Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 132 (2021) 135–163

stochastic processes and their applications

www.elsevier.com/locate/spa

# On the identifiability of interaction functions in systems of interacting particles

Zhongyang Li[a], Fei Lu[b,c,*], Mauro Maggioni[b,c], Sui Tang[d], Cheng Zhang[e]

[a] *Department of Mathematics, University of Connecticut, Storrs, CT, 06269, USA*
[b] *Department of Mathematics, Johns Hopkins University, Baltimore, MD, 21218, USA*
[c] *Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, MD, 21218, USA*
[d] *Department of Mathematics, University of California Santa Barbara, Isla Vista, 93117, USA*
[e] *Department of Mathematics, University of Rochester, Rochester, NY 14627, USA*

Received 18 February 2020; received in revised form 1 October 2020; accepted 14 October 2020
Available online 3 November 2020

**Abstract**

We address a fundamental issue in the nonparametric inference for systems of interacting particles: the identifiability of the interaction functions. We prove that the interaction functions are identifiable for a class of first-order stochastic systems, including linear systems with general initial laws and nonlinear systems with stationary distributions. We show that a coercivity condition is sufficient for identifiability and becomes necessary when the number of particles approaches infinity. The coercivity is equivalent to the strict positivity of related integral operators, which we prove by showing that their integral kernels are strictly positive definite by using Müntz type theorems.
© 2020 Elsevier B.V. All rights reserved.

*Keywords:* Interacting particle systems; Nonparametric regression; Positive-definite kernel; Positive operator; Identifiability

## 1. Introduction

Dynamical systems of interacting particles or agents are widely used in many areas in science and engineering, such as physics [12], biology [3], social science [6,31]; we refer to [7,36] for reviews. With the recent advancement of technology in data collection and

* Corresponding author at: Department of Mathematics, Johns Hopkins University, Baltimore, MD, 21218, USA.
*E-mail addresses:* zhongyang.li@uconn.edu (Z. Li), feilu@math.jhu.edu (F. Lu), mauromaggionijhu@icloud.com (M. Maggioni), suitang@math.ucsb.edu (S. Tang), czhang77@ur.rochester.edu (C. Zhang).

https://doi.org/10.1016/j.spa.2020.10.005
0304-4149/© 2020 Elsevier B.V. All rights reserved.

computation, inference of such systems from data has attracted increasing attention [8,17,19]. In general, such systems are high-dimensional and there is no natural parametric form for the interaction laws, so their inference tends to be statistically and computationally infeasible due to the curse of dimensionality. When the particles interact according to a function that depends only on pairwise distances, one only needs to estimate such interaction function, opening the possibility of statistically and computationally efficient inference techniques [5,26,28]. However, a fundamental challenge arises: the interaction function may be non-identifiable, because its values are under-determined from the observation data consisting of trajectories. To ensure the identifiability of the interaction function, a coercivity condition is introduced in [5,26,28]. In this study, we show that the coercivity condition is sufficient for the identifiability, that it becomes necessary when the number of particles goes to infinity, and that it holds true for linear systems and for a class of three-particle nonlinear systems with a stationary distribution.

More precisely, we consider a first-order stochastic gradient system of interacting particles in the form

$$\begin{cases} d\boldsymbol{X}_i^t = \dfrac{1}{N} \displaystyle\sum_{1\leqslant j\leqslant N, j\neq i} \phi(|\boldsymbol{X}_j^t - \boldsymbol{X}_i^t|) \dfrac{\boldsymbol{X}_j^t - \boldsymbol{X}_i^t}{|\boldsymbol{X}_j^t - \boldsymbol{X}_i^t|} dt + \sigma d\boldsymbol{B}_i^t, \quad \text{for } i = 1, \dots, N, \\ \boldsymbol{X}^0 \sim \mu_0, \end{cases} \tag{1.1}$$

where $\boldsymbol{X}_i^t \in \mathbb{R}^d$ represents the position of particle $i$ at time $t$, $\{\boldsymbol{B}_i^t\}_{i=1}^N$ are independent Brownian motions in $\mathbb{R}^d$ representing the random environment, $|\cdot|$ denotes the Euclidean norm, $\sigma > 0$ is the strength of the noise. Without loss of generality, we assume $\sigma = 1$ in (1.1). The function $\phi : \mathbb{R}^+ = [0, \infty) \to \mathbb{R}$ models the pairwise interaction between particles, which is referred to as the *interaction function*. We assume that the initial condition $\boldsymbol{X}^0$ has an exchangeable absolutely continuous distribution $\mu_0$ on the state space $\mathbb{R}^{dN}$, that is, the joint distribution of $(\boldsymbol{X}_{i_1}^0, \dots, \boldsymbol{X}_{i_N}^0)$ is $\mu_0$ for any ordering of the index set $\{i_1, \dots, i_N\} = \{1, \dots, N\}$. As a consequence, combining with the fact that the system is equivalent under permutations of the indices of the agents, the distribution of $\boldsymbol{X}^t$ is exchangeable for any $t \in [0, T]$.

We consider the identifiability of the interaction function $\phi$ from many independent trajectories on a time interval $[0, T]$, denoted by $\{\boldsymbol{X}^{[0,T],m}\}_{m=1}^M$, in the likelihood-based nonparametric inference setting. We focus on the case of infinitely many trajectories (i.e. $M = \infty$) for our analysis. Clearly, the function space for inference must depend on the information from the process defined by the system (1.1), and in particular we can hope to estimate the interaction function only on the interval explored by pairwise distances. A natural choice is the space $L^2(\overline{\rho}_T)$ (or a subspace thereof), where $\overline{\rho}_T$ is the average-in-time distribution of all the pairwise distances $\{|\boldsymbol{X}_i^t - \boldsymbol{X}_j^t|, t \in [0, T]\}_{i,j=1}^N$. By the exchangeability of the distribution of $\boldsymbol{X}^t$, the distribution $\rho_t$ of $|\boldsymbol{X}_i^t - \boldsymbol{X}_j^t|$ is the same for all $(i, j)$ pairs (which is why we may abuse notation and avoid writing $\rho_{t,i,j}$), thus $\overline{\rho}_T$ can be written as

$$\overline{\rho}_T(dr) := \frac{1}{T}\int_0^T \rho_t(dr)dt, \qquad \text{with} \qquad \rho_t(dr) := \mathbb{E}[\delta(|\boldsymbol{X}_i^t - \boldsymbol{X}_j^t| \in dr)], \tag{1.2}$$

In other words, $\overline{\rho}_T$ is the average of the measures $\{\rho_t, t \in [0, T]\}$. Note that $\overline{\rho}_T$ depends on both the initial distribution and the interaction function $\phi$.

We define the identifiability of the interaction function as follows.

**Definition 1.1** (*Identifiability*)**.** The interaction function $\phi$ of the system (1.1), which defines the process $\boldsymbol{X}^{[0,T]}$, is said to be identifiable in a linear subspace $\mathcal{H}$ of $L^2(\overline{\rho}_T)$, if it is the unique maximizer of the expectation of the log-likelihood ratio of the process (see $\mathbb{E}\mathcal{E}_{\boldsymbol{X}^{[0,T]}}(\cdot)$ in (2.5)).

In practice, the above identifiability requires a unique global maximizer for the expectation of the log-likelihood ratio (a functional on the high- or infinite-dimensional subspace $\mathcal{H}$, see Section 2.1 for details), which is difficult to verify. The following coercivity condition, introduced in [5,26–28], provides an appealing alternative because it can be numerically verified from data. It ensures the uniqueness of the maximizer of the empirical likelihood ratio on finite dimensional hypothesis spaces by ensuring the Hessian to be strictly negative definite.

**Definition 1.2** (*Coercivity Condition on a Time Interval*)**.** The system (1.1) on $[0, T]$ is said to satisfy a coercivity condition on a finite-dimensional linear subspace $\mathcal{H} \subset L^2(\bar{\rho}_T)$ with $\bar{\rho}_T$ defined in (1.2) if

$$c_{\mathcal{H},T} := \inf_{h\in\mathcal{H},\, \|h\|_{L^2(\overline{\rho}_T)}=1} \frac{1}{T}\int_0^T \mathbb{E}[h(|\boldsymbol{r}_{12}^t|)h(|\boldsymbol{r}_{13}^t|)\frac{\langle \boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t\rangle}{|\boldsymbol{r}_{12}^t||\boldsymbol{r}_{13}^t|}]dt > 0, \tag{1.3}$$

where $\boldsymbol{r}_{ij}^t = \boldsymbol{X}_i^t - \boldsymbol{X}_j^t$. When $\mathcal{H} \subseteq L^2(\overline{\rho}_T)$ is infinite-dimensional, we say that the system satisfies a coercivity condition on $\mathcal{H}$ if the coercivity condition holds on each finite dimensional linear subspace of $\mathcal{H}$.

