Near-Perfect Load Balancing by Randomized Rounding

Tobias Friedrich
International Computer Science Institute
1947 Center St., Suite 600
94704 Berkeley, CA, USA

Thomas Sauerwald
International Computer Science Institute
1947 Center St., Suite 600
94704 Berkeley, CA, USA

ABSTRACT

We consider and analyze a new algorithm for balancing indivisible loads on a distributed network with n processors. The aim is minimizing the discrepancy between the maximum and minimum load. In every time-step paired processors balance their load as evenly as possible. The direction of the excess token is chosen according to a randomized rounding of the participating loads.

We prove that in comparison to the corresponding model of Rabani, Sinclair, and Wanka (1998) with arbitrary roundings, the randomization yields an improvement of roughly a square root of the achieved discrepancy in the same number of time-steps on all graphs. For the important case of expanders we can even achieve a constant discrepancy in $\mathcal{O}(\log n\,(\log\log n)^3)$ rounds. This is optimal up to $\log\log n$-factors while the best previous algorithms in this setting either require $\Omega(\log^2 n)$ time or can only achieve a logarithmic discrepancy. This result also demonstrates that with randomized rounding the difference between discrete and continuous load balancing vanishes almost completely.

Categories and Subject Descriptors: F.2 [Theory of Computation]: Analysis of Algorithms and Problem Complexity

General Terms: Algorithms, Theory

1. INTRODUCTION

Consider an application running on a parallel or distributed network consisting of n processors connected in an arbitrary topology. Each processor has initially a collection of jobs (which we call *tokens*). The goal of load-balancing is to reallocate the tokens by transmitting them along edges so that each processor has nearly the same amount of load. The problem has manifold applications in job scheduling, routing, adaptive mesh partitioning, finite element computations, and in simulations of physical phenomena.

STOC'09, May 31–June 2, 2009, Bethesda, Maryland, USA.
Copyright 2009 ACM 978-1-60558-506-2/09/05 ...$5.00.

There are various models for load balancing. A common simplifying assumption is that the tokens are divisible. This idealized process is very well understood [18]. However, the divisibility assumption is invalid for many applications [21]. It has been shown that the deviation can be quite significant [11, 21] and the question of a precise quantitative relationship between the discrete and the idealized process has been posed by several authors [10, 11, 15, 18, 21].

Existing models also differ in the assumptions regarding communication in the underlying network. Some models restrict the number of tokens send across a link at a time [1, 10, 16]. On the other hand, the diffusion model allows load to be moved from each processor to all its neighbors in parallel in each time step [8, 18, 21]. As pointed out by Ghosh and Muthukrishnan [11], it is more efficient to send a stream of many tokens to one neighbor than to send one token to each neighbor. This motivates us to study the *balancing circuit model* [3] where each vertex transfers an arbitrary number of tokens to exactly *one* neighbor in each round [5, 11, 19].

The load-balancing process of a balancing circuit is governed by a sequence of (not necessarily perfect) matchings together with an orientation of each edge[1]. In each round, two paired vertices balance their loads as evenly as possible. If this is not possible, the excess token is send in the direction of the edge. For an arbitrary graph it is not clear how to choose a good sequence of matchings that balance the load quickly. This is different for highly structured graphs. There is for example a canonical choice of matchings for the hypercube that uses in round i all edges across dimension $(i \bmod \log n)$ [2]. In general such matching sequences with a fixed period d are called periodic balancing circuits (note that typically d is of order the maximum degree of G).

Let $x^{(0)} \in \mathbb{R}^n$ be the initial load vector of the network. The aim of load-balancing algorithms is to reduce the initial discrepancy $K = \max_{ij} |x_i^{(0)} - x_j^{(0)}|$ within a certain number of steps. For this, Rabani, Sinclair and Wanka [19] introduced the so-called local divergence that provides an upper bound on the deviation between the idealized and the discrete model over all time steps. They [19, Corollary 5] showed in the periodic balancing circuits model that within $\mathcal{O}(d\log(Kn)/(1-\lambda_2))$ steps the discrepancy can be reduced to $\mathcal{O}(d\log n/(1-\lambda_2))$, where $(1-\lambda_2)$ is the eigenvalue gap of the balancing matrix.

[1] Note that the same edge may have different orientation over time. The precise model is described in Section 2.

Graph class	Rounds	Discrepancy	Orientation	Reference
General Graph (Balancing Circuit)	t	$\Omega(\max\{\mathrm{diam}(G), \log n/(\log t)\})$	arb.	Proposition 2.3
	$\mathcal{O}(\log(Kn)/(1-\lambda_2))$	$\mathcal{O}(d\log n/(1-\lambda_2))$	arb.	[19, Corollary 5]
	$\mathcal{O}(\log(Kn)/(1-\lambda_2))$	$\mathcal{O}(\sqrt{d\log n/(1-\lambda_2)})$	rand.	Corollary 4.5
	$\mathcal{O}(\log(Kn)/(1-\lambda_2))$	$\mathcal{O}(d\log\log n/(1-\lambda_2))$	rand.	Corollary 4.7
Constant-Degree Expander (Balancing Circuit)	$\mathcal{O}(\log(Kn))$	$\mathcal{O}(\log n)$	arb.	[19, Corollary 5]
	$\mathcal{O}(\log(Kn))$	$\mathcal{O}(\log\log n)$	rand.	Proposition 4.8
	$\mathcal{O}(\log(Kn)(\log\log n)^3)$	$\mathcal{O}(1)$	rand.	Corollary 5.22
d-dim. Torus (Balancing Circuit)	$\mathcal{O}(\log(Kn)\, n^{2/d})$	$\mathcal{O}(d\, n^{1/d})$	arb.	[19, Theorem 8]
	$\mathcal{O}(\log(Kn)\, n^{2/d})$	$\mathcal{O}(\sqrt{d\, n^{1/d}}\,\log n)$	rand.	Corollary 4.9
	$\mathcal{O}(n^{2/d}\log(Kn) + d\, n^{1+1/d})$	1	std.	[19, Corollary 10]
General Graph (Random matching)	$\mathcal{O}(\log(Kn)/(1-\lambda_2))$	$\mathcal{O}(\log(n)/(1-\lambda_2))$	arb.	[19][2]
	$\mathcal{O}(\log(Kn)/(1-\lambda_2))$	$\mathcal{O}(\sqrt{\log(n)/(1-\lambda_2)}\,\log n)$	rand.	Corollary 5.4
Expander (Random matching)	$\mathcal{O}(\log(Kn))$	$\mathcal{O}(\log(n))$	arb.	[19][2]
	$\mathcal{O}(\log(Kn)(\log\log n)^3)$	$\mathcal{O}(1)$	rand.	Theorem 5.21

Table 1. Summary and comparison of our new upper and lower bounds on the discrepancy for different graphs. Arb., stand., and rand. refer to an arbitrary, standard and randomized orientation of all edges, respectively. K is the initial discrepancy of the load vector. In the balancing circuit model, λ_2 is the second largest eigenvalue of the round matrix $\mathbf{P} = \prod_{k=1}^{d} \mathbf{P}^{(k)}$. In case of random matchings, λ_2 is the second largest eigenvalue of the standard diffusion matrix $\mathbf{Q}$ (cf. Section 2.2) in absolute value.

Our results

When balancing an odd number of indivisible tokens, it is crucial to decide in which direction to send the excess token. The results of Muthukrishnan, Ghosh and Schultz [18] and most results of Rabani *et al.* [19] hold regardless of the orientation of the matching edges. We observe in Section 2.3 that in this case on every graph in $\log n$ time-steps the discrepancy cannot be reduced below $\mathcal{O}(\log n/\log\log n)$ and that n time-steps are necessary for constant discrepancy. This is rather unsatisfying as the idealized process requires only a logarithmic number of rounds to reach a constant discrepancy on many important graphs like expanders, complete graphs and hypercubes.

In order to reduce the deviation between the idealized and the discrete process we follow a suggestion of Rabani *et al.* [19] and distribute the excess token in a more balanced manner, that is, in a random direction. Intuitively it is clear that this discrete randomized model should be closer to the idealized process. However, it is surprising that this small difference in the model results in a such a vast reduction of the discrepancy. In more detail, our results that are summarized in Table 1 are as follows.

For general graphs we reduce the discrepancy to $\mathcal{O}\left(\min\left\{\sqrt{\frac{d\,\log n}{1-\lambda_2}}, \frac{d\log\log n}{1-\lambda_2}\right\}\right)$ in $\mathcal{O}\left(\frac{d\cdot\log(Kn)}{1-\lambda_2}\right)$ steps w.h.p. (cf. Corollary 4.5 and 4.7). The analogous result of [19, Corollary 5] only achieves a discrepancy of $\mathcal{O}\left(\frac{d\,\log n}{1-\lambda_2}\right)$ in the same number of rounds. While this gives a quadratic improvement for many graph, the improvement for constant-degree expanders from $\mathcal{O}(\log n)$ to $\mathcal{O}(\log\log n)$ is even exponential. Interestingly, our proof also reveals that the deviation between the discrete and idealized model is $\mathcal{O}(\log\log n)$ for constant-degree expanders, which we prove to be tight on any constant-degree expander (cf. Theorem 4.10). For the d-dimensional torus graph we achieve a discrepancy of $\mathcal{O}(\sqrt{d\,n^{1/d}}\,\log n)$ (cf. Corollary 4.9) in a number of rounds where [19, Theorem 8] only achieves $\mathcal{O}(d\,n^{1/d})$ w.h.p.

