Proving Correctness for Balancing Networks*

Costas Busch†
Department of Computer Science
University of Crete
Heraklion 71110, Greece
and
Institute of Computer Science
Foundation for Research and Technology – Hellas
Heraklion 71110, Greece

Marios Mavronicolas‡
Department of Computer Science
University of Cyprus
Nicosia 1678, Cyprus

December 1994

*This work has been partially supported by ESPRIT III Basic Research Project # 8144 – LYDIA (Load Balancing on High-Performance Parallel and Distributed Systems). A preliminary version of this work has been presented in the *DIMACS Workshop on Parallel Processing of Discrete Optimization Problems,* DIMACS (Center for Discrete Mathematics and Theoretical Computer Science), Rutgers University, New Jersey, April 1994.

†E-mail address: mpous@csi.forth.gr

‡Part of the work of this author was performed while visiting Institute of Computer Science, Foundation for Research and Technology – Hellas. Partially supported by the fund for the promotion of research at University of Cyprus (Research Project "Load Balancing Problems in Shared-Memory Multiprocessor Architectures"). E-mail address: mavronic@csi.forth.gr

Abstract

Balancing networks have recently been proposed by Aspnes, Herlihy and Shavit (*Proc. of the 23rd Annual ACM Symp. on Theory of Computing,* pp. 348–358, May 1991) as a new class of distributed, low-contention data structures suitable for solving a variety of multi-processor coordination problems that can be expressed as *balancing problems*.

In a recent work (*Proc. of the 13th Annual ACM Symp. on Principles of Distributed Computing,* pp. 206–215, August 1994), Busch and Mavronicolas develop a mathematical theory of the combinatorial structure of balancing networks. In this work, a paradigmatic methodology for showing correctness of balancing networks is developed as a direct consequence of this combinatorial theory. This methodology is applied to yield a transparent correctness proof for the *bitonic* counting network introduced by Aspnes *et al.*, whose layout is isomorphic to that of the classical, bitonic sorting network of Batcher; our proof provides an interesting complement to the one given by Aspnes *et al.* in terms of modularity and simplicity.

This new use of the combinatorial theory created and tuned by Busch and Mavronicolas, along with its original uses in deriving impossibility results and designing verification algorithms for balancing networks, strengthens the evidence that this theory provides the right framework for a systematic study of balancing networks.

1 Introduction

Consider a situation where we have p producers and c consumers. The producers produce jobs at some arbitrary rate; the jobs should be performed by the consumers. One would like to distribute the jobs as evenly as possible among the consumers. A very simple way to solve this problem makes use of a *counter*. Each time a new job is produced, the producer accesses the counter, increases it, and places the new job in a given array according to the value obtained from the counter. Consumer numbered l periodically checks the array locations $l, c+l, 2c+l, \ldots$, and whenever any of them contains a new job, the consumer performs it. Clearly, the difference in the total number of jobs eventually performed by any two consumers is at most one.

A counter can be easily implemented using a single shared *Fetch&Increment* variable. However, empirically, the time to access a shared variable grows at least linearly with the *contention*, the extent to which concurrent processors simultaneously access the variable.* In a seminal paper, Aspnes, Herlihy and Shavit [5] suggest a completely different approach to such counting problems. Their idea is to use a collection of shared variables called *balancers*, each having low expected contention, in a way that a processor needs to access only a few variables in order to obtain a value from the counter. Loosely speaking, a balancer can be thought of as a two-input, two-output toggle. When an input appears on one of its input wires, it takes the output wire to which the toggle is set, and toggles the gate so that the input next to come will leave on the other output wire. If the balancer is initialized so that the first input to pass through will exit on the top output wire, then, after m inputs have passed through the toggle, exactly $\lceil m/2 \rceil$ will exit on the top output wire, and $\lfloor m/2 \rfloor$ will exit on the bottom output wire. On a shared-memory multi-processor machine, a balancer can be implemented by a single bit *Compare&Swap* variable, and a wire can be implemented by a memory address pointer.

One can "connect" a collection of balancers to form a *balancing network,* much in the same way a sorting network is obtained by connecting a collection of comparators (see, e.g., [17]). This is done by connecting output wires from some balancers to input wires of others. The remaining unconnected input and output wires are the input and output wires, respectively, of the network. Each request for a counter value corresponds to a traversal of the network by a *token*, starting from some input wire, following the pointer obtained by accessing the first balancer to the next one, and so on. Let x_i and y_j denote the number of tokens that have entered the network on the ith input wire and left the network on the jth output wire, $0 \leq i, j \leq w-1$, respectively, where w is the *width* of the network. A balancing network of width w is a *counting network* if each time the network becomes free of tokens, i.e., all entering tokens have exited, $0 \leq y_i - y_j \leq 1$, for any i, j, $0 \leq i < j \leq w-1$; that is, the output has the *step* property. It is often only required that the output have the step property just in case the input is *block-step,* that is, in case each of two specified input subsequences has the step property. Networks satisfying this weaker property are called *merging networks,* in direct

*The cost of contention varies according to the architecture of the system and the specific arbitration protocols used (cf. [4]).

analogy to corresponding comparator networks that "merge" two sorted input sequences [17].

Aspnes *et al.* [5] present the first constructions of counting networks, both with width 2^k for any integer $k \geq 1$; these constructions have layouts isomorphic to the *bitonic* sorting network of Batcher [6] and the *periodic* sorting network of Dowd *et al.* [11], respectively. The correctness proof for each of these constructions have been carried out through considering all possible executions of tokens in the network; these proofs do not appear to provide much insight into any possible structural or combinatorial properties of these networks. Subsequently, many other constructions of counting networks and their variations have been presented (see, e.g., [1, 2, 12, 13, 15, 16]), but each of the correctness proofs for these constructions seems to require a different argument about patterns of token executions for each specific case.

In an effort towards understanding how "external" properties of balancing networks, like, e.g., the step property on outputs, come out as a result of "internal" combinatorial structure, the present authors develop in [8] a systematic, mathematical theory of the combinatorial structure of balancing networks. They propose a matrix representation of a balancing network which relies on its relative interconnections. More specifically, they introduce the *connection matrix* and *order vector* to describe the relation between inputs and outputs for each of the balancers in a *layer,* a balancing network of depth one.[†] In this way, a balancing network is represented by a collection of pairs of a connection matrix and an order vector, one pair for each layer.

For a wide spectrum of properties of balancing networks, Busch and Mavronicolas [8] provide tight combinatorial characterization theorems for classes of balancing networks possessing each of the properties. These characterizations theorems provide necessary and sufficient conditions on the connection matrices and order vectors for the property to hold. In most of the cases, these conditions say that, roughly speaking, the network uniformly assigns the input part corresponding to the most significant digits of inputs on its output wires, while the property is inherited down to the network's response to the input part corresponding to the least significant digits of inputs. In turn, these conditions have given rise to impossibility results for corresponding classes of balancing networks [8, Section 5], and formal algorithms to mathematically verify that a balancing network in hand belongs to a certain class of networks [8, Section 6].

In this work, we present yet another application of the combinatorial theory presented in [8]. We suggest a paradigmatic methodology for showing correctness of general constructions of balancing networks that belong to a certain class. This methodology consists of verifying that a general construction satisfies the necessary and sufficient conditions involved in the combinatorial characterization theorems for the corresponding class of networks, shown in [8].

We apply our methodology on the concrete example of the bitonic network construction [5]. We obtain a new proof that the bitonic network is a counting network; this proof employs a routine verification of the necessary and sufficient conditions involved in the combinatorial characterization theorems for counting and merging networks, shown in [8]. Although Aspnes

[†]The *depth* of a balancing network is the length of the longest path from an input wire to an output wire.

et al. [5] already present a corresponding proof that the bitonic network is a counting network, we feel that our proof has some additional interesting features compared to the one in [5]: first, it is simple and modular, while the one in [5] is rather ad-hoc and not so structured; most important, our proof yields significant insight into the combinatorial structure of the bitonic network. It precisely determines the combinatorial *transfer parameters* [8] of the bitonic network and pins down the exact properties of its construction that enforce these parameters to satify the conditions in the combinatorial characterization theorems for counting and merging networks shown in [8]. Furthermore, showing that these theorems hold for the bitonic network construction reveals that the *bitonic merger* network, a building block of the bitonic network, actually satisfies properties additional to those known so far: our analysis precisely identifies classes of inputs which, though not block-step, result in a step output when filtered through the bitonic merger.

The rest of this paper is organized as follows. Section 2 introduces some definitions and preliminary facts. In Sections 3 and 4, we present some mathematical preliminaries, in particular, some combinatorial properties of step vectors and block-step vectors, respectively. It turns out that some of the main arguments in our later correctness proof are nothing but restatements of these general properties. In Section 5, we provide an outline of the combinatorial theory of balancing networks presented in [8]. In Section 6, we present the bitonic network and some preliminary properties of it, while the formal proof that the bitonic network is a counting network appears in Section 7. We conclude, in Section 8, with a discussion of our work and directions for further research.

2 Definitions and Preliminaries

Fix throughout any integer $w \geq 2$; $\mathbf{X}^{(w)}$ will denote the vector $\langle x_0, x_1, \ldots, x_{w-1} \rangle^T$, and $\lceil \mathbf{X}^{(w)} \rceil$ and $\lfloor \mathbf{X}^{(w)} \rfloor$ will denote the vectors $\langle \lceil x_0 \rceil, \lceil x_1 \rceil, \ldots, \lceil x_{w-1} \rceil \rangle^T$ and $\langle \lfloor x_0 \rfloor, \lfloor x_1 \rfloor, \ldots, \lfloor x_{w-1} \rfloor \rangle^T$, respectively. Denote by $[w]$ the index set $\{0, 1, \ldots, w-1\}$. The *1-norm function* $\|.\|_1 : \Re^w \rightarrow \Re$ is defined as: $\|\mathbf{X}^{(w)}\|_1 = \sum_{i=0}^{w-1} |x_i|$.

Fix an integer $k \geq 1$. For any integer $x \geq 0$, define

$$x \downarrow_2 k \;=\; x - \left\lfloor \frac{x}{2^k} \right\rfloor 2^k,$$

and

$$x \uparrow_2 k \;=\; \left\lfloor \frac{x}{2^k} \right\rfloor 2^k \;.$$

Notice that $x \downarrow_2 k$ is the integer represented by the k least significant binary digits of x, while $x \uparrow_2 k$ is the integer obtained from x by setting each of these digits to zero. Clearly, $x \downarrow_2 k + x \uparrow_2 k = x$. Define also

$$x \updownarrow_2 k \;=\; x \downarrow_2 (k+1) - x \downarrow_2 k \;=\; x \uparrow_2 k - x \uparrow_2 (k+1) \;=\; \left\lfloor \frac{x}{2^k} \right\rfloor 2^k - \left\lfloor \frac{x}{2^{k+1}} \right\rfloor 2^{k+1} \;.$$

Notice that $x \updownarrow_2 k$ is the integer represented by the kth least significant binary digit of x. Thus, either $x \updownarrow_2 k = 0$ or $x \updownarrow_2 k = 2^{k-1}$, according to whether the kth least significant binary digit of x is 0 or 1, respectively. We will sometimes abuse notation and use $x \downarrow_2 k$, $x \uparrow_2 k$ and $x \updownarrow_2 k$ to denote the corresponding binary representations.

Extend the definitions for $x \downarrow_2 k$, $x \uparrow_2 k$ and $x \updownarrow_2 k$ from integers to any vector of integers $\mathbf{X}^{(w)}$ to obtain:

$$\begin{aligned}
\mathbf{X}^{(w)} \downarrow_2 k &= \langle x_0 \downarrow_2 k, x_1 \downarrow_2 k, \ldots, x_{w-1} \downarrow_2 k \rangle^T, \\
\mathbf{X}^{(w)} \uparrow_2 k &= \langle x_0 \uparrow_2 k, x_1 \uparrow_2 k, \ldots, x_{w-1} \uparrow_2 k \rangle^T, \text{ and} \\
\mathbf{X}^{(w)} \updownarrow_2 k &= \langle x_0 \updownarrow_2 k, x_1 \updownarrow_2 k, \ldots, x_{w-1} \updownarrow_2 k \rangle^T.
\end{aligned}$$

Assume henceforth that w is a power of two and fix any vector $\mathbf{X}^{(w)}$. Let $\mathbf{X}_{up}^{(w/2)}$ and $\mathbf{X}_{down}^{(w/2)}$ denote the vectors $\langle x_0, x_1, \ldots, x_{w/2-1} \rangle^T$ and $\langle x_{w/2}, x_{w/2+1}, \ldots, x_{w-1} \rangle^T$, respectively. Let also $\mathbf{X}_{e}^{(w/2)}$ and $\mathbf{X}_{o}^{(w/2)}$ denote the vectors $\langle x_0, x_2, \ldots, x_{w-2} \rangle$ and $\langle x_1, x_3, \ldots, x_{w-1} \rangle$, respectively. That is, the vectors $\mathbf{X}_{e}^{(w/2)}$ and $\mathbf{X}_{o}^{(w/2)}$ contain the even and odd, respectively, entries of $\mathbf{X}^{(w)}$. Finally, let $\mathbf{X}_{eo}^{(w/2)}$ and $\mathbf{X}_{oe}^{(w/2)}$ denote the vectors

$$\langle x_0, x_2, \ldots, x_{w/2-2}, x_{w/2+1}, x_{w/2+3}, \ldots, x_{w-1} \rangle$$

and

$$\langle x_1, x_3, \ldots, x_{w/2-1}, x_{w/2}, x_{w/2+2}, \ldots, x_{w-2} \rangle \ ,$$

respectively. That is, the vector $\mathbf{X}_{eo}^{(w/2)}$ represents the concatenation of the vector of even entries of $\mathbf{X}_{up}^{(w/2)}$ with the vector of odd entries of $\mathbf{X}_{down}^{(w/2)}$, while the vector $\mathbf{X}_{oe}^{(w/2)}$ represents the concatenation of the vector of odd entries of $\mathbf{X}_{up}^{(w/2)}$ with the vector of even entries of $\mathbf{X}_{down}^{(w/2)}$.

