ON THE RARE-EVENT SIMULATIONS OF DIFFUSION PROCESSES PERTAINING TO A CHAIN OF DISTRIBUTED SYSTEMS WITH SMALL RANDOM PERTURBATIONS∗

GETACHEW K. BEFEKADU†

Abstract. In this paper, we consider an importance sampling problem for a certain rare-event simulations involving the behavior of a diffusion process pertaining to a chain of distributed systems with random perturbations. We also assume that the distributed system formed by \( n \)-subsystems – in which a small random perturbation enters in the first subsystem and then subsequently transmitted to the other subsystems – satisfies an appropriate Hörmander condition. Here we provide an efficient importance sampling estimator, with an exponential variance decay rate, for the asymptotics of the probabilities of the rare events involving such a diffusion process that also ensures a minimum relative estimation error in the small noise limit. The framework for such an analysis basically relies on the connection between the probability theory of large deviations and the values functions for a family of stochastic control problems associated with the underlying distributed system, where such a connection provides a computational paradigm – based on an exponentially-tilted biasing distribution – for constructing efficient importance sampling estimators for the rare-event simulation. Moreover, as a by-product, the framework also allows us to derive a family of Hamilton-Jacobi-Bellman for which we also provide a solvability condition for the corresponding optimal control problem.

Key words. Diffusion processes, distributed systems, exit probability, HJB equations, importance sampling, large deviations, rare-event simulation, stochastic control problem

AMS subject classifications. 37H10, 35F21, 49K15, 49L20, 60F10, 65C05

1. Introduction. In this paper, we consider an importance sampling problem for a certain rare-event simulations involving the behavior of a diffusion process pertaining to a chain of distributed systems with random perturbations. We also assume that the distributed system formed by \( n \)-subsystems – in which a small random perturbation enters only in the first subsystem and then subsequently transmitted to the other subsystems – satisfies an appropriate Hörmander condition and, as a result of this, the corresponding diffusion process is assumed to have a transition probability density with strong Feller property (e.g., see \([19, 21, 2]\) and \([8]\) for related discussions from the physical point of view). Here we provide an efficient importance sampling estimator, with an exponential variance decay rate, for the asymptotics of the probabilities of the rare-event simulations involving the behavior of such a diffusion process which also ensures a minimum relative estimation error in the small noise limit case. The framework for such an analysis basically relies on the connection between the probability theory of large deviations (i.e., the asymptotics for exit probabilities from a given bounded open domain) and that of the values functions for a family of stochastic control problems associated with the underlying distributed system, where such a connection also provides a desirable computational paradigm – based on an exponentially-tilted biasing distribution – for constructing an efficient importance sampling estimator for rare-event simulations in the context of such a diffusion process. Moreover, as a by-product, the framework also allows us to derive a family of Hamilton-Jacobi-Bellman for which we also provide a solvability condition for the corresponding optimal control problem.

∗Version 1.0 – August, 2020.
†Department of Electrical & Computer Engineering, Morgan State University, 1700 E. Cold Spring Lane, Schaefer Engr. Bldg. 331, Baltimore, MD 21251, USA (getachew.befekadu@morgan.edu)

1
Here, it is worth mentioning that a number of interesting studies based on various approximations techniques from the theory of large deviations have provided a framework for constructing efficient importance sampling estimators for rare-event simulation problems involving the behavior of diffusion processes (e.g., [4], [6], [7], [18] and [24] for additional discussions). The main approach followed in these studies is to construct exponentially-tilted biasing distributions, which was originally introduced for proving Cramér’s theorem and its extension, and later on it was found to be an efficient importance sampling distribution for certain problems with various approximations involving rare-events (e.g., see [20], [1], [15] or [6] for detailed discussions). Note that the rationale behind our framework follows in some sense the settings of these papers. However, to our knowledge, the problem of rare-event simulations involving the behavior of a diffusion process pertaining to a chain of distributed systems has not been addressed in the context of stochastic control arguments along with asymptotic analysis in the small noise limit; and it is important because it provides a framework for constructing an efficient importance sampling estimator for rare-event simulations in the context of such a diffusion process.

This paper is organized as follows. In Section 2, we discuss our problem formulation, where we start by considering a diffusion process pertaining to a distributed systems with small random perturbations and we formalize our problem formulation. In this section, we also provide an asymptotic estimate on the exit probability using the Ventcel-Freidlin theorem [27] (see also [13, Chapter 4]) and the stochastic control arguments from Fleming [12] (see also [11]). In particular, as we discuss in detail in Subsection 2.2 such an asymptotic estimate relies on the interpretation of the exit probability function as a value function for a family of stochastic control problems that can be associated with the underlying distributed control system with small random perturbations. In Section 3, we provide our main result for an efficient importance sampling estimator of the rare-events involving the behavior of such a diffusion process. Finally, Section 4 provides some concluding remarks.

