A Uniform Min-Max Theorem with Applications in Cryptography*

Salil Vadhan and Colin Jia Zheng

School of Engineering and Applied Sciences
Harvard University
Cambridge, Massachusetts
{salil,colinz}@seas.harvard.edu

Abstract. We present a new, more constructive proof of von Neumann's Min-Max Theorem for two-player zero-sum game — specifically, an algorithm that builds a near-optimal mixed strategy for the second player from several best-responses of the second player to mixed strategies of the first player. The algorithm extends previous work of Freund and Schapire (Games and Economic Behavior '99) with the advantage that the algorithm runs in poly(n) time even when a pure strategy for the first player is a distribution chosen from a set of distributions over $\{0,1\}^n$. This extension enables a number of additional applications in cryptography and complexity theory, often yielding uniform security versions of results that were previously only proved for nonuniform security (due to use of the non-constructive Min-Max Theorem).

We describe several applications, including a more modular and improved uniform version of Impagliazzo's Hardcore Theorem (FOCS '95), showing impossibility of constructing succinct non-interactive arguments (SNARGs) via black-box reductions under uniform hardness assumptions (using techniques from Gentry and Wichs (STOC '11) for the nonuniform setting), and efficiently simulating high entropy distributions within any sufficiently nice convex set (extending a result of Trevisan, Tulsiani and Vadhan (CCC '09)).

1 Introduction

Von Neumann's Min-Max Theorem (or Linear Programming Duality, finite-dimensional Hahn-Banach Theorem) has proved to be an extremely useful tool in theoretical computer science. Consider a zero-sum game between two players where for every mixed strategy V for Player 1 (as a distribution over his strategy space $\mathcal{V}$), Player 2 has a response $W \in \mathcal{W}$ that guarantees $\mathbb{E}[F(V,W)] \geq 0$, where F (payoff) can be an arbitrary function. The Min-Max Theorem says that there must exist a Player 2's mixed strategy W^* (as a distribution over his strategy space $\mathcal{W}$) that guarantees $\mathbb{E}[F(V,W^*)] \geq 0$ for *all* strategies $V \in \mathcal{V}$ of Player 1.

* Supported by NSF grant CCF-1116616 and US-Israel BSF grant 2010196. A full version of this paper [VZ2] to appear on the Cryptology ePrint Archive.

R. Canetti and J.A. Garay (Eds.): CRYPTO 2013, Part I, LNCS 8042, pp. 93–110, 2013.

The Min-Max Theorem gives rise to a number of results in cryptography and complexity theory such as Impagliazzo's Hardcore Theorem [Imp], equivalence of different notions of computational entropy [BSW], the Dense Model Theorem [RTTV], leakage-resilient cryptography [DP,FR], efficient simulation of high entropy distributions [TTV], impossibility of constructing succinct non-interactive arguments (SNARGs) via black-box reductions [GW], and simple construction of pseudorandom generators from one-way functions [VZ1]. In a typical application like these, Player 1 chooses V from a convex set $\mathcal{V}$ of distributions over $\{0,1\}^n$, and Player 2 chooses W from a set $\mathcal{W}$ of (possibly randomized) boolean functions $\{0,1\}^n \to \{0,1\}$ and receives payoff $F(V,W) = \mathbb{E}[W(V)]$ i.e. function W's expected output when input is drawn from the distribution V. For example, $\mathcal{V}$ contains all high entropy distributions over $\{0,1\}^n$ and $\mathcal{W}$ contains all boolean functions of small circuit size.

A limitation of the Min-Max Theorem is that it is highly non-constructive; it only asserts the existence of the optimal strategy W^* but does not say how it can be found (algorithmically). Consequently, applications of the Min-Max Theorem only give rise to results about nonuniform boolean circuits, rather than uniform algorithms (e.g. we set cryptographic protocols based on nonuniform hardness rather than uniform hardness assumptions).

To overcome this, we consider the natural algorithmic task of constructing such an optimal strategy W^* for Player 2, given an efficient algorithm for F. When the sizes of strategy spaces $\mathcal{V}$ and $\mathcal{W}$ are small (e.g. polynomial) this can be done by linear programming, for which efficient algorithms are well-known. However, applications in cryptography and complexity theory such as ones just mentioned involve exponentially large strategy spaces, and an optimal strategy W^* cannot be found in polynomial time in general. Thus we also require that, given any mixed strategy V for Player 1, not only does there exist a strategy $W \in \mathcal{W}$ for Player 2 with $\mathbb{E}[F(V,W)] \geq 0$, but such response W can be obtained efficiently by an oracle (or an efficient uniform algorithm).

Assuming such an oracle, Freund and Schapire [FS] show how to find an approximately optimal W^* for Player 2 in polynomial time and by making $O((\log|\mathcal{V}|)/\epsilon^2)$ adaptive oracle queries, using the idea of multiplicative weight updates. However, their algorithm still falls short in some of aforementioned applications where $\mathcal{V}$ is a set of distributions over $\{0,1\}^n$, and thus $\mathcal{V}$ can have doubly-exponentially many vertices. For example, consider the set of distributions on $\{0,1\}^n$ of min-entropy at least k; the vertices of $\mathcal{V}$ are uniform distributions on a subset of size 2^k, and there are $\binom{2^n}{2^k}$ such subsets.

We present a Uniform Min-Max Theorem that efficiently finds an approximately optimal strategy W^* for Player 2, given an oracle that for any of Player 1's mixed strategy $V \in \mathcal{V}$ returns some Player 2's strategy that guarantees reasonable payoff, even when $\mathcal{V}$ is a (sufficiently nice) set of distributions over $\{0,1\}^n$. Our algorithm is inspired by the proof of Uniform Hardcore Theorem of Barak, Hardt, and Kale [BHK]. Like [BHK], our algorithm uses "relative entropy (KL) projections" together with multiplicative weight updates (a technique originally due to Herbster and Warmuth [HW]). Our contribution is the formulation

of this algorithm as providing a Uniform Min-Max Theorem. An advantage of this formulation is that it is more modular, and not specific to the Hardcore Theorem. Consequently it immediately enables a number of applications, including deriving uniform versions of many of the aforementioned results, where we now deal with algorithms rather than nonuniform boolean circuits. Even for the Hardcore Theorem, where the uniform version was already known [Hol1,BHK], there are several advantages to deducing it using the Uniform Min-Max Theorem.

Uniform Hardcore Theorem. Impagliazzo's Hardcore Theorem ([Imp] and later strengthened in [KS,Hol1,BHK]) is a fundamental result in complexity theory that says if a boolean function f is somewhat hard on average, then there must be a subset of inputs (the hardcore) on which f is extremely hard, and outside of which f is easy. There are two approaches to proving the theorem. One is constructive [Imp,KS,Hol1,BHK] and leads to a *Uniform Hardcore Theorem* where hardness of f is measured against uniform algorithms, rather than nonuniform boolean circuits, and has found several applications in cryptography [KS,Hol1,Hol2,HHR,HRV]. However, the existing proofs turn out to be adhoc and do not achieve all of the optimal parameters simultaneously for a Uniform Hardcore Theorem. Another approach due to Nisan [Imp] (and strengthened in [Hol1]) uses the (non-constructive) Min-Max Theorem and has the advantage of simplicity, but is restricted to the nonuniform measure of hardness.

In Section 4, we show that by replacing the use of Min-Max Theorem in the proof of Nisan [Imp] or Holenstein [Hol1] with our Uniform Min-Max Theorem, we obtain a new proof of the Uniform Hardcore Theorem with the advantages of (i) optimal hardcore density; (ii) optimal complexity blow-up; and (iii) modularity and simplicity.