Corresponding to the definitions for $\mathbf{X}_{up}^{(w/2)}$, $\mathbf{X}_{down}^{(w/2)}$, $\mathbf{X}_{e}^{(w/2)}$, $\mathbf{X}_{o}^{(w/2)}$, $\mathbf{X}_{eo}^{(w/2)}$ and $\mathbf{X}_{oe}^{(w/2)}$, we define index sets $up[w] = \{0, 1, \ldots, w/2 - 1\}$, $down[w] = \{w/2, w/2 + 1, \ldots, w - 1\}$, $e[w] = \{0, 2, \ldots, w-2\}$, $o[w] = \{1, 3, \ldots, w-1\}$, $eo[w] = \{0, 2, \ldots, w/2-2, w/2+1, w/2+3, \ldots, w-1\}$ and $oe[w] = \{1, 3, \ldots, w/2 - 1, w/2, w/2 + 2, \ldots, w - 2\}$.

For a vector $\mathbf{X}^{(w/2)} = \langle x_0, x_1, x_{w/2-1} \rangle^T$, denote by $\mathbf{2X}^{(w)}$ the vector

$$\langle x_0, x_0, x_1, x_1, \ldots, x_{w/2-1}, x_{w/2-1} \rangle^T;$$

that is, the vector $\mathbf{2X}^{(w)}$ results from $\mathbf{X}^{(w/2)}$ by appending in order each of the entries of $\mathbf{X}^{(w/2)}$ to itself.

Define the vectors $\mathbf{0}^{(w)}$ and $\mathbf{1}^{(w)}$ to be $\langle 0, 0, \ldots, 0 \rangle^T$ and $\langle 1, 1, \ldots, 1 \rangle^T$, respectively, each with w entries. Define also the vectors $\mathbf{E}^{(w)}$ and $\mathbf{O}^{(w)}$ as follows: $e_i = 1/2$ for $i \in e[w]$ and 0 for $i \in o[w]$, while $o_i = 0$ for $i \in e[w]$ and $1/2$ for $i \in o[w]$; that is, all non-zero entries of $\mathbf{E}^{(w)}$ and $\mathbf{O}^{(w)}$, the even and odd ones, respectively, are equal to $1/2$.

We will sometimes adopt Iverson's notation (see, e.g. [18]) and write logical expressions in parentheses to mean one if true and zero if false. Thus, for example, for any number x and condition $\mathrm{P}(x)$ on x, the quantity $x + (\mathrm{P}(x))$ is equal to either $x + 1$ if x satisfies $\mathrm{P}(x)$ or x if x does not satisfy $\mathrm{P}(x)$.

Our last Proposition provides elementary identities involving the ceil function that follow immediately from the definition of this function.

Proposition 2.1 *For any integers x and i,*

(1) $\left\lceil x - \frac{1}{2}(i \text{ is odd}) \right\rceil = x;$

(2) $\left\lceil x - \frac{1}{2} - \frac{1}{2}(i \text{ is odd}) \right\rceil = x - (i \text{ is odd});$

(3) $\left\lceil x + \frac{1}{2}(i \text{ is even}) \right\rceil = x - (i \text{ is even});$

(4) $\left\lceil x - \frac{1}{2} + \frac{1}{2}(i \text{ is even}) \right\rceil = x,$ *and:*

(5) $\left\lceil x + \frac{1}{2} + \frac{1}{2}(i \text{ is even}) \right\rceil = x + 1.$

3 Step Vectors

In this Section, we formally define step vectors and show several combinatorial properties of them. The reader may prefer to skip this Section for now, returning to it later when its results are required.

We say that an integer vector $\mathbf{X}^{(w)}$ is *step* if for any i and k, $0 \leq i < k \leq w - 1$,

$$0 \leq x_i - x_k \leq 1 .$$

Denote by $step(\mathbf{N}^w)$ the set of all step vectors with w entries. The next Proposition provides an equivalent condition for a step vector:

Proposition 3.1 (Aspnes, Herlihy and Shavit [5]) *An integer vector $\mathbf{X}^{(w)}$ is step if and only if there exists some index c_x, $0 < c_x \leq w$, such that $x_i = x_0$ for all i such that $0 \leq i < c_x$ and $x_i = x_{c-1} - 1$ for all i such that $c_x \leq i \leq w - 1$.*

Call the quantities x_0 and c_x in Proposition 3.1 the *step value* and *step index,* respectively, of the vector $\mathbf{X}^{(w)}$. Clearly, the special case where the index c_x is equal to w corresponds to the case where $x_i = x_0$ for all $i \in [w]$. Such a step vector is called a *constant* vector.

Consider step vectors $\mathbf{X}^{(w)}$ and $\mathbf{Z}^{(w)}$ with step values x_0 and z_0, and step indices c_x and c_z, respectively. Since, by Proposition 3.1, $0 < c_x \leq w$ and $0 < c_z \leq w$, it follows that $|c_x - c_z| \leq w - 1$; it also follows that

$$\|\mathbf{X}^{(w)}\|_1 \;=\; c_x x_0 + (w - c_x)(x_0 - 1) \;=\; w(x_0 - 1) + c_x \;,$$

and

$$\|\mathbf{Z}^{(w)}\|_1 \;=\; c_z z_0 + (w - c_z)(z_0 - 1) \;=\; w(z_0 - 1) + c_z \;,$$

so that, in conclusion:

Claim 3.2 *For any vectors* $\mathbf{X}^{(w)}$ *and* $\mathbf{Z}^{(w)}$,

(1) $|\|\mathbf{X}^{(w)}\|_1 - \|\mathbf{Z}^{(w)}\|_1| \;=\; |w(x_0 - z_0) + c_x - c_z|$, *and*

(2) $|c_x - c_z| \leq w - 1$.

Our next two Propositions show certain interesting dependencies between two step vectors whose 1-norms' difference may only attain a value in a specific set. We first prove:

Proposition 3.3 *Assume each of the vectors* $\mathbf{X}^{(w)}$ *and* $\mathbf{Z}^{(w)}$ *is step and* $|\|\mathbf{X}^{(w)}\|_1 - \|\mathbf{Z}^{(w)}\|_1| \in \{0, 1\}$. *Then, either* $x_0 = z_0$ *and* $|c_x - c_z| \in \{0, 1\}$, *or* $|x_0 - z_0| = 1$ *and* $|c_x - c_z| = w - 1$.

Proof: We start by proving that no case other than $x_0 = z_0$ and $|x_0 - z_0| = 1$ is possible regarding x_0 and z_0.

Lemma 3.4 $|x_0 - z_0| \leq 1$

Proof: Assume, by way of contradiction, that $|x_0 - z_0| > 1$. There are two cases.

Take first $x_0 - z_0 > 1$. Since, by Claim 3.2(2), $c_x - c_z \geq -w + 1$, this implies that $w(x_0 - z_0) + c_x - c_z > w + -w + 1 = 1$. Hence, it follows by Claim 3.2(1) that $|\|\mathbf{X}^{(w)} - \mathbf{Y}^{(w)}\|_1| = |w(x_0 - z_0) + c_x - c_z| > 1$, a contradiction.

Take now $x_0 - z_0 < -1$. Since, by Claim 3.2(2), $c_x - c_z \leq w - 1$, this implies that $w(x_0 - z_0) + c_x - c_z < -w + w - 1 = -1$. Hence, it follows by Claim 3.2(1) that $|\|\mathbf{X}^{(w)} - \mathbf{Y}^{(w)}\|_1| = |w(x_0 - z_0) + c_x - c_z| > 1$, a contradiction. This completes our proof. ■

Lemma 3.4 implies that either $|x_0 - z_0| = 0$ or $|x_0 - z_0| = 1$. We proceed by case analysis.

1. Assume first that $|x_0 - z_0| = 0$ so that $x_0 = z_0$. By Claim 3.2(1), it follows that $|\|\mathbf{X}^{(w)}\|_1 - \|\mathbf{Z}^{(w)}\|_1| = |c_x - c_z|$. Since $|\|\mathbf{X}^{(w)}\|_1 - \|\mathbf{Z}^{(w)}\|_1 \in \{0, 1\}$, it follows that $|c_x - c_z| \in \{0, 1\}$, as needed.

2. Assume now that $|x_0 - z_0| = 1$. Since $|\|\mathbf{X}^{(w)}\|_1 - \|\mathbf{Z}^{(w)}\|_1 \in \{0,1\}$, it follows by Claim 3.2(1) that $-1 - w(x_0 - z_0) \leq c_x - c_z \leq 1 - w(x_0 - z_0)$.

   Take first $x_0 - z_0 = 1$ so that $c_x - c_z \leq 1 - w$. Since, by Claim 3.2(2), $|c_x - c_z| \leq w - 1$, it follows that $c_x - c_z = 1 - w$. Take now $x_0 - z_0 = -1$ so that $c_x - c_z \geq w - 1$. Since, by Claim 3.2(2), $|c_x - c_z| \leq w - 1$, it follows that $c_x - c_z = w - 1$. Thus, in either case $|x_0 - z_0| = w - 1$, as needed.

■

We continue by showing:

Proposition 3.5 *Assume each of the vectors* $\mathbf{X}^{(w)}$ *and* $\mathbf{Z}^{(w)}$ *is step and* $|\|\mathbf{X}^{(w)}\|_1 - \|\mathbf{Z}^{(w)}\|_1| \in \{w-1, w\}$. *Then, either* $x_0 = z_0$ *and* $|c_x - c_z| = w - 1$, *or* $|x_0 - z_0| = 1$ *and* $|c_x - c_z| \in \{0, 1\}$.

Proof: We start by proving that no case other than $x_0 = z_0$ and $|x_0 - z_0| = 1$ is possible regarding x_0 and z_0.

Lemma 3.6 $|x_0 - z_0| \leq 1$

Proof: Assume, by way of contradiction, that $|x_0 - z_0| > 1$. There are two cases.

Take first $x_0 - z_0 \geq 2$. Since, by Claim 3.2(2), $c_x - c_z \geq -w + 1$, it follows that $w(x_0 - z_0) + c_x - c_z \geq 2w - w + 1 = w + 1$. Hence, it follows by Claim 3.2(1) that $|\|\mathbf{X}^{(w)}\|_1 - \mathbf{Y}^{(w)}\|_1| \geq w + 1$, a contradiction.

Take now $x_0 - z_0 \leq -2$. Since, bu Claim 3.2(2), $c_x - c_z \leq w - 1$, it follows that $w(x_0 - z_0) + c_x - c_z \leq -2w + w - 1 = -w - 1$. Hence, it follows by Claim 3.2(1) that $|\|\mathbf{X}^{(w)}\|_1 - \mathbf{Y}^{(w)}\|_1| \geq w + 1$, a contradiction. This completes our proof. ■

Lemma 3.6 implies that either $|x_0 - z_0| = 0$ or $|x_0 - z_0| = 1$. We proceed by case analysis.

1. Assume first that $|x_0 - z_0| = 0$, i.e., $x_0 = z_0$. Since $|\|\mathbf{X}^{(w)}\|_1 - \|\mathbf{Y}^{(w)}\|_1| \in \{w-1, w\}$, it follows by Claim 3.2(1) that $|c_x - c_z| \in \{w-1, w\}$. Since, by Claim 3.2(2), $|c_x - c_z| \leq w - 1$, it follows that $|c_x - c_z| = w - 1$, as needed.

2. Assume now that $|x_0 - z_0| = 1$. Since $|\|\mathbf{X}^{(w)}\|_1 - \|\mathbf{Z}^{(w)}\|_1 \in \{w-1, w\}$, it follows by Claim 3.2(1) that $|w(x_0 - z_0) + c_x - c_z| \in \{w-1, w\}$, where $x_0 - z_0 \in \{-1, 1\}$. There are two cases.

   Take first $x_0 - z_0 = -1$ so that $|-w + c_x - c_z| \in \{w-1, w\}$. If $|-w + c_x - c_z| = w - 1$, then either $-w + c_x - c_z = w - 1$, implying $c_x - c_z = 2w - 1$, which is not possible because, by Claim 3.2(2), $|c_x - c_z| \leq w - 1$, or $-w + c_x - c_z = -w + 1$, implying $c_x - c_z = 1$. If

	$\|c_x - c_z\| \in \{0,1\}$	$\|c_x - c_z\| = w-1$
$\|x_0 - z_0\| = 0$	$\{0,1\}$	$\{w-1, w\}$
$\|x_0 - z_0\| = 1$	$\{w-1, w\}$	$\{0,1\}$

Table 1: Summary of Propositions 3.3 and 3.5

$|-w + c_x - c_z| = w$, then either $-w + c_x - c_z = w$, implying $c_x - c_z = 2w$, which is not possible because, by Claim 3.2(2), $|c_x - c_z| \leq w-1$, or $-w + c_x - c_z = -w$, implying $c_x - c_z = 0$. Thus, for $x_0 - z_0 = -1$, $|c_x - c_z| \in \{0,1\}$.

Take now $x_0 - z_0 = 1$ so that $|w + c_x - c_z| \in \{w-1, w\}$. If $|w + c_x - c_z| = w-1$, then either $w + c_x - c_z = w-1$, implying $c_x - c_z = -1$, or $w + c_x - c_z = -w+1$, implying $c_x - c_z = -2w+1$, which is not possible because, by Claim 3.2(2), $|c_x - c_z| \leq w-1$. If $|w + c_x - c_z| = w$, then either $w + c_x - c_z = w$, implying $c_x - c_z = 0$, or $w + c_x - c_z = -w$, implying $c_x - c_z = -2w$, which is not possible because, by Claim 3.2(2), $|c_x - c_z| \leq w-1$. Thus, for $x_0 - z_0 = 1$, $|c_x - c_z| \in \{0,1\}$.

Hence, for $|x_0 - z_0| = 1$, $|c_x - c_z| \in \{0,1\}$, as needed.

■

Table 1 summarizes Propositions 3.3 and 3.5. For each condition on $|x_0 - z_0|$ along the left side and each condition on $|c_x - c_z|$ across the top, the appropriate entry provides the range of values of $|\|\mathbf{X}^{(w)}\|_1 - \|\mathbf{Y}^{(w)}\|_1|$ for which these conditions simultaneously hold, assuming $\mathbf{X}^{(w)}$ and $\mathbf{Y}^{(w)}$ are step.

