Check for updates

J. Appl. Probab. **63,** 890–918 (2026)
*doi:*10.1017/jpr.2025.10062

EULER–MARUYAMA APPROXIMATION FOR STABLE STOCHASTIC DIFFERENTIAL EQUATIONS WITH MARKOVIAN SWITCHING AND RELATED ASYMPTOTICS

XINGHU JIN,* ** AND
XUN XU,* *Hefei University of Technology*

Abstract

We investigate the EM approximation for $\mathbb{R}^d$-valued ergodic stochastic differential equations (SDEs) driven by rotationally invariant α-stable processes ($\alpha \in (1, 2)$) with Markovian switching. The coefficient g violates the dissipative condition for certain states of the switching process. Using the Lindeberg principle, we establish quantitative error bounds between the original process $(X_t, R_t)_{t\geqslant 0}$ and its Euler–Maruyama (EM) scheme under a specially designed metric. Furthermore, we derive both a central limit theorem and a moderate derivation principle for the empirical measures of both the SDE and its EM scheme. The theoretical results are subsequently validated through a concrete example.

Keywords: Euler–Maruyama scheme; rotationally invariant α-stable process; Markovian switching; central limit theorem; moderate deviation principle

2020 Mathematics Subject Classification: Primary 60B10
Secondary 60G51; 60J25; 60J76

1. Introduction

1.1. Setting

Since the early 1990s, stochastic differential equations (SDEs) with Markovian switching have attracted considerable interest. These models comprise two key components: (i) a continuous-time dynamical system that follows one of finitely many underlying Markovian dynamics, with transitions dictated by (ii) an underlying continuous-time Markov chain. Such random switching processes find broad applications in fields like financial engineering, risk management, and operations research [3, 8, 12, 18, 23, 30, 31, 36, 37, 38, 39, 41].

We study the following stable SDE with Markovian switching: the first component $(X_t)_{t\geqslant 0}$ satisfies

$$\mathrm{d}X_t = g(X_t, R_t)\,\mathrm{d}t + \sigma\,\mathrm{d}Z_t, \quad (X_0, R_0) \in \mathbb{R}^d \times \mathbb{S}, \tag{1.1}$$

where $(Z_t)_{t\geqslant 0}$ is a d-dimensional rotationally invariant α-stable process with $\alpha \in (1, 2)$ and σ is a nondegenerate constant matrix, the switching part $(R_t)_{t\geqslant 0}$ is a continuous-time Markov chain with a finite state space $\mathbb{S} = \{1, 2, \ldots, N\}$ and an irreducible transition rate matrix

Received 25 February 2025; accepted 30 October 2025.
* Postal address: School of Mathematics, Hefei University of Technology, Hefei, Anhui, China.
** Email address: xinghujin@hfut.edu.cn

© The Author(s), 2026. Published by Cambridge University Press on behalf of Applied Probability Trust.

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

$Q = (q_{ij})_{1 \leqslant i,j \leqslant N}$. The entries of Q satisfy $q_{ij} \geqslant 0$ for all $i \neq j$ and $q_{ii} = -\sum_{j \neq i} q_{ij}$ for all $i \in \mathbb{S}$, and the transition probability for $(R_t)_{t \geqslant 0}$ is given by e^{Qt} [23, Theorem 1.44], which implies that, for any $u, v \in \mathbb{S}$,

$$\mathbb{P}(R_{t+\Delta} = v \mid R_t = u) = \begin{cases} q_{uv}\Delta + o(\Delta) & \text{if } v \neq u, \\ 1 + q_{uv}\Delta + o(\Delta) & \text{if } v = u \end{cases}$$

for some small enough $\Delta > 0$, where $a = o(\Delta)$ means $\lim_{\Delta \to 0} a/\Delta = 0$.

We impose the following assumptions on the coefficient g in SDE (1.1).

Assumption 1.1.

(i) *There exist* $\{c_i, i \in \mathbb{S}\}$ *with* $c_i \in \mathbb{R}$ *such that, for all* $x, y \in \mathbb{R}^d$ *and* $i \in \mathbb{S}$,

$$\langle x, \nabla g(y, i)x \rangle \leqslant c_i |x|^2. \tag{1.2}$$

(ii) *For* $i \in \mathbb{S}$, $g(\cdot, i) \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^d)$ *and* $\nabla^m g(\cdot, i)$ $(m = 0, 1, 2)$ *are Lipschitz continuous.*

Under Assumption 1.1, the system in (1.1) admits a unique strong solution $(X_t, R_t)_{t \geqslant 0}$ for any initial value $(X_0, R_0) \in \mathbb{R}^d \times \mathbb{S}$ (see [23] and [38, Theorem 2.1]). Conditioned on $R_t = i$, the first component, denoted by $X_t^{(i)}$, follows $\mathrm{d}X_t^{(i)} = g(X_t^{(i)}, i)\,\mathrm{d}t + \sigma\,\mathrm{d}Z_t$. Crucially, the constants $\{c_\ell, \ell \in \mathbb{S}\}$ take values in $\mathbb{R}$, rather than $(-\infty, 0)$; this is a pivotal condition for the coefficient g: for some switching state $i \in \mathbb{S}$ with $c_i > 0$, the corresponding $X_t^{(i)}$ may exhibit explosive behavior as $t \to \infty$, but the switching component $(R_t)_{t \geqslant 0}$ can transition to a state $j \neq i$ with $c_j < 0$, which acts to stabilize the process $(X_t)_{t \geqslant 0}$. Thus, the switching component $(R_t)_{t \geqslant 0}$ regulates the dynamics by alternating between different stable SDEs $(X_t^{(i)})_{t \geqslant 0}$ with $i \in \mathbb{S}$.

To approximate the law of the process $(X_t, R_t)_{t \geqslant 0}$ in SDE (1.1), we construct an Euler–Maruyama (EM) scheme with step size γ. For $(Y_0, R_0) = (X_0, R_0)$ and $n \in \mathbb{N}_0 = \mathbb{N}_+ \cup \{0\}$ with $\mathbb{N}_+ = \{1, 2, \ldots\}$, the EM scheme reads $Y_{t_{n+1}} = Y_{t_n} + \gamma g(Y_{t_n}, R_{t_n}) + \sigma \Delta Z_\gamma$, where $t_n = n\gamma$, $\Delta Z_\gamma = Z_{t_{n+1}} - Z_{t_n}$, and $(R_{t_n})_{n \in \mathbb{N}_0}$ is a Markov chain with one-step transition probability matrix $\mathrm{e}^{\gamma Q}$.

1.2. Related works

Since the transition rate matrix Q is constant, the switching process $(R_t)_{\geqslant 0}$ is a continuous-time Markov chain that is independent of the rotationally invariant α-stable process $(Z_t)_{t \geqslant 0}$. For more details about SDEs with Markovian switching, we refer the reader to [23, 39]. Extensive research has established fundamental results on the stability, ergodicity, and asymptotic behavior of such processes [2, 8, 18, 19, 22, 27, 30, 31, 35–38, 40, 41].

Recent studies have made significant progress in analyzing regime-switching processes through various approaches. [18, 38] employed coupling techniques and Radon–Nikodym derivatives to investigate both Feller and strong Feller properties along with exponential ergodicity for regime-switching jump processes with countable state spaces. Based on the strong Feller property and irreducibility, [8, 19, 30] developed verifiable criteria for the ergodicity of random switching processes driven by Brownian motion or pure jump processes. Stability properties were thoroughly examined in [12, 22, 23, 39], while asymptotic behavior was considered in [3, 36, 37, 39]. [35] used the Radon–Nikodym derivative and constructed an auxiliary process to derive fundamental properties, including Feller and strong Feller properties and ergodicity, for state-dependent switching diffusion processes.

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

1.2.1. *EM scheme for SDEs with regime switching.* Closed-form expressions for the laws or ergodic measures of SDEs are generally unavailable, with notable exceptions including one-dimensional diffusions [16, p. 171] and Langevin SDEs [10, p. 555]. The EM scheme provides a practical approximation for these measures. Recent advances include the following: [10] used Malliavin calculus and Stein's equation to investigate the Wasserstein-1 distance for diffusions and their constant step-size EM schemes; [4, 28] applied Malliavin calculus and the Lindeberg principle to discuss the Wasserstein-1 and total variation distances for diffusions and their various step-size EM schemes; and [5, 7] were concerned with constant and variable step-size EM schemes for stable SDEs and obtained the convergence rate in Wasserstein-1 distance. Additionally, strong convergence properties were established in [20, 31, 40].

While explicit solutions for SDEs with Markovian switching remain elusive, the EM scheme provides an effective approximation tool. Previous studies [2, 31, 39, 40] established convergence results for random switching diffusions using the EM scheme. Recently, [6, 14] used constant and decreasing step-size EM schemes to approximate the law for regime-switching jump processes, though limited to dissipative drift conditions and excluding rotationally invariant α-stable processes. This paper extends their framework to SDEs driven by d-dimensional rotationally invariant α-stable processes ($\alpha \in (1, 2)$), analyzing both the EM approximation and its asymptotic properties.

1.2.2. *Central limit theorem and moderate deviation principle for empirical measures.* Propositions A.1 and A.2 establish exponential ergodicities for both the original process $(X_t, R_t)_{t\geqslant 0}$ and its EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$ with ergodic measures μ and μ_γ respectively. By the Birkhoff ergodic theorem, the empirical measures $\mathcal{E}_T^{(x,k)}$ and $\mathcal{E}_{n,\gamma}^{(x,k)}$ (definitions in (2.1) and (2.2) respectively) converge almost surely to μ and μ_γ, i.e. $\lim_{T\to\infty} \mathcal{E}_T^{(x,k)} = \mu$ and $\lim_{n\to\infty} \mathcal{E}_{n,\gamma}^{(x,k)} = \mu_\gamma$ almost surely, as shown in [9]. For diffusion processes, prior work has investigated limit theorems for empirical measures: [13] derived central limit theorems (CLTs) and moderate deviation principles (MDPs) for both the diffusion and its EM scheme, while [21] established the CLT and Cramér-type moderate deviation for the EM approximation. In the case of $\alpha \in (1, 2)$, it is natural to investigate whether analogous CLT and MDP results hold for the regime-switching process $(X_t, R_t)_{t\geqslant 0}$ and its EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$.

1.3. Preliminaries

For any given real number κ, define the matrix $Q_{c,\kappa}$ as

$$Q_{c,\kappa} = Q + \kappa \operatorname{diag}(c_1, \ldots, c_N), \tag{1.3}$$

where $\operatorname{diag}(c_1, \ldots, c_N)$ is the diagonal matrix formed by the vector $(c_1, \ldots, c_N)$, and the constants $c_1, \ldots, c_N$ are defined in (1.2).

Assumption 1.2. *There exist two constants $\kappa_1 \in (1, \alpha)$ and $\kappa_2 \in (0, \kappa_1/4]$ such that the matrices $-Q_{c,\kappa_1}$ and $-Q_{c,\kappa_2}$ are both nonsingular M-matrices.*

It follows from [31, Lemma 2.1] or [3, Proposition 4.1] that the following lemma holds.

Lemma 1.1. *Under Assumptions 1.1 and 1.2, and for the matrix $Q_{c,4}$ defined in (1.3), let*

$$\eta = -\max_{\theta\in\operatorname{spec}(Q_{c,4})} \operatorname{Re}(\theta), \tag{1.4}$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

where $\operatorname{spec}(Q_{c,4})$ *denotes the spectrum of matrix* $Q_{c,4}$, *and* $\operatorname{Re}(\theta)$ *is the real part of* θ. *Assume that* $\eta > 0$. *Then, for* $\{c_\ell, \ell \in \mathbb{S}\}$ *in (1.2), there exists some constant* $C > 0$ *such that*

$$\mathbb{E}\exp\left(4\int_0^t c_{R_s}\,\mathrm{d}s\right) \leqslant C\mathrm{e}^{-\eta t} \quad \textit{for all } t \geqslant 0.$$

Remark 1.1. The parameter κ_1 in Assumption 1.2 is crucial for obtaining the convergence rate of the EM approximation (see more details in Theorem 2.1) and for the moment estimates (see Lemma 3.2). Assume that $\eta > 0$. Lemma 1.1 plays an important role in estimating Jacobi flows for the process $(X_t, R_t)_{t\geqslant 0}$ (see Lemma C.1 for further details).

1.4. Notation

Denote by $(X_t^{x,k}, R_t^k)$ the solution to the system (1.1) with initial value $(X_0, R_0) = (x, k)$, and define its semigroup P_t by $P_t f(x,k) = \mathbb{E}f(X_t^{x,k}, R_t^k)$ for any bounded Borel function $f \in \mathcal{B}_{\mathrm{b}}(\mathbb{R}^d \times \mathbb{S}, \mathbb{R})$ and $(x,k) \in \mathbb{R}^d \times \mathbb{S}$. For the EM scheme $(Y_{t_n}^{x,k}, R_{t_n}^k)$ initialized at $(Y_0, R_0) = (x,k)$, let $\bar{P}_\gamma$ be its transition semigroup $\bar{P}_\gamma f(x,k) = \mathbb{E}f(Y_\gamma^{x,k}, R_\gamma^k)$, with $\bar{P}_\gamma^m = \bar{P}_\gamma \circ \bar{P}_\gamma^{m-1}$ for integers $m \geqslant 2$.

Let M be a matrix, and denote its transpose by M^*. We write $M \geqslant 0$ if all entries of M are nonnegative, and $M \gg 0$ if all entries are strictly positive. For a vector M, M_i represents its ith component. The Hilbert–Schmidt norm of an $\mathbb{R}^{d\times d}$ matrix $M = (M_{ij})_{d\times d}$ is given by $\|M\|_{\mathrm{HS}} = \sqrt{\sum_{i,j=1}^d M_{ij}^2}$.

Since the state space for the process $(X_t, R_t)_{t\geqslant 0}$ is $\mathbb{R}^d \times \mathbb{S}$, and for $f(\cdot,\cdot) \in \mathcal{C}^m(\mathbb{R}^d \times \mathbb{S})$ with integers $m \geqslant 1$, $\nabla^m f(x,k)$ refers to the mth-order partial derivative with respective to x (with k fixed), and $\nabla_u f$ represents the directional derivative along the direction $u \in \mathbb{R}^d$. For $f \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{S}, \mathbb{R})$, define

$$\|\nabla^i f\|_\infty = \sup_{k\in\mathbb{S}} \|\nabla^i f(\cdot,k)\|_\infty = \sup_{(x,k)\in\mathbb{R}^d\times\mathbb{S}} |\nabla^i f(x,k)|, \quad i = 0, 1;$$

$$\|\nabla^2 f\|_\infty = \sup_{k\in\mathbb{S}} \|\nabla^2 f(\cdot,k)\|_{\mathrm{HS}} = \sup_{(x,k)\in\mathbb{R}^d\times\mathbb{S}} \|\nabla^2 f(x,k)\|_{\mathrm{HS}}.$$

For a τ-Hölder continuous function f defined on A with $\tau \in (0,1]$, define

$$[f]_{\tau;A} = \sup_{x,y\in A, x\neq y} \frac{|f(x)-f(y)|}{|x-y|^\tau}.$$

Let $\mathcal{G} = \{h \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{S}, \mathbb{R})\colon \|\nabla^i h\|_\infty < \infty \text{ with } i = 1, 2\}$. We introduce a metric $\operatorname{dist}(\cdot,\cdot)$ on $\mathbb{R}^d \times \mathbb{S}$ as follows (see [39, (9.6)] or [31, p. 3229]):

$$\operatorname{dist}((x,i),(y,j)) = |x-y| + d(i,j), \qquad \text{where } d(i,j) = \begin{cases} 1 & \text{if } j \neq i, \\ 0 & \text{if } j = i. \end{cases}$$

Let $\mathcal{B}(\mathbb{R}^d \times \mathbb{S})$ be the collection of Borel sets on $\mathbb{R}^d \times \mathbb{S}$, and $\mathcal{P}$ denote the space of probability measures on $(\mathbb{R}^d \times \mathbb{S}, \mathcal{B}(\mathbb{R}^d \times \mathbb{S}))$; define

$$\mathcal{P}_0 = \left\{\mu_1 \in \mathcal{P}\colon \mu_1(h) = \int_{\mathbb{R}^d\times\mathbb{S}} h(x,k)\,\mathrm{d}\mu_1(x,k) < \infty \text{ for any } h \in \mathcal{G}\right\}.$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

For $\mu_1, \mu_2 \in \mathcal{P}_0$, we define the distance $d_{\mathcal{G}}(\mu_1, \mu_2)$ as

$$d_{\mathcal{G}}(\mu_1, \mu_2) = \sup_{h\in\mathcal{G}} \left\{ \int_{\mathbb{R}^d\times\mathbb{S}} h(x,k)\,\mathrm{d}\mu_1(x,k) - \int_{\mathbb{R}^d\times\mathbb{S}} h(x,k)\,\mathrm{d}\mu_2(x,k) \right\}. \tag{1.5}$$

Clearly, $d_{\mathcal{G}}(\cdot,\cdot)$ is a metric on $\mathcal{P}_0$. For any random vector X, denote its law by $\mathcal{L}(X)$.

