-j)}\right) - R\left(\mathcal{D}^{\otimes t}\right) \right|\right|_{\mathrm{stat}} \right]^2 ,$$

where the equality follows from the distinguishability interpretation of statistical distance. Using this and Jensen's inequality, we find

$$\left(\frac{1}{t} \sum_{j \in [t]} \mathbb{E}_{x^j \sim \mathcal{D}}[||R\left(\mathcal{D}^{\otimes(j-1)}, x^j, \mathcal{D}^{\otimes(t-j)}\right) - R\left(\mathcal{D}^{\otimes t}\right)||_{\mathrm{stat}}]\right)^2$$

$$\leq \frac{1}{t}\sum_{j\in[t]} \mathbb{E}_{x^j\sim\mathcal{D}}[||R\left(\mathcal{D}^{\otimes(j-1)}, x^j, \mathcal{D}^{\otimes(t-j)}\right) - R\left(\mathcal{D}^{\otimes t}\right)||_{\text{stat}}]^2,$$

which is less than $\frac{\ln 2}{2}\cdot\frac{t'+1}{t}$. Thus, R is $\sqrt{\frac{\ln 2}{2}\cdot\frac{t'+1}{t}}$-distributionally stable with respect to $\mathcal{D}$. This proves item 1 of the Lemma.

For item 2, we apply the alternative bound, Vajda's inequality (Theorem II.4), to each $j\in[t]$, to find

$$D_{\text{KL}}\left((X^j,\mathbf{R})\,||\,(Y^j,\mathbf{R})\right) \geq$$

$$\frac{1}{\ln 2}\left(\ln\left((1-||(X^j,\mathbf{R})-(Y^j,\mathbf{R})||_{\text{stat}})^{-1}\right)-1\right)$$

$= \frac{1}{\ln 2}\left(\ln\left(\varepsilon_j^{-1}\right)-1\right)$, where we define $\varepsilon_j := 1 - \mathbb{E}_{x^j\sim\mathcal{D}}\left[\left|\left|R\left(\mathcal{D}^{\otimes(j-1)}, x^j, \mathcal{D}^{\otimes(t-j)}\right) - R\left(\mathcal{D}^{\otimes t}\right)\right|\right|_{\text{stat}}\right]$, and note that $\varepsilon_j > 0$ since the support of $(Y^j,\mathbf{R})$ contains that of $(X^j,\mathbf{R})$. Averaging over $j\in[t]$ and applying Eqs. (1) and (2),

$$\frac{t'+1}{t} \;\geq\; \frac{1}{t}\sum_{j\in[t]}\frac{1}{\ln 2}\left(\ln\left(\varepsilon_j^{-1}\right)-1\right)\;,$$

i.e.,

$$\frac{1}{t}\sum_{j\in[t]}\ln\left(\varepsilon_j^{-1}\right) \;\leq\; \frac{(\ln 2)(t'+1)}{t}+1\;.$$

The function $f(x)=\ln(1/x)$ has $f''(x)=x^{-2}>0$ for $x>0$, and so Jensen's inequality gives $\ln((\frac{1}{t}\sum_{j\in[t]}\varepsilon_j)^{-1}) \leq \frac{(\ln 2)(t'+1)}{t}+1$. Thus $\frac{1}{t}\sum_{j\in[t]}\varepsilon_j \;\geq\; (e^{\frac{(\ln 2)(t'+1)}{t}+1})^{-1} \;\geq\; 2^{-\frac{t'}{t}-3}$, giving item 2. ∎

Next we need a technical lemma stating that if a mapping F is distributionally stable with respect to i.i.d. inputs, then F also obeys a slightly different stability property, in which we replace an input distribution $\mathcal{D}$ with a "sparsified" version of $\mathcal{D}$. The proof is in the full version.