As it might be hard to define canonical matchings for non-structured graphs, it is popular to use random matchings instead, see e.g. [4, 5, 10, 11]. We prove results that hold for a large class of randomly generated matchings including the models of [5, 10, 11]. For arbitrary graphs we prove a bound of $\mathcal{O}\big(\sqrt{\frac{\log n}{1-\lambda_2}}\,\log n\big)$ after $\mathcal{O}\big(\frac{\log(Kn)}{1-\lambda_2}\big)$ rounds w.h.p. (cf. Corollary 5.4). If $1/(1-\lambda_2)$ is the dominant term, we get again a quadratic improvement over the model with arbitrary directions of the matching edges. For the important case of an expander where $1/(1-\lambda_2)$ is constant, we do a separate analysis and show that the discrepancy gets down to $\mathcal{O}(1)$ in nearly optimal time $\mathcal{O}(\log(Kn)(\log\log n)^3)$ w.h.p. (cf. Theorem 5.21). This result can also be extended to constant-degree expanders in the balancing circuit model with appropriate deterministic matchings (cf. Corollary 5.22).

Overall, our results demonstrate that with an appropriate randomization the gap between the discrete and idealized model decreases significantly. For the important class of expander graphs the difference disappears almost completely. On these graphs it is interesting to look at the time-discrepancy trade-off, which can be measured as the product of time and the achieved discrepancy. All previous trade-offs had a time-discrepancy product of $\Omega(\log^2 n + \log(Kn))$ [1, 8, 10, 18, 19, 6] while we achieve a product $\mathcal{O}(\log(Kn)(\log\log n)^3)$ which is very close to the natural lower bound of $\Omega(\log(Kn))$.

Related work

Balancing circuits were introduced by Aspnes, Herlihy and Shavit [3]. They constructed a sequence of $\Theta(\log^2 n)$ matchings that achieve a discrepancy of one for all inputs for a specific orientation of the edges. This result was improved by Klugerman and Plaxton [13, 14] who constructed for the same problem a sequence of only $\Theta(\log n)$ matchings. Note

[2] Rabani *et al.* [19] did not consider random matchings, but a straightforward adaptation of their techniques yields this bound.

that in contrast to the model we have described before, the orientation of all edges must be fixed and there is no restriction on the set of matching edges (this can be viewed as a balancing circuit on complete graphs).

Recently, a special balancing circuit called block network was examined under the assumption that edges are oriented uniformly at random [12]. In [17] the authors showed that the cascade of two block networks gives a discrepancy of 17 in $2\log n$ steps. However, the analysis of [17] is rather tailored for this special network, while our results apply to all graphs.

Another very surprising relationship between a discrete process and its idealized (continuous) counterpart appears for so-called deterministic random walks. Cooper and Spencer [7] show a remarkable similarity between the expectation of a random walk (the idealized process) and a deterministic analogue where instead of distributing tokens randomly, each vertex serves its neighbors in a fixed order. If an (almost) arbitrary distribution of token is placed on the vertices of an infinite grid $\mathbb{Z}^d$ and does a simultaneous walk in the deterministic random walk model, then at all times and on each vertex, the number of tokens deviates from the expected number the standard random walk would have gotten there, by at most a constant.

Rounding the number of tokens $(x_i^{(t)} + x_j^{(t)})/2$ shared at time t by two paired vertices i and j at random can be seen as dependent randomized rounding of half-integral numbers. This very general approach of randomized rounding [20] up or down with probability depending on the fractional part is a standard method for approximating the solution of a discrete problem by rounding the solution of an idealized (continuous) problem.

Organization of the paper

In Section 2 we give a more formal description of our load-balancing model. Section 3 presents the basic method and proves some general results. In Section 4 we introduce the local p-divergence Ψ_p to bound the difference between the discrete and idealized process in the periodic setting. In Section 5 we study random matchings and give in Section 5.1 results for arbitrary graphs. In Section 5.2 we show that we can also achieve a discrepancy of $\mathcal{O}(1)$ on any expander at the cost of an $(\log\log n)^3$ factor in the runtime.

2. PRELIMINARIES AND DEFINITIONS

2.1 Our Model and Notations

Let $G = (V, E)$ be a graph with vertices $V = [n]$, edges $E \subseteq V^2$, min-degree $\delta = \min_{v \in V} \deg(v)$, and max-degree $\Delta = \max_{v \in V} \deg(v)$. For any vertex $u \in V$ and integer $k \in \mathbb{N}$, we define $B_k(u) := \{v \in V \colon \operatorname{dist}(v,u) \leqslant k\}$. All logarithms are to the base 2. By w.h.p. (with high probability) we refer to an event that holds with probability at least $1 - n^{-c}$ for some constant $c > 1$.

The iterative load-balancing process is governed by a sequence of (not necessarily perfect) matchings $M^{(1)}, M^{(2)}, \ldots$ Every matching corresponds to a doubly stochastic communication matrix $\mathbf{P}^{(t)}$ with $\mathbf{P}_{ij}^{(t)} = 1/2$ if i and j are matched in $M^{(t)}$, $\mathbf{P}_{ii}^{(t)} = 1$ if i is not matched in $M^{(t)}$, and $\mathbf{P}_{ij}^{(t)} = 0$ otherwise. The special case of periodic balancing circuits was introduced by Aspnes *et al.* [3]. In the language of the model described above they refer to a sequence of communication matrices $\mathbf{P}^{(t)}$ where two matrices $\mathbf{P}^{(t_1)}$ and $\mathbf{P}^{(t_2)}$ are identical if $t_1 \equiv t_2 \bmod d$.

We start with an arbitrary load vector $\xi^{(0)} \in \mathbb{R}_+^n$. In round t, if two vertices i, j are matched, they balance their loads as closely as possible. For divisible tokens the process is just a Markov chain and can be described by

$$\xi^{(t)} = \xi^{(t-1)}\, \mathbf{P}^{(t)}. \tag{1}$$

In the discrete process with indivisible tokens, we have to decide where to send the excess token if the sum of the tokens of two matched vertices is odd. Most of the results [19] hold for an arbitrary orientation of all edges. However, for their results about perfect load balancing, they require the so-called standard orientation where each edge $\{i, j\}$ with $i < j$ is oriented towards i, that is an excess token is sent to i. In spite of the huge bias, this orientation is particularly useful for the reduction to sorting networks [3, 14, 19]. As a drawback, the standard orientation requires global consistency. Therefore it is intuitively appealing to consider a random orientation of all edges $\{i, j\}$. That way, the load is distributed more evenly (in expectation). Moreover, a random orientation can be computed locally and offers fault-tolerance against crashes or replacement of the communication links [12].

The precise model is as follows. Whenever the vertices i and j perform a balancing operation in round t, the vertex i with $i < j$ flips an unbiased coin at step t and according to the outcome it gets either $\lfloor (x_i^{(t)} + x_j^{(t)})/2 \rfloor$ or $\lceil (x_i^{(t)} + x_j^{(t)})/2 \rceil$ tokens. The remaining number of tokens is sent to j. Note that this corresponds to a randomized rounding [20] of the idealized process which sends $(x_i^{(t)} + x_j^{(t)})/2$ to both vertices.

To describe the direction of the excess token in time step t, we use $\Phi_{i,j}^{(t)}$ to specify at edge $\{i, j\} \in M^{(t)}$ where it is send. $\Phi_{i,j}^{(t)}$ is $+1$ if the excess token is send to i and -1 if it is send to j. In this setting, the standard orientation of Rabani *et al.* [19] corresponds to $\Phi_{i,j}^{(t)} = 1$ for $i < j$ and $\Phi_{i,j}^{(t)} = -1$ otherwise. The randomized rounding implies that for each $\{i, j\} \in M^{(t)}$, $\Phi_{i,j}^{(t)}$ is chosen uniformly and independently at random from $\{-1, 1\}$. We will refer to a matching edge $\{i, j\} \in M^{(t)}$ with $i < j$ shortly as $[i : j]$.

2.2 Preliminaries

Following the approach by [19] of relating the discrete and idealized process, we set $x^{(0)} = \xi^{(0)}$, and therefore

$$\xi^{(t)} = \xi^{(0)} \prod_{k=1}^{t} \mathbf{P}^{(k)}. \tag{2}$$

Note that each matching matrix $\mathbf{P} = \mathbf{P}^{(k)}$ is a symmetric, doubly stochastic matrix that satisfies $\mathbf{P}^2 = \mathbf{P}$. Hence all eigenvalues of $\mathbf{P}^{(k)}$ are real and non-negative and so are the ones of $\prod_{k=1}^{t} \mathbf{P}^{(k)}$.

Note that the idealized process is well-understood, in particular if all $\mathbf{P}^{(k)}$'s are the same. We define average load to be $\bar{x} = \sum_{i=1}^{n} x_i^{(0)}/n$. We restate the following two well-known results.

Lemma 2.1 (e.g. [18, Lemma 2]). *For any time step $t \geqslant 0$, $\|\xi^{(t)} - \bar{x}\|_2^2 \leqslant \|\xi^{(0)} - \bar{x}\|_2^2 \cdot (\lambda_2(\mathbf{P}))^{2t}$.*

Theorem 2.2 (e.g. [19, Theorem 1]). *In the idealized process, the number of rounds for achieving a discrepancy of ℓ for an initial load vector $\xi^{(0)}$ with discrepancy K is bounded above by $\mathcal{O}\left(\frac{1}{1-\lambda_2(\mathbf{P})} \cdot \log\left(\frac{Kn}{\ell}\right)\right)$.*

Finally, we define $\mathbf{Q} := \mathbf{I} - \frac{1}{\Delta+1}\mathbf{L}$ with $\mathbf{L}$ being the Laplacian matrix of G. Note that $\mathbf{Q}$ is a diffusion-matrix and corresponds to a natural random walk with loops that moves to each neighbor with the same probability. We call a graph with $\Delta/\delta = \mathcal{O}(1)$ an expander graph, if $1/(1 - \lambda_2(\mathbf{Q}))$ is bounded above by a constant. Unless otherwise stated, we do not require an expander to be of constant-degree.