A sequence of random variables $\{Y_n, n \geqslant 1\}$ is said to converge weakly (or converge in distribution) to a limit Y_∞; we write $Y_n \Rightarrow Y_\infty$ if $\lim_{n\to\infty} \mathbb{E}f(Y_n) = \mathbb{E}f(Y_\infty)$ for all bounded continuous functions f. Additionally, $Y_n \xrightarrow{\mathrm{p}} Y_\infty$ means convergence in probability, namely, $\lim_{n\to\infty} \mathbb{P}(|Y_n - Y_\infty| > \delta) = 0$ for all $\delta > 0$.

For a given probability measure ν, define $L^2(\nu)$ as the Hilbert space induced by ν with inner product $\langle f_1, f_2\rangle_\nu = \int f_1 f_2\,\mathrm{d}\nu$ for $f_1, f_2 \in L^2(\nu)$.

1.5. Strategy

We conclude this section by outlining the proof strategy for our main results. The main results presented in Section 2 include the EM approximation, CLT, and MDP for empirical measures of an SDE and its EM scheme, with detailed proofs provided in Sections 3, 4, and 5 respectively. Our approach involves approximating the law of the process $(X_t, R_t)_{t\geqslant 0}$ by its EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$, and establishing the convergence rate under the metric $d_{\mathcal{G}}$, as shown in Theorem 2.1. Applying the Lindeberg principle from [28], we know that

$$\begin{aligned} d_{\mathcal{G}}\big(\mathcal{L}\big(X_{t_n}^{x,k}, R_{t_n}^k\big), \mathcal{L}\big(Y_{t_n}^{x,k}, R_{t_n}^k\big)\big) &= \sup_{h\in\mathcal{G}}\{P_{t_n}h(x,k) - \bar{P}_{t_n}h(x,k)\} \\ &= \sup_{h\in\mathcal{G}} \left\{ \sum_{i=1}^{n} \bar{P}_{t_{i-1}} \circ (P_\gamma - \bar{P}_\gamma) \circ P_{t_n - t_i} h(x,k) \right\}. \end{aligned}$$

The term $(P_\gamma - \bar{P}_\gamma)f(x,k)$ with $f(x,k) = P_{t_n - t_i}h(x,k)$ plays a crucial role in estimating $d_{\mathcal{G}}\big(\mathcal{L}\big(X_{t_n}^{x,k}, R_{t_n}^k\big), \mathcal{L}\big(Y_{t_n}^{x,k}, R_{t_n}^k\big)\big)$. We refer to this term as the *one-step difference* between the process $(X_t, R_t)_{t\geqslant 0}$ and its EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$; this is given in Lemma 3.3, whose proof is provided in Appendix B. After estimating the one-step difference, *gradient estimates* for the semigroup P_t are also essential (see Lemma 3.1), and their proofs can be found in Appendix C. Additionally, *moment estimates* for the process $(X_t, R_t)_{t\geqslant 0}$ and its EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$ are given in Lemma 3.2 with proofs in Appendix A, obtained by the Lyapunov condition and the generator comparison. Finally, we employ various types of Stein's equations to prove the CLT and MDP for empirical measures of the SDE and its EM scheme.

2. Main results and an example

2.1. Main results

Our first main result concerns the EM approximation, which quantifies the error between the laws of the SDE $(X_t, R_t)_{t\geqslant 0}$ and its EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$; the proof is provided in Section 3.

Theorem 2.1. *Under Assumptions 1.1 and 1.2, with $\eta > 0$ defined in (1.4), there exists some positive constant C, independent of γ, x, k, and t_n, such that*

$$d_{\mathcal{G}}\big(\mathcal{L}\big(X_{t_n}^{x,k}, R_{t_n}^k\big), \mathcal{L}\big(Y_{t_n}^{x,k}, R_{t_n}^k\big)\big) \leqslant C(1 + |x|^{\kappa_1})\gamma^{\kappa_1/\alpha},$$

where the distance $d_{\mathcal{G}}$ is defined in (1.5), and the constant κ_1 is from Assumption 1.2.

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

Remark 2.1. When the state space $\mathbb{S} = \{1\}$ and $c_1 < 0$ with c_1 defined in (1.2), we know that Assumption 1.2 holds for any $\kappa_1 \in (1, \alpha)$. In this case, the process $(X_t, R_t)_{t\geqslant 0}$ reduces to a stable SDE $(X_t)_{t\geqslant 0}$ without Markovian switching. The convergence rate for the EM scheme aligns precisely with those in [5, Theorem 1.3] and [7, Theorem 1.9].

Our second set of main results establishes the CLT and MDP for the long-term behavior of the process $(X_t, R_t)_{t\geqslant 0}$ and its EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$; proofs of these results can be found in Sections 4 and 5 respectively.

For any $(x, k) \in \mathbb{R}^d \times \mathbb{S}$ and $T > 0$, the empirical measure $\mathcal{E}_T^{(x,k)}$ of the process $(X_t, R_t)_{t\geqslant 0}$ is defined by

$$\mathcal{E}_T^{(x,k)}(A) = \frac{1}{T}\int_0^T \delta_{(X_s^{x,k}, R_s^k)}(A)\,\mathrm{d}s, \quad A \in \mathcal{B}(\mathbb{R}^d \times \mathbb{S}), \tag{2.1}$$

where $\delta_y(\cdot)$ is the Dirac delta measure defined as $\delta_y(A) = 1$ if $y \in A$ and $\delta_y(A) = 0$ otherwise. It is straightforward to verify that, for any continuously differentiable bounded Borel-measurable function $h \in \mathcal{C}_{\mathrm{b}}^1(\mathbb{R}^d \times \mathbb{S}, \mathbb{R})$, $\mathcal{E}_T^{(x,k)}(h) = (1/T)\int_0^T h(X_s^{x,k}, R_s^k)\,\mathrm{d}s$.

For any $(x, k) \in \mathbb{R}^d \times \mathbb{S}$ and $n \in \mathbb{N}_+$, the empirical measure $\mathcal{E}_{n,\gamma}^{(x,k)}$ of the EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$ is defined by

$$\mathcal{E}_{n,\gamma}^{(x,k)}(A) = \frac{1}{n}\sum_{i=0}^{n-1} \delta_{(Y_{t_i}^{x,k}, R_{t_i}^k)}(A), \quad A \in \mathcal{B}(\mathbb{R}^d \times \mathbb{S}). \tag{2.2}$$

It is easy to verify that, for any $h \in \mathcal{C}_{\mathrm{b}}^1(\mathbb{R}^d \times \mathbb{S}, \mathbb{R})$, $\mathcal{E}_{n,\gamma}^{(x,k)}(h) = (1/n)\sum_{i=0}^{n-1} h\big(Y_{t_i}^{x,k}, R_{t_i}^k\big)$.

Theorem 2.2. (Central limit theorem.) *Under Assumptions 1.1 and 1.2, with $\eta > 0$ defined in (1.4), for $h \in \mathcal{C}_{\mathrm{b}}^1(\mathbb{R}^d \times \mathbb{S}, \mathbb{R})$ and $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, the term $\sqrt{t}\big[\mathcal{E}_t^{(x,k)}(h) - \mu(h)\big]$ converges weakly to the Gaussian distribution $\mathcal{N}(0, \mathcal{V}(f_h))$ as $t \to \infty$, where $\mathcal{V}(f_h)$ is given by*

$$\begin{aligned}\mathcal{V}(f_h) = &\int_{\mathbb{R}^d\times\mathbb{S}}\int_{\mathbb{R}_0^d} |f_h(x + \sigma z, k) - f_h(x, k)|^2\,\nu(\mathrm{d}z)\,\mathrm{d}\mu(x, k)\\ &+ \int_{\mathbb{R}^d\times\mathbb{S}}\int_{\mathbb{R}_0} |f_h(x, k + \ell(k, \tilde{z})) - f_h(x, k)|^2\,\mathrm{d}\tilde{z}\,\mathrm{d}\mu(x, k),\end{aligned} \tag{2.3}$$

where μ is the ergodic measure for the process $(X_t, R_t)_{t\geqslant 0}$, f_h is the solution to Stein's equation (4.5), ℓ is defined in (4.1), and ν is the Lévy measure in (4.2).

Theorem 2.3. (Moderate deviation principle.) *Under Assumptions 1.1 and 1.2, with $\eta > 0$ defined in (1.4), for any $h \in \mathcal{C}_{\mathrm{b}}^1(\mathbb{R}^d \times \mathbb{S}, \mathbb{R})$, $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, and any measurable set $A \subseteq \mathbb{R}$, we have*

$$\begin{aligned}-\inf_{z\in\mathring{A}} \frac{z^2}{2\mathcal{V}(f_h)} &\leqslant \liminf_{t\to\infty} \frac{1}{a_t^2}\log\mathbb{P}\left(\frac{\sqrt{t}}{a_t}\big[\mathcal{E}_t^{(x,k)}(h) - \mu(h)\big] \in A\right)\\ &\leqslant \limsup_{t\to\infty} \frac{1}{a_t^2}\log\mathbb{P}\left(\frac{\sqrt{t}}{a_t}\big[\mathcal{E}_t^{(x,k)}(h) - \mu(h)\big] \in A\right) \leqslant -\inf_{z\in\bar{A}} \frac{z^2}{2\mathcal{V}(f_h)},\end{aligned}$$

where μ is the ergodic measure for the process $(X_t, R_t)_{t\geqslant 0}$, sets $\bar{A}$ and $\mathring{A}$ denote the closure and interior of A respectively, the function a_t satisfies $a_t \to \infty$ and $t^{-1/2}a_t \to 0$ as $t \to \infty$, and $\mathcal{V}(f_h)$ is defined in (2.3).

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

Theorem 2.4. (CLT and MDP.) *Under Assumptions 1.1 and 1.2, with $\eta > 0$ defined in (1.4), let μ_γ be the ergodic measure for the EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$. For $h \in \mathcal{C}_b^1(\mathbb{R}^d \times \mathbb{S}, \mathbb{R})$ and $(x, k) \in \mathbb{R}^d \times \mathbb{S}$,*

(i) *The term $\sqrt{n}\big[\mathcal{E}_{n,\gamma}^{(x,k)}(h) - \mu_\gamma(h)\big]$ converges weakly to the Gaussian distribution $\mathcal{N}(0, \widetilde{\mathcal{V}}(f_{\gamma,h}))$ as $n \to \infty$, where $\widetilde{\mathcal{V}}(f_{\gamma,h})$ is given by*

$$\widetilde{\mathcal{V}}(f_{\gamma,h}) = \langle f_{\gamma,h}, f_{\gamma,h}\rangle_{\mu_\gamma} - \langle \bar{P}_\gamma f_{\gamma,h}, \bar{P}_\gamma f_{\gamma,h}\rangle_{\mu_\gamma}, \tag{2.4}$$

with $f_{\gamma,h}$ the solution to Stein's equation (5.1). Moreover, the term $\widetilde{\mathcal{V}}(f_{\gamma,h})$ can also be expressed as

$$\widetilde{\mathcal{V}}(f_{\gamma,h}) = \langle h - \mu_\gamma(h), h - \mu_\gamma(h)\rangle_{\mu_\gamma} + \sum_{i=1}^{\infty} \langle \bar{P}_\gamma^i h, h - \mu_\gamma(h)\rangle_{\mu_\gamma}.$$

(ii) *For any measurable set $A \subseteq \mathbb{R}$, we know that*

$$-\inf_{z\in\mathring{A}} \frac{z^2}{2\widetilde{\mathcal{V}}(f_{\gamma,h})} \leqslant \liminf_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(\frac{\sqrt{n}}{a_n}\big[\mathcal{E}_{n,\gamma}^{(x,k)}(h) - \mu_\gamma(h)\big] \in A\right)$$
$$\leqslant \limsup_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(\frac{\sqrt{n}}{a_n}\big[\mathcal{E}_{n,\gamma}^{(x,k)}(h) - \mu_\gamma(h)\big] \in A\right) \leqslant -\inf_{z\in\bar{A}} \frac{z^2}{2\widetilde{\mathcal{V}}(f_{\gamma,h})},$$

where $\bar{A}$ and $\mathring{A}$ are the closure and interior of the set A respectively, the function a_n satisfies $a_n \to \infty$ and $n^{-1/2}a_n \to 0$ as $n \to \infty$, and $\widetilde{\mathcal{V}}(f_{\gamma,h})$ is defined in (2.4).

2.2. An example

We begin by recalling several equivalent characterizations of nonsingular M-matrices; detailed treatments can be found in [30, Proposition 2.4] and [22, Lemma 4.2].

Lemma 2.1. *The following statements are equivalent:*

(i) *The matrix A is a nonsingular $d \times d$ M-matrix.*

(ii) *All of the principal minors of matrix A are positive, i.e.*

$$\begin{vmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} \end{vmatrix} > 0 \quad \text{for every } i = 1, 2, \ldots, d.$$

(iii) *Every real eigenvalue of matrix A is positive.*

(iv) *The matrix A is semi-positive, i.e. there exists a vector $x \gg 0$ such that $Ax \gg 0$.*

Example 2.1. We consider a one-dimensional SDE driven by the rotationally invariant α-stable process ($\alpha \in (1, 2)$) with Markovian switching:

$$\mathrm{d}X_t = g(X_t, R_t)\,\mathrm{d}t + \mathrm{d}Z_t, \quad (X_0, R_0) \in \mathbb{R} \times \mathbb{S}, \tag{2.5}$$

where the state space of $(R_t)_{t\geqslant 0}$ is $\mathbb{S} = \{1, 2\}$ and the corresponding matrix Q is

$$Q = \begin{pmatrix} -1 & 1 \\ 3 & -3 \end{pmatrix}.$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

Let the coefficients be $g(x, 1) = x/20$ and $g(x, 2) = -x$. For the parameters $c_1 = 1/20$ and $c_2 = -1$ derived from (1.2), the matrix $-Q_{c,\kappa}$ in (1.3) takes the form

$$-Q_{c,\kappa_1} = \begin{pmatrix} 1 - (\kappa_1/20) & -1 \\ -3 & 3 + \kappa_1 \end{pmatrix},$$

which constitutes a nonsingular M-matrix for any $\kappa_1 \in (1, \alpha)$ by Lemma 2.1, verifying Assumption 1.2. Moreover, spectral analysis reveals that $Q_{c,4}$ has eigenvalues with strictly negative real parts, ensuring the existence of $\eta > 0$ as defined in (1.4).

The EM scheme is defined by $(Y_0, R_0) = (X_0, R_0)$ and

$$Y_{t_{n+1}} = Y_{t_n} + \gamma g(Y_{t_n}, R_{t_n}) + \sigma \Delta Z_\gamma, \tag{2.6}$$

while the embedded Markov chain $(R_{t_n})_{n\in\mathbb{N}_0}$ evolves according to the transition matrix $e^{\gamma Q}$.

Thus, all the conditions in Theorem 2.1 hold, so we can get the approximation between the process $(X_t, R_t)_{t\geqslant 0}$ in (2.5) and its EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$ in (2.6). For any $(x, k) \in \mathbb{R} \times \mathbb{S}$ and $n \in \mathbb{N}_0$, there exists a positive constant C, independent of γ, x, k, and t_n, such that

$$d_{\mathcal{G}}\big(\mathcal{L}\big(X^{x,k}_{t_n}, R^k_{t_n}\big), \mathcal{L}\big(Y^{x,k}_{t_n}, R^k_{t_n}\big)\big) \leqslant C(1 + |x|^{\kappa_1})\gamma^{\kappa_1/\alpha},$$

where the parameter $\kappa_1 \in (1, \alpha)$.

3. Proof of Theorem 2.1

We begin by establishing gradient estimates for the semigroup P_t associated with the process $(X_t, R_t)_{t\geqslant 0}$, whose proofs via Jacobi flow techniques are detailed in Appendix C.

Lemma 3.1. *Under Assumptions 1.1 and 1.2, with $\eta > 0$ defined in (1.4), for $h \in \mathcal{G}$ there exists some positive constant C, independent of x, k, and t, such that $\|\nabla^i P_t h(x, k)\|_\infty \leqslant Ce^{-\eta t/8}$ for $i = 1, 2$.*

We next establish moment estimates for both the process $(X_t, R_t)_{t\geqslant 0}$ and its EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$.

