

First-passage time in space-dependent stochastic resetting

Johannes Aspman ¹, Daniel Mastropietro ^{1,2,3} and Jakub Mareček ¹

¹*Department of Computer Science, Czech Technical University in Prague, Czech Republic*

²*Université de Toulouse INP, 31071 Toulouse, France*

³*CNRS-IRIT, 31071 Toulouse, France*



(Received 16 September 2025; accepted 13 July 2026; published 7 August 2026)

We consider the mean first-passage time (MFPT) through a target of interest for a diffusive particle of Langevin type, with the added condition that the particle is reset to its original position with some rate r . We study both smooth and nonsmooth, nonconvex potentials, focusing on the case where the reset rate depends on the space coordinate. For quadratic and piecewise-quadratic potentials, we show that the benefits of resetting depend on the ratio between drift and noise, and become more important as the drift potential becomes smaller compared to the noise. When the target is a local optimum of the potential, we further show that it is beneficial to use a space-dependent resetting where the reset rate is lower when the particle is closer to the target.

DOI: [10.1103/w2yj-91m8](https://doi.org/10.1103/w2yj-91m8)

I. INTRODUCTION

Searching for a given target is something we do every day, and it is a ubiquitous problem in modern technology and research. For example, enzymes searching for a target site in a DNA strand [1], neural network training algorithms searching for an optimum [2], or simply when we are looking for our keys in the morning. Resetting strategies for searching are built on the simple idea that after performing the search for some amount of time a natural step might be to return to a given point and start your search again.

A natural question is to ask whether resetting provably speeds up the search. In their seminal papers [3–5], Evans and Majumdar answer the question in a model, where a diffusion process has a nonzero probability rate of resetting to a given fixed point. Similar models have also been proposed by other authors [6–8], and many extensions of Evans and Majumdar’s approach have been studied recently [9–17], largely within statistical mechanics. Resetting has also been studied in mathematical optimization for many years as a way of speeding up algorithms [18,19]. In AI and computer science, there is a long history of engineering local-search algorithms with resets [20–25]. In particular, much of the literature [26–31] focuses on the case where the drift potential is convex and smooth. For example, [28,30] study various first-order methods and show that resetting can be used to accelerate these. Reference [29] discusses Luby’s [18] reset strategy in a continuous setting and [31] derives universal performance bounds on general resetting strategies for stochastic processes. The nonsmooth nonconvex setting [32,33], while clearly relevant in AI and machine learning due to the ubiquitous use of neural network modeling via nonsmooth neuron activation functions, has not been given much attention in the resetting literature.

The noteworthy exception is Ahmad *et al.* [17], who study piecewise-linear potentials as well as smooth, but nonconvex, potentials.

In this paper, we extend the work of Evans and Majumdar [3–5,11], as well as the work of Ahmad *et al.* [17] to nonsmooth, nonconvex potentials and reset rates conditional on the gradient of the objective function, whose global optimum is sought. In the case where the optimum is known, it may seem intuitive to reset less often, once a threshold close to the optimum is reached, for instance. Note that this is an important special case, considering that the empirical risk (the error of a model’s prediction of a target of interest) is bounded from below by 0 in most machine learning applications, but it is also bounded from above by 0 in the overparametrized regime [2], typical of many modern models where the number of parameters is far larger than the number of training cases. In the more general case where the optimum is unknown, an indication of such a situation is partially given by a threshold on the gradient of the objective function: Assuming the optimum occurs in a smooth region of the parameter space, a “small” value of the gradient norm is an indication of being close to a local optimum. This is the case we focus on in this work. Specifically, we consider a piecewise quadratic potential for the diffusion process, and we introduce a piecewise-constant reset rate that depends on the gradient of the objective function. We show that the mean first-passage time (MFPT) can be reduced when one considers the piecewise-constant reset rate, conditional on the gradient norm of the objective function, rather than any single reset rate, or no reset at all. For simplicity, we focus mainly on the one-dimensional case, but provide a brief comparison with the higher-dimensional situation. We also present numerical results in both the one- and two-dimensional cases that support the theoretical findings.

II. PROBLEM STATEMENT AND BACKGROUND

As mentioned in the introduction, we will not consider a pure diffusion, but rather diffusion with drift, or diffusion in

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

a potential landscape. In other words, we are considering a process $x(t) \in \mathbb{R}^d$ satisfying

$$\dot{x} = \mu F(x) + \eta(t), \quad (1)$$

where $\dot{x}$ denotes the temporal derivative of $x(t)$ and $F(x)$ is given by the negative spatial gradient of a potential $V(x) : \mathbb{R}^d \rightarrow \mathbb{R}$,

$$F(x) = -\nabla V(x) \quad (2)$$

or a suitable generalization thereof in the nonsmooth setting (e.g., the subgradient). In general, parameter μ suggests the mobility of the particle, although we consider only $\mu = 1$ in the rest of the paper. The noise $\eta(t) := (\eta_i(t))_{i=1,\dots,d}$ is assumed to be a white Gaussian vector with the property

$$\begin{aligned} \langle \eta_i(t) \rangle &= 0, \quad \langle \eta_i(t) \eta_j(t') \rangle = 2D \delta_{ij} \delta(t - t'), \\ \text{for } i, j &= 1, \dots, d, \end{aligned}$$

where $D > 0$ is called the diffusion coefficient, δ_{ij} is the Kronecker delta, and $\delta(t)$ the Dirac delta function [34].

We will focus our analysis on the mean first-passage time (MFPT) $T(x)$ for the particle to reach a given target L from the point x . It is well known [35], that the MFPT is infinite for a pure diffusive particle (or random search), while [3] showed that in the presence of resetting, it becomes finite. The standard starting point for MFPT analysis is the backward Chapman-Kolmogorov equation, which governs the probability $Q(x, t)$ that the particle has not reached the desired target up to time t , starting from position x . For $x \in \mathbb{R}$, the equation reads

$$\begin{aligned} \frac{\partial Q(x, t)}{\partial t} &= D \frac{\partial^2 Q(x, t)}{\partial x^2} - V'(x) \frac{\partial Q(x, t)}{\partial x} - r(x) Q(x, t) \\ &\quad + r(x) Q(x_0, t), \end{aligned} \quad (3)$$

with boundary conditions $Q(L, t) = 0$, $Q(x, 0) = 1$, where x_0 is the initial point, as well as the point where the process restarts. Noting that $-\frac{\partial Q(x, t)}{\partial t} dt$ measures the probability that the particle hits the target in the time interval $[t, t + dt]$, we find the MFPT by integrating (3) and using that

$$T(x) = - \int_0^\infty t \frac{\partial Q(x, t)}{\partial t} dt = \int_0^\infty Q(x, t) dt$$

to obtain the equation

$$-1 = D \frac{\partial^2 T(x)}{\partial x^2} - V'(x) \frac{\partial T(x)}{\partial x} - r(x) T(x) + r(x) T(x_0) \quad (4)$$

for the MFPT, with boundary conditions $T(L) = 0$ and $T(x)$ remains finite as $x \rightarrow \infty$ [4, 11].

Ahmad *et al.* [12] study the situation where the potential is a polynomial, $V(x) = kx^n$ with $k, n > 0$, and where the reset rate is constant. (See also [14], for the linear case, and [15] for the case of a logarithmic potential, $V(x) = V_0 \ln(|x|)$. A piecewise linear potential was studied in [17].) We will expand on this in two ways. Firstly, instead of a constant reset rate, we will consider having a piecewise-constant reset rate, i.e., where the reset rate is allowed to change when we

move around in space. Secondly, rather than smooth polynomial potentials [12] or piecewise linear potentials [17], we will consider piecewise polynomial potentials, in general, and piecewise quadratic potentials, in particular. The motivation for studying the piecewise quadratic potential comes from modern machine learning applications, where this is the predominant structure of the objective function [33]. The inspiration for studying the effect of having a piecewise-constant reset rate comes from different types of informed reset strategies, such as gradient-informed reset [36] which have been shown to be beneficial in many situations.

Let us review some of the previous results from [12]. For brevity, and as previously mentioned, we assume that the particle will reset from the initial position $x = x_0$. In general, Eq. (4) is very difficult to solve, but when $r(x)$ and $V(x)$ take some simple forms, we can solve it analytically. When we have a constant reset rate $r(x) = r$ and polynomial potential $V(x) = kx^n$ we get the equation

$$-1 = D \frac{\partial^2 T(x)}{\partial x^2} - knx^{n-1} \frac{\partial T(x)}{\partial x} - rT(x) + rT_0, \quad (5)$$

where we introduced the notation $T_0 := T(x_0)$. Analytical solutions do not exist for $n > 3$, so we focus on the case of $n = 2$, which was also studied in [12]. The general solutions are now

$$T(x) = AH_{-\frac{r}{2k}} \left(\sqrt{\frac{k}{D}} x \right) + BF_1 \left[\frac{r}{4k}, \frac{1}{2}; \frac{k}{D} x^2 \right] + \frac{1}{r} + T_0, \quad (6)$$

where $H_n(x)$ are the Hermite polynomials, ${}_1F_1[a, b; x]$ is the Kummer confluent hypergeometric function and A, B are scalars to be determined. The boundary conditions constrain the solution (6) to be finite for $x \rightarrow \infty$ and $T(L) = 0$, for some target $x = L$. These conditions determine A and B , and, after some algebraic manipulations, we find the solution of the MFPT for the smooth quadratic potential using a constant reset rate:

$$T_0 = \frac{1}{r} \left(\frac{H_{-\frac{r}{2k}} \left(\sqrt{\frac{k}{D}} L \right)}{H_{-\frac{r}{2k}} \left(\sqrt{\frac{k}{D}} x_0 \right)} - 1 \right). \quad (7)$$

This agrees with [12]. When the potential is weak, i.e., for low signal-to-noise ratio represented by a small initial drift compared to the noise level, one can find that there is an optimal nonzero reset rate, cf. [12], which means that it is beneficial to use resetting for this scenario. As can be seen in Fig. 1 on page 3, there are nonzero local minima of the MFPT for a large-enough value of D . In [12], the phase transition is found numerically and expressed in terms of the dimensionless variable $K := |V'(x_0)|/\sqrt{4D} = kx_0/\sqrt{D}$, whose critical value is $K_c \approx 0.7393$. That is, an optimal reset rate exists only when $K < K_c$. This is verified in Fig. 1, which plots the MFPT, T_0 in (7), for selected values of D ranging from 20 to 40 when $k = 1$, $x_0 = 4$, $L = 0.01$: The minimum disappears when D goes below 30, which is consistent with the value of K_c above since, for $k = 1$ and $x_0 = 4$, the critical value corresponds to $D \approx 29.3$.

We end the section by noting that here, as in later sections, it would also make sense to consider only the case $L = 0$, as

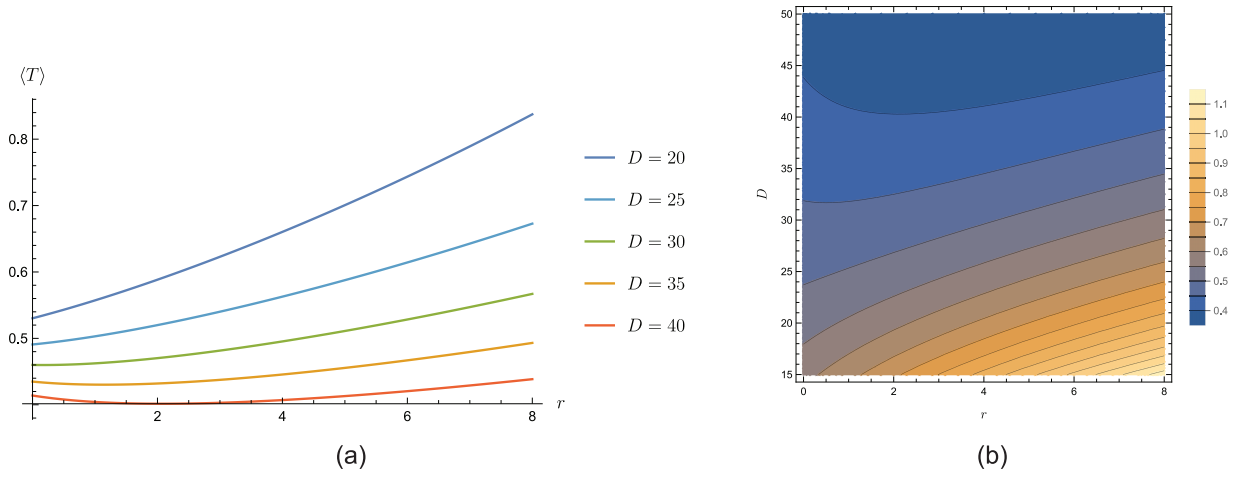


FIG. 1. (1d smooth potential/constant reset/varying D) MFPT, $\langle T \rangle := T_0$, of reaching a target $x = L$ for the quadratic potential $V(x) = x^2$, computed using Eq. (7). Here we have fixed $L = 0.01$ and $x_0 = 4$. (a) MFPT as a function of the constant reset rate r , for values of $D \in \{20, 25, 30, 35, 40\}$. (b) Contour plot of MFPT as a function of r and D . All numerical evaluations were done using Mathematica [37].

this is the global minimum, with the advantage that the expression for T_0 is further simplified. However, to keep expressions general, we choose L small compared to x_0 , but nonzero.

III. THEORETICAL RESULTS

Let us now turn to our analysis. We start by considering the smooth quadratic potential as above, but with a piecewise-constant reset rate. After this we consider the situations where we have a nonsmooth, piecewise quadratic potential both with a constant and with a piecewise-constant reset rate.

