SIMULATIONS OF TWO COUPLED UNIFORM SPLIT-MERGE PROCESSES

PETER WINDRIDGE

Abstract. Uniform split-merge is a Markov chain on the space of (countable) partitions of the unit interval. The update step begins by size-biased sampling two parts (with repetition). If the chosen parts are distinct then they are merged with probability $\beta_{\mathrm{m}} \in (0, 1]$. If the same part is chosen twice then it is split at a uniformly chosen point with probability $\beta_{\mathrm{s}} \in (0, 1]$.

The Poisson-Dirichlet($\beta_{\mathrm{s}}/\beta_{\mathrm{m}}$) distribution is known to be invariant for this process (Tsilevich [Tsi00], Pitman [Pit02] and Mayer-Wolf et al. [MWZZ02]). Unicity is known for $\beta_{\mathrm{s}} = \beta_{\mathrm{m}}$ (Diaconis et al. [DMWZZ04] and Schramm [Sch05]).

Schramm's proof involves coupling two copies of the process such that they gradually match up (in a suitable sense). The argument readily generalises to the case $\beta_{\mathrm{s}} < \beta_{\mathrm{m}}$. The generisalation to the case $\beta_{\mathrm{m}} < \beta_{\mathrm{s}}$ is not obvious, but that regime is relevant for quantum spin models (at least on a heuristic level [GUW11]).

In these notes we look at simulations of Schramm's coupling when $\beta_{\mathrm{m}} < \beta_{\mathrm{s}}$. We find that the two copies still gradually match up, as in the case $\beta_{\mathrm{m}} \geqslant \beta_{\mathrm{s}}$. (We do not prove anything, however.)

1. Introduction

Uniform split-merge is a discrete time Markov process on countable partitions of the unit interval,

$$\mathcal{P} = \left\{ y_1 \geqslant y_2 \geqslant \ldots \geqslant 0, \ \sum_{i=1}^{\infty} y_i = 1 \right\} \subset \ell^1. \tag{1.1}$$

The update step begins by size-biased sampling two parts, independently (with repetition). If distinct parts are chosen then they are merged. If the same part is chosen twice then it is split in two. We then order the parts according to size to obtain the updated partition.

The Poisson-Dirichlet distribution with parameter $\theta = 1$ (as defined e.g. in [Hol01]) is invariant for this process, as shown by Tsilevich [Tsi00], Pitman [Pit02] and Mayer-Wolf, Zeitouni and Zerner [MWZZ02].

Tsilevich [Tsi00] proves uniqueness amongst all measures satisfying a certain symmetry condition, Pitman [Pit02] amongst measures satisfying a size-biased permutation invariance condition, and Mayer-Wolf et al. [MWZZ02] amongst measures satisfying an analyticity condition. Diaconis et al. [DMWZZ04] prove uniqueness with no restriction, by considering a discrete analogue of the process and using representation theory.

Schramm [Sch05] gives a different uniqueness proof. The idea is to couple two copies of the split-merge process, one starting off with Poisson-Dirichlet(1) distribution and the other with an arbitrary invariant measure. The coupling is constructed so that both copies of the process gradually match up, in an appropriate sense. This shows that the two initial invariant distributions are actually equal.

2010 *Mathematics Subject Classification.* ...

Key words and phrases. uniform split-merge, coagulation-fragmentation, simulations, Poisson-Dirichlet, invariant measure, coupling.

We can extend Schramm's coupling to the generalised uniform split-merge process studied in [MWZZ02]. There, a proposed split, or merge, is only accepted with probability $\beta_{\mathrm{s}} \in (0,1]$, or $\beta_m \in (0,1]$, respectively. (The partition is left unchanged in the case of a rejection.) It is known that Poisson-Dirichlet(β_s/β_m) is invariant for this generalised process [MWZZ02; GUW11]. We would like to know it is the only one, as this question arises in quantum spin models [GUW11] and also seems interesting in its own right.