Construction of Pseudorandom Generators from One-Way Functions. Recently, we [VZ1] obtained a simplified and more efficient construction of pseudorandom generators from arbitrary one-way functions, building on the work of [HRV]. Key to the simplification is a new characterization of a computational analogue of Shannon entropy, whose proof in the nonuniform setting involves the Min-Max Theorem. Using the Uniform Min-Max Theorem instead, we proved our characterization of pseudoentropy in the uniform setting, and hence obtain (simpler) pseudorandom generator from arbitrary one-way functions that are secure against efficient algorithms. We refer to the full version [VZ2] for a more detailed discussion.

Impossibility of Black-Box Construction of Succinct Non-interactive Argument. A result of Gentry and Wichs [GW] shows that there is no black-box construction of succinct non-interactive arguments (SNARGs) from any natural cryptographic assumption. Their result relies on the (mild) assumption that there exist *hard subset membership problems*, which is equivalent to the existence of subexponentially hard one-way functions. One limitation is that they need to assume nonuniformly secure one-way functions, in part due to their use of the non-constructive Min-Max theorem (in [GW] Lemma 3.1).

In Section 5, we show how to obtain the analogous result in the *uniform setting* by using the Uniform Min-Max Theorem. More specifically, assuming that there exist subexponentially hard one-way functions that are secure against uniform algorithms, we show that there is no construction of SNARGs whose security can be reduced in a black-box way to a cryptographic assumption against uniform algorithms (unless the assumption is already false).

Simulating Arbitrary Distributions within a Convex Set. In the full version [VZ2], we apply the Uniform Min-Max Theorem to show a result analogous to the main result of Trevisan, Tulsiani, and Vadhan [TTV], which (informally) says that any high min-entropy distribution X is indistinguishable from some high min-entropy distribution Y of *low complexity*. It is shown in [TTV] that such results can be used to deduce (versions of) the Dense Model Theorem [GT,TZ,RTTV], the Hardcore Theorem [Imp], and the Weak Regularity Lemma [FK], by translating the problem to a simpler one where the unknown distribution X is replaced with the low complexity distribution Y that can be efficiently analyzed and manipulated.

Our result is more general than [TTV] in the sense that we are no longer restricted to distributions of high min-entropy. We show that for any sufficiently "nice" convex set of distributions $\mathcal{V}$, any distribution $X \in \mathcal{V}$ is indistinguishable from some distribution $Y \in \mathcal{V}$ where Y has "low complexity" (for several slightly different definitions of complexity than [TTV]). One application of this result is a slight strengthening of the Weak Regularity Lemma of Frieza and Kannan [FK] that achieves better parameters for graphs that are not dense. Another application is deducing an "efficient" version of a technical lemma of [GW]. (The efficient version has been independently proved by Chung, Lui, and Pass [CLP] and applied in the context of distributional zero-knowledge). We note that our result has an average-case variant, which contains as special case a recent result of Pietrzak and Jetchev [PJ] on leakage-resilient cryptography.

1.1 Paper Organization

Basic notions from information theory including KL projection are defined in Section 2. In Section 3 we state and prove the Uniform Min-Max Theorem, and show that it also implies the standard Min-Max Theorem. In Section 4, 5, we describe two applications of the Uniform Min-Max Theorem (other applications can be found in the full version [VZ2]).

2 Preliminaries

Notations. For a natural number n, $[n]$ denotes the set $\{1, \ldots, n\}$, U_n denotes the uniform distribution on binary strings of length n. For a finite set Σ, U_Σ denotes the uniform distribution on Σ. For a distribution X, $\mathrm{supp}(X)$ denotes the support of X, and $x \leftarrow X$ means x is a random sample drawn from distribution X. We write $\mathrm{Avg}_{a \le i \le b}$ as a shorthand for the average over all $i \in \{a, \ldots, b\}$. $\mathrm{Conv}(\cdot)$ denotes the convex hull.

For more background on entropy and proofs of the lemmas below, see [CT].

Definition 2.1 (Entropy). *For a random variable X, the* (Shannon) entropy *of X is defined to be*

$$\mathrm{H}(X) = \mathop{\mathbb{E}}_{x \leftarrow X}\left[\log \frac{1}{\Pr[X = x]}\right].$$

The min-entropy *of X is defined to be*

$$\mathrm{H}_\infty(X) = \min_{x \in \mathrm{supp}(X)}\left(\log \frac{1}{\Pr[X = x]}\right).$$

The notion of *KL divergence* from random variable A to random variable B is closely related to Shannon entropy; intuitively it measures how dense A is within B, on average (with 0 divergence representing maximum density, i.e. $A = B$, and large divergence meaning that A is concentrated in a small portion of B).

Definition 2.2 (KL divergence). *For random variables A and B, the* KL divergence from A to B *is defined to be*

$$\mathrm{KL}(A \parallel B) = \mathop{\mathbb{E}}_{a \leftarrow A}\left[\log \frac{\Pr[A = a]}{\Pr[B = a]}\right],$$

or conventionally $+\infty$ if $\mathrm{supp}(A) \not\subseteq \mathrm{supp}(B)$.

For random variables (X, A) and (Y, B), the conditional KL divergence from $A|X$ to $B|Y$ *is defined to be*

$$\mathrm{KL}((A|X) \parallel (B|Y)) = \mathop{\mathbb{E}}_{(x,a) \leftarrow (X,A)}\left[\log \frac{\Pr[A = a|X = x]}{\Pr[B = a|Y = x]}\right].$$

Thus, conditional KL divergence captures the expected KL divergence from $A|_{X=x}$ to $B|_{Y=x}$, over $x \leftarrow X$. Like Shannon entropy, it has a chain rule:

Proposition 2.1 (Chain rule for KL divergence). $\mathrm{KL}(X, A \parallel Y, B) = \mathrm{KL}(X \parallel Y) + \mathrm{KL}((A|X) \parallel (B|Y))$.

Note however, that the KL divergence is *not* a metric; it is not symmetric and does not satisfy the triangle inequality.

Definition 2.3 (KL projection). *Let X be a distribution on Σ, and $\mathcal{V}$ be a non-empty closed convex set of distributions on Σ. $Y^* \in \mathcal{V}$ is called a* KL projection of X on $\mathcal{V}$ *if*

$$Y^* = \arg\min_{Y \in \mathcal{V}} \mathrm{KL}(Y \parallel X).$$

A nice property of KL projection is the following geometric structure (see [CT], Chap 11, Section 6):

Theorem 2.1 (Pythagorean theorem). *Let $\mathcal{V}$ be a non-empty closed convex set of distributions on Σ. Let Y^* be a KL projection of X on $\mathcal{V}$. Then for all $Y \in \mathcal{V}$,*

$$\mathrm{KL}(Y \parallel Y^*) + \mathrm{KL}(Y^* \parallel X) \leq \mathrm{KL}(Y \parallel X).$$

In particular,

$$\mathrm{KL}(Y \parallel Y^*) \leq \mathrm{KL}(Y \parallel X).$$

Assuming $\mathrm{KL}(Y^* \parallel X)$ is finite, then Pythagorean theorem implies that the KL projection Y^* is unique: for any $Y \in \mathcal{V}$ which is also a KL projection, the theorem implies $\mathrm{KL}(Y \parallel Y^*) = 0$, which holds only when $Y = Y^*$.

