On goodness-of-fit testing for volatility in McKean–Vlasov models

Akram Heidari [*] Mark Podolskij [†]

Abstract

This paper develops a statistical framework for goodness-of-fit testing of volatility functions in McKean–Vlasov stochastic differential equations, which describe large systems of interacting particles with distribution-dependent dynamics. While integrated volatility estimation in classical SDEs is now well established, formal model validation and goodness-of-fit testing for McKean–Vlasov systems remain largely unexplored, particularly in regimes with both large particle limits and high-frequency sampling. We propose a test statistic based on discrete observations of particle systems, analysed in a joint regime where both the number of particles and the sampling frequency increase. The estimators involved are proven to be consistent, and the test statistic is shown to satisfy a central limit theorem, converging in distribution to a centred Gaussian law.

Keywords: Asymptotic distribution, goodness-of-fit testing, high-frequency data, McKean–Vlasov diffusions, mean-field models, nonlinear diffusion, volatility specification.

AMS subject classifications: 62E20, 62G10, 60F15, 60H10.

Contents

[*]Department of Mathematics, University of Luxembourg, Maison du Nombre, 6 Avenue de la Fonte, 4364 Esch-sur-Alzette, Luxembourg, Email: akram.heidari@uni.lu

[†]Department of Mathematics, University of Luxembourg, Maison du Nombre, 6 Avenue de la Fonte, 4364 Esch-sur-Alzette, Luxembourg, Email: mark.podolskij@uni.lu

1 Introduction

McKean–Vlasov stochastic differential equations (SDEs) have gained increasing prominence as a widely used modeling framework for complex systems consisting of large populations of interacting agents. Unlike classical SDEs, the dynamics of each particle depend not only on its individual state but also on the statistical distribution of the entire system. This distinctive feature makes McKean–Vlasov models particularly well-suited for capturing systemic interactions and emergent behavior in diverse applications ranging from economics and finance to physics and engineering [8, 18, 22, 29].

In the context of financial modeling, McKean–Vlasov dynamics have been used to study systemic risk [19], mean-field interactions in portfolio management [9], and the evolution of agent-based financial markets. More broadly, their use has expanded to problems involving optimal control, equilibrium analysis, and the dynamics of large-scale interacting systems, as surveyed in [10, 24]. These models allow for a nuanced representation of endogenous feedback effects in financial systems, where local decisions and aggregate dynamics are tightly coupled through distributional dependencies.

Given the growing importance of McKean–Vlasov models in applications, there is an increasing need for statistical methods that can rigorously validate their structural components, particularly the volatility function, which governs the system's stochastic fluctuations. While classical diffusion models often rely on volatility functions depending solely on the state or time variables, this assumption breaks down in systems where interaction between agents drives the evolution. In such settings, volatility may depend on the entire population distribution, and any mis-specification can significantly affect downstream predictions and risk measures.

Despite the importance of model validation, the literature on goodness-of-fit testing for McKean–Vlasov equations remains scarce. Existing parametric testing procedures for volatility, such as those developed for standard SDEs [2, 14, 16, 17] and fractional SDEs [27], are tailored to non-interacting systems only. Parametric and non-parametric estimation methods for McKean-Vlasov SDEs have been investigated in [3, 4, 6, 7, 11, 12, 13, 15, 20, 21, 23, 25, 26, 28, 30], but they mostly focus on the drift function. This creates a critical methodological gap: there is currently no general method for assessing whether a given volatility structure adequately captures the behavior of a McKean–Vlasov system based on empirical data.

We address this gap by developing a statistical testing procedure for McKean–Vlasov particle systems. Our focus lies on constructing a goodness-of-fit test for the volatility function under high-frequency and large-population asymptotics. In this work, we consider a system of N interacting particles $(X^i)^{i=1,\dots,N}$, defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, and evolving over a fixed time interval $[0,T]$. The particles are modeled as independent copies of a non-linear process satisfying the McKean–Vlasov SDE:

$$\begin{cases} dX_t^i = b\big(X_t^i, \mu_t\big)dt + a\big(X_t^i, \mu_t\big)dW_t^i \qquad i = 1, \dots, N, \quad t \in [0,T] \\ \mathrm{Law}\big(X_0^1, \dots, X_0^N\big) := \mu_0 \times \dots \times \mu_0 \end{cases} \tag{1.1}$$

where the processes $(W_t^i)_{t\in[0,T]}$, $i = 1, \dots, N$, are independent Brownian motions, and μ_t denotes the law of X_t^i. The model coefficients

$$b : \mathbb{R} \times \mathcal{P}_2 \to \mathbb{R}, \qquad a : \mathbb{R} \times \mathcal{P}_2 \to \mathbb{R}$$

are measurable functions that depend on the current state and the current distribution of the solution. Here, $\mathcal{P}_2$ denotes the space of probability measures on $\mathbb{R}$ with finite second moments. This space is equipped with the Wasserstein 2-metric, defined by

$$W_2(\mu,\nu) = \Big(\inf_{m\in\Gamma(\mu,\nu)} \int_{\mathbb{R}^2} |x-y|^2 m(dx,dy)\Big)^{\frac{1}{2}},$$

where $\Gamma(\mu,\nu)$ denotes the set of probability measures on $\mathbb{R}^2$ with marginals μ and ν.