2.3 Lower Bounds on Arbitrary Rounding

By relatively simple arguments we show that without randomly chosen directions for the excess tokens, the discrepancy can still be large.

Proposition 2.3. *Let G be any graph and let $M^{(1)}, M^{(2)}, \ldots, M^{(T)}$ a sequence of T matchings. Then there is an orientation of each matching edge and an initial load vector $x^{(0)}$ such that the discrepancy of $x^{(T)}$ is at least $\max\{\operatorname{diam}(G), \frac{\log n}{\log T}\}$.*

Hence, for $T = \mathcal{O}(\log n)$ the best discrepancy we can get with an arbitrary orientation is $\Omega(\frac{\log n}{\log\log n})$. Note that this result matches the result of [19, Corollary 5] for constant-degree expanders up to a factor of $\log\log n$ (assuming that the initial discrepancy K is polynomially bounded). To get down to a constant discrepancy for an arbitrary input and orientation of the matching edges, even $\Omega(n)$ rounds are necessary. In sharp contrast, our general result achieves for any expander a constant discrepancy in only $\mathcal{O}(\log n\,(\log\log n)^3)$ rounds.

With the same construction as in Proposition 2.3 we obtain the following result.

Corollary 2.4. *There is a graph G such that for any sequence of matchings $M^{(1)}, M^{(2)}, \ldots, M^{(T)}$ with the standard orientation, there is an initial load vector $x^{(0)}$ such that the discrepancy of $x^{(T)}$ is at least $\max\{\operatorname{diam}(G), \frac{\log n}{\log T}\}$.*

3. THE BASIC METHOD

We have seen that if the $\Phi_{i,j}^{(t)}$ are arbitrary (or all 1), we cannot hope for a discrepancy that is significantly less than logarithmic. That is why we incorporate the idea of randomized rounding that can be described as follows.

We have

$$x^{(t)} = x^{(t-1)}\,\mathbf{P}^{(t)} + e^{(t)} \tag{3}$$

for $t \geqslant 1$. $e^{(t)}$ is the excess load allocated as a result of rounding up and down. More precisely, for all $t \geqslant 1$ and $i \in [n]$, $e^{(t)}$ is given by

$$e_i^{(t)} = \begin{cases} \frac{1}{2}\mathsf{Odd}(x_i^{(t-1)} + x_j^{(t-1)})\,\Phi_{i,j}^{(t)} & \text{if } \{i,j\} \in M^{(t)} \\ 0 & \text{otherwise} \end{cases},$$

where we define $\mathsf{Odd}(i)$ to be one if i is odd and zero otherwise. With $M_{\mathsf{Odd}}^{(t)} := \left\{\{i,j\} \in M^{(t)} \mid \mathsf{Odd}(x_i^{(t-1)} + x_j^{(t-1)}) = 1\right\}$, this is the same as

$$\begin{aligned} e^{(t)} &= \tfrac{1}{2}\textstyle\sum_{[i:j]\in M^{(t)}} \mathsf{Odd}(x_i^{(t-1)} + x_j^{(t-1)})\,\Phi_{i,j}^{(t)}\left(u_i - u_j\right) \\ &= \tfrac{1}{2}\textstyle\sum_{[i:j]\in M_{\mathsf{Odd}}^{(t)}} \Phi_{i,j}^{(t)}\left(u_i - u_j\right) \end{aligned}$$

with u_i denoting the i-th n-dimensional row unit vector.

Unwinding equation (3) yields

$$x^{(t)} = x^{(0)} \prod_{k=1}^{t} \mathbf{P}^{(k)} + \sum_{\ell=1}^{t} e^{(\ell)} \prod_{k=\ell+1}^{t} \mathbf{P}^{(k)}. \tag{4}$$

By equations (4) and (2),

$$\begin{aligned} x^{(t)} - \xi^{(t)} &= \textstyle\sum_{\ell=1}^{t} e^{(\ell)} \prod_{k=\ell+1}^{t} \mathbf{P}^{(k)} \\ &= \tfrac{1}{2}\textstyle\sum_{\ell=1}^{t} \sum_{[i:j]\in M_{\mathsf{Odd}}^{(\ell)}} \Phi_{i,j}^{(\ell)}\left(u_i - u_j\right) \prod_{k=\ell+1}^{t} \mathbf{P}^{(k)} \end{aligned}$$

$$\left(x^{(t)} - \xi^{(t)}\right)_v = \tfrac{1}{2}\textstyle\sum_{\ell=1}^{t} \sum_{[i:j]\in M_{\mathsf{Odd}}^{(\ell)}} \Phi_{i,j}^{(\ell)} w_{i,j}^{[\ell,t]}(v) \tag{5}$$

with

$$w_{i,j}^{[\ell,t]}(v) := \left(\left(\textstyle\prod_{k=\ell+1}^{t} \mathbf{P}^{(k)}\right)_{i,v} - \left(\textstyle\prod_{k=\ell+1}^{t} \mathbf{P}^{(k)}\right)_{j,v}\right).$$

To bound the summand (or more generally, a subset of the summand) of equation (5) we need the following lemma. This lemma allows us to assume that all matching edges receive an odd number of tokens and therefore we may deal with a sum of independent random variables. A less general, but similar lemma has been shown in [12].

Lemma 3.1. *For any triple of time steps $1 \leqslant t_1 \leqslant t_2 \leqslant t \leqslant T$ let $W_{\mathsf{Odd}} := \sum_{\ell=t_1}^{t_2} \sum_{[i:j]\in M_{\mathsf{Odd}}^{(\ell)}} \Phi_{i,j}^{(\ell)}\, w_{i,j}^{[\ell,t]}$, and $W := \sum_{\ell=t_1}^{t_2} \sum_{[i:j]\in M^{(\ell)}} \Phi_{i,j}^{(\ell)}\, w_{i,j}^{[\ell,t]}$. Then for any $\delta \geqslant 0$, $\mathbf{Pr}\left[|W_{\mathsf{Odd}}| \geqslant \delta\right] \leqslant 2\,\mathbf{Pr}\left[|W| \geqslant \delta\right]$.*

We will also frequently use the following basic lemma.

Lemma 3.2. *For any pairs of steps ℓ, t with $1 \leqslant \ell \leqslant t \leqslant T$ and any vertex v, we have $\sum_{[i:j]\in M^{(\ell)}} w_{i,j}^{[\ell,t]}(v) \leqslant 2\,\|(\prod_{k=\ell+1}^{t} \mathbf{P}^{(k)})u_v\|_2^2$.*

We state two results for general sequences of matching matrices that are used later on.

Theorem 3.3. *For any graph G and matchings $M_1, M_2, \ldots, M_T$ fix two time steps t_1, t with $1 \leqslant t_1 \leqslant t \leqslant T$ and any vertex v. Then,*

$$\mathbf{Pr}\left[\left|\tfrac{1}{2}\textstyle\sum_{\ell=1}^{t_1} \sum_{[i:j]\in M_{\mathsf{Odd}}^{(\ell)}} \Phi_{i,j}^{(\ell)} w_{i,j}^{[\ell,t]}(v)\right| \geqslant \sqrt{8t_1 \left\|\left(\textstyle\prod_{k=t_1}^{t} \mathbf{P}^{(k)}\right)u_v\right\|_2^2 \log n}\,\right] \leqslant n^{-4}.$$

Theorem 3.3 roughly states that the rounding errors that have small impact (small norm) contribute only little to $x^{(t)} - \xi^{(t)}$. For the special case $t_1 = t$, the theorem above directly gives us the following general result.

Corollary 3.4. *For any graph G, $v \in V$ and step t, the deviation between the discrete and idealized process after t rounds is at most $\sqrt{8\log n\, t}$ w.h.p.*

4. PERIODIC BALANCING CIRCUITS

4.1 Upper Bounds

We now consider the periodic balancing circuit model. In this case, two matching matrices $\mathbf{P}^{(t_1)}$ and $\mathbf{P}^{(t_2)}$ are identical if $t_1 \equiv t_2 \bmod d$. Hence, it makes sense to define the round matrix $\mathbf{P} = \prod_{k=1}^{d} \mathbf{P}^{(k)}$.

For bounding the deviation between the discrete and idealized process, we state the following definition that generalizes the one of [19] to arbitrary p-norms.

Definition 4.1. *The local p-divergence is defined by*

$$\Psi_p = \max_{v \in V} \left(\sum_{t=1}^{\infty} \sum_{[i:j] \in M^{(\ell)}} \left| \Big(\prod_{k=\ell+1}^{t} \mathbf{P}^{(k)} \Big)_{i,v} - \Big(\prod_{k=\ell+1}^{t} \mathbf{P}^{(k)} \Big)_{j,v} \right|^p \right)^{1/p}.$$

Note that $\Psi_2(\mathbf{P}) \leqslant \Psi_1(\mathbf{P})$, but also $\Psi_2(\mathbf{P})^2 \leqslant \Psi_1(\mathbf{P})$, since $\prod_{k=\ell+1}^{t} \mathbf{P}^{(k)}$ is a stochastic matrix. Rabani *et al.* [19] expressed the deviation between the discrete and idealized process in terms of $\Psi_1(\mathbf{P})$ and showed that reducing the discrepancy to $\mathcal{O}(\Psi_1(\mathbf{P}))$ for any initial vector with discrepancy K can be achieved within $\mathcal{O}\left(d\,\frac{\log(Kn)}{1-\lambda_2(\mathbf{P})}\right)$ steps[3].

