Stochastic coalescence in logarithmic time*

Po-Shen Loh† Eyal Lubetzky‡

Abstract

The following distributed coalescence protocol was introduced by Dahlia Malkhi in 2006 motivated by applications in social networking. Initially there are n agents wishing to coalesce into one cluster via a decentralized stochastic process, where each round is as follows: Every cluster flips a fair coin to dictate whether it is to issue or accept requests in this round. Issuing a request amounts to contacting a cluster randomly chosen *proportionally to its size*. A cluster accepting requests is to select an incoming one *uniformly* (if there are such) and merge with that cluster. Empirical results by Fernandess and Malkhi suggested the protocol concludes in $O(\log n)$ rounds with high probability, whereas numerical estimates by Oded Schramm, based on an ingenious analytic approximation, suggested that the coalescence time should be super-logarithmic.

Our contribution is a rigorous study of the stochastic coalescence process with two consequences. First, we confirm that the above process indeed requires super-logarithmic time w.h.p., where the inefficient rounds are due to oversized clusters that occasionally develop. Second, we remedy this by showing that a simple modification produces an essentially optimal distributed protocol: If clusters favor their *smallest* incoming merge request then the process *does* terminate in $O(\log n)$ rounds w.h.p., and simulations show that the new protocol readily outperforms the original one. Our upper bound hinges on a potential function involving the logarithm of the number of clusters and the cluster-susceptibility, carefully chosen to form a supermartingale. The analysis of the lower bound builds upon the novel approach of Schramm which may find additional applications: Rather than seeking a single parameter that controls the system behavior, instead one approximates the system by the Laplace transform of the entire cluster-size distribution.

1 Introduction

The following stochastic distributed coalescence protocol was proposed by Dahlia Malkhi in 2006, motivated by applications in social networking and the reliable formation of peer-to-peer networks (see [10] for more on these applications). The objective is to coalesce n participating agents into a single hierarchal cluster reliably and efficiently. To do so without relying on a centralized authority, the protocol first identifies each agent as a cluster (a singleton), and then proceeds in rounds as follows:

(i) Each cluster flips a fair coin to determine whether it will be *issuing* a merge-request or *accepting* requests in the upcoming round.

(ii) Issuing a request amounts to selecting another cluster randomly *proportionally to its size*.

(iii) Accepting requests amounts to choosing an incoming request (if there are any) *uniformly* at random and proceeding to merge with that cluster.

In practice, each cluster is in fact a layered tree whose root is entrusted with running the protocol, e.g. each root decides whether to issue/accept requests in a given round etc. When attempting to merge with another cluster, the root of cluster $\mathcal{C}_i$ simply chooses a vertex v uniformly out of $[n]$, which then propagates the request to its root. This therefore corresponds to choosing the cluster $\mathcal{C}_j$ proportionally to $|\mathcal{C}_j|$. This part of the protocol is well-justified by the fact that agents within a cluster typically have no information on the structure of other clusters in the system.

A second feature of the protocol is the symmetry between the roles of issuing/accepting requests played by the clusters. Clearly, every protocol enjoying this feature would have (roughly) at most half of its clusters become acceptors in any given round, and as such could terminate within $O(\log n)$ rounds. Furthermore, on an intuitive level, as long as all clusters are of roughly the same size (as is the case initially), there are few "collisions" (multiple clusters issuing a request to the same cluster) each round and hence the effect of a round is similar to that of merging clusters according to a random perfect matching. As such, one might expect that the protocol should conclude with a roughly balanced binary tree in logarithmic time.

Indeed, empirical evidence by Fernandess and Malkhi [11] showed that this protocol seems highly ef-

*Research supported in part by a Fannie and John Hertz Foundation Fellowship, an NSF Graduate Research Fellowship, and a Princeton Centennial Fellowship.

†Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213. Email: ploh@cmu.edu.

‡Microsoft Research, One Microsoft Way, Redmond, WA 98052. Email: eyal@microsoft.com.

ficient, typically taking a logarithmic number of rounds to coalesce. However, rigorous performance guarantees for the protocol were not available.

While there are numerous examples of stochastic processes that have been successfully analyzed by means of identifying a single tractable parameter that controls their behavior, here it appears that the entire distribution of the cluster-sizes plays an essential role in the behavior of the system. Demonstrating this is the following example: Suppose that the cluster $\mathcal{C}_1$ has size $n - o(\sqrt{n})$ while all others are singletons. In this case it is easy to see that with high probability all of the merge-requests will be issued to $\mathcal{C}_1$, who will accept at most one of them (we say an event holds *with high probability*, or w.h.p. for brevity, if its probability tends to 1 as $n \to \infty$). Therefore, starting from this configuration, coalescence will take at least $n^{1/2-o(1)}$ rounds w.h.p., a polynomial slowdown. Of course, this scenario is extremely unlikely to arise when starting from n individual agents, yet possibly other mildly unbalanced configurations *are* likely to occur and slow the process down.

In 2007, Oded Schramm proposed a novel approach to the problem, approximately reducing it to an analytic problem of determining the asymptotics of a recursively defined family of real functions. Via this approximation framework Schramm then gave numerical estimates suggesting that the running time of the stochastic coalescence protocol is w.h.p. super-logarithmic. Unfortunately, the analytical problem itself seemed highly nontrivial and overall no bounds for the process were known.

1.1 New results In this work we study the stochastic coalescence process with two main consequences. First, we provide a rigorous lower bound confirming that this process w.h.p. requires a super-logarithmic number of rounds to terminate. Second, we identify the vulnerability in the protocol, namely the choice of which merge-request a cluster should approve: While the original choice seems promising in order to maintain the balance between clusters, it turns out that typical deviations in cluster-sizes are likely to be amplified by this rule and lead to irreparably unbalanced configurations. On the other hand, we show that a simple modification of this rule to favor the smallest incoming request is already enough to guarantee coalescence in $O(\log n)$ rounds w.h.p. (Here and in what follows we let $f \lesssim g$ denote that $f = O(g)$ while $f \asymp g$ is short for $f \lesssim g \lesssim f$.)