Our primary objective is to develop a goodness-of-fit testing framework for the volatility function $a(x,\mu)$, based on discrete-time observations of the system. We consider observations of the form

$$\Big(X^i_{t_j}\Big)^{i=1,\dots,N}_{j=1,\dots,n}, \qquad \text{with } t_j = Tj/n, \tag{1.2}$$

and study the regime where both the observation frequency increases ($\Delta_n := T/n \to 0$) and the number of particles grows ($N \to \infty$), with a fixed time horizon $T > 0$.

The main statistical goal of this work is to develop a goodness-of-fit test for the volatility function $a(x,\mu)$, under the null hypothesis that it belongs to a given parametric family. To this end, we introduce a test statistic based on discrete-time observations of an interacting particle system, constructed through an appropriate distance measure. We prove consistency of the underlying estimators and establish a central limit theorem for the proposed statistic. This yields a testing procedure that maintains the correct asymptotic level and is consistent against any fixed alternative. The main methodological challenge lies in the measure dependence of $a(x,\mu)$, which generates non-linear and path-dependent effects that render standard techniques from classical SDE analysis inapplicable.

To the best of our knowledge, this work provides the first rigorous statistical testing framework for volatility structures in McKean–Vlasov models based on discrete-time observations of interacting particle systems. By combining high-frequency asymptotics with the mean-field structure, our approach extends the scope of model validation to complex stochastic systems with distribution-dependent dynamics and lays the theoretical groundwork for hypothesis testing in nonlinear diffusion models.

The structure of the paper is as follows. Section 2 introduces the framework and sets out the standing assumptions for model (1.1). In Section 3, we develop the proposed goodness-of-fit testing procedure, detailing the construction of the test statistic. Section 4 presents the main theoretical results, establishing the consistency and asymptotic normality of the estimators, and deriving the limiting distribution of the test statistic under the null hypothesis. All proofs and supporting technical arguments are collected in Section 5.

2 Assumptions

In this section, we introduce the main assumptions associated with the model (1.1), which are satisfied by a wide class of stochastic volatility models.

Assumption 1. *For all $k \geq 1$,*

$$\int_{\mathbb{R}} |x|^k \mu_0(dx) \leq C_k.$$

The following assumption ensures the existence and uniqueness of a strong solution to (1.1), guaranteeing well-posedness of the model.

Assumption 2. *The drift and diffusion coefficients satisfy Lipschitz continuity and a linear growth condition. Specifically, there exists a constant $C > 0$ such that for all $(x,\mu),(y,\lambda) \in \mathbb{R} \times \mathcal{P}_2$:*

$$\begin{aligned}|b(x,\mu) - b(y,\lambda)| + |a(x,\mu) - a(y,\lambda)| &\le C\big(|x-y| + W_2(\mu,\lambda)\big),\\ |b(x,\mu)|^2 + |a(x,\mu)|^2 &\le C\big(1 + |x|^2 + W_2^2(\mu,\delta_0)\big),\end{aligned}$$

Regarding the regularity of the diffusion function $a : \mathbb{R} \times \mathcal{P}_2 \to \mathbb{R}$, we adopt the notion of linear differentiability, which is widely used in the literature on McKean–Vlasov equations and mean-field games to characterize the smoothness of the mapping $\mu \mapsto a(x,\mu)$ from $\mathcal{P}_2 \to \mathbb{R}$. This concept is particularly well-suited to our framework, and we refer the reader to Section 2 of [15] and the references therein for a detailed exposition.

Definition 1. *A mapping $f : \mathcal{P}_2 \to \mathbb{R}$ is said to have a linear functional derivative, if there exists $\partial_\mu f : \mathbb{R} \times \mathcal{P}_2 \to \mathbb{R}$ such that*

$$f(\mu) - f(\mu') = \int_0^1 \int_{\mathbb{R}} \partial_\mu f(y, \lambda\mu + (1-\lambda)\mu')(\mu - \mu')(dy)d\lambda$$

for every $(\mu,\mu') \in \mathcal{P}_2$ and $\partial_\mu f$ satisfies additional smoothness properties, which will be provided in the following assumption.

Assumption 3. *The map $\mu \mapsto a(x,\mu)$ admits a functional derivative in the sense of Definition 1. Furthermore, there exists a constant $C > 0$ such that for all $(x,\mu),(x',\mu') \in \mathbb{R} \times \mathcal{P}_2$,*

$$\big|\partial_\mu a(x,y,\mu) - \partial_\mu a(x',y',\mu')\big| \le C(|x-x'| + |y-y'| + W_2(\mu,\mu')).$$

Additionally it holds that

$$\begin{aligned}|\partial_y \partial_\mu a(x,y,\mu)| &\le C \qquad \forall (x,y,\mu),\\ |\partial_y \partial_\mu a(x,y,\mu) - \partial_y \partial_\mu a(x,y,\mu')| &\le C W_2(\mu,\mu').\end{aligned}$$

Finally, the function $a(x,t) := a(x,\mu_t)$ is in $C^{2,1}(\mathbb{R} \times [0,T])$.