Theorem 4.2. *For any time step $t \geqslant 0$, the maximum deviation between the discrete and idealized process at step t is at most $\mathcal{O}(\Psi_2(\mathbf{P})\sqrt{\log n})$ w.h.p. If the initial discrepancy K is polynomial in n, the maximum deviation over all time steps is at most $\mathcal{O}(\Psi_2(\mathbf{P})\sqrt{\log n})$ w.h.p.*

Using this and Theorem 2.2, we immediately get

Corollary 4.3. *We achieve a discrepancy of $\mathcal{O}\left(\Psi_2(\mathbf{P})\sqrt{\log n}\right)$ for any vector with discrepancy K within $\mathcal{O}\left(\frac{\log(Kn)}{1-\lambda_2(\mathbf{P})}\right)$ steps w.h.p.*

Let us now bound $\Psi_2(\mathbf{P})$ in terms of the second largest eigenvalue.

Theorem 4.4. *For any round matrix $\mathbf{P}$, $\Psi_2(\mathbf{P}) = \mathcal{O}\left(\sqrt{\frac{d}{1-\lambda_2(\mathbf{P})}}\right)$.*

Compared to the result $\Psi_1(\mathbf{P}) = \mathcal{O}\left(\frac{d \log n}{1-\lambda_2(\mathbf{P})}\right)$ of [19, Theorem 4], our bound on $\Psi_2(\mathbf{P})$ is much smaller, because the ℓ_2-convergence of $\mathbf{P}^{(t)}$ is faster than the ℓ_1-convergence. A direct application gives the following result.

Corollary 4.5. *We achieve a discrepancy of $\mathcal{O}\left(\sqrt{\frac{d}{1-\lambda_2(\mathbf{P})}\log n}\right)$ after $\mathcal{O}\left(\frac{d \cdot \log(Kn)}{1-\lambda_2(\mathbf{P})}\right)$ steps w.h.p.*

By a more subtle analysis, we get a much better result for small values of $1-\lambda_2(\mathbf{P})$ and d. Again we first consider the deviation between the discrete and continuous process.

Theorem 4.6. *For any time step t, the maximum deviation between the discrete and idealized model at step t is at most $\mathcal{O}\left(\frac{d \log\log n}{1-\lambda_2(\mathbf{P})}\right)$ w.h.p. If the initial discrepancy K is polynomial in n, the maximum deviation over all time steps is at most $\mathcal{O}\left(\frac{d \log\log n}{1-\lambda_2(\mathbf{P})}\right)$ w.h.p.*

Proof. We proceed similarly as in Theorem 4.2, but here we split the sum and bound the critical summand directly in terms of $\lambda_2 = \lambda_2(\mathbf{P})$. Let $t \geqslant 1$ be an arbitrary, but fixed time step.

$$\begin{aligned} x_v^{(t)} - \xi_v^{(t)} =& \frac{1}{2} \sum\nolimits_{\ell=1}^{t} \sum\nolimits_{[i:j] \in M_{\mathsf{Odd}}^{(\ell)}} \Phi_{i,j}^{(\ell)} w_{i,j}^{[\ell,t]}(v) \\ =& \frac{1}{2} \sum\nolimits_{\ell=1}^{t-\frac{8d\log\log n}{1-\lambda_2}} \sum\nolimits_{[i:j] \in M_{\mathsf{Odd}}^{(\ell)}} \Phi_{i,j}^{(\ell)} w_{i,j}^{[\ell,t]}(v) \\ &+ \frac{1}{2} \sum\nolimits_{\ell=t-\frac{8d\log\log n}{1-\lambda_2}+1}^{t} \sum\nolimits_{[i:j] \in M_{\mathsf{Odd}}^{(\ell)}} \Phi_{i,j}^{(\ell)} w_{i,j}^{[\ell,t]}(v). \end{aligned}$$

The second term can be bounded trivially by $\frac{8d\log\log n}{1-\lambda_2}$, since $|\sum_{[i:j] \in M_{\mathsf{Odd}}^{(\ell)}} w_{i,j}^{[\ell,t]}(v)| \leqslant 1$ for every ℓ. It remains to bound the first term. Note that for any $1 \leqslant \ell \leqslant t$,

$$\begin{aligned} &\sum\nolimits_{[i:j] \in M_{\mathsf{Odd}}^{(\ell)}} \mathbf{Range}\left[\Phi_{i,j}^{(\ell)} w_{i,j}^{[\ell,t]}(v)\right]^2 \\ &\leqslant 2 \sum\nolimits_{[i:j] \in M_{\mathsf{Odd}}^{(\ell)}} \left(w_{i,j}^{[\ell,t]}(v)\right)^2 \\ &= 2 \sum\nolimits_{[i:j] \in M_{\mathsf{Odd}}^{(\ell)}} \left(\big(\prod\nolimits_{k=\ell+1}^{t} \mathbf{P}^{(k)}\big)_{i,v} - \big(\prod\nolimits_{k=\ell+1}^{t} \mathbf{P}^{(k)}\big)_{j,v} \right)^2 \\ &\leqslant 2\, \|(\prod\nolimits_{k=\ell+1}^{t} \mathbf{P}^{(k)}) u_v\|_2^2 \leqslant 2\lambda_2^{\lfloor \frac{t-\ell}{d} \rfloor} \end{aligned}$$

by Lemmas 3.2 and 2.1 and therefore

$$\begin{aligned} &\sum\nolimits_{\ell=1}^{t-\frac{8d\log\log n}{1-\lambda_2}} \sum\nolimits_{[i:j] \in M_{\mathsf{Odd}}^{(\ell)}} \mathbf{Range}\left[\Phi_{i,j}^{(\ell)} w_{i,j}^{[\ell,t]}(v)\right]^2 \\ &\leqslant 2 \sum\nolimits_{\ell=1}^{t-\frac{8d\log\log n}{1-\lambda_2}} \lambda_2^{\lfloor \frac{t-\ell}{d} \rfloor} \\ &= \mathcal{O}\left(d\, \frac{\lambda_2^{\frac{8\log\log n}{1-\lambda_2}-1}}{1-\lambda_2} \right) = \mathcal{O}\left(d\, \frac{\frac{1}{\log^4 n}}{1-\lambda_2} \right). \end{aligned}$$

Applying Lemma 3.1 and Hoeffding's bound we obtain,

$$\begin{aligned} &\mathbf{Pr}\left[\left| \sum\nolimits_{\ell=1}^{t-\frac{8\log\log n}{1-\lambda_2}} \sum\nolimits_{[i:j] \in M_{\mathsf{Odd}}^{(\ell)}} \Phi_{i,j}^{(\ell)} w_{i,j}^{[\ell,t]}(v) \right| \geqslant \delta \right] \\ &\leqslant 2\exp\left(\frac{-\delta^2}{\sum\nolimits_{\ell=1}^{t-\frac{8\log\log n}{1-\lambda_2}} \sum\nolimits_{[i:j] \in M_{\mathsf{Odd}}^{(\ell)}} \mathbf{Range}\left[\Phi_{i,j}^{(\ell)} w_{i,j}^{[\ell,t]}(v)\right]^2} \right) \\ &\leqslant 2\exp\left(-\delta^2 \Big/ \mathcal{O}\left(d\, \frac{\frac{1}{\log^4 n}}{1-\lambda_2} \right) \right). \end{aligned}$$

Choosing $\delta = \Theta\left(\sqrt{\frac{d}{1-\lambda_2(\mathbf{P})}}\right)$, we get that the deviation is at most $\delta + \frac{8d\log\log n}{1-\lambda_2} = \mathcal{O}(\frac{d\log\log n}{1-\lambda_2})$ at time t with probability $1-n^{-2}$.

For the second claim, assume that the initial discrepancy satisfies $K \leqslant n^C$ for a constant $C \geqslant 1$. By Theorem 2.2, after $t = \mathcal{O}(\frac{\log n}{1-\lambda_2})$ rounds the discrepancy of the idealized process is less than 1. Moreover, by the argument above and using a union bound, we find that the maximum deviation up to step t is at most $\mathcal{O}\left(\frac{d\log\log n}{1-\lambda_2}\right)$ with probability $1-n^{-2}$. Now at step t, the discrepancy of the discrete process is at most the deviation to the idealized process plus one, which is $\mathcal{O}\left(\frac{d\log\log n}{1-\lambda_2}\right)$. Since the discrepancy in the discrete process may not increase, and the discrepancy of the idealized process is also non-increasing and at most 1, the maximum deviation is $\mathcal{O}\left(\frac{d\log\log n}{1-\lambda_2}\right)$ also for all time steps larger than t. □

[3] Note that in contrast to Rabani *et al.* [19] we count the factor of d to make our results more comparable to the ones in Section 5.

We will see in Theorem 4.10 that the maximum deviation is indeed $\Theta(\log\log n)$ for constant-degree expanders. Before we prove this, we state the following implication of Theorem 4.6,

Corollary 4.7. *We achieve a discrepancy of $\mathcal{O}\left(\frac{d\log\log n}{1-\lambda_2(\mathbf{P})}\right)$ after $\mathcal{O}\left(\frac{d\log(Kn)}{1-\lambda_2(\mathbf{P})}\right)$ steps w.h.p.*

Proposition 4.8. *For any constant-degree expander G there is a round-matrix $\mathbf{P}$ consisting of at most $d+1$ matchings such that $1/(1-\lambda_2(\mathbf{P})) = \Theta(1)$. Hence, we can achieve a discrepancy of $\mathcal{O}(\log\log n)$ in time $\mathcal{O}(\log(Kn))$, where K is the initial discrepancy.*

Hence for the important class of constant-degree expanders, we achieve a discrepancy of $\mathcal{O}(\log\log n)$ in optimal time. We now apply our results to d-dimensional torus graphs. [19, Theorem 8] showed $\Psi_1(\mathbf{P}) = \Theta(d\,n^{1/d})$. Applying Corollary 4.3 and using the fact that $\Psi_2(\mathbf{P}) \leqslant \sqrt{\Psi_1(\mathbf{P})}$ gives the following.