A. Smooth quadratic potential: Piecewise-constant reset rate

As discussed in the introduction, it is natural to consider having a reset rate that depends on the space time coordinate x . Some simple situations have been studied in the literature; see, for example, [4,10]. We will consider a version of this in which $r(x)$ changes between two different constant rates depending on whether the value of some function $f(x)$ is

within a threshold. In particular, we consider

$$r(x) = \begin{cases} r_1, & |f(x)| > \beta, \\ r_2, & |f(x)| \leq \beta, \end{cases} \quad (8)$$

for some function $f(x)$ and some constant $\beta > 0$. We will restrict ourselves to the case where the condition is governed by the magnitude of the gradient (such as $f(x) = V'(x) = 2kx$ in the smooth case) or a suitable generalization (to the nonsmooth case). This means that we now have two copies of Eq. (5), one copy for the case with r_1 and one with r_2 . The general solutions will also look the same in each window. As extra boundary conditions, we should now also demand that both $T(x)$ and $T'(x)$ be continuous on the boundary $|f(x)| = \beta$. The solution will depend on whether we set the target $x = L$ and the reset point $x = x_0$ below or above the threshold $|x| = \frac{\beta}{2k}$. We focus on the case where L is close to the origin, so we assume $L < \frac{\beta}{2k}$, and where the resetting is further away, so that $x_0 > \frac{\beta}{2k}$. After some algebraic manipulations, we find the solution of the MFPT for the smooth quadratic potential using a piecewise-constant reset rate:

$$T_0 = \frac{1}{r_1 r_2 H_{-\frac{r_1}{2k}}\left(\sqrt{\frac{k}{D}}x_0\right)\left(2\sqrt{kD}H_{-1-\frac{r_2}{2k}}\left(\frac{\beta}{2\sqrt{kD}}\right)_1F_1\left[\frac{r_2}{4k}, \frac{1}{2}; \frac{\beta^2}{4kD}\right] + \beta H_{-\frac{r_2}{2k}}\left(\frac{\beta}{2\sqrt{kD}}\right)_1F_1\left[1 + \frac{r_2}{4k}, \frac{3}{2}; \frac{\beta^2}{4kD}\right]\right)} \\ \times \left[2r_1\sqrt{kD}H_{-1-\frac{r_1}{2k}}\left(\frac{\beta}{2\sqrt{kD}}\right)\left(H_{-\frac{r_2}{2k}}\left(\sqrt{\frac{k}{D}}L\right)_1F_1\left[\frac{r_2}{4k}, \frac{1}{2}; \frac{\beta^2}{4kD}\right]\right) - H_{-\frac{r_2}{2k}}\left(\frac{\beta}{2\sqrt{kD}}\right)_1F_1\left[\frac{r_2}{4k}, \frac{1}{2}; \frac{kL^2}{D}\right] \right. \\ \left. + r_2H_{-\frac{r_1}{2k}}\left(\frac{\beta}{2\sqrt{kD}}\right)\left(2\sqrt{kD}H_{-1-\frac{r_2}{2k}}\left(\frac{\beta}{2\sqrt{kD}}\right)_1F_1\left[\frac{r_2}{4k}, \frac{1}{2}; \frac{kL^2}{D}\right] + \beta H_{-\frac{r_2}{2k}}\left(\sqrt{\frac{k}{D}}L\right)_1F_1\left[1 + \frac{r_2}{4k}, \frac{3}{2}; \frac{\beta^2}{4kD}\right]\right) \right. \\ \left. + H_{-\frac{r_1}{2k}}\left(\sqrt{\frac{k}{D}}x_0\right)\left(2\sqrt{kD}H_{-1-\frac{r_2}{2k}}\left(\frac{\beta}{2\sqrt{kD}}\right)\left((r_1 - r_2)_1F_1\left[\frac{r_2}{4k}, \frac{1}{2}; \frac{kL^2}{D}\right] - r_{11}F_1\left[\frac{r_2}{4k}, \frac{1}{2}; \frac{\beta^2}{4kD}\right]\right) \right. \right. \\ \left. \left. + \beta\left((r_1 - r_2)H_{-\frac{r_2}{2k}}\left(\sqrt{\frac{k}{D}}L\right) - r_1H_{-\frac{r_2}{2k}}\left(\frac{\beta}{2\sqrt{kD}}\right)\right)_1F_1\left[1 + \frac{r_2}{4k}, \frac{3}{2}; \frac{\beta^2}{4kD}\right]\right) \right]. \quad (9)$$

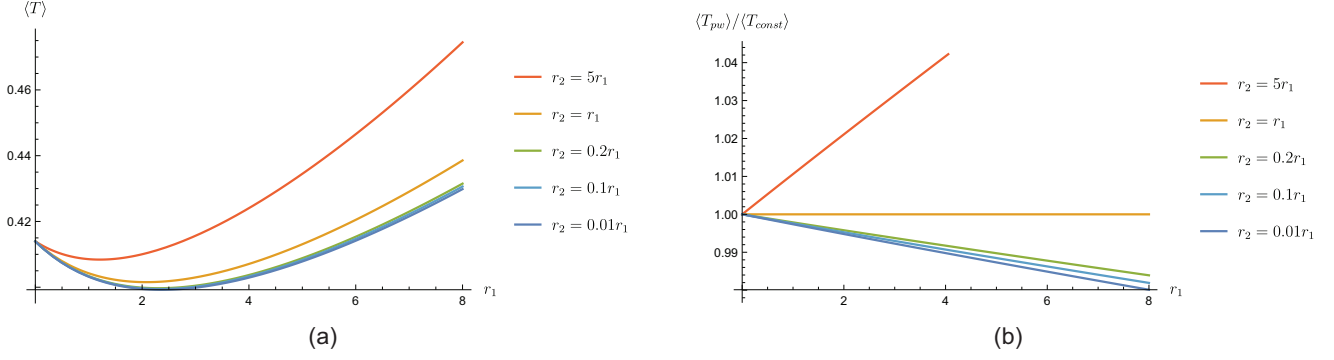


FIG. 2. (1d smooth potential/piecewise vs constant reset/ D fixed). Comparison of the MFPT, $\langle T \rangle := T_0$ in Eq. (9), between using a constant reset rate vs the piecewise-constant reset rate defined in Eq. (8), for the smooth quadratic potential $V(x) = x^2$. Here we have fixed $L = 0.01$, $x_0 = 4$, $D = 40$, $\beta = 1$, and r_2 is set as a fixed multiple of r_1 , i.e., the MFPT is plotted as a function of r_1 with $r_2 = cr_1$, for values of $c \in \{0.01, 0.1, 0.2, 1, 5\}$. (a) MFPT, labeled by c . (b) Ratio between the piecewise-constant reset rate, $\langle T_{pw} \rangle$, and the constant reset rate, $\langle T_{const} \rangle$, labeled by c . All numerical evaluations were done using Mathematica [37].

An interesting analysis is to compare the results for our two cases, i.e., having one constant reset rate, as in Eq. (7), and having two different resets as in Eq. (9). This comparison is shown in Figs. 2 and 3. There, we fix r_2 to be a multiple of r_1 for the case where we have two resets and plot T_0 as a function of r_1 . We see that when $r_2 > r_1$ it is always better to just keep the constant reset rate, while when $r_2 < r_1$ it is better to switch the rate. This can be interpreted as saying that whenever we get closer to the target, it is better to not reset as often.

In Fig. 4, we further see that the benefit of resetting is dependent on the signal-to-noise ratio, in the sense that there is no optimal reset rate when the diffusion constant D is small enough, $D \lesssim 25$. We also observe that larger changes in MFPT behavior occur when switching from a constant to a piecewise-constant reset rate with $r_2 > r_1$ than when switching to a piecewise-constant reset rate with $r_2 < r_1$.

B. Nonsmooth piecewise quadratic potential: Constant and piecewise-constant reset rates

Let us now turn to the second case of interest to us, namely that of having a piecewise-polynomial potential. As a working

example, we will consider the piecewise-quadratic potential

$$V(x) = \begin{cases} x^2, & x \leq 2, \\ 2 + \frac{1}{2}(x-4)^2, & x > 2. \end{cases} \quad (10)$$

This has two minima, a global one at $x = 0$ and a local one at $x = 4$. See Fig. 5.

We now have the two equations

$$-1 = \begin{cases} D \frac{\partial^2 T(x)}{\partial x^2} - 2x \frac{\partial T(x)}{\partial x} - r(x)T(x) + r(x)T_0, & \text{for } x < 2, \\ D \frac{\partial^2 T(x)}{\partial x^2} - (x-4) \frac{\partial T(x)}{\partial x} - r(x)T(x) + r(x)T_0, & \text{for } x > 2. \end{cases}$$

Let us again look at the two cases separately, i.e., having a constant reset rate or having two different reset rates.

a. *Constant reset.* The general solutions are

$$T_1(x) = A_1 H_{-\frac{r}{2}}\left(\frac{x}{\sqrt{D}}\right) + B_1 {}_1F_1\left[\frac{r}{4}, \frac{1}{2}, \frac{x^2}{D}\right] + \frac{1}{r} + T_0, \quad x < 2 \quad (11)$$

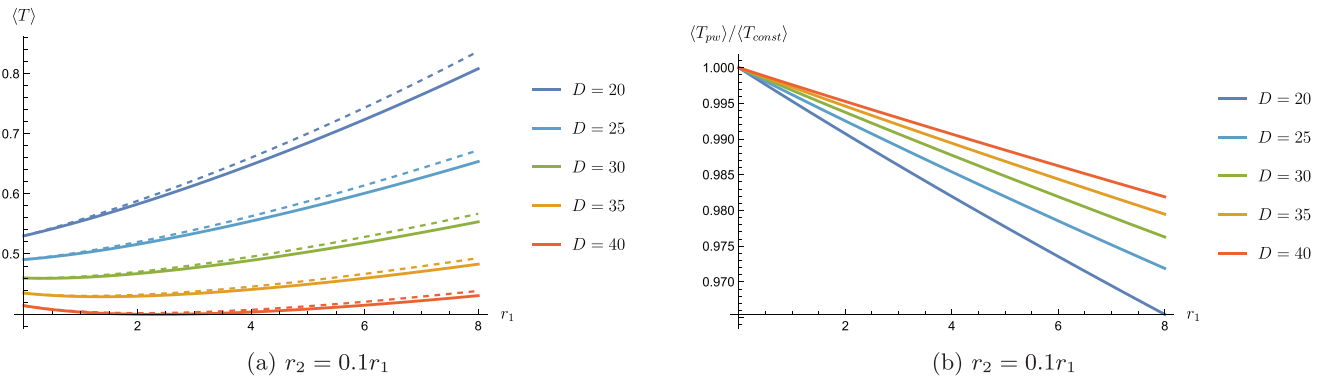


FIG. 3. (1d smooth potential/piecewise vs constant reset/varying D). Comparison of the MFPT, $\langle T \rangle := T_0$ in Eq. (9), between using a constant reset rate vs the piecewise-constant reset rate defined in Eq. (8) with $r_2 = 0.1r_1$, for the smooth quadratic potential $V(x) = x^2$, for values of $D \in \{20, 25, 30, 35, 40\}$. Here we have fixed $L = 0.01$, $x_0 = 4$, $\beta = 1$. (a) MFPT for constant reset rate ($\langle T_{const} \rangle$, dashed) and piecewise-constant reset rate ($\langle T_{pw} \rangle$, solid), labeled by D . (b) Ratio $\langle T_{pw} \rangle / \langle T_{const} \rangle$, labeled by D . All numerical evaluations were done using Mathematica [37].

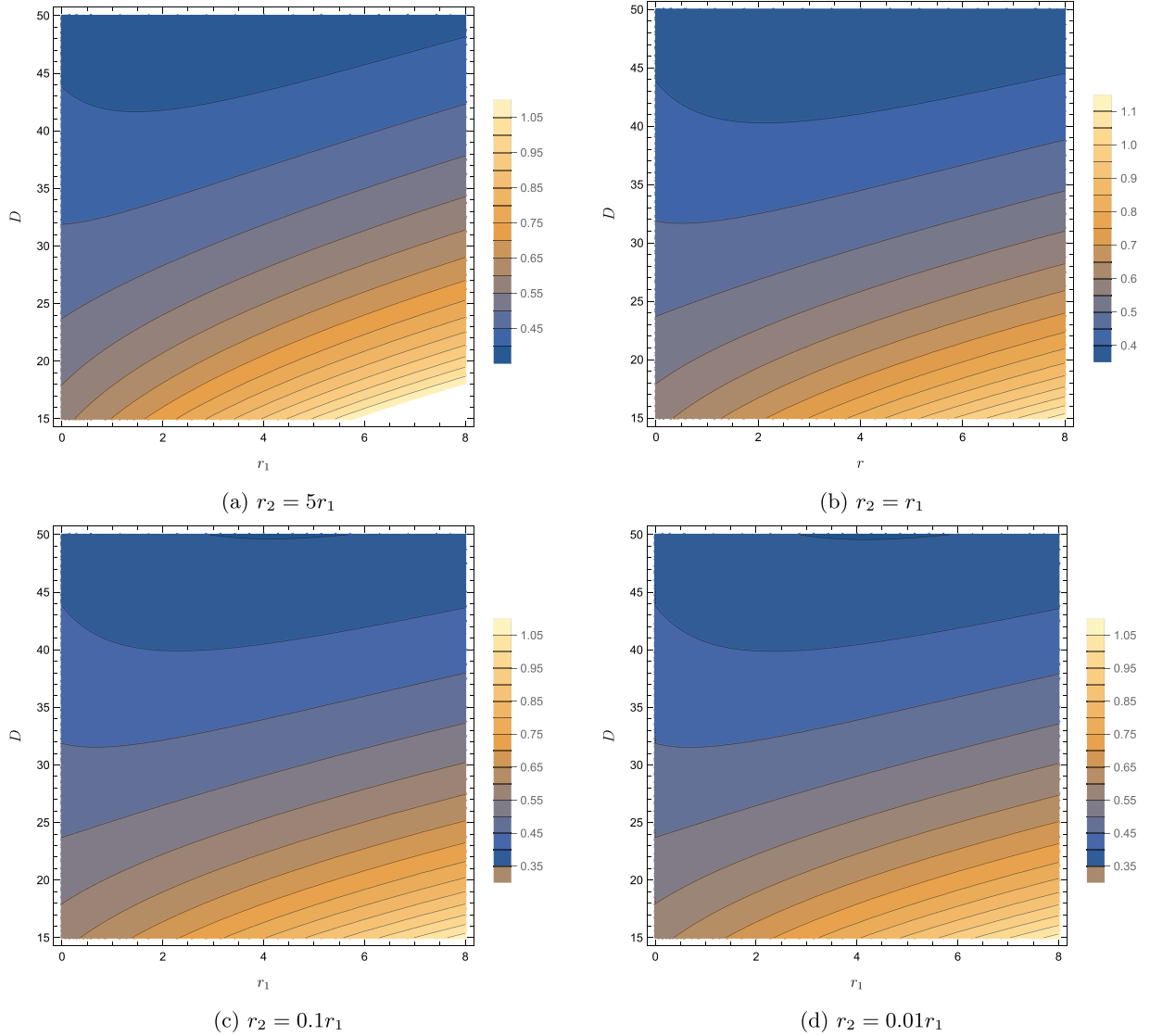


FIG. 4. (1d smooth potential/piecewise reset). Contour plots for the MFPT, T_0 in Eq. (9), as a function of the reset rate r_1 and diffusion coefficient D , for the smooth quadratic potential $V(x) = x^2$ and the piecewise-constant reset rate defined in Eq. (14). Four different values $c \in \{0.01, 0.1, 1, 5\}$ are considered in scaling the reset rate $r_2 = cr_1$. Here we have fixed $L = 0.01$, $x_0 = 4$, and $\beta = 1$. All numerical evaluations were done using Mathematica [37].

