Estimating the most probable transition time for stochastic dynamical systems

Yuanfei Huang¹, Ying Chao², Wei Wei¹ and Jinqiao Duan³,∗

¹ School of Mathematics and Statistics & Center for Mathematical Sciences & Hubei National Center for Applied Mathematics, Huazhong University of Science and Technology, Wuhan 430074, People’s Republic of China
² School of Mathematics and Statistics, Xi’an Jiaotong University, Xi’an 710049, People’s Republic of China
³ Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616, United States of America

E-mail: yfhuang@hust.edu.cn, yingchao1993@xjtu.edu.cn, weiw16@hust.edu.cn and duan@iit.edu

Received 24 June 2020, revised 8 March 2021
Accepted for publication 29 March 2021
Published 21 June 2021

Abstract
This work is devoted to the investigation of the most probable transition time between metastable states for stochastic dynamical systems with non-vanishing Brownian noise. Instead of minimizing the Onsager–Machlup action functional, we examine the maximum probability that the solution process of the system stays in a neighbourhood (or a tube) of a transition path, in order to characterize the most probable transition path. We first establish the exponential decay lower bound and a power law decay upper bound for the maximum of this probability. Based on these estimates, we further derive the lower and upper bounds for the most probable transition time, under suitable conditions. Finally, we illustrate our results in simple stochastic dynamical systems, and highlight the relation with some relevant works.

Keywords: stochastic differential equations, most probable transition time, Onsager–Machlup action functional, metastable states, rare events

Mathematics Subject Classification numbers: 60H10.

(Some figures may appear in colour only in the online journal)
1. Introduction

Stochastic differential equations (SDEs) are models for complex phenomena in physical, chemical, biological, and engineering systems under random fluctuations. In particular, transition phenomena between dynamically significant states, such as climate change and gene transcription, occur under the interaction of nonlinearity and uncertainty. It has been an interesting issue to quantify the transition behaviour between two metastable states for stochastic dynamical systems. Indeed, this has been actively investigated [1–16]. The Onsager–Machlup (OM) method and the related path integrals formulation have been set up to study this problem through minimizing a so-called OM action functional. It is difficult to predict and describes state changes or transitions in, for example, climate systems and Arctic sea [17–19]. It was reported [20, 21] that transition phenomena in climate systems complete in finite (or even relatively short) time scales. In molecule dynamics, the transitions between two molecule states (or protein configurations) may occur in finite or short time [22, 23]. In gene regulation systems, transitions between different concentration levels of transcription factors may happen frequently [24, 25].

Therefore, as a part of investigation on the transition pathways, it is crucial to characterize the transition time. The purpose of this paper is to estimate the most probable transition time. We consider the following SDE in Euclidean space $\mathbb{R}^k$:

$$dX_t = b(X_t)dt + c\, dB_t, \quad t \geq 0. \quad (1.1)$$

Here $B_t$ is a standard Brownian motion in $\mathbb{R}^k$. The noise intensity $c$ is a positive constant. We suppose that $b(x)$ is in the space $C^2(\mathbb{R}^k, \mathbb{R}^k)$ of functions having all continuous derivatives of order up to 2. There exist at least two metastable states of system (1.1). Let $x_0$ and $x_f$ be two distinct metastable states of the system (1.1): $b(x_0) = 0$ and $b(x_f) = 0$.

For system (1.1) with a given transition time $T$, this is the usual setup for studying transition paths between two metastable states [1, 26]: among all possible smooth paths connecting two metastable states ($X_0 = x_0$ and $X_T = x_f$), which one is the most probable for the solution process of (1.1)? The solution process of system (1.1) is almost surely nowhere differentiable. So to quantify which smooth path is the most probable one, an usual way is to compare the probabilities that the solution process stays in the neighbourhood or ‘tube’ of such a smooth path. It was proved in [1, 14, 15, 26] that, the probability of the solution process of (1.1) in a tube of a smooth path $\psi(t)$, as the tube size scaling $\delta \to 0$, is

$$\mathbb{P}^x_0 \{ \| X - \psi \|_T \leq \delta \} \approx \exp \left( -\frac{S_{OM}(\psi)}{c^2} \right) \cdot \mathbb{P}^x_0 \{ \| B^c \|_T \leq \delta \}, \quad (1.2)$$

where $\mathbb{P}^x_0$ denotes the probability conditional on the initial position $X_0 = x_0$, $\cdot$ is the Euclidean norm, and $\| \cdot \|_T$ is the uniform norm:

$$\| \psi \|_T = \sup_{0 \leq t \leq T} | \psi(t) |. \quad (1.3)$$

Here $B^c_t = c(B_t - B_0) + x_0$ (a shifted Brownian motion with magnitude $c$). The OM action functional is

$$S_{OM}^T(\psi) = \frac{1}{2} \int_0^T \left[ \langle \dot{\psi}(t) - b(\psi(t)), (\dot{\psi}(t) - b(\psi(t))) \rangle + c^2 \nabla \cdot b(\psi(t)) \right] dt, \quad (1.4)$$
where $\langle \cdot, \cdot \rangle$ denotes the Euclidean scalar product in $\mathbb{R}^k$. The ‘tube probability’ estimate (1.2) shows that for a given transition time $T$, in order to compare the probabilities that the solution process of (1.1) stays in the neighbourhood of a smooth path, it can be approximately performed by comparing their corresponding OM action functionals. Thus the most probable transition path (MPTP), when the transition time $T$ is given or known, is defined as the one which minimizes the OM action functional among a class of smooth paths [1, 14, 15, 26]:

$$
\sup_{\psi \in C^2([0,T],x_0,x_f)} \mathbb{P}^{\psi_0}\{\|X - \psi\|_T \leq \delta\} \approx \sup_{\psi \in C^2([0,T],x_0,x_f)} \exp(-\frac{S_{\text{OM}}(\psi)}{c^2}) \cdot \mathbb{P}^{\psi_0}\{\|B^c - x_0\|_T \leq \delta\}
$$

where $C^2([0,T],x_0,x_f)$ denotes the set of functions $g : [0,T] \rightarrow \mathbb{R}^k$ such that $g(0) = x_0$, $g(T) = x_f$ and $g'$, $g''$ exist and are continuous. Note that when the transition time $T$ is known, the factor $\mathbb{P}^{\psi_0}\{\|B^c - x_0\|_T \leq \delta\}$ in (1.5) does not affect the minimization.

The aforementioned works focus on the case that the transition time $T$ is known. However, the transition time $T$ varies or is not known in advance in many stochastic dynamical systems in mathematical modelling. It is thus desirable to estimate the transition time $T$. In this paper, we focus on the transition behaviour of the stochastic system in a bounded domain. Hence we aim to investigate the following double optimization problem on the tube probability, provided the tube size $\delta$ satisfies the appropriate condition $0 < \delta < |x_f - x_0|$:  

$$
\sup_{T > 0} \sup_{\psi \in C_D([0,T],x_0,x_f)} \mathbb{P}^{\psi_0}\{\|X - \psi(\cdot)\|_T \leq \delta\},
$$

where $D$ is a bounded connected region containing $x_0$, $x_f$, and $C_D([0,T],x_0,x_f)$ denotes the space of all continuous functions $g : [0,T] \rightarrow D$ such that $g(0) = x_0$ and $g(T) = x_f$. As the two metastable states $x_0$ and $x_f$ are given, we use the notation $C_D([0,T],x_0,x_f)$ to denote $C_D([0,T],x_0,x_f)$. The set $D$ can be considered as a transition region. For instance, in one dimensional cases, the set $D$ is selected as the interval with ending points $x_0$ and $x_f$. The most probable transition time for the stochastic dynamical system (1.1) transits from metastable state $x_0$ to another metastable state $x_f$ is defined as the time at which this double optimization problem achieves its double maximum value. The corresponding MPTP may be regarded as global. We denote the most probable transition time as $T_{x_0 \rightarrow x_f}^{\text{MPTP}}$.

This paper is organized as follows. After recalling some preliminaries (section 2), we investigate the probability that the solution process of the SDE stays in a neighbourhood of a smooth path (section 3). We establish the exponential decay lower bound and power law decay upper bound of the maximum of the probability for a class of transition paths. Furthermore, we derive the bounds for the most probable transition time. Then we present two examples to illustrate our results (section 4). Finally, in section 5 we summarize our main results, and discuss the connection and difference between our work and some relevant works.

2. Preliminaries

We recall some basic concepts about Wiener measure induced by an SDE and exit properties for the solution paths of the SDE.
2.1. Measure induced by the solution process

Let $X$ denote a nonexploding diffusion process on $[0, T]$ defined by the scalar SDE in the probability space $(\Omega, \mathcal{F}, P)$

$$dX_t = b(X_t)dt + c dB_t, \quad X_0 = x_0 \in \mathbb{R}^k. \quad (2.1)$$

The space of paths of such a diffusion process is the space $C([0, T], x_0)$ of continuous functions

$$C([0, T], x_0) = \{ \psi | \psi : [0, T] \rightarrow \mathbb{R}^k, \psi(t) \text{ is continuous, } \psi(0) = x_0 \}, \quad (2.2)$$

with the uniform norm $\| \cdot \|_T$. In this norm, we have the Borel field $\mathcal{B}_{[0,T]}^0$ of $C([0, T], x_0)$.

A subset $I_\delta$ of $C([0, T], x_0)$ in the following form is called an $n$-dimensional cylinder set:

$$I_\delta = \{ \psi \in C([0, T], x_0)|(\psi(t_1), \ldots, \psi(t_n)) \in H = (H_1, \ldots, H_n) \}, \quad (2.3)$$

where $0 < t_1 < \cdots < t_n \leq T$ and $H_i$ is a Borel set in $k$-dimensional Euclidean space. The collection of all $n$-dimensional cylinder sets is a $\sigma$-field and the class of all finite-dimensional cylinder sets is a field, which is denoted by $I$. It is known that the $\sigma$-field $\sigma(I)$, generated by $I$, is the Borel field $\mathcal{B}_{[0,T]}^0$. That is, $\sigma(I) = \mathcal{B}_{[0,T]}^0$.