The coercivity condition on $\mathcal{H}$ defined here is slightly different than the previous one in [26–28], which requires $c_{\mathcal{H},T} + \frac{1}{N-2} > 0$ and ensures the uniform concavity of the expectation of the log-likelihood ratio. This new definition has the advantage of being independent of $N$, and requires a positive coercivity constant $c_{\mathcal{H},T}$ only on each finite-dimensional hypothesis space $\mathcal{H}$, rather than on any compact set of $L^2(\bar{\rho}_T)$, making it suitable for studying the mean-field limit when $N \to \infty$. This new definition also highlights the dependence on the joint distribution of $(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{12}^t)$ only, and the connection with positive integral operators (see Section 2.2). A drawback is that it can be slightly more restrictive than the previous one for finite $N$, with such difference vanishing as $N \to \infty$ (see Remark 2.3 for details).

We show that the coercivity condition is sufficient for the identifiability, and it holds for certain classes of interaction functions, including $\phi(r) = r^{2\beta-1}$ with $\beta \in [\frac{1}{2}, 1]$:

**Theorem 1.3.** *Consider the system (1.1) on $[0, T]$ with interaction function $\phi$, and the average distribution of the pairwise distances between particles $\overline{\rho}_T$ as in (1.2).*

*(a). The interaction function $\phi$ is identifiable in a linear subspace $\mathcal{H}$ of $L^2(\overline{\rho}_T)$ if the coercivity condition holds on $\mathcal{H}$.*

*(b). The coercivity condition holds on $L^2(\overline{\rho}_T)$ if $\phi(r) = \theta r$, i.e. when the system is linear, and the initial distribution $\mu_0$ of $\boldsymbol{X}^0$ is a non-degenerate exchangeable Gaussian.*

*(c). The coercivity condition holds on $L^2(\overline{\rho}_T)$ for nonlinear systems with three particles and with the following interaction functions and initial distribution:*

1. *the interaction function $\phi$ is of the form*

$$\phi(r) = \Phi'(r), \qquad \text{where } \Phi(r) := ar^{2\beta} + \Phi_0(r), \quad a > 0,\ \beta \in [\frac{1}{2}, 1], \tag{1.4}$$

*where $\Phi_0 \in C^2(\mathbb{R}^+, \mathbb{R})$ satisfies that $f(u, v) := \Phi_0(|u - v|) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is a negative definite function and that $\lim_{r\to\infty} \Phi(r) = +\infty$;*

2. *the joint probability density of $(\boldsymbol{X}_1^0 - \boldsymbol{X}_2^0, \boldsymbol{X}_1^0 - \boldsymbol{X}_3^0)$ is (with $Z$ being a normalizing constant)*

$$p(u, v) = \frac{1}{Z} e^{-\frac{1}{3}[\Phi(|u|)+\Phi(|v|)+\Phi(|u-v|)]}, \tag{1.5}$$

*i.e., an invariant density of the process $(\boldsymbol{X}_1^t - \boldsymbol{X}_2^t, \boldsymbol{X}_1^t - \boldsymbol{X}_3^t)$.*

Part (a) of Theorem 1.3 is proved in Proposition 2.1. In addition to being sufficient for identifiability, the coercivity condition also becomes necessary when $N$, the number of particles in the system, is infinity. In particular, the coercivity constant $c_{\mathcal{H},T}$ in (1.3) is independent of $N$ and it depends only on the distribution of the process $(\boldsymbol{X}_1^t - \boldsymbol{X}_2^t, \boldsymbol{X}_1^t - \boldsymbol{X}_3^t)$. We prove Part (b) in Theorem 3.6 and Part (c) in Theorem 4.1.

We show that the coercivity condition is equivalent to the strict positivity of an integral operator arising from the expectation in (1.3) (see Section 2.2). Then, to prove the strict positivity of the operator, we show that its integral kernel is strictly positive-definite, by introducing a series representation of the integral kernel and resorting to Müntz-type theorems for the completeness of polynomials in $L^2(\overline{\rho_T})$ (see Section 3). In particular, in the treatment of nonlinear systems, we develop a "comparison to a Gaussian kernel" technique (Sections 4.2–4.3) to prove the strictly positive-definiteness of integral kernels.

This study serves as a starting point towards understanding the identifiability of the interaction function for particle or agent systems. While providing a full characterization for linear systems, i.e., the coercivity condition holds for general initial distributions, Theorem 1.3 provides limited results for nonlinear systems, covering only stationary initial distribution for systems with $N = 3$ particles and with polynomial dominated interaction functions. The constraint $N = 3$ arises because our series representation of the integral kernel is based on the explicit expression of the joint distribution of $(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t)$, which is currently unknown to us when $N > 3$, albeit we are hopeful to eventually be able to remove this constraint in future work. The constraint of stationary initial distribution may be removable by perturbation-type arguments. Future directions of research include, to name just a few, first-order nonlinear systems with more general interaction functions that are regular [19,28] or singular [24,25], second-order systems and systems with multiple types of particles or agents [26], and mean-field equations [10,20,30].

Positive-definite integral kernels play an increasingly prominent role in many applications in science, in particular in statistical learning theory and in reproducing kernel Hilbert space (RKHS) representations [11,13,35]. As a by-product, our results lead to a new class of positive-definite integral kernels from particle systems, and our technique of comparison to a Gaussian kernel may be of broader interest, for example in establishing identifiability of statistical learning problems.

The organization of the paper is as follows: we summarize in Table 1 the frequently used notations. In Section 2, we introduce the coercivity conditions in inference, and establish the connections between identifiability, the coercivity condition and positive integral operators. In Section 3 we prove the coercivity condition for linear systems and Section 4 is devoted to a class of three-particle nonlinear systems with stationary distributions. We list in the Appendix the preliminaries, such as properties of positive-definite kernels, a Müntz-type theorem on the half-line, and a stationary measure for gradient systems.

## 2. The coercivity conditions and strictly positive integral operators

In the context of likelihood-based nonparametric inference of the interaction function, we show that the coercivity condition is sufficient for identifiability, and it is equivalent to the strict positive-definiteness of an integral operator. Also, we introduce a coercivity condition at a single time, which suggests that the interaction function can be identifiable from many samples at a single time.

In vector format, we can write the system (1.1) as

$$d\boldsymbol{X}^t = -\nabla J_\phi(\boldsymbol{X}^t)dt + d\boldsymbol{B}^t \tag{2.1}$$

**Table 1**
Notations.

| Notation | Description |
|---|---|
| $\phi$ and $\varphi$ | The true and, respectively, a generic interaction function |
| $\Phi$ | The true interaction potential, such that $\Phi'(r) = \phi(r)$ as in (1.4) |
| $\boldsymbol{X}_i^t$ and $\boldsymbol{X}^{[0,T]}$ | Position of the $i$th particle at time $t$ and, resp., trajectory on $[0, T]$ |
| $\boldsymbol{r}_{ij}^t = \boldsymbol{X}_i^t - \boldsymbol{X}_j^t$ | Position difference from particle $j$ to particle $i$ at time $t$ |
| $\rho_t$ and $\overline{\rho}_T$ | Probability distribution of $\lvert\boldsymbol{r}_{12}^t\rvert$ and, resp., its average on $[0, T]$ in (1.2) |
| $L^2(\rho_t)$ and $L^2(\overline{\rho}_T)$ | The function spaces $L^2(\mathbb{R}^+, \rho_t)$ and $L^2(\mathbb{R}^+, \overline{\rho}_T)$ |
| $p_t(u, v)$ and $p(u, v)$ | The joint density of $(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t)$ and, resp., the stationary density, as in (1.5) |

where $\boldsymbol{X}^t := (\boldsymbol{X}_i^t)_{i=1}^N \in \mathbb{R}^{Nd}$, and the potential function $J_\phi : \mathbb{R}^{Nd} \to \mathbb{R}$ is

$$J_\phi(\boldsymbol{x}) = \frac{1}{2N} \sum_{i,j=1, j\neq i}^{N} \Phi(|\boldsymbol{x}_i - \boldsymbol{x}_j|), \quad \boldsymbol{x} \in \mathbb{R}^{Nd}, \text{ with } \Phi(r) = \int_0^r \phi(s)ds. \tag{2.2}$$

Note that the constant $\Phi(0)$ does not affect the gradient system, and it can be arbitrary. We assume for simplicity that $\Phi(0) = 0$.

