Combinatorics, Probability and Computing (2021), **30**, pp. 513–525
doi:10.1017/S0963548320000541

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Lipschitz bijections between boolean functions

Tom Johnston* and Alex Scott

Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK
*Corresponding author. Email: thomas.johnston@maths.ox.ac.uk

(Received 21 December 2018; revised 11 September 2020; accepted 30 May 2019; first published online 16 November 2020)

Abstract

We answer four questions from a recent paper of Rao and Shinkar [17] on Lipschitz bijections between functions from $\{0,1\}^n$ to $\{0,1\}$. (1) We show that there is no $O(1)$-bi-Lipschitz bijection from Dictator to XOR such that each output bit depends on $O(1)$ input bits. (2) We give a construction for a mapping from XOR to Majority which has average stretch $O(\sqrt{n})$, matching a previously known lower bound. (3) We give a 3-Lipschitz embedding $\phi\colon \{0,1\}^n \to \{0,1\}^{2n+1}$ such that $\mathrm{XOR}(x) = \mathrm{Majority}(\phi(x))$ for all $x \in \{0,1\}^n$. (4) We show that with high probability there is an $O(1)$-bi-Lipschitz mapping from Dictator to a uniformly random balanced function.

2020 MSC Codes: Primary: 68R05; Secondary: 05D40, 94D10

1. Introduction

Given two boolean functions $f, g\colon \{0,1\}^n \to \{0,1\}$, we say that a bijection $\phi\colon \{0,1\}^n \to \{0,1\}^n$ is a *mapping from f to g* if, for every $x \in \{0,1\}^n$, we have $f(x) = g(\phi(x))$. The analysis of boolean functions, in particular their Fourier coefficients and noise stability, is widely studied and has applications in many areas of mathematics including the theory of social choice, complexity theory and in the hardness of approximations (see *e.g.* [8], [13], [14], [16], [7], [4], [9], [1], [12], [3]). A frequent theme in the literature is the analysis of the similarities and differences between boolean functions with different geometric or structural properties, for example between functions such as Dictator that are determined by a small number of coordinates, and functions such as Majority or XOR that are not. One measure of similarity between functions is the existence of a Lipschitz mapping (with small constant) between them. In this paper we continue the study of Rao and Shinkar [17] on Lipschitz mappings between the boolean functions Dictator, XOR, Majority and a uniformly random balanced function, and answer several of the questions they pose.

Write a point $x \in \{0,1\}^n$ as $x = (x_1, \ldots, x_n)$ and, for $\phi\colon \{0,1\}^n \to \{0,1\}^n$, write $\phi = (\phi_1, \ldots, \phi_n)$. For $x, y \in \{0,1\}^n$ let $|x| = \sum_{i=1}^n x_i$ denote the Hamming weight of x and $\mathrm{dist}(x,y) = \sum_{i=1}^n |x_i - y_i|$ the Hamming distance between x and y.

A bijection $\phi\colon \{0,1\}^n \to \{0,1\}^n$ is said to be *C-Lipschitz* if, for all $x, y \in \{0,1\}^n$, $\mathrm{dist}(\phi(x), \phi(y)) \leqslant C\,\mathrm{dist}(x,y)$, and ϕ is said to be *C-bi-Lipschitz* if both ϕ and ϕ^{-1} are C-Lipschitz. As a relaxation from being Lipschitz we define the *average stretch* of a mapping $\phi\colon \{0,1\}^n \to \{0,1\}^n$ by

$$\mathrm{avgStretch}(\phi) = \mathbb{E}_{x,i}[\mathrm{dist}(\phi(x), \phi(x + e_i))],$$

where $x \in \{0,1\}^n$ and $i \in [n]$ are both chosen independently and uniformly at random.

© The Author(s), 2020. Published by Cambridge University Press.

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press

Given a bijection $\phi\colon \{0,1\}^n \to \{0,1\}^n$ we say that the jth output bit ϕ_j *depends on* the ith input bit x_i if there exists $x \in \{0,1\}^n$ such that $\phi_j(x) \neq \phi_j(x + e_i)$. If every output bit depends on at most k input bits, we say the map ϕ is *k-local*.

In Sections 2 and 3 we study mappings between three boolean functions, Dictator, XOR and Majority, which we define by

- $\text{Dictator}(x) = x_1$,
- $\text{XOR}(x) = \sum_{i=1}^{n} x_i \mod 2$,
- $\text{Majority}(x) = 1$ if $|x| > n/2$ and $\text{Majority}(x) = 0$ otherwise.

In Section 4 we consider a uniformly random balanced function $f\colon \{0,1\}^n \to \{0,1\}$, where we say a boolean function $g\colon \{0,1\}^n \to \{0,1\}$ is *balanced* if $g^{-1}(1)$ and $g^{-1}(0)$ are of the same size. Clearly, both XOR and Dictator are always balanced while Majority is only balanced when n is odd (and so a bijection from Dictator or XOR to Majority can only exist for odd n).

1.1 Bijections between Dictator *and* Majority

The functions Dictator and Majority are in many ways opposites of each other. For example, the first coordinate clearly has a large influence on the value of Dictator, while for Majority every coordinate has the same small amount of influence. There are many results which show that functions which differ from Majority in some way must have influential coordinates and are therefore similar to Dictator functions. For example, the 'Majority Is Stablest' theorem of Mossel, O'Donnell and Oleszkiewicz in [16] shows that if a balanced boolean function is essentially more noise-stable than Majority, then it must have an influential coordinate.

It is straightforward to see that no bijection ϕ from Dictator to Majority can be C-Lipschitz for any $C < n/2$. Indeed, suppose ϕ is such a bijection and let $x \in \{0,1\}^n$ be such that $\phi(x) = \sum_{i=1}^{n} e_i$. Clearly $y := x - e_1$ differs from x in the first coordinate so $\text{Dictator}(y) \neq \text{Dictator}(x)$. By the definition of ϕ, we must have $\text{Majority}(\phi(y)) = 0$ and $|\phi(y)|$ is at most $n/2$. Hence $\text{dist}(\phi(x), \phi(y)) \geqslant n/2$, which gives a contradiction.

For maps from Majority to Dictator, the situation is better. Rao and Shinkar [17] showed that, for all odd n, there is a mapping ψ from Majority to Dictator which is 11-Lipschitz. As noted above, the function ψ^{-1} cannot be $O(1)$-Lipschitz. However, it has the weaker property that ψ^{-1} has $O(1)$-average stretch. Indeed, Rao and Shinkar's construction gives a Lipschitz function that is in fact an $O(1)$-bi-Lipschitz bijection from the upper half of $\{0,1\}^n$ to the half-cube with first coordinate 1; similarly ψ induces an $O(1)$-bi-Lipschitz bijection between the lower half of $\{0,1\}^n$ and the half-cube with first co-ordinate 0. As there are only 2^{n-1} edges between the two half-cubes, the average stretch of ψ^{-1} is $O(1)$.