We note that Assumption 3 ensures, in particular, that the process $(a(X_t^i,t))_{t\in[0,T]}$ is a continuous semimartingale. Indeed, by Itô's formula one obtains

$$\begin{aligned}a(X_t^i,t) = a(X_0,0) &+ \int_0^t \Big(\partial_t a(X_s^i,s) + b(X_s^i,\mu_s)\,\partial_x a(X_s^i,s) + \tfrac{1}{2}\,\partial_{xx} a(X_s^i,s)\,a^2(X_s^i,s)\Big)\,ds\\ &+ \int_0^t \partial_x a(X_s^i,s) a(X_s^i,s)\,dW_s^i. \end{aligned} \tag{2.1}$$

This representation plays a key role in deriving error estimates for high-frequency statistics of the process $(X_t^i)_{t\in[0,T]}$ (cf. [5]).

3 Testing parametric hypotheses for the volatility

In this section, we develop a goodness-of-fit testing framework for the volatility structure in McKean–Vlasov SDEs. Our goal is to assess whether a given parametric form of the volatility function is consistent with the observed behavior of a discretely sampled particle system. We begin by formally stating the parametric hypothesis. Then, we introduce our proposed test statistic $\widehat{S}^N$ and describe its construction under high-frequency and large-population asymptotics.

Let

$$a_1^2, \ldots, a_d^2 : \mathbb{R} \times \mathcal{P}_2 \to \mathbb{R}^+$$

be a collection of known functions, assumed to be linearly independent and to satisfy the same regularity conditions as the volatility function $a(x, \mu)$. Our objective is to test whether the squared volatility function $a^2(x, \mu)$ belongs to the linear span of the basis functions $a_1^2, \ldots, a_d^2$. More precisely, the null hypothesis is given by

$$H_0 : L := \min_{(\lambda_1,\ldots,\lambda_d)\in\mathbb{R}^d} \int_0^T \int_{\mathbb{R}} \Big(a^2(x, \mu_t) - \sum_{k=1}^d \lambda_k a_k^2(x, \mu_t) \Big)^2 \mu_t(dx)\, dt = 0, \tag{3.1}$$

with the alternative hypothesis $H_1 : L > 0$. Here, μ_t denotes the distribution of the underlying particle system $(X_t^1, \ldots, X_t^N)$, which is not directly observable. In practice, we approximate μ_t by the empirical distribution of the observed particles. The criterion in (3.1) serves as a natural foundation for our test construction, since it directly measures model discrepancy in a Hilbert space framework. Moreover, it admits a discretized version that can be readily implemented using particle observations.

Remark 3.1. The distance measure L introduced in (3.1) is conceptually related to the distance proposed in [16, 17] for classical SDEs, although the two approaches differ in essential aspects. To clarify this, recall that [16, 17] study the one-dimensional diffusion model

$$dX_t = b(X_t)\, dt + a(X_t)\, dW_t,$$

observed at discrete time points t_j. They introduce the *random* distance measure

$$M^2 := \min_{(\lambda_1,\ldots,\lambda_d)\in\mathbb{R}^d} \int_0^T \Big(a^2(X_t) - \sum_{k=1}^d \lambda_k a_k^2(X_t) \Big)^2 dt,$$

and consider the hypothesis test $H_0 : M^2 = 0$ versus $H_1 : M^2 > 0$. A key point is that, under high-frequency observations of a single trajectory of $(X_t)_{t\in[0,T]}$, one can only verify whether

$$a^2(X_t) = \sum_{k=1}^d \lambda_k a_k^2(X_t)$$

holds for some choice of λ_k, *along the realized path* $(X_t(\omega))_{t\in[0,T]}$. There is no possibility to test this identity outside the observed trajectory, i.e., for $x \notin (X_t(\omega))_{t\in[0,T]}$.

In contrast, in our setting with N independent trajectories as in (1.1), the condition $L = 0$ entails that

$$a^2(x, \mu_t) = \sum_{k=1}^d \lambda_k a_k^2(x, \mu_t)$$

for some λ_k, holding for μ_t-almost every $x \in \mathbb{R}$, $t \in [0, T]$. Nevertheless, this identity is testable only with respect to the distributions μ_t of the observed particles (X_t^i), and not for arbitrary distributions. □

Standard arguments (cf. [1]) show that this L^2-distance admits the closed-form expression:

$$L = g(\Gamma_1, \dots, \Gamma_d, \mathcal{B}, \Lambda) = \mathcal{B} - (\Gamma_1, \dots, \Gamma_d)\Lambda^{-1}(\Gamma_1, \dots, \Gamma_d)^\top \tag{3.2}$$

where the quantities $\mathcal{B}, \Gamma_1, \dots, \Gamma_d$ and the matrix $\Lambda = (\Lambda_{k,l})_{1\le k,l\le d}$ are given by

$$\begin{aligned} \mathcal{B} &:= \int_0^T \int_{\mathbb{R}} a^4(x, \mu_t)\mu_t(dx)dt, \\ \Gamma_k &:= \int_0^T \int_{\mathbb{R}} a_k^2(x, \mu_t)a^2(x, \mu_t)\mu_t(dx)dt, \qquad k = 1, \dots, d \\ \Lambda_{k,l} &:= \int_0^T \int_{\mathbb{R}} a_k^2(x, \mu_t)a_l^2(x, \mu_t)\mu_t(dx)dt, \qquad k, l = 1, \dots, d \end{aligned} \tag{3.3}$$