THEOREM 1.1. *The uniform coalescence process $\mathcal{U}$ coalesces in $\tau_c(\mathcal{U}) \gtrsim \log n \cdot \frac{\log\log n}{\log\log\log n}$ rounds w.h.p. Consider a modified* size-biased *process $\mathcal{S}$ where every accepting cluster $\mathcal{C}_i$ has the rule:*

- *Ignore requests from clusters of size larger than $|\mathcal{C}_i|$.*
- *Among other requests (if any) select one issued by a cluster $\mathcal{C}_j$ of smallest size.*

Then the coalescence time of the size-biased process satisfies $\tau_c(\mathcal{S}) \asymp \log n$ w.h.p.

Observe that the new protocol is easy to implement efficiently in practice as each root can keep track of the size of its cluster and can thus include it as part of the merge-request.

1.2 Empirical results Our simulations show that the running time of the size-biased process is approximately $5 \log_2 n$. Moreover, they further demonstrate that the new size-biased process empirically performs substantially better than the uniform process even for fairly small values of n, i.e. the improvement appears not only asymptotically in the limit but already for ordinary input sizes. These results are summarized in Figure 1, where the plot on the left clearly shows how the uniform process diverges from the linear (in logarithmic scale) trend corresponding to the runtime of the size-biased process. The rightmost plot identifies the crux of the matter: the uniform process rapidly produces a highly skewed cluster-size distribution, which slows it down considerably.

1.3 Related work There is extensive literature on stochastic coalescence processes whose various flavors fit the following scheme: The clusters act via a continuous-time process where the coalescence rate of two clusters with given masses x, y (which can be either discrete or continuous) is dictated up to re-scaling by a rate kernel K. A notable example of this is Kingman's coalescent [18], which corresponds to the kernel $K(x, y) = 1$ and has been intensively studied in mathematical population genetics; see e.g. [7] for more on Kingman's coalescent and its applications in genetics. Other rate kernels that have been thoroughly studied include the additive coalescent $K(x, y) = x + y$ which corresponds to Aldous's continuum random tree [1], and the multiplicative coalescent $K(x, y) = xy$ that corresponds to Erdős-Rényi random graphs [9] (see the books [4, 17]). For further information on these as well as other coalescence processes, whose applications range from physics to chemistry to biology, we refer the reader to the excellent survey of Aldous [2].

A major difference between the classical stochastic coalescence processes mentioned above and those studied in this work is the synchronous nature of the latter ones: Instead of individual merges whose occurrences are governed by independent exponentials, here

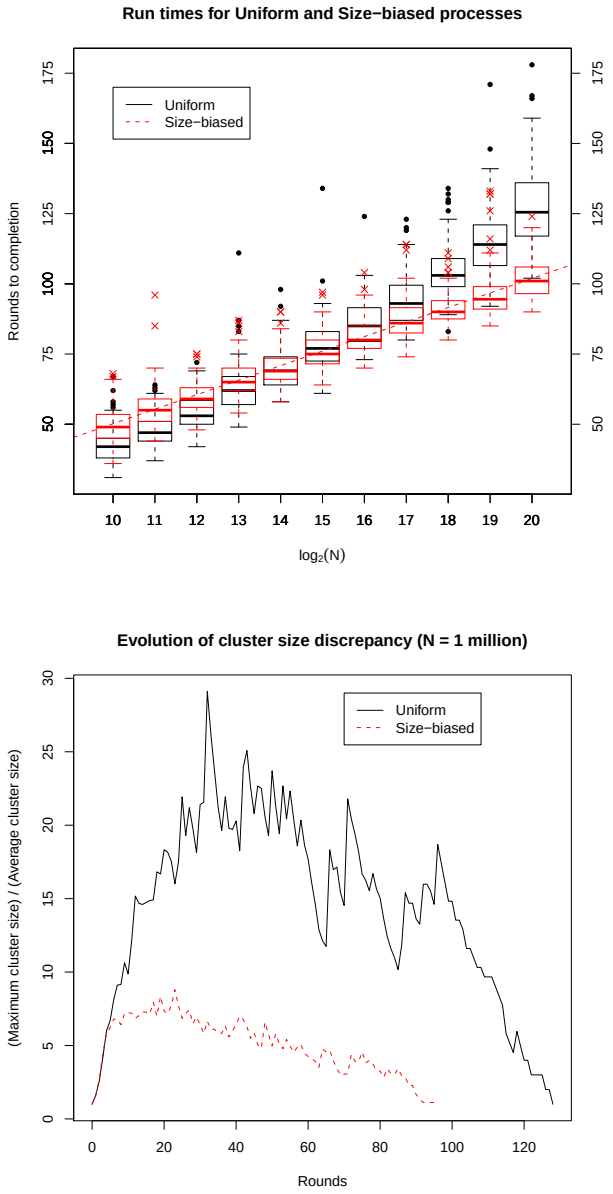

Figure 1: The top plot compares the running times for the two processes. Statistics are derived from 100 independent runs of each process, for each $n \in \{1024, 2048, \ldots, 2^{20}\}$. The bottom plot tracks the ratio between the maximum and average cluster-sizes, through a single run of each process, for $n = 10^6$. There, the uniform process took 128 rounds, while the size-biased process finished in 96.

the process is comprised of rounds where all clusters act simultaneously and the outcome of a round (multiple disjoint merges) is a function of these combined actions. This framework introduces delicate dependencies between the clusters, and rather than having the coalescence rate of two clusters be given by the rate kernel K as a function of their masses, here it is a function of the entire cluster distribution. For instance, suppose nearly all of the mass is in one cluster $\mathcal{C}_i$ (which thus attracts almost all merge requests); its coalescence rate with a given cluster $\mathcal{C}_j$ in the uniform coalescence process $\mathcal{U}$ clearly depends on the total number of clusters at that given moment, and similarly in the size-biased coalescence process $\mathcal{S}$ it depends on the sizes of all other clusters, viewed as competing with $\mathcal{C}_j$ over this merge. In face of these mentioned dependencies, the task of analyzing the evolution of the clusters along the high-dimensional stochastic processes $\mathcal{U}$ and $\mathcal{S}$ becomes highly nontrivial.

In terms of applications and related work in Computer Science, the processes studied here have similar flavor to those which arose in the 1980's, most notably the Random Mate algorithm introduced by Reif, and used by Gazit [15] for parallel graph components and by Miller and Reif [20] for parallel tree contraction. However, as opposed to the setting of those algorithms, a key difference here is the fact that as the process evolves through time each cluster is oblivious to the distribution of its peers at any given round (including the total number of clusters for that matter). Therefore for instance it is impossible for a cluster to sample from the uniform distribution over the other clusters when issuing its merge request.