The measure $\mu_X$ on $\mathcal{B}_{[0,T]}^0$ induced by the solution process $X$ of the SDE (2.1) is defined by

$$\mu_X(B) = \mathbb{P}(w \in \Omega | X(w) \in B), \quad B \in \mathcal{B}_{[0,T]}^0. \quad (2.4)$$

Recall that such a measure induced by Brownian motion $B_t$ is the Wiener measure. For convenience, we also call $\mu_X$ the Wiener measure induced by solution process $X$. Let $K_T(\psi, \delta) = \{ x \in C([0, T], x_0)|\| x - \psi \|_T < \delta \}$ denote the open tube of a path $\psi$ with tube size $\delta$ (i.e. neighbourhood size), for $\psi \in C([0, T], x_0)$. And $\bar{K}_T(\psi, \delta) = \{ x \in C([0, T], x_0)|\| x - \psi \|_T \leq \delta \}$ denotes the corresponding closed tubes. Thus the probability that the solution process of (2.1) stays in the closed $\delta$-tube of a path $\psi \in C([0, T], x_0)$ if $\mu_X(\bar{K}_T(\psi, \delta))$.

**Remark 2.1.** By the definition (2.4), the tube probability in the double optimization problem (1.6) will be estimated via the Wiener measure $\mu_X$ for the rest of this paper.

2.2. Mean exit time and non-exit probability of a diffusion process

In this subsection, we recall the mean exit time of system (1.1). The following lemma from [27, 28] is needed later. Suppose that $M$ is a domain containing $x_0$ and define that $\tau^M_0(X) = \inf \{ t | X_t \notin M, X_0 = x \in M \}$. By [27], it is known that $\tau^M_0(X)$ is a stopping time.

**Lemma 2.2.1 (Mean exit time).** The mean exit time $u(x) = E_T^{\tau^M_0}(X)$ of the stochastic system (1.1) with $c = 1$, for an orbit (i.e., a trajectory) starting at $x \in M$, satisfies the following elliptic partial differential equation

$$Au = -1, u|_{\partial M} = 0, \quad (2.5)$$

where $\partial M$ is the boundary of $M$ and $A$ is the generator

$$Au = b \cdot \nabla u + \frac{1}{2} \Delta u. \quad (2.6)$$

Moreover, if the domain $M$ has $C^{1,\gamma}$ boundary and the drift $b$ is in $C^\gamma(M)$ for some $\gamma \in (0, 1)$, then the mean exit time $u(x)$ uniquely exists and is in $C^{2,\gamma}(M)$.
Recall that $C^\gamma(M)$ is the Hölder space consisting of functions in $M$ which are locally Hölder continuous with exponent $\gamma$. In particular, Hölder space $C^{2,\gamma}(M)$ is the subspace of $C^\gamma(M)$ consisting of functions whose second order derivatives are locally Hölder continuous with exponent $\gamma$. A bounded domain $M$ is called a $C^{2,\gamma}$ domain if each point of its boundary $\partial M$ has a neighbourhood in which $\partial M$ is the graph of a $C^{2,\gamma}$ function. We also say that $M$ has a $C^{2,\gamma}$ boundary.

Now we turn to discuss the probability that a diffusion process $B^t = c(B_t - B_0)$ (where $B_t$ is a standard Brownian motion and $c$ is a positive constant) stays in the open $\delta$-tube of the origin $0$ during the time period $[0, T]$. When the tube size $\delta$ is fixed, according to lemma 2.2.1, the mean exit time of $B^\gamma$ from the $\delta$-neighbourhood of the origin is finite, so the non-exit probability $\mu_{B^\gamma}(K_T(0, \delta))$ tends to $0$ when $T$ tends to infinity. Thus the probability $\mu_{B^\gamma}(K_T(0, \delta))$ decreases monotonically from $1$ to $0$ in $T$.

In one dimensional cases, this probability was studied in lemma 8.1 of [26] with unit noise intensity. We obtain the probability for different noise intensity (appendix A.1):

$$
\mu_{B^\gamma}(K_T(0, \delta)) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)\pi} \exp \left\{ \frac{(2n+1)^2 \pi^2 T}{8\delta^2} \right\}.
$$

\[2.7\]

3. Bounds for the most probable transition time

Without loss of generality, we assume that the set $D$ has $C^{2,\gamma}$ boundary for some $\gamma \in (0, 1)$. For a given tube size $\delta$, the upper and lower bounds of $\sup_{\psi \in C[0, T]} \mu_X(K_T(\psi, \delta))$ are estimated in the following two subsections. Then we establish the bounds for the most probable transition time in subsections 3.3. Finally, we examine the connection between the double optimization problems on the tube probability (1.6) and on the OM action functional $S_{\text{OM}}(\psi)$, in subsection 3.4.

3.1. Upper bound of $\sup_{\psi \in C[0, T]} \mu_X(K_T(\psi, \delta))$

Define an ‘enlarged’ set $D^* = \{ x + 2\delta | x \in D^* \}$ (here $D^*$ denotes the interior of $D$) and let $\tau_{D^*}^{x_0}$ be the first time that the solution process $X$ of system (2.1) escapes from the domain $D^*$:

$$
\tau_{D^*}^{x_0} = \inf \{ t > 0 | x_0, x_t \in \partial D^* \}.
$$

(3.1)

For every time instant $t > 0$, the mean exit time is

$$
\mathbb{E}_{D^*}^{x_0} = \int_{\Omega} \int_{\Gamma_{D^*}^{x_0}(w)} \tau_{D^*}^{x_0}(w) dP^{x_0}(w)
$$

$$
= \int_{\{ \tau_{D^*}^{x_0}(w) \leq t \}} \tau_{D^*}^{x_0}(w) dP^{x_0}(w) + \int_{\{ \tau_{D^*}^{x_0}(w) > t \}} \tau_{D^*}^{x_0}(w) dP^{x_0}(w)
$$

$$
= \int_{\{ \tau_{D^*}^{x_0}(w) \leq t \}} \tau_{D^*}^{x_0}(w) dP^{x_0}(w) + t P^{x_0}(\{ w | \tau_{D^*}^{x_0}(w) > t \}).
$$

(3.2)

According to lemma 2.2.1 and the regularities of the domain $D^*$ and the drift term $b$, we know that $\mathbb{E}_{D^*}^{x_0} \tau_{D^*}^{x_0}$ is finite. Letting $t$ tend to infinity, we obtain that

$$
\lim_{t \to \infty} dP^{x_0}(\{ w | \tau_{D^*}^{x_0}(w) > t \}) = 0.
$$

(3.3)
So for a positive constant $\epsilon$, there exists a positive constant $t_\epsilon$ such that for each $t > t_\epsilon$:

$$P_0(\{ |w| > t \} \in \mathcal{A}) < T < \epsilon. \quad (3.4)$$

Notice that for every $\psi \in C_0([0, T])$ with $T > t_\epsilon$,

$$\mu_X(\mathcal{K}_T(\psi, \delta)) = \mathbb{P}_0(\{ |X_t(w) - \psi(t)| < \delta, \quad \forall t \in [0, T] \})$$

$$\leq \mathbb{P}_0(\{ |w| \psi(t) - \delta < X_t(w) < \psi(t) + \delta, \quad \forall t \in [0, T] \})$$

$$= \mathbb{P}_0(\{ |w| X_t(w) \in D^*, \quad \forall t \in [0, T] \})$$

$$\leq \mathbb{P}_0(\{ |w| \tau(w) > T \})$$

$$< \frac{\epsilon}{T}. \quad (3.5)$$

Thus for $T > t_\epsilon$,

$$\sup_{\psi \in C_0([0, T])} \mu_X(\mathcal{K}_T(\psi, \delta)) \leq \frac{\epsilon}{T}. \quad (3.6)$$

### 3.2. Lower bound of $\sup_{\psi \in C_0([0, T])} \mu_X(\mathcal{K}_T(\psi, \delta))$

For simplicity, we deal with one dimensional cases. The higher dimensional cases are similar with slight modification of the calculation. For a path $\psi \in C^2_0([0, T])$ and tube size $\delta$, there exist constants $h_0, h_1, h_2$ such that the Wiener measure $\mu_X(\mathcal{K}_T(\psi, \delta))$ satisfies the following inequality (appendix A.2)

$$\mu_X(\mathcal{K}_T(\psi, \delta)) \geq \exp \left\{-\frac{1}{c^2} \left[ |\dot{\psi}(T)| \delta + h_0 + \left( h_1 + \frac{1}{2} \delta^2 h_2 + k \|\dot{\psi}\|_T \right) T \delta \right] \right\} \cdot \exp \left\{-\frac{1}{c^2} S_T(\psi) \right\} \cdot \mathbb{P}_0(\mathcal{K}_T(0, \delta)). \quad (3.7)

To give a clear and intuitive approximate form of the lower bound, we consider the following family of paths $\{\phi_T(t)\}_{T > 0}$:

$$\phi_T(t) := \frac{x_T - x_0}{T} t + x_0, \quad t \in [0, T]. \quad (3.8)$$

Then we have

$$\sup_{\psi \in C_0([0, T])} \mu_X(\mathcal{K}_T(\psi, \delta)) \geq \mu_X(\mathcal{K}_T(\phi_T, \delta))$$

$$\geq \exp \left\{-\frac{1}{c^2} \left[ |\phi_T(T)| \delta + h_0 + \left( h_1(\phi_T, \delta) + \frac{1}{2} \delta^2 h_2(\phi_T, \delta) \right) \right] \right\} \cdot \exp \left\{-\frac{1}{c^2} S_T(\phi_T, \phi_T) \right\} \cdot \mathbb{P}_0(\mathcal{K}_T(0, \delta))$$

$$= \exp \left\{-\frac{h_0}{c^2} \cdot \exp \left\{-\frac{1}{c^2} \left[ \frac{|x_T - x_0|}{T} \delta + \left( h_1(\phi_T, \delta) + \frac{1}{2} \delta^2 h_2 \right) \right] \right\} \cdot \exp \left\{-\frac{1}{c^2} S_T(\phi_T, \phi_T) \right\} \cdot \mathbb{P}_0(\mathcal{K}_T(0, \delta)) \right\}$$
\[ := c_0 \exp \left\{ -\frac{1}{c^2} \mathcal{S}_{OM}^{GT}(\phi_T, \dot{\phi}_T) - c_1 T + \ln(\mu_{\mathcal{B}}(\bar{K}_T(0, \delta))) - \frac{|x_f - x_0| \delta}{T c^2} \right\}, \]  
(3.9)

where
\[ c_0 = \exp \left\{ -\frac{h_0}{c^2} \right\}, \quad c_1 = \frac{1}{c^2} \left[ \left( h_1(\phi_T, \delta) + \frac{1}{2} c^2 h_2(\phi_T, \delta) \right) \delta \right], \]  
(3.10)

are positive constants depending on the tube \( \bar{K}_T(\phi_T, \delta) \). Notice that all the paths \( \{\phi_T(t)\}_{T>0} \) share the same curve in the state space \( \mathbb{R} \) which means the two constants \( c_0, c_1 \) are independent of time \( T \). Here we have used the fact that \( x_f \) and \( x_0 \) are distinct metastable states. Furthermore, we obtain that (see appendix A.3 for details)
\[ c_0 \exp \left\{ -\frac{1}{c^2} \mathcal{S}_{OM}^{GT}(\phi_T, \dot{\phi}_T) - c_1 T + \ln(\mu_{\mathcal{B}}(\bar{K}_T(0, \delta))) - \frac{|x_f - x_0| \delta}{T c^2} \right\} = c_0 \exp \left\{ -\frac{h_0}{c^2} + \frac{1}{2} \int_0^{1} b(\phi_1) d\tau \right\} \]  
(3.11)

