SIAM J. DISCRETE MATH.
Vol. 27, No. 1, pp. 363–385

ADVERSARIAL LEAKAGE IN GAMES*

NOGA ALON†, YUVAL EMEK‡, MICHAL FELDMAN§, AND MOSHE TENNENHOLTZ¶

Abstract. While the minimax (or maximin) strategy has become the standard and most agreed-upon solution for decision making in adversarial settings, as discussed in game theory, computer science, and other disciplines, its power arises from the use of mixed strategies, also known as probabilistic algorithms. Nevertheless, in adversarial settings we face the risk of information leakage about the actual strategy instantiation. Hence, real robust algorithms should take information leakage into account. In this paper we introduce the notion of *adversarial leakage* in games, namely, the ability of a player to learn the value of b binary predicates about the strategy instantiation of her opponent. Different leakage models are suggested and tight bounds on the effect of adversarial leakage as a function of the level of leakage (captured by b) are established. The complexity of computing optimal strategies under these adversarial leakage models is also addressed. Together, our study introduces a new framework for robust decision making and provides rigorous fundamental understanding of its properties.

Key words. two-player zero-sum games, adversarial information leakage

AMS subject classifications. 91A05, 91A10, 91A40

DOI. 10.1137/110858021

1. Introduction.

Decision making is at the foundations of fields such as economics, operations research, and artificial intelligence. The question of what action should be taken by a decision maker when facing an uncertain environment, potentially consisting of other decision makers, is a fundamental problem which has led to a wide variety of models and solutions. The only type of situation for which this question has had an agreed-upon answer is in the context of two-player zero-sum games. This setting can model any situation in which a decision maker aims at maximizing her guaranteed payoff. When mixed strategies are allowed, such desired behavior, termed an agent's maximin (or safety level) strategy, leads to a well-defined expected payoff (known as the *value* of the game). Moreover, when presented explicitly in a matrix form, the computation of a maximin strategy is polynomial (by solving a linear program). Various equilibrium concepts have been considered in the game theoretic literature, but none of them provides prescriptive advice to a decision maker which will be as acceptable as the maximin strategy solution in adversarial settings. Since the introduction of the study of two-person zero-sum games [24], maximin strategies have received very little criticism (see [6] for an exception). Moreover, the safety level strategy has been advocated for some non-zero-sum settings as well (see [20], following observations by [5]).

Much of the power of a maximin strategy is associated with the use of mixed strategies, also known as randomized algorithms. In such algorithms the randomiza-

*Received by the editors December 6, 2011; accepted for publication (in revised form) November 29, 2012; published electronically February 12, 2013.
http://www.siam.org/journals/sidma/27-1/85802.html

†Tel Aviv University, Tel Aviv, Israel, and Microsoft Research, Herzelia, Israel (nogaa@post.tau.ac.il). This author's research was supported in part by an ERC advanced grant.

‡ETH Zurich, Zurich, Switzerland (yuval.emek@tik.ee.ethz.ch). This work was partially done while the author was at Microsoft Research, Herzelia, Israel.

§Hebrew University of Jerusalem, Jerusalem, Israel (mfeldman@huji.ac.il). This work was partially done while the author was at Microsoft Research, Herzelia, Israel.

¶Microsoft Research, Herzelia, Israel, and Technion–Israel Institute of Technology, Haifa, Israel (moshet@microsoft.com).

tion phase is assumed to be done in a private manner by the decision maker, and no information about the instantiation selected in that phase is assumed to be revealed. In reality, however, nothing is really private; for example, competitors will always strive to obtain the private actions of a business, possibly by means of industrial espionage [16] (see also [23]); hence, information leakage should be considered. As a result, it may be of interest to study the effects of adversarial leakage, where a limited amount of information on an agent's instantiation of her mixed strategy may leak in an adversarial manner. We believe that only by considering this situation will it be possible to construct robust strategies when acting in an adversarial setting.

Information leakage appeared in game theory in the context of conditioning a player's strategy about the other player's strategy [22, 14]; however, that work did not consider the leakage of mixed strategy instantiations nor its effects on designing robust algorithms in adversarial settings taking information leakage into account. Other game theoretic papers whose focus is related to information leakage studied purchasing noisy information [19], partial exposure in games with many players [9], spies in matrix games [17], and settings with costly randomness [7] or computation [21, 18, 10]. Another related topic is that of *leakage-resilient cryptography* [8], where the goal is to design cryptographic protocols that remain secure even if a bounded amount of information regarding the internal state leaks to the adversary.

Our model of adversarial leakage considers a two-player zero-sum game in strategic form (matrix form), depicted by an m by n matrix M: the MAX player (MAX) chooses some row $1 \leq i \leq m$ of M and the MIN player (MIN) chooses some column $1 \leq j \leq n$ of M; the outcome of the game is then dictated by the corresponding entry $M_{i,j}$ in M. This outcome is viewed as the payoff that MAX receives from MIN. The focus of the current paper is restricted to $\{0,1\}$-matrices M. This is known to be a highly applicable model, as it captures games in which a goal is either achieved or not.

We first study the *unidirectional leakage* model, where MAX is our decision maker and MIN is the adversary. MAX chooses a mixed strategy, that is, a distribution vector over her pure strategies. MIN may base her action on the value of b binary predicates defined on MAX's pure strategies; each such predicate is a Boolean formula on the set of strategies whose value is determined according to the actual instantiation of MAX's mixed strategy. The parameter b can be thought of as the amount of information leakage (or number of leaking bits) regarding the instance of MAX's mixed strategy. MAX would like to maximize her guaranteed expected payoff against any choice of such b binary predicates.

Two settings are studied, distinguished by the information structure assumed in them. Under the *strong* leakage setting, MAX chooses a mixed strategy, which is observable by MIN, who can then act upon it in determining the b predicates. On the other hand, under the *weak* leakage setting, MIN chooses the b predicates first, and MAX can observe it and act upon it in choosing her mixed strategy. (Refer to section 2 for formal definitions.)

Note that by von Neumann's minimax theorem, if MIN is allowed to choose the b predicates probabilistically, then weak leakage becomes equivalent to strong leakage in terms of the outcome that MAX (and MIN) can guarantee. However, the information structure we consider is such that MIN is restricted to deterministically choosing the b predicates. This clearly provides MAX with a potential advantage (compared to the strong leakage setting) and we are interested in understanding and quantifying this advantage. Other intriguing questions arise in this setting of unidirectional adversarial leakage: What would be the best mixed strategy for the MAX player? How well will

the original maximin strategy of the game perform? What is the computational complexity of finding the optimal strategy under information leakage?

We then turn to study a *bidirectional* leakage model. Here, we no longer assume a decision maker/adversary dichotomy; instead, each player has access to the answers of b binary predicates defined over the instantiation of her opponent's mixed strategy. To avoid a circular definition, we actually take the domain of these binary predicates to be the source of randomness used by the opponent. The main questions we study under this model are, Do two-player zero-sum games with bidirectional leakage admit a value? How does the number b of leaking bits affect the answer to the previous question?

Our results. For the strong unidirectional leakage, if the value of the game is $q = 1-\epsilon$ (for small positive ϵ) and 2^b is much smaller than $1/\epsilon$, then MAX can ensure an outcome close to 1 (at least $1 - 2^b\epsilon$), and this is tight. To do so, she simply uses the maximin strategy (that is, the optimal mixed strategy for the original game with no predicates). On the other hand, if 2^b is much bigger than $1/\epsilon$, then for every mixed strategy of MAX, the MIN player can ensure an outcome close to zero (at most $e^{-2^b\epsilon}$). Therefore, for every such game with value $1-\epsilon$, which is close to 1, a sharp transition occurs at b which is about[1] $\log(1/\epsilon)$: if b is slightly smaller, the outcome stays close to 1; if it is slightly larger, the outcome drops to nearly zero.

For games with value q bounded away from 1, even one bit enables MIN to square the outcome and drop it to at most q^2, and every additional bit squares the outcome again. There are also examples showing that this is essentially tight. Finally, for any fixed value $q < 1$, $\log\log m + O_q(1)$ bits suffice to enable MIN to drop the outcome to precisely 0.

The situation is different for weak unidirectional leakage. Clearly, here MAX is in a better shape; hence if the value of the game is $q = 1 - \epsilon$ (for small positive ϵ), MAX can still ensure an outcome close to 1 if the number of bits is much smaller than $\log(1/\epsilon)$ as in the setting of strong unidirectional leakage. For games with value q bounded away from 1, however, there are examples in which she can do much better than under strong unidirectional leakage and in fact can ensure no essential drop in the outcome as long as the number of leaking bits is somewhat smaller than $\log\log m$. More precisely, for any fixed value $0 < q < 1$ and for every large polynomially related[2] m, n, there are examples of games represented by a binary m by n matrix with value $q + o(1)$, so that even if $b = \log\log m - O(1)$, MAX can ensure that the outcome will stay roughly q. This should be contrasted with the strong leakage setting, where every additional bit squares MAX's outcome.

Somewhat surprisingly, once the number of leaking bits is slightly larger, that is, $b = \log\log m + O(1)$, the MIN player can already ensure a 0 outcome in any game with a fixed value $q < 1$. Thus, in the examples above a sharp transition occurs at nearly $b = \log\log m$ under weak unidirectional leakage: nearly $\log\log m$ bits have essentially no effect on the outcome, while slightly more bits already suffice to drop the outcome to 0.

Note that in contrast to leakage-free settings, where no advantage is gained by observing the opponent's mixed strategy (due to the minimax theorem), in settings of adversarial leakage, such information can contribute a great deal to the informed player, reflected by the advantage obtained by MAX in the weak leakage setting compared with the strong leakage setting.

[1] Unless stated otherwise, all logarithms are in base 2.

[2] We say that m and n are polynomially related if there exists some constant c such that $m \le n^c$ and $n \le m^c$.

With respect to computational complexity, computing the optimal strategy (for the MAX player) against b strongly leaking bits is poly-time for any fixed b, while this problem becomes NP-hard to compute, or even to approximate within any factor, for a general b. Under weak unidirectional leakage, the optimal strategy of MAX can be computed in polynomial time for every b. Computing the optimal leakage strategy of MIN under the weak leakage setting is NP-hard for any b.

Our last result concerns the bidirectional leakage model. In this context, we show that for every $\epsilon > 0$, there exists some $b_0 = b_0(\epsilon)$ such that in any game with $b \geq b_0$ bidirectional leaking bits, MAX (resp., MIN) can guarantee an expected payoff of at least $q - \epsilon$ (resp., at most $q + \epsilon$), where q is the value of the game, even if she reveals her strategy first.

It is important to point out that under the (strong and weak) unidirectional leakage models, a game typically does not admit a value in the sense that if we switch the roles of MAX and MIN so that MAX enjoys the benefit of having access to information leaking from MIN's strategy instantiation, then the expected payoff may change dramatically. (Clearly, in that case MAX can always guarantee an expected payoff of at least the value q.) This is in contrast to the bidirectional leakage model, where our result essentially demonstrates the existence of a value.

2. Model. We consider two-player zero-sum games defined by an m by n matrix M with $\{0,1\}$ entries,[3] where the rows correspond to MAX's pure strategies and the columns correspond to MIN's pure strategies: $M_{i,j}$ is the payoff that MAX receives from MIN if MAX and MIN play row $i \in [m]$ and column $j \in [n]$, respectively.[4] The matrix M is known to both players.

Unidirectional leakage. Given a matrix M and an integer $b \geq 0$, we describe a precise setting of adversarial *strong* unidirectional leakage, as follows:

(1) MAX chooses a distribution vector $\mathbf{p} = (p_1, \ldots, p_m)$ on $[m]$.
(2) MIN observes $\mathbf{p}$ and chooses a b-bit leakage function $f : [m] \to \{0,1\}^b$.
(3) MAX realizes $i \in [m]$ according to $\mathbf{p}$, i.e., chooses row i with probability p_i.
(4) MIN observes $f(i)$ and chooses an action $j \in [n]$.
(5) MAX receives a payoff of $M_{i,j}$ from MIN.

