Eur. J. Stat. 3 (2023) 11
doi: 10.28924/ada/stat.3.11

# Bernstein-von Mises Theorem and Bayes Estimation in Interacting Particle Systems of Diffusions

Jaya P. N. Bishwal

Department of Mathematics and Statistics, University of North Carolina at Charlotte, 376 Fretwell Bldg., 9201 University City Blvd. Charlotte, NC 28223, USA

Correspondence: J.Bishwal@uncc.edu

Abstract. Consistency and asymptotic normality of the Bayes estimator of the drift coefficient of an interacting particles of diffusions are studied. For the Bayes estimator, observations are taken on a fixed time interval $[0, T]$ and asymptotics are studied in the mean-field limit as the number of interacting particles increases. *Interalia*, the Bernstein-von Mises theorem concerning the convergence in the mean-field limit of the posterior distribution, for smooth prior distribution and loss function, to normal distribution is proved.

## 1. Introduction

Finite dimensional parameter estimation in one-dimensional stochastic differential equations from continuous and discrete observations by maximum likelihood and Bayes methods are extensively studied in Bishwal (2008). Parameter estimation for partially observed SDE system which is a factor model of multiple correlated SDEs is studied in Bishwal (2022). Interacting particle systems of diffusions which are generalizations of these factor models, are important for modeling many complex phenomena, see Dawson (1983) and Ligget (1985). Interacting particle systems are useful in constructing particle filter algorithms for finance and computation of credit portfolio losses, see Carmona *et al.* (2009).

McKean (1966) studied a class of Markov processes associated with nonlinear parabolic equations and introduced stochastic systems of interacting particles and the associated non-linaer Markov processes starting from statistical physics to model the dynamics of plasma. McKean (1967) studied propagation of chaos for a class of non-linear parabolic equations. Lot of probabilistic tools have been developed in this context. However, statistics for interacting particle models has not received much attention. Maximum likelihood estimation in interacting particle system of stochastic differential equations was studied in Kasonga (1990) in the mean-field limit where the

Received: 11 Jan 2023.
*Key words and phrases.* stochastic differential equations; mean-field model; large interacting systems; diffusion process; maximum likelihood estimation; Bayes estimation; Bernstein-von Mises theorem; social network.

particles $n \to \infty$. This is in a sense infinite factor models where the factors are correlated. Amorino *et al.* (2022) studied minimum contrast estimation for discretely observed interacting particle systems of McKean-Vlasov type where the particles $n \to \infty$ and the number of discrete time points of observations $N \to \infty$. Sharrock *et al.* (2021) studied parameter estimation for the McKean-Vlasov stochastic differential equation. Della Maestra(2022a) studied nonparametric estimation for McKean-Vlasov models of interacting particle systems. Belomestny *et al.* (2021) studied semi-parametric estimation of McKean-Vlasov SDEs. Geisecke *et al.* (2020) studied inference for large financial systems. Gomes *et al.* (2019) studied parameter estimation for macroscopic pedestrian dynamics models from microscopic data. Baladron *et al.* (2012) studied mean-field description of propagation of chaos in networks of Hugkin-Huxley and FitzHugh-Nagumo neurons. Cardaliaguet and Lehalle (2019) studied mean field game of controls and application to trade crowding. Chazelle (2015a) studied diffusive influence systems. Chazelle (2015b) studied an algorithmic approach to collective behavior. Garnier *et al.* (2017) studied consensus convergence with stochastic effects. Hegselmann and Krause (2002) studied analysis and simulation opinion dynamics models and bounded confidence. Liu and Qiao (2022) studied parameter estimation of path dependent McKean-Vlasov stochastic differential equations. Mahato *et al.* (2018) studied particle method for multi-group pedestrian flow. Molginer and Edelstein-Keshet (1999) studied a non-local model for a swarm. Motsch and Tadmor (2014) showed that heterophilious dynamics enhances concensus. Pavliotis and Zanoni (2022) studied eigenfunction martingale estimators for interacting particle systems and their mean field limit. Suzuki(2005a) studied free energy and self-interacting particles. Suzuki (2005b) studied chemotaxis, reaction and network as models for self-organization. Sznitman (1991) studied propagation of chaos. Toscani (2006) studied kinematic models of opinion formation Wen *et al.*(2016) studied maximum likelihood estimation of McKean-Vlasov stochastic differential equations. Yao *et al.* (2022) studied mean-field nonparametric estimation of interacting particle systems. Benachour *et al.* (1998a, 1998b) studied existence, convergence to invariant probability and propagation of chaos of nonlinear self-stabilizing processes.