In our next two Propositions, we still consider a pair of step vectors whose 1-norms' difference may only attain a value in a specific set, and show that certain combinations of these step vectors are also step. We start by proving:

Proposition 3.7 *Assume each of the vectors* $\mathbf{X}^{(w/2)}$ *and* $\mathbf{Z}^{(w/2)}$ *is step and* $|\|\mathbf{X}^{(w/2)}\|_1 - \|\mathbf{Z}^{(w/2)}\|_1| \in \{0,1\}$. *Then, the vector* $\left\lceil \frac{1}{2}\mathbf{2}(\mathbf{X}^{(w/2)} + \mathbf{Z}^{(w/2)}) - \mathbf{O}^{(w)} \right\rceil$ *is step.*

Proof: By Proposition 3.3, either $x_0 = z_0$ and $|c_x - c_z| \in \{0,1\}$, or $|x_0 - z_0| = 1$ and $|c_x - c_z| = w/2 - 1$. We proceed by case analysis.

1. Assume first that $x_0 = z_0$ and $|c_x - c_z| \in \{0,1\}$. There are two cases.

   (a) Take first $c_x = c_z$ so that $|c_x - c_z| = 0$ Then,

$$(x+z)_i \quad = \quad \begin{cases} 2x_0, & 0 \leq i < c_x \\ 2x_0 - 2, & c_x \leq i \leq w/2 - 1 \end{cases},$$

so that

$$\frac{1}{2}(2(x+z))_i = \begin{cases} x_0, & 0 \le i \le 2c_x - 1 \\ x_0 - 1, & 2c_x - 1 < i \le w - 1 \end{cases},$$

and

$$\left\lceil (\frac{1}{2}(2(x+z)) - o)_i \right\rceil = \begin{cases} \left\lceil x_0 - \frac{1}{2}(i \text{ is odd}) \right\rceil, & 0 \le i \le 2c_x - 1 \\ \left\lceil x_0 - 1 - \frac{1}{2}(i \text{ is odd}) \right\rceil, & 2c_x - 1 < i \le w - 1 \end{cases}.$$

$$= \begin{cases} x_0, & 0 \le i \le 2c_x - 1 \\ x_0 - 1, & 2c_x - 1 < i \le w - 1 \end{cases},$$

by Proposition 2.1(1).

(b) Take now $|c_x - c_z| = 1$. Without loss of generality, let $c_x - c_z = -1$. We have:

$$(x+z)_i = \begin{cases} 2x_0, & 0 \le i < c_x \\ 2x_0 - 1, & i = c_x \\ 2x_0 - 2, & c_x < i \le \frac{w}{2} - 1 \end{cases},$$

so that

$$\frac{1}{2}(2(x+z))_i = \begin{cases} x_0, & 0 \le i \le 2c_x - 1 \\ x_0 - \frac{1}{2}, & 2c_x \le i \le 2c_x + 1 \\ x_0 - 1, & 2c_x + 1 < i \le w - 1 \end{cases},$$

and

$$\left\lceil (\frac{1}{2}(2(x+z)) - o)_i \right\rceil = \begin{cases} \left\lceil x_0 - \frac{1}{2}(i \text{ is odd}) \right\rceil, & 0 \le i \le 2c_x - 1 \\ \left\lceil x_0 - \frac{1}{2} - \frac{1}{2}(i \text{ is odd}) \right\rceil, & 2c_x \le i \le 2c_x + 1 \\ \left\lceil x_0 - 1 - \frac{1}{2}(i \text{ is odd}) \right\rceil, & 2c_x + 1 < i \le w - 1 \end{cases},$$

$$= \begin{cases} x_0, & 0 \le i \le 2c_x \\ x_0 - 1, & 2c_x + 1 \le i \le w - 1 \end{cases},$$

by Proposition 2.1(1) and (2).

2. Assume now that $|x_0 - z_0| = 1$. Inspecting the proof of Proposition 3.3 (case 2) reveals that either $x_0 - z_0 = 1$ and $c_x - c_z = 1 - w/2$, or $x_0 - z_0 = -1$ and $c_x - c_z = w/2 - 1$. The two cases being symmetrical, we consider, without loss of generality, only the case where $x_0 - z_0 = 1$ and $c_x - c_z = 1 - w/2$. Since, by Claim 3.2(2), $|c_x - c_z| \le w/2 - 1$, the latter equality implies that $c_z = w/2$ and $c_x = 1$, so that $z_i = z_0$ for all $i \in [w/2]$, and $x_i = x_0 - 1$ for all i such that $1 \le i \le w/2 - 1$. Thus, we have:

$$(x+z)_i = \begin{cases} x_0 + z_0, & i = 0 \\ x_0 + z_0 - 1, & 1 \le i \le w/2 - 1 \end{cases}$$

$$= \begin{cases} 2x_0 - 1, & i = 0 \\ 2x_0 - 2, & 1 \le i \le w/2 - 1 \end{cases}$$

since $z_0 = x_0 - 1$, so that

$$\frac{1}{2}(2(x+z))_i = \begin{cases} x_0 - \frac{1}{2}, & 0 \le i \le 1 \\ x_0 - 1, & 1 < i \le w-1 \end{cases},$$

and

$$\begin{aligned} \left\lceil (\frac{1}{2}(2(x+z)) - o)_i \right\rceil &= \begin{cases} \left\lceil x_0 - \frac{1}{2} - \frac{1}{2}(i \text{ is odd}) \right\rceil, & 0 \le i \le 1 \\ \left\lceil x_0 - 1 - \frac{1}{2}(i \text{ is odd}) \right\rceil, & 1 < i \le w-1 \end{cases} \\ &= \begin{cases} x_0, & i = 0 \\ x_0 - 1, & 1 \le i \le w-1 \end{cases}, \end{aligned}$$

by Proposition 2.1(1) and (2).

The previous case analysis reveals that the vector $\left\lceil \frac{1}{2}\mathbf{2}(\mathbf{X}^{(w/2)} + \mathbf{Z}^{(w/2)}) - \mathbf{O}^{(w)} \right\rceil$ is step in all cases, as needed. ■

We continue by proving:

Proposition 3.8 *Assume each of the vectors $\mathbf{X}^{(w/2)}$ and $\mathbf{Z}^{(w/2)}$ is step and $|\|\mathbf{X}^{(w/2)}\|_1 - \|\mathbf{Z}^{(w/2)}\|_1| \in \{w/2-1, w/2\}$. Then, the vector $\left\lceil \frac{1}{2}\mathbf{2}(\mathbf{X}^{(w/2)} + \mathbf{Z}^{(w/2)}) + \mathbf{E}^{(w)} \right\rceil$ is step.*

Proof: By Proposition 3.5, either $x_0 = z_0$ and $|c_x - c_z| = w/2 - 1$, or $|x_0 - z_0| = 1$ and $|c_x - c_z| \in \{0, 1\}$. We proceed by case analysis.

1. Assume first that $x_0 = z_0$ and $|c_x - c_z| = w/2 - 1$. Without loss of generality, let $c_x - c_z = w/2 - 1$. Since, by Claim 3.2(2), $|c_x - c_z| \le w/2$, this implies that $c_x = w/2$ and $c_z = 1$, so that $x_i = x_0$ for all $i \in [w/2]$ and $z_i = x_0 - 1$ for all i such that $1 \le i \le w/2 - 1$. Thus, we have:

$$\begin{aligned} (x+z)_i &= \begin{cases} x_0 + z_0, & i = 0 \\ x_0 + z_0 - 1, & 1 \le i \le w/2 - 1 \end{cases} \\ &= \begin{cases} 2x_0, & i = 0 \\ 2x_0 - 1, & 1 \le i \le w/2 - 1 \end{cases}, \end{aligned}$$

since $z_0 = x_0$, so that

$$\frac{1}{2}(2(x+z))_i = \begin{cases} x_0, & 0 \le i \le 1 \\ x_0 - 1/2, & 1 < i \le w-1 \end{cases},$$

and

$$\left\lceil (\frac{1}{2}(2(x+z)) - o)_i \right\rceil = \begin{cases} \left\lceil x_0 + \frac{1}{2}(i \text{ is even}) \right\rceil, & 0 \leq i \leq 1 \\ \left\lceil x_0 - \frac{1}{2} + \frac{1}{2}(i \text{ is even}) \right\rceil, & 1 < i \leq w-1 \end{cases}$$
$$= \begin{cases} x_0 + 1, & i = 0 \\ x_0, & 1 \leq i \leq w-1 \ , \end{cases}$$

by Proposition 2.1(3) and (4).

2. Assume now that $|x_0 - z_0| = 1$ and $|c_x - c_z| \in \{0,1\}$. Inspecting the proof of Proposition 3.5 (case 2) reveals that either $x_0 - z_0 = -1$ and $c_x - c_z \in \{0,1\}$, or $x_0 - z_0 = 1$ and $c_z - c_x \in \{0,1\}$. The two cases being symmetrical, we consider, without loss of generality, only the case where $x_0 - z_0 = -1$ and $c_x - c_z \in \{0,1\}$. We proceed by case analysis on the value taken by $c_x - c_z$.

   (a) Assume first that $c_x - c_z = 0$. We have:

$$(x+z)_i = \begin{cases} x_0 + z_0, & 0 \leq i < c_x \\ x_0 + z_0 - 2, & c_x \leq i \leq w/2 - 1 \ , \end{cases}$$
$$= \begin{cases} 2x_0 + 1, & 0 \leq i < c_x \\ 2x_0 - 1, & c_x \leq i \leq w/2 - 1 \ , \end{cases}$$

   since $x_0 - z_0 = -1$, so that

$$\frac{1}{2}(2(x+z))_i = \begin{cases} x_0 + \frac{1}{2}, & 0 \leq i \leq 2c_x - 1 \\ x_0 - \frac{1}{2}, & 2c_x - 1 < i \leq w-1 \ , \end{cases}$$

   and

$$\left\lceil (\frac{1}{2}(2(x+z)) - o)_i \right\rceil = \begin{cases} \left\lceil x_0 + \frac{1}{2} + \frac{1}{2}(i \text{ is even}) \right\rceil, & 0 \leq i \leq 2c_x - 1 \\ \left\lceil x_0 - \frac{1}{2} + \frac{1}{2}(i \text{ is even}) \right\rceil, & 2c_x - 1 < i \leq w-1 \ , \end{cases}$$
$$= \begin{cases} x_0 + 1, & 0 \leq i \leq 2c_x - 1 \\ x_0, & 2c_x \leq i \leq w-1 \ , \end{cases}$$

   by Proposition 2.1(5) and (4).

   (b) Assume now that $c_x - c_z = 1$. We have:

$$(x+z)_i = \begin{cases} x_0 + z_0, & 0 \leq i < c_z \\ x_0 + z_0 - 1, & i = c_z \\ x_0 + z_0 - 2, & c_z < i \leq w/2 - 1 \ , \end{cases}$$
$$= \begin{cases} 2z_0 - 1, & 0 \leq i < c_z \\ 2z_0 - 2, & i = c_z \\ 2z_0 - 3, & c_z < i \leq w/2 - 1 \ , \end{cases}$$

since $x_0 = z_0 - 1$, so that

$$\frac{1}{2}(2(x+z))_i \;=\; \begin{cases} z_0 - \frac{1}{2}, & 0 \le i \le 2c_z - 1 \\ z_0 - 1, & 2c_z \le i \le 2c_z + 1 \\ z_0 - \frac{3}{2}, & 2c_z + 1 < i \le w - 1 \ , \end{cases}$$

and

$$\begin{aligned} \left\lceil (\frac{1}{2}(2(x+z)) + e)_i \right\rceil \;&=\; \begin{cases} \left\lceil z_0 - \frac{1}{2} + \frac{1}{2}(i \text{ is even}) \right\rceil, & 0 \le i \le 2c_z - 1 \\ \left\lceil z_0 - 1 + \frac{1}{2}(i \text{ is even}) \right\rceil, & 2c_z \le i \le 2c_z + 1 \\ \left\lceil x_0 - \frac{3}{2} + \frac{1}{2}(i \text{ is even}) \right\rceil, & 2c_z + 1 < i \le w - 1 \ , \end{cases} \\ &=\; \begin{cases} z_0, & 0 \le i \le 2c_z \\ z_0 - 1, & 2c_z + 1 \le i \le w - 1 \ , \end{cases} \end{aligned}$$

by Proposition 2.1(4) and (3).

The previous case analysis reveals that the vector $\left\lceil \frac{1}{2}\mathbf{2}(\mathbf{X}^{(w/2)} + \mathbf{Z}^{(w/2)}) + \mathbf{E}^{(w)} \right\rceil$ is step in all cases, as needed. ■

We continue by showing a necessary condition for $\mathbf{X}^{(w)} \uparrow_2 (d-1)$ in case both $\mathbf{X}^{(w)}$ and $\mathbf{X}^{(w)} \downarrow_2 (d-1)$ are step.

Proposition 3.9 *For any integer $d \ge 2$, assume both $\mathbf{X}^{(w)}$ and $\mathbf{X}^{(w)} \downarrow_2 (d-1)$ are step. Then, $\mathbf{X}^{(w)} \uparrow_2 (d-1)$ is a constant vector.*

Proof: Assume, by way of contradiction, that there are indices i and k, $i < k$, such that $x_i \uparrow_2 (d-1) \neq x_k \uparrow_2 (d-1)$.

By their definition, each of $x_i \uparrow_2 (d-1)$ and $x_k \uparrow_2 (d-1)$ is a multiple of 2^{d-1}. Hence, it follows that $x_i \uparrow_2 (d-1) - x_k \uparrow_2 (d-1) = f \cdot 2^{d-1}$, for some integer $f \neq 0$. We have:

$$\begin{aligned} x_i - x_k \;&=\; x_i \uparrow_2 (d-1) + x_i \downarrow_2 (d-1) - x_k \uparrow_2 (d-1) - x_k \downarrow_2 (d-1) \\ &=\; f \cdot 2^{d-1} + x_i \downarrow_2 (d-1) - x_k \downarrow_2 (d-1) \end{aligned}$$

Since $\mathbf{X}^{(w)} \downarrow_2 (d-1)$ is step, $0 \le x_i \downarrow_2 (d-1) - x_k \downarrow_2 (d-1) \le 1$. Hence, it follows that:

$$f \cdot 2^{d-1} \le x_i - x_k \;\le\; f \cdot 2^{d-1} + 1$$

We proceed by case analysis. Assume first that $f \ge 1$. It follows that $x_i - x_k \ge 2^{d-1} > 1$, since $d \ge 2$. Since $\mathbf{X}^{(w)}$ is step and $i < k$, $x_i - x_k \le 1$. A contradiction.

