Michael Thoennessen, Editor in Chief
CC: Terre Arena, Senior Editorial Assistant
Physical Review E

Michael Thoennessen,

Thank you for the opportunity to revise our manuscript, *Capacities and the free passage of entropic barriers*. We appreciate the careful reviews and the constructive suggestions made by the two reviewers. The revision addresses almost all of the points raised by the reviewers. We have attached an overview of the revisions, along with a point-by-point response to each of the reviewer comments.

Sincerely,
  Jackson Loper, Guangyao Zhou, and Stuart Geman

1 Revision overview

As both referees noticed, we the authors of this paper have a mathematics-heavy background, and in our first draft we struggled to make the significance of our work clear to a Physical Review E audience. We did not succeed. Although both referees acknowledged the interesting mathematical aspects of the work, neither of them found the computational motivation compelling. Over the past few months we have tried to make a better study of the the style and language of the field, guided by the pointers from reviewer 2. We think that the new draft is more accessible and better motivated.

There are two major changes:

1. We rewrote the introduction to:

    - clarify what problems this method is good for,
    - clarify how this work relates to existing literature,
    - make our terminology more consistent with standard use.

2. We conducted more comprehensive numerical simulations of the technique's accuracy. In the previous draft we used the numerical results mostly to verify the theory. The new draft tries to take a more methodical look, using a wider variety of problems and quantifying the approximation error in a wider variety of ways.

2 Point-by-point

Reviewer 1. This paper is a mathematical theoretical study...

This paper does contain a fair amount of theory, but it also contains a computational aspect which was easy to miss in the first draft. In fact, in our view this work centers around a new algorithm for estimating hitting probabilities in reversible diffusions. To make this clear, we now devote the second paragraph to focusing on the computational problem we're interested in. We also clarify that this paper includes a practical algorithm for tackling this computational problem.

Reviewer 1. ...that is not easy to follow for the non-mathematician

In the new revision we have tried to make the main ideas clearer:

- Reversible diffusions – in the first paragraph we take a moment to actually describe reversible diffusions, giving a few examples from the sciences. In the second section we give the technical definition.

- Hitting probabilities – in the second paragraph we define these and give an example to clarify. In the second section we give the technical definition.

- When does this algorithm apply – Here we had a bit of difficulty, because there are two ways of thinking about this:
  - On the one hand, there is a technical condition we actually use (ε-flatness). The algorithm is only guaranteed to work if that condition holds
  - On the other hand, we have a more the intuitive and vague condition which implies the technical condition.

  In the third paragraph we try to explain both in progressive order from simple to sophisticated, so that the casual reader can skip unnecessary details. In the second section we give a technical definition.

Reviewer 1. (also because of the notation that is not common for theoretical chemists/physicists).

This was one issue we unfortunately were not able to substantially improve. We tried to find more standard notations for our various variables and concepts, but were unable to figure out what was really standard. If there is anything specific we should change in the notation we would of course be happy to do so. We note there is also a related issue with our first draft, which is that some of our choice of vocabulary was non-standard. We discuss this issue below, with respect to a point from Reviewer 2.

Reviewer 1. References are not order and superscript citations should follow after the punctuation. Last Eq on p6 (not numbered): I think it should be $\rho(x)$ and not $\rho(dx)$

Fixed.

Reviewer 2. The theory well done and the proofs are carried out (and presented) in an understandable manner, but I failed to recognize an immediate relevance for the proposed algorithm ... The authors might therefore reconsider taking some stress from MD and focusing it more on mesoscopic bio-systems where the algorithm could have a much higher applicability.

Following Reviewer 2's suggestion of looking at mesoscopic systems we now include a description of the "narrow escape problem" in the introduction. This is at least one area of study where these methods should find immediately application. For narrow escape problems, the hitting probabilities are proportional to the hitting times, so these methods can be applied to determine the relative hitting times to all kinds of different odd-shaped targets that that literature is interested in.

Narrow escape notwithstanding, Reviewer 2 does have a valid point about the immediate relevance of this work. Frankly, we do not expect readers to apply

our code to their problems. Our lack of expertise in the physical sciences makes it difficult for us to understand the kind of toolchain that would be genuinely useful to science. It is our hope that by publishing in Physical Review E, we could begin to have exactly that conversation. Indeed, this review process has already gone a long way in helping us understand how to usefully apply the theoretical machinery that we are so familiar with. We are grateful to both reviewers for their thoughts in this regard.

> *Reviewer 2.* I could not judge on the novelty of the proofs and also did not check the existing literature, but some aspects did seem vaguely familiar yet were unreferenced.