1.2 Bijections between Dictator *and* XOR

Rao and Shinkar note that the map ϕ given by $\phi(x) = (\text{XOR}(x), x_2, x_3, \ldots, x_n)$ is a mapping from Dictator to XOR, and it is easy to check that ϕ is 2-bi-Lipschitz. This leads them to consider maps with stronger properties. In the above map, the first output bit depends on all n input bits and the map is not k-local for any $k < n$. However, one can easily find a map which is 2-local: Rao and Shinkar give the example $\phi(x) = (x_1 + x_2, x_2 + x_3, \ldots, x_{n-1} + x_n, x_n)$ which is 2-Lipschitz and 2-local, although the inverse of this map is not even $(n-1)$-Lipschitz. In [17] Rao and Shinkar construct a map ϕ which is 2-Lipschitz and 3-local and where the inverse is $O(\log n)$-Lipschitz. This leads them to ask the following question.

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press

Question 1.1. **([17, Question 6.1]).** Is there a mapping from Dictator to XOR that is $O(1)$-local and $O(1)$-bi-Lipschitz?

We answer this in the negative with the following theorem.

Theorem 1.2. *Let* $\phi\colon \{0,1\}^n \to \{0,1\}^n$ *be a mapping from* Dictator *to* XOR *which is C-Lipschitz and where each output bit depends on at most k input bits. Then there is a constant* $\delta = \Omega(1/(k + \log(C)))$ *such that the inverse map* ϕ^{-1} *is not* $\delta \log n$*-Lipschitz.*

Furthermore, if ϕ *is a linear mapping, then* $\delta = \Omega(1/\log(C + k))$.

It follows that the map constructed by Rao and Shinkar in [17] is essentially best possible: if ϕ is a mapping from Dictator to XOR which is $O(1)$-Lipschitz and $O(1)$-local, then ϕ^{-1} cannot be $o(\log n)$-Lipschitz. We prove Theorem 1.2 in Section 2.

1.3 Bijections between XOR *and* Majority

Since we have a 2-bi-Lipschitz map from Dictator to XOR, we might expect the maps between XOR and Majority to behave similarly to those between Dictator and Majority. Indeed, composing this 2-bi-Lipschitz map from Dictator to XOR with the $O(1)$-Lipschitz map from Majority to Dictator gives an $O(1)$-Lipschitz map from Majority to XOR, and the argument used to show that there is no $O(1)$-Lipschitz mapping from Dictator to Majority can also be applied to mappings from XOR to Majority to show that they cannot be C-Lipschitz for $C < n/2$. However, while there is a map from XOR to Dictator with constant average stretch, Rao and Shinkar [17] show that for any mapping ϕ from XOR to Majority the average stretch is $\Omega(\sqrt{n})$. They then ask if this lower bound is tight.

Question 1.3 **([17, Question 6.3]).** Is there a mapping ϕ from XOR to Majority such that $\text{avgStretch}(\phi) = O(\sqrt{n})$?

In Section 3 we show that the bound is indeed tight by giving a map which has average stretch $O(\sqrt{n})$. This result also follows from a recent result of Boczkowski and Shinkar [5] where they prove, for any two sets $A, B \subseteq \{0,1\}^n$ with $|A| = |B| = 2^{n-1}$, there is a mapping $\phi\colon A \to B$ with $\mathbb{E}[\,\text{dist}(x, \phi(x))] \leqslant \sqrt{2n}$.

Theorem 1.4. *For odd n there is a mapping* ϕ *from* XOR *to* Majority *such that*

$$\text{avgStretch}(\phi) = O(\sqrt{n}).$$

Given the need for large average stretch of such a mapping, Rao and Shinkar also ask what happens if we relax the problem from finding a bijection to finding an embedding to a larger space.

Question 1.5 **([17, Question 6.4]).** Is there a Lipschitz embedding $\phi\colon \{0,1\}^n \to \{0,1\}^{\text{poly}(n)}$ such that $\text{XOR}(x) = \text{Majority}(\phi(x))$ for all $x \in \{0,1\}^n$?

We give a simple construction that gives a positive answer.

Theorem 1.6. *For every n there exists a* 3*-Lipschitz embedding* $\phi\colon \{0,1\}^n \to \{0,1\}^{2n+1}$ *such that* $\text{XOR}(x) = \text{Majority}(\phi(x))$ *for all* $x \in \{0,1\}^n$.

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press

1.4 Bijections from Dictator *to a random f*

Rao and Shinkar also consider random functions. Building on a construction of Hastad, Leighton and Newman [11], they show that with high probability there is a bijection from Dictator to a uniformly random balanced function f which has average stretch bounded by an absolute constant [17]. They ask whether we can in fact ask for more.

Question 1.7 ([17, Question 6.2]). Is it true that with high probability there is an $O(1)$-bi-Lipschitz mapping from Dictator to a uniformly random balanced function f?

In Section 4 we give a positive answer to this question. In fact we prove a stronger statement.

Theorem 1.8. *Let $f\colon \{0,1\}^n \to \{0,1\}$ be a balanced boolean function chosen uniformly at random. Then, with high probability, there exists a mapping $\phi\colon \{0,1\}^n \to \{0,1\}^n$ from* Dictator *to f such that for all $x \in \{0,1\}^n$,*

$$\operatorname{dist}(x, \phi(x)) \leqslant C,$$

where C is an absolute constant. In particular, ϕ is $(2C+1)$-bi-Lipschitz.

It is straightforward to extend this to bijections between two independent uniformly random balanced boolean functions f and g.

Corollary 1.1. *Let $f, g\colon \{0,1\}^n \to \{0,1\}$ be two balanced boolean functions chosen independently and uniformly at random. Then, with high probability, there exists a mapping $\phi\colon \{0,1\}^n \to \{0,1\}^n$ from f to g such that for all $x \in \{0,1\}^n$,*

$$\operatorname{dist}(x, \phi(x)) \leqslant C,$$

where C is an absolute constant. In particular, ϕ is $(2C+1)$-bi-Lipschitz.

