2014 IEEE Annual Symposium on Foundations of Computer Science

# Bi-Lipschitz Bijection between the Boolean Cube and the Hamming Ball

Itai Benjamini
Department of Mathematics,
Weizmann Institute of Science,
Rehovot, Israel
itai.benjamini@weizmann.ac.il

Gil Cohen
Department of Computer Science
Weizmann Institute of Science,
Rehovot, Israel.
gil.cohen@weizmann.ac.il

Igor Shinkar
Department of Computer Science
Weizmann Institute of Science,
Rehovot, Israel.
igor.shinkar@weizmann.ac.il

***Abstract*—We construct a bi-Lipschitz bijection from the Boolean cube to the Hamming ball of equal volume. More precisely, we show that for all even $n \in \mathbb{N}$ there exists an explicit bijection $\psi\colon \{0,1\}^n \to \{x \in \{0,1\}^{n+1}\colon |x| > n/2\}$ such that for every $x \neq y \in \{0,1\}^n$ it holds that**

$$\frac{1}{5} \leq \frac{\mathsf{dist}(\psi(x),\psi(y))}{\mathsf{dist}(x,y)} \leq 4,$$

**where $\mathsf{dist}(\cdot,\cdot)$ denotes the Hamming distance. In particular, this implies that the Hamming ball is bi-Lipschitz transitive.**

**This result gives a strong negative answer to an open problem of Lovett and Viola [CC 2012], who raised the question in the context of sampling distributions in low-level complexity classes. The conceptual implication is that the problem of proving lower bounds in the context of sampling distributions requires ideas beyond the sensitivity-based structural results of Boppana [IPL 97].**

**We study the mapping $\psi$ further and show that it (and its inverse) are computable in DLOGTIME-uniform $\mathbf{TC^0}$, but not in $\mathbf{AC^0}$. Moreover, we prove that $\psi$ is "approximately local" in the sense that all but the last output bit of $\psi$ are essentially determined by a single input bit.**

## I. Introduction

The Boolean cube $\{0,1\}^n$ and the Hamming ball $\mathcal{B}_n = \{x \in \{0,1\}^{n+1}\colon |x| > n/2\}$, equipped with the Hamming distance, are two fundamental combinatorial structures that exhibit, in some aspects, different geometric properties. As a simple illustrative example, for an even integer $n \in \mathbb{N}$, consider the vertex and edge boundaries[1] of $\{0,1\}^n$ and $\mathcal{B}_n$, when viewed as subsets of $\{0,1\}^{n+1}$ of equal density $1/2$. The Boolean cube is easily seen to maximize vertex boundary among all subsets of equal density (since all its vertices lie on the boundary), whereas Harper's vertex-isoperimetric inequality [Har66] implies that the Hamming ball is in fact the unique minimizer. The same phenomena occurs for edge boundary, though interestingly, the roles are reversed: among all monotone sets[2] of density $1/2$, the Poincaré inequality implies that the Boolean cube is the unique minimizer of edge boundary, whereas a classical result of Hart shows that the Hamming ball is the unique maximizer [Har76]. From the Boolean functions perspective, the indicator of $\{0,1\}^n$ embedded in $\{0,1\}^{n+1}$ is commonly referred to as the *dictator* function, and the indicator of $\mathcal{B}_n \subset \{0,1\}^{n+1}$ is the *majority* function, and it is a recurring theme in the analysis of Boolean functions that they are, in some senses, opposites of one another.

Lovett and Viola [LV12] suggested to utilize the opposite structure of the Boolean cube and the Hamming ball for proving lower bounds on sampling by low-level complexity classes such as $\mathbf{AC^0}$ and $\mathbf{TC^0}$. In particular, Lovett and Viola were interested in proving that for any even $n$, any bijection $f\colon \{0,1\}^n \to \mathcal{B}_n$ has a large average stretch, where

$$\mathsf{avgStretch}(f) = \mathop{\mathbf{E}}_{\substack{x\sim\{0,1\}^n \\ i\sim[n]}} \left[\mathsf{dist}(f(x), f(x+e_i))\right],$$

and $\mathsf{dist}(\cdot,\cdot)$ denotes the Hamming distance. To be more precise, Lovett and Viola raised the following open problem.

**Problem I.1** ([LV12], Open Problem 4.1)**.** *Let $n \in \mathbb{N}$ be an even integer. Prove that for any bijection $f\colon \{0,1\}^n \to \mathcal{B}_n$, it holds that*

$$\mathsf{avgStretch}(f) = (\log n)^{\omega(1)}. \tag{1}$$

A positive answer to Problem I.1 would demonstrate yet another scenario in which the Boolean cube and the Hamming ball have a different geometric structure – any bijection from the former to the latter does not respect distances. Furthermore, a positive answer to Problem I.1 would have applications to lower bounds for sampling in $\mathbf{AC^0}$; even a weaker claim, where the right hand side in Equation (1) is replaced with $\omega(1)$, would have implications for sampling in the lower class $\mathbf{NC^0}$. We discuss this further in Section I-B.

Arguably, the simplest and most natural bijection $\varphi\colon \{0,1\}^n \to \mathcal{B}_n$ to consider is the following.

$$\varphi(x) = \begin{cases} \mathsf{flip}(x) \circ 1 & \text{if } |x| \leq n/2 \\ x \circ 0 & \text{otherwise,} \end{cases}$$

where $\mathsf{flip}(x)$ denotes the bit-wise complement of $x$. It is straightforward to verify that $\mathsf{avgStretch}(\varphi) = \Theta(\sqrt{n})$. To see this, note that any edge $(x,y)$ in $\{0,1\}^n$, where $|x| = n/2$ and $|y| = n/2+1$, contributes $n$ to the average stretch, whereas

[1]The edge boundary of a subset $A \subset \{0,1\}^{n+1}$ is set of edges with one endpoint in $A$ and one outside $A$. The vertex boundary of $A$ is the set of vertices outside $A$ that are endpoints of boundary edges.

[2]Recall that a subset $A \subset \{0,1\}^{n+1}$ is *monotone* if $x \in A$ implies $y \in A$ for all $y \succeq x$.

0272-5428/14 $31.00 © 2014 IEEE
DOI 10.1109/FOCS.2014.17

all other edges contribute 1. The assertion then follows since $\Theta(1/\sqrt{n})$ fraction of the edges are of the first type. In fact, the maximum stretch of $\varphi$ is $n$, where

$$\mathsf{maxStretch}(\varphi) = \max_{\substack{x\in\{0,1\}^n \\ i\in[n]}} \mathsf{dist}(\varphi(x), \varphi(x+e_i)).$$

As far as we know, prior to our work this simple bijection achieved the best-known upper bound on the average stretch between $\{0,1\}^n$ and $\mathcal{B}_n$, and no non-trivial upper bounds (*i.e.*, sublinear) on maximum stretch were known. For a survey on metric embeddings of finite spaces see [Lin02]. In particular, a lot of research has been done on the question of embedding into the Boolean cube. For example, see [AB07], [HLN87] for some work on embeddings between random subsets of the Boolean cube, and [Gra88] for isometric embeddings of arbitrary graphs into the Boolean cube.

### A. Our Results

The main result of this paper is a strong *negative* answer to Problem I.1.

**Theorem 1** (Main theorem)**.** *For all even integers $n$, there exists a bijection $\psi\colon \{0,1\}^n \to \mathcal{B}_n$ with*

$$\mathsf{maxStretch}(\psi) \le 4$$

*and*

$$\mathsf{maxStretch}(\psi^{-1}) \le 5.$$

We believe that Theorem 1 will find other applications in theoretical computer science on top of the original motivation for studying the problem, as it highlights a surprising and counter-intuitive geometric resemblance between two well-studied objects in theory – the Boolean cube and the Hamming ball. In the language of metric geometry, Theorem 1 says that there is a bi-Lipschitz bijection between the two spaces.

