

# A Randomized, $O(\log w)$ -Depth 2-Smoothing Network\*

Marios Mavronicolas  
Department of Computer Science  
University of Cyprus  
CY-1678 Nicosia, Cyprus  
mavronic@cs.ucy.ac.cy

Thomas Sauerwald  
International Computer Science Institute (ICSI)  
1947 Center Street, Suite 600  
Berkeley, CA 94704, USA  
sauerwal@upb.de

## ABSTRACT

A  $K$ -smoothing network is a distributed, low-contention data structure where *tokens* arrive arbitrarily on  $w$  input wires and reach  $w$  output wires via their completely asynchronous propagation through the network. The maximum discrepancy among the numbers of tokens arriving at the output wires, called *smoothness*, is at most  $K$ . It has been a long-standing open problem to construct a  $K$ -smoothing network with (i) optimal  $K$ , (ii) optimal  $\Theta(\lg w)$  depth (called *small-depth*), (iii) no use of the AKS sorting network, and (iv) no reliance on global initialization.

In this work, we present a very simple, *randomized* network which meets all four desiderata:

- It is the cascade of a reasonably small number (about 150) of copies of the simple *block network* [6]; hence, it is small-depth and does not use the AKS sorting network.
- It achieves smoothness  $K = 2$ ; hence, it is optimal with respect to smoothness due to a recent improbability result about randomized, small-depth, 1-smoothing networks from [14].
- The network is *randomized*: each balancer is oriented independently and uniformly at random, thus requiring no global initialization.

Cascaded before the  $\Theta(\lg w)$ -depth 2-counter network due to Klugerman and Plaxton [13], which does use the AKS sorting network as a building block, our 2-smoothing network yields a new, *randomized* counting network with depth  $\Theta(\lg w)$ . The new network is a much simpler alternative to the classical, small-depth counting networks from [12, 13].

---

\*This work has been partially supported by the IST Program of the European Union under contract number 15964 (AEOLUS) and by the German Academic Exchange Service (DAAD)

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SPAA'09, August 11–13, 2009, Calgary, Alberta, Canada.  
Copyright 2009 ACM 978-1-60558-606-9/09/08 ...\$5.00.

## Categories and Subject Descriptors

G.2.1 [Mathematics of Computing]: Combinatorics—*Combinatorial Algorithms*; G.3 [Probability and Statistics]: Stochastic Processes; C.2.4 [Distributed Systems]: Distributed Applications—*Load Balancing*

## 1. INTRODUCTION

*Smoothing networks* (together with *ceibling counting networks* [4]) have been studied extensively in Distributed Computing Theory since their introduction in the seminal paper by Aspnes, Herlihy and Shavit [4]— see, e.g., [1, 2, 4, 8, 9, 11, 12, 13, 14]. Smoothing networks offer a modern, wait-free approach for important applications in asynchronous concurrent computing, such as *load balancing* and *producers-consumers*, which alleviates the latency due to contention.

A smoothing network [4] is a distributed data structure which acyclically interconnects *balancers* and *wires*. A balancer typically receives *tokens* on its two *input wires* and forwards them out to its two *output wires*, called *top* and *bottom*, in a round-robin fashion. A token represents a request for a service by a process. Each token arrives at one of the network's  $w$  *input wires*, propagates asynchronously through the network and exits through one of the network's *output wires*. The dispersion through the network helps reducing the contention and, thereby, the latency due to contention. On the other hand, the network should have *small depth* in order to reduce the latency due to propagation. We are interested in the pairwise difference of the numbers of tokens exiting on the network's output wires; the maximum (over all executions) of these differences is called *smoothness*. A  $K$ -smoothing network [1, 4] has smoothness  $K$ . A **perfect** smoothing network should meet four desiderata:

- (1) It should guarantee **optimal smoothness** in order to optimize the performance of load balancing applications running on top of the smoothing network.
- (2) It should have **optimal  $\mathcal{O}(\lg w)$  depth**;  $\Omega(\lg w)$  has been observed to be a lower bound on the depth of a (1-)smoothing network [15].
- (3) It should be **simple** in order to allow for practical implementations; this excludes the use of the famous AKS sorting network due to Ajtai, Komlós and Szemerédi [3] (with depth  $\Theta(\lg w)$ ) as a building block in the construction due to its huge constants hidden in the  $\Theta(\lg w)$  notation.
- (4) It should permit for **local initialization** [2, 8, 9] in order to increase robustness against balancer failures.

Despite a large research effort in the last eighteen years (see, for example, [2, 4, 8, 9, 11, 12, 13, 14]), no perfect smoothing network has been yet known. As early as 1992, Klugerman and Plaxton envisioned [13]:

“In view of the fact that every smoothing network produced thus far has incorporated a sorting network as a primary component, it would be interesting to bound the depth complexity of sorting by a small constant times the depth complexity of smoothing, or to construct a small-depth smoothing network that makes no use of sorting networks.”

In this work, we shall present a smoothing network that employs *randomization*. As we shall explain, the smoothing network we shall present comes as close as possible to the second vision of Klugerman and Plaxton [13] while simultaneously providing additional benefits.

We present a very simple, *randomized*,  $\mathcal{O}(\lg w)$ -depth 2-smoothing network which meets all four desiderata. The network is the cascade of a reasonably small number of copies of the very simple *block network* introduced in [6] and used in many constructions of smoothing and counting networks (such as the *periodic counting network* [4]); for example, the required number of copies is no more than 323 for  $w \geq 2^{12}$ , and no more than 102 for  $w \geq 2^{30}$ . Since it has been observed in [14, Section 4.4] that the block network is topologically equivalent to the popular *cube-connected-cycles network* [16], we shall often refer to the latter in our discussion.

The network uses *randomized initialization*, where each balancer is oriented either *top* or *bottom* independently and uniformly at random in some local initialization phase; so desideratum (4) is met. Since the block network is very simple (and, in particular, it makes no use of the AKS sorting network), desideratum (3) is also met. The block network has depth (exactly)  $\lg w$ , so that desideratum (2) is also met. Finally, we recall a recent improbability result due to Mavronicolas and Sauerwald [14, Theorem 7.1 and Corollary 7.2], implying that there is no  $\mathcal{O}(\lg w)$ -depth, randomized 1-smoothing network with constant probability. This implies that  $K = 2$  is the **optimal** smoothness one could hope for when restricted to randomized networks guaranteeing smoothness with probability no smaller than constant; hence, desideratum (1) is met.

Klugerman and Plaxton [13, Section 4.3] present an explicit construction of a *2-counter*: a network guaranteeing that its output will have the *step property* [4] when its input is 2-smooth. (Formally, a balancing network  $\mathbf{B}_w$  is a *2-counter* if the assumption that its input vector  $\mathbf{x}$  is 2-smooth implies that its output vector  $\mathbf{y}$  is *step*: for any pair of indices  $0 \leq i < j \leq w - 1$ ,  $0 \leq y_i - y_j \leq 1$ ; so, a 2-counter is a conditional version of a *counting network* [4].) The 2-counter of Klugerman and Plaxton [13] is deterministic and it achieves depth  $\Theta(\lg w)$  and uses the AKS sorting network [3] as a building block. Now, the cascade of the (randomized) 2-smoothing network from this paper and the 2-counter from [13] yields a randomized counting network which achieves depth  $\Theta(\lg w)$  (and uses the AKS sorting network). This complements nicely the existence result of a counting network with  $\Theta(\lg w)$  depth from Klugerman and Plaxton [13]; they provided a random construction (implying the existence of a deterministic network), which was later derandomized by Klugerman [12], thus yielding an ex-

PLICIT construction of a deterministic counting network with these properties. We feel that the cascade of our randomized 2-smoothing network with the 2-counter from [13] provides a much simpler and transparent, explicit construction of a small-depth counting network, albeit randomized, than the ones in [12, 13].

In the reverse direction, the  $\Theta(\lg w)$ -depth, 2-smoothing network constructed in this work offers a revival to the first vision of Klugerman and Plaxton [13]: now, a way to construct a (randomized)  $\Theta(\lg w)$ -depth sorting network improving on the AKS sorting network (in terms of the hidden constants) is to bound the depth complexity of sorting by a *small* constant times the depth complexity of 2-smoothing (and then use the  $\Theta(\lg w)$ -depth 2-smoothing network from this paper).

Our analysis uses (in Section 3) the following two ingredients.

- We use the notion of *maximum-survive* (*minimum-survive*, resp.) *path* as a variant of a similar one introduced recently in [7]. Roughly speaking, a maximum-survive (resp., minimum-survive) path traverses a network starting from an input wire; it continues layer by layer as long as the maximum (resp., minimum) number of input tokens is not “destroyed” due to meeting at some balancer with an input wire carrying a sufficiently smaller (resp., larger) number of tokens. See Definition 3.1 (resp., Definition 3.2) for the formal details. Basic combinatorial properties of maximum-survive and minimum-survive paths are stated in Observation 3.5.

The use of maximum- and minimum-survive paths is essential for our analysis, since one block network may not be alone sufficient to reduce the smoothness of its input tokens by 1. However, we shall prove that after each block network, the number of indices in the vector of input tokens that have the maximum (or minimum) number of tokens decreases significantly as tokens proceed to traverse the layers. This implies that after sufficiently many block networks, the smoothness does decrease by one.