For another related line of works in Computer Science, recall that the coalescence processes studied in this work organize n agents in a hierarchic tree, where each merged cluster reports to its acceptor cluster. This is closely related to the rich and intensively studied topic of Randomized Leader Elections (see e.g. [6, 12, 22, 23, 27]), where a computer network comprised of n processors attempts to single out a leader (in charge of communication, etc.) by means of a distributed randomized process generating the hierarchic tree. Finally, studying the dynamics of randomly merging sets is also fundamental to understanding the average-case performance of disjoint-set data structures (see e.g. the works of Bollobás and Simon [5], Knuth and Schönhage [19] and Yao [26]). These structures, which are of fundamental importance in computer science, store collections of disjoint sets, and support two operations: (i) taking the union of a pair of sets, and (ii) determining which set a particular element is in. See e.g. [14] for a survey of these data structures. The pro-

cesses studied here precisely consider the evolution of a collection of disjoint sets under random merge operations, and it is plausible that the tools used here could contribute to advances in that area.

1.4 Main techniques As we mentioned above, the main obstacle in the coalescence processes studied here is that since requests go to other clusters with probability proportional to their size, the largest clusters can create a bottleneck, absorbing all requests yet each granting only one per round. An intuitive approach for analyzing the size-biased process $\mathcal{S}$ would be to track a statistic that would warn against this scenario, with the most obvious candidate being the size of the largest cluster. However, simulations indicate that this alone will be insufficient as the largest cluster does in fact grow out of proportion in typical runs of the process. Nevertheless, the distribution of large clusters turns out to be sparse. The key idea is then to track a smoother parameter involving the *susceptibility*, which is essentially the second moment of the cluster-size distribution.

To simplify notation normalize the cluster-sizes to sum to 1 so that the initial distribution consists of n clusters of size $\frac{1}{n}$ each. With this normalization, the susceptibility χ_t is defined as the sum of squares of cluster-sizes after the t-th round. (We note in passing that this parameter has played a central role in the study of the phase-transition in Percolation and Random Graphs, see e.g. [16, 25].) The proof that the size-biased protocol is optimal hinges on a carefully chosen potential function $\Phi_t = \chi_t\kappa_t + C\log\kappa_t$, where κ_t denotes the number of clusters after the t-th round and C is an absolute constant chosen to turn Φ_t into a supermartingale. We will control the evolution of Φ_t and prove our upper bound on the running time of the size-biased process.

The analysis of the uniform process $\mathcal{U}$ is delicate and relies on rigorizing and analyzing the novel framework of Schramm [24] for approximating the problem by an analytic one. We believe this technique is of independent interest and may find additional applications in analyzing high-dimensional stochastic processes. Instead of seeking a single parameter to summarize the system behavior, one instead measures the system via a Laplace transform of the entire cluster-size distribution:

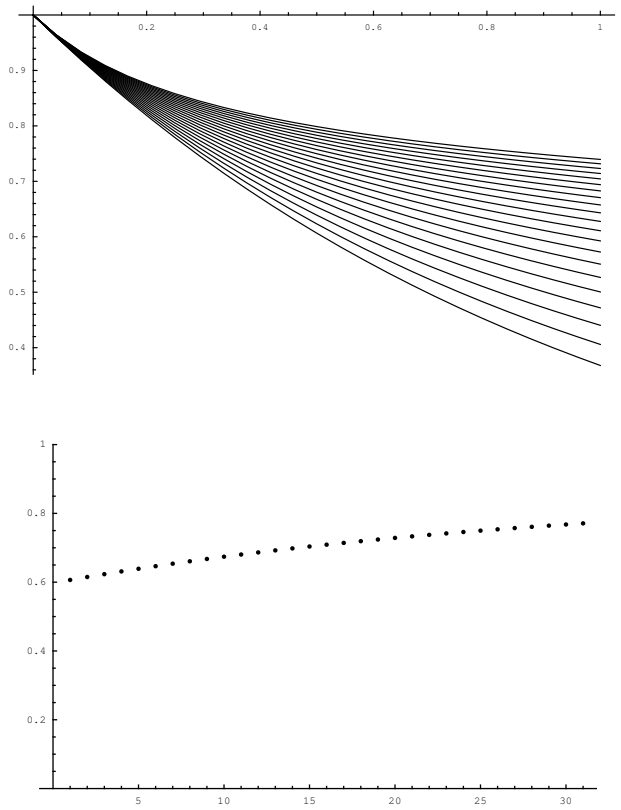

Figure 2: Numerical estimations by Oded Schramm for the functions $G_t(s)$ from his analytic approximation of the uniform coalescence process. The top plot features $G_t(s)$ for $t = \{0, 2, \ldots, 40\}$ and $s \in [0, 1]$ and demonstrates how these increase with t. The bottom plot focuses on $G_t(\frac{1}{2})$ and suggests that $G_t(\frac{1}{2}) \to 1$ and that in turn the coalescence rate should be super-logarithmic.

DEFINITION 1. *For any integer $t \geq 0$ let $\mathcal{F}_t$ be the σ-algebra generated by the first t rounds of the process. Conditioned on $\mathcal{F}_t$, define the functions $F_t(s)$ and $G_t(s)$ on the domain $\mathbb{R}$ as follows. Let κ be the number of clusters and let $w_1, \ldots, w_\kappa$ be the normalized cluster-sizes after t rounds. Set*

$$F_t(s) = \sum_{i=1}^{\kappa} \exp(-w_i\, s)\,, \quad G_t(s) = \frac{1}{\kappa} F_t(\kappa\, s)\,. \tag{1.1}$$

As we will further explain in Section 2, the Laplace transform F_t simultaneously captures all the moments of the cluster-size distribution, in a manner analogous to the moment generating function of a random variable. This form is particularly useful in our application as it turns out that the specific evaluation $G_t\big(\frac{1}{2}\big)$ governs the expected coalescence rate. Furthermore, it turns out that it is possible to estimate values of F_t (and G_t) recursively. Although the resulting recursion is nonstandard and highly complex, a somewhat intricate analysis eventually produces a lower bound for the uniform process.