**Corollary I.2** (A bi-Lipschitz bijection between $\{0,1\}^n$ and $\mathcal{B}_n$)**.** *For all even integers $n$, there exists a bijection $\psi\colon \{0,1\}^n \to \mathcal{B}_n$, such that for every $x \neq y \in \{0,1\}^n$ it holds that*

$$\frac{1}{5} \le \frac{\mathsf{dist}(\psi(x),\psi(y))}{\mathsf{dist}(x,y)} \le 4.$$

As a corollary from Theorem 1, we obtain that the subgraph of $\{0,1\}^{n+1}$ induced by the vertices of $\mathcal{B}_n$ is *bi-Lipschitz transitive*. Informally speaking, this says that any two points in $\mathcal{B}_n$ have roughly the same "view" – even the unique point with Hamming weight $n+1$ and the boundary points which have weight $n/2+1$. More formally,

**Corollary I.3** (The Hamming balls are uniformly bi-Lipschitz transitive)**.** *For all even integers $n$, and for every two vertices $x, y \in \mathcal{B}_n$ there is a bijection $f\colon \mathcal{B}_n \to \mathcal{B}_n$ such that $f(x) = y$, $f(y) = x$, and for every $z \neq w \in \mathcal{B}_n$, it holds that*

$$\frac{1}{20} \le \frac{\mathsf{dist}(f(z),f(w))}{\mathsf{dist}(z,w)} \le 20.$$

To see this, first note that $\mathcal{B}_n$ is a convex subset of $\{0,1\}^{n+1}$, and thus, the distances between vertices in $\mathcal{B}_n$ are the same as their distances as a subset of the cube. Now, for a given pair $x, y \in \mathcal{B}_n$, let $x' = \psi^{-1}(x)$ and $y' = \psi^{-1}(y)$, where $\psi$ is the function from Theorem 1. Define $f\colon \mathcal{B}_n \to \mathcal{B}_n$ as $f(z) = \psi(\psi^{-1}(z) \oplus x' \oplus y')$. It is easy to see that $f$ indeed satisfies the requirements of Corollary I.3.

*Approximating $\psi_i$:* We highlight another property of the bijection $\psi$ from our main theorem.

**Proposition I.4.** *Let $\psi$ be the function from Theorem 1. Then, for all $i \in [n]$ it holds that*

$$\Pr_x[\psi_i(x) = x_i] > 1 - O(1/\sqrt{n}).$$

That is, all but the last output bit of $\psi$ are essentially determined by a *single* input bit. In Section I-C we show that any bijection $\psi\colon \{0,1\}^n \to \mathcal{B}_n$ with constant average stretch satisfies a similar, though somewhat weaker, locality property.

*The complexity of $\psi$:* Since the original motivation for constructing $\psi$ comes from efficient sampling of distributions, Theorem 1 is of larger interest if the bijection $\psi$ (and $\psi^{-1}$) can be computed by low-level circuits.

**Proposition I.5.** *The bijections $\psi$ and $\psi^{-1}$ are computable in DLOGTIME-uniform $\mathbf{TC^0}$.*

**Remark I.6.** *In fact, we show that for all $i \in [n+1]$ there is an $\mathbf{NC^0}$-reduction from* majority *to $\psi_i$. That is, $\mathbf{TC^0}$ is the "correct" complexity of $\psi$, and in particular, $\psi$ is not in $\mathbf{AC^0}$. See Proposition III.1 and Remark III.2 for details.*

### B. The Complexity of Distributions

Lower bounds in circuit complexity are usually concerned with showing that a family of functions $\{f_n\colon \{0,1\}^n \to \{0,1\}\}_{n\in\mathbb{N}}$ cannot be computed by a family of circuits $\{C_n\}_{n\in\mathbb{N}}$ belonging to some natural class of circuits such as $\mathbf{AC^0}$ or $\mathbf{TC^0}$. Taking a broader interpretation of computation, it is often interesting to show that a class of circuits cannot perform a certain natural task beyond just computing a function.

One such natural task, introduced by Goldreich et al. [GGN10] and further advocated by Viola [Vio12], is that of sampling distributions. In this problem, for a given distribution $\mathcal{D}$ supported on $\{0,1\}^n$, we are looking for a function $f\colon \{0,1\}^m \to \{0,1\}^n$ that samples (or approximates) $\mathcal{D}$, that is, for a uniformly random $x \sim \{0,1\}^m$, the distribution $f(x)$ is equal (or close to) $\mathcal{D}$, and furthermore, each output bit $f_i$ of the function $f$ belongs to some low-level complexity class, such as $\mathbf{AC^0}$ or $\mathbf{TC^0}$.

As a concrete example, let $U_\oplus$ be the uniform distribution over the set $\{(x, \mathsf{parity}(x))\colon x \in \{0,1\}^{n-1}\} \subseteq \{0,1\}^n$. Note that although the parity function is not computable in $\mathbf{AC^0}$ (see [Has86] and references therein) there is a function $f\colon \{0,1\}^{n-1} \to \{0,1\}^n$ that samples $U_\oplus$, such that each output bit depends on only two input bits:

$$f(x_1,\ldots,x_{n-1}) = (x_1, x_1+x_2, x_2+x_3, \ldots, x_{n-1}).$$

Motivated by the foregoing somewhat surprising example, Viola [Vio12] suggested to replace parity above with majority

– the other notoriously hard function for $\mathbf{AC^0}$. The following two problems have been stated in [LV12].

**Problem I.7.** *Let $n \in \mathbb{N}$ be even. Does there exist a bijection $g\colon \{0,1\}^n \to \mathcal{B}_n$ such that each output bit of $g$ is computable in $\mathbf{AC^0}$?*

**Problem I.8.** *Let $n \in \mathbb{N}$ be odd. Does there exist a bijection $h\colon \{0,1\}^n \to \{(x, \mathsf{majority}(x))\colon x \in \{0,1\}^n\}$ such that each output bit of $h$ is computable in $\mathbf{AC^0}$?*

Note that a positive answer to Problem I.7 implies a positive answer to Problem I.8. Indeed, if $g\colon \{0,1\}^n \to \mathcal{B}_n$ is an embedding from Problem I.7, then the function $h\colon \{0,1\}^{n+1} \to \{0,1\}^{n+2}$ defined as

$$h(x_1,\ldots,x_{n+1}) = \begin{cases} g(x_1,\ldots,x_n)\circ 1 & x_{n+1}=1 \\ \mathsf{flip}(g(x_1,\ldots,x_n))\circ 0 & x_{n+1}=0 \end{cases}$$

gives an embedding for Problem I.8. Therefore, a negative answer to Problem I.8 implies a negative answer to Problem I.7. In the other direction, if a function $h\colon \{0,1\}^n \to \{(x, \mathsf{majority}(x))\colon x \in \{0,1\}^n\}$ gives a positive answer to Problem I.8, then the function $g\colon \{0,1\}^n \to \{0,1\}^n$ defined as[3]

$$g(x_1,\ldots,x_n) = \begin{cases} (h(x))_{[1,\ldots,n]} & (h(x))_{n+1}=1 \\ \mathsf{flip}(h(x))_{[1,\ldots,n]} & \text{otherwise} \end{cases}$$

samples $\mathcal{B}_{n-1}$ using input of length $n$, which almost[4] answers Problem I.8.

Problem I.1 was raised by Lovett and Viola [LV12] in an attempt to prove a lower bound for Problem I.7. A positive answer to Problem I.1 would imply a lower bound for Problem I.7, since by the result of [Bop97], any function $f\colon \{0,1\}^n \to \{0,1\}^{n+1}$ computable by a polynomial size Boolean circuit of constant depth has average stretch at most $\log^{O(1)}(n)$.

As we resolved Problem I.1 negatively, it seems that ideas beyond the sensitivity-based structural results of [Bop97] are required in order to resolve Problems I.7 and I.8.

On the positive side Viola [Vio12] showed an explicit $\mathbf{AC^0}$ circuit $C\colon \{0,1\}^{\mathrm{poly}(n)} \to \{0,1\}^n$ of size $\mathrm{poly}(n)$ whose output distribution has statistical distance $2^{-n}$ from the uniform distribution on $\{(x, \mathsf{majority}(x))\colon x \in \{0,1\}^n\}$. It is an open problem to improve either the input length or to reduce the statistical distance to zero. Theorem 1 gives a sign of hope in this direction.

For a lower bound Viola [Vio11] gave an explicit construction of a function $b\colon \{0,1\}^n \to \{0,1\}$ such that $(x, b(x))$ cannot be sampled by $\mathbf{AC^0}$ circuits. That is, it gives a negative answer to Problem I.8 if we replace majority by the function $b$. Nonetheless, we feel that it would still be interesting to give a negative answer to Problem I.8 for majority function, since this is a more natural function.

[3]We use the following notation: for a string $s \in \{0,1\}^n$ and integers $i \le j$ in $[n]$, the string $s_{[i,\ldots,j]}$ denotes the substring $s_i s_{i+1}\cdots s_j$.

[4]Problem I.8 asks for a function that takes $n-1$ bits as input.