In order to construct a consistent estimator of L in (3.2), we replace the quantities in (3.3) with their empirical counterparts, based on the discrete-time observations introduced in (1.2). Accordingly, we define the following estimators:

$$\begin{aligned} \widehat{\mathcal{B}} &:= \frac{1}{3N\Delta_n} \sum_{i=1}^N \sum_{j=1}^n |X^i_{t_{j+1}} - X^i_{t_j}|^4 \\ \widehat{\Gamma}_k &:= \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^n a_k^2(X^i_{t_j}, \mu^N_{t_j})|X^i_{t_{j+1}} - X^i_{t_j}|^2 \qquad k = 1, \dots, d \\ \widehat{\Lambda}_{k,l} &:= \frac{\Delta_n}{N} \sum_{i=1}^N \sum_{j=1}^n a_k^2(X^i_{t_j}, \mu^N_{t_j})a_l^2(X^i_{t_j}, \mu^N_{t_j}) \qquad k, l = 1, \dots, d \end{aligned} \tag{3.4}$$

(For simplicity of notation we suppress the dependence of estimators on n and N). Here, μ_t^N denotes the empirical measure of the system at time t, i.e.

$$\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{X^i_t}.$$

We then introduce the following test statistic as the empirical analogue of the quantity (3.2):

$$\widehat{S}^N = g(\widehat{\Gamma}_1, \dots, \widehat{\Gamma}_d, \widehat{\mathcal{B}}, \widehat{\Lambda}) = \widehat{\mathcal{B}} - \widehat{\mathbf{\Gamma}}^\top \widehat{\Lambda}^{-1} \widehat{\mathbf{\Gamma}}$$

where the vector $\widehat{\mathbf{\Gamma}} = (\widehat{\Gamma}_1, \dots, \widehat{\Gamma}_d)^\top$ is defined by the components in (3.4). This statistic captures the deviation from the null hypothesis and forms the basis of our goodness-of-fit test.

In the following section, we briefly summarise the key statistical properties of our proposed estimator $\widehat{S}^N$. Specifically, building on the consistent estimation of $\mathcal{B}$, Γ_k, and $\Lambda_{k,l}$ by $\widehat{\mathcal{B}}$, $\widehat{\Gamma}_k$, and $\widehat{\Lambda}_{k,l}$ (as established in Theorems 4.1 and 4.2, respectively), we show that $\widehat{S}^N$ is a consistent estimator of L as $N \to \infty$ and $\Delta_n \to 0$, and provide the associated central limit theorem.

4 Main results

In this section, we present the core theoretical contributions of this paper, focusing on the asymptotic properties of our proposed estimators and their associated test statistics. We begin by demonstrating the consistency and convergence rates of our estimators. Specifically, Theorem 4.1 establishes the consistent approximation of the limiting matrix Λ by its empirical counterpart $\widehat{\Lambda}$. Similarly, Theorem 4.2 provides stochastic expansions for the quantities $\widehat{\mathcal{B}}$ and $\widehat{\Gamma}_k$ (for $k = 1, \ldots, d$), which are essential for deriving the limiting distribution of our test statistic. Corollary 4.3 then establishes the joint asymptotic normality of these key components. These results together characterize the limiting distribution and its asymptotic covariance structure. Finally, combining these foundational results, we present the main asymptotic normality result for our proposed estimator $\widehat{S}^N$.

In what follows, we introduce two theorems that provide stochastic expansions for the quantities $\widehat{\Lambda}$, $\widehat{\mathcal{B}}$, and $\widehat{\Gamma}_k$. Throughout the sequel, we will frequently use the notation $o_{\mathbb{P}}(1)$ to denote terms that converge to 0 in probability.

Theorem 4.1. *Assume that Assumptions 1-3 hold and $N\Delta_n^2 \to 0$. Then*

$$\sqrt{N}\big(\widehat{\Lambda} - \Lambda\big) = \sqrt{N} M_\Lambda + o_{\mathbb{P}}(1)$$

where M_Λ is a $(d \times d)$-matrix with elements

$$\begin{aligned} M_{\Lambda,k,l} &:= \frac{1}{N}\sum_{i=1}^{N}\big(Z^i_{\Lambda,k,l} - \mathbb{E}(Z^1_{\Lambda,k,l})\big) \\ Z^i_{\Lambda,k,l} &:= \int_0^T a_k^2(X_s^i, \mu_s) a_l^2(X_s^i, \mu_s) ds \qquad\qquad k, l = 1, \ldots, d. \end{aligned}$$

Theorem 4.2. *Assume that Assumptions 1-3 hold and $N\Delta_n^2 \to 0$. Then*

$$\begin{aligned} \sqrt{N}\Big(\widehat{\Gamma}_k - \Gamma_k\Big) &= \sqrt{N} M_k + o_{\mathbb{P}}(1) \\ \sqrt{N}\Big(\widehat{\mathcal{B}} - \mathcal{B}\Big) &= \sqrt{N} M_{\mathcal{B}} + o_{\mathbb{P}}(1) \end{aligned}$$

with

$$\begin{aligned} M_k &:= \frac{1}{N}\sum_{i=1}^{N}\big(Z^i_k - \mathbb{E}(Z^1_k)\big), \qquad Z^i_k := \int_0^T a_k^2(X_s^i, \mu_s) a^2(X_s^i, \mu_s) ds \\ M_{\mathcal{B}} &:= \frac{1}{N}\sum_{i=1}^{N}\big(Z^i_{\mathcal{B}} - \mathbb{E}(Z^1_{\mathcal{B}})\big), \qquad Z^i_{\mathcal{B}} := \int_0^T a^4(X_s^i, \mu_s) ds \end{aligned}$$

