Chapter

Rare Event Simulation in a Dynamical Model Describing the Spread of Traffic Congestions in Urban Network Systems

Getachew K. Befekadu

Abstract

In this chapter, we present a mathematical framework that provides a new insight for understanding the spread of traffic congestions in an urban network system. In particular, we consider a dynamical model, based on the well-known susceptible-infected-recovered (SIR) model from mathematical epidemiology, with small random perturbations, that describes the process of traffic congestion propagation and dissipation in an urban network system. Here, we provide the asymptotic probability estimate based on the Freidlin-Wentzell theory of large deviations for certain rare events that are difficult to observe in the simulation of urban traffic network dynamics. Moreover, the framework provides a computational algorithm for constructing efficient importance sampling estimators for rare event simulations of certain events associated with the spread of traffic congestions in the dynamics of the traffic network.

Keywords: diffusion processes, exit probability, HJB equations, importance sampling, large deviations, rare-event simulation, SIR model, traffic network dynamics

1. Introduction

In recent years, there have been a number of interesting studies related to modeling the spread of traffic congestion propagation and traffic dissipation in urban network systems (e.g., see [1–5] in the context of macroscopic traffic model involving traffic flux and traffic density; see [6, 7] in the context of percolation theory; see [8] for results based on machine-learning methods; and see [9, 10] for studies based on queuing theory). In this paper, without attempting to give a literature review, we consider a dynamical model, based on the well-known susceptible-infected-recovered (SIR) model from mathematical epidemiology, with small random perturbation, that describes the spread of traffic congestion propagation and dissipation in an urban network system, i.e.,

$$\begin{aligned} dc^{\varepsilon}(t) &= (-\mu + \beta k(1 - r^{\varepsilon}(t) - c^{\varepsilon}(t)))c^{\varepsilon}(t)dt \\ &\quad + \sqrt{\varepsilon}\sqrt{(\mu + \beta k(1 + r^{\varepsilon}(t) + c^{\varepsilon}(t)))c^{\varepsilon}(t)}\,dW_1(t) \end{aligned} \tag{1}$$

$$dr^{\varepsilon}(t) = \mu r^{\varepsilon}(t) + \sqrt{\varepsilon}\sqrt{\mu r^{\varepsilon}(t)}\,dW_2(t) \tag{2}$$

$$df^{\varepsilon}(t) = (-\beta k(1 - r^{\varepsilon}(t) - c^{\varepsilon}(t)))c^{\varepsilon}(t)dt \\ + \sqrt{\varepsilon}\sqrt{(\beta k(1 + r^{\varepsilon}(t) + c^{\varepsilon}(t)))c^{\varepsilon}(t)}dW_3(t) \tag{3}$$

where

- $c^{\varepsilon}(t)$ represents the fraction of congested links in the network
- $r^{\varepsilon}(t)$ represents the fraction of recovered links in the network
- $f^{\varepsilon}(t)$ represents the fraction of free flow links in the network
- the parameters β and μ represent respectively the propagation and recovery rates considering that a certain fraction of congested links will eventually recover as the demand for travel diminishes
- the quantity $k\beta/\mu$ represents the average number of newly congested links that, in a fully freely flowing traffic network, each already congested link can potentially create,
- $W_1(t)$, $W_2(t)$ and $W_3(t)$ are three independent standard (one-dimensional) Wiener processes, and
- ε is a small positive number that represents the level of random perturbation in the network.

Notice that Eq. (1) describes the rate at which the fraction of congested links, i.e., $c^{\varepsilon}(t)$, changes over time given the propagation rate β and recovery rate μ considering that a fraction of congested links will eventually recover as the demand for the travel volume diminishes. Moreover, Eq. (2) describes the rate at which congested links normally recover given the recovery rate μ. Finally, Eq. (3) represents how the fraction of free flow links $f^{\varepsilon}(t)$ in the network changes over time given $c^{\varepsilon}(t)$ and $r^{\varepsilon}(t)$. Note that, for a normalized SIR based traffic network dynamic model, the following mathematical condition $c^{\varepsilon}(t) + r^{\varepsilon}(t) + f^{\varepsilon}(t) = 1$ holds true for all $t > 0$, where $f^{\varepsilon}(t)$ represents links that have remained in a free flow state starting from $t = 0$ (e.g., see Saberi et al. [11] for detailed discussions related to deterministic models).

In this chapter, we provide the asymptotic probability estimate based on the Freidlin-Wentzell theory of large deviations for certain rare events that are difficult to observe in the simulation of urban traffic network dynamics. The framework considered in this study basically relies on the connection between the probability theory of large deviations and that of the values functions for a family of stochastic control problems, where such a connection also provides a desirable computational algorithm for constructing an efficient importance sampling estimator for rare event simulations of certain events associated with the spread of traffic congestions in the dynamics of the traffic network. 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., [12–16] for additional discussions). The 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 [17–19] or [13] for detailed discussions). 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 spread of traffic congestions in an urban network system has not been addressed in the context of large deviations and stochastic control arguments in the small noise limit; and it is important because it provides a new insight for understanding the spread of traffic congestions in an urban network system.

This chapter is organized as follows. In Section 2, we provide an asymptotic estimate on the exit probability using the Freidlin-Wentzell theory of large deviations [20] (see also [21], Chapter 4) and the stochastic control arguments from Fleming [22] (see also [23]), where 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 SIR based traffic network dynamic model with small random perturbations. In Section 3, we discuss importance sampling and the necessary background upon which our main results rely. In Section 4, we provide our main results for an efficient importance sampling estimator for rare event simulations of certain events associated with the spread of traffic congestions in the dynamics of the traffic network. Finally, Section 5 provides some concluding remarks.

2. The Freidlin-Wentzell theory

In this section, we briefly review the classical Freidlin-Wentzell theory of large deviations for the stochastic differential equations (SDEs) with small noise terms. In what follows, let us denote the solution of the SDEs in Eqs. (1)–(3) by a bold face letter $\left(\mathbf{x}_t^\varepsilon\right)_{t\ge 0} = \left(x_t^{\varepsilon,1}, x_t^{\varepsilon,2}, x_t^{\varepsilon,3}\right)_{t\ge 0} \triangleq \left(c^\varepsilon(t), r^\varepsilon(t), f^\varepsilon(t)\right)_{t\ge 0}$ as an $\mathbb{R}^3$-valued diffusion process and rewrite the above equations as follows

$$d\mathbf{x}_t^\varepsilon = \mathbf{f}\left(\mathbf{x}_t^\varepsilon\right)dt + \sqrt{\varepsilon}\,\sigma\left(\mathbf{x}_t^\varepsilon\right)dW_t, \tag{4}$$

where $\mathbf{f}\left(\mathbf{x}_t^\varepsilon\right) = \left[f_1\left(\mathbf{x}_t^\varepsilon\right), f_2\left(\mathbf{x}_t^\varepsilon\right), f_3\left(\mathbf{x}_t^\varepsilon\right)\right]^T$ with