Since the pairwise potential $\Phi$ in (1.4) is $C^2(\mathbb{R}^+)$ and $\lim_{r\to\infty} \Phi(r) = \infty$, the drift term in (2.1) is locally Lipschitz and the total potential $J_\phi$ is a Lyapunov function for the system. Thus, a global solution exists (see e.g. [22, Theorem 1.3]). For the existence and properties of the solutions in systems with singular potentials, we refer to [1,24,25] and the reference therein.

### 2.1. Identifiability and the coercivity condition

Consider the likelihood-based inference of the interaction function $\phi$ from observation data consisting of many independent trajectories $\{\boldsymbol{X}^{[0,T],m}\}_{m=1}^M$. The maximum likelihood estimator (MLE) is a maximizer of the log-likelihood ratio of these trajectories over a hypothesis space $\mathcal{H}$:

$$\widehat{\phi}_{\mathcal{H},M} := \arg\max_{\varphi\in\mathcal{H}} \mathcal{E}^M(\varphi), \qquad \text{with } \mathcal{E}^M(\varphi) := \frac{1}{M}\sum_{m=1}^M \mathcal{E}_{\boldsymbol{X}^{[0,T],m}}(\varphi)$$

where $\mathcal{E}_{\boldsymbol{X}^{[0,T],m}}(\varphi)$ denotes the average log-likelihood ratio of the trajectory $\boldsymbol{X}^{[0,T],m}$, which is given, by the Girsanov theorem (see e.g. [23, Section 1.1.4] and [21, Section 3.5]), by

$$\mathcal{E}_{\boldsymbol{X}^{[0,T]}}(\varphi) = -\frac{1}{2TN}\int_0^T \left(|\nabla J_\varphi(\boldsymbol{X}^t)|^2 dt + 2\langle \nabla J_\varphi(\boldsymbol{X}^t), d\boldsymbol{X}^t\rangle\right). \tag{2.3}$$

Note that $\mathcal{E}^M(\cdot)$ is a quadratic functional. When $\mathcal{H}$ is a finite dimensional linear space, $\mathcal{E}^M(\cdot)$ becomes a quadratic function on $\mathcal{H}$, and an estimator $\widehat{\phi}_{\mathcal{H},M}$ can be obtained by solving a least squares problem (which has multiple solutions if the matrix of the normal equations is singular.

A key assumption in the definition of MLE is the uniqueness of the maximizer of the log-likelihood ratio $\mathcal{E}^M(\cdot)$. When $M \to \infty$, by the Law of Large Numbers,

$$\mathbb{E}\mathcal{E}_{\boldsymbol{X}^{[0,T]}}(\varphi) = \lim_{M\to\infty} \mathcal{E}^M(\varphi) \quad a.s.,$$

and the uniqueness assumption leads to the identifiability in Definition 1.1: the interaction function is identifiable if it is the unique maximizer of the expectation of the log-likelihood ratio.

The coercivity condition in Definition 1.2 can be readily verified from data on any finite dimensional space $\mathcal{H}$ and it provides guidance on the choice of basis functions for $\mathcal{H}$. In fact, with the choice of an orthonormal basis for $\mathcal{H} \subset L^2(\overline{\rho}_T)$, the coercivity constant $c_{\mathcal{H},T}$ provides a lower bound for the smallest eigenvalue of the normal matrix [26,27], thus ensuring the uniqueness of the estimator. More importantly, if the coercivity condition holds true, the estimator is proved to be consistent and converge at a rate (in $M$) equal to the minimax rate of nonparametric regression in 1 dimension (see [26, Theorem 5–6] and [27, Theorem 3.1–3.2]), and the estimation errors can be, under possibly further assumptions, dimension independent (see [26, Theorem 9]).

The next proposition shows that the coercivity condition is sufficient for the identifiability and it also becomes necessary when the number of particles in the system goes to infinity. This implies Theorem 1.3(a).

**Proposition 2.1.** *Consider the system (1.1) on $[0, T]$ with interaction function $\phi$. Let $\overline{\rho}_T$ be as defined in (1.2). Then, the interaction function is identifiable on a subspace $\mathcal{H} \subset L^2(\overline{\rho}_T)$ if and only if*

$$\frac{1}{T}\int_0^T \mathbb{E}[h(|\boldsymbol{r}_{12}^t|)h(|\boldsymbol{r}_{13}^t|)\frac{\langle \boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t\rangle}{|\boldsymbol{r}_{12}^t||\boldsymbol{r}_{13}^t|}](t)dt > -\frac{1}{N-2}\|h\|^2_{L^2(\overline{\rho}_T)}, \quad \textit{for all } h \neq 0 \in \mathcal{H}. \tag{2.4}$$

*Thus, it is identifiable on a linear subspace $\mathcal{H} \subset L^2(\overline{\rho}_T)$ if the coercivity condition holds on $\mathcal{H}$. Furthermore, when $N \to \infty$, the interaction function is identifiable on $L^2(\overline{\rho}_T)$ if and only if the coercivity condition holds on $L^2(\overline{\rho}_T)$.*

**Proof.** Noting that $d\boldsymbol{X}^t = -\nabla J_\phi(\boldsymbol{X}^t)dt + d\boldsymbol{B}^t$ and that $J_\varphi$ is linear in $\varphi$, we have

$$\begin{aligned}
&\int_0^T |\nabla J_\varphi(\boldsymbol{X}^t)|^2 dt + 2\langle \nabla J_\varphi(\boldsymbol{X}^t), d\boldsymbol{X}^t\rangle \\
&= \int_0^T |\nabla J_\varphi(\boldsymbol{X}^t)|^2 dt - 2\langle \nabla J_\varphi(\boldsymbol{X}^t), \nabla J_\phi(\boldsymbol{X}^t)\rangle dt + \langle \nabla J_\varphi(\boldsymbol{X}^t), d\boldsymbol{B}^t\rangle \\
&= \int_0^T |\nabla J_{\varphi-\phi}(\boldsymbol{X}^t)|^2 dt - |\nabla J_\phi(\boldsymbol{X}^t)|^2 dt + \langle \nabla J_\varphi(\boldsymbol{X}^t), d\boldsymbol{B}^t\rangle,
\end{aligned}$$

where the last inequality follows from completing the squares. Note first that from (2.3) with $\varphi = \phi$ we have $\mathbb{E}\mathcal{E}_{\boldsymbol{X}^{[0,T]}}(\phi) = -\frac{1}{TN}\int_0^T |\nabla J_\phi(\boldsymbol{X}^t)|^2 dt$. Then, the expectation of the log-likelihood ratio in (2.3) is

$$\mathbb{E}\mathcal{E}_{\boldsymbol{X}^{[0,T]}}(\varphi) = -\frac{1}{2TN}\int_0^T \mathbb{E}|\nabla J_{\varphi-\phi}(\boldsymbol{X}^t)|^2 dt - \mathbb{E}\mathcal{E}_{\boldsymbol{X}^{[0,T]}}(\phi). \tag{2.5}$$

Thus, $\phi$ is the unique maximizer of $\mathbb{E}\mathcal{E}_{\boldsymbol{X}^{[0,T]}}(\cdot)$ on $\mathcal{H}$, i.e., is identifiable on $\mathcal{H}$, if and only if

$$\frac{1}{TN}\int_0^T \mathbb{E}|\nabla J_h(\boldsymbol{X}^t)|^2 dt > 0 \tag{2.6}$$

whenever $h = \varphi - \phi \neq 0$ in $\mathcal{H}$.