In the strong unidirectional leakage setting, MIN knows MAX's mixed strategy $\mathbf{p}$ and may base the choice of the leakage function f on that knowledge. We would also like to study the setting in which MIN is forced to choose the leakage function prior to MAX's choice of mixed strategy. This setting, referred to as *weak* unidirectional leakage, is cast in the following description:

(1) MIN chooses a b-bit leakage function $f : [m] \to \{0,1\}^b$.
(2) MAX observes f and chooses a distribution vector $\mathbf{p} = (p_1, \ldots, p_m)$ on $[m]$.
(3) MAX realizes $i \in [m]$ according to $\mathbf{p}$, i.e., chooses row i with probability p_i.
(4) MIN observes $\mathbf{p}$ and $f(i)$ and chooses an action $j \in [n]$.
(5) MAX receives a payoff of $M_{i,j}$ from MIN.

It will be convenient to formalize the choice of (pure) action made by MIN in step (4) of both the strong and weak unidirectional leakage descriptions as a function $g : \{0,1\}^b \to [n]$. Note that MIN decides on g when it already knows the mixed strategy $\mathbf{p}$ of MAX. This is less important under strong unidirectional leakage, where it can be assumed that MIN chooses g simultaneously with her choice of f; however,

[3] While we focus on the natural binary case, some of our results (specifically, Proposition 3.1 and Theorem 5.1) hold for any matrix with entries in $[0,1]$ as well, while others become noninteresting or are easily seen to be false.

[4] We use the standard notation $[k] = \{1, \ldots, k\}$, where k is a positive integer.

under weak unidirectional leakage the choice of g must be made at a later stage (when MIN already knows $\mathbf{p}$).

Given a matrix M, a nonnegative integer b, a distribution vector $\mathbf{p}$ on $[m]$, a function $f : [m] \to \{0,1\}^b$, and a function $g : \{0,1\}^b \to [n]$, let

$$u(M, b, \mathbf{p}, f, g) = \sum_{i \in [m]} p_i M_{i,g(f(i))} = \sum_{w \in \{0,1\}^b} \sum_{i : f(i) = w} p_i M_{i,g(w)}$$

denote the expected payoff of MAX in M with respect to b, $\mathbf{p}$, f, and g. Denote

$$u_{\mathbf{p}}(M, b) = \min_{f:[m]\to\{0,1\}^b \text{ and } g:\{0,1\}^b\to[n]} u(M, b, \mathbf{p}, f, g).$$

The (expected) payoff guaranteed by MAX in M against b unidirectional strongly leaking bits is defined as[5]

$$v^{\text{strong}}(M, b) = \max_{\mathbf{p} \in \Delta(m)} u_{\mathbf{p}}(M, b).$$

We denote by $\mathbf{p}_b^*$ a distribution vector that realizes $v^{\text{strong}}(M, b)$, i.e., $u_{\mathbf{p}_b^*}(M, b) = v^{\text{strong}}(M, b)$. The (expected) payoff guaranteed by MAX in M against b unidirectional weakly leaking bits is defined as

$$v^{\text{weak}}(M, b) = \min_{f:[m]\to\{0,1\}^b} \max_{\mathbf{p}\in\Delta(m)} \min_{g:\{0,1\}^b\to[n]} u(M, b, \mathbf{p}, f, g).$$

Observe that under this notation, $v^{\text{strong}}(M, 0) = v^{\text{weak}}(M, 0)$ is the classical value of (the two-player zero-sum game defined by) M. We shall denote this value by $v(M)$.

Bidirectional leakage. We now describe the bidirectional leakage model. An attempt to use the action space of one player as the domain of the leakage function of her opponent (as we did with the function f in the unidirectional setting) would result in a circular definition as each player should base her choice of action on some information regarding her opponent's choice of action. To avoid this obstacle, we first have to reformulate the way we address mixed strategies.

Let $\mathcal{R}_{\max}$ (resp., $\mathcal{R}_{\min}$) be the *source of randomness* of MAX (resp., MIN) out of which *nature* picks some element $r_{\max}$ (resp., $r_{\min}$) uniformly at random and hands it to MAX (resp., MIN); the random elements $r_{\max}$ and $r_{\min}$ are independent of each other. A mixed strategy of MAX is therefore defined as a function from $\mathcal{R}_{\max}$ to MAX's action space $[m]$, while a mixed strategy of MIN is defined as a function from $\mathcal{R}_{\min}$ to MIN's action space $[n]$.[6] It will be convenient to think of the sources of randomness $\mathcal{R}_{\max}$ and $\mathcal{R}_{\min}$ as the continuous interval $[0, 1]$. As such, we consider the elements $r_{\max}$ and $r_{\min}$, randomly picked by nature, as infinite length bit strings (standing for the binary representation of real numbers in $[0, 1]$). Having said that, it is important to point out that the mixed strategies designed throughout this paper use a finite number of random bits, so taking $\mathcal{R}_{\max}$ and $\mathcal{R}_{\min}$ to be sufficiently large finite sets is, in that sense, just as good.

[5]Given some set S, we use the standard notation $\Delta(S)$ to denote the collection of all probability distributions over S. For a positive integer k, we slightly abuse the notation and write $\Delta(k)$ instead of $\Delta([k])$ to denote the collection of all probability distributions over the set $[k] = \{1, \ldots, k\}$.

[6]This alternative view of mixed strategies can be easily adapted to arbitrary games with any number of players.

Now, in a bidirectional leakage scenario, the strategy of each player is augmented with a leakage function. Formally, the strategy of MAX with b bidirectional leaking bits consists of a function $f_{\max} : \mathcal{R}_{\min} \to \{0,1\}^b$ and a function $g_{\max} : \mathcal{R}_{\max} \times \{0,1\}^b \to [m]$, while the strategy of MIN with b bidirectional leaking bits consists of a function $f_{\min} : \mathcal{R}_{\max} \to \{0,1\}^b$ and a function $g_{\min} : \mathcal{R}_{\min} \times \{0,1\}^b \to [n]$. MAX's action is $g_{\max}(r_{\max}, f_{\max}(r_{\min}))$ and MIN's action is $g_{\min}(r_{\min}, f_{\min}(r_{\max}))$, where $r_{\max}$ and $r_{\min}$ are the elements randomly picked by nature from $\mathcal{R}_{\max}$ and $\mathcal{R}_{\min}$, respectively.

Remark. We could have, in fact, formulated the bidirectional leakage model with a different number of leaking bits for each of the two players. However, this would have complicated the model, the statement of the results, and the corresponding analysis without contributing any additional meaningful insight.

3. Strong unidirectional leakage.

Games with high value. We first show that for any m by n matrix with $\{0,1\}$ entries of value $q = 1 - \epsilon$, the MAX player can guarantee herself at least a payoff of $1 - 2^b\epsilon$. This can be done, in particular, by playing the maximin strategy.

PROPOSITION 3.1. *Let M be an m by n matrix with $\{0,1\}$ entries. Let $q = 1 - \epsilon$ be the value of the game defined by M, that is, $v(M) = 1 - \epsilon$. Then, for every $b \geq 0$, $u_{\mathbf{p}_0^*}(M, b) \geq 1 - 2^b\epsilon$.*

Proof. Let $\mathbf{p}_0^* = (p_1, \ldots, p_m)$. For every $w \in \{0,1\}^b$, let $S^w = \{i \in [m] | f(i) = w\}$, and let $p^w = \sum_{i \in S^w} p_i$. Fix some column j. Since $1 - \epsilon$ is the value of the game, it holds that for every w, $\sum_{i \in S^w} p_i M_{i,j} + \sum_{i \notin S^w} p_i M_{i,j} \geq 1 - \epsilon$. As $M_{i,j} \leq 1$ for every i, j, we have $\sum_{i \in S^w} p_i M_{i,j} + \sum_{i \in [m] \setminus S^w} p_i \geq 1 - \epsilon$. Substituting $\sum_{i \notin S^w} p_i = 1 - p^w$ and rearranging the last inequality yields

$$\sum_{i \in S^w} p_i M_{i,j} \geq p^w - \epsilon. \tag{1}$$

The expected payoff of MAX is given by the expression $\sum_{w \in \{0,1\}^b} \sum_{i:f(i)=w} p_i \cdot M_{i,g(w)}$ and the expected payoff of MAX conditioned on the event that some row $i \in S^w$ is played is given by the expression $\sum_{i:f(i)=w} \frac{p_i}{p^w} \cdot M_{i,g(w)}$, which is at least $\frac{1}{p^w}(p^w - \epsilon)$, by (1). Therefore the expected payoff of MAX is at least $\sum_{w \in \{0,1\}^b} p^w \frac{1}{p^w}(p^w - \epsilon) = 1 - 2^b\epsilon$. ∎

The above bound is tight, as established in the following proposition.

PROPOSITION 3.2. *For every $\epsilon > 0$ and every $b \geq 0$, there exists a matrix M with $\{0,1\}$ entries so that (1) $v(M) = 1 - \epsilon$ and (2) $u_{\mathbf{p}_0^*}(M, b) = u_{\mathbf{p}_b^*}(M, b) = 1 - 2^b\epsilon$.*

Proof. Let $n = 1/\epsilon$ and consider the n by n matrix M in which $M_{i,i} = 0$ for every i and $M_{i,j} = 1$ for every $i \neq j$. From symmetry considerations, both the maximin strategy and the optimal strategy against b leaking bits is the uniform distribution over the rows. Let f be a function which imposes the following partition on the rows: each one of the first $2^b - 1$ rows constitutes its own subset, and the remaining rows constitute the last subset. In this case, if one of the first $2^b - 1$ rows is chosen (each with probability ϵ), then MAX's payoff is 0, while if one of the remaining rows is chosen (with a total probability of $1 - (2^b - 1)\epsilon$), then the payoff obtained by the MAX player is $\frac{\frac{1}{\epsilon} - 2^b}{\frac{1}{\epsilon} - (2^b - 1)}$. The expected payoff of the MAX player is therefore $(1 - (2^b - 1)\epsilon) \cdot \frac{\frac{1}{\epsilon} - 2^b}{\frac{1}{\epsilon} - (2^b - 1)} = 1 - 2^b\epsilon$. ∎

Games with arbitrary value. The above two propositions essentially say that for games with value $q = 1-\epsilon$ and b such that $2^b\epsilon = o(1)$, MAX can guarantee a payoff of about q^{2^b} by playing the maximin strategy, and this is optimal. The case of general q and b, however, requires more work, and this is the focus of the following statement.

THEOREM 3.3. *Let M be an m by n matrix with $\{0,1\}$ entries. Let q be the value of the game defined by M, that is, $q = v(M)$. Then, for every $b \geq 0$ and every distribution vector $\mathbf{p}$ of the MAX player, $u_{\mathbf{p}}(M,b) \leq q^{2^b}$.*

Proof. Put $\mathbf{p}^{(1)} = \mathbf{p}$, and let $j_1 \in [n]$ be a pure strategy of MIN (a column of M) ensuring an expected payoff of $q_1 \leq q$ against the mixed strategy $\mathbf{p}^{(1)}$ of MAX. Such a pure strategy must exist since q is the value of the game. Define $S_1 = \{i \in [m] \mid M_{i,j_1} = 0\}$. It holds that $\sum_{i\in S_1} \mathbf{p}_i^{(1)} M_{i,j_1} + \sum_{i\in[m]-S_1} \mathbf{p}_i^{(1)} M_{i,j_1} = q_1$, hence $\sum_{i\in[m]-S_1} \mathbf{p}_i^{(1)} = q_1$.