2. Problem setup and preliminary results. In this section, we consider the following distributed system, which is formed by a chain of $n$ subsystems, with small random perturbations, i.e.,

\[
\begin{align*}
    dx_{i}^{\varepsilon,1} &= f_1(t, x_{i}^{\varepsilon,1}, x_{i}^{\varepsilon,2}, \ldots, x_{i}^{\varepsilon,n})dt + \sqrt{\varepsilon} \sigma(t, x_{i}^{\varepsilon,1}, x_{i}^{\varepsilon,2}, \ldots, x_{i}^{\varepsilon,n})dW_t \\
    dx_{i}^{\varepsilon,2} &= f_2(t, x_{i}^{\varepsilon,1}, x_{i}^{\varepsilon,2}, \ldots, x_{i}^{\varepsilon,n})dt \\
    dx_{i}^{\varepsilon,3} &= f_3(t, x_{i}^{\varepsilon,2}, x_{i}^{\varepsilon,3}, \ldots, x_{i}^{\varepsilon,n})dt \\
    &\vdots \\
    dx_{i}^{\varepsilon,n} &= f_n(t, x_{i}^{\varepsilon,n-1}, x_{i}^{\varepsilon,n})dt, \quad \text{with} \quad (x_{s}^{\varepsilon,1}, x_{s}^{\varepsilon,2}, \ldots, x_{s}^{\varepsilon,n}), \quad 0 \leq s \leq t \leq T,
\end{align*}
\]

where

- $x^{\varepsilon,i}$ is an $\mathbb{R}^d$-valued diffusion process that corresponds to the $i$th subsystem,
- the functions $f_i$ are uniformly Lipschitz, with bounded first derivatives,
- $\sigma: [0, \infty) \times \mathbb{R}^{nd} \to \mathbb{R}^{d \times d}$ is Lipschitz with the least eigenvalue of $\sigma \sigma^T$ uniformly bounded away from zero, i.e.,

$$\sigma(t, x) \sigma^T(t, x) \geq \lambda I_{d \times d}, \quad \forall x \in \mathbb{R}^{nd}, \quad \forall t \geq 0,$$

for some $\lambda > 0$,
- $W$ (with $W_0 = 0$) is a $d$-dimensional standard Wiener process,
\(-\epsilon\) is a small positive number that represents the level of random perturbation in the distributed system, for \(i = 1, 2, \ldots, n\).

Note that, from the physical point of view (e.g., see [19], [21], [2] or [8]), the functions \(f_i\), for \(i \geq 3\), cannot be dependent on the position variables \(x_1^{\epsilon}, \ldots, x_{i-1}^{\epsilon}\) and, as a result of this, the random perturbation enters only in the first subsystem and then it subsequently passes through the second subsystem, the third subsystem, and so on to the \(n\)th subsystem. Hence, such a distributed system is described by an \(\mathbb{R}^{nd}\)-valued diffusion process and further rewrite the above SDE as follows

\[
dx_i^\epsilon = f(t, x_i^\epsilon) + \sqrt{\sigma(t, x_i^\epsilon)} \, dW_t, \tag{2.2}
\]

where \(f = (f_1, f_2, \ldots, f_n)\) is an \(\mathbb{R}^{nd}\)-valued function, \(\sigma\) is a \(nd \times d\) constant matrix that embeds \(\mathbb{R}^d\) into \(\mathbb{R}^{nd}\), i.e., \(\sigma = (I_{d \times d}, 0, \ldots, 0)^T\) and, moreover, for \(x = (x_1, x_2, \ldots, x_n)\), we set \(x_i^{\epsilon,n} = (x_i, \ldots, x_n)\). Note that the corresponding backward operator for the diffusion process \(x_i^\epsilon\), when applied to a certain function \(v^\epsilon(t, x)\), is given by

\[
\frac{\partial v^\epsilon}{\partial t} + \mathcal{L}^\epsilon v^\epsilon = \frac{\epsilon}{2} \text{tr}\left\{a(v^\epsilon x_i^\epsilon)\right\} + \sum_{i=1}^{n} \langle f_i, v^\epsilon x_i^\epsilon \rangle, \tag{2.3}
\]

where \(a(t, x) = \sigma(t, x) \sigma^T(t, x)\).

Let \(\Omega \in \mathbb{R}^{nd}\) be bounded open domains with smooth boundary (i.e., \(\partial \Omega\) is a manifold of class \(C^2\)) and let \(\Omega^T\) be an open set defined by

\[
\Omega^T = (0, T) \times \Omega.
\]

Furthermore, let us denote by \(C^\infty(\Omega^T)\) the spaces of infinitely differentiable functions on \(\Omega^T\) and by \(C_0^\infty(\Omega^T)\) the space of the functions \(\phi \in C^\infty(\Omega^T)\) with compact support in \(\Omega^T\).

A locally square integrable function \(v^\epsilon(t, x)\) on \(\Omega^T\) is said to be a distribution solution to the following equation

\[
\frac{\partial v^\epsilon}{\partial t} + \mathcal{L}^\epsilon v^\epsilon = 0, \tag{2.4}
\]

if, for any test function \(\phi \in C_0^\infty(\Omega^T)\), the following holds true

\[
\int_{\Omega^T} \left( -\frac{\partial \phi}{\partial t} + \mathcal{L}^\epsilon \phi \right) v^\epsilon \, d\Omega^T = 0, \tag{2.5}
\]

where \(d\Omega^T\) denotes the Lebesgue measure on \(\mathbb{R}^{nd+1}\) and \(\mathcal{L}^\epsilon\) is an adjoint operator corresponding to the infinitesimal generator \(\mathcal{L}^\epsilon\) of the process \(x_i^\epsilon\), i.e.,

\[
\mathcal{L}^\epsilon \phi = \frac{\epsilon}{2} \sum_{j=1}^{d} \sum_{m=1}^{d} (a_{j,m} \phi) x_j^\epsilon x_m^\epsilon - \sum_{i=1}^{n} \sum_{m=1}^{d} (f_i \phi) x_i^\epsilon x_m^\epsilon. \tag{2.6}
\]

Moreover, we also assume that the following statements hold for the distributed system in (2.1).