Lemma 3.2. *Under Assumptions 1.1 and 1.2, there exists some positive constant C, independent of n, γ, t, x, and k, such that, for $0 \leqslant p \leqslant \kappa_1 < \alpha$,*

(i) $\mathbb{E}|X^{x,k}_t|^p \leqslant C(1 + |x|^p)$ *for all* $t \geqslant 0$;

(ii) $\mathbb{E}|Y^{x,k}_{t_n}|^p \leqslant C(1 + |x|^p)$ *for all* $n \in \mathbb{N}_0$;

(iii) $\mathbb{E}|X^{x,k}_t - x|^p \leqslant C(1 + |x|^p)(t^{p/\alpha} \vee t^p)$ *for all* $t \geqslant 0$;

(iv) $\mathbb{E}|X^{x,k}_\gamma - Y^{x,k}_\gamma|^p \leqslant C(1 + |x|^p)\gamma^{p(1+(1/\alpha))}$;

(v) $\mathbb{E}|Y^{x,k}_\gamma - x|^p \leqslant C(1 + |x|^p)\gamma^{p/\alpha}$.

The one-step difference between the process $(X_t, R_t)_{t\geqslant 0}$ and its EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$ plays a crucial role in proving our main results; its proof can be found in Appendix B.

Lemma 3.3. *Under Assumptions 1.1 and 1.2, for $f \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{S}, \mathbb{R})$ satisfying $\|\nabla^j f\|_\infty \leqslant C$ with $j = 1, 2$, there exists some positive constant C. independent of γ, x, and k, such that*

$$|P_\gamma f(x, k) - \bar{P}_\gamma f(x, k)| \leqslant C(\|\nabla f\|_\infty + \|\nabla^2 f\|_\infty)(1 + |x|^{\kappa_1})\gamma^{1+(\kappa_1/\alpha)}.$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

Proof of Theorem 2.1. From the definition of $d_{\mathcal{G}}$ in (1.5) and the Lindeberg principle in [28], we know that

$$d_{\mathcal{G}}\big(\mathcal{L}\big(X_{t_n}^{x,k}, R_{t_n}^{k}\big), \mathcal{L}\big(Y_{t_n}^{x,k}, R_{t_n}^{k}\big)\big) = \sup_{h\in\mathcal{G}}\{P_{t_n}h(x,k) - \bar{P}_{t_n}h(x,k)\}$$
$$= \sup_{h\in\mathcal{G}}\left\{\sum_{i=1}^{n}\bar{P}_{t_{i-1}}\circ(P_\gamma - \bar{P}_\gamma)\circ P_{t_n-t_i}h(x,k)\right\}.$$

Let $f(x,k) = P_{t_n-t_i}h(x,k)$ with $\|\nabla h\|_\infty + \|\nabla^2 h\|_\infty \leqslant C$. It follows from Lemma 3.1 that $\|\nabla^j f\|_\infty \leqslant C$ with $j = 1, 2$; combining this with Lemma 3.3, we obtain

$$|(P_\gamma - \bar{P}_\gamma)\circ P_{t_n-t_i}h(x,k)| = |(P_\gamma - \bar{P}_\gamma)f(x,k)|$$
$$\leqslant C(\|\nabla f\|_\infty + \|\nabla^2 f\|_\infty)(1+|x|^{\kappa_1})\gamma^{1+(\kappa_1/\alpha)}.$$

Using Lemma 3.1 again implies that

$$|P_{t_n}h(x,k) - \bar{P}_{t_n}h(x,k)| \leqslant \sum_{i=1}^{n} C\mathrm{e}^{-\eta(t_n-t_i)/8}\mathbb{E}\big(1+\big|Y_{t_{i-1}}^{x,k}\big|^{\kappa_1}\big)\gamma^{1+(\kappa_1/\alpha)} \leqslant C(1+|x|^{\kappa_1})\gamma^{\kappa_1/\alpha},$$

where the estimate for $\mathbb{E}\big|Y_{t_{i-1}}^{x,k}\big|^{\kappa_1}$ is from Lemma 3.2, and the last inequality holds from the fact that

$$\sum_{i=1}^{n}\mathrm{e}^{-\eta(t_n-t_i)/8}\gamma^{1+(\kappa_1/\alpha)} = \sum_{i=1}^{n}\mathrm{e}^{-\eta(n-i)\gamma/8}\gamma^{1+(\kappa_1/\alpha)} \leqslant C\gamma^{\kappa_1/\alpha}.$$

The proof is complete. □

4. Proof of Theorems 2.2 and 2.3

We employ Skorokhod's representation for the continuous-time Markov chain $(R_t)_{t\geqslant 0}$ in terms of the Poisson random measure, as detailed in [32, pp. 103–104], [39, p. 29], or [31, p. 3219]. We construct a family of intervals $\{\Delta_{ij}\colon i,j\in\mathbb{S}\}$ on the half line in the following manner:

$$\Delta_{12} = [0, q_{12}), \quad \Delta_{13} = [q_{12}, q_{12}+q_{13}), \quad \ldots, \quad \Delta_{1N} = \big[\textstyle\sum_{j=2}^{N-1} q_{1j}, q_1\big),$$
$$\Delta_{21} = [q_1, q_1+q_{21}), \quad \Delta_{23} = [q_1+q_{21}, q_1+q_{21}+q_{23}), \quad \ldots,$$
$$\vdots$$

where $q_i = -q_{ii}$ for all $i\in\mathbb{S}$. Therefore, we obtain a sequence of consecutive, left-closed, right-open intervals $\{\Delta_{ij}, i,j\in\mathbb{S}\}$, each having length q_{ij}. For convenience, we set $\Delta_{ii} = \emptyset$ and $\Delta_{ij} = \emptyset$ if $q_{ij} = 0$. Define a function $\ell\colon \mathbb{S}\times\mathbb{R}\to\mathbb{R}$ by

$$\ell(i,z) = \sum_{j\in\mathbb{S}}(j-i)\mathbf{1}_{\{z\in\Delta_{ij}\}}. \tag{4.1}$$

Then the continuous-time Markov chain $(R_t)_{t\geqslant 0}$ can be expressed (see [31, (2.1)]) as

$$\mathrm{d}R_t = \int_{\mathbb{R}_0}\ell(R_{t-},z)\,N_1(\mathrm{d}t,\mathrm{d}z) = \int_{[0,M]}\ell(R_{t-},z)\,N_1(\mathrm{d}t,\mathrm{d}z),$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

where $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$, $N_1(\cdot, \cdot)$ is the Poisson process with $\mathbb{E}N_1(\mathrm{d}t, \mathrm{d}z) = \mathrm{d}z\,\mathrm{d}t$, and $M = N(N-1)H$ with $H = \max_{i\in\mathbb{S}}\{-q_{ii}\}$.

It follows from the Lévy–Itô decomposition [1, Theorem 2.4.16] that the continuous part $(X_t)_{t\geqslant 0}$ can be expressed as

$$\mathrm{d}X_t = g(X_t, R_t)\,\mathrm{d}t + \int_{0<|z|<1} \sigma z\, \widetilde{N}_2(\mathrm{d}t, \mathrm{d}z) + \int_{|z|\geqslant 1} \sigma z\, N_2(\mathrm{d}t, \mathrm{d}z),$$

where $N_2(\cdot, \cdot)$ is also a Poisson process with Lévy measure

$$\nu(\mathrm{d}z) = \frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z \tag{4.2}$$

and $d_\alpha = \alpha 2^{\alpha-1}\pi^{-d/2}\Gamma((d+\alpha)/2)/\,\Gamma(1-(\alpha/2))$, and $\widetilde{N}_2(\cdot, \cdot)$ is the corresponding compensated Poisson measure.

Denote by $\mathcal{A}$ the generator for the process $(X_t, R_t)_{t\geqslant 0}$ [8, 23, 31, 37, 39]. For a function $\psi \in \mathcal{D}(\mathcal{A})$, where $\mathcal{D}(\mathcal{A})$ is the domain of generator $\mathcal{A}$, and for any $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, we know that

$$\mathcal{A}\psi(x, k) = \langle \nabla\psi(x, k), g(x, k)\rangle + \mathcal{J}\psi(x, k) + Q\psi(x, k), \tag{4.3}$$

where the operator $\mathcal{J}$ is defined as

$$\mathcal{J}\psi(x, k) = \int_{\mathbb{R}_0^d} [\psi(x+\sigma z, k) - \psi(x, k) - \langle \nabla\psi(x, k), \sigma z\rangle \mathbf{1}_{\{|z|\leqslant 1\}}] \frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z, \tag{4.4}$$

with $\mathbb{R}_0^d = \mathbb{R}^d \setminus \{0\}$, and $Q\psi(x, k)$ given by

$$Q\psi(x, k) = \sum_{j\neq k} q_{kj}[\psi(x, j) - \psi(x, k)] = \sum_{j\in\mathbb{S}} q_{kj}\psi(x, j) = [Q\psi(x, \cdot)]_k,$$

where $[Q\psi(x, \cdot)]_k$ is the kth element of the vector $Q\psi(x, \cdot)$.

We now consider the following Stein equation: for any function $h \in \mathcal{C}_{\mathrm{b}}^1(\mathbb{R}^d \times \mathbb{S}, \mathbb{R})$,

$$\mathcal{A}f_h(x, k) = h(x, k) - \mu(h), \tag{4.5}$$

where $\mathcal{A}$ and μ are respectively the infinitesimal generator and invariant measure for the process $(X_t, R_t)_{t\geqslant 0}$. The following regularity results hold for the solution f_h.

Lemma 4.1. *Under Assumptions 1.1 and '1.2, with $\eta > 0$ defined in (1.4), let $h \in \mathcal{C}_{\mathrm{b}}^1(\mathbb{R}^d \times \mathbb{S}, \mathbb{R})$ and f_h be the solution to Stein's equation (4.5). For any $(x, k) \in \mathbb{R}^d \times \mathbb{S}$, there exists some positive constant C such that $|f_h(x, k)| \leqslant C(1 + |x|^{\kappa_2})$ and $\|\nabla f_h(x, k)\|_\infty \leqslant C$, where the constant κ_2 is specified in Assumption 1.2.*

4.1. Proof of Theorem 2.2

We prove Theorem 2.2 by Itô's formula and the martingale CLT. The verification of the martingale CLT conditions is highly nontrivial.

Proof of Theorem 2.2. For the process $(X_t, R_t)_{t\geqslant 0}$ in the system in (1.1) with $(X_0, R_0) = (x, k)$, by Itô's formula for f_h, which is the solution of Stein's equation (4.5), we have

$$\begin{aligned} f_h(X_t^{x,k}, R_t^k) - f_h(x, k) &= \int_0^t \mathcal{A}f_h(X_s^{x,k}, R_s^k)\,\mathrm{d}s + \mathcal{M}_{1,t} + \mathcal{M}_{2,t} \\ &= \int_0^t [h(X_s^{x,k}, R_s^k) - \mu(h)]\,\mathrm{d}s + \mathcal{M}_{1,t} + \mathcal{M}_{2,t}, \end{aligned}$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

where $\mathcal{M}_{1,t}$ and $\mathcal{M}_{2,t}$ are martingales:

$$\mathcal{M}_{1,t} = \int_0^t \int_{\mathbb{R}_0^d} \big[f_h\big(X_{s-}^{x,k} + \sigma z, R_{s-}^k\big) - f_h\big(X_{s-}^{x,k}, R_{s-}^k\big)\big]\, \widetilde{N}_2(\mathrm{d}s, \mathrm{d}z),$$

$$\mathcal{M}_{2,t} = \int_0^t \int_{\mathbb{R}_0} \big[f_h\big(X_{s-}^{x,k}, R_{s-}^k + \ell(R_{s-}^k, \tilde{z})\big) - f_h\big(X_{s-}^{x,k}, R_{s-}^k\big)\big]\, \widetilde{N}_1(\mathrm{d}s, \mathrm{d}\tilde{z}),$$

with $\widetilde{N}_1$ the compensated Poisson measure for N_1, and ℓ as defined in (4.1). This implies that

$$\begin{aligned}\sqrt{t}\bigg[\frac{1}{t}\int_0^t \delta_{(X_s^{x,k}, R_s^k)}(h)\,\mathrm{d}s - \mu(h)\bigg] &= \frac{1}{\sqrt{t}}\int_0^t \big[h\big(X_s^{x,k}, R_s^k\big) - \mu(h)\big]\,\mathrm{d}s \\ &= \frac{1}{\sqrt{t}}\big[f_h\big(X_t^{x,k}, R_t^k\big) - f_h(x,k)\big] - \frac{1}{\sqrt{t}}(\mathcal{M}_{1,t} + \mathcal{M}_{2,t}).\end{aligned}$$

From the estimates for $|f_h(x,k)|$ and $\mathbb{E}|X_t^{x,k}|^{\kappa_2}$ in Lemmas 4.1 and 3.2 respectively, we get

$$|\mathbb{E} f_h\big(X_t^{x,k}, R_t^k\big)| \leqslant \mathbb{E}(1 + |X_t^{x,k}|^{\kappa_2}) \leqslant C(1 + |x|^{\kappa_2}),$$

which implies that $\mathbb{E}|(1/\sqrt{t})[f_h(X_t^{x,k}, R_t^k) - f_h(x,k)]| \to 0$ as $t \to \infty$. It follows from the estimates for $|f_h(x,k)|$ and $\|\nabla f_h\|_\infty$ in Lemma 4.1 that

$$\begin{aligned}&\int_{\mathbb{R}_0^d} |f_h(x+\sigma z) - f_h(x)|^2\, \nu(\mathrm{d}z) \\ &\quad= \int_{0<|z|\leqslant 1} \bigg| \int_0^1 \langle \nabla f_h(x + r\sigma z), \sigma z\rangle\, \mathrm{d}r\bigg|^2 \nu(\mathrm{d}z) + \int_{|z|>1} |f_h(x+\sigma z) - f_h(x)|^2\, \nu(\mathrm{d}z) \\ &\quad\leqslant C\int_{0<|z|\leqslant 1} |z|^2\, \nu(\mathrm{d}z) + C\int_{|z|>1} (1 + |x|^{2\kappa_2} + |z|^{2\kappa_2})\, \nu(\mathrm{d}z) \\ &\quad\leqslant C(1 + |x|^{2\kappa_2}),\end{aligned}$$

where the last two inequalities hold since $2\kappa_2 < \kappa_1 < \alpha$ and $\int_{\mathbb{R}_0^d} (1 \wedge |z|^{\kappa_1})\, \nu(\mathrm{d}z) < \infty$. Combining this with

$$\int_{[0,M]} |f_h(x, k + \ell(k,\tilde{z})) - f_h(x,k)|^2\, \mathrm{d}\tilde{z} \leqslant C(1 + |x|^{2\kappa_2})$$

implies that

$$\begin{aligned}\mathcal{V}(f_h) :=& \int_{\mathbb{R}^d\times\mathbb{S}} \int_{\mathbb{R}_0^d} |f_h(x+\sigma z, k) - f_h(x,k)|^2\, \nu(\mathrm{d}z)\, \mathrm{d}\mu(x,k) \\ &+ \int_{\mathbb{R}^d\times\mathbb{S}} \int_{\mathbb{R}_0} |f_h(x, k + \ell(k,\tilde{z})) - f_h(x,k)|^2\, \mathrm{d}\tilde{z}\, \mathrm{d}\mu(x,k) \\ &\leqslant C\mu(1 + |x|^{2\kappa_2}) < \infty, \end{aligned} \tag{4.6}$$

where the inequality holds since $2\kappa_2 < \kappa_1$ and $\mu(|x|^{\kappa_1}) < \infty$ in Proposition A.1.

