

Dear Editors,

We write to appeal the decision to reject publication of our manuscript in PRL. The three responding referees were unanimous in agreement that our work presents a compelling theoretical and numerical work and raise no objections on that front. The sole objection raised, chiefly by Referee B, relates to the referees concerns about the feasibility of a handful of loosely proposed experimental tests of our model. We argue that this is an unjust standard to apply to theoretical and numerical work. Judged on its own merits this work meets the criteria of PRL for validity, importance, and broad interest. It was never our intention to offer this paper as experimental confirmation of such a theory (indeed, that would be an entirely different paper) but instead to develop theoretical predictions for the extreme value statistics of many-particle diffusion that we subsequently numerically verify down to realistic physical system sizes.

Referee A states that this is “a beautiful theoretical work”. Referee B remarks upon the, “innovative aspects of the work, its high quality, and its potential appeal for statistical physicists.” Referee C finds that our prior response, “clearly answer[s] the objections and successfully motivate[s] the relevance of their work,” and goes on to find the manuscript to be well performed and of general interest to a broad audience. Finally, referee C agrees that, in opposition to referee B’s findings, “The experiment proposed appears doable, and it could stimulate new further research.” We reiterate that we do not believe that the plausibility of these experiments should be the determining factor in the publication decision. Nevertheless, even if it were, we believe that we would pass that bar.

Below, we respond point-by-point to the referees’ concerns. We hope that in doing so we have demonstrated the appropriateness of this work for publication in PRL.

Second Report of Referee A -- LS18194/Hass

The authors provided very detailed and satisfactory replies to all former comments. They considerably improved the readability of the manuscript. While this is a beautiful theoretical work, I still doubt on its practical relevance; in particular, the discovered intermediate regime in Eq. (5) is hardly observable, as fairly witnessed by the authors:

"Although the full characterization of the short-time regime is in excellent agreement with the numerical results, the $t^{1/3}$ power-law is difficult to capture since the transitional window of $\log(N)$ to $(\log(N))^2$ is too narrow for realistic sizes of N , even up to $N = 10^{300}$."

We agree with the referee that it is unlikely that this transient $t^{1/3}$ power law will be observed in experiments. Such behavior is expected only within a narrow range of times, set by $\text{Log}(N)$. The practical value of this work lies, instead, in the long-time universal regime. Our work demonstrates that such a regime, marked by a very clear power-law, exists even in the limit of

very small system sizes, such as $N \sim 100$. Moreover, we provide a roadmap for extracting previously hidden information about the correlation structure of the environment from measurements of the extreme value statistics.

As a consequence, the manuscript might be more appropriate to Physical Review E but I would also be glad to read it in PRL.

Our work highlights that there is much more to be discovered in simple many-particle diffusion than previously thought. This opens the way to a wide range of future numerical, theoretical and experimental investigations. As such, and due to the wide interest in diffusion across science, it seems quite appropriate for the broad readership of PRL.

The authors are encouraged to address the following minor points:

1) Page 1, column 2:

"...even when the number of particles is very small".

I suppose that the authors refer here to $N = 10^2$. This is of course very small as compared to 10^{300} , but it is not so small for many realistic systems. It is therefore suggested to modify this sentence as

**"...even when the number of particles is as small as a few hundred".
(or similar)**

Done. The sentence now reads:

"This residual environmental variance is characterized by a novel power law that we demonstrate holds even when the number of particles is as small as a few hundred."

2) Caption of Fig. 1: The sentence

"We also plot in green the asymptotic theory mean location for the maximum particle location Around this is a shaded region with width two standard deviations based on the asymptotic theory variance."

sounds a little weird. Perhaps, the dot is missing after "particle location". Perhaps, "with width two standard deviations" can be replaced by "with width of two standard deviations"?

Thank you for catching our mistake. The caption now reads:

"We also plot in green the asymptotic theory mean location for the maximum particle location. Around this is a shaded region with a width of two standard deviations based on the asymptotic theory variance."