Let $\mathbf{p}^{(2)}$ be the distribution vector defined by restricting $\mathbf{p}^{(1)}$ to the rows in $[m] - S_1$, namely,

$$\mathbf{p}_i^{(2)} = \begin{cases} \mathbf{p}_i^{(1)}/q_1 & \text{if } i \in [m] - S_1, \\ 0 & \text{otherwise.} \end{cases}$$

Let j_2 be a pure strategy of MIN ensuring an expected payoff of $q_2 \leq q$ against the mixed strategy $\mathbf{p}^{(2)}$ of MAX. Once again, such a pure strategy must exist since q is the value of the game. Define $S_2 = \{i \in [m] - S_1 \mid M_{i,j_2} = 0\}$. As before, it holds that $\sum_{i\in S_2} \mathbf{p}_i^{(2)} M_{i,j_2} + \sum_{i\in[m]-S_1-S_2} \mathbf{p}_i^{(2)} M_{i,j_2} = q_2$, hence $\sum_{i\in[m]-S_1-S_2} \mathbf{p}_i^{(2)} = q_2$.

Continuing in this manner for 2^b steps, we obtain 2^b pairwise disjoint subsets $S_1, \ldots, S_{2^b}$ of $[m]$ with corresponding columns $j_1, \ldots, j_{2^b}$ such that $M_{i,j_k} = 0$ for every $1 \leq k \leq 2^b$ and $i \in S_k$. For convenience we index the words in $\{0,1\}^b$ by $w_1, \ldots, w_{2^b}$ and fix

$$f(i) = \begin{cases} w_k & \text{if } 1 \leq k < 2^b \text{ and } i \in S_k, \\ w_{2^b} & \text{if } i \notin \cup_{k<2^b} S_k \end{cases}$$

and

$$g(w_k) = j_k \text{ for every } 1 \leq k \leq 2^b.$$

The above construction guarantees that when MAX plays according to $\mathbf{p}$ and MIN follows f and g, the payoff is 1 with probability at most $q_1 q_2 \cdots q_{2^b} \leq q^{2^b}$. It follows that $u_{\mathbf{p}}(M,b) = q_1 \cdots q_{2^b} \leq q^{2^b}$ as required. □

Using Theorem 3.3, we show that with $b = \log\log m + O(1)$ leaking bits, MIN can always ensure a 0 outcome; this relies on the following lemma.

LEMMA 3.4. *Let M be an m by n matrix with $\{0,1\}$ entries, and let q be the value of the game defined by M. If $q^{2^b} < 1/m$, then there exist functions $f : [m] \to \{0,1\}^b$ and $g : \{0,1\}^b \to [n]$ such that $M_{i,g(f(i))} = 0$ for every $i \in [m]$.*

Proof. Let $\mathbf{p}$ be the uniform distribution on $[m]$. Take f and g to be the functions promised to MIN against $\mathbf{p}$ by Theorem 3.3, that is, if MAX plays according to $\mathbf{p}$ and MIN follows f and g, then the expected payoff is at most $q^{2^b} < 1/m$. We argue that, in fact, the expected payoff in this case must be 0. Indeed, since $p_i = 1/m$ for every $i \in [m]$, a positive expected payoff is possible only if it is at least $1/m$, which derives a contradiction. Therefore, for every $w \in \{0,1\}^b$ and for every $i \in [m]$ such that $f(i) = w$, we must have $M_{i,g(w)} = 0$. The assertion follows. □

Clearly, if MIN plays according to the functions f and g promised by Lemma 3.4, then the expected payoff drops to 0 regardless of the mixed strategy of MAX.

COROLLARY 3.5. *Let M be an m by n matrix with $\{0,1\}$ entries, and let q be the value of the game defined by M. If $q^{2^b} < 1/m$, then for every distribution vector $\mathbf{p}$ of the MAX player, $u_{\mathbf{p}}(M,b) = 0$. Therefore, for every fixed $0 < q < 1$, taking $b = \log\log m + O_q(1)$ suffices for MIN to to ensure a 0 outcome.*

Remark. The corollary is essentially the known simple fact (proved in [11, 13]) that the ratio between the fractional cover and the integer cover of a hypergraph with m edges is at most $\ln m$.

The following theorem shows that both Theorem 3.3 and Corollary 3.5 are essentially tight.

THEOREM 3.6. *For every real $0 < q < 1$, for every integer $b \geq 0$, and for every large polynomially related m and n satisfying $q^{2^b} m > 2^b \log n$, there exists an m by n $\{0,1\}$-matrix M that satisfies*

(i) *$v(M) = q \pm o(1)$, where the $o(1)$-term tends to 0 as m and n grow; and*
(ii) *if $\mathbf{p} = (p_1, p_2, \ldots, p_m)$ is the uniform distribution on the rows, then $u_{\mathbf{p}}(M,b) \geq (1 - o(1))q^{2^b}$ (and thus $u_{\mathbf{p}}(M,b) = (1 \pm o(1))q^{2^b}$, by Theorem 3.3).*

In particular, for, say, $m = n^2$ and $b \leq \log\log m - \Theta(1)$, $u_{\mathbf{p}}(M,b) > 0$.

Proof. Let M be a random m by n matrix with $\{0,1\}$-entries obtained by choosing each entry $M_{i,j}$, randomly and independently, to be 1 with probability q and 0 with probability $1-q$. We show that M satisfies the assertion of the theorem with positive probability.

Since m, n are large and are polynomially related, almost surely (that is, with probability that tends to 1 as m, n tend to infinity) every row of M has $(1 \pm o(1))qn$ 1-entries, and every column of M has $(1 \pm o(1))qm$ 1 entries. This follows easily by the standard known estimates for Binomial distributions; see, for example, [4]. This implies that the value of the game is $(1 \pm o(1))q$: indeed, if MAX (respectively, MIN) plays according to the uniform distribution on the rows (resp., columns), then it guarantees an expected payoff of at least (resp., at most) $(1 \pm o(1))q$. Thus (i) holds almost surely.

We establish the assertion by showing that (ii) holds with high probability as well. For that purpose, we argue that for every choice of a set $J \subset [n]$ of size $|J| = 2^b$, the number of indices $i \in [m]$ so that $M_{i,j} = 1$ for all $j \in J$ is, almost surely, $(1 \pm o(1))q^{2^b} m$. Indeed, for a fixed choice of a set J, the random variable X that counts the number of such indices i is a Binomial random variable with parameters m and q^{2^b}. Therefore the probability that X is not $(1 \pm o(1))q^{2^b} m$ decreases exponentially with $q^{2^b} m > 2^b \log n = \log(n^{2^b})$. The assertion is established by applying the union bound over all $\binom{n}{2^b} < n^{2^b}$ choices of the set J. ∎

Remark. Theorem 3.6 is proved using a probabilistic argument: we show that a random matrix M satisfies the promise of the theorem with positive probability. In Appendix A we present two explicit constructions of the desired matrix M using either finite geometries or character sum estimates (Weil's theorem [25]). In general, one can use any construction of small sample spaces supporting nearly 2^b-wise independent binary random variables to supply additional examples. Indeed, this guarantees that in a random row, every subset of 2^b entries appears nearly independent, and in particular, they are all 1 with probability close to q^{2^b}. (Recall that this is the main argument in Theorem 3.6's proof.) General constructions of small sample spaces supporting nearly 2^b-wise independent random variables are developed in [15, 2].

4. Weak unidirectional leakage.

The adversarial weakness. Arbitrarily large matrices M with the property that MIN can significantly decrease the expected payoff with a small number of leaking bits are easy to construct even under the weak leakage setting as long as $v(M)$ is bounded away from 1. For example, if M is obtained from a constant size binary matrix M' (that does not admit any row of 1s) by replacing each 1-entry (resp., 0-entry) with an arbitrarily large block of 1s (resp., 0s), then clearly, $v(M) = v(M')$, i.e., a constant value, whereas a constant number of leaking bits is sufficient for MIN to guarantee a 0 outcome. In that regard, the weak leakage setting is no different from the strong leakage setting.

The interesting question, though, concerns the guarantees that hold for arbitrary matrices, specifically those promised by Theorem 3.3 and Corollary 3.5. The following result shows that in sharp contrast to the situation with the strong leakage setting, under the weak leakage setting there are examples in which $\log\log m - O(1)$ bits of information do not enable the MIN player to gain any significant advantage.

THEOREM 4.1. *For every real q, $0 < q < 1$, for every positive δ, and for all large polynomially related n, m satisfying*

$$\left[\left(\frac{10}{\delta}\right)\right]^{n2^b} < \delta^{-m/\sqrt{n}} \quad \text{and} \quad \left[\frac{q(1-q)}{10}\right]^{2^b} \geq \frac{1}{\sqrt{n}},$$

there is an m by n matrix M with $\{0,1\}$-entries so that the value $v(M)$ of the game it determines is $q \pm o(1)$ and $v^{weak}(M,b) \geq q - \delta$. In particular, if $n = m^2$ and $b = \log\log m - \Theta(1)$, $v^{weak}(M,b)$ is essentially equal to $v(M)$.

The proof of the above theorem is more complicated than the ones in the previous section and requires several preparations. We need the following known result.

LEMMA 4.2 (see [3, Lemma 3.2]). *Let Y be a random variable with expectation $\mathbb{E}[Y] = 0$, variance $\mathbb{E}[Y^2]$, and fourth moment $\mathbb{E}[Y^4] \leq k(\mathbb{E}[Y^2])^2$, where k is a positive real. Then $\mathbb{P}[Y \geq 0] \geq \frac{1}{2^{4/3}k}$.*

Using the above lemma, we prove the following.

LEMMA 4.3. *Consider some real $0 < q < 1$ and let $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ be a distribution vector on $[n]$. Let $X_1, X_2, \ldots, X_n$ be independent identically distributed indicator random variables, where each X_j is 1 with probability q (and 0 with probability $1-q$). Define $X = \sum_{j=1}^n X_j p_j$. Then the probability that X is at least its expectation (which is q) is bigger than $\frac{q(1-q)}{10}$.*

Proof. Define $Y_j = X_j p_j - \mathbb{E}[X_j p_j] = X_j p_j - q p_j$, and $Y = \sum_j Y_j$. By linearity of expectation, $Y = X - \mathbb{E}[X]$, and $\mathbb{E}[Y] = 0$. In order to apply the previous lemma, we compute the variance of Y and estimate its fourth moment:

$$\begin{aligned}
\mathrm{Var}[Y] &= \sum_j \mathrm{Var}[Y_j] \\
&= \sum_j [q(1-q)^2 p_j^2 + (1-q) q^2 p_j^2] \\
&= q(1-q) \sum_j p_j^2;
\end{aligned}$$

similarly,

$$\begin{aligned}\mathbb{E}[Y^4] &= \sum_j \mathbb{E}[Y_j^4] + 6\sum_{i<j}\mathbb{E}[Y_i^2]\mathbb{E}[Y_j^2] \\ &= \sum_j [q(1-q)^4 p_j^4 + (1-q)q^4 p_j^4] + 6\sum_{i<j} q^2(1-q)^2 p_i^2 p_j^2 \\ &\le q(1-q)\sum_j p_j^4 + 6\sum_{i<j} q^2(1-q)^2 p_i^2 p_j^2 \\ &\le \frac{1}{q(1-q)}\left[q^2(1-q)^2\sum_j p_j^4 + 6\sum_{i<j} q^2(1-q)^2 p_i^2 p_j^2\right] \\ &\le \frac{3}{q(1-q)}(\mathrm{Var}[Y])^2.\end{aligned}$$

The desired result now follows from Lemma 4.2 (using the fact that $2^{4/3}\cdot 3 < 10$). □

Remark. For $q \le 1/2$ the estimate in the lemma is tight, up to a constant factor. Indeed, for $\mathbf{p} = (1, 0, 0, \ldots, 0)$ the probability that X is at least q is precisely the probability that $X_1 = 1$, which is q. For $q = 1/k$ with k being an integer there is a simpler argument showing that in this case the probability that X is at least its expectation is at least q (which is precisely tight). The idea is to choose the random vector $(X_1, X_2, \ldots, X_n)$ by first choosing, for each $1 \le j \le n$, a random uniform number n_j in $\{1, 2, , \ldots, k\}$ with all choices being independent and then by selecting a uniform random $Z \in \{1, 2, \ldots, k\}$, defining X_j to be 1 if and only if $n_j = Z$. Since the sum $\sum_{Z\in[k]}(\sum_{j:n_j=Z} p_j) = 1$, it follows that for each choice of the values n_j, there is at least one Z so that $\sum_{j:n_j=Z} p_j \ge 1/k$, and therefore the probability that the obtained random sum is at least $q = 1/k$ is at least $1/k$, as claimed. Note that for some values of q the probability that X is at least q is strictly smaller than q. Indeed, for example, if $q = 0.501$ and the vector $\mathbf{p}$ is $(0.5, 0.5, 0, 0, \ldots, 0)$, then the probability that X is at least q is the probability that $X_1 = X_2 = 1$, which is q^2, that is, roughly $q/2$.