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

To make the notation simpler, for $i = 1, 2, \ldots, \lfloor t \rfloor$ we write

$$\begin{aligned} U_i &= \int_{i-1}^{i} \int_{\mathbb{R}_0^d} \Big[f_h\Big(X_{s-}^{x,k} + \sigma z, R_{s-}^k\Big) - f_h\Big(X_{s-}^{x,k}, R_{s-}^k\Big) \Big] \widetilde{N}_2(\mathrm{d}s, \mathrm{d}z) \\ &\quad + \int_{i-1}^{i} \int_{\mathbb{R}_0} \Big[f_h\Big(X_{s-}^{x,k}, R_{s-}^k + \ell(R_{s-}^k, \tilde{z})\Big) - f_h\Big(X_{s-}^{x,k}, R_{s-}^k\Big) \Big] \widetilde{N}_1(\mathrm{d}s, \mathrm{d}\tilde{z}), \\ U_{\lfloor t \rfloor + 1} &= \int_{\lfloor t \rfloor}^{t} \int_{\mathbb{R}_0^d} \Big[f_h\Big(X_{s-}^{x,k} + \sigma z, R_{s-}^k\Big) - f_h\Big(X_{s-}^{x,k}, R_{s-}^k\Big) \Big] \widetilde{N}_2(\mathrm{d}s, \mathrm{d}z) \\ &\quad + \int_{\lfloor t \rfloor}^{t} \int_{\mathbb{R}_0} \Big[f_h\Big(X_{s-}^{x,k}, R_{s-}^k + \ell(R_{s-}^k, \tilde{z})\Big) - f_h\Big(X_{s-}^{x,k}, R_{s-}^k\Big) \Big] \widetilde{N}_1(\mathrm{d}s, \mathrm{d}\tilde{z}). \end{aligned}$$

Then,

$$\frac{1}{\sqrt{t}}(\mathcal{M}_{1,t} + \mathcal{M}_{2,t}) = \frac{1}{\sqrt{t}} \sum_{i=1}^{\lfloor t \rfloor + 1} U_i.$$

We know that the U_i are martingale differences and $\mathbb{E}U_i^2 < \infty$, which can be obtained similarly with (4.6) for all $i = 1, 2, \ldots, \lfloor t \rfloor + 1$. We claim that

$$\lim_{t \to \infty} \mathbb{E}\Big(\max_{1 \leqslant i \leqslant \lfloor t \rfloor + 1} \frac{1}{\mathcal{V}(f_h)t} |U_i|^2 \Big) = 0, \tag{4.7}$$

$$\lim_{t \to \infty} \mathbb{E} \left| \sum_{i=1}^{\lfloor t \rfloor + 1} \frac{1}{\mathcal{V}(f_h)t} |U_i|^2 - 1 \right|^2 = 0. \tag{4.8}$$

We know that (4.7) and (4.8) imply that

$$\mathbb{E}\Big(\max_{1 \leqslant i \leqslant \lfloor t \rfloor + 1} \frac{1}{\sqrt{\mathcal{V}(f_h)t}} |U_i| \Big) \to 0 \quad \text{and} \quad \sum_{i=1}^{\lfloor t \rfloor + 1} \frac{1}{\mathcal{V}(f_h)t} |U_i|^2 \xrightarrow{\mathrm{p}} 1$$

as t tends to infinity. Applying the martingale CLT in [29, Theorem 2], due to [24], we have

$$\frac{1}{\sqrt{\mathcal{V}(f_h)t}} \sum_{i=1}^{\lfloor t \rfloor + 1} U_i \Rightarrow \mathcal{N}(0, 1) \quad \text{as } t \to \infty.$$

Then, we know that

$$\sqrt{t}\left[\frac{1}{t} \int_0^t \delta_{(X_s^{x,k}, R_s^k)}(h)\, \mathrm{d}s - \mu(h) \right] \Rightarrow \mathcal{N}(0, \mathcal{V}(f_h)) \quad \text{as } t \to \infty.$$

It remains to show (4.7) and (4.8). For (4.7), we have

$$\begin{aligned} \mathbb{E}\Big(\max_{1 \leqslant i \leqslant \lfloor t \rfloor + 1} |U_i|^2 \Big) &= \mathbb{E}\Big(\max_{1 \leqslant i \leqslant \lfloor t \rfloor + 1} (|U_i|^2 \mathbf{1}_{\{|U_i|^2 \leqslant \sqrt{t}\}} + |U_i|^2 \mathbf{1}_{\{|U_i|^2 > \sqrt{t}\}}) \Big) \\ &\leqslant \mathbb{E}\Big(\max_{1 \leqslant i \leqslant \lfloor t \rfloor + 1} |U_i|^2 \mathbf{1}_{\{|U_i|^2 \leqslant \sqrt{t}\}} \Big) + \mathbb{E}\Big(\max_{1 \leqslant i \leqslant \lfloor t \rfloor + 1} |U_i|^2 \mathbf{1}_{\{|U_i|^2 > \sqrt{t}\}} \Big) \\ &\leqslant \sqrt{t} + (\lfloor t \rfloor + 1) \max_{1 \leqslant i \leqslant \lfloor t \rfloor + 1} \mathbb{E}\big[|U_i|^2 \mathbf{1}_{\{|U_i|^2 > \sqrt{t}\}} \big]. \end{aligned}$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

Thus, as $t \to \infty$, we obtain

$$\mathbb{E}\Big(\max_{1\leqslant i\leqslant \lfloor t\rfloor+1} \frac{1}{\mathcal{V}(f_h)t}|U_i|^2\Big) \leqslant \frac{1}{\mathcal{V}(f_h)t}\Big(\sqrt{t}+(\lfloor t\rfloor+1)\max_{1\leqslant i\leqslant \lfloor t\rfloor+1}\mathbb{E}[|U_i|^2\mathbf{1}_{\{|U_i|^2>\sqrt{t}\}}]\Big)\to 0.$$

It is easy to verify (4.8) by proving the equation

$$\lim_{t\to\infty}\mathbb{E}\left|\sum_{i=1}^{\lfloor t\rfloor+1}\frac{1}{\mathcal{V}(f_h)t}|U_i|^2-\frac{\lfloor t\rfloor+1}{t}\right|^2=0. \tag{4.9}$$

Observe that

$$\begin{aligned}
&\mathbb{E}\left|\sum_{i=1}^{\lfloor t\rfloor+1}\frac{1}{\mathcal{V}(f_h)t}|U_i|^2-\frac{\lfloor t\rfloor+1}{t}\right|^2\\
&=\mathbb{E}\left|\frac{1}{t}\sum_{i=1}^{\lfloor t\rfloor+1}\left(\frac{1}{\mathcal{V}(f_h)}|U_i|^2-1\right)\right|^2\\
&=\underbrace{\frac{1}{t^2}\sum_{i=1}^{\lfloor t\rfloor+1}\mathbb{E}\left(\frac{1}{\mathcal{V}(f_h)}|U_i|^2-1\right)^2}_{\mathrm{I}}+\underbrace{\frac{2}{t^2}\sum_{i<j}\mathbb{E}\left[\left(\frac{1}{\mathcal{V}(f_h)}|U_i|^2-1\right)\left(\frac{1}{\mathcal{V}(f_h)}|U_j|^2-1\right)\right]}_{\mathrm{II}}.
\end{aligned}$$

Using Kunita's inequality (see [1, Theorem 4.4.23]), with similar calculations for (4.6), there exists some constant $C>0$ such that

$$\begin{aligned}
\mathbb{E}|U_i|^4 &\leqslant C\mathbb{E}\left|\int_{i-1}^{i}\int_{\mathbb{R}_0^d}\left|f_h\big(X_{s-}^{x,k}+\sigma z,R_{s-}^k\big)-f_h\big(X_{s-}^{x,k},R_{s-}^k\big)\right|^2\nu(\mathrm{d}z)\,\mathrm{d}s\right|^2\\
&\quad+C\mathbb{E}\int_{i-1}^{i}\int_{\mathbb{R}_0^d}\left|f_h\big(X_{s-}^{x,k}+\sigma z,R_{s-}^k\big)-f_h\big(X_{s-}^{x,k},R_{s-}^k\big)\right|^4\nu(\mathrm{d}z)\,\mathrm{d}s\\
&\quad+C\mathbb{E}\left|\int_{i-1}^{i}\int_{[0,M]}\left|f_h\big(X_{s-}^{x,k},R_{s-}^k+\ell\big(R_{s-}^k,\tilde{z}\big)\big)-f_h\big(X_{s-}^{x,k},R_{s-}^k\big)\right|^2\mathrm{d}\tilde{z}\,\mathrm{d}s\right|^2\\
&\quad+C\mathbb{E}\int_{i-1}^{i}\int_{[0,M]}\left|f_h\big(X_{s-}^{x,k},R_{s-}^k+\ell\big(R_{s-}^k,\tilde{z}\big)\big)-f_h\big(X_{s-}^{x,k},R_{s-}^k\big)\right|^4\mathrm{d}\tilde{z}\,\mathrm{d}s\\
&\leqslant C\mathbb{E}\big(1+|X_{i-1}^{x,k}|^{4\kappa_2}\big)\leqslant C(1+|x|^{4\kappa_2}),
\end{aligned}$$

where the second inequality holds from the fact $4\kappa_2\leqslant\kappa_1$ and the last inequality holds from the estimate $\mathbb{E}|X_{i-1}^{x,k}|^{4\kappa_2}$ in Lemma 3.2. Combining this with Lemma 3.2 and the fact that $4\kappa_2\leqslant\kappa_1<\alpha$, we know that

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{\mathcal{V}(f_h)}|U_i|^2-1\right)^2&=\mathbb{E}\left[\frac{1}{\mathcal{V}^2(f_h)}|U_i|^4-\frac{2}{\mathcal{V}(f_h)}|U_i|^2+1\right]\\
&\leqslant C\big(1+\mathbb{E}|X_i^{x,k}|^{4\kappa_2}\big)\leqslant C\big(1+|x|^{4\kappa_2}\big),
\end{aligned}$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

which implies that

$$\mathrm{I} = \frac{1}{t^2}\sum_{i=1}^{\lfloor t\rfloor+1}\mathbb{E}\bigg(\frac{1}{\mathcal{V}(f_h)}|U_i|^2-1\bigg)^2 \to 0 \quad \text{as } t\to\infty.$$

As for II, for each i, it holds that, for some positive constant C,

$$\begin{aligned}\bigg|\sum_{j=i+1}^{\lfloor t\rfloor+1}\mathbb{E}\bigg[\bigg(\frac{1}{\mathcal{V}(f_h)}|U_j|^2-1\bigg)\mid\mathcal{F}_i\bigg]\bigg| &= \frac{1}{\mathcal{V}(f_h)}\bigg|\sum_{j=i+1}^{\lfloor t\rfloor+1}\mathbb{E}\big[(|U_j|^2-\mathcal{V}(f_h))\mid\mathcal{F}_i\big]\bigg|\\ &\leqslant \int_0^{\lfloor t\rfloor+1}\Big(1+|X_i^{x,k}|^{2\kappa_2}\Big)\mathrm{e}^{-cs}\,\mathrm{d}s \leqslant C\Big(1+|X_i^{x,k}|^{2\kappa_2}\Big),\end{aligned}$$

where $\mathcal{F}_i$ is the filtration generated by the process $(X_t, R_t)_{t\geqslant 0}$ at time $t=i$, and the first inequality is by Proposition A.1 and the fact that $2\kappa_2<\kappa_1$. Hence, we can obtain that

$$\begin{aligned}|\mathrm{II}| &= \frac{2}{t^2}\bigg|\sum_{i=1}^{\lfloor t\rfloor}\sum_{j=i+1}^{\lfloor t\rfloor+1}\mathbb{E}\bigg\{\bigg(\frac{1}{\mathcal{V}(f_h)}|U_i|^2-1\bigg)\mathbb{E}\bigg[\bigg(\frac{1}{\mathcal{V}(f_h)}|U_j|^2-1\bigg)\mid\mathcal{F}_i\bigg]\bigg\}\bigg|\\ &\leqslant \frac{C}{t^2}\sum_{i=1}^{\lfloor t\rfloor}\mathbb{E}\bigg\{\bigg|\frac{1}{\mathcal{V}(f_h)}|U_i|^2-1\bigg|\Big(1+|X_i^{x,k}|^{2\kappa_2}\Big)\bigg\}\\ &\leqslant \frac{C}{t^2}\sum_{i=1}^{\lfloor t\rfloor}\bigg[\mathbb{E}\bigg|\frac{1}{\mathcal{V}(f_h)}|U_i|^2-1\bigg|^2\bigg]^{1/2}\big[1+\mathbb{E}|X_i^{x,k}|^{4\kappa_2}\big]^{1/2}\to 0 \quad \text{as } t\to\infty.\end{aligned}$$

Combining the estimates for I and II, (4.9) holds. The proof is complete. □

4.2. Proof of Theorem 2.3

We shall prove the MDP by the criterion in [34, Theorem 2.6].

Proof of Theorem 2.3. According to the criterion for MDP in [34, Theorem 2.6], we need to show that the semigroup P_t associated with the process $(X_t, R_t)_{t\geqslant 0}$ is strong Feller, topologically irreducible, and satisfies a Lyapunov condition. The process $(X_t, R_t)_{t\geqslant 0}$ defined in (1.1) is strong Feller and irreducible, which can be obtained from the discussions in [35, Theorems 3.2 and 4.5], [39, Theorem 2.24], and [37, Theorems 5.6 and 6.3], and the Lyapunov condition is implied by (A.4).

For $h\in\mathcal{C}_{\mathrm{b}}^1(\mathbb{R}^d\times\mathbb{S},\mathbb{R})$, it follows from [34, Theorem 2.6] that

$$\mathbb{P}\bigg(\frac{1}{a_t\sqrt{t}}\int_0^t[h(X_s^{x,k},R_s^k)-\mu(h)]\,\mathrm{d}s\in\cdot\bigg)$$

satisfies the large deviation principle with speed a_t^{-2} and rate function $I_h(z)=z^2/2\mathcal{H}(h)$, with $\mathcal{H}(h)=2\int_0^\infty\langle P_t h, h-\mu(h)\rangle_\mu\,\mathrm{d}t$, that is,

$$\begin{aligned}-\inf_{z\in\mathring{A}}I_h(z) &\leqslant \liminf_{t\to\infty}\frac{1}{a_t^2}\log\mathbb{P}\bigg(\frac{\sqrt{t}}{a_t}\big[\mathcal{E}_t^{x,k}(h)-\mu(h)\big]\in A\bigg)\\ &\leqslant \limsup_{t\to\infty}\frac{1}{a_t^2}\log\mathbb{P}\bigg(\frac{\sqrt{t}}{a_t}\big[\mathcal{E}_t^{x,k}(h)-\mu(h)\big]\in A\bigg)\leqslant -\inf_{z\in\bar{A}}I_h(z),\end{aligned}$$

where $\bar{A}$ and $\mathring{A}$ are the closure and interior of set A respectively.

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

We claim that

$$\mathcal{H}(h) = \mathcal{V}(f_h), \tag{4.10}$$

where f_h is the solution to Stein's equation (4.5). Then the desired result holds.

Now, we show that the claim (4.10) holds. With some calculations, we have

$$\begin{aligned}
&\frac{1}{t}\mathbb{E}^{\mu}\bigg(\int_0^t \Big[h\big(X_s^{x,k}, R_s^k\big) - \mu(h)\Big]\,\mathrm{d}s\bigg)^2 \\
&\quad = \frac{2}{t}\mathbb{E}^{\mu}\bigg(\int_0^t \int_0^u \Big[h\big(X_s^{x,k}, R_s^k\big) - \mu(h)\Big]\Big[h\big(X_u^{x,k}, R_u^k\big) - \mu(h)\Big]\,\mathrm{d}s\,\mathrm{d}u\bigg) \\
&\quad = \frac{2}{t}\int_0^t \int_0^u \mathbb{E}^{\mu}\Big\{h\big(X_s^{x,k}, R_s^k\big)\mathbb{E}\Big\{\Big[h\big(X_u^{x,k}, R_u^k\big) - \mu(h)\Big] \mid \mathcal{F}_s\Big\}\Big\}\,\mathrm{d}s\,\mathrm{d}u,
\end{aligned}$$

where $\mathcal{F}_s$ is the filtration generated by the process $(X_t, R_t)_{t\geqslant 0}$ at time $t = s$; the second equality holds from conditional probability and the fact that

$$\begin{aligned}
&\int_0^t \int_0^u \mathbb{E}^{\mu}\Big\{\mu(h)\mathbb{E}\Big\{\Big[h\big(X_u^{x,k}, R_u^k\big) - \mu(h)\Big] \mid \mathcal{F}_s\Big\}\Big\}\,\mathrm{d}s\,\mathrm{d}u \\
&\qquad = \mu(h)\int_0^t \int_0^u \mathbb{E}^{\mu}\Big\{h\big(X_u^{x,k}, R_u^k\big) - \mu(h)\Big\}\,\mathrm{d}s\,\mathrm{d}u = 0.
\end{aligned}$$

Furthermore, for all $0 \leqslant s \leqslant u < \infty$, we have

$$\begin{aligned}
&\mathbb{E}^{\mu}\Big\{h\big(X_s^{x,k}, R_s^k\big)\mathbb{E}\Big\{\Big[h\big(X_u^{x,k}, R_u^k\big) - \mu(h)\Big] \mid \mathcal{F}_s\Big\}\Big\} \\
&\quad = \mathbb{E}^{\mu}\Big\{h\big(X_s^{x,k}, R_s^k\big)\Big[P_{u-s}h\big(X_s^{x,k}, R_s^k\big) - \mu(h)\Big]\Big\} \\
&\quad = \int_{\mathbb{R}^d\times\mathbb{S}} h(y,k)\Big[P_{u-s}h(y,k) - \mu(h)\Big]\,\mathrm{d}\mu(y,k),
\end{aligned}$$

which implies that

$$\begin{aligned}
&\frac{2}{t}\int_0^t \int_0^u \mathbb{E}^{\mu}\Big\{h\big(X_s^{x,k}, R_s^k\big)\mathbb{E}\Big\{\Big[h\big(X_u^{x,k}, R_u^k\big) - \mu(h)\Big] \mid \mathcal{F}_s\Big\}\Big\}\,\mathrm{d}s\,\mathrm{d}u \\
&\quad = \int_{\mathbb{R}^d\times\mathbb{S}} h(y,k)\frac{2}{t}\int_0^t \int_0^u \Big[P_{u-s}h(y,k) - \mu(h)\Big]\,\mathrm{d}s\,\mathrm{d}u\,\mathrm{d}\mu(y,k) \\
&\quad = \int_{\mathbb{R}^d\times\mathbb{S}} h(y,k)\frac{2}{t}\int_0^t \int_0^u \Big[P_{\tilde{s}}h(y,k) - \mu(h)\Big]\,\mathrm{d}\tilde{s}\,\mathrm{d}u\,\mathrm{d}\mu(y,k).
\end{aligned}$$