In Schramm's coupling it turns out that big parts are easier to handle. So, merges are good. In fact, if $\beta_{\mathrm{m}} \geqslant \beta_{\mathrm{s}}$ then it is straightforwards to modify Schramm's arguments.

quenessthm **Theorem 1.1** (Schramm [Sch05])**.** *Poisson-Dirichlet(β_s) is the unique invariant measure for uniform split-merge provided $\beta_{\mathrm{m}} \geqslant \beta_{\mathrm{s}}$.*

The rest of the notes are laid out as follows. We explain the coupling in Section 2. Then, in Section 3, we discuss simulations for the case $\beta_{\mathrm{m}} \geqslant \beta_{\mathrm{s}}$ (to give some feel for what happens, and what may be proven). We finally consider the physically relevant case $\beta_{\mathrm{m}} = 1/2$ and $\beta_{\mathrm{s}} = 1$ in Section 4.

To close this section, note that it suffices to assume either $\beta_{\mathrm{m}} = 1$ or $\beta_{\mathrm{s}} = 1$, since we can just add a lazy step to the chain otherwise. Thus we may accept the dichotomy $\beta_{\mathrm{m}} = 1$ (and $\beta_{\mathrm{s}} \leqslant 1$) or $\beta_{\mathrm{m}} < 1$ (and $\beta_{\mathrm{s}} = 1$). This matter recieves no further comment below.

Acknowledgements. Daniel Ueltschi's suggestions and encouragement in a related project are gratefully acknowledged.

2. Definition of the coupling

s:coupling

In this section we define the coupling; that is, a construction of two copies, denoted $(Y^t;\ t = 0,1,2,\ldots)$ and $(Z^t;\ t = 0,1,2,\ldots)$, of the uniform split-merge process on the same probalitity space. We retain the notation of Schramm, who writes that:

> "The basic idea [of the coupling] is that if we have entries in Y^t that are equal to entries in Z^t, then we don't want to ruin this. Consequently, if we make a change to such an entry in Y^t, we want to make [the same change] to the corresponding entry in Z^t. On the other hand, as much as we can, we do want to produce new entries in Y^t and Z^t that match."

To define the coupling we need some notation. Suppose $(Y,Z) \in \mathcal{P}^2$. We say the i:th part Y_i of Y is *matched* in Z if there exists $j \in \mathbb{N}$ such that $Y_i = Z_j$. There is at most one such j in our setting, since all the Y_i, and all the Z_i are distinct almost surely for us. The set of indices of matched parts of Y is denoted by $I(Y,Z)$. Likewise, $I(Z,Y)$ is the set of indices of parts of Z that are matched in Y.

The total mass of matched parts is denoted by

$$Q(Y,Z) := \sum_{i \in I(Y,Z)} Y_i.$$

Each matched part in Y corresponds to exactly one part in Z so it follows that $Q(Y,Z) = Q(Z,Y)$. We denote the common value by Q.

Given (Y,Z), we perform an update step to generate new partitions Y' and Z' as follows, and as we have attempted to depict in Figure 1. The first thing to do is make a size-biased sample from both Y and Z. We want a matched part in Y to be chosen iff a matched part in Z is chosen. To achieve this, we imagine shuffling the parts in Y and Z around so that all the matched parts are aligned. In more detail, define $\tilde{Y}$ by moving unmatched parts of Y so that they partition $[0, 1-Q]$ and matched parts so they

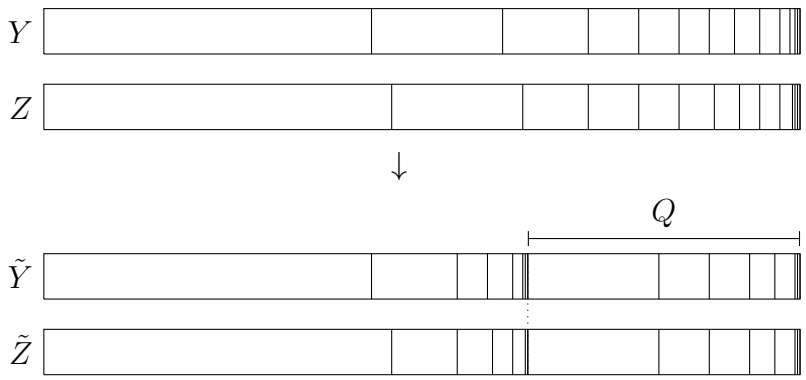


FIGURE 1. $\tilde{Y}$ and $\tilde{Z}$ are obtained by shuffling the parts in Y and Z so that the matched parts are aligned on the right. Q is the mass of matched parts. f:coupling

