A Note on the Probability of Rectangles for Correlated Binary Strings

Or Ordentlich\textsuperscript{◊}, Member, IEEE, Yury Polyanskiy, Senior Member, IEEE, and Ofer Shayevitz\textsuperscript{◦}, Senior Member, IEEE

Abstract—Consider two sequences of $n$ independent and identically distributed fair coin tosses, $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$, which are $\rho$-correlated for each $j$, i.e., $\mathbb{P}[X_{j} = Y_{j}] = \frac{1+\rho}{2}$. We study the question of how large (small) the probability $\mathbb{P}[X \in A, Y \in B]$ can be among all sets $A,B \subset \{0, 1\}^n$ of a given cardinality. For sets $|A|,|B| = \Theta(2^n)$ it is well known that the largest (smallest) probability is approximately attained by concentric (anti-concentric) Hamming balls, and this can be proved via the hypercontractive inequality (reverse hypercontractivity). Here we consider the case of $|A|,|B| = 2^{\Theta(n)}$. By applying a recent extension of the hypercontractive inequality of Polyanskiy-Samorodnitsky (J. Functional Analysis, 2019), we show that Hamming balls of the same size approximately maximize $\mathbb{P}[X \in A, Y \in B]$ in the regime of $\rho \to 1$. We also prove a similar tight lower bound, i.e., show that for $\rho \to 0$ the pair of opposite Hamming balls approximately minimizes the probability $\mathbb{P}[X \in A, Y \in B]$.

Index Terms—Isoperimetric inequalities, hypercontractivity, binary adder multiple access channel (MAC).

I. INTRODUCTION

Let $X \sim \text{Uniform}(\{0,1\}^n)$ and $Y \in \{0,1\}^n$ be a $\rho$-correlated copy of $X$, where $0 \leq \rho < 1$, i.e.,

$$\text{Pr}(Y = y | X = x) = \prod_{i=1}^{n} \left( \frac{1-\rho}{2} \right)^{d(x_i,y_i)} \left( \frac{1+\rho}{2} \right)^{1-d(x_i,y_i)} = \left( \frac{1+\rho}{2} \right)^n \cdot \left( \frac{1-\rho}{1+\rho} \right)^{d(x,y)} ,$$

(1)

where $d(x_i,y_i) = \mathbb{I}_{x_i \neq y_i}$ and $d(x,y) = \sum_{i=1}^{n} d(x_i,y_i)$. For $A,B \subset \{0,1\}^n$, we denote $P_{XY}(A \times B) = \text{Pr}(X \in A, Y \in B)$ – probability of a rectangle with sides $A$ and $B$. In this article we are interested in the following question: Among all sets of a given size, how large/small can the probability of a rectangle be? Previous works addressing similar questions relied on hypercontractive and reverse hypercontractive inequalities, as we describe below. Our main innovation is applying a new tool from [1] that is a refinement of the direct hypercontractive inequality to functions with sparse support.

A direct application of the hypercontractive inequality [2]–[6] (see Section III for more details) yields that for $A$ and $B$ of equal cardinalities, i.e. $|A| = |B| \triangleq \eta \cdot 2^n$, we have

$$P_{XY}(A \times B) \leq \eta^{\frac{2}{1+\rho} \rho^n},$$

(2)

whereas the reverse hypercontractive inequality of [7] was applied in [8] to obtain

$$P_{XY}(A \times B) \geq \eta^{\frac{2}{1+\rho} \rho^n}.$$

(3)

Both bounds become quite tight for the regime of $\eta = \Theta(1)$, i.e. for very large sets of cardinalities $|A| = |B| = \Theta(2^n)$. In particular, (2) is approximately attained by taking $A$ and $B$ as the zero-centered Hamming balls containing all vectors with Hamming weight smaller than $\frac{s}{2} - \sqrt{s/\eta}$, for large $s$ independent of $n$, whereas (3) is approximately attained by taking $A$ as such zero-centered ball and $B$ as the same ball shifted such that its center is the all-ones vector. A special case of the construction in [9] also gives more constructions of sets approximately attaining (2); namely, for any $k \in \mathbb{Z}_+$ and all sufficiently large $n \geq n_0(k)$ they constructed sets $A = B$ of cardinality $2^{n-k}$ such that

$$P_{XY}(A \times B) \geq \Omega_{\rho}(1/\sqrt{k}) 2^{-k \frac{1+\rho}{2}},$$

(4)

thus showing that the estimate (2) is tight (up to a polylog factor $(\log \frac{\rho}{2})^{-\frac{1}{2}}$).

In this article we are interested in estimating the probability of rectangles for sets $A, B$ of much smaller cardinalities (such as those frequently encountered in information and coding theories), namely $|A| = 2^{n\alpha}, |B| = 2^{n\beta}$ for $\alpha, \beta < 1$. Our original motivation stems from the bounds on the adder multiple access channel (MAC) zero-error capacity, obtained in [10]. Sets $A, B \subset \{0,1\}^n$ are called a zero-error code for the adder MAC, if $|A + B| = |A| \cdot |B|$, where $A + B \subset \{0,1\}^n$ is the Minkowski sum (over the reals) of the sets $A$ and $B$. The problem of finding all pairs $(R_1, R_2) \subset [0,1]^2$ for which there exist a zero-error code with sizes $|A| = 2^{nR_1}$,
\[ |B| = 2^{nR_2} \] is a long standing open problem [11]-[18]. One of the first results in the area, due to van Tilborg [12], states that if \( A, B \) form a zero error code, then
\[
W_d(A, B) \triangleq \frac{1}{n} \log \left| \{ (a, b) \in A \times B : d(a, b) = nd \} \right| \quad (5)
\]
\[
\leq \frac{1}{n} \log \left( \frac{n}{nd} \right) + \min(d, 1 - d),
\]
for all \( d \in \{0, \frac{1}{n}, \ldots, 1\} \). The basic idea in [10] was to use (6) for upper bounding
\[
P_{XY}(A \times B)
\]
\[
= 2^{-n \left( \frac{1 + \rho}{2} \right) n} \sum_{d=0}^{1} 2^{nW_d(A, B)} \left( \frac{1 - \rho}{1 + \rho} \right)^{nd}
\]
for any zero-error code \( (A, B) \), and to contrast this with lower bounds on \( P_{XY}(A \times B) \) for sets \( |A| = 2^{nR_1}, |B| = 2^{nR_2} \) obtained in [8] (see Remark 4 below). A simple modification of this approach [10] yielded the best known outer bounds on \( (R_1, R_2) \) to date, and possibly replacing the lower bound from [8] on \( P_{XY}(A \times B) \) with a sharper one, could yield stronger bounds on \( (R_1, R_2) \). For instance, if our main conjecture, stated below, turns out to be true, repeating the arguments in [10] with the improved bounds will yield that as \( R_1 \) approaches 1 we must have that \( R_2 < 0.4177 \), improving upon \( R_2 < 0.4228 \) established in [10], which is the best known bound to date.