Notice that the integral term \( \int_0^{1} b(\phi_1) d\tau \) is a constant. This lower bound can be rewritten as
\[ k_0 \exp \left\{ -\frac{k_1}{T} - k_2 T + \ln(\mu_{\mathcal{B}}(\bar{K}_T(0, \delta))) \right\}, \]  
(3.12)

where
\[ k_0 = c_0 \exp \left\{ \frac{1}{c^2} \int_0^{1} b(\phi_1) d\tau \right\}, \]  
\[ k_1 = \frac{(x_f - x_0)^2 + 2|x_f - x_0| \delta}{2c^2}, \]  
\[ k_2 = \left( \frac{1}{2} \int_0^{1} \frac{b^2(\phi_1)}{c^2} d\tau + c_1 \right). \]  
(3.13)

Noting that \( \mu_{\mathcal{B}}(\bar{K}_T(0, \delta)) \leq \mu_{\mathcal{B}}(\bar{K}_T(0, \delta)) \), now we have
\[ \sup_{\psi \in C_0[0,T]} \mu_{\mathcal{H}}(\bar{K}_T(\psi, \delta)) \geq k_0 \exp \left\{ -\frac{k_1}{T} - k_2 T + \ln(\mu_{\mathcal{B}}(\bar{K}_T(0, \delta))) \right\}. \]  
(3.14)

3.3. Estimation of the most probable transition time

We now summarise the estimation on the most probable transition time in the following theorem.

**Theorem 3.3.1 (Upper and lower bounds of the most probable transition time).**

For the stochastic dynamical system (2.1) with two distinct metastable states \( x_0 \) and \( x_f \). Suppose that \( D \) is a bounded connected transition region with \( C^{2,\gamma} \) boundary, and the drift \( b \) is of
Then for every \( \delta > 0 \) with \( 0 < \delta < |x_f - x_0| \), there exist strictly positive upper and lower bounds for the most probable transition time \( T_{x_0 \to x_f}^\delta \).

**Proof.** We derive the upper and lower bounds for the most probable transition time separately.

**Step 1: upper bound.**

Combining the upper and lower bounds in previous sections and for every positive constant \( \epsilon \), we have

\[
k_0 \exp \left\{ -\frac{k_1}{T} - k_2 T + \ln(\mu_B(K_T(0, \delta))) \right\} \leq \sup_{\psi \in C_0[0,T]} \mu_X(\bar{K}_T(\psi, \delta)) \leq \frac{\epsilon}{T},
\]

where the notation \( \leq \) means the inequality holds when \( T > t_\epsilon \) and \( \{ \phi_T \}_{T \geq 0} \) is the family of paths which have been introduced earlier. Define a function \( \Theta \) as

\[
\Theta(t) = k_0 \exp \left\{ -\frac{k_1}{T} - k_2 t + \ln(\mu_B(K_t(\psi, \delta))) \right\}.
\]

Recalling that the probability \( \mu_B(K_T(0, \delta)) \) decreases monotonically from 1 to 0, we know that there exist constants \( T_1^\delta \) and \( T_2^\delta \), with \( 0 < T_1^\delta \leq T_2^\delta < \infty \), such that \( \Theta(t) \) monotonically increases in \( [0, T_1^\delta) \) starting at 0, and monotonically decreases in \( [T_2^\delta, +\infty) \) tending to 0 as \( t \to \infty \). Thus the function \( \Theta(t) \) achieves its maximum value for some \( T' \in [T_1^\delta, T_2^\delta] \).

Define a function \( \Xi(t) \):

\[
\Xi(t) = \sup_{\psi \in C_0[0,T]} \mu_X(\bar{K}_t(\psi, \delta)).
\]

Since for every path \( \psi \in C_0[0,T] \), the probability \( \mu_X(\bar{K}_t(\psi, \delta)) \) is well defined (this has been introduced in subsection 2.1). Thus for each \( t \in [0, \infty) \), the function \( \Xi(t) \) is well defined and it is easy to see that \( 0 \leq \Xi(t) \leq 1 \).

Therefore

\[
\Theta(t) \leq \Xi(t) \leq \frac{\epsilon}{T}, \quad t \in [0, \infty).
\]

For a given constant \( \epsilon > 0 \), denote \( T^\star = \max\{ \frac{\epsilon}{\Xi(T)}, t_\epsilon \} \). Then for every \( t \in [T^\star, \infty) \), we have

\[
\Xi(t) \leq \Theta(T^\star) \leq \Xi(T^\star).
\]

Thus the function \( \Xi(t) \) achieves its maximum value in finite interval \([0, T^\star]\), and the upper bound \( T^\star \) for the MPTP is thus established.

**Step 2: lower bound.**

We now prove that the most probable transition time has a positive lower bound. Notice that \( |x_f - x_0| > \delta \) and the solution process \( X \) of (2.1) is almost surely continuous. If such a lower bound does not exist, then

\[
\sup_{T' \to 0} \sup_{\psi \in C_0[0,T]} P^{x_0}_t \{ \| X - \psi(\cdot) \|_F \leq \delta \}
\]

\[
= \lim_{T' \to 0} \sup_{\psi \in C_0[0,T]} P^{x_0}_t \{ \| X - \psi(\cdot) \|_F \leq \delta \}
\]

\[
\leq \lim_{T' \to 0} P^{x_0}_t \{ \| x_T - x_f \| \leq \delta \}
\]

\[
= 0.
\]
This contradicts the fact that the double supreme value is strictly positive. So the proof is complete.

Theorem 3.3.1 provides a rough range for the most probable transition time:

\[ 0 < \varrho \leq T_{w,x_0}^{0,\delta} \leq T^* < \infty. \]  

(3.21)

Figure 1 offers an intuitive representation of our idea. The left graph of figure 1 shows a schematic plot of the relationship (3.15). The right graph shows the minus logarithm form of the relationship.

### 3.4. Modified action functional for estimating the most probable transition time

In this subsection, we try to find the connection between the double optimization problems on the tube probability (1.6) and on the OM action functional \( S_{OM}^T(\psi) \). Based on this connection, we will present an estimation method for the most probable transition time, under additional assumptions. To achieve this we start with the framework and assumptions of [29].

We focus on a one dimensional gradient system with the drift term \( b(x) = -V'(x) \), for a potential energy \( V(x) \in C^3(\mathbb{R}) \), although some results can be readily extended to the non-gradient case. The path potential \( U(x) \) is given by

\[ U(x) = c^2V''(x) - \frac{1}{2}(V'(x))^2. \]  

(3.22)

Then \( U(x) \in C^1(\mathbb{R}) \) by the smoothness assumption on \( V \). Note that

\[
S_{OM}^T(\psi) = \int_0^T \left( \frac{1}{2}(\dot{\psi})^2 + V'(x)\dot{\psi} - U(\psi) \right) dt \\
= \int_0^T \left( \frac{1}{2}(\dot{\psi})^2 - U(\psi) \right) dt + V(x_f) - V(x_0). 
\]  

(3.23)

We make the following assumptions on the potentials \( V \) and \( U \).
Assumption 3.4.1.

(a) There exists a local minimizer \( x_m \) of \( V(x) \), such that \( V'(x_m) = 0 \), and \( V''(x_m) \) is strictly positive.

(b) The maximizers of \( U(x) \) are contained in a bounded domain.

(c) For every \( E \in \mathbb{R} \), the level set \( \mathcal{L}_E = \{ x \in \mathbb{R} | U(x) = E \} \) can be decomposed into a finite number of closed and connected subsets, i.e.

\[
\mathcal{L}_E = \bigcup_{k=1}^N B_k, \tag{3.24}
\]

where the subsets \( B_k \) are closed and connected, and (pair-wise disjoint) \( B_j \cap B_k = \emptyset \) if \( j \neq k \). The point \( x^* \in \mathbb{R} \) is called a critical point if

\[
\min_{y \in \mathbb{R}} U(y).
\tag{3.25}
\]

(d) The set of critical points, \( \Lambda \), is discrete and has no accumulation points.

(e) For given metastable states \( x_0, x_f \) and final time \( T \), the OM action functional \( S^\text{OM}_T(\psi) \) has a unique minimizer \( \psi_T \).