Proof. By the above there exists a mapping ϕ_f from Dictator to f such that $\operatorname{dist}(x, \phi_f(x)) \leqslant c$ and a mapping ϕ_g from Dictator to g such that $\operatorname{dist}(x, \phi_g(x)) \leqslant c$ for some absolute constant c, both with high probability. Then $\phi = \phi_g \circ \phi_f^{-1}$ is a mapping from f to g such that $\operatorname{dist}(x, \phi(x)) \leqslant c^2$. This proves the first part of the corollary with $C = c^2$.

The second part follows as in the proof of Theorem 1.8. Fix $x \in \{0,1\}^n$ and let $y = x + e_i$ for some i. Then

$$\operatorname{dist}(\phi(x), \phi(y)) \leqslant \operatorname{dist}(x, \phi(x)) + \operatorname{dist}(x, y) + \operatorname{dist}(y, \phi(y)) \leqslant 2C + 1$$

and the triangle inequality shows that ϕ is $(2C+1)$-Lipschitz. A similar argument shows that ϕ^{-1} is also $(2C+1)$-Lipschitz. □

The rest of this paper is organized as follows. In Section 2 we prove Theorem 1.2, and in Section 3 we prove Theorem 1.4 and Theorem 1.6. We then prove Theorem 1.8 in Section 4.

2. Bijections from Dictator to XOR

Given a mapping $\phi\colon \{0,1\}^n \to \{0,1\}^n$ from Dictator to XOR, we define the *dependency graph* D_ϕ as follows. Let D_ϕ be the bipartite graph with vertex sets $A = \{a_1, \dots, a_n\}$ and $B = \{b_1, \dots, b_n\}$, where there is an edge $a_i b_j$ if and only if the jth output bit of ϕ depends on the ith input bit. The following lemma shows that if there is an output bit ϕ_j which is at a large distance from a_1 in the dependency graph, then changing ϕ_j must cause many input bits to change.

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press

Lemma 2.1. *Suppose ϕ is a mapping from* Dictator *to* XOR *such that the distance between a_1 and b_j in D_ϕ is at least d. Then, for any $y \in \{0,1\}^n$,*

$$\operatorname{dist}(\phi^{-1}(y), \phi^{-1}(y+e_j)) \geqslant \frac{d+1}{2}.$$

Proof. Suppose that $\phi^{-1}(y) = x$ and $\phi^{-1}(y+e_j) = x'$ and let the set of coordinates in which they differ be X. Now let D'_ϕ be the subgraph of D_ϕ induced by $A_X := \{a_i : i \in X\}$ and its neighbours. Since $\phi(x)$ and $\phi(x')$ differ in the jth bit, one of the neighbours of b_j must be in A_X so that $b_j \in V(D'_\phi)$. As ϕ is a mapping from Dictator to XOR and $\mathrm{XOR}(y) \neq \mathrm{XOR}(y+e_j)$, we must have $\mathrm{Dictator}(x) \neq \mathrm{Dictator}(x')$ and so $a_1 \in A_X$. If a_1 and b_j are in the same component of D'_ϕ, then D'_ϕ must contain a path from a_1 to b_j, which must have length at least d. Since D'_ϕ is bipartite, $|X| \geqslant (d+1)/2$ and we are done.

Otherwise, let the component of D'_ϕ containing a_1 have vertices $A_Y \cup B_{Y'}$, where $Y, Y' \subset [n]$. Consider the input $x = e_Y$, where $e_Y = \sum_{i \in Y} e_i$: we will show that $\phi(x+e_Y) = \phi(x)$. If $y' \in Y'$, then the output bit $\phi_{y'}$ depends only on the input bits in Y and input bits that are not in X, so $\phi_{y'}(x+e_Y) = \phi_{y'}(x') = \phi_{y'}(x)$ (since $\phi(x)$ and $\phi(x')$ differ only in the jth bit and $j \notin Y'$). Changing the value of x on Y can only change $\phi(x)$ in the output bits in Y', so $\phi_i(x) = \phi_i(x+e_Y)$ for every $i \notin Y'$. This means that $\phi(x) = \phi(x+e_Y)$, which gives a contradiction as ϕ is a bijection. □

Now that we have related distance in the dependency graph D_ϕ to the Hamming distance of the inverse map, we prove Theorem 1.2 by showing that the conditions on ϕ imply that there is a vertex at least logarithmically far from a_1 in D_ϕ.

Proof of Theorem 1.2. As each output bit of $\phi(x)$ depends on at most k input bits of x, the degree of a vertex b_j in D_ϕ is at most k.

Let us now show that the degrees of the a_i are also bounded. Fix a vertex a_i in D_ϕ and let it have neighbours $b_{j_1}, \ldots, b_{j_m}$. The value of the output bit j_1 depends on i and at most $k-1$ other bits which we denote X. By definition, there is some assignment x of the $k-1$ bits X such that $\phi(x)$ and $\phi(x+e_i)$ differ in the bit j_1. Let $U \in \{0,1\}^n$ be a bit string chosen uniformly at random: then the probability that this is equal to x on X is 2^{1-k} and, hence, the probability that $\phi(U)$ and $\phi(U+e_i)$ differ in bit j_1 is at least 2^{1-k}. The same holds for bits $j_2, \ldots, j_m$ and so, by the linearity of expectation, the expected number of bits in which $\phi(U)$ and $\phi(U+e_i)$ differ is at least $m2^{1-k}$. There must be some value for U for which the number of bits that differ is at least $m2^{1-k}$ and, as ϕ is C-Lipschitz, we must have $m2^{1-k} \leqslant C$. In particular, the degree of a vertex a_i is bounded by $C2^{k-1}$.

Define $\Delta := \max\{k, C2^{k-1}\}$ so that every vertex has degree at most Δ. The number of vertices at distance at most l from a_1 is at most $(\Delta+1)^l$ and hence, if the distance from a_1 to any vertex is at most d, $(\Delta+1)^d \geqslant 2n$ and $d \geqslant (\log 2n)/\log(\Delta+1)$. As the graph is bipartite, there must be a vertex b_j at distance at least $d-1$ from a_1 and hence, by Lemma 2.1, ϕ is not $\delta \log(n)$-Lipschitz for $\delta = 1/2\log(\Delta+1)$.

If ϕ is a linear map, then $\phi(x)$ and $\phi(x+e_i)$ differ in the bits $j_1, \ldots, j_m$ for every x. Hence the degree of a vertex a_i is bounded by C (compared to $C2^{k-1}$ in the general case), and the bound on δ follows from bounding the maximum degree by $C+k$. □

The construction presented by Rao and Shinkar in [17] gives a linear map which shows $\delta(2,k) = O(1/\log(k))$. Theorem 1.2 shows that this is tight for linear maps, but only gives the bound $\delta(2,k) = \Omega(1/k)$ for general maps. This raises the following question: Does there exist a map ϕ from Dictator to XOR which is C-Lipschitz, where each output bit depends on at most k input bits and such that ϕ^{-1} is $o((\log n)/\log(k))$-Lipschitz as $n, k \to \infty$?