Our interest is in the greatest and smallest exponential decay rate of \( P_{XY}(A \times B) \) among all possible sets \( A, B \) of sizes \( 2^{n\alpha} \) and \( 2^{n\beta} \), respectively. To that end, for fixed \( 0 < \alpha, \beta < 1 \) we define
\[
\overline{E}(\alpha, \beta, \rho) \triangleq -\limsup_{n \to \infty} \frac{1}{n} \log P_{XY}(A \times B),
\]
\[
\underline{E}(\alpha, \beta, \rho) \triangleq -\liminf_{n \to \infty} \frac{1}{n} \log P_{XY}(A \times B),
\]
where \( \max\{A\}, \{B\} \) and \( \min\{A\}, \{B\} \) denote optimizations over the sequences of sets \( A_n \subset \{0, 1\}^n, B_n \subset \{0, 1\}^n, n \in \mathbb{Z}_+ \) such that
\[
|A_n| = 2^{n\alpha + o(n)}, \quad |B_n| = 2^{n\beta + o(n)}.
\]

Our main conjecture is that both \( \overline{E}(\alpha, \beta, \rho) \) and \( \underline{E}(\alpha, \beta, \rho) \) are optimized by concentric (resp., anti-concentric) Hamming balls. In this work we show partial progress towards establishing this conjecture. Our conjecture is in line with the well-known facts that among all pairs of sets \( A, B \subset \{0, 1\}^n \) of given sizes, the maximal distance \( d_{\max}(A, B) = \max_{a \in A, b \in B} d(a, b) \) is minimized by concentric Hamming (quasi) balls [19], [20], whereas the minimum distance \( d_{\min}(A, B) = \min_{a \in A, b \in B} d(a, b) \) is maximized by anti-concentric Hamming (quasi) balls [21].

Notation: Logarithms are taken to base 2 throughout, unless stated otherwise. We denote the Shannon entropy of a random variable \( V \) by \( H(V) \). For a binary random variable \( V \sim \text{Ber}(p) \) we denote the entropy by \( h(p) \triangleq -p \log p - (1 - p) \log (1 - p) \). For all \( p \leq q \leq 1 \) we denote \( p \ast q \triangleq p(1 - q) + q(1 - p) \).

Our main results characterize \( \overline{E}(\alpha, \alpha, \rho) \) in the low noise (large \( \rho \)) regime, and \( \overline{E}(\alpha, \beta, \rho) \) in the high noise (small \( \rho \)) regime, as follows.

**Theorem 1:** As \( \rho \to 1 \) we have
\[
\overline{E}(\alpha, \alpha, \rho) = (1 - \alpha) + \frac{1}{2} h^{-1}(\alpha) (1 + h^{-1}(\alpha)) + o(1) \quad (10)
\]
for all \( 0 < \alpha < 1 \). Theorem 1 will follow from combining Proposition 1 and Proposition 3, proved in Section II and Section III, respectively.

**Theorem 2:** As \( \rho \to 1 \) we have
\[
\underline{E}(\alpha, \beta, \rho) = (1 - \alpha) + (1 - \beta) + \rho \log e \left( 1 - 2h^{-1}(\alpha) h^{-1}(\beta) \right) + o(\rho) \quad (11)
\]
for all \( 0 < \alpha, \beta < 1 \). Theorem 2 will follow from combining Proposition 2 and Proposition 6, proved in Section II and Section IV, respectively.

In both cases, the optimal exponents are obtained (up to \( o(\rho) \) and \( o(1 - \rho) \) terms) by taking \( A, B \) to be Hamming spheres. In Section II we compute \( P_{XY}(A \times B) \) for Hamming spheres, and prove the corresponding upper bound for \( \overline{E}(\alpha, \alpha, \rho) \) obtained by concentric spheres, and the lower bound on \( \overline{E}(\alpha, \beta, \rho) \) obtained by spheres with opposite centers. In Section III we prove the lower bound on \( \overline{E}(\alpha, \alpha, \rho) \), which is obtained by applying a recent improvement [1] of this inequality for functions of small support (cf. Section III). In Section IV we prove the upper bound on \( \underline{E}(\alpha, \beta, \rho) \) by bounding the maximal average Hamming distance between members of \( A, B \), subject to the cardinality constraint – another combinatorial optimization problem of possible interest.

**Remark 1:** After this work had been completed, we have learned from Naomi Kirshner and Alex Samorodnitsky about their concurrent work [22] in which, among other things, they were able to show that \( \overline{E}(\alpha, \alpha, \rho) \) is attained by concentric spheres for all \( 0 < \rho < 1 \). Their result subsumes our Theorem 1 and relies on a different strengthening of a hypercontractive inequality.1 The problems of characterizing \( \overline{E}(\alpha, \beta, \rho) \) for \( \alpha \neq \beta \) and that of \( \underline{E}(\alpha, \beta, \rho) \) remain open.

## II. Bounds via Spheres

For \( x = (x_1, \ldots, x_n) \in \{0, 1\}^n \) denote the Hamming weight of \( x \) and the Hamming sphere centered at zero as
\[
|x| \triangleq \| j : x_j = 1 \| \quad (12)
\]
\[
S_j \triangleq \{ x : |x| = j \}. \quad (13)
\]
For the size of Hamming spheres we have [23, Exc. 5.8]
\[
|S_{[dn]}| = \left( \frac{n}{[dn]} \right) = 2^{hn(d) - \frac{1}{2} \log n + O(1)} \quad (14)
\]

1In the notation of Section III, our work leverages the inequality \( \| T_n f \|_{\| \cdot \|_1} \leq \| f \|_1 \) among all support-constrained functions \( f \) (with the best possible \( q \)), whereas the work [22] uses the inequality \( \| T_n f \|_{\| \cdot \|_1} \leq e^{-n\lambda} \| f \|_{1+(\rho_0-1)\varepsilon^2} \) with the largest possible \( \lambda \), which depends on the support size of \( f \).
where the estimate is a consequence of Stirling’s formula, \( O(1) \) is uniform in \( \delta \) on compact subsets of \((0,1)\).