Assume now that $f \le -1$. It follows that $x_i - x_k \le -2^{d-1} + 1 \le -1$, since $d \ge 2$. Since $\mathbf{X}^{(w)}$ is step and $i < k$, $x_i - x_k \ge 0$. A contradiction. This completes the proof that $\mathbf{X}^{(w)} \uparrow_2 (d-1)$ is a constant vector. ■

In particular, Proposition 3.9 implies:

Corollary 3.10 *For any integer $d \geq 2$, assume both $\mathbf{X}^{(w)}$ and $\mathbf{X}^{(w)} \downarrow_2 (d-1)$ are step. Then, either $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1 = 0$ or $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1 = w2^{d-1}$.*

We continue with a necessary condition for $\mathbf{X}^{(w)} \downarrow_2 (d-1)$ in case $\mathbf{X}^{(w)}$ is step but $\mathbf{X}^{(w)} \downarrow_2 (d-1)$ is not step.

Proposition 3.11 *For any integer $d \geq 2$, assume $\mathbf{X}^{(w)}$ is step but $\mathbf{X}^{(w)} \downarrow_2 (d-1)$ is not step. Then, there exists some index c, $0 < c \leq w-1$, such that either*

(1) $x_0 \downarrow_2 d = \ldots = x_{c-1} \downarrow_2 d = 00\ldots0$ *and* $x_c \downarrow_2 d = \ldots = x_{w-1} \downarrow_2 d = 11\ldots1$, *or*

(2) $x_0 \downarrow_2 d = \ldots = x_{c-1} \downarrow_2 d = 100\ldots0$ *and* $x_c \downarrow_2 d = \ldots = x_{w-1} \downarrow_2 d = 011\ldots1$.

Proof: Since $\mathbf{X}^{(w)}$ is step but $\mathbf{X}^{(w)} \downarrow_2 (d-1)$ is not step, it follows by Proposition 3.1 that there exists some index c, $0 < c \leq w-1$, such that $x_i = x_0$ for all i such that $0 \leq i < c$ and $x_i = x_{c-1} - 1$ for all i such that $c \leq i \leq w-1$. It must be that $x_0 \downarrow_2 (d-1) = \ldots = x_{c-1} \downarrow_2 (d-1) = 00\ldots0$, since otherwise $x_c \downarrow_2 (d-1) = \ldots = x_{w-1} \downarrow_2 (d-1) = x_{c-1} \downarrow_2 (d-1) - 1$, contradicting the assumption that $\mathbf{X}^{(w)} \downarrow_2 (d-1)$ is not step.

We proceed by case analysis on the dth least significant binary digit of $x_0 = \ldots = x_{c-1}$. Assume first that this digit is 0, so that $x_0 \downarrow_2 d = \ldots = x_{c-1} \downarrow_2 d = 00\ldots0$. Since $x_i = x_{c-1} - 1$ for all i such that $i \leq c \leq w-1$, this implies that $x_c \downarrow_2 d = \ldots = x_{w-1} \downarrow_2 d = 11\ldots1$, as needed.

Assume now that the dth least significant binary digit of $x_0 = \ldots = x_{c-1}$ is 1, so that $x_0 \downarrow_2 d = \ldots = x_{c-1} \downarrow_2 d = 100\ldots0$. Since $x_i = x_{c-1} - 1$ for all i such that $c \leq i \leq w-1$, this implies that $x_c \downarrow_2 d = \ldots = x_{w-1} \downarrow_2 d = 011\ldots1$, as needed. This completes our proof. ■

4 Block-Step Vectors

In this Section, we formally define block-step vectors and show several combinatorial properties of them. The reader may prefer to skip this Section for now, returning to it later when its results are required.

We say that an integer vector $\mathbf{X}^{(w)}$ is *block-step* if both vectors $\mathbf{X}_{up}^{(w/2)}$ and $\mathbf{X}_{down}^{(w/2)}$ are step. Denote by $blockstep(\mathbf{N}^w)$ the set of all block-step vectors with w entries.

We first notice that certain sub-vectors of a block-step vector are also block-step.

Proposition 4.1 *Assume $\mathbf{X}^{(w)}$ is block-step. Then, both $\mathbf{X}_{eo}^{(w/2)}$ and $\mathbf{X}_{oe}^{(w/2)}$ are block-step.*

Proof: By definition of a block-step vector, both $\mathbf{X}_{up}^{(w/2)}$ and $\mathbf{X}_{down}^{(w/2)}$ are step. Proposition 3.1 immediately implies that both $(\mathbf{X}_{up})_e^{(w/4)}$ and $(\mathbf{X}_{up})_o^{(w/4)}$ (resp., $(\mathbf{X}_{down})_e^{(w/4)}$ and $(\mathbf{X}_{down})_o^{(w/4)}$) are step. Since $\mathbf{X}_{eo}^{(w/2)}$ (resp., $\mathbf{X}_{oe}^{(w/2)}$) is the concatenation of $(\mathbf{X}_{up})_e^{(w/4)}$ and $(\mathbf{X}_{down})_o^{(w/4)}$ (resp., $(\mathbf{X}_{up})_o^{(w/4)}$ and $(\mathbf{X}_{down})_e^{(w/4)}$), it follows by definition of a block-step vector that $\mathbf{X}_{eo}^{(w/2)}$ (resp., $\mathbf{X}_{oe}^{(w/2)}$) is block-step, as needed. ■

We next show that the 1-norms of certain subvectors of a block-step vector come close to each other.

Proposition 4.2 *Assume $\mathbf{X}^{(w)}$ is block-step. Then,*

$$|\|\mathbf{X}_{eo}^{(w/2)}\|_1 - \|\mathbf{X}_{oe}^{(w/2)}\|_1| \ \in \ \{0, 1\} \ .$$

Proof: Recall that by definitions of $\mathbf{X}_{eo}^{(w/2)}$ and $\mathbf{X}_{oe}^{(w/2)}$,

$$\|\mathbf{X}_{eo}^{(w/2)}\|_1 = \|(\mathbf{X}_{up})_e^{(w/4)}\|_1 + \|(\mathbf{X}_{down})_o^{(w/4)}\|_1 \ ,$$

and

$$\|\mathbf{X}_{oe}^{(w/2)}\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)}\|_1 + \|(\mathbf{X}_{down})_e^{(w/4)}\|_1 \ .$$

By definition of a block-step vector, both $\mathbf{X}_{up}^{(w/2)}$ and $\mathbf{X}_{down}^{(w/2)}$ are step. By Proposition 3.1, there are indices c_{up} and c_{down}, $0 < c_{up}, c_{down} \leq w/2$, such that:

- $(x_{up})_i = (x_{up})_0$ for all i such that $0 \leq i < c_{up}$ and $(x_{up})_i = (x_{up})_{c_{up}-1} - 1$ for all i such that $c_{up} \leq i \leq w/2 - 1$, and
- $(x_{down})_i = (x_{down})_0$ for all i such that $0 \leq i < c_{down}$ and $(x_{down})_i = (x_{down})_{c_{down}-1} - 1$ for all i such that $c_{down} \leq i \leq w/2 - 1$.

We proceed by case analysis on the parities of c_{up} and c_{down}:

1. Assume both c_{up} and c_{down} are even. Then, clearly, $\|(\mathbf{X}_{up})_e^{(w/4)}\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)}\|_1$ and $\|(\mathbf{X}_{down})_e^{(w/4)}\|_1 = \|(\mathbf{X}_{down})_o^{(w/4)}\|_1$, so that $\|\mathbf{X}_{eo}^{(w/2)}\|_1 = \|\mathbf{X}_{oe}^{(w/2)}\|_1$.
2. Assume both c_{up} and c_{down} are odd. Then, clearly, $\|(\mathbf{X}_{up})_e^{(w/4)}\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)}\|_1 + 1$ and $\|(\mathbf{X}_{down})_e^{(w/4)}\|_1 = \|(\mathbf{X}_{down})_o^{(w/4)}\|_1 + 1$, so that $\|\mathbf{X}_{eo}^{(w/2)}\|_1 = \|\mathbf{X}_{oe}^{(w/2)}\|_1$.
3. Assume c_{up} is odd and c_{down} is even. Then, clearly, $\|(\mathbf{X}_{up})_e^{(w/4)}\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)}\|_1 + 1$ and $\|(\mathbf{X}_{down})_e^{(w/4)}\|_1 = \|(\mathbf{X}_{down})_o^{(w/4)}\|_1$, so that $\|\mathbf{X}_{eo}^{(w/2)}\|_1 = \|\mathbf{X}_{oe}^{(w/2)}\|_1 + 1$.
4. Assume c_{up} is even and c_{down} is odd. Then, clearly, $\|(\mathbf{X}_{up})_e^{(w/4)}\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)}\|_1$ and $\|(\mathbf{X}_{down})_e^{(w/4)}\|_1 = \|(\mathbf{X}_{down})_o^{(w/4)}\|_1 + 1$, so that $\|\mathbf{X}_{eo}^{(w/2)}\|_1 + 1 = \|\mathbf{X}_{oe}^{(w/2)}\|_1$.

Thus, in all cases, $|\|\mathbf{X}_{eo}^{(w/2)}\|_1 - \|\mathbf{X}_{oe}^{(w/2)}\|_1| \in \{0,1\}$, as needed. ■

For each integer k, denote by $(\mathit{blockstep}(\mathbf{N}^w)) \downarrow_2 k$ the set of all integer vectors $\mathbf{X}^{(w)} \in [2^k]^w$ such that $\mathbf{X}^{(w)} = \mathbf{Y}^{(w)} \downarrow_2 k$ for some block-step vector $\mathbf{Y}^{(w)}$; that is, $(\mathit{blockstep}(\mathbf{N}^w)) \downarrow_2 k$ is the set of the restrictions to their k least significant binary digits of blockstep vectors with w entries. Notice that $(\mathit{blockstep}(\mathbf{N}^w)) \downarrow_2 k \subseteq [2^k]^w$.

Our first Proposition provides an analog to block-step vectors of Corollary 3.10.

Proposition 4.3 *For $w \geq 4$, assume $\mathbf{X}^{(w)} \in (\mathit{blockstep}(\mathbf{N}^w)) \downarrow_2 \lg w$ and $\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)$ is block-step. Then, $\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 / w$ is an integer.*

Proof: By definition of $(\mathit{blockstep}(\mathbf{N}^w)) \downarrow_2 \lg w$, there exists some block-step vector $\mathbf{V}^{(w)}$ such that $\mathbf{V}^{(w)} \downarrow_2 \lg w = \mathbf{X}^{(w)}$. By definition of a block-step vector, $\mathbf{V}_{up}^{(w)}$ (resp., $\mathbf{V}_{down}^{(w)}$) is step.

Since $\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)$ is block-step, it follows, by definition of a block-step vector, that $\mathbf{X}_{up}^{(w)} \downarrow_2 (\lg w - 1)$ (resp., $\mathbf{X}_{down}^{(w)} \downarrow_2 (\lg w - 1)$) is step.

Since $\mathbf{V}^{(w)} \downarrow_2 \lg w = \mathbf{X}^{(w)}$, it clearly follows that $\mathbf{V}_{up}^{(w)} \downarrow_2 (\lg w - 1) = \mathbf{X}_{up}^{(w)} \downarrow_2 (\lg w - 1)$ (resp., $\mathbf{V}_{down}^{(w)} \downarrow_2 (\lg w - 1) = \mathbf{X}_{down}^{(w)} \downarrow_2 (\lg w - 1)$) and $\|\mathbf{V}_{up}^{(w)} \updownarrow_2 \lg w\|_1 = \|\mathbf{X}_{up}^{(w)} \updownarrow_2 \lg w\|_1$ (resp., $\|\mathbf{V}_{down}^{(w)} \updownarrow_2 \lg w\|_1 = \|\mathbf{X}_{down}^{(w)} \updownarrow_2 \lg w\|_1$).

Hence, it follows by Corollary 3.10 that either $\|\mathbf{X}_{up}^{(w/2)} \updownarrow_2 \lg w\|_1 = 0$ or $\|\mathbf{X}_{up}^{(w/2)} \updownarrow_2 \lg w\|_1 = (w/2)2^{\lg w - 1} = w^2/4$ (resp., either $\|\mathbf{X}_{down}^{(w/2)} \updownarrow_2 \lg w\|_1 = 0$ or $\|\mathbf{X}_{down}^{(w/2)} \updownarrow_2 \lg w\|_1 = (w/2)2^{\lg w - 1} = w^2/4$). Since

$$\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 \;=\; \mathbf{X}_{up}^{(w/2)} \updownarrow_2 \lg w\|_1 + \mathbf{X}_{down}^{(w/2)} \updownarrow_2 \lg w\|_1 \;,$$

it follows that $\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 = 0$, $w^2/4$, or $w^2/2$. Since w is a power of two and at least four, it follows that $\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 / w$ is an integer, as needed. ■

Our next Proposition provides an analog to block-step vectors of Proposition 3.11:

Proposition 4.4 *For any integer $d \geq 2$, assume $\mathbf{X}^{(w)}$ is block-step, but $\mathbf{X}^{(w)} \downarrow_2 (d-1)$ is not block-step. Then, either*

$$|\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (d-1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (d-1)\|_1| \;\in\; \{0,1\}$$

and $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1 / 2^d = l$ for some integer l, or

$$|\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (d-1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (d-1)\|_1| \;\in\; \{w/2-1, w/2\}$$

and $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1 / 2^d = l + 1/2$ for some integer l.

Proof: By definition of a block-step vector, both $\mathbf{X}_{up}^{(w/2)}$ and $\mathbf{X}_{down}^{(w/2)}$ are step, but at least one of $\mathbf{X}_{up}^{(w/2)} \downarrow_2 (d-1)$ and $\mathbf{X}_{down}^{(w/2)} \downarrow_2 (d-1)$ is not step. Without loss of generality, assume $\mathbf{X}_{up}^{(w/2)} \downarrow_2 (d-1)$ is not step. By Proposition 3.11, it follows that there exists some index c_{up}, $0 < c_{up} \leq w/2 - 1$, such that either

- $(x_{up})_0 \downarrow_2 d = \ldots = (x_{up})_{c_{up}-1} \downarrow_2 d = 00\ldots0$ and $(x_{up})_{c_{up}} \downarrow_2 d = \ldots = (x_{up})_{w/2-1} \downarrow_2 d = 11\ldots1$, so that $\|\mathbf{X}_{up}^{(w/2)} \updownarrow_2 d\|_1 = (w/2 - c_{up})2^{d-1}$, or
- $(x_{up})_0 \downarrow_2 d = \ldots = (x_{up})_{c_{up}-1} \downarrow_2 d = 100\ldots0$ and $(x_{up})_{c_{up}} \downarrow_2 d = \ldots = (x_{up})_{w-1} \downarrow_2 d = 011\ldots1$, so that $\|\mathbf{X}_{up}^{(w/2)} \updownarrow_2 d\|_1 = c_{up}2^{d-1}$.