Also, by exchangeability, with notation $\boldsymbol{r}^t_{ji} = \boldsymbol{X}^t_j - \boldsymbol{X}^t_i$, we have

$$\frac{1}{N}\mathbb{E}|\nabla J_h(\boldsymbol{X}^t)|^2 = \sum_{i=1}^N \frac{1}{N^3} \sum_{\substack{j,k=1,\\ j\neq i, k\neq i}}^N \underbrace{\mathbb{E}[h(|\boldsymbol{r}^t_{ji}|)h(|\boldsymbol{r}^t_{ki}|)\frac{\langle \boldsymbol{r}^t_{ji}, \boldsymbol{r}^t_{ki}\rangle}{|\boldsymbol{r}^t_{ji}||\boldsymbol{r}^t_{ki}|}]}_{I_{ijk}(t)}$$
$$= \frac{(N-1)[(N-2)I_{123} + I_{122}]}{N^2},$$

where the equality follows from that $I_{ijk} = I_{123}$ for all triplets $\{(i, j, k), j \neq i, k \neq i, j \neq k\}$, contributing $N(N-1)(N-2)$ copies of $I_{123}$; and that $I_{ijk} = I_{122}$ for all triplets $\{(i, j, k), j = k \neq i\}$, contributing $N(N-1)$ copies of $I_{122}$. Therefore,

$$\frac{1}{TN}\int_0^T \mathbb{E}|\nabla J_h(\boldsymbol{X}^t)|^2 dt = \frac{(N-1)(N-2)}{N^2}\frac{1}{T}\int_0^T I_{123}(t)dt + \frac{N-1}{N^2}\frac{1}{T}\int_0^T I_{122}(t)dt. \tag{2.7}$$

Then, Eq. (2.4) is equivalent to Eq. (2.6) by noting that $\frac{1}{T}\int_0^T I_{122}(t)dt = \frac{1}{T}\int_0^T \mathbb{E}[h(|\boldsymbol{r}^t_{12}|)^2]dt = \|h\|^2_{L^2(\overline{\rho}_T)}$.

If the coercivity condition holds on $\mathcal{H}$, i.e. $\frac{1}{T}\int_0^T \mathbb{E}[h(|\boldsymbol{r}^t_{12}|)h(|\boldsymbol{r}^t_{13}|)\frac{\langle \boldsymbol{r}^t_{12}, \boldsymbol{r}^t_{13}\rangle}{|\boldsymbol{r}^t_{12}||\boldsymbol{r}^t_{13}|}](t)dt \geqslant c_{\mathcal{H},T}\|h\|^2_{L^2(\overline{\rho}_T)}$ for all $h \in \mathcal{H}$, so Eq. (2.4) holds and $\phi$ is identifiable on $\mathcal{H}$.

When $N \to \infty$, by (2.5), the maximizer is unique iff (together with (2.7))

$$\lim_{N\to\infty}\frac{1}{TN}\int_0^T \mathbb{E}|\nabla J_h(\boldsymbol{X}^t)|^2 dt = \frac{1}{T}\int_0^T \mathbb{E}[h(|\boldsymbol{r}^t_{12}|)h(|\boldsymbol{r}^t_{13}|)\frac{\langle \boldsymbol{r}^t_{12}, \boldsymbol{r}^t_{13}\rangle}{|\boldsymbol{r}^t_{12}||\boldsymbol{r}^t_{13}|}](t)dt > 0,$$

for each $h \neq 0 \in \mathcal{H}$. Hence, for any finite dimensional linear subspace $\mathcal{H}$, Eq. (1.3) holds. Thus, identifiability on $L^2(\overline{\rho}_T)$ is equivalent to the coercivity condition on $L^2(\overline{\rho}_T)$ when $N \to \infty$. ■

**Remark 2.2.** When the related integral operator introduced below is compact and strictly positive (see Remark 2.6), the coercivity constant $c_{\mathcal{H},T}$ will converge to zero as the dimension of $\mathcal{H}$ increases to infinity, because it is the smallest eigenvalue of the integral operator on $\mathcal{H} \subset L^2(\overline{\rho}_T)$. Thus, when performing nonparametric inference of the interaction function, even when the coercivity condition holds true on $L^2(\overline{\rho}_T)$, regularization becomes necessary to control the condition number of the normal matrix when the dimension of the hypothesis space $\mathcal{H}$ becomes large [26–28].

**Remark 2.3.** The coercivity in Definition 1.2 is sightly more restrictive than the one in the previous studies [26–28], which is almost equivalent to the identifiability. More precisely, the definition in [27, Definition 3.1] says that the coercivity condition holds on a set $\mathcal{H}$ if there is a positive constant $c_{\mathcal{H}}$ such that

$$c_{\mathcal{H}}\|\varphi\|_{L^2(\overline{\rho}_T)} \leqslant \frac{1}{2NT}\int_0^T \sum_{i=1}^N \mathbb{E}|\nabla J_\varphi(\boldsymbol{X}^t)|^2 dt. \tag{2.8}$$

for all $\varphi \in \mathcal{H}$. This is almost the identifiability in (2.6), except requiring a uniform bound $c_{\mathcal{H}}$ over $\mathcal{H}$. Definition 1.2 is more restrictive than the previous one in the sense that its coercivity constant $c_{\mathcal{H},T}$ in (1.3) may be zero while the constant $c_{\mathcal{H}}$ above is positive, as in (2.9). In fact,

by (2.7) with $h$ replaced by $\varphi$ and the fact that $\frac{1}{T}\int_0^T I_{122}(t)dt = \|\varphi\|^2_{L^2(\overline{\rho}_T)}$, we can write (2.8) as

$$c_{\mathcal{H}}\|\varphi\|_{L^2(\overline{\rho}_T)} \leqslant \frac{(N-1)(N-2)}{2N^2}\frac{1}{T}\int_0^T I_{123}(t)dt + \frac{N-1}{2N^2}\|h\|^2_{L^2(\overline{\rho}_T)}$$

$$\iff \quad (c_{\mathcal{H}} - \frac{N-1}{2N^2})\|\varphi\|_{L^2(\overline{\rho}_T)} \leqslant \frac{(N-1)(N-2)}{2N^2}\frac{1}{T}\int_0^T I_{123}(t)dt,$$

Combining with (1.3), we obtain

$$c_{\mathcal{H}} - \frac{N-1}{2N^2} \leqslant \frac{(N-1)(N-2)}{2N^2}c_{\mathcal{H},T}. \tag{2.9}$$

Hence, we can have $c_{\mathcal{H},T} \leqslant 0$ while having $0 < c_{\mathcal{H}} \leqslant \frac{(N-1)(N-2)}{2N^2}c_{\mathcal{H},T} + \frac{N-1}{2N^2}$. When $N \to \infty$, we have $c_{\mathcal{H}} \leqslant c_{\mathcal{H},T}$ and the two definitions are equivalent. Definition 1.2 has the advantage of providing a coercivity constant independent of $N$. Also, by requiring a positive coercivity constant $c_{\mathcal{H},T}$ only on each finite-dimensional hypothesis space $\mathcal{H}$, it can hold on infinite dimensional spaces such as $L^2(\overline{\rho}_T)$ when $N \to \infty$, while the previous definition cannot. Thus, the new definition is suitable for studying the identifiability of the interaction function in the mean-field equation.

While numerically easy to verify, the above coercivity condition on a time interval is difficult to analyze directly, because it involves the average-in-time distribution $\overline{\rho}_T$ which is complicated in general, unless it is an invariant measure of the system, either when the system starts from the invariant measure, or when we consider the large time limit. The following single-time version of the coercivity condition can be analyzed directly, and involves only the single-time distribution $\rho_t$.

**Definition 2.4** (*Coercivity Condition at Time $t$*)**.** The dynamical system (1.1) is said to satisfy the coercivity condition at time $t$ on a linear subspace $\mathcal{H} \subset L^2(\rho_t)$, where $\rho_t$ is defined in (1.2), if

$$c_{\mathcal{H}}(t) := \inf_{h\in\mathcal{H},\, \|h\|_{L^2(\rho_t)}=1} \mathbb{E}[h(|\boldsymbol{r}^t_{12}|)h(|\boldsymbol{r}^t_{13}|)\frac{\langle \boldsymbol{r}^t_{12}, \boldsymbol{r}^t_{13}\rangle}{|\boldsymbol{r}^t_{12}||\boldsymbol{r}^t_{13}|}] > 0, \tag{2.10}$$

where $\boldsymbol{r}^t_{ij} = \boldsymbol{X}^t_i - \boldsymbol{X}^t_j$. If the coercivity condition holds true on every finite dimensional subspace $\mathcal{H} \subset L^2(\rho_t)$, we say the system satisfies the coercivity condition on $L^2(\rho_t)$ at time $t$.

Similar to Proposition 2.1, if the coercivity condition holds on $\mathcal{H}$ at $t_0$, then the interaction function can be identified on $\mathcal{H}$ from a large size of samples at time $t_0$. This explains the observation in [26–28] that the interaction function can be learned from multiple very short-time trajectories.

### 2.2. *Relation to strictly positive integral operators*

We show in this subsection that the coercivity condition on an interval $[0, T]$ (or at single-time $t$) discussed above is equivalent to the strict positivity of related integral operators on $L^2(\overline{\rho}_T)$ (or $L^2(\rho_t)$).