- To prove that at each such maximum-survive (resp., minimum-survive) path, the maximum (resp., minimum) number of input tokens is “destroyed” at a certain layer, we first present an improvement of the so-called *Concentration-to-Average-Lemma* [14, Lemma 6.2]. This technical claim concerns the probability that some well-determined subnetwork of the cube-connected-cycles network receives tokens whose average number (for the particular subnetwork) is within a small fraction (precisely,  $\frac{1}{4}$ ) of the average number of tokens with respect to the entire network. The precise improvement is recorded in Lemma 4.1.

We then continue to derive a new deviation inequality (Lemma 4.2), establishing that with reasonable probability, the survive-maximum (resp., survive-minimum) path will eventually terminate. Deviation inequalities of this kind, which were previously employed in [7, 8, 14], were essentially based on *Hoeffding’s Bound* [10]. However, for small deviations as the ones required for eventually establishing a smoothness of 2, such deviation inequalities may only provide trivial bounds:

namely, an upper bound on the probability which is larger than 1.

The main result is established as a simple consequence of Lemma 5.2. In more detail, this establishes that the cascade of a reasonably small number of copies of the cube-connected-cycles network suffices to reduce the smoothness by 1 (Proposition 5.1). Repeating this cascading over and over yields eventually a smoothness of 2 (with high probability).

The randomized  $\mathcal{O}(\lg w)$ -depth smoothing network we are presenting is the *first* known network that simultaneously meets all four desiderata for smoothing networks. A summary of all known results on smoothing networks with constant smoothness appears in Table 1; this is based on [14, Table 1], which is extended to incorporate the present result and a recent related result from [7].

Very recently, Friedrich and Sauerwald [7] identify a large class of smoothing networks with constant smoothness. Specifically, they prove any *expander graph* with  $w$  vertices induces a smoothing network with depth  $\mathcal{O}(\lg w (\lg \lg w)^3)$  that guarantees constant smoothness (but no less than 10) with high probability, provided that the smoothness of the input vector is polynomial in  $w$ . This result is *orthogonal* to the result presented in this paper: while it is more general in applying to all expanders, our result provides an optimal constant (2) for smoothness and optimal  $\Theta(\lg w)$ -depth for a *specific* network (namely, the cascade of some copies of the block network). Moreover, for our result, no assumption on the initial smoothness is made.

The simple randomized two-blocks network from [14, Section 6] achieves smoothness of 17. We consider that the improvement from 17 to 2 is major, especially because smoothness of 2 is optimal due to the improbability result from [14, Section 7]. Even more so, our network is the *first* network with smoothness 2 that does not rely on global initialization. An earlier construction of a 2-smoothing network by Aiello *et al.* [2, Theorem 3.1] relies (partially) on global initialization, and so it fails to meet desideratum (3); furthermore, it uses a less simple construction involving the *butterfly network* and the *bitonic network* due to Batcher [5], while our construction is much more simple and transparent. Finally, we remark that there are known 1-smoothing networks requiring global initialization, which either have  $\Theta(\lg^2 w)$  depth [4] or use the AKS sorting network [3] to achieve depth  $\Theta(\lg w)$  [12, 13].

## 2. PRELIMINARIES AND NOTATION

Our presentation follows closely the one in [14, Sections 2 & 3]. All logarithms are to the base 2. Given a fixed (power of two) integer  $w = 2^{\lg w}$ , we identify each integer  $i$  with  $0 \leq i \leq w - 1$  with its binary representation  $i_1 i_2 \dots i_{\lg w}$ . Moreover, for any integer  $j \geq 1$ , we define  $[j] = \{0, \dots, j - 1\}$ . For a vector  $\mathbf{x}$  with  $w$  entries, denote  $x_{\min} := \min_{i \in [w]} x_i$  and  $x_{\max} := \max_{i \in [w]} x_i$ ;  $\mathbf{x}$  is  $\gamma$ -smooth if  $\max_{i, j \in [w]} |x_i - x_j| \leq \gamma$ .

For a random variable  $v$ , we shall denote as  $\mathbb{E}[v]$  the *expectation* of  $v$ . In some later proofs, we shall use the *Union Bound*, *Markov's Inequality* and an elementary rule about *conditional expectations*.

LEMMA 2.1 (UNION BOUND). *For a finite sequence of events  $\mathcal{E}_1, \mathcal{E}_2, \dots$ ,  $\mathbb{P}[\bigvee_{i \geq 1} \mathcal{E}_i] \leq \sum_{i \geq 1} \mathbb{P}[\mathcal{E}_i]$ .*

LEMMA 2.2 (MARKOV'S INEQUALITY). *Let  $v$  be a non-negative random variable. Then for any number  $c > 0$ ,  $\mathbb{P}[v \geq c \cdot \mathbb{E}[v]] \leq \frac{1}{c}$ .*

LEMMA 2.3. *Let  $v$  be a random-variable, and let  $C_1, C_2, \dots$  be a countable set of events such that  $\mathbb{P}[\bigvee_{i=1}^{\infty} C_i] = 1$ . Then,  $\mathbb{E}[v] = \sum_{i=1}^{\infty} \mathbb{P}[C_i] \cdot \mathbb{E}[v \mid C_i]$ .*

## 3. SMOOTHING NETWORKS

Roughly speaking, a smoothing (balancing) network [4] is a collection of interconnected balancers. A *balancer* [4] is an asynchronous switch with two *input wires* and two *output wires* denoted as  $i_1(\mathbf{b})$  and  $i_2(\mathbf{b})$ . Each balancer is always in one of two states, *top* or *bottom*. During an *initialization* phase, each balancer is oriented either *top* or *bottom*. After the initialization, a stream of tokens enters the network at the input wires in an arbitrary way. The tokens propagate through the network by following the orientation of the balancers; each time a token passes through a balancer, the balancer instantaneously changes its orientation. This guarantees a fair distribution on the balancer's output wires if the number of arriving tokens is even. However, if the number is odd, an excess token arises which is forwarded to the output wire the balancer is oriented to. For a balancer  $\mathbf{b}$ , we shall write  $i_1(\mathbf{b}) \stackrel{\leftarrow}{\leftarrow} i_2(\mathbf{b})$  exactly when the excess token (if any) is forwarded to its top output wire  $i_1(\mathbf{b})$ , and  $i_1(\mathbf{b}) \stackrel{\rightarrow}{\rightarrow} i_2(\mathbf{b})$  otherwise.

A *balancing network*  $\mathbf{B}_w$  [4] is an acyclic network of balancers, where output wires of balancers are connected to input wires of (other) balancers. The *input wires*  $0, 1, \dots, w - 1$  may not be connected from any output wires; the *output wires*  $0, 1, \dots, w - 1$  may not be connected to any input wires. So, we shall consider a balancing network  $\mathbf{B}_w$  with the same number  $w$  of input and output wires, called the network's *width*. By the assumption of acyclicity, each balancer is assigned a unique integer called *layer*, which is the length of the longest path from an input wire to that balancer. The *depth*, denoted as  $d(\mathbf{B}_w)$ , is the maximum layer in the network. We denote by  $\mathbf{B}_w \setminus [\ell_1, \ell_2]$  the restriction of  $\mathbf{B}_w$  to the layers  $\ell_1 + 1, \dots, d(\mathbf{B}_w) - \ell_2$ .

We shall always consider a balancing network in a *quiescent* state where all tokens have exited. For any balancer  $\mathbf{b}$ , denote as  $x_1$  and  $x_2$  the numbers of tokens entering the input wires  $i_1(\mathbf{b})$  and  $i_2(\mathbf{b})$ , respectively, of  $\mathbf{b}$ . Denote as  $y_1$  and  $y_2$  the number of tokens exiting through the output wires  $i_1(\mathbf{b})$  and  $i_2(\mathbf{b})$ , respectively, of  $\mathbf{b}$ . If reference to  $\mathbf{b}$  is necessary, we shall also write  $x_1(\mathbf{b})$ ,  $x_2(\mathbf{b})$ ,  $y_1(\mathbf{b})$  and  $y_2(\mathbf{b})$ . If  $\mathbf{b}$  is oriented *top* (resp., *bottom*), then  $y_1 = \lceil \frac{x_1 + x_2}{2} \rceil$  and  $y_2 = \lfloor \frac{x_1 + x_2}{2} \rfloor$  (resp.,  $y_1 = \lfloor \frac{x_1 + x_2}{2} \rfloor$  and  $y_2 = \lceil \frac{x_1 + x_2}{2} \rceil$ ).

There are three natural ways of choosing an orientation for each balancer. The first is to allow each balancer to be oriented arbitrarily, which was considered in [9]. A second way is to consider a *global orientation* [4], where each balancer must be oriented in some certain way (for example, all balancers must be oriented *top*). In this work, we consider *random orientation*, where each balancer chooses its orientation uniformly and independently at random [2, 7, 8, 14].