The next ingredient required for our proof is a special case of the FKG inequality (see, e.g., [4, Chapter 6]). Let U be the collection of all length d binary vectors. Consider the following natural partial order on U: for two vectors $x, y \in U$, the relation $x \le y$ holds if $x(i) \le y(i)$ for every $i \in [d]$. A binary function $f : U \to \{0, 1\}$ is said to be *increasing* if $f(x) = 1$ implies $f(y) = 1$ for every two vectors $x, y \in U$ satisfying $x \le y$.

LEMMA 4.4 (see [4]). *Let $\mathcal{F}$ be a finite collection of increasing binary functions on U and let x be a vector chosen randomly from U by picking each coordinate $x(i)$ according to some specified probability distribution P_i, independently of all other coordinates. Then*

$$\mathbb{P}_x\left[\prod_{f\in\mathcal{F}} f(x) = 1\right] \ge \prod_{f\in\mathcal{F}} \mathbb{P}_x\left[f(x) = 1\right].$$

We are now ready to establish the following corollary.

COROLLARY 4.5. *Let w be a random vector of length n with $\{0, 1\}$ entries obtained by selecting each entry, randomly and independently, to be 1 with probability q and 0*

with probability $1-q$. *Let* $\mathbf{P}$ *be any fixed set of distribution vectors of length* n. *Then, the probability that the inner product of the vector* w *with each of the vectors* $\mathbf{p} \in \mathbf{P}$ *is at least* q *is at least* $(\frac{q(1-q)}{10})^{|\mathbf{P}|}$.

Proof. The proof relies on the following application of Lemma 4.4. Take $d = n$. Define the probability distribution P_i, $i \in [d]$, so that 1 is chosen with probability q and 0 is chosen with probability $1-q$ and identify the random vector $x \in U$ from the statement of Lemma 4.4 with the random vector $w \in \{0,1\}^n$ from the statement of the current lemma. For each $\mathbf{p} \in \mathbf{P}$, define the binary function $f_{\mathbf{p}} : \{0,1\}^n \to \{0,1\}$ by setting

$$f_{\mathbf{p}}(v) = \begin{cases} 1 & \text{if } \langle v, \mathbf{p} \rangle \geq q, \\ 0 & \text{otherwise} \end{cases}$$

for every vector $v \in \{0,1\}^n$, observing that all functions $f_{\mathbf{p}}$ are increasing. Applying Lemma 4.4, we conclude that

$$\mathbb{P}_{w \in \{0,1\}^n} \left[\prod_{\mathbf{p} \in \mathbf{P}} f_{\mathbf{p}}(w) = 1 \right] \geq \prod_{\mathbf{p} \in \mathbf{P}} \mathbb{P}_{w \in \{0,1\}^n} \left[f_{\mathbf{p}}(w) = 1 \right],$$

that is, the probability that the inner product of w with each of the vectors $\mathbf{p} \in \mathbf{P}$ is at least q is greater than or equal to the product of these probabilities. But according to Lemma 4.3, for every $\mathbf{p} \in \mathbf{P}$, the probability that the inner product of w with $\mathbf{p}$ is at least q is greater than $\frac{q(1-q)}{10}$. The desired result follows. ∎

We are now ready to state the proof of Theorem 4.1.

Proof of Theorem 4.1. Let M be a random m by n matrix with $\{0,1\}$-entries obtained by choosing each entry $M_{i,j}$, randomly and independently, to be 1 with probability q and 0 with probability $1-q$. As specified in the proof of Theorem 3.6, the value of this game is almost surely $v(M) = q \pm o(1)$.

Recall that in a weak unidirectional leakage setting, MIN first chooses a function $f : [m] \to \{0,1\}^b$. This can be thought of as partitioning the rows of M into 2^b disjoint subsets, each defining a submatrix of M—denote these submatrices by $A_1, \ldots, A_{2^b}$. Each submatrix A_i corresponds to a two-player zero-sum game by its own right. Given the choice of f, the expected payoff guaranteed by MAX in M against b weakly leaking bits is $v^{\text{weak}}(M,b) = \max_{1 \leq i \leq 2^b} v(A_i)$. Therefore, together with the choice of the function f, MIN essentially chooses a set $\mathbf{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_{2^b}\}$ of distribution vectors of length n, where $\mathbf{p}_i$ realizes $v(A_i)$. In order to show that $v^{\text{weak}}(M,b) \geq q - \delta$, it suffices to prove that in each submatrix A_i, there is some row r_i such that the inner product of r_i and $\mathbf{p}_i$ is at least $q - \delta$. This will be established by arguing that almost surely, for every choice of a set $\mathbf{P}$, there exists some row M_i in M so that $M_i \cdot \mathbf{p} \geq q - \delta$ for every $\mathbf{p} \in \mathbf{P}$. To that end, we take some δ-net $\mathcal{N}$ of distributions of length n with respect to the ℓ_1-norm and show that almost surely, for every $\mathbf{P} \subseteq \mathcal{N}$, $|\mathbf{P}| = 2^b$, there exists some row M_i in M so that $M_i \cdot \mathbf{p} \geq q$ for every $\mathbf{p} \in \mathbf{P}$.

It is a standard fact that there exists a δ-net $\mathcal{N}$ of distribution vectors of length n with respect to the ℓ_1-norm so that $|\mathcal{N}| \leq \left(\frac{10}{\delta}\right)^n$. This δ-net exhibits at most $\left(\frac{10}{\delta}\right)^{n2^b}$ ways to choose a subset of size 2^b. Fix some set $\mathbf{P} \subseteq \mathcal{N}$, $|\mathbf{P}| = 2^b$, and fix some row M_i in M. By the last corollary, the probability that $M_i \cdot \mathbf{p} \geq q$ for every $\mathbf{p} \in \mathbf{P}$ is at least $(\frac{q(1-q)}{10})^{2^b}$. Hence, the probability that none of the m rows ensures MAX a payoff of at least q with each of the mixed strategies $\mathbf{p} \in \mathbf{P}$ is at most $(1 - (\frac{q(1-q)}{10})^{2^b})^m$. Applying the union bound, with probability at least $1 - (1 - (\frac{q(1-q)}{10})^{2^b})^m (\frac{10}{\delta})^{n2^b}$, for

every set $\mathbf{P} \subseteq \mathcal{N}$, $|\mathbf{P}| = 2^b$, there exists some row M_i in M such that $M_i \cdot \mathbf{p} \geq q$ for every $\mathbf{p} \in \mathbf{P}$. The assertion follows by applying the inequalities specified in the statement of the theorem. □

Paley games. The proof of Theorem 4.1 relies on a probabilistically constructed matrix $M \in \{0,1\}^{m \times n}$ that satisfies the following property almost surely. For any collection $\mathbf{P}$ of 2^b distribution vectors of length n, there exists some $i \in [m]$ such that for every distribution vector $\mathbf{p} \in \mathbf{P}$, the inner product of the ith row in M and $\mathbf{p}$ is almost as high as the value $v(M)$ of the game defined by M. This result raises two questions. First, in light of the bidirectional leakage model discussed in section 6, does there exist such a matrix M with a similar guarantee for both MAX and MIN? Second, does there exist an explicit construction of such a matrix M? It turns out that these two questions can be answered in the affirmative.

Let n be an odd prime, let Z_n denote the finite field with n elements, and let $\chi(x)$ denote the quadratic character function defined on Z_n, namely, $\chi(0) = 0$; $\chi(x) = 1$ if $x \neq 0$ is a quadratic residue modulo n and $\chi(x) = -1$ if $x \neq 0$ is a quadratic nonresidue. Let $A = A_n$ denote the n by n matrix defined by $A[i,j] = \chi(i-j)$. Note that A_n is symmetric if $n \equiv 1 \pmod 4$ and A_n is antisymmetric if $n \equiv 3 \pmod 4$.

Let $M = M_n \in \{0,1\}^{n \times n}$ be the matrix obtained from A_n by arbitrarily changing the 0-entries on the diagonal of A to be either 1 or -1 and then setting $M[i,j] = \frac{A[i,j]+1}{2}$ for every $i,j \in [n]$; in other words, the -1-entries of A turn into 0-entries in M and the 1-entries in A turn into 1-entries in M. As the number of 0-entries and the number of 1-entries in each row and each column of M are $(n \pm 1)/2$, it follows that the value $v(M)$ of M is $1/2 \pm o(1)$.

THEOREM 4.6. *For every integer $B \geq 1$ and any real $\epsilon > 0$, there is an $n_0 = n_0(\epsilon, B)$ so that the following holds for every prime $n > n_0$. For any collection $\mathbf{P}$ of B distribution vectors of length n, there is a row of M whose inner product with each of the vectors in $\mathbf{P}$ is at least $1/2 - \epsilon$ and a column of M whose inner product with each of the vectors in $\mathbf{P}$ is at most $1/2 + \epsilon$.*

In order to establish Theorem 4.6, we first have to state the following two lemmas.

LEMMA 4.7. *Every vector $r = (r_1, \ldots, r_n)$ of nonnegative reals whose sum of coordinates is at most 1 satisfies $||Ar||_2^2 \leq n \cdot \max_i\{r_i\}$.*

Proof. Note that $A^t A = nI - J$, where I is the identity matrix of dimension n and J is the $n \times n$ all 1 matrix. Therefore,

$$\|Ar\|_2^2 = r^t A^t A r = r^t(nI - J)r = \sum_{i=1}^{n} n r_i^2 - \left(\sum_{i=1}^{n} r_i\right)^2$$
$$\leq n \cdot \max_i\{r_i\} \cdot \left(\sum_{i=1}^{n} r_i\right) \leq n \cdot \max_i\{r_i\}.$$

The assertion follows. □

The next lemma is a known consequence of Weil's theorem [25]; see, e.g., [1] for a proof.

LEMMA 4.8. *For any positive integer $k \leq 0.5 \log n$, any subset $K \subset [n]$ of size $|K| = k$, and any values $\delta_j \in \{-1, 1\}$ for $j \in K$, the number of rows i of A so that $A[i,j] = \delta_j$ for all $j \in K$ deviates from $\frac{n}{2^k}$ by at most $k \cdot n^{1/2}$.*

We are now ready to establish Theorem 4.6.

Proof of Theorem 4.6. We prove the existence of a row with the required properties; the proof of existence of a column as needed is essentially identical. The proof relies on showing that for any collection $\mathbf{P}$ as in the statement of the theorem, there

is a row of $A = A_n$ whose inner product with each of the vectors in $\mathbf{P}$ is at least $-\epsilon$. The assertion follows by the construction of M.