In this paper we study Bayes estimation in interacting particle system of stochastic differential equations. Consider the model of $n$ interacting particles of diffusions satisfying the Itô stochastic differential equations

$$dX_j(t) = \sum_{l=1}^{p} \theta_l \mu_{jl}(X(t)) + \sigma_j(X(t))dW_j(t), \quad j = 1, 2, \cdots, n \tag{1.1}$$

where $X(t) = (X_1(t), X_2(t), \cdots, X_n(t))'$ and $(W_j(t); t \geq 0)$, $j = 1, 2, \cdots, n$ are independent Wiener processes. Here $\theta_l(\cdot) \in L^2([0, T], dt), l = 1, \ldots, p$ are unknown functions to be estimated based on observation of the process $X$ in the time interval $[0, T]$. Let $\theta = (\theta_1, \theta_2, \ldots, \theta_p)$ and $\mu_j(x) = (\mu_{j1}(x), \mu_{j2}(x), \ldots, \mu_{jp}(x))'$. The processes $X_j(t), j = 1, 2, \cdots, n$ are observed on $[0, T]$.

The functions $\mu_j, \sigma_j;\ \ j = 1, 2, \cdots, n$ are assumed to be known such that the system has a unique solution.

We need the following assumption and results to prove the main results.

**(A1)** Suppose that $b_{jl} := \mu_{jl}(s)\sigma_j^{-1}(s); j = 1, 2, \cdots, n; l = 1, 2 \ldots, p$ are measurable and adapted processes satisfying

$$\frac{1}{n}\sum_{j=1}^{n}\int_0^t b_{jl}(s)b_{jm}(s)ds \to c_{lm}(t) \text{ a.s.} \quad \text{as } n \to \infty$$

$l, m = 1, 2 \ldots, p$ where $c_{lm}(t)$ are finite and continuous nonrandom functions of $t \in [0, T]$. The limiting matrix $I(t) = (c_{lm}(t))_{l,m=1,2\ldots,p}$ is positive definite, $\delta' I(t)\delta$ is increasing for all $\delta \in \mathbb{R}^p$ and $I(0) = 0$.

In the exchangeable case, (A1) follows from McKean-Vlasov Law of Large Numbers. In particular, (A1) will be satisfied when $\mu_{jl}(X) = \mu_l X_j$ and $\sigma_j(X) = \sigma(X_j)$ which corresponds to the independent replicated sampling on $[0, T]$. See Oelschlager (1984).

We also need the following version of Rebolledo's Central Limit Theorem for Martingales, see Rebolledo (1980):

*Let $M_n, n \in \mathbb{Z}_+$ be a sequence of locally square integrable martingales with $M_n(0) = 0$. Suppose the following condition holds: $\sum_{s\le t} E\{|\Delta M_n(s)|^2 I(|\Delta M_n(s)| > \epsilon)\} \to 0$ for all $t \in [0, T]$, $\epsilon > 0$; and $\langle M_n\rangle(t) \to c(t)$ a.s. for all $t \in [0, T]$, where $c(t)$ is a continuous increasing function with $c(0) = 0$. Then $M_n \to^{\mathcal{D}} M$, a continuous Gaussian martingale with zero mean and covariance function $K(s, t) = c(s \wedge t), s, t \in [0, T]$ where $\Delta M_s = M_s - M_{s-}$ denotes the jump of $M$ at the point $s$.*

## 2. Maximum Likelihood Estimation

The model is given by

$$dX_j(t) = \sum_{l=1}^{p}\theta_l\mu_{jl}(X(t)) + \sigma_j(X(t))dW_j(t), \quad j = 1, 2, \cdots, n \tag{2.1}$$

where $X(t) = (X_1(t), X_2(t), \cdots, X_n(t))'$ and $(W_j(t); t \ge 0)$, $j = 1, 2, \cdots, n$ are independent Wiener processes. Here $\theta = (\theta_1, \theta_2, \ldots, \theta_p)$ is the unknown parameter. The functions $\mu_{jl}, \sigma_j, j = 1, \ldots, n; l = 1, \ldots, p$ are assumed to be known such that there exists a unique solution $X(t)$ to the above SDE.