### C. Low Stretch vs. Locality

Below we discuss some more aspects of the problem of sampling the uniform distribution over $\mathcal{B}_n$. In [Vio12] Viola showed that local samplers (say, $\mathbf{NC^0}$ circuits) cannot do the job, as they always have $\Omega(1)$ statistical distance from $\mathcal{B}_n$. Theorem 1 shows that low-stretch samplers can do the job, with no error, and even with a constant worst-case stretch. We find it somewhat unexpected because local samplers appear very similar to low-stretch samplers. Indeed, it is not hard to see that a local sampler has low average stretch, and the reverse direction follows from Friedgut's Junta Lemma [Fri98] as we explain next. However, the connection between low stretch samplers and local samplers only holds in the *average case*, in the sense that for every sampler $X$ of one type there is a sampler $Y$ of the other type that does the same with high probability over the input and the flipped bit. Theorem 1 states that the picture changes completely when moving to the *worst case* computation.

We now sketch a proof for the fact that low stretch sources are local in the average case. This is a direct consequence of Friedgut's Junta Lemma (a similar argument has appeared recently in the paper of Austin [Aus13], where he studies Bi-Lipschitz functions $F : [0,1]^N \to [0,1]^M$). Let $f\colon \{0,1\}^n \to \mathcal{B}_n$ be a bijection of constant average stretch. We think of $f$ as a vector of Boolean functions $\langle f_1,\ldots,f_{n+1}\rangle$, where $f_i(x)$ is the $i^{\text{th}}$ output bit of $f$ on input $x \in \{0,1\}^n$. Recall that the *total influence* of a Boolean function $f_i\colon \{0,1\}^n \to \{0,1\}$ is the quantity

$$\mathbf{Inf}[f_i] = \mathop{\mathbf{E}}_{x\sim\{0,1\}^n}\left[\#\{j\in[n]\colon f_i(x)\neq f_i(x+e_j)\}\right].$$

By linearity of expectation, we have that

$$\mathop{\mathbf{E}}_{i\sim[n+1]}[\mathbf{Inf}[f_i]] = O(1). \tag{2}$$

Next we recall Friedgut's Junta Lemma, which states that a Boolean function with constant total influence is well-approximated by another Boolean function that only depends on a *constant* number of input bits. More precisely,

**Friedgut's Junta Theorem** ([Fri98])**.** *Let $h\colon \{0,1\}^n \to \{0,1\}$ be a Boolean function. For every $\varepsilon > 0$ there exists a Boolean function $g\colon \{0,1\}^n \to \{0,1\}$ such that $g$ is a $2^{O(\mathbf{Inf}[h]/\varepsilon)}$-junta[5] and $\mathbf{Pr}[h(x)\neq g(x)] \le \varepsilon$.*

Combining Equation (2) with Friedgut's Junta Theorem, we see that for any constants $\delta,\varepsilon > 0$, all but a $\delta$-fraction of the $f_i$'s are $\varepsilon$-approximated by $O(1)$-juntas.

### D. Proof Overview

In this section we describe in high-level the proof of Theorem 1. A full proof is given in Section II. Let $n \in \mathbb{N}$ be an even integer. Our goal is to map $\{0,1\}^n$ to $\mathcal{B}_n$ in a way that the two endpoints of every edge in $\{0,1\}^n$ are mapped to close vertices in $\mathcal{B}_n$. The key building block we use is a classical

[5]Recall that a $k$-junta is a Boolean function that only depends on at most $k$ of its input bits.

partition of the vertices of $\{0,1\}^n$ to symmetric chains, due to De Bruijn, Tengbergen, and Kruyswijk [BvETK51], where a symmetric chain is a path $\{c_k, c_{k+1}, \ldots, c_{n-k}\}$ in $\{0,1\}^n$, such that each $c_i$ has Hamming weight $i$.

As a first step, we study the chains in the partition of De Bruijn et al. Roughly speaking, we show[6] that adjacent chains move closely to each other. More precisely, if two adjacent vertices $x$ and $y$ belong to different chains, then $x$ and $y$ have the same distance from the top of their respective chains, up to some additive constant. Moreover, the lengths of the two chains differ by at most some additive constant, and the $i^{\text{th}}$ vertex in one chain, when counting from the top, is $O(1)$-close to the $i^{\text{th}}$ vertex in the other chain (if such exists).

We now describe how to map $\{0,1\}^n$ to $\mathcal{B}_n$ based on the partition of De Bruijn et al. Consider a chain $c_k, c_{k+1}, \ldots, c_{n-k}$. Our mapping will "squeeze" the vertices to the top half of the cube while exploiting the extra dimension. In particular, every vertex will climb up its chain half the distance it has from the top, and then, the collision between two vertices is resolved by setting the extra last bit to 1 for the first vertex and to 0 for the second vertex. More precisely, the vertex $c_{n-k}$, which is at the top of its chain, is mapped to $c_{n-k} \circ 1$, while $c_{n-k-1}$ is mapped to $c_{n-k} \circ 0$. The third vertex from the top $c_{n-k-2}$ is mapped to $c_{n-k-1} \circ 1$ while $c_{n-k-3}$ is mapped to $c_{n-k-1} \circ 0$ and so on. The vertex $c_k$ at the bottom of the chain is mapped to $c_{n/2} \circ 1$, which is indeed in $\mathcal{B}_n$.

Consider now two adjacent vertices $x, y$ in $\{0,1\}^n$. By the above, these vertices reside in "close" chains with roughly the same length and have roughly the same distance from the top of their respective chains. Thus, in the climbing process, both $x$ and $y$ will be mapped to vertices that have roughly the same distance from the top of their respective chains, and hence, from the discussion above, their images will be $O(1)$-close.

## II. Proof of the Main Theorem

In this section we prove the main theorem. In Section II-A we describe the De Bruijn-Tengbergen-Kruyswijk partition. In Section II-B we define the mapping $\psi$ and prove basic facts about it. In Section II-C we give the proof for Theorem 1, omitting some technical details that can be found in Section II-D.

### A. The De Bruijn-Tengbergen-Kruyswijk Partition

**Definition II.1.** *Let $n$ be an even integer. A* symmetric chain *in $\{0,1\}^n$ is a sequence of vertices $C = \{c_k, c_{k+1}, \ldots, c_{n-k}\}$ such that $|c_i| = i$ for $i = k, k+1, \ldots, n-k$, and $\mathsf{dist}(c_i, c_{i+1}) = 1$ for $i = k, k+1, \ldots, n-k-1$. We say that a symmetric chain is* monotone *if it satisfies the following property: if $c_{i-1}$ and $c_i$ differ in the $j^{th}$ coordinate, and $c_i$ and $c_{i+1}$ differ in the $(j')^{th}$ coordinate, then $j < j'$.*

We shall represent a monotone symmetric chain as follows. Let $y \in \{0,1,\sqcup\}^n$ be such that $m = |\{i\colon y_i = \sqcup\}|$ satisfies $m \equiv n \pmod 2$, and let $k = (n-m)/2$. The monotone symmetric chain $C_y = \{c_k, c_{k+1}, \ldots, c_{n-k}\}$ is specified by $y$ as follows. For $i = k, k+1, \ldots, n-k$, the string $c_i$ is obtained by replacing the $m-(i-k)$ leftmost symbols $\sqcup$ of $y$ by 0 and the remaining $i-k$ symbols $\sqcup$ by 1. Note that $C_y$ is indeed a monotone symmetric chain.

De Bruijn, Tengbergen, and Kruyswijk [BvETK51] suggested a recursive algorithm that partitions $\{0,1\}^n$ to monotone symmetric chains. We will follow the presentation of the algorithm described in [vLW01] (see Problem 6E in Chapter 6). The algorithm gets as input a string $x \in \{0,1\}^n$, and computes a string $y \in \{0,1,\sqcup\}^n$ which encodes the monotone symmetric chain $C_y$ that contains $x$.

The algorithm is iterative. During the running of the algorithm, every coordinate of $x$ is either marked or unmarked, where we denote a marked 0 by $\hat{0}$ and a marked 1 by $\hat{1}$. In each step, the algorithm chooses a consecutive pair 10, marks it by $\hat{1}\hat{0}$, temporarily deletes it, and repeats the process. The algorithm halts once there is no such pair, that is the remaining string is of the form $00\ldots01\ldots11$. We call this stage of the algorithm the *marking stage*, and denote the marked string by $\mathsf{mark}(x) \in \{0,1,\hat{0},\hat{1}\}^n$. The string $y$ is then defined as follows: if the $i^{\text{th}}$ bit of $x$ was marked then $y_i = x_i$. Otherwise, $y_i = \sqcup$.