Given a collection $\mathbf{P}$ as in the statement of the theorem, let n be some sufficiently large prime. Specifically, we assume that n is sufficiently large with respect to B and ϵ; the exact dependency on B and ϵ will be revealed in the course of the proof. Let $Q = \lfloor \frac{\log n}{2B} \rfloor$. For simplicity, we omit all floor signs from now on; it is easy to check that this is not essential. For each vector $p = (p_1, \ldots, p_n) \in \mathbf{P}$ let $K(p)$ be the (possibly empty) set of coordinates j for which $p_j \geq \frac{1}{Q}$. Note that $|K(p)| \leq Q$ for every $p \in \mathbf{P}$. Let p^{large} be the vector obtained from p by keeping its coordinates p_j with $j \in K(p)$ and replacing each of its other coordinates by 0; let $p^{\text{small}} = p - p^{\text{large}}$.

Fix $S = \cup_{p \in \mathbf{P}} K(p)$ and recall that $p_j^{\text{large}} = 0$ for every $p \in \mathbf{P}$ and $j \notin S$. Let $\{-1,1\}^S$ be the set of all vectors over $\{-1,1\}$ with $|S|$ coordinates identified with the indices in S and consider some random vector $w \in \{-1,1\}^S$. Given an arbitrary vector $p \in \mathbf{P}$, the probability that

$$\sum_{j \in S} w(j) \cdot p_j^{\text{large}} \geq 0 \tag{2}$$

is at least $1/2$ since from any pair consisting of a vector $w \in \{-1,1\}^S$ and its additive inverse, at least one of the vectors satisfies (2).

We now apply Lemma 4.4 as follows. Put $d = |S|$. Take the probability distribution P_i, $i \in [d]$, to be the uniform distribution over $\{-1,1\}$ and identify the random vector $x \in U$ from the statement of Lemma 4.4 with the aforementioned random vector $w \in \{-1,1\}^S$. For each vector $p \in \mathbf{P}$, define the binary function $f_p : \{-1,1\}^S \to \{0,1\}$ by setting

$$f_p(v) = \begin{cases} 1 & \text{if } \sum_{j \in S} v(j) \cdot p_j^{\text{large}} \geq 0, \\ 0 & \text{otherwise} \end{cases}$$

for every vector $v \in \{-1,1\}^S$, observing that all functions f_p are increasing. Applying Lemma 4.4, we conclude that

$$\mathbb{P}_{w \in \{-1,1\}^S} \left[\prod_{p \in \mathbf{P}} f_p(w) = 1 \right] \geq \prod_{p \in \mathbf{P}} \mathbb{P}_{w \in \{-1,1\}^S} \left[f_p(w) = 1 \right],$$

that is, the probability that w satisfies (2) with the vector p^{large} for all $p \in \mathbf{P}$ is at least $\frac{1}{2^B}$. We refer to such a vector $w \in \{-1,1\}^S$ as a *good* vector. Thus, there exist at least $2^{|S|-B}$ good vectors.

A row A_i of A is said to be *good* if there exists some good vector $w \in \{-1,1\}^S$ such that A_i agrees with w on the coordinates in S. Since $|S| \leq QB \leq \log(n)/2$, Lemma 4.8 guarantees that each good vector $w \in \{-1,1\}^S$ contributes at least $\frac{n}{2^{|S|}} - |S| n^{1/2}$ good rows, which sums up to at least

$$2^{|S|-B} \left(\frac{n}{2^{|S|}} - |S| n^{1/2} \right) = \frac{n}{2^B} - 2^{QB-B} QB n^{1/2} > \frac{n}{2^{B+1}}$$

good rows, where the last inequality holds provided that n is sufficiently large.

Lemma 4.7 implies that

$$\sum_{p \in \mathbf{P}} \sum_{i \in [n]} \langle A_i, p^{\text{small}} \rangle^2 = \sum_{p \in \mathbf{P}} \| A p^{\text{small}} \|_2^2 \leq \frac{Bn}{Q},$$

and thus there is a good row A_i so that

$$\sum_{p\in\mathbf{P}}\langle A_i, p^{\text{small}}\rangle^2 \le \frac{Bn/Q}{n/2^{B+1}} = \frac{B2^{B+1}}{Q} = \frac{4B^2\cdot 2^{B+1}}{\log n}.$$

The last quantity is at most ϵ^2 provided n is sufficiently large, hence $\langle A_i, p^{\text{small}}\rangle \ge -\epsilon$ for all $p \in \mathbf{P}$. The proof is completed by recalling that $\langle A_i, p^{\text{large}}\rangle \ge 0$ as A_i is a good row. □

A slightly increased leakage. Somewhat surprisingly, even under the weak unidirectional leakage setting, although there are examples in which the MIN player cannot decrease the expected payoff by much using at most $\log\log m - O(1)$ bits of information, if she is allowed to use $\log\log m + O(1)$ bits, she can always decrease the expected payoff to 0. This is described in the next (simple) result, whose proof follows directly from Lemma 3.4. Together with Theorem 4.1, this exhibits an unexpected sharp phase transition at $b = \log\log m$.

THEOREM 4.9. *Let M be an m by n matrix with $\{0,1\}$ entries, and let q be the value of the game defined by M. If $q^{2^b} < 1/m$, then $v^{weak}(M,b) = 0$. Therefore, for every fixed $0 < q < 1$, taking $b = \log\log m + O_q(1)$ suffices for MIN to ensure a 0 outcome, even under the weak unidirectional leakage setting.*

5. Unidirectional leakage—optimal strategy computation. In this section we study the complexity of various computations in the adversarial unidirectional leakage model. We begin with a simple example of a $\{0,1\}$ matrix M with a maximin strategy $\mathbf{p}_0^*$ that satisfies (i) $u_{\mathbf{p}_0^*}(M,0) = 1/2$ and (ii) $u_{\mathbf{p}_0^*}(M,1) = 0$. On the other hand, there exists another mixed strategy $\mathbf{p}$ of MAX that satisfies (i) $u_{\mathbf{p}}(M,0) = 3/7$ and (ii) $u_{\mathbf{p}_1^*}(M,1) \ge 1/7$. This shows that playing the maximin strategy may be a naive behavior for $b > 0$ and hence motivates the computation of better strategies. The matrix M showing the above is of dimension 9×14 and is depicted in Table 1. The main ingredient in the construction is a 7 by 7 matrix T_7 with $\{0,1\}$ entries that satisfies the following properties: (1) every row and every column of T_7 contain exactly three 1-entries, and (2) for every choice of $1 \le j \le j' \le 7$, there exists some $1 \le i \le 7$ such that $M_{i,j} = M_{i,j'} = 1$. (Refer to Example 1 in Appendix A.) Playing the first two rows with probability 1/2 each yields an expected payoff of 1/2 for MAX. One can easily verify that this is a unique optimal strategy when $b = 0$, while its expected payoff is clearly 0 when $b = 1$. Yet, by playing the uniform distribution on the bottom seven rows, MAX ensures an expected payoff of 3/7 when $b = 0$ and an expected payoff of at least 1/7 when $b = 1$.

MAX's optimal strategy. We now turn to consider the computational complexity of finding the optimal strategy for the MAX player under the strong leakage setting. The following theorem shows that computing the optimal strategy against b bits is poly-time for any fixed b.

THEOREM 5.1. *Given an m by n matrix M with $\{0,1\}$ entries and a fixed $b \ge 0$, computing the optimal strategy against b strongly leaking bits $(\mathbf{p}_b^*)$ is poly-time.*

TABLE 1
$M \in \{0,1\}^{9\times 14}$ of value 1/2, satisfying (i) $u_{\mathbf{p}_0^*}(M,1) = 0$ *and* (ii) $u_{\mathbf{p}_1^*}(M,1) \ge 1/7$.

$1\cdots 1$	$0\cdots 0$
$0\cdots 0$	$1\cdots 1$
T_7	T_7

Proof. An optimal strategy $\mathbf{p}_b^* = (p_1, \ldots, p_m)$ can be computed by solving the following linear program:

$$\begin{aligned}
&\text{maximize } z \text{ s.t.}\\
&\sum_{w\in\{0,1\}^b} \sum_{i:f(i)=w} p_i M_{i,g(w)} \geq z \quad \forall f : [m] \to \{0,1\}^b, \forall g : \{0,1\}^b \to [n],\\
&\sum_{i\in[m]} p_i = 1,\\
&p_i \geq 0 \quad \forall i \in [m].
\end{aligned}$$

Since there are 2^{bm} possible functions f and n^{2^b} possible functions g, this linear program admits a polynomial number of variables but an exponential number of constraints. However, a closer analysis shows that it can be rewritten with a polynomial number of constraints.

The composition of f and g is essentially a mapping $h : [m] \to [n]$ with image of cardinality at most 2^b. Fixing some subset $J \subseteq [n]$, $|J| \leq 2^b$, it is easy to compute a mapping h_J that minimizes the expected payoff of MAX over all mappings h with image J: h_J simply maps each row $i \in [m]$ to a column $j \in J$ that minimizes $M_{i,j}$. Therefore an optimal strategy $\mathbf{p}_b^*$ can be computed by solving the linear program

$$\begin{aligned}
&\text{maximize } z \text{ s.t.}\\
&\sum_{j\in J} \sum_{i:h_J(i)=j} p_i M_{i,j} \geq z \quad \forall J \subseteq [n], |J| \leq 2^b, \qquad (3)\\
&\sum_{i\in[m]} p_i = 1,\\
&p_i \geq 0 \quad \forall i \in [m],
\end{aligned}$$

whose size is polynomial as long as b is a constant. □

We next show that for general b, computing the optimal strategy against b bits is NP-hard. Moreover, we show that it is NP-hard to approximate the expected payoff guaranteed by MAX to within any factor.

THEOREM 5.2. *Given an m by n matrix M with $\{0,1\}$ entries, it is NP-hard to approximate $v^{strong}(M,b)$ by any factor under the strong leakage setting.*

Proof. We show that given an m by n matrix M with $\{0,1\}$ entries and some $b \geq 0$, it is NP-hard to decide whether $v^{\text{strong}}(M,b) > 0$. This is done by reduction from set cover (SC). An instance of SC is composed of a finite set of elements $U = \{1, \ldots, m\}$, a collection $\mathcal{C} = \{C_1, \ldots, C_r\}$ of subsets of U and an integer k. The question is whether there is a subcollection $\mathcal{C}' \subseteq \mathcal{C}$, $|\mathcal{C}'| \leq k$, such that every element in U belongs to at least one member of $\mathcal{C}'$.

Given an instance of SC, $\langle U, \mathcal{C}, k\rangle$, we construct the following instance of our problem. Let M be a binary matrix with $m = |U|$ rows and $n = r$ columns such that $M_{i,j} = 0 \Leftrightarrow i \in C_j$. Fix $b = \log k$. We show that there is a set cover of size at most k if and only if $v^{\text{strong}}(M,b) = 0$.

Sufficiency. Suppose the size of the set cover is greater than k. Then, we show that taking the uniform distribution over the whole action set (i.e., setting $p_i = \frac{1}{m}$ $\forall i \in [m]$) yields $v^{\text{strong}}(M,b) > 0$.

Consider inequality (3) and let $\mathbf{p}$ be the uniform distribution as described above. For every choice of $J \subseteq [n]$, the left-hand side of the inequality is composed of a finite set of summands. In order to show that the obtained payoff is greater than zero, it

is sufficient to show that at least one summand is greater than zero. Indeed, since the set cover is greater than $k = 2^b$, there must exist some row $i \in [m]$, call it i', such that $M_{i',j} = 1$ for every $j \in J$, and also $\mathbf{p}(i') > 0$ (since $\mathbf{p}$ has a full support). Consequently, $v^{\text{strong}}(M, b) > 0$.