Our aim is to estimate the parameter $\theta$ based on $n$ particles $X_1(\cdot), X_2(\cdot), \cdots, X_n(\cdot)$ of $X(t)$ on $[0, T]$. We denote this data by $X^{n,T}$

The Radon-Nikodym derivative (likelihood) is given by

$$\begin{aligned} Z_n^\theta(X^{n,T}) &:= \tfrac{dP_\theta}{dP_0}(X^{n,T}) = \exp\Big\{\textstyle\sum_{l=1}^p \theta_l \sum_{j=1}^n \int_0^T \mu_{jl}(X(t))\sigma_j^{-2}(X(t))dX_j(t) \\ &-\tfrac{1}{2}\textstyle\sum_{l=1}^p\sum_{m=1}^p \theta_l\theta_m \sum_{j=1}^n \int_0^T \mu_{jl}(X(t))\sigma_j^{-2}(X(t))\mu_{jm}(X(t))dt\Big\}. \end{aligned} \tag{2.2}$$

The consistency and asymptotic normality of the approximate maximum likelihood estimator are given below from Kasonga (1990):

**Theorem 2.1 (Consistency)** Under (A1), we have

$\hat{\theta}^n \to^P \theta$ as $n \to \infty$.

**Theorem 2.2 (Asymptotic Normality)** Under (A1), we have

$\sqrt{n}(\hat{\theta}^n - \theta) \to^{\mathcal{D}} \mathcal{N}(0, I^{-1}(T))$ as $n \to \infty$ where $I(T)$ is the Fisher information.

## 3. The Bernstein-von Mises Theorem

Suppose that $\Pi$ is a prior probability measure on $(\Theta, \mathcal{D})$, where $\mathcal{D}$ is the $\sigma$-algebra of Borel subsets of $\Theta$. Assume that $\Pi$ has a density $\pi(\cdot)$ w.r.t. the Lebesgue measure and the density is continuous and positive in an open neighbourhood of $\theta_0$.

The posterior density of $\theta$ given in $X^{n,T}$ is given by

$$p(\theta|X^n) := \frac{Z_n^{\theta}(X^{n,T})\pi(\theta)}{\int_{\Theta} Z_n^{\theta}(X^{n,T})\pi(\theta)d\theta}.$$

Let $\tau := n^{1/2}(\theta - \hat{\theta}^n)$. Then the posterior density of $n^{1/2}(\theta - \hat{\theta}^n)$ is given by

$$p^*(\tau|X^{n,T}) := n^{-1/2}p(\hat{\theta}^n + \psi_n^{-1/2}\tau|X^{n,T}).$$

Let

$$\nu_n(\tau) := \frac{dP^n_{\hat{\theta}^n+n^{-1/2}\tau}/dP^n_{\theta_0}}{dP^n_{\hat{\theta}^n}/dP^n_{\theta_0}} = \frac{dP^n_{\hat{\theta}^n+n^{-1/2}\tau}}{dP^n_{\hat{\theta}^n}}, \quad C_n := \int_{-\infty}^{\infty} \nu_n(\tau)\pi(\hat{\theta}^n + n^{-1/2}\tau)d\tau.$$

Clearly

$$p^*(\tau|X^{n,T}) = C_n^{-1}\nu_n(\tau)\pi(\hat{\theta}^n + n^{-1/2}\tau).$$

Let $K(\cdot)$ be a non-negative measurable function satisfying the following two conditions :

**(K1)** There exists a number $\eta$, $0 < \eta < 1$, for which

$$\int_{-\infty}^{\infty} K(\tau)\exp\{-\frac{1}{2}\tau^2(1-\eta)\}d\tau < \infty.$$

**(K2)** For every $\epsilon > 0$ and $\delta > 0$

$$e^{-\epsilon n}\int_{|\tau|>\delta} K(\tau n^{1/2})\pi(\hat{\theta}^n + \tau)d\tau \to 0 \text{ a.s. } [P_{\theta_0}] \text{ as } n \to \infty.$$

We need the following Lemma to prove the Bernstein-von Mises theorem.

**Lemma 3.1** Under the assumptions (A1) and (K1) - (K2),
(i) There exists a $\delta_0 > 0$ such that

$$\lim_{n\to\infty}\int_{|\tau|\le\delta_0 n^{1/2}} K(\tau)\left|\nu_n(\tau)\pi(\hat{\theta}^n + n^{-1/2}\tau) - \pi(\theta_0)\exp(-\frac{1}{2}\tau^2)\right| d\tau = 0 \text{ a.s. } [P_{\theta_0}].$$

(ii) For every $\delta > 0$,

$$\lim_{n\to\infty}\int_{|\tau|\geq\delta n^{1/2}} K(\tau)\left|\nu_n(\tau)\pi(\hat{\theta}^n + n^{-1/2}\tau) - \pi(\theta_0)\exp(-\frac{1}{2}\tau^2)\right| d\tau = 0 \ \text{ a.s. } [P_{\theta_0}].$$

**Proof.** From (2.2), it is easy to check that

$$\log\nu_n(\tau) = -\frac{1}{2}\sum_{l=1}^{p}\sum_{m=1}^{p}\sum_{j=1}^{n}\int_0^T \mu_{jl}(X(t))\sigma_j^{-2}(X(t))\mu_{jm}(X(t))dt.$$

Now (i) follows by an application of the dominated convergence theorem.