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press

3. Bijections from XOR to Majority

It was shown by Rao and Shinkar in [17] that any mapping $\phi\colon \{0,1\}^n \to \{0,1\}^n$ from XOR to Majority must have average stretch $\Omega(\sqrt{n})$ in any direction. In this section we prove Theorem 1.4 by constructing a map with average stretch $O(\sqrt{n})$. Our strategy is to map each $x \in \{0,1\}^n$ for which $\mathrm{XOR}(x) = \mathrm{Majority}(x)$ to itself, and otherwise to swap x with a y chosen such that $|x| + |y| = n$. The problem is then to find a matching so that when the matched elements are switched, they are not too far from their neighbours on average. We do this by matching according to a symmetric chain decomposition so the Hamming distance between switched elements is minimized. This is similar to the proof of Theorem 1.2 in [17], which defines a mapping from Majority to Dictator by permuting elements along the chains in a particular symmetric chain decomposition. Recall that a *symmetric chain* is a path $C = \{c_k, \dots, c_{n-k}\}$ in the hypercube such that $|c_i| = i$. A *symmetric chain decomposition* is a partition of the hypercube into symmetric chains. It is well known that these decompositions exist for all n (see *e.g.* [6]).

Proof of Theorem 1.4. Suppose we have a symmetric chain decomposition of $\{0,1\}^n$ and, for a point $x \in \{0,1\}^n$, let y_x be the unique point in the chain containing x such that $|x| + |y_x| = n$. We note that $\mathrm{XOR}(x) \neq \mathrm{XOR}(y_x)$ and $\mathrm{Majority}(x) \neq \mathrm{Majority}(y_x)$, so the bijection $\phi\colon \{0,1\}^n \to \{0,1\}^n$ defined by

$$\phi(x) = \begin{cases} x & \text{if } \mathrm{XOR}(x) = \mathrm{Majority}(x), \\ y_x & \text{if } \mathrm{XOR}(x) \neq \mathrm{Majority}(x) \end{cases}$$

is a mapping from XOR to Majority.

Claim 1. *Suppose $x \in \{0,1\}^n$ has Hamming weight m. Then, for any i,*

$$\mathrm{dist}(\phi(x), \phi(x + e_i)) \leqslant |n - 2m| + 3.$$

Proof of Claim 1. The proof of this claim is a case analysis over $\mathrm{XOR}(x)$, $\mathrm{Majority}(x)$, whether $e_i \leqslant x$ and, where necessary, $\mathrm{Majority}(x + e_i)$. Let us fix i and use x' to denote $x + e_i$.

First consider the case where $\mathrm{XOR}(x) = 1$, $\mathrm{Majority}(1) = 1$ and $x_i = 0$. Then we have $|x'| = m + 1$, $\mathrm{XOR}(x') = 0$ and $\mathrm{Majority}(x') = 1$. Hence ϕ keeps x constant and switches x' to $y_{x'}$. We know that $y_{x'} \leqslant x'$ and is of weight $n - m - 1$, so x and $y_{x'}$ must agree in at least $n - m - 2$ places. Hence the distance between x and $y_{x'}$ is bounded by $1 + m - (n - m - 2) = 2m - n + 3$.

Now consider the case $\mathrm{XOR}(x) = 1$, $\mathrm{Majority}(1) = 1$ and $x_i = 1$. In this case it is possible for $\mathrm{Majority}(x') = 1$ or $\mathrm{Majority}(x') = 0$, so we must consider these cases separately. Suppose $\mathrm{Majority}(x') = 1$. Then ϕ keeps x constant and switches x' to $y_{x'}$, where $|y_{x'}| = n - m + 1$. We also have $y_{x'} \leqslant x + e_i \leqslant x$, so the stretch is $2m - n - 1$. If $\mathrm{Majority}(x + e_i) = 0$, then ϕ keeps both x and $x + e_i$ constant and the stretch is 1.

The other cases follow similarly. □

For a uniformly chosen x and an arbitrarily chosen direction, we have

$$\mathbb{E}_{x,i}[\mathrm{dist}\,(\phi(x), \phi(x + e_i))] \leqslant 3 + 2^{-n} \sum_{m=0}^{n} |2m - n| \binom{n}{m}. \tag{3.1}$$

We can bound $|2n - m|$ by n and the binomial coefficient $\binom{n}{m}$ by $\binom{n+1}{(n+1)/2}$. Applying standard results, we get

$$\mathbb{E}_{x,i}[\mathrm{dist}\,(\phi(x), \phi(x + e_i))] \leqslant 3 + \frac{n}{2^n} \binom{n+1}{(n+1)/2} \sim \sqrt{\frac{8n}{\pi}}.$$ □

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press

Given that the average stretch for a mapping from XOR to Majority must be $\Omega(\sqrt{n})$, Rao and Shinkar ask if it is possible to relax the problem by increasing the size of the codomain and asking instead for an $O(1)$-Lipschitz injection. In the proof of Theorem 1.6 below we give one such example.

Proof of Theorem 1.6. Define $\phi(x)\colon \{0,1\}^n \to \{0,1\}^{2n+1}$ by

$$\phi(x) = (x^c, \mathrm{XOR}(x), x),$$

where x^c is the bitwise complement of x (and where we have written (a, b, c) for the concatenation of vectors a, b and c). As we can read off x from the last section, this is clearly a one-to-one mapping. If x has Hamming weight m, x^c has Hamming weight $n-m$, so overall $\phi(x)$ has Hamming weight $n + \mathrm{XOR}(x)$. Hence $\mathrm{Majority}(\phi(x)) = \mathrm{XOR}(x)$ for all $x \in \{0,1\}^n$.