Using L'Hôpital's rule, we have

$$\begin{aligned}
&\lim_{t\to\infty}\int_{\mathbb{R}^d\times\mathbb{S}} h(y,k)\frac{2}{t}\int_0^t \int_0^u [P_{\tilde{s}}h(y,k) - \mu(h)]\,\mathrm{d}\tilde{s}\,\mathrm{d}u\,\mathrm{d}\mu(y,k) \\
&\qquad = 2\int_{\mathbb{R}^d\times\mathbb{S}} h(y,k)\int_0^{\infty} [P_{\tilde{s}}h(y,k) - \mu(h)]\,\mathrm{d}\tilde{s}\,\mathrm{d}\mu(y,k),
\end{aligned}$$

that is,

$$\frac{1}{t}\mathbb{E}^{\mu}\bigg(\int_0^t \Big[h\big(X_s^{x,k}, R_s^k\big) - \mu(h)\Big]\,\mathrm{d}s\bigg)^2 \to \mathcal{H}(h) \quad \text{as } t\to\infty. \tag{4.11}$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

Indeed, $\mathcal{H}(h)$ can also be expressed as

$$\mathcal{H}(h) = 2\int_0^\infty \langle P_t h,\, h-\mu(h)\rangle_\mu \,\mathrm{d}t = 2\int_0^\infty \langle P_t h-\mu(h),\, h\rangle_\mu \,\mathrm{d}t.$$

Using Itô's formula and Stein's equation (4.5), we have

$$\begin{aligned} f_h(X_t^{x,k}, R_t^k) - f_h(x,k) &= \int_0^t \mathcal{A} f_h\big(X_s^{x,k}, R_s^k\big)\,\mathrm{d}s + \mathcal{M}_{1,t} + \mathcal{M}_{2,t} \\ &= \int_0^t \big[h\big(X_s^{x,k}, R_s^k\big) - \mu(h)\big]\,\mathrm{d}s + \mathcal{M}_{1,t} + \mathcal{M}_{2,t}. \end{aligned}$$

Combining this with the ergodic theorem and the fact that $\mathbb{E}\big|[f_h(X_t^{x,k}, R_t^k) - f_h(x,k)]/\sqrt{t}\big| \to 0$ as $t\to\infty$ yields

$$\begin{aligned} &\frac{1}{t}\mathbb{E}^\mu\bigg(\int_0^t \big[h\big(X_s^{x,k}, R_s^k\big) - \mu(h)\big]\,\mathrm{d}s\bigg)^2 \\ &\quad = \mathbb{E}^\mu\bigg(\frac{1}{\sqrt{t}}\big[f_h\big(X_t^{x,k}, R_t^k\big) - f_h(x,k) - \mathcal{M}_{1,t} - \mathcal{M}_{2,t}\big]\bigg)^2 \to \mathcal{V}(f_h) \quad \text{as } t\to\infty. \end{aligned} \tag{4.12}$$

Thus, we know that $\mathcal{H}(h) = \mathcal{V}(f_h)$ by (4.11) and (4.12), so the claim (4.10) holds. The proof is complete. □

5. Proof of Theorem 2.4

In order to determine the variance in Theorem 2.4, we need to study the second Stein equation

$$\mathcal{A}_\gamma f_{\gamma,h}(x,k) = h(x,k) - \mu_\gamma(h), \tag{5.1}$$

where $\mathcal{A}_\gamma$ is defined as

$$\mathcal{A}_\gamma f_{\gamma,h}(x,k) = \bar{P}_\gamma f_{\gamma,h}(x,k) - f_{\gamma,h}(x,k). \tag{5.2}$$

It is easy to verify that the solution of the equation (5.1) is

$$f_{\gamma,h}(x,k) = -\sum_{i=0}^\infty \bar{P}_\gamma^i (h(x,k) - \mu_\gamma(h)), \tag{5.3}$$

because $\bar{P}_\gamma f_{\gamma,h}(x,k) = -\sum_{i=1}^\infty \bar{P}_\gamma^i (h(x,k) - \mu_\gamma(h))$.

Proof of Theorem 2.4. (i) From Proposition A.2, we know that the EM scheme $(Y_{t_n}^{x,k}, R_{t_n}^k)_{n\in\mathbb{N}_0}$ is exponentially ergodic under the weighted total variation distance $\|\cdot\|_{1+V^{(1)}}$ and $\mu_\gamma(V^{(1)}) \leqslant C$ with $V^{(1)}$ as in (A.1). It follows from [15, Theorem 9] that, for any function h satisfying $|h| \leqslant V^{(1)}$ and any initial distribution (Y_0, R_0), $\sqrt{n}\big[\mathcal{E}_{n,\gamma}^{x,k}(h) - \mu_\gamma(h)\big] \Rightarrow \mathcal{N}(0, \sigma_h^2)$ with

$$\sigma_h^2 = \langle h-\mu_\gamma(h),\, h-\mu_\gamma(h)\rangle_{\mu_\gamma} + 2\sum_{i=1}^\infty \langle \bar{P}_\gamma^i h,\, h-\mu_\gamma(h)\rangle_{\mu_\gamma}. \tag{5.4}$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

We give some calculations for the variance of $\sqrt{n}\big[\mathcal{E}^{x,k}_{n,\gamma}(h)-\mu_\gamma(h)\big]$. Since $\mathbb{E}\big[\sqrt{n}\big(\mathcal{E}^{x,k}_{n,\gamma}(h)-\mu_\gamma(h)\big)\big]=0$, we have

$$\begin{aligned}
&\mathbb{E}\big|\sqrt{n}\big[\mathcal{E}^{x,k}_{n,\gamma}(h)-\mu_\gamma(h)\big]\big|^2\\
&\quad=\mathbb{E}\left|\frac{1}{\sqrt{n}}\sum_{i=0}^{n-1}\big[h\big(Y^{x,k}_{t_i},R^k_{t_i}\big)-\mu_\gamma(h)\big]\right|^2\\
&\quad=\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\big|h\big(Y^{x,k}_{t_i},R^k_{t_i}\big)-\mu_\gamma(h)\big|^2\\
&\qquad+\frac{2}{n}\sum_{i=1}^{n-1}\sum_{j=0}^{i-1}\mathbb{E}\big\{\big[h\big(Y^{x,k}_{t_i},R^k_{t_i}\big)-\mu_\gamma(h)\big]\big[h\big(Y^{x,k}_{t_j},R^k_{t_j}\big)-\mu_\gamma(h)\big]\big\}. \qquad (5.5)
\end{aligned}$$

We aim to deal with the second term on the right-hand side of (5.5). By the conditional expectation, we have

$$\begin{aligned}
&\frac{2}{n}\sum_{i=1}^{n-1}\sum_{j=0}^{i-1}\mathbb{E}\big\{\big[h\big(Y^{x,k}_{t_i},R^k_{t_i}\big)-\mu_\gamma(h)\big]\big[h\big(Y^{x,k}_{t_j},R^k_{t_j}\big)-\mu_\gamma(h)\big]\big\}\\
&\quad=\frac{2}{n}\sum_{i=1}^{n-1}\sum_{j=0}^{i-1}\mathbb{E}\big\{\big[h\big(Y^{x,k}_{t_j},R^k_{t_j}\big)-\mu_\gamma(h)\big]\mathbb{E}\big[\big(h\big(Y^{x,k}_{t_i},R^k_{t_i}\big)-\mu_\gamma(h)\big)\mid\bar{\mathcal{F}}_{t_j}\big]\big\}\\
&\quad=\frac{2}{n}\sum_{j=0}^{n-2}\mathbb{E}\left\{\big[h\big(Y^{x,k}_{t_j},R^k_{t_j}\big)-\mu_\gamma(h)\big]\mathbb{E}\left[\sum_{i=j+1}^{n-1}\big(h\big(Y^{x,k}_{t_i},R^k_{t_i}\big)-\mu_\gamma(h)\big)\mid\bar{\mathcal{F}}_{t_j}\right]\right\}\\
&\quad=\frac{2}{n}\sum_{j=0}^{n-2}\mathbb{E}\left\{\big[h\big(Y^{x,k}_{t_j},R^k_{t_j}\big)-\mu_\gamma(h)\big]\left[\sum_{i=1}^{n-1-j}\bar{P}^i_\gamma\big(h\big(Y^{x,k}_{t_j},R^k_{t_j}\big)-\mu_\gamma(h)\big)\mid\bar{\mathcal{F}}_{t_j}\right]\right\}\\
&\quad\to 2\sum_{i=1}^{\infty}\langle\bar{P}^i_\gamma[h(x,k)-\mu_\gamma(h)],h(x,k)-\mu_\gamma(h)\rangle_{\mu_\gamma}\quad\text{as } n\to\infty, \qquad (5.6)
\end{aligned}$$

where $\bar{\mathcal{F}}_s$ is the filtration generated by the EM scheme $(Y_{t_n},R_{t_n})_{n\in\mathbb{N}_0}$ at time s.

Combining (5.5) and (5.6), as $n\to\infty$ we know that

$$\mathbb{E}\big|\sqrt{n}\big[\mathcal{E}^{x,k}_{n,\gamma}(h)-\mu_\gamma(h)\big]\big|^2\to\langle h-\mu_\gamma(h),h-\mu_\gamma(h)\rangle_{\mu_\gamma}+2\sum_{i=1}^{\infty}\langle\bar{P}^i_\gamma h,h-\mu_\gamma(h)\rangle_{\mu_\gamma},$$

which implies that (5.4) holds.

Since $\langle\mu_\gamma(h),h-\mu_\gamma(h)\rangle_{\mu_\gamma}=0$, using (5.3) it is easy to verify that

$$\begin{aligned}
\widetilde{\mathcal{V}}(f_{\gamma,h})&=\langle f_{\gamma,h},f_{\gamma,h}\rangle_{\mu_\gamma}-\langle\bar{P}_\gamma f_{\gamma,h},\bar{P}_\gamma f_{\gamma,h}\rangle_{\mu_\gamma}\\
&=\langle h-\mu_\gamma(h),h-\mu_\gamma(h)\rangle_{\mu_\gamma}+2\sum_{i=1}^{\infty}\big\langle\bar{P}^i_\gamma h,h-\mu_\gamma(h)\big\rangle_{\mu_\gamma}=\sigma_h^2.
\end{aligned}$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

(ii) It follows from Proposition A.2 that the EM scheme $\big(Y^{x,k}_{t_n}, R^k_{t_n}\big)_{n\in\mathbb{N}_0}$ has exponential ergodicity. It follows from [34, Remark 2.17] that $\bar{P}^n_\gamma$ has a spectral gap near its largest eigenvalue 1, which implies that 1 is an isolated eigenvalue. Since $\bar{P}^i_\gamma c = c$ for all $i\in\mathbb{N}_0$, 1 is an eigenvalue of $\bar{P}^i_\gamma$ for all $i\in\mathbb{N}_0$. If there exists some function $\hat{f}$ satisfying $0\neq\hat{f}(x,k)\leqslant 1+V^{(1)}(x,k)$ for all $x\in\mathbb{R}^d$ such that $\bar{P}^{k_1}_\gamma\hat{f}=\lambda\hat{f}$ for some $k_1\in\mathbb{N}_0$ and $\lambda\in\mathbb{C}$ with $|\lambda|=1$, then $\lambda^{\tilde{n}}\hat{f}=\bar{P}^{\tilde{n}k_1}_\gamma\hat{f}\to\mu_\gamma(\hat{f})$ as $\tilde{n}\to\infty$, so that $\hat{f}$ has to be constant and $\lambda=1$. Thus, 1 is a simple eigenvalue and the only eigenvalue with modulus 1 for $\bar{P}^n_\gamma$.

For $h\in\mathcal{C}^1_{\mathrm{b}}(\mathbb{R}^d\times\mathbb{S},\mathbb{R})$, it follows from [33, Theorem 2.1] that

$$\mathbb{P}\bigg(\frac{\sqrt{n}}{a_n}\big[\mathcal{E}^{x,k}_{n,\gamma}(h)-\mu_\gamma(h)\big]\in\cdot\bigg)$$

satisfies the large deviation principle with speed a_n^{-2} and rate function $I_h(z)=z^2/2\widetilde{\mathcal{V}}(f_{\gamma,h})$, with $\widetilde{\mathcal{V}}(f_{\gamma,h})$ as in (2.4), that is,

$$\begin{aligned}-\inf_{z\in\mathring{A}}\frac{z^2}{2\widetilde{\mathcal{V}}(f_{\gamma,h})}&\leqslant\liminf_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\bigg(\frac{\sqrt{n}}{a_n}\big[\mathcal{E}^{x,k}_{n,\gamma}(h)-\mu_\gamma(h)\big]\in A\bigg)\\&\leqslant\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\bigg(\frac{\sqrt{n}}{a_n}\big[\mathcal{E}^{x,k}_{n,\gamma}(h)-\mu_\gamma(h)\big]\in A\bigg)\leqslant-\inf_{z\in\bar{A}}\frac{z^2}{2\widetilde{\mathcal{V}}(f_{\gamma,h})},\end{aligned}$$

where $\bar{A}$ and $\mathring{A}$ are the closure and interior of set A respectively. The proof is complete. □

Appendix A. Exponential ergodicity and moment estimates

In this section we first show that the process $(X_t,R_t)_{t\geqslant0}$ and its EM scheme $(Y_{t_n},R_{t_n})_{n\in\mathbb{N}_0}$ both have exponential ergodicities under the weighted total variation distance, and then give the moment estimates.

A.1. Exponential ergodicity

Proposition A.1. *Under Assumptions 1.1 and 1.2, the process $(X_t,R_t)_{t\geqslant0}$ is exponentially ergodic with ergodic measure μ. More precisely, for any $(x,k)\in\mathbb{R}^d\times\mathbb{S}$, there exist some positive constants c and C, independent of x, k, and t, satisfying*

$$\sup_{\psi\leqslant1+|\cdot|^{\kappa_j}}\{P_t\psi(x,k)-\mu(\psi)\}\leqslant C(1+|x|^{\kappa_j})\mathrm{e}^{-ct},$$

where the κ_j with $j\in\{1,2\}$ are as in Assumption 1.2. In addition, $\mu(|x|^{\kappa_j})\leqslant C$.

To get the exponential ergodicity for the EM scheme $(Y_{t_n},R_{t_n})_{n\in\mathbb{N}_0}$, we need the following relationship between $\mathcal{A}$ and $\mathcal{A}_\gamma$ defined in (4.3) and (5.2) respectively.

Lemma A.1. *For $\psi\in\mathcal{C}^2(\mathbb{R}^d,\mathbb{R})\cap\mathcal{D}(\mathcal{A})\cap\mathcal{D}(\mathcal{A}_\gamma)$, $\mathcal{A}_\gamma\psi(x,k)=\gamma\mathcal{A}\psi(x,k)+\mathcal{R}_1+\mathcal{R}_2+\mathcal{R}_3+\mathcal{R}_4$, where the remainders are defined as*

$$\begin{aligned}\mathcal{R}_1&=(\mathrm{e}^{\gamma Q}-\mathrm{I}_d-\gamma Q)\psi(x,k),\\\mathcal{R}_2&=(\mathrm{e}^{\gamma Q}-\mathrm{I}_d)\mathbb{E}\big[\psi\big(Y^{x,k}_\gamma,k\big)-\psi(x,k)\big],\\\mathcal{R}_3&=\int_0^1\langle\nabla\psi(x+r\gamma g(x,k),k)-\nabla\psi(x,k),\gamma g(x,k)\rangle\,\mathrm{d}r,\\\mathcal{R}_4&=\mathbb{E}\int_0^\gamma[\mathcal{J}\psi(x+\gamma g(x,k)+\sigma Z_s,k)-\mathcal{J}\psi(x,k)]\,\mathrm{d}s,\end{aligned}$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

with the operator $\mathcal{J}$ defined in (4.4), and

$$\mathbb{E}\big[\psi\big(Y_\gamma^{x,k}, k\big) - \psi(x,k)\big] = \gamma\mathcal{J}\psi(x,k) + \gamma\langle\nabla\psi(x,k), g(x,k)\rangle + \mathcal{R}_3 + \mathcal{R}_4.$$

Proposition A.2. *Under Assumptions 1.1 and 1.2, the EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$ is exponentially ergodic with ergodic measure μ_γ. More precisely, for any $(x,k)\in\mathbb{R}^d\times\mathbb{S}$, there exist some positive constants c and C, independent of x, k, and t_n, satisfying*

$$\sup_{\psi\leqslant 1+|\cdot|^{\kappa_1}}\{\bar{P}_{t_n}\psi(x,k) - \mu_\gamma(\psi)\} \leqslant C(1+|x|^{\kappa_1})\mathrm{e}^{-ct_n},$$

where the constant κ_1 is as in Assumption 1.2.