Recall that a linear operator $Q$ on a Hilbert space $H$ is *positive* if $\langle Qf, f\rangle \geqslant 0$ for any $f \in H$. It is said to be *strictly positive* if $\langle Qf, f\rangle > 0$ whenever $f \neq 0 \in \mathcal{H}$. Hereafter, by an

integral operator $Q$ with kernel $K$ on $L^2(\rho)$, we mean the bounded linear operator defined by

$$[Qh](r) = \int K(r,s)h(s)\rho(s)ds$$

for any $h \in L^2(\rho)$.

**Proposition 2.5.** *The system (1.1) on $[0,T]$ satisfies the coercivity condition on $L^2(\overline{\rho}_T)$ iff the integral operator $\overline{Q}_T$ on $L^2(\overline{\rho}_T)$ with the integral kernel*

$$\overline{K}_T(r,s) := \frac{1}{\overline{q}_T(r)\overline{q}_T(s)}(rs)^{d-1}\frac{1}{T}\int_0^T\int_{S^{d-1}}\int_{S^{d-1}}\langle\xi,\eta\rangle p_t(r\xi,s\eta)d\xi d\eta dt \tag{2.11}$$

*is strictly positive, where $\overline{q}_T(r)$ denotes the density of the measure $\overline{\rho}_T$ and $p_t(u,v)$ denotes the density function of the random vector $(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t)$.*

**Proof.** By definition, we have

$$[\overline{Q}_T h](r) = \int \overline{K}_T(r,s)h(s)\overline{\rho}_T(s)ds. \tag{2.12}$$

Note first that for any $h, g \in L^2(\overline{\rho}_T)$, by a change of variable to the polar coordinates, we have

$$\begin{aligned}
&\frac{1}{T}\int_0^T \mathbb{E}[h(|\boldsymbol{r}_{12}^t|)g(|\boldsymbol{r}_{13}^t|)\frac{\langle \boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t\rangle}{|\boldsymbol{r}_{12}^t||\boldsymbol{r}_{13}^t|}]dt = \frac{1}{T}\int_0^T\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(u)g(v)\frac{\langle u,v\rangle}{|u||v|}p_t(u,v)dudvdt\\
=&\frac{1}{T}\int_0^T\int_{\mathbb{R}^+}\int_{\mathbb{R}^+}\int_{S^{d-1}}\int_{S^{d-1}} h(r)g(s)p(r\xi,s\eta)\langle\xi,\eta\rangle(rs)^{d-1}d\xi d\eta drdsdt\\
=&\int_{\mathbb{R}^+}\int_{\mathbb{R}^+}\overline{K}_T(r,s)h(r)g(s)\overline{\rho}_T(dr)\overline{\rho}_T(ds) = \langle\overline{Q}_T h, g\rangle_{L^2(\overline{\rho}_T)}.
\end{aligned} \tag{2.13}$$

Then, to show the equivalence between the strictly positive-definiteness of $\overline{Q}_T$ and the coercivity condition on $L^2(\overline{\rho}_T)$, it suffices to note that by the above equality, the coercivity constant in (1.3) satisfies

$$c_{\mathcal{H},T} = \inf_{h\in\mathcal{H},\, \|h\|_{L^2(\overline{\rho}_T)}=1}\langle\overline{Q}_T h, h\rangle_{L^2(\overline{\rho}_T)} > 0,$$

for any finite dimensional linear subspace $\mathcal{H} \subset L^2(\overline{\rho}_T)$. ■

**Remark 2.6.** It follows from (2.13) that $\overline{Q}_T$ is a symmetric bounded linear operator on $L^2(\overline{\rho}_T)$:

$$\langle\overline{Q}_T h, g\rangle_{L^2(\overline{\rho}_T)} \leqslant \frac{1}{T}\int_0^T \mathbb{E}[h(|\boldsymbol{r}_{12}^t|)g(|\boldsymbol{r}_{13}^t|)]dt \leqslant \|h\|_{L^2(\overline{\rho}_T)}\|g\|_{L^2(\overline{\rho}_T)},$$

where the second inequality follows from the Cauchy–Schwarz inequality and the fact that $\frac{1}{T}\int_0^T \mathbb{E}[h(|\boldsymbol{r}_{12}^t|)^2]dt = \|h\|^2_{L^2(\overline{\rho}_T)}$. When the integral kernel $\overline{K}_T$ is in $L^2(\overline{\rho}_T \otimes \overline{\rho}_T)$, the operator $\overline{Q}_T$ is compact and its eigen-functions form an orthonormal basis of $L^2(\overline{\rho}_T)$. Then, for any linear subspace $\mathcal{H} \in L^2(\overline{\rho}_T)$, the coercivity constant $c_{\mathcal{H},T}$ is the smallest eigenvalue of $\overline{Q}_T$ on $\mathcal{H}$.

Similarly, we have the following proposition for the coercivity condition at a single time.

**Proposition 2.7.** *The system (1.1) satisfies the coercivity conditions on $L^2(\rho_t)$ at time $t$ iff the integral operator $Q_t$ with kernel*

$$K_t(r,s) := \frac{1}{q_t(r)q_t(s)}(rs)^{d-1}\int_{S^{d-1}}\int_{S^{d-1}}\langle \xi,\eta\rangle p_t(r\xi, s\eta)d\xi d\eta. \tag{2.14}$$

*is strictly positive on $L^2(\rho_t)$, where $q_t$ denotes the density of the measure $\rho_t$ and $p_t(u,v)$ denotes the density function of the random vector $(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t)$.*

**Proof.** Note that

$$\mathbb{E}\left[h(|\boldsymbol{r}_{12}^t|)h(|\boldsymbol{r}_{13}^t|)\frac{\langle \boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t\rangle}{|\boldsymbol{r}_{12}^t||\boldsymbol{r}_{13}^t|}\right] = \int_0^\infty\int_0^\infty h(r)h(s)K_t(r,s)\rho_t(dr)\rho_t(ds) = \langle Q_t h, h\rangle_{L^2(\rho_t)}. \tag{2.15}$$

Then, $c_{\mathcal{H}}(t) = \inf_{h\in\mathcal{H},\, \|h\|_{L^2(\rho_t)}=1}\langle Q_t h, h\rangle_{L^2(\rho_t)}$. Hence, $Q_t$ is strictly positive iff $c_{\mathcal{H}}(t) > 0$ for any finite dimensional linear subspace $\mathcal{H} \subset L^2(\rho_t)$. ■

The positivity of $Q_t$ on $L^2(\rho_t)$ will be proved in later sections by showing that its integral kernel $K_t$ is strictly positive-definite, using the special structure of $p_t(u,v)$ (the density function of the random vector $(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t)$). However, the positivity of $\overline{Q}_T$ on $L^2(\overline{\rho}_T)$ cannot be proved in the same way because the kernel $\overline{K}_T$ in (2.11) involves the average of $p_t(u,v)$. The following proposition shows that, while $Q_t$ and $\overline{Q}_T$ are defined on different spaces (unless $\rho_t$ is the same as $\overline{\rho}_T$), the strict positivity of $Q_t$ for all $t \in [0,T]$ implies the positivity of $\overline{Q}_T$.

**Proposition 2.8.** *The integral operator $\overline{Q}_T$ on $L^2(\overline{\rho}_T)$ with kernel $\overline{K}_T$ in (2.11) is strictly positive on $\mathcal{H}$ if the family of operators $Q_t$ on $L^2(\rho_t)$ with kernel $K_t$ in (2.14) is strictly positive for each $t \in [0,T]$.*

**Proof.** Note that a combination of (2.13) and (2.15) leads to

$$\langle \overline{Q}_T h, h\rangle_{L^2(\overline{\rho}_T)} = \frac{1}{T}\int_0^T \langle Q_t h, h\rangle_{L^2(\rho_t)}dt. \tag{2.16}$$

For any $h \neq 0$, we have $\langle Q_t h, h\rangle_{L^2(\rho_t)} > 0$ for all $t$ because $Q_t$ is strictly positive. Thus, $\langle \overline{Q}_T h, h\rangle_{L^2(\overline{\rho}_T)} > 0$. This implies that $\overline{Q}_T$ is strictly positive on $L^2(\overline{\rho}_T)$. ■

**Remark 2.9.** If we let $\mathcal{H} \subseteq C_b(\mathbb{R}^+)$, then $\overline{Q}_T$ is strictly positive on $\mathcal{H}$ if $Q_t$ is non-negative for each $t \in [0,T]$ and strictly positive on $\mathcal{H}$ for some $t_0 \in [0,T]$, because the integrand $g(t) = \langle Q_t h, h\rangle_{L^2(\rho_t)}$ in (2.16) is continuous in time $t$ for each $h \in \mathcal{H}$. Indeed, note that $\mathbb{E}[f(\boldsymbol{X}^t)]$ is continuous in $t$ for any $f \in C_b(\mathbb{R}^{Nd}, \mathbb{R})$ because $\boldsymbol{X}^t$ is a diffusion process whose density satisfies the Kolmogorov forward equation. Thus, for any function $h \in C_b\mathbb{R}^+$, the function $f(\boldsymbol{X}^t) := h(|\boldsymbol{r}_{12}^t|)h(|\boldsymbol{r}_{13}^t|)\frac{\langle \boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t\rangle}{|\boldsymbol{r}_{12}^t||\boldsymbol{r}_{13}^t|}$ is continuous and bounded, so $\langle Q_t h, h\rangle_{L^2(\rho_t)} = \mathbb{E}[f(\boldsymbol{X}^t)]$ in (2.15) is continuous in $t \in [0,T]$.