Existential results (an upper bound on \( F \) and a lower bound on \( E \)) follow from taking \( A \) and \( B \) as Hamming spheres \( S_{\alpha} \), \( S_{\beta} \) for a suitable \( i, j \). Here we compute the probability of such spherical rectangles.

For any two sets \( A, B \subset \{0,1\}^n \), we have

\[
P_{XY}(A \times B) = \sum_{x \in A, y \in B} 2^{-n} \left[ 2 \log(1 + \rho) - d \log \left( \frac{1 - \rho}{1 + \rho} \right) \right]
\]

where

\[
E(A, B, \rho) \triangleq \min_{0 \leq d \leq 1} \left[ 2 \log(1 + \rho) - d \log \left( \frac{1 - \rho}{1 + \rho} \right) - W_d(A, B) \right],
\]

and \( W_d(A, B) \) is as defined in (5). Note that if \( nd \notin \mathbb{N} \) we have that \( W_d(A, B) = -\infty \), and therefore the minimization in (16) can indeed be performed on \([0,1]\) and need not be restricted to \( d \in \{0,1/n, \ldots, 1\} \).

For two natural numbers \( j \geq i \) and \( d \in [0,1] \) such that \( j - i + nd \) is even, we have that

\[
W_d(S_{\alpha}, S_{\beta}) = \frac{1}{n} \log \left( \binom{n}{i} \frac{1}{2} \right) \left( j + i - nd \right) + \frac{n - i}{2} \right) \right),
\]

for \( j - i \leq nd \leq j + i \), and \( W_d(S_{\alpha}, S_{\beta}) = -\infty \) otherwise. Let \( 0 < \alpha \leq \beta \leq 1 \) and \( d \in [0,1] \) be such that \( i = nh^{-1}(\alpha) \) and \( j = nh^{-1}(\beta) \) are integers and \( j - i + nd \) is an even integer. Approximating \( \binom{n}{k} \approx 2^{n(h(k/n) + o(1))} \) as in (14), we have (18), (19) and (20), shown at the bottom of the page, and it therefore follows from (17) that

\[
W_d(S_{\alpha}, S_{\beta}) = w_d(\alpha, \beta) + o(1),
\]

dense as \( n \) grows, by continuity of \( d \mapsto w_d(\alpha, \beta) \), we have that

\[
E(S_{nh^{-1}(\alpha)}, S_{nh^{-1}(\beta)}, \rho) = \min_{0 \leq d \leq 1} \left[ 2 \log(1 + \rho) - w_d(\alpha, \beta) - d \log \left( \frac{1 - \rho}{1 + \rho} \right) \right] + o(1).
\]

**Proposition 1:** For large \( \rho \) we have

\[
E(\alpha, \alpha, \rho) \leq (1 - \alpha) + \frac{1}{r} - \rho \log \left( \frac{1}{1 - \rho} \right) + o(1).
\]

**Proof:** Let \( 0 < \alpha \leq 1 \). We establish the claim by evaluating \( P_{XY}(A \times B) \) for \( A = B = S_{nh^{-1}(\alpha)} \) and \( \rho = 1 - \epsilon \).

Denoting \( r = r_{\alpha} = h^{-1}(\alpha) \), we have that

\[
w_d(\alpha, \alpha) = h(r) + r \cdot h \left( \frac{d/2}{r} \right) + (1 - r) \cdot h \left( \frac{d/2}{1 - r} \right).
\]

The function \( d \mapsto w_d(\alpha, \alpha) - d \log \left( \frac{d/2}{r} \right) + d \frac{2}{r} \log(e) \) is concave and its derivative

\[
\frac{1}{2} \log \left( 1 + \frac{d/2}{r} \right) + (1 - r) \cdot h \left( \frac{d/2}{1 - r} \right) - \log \left( \frac{2}{r} \right) + \frac{2}{r} \log(e)
\]

equals zero at \( d^* = \epsilon \sqrt{r(1 - r)} + o(\epsilon) \). Thus, the optimizing \( d \) in (25) is \( d^* = \epsilon \sqrt{r(1 - r)} + o(\epsilon) \), and therefore

\[
E(S_{nh^{-1}(\alpha)}, S_{nh^{-1}(\alpha)}, 1 - \epsilon) = 1 - h(r) + \frac{1}{r} \log(e) + \epsilon \sqrt{r(1 - r)} + o(\epsilon)
\]

\[
- \left[ r \cdot h \left( \sqrt{\frac{r - 1}{1 - r} \epsilon} \right) + (1 - r) \cdot h \left( \sqrt{\frac{r}{1 - r} \epsilon} \right) \right] + o(\epsilon) + o(1).
\]
We approximate the term in the square brackets in equations (29), (30) and (31), shown at the bottom of the page. Substituting (31) into (28), we obtain
\[ E\left(S_{nh^{-1}(\alpha)}, S_{nh^{-1}(\alpha)}, 1 - \epsilon \right) = 1 - h(r) \]
\[ + \left( \frac{1}{2} - \sqrt{r(1 - r)} \right) \epsilon \log(e) + o(\epsilon) + o(1). \] (32)

The claim now follows by definition of \( E^\text{ord}(\alpha, \alpha, \rho) \).

**Proposition 2:** For small \( \rho \) we have that
\[ E^\text{ord}(\alpha, \beta, \rho) \geq (1 - \alpha) + (1 - \beta) + \rho \log e \left( 1 - 2h^{-1}(\alpha) * h^{-1}(\beta) \right) + o(\rho). \] (33)