Finding the exact KL projection is often computationally infeasible, so we consider *approximate KL projection*:

Definition 2.4 (Approximate KL projection). *We say Y^* is a* σ-approximate KL projection *of X on $\mathcal{V}$, if $Y^* \in \mathcal{V}$ and for all $Y \in \mathcal{V}$,*

$$\mathrm{KL}(Y \parallel Y^*) \leq \mathrm{KL}(Y \parallel X) + \sigma.$$

3 A Uniform Min-Max Theorem

Consider a zero-sum game between two players, where the space of pure strategies for Player 1 is $\mathcal{V}$, the space of pure strategies for Player 2 is $\mathcal{W}$, and $\mathcal{V}$ is an arbitrary subset of distributions over $[N]$. In this section we present a Uniform Min-Max Theorem that efficiently finds an approximately optimal strategy $W^* \in \mathrm{Conv}(\mathcal{W})$ for Player 2, given an oracle which, when fed any of Player 1's mixed strategies $V \in \mathrm{Conv}(\mathcal{V})$, returns a strategy for Player 2 that guarantees good payoff. Our algorithm is inspired by the proof of Uniform Hardcore Theorem of Barak, Hardt, and Kale [BHK]. Like [BHK], our algorithm uses "relative entropy (KL) projections" together with multiplicative weight updates (a technique originally due to Herbster and Warmuth [HW]).

We first state the theorem and mention how it implies standard Min-Max Theorem.

Theorem 3.1 (A Uniform Min-Max Theorem). *Consider a two-player zero-sum game where the sets of pure strategies for Player 1 and Player 2 are $\mathcal{V} \subseteq \{$distributions over $[N]\}$ and $\mathcal{W}$, and the payoff to Player 2 is defined to be $F(V, W) = \mathbb{E}_V[f(V, W)]$ for some function $f : [N] \times \mathcal{W} \to [-k, k]$. Then for every $0 < \epsilon \leq 1$ and $S \geq \max_{V \in \mathrm{Conv}(\mathcal{V})} \mathrm{KL}(V \parallel V^{(1)})/\epsilon^2$, after S iterations Algorithm 3.1 (Finding Universal Strategy) always outputs a mixed strategy W^* for Player 2 such that*

$$\min_{V \in \mathcal{V}} F(V, W^*) \geq \operatorname*{Avg}_{1 \leq i \leq S} F(V^{(i)}, W^{(i)}) - O(k\epsilon).$$

(This holds regardless of the arbitrary choice of $W^{(i)}$ and $V^{(i+1)}$ in the algorithm.)

In particular, it suffices to take $S \geq (\log N - \min_{V \in \mathcal{V}} \mathrm{H}(V))/\epsilon^2$ if we set $V^{(1)} = U_{[N]} \in \mathrm{Conv}(\mathcal{V})$.

Choose an initial strategy $V^{(1)} \in \mathrm{Conv}(\mathcal{V})$ for Player 1
for $i \leftarrow 1$ **to** S **do**
  Obtain an arbitrary strategy $W^{(i)} \in \mathcal{W}$ for Player 2
  Weight Update:
  Let $V^{(i)'}$ be such that $\Pr[V^{(i)'} = x] \propto e^{-\epsilon \cdot f(x,W^{(i)})/2k} \cdot \Pr[V^{(i)} = x]$
  Projection:
  $V^{(i+1)} \leftarrow$ an arbitrary ϵ^2-approx KL projection of $V^{(i)'}$ on $\mathrm{Conv}(\mathcal{V})$
end
Let W^* be the mixed strategy for Player 2 uniform over $W^{(1)}, \ldots, W^{(S)}$
return W^*

Algorithm 3.1. Finding Universal Strategy

We now describe how Theorem 3.1 implies the original Min-Max Theorem, which says

$$\max_{W \in \mathrm{Conv}(\mathcal{W})} \min_{V \in \mathcal{V}} F(V,W) = \min_{V \in \mathrm{Conv}(\mathcal{V})} \max_{W \in \mathcal{W}} F(V,W).$$

For each i, take $W^{(i)}$ to be Player 2's best response to Player 1's mixed strategy $V^{(i)}$, i.e. $F(V^{(i)}, W^{(i)}) = \max_{W \in \mathcal{W}} F(V^{(i)}, W)$. Theorem 3.1 says for every $\lambda = O(k\epsilon) > 0$, by setting an appropriate $V^{(1)}$ and sufficiently large S, there exists $W^* \in \mathrm{Conv}(\mathcal{W})$ with

$$\begin{aligned}
\min_{V \in \mathcal{V}} F(V,W^*) &\geq \operatorname*{Avg}_{1 \leq i \leq S} F(V^{(i)}, W^{(i)}) - \lambda \\
&= \operatorname*{Avg}_{1 \leq i \leq S} \max_{W \in \mathcal{W}} F\left(V^{(i)}, W\right) - \lambda \\
&\geq \min_{V \in \mathrm{Conv}(\mathcal{V})} \max_{W \in \mathcal{W}} F\left(V, W\right) - \lambda,
\end{aligned}$$

where the last inequality holds because for every i, $\max_{W \in \mathcal{W}} F\left(V^{(i)}, W\right) \geq \min_{V \in \mathrm{Conv}(\mathcal{V})} \max_{W \in \mathcal{W}} F\left(V, W\right)$. Thus, for every $\lambda > 0$,

$$\max_{W \in \mathrm{Conv}(\mathcal{W})} \min_{V \in \mathcal{V}} F(V,W) \geq \min_{V \in \mathrm{Conv}(\mathcal{V})} \max_{W \in \mathcal{W}} F(V,W) - \lambda$$

Taking $\lambda \to 0$ gives the Min-Max Theorem.

Proof (of Theorem 3.1). Consider any $V \in \mathcal{V}$. It follows from Lemma A.1 that

$$\mathrm{KL}(V \parallel V^{(i)'}) \leq \mathrm{KL}(V \parallel V^{(i)}) - (\log e)\epsilon \left(\frac{F(V^{(i)}, W^{(i)}) - F(V, W^{(i)})}{2k} - \epsilon \right).$$

Since $V^{(i+1)}$ is a σ-approximate KL projection of $V^{(i)'}$ on $\mathrm{Conv}(\mathcal{V})$,

$$\mathrm{KL}(V \parallel V^{(i+1)}) \leq \mathrm{KL}(V \parallel V^{(i)'}) + \sigma.$$

Therefore

$$\mathrm{KL}(V \parallel V^{(i)}) - \mathrm{KL}(V \parallel V^{(i+1)}) \geq (\log e)\epsilon \left(\frac{F(V^{(i)}, W^{(i)}) - F(V, W^{(i)})}{2k} - \epsilon \right) - \sigma.$$

Summing over $i = 1, \ldots, S$ and telescoping, we obtain

$$\begin{aligned} &\mathrm{KL}(V \parallel V^{(1)}) - \mathrm{KL}(V \parallel V^{(S+1)}) \\ &\geq (\log e)\epsilon \sum_{i=1}^{S} \left(\frac{F(V^{(i)}, W^{(i)}) - F(V, W^{(i)})}{2k} - \epsilon \right) - S\sigma \\ &= (\log e) S\epsilon \left(\frac{\mathrm{Avg}_{1 \leq i \leq S} F(V^{(i)}, W^{(i)}) - F(V, W^*)}{2k} - \epsilon \right) - S\sigma. \end{aligned}$$

Since $\mathrm{KL}(V \parallel V^{(S+1)}) \geq 0$, rearranging gives

$$\frac{\mathrm{Avg}_{1 \leq i \leq S} F(V^{(i)}, W^{(i)}) - F(V, W^*)}{2k} \leq \frac{\mathrm{KL}(V \parallel V^{(1)}) + S\sigma}{(\log e) S\epsilon} + \epsilon = O(\epsilon)$$

for $\sigma = \epsilon^2$, $S = \mathrm{KL}(V \parallel V^{(1)})/\epsilon^2$.