## 3. The case of linear systems

We prove in this section that the coercivity condition holds true for linear systems with general non-degenerate exchangeable Gaussian, covering Theorem 1.3(b). We start with a macro–micro decomposition (with the average position of all particles $\boldsymbol{X}_c^t$ being the macro-state and with $\boldsymbol{Y}^t = \boldsymbol{X}^t - \boldsymbol{X}_c^t$ being the micro-state) to transform the system into a system of

decentralized positions, which is ergodic. Then, we prove the coercivity condition when the initial distribution has a special structure similar to the covariance of the invariant measure (see Theorem 3.3), by using a series representation of the integral kernel and by a Müntz-type theorem about polynomials with even degrees being dense in $L^2(\mathbb{R}^+, \mu)$ for a proper $\mu$. Lastly, we prove the coercivity condition for general initial distributions by extending the arguments to non-radial interaction kernels (see Theorem 3.6).

### 3.1. A macro–micro decomposition

We first consider linear systems for which we have $\phi(r) = \theta r$ with $\theta > 0$ (i.e., $\Phi(r) = \frac{1}{2}\theta r^2$, a quadratic potential). The system (1.1) can be written as

$$d\boldsymbol{X}^t = -\theta \boldsymbol{A}\boldsymbol{X}^t dt + d\boldsymbol{B}^t, \tag{3.1}$$

where the matrix $\boldsymbol{A} \in \mathbb{R}^{Nd\times Nd}$ is given by (with $I_d$ being the identity matrix on $\mathbb{R}^d$)

$$\boldsymbol{A} = \frac{1}{N}\begin{pmatrix} (N-1)I_d & -I_d & \cdots & -I_d \\ -I_d & (N-1)I_d & \cdots & -I_d \\ \vdots & \vdots & \ddots & \vdots \\ -I_d & -I_d & \cdots & (N-1)I_d \end{pmatrix}. \tag{3.2}$$

It is straightforward to see that $\boldsymbol{A}^2 = \boldsymbol{A}$, and that the matrix $\boldsymbol{A}$ has eigenvalue 1 of multiplicity $(N-1)d$ and eigenvalue 0 of multiplicity $d$ (with a null space $\{\boldsymbol{x} = c(\mathbf{v}, \mathbf{v}, \ldots, \mathbf{v}) : c \in \mathbb{R}, \mathbf{v} \in \mathbb{R}^d\}$).

By a macro–micro decomposition of the system as in [9,29], the next lemma shows that the center of the particles moves like a Brownian motion, and the particles concentrate around the center with a distribution close to Gaussian.

**Lemma 3.1.** *(i) The solution $\boldsymbol{X}^t$ of Eq. (3.1) can be explicitly written as*

$$\boldsymbol{X}^t = e^{-\theta t}\boldsymbol{A}\boldsymbol{X}^0 + \int_0^t e^{-\theta(t-s)}\boldsymbol{A}d\boldsymbol{B}^s + \boldsymbol{X}_c^t, \tag{3.3}$$

*where $\boldsymbol{X}_c^t = (\mathbf{v}^\mathbf{t}, \mathbf{v}^\mathbf{t}, \ldots, \mathbf{v}^\mathbf{t})'$ with $\mathbf{v}^\mathbf{t} := \frac{1}{N}\sum_{i=1}^N \boldsymbol{X}_i^t = \frac{1}{N}\sum_{i=1}^N(\boldsymbol{X}_i^0 + \boldsymbol{B}_i^t)$.*

*(ii) Conditional on $\boldsymbol{X}^0$, the centralized process*

$$\boldsymbol{Y}^t = \boldsymbol{X}^t - \boldsymbol{X}_c^t$$

*is an Ornstein–Uhlenbeck process with distribution $\mathcal{N}\left(e^{-\theta t}\boldsymbol{A}\boldsymbol{X}^0, \frac{1}{2\theta}(1-e^{-\theta t})\boldsymbol{A}\right)$ for each $t$. In particular, if $\boldsymbol{X}^0$ is Gaussian with variance $\boldsymbol{\Sigma}$, then for each $t$, $\boldsymbol{Y}^t$ has a distribution $\mathcal{N}\left(\mathbf{0}, e^{-2\theta t}\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A} + \frac{1}{2\theta}(1-e^{-\theta t})\boldsymbol{A}\right)$.*

**Proof.** The fact that $\mathbf{v}^\mathbf{t} := \frac{1}{N}\sum_{i=1}^N \boldsymbol{X}_i^t = \frac{1}{N}\sum_{i=1}^N(\boldsymbol{X}_i^0 + \boldsymbol{B}_i^t)$ follows directly from the equation

$$d\mathbf{v}^\mathbf{t} = \frac{1}{N}\sum_{i=1}^N d\boldsymbol{X}_i^t = \frac{1}{N}\sum_{i=1}^N d\boldsymbol{B}_i^t.$$

Next, note that $\boldsymbol{Y}^t = \boldsymbol{X}^t - \boldsymbol{X}_c^t = \boldsymbol{A}\boldsymbol{X}^t$ and

$$d\boldsymbol{Y}^t = \boldsymbol{A}d\boldsymbol{X}^t = -\theta\boldsymbol{A}^2\boldsymbol{X}^t dt + \boldsymbol{A}d\boldsymbol{B}^t = -\theta\boldsymbol{Y}^t dt + \boldsymbol{A}d\boldsymbol{B}^t,$$

where we used $\boldsymbol{A}^2 = \boldsymbol{A}$ in the third equality. Therefore, $(\boldsymbol{Y}^t)$ is an Ornstein–Uhlenbeck process

$$\boldsymbol{Y}^t = e^{-\theta t}\boldsymbol{Y}^0 + \int_0^t e^{-\theta(t-s)}\boldsymbol{A}d\boldsymbol{B}^s.$$

Therefore, conditional on $\boldsymbol{X}^0$, with $\boldsymbol{Y}^0 = \boldsymbol{A}\boldsymbol{X}^0$ and $\boldsymbol{A}^2 = \boldsymbol{A}$, we have that the distribution of $\boldsymbol{Y}^t$ is $\mathcal{N}\left(e^{-\theta t}\boldsymbol{A}\boldsymbol{X}^0, \frac{1}{2\theta}(1-e^{-\theta t})\boldsymbol{A}\right)$ and that $\boldsymbol{X}^t = \boldsymbol{X}^t_c + \boldsymbol{Y}^t$ can be written as in (3.3).

If the initial distribution $\boldsymbol{X}^0$ is exchangeable, then $\mathbb{E}[\boldsymbol{Y}^0] = \boldsymbol{A}\mathbb{E}[\boldsymbol{X}^0] = \boldsymbol{0}$ because $\mathbb{E}[\boldsymbol{X}^0_i] = \mathbb{E}[\boldsymbol{X}^0_j]$ for any $(i, j)$. Thus, if $\boldsymbol{X}^0$ is Gaussian and exchangeable, then $\boldsymbol{Y}^t$ is Gaussian with mean $\boldsymbol{0}$. The variance of $\boldsymbol{Y}^t$ follows directly from the above integral representation. ■

### 3.2. Coercivity condition for linear systems

We begin with a technical lemma for generic Gaussian random vectors. We denote by $\mathrm{cov}(X, Y)$ the covariance of $X$ and $Y$, with the convention that $\mathrm{cov}(X) = \mathrm{cov}(X, X)$.