Reviewer 2 is quite correct: our work contains many propositions and lemmas, and some of these are certainly known by at least some people. For example, the results about capacity and flux follows quite quickly from Green's first identity applied to the relevant equations. We also include an inequality about Bessel functions that is doubtless obvious to those that specialize in such things. Unfortunately, we had a difficult time finding sources in the literature which proved exactly the things we needed. For this reason we were unable to source many of the proofs, even though they are certainly known.

To try to address the reviewer's valid point, we now include several places where we mention that a result is known.

Our main results (Theorem 1 and 2) are substantially new to the literature, as far as we know.

> *Reviewer 2.* (i) The motivation builds heavily on the analysis of molecular dynamics data yet the algorithm will presumably be rather useless for the analysis of macromolecular systems because the central assumption (of small absorbing points in a mostly flat landscape) is mostly incompatible with the respective landscapes.

First, perhaps most importantly: in light of this point of view, we completely dropped molecular dynamics from our motivation. We instead included some references to the narrow escape problem (which does have implications in mesoscopic bio-systems) for which the proposed algorithm is more obviously relevant.

Second: we actually think that the central assumption of small absorbing points in a mostly flat landscape *is* quite relevant to macromolecular systems. Our expertise is limited and we may be wrong. However, we note TCB McLeish has said that "folding rates are controlled by the rate of diffusion of pieces of the open coil in their search for favorable contacts" and "the vast majority of the space covered by the energy landscape must actually be flat." There is also some experimental evidence that the exploration of regions with golf-course potentials is the rate-limiting step for a variety of processes. We would be curious to hear Reviewer 2's point of view on this; it is possible we may have missed something

fundamental.

Some references we looked at that gave us this impression:

- T C B McLeish. Protein folding in High-Dimensional spaces: Hypergutters and the role of nonnative interactions. Biophys. J., 88(1):172183, January 2005.

- J M Goldberg and R L Baldwin. A specific transition state for s-peptide combining with folded s-protein and then refolding. Proc. Natl. Acad. Sci. U. S. A., 96(5):20192024, March 1999

- M Jacob, M Geeves, G Holtermann, and F X Schmid. Diffusional barrier crossing in a two-state protein folding reaction. Nat. Struct. Biol., 6(10):923926, October 1999.

- Kevin W Plaxco and David Baker. Limited internal friction in the rate-limiting step of a two-state protein folding reaction. Proc. Natl. Acad. Sci. U. S. A., 95(23):1359113596, November 1998.

- David J Wales. Energy landscapes: calculating pathways and rates. Int. Rev. Phys. Chem., 25(1-2):237282, January 2006

Reviewer 2. ...But with the view of what is to be achieved algorithmically, I could not detect a breakthrough with respect to what was done by Schuss, Keller or Ward, partly decades ago...In particular, the golf-course problem (as the authors called it) has been solved asymptotically by singular perturbation theory and matched asymptotics.

For several special cases, Reviewer 2 is correct that there are existing methods which could be used to get reasonably accurate estimates for hitting probabilities. We now include a section comparing our work with these asymptotic methods.

In brief, unlike the existing work, there are three differences between this work and the approach taken by Schuss/Keller/Ward:

1. Not all of our results are asymptotic. One of our results is, but one of our results is not:

    - Our main theorem really isn't asymptotic at all: it proposes a method for approximating hitting probabilities, and shows that the error of this approximation is bounded in terms of a value ε which quantifies how much the hitting probabilities vary inside a set M.

    - We also have a second theorem which shows that this ε vanishes as the target sets become smaller. This is indeed an asymptotic result, although qualitatively different from the kinds of results found in the narrow escape literature. Showing that ε vanishes is much different from matching the asymptotics. To show ε vanishes it suffices to

get a rate of uniformly ergodicity inside a set M and show that it is fast relative to the typical escape time from M (e.g. it is sufficient to lower bound $E[\tau^\alpha]$ for any $\alpha > 1$). By contrast, matching the asymptotics requires a careful treatment of the entire process.

2. Second, taking a step back, the the proposed algorithm is posed to handle a much wider variety of problems, because it requires very little special analysis to deal with the particular shape of the target. By reducing the problem to a small region around each target, the proposed algorithm makes it feasible to compute the relevant properties numerically; no asymptotic analysis of the shape are necessary. Moreover, the proposed method applies even in the case of nontrivial energy landscapes, which should be helpful. For example, the axons studied in the narrow escape literature certainly contain large-scale fields, and this method should be able to handle those kinds of issues easily, whereas existing work focuses almost exclusively on the Brownian Motion case.