Lemma III.3. *Let U be a finite set, and let $F(x^1,\ldots,x^T): \{0,1\}^{T\times n}\to U$ be given. Suppose F is δ-distributionally stable with respect to input distribution $\mathcal{D}^{\otimes T}$, for every distribution $\mathcal{D}$ over $\{0,1\}^n$. Fix some distribution $\mathcal{D}$ over $\{0,1\}^n$, and let $x^1,\ldots,x^d$ be independently sampled from $\mathcal{D}$. Let $k^*\sim\mathcal{U}_{[d]}$. Let $\widehat{\mathcal{D}}$ denote the distribution defined by sampling uniformly from the multiset $\{x^k\}_{k\neq k^*}$. (This distribution is itself a random variable, determined by $x^1,\ldots,x^d$ and by k^*.) Define β_j as*

$$\mathop{\mathbb{E}}_{k^*,x^1,\ldots,x^d}\left[\;\left|\left|F\left(\widehat{\mathcal{D}}^{\otimes(j-1)}, x^{k^*}, \widehat{\mathcal{D}}^{\otimes(T-j)}\right) \;-\; F\left(\widehat{\mathcal{D}}^{\otimes T}\right)\right|\right|_{\text{stat}}\;\right],$$

where the $\widehat{\mathcal{D}}$s are to be mutually independent (for fixed values of $x^1,\ldots,x^d$ and k^*). Then, $\frac{1}{T}\sum_{j=1}^t\beta_j \;\leq\; \delta+2T/d$.

Lemma III.4. *Let U be a finite set, and let $F(x^1,\ldots,x^T): \{0,1\}^{T\times n}\to U$ be given. Suppose F is δ-distributionally stable with respect to input distribution $\mathcal{D}^{\otimes T}$, for every distribution $\mathcal{D}$ over $\{0,1\}^n$.*

Let $S\subseteq\{0,1\}^n$, and fix $d>0$. Given any $\varepsilon>0$, let $s := \lceil(.5\ln 2)n/\varepsilon^2\rceil$. Then there exists a collection $K_1,\ldots,K_s$ of size-d multisets contained in S, such that for every $y\in S$ the following holds:

$$\mathop{\mathbb{E}}_{a\sim\mathcal{U}_{[s]},j^*\sim\mathcal{U}_{[t]}}\left[\;\left|\left|F\left(\mathcal{U}_{K_a}^{\otimes(j^*-1)}, y, \mathcal{U}_{K_a}^{\otimes(T-j^*)}\right) \;-\; F\left(\mathcal{U}_{K_a}^{\otimes T}\right)\right|\right|_{\text{stat}}\;\right]$$
$$\leq\; \delta + 2T/(d+1)+\varepsilon.$$

Proof: Consider the following two-player, simultaneous-move, zero-sum game:

- **Player 1:** chooses a size-d multiset $K\subseteq S$.
- **Player 2:** chooses a string $y\in S$.
- **Payoff:** Player 2 receives a payoff equal to

$$\mathop{\mathbb{E}}_{j^*\sim\mathcal{U}_{[T]}}\left[\;\left|\left|F\left(\mathcal{U}_K^{\otimes(j^*-1)}, y, \mathcal{U}_K^{\otimes(T-j^*)}\right) \;-\; F\left(\mathcal{U}_K^{\otimes T}\right)\right|\right|_{\text{stat}}\;\right].$$

Note that the payoff is a determinate value, given (K,y).

Consider any randomized strategy by Player 2, specified by a distribution $y\sim Y$ over S. In response, let $\mathcal{K}_Y$ be the randomized Player-1 strategy that chooses a size-d multiset K of elements sampled independently from Y.

To bound the expected payoff under the strategy-pair $(\mathcal{K}_Y, Y)$, note that we can equivalently generate $(K,y)\sim(\mathcal{K}_Y,Y)$ as follows. First, sample $x^1,\ldots,x^{d+1}$ independently from Y. Sample $k^*\sim\mathcal{U}_{[d+1]}$, set $y := x^{k^*}$, and let

$$K := \{x^1,\ldots,x^{k^*-1},x^{k^*+1},\ldots,x^{d+1}\}\;.$$

It is easily verified that $(K,y)\sim(\mathcal{K}_Y,Y)$ as desired.