**Proof:** We establish the claim by evaluating \( P_{XY}(A \times B) \) for \( A = S_{nh^{-1}(\alpha)} \) and \( B = 1^n + S_{nh^{-1}(\beta)} \), i.e., a zero-centered Hamming sphere and a Hamming sphere centered around the all-ones vector \( 1^n \). First, note that for any \( A, B \subset \{0, 1\}^n \) it holds that
\[ W_d(A, 1^n + B) = W_{1-d}(A, B). \] (34)
Thus, applying (23), we see that for \( 0 < \alpha \leq \beta \leq 1 \) it holds that
\[ E\left(S_{nh^{-1}(\alpha)}, 1^n + S_{nh^{-1}(\beta)}, \rho \right) = \min_{0 \leq d \leq 1} \left( 2 - \log(1 + \rho) - w_d(\alpha, \beta) \right) \]
\[ - (1 - d) \log \left( \frac{1 - \rho}{1 + \rho} \right) + o(1) \]
\[ = 2 - \log(1 + \rho) - \max_{0 \leq d \leq 1} \left( w_d(\alpha, \beta) - d \log \left( \frac{1 - \rho}{1 + \rho} \right) \right) + o(1) \] (35)
Let us consider the case of \( \rho \ll 1 \). In this case, we have that
\[ \log(1 + \rho) = \rho \log e + o(\rho) \], so that (35) reads
\[ E\left(S_{nh^{-1}(\alpha)}, 1^n + S_{nh^{-1}(\beta)}, \rho \right) = 2 + \rho \log e \]
\[ - \max_{d} \left( w_d(\alpha, \beta) + 2d \log e + o(\rho) + o(1) \right). \] (36)
The function \( d \mapsto w_d(\alpha, \beta) \equiv g(d) \) is strictly concave, and it is straightforward to verify that \( g'(h^{-1}(\alpha) * h^{-1}(\beta)) = 0 \) and that \( g(h^{-1}(\alpha) * h^{-1}(\beta)) = \alpha + \beta \). Denoting \( c = 2g''(h^{-1}(\alpha) * h^{-1}(\beta)) < 0 \) and setting \( \delta = d - h^{-1}(\alpha) * h^{-1}(\beta) \), we therefore have
\[ g(d) = \alpha + \beta + c\delta^2 + o(\delta^2). \] (37)
Consequently,
\[ w_d(\alpha, \beta) + 2d \log e = g(d) + 2d \rho \log e \]
\[ = \alpha + \beta + c\delta^2 + 2h^{-1}(\alpha) * h^{-1}(\beta) + \delta \rho \log e + o(\delta^2) \]
\[ = \alpha + \beta + \rho \log e \cdot 2h^{-1}(\alpha) * h^{-1}(\beta) + \delta \rho \log e + o(\delta^2) \]
\[ \leq \alpha + \beta + \rho \log e \cdot 2h^{-1}(\alpha) * h^{-1}(\beta) + o(\rho), \] (38)
where the last inequality follows since \( c < 0 \). Substituting (38) into (36) we obtain
\[ E\left(S_{nh^{-1}(\alpha)}, 1^n + S_{nh^{-1}(\beta)}, \rho \right) \geq (1 - \alpha) + (1 - \beta) + \rho \log e \left( 1 - 2h^{-1}(\alpha) * h^{-1}(\beta) \right) + o(\rho) + o(1). \] (39)
The claim now follows by definition of \( E^\text{ord}(\alpha, \beta, \rho) \).

**III. LOWER BOUND ON \( T^\text{ord}_{\alpha, \alpha, \rho} \)**

For a function \( f : \{0, 1\}^n \rightarrow \mathbb{R}^+ \) and \( \rho \geq 1 \) we define \( \|f\|_p = E^{1/p}[|f(X)|^p] \). For a set \( A \subset \{0, 1\}^n \) denote
\[ 1_A(x) \triangleq \begin{cases} 0, & x \not\in A \\ 1, & x \in A \end{cases} \]
We have that
\[ P_{XY}(A \times B) = E[1_A(X)1_B(Y)] \]
\[ = E[1_B(Y)E[1_A(X)|Y]] \]
\[ = E[1_B(Y)(T_{\rho}1_A)(Y)], \] (40)
where
\[ (T_{\rho}f)(y) \triangleq E[f(X)|Y = y]. \] (41)
Denoting the inner-product \( (f, g) = E[f(Y)g(Y)] \) and noticing that \( T_{\rho} \) is self-adjoint and satisfies the semigroup property
\[ r \cdot h \left( \sqrt{1 - \frac{r}{2}} \right) = -\sqrt{r(1 - r)} \frac{\epsilon}{2} \log \left( \sqrt{1 - \frac{r}{2}} \right) - r \left( 1 - \sqrt{1 - \frac{r}{2}} \right) \log \left( 1 - \sqrt{1 - \frac{r}{2}} \right) \]
\[ = -\sqrt{r(1 - r)} \frac{\epsilon}{2} \log \left( \sqrt{1 - \frac{r}{2}} \right) + \frac{\epsilon}{2} \sqrt{r(1 - r)} \log(e) + o(\epsilon), \] (29)
\[ (1 - r) \cdot h \left( \sqrt{\frac{r}{1 - r}} \right) = -\sqrt{r(1 - r)} \frac{\epsilon}{2} \log \left( \sqrt{\frac{r}{1 - r}} \right) - (1 - r) \left( 1 - \sqrt{\frac{r}{1 - r}} \right) \log \left( 1 - \sqrt{\frac{r}{1 - r}} \right) \]
\[ = -\sqrt{r(1 - r)} \frac{\epsilon}{2} \log \left( \sqrt{\frac{r}{1 - r}} \right) + \frac{\epsilon}{2} \sqrt{r(1 - r)} \log(e) + o(\epsilon), \] (30)
\[ r \cdot h \left( \sqrt{\frac{1 - r}{r}} \right) + (1 - r) \cdot h \left( \sqrt{\frac{r}{1 - r}} \right) = -\sqrt{r(1 - r)} \epsilon \log \left( \frac{\epsilon}{2} \right) - \epsilon \sqrt{r(1 - r)} \log(e) + o(\epsilon) \]
\[ = \sqrt{r(1 - r)} \epsilon \log \left( \frac{1}{\epsilon} \right) + \epsilon \sqrt{r(1 - r)} + \epsilon \sqrt{r(1 - r)} \log(e) + o(\epsilon). \] (31)
some indicator functions. smooth, convex and strictly increasing bijection.

Differential equation (ODE) exists a function \( C \): \([0, 1) \rightarrow \mathbb{R}^+ \) with \(|f|_{p_0} \geq e^{\lambda e \|f\|_1} \) we have

\[
\|T_{t \rightarrow f}\|_{p(t)} \leq \|f\|_{p_0}, \quad p(t) = 1 + e^{u(t)},
\]

where \( u(t) \) is the unique solution on \([0, \infty)\) of the following ODE with initial condition \( u(0) = \ln(p_0 - 1) \)

\[
u'(t) = C \left( \lambda_0 (1 + e^{-u(t)}) \right)
\]

\[
C(\ln(2(1 - h(y))) = \frac{-2 + \sqrt{1 + 4(\int (1 - y) \ln(2(1 - h(y))}}{\ln 2}
\]

Furthermore, the function \( C : [0, \ln 2] \rightarrow [2, 2/\ln 2] \) is a smooth, convex and strictly increasing bijection.

From this result we derive the following implication for indicator functions.