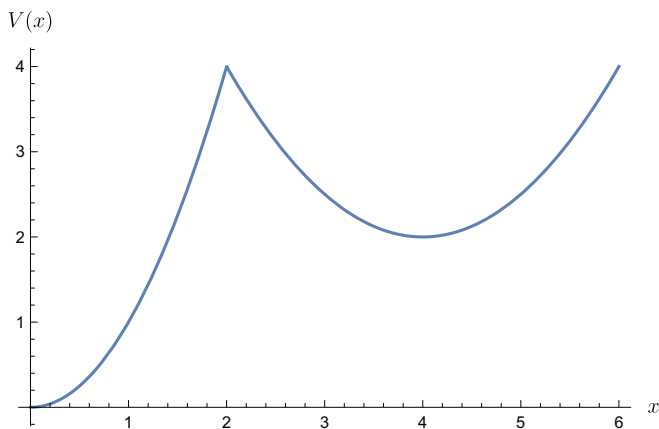


FIG. 5. The piecewise quadratic potential $V(x)$ studied in this paper. See Eq. (10) for the definition.

and

$$T_2(x) = A_2 H_{-r} \left(\frac{x-4}{\sqrt{2D}} \right) + B_2 {}_1F_1 \left[\frac{r}{2}, \frac{1}{2}, \frac{(x-4)^2}{2D} \right] + \frac{1}{r} + T_0, \quad (12)$$

$$x > 2,$$

where A_1, A_2, B_1 , and B_2 are scalar constants to be determined. The boundary conditions are that $T(x)$ should be finite for $x \rightarrow \infty$, $T(L) = 0$ and both $T(x)$ and $T'(x)$ should be continuous at the boundary of the two pieces $x = 2$. This fixes all the constants A_j, B_j and the final solution will again depend on whether we assume that L and x_0 are larger than or smaller than 2.

As before, we will focus on the case where $L < 2$ and $x_0 > 2$, i.e., where we start (and reset) in the valley of the local minimum, but the target is in the valley of the global one. After some algebraic manipulation, we find the solution of

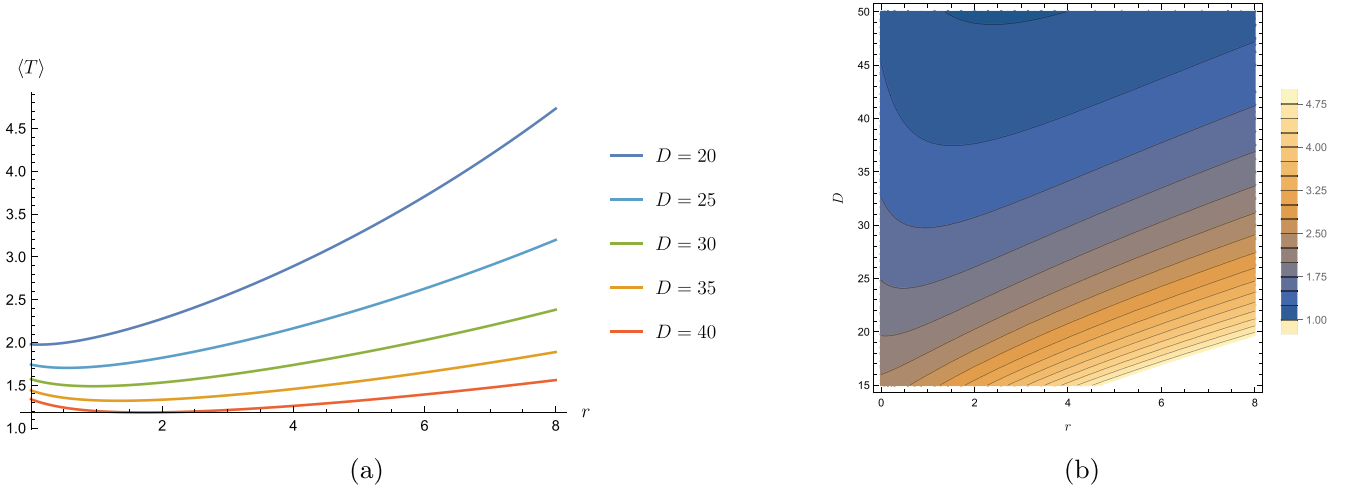


FIG. 6. (1d nonsmooth potential/constant reset/varying D) MFPT, $\langle T \rangle := T_0$, of reaching a target $x = L$ for the nonsmooth piecewise quadratic potential defined in Eq. (10), computed using Eq. (13). Here we have fixed $L = 0.01$ and $x_0 = 6$. (a) MFPT as a function of the constant reset rate r for values of $D \in \{20, 25, 30, 35, 40\}$. (b) Contour plot of MFPT as a function of r and D . All numerical evaluations were done using Mathematica [37].

the MFPT for the nonsmooth piecewise quadratic potential and constant reset rate:

$$\begin{aligned}
 T(x_0) = & \left(r H_{-r} \left(\frac{x_0 - 4}{\sqrt{2D}} \right) \left(2 H_{-\frac{r}{2}} \left(\frac{2}{\sqrt{D}} \right) {}_1F_1 \left[1 + \frac{r}{4}, \frac{3}{2}, \frac{4}{D} \right] + \sqrt{D} H_{-1-\frac{r}{2}} \left(\frac{2}{\sqrt{D}} \right) {}_1F_1 \left[\frac{r}{4}, \frac{1}{2}, \frac{4}{D} \right] \right) \right)^{-1} \\
 & \times \left(-H_{-r} \left(\frac{x_0 - 4}{\sqrt{2D}} \right) \left(2 H_{-\frac{r}{2}} \left(\frac{2}{\sqrt{D}} \right) {}_1F_1 \left[1 + \frac{r}{4}, \frac{3}{2}, \frac{4}{D} \right] + \sqrt{D} H_{-1-\frac{r}{2}} \left(\frac{2}{\sqrt{D}} \right) {}_1F_1 \left[\frac{r}{4}, \frac{1}{2}, \frac{4}{D} \right] \right) \right. \\
 & + H_{-r} \left(-\sqrt{\frac{2}{D}} \right) \left(2 H_{-\frac{r}{2}} \left(\frac{L}{\sqrt{D}} \right) {}_1F_1 \left[1 + \frac{r}{4}, \frac{3}{2}, \frac{4}{D} \right] + \sqrt{D} H_{-1-\frac{r}{2}} \left(\frac{2}{\sqrt{D}} \right) {}_1F_1 \left[\frac{r}{4}, \frac{1}{2}, \frac{L^2}{D} \right] \right) \\
 & \left. + \sqrt{2D} H_{-1-r} \left(-\sqrt{\frac{2}{D}} \right) \left(H_{-\frac{r}{2}} \left(\frac{L}{\sqrt{D}} \right) {}_1F_1 \left[\frac{r}{4}, \frac{1}{2}, \frac{4}{D} \right] - H_{-\frac{r}{2}} \left(\frac{2}{\sqrt{D}} \right) {}_1F_1 \left[\frac{r}{4}, \frac{1}{2}, \frac{L^2}{D} \right] \right) \right). \quad (13)
 \end{aligned}$$

This is plotted as a function of r in Fig. 6(a), for fixed $x_0 = 6$, $L = 0.01$, and for the same values of D considered in the smooth quadratic potential case. As before, we see that for a sufficiently large diffusion coefficient D , there is an optimal reset rate. It is noteworthy that the signal-to-noise ratio, defined by $|V'(x_0)|/\sqrt{4D}$ as before, is four times larger than in the smooth quadratic potential case, because $V'(x_0) = 2$ instead of 8. Despite this, the MFPT curves are similar to those in the smooth case, because the algorithm needs to overcome the local optimum valley centered at $x = 4$ in order to reach the global optimum at $x = 0$. It is clear that a higher noise level is needed to reach the global minimum in the nonsmooth case, compared to the smooth case. In Fig. 6(b), we also see that the benefits of resetting disappear when the diffusion coefficient is small enough, $D \lesssim 20$.

b. Piecewise-constant reset rate. We can again consider the case where we switch between two different reset rates r_1 and r_2 as in Eq. (8). There are now many options for the condition function $f(x)$. As before, one natural condition is that it depends on the gradient of the potential, i.e.,

$$f(x) = V'(x) = \begin{cases} 2x, & x \leq 2, \\ x - 4, & x > 2. \end{cases}$$

This would mean that

$$r(x) = \begin{cases} r_1, & |x| > \frac{\beta}{2}, x \leq 2, \\ r_2, & |x| \leq \frac{\beta}{2}, x \leq 2, \\ r_1, & |x - 4| > \beta, x > 2, \\ r_2, & |x - 4| \leq \beta, x > 2, \end{cases} \quad (14)$$

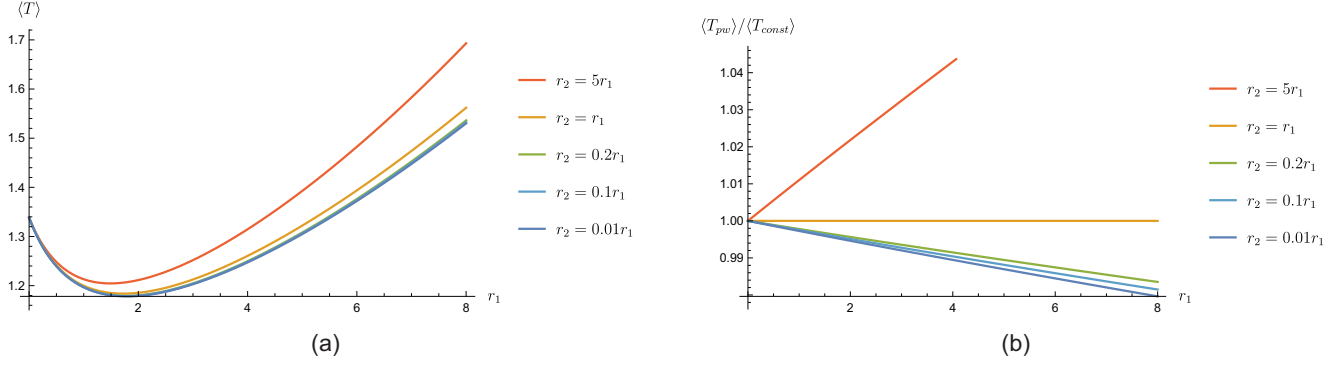


FIG. 7. (1d nonsmooth potential/piecewise reset/ D fixed). Comparison of the MFPT, $\langle T \rangle := T_0$ in Eq. (A1), between using a constant reset rate and the piecewise-constant reset rate defined in Eq. (8), for the nonsmooth piecewise-quadratic potential defined in Eq. (10). Here we have fixed $L = 0.01$, $x_0 = 6$, $\beta = 1$, and r_2 is set as a fixed multiple of r_1 , i.e., the MFPT is plotted as a function of r_1 with $r_2 = cr_1$, for values of $c \in \{0.01, 0.1, 0.2, 1, 5\}$. (a) MFPT, labeled by c . (b) Ratio between the piecewise-constant reset rate, $\langle T_{pw} \rangle$, and the constant reset rate, $\langle T_{const} \rangle$, labeled by c . All numerical evaluations were done using Mathematica [37].

and we have four different solutions to the differential equations to consider, namely Eqs. (11) and (12), each with the two different reset rates:

$$T(x) = \begin{cases} A_1 H_{-\frac{r_1}{2}}\left(\frac{x}{\sqrt{D}}\right) + B_1 {}_1F_1\left[\frac{r_1}{4}, \frac{1}{2}, \frac{x^2}{D}\right] + \frac{1}{r_1} + T_0, & x \leq 2, |x| > \beta/2, \\ A_2 H_{-\frac{r_2}{2}}\left(\frac{x}{\sqrt{D}}\right) + B_2 {}_1F_1\left[\frac{r_2}{4}, \frac{1}{2}, \frac{x^2}{D}\right] + \frac{1}{r_2} + T_0, & x \leq 2, |x| \leq \beta/2, \\ A_3 H_{-r_1}\left(\frac{x-4}{\sqrt{2D}}\right) + B_3 {}_1F_1\left[\frac{r_1}{2}, \frac{1}{2}, \frac{(x-4)^2}{2D}\right] + \frac{1}{r_1} + T_0, & x > 2, |x-4| > \beta, \\ A_4 H_{-r_2}\left(\frac{x-4}{\sqrt{2D}}\right) + B_4 {}_1F_1\left[\frac{r_2}{2}, \frac{1}{2}, \frac{(x-4)^2}{2D}\right] + \frac{1}{r_2} + T_0, & x > 2, |x-4| \leq \beta. \end{cases} \quad (15)$$