Then Lemma III.3, applied to our initial distributional-stability assumption on F, informs us that

$$\mathop{\mathbb{E}}_{j^*\sim\mathcal{U}_{[t]},K,y}\left[\;\left|\left|F\left(\mathcal{U}_K^{\otimes(j^*-1)}, y, \mathcal{U}_K^{\otimes(T-j^*)}\right) - F\left(\mathcal{U}_K^{\otimes T}\right)\right|\right|_{\text{stat}}\;\right]$$
$$\leq \delta + 2T/(d+1).$$

Thus Player 2's expected payoff against $\mathcal{K}_Y$ is at most $\delta+2T/(d+1)$. As Y was arbitrary, the minimax theorem tells us that there exists a distribution $\mathcal{K}$ over Player-1 moves that forces Player 2's expected payoff under *every* strategy to be at most $\delta+2T/(d+1)$. Lemma III.4 then follows, as an immediate application of a general result due to [26, Theorem 2] and, independently, to [27] to our game above—a result showing that all two-player, zero-sum games have sparsely-supported, nearly-optimal player strategies. ∎

Lemma III.5 (Disguising-Distribution Lemma). *Let $R(x^1,\ldots,x^t): \{0,1\}^{t\times n}\to\{0,1\}^{\leq t'}$ be any possibly-randomized mapping, for $t,t'\in\mathbb{N}^+$. Let $S\subseteq\{0,1\}^n$, and fix $d>0$. Given any $\varepsilon>0$, let $s := \lceil(.5\ln 2)n/\varepsilon^2\rceil$. Let*

$$\widehat{\delta} := \min\left\{\sqrt{(\ln 2)\cdot(t'+1)/(2t)},\;\; 1-2^{-\frac{t'}{t}-3}\right\}\;.$$

Then there exists a collection $K_1,\ldots,K_s$ of size-d multisets contained in S, such that for every $y\in S$, we have

$$\mathop{\mathbb{E}}_{a\sim\mathcal{U}_{[s]},j^*\sim\mathcal{U}_{[t]}}\left[\;\left|\left|R\left(\mathcal{U}_{K_a}^{\otimes(j^*-1)},\; y,\; \mathcal{U}_{K_a}^{\otimes(t-j^*)}\right) \;-\; R\left(\mathcal{U}_{K_a}^{\otimes t}\right)\right|\right|_{\text{stat}}\;\right]$$

$$\leq \widehat{\delta} + 2t/(d+1) + \varepsilon.$$

Proof: This follows immediately from the combination of Lemmas III.2 and III.4, applied with $F := R, T := t$. ■

IV. Limits to efficient (classical) compression

Theorem IV.1. *Let L be any language. Suppose $t_1(n), t_2(n) : \mathbb{N}^+ \to \mathbb{N}^+$ are (not necessarily computable) functions. Suppose that there exists a PPT-*OR*-compression reduction $R(x^1, \ldots, x^t) : \{0,1\}^{t_1(n)\times n} \to \{0,1\}^{\leq t_2(n)}$ for L with parameters $t_1(n), t_2(n)$, error bound $\xi(n) < .5$, and some target language L'. Let*

$$\widehat{\delta} := \min\{\sqrt{(\ln 2)(t_2(n)+1)/(2t_1(n))},\ \ 1 - 2^{-\frac{t_2(n)}{t_1(n)} - 3}\}\ .$$

1) If for some constant $c > 0$ we have

$$1 - 2\xi(n) - \widehat{\delta}\ \geq\ \frac{1}{n^c}\ , \tag{3}$$

then $L \in \mathsf{NP/poly}$.

2) If for some $c > 0$ we have the (stronger) bound

$$(1 - 2\xi(n))^2 - \widehat{\delta}\ \geq\ \frac{1}{n^c}\ , \tag{4}$$

then there is a many-to-one reduction, computable in non-uniform polynomial time, from L to a promise problem in pr-SZK. *Thus $L \in \mathsf{NP/poly} \cap \mathsf{coNP/poly}$.*

Proof of Theorem IV.1: We will use the same basic reduction to prove items 1 and 2. First, with non-uniformity it is easy to handle length-n inputs whenever $L_n = \{0,1\}^n$, so let us assume from this point on that $\overline{L}_n$ is nonempty.