**Assumption 2.1.**
(a) The function $f$ is a bounded $C^\infty((0, \infty) \times \Omega)$-function, with bounded first derivatives. Moreover, $\sigma$ and $\sigma^{-1}$ are bounded $C^\infty((0, \infty) \times \mathbb{R}^{nd})$-functions, with bounded first derivatives.

(b) The backward operator in (2.3) is hypoelliptic in $C^\infty((0, \infty) \times \Omega)$ (e.g., see [10] or [23]).

(c) Let $n(x)$ be the outer normal vector to $\partial\Omega$ and, further, let $\Gamma^+$ and $\Gamma^0$ denote the sets of points $(t, x)$, with $x \in \partial\Omega$, such that

$$\langle f(t, x), n(x) \rangle$$

is positive and zero, respectively.

**Remark 2.2.** Here we remark that, in general, the hypoellipticity assumption (i.e., the backward operator in (2.3)) is related to a strong accessibility property of controllable nonlinear systems that are driven by white noise (e.g., see [22] concerning the controllability of nonlinear systems, which is closely related to [22] and [17]; see also [9, Section 3]). That is, the hypoellipticity assumption implies that the diffusion process $x^e_\tau$ has a transition probability density with a strong Feller property.

### 2.1. Importance sampling.

In this paper, we are mainly concerned with estimating the following quantity

$$E_{s, x^e_\tau}^\epsilon \left[ \exp \left( -\frac{1}{\epsilon} \Phi^\epsilon(x^e) \right) \right],$$

where $\Phi^\epsilon$ is an appropriate functional on $C([0, T] ; \mathbb{R}^{nd})$ and $x^e$ is a solution of the SDE in (2.1) and our analysis is in the situation where the level of the random perturbation is small, i.e., $\epsilon \ll 1$, and the functional $E_{s, x^e_\tau}^\epsilon \left[ \exp \left( -\frac{1}{\epsilon} \Phi^\epsilon(x^e) \right) \right]$ is rapidly varying in $x^e$. Note that the challenge presented by such an analysis of rare-event probabilities is well documented (see [3], [11] and [13] for additional discussions). In the following (and also in Subsection 2.1 and Section 3), we specifically consider the case when the functional $\Phi^\epsilon$ is bounded and nonnegative Lipschitz, with $\Phi^\epsilon = 0$, if $x^e_\tau \in \Omega^T \subset C([0, T] ; \mathbb{R}^{nd})$ and $\Phi^\epsilon = \infty$ otherwise; and we further consider analysis on the asymptotic estimates for exit probabilities from a given bounded open domain in the small noise limit case.

Consider the following simple estimator for the quantity of interest in (2.7)

$$\rho(\epsilon) = \frac{1}{N} \sum_{j=1}^{N} \exp \left( -\frac{1}{\epsilon} \Phi^\epsilon(x^e(j)) \right),$$

where $\{x^e(j)\}_{j=1}^{N}$ are $N$-copies of independent samples of $x^e$. Here we remark that such an estimator is unbiased in the sense that

$$E_{s, x^e_\tau}^\epsilon \left[ \rho(\epsilon) \right] = E_{s, x^e_\tau}^\epsilon \left[ \exp \left( -\frac{1}{\epsilon} \Phi^\epsilon(x^e) \right) \right],$$

---

Note that

$$P_{s, x^e_\tau}^\epsilon \left\{ (\tau^e, x^e_{\tau^e}) \in \Gamma^+ \cup \Gamma^0, \tau^e < \infty \right\} = 1, \quad \forall (s, x^e_\tau) \in \Omega^\infty_0,$$

where $\tau^e = \inf \{ t > s \mid x^e_t \in \partial\Omega \}$. Moreover, if

$$P_{s, x^e_\tau}^\epsilon \left\{ (t, x^e_t) \in \Gamma^0 \text{ for some } t \in [s, T] \right\} = 0, \quad \forall (s, x^e_\tau) \in \Omega^\infty_0,$$

and $\tau^e \leq T$, then we have $(\tau^e, x^e_{\tau^e}) \in \Gamma^+$, almost surely (see [24] Section 7).
Moreover, its variance is given by
\[
\text{Var} (\rho (\epsilon)) = \frac{1}{N} \left( \mathbb{E}^\epsilon_{s, x^\epsilon} \left[ \exp \left( - \frac{2}{\epsilon} \Phi^\epsilon (x^\epsilon) \right) \right] - \mathbb{E}^\epsilon_{s, x^\epsilon} \left[ \exp \left( - \frac{1}{\epsilon} \Phi^\epsilon (x^\epsilon) \right) \right]^2 \right). \tag{2.10}
\]

Then, we have the following for the relative estimation error
\[
R_{\text{err}} (\rho (\epsilon)) = \sqrt{\text{Var} (\rho (\epsilon))} \frac{\mathbb{E}^\epsilon_{s, x^\epsilon} [\rho(\epsilon)]}{\mathbb{E}^\epsilon_{s, x^\epsilon} [\rho(\epsilon)]} \tag{2.11}
\]
which can be further rewritten as follows
\[
R_{\text{err}} (\rho (\epsilon)) = \left( \frac{1}{\sqrt{N}} \right) \sqrt{\Delta (\rho(\epsilon)) - 1}, \tag{2.12}
\]
where
\[
\Delta (\rho(\epsilon)) = \frac{\mathbb{E}^\epsilon_{s, x^\epsilon} \left[ \exp \left( - \frac{2}{\epsilon} \Phi^\epsilon (x^\epsilon) \right) \right]}{\mathbb{E}^\epsilon_{s, x^\epsilon} \left[ \exp \left( - \frac{1}{\epsilon} \Phi^\epsilon (x^\epsilon) \right) \right]^2} \tag{2.13}
\]