1.5 Organization In Section 2, we will describe Schramm's analytic approach for approximating the uniform process $\mathcal{U}$. Section 3 outlines the argu-

ments showing that the size-biased process $\mathcal{S}$ satisfies $\mathbb{E}[\tau_c(\mathcal{S})] = O(\log n)$. The fact that $\tau_c(\mathcal{S}) = O(\log n)$ w.h.p. requires additional ideas and a delicate martingale analysis and will appear in the journal version of this paper. The final section, Section 4, builds upon Schramm's aforementioned framework to produce a super-logarithmic lower bound for $\tau_c(\mathcal{U})$.

2 Oded Schramm's analytic approximation framework for the uniform process

In this section we describe Schramm's analytic approach as it was presented in [24] for analyzing the uniform coalescence process $\mathcal{U}$, as well as the numerical evidence that Schramm obtained based on this approach suggesting that $\tau_c(\mathcal{U})$ is super-logarithmic. Throughout this section we write approximations loosely as they were sketched by Schramm and postpone any arguments on their validity (including concentration of random variables, etc.) to Section 4, where we will turn elements from this approach into a rigorous lower bound on $\tau_c(\mathcal{U})$.

Let $\mathcal{F}_t$ denote the σ-algebra generated by the first t rounds of the coalescence process $\mathcal{U}$. The starting point of Schramm's approach was to examine the following function conditioned on $\mathcal{F}_t$:

$$F_t(s) = \sum_{i=1}^{\kappa_t} \exp(-w_i\, s)\,,$$

where κ_t is the number of clusters after t rounds and $w_1, \ldots, w_{\kappa_t}$ denote the normalized cluster-sizes at that time (see Definition 1). The benefit that one could gain from understanding the behavior of $F_t(s)$ is obvious as $F_t(0)$ recovers the number of clusters at time t.

More interesting is the following observation of Schramm regarding the role that $F_t(\kappa_t/2)$ plays in the evolution of the clusters. Conditioned on $\mathcal{F}_t$, the probability that the cluster $\mathcal{C}_i$ receives a merge request from another cluster $\mathcal{C}_j$ is $\frac{1}{2}w_i$ (the factor $\frac{1}{2}$ accounts for the choice of $\mathcal{C}_j$ to issue rather than accept requests). Thus, the probability that $\mathcal{C}_i$ will receive *any* incoming request in round $t+1$ and independently decide to be an acceptor is

$$\frac{1}{2}\left[1 - (1 - w_i/2)^{\kappa_t - 1}\right] \approx \frac{1}{2}\left[1 - \exp(-w_i\kappa_t/2)\right]\,.$$

On this event, $\mathcal{C}_i$ will account for one merge at time $t+1$, and summing this over all clusters yields

$$\begin{aligned}\mathbb{E}[\kappa_{t+1} \mid \mathcal{F}_t] &\approx \kappa_t - \frac{1}{2}\sum_{i=1}^{\kappa_t}\left[1 - \exp(-w_i\kappa_t/2)\right] \\ &= \frac{1}{2}\left[\kappa_t + F_t(\kappa_t/2)\right]\,,\end{aligned}$$

or equivalently, if we re-scale $F_t(s)$ into $G_t(s) = (1/\kappa_t)F_t(\kappa_t s)$ as in Eq. (1.1),

$$\mathbb{E}\left[\kappa_{t+1}/\kappa_t \mid \mathcal{F}_t\right] \approx \frac{1 + G_t(\frac{1}{2})}{2}\,. \tag{2.2}$$

In order to have $\tau_c(\mathcal{U}) \asymp \log n$ the number of clusters would need to typically drop by at least a constant factor at each round. This would require the ratio in (2.2) to be bounded away from 1, or equivalently, $G_t(\frac{1}{2})$ should be bounded away from 1.

Unfortunately, the evolution of the sequence $G_t(\frac{1}{2}) = (1/\kappa_t)F_t(\kappa_t/2)$ appears to be quite complex and there does not seem to be a simple way to determine its limiting behavior. Nevertheless, Schramm was able to write down an approximate recursion for the expected value of F_{t+1} in terms of multiple evaluations of F_t by observing the following: On the above event that $\mathcal{C}_i$ chooses to accept the merge request of some other cluster $\mathcal{C}_j$, by definition of the process $\mathcal{U}$ the identity of the cluster $\mathcal{C}_j$ is uniformly distributed over all $\kappa_t - 1$ clusters other than $\mathcal{C}_i$. Hence,

$$\begin{aligned}&\mathbb{E}\big[F_{t+1}(s) - F_t(s) \mid \mathcal{F}_t\big] \\ &\qquad\approx \sum_i \frac{1}{2}\left(1 - e^{-w_i\kappa_t/2}\right)\frac{1}{\kappa_t}\times \\ &\qquad\qquad \sum_{j\neq i}\left(e^{-(w_i+w_j)s} - e^{-w_i s} - e^{-w_j s}\right)\,.\end{aligned}$$

Ignoring the fact that the last sum in the approximation skips the diagonal terms $j = i$, one arrives at a summation over all $1 \le i, j \le \kappa_t$ of exponents similar to those in the definition of F_t with an argument of either s, $\kappa_t/2$, or $s + \kappa_t/2$, which after rearranging gives

$$\begin{aligned}&\mathbb{E}[F_{t+1}(s) \mid \mathcal{F}_t] \\ &\quad\approx \frac{1}{2}F_t(s + \kappa_t/2) \\ &\quad + \frac{1}{2\kappa_t}F_t(s)\Big[F_t(s) + F_t(\kappa_t/2) - F_t(s + \kappa_t/2)\Big]\,.\end{aligned}$$

To turn the above into an expression for $G_{t+1}(s)$ one needs to evaluate $F_{t+1}(\kappa_{t+1}s)$ rather than $F_{t+1}(\kappa_t s)$, to which end the approximation $\kappa_{t+1} \approx \frac{1}{2}[1 + G_t(\frac{1}{2})]\kappa_t$ can be used based on (2.2). Additionally, for the starting point of the recursion, note that the initial configuration of $w_i = 1/\kappa_0$ for all $1 \le i \le \kappa_0$ has $G_0(s) = \exp(-s)$. Altogether, Schramm obtained the following deterministic analytic recurrence, whose behavior should (approximately) dictate the coalescence

rate:

$$\begin{cases} g_0(s) = & \exp(-s)\,, \\ g_{t+1}(s) = & \frac{1}{2\alpha}\left[\begin{array}{l} g_t(\alpha s)^2 - g_t(\alpha s + \frac{1}{2})g_t(\alpha s) \\ +g_t(\alpha s + \frac{1}{2}) + g_t(\frac{1}{2})g_t(\alpha s) \end{array}\right], \\ & \text{where } \alpha = \frac{1}{2}[1 + g_t(\frac{1}{2})]\,. \end{cases}$$

In light of this, aside from the task of assessing how good of an approximation the above defined functions g_t provide for the random variables G_t along the uniform coalescence process $\mathcal{U}$, the other key question is whether the sequence $g_t(\frac{1}{2})$ converges to 1 as $t \to \infty$, and if so, at what rate.