A *path*  $\pi = (i_1, i_2, \dots, i_\ell)$  is a sequence of interconnected wires from layer 1 to layer  $\ell$ ,  $1 \leq \ell \leq d(\mathbf{B}_w)$ ; so, for each layer  $r$  with  $1 \leq r \leq \ell - 1$ ,  $i_r$  is connected to a balancer in layer  $r$  which has  $i_{r+1}$  as one of its two output wires. For

Network	Depth	Type	GI	Smoothness	AKS	With Probability	Reference
Bitonic	$\Theta(\lg^2 w)$	D	✓	1	X	Not applicable	[4, Theorem 3.6]
Periodic	$\Theta(\lg^2 w)$	D	✓	1	X	Not applicable	[4, Theorem 4.4]
KP	$\Theta(\lg w)$	D	✓	1	✓	Not applicable	[13, Theorem 5.2]
$r$ -Butterfly	$(1 + o(1)) \lg w$	D/R	✓	2	X	$\geq 1 - \frac{1}{\omega(w^k)}$	[2, Theorem 3.1]
Two Blocks	$2 \lg w$	R	X	17	X	$\geq 1 - 2 \cdot \frac{4 \lg \lg w - 39}{w}$	[14, Theorem 6.1]
Any	$d$	R	X	1	X	$\leq \frac{d}{w-1}$	[14, Theorem 7.1]
Expander	$\mathcal{O}(\lg w (\lg \lg w)^3)$	R	X	$\mathcal{O}(1)$	X	$\geq 1 - \mathcal{O}(\frac{1}{w})$	[7, Theorem 5.19]
449 Blocks	$449 \lg w$	R	X	2	X	$\geq 1 - \mathcal{O}(\frac{\lg \lg w}{w})$	Theorem 5.5

**Table 1:** Summary of results about smoothing networks with constant smoothness. D and R stand for deterministic (that is, globally initialized) and randomized balancers, respectively; D/R stands for a combination of deterministic and randomized balancers. GI stands for global initialization; the corresponding column indicates whether GI is required or not. AKS stands for the sorting network of Ajtai, Komlós and Szemerédi [3]; the corresponding column indicates whether the smoothing network uses the AKS network or not as a block in the construction. KP stands for Klugerman and Plaxton [12, 13]. The result of [7] requires additionally a smoothness of  $\mathcal{O}(\text{poly}(w))$  of the input vector.

each wire  $i_r$  with  $1 < r \leq \ell$ , let  $b_r$  the balancer in layer  $r$  which has input wire  $i_r$ . Then,  $x_{i_r}$  is the number of tokens at the input wire to  $b_r$  which is connected to the output wire  $i_{r-1}$  of the previous layer;  $\hat{x}_{i_r}$  is defined as the other input to  $b_r$ . For layers  $\ell_1 \leq \ell_2$ ,  $\Pi_{j_0}[\ell_1, \ell_2]$  denotes the set of all possible paths from a wire  $j_0$  in layer  $\ell_1$  to any wire in layer  $\ell_2$ .

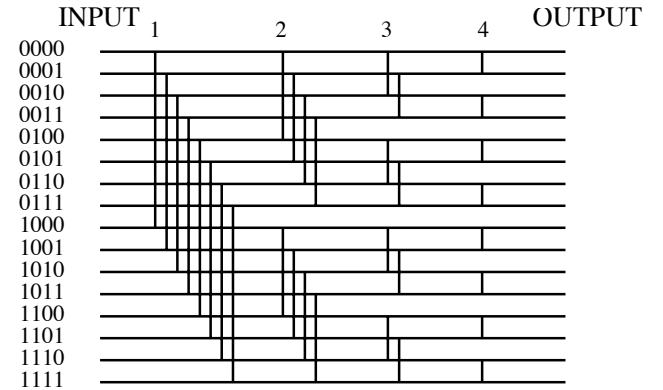
A balancer  $b$  in layer  $\ell$  depends on balancer  $b'$  in layer  $\ell' \leq \ell$  if there is a path from an output wire of  $b'$  to an input wire of  $b$ ; by convention, the balancer  $b$  depends trivially on itself. Dependencies among wires are defined in the same way. Two balancers  $b_1$  and  $b_2$  in layer  $\ell$  are independent, if there is no balancer  $b$  in an earlier layer on which both  $b_1$  and  $b_2$  depend. More specifically, two balancers  $b_1$  and  $b_2$  are *independent up to layer*  $\ell' \leq \ell$  if there is no balancer  $b$  in a layer between  $\ell'$  and  $\ell$  on which both depend. The *dependency set* of a balancer  $b$  in layer  $\ell$  is the set of all balancers in previous layers on which  $b$  depends. Dependency sets of wires (up to layers  $\ell' \leq \ell$ ) are defined in the same way.

A *randomized balancing network* [2, 8] or *randomized network* for short, is a balancing network with a random orientation. So, each balancer is initialized to each of *top* and *bottom* with probability  $\frac{1}{2}$ . To this end, we associate to each balancer  $b$  a random variable  $r_b$  taking values  $\frac{1}{2}$  and  $-\frac{1}{2}$  with equal probability (cf. [8]). (Clearly,  $\mathbb{E}[r_b] = 0$ .) Define also  $\chi_b = \text{Odd}(\mathbf{x}_b) \cdot r_b$  (cf. [8]). Then, the number of tokens at the two output wires  $y_1$  and  $y_2$  can be expressed as  $y_1 = \frac{x_1 + x_2}{2} + \text{Odd}(\mathbf{x}_b) \cdot r_b$  and  $y_2 = \frac{x_1 + x_2}{2} - \text{Odd}(\mathbf{x}_b) \cdot r_b$ . Fixing an input vector  $\mathbf{x}$  to a randomized balancing network induces a probability measure  $\mathbb{P}$  on associated events. In particular, it induces for each layer  $\ell$  with  $1 \leq \ell \leq d(\mathbf{B}_w)$  a *random vector*  $\mathbf{y}(\ell)$ ;  $\mathbf{y}(\ell)$  is determined by (i) the (random) input vector  $\mathbf{x}(\ell)$  and (ii) the random variables  $r_b$  corresponding to the orientation of the balancers in layer  $\ell$ . Write  $\mathbf{y} = \mathbf{y}(d(\mathbf{B}_w))$  to denote the (random) output vector of  $\mathbf{B}_w$ ; hence,  $\mathbf{B}_w(\mathbf{x}) = \mathbf{y}$ .

For some integer  $\gamma \geq 1$ , say that  $\mathbf{B}_w$  is a  $\gamma$ -*smoothing network with probability*  $\delta$  [2, 8], where  $0 \leq \delta \leq 1$ , if for each input vector  $\mathbf{x}$ ,  $\mathbb{P}[\mathbf{B}_w(\mathbf{x}) \text{ is } \gamma\text{-smooth}] \geq \delta$ ; that is, the probability that for each pair of output wires  $j, k \in [w]$ ,  $|y_j - y_k| \leq \gamma$  is at least  $\delta$ .

Henceforth, denote as  $\text{CCC}_w$  the *cube-connected-cycles net-*

*work* [16] of width  $w$  (where  $w$  is a power of 2) which consists of  $\lg w$  layers. In each layer  $\ell$  with  $1 \leq \ell \leq \lg w$ , for each wire  $u \in \{0, 1\}^{\lg w}$ , there is a balancer  $b$  connecting wire  $u$  and wire  $u(\ell)$ , where  $u(\ell)$  denotes the wire obtained by flipping the  $\ell$ -th bit of  $u$ . See Figure 1 for an illustration.



**Figure 1:** The  $\text{CCC}_{16}$  network.

We denote by  $\text{CCC}_w^2$  the sequential cascade of two  $\text{CCC}_w$  networks. We recall:

**LEMMA 3.1** ([14, LEMMA 4.4]). *Fix a pair of integers  $\ell_1$  and  $\ell_2$  with  $\ell_1 + \ell_2 < \lg w$ , and a corresponding pair of binary strings  $l_1 \in \{0, 1\}^{\ell_1}$  and  $l_2 \in \{0, 1\}^{\ell_2}$ . Then, the network  $\text{CCC}_w \setminus [\ell_1, \ell_2]$  restricted to the set of wires  $\{l_1 u l_2 \mid u \in \{0, 1\}^{\lg w - \ell_1 - \ell_2}\}$  is a cube-connected-cycles network  $\text{CCC}_{2^{\lg w - \ell_1 - \ell_2}}$ .*

The cube-connected-cycles network  $\text{CCC}_w$  is topologically equivalent to the block network  $\text{Block}_w$  (cf. [14, Section 4.4] for more details). This allows us to apply previous results stated (and proved) for the  $\text{Block}_w$  network to the  $\text{CCC}_w$  network.