We proceed by case analysis on whether or not $\mathbf{X}_{down}^{(w/2)} \downarrow_2 (d-1)$ is step.

1. Assume first that $\mathbf{X}_{down}^{(w/2)} \downarrow_2 (d-1)$ is step.

   By Proposition 3.1, there exists some index c_{down}, $0 < c_{down} \leq w/2$, such that $(x_{down})_0 \downarrow_2 (d-1) = \ldots = (x_{down})_{c_{down}-1} \downarrow_2 (d-1)$ and $(x_{down})_{c_{down}} \downarrow_2 (d-1) = \ldots = (x_{down})_{w/2-1} \downarrow_2 (d-1) = (x_{down})_{c_{down}-1} \downarrow_2 (d-1) - 1$.

   Since $\mathbf{X}_{down}^{(w/2)}$ is also step, it follows by Corollary 3.10 that either $\|\mathbf{X}_{down}^{(w/2)} \updownarrow_2 d\|_1 = 0$, or $\|\mathbf{X}_{down}^{(w/2)} \updownarrow_2 d\|_1 = (w/2)2^{d-1}$. Hence, we have: Since $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1 = \|\mathbf{X}_{up}^{(w/2)} \updownarrow_2 d\|_1 + \|\mathbf{X}_{down}^{(w/2)} \updownarrow_2 d\|_1$, it follows that the possible values of $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d$ are $c_{up}/2$, $c_{up}/2 + w/4$, $w/4 - c_{up}/2$ and $w/2 - c_{up}/2$.

   Since $w \geq 4$ is a power of two, it follows that either $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l$ for some integer l in case c_{up} is even, or $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l + 1/2$ for some integer l in case c_{up} is odd.

   We proceed by case analysis on combinations of the parities of c_{up} and c_{down}.

   (a) Assume both c_{up} and c_{down} are even.

   Then, clearly, $\|(\mathbf{X}_{up})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)} \downarrow_2 (d-1)\|_1$ and $\|(\mathbf{X}_{down})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{down})_o^{(w/4)} \downarrow_2 (d-1)\|_1$, so that $\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (d-1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (d-1)\|_1 = 0$.

   Since c_{up} is even, $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l$ for some integer l.

   (b) Assume both c_{up} and c_{down} are odd.

   Then, clearly, $\|(\mathbf{X}_{up})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)} \downarrow_2 (d-1)\|_1 - (2^{\lg w - 1} - 1) = \|(\mathbf{X}_{up})_o^{(w/4)}\|_1 - w/2 + 1$ and $\|(\mathbf{X}_{down})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{down})_o^{(w/4)} \downarrow_2 (d-1)\|_1 + 1$, so that $\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (d-1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (d-1)\|_1 = -w/2$.

   Since c_{up} is odd, $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l + 1/2$ for some integer l.

   (c) Assume c_{up} is even and c_{down} is odd.

Then, clearly, $\|(\mathbf{X}_{up})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)} \downarrow_2 (d-1)\|_1$ and $\|(\mathbf{X}_{down})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{down})_o^{(w/4)} \downarrow_@ (d-1)\|_1 + 1$, so that $\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (d-1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (d-1)\|_1 = -1$.

Since c_{up} is even, $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l$ for some integer l.

(d) Assume c_{up} is odd and c_{down} is even.

Then, clearly, $\|(\mathbf{X}_{up})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)} \downarrow_2 (d-1)\|_1 - (2^{\lg w - 1} - 1) = \|(\mathbf{X}_{up})_o^{(w/4)} \downarrow_2 (d-1)\|_1 - w/2 + 1$ and $\|(\mathbf{X}_{down})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{down})_o^{(w/4)} \downarrow_2 (d-1)\|_1$, so that $\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (d-1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (d-1)\|_1 = -w/2 + 1$.

Since c_{up} is odd, $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l + 1/2$ for some integer l.

2. Assume now that $\mathbf{X}_{down}^{(w/2)} \downarrow_2 (d-1)$ is not step.

By Proposition 3.11, there exists some index c_{down}, $0 < c_{down} \leq w/2 - 1$, such that either

- $(x_{down})_0 \downarrow_2 d = \ldots = (x_{down})_{c_{down}-1} \downarrow_2 d = 00\ldots1$ and $(x_{down})_{c_{down}} \downarrow_2 d = \ldots = (x_{down})_{w/2-1} \downarrow_2 d = 11\ldots1$, so that $\|\mathbf{X}_{down}^{(w/2)} \updownarrow_2 d\|_1 = (w/2 - c_{down})2^{d-1}$, or
- $(x_{down})_0 \downarrow_2 d = \ldots = (x_{down})_{c_{down}-1} \downarrow_2 d = 100\ldots0$ and $(x_{down})_{c_{down}} \downarrow_2 d = \ldots = (x_{down})_{c_{w-1}} \downarrow_2 d = 011\ldots1$, so that $\|\mathbf{X}_{down}^{(w/2)} \updownarrow_2 d\|_1 = c_{down}2^{d-1}$.

Since $\mathbf{X}^{(w)} \updownarrow_2 d\|_1 = \|\mathbf{X}_{up}^{(w/2)} \updownarrow_2 d\|_1 + \|\mathbf{X}_{down}^{(w/2)} \updownarrow_2 d\|_1$, it follows that the possible values of $\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d$ are $(c_{up} + c_{down})/2$, $(w/2 + c_{up} - c_{down})/2$, $(w/2 + c_{down} - c_{up})/2$ and $(w - c_{up} - c_{down})/2$.

Since $w \geq 4$ is a power of two, it follows that either $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l$ for some integer l in case c_{up} and c_{down} are of the same parity, or $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l + 1/2$ in case c_{up} and c_{even} are of different parity.

We proceed by case analysis on combinations of the parities of c_{up} and c_{down}.

(a) Assume both c_{up} and c_{down} are even.

Then, clearly, $\|(\mathbf{X}_{up})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)} \downarrow_2 (d-1)\|_1$ and $\|(\mathbf{X}_{down})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{down})_o^{(w/4)} \downarrow_2 (d-1)\|_1$, so that $\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (d-1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (d-1)\|_1 = 0$.

Since c_{up} and c_{down} are of the same parity, $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l$ for some integer l.

(b) Assume both c_{up} and c_{down} are odd.

Then, clearly, $\|(\mathbf{X}_{up})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)} \downarrow_2 (d-1)\|_1 - (2^{\lg w - 1} - 1) = \|(\mathbf{X}_{up})_o^{(w/4)}\|_1 - w/2 + 1$ and $\|(\mathbf{X}_{down})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{down})_o^{(w/4)} \downarrow_2 (d-1)\|_1 - (2^{\lg w - 1} - 1) = \|(\mathbf{X}_{down})_o^{(w/4)}\|_1 - w/2 + 1$, so that $\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 2(d-1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (d-1)\|_1 = 0$.

Since c_{up} and c_{down} are of the same parity, $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l$ for some integer l.

(c) Assume c_{up} is even and c_{down} is odd.

Then, clearly, $\|(\mathbf{X}_{up})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)} \downarrow_2 (d-1)\|_1$ and $\|(\mathbf{X}_{down})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{down})_o^{(w/4)} \downarrow_2 (d-1)\|_1 - (2^{\lg w-1}-1) = \|(\mathbf{X}_{down})_o^{(w/4)} \downarrow_2 (d-1)\|_1 - w/2+1$, so that $\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (d-1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (d-1)\|_1 = w/2-1$.

Since c_{up} and c_{down} are of different parities, $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l + 1/2$ for some integer l.

(d) Assume c_{up} is odd and c_{down} is even.

Then, clearly, $\|(\mathbf{X}_{up})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{up})_o^{(w/4)} \downarrow_2 (d-1)\|_1 - (2^{\lg w-1}-1) = \|(\mathbf{X}_{up})_o^{(w/4)} \downarrow_2 (d-1)\|_1 - w/2+1$ and $\|(\mathbf{X}_{down})_e^{(w/4)} \downarrow_2 (d-1)\|_1 = \|(\mathbf{X}_{down})_o^{(w/4)} \downarrow_2 (d-1)\|_1$, so that $\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (d-1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (d-1)\|_1 = -w/2+1$.

Since c_{up} and c_{down} are of different parities, $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l + 1/2$ for some integer l.

Inspecting the previous case analysis reveals that either $|\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (d-1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (d-1)\|_1| \in \{0,1\}$ and $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l$ for some integer l (cases 1(a), 1(c), 2(a), 2(b)), or $|\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (d-1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (d-1)\|_1| \in \{w/2-1, w\}$ and $\|\mathbf{X}^{(w)} \updownarrow_2 d\|_1/2^d = l + 1/2$ for some integer l (cases 1(b), 1(d), 2(c), 2(d)), as needed. ■

5 Balancing Networks

This Section is organized as follows. Section 5.1 presents definitions for and preliminary properties of balancing networks. In Section 5.2, we define interesting classes of balancing networks, along with corresponding combinatorial characterization theorems. Our presentation closely follows the one in [8], where the reader is referred for a more detailed treatment.

5.1 Balancing Networks

Balancing networks are constructed from wires and computing elements called balancers. Formally, a *balancer* $b : \mathbf{X}^{(2)} \rightarrow \mathbf{Y}^{(2)}$ [5] is a computing element which receives integer inputs x_0 and x_1 on input wires 0 and 1, respectively, and computes integer outputs y_0 and y_1 on output wires 0 and 1, respectively, so that for each j, $0 \leq j \leq 1$,

$$y_j = \left\lceil \frac{\sum_{i=0}^{1} x_i - j}{2} \right\rceil .$$

For each j, $0 \leq j \leq 1$, the *order* of output wire j is defined to be $j/2$; thus, output wires 0 and 1 have orders 0 and $1/2$, respectively.

An immediate consequence of the definition of a balancer is that the output vector $\mathbf{Y}^{(2)}$ is step:

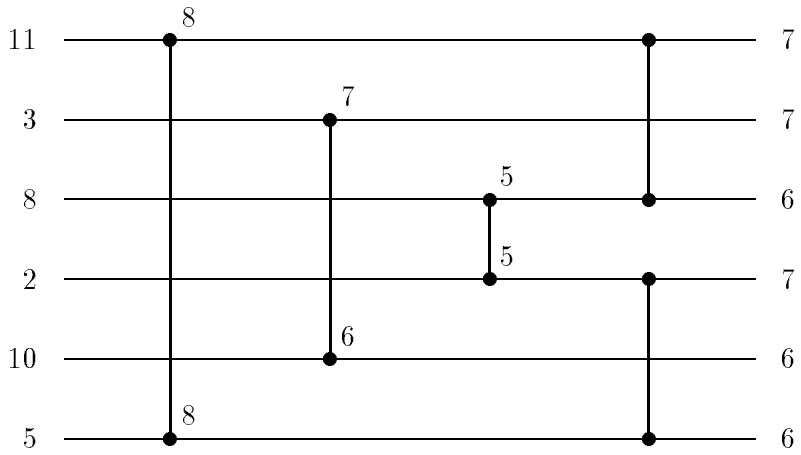

Figure 1: A balancing network

Proposition 5.1 *For a balancer* $b : \mathbf{X}^{(2)} \rightarrow \mathbf{Y}^{(2)}$*, the output vector* $\mathbf{Y}^{(2)}$ *is step.*

The *sum preservation property* for a balancer is another immediate consequence of its definition.

Proposition 5.2 *For a balancer* $b : \mathbf{X}^{(2)} \rightarrow \mathbf{Y}^{(2)}$, $\|\mathbf{Y}^{(2)}\|_1 = \|\mathbf{X}^{(2)}\|_1$.

For any integer $w \geq 2$, a *balancing network* $\mathcal{B} : \mathbf{X}^{(t)} \rightarrow \mathbf{Y}^{(w)}$ *of width* w is a collection of balancers, where output wires are connected to input wires, having w designated input wires $0, 1, \ldots, w-1$ (which are not connected to output wires of balancers), w designated output wires $0, 1, \ldots, w-1$ (similarly not connected to input wires of balancers), and containing no cycles. Integer inputs $x_0, x_1, \ldots, x_{w-1}$ are received on input wires $0, 1, \ldots, w-1$, respectively, and integer outputs $y_0, y_1, \ldots, y_{w-1}$ are computed on output wires $0, 1, \ldots, w-1$, respectively, in the natural way. Throughout the paper, we will often abuse notation and use x_i (resp., y_j) as the name of the ith input (resp., jth output) wire. Figure 1 depicts a balancing network, with wires drawn as horizontal lines and balancers stretched vertically, and the outputs computed on all output wires of its balancers on a specific input.

The sum preservation property for a balancing network $\mathcal{B}$ follows naturally from its definition and the corresponding property for a balancer:

Proposition 5.3 (Sum Preservation Property) *For a balancing network* $\mathcal{B} : \mathbf{X}^{(w)} \rightarrow \mathbf{Y}^{(w)}$, $\|\mathbf{Y}^{(w)}\|_1 = \|\mathbf{X}^{(w)}\|_1$.

For a balancing network $\mathcal{B}$, the *depth* of $\mathcal{B}$, $depth(\mathcal{B})$, is defined to be the maximal depth of any of its wires, where the depth of a wire is defined to be zero for an input wire of $\mathcal{B}$, and $\max_{i \in \{0,1\}} depth(x_i) + 1$, for the output wires of a balancer with input wires x_0, x_1.

In case $depth(\mathcal{B}) = 1$, $\mathcal{B}$ will be called a *layer* and will be uniquely represented by a square $w \times w$ matrix $\mathbf{C}_{\mathcal{B}}$, called *connection matrix* and determining the connections between input and output wires, and a $w \times 1$ column vector $\mathbf{O}_{\mathcal{B}}$, called *order vector* and determining the order of each output wire. Formally, we set:

- for any i and j, $0 \leq i, j \leq w-1$, $\mathbf{C}_{\mathcal{B}}[ji] = 1/2$ if input wire i and output wire j are connected via a balancer, else $\mathbf{C}_{\mathcal{B}}[ji] = 1$ if output wire j coincides with input wire i, and 0 otherwise.
- For any j, $0 \leq j \leq w-1$, $\mathbf{O}_{\mathcal{B}}[j] = o$ if output wire j is the output wire of a balancer and has order o, else (output wire j is not the output wire of a balancer) $\mathbf{O}_{\mathcal{B}}[j] = 0$.