Corollary 4.9. *Consider the d-dimensional torus with constant d. For any load vector with initial discrepancy K, we achieve after $\mathcal{O}(d\,n^{2/d}\log(Kn))$ rounds a discrepancy of $\mathcal{O}\left(\sqrt{d\,n^{1/d}\log n}\right)$ w.h.p.*

4.2 Lower Bounds

We now prove the following lower bound that matches Theorem 4.6.

Theorem 4.10. *Let G be an arbitrary d-regular expander graph with $d = \mathcal{O}(1)$. Then there is an initial load vector with discrepancy $K = \Theta(\log\log n)$ such that the maximum deviation between the discrete and idealized process is at least $\Omega(\log\log n)$ w.h.p.*

Proof. Choose a subset $S \subseteq V$ such that two vertices in S have a distance of at least $4c_1\log\log n$ to each other, where c_1 is a large constant to be determined later. Since for every vertex $v \in V$, $|B_{2c_1\log\log n}(v)| = \mathcal{O}(\log n)$, we can find such a subset S of size at least $\Omega(n/\log n)$ (cf. [9]). Define the load vector $x_i^{(0)} = \max\{0, c_2\log\log n - \mathrm{dist}(i,S)\}$, where $c_2 \leqslant c_1$ is a small constant that is specified later. Clearly, the initial discrepancy equals $c_2\log\log n$.

We start by examining the idealized process. Note that if we run the idealized process for less than $c_1\log\log n$ rounds, the load balancing processes within $B_{2c_1\log\log n}(s_1)$ and $B_{2c_1\log\log n}(s_2)$ for $s_1, s_2 \in S, s_1 \neq s_2$ are independent. Hence to compute $\xi^{c_1\log\log n}$ within $B_{2c_1\log\log n}(s)$ for a fixed $s \in S$, we may also replace $\xi^{(0)}$ by a vector $y^{(0)}$ which coincides with $x^{(0)}$ within $B_{2c_1\log\log n}(s)$, but is 0 elsewhere. Now the initial quadratic error can be bounded above by

$$\begin{aligned}
&\|y^{(0)} - \overline{y}\|_2^2 \\
&\leqslant \sum_{i=0}^{(c_2/4)\log\log n} 2\delta^i \left(\frac{c_2}{4}\log\log n - i - \frac{1}{n}\right)^2 + n\cdot\left(\frac{1}{n}\right)^2 \\
&= \mathcal{O}(\log n).
\end{aligned}$$

By Lemma 2.1, we define $c_1 > 0$ to be the constant such that after $c_1\log\log n$ rounds, $\|y^{(t)} - \overline{y}\|_2 \leqslant 1$. Hence by Lemma 2.1, we have for $t = c_1\log\log n, c_1 \in \mathcal{O}(1)$ that $\|y^{(t)} - \overline{y}\|_2 = \mathcal{O}(1)$. Since $\overline{y} = \mathcal{O}(1)$, we have for all $i \in B_{(c_1/2)\log\log n}(s_1)$, $\xi_i^{(t)} = \mathcal{O}(1)$.

Let us now consider the discrete process. Since each two different vertices $s_i, s_j \in S$ have a distance of at least $4c_1\log\log n$ to each other, the balancing processes within $B_{2c_1\log\log n}(s_i)$ and $B_{2c_1\log\log n}(s_j)$ are independent during the first $c_1\log\log n$ steps. Now the probability that in $B_{c_2\log\log n}(s), s \in S$ all matching edges are oriented towards s during the first $c_1\log\log n$ rounds is at least

$$2^{-|B_{c_2\log\log n}(s_i)|c_1\log\log n} \leqslant \frac{1}{\sqrt{n}},$$

if c_2 is sufficiently small chosen. Since we have at least $\Omega(n/\log n)$ independent events, there is at least one $s \in S$ such that all matching edges in $B_{c_2\log\log n}(s)$ are oriented towards s during the first $c_1\log\log n$ steps w.h.p. In this case,

$$|\xi_s^{(c_1\log\log n)} - x_s^{(c_1\log\log n)}| \geqslant c_2\log\log n - \mathcal{O}(1),$$

and the claim follows. □

The same proof technique in Theorem 4.10 also works for torus graphs.

Theorem 4.11. *For the d-dimensional torus graph with $d = \mathcal{O}(1)$, there is an initial load vector with discrepancy $K = \Theta(\mathrm{polylog}(n))$ such that the maximum deviation between the discrete and idealized process is $\Omega(\mathrm{polylog}(n))$ w.h.p.*

While the actual lower bound may very well be polynomial in n, this lower bound together with Theorem 4.6 still demonstrates an exponential gap between torus graphs and constant-degree expanders.

5. RANDOM MATCHINGS

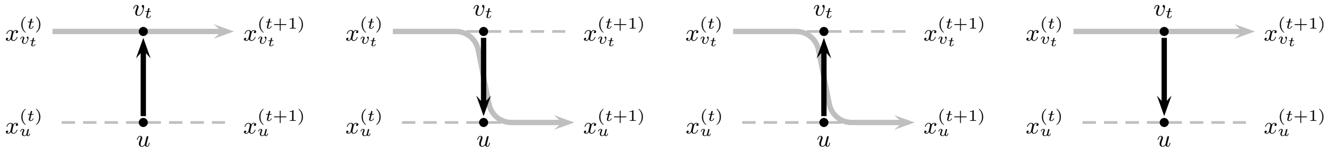
For certain graphs, it might be non-trivial to construct explicit matching matrices such that their eigenvalue gap (or local p-divergence) can be bounded. Also, for many graphs it is not obvious which matching sequences are to be considered "good". This motivates us to examine sequences of random matchings. We choose independently for each $\mathbf{P}^{(i)}$ a random matching by a local algorithm. There are many such algorithms available, e.g. the LR algorithm of Ghosh and Muthukrishnan [11] or the distributed synchronous algorithm of Boyd, Ghosh, Prabhakar and Shah [5].

Definition 5.1. *We call a sequence of matchings $M^{(t)}$, $t \in [T]$, a random matching if*

(i) There is a constant $\alpha > 0$ such that $\mathbf{Pr}\left[\{i,j\} \in M^{(t)}\right] \geqslant \alpha/\Delta$ for all $t \in [T]$ and $\{i,j\} \in E$.

(ii) The random decisions within one time-step do not pairwise correlate negatively, that is, $\mathbf{Pr}\left[\{i,j\} \in M^{(t)} \mid \{u,v\} \in M^{(t)}\right] \geqslant \mathbf{Pr}\left[\{i,j\} \in M^{(t)}\right]$ for all $t \in [T]$ and $\{i,j\} \cap \{u,v\} = \emptyset$.

(iii) All random decisions between different time-steps are independent.



(a) $x_{v_t}^{(t)} \geqslant x_u^{(t)}$. (b) $x_{v_t}^{(t)} \geqslant x_u^{(t)}$. (c) $x_{v_t}^{(t)} < x_u^{(t)}$. (d) $x_{v_t}^{(t)} < x_u^{(t)}$.

Figure 1. Illustration of the four cases of Definition 5.6 with $\{v_t, u\} \in M^{(t)}$.

Throughout this section, we will denote by λ_2 the second largest eigenvalue of $\mathbf{Q}$ in absolute value. Our aim is to state bounds in terms of the (explicit) eigenvalue $\lambda_2(\mathbf{Q})$ instead of the eigenvalues of the matching matrices (which are random variables).

5.1 Coarse balancing and Preliminaries

Let us first consider the idealized process. Consider an arbitrary step t. Let $\xi^{(t)}$ be any n-dimensional load vector and let $\xi^{(t+1)} = \xi^{(t)}\mathbf{P}^{(t+1)}$, where $\mathbf{P}^{(t)}$ represents the random matching matrix of round t. The following basic lemma asserts that the quadratic error decreases exponentially in $\lambda_2 = \lambda_2(\mathbf{Q})$, if α is a constant greater than zero.

Lemma 5.2. *(from [11]) Consider a randomly generated matching with the property that each $\{i,j\} \in E$ is chosen with probability at least $\frac{\alpha}{\Delta}$ with $0 < \alpha \leqslant 1$. Then,*

$$\mathbf{E}\left[\|\xi^{(t-1)} - \bar{x}\|_2^2 - \|\xi^{(t)} - \bar{x}\|_2^2\right] \geqslant \frac{\alpha}{3}\,(1 - \lambda_2(\mathbf{Q}))\,\|\xi^{(t-1)} - \bar{x}\|_2^2.$$

Using conditional expectations and Markov's inequality, we obtain the following theorem (cf. [5] for a similar statement).

Theorem 5.3. *For any initial vector $\xi^{(0)}$ and any time step $t \geqslant 1$,*

$$\mathbf{Pr}\left[\left\|\xi^{(t)} - \bar{x}\right\|_2^2 \geqslant \left(1 - \frac{\alpha}{3}\,(1 - \lambda_2(\mathbf{Q}))\right)^{t/2}\left\|\xi^{(0)} - \bar{x}\right\|_2^2\right] \leqslant \left(1 - \frac{\alpha}{3}\,(1 - \lambda_2(\mathbf{Q}))\right)^{t/2}.$$

Using the theorem above and Corollary 3.4 we immediately obtain

Corollary 5.4. *The discrepancy after $\mathcal{O}\left(\frac{\log(Kn)}{1-\lambda_2}\right)$ rounds is at most $\mathcal{O}\left(\sqrt{\frac{\log n}{1-\lambda_2}}\,\log n\right)$ w.h.p.*

For expanders, where $1/(1-\lambda_2)$ is a constant, this result is rather weak, as it will be shown by a more involved analysis below. To this end we state the following basic lemma that relates the short-term decrease of $\|\xi^t\|_2^2$ to the eigenvalue gap.