**Lemma 3.2.** *Let $(X, Y, Z)$ be a Gaussian vector in $\mathbb{R}^{3d}$ with an exchangeable joint distribution and $\mathrm{cov}(X) - \mathrm{cov}(X, Y) = \lambda I_d$ for some $\lambda > 0$. Denote $p^\lambda(u, v)$ the joint distribution of $(X - Y, X - Z)$ and denote $\rho^\lambda$ the distribution (and $q^\lambda(r)$ the density function) of $|X - Y|$. Then*

*(i) The function $K^\lambda(r, s) : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ defined by*

$$K^\lambda(r, s) := \frac{1}{q^\lambda(r)q^\lambda(s)}(rs)^{d-1}\int_{S^{d-1}}\int_{S^{d-1}}\langle\xi, \eta\rangle p^\lambda(r\xi, s\eta)d\xi d\eta \tag{3.4}$$

*is uniformly bounded and is in $L^2(\rho^\lambda \otimes \rho^\lambda)$.*

*(ii) The integral operator $Q^\lambda$ with kernel $K^\lambda$ is strictly positive on $L^2(\rho^\lambda)$, i.e., for any $0 \neq h \in L^2(\rho^\lambda)$,*

$$\langle Q^\lambda h, h\rangle_{L^2(\rho^\lambda)} = \mathbb{E}\left[h(|X - Y|)h(|X - Z|)\frac{\langle X - Y, X - Z\rangle}{|X - Y||X - Z|}\right] > 0. \tag{3.5}$$

**Proof.** We start with explicit expressions for $p^\lambda(u, v)$, $\rho^\lambda(r)$ and $K^\lambda(r, s)$. By exchangeability, the random vector $(X - Y, X - Z)$ is centered Gaussian with covariance matrix $\lambda\begin{pmatrix}2I_d & I_d\\ I_d & 2I_d\end{pmatrix}$, whose inverse is $\frac{1}{3\lambda}\begin{pmatrix}2I_d & -I_d\\ -I_d & 2I_d\end{pmatrix}$. Thus, the joint distribution is

$$p^\lambda(u, v) = (2\sqrt{3}\pi\lambda)^{-d}e^{-c_\lambda(|u|^2+|v|^2-\langle u,v\rangle)}, \quad \text{with } c_\lambda = \frac{1}{3\lambda},$$

and a direct computation yields that the density of $|X - Y|$ is

$$q^\lambda(r) = \frac{1}{C_\lambda}r^{d-1}e^{-\frac{r^2}{4\lambda}}\mathbf{1}_{r\geqslant 0}, \quad \text{with } C_\lambda = \frac{1}{2}(4\lambda)^{\frac{d}{2}}\Gamma(\frac{d}{2}),$$

where the constant $\Gamma(\frac{d}{2})$ comes from the surface area of the unit sphere in $\mathbb{R}^d$: $|S^{d-1}| = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}$.

Then, the integral kernel in (3.4) can be written as

$$K^\lambda(r, s) = C_d e^{-\frac{1}{12\lambda}(r^2+s^2)}\int_{S^{d-1}}\int_{S^{d-1}}\langle\xi, \eta\rangle e^{c_\lambda rs\langle\xi,\eta\rangle}\frac{d\xi d\eta}{|S^{d-1}|^2}, \quad \text{with } C_d = (\frac{\sqrt{3}}{2})^{-d}.$$

Here when $d = 1$, the above spherical measure on $S^0 = \{-1, 1\}$ is interpreted as $\mathbb{P}(\xi = 1) = \mathbb{P}(\xi = -1) = \frac{1}{2}$, or equivalently, $\int_{S^{d-1}} \int_{S^{d-1}} \langle \xi, \eta \rangle e^{\frac{1}{3\lambda}(rs\langle \xi, \eta \rangle)} \frac{d\xi d\eta}{|S^{d-1}|^2} = \frac{1}{2}(e^{c_\lambda rs} - e^{-c_\lambda rs})$.

To prove (i), note that $\langle \xi, \eta \rangle e^{c_\lambda rs\langle \xi, \eta \rangle} \leqslant e^{c_\lambda rs}$. Then,

$$K^\lambda(r, s) \leqslant C_d e^{-\frac{1}{12\lambda}(r^2+s^2)+c_\lambda rs},$$

and it follows that $K^\lambda(r, s)$ is ~~uniformly bounded above and is~~ in $L^2(\rho^\lambda \otimes \rho^\lambda)$.

To prove (ii), we first represent $K^\lambda(r, s)$ in terms a series of polynomials and then apply a Müntz-type theorem. Note that by Taylor expansion, $\langle \xi, \eta \rangle e^{c_\lambda rs\langle \xi, \eta \rangle} = \sum_{k=1}^\infty \frac{1}{(k-1)!} c_\lambda^{k-1} (rs)^{k-1} \langle \xi, \eta \rangle^k$, and that

$$b_k = \int_{S^{d-1}} \int_{S^{d-1}} \langle \xi, \eta \rangle^k \frac{d\xi d\eta}{|S^{d-1}|^2} \begin{cases} = 0, & \text{for odd } k, \\ \in (0, 1), & \text{for even } k, \end{cases}$$

due to symmetry. We have

$$K^\lambda(r, s) = C_d e^{-\frac{1}{12\lambda}(r^2+s^2)} \sum_{k=0}^\infty \frac{1}{k!} c_\lambda^k b_{k+1} (rs)^k = C_d e^{-\frac{1}{12\lambda}(r^2+s^2)} \sum_{k=1, k \text{ odd}}^\infty \frac{1}{k!} c_\lambda^k b_{k+1} (rs)^{k-1}.$$

Then, for any $h \in L^2(\rho^\lambda)$, we have

$$\begin{aligned} \langle Q^\lambda h, h \rangle_{L^2(\rho^\lambda)} &= \int_0^\infty h(r)h(s)K^\lambda(r, s)\rho^\lambda(r)\rho^\lambda(s)drds \\ &= C_d \sum_{k=1, k \text{ odd}}^\infty \frac{1}{k!} c_\lambda^k b_{k+1} \left( \int_0^\infty h(r) r^{k-1} e^{-\frac{1}{12\lambda}r^2} \rho^\lambda(r) dr \right)^2 \geqslant 0. \end{aligned}$$

Note that

$$\int_0^\infty h(r) r^{k-1} e^{-\frac{1}{12\lambda}r^2} \rho^\lambda(r) dr = C_\lambda^{-1} \int_0^\infty h(r) r^{k+d-2} e^{-\frac{1}{3\lambda}r^2} dr.$$

By Lemma A.9, a variation of the Müntz Theorem, the space $\text{span}\{1, r^2, r^4, r^6, \ldots\}$ is dense in $L^2(r^{d-1}e^{-\frac{1}{3\lambda}r^2})$. Thus, $\langle Q^\lambda h, h \rangle_{L^2(\rho^\lambda)} = 0$ only if $h \equiv 0$, and $Q^\lambda$ is strictly positive. Eq. (3.5) follows as in Eq. (2.13). ■

The next theorem implies Theorem 1.3(b) under the additional assumption that the initial distribution of $\boldsymbol{X}^0$ is exchangeable Gaussian with a covariance satisfying $\text{cov}(\boldsymbol{X}_i^0) - \text{cov}(\boldsymbol{X}_i^0, \boldsymbol{X}_j^0) = \lambda_0 I_d$ for any $1 \leqslant i < j \leqslant N$. Intuitively, we may decompose each component of $\boldsymbol{X}^0$ as the sum of a common variable and an independent variable, i.e., $\boldsymbol{X}_i^0 = Z_i + W$, where $\{Z_i\}_{i=1}^N$ are i.i.d. $\mathcal{N}(0, \lambda_0 I_d)$ and $W$ is a common Gaussian random variable, and this implies that the particles are initially scattered randomly around a random position.

**Theorem 3.3.** *Suppose the linear system (3.1) starts from $\boldsymbol{X}^0$ whose distribution is exchangeable Gaussian with covariance satisfying $\text{cov}(\boldsymbol{X}_1^0) - \text{cov}(\boldsymbol{X}_1^0, \boldsymbol{X}_2^0) = \lambda_0 I_d$ for some $\lambda_0 > 0$. Then,* (i) *the coercivity condition holds on $L^2(\rho_t)$ at each time $t \in [0, T]$ as in Definition 2.4;* (ii) *the coercivity condition holds on $L^2(\overline{\rho_T})$ on $[0, T]$ as in Definition 1.2.*

**Proof.** Let $\boldsymbol{Y}^t = \boldsymbol{X}^t - \boldsymbol{X}_c^t$. Note that

$$\boldsymbol{r}_{ij}^t = \boldsymbol{X}_i^t - \boldsymbol{X}_j^t = \boldsymbol{Y}_i^t - \boldsymbol{Y}_j^t.$$

Thus, the coercivity conditions for the process $(\boldsymbol{X}^t)$ is equivalent to those for the process $(\boldsymbol{Y}^t)$.