The boundary conditions are the same as before, finiteness at infinity, zero at the target, and continuous over the various break points. The final solution is rather tedious, so we produce the full expression only in Appendix A. It does depend on assumptions on the values of L , β , and x_0 , again. The general behavior is similar to before, and we can compare this solution for various ratios between the reset rates. The results

are equivalent to what we found in the smooth case, and are shown in Figs. 7–9. We fix $L = 0.01$, $x_0 = 6$, $\beta = 1$ and consider $r_2 = cr_1$. In Fig. 7, we further fix $D = 40$ and plot the MFPT as a function of r_1 for different ratios $r_2/r_1 = c$. As in the smooth case, we see that it is beneficial to use the piecewise-resetting whenever $r_2 < r_1$. In Fig. 8, we vary diffusion coefficient D and fix $r_2 = 0.1r_1$. Here, we see that

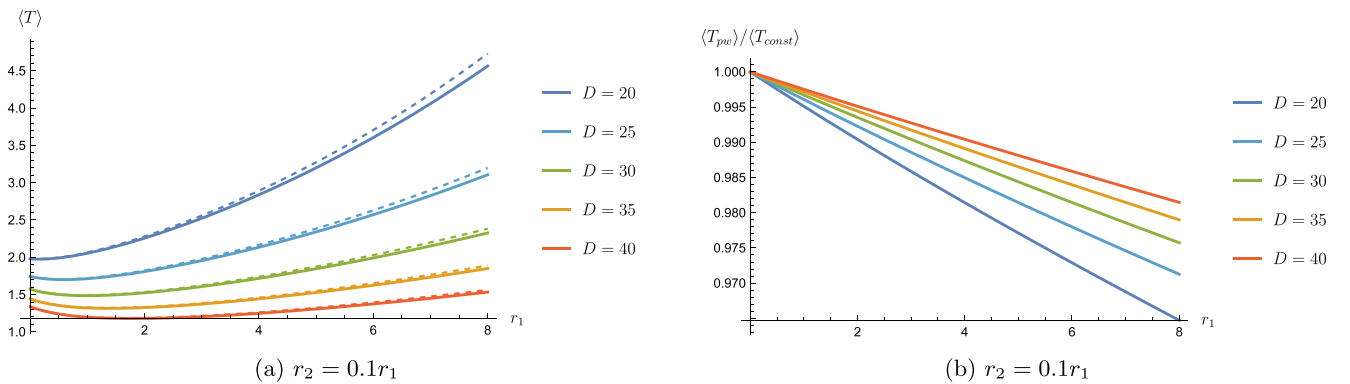


FIG. 8. (1d nonsmooth potential/piecewise reset/varying D). Comparison of the MFPT, $\langle T \rangle := T_0$ in Eq. (A1), between using a constant reset rate and the piecewise-constant reset rate defined in Eq. (8) with $r_2 = 0.1r_1$, for the nonsmooth piecewise-quadratic potential defined in Eq. (10), for values of $D \in \{20, 25, 30, 35, 40\}$. Here we have fixed $L = 0.01$, $x_0 = 6$, $\beta = 1$. (a) MFPT for constant reset rate ($\langle T_{const} \rangle$, dashed) and piecewise-constant reset rate ($\langle T_{pw} \rangle$, solid), labeled by D . (b) Ratio $\langle T_{pw} \rangle / \langle T_{const} \rangle$, labeled by D . All numerical evaluations were done using Mathematica [37].

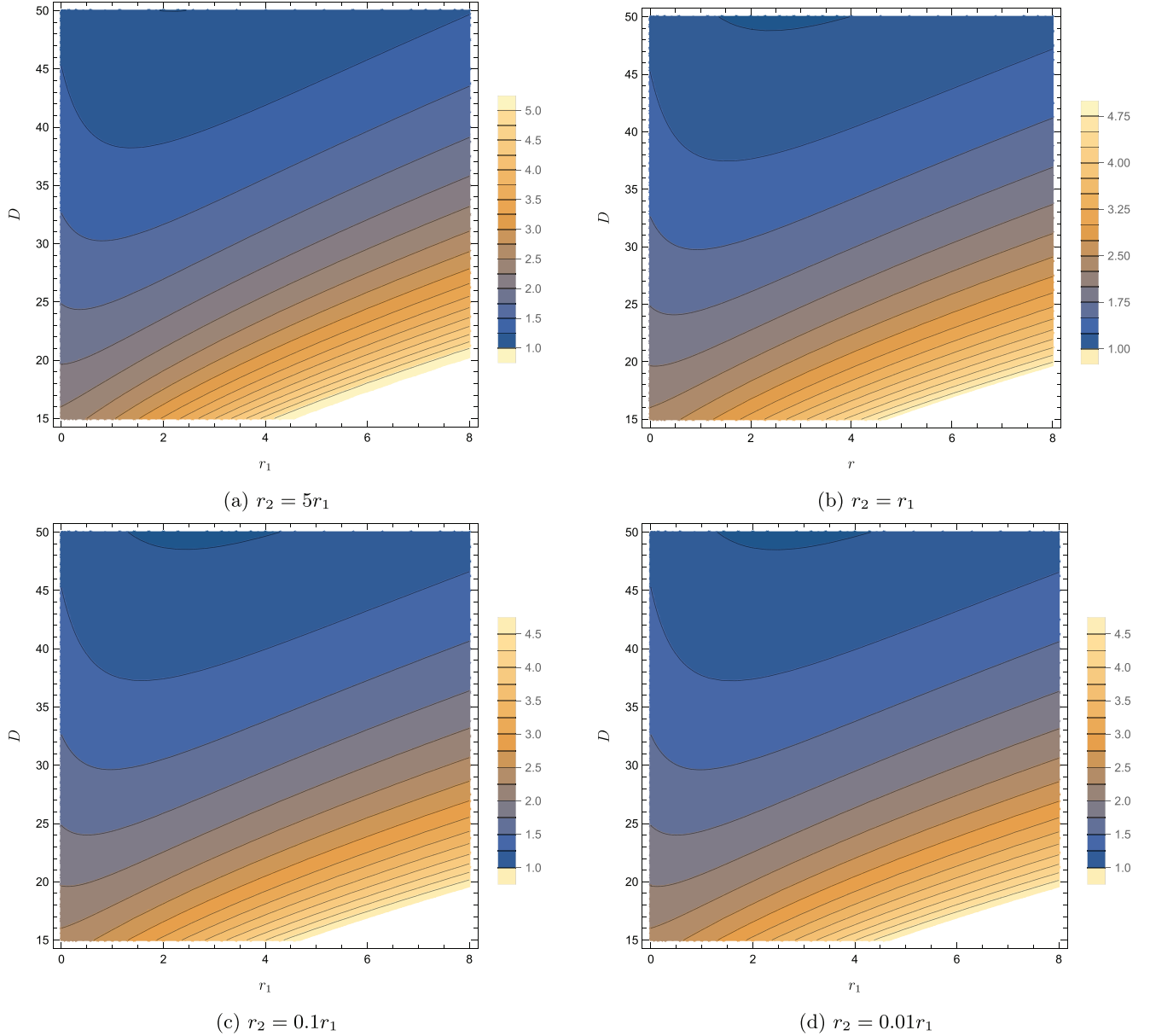


FIG. 9. (1d nonsmooth potential/piecewise reset). Contour plot for the MFPT, T_0 in Eq. (A1), as a function of the reset rate r_1 and diffusion coefficient D , for the nonsmooth piecewise quadratic potential defined in Eq. (10) and the piecewise-constant reset rate defined in Eq. (14). Four different $r_2 = cr_1$ values are considered, for selected values of $c \in \{0.01, 0.1, 1, 5\}$. Here we have fixed $L = 0.01$, $x_0 = 6$ and $\beta = 1$. All numerical evaluations were done using Mathematica [37].

the benefits of using the piecewise-resetting are larger when diffusion coefficient D is smaller. Finally, in Fig. 9, we plot the MFPT as a function of both the resetting rate r_1 and the diffusion coefficient D . We can see that, for all r_2/r_1 ratios considered, the optimal reset strategy disappears when the diffusion coefficient goes below a certain threshold, around $D \lesssim 20$. As in the 1d smooth potential, we observe that larger changes in MFPT behavior occur when switching from a constant to a piecewise-constant reset rate with $r_2 > r_1$ than when switching to a piecewise-constant reset rate with $r_2 < r_1$.

We further note that, although the MFPT values are considerably lower for the smooth quadratic potential compared to the nonsmooth piecewise quadratic potential, the *relative* impact on the MFPT of using a piecewise-constant reset rate

over a constant reset rate is strikingly similar across the two potentials, as becomes apparent by comparing Fig. 2(b) with Fig. 7(b) and Fig. 3(b) with Fig. 8(b).

c. Higher dimensions. It is of interest to consider how these questions generalize to higher dimensions. We will not explore this at length here, however, it is worthwhile to discuss some of the possible generalizations. We leave it for future research to explore this further.

When working in $\mathbb{R}^d$, the MFPT $T(x, L)$ of reaching a target $x = L \in \mathbb{R}^d$ will satisfy the following diffusion equation [3–5,12]:

$$-1 = D\nabla^2 T(x, L) - \nabla V(x) \cdot \nabla T(x, L) - r(x)T(x, L) + r(x)T(x_0, L),$$

with boundary condition $T(L, L) = 0$.

We again restrict our analysis to quadratic potentials of the form generalizing Eq. (10):

$$V(x) = a + b(\|x\| - c)^2, \quad (16)$$

for some constants a, b, c , or rather, more generally, piecewise quadratic potentials where each piece is given by (16) for different values of the constants, such that it still is continuous, but not smooth. Note that this is highly symmetric and, in particular, depends only on the radial coordinate $R = \|x\|$. The diffusion equation for the MFPT then simplifies to

$$\begin{aligned} -1 = D \left(\partial_R^2 T(R) + \frac{d-1}{R} \partial_R T(R) \right) - 2b(R-c) \partial_R T(R) \\ - r(R)T(R) + rT_0, \end{aligned} \quad (17)$$

where we introduced the notation $T_0 := T(R_0, L)$ as before and further removed the dependence on L in the notation for T , i.e., defined $T(R) := T(R, L)$. In the following, we will also often drop the dependence on R and simply write $T = T(R)$.

In general, Eq. (17) is, of course, very difficult to solve. When $r(x) = r$ we can define the variables $w(R) = T(R) - T_0 - 1/r$ and $y = \sqrt{\frac{b}{D}}R$ to write Eq. (17) on the form

$$\begin{aligned} yw'' + (1 + \alpha - \beta y - 2y^2)w' + ((\gamma - 2 - \alpha)y \\ - \frac{1}{2}(\delta + \beta(1 + \alpha)))w = 0, \end{aligned}$$

with

$$\begin{aligned} \alpha &:= d - 2, \\ \beta &:= -2c\sqrt{\frac{b}{D}}, \\ \gamma &:= d - \frac{r}{b}, \\ \delta &:= 2c\sqrt{\frac{b}{D}}(d - 1). \end{aligned}$$

This is the biconfluent Heun equation [38]. Unfortunately, it seems that the solutions to this equation when $\alpha \in \mathbb{Z}$ have not been studied in much detail. For example, the general solutions used in much of the literature [38], do not work in this situation.

However, if we consider a smooth quadratic potential around the origin, $V(R) = kR^2, k > 0$, the situation simplifies. For a constant resetting rate, we can solve this analytically for any dimension. The procedure is the same as before using the boundary conditions together with some known relations for hypergeometric functions, we end up with

$$T(R_0) = \frac{1}{r} \left(\frac{U\left[\frac{r}{4k}, \frac{d}{2}, \frac{kL^2}{D}\right]}{U\left[\frac{r}{4k}, \frac{d}{2}, \frac{kR_0^2}{D}\right]} - 1 \right), \quad (18)$$

where $U[\cdot, \cdot; \cdot]$ is Tricomi's confluent hypergeometric function [39].

When we consider the piecewise-constant reset rate, the solution becomes more complicated. We can simplify the analysis by focusing on a fixed dimension. Let us concentrate on $d = 2$. We again introduce a condition for the switch

between rates. The natural generalization of our previous condition is

$$r(x) = \begin{cases} r_1, & \|\nabla V(x)\| > \beta, \\ r_2, & \|\nabla V(x)\| \leq \beta. \end{cases} \quad (19)$$

The solution for the MFPT is again quite messy, so we produce the full expression only in Appendix B. There are some seemingly singular points, for example for $r_2 = r_1 = 4$, this is however only an artifact of the particular solution techniques used and would not be a reflection of the actual MFPT. The results are similar as in the one-dimensional case, e.g., we again see that the benefit of using multiple reset rates is present when $r_2 < r_1$. This is highlighted in Fig. 10. We leave a more in depth analysis of the multidimensional case for future research.

IV. NUMERICAL EXPERIMENTS

In order to corroborate the theoretical results obtained in the previous section, we consider a time-discretized version of the diffusion equation (1) using the Euler-Maruyama approximation [40], which is used to run the numerical experiments. The Euler-Maruyama approximation gives rise to a Markov process in $\mathbb{R}^d$, $\{x_k := x(k\Delta t)\}_{k \in \mathbb{N}}$ which, given the start value x_0 , is defined by

$$x_{k+1} = x_k + \mu F(x_k)\Delta t + \sqrt{2D\Delta t}z_k, \quad (20)$$

where Δt is the step size resulting from a positive time discretization interval, and $\{z_k\}_{k \in \mathbb{N}}$ is an uncorrelated standard Gaussian process in $\mathbb{R}^d$, independent of the $\{x_k\}$ process. As before, $F(x) = -\nabla V(x)$ where $V(x)$ is some potential in $\mathbb{R}^d$, and we set $\mu = 1$.

Each experiment consists of using Eq. (20) to generate the iterates x_k , except when a resetting event occurs, in which case x_k is reset to x_0 . In addition, when the target L is reached, the value $k\Delta t$ is recorded as a first-passage time and the pair (k, x_k) is reset to $(0, x_0)$. If $r(x)$ is the reset rate as a function of the position, the reset event is defined at each iteration k as a Bernoulli variable with event probability $r(x_k)\Delta t$.