3) Page 2, column 1: The sentence

"Barraquand and Corwin [62] discovered the exactly solvable Beta RWRE discussed extensively below and uncovered a remarkable connection between its large deviations and the Kardar-Parisi-Zhang (KPZ) universality class [63, 64] Gaussian Unitary Ensemble (GUE) Tracy-Widom distribution [65] which holds for times comparable to $\log(N)$."

perhaps misses a dot (after [63,64]) or some linking words.

Has been replaced as

"Barraquand and Corwin [62] discovered the exactly solvable Beta RWRE discussed extensively below and uncovered a remarkable connection between its large deviations for times of order $\log(N)$ and the statistics of the Kardar-Parisi-Zhang (KPZ) universality class [63, 64], namely the Gaussian Unitary Ensemble (GUE) Tracy-Widom distribution [65]."

4) In bibliography, some authors are provided with full names (like Ref. [6]) and some others with abbreviated names (like Ref. [7]). Please homogenize.

The referee is correct and we will work with the editorial staff to ensure the bibliography is homogenized.

5) Refs [28,30] are identical, please remove one.

Thank you for pointing this out. We have removed the duplicate.

Second Report of Referee B -- LS18194/Hass

I thank that the authors for their detailed answers to my questions, in particular, my concerns about the relevance for physics. I am also glad that the authors followed my recommendation to make their numerics code publicly available, which fosters the aims of reproducible numerical research.

I acknowledge that the authors have initiated collaborations to connect their mathematical results to actual experiments. However, these new projects appear to be in an early phase. It is not clear at all that these experiments can be conducted in such a way that the predicted phenomena are likely to be observed, simply for practical limitations (see below for details).

We appreciate the referee's concerns, however we strongly disagree that this work should be judged based on a detailed evaluation of our future experimental plans, which we've only very broadly sketched out in our previous response. Clearly, experimental tests of the time of first passage of diffusing objects are in principle possible. Such tests will surely be the subject of future papers from our group and (we hope) others. The present work exists to lay out a firm theoretical and numerical foundation upon which experimental (and further theoretical) work may be laid to construct this new field of extreme diffusion.

Overall, I have the impression that, conceptually, it remains a considerable experimental challenge to collect sufficient statistics of the extremes that are needed to draw value from the theory for a concrete physical system. Therefore, I regret that I cannot conclude that the present manuscript should be published in Physical Review Letters. I recommend to transfer the work to Physical Review E, where it may qualify, given the innovative aspects of the work, its high quality, and its potential appeal for statistical physicists.

We thank the referee for their kind and very positive assessment of the innovation and high quality of our work. As previously stated, we strenuously disagree with the assertion that this work should be judged on the strength of our proposal for experimental validation. Undoubtedly, such experiments will offer challenges, as all new experiments must. However, we firmly believe that there is strong evidence that such experiments will be possible.

Experiments:

Specifically, colloidal motion in quasi-1D confinement (including circular confinement) has been studied some years ago, with focus on the intriguing phenomenon of single-file diffusion. In such an experimental setup, the colloidal particles cannot pass through each other and thus the front is always marked by the same particle. This is a completely different situation than that of non-interacting particles covered in the manuscript. In order to make a link one would need to repeat the experiment N times for the same environment. However, the slowness of colloidal motion appears to drastically limit the number of measurements that the experimental staff can obtain realistically to perhaps $N \sim 100$, certainly less than 1000. This is in stark contrast to huge values of N used in the simulations, needed to identify the scaling features of $\text{Var}(\text{Env}_t)$.

There are innumerable potential experimental systems involving the extreme value statistics of diffusion. Thus, we are somewhat surprised to see the referee describe a very different sort of experimental system, one for which our model is likely ill-suited as the described system very explicitly breaks the assumption of small inter-particle interactions. In the referee's system, particles are unable to move past one another and thus the ordering is maintained. One can consider this as an extreme limit of a highly interacting and truly one dimensional model. As such, it would not make sense to model it as a non-interacting diffusive system. We have no doubts, however, that it is an interesting system in its own rights.