Lemma 5.5. *Consider a randomly generated matching with the property that each $\{i,j\} \in E$ is chosen with probability at least $\frac{\alpha}{\Delta}$ with $0 < \alpha \leqslant 1$. Then if $\|\xi^{(t)}\|_1 = 1$,*

$$\mathbf{Pr}\left[\left\|\xi^{(t)}\right\|_2^2 \leqslant \left(1 - \frac{\alpha^2}{9}\,(1 - \lambda_2(\mathbf{Q}))^2\right)\left\|\xi^{(t-1)}\right\|_2^2 + \frac{2}{n}\right] \geqslant 1 - \left(1 + \frac{\alpha}{3}\,(1 - \lambda_2(\mathbf{Q}))\right)^{-1}.$$

5.2 Near-Perfect Balancing

In this section, we address the problem of near-perfect load balancing. By this we mean a state where the discrepancy has been reduced to a constant value. Note that [19] gave results for perfect balancing where the discrepancy is reduced to 1. However, it is not too difficult to see that with a random orientation of all edges, a discrepancy of 1 can only be achieved in $\Omega(n)$ rounds for any graph (and any matchings) (cf. [17]). Therefore, we confine ourselves with achieving a constant discrepancy, where the constant is at least 2. We will see that on expanders for which $\Delta/\delta = \mathcal{O}(1)$, we can achieve a constant discrepancy in $\mathcal{O}(\log(kn)(\log\log n)^3)$ steps.

To describe how packages of tokens move through the network over time, we define *canonical paths*. This allows us to fix a certain package "routing" through the network and analyze the contribution to the load deviation only on this path. This fixed-view of the w paths is crucial for the analysis, as it avoids the necessity of a union bound over *all possible* paths.

Definition 5.6. *The sequence $(v_{t_1}, v_{t_1+1}, \ldots, v_{t_2})$ is called the* canonical path *of v_{t_1} from time t_1 to t_2 if for all times t with $t_1 < t \leqslant t_2$ the following holds: If v_t is unmatched in $M^{(t)}$, then $v_{t+1} = v_t$. Otherwise, let $u \in V$ be such that $\{v_t, u\} \in M^{(t)}$. Then,*

- *if $x_{v_t}^{(t)} \geqslant x_u^{(t)}$ and $\Phi_{v_t,u}^{(t)} = 1$ then $v_{t+1} = v_t$,*
- *if $x_{v_t}^{(t)} \geqslant x_u^{(t)}$ and $\Phi_{v_t,u}^{(t)} = -1$ then $v_{t+1} = u$,*
- *if $x_{v_t}^{(t)} < x_u^{(t)}$ and $\Phi_{v_t,u}^{(t)} = 1$ then $v_{t+1} = u$,*
- *if $x_{v_t}^{(t)} < x_u^{(t)}$ and $\Phi_{v_t,u}^{(t)} = -1$ then $v_{t+1} = v_t$.*

We will denote such a canonical path from t_1 to t_2 that starts at v as $\mathcal{P}_v^{[t_1,t_2]}$. Observe that the endpoint $d(\mathcal{P}_v)$ of the path depends on *(i)* the randomly chosen matched edges, *(ii)* their randomly chosen orientations and *(iii)* the load vector at step t_1. Also note that if two equal loads are balanced, the canonical path stays where it is iff the randomly chosen orientation of the edge points upwards.

We state two more facts about canonical paths.

Fact 5.7. *Given the load vector $x^{(t_1)}$ at step t_1, a vertex v and all chosen matching edges between step t_1 and t_2, a path $\mathcal{P}_v^{[t_1,t_2]}$ performs a simple random walk on G, i.e., whenever there is an adjacent matching edge, the path switches along this matching edge with probability $1/2$, and otherwise stays at the current vertex.*

Fact 5.8. *Two canonical paths $\mathcal{P}_{u_1}^{[t_1,t_2]} = (u_1, u_2, \ldots, u_\kappa)$ and $\mathcal{P}_{v_1}^{[t_1,t_2]} = (v_1, v_2, \ldots, v_\kappa)$ do not intersect on a vertex on the same time, that is, $u_i \neq v_i$ for all $i \in [\kappa]$.*

The statements and proofs of this section use different constants which are defined as follows.

DEFINITION 5.9. *Our proofs use the following constants.*

$$
\begin{aligned}
C_1 &:= (\alpha/20)^5 \cdot \gamma^{-5} > 0,\\
C_2 &:= -\tfrac{50}{C_1} \ln\left(\tfrac{1-\exp(-C_1/50)}{8}\right) > 0,\\
C_3 &:= 1 - \tfrac{\alpha^2}{324}\,(1-\lambda_2(\mathbf{Q}))^2 \in (0,1),\\
C_4 &:= 1/(1 + \tfrac{\alpha}{54}\,(1-\lambda_2(\mathbf{Q}))) \in (0,1),\\
C_5 &:= C_1\,(1-C_4)/2 > 0,\\
C_6 &:= \max\left\{0,\ \tfrac{24\ln(2)-8\,\ln\left(1-\exp(-\frac{C_5}{8})\right)}{C_5} - C_2\right\} \geqslant 0,\\
C_7 &:= \tfrac{1}{1-C_4} + \tfrac{2\left(-4\ln\left((1-C_4^{1/3})\right)\right)^{1/2}}{1-C_4^{1/3}} + 1 > 0,\\
C_8 &:= -8/\log C_4 > 0,\\
C_9 &:= \sqrt{32((\lceil C_7 + C_6 + C_2 + 2\rceil)^2 + 1/4)}\,C_8 > 0,\\
\vartheta &:= \lceil C_7 + C_2 + C_6 + 2\rceil > 2.
\end{aligned}
$$

For a canonical path $\mathcal{P}$, let $d(\mathcal{P})$ the destination of the path, that is the vertex at step t_2 of $\mathcal{P}$. It is easy to verify that a canonical path $\mathcal{P}_v^{[t_1,t_2]}$ visits in expectation at least $\frac{C_1}{6}(t_2 - t_1 + 1)$ times six different vertices in a row. The following definition will be important in the remainder.

DEFINITION 5.10. *For a canonical path $\mathcal{P}_v^{[t_1,t_2]}$, we define for $t_1 \leqslant t \leqslant t_2$, $Y_t = 1$, if $t \bmod 6 = 0$ and the canonical path visits six different vertices within the time-interval $[t, t+5]$, and $Y_t = 0$ otherwise.*

For a fixed time step t_2, an arbitrary time step $t_1 \leqslant t_2$, and a vertex v, we set $\Lambda_v^{(t_1)} := \big(\prod_{k=t_1}^{t_2} \mathbf{P}^{(k)}\big)u_v$, where u_v is the v-th unit-vector.

DEFINITION 5.11. *For every canonical path $\mathcal{P} = \mathcal{P}_v^{[t_1,t_2]}$ with destination $d(\mathcal{P})$, we define the following three events,*

$$
\begin{aligned}
\textsc{Trans}(\mathcal{P}) &:= \left\{\bigwedge_{\ell=t_1}^{t_2-C_2} \sum_{t=\ell+1}^{t_2} Y_t \geqslant \frac{C_1}{2}(t_2-\ell)\right\},\\
\textsc{Bal}(\mathcal{P}) &:= \left\{\bigwedge_{t=t_1}^{t_2-C_2-C_6} \left(\big\|\Lambda_{u_{d(\mathcal{P})}}^{(t)}\big\|_2^2 \leqslant {C_4}^{(t_2-t)} + \frac{t_2-t}{n}\right)\right\},\\
\textsc{Success}(\mathcal{P}) &:= \left\{\left|\sum_{\ell=t_1}^{t_2} \sum_{[i:j]\in M_{\mathsf{Odd}}^{(\ell)}} \Phi_{i,j}^{(\ell)} w_{i,j}^{[\ell,t_2]}(d(\mathcal{P}))\right| \leqslant C_7 + C_2 + C_6\right\}.
\end{aligned}
$$

Let us describe the intuitions behind these definitions. $\textsc{Trans}(\mathcal{P})$ denotes the event that the canonical path visits approximately the correct number of times six different vertices. $\textsc{Bal}(\mathcal{P})$ means that ℓ_2-norm of the weights of the random variables (representing the rounding errors) at round t that contribute to $x_{d(\mathcal{P})}^{(t_2)}$ is exponentially decreasing in $t - t_2$. Finally, $\textsc{Success}(\mathcal{P})$ denotes the event that the contribution of the rounding errors "close" to $x_{d(\mathcal{P})}^{(t_2)}$ is small. Our strategy is first to prove that $\textsc{Trans}(\mathcal{P})$ holds with constant probability and then prove the implications

$$\textsc{Trans}(\mathcal{P}) \Rightarrow \textsc{Bal}(\mathcal{P}) \Rightarrow \textsc{Success}(\mathcal{P}).$$

LEMMA 5.12. *For every canonical path $\mathcal{P}$,* $\mathbf{Pr}\left[\textsc{Trans}(P)\right] \geqslant \frac{7}{8}$.

LEMMA 5.13. *For every canonical path $\mathcal{P}$,* $\mathbf{Pr}\left[\textsc{Bal}(\mathcal{P}) \mid \textsc{Trans}(\mathcal{P})\right] \geqslant \frac{7}{8}$.

LEMMA 5.14. *For arbitrary t_1, t_2 with $t_1 < t_2$, $t_2 - t_1 \leqslant \log n$, let $\mathcal{P} = \mathcal{P}_v^{[t_1,t_2]}$ be a canonical path. Then,* $\mathbf{Pr}\left[\textsc{Success}(\mathcal{P}) \wedge \textsc{Bal}(\mathcal{P})\right] \geqslant 1/2$.

For the proof of our main result, we next introduce the following potential function. Then we prove that if the events $\textsc{Success}(\mathcal{P})$ and $\textsc{Bal}(\mathcal{P})$ hold, the improvement of the potential along a fixed canonical path is significant.