We consider different hyperparameter settings in order to analyze their impact on the MFPT. Each setting defines an experiment, which is run five times for up to one million iterations each. The same five seeds are used across experiments so that any observed differences are only due to the different hyperparameter values, not to different seeds. The MFPT of each experiment is estimated as the average of all recorded first-passage times by the target, across all replications, giving the following expression for the estimate of the MFPT:

$$\widehat{\text{MFPT}} = \frac{\sum_{i=1}^I \sum_{n=1}^{N_i} \tau_{i,n}}{\sum_{i=1}^I N_i},$$

where N_i is the number of first-passage times after reset for replication i of I , and $\tau_{i,n}$ is the n -th first-passage time in the i -th replication, computed as the number of iterations to reach the target since reset, multiplied by Δt . The event of reaching the target is triggered by the condition that the absolute difference between the observed potential at time step k , $V(x_k)$, and the potential value at the target, $V(L)$, is less than a tolerance defined separately for each case analyzed below.

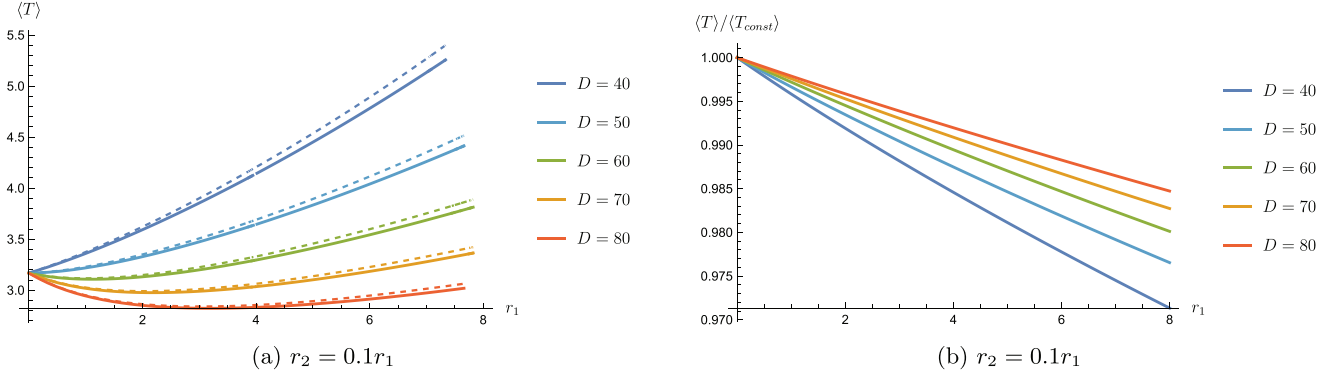


FIG. 10. (2d smooth potential/piecewise reset/varying D). Comparison of the MFPT, $\langle T \rangle =: T_0$ of Eq. (B1), between using a constant reset rate and the piecewise-constant reset rate defined in Eq. (19) with $r_2 = 0.1r_1$, for a two-dimensional smooth quadratic potential $V(R) = R^2$, for values of $D \in \{40, 50, 60, 70, 80\}$. Here we have fixed $L = 0.01$, $R_0 = 4\sqrt{2}$, $\beta = 1$. (a) MFPT for constant reset rate ($\langle T_{const} \rangle$, dashed) and piecewise-constant reset rate ($\langle T_{pw} \rangle$, solid), labeled by D . (b) Ratio $\langle T_{pw} \rangle / \langle T_{const} \rangle$, labeled by D . All numerical evaluations were done using Mathematica [37].

As potential $V(x)$, we consider the functions analyzed in the theoretical analysis of Sec. III: the smooth quadratic potential in one and two dimensions, and the 1d nonsmooth piecewise quadratic potential with its 2d analog found by revolving the 1d potential around the third axis, to obtain a symmetric piecewise parabolic function. In all cases, the step size is set to $\Delta t = 10^{-4}$ and the target is set to the position of the global optimum of the potential, i.e., $L = 0$. The reset rate function $r(x)$ is defined by the piecewise-constant function (19) for all cases, parameter $\beta = 1$, and parameters $r_1 \neq r_2$. Note that the constant reset rate strategy uses $r_1 = r_2$. For the purpose of comparing results with those of Sec. III, the effect on the MFPT with four different reset rates, $r_1 = 0.1, 1, 5, 10$, three reset rate ratios $r_2/r_1 = 0.1, 1, 10$, and five diffusion coefficients D (to be defined in each case) is analyzed. The value of x_0 depends on the potential and is defined in each subsection below.

A. 1d experiments

With the purpose of verifying the theoretical results of Figs. 3 and 8 in Sec. III, as well as Figs. 13 and 14, the same five diffusion coefficients D and reset position x_0 used therein were considered in the 1d experiments. This means $D = [20, 25, 30, 35, 40]$, $x_0 = 4$ for the smooth potential and $x_0 = 6$ for the nonsmooth potential. The only difference with the analytic figures is that the target L is set to 0 instead of 0.01.

Only two parameters belong exclusively to the numerical experiments context: the step size Δt (already fixed at 10^{-4}), and the tolerance ΔV_{tol} , which is set to 0.01 for the 1d potentials.

With this parameter setup, the discrete-time version of the signal-to-noise ratio, defined in Sec. II as $|V'(x_0)|/\sqrt{4D}$, becomes $|V'(x_0)|\sqrt{\Delta t}/\sqrt{4D}$ and varies from a lowest value of ≈ 0.006 at $D = 40$ to a highest value of ≈ 0.009 at $D = 20$. We note that the discrete-time signal-to-noise ratio decreases as Δt decreases, for a fixed value of D .

B. 2d experiments

In order to compare the 2d experiments with the theoretical results of Sec. III, we consider the equivalent reset points

of $x_0 = (4, 4)$ for the smooth quadratic potential case. For the nonsmooth piecewise quadratic case we pick the point $x_0 = (6, 6)$. We further enforce the same initial signal-to-noise ratio as in the respective 1d counterparts. Out of the two parameters exclusive to the numerical experiments, Δt and ΔV_{tol} , Δt is kept equal to 10^{-4} as in the 1d case, but ΔV_{tol} is increased tenfold to 0.1, considering the 2d excursions of the x_k iterates have more wandering space available, which increases the difficulty of finding the target compared to the 1d case.

The condition on using the same signal-to-noise ratio as in the 1d experiments is enforced through the selected values of D in these 2d experiments. This is done with the help of Table I, which shows the expressions for the signal-to-noise ratio (SNR) for all four potentials considered, computed at their respective reset point. The D values to use for the 2d problems, given the D values of the respective 1d problems, is obtained by equating the two columns in each of the two rows of the table.

Figure 16 in Appendix D shows an example of a diffusion trajectory on the 2d nonsmooth piecewise quadratic potential.

C. Analysis of results

Figures 11 and 12 present the results in terms of the MFPT as a function of the reset rate r_1 with curves labeled

TABLE I. Signal-to-noise ratio ($\text{SNR} := \frac{\| \nabla V(x_0) \|}{\sqrt{4D}}$) for the four different potentials $V(x)$ considered in the discrete-time diffusion process defined in Eq. (20), with $F(x) := -\nabla V(x)$ and x_0 the reset point.

SNR	1d	2d
smooth quadratic (defined for 1d case in Sec. II with $k = 1$)	$x_0 \sqrt{\frac{\Delta t}{D}}, x_0 = 4$	$\ x_0\ \sqrt{\frac{\Delta t}{D}},$ $x_0 = (4, 4)$
piecewise quadratic [defined for 1d case in Eq. (10)]	$(x_0 - 4) \sqrt{\frac{\Delta t}{4D}}, x_0 = 6$	$(\ x_0\ - 4) \sqrt{\frac{\Delta t}{4D}},$ $x_0 = (6, 6)$

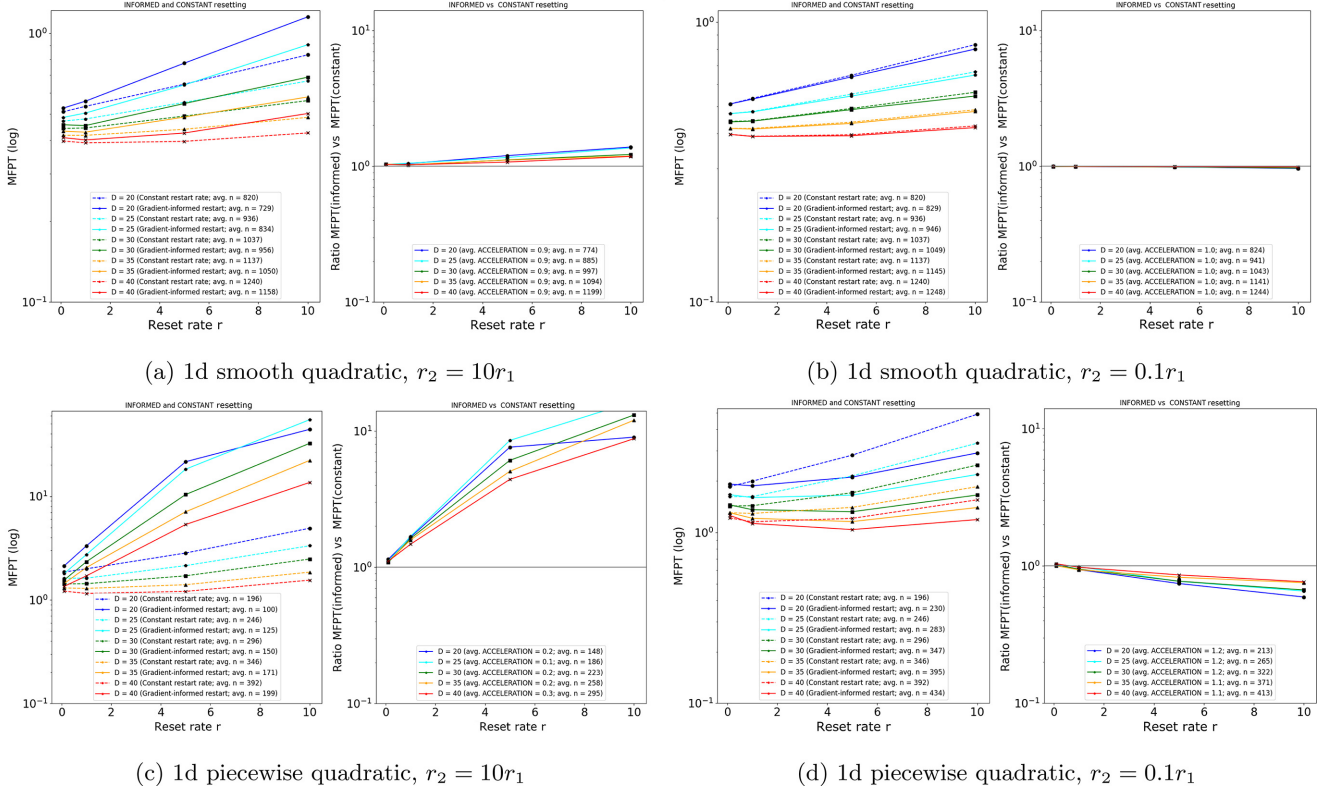


FIG. 11. Average MFPT (measured in time units) obtained on two 1d quadratic potentials, $V(x)$, using both the constant reset rate strategy and the piecewise-constant reset rate defined in Eq. (8) by setting $f(x) = V'(x)$. (a) and (b): 1d smooth quadratic potential, $V(x) = x^2$. (c) and (d): 1d nonsmooth piecewise quadratic potential, with $V(x)$ defined in Eq. (10). Reset rate r_2 is used when the gradient norm is smaller than the threshold $\beta = 1$, and otherwise reset rate r_1 is used. For visual comparison, all right plots across panels, showing the ratio between using a piecewise-constant reset rate vs a constant reset rate, have the same vertical scales.

by D , respectively for the 1d and the 2d experiments. Each set of results is arranged in a 2×2 panel where the top and bottom rows correspond, respectively, to the smooth and nonsmooth potentials, and the left and right columns correspond, respectively, to the reset rate ratios $r_2/r_1 = 10$ and 0.1 . The MFPT values obtained for the constant reset rate strategy are plotted as dashed curves in each of the four cells.

Following are the main observations that we can infer from the MFPT plots obtained for the numerical experiments:

(1) 1d numerical experiments confirm the following aspects of the theoretical results:

(a) The MFPT tends to decrease as the diffusion coefficient D increases, for a given reset rate.

(b) An optimum reset rate exists as long as the diffusion coefficient D is large enough, i.e., for lower signal-to-noise ratios computed at the reset position.

(c) The curvature of the curve increases as the diffusion coefficient D increases from 20 to 40, i.e., from blue to red curves.

(d) The regime $r_2 < r_1$ is better than $r_2 > r_1$ in terms of obtaining a smaller MFPT than in the constant reset rate strategy $r_2 = r_1$, for a given diffusion coefficient. This is verified, for $D = 40$, in Figs. 11(a) and 11(b), which confirm the theoretical behavior plotted in Fig. 2 for the smooth quadratic potential. Figs. 11(c) and 11(d) confirm

the theoretical behavior plotted in Fig. 7 for the nonsmooth piecewise quadratic potential.

(e) The relative impact on the MFPT of the piecewise-constant reset rate becomes larger as the diffusion coefficient D decreases, for a given reset rate. This is true regardless of whether r_2 is smaller or larger than r_1 , and is seen in Figs. 11(a)–11(d) for the numerical experiments, and respectively in Figs. 13(b), 3(b), 14(b), 8(b) for the theoretical results.

(2) However, differences are observed with the theoretical results in terms of the *relative* decrease or increase in MFPT of using a piecewise-constant reset rate over a constant reset rate. The theoretical results presented in Figs. 2(b) and 7(b) show that the relative MFPT change is about *the same* in both the smooth and nonsmooth potentials. In contrast, the corresponding numerical results presented in Fig. 11 show that the relative MFPT change is *larger* in the nonsmooth piecewise quadratic potential than in the smooth quadratic potential.