$$\begin{aligned} f_1\left(x_t^{\varepsilon,1}, x_t^{\varepsilon,2}, x_t^{\varepsilon,3}\right) &= \left(-\mu + \beta k\left(1 - x_t^{\varepsilon,2} - x_t^{\varepsilon,1}\right)\right)x_t^{\varepsilon,1} \\ f_2\left(x_t^{\varepsilon,1}, x_t^{\varepsilon,2}, x_t^{\varepsilon,3}\right) &= \mu x_t^{\varepsilon,2} \\ f_3\left(x_t^{\varepsilon,1}, x_t^{\varepsilon,2}, x_t^{\varepsilon,3}\right) &= \left(-\beta k\left(1 - x_t^{\varepsilon,2} - x_t^{\varepsilon,1}\right)\right)x_t^{\varepsilon,1} \end{aligned} \tag{5}$$

and $\sigma\left(\mathbf{x}_t^\varepsilon\right) = \left[\sigma_1\left(\mathbf{x}_t^\varepsilon\right), \sigma_2\left(\mathbf{x}_t^\varepsilon\right), \sigma_3\left(\mathbf{x}_t^\varepsilon\right)\right]^T$ with

$$\begin{aligned} \sigma_1\left(x_t^{\varepsilon,1}, x_t^{\varepsilon,2}, x_t^{\varepsilon,3}\right) &= \sqrt{\left(\mu + \beta k\left(1 + x_t^{\varepsilon,2} + x_t^{\varepsilon,1}\right)\right)x_t^{\varepsilon,1}} \\ \sigma_2\left(x_t^{\varepsilon,1}, x_t^{\varepsilon,2}, x_t^{\varepsilon,3}\right) &= \sqrt{\mu x_t^{\varepsilon,2}} \\ \sigma_3\left(x_t^{\varepsilon,1}, x_t^{\varepsilon,2}, x_t^{\varepsilon,3}\right) &= \sqrt{\left(\beta k\left(1 + x_t^{\varepsilon,2} + x_t^{\varepsilon,1}\right)\right)x_t^{\varepsilon,1}}. \end{aligned} \tag{6}$$

Moreover, W_t is a standard three-dimensional Wiener process. Note that the corresponding backward operator for the diffusion process $\mathbf{x}_t^\varepsilon$, when applied to a certain function $v^\varepsilon(t, \mathbf{x})$, is given by

$$\partial_t v^\varepsilon + \mathcal{L}^\varepsilon v \triangleq \frac{\partial v^\varepsilon(t,\mathbf{x})}{\partial t} + \frac{\varepsilon}{2}\sum_{i,j=1}^{3} a_{i,j}(\mathbf{x})\frac{\partial^2 v^\varepsilon(t,\mathbf{x})}{\partial x^i \partial x^j} + \mathbf{f}(\mathbf{x})\cdot\nabla_{\mathbf{x}} v^\varepsilon(t,\mathbf{x}), \tag{7}$$

where $a(\mathbf{x}) = \sigma(\mathbf{x})\sigma^T(\mathbf{x})$.

Let $\Omega \in \mathbb{R}^3$ be bounded open domains with smooth boundary (i.e., $\partial\Omega$ is a manifold of class C^2) and let Ω^T be an open set defined by

$$\Omega^T = (0,T)\times\Omega. \tag{8}$$

Furthermore, let us denote by $C^\infty(\Omega^T)$ the spaces of infinitely differentiable functions on Ω^T and by $C_0^\infty(\Omega^T)$ the space of the functions $\phi \in C^\infty(\Omega^T)$ with compact support in Ω^T. A locally square integrable function $v^\varepsilon(t,\mathbf{x})$ on Ω^T is said to be a distribution solution to the following equation

$$\partial_t v^\varepsilon + \mathcal{L}^\varepsilon v^\varepsilon = 0, \tag{9}$$

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

$$\int_{\Omega^T}(-\partial_t\phi + \mathcal{L}^{\varepsilon *}\phi)v^\varepsilon d\Omega^T = 0, \tag{10}$$

where $d\Omega^T$ denotes the Lebesgue measure on $\mathbb{R}^3\times\mathbb{R}_+$ and $\mathcal{L}^{\varepsilon *}$ is an adjoint operator corresponding to the infinitesimal generator $\mathcal{L}^\varepsilon$ of the process $\mathbf{x}_t^\varepsilon$.

Moreover, we also assume that the following statements hold for the SDE in (4).

Assumption 1

a. The function $\mathbf{f}$ is a bounded $C^\infty((0,\infty)\times\Omega)$-function, with bounded first derivatives. Moreover, σ and σ^{-1} are bounded $C^\infty\left((0,\infty)\times\mathbb{R}^3\right)$-functions, with bounded first derivatives.

b. Let $n(\mathbf{x})$ be the outer normal vector to $\partial\Omega$ and, further, let Γ^+ and Γ^0 denote the sets of points $(t,\mathbf{x})$, with $\mathbf{x}\in\partial\Omega$, such that

$$\langle \mathbf{f}(t,\mathbf{x}), n(\mathbf{x})\rangle \tag{11}$$

is positive and zero, respectively.

Remark 1 *Note that*

$$\mathbb{P}^\varepsilon_{s,\mathbf{x}_s^\varepsilon}\left\{\left(\tau^\varepsilon, \mathbf{x}^\varepsilon_{\tau^\varepsilon}\right)\in\Gamma^+\bigcup\Gamma^0,\ \tau^\varepsilon<\infty\right\} = 1, \quad \forall\left(s,\mathbf{x}_s^\varepsilon\right)\in\Omega_0^\infty. \tag{12}$$

where $\tau^\varepsilon = \inf\left\{t>s \,|\, \mathbf{x}_t^\varepsilon\in\partial\Omega\right\}$. Moreover, if

$$\mathbb{P}^\varepsilon_{s,\mathbf{x}_s^\varepsilon}\left\{\left(t,\mathbf{x}_t^\varepsilon\right)\in\Gamma^0 \text{ for some } t\in[s,T]\right\} = 0, \quad \forall\left(s,\mathbf{x}_s^\varepsilon\right)\in\Omega_0^\infty, \tag{13}$$

and $\tau^\varepsilon\le T$, then we have $\left(\tau^\varepsilon,\mathbf{x}^\varepsilon_{\tau^\varepsilon}\right)\in\Gamma^+$, almost surely (see [24], Section 7).

In what follows, let $\mathbf{x}_t^\varepsilon$, for $0\le t\le T$, be the diffusion process associated with (4) (or Eqs. (1)–(3)) and consider the following boundary value problem

$$\left.\begin{array}{c}\partial_s v^\varepsilon + \mathcal{L}^\varepsilon v^\varepsilon = 0 \quad \text{in} \quad \Omega^T \\ v^\varepsilon(s,\mathbf{x}) = 1 \quad \text{on} \quad \Gamma_T^+ \\ v^\varepsilon(s,\mathbf{x}) = 0 \quad \text{on} \quad \{T\}\times\Omega\end{array}\right\} \tag{14}$$

where $\mathcal{L}^\varepsilon$ is the backward operator in (7) and

$$\Gamma_T^+ = \{(s, \mathbf{x}) \in \Gamma^+ \,|\, 0 < s \leq T\}. \tag{15}$$

Further, let Ω^{0T} be the set consisting of $\Omega^T \cup \{T\} \times \Omega$, together with the boundary points $(s, \mathbf{x}) \in \Gamma^+$, with $0 < s < T$. Then, the following proposition, whose proof is given in [25], provides a solution to the exit probability $\mathbb{P}^\varepsilon_{s,\mathbf{x}^\varepsilon_s}\{\tau^\varepsilon \leq T\}$ with which the diffusion process $\mathbf{x}^\varepsilon_t$ exits from the domain Ω.