Using R, we define a deterministic, non-uniform polynomial-time reduction $\mathcal{R}$ that, on input $y \in \{0,1\}^n$, builds a description of two circuits C, C'. The aim is that $||\mathcal{D}_C - \mathcal{D}_{C'}||_{\text{stat}}$ should be large if $y \in L$, and small if $y \notin L$. $\mathcal{R}$ works as follows:

• **Non-uniform advice for length n:** a description of the value $t_1(n)$, and the multisets $K_1, \ldots, K_s \subseteq \overline{L}_n$ given by Lemma III.4 with $(t, t') := (t_1(n), t_2(n))$, $S := \overline{L}_n$, $d := \lceil 8t_1(n) \cdot n^c \rceil$, $\varepsilon := \frac{1}{4n^c}$. (Here $c > 0$ is as in Eq. (3) or Eq. (4), according to which item of the Theorem we are proving.) Note that d and the value s given by Lemma III.5 are both $\leq \text{poly}(n)$ under these settings, so our advice is of polynomial length.

• **On input $y \in \{0,1\}^n$:** let $\mathcal{R}$ output descriptions $\langle C, C'\rangle$ of the following two randomized circuits:

Circuit C: samples $a \sim \mathcal{U}_{[s]}$, then samples $\overline{x} = (x^1, \ldots, x^{t_1(n)}) \sim \mathcal{U}_{K_a}^{\otimes t_1(n)}$, and outputs $z := R(\overline{x})$.

Circuit C': samples values $a \sim \mathcal{U}_{[s]}$, $j^* \sim \mathcal{U}_{[t_1(n)]}$; then, samples $\overline{x} \sim \left(\mathcal{U}_{K_a}^{\otimes(j^*-1)},\ y,\ \mathcal{U}_{K_a}^{\otimes(t_1(n)-j^*)}\right)$, and outputs $z := R(\overline{x})$.

Claim IV.2. *The following holds:*

1) *If $y \in L$, then $||\mathcal{D}_C - \mathcal{D}_{C'}||_{\text{stat}} \geq D(n) := 1 - 2\xi(n)$;*
2) *If $y \notin L$, then $||\mathcal{D}_C - \mathcal{D}_{C'}||_{\text{stat}} \leq d(n) := \widehat{\delta} + \frac{1}{2n^c}$.*

The first part of Claim IV.2 follows from the "OR-respecting" property of R; the second part follows from Lemma III.5. See the full version for details.

For item 1 of Theorem IV.1, if Eq. (3) holds (for sufficiently large n), then $D(n) - d(n) \geq \frac{1}{n^c}$. Now $D(n), d(n)$ were parametrized in terms of $n = |y|$, but the gap $D(n) - d(n)$ is also at least inverse-polynomial in the length $N \leq \text{poly}(n)$ of the output description $\langle C, C'\rangle$. Thus our reduction $\mathcal{R}$ reduces any instance y of the decision problem for L, to an equivalent instance $\mathcal{R}(y) = \langle C, C'\rangle$ of the promise problem $\text{SD}^{\geq D'(N)}_{\leq d'(N)}$, with different parameters $D'(N), d'(N)$ still satisfying the gap condition $D' - d' \geq \frac{1}{\text{poly}(N)}$.

By item 1 of Theorem II.5, $\text{SD}^{\geq D'}_{\leq d'} \in$ pr-$\mathsf{NP/poly}$. Let $(A, \{a_N\}_{N>0})$ be an nondeterministic, non-uniform polynomial-time algorithm and advice family solving $\text{SD}^{\geq D'}_{\leq d'}$. By applying $(A, \{a_N\})$ to $\mathcal{R}(y)$, we get a nondeterministic, non-uniform polynomial-time algorithm solving L. Thus $L \in \mathsf{NP/poly}$, proving item 1 of the Theorem.