We believe the difference is a consequence of both the time discretization of the original continuous-time problem and the tolerance hyperparameter ΔV_{tol} used to signal the occurrence of reaching the target, as described above. More precisely, due to time discretization, there is no longer a continuous evolution of the $x(t)$ solution, but instead a sequence of random discrete steps $x_{k+1} - x_k$ computed according to Eq. (20). Therefore, the time to reach the target

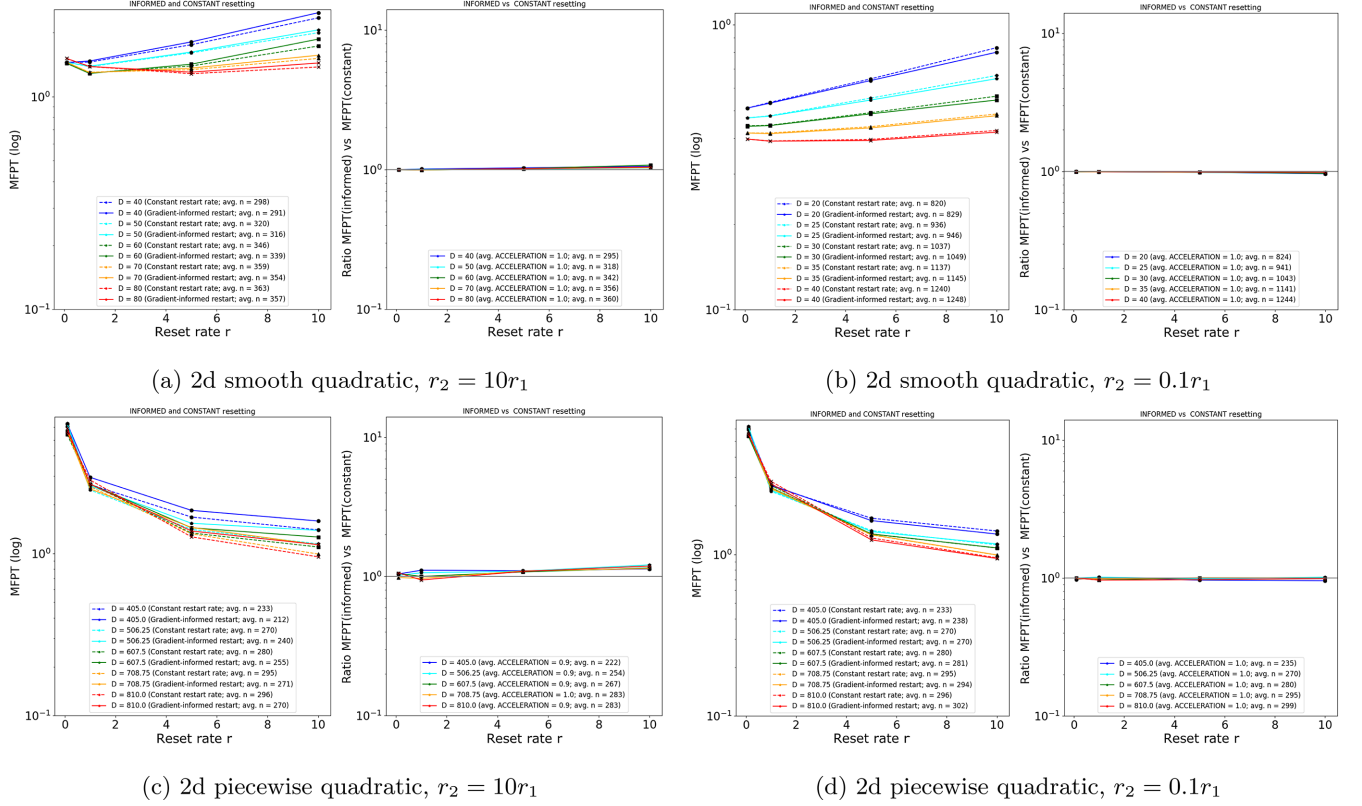


FIG. 12. Average MFPT (measured in time units) obtained on two 2d quadratic potentials, $V(x)$ with $x := (x_1, x_2)$, using either the constant reset rate or the piecewise-constant reset rate defined in (19). Reset rate r_2 is used when the gradient norm is smaller than the threshold $\beta = 1$, and otherwise reset rate r_1 is used. (a) and (b) smooth quadratic potential, $V(x_1, x_2) = x_1^2 + x_2^2$. (c) and (d) nonsmooth piecewise quadratic potential, obtained by revolving the $V(x)$ function defined in Eq. (10) around the V axis. For visual comparison, all right plots across panels, showing the ratio between using a piecewise-constant reset rate vs a constant reset rate, have the same vertical scales.

is clearly a function of the ΔV_{tol} tolerance. Additionally, the particle may have actually *passed* by the target from one particular $x_{k-1} \rightarrow x_k$ iteration, but with no detection of the target just because $V(x_k)$ does not satisfy the specified tolerance. Thus, the discrete-time algorithm needs to spend some time near the target before triggering a reset event, as this will increase the chances that $V(x_l)$ satisfies

the tolerance at some future iteration l . This is why using a gradient-based piecewise-constant reset rate with $r_2 < r_1$ is beneficial, as observed both theoretically and numerically. However, in the discrete-time context, this benefit is expected to be larger in the nonsmooth piecewise-quadratic potential than in the smooth quadratic potential because a reset event sends the particle back to its original position, where the

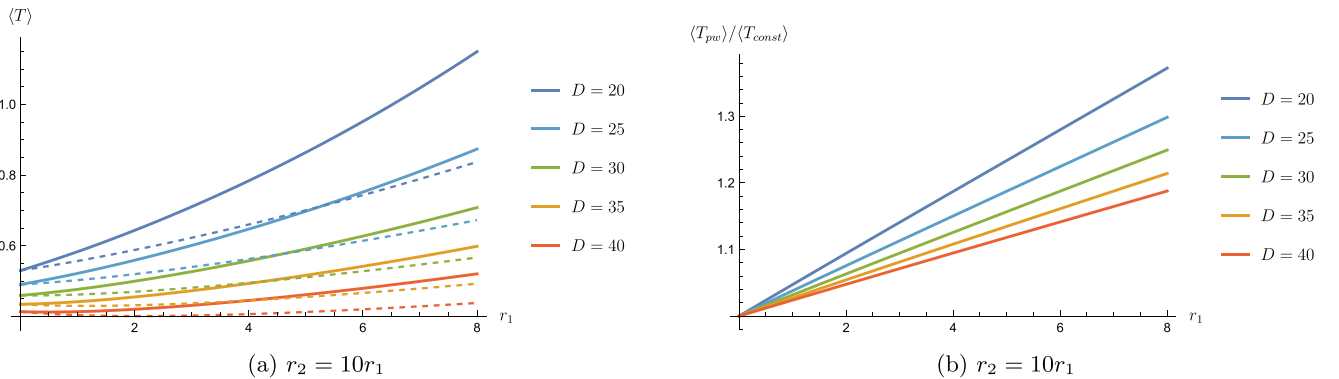


FIG. 13. (1d smooth potential/piecewise vs constant reset/varying D). Comparison of the MFPT, $\langle T \rangle := T_0$ in Eq. (9), between using a constant reset rate vs the piecewise-constant reset rate defined in Eq. (8) with $r_2 = 10r_1$, for the smooth quadratic potential $V(x) = x^2$, for values of $D \in \{20, 25, 30, 35, 40\}$. Here we have fixed $L = 0.01$, $x_0 = 4$, $\beta = 1$. (a) MFPT for constant reset rate ($\langle T_{\text{const}}$, dashed) and piecewise-constant reset rate ($\langle T_{\text{pw}}$, solid), labeled by D . (b) Ratio $\langle T_{\text{pw}} \rangle / \langle T_{\text{const}} \rangle$, labeled by D . All numerical evaluations were done using Mathematica [37].

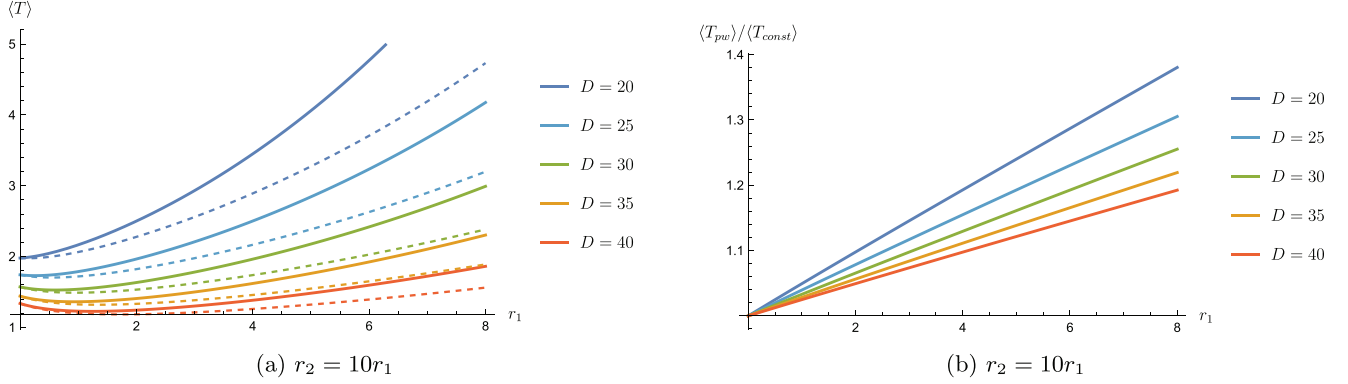


FIG. 14. (1d nonsmooth potential/piecewise vs constant reset/varying D). Comparison of the MFPT, $\langle T \rangle := T_0$ in Eq. (A1), between using a constant reset rate and the piecewise-constant reset rate defined in Eq. (8) with $r_2 = 10r_1$, for the nonsmooth piecewise-quadratic potential defined in Eq. (10), for values of $D \in \{20, 25, 30, 35, 40\}$. Here we have fixed $L = 0.01$, $x_0 = 6$, $\beta = 1$. (a) MFPT for constant reset rate ($\langle T_{const} \rangle$, dashed) and piecewise-constant reset rate ($\langle T_{pw} \rangle$, solid), labeled by D . (b) Ratio $\langle T_{pw} \rangle / \langle T_{const} \rangle$, labeled by D . All numerical evaluations were done using Mathematica [37].

signal-to-noise ratio is smaller than in the smooth case (as mentioned previously), and the particle needs to overcome the local optimum valley before being able to reach the global optimum target.

(3) The 2d experiments show similar results as the 1d counterparts, but the benefits or hindrances of using a piecewise-constant reset rate are not as apparent as in the 1d case, even when the same initial signal-to-noise ratios are used in the respective cases.

Finally, it is interesting to note that using a piecewise-constant reset rate with $r_2 > r_1$ can make the optimal reset rate (of the constant reset rate strategy) vanish. This can be clearly seen for the nonsmooth piecewise quadratic potential in the left subplot of Fig. 11(c).

V. DISCUSSION

In this paper, we have considered diffusions with drift given by the gradient of a potential function and with stochastic resetting to a fixed position. We have analyzed the mean first-passage time (MFPT) for the scenario when the potential of the drift is nonsmooth, in particular piecewise quadratic, as well as the scenario where the reset rate is piecewise constant.

The theoretical analysis of the nonsmooth case shows that the benefit of resetting is dependent on the interplay between drift and noise, e.g., when the drift at the reset position is weak (small potential) compared to the noise (large diffusion coefficient), the resetting is more likely to be beneficial. This is also consistent with previous literature on the smooth case [12,14]. We further find that having a piecewise-constant reset rate is sometimes beneficial, but not always. This depends on the condition for switching between the different rates as well as the ratio between the rates. In our analysis, with a gradient dependent resetting rate, we found that it is detrimental to use a higher reset rate for smaller gradient regions, but beneficial to switch to a lower rate in those same regions (some of which may occur near the given target). See plots of theoretical curves in Figs. 3 and 7, as well as the numerical experiment curves in Figs. 11 and 12. This may suggest iterative gradient-based optimization algorithms with reset rates conditional

on both the objective-function value and some norm of the subgradient at each iteration.

Additionally, we have confirmed with numerical experiments all theoretical results developed on the 1d case, except for the differences described in item (2) of Sec. IV C, which regard the *relative* impact on the MFPT of using a piecewise-constant reset rate compared to using a constant one. In the 1d numerical experiments, this relative change is stronger in the nonsmooth piecewise quadratic potential than in the smooth quadratic potential, whereas it is the same from the theoretical standpoint. The conjectured reasons for this discrepancy are explained in Sec. IV C. Although this discrepancy is less apparent in the 2d case, it suggests that space-dependent resetting may be more beneficial in practice than predicted by theory, especially in nonconvex problems.

Clearly, the analysis and experiments of our paper are limited to very simple nonsmooth potentials. In particular, the gradient threshold β used to define the piecewise-constant reset rate has been set to 1 throughout, based on the knowledge of the potentials analyzed. In practical settings, where the potential function is generally not known but instead learned from data collected as the discrete diffusion algorithm progresses, sensible strategies to define the threshold are called for. For instance, the authors in [36] consider a burn-in exploration period, with no reset events, which is used to estimate an average gradient value of which a fraction is used as the β parameter defining the piecewise-constant resetting. Robustness against noise could be incorporated by choosing an appropriate quantile of the gradient norm distribution, albeit at higher storage costs. It could also be updated with a certain frequency based on updated gradient norm distributions, computed on a moving window. Information about the variability of the observed gradient values could also be leveraged, as done in many gradient-based algorithms, such as Nesterov's accelerated gradient method [41].

It should be noted that a few optimization algorithms typically used in machine learning, e.g., the Adam algorithm [42], implement adaptive step sizes Δt that are adjusted based on the past values of the function and its gradient. Even though such approaches could be framed into our context by keeping

a constant Δt value and absorbing any adjustments into a time-dependent mobility parameter μ , they fall outside the scope of this work for two main reasons: (i) The μ parameter is assumed to be constant (equal to 1) in the Chapman-Kolmogorov equation (3), and (ii) the system's condition after reset is not the same as at the onset of the optimization algorithm because new function and gradient information has been collected by then and thus impact the mobility parameter μ after reset. The theoretical analysis would therefore need adjusting in order to account for such settings. This would be an interesting direction for future research.