For every $\delta > 0$, there exists $\epsilon > 0$ depending on $\delta$ and $\beta$ such that

$$\begin{aligned}
&\int_{|\tau|\geq\delta n^{1/2}} K(\tau)\left|\nu_n(\tau)\pi(\hat{\theta}^n + n^{-1/2}\tau) - \pi(\theta_0)\exp(-\frac{1}{2}\tau^2)\right| d\tau \\
\leq\ & \int_{|\tau|\geq\delta n^{1/2}} K(\tau)\nu_n(\tau)\pi(\hat{\theta}^n + n^{-1/2}\tau)d\tau + \int_{|\tau|\geq\delta n^{1/2}} \pi(\theta_0)\exp(-\frac{1}{2}\tau^2)d\tau \\
\leq\ & e^{-\epsilon n}\int_{|\tau|\geq\delta\psi_n^{1/2}} K(\tau)\pi(\hat{\theta}^n + n^{-1/2}\tau)d\tau + \pi(\theta_0)\int_{|\tau|\geq\delta n^{1/2}} \exp(-\frac{1}{2}\tau^2)d\tau \\
=:\ & F_n + G_n.
\end{aligned}$$

By condition (K2), it follows that $F_n \to 0$ a.s. $[P_{\theta_0]}$ as $n\to\infty$ for every $\delta > 0$. Condition K(1) implies that $G_n \to 0$ as $n\to\infty$. This completes the proof of the Lemma. □

Now we are ready to prove the generalized version of the Bernstein-von Mises theorem for IPS of diffusions.

**Theorem 3.1** Under the assumptions (A1) and (K1) – (K2), we have

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} K(\tau)\left|p^*(\tau|X^{n,T}) - (\frac{1}{2\pi})^{1/2}\exp(-\frac{1}{2}\tau^2)\right| d\tau = 0 \ \text{ a.s. } [P_{\theta_0}].$$

**Proof** From Lemma 3.1, we have

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} K(\tau)\left|\nu_n(\tau)\pi(\hat{\theta}^n + n^{-1/2}\tau) - \pi(\theta_0)\exp(-\frac{1}{2}\tau^2)\right| d\tau = 0 \ \text{ a.s. } [P_{\theta_0}]. \tag{3.1}$$

Putting $K(\tau) = 1$ which trivially satisfies (K1) and (K2), we have

$$C_n = \int_{-\infty}^{\infty}\nu_n(\tau)\pi(\hat{\theta}^n + n^{-1/2}\tau)d\tau \to \pi(\theta_0)\int_{-\infty}^{\infty}\exp(-\frac{1}{2}\tau^2)d\tau \ \text{ a.s. } [P_{\theta_0}]. \tag{3.2}$$

Therefore, by (3.1) and (3.2), we have

$$\begin{aligned}
&\int_{-\infty}^{\infty} K(\tau)\left|p^*(\tau|X^{n,\theta}) - (\frac{1}{2\pi})^{1/2}\exp(-\frac{1}{2}\tau^2)\right| d\tau \\
\leq\ & \int_{-\infty}^{\infty} K(\tau)\left|C_n^{-1}\nu_n(\tau)\pi(\hat{\theta}^n + n^{-1/2}\tau) - C_n^{-1}\pi(\theta_0)\exp(-\frac{1}{2}\tau^2)\right| d\tau \\
&+ \int_{-\infty}^{\infty} K(\tau)\left|C_n^{-1}\pi(\theta_0)\exp(-\frac{1}{2}\tau^2) - (\frac{1}{2\pi})^{1/2}\exp(-\frac{1}{2}\tau^2)\right| d\tau \\
\longrightarrow\ & 0 \ \text{ a.s. } [P_{\theta_0}] \text{ as } n\to\infty. \quad \square
\end{aligned}$$

**Theorem 3.2** Suppose (A1) and $\int_{-\infty}^{\infty} |\theta|^r \pi(\theta) d\theta < \infty$ for some non-negative integer $r$ hold. Then

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} |\tau|^r \left| p^*(\tau|X^n) - (\frac{1}{2\pi})^{1/2} \exp(-\frac{1}{2}\tau^2) \right| d\tau = 0 \quad \text{a.s. } [P_{\theta_0}].$$