By contrast, the type of (quasi-1d) system which we could imagine our model applying appropriately to is a long thin capillary whose width is still many times the size of an individual particle. Such a system would allow particles to freely pass by one another while still being subject to local and correlated forces within the fluid.

Beyond the issue of applicability of our model to the referee's system, the referee is concerned about the number of repeated measurements that can be performed in light of the fact that our simulations were done using repeats ranging from 500 to 10,000. We wish to point out that such large numbers of repeats were performed solely to aid in the clarity and impact of our figures. One can extract meaningful results with as few as 1000 repeats, as shown in the below figure.

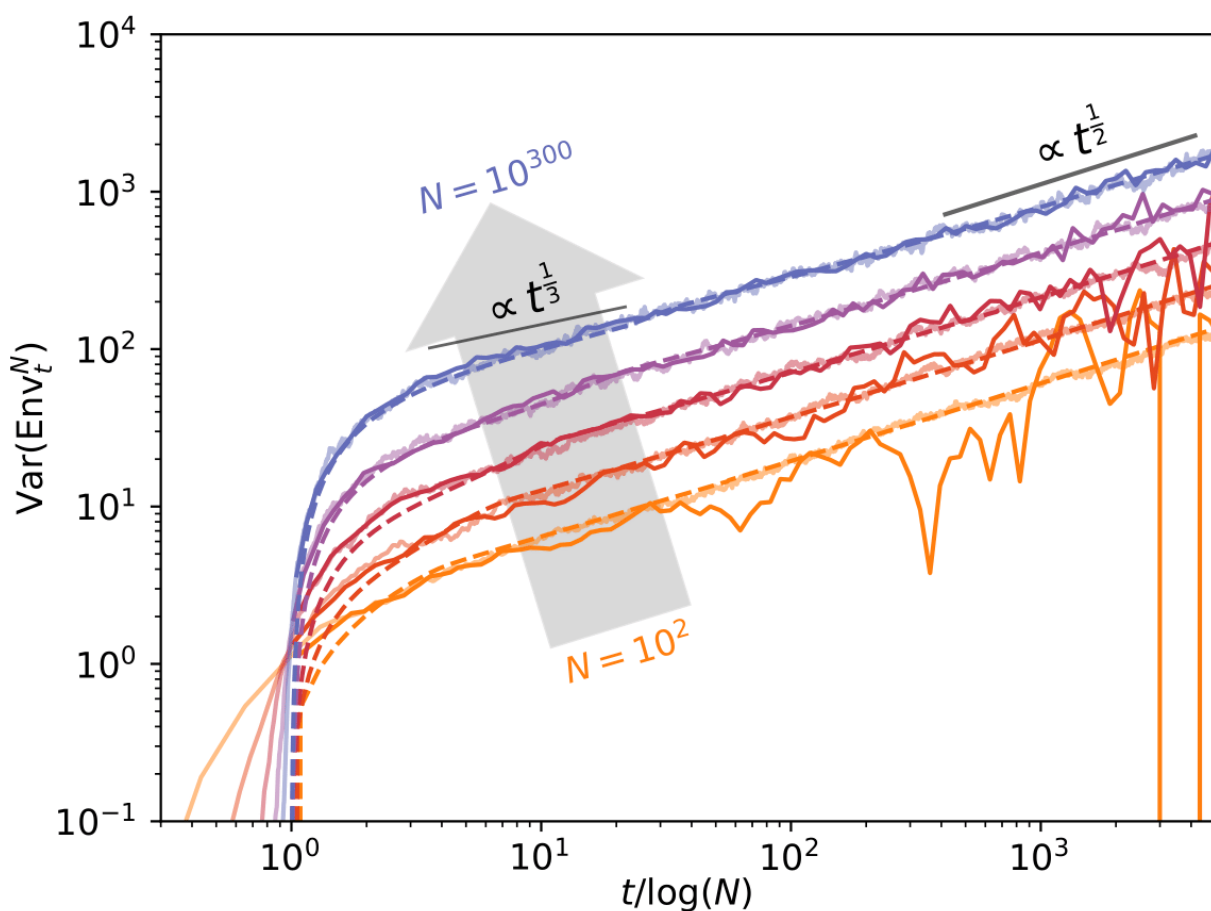


Figure 3. Plots of $Var^{num}(Env_t^N)$ (transparent solid) computed over 500 environments, $Var^{asy}(Env_t^N)$ (dashed), and $Var^{num}(Max_t^N) - Var^{asy}(Sam_t^N)$ (dark solid) smoothed in each $1/25^{th}$ of a decade and computed over 1000, 1000, 1000, 500, 500 environments (respectively) for $N = 10^2, 10^7, 10^{24}, 10^{85}, 10^{300}$.

Experimental systems can easily be multiplexed to achieve rapid and parallel measurements, allowing for a large number of repeats to be economically performed. Back of the envelope

calculations suggest that one could measure the first passage of micron-scale particles traveling mm-scale differences in time periods of the order of days, making such experiments easily accessible.

Furthermore, the referee is inaccurate in claiming that huge values of repeats are needed to identify the scaling features. Scaling is such a nice framework to work within because relatively few measurements are needed provided that they are done across a wide range of parameter space. The scaling is then determined from the collapse of the data across parameter space.

The referee also asks about the repetition of the same measurements for a fixed environment. This is of course, experimentally impossible, as the microscopic environment is a dynamically changing thing. In fact, a punchline of our work is that it should not be necessary to do such a repeated measurement. Since this is subtle (though explained in detail in the paper) let us unpack this a bit more.

For our numerical work we do make a direct measurement of the variance of Env^N_t – this requires finding the $1/N$ quantile of the CDF of the RWRE for a particular environment and then measuring how much that varies after the environment is replaced by a fresh one. This type of direct measurement would be impossible in an experimental setting since one cannot find that CDF for a given environment – doing so would require somehow quenching an environment and then repeating an experiment in it multiple times. Perhaps this is what the referee believes we are proposing? In any case, it is not.