By symmetry, all random variables  $y_j$  with  $j \in [w]$  are identically distributed (cf. [8, proof of Theorem 10]). In particular, for each output wire  $j \in [w]$  of  $\text{CCC}_w$ :  $y_j = \frac{\sum \mathbf{x}}{w} + \sum_{\ell=1}^{\lg w} \frac{1}{2^{\lg w - \ell}} \sum_{b \in v_{\ell,1}} \chi_b$  and  $\mathbb{E}[y_j] = 0$ . We will use the following lemma, which is an immediate consequence of the disjointness among dependency sets:

LEMMA 3.2 (CF. [14, LEMMA 3.3]). *Fix an input vector  $\mathbf{x}$  to a  $\text{CCC}_w$  network. Consider any path  $\pi = (i_1, \dots, i_\ell)$  from layer 1 to  $\ell$ . Let  $\lambda_1, \dots, \lambda_\ell$  and  $\tau_1, \dots, \tau_\ell$  be a collection of  $2\ell$  arbitrary, fixed numbers. Then,  $\mathbb{P}[\widehat{x}_{i_\ell} = \tau_\ell \mid (\bigwedge_{r=1}^\ell x_{i_r} = \lambda_r) \wedge (\bigwedge_{r=1}^{\ell-1} \widehat{x}_{i_r} = \tau_r)]$ .*

So, roughly speaking, the claim asserts that  $\widehat{x}_{i_\ell}$  is *independent* of any events associated to all previous inputs to the path; note, however, that  $x_{i_1}$  and  $x_{i_2}$  could be dependent. We shall use the following previous results:

THEOREM 3.3 ([14, THM. 5.1]). *The  $\text{CCC}_w$  network is a  $(\lceil \lg \lg w \rceil + 3)$ -smoothing network with probability at least  $1 - \frac{4}{w^3}$ .*

THEOREM 3.4 ([14, THM. 6.7]). *The  $\text{CCC}_w^2$  network is a 17-smoothing network with prob. at least  $1 - 2 \cdot \frac{4 \lg \lg w - 39}{w}$ .*

We now define certain paths in a network that can be seen as maximal trajectories of the maximum number and minimum number of tokens on the network's input wires, respectively. The next definition is a variant of one from [7].

DEFINITION 3.1. *Let  $\mathbf{B}_w$  be a smoothing network. For each wire  $j_1 \in [w]$  in layer 1 with  $x_{j_1}(1) = x_{\max}(1)$ , a **maximum-survive path**  $\pi_{j_1}^{\max}$  is defined by induction:*

- For the basis case where  $\ell = 1$ ,  $\pi_{j_1}^{\max} := (j_1)$ .
- Assume  $\pi_{j_1}^{\max} = (j_1, j_2, \dots, j_\ell)$  is a maximum-survive path from layer 1 to layer  $\ell$ .  
For the induction step, consider layer  $\ell + 1$ .

– If  $\widehat{x}_{j_\ell} = x_{\max}(1) - 1$ , then

$$j_{\ell+1} := \begin{cases} j_\ell, & \text{if } j_\ell \stackrel{\text{b}}{\leftarrow} j_\ell(\ell), \\ j_\ell(\ell), & \text{if } j_\ell \stackrel{\text{b}}{\rightarrow} j_\ell(\ell), \end{cases}$$

– If  $\widehat{x}_{j_\ell} = x_{\max}(1)$ , then  $j_{\ell+1} := j_\ell$ .

– Otherwise, the path  $\pi_{j_1}^{\max}$  terminates at layer  $\ell$  (with length  $|\pi_{j_1}^{\max}| = \ell$ ).

An illustration of the definition for the  $\text{CCC}_w$  network is given in Figure 2. We continue to define:

DEFINITION 3.2. *Let  $\mathbf{B}_w$  be a smoothing network. For each wire  $j_1 \in [w]$  in layer 1 with  $x_{j_1}(1) = x_{\min}(1)$ , a **minimum-survive path**  $\pi_{j_1}^{\min}$  is defined by induction:*

- For the basis case where  $\ell = 1$ ,  $\pi_{j_1}^{\min} := (j_1)$ .
- Assume  $\pi_{j_1}^{\min} = (j_1, j_2, \dots, j_\ell)$  is a minimum-survive path from layer 1 to layer  $\ell$ .  
For the induction step, consider layer  $\ell + 1$ .

– If  $\widehat{x}_{j_\ell} = x_{\min}(1) + 1$ , then

$$j_{\ell+1} := \begin{cases} j_\ell, & \text{if } j_\ell \stackrel{\text{b}}{\rightarrow} j_\ell(\ell), \\ j_\ell(\ell), & \text{if } j_\ell \stackrel{\text{b}}{\leftarrow} j_\ell(\ell), \end{cases}$$

– If  $\widehat{x}_{j_\ell} = x_{\min}(1)$ , then  $j_{\ell+1} := j_\ell$ .

– Otherwise, the path  $\pi_{j_1}^{\min}$  terminates at layer  $\ell$  (with length  $|\pi_{j_1}^{\min}| = \ell$ ).

So, a path  $\pi_{j_1}^{\max}$  (resp.,  $\pi_{j_1}^{\min}$ ) is a sequence of wires which receive the maximum (resp., minimum) number of tokens; the maximum (resp., minimum) is with respect to the input vector to the first layer. If the last input wire of the path is connected to a balancer which receives less than  $x_{\max}(1) - 1$  (resp., more than  $x_{\min}(1) + 1$ ) tokens, the path terminates; this happens because, by definition of a balancer both outputs of a balancer will then be less than  $x_{\max}(1)$  (resp., more than  $x_{\min}(1)$ ). Moreover, we define:

DEFINITION 3.3. *Let  $\mathbf{B}_w$  be a smoothing network. For any layer  $\ell$  with  $1 \leq \ell \leq \mathbf{d}(\mathbf{B}_w)$ , define*

$$|\Pi^{\max}(\ell)| := \left| \left\{ i \in [w] : |\pi_i^{\max}| > \ell \right\} \right|,$$

and

$$|\Pi^{\min}(\ell)| := \left| \left\{ i \in [w] : |\pi_i^{\min}| > \ell \right\} \right|.$$

So,  $|\Pi^{\max}(\ell)|$  ( $|\Pi^{\min}(\ell)|$ ) counts the number of maximum-survive paths  $\pi^{\max}$  (resp., minimum-survive paths  $\pi^{\min}$ ) surviving layer  $\ell$ , i.e., reaching layer  $\ell + 1$  (if  $\ell + 1 = \mathbf{d}(\mathbf{B}_w)$  this means that the maximum (resp., minimum) reaches an output wire of  $\mathbf{B}_w$ ). By Definition 3.1, for each wire  $i \in [w]$  in a layer  $\ell$  with  $x_i(\ell) = x_{\max}(1)$ , there is exactly one wire  $j(i) \in [w]$  which induces (exactly one) path  $\pi_j^{\max}$ ,  $j \in [w]$ , starting at layer 1 that reaches  $i(\ell)$  at layer  $\ell$ . Hence, we obtain:

OBSERVATION 3.5. *Let  $\mathbf{B}_w$  be a smoothing network. For every layer  $\ell$  with  $1 \leq \ell \leq \mathbf{d}(\mathbf{B}_w)$  it holds that*

$$|\{i \in [w] : x_i(\ell) = x_{\max}(1)\}| = |\Pi^{\max}(\ell)|$$

and

$$|\{i \in [w] : x_i(\ell) = x_{\min}(1)\}| = |\Pi^{\min}(\ell)|.$$

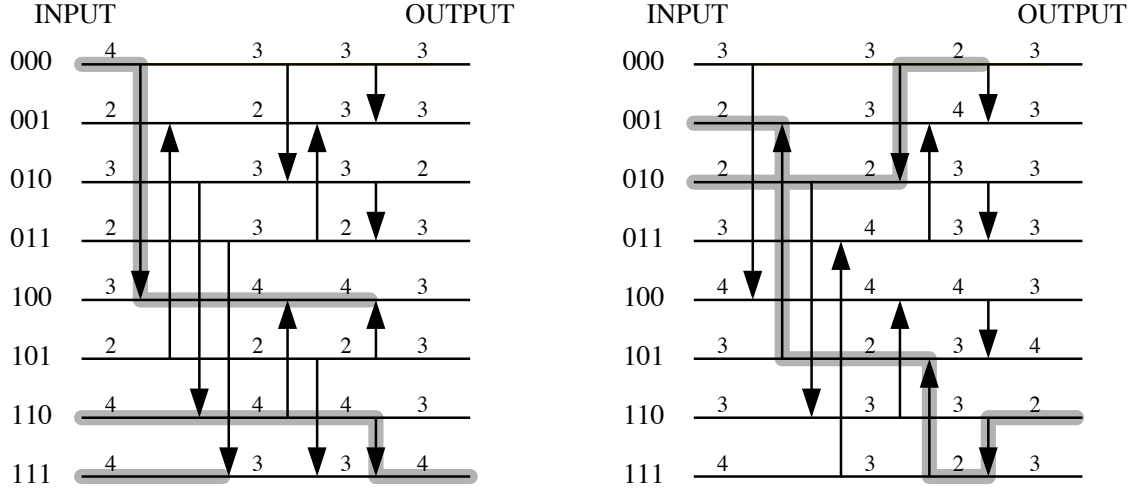
## 4. IMPROVED CONCENTRATION

We shall use the following definition slightly adapted from [14]; we shall require a stronger concentration by replacing the constant 2 in [14, Definition 6.1] with  $\frac{1}{4}$ , while slightly increasing  $\zeta$ .