**Theorem 4**: Fix \( 0 < \alpha < 1 \) and \( 1 < q_0 < \infty \). Then there exists a function \( q = q(t) \) defined on an interval \( t \in [0, \epsilon) \) for some \( \epsilon > 0 \) such that for all sets \( A \subset \{0, 1\}^n \) with \(|A| \leq 2^{n\alpha}\) we have

\[
\|T_{t \rightarrow f}\|_{q(t)} \leq \|A\|_{q(t)} \quad \forall t \in [0, \epsilon). \tag{46}
\]

The function \( q(t) \) satisfies

\[
q(t) = q_0 - (q_0 - 1)C((-1 - \alpha)\ln 2)t + O(t^2) \quad \text{as} \quad t \to 0. \tag{47}
\]

**Remark 2**: Note that the standard hypercontractivity estimate \([2]-[5]\) yields the same result without restriction on the size of the set \( A \) but with a strictly worse (larger) function \( q(t) = (q_0 - 1)e^{-2t} + 1 \). See [1, Remark 3].

**Proof**: Denote by \( u_f(a, b, t) \) the solution of the ordinary differential equation (ODE)

\[
\frac{du}{dt}(t) = C(b(1 + e^{-u(t)})),
\]

with \( u(0) = a \). Here \( C(\cdot) \) is a function defined in (60), \( a \in \mathbb{R} \) and \( 0 < b < (1 + e^{-\alpha})^{-1} \ln 2 \). For a fixed \( a, b \) the standard results on ODEs imply that this solution exists and is unique in some neighborhood \(-\epsilon < t < \epsilon \) of zero. Furthermore, for any \( a_0, b_0 \) satisfying \( 0 < b_0 < (1 + e^{-\alpha_0})^{-1} \ln 2 \) there exists an \( \epsilon_1 > 0 \) such that the map

\[
(a, b, t) \mapsto u_f(a, b, t)
\]

is smooth for \( |a - a_0| < \epsilon_1, |b - b_0| < \epsilon_1, |t| < \epsilon_1 \) (for both of these results, cf. [24, Chapter 2, Section 7, Corollary 6]. We set \( a_0 = \ln(q_0 - 1) \) and \( b_0 = (1 - \alpha)(1 - q_0^{-1}) \ln 2 \). We will call triplets \((a, b, t)\) in the above neighborhood of \((a_0 ,b_0 ,0)\) admissible.

From (44) we have for any admissible \((a, b, s)\) with \( s \geq 0 \) and any \( A \) with \(|A| \leq 2^{n\alpha}\):

\[
\|T_{t \rightarrow f}\|_{1 + e^{\alpha t}} \leq \|f\|_{1 + e^{\alpha t}}, \tag{48}
\]

provided that \( b(1 + e^{-\alpha}) \leq (1 - \alpha)\ln 2 \) (this is just the condition \(|f|_{p_0} \geq e^{\lambda e \|f\|_1} \) of Theorem 3).

Our aim is to set \( s = t \) in (48) and show that there exists a choice of \( a = a(t) \) and \( b = b(t) \) and \( \epsilon < \epsilon_1 \) such that the following conditions are satisfied: (C1) \( a(0) = a_0, b(0) = b_0 \) and both functions are smooth on \(|t| < \epsilon \); (C2) for any \(|t| < \epsilon \) the triplet \((a(t), b(t), t)\) is admissible; (C3) for each \(|t| < \epsilon \)

\[
\left\{ \begin{array}{l}
(b(t) + e^{-\alpha(t)})(1 - \alpha) = \ln 2 \\
u_f(a(t), b(t), t) = \ln(q_0 - 1)
\end{array} \right. \quad \text{as} \quad t \to 0.
\]

It is clear that if indeed such a choice of \((a(t), b(t))\) were found we get from (48) with \( s = t \) the statement of the Theorem with \( q(t) = 1 + e^{\alpha(t)} \).

We claim that it is sufficient to show that the system of equations

\[
\left\{ \begin{array}{l}
f(a, b) = 0, \\
u_f(a, b, t) = \ln(q_0 - 1)
\end{array} \right. \quad \text{as} \quad t \to 0.
\]

has non-trivial Jacobian at \((a_0, b_0, 0)\). Indeed, denoting \( \partial_x = \frac{\partial}{\partial x} \) the Jacobian is given by

\[
Jac(a, b, t) = (\partial_x f)(\partial_x u_f) - (\partial_y f)(\partial_y u_f).
\]

To evaluate this we note an identity \( u_f(a, b, 0) = a \) and thus

\[
\begin{align*}
\frac{\partial}{\partial a}|_{t=0} u_f(a, b, t) &= 1, \\
\frac{\partial}{\partial b}|_{t=0} u_f(a, b, t) &= 0, \\
\frac{\partial}{\partial t}|_{t=0} u_f(a, b, t) &= C(b(1 + e^{-\alpha})).
\end{align*}
\]

Therefore, at \((a = a_0, b = b_0, t = 0)\) the Jacobian evaluates to

\[
Jac(a_0, b_0, 0) = -1 \neq 0.
\]

Since the Jacobian is non-zero in some neighborhood of \((a_0, b_0, 0)\), the map (51) can be locally inverted, and we take for \((a(t), b(t))\) the pre-image of \((0, 0, t)\) under (51).

Finally, we need to show that \( q(t) = 1 + e^{\alpha(t)} \) satisfies the expansion (47). To that end, we differentiate over \( t \) the identity

\[
u_f(a(t), b(t), t) = \ln(q_0 - 1)
\]
This clearly implies that $q(t) = 1 + e^{a(t)}$ satisfies (47).

The following application of the previous result establishes the hard direction of Theorem 1.

Proposition 3: Fix $\rho \in (0,1)$. Then for any sets $A, B$ with $|A| \leq 2^{n\alpha}$, $|B| \leq 2^{n\alpha}$ we have

$$P_{XY}(A \times B) \leq 2^{-n\psi(\alpha, \rho)},$$

where as $\rho \to 1$ we have

$$\psi(\alpha, \rho) = (1 - \alpha) + \frac{1}{\ln 2}(1/2 - \sqrt{h^{-1}(\alpha)(1 - h^{-1}(\alpha))})(1 - \rho) + o(1 - \rho).$$

**Remark 3:** For bounding $E(\alpha, \beta, \rho)$ with $\alpha \neq \beta$ this method does not give a bound matching that attained by Hamming spheres. The main reason is that if we take $A, B$ as concentric (but grossly unequal) Hamming balls the Cauchy-Schwarz inequality (43) is applied to functions $T_{p,1}A$, $T_{p,1}B$ which have effectively disjoint supports for $\rho \to 1$.