Note that, as we might expect, the relative estimation error may decrease with increasing the number of the sample size $N$. However, from Varahian's lemma (e.g., see [25]; see also [26] and [27]), under suitable assumptions, we also have the following conditions
\[
\limsup_{\epsilon \to 0} \epsilon \log \mathbb{E}^\epsilon_{s, x^\epsilon} \left[ \exp \left( - \frac{1}{\epsilon} \Phi^\epsilon (x^\epsilon) \right) \right] = - \inf_{\varphi \in C_{sT} ([s, T], \mathbb{R}^n)} \left\{ I(\varphi) + \Phi^\epsilon(\varphi) \right\}, \tag{2.14}
\]
and
\[
\limsup_{\epsilon \to 0} \epsilon \log \mathbb{E}^\epsilon_{s, x^\epsilon} \left[ \exp \left( - \frac{2}{\epsilon} \Phi^\epsilon (x^\epsilon) \right) \right] = - \inf_{\varphi \in C_{sT} ([s, T], \mathbb{R}^n)} \left\{ I(\varphi) + 2\Phi^\epsilon(\varphi) \right\}, \tag{2.15}
\]
where $C_{sT} ([s, T], \mathbb{R}^n)$ is the set of absolutely continuous functions from $[s, T]$ into $\mathbb{R}^n$, with $0 \leq s \leq t \leq T$, and $I(\varphi)$ is the rate functional for the diffusion process $x^\epsilon_t$. From Jensen's inequality, the above equations in (2.14) and (2.15) also imply the following condition $\Delta (\rho(\epsilon)) \geq 1$.

In the following subsection, i.e., Subsection 2.2 (see also Section 3) the framework that we use to address rare-event simulation problem basically relies on the connection between the probability theory of large deviations (i.e., the asymptotics for exit probabilities) and the values functions for a family of stochastic control problems associated with the underlying distributed system. Note that such a connection provides a computational paradigm – based on an exponentially-tilted biasing distribution – for constructing an efficient importance sampling estimators for rare-event simulations that further improves the efficiency of Monte Carlo simulations.

### 2.2. Exit probabilities and stochastic optimal control.

Let $x^\epsilon_t$, for $0 \leq t \leq T$, be the diffusion process associated with (2.1) and consider the following boundary value problem
\[
\begin{align*}
\partial_s \nu^\epsilon + \mathcal{L}^\epsilon \nu^\epsilon &= 0 & \text{in } \Omega^T \\
\nu^\epsilon (s, \mathbf{x}) &= 1 & \text{on } \Gamma^+_T \\
\nu^\epsilon (s, \mathbf{x}) &= 0 & \text{on } \{T\} \times \Omega
\end{align*}
\tag{2.16}
\]
where $L^\epsilon$ is the backward operator in (2.3) and

$$\Gamma_+^\epsilon = \{(s, x) \in \Gamma^+ \mid 0 < s \leq T\}.$$ 

Further, let $\Omega_0^\epsilon$ be the set consisting of $\Omega^T \cup \{T\} \times \Omega$, together with the boundary points $(s, x) \in \Gamma^+$, with $0 < s < T$. Then, the following proposition, whose proof is given in [2], provides a solution to the exit probability $P_{s, x}^\epsilon\{\tau^\epsilon \leq T\}$ with which the diffusion process $x^\epsilon_t$ exits from the domain $\Omega$.

**Proposition 2.3.** Suppose that the statements (a)-(c) in the above assumption (i.e., Assumption 2.1) hold true. Then, the exit probability $q^\epsilon(s, x^\epsilon) = P_{s, x}^\epsilon\{\tau^\epsilon \leq T\}$ is a smooth solution to the boundary value problem in (2.16) and, moreover, it is a continuous function on $\Omega_0^\epsilon$.

Note that, from Proposition 2.3, the exit probability $q^\epsilon(s, x^\epsilon)$ is a smooth solution to the boundary value problem in (2.16). Further, if we introduce the following logarithmic transformation (e.g., see [12], [10] or [11])

$$I^\epsilon(s, x^\epsilon) = -\epsilon \log q^\epsilon(s, x^\epsilon). \quad (2.17)$$

Then, using ideas from stochastic control theory (see [12] for similar arguments), we present results useful for proving the following asymptotic property

$$I^\epsilon(s, x^\epsilon) \to I^0(s, x^\epsilon) \quad \text{as} \quad \epsilon \to 0. \quad (2.18)$$

The starting point for such an analysis is to introduce a family of related stochastic control problems whose dynamic programming equation, for $\epsilon > 0$, is given below by (2.22). Then, this also allows us to reinterpret the exit probability function as a value function for a family of stochastic control problems that are associated with the underlying distributed system with small random perturbation.