Next we describe an average case variant where the set $\mathcal{V}$ of strategies for Player 1 is a set of distributions of the form (X, C) where C may vary, but the marginal distribution of X is fixed. This is convenient for a number of applications (e.g. Section 5, and simple construction of pseudorandom generators from one-way functions [VZ1]) that involve distinguishers on such joint distributions (X, C).

Theorem 3.2 (Uniform Min-Max Theorem – Average Case). *Let $\mathcal{V}$ be a subset of distributions over $[N] \times [q]$ of the form (X, C) where C may vary, but the marginal distribution of X is fixed. That is, for every (X, C), $(X', C') \in \mathcal{V}$ and every $x \in [N]$ we have $\sum_c \Pr[(X, C) = (x, c)] = \sum_c \Pr[(X', C') = (x, c)]$.*

Consider a two-player zero-sum game where the sets of pure strategies for Player 1 and Player 2 are $\mathcal{V}$ and $\mathcal{W}$, and the payoff to Player 2 is defined to be $F((X, C), W) = \mathbb{E}_{X,C}[f((X, C), W)]$ for some function $f : [N] \times [q] \times \mathcal{W} \to [-k, k]$. Then for every $0 < \epsilon \leq 1$ and $S \geq \max_{(X,C) \in \mathrm{Conv}(\mathcal{V})} \mathrm{KL}(X, C \parallel X, C^{(1)})/\epsilon^2$, after S iterations Algorithm 3.2 (Finding Universal Strategy – Average Case) always outputs a mixed strategy W^* for Player 2 such that

$$\min_{(X,C) \in \mathcal{V}} F((X, C), W^*) \geq \operatorname*{Avg}_{1 \leq i \leq S} F((X, C^{(i)}), W^{(i)}) - O(k\epsilon).$$

(This holds regardless of the arbitrary choice of $W^{(i)}$ and $C^{(i+1)}$ in the algorithm.)

In particular, it suffices to take $S \geq \left(\log q - \min_{(X,C) \in \mathcal{V}} \mathrm{H}(C|X)\right)/\epsilon^2$ if we set $(X, C^{(1)}) = (X, U_{[q]}) \in \mathrm{Conv}(\mathcal{V})$ (where $U_{[q]}$ is independent of X).

Choose an initial strategy $(X, C^{(1)}) \in \mathrm{Conv}(\mathcal{V})$ for Player 1
for $i \leftarrow 1$ **to** S **do**
  Obtain an arbitrary strategy $W^{(i)} \in \mathcal{W}$ for Player 2
  Weight Update:
  Let $C^{(i)'}$ be such that $\forall x, a$,
    $\Pr[C^{(i)'} = a | X = x] \propto e^{-\epsilon \cdot f(x,a,W^{(i)})/2k} \cdot \Pr[C^{(i)} = a | X = x]$
  Projection:
  $(X, C^{(i+1)})$
    $\leftarrow$ an arbitrary ϵ^2-approx KL projection of $(X, C^{(i)'})$ on $\mathrm{Conv}(\mathcal{V})$
end
Let W^* be the mixed strategy for Player 2 uniform over $W^{(1)}, \ldots, W^{(S)}$
return W^*

Algorithm 3.2. Finding Universal Strategy – Average Case

Proof. Note that Algorithm 3.2 is the same as Algorithm 3.1, except for the difference that here we update $C^{(i)}$ instead of $V^{(i)}$. We show that the combined effect of the update and KL projection steps is identical in the two algorithms. Note that we can write $V^{(i)'}$ as $(X^{(i)'}, g_i(X^{(i)'}))$ for the randomized function g_i where $\Pr[g_i(x) = a] \propto e^{\epsilon \cdot f(x,a,W^{(i)})/2k} \cdot \Pr[C^{(i)} = a | X = x]$ for every x and a. For the same function g_i, we have $(X, g_i(X)) = (X, C^{(i)'})$. Thus, we can apply the following lemma.

Lemma 3.1. *Let X' be a distribution on $[N]$ with $\mathrm{supp}(X') \supseteq \mathrm{supp}(X')$, and let $g : [N] \to [q]$ be a randomized function. Then the KL projection of $(X', g(X'))$ on $\mathrm{Conv}(\mathcal{V})$ equals the KL projection of $(X, g(X))$ on $\mathrm{Conv}(\mathcal{V})$.*

Proof. Consider any $(X, C) \in \mathrm{Conv}(\mathcal{V})$. We have

$$
\begin{aligned}
&\mathrm{KL}(X, C \parallel X', g(X')) \\
&= \mathrm{KL}(X \parallel X') + \mathrm{KL}((C|X) \parallel (g(X')|X')) && \text{(by chain rule for KL divergence)} \\
&= \mathrm{KL}(X \parallel X') + \mathrm{KL}((C|X) \parallel (g(X)|X)) && \text{(by def of conditional KL divergence)} \\
&= \mathrm{KL}(X \parallel X') + \mathrm{KL}(X, C \parallel X, g(X)). && \text{(by chain rule for KL divergence)}
\end{aligned}
$$

Thus the KL projections are the same.

4 Application: Uniform Hardcore Theorem

A fundamental result in complexity theory is Impagliazzo's Hardcore Theorem [Imp], which, in the strengthened version due to Klivans and Servedio [KS] and Holenstein [Hol1], says that every function $f : \{0,1\}^n \to \{0,1\}$ that is δ-hard for poly-sized boolean circuits (that is, every poly-sized circuit fails to compute f on at least δ fraction of inputs) must be *extremely* hard on a subset of inputs of density at least 2δ (the *hardcore set*) (and may be easy elsewhere). In this

section, we provide a simplified proof of a hardcore theorem with optimal parameters, where hardness is defined with respect to *uniform* algorithms rather than boolean circuits. Following [Imp], we will deal with hardcore distributions instead of hardcore sets, which are equivalent up to a negligible additive difference in density, where density of a distribution is defined as follows:

Definition 4.1 (Density of distribution). *Let X and Y be distributions over* some finite set Σ. *We say X is* δ-dense in Y *if* $\Pr[Y = x] \geq \delta \cdot \Pr[X = x]$ *for all* $x \in \Sigma$. We say X is δ-dense if it is δ-dense in U_Σ (equivalently, having min-entropy at least $\log|\Sigma| - \log(1/\delta)$). *We denote by $\mathcal{C}_{m,\delta}$ the set of* all δ-dense *distributions* on $\{0,1\}^m$.

The (nonuniform) hardcore theorem with optimal hardcore density 2δ and optimal complexity blow-up $O(\log(1/\delta)/\epsilon^2)$, is due to [KS] using techniques from boosting, and an idea of iteratively increasing hardcore size due to Wigderson. The theorem can be stated as follows:

Theorem 4.1 (Hardcore Theorem [KS]). *Let (X, B)[1] be a joint distribution over $\{0,1\}^n \times \{0,1\}$ and $\epsilon > 0$. Let B be (t, δ)-hard given X, i.e. for every size t circuit P it holds that* $\Pr[P(X) = B] \leq 1 - \delta$. *Then there is a joint distribution $(\hat{X}, \hat{B})$ that is 2δ-dense in (X, B), such that for every size $t' = t/O(\log(1/\delta)/\epsilon^2)$ circuit A it holds that* $\Pr[A(\hat{X}) = \hat{B}] \leq (1 + \epsilon)/2$.