(f) For given metastable states \( x_0, x_f \) and final time \( T \), there exists a positive constant \( M = M(x_0, x_f) \) such that the minimizer \( \psi_T \) of the OM action functional \( S^\text{OM}_T(\psi) \) satisfies an integral condition, i.e., the ‘speed’ of the minimizing transition path has bounded integral or say the path has finite length:

\[
\int_0^T |\dot{\psi}_T| \, dt \leq M.
\]

Under assumption 3.4.1 the minimizer \( \psi_T \in C^2([0, T], x_0, x_f) \) and satisfies the following Euler–Lagrangian equation:

\[
\begin{cases}
\dddot{\psi}_T + U'(\psi_T) = 0, \\
\psi_T(0) = x_0, \quad \psi_T(T) = x_f.
\end{cases}
\tag{3.26}
\]

The proof of the smoothness of the minimizer can be found in [30]. It is a classical result that the energy of this Euler–Lagrangian equation is conserved along the path \( \psi_T \), i.e.

\[
\frac{1}{2}(\dot{\psi}_T)^2 + U(\psi_T) \equiv E, \quad t \in [0, T].
\tag{3.27}
\]

With assumption 3.4.1 the value \( E \) (or denoted by \( E(T) \)) is uniquely determined by the initial and terminal states \( x_0, x_f \) and the transition time \( T \). For fixed \( x_0 \) and \( x_f \), the value \( E \) is a function of \( T \) only. In this case, \( E \) and \( T \) are related by the equation

\[
T = \int_{\gamma(\psi_T)} \frac{|d\psi_T|}{\sqrt{2E - 2U(\psi_T))}}, \tag{3.28}
\]

where \( \gamma(\psi_T) \) is the graph of \( \psi_T \).

It was proved in proposition 3 of [29] that for every \( T > 0 \), there exists a \( t_c \in [0, T] \) such that

\[
U(\psi_T(t_c)) + \frac{|x_0 - x_f|^2}{2T^2} \leq E(T) \leq U(\psi_T(t_c)) + \frac{M^2}{2T^2}. \tag{3.29}
\]
As also shown in [29], \( \{ \psi_T \}_{T>0} \) is uniformly bounded. Since \( U(x) \) and \( U'(x) \) are continuous, \( \{ U(\psi_T) \}_{T>0} \) and \( \{ U'(\psi_T) \}_{T>0} \) are also uniformly bounded. Therefore, by further combining (3.26), (3.27) and (3.29), we have the following result.

**Theorem 3.4.1 (Uniform boundedness for the most probable transition path).**

(a) The velocity for the MPTP is uniformly bounded after a positive time: for every positive \( \varrho \), the set of ‘speed’ \( \{ \| \dot{\psi}_T \|_T \}_{T>\varrho} \) is uniformly bounded by a constant \( M_{1,\varrho} \).

(b) The acceleration for the MPTP is uniformly bounded: the set of magnitude for the acceleration \( \{ \| \ddot{\psi}_T \|_T \}_{T>0} \) is uniformly bounded by a constant \( M_{2} \).

Recalling the results from the previous section, we can further provide an exact lower bound of \( \sup_{\psi \in C[0,T]} \mu_X(\mathcal{K}_T(\psi, \varrho)) \) by introducing a family of paths \( \{ \phi_T \}_{T>0} \). We use the property that \( \{ \phi_T \}_{T>0} \) and \( \{ \dot{\phi}_T \}_{T>0} \) (in fact \( \dot{\phi}_T \equiv 0 \)) are uniformly bounded. Thus, according to theorems 3.3.1 and 3.4.1 we give a lower bound by replacing \( \{ \phi_T \}_{T>\varrho} \) by \( \{ \psi_T \}_{T>\varrho} \) for some \( \varrho > 0 \):

\[
\sup_{T>\varrho} \sup_{\psi \in C[0,T]} \mathbb{P}^{\mu_0} \left\{ \| X - \psi(\cdot) \|_T \leq \delta \right\} \\
= \sup_{T>\varrho} \sup_{\psi \in C[0,T]} \mathbb{P}^{\mu_0} \left\{ \| X - \psi(\cdot) \|_T \leq \delta \right\} \\
\geq \sup_{T>\varrho} \mathbb{P}^{\mu_0} \left\{ \| X - \psi_T(\cdot) \|_T \leq \delta \right\} \\
\geq \sup_{T>\varrho} \text{exp} \left\{ -M_{1,T} \right\} \\
\times \exp \left\{ -\frac{1}{c_2} S^0_T(\psi_T) - \bar{c}_1 T + \ln(\mu_{B}(\mathcal{K}_0,0,\varrho)) \right\}.
\]

where the coefficients \( \bar{c}_0, \bar{c}_1 \) can be determined in a similar way like \( c_0, c_1 \) in (A.13) and (3.10).

Recall from references [1, 14, 15, 26], the estimation of probability \( \mu_X(\mathcal{K}_T(\psi, \varrho)) \) is

\[
\mu_X(\mathcal{K}_T(\psi, \varrho)) \sim \exp \left\{ -\frac{1}{2} \int_0^T \left[ \frac{\left( \dot{\psi} - b(\psi) \right)^2}{c^2} + b'(\psi) \right] \mu_B(\mathcal{K}_0,0,\varrho) \right\} \mu_{s} T \psi_T(0,\varrho) \)

\[
\sim \frac{4}{\pi} \exp \left\{ -\frac{1}{2} \int_0^T \left[ \frac{\left( \dot{\psi} - b(\psi) \right)^2}{c^2} + b'(\psi) \right] \mu_{s} T \psi_T(0,\varrho) \right\} \delta \downarrow 0.
\]

This approximation has the similar form with the lower bound in (3.30). Note that \( \frac{4}{\pi} \exp \left\{ -\frac{1}{2} S^0_T(\psi_T) \right\} \) is the first term of the infinite series representation of the probability \( \mu_{s} (K_T(0,\varrho)) \) in (2.7).

Moreover, for a fixed constant \( \varrho > 0 \) and when \( T > \varrho \), we have

\[
S^0_T(\psi_T) = \int_0^T \left( \frac{1}{2} \dot{\psi}_T^2 - U(\psi_T) \right) dt + V(x_f) - V(x_0)
\]

\[
= \int_0^T \dot{\psi}_T^2 dt + V(x_f) - V(x_0) - E(T) T.
\]

It was proved in lemma 2 of [29] that

\[
\lim_{T \to +\infty} E(T) > 0.
\]
Hence, if
\[
\frac{\pi^2 c^4}{8\delta^2} > \limsup_{T\to+\infty} E(T),
\] (3.34)
then for each \( \epsilon > 0 \) and when \( T \) is large enough we have
\[
\sup_{v \in C_{0}[0,T]} \frac{4}{\pi} \exp \left\{ -\frac{1}{c^2} S_{T}^{OM}(\psi) - \frac{\pi^2 c^2 T}{8\delta^2} \right\} = \frac{4}{\pi} \exp \left\{ -\frac{1}{c^2} S_{\psi}^{OM}(\psi) - \frac{\pi^2 c^2 T}{8\delta^2} \right\} \leq \epsilon. \] (3.35)
and the following inequality holds for every \( T > 0 \):
\[
k_0 \exp \left\{ -\frac{k_1}{T} - k_2 T + \ln(\mu_{B}(K_{T}(0, \delta))) \right\} \leq \frac{4}{\pi} \exp \left\{ -\frac{1}{c^2} S_{T}^{OM}(\psi) - \frac{\pi^2 c^2 T}{8\delta^2} \right\} \leq \sup_{v \in C_{0}[0,T]} \frac{4}{\pi} \exp \left\{ -\frac{1}{c^2} S_{T}^{OM}(\psi) - \frac{\pi^2 c^2 T}{8\delta^2} \right\}. \] (3.36)
That is
\[
k_0 \exp \left\{ -\frac{k_1}{T} - k_2 T \right\} \mu_{B}(K_{T}(0, \delta)) \leq \sup_{v \in C_{0}[0,T]} \frac{4}{\pi} \exp \left\{ -\frac{1}{c^2} S_{T}^{OM}(\psi) - \frac{\pi^2 c^2 T}{8\delta^2} \right\} \leq \epsilon. \] (3.37)
This inspires us to use the estimation (3.31) to approximately calculate the probability of the solution process \( X \) staying in the neighbourhood of a transition path (although this estimate is quite rough for a fixed \( \delta \)):
\[
P^{\alpha_{0}} \{ \|X - \psi(\cdot)\| \leq \delta \} \approx \frac{4}{\pi} \exp \left\{ -\frac{1}{2} \int_{0}^{T} \left[ \frac{(\dot{\psi} - b(\psi))^2}{c^2} + b'(\psi) \right] \, dt - \frac{\pi^2 c^2 T}{8\delta^2} \right\}. \] (3.38)
We now use it to find the most probable transition time. Define a modified Lagrangian functional \( (L_{mOM}) \) by
\[
L_{mOM}(\psi) := \frac{1}{2} \left[ (\dot{\psi} - b(\psi))^2 + c^2 b'(\psi) + \frac{\pi^2 c^4}{4\delta^2} \right]
\] (3.39)
and the corresponding modified action functional is
\[
S_{T}^{mOM}(\psi) = \frac{1}{2} \int_{0}^{T} \left[ (\dot{\psi} - b(\psi))^2 + c^2 b'(\psi) + \frac{\pi^2 c^4}{4\delta^2} \right] \, dt. \] (3.40)
Thus the double optimization problem on the tube probability (1.6) in this case is approximately equivalent to the following double optimization problem on the modified OM action functional
\[
\inf_{T > 0} \inf_{v \in C_{0}[0,T]} S_{T}^{mOM}(\psi). \] (3.41)
Remark 3.1. We should notice that the original OM action functional comes from the path density functions, while the modified OM action functional is derived from the estimation of the probability that the diffusion process stays in a tube surrounding the transition path.

Remark 3.2. The condition (3.34) is indeed valid for some noise intensity $c$ and tube size $\delta$. This is shown as follows.