What we instead learn from our work is that for our numerical model it is possible to recover the variance of Env^N_t as the difference between the variance of Max^N_t and Sam^N_t . Each of these variances, in contrast to that of Env^N_t , are potentially measurable experimentally. The first requires studying the maximum of N particles at time t and then computing the variance over multiple different environments of this maximum; the second can be computed by studying the diffusion coefficient for a single particle and using the formula we provide in the conclusion that $\text{Var}(\text{Sam}^N_t) = (\pi^2 D/6) t/\log(N)$. We numerically verify that this subtraction of the Sam^N_t variance from that of Max^N_t does recover the true variance of Env^N_t . This conclusion will inform our future experimental work in that we will study the same difference of experimentally measurable variances and then be able to compare the result to the predicted power-laws and coefficients.

Finally, the scope of this paper was not intended to include instructions for fully realized experiments, but provide an alternate framework of diffusion which could be measured (even if this is difficult).

A (perhaps naive) setup that would more closely resemble the situation described in the manuscript is as follows: Consider the spread of a droplet of ink in a highly viscous, random medium. Here, I can imagine that it is possible to obtain the (locally averaged) concentration profile by imaging and to infer the mean position of the front from it. This corresponds to the green solid line in Fig. 1. However, the relevant information is

contained in the variance of the leading ink molecule (the green shaded region in Fig. 1), and it is not obvious to me how to get hold of this quantity in the experiment.

We agree with the referee that there are in fact an enormous number of ways in which one could devise a measurement system to extract the extreme statistics of diffusion. The systems that we have suggested are ones for which we have a good understanding of how to control the sources of error and systematics. However, as the referee ably demonstrates there are in fact myriad other systems available. We would argue that this strengthens our claims of the universal appeal and applicability of this work.

To address the specific points raised, if one can measure the mean position of the front of diffusing ink then repeated measurements would yield a measurement of the variance. Additionally, one could improve this particular system by using fluorescent dye molecules for which single-molecule imaging is easily performed.