Under Assumption 1.2, the matrices $-Q_{c,\kappa_j}$ with $j=1,2$ are both nonsingular M-matrices with some constants $\kappa_1\in(1,\alpha)$ and $\kappa_2\in(0,\kappa_1/4]$; it follows from Lemma 2.1(iv) that there exist vectors $\xi^{(j)} = \big(\xi_1^{(j)},\dots,\xi_N^{(j)}\big) \gg 0$ such that the vectors $\beta^{(j)} := -Q_{c,\kappa_j}\xi^{(j)} \gg 0$.

Letting $\bar{c}^{(j)} = \min_{i\in\mathbb{S}}\big\{\beta_i^{(j)}/2\xi_i^{(j)}\big\}$, we know that $\bar{c}^{(j)} > 0$ for $j=1,2$. For the constants $\kappa_1\in(1,\alpha)$ and $\kappa_2\in(0,\kappa_1/4]$ in Assumption 1.2, let the Lyapunov function be

$$V^{(j)}(x,i) = (1+|x|^2)^{\kappa_j/2}\xi_i^{(j)} = V^{(j)}(x)\xi_i^{(j)}, \tag{A.1}$$

with $V^{(j)}(x) = (1+|x|^2)^{\kappa_j/2}$ and $j\in\{1,2\}$.

Proof of Proposition A.1. First, we show that the Lyapunov condition holds. For $V^{(j)}(x,i)$ in (A.1), using Taylor's expansion and Young's inequality, for any $\varepsilon>0$ there exists some positive constant $C(\varepsilon)$ such that

$$\begin{aligned}
&\int_{\mathbb{R}_0^d}\big[V^{(j)}(x+\sigma z,i) - V^{(j)}(x,i) - \langle\nabla V^{(j)}(x,i),\sigma z\rangle\mathbf{1}_{\{|z|\leqslant 1\}}\big]\frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z\\
&\quad= \int_{|z|>1}[V^{(j)}(x+\sigma z,i) - V^{(j)}(x,i)]\frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z\\
&\qquad+\int_{0<|z|\leqslant 1}\big[V^{(j)}(x+\sigma z,i) - V^{(j)}(x,i) - \langle\nabla V^{(j)}(x,i),\sigma z\rangle\mathbf{1}_{\{|z|\leqslant 1\}}\big]\frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z\\
&\quad\leqslant \frac{d_\alpha\kappa_j(3-\kappa_j)\sqrt{d}\sigma_{d-1}}{2(2-\alpha)}\xi_i^{(j)} + d_\alpha\kappa_j\sigma_{d-1}\bigg(\frac{|x|^{\kappa_j-1}}{\alpha-1}+\frac{1}{\alpha-\kappa_j}\bigg)\xi_i^{(j)}\\
&\quad\leqslant \varepsilon V^{(j)}(x,i) + C(\varepsilon),
\end{aligned}\tag{A.2}$$

where $\sigma_{d-1} = 2\pi^{d/2}/\Gamma(d/2)$. In addition, under Assumption 1.1 and using Young's inequality, for any $\varepsilon>0$ there exists some positive constant $C(\varepsilon)$ such that

$$\begin{aligned}
&\langle g(x,i),\nabla V^{(j)}(x,i)\rangle\\
&\quad= \kappa_j(1+|x|^2)^{\kappa_j/2-1}\xi_i^{(j)}\langle g(x,i)-g(0,i),x\rangle + \kappa_j(1+|x|^2)^{\kappa_j/2-1}\xi_i^{(j)}\langle g(0,i),x\rangle\\
&\quad\leqslant \kappa_j(1+|x|^2)^{\kappa_j/2-1}\xi_i^{(j)}c_i|x|^2 + |g(0,i)|\kappa_j(1+|x|^2)^{\kappa_j/2-1}|x|\xi_i^{(j)}\\
&\quad= \kappa_j(1+|x|^2)^{\kappa_j/2}\xi_i^{(j)}c_i - \kappa_j(1+|x|^2)^{\kappa_j/2-1}\xi_i^{(j)}c_i + |g(0,i)|\kappa_j(1+|x|^2)^{\kappa_j/2-1}|x|\xi_i^{(j)}\\
&\quad\leqslant \kappa_j c_i V^{(j)}(x,i) + \varepsilon V^{(j)}(x,i) + C(\varepsilon).
\end{aligned}\tag{A.3}$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

Combining (4.3), (A.2), and (A.3) yields

$$\begin{aligned}\mathcal{A}V^{(j)}(x,i) &\leqslant \kappa_j c_i V^{(j)}(x,i) + QV^{(j)}(x,i) + 2\varepsilon V^{(j)}(x,i) + 2C(\varepsilon)\\ &\leqslant \big[(Q+\kappa_j \operatorname{diag}(c_1,\dots,c_N))\xi^{(j)}\big]_i V^{(j)}(x) + 2\varepsilon V^{(j)}(x,i) + 2C(\varepsilon)\\ &= -\beta_i^{(j)} V^{(j)}(x) + 2\varepsilon V^{(j)}(x,i) + 2C(\varepsilon)\\ &\leqslant -\bar{c}^{(j)} V^{(j)}(x,i) + C, \end{aligned} \tag{A.4}$$

where the last inequality holds since $\varepsilon = \beta_i^{(j)}/4\xi_i^{(j)}$. This implies (see [11, Proof of Lemma 7.2]) that

$$\mathbb{E}V^{(j)}\big(X_t^{x,k}, R_t^k\big) \leqslant \mathrm{e}^{-\bar{c}^{(j)}t} V^{(j)}(x,k) + C\big(1-\mathrm{e}^{-\bar{c}^{(j)}t}\big). \tag{A.5}$$

Second, the process $(X_t, R_t)_{t\geqslant 0}$ defined in (1.1) is strong Feller and irreducible, which can be obtained from the discussions in [35, Theorems 3.2 and 4.5], [39, Theorem 2.24], and [37, Theorems 5.6 and 6.3]. Combining the above Lyapunov condition, strong Feller, and irreducibility, it follows from [26, Theorem 6.1] or [35, Theorem 5.4] that the process $(X_t, R_t)_{t\geqslant 0}$ is exponentially ergodic with ergodic measure μ. Furthermore, there exist some positive constants c and C, both independent of x, k, and t, satisfying

$$\sup_{\psi\leqslant 1+V^{(j)}} \{P_t\psi(x,k) - \mu(\psi)\} \leqslant C\big(1+V^{(j)}(x,k)\big)\mathrm{e}^{-ct}.$$

Finally, we take expectations with respect to the ergodic measure μ on both sides of (A.5), and knowing that, for $t=1$, $\mu(V^{(j)}) \leqslant \mathrm{e}^{-\bar{c}^{(j)}}\mu(V^{(j)}) + C$, this implies that $\mu(V^{(j)}) \leqslant C$ for some constant $C>0$. The proof is complete. □

Proof of Proposition A.2. Let $V^{(1)}$ be as defined in (A.1). It follows from Lemma A.1 that

$$\mathcal{A}_\gamma V^{(1)}(x,k) = \gamma \mathcal{A}V^{(1)}(x,k) + \mathcal{R}_{V^{(1)},1} + \mathcal{R}_{V^{(1)},2} + \mathcal{R}_{V^{(1)},3} + \mathcal{R}_{V^{(1)},4},$$

where

$$\begin{aligned}\mathcal{R}_{V^{(1)},1} &= (\mathrm{e}^{\gamma Q} - \mathrm{I}_d - \gamma Q)V^{(1)}(x,k),\\ \mathcal{R}_{V^{(1)},2} &= (\mathrm{e}^{\gamma Q} - \mathrm{I}_d)\mathbb{E}\big[V^{(1)}\big(Y_\gamma^{x,k},k\big) - V^{(1)}(x,k)\big],\\ \mathcal{R}_{V^{(1)},3} &= \int_0^1 \langle \nabla V^{(1)}(x + r\gamma g(x,k),k) - \nabla V^{(1)}(x,k), \gamma g(x,k)\rangle \,\mathrm{d}r,\\ \mathcal{R}_{V^{(1)},4} &= \mathbb{E}\int_0^\gamma [\mathcal{J}V^{(1)}(x+\gamma g(x,k) + \sigma Z_s, k) - \mathcal{J}V^{(1)}(x,k)]\,\mathrm{d}s\end{aligned}$$

and

$$\mathbb{E}\big[V^{(1)}\big(Y_\gamma^{x,k},k\big) - V^{(1)}(x,k)\big] = \gamma \mathcal{J}V^{(1)}(x,k) + \gamma\langle \nabla V^{(1)}(x,k), g(x,k)\rangle + \mathcal{R}_{V^{(1)},3} + \mathcal{R}_{V^{(1)},4}.$$

We need to give estimates for the remainders $\mathcal{R}_{V^{(1)},i}$, $i=1,2,3,4$.

For the term $\mathcal{R}_{V^{(1)},1}$ and for small enough $\gamma>0$, we know that

$$|\mathcal{R}_{V^{(1)},1}| \leqslant C\gamma^2 V^{(1)}(x,k). \tag{A.6}$$

For $\kappa_1 \in (1,\alpha)$ in Assumption 1.2 and using the interpolation inequality in [17, Theorem 3.2.1], we know that, for any $x\in\mathbb{R}^d$,

$$[\nabla V]_{\kappa_1-1;B_1(x)} \leqslant C\varepsilon^{1-\kappa_1}[V^{(1)}]_{1;B_1(x)} + C\varepsilon^{2-\kappa}\|\nabla^2 V^{(1)}\|_\infty,$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

where $B_1(x) = \{z\colon |z-x| \leqslant 1\}$. Taking $\varepsilon = 1 + |x|$, we have $[\nabla V^{(1)}]_{\kappa_1-1;\mathbb{R}^d} \leqslant C$, so

$$\begin{aligned}|\mathcal{R}_{V^{(1)},3}| &= \left| \int_0^1 \left\langle \frac{\nabla V^{(1)}(x + r\gamma g(x,k),k) - \nabla V^{(1)}(x,k)}{|r\gamma g(x,k)|^{\kappa_1-1}} |r\gamma g(x,k)|^{\kappa_1-1}, \gamma g(x,k) \right\rangle \mathrm{d}r \right| \\ &\leqslant C|\gamma g(x,k)|^{\kappa_1} \leqslant C\gamma^{\kappa_1} V^{(1)}(x,k). \end{aligned} \tag{A.7}$$

Since

$$\mathcal{R}_{V^{(1)},4} = \mathbb{E}\int_0^\gamma \frac{\mathcal{J}V^{(1)}(x + \gamma g(x,k) + \sigma Z_s, k) - \mathcal{J}V^{(1)}(x,k)}{|\gamma g(x,k) + \sigma Z_s|^{2-\alpha}} |\gamma g(x,k) + \sigma Z_s|^{2-\alpha}\, \mathrm{d}s,$$

for small enough $\gamma > 0$ we have

$$|\mathcal{R}_{V^{(1)},4}| \leqslant C(1 + |x|^{2-\alpha})\gamma^{2/\alpha} \leqslant C\gamma^{2/\alpha} V^{(1)}(x,k), \tag{A.8}$$

where the second inequality holds from [7, Lemma 3.1] and the fact that $3 - \alpha > 2/\alpha$ for $\alpha \in (1, 2)$.

Since

$$\begin{aligned}\mathcal{R}_{V^{(1)},2} &= (\mathrm{e}^{\gamma Q} - \mathrm{I}_d)\mathbb{E}\big[V^{(1)}\big(Y_\gamma^{x,k}, k\big) - V^{(1)}(x,k)\big] \\ &= (\mathrm{e}^{\gamma Q} - \mathrm{I}_d)\mathbb{E}\{\gamma \mathcal{J}V^{(1)}(x,k) + \gamma\langle \nabla V^{(1)}(x,k), g(x,k)\rangle + \mathcal{R}_{V^{(1)},3} + \mathcal{R}_{V^{(1)},4}\},\end{aligned}$$

it follows from (A.2) and the estimates for $|\mathcal{R}_{V^{(1)},3}|$ and $|\mathcal{R}_{V^{(1)},4}|$ that, for small enough $\gamma > 0$,

$$\begin{aligned}|\mathcal{R}_{V^{(1)},2}| &\leqslant C\gamma\{\gamma[\varepsilon V^{(1)}(x,k) + C(\varepsilon)] + \gamma V^{(1)}(x,k) + \gamma^{\kappa_1} V^{(1)}(x,k) + \gamma^{2/\alpha} V^{(1)}(x,k)\} \\ &\leqslant C\gamma^2 V^{(1)}(x,k) + C\gamma^2. \end{aligned} \tag{A.9}$$

Combining (A.6), (A.7), (A.8), and (A.9), for small enough $\gamma > 0$ we know that

$$\begin{aligned}|(\mathcal{A}_\gamma - \gamma\mathcal{A})V^{(1)}(x,k)| &\leqslant |\mathcal{R}_{V^{(1)},1}| + |\mathcal{R}_{V^{(1)},2}| + |\mathcal{R}_{V^{(1)},3}| + |\mathcal{R}_{V^{(1)},4}| \\ &\leqslant C(\gamma^2 + \gamma^{\kappa_1} + \gamma^{2/\alpha})V^{(1)}(x,k) + C\gamma^2,\end{aligned}$$

and thus $|(\mathcal{A}_\gamma - \gamma\mathcal{A})V^{(1)}(x,k)| \leqslant C\gamma^{\kappa_1\wedge 2/\alpha} V^{(1)}(x,k) + C\gamma^2$. It then follows from (A.5) that

$$\begin{aligned}\mathbb{E}V^{(1)}\big(Y_{t_{n-1},t_n}^{x,k}, R_{t_{n-1},t_n}^k\big) &= \mathbb{E}\big[V^{(1)}\big(Y_{t_{n-1},t_n}^{x,k}, R_{t_{n-1},t_n}^k\big) - V^{(1)}(x,k)\big] + V^{(1)}(x,k) \\ &= \mathbb{E}\mathcal{A}_\gamma V^{(1)}(x,k) + V^{(1)}(x,k) \\ &\leqslant \mathbb{E}\big[\gamma\mathcal{A}V^{(1)}(x,k) + C\gamma^{\kappa_1\wedge 2/\alpha} V^{(1)}(x,k) + C\gamma^2\big] + V^{(1)}(x,k) \\ &\leqslant (1 - \bar{c}^{(1)}\gamma + C\gamma^{\kappa_1\wedge 2/\alpha})V^{(1)}(x,k) + C\gamma.\end{aligned}$$

Then, for small enough $\gamma > 0$, there exists some positive constant C, independent of n, γ, x, and k, such that

$$\mathbb{E}V^{(1)}\big(Y_{t_n}^{x,k}, R_{t_n}^k\big) \leqslant \left(1 - \frac{\bar{c}^{(1)}}{2}\gamma\right)\mathbb{E}V^{(1)}\big(Y_{t_{n-1}}^{x,k}, R_{t_{n-1}}^k\big) + C\gamma. \tag{A.10}$$

From standard discussions in [25, 26], there exist some positive constants c and C such that

$$\sup_{\psi \leqslant 1 + V^{(1)}} \{\bar{P}_{t_n}\psi(x,k) - \mu_\gamma(\psi)\} \leqslant C(1 + V^{(1)}(x))\mathrm{e}^{-ct_n},$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

which implies the desired result. □

Proof of Lemma A.1. For $\psi \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}) \cap \mathcal{D}(\mathcal{A}) \cap \mathcal{D}(\mathcal{A}_\gamma)$, we have

$$\begin{aligned}\mathcal{A}_\gamma \psi(x,k) &= \mathbb{E}\big[\psi\big(Y_\gamma^{x,k}, R_\gamma^k\big) - \psi(x,k)\big] \\ &= \mathbb{E}\big[\psi\big(Y_\gamma^{x,k}, R_\gamma^k\big) - \psi\big(Y_\gamma^{x,k}, k\big)\big] + \mathbb{E}[\psi\big(Y_\gamma^{x,k}, k\big) - \psi(x,k)]. \end{aligned} \tag{A.11}$$