partition $[1-Q, 1]$. Within $[0, 1-Q]$ and $[1-Q, 1]$ we arrange the parts in size order. Construct $\tilde{Z}$ in the same way.

Now, let U be a uniform random variable on $[0, 1]$, so $U \in \tilde{Y}_i$ and $U \in \tilde{Z}_j$ for a unique choice of $i, j \in \mathbb{N}$. In this way, U defines a size-biased sample from Y and from Z. We refer to the chosen parts as Y_i and Z_j, ignoring labelling nuances. This method of sampling ensures that matched parts are indeed chosen together.

Next we need a second size-biased sample from Y and from Z, made independently of the first. Again, we want to ensure that if a matched part is chosen in Y, then its matching part in Z is also chosen. The method used to do this is similar to that above. Let $\hat{Y}$ be obtained from $\tilde{Y}$ by moving the chosen part Y_i to the beginning of $[0, 1]$, but leaving the other parts in the same order. Similarly, let $\hat{Z}$ be obtained from $\tilde{Z}$ by shifting Z_j to the beginning of $[0, 1]$.

Let V be uniform on $[0, 1]$ and independent of U. Then $V \in \hat{Y}_{i'}$ and $V \in \hat{Z}_{j'}$ for some $i', j' \in \mathbb{N}$. This gives the second sized biased samples $Y_{i'}$ and $Z_{j'}$ (again we abuse notation in keeping the same labels). As before, matched parts are chosen together. If the same part is chosen twice (i.e. $i = i'$) we propose splitting Y_i uniformly. By our choice of $\hat{Y}$, this happens iff $V \in [0, Y_i]$. But, conditional on this, V is uniform in $[0, Y_i]$ and so we can split Y_i into a part of length V and a part of length $Y_i - V$. Similarly, the same part is chosen in Z (i.e. $j = j'$) iff $V \in [0, Z_j]$, in which case we propose splitting Z_j into parts of length V and $Z_j - V$. If splits are proposed in both and Y and Z, we accept them both (together) with probability β_s, or reject them both. (That is, we never accept one split and reject the other.) Thus, if splits are proposed in both Y and Z, then either a new matching part of length V is generated or nothing happens.

If the two parts of Y chosen are distinct (i.e. $i \neq i'$) then we merge Y_i and Y_j with probability β_{m}, and similarly for Z. If merges are proposed in both Y and Z, they are both accepted or both rejected. If a merge is proposed in Y and a split proposed in Z, or vice versa, the proposals are accepted independently. (There is some freedom with these choices, and possibly a variation of these rules is better for actually proving results).

The new configuration (Y', Z') is formed by resorting the modified Y and Z into size order. This concludes the definition of the coupling.

It is worth reiterating a couple of key features. If a split is proposed (and accepted) in both Y and Z then a new matched part is created. On the other hand, the mass of unmatched parts increases is if a matched part merges with an unmatched part. In the remainder of these notes we'll see (via simulations) how these two effects trade off.

s:bm1

3. Discussion and simulations for $\beta_{\rm m} \geqslant \beta_{\rm s}$

In this section we consider $\beta_{\rm m} = 1$ (and any $\beta_{\rm s} \leqslant 1$). Let $(Y^t, Z^t;\ t = 0, 1, 2 \ldots)$ be a realisation of the coupling. For the theoretical discussions we take Y^0 distributed as Poisson-Dirichlet($\beta_{\rm s}/\beta_{\rm m} = \beta_{\rm s}$) and Z^0 an arbitrary invariant distribution. For convenience we may assume Z^0 is independent of Y^0 (to make some of the facts used below more obvious), but this is not necessary.

