



# Response to Cho and Liu, “Sampling from complicated and unknown distributions: Monte Carlo and Markov chain Monte Carlo methods for redistricting”

William T. Adler<sup>\*</sup>, Samuel S.-H. Wang

Princeton Gerrymandering Project, Center for Information Technology Policy, Princeton University, Princeton, NJ, United States



## HIGHLIGHTS

- Markov chain Monte Carlo (MCMC) methods are used to study redistricting.
- A previously published paper critiques MCMC redistricting work in general, but makes significant omissions.
- MCMC methods for redistricting are in need of further development, but more progress has been made than is implied by Cho and Liu.

## ARTICLE INFO

### Article history:

Received 23 May 2018

Received in revised form 5 October 2018

Available online 3 November 2018

## ABSTRACT

A question of legal significance is whether an enacted map of political districts is “typical.” Recent work has used Markov chain Monte Carlo (MCMC) methods to produce null distributions of maps in order to answer this question. A recent article by Cho and Liu critiques one particular implementation of MCMC for redistricting, that of Fifield et al. The goal of the present commentary is to draw attention to two facts omitted by Cho and Liu that, if included, would have severely weakened their conclusions. In particular, Cho and Liu point out that Fifield et al.’s algorithm fails to approximate a known target distribution, but neglect Fifield et al.’s use of parallel and simulated tempering, which greatly improves the approximation. Secondly, Cho and Liu argue that it is overly difficult to detect when Markov chains have mixed; they neglect to mention diagnostics used for this exact purpose in Fifield et al.

© 2018 Elsevier B.V. All rights reserved.

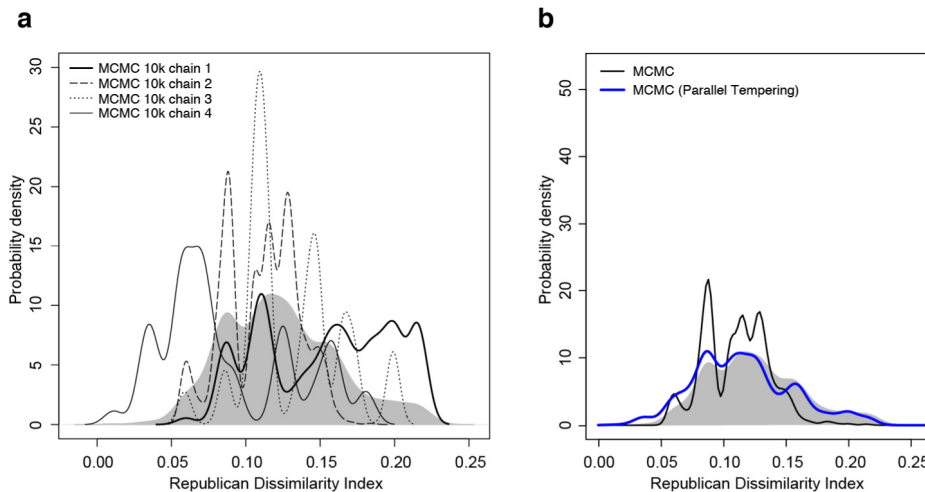
The quantitative analysis of political redistricting is a growing field with substantial real-world implications. This rapid growth requires careful analysis. For this reason, we note with concern the minireview recently published in *Physica A* by Cho and Liu, “Sampling from complicated and unknown distributions: Monte Carlo and Markov chain Monte Carlo methods for redistricting” [1]. This minireview’s main focus is to offer criticism of an approach adopted by Fifield et al. (2017) [2] for sampling from the space of possible district maps. However, Cho and Liu omit facts that, if included, would have severely weakened their conclusions. Oddly, the omitted facts are present in Fifield et al.<sup>1</sup>

The Cho and Liu paper (which we will refer to as CL2018) notes at length that, in most states, the space of possible Congressional district maps is large and difficult to explore, a fact which is well-known [5,6]. The original contribution of the minireview, starting on page 8, is to partially replicate the validation used in Fifield et al. (which we will refer to as F2017). However, the Cho and Liu replication is incomplete in two critical ways.

<sup>\*</sup> Corresponding author.

E-mail address: [wtadler@princeton.edu](mailto:wtadler@princeton.edu) (W.T. Adler).

<sup>1</sup> Cho and Liu cite a version of Fifield et al. from 2017, month and day unspecified. Although a newer draft is available online, we refer to the earliest publicly shared version from 2017 (March 15, 2017). The omitted facts are present in this version as well as in versions of the paper going back to 2014 [3,4].



**Fig. 1.** MCMC chain approximations of a known target distribution (gray, filled). The target distribution consists of all possible three-district maps, from a set of 25 precincts, where each district is contiguous and within a 10% population deviation. The target distributions in a and b are identical but displayed with different scaling, as in the original articles. (a) CL2018 use this figure as evidence that MCMC fails to approximate the (gray, filled) target distribution. Reproduced from CL2018. (b) F2017 also show that standard MCMC (black line) fails to approximate the target distribution, but succeeds when using parallel tempering (blue line), a fact which was omitted in CL2018. Adapted from F2017 (outputs of other algorithms were removed).  
 Source: (a) Reproduced by permission from Elsevier: *Physica A*, Cho and Liu, 2018. (b) Reproduced by permission from B. Fifield: Fifield et al., 2017.