DEFINITION 5.15. *We define the potential in round t as $\Upsilon^{(t)} = \sum_{i=1}^n \Upsilon_i^{(t)}$, where $\Upsilon_i^{(t)} = \big(\big|x_i^{(t)} - \lfloor\bar{x}\rfloor\big| - \vartheta\big)^2$ if $|x_i^{(t)} - \lfloor\bar{x}\rfloor| \geqslant \vartheta$, and $\Upsilon_i^{(t)} = 0$ otherwise. Moreover, we define the improvement of the potential by $\Delta^{(t)} := \Upsilon^{(t)} - \Upsilon^{(t+1)}$ and the local change of a vertex i by $\Delta_i^{(t)} := \Upsilon_i^{(t)} - \Upsilon_i^{(t+1)}$.*

For simplicity, we may assume that $\lfloor\bar{x}\rfloor = 0$ when dealing with $\Upsilon^{(t)}$ in the following, which can be justified as follows. Clearly, for any initial load vector $x^{(0)}$ there is a number $\gamma \in \mathbb{Z}$ such that the average of $\tilde{x}^{(0)} = x^{(0)} + \gamma \cdot \mathbf{1}$ is between 0 and 1, where $\mathbf{1}$ denotes the all-ones vector. By definition of our load balancing algorithm, $\tilde{x}^{(t)} = x^{(t)}$ for any step t, in particular, the discrepancy of $\tilde{x}^{(t)}$ and $x^{(t)}$ are the same.

We further observe the following properties of Υ.

LEMMA 5.16. *For any graph, any initial load vector $x^{(0)}$, and time $t \in \mathbb{N}$,*

(i) $\Upsilon^{(t)} \in \mathbb{N}$,
(ii) $\Upsilon^{(t+1)} \leqslant \Upsilon^{(t)}$,
(iii) *If $\{i,j\} \in M^{(t)}$, then* $\Delta_i^{(t)} + \Delta_j^{(t)} \geqslant \frac{(x_i^{(t)} - x_j^{(j)})^2}{4} - \vartheta^2$.
(iv) *If $\vartheta - 1 \leqslant x_i^{(t)} \leqslant x_j^{(t)} + 2$ and $\{i,j\} \in M^{(t)}$, then* $\Delta_i^{(t)} + \Delta_j^{(t)} \geqslant 1$.
(v) $\Upsilon^{(t)} = 0$ *implies that $x^{(t)}$ has discrepancy at most* $2\vartheta + 1$.

The definition of the improvement of the potential is naturally extended to canonical paths as follows. Let $\mathcal{P} = \mathcal{P}_v^{[t_1,t_2]}$ be a canonical path and for each $t \in [t_1, t_2]$, let v_t be the vertex at step t visited by $\mathcal{P}$. Define

$$\Delta(\mathcal{P})^{(t)} := \begin{cases} 0 & \text{if } v_t \text{ is unmatched at step } t,\\ \frac{1}{2}\left(\Delta_{v_t}^{(t)} + \Delta_u^{(t)}\right) & \text{if } v_t \text{ is matched with } u \text{ at step } t.\end{cases}$$

LEMMA 5.17. *Consider a canonical path $\mathcal{P}_v^{[t_1,t_2]}$ of length $C_8 \log\log n$. If $|x_v^{(t_1)}| \geqslant \vartheta + 1$, $|x_{d(\mathcal{P})}^{(t_2)}| \leqslant \vartheta$, then $\Delta(\mathcal{P}) = \sum_{t=t_1}^{t_2} \Delta(\mathcal{P})^{(t)} \geqslant \frac{1}{4}\big(\frac{|x_v^{(t_1)}| - \vartheta}{C_9 \log\log n}\big)^2$.*

Proof. Fix a vertex v and consider the canonical path $\mathcal{P}_v^{[t_1,t_2]}$. We proceed by a case distinction on $|x_v^{(t_1)}|$. First, suppose that $\vartheta + 1 \leqslant |x_v^{(t_1)}| \leqslant \sqrt{32(\vartheta^2 + 1/4)}\,C_8 \log\log n + \vartheta$. By $|x_{d(\mathcal{P})}^{(t_2)}| \leqslant \vartheta$ and Definition 5.10 there is at least one step t on the path $\mathcal{P}$ on which $x_{v'}^{(t)}$ is matched with another vertex whose load is at least two tokens closer to $\lfloor\bar{x}\rfloor = 0$ and $|x_{v'}^{(t)}| \geqslant \vartheta + 1$. Hence in this case by Lemma 5.16,

$$\sum_{t=t_1}^{t_2} \Delta(\mathcal{P})^{(t)} \geqslant 1 \geqslant \left(\frac{|x_v^{(t_1)}| - \vartheta}{\sqrt{32(\vartheta^2 + 1/4)}\,C_8 \log\log n}\right)^2.$$

For the second case, let $|x_v^{(t_1)}| \geqslant \sqrt{32(\vartheta^2 + 1/4)}\,C_8 \log\log n + \vartheta$. For simplicity, we will assume in the following calculations that each vertex v_t on $\mathcal{P}_v^{[t_1,t_2]}$ is matched at step t with another vertex u_t when $\mathcal{P}$ is located on v_t.

$$\begin{aligned}
2\sum_{t=t_1}^{t_2}\Delta(\mathcal{P})^{(t)} &= \sum_{t=t_1}^{t_2}\Delta_{v_t}^{(t)}+\Delta_{u_t}^{(t)}\\
&\geqslant \sum_{t=t_1}^{t_2}\frac{(x_{v_t}^{(t)}-x_{u_t}^{(t)})^2}{4}-\vartheta^2\\
&\geqslant \sum_{t=t_1}^{t_2}\frac{(2x_{v_t}^{(t)}-2x_{v_t}^{(t+1)}\pm 1)^2}{4}-\vartheta^2\\
&\geqslant \sum_{t=t_1}^{t_2}\Big((x_{v_t}^{(t)}-x_{v_t}^{(t+1)})^2-(x_{v_t}^{(t)}-x_{v_t}^{(t+1)})/2\Big)\\
&\qquad -(\vartheta^2+1/4)(t_2-t_1)\\
&\geqslant \sum_{t=t_1}^{t_2}\Big((x_{v_t}^{(t)}-x_{v_t}^{(t+1)})^2/2\Big)-(\vartheta^2+1/4)(t_2-t_1)
\end{aligned}$$

as $(x_{v_t}^{(t)}-x_{v_t}^{(t+1)})\in\mathbb{N}$. The sum in the previous expression is minimized when all $x_{v_t}^{(t)}-x_{v_t}^{(t+1)}$ are the same. Therefore,

$$\begin{aligned}
&2\sum_{t=t_1}^{t_2}\Delta(\mathcal{P})^{(t)}\\
&\geqslant \sum_{t=t_1}^{t_2}\left(\frac{|x_v^{(t_1)}|-\vartheta-2C_8\log\log n}{2C_8\log\log n}\right)^2-(\vartheta^2+1/4)(t_2-t_1)\\
&\geqslant \sum_{t=t_1}^{t_2}\left(\frac{|x_v^{(t_1)}|-\vartheta}{4C_8\log\log n}\right)^2-(\vartheta^2+1/4)(t_2-t_1)\\
&\geqslant (t_2-t_1)\left(\left(\frac{|x_v^{(t_1)}|-\vartheta}{4C_8\log\log n}\right)^2-(\vartheta^2+1/4)\right)\\
&\geqslant \left(\frac{|x_v^{(t_1)}|-\vartheta}{4C_8\log\log n}\right)^2-(\vartheta^2+1/4)\\
&\geqslant \left(\frac{|x_v^{(t_1)}|-\vartheta}{\sqrt{32}\,C_8\log\log n}\right)^2,
\end{aligned}$$

where the last inequality holds since $|x_v^{(t_1)}|\geqslant\sqrt{32(\vartheta^2+1/4)}\,C_8\log\log n+\vartheta$ implies that

$$\left(\frac{|x_v^{(t_1)}|-\vartheta}{4C_8\log\log n}\right)^2\geqslant 2(\vartheta^2+1/4). \qquad\square$$

The next lemma follows directly by Theorem 3.3 and a union bound over all n vertices and $(\log n)^2$ time steps.

Lemma 5.18. *Let $\mathcal{E}_1$ be the event that for all vertices $v\in[n]$ and all timesteps $t_1<t=\mathcal{O}(\log^2 n)$ with $\left\|\big(\prod_{k=t_1+1}^{t}\mathbf{P}^{(k)}\big)u_v\right\|_2^2\leqslant(\log n)^{-4}$ it holds that $\left|\sum_{\ell=1}^{t_1}\sum_{[i:j]\in M_{\mathsf{Odd}}^{(\ell)}}\Phi_{i,j}^{(\ell)}w_{i,j}^{(\ell)}(v)\right|\leqslant 1$. Then, $\mathbf{Pr}[\mathcal{E}_1]\geqslant 1-n^{-2}$.*

Intuitively, the lemma says that the rounding errors from step 1 to t_1 for the load of a vertex v at step t never cause a large deviation, provided that the load is "well distributed" at the neighborhood of v during the steps t_1 and t_2.

The next lemma is a simple application of Theorem 5.3.

Lemma 5.19. *Let $\mathcal{E}_2$ denote the event that the discrepancy of the idealized process is 1 after $C\log(Kn)/(1-\lambda_2)$ steps, where C is a sufficiently large constant. Then, $\mathbf{Pr}[\mathcal{E}_2]\geqslant 1-n^{-4}$.*

Basically, we now prove that the event Success($\mathcal{P}$) implies that the vertex $w=d(\mathcal{P})$ at the end of a canonical path $\mathcal{P}$ deviates from the average by only a constant. Then we use this fact together with Lemma 5.17 to show that the potential along each $\mathcal{P}$ with Success($\mathcal{P}$) decreases.