**Proof:** Let $\rho = e^{-2t}$ for some fixed $t$. Suppose the sets $A, B$ both have sizes at most $2^{n\alpha}$. Then from Theorem 4 we obtain

$$||T_{e^{-t}}1_A||_2 \leq ||1_A||_{p(t)}$$

$$||T_{e^{-t}}1_B||_2 \leq ||1_B||_{p(t)}$$

$$p(t) = 2 - (2 - 1)C((1 - \alpha) \ln 2)t + o(t).$$

Since $||1_A||_q = 2^{-n(1-\alpha)/q}$ we get from (43) the following:

$$\frac{1}{n} \log P_{XY}(A \times B) \leq - \frac{2}{p(t)}(1 - \alpha)$$

$$= -(1 - \alpha) \left(1 + \frac{t}{2}C((1 - \alpha) \ln 2) + o(t)\right)$$

$$= -(1 - \alpha) - \frac{t(1 - \alpha)}{2} \frac{2 - 4\sqrt{h^{-1}(\alpha)(1 - h^{-1}(\alpha))}}{(1 - \alpha) \ln 2} + o(t).$$

The statement now follows since $t = \frac{1-\rho}{1+\rho} + o(1 - \rho)$.

**IV. Upper Bound on $E(\alpha, \beta, \rho)$**

Note that

$$P_{XY}(A \times B) = \sum_{a \in A, b \in B} \Pr(X = a, Y = b)$$

$$= |A| \cdot |B|$$

$$\cdot \frac{1}{|A| \cdot |B|} \sum_{a \in A, b \in B} 2^{-n} \left(\frac{1 + \rho}{2}\right)^n \cdot \left(\frac{1 - \rho}{1 + \rho}\right)^d(a,b)$$

$$\geq |A| \cdot |B|$$

$$\cdot 2^{-n} \left(\frac{1 + \rho}{2}\right)^n \cdot \left(\frac{1 - \rho}{1 + \rho}\right)^d(a,b)$$

$$= 2 \left(\frac{2 - \log(|A| \cdot |B|)}{n} - \log(1 + \rho) - \frac{\log(1 + p)}{|A| \cdot |B|}\right) \sum_{a \in A, b \in B} d(a,b),$$

(64)

where we have used Jensen’s inequality in (64). As $\frac{1 - p}{1 + p} < 1$, we need to upper bound $\frac{1 - p}{1 + p} \sum_{a \in A, b \in B} d(a,b)$ in terms of $|A|$ and $|B|$ in order to further lower bound (65). Consequently, we define

$$\tilde{d}(n, \alpha, \beta)$$

$$= \frac{1}{n} \max_{A, B; |A| = 2^{n\alpha}, |B| = 2^{n\beta}} |A| \cdot |B| \sum_{a \in A, b \in B} d(a,b)$$

$$= \frac{1}{n} \min_{A, B; |A| = 2^{n\alpha}, |B| = 2^{n\beta}} |A| \cdot |B| \sum_{a \in A, b \in B} d(a,b).$$

(67)

With these definitions we relax (65) to

$$P_{XY}(A \times B) \geq 2^{-n((1-\alpha) + (1-\beta) - \log(1 + \rho) - \tilde{d}(n, \alpha, \beta) \log(1 + \rho))}.$$

(68)

It is obvious that $\tilde{d}(n, \alpha, \beta) = 1 - \tilde{d}(n, \alpha, \beta)$, since if the sets $(A, B)$ achieve the minimal average distance, the sets $(A, B') = (A, B)$ must achieve the maximal average distance. A quantity similar to $\tilde{d}(n, \alpha, \beta)$, where the optimization in (66) is performed over all families $A$ of size $2^{n\alpha}$ while $B = A$ was defined in [20, p.10 eq. 1], and its asymptotic (in $n$) value, was characterized in [25]. Below we prove a lower bound on $\tilde{d}(n, \alpha, \beta)$. The technique is quite similar to that of [25], and requires the following simple proposition.

Proposition 4: The function $\varphi(x, y) = h^{-1}(x) \ast h^{-1}(y)$ is jointly convex in $(x, y) \in [0,1]^2$.

The function $\varphi(x, y)$ is plotted in Figure 1. To prove Proposition 4, we will rely on the following simpler statement, which is essentially proved in [25]. For completeness we provide the proof in the appendix.

Proposition 5: The function $x \mapsto h^{-1}(x) \cdot (1 - h^{-1}(x))$ is convex in $[0,1]$.

**Proof of Proposition 4:** Let $(X, Y)$ be two (possibly dependent) random variables on $[0,1]^2$. We use the identity
\[ a \cdot b = \frac{1}{2} (1 - (1 - 2a)(1 - 2b)) \]

to write

\[ \mathbb{E}[\varphi(X, Y)] \]
\[ = \frac{1}{2} (1 - \mathbb{E} [(1 - 2h^{-1}(X))(1 - 2h^{-1}(Y))]) \]
\[ \geq \frac{1}{2} \left( 1 - \sqrt{\mathbb{E} [(1 - 2h^{-1}(X))^2]} \right) \]
\[ \sqrt{\mathbb{E} [(1 - 2h^{-1}(Y))^2]} \]
\[ \geq \frac{1}{2} \left( 1 - \sqrt{(1 - 2h^{-1}(\mathbb{E}[X]))^2} \right) \]
\[ \sqrt{(1 - 2h^{-1}(\mathbb{E}[Y]))^2} \]
\[ = \varphi(\mathbb{E}[X], \mathbb{E}[Y]), \] (72)

where (70) follows from the Cauchy-Schwarz inequality, and (71) from Jensen’s inequality and the fact that \( t \mapsto (1 - 2h^{-1}(t))^2 = 1 - 4h^{-1}(t)(1 - h^{-1}(t)) \) is concave due to Proposition 5.