The original paper of Impagliazzo [Imp] contains both a non-trivial constructive proof, as well as a much simpler, yet non-constructive proof due to Nisan that uses the Min-Max Theorem. Nisan's proof has an appealing simplicity: Assume for contradiction that there is no hardcore distribution of high density. Then, by the Min-Max Theorem there is a *universal* predictor A^* such that for every $(\hat{X}, \hat{B})$ that is dense in (X, B) it holds that $\Pr\left[A^*(\hat{X}) = \hat{B}\right] > (1 + \epsilon)/2$. A^* is a distribution over circuits of size t, and its prediction probability is taken over this distribution as well as $(\hat{X}, \hat{B})$. By subsampling we can assume that A^* is uniform over a multiset of $S = O(\log(1/\epsilon\delta)/\epsilon^2)$ circuits of size t, while changing the advantage ϵ by at most a constant fraction. Given the universal predictor A^*, one can build a good predictor for B, contradicting the hardness of B given X, as formalized in Lemma 4.1:

Lemma 4.1 (From universal circuit to predictor [Imp]). *Let (X, B) be a joint distribution on $\{0,1\}^n \times \{0,1\}$. Let A^* be the uniform distribution over a multiset of S circuits of size t. Suppose for every joint distribution $(\hat{X}, \hat{B})$ that is δ-dense in (X, B) it holds that* $\Pr\left[A^*(\hat{X}) = \hat{B}\right] > (1 + \epsilon)/2$. *Then there is a circuit P of size $O(S \cdot t)$ such that* $\Pr[P(X) = B] > 1 - \delta$.

Specifically, we can let $P(x) = \text{majority}\{A(x) : A \in A^*\}$. Equivalently, $P(x)$ outputs 1 with probability

$$\frac{1}{2}\left(1 + \text{sign}\left(\Pr[A^*(x) = 1] - \frac{1}{2}\right)\right).$$

[1] The version we state is a slight generalization of the version in [KS], which only allows B to be a deterministic boolean function of X. However, the more general version follows readily from almost the same proof.

Unfortunately, both proofs in [Imp] yield a non-optimal hardcore density of δ. Following Nisan's proof using Min-Max Theorem, Holenstein [Hol1] proves the hardcore theorem with optimal hardcore density of 2δ (Theorem 4.1), by strengthening the above lemma to Lemma 4.2 below (using a trick from Levin's proof of the XOR Lemma).

Lemma 4.2 (From universal circuit to *optimal* predictor [Hol1]). *Let (X, B) be a joint distribution on $\{0,1\}^n \times \{0,1\}$. Let A^* be the uniform distribution over a multiset of S circuits of size t. Suppose for every joint distribution $(\hat{X}, \hat{B})$ that is 2δ-dense in (X, B) it holds that $\Pr\left[A^*(\hat{X}) = \hat{B}\right] > (1+\epsilon)/2$. Then there is a circuit P of size $O(S \cdot t)$ such that $\Pr[P(X) = B] > 1 - (1-\epsilon)\,\delta$.*

Specifically, we can let $P(x)$ output 1 with probability $p(x)$ truncated at 0 and 1 (i.e. $\min\{\max\{p(x), 0\}, 1\}$), for

$$p(x) = \frac{1}{2}\left(1 + \frac{\Pr[A^*(x) = 1] - \frac{1}{2}}{\phi}\right)$$

where ϕ is the least number s.t. $\Pr_{X,B}\left[\Pr_{A^*}\left[A^*(X) = B\right] \leq 1/2 + \phi\right] \geq 2\delta$. (WLOG ϕ is a multiple of $1/S$.)

The drawback of these proofs based on the standard Min-Max Theorem is that they are non-constructive, and that the complexity blow-up is non-optimal (with non-optimal settings of S due to probabilistic construction of the multiset).

A constructive proof such as the one by Impagliazzo [Imp] can be interpreted as a hardcore theorem for the *uniform* setting of hardness, where the hardness is with respect to efficient algorithms rather than small circuits. (See Theorem 4.2 below for the exact formulation). This *Uniform Hardcore Theorem* is needed for several important applications ([KS,Hol1,Hol2,HHR,HRV]). Building on the constructive proof in [Imp], Holenstein [Hol1] also shows a *uniform* hardcore theorem with optimal hardcore density, but is rather involved and fails to achieve the optimal complexity blow-up $O(\log(1/\delta)/\epsilon^2)$. Subsequently, Barak, Hardt, and Kale ([BHK]) gave an alternative proof of uniform hardcore theorem achieving optimal complexity blow-up of $O(\log(1/\delta)/\epsilon^2)$ (but without optimal hardcore density), based on ideas of multiplicative weights and Bregman projection.

As an application of the Uniform Min-Max Theorem (which itself is inspired by [BHK]), we offer a new proof of the Uniform Hardcore Theorem. Essentially, our proof simply replaces the use of Min-Max Theorem in Holenstein's proof (of the non-uniform hardcore theorem, Theorem 4.1) with the Uniform Min-Max Theorem. Consequently it has the advantages of (i) optimal hardcore density 2δ; (ii) optimal complexity blow-up $O(\log(1/\delta)/\epsilon^2)$; (iii) being more modular (e.g. compared to [BHK]) and simpler (e.g. compared to Holenstein's uniform proof [Hol1]).

Notation. For a distribution Z, let O_Z denote the oracle that gives a random sample from Z when queried.

Theorem 4.2 (Uniform Hardcore Theorem). *Let n be a security parameter, $m = m(n) = \text{poly}(n)$, $\delta = \delta(n)$, $\epsilon' = \epsilon'(n)$, $q = q(n)$ all computable in*

$\mathrm{poly}(n)$ *time, and* $(X, B) = g(U_m)$ *be a joint distribution where* $g : \{0,1\}^m \to \{0,1\}^n \times \{0,1\}$ *is computable in* $\mathrm{poly}(n)$ *time. Suppose that* (X, B) *has no hard-core distribution of density at least* 2δ*, i.e. there is a time* t *oracle algorithm* A *and infinitely many* n*, such that for every* $C \in \mathcal{C}_{m,2\delta}$*,*

$$\Pr_{(x,b)\leftarrow g(C)}\left[A^{O_C}(x) = b\right] > \frac{1}{2} + \epsilon'.$$

Then there is a time $\mathrm{poly}(t, n, 1/\delta, 1/\epsilon')$ *randomized algorithm* P *such that for infinitely many* n*,*

$$\Pr[P(X) = B] > 1 - \delta.$$

Moreover, P *is computable with* $O(\log(1/\delta)/\epsilon'^2)$ *oracle queries to* A*.*

Proof (Sketch). (See the full version [VZ2] for a complete proof). We will apply Theorem 3.1 (Uniform Min-Max Theorem), with

- $\mathcal{V} = \mathcal{C}_{m,2\delta}$;
- $\mathcal{W} = \{$(deterministic) circuits of size $tm + \mathrm{poly}(t)\}$;
- $f(z, W) = I(W(x) = b)$, where $(x, b) = g(z)$ and $I(\cdot)$ is the indicator function.

This corresponds to the two-player zero-sum game where Player 1 chooses some distribution $C \in \mathcal{C}_{m,2\delta}$, and Player 2 chooses a $tm+\mathrm{poly}(t)$ sized circuit W, with expected payoff $F(C, W) = \mathbb{E}[f(C, W)] = \Pr_{(x,b)\leftarrow g(C)}[W(x) = b]$ for Player 2. It turns out that using A, Algorithm 3.1 (Finding Universal Strategy) with KL projection on the set $\mathcal{V} = \mathcal{C}_{m,2\delta}$ can be implemented efficiently, such that for infinitely many n, in each iteration we obtain (from running A) some W with good prediction probability. This gives us an efficient universal predictor A^* of B given X, by the Uniform Min-Max Theorem. From the universal predictor, we then obtain a $(1 - \delta)$-predictor of B using Lemma 4.2, by searching for the correct ϕ.