Recall the inequalities (3.29):

$$U(\psi_T(t_\epsilon)) + \frac{|x_0 - x_f|^2}{2T^2} \leq E(T) \leq U(\psi_T(t_\epsilon)) + \frac{M^2}{2T^2},$$

thus

$$E(T) \leq c^2 V''(\psi_T(t_\epsilon)) - \frac{1}{2}(V'(\psi_T(t_\epsilon)))^2 + \frac{M^2}{2T^2}.$$  \hspace{1cm} (3.43)

Since $\{\psi_T\}_{T>0}$ are uniformly bounded and the potential $V(x)$ is smooth enough, thus when $T$ is large enough, there exist some $c$ and $\delta$ such that

$$E(T) \leq c^2 V''(\psi_T(t_\epsilon)) - \frac{1}{2}(V'(\psi_T(t_\epsilon)))^2 + \frac{M^2}{2T^2} < \frac{\pi^2c^4}{8\delta^2}.$$  \hspace{1cm} (3.44)

Furthermore we have

$$\inf_{T>0} \inf_{\psi \in \bar{C}[0,T]} S_{T}^{\text{SMOM}}(\psi) = \inf_{T>0} S_{T}^{\text{OM}}(\psi_T) = \inf_{T>0} \left( S_{T}^{\text{OM}}(\psi_T) + \frac{\pi^2c^4T}{8\delta^2} \right),$$

and

$$\frac{d}{dT} \left( S_{T}^{\text{OM}}(\psi_T) + \frac{\pi^2c^4T}{8\delta^2} \right) = \frac{dS_{T}^{\text{SMOM}}(\psi_T)}{dT} + \frac{\pi^2c^4}{8\delta^2} = -E(T) + \frac{\pi^2c^4}{8\delta^2}.$$  \hspace{1cm} (3.46)

Here the relation $\frac{dS_{T}^{\text{OM}}(\psi_T)}{dT} = -E(T)$ is a classical result in Hamilton–Jacobi theory whose proof can be found in [29, 30]. Thus if $E(T)$ is monotonic, then $T^{\text{opt}}_{\text{MPTP}} = E^{-1}(\frac{\pi^2c^4}{8\delta^2})$. This implies that the global MPTP roughly lies on the energy shell $E = \frac{\pi^2c^4}{8\delta^2}$ in phase space.

4. Examples

In this section we present two examples to illustrate our results.

Example 1. One-dimensional Brownian motion

Consider a scalar SDE without drift:

$$dX_t = c \, dB_t, \quad 0 \leq t \leq T,$$

where $c$ is a positive constant. The modified Lagrange function is

$$L^{\text{mOM}}(\dot{\psi}, \psi) = \frac{1}{2}(\dot{\psi})^2 + \frac{\pi^2c^4}{8\delta^2}.$$  \hspace{1cm} (4.2)

and the modified OM action functional is

$$S_{T}^{\text{mOM}}(\psi_T) = \left( \frac{x_f - x_0}{T} \right)^2 + \frac{\pi^2c^4T}{8\delta^2}.$$  \hspace{1cm} (4.3)
Figure 2. Top: sample paths of a Brownian motion. Middle: sample paths of the stochastic double well system (4.5). Bottom: the solutions of the Euler–Lagrange equations (4.9) (finite noise) and (4.10) (vanishing noise) for the stochastic double well system by a shooting method, with $x_0 = -1, x_f = 1$ and $T = 10$.

So by minimizing this functional (setting its first derivative with respect to time $T$ to be zero), we obtain the estimation for the most probable transition time $T_d^{\delta}$

$$T_d^{\delta} = \frac{2\delta |x_f - x_0|}{\pi c^2}. \quad (4.4)$$

Example 2. A stochastic double well system

Consider a nonlinear scalar SDE:

$$dX_t = (X_t - X_t^3)dt + c dB_t, \quad (4.5)$$

with $c$ is a positive constant (without loss of generality we set $c = 1$). The corresponding undisturbed system has three equilibrium points: $-1, 0, 1$ (we know that $-1$ and $1$ are stable equilibrium points, and $0$ is an unstable equilibrium point). In figure 2, the middle graph
Figure 3. Two sample transition paths of the stochastic double well system (4.5). Black: the transition time of sample path 1 is 0.429 and the tube size is \( \delta = 0.46 \); blue: the transition time of sample path 2 is 1.449 and the tube size is \( \delta = 1.21 \).

shows two sample paths of system (4.5) with initial position \( x_0 = -1 \). The top graph of figure 2 shows two sample paths of Brownian motion staring at 0 as a contrast. From the comparison of these two graphs it is easy to see the difference of the behaviour of a diffusion process with different drift terms. The diffusion process (4.5) fluctuates between two metastable states \(-1\) and \(1\). In figure 3, there are two sample transition paths of this stochastic double well system. The red and yellow curves are the corresponding MPTPs, and the dash curves are the boundaries of the transition tube.

The corresponding modified Lagrangian \( L^{mOM} \) is

\[
L^{mOM}(\dot{\psi}, \psi) = \frac{1}{2} \left[ (\dot{\psi} - \psi + \psi^3)^2 + c^2(1 - 3\psi^2) + \frac{\pi^2 c^4}{4\delta^2} \right],
\]

(4.6)

and the modified OM action functional is

\[
S^{mOM}_T(\psi) = \frac{1}{2} \int_0^T \left[ (\dot{\psi} - \psi + \psi^3)^2 + c^2(1 - 3\psi^2) + \frac{\pi^2 c^4}{4\delta^2} \right] ds.
\]

(4.7)

We use Euler method to generate sample solution paths with time step size \( \Delta t = 10^{-4} \). To solve the optimization problem, we assume that the minimizer is twice differentiable and thus we obtain the Euler–Lagrange equation:

\[
\frac{d}{dr} \frac{\partial L^{mOM}(\dot{\psi}, \psi)}{\partial \dot{\psi}} = \frac{\partial L^{mOM}(\dot{\psi}, \psi)}{\partial \psi}.
\]

(4.8)
Hence the optimization problem turns into a second order ordinary differential equation with two boundary values, which can be solved numerically by a shooting method:

$$\begin{cases} \dddot{\psi} = b'(\psi)b(\psi) + \frac{c^2}{2}b''(\psi), \\ \psi(0) = x_0, \quad \psi(T) = x_f. \end{cases}$$

(4.9)

As a contrast, the FW Lagrangian is $L_{\text{FW}}(\dot{\psi}, \psi) = \frac{1}{2}(\dot{\psi}^2 + \psi^3)$ (which will be introduced in next section 5.3), thus the Euler–Lagrangian equation for this double well system with vanishing noise is

$$\begin{cases} \ddot{\psi} = b'(\psi)b(\psi), \\ \psi(0) = x_0, \quad \psi(T) = x_f. \end{cases}$$

(4.10)

We should note that the Euler–Lagrangian equation determines the local minimizer. That is the Euler–Lagrangian equation is a necessary but not sufficient description for the MPTP. A shortcoming of the shooting method is restricted by the selection of the initial iteration parameter (see [31]). In [31], a machine learning framework of the shooting method was developed to overcome this shortcoming. The bottom graph of figure 2 shows the numerical solutions of these two Euler–Lagrangian equations (4.9) and (4.10) calculated by a shooting method with $T = 10$. This graph shows an interesting difference between the two cases: the solution of (4.9) exhibits three forward and two backward transitions but the solution of (4.10) has only one transition. This difference seems indicates that the transition occurs in relative short time interval (or say transition is more easy to occur) for the stochastic dynamical systems with non-vanishing noise.

In this example, we focus on the tube sizes that smaller than 1. Since the domain $B_{\delta}(x_f)$ only contains one equilibrium point of the undisturbed system when $\delta < 1$. And from the subsection 3.4 we know that our estimations for the most probable transition time make no sense when the tube size $\delta$ is too large. Figure 4 shows the graphs of $S_p^{\text{OM}}(\psi_T)$, Figure 5 shows the MPTPs for different tube sizes $\delta$ and transition time $T$. It indicates that for tube sizes smaller than 1, the estimation of most probable transition time is bounded by 1.5. So we restrict our attention in relative short time interval $[0, 1.5]$ and the first transition behaviour. The paths in
Figure 5. The ‘MPTPs’ with different transition time calculated by shooting method.

Figure 6. The transition times and tube sizes of 3135 transition paths.

figure 5 are uniformly bounded in domain $[-1, 1]$, so in this example we use modified action functional to characterize the transition behaviour. The numerical result of the most probable transition times $T^D_{s_i s_j}$ for different $\delta$ is obtained which are shown in figure 6 in red spots.

Furthermore, we simulated 30000 sample paths starting at $-1$ and there are 3135 transition paths (the first transition occurred before time 1.5). For every transition path we recorded its transition time $T$ and calculated its corresponding MPTP and tube size $\delta$ (i.e. this transition path can be contained in the tube of this MPTP with this tube size). All the pairs $(T, \delta)$ were draw in figure 6 by blue stars. The yellow squares are the mean values of the tube sizes. We separated the time interval $[0, 1.5]$ into subintervals $[0, 0.25), [0.25, 0.3), [0.3, 0.35), [0.35, 0.4), \ldots, [1.45, 1.5]$ and calculated the mean tube size values in every subinterval. As shown in figure 6, this numerical simulation shows that our results characterize the transition behaviour.

5. Discussion

We now summarize our work and highlight the differences with relevant works.

5.1. Our contribution

Under some mild assumptions, we have estimated the most probable transition time between metastable states, for stochastic dynamical systems with non-vanishing Brownian noise. The
problem is represented by a double optimization on the probability that sample paths staying in a tube surrounding the most probable transition pathway. We have provided estimates for the most probable transition time.

In our framework, we have adopted the original idea of OM framework [3], using the concept of tubes surrounding the transition path to study the transition time. Instead of letting the tube size $\delta$ tending to 0, we require this tube size $\delta$ to be a positive constant. This promises the double optimization problem is well-defined. Since when $\delta \to 0$ all tube probabilities are 0. This requirement makes sense as the noise intensity is non-vanishing, which is different from the case in Freidlin–Wentzell’s (FW) large deviation theory [32, 33].

Our estimate can be extended to higher dimensional systems. Similarly, lemma 2.2.1 shows that the probability that the solution process stays in the tube of a transition path has a power decay law upper bound. It is similar to the one-dimensional case that the most probable transition time has upper and lower bounds. The probability of the solution process staying in the tube of a transition path can be computed approximately in higher dimensional case:

$$\mathbb{P}^x \{ \| X - \psi(\cdot) \|_T \leq \delta \} \approx \exp \left( -\frac{S^{OM}(\psi)}{\epsilon^2} \right) \cdot \mathbb{P}^x \{ \| B^c - x_0 \|_T \leq \delta \}, \quad \delta \downarrow 0. \quad (5.1)$$

The probability $\mathbb{P}^x \{ \| B^c - x_0 \|_T \leq \delta \}$ monotonically decreases in $T$. So we could define a modified OM action functional, if we have an appropriate analytical estimation for the probability $\mathbb{P}^x \{ \| B^c - x_0 \|_T \leq \delta \}$.

There are some works related to the heuristic discussions and measurements of transition time. These include the transition time distribution [34–37] and the expected transition time under small noise intensity [38].