Initial conditions for the simulations are obtained by performing 5000 updates of the split-merge process for Y and Z independently, starting from $(1, 0, 0, \ldots)$. (We could do more updates than this, but it seems unnecessary).

250,000 update steps were performed for the coupling. However we only plot results for the first 50,000 updates, as the coupled processes seem fairly relaxed by then.

Ultimately we want control on the size of largest unmatched part, $y_1^t = \max\{Y_i^t : i \notin I(Y^t, Z^t)\}$, of Y^t and that in Z^t, denoted z_1^t. (Other parts are either smaller than $\max\{y_1^t, z_1^t\}$ or matched.) There is a crucial, but perhaps subtle point that makes things more interesting. The point is that $\max\{y_1^t, z_1^t\}$ does not tend to zero (in probability, or other usual notion) as $t \to \infty$. This is because bad unmatched mass can always 'flare up' over long enough periods of time, via matched mass merging with unmatched mass (provided there is with our initial conditions).

The key to combating this is to examine (Y^t, Z^t) at a random time T chosen uniformly and independently of everything in the coupling, so that Y^T and Z^T remain distributed as Y^0 and Z^0. (Schramm denotes the random time by 'q' but 'T' seems more descriptive. This is the only point we deem it worth diverging from Schramm's notation).

Take any $\varepsilon > 0$. Note there are infinitely many unmatched parts smaller than ε just in Y^0 (by elementary properties of the Poisson-Dirichlet distribution), and we'd like to ignore those. (A benefit of working with unmatched parts bigger than ε is we can lower bound the probability that they are involved in splits or merges.)

Schramm's proof also involves the *number* of unmatched parts bigger than ε as well as the mass. The number of parts of a partition $Y \in \mathcal{P}$ that are not matched by parts in $Z \in \mathcal{P}$ and have size bigger than ε is denoted

$$N_\varepsilon(Y, Z) := |\{i \notin I(Y, Z) : Y_i > \varepsilon\}|\,.$$

Thus the total number of non-negligible unmatched parts (in either process) at time t is,

$$N_\varepsilon^t := N_\varepsilon(Y^t, Z^t) + N_\varepsilon(Z^t, Y^t),$$

and we suppress the 'ε' subscript. As mentioned above, we want to ignore splits and merges of parts with size less than ε. The mass of such parts in a partition $Y \in \mathcal{P}$ is denoted by

$$\Upsilon(\varepsilon, Y^0) := \sum_{i: Y_i^0 < \varepsilon} Y_i^0.$$

Part of Schramm's proof involves showing $\Upsilon(\varepsilon, Y^0)$ is small, so small parts can be thrown in with ε and it is more relevant to work with

$$\bar{\varepsilon} := \varepsilon + \Upsilon(\varepsilon, Y^0) + \Upsilon(\varepsilon, Z^0)$$

to control bad events. Note that this quantity is also random.

It is surprisingly easy to connect $N^t, \max\{y_1^t, z_1^t\}, Q^t$ and $\bar{\varepsilon}$, at least when $\beta_{\rm m} = 1$. Roughly, N^t decreases by 2 only when two unmatched parts in both Y^t and Z^t merge. The probability for this to happen is bounded below by $\beta_{\rm m}(1-Q^t)(1-Q^t-\max\{y_1^t, z_1^t\})$,

FIGURE 2. The size of the largest unmatched component in Y.

i.e. if the latter is large, there is a good chance that N^t will decrement. Also, N^t is non-increasing (even though $N_\varepsilon(Y^t, Z^t)$ can increase). This can be seen by examining the different combinations of proposed split or merge in the different processes, remembering that any proposed merge is accepted when $\beta_{\mathrm{m}} = 1$. These two simple points are the key idea for proving the following Lemma.