Necessity. Suppose there exists a set cover of size at most $k = 2^b$. Then, there is a set of columns S, $|S| \leq 2^b$, such that for every $i \in [m]$ there exists $j \in |S|$ for which $M_{i,j} = 0$. Let g be a function that maps every $w \in \{0,1\}^b$ to a different column in S (arbitrarily). By the choice of S, it must hold that for every $i \in [m]$, $M_{i,g(f_g(i))} = 0$. Therefore, for every distribution vector $\mathbf{p}$, every summand in inequality (3) equals zero. Consequently, $v^{\text{strong}}(M, b) = 0$. □

In contrast to the last theorem, under the weak leakage setting, computing the optimal strategy for a given $f : [m] \to \{0,1\}^b$ is trivially poly-time: For every $w \in \{0,1\}^b$, let $S^w = \{i : f(i) = w\}$, and let M^w denote the submatrix $M^w \in \mathbb{R}^{|S^w| \times n}$ induced by S^w. MAX will choose w that maximizes $v(M^w)$ and play the corresponding maximin strategy.

MIN's optimal strategy. Finally, we consider the computational complexity of finding the optimal strategy (i.e., the optimal f function) for MIN under the weak leakage setting. For a general b, the exact same reduction from set cover, presented in the proof of Theorem 5.2, shows that computing the optimal f function is NP-hard and that it is NP-hard to find an f function that approximates the optimal expected payoff (for MIN) within any factor. We next show that given an integer matrix, computing the optimal f function (for MIN) is NP-hard for any number b of weakly leaking bits. The following theorem establishes the NP-hardness of this problem under a single bit. This proof can be easily extended to any b.

THEOREM 5.3. *Given an m by n integer matrix M and some real $0 < v < 1$, it is NP-hard to determine whether MIN can guarantee an expected payoff of at most v, even under a single weakly leaking bit.*

Proof. An expected payoff of at most v can be guaranteed by MIN if and only if there exists a pair of probability distributions $\mathbf{p^1}, \mathbf{p^2} \in \Delta(n)$ so that for every row M_i of M, the inner product of M_i with either $\mathbf{p^1}$ or $\mathbf{p^2}$ is at most v, i.e., either $M_i \cdot \mathbf{p^1} \leq v$ or $M_i \cdot \mathbf{p^2} \leq v$. (To see the correctness of the last statement, refer to the proof of Theorem 4.1.)

We show that given an m by n matrix M and an integer v, it is NP-hard to determine whether M admits a pair of probability distributions satisfying the last requirement. This is done by a reduction from balanced 3-uniform hypergraph 2-coloring (BH2C). An instance of BH2C is composed of a 3-uniform hypergraph $H = (V, E)$, where $V = \{u_1, \ldots, u_{2k}\}$ is the set of vertices and E is the set of edges. Every edge $e \in E$ is composed of a set of three vertices, denoted V_e. A 2-coloring function is denoted by $c : V \to \{1, 2\}$, mapping each vertex to either 1 or 2. The question is whether there exists a 2-coloring function c such that no edge of H is monochromatic and exactly k vertices are colored with each color.

It is well known that the problem of deciding whether a 3-uniform hypergraph G is 2-colorable is NP-complete (see [12]), and a moment's reflection shows that a vertex disjoint union of two copies of G has a balanced 2-coloring if and only if G has a 2-coloring.

Given an instance of BH2C, $H = (V, E)$, we construct a matrix M with $n = 2k+2$ columns and $m = 2|E| + 2k + 2$ rows. Columns $j = 1, \ldots, 2k$ correspond to vertices $u_1, \ldots, u_{2k}$, respectively. As for the rows, the first $2k$ rows correspond to the vertices and the next $2|E|$ rows correspond to the edges (two for each edge). The entries of M are given below.

For $i = 1, \ldots, 2k$, $M_{i,i} = 2$ and $M_{i,j} = 6k+2$ for $j \neq i$. Rows $i = 2k+1, \ldots, 2k+2|E|$ correspond to the edges, two for each edge. For both rows i that correspond to some edge e: for $1 \leq j \leq 2k$, $M_{i,j} = 2$ if $u_j \in V_e$ and $M_{i,j} = 6k+2$ if $u_j \in V \setminus V_e$. As for the last two columns, if i is the first row corresponding to the edge, then $M_{i,2k+1} = 6k+2$ and $M_{i,2k+2} = 12k$, and otherwise it is reversed; i.e., $M_{i,2k+1} = 12k$ and $M_{i,2k+2} = 6k+2$. Finally, in the last two rows, all the entries equal $12k$ except the following two entries: $M_{2k+2|E|+1,2k+1} = 3k$ and $M_{2k+2|E|+2,2k+2} = 3k$.

We show that H admits a balanced 2-coloring if and only if there exists a pair of probability distributions $\mathbf{p^1}, \mathbf{p^2}$ such that for every $i \in [m]$ either $M_i \cdot \mathbf{p^1} \leq 6k$ or $M_i \cdot \mathbf{p^2} \leq 6k$.

Necessity. Suppose H admits a balanced 2-coloring, and let V_1, V_2 denote the respective sets of vertices colored 1 and 2, i.e., $V_1 = \{u \in V : c(u) = 1\}$ and $V_2 = \{u \in V : c(u) = 2\}$. We construct the required pair of probability distributions $\mathbf{p^1}, \mathbf{p^2}$ as follows. $\mathbf{p^1}_j = 1/3k$ for every j such that $u_j \in V_1$, $\mathbf{p^1}_j = 0$ for every j such that $u_j \in V_2$ and for $j = 2k+2$, and $\mathbf{p^1}_{2k+1} = 2/3$. Similarly, $\mathbf{p^2}_j = 1/3k$ for every j such that $u_j \in V_2$, $\mathbf{p^2}_j = 0$ for every j such that $u_j \in V_1$ and for $j = 2k+1$, and $\mathbf{p^2}_{2k+2} = 2/3$. We show that for every row $i \in [m]$, either $M_i \cdot \mathbf{p^1} \leq 6k$ or $M_i \cdot \mathbf{p^2} \leq 6k$.

Consider first rows $i = 1, \ldots, 2k$ corresponding to the $2k$ vertices. If $u_i \in V_1$, then $M_i \cdot \mathbf{p^1} = 1/3k \cdot 2 + (1 - 1/3k) \cdot (6k+2) = 6k$. Otherwise, $u_i \in V_2$ and similarly $M_i \cdot \mathbf{p^2} = 6k$.

Consider next a row i that corresponds to edge e. Since every edge has at least one vertex of each color and $M_{i,j} = 2$ for every j such that $u_j \in V_e$, then for every $\ell \in \{1, 2\}$, there exists $1 \leq j \leq 2k$ such that $\mathbf{p^\ell}_j = 1/3k$ and $M_{i,j} = 2$. If i is odd, then $M_{i,2k+1} = 6k+2$ and it follows that $M_i \cdot \mathbf{p^1} \leq 1/3k \cdot 2 + (1 - 1/3k) \cdot (6k+2) = 6k$. Otherwise, $M_{i,2k+2} = 6k+2$, and we similarly have $M_i \cdot \mathbf{p^2} \leq 6k$.

Finally, for the last two rows, $M_{2k+2|E|+1} \cdot \mathbf{p^1} = 2/3 \cdot 3k + 1/3 \cdot 12k = 6k$, and similarly $M_{2k+2|E|+2} \cdot \mathbf{p^2} = 6k$. This establishes necessity.

Sufficiency. Suppose there exists a pair of probability distributions $\mathbf{p^1}, \mathbf{p^2} \in \Delta(n)$ such that for every $i \in [m]$ either $M_i \cdot \mathbf{p^1} \leq 6k$ or $M_i \cdot \mathbf{p^2} \leq 6k$. Then, we construct a balanced 2-coloring of H as follows.

Suppose without loss of generality that $\mathbf{p^1}$ satisfies $M_{2k+2|E|+1} \cdot \mathbf{p^1} \leq 6k$. Simple arithmetic reveals that $\mathbf{p^1}_{2k+1} \geq 2/3$. Similar reasoning for row $i = 2k + 2|E| + 2$ implies that for some $\ell \in \{1, 2\}$ it must hold that $\mathbf{p^\ell}_{2k+2} \geq 2/3$, but since $\mathbf{p^1}_{2k+1} \geq 2/3$, it follows that $\mathbf{p^2}_{2k+2} \geq 2/3$.

Consider next a row $i \in \{1, \ldots, 2k\}$. Since there exists an $\ell \in \{1, 2\}$ such that $M_i \cdot \mathbf{p^\ell} \leq 6k$, $\mathbf{p^\ell}_i$ should satisfy $\mathbf{p^\ell}_i \cdot 2 + (1 - \mathbf{p^\ell}_i) \cdot (6k+2) \leq 6k$, implying that for every $i \in \{1, \ldots, 2k\}$, there exists an $\ell \in \{1, 2\}$ such that $\mathbf{p^\ell}_i \geq 1/3k$. Combining this with the inequalities $\mathbf{p^1}_{2k+1} \geq 2/3$ and $\mathbf{p^2}_{2k+2} \geq 2/3$ we have that $\mathbf{p^1}$ and $\mathbf{p^2}$ essentially partition the indices $j = 1, \ldots, 2k$ into two disjoint equal-size sets X_1, X_2 such that $\mathbf{p^1}_j = 1/3k$ for every $j \in X_1$ and $\mathbf{p^2}_j = 1/3k$ for every $j \in X_2$. In addition, $\mathbf{p^1}_{2k+1} = 2/3$, $\mathbf{p^2}_{2k+2} = 2/3$, and all other entries are zero.

Consider the following coloring. For $i = 1, \ldots, 2k$, if $i \in X_1$, then $c(u_i) = 1$; else (i.e., if $i \in X_2$), $c(u_i) = 2$. c is balanced since $|X_1| = |X_2| = k$. It remains to show that none of the edges is monochromatic. We show that for every edge e, there exists $u \in V_e$ such that $c(u) = 1$. One can analogously show that for every edge e, there exists $u \in V_e$ such that $c(u) = 2$.

Suppose toward contradiction that there exists an edge e such that $c(u_j) = 2$ for all j such that $u_j \in V_e$. This implies that $\mathbf{p^2}_j = 1/3k$ for all such j. Consider the odd row i corresponding to edge e (i.e., row i in which $M_{i,2k+1} = 6k+2$ and $M_{i,2k+2} = 12k$). It holds that $M_i \cdot \mathbf{p^2} \geq 2/3 \cdot 12k = 8k > 6k$ and $M_i \cdot \mathbf{p^1} = 6k+2 > 6k$. That is, the

inner product of row M_i with both $\mathbf{p^1}$ and $\mathbf{p^2}$ is greater than $6k$, in contradiction. It follows that every edge has at least one vertex colored in 1. This concludes the proof. □

6. Bidirectional leakage. In this section we show that under the bidirectional leakage model, if b is a large constant, then both players can guarantee a payoff very close to that of the original game.

THEOREM 6.1. *For every small $\epsilon > 0$, there exists some $b_0 = b_0(\epsilon)$ so that the following hold for every m, n, and $M \in \{0,1\}^{m \times n}$:*

(1) *There exists a strategy of MAX consisting of a function $f_{max} : \mathcal{R}_{min} \to \{0,1\}^b$ and a function $g_{max} : \mathcal{R}_{max} \times \{0,1\}^b \to [m]$ so that for every strategy of MIN consisting of a function $f_{min} : \mathcal{R}_{max} \to \{0,1\}^b$ and a function $g_{min} : \mathcal{R}_{min} \times \{0,1\}^b \to [n]$, the expected payoff of the game played on M with $b \geq b_0$ bidirectional leaking bits is at least $v(M) - \epsilon$.*

(2) *There exists a strategy of MIN consisting of a function $f_{min} : \mathcal{R}_{max} \to \{0,1\}^b$ and a function $g_{min} : \mathcal{R}_{min} \times \{0,1\}^b \to [n]$ so that for every strategy of MAX consisting of a function $f_{max} : \mathcal{R}_{min} \to \{0,1\}^b$ and a function $g_{max} : \mathcal{R}_{max} \times \{0,1\}^b \to [m]$, the expected payoff of the game played on M with $b \geq b_0$ bidirectional leaking bits is at most $v(M) + \epsilon$.*

We will soon prove Theorem 6.1, but first let us introduce the following notation. Given some probability distribution $\mathbf{p} = (p_1, \ldots, p_n)$, let $\mathrm{H}(\mathbf{p})$ denote the *entropy* of $\mathbf{p}$, defined as $\mathrm{H}(\mathbf{p}) = -\sum_{i=1}^n p_i \log(p_i)$. Given two probability distributions $\mathbf{p} = (p_1, \ldots, p_n)$ and $\mathbf{p'} = (p'_1, \ldots, p'_n)$, let $\mathrm{d}(\mathbf{p}, \mathbf{p'})$ denote the *variation distance* between $\mathbf{p}$ and $\mathbf{p'}$, defined as $\sum_{i=1}^n |p_i - p'_i|$. Let $\mathbf{u} = (1/n, \ldots, 1/n)$ denote the uniform distribution on $[n]$.