DEFINITION 4.1 (CONCENTRATION-TO-AVERAGE). *Fix an input vector  $\mathbf{x}$  and a layer  $\ell$  with  $\lg w + 1 \leq \ell \leq 2 \lg w$  in the network  $\text{CCC}_w^2$ . Denote as  $\mathcal{E}(\mathbf{x}, \ell)$  the event that for all integers  $\zeta$  with  $6 \leq \zeta \leq 2 \lg w - \lceil \lg \lg w \rceil - \ell + 1$ , and for all pairs of binary strings  $u_1 \in \{0, 1\}^{\ell - \lg w - 1}$  and  $u_2 \in \{0, 1\}^{2 \lg w - \ell - \lceil \lg \lg w \rceil - \zeta + 1}$ ,*

$$\left| \frac{\sum_{\{u_1 u_2 \mid u \in \{0, 1\}^{\lceil \lg \lg w \rceil + \zeta}\}} \mathbf{x}(\ell)}{2^{\lceil \lg \lg w \rceil + \zeta}} - \frac{\sum \mathbf{x}}{w} \right| \leq \frac{1}{4}.$$

Lemma 3.1 implies that for each integer  $\zeta$  with  $6 \leq \zeta \leq 2 \lg w - \lceil \lg \lg w \rceil - \ell + 1$ , and for each pair of binary strings  $u_1 \in \{0, 1\}^{\ell - \lg w - 1}$  and  $u_2 \in \{0, 1\}^{2 \lg w - \ell - \lceil \lg \lg w \rceil - \zeta + 1}$ , the network  $\text{CCC}_w^2 \setminus [\ell - 1, \lg w - \lceil \lg \lg w \rceil - \zeta + 1]$  restricted to the set of wires  $\{u_1 u_2 \mid u \in \{0, 1\}^{\lceil \lg \lg w \rceil + \zeta}\}$  is a cube-connected-cycles network  $\text{CCC}_{2^{\lceil \lg \lg w \rceil + \zeta}}$ . Hence, observe that the event  $\mathcal{E}(\mathbf{x}, \ell)$  refers to each separate restriction of the inputs (to  $\text{CCC}_w^2$ ) on the input wires of such a smaller cube-connected-cycles network; it asserts that each such cube-connected-cycles network receives an average number of tokens (over its input wires) which is within  $\frac{1}{4}$  of the average  $\frac{\sum \mathbf{x}}{w}$ . We continue to establish that the event  $\mathcal{E}(\mathbf{x}, \ell)$  occurs with high probability. As in [14, Lemma 6.1], we prove:



**Figure 2:** An illustration of maximum- and minimum-survive paths for two particular inputs to the  $\text{CCC}_8$  network. The three maximum-survive paths on the left side are  $\pi_{000}^{\max} = (000, 100, 100)$ ,  $\pi_{110}^{\max} = (110, 110, 110, 111)$  and  $\pi_{111}^{\max} = (111)$ . The two minimum-survive paths on the right side are  $\pi_{001}^{\min} = (001, 101, 111, 110)$  and  $\pi_{010}^{\min} = (010, 010, 000)$ . All inputs and outputs to balancers are indicated. Balancers are depicted as oriented vertical line segments joining two wires; wires are depicted as horizontal lines.

LEMMA 4.1 (CONCENTRATION-TO-AVERAGE LEMMA). *Consider the randomized  $\text{CCC}_w^2$  network. Fix an input vector  $\mathbf{x}$  and a layer  $\ell$  with  $\lg w + 1 \leq \ell \leq 2 \lg w$ . Then,  $\mathbb{P}[\mathcal{E}(\mathbf{x}, \ell)] \geq 1 - \frac{4}{w^3}$ .*

The convenience provided by above Lemma is revealed in the following conditional concentration property.

LEMMA 4.2. *Consider the randomized  $\text{CCC}_w$  network and fix a layer  $\ell$  with  $1 \leq \ell \leq \lg w - \lceil \lg \lg w \rceil$ , an integer  $\zeta$  with  $6 \leq \zeta \leq \lg w - \lceil \lg \lg w \rceil - \ell + 1$ , and a pair of binary strings  $u_1 \in \{0, 1\}^{\ell-1}$  and  $u_2 \in \{0, 1\}^{\lg w - \lceil \lg \lg w \rceil - \ell - \zeta + 1}$ . Fix an input vector  $\mathbf{x}$  such that the event  $\mathcal{E}(\mathbf{x}, \ell)$  is satisfied. Consider a balancer  $\mathbf{b}$  in layer  $\ell(\mathbf{b}) = \ell + \lceil \lg \lg w \rceil + \zeta$  with input wires  $i = u_1 \hat{u} u_2$  and  $i(\ell(\mathbf{b}))$ , for some binary string  $\hat{u} \in \{0, 1\}^{\lceil \lg \lg w \rceil + \zeta}$ . Then, for any integer  $\delta > x_{\min}(\ell)$ ,*

$$\mathbb{P}[x(\mathbf{b}) \geq \delta] \leq \frac{\sum \mathbf{x} - x_{\min}(\ell) + \frac{1}{4}}{\delta - x_{\min}(\ell)},$$

and for any integer  $\delta < x_{\max}(\ell)$

$$\mathbb{P}[x(\mathbf{b}) \leq \delta] \leq \frac{x_{\max}(\ell) - \sum \mathbf{x} + \frac{1}{4}}{x_{\max}(\ell) - \delta}.$$

*Proof.* First observe that  $x(\mathbf{b})$  is connected to the output of a CCC subnetwork of depth  $\lceil \lg \lg w \rceil + \zeta$  with input wires  $\{u_1 u u_2 \mid u \in \{0, 1\}^{\lceil \lg \lg w \rceil + \zeta}\}$ , and hence

$$\mathbb{E} \left[ x(\mathbf{b}) - \frac{\sum \{u_1 u u_2 \mid u \in \{0, 1\}^{\lceil \lg \lg w \rceil + \zeta}\} \mathbf{x}}{2^{\lceil \lg \lg w \rceil + \zeta}} \right] = 0.$$

Using that  $\mathbf{x}$  satisfies  $\mathcal{E}(\mathbf{x}, \ell)$  and linearity of expectations

we obtain that

$$\begin{aligned} & \mathbb{E} \left[ x(\mathbf{b}) - \frac{\sum \{u_1 u u_2 \mid u \in \{0, 1\}^{\lceil \lg \lg w \rceil + \zeta}\} \mathbf{x}}{2^{\lceil \lg \lg w \rceil + \zeta}} \right] \\ & \geq \mathbb{E} \left[ x(\mathbf{b}) - \frac{\sum \mathbf{x}}{w} \right] - \frac{1}{4} \\ & = \sum_{k=x_{\min}(\ell)}^{\delta-1} k \cdot \mathbb{P}[x(\mathbf{b}) = k] + \sum_{k=\delta}^{x_{\max}(\ell)} k \cdot \mathbb{P}[x(\mathbf{b}) = k] - \frac{\sum \mathbf{x}}{w} - \frac{1}{4}. \end{aligned}$$

Clearly,

$$\begin{aligned} & \sum_{k=x_{\min}(\ell)}^{\delta-1} k \cdot \mathbb{P}[x(\mathbf{b}) = k] \\ & + \sum_{k=\delta}^{x_{\max}(\ell)} k \cdot \mathbb{P}[x(\mathbf{b}) = k] - \frac{\sum \mathbf{x}}{w} - \frac{1}{4} \\ & \geq \sum_{k=x_{\min}(\ell)}^{\delta-1} x_{\min}(\ell) \cdot \mathbb{P}[x(\mathbf{b}) = k] \\ & + \sum_{k=\delta}^{x_{\max}(\ell)} \delta \cdot \mathbb{P}[x(\mathbf{b}) = k] - \frac{\sum \mathbf{x}}{w} - \frac{1}{4} \\ & = x_{\min}(\ell) \cdot (1 - \mathbb{P}[x(\mathbf{b}) \geq \delta]) + \delta \cdot \mathbb{P}[x(\mathbf{b}) \geq \delta] - \frac{\sum \mathbf{x}}{w} - \frac{1}{4} \\ & = (\delta - x_{\min}(\ell)) \cdot \mathbb{P}[x(\mathbf{b}) \geq \delta] + x_{\min}(\ell) - \frac{\sum \mathbf{x}}{w} - \frac{1}{4}. \end{aligned}$$

Combining this inequality with the equality above yields

$$\begin{aligned} & (\delta - x_{\min}(\ell)) \cdot \mathbb{P}[x(\mathbf{b}) \geq \delta] + x_{\min}(\ell) - \frac{\sum \mathbf{x}}{w} - \frac{1}{4} \\ & \leq \mathbb{E} \left[ x(\mathbf{b}) - \frac{\sum \{u_1 u u_2 \mid u \in \{0, 1\}^{\lceil \lg \lg w \rceil + \zeta}\} \mathbf{x}}{2^{\lceil \lg \lg w \rceil + \zeta}} \right] = 0 \end{aligned}$$

Rearranging gives  $\mathbb{P}[x(\mathbf{b}) \geq \delta] \leq \frac{\sum \mathbf{x} - x_{\min}(\ell) + \frac{1}{4}}{\delta - x_{\min}(\ell)}$ , as needed. The second claim is established by identical arguments.  $\square$

The proof used an elementary argument based on the expectation of  $x(\mathbf{b})$ ; this argument is essential since claims based on *Hoeffding bound* (as those used in [7, 8] and [14, Lemma 4.3, Lemma 4.7]) could only yield a probability larger than 1 for very small  $\delta$ . However, as it will turn out in the proof of Lemma 5.2, Lemma 4.2 still gives non-trivial bounds for  $\delta$  very close to  $\frac{\sum \mathbf{x}}{w}$ .