Proposition 1 Suppose that the statements in Assumption 1 hold true. Then, the exit probability $q^\varepsilon(s, \mathbf{x}^\varepsilon) = \mathbb{P}^\varepsilon_{s,\mathbf{x}^\varepsilon_s}\{\tau^\varepsilon \leq T\}$ is a smooth solution to the boundary value problem in (14) and, moreover, it is a continuous function on Ω^{0T}.

Note that, from Proposition 1, the exit probability $q^\varepsilon(s, \mathbf{x}^\varepsilon)$ is a smooth solution to the boundary value problem in (14). Further, if we introduce the following logarithmic transformation (e.g., see [22, 26] or [23])

$$I^\varepsilon(s, \mathbf{x}^\varepsilon) = -\varepsilon \log q^\varepsilon(s, \mathbf{x}^\varepsilon). \tag{16}$$

Then, using ideas from stochastic control theory (see [22] for similar arguments), we present results useful for proving the following asymptotic property

$$I^\varepsilon(s, \mathbf{x}^\varepsilon) \to I^0(s, \mathbf{x}^\varepsilon) \quad \text{as} \quad \varepsilon \to 0. \tag{17}$$

The starting point for such an analysis is to introduce a family of related stochastic control problems whose dynamic programming equation, for $\varepsilon > 0$, is given below by (21). Then, this also allows us to reinterpret the exit probability function as a value function for a family of stochastic control problems associated with the underlying urban traffic network dynamics with small random perturbation. Moreover, as discussed later in Section 5, 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.

Then, we consider the following boundary value problem

$$\left.\begin{array}{r} \partial_s g^\varepsilon + \dfrac{\varepsilon}{2}\mathcal{L}^\varepsilon = 0 \quad \text{in} \quad \Omega^T \\ g^\varepsilon = \mathbb{E}^\varepsilon_{s,\mathbf{x}}\left\{ \exp\left(-\dfrac{1}{\varepsilon}\Phi^\varepsilon\right)\right\} \quad \text{on} \quad \partial^*\Omega^T \end{array}\right\} \tag{18}$$

where $\Phi^\varepsilon(s, \mathbf{x}^\varepsilon)$ is a bounded, nonnegative Lipschitz function such that

$$\Phi^\varepsilon(s, \mathbf{x}^\varepsilon) = 0, \quad \forall (s, \mathbf{x}^\varepsilon) \in \Gamma_T^+. \tag{19}$$

Observe that the function $g^\varepsilon(s, \mathbf{x}^\varepsilon)$ is a smooth solution in Ω^T to the backward operator in (9); and it is also continuous on $\partial^*\Omega^T$. Moreover, if we introduce the following logarithm transformation

$$J^\varepsilon(s, \mathbf{x}^\varepsilon) = -\varepsilon \log g^\varepsilon(s, \mathbf{x}^\varepsilon). \tag{20}$$

Then, $J^\varepsilon(s, \mathbf{x}^\varepsilon)$ satisfies the following dynamic programming equation (i.e., the Hamilton-Jacobi-Bellman equation)

$$\partial_s J^\varepsilon + \frac{\varepsilon}{2}\sum_{i,j=1}^{3} a_{i,j}\,\frac{\partial^2 J^\varepsilon}{\partial x^i \partial x^j} + H^\varepsilon = 0, \quad \text{in} \quad \Omega^T, \tag{21}$$

where $H^\varepsilon = H^\varepsilon(s, \mathbf{x}^\varepsilon, \nabla_\mathbf{x} J^\varepsilon)$ is given by

$$H^\varepsilon(s, \mathbf{x}^\varepsilon, \nabla_\mathbf{x} J^\varepsilon) = \mathbf{f}(\mathbf{x}^\varepsilon) \cdot \nabla_\mathbf{x} J^\varepsilon(s, \mathbf{x}^\varepsilon) - \frac{1}{2}(\nabla_\mathbf{x} J^\varepsilon(s, \mathbf{x}^\varepsilon))^T a(\mathbf{x}^\varepsilon) \nabla_\mathbf{x} J^\varepsilon(s, \mathbf{x}^\varepsilon). \tag{22}$$

Note that the duality relation between $H^\varepsilon(s, \mathbf{x}^\varepsilon, \cdot)$ and $L^\varepsilon(s, \mathbf{x}^\varepsilon, \cdot)$, i.e.,

$$H^\varepsilon(s, \mathbf{x}^\varepsilon, \nabla_\mathbf{x} J^\varepsilon) = \inf_{\hat{u}} \{L^\varepsilon(s, \mathbf{x}^\varepsilon, \hat{u}) + \ \nabla_\mathbf{x} J^\varepsilon \cdot \hat{u}\rangle\}, \tag{23}$$

with

$$L^\varepsilon(s, \mathbf{x}^\varepsilon, \hat{u}) = \frac{1}{2}\|\mathbf{f}(\mathbf{x}^\varepsilon) - \hat{u}\|^2_{[a(\mathbf{x}^\varepsilon)]^{-1}}, \tag{24}$$

where $\|\cdot\|^2_{[a(\mathbf{x}^\varepsilon)]^{-1}}$ denotes the Riemannian norm of a tangent vector.

Then, it is easy to see that $J^\varepsilon(s, \mathbf{x}^\varepsilon)$ is a solution in Ω^T, with $J^\varepsilon = \Phi^\varepsilon$ on $\partial^*\Omega^T$, to the dynamic programming in (21), where the latter is associated with the following stochastic control problem

$$J^\varepsilon(s, \mathbf{x}^\varepsilon) = \inf_{\hat{u} \in \hat{U}(s, \mathbf{x}_s^\varepsilon)} \mathbb{E}_{s, \mathbf{x}_s^\varepsilon}\left\{\int_s^\theta L^\varepsilon(s, \mathbf{x}^\varepsilon, \hat{u})dt + \Phi^\varepsilon(\theta, \mathbf{x}^\varepsilon)\right\} \tag{25}$$

that corresponds to the following system of SDEs

$$d\mathbf{x}_t^\varepsilon = \hat{u}(t)dt + \sqrt{\varepsilon}\sigma(\mathbf{x}_t^\varepsilon)dW_t, \tag{26}$$

with an initial condition $\mathbf{x}_s^\varepsilon = \mathbf{x}^\varepsilon$ and $\hat{U}(s, \mathbf{x}^\varepsilon)$ is a class of continuous functions for which $\theta \le T$ and $(\theta, x_\theta^\varepsilon) \in \Gamma_T^+$.