For example, for the layer $\mathcal{B}$ depicted in Figure 2 using the same conventions as for Figure 1, we have:

$$\mathbf{C}_{\mathcal{B}} = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 0 & 1/2 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 1/2 \end{pmatrix},$$

and

$$\mathbf{O}_{\mathcal{B}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}.$$

By definitions for balancers, the connection matrix $\mathbf{C}_{\mathcal{B}}$, and the order vector $\mathbf{O}_{\mathcal{B}}$, it immediately follows:

Proposition 5.4 *For a layer* $\mathcal{B} : \mathbf{X}^{(w)} \rightarrow \mathbf{Y}^{(w)}$,

$$\mathbf{Y}^{(w)} = \lceil \mathbf{C}_{\mathcal{B}} \cdot \mathbf{X}^{(w)} - \mathbf{O}_{\mathcal{B}} \rceil$$

Notice that the connection matrix $\mathbf{C}_{\mathcal{B}}$ is a *doubly stochastic* matrix, i.e., all rows and columns sum to one (see, e.g., [3, Chapter VIII, Section 2] for an account on doubly stochastic matrices).

If $depth(\mathcal{B}) = d > 1$, then $\mathcal{B}$ can be uniquely partitioned into layers $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_d$ from left to right in the obvious way. The connection matrix $\mathbf{C}_{\mathcal{B}_i}$ and the order vector $\mathbf{O}_{\mathcal{B}_i}$ are associated with layer $\mathcal{B}_i$, $1 \leq i \leq d$. We represent $\mathcal{B}$ by the sequence of d connection matrices $\mathbf{C}_{\mathcal{B}_1}, \mathbf{C}_{\mathcal{B}_2}, \ldots, \mathbf{C}_{\mathcal{B}_d}$, and the sequence of d order vectors $\mathbf{O}_{\mathcal{B}_1}, \mathbf{O}_{\mathcal{B}_2}, \ldots, \mathbf{O}_{\mathcal{B}_d}$.

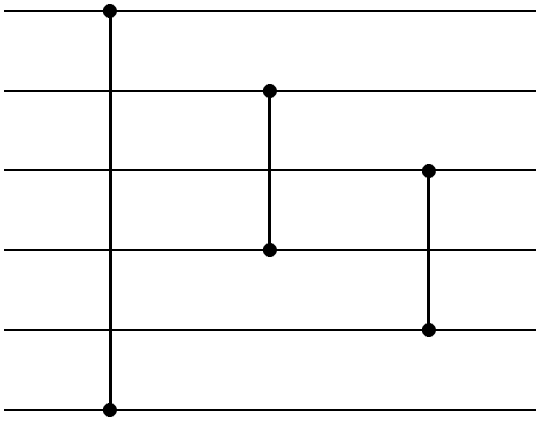

Figure 2: The layer $\mathcal{B}$

The next Theorem shows that for any balancing network, the outputs take a particular algebraic form as a function of the inputs, depending on the network's depth and the topology of the network.

Theorem 5.5 (Busch and Mavronicolas [8]) *Let $\mathcal{B} : \mathbf{X}^{(w)} \rightarrow \mathbf{Y}^{(w)}$ be a balancing network of depth d with associated connection matrices and order vectors $\mathbf{C}_{\mathcal{B}_1}, \mathbf{C}_{\mathcal{B}_2}, \ldots, \mathbf{C}_{\mathcal{B}_d}$ and $\mathbf{O}_{\mathcal{B}_1}, \mathbf{O}_{\mathcal{B}_2}, \ldots, \mathbf{O}_{\mathcal{B}_d}$, respectively. Then:*

$$\mathbf{Y}^{(w)} \;=\; \mathbf{C}_{\mathcal{B}} \cdot \mathbf{X}^{(w)} \uparrow_2 d + \mathbf{F}_{\mathcal{B}}(\mathbf{X}^{(w)} \downarrow_2 d),$$

for some matrix $\mathbf{C}_{\mathcal{B}}$ and vector function $\mathbf{F}_{\mathcal{B}} : [2^d]^w \rightarrow \mathbf{N}^w$, such that:

(1) $\mathbf{C}_{\mathcal{B}} \;=\; \mathbf{C}_{\mathcal{B}_d} \cdot \mathbf{C}_{\mathcal{B}_{d-1}} \cdot \ldots \cdot \mathbf{C}_{\mathcal{B}_1}$, *and:*

(2) $\mathbf{F}_{\mathcal{B}} = \mathbf{F}_{\mathcal{B}_d}$, *where the vector functions $\mathbf{F}_{\mathcal{B}_l} : [2^l]^w \rightarrow \mathbf{N}^w$, $1 \leq l \leq d$, are defined recursively as follows:*

$$\begin{aligned} & \mathbf{F}_{\mathcal{B}_l}(\mathbf{X}^{(w)} \downarrow_2 l) \\ = \; & \begin{cases} \left\lceil \mathbf{C}_{\mathcal{B}_l} \cdot \mathbf{C}_{\mathcal{B}_{l-1}} \cdot \ldots \cdot \mathbf{C}_{\mathcal{B}_1} \cdot \mathbf{X}^{(w)} \uparrow_2 l + \mathbf{C}_{\mathcal{B}_l} \cdot \mathbf{F}_{\mathcal{B}_{l-1}}(\mathbf{X}^{(w)} \downarrow_2 (l-1)) - \mathbf{O}_{\mathcal{B}_l} \right\rceil, & l > 1 \\ \left\lceil \mathbf{C}_{\mathcal{B}_1} \cdot \mathbf{X}^{(w)} \uparrow_2 1 - \mathbf{O}_{\mathcal{B}_1} \right\rceil, & l = 1 \end{cases} \end{aligned}$$

Call the matrix $\mathbf{C}_{\mathcal{B}}$ the *steady transfer matrix* of $\mathcal{B}$. Call the vector function $\mathbf{F}_{\mathcal{B}}$ the *transient transfer function* of $\mathcal{B}$.

Theorem 5.5 shows that the output vector of a balancing network is the sum of two terms. The first term $\mathbf{C}_{\mathcal{B}} \cdot \mathbf{X}^w \uparrow_2 d$, called the *steady output* term, involves the most significant part

$\mathbf{X}^{(w)} \uparrow_2 d$ of the input vector; this part is obtained by setting the d least significant binary digits of each entry of the input vector to zero. The steady output term is a linear transformation, defined by the steady transfer matrix $\mathbf{C}_{\mathcal{B}}$, of the most significant part of the input vector.

The second term $\left[\mathbf{C}_{\mathcal{B}} \cdot \mathbf{X}^{(w)} \updownarrow_2 d + \mathbf{C}_{\mathcal{B}_d} \cdot \mathbf{F}_{\mathcal{B}_{d-1}}(\mathbf{X}^{(w)} \downarrow_2 (d-1)) - \mathbf{O}_{\mathcal{B}_d}\right]$, called the *transient output* term, involves the least significant part $\mathbf{X}^{(w)} \downarrow_2 d$ of the input vector; this part corresponds to the d least significant binary digits of each entry. The transient output term is the image, under the transient transfer function $\mathbf{F}_{\mathcal{B}}$ of $\mathcal{B}$, of the least significant part of the input vector; apparently, the least significant part of the input vector undergoes a non-linear transformation defined by $\mathbf{F}_{\mathcal{B}}$.

Thus, the steady transfer matrix $\mathbf{C}_{\mathcal{B}}$ is determined by the relative connections of the network and shapes the steady output term, while the transient transfer function $\mathbf{F}_{\mathcal{B}}$ is determined by both the connections and the relative order of outputs for each balancer and shapes the transient output term. Call the steady transfer matrix $\mathbf{C}_{\mathcal{B}}$ and the transient transfer function $\mathbf{F}_{\mathcal{B}}$ the *transfer parameters* of $\mathcal{B}$.

Since the product of two doubly stochastic matrices is doubly stochastic (see, e.g., [3, Theorem 8.40]), Theorem 5.5(1) immediately implies:

Proposition 5.6 *The steady transfer matrix $\mathbf{C}_{\mathcal{B}}$ is doubly stochastic.*

Proposition 5.6 immediately implies:

Corollary 5.7 *For any vector $\mathbf{X}^{(w)}$, $\|\mathbf{C}_{\mathcal{B}} \cdot \mathbf{X}^{(w)}\|_1 = \|\mathbf{X}^{(w)}\|_1$.*

Our last Proposition establishes an intuitive property of the transient transfer function.

Proposition 5.8 *Let $\mathcal{B} : \mathbf{X}^{(w)} \rightarrow \mathbf{Y}^{(w)}$ be a balancing network of depth d. Then,*

$$\|\mathbf{F}_{\mathcal{B}}(\mathbf{X}^{(w)} \downarrow_2 d)\|_1 = \|\mathbf{X}^{(w)} \downarrow_2 d\|_1 \ .$$

Proof: By Theorem 5.5 and definition of 1-norm,

$$\begin{aligned} \|\mathbf{Y}^{(w)}\|_1 &= \|\mathbf{C}_{\mathcal{B}} \cdot \mathbf{X}^{(w)} \uparrow_2 d\|_1 + \|\mathbf{F}_{\mathcal{B}}(\mathbf{X}^{(w)} \downarrow_2 d)\|_1 \\ &= \|\mathbf{X}^{(w)} \uparrow_2 d\|_1 + \|\mathbf{F}_{\mathcal{B}}(\mathbf{X}^{(w)} \downarrow_2 d)\|_1 \ , \end{aligned}$$

since, by Proposition 5.6, $\mathbf{C}_{\mathcal{B}}$ is a doubly stochastic matrix.

By Proposition 5.3 and linearity of the 1-norm function,

$$\|\mathbf{Y}^{(w)}\|_1 = \|\mathbf{X}^{(w)}\|_1 = \|\mathbf{X}^{(w)} \uparrow_2 d\|_1 + \|\mathbf{X}^{(w)} \downarrow_2 d\|_1 \ .$$

Hence, it follows that

$$\|\mathbf{F}_{\mathcal{B}}(\mathbf{X}^{(w)} \downarrow_2 d)\|_1 = \|\mathbf{X}^{(w)} \downarrow_2 d\|_1 \ ,$$

as needed. ■

5.2 Counting and Merging Networks

Counting and merging networks [5] are among the most well studied classes of balancing networks (see, e.g., [1, 2, 7, 9, 10, 12, 13, 14, 15, 16]).

Definition 5.1 (Aspnes, Herlihy and Shavit [5]) *A* counting network *is a balancing network* $\mathcal{B} : \mathbf{X}^{(w)} \rightarrow \mathbf{Y}^{(w)}$ *such that the output vector* $\mathbf{Y}^{(w)}$ *is step.*

Counting networks have been shown suitable for implementing *shared counters* and *producer/consumer buffers* for multiprocessor architectures [5, 14].

A way of relaxing definition 5.1 is to require the step property for the output sequence only if the inputs have some kind of a step property.

Definition 5.2 (Aspnes, Herlihy and Shavit [5]) *A* merging network *is a balancing network* $\mathcal{B} : \mathbf{X}^{(w)} \rightarrow \mathbf{Y}^{(w)}$ *such that if the input vector* $\mathbf{X}^{(w)}$ *is block-step, then the output vector* $\mathbf{X}^{(w)}$ *is step.*

That is, the set of inputs is partitioned into two blocks, each of size $w/2$, and the output vector is step whenever each of the two corresponding vectors of inputs, one for each of these blocks, is. Merging networks have been used as building blocks of counting networks [5, 7, 10, 12].

A number of combinatorial characterizations may be derived from Theorem 5.5 for counting and merging networks. These characterizations are stated as necessary and sufficient conditions on the transfer parameters of a network. We start with a necessary and sufficient condition for a counting network.

Theorem 5.9 (Busch and Mavronicolas [8]) *The network* $\mathcal{B}$ *is a counting network if and only if:*

(1) $\mathbf{C}_{\mathcal{B}}[ji] = 1/w$ *for all* $i, j \in [w]$, *and*

(2) *the vector function* $\mathbf{F}_{\mathcal{B}}$ *is step on* $[2^d]^w$.

Theorem 5.5 can be used to show a *conditional* combinatorial characterization for merging networks.

Theorem 5.10 (Busch and Mavronicolas [8]) *For a network* $\mathcal{B} : \mathbf{X}^{(w)} \rightarrow \mathbf{Y}^{(w)}$ *of depth* d, *assume* $\mathbf{C}_{\mathcal{B}}[ji] = 1/w$ *for all* $i, j \in [w]$. *Then,* $\mathcal{B}$ *is a merging network if and only if the vector function* $\mathbf{F}_{\mathcal{B}}$ *is step on* $(blockstep(\mathbf{N}^w)) \downarrow_2 d$.

Theorem 5.10 reveals that certain merging networks may actually do more than what the formal definition of a merging network requires: a merging network with a constant steady transfer matrix produces a step output vector on an input vector which is *not* step, but has all of its entries no more than 2^{d-1} (i.e., each of its entries can be represented with d binary digits), where d is the depth of the network, and can be extended to a step vector by "sticking" most significamt binary digits to the left of each of its entries.

Theorems 5.9 and 5.10 suggest corresponding methodologies for proving correctness of counting and merging networks. More specifically, to show that a balancing network $\mathcal{B}$ is a counting or merging network, one computes expressions for the transfer parameters $\mathbf{C}_{\mathcal{B}}$ and $\mathbf{F}_{\mathcal{B}}$ and verifies inductively that the (conditional) necessary and sufficient conditions involved in Theorems 5.9 and 5.10, respectively, hold. A concrete application of this general methodology on a specific example of a merging network appears in Section 7.

6 The Bitonic Network

In this Section, we describe the construction and show some preliminary properties of the bitonic network [6].

Fix throughout w to be any power of two. The construction of the bitonic network $\mathcal{B}^{(w)}$ is inductive and identical to the one in [5]; it uses the bitonic merger network $\mathcal{M}^{(w)}$, whose construction is described next, as a basic module.