5.2. The difference between double optimizations of tube probability and of OM action functional

A double optimization problem on the OM action functional was investigated in [29]:

$$\inf_{T > 0} \inf_{\psi \in \bar{C}[0,T]} S^{OM}_T(\psi), \quad (5.2)$$

where $\bar{C}[0, T]$ denotes the space of all absolutely continuous functions $g : [0, T] \to D$ such that $g(0) = x_0$ and $g(T) = x_f$. This problem aims to study the property of the OM action functional. Although [29] and we focus on the same stochastic dynamical systems, our result is different. Du et al [29] focused on the OM action functional in the $\delta \to 0$ scaling, and the effect of the probability $\mu_{B^c}(\text{KL}(0, \delta))$ is ignored when $T$ varies.

When transition time $T$ is fixed, the MPTP is determined by the OM functional, $S^{OM}_T$ (finite noise cases) or the FW function (vanishing noise cases). Note that the OM functional was derived from Radon–Nikodym derivative of two measures by Girsanov theorem [1, 14, 15, 26]. Thus $\exp(-S^{OM}_T)$ can be considered as the path densities in path space. However when the transition time $T$ varies, the associate path spaces are different, so we avoid minimizing OM function in $T$ directly and minimize the probabilities of the transition tubes. Since the comparisons between probabilities make sense when $T$ varies. This is one of our motivations to study the double optimization problem (1.6).

Under the setting in [29] (see the assumption 3.4.1), the OM functional is unbounded from below (proposition 3 in [29]):

$$\inf_{T > 0} \inf_{\psi \in \bar{C}[0,T]} S^{OM}_T(\psi) = \lim_{T \to +\infty} \inf_{\psi \in \bar{C}[0,T]} S^{OM}_T(\psi) = -\infty. \quad (5.3)$$
However the assumption 3.4.1-(f) assumes that all the MPTPs \( \{\psi_T\}_{T>0} \) have uniform finite length. Thus they can be contained in a bounded domain \( D \). As shown in section 3, we know that the probabilities of the diffusion process \( X \) stays in the \( \delta \)-neighbourhoods of these MPTPs tend to 0 when \( T \) tends to infinity. Thus our results emphasize that for stochastic dynamical systems with non-vanishing noise, the transition time problem should also focus on finite time cases.

In the estimation (3.31), it can be seen that the OM action functional term characterizes the geometrical features of a transition path, while the term \( \mu_{Bc}(K_T(0, \delta)) \) characterizes the diffusion ability of the path. So for a stochastic system with non-vanishing Brownian noise, the MPTP and the most probable transition time are supposed to be determined by these two terms (in the original OM context). Although [29] ignored the effect of \( \mu_{Bc}(K_T(0, \delta)) \), it provides valuable insights on the OM action functional.

5.3. The difference between our work and the large deviation theory

Transition phenomena have been treated in the large deviation theory (i.e., under sufficiently small noise). The large deviation theory focuses on the following system [32]:

\[
dX_t^\varepsilon = b(X_t^\varepsilon)dt + \sqrt{\varepsilon} dB_t, \quad t \geq 0, \quad X_0^\varepsilon = x_0 \in \mathbb{R}^k, \quad (5.4)
\]

where \( \varepsilon \to 0 \). The FW action functional \( S_{FW}^T(\psi) \) is

\[
S_{FW}^T(\psi) = \int_0^T L_{FW}(\dot{\psi}, \psi)dt,
\]

if \( \psi \in C(0, T) \) is absolutely continuous and the integral converges, otherwise denote that \( S_{FW}^T(\psi) = \infty \). Here the Lagrangian \( L_{FW}(x, y) \) is given by

\[
L_{FW}(x, y) = \langle y - b(x), y - b(x) \rangle.
\]

When the transition time was considered as a factor in transition phenomena, the quasi-potential in [33] was defined as

\[
V(x_0, x_f) = \inf_{T>0, \psi \in C[0, T]} \inf_{\psi \in \bar{C}[0, T]} S_{FW}^T(\psi),
\]

which is used to find the ‘global’ MPTP and the corresponding most probable transition time. The large derivation theory asserts that, for \( \delta \) and \( \varepsilon \) sufficiently small,

\[
P_{x_0}^\varepsilon \left\{ \sup_{0 \leq t \leq T} |X_t^\varepsilon - \psi(t)| \leq \delta \right\} \approx \exp \left( -\varepsilon^{-1} S_{FW}^T(\psi) \right).
\]

It was shown in [33] that the most probable transition time between two metastable states of system (5.4) is infinite. Since the noise intensity \( \sqrt{\varepsilon} \) tends to 0, the behaviour of the system is closely influenced by that of the deterministic system. Hence the transition between two distinct metastable states needs infinite time to make it happen most likely.

The modified OM functional in this small noise case is

\[
\lim_{c \to 0} S_{cOM}^T(\psi) = \lim_{c \to 0} \frac{1}{2} \int_0^T \left[ (\dot{\psi} - b(\psi))^2 + c^2 b'(\psi) + \frac{\pi^2 c^4}{4 \delta^2} \right] dt = S_{FW}^T(\psi).
\]

(5.9)
In particular, the estimation of the most probable transition time \( T_{x_0 \to x_f} \) for example 1 in section 4 is thus

\[
T_{x_0 \to x_f}^{D, \delta} = \frac{2\delta |x_f - x_0|}{\pi c^2}.
\]  

(5.10)

When \( c \to 0 \) then \( T_{x_0 \to x_f}^{D, \delta} \to \infty \) for every \( \delta > 0 \). This result is consistent with the large deviation theory.

Acknowledgments

The authors would like to thank Jianyu Hu, Pingyuan Wei, Xiujun Cheng, Qiao Huang, Ao Zhang and Yancai Liu for helpful discussions. This work was partly supported by the NSFC Grants 11531006 and 11771449.

Appendix A

A.1. Non-exit probability for scalar Brownian motions

Define \( G = \{ x \in \mathbb{R} | |x| < 1 \} \) and \( \tau_{c}^{0}(B_{t}^{c}) = \inf \{ t | B_{t}^{c} \notin G, B_{0}^{c} = x \in G \} \). Then \( u(t,x) = E^{x}[f(B_{t}^{c}) | \tau_{c}^{0}(B_{t}^{c}) > t] \), \( x \in G, t > 0 \), is the solution of the initial value problem

\[
\begin{cases}
\frac{\partial u}{\partial t} = \frac{1}{2} c^2 \Delta u, & u \in G, \\
|u|_{G} = 0, & u|_{t=0} = f.
\end{cases}
\]  

(A.1)

Consequently

\[
u(t,x) = \sum_{n=1}^{\infty} e^{-\lambda_n t} \phi_n(x) \int_{G} \phi_n(y) f(y) dy,
\]  

(A.2)

where \( 0 < \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \) are eigenvalues and \( \{ \phi_n(x) \} \) are corresponding eigenfunctions of the eigenvalue problem

\[
\begin{cases}
\frac{1}{2} c^2 \triangle \phi + \lambda \phi = 0 & \text{in } G, \\
\phi|_{G} = 0.
\end{cases}
\]  

(A.3)

The eigenfunctions (which need to be nonzero) are \( \{ \sin(m \pi x) \}_{m=1}^{\infty} \) and \( \{ \cos(\frac{1}{2} m \pi x) \}_{m=0}^{\infty} \), with the corresponding eigenvalues

\[
\begin{align}
\left\{ \frac{m \pi^2}{2} c^2 \right\}_{m=1}^{\infty}, & \quad \left\{ \frac{(m + \frac{1}{2}) \pi^2}{2} c^2 \right\}_{m=0}^{\infty}.
\end{align}
\]

The set of normalized eigenfunctions form an orthonormal basis for the Hilbert space \( \mathcal{H} = L^2(-1, 1) \).

In particular,

\[
\mu_{B}(K_{T}(0, \delta)) = \mathbb{P}_{0}(\|B_{T}^{c} - 0\| < \delta)
\]

\[
= \sum_{n=0}^{\infty} e^{-\lambda_n T/\delta^2} \phi_n(0) \int_{G} \phi_n(y) dy
\]

4562
\[
\sum_{n=0}^{\infty} \frac{(-1)^n 4}{(2n+1)\pi} \exp \left\{ -\frac{(2n+1)^2\pi^2 c^2 T}{8\delta^2} \right\}.
\]

(A.4)

### A.2. The Wiener measure of a closed tube set

We start with the following lemma.

**Lemma A.2.1 (Approximation on cylinder sets for the measure \( \mu_X \) of a closed tube).** For a fixed transition time \( T \), a smooth path \( \psi \in C([0, T], x_0) \) and a tube size \( \delta \), there exists a sequence of sets \( \{J_n\}_{n=1}^{\infty} \) in the field \( I \), such that

\[
\mu_X(K_T(\psi, \delta)) = \lim_{n\to\infty} \mathbb{P}^\delta(J_n).
\]

(A.5)

**Proof.** Let \( Q \) be the countable set of rational numbers in \( \mathbb{R} \). So for a fixed transition time \( T > 0 \), the set \( Q \cap [0, T] \) is also countable. Define

\[
Q \cap [0, T] = \{q_0, q_1, \ldots, q_n, \ldots\}.
\]

(A.6)

If \( T \) is a rational number, we switch the positions of 0 and \( T \) with the ones of \( q_0 \) and \( q_1 \) respectively and still represent the sequence as \( \{q_0 = 0, q_1 = T, q_2, \ldots, q_n, \ldots\} \). If \( T \) is an irrational number we switch the position of 0 with the one of \( q_0 \) and add \( T \) into the sequence and denote the sequence as the same representation \( \{q_0 = 0, q_1 = T, q_2, \ldots, q_n, \ldots\} \).

Introduce the subsets

\[
O_n = \{q_0, q_1, \ldots, q_n\}, J_n = \{X \in C([0, T], x_0) | |X_q - \psi(q)| \leq \delta, \ \forall \ q \in O_n \} \subset I.
\]

(A.7)

Note that \( \{J_n\}_{n=1}^{\infty} \) is a decreasing sequence. Recall that the solution process of (2.1) is almost surely continuous. So

\[
\mu_X(K_T(\psi, \delta)) = \mathbb{P}^\delta(\{w \in \Omega | \sup_{t \in [0, T]} |X_t(w) - \psi(t)| \leq \delta\})
\]

\[
= \mathbb{P}^\delta(\{w \in \Omega | \sup_{t \in Q \cap [0, T]} |X_t(w) - \psi(t)| \leq \delta\})
\]

\[
= \mathbb{P}^\delta(\cap_{n=1}^{\infty} J_n)
\]

\[
= \lim_{n \to \infty} \mathbb{P}^\delta(J_n).
\]

(A.8)

This completes the proof of this lemma.