By the conditional expectation, it follows from [23, Theorem 1.44] that

$$\begin{aligned}&\mathbb{E}\big[\psi\big(Y_\gamma^{x,k}, R_\gamma^k\big) - \psi\big(Y_\gamma^{x,k}, k\big)\big] \\ &\quad= \mathbb{E}\big\{\mathbb{E}\big[\psi\big(Y_\gamma^{x,k}, R_\gamma^k\big) - \psi\big(Y_\gamma^{x,k}, k\big)\big] \mid \hat{\mathcal{F}}_\gamma\big\} \\ &\quad= (\mathrm{e}^{\gamma Q} - \mathrm{I}_d)\mathbb{E}\psi\big(Y_\gamma^{x,k}, k\big) \\ &\quad= (\mathrm{e}^{\gamma Q} - \mathrm{I}_d)\psi(x,k) + (\mathrm{e}^{\gamma Q} - \mathrm{I}_d)\mathbb{E}\big[\psi\big(Y_\gamma^{x,k}, k\big) - \psi(x,k)\big] \\ &\quad= \gamma Q\psi(x,k) + (\mathrm{e}^{\gamma Q} - \mathrm{I}_d - \gamma Q)\psi(x,k) + (\mathrm{e}^{\gamma Q} - \mathrm{I}_d)\mathbb{E}\big[\psi\big(Y_\gamma^{x,k}, k\big) - \psi(x,k)\big] \\ &\quad= \gamma Q\psi(x,k) + \mathcal{R}_1 + \mathcal{R}_2, \end{aligned} \tag{A.12}$$

where $\hat{\mathcal{F}}_\gamma$ is the filtration generated by the α-stable process $(Z_t)_{t\geqslant 0}$ at time $t = \gamma$. Using Itô's formula and Taylor's expansion, we obtain

$$\begin{aligned}&\mathbb{E}\big[\psi\big(Y_\gamma^{x,k}, k\big) - \psi(x,k)\big] \\ &\quad= \mathbb{E}[\psi(x + \gamma g(x,k) + \sigma Z_\gamma, k) - \psi(x + \gamma g(x,k), k) + \psi(x + \gamma g(x,k), k) - \psi(x,k)] \\ &\quad= \mathbb{E}\int_0^\gamma \mathcal{J}\psi(x + \gamma g(x,k) + \sigma Z_s, k)\,\mathrm{d}s + \gamma\langle\nabla\psi(x,k), g(x,k)\rangle \\ &\qquad+ \int_0^1 \langle\nabla\psi(x + r\gamma g(x,k), k) - \nabla\psi(x,k), \gamma g(x,k)\rangle\,\mathrm{d}r \\ &\quad= \gamma\mathcal{J}\psi(x,k) + \mathbb{E}\int_0^\gamma [\mathcal{J}\psi(x + \gamma g(x,k) + \sigma Z_s, k) - \mathcal{J}\psi(x,k)]\,\mathrm{d}s \\ &\qquad+ \gamma\langle\nabla\psi(x,k), g(x,k)\rangle + \mathcal{R}_3. \end{aligned} \tag{A.13}$$

Combining (4.3), (A.11), (A.12), and (A.13), we get the desired result. □

A.2. Moment estimates

Moment estimates for the process $(X_t, R_t)_{t\geqslant 0}$ can be obtained from the Lyapunov condition in Appendix A.1. Combining this with the generator comparison in Lemma A.1 yields moment estimates for the EM scheme $(Y_{t_n}, R_{t_n})_{n\in\mathbb{N}_0}$.

Proof of Lemma 3.2.

(i) The moment estimate for the process $(X_t^{x,k}, R_t^k)_{t\geqslant 0}$ holds from (A.5), i.e. for $\kappa_1 \in (1, \alpha)$ in Assumption 1.2,

$$\mathbb{E}|X_t^{x,k}|^{\kappa_1} \leqslant C\mathbb{E}V^{(1)}(X_t^{x,k}, R_t^k) \leqslant C\mathrm{e}^{-\bar{c}^{(1)}t}V^{(1)}(x,k) + C(1 - \mathrm{e}^{-\bar{c}^{(1)}t}), \tag{A.14}$$

which implies the desired estimate.

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

(ii) By induction, it follows from (A.10) that there exists some positive constant C, independent of n, such that

$$\mathbb{E}V^{(1)}\big(Y_{t_n}^{x,k}, R_{t_n}^k\big) \leqslant \left(1 - \frac{\bar{c}}{2}\gamma\right)^n V^{(1)}(x,k) + C\gamma \sum_{j=1}^{n-1} \left(1 - \frac{\bar{c}}{2}\gamma\right)^j \leqslant C(1 + V^{(1)}(x,k)).$$

(iii) Recall that $X_t^{x,k} = x + \int_0^t g(X_s^{x,k}, R_s^k)\,\mathrm{d}s + \sigma Z_t$, so for $0 \leqslant p \leqslant \kappa_1 < \alpha$, using the Hölder and Cramér–Rao inequalities, we get

$$\mathbb{E}|X_t^{x,k} - x|^p = \mathbb{E}\left|\int_0^t g\big(X_s^{x,k}, R_s^k\big)\,\mathrm{d}s + \sigma Z_t\right|^p \leqslant C(1 + |x|^p)(t^{p/\alpha} \vee t^p), \tag{A.15}$$

where the inequality holds from the linear growth condition for g and the estimate for $\mathbb{E}|X_t^{x,k}|^{\kappa_1}$ in (A.14).

(iv) Since $X_\gamma^{x,k} - Y_\gamma^{x,k} = \int_0^\gamma [g(X_s^{x,k}, R_s^k) - g(x,k)]\,\mathrm{d}s$, using the Hölder inequality we get, for $0 \leqslant p \leqslant \kappa_1 < \alpha$,

$$\begin{aligned}
\mathbb{E}\big|X_\gamma^{x,k} - Y_\gamma^{x,k}\big|^p &= \mathbb{E}\left|\int_0^\gamma \big[g(X_s^{x,k}, R_s^k) - g\big(x, R_s^k\big) + g\big(x, R_s^k\big) - g\big(x,k\big)\big]\,\mathrm{d}s\right|^p \\
&\leqslant C\mathbb{E}\left|\int_0^\gamma |X_s^{x,k} - x|\,\mathrm{d}s\right|^p + C(1+|x|^p)\mathbb{E}\left|\int_0^\gamma \mathbf{1}_{\{R_s^k \neq k\}}\,\mathrm{d}s\right|^p \\
&\leqslant C\gamma^{p-1}\int_0^\gamma \mathbb{E}|X_s^{x,k} - x|^p\,\mathrm{d}s + C(1+|x|^p)\gamma^{p-1}\int_0^\gamma \mathbb{P}(R_s^k \neq k)\,\mathrm{d}s \\
&\leqslant C(1+|x|^p)\gamma^{p(1+(1/\alpha))},
\end{aligned}$$

where the last inequality holds from the estimate for $\mathbb{E}|X_s^{x,k} - x|^p$ in (A.15) and the fact $\mathbb{P}(R_s^k \neq k) \leqslant Cs$ with $s \leqslant \gamma$; indeed, $\mathbb{P}(R_s^k \neq k) = 1 - \mathbb{P}(R_s^k = k) = 1 - (\mathrm{e}^{s\mathcal{Q}})_{kk} \leqslant Cs$ for some small $s \leqslant \gamma$.

(v) Recalling that $Y_\gamma^{x,k} = x + \gamma g(x,k) + \sigma Z_\gamma$, it is obvious that for $0 \leqslant p \leqslant \kappa_1 < \alpha$ there exits some positive constant C such that $\mathbb{E}|Y_\gamma^{x,k} - x|^p \leqslant C\gamma^{p/\alpha}(1+|x|^p)$. The proof is complete. □

Appendix B. One-step difference

Proof of Lemma 3.3. For $f \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ satisfying $\|\nabla f\|_\infty + \|\nabla^2 f\|_\infty < \infty$, we know that

$$\begin{aligned}
&(P_\gamma - \bar{P}_\gamma)f(x,k) \\
&\quad = \mathbb{E}f\big(X_\gamma^{x,k}, R_\gamma^k\big) - \mathbb{E}f\big(Y_\gamma^{x,k}, R_\gamma^k\big) \\
&\quad = \mathbb{E}\big\langle \nabla f(x,k), X_\gamma^{x,k} - Y_\gamma^{x,k}\big\rangle + \mathbb{E}\big\langle \nabla f\big(x, R_\gamma^k\big) - \nabla f(x,k), X_\gamma^{x,k} - Y_\gamma^{x,k}\big\rangle \\
&\qquad + \mathbb{E}\int_0^1 \big\langle \nabla f\big(Y_\gamma^{x,k} + r\big[X_\gamma^{x,k} - Y_\gamma^{x,k}\big], R_\gamma^k\big) - \nabla f\big(x, R_\gamma^k\big), X_\gamma^{x,k} - Y_\gamma^{x,k}\big\rangle\,\mathrm{d}r \\
&\quad = \mathrm{I} + \mathrm{II} + \mathrm{III}.
\end{aligned} \tag{B.1}$$

We first give the estimate for I. Recall that

$$\mathbb{E}\big[X_\gamma^{x,k} - Y_\gamma^{x,k}\big] = \int_0^\gamma \mathbb{E}\big[g\big(X_s^{x,k}, R_s^k\big) - g(x,k)\big]\,\mathrm{d}s = \int_0^\gamma \int_0^s \mathbb{E}\mathcal{A}g\big(X_u^{x,k}, R_u^k\big)\,\mathrm{d}u\,\mathrm{d}s, \tag{B.2}$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

where $\mathcal{A}$ is defined in (4.3). Under Assumption 1.1(ii), we claim that

$$|\mathcal{A}g(x,k)| \leqslant C(1+|x|), \tag{B.3}$$

which implies that

$$\big|\mathbb{E}\big[X_\gamma^{x,k} - Y_\gamma^{x,k}\big]\big| \leqslant C\int_0^\gamma \int_0^s \mathbb{E}\big(1+|X_u^{x,k}|\big)\,\mathrm{d}u\,\mathrm{d}s \leqslant C(1+|x|)\gamma^2,$$

where the moment estimate for $\mathbb{E}|X_u^{x,k}|$ is from Lemma 3.2. This yields that

$$|\mathrm{I}| \leqslant C\|\nabla f\|_\infty(1+|x|)\gamma^2. \tag{B.4}$$

For the second term II, by the conditional expectation we know that

$$\begin{aligned}&\mathbb{E}\big\langle \nabla f\big(x, R_\gamma^k\big) - \nabla f(x,k), X_\gamma^{x,k} - Y_\gamma^{x,k}\big\rangle\\ &\qquad = \mathbb{E}\big\{\big\langle \nabla f\big(x, R_\gamma^k\big) - \nabla f(x,k), \mathbb{E}\big[X_\gamma^{x,k} - Y_\gamma^{x,k} \mid R_\gamma^k\big]\big\rangle\big\};\end{aligned}$$

it follows from (B.2) and the linear growth condition for g that

$$\big|\mathbb{E}\big[X_\gamma^{x,k} - Y_\gamma^{x,k} \mid R_\gamma^k\big]\big| = \left|\int_0^\gamma C\mathbb{E}\big(1+\big|X_s^{x,k}\big|\big)\,\mathrm{d}s\right| \leqslant C(1+|x|)\gamma.$$

In addition, $|\mathbb{E}[\nabla f(x, R_\gamma^k) - \nabla f(x,k)]| = |[\mathrm{e}^{\gamma Q} - \mathrm{I}_d]\nabla f(x,k)| \leqslant C\|\nabla f\|_\infty\gamma$. Thus, we have

$$|\mathrm{II}| \leqslant C\|\nabla f\|_\infty(1+|x|)\gamma^2. \tag{B.5}$$

For the last term III, we know that

$$\begin{aligned}&\left|\mathbb{E}\int_0^1 \big\langle \nabla f\big(Y_\gamma^{x,k} + r\big[X_\gamma^{x,k} - Y_\gamma^{x,k}\big], R_\gamma^k\big) - \nabla f\big(x, R_\gamma^k\big), X_\gamma^{x,k} - Y_\gamma^{x,k}\big\rangle\,\mathrm{d}r\right|\\ &\quad\leqslant \mathbb{E}\int_0^1 \frac{\big|\nabla f\big(Y_\gamma^{x,k} + r\big[X_\gamma^{x,k} - Y_\gamma^{x,k}\big], R_\gamma^k\big) - \nabla f\big(x, R_\gamma^k\big)\big|}{\big|Y_\gamma^{x,k} - x + r\big[X_\gamma^{x,k} - Y_\gamma^{x,k}\big]\big|^{\kappa_1-1}}\\ &\qquad\times \big(\big|Y_\gamma^{x,k} - x + r\big[X_\gamma^{x,k} - Y_\gamma^{x,k}\big]\big|^{\kappa_1-1}\big)\big|X_\gamma^{x,k} - Y_\gamma^{x,k}\big|\,\mathrm{d}r\\ &\quad\leqslant C(\|\nabla f\|_\infty + \|\nabla^2 f\|_\infty)\mathbb{E}\big(\big|Y_\gamma^{x,k} - x\big|^{\kappa_1-1}\big|X_\gamma^{x,k} - Y_\gamma^{x,k}\big| + \big|X_\gamma^{x,k} - Y_\gamma^{x,k}\big|^{\kappa_1}\big),\end{aligned}$$

where the last inequality holds from the fact that, for any $x, y \in \mathbb{R}^d$, $k \in \mathbb{S}$, and $\tau \in (0,1)$,

$$|\nabla f(x,k) - \nabla f(y,k)| \leqslant C(\|\nabla f\|_\infty + \|\nabla^2 f\|_\infty)|x-y|^\tau.$$

By the Hölder inequality, we obtain

$$\begin{aligned}\mathbb{E}\big[\big|Y_\gamma^{x,k} - x\big|^{\kappa_1-1}\big|X_\gamma^{x,k} - Y_\gamma^{x,k}\big|\big] &\leqslant \big[\mathbb{E}\big|Y_\gamma^{x,k} - x\big|^{p(\kappa_1-1)}\big]^{1/p}\big[\mathbb{E}\big|X_\gamma^{x,k} - Y_\gamma^{x,k}\big|^q\big]^{1/q}\\ &\leqslant C(1+|x|^{\kappa_1})\gamma^{1+(\kappa_1/\alpha)},\end{aligned}$$

where

$$p = \frac{1}{2}\left(\frac{\alpha}{\alpha-1} + \frac{\alpha}{\kappa_1-1}\right).$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

This implies that

$$|\mathrm{III}| \leqslant C(\|\nabla f\|_\infty + \|\nabla^2 f\|_\infty)(1+|x|^{\kappa_1})\gamma^{1+(\kappa_1/\alpha)}. \tag{B.6}$$

The desired inequality holds from (B.1), (B.4), (B.5), and (B.6).

It remains to show the claim (B.3). Recall that

$$\begin{aligned}\mathcal{A}g(x,k) = {}& \langle \nabla g(x,k), g(x,k)\rangle + Qg(x,k) + \int_{|z|>1} [g(x+\sigma z,k) - g(x,k)]\frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z \\ & + \int_{0<|z|\leqslant 1} [g(x+\sigma z,k) - g(x,k) - \langle \nabla g(x,k), \sigma z\rangle]\frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z. \end{aligned}\tag{B.7}$$

Since $\|\nabla g\|_\infty \leqslant C$, we know that g satisfies the linear growth condition and

$$|\langle \nabla g(x,k), g(x,k)\rangle + Qg(x,k)| \leqslant C(1+|x|). \tag{B.8}$$

In addition, under $\|\nabla g\|_\infty + \|\nabla^2 g\|_\infty \leqslant C$, by the Taylor expansion,

$$\int_{|z|>1} [g(x+\sigma z,k) - g(x,k)]\frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z = \int_{|z|>1}\int_0^1 \langle \sigma z, \nabla g(x+r\sigma z,k)\rangle\,\mathrm{d}r\frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z$$

and

$$\begin{aligned}&\int_{0<|z|\leqslant 1} [g(x+\sigma z,k) - g(x,k) - \langle \nabla g(x,k), \sigma z\rangle]\frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z \\ &\qquad = \int_{0<|z|\leqslant 1}\int_0^1\int_0^1 \langle r\sigma z(\sigma z)^*, \nabla^2 g(x+\tilde{r}r\sigma z,k)\rangle_{\mathrm{HS}}\,\mathrm{d}\tilde{r}\,\mathrm{d}r\frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z.\end{aligned}$$

This implies that

$$\begin{aligned}\left|\int_{|z|>1} [g(x+\sigma z,k) - g(x,k)]\frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z\right| &\leqslant C, \\ \left|\int_{0<|z|\leqslant 1} [g(x+\sigma z,k) - g(x,k) - \langle \nabla g(x,k), \sigma z\rangle]\frac{d_\alpha}{|z|^{d+\alpha}}\,\mathrm{d}z\right| &\leqslant C.\end{aligned}$$

Combining these with (B.7) and (B.8), the claim in (B.3) holds. The proof is complete. □

Appendix C. Regularities for Stein's equation

Letting $(X_0, R_0) = (x,k)$, the Jacobi flow of X_t along the direction $u \in \mathbb{R}^d$ is defined as

$$\nabla_u X_t^{x,k} = \lim_{\varepsilon\to 0}\frac{X_t^{x+\varepsilon u,k} - X_t^{x,k}}{\varepsilon}, \quad t \geqslant 0.$$

Lemma C.1. *Under Assumptions 1.1 and 1.2, with $\eta > 0$ defined in (1.4), for $u_1, u_2 \in \mathbb{R}^d$ satisfying $|u_1| = |u_2| = 1$ and $(x,k) \in \mathbb{R}^d \times \mathbb{S}$, there exists some positive constant C, independent of x, k, and t, such that $\mathbb{E}|\nabla_{u_1} X_t^{x,k}|^4 \leqslant C\exp(-\eta t)$ and $\mathbb{E}|\nabla_{u_2}\nabla_{u_1} X_t^{x,k}|^2 \leqslant C\exp(-\eta t/4)$.*

The proof of Lemma C.1 is deferred to the end of this section.