In what follows, we consider the following boundary value problem

$$\partial_s g^\epsilon + \frac{\epsilon}{2} \text{tr}\left\{a g^\epsilon_{x, x} \right\} + \sum_{j=1}^n (f_j, g^\epsilon_{x, j}) = 0 \quad \text{in} \quad \Omega^T,$$

$$g^\epsilon = \mathbb{E}_{s, x}^\epsilon \left\{\exp\left(-\frac{1}{\epsilon} \Phi^\epsilon\right) \right\} \quad \text{on} \quad \partial^* \Omega^T \quad (2.19)$$

where $\Phi^\epsilon(s, x^\epsilon)$ is a bounded, nonnegative Lipschitz function such that

$$\Phi^\epsilon(s, x^\epsilon) = 0, \quad \forall (s, x^\epsilon) \in \Gamma_+^\epsilon. \quad (2.20)$$

Observe that the function $g^\epsilon(s, x^\epsilon)$ is a smooth solution in $\Omega^T$ to the backward operator in (2.3); and it is also continuous on $\partial^* \Omega^T$. Moreover, if we introduce the following logarithm transformation (cf. Equation 2.17)

$$J^\epsilon(s, x^\epsilon) = -\epsilon \log g^\epsilon(s, x^\epsilon). \quad (2.21)$$

Then, $J^\epsilon(s, x^\epsilon)$ satisfies the following dynamic programming equation

$$\partial_s J^\epsilon + \frac{\epsilon}{2} \text{tr}\left\{a J^\epsilon_{x, x} \right\} + \sum_{j=1}^n (f_j, J^\epsilon_{x, j}) + H^\epsilon = 0 \quad \text{in} \quad \Omega^T, \quad (2.22)$$

where $H^\epsilon = H^\epsilon(s, x^\epsilon, J^\epsilon_{x, 1})$ is given by

$$H^\epsilon(s, x^\epsilon, J^\epsilon_{x, 1}) = (f_1(s, x^\epsilon), J^\epsilon_{x, 1}) - \frac{1}{2} \left(J^\epsilon_{x, 1}\right)^T a(s, x^\epsilon) J^\epsilon_{x, 1}. \quad (2.23)$$
Note that the duality relation between $H^\epsilon(s, \mathbf{x}', \cdot)$ and $L^\epsilon(s, \mathbf{x}', \cdot)$, i.e.,

$$H^\epsilon(s, \mathbf{x}', J_{x,s}^\epsilon) = \inf_{\hat{u}} \left\{ L^\epsilon(s, \mathbf{x}', \hat{u}) + \langle J_{x,s}^\epsilon, \hat{u} \rangle \right\},$$

(2.24)

with

$$L^\epsilon(s, \mathbf{x}', \hat{u}) = \frac{1}{2} \left\| f_1(s, \mathbf{x}') - \hat{u}(\cdot) \right\|_{\mathcal{L}^1}^2$$

where $\left\| \cdot \right\|_{\mathcal{L}^1}^2$ denotes the Riemannian norm of a tangent vector.

Then, it is easy to see that $J^\epsilon(s, \mathbf{x}')$ is a solution in $\Omega^T$, with $J^\epsilon = \Phi^\epsilon$ on $\partial^* \Omega^T$, to the dynamic programming in (2.22), where the latter is associated with the following stochastic control problem

$$J^\epsilon(s, \mathbf{x}') = \inf_{\hat{u} \in \hat{U}(s, \mathbf{x}')} \mathbb{E}_{s, \mathbf{x}'} \left\{ \int_s^\theta L^\epsilon(s, \mathbf{x}', \hat{u}) dt + \Phi^\epsilon(\theta, \mathbf{x}') \right\}$$

(2.25)

that corresponds to the following system of SDEs

$$dx_{x,s}^{\epsilon, 1} = \hat{u}(t) dt + \sqrt{\epsilon} \sigma(t, \mathbf{x}(t)) dW_t$$
$$dx_{x,s}^{\epsilon, 2} = f_2(t, x_{x,s}^{\epsilon, 1}) dt$$
$$dx_{x,s}^{\epsilon, 3} = f_3(t, x_{x,s}^{\epsilon, 2, n}) dt$$
$$\vdots$$
$$dx_{x,s}^{\epsilon, n} = f_n(t, x_{x,s}^{\epsilon, n-1, n}) dt$$

(2.26)

with an initial condition $\mathbf{x}_{x,s}^\epsilon$

where $x_{x,s}^{\epsilon, i, n} = (x_{x,s}^{\epsilon, i}, \ldots, x_{x,s}^{\epsilon, n})$, with $i \geq 3$, and $\hat{U}(s, \mathbf{x}')$ is a class of continuous functions for which $\theta \leq T$ and $(\theta, x_{\theta}^\epsilon) \in \Gamma^+_T$.

In what follows, we provide bounds (i.e., the asymptotic lower and upper bounds) on the exit probability $q^\epsilon(s, \mathbf{x}')$.

Define

$$I^\epsilon_{\Omega}(s, \mathbf{x}'); \partial \Omega) = - \lim_{\epsilon \to 0} \epsilon \log \mathbb{P}^\epsilon_{s, \mathbf{x}'} \left\{ \mathbf{x}_{\theta}^\epsilon \in \partial \Omega \right\},$$

(2.27)

where $\theta$ (or $\theta = \tau^\epsilon \wedge T$) is the first exit-time of $\mathbf{x}_{s}^\epsilon$ from the domain $\Omega$. Furthermore, let us introduce the following supplementary minimization problem

$$\tilde{I}^\epsilon_{\Omega}(s, \varphi, \theta) = \inf_{\varphi \in C_{s,T}([s, T], \mathbb{R}^d)} \int_s^\theta L^\epsilon(t, \varphi(t), \dot{\varphi}(t)) dt,$$