Next, we provide bounds, i.e., the asymptotic lower and upper bounds, on the exit probability $q^\varepsilon(s, \mathbf{x}^\varepsilon)$.

Define

$$\begin{aligned} I_\Omega^\varepsilon((s, \mathbf{x}^\varepsilon); \partial\Omega) &= -\lim_{\varepsilon \to 0} \varepsilon \log \mathbb{P}_{s, \mathbf{x}_s^\varepsilon}^\varepsilon\{\mathbf{x}_\theta^\varepsilon \in \partial\Omega\}, \\ &\triangleq -\lim_{\varepsilon \to 0} \varepsilon \log q^\varepsilon(s, \mathbf{x}^\varepsilon), \end{aligned} \tag{27}$$

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

$$\tilde{I}_\Omega^\varepsilon(s, \varphi, \theta) = \inf_{\varphi \in C_{sT}([s, T], \mathbb{R}^3), \theta \ge s} \int_s^\theta L^\varepsilon(t, \varphi(t), \dot{\varphi}(t))dt, \tag{28}$$

where the infimum is taken among all $\varphi(\cdot) \in C_{sT}([s, T], \mathbb{R}^3)$ (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 $\theta \ge s > 0$ such that $\varphi(s) \in \Omega^T$, for all $t \in [s, \theta)$, and $(\theta, \varphi(\theta)) \in \Gamma_T^+$. Then, it is easy to see that

$$\tilde{I}_\Omega^\varepsilon(s, \varphi, \theta) = I_\Omega^\varepsilon((s, \mathbf{x}^\varepsilon); \partial\Omega). \tag{29}$$

Next, we state the following lemma that will be useful for proving Proposition 2 (cf. [22], Lemma 3.1).

Lemma 1 If $\varphi \in C_{sT}([s,T], \mathbb{R}^3)$, for $s > 0$, and $\varphi(s) = \mathbf{x}_s^\varepsilon$, $(t, \varphi(t)) \in \Omega^T$, for all $t \in [s, T)$, then $\lim_{T\to\infty} \int_s^T L^\varepsilon(t, \varphi(t), \dot{\varphi}(t))dt = +\infty$.

Consider again the stochastic control problem in (25) together with (26). Suppose that Φ_M^ε (with $\Phi_M^\varepsilon \geq 0$) is class C^2 such that $\Phi_M^\varepsilon \to +\infty$ as $M \to \infty$ uniformly on any compact subset of $\Omega^T \backslash \overline{\Gamma}_T^+$ and Φ_M^ε on Γ_T^+. Further, if we let $J^\varepsilon = J_{\Phi_M}^\varepsilon$, when $\Phi^\varepsilon = \Phi_M^\varepsilon$, then we have the following lemma.

Lemma 2 Suppose that Lemma 1 holds, then we have

$$\liminf_{\substack{M\to\infty \\ (t, \mathbf{x}_t^\varepsilon)\to(s, \mathbf{x}_s^\varepsilon)}} \quad J_{\Phi_M}^\varepsilon((s, \mathbf{x}^\varepsilon)) \geq I^\varepsilon(s, \mathbf{x}^\varepsilon). \tag{30}$$

Then, we have the following result.

Proposition 2 [25, Proposition 2.8] Suppose that Lemma 1 holds, then we have

$$I^\varepsilon(s, \mathbf{x}^\varepsilon) \to I^0(s, \mathbf{x}^\varepsilon) \quad \text{as} \quad \varepsilon \to 0, \tag{31}$$

uniformly for all $(s, \mathbf{x}_s^\varepsilon)$ in any compact subset $\overline{\Omega}^T$.

Proof: It is suffices to show the following conditions

$$\limsup_{\varepsilon\to 0} \varepsilon \, \log \mathbb{P}_{s,\mathbf{x}_s^\varepsilon}^\varepsilon \{\mathbf{x}_\theta^\varepsilon \in \partial\Omega\} \leq -I_\Omega^\varepsilon((s, \mathbf{x}^\varepsilon); \partial\Omega) \tag{32}$$

and

$$\liminf_{\varepsilon\to 0} \varepsilon \, \log \mathbb{P}_{s,\mathbf{x}_s^\varepsilon}^\varepsilon \{\mathbf{x}_\theta^\varepsilon \in \partial\Omega\} \geq -I_\Omega^\varepsilon((s, \mathbf{x}^\varepsilon); \partial\Omega), \tag{33}$$

uniformly for all $(s, \mathbf{x}_s^\varepsilon)$ in any compact subset $\overline{\Omega}^T$. Note that $I_\Omega^\varepsilon((s, \mathbf{x}^\varepsilon); \partial\Omega) = I^\varepsilon(s, \mathbf{x}^\varepsilon)$ (cf. Eq. (29)), then the upper bound in (32) can be verified using the Freidlin-Wentzell asymptotic estimates (e.g., see [27], pp. 332–334, [20] or [28]).

On the other hand, to prove the lower bound in (33), we introduce a penalty function Φ_M^ε (with $\Phi_M^\varepsilon(t, \mathbf{y}) = 0$ for $(t, \mathbf{y}) \in \Gamma_T^+$); and write $g^\varepsilon = g_M^\varepsilon$ $(\equiv \mathbb{E}_{s,\mathbf{x}_s^\varepsilon}^\varepsilon\{\exp(-\frac{1}{\varepsilon}\Phi_M^\varepsilon)\})$ and $J^\varepsilon = J_{\Phi_M}^\varepsilon$, with $\Phi^\varepsilon = \Phi_M^\varepsilon$. From the boundary condition in (18), then, for each M, we have

$$g^\varepsilon(s, \mathbf{x}^\varepsilon) \leq g_M^\varepsilon(s, \mathbf{x}^\varepsilon). \tag{34}$$

Using Lemma 2 and noting further the following

$$J_{\Phi_M}^\varepsilon(s, \mathbf{x}^\varepsilon) \geq I_\Omega^\varepsilon((s, \mathbf{x}^\varepsilon); \partial\Omega). \tag{35}$$

Then, the lower bound in (33) holds uniformly for all $(s, \mathbf{x}_s^\varepsilon)$ in any compact subset $\overline{\Omega}^T$. This completes the proof of Proposition 2. □