For the latter, as the complicated definition of g_{t+1} attests, analyzing the recursion of g_t seems highly non-trivial. Moreover, a naive evaluation of $g_t(\frac{1}{2})$ involves exponentially many terms, making numerical simulations already challenging. The computer-assisted numerical estimates performed by Schramm for the above recursion, shown in Figure 2, seemed to suggest that indeed $g_t(\frac{1}{2}) \to 1$ (albeit very slowly), which should lead to a super-logarithmic coalescence time for $\mathcal{U}$. However, no rigorous results were known for the limit of $g_t(\frac{1}{2})$ or its stochastic counterpart $G_t(\frac{1}{2})$.

In order to turn Schramm's argument into a rigorous lower bound on $\tau_c(\mathcal{U})$, we will need to move our attention away from the sought value of $G_t(\frac{1}{2})$ and focus instead on $G_t(1)$. By manipulating Schramm's recursion for G_t and combining it with additional analytic arguments and appropriate concentration inequalities, we show that as long as κ_t is large enough and $G_t(\frac{1}{2}) < 1-\delta$ for some fixed $\delta > 0$, then typically $G_{t+1}(1) > G_t(1) + \epsilon$ for some $\epsilon(\delta) > 0$. Since by definition $0 \le G_t(1) \le 1$ this can be used to show that ultimately $G_t(\frac{1}{2}) \to 1$ w.h.p., and a careful quantitative version of this argument produces the rigorous lower bound on $\tau_c(\mathcal{U})$ stated in Theorem 1.1.

3 Expected running time of the size-biased process

The goal of this section is to prove that the expected time for the size-biased process to complete has logarithmic order, as stated in Proposition 3.1. Following a few simple observations on the process we will prove this proposition using two key lemmas, Lemmas 3.2 and 3.3, whose proofs will appear in the journal version of this paper. There, we also extend the proof of this proposition using some additional ideas to establish that the coalescence time is bounded by $O(\log n)$ w.h.p.

PROPOSITION 3.1. *Let $\tau_c = \tau_c(\mathcal{S})$ denote the coalescence time of the size-biased process $\mathcal{S}$. Then there exists an absolute constant $C > 0$ such that $\mathbb{E}_1[\tau_c] \le C \log n$, where $\mathbb{E}_1[\cdot]$ denotes expectation w.r.t. an initial cluster distribution comprised of n singletons.*

Throughout this section we refer only to the size-biased process and use the following notation. Define the filtration $\mathcal{F}_t$ to be the σ-algebra generated by the process up to and including the t-th round. Let κ_t denote the number of clusters after the conclusion of round t, noting that with these definitions we are interested in bounding the expected value of the stopping time

$$\tau_c = \min\{t : \kappa_t = 1\}\,. \tag{3.3}$$

As mentioned in the introduction, we normalize the cluster-sizes so that they sum to 1. Finally, the susceptibility χ_t denotes the sum of squares of the cluster-sizes at the end of round t.

Observe that by Cauchy-Schwarz, if $w_1, \ldots, w_{\kappa_t}$ are the cluster-sizes at the end of round t then we always have

$$\chi_t\, \kappa_t \ge \left(\sum_{i=1}^{\kappa_t} w_i\right)^2 = 1\,, \tag{3.4}$$

with equality iff all clusters have the same size. Indeed, the susceptibility χ_t measures the variance of the cluster-size distribution. When χ_t is smaller (closer to κ_t^{-1}), the distribution is more uniform. We further claim that

$$\chi_{t+1} \le 2\chi_t \quad \text{for all } t\,. \tag{3.5}$$

To see this, note that if a cluster of size a merges with a cluster of size b the susceptibility increases by exactly $(a+b)^2 - (a^2+b^2) = 2ab \le a^2+b^2$. Since each round only involves merges between disjoint pairs of clusters, this immediately implies that the total additive increase in susceptibility is bounded by the current sum of squares of the cluster sizes, i.e., the current susceptibility χ_t.

Before commencing with the proof of Proposition 3.1, we present a trivial linear bound for the expected running time of the coalescence process, which will later serve as the final step in our proof. Here and in what follows, $\mathbb{P}_w$ and $\mathbb{E}_w$ denote probability and expectation given the initial cluster distribution w. While the estimate featured here appears to be quite crude when w is uniform, recall that in general τ_c can in fact be linear in the initial number of clusters w.h.p., e.g. when w is comprised of one cluster of mass $1 - 1/\sqrt{n}$ and $\sqrt{n}$ other clusters of mass $1/n$ each.

LEMMA 3.1. *Starting from κ clusters with an arbitrary cluster distribution $w = (w_1, \ldots, w_\kappa)$ we have $\mathbb{E}_w[\tau_c] \le 8\kappa$. Furthermore, $\mathbb{P}_w(\tau_c > 16\kappa) \le e^{-\kappa/4}$.*

Proof. Consider an arbitrary round in which at least 2 clusters still remain. We claim that the probability that there is at least one merge in this round is at least $\frac{1}{8}$. Indeed, let $\mathcal{C}_1$ be a cluster of minimal size: The probability that it decides to send a request is $\frac{1}{2}$, and since there are at least two clusters and $\mathcal{C}_1$ is the smallest one, the probability that this request goes to some $\mathcal{C}_j$ with $j \neq 1$ is at least $\frac{1}{2}$. Finally, the probability that $\mathcal{C}_j$ is accepting requests is again $\frac{1}{2}$. Conditioned on these events, $\mathcal{C}_j$ will definitely accept some request (possibly not the one from $\mathcal{C}_1$ as another cluster of the same size as $\mathcal{C}_1$ may have sent it a request) leading to at least one merge, as claimed.