(2.28)

where the infimum is taken among all $\varphi(\cdot) \in C_{s,T}([s, T], \mathbb{R}^d)$ (i.e., from the space of $\mathbb{R}^d$-valued (locally) absolutely continuous functions, with $\int_s^T |\dot{\varphi}(t)|^2 dt < \infty$ for each $T > s$) and $\theta \geq s > 0$ such that $\varphi(s) = x_{s}^{\epsilon, 1}, (t, \varphi(t), x_{t}^{\epsilon, 2, n}) \in \Omega^T$, for all $t \in [s, \theta)$, and $(\theta, \varphi(\theta), x_{\theta}^{\epsilon, 2, n}) \in \Gamma^+_T$. Then, it is easy to see that

$$\tilde{I}^\epsilon_{\Omega}(s, \varphi, \theta) = I^\epsilon_{\Omega}(s, \mathbf{x}'); \partial \Omega).$$

(2.29)
Next, we state the following lemma that will be useful for proving Proposition 2.6 (cf. [12, Lemma 3.1]).

**Lemma 2.4.** If \( \varphi \in C^\infty_T ([s, T], \mathbb{R}^d) \), for \( s > 0 \), and \( \varphi(s) = x_s^{\epsilon, 1}, (t, \varphi(t), x_t^{\epsilon, 2,n}) \in \Omega_T \), for all \( t \in [s, T] \), then \( \lim_{T \to \infty} \int_0^T L^{\epsilon, k}(t, \varphi(t), \dot{\varphi}(t)) \, dt = +\infty \).

Consider again the stochastic control problem in (2.25) together with (2.26). Suppose that \( \Phi_M' \) (with \( \Phi_M' \geq 0 \)) is class \( C^2 \) such that \( \Phi_M' \to +\infty \) as \( M \to \infty \) uniformly on any compact subset of \( \Omega^+_T \setminus \bar{\Gamma}_T^+ \) and \( \Phi_M' \) on \( \Gamma_T^+ \). Further, if we let \( J^\epsilon = J_{\Phi_M}^\epsilon \), when \( \Phi^\epsilon = \Phi_M^\epsilon \), then we have the following lemma.

**Lemma 2.5.** Suppose that Lemma 2.4 holds, then we have

\[
\lim_{M \to \infty} J_{\Phi_M}^\epsilon ((s, x^\epsilon)) \geq I^\epsilon (s, x^\epsilon). \tag{2.30}
\]

Then, we have the following result.

**Proposition 2.6.** Suppose that Lemma 2.4 holds, then we have

\[
I^\epsilon (s, x^\epsilon) \to I^0 (s, x^\epsilon) \quad {\text{as}} \quad \epsilon \to 0, \tag{2.31}
\]

uniformly for all \( (s, x^{\epsilon, 1}(s), x^{\epsilon, 2}(s), \ldots, x^{\epsilon, k}(s)) \) in any compact subset \( \bar{\Omega}_T^\epsilon \).

**Proof.** It is sufficient to show the following conditions

\[
\limsup_{\epsilon \to 0} \epsilon \log \mathbb{P}_{s, x^\epsilon}^{\epsilon} \left\{ x^\epsilon_0 \in \partial \Omega \right\} \leq -I_{\Omega}^\epsilon ((s, x^\epsilon); \partial \Omega) \tag{2.32}
\]

and

\[
\liminf_{\epsilon \to 0} \epsilon \log \mathbb{P}_{s, x^\epsilon}^{\epsilon} \left\{ x^\epsilon_0 \in \partial \Omega \right\} \geq -I_{\Omega}^\epsilon ((s, x^\epsilon); \partial \Omega), \tag{2.33}
\]

uniformly for all \( (s, x^\epsilon_0) \) in any compact subset \( \bar{\Omega}_T^\epsilon \).

Note that \( I_{\Omega}^\epsilon ((s, x^\epsilon); \partial \Omega) = I^\epsilon (s, x^\epsilon) \) (cf. Equation (2.29)), then the upper bound in (2.32) can be verified using the Ventcel-Freidlin asymptotic estimates (see [14, pp. 332–334], [27] or [26]).

On the other hand, to prove the lower bound in (2.33), we introduce a penalty function \( \Phi_M' \) (with \( \Phi_M'(t, y) = 0 \) for \( (t, y) \in \bar{\Gamma}_T^+ \)); and write \( g' = g_M' (\equiv E_s \exp (-\frac{1}{2} \Phi_M') \)) and \( J' = J_{\Phi_M'}^\epsilon \), with \( \Phi^\epsilon = \Phi_M' \). From the boundary condition in (2.19), then, for each \( M \), we have

\[
g' (s, x^\epsilon) \leq g_M' (s, x^\epsilon). \tag{2.34}
\]

Using Lemma 2.5 and noting further the following

\[
J_{\Phi_M}^\epsilon (s, x^\epsilon) \geq I_{\Omega}^\epsilon ((s, x^\epsilon); \partial \Omega). \tag{2.35}
\]

Then, the lower bound in (2.33) holds uniformly for all \( (s, x^\epsilon_0) \) in any compact subset \( \bar{\Omega}_T^\epsilon \). This completes the proof of Proposition 2.6.
3. Main results. Let $\hat{x}^\varepsilon_t$ be the solution to the following SDE

$$d\hat{x}^\varepsilon_t = f(t, \hat{x}^\varepsilon_t) + b\sigma(t, \hat{x}^\varepsilon_t) + \sqrt{\varepsilon}b\sigma(t, \hat{x}^\varepsilon_t)\,dW_t,$$

with an initial condition $\hat{x}^\varepsilon_s = x^\varepsilon_s$, (3.1)

where $v^\varepsilon$ is an appropriate control function (which also depends on $\varepsilon$) to be chosen so as to reduce the variance of the importance sampling estimator.