6.1 The Bitonic Merger Network

The balancing network $\mathcal{M}^{(w)} : \mathbf{X}^{(w)} \rightarrow \mathbf{Y}^{(w)}$, called *bitonic merger,* is defined inductively as follows: For the base case, where $w = 2$, $\mathcal{M}^{(2)}$ consists of a single balancer. Assume inductively that we have constructed $\mathcal{M}^{(w/2)}$, where $w \geq 4$; we show how to construct $\mathcal{M}^{(w)}$. The network $\mathcal{M}^{(w)}$ is the "cascade" of:

- a network $\mathcal{N}^{(w)} : \mathbf{X}^{(w)} \rightarrow \mathbf{Z}^{(w)}$, which is the "parallel composition" of two networks $\mathcal{M}_{eo}^{(w/2)} : \mathbf{X}_{eo}^{(w/2)} \rightarrow \mathbf{Z}_{e}^{(w/2)}$ and $\mathcal{M}_{oe}^{(w/2)} : \mathbf{X}_{oe}^{(w/2)} \rightarrow \mathbf{Z}_{o}^{(w/2)}$;
- a layer $\mathcal{L}^{(w)} : \mathbf{Z}^{(w)} \rightarrow \mathbf{Y}^{(w)}$ consisting of $w/2$ balancers $b_0, b_1, \ldots, b_{w/2-1}$, where balancer b_i receives inputs z_{2i} and z_{2i+1} and produces outputs y_{2i} and y_{2i+1}, $i \in [w/2]$.

Notice that the construction of $\mathcal{M}^{(w)}$ implies that $depth(\mathcal{M}_2) = 1$, while for $w > 2$,

$$depth(\mathcal{M}^{(w)}) \;=\; depth(\mathcal{N}^{(w)}) \;+ depth(\mathcal{L}^{(w)}) \;=\; depth(\mathcal{M}^{(w/2)}) + 1 \;,$$

implying:

Proposition 6.1 *For all $w \geq 2$, $depth(\mathcal{M}^{(w)}) = \lg w$.*

Notice that, by definition of the network $\mathcal{L}^{(w)}$, $\mathbf{C}_{\mathcal{L}^{(w)}}[ji] = 1/2$ if $\{i,j\} \subseteq \{2l, 2l+1\}$ for some $l \in [w/2]$ and 0 otherwise, i.e.,

$$\mathbf{C}_{\mathcal{L}^{(w)}} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & \dots & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & \dots & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & \dots & 1/2 & 1/2 \end{pmatrix}.$$

Also, by definition of the network $\mathcal{L}^{(w)}$, $\mathbf{O}_{\mathcal{L}^{(w)}}[i] = 0$ for $i \in e[w]$, and $1/2$ for $i \in o[w]$, i.e.,

$$\mathbf{O}_{\mathcal{L}^{(w)}} = \mathbf{O}^{(w)} = \begin{pmatrix} 0 \\ 1/2 \\ 0 \\ 1/2 \\ \vdots \\ 0 \\ 1/2 \end{pmatrix}.$$

We continue by showing that the steady transfer matrix of the bitonic merger is a constant matrix.

Proposition 6.2 *For all $i, j \in [w]$, $\mathbf{C}_{\mathcal{M}^{(w)}}[ji] = 1/w$.*

Proof: By induction on w. For the base case, where $w = 2$, the statement holds trivially by definitions of balancer and connection matrix. Assume inductively that for each $k \leq w/2$ that is a power of two, $\mathbf{C}_{\mathcal{M}^{(k)}}[ji] = 1/k$ for all $i, j \in [k]$. We show that $\mathbf{C}_{\mathcal{M}^{(w)}}[ji] = 1/w$ for all $i, j \in [w]$.

By definition of the network $\mathcal{N}^{(w)}$ and induction hypothesis, $\mathbf{C}_{\mathcal{N}^{(w)}}[ji] = 1/(w/2) = 2/w$ if either $j \in e[w]$ and $i \in eo[w]$, or $j \in o[w]$ and $i \in oe[w]$, and 0 otherwise, so that:

$$\mathbf{C}_{\mathcal{N}^{(w)}} = \begin{pmatrix} 2/w & 0 & 2/w & 0 & \dots & 2/w & 0 & 0 & 2/w & 0 & 2/w & \dots & 0 & 2/w \\ 0 & 2/w & 0 & 2/w & \dots & 0 & 2/w & 2/w & 0 & 2/w & 0 & \dots & 2/w & 0 \\ 2/w & 0 & 2/w & 0 & \dots & 2/w & 0 & 0 & 2/w & 0 & 2/w & \dots & 0 & 2/w \\ 0 & 2/w & 0 & 2/w & \dots & 0 & 2/w & 2/w & 0 & 2/w & 0 & \dots & 2/w & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 2/w & 0 & 2/w & 0 & \dots & 2/w & 0 & 0 & 2/w & 0 & 2/w & \dots & 0 & 2/w \\ 0 & 2/w & 0 & 2/w & \dots & 0 & 2/w & 2/w & 0 & 2/w & 0 & \dots & 2/w & 0 \end{pmatrix}.$$

Hence, Theorem 5.5(1) implies:

$$\begin{aligned}
\mathbf{C}_{\mathcal{M}^{(w)}} &= \mathbf{C}_{\mathcal{L}^{(w)}} \cdot \mathbf{C}_{\mathcal{N}^{(w)}} \\
&= \begin{pmatrix}
1/w & 1/w & 1/w & 1/w & \ldots & 1/w & 1/w \\
1/w & 1/w & 1/w & 1/w & \ldots & 1/w & 1/w \\
1/w & 1/w & 1/w & 1/w & \ldots & 1/w & 1/w \\
1/w & 1/w & 1/w & 1/w & \ldots & 1/w & 1/w \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
1/w & 1/w & 1/w & 1/w & \ldots & 1/w & 1/w \\
1/w & 1/w & 1/w & 1/w & \ldots & 1/w & 1/w
\end{pmatrix},
\end{aligned}$$

as needed. ■

By appealing to Proposition 6.2, Theorem 5.10 calls for computing expressions for $\mathbf{F}_{\mathcal{M}^{(w)}}$ on $\mathbf{X}^{(w)} \in (\mathit{blockstep}(\mathbf{N}^w)) \downarrow_2 \mathit{depth}(\mathcal{M}^{(w)})$ which, by Proposition 6.1, equals $(\mathit{blockstep}(\mathbf{N}^w)) \downarrow_2 \lg w$. We first introduce some notation. Define

$$\begin{aligned}
&\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)},\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) \\
= \;& \frac{1}{2}(\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1)) + \mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1)) .)
\end{aligned}$$

Our next proposition shows that the transient transfer function of $\mathcal{M}^{(w)}$ takes a particular algebraic form for $\mathbf{X}^{(w)} \in (\mathit{blockstep}(\mathbf{N}^w)) \downarrow_2 \lg w$.

Proposition 6.3 *Assume* $\mathbf{X}^{(w)} \in (\mathit{blockstep}(\mathbf{N}^w)) \downarrow_2 \lg w$. *Then,*

$$\begin{aligned}
&\mathbf{F}_{\mathcal{M}^{(w)}}(\mathbf{X}^{(w)}) \\
= \;& \left\lceil \frac{1}{w} \| \mathbf{X}^{(w)} \updownarrow_2 \lg w \|_1 \mathbf{1}^{(w)} + \mathbf{2F}_{\mathcal{M}_{eo}^{(w/2)},\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) - \mathbf{O}_{\mathcal{L}^{(w)}} \right\rceil .
\end{aligned}$$

Proof: Since $\mathbf{X}^{(w)} \in (\mathit{blockstep}(\mathbf{N}^w)) \downarrow_2 \lg w \subseteq \mathbf{N}^w \downarrow_2 \lg w$, $\mathbf{X}^{(w)} \uparrow_2 (\lg w - 1) = \mathbf{X}^{(w)} \updownarrow_2 \lg w$ and $\mathbf{X}^{(w)} \uparrow_2 \lg w = \mathbf{O}^{(w)}$.

By construction of the network $\mathcal{N}^{(w)}$ and Theorem 5.5,

$$\begin{aligned}
\mathbf{Z}^{(w)} &= \mathbf{C}_{\mathcal{N}^{(w)}} \cdot \mathbf{X}^{(w)} \uparrow_2 (\lg w - 1) + \mathbf{F}_{\mathcal{N}^{(w)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) \\
&= \mathbf{C}_{\mathcal{N}^{(w)}} \cdot \mathbf{X}^{(w)} \updownarrow_2 \lg w + \mathbf{F}_{\mathcal{N}^{(w)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) ,
\end{aligned}$$

where

$$(\mathbf{F}_{\mathcal{N}^{(w)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)))_e = \mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1)) ,$$

and

$$(\mathbf{F}_{\mathcal{N}^{(w)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)))_o = \mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1)) .$$

Applying Proposition 5.4 on layer $\mathcal{L}^{(w)}$ yields:

$$\begin{aligned}
\mathbf{Y}^{(w)} &= \lceil \mathbf{C}_{\mathcal{L}^{(w)}} \cdot \mathbf{Z}^{(w)} - \mathbf{O}_{\mathcal{L}^{(w)}} \rceil \\
&= \lceil \mathbf{C}_{\mathcal{L}^{(w)}} \cdot (\mathbf{C}_{\mathcal{N}^{(w)}} \cdot \mathbf{X}^{(w)} \updownarrow_2 \lg w + \mathbf{F}_{\mathcal{N}^{(w)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1))) - \mathbf{O}_{\mathcal{L}^{(w)}} \rceil \\
&= \lceil \mathbf{C}_{\mathcal{L}^{(w)}} \cdot \mathbf{C}_{\mathcal{N}^{(w)}} \cdot \mathbf{X}^{(w)} \updownarrow_2 \lg w + \mathbf{C}_{\mathcal{L}^{(w)}} \cdot \mathbf{F}_{\mathcal{N}^{(w)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) - \mathbf{O}_{\mathcal{L}^{(w)}} \rceil \\
&= \lceil \mathbf{C}_{\mathcal{M}^{(w)}} \cdot \mathbf{X}^{(w)} \updownarrow_2 \lg w + \mathbf{C}_{\mathcal{L}^{(w)}} \cdot \mathbf{F}_{\mathcal{N}^{(w)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) - \mathbf{O}_{\mathcal{L}^{(w)}} \rceil \ ,
\end{aligned}$$

by Theorem 5.5(1). Proposition 6.2 implies that

$$\mathbf{C}_{\mathcal{M}^{(w)}} \cdot \mathbf{X}^{(w)} \updownarrow_2 \lg w \ = \ \frac{1}{w} \| \mathbf{X}^{(w)} \updownarrow_2 \lg w \|_1 \mathbf{1}^{(w)} \ .$$

Notice also that

$$\begin{aligned}
& \mathbf{C}_{\mathcal{L}^{(w)}} \cdot \mathbf{F}_{\mathcal{N}^{(w)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) \\
= & \begin{pmatrix}
1/2 & 1/2 & 0 & 0 & \dots & 0 & 0 \\
1/2 & 1/2 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 1/2 & 1/2 & \dots & 0 & 0 \\
0 & 0 & 1/2 & 1/2 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \dots & 1/2 & 1/2 \\
0 & 0 & 0 & 0 & \dots & 1/2 & 1/2
\end{pmatrix} \cdot \mathbf{F}_{\mathcal{N}^{(w)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) \\
= & \frac{1}{2} \mathbf{2} (\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1)) + \mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))) \ .
\end{aligned}$$

Hence, we have:

$$\begin{aligned}
& \mathbf{F}_{\mathcal{M}^{(w)}}(\mathbf{X}^{(w)}) \\
= & \left\lceil \frac{1}{w} \| \mathbf{X}^{(w)} \updownarrow_2 \lg w \|_1 \mathbf{1}^{(w)} + \mathbf{2} \mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}, \mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) - \mathbf{O}_{\mathcal{L}^{(w)}} \right\rceil \ .
\end{aligned}$$

Since $\mathbf{X}^{(w)} \in (\mathit{blockstep}(\mathbf{N}^w)) \downarrow_2 \lg w \subseteq \mathbf{N}^w \downarrow_2 \lg w = [2^{\lg w}]^w$, it follows, by Theorem 5.5, that

$$\begin{aligned}
& \mathbf{F}_{\mathcal{M}^{(w)}}(\mathbf{X}^{(w)}) \\
= & \left\lceil \frac{1}{w} \| \mathbf{X}^{(w)} \updownarrow_2 \lg w \|_1 \mathbf{1}^{(w)} + \mathbf{2} \mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}, \mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) - \mathbf{O}_{\mathcal{L}^{(w)}} \right\rceil \ ,
\end{aligned}$$

as needed. ■

As we will show, $\mathcal{M}^{(w)}$ is a merging network, i.e., it guarantees the step property on its output vector when each of the vectors $\mathbf{X}_{up}^{(w/2)}$ and $\mathbf{X}_{down}^{(w/2)}$ is step – but this can be ensured by filtering each of these vectors through smaller counting networks.

6.2 The Bitonic Network

The balancing network $\mathcal{B}^{(w)} : \mathbf{X}^{(w)} \rightarrow \mathbf{Y}^{(w)}$, called *bitonic,* is defined inductively. For the base case, where $w = 2$, $\mathcal{B}^{(2)}$ consists of a single balancer. Assume inductively that we have constructed $\mathcal{B}^{(w/2)}$, where $w \geq 4$; we show how to construct $\mathcal{B}^{(w)}$. The network $\mathcal{B}^{(w)}$ is the cascade of:

- a network $\mathcal{R}^{(w)} : \mathbf{X}^{(w)} \rightarrow \mathbf{Z}^{(w)}$, which is the "parallel composition" of two networks $\mathcal{B}_{up}^{(w/2)} : \mathbf{X}_{up}^{(w/2)} \rightarrow \mathbf{Z}_{up}^{(w/2)}$ and $\mathcal{B}_{down}^{(w/2)} : \mathbf{X}_{down}^{(w/2)} \rightarrow \mathbf{Z}_{down}^{(w/2)}$;
- a bitonic merger network $\mathcal{M}^{(w)} : \mathbf{Z}^{(w)} \rightarrow \mathbf{Y}^{(w)}$.

7 A Correctness Proof

In this Section, we present a formal correctness proof that the bitonic network is a counting network.