## 5. MAIN RESULT

We first prove:

**PROPOSITION 5.1 (SMOOTHNESS REDUCTION LEMMA).** *Fix an integer  $\gamma \geq 3$ . Consider the cascade of*

$$1 + \left[ \left( \lg 2 / \lg \left( \frac{\gamma - 1}{2} + \frac{1}{4} \right) \right) \cdot \left( \frac{2 \lg w}{\lg w - \lceil \lg \lg w \rceil - 6} \right) \right]$$

*CCC<sub>w</sub> networks. Assume that the input vector  $\mathbf{x}$  is  $\gamma$ -smooth. Then, the output vector  $\mathbf{y}$  is  $(\gamma - 1)$ -smooth with probability at least  $1 - \frac{2}{w}$ .*

*Proof.* We start with a technical claim:

**LEMMA 5.2.** *Fix an arbitrary integer  $\gamma \geq 3$ . Consider a CCC<sub>w</sub> network with an input vector  $\mathbf{x}$  which is  $\gamma$ -smooth and satisfies  $\mathcal{E}(\mathbf{x}, 1)$ . Then, for any pair of a wire  $j_1 \in [w]$ , and a layer  $\ell$  with  $\lceil \lg \lg w \rceil + 7 \leq \ell \leq \lg w$  it holds that*

$$\mathbb{P} \left[ |\pi_{j_1}^{\max}| > \ell \right] \leq \left( \frac{\sum \mathbf{x} - x_{\min}(1) + \frac{1}{4}}{\gamma - 1} \right)^{\ell - \lceil \lg \lg w \rceil - 6},$$

and

$$\mathbb{P} \left[ |\pi_{j_1}^{\min}| > \ell \right] \leq \left( \frac{x_{\max}(1) - \frac{\sum \mathbf{x}}{w} + \frac{1}{4}}{\gamma - 1} \right)^{\ell - \lceil \lg \lg w \rceil - 6}.$$

*Proof.* We start with the first claim. We may assume that  $x_{\max}(1) - x_{\min}(1) = \gamma$ ; otherwise, the claim holds trivially, as in this case  $\mathbf{x}(1)$  is already  $(\gamma - 1)$ -smooth. To shorten the notation, we define  $\pi := \pi_{j_1}^{\max}$ ; recall that  $\pi$  is a random path depending on the load vector  $\mathbf{x}(1)$  and the balancers' orientation. Clearly,

$$\begin{aligned} \mathbb{P} [|\pi| > \ell] &= \prod_{r=\lceil \lg \lg w \rceil + 7}^{\ell} \mathbb{P} [|\pi| > r \mid |\pi| \geq r] \\ &= \prod_{r=\lceil \lg \lg w \rceil + 7}^{\ell} \frac{\mathbb{P} [|\pi| > r \wedge |\pi| \geq r]}{\mathbb{P} [|\pi| \geq r]}. \end{aligned}$$

Recall that  $\Pi_{j_1}[1, r]$  is the set of all paths from wire  $j_1$  in layer 1 to any wire in layer  $r$ . For a specific path  $\hat{\pi} \in \Pi_{j_1}[1, r]$ , we write  $\hat{\pi} \subseteq \pi$  to denote event that the (random path)  $\pi$  coincides with the (specific) path  $\hat{\pi}$  on the layers 1

to  $r$ . It follows by the law of Conditional Probabilities that

$$\begin{aligned} &\frac{\mathbb{P} [|\pi| > r \wedge |\pi| \geq r]}{\mathbb{P} [|\pi| \geq r]} \\ &= \frac{\sum_{\hat{\pi} \in \Pi_{j_1}[1, r]} \mathbb{P} [|\pi| > r \wedge |\pi| \geq r \wedge \hat{\pi} \subseteq \pi]}{\mathbb{P} [|\pi| \geq r]} \\ &= \frac{1}{\mathbb{P} [|\pi| \geq r]} \cdot \sum_{\hat{\pi} \in \Pi_{j_1}[1, r]} \left( \mathbb{P} [|\pi| > r \mid \hat{\pi} \subseteq \pi \wedge |\pi| \geq r] \cdot \mathbb{P} [|\pi| \geq r \mid \hat{\pi} \subseteq \pi] \cdot \mathbb{P} [\hat{\pi} \subseteq \pi] \right). \end{aligned}$$

By definition, the event  $\hat{\pi} \subseteq \pi$  implies  $|\pi| \geq r$  and hence

$$\begin{aligned} &\frac{1}{\mathbb{P} [|\pi| \geq r]} \cdot \sum_{\hat{\pi} \in \Pi_{j_1}[1, r]} \left( \mathbb{P} [|\pi| > r \mid \hat{\pi} \subseteq \pi \wedge |\pi| \geq r] \cdot \mathbb{P} [|\pi| \geq r \mid \hat{\pi} \subseteq \pi] \cdot \mathbb{P} [\hat{\pi} \subseteq \pi] \right) \\ &= \frac{1}{\mathbb{P} [|\pi| \geq r]} \cdot \sum_{\hat{\pi} \in \Pi_{j_1}[1, r]} \left( \mathbb{P} [|\pi| > r \mid \hat{\pi} \subseteq \pi] \cdot \mathbb{P} [|\pi| \geq r \mid \hat{\pi} \subseteq \pi] \cdot \mathbb{P} [\hat{\pi} \subseteq \pi] \right). \end{aligned}$$

Conditioned on the event  $\hat{\pi} \subseteq \pi$ , the path  $\pi = (j_1, \dots, j_r)$  survives layer  $r$  (that is,  $\pi$  extended to layer  $r + 1$ ) if and only if  $\hat{x}_{j_r} \geq x_{\max}(1) - 1$ . Hence Lemma 3.2 and Lemma 4.2 imply together that

$$\begin{aligned} \mathbb{P} [|\pi| > r \mid \hat{\pi} \subseteq \pi] &= \mathbb{P} [\hat{x}_{j_r} \geq x_{\max}(1) - 1] \\ &\leq \frac{\frac{\sum \mathbf{x}}{w} - x_{\min}(1) + \frac{1}{4}}{x_{\max}(1) - 1 - x_{\min}(1)} \\ &= \frac{\frac{\sum \mathbf{x}}{w} - x_{\min}(1) + \frac{1}{4}}{\gamma - 1}. \end{aligned}$$

Hence, it follows that

$$\begin{aligned} &\prod_{r=\lceil \lg \lg w \rceil + 7}^{\ell} \left( \frac{1}{\mathbb{P} [|\pi| \geq r]} \cdot \sum_{\hat{\pi} \in \Pi_{j_1}[1, r]} \left( \mathbb{P} [|\pi| > r \mid \hat{\pi} \subseteq \pi] \cdot \mathbb{P} [|\pi| \geq r \mid \hat{\pi} \subseteq \pi] \cdot \mathbb{P} [\hat{\pi} \subseteq \pi] \right) \right) \\ &\leq \left( \frac{\frac{\sum \mathbf{x}}{w} - x_{\min}(1) + \frac{1}{4}}{\gamma - 1} \right)^{\ell - \lceil \lg \lg w \rceil - 6} \\ &\cdot \prod_{r=\lceil \lg \lg w \rceil + 6}^{\ell} \frac{\sum_{\hat{\pi} \in \Pi_{j_1}[1, r]} \mathbb{P} [\hat{\pi} \subseteq \pi] \cdot \mathbb{P} [|\pi| \geq r \mid \hat{\pi} \subseteq \pi]}{\mathbb{P} [|\pi| \geq r]} \\ &= \left( \frac{\frac{\sum \mathbf{x}}{w} - x_{\min}(1) + \frac{1}{4}}{\gamma - 1} \right)^{\ell - \lceil \lg \lg w \rceil - 6} \\ &\cdot \prod_{r=\lceil \lg \lg w \rceil + 6}^{\ell} \frac{\sum_{\hat{\pi} \in \Pi_{j_1}[1, r]} \mathbb{P} [|\pi| \geq r \wedge \hat{\pi} \subseteq \pi]}{\mathbb{P} [|\pi| \geq r]} \\ &= \left( \frac{\frac{\sum \mathbf{x}}{w} - x_{\min}(1) + \frac{1}{4}}{\gamma - 1} \right)^{\ell - \lceil \lg \lg w \rceil - 6}, \end{aligned}$$

which immediately implies that

$$\mathbb{P} [|\pi| > \ell] \leq \left( \frac{\frac{\sum \mathbf{x}}{w} - x_{\min}(1) + \frac{1}{4}}{\gamma - 1} \right)^{\ell - \lceil \lg \lg w \rceil - 6}.$$

Using identical arguments, we prove the second claim.  $\square$

Consider now the cascade of

$$1 + \left[ \left( \lg 2 / \lg \left( \frac{\gamma - 1}{\frac{\gamma}{2} + \frac{1}{4}} \right) \right) \cdot \left( \frac{2 \lg w}{\lg w - \lceil \lg \lg w \rceil - 6} \right) \right]$$

CCC<sub>w</sub> networks. It follows by linearity of expectation and Lemma 2.3 that for any layer  $\ell$  with  $\lg w + 1 \leq \ell \leq 2 \lg w + 1$ ,

$$\begin{aligned} & \mathbb{E} [|\Pi^{\max}(\ell)|] \\ &= \mathbb{P}[\mathcal{E}(\mathbf{x}, \lg w + 1)] \cdot \mathbb{E} [|\Pi^{\max}(\ell)| \mid \mathcal{E}(\mathbf{x}, \lg w + 1)] \\ & \quad + (1 - \mathbb{P}[\mathcal{E}(\mathbf{x}, \lg w + 1)]) \cdot \mathbb{E} [|\Pi^{\max}(\ell)| \mid \neg \mathcal{E}(\mathbf{x}, \lg w + 1)] \\ &\leq \mathbb{E} [|\Pi^{\max}(\ell)| \mid \mathcal{E}(\mathbf{x}, \lg w + 1)] + \frac{4}{w^3} \cdot w, \end{aligned}$$

where the last inequality is due to Lemma 4.1.