The process terminates when the total cumulative number of merges reaches $\kappa - 1$. Therefore, the time of completion is stochastically dominated by the sum of $\kappa - 1$ geometric random variables with success probability $\frac{1}{8}$, and in particular $\mathbb{E}_w[\tau_c] \leq 8(\kappa - 1)$.

By the same reasoning, the total number of merges that occurred in the first t rounds clearly stochastically dominates a binomial variable $\mathrm{Bin}(t, \frac{1}{8})$ as long as $t \leq \tau_c$. Therefore,

$$\mathbb{P}_w(\tau_c > 16\kappa) \leq \mathbb{P}\left(\mathrm{Bin}(16\kappa, \tfrac{1}{8}) \leq \kappa - 1\right) \leq e^{-\kappa/4},$$

where the last inequality used the well-known Chernoff bounds (see e.g. [17, Theorem 2.1]). □

3.1 Proof of Proposition 3.1 via two key lemmas We next present the two main lemmas on which the proof of the proposition hinges. The key idea is to design a potential function comprised of two parts Φ_1, Φ_2 while identifying a certain event A_t such that the following holds: $\mathbb{E}[\Phi_1(t+1) - \Phi_1(t) \mid \mathcal{F}_t, A_t] < c_1 < 0$ and $\mathbb{E}[\Phi_2(t+1) - \Phi_2(t) \mid \mathcal{F}_t] < c_2$ where c_1, c_2 are absolute constants, and a similar statement holds conditioned on A_t^c when reversing the roles of Φ_1 and Φ_2. At this point we will establish that an appropriate linear combination of Φ_1, Φ_2 is a supermartingale, and the required bound on τ_c will follow from Optional Stopping. Note that throughout the proof we make no attempt to optimize the absolute constants involved. The event A_t of interest is defined as follows:

Definition 2. *Let A_t be the event that the following two properties hold after the t-th round:*

(i) *At least $\kappa_t/2$ clusters have size at most $1/(600\kappa_t)$.*

(ii) *The cluster-size distribution satisfies the inequality $\sum_i w_i \mathbb{1}_{\{w_i < 41/\kappa_t\}} < 4 \cdot 10^{-5}$.*

The intuition behind this definition is that Property (2) boosts the number of tiny clusters, thereby severely retarding the growth of the largest clusters, which will tend to see incoming requests from these tiny clusters. Property (2) ensures that most of the mass of the cluster-size distribution is on relatively large clusters, of size at least 41 times the average.

Examining the event A_t will aid in tracking the variable $\chi_t \kappa_t$, the normalized susceptibility (recall from (3.4) that this quantity is always at least 1 and it equals 1 whenever all clusters are of the same size). The next lemma, whose proof appears in the journal version of this paper, estimates the expected change in this quantity and most notably shows that it is at most $-\frac{1}{200}$ if we condition on A_t.

Lemma 3.2. *Let $\Phi_1(t) = \chi_t \kappa_t$ and suppose that at the end of the t-th round one has $\kappa_t \geq 2$. Then*

$$\mathbb{E}[\Phi_1(t+1) - \Phi_1(t) \mid \mathcal{F}_t] \leq 5 \tag{3.6}$$

and furthermore

$$\mathbb{E}\left[\Phi_1(t+1) - \Phi_1(t) \mid \mathcal{F}_t, A_t, \chi_t < 3 \cdot 10^{-7}\right] \leq -\tfrac{1}{200}. \tag{3.7}$$

Fortunately, when A_t does not hold the behavior in the next round can still be advantageous in the sense that in this case the number of clusters tends to fall by at least a constant fraction. This is established by the following lemma, whose proof is postponed to the journal version of this paper.

Lemma 3.3. *Let $\Phi_2(t) = \log \kappa_t$ and suppose that after the t-th round one has $\kappa_t \geq 2$. Then*

$$\mathbb{E}\left[\Phi_2(t+1) - \Phi_2(t) \mid \mathcal{F}_t, A_t^c\right] < -2 \cdot 10^{-7}. \tag{3.8}$$

We are now in a position to derive Proposition 3.1 from the above two lemmas.

Proof of Proposition 3.1. Define the stopping time τ to be

$$\tau = \min\left\{i : \chi_t \geq 3 \cdot 10^{-7}\right\}.$$

Observe that the susceptibility is initially $1/n$, its value is 1 once the process arrives at a single cluster (i.e. at time τ_c) and until that point it is nondecreasing, hence $\mathbb{E}\tau \leq \mathbb{E}\tau_c < \infty$ by Lemma 3.1. Further define the random variable

$$Z_t = \chi_t \kappa_t + 3 \cdot 10^7 \log \kappa_t + \frac{t}{200}.$$

We claim that $(Z_{t \wedge \tau})$ is a supermartingale. Indeed, consider $\mathbb{E}[Z_{t+1} \mid \mathcal{F}_t, \tau > t]$ and note that the fact that $\tau > t$ implies in particular that $\kappa_t \geq 2$ since in that case $\chi_t < 3 \cdot 10^{-7} < 1$.

- If A_t holds then by (3.7) the conditional expected change in $\chi_t \kappa_t$ is below $-\frac{1}{200}$, while $\log \kappa_t$ can only decrease (as κ_t is non-increasing), hence $\mathbb{E}[Z_{t+1} \mid \mathcal{F}_t, A_t, \tau > t] \leq Z_t$.

- If A_t does not hold then by (3.6) the conditional expected change in $\chi_t\kappa_t$ is at most $+5$ whereas the conditional expected change in $\log\kappa_t$ is below $-2\cdot 10^{-7}$ due to (3.8). By the scaling in the definition of Z_t these add up to give $\mathbb{E}[Z_{t+1} \mid \mathcal{F}_t, A_t^c, \tau > t] \le Z_t - \frac{199}{200}$.