We start by showing:

Theorem 7.1 *The network $\mathcal{M}^{(w)}$ is a merging network.*

Proof: The proof appeals to Theorem 5.10. By Propositions 6.2 and 6.1, it suffices to show:

Lemma 7.2 *The function $\mathbf{F}_{\mathcal{M}^{(w)}}$ is step on $(blockstep(\mathbf{N}^w)) \downarrow_2 \lg w$.*

Proof: By induction on w. For the base case, where $w = 2$ and $\mathcal{M}^{(2)}$ consists of a single balancer, notice that $(blockstep(\mathbf{N}^2)) \downarrow_2 \lg 2 = \mathbf{N}^2 \downarrow_2 1$ and take any $\mathbf{X}^{(2)} \in \mathbf{N}^2 \downarrow_2 1$. By definition of a balancer, $\mathbf{Y}^{(2)}$ is step, while, by Theorem 5.5, $\mathbf{Y}^{(2)} = \mathbf{F}_{\mathcal{B}}(\mathbf{X}^{(2)})$. It follows that $\mathbf{F}_{\mathcal{B}}(\mathbf{X}^{(2)})$ is step, as needed.

Assume inductively that $\mathbf{F}_{\mathcal{M}^{(k)}}$ is step on $(blockstep(\mathbf{N}^k)) \downarrow_2 \lg k$ for each $k \leq w/2$ that is a power of two. Take any $\mathbf{X}^{(w)} \in (blockstep(\mathbf{N}^w)) \downarrow_2 \lg w$. We show that $\mathbf{F}_{\mathcal{M}^{(w)}}(\mathbf{X}^{(w)})$ is step.

By definition of $(blockstep(\mathbf{N}^w)) \downarrow_2 \lg w$, there exists a block-step vector $\mathbf{V}^{(w)}$ such that $\mathbf{V}^{(w)} \downarrow_2 \lg w = \mathbf{X}^{(w)}$. It follows by Proposition 4.1 that each of $\mathbf{V}_{eo}^{(w/2)}$ and $\mathbf{V}_{oe}^{(w/2)}$ is block-step. Notice that $\mathbf{V}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1) = \mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1)$ and $\mathbf{V}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1) = \mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1)$. It follows, by definition of $(blockstep(\mathbf{N}^{w/2})) \downarrow_2 (\lg w - 1)$, that both $\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1) \in (blockstep(\mathbf{N}^{w/2})) \downarrow_2 (\lg w - 1)$ and $\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1) \in (blockstep(\mathbf{N}^{w/2})) \downarrow_2 (\lg w - 1)$. Hence, by induction hypothesis, each of $\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1))$ and $\mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))$ is step.

We proceed by case analysis on whether or not $\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)$ is block-step.

1. **Case 1: $\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)$ is block-step.**

   By Proposition 4.3, $\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 / w$ is an integer. Hence, by Proposition 6.3:

$$\begin{aligned} & \mathbf{F}_{\mathcal{M}^{(w)}}(\mathbf{X}^{(w)}) \\ = \; & \frac{1}{w}\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 \mathbf{1}^{(w)} + \left[\mathbf{2F}_{\mathcal{M}_{eo}^{(w/2)},\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) - \mathbf{O}_{\mathcal{L}^{(w)}}\right] . \end{aligned}$$

   By Proposition 4.2,

$$|\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1)\|_1| \in \{0, 1\} .$$

   By Corollary 5.8,

$$\|\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1 = \|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1)\|_1$$

   and

$$\|\mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1 = \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1)\|_1 .$$

   Hence, it follows that

$$|\|\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1 - \|\mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1| \in \{0, 1\} .$$

   Since each of $\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1))$ and $\mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))$ is step, it follows by Proposition 3.7 that the vector

$$\begin{aligned} & \left[\frac{1}{2}\mathbf{2}(\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1)) + \mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))) - \mathbf{O}^{(w)}\right] \\ = \; & \left[\mathbf{2F}_{\mathcal{M}_{eo}^{(w/2)},\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}^{(w/2)} \downarrow_2 (\lg w - 1)) - \mathbf{O}_{\mathcal{L}^{(w)}}\right] \end{aligned}$$

   is step. Since the sum of a constant vector and a step vector is a step vector, it follows that the vector $\mathbf{F}_{\mathcal{M}^{(w)}}(\mathbf{X}^{(w)})$ is step, as needed.

2. **Case 2: $\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)$ is not block-step.**

   Setting $d = \lg w$ in Proposition 4.4, so that $2^d = w$, yields that either

$$|\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1)\|_1| \in \{0, 1\}$$

   and $\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 / w = l$ for some integer l, or

$$|\|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1)\|_1 - \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1)\|_1| \in \{w/2 - 1, w/2\}$$

   and $\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 / w = l + 1/2$ for some integer l.

By Corollary 5.8,

$$\|\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1 \;=\; \|\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1)\|_1$$

and

$$\|\mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1 = \|\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1)\|_1 \; .$$

Hence, it follows that either

$$|\|\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1 - \|\mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1| \;\in\; \{0, 1\}$$

and $\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 / w = l$ for some integer l, or

$$|\|\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1 - \|\mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1| \;\in\; \{w/2 - 1, w/2\}$$

and $\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 / w = l + 1/2$ for some integer l.
We proceed by case analysis.

(a) Assume first that

$$|\|\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1 - \|\mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1| \;\in\; \{0, 1\}$$

and $\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 / w = l$ for some integer l.
By Proposition 6.3, it follows that

$$\begin{aligned} & \mathbf{F}_{\mathcal{M}^{(w)}}(\mathbf{X}^{(w)} \downarrow_2 \lg w) \\ = \;& \frac{1}{w}\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 \mathbf{1}^{(w)} + \left\lceil \mathbf{2F}_{\mathcal{M}_{eo}^{(w/2)}, \mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) - \mathbf{O}_{\mathcal{L}^{(w)}} \right\rceil \; . \end{aligned}$$

Since each of $\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1))$ and $\mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))$ is step, it follows by Proposition 3.7 that the vector

$$\begin{aligned} & \left\lceil \frac{1}{2}\mathbf{2}(\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1)) + \mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))) - \mathbf{O}^{(w)} \right\rceil \\ = \;& \left\lceil \mathbf{2F}_{\mathcal{M}_{eo}^{(w/2)}, \mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) - \mathbf{O}_{\mathcal{L}^{(w)}} \right\rceil \end{aligned}$$

is step. Since the sum of a constant vector and a step vector is a step vector, it follows that the vector $\mathbf{F}_{\mathcal{M}^{(w)}}(\mathbf{X}^{(w)})$ is step, as needed.

(b) Assume now that

$$|\|\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1 - \|\mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))\|_1| \;\in\; \{w/2 - 1, w/2\}$$

and $\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1/w = l + 1/2$ for some integer l. By Proposition 6.3:

$$\begin{aligned} & \mathbf{F}_{\mathcal{M}^{(w)}}(\mathbf{X}^{(w)}) \\ = \; & \frac{1}{w}\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 \mathbf{1}^{(w)} + \left\lceil \frac{1}{2}\mathbf{1}^{(w)} + \mathbf{2F}_{\mathcal{M}_{eo}^{(w/2)},\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) - \mathbf{O}_{\mathcal{L}^{(w)}} \right\rceil \; . \end{aligned}$$

By definition of the vectors $\mathbf{E}^{(w)}$ and $\mathbf{O}^{(w)}$,

$$\frac{1}{2}\mathbf{1}^{(w)} - \mathbf{O}_{\mathcal{L}^{(w)}} \; = \; \frac{1}{2}\mathbf{1}^{(w)} - \mathbf{O}^{(w)} \; = \; \mathbf{E}^{(w)} \; .$$

It follows that

$$\begin{aligned} & \mathbf{F}_{\mathcal{M}^{(w)}}(\mathbf{X}^{(w)}) \\ = \; & \frac{1}{w}\|\mathbf{X}^{(w)} \updownarrow_2 \lg w\|_1 \mathbf{1}^{(w)} + \left\lceil \mathbf{2F}_{\mathcal{M}_{eo}^{(w/2)},\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) + \mathbf{E}^{(w)} \right\rceil \; . \end{aligned}$$

Since each of $\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1))$ and $\mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))$ is step, it follows by Proposition 3.8 that the vector

$$\begin{aligned} & \left\lceil \frac{1}{2}\mathbf{2}(\mathbf{F}_{\mathcal{M}_{eo}^{(w/2)}}(\mathbf{X}_{eo}^{(w/2)} \downarrow_2 (\lg w - 1)) + \mathbf{F}_{\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}_{oe}^{(w/2)} \downarrow_2 (\lg w - 1))) + \mathbf{E}^{(w)} \right\rceil \\ = \; & \left\lceil \mathbf{2F}_{\mathcal{M}_{eo}^{(w/2)},\mathcal{M}_{oe}^{(w/2)}}(\mathbf{X}^{(w)} \downarrow_2 (\lg w - 1)) + \mathbf{E}^{(w)} \right\rceil \end{aligned}$$

is step. Since the sum of a constant vector and a step vector is a step vector, it follows that the vector $\mathbf{F}_{\mathcal{M}^{(w)}}(\mathbf{X}^{(w)})$ is step, as needed.

■

By Lemmas 6.2 and 7.2, it follows from Theorem 5.10 that the network $\mathcal{M}^{(w)}$ is a merging network, as needed. ■

We finally argue:

Theorem 7.3 *The network $\mathcal{B}^{(w)} : \mathbf{X}^{(w)} \rightarrow \mathbf{Y}^{(w)}$ is a counting network.*

Proof: By induction on w. For the base case, where $w = 2$, $\mathcal{B}^{(2)}$ consists of a single balancer which, by Proposition 5.1, produces a step output vector.

Assume inductively that $\mathcal{B}^{(k)}$ is a counting network for each $k \leq w/2$ that is a power of two. We show that $\mathcal{B}^{(w)}$ is a counting network. By construction of $\mathcal{B}^{(w)}$ (Section 6.2) and induction hypothesis, each of $\mathbf{Z}_{up}^{(w/2)}$ and $\mathbf{Z}_{down}^{(w/2)}$ is a step vector. Since, by Theorem 7.1, the network $\mathcal{M}^{(w)}$ is a merging network, it follows that the output vector $\mathbf{Y}^{(w)}$ is step, as needed. ■

8 Concluding Remarks

We presented a new proof that the bitonic network is a counting network. Our proof consists of a routine verification of the necessary and sufficient conditions involved in combinatorial characterization theorems for counting and merging networks shown in [8]. This proof is a concrete instance of a paradigmatic methodology, suggested in this work, for showing correctness of balancing networks. For other instances where this methodology has already been applied and yielded corresponding proofs of comparable modularity and simplicity, we refer the reader to [7, 10].

It would yet be interesting to further investigate the generality of our proof technique by applying it to other constructions of counting networks. Good candidates would be the periodic counting network with fan-out 2^k [5] and the periodic smoothing network with fan-out p^k [13]. Finally, recent randomized constructions of counting and smoothing networks [2], using *randomized balancers,* call for corresponding methodologies for proving probabilistic properties of randomized constructions.

Acknowledgements:

We have had helpful discussions with Maurice Herlihy and Nancy Lynch on our initial ideas on investigating techniques for proving correctness of constructions of balancing networks, based on the theory developed in [8], in the context of a concrete example. Special thanks go to Mauricio Resende and an anonymous referee for many helpful editorial and wording comments.

References

[1] E. Aharonson and H. Attiya, "Counting Networks with Arbitrary Fan-Out," *Proceedings of the 3rd Annual ACM–SIAM Symposium on Discrete Algorithms,* pp. 104–113, January 1992.

[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, August 1994.

[3] M. Aigner, *Combinatorial Theory,* Springer-Verlag, 1979.

[4] T. E. Anderson, "The Performance of Spin Lock Alternatives for Shared-Memory Multiprocessors," *IEEE Transactions on Parallel and Distributed Systems,* Vol. 1, No. 1, pp. 6–16, January 1990.

[5] J. Aspnes, M. Herlihy and N. Shavit, "Counting Networks and Multi-Processor Coordination," *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing,* pp. 348–358, May 1991.

Expanded versions: "Counting Networks," Technical Memo MIT/LCS/TM-451, Laboratory of Computer Science, MIT, June 1991, and: "Counting Networks," Technical Report CRL 93/11, Digital Equipment Corporation, Cambridge Research Laboratory, August 1993.

[6] K. E. Batcher, "Sorting Networks and their Applications," *Proceedings of AFIPS Spring Joint Computer Conference,* pp. 307–314, 1968.

[7] C. Busch, N. Hardavellas and M. Mavronicolas, "Contention in Counting Networks," *Proceedings of the 13th Annual ACM Symposium on Principles of Distributed Computing,* pp. 404, August 1994.

[8] C. Busch and M. Mavronicolas, "A Combinatorial Treatment of Balancing Networks," *Proceedings of the 13th Annual ACM Symposium on Principles of Distributed Computing,* pp. 206–215, August 1994.

[9] C. Busch and M. Mavronicolas, "A Depth-Contention Optimal Counting Network," submitted for publication.

[10] C. Busch and M. Mavronicolas, "Odd-Even Counting Networks," in preparation.

[11] 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, October 1989.

[12] E. W. Felten, A. LaMarca and R. Ladner, "Building Counting Networks from Larger Balancers," Technical Report 93-04-09, Department of Computer Science and Engineering, University of Washington, April 1993.

[13] N. Hardavellas, D. Karakos and M. Mavronicolas, "Notes on Sorting and Counting Networks," *Proceedings of the 7th International Workshop on Distributed Algorithms (WDAG-93),* Lecture Notes in Computer Science, Vol. # 725 (A. Schiper, ed.), Springer-Verlag, pp. 234–248, September 1993.

[14] M. Herlihy, B.-C. Lim and N. Shavit, "Low Contention Load Balancing on Large-Scale Multiprocessors," *Proceedings of the 4th Annual ACM Symposium on Parallel Algorithms and Architectures,* pp. 219–227, July 1992.

[15] M. Herlihy, N. Shavit and O. Waarts, "Low Contention Linearizable Counting Networks," *Proceedings of the 32nd Annual IEEE Symposium on Foundations of Computer Science,* pp. 526–535, October 1991.

[16] M. Klugerman and C. Plaxton, "Small-Depth Counting Networks," *Proceedings of the 24th Annual ACM Symposium on Theory of Computing,* pp. 417–428, May 1992.

[17] D. Knuth, *Sorting and Searching*, Volume 3 of *The Art of Computer Programming*, Addison-Wesley, 1973.

[18] D. Loeb and G.-C. Rota, "Formal Power Series of Logarithmic Type," *Advances in Mathematics,* Vol. 75, No. 1, pp. 1–118, May 1989.