Lemma 3.1. *Suppose $\beta_s \in (0, 1]$, $\beta_m = 1$ and T is a random variable on $\mathbb{N}$ which is independent of (Y^t, Z^t). Let*

$$\eta = \max_{t \in \mathbb{N}} \mathbb{P}(T = t).$$

Then (conditional on the initial conditions),

$$\mathbb{E}\left[(1 - Q^T)(1 - Q^T - \max\{y_1^T, z_1^T\})\right] \leqslant \tfrac{1}{2}\eta N^0 + 4\bar{\varepsilon}\,\mathbb{E}[T + 1]. \tag{3.1}$$

e:largestu

This is [Sch05, Lemma 3.1], but generalised to allow splits. The term on the right-hand side can be bounded in expectation for any β_{m} and β_{s}, by modifying the other arguments of Schramm. (The second term is due to ignoring parts of size less than ε). Schramm's argument to get from Lemma 3.1 to the conclusion that $\max\{y_1^T, z_1^T\}$ is small can also be modified easily.

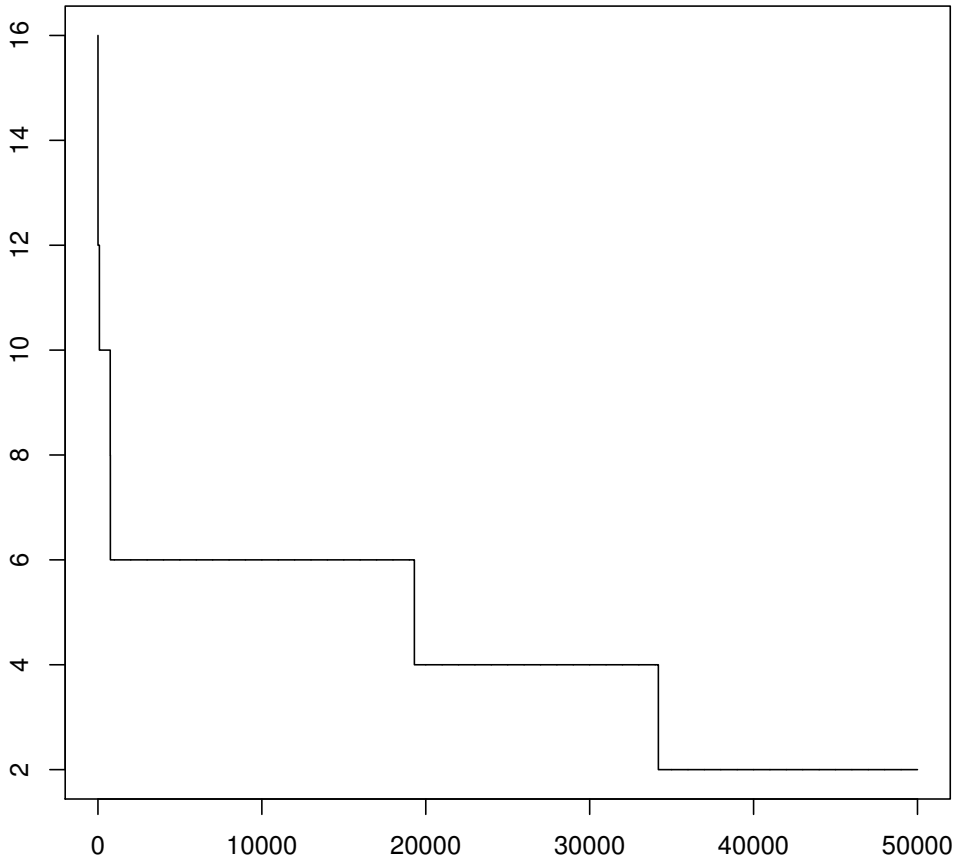


FIGURE 3. The number of unmatched components larger than ε small. (Each jump is by -2 a.s.).

The main problem then, is to prove something like Lemma 3.1 for the case $\beta_{\rm m} < 1$. The proof breaks down because N^t increases if a split of unmatched parts is accepted in Y but a merge is rejected in Z (or vice versa). Thus N^t can fluctuate, and it is more difficult to control the number of decrements with N^0.