Let

$$z^\varepsilon = \exp\left(-\frac{1}{\sqrt{\varepsilon}}\int_s^T \langle v^\varepsilon(t, \hat{x}^\varepsilon_t), dW_t \rangle - \frac{1}{2\varepsilon} \int_s^T |v^\varepsilon(t, \hat{x}^\varepsilon_t)|^2 dt \right).$$

(3.2)

The corresponding importance sampling estimator is then given by

$$\hat{\rho}(\varepsilon) = \frac{1}{N} \sum_{j=1}^N \exp\left(-\frac{1}{\varepsilon} \Phi^\varepsilon(\hat{x}^\varepsilon)\right)z^{\varepsilon(j)},$$

(3.3)

where $\{(\hat{x}^{\varepsilon(j)}, z^{\varepsilon(j)})\}_{j=1}^N$ are $N$-copies of independent samples of $(\hat{x}^\varepsilon, z^\varepsilon)$. Note that, for an appropriately chosen control function $v^\varepsilon$, the above importance sampling estimator in (3.3) is an unbiased estimator for (2.8), i.e.,

$$E_{s, x^\varepsilon_s} [\hat{\rho}(\varepsilon)] = E_{s, x^\varepsilon_s} [\rho(\varepsilon)].$$

(3.4)

Moreover, the relative estimation error is given by

$$R_{err}(\hat{\rho}(\varepsilon)) = \frac{\sqrt{\text{Var}(\hat{\rho}(\varepsilon))}}{E_{s, x^\varepsilon_s} [\hat{\rho}(\varepsilon)]}$$

(3.5)

which can be rewritten as follows

$$R_{err}(\hat{\rho}(\varepsilon)) = \frac{1}{\sqrt{N}} \sqrt{\Delta(\hat{\rho}(\varepsilon))} - 1,$$  

(3.6)

where

$$\Delta(\hat{\rho}(\varepsilon)) = \frac{E_{s, x^\varepsilon_s} \left[ \exp\left(-\frac{2}{\varepsilon} \Phi^\varepsilon(\hat{x}^\varepsilon)\right) \right] (z^\varepsilon)^2}{E_{s, x^\varepsilon_s} \left[ \exp\left(-\frac{1}{\varepsilon} \Phi^\varepsilon(\hat{x}^\varepsilon)\right) \right]^2}.$$  

(3.7)

Hence, in order to reduce the relative estimation error $R_{err}(\hat{\rho}(\varepsilon))$, we need to control the term $\Delta(\hat{\rho}(\varepsilon))$ in (3.6). Note that, from Jensen’s inequality, we have the following condition

$$\lim_{\varepsilon \to 0} \epsilon \log E_{s, x^\varepsilon_s} \left[ \exp\left(-\frac{2}{\varepsilon} \Phi^\varepsilon(\hat{x}^\varepsilon)\right) \right] \leq 2 \lim_{\varepsilon \to 0} \epsilon \log E_{s, x^\varepsilon_s} \left[ \exp\left(-\frac{1}{\varepsilon} \Phi^\varepsilon(\hat{x}^\varepsilon)\right) \right]$$

(3.8)

which also implies $\Delta(\hat{\rho}(\varepsilon)) \geq 1$ with $\lim_{\varepsilon \to 0} \Delta(\hat{\rho}(\varepsilon)) = 1$. Moreover, the statement in (3.5) further implies the following

$$R_{err}(\hat{\rho}(\varepsilon)) = \frac{1}{\sqrt{N}} \exp(o(1)/\varepsilon) \quad \text{as} \quad \varepsilon \to 0,$$
which is generally referred as asymptotic efficiency or optimality. In this paper, our main objective is to choose appropriately the control function $v^\epsilon$ in (3.1), so that the resulting importance sampling estimator achieves a minimum rate of error growth. For this reason, we introduce the following standard definition from simulation theory (e.g., see [3] or [4]) which is useful for interpreting our result.

**Definition 3.1.** An importance sampling estimator of the form (3.3) is log-efficient (i.e., asymptotic efficiency or optimal) if

$$\lim_{\epsilon \to 0} -\epsilon \log \Delta (\hat{\rho}(\epsilon)) = 0. \quad (3.9)$$

Then, we state our main result as follows.

**Proposition 3.2.** Suppose that the importance sampling estimator $\hat{\rho}(\epsilon)$ in (3.3), with $v^\epsilon(t, x) = -\sigma^T(t, x)J^\epsilon_x(t, x)$, is uniformly log-efficient (i.e., asymptotic efficient), where $J^\epsilon(t, x)$ satisfies the corresponding dynamic programming equation in $\Omega^T$, w.r.t. the system in (3.1), with $J^\epsilon = \Phi^\epsilon$ on $\partial^*\Omega^T$. Then, there exits a set $\mathbb{A} \subset \mathbb{R}^{nd}$ such that the Hausdorff dimension of $\mathbb{A}^c$ is zero and

$$\lim_{\epsilon \to 0} R_{err} (\hat{\rho}(\epsilon)) = 0, \quad (3.10)$$

for all $x \in \mathbb{A}$.