For simplicity we demonstrate the calculation in one dimensional case. Now we discretize \( X_t \) of (2.1) in the following way:

\[
\begin{cases}
X_i - X_{i-1} = [\kappa b(X_i) + (1 - \kappa)b(X_{i-1})] \Delta t + \epsilon(B_i - B_{i-1}), i = 1, \ldots, n, \\
X_0 = x_0,
\end{cases}
\]

(A.9)

For simplicity we consider the following discretized version which has the same distribution of (A.9):

\[
\begin{cases}
X_i - X_{i-1} = [\kappa b(X_i) + (1 - \kappa)b(X_{i-1})] \Delta t + B_i - B_{i-1}, i = 0, 1, \ldots, n, \\
X_0 = x_0,
\end{cases}
\]

(A.10)
where \((B'_t(u) - B'_t(w)) \sim N(0, c^2(t-s))\) and the time partition is \(\Pi_n\): \(0 = t_0 < t_1 = \ldots < t_n = T\) and \(\kappa \in (0,1]\). For every \(\psi \in C^2_T[0, T] := \{\psi \in C^2_T[0, T]|\psi\) has all continuous derivatives of order no more than 2\}, and using lemma A.2.1, the probability \(\mu_{\kappa}(\mathcal{K}_T(\psi, \delta))\) is calculated in the following way (for every \(O_n\) in lemma A.2.1, we reorder the elements of \(O_n\) from small to large and use the new sequence for the time partition \(\Pi_n\)):

\[
\mu_{\kappa}(\mathcal{K}_T(\psi, \delta)) = \mu_{\kappa}(\sup_{t \in [0, T]} |X_t - \psi_t| \leq \delta)
\]

\[
= \lim_{n \to \infty} \mathcal{D}_n(w)[X_{t_0}, X_{t_1}, \ldots, X_{t_n}] \in \mathcal{I}_n)
\]

\[
= \lim_{n \to \infty} \int \mathcal{D}_n \mathcal{B}_x \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \left( \frac{(B'_i - B'_{i-1})^2}{c^2 \Delta t} \right) \right\}
\]

\[
\approx \lim_{n \to \infty} \int \mathcal{D}_n \mathcal{B}_x \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \left( \frac{(x_i - x_{i-1} - \kappa b(x_i + \psi_i) \Delta t)^2}{2c^2 \Delta t} \right) \right\}
\cdot \prod_{i=1}^n \left( 1 + \kappa b(x_i) \Delta t \right)
\]

\[
= \lim_{n \to \infty} \int \mathcal{D}_n \mathcal{B}_x \prod_{i=1}^n \left( 1 + \kappa b(y_i + \psi_i) \Delta t \right)
\cdot \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \left[ \frac{(y_i - y_{i-1} + \psi_i - \psi_{i-1} - \kappa b(y_i + \psi_i) + (1 - \kappa) b(y_{i-1} + \psi_{i-1}) \Delta t)^2}{c^2 \Delta t} \right] \right\}
\]

\[
= \int_{y \leq \delta} \mathcal{D}_n \mathcal{Y}_x \exp \left\{ -\frac{1}{2} \int_0^T \left[ \frac{(y + \psi - b(y + \psi))^2}{c^2} + 2 \kappa b'(y + \psi) \right] \right\}
\]

\[
= \int_{y \leq \delta} \mathcal{D}_n \mathcal{Y}_x \exp \left\{ \frac{U(y + \psi)|y < \delta|}{c^2} - \frac{1}{2} \int_0^T \left[ \frac{y^2 + 2y\psi + \psi^2 + b^2(y + \psi)}{c^2} + 2 \kappa b'(y + \psi) \right] \right\}
\]

\[
= \exp \left\{ \frac{1}{2} \int_0^T \left[ (\psi - b(\psi))^2 + 2 \kappa \bar{b}'(\psi) \right] \right\} \int_{y < \delta} \mathcal{D}_n \mathcal{Y}_x \exp \left\{ \frac{U(y(T) + \psi)}{c^2} - \frac{1}{2} \int_0^T \left[ \frac{y^2 + 2y\psi + \psi^2 + b^2(y + \psi)}{c^2} + 2 \kappa b'(y + \psi) \right] \right\} \),
\]

where \(\mathcal{I}_n = [\psi_0 - \delta, \psi_0 + \delta] \times [\psi_1 - \delta, \psi_1 + \delta] \times \cdots \times [\psi_n - \delta, \psi_n + \delta]\) and \(\mathcal{D}_n \mathcal{B}_x = \prod_{i=1}^n \frac{d\mathcal{B}_x}{\sqrt{2c^2 \Delta t}}, \mathcal{D}_n \mathcal{Y}_x = \lim_{n \to \infty} \prod_{i=1}^n \frac{d\mathcal{Y}_x}{\sqrt{2c^2 \Delta t}}\) (the notation \(\mathcal{D}_n \mathcal{B}_x\) and \(\mathcal{D}_n \mathcal{Y}_x\) are defined in a similar way) and \(U(x) = \int_0^x b(y) dy\).

**Remark 1.** In (A.11) we have used the results of the path integral method for more mathematical details, such as the existence of the limitation and the links between discrete approximation and continuous functions, see [40].
For a path $\psi \in C_D([0, T])$, denote that

\[
\gamma(\psi, \delta) := \bigcup_{t \in [0, T]} \{ x \in \mathbb{R} \mid |x - \psi(t)| < \delta \} \subseteq \mathbb{R},
\]

\[
h_0 := \sup_{|x - x_f| < \delta} |\mathcal{U}(x) - \mathcal{U}(x_f)|,
\]

\[
h_1(\psi, \delta) := \sup_{x \in \gamma(\psi, \delta)} |b(x)b'(x)|,
\]

\[
h_2(\psi, \delta) := \sup_{x \in \gamma(\psi, \delta)} |b''(x)|.
\]

(A.12)

Now we give the following estimations:

\[
\left| \int_0^T \dot{y} \psi \, dt \right| = \left| y(T)\dot{\psi}(T) - y(0)\dot{\psi}(0) - \int_0^T y\ddot{\psi} \, dt \right|
\]

\[
= \left| y(T)\dot{\psi}(T) - \int_0^T y\ddot{\psi} \, dt \right| \leq \|\ddot{\psi}\|_T \delta \delta + \|\ddot{\psi}\|_T \delta,
\]

\[
\left| \int_0^T b^2(y + \psi) - b^2(\psi) \, dt \right| = \lim_{n \to \infty} \sum_{i=1}^n \left| b^2(y_{i-1} + \psi_{i-1}) - b^2(\psi_{i-1}) \right| \Delta t
\]

\[
= \lim_{n \to \infty} \sum_{i=1}^n \left[ 2b(\theta^1_{i-1}y_{i-1} + \psi_{i-1}) \times b'(\theta^1_{i-1}y_{i-1} + \psi_{i-1}) \right] \Delta t
\]

\[
\leq \lim_{n \to \infty} \sum_{i=1}^n \left[ 2b(\theta^1_{i-1}y_{i-1} + \psi_{i-1}) \times b'(\theta^1_{i-1}y_{i-1} + \psi_{i-1}) \right] \Delta t
\]

\[
\leq 2h_1 \delta \lim_{n \to \infty} \sum_{i=1}^n \Delta t
\]

\[
= 2h_1 \delta T,
\]

(A.13)

\[
\left| \int_0^T b'(y + \psi) - b'(\psi) \, dt \right| = \lim_{n \to \infty} \sum_{i=1}^n \left| b'(y_{i-1} + \psi_{i-1}) - b'(\psi_{i-1}) \right| \Delta t
\]

\[
= \lim_{n \to \infty} \sum_{i=1}^n \left[ b''(\theta^2_{i-1}y_{i-1} + \psi_{i-1}) \left( y_{i-1} + \psi_{i-1} \right) \right] \Delta t
\]

\[
\leq \lim_{n \to \infty} \sum_{i=1}^n \left[ b''(\theta^2_{i-1}y_{i-1} + \psi_{i-1}) \left( y_{i-1} + \psi_{i-1} \right) \right] \Delta t
\]

\[
\leq h_2 \delta \lim_{n \to \infty} \sum_{i=1}^n \Delta t
\]

\[
= h_2 \delta T,
\]

4565
where \( \theta_i \in [0, 1] \) for \( j = 1, 2 \) and \( i = 1, 2, \ldots, n \), and we have used the well known mean value theorem.

Substituting the estimations (A.13) into (A.11), we obtain that

\[
\mu_X(K_T(\psi, \delta)) \geq \exp \left\{ -\frac{1}{c^2} \left[ \| \dot{\psi}(T) \| \delta + h_0 + \left( h_1 + \frac{1}{2} c^2 h_2 + \| \ddot{\psi} \| T \delta \right) \right] \right\} \cdot \exp \left\{ -\frac{1}{c^2} S_{OM}(\psi) \right\} \mu_{B^c}(K_T(0, \delta)),
\]

where \( S_{OM}(\psi) = \int_0^T L_{OM}(\psi, \dot{\psi}) dt = \int_0^T [\| \dot{\psi} - b(\psi) \| + 2 \kappa c^2 b'(\psi)] dt. \)

**Remark 2.** When \( \kappa = \frac{1}{2} \), the functional \( S_{OM}(\psi) \) is the OM action functional. The parameter \( \kappa \) sometimes plays an important role in studying the dynamical behaviour of stochastic systems. For example, to describe time-reversible dynamics, the effective action based on the symmetrical (Stratonovich’s, \( \kappa = \frac{1}{2} \)) interpretation is applied to the system with additive noise [4, 5]. In [39], an experiment on a Brownian particle near a wall suggests that the system favours the anti-Itô’s (\( \kappa = 1 \)) interpretation rather than the Stratonovich’s, to ensure the Boltzmann–Gibbs distribution for the final steady state.