4. Simulations for $\beta_{\rm m} = 1/2, \beta_{\rm s} = 1$

s:bm1/2

In this section we simulate the physically relevant case $\beta_{\rm s}/\beta_{\rm m} = 2$, as an example of what happens when $\beta_{\rm m} < \beta_{\rm s}$. Again, 250,000 updates were performed, with Y^0 and Z^0 independent.

We see again, it does look like $\max\{y_1^T, z_1^T\}$ is small in probability, see Figure **??**. Note that it looks like there are longer spells between 'flares' of unmatched bad mass. Presumably this is due to the fact that unmatched mass is generated by merges between matched and unmatched parts. So, if those merges are rejected, it actually takes longer to generate unmatched mass.

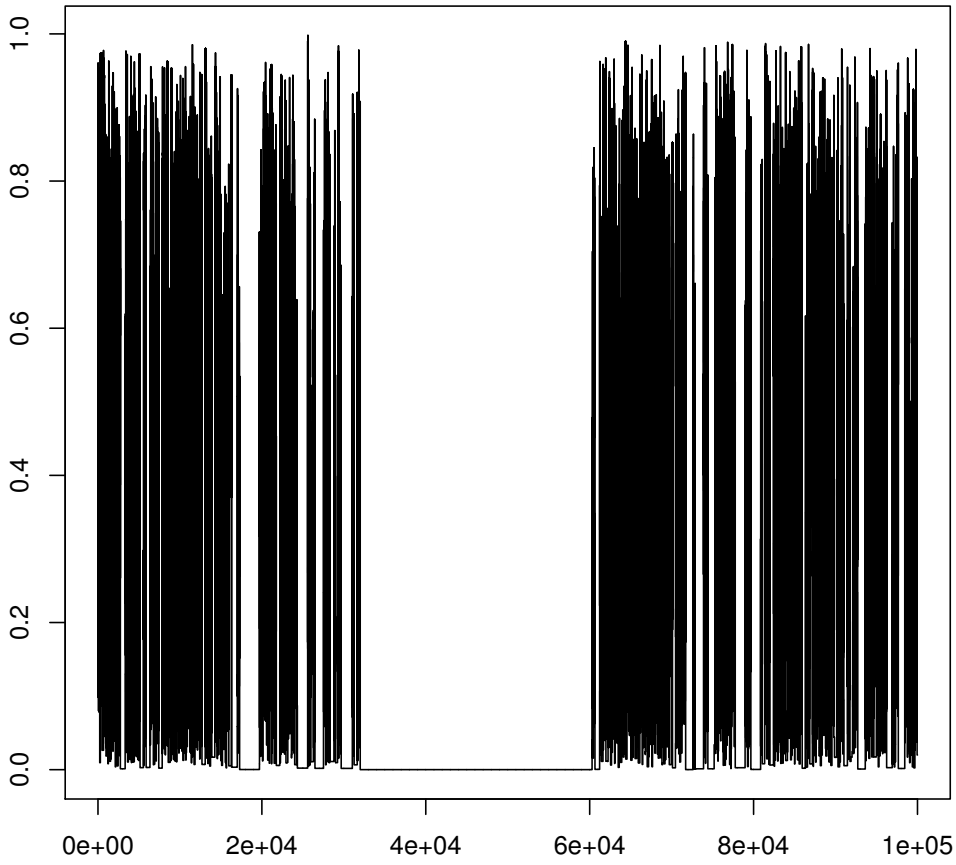


FIGURE 4. Size of the largest unmatched component in Y^t (first 100,000 updates).

As remarked above, if we want to follow the general flow of Schramm's proof, we need Lemma 3.1 to hold. Although N^t can increase, it seems that the increases are few and far between, see Figure 4.