5 Application: Impossibility of Black-Box Construction of Succinct Non-interactive Argument

A result of Gentry and Wichs [GW] shows that there is no black-box construction of succinct non-interactive arguments (SNARGs) from any natural cryptographic assumption (formally, they consider *falsifiable* cryptographic assumptions: ones that are defined by a polynomial-time security game). Their result relies on the (mild) assumption that there exist *hard subset membership problems*, which is equivalent to the existence of subexponentially hard one-way functions. One limitation is that they need to work in the non-uniform setting, in part due to their use of the Min-Max Theorem (in [GW] Lemma 3.1). In this section we show how to obtain the analogous result in the *uniform setting* by using the Uniform Min-Max Theorem. More specifically, assuming that there exist subexponentially hard one-way functions that are secure against uniform algorithms,

we show that there is no black-box construction of SNARGs based on cryptographic assumptions where security is measured against uniform algorithms (unless the assumption is already false).

A succinct non-interactive argument (SNARG) is a non-interactive argument system where the proof size is bounded by a fixed polynomial, for all instances and witnesses whose size can be an arbitrarily large polynomial. Formally,

Definition 5.1 (SNARG). *Let L be an* **NP** language associated with relation R. *We say that a tuple (G, P, V) of probabilistic polynomial-time (PPT) algorithms is a* succinct non-interactive argument for R *if the following properties hold:*

- ***Completeness:*** *For all $(x, w) \in R$, if we choose $(\mathsf{CRS}, \mathsf{PRIV}) \leftarrow G(1^n), \Pi \leftarrow P(\mathsf{CRS}, x, w)$,* then

$$\Pr\left[V\left(\mathsf{PRIV}, x, \Pi\right) = 0\right] = \mathrm{negl}(n).$$

- ***Soundness:*** *For every PPT algorithm (efficient* adversary*) A, if we choose $(\mathsf{CRS}, \mathsf{PRIV}) \leftarrow G(1^n), (X, \Pi) \leftarrow A(1^n, \mathsf{CRS})$,* then

$$\Pr\left[V\left(\mathsf{PRIV}, X, \Pi\right) = 1 \wedge X \notin L\right] = \mathrm{negl}(n).$$

- ***Succinctness:*** *For all $(x, w) \in \mathrm{supp}(X, W)$ and $\mathsf{crs} \in \mathrm{supp}(\mathsf{CRS})$, the length of the proof $\pi = P(\mathsf{crs}, x, w)$ is $|\pi| = \mathrm{poly}(n)(|x| + |w|)^{o(1)}$. We also consider a weaker variant called* slightly succinct*, where we require the length of a proof to be $|\pi| = \mathrm{poly}(n)(|x| + |w|)^{\alpha} + o(|x| + |w|)$ for some constant $\alpha < 1$.*[2]

Our notion of a falsifiable cryptographic assumption is analogous to [GW], except that the adversary A is a *uniform* algorithm instead of circuit:

Definition 5.2 (Falsifiable assumption). *Given an interactive PPT algorithm* Chal *(the challenger), the* uniform falsifiable (cryptographic) assumption (associated with) Chal *states that for all (uniform) PPT algorithms H, the probability that* $\mathrm{Chal}(1^n)$ *outputs a special symbol* win *after interacting with $H(1^n)$ is at most* $\mathrm{negl}(n)$ *for all sufficiently large n.*

For any randomized (possibly inefficient) function H, we let $\mathrm{Break}_H(n)$ *denote the above probability and say that H* breaks the assumption *if* $\mathrm{Break}_H(n) \geq 1/\mathrm{poly}(n)$ *for infinitely many n.*

Remark 5.1. An alternative definition of falsifiable assumption allows specifying a constant β, and says that the probability $\mathrm{Chal}(1^n)$ outputs win is at most $\beta + \mathrm{negl}(n)$. However, it turns out that setting $\beta = 0$, i.e. our definition above, is without loss of generality [HH]. We adopt the simpler definition because it is convenient for our proof.

[2] Earlier versions of [GW] contained a minor bug in the definition of slight succinctness. We use the corrected definition from the current version of their paper.

Next we define black-box reductions:

Definition 5.3 (Adversary and reduction). *For a randomized function A and a constant $c \in \mathbb{N}$, we say (A, c) is a* (G, P, V)-adversary *if $|A(1^n, \mathsf{crs})| \leq n^c$ and A violates the soundness condition infinitely often, i.e. if we choose $(\mathsf{CRS}, \mathsf{PRIV}) \leftarrow G(1^n), (X, \Pi) \leftarrow A(1^n, \mathsf{CRS})$, then*

$$\Pr\left[V\left(\mathsf{PRIV}, X, \Pi\right) = 1 \wedge X \notin L\right] \geq n^{-c}$$

for infinitely many n. We say (A, c) is an a.e. (G, P, V)-adversary *if A violates soundness for all sufficiently large n.*

A uniform black-box reduction showing the soundness of (G, P, V) based on a falsifiable assumption Chal *is a family of (uniform) probabilistic oracle algorithms $\{\mathrm{Red}_c\}$ (one for each $c \in \mathbb{N}$) such that for every (G, P, V)-adversary (A, c), $\mathrm{Red}_c^A(1^n)$ breaks the assumption and runs in time $\mathrm{poly}_c(n)$ (i.e. a polynomial that depends on c).*

For a probabilistic oracle algorithm Red*, we say a query $(1^m, \mathsf{crs})$ of $\mathrm{Red}(1^n)$ has* length m*. In general, $\mathrm{Red}(1^n)$ may make queries of various lengths. We say* Red *is* length-mapping *if for all n, all queries of $\mathrm{Red}(1^n)$ are of the same length $m = m(n)$;* denote this m by $\mathrm{query}_{\mathrm{Red}}(n)$*. Most reductions in cryptography set $m = n$ i.e. preserve length; that is, the security parameter of (G, P, V) is equal to that of the assumption.*

Following [GW], our results assume the existence of *hard subset membership problem.*

Definition 5.4 (Uniformly hard subset membership problem). *Let n be a security parameter, L be an* **NP** *language associated with relation R. We say $((X, W), U)$ is a* subset membership problem for R *if $(X, W) = (X, W)(n)$ is a $\mathrm{poly}(n)$-time samplable joint distribution whose support lies in R, and $U = U(n)$ a $\mathrm{poly}(n)$-time samplable distribution with $\Pr[U \notin L] \geq n^{-O(1)}$.*

A subset membership problem $((X, W), U)$ is a subexponentially hard *if X and U are $(2^{\Omega(n^\delta)}, 2^{-\Omega(n^\delta)})$-indistinguishable for a constant $\delta > 0$. We say it is* exponentially hard *if the above occurs and $|x| + |w| = O(n^\delta)$ for every $(x, w) \in \mathrm{supp}(X, W)$.*

This is a relatively mild assumption; for subexponentially hard subset membership problems, their existence is equivalent to the existence of subexponentially hard one-way functions.

Remark 5.2. Our definition of a hard subset membership problem is a variant of [GW] that is needed in the uniform setting, but also can be used in the nonuniform setting of [GW]. In [GW], they require that X is indistinguishable from a (not necessarily samplable) distribution U whose support is disjoint from L, whereas we require that U is samplable and allow it to hit L with negligible probability.