For example, consider the string $x = 01100110$. At the first iteration, the algorithm may mark the third and fourth bits to obtain $01\hat{1}\hat{0}0110$. Then, the second and fifth bits are marked $0\hat{1}\hat{1}\hat{0}\hat{0}110$. Lastly, the rightmost two bits are marked, and we obtain the marked string $\mathsf{mark}(x) = 0\hat{1}\hat{1}\hat{0}\hat{0}1\hat{1}\hat{0}$. Hence $y = \sqcup 1100 \sqcup 10$ and $C_y = \{\underline{0}1100\underline{0}10, \underline{0}1100\underline{1}10, \underline{1}1100\underline{1}10\}$.

Readily, the algorithm induces a partition of $\{0,1\}^n$ to monotone symmetric chains. We stress that although the algorithm has some degree of freedom when choosing a 10 pair out of, possibly, many pairs in a given iteration, the output of the algorithm, $y$, is independent of the specific choices that were made. That is, $y$ is a function of $x$, and does not depend on the specific order in which the algorithm performs the marking. This assertion can be proven easily by induction on $n$. As a consequence, we may choose the ordering of the 10 pairs as we wish. We will use this fact in the proof of Theorem 1.

### B. The Bijection $\psi$

We define the mapping $\psi$ as follows. Let $n \in \mathbb{N}$ be an even integer. For an input $x \in \{0,1\}^n$, let $C = \{c_k, c_{k+1}, \ldots, c_{n-k}\}$ be the symmetric chain from the partition of De Bruijn et al., that contains $x$. Let $j$ be the index such that $x = c_j$. Define

$$\psi(x) \stackrel{\text{def}}{=} \begin{cases} c_{\frac{(n-k)+j}{2}} \circ 1 & j \equiv (n-k) \pmod 2; \\ c_{\frac{(n-k)+j+1}{2}} \circ 0 & j \not\equiv (n-k) \pmod 2. \end{cases} \tag{3}$$

**Claim II.2.** *The mapping $\psi$ is a bijection from $\{0,1\}^n$ to $\mathcal{B}_n$.*

*Proof.* We first show that the range of $\psi$ is $\mathcal{B}_n$. Consider $x \in \{0,1\}^n$ and let $C = \{c_k, c_{k+1}, \ldots, c_{n-k}\}$ be the symmetric

[6] This is somewhat implicit in our proofs, and is mentioned here mainly in order to build an intuition.

chain that contains $x$. Suppose that $x = c_j$ for some $k \leq j \leq n-k$. If $j \equiv (n-k) \pmod 2$, then using the fact that $j \geq k$,

$$|\psi(x)| = \left| c_{\frac{n-k+j}{2}} \circ 1 \right| = \frac{n-k+j}{2} + 1 > \frac{n}{2}.$$

Otherwise, $j \not\equiv (n-k) \pmod 2$. Since $n$ is even, it follows that $j \not\equiv k \pmod 2$, and thus $j \geq k+1$. Hence,

$$|\psi(x)| = \left| c_{\frac{n-k+j+1}{2}} \circ 0 \right| > \frac{n}{2}.$$

In both cases $\psi(x) \in \mathcal{B}_n$.

We conclude the proof by describing the inverse mapping $\psi^{-1} \colon \mathcal{B}_n \to \{0,1\}^n$. For $z \in \mathcal{B}_n$, write $z = x \circ z_{n+1}$, where $x \in \{0,1\}^n$ and $z_{n+1}$ is the $(n+1)^{\text{st}}$ bit of $z$. Let $C = \{c_k, c_{k+1}, \ldots, c_{n-k}\}$ be the symmetric chain that contains $x$, and let $j$ be the index such that $x = c_j$ (note that $j \geq n/2$). Then,

$$\psi^{-1}(z) = \begin{cases} c_{2j-(n-k)} & \text{if } z_{n+1} = 1; \\ c_{2j-(n-k)-1} & \text{if } z_{n+1} = 0. \end{cases} \tag{4}$$

It is straightforward to verify that this is indeed the inverse mapping of $\psi$. □

In order to understand the mapping $\psi$ better, consider $x \in \{0,1\}^n$ and let $y \in \{0,1,\sqcup\}^n$ be the encoding of the chain that contains $x$. Note that if $1 \leq i_1 < i_2 < \cdots < i_t \leq n$ are the coordinates in which $y$ contains $\sqcup$, then there exists some $0 \leq \ell \leq t$ such that $x_{i_1} = \cdots = x_{i_\ell} = 0$ and $x_{i_{\ell+1}} = \cdots = x_{i_t} = 1$. That is, $x$ is located at the $(\ell+1)^{\text{st}}$ position of the chain $C_y$, when counting from the top. The function $\psi$ outputs the vertex located at the $(\lfloor \ell/2 \rfloor + 1)^{\text{st}}$ position in the chain, concatenated with 1 or 0, depending on the parity of $\ell$. In other words, we obtain $\psi(x)$ by keeping intact all the bits of $x$ in the coordinates other than $i_{\lfloor \ell/2 \rfloor + 1}, \ldots, i_\ell$, and by setting $\psi(x)_{i_{\lfloor \ell/2 \rfloor + 1}} = \cdots = \psi(x)_{i_\ell} = 1$. Then, we append 1 to the obtained string if $\ell$ is even, and append 0 otherwise.

The following claim is immediate from the definition of $\psi$.

**Claim II.3.** *Fix a string $x \in \{0,1\}^n$. Let $M \subseteq [n]$ be the set of marked coordinates in $\mathsf{mark}(x)$. Then,*

- *For every $i \in M$ it holds that $\psi(x)_i = x_i$.*
- *For every $j \in [n] \setminus M$, the $j^{th}$ coordinate of $\psi(x)$ does not depend on any of the bits $\{x_i\}_{i \in M}$.*

We are now ready to prove Theorem 1.

### C. Proof of Theorem 1

*Proof of Theorem 1.* We first show that $\mathsf{maxStretch}(\psi) \leq 4$. Take $x \in \{0,1\}^n$ and $i \in [n]$ such that $x_i = 0$. Our goal is to show that $\mathsf{dist}(\psi(x), \psi(x+e_i)) \leq 4$. As mentioned in Section II-A, the output of the algorithm on input $x$ is independent of the order in which the algorithm marks the 10 pairs. Therefore, given an input $x$, we may perform the marking stage in three steps:

1) Perform the marking stage on the prefix of $x$ of length $i-1$.
2) Perform the marking stage on the suffix of $x$ of length $n-i$.
3) Perform the marking stage on the resulting, partially marked, string.

Since $x$ and $x+e_i$ agree on all but the $i^{\text{th}}$ coordinate, the running of the marking stage in steps 1 and 2 yield the same marking. That is, prior to the third step the strings $x$ and $x+e_i$ have the same bits marked. Denote by $s \in \{0,1,\hat{0},\hat{1}\}^{i-1}$ and $t \in \{0,1,\hat{0},\hat{1}\}^{n-i}$ the two partially marked strings such that the resulted strings after the second step on inputs $x$ and $x+e_i$ are $s \circ 0 \circ t$ and $s \circ 1 \circ t$ respectively. Let us suppose for concreteness that the string $s$ contains $a$ unmarked zeros and $b$ unmarked ones, and the string $t$ contains $c$ unmarked zeros and $d$ unmarked ones. Recall that at the end of the marking stage, all unmarked zeros are to the left of all unmarked ones in both $s$ and $t$.

By Claim II.3, the only coordinates that may contribute to $\mathsf{dist}(\psi(x), \psi(x+e_i))$ are the unmarked coordinates prior to the third step, and so $\mathsf{dist}(\psi(x), \psi(x+e_i))$ is equal to

$$\mathsf{dist}(\psi(0^a 1^b \circ 0 \circ 0^c 1^d), \psi(0^a 1^b \circ 1 \circ 0^c 1^d)).^7$$

Therefore, it is enough to bound this from above by 4. At this point, it is fairly easy to be convinced that the right hand side is bounded by *some* constant. Proving that the constant is 4 is done by a somewhat tedious case analysis, according to the relations between $a, b, c$ and $d$. We defer the proof of the following claim to Section II-D.

**Claim II.4.** *For every $a, b, c, d \in \mathbb{N}$, we have*

$$\mathsf{dist}(\psi(0^a 1^b \circ 0 \circ 0^c 1^d), \psi(0^a 1^b \circ 1 \circ 0^c 1^d)) \leq 4.$$

This completes the proof for $\mathsf{maxStretch}(\psi) \leq 4$.