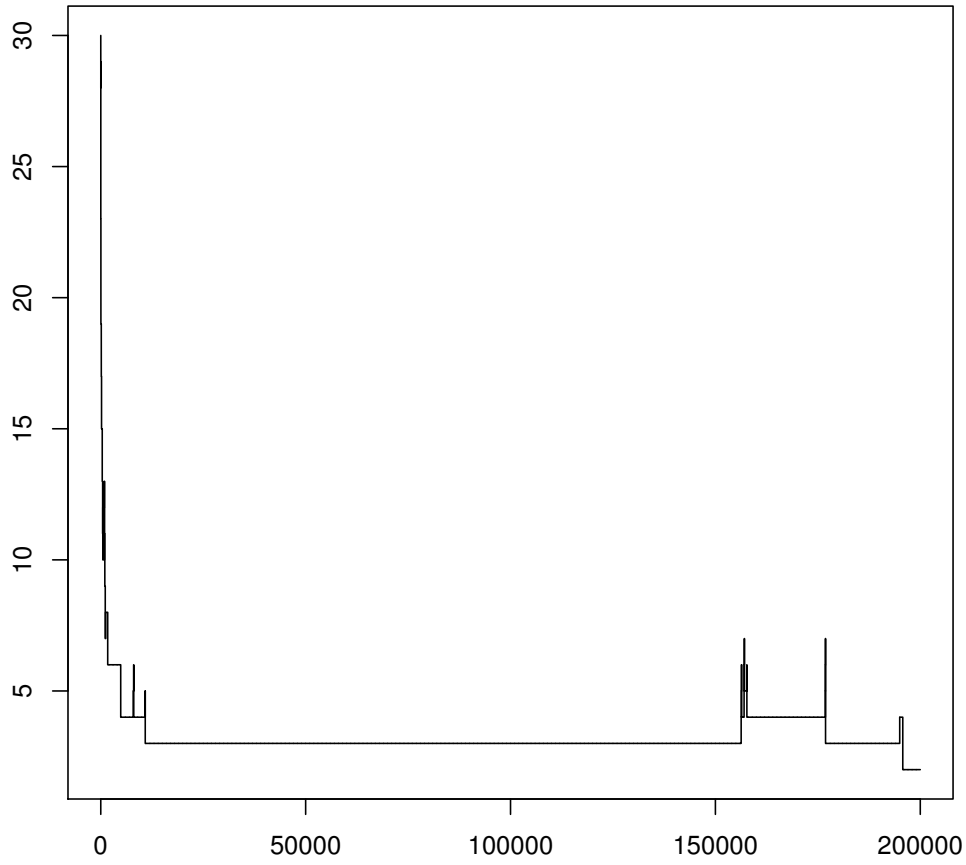
REFERENCES

MR2044670 [DMWZZ04] Persi Diaconis, Eddy Mayer-Wolf, Ofer Zeitouni, and Martin P. W. Zerner. The Poisson-Dirichlet law is the unique invariant distribution for uniform split-merge transformations. *Ann. Probab.*, 32(1B):915–938, 2004.

MR2868048 [GUW11] Christina Goldschmidt, Daniel Ueltschi, and Peter Windridge. Quantum Heisenberg models and their probabilistic representations. In *Entropy and the quantum II*, volume 552 of *Contemp. Math.*, pages 177–224. Amer. Math. Soc., Providence, RI, 2011.

001poisson [Hol01] Lars Holst. The Poisson-Dirichlet distribution and its relatives revisited. *Preprint of the Royal Institute of Technology, Stockholm, Sweden*, 2001.

MR1902841 [MWZZ02] Eddy Mayer-Wolf, Ofer Zeitouni, and Martin P. W. Zerner. Asymptotics of certain coagulation-fragmentation processes and invariant Poisson-Dirichlet measures. *Electron. J. Probab.*, 7:no. 8, 25 pp. (electronic), 2002.



f:bm5N

FIGURE 5. The number of unmatched components larger than ε small.

MR1930355 [Pit02] Jim Pitman. Poisson-Dirichlet and GEM invariant distributions for split-and-merge transformation of an interval partition. *Combin. Probab. Comput.*, 11(5):501–514, 2002.

Sch05 [Sch05] Oded Schramm. Compositions of random transpositions. *Israel J. Math.*, 147:221–243, 2005.

tsilevich [Tsi00] N. Tsilevich. Stationary random partitions of positive integers. *Theor. Probab. Appl.*, 44(1):60–74, 2000.