Building upon the individual asymptotic properties established in Theorems 4.1 and 4.2, we are now in a position to derive the joint asymptotic distribution of the involved components $(M_1, \ldots, M_d, M_{\mathcal{B}}, M_\Lambda)$. The following statement is a simple consequence of the standard central limit theorem and the δ-method.

Corollary 4.3. *Assume that Assumptions 1-3 are satisfied and $N\Delta_n^2 \to 0$.*

(i) It holds that

$$\widehat{Z} := \sqrt{N}\begin{pmatrix} M_1 \\ \vdots \\ M_d \\ M_{\mathcal{B}} \\ \text{vec}\,(M_\Lambda)_{k,l} \end{pmatrix} \xrightarrow{\mathcal{L}} \mathbf{Z}^*, \qquad \mathbf{Z}^* \stackrel{d}{=} \mathcal{N}_{d^2+d+1}(0, \Sigma)$$

where the components of the covariance matrix Σ *are given by*

$$\Sigma_{p,q} = \mathrm{Cov}\left(\widehat{Z}_p, \widehat{Z}_q\right).$$

(ii) It holds that

$$\sqrt{N}(\widehat{S}^N - L) \xrightarrow{\mathcal{L}} \mathcal{G} \sim \mathcal{N}(0, \tau^2)$$

where the asymptotic variance τ^2 *is defined as*

$$\tau^2 = \nabla g^\top(\Gamma_1, \dots, \Gamma_d, \mathcal{B}, \Lambda)\Sigma\nabla g(\Gamma_1, \dots, \Gamma_d, \mathcal{B}, \Lambda).$$

The asymptotic result of Corollary 4.3 forms the theoretical basis of our goodness-of-fit test. Under the null hypothesis $H_0 : L = 0$, it yields

$$\sqrt{N}\,\widehat{S}^N \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tau^2).$$

Suppose we can construct a consistent estimator $\widehat{\tau}^2$ of the asymptotic variance τ^2, that is,

$$\widehat{\tau}^2 \xrightarrow{\mathbb{P}} \tau^2.$$

Then, for a given significance level $\alpha \in (0, 1)$, the null hypothesis $H_0 : L = 0$ is rejected whenever

$$\frac{\sqrt{N}\,\widehat{S}^N}{\widehat{\tau}} > z_{1-\alpha},$$

where $z_{1-\alpha}$ denotes the $(1-\alpha)$-quantile of the standard normal distribution. By construction, this test attains the correct asymptotic size α. Moreover, under the alternative $H_1 : L > 0$, we have $\sqrt{N}\,\widehat{S}^N \xrightarrow{\mathbb{P}} +\infty$, which ensures that the procedure is consistent against any fixed alternative.

In the final step we construct a consistent estimator $\widehat{\tau}^2$ of τ^2. We introduce the vector

$$\widehat{V}^i := \left(\widehat{Z}^i_1, \dots, \widehat{Z}^i_d,\ \widehat{Z}^i_{\mathcal{B}},\ \mathrm{vec}\,(\widehat{Z}^i_\Lambda)_{k,l}\right)^\top$$

with the estimators given by

$$\widehat{Z}^i_k := \sum_{j=1}^n a_k^2(X^i_{t_j}, \mu^N_{t_j})|X^i_{t_{j+1}} - X^i_{t_j}|^2 \qquad k = 1, \dots, d$$

$$\widehat{Z}^i_{\mathcal{B}} := \frac{1}{3\Delta_n}\sum_{j=1}^n |X^i_{t_{j+1}} - X^i_{t_j}|^4$$

$$(\widehat{Z}^i_\Lambda)_{k,l} := \Delta_n \sum_{j=1}^n a_k^2(X^i_{t_j}, \mu^N_{t_j}) a_l^2(X^i_{t_j}, \mu^N_{t_j}) \qquad k, l = 1, \dots, d$$

Then the empirical covariance estimator of the covariance matrix Σ is defined as

$$\widehat{\Sigma} = \frac{1}{N}\sum_{i=1}^N \left(\widehat{V}^i - \overline{V}\right)\left(\widehat{V}^i - \overline{V}\right)^\top, \qquad \text{with } \overline{V} = \frac{1}{N}\sum_{i=1}^N \widehat{V}^i.$$

Applying similar methods as in Theorem 4.1, we conclude that $\widehat{\Sigma} \xrightarrow{\mathbb{P}} \Sigma$ and consequently the estimator

$$\widehat{\tau}^2 := \nabla g^\top\left(\widehat{\Gamma_1}, \dots, \widehat{\Gamma_d}, \widehat{\mathcal{B}}, \widehat{\Lambda}\right)\widehat{\Sigma}\nabla g\left(\widehat{\Gamma_1}, \dots, \widehat{\Gamma_d}, \widehat{\mathcal{B}}, \widehat{\Lambda}\right)$$

satisfies $\widehat{\tau}^2 \xrightarrow{\mathbb{P}} \tau^2$.