3. Importance sampling

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

$$\mathbb{E}_{s,\mathbf{x}_s^\varepsilon}^\varepsilon\left[\exp\left(-\frac{1}{\varepsilon}\Phi^\varepsilon(\mathbf{x}^\varepsilon)\right)\right], \tag{36}$$

where Φ^ε is an appropriate functional on $C([0,T];\mathbb{R}^3)$ and $\mathbf{x}^\varepsilon$ is a solution of the SDE in (4) and our analysis is in the situation where the level of the random perturbation is small, i.e., $\varepsilon \ll 1$, and the functional $\mathbb{E}^\varepsilon_{s,\mathbf{x}^\varepsilon_s}\left[\exp\left(-\frac{1}{\varepsilon}\Phi^\varepsilon(\mathbf{x}^\varepsilon)\right)\right]$ is rapidly varying in $\mathbf{x}^\varepsilon$. Note that the challenge presented by such an analysis of rare event probabilities is well documented (see [12, 18, 29] for additional discussions). In the following (and see also Section 4), we specifically consider the case when the functional Φ^ε is bounded and nonnegative Lipschitz, with $\Phi^\varepsilon = 0$, if $\mathbf{x}^\varepsilon_t \in \Omega^T \subset C([0,T]:\mathbb{R}^3)$ and $\Phi^\varepsilon = \infty$ 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 (36)

$$\rho(\varepsilon) = \frac{1}{N}\sum_{j=1}^{N} \exp\left(-\frac{1}{\varepsilon}\Phi^\varepsilon\left(\mathbf{x}^{\varepsilon(j)}\right)\right), \tag{37}$$

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

$$\mathbb{E}^\varepsilon_{s,\mathbf{x}^\varepsilon_s}[\rho(\varepsilon)] = \mathbb{E}^\varepsilon_{s,\mathbf{x}^\varepsilon_s}\left[\exp\left(-\frac{1}{\varepsilon}\Phi^\varepsilon(\mathbf{x}^\varepsilon)\right)\right], \tag{38}$$

Moreover, its variance is given by

$$\mathrm{Var}(\rho(\varepsilon)) = \frac{1}{N}\left(\mathbb{E}^\varepsilon_{s,\mathbf{x}^\varepsilon_s}\left[\exp\left(-\frac{2}{\varepsilon}\Phi^\varepsilon(\mathbf{x}^\varepsilon)\right)\right] - \mathbb{E}^\varepsilon_{s,\mathbf{x}^\varepsilon_s}\left[\exp\left(-\frac{1}{\varepsilon}\Phi^\varepsilon(\mathbf{x}^\varepsilon)\right)\right]^2\right). \tag{39}$$

Then, we have the following for the relative estimation error

$$R_{\mathrm{err}}(\rho(\varepsilon)) = \frac{\sqrt{\mathrm{Var}(\rho(\varepsilon))}}{\mathbb{E}^\varepsilon_{s,\mathbf{x}^\varepsilon_s}[\rho(\varepsilon)]} \tag{40}$$

which can be further rewritten as follows

$$R_{\mathrm{err}}(\rho(\varepsilon)) = \left(1/\sqrt{N}\right)\sqrt{\Delta(\rho(\varepsilon)) - 1}, \tag{41}$$

where

$$\Delta(\rho(\varepsilon)) = \frac{\mathbb{E}^\varepsilon_{s,\mathbf{x}^\varepsilon_s}\left[\exp\left(-\frac{2}{\varepsilon}\Phi^\varepsilon(\mathbf{x}^\varepsilon)\right)\right]}{\mathbb{E}^\varepsilon_{s,\mathbf{x}^\varepsilon_s}\left[\exp\left(-\frac{1}{\varepsilon}\Phi^\varepsilon(\mathbf{x}^\varepsilon)\right)\right]^2}. \tag{42}$$

Note that, as we might expect, the relative estimation error may decrease with increasing the number of the sample size N. However, from Varahhan's lemma (e.g., see [30]; see also [20, 28]), under suitable assumptions, we also have the following conditions

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{E}^\varepsilon_{s,\mathbf{x}^\varepsilon_s}\left[\exp\left(-\frac{1}{\varepsilon}\Phi^\varepsilon(\mathbf{x}^\varepsilon)\right)\right] = -\inf_{\substack{\varphi \in C_{sT}\left([s,T],\mathbb{R}^{nd}\right)\\ \varphi(s)=\mathbf{x}_s}} \{I(\varphi) + \Phi^\varepsilon(\varphi)\} \tag{43}$$

and

$$\limsup_{\varepsilon \to 0} \varepsilon \log \mathbb{E}^{\varepsilon}_{s,\mathbf{x}^{\varepsilon}_s}\left[\exp\left(-\frac{2}{\varepsilon}\Phi^{\varepsilon}(\mathbf{x}^{\varepsilon})\right)\right] = -\inf_{\substack{\varphi \in C_{sT}\left([s,T],\mathbb{R}^{nd}\right)\\ \varphi(s)=\mathbf{x}_s}} \{I(\varphi) + 2\Phi^{\varepsilon}(\varphi)\} \tag{44}$$

where $C_{sT}\left([s,T],\mathbb{R}^3\right)$ is the set of absolutely continuous functions from $[s,T]$ into $\mathbb{R}^3$, with $0 \le s \le t \le T$, and $I(\varphi)$ is the rate functional for the diffusion process $\mathbf{x}^{\varepsilon}_t$. From Jensen's inequality, the above equations in (43) and (44) also imply the following condition $\Delta(\rho(\varepsilon)) \ge 1$.

4. Main results

In this section, we present our main result that asserts the relative error decreases to zero as the small random perturbation tends to zero, which in turn implies the uniform log-efficiency for the estimation problem in (36).

In what follows, let $\hat{\mathbf{x}}^{\varepsilon}_t$ be the solution to the following SDE

$$\begin{aligned} d\hat{\mathbf{x}}^{\varepsilon}_t &= \mathbf{f}\left(t,\hat{\mathbf{x}}^{\varepsilon}_t\right)dt + \mathbf{b}\sigma\left(t,\hat{\mathbf{x}}^{\varepsilon}_t\right)v^{\varepsilon}\left(t,\hat{\mathbf{x}}^{\varepsilon}_t\right)dt + \sqrt{\varepsilon}\mathbf{b}\sigma\left(t,\hat{\mathbf{x}}^{\varepsilon}_t\right)dW_t, \\ &\text{with an initial condition} \quad \hat{\mathbf{x}}^{\varepsilon}_s = \mathbf{x}^{\varepsilon}_s, \end{aligned} \tag{45}$$

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

Let

$$z^{\varepsilon} = \exp\left(-\frac{1}{\sqrt{\varepsilon}}\int_s^T \langle v^{\varepsilon}\left(t,\hat{\mathbf{x}}^{\varepsilon}_t\right), dW_t\rangle - \frac{1}{2\varepsilon}\int_s^T \left|v^{\varepsilon}\left(t,\hat{\mathbf{x}}^{\varepsilon}_t\right)\right|^2 dt\right). \tag{46}$$

Then, the corresponding importance sampling estimator is given by

$$\hat{\rho}(\varepsilon) = \frac{1}{N}\sum_{j=1}^{N} \exp\left(-\frac{1}{\varepsilon}\Phi^{\varepsilon}\left(\hat{\mathbf{x}}^{\varepsilon(j)}\right)\right)z^{\varepsilon(j)}, \tag{47}$$

where $\left\{\left(\hat{\mathbf{x}}^{\varepsilon(j)}, z^{\varepsilon(j)}\right)\right\}_{j=1}^{N}$ are N-copies of independent samples of $(\hat{\mathbf{x}}^{\varepsilon}, z^{\varepsilon})$. Note that, for an appropriately chosen control function v^{ε}, the above importance sampling estimator in (47) is an unbiased estimator for (37), i.e.,

$$\begin{aligned} \mathbb{E}^{\varepsilon}_{s,\mathbf{x}^{\varepsilon}_s}[\hat{\rho}(\varepsilon)] &= \mathbb{E}^{\varepsilon}_{s,\mathbf{x}^{\varepsilon}_s}\left[\exp\left(-\frac{1}{\varepsilon}\Phi^{\varepsilon}(\mathbf{x}^{\varepsilon})\right)\right] \\ &\equiv \mathbb{E}^{\varepsilon}_{s,\mathbf{x}^{\varepsilon}_s}[\rho(\varepsilon)]. \end{aligned} \tag{48}$$