We now prove that $\mathsf{maxStretch}(\psi^{-1}) \leq 5$, where we use the description of $\psi^{-1}$ given in Equation (4). In order to bound $\mathsf{maxStretch}(\psi^{-1})$, let us fix an edge in $\mathcal{B}_n$, that is, take $z \in \mathcal{B}_n$ and $i \in [n+1]$ such that $z_i = 0$ and show that $\mathsf{dist}(\psi^{-1}(z), \psi^{-1}(z+e_i)) \leq 5$. By the proof of Claim II.2, if $i = n+1$ then $\psi^{-1}(z)$ and $\psi^{-1}(z+e_i)$ are consecutive vertices in some monotone symmetric chain, and thus $\mathsf{dist}(\psi^{-1}(z), \psi^{-1}(z+e_i)) = 1$.

Therefore, we shall assume henceforth that $i \neq n+1$. Let $z = x \circ z_{n+1}$ and $z + e_i = (x+e_i) \circ z_{n+1}$ for some $x \in \{0,1\}^n$ and $z_{n+1} \in \{0,1\}$. Similarly to the proof for $\mathsf{maxStretch}(\psi) \leq 4$, we perform the marking stage by first performing the marking stage on the prefix of $x$ of length $i-1$, then perform the marking stage on the suffix of $x$ of length $n-i$, and finally, perform the marking stage on the resulting, partially marked string. Denote by $s \in \{0,1,\hat{0},\hat{1}\}^{i-1}$ and $t \in \{0,1,\hat{0},\hat{1}\}^{n-i}$ the two partially marked strings such that the resulted strings after the second step on inputs $x$ and $x+e_i$ are $s \circ 0 \circ t$ and $s \circ 1 \circ t$ respectively. Suppose again for

[7] Note that $\psi$ is applied to inputs whose length is not necessarily $n$. However, for the sake of readability, we do not indicate the input length when applying $\psi$. In other words, $\psi$ is a shorthand for a family of functions $\{\psi_n\}_{n \in \mathbb{N}}$.

concreteness that the string $s$ contains $a$ unmarked zeros and $b$ unmarked ones, and the string $t$ contains $c$ unmarked zeros and $d$ unmarked ones.

By Claim II.3, the only coordinates that may contribute to $\mathsf{dist}(\psi^{-1}(z), \psi^{-1}(z+e_i))$ are the unmarked coordinates in $s$ and $t$, and so $\mathsf{dist}(\psi^{-1}(z), \psi^{-1}(z+e_i))$ is equal to

$$\mathsf{dist}(\psi^{-1}(0^a1^b \circ 0 \circ 0^c1^d \circ z_{n+1}), \psi^{-1}(0^a1^b \circ 1 \circ 0^c1^d \circ z_{n+1})).$$

Thus, it is enough to upper bound this by 5. We first note that $a+c+1 \leq b+d$. To see this recall that $|z| > n/2$ and $0^a1^b \circ 0 \circ 0^c1^d$ was obtain from $x = z_1 \dots z_n$ (that is, $z$ without its last bit $z_{n+1}$) by deleting the same number of zeros and ones.

**Claim II.5.** *For every $a, b, c, d \in \mathbb{N}$ such that $a+c+1 \leq b+d$, and for every $z_{n+1} \in \{0,1\}$ it holds that* $\mathsf{dist}(\psi^{-1}(0^a1^b \circ 0 \circ 0^c1^d \circ z_{n+1}), \psi^{-1}(0^a1^b \circ 1 \circ 0^c1^d \circ z_{n+1})) \leq 5$.

Therefore, by Claim II.5 we have $\mathsf{maxStretch}(\psi^{-1}) \leq 5$. This completes the proof of Theorem 1. □

### D. Proof of Claim II.4

We now return to the proof of Claim II.4. The proof of Claim II.5 is very similar, and is omitted it in this version.

*Proof of Claim II.4.* Let $w = 0^a1^b \circ 0 \circ 0^c1^d$ and $w' = 0^a1^b \circ 1 \circ 0^c1^d$. We prove the claim using the following case analysis. It will be convenient to introduce the function $\mathsf{even}\colon \mathbb{N} \to \{0,1\}$ defined as $\mathsf{even}(n) = 1$ if $n$ is even, and $\mathsf{even}(n) = 0$ otherwise.

*a) Case 1 ($\boldsymbol{b = c}$).:* In this case we have $w = 0^a \circ 1^b0^b \circ 01^d$ and $w' = 0^a1 \circ 1^b0^b \circ 1^d$. After the marking stage we get $\mathsf{mark}(w) = \underline{0}^a \circ \hat{1}^b\hat{0}^b \circ \underline{01}^d$ and $\mathsf{mark}(w') = \underline{0}^a\underline{1} \circ \hat{1}^b\hat{0}^b \circ \underline{1}^d$. Therefore,

$$\psi(w) = 0^{\lfloor \frac{a+1}{2} \rfloor}1^{a-\lfloor \frac{a+1}{2} \rfloor} \circ 1^b0^b \circ 1^{d+1} \circ \mathsf{even}(a+1)$$

and

$$\psi(w') = 0^{\lfloor \frac{a}{2} \rfloor}1^{\lceil \frac{a}{2} \rceil+1} \circ 1^b0^b \circ 1^d \circ \mathsf{even}(a).$$

By inspection, one can now easily verify that $\mathsf{dist}(\psi(w), \psi(w')) \leq 4$ in this case.

*b) Case 2 ($\boldsymbol{b > c}$).:* In this case we have $w = 0^a \circ 1^{b-c-1} \circ 1^{c+1}0^{c+1} \circ 1^d$ and $w' = 0^a \circ 1^{b-c+1} \circ 1^c0^c \circ 1^d$. After the marking stage we get $\mathsf{mark}(w) = \underline{0}^a\underline{1}^{b-c-1} \circ \hat{1}^{c+1}\hat{0}^{c+1} \circ \underline{1}^d$ and $\mathsf{mark}(w') = \underline{0}^a\underline{1}^{b-c+1} \circ \hat{1}^c\hat{0}^c \circ \underline{1}^d$. Therefore,

$$\psi(w) = 0^{\lfloor \frac{a}{2} \rfloor}1^{\lceil \frac{a}{2} \rceil+b-c-1} \circ 1^{c+1}0^{c+1} \circ 1^d \circ \mathsf{even}(a)$$

and

$$\psi(w') = 0^{\lfloor \frac{a}{2} \rfloor}1^{\lceil \frac{a}{2} \rceil+b-c+1} \circ 1^c0^c \circ 1^d \circ \mathsf{even}(a).$$

Therefore, in this case, $\mathsf{dist}(\psi(w), \psi(w')) \leq 1$.

*c) Case 3 ($\boldsymbol{b < c}$ and $\boldsymbol{a \geq c-b}$).:* In this case we have $w = 0^a \circ 1^b0^b \circ 0^{c-b+1} \circ 1^d$ and $w' = 0^a \circ 1^{b+1}1^{b+1} \circ 0^{c-b-1} \circ 1^d$. After the marking stage we get $\mathsf{mark}(w) = \underline{0}^a \circ \hat{1}^b\hat{0}^b \circ \underline{0}^{c-b+1}\underline{1}^d$ and $\mathsf{mark}(w') = \underline{0}^a \circ \hat{1}^{b+1}\hat{0}^{b+1} \circ \underline{0}^{c-b-1}\underline{1}^d$. By the assumption that $a \geq c-b$ we have $a \geq \lfloor \frac{a+c-b+1}{2} \rfloor$, and so

$$\psi(w) = 0^{\lfloor \ell/2 \rfloor}1^{a-\lfloor \ell/2 \rfloor} \circ 1^b0^b \circ 1^{d+c-b+1} \circ \mathsf{even}(\ell),$$

where $\ell = a+c-b+1$, and

$$\psi(w') = 0^{\lfloor \ell'/2 \rfloor}1^{a-\lfloor \ell'/2 \rfloor} \circ 1^{b+1}0^{b+1} \circ 1^{d+c-b-1} \circ \mathsf{even}(\ell'),$$

where $\ell' = a+c-b-1$. Therefore, by inspection we have $\mathsf{dist}(\psi(w), \psi(w')) \leq 4$ for this case.