Suppose $x, y \in \{0,1\}^n$ are distinct. Then

$$\begin{aligned}\mathrm{dist}(\phi(x), \phi(y)) &= \mathrm{dist}(x^c, y^c) + \mathrm{dist}(\mathrm{XOR}(x), \mathrm{XOR}(y)) + \mathrm{dist}(x, y)\\ &= 2\,\mathrm{dist}(x, y) + \mathrm{XOR}(x+y)\\ &\leqslant 3\,\mathrm{dist}(x, y)\end{aligned}$$

and the map is 3-Lipschitz. □

4. Bijections from Dictator to a random *f*

Let f be a boolean function. We say a point $x \in \{0,1\}^n$ is a 1 *of* f if $f(x) = 1$ and similarly say x is a 0 *of* f if $f(x) = 0$. We say a set B is *balanced* if B contains an equal number of 1s and 0s of f, is *positively imbalanced* if B contains more 1s of f than 0s of f and *negatively imbalanced* if B contains more 0s of f than 1s of f. The *imbalance* of f over a set B is the unsigned difference between the number of 1s of f and $|B|/2$.

Before we prove Theorem 1.8, let us first sketch the ideas behind the proof without worrying about the technicalities.

Sketch proof of Theorem 1.8. We first partition the hypercube $\{0,1\}^n$ into 'blocks' which contain a large polynomial number of points (say n^a), have bounded diameter and contain an equal number of 1s and 0s of Dictator. These blocks are small in comparison to the cube so, in a given block B, we expect the number of 1s of f to be distributed similarly to a binomial distribution with $|B|$ trials and success probability $1/2$. This means we expect every block to have an imbalance not much more than $n^{a/2}$ and a large proportion (tending to 1) of the blocks to have an imbalance not much less than $n^{a/2}$. We call a sequence of distinct blocks $B_1, \ldots, B_k$ a *block path* if, for $i = 1, \ldots, k-1$, there is an edge between a vertex of B_i and a vertex of B_{i-1}. In our construction, we take a large set of random block paths in the hypercube and use these to move the imbalance. If a block B_1 has too many 1s of f, we find a block path $B_1, \ldots, B_k$ from B_1 to a block B_k with too many 0s and map a 0 of Dictator from B_1 to a 0 of f from B_2, a 0 of Dictator from B_2 to a 0 of f from B_3, and so on. We also do the same with the 1s but in the opposite direction. That is, we map a 1 of Dictator from B_2 to a 1 of f from B_1, map a 1 of Dictator from B_3 to a 1 of f from B_2, and so on. By doing this along enough paths, we can even out the sets and then arbitrarily match within them.

For this to work we need to make sure that we do not try to map too many points to a block, which means ensuring there are not too many paths through any single block. However, we also need to make sure there are enough paths between the blocks to spread the imbalance around. Once we have chosen our paths (at random) we construct a bipartite graph between the positive and negative blocks and find a subgraph with suitable degrees. This corresponds to the paths that will actually be used to move the imbalance around. The imbalance is not much more than $n^{a/2}$ in any block, so we do not need too many paths; at the same time, the imbalance is not

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press

much less than $n^{a/2}$ in most blocks, which ensures that there are enough random paths between imbalanced sets. □

Our proof of Theorem 1.8 uses the fact that, with high probability, there is a perfect matching in our random bipartite graph. The following lemma is an immediate consequence of a result of Erdős and Rényi [10]. (The lemma can easily be strengthened but this form is sufficient for our case.)

Lemma 4.1. *Let A be a multi-set of at least $2n \log n$ pairs (i, j), each chosen uniformly and independently at random from $[n] \times [n]$. Let G be the random bipartite graph with vertex set $\{v_1^+, \dots, v_n^+, v_1^-, \dots, v_n^-\}$, where $v_i^+ v_j^-$ is an edge if and only if $(i, j) \in A$. Then the probability that G contains a perfect matching tends to 1 as $n \to \infty$.*

We will also make use of the following Chernoff bounds; see [15] for a discussion and derivation of these bounds.

Lemma 4.2. *Let $X \sim \mathrm{Bin}(n, p)$ and $t > 0$. Then*

$$\mathbb{P}(X \geqslant np + t) \leqslant e^{-2t^2/n}.$$

Also, for $0 \leqslant \varepsilon \leqslant 1$ we have

$$\mathbb{P}(X \geqslant (1+\varepsilon)np) \leqslant e^{-\varepsilon^2 np/4}$$

and

$$\mathbb{P}(X \leqslant (1-\varepsilon)np) \leqslant e^{-\varepsilon^2 np/2}.$$

We are now ready to prove the main theorem. Throughout this proof we will use the constants a, b, c and d, to which we assign explicit values only at the end of the proof.

Proof of Theorem 1.8. Suppose first that we have a mapping $\phi\colon \{0,1\}^n \to \{0,1\}^n$ from Dictator to f such that $\mathrm{dist}(x, \phi(x)) \leqslant C$ for all $x \in \{0,1\}^n$. Fix $x \in \{0,1\}^n$ and let $y = x + e_i$ for some i. Then

$$\mathrm{dist}(\phi(x), \phi(y)) \leqslant \mathrm{dist}(x, \phi(x)) + \mathrm{dist}(x, y) + \mathrm{dist}(y, \phi(y)) \leqslant 2C + 1$$

and the triangle inequality shows that ϕ is $(2C+1)$-Lipschitz. A similar argument shows that ϕ^{-1} is also $(2C+1)$-Lipschitz and this proves the latter part of the theorem. Thus it remains to show that such a ϕ exists with high probability.

We shall assume for the rest of the argument that n is large enough for our estimates to hold. We start by partitioning the hypercube $\{0,1\}^n$ into 'blocks' with diameter at most $4a$ which contain between n^{a-2} and $4n^a$ elements, exactly half of which are 1s of Dictator. Let B be a maximal set of points in $\{0,1\}^{n-1}$ such that the pairwise distance between any two points is at least a. For each point $v \in B$, we define the set of points A_v to be the points in $\{0,1\}^{n-1}$ which are closer to v than to any other point in B (settling ties arbitrarily). As the set B is maximal, the radius of A_v must be less than a, or else we could add a point to B. As the distance from v to any other point in B is at least a, A_v contains a Hamming ball of radius $a-1$ centred at v and at least n^{a-2} points. If a set contains N elements with $N > 2n^a$, we arbitrarily split the set into $\lfloor N/n^a \rfloor$ sets, each containing between n^a and $2n^a$ points. For each of our sets A, we define a corresponding 'block' in $\{0,1\}^n$ to be $\{0,1\} \times A$. Clearly this gives a set $\mathcal{B}$ of blocks with the desired properties. Note that $2^n/(4n^a) \leqslant |\mathcal{B}| \leqslant 2^n/n^{a-2}$.

The following claim shows that, with probability tending to 1, no block $B \in \mathcal{B}$ has an imbalance much more than $n^{a/2}$.