Altogether, $(Z_{t\wedge\tau})$ is indeed a supermartingale. As its increments are bounded and the stopping time τ is integrable we can apply the Optional Stopping Theorem (see e.g. [8, Chapter 5]) and get

$$\mathbb{E}Z_\tau \le Z_0 = \chi_0\kappa_0 + 3\cdot 10^7 \log\kappa_0 = O(\log n)\,. \tag{3.9}$$

At the same time, by definition of τ we have $\chi_\tau \ge 3\cdot 10^{-7}$ and so

$$Z_\tau = \chi_\tau\kappa_\tau + 3\cdot 10^7\log\kappa_\tau + \frac{\tau}{200} \ge 3\cdot 10^{-7}\left(\kappa_\tau + \tau/8\right)\,. \tag{3.10}$$

Taking expectation in (3.10) and combining it with (3.9) we find that

$$\mathbb{E}\left[\tau + 8\kappa_\tau\right] \le O(\log n)\,.$$

Finally, conditioned on the cluster distribution at time τ we know by Lemma 3.1 that the expected number of additional rounds it takes the process to conclude is at most $8\kappa_\tau$, thus $\mathbb{E}[\tau_c] \le \mathbb{E}[\tau + 8\kappa_\tau]$. We can now conclude that $\mathbb{E}[\tau_c] = O(\log n)$, as required. □

4 Super-logarithmic lower bound for the uniform process

In this section we use the analytic approximation framework introduced by Schramm to prove the super-logarithmic lower bound stated in Theorem 1.1 for the coalescence time of the uniform process. Recall that a key element in this framework is the normalized Laplace transform of the cluster-size distribution, namely $G_t(s) = (1/\kappa_t)F_t(\kappa_t s)$ where $F_t(s) = \sum_{i=1}^{\kappa_t} e^{-w_i s}$ (see Definition 1). The following proposition, whose proof entails most of the technical difficulties in our analysis of the uniform process, demonstrates the effect of $G_t(\frac{1}{2})$ and $G_t(1)$ on the coalescence rate.

PROPOSITION 4.1. *Let $\epsilon_t = 1 - G_t(\frac{1}{2})$ and $\zeta_t = G_t(1)$. There exists an absolute constant $C > 0$ such that, conditioned on $\mathcal{F}_t$, with probability at least $1 - C\kappa_t^{-100}$ we have*

$$\left|\kappa_{t+1} - (1-\epsilon_t/2)\kappa_t\right| \le \kappa_t^{2/3}\,, \tag{4.11}$$

$$\zeta_{t+1} \ge \zeta_t + \epsilon_t^{13/\epsilon_t} - 8\kappa_t^{-1/3}\,. \tag{4.12}$$

We postpone the proof of this proposition to the journal version in favor of showing how the relations that it establishes between $\kappa_t, G_t(1), G_t(\frac{1}{2})$ can be used to derive the desired lower bound on τ_c. We claim that as long as $\kappa_t, G_t(\frac{1}{2}), G_t(1)$ satisfy Eq. (4.11),(4.12) and $t = O\big(\log n \cdot \frac{\log\log\log n}{\log\log n}\big)$ then $\kappa_t \ge n^{3/4}$; this deterministic statement is given by the following lemma:

LEMMA 4.1. *Set $T = \frac{1}{75}\log n \cdot \frac{\log\log n}{\log\log\log n}$ for a sufficiently large n and let $\kappa_0, \ldots, \kappa_T$ be a sequence of integers in $\{1,\ldots,n\}$ with $\kappa_0 = n$. Further let ϵ_t and ζ_t for $t = 0,\ldots,T$ be two sequences of reals in $[0,1]$ and suppose that for all $t < T$ the three sequences satisfy inequalities (4.11) and (4.12). Then $\kappa_t > n^{3/4}$ for all $t \le T$.*

Observe that the desired lower bound on the coalescence time of the uniform process $\mathcal{U}$ is an immediate corollary of Proposition 4.1 and Lemma 4.1. Indeed, condition on the first t rounds where $0 \le t < T = \frac{1}{75}\log n \cdot \frac{\log\log n}{\log\log\log n}$ and assume $\kappa_t > n^{3/4}$. Proposition 4.1 implies that Eqs. (4.11), (4.12) hold except with probability $O(\kappa_t^{-100}) = o(n^{-1})$. On this event Lemma 4.1 yields $\kappa_{t+1} > n^{3/4}$, extending our assumption to the next round. Accumulating these probabilities for all $t < T$ now shows that $\mathbb{P}(\kappa_T > n^{3/4}) = 1 - o(T/n)$ and in particular $\tau_c > T$ w.h.p., as required.

Proof of lemma. The proof proceeds by induction: Assuming that $\kappa_i > n^{3/4}$ for all $i \le t < T$ we wish to deduce that $\kappa_{t+1} > n^{3/4}$.

Repeatedly applying Eq. (4.12) and using the induction hypothesis we find that

$$\zeta_{t+1} \ge \zeta_0 + \sum_{i=0}^{t}\left(\epsilon_i^{13/\epsilon_i} - 8\kappa_i^{-1/3}\right) \tag{4.13}$$

$$> \sum_{i=0}^{t}\left(\epsilon_i^{13/\epsilon_i}\right) - 8(t+1)\left(n^{3/4}\right)^{-1/3} \tag{4.14}$$

$$= \sum_{i=0}^{t}\left(\epsilon_i^{13/\epsilon_i}\right) - n^{-\frac{1}{4}+o(1)} \tag{4.15}$$

since $t \le T = n^{o(1)}$. Following this, we claim that the set $I = \left\{0 \le i \le t : \epsilon_i \ge 15\frac{\log\log\log n}{\log\log n}\right\}$ has size at most $(\log n)^{\frac{9}{10}}$. Indeed, as $x^{1/x}$ is monotone increasing for all $x \le e$, every such $i \in I$ has

$$\begin{aligned}\epsilon_i^{13/\epsilon_i} &\ge \left(15\frac{\log\log\log n}{\log\log n}\right)^{\frac{13}{15}\frac{\log\log n}{\log\log\log n}}\\ &= (\log n)^{-\frac{13}{15}+o(1)}\\ &> (\log n)^{-\frac{9}{10}}\,,\end{aligned}$$

where the last inequality holds for large n. Hence, if we had $|I| > 2(\log n)^{9/10}$ then it would follow from (4.14)

that $\zeta_{t+1} > 2 - o(1)$, contradicting the assumption of the lemma for large enough n.