**Lemma 1:** For any two independent \( n \)-dimensional random binary vectors \( V \) and \( W \)

\[ h^{-1} \left( \frac{n}{n} (H(V) - h^{-1} \left( \frac{n}{n} H(W) - h^{-1} \left( \frac{n}{n} \right) \right) \right) \]
\[ \leq 1 - h^{-1} \left( \frac{n}{n} H(V) \right) \times h^{-1} \left( \frac{n}{n} H(W) \right) \] (73)

**Proof:** Let \( V \) and \( W \) be two independent random vectors with \( H(V) = n\alpha \) and \( H(W) = n\beta \). Further, let \( a_i \triangleq \Pr(V_i = 1), b_i \triangleq \Pr(W_i = 1) \), be the induced marginal distributions for each coordinate. Our goal is to minimize and maximize \( \sum_{i=1}^{n} a_i \cdot b_i \) under the entropy constraints \( H(V) = n\alpha, H(W) = n\beta \). We may and will assume without loss of generality that \( a_i, b_i \leq 1/2 \) for all \( i \). We have

\[ \inf_{V, W:\ H(V) = n\alpha, H(W) = n\beta} \sum_{i=1}^{n} a_i \cdot b_i \]
\[ = \inf_{\{a_i, b_i\}_{i=1}^{n}} \sum_{i=1}^{n} a_i \cdot b_i \] (74)
\[ \geq \sum_{i=1}^{n} h^{-1}(\alpha_i) \cdot h^{-1}(\beta_i) \] (75)

where (74) follows since the cost function \( \sum_{i=1}^{n} a_i \cdot b_i \) depends only on the marginal distributions, and for every feasible distribution \( V, W \) the product of the marginalized distributions is also feasible. Our lower bound now immediately follows from Proposition 4. For the upper bound, note that if \( V \) and \( W \) minimize \( \mathbb{E}[D(V, W)] \) under the entropy constraints, \( V \) and \( W' = W + 1^n \) maximizes the expected distance under the same entropy constraints.

Taking \( V \sim \text{Uniform}(A) \) and \( W \sim \text{Uniform}(B) \), we immediately get the following.

**Corollary 1:**

\[ \bar{d}(n, \alpha, \beta) \leq n \left( 1 - h^{-1}(\alpha) \cdot h^{-1}(\beta) \right), \] (76)
\[ \bar{d}(n, \alpha, \beta) \geq nh^{-1}(\alpha) \cdot h^{-1}(\beta). \] (77)
Combining (68) and Corollary 1, gives
\[
E(\alpha, \beta, \rho) \leq (1 - \alpha) + (1 - \beta) - \log(1 + \rho) \\
- (1 - h^{-1}(\alpha) * h^{-1}(\beta)) \log \frac{1 - \rho}{1 + \rho} \\
= (1 - \alpha) + (1 - \beta) - \log(1 - \rho) \\
+ (h^{-1}(\alpha) * h^{-1}(\beta)) \log \frac{1 - \rho}{1 + \rho}.
\]
(78)

We have therefore obtained the following.

**Proposition 6:** We have
\[
E(\alpha, \beta, \rho) \leq (1 - \alpha) + (1 - \beta) + \rho \log(e) \left(1 - 2h^{-1}(\alpha) * h^{-1}(\beta)\right) + o(\rho).
\]
(79)

**Remark 4:** In [8] the bound
\[
E(\alpha, \beta, \rho) \leq \frac{(1 - \alpha) + (1 - \beta) + 2\rho \sqrt{(1 - \alpha)(1 - \beta)}}{1 - \rho^2}
\]
(80)

was proved, using reverse hypercontractivity. It is easy to verify that for \(\alpha = \beta\) the bound (78) is strictly better than (80) for all \(\alpha < 1 - \frac{1 - \rho^2}{2p} \log \left(\frac{1 - \rho}{1 + \rho}\right)\). Moreover, for any \(0 < \alpha, \beta < 1\) the bound (78) is better than (80) for \(\rho\) large enough. The reverse hypercontractivity bound states that for \(p < 1\) we have \(\|T_p f\|_{q(p, \rho)} \geq \|f\|_p\) where \(q(p, \rho) = 1 - \frac{1 - \rho^2}{p}\) for \(\rho < 1\). The weakness of this bound in our setup is that the function \(q(p, \rho)\) does not depend on the support of \(f\), which is exponentially small. It is quite plausible that deriving support dependent reverse hypercontractivity bounds, analogous to the support dependent hypercontractivity bounds of [1], would result in tighter upper bounds on \(E(\alpha, \beta, \rho)\) in the high-correlation regime.

**APPENDIX**

Let \(\phi(x) = (1 - 2h^{-1}(x))^2\). Since \(h^{-1}(1 - h^{-1}(x)) = 1 - \frac{4}{\log e}v(h^{-1}(x))\), it suffices to show that \(x \mapsto \phi(x)\) is concave. We have
\[
\phi'(x) = \frac{4}{\log e}v(h^{-1}(x)),
\]
(81)

where
\[
v(t) = \frac{1 - 2t}{\ln(\frac{1}{1 - t})}.
\]
(82)

Showing that \(x \mapsto \phi'(x)\) is concave is equivalent to showing that \(x \mapsto \phi''(x)\) is decreasing, which in turn is equivalent to showing that \(t \mapsto v'(t)\) is increasing in \((0, 1/2)\), due to monotonicity of \(x \mapsto h^{-1}(x)\). Thus, it remains to show that \(v'(t) \geq 0\) for \(t \in (0, 1/2)\). Let \(y = y_t = \frac{1 - 2t}{\ln(\frac{1}{1 - t})} \in (1, \infty)\).

We have that \(v'(t) = \frac{1 - 2t \ln(\frac{1}{1 - t})}{\frac{1 - 2t}{\ln(\frac{1}{1 - t})} - 2 \ln(y)}\) and since \(\frac{1 - 2t}{\ln(\frac{1}{1 - t})} = \frac{y^2 - 1}{y^2}\), it suffices to show that \(g(y) = y^2 - 1 - 2 \ln(y) \geq 0\) for all \(y > 1\). Noting that \(g(1) = 0\) and \(g'(y) = 1 + \frac{1}{y^2} - 2 = \frac{y^2 - 3}{y^2} \geq 0\) for all \(y \geq 1\), we see that indeed \(g(y) \geq 0\) for all \(y \geq 1\), which establishes our claim.

**ACKNOWLEDGMENT**

The authors would like to thank the anonymous reviewers and the associate editor for their excellent suggestions, and in particular for a simplification of the proof of Proposition 4.

**REFERENCES**

[1] Y. Polyanskiy and A. Samorodnitsky, “Improved log-Sobolev inequalities, hypercontractivity and uncertainty principle on the hypercube,” J. Funct. Anal., vol. 277, no. 11, Dec. 2019, Art. no. 108280.

[2] E. Nelson, “A quartic interaction in two dimensions,” in *Mathematical Theory of Elementary Particles*, R. Goodman and I. Segal, Eds. Cambridge, MA, USA: MIT Press, 1966.

[3] A. Bonami, “Cône des dés de Fourier des fonctions de Ip(g),” Annales de l’Institut Fourier, vol. 20, no. 2, pp. 335–402, 1970.