3. Third, our method says *nothing* about the overall timescale of the process. The existing asymptotic analysis are able to give information about these timescale because they make strong assumptions about the process (e.g. that it is a Brownian motion in a sphere). The algorithm proposed here is much more general, but it comes at a steep cost: we can't get timing information.

We try to make these points clear in the new draft of the introduction. The last point we also revisit in the discussion.

Reviewer 2. (iii) the terminology is non-standard, occasionally unclear and often even contrasting the standard use. For example, the term entropic barrier...It was also not entirely clear whether capacity is used in the same context as in probabilistic potential theory

We tried to tighten this up as much as possible. For example,

- We removed all reference to "entropic barriers."

- What we used to call "capacity" we now call "first passage capacity" to differentiate it from all the other notions of set capacity. We compare it with Harmonic capacity and cite the work of Anton Bovier (who uses essentially the same notion of capacity). These capacities are indeed to be understood in the context of probabilistic potential theory, a fact we now mention.

- We mention that our "hitting probabilities" are sometimes called "splitting probabilities" (especially in the case when there are exactly two targets; this is the case we mostly discuss though the algorithm holds equally well for any finite number of targets).

We are certainly open to any further suggestions in this regard.

> *Reviewer 2.* The existing literature is extremely poorly described.

We have tried to substantially bulk out our description of what is known. We now describe four different literatures and relate our approach to each one.

> *Reviewer 2.* However, the assumption of MSM is actually orthogonal to the assumptions made in the present work (yet this appears to have slipped under the radar). Namely, MSM are built on the assumption that the residence time in the metastable cores (and hence the escape from a core) is much longer than the time spent outside of the cores...splitting probabilities from a point located initially *not* inside a stable core seem to bear little relevance in describing conformational dynamics...Splitting probabilities in that sense (even if determined with perfect accuracy) as a whole would e relevant for the dynamics in-between the cores only upon the further assumption of renewal dynamics between the cores.

This is a subtle point. We will first try to explain our point of view here, then describe what we did to address this remark in our revision.

Let $\{X_t\} \subset \Omega$ denote a diffusion.

Markov State Models. Let $f : \Omega \to \{1, 2, \cdots n\}$. MSM is built on the assumption that $Y_t = f(X_t)$ is "nearly Markov," i.e. one can find a Markov process $\tilde{Y}$ which has nearly the same distribution as Y. This often happens, for example, if the rate of ergodicity inside each partition is high relative to the escape time from each partition. When these conditions hold, one can approximate the gist of the dynamics of X efficiently by simulating from $\tilde{Y}$.

The approach proposed here. We assume $A_1, A_2 \cdots A_n$ are a collection of targets and there is a region M such that the chance we hit A_i first is nearly the same for every initial condition $x \in M$ and every target i. This often happens, for example, if the rate of ergodicity inside each partition is fast relative to the escape time from M. When this condition holds, one can approximate the hitting probabilities efficiently by calculating first-passage capacities.

Now there are (at least) two ways in which these two approaches are related.

- First, one could use MSM to compute hitting probabilities. Indeed, one could accelerate the computation of the hitting probabilities calculated in our approach by subdividing M into a bunch of different regions and then performing approximate simulations on M.

- Second, one could use the approach proposed here to perform something very similar to what MSM does. To do so, one needs an extra assumption: that every path between A_i and A_j passes through M. Under this further assumption, it follows that the process will alternate between two phases: starting from M it will go to some target, then starting from some target

it will go to M. We can therefore simulate the order in which the process explores targets by modeling the first-passage capacities.

Our original draft described these two connections in a fairly sloppy way. In the new version, we try to focus very clearly on the task that we actually solve, namely computing hitting probabilities. Thus in the introduction we mention MSM *only* in the context of the first way that the methods are two related (i.e. insofar as MSM could help us compute hitting probabilities). We describe how the proposed method is applicable in cases where MSM would be hard to pull off. We do also make some mention of the second method, but only at the very end of the discussion.

Hopefully the new revision clarifies this somewhat confusing tangle of ideas.

Reviewer 2. I failed to see any quantitative evidence for the quality of the approximations, as advertised in section 3. Nor did I could recognize the gauge on what 'good agreement' is supposed to mean.

This point is well-taken. In the previous draft we used the numerical results mostly to verify the theory. The new draft tries to take a more methodical look, using a wider variety of problems and quantifying the approximation error in a wider variety of ways.