As before, we may assume that  $x_{\max}(1) - x_{\min}(1) = \gamma$ , since otherwise  $x_{\max}(1) - x_{\min}(1) \leq \gamma - 1$  and Proposition 5.1 holds trivially. Furthermore, we assume that  $\frac{\sum \mathbf{x}}{w} - x_{\min}(1) \leq \frac{\gamma}{2}$  (otherwise,  $x_{\max}(1) - \frac{\sum \mathbf{x}}{w} \leq \frac{\gamma}{2}$  and we proceed by identical arguments bounding  $|\Pi^{\min}(\ell)|$  instead of  $|\Pi^{\max}(\ell)|$ ). Linearity of expectations, Observation 3.5 and Lemma 5.2 imply that for every layer  $\ell$  with  $\lg w + \lceil \lg \lg w \rceil + 7 \leq \ell \leq 2 \lg w + 1$ ,

$$\begin{aligned} & \mathbb{E} [|\Pi^{\max}(\ell)|] \\ &\leq \mathbb{E} [|\Pi^{\max}(\ell)| \mid \mathcal{E}(\mathbf{x}, \lg w + 1)] + \frac{4}{w^2} \\ &= \sum_{\substack{j_1: \\ x_{j_1}(1) = x_{\max}(1)}} \mathbb{P} [|\pi_{j_1}^{\max}| > \ell \mid \mathcal{E}(\mathbf{x}, \lg w + 1)] + \frac{4}{w^2} \\ &\leq |\Pi^{\max}(1)| \cdot \left( \frac{\frac{\gamma}{2} + \frac{1}{4}}{\gamma - 1} \right)^{\ell - \lg w - \lceil \lg \lg w \rceil - 6} + \frac{4}{w^2}. \end{aligned}$$

More generally, we obtain this way for any integer  $\alpha \geq 2$ ,

$$\begin{aligned} & \mathbb{E} [|\Pi^{\max}(\alpha \lg w + 1)|] \\ &\leq |\Pi^{\max}((\alpha - 1) \lg w + 1)| \cdot \left( \frac{\frac{\gamma}{2} + \frac{1}{4}}{\gamma - 1} \right)^{\lg w - \lceil \lg \lg w \rceil - 6} + \frac{4}{w^2}. \end{aligned}$$

We continue to prove:

LEMMA 5.3. *For each integer  $\alpha \in \mathbb{N}$ , it holds that*

$$\begin{aligned} \mathbb{E} [|\Pi^{\max}(\alpha \lg w + 1)|] &\leq w \cdot \left( \frac{\frac{\gamma}{2} + \frac{1}{4}}{\gamma - 1} \right)^{(\alpha - 1) (\lg w - \lceil \lg \lg w \rceil - 6)} \\ & \quad + \frac{4(\alpha - 1)}{w^2}. \end{aligned}$$

*Proof.* By induction on  $\alpha$ . The basis case  $\alpha = 1$  holds vacuously. Assume inductively that the claim holds for  $\alpha - 1$ .

For the induction step, Lemma 2.3 gives

$$\begin{aligned} & \mathbb{E} [|\Pi^{\max}((\alpha + 1) \lg w + 1)|] \\ &= \sum_{j=1}^w \mathbb{P} [|\Pi^{\max}(\alpha \lg w + 1)| = j] \\ & \quad \cdot \mathbb{E} [|\Pi^{\max}((\alpha + 1) \lg w + 1)| \mid |\Pi^{\max}(\alpha \lg w + 1)| = j] \\ &\leq \sum_{j=1}^w \mathbb{P} [|\Pi^{\max}(\alpha \lg w + 1)| = j] \\ & \quad \cdot \left( j \cdot \left( \frac{\frac{\gamma}{2} + \frac{1}{4}}{\gamma - 1} \right)^{\lg w - \lceil \lg \lg w \rceil - 6} + \frac{4}{w^2} \right) \\ &= \left( \frac{\frac{\gamma}{2} + \frac{1}{4}}{\gamma - 1} \right)^{\lg w - \lceil \lg \lg w \rceil - 6} \cdot \mathbb{E} [|\Pi^{\max}(\alpha \lg w + 1)|] + \frac{4}{w^2} \\ &\leq \left( \frac{\frac{\gamma}{2} + \frac{1}{4}}{\gamma - 1} \right)^{\lg w - \lceil \lg \lg w \rceil - 6} \\ & \quad \cdot \left( w \cdot \left( \frac{\frac{\gamma}{2} + \frac{1}{4}}{\gamma - 1} \right)^{(\alpha - 1) (\lg w - \lceil \lg \lg w \rceil - 6)} + \frac{4(\alpha - 1)}{w^2} \right) + \frac{4}{w^2} \\ &\leq w \cdot \left( \frac{\frac{\gamma}{2} + \frac{1}{4}}{\gamma - 1} \right)^{\alpha (\lg w - \lceil \lg \lg w \rceil - 6)} + \frac{4(\alpha - 1)}{w^2} + \frac{4}{w^2} \\ &= w \cdot \left( \frac{\frac{\gamma}{2} + \frac{1}{4}}{\gamma - 1} \right)^{\alpha (\lg w - \lceil \lg \lg w \rceil - 6)} + \frac{4\alpha}{w^2}, \end{aligned}$$

where the second last inequality holds due to the induction hypothesis.  $\square$

For any  $3 \leq \gamma \leq 17$ , fix now the value  $\alpha = \alpha(\gamma) := 1 + \left[ \left( \lg 2 / \lg \left( \frac{\gamma - 1}{\frac{\gamma}{2} + \frac{1}{4}} \right) \right) \cdot \left( \frac{2 \lg w}{\lg w - \lceil \lg \lg w \rceil - 6} \right) \right]$ . We observe:

OBSERVATION 5.4. *For any  $w \geq 2^{12}$  and any  $\gamma \geq 3$ , it holds that  $4(\alpha - 1) \leq w$ .*

By the choice of  $\alpha$  and Lemma 5.3,

$$\begin{aligned} & \mathbb{E} [|\Pi^{\max}(\alpha \lg w + 1)|] \\ &\leq w \cdot \left( \frac{\frac{\gamma}{2} + \frac{1}{4}}{\gamma - 1} \right)^{(\alpha - 1) (\lg w - \lceil \lg \lg w \rceil - 6)} + \frac{4(\alpha - 1)}{w^2} \\ &\leq w \cdot 2^{-\frac{2 \lg w}{\lg w - \lceil \lg \lg w \rceil - 6} (\lg w - \lceil \lg \lg w \rceil - 6)} + \frac{4(\alpha - 1)}{w^2} \\ &\leq w \cdot \frac{1}{w^2} + \frac{w}{w^2} = 2 \cdot \frac{1}{w}, \end{aligned}$$

where the last inequality is due to Observation 5.4. Since  $|\Pi^{\max}(\alpha \lg w + 1)|$  is an integer random variable, Markov's inequality (Lemma 2.2) implies  $\mathbb{P} [|\Pi^{\max}(\alpha \lg w + 1)| > 0] \leq \frac{2}{w}$ . By Observation 3.5,  $|\Pi^{\max}(\alpha \lg w + 1)| = 0$  implies that the output vector  $\mathbf{y}(\alpha \lg w + 1)$  is  $(\gamma - 1)$ -smooth, as needed.  $\square$

We are now ready to prove:

THEOREM 5.5. *Fix a value  $w \geq 2^{12}$ . Then the cascade of  $2 + \sum_{\gamma=3}^{17} 1 + \left[ \left( \lg 2 / \lg \left( \frac{\gamma - 1}{\frac{\gamma}{2} + \frac{1}{4}} \right) \right) \cdot \left( \frac{2 \lg w}{\lg w - \lceil \lg \lg w \rceil - 6} \right) \right]$  CCC<sub>w</sub> networks is a 2-smoothing network with probability at least  $1 - 30 w^{-1} - \max \{4w^{-3}, 2 \frac{4 \lg \lg w - 39}{w}\}$ .*



$w$	$2^{12}$	$2^{13}$	$2^{14}$	$2^{15}$	$2^{16}$	$2^{17}$	$2^{18}$	$2^{19}$	$2^{20}$	$2^{25}$	$2^{30}$
# CCC $_w$	323	240	200	174	157	165	153	142	137	114	102

**Table 2:** Number of required CCC $_w$  networks to get a 2-smoothing network for various  $w$ . Note that due to rounding issues, the given upper bound is not strictly decreasing in  $w$ .

*Proof.* We shall prove by backward induction on  $k$  with  $2 \leq k \leq 17$  that the cascade of

$$2 + \sum_{\gamma=k+1}^{17} 1 + \left[ \left( \lg 2 / \lg \left( \frac{\gamma-1}{\frac{\gamma}{2} + \frac{1}{4}} \right) \right) \cdot \left( \frac{2 \lg w}{\lg w - \lfloor \lg \lg w \rfloor - 6} \right) \right]$$

CCC $_w$  networks is a  $k$ -smoothing network with probability  $1 - 2(17-k)w^{-1} - \max \{4w^{-3}, 2 \frac{4 \lg \lg w - 39}{w}\}$ . For the basis case where  $k = 17$ , we distinguish between two cases on  $w$ .