I cannot comment on the feasibility of the second set of experiments, based on time-of-flight measurements of photons. Again, neither a detailed description nor convincing quantitative estimates of the outcomes (also taking into account expected measurement errors) were provided by the authors.

Again, the purpose of the paper was not to provide fully fleshed out experiments but provide a theoretical framework for diffusion which could be verified (or disproved!). We believe the experimental setups mentioned in the paper are sufficient to argue that the proposed model can be feasibly measured. Indeed, we are currently working towards making precisely these measurements using colloids diffusing in a quasi-1D lattice and photons scattering through a random medium.

Interacting particles:

It is true that the density of particles is low at the edge of a spreading front. However, the statistics of the front sensitively depends on the motion at previous times, when the leading particle may have been inside the bulk. (This is the whole point of the paper.) So I cannot accept that argument that interactions would be irrelevant for the present predictions of the extreme statistics.

The referee is correct that the particles at the leading edge must move through the bulk which has a high density to reach the edge. However, ample evidence exists to demonstrate that the bulk of the system is accurately modeled by classical diffusion, i.e. diffusion coefficients, which completely ignores particle interactions. It has been proven our proposed model is consistent with classical diffusion in the bulk. Therefore, it is reasonable to assume particle interactions to the next leading order can be ignored. Furthermore, it is very reasonable to consider the behavior of the leading particle as independent as one can readily see that the leading particles quickly separate from the bulk and thus spend nearly all of their time experiencing considerably lower densities.

Like any physical model, we have made simplifying assumptions, namely that particles don't interact. However, our model makes fewer assumptions than those in classical diffusion which ignores any effects due to the environment. We have provided a compelling argument that particle interactions are less important than environmental effects and outlined experiments that would verify if this assumption is justified or not. Therefore, we believe this paper provides a large advancement in our understanding of diffusive behavior.

Connection to the KPZ equation:

Readers may appreciate if the equations given in the answer to Referee A would also be included in the manuscript.

As requested by the referee we have revised the bottom of the first column of page 2 to include the multiplicative noise stochastic heat equation and KPZ equation in our discussion of why universality should hold in the current RWRE setting. As we have now stated the KPZ equation at this point, we also removed its definition later in the text in the RWRE Env^N_t subsection and now reference back to the earlier definition.

The revised text on the bottom of the first column of page 2 now reads:

The recursion relation (1) for RWRE transition probabilities solves a discrete version of the multiplicative noise stochastic heat equation (mSHE)

$$d_t Z(x,t) = \frac{1}{2} d_x^2 Z(x,t) + \xi(x,t)Z(x,t)$$

with ξ space-time white noise. The logarithm of the mSHE $h(x,t) = \log Z(x,t)$ solves the KPZ equation

$$d_t h(x,t) = \frac{1}{2} d_x^2 h(x,t) + \frac{1}{2} (d_x h(x,t))^2 + \xi(x,t).$$

Hence, large deviations for RWREs, in particular beyond the solvable model and even in experimental settings, may relate to the KPZ equation and its universality class

Minor issue: The meaning of the symbol ' \ll ' is presumably clear to many physics readers. Yet, it is still confusing that the strict inequality sign " $<$ " is used to express an asymptotic relation. For example, $t / \log(N) < 1 / \log(2)$ implies merely that $N > 2^t$, but not necessarily that $N \gg 2^t$. As an example, the former relation is already fulfilled by $N = (1+\epsilon) 2^t$ for any $\epsilon > 0$, but in this case $N / 2^t = 1 + \epsilon$ does not tend to infinity in whatever limit is taken.