Moreover, by the assumption that $\epsilon_i \in [0,1]$ we have $\frac{1}{2} \leq (1-\epsilon_i/2) \leq 1$ for all i. Together with the facts that $\kappa_{i+1} \geq (1-\epsilon_i/2)\kappa_i - \kappa_i^{2/3}$ for all $i \leq t$ due to (4.11) while $\kappa_i \leq n$ for all i we now get

$$\begin{aligned}\kappa_{t+1} &\geq \kappa_0 \prod_{i=0}^{t}(1-\epsilon_i/2) - \sum_{i=0}^{t}\kappa_i^{2/3} \\ &\geq \left(1 - \frac{1}{2}\cdot\frac{15\log\log\log n}{\log\log n}\right)^t 2^{-|I|}\, n - (t+1)n^{2/3} \\ &\geq e^{-15\frac{\log\log\log n}{\log\log n}T} 2^{-|I|}\, n - Tn^{2/3}\,,\end{aligned}$$

where the last inequality used the fact that $t < T$ as well as the inequality $1 - x/2 > e^{-x}$, valid for all $0 < x < 1$. Now, $2^{-|I|} = n^{-o(1)}$ since $|I| \leq 2(\log n)^{9/10}$ and by the definition of T we obtain that

$$\kappa_{t+1} \geq e^{-\frac{1}{5}\log n} n^{1-o(1)} - n^{\frac{2}{3}+o(1)} = n^{\frac{4}{5}-o(1)} > n^{3/4}$$

for sufficiently large n, as claimed. This concludes the proof. □

Acknowledgments

We thank Yuval Peres and Dahlia Malkhi for suggesting the problem and for useful discussions. The starting point of our work is attributed to the analytic approximation framework of Oded Schramm, and E. L. is indebted to Oded for enlightening and fruitful discussions on his approach.

This work was initiated while P.-S. L. was an intern at the Theory Group of Microsoft Research, and he thanks the Theory Group for its hospitality.

References

[1] D. Aldous, *Asymptotic fringe distributions for general families of random trees*, Ann. Appl. Probab. **1** (1991), no. 2, 228–266.

[2] D. Aldous, *Deterministic and stochastic models for coalescence (aggregation, coagulation): a review of the mean-field theory for probabilists*, Bernoulli **5** (1997), 3–48.

[3] N. Alon and J. H. Spencer, *The probabilistic method*, 3rd ed., John Wiley & Sons Inc., Hoboken, NJ, 2008.

[4] B. Bollobás, *Random graphs*, 2nd ed., Cambridge University Press, 2001.

[5] B. Bollobás and I. Simon, *Probabilistic analysis of disjoint set union algorithms*, SIAM J. Comput. **22** (1993), no. 5, 1053–1074.

[6] J. Cooper and N. Linial, *Fast perfect-information leader-election protocols with linear immunity*, Combinatorica **15** (1995), 319–332. Early version appeared in Proc. of the the 25th Annual ACM Symposium on Theory of Computing (STOC 1993), 662–671.

[7] R. Durrett, *Probability models for DNA sequence evolution*, 2nd ed., Probability and its Applications (New York), Springer, New York, 2008.

[8] R. Durrett, *Probability: theory and examples*, 3rd ed., Duxbury Press, 2004.

[9] P. Erdős and A. Rényi, *On the evolution of random graphs*, Publ. Math. Inst. Hung. Acad. Sci. **5** (1960), 17–61.

[10] Y. Fernandess and D. Malkhi, *On spreading recommendations via social gossip*, Proc. of the 20th annual symposium on Parallelism in algorithms and architectures (SPAA 2008), 91–97.

[11] Y. Fernandess and D. Malkhi, *private communication* (2007).

[12] J. A. Fill, H. M. Mahmoud, and W. Szpankowski, *On the distribution for the duration of a randomized leader election algorithm*, Ann. Appl. Probab. **6** (1996), no. 4, 1260–1283.

[13] D. A. Freedman, *On tail probabilities for martingales*, Ann. Probability **3** (1975), 100–118.

[14] Z. Galil and G. F. Italiano, *Data structures and algorithms for disjoint set union problems*, ACM Comput. Surv. **23** (1991), 319–344.

[15] H. Gazit, *An optimal randomized parallel algorithm for finding connected components in a graph*, Proc. of the 27th Annual Symposium on Foundations of Computer Science (FOCS 1986), 492–501.

[16] G. Grimmett, *Percolation*, 2nd ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 321, Springer-Verlag, Berlin, 1999.

[17] S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.

[18] J. F. C. Kingman, *The coalescent*, Stochastic Process. Appl. **13** (1982), no. 3, 235–248.

[19] D. E. Knuth and A. Schönhage, *The expected linearity of a simple equivalence algorithm*, Theoret. Comput. Sci. **6** (1978), no. 3, 281–315.

[20] G. Miller and J. H. Reif, *Parallel tree contraction and its applications*, Proc. of the 26th Annual Symposium on Foundations of Computer Science (FOCS 1985), 478–489.

[21] M. Molloy and B. Reed, *Graph colouring and the probabilistic method*, Springer-Verlag, Berlin, 2002.

[22] R. Ostrovsky, S. Rajagopalan, and U. Vazirani, *Simple and Efficient Leader Election in the Full Information Model*, Proc. of the 26th Annual ACM Symposium on Theory of Computing (STOC 1994), 234–242.

[23] A. Russell and D. Zuckerman, *Perfect information leader election in* $\log^* n + O(1)$ *rounds*, J. Comput. System Sci. **63** (2001), 612–626. Special issue on FOCS 1998 (Palo Alto, CA).

[24] O. Schramm, *An analytic approach to Malkhi's stochastic coalescence process*, Microsoft Research Theory Seminar (Redmond, WA, August 2008). Also: *private communication* (2007).

[25] J. H. Spencer and N. Wormald, *Birth control for giants*, Combinatorica **27** (2007), 587–628.

[26] A. C.-C. Yao, *On the average behavior of set merging algorithms (extended abstract)*, Proc. of the 8th annual ACM symposium on Theory of Computing (STOC 1976), 192–195.

[27] D. Zuckerman, *Randomness-optimal sampling, extractors, and constructive leader election*, Proc. of the 28th Annual ACM Symposium on Theory of Computing (STOC 1996), 286–295.