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press

Claim 2. *Provided that $b > 1/2$, with probability tending to 1, none of the sets $B \in \mathcal{B}$ have an imbalance of more than $n^{a/2+b}$.*

Proof. Say that a block B is *bad* under a function f if B has an imbalance between the number of 1s and 0s of f greater than $n^{a/2+b}$. Let $g\colon \{0,1\}^n \to \{0,1\}$ be a random boolean function where $g(x)$ is chosen independently and uniformly at random for each $x \in \{0,1\}^n$. Then the probability that B is bad under f is the same as the probability that B is bad under g conditioned on the event that g is balanced, so

$$\mathbb{P}(B \text{ bad under } f) = \mathbb{P}(B \text{ bad under } g \mid g \text{ balanced}) \leqslant \frac{\mathbb{P}(B \text{ bad under } g)}{\mathbb{P}(g \text{ balanced})}.$$

The number of 1s of g in B is a binomial random variable, so the first Chernoff bound in Lemma 4.2 shows the probability that B is bad under g is at most $2e^{-2n^{a+2b}/N} \leqslant 2e^{-n^{2b}/2}$. The probability that g is balanced is

$$2^{-n}\binom{2^n}{2^{n-1}} \sim (\pi 2^{n-1})^{-1/2}.$$

Using the union bound over all blocks $B \in \mathcal{B}$, the probability that at least one of the sets B is bad is at most $2^{3n/2}e^{-n^{2b}/2}$ for all sufficiently large n, and this is $o(1)$ provided $b > 1/2$. □

We will also need that most sets have an imbalance not much less than $n^{a/2}$.

Claim 3. *Provided that $c > 2$, with probability tending to 1, the sum over $B \in \mathcal{B}$ of the imbalance of B is at least $2^n/(8n^{a/2+c})$.*

Proof. Fix B and let $|B| = N$. We can then write down the probability that the imbalance of B is at most $n^{a/2-c}$ as

$$\sum_{i=-n^{a/2-c}}^{n^{a/2-c}} \frac{\binom{N}{N/2+i}\binom{2^n-N}{2^{n-1}-N/2-i}}{\binom{2^n}{2^{n-1}}}.$$

Using $\binom{2n}{n} \sim 4^n/\sqrt{\pi n}$ and that $N = o(2^n)$, the largest term ($i = 0$) is at most $O(N^{-1/2})$. As $N \geqslant n^{a-2}$, each term is $O(n^{-(a/2-1)})$ and the sum is $O(1/n^{c-1})$. Let X be the number of blocks with imbalance at most $n^{a/2-c}$. Then

$$\mathbb{E}[X] = O\left(\frac{|\mathcal{B}|}{n^{c-1}}\right)$$

and so, by Markov's inequality, $\mathbb{P}(X \geqslant |\mathcal{B}|/n) = O(1/n^{c-2})$. This is $o(1)$ provided $c > 2$, and in this case, with probability tending to 1, the sum of the imbalances is at least

$$(2^n/(4n^a)) \cdot (1 - 1/n) \cdot n^{a/2-c} \geqslant 2^n/(8n^{a/2+c}) \quad \text{for large } n.$$ □

Now suppose that every block has an imbalance of at most $n^{a/2+b}$ and that the total imbalance is at least $2^n/(8n^{a/2+c})$.

Independently sample $K = \lfloor 2^n/n^d \rfloor$ uniformly random pairs of vertices in the hypercube and a random shortest path between them.

Claim 4. *Provided that $a + 1 - d > 1$, with high probability, the number of paths in K which intersect a given block $B \in \mathcal{B}$ is at most $8n^{a+1-d}$.*

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press

Proof. By symmetry, every point is equally likely to be on this random shortest path and, as any shortest path has at most n vertices, the probability that a given vertex is on each sampled path is at most $n/2^n$. Using the union bound, the probability that a given random path goes through a fixed $B \in \mathcal{B}$ is at most $n|B|/2^n$. As each path is sampled independently, the number of paths through a fixed B is dominated by a binomial distribution with K trials and success probability $4n^{a+1}/2^n$. Applying the second Chernoff bound from Lemma 4.2 with $\varepsilon = 1$, the probability that there are more than $8n^{a+1-d}$ paths through a given block is at most $\exp(-n^{a+1-d}/2)$ and, taking the union bound, the probability that any B intersects more than $8n^{a+1-d}$ paths is at most $\exp(-n^{a+1-d}/2) \cdot 2^n/n^{a-2}$. This tends to 0 as $n \to \infty$ provided that $a + 1 - d > 1$. □

We now form a random bipartite graph G with vertex sets V^+ and V^- as follows. If $B \in \mathcal{B}$ has a positive imbalance of i, we add i *positive vertices* $v_1^B, \ldots, v_i^B$ to V^+, and similarly, if B has a negative imbalance of i, we add i *negative vertices* to V^-. We generate the edge set of G by examining each of our random paths in turn. Given a path, say between $v_1 \in B_1$ and $v_2 \in B_2$, we do the following.

- Discard the path with probability $1 - n^{2a-4}/(|B_1| \cdot |B_2|)$.
- If the path has not been discarded, independently choose two numbers i, j in $[n^{a/2+b}]$ and say that the edge $v_i^{B_1} v_j^{B_2}$ is present if both vertices exist and have opposite signs.

Claim 5. *The graph G has a perfect matching M with high probability provided*

$$a/2 > 3b + 2c + d + 7. \tag{4.1}$$

Proof. The probability that a random path gives any particular edge between V^+ and V^- is $p = 2(n^{a/2-2-b}/2^n)^2$ and, in particular, every edge is equally likely. Hence the number of edges in the multi-set induced by the sampled random paths has a binomial distribution with K trials and success probability $|V^+|^2 p$. By assumption $|V^+| \geqslant 2^n/(16n^{a/2+c})$ and so $|V^+|^2 p \geqslant (128n^{2b+2c+4})^{-1}$. Using the third Chernoff bound in Lemma 4.2, the probability the multi-set contains at most $2^{n-9}/n^{4+2b+2c+d}$ pairs is at most $\exp(-2^{n-11}/n^{4+2b+2c+d})$. There are at most $2^n/n^{a-2}$ blocks, each with imbalance at most $n^{a/2+b}$, so $|V^+| \leqslant 2^n/n^{a/2-2-b}$. Hence $2|V^+| \log |V^+| \leqslant 2^{n+1}/n^{a/2-b-3}$. From (4.1), for large enough n,

$$2^{n+1}/n^{a/2-b-3} \leqslant 2^{n-9}/n^{4+2b+2c+d},$$

and we have that G contains a perfect matching M with high probability by Lemma 4.1. □

Given the matching M we can greedily construct the map ϕ. We start with all vertices as both unset and unused. During the construction we will let $\phi(x) = y$ for some $x, y \in \{0,1\}^n$ and we will then say that x is *set* and y is *used*. The vertices for which we still have to define ϕ are exactly the unset vertices, and the vertices which are not yet in the image of ϕ are exactly the unused vertices.