We now state the uniform analogue of the main result of [GW]. Compared to [GW], our Theorem 5.1 makes the weaker assumption of subexponentially hard

subset membership problem with respect to *uniform* algorithms, with the conclusion that a *uniform* falsifiable assumption cannot be broken also being weaker (unless the assumption is false).

Theorem 5.1 (Main theorem). *Let L be an* **NP** *language associated with relation R that has* a *subexponentially hard subset membership problem, and (G, P, V) be an non-interactive proof system for R that satisfies the completeness and succinctness properties. Then for every uniform falsifiable assumption* Chal*, one of the following must hold:*

- *The assumption* Chal *is false, or*
- *There is no uniform black-box reduction showing the soundness of (G, P, V) based on* Chal*.*

The same conclusion also holds if we assume an exponentially *hard subset membership problem, and (G, P, V) is only* slightly *succinct.*

To prove it in the nonuniform setting, the main idea of [GW] is showing that any SNARG (G, P, V) has an inefficient adversary A that can be (efficiently) "simulated" i.e. there exists an efficient algorithm Sim (the simulator) such that $\mathrm{Red}^A(1^n) \approx \mathrm{Red}^{\mathsf{Sim}}(1^n)$ for all PPT oracle algorithms Red (cf. [GW] Lemma 4.1). Thus, if there were a black-box reduction Red showing the soundness of (G, P, V) based on a falsifiable assumption, then Red^A would break the falsifiable assumption (since A is an adversary) and so would $\mathrm{Red}^{\mathsf{Sim}}$ (since $\mathrm{Red}^A(1^n) \approx \mathrm{Red}^{\mathsf{Sim}}(1^n)$). In other words, the assumption would be false.

To prove it in the uniform setting, we do the same showing that there is an adversary (A, c) that can be simulated by a *uniform* algorithm Sim, with several necessary tweaks:

Lemma 5.1 (Existence of simulatable adversary). *Let L be an* **NP** *language associated with relation R that has* a *subexponentially hard subset membership problem $((X, W), U)$, and (G, P, V) be a non-interactive proof system for R that satisfies the completeness and succinctness properties. Let n be a security parameter, $(\mathsf{PRIV}, \mathsf{CRS}) = G(1^n)$, $((X, W), U) = ((X, W), U)(n)$,* and *$\Pi = P(\mathsf{CRS}, X, W)$. Let $\ell = \ell(n) \geq n$ be a polynomial bound on the running time of $G(1^n)$ as well as the proof size $|\Pi|$, and c be a constant such that $|X| + |\Pi| \leq n^c$.*

Then for every length-mapping PPT oracle algorithm Red *such that $\mathrm{query}_{\mathrm{Red}}(k) = \omega(1)$, there is a PPT algorithm* Sim *and randomized function A satisfying:*

- *(A, c) is an a.e. (G, P, V)-adversary; and*
- Sim *simulates A: For all sufficiently large k, w.p. at least $1/\mathrm{poly}(k)$, $\mathsf{Sim}(1^k)$ outputs a randomized circuit B such that*

$$\mathrm{Break}_{\mathrm{Red}^A}(k) - \mathrm{Break}_{\mathrm{Red}^B}(k) = \mathrm{negl}(k).$$

(WLOG B only takes inputs $(1^n, \cdot)$ where $n = \mathrm{query}_{\mathrm{Red}}(k)$.)

The same conclusion also holds if we assume an exponentially *hard subset membership problem, and that* (G, P, V) *is only* slightly *succinct.*

Note that Lemma 5.1 is only stated for length-mapping reductions (unlike [GW]). We remove this restriction in the full version [VZ2] where we prove the main theorem (for which we use the fact that the simulatable adversary (A, c) is an a.e. adversary).

We defer the complete proof of Lemma 5.1 to the full version [VZ2], and offer an overview below.

Overview of Proof of Lemma 5.1. The proof is set up as follows. Given a subexponentially hard subset membership problem $((X, W), U)$, we can WLOG assume that X and U are $(2^{d\ell}, 2^{-d\ell})$-indistinguishable for a sufficiently large constant d, where $\ell = \ell(n)$ is a bound on the length of the proof output by $P(\mathsf{crs}, x, w)$ for $(x, w) \in \mathrm{supp}(X, W)$ and $\mathsf{crs} \in \mathrm{supp}(\mathsf{CRS})$. (If X and U are only $(2^{n^\delta}, 2^{-n^\delta})$-indistinguishable for some $\delta > 0$, we simply re-index, replacing $X(n)$ with $X((d\ell)^{1/\delta})$.) If $((X, W), U)$ is exponentially hard, we can also ensure that X and U are $(2^{d\ell}, 2^{-d\ell})$-indistinguishable by re-indexing so that $\ell \leq \mathrm{poly}(n) \cdot (|x|+|w|)^\alpha + o(|x|+|w|) = O(|x|+|w|)/d$ for all $(x, w) \in \mathrm{supp}(X, W)$ and $\mathsf{crs} \in \mathrm{supp}(\mathsf{CRS})$.

Consider the joint distribution (CRS, X, Π) where $\mathsf{CRS} = \mathsf{CRS}(n)$ is the distribution of common reference string, and $\Pi = \Pi(n)$ is the ℓ-bit proof produced by P for the instance/witness pair (X, W). Using the fact that Π is short (by succinctness), it turns out that the $2^{-d\ell}$-indistinguishability of X and U — and hence of (CRS, X) and (CRS, U), by samplability of CRS — implies there is no universal distinguisher D^* (as a $2^{O(\ell)}$ time algorithm) that $2^{-O(\ell)}$-distinguishes (CRS, X, Π) from all (CRS, U, Π'), where Π' is an ℓ-bit string arbitrarily jointly distributed with (CRS, U). This is extracted from the proof of a technical lemma of Gentry and Wichs ([GW] Lemma 3.1) and doesn't require the use of the Min-Max Theorem.

We consider the two-player zero-sum game where Player 1 selects a distribution Π' on $\{0,1\}^\ell$ jointly distributed with (CRS, U), then Player 2 selects a small circuit D and receives (expected) payoff $\mathbb{E}[D(\mathsf{CRS}, X, \Pi)] - \mathbb{E}[D(\mathsf{CRS}, U, \Pi')]$. Recall that the Uniform Min-Max Theorem – Average Case (Theorem 3.2) builds a sequence — which we denote by List_n — of Π' jointly distributed with (CRS, U), and says that if for each $\Pi' \in \mathsf{List}_n$ we can obtain a $2^{-O(\ell)}$-distinguisher D between (CRS, X, Π) and (CRS, U, Π') e.g. by some $2^{O(\ell)}$ time algorithm $\mathsf{FindDist}$, then we can obtain a universal $2^{-O(\ell)}$-distinguisher D^* (as a $2^{-O(\ell)}$ time algorithm) for *all* possible Π'. Since such D^* cannot exist (by previous discussion), it must be that for every $2^{O(\ell)}$ time algorithm $\mathsf{FindDist}$ there is some $\Pi' \in \mathsf{List}_n$ for which $\mathsf{FindDist}$ fails to produce a $2^{-O(\ell)}$-distinguisher D. Note that List_n actually depends on $\mathsf{FindDist}$ (indeed it is obtained by running Algorithm 3.2 using $\mathsf{FindDist}$ to select the actions for Player 2).