**Proof.** For $r = 0$, the verification of (K1) and (K2) is easy and the theorem follows from Theorem 3.1. Suppose $r \geq 1$. Let $K(\tau) = |\tau|^r, \delta > 0$ and $\epsilon > 0$. Using $|a + b|^r \leq 2^{r-1}(|a|^r + |b|^r)$, we have

$$\begin{aligned}
& e^{-\epsilon n} \int_{|\tau|>\delta} K(\tau n^{1/2}) \pi(\hat{\theta}^n + \tau) d\tau \\
\leq \quad & n^{r/2} e^{-\epsilon n} \int_{|\tau - \hat{\theta}^n|>\delta} \pi(\tau) |\tau - \hat{\theta}^n|^r d\tau \\
\leq \quad & 2^{r-1} n^{r/2} e^{-\epsilon n} [\int_{|\tau - \hat{\theta}^n|>\delta} \pi(\tau)|\tau|^r d\tau + \int_{|\tau - \hat{\theta}^n|>\delta} \pi(\tau) |\hat{\theta}^n|^r d\tau] \\
\leq \quad & 2^{r-1} n^{r/2} e^{-\epsilon n} [\int_{-\infty}^{\infty} \pi(\tau)|\tau|^r d\tau + |\hat{\theta}^n|^r] \\
\longrightarrow \quad & 0 \text{ a.s. } [P_{\theta_0}] \text{ as } n \to \infty
\end{aligned}$$

from the strong consistency of $\hat{\theta}^n$ and hypothesis of the theorem. Thus the theorem follows from Theorem 3.1. □

**Remark 3.1** For $r = 0$ in Theorem 3.2, we have

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \left| p^*(\tau|X^n) - (\frac{1}{2\pi})^{1/2} \exp(-\frac{1}{2}\tau^2) \right| d\tau = 0 \quad \text{a.s. } [P_{\theta_0}].$$

This is the classical form of Bernstein-von Mises theorem for interacting SDEs in its simplest form.

As a special case of Theorem 3.2, we obtain

$$E_{\theta_0}[n^{1/2}(\hat{\theta}^n - \theta_0)]^r \to E[\xi^r] \text{ as } n \to \infty \text{ where } \xi \sim \mathcal{N}(0, 1).$$

## 4. Bayes Estimation

As an application of Theorem 3.1, we obtain the asymptotic properties of a regular Bayes estimator of $\theta$. Suppose $l(\theta, \phi)$ is a loss function defined on $\Theta \times \Theta$. Assume that $l(\theta, \phi) = l(|\theta - \phi|) \geq 0$ and $l(\cdot)$ is non decreasing. Suppose that $J$ is a non-negative function on $\mathbb{N}$ and $K(\cdot)$ and $G(\cdot)$ are functions on $\mathbb{R}$ such that

**(B1)** $J(n)l(\tau n^{-1/2}) \leq G(\tau)$ for all $n$,

**(B2)** $J(n)l(\tau n^{-1/2}) \to K(\tau)$ as $n \to \infty$ uniformly on bounded subsets of $\mathbb{R}$.

**(B3)** $\int_{-\infty}^{\infty} K(\tau + s) \exp\{-\frac{1}{2}\tau^2\} d\tau$ has a strict minimum at $s = 0$.

**(B4)** $G(\cdot)$ satisfies (K1) and (K2).

Let

$$B_n(\phi) = \int_{\theta} l(\theta, \phi) p(\theta|X^n) d\theta.$$

A regular Bayes estimator $\tilde{\theta}^n$ based on $X^{n,T}$ is defined as

$$\tilde{\theta}^n := \arg \inf_{\phi \in \Theta} B_n(\phi).$$

Assume that such an estimator exists.

The following Theorem shows that MLE and Bayes estimators are asymptotically equivalent as $n \to \infty$.

**Theorem 4.1** Assume that (A1), (K1) – (K2) and (B1) – (B4) hold. Then we have

(i) $n^{1/2}(\tilde{\theta}^n - \hat{\theta}^n) \to 0$ a.s.-$[P_{\theta_0}]$ as $n \to \infty$,

(ii) $\lim_{n\to\infty} J(n)B_n(\tilde{\theta}^n) = \lim_{n\to\infty} J(n)B_n(\hat{\theta}^n) = (\frac{1}{2\pi})^{1/2} \int_{-\infty}^{\infty} K(\tau)\exp(-\frac{1}{2}\tau^2)d\tau$ a.s. $[P_{\theta_0}]$.

**Proof.** The proof is analogous to Theorem 4.1 in Borwanker *et al.* (1972). We omit the details. □

**Corollary 4.2** Under the assumptions of Theorem 4.1, we have

(i) $\tilde{\theta}^n \to \theta_0$ a.s. $[P_{\theta_0}]$ as $n \to \infty$.

(ii) $n^{1/2}(\tilde{\theta}^n - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$ as $n \to \infty$.