Each edge in M corresponds to a path in the hypercube, and this induces a walk between two blocks in the block decomposition. By removing any loops, we can say the edge corresponds to a path $P = B_{i_1} \ldots B_{i_k}$ in the block decomposition, where B_{i_1} has positive imbalance and B_{i_k} has negative imbalance. For $j = 1, \ldots, k-1$, we choose any point $x \in B_{i_j}$ with $\text{Dictator}(x) = 0$ which is currently unset and any point $y \in B_{i_{j+1}}$ with $f(y) = 0$ which is currently unused, and set $\phi(x) = y$, setting x and using y. We also choose any point in $x' \in B_{i_{j+1}}$ with $\text{Dictator}(x) = 1$ which is currently unset and any point $y' \in B_{i_j}$ with $f(y) = 1$ and set $\phi(x') = y'$. In order to guarantee that there will always be enough points to do the above, we will enforce that $8n^{a+1-d} \leqslant n^{a-2}$.

If B has positive imbalance i, there are i paths starting at B (those corresponding to the edges in the matching using $v_1^B, \ldots, v_i^B$) and hence, after processing all the paths which start at B, the

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press

number of unset 1s of Dictator in B is still $|B|/2$ while the number of 0s is now $|B|/2 - i$. Similarly the number of unused 1s of f in B is now $|B|/2$ while the number of 0s is still $|B|/2 - i$. The paths for which B is an internal block reduce all four quantities by 1 so, after processing all the paths, the number of 1s and 0s of Dictator which are unset and the number of 1s and 0s of f which are unused, are equal. This means we can greedily complete ϕ inside B by taking any unset x and any unused y with $\text{Dictator}(x) = f(y)$ and setting $\phi(x) = y$. A similar argument works when B is balanced or has negative imbalance.

It is clear that, provided nothing goes wrong, this construction gives a bijection $\phi\colon \{0,1\}^n \to \{0,1\}^n$ which maps Dictator to f. If $x \in B$, then either $\phi(x) \in B$ or $\phi(x) \in B'$ for some B' for which there is an edge between B and B' in the hypercube. As the diameter of a block is bounded by $4a$, the distance between x and $\phi(x)$ is bounded by $8a + 1$.

This construction can fail if any block has too large an imbalance, the total imbalance is too small, there are too many paths through a block, there are not enough edges in the multi-set or there is not a matching in G; but all of these events happen with probability tending to 0 (for suitable choices of a, b, c and d).

To complete the proof it remains to find suitable values for the constants a, b, c and d. From Claim 2 we must have $b > 1/2$ and from Claim 3 we must have $c > 2$. We will also need $8n^{a+1-d} \leqslant n^{a-2}$, which follows for large enough n provided $d > 3$. Finally, we require $a \geqslant 2$, $a + 1 - d > 1$ and, from Claim 4, that $a/2 > 3b + 2c + d + 7$. All these constraints can be satisfied by taking $a = 42$, $b = 1$, $c = 3$ and $d = 4$. □

5. Open problems

We saw in Theorem 1.2 that it is not possible for a map from Dictator to XOR to be $O(1)$-Lipschitz, $O(1)$-local and also have an inverse that is $O(1)$-Lipschitz. On the other hand we saw that maps exist if we drop the condition on the inverse or on the locality. Can we drop the first condition while keeping the other two?

Problem 5.1. Is there a mapping $\phi\colon \{0,1\}^n \to \{0,1\}^n$ from Dictator to XOR such that each output bit depends on $O(1)$ input bits and such that ϕ^{-1} is $O(1)$-Lipschitz?

Corollary 1.1 shows that, with high probability, if $A, B \subseteq \{0,1\}^n$ are two independent uniformly random sets with $|A| = |B| = 2^{n-1}$, there exists an $O(1)$-bi-Lipschitz bijection $\phi\colon \{0,1\} \to \{0,1\}$ such that $\phi(A) = B$, but we made no attempt to optimize the constant. We expect a much smaller constant to hold in this corollary (and in Theorem 1.8), possibly even as small as 2 or 3. How small could the constant be?

Problem 5.2. Let $A, B \subseteq \{0,1\}^n$ be two independent uniformly random sets with $|A| = |B| = 2^{n-1}$. What is the smallest constant C such that, with high probability, there is a C-bi-Lipschitz bijection $\phi\colon \{0,1\}^n \to \{0,1\}^n$ such that $\phi(A) = B$?

There are many interesting variations of this problem. For example, what happens if we let $|A| = |B| = f(n)$ for some function $f(n)$?

Problem 5.3. Let $A, B \subseteq \{0,1\}^n$ be two independent uniformly random sets with $|A| = |B| = f(n)$. For what functions $f\colon \mathbb{N} \to \mathbb{N}$ does there exist a C such that, with high probability, there is a C-bi-Lipschitz bijection $\phi\colon \{0,1\}^n \to \{0,1\}^n$ with $\phi(A) = B$? For example, what happens when $f(n) = \Theta(2^n/n)$?

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press

A closely related problem concerns colourings. If we view the partition of $\{0,1\}^n$ into two parts (A and A^c) as a partition into two colour classes, then we can view mappings as relabellings of the cube which respect two balanced 2-colourings. What happens if we instead colour the cube with k colours? If k is a constant, then a modification of the argument used to prove Theorem 1.8 will show that there is an $O(1)$-bi-Lipschitz mapping with high probability. But what happens if we let $k \to \infty$ as $n \to \infty$?

We say that a partition of $\{0,1\}^n$ into k parts $A_1, \dots, A_k$ is a *balanced partition* if $|A_1| \leqslant |A_2| \leqslant \cdots |A_k| \leqslant |A_1| + 1$.

Problem 5.4. Let $A_1, \dots, A_k$ and $B_1, \dots, B_k$ be two independent uniformly random balanced partitions of $\{0,1\}^n$ into $k(n)$ parts. For what functions $k\colon \mathbb{N} \to \mathbb{N}$ does there exist a constant C such that, with high probability, there is a C-bi-Lipschitz bijection $\phi\colon \{0,1\}^n \to \{0,1\}^n$ with $\phi(A_i) = B_i$ for all i? How small can C be?