2.1   Reviewer 2

In their work the authors combine clean mathematics with the attempt of creating a useful computational algorithm for estimating splitting probabilities to small absorbing targets for overdamped Langevin dynamics in higher dimensions. The theory well done and the profs are carried out (and presented) in an understandable manner, but I failed to recognize an immediate relevance for the proposed algorithm. I could not judge on the novelty of the proofs and also did not check the existing literature, but some aspects did seem vaguely familiar yet were unreferenced. In case not all proofs are indeed completely new (the referee makes no judgment on that), the known parts/results should be appropriately acknowledged.

In general, the paper is unbalanced on all fronts.

(i) The motivation builds heavily on the analysis of molecular dynamics data yet the algorithm will presumably be rather useless for the analysis of macromolecular systems, because the central assumption (of small absorbing points in a mostly flat landscape) is mostly incompatible with the respective landscapes. One could used it in the context of milestoning, core-set methods (and other computational technologies), but then one needs to reconsider some additional points given in (v) below. The authors might therefore reconsider taking some stress from MD and focusing it more on mesoscopic bio-systems where the algorithm could have a much higher applicability.

(ii) The theory is quite rigorous (for a physics audience in particular) but carried out very well and presented clearly. But with the view of what is to be achieved algorithmically, I could not detect a breakthrough with respect to what was done by Schuss, Keller or Ward, partly decades ago. In particular, why the need for rigorous formal asymptotics if the elaborate/convoluted numerics completely lack an accuracy/error analysis. In particular, the golf-course problem (as the authors called it) has been solved asymptotically by singular perturbation theory and matched asymptotics. Without a proper (quantitative) error analysis the combination of clean theory and uncontrolled numerics is unreasonable.

(iii) the terminology is non-standard, occasionally unclear and often even contrasting the standard use. For example, the term entropic barrier is canonically used in the context of dynamics on effective potential hyper-surfaces arising in dimensional reduction, when the potential energy along a reduced coordinate is flat but included a narrow opening creating an *entropic barrier* depending on the noise intensity/temperature. True, the curvature around a small domain creates a geometrical barrier along a 'proximity coordinate', but this is typically called a strong localized perturbation. It was also not entirely clear whether capacity is used in the same context as in probabilistic potential theory – some additional clarifying remarks are needed. The authors can keep the terminology but then must also explain the canonical one (and also why they decided to choose another one).

(iv) The existing literature is extremely poorly described. In particular, at lest the abundant works of Michael Ward, Joe Keller and Zeeev Schuss as well as their 'offspring' should be appropriately acknowledged. I would also suggest to have a look at Morro, J. Chem. Phys. 103, 7514 (1995) on a very intuitive analysis in a macromolecule-relevant setting written for a physics audience. Also, the advantages and disadvantages of the proposed method with respective to existing computational methods (which are superficially mentioned) should be included.

(v) Physics: Although I am not a fan myself, Markov-state-models (MSM) did emerge as very successful reduced models of bio-macromolecular dynamics. However, the assumption of MSM is actually orthogonal to the assumptions made in the present work (yet this appears to have slipped under the radar). Namely, MSM are built on the assumption that the residence time in the metastable cores (and hence the escape from a core) is much longer than the time spent outside of the cores. The fact that MSM often work well, but most importantly the abundant experimental evidence that macromolecules have identifiably long-lived conformational states, supports that this assumption is physically correct. Therefore, splitting probabilities from a point located initially *not* inside a stable core seem to bear little relevance in describing conformational dynamics, even if 'entropic barriers' (in the sense defied by the authors) are present.

Starting from a core, there is one additional aspect. The underlying landscape is allowed not to be flat near the 'attractive' cores and the potential is flat sufficiently far away from the cores (this assumption is actually potentially

more restrictive than it might appear at first). Splitting probabilities in that sense (even if determined with perfect accuracy) as a whole would e relevant for the dynamics in-between the cores only upon the further assumption of renewal dynamics between the cores.

(vi) I failed to see any quantitative evidence for the quality of the approximations, as advertised in section 3. Nor did I could recognize the gauge on what 'good agreement' is supposed to mean. The difference with respect to simulation is as large as the scatter within the simulations despite only 2000 trajectories have bin considered. If the authors decide that this is a good agreement (the referee reserves all judgment whether this is indeed the case) then they could provide some indication what comparably a bad agreement would correspond to, or at least what is the basis for this assessment.

As a bottom line, I recommend to ante-up the simulations/computational aspect of the work to match the well-done theoretical part, as well as reconsidering the presentation-style if the computational physics audience is to be targeted as the main audience.