**Proof.** (i) and (ii) follow easily by combining Theorem 4.1 and the strong consistency and asymptotic normality results of the MLE in Kasonga (1990). □

The following theorem shows that Bayes estimators are *locally asymptotically minimax* (LAM) in the Hajek-Le Cam sense, i.e., equality is achieved in the Hajek-Le Cam inequality.

**Theorem 4.3** Under the assumptions of Theorem 4.1, we have

$$\lim_{\delta\to\infty} \lim_{n\to\infty} \sup_{|\theta-\theta_0|<\delta} E\omega\left(I_n^{1/2}(\tilde{\theta}^n - \theta_0)\right) = E\omega(\xi), \quad \mathcal{L}(\xi) = \mathcal{N}(0, 1),$$

where $\omega(\cdot)$ is a loss function as defined in Theorem earlier and $I_n$ is the Fisher information.

**Proof.** The Theorem follows from Theorem III.2.1 in Ibragimov-Has'minskii (1981) since here conditions (N1) - (N4) of the said theorem are satisfied using Lemma 3.1-3.3 and local asymptotic normality (LAN) property obtained in Della Maestra and Hoffmann (2022). □

## 5. Practical Examples

**Example 1: Mean-Field Model**

Let us consider maximum likelihood estimator (MLE) for the simple mean-field model

$$dX_j(t) = \alpha X_j(t)dt - \beta(X_j(t) - \bar{X}_n(t))dt + dW_j(t), \quad X_j(0) = x_j(0), \quad j = 1, 2, \cdots, n \qquad (5.1)$$

where $\bar{X}_n(t)) = n^{-1}\sum_{j=1}^n X_j(t)$, $\beta \neq \alpha$, and $\alpha \neq 0$. The middle term on the right side of (5.1) can be viewed as an interaction among the subsystems which create a tendency for the subsystems to relax towards the center of gravity of the ensemble. Thus the system provides a simple example of a cooperative interaction. Mean-field type models have applications in physics, biology and economics, see Dawson (1983). The case $\beta = 0$ corresponds to sampling independent replications of Ornstein-Uhlenbeck processes on $[0, T]$. Our parameter here is $\theta = (\alpha, \beta)$.

Suppose $\frac{1}{n}\sum_{j=1}^{n} x_j(0) \to \nu_0$ almost surely and $\frac{1}{n}\sum_{j=1}^{n} x_j^2(0) \to \gamma_0^2 + \nu_0^2$ almost surely as $n \to \infty$. Then the estimator $\hat{\theta}^n \to^P \theta$ as $n \to \infty$ and $\sqrt{n}(\hat{\theta}^n - \theta) \to^{\mathcal{D}} \mathcal{N}(0, I^{-1}(T))$ as $n \to \infty$

where

$$I(T) = \begin{pmatrix} A(T) & -B(T) \\ -B(T) & B(T) \end{pmatrix}$$

with

$$A(T) := \frac{\nu_0^2}{2\alpha}(e^{2\alpha T} - 1) + B(T), \quad B(T) := \frac{e^{2(\alpha-\beta)T} - 1}{4(\alpha-\beta)^2} - \frac{T}{2(\alpha-\beta)} + \frac{\gamma_0^2(e^{2(\alpha-\beta)T} - 1)}{2(\alpha-\beta)}.$$

**Example 2: Independent Sampling**

The case $\beta = 0$ corresponds to sampling independent replications of the same process given below:

$$dX_j(t) = \alpha X_j(t)dt + X_j(t)dW_j(t), \quad j = 1, 2, \cdots, n \tag{5.2}$$

In the classical case when $\beta = 0$, the MLE is given by

$$\widehat{\alpha}^n = \frac{\sum_{j=1}^{n} \int_0^T X_j(t)dX_j(t)}{\sum_{j=1}^{n} \int_0^T (X_j(t))^2 dt}.$$

Sampling $n$ independent Ornstein-Uhlenbeck processes on $[0, T]$ and letting $n \to \infty$ give weak consistency and asymptotic normality of the MLE: $\widehat{\alpha}^n \to^P \alpha$ and $\sqrt{n}(\widehat{\alpha}^n - \alpha) \to^{\mathcal{D}} \mathcal{N}(0, \frac{2\alpha}{\nu_0^2(e^{2\alpha T}-1)})$ as $n \to \infty$.

See also Bishwal (2010) for independent sampling case.

**Example 3: Kuramoto Model** (Chazelle *et al.* (2017))

This is the most classical model for synchronization phenomenon in large populations of coupled oscillators such as clapping crowd, a population of fireflies or a system of neurons). The $n$ oscillators are defined by $n$ angles

$$dX_j(t) = \alpha \frac{1}{n}\sum_{j=1}^{n} \sin(X_i(t) - X_j(t)))dt + X_j(t)dW_j(t), \quad j = 1, 2, \cdots, n \tag{5.3}$$

The minimum contrast estimator is consistent as $n \to \infty$ and $N \to \infty$, see Amorino *et al.* (2022).