Remark 4.4. For practical applications, the null hypothesis will almost never hold exactly. It is therefore natural to ask how well the linear span of the functions $a_1^2, \dots, a_d^2$ can approximate the true squared volatility coefficient a^2. The distance measure L is not ideal in this context, since its numerical size is difficult to interpret. A more convenient criterion was introduced in [27] for fractional diffusion models and, in our setting, takes the form

$$G := \frac{L}{\mathcal{B}}.$$

In contrast to L, the statistic G enjoys the appealing property $G \in [0,1]$, which follows directly from Pythagoras' theorem. This normalization allows deviations from the null hypothesis to be expressed in relative terms rather than in absolute units. Moreover, for any fixed $\delta \in (0,1)$, one can test

$$H_0: \; G \in [0,\delta] \qquad \text{vs.} \qquad H_1: \; G \in (\delta, 1],$$

and the asymptotic normality of G follows directly from Corollary 4.3. □

5 Proofs

As a preliminary step, we recall a collection of moment bounds, adapted from [4], that will serve as a foundation for establishing the main results of this paper.

Lemma 5.1. *Assumptions 1–3 hold. Then, for any $p \geq 1$, there exists a constant $C > 0$ such that the following bounds hold uniformly over all particles $i \in \{1, \dots, N\}$, for all $N \in \mathbb{N}$, and for all times $t \in [0,T]$:*

(i) $\sup_{t\in[0,T]} \mathbb{E}[|X_t^i|^p] < C$, *and moreover,* $\sup_{t\in[0,T]} \mathbb{E}[W_p^q(\mu_t^N, \delta_0)] < C$ *for* $p \leq q$.

(ii) $\mathbb{E}[|X_{t_{j+1}}^i - X_{t_j}^i|^p] \leq C\Delta_n^{p/2}$.

(iii) $\mathbb{E}[W_2^2(\mu_t^N, \mu_t)] \leq CN^{-1}$.

5.1 Proof of Theorem 4.1

We restrict our attention to the case $d = 1$ and set $f := a_1^4$. The extension to higher dimensions $d > 1$ is straightforward and involves only additional notational complexity. We consider the following decomposition

$$\begin{aligned}
\widehat{\Lambda} - \Lambda &= \frac{\Delta_n}{N}\sum_{i=1}^N\sum_{j=1}^n f(X_{t_j}^i, \mu_{t_j}^N) - \int_0^T\int_{\mathbb{R}} f(x,\mu_s)\mu_s(dx)ds \\
&= \frac{1}{N}\sum_{i=1}^N\int_0^T f(X_s^i,\mu_s)ds - \int_0^T\int_{\mathbb{R}} f(x,\mu_s)\mu_s(dx)ds \\
&+ \frac{\Delta_n}{N}\sum_{i=1}^N\sum_{j=1}^n f(X_{t_j}^i, \mu_{t_j}) - \frac{1}{N}\sum_{i=1}^N\int_0^T f(X_s^i,\mu_s)ds \\
&+ \frac{\Delta_n}{N}\sum_{i=1}^N\sum_{j=1}^n \Big[f(X_{t_j}^i, \mu_{t_j}^N) - f(X_{t_j}^i, \mu_{t_j})\Big] \\
&=: M_\Lambda + H_{(1)} + H_{(2)}
\end{aligned}$$

The term M_Λ corresponds to the deviation between the empirical mean and its deterministic counterpart. We thus need to show that $\sqrt{N}(|H_{(1)}| + |H_{(2)}|) = o_{\mathbb{P}}(1)$.

First of all, we note that the term $H_{(1)}$ corresponds to the Riemann sum approximation error associated with stochastic process $f_t := f(X_t^i, \mu_t)$. According to (2.1), this stochastic process is a continuous Itô semimartingale, we know from [5, Section 7 and 8] that

$$\Delta_n \sum_{j=1}^n f(X_{t_j}^i, \mu_{t_j}) - \int_0^T f(X_s^i, \mu_s) ds = \Delta_n A_i^n$$

with

$$\frac{1}{N} \sum_{i=1}^N A_i^n = O_{\mathbb{P}}(1).$$

In other words, $\sqrt{N} H_{(1)} = O_{\mathbb{P}}(\Delta_n \sqrt{N})$ and the latter is negligible as $N\Delta_n^2 \to 0$.