Moreover, the relative estimation error is given by

$$\mathrm{R}_{\mathrm{err}}(\hat{\rho}(\varepsilon)) = \frac{\sqrt{\mathrm{Var}(\hat{\rho}(\varepsilon))}}{\mathbb{E}^{\varepsilon}_{s,\mathbf{x}^{\varepsilon}_s}[\hat{\rho}(\varepsilon)]} \tag{49}$$

which can be rewritten as follows

$$\mathrm{R}_{\mathrm{err}}(\hat{\rho}(\varepsilon)) = \left(1/\sqrt{N}\right)\sqrt{\Delta(\hat{\rho}(\varepsilon)) - 1}, \tag{50}$$

where

$$\Delta(\hat{\rho}(\varepsilon)) = \frac{\mathbb{E}^{\varepsilon}_{s,\mathbf{x}^{\varepsilon}_s}\left[\exp\left(-\frac{2}{\varepsilon}\Phi^{\varepsilon}(\hat{\mathbf{x}}^{\varepsilon})\right)\right](z^{\varepsilon})^2}{\mathbb{E}^{\varepsilon}_{s,\mathbf{x}^{\varepsilon}_s}\left[\exp\left(-\frac{1}{\varepsilon}\Phi^{\varepsilon}(\mathbf{x}^{\varepsilon})\right)\right]^2}. \tag{51}$$

Hence, in order to reduce the relative estimation error $\mathrm{R_{err}}(\hat{\rho}(\varepsilon))$, we need to control the term $\Delta(\hat{\rho}(\varepsilon))$ in (50). Note that, from Jensen's inequality, we have the following condition

$$\limsup_{\varepsilon\to 0} -\varepsilon\,\log \mathbb{E}^{\varepsilon}_{s,\mathbf{x}^{\varepsilon}_s}\left[\exp\left(-\frac{2}{\varepsilon}\Phi^{\varepsilon}(\hat{\mathbf{x}}^{\varepsilon})\right)\right] \le 2\lim_{\varepsilon\to 0} -\varepsilon\,\log \mathbb{E}^{\varepsilon}_{s,\mathbf{x}^{\varepsilon}_s}\left[\exp\left(-\frac{1}{\varepsilon}\Phi^{\varepsilon}(\hat{\mathbf{x}}^{\varepsilon})\right)\right] \tag{52}$$

which also implies $\Delta(\hat{\rho}(\varepsilon)) \ge 1$ with $\lim_{\varepsilon\to 0}\Delta(\hat{\rho}(\varepsilon)) = 1$. Moreover, the statement in (49) further implies the following

$$\mathrm{R_{err}}(\hat{\rho}(\varepsilon)) = \frac{1}{\sqrt{N}}\exp\left(o(1)/\varepsilon\right) \quad \text{as} \quad \varepsilon \to 0, \tag{53}$$

which is generally referred as asymptotic efficiency or optimality. In this paper, our main objective is to choose appropriately the control function v^{ε} in (45), 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 [29] or [12]) which is useful for interpreting our main result.

Definition 1 An importance sampling estimator of the form (47) is log-efficient (i.e., asymptotic efficiency or optimal) if

$$\lim_{\varepsilon\to 0} -\varepsilon \log\Delta(\hat{\rho}(\varepsilon)) = 0. \tag{54}$$

Then, we state the following result as follows.

Proposition 3 Suppose that the importance sampling estimator $\hat{\rho}(\varepsilon)$ in (47), with $v^{\varepsilon}(t,\mathbf{x}) = -\sigma^T(\mathbf{x})\nabla_{\mathbf{x}}J^{\varepsilon}(t,\mathbf{x})$, is uniformly log-efficient (i.e., asymptotic efficient), where $J^{\varepsilon}(t,\mathbf{x})$ satisfies the corresponding dynamic programming equation in Ω^T with respect to the system in (45), with $J^{\varepsilon} = \Phi^{\varepsilon}$ on $\partial^*\Omega^T$. Then, there exits a set $\mathbb{A}\subset\mathbb{R}^3$ such that the Hausedorf dimension of $\mathbb{A}^c$ is zero and

$$\lim_{\varepsilon\to 0} R_{\mathrm{err}}(\hat{\rho}(\varepsilon)) = 0, \tag{55}$$

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

Proof: The above proposition basically asserts that the relative error $\mathrm{R_{err}}(\hat{\rho}(\varepsilon))$ decreases to zero as the small random perturbation level ε tends to zero. Note that, if $J^{\varepsilon}(s,\mathbf{x}^{\varepsilon})$ satisfies the dynamic programming equation in (21), then, with $v^{\varepsilon}(t,\mathbf{x}) = -\sigma^T(\mathbf{x})\nabla_{\mathbf{x}}J^{\varepsilon}(t,\mathbf{x})$, the importance sampling for the estimation problem in (36), i.e., $\mathbb{E}^{\varepsilon}_{s,\mathbf{x}^{\varepsilon}_s}\left[\exp\left(-\frac{1}{\varepsilon}\Phi^{\varepsilon}(\mathbf{x}^{\varepsilon})\right)\right]$, is uniformly log-efficient if the point $\left(s,\mathbf{x}^{\varepsilon}_s\right)$ is contained in a region of sufficient regularity that encompasses almost all $\mathbb{R}^3$. As a result of this, it only suffices to show that

$$\lim_{\varepsilon\to 0}\frac{\mathbb{E}^{\varepsilon}_{s,\mathbf{x}^{\varepsilon}_s}\left[\exp\left(-\frac{2}{\varepsilon}\Phi^{\varepsilon}(\hat{\mathbf{x}}^{\varepsilon})\right)(z^{\varepsilon})^2\right]}{\mathbb{E}^{\varepsilon}_{s,\mathbf{x}^{\varepsilon}_s}\left[\exp\left(-\frac{1}{\varepsilon}\Phi^{\varepsilon}(\mathbf{x}^{\varepsilon})\right)\right]^2} = 1 \tag{56}$$

holds uniformly for all $\left(s,\mathbf{x}^{\varepsilon}_s\right)$ in any compact subset $\overline{\Omega}^T$.