If $n\Delta_N \to 0$, then the minimum contrast estimator is asymptotically normal, see Amorino *et al.* (2022). Here $\Delta_N = T/N$ is the length of the observation time interval. If $T$ is fixed, we need $n/N \to 0$.

**Example 4: Opinion Dynamics** (Toscani (2006))

$$dX_j(t) = \alpha \frac{1}{n} \sum_{j=1}^{n} \phi(|X_i(t) - X_j(t)|)(X_i(t) - X_j(t))dt + X_j(t)dW_j(t), \quad j = 1, 2, \cdots, n. \quad (5.4)$$

The influence function $\phi$ acts on the "difference of opinions" between agents.

**Example 5: Pearson System** (Forman and Sørensen (2008))

$$dX_j(t) = \alpha \frac{1}{n} \sum_{j=1}^{n} (X_j(t) - \beta X_i(t))dt + \gamma \sqrt{1 + X_j^2(t)} dW_j(t), \quad j = 1, 2, \cdots, n \quad (5.5)$$

Other examples of IPS are Crowd Dynamics (Chazelle *et al.* (2017)), Urban Modeling (Chazelle *et al.* (2017)), Chemotaxis (Suzuki (2005)), Pedestrian Dynamics (Gomes *et al.* (2019)), Collective Behavior (Chazelle (2015)), Molginer Swarm Model (Molgiener and Edelstein-Keshet (1999)), Consensus Dynamics (Hegselmann and Krause (2002)) and Curie-Wiess Model (Dawson (1983)). Curie-Wiess model has quadratic Interaction. The limit of this model is the mean-field model.

## References

[1] C. Amorino, A. Heidari, V. Pilipauskaite, M. Podolskij, Parameter estimation for discretely observed interacting particle systems, *ArXiv.* (2022). https://doi.org/10.48550/arXiv.2208.11965.

[2] J. Baladron, D. Fasoli, O. Faugeras, J. Touboul, Mean-field description of propagation of chaos in networks of Hugkin-Huxley and FitzHugh-Nagumo neurons, J. Math. Neurosci. 2 (2012) 1-50.

[3] D. Belomestny, V. Pilipauskaite, M. Podloskij, Semiparametric estimation of McKean-Vlasov SDEs, ArXiv. (2021). https://doi.org/10.48550/arXiv.2107.00539.

[4] S. Benachour, B. Roynette, D. Talay, P. Vallois, Nonlinear self-stabilizing processes-I : Existence, invariant probability, propagation of chaos, Stoch. Process. Appl. 75 (1998a) 173-201.

[5] S. Benachour, B. Roynette, P. Vallois, Nonlinear self-stabilizing processes-II : Convergence to invariant probability, Stoch. Process. Appl. 75 (1998b) 203-224.

[6] J.P.N. Bishwal, Parameter Estimation in Stochastic Differential Equations, Lecture Notes in Mathematics, 1923, Springer-Verlag. (2008).

[7] J.P.N. Bishwal, : Maximum likelihood estimation in Skorohod stochastic differential equations, Proc. Amer. Math. Soc. 138 (2010) 1471-1478.

[8] J.P.N. Bishwal, Milstein approximation of posterior density of diffusions, Int. J. Pure Appl. Math. 68 (2011a) 403-414.

[9] J.P.N. Bishwal, Estimation in interacting diffusions: continuous and discrete sampling, Appl. Mathe. 2 (2011b) 1154-1158.

[10] J.P.N. Bishwal, Benstein-von Mises theorem and small noise Bayesian asymptotics for parabolic stochastic partial differential equations, Theory Stoch. Processes. 23 (2018) 6-17.

[11] J.P.N. Bishwal, Berry-Esseen bounds of approximate Bayes estimators for the discretely observed Ornstein-Uhlenbeck process, Asian J. Stat. Sci. 1 (2021) 83-122.

[12] J.P.N. Bishwal, Parameter Estimation in Stochastic Volatility Models, Springer Nature, Cham. (2022).

[13] P. Cardaliaguet, C. Lehalle, Mean field game of controls and application to trade crowding, Math. Financ. Econ. 12 (2019) 335-363.

[14] R. Carmona, J.P. Fouque, D. Vestal, Interacting particle systems for the computation of rare credit portfolio losses, Finance Stoch. 13 (2009) 613-633.

[15] B. Chazelle, Diffusive influence systems, SIAM J. Comput. 44 (2015a) 1403-1442.