Now, we focus on the term $H_{(2)}$. Here we use the notion of the linear functional derivative and apply it to the function f:

$$H_{(2)} = H_{(2.1)} + H_{(2.2)},$$

where the terms $H_{(2.1)}$ and $H_{(2.2)}$ are defined via

$$\begin{aligned}
H_{(2.1)} &:= \frac{\Delta_n}{N} \sum_{i=1}^N \sum_{j=1}^n \int_{\mathbb{R}} \partial_\mu f(X_{t_j}^i, y, \mu_{t_j})(\mu_{t_j}^N - \mu_{t_j})(dy) \\
H_{(2.2)} &:= \frac{\Delta_n}{N} \sum_{i=1}^N \sum_{j=1}^n \int_0^1 \int_{\mathbb{R}} \left\{ \partial_\mu f(X_{t_j}^i, y, \lambda \mu_{t_j}^N + (1-\lambda)\mu_{t_j}) - \partial_\mu f(X_{t_j}^i, y, \mu_{t_j}) \right\} \\
&\quad \times (\mu_{t_j}^N - \mu_{t_j})(dy) d\lambda
\end{aligned}$$

For the term $H_{(2.1)}$ we obtain the identity

$$H_{(2.1)} = \frac{\Delta_n}{N^2} \sum_{i,k=1}^N \sum_{j=1}^n \left(\partial_\mu f(X_{t_j}^i, X_{t_j}^k, \mu_{t_j}) - \int_{\mathbb{R}} \partial_\mu f(X_{t_j}^i, y, \mu_{t_j}) \mu_{t_j}(dy) \right).$$

If we write $H_{(2.1)} = N^{-2} \sum_{i,k=1}^N R_n(i,k)$ and use the arguments from Hoeffding decomposition for U-statistics to compute the variance of $H_{(2.1)}$, we immediately conclude that

$$H_{(2.1)} = O_{\mathbb{P}}(1/N).$$

Now, we move on to handling the term $H_{(2.2)}$. For a function $g : \mathbb{R} \to \mathbb{R}$ with bounded derivative we know that

$$\left| \int_{\mathbb{R}} g(x)(\mu - \nu)(dx) \right| \le \|g'\|_\infty W_1(\mu, \nu).$$

Applying this inequality, Lemma 5.1(iii) and Assumption 3, we obtain for the term $H_{(2.2)}$:

$$\begin{aligned}
|H_{(2.2)}| &\le \frac{\Delta_n}{N} \sum_{i=1}^N \sum_{j=1}^n \int_0^1 \sup_{y \in \mathbb{R}} \left| \partial_y \partial_\mu f(X_{t_j}^i, y, \lambda \mu_{t_j}^N + (1-\lambda)\mu_{t_j}) - \partial_y \partial_\mu f(X_{t_j}^i, y, \mu_{t_j}) \right| d\lambda \\
&\quad \times W_1(\mu_{t_j}^N, \mu_{t_j}) \\
&\le C \Delta_n \sum_{j=1}^n W_2(\mu_{t_j}^N, \mu_{t_j}) W_1(\mu_{t_j}^N, \mu_{t_j}) = O_{\mathbb{P}}(N^{-1}).
\end{aligned}$$

Hence, we conclude that $H_{(2)} = O_{\mathbb{P}}(N^{-1})$, which completes the proof of Theorem 4.1.

5.2 Proof of Theorem 4.2

We start with the term $\widehat{\Gamma}_k$. We obtain the decomposition

$$\widehat{\Gamma}_k - \Gamma_k = \Gamma_{k.1} + \Gamma_{k.2}$$

where the terms $\Gamma_{k.1}$ and $\Gamma_{k.2}$ are defined as

$$\Gamma_{k.1} := \frac{\Delta_n}{N} \sum_{i=1}^{N} \sum_{j=1}^{n} a_k^2(X_{t_j}^i, \mu_{t_j}^N) a^2(X_{t_j}^i, \mu_{t_j}^N) - \int_0^T \int_{\mathbb{R}} a_k^2(x, \mu_t) a^2(x, \mu_t) \mu_t(dx) dt,$$

$$\Gamma_{k.2} := \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{n} a_k^2(X_{t_j}^i, \mu_{t_j}^N) |X_{t_{j+1}}^i - X_{t_j}^i|^2 - \frac{\Delta_n}{N} \sum_{i=1}^{N} \sum_{j=1}^{n} a_k^2(X_{t_j}^i, \mu_{t_j}^N) a^2(X_{t_j}^i, \mu_{t_j}^N).$$

Analogously to the proof of Theorem 4.1 we conclude that

$$\Gamma_{k.1} = M_k + o_{\mathbb{P}}(N^{-1/2}).$$

Applying the methods of [5, Section 7 and 8], we deduce the decomposition

$$\Gamma_{k.2} = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{n} a_k^2(X_{t_j}^i, \mu_{t_j}^N) a^2(X_{t_j}^i, \mu_{t_j}^N) \left((W_{t_{j+1}}^i - W_{t_j}^i)^2 - \Delta_n \right) + O_{\mathbb{P}}(N^{-1}).$$

By martingale methods we immediately obtain that

$$\mathrm{var}(\Gamma_{k.2}) = N^{-1} \Delta_n.$$

This implies that $\sqrt{N}\Gamma_{k.1} = o_{\mathbb{P}}(1)$.

The term $\widehat{\mathcal{B}}$ is handled the same way as $\widehat{\Gamma}_k$, which completes the proof of Theorem 4.2.

□

References

[1] Achieser, N. J. (1956). *Theory of approximation.* Dover Publications Inc., New York.

[2] Ait-Sahalia, Y. (1996). Testing continuous-time models of the spot interest rate. *Review of Financial Studies*, 9, 385–426.