Let us define following two functions

$$\psi_1^\varepsilon\left(s, \mathbf{x}_s^\varepsilon\right) = -\varepsilon \log \mathbb{E}_{s,\mathbf{x}_s^\varepsilon}^\varepsilon \left[\exp\left(-\frac{2}{\varepsilon}\Phi^\varepsilon\left(\hat{\mathbf{x}}^\varepsilon\right)\right)\right] \tag{57}$$

and

$$\begin{aligned} \psi_2^\varepsilon\left(s, \mathbf{x}_s^\varepsilon\right) &= -\varepsilon \log \mathbb{E}_{s,\mathbf{x}_s^\varepsilon}^\varepsilon \left[\exp\left(-\frac{2}{\varepsilon}\Phi^\varepsilon\left(\hat{\mathbf{x}}^\varepsilon\right)\right)\left(z^\varepsilon\right)^2\right] \\ &= -\varepsilon \log \mathbb{E}_{s,\mathbf{x}_s^\varepsilon}^\varepsilon \left[\exp\left(-\frac{2}{\varepsilon}\Phi^\varepsilon\left(\hat{\mathbf{x}}^\varepsilon\right) - \frac{2}{\sqrt{\varepsilon}}\int_s^T \left\langle v^\varepsilon\left(t, \hat{\mathbf{x}}_t^\varepsilon\right), dW_t\right\rangle \right.\right. \\ &\qquad \left.\left. -\frac{1}{\varepsilon}\int_s^T \left|v^\varepsilon\left(t, \hat{\mathbf{x}}_t^\varepsilon\right)\right|^2 dt\right)\right]. \end{aligned} \tag{58}$$

Note that, from the large deviations results for the diffusion process $\hat{\mathbf{x}}_t^\varepsilon$ (e.g., see [21], Chapter 4, [30] or [27], pp.332–334; and see also the asymptotic estimates in Proposition 2 of Section 3), then there exists a constant C, $\gamma > 0$ and ε_0, with $\varepsilon \in (0, \varepsilon_0)$, such that

$$\begin{aligned} \mathbb{E}_{0,\mathbf{x}_0^\varepsilon}^\varepsilon \left[\exp\left(-\frac{1}{\varepsilon}\left(\psi_2^\varepsilon\left(\hat{\tau}^\varepsilon, \hat{\mathbf{x}}_{\hat{\tau}^\varepsilon}^\varepsilon\right) - 2\psi_1^\varepsilon\left(\hat{\tau}^\varepsilon, \hat{\mathbf{x}}_{\hat{\tau}^\varepsilon}^\varepsilon\right)\right) - \int_0^{\hat{\tau}^\varepsilon} \sum_{i,j=1}^3 a_{i,j}\left(\mathbf{x}_s^\varepsilon\right) \frac{\partial^2 {\psi_1}^\varepsilon\left(s, \hat{\mathbf{x}}_s^\varepsilon\right)}{\partial x^i \partial x^j} ds\right)\right] \\ \le C \exp\left(-\gamma/2\varepsilon\right), \end{aligned} \tag{59}$$

where $\hat{\tau}^\varepsilon = \inf\left\{t > s \,|\, \hat{\mathbf{x}}_t^\varepsilon \in \partial\Omega\right\} \wedge T$. Note that the above relation further implies that

$$\lim_{\varepsilon \to 0} \exp\left(-\frac{1}{\varepsilon}\left(\psi_2^\varepsilon(0, \mathbf{x}_0) - 2\psi_1^0(0, \mathbf{x}_0)\right)\right) = \exp\left(\int_0^T \sum_{i,j=1}^3 a_{i,j}\left(\mathbf{x}_s^\varepsilon\right) \frac{\partial^2 {\psi_1}^\varepsilon\left(s, \hat{\mathbf{x}}_s^\varepsilon\right)}{\partial x^i \partial x^j} ds\right). \tag{60}$$

Moreover, in the same way, we can also show the following relation

$$\lim_{\varepsilon \to 0} \exp\left(-\frac{1}{\varepsilon}\left(\psi_1^\varepsilon(0, \mathbf{x}_0) - \psi_1^0(0, \mathbf{x}_0)\right)\right) = \exp\left(\int_0^T \sum_{i,j=1}^3 a_{i,j}\left(\mathbf{x}_s^\varepsilon\right) \frac{\partial^2 {\psi_1}^\varepsilon\left(s, \hat{\mathbf{x}}_s^\varepsilon\right)}{\partial x^i \partial x^j} ds\right). \tag{61}$$

Finally, if we combine the above two equations, then we have the condition following

$$\lim_{\varepsilon \to 0} \exp\left(-\frac{1}{\varepsilon}\left(\psi_2^\varepsilon(0, \mathbf{x}_0) - \psi_1^\varepsilon(0, \mathbf{x}_0)\right)\right) = 1, \tag{62}$$

which implies the uniform log-efficiency for the estimation problem in (36). This completes the proof of Proposition 3.

Remark 2 The above proposition basically ensures a minimum relative estimation error in the small noise limit case for the estimation problem in (36). Note that, if $J^\varepsilon(t, \mathbf{x}^\varepsilon)$ satisfies the dynamic programming equation in (21) (i.e., if it is the solution for the family of stochastic control problems that are associated with the underlying distributed system with small random perturbation). Then, with $v^{\varepsilon}(t, \mathbf{x}) = -\sigma^{T}(\mathbf{x})\nabla_{\mathbf{x}} J^{\varepsilon}(t, \mathbf{x})$, one can provide a numerical computational framework for constructing efficient importance sampling estimators, with an exponential variance decay rate – based on an exponentially-tilted biasing distribution – for rare-event simulations involving the behavior of the diffusion process $\mathbf{x}^{\varepsilon}$.

Remark 3 Here, our primary intent is to provide a theoretical framework, rather than considering some specific numerical simulation results with respect to system parameters (such as the propagation rate β and recovery rate μ of the network), which is an ongoing research area.

5. Concluding remarks

In this chapter, we presented a mathematical framework that provides a new insight for understanding the spread of traffic congestions in an urban network system. In particular, we considered a dynamical model, based on the well-known susceptible-infected-recovered (SIR) model from mathematical epidemiology, with small random perturbations, that describes the process of traffic congestion propagation and dissipation in an urban network system. Moreover, we also provided the asymptotic probability estimate based on the Freidlin-Wentzell theory of large deviations for certain rare events that are difficult to observe in the simulation of an urban traffic network dynamic, where such a framework provides a computational algorithm for constructing efficient importance sampling estimators for rare event simulations of certain events associated with the spread of traffic congestions in the traffic network.