First, Cho and Liu use the algorithm described in F2017 to run Markov chain Monte Carlo (MCMC) chains, noting that the chains do not approximate a known target distribution well (Fig. 1a). The authors use this supposed failure as evidence that the space of district maps is so large that it is nearly impossible for MCMC to explore. The authors then note that “there are various strategies for improving MCMC, including simulated tempering [and] parallel tempering” and that it might be possible to extend MCMC “by, perhaps, using MPI [...] to implement an asynchronous inter-process communication framework where parallel MCMC instances can be efficiently deployed”. But this exact approach was adopted by Fifield et al. to improve their sampling. Fifield et al. themselves note that MCMC fails to approximate their target distribution well (Fig. 1b, black line), but greatly improves with the introduction of MPI and parallel tempering (Fig. 1b, blue line) or simulated tempering (F2017, Fig. 3), the exact methods suggested by Cho and Liu. In short, Cho and Liu present and criticize part of F2017 and fail to mention a following part that specifically addresses the criticism. Additionally, there are papers from another research group [7,8] (which Cho and Liu cite) describing a different MCMC procedure that uses simulated annealing in order to better explore the space of district maps. Cho and Liu write that an effective “MCMC implementation must embody a mechanism that allows it to visit different and disconnected portions of the feasible space”, but they neglect to mention that the very literature they cite contains examples of such mechanisms that have been used with apparent success.

Second, Cho and Liu make a false claim about the difficulties of diagnosing MCMC convergence. After running MCMC chains, without tempering, that fail to approximate the target distribution, they plot the chain traces in CL2018, Figure 6. The authors claim that “diagnostics for the chains indicate that they have mixed when they have not”. However, a brief glance at the traces reveals that the chains have different means and variances; this between-chain variance would be revealed by a common [9] test, the Gelman–Rubin Diagnostic [10]. In fact, Fifield et al. use this diagnostic to demonstrate that they achieve convergence when using parallel tempering (F2017, Figure 5, right). Here, again, Cho and Liu suggest a major problem, omitting that the problem is addressed in the very paper that they criticize. Cho and Liu are correct that convergence checks can never guarantee that one is sampling from the stationary distribution, but it would have been pertinent for Cho and Liu to mention specific checks such as the Gelman–Rubin Diagnostic that would have flagged their plotted traces, especially since the diagnostic was used in F2017. Cho and Liu also make a general mischaracterization of how to determine that a chain is not mixed after some number of steps. They use, as evidence for their claim that their chains are not mixed after 10,000 steps, the fact that the empirical distribution of all 10,000 samples from a chain does not match the known target distribution. This suggests that a chain is only mixed once the empirical distribution of *all* samples taken on the chain matches the target distribution. However, chain mixing only refers to the final state of the chain; it may be that 10,000 steps is sufficient for mixing, in which case the final samples of many 10,000-length chains will be distributed according to the target distribution. Thus, the fact that the empirical distribution of all samples in a 10,000-length chain does not approximate the target distribution is not *per se* evidence that the chain is not mixed.

We agree with Cho and Liu that MCMC is not a silver bullet. However, publication of their unsubstantiated negative claims in a peer-reviewed journal may carry significant weight in court. It is essential for the peer-reviewed literature to contain an accurate account of the intellectual advances being made, such as the efforts toward proper sampling taken by Fifield et al. and others. We therefore offer this correspondence as a means of appending the scientific record.

## CRediT authorship contribution statement

**William T. Adler:** Conceptualization, Writing - original draft, Writing - review & editing. **Samuel S.-H. Wang:** Writing - review & editing, Supervision, Funding acquisition.

## Acknowledgment

This work was supported by Educational Ventures, USA.

## References

- [1] W.K.T. Cho, Y.Y. Liu, *Sampling from complicated and unknown distributions: Monte Carlo and Markov chain Monte Carlo methods for redistricting*, *Physica A* (2018).
- [2] B. Fifield, M. Higgins, K. Imai, A. Tarr, A new automated redistricting simulator using Markov chain Monte Carlo. Working Paper, Princeton University, Princeton NJ. Retrieved August 10, 2018, 2017.
- [3] B. Fifield, personal communication, August 10, 2018.
- [4] B. Fifield, personal website. Retrieved from <https://www.benfield.com/research/>, August 10, 2018.
- [5] Y.Y. Liu, W.K.T. Cho, S. Wang, PEAR: a massively parallel evolutionary computation approach for political redistricting optimization and analysis, *Swarm Evol. Comput.* 30 (2016) 78–92.
- [6] M. Altman, Is automation the answer: The computational complexity of automated redistricting, *Rutgers Comput. Technol. Law J.* 23 (1997).
- [7] S. Bangia, C.V. Graves, G. Herschlag, H.S. Kang, J. Luo, J.C. Mattingly, R. Ravier, Redistricting: Drawing the line. *arXiv:1704.03370*, 2017.
- [8] G. Herschlag, R. Ravier, J.C. Mattingly, Evaluating partisan gerrymandering in Wisconsin. *arXiv:1709.01596*, 2017.
- [9] M.K. Cowles, B.P. Carlin, Markov chain Monte Carlo convergence diagnostics: A comparative review, *J. Amer. Statist. Assoc.* 91 (434) (1996) 883–904.
- [10] A. Gelman, D.B. Rubin, Inference from iterative simulation using multiple sequences, *Statist. Sci. Rev. J. Inst. Math. Statist.* 7 (4) (1992) 457–472.