Using the parameter \( \kappa = \frac{1}{2} \),

\[
\mu_X(K_T(\psi, \delta)) \geq \exp \left\{ -\frac{1}{c^2} \left[ \| \dot{\psi}(T) \| \delta + h_0 + \left( h_1 + \frac{1}{2} c^2 h_2 + \| \ddot{\psi} \| T \delta \right) \right] \right\} \cdot \exp \left\{ -\frac{1}{c^2} S_{OM}(\psi) \right\} \mu_{B^c}(K_T(0, \delta)).
\]

This coincides with the results in [1, 26]. For higher dimensional cases (\( k > 1 \)), we are able to obtain that

\[
\mu_X(K_T(\psi, \delta)) \geq \exp \left\{ -\frac{k}{c^2} \left[ \| \dot{\psi}(T) \| \delta + h_0 + \left( h_1 + \frac{1}{2} c^2 h_2 + \| \ddot{\psi} \| T \delta \right) \right] \right\} \cdot \exp \left\{ -\frac{1}{c^2} S_{OM}(\psi) \right\} \mu_{B^c}(K_T(0, \delta)).
\]

**A.3. Calculation of the lower bound**

\[
c_0 \exp \left\{ -\frac{1}{c^2} S_{OM}(\phi_T, \dot{\phi}_T) - c_1 T + \ln(\mu_{B^c}(\bar{K}_T(0, \delta))) - \frac{|x_f - x_0| \delta}{T c^2} \right\}
\]

\[
= c_0 \exp \left\{ -\frac{1}{2} \int_0^T \left[ \frac{(\dot{\phi}_T - b(\phi_T))^2}{c^2} + b'(\dot{\phi}_T) \right] dt - c_1 T \right.
\]

\[
+ \ln(\mu_{B^c}(K_T(0, \delta))) - \frac{|x_f - x_0| \delta}{T c^2} \right\}
\]

\[
= c_0 \exp \left\{ -\frac{1}{2} \int_0^T \left[ \frac{(\dot{\phi}_T - b(\phi_T))^2}{c^2} + b'(\dot{\phi}_T) \right] dt \right. \]
\[-c_1 T + \ln (\mu_B (K_T(0, \delta))) = \frac{|x_f - x_0| \delta}{T c^2}\]

\[
= c_0 \exp \left\{ -\frac{(x_f - x_0)^2}{2 T c^2} + \frac{(x_f - x_0)}{T c^2} \int_0^T b(\phi_T) dt \right. \\
- \frac{1}{2} \int_0^T \left[ \frac{b^2(\phi_T)}{c^2} + b'(\phi_T) \right] dt - c_1 T \\
+ \ln (\mu_B (K_T(0, \delta))) - \frac{|x_f - x_0| \delta}{T c^2} \right\}
\]

\[
= c_0 \exp \left\{ -\frac{(x_f - x_0)^2}{2 T c^2} + \frac{(x_f - x_0)}{c^2} \int_0^1 b(\phi_1) dt \\
- \frac{T}{2} \int_0^1 \left[ \frac{b^2(\phi_1)}{c^2} + b'(\phi_1) \right] dt - c_1 T \\
+ \ln (\mu_B (K_T(0, \delta))) - \frac{|x_f - x_0| \delta}{T c^2} \right\}
\]

\[
= c_0 \exp \left\{ -\frac{(x_f - x_0)^2}{2 T c^2} + \frac{(x_f - x_0)}{c^2} \int_0^1 b(\phi_1) dt \\
- \frac{1}{2} \int_0^1 \frac{b^2(\phi_1)}{c^2} dt - \left( \frac{1}{2} \int_0^1 b(\phi_1) dt + c_1 \right) T + \ln (\mu_B (K_T(0, \delta))) \right\}. \tag{A.17}
\]

Here we use the fact that

\[
\phi_T(t) = \phi_1 \left( \frac{t}{T} \right), \quad t \in [0, T], \tag{A.18}
\]

and

\[
\int_0^T b(\phi_T(t)) dt = T \int_0^1 b \left( \phi_1 \left( \frac{t}{T} \right) \right) d \frac{t}{T} = T \int_0^1 b(\phi_1(t)) dt. \tag{A.19}
\]

**ORCID iDs**

Jinqiao Duan \(\text{https://orcid.org/0000-0002-2077-990X}\)

**References**

[1] Dürr D and Bach A 1978 The Onsager–Machlup function as Lagrangian for the most probable path of a diffusion process *Commun. Math. Phys.* 60 153
[2] Hunt K L C and Ross J 1981 Path integral solutions of stochastic equations for nonlinear irreversible processes: the uniqueness of the thermodynamic Lagrangian J. Chem. Phys. 75 976
[3] Machlup S and Onsager L 1953 Fluctuations and irreversible process II. Systems with kinetic energy Phys. Rev. 91 1512
[4] Faccioli P, Sega M, Pederiva F and Orland H 2006 Dominant pathways in protein folding Phys. Rev. Lett. 97 108101
[5] Zuckerman D M and Woolf T B 2000 Efficient dynamic importance sampling of rare events in one dimension Phys. Rev. E 63 016702
[6] Wang J, Zhang K, Lu H and Wang E 2006 Dominant kinetic paths on biomolecular binding-folding energy landscape Phys. Rev. Lett. 96 168101
[7] Gobbo G, Laio A, Maleki A and Baroni S 2012 Absolute transition rates for rare events from dynamical decoupling of reaction variables Phys. Rev. Lett. 109 150601
[8] Langouche F, Rockaerts D and Tirapegui E 1982 Functional Integration and Semiclassical Expansions (Berlin: Springer)
[9] Schulman L S 1981 Techniques and Applications of Path Integration (New York: Wiley)
[10] Wiegel F W 1986 Introduction to Path Integral Methods in Physics and Polymer Science (Singapore: World Scientific)
[11] Wio H S 2013 Path Integrals for Stochastic Processes: An Introduction (Singapore: World Scientific)
[12] Khandekar D C, Lawande S V and Baghwat K V 2000 Path Integral Methods and Their Applications (Singapore: World Scientific)
[13] Huang Y, Chao Y, Yuan S and Duan J 2019 Characterization of the most probable transition paths of stochastic dynamical systems with stable Lévy noise J. Stat. Mech. 063204
[14] Zeitouni O and Dembo A 1987 A maximum a posteriori estimator for trajectories of diffusion processes Stochastics 20 221–46
[15] Zeitouni O and Dembo A 1988 An existence theorem and some properties of maximum a posteriori estimators of trajectories of diffusions Stochastics 23 197–218
[16] Chao Y and Duan J 2019 The Onsager–Machlup function as Lagrangian for the most probable path of a jump-diffusion process Nonlinearity 32 3715
[17] Field C B, Barros V, Stocker T F and Dahe Q 2012 Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change (Cambridge: Cambridge University Press)
[18] Ragone F, Wouters J and Bouchet F 2018 Computation of extreme heat waves in climate models using a large deviation algorithm Proc. Natl. Acad. Sci. USA 115 24
[19] Moon W and Wettlaufer J S 2017 A stochastic dynamical model of Arctic sea ice J. Climate 30 5119–40
[20] Ditllevsen P D 1999 Observation of α-stable noise induced millennial climate changes from an ice-core record Geophys. Res. Lett. 26 1441–4
[21] Lucarini V and Bda T 2019 Transitions across Melancholia states in a climate model: reconciling the deterministic and stochastic points of view Phys. Rev. Lett. 122 158701
[22] Sturzenegger F, Zosel F, Holmstrom E D, Buholzer K J, Makarov D E, Nettels D and Schuler B 2018 Transition path times of coupled folding and binding reveal the formation of an encounter complex Nat. Commun. 9 4708
[23] Hofler N Q, Neupane K, Pyo A G T and Woodside M T 2019 Measuring the average shape of transition paths during the folding of a single biological molecule Proc. Natl. Acad. Sci. USA 116 8125–30
[24] Turcotte M, Garcia-Ojalvo J and Suel G M 2008 A genetic timer through noise-induced stabilization of an unstable state Proc. Natl. Acad. Sci. 105 15732–7
[25] Stefan D, Pinel C, Pinhal S, Cinquemani E, Geiselmann J and de Jong H 2015 Inference of quantitative models of bacterial promoters from time-series reporter gene data PLoS Comput. Biol. 11 e1004028
[26] Ikeda N and Watanabe S 1980 Stochastic Differential Equations and Diffusion Processes (New York: Wiley)
[27] Oksendal B 2003 Stochastic Differential Equations (Berlin: Springer)
[28] Duan J 2015 An Introduction to Stochastic Dynamics (Cambridge: Cambridge University Press)
[29] Du Q, Li T, Li X and Ren W 2020 The graph limit of the minimizer of the Onsager–Machlup functional and its computation Sci. China Math. 63
[30] Giaquinta M and Hildebrandt S 2004 Calculus of Variations I (Berlin: Springer)
[31] Li Y, Duan J and Liu X 2021 Machine learning framework for computing the most probable paths of stochastic dynamical systems Phys. Rev. E 103 012124
[32] Friedlin M I and Wentzell A D 1998 Random Perturbation of Dynamical Systems 2nd edn (Berlin: Springer)
[33] Heymann M and Vanden-Eijnden E 2008 The geometric minimum action method: a least action principle on the space of curves Commun. Pure Appl. Math. 61 1052
[34] Malinin S V and Chernyak V Y 2010 Transition times in the low-noise limit of stochastic dynamics J. Chem. Phys. 132 014504
[35] Carlon E, Orland H, Sakaue T and Vanderzande C 2018 Effect of memory and active forces on transition path time distributions J. Phys. Chem. B 122 11186–94
[36] Laleman M, Carlon E and Orland H 2017 Transition path time distributions J. Chem. Phys. 147 214103
[37] Janakiraman D 2018 Transition path time distributions for Lévy flights J. Phys. A: Math. Theor. 51 285001
[38] Barret F 2015 Sharp asymptotics of metastable transition times for one dimensional SPDEs Ann. Inst. Henri Poincare Probab. Stat. 51 129C166
[39] Volpe G, Helden L, Brettschneider T, Wehr J and Bechinger C 2010 Influence of noise on force measurements Phys. Rev. Lett. 104 170602
[40] Chaichian M and Demichev A 2001 Path Integrals in Physics (Stochastic Processes and Quantum Mechanics vol 1) (Bristol: IOP Publishing)