[3] Amorino, C., Belomestny, D. Pilipauskaitė, V., Podolskij, M. & Zhou, S.-Y. (2024). Polynomial rates via deconvolution for nonparametric estimation in McKean-Vlasov SDEs. To appear in *Probability Theory and Related Fields.* arXiv preprint arxiv: 2208.11965.

[4] Amorino, C., Heidari, A., Pilipauskaitė, V. &, Podolskij, M. (2023). Parameter estimation of discretely observed interacting particle systems. *Stochastic processes and Their Applications*, 163, 350–386.

[5] Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J., Podolskij, M., & Shephard, N. (2006). A central limit theorem for realised power and bipower variations of continuous semimartingales. *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev*, Springer.

[6] Belomestny, D. Pilipauskaitė, V. & Podolskij, M. (2022). Semiparametric estimation of McKean-Vlasov SDEs. *Annales de l'Institut Henri Poincaré,* 59(1), 79–96.

[7] Bishwal, J.P.N. (2011). Estimation in interacting diffusions: Continuous and discrete sampling. *Applied Mathematics*, 2(9), 1154–1158.

[8] Carmona, R., & Delarue, F. (2018). *Probabilistic theory of mean field games with applications II: Mean field games with common noise and stochastic control.* Springer.

[9] Carmona, R., & Webster, B. (2017). A mean field game approach to competitive electricity markets. *Applied Mathematical Finance*, 24(2), 150-170.

[10] Carmona, R., & Webster, A. (2020). *Mean field games in financial economics.* Springer.

[11] Chen, X. (2021). Maximum likelihood estimation of potential energy in interacting particle systems from single-trajectory data. *Electronic Communications in Probability*, 26, 1–13.

[12] Comte, F. & Genon-Catalot, V. (2024): Parametric inference for ergodic McKean-Vlasov stochastic differential equations. *Bernoulli*, 30(3), 1971–1997.

[13] Comte, F, Genon-Catalot, V. & Laredo, C. (2024). Nonparametric moment method for McKean-Vlasov stochastic differential equations. *Working paper.*

[14] Corradi, V. & White, H. (1999). Specification tests for the variance of a diffusion. *J. Time Ser. Anal.* 20, 253–270.

[15] Della Maestra, L., & Hoffmann, M. (2022). Nonparametric estimation for interacting particle systems: McKean–Vlasov models. *Probability Theory and Related Fields*, 182(1), 551-613.

[16] Dette, H. & Podolskij, M. (2008). Testing the parametric form of the volatility in continuous time diffusion models—a stochastic process approach. *Journal of Econometrics*, 143, 56–73.

[17] Dette, H., Podolskij, M. & Vetter, M. (2006). Estimation of integrated volatility in continuous-time financial models with applications to goodness-of-fit testing. *Scandinavian Journal of Statistics*, 33, 259–278.

[18] B. Djehiche, F. Gozzi, G. Zanco & M. Zanella (2022): Optimal portfolio choice with path dependent benchmarked labor income: a mean field model. *Stochastic Processes and Their Applications*, 145, 48–85.

[19] Fouque, J. P., & Sun, X. (2013). Systemic risk with feedback. *Quantitative Finance*, 13(12), 1845-1856.

[20] Genon-Catalot, V. & Larédo, C. (2021). Parametric inference for small variance and long time horizon McKean-Vlasov diffusion models. *Electronic Journal of Statistics,* 15(2), 5811–5854.

[21] Genon-Catalot, V. & Larédo, C. (2021). Probabilistic properties and parametric inference of small variance nonlinear self-stabilizing stochastic differential equations. *Stochastic Processes and their Applications,* 142, 513–548.

[22] Jourdain, B., Malrieu, F., & Woyczynski, W. A. (2009). Poisson approximation for the solutions of McKean-Vlasov equations. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 45(4), 865-882.

[23] Kasonga, R.A. (1990). Maximum likelihood theory for large interacting systems. *SIAM Journal on Applied Mathematics,* 50(3), 865–875.

[24] Lacker, H. (2020). *Mean field games and control in economics.* Cambridge University Press.

[25] Liu, M. & Qiao, H. (2022). Parameter estimation of path-dependent McKean-Vlasov stochastic differential equations. *Acta Mathematica Scientia,* 42, 876–886.

[26] Nickl, R., Pavliotis, G.A. & Ray, K. (2025). Bayesian nonparametric inference in McKean-Vlasov models. *Annals of Statistics,* 53(1), 170–193.

[27] Podolskij, M. & Wasmuth, K. (2013). Goodness-of-fit testing for fractional diffusions. *Statistical Inference for Stochastic Processes*, 16(2), 147–159.

[28] Sharrock, L., Kantas, N., Parpas, P. & Pavliotis, G.A. (2021). Online parameter estimation for the McKean-Vlasov stochastic differential equation. *Stochastic Processes and their Applications,* 162, 481–546.

[29] Sznitman, A.-S. (1991). *Topics in propagation of chaos.* In Ecole d'eté de probabilités de Saint-Flour XIX-1989 (pp. 165-251). Springer, Berlin, Heidelberg.

[30] Wen, J. Wang, X., Mao, S. & Xiao, X. (2016). Maximum likelihood estimation of McKean-Vlasov stochastic differential equation and its application. *Applied Mathematics and Computation,* 274, 237-246.