*d) Case 4 ($\boldsymbol{b < c}$ and $\boldsymbol{a < c-b}$).:* Just like in the previous case, we have $\mathsf{mark}(w) = \underline{0}^a \circ \hat{1}^b\hat{0}^b \circ \underline{0}^{c-b+1}\underline{1}^d$ and $\mathsf{mark}(w') = \underline{0}^a \circ \hat{1}^{b+1}\hat{0}^{b+1} \circ \underline{0}^{c-b-1}\underline{1}^d$. By the assumption that $a < c-b$, we have $a \leq \lfloor \frac{a+c-b-1}{2} \rfloor$, and so

$$\psi(w) = 0^a \circ 1^b0^b \circ 0^{\lfloor \ell/2 \rfloor - a}1^{c-b+1+d-\lfloor \ell/2 \rfloor+a} \circ \mathsf{even}(\ell),$$

where $\ell = a+c-b+1$, and

$$\psi(w') = 0^a \circ 1^{b+1}0^{b+1} \circ 0^{\lfloor \ell'/2 \rfloor - a}1^{c-b-1+d-\lfloor \ell'/2 \rfloor+a} \circ \mathsf{even}(\ell').$$

where $\ell' = a+c-b-1$. Therefore, in this case, $\mathsf{dist}(\psi(w), \psi(w')) \leq 2$.

This completes the proof of Claim II.4. □

## III. The Mapping $\psi$ is Computable in DLOGTIME-uniform $\mathbf{TC^0}$

In this section we analyze the complexity of the bijection $\psi$ described in the proof of Theorem 1. We first claim that each output bit of $\psi$ (and of $\psi^{-1}$) can be computed in DLOGTIME-uniform $\mathbf{TC^0}$. In Proposition III.1 and in the remark following it, we show that indeed $\mathbf{TC^0}$ is the "correct" class for $\psi$.

*e) Proposition I.5 (restated).: The bijections $\psi$ and $\psi^{-1}$ are computable in DLOGTIME-uniform $\mathbf{TC^0}$.*

We prove the proposition only for $\psi$. The proof for $\psi^{-1}$ is very similar, and we omit it.

*Proof.* We divide the proof into two steps. First we show that the marking stage can be implemented in $\mathbf{TC^0}$. Then, given the marking of an input, we show how to compute $\psi$ in $\mathbf{TC^0}$. Both steps can be easily seen to be DLOGTIME-uniform.

Throughout the proof, the output of the marking stage is represented by two bits for each coordinate, encoding a symbol in $\{0, 1, \hat{0}, \hat{1}\}$, where one bit represents the Boolean symbol, and the other indicates whether the coordinate is marked or not.

*f) Implementing the marking stage.:* Let $x \in \{0,1\}^n$. In order to implement the marking stage in $\mathbf{TC^0}$, we observe that the $i^{\text{th}}$ coordinate in $x$ is marked if and only if there are coordinates $s_i \leq i \leq e_i$ such that

1) The number of ones in $x_{[s_i,\dots,e_i]}$ is equal to the number of zeros in $x_{[s_i,\dots,e_i]}$.
2) For every $k \in \{s_i, , \dots, e_i\}$, the number of ones in the prefix $x_{[s_i,\dots,k]}$ is greater or equal to the number of zeros in $x_{[s_i,\dots,k]}$.

Fix $i \in [n]$ and fix $s_i, e_i \in [n]$ such that $s_i \leq i \leq e_i$. Thinking of the bit 1 as '(' and 0 as ')', the above two conditions are equivalent to checking whether the string $x_{[s_i,\ldots,e_i]}$ of parentheses is balanced, or in other words, deciding whether $x_{[s_i,\ldots,e_i]}$ is in Dyck language. It is well-known that Dyck language can be recognized in $\mathbf{TC^0}$ [Lyn77]. In fact, it is not hard to show that deciding whether a string of length $m$ is in Dyck language can be carried out by a DLOGTIME-uniform $\mathbf{TC^0}$ circuit with size $O(m)$.

Now, for each $i \in [n]$, we go over all choices for $s_i, e_i$ in parallel, and take the OR of the $O(n^2)$ results. Thus, for each $i \in [n]$, there is a DLOGTIME-uniform $\mathbf{TC^0}$ circuit with size $O(n^3)$ that decides whether the $i^{\text{th}}$ coordinate is marked or not.

*g) Computing $\psi(x)$ from $\mathsf{mark}(x)$::* In order to compute $\psi(x)$, let $\mathsf{mark}(x) \in \{0, 1, \hat{0}, \hat{1}\}^n$ be the marking of $x$. Since every marked coordinate will remain unchanged, we need to consider only of the unmarked coordinates. Recall also that the unmarked bits form a sequence of zeros followed by a sequence of ones. That is, if we ignore the marked coordinates, then we get a string of the form $0^a 1^b$ for some $a = a(x), b = b(x)$, and the output should be $0^{\lfloor \frac{a}{2} \rfloor} 1^{\lceil \frac{a}{2} \rceil + b} \circ \mathsf{even}(a)$ (recall that $\mathsf{even}(a) = 1$ if $a$ is even, and $\mathsf{even}(a) = 0$ otherwise). This can be implemented as follows.

1) Let $a$ be the number of unmarked zeros in $\mathsf{mark}(x)$.
2) For each $i \in [n]$, let $u_i = u_i(x)$ be the number of unmarked coordinates among $\{1, \ldots, i\}$.
3) For all unmarked coordinates $i \in [n]$, if $2u_i < a$, then set the $i^{\text{th}}$ bit of the output to 0. Otherwise, set the $i^{\text{th}}$ bit to 1.
4) Set the $(n+1)^{\text{st}}$ bit of the output to $\mathsf{even}(a)$.

It is easy to verify that given $\mathsf{mark}(x)$, checking whether the inequality $2u_i < a$ holds can be done in $\mathbf{TC^0}$, and so the entire second step can be carried out by a $\mathbf{TC^0}$ circuit. □

We remark that the bijection $\psi$ cannot be computed in $\mathbf{AC^0}$. For example, we prove that the first output bit of $\psi$ cannot be computed in $\mathbf{AC^0}$.

**Proposition III.1.** *The function* $\mathsf{majority}$ *is* $\mathbf{NC^0}$*-reducible to* $\psi_1$*, i.e.,* $\mathsf{majority} \leq_{\mathbf{NC^0}} \psi_1$*. In particular* $\psi_1 \notin \mathbf{AC^0}$.

*Proof.* We first note that $\psi_1(x) = 0$ if and only if $x_1 = 0$ and $\mathsf{mark}(x)$ contains at least two unmarked zeros. For odd $n$, we construct a reduction $r\colon \{0,1\}^n \to \{0,1\}^{3n+1}$ that for input $x \in \{0,1\}^n$ outputs a string $r(x) \in \{0,1\}^{3n+1}$ as follows. Let $x' \in \{0,1\}^{2n}$ be the string obtained from $x$ by replacing each 0 of $x$ with 10, and by replacing each 1 of $x$ with 00. Define $r(x) = 0 \circ 1^n \circ x'$. For example, if $x = 01101$, then $x' = 10\circ00\circ00\circ10\circ00$, and $r(x) = 0\circ1^5\circ10\circ00\circ00\circ10\circ00$. By the definition of $r$, it is clear that each bit of $r(x)$ depends on at most one bit of $x$. It is straightforward to check that $\mathsf{majority}(x) = \psi_1(r(x))$, and the assertion, then, follows. □

**Remark III.2.** *Note that the reduction above also gives* $\mathsf{majority} \leq_{\mathbf{NC^0}} \psi_{n+1}$*. A similar proof also shows that* $\mathsf{majority} \leq_{\mathbf{NC^0}} \psi_i$ *for all* $i \in [n+1]$.

## IV. All But the Last Output Bit Depend Essentially on a Single Input Bit

In this section we prove Proposition I.4. We recall it here for convenience.

*h) Proposition I.4 (restated).: For all $i \in [n]$ it holds that*

$$\Pr_x[\psi_i(x) = x_i] > 1 - O(1/\sqrt{n}).$$

Before proving the proposition, we need to further study the structure of the De Bruijn-Tengbergen-Kruyswijk partition described in Section II-A. We start with the following claim.

**Claim IV.1.** *Let $n$ be an integer, and let $\mathcal{P}$ be a partition of $\{0,1\}^n$ into symmetric chains. For every $1 \leq t \leq n+1$, let $M_t$ be the number of symmetric chains of length $t$ in $\mathcal{P}$. Then,*

$$M_t = \begin{cases} \binom{n}{\frac{n-t+1}{2}} - \binom{n}{\frac{n-t-1}{2}} & t \not\equiv n \pmod 2; \\ 0 & \text{otherwise.} \end{cases}$$

*Proof.* Note first that if $C = \{c_k, c_{k+1}, \ldots, c_{n-k}\}$ is a symmetric chain, then its length is $n - 2k + 1$. In particular, this implies that there are no symmetric chains of length $t$ where $t \equiv n \pmod 2$, and hence $M_t = 0$ for such $t$.