In sum, as seen by the results of this paper which confirm the classic results of Evans and Majumdar [3], resetting can either benefit or hinder the search depending on the parameters of the process. Pal *et al.* [13], developed a Landau-like theory to characterize these phase transitions in resetting processes. This is further studied in [17]. It would be interesting to study these aspects further for the settings discussed in our paper. It would also be of interest to explore *restarting* strategies, i.e., where the reset position is selected at random in the solution

space, as this will promote exploration of other regions and thus may help avoid local optima and find the global optimum faster. We plan to explore this approach in future work as well as to generalize the choice of the gradient threshold for practical optimization problems where the potential function is unknown.

ACKNOWLEDGMENTS

We are grateful to the two anonymous reviewers, whose comments helped improve the paper. This work has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101070568.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: THE MFPT FOR PIECEWISE QUADRATIC POTENTIAL WITH TWO RESET RATES

For completeness, we give the full expression for the solutions of the MFPT in the case when we have a one-dimensional piecewise quadratic potential and two different reset rates. In other words, the solution to the boundary value problem in (15). The exact solution depends on the assumptions on the values of the target L , the cutoff β , and the starting position x_0 . The case of most interest, which is also the one studied in the main text is for $0 < \beta \leq 1$, $L \leq \frac{\beta}{2}$ and $x_0 > 4$. With these conditions we find the expression of the MFPT for the nonsmooth piecewise quadratic potential with piecewise-constant reset rate from the boundary value problem (15):

$$\begin{aligned}
 T_0 = & \frac{1}{r_1 r_2 H_{-r_1} \left(\frac{x_0 - 4}{\sqrt{2D}} \right) \left(2H_{-\frac{r_1}{2}} \left(\frac{2}{\sqrt{D}} \right) {}_1F_1 \left[1 + \frac{r_1}{4}, \frac{3}{2}, \frac{4}{D} \right] + \sqrt{D} H_{-1-\frac{r_1}{2}} \left(\frac{2}{\sqrt{D}} \right) {}_1F_1 \left[\frac{r_1}{4}, \frac{1}{2}, \frac{4}{D} \right] \right)} \\
 & \times \frac{1}{\beta H_{-\frac{r_2}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[1 + \frac{r_2}{4}, \frac{3}{2}, \frac{\beta^2}{4D} \right] + 2\sqrt{D} H_{-1-\frac{r_2}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{\beta^2}{4D} \right]} \\
 & \times \left(H_{-r_1} \left(\frac{x_0 - 4}{\sqrt{2D}} \right) \left(2H_{-\frac{r_1}{2}} \left(\frac{2}{\sqrt{D}} \right) {}_1F_1 \left[1 + \frac{r_1}{4}, \frac{3}{2}, \frac{4}{D} \right] + \sqrt{D} H_{-1-\frac{r_1}{2}} \left(\frac{2}{\sqrt{D}} \right) {}_1F_1 \left[\frac{r_1}{4}, \frac{1}{2}, \frac{4}{D} \right] \right) \right. \\
 & \times \left((r_1 - r_2) \beta H_{-\frac{r_2}{2}} \left(\frac{L}{\sqrt{D}} \right) {}_1F_1 \left[1 + \frac{r_2}{4}, \frac{3}{2}, \frac{\beta^2}{4D} \right] - r_1 \beta H_{-\frac{r_2}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[1 + \frac{r_2}{4}, \frac{3}{2}, \frac{\beta^2}{4D} \right] \right. \\
 & + 2\sqrt{D} H_{-1-\frac{r_2}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) \left((r_1 - r_2) {}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{L^2}{D} \right] - r_{11} {}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{\beta^2}{4D} \right] \right) + \sqrt{2D} H_{-1-r_1} \left(-\sqrt{\frac{2}{D}} \right) \\
 & \times \left(r_2 H_{-\frac{r_1}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[\frac{r_1}{4}, \frac{1}{2}, \frac{4}{D} \right] \left(\beta H_{-\frac{r_2}{2}} \left(\frac{L}{\sqrt{D}} \right) {}_1F_1 \left[1 + \frac{r_2}{4}, \frac{3}{2}, \frac{\beta^2}{4D} \right] + 2\sqrt{D} H_{-1-\frac{r_2}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{L^2}{D} \right] \right) \right. \\
 & + 2r_1 \sqrt{D} H_{-1-\frac{r_1}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[\frac{r_1}{4}, \frac{1}{2}, \frac{4}{D} \right] \left(H_{-\frac{r_2}{2}} \left(\frac{L}{\sqrt{D}} \right) {}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{\beta^2}{4D} \right] - H_{-\frac{r_2}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{L^2}{D} \right] \right) \\
 & - H_{-\frac{r_1}{2}} \left(\frac{2}{\sqrt{D}} \right) \left({}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{L^2}{D} \right] \left(r_1 \beta H_{-\frac{r_2}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[1 + \frac{r_1}{4}, \frac{3}{2}, \frac{\beta^2}{4D} \right] + 2r_2 \sqrt{D} H_{-1-\frac{r_2}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{\beta^2}{4D} \right] \right) \right. \\
 & + \beta H_{-\frac{r_2}{2}} \left(\frac{L}{\sqrt{D}} \right) \left(r_{21} {}_1F_1 \left[\frac{r_1}{4}, \frac{1}{2}, \frac{\beta^2}{4D} \right] {}_1F_1 \left[1 + \frac{r_2}{4}, \frac{3}{2}, \frac{\beta^2}{4D} \right] - r_{11} {}_1F_1 \left[1 + \frac{r_1}{4}, \frac{3}{2}, \frac{\beta^2}{4D} \right] {}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{\beta^2}{4D} \right] \right) \left. \right) \\
 & + H_{-r_1} \left(-\sqrt{\frac{2}{D}} \right) \left(2r_2 H_{-\frac{r_1}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[1 + \frac{r_1}{4}, \frac{3}{2}, \frac{4}{D} \right] \left(\beta H_{-\frac{r_2}{2}} \left(\frac{L}{\sqrt{D}} \right) {}_1F_1 \left[1 + \frac{r_2}{4}, \frac{3}{2}, \frac{\beta^2}{4D} \right] \right. \right. \\
 & \left. \left. + 2\sqrt{D} H_{-1-\frac{r_2}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{L^2}{D} \right] \right) + \sqrt{D} \left(4r_1 H_{-1-\frac{r_1}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[1 + \frac{r_1}{4}, \frac{3}{2}, \frac{4}{D} \right] \right. \right.
 \end{aligned}$$

$$\begin{aligned}
& \times \left(H_{-\frac{r_2}{2}} \left(\frac{L}{\sqrt{D}} \right) {}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{\beta^2}{4D} \right] - H_{-\frac{r_2}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{L^2}{D} \right] \right. \\
& + H_{-1-\frac{r_1}{2}} \left(\frac{2}{\sqrt{D}} \right) \left({}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{L^2}{D} \right] \left(r_1 \beta H_{-\frac{r_2}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[1 + \frac{r_1}{4}, \frac{3}{2}, \frac{\beta^2}{4D} \right] + 2r_2 \sqrt{D} H_{-1-\frac{r_2}{2}} \left(\frac{\beta}{\sqrt{4D}} \right) {}_1F_1 \left[\frac{r_1}{4}, \frac{1}{2}, \frac{\beta^2}{4D} \right] \right) \right. \\
& \left. \left. + \beta H_{-\frac{r_2}{2}} \left(\frac{L}{\sqrt{D}} \right) \left(r_{21} {}_1F_1 \left[\frac{r_1}{4}, \frac{1}{2}, \frac{\beta^2}{4D} \right] {}_1F_1 \left[1 + \frac{r_2}{4}, \frac{3}{2}, \frac{\beta^2}{4D} \right] - r_{11} {}_1F_1 \left[1 + \frac{r_1}{4}, \frac{3}{2}, \frac{\beta^2}{4D} \right] {}_1F_1 \left[\frac{r_2}{4}, \frac{1}{2}, \frac{\beta^2}{4D} \right] \right) \right) \right). \quad (\text{A1})
\end{aligned}$$

APPENDIX B: THE MFPT FOR PIECEWISE RESETTING IN TWO DIMENSIONS, WITH SMOOTH QUADRATIC POTENTIAL

The MFPT for resetting in a smooth two-dimensional potential, $V(R) = R^2$, with a piecewise-constant resetting rate given by Eq. (19), is given by

$$\begin{aligned}
T_0 = & \frac{1}{r_1 r_2 \left(e^{\frac{i\pi r_1}{4}} \Gamma\left(\frac{r_1}{4}\right) {}_1F_1\left(\frac{r_1}{4}; 1; \frac{R_0^2}{D}\right) + G_{1,2}^{2,0}\left(-\frac{R_0^2}{D} \middle| \frac{1-\frac{r_1}{4}}{0}, 0\right) \right)} \\
& \times \frac{1}{\left(r_1 {}_1F_1\left(\frac{r_1}{4} + 1; 2; \frac{\beta^2}{4D}\right) G_{1,2}^{2,0}\left(-\frac{\beta^2}{4D} \middle| \frac{1-\frac{r_1}{4}}{0}, 0\right) - 4 {}_1F_1\left(\frac{r_1}{4}; 1; \frac{\beta^2}{4D}\right) G_{1,2}^{2,0}\left(-\frac{\beta^2}{4D} \middle| -1, 0\right) \right)} \\
& \times \left(r_2 \left(e^{\frac{i\pi r_1}{4}} r_1 \Gamma\left(\frac{r_1}{4}\right) {}_1F_1\left(\frac{r_1}{4}; 1; \frac{R_0^2}{D}\right) {}_1F_1\left(\frac{r_1}{4} + 1; 2; \frac{\beta^2}{4D}\right) - e^{\frac{i\pi r_1}{4}} r_1 \Gamma\left(\frac{r_1}{4}\right) {}_1F_1\left(\frac{r_1}{4}; 1; \frac{R_0^2}{D}\right) {}_1F_1\left(\frac{r_1}{4} + 1; 2; \frac{\beta^2}{4D}\right) \right. \right. \\
& + e^{\frac{i\pi r_1}{4}} r_1 \Gamma\left(\frac{r_1}{4}\right) {}_1F_1\left(\frac{r_1}{4}; 1; \frac{\beta^2}{4D}\right) {}_1F_1\left(\frac{r_1}{4} + 1; 2; \frac{\beta^2}{4D}\right) + (r_1 - r_1) G_{1,2}^{2,0}\left(-\frac{R_0^2}{D} \middle| \frac{1-\frac{r_1}{4}}{0}, 0\right) {}_1F_1\left(\frac{r_1}{4} + 1; 2; \frac{\beta^2}{4D}\right) \\
& + r_1 G_{1,2}^{2,0}\left(-\frac{\beta^2}{4D} \middle| \frac{1-\frac{r_1}{4}}{0}, 0\right) {}_1F_1\left(\frac{r_1}{4} + 1; 2; \frac{\beta^2}{4D}\right) - e^{\frac{i\pi r_1}{4}} r_1 \Gamma\left(\frac{r_1}{4}\right) {}_1F_1\left(\frac{r_1}{4} + 1; 2; \frac{\beta^2}{4D}\right) {}_1F_1\left(\frac{r_1}{4}; 1; \frac{\beta^2}{4D}\right) \\
& - 4 {}_1F_1\left(\frac{r_1}{4}; 1; \frac{\beta^2}{4D}\right) G_{1,2}^{2,0}\left(-\frac{\beta^2}{4D} \middle| -1, 0\right) \left. \right) G_{1,2}^{2,0}\left(-\frac{L^2}{D} \middle| \frac{1-\frac{r_1}{4}}{0}, 0\right) \\
& - r_1 \left(e^{\frac{i\pi r_1}{4}} r_1 \Gamma\left(\frac{r_1}{4}\right) \left({}_1F_1\left(\frac{r_1}{4}; 1; \frac{R_0^2}{D}\right) {}_1F_1\left(\frac{r_1}{4} + 1; 2; \frac{\beta^2}{4D}\right) - {}_1F_1\left(\frac{r_1}{4} + 1; 2; \frac{\beta^2}{4D}\right) {}_1F_1\left(\frac{r_1}{4}; 1; \frac{L^2}{D}\right) \right) \right. \\
& + r_1 {}_1F_1\left(\frac{r_1}{4} + 1; 2; \frac{\beta^2}{4D}\right) G_{1,2}^{2,0}\left(-\frac{R_0^2}{D} \middle| \frac{1-\frac{r_1}{4}}{0}, 0\right) - 4 {}_1F_1\left(\frac{r_1}{4}; 1; \frac{L^2}{D}\right) G_{1,2}^{2,0}\left(-\frac{\beta^2}{4D} \middle| -1, 0\right) \left. \right) G_{1,2}^{2,0}\left(-\frac{\beta^2}{4D} \middle| \frac{1-\frac{r_1}{4}}{0}, 0\right) \\
& - 4 \left(e^{\frac{i\pi r_1}{4}} \Gamma\left(\frac{r_1}{4}\right) \left(r_1 {}_1F_1\left(\frac{r_1}{4}; 1; \frac{\beta^2}{4D}\right) {}_1F_1\left(\frac{r_1}{4}; 1; \frac{L^2}{D}\right) + {}_1F_1\left(\frac{r_1}{4}; 1; \frac{R_0^2}{D}\right) \left((r_1 - r_1) {}_1F_1\left(\frac{r_1}{4}; 1; \frac{L^2}{D}\right) \right. \right. \\
& \left. \left. - r_1 {}_1F_1\left(\frac{r_1}{4}; 1; \frac{\beta^2}{4D}\right) \right) \right) + \left((r_1 - r_1) {}_1F_1\left(\frac{r_1}{4}; 1; \frac{L^2}{D}\right) - r_1 {}_1F_1\left(\frac{r_1}{4}; 1; \frac{\beta^2}{4D}\right) \right) G_{1,2}^{2,0}\left(-\frac{R_0^2}{D} \middle| \frac{1-\frac{r_1}{4}}{0}, 0\right) \\
& + r_1 {}_1F_1\left(\frac{r_1}{4}; 1; \frac{L^2}{D}\right) G_{1,2}^{2,0}\left(-\frac{\beta^2}{4D} \middle| \frac{1-\frac{r_1}{4}}{0}, 0\right) \left. \right) G_{1,2}^{2,0}\left(-\frac{\beta^2}{4D} \middle| -1, 0\right). \quad (\text{B1})
\end{aligned}$$

Here

$$G_{p,q}^{m,n} \left(z \middle| \begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \right)$$

is the Meijer G -function.