[16] B. Chazelle, An algorithmic approach to collective behavior, J. Stat. Phys. 158 (2015b) 514-548.

[17] B. Chazelle, Q. Jiu, Q. Li, C. Wang, Well-posedness of the limiting equation of a noisy consensus model in opinion dynamics, J. Diff. Equ. 263 (2017) 365-397.

[18] D. Dawson, Critical dynamics and fluctuations for a mean-field model of cooperative behavior, J. Stat. Phys. 31 (1983) 29-85.

[19] L. Della Maestra, M. Hoffmann, Nonparametric estimation for interacting particle systems: McKean-Vlasov models, Prob. Theory Related Fields. 182 (2022a) 551-613.

[20] L. Della Maestra, M. Hoffmann, The LAN property for McKean-Vlasov models in a mean-field regime, ArXiv. (2022b). https://doi.org/10.48550/arXiv.2205.05932.

[21] J.L. Forman, M. Sørensen, The Pearson diffusions: a class of statistically tractable diffusion processes, Scandinavian J. Stat. 35 (2008) 438-465.

[22] J. Garnier, G. Papanicolaou, T. Yang, Consensus convergence with stochastic effects, Vietnam J. Math. 45 (2017) 51-75.

[23] K. Geisecke, G. Schwenkler, J.A. Sirigano, Inference for large financial systems, Math. Finance. 30 (2020) 3-46.

[24] S.N. Gomes, A.M. Start, M.T. Wolfram, Parameter estimation for macroscopic pedestrian dynamics models from microscopic data, SIAM J. Appl. Math. Appl. 79 (2019) 1475-1500.

[25] R. Hegselmann, U. Krause, Opinion dynamics and bounded confidence: models, analysis and simulation, J. Artif. Soc. Soc. Sim. 5 (2002) 1-33.

[26] R.A. Kasonga, Maximum likelihood theory for large interacting systems, SIAM J. Appl. Math. 50 (1990) 865-875.

[27] T. Ligget, Interacting Particle Systems, Springer-Verlag, New York. (1985).

[28] M. Liu, H. Qiao, Parameter estimation of path dependent McKean-Vlasov stochastic differential equations, Acta Math. Sci. 42 (2022) 876-886.

[29] N.K. Mahato, A. Klar, K.S. Tiwari, Particle method for multi-group pedestrian flow, Appl. Math. Model. 53 (2018) 447-461.

[30] H.P. McKean, A class of Markov processes associated with nonlinear parabolic equations, Proc. Nat. Acad. Sci. 56 (1966) 1907-1911.

[31] H.P. McKean, Propagation of chaos for a class of non-linear parabolic equations, Stochastic Differential Equations, Lecture Series in Differential Equations, Session 7, Catholic University, 41-57. (1967).

[32] A. Molginer, L. Edelstein-Keshet, A non-local model for a swarm, J. Math. Biol. 38 (1999) 534-570.

[33] S. Motsch, E. Tadmor, Heterophilious dynamics enhances concensus, SIAM Rev. 56 (2014) 577-621.

[34] K. Oelschlager, A martingale approach to the law of large numbers for weakly interacting particle stochastic processes, Ann. Prob. 12 (1984) 458-479.

[35] G.A. Pavliotis, A. Zanoni, Eigenfunction martingale estimators for interacting particle systems and their mean field limit, ArXiv. (2022). https://doi.org/10.48550/arXiv.2112.04870.

[36] R. Rebolledo, Central limit theorems for local martingales, Zeit. Wahr. Verw. Gebiete. 51 (1980) 269-286.

[37] L. Sharrock, N. Kantas, P. Parpas, G.A. Pavliotis, Parameter estimation for the McKean-Vlasov stochastic differential equation, ArXiv. (2021). https://doi.org/10.48550/arXiv.2106.13751.

[38] T. Suzuki, Free Energy and Self-interacting Particles, Birkhauser, Boston. (2005a).

[39] T. Suzuki, Chemotaxis, Reaction, Network: mathematics for Self-Organization, World Scientific, Singapore. (2005b).

[40] A.S. Sznitman, Topics in Propagation of Chaos, Lecture Notes in Mathematics **1464**, Springer, Berlin, pp. 165-251. (1991).

[41] G. Toscani, Kinematic models of opinion formation, Commun. Math. Sci. 4 (2006) 481-496.

[42] J. Wen, X. Wang, S. Mao, X. Xiao, Maximum likelihood estimation of McKean-Vlasov stochastic differential equations and its applications, Appl. Math. Comp. 274 (2016) 237-246.

[43] R. Yao, X. Chen, Y. Yang, Mean-field nonparametric estimation of interacting particle systems, Proc. Mach. Learn. Res. 178 (2022) 1-34.