LEMMA 6.2. *There exists an absolute constant $c > 0$ so that if $\mathbf{p}$ is a probability distribution on $[n]$ and $\mathrm{d}(\mathbf{p}, \mathbf{u}) = \epsilon$ for some $\epsilon < 1/5$, then $\mathrm{H}(\mathbf{p}) \leq \log n - c\epsilon^2$.*

Proof. For convenience, we will prove that $\mathrm{H}_e(\mathbf{p}) \leq \ln n - c\epsilon^2$ for some constant $c > 0$, where e is Euler's number, $\ln(\cdot) = \log_e(\cdot)$ is the natural base logarithm, and $\mathrm{H}_e(\mathbf{p}) = -\sum_{i=1}^n p_i \ln(p_i)$ is the entropy of $\mathbf{p}$ measured in units of e. The assertion follows as $\mathrm{H}(\mathbf{p}) = \mathrm{H}_e(\mathbf{p}) \cdot \log(e)$.

By compactness, there is a probability distribution $\mathbf{p} = (p_1, \ldots, p_n)$ satisfying $d(\mathbf{p}, \mathbf{u}) = \epsilon$ for which $\mathrm{H}_e(\mathbf{p})$ is maximum. It is easy to see that $\mathbf{p}$ cannot contain two distinct components $p_i \neq p_j$ so that $p_i, p_j \leq 1/n$ since replacing both by their average will increase the value of $\mathrm{H}_e(\mathbf{p})$ without changing $\mathrm{d}(\mathbf{p}, \mathbf{u})$. Similarly, there cannot be distinct $p_i \neq p_j$ so that $p_i, p_j > 1/n$. It thus follows that there exists an integer k so that $\mathbf{p}$ has k coordinates whose value is $1/n + \frac{\epsilon}{2k}$ and $n - k$ coordinates, each equal to $1/n - \frac{\epsilon}{2(n-k)}$. Notice that k must satisfy $\frac{\epsilon}{2(n-k)} < 1/n$.

The proof continues by considering two possible cases.

Case 1. $\frac{\epsilon}{2k} \leq \frac{5}{n}$.

We argue that for any real x satisfying $-1/n < x \leq \frac{5}{n}$,

$$-\left(\frac{1}{n} + x\right) \ln\left(\frac{1}{n} + x\right) \leq \frac{\ln n}{n} + (\ln n - 1)x - \frac{1}{12}x^2 n. \tag{4}$$

To establish this argument, define

$$g(x) = \frac{\ln n}{n} + (\ln n - 1)x - \frac{1}{12}x^2 n + \left(\frac{1}{n} + x\right) \ln\left(\frac{1}{n} + x\right).$$

Then $g'(x) = \ln n - 1 - \frac{xn}{6} + \ln(\frac{1}{n} + x) + 1$ and $g''(x) = \frac{-n}{6} + \frac{1}{(1/n)+x}$. Since $g(0) = 0$,

$g'(0) = 0$, and $g''(x) \geq 0$ for all $-1/n < x \leq \frac{5}{n}$, it follows that $g(x) \geq 0$ for all $-1/n < x \leq \frac{5}{n}$ which establishes our argument.

Plugging (4) with $x = \frac{\epsilon}{2(n-k)}$ and with $x = \frac{\epsilon}{2k}$ in the expression for $\mathrm{H}_e(\mathbf{p})$, we conclude that

$$\begin{aligned}\mathrm{H}_e(\mathbf{p}) &= -k\left(\frac{1}{n}+\frac{\epsilon}{2k}\right)\ln\left(\frac{1}{n}+\frac{\epsilon}{2k}\right)\\ &\quad -(n-k)\left(\frac{1}{n}-\frac{\epsilon}{2(n-k)}\right)\ln\left(\frac{1}{n}-\frac{\epsilon}{2(n-k)}\right)\\ &\leq \ln n - k\frac{1}{12}\cdot\frac{\epsilon^2}{4k^2}n-(n-k)\frac{1}{12}\cdot\frac{\epsilon^2}{4(n-k)^2}n\\ &= \ln n - \frac{1}{48}\epsilon^2\left(\frac{n}{k}+\frac{n}{n-k}\right).\end{aligned}$$

The assertion follows since $\frac{n}{k}+\frac{n}{n-k} \geq 4$.

Case 2. $\frac{\epsilon}{2k} > \frac{5}{n}$.

In this case

$$\begin{aligned}\mathrm{H}_e(\mathbf{p}) &= -k\left(\frac{1}{n}+\frac{\epsilon}{2k}\right)\ln\left(\frac{1}{n}+\frac{\epsilon}{2k}\right)\\ &\quad -(n-k)\left(\frac{1}{n}-\frac{\epsilon}{2(n-k)}\right)\ln\left(\frac{1}{n}-\frac{\epsilon}{2(n-k)}\right)\\ &= -k\left(\frac{1}{n}+\frac{\epsilon}{2k}\right)\left[\ln\left(1+\frac{\epsilon n}{2k}\right)-\ln(n)\right]\\ &\quad -(n-k)\left(\frac{1}{n}-\frac{\epsilon}{2(n-k)}\right)\left[\ln\left(1-\frac{\epsilon n}{2(n-k)}\right)-\ln(n)\right]\\ &= \ln(n) - k\left(\frac{1}{n}+\frac{\epsilon}{2k}\right)\ln\left(1+\frac{\epsilon n}{2k}\right)\\ &\quad -(n-k)\left(\frac{1}{n}-\frac{\epsilon}{2(n-k)}\right)\ln\left(1-\frac{\epsilon n}{2(n-k)}\right).\end{aligned}$$

Fix

$$h = k\left(\frac{1}{n}+\frac{\epsilon}{2k}\right)\ln\left(1+\frac{\epsilon n}{2k}\right)+(n-k)\left(\frac{1}{n}-\frac{\epsilon}{2(n-k)}\right)\ln\left(1-\frac{\epsilon n}{2(n-k)}\right),$$

so it remains to show that $h \geq c\epsilon^2$ for some constant c. Since $\frac{\epsilon n}{2k} \geq 5$, it follows that

$$\begin{aligned}h &\geq \frac{2k+\epsilon n}{2n}\ln(6)+\frac{2(n-k)-\epsilon n}{2n}\ln\left(1-\frac{\epsilon n}{2(n-k)}\right)\\ &\geq \left(\frac{k}{n}+\frac{\epsilon}{2}\right)\ln(6)+\ln\left(1-\frac{\epsilon n}{2(n-k)}\right).\end{aligned}$$

Using the fact that $\ln(6) > 3/2$ and the fact that for $y \leq 1/2$, $\ln(1-y) \geq -y-y^2$, we conclude that

$$h \geq \frac{3}{2}\left(\frac{k}{n}+\frac{\epsilon}{2}\right)-\frac{\epsilon n}{2(n-k)}-\left(\frac{\epsilon n}{2(n-k)}\right)^2. \tag{5}$$

Since $\frac{\epsilon}{2k} > \frac{5}{n}$ and since $\epsilon \leq 1/5$, it follows that $k < n/2$. Therefore, if $k \geq n/5$, then (5) implies that

$$h \geq \frac{3}{2}\left(\frac{1}{5}+\frac{\epsilon}{2}\right) - \frac{\epsilon n}{2(n-n/2)} - \left(\frac{\epsilon n}{2(n-n/2)}\right)^2 \geq \frac{3}{2}\epsilon + \frac{3}{4}\epsilon - \epsilon - \epsilon^2 = \frac{5}{4}\epsilon - \epsilon^2.$$

On the other hand, if $k < n/5$, then (5) implies that

$$h > \frac{3}{4}\epsilon - \frac{\epsilon n}{2(n-n/5)} - \left(\frac{\epsilon n}{2(n-n/5)}\right)^2 = \frac{1}{8}\epsilon - \frac{25}{64}\epsilon^2 \geq \frac{5}{8}\epsilon^2 - \frac{25}{64}\epsilon^2 = \frac{15}{64}\epsilon^2.$$

The assertion follows. □

Remark. For small $\epsilon > 0$ the best possible value of the constant c promised by Lemma 6.2 is $1/2$, but we make no attempt to optimize it here.

Lemma 6.2 is the key ingredient in the proof of the following lemma (which is a simple, though not very efficient, construction of a strong extractor).

LEMMA 6.3. *Let b be a positive integer and let $X = (X_1, \ldots, X_{2^b})$ be a random uniform bit string of length $k \cdot 2^b$, consisting of 2^b blocks X_i, each being a random uniform bit string of length k. Let $f : \{0,1\}^{k \cdot 2^b} \to \{0,1\}^b$ be an arbitrary function. If i is chosen uniformly at random from $[2^b]$, then the expected (over the random choice of* i*) variation distance between X_i given the value of $f(X)$ and a uniform distribution on $\{0,1\}^k$ is at most $O(\sqrt{b/2^b})$.*

Proof. For $i = 1, \ldots, 2^b$, let $\mathbf{p}^i$ denote the probability distribution of the (random) block X_i given the value of $f(X)$. Since the conditional entropy $\mathrm{H}(X|f(X))$ satisfies $\mathrm{H}(X|f(X)) = \mathrm{H}(X, f(X)) - \mathrm{H}(f(X)) = \mathrm{H}(X) - \mathrm{H}(f(X)) \geq k \cdot 2^b - b$ and since the subadditivity of the entropy function implies that $\mathrm{H}(X|f(X)) \leq \sum_{i=1}^{2^b} \mathrm{H}(X_i|f(X)) = \sum_{i=1}^{2^b} \mathrm{H}(\mathbf{p}^i)$, it follows that

$$\frac{1}{2^b}\sum_{i=1}^{2^b} \mathrm{H}(\mathbf{p}^i) \geq k - \frac{b}{2^b}.$$

Lemma 6.2 guarantees that

$$\frac{1}{2^b}\sum_{i=1}^{2^b} \mathrm{d}(\mathbf{p}^i, \mathbf{u}) \leq \frac{1}{2^b}\sum_{i=1}^{2^b} O\left(\sqrt{k - \mathrm{H}(\mathbf{p}^i)}\right),$$

hence the concavity of $\sqrt{\cdot}$ implies that

$$\frac{1}{2^b}\sum_{i=1}^{2^b} \mathrm{d}(\mathbf{p}^i, \mathbf{u}) \leq O\left(\sqrt{\frac{1}{2^b}\sum_{i=1}^{2^b}(k - \mathrm{H}(\mathbf{p}^i))}\right) = O\left(\sqrt{b/2^b}\right).$$

The assertion follows. □

We are now ready to establish the main theorem of this section.