The problems above have been concerned with functions on the hypercube, but we can ask similar questions for functions on other graphs and, in particular, Cayley graphs. For example, what happens for $\mathbb{Z}_n \times \mathbb{Z}_n$ generated by the elements $(1,0)$ and $(0,1)$? In this case we say that a map $\phi\colon \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n \times \mathbb{Z}_n$ is *C-Lipschitz* if, for every $x, y \in V$, $d(\phi(x), \phi(y)) \leqslant Cd(x,y)$, where d is the graph distance in the Cayley graph. The proof of Theorem 1.8 relied on a decomposition of $\{0,1\}^n$ into 'blocks', each with bounded radius but containing a large polynomial number of points. In $\mathbb{Z}_n \times \mathbb{Z}_n$ any subset with bounded radius must contain a bounded number of points, so the 'blocks' will not tend towards being relatively balanced.

Problem 5.5. Let $A, B \subseteq \mathbb{Z}_n \times \mathbb{Z}_n$ be two independent uniformly random sets with $|A| = |B| = \lceil n^2/2 \rceil$. What is the best $C = C(n)$ such that, with high probability, there is a C-bi-Lipschitz bijection $\phi\colon \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n \times \mathbb{Z}_n$ with $\phi(A) = \phi(B)$?

In a different direction, we observe that there is a natural symmetric measure of the difference between two boolean functions. Define the *Lipschitz constant* of a mapping ϕ by

$$\|\phi\|_{\mathrm{Lip}} = \inf\{C \geqslant 0\colon \|\phi(x) - \phi(y)\| \leqslant C\|x - y\| \text{ for all } x, y \in \{0,1\}^n\},$$

and define the *dissimilarity* between two boolean functions f and g by

$$\mathrm{diss}(f,g) := \inf\{\|\phi\|_{\mathrm{Lip}} \cdot \|\phi^{-1}\|_{\mathrm{Lip}}\colon \phi \text{ is a mapping from } f \text{ to } g\}.$$

The 2-bi-Lipschitz mapping between Dictator and XOR we saw earlier implies that $\mathrm{diss}(\mathrm{Dictator}, \mathrm{XOR}) \leqslant 4$, and Theorem 1.8 states that if f and g are uniformly random balanced boolean functions, $\mathrm{diss}(f,g) = O(1)$ with high probability. On the other hand we know that any mapping from Dictator to Majority is not C-Lipschitz for any $C < n/2$, so $\mathrm{diss}(\mathrm{Dictator}, \mathrm{Majority}) = \Omega(n)$. Rao and Shinkar [17] give an $O(1)$-Lipschitz mapping from Majority to Dictator, so we can strengthen this to $\mathrm{diss}(\mathrm{Dictator}, \mathrm{Majority}) = \Theta(n)$. In general we know that every mapping is n-Lipschitz, which gives a trivial upper bound of n^2, but is it possible to do better?

Problem 5.6. Is there a constant $\alpha < 2$ such that $\mathrm{diss}(f,g) = O(n^\alpha)$ for all balanced boolean functions f and g?

Finally, we note that the paper by Rao and Shinkar is motivated in part by a problem from Benjamini, Cohen and Shinkar in [2], which remains open and seems very interesting.

Problem 5.7. Is there a set $A \subseteq \{0,1\}^n$ of size $|A| = 2^{n-1}$ such that any bijection from $\{0,1\}^{n-1}$ to A has average stretch $\omega(1)$?

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press

Acknowledgements

We would like to thank an anonymous referee for their helpful comments. The second author is supported by a Leverhulme Trust Research Fellowship.

References

[1] Ben-Or, M. and Linial, N. (1989) Collective coin flipping. *Adv. Comput. Research* **5** 91–115.
[2] Benjamini, I., Cohen, G. and Shinkar, I. (2016) Bi-Lipschitz bijection between the boolean cube and the Hamming ball. *Israel J. Math.* **212** 677–703.
[3] Benjamini, I., Kalai, G. and Schramm, O. (1999) Noise sensitivity of boolean functions and applications to percolation. *Publ. Math. Inst. Hautes Études Sci.* **90** 5–43.
[4] Bernasconi, A. (1998) Mathematical techniques for the analysis of boolean functions. *Bull. Eur. Assoc. Theor. Comput. Sci.* **65** 228–230.
[5] Boczkowski, L. and Shinkar, I. (2019) On mappings on the hypercube with small average stretch. arXiv:1905.11350
[6] Bollobás, B. (1986) *Combinatorics: Set Systems, Hypergraphs, Families of Vectors, and Combinatorial Probability*. Cambridge University Press.
[7] Boppana, R. B. (1997) The average sensitivity of bounded-depth circuits. *Inform. Process. Lett.* **63** 257–261.
[8] Bourgain, J. (2002) On the distribution of the Fourier spectrum of boolean functions. *Israel J. Math.* **131** 269–276.
[9] Dinur, I. and Safra, S. (2005) On the hardness of approximating minimum vertex cover. *Ann. of Math.* **162** 439–485.
[10] Erdős, P. and Rényi, A. (1963) On random matrices. *Magyar Tud. Akad. Mat. Kutató Int. Közl* **8** 455–461.
[11] Hastad, J., Leighton, T. and Newman, M. (1987) Reconfiguring a hypercube in the presence of faults. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing (STOC '87)*, pp. 274–284. ACM.
[12] Kahn, J., Kalai, G. and Linial, N. (1988) *The Influence of Variables on Boolean Functions*. IEEE.
[13] Kindler, G., Kirshner, N. and O'Donnell, R. (2018) Gaussian noise sensitivity and Fourier tails. *Israel J. Math.* **225** 71–109.
[14] Lovett, S. and Viola, E. (2011) Bounded-depth circuits cannot sample good codes. In *26th Annual Conference on Computational Complexity (CCC 2011)*, pp. 243–251. IEEE.
[15] Mitzenmacher, M. and Upfal, E. (2005) *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press.
[16] Mossel, E., O'Donnell, R. and Oleszkiewicz, K. (2010) Noise stability of functions with low influences: invariance and optimality. *Ann. of Math.* **171** 295–341.
[17] Rao, S. and Shinkar, I. (2018) On Lipschitz bijections between boolean functions. *Combin. Probab. Comput.* **27** 411–426.

Cite this article: Johnston T and Scott A (2021). Lipschitz bijections between boolean functions. *Combinatorics, Probability and Computing* **30**, 513–525. https://doi.org/10.1017/S0963548320000541

https://doi.org/10.1017/S0963548320000541 Published online by Cambridge University Press