We will use $\mathsf{FindDist}$ to construct the simulatable adversary A as follows. Consider any $\Pi' \in \mathsf{List}_n$ for which $\mathsf{FindDist}$ fails to produce a $2^{-O(\ell)}$-distinguisher. We let A be the randomized function such that $A(1^n, \mathsf{CRS}) = (U, \Pi')$. For an appropriate choice of $\mathsf{FindDist}$, such an A will always be an a.e. adversary. Indeed,

if A is not an a.e. adversary then (PRIV, U, Π') does not pass the soundness test, whereas (PRIV, X, Π) passes the completeness test, hence we can use the verifier V to construct a distinguisher between (CRS, X, Π) and (CRS, U, Π'). Choosing $\mathsf{FindDist}$ to produce this distinguisher yields an a.e. adversary A.

Thus we only need to argue that, for an appropriate choice of $\mathsf{FindDist}$, A is also simulatable. Our simulation is the algorithm S such that $S(1^n, \mathsf{CRS}) = (X, \Pi)$. If we appropriately construct $\mathsf{FindDist}$ from the reduction Red and challenger Chal, then we can show that

$$\mathrm{Break}_{\mathrm{Red}^A}(k) - \mathrm{Break}_{\mathrm{Red}^S}(k) \leq 1/\mathrm{poly}(k) \cdot 2^{-O(\ell)},$$

where $\ell = \ell(n)$ for $n = \mathrm{query}_{\mathrm{Red}}(k)$. (Otherwise, we could use Red and Chal to construct a $2^{-O(\ell)}$-distinguisher between $(\mathsf{CRS}, A(1^n, \mathsf{CRS})) = (\mathsf{CRS}, X, \Pi)$ and $(\mathsf{CRS}, S(1^n, \mathsf{CRS})) = (\mathsf{CRS}, U, \Pi')$.) This completes the proof provided that $2^{-O(\ell)} \leq 1/\mathrm{poly}(k)$, which follows if Red does not make queries that are too short. If instead $2^{-O(\ell)} > 1/\mathrm{poly}(k)$, then we construct a simulator for A differently — by simply outputting a random element of List_n, which will equal A and be a perfect simulator w.p. $1/|\mathsf{List}_n| = 1/2^{O(\ell)} \geq 1/\mathrm{poly}(k)$. (Gentry and Wichs [GW] handle short queries using nonuniformity, by hardcoding all the answers.)

Acknowledgments. We thank Kai-Min Chung for many helpful discussions, especially on SNARGs.

References

BHK. Barak, B., Hardt, M., Kale, S.: The uniform hardcore lemma via approximate bregman projections. In: SODA 2009: Proceedings of the Nineteenth Annual ACM -SIAM Symposium on Discrete Algorithms, Philadelphia, PA, USA, pp. 1193–1200. Society for Industrial and Applied Mathematics (2009)

BSW. Barak, B., Shaltiel, R., Wigderson, A.: Computational analogues of entropy. In: Arora, S., Jansen, K., Rolim, J.D.P., Sahai, A. (eds.) RANDOM 2003 and APPROX 2003. LNCS, vol. 2764, pp. 200–215. Springer, Heidelberg (2003)

CLP. Chung, K.-M., Lui, E., Pass, R.: From weak to strong zero knowledge using a new non-black-box simulation technique (unpublished manuscript)

CT. Cover, T.M., Thomas, J.A.: Elements of information theory, 2nd edn. Wiley (2006)

DP. Dziembowski, S., Pietrzak, K.: Leakage-resilient cryptography. In: FOCS, pp. 293–302. IEEE Computer Society (2008)

FK. Frieze, A., Kannan, R.: Quick approximation to matrices and applications. Combinatorica 19(2), 175–220 (1999)

FR. Fuller, B., Reyzin, L.: Computational entropy and information leakage (2011), `http://www.cs.bu.edu/fac/reyzin`

FS. Freund, Y., Schapire, R.E.: Adaptive game playing using multiplicative weights. Games and Economic Behavior 29, 79–103 (1999)

GT. Green, B., Tao, T.: The primes contain arbitrarily long arithmetic progressions. Ann. of Math. 167(2), 481–547 (2008)

GW. Gentry, C., Wichs, D.: Separating succinct non-interactive arguments from all falsifiable assumptions. In: Fortnow, L., Vadhan, S.P. (eds.) STOC, pp. 99–108. ACM (2011)

HH. Haitner, I., Holenstein, T.: On the (Im)Possibility of key dependent encryption. In: Reingold, O. (ed.) TCC 2009. LNCS, vol. 5444, pp. 202–219. Springer, Heidelberg (2009)
HHR. Haitner, I., Harnik, D., Reingold, O.: Efficient pseudorandom generators from exponentially hard one-way functions. In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (eds.) ICALP 2006. LNCS, vol. 4052, pp. 228–239. Springer, Heidelberg (2006)
Hol1. Holenstein, T.: Key agreement from weak bit agreement. In: Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC), pp. 664–673 (2005)
Hol2. Holenstein, T.: Pseudorandom generators from one-way functions: A simple construction for any hardness. In: Halevi, S., Rabin, T. (eds.) TCC 2006. LNCS, vol. 3876, pp. 443–461. Springer, Heidelberg (2006)
HRV. Haitner, I., Reingold, O., Vadhan, S.: Efficiency improvements in constructing pseudorandom generators from one-way functions. In: Proceedings of the 42nd Annual ACM Symposium on Theory of Computing (STOC), pp. 437–446 (2010)
HW. Herbster, M., Warmuth, M.: Tracking the best linear predictor. Journal of Machine Learning Research 1, 281–309 (2001)
Imp. Impagliazzo, R.: Hard-core distributions for somewhat hard problems. In: Proceedings of the 36th Annual Symposium on Foundations of Computer Science (FOCS), pp. 538–545 (1995)
KS. Klivans, A.R., Servedio, R.A.: Boosting and hard-core set construction. Machine Learning 51(3), 217–238 (2003)
PJ. Pietrzak, K., Jetchev, D.: How to fake auxiliary input. In: ICITS 2012 Invited Talk (2012)
RTTV. Reingold, O., Trevisan, L., Tulsiani, M., Vadhan, S.: Dense subsets of pseudorandom sets. In: Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2008), October 26-28, pp. 76–85. IEEE (2008)
TTV. Trevisan, L., Tulsiani, M., Vadhan, S.: Regularity, boosting, and efficiently simulating every high-entropy distribution. In: Proceedings of the 24th Annual IEEE Conference on Computational Complexity (CCC 2009), July 15-18, pp. 126–136 (2009); Preliminary version posted as ECCC TR08-103
TZ. Tao, T., Ziegler, T.: The primes contain arbitrarily long polynomial progressions. Acta Math. 201(2), 213–305 (2008)
VZ1. Vadhan, S., Zheng, C.J.: Characterizing pseudoentropy and simplifying pseudorandom generator constructions. In: Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC 2012), May 19-22, pp. 817–836 (2012)
VZ2. Vadhan, S.P., Zheng, C.J.: A uniform min-max theorem with applications in cryptography. To appear on the Cryptology ePrint Archive (in preparation, 2013)

A Omitted Lemmas

Lemma A.1 (Multiplicative weight update decreases KL). *Let A, B be distributions over $[N]$ and $f : [N] \to [0,1]$ any function. Define random variable A' such that*

$$\Pr[A' = x] \propto e^{\epsilon \cdot f(x)} \Pr[A = x]$$

for $0 \leq \epsilon \leq 1$. Then $\mathrm{KL}(B \parallel A') \leq \mathrm{KL}(B \parallel A) - (\log e)\epsilon\,(\mathbb{E}[f(B)] - \mathbb{E}[f(A)] - \epsilon)$.

Proof. See the full version [VZ2].