Next, we prove the claim for $t \not\equiv n \pmod 2$. This is done by backward induction on $t$. For $t = n+1$ we clearly have a unique symmetric chain starting at $0^n$ and ending at $1^n$, and hence $M_{n+1} = 1$, as claimed.

Before actually doing the induction step, let us consider the next case, namely, $t = n-1$. Note that only one of the vertices of Hamming weight 1 is contained in the unique chain of length $n+1$, and so, since distinct vertices with equal weight are contained in distinct symmetric chains, there are $n-1$ chains with bottom vertex of Hamming weight 1. Therefore $M_{n-1} = n - 1$, as claimed.

For the general induction step, suppose that the claim holds for all $t'$ larger than $t$. We prove the assertion for $t \not\equiv n \pmod 2$. Every symmetric chain of length $t$ must be of the form $C = \{c_k, c_{k+1}, \ldots, c_{n-k}\}$, where $k = \frac{n-t+1}{2}$. Since the chains of length greater than $t$ are disjoint, and each contains a vertex with Hamming weight $k$, it follows that the number of vertices with Hamming weight $k$ that are contained in chains of length greater than $t$ is $\sum_{t'>t} M_{t'} = \binom{n}{k-1}$. The remaining $\binom{n}{k} - \binom{n}{k-1}$ vertices must be contained in chains of length $t$, and so, since distinct vertices of Hamming weight $k$ are contained in distinct symmetric chains, it follows that there are $\binom{n}{k} - \binom{n}{k-1}$ chains of length $t$. □

The following corollary is immediate from the observation that any $x \in \{0,1\}^n$ such that $\mathsf{mark}(x)$ contains exactly $a$ unmarked zeros and $b$ unmarked ones is contained in a unique chain of length $a+b+1$ in the De Bruijn-Tengbergen-Kruyswijk partition.

**Corollary IV.2.** *Let $n, a, b \in \mathbb{N}$ such that $a+b \equiv n \pmod 2$, and $a+b \leq n$. Then,*

1) *The number of $x \in \{0,1\}^n$ such that $\mathsf{mark}(x)$ contains exactly $a$ unmarked zeros and $b$ unmarked ones is $\binom{n}{\frac{n-a-b}{2}} - \binom{n}{\frac{n-a-b-2}{2}}$.*

2) *The number of $x \in \{0,1\}^n$ such that* $\mathsf{mark}(x)$ *contains exactly $a$ unmarked zeros (and any number of unmarked ones) is* $\binom{n}{\lfloor \frac{n-a}{2} \rfloor}$.

We are now ready to prove Proposition I.4.

*Proof of Proposition I.4.* Let $x \in \{0,1\}^n$, and let $\mathsf{mark}(x)$ be its marking. Suppose that the unmarked coordinates in $\mathsf{mark}(x)$ are $i_1 < i_2 < \cdots < i_t$, and let $0 \leq \ell \leq t$ be such that $x_{i_1} = \cdots = x_{i_\ell} = 0$ and $x_{i_{\ell+1}} = \cdots = x_{i_t} = 1$. Note that $\psi_i(x) \neq x_i$ if and only if the $i^{\text{th}}$ coordinate is unmarked in $\mathsf{mark}(x)$ and $i = i_j$ for some $j \in \{\lfloor \frac{\ell}{2} \rfloor + 1, \ldots, \ell\}$.

As in the proof of Theorem 1, it will be convenient to perform the following partial marking of $x$. First perform the marking stage on the prefix of $x$ of length $i-1$, and denote the resulting string by $s \in \{0,1,\hat{0},\hat{1}\}^{i-1}$. Then, perform the marking stage on the suffix of $x$ of length $n-i$, and denote the result string by $t \in \{0,1,\hat{0},\hat{1}\}^{n-i}$. Suppose for concreteness that the string $s$ contains $a$ unmarked zeros and $b$ unmarked ones, and the string $t$ contains $c$ unmarked zeros and $d$ unmarked ones. By the definition of $\psi$ we have $\psi_i(x) \neq x_i$ if and only if $x_i = 0$, $b = 0$ and $a \geq c$. Therefore,

$$\begin{aligned}\mathbf{Pr}[\psi_i(x) \neq x_i] &= \mathbf{Pr}[x_i = 0] \cdot \mathbf{Pr}[b = 0, a \geq c] \\ &= \frac{1}{2} \sum_{k=0}^{n-i} \sum_{j=k}^{i} \mathbf{Pr}[a = j, b = 0, c = k].\end{aligned}$$

Note that since each bit of $x$ is chosen independently, the partially marked strings $s, t$ and the bit $x_i$ are also independent, and so

$$\mathbf{Pr}[a = j, b = 0, c = k] = \mathbf{Pr}[a = j, b = 0]\,\mathbf{Pr}[c = k]$$

for all $j$ and $k$. Next we compute each of $\mathbf{Pr}[a = j, b = 0]$ and $\mathbf{Pr}[c = k]$ independently. By Corollary IV.2, for $j \not\equiv i \pmod 2$ we have

$$\mathbf{Pr}[a = j, b = 0] = \frac{1}{2^{i-1}} \cdot \left( \binom{i-1}{\frac{i-j-1}{2}} - \binom{i-1}{\frac{i-j-3}{2}} \right),$$

and

$$\mathbf{Pr}[c = k] = \frac{1}{2^{n-i}} \cdot \binom{n-i}{\lfloor \frac{n-i-k}{2} \rfloor}.$$

Therefore, for every $k \leq i$ we have

$$\sum_{j=k}^{i} \mathbf{Pr}[a = j, b = 0] = \frac{1}{2^{i-1}} \cdot \binom{i-1}{\lfloor \frac{i-k-1}{2} \rfloor},$$

and so

$$\mathbf{Pr}[\psi_i(x) \neq x_i] \leq \frac{1}{2^{n+1}} \cdot \sum_{k=0}^{\min(i,n-i)} \binom{n-i}{\lfloor \frac{n-i-k}{2} \rfloor} \binom{i-1}{\lfloor \frac{i-k-1}{2} \rfloor}.$$

Let us assume that $i \geq n/2$ (the case of $i < n/2$ is handled similarly). Then, using the fact that $\binom{i-1}{\lfloor \frac{i-k-1}{2} \rfloor} \leq O(2^i/\sqrt{i})$ for all $k$, we have

$$\mathbf{Pr}[\psi_i(x) \neq x_i] = O\left(\frac{1}{\sqrt{i}}\right) \cdot \frac{1}{2^{n-i}} \cdot \sum_{k=0}^{n-i} \binom{n-i}{\lfloor \frac{n-i-k}{2} \rfloor}.$$

By the identity

$$\sum_{k=0}^{n-i} \binom{n-i}{\lfloor \frac{n-i-k}{2} \rfloor} = \sum_{j=0}^{n-i} \binom{n-i}{j} = 2^{n-i},$$

we get $\mathbf{Pr}[\psi_i(x) \neq x_i] = O(1/\sqrt{i})$, and so, since we assumed that $i \geq n/2$ we get that $\mathbf{Pr}[\psi_i(x) \neq x_i] = O(1/\sqrt{n})$, as required. □

## V. Concluding Remarks and Open Problems

*Bi-Lipschitz bijection between balanced halfspaces.:* Let $a_0, \ldots, a_n \in \mathbb{R}$. The halfspace determined by the $a_i$'s is the set of all points $(x_1, \ldots, x_n) \in \{-1,1\}^n$ such that $a_0 + a_1 x_1 + \cdots + a_n x_n \geq 0$.[8] A balanced halfspace is a halfspace with $a_0 = 0$. The Boolean cube $\{-1,1\}^n$ embedded in the natural way in $\{-1,1\}^{n+1}$ and the Hamming ball $\{x \in \{-1,1\}^{n+1} : x_1 + \cdots + x_{n+1} \geq 0\}$ are two examples of balanced halfspaces. We showed a bi-Lipschitz bijection between them. It is therefore natural to ask the following question.

**Problem V.1.** *Is there a bi-Lipschitz bijection between* any *two balanced halfspaces? Or even a bijection with constant average stretch from the Boolean cube $\{-1,1\}^n$ to any balanced halfspace in $\{-1,1\}^{n+1}$?*

In functions terminology, the Boolean cube $\{-1,1\}^n$ embedded in $\{-1,1\}^{n+1}$ is indicated by the dictator function, while the Hamming ball is indicated by the majority function. Problem V.1 refers more generally to linear threshold functions. One attempt at solving Problem V.1 positively, would be to generalize the partition of De Bruijn et al. to general halfspaces.