Lemma 5.20. *Let $t_1\geqslant\Omega(\log(Kn)/(1-\lambda_2))$. Assume that $\mathcal{E}_1$ and $\mathcal{E}_2$ hold and consider a canonical path $\mathcal{P}=\mathcal{P}_v^{[t_1,t_2]}$, $t_2=t_1+\frac{8\log\log n}{-\log C_4}$ with destination $d(\mathcal{P})=w$ such that* Success($\mathcal{P}$) *and* Bal($\mathcal{P}$) *holds. Then, $\left|x_w^{t_2}-\bar{x}\right|\leqslant C_7+C_6+C_2+2$.*

Now we are ready to prove the main result of this section.

Theorem 5.21. *For expanders with a constant degree-ratio Δ/δ, the discrepancy of any initial load vector of discrepancy K is reduced to $2\vartheta+2$ within $\mathcal{O}(\log(Kn)(\log\log n)^3)$ steps.*

Proof. Choose $t_2=t_1+\frac{8\log\log n}{-\log C_4}$. Moreover assume that we know $\Upsilon^{(t_1)}$. Then by linearity of expectations, the fact that two canonical paths are vertex-disjoint and Lemmas 5.14, 5.18, 5.19, and 5.20,

$$\begin{aligned}
&\mathbf{E}\left[\Upsilon^{(t_1)}-\Upsilon^{(t_2)}\mid\mathcal{E}_1\wedge\mathcal{E}_2\right]\\
&=\frac{1}{2}\sum_{v\in[n]}\mathbf{E}\left[\Delta(\mathcal{P}_v^{[t_1,t_2]})\mid\mathcal{E}_1\wedge\mathcal{E}_2\right]\\
&\geqslant\frac{1}{2}\mathbf{Pr}[\mathcal{E}_1\wedge\mathcal{E}_2]\sum_{v\in[n]}\mathbf{Pr}[\textsc{Success}(\mathcal{P}_v)\wedge\textsc{Bal}(\mathcal{P}_v)\mid\mathcal{E}_1\wedge\mathcal{E}_2]\\
&\quad\cdot\mathbf{E}\left[\Delta(\mathcal{P}_v^{[t_1,t_2]})\mid\textsc{Success}(\mathcal{P}_v)\wedge\textsc{Bal}(\mathcal{P}_v)\wedge\mathcal{E}_1\wedge\mathcal{E}_2\right]\\
&\geqslant\frac{1}{4}\cdot(1-4n^{-2})\cdot\sum_{v\in[n]}\left(\frac{|x_v^{(t_1)}|-\vartheta}{C_9\log\log n}\right)^2\\
&\geqslant\frac{1}{8(C_9\log\log n)^2}\Upsilon^{(t_1)}.
\end{aligned}$$

It follows from Corollary 5.4 that $\mathbf{E}\left[\Upsilon^{(t_1)}\right]\leqslant n^2$ w.h.p., if $t_1=\Theta\big(\frac{\log(Kn)}{1-\lambda_2}\big)$. As shown above for every pair of timesteps $t_1,t_2=t_1+\frac{8\log\log n}{-\log C_4}$,

$$\mathbf{E}\left[\Upsilon^{(t_2)}\right]\leqslant\left(1-\frac{1}{8(C_9\log\log n)^2}\right)\Upsilon^{(t_1)}.$$

Iterating this argument yields

$$\begin{aligned}
&\mathbf{E}\left[\Upsilon^{(t_1+8(C_9\log\log n)^2C_8\log\log n\cdot 6\ln n)}\right]\\
&\qquad\leqslant\left(1-\frac{1}{8(C_9\log\log n)^2}\right)^{8(C_9\log\log n)^2\cdot 6\ln n}\Upsilon^{(t_1)}\\
&\qquad\leqslant n^{-6}\,n^2=n^{-4}.
\end{aligned}$$

Finally, using the integrality of Υ we arrive at

$$\begin{aligned}
&\mathbf{Pr}\left[|x^{(t_1+8(C_9 \log\log n)^2 C_8 \log\log n\cdot 6\ln n)} - \bar{x}|_\infty > \vartheta + 1\right] \\
&\leqslant \mathbf{Pr}\left[\Upsilon^{(t_1+8(C_9 \log\log n)^2 C_8 \log\log n\cdot 6\ln n)} > 0\right] \\
&\leqslant \mathbf{E}\left[\Upsilon^{(t_1+8(C_9 \log\log n)^2 C_8 \log\log n\cdot 6\ln n})\right] \\
&\leqslant n^{-4}. \qquad \square
\end{aligned}$$

Our techniques also apply to constant-degree expanders in the balancing circuit model.

COROLLARY 5.22. *Let G be a constant-degree expander and consider the balancing circuit model with a round matrix $\mathbf{P}$ satisfying $(1-\lambda_2(\mathbf{P}))^{-1} = \mathcal{O}(1)$. Then we can achieve a discrepancy of $\mathcal{O}(1)$ in $\mathcal{O}(\frac{\log(Kn)}{1-\lambda_2(\mathbf{P})}(\log\log n)^3)$ rounds w.h.p.*

6. CONCLUSIONS

We present the first analysis of a natural local load balancing algorithm that directs the excess tokens at random. It is shown that on many important graphs this simple randomization improves over its deterministic counterpart [19] by at least a quadratic factor for many graphs. For the important case of expanders, we show that the load is balanced almost perfectly within $\mathcal{O}(\log(Kn)(\log\log n)^3)$ rounds. This result is optimal up to a factor of $\mathcal{O}(\log\log n)^3)$ while all previous approaches were only optimal up to a factor of $\mathcal{O}(\log n)$. Nevertheless, an interesting open question is to close the gap between our upper bound of $\mathcal{O}(\log(Kn)(\log\log n)^3)$ and the (trivial) lower bound of $\Omega(\log(Kn))$.

Acknowledgments

We are thankful to Robert Elsässer, Alistair Sinclair and Danny Vilenchik for helpful discussions. This work was supported by postdoctoral fellowships from the German Academic Exchange Service (DAAD).

References

[1] W. Aiello, B. Awerbuch, B. M. Maggs, and S. Rao. Approximate load balancing on dynamic and asynchronous networks. In *25th Annual ACM Symposium on Theory of Computing (STOC'93)*, pages 632–641, 1993.

[2] H. Arndt. Load balancing: Dimension exchange on product graphs. In *18th International Parallel and Distributed Processing Symposium (IPDPS'04)*. IEEE Computer Society, 2004.

[3] J. Aspnes, M. Herlihy, and N. Shavit. Counting networks. *J. ACM*, 41(5):1020–1048, 1994.

[4] P. Berenbrink, T. Friedetzky, and Z. Hu. A new analytical method for parallel, diffusion-type load balancing. *J. Parallel Distrib. Comput.*, 69(1):54–61, 2009.

[5] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory and IEEE/ACM Transactions on Networking*, 52(6):2508–2530, 2006.

[6] F. Cedo, A. Cortes, A. Ripoll, M. A. Senar, and E. Luque. The convergence of realistic distributed load-balancing algorithms. *Theor. Comp. Sys.*, 41(4):609–618, 2007.

[7] J. Cooper and J. Spencer. Simulating a random walk with constant error. *Comb. Probab. Comput.*, 15:815–822, 2006.

[8] R. Elsässer, B. Monien, and S. Schamberger. Distributing unit size workload packages in heterogenous networks. *J. Graph Algorithms & Applications*, 10(1):51–68, 2006.

[9] U. Feige, D. Peleg, P. Raghavan, and E. Upfal. Randomized broadcast in networks. *Random Structures & Algorithms*, 1(4):447–460, 1990.

[10] B. Ghosh, F. T. Leighton, B. M. Maggs, S. Muthukrishnan, C. G. Plaxton, R. Rajaraman, A. W. Richa, R. E. Tarjan, and D. Zuckerman. Tight analyses of two local load balancing algorithms. *SIAM J. Comput.*, 29(1):29–64, 1999.

[11] B. Ghosh and S. Muthukrishnan. Dynamic load balancing by random matchings. *Journal of Computer and System Sciences*, 53(3):357–370, 1996.

[12] M. Herlihy and S. Tirthapura. Randomized smoothing networks. *J. Parallel and Distributed Computing*, 66(5):626–632, 2006.

[13] M. Klugerman. *Small-depth counting networks and related topics*. PhD thesis, Department of Mathematics, Massachusetts Institute of Technology, September 1994.

[14] M. Klugerman and C. Plaxton. Small-depth counting networks. In *24th Annual ACM Symposium on Theory of Computing (STOC'92)*, pages 417–428, 1992.

[15] L. Lovász and P. Winkler. Mixing of random walks and other diffusions on a graph. *Surveys in combinatorics*, pages 119–154, 1995.

[16] R. Lüling and B. Monien. A dynamic distributed load balancing algorithm with provable good performance. In *5th ACM Symposium on Parallel Algorithms and Architectures (SPAA'93)*, pages 164–172, 1993.

[17] M. Mavronicolas and T. Sauerwald. The impact of randomization in smoothing networks. In *27th Annual ACM Principles of Distributed Computing (PODC'08)*, pages 345–354, 2008.

[18] S. Muthukrishnan, B. Ghosh, and M. H. Schultz. First- and second-order diffusive methods for rapid, coarse, distributed load balancing. *Theory Comput. Syst.*, 31(4):331–354, 1998.

[19] Y. Rabani, A. Sinclair, and R. Wanka. Local divergence of Markov chains and the analysis of iterative load balancing schemes. In *39th Annual IEEE Symposium on Foundations of Computer Science (FOCS'98)*, pages 694–705, 1998.

[20] P. Raghavan and C. Thompson. Randomized rounding: a technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7:365–374, 1987.

[21] R. Subramanian and I. D. Scherson. An analysis of diffusive load-balancing. In *6th ACM Symposium on Parallel Algorithms and Architectures (SPAA'94)*, pages 220–225, 1994.