For item 2 of Theorem IV.1, if Eq. (4) holds for sufficiently large n, then $D(n)^2 - d(n) \geq \frac{1}{n^c}$. Arguing as in the previous case, but this time applying item 2 of Theorem II.5, we get a nonuniform polynomial-time reduction from L to $\text{SD}^{\geq D'}_{\leq d'}$, where this time $D'(N)^2 - d'(N) \geq \frac{1}{\text{poly}(N)}$. This problem can in turn be reduced to $\text{SD}^{\geq 2/3}_{\leq 1/3} \in$ pr-SZK in non-uniform polynomial time. This also yields $L \in \mathsf{NP/poly} \cap \mathsf{coNP/poly}$, and completes the proof of Theorem IV.1. ■

The assertions of Theorems I.1 and I.2 for AND(SAT) follow easily from item 2 of Theorem IV.1, after noting that an AND-compression reduction for SAT is also an OR-compression reduction for $L := \overline{\text{SAT}}$. The assertions of those Theorems for OR(SAT) follow from item 2 of Theorem IV.1 applied to $L := \text{SAT}$. See the full version for the calculations involved.

V. Questions for further study

(1.) Can the limitations we show on efficient compression for AND(SAT) and OR(SAT) be extended to the *oracle communication model* studied in [8]?

(2.) Using our results on the infeasibility of compression for AND(SAT), can we extend the work of [8] to prove new kernel-size lower bounds for interesting problems *with* polynomial kernels, under the assumption $\mathsf{NP} \not\subseteq \mathsf{coNP/poly}$?

(3.) Can we obtain a better quantitative understanding of the limits to efficient f-compression of NP-complete languages, where f is a combining function other than OR or AND? The case $f = \bigvee_{i=1}^{m}(\bigwedge_{j=1}^{m} x^{i,j})$ is an interesting candidate for study.

(4.) Find more applications for "disguising distributions."

Acknowledgment

I thank Hans Bodlaender, Holger Dell, Lance Fortnow, Russell Impagliazzo, James Lee, Dieter van Melkebeek,

Ashwin Nayak, Karolina Soltys, Salil Vadhan, Avi Wigderson, Ryan Williams, and several anonymous reviewers for helpful comments. Thanks especially to Russell, Ashwin, and Salil for allowing me to include their alternative proof suggestions.

REFERENCES

[1] R. G. Downey and M. Fellows, *Parametrized Complexity*, 1st ed. Springer (Monographs in Computer Science), 1999.

[2] J. Guo and R. Niedermeier, "Invitation to data reduction and problem kernelization," *SIGACT News*, vol. 38, no. 1, pp. 31–45, 2007.

[3] D. Harnik and M. Naor, "On the compressibility of NP instances and cryptographic applications," *SIAM J. Comput.*, vol. 39, no. 5, pp. 1667–1713, 2010.

[4] L. Fortnow and R. Santhanam, "Infeasibility of instance compression and succinct PCPs for NP," *J. Comput. Syst. Sci.*, vol. 77, no. 1, pp. 91–106, 2011.

[5] H. L. Bodlaender, R. G. Downey, M. R. Fellows, and D. Hermelin, "On problems without polynomial kernels," *J. Comput. Syst. Sci.*, vol. 75, no. 8, pp. 423–434, 2009.

[6] H. L. Bodlaender, B. M. P. Jansen, and S. Kratsch, "Cross-composition: A new technique for kernelization lower bounds," in *STACS*, 2011, pp. 165–176.

[7] M. Dom, D. Lokshtanov, and S. Saurabh, "Incompressibility through colors and IDs," in *36th ICALP*, 2009, pp. 378–389.

[8] H. Dell and D. van Melkebeek, "Satisfiability allows no nontrivial sparsification unless the polynomial-time hierarchy collapses," in *42nd ACM STOC*, 2010, pp. 251–260.

[9] H. L. Bodlaender, S. Thomassé, and A. Yeo, "Kernel bounds for disjoint cycles and disjoint paths," *Theor. Comput. Sci.*, vol. 412, no. 35, pp. 4570–4578, 2011.