Proof of Lemma 3.1. For any $u_1 \in \mathbb{R}^d$ satisfying $|u_1| = 1$,

$$\nabla_{u_1} P_t h(x,k) = \nabla_{u_1}\mathbb{E}h\big(X_t^{x,k}, R_t^k\big) = \mathbb{E}\big[\nabla h\big(X_t^{x,k}, R_t^k\big)\nabla_{u_1} X_t^{x,k}\big].$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

Then, using Lemma C.1 and $\|\nabla h\|_\infty < \infty$, there exists some positive constant C, independent of t, x, and k, such that

$$\begin{aligned}\|\nabla P_t h(x,k)\|_\infty &= \sup_{(x,k)\in\mathbb{R}^d\times\mathbb{S}} |\nabla_{u_1} P_t h(x,k)| \\ &\leqslant \|\nabla h\|_\infty \sup_{(x,k)\in\mathbb{R}^d\times\mathbb{S}} \mathbb{E}[|\nabla_{u_1} X_t^{x,k}|] \leqslant C\mathrm{e}^{-\eta t/4}.\end{aligned}$$

For any $u_i \in \mathbb{R}^d$ with $i = 1, 2$ satisfying $|u_1| = |u_2| = 1$, we know that

$$\nabla_{u_2}\nabla_{u_1} P_t h(x,k) = \mathbb{E}\big[\nabla h\big(X_t^{x,k}, R_t^k\big)\nabla_{u_2}\nabla_{u_1} X_t^{x,k} + \nabla^2 h\big(X_t^{x,k}, R_t^k\big)\nabla_{u_2} X_t^{x,k}\nabla_{u_1} X_t^{x,k}\big].$$

Using Lemma C.1 and $\|\nabla^i h\|_\infty < \infty$ with $i = 1, 2$, there then exists some positive constant C, independent of t, x, and k, such that $|\nabla_{u_2}\nabla_{u_1} P_t h(x,k)| \leqslant C\mathrm{e}^{-\eta t/8}$, which implies the estimate for $\|\nabla^2 P_t h(x,k)\|_\infty$. The proof is complete. □

Proof of Lemma 4.1. With arguments similar to [14, Proposition 3.2], the solution f_h of the Stein equation (4.5) can be expressed as $f_h(x,k) = -\int_0^{+\infty} [P_t h(x,k) - \mu(h)]\,\mathrm{d}t$, where P_t is the semigroup of the process $(X_t, R_t)_{t\geqslant 0}$.

By the exponential ergodicity of the process $(X_t, R_t)_{t\geqslant 0}$ in Proposition A.1, there exists some positive constant C such that

$$|f_h(x,k)| \leqslant \int_0^{+\infty} |P_t h(x,k) - \mu(h)|\,\mathrm{d}t \leqslant C(1 + |x|^{\kappa_2}).$$

For any $u_1 \in \mathbb{R}^d$, we know that

$$\nabla_{u_1} f_h(x,k) = -\int_0^{+\infty} \nabla_{u_1}\mathbb{E}h\big(X_t^{x,k}, R_t^k\big)\,\mathrm{d}t = -\int_0^{+\infty} \mathbb{E}\big[\nabla h\big(X_t^{x,k}, R_t^k\big)\nabla_{u_1} X_t^{x,k}\big]\,\mathrm{d}t.$$

Combining this with Lemma C.1 and the fact that $\|\nabla h\|_\infty \leqslant C$, we get that $\|\nabla f_h(x,k)\|_\infty \leqslant C$. The proof is complete. □

Proof of Lemma C.1. For any $u_1 \in \mathbb{R}^d$ with $|u_1| = 1$, we have $\nabla_{u_1} X_0^{x,k} = u_1$ and

$$\mathrm{d}\nabla_{u_1} X_t^{x,k} = \nabla g\big(X_t^{x,k}, R_t^k\big)\nabla_{u_1} X_t^{x,k}\,\mathrm{d}t,$$

which implies that

$$\mathrm{d}|\nabla_{u_1} X_t^{x,k}|^4 = \langle 4\nabla_{u_1} X_t^{x,k}|\nabla_{u_1} X_t^{x,k}|^2, \nabla g(X_t^{x,k}, R_t^x)\nabla_{u_1} X_t^{x,k}\rangle\,\mathrm{d}t \leqslant 4c_{R_t^k}|\nabla_{u_1} X_t^{x,k}|^4\,\mathrm{d}t,$$

where the last inequality holds from (1.2). This can be estimated as

$$|\nabla_{u_1} X_t^{x,k}|^4 \leqslant \exp\bigg(\int_0^t 4c_{R_u^k}\,\mathrm{d}u\bigg), \tag{C.1}$$

and it follows from Lemma 1.1 that $\mathbb{E}|\nabla_{u_1} X_t^{x,k}|^4 \leqslant C\exp(-\eta t)$.

For any $u_1, u_2 \in \mathbb{R}^d$ with $|u_1| = |u_2| = 1$, we know that $\nabla_{u_2}\nabla_{u_1} X_0^{x,k} = 0$ and

$$\mathrm{d}\nabla_{u_2}\nabla_{u_1} X_t^{x,k} = \nabla g\big(X_t^{x,k}, R_t^k\big)\nabla_{u_2}\nabla_{u_1} X_t^{x,k}\,\mathrm{d}t + \nabla^2 g\big(X_t^{x,k}, R_t^k\big)\nabla_{u_2} X_t^{x,k}\nabla_{u_1} X_t^{x,k}\,\mathrm{d}t.$$

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

using Itô's formula and Young's inequality, for any $a > 0$ we have

$$\begin{aligned}\mathrm{d}|\nabla_{u_2}\nabla_{u_1}X_t^{x,k}|^2 &= \big\langle 2\nabla_{u_2}\nabla_{u_1}X_t^{x,k},\, \nabla g\big(X_t^{x,k}, R_t^k\big)\nabla_{u_2}\nabla_{u_1}X_t^{x,k} \\ &\quad + \nabla^2 g\big(X_t^{x,k}, R_t^k\big)\nabla_{u_2}X_t^{x,k}\nabla_{u_1}X_t^{x,k}\big\rangle\,\mathrm{d}t \\ &\leqslant (2c_{R_t^k} + a)|\nabla_{u_2}\nabla_{u_1}X_t^{x,k}|^2\,\mathrm{d}t + C|\nabla_{u_2}X_t^{x,k}|^2|\nabla_{u_1}X_t^{x,k}|^2\,\mathrm{d}t,\end{aligned}$$

which can be estimated as

$$|\nabla_{u_2}\nabla_{u_1}X_t^{x,k}|^2 \leqslant C\int_0^t \exp\bigg(\int_s^t [2c_{R_u^k} + a]\,\mathrm{d}u\bigg)|\nabla_{u_2}X_s^{x,k}|^2|\nabla_{u_1}X_s^{x,k}|^2\,\mathrm{d}s.$$

Combining this with (C.1), we know that

$$\begin{aligned}|\nabla_{u_2}\nabla_{u_1}X_t^{x,k}|^2 &\leqslant C\int_0^t \exp\bigg(\int_s^t [2c_{R_u^k} + a]\,\mathrm{d}u\bigg)\exp\bigg(\int_0^s 4c_{R_u^k}\,\mathrm{d}u\bigg)\mathrm{d}s \\ &= C\int_0^t \mathrm{e}^{a(t-s)}\exp\bigg(\int_0^t 2c_{R_u^k}\,\mathrm{d}u\bigg)\exp\bigg(\int_0^s 2c_{R_u^k}\,\mathrm{d}u\bigg)\mathrm{d}s,\end{aligned}$$

which implies that

$$\begin{aligned}\mathbb{E}|\nabla_{u_2}\nabla_{u_1}X_t^{x,k}|^2 &\leqslant C\int_0^t \mathrm{e}^{a(t-s)}\bigg[\mathbb{E}\exp\bigg(\int_0^t 4c_{R_u^k}\,\mathrm{d}u\bigg)\bigg]^{1/2}\bigg[\mathbb{E}\exp\bigg(\int_0^s 4c_{R_u^k}\,\mathrm{d}u\bigg)\bigg]^{1/2}\mathrm{d}s \\ &\leqslant C\int_0^t \mathrm{e}^{a(t-s)}\exp\bigg(-\frac{\eta}{2}t\bigg)\exp\bigg(-\frac{\eta}{2}s\bigg)\mathrm{d}s \leqslant C\mathrm{e}^{-\eta t/4},\end{aligned}$$

where the second inequality holds from Lemma 1.1 and the last inequality holds by $a = \eta/4$. The proof is complete. □

Acknowledgements

We wish to thank the editors and referees for their helpful comments.

Funding Information

X. Jin is supported by the National Natural Science Foundation of China (grant no. 12401176), and the Fundamental Research Funds for the Central Universities (grant no. JZ2025HGTB0178).

Competing Interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus*. Cambridge University Press.
[2] Bao, J., Shao, J. and Yuan, C. (2016). Approximation of invariant measures for regime-switching diffusions. *Potential Anal.* **44**, 707–727.
[3] Bardet, J. B., Guérin, H. and Malrieu, F. (2010). Long time behavior of diffusions with Markov switching. *Alea* **7**, 151–170.
[4] Bras, P., Pagés, G. and Panloup, F. (2022). Total variation distance between two diffusions in small time with unbounded drift: Application to the Euler–Maruyama scheme. *Electron. J. Prob.* **27**, 1–19.

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

[5] Chen, P., Deng, C. S., Schilling, R. L. and Xu, L. (2023). Approximation of the invariant measure of stable SDEs by an Euler–Maruyama scheme. *Stoch. Process. Appl.* **163**, 136–167.
[6] Chen, P., Jin, X., Shen, T. and Su, Z. (2024). Variable-step Euler–Maruyama's approximations of regime-switching jump diffusion processes. *J. Theoret. Prob.* **37**, 1597–1626.
[7] Chen, P., Jin, X., Xiao, Y. and Xu, L. (2025). Approximation of the invariant measure for stable SDE by the Euler–Maruyama scheme with decreasing step-sizes. *Adv. Appl. Prob.* **57**, 1000–1030.
[8] Cloez, B. and Hairer, M. (2015). Exponential ergodicity for Markov processes with random switching. *Bernoulli* **21**, 505–536.
[9] Da, P. G. and Zabczyk, J. (1996). *Ergodicity for Infinite Dimensional Systems* (London Math. Soc. Lect. Note Ser. **229**). Cambridge University Press.
[10] Fang, X., Shao, Q. M. and Xu, L. (2019). Multivariate approximations in Wasserstein distance by Stein's method and Bismut's formula. *Prob. Theory Relat. Fields* **174**, 945–979.
[11] Gurvich, I. (2014). Diffusion models and steady state approximations for exponentially ergodic Markovian queues. *Ann. Appl. Prob.* **24**, 2527–2559.
[12] Ji, Y. and Chizeck, H. J. (1990). Controllability, stabilizability, and continuous-time Markovian jump linear quadratic control. *IEEE Trans. Automatic Control* **35**, 777–788.
[13] Jin, X., Pang, G., Xu, L. and Xu, X. (2025). An approximation to the invariant measure of the limiting diffusion of G/PH/(n + GI) queues in the Halfin–Whitt regime and related asymptotics. *Math. Operat. Res.* **50**, 783–812.
[14] Jin, X., Shen, T. and Su, Z. (2023). Using Stein's method to analyse Euler–Maruyama's approximations of regime-switching jump diffusion processes. *J. Theoret. Prob.* **36**, 1797–1828.
[15] Jones, G. L. (2004). On the Markov chain central limit theorem. *Prob. Surv.* **1**, 299–320.
[16] Klebaner, F. C. (2005). *Introduction to Stochastic Calculus with Applications*. Imperial College Press, London.
[17] Krylov, N. V. (1996). *Lectures on Elliptic and Parabolic Equations in Hölder Spaces* (Graduate Studies Math. **12**). American Mathematical Society, Providence, RI.
[18] Kunwai, K. and Zhu, C. (2020). On Feller and strong Feller properties and irreducibility of regime-switching jump diffusion processes with countable regimes. *Nonlinear Anal. Hybrid Syst.* **38**, 100946.
[19] Li, J. and Xi, F. (2021). Convergence, boundedness, and ergodicity of regime-switching diffusion processes with infinite memory. *Frontiers Math. China* **16**, 499–523.
[20] Liu, W. and Mao, X. (2013). Strong convergence of the stopped Euler–Maruyama method for nonlinear stochastic differential equations. *Appl. Math. Comput.* **223**, 389–400.
[21] Lu, J., Tan, Y. and Xu, L. (2022). Central limit theorem and self-normalized Cramér-type moderate deviation for Euler–Maruyama scheme. *Bernoulli* **28**, 937–964.
[22] Mao, X. (1999). Stability of stochastic differential equations with Markovian switching. *Stoch. Process. Appl.* **79**, 45–67.
[23] Mao, X. and Yuan, C. (2006). *Stochastic Differential Equations with Markovian Switching*. Imperial College Press, London.
[24] Mcleish, D. L. (1974). Dependent central limit theorems and invariance principles. *Ann. Prob.* **2**, 620–628.
[25] Meyn, S. P. and Tweedie, R. L. (1992). Stability of Markovian processes I: Criteria for discrete-time chains. *Adv. Appl. Prob.* **24**, 542–574.
[26] Meyn, S. P. and Tweedie, R. L. (1993). Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes. *Adv. Appl. Prob.* **25**, 518–548.
[27] Nguyen, D. H., Yin, G. and Zhu, C. (2017). Certain properties related to well posedness of switching diffusions. *Stoch. Process. Appl.* **127**, 3135–3158.
[28] Pagès, G. and Panloup, F. (2023). Unajusted Langevin algorithm with multiplicative noise: Total variation and Wasserstein bounds. *Ann. Appl. Prob.* **33**, 726–779.
[29] Sethuraman, S. (2002). A martingale central limit theorem. Available at: https://math.arizona.edu/~sethuram/notes/wi_mart1.pdf.
[30] Shao, J. (2015). Criteria for transience and recurrence of regime-switching diffusion processes. *Electron. J. Prob.* **20**, 1–15.
[31] Shao, J. (2018). Invariant measures and Euler–Maruyama's approximations of state-dependent regime-switching diffusions. *SIAM J. Control Optim.* **56**, 3215–3238.
[32] Skorokhod, A. V. (2009). *Asymptotic Methods in the Theory of Stochastic Differential Equations* (Trans. Math. Monographs **78**). American Mathematical Society, Providence, RI.
[33] Wu, L. (1995). Moderate deviations of dependent random variables related to CLT. *Ann. Prob.* **23**, 420–445.
[34] Wu, L. (2001). Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stoch. Process. Appl.* **91**, 205–238.
[35] Xi, F. (2008). Feller property and exponential ergodicity of diffusion processes with state-dependent switching. *Science China Ser A: Math.* **51**, 329–342.

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press

[36] XI, F (2008). On the stability of jump-diffusions with Markovian switching. *J. Math. Anal. Appl.* **341**, 588–600.
[37] XI, F. (2009). Asymptotic properties of jump-diffusion processes with state-dependent switching. *Stoch. Process. Appl.* **119**, 2198–2221.
[38] XI, F. AND ZHU, C. (2017). On Feller and strong Feller properties and exponential ergodicity of regime-switching jump diffusion processes with countable regimes. *SIAM J. Control Optim.* **55**, 1789–1818.
[39] YIN, G. AND ZHU, C. (2009). *Hybrid Switching Diffusions: Properties and Applications* (Prob. Theory Stoch. Model. **63**). Springer, New York.
[40] YUAN, C. AND MAO, X. (2004). Convergence of the Euler–Maruyama method for stochastic differential equations with Markovian switching. *Math. Comput. Simul.* **64**, 223–235.
[41] ZONG, X., WU, F., YIN, G. AND JIN, Z. (2014). Almost sure and *p*th-moment stability and stabilization of regime-switching jump diffusion systems. *SIAM J. Control Optim.* **52**, 2595–2622.

https://doi.org/10.1017/jpr.2025.10062 Published online by Cambridge University Press