Proof of Theorem 6.1. Given some $\epsilon > 0$, take $b_0 = b_0(\epsilon)$ to be sufficiently large so that $c\sqrt{b_0/2^{b_0}} \leq \epsilon$, where c is the hidden constant in the O-notation in Lemma 6.3. Let m and n be some positive integers and let M be an arbitrary binary $m \times n$ matrix. Consider the game defined by M with $b \geq b_0$ bidirectional leaking bits. We will construct a strategy for MAX consisting of a function $f_{\max} : \mathcal{R}_{\min} \to \{0,1\}^b$ and a function $g_{\max} : \mathcal{R}_{\max} \times \{0,1\}^b \to [m]$ so that for every strategy of MIN, the expected payoff of the game is at least $v(M) - \epsilon$. The construction of such a strategy for MIN

that guarantees an expected payoff of at most $v(M)+\epsilon$ for every strategy of MAX is analogous.

Let s be an optimal (mixed) strategy of MAX in the two-player zero-sum game defined by M (with no leakage). For simplicity, we assume that s can be implemented with k random bits; in particular, it will be convenient to think of s as a function $s : \{0,1\}^k \to [m]$ that maps a random bit string of length k to some action (row) in $[m]$.

Recall that the elements $r_{\max}$ and $r_{\min}$, randomly picked by nature from $\mathcal{R}_{\max}$ and $\mathcal{R}_{\min}$, respectively, are assumed to be infinite length bit strings. In particular, our proof relies on $r_{\max}$ being of length at least $k \cdot 2^b$ and on $r_{\min}$ being of length at least b. The latter assumption naturally raises the following question: Is it valid on behalf of MAX (and on our behalf) to assume that MIN uses at least b random bits? We believe that in a setting involving the leakage of b bits from the players' random sources, one can safely assume that the players employ significantly more than b random bits. Indeed, if MIN restricts herself to using b (or less) random bits, then when she plays first, she cannot expect the payoff to be better than the game's *deterministic* minimax, which is typically much worse than the game's value $v(M)$. Of course, this question becomes redundant once we recall the assumption that both $r_{\max}$ and $r_{\min}$ contain infinite many random bits, regardless of how many random bits the players choose to employ.

The mixed strategy of MAX is defined as follows. Take $f_{\max}$ to be the function that returns the first b bits in $r_{\min}$ and let $0 \le \lambda \le 2^b - 1$ be the integer interpretation of $f_{\max}(r_{\min})$. For the construction of $g_{\max}$, we partition the first $k \cdot 2^b$ bits in $r_{\max}$ into 2^b *blocks*, each containing k consecutive bits, denoted

$$B_\ell = (r_{\max}[\ell k + 1], \ldots, r_{\max}[\ell k + k]), \qquad \ell = 0, 1, \ldots, 2^b - 1.$$

MAX then employs the random bits in block B_λ to implement the mixed strategy s, namely,

$$g_{\max}(r_{\max}, f_{\max}(r_{\min})) = s(B_\lambda).$$

Informally speaking, MIN knows that MAX realizes s based on some block B_λ and she could have demolished the strategy of MAX if she would have known which block B_λ is used. However, the index λ is determined by $r_{\min}$ which, by the definition of our model, is revealed to MIN only after she already committed to her leakage function $f_{\min}$.[7] Hence, MIN can obtain (via $f_{\min}$) very limited information regarding the block B_λ and this limited information does not suffice to harm the payoff of MAX by more than ϵ.

More formally, consider an arbitrary strategy of MIN consisting of a function $f_{\min} : \mathcal{R}_{\max} \to \{0,1\}^b$ and a function $g_{\min} : \mathcal{R}_{\min} \times \{0,1\}^b \to [n]$. Our goal is to show that if MAX plays row $s(B_\lambda) = g_{\max}(r_{\max}, f_{\max}(r_{\min})) \in [m]$ and MIN plays column $g_{\min}(r_{\min}, f_{\min}(r_{\max})) \in [n]$, then the expected payoff of the game is at least $v(M) - \epsilon$.

To that end, let $\mathbf{p}^\lambda$ be the probability distribution of B_λ given the value of $f_{\min}(r_{\max})$. Note that in the language of the aforementioned informal discussion, $\mathbf{p}^\lambda$ is what MIN "sees" as the probability distribution MAX uses to realize s. The key observation here is that by the choice of $b_0 = b_0(\epsilon)$, Lemma 6.3 guarantees that the expected variation distance between $\mathbf{p}^\lambda$ and the uniform distribution $\mathbf{u}_{2^k}$ on the set of all length k bit strings is at most ϵ. Therefore, in expectation, the payoff of the

[7] This feature of our bidirectional leakage model can be thought of as if the players' strategies are mixed, but the actual choice of the leakage functions is assumed to be pure.

game cannot differ from $v(M)$, which is obtained by MAX if she uses $\mathbf{u}_{2^k}$ to realize s, by more than ϵ. The assertion follows. □

Appendix A. Explicit constructions. In this appendix we describe several explicit constructions of n by n $\{0,1\}$-matrices representing games with value q, such that if the MIN player has b bits and b is smaller than $\log\log n + O(1)$, then the MAX player can guarantee a payoff of at least roughly q^{2^b}. This shows (by explicit examples) that the statements of Theorem 3.3 and Corollary 3.5 are essentially tight.

Example 1. Let p be a prime power and let r be a positive integer. Fix $n = p^r - 1$. Let $M = (M_{u,v})$ be the following n by n binary matrix whose rows and columns are indexed by the set N of all nonzero vectors of length r over $GF(p)$. For each such u, v, $M_{u,v} = 1$ if and only if the two vectors u and v are orthogonal over $GF(p)$ (namely, their inner product over $GF(p)$ is zero). Note that M is a symmetric matrix, and every row and every column of it contains exactly $p^{r-1} - 1$ 1-entries. Indeed, this is the number of nonzero solutions of a single linear equation in r variables over $GF(p)$. It is easy to check that the maximin strategy of the game determined by M is the uniform distribution over N, yielding a value of $q = \frac{p^{r-1}-1}{p^r-1}$. Note that for large $n = p^r - 1$ this is very close to $1/p$.

We claim that for every set $J \subseteq N$ of at most $\log_p n$ columns, there are at least $p^{r-|J|} - 1$ rows u so that $M_{u,v} = 1$ for every $v \in J$. Note that if $p^{r-|J|}$ is large, then this number is very close to $q^{|J|}n$, implying that by playing the uniform distribution on the rows of M, MAX can ensure a value close to $q^{|J|}$. Note also that if $b \leq \log r - O(1) = \log\log n - O_p(1)$, then 2^b is much smaller than r, and hence $p^{r-|J|}$ is large provided $|J| \leq 2^b$. Fix a subset $J \subseteq N$ of cardinality at most $\log_p n$. By definition, row u satisfies $M_{u,v} = 1$ for every $v \in J$ if and only if the inner product of u and v over $GF(p)$ is zero for every $v \in J$. This is a homogeneous system of $|J|$ linear equations in the r variables representing the coordinates of u. This system clearly admits at least $p^{r-|J|} - 1$ nontrivial solutions; each such nontrivial solution corresponds to a row with the desired properties, proving the claim.

This completes the description of the first set of examples. Note that it works for every value q which is about $1/p$, where p is a prime power.

Example 2 (sketch). Let p be a prime and let M be a $p \times p$ binary matrix, where $M_{i,j} = 1$ if and only if $i - j$ is a quadratic residue modulo p (where here zero is considered a quadratic residue). The value of the game represented by M is $(p+1)/(2p)$, which, for large p, is roughly $1/2$. Using Weil's theorem, it is not difficult to show that for every subset S of Z_p of size at most $(0.5 - \delta)\log p$, the number of rows i such that $M_{i,j} = 1$ for all $j \in S$ is $(1 + o(1))\frac{p}{2^{|S|}}$. A similar example holds for characters of other orders instead of the quadratic character, providing examples with values close to $1/d$ for any desired positive integer $d > 1$ (where here we have to choose a prime p so that d divides $p - 1$—by Dirichlet's theorem on primes in arithmetic progressions it is known that there are infinitely many such primes for any such d).

REFERENCES

[1] N. Alon, *Tools from higher algebra*, in Handbook of Combinatorics, R.L. Graham, M. Grötschel, and L. Lovász, eds, North-Holland, Amsterdam, 1995, pp. 1749–1783.

[2] N. Alon, O. Goldreich, J. Håstad, and R. Peralta, *Simple constructions of almost k-wise independent random variables*, in Proceedings of the 31st IEEE FOCS, St. Louis, MO, IEEE, 1990, pp. 544–553.

[3] N. Alon, G. Gutin, and M. Krivelevich, *Algorithms with large domination ratio*, J. Algorithms, 50 (2004), pp. 118–131.
[4] N. Alon and J.H. Spencer, *The Probabilistic Method*, 3rd ed., John Wiley, New York, 2008.
[5] R.J. Aumann, *On the non-transferable utility value: A comment on the Roth-Shaper examples*, Econometrica, 53 (1985), pp. 667–677.
[6] R.J. Aumann and M. Maschler, *Some thoughts on the minimax principle*, J. Management Sci., 18 (1972), pp. 54–63.
[7] M. Budinich and L. Fortnow, *Repeated matching pennies with limited randomness*, in Proceedings of the 12th ACM Conference on Electronic Commerce, 2011, pp. 111–118.
[8] S. Dziembowski and K. Pietrzak, *Leakage-resilient cryptography*, in Proceedings of the 49th IEEE Symposium on Foundations of Computer Science, 2008, pp. 293–302.
[9] R. Gradwohl and O. Reingold, *Partial exposure in large games*, Games Econom. Behav., 68 (2010), pp. 602–613.
[10] J.Y. Halpern and R. Pass, *Algorithmic rationality: Adding cost of computation to game theory*, ACM SIGecom Exchanges, 10 (2011), pp. 9–15.
[11] D.S. Johnson, *Approximation algorithms for combinatorial problems*, J. Comput. System Sci., 9 (1974), pp. 256–278.
[12] L. Lovász, *Coverings and coloring of hypergraphs*, in Proceedings of the Fourth Southeastern Conference on Combinatorics, Graph Theory, and Computing, Florida Atlantic University, Boca Raton, FL, Utilitas Mathematica, Winnipeg, MB, Canada, 1973, pp. 3–12.
[13] L. Lovász, *On the ratio of optimal and fractional covers*, Discrete Math., 13 (1975), pp. 383–390.
[14] A. Matsui, *Information Leakage Forces Cooperation*, Discussion paper, Northwestern University, 1988.
[15] J. Naor and M. Naor, *Small-bias probability spaces: Efficient constructions and applications*, in Proceedings of the 22nd Annual ACM STOC, ACM Press, 1990, pp. 213–223.
[16] H. Nasheri, *Economic Espionage and Industrial Spying*, Cambridge University Press, Cambridge, UK, 2005.
[17] J.S. Provan, *The Use of Spies in Strategic Situations: Preliminary Report*, UNC/STOR/07/01, University of North Carolina at Chapel Hill, NC, 2008.
[18] A. Rubinstein, *Finite automata play the repeated prisoner's dilemma*, J. Econom. Theory, 39 (1986), pp. 83–96.
[19] E. Solan and L.Yariv, *Games with espionage*, Games Econom. Behav., 47 (2004), pp. 172–199.
[20] M. Tennenholtz, *Competitive safety analysis: Robust decision-making in multi-agent systems*, J. Artificial Intelligence Res., 17 (2002), pp. 363–378.
[21] H.A. Simon, *A behavioral model of rational choice*, Quart. J. Econom., 49 (1955), pp. 99–118.
[22] M. Tennenholtz, *Program equilibrium*, Games Econom. Behav., 49 (2004), pp. 363–373.
[23] *The Eudaemons*, http://physics.ucsc.edu/people/eudaemons/eudaemons.html.
[24] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, 1944.
[25] A. Weil, *Sur les courbes algebriques et les varietes qui s'en deduisent*, Actualités Sci. Indust. 1041, Herman, Paris, 1948.