Author details

Getachew K. Befekadu
Department of Electrical and Computer Engineering, Morgan State University, Baltimore, USA

*Address all correspondence to: getachew.befekadu@morgan.edu

References

[1] Meead Saberi, Mahmassani HS. Exploring properties of network-wide flow-density relations in a freeway network. Transp. Res. Rec. 2012;**2315**: 153-163. https://doi.org/10.3141/2315-16 Accessed: 27 December 2020

[2] Meead Saberi, Hani S. Mahmassani: Hysteresis and capacity drop phenomena in freeway networks: empirical characterization and interpretation. Transp. Res. Rec. 2013; 2391:44–55. https://doi.org/10.3141/2391-05 [Accessed: 27 December 2020]

[3] Nikolas Geroliminis, Carlos F. Daganzo: Existence of urban-scale macroscopic fundamental diagrams: some experimental findings. Transp. Res. B. 2008; 42:759–770. https://doi.org/10.1016/j.trb.2008.02.002 [Accessed: 27 December 2020]

[4] Yuxuan Ji, Nikolas Geroliminis: On the spatial partitioning of urban transportation networks. Transp. Res. B. 2012; 46:1639–1656. https://doi.org/10.1016/j.trb.2012.08.005 [Accessed: 27 December 2020]

[5] Mohammadreza Saeedmanesh, Nikolas Geroliminis: Dynamic clustering and propagation of congestion in heterogeneously congested urban traffic networks. Transp. Res. Procedia. 2017; 23:962–979. https://doi.org/10.1016/j.trb.2017.08.021 [Accessed: 27 December 2020]

[6] Guanwen Zeng, Daqing Li, Shengmin Guo, Liang Gao, Ziyou Gao, H. Eugene Stanley, Shlomo Havlin: Switch between critical percolation modes in city traffic dynamics. Proc. Natl Acad Sci. USA. 2019; 116:23–28. https://doi.org/10.1073/pnas.1801545116 [Accessed: 27 December 2020]

[7] Daqing Li, Bowen Fu, Yunpeng Wang, Guangquan Lu, Yehiel Berezin, H. Eugene Stanley, Shlomo Havlin: Percolation transition in dynamical traffic network with evolving critical bottlenecks. Proc. Natl Acad. Sci. USA. 2015; 112:669–672. https://doi.org/10.1073/pnas.1419185112 [Accessed: 27 December 2020]

[8] M.T. Asif, J. Dauwels, C.Y. Goh, A. Oran, E. Fathi, M. Xu, M.M. Dhanya, N. Mitrovic and P. Jaillet: Spatio-temporal patterns in large-scale traffic speed prediction. IEEE Trans. Intell. Transp. Syst. 2014; 15:794–804. https://doi.org/10.1109/TITS.2013.2290285 [Accessed: 27 December 2020]

[9] Richard Arnott: A bathtub model of downtown traffic congestion. J. Urban Econ. 2013; 76:110–121. https://doi.org/10.1016/j.jue.2013.01.001 [Accessed: 27 December 2020]

[10] William S. Vickrey: Congestion theory and transport investment. Am. Econ. Rev. 1969; 59:251–260 https://www.jstor.org/stable/1823678 [Accessed: 27 December 2020]

[11] Meead Saberi, H. Hamedmoghadam, M. Ashfaq, et al. : A simple contagion process describes spreading of traffic jams in urban networks. Nat Commun. 2020; 11:1616. https://doi.org/10.1038/s41467-020-15353-2 [Accessed: 27 December 2020]

[12] Amarjit Budhiraja, Paul Dupuis: Analysis and approximation of rare events: representations and weak convergence methods. Prob. Theory and Stoch. Modelling Series, 94, Springer, 2019. https://doi.org/10.1007/978-1-4939-9579-0 [Accessed: 27 December 2020]

[13] Paul Dupuis Richard S. Ellis: A weak convergence approach to the theory of large deviations. Wiley, New York, 1997.

[14] Paul Dupuis, Hui Wang: Importance sampling, large deviations, and

differential games. Stoch. Stoch. Rep. 2004; 76:481–508. https://doi.org/10.1080/10451120410001733845 [Accessed: 27 December 2020]

[15] Weinan E, Weiqing Ren, Eric Vanden-Eijnden: Minimum action method for the study of rare events. Comm. Pure Appl. Math. 2004; 57:637–656. https://doi.org/10.1002/cpa.20005 [Accessed: 27 December 2020]

[16] Eric Vanden-Eijnden, Jonathan Weare: Rare event simulation of small noise diffusions. Comm. Pure Appl. Math. 2012; 65:1770–1803. https://doi.org/10.1002/cpa.21428 [Accessed: 27 December 2020]

[17] D. Siegmund: Importance sampling in the monte carlo study of sequential tests. The Annals of Statistics. 1976; 673–684. https://doi.org/10.1214/aos/1176343541 [Accessed: 27 December 2020]

[18] Soren Asmussen, Peter W. Glynn: Stochastic simulation: algorithms and analysis. Springer, New York, 2010. https://doi.org/10.1007/978-0-387-69033-9 [Accessed: 27 December 2020]

[19] Jügen Gärtner: On large deviations from the invariant measure. Theory Probab. Appl. 1977; 22(1), 24–39. https://doi.org/10.1137/1122003 [Accessed: 27 December 2020]

[20] A. D. Ventsel, M. I. Freidlin: On small random perturbations of dynamical systems. Russian Math. Surv. 1970; 25:1–55. https://doi.org/10.1070/RM1970v025n01ABEH001254 [Accessed: 27 December 2020]

[21] Mark I. Freidlin, and Alexander D. Wentzell: Random perturbations of dynamical systems. Springer, Berlin, 1984. https://doi.org/10.1007/978-3-642-25847-3 [Accessed: 27 December 2020]

[22] Wendell H. Fleming: Exit probabilities and optimal stochastic control. Appl. Math. Optim. 1978; 4:329–346. https://doi.org/10.1007/BF01442148 [Accessed: 27 December 2020]

[23] Wendell Fleming, Raymond Rishel: Deterministic and stochastic optimal control. Springer-Verlag, New York, 1975. https://doi.org/10.1007/978-1-4612-6380-7 [Accessed: 27 December 2020]

[24] D. Stroock, S. R. S. Varadhan: On degenerate elliptic-parabolic operators of second order and their associated diffusions. Comm. Pure Appl. Math., 1972; 25:651–713. https://doi.org/10.1002/cpa.3160250603 [Accessed: 27 December 2020]

[25] Getachew K. Befekadu, Panos 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. 2015; 53:2297–2318. https://doi.org/10.1137/140990322 [Accessed: 27 December 2020]

[26] L. C. Evans, H. Ishii: A PDE approach to some asymptotic problems concerning random differential equations with small noise intensities. Ann. Inst. H. Poincaré Anal. Non Linearé. 1985; 2:1–20. https://doi.org/10.1016/S0294-1449(16)30409-7 [Accessed: 27 December 2020]

[27] Avner Friedman: Stochastic differential equations and applications. Dover Publisher, Inc. Mineola, New York, 2006.

[28] A. D. Ventcel: Limit theorems on large deviations for stochastic processes. Theo. Prob. Appl. 1973; 18:817–821. https://doi.org/10.1137/1121030 [Accessed: 27 December 2020]

[29] James A. Bucklew: Introduction to rare-event simulation. Springer Series in

Statistics. Springer, New York, 2004. https://doi.org/10.1007/978-1-4757-4078-3 [Accessed: 27 December 2020]

[30] S. R. S. Varadhan: Large deviations and applications, CBMS-NSF Regional Conference Series in Applied Mathematics, 46. SIAM, Philadelphia, 1985. https://doi.org/10.1137/1.9781611970241 [Accessed: 27 December 2020]