Commun. Math. Anal. Appl.
doi: 10.4208/cmaa.2025-0014

Vol. **4**, No. 4, pp. 495-524
November 2025

Asymptotics in Wasserstein Distance for Empirical Measures of Markov Processes

Feng-Yu Wang*

Center for Applied Mathematics, Tianjin University, Tianjin 300072, China.

Received 20 November 2024; Accepted 20 July 2025

Abstract. In this paper we introduce some recent progresses on the convergence rate in Wasserstein distance for empirical measures of Markov processes. For diffusion processes on compact manifolds possibly with reflecting or killing boundary conditions, the sharp convergence rate as well as renormalization limits are presented in terms of the dimension of the manifold and the spectrum of the generator. For general ergodic Markov processes, explicit estimates are presented for the convergence rate by using a nice reference diffusion process, which are illustrated by some typical examples. Finally, some techniques are introduced to estimate the Wasserstein distance between empirical measures.

AMS subject classifications: 60B05, 60B10

Key words: Empirical measure, Wasserstein distance, Markov process, Riemannian manifold, eigenvalue.

1 Introduction

The empirical measure is a fundamental statistic to estimate the stationary distribution of a Markov process. In this paper, we study the long time behavior of empirical measures for Markov processes under the Wasserstein distance. For two nonnegative functions f,g on a space E, we write $f \lesssim g$ if there exists a constant $c>0$ such that $f \le cg$ holds on E, and write $f \asymp g$ if $f \lesssim g$ and $g \lesssim f$.

Let (M,ρ) be a Polish space, let $\mathscr{P}$ be the space of all probability measures on M. For a Markov process X_t on M, the empirical measure is defined as

*Corresponding author. *Email address:* `wangfy@tju.edu.cn` (F.-Y. Wang)

$$\mu_t := \frac{1}{t}\int_0^t \delta_{X_s} \mathrm{d}s$$

for $t>0$, where δ_{X_s} is the Dirac measure at X_s. We intend to study the long time behavior of μ_t under the p-Wasserstein distance for any $p\in[1,\infty)$,

$$\mathbb{W}_p(\mu_1,\mu_2) := \inf_{\pi\in\mathscr{C}(\mu_1,\mu_2)} \left(\int_{M\times M} \rho^p \mathrm{d}\pi\right)^{\frac{1}{p}}, \quad \mu_1,\mu_2\in\mathscr{P},$$

where $\mathscr{C}(\mu_1,\mu_2)$ is the set of all couplings of μ_1 and μ_2.

To this end, we will consider the ergodicity and quasi-ergodicity cases respectively.

1.1 Ergodicity case

For any $\nu\in\mathscr{P}$, let $\mathbb{E}^\nu$ denote the expectation for the Markov process X_t with initial distribution ν, and denote by $P_t^*\nu$ the law of X_t with initial distribution ν. We have

$$\int_M f \mathrm{d}(P_t^*\nu) = \mathbb{E}^\nu[f(X_t)] = \int_M P_t f(x)\nu(\mathrm{d}x), \quad f\in\mathbf{B}_b(M),$$

where $P_t f(x) := \mathbb{E}^x[f(X_t)]$, and $\mathbf{B}_b(M)$ the class of all bounded measurable functions on M.

If $\mu\in\mathscr{P}$ satisfies $P_t^*\mu=\mu$ for all $t\ge 0$, we call μ an invariant probability measure of the Markov process. If furthermore

$$\lim_{t\to\infty} P_t^*\nu = \mu \quad \text{weakly}$$

holds for any $\nu\in\mathscr{P}$, we call the Markov process ergodic.

In particular, when the Markov process is exponentially ergodic in $L^2(\mu)$, i.e.

$$\|P_t-\mu\|_{L^2(\mu)} \le c\mathrm{e}^{-\lambda t}, \quad t\ge 0$$

holds for some constants $c,\lambda>0$, where $\|\cdot\|_{L^2(\mu)}$ is the operator norm in $L^2(\mu)$ and $\mu(f) := \int_E f\mathrm{d}\mu$ for $f\in L^2(\mu)$, we have

$$\mathbf{V}_f := \int_0^\infty \mu(\hat{f} P_t \hat{f})\mathrm{d}t \in (0,\infty), \quad 0\neq \hat{f} := f-\mu(f), \quad f\in L^2(\mu).$$

Moreover, according to [39], for any $f\in L^2(\mu)$, $\mathbb{P}$-a.s.

$$\lim_{t\to\infty}\mu_t(f) = \mu(f),$$