First assume that  $w \leq 2^{2^{13}}$ . By Theorem 3.3, a single CCC $_w$  network is a  $(\lfloor \lg \lg w \rfloor + 3)$ -smoothing network with probability at least  $1 - 4w^{-3}$ . Observe that by assumption on  $w$ ,  $\lfloor \lg \lg w \rfloor + 3 \leq 17$ . For the second case, assume that  $w > 2^{2^{13}}$ . Then, Theorem 3.4 implies that the output of the cascade of two CCC $_w$  networks is a 17-smoothing network with probability at least  $1 - 2 \frac{4 \lg \lg w - 39}{w}$ . This completes the basis case.

For the induction step, consider the cascade of  $2 + \sum_{\gamma=k}^{17} 1 + \left[ \left( \lg 2 / \lg \left( \frac{\gamma-1}{\frac{\gamma}{2} + \frac{1}{4}} \right) \right) \cdot \left( \frac{2 \lg w}{\lg w - \lfloor \lg \lg w \rfloor - 6} \right) \right]$  CCC $_w$  networks. By induction hypothesis, we obtain that the cascade of  $2 + \sum_{\gamma=k+1}^{17} 1 + \left[ \left( \lg 2 / \lg \left( \frac{\gamma-1}{\frac{\gamma}{2} + \frac{1}{4}} \right) \right) \cdot \left( \frac{2 \lg w}{\lg w - \lfloor \lg \lg w \rfloor - 6} \right) \right]$  CCC $_w$  networks is a  $k$ -smoothing network with probability at least  $1 - 2(17-k)w^{-1} - \max \{4w^{-3}, 2 \frac{4 \lg \lg w - 39}{w}\}$ . By an application of Proposition 5.1, it follows that the cascade of  $1 + \left[ \left( \lg 2 / \lg \left( \frac{k-1}{\frac{k}{2} + \frac{1}{4}} \right) \right) \cdot \left( \frac{2 \lg w}{\lg w - \lfloor \lg \lg w \rfloor - 6} \right) \right]$  CCC $_w$  networks reduces the smoothness from  $k$  to  $k-1$  with probability at least  $1 - 2w^{-1}$ . Hence by the Union Bound, the cascade of  $2 + \sum_{\gamma=k}^{17} 1 + \left[ \left( \lg 2 / \lg \left( \frac{\gamma-1}{\frac{\gamma}{2} + \frac{1}{4}} \right) \right) \cdot \left( \frac{2 \lg w}{\lg w - \lfloor \lg \lg w \rfloor - 6} \right) \right]$  CCC $_w$  networks is a  $(k-1)$ -smoothing network with probability at least  $1 - 2w^{-1} - 2(17-k)w^{-1} - \max \{4w^{-3}, 2 \frac{4 \lg \lg w - 39}{w}\} = 1 - 2(17-(k-1))w^{-1} - \max \{4w^{-3}, 2 \frac{4 \lg \lg w - 39}{w}\}$ .  $\square$

To present a more concrete upper bound on the number of cascaded CCC $_w$  networks, we calculate:

OBSERVATION 5.6. *We have (i)*

$$\sum_{\gamma=3}^{17} \left[ \left( \lg 2 / \lg \left( \frac{\gamma-1}{\frac{\gamma}{2} + \frac{1}{4}} \right) \right) \right] \leq 36,$$

and (ii), for any  $w \geq 2^{12}$ ,

$$\left[ \frac{2 \lg w}{\lg w - \lfloor \lg \lg w \rfloor - 6} \right] \leq 12.$$

*Proof.* The first estimate is independent of  $w$ , so it can be verified numerically. For the second estimate, we observe that for any  $w \geq 2^{12}$ ,  $\lfloor \lg \lg w \rfloor \leq \frac{1}{3} \lg w$  and  $6 \leq \frac{1}{2} \lg w$ . Hence,  $\frac{2 \lg w}{\lg w - \lfloor \lg \lg w \rfloor - 6} \leq 12$ .  $\square$

Using Theorem 5.5 and Observation 5.6 we immediately obtain:

COROLLARY 5.7. *For any  $w \geq 2^{12}$ , the cascade of 449 CCC $_w$  networks is a 2-smoothing network with probability at least  $1 - 30w^{-1} - \max \{4w^{-3}, 2 \frac{4 \lg \lg w - 39}{w}\}$ .*

*Proof.* Theorem 5.5 and Observation 5.6 imply that the cascade of

$$2 + \sum_{\gamma=3}^{17} 1 + \left[ \left( \lg 2 / \lg \left( \frac{\gamma-1}{\frac{\gamma}{2} + \frac{1}{4}} \right) \right) \cdot \left( \frac{2 \lg w}{\lg w - \lfloor \lg \lg w \rfloor - 6} \right) \right] \\ \leq 17 + \sum_{\gamma=3}^{17} \left[ \left( \lg 2 / \lg \left( \frac{\gamma-1}{\frac{\gamma}{2} + \frac{1}{4}} \right) \right) \right] \cdot \left[ \left( \frac{2 \lg w}{\lg w - \lfloor \lg \lg w \rfloor - 6} \right) \right] \\ \leq 17 + 36 \cdot 12 = 449.$$

CCC $_w$  networks is a 2-smoothing network with probability at least  $1 - 30w^{-1} - \max \{4w^{-3}, 2 \frac{4 \lg \lg w - 39}{w}\}$ .  $\square$

For larger  $w$  and a more careful calculation, we can obtain much smaller upper bounds on the required number of CCC $_w$  networks to achieve 2-smoothness (Table 2).

## 6. EPILOGUE

In this work we presented a simple, randomized  $\Theta(\lg w)$ -depth 2-smoothing network which meets all four desiderata on smoothing network, thus resolving a long-standing open problem dating back at least to the early work of Klugerman and Plaxton [12, 13]. Improving the constant number of required block networks remains an interesting open problem; the current record is about 150 for reasonably large  $w$ .

## 7. REFERENCES

- [1] E. Aharonson and H. Attiya, "Counting Networks with Arbitrary Fan-Out," *Distributed Computing*, Vol. 8, No. 4, pp. 163–169, 1995.
- [2] W. Aiello, R. Venkatesan and M. Yung, "Coins, Weights and Contention in Balancing Networks," *Proceedings of the 13th Annual ACM Symposium on Principles of Distributed Computing*, pp. 193–205, 1994.
- [3] M. Ajtai, J. Komlós and E. Szemerédi, "Sorting in  $c \log n$  Parallel Steps," *Combinatorica*, Vol. 3, pp. 1–19, 1983.
- [4] J. Aspnes, M. Herlihy and N. Shavit, "Counting Networks," *Journal of the ACM*, Vol. 41, No. 5, pp. 1020–1048, 1994.
- [5] K. E. Batcher, "Sorting Networks and their Applications," *Proceedings of the AFIPS Joint Computer Conference*, pp. 334–338, 1968.
- [6] M. Dowd, Y. Perl, L. Rudolph, and M. Saks, "The Periodic Balanced Sorting Network," *Journal of the ACM*, Vol. 36, No. 4, pp. 738–757, 1989.
- [7] T. Friedrich and T. Sauerwald, "Near-Perfect Load Balancing by Randomized Rounding," *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, to appear, 2009.

- [8] M. Herlihy and S. Tirthapura, "Randomized Smoothing Networks," *Journal of Parallel and Distributed Computing*, Vol. 66, No. 5, pp. 626–632, 2006.
- [9] M. Herlihy and S. Tirthapura, "Self-Stabilizing Smoothing and Counting Networks," *Distributed Computing*, Vol. 18, No. 5, pp. 345–357, 2006.
- [10] W. Hoeffding. "Probability Inequalities for Sums of Bounded Random Variables," *Journal of the American Statistical Association*, Vol. 53, No. 301, pp. 13–30, 1963.
- [11] S. Kapidakis and M. Mavronicolas, "Distributed, Low Contention Task Allocation," *Proceedings of the 8th IEEE Symposium on Parallel and Distributed Processing*, pp. 358–365, 1996.
- [12] M. Klugerman, *Small-Depth Counting Networks and Related Topics*, Ph.D. Thesis, Department of Mathematics, Massachusetts Institute of Technology, 1994.
- [13] M. Klugerman and C. G. Plaxton, "Small-Depth Counting Networks," *Proceedings of the 24th Annual ACM Symposium on Theory of Computing*, pp. 417–428, 1992.
- [14] M. Mavronicolas and T. Sauerwald, "The Impact of Randomization in Smoothing Networks," *Proceedings of the 27th Annual ACM Symposium on Principles of Distributed Computing*, pp. 345–354, 2008.  
available at <http://www.cs.ucy.ac.cy/~mavronic>
- [15] S. Moran and G. Taubenfeld "A Lower Bound on Wait-Free Counting," *Journal of Algorithms*. Vol. 24, No. 1, pp. 1–17, 1997.
- [16] F. Preparata and J. Vuillemin, "The Cube-Connected-Cycles: A Versatile Network for Parallel Computation," *Communications of the ACM*. Vol. 24, No. 5, pp. 300–309, 1981.