We appreciate the referee raising the concern around this notation and offering an example to illustrate their confusion. The notation \ll or \gg was defined in the first few lines of the "Asymptotic Theory Results" subsection where we stated that given a relationship between t and $\log(N)$ such as $t/\log(N) = \hat{t}$, we say that $f(N,t) \gg g(N,t)$ if $f(N,t)/g(N,t)$ tends to infinity as t and N do while respecting their given relationship. It is this point about the relationship being given where the referee seems to have gotten confused. In the referee's proposed counterexample the relation $t/\log(N) < 1/\log(2)$ does hold. However, their example does not

correspond to $t / \log(N) = \hat{t}$ for some $\hat{t} < 1 / \log(2)$ fixed. Instead, their effective \hat{t} would be converging to $1 / \log(2)$ in contradiction to our assumption of a fixed ratio of $t / \log(N)$. In any case, we understand that this point created some confusion and thus we have emphasized in our text that the ratio $t / \log(N)$ is fixed to equal some \hat{t} (or at another point that $t / \log(N)^2 = \hat{t}$).

Report of Referee C -- LS18194/Hass

The authors study a model of random walk in a random environment in One dimension, focusing on the case of uniform distribution for the stochastic environment. This corresponds to a well-defined choice of parameters in the Beta distribution, for which the model is exactly solvable. Combining high-precision numerical simulations and analytical predictions, they analyze the behavior of extreme particles that are affected by the random environment unlike particles in the bulk. Most of the analytical calculations on asymptotic behavior are reported earlier or derived thanks to already published results, as also specified by the authors. However, the authors' main claim is to interpolate between these asymptotic phases, in such a way to acquire predictive power for realistic physical sizes, thus testable in experiments. As they explain in referee's answers and in the revised version of the manuscript the relevance of this study lays in the degree of the universality this model has, namely in the asymptotic power-law behaviors of the outliers and in the fact that the variance of the extreme particles always decouples in the sum of the environment and of the sampling one. According to them, this division can allow to infer information on the statistics of in-homogeneous environments of physical systems, just by analyzing the behavior of extreme particles.

Only the revised version of the manuscript was available to me. This manuscript appears well written and clear, it is also visible how the authors modified the text in order to supply referees' arguments. My opinion is that they clearly answer the objections and successfully motivate the relevance of their work. Moreover, the work appears well done and of general interest for physicists, since it concerns a very basic model that eventually can interest a broad audience. One of the main concerns of Referee A, namely the experimental applications, has been discussed by the authors and clarified in the main text in a more quantitative than speculative way. The experiment proposed appears doable, and it could stimulate new further research. For these reasons, I think this paper is suitable for publication in PRL.

Just some additional comments: In several points of the paper, authors highlight the relation between the probability distribution of the environment and the KPZ equation. The sentences they added in this revised version make it mathematically clear to understand why. However, I would appreciate it if this relation can be made more quantitative. KPZ universality class is identified by scaling laws and exponents, then which of them authors think can be measured in this type of systems? Do they think is it related to the universality of their asymptotic regimes? They can clarify it a bit more in

the text.

The referee makes an important point that we clearly haven't explained enough in the paper. We expect the long time behavior of our system, corresponding to a $t^{1/2}$ power law, to be experimentally accessible. Measuring this power law corresponds to the short time behavior of the KPZ equation up to a prefactor by Eq. 4 in the text. Therefore, by a simple transform, the short-time behavior of the KPZ equation could be experimentally measured with our system. We have added the following to the “*Comparison of Numerical and Theoretical Results*” section

“By measuring the long-time $t^{1/2}$ power law, we measure the short-time scaling behavior of the KPZ equation up to a prefactor using Eq. 6.”

Do they think long-range correlated environments would behave completely differently? Can they give an insight of why and how?

This is a really great question that points towards important future theoretical considerations. Long-range correlated environments can come in different flavors – correlations that decay slowly in space or in time; and those that vary with the system size versus are fixed relative to that. In the latter case where the correlation lengths are fixed as the system size grows, we expect by coarse-graining that the short-range approximation should become approximately accurate. However, in the former case where the lengths grow with the system size or are suitably heavy-tailed, we expect other behavior may emerge and significantly affect the behavior of extreme particles. This question definitely warrants further investigation in subsequent work.