Besides being a natural problem, a positive solution to Problem V.1 may have implications to fully polynomial approximation scheme for counting solutions of the 0-1 knapsack problem [MS04].

Another interesting problem, inspired by Corollary I.3, is the following.

**Problem V.2.** *Is it true that any halfspace is bi-Lipschitz transitive?*

*Bi-Lipschitz bijection of the hypercube mapping the half cube to the Hamming ball:* The following problem has been suggested to us by Daniel Varga. It asks whether the bijection given in Theorem 1 can be strengthened in the following way.

**Problem V.3.** *Let $n$ be even. Is there a bi-Lipschitz bijection $f : \{0,1\}^{n+1} \to \{0,1\}^{n+1}$ that maps the half cube to the Hamming ball? That is, for all $x \in \{0,1\}^{n+1}$ such that $x_1 = 1$ the bijection satisfies $f(x) \in \mathcal{B}_n$.*

[8] The $\{-1,1\}^n$ representation of the Boolean cube is more natural in the context of halfspaces.

*Tightness of the stretch from the Boolean cube to the Hamming ball:* One may ask whether the constants 4 and 5 in Theorem 1 are tight. By a slight variation on the proof of Theorem 1, we can show that there exists a bijection $\phi\colon \{0,1\}^n \to \mathcal{B}_n$ with $\mathsf{maxStretch}(\phi) \leq 3$, improving on Theorem 1 in this respect. However, the maximum stretch of $\phi^{-1}$ is unbounded.

**Theorem 2.** *For all even integers $n$, define the bijection $\phi\colon \{0,1\}^n \to \mathcal{B}_n$ as follows. Let $x \in \{0,1\}^n$, and let $C = \{c_k, c_{k+1}, \ldots, c_{n-k}\}$ be the symmetric chain from the partition of De Bruijn et al., that contains $x$. Let $j$ be the index such that $x = c_j$. Define,*

$$\phi(x) \stackrel{\text{def}}{=} \begin{cases} c_{n-j} \circ 1 & j \leq n/2; \\ c_j \circ 0 & \textit{otherwise}. \end{cases}$$

*Then,* $\mathsf{maxStretch}(\phi) = 3$ *and* $\mathsf{avgStretch}(\phi^{-1}) = 2 + o(1)$.

The proof of Theorem 2 is similar to the proof of Theorem 1, and thus we omit it. One can easily see that any bijection $f\colon \{0,1\}^n \to \mathcal{B}_n$ has maximum stretch at least 2. Indeed, let $y = f(x) \in \mathcal{B}_n$ be a point with Hamming weight $n/2+1$. Then $y$ has only $n/2$ neighbors in $\mathcal{B}_n$, which cannot accommodate all $n$ neighbors of $x \in \{0,1\}^n$. We do not know whether the stretch 3 of $\phi$ in Theorem 2 is tight or not, and leave it as an open problem. What is the smallest possible stretch of a bijection from $\mathcal{B}_n$ to $\{0,1\}^n$? Are the constants 4 and 5 optimal if one considers only bi-Lipschitz bijections? Is the constant 20 in Corollary I.3 optimal?

*Lower bounds on average and maximum stretch:*

**Problem V.4.** *Exhibit an explicit subset $A \subset \{0,1\}^{n+1}$ of density $1/2$ such that any bijection $f\colon \{0,1\}^n \to A$ has $\mathsf{avgStretch}(f) = \omega(1)$, or prove that no such subset exists.*

As a concrete candidate, we suggest to consider sets $A = \{x\colon f(x) = 1\}$, where $f$ is a monotone noise-sensitive function (e.g., Tribes[9] or Recursive-Majority-of-Three). A sufficiently strong positive answer to this question would imply a lower bound for sampling the uniform distribution on $A$ by low-level complexity classes.

*Bijections from the Gale-Shapley algorithm for the stable marriage problem:* Let $A, B$ be two subsets of $\{0,1\}^{n+1}$ with density $1/2$. Consider the Gale-Shapley algorithm for the stable marriage problem, where each vertex $v \in A$ ranks all the vertices in $B$ according to their distance to $v$ (breaking ties according some rule). What can be said about the average stretch of the bijection obtained from this algorithm? Two interesting settings are (1) $A = \{0,1\}^n, B = \mathcal{B}_n$ and (2) $A, B$ are random subsets of $\{0,1\}^n$ of density $1/2$. For related work in this direction see Holroyd [Hol11]. Another natural bijection to consider, suggested to us by Avishay Tal, is the one induced by the Hungarian method for the assignment problem [Kuh55].


## ACKNOWLEDGEMENT

We thank Li-Yang Tan for introducing us Problem I.1, and for helpful discussions. We thank Ehud Friedgut for suggesting to use the De Bruijn–Tengbergen–Kruyswijk partition, which turned out to be the key step in the proof of Theorem 1. We also thank Emanuele Viola for referring us to [Vio11]. Lastly, we thank the anonymous referees for their helpful comments. In particular we are thankful to the referee who shed light on the relation between local and low stretch sources. Section I-C is based on his/her insights. Itai Benjamini would also like to thank Microsoft Research New England, where this research was started. Gil Cohen's research is supported by Israel Science Foundation (ISF) grant. Igor Shinkar's research is supported by ERC grant number 239985.


[9]We note that Tribes has density close to $1/2$.

## REFERENCES


[AB07] O. Angel and I. Benjamini. A phase transition for the metric distortion of percolation on the hypercube. *Combinatorica*, 27(6):645–658, 2007.

[Aus13] T. Austin. On the failure of concentration for the $\ell_\infty$-ball. 2013. http://arxiv.org/abs/1309.3315.

[Bop97] R. Boppana. The average sensitivity of bounded-depth circuits. *Information Processing Letters*, 63(5):257–261, 1997.

[BvETK51] N. G. De Bruijn, C. van Ebbenhorst Tengbergen, and D. Kruyswijk. On the set of divisors of a number. *Nieuw Arch. Wiskunde (2)*, 23:191–193, 1951.

[Fri98] E. Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):27–35, 1998.

[GGN10] O. Goldreich, S. Goldwasser, and Nussboim. On the implementation of huge random objects. *SIAM Journal on Computing*, 39(7):2761–2822, 2010.

[Gra88] R. L. Graham. Isometric embeddings of graphs. *Selected Topics in Graph Theory*, 3:133–150, 1988.

[Har66] L. H. Harper. Optimal numbering and isoperimetric problems on graphs. *Journal of Combinatorial Theory*, (1):385–393, 1966.

[Har76] S. Hart. A note on the edges of the $n$-cube. *Discrete Mathamatics*, 14(2):157–163, 1976.

[Has86] J. Hastad. Almost optimal lower bounds for small depth circuits. In *Proceedings of the eighteenth annual ACM Symposium on Theory of Computing*, pages 6–20. ACM, 1986.

[HLN87] J. Hastad, T. Leighton, and M. Newman. Reconfiguring a hypercube in the presence of faults. In *Proceedings of the nineteenth annual ACM Symposium on Theory of Computing*, pages 274–284, 1987.

[Hol11] A. E. Holroyd. Geometric properties of poisson matchings. *Probability Theory and Related Fields*, 150(3–4):511–527, 2011.

[Kuh55] H. W. Kuhn. The Hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2):83–97, 1955.

[Lin02] N. Linial. Finite metric spaces - combinatorics, geometry and algorithms. In *Proceedings of the International Congress of Mathematicians III*, pages 573–586, 2002.

[LV12] S. Lovett and E. Viola. Bounded-depth circuits cannot sample good codes. *Computational Complexity*, 21(2):245–266, 2012.

[Lyn77] N. Lynch. Log space recognition and translation of parenthesis languages. *Journal of the ACM*, 24(4):583–590, 1977.

[MS04] B. Morris and A. Sinclair. Random walks on truncated cubes and sampling 0-1 knapsack solutions. *SIAM Journal on Computing*, 34(1):195–226, 2004.

[Vio11] E. Viola. Extractors for circuit sources. In *IEEE 52nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 220–229. IEEE, 2011.

[Vio12] E. Viola. The complexity of distributions. *SIAM Journal on Computing*, 41(1):191–218, 2012.

[vLW01] J. H. van Lint and R.M. Wilson. *A Course in Combinatorics*. Cambridge University Press, Cambridge, 2001.