[10] H. L. Bodlaender, B. M. P. Jansen, and S. Kratsch, "Kernel bounds for path and cycle problems," in *IPEC*, 2011, pp. 145–158.

[11] ——, "Preprocessing for treewidth: A combinatorial analysis through kernelization," in *38th ICALP*, 2011, pp. 437–448.

[12] Y. Chen, J. Flum, and M. Müller, "Lower bounds for kernelizations and other preprocessing procedures," *Theory Comput. Syst.*, vol. 48, no. 4, pp. 803–839, 2011.

[13] D. Hermelin and X. Wu, "Weak compositions and their applications to polynomial lower bounds for kernelization," in *23rd ACM-SIAM SODA*, 2012, pp. 104–113.

[14] H. Dell and D. Marx, "Kernelization of packing problems," in *23rd ACM-SIAM SODA*, 2012, pp. 68–81.

[15] S. Kratsch, "Co-nondeterminism in compositions: a kernelization lower bound for a Ramsey-type problem," in *23rd ACM-SIAM SODA*, 2012, pp. 114–122.

[16] H. Buhrman and J. M. Hitchcock, "NP-hard sets are exponentially dense unless coNP $\subseteq$ NP/poly," in *23rd IEEE Conference on Computational Complexity*, 2008, pp. 1–7.

[17] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[18] J. Watrous, "PSPACE has constant-round quantum interactive proof systems," *Theor. Comput. Sci.*, vol. 292, no. 3, pp. 575–588, 2003.

[19] R. Jain, Z. Ji, S. Upadhyay, and J. Watrous, "QIP = PSPACE," *J. ACM*, vol. 58, no. 6, p. 30, 2011.

[20] A. Sahai and S. P. Vadhan, "A complete problem for statistical zero knowledge," *J. ACM*, vol. 50, no. 2, pp. 196–249, 2003.

[21] O. Regev, "Entropy-based bounds on dimension reduction in L_1," 2011, arXiv:1108.1283v3.

[22] A. A. Fedotov, P. Harremoës, and F. Topsøe, "Refinements of pinsker's inequality," *IEEE Transactions on Information Theory*, vol. 49, no. 6, pp. 1491–1498, 2003.

[23] M. D. Reid and R. C. Williamson, "Generalised pinsker inequalities," in *COLT*, 2009.

[24] R. Raz, "A parallel repetition theorem," *SIAM J. Comput.*, vol. 27, no. 3, pp. 763–803, 1998.

[25] H. Klauck, A. Nayak, A. Ta-Shma, and D. Zuckerman, "Interaction in quantum communication," *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp. 1970–1982, 2007.

[26] R. J. Lipton and N. E. Young, "Simple strategies for large zero-sum games with applications to complexity theory," in *26th ACM STOC*, 1994, pp. 734–740.

[27] I. Althöfer, "On sparse approximations to randomized strategies and convex combinations," *Linear Algebra and its Applications*, vol. 199, Supplement 1, no. 0, pp. 339 – 355, 1994.

[28] A. Nayak, "Optimal lower bounds for quantum automata and random access codes," in *40th IEEE FOCS*, 1999, pp. 369–377.

[29] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.

[30] L. M. Adleman, "Two theorems on random polynomial time," in *19th IEEE FOCS*, 1978, pp. 75–83.

[31] L. Fortnow, "The complexity of perfect zero-knowledge (extended abstract)," in *19th ACM STOC*, A. V. Aho, Ed., 1987, pp. 204–209.

[32] W. Aiello and J. Håstad, "Statistical zero-knowledge languages can be recognized in two rounds," *J. Comput. Syst. Sci.*, vol. 42, no. 3, pp. 327–345, 1991.

[33] T. Okamoto, "On relationships between statistical zero-knowledge proofs," *J. Comput. Syst. Sci.*, vol. 60, no. 1, pp. 47–108, 2000.

[34] O. Goldreich and S. P. Vadhan, "On the complexity of computational problems regarding distributions (a survey)," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. TR11-004, p. 4, 2011.