[4] W. Beckner, “Inequalities in Fourier analysis on \(\mathbb{R}^n\),” *Proc. Nat. Acad. Sci. USA*, vol. 72, no. 2, pp. 638–641, 1975.

[5] L. Gross, “Logarithmic Sobolev inequalities,” *Ann. Math.*, vol. 97, no. 4, pp. 1061–1083, 1973.

[6] R. O’Donnell, *Analysis of Boolean Functions*. Cambridge, U.K.: Cambridge Univ. Press, 2014.

[7] C. Borell, “Positive improving operators and hypercontractivity,” *Mathematische Zeitschrift*, vol. 180, pp. 225–234, Sep. 1982.

[8] E. Mossel, R. O’Donnell, O. Regev, J. E. Steif, and B. Sudakov, “Non-interactive correlation distillation, inhomogeneous Markov chains, and the reverse Bonami-Beckner inequality,” *Isr. J. Math.*, vol. 154, no. 1, pp. 299–336, Dec. 2006.

[9] A. Bogdanov and E. Mossel, “On extracting common random bits from correlated sources,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6351–6355, Oct. 2011.

[10] E. Mossel, O. Regev, J. E. Steif, and B. Sudakov, “Non-interactive correlation distillation, inhomogeneous Markov chains, and the reverse Bonami-Beckner inequality,” *Isr. J. Math.*, vol. 154, no. 1, pp. 299–336, Dec. 2006.

[11] R. A. K. J. W. A. S. R. A. D. B. W., “Bounds on the achievable rates of block coding for a memoryless multiple-access channel,” *IEEE Trans. Inf. Theory*, vol. IT-24, no. 1, pp. 112–116, Jan. 1978.

[12] N. B. Y. C. N. M. K. P. A. V. T. K. S. L. W. V. T. W. S. W., “Bounds on the achievable rates of block coding for a memoryless multiple-access channel,” *IEEE Trans. Inf. Theory*, vol. IT-24, no. 2, pp. 187–197, Mar. 1978.

[13] E. Weldon, “Coding for a multiple-access channel,” *Inf. Control*, vol. 36, no. 3, pp. 256–274, 1978.

[14] T. Kasami, S. Lin, W. Wei, and S. Yamamura, “Graph theoretic approaches to the code construction for the two-user multiple-access binary adder channel,” *IEEE Trans. Inf. Theory*, vol. IT-29, no. 1, pp. 114–130, Jan. 1983.

[15] R. Urbanke and Q. Li, “The zero-error capacity region of the 2-user synchronous BAC is strictly smaller than its Shannon capacity region,” in *Proc. Inf. Theory Workshop*, Jun. 1998, p. 61.

[16] O. Ordentlich and O. Shayevitz, “An upper bound on the sizes of multiset-union-free families,” *SIAM J. Discrete Math.*, vol. 30, no. 2, pp. 1032–1045, Jan. 2016.

[17] G. Ajianagadde and Y. Polyanskiy, “Adder MAC and estimates for Rényi entropy,” in *Proc. 53rd Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2015, pp. 434–441.

[18] D. J. Kleinberg, “On a combinatorial conjecture of Erdős,” *J. Combinat. Theory*, vol. 1, no. 2, pp. 209–214, 1966.

[19] R. Ahlswede and G. O. H. Katona, “Contributions to the geometry of Hamming spaces,” *Discrete Math.*, vol. 17, no. 1, pp. 1–22, 1977.

[20] P. Frankl and Z. Füredi, “A short proof for a theorem of harper about Hamming-spheres,” *Discrete Math.*, vol. 34, no. 3, pp. 311–313, 1981.

[21] A. A. A. S. A., “A moment ratio bound for polynomials and some extremal properties of Krawchouk polynomials and Hamming spheres,” 2019, arXiv:1909.11192. [Online]. Available: http://arxiv.org/abs/1909.11192

[22] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.

[23] V. Arnold, *Ordinary Differential Equations*. Springer, 1992.

[24] R. Ahlswede and I. A. A., “The asymptotic behavior of diameters in the average,” *J. Combinat. Theory, B*, vol. 61, no. 2, pp. 167–177, Jul. 1994.
Or Ordentlich (Member, IEEE) received the B.Sc. (*cum laude*), M.Sc. (*summa cum laude*), and Ph.D. degrees in electrical engineering from Tel Aviv University, Israel, in 2010, 2011, and 2016, respectively. From 2015 to 2017, he was a Post-Doctoral Fellow with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology (MIT), and the Department of Electrical and Computer Engineering, Boston University. He is currently a Senior Lecturer (Assistant Professor) with the School of Computer Science and Engineering, Hebrew University of Jerusalem.

Yury Polyanskiy (Senior Member, IEEE) received the M.S. degree in applied mathematics and physics from the Moscow Institute of Physics and Technology, Moscow, Russia, in 2005, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 2010. He is currently an Associate Professor of electrical engineering and computer science and a member of IDSS and LIDS with MIT. His research interests include information theory, statistical learning, error-correcting codes, wireless communication, and fault tolerance. He received the 2020 IEEE Information Theory Society James Massey Award, the 2013 NSF CAREER Award, and the 2011 IEEE Information Theory Society Paper Award.

Ofer Shayevitz (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from the Technion Institute of Technology, Haifa, Israel, in 1997, and the M.Sc. and Ph.D. degrees in electrical engineering from Tel-Aviv University, Tel Aviv, Israel, in 2004 and 2009, respectively. He is currently an Associate Professor with the Department of EE-Systems, Tel Aviv University, and serves as the Head of the Advanced Communication Center (ACC). His research interests include wide cross-section of problems in information theory, statistical signal processing, and discrete mathematics. He was a recipient of the European Research Council (ERC) Starting Grant from 2015 to 2020, and his group’s research activities have further been supported by Israel Science Foundation grants from 2014 to 2022, and the Marie Curie Grant from 2014 to 2018. He is also actively involved in the Israeli hi-tech industry, and regularly consults to various startup companies. Before joining Tel Aviv University, he was a Post-Doctoral Fellow with the Information Theory and Applications (ITA) Center, University of California, San Diego, from 2008 to 2011, and worked as a Quantitative Analyst with the D.E. Shaw Group, New York, from 2011 to 2013. Prior to his graduate studies, he served as a Digital Communication Engineer and a Team Leader at the Israeli Defense Forces from 1997 to 2003, and worked in statistical signal processing at CellGuide from 2003 to 2004.