4. **Concluding remarks.** In this paper, we considered an importance sampling problem for a certain rare-event simulations involving the behavior of a diffusion process pertaining to a chain of distributed systems with random perturbations. In particular, we have provided an efficient importance sampling estimator, with an exponential variance decay rate, for the asymptotics of the probabilities of the rare-events involving the behavior of such a diffusion process on finite time intervals which also ensures a minimum relative estimation error in the small noise limit. The framework for such an analysis basically relies on the connection between the probability theory of large deviations and the values functions for a family of stochastic control problems associated with the underlying distributed system, where such a connection provides a computational paradigm – based on an exponentially-tilted biasing distribution – for constructing efficient importance sampling estimators for the rare-event simulation. Moreover, as a by-product, the framework also allows us to derive a family of Hamilton-Jacobi-Bellman for which we also provide a solvability condition for the corresponding optimal control problem.

**REFERENCES**

[1] S. Asmussen and P. W. Glynn, *Stochastic simulation: algorithms and analysis*, Springer, New York, 2010.
[2] G.K. Befekadu and P.J. Antsaklis, *On the asymptotic estimates for exit probabilities and minimum exit rates of diffusion processes pertaining to a chain of distributed control systems*, SIAM J. Control Optim. 53 (2015), pp. 2297–2318.
[3] J.A. Bucklew, *Introduction to rare-event simulation*, Springer Series in Statistics. Springer, New York, 2004.
[4] A. Budhiraja and P. Dupuis, *Analysis and approximation of rare events: representations and weak convergence methods*, Prob. Theory and Stoch. Modelling Series, 94, Springer, 2019.
[5] F. Delarue and S. Menozzi, *Density estimates for a random noise propagating through a chain of differential equations*, J. Funct. Anal., 259 (2010), pp. 1577–1630.
On the rare-event simulations of diffusion processes

[6] P. Dupuis and R. Ellis, *A weak convergence approach to the theory of large deviations*, Wiley, New York, 1997.

[7] P. Dupuis and H. Wang, *Importance sampling, large deviations, and differential games*, Stoch. Stoch. Rep. 76 (2004), pp. 481–508.

[8] J.P. Eckmann, C.A. Pillet and L. Rey-Bellet, *Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures*, Comm. Math. Phys. 201 (1999), pp. 657–697.

[9] D.L. Elliott, *Diffusions on manifolds arising from controllable systems*, in Geometric Methods in System Theory, D. Q. Mayne and R.W. Brockett, eds. Reidel Publ. Co., Dordrecht, Holland, 1973, pp. 285–294.

[10] L.C Evans and H. Ishii, *A PDE approach to some asymptotic problems concerning random differential equations with small noise intensities*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 2 (1985), pp. 1–20.

[11] W.H. Fleming and R.W. Rishel, *Deterministic and stochastic optimal control*, Springer-Verlag, New York, 1975.

[12] W.H. Fleming, *Exit probabilities and optimal stochastic control*, Appl. Math. Optim., 4 (1978), pp. 329–346.

[13] M.I. Freidlin and A.D. Wentzell, *Random perturbations of dynamical systems*, Springer, Berlin, 1984.

[14] A. Friedman, *Stochastic differential equations and applications*, Dover Publisher, Inc., Mineola, New York, 2006.

[15] J. Gärtner, *On large deviations from the invariant measure*, Th. Prob. Appl. 22 (1977), pp. 24-39.

[16] L. Hörmander, *Hypoelliptic second order differential operators*, Acta Math. 119 (1967), pp. 147–171.

[17] K. Ichihara and H. Kunita, *A classification of the second order degenerate elliptic operators and its probabilistic characterization*, Z. Wahrscheinlichkeitstheor. Verw. Geb., 30 (1974), pp. 253–254.

[18] E.W. Ellem and E. Vanden-Eijnden, *Minimum action method for the study of rare events*, Comm. Pure Appl. Math. 57 (2004), pp. 637–656.

[19] L. Rey-Bellet and L.E. Lawrence, *Asymptotic behavior of thermal nonequilibrium steady states for a driven chain of anharmonic oscillators*, Comm. Math. Phys. 215 (2000), pp. 1–24.

[20] D. Siegmund, *Importance sampling in the monte carlo study of sequential tests*, The Annals of Statistics, (1976), pp. 673–684.

[21] C. Soize, *The Fokker-Planck equation for stochastic dynamical systems and its explicit steady state solutions*, Ser. Adv. Math. Appl. Sci., vol. 17, World Scientific Publishing Co., Inc., River Edge, NJ, 1994.

[22] D. Stroock and S.R.S. Varadhan, *On degenerate elliptic-parabolic operators of second order and their associated diffusions*, Comm. Pure Appl. Math., 25 (1972), pp. 651–713.

[23] H.J. Sussmann and V. Jurjevic, *Controllability of nonlinear systems*, J. Diff. Equ. 12 (1972), pp. 95–116.

[24] E. Vanden-Eijnden and J. Weare, *Rare event simulation of small noise diffusions*, Comm. Pure Appl. Math. 65 (2012), pp. 1770–1803.

[25] S.R.S. Varadhan, *Large deviations and applications*, CBMS-NSF Regional Conference Series in Applied Mathematics, 46. SIAM, Philadelphia, 1985.

[26] A.D. Ventcel, *Limit theorems on large deviations for stochastic processes*, Theo. Prob. Appl., 18 (1973), pp. 817–821.

[27] A.D. Ventcel and M.I. Freidlin, *On small random perturbations of dynamical systems*, Russian Math. Surv. 25 (1970), pp. 1–55.