APPENDIX C: ADDITIONAL RESULTS ON THE MFPT WHEN $r_2 > r_1$

For completeness of the discussion, we add here a few more figures, Figs. 13–15, which correspond to Figs. 3, 8, and 10, but with $r_2 = 10r_1$ instead. This highlights the fact that it is only when $r_2 < r_1$ that we see the benefit of having multiple reset rates.

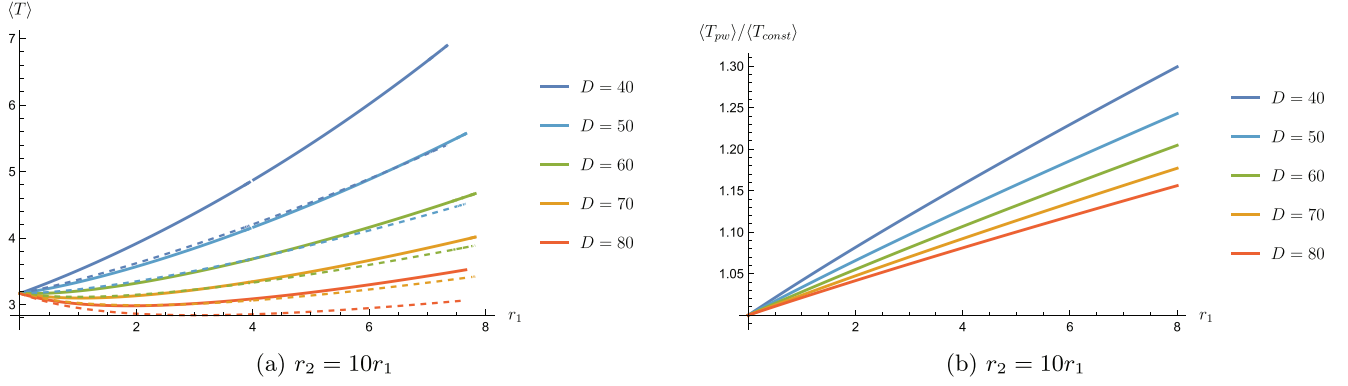


FIG. 15. (2d smooth potential/piecewise vs constant reset/varying D). Comparison of the MFPT, $\langle T \rangle =: T_0$ of Eq. (B1), between using a constant reset rate and the piecewise-constant reset rate defined in Eq. (19) with $r_2 = 10r_1$, for a two-dimensional smooth quadratic potential $V(R) = R^2$, for values of $D \in \{40, 50, 60, 70, 80\}$. Here we have fixed $L = 0.01$, $R_0 = 4\sqrt{2}$, $\beta = 1$. (a) MFPT for constant reset rate ($\langle T_{const} \rangle$, dashed) and piecewise-constant reset rate ($\langle T_{pw} \rangle$, solid), labeled by D . (b) Ratio $\langle T_{pw} \rangle / \langle T_{const} \rangle$, labeled by D . All numerical evaluations were done using Mathematica [37].

APPENDIX D: EXAMPLE TRAJECTORY OF THE DISCRETE-TIME 2D DIFFUSION PROCESS

Figure 16 shows an example of a trajectory of a particle driven by the discrete-time diffusion equation (20) with $F(x) = -\nabla V(x)$, where $V(x)$ is a 2d symmetric nonsmooth piecewise quadratic potential defined by Eq. (10) in each

radial direction. A piecewise-constant reset rate strategy is used, as defined in Eq. (19), $\beta = 1$, $r_1 = 10$, and $r_2 = 10r_1$. The reset point is set to $x_0 = (6, 6)$, the diffusion coefficient is $D = 810$. The discretization interval of the original continuous-time diffusion equation (1) is set to $\Delta t = 10^{-4}$. The particle initially tends to drift away but finally gets closer to the potential's global optimum at $(0, 0)$.

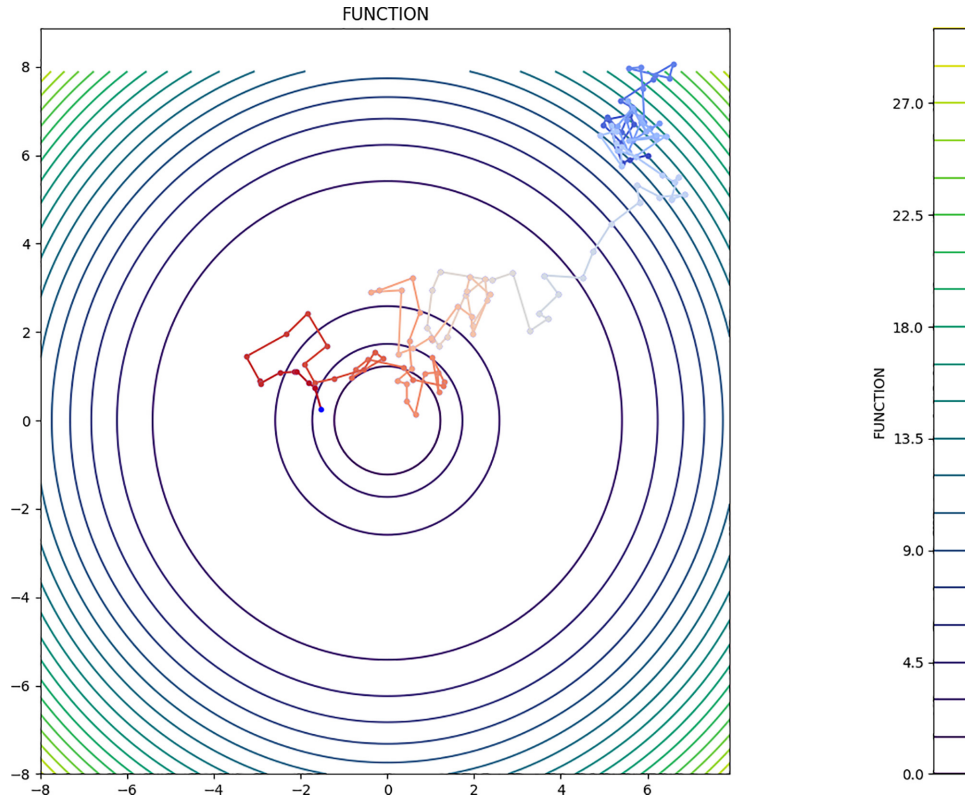


FIG. 16. Example of a particle diffusing on a 2d nonsmooth piecewise quadratic potential having a gradient discontinuity at a radius $R = 2$, overlaid on the potential's contour plot. The diffusion and potential characteristics are described in Appendix D. Time is represented by the trajectory color, as it turns from blue to red.

- [1] I. Eliazar, T. Koren, and J. Klafter, Searching circular DNA strands, *J. Phys.: Condens. Matter* **19**, 065140 (2007).
- [2] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, *Nature (London)* **521**, 436 (2015).
- [3] Martin R. Evans and Satya N. Majumdar, Diffusion with stochastic resetting, *Phys. Rev. Lett.* **106**, 160601 (2011).
- [4] M. R. Evans and S. N. Majumdar, Diffusion with optimal resetting, *J. Phys. A: Math. Theor.* **44**, 435001 (2011).
- [5] M. R. Evans and S. N. Majumdar, Diffusion with resetting in arbitrary spatial dimension, *J. Phys. A: Math. Theor.* **47**, 285001 (2014).
- [6] O. Bénichou, M. Moreau, P.-H. Suet, and R. Voituriez, Intermittent search process and teleportation, *J. Chem. Phys.* **126**, 234109 (2007).
- [7] E. Gelenbe, Search in unknown random environments, *Phys. Rev. E* **82**, 061112 (2010).
- [8] S. Janson and Y. Peres, Hitting times for random walks with restarts, *SIAM J. Discrete Math.* **26**, 537 (2012).
- [9] A. Pal and S. Reuveni, First passage under restart, *Phys. Rev. Lett.* **118**, 030603 (2017).
- [10] É. Roldán and S. Gupta, Path-integral formalism for stochastic resetting: Exactly solved examples and shortcuts to confinement, *Phys. Rev. E* **96**, 022130 (2017).
- [11] M. R. Evans, S. N. Majumdar, and G. Schehr, Stochastic resetting and applications, *J. Phys. A: Math. Theor.* **53**, 193001 (2020).
- [12] S. Ahmad, I. Nayak, A. Bansal, A. Nandi, and D. Das, First passage of a particle in a potential under stochastic resetting: A vanishing transition of optimal resetting rate, *Phys. Rev. E* **99**, 022130 (2019).
- [13] A. Pal and V. V. Prasad, Landau-like expansion for phase transitions in stochastic resetting, *Phys. Rev. Res.* **1**, 032001(R) (2019).
- [14] S. Ray, D. Mondal, and S. Reuveni, Péclet number governs transition to acceleratory restart in drift-diffusion, *J. Phys. A: Math. Theor.* **52**, 255002 (2019).
- [15] S. Ray and S. Reuveni, Diffusion with resetting in a logarithmic potential, *J. Chem. Phys.* **152**, 234110 (2020).
- [16] O. L. Lauber Bonomo and A. Pal, First passage under restart for discrete space and time: Application to one-dimensional confined lattice random walks, *Phys. Rev. E* **103**, 052129 (2021).
- [17] S. Ahmad, K. Rijal, and D. Das, First passage in the presence of stochastic resetting and a potential barrier, *Phys. Rev. E* **105**, 044134 (2022).
- [18] M. Luby, A. Sinclair, and D. Zuckerman, Optimal speedup of Las Vegas algorithms, *Inf. Process. Lett.* **47**, 173 (1993).
- [19] H. Alt, L. Guibas, K. Mehlhorn, R. Karp, and A. Wigderson, A method for obtaining randomized algorithms with small tail probabilities, *Algorithmica* **16**, 543 (1996).
- [20] H. Kautz, E. Horvitz, Y. Ruan, C. Gomes, and B. Selman, Dynamic restart policies, *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 97 (AAAI, 2002), p. 674.
- [21] H. Tong, C. Faloutsos, and J.-Y. Pan, Random walk with restart: Fast solutions and applications, *Knowl. Inf. Syst.* **14**, 327 (2008).
- [22] O. V. Shylo and O. A. Prokopyev, *Restart Strategies* (Springer International Publishing, Cham, 2018), pp. 205–220.
- [23] H. Yu, S. Yang, and S. Zhu, Parallel restarted SGD with faster convergence and less communication: Demystifying why model averaging works for deep learning, *Proc. AAAI Conf. Artif. Intell.* **33**, 5693 (2019).
- [24] I. Loshchilov and F. Hutter, SGDR: Stochastic gradient descent with warm restarts, in *International Conference on Learning Representations* (ICLR, 2017).
- [25] Y. Bae, Y. Song, and H. Jeong, Stochastic restarting to overcome overfitting in neural networks with noisy labels, *Mach. Learn.: Sci. Technol.* **6**, 015062 (2025).
- [26] M. Gagliolo and J. Schmidhuber, Learning restart strategies, In *Proceedings of the 20th International Joint Conference on Artificial Intelligence* (Morgan Kaufmann Publishers Inc., San Francisco, CA, 2007), pp. 792–797.
- [27] N. A. Wedge and M. S. Branicky, On heavy-tailed runtimes and restarts in rapidly-exploring random trees, in *Twenty-Third AAAI Conference on Artificial Intelligence* (AAAI, 2008), pp. 127–133.
- [28] V. Roulet and A. d’Aspremont, Sharpness, restart and acceleration, *Advances in Neural Information Processing Systems*, Vol. 30 (NIPS) (Curran Associates, Inc., 2017).
- [29] J.-H. Lorenz, Restart strategies in a continuous setting, *Theory Comput. Syst.* **65**, 1143 (2021).
- [30] J. Renegar and B. Grimmer, A simple nearly optimal restart scheme for speeding up first-order methods, *Found. Comput. Math.* **22**, 211 (2022).
- [31] D. Starkov and S. Belan, Universal performance bounds of restart, *Phys. Rev. E* **107**, L062101, (2023).
- [32] D. Davis, D. Drusvyatskiy, S. Kakade, and J. D. Lee, Stochastic subgradient method converges on tame functions, *Found. Comput. Math.* **20**, 119 (2020).
- [33] J. Berner, P. Grohs, G. Kutyniok, and P. Petersen, *The Modern Mathematics of Deep Learning*, edited by P. Grohs and G. Kutyniok (Cambridge University Press, Cambridge, 2022), pp. 1–11.
- [34] A. Pal, Diffusion in a potential landscape with stochastic resetting, *Phys. Rev. E* **91**, 012113 (2015).
- [35] S. Redner, *A Guide to First-Passage Processes* (Cambridge University Press, Cambridge, 2001).
- [36] D. Mastropietro, G. Korpas, V. Kungurtsev, and J. Marecek, Parallel variational quantum algorithms with gradient-informed restart to speed up optimization in the presence of barren plateaus, *IEEE Trans. Quantum Eng.* **7**, 1 (2026).
- [37] Wolfram Research, Inc., Mathematica, Version 13.1, Champaign, IL (2022).
- [38] F. Batola, Une généralisation d’une formule d’Erdelyi-Tricomi, *Arkiv för Matematik* **20**, 87 (1982).
- [39] F. Tricomi, Sulle funzioni ipergeometriche confluenti, *Ann. Mat. Pura Appl.* **26**, 141 (1947).
- [40] W. Mou, N. Flammarion, M. J. Wainwright, and P. L. Bartlett, Improved bounds for discretization of Langevin diffusions: Near-optimal rates without convexity, *Bernoulli* **28**, 1577 (2022).
- [41] Y. Nesterov, A method for unconstrained convex minimization problem with the rate of convergence $o(1/k^2)$, *Dokl. Akad. Nauk. SSSR* **269**, 543 (1983).
- [42] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, in *International Conference on Learning Representations* (ICLR, 2015).