With $M := \mathrm{cov}(\boldsymbol{X}_i^0, \boldsymbol{X}_j^0)$, the covariance of $\boldsymbol{X}^0$ can be written as

$$\boldsymbol{\Sigma} = \begin{pmatrix} M + \lambda_0 I_d & M & \cdots & M \\ M & M + \lambda_0 I_d & \cdots & M \\ \vdots & \vdots & \ddots & \vdots \\ M & M & \cdots & M + \lambda_0 I_d \end{pmatrix}.$$

A direct computation shows that $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A} = \lambda_0 \boldsymbol{A}$ for $\boldsymbol{A}$ defined in (3.2). By Lemma 3.1(ii), this special covariance implies that the centralized process $\boldsymbol{Y}^t = \boldsymbol{X}^t - \boldsymbol{X}_c^t$ has a covariance $\lambda(t)\boldsymbol{A}$ with $\lambda(t) = \left[e^{-2\theta t}\lambda_0 + \frac{1}{2\theta}(1 - e^{-\theta t})\right]$. Then, $(\boldsymbol{Y}_1^t, \ldots, \boldsymbol{Y}_N^t)$ is exchangeable Gaussian with covariance satisfying $\mathrm{cov}(\boldsymbol{Y}_i^t) - \mathrm{cov}(\boldsymbol{Y}_i^t, \boldsymbol{Y}_j^t) = \lambda(t) I_d$, particularly for the vector $(\boldsymbol{Y}_1^t, \boldsymbol{Y}_2^t, \boldsymbol{Y}_3^t)$. Then, applying Lemma 3.2 to the vector $(\boldsymbol{Y}_1^t, \boldsymbol{Y}_2^t, \boldsymbol{Y}_3^t)$, we obtain that the integral operator $Q^{\lambda(t)}$ with kernel $K^{\lambda(t)}$ defined in (3.4) is strictly positive on $L^2(\rho_t)$ for each $t$. In the notation of Proposition 2.7, this implies that integral operator $Q_t = Q^{\lambda(t)}$ with kernel $K_t = K^{\lambda(t)}$ is strictly positive on $L^2(\rho_t)$. Part (i) then follows.

Since $Q_t$ is strictly positive on $L^2(\rho_t)$ for each $t \in [0, T]$, so is $\overline{Q}_T$ on $L^2(\overline{\rho}_T)$ by Proposition 2.8. Then, the coercivity condition holds on $L^2(\overline{\rho}_T)$ by Proposition 2.5. ■

**Remark 3.4.** When the system is deterministic, i.e. there is no stochastic force, the coercivity conditions hold true on $L^2(\overline{\rho}_T)$ when the initial distribution is exchangeable Gaussian with $\mathrm{cov}(\boldsymbol{X}_i^0) - \mathrm{cov}(\boldsymbol{X}_i^0, \boldsymbol{X}_j^0) = \lambda_0 I_d$. In fact, we have $\boldsymbol{X}^t = e^{-\theta t}\boldsymbol{A}\boldsymbol{X}^0 + \boldsymbol{X}_c^t$ and $\boldsymbol{Y}^t = e^{-\theta t}\boldsymbol{A}\boldsymbol{X}^0$. Then the vector $(\boldsymbol{Y}_1^t, \boldsymbol{Y}_2^t, \boldsymbol{Y}_3^t)$ is exchangeable Gaussian with $\mathrm{cov}(\boldsymbol{Y}_i^t) - \mathrm{cov}(\boldsymbol{Y}_i^t, \boldsymbol{Y}_j^t) = e^{-2\theta t}\lambda_0 I_d$. The coercivity condition follows again from Lemma 3.2 and Proposition 2.5. In particular, it holds when the initial distribution is standard Gaussian, i.e., the initial position of the particles are i.i.d. Gaussian.

### *3.3. Coercivity condition for non-radial interaction functions*

The covariance constraint $\mathrm{cov}(\boldsymbol{X}_i^0) - \mathrm{cov}(\boldsymbol{X}_i^0, \boldsymbol{X}_j^0) = \lambda_0 I_d$ in Theorem 3.3 is necessary for the above proof, due to the need of a series representation of the radial integral kernel $K_t$. Here we remove this constraint by using a series representation of the corresponding non-radial integral kernel. More importantly, we show that the coercivity condition holds true on $L^2$ for interaction functions that are non-radial.

More precisely, consider the system with a non-radial interaction kernel $\phi : \mathbb{R}^d \to \mathbb{R}$,

$$d\boldsymbol{X}_i^t = \frac{1}{N} \sum_{1 \leqslant j \leqslant N, j \neq i} \phi(\boldsymbol{X}_j^t - \boldsymbol{X}_i^t) \frac{\boldsymbol{X}_j^t - \boldsymbol{X}_i^t}{|\boldsymbol{X}_j^t - \boldsymbol{X}_i^t|} dt + \sigma d\boldsymbol{B}_i^t, \quad \text{for } i = 1, \ldots, N, \tag{3.6}$$

with initial condition $\boldsymbol{X}^0$. We will extend the above coercivity condition to non-radial interaction functions.

It is straightforward to see from Section 2.1 that for non-radial interaction functions, the function space of learning is $L^2(\mathbb{R}^d, \overline{\rho}_T)$ or $L^2(\mathbb{R}^d, \rho_t)$ with

$$\overline{\rho}_T(du) = \frac{1}{T} \int_0^T \rho_t(du) dt, \qquad \text{with } \rho_t(du) = \mathbb{E}[\delta(\boldsymbol{X}_i^t - \boldsymbol{X}_j^t \in du)]. \tag{3.7}$$

Correspondingly, the coercivity condition is on $L^2(\mathbb{R}^d, \overline{\rho}_T)$ or $L^2(\mathbb{R}^d, \rho_t)$. For simplicity of notation, we denote them by $L^2(\overline{\rho}_T)$ and $L^2(\rho_t)$.

**Definition 3.5** (*Coercivity Condition for Non-radial Functions*)**.** The dynamical system (3.6) on $[0, T]$ with an initial condition $\boldsymbol{X}^0$ and an interaction function $\phi : \mathbb{R}^d \to \mathbb{R}$ is said to satisfy the coercivity condition on a finite dimensional linear subspace $\mathcal{H} \subset L^2(\overline{\rho_T})$, with $\bar{\rho}_T$ defined in (3.7), if

$$c_{\mathcal{H},T} := \inf_{h\in\mathcal{H},\, \|h\|_{L^2(\overline{\rho}_T)}=1} \frac{1}{T}\int_0^T \mathbb{E}[h(\boldsymbol{r}_{12}^t)h(\boldsymbol{r}_{13}^t)\frac{\langle \boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t\rangle}{|\boldsymbol{r}_{12}^t||\boldsymbol{r}_{13}^t|}]dt > 0, \tag{3.8}$$

where $\boldsymbol{r}_{ij}^t = \boldsymbol{X}_i^t - \boldsymbol{X}_j^t$. If the coercivity condition holds true on every finite dimensional linear subspace $\mathcal{H} \subset L^2(\overline{\rho_T})$, we say the system satisfies the coercivity condition. Similarly, we can define the coercivity condition at a single time $t$ on $L^2(\rho_t)$.

The next theorem shows that the coercivity condition holds true for the linear system when the distribution of $\boldsymbol{X}^0$ is non-degenerate exchangeable Gaussian.

**Theorem 3.6.** *Suppose the linear system (3.1) starts with an initial condition $(\boldsymbol{X}_1^0, \ldots, \boldsymbol{X}_N^0)$ whose distribution is non-degenerate exchangeable Gaussian. Then the coercivity condition holds true at each time $t \geqslant 0$, as well as on $[0, T]$, in the sense of Definition 3.5.*

As in the previous section, we prove the coercivity condition by showing that the corresponding integral operator is strictly positive, based on a series representation of the non-radial integral kernel. Propositions 2.5 and 2.7 can be directly generalized to a non-radial version. We begin with the following lemma, which is a counterpart of the Müntz-type theorem, showing that polynomials are dense in weighted $L^2$ spaces.

**Lemma 3.7** (*[33, Lemma 1.1]*)**.** *Let $\mu$ be a measure on $\mathbb{R}^d$ satisfying*

$$\int e^{c|x|} d\mu(x) < \infty$$

*for some $c > 0$, where $|x| = \sum_{j=1}^d |x_j|$. Then the polynomials are dense in $L^2(\mu)$.*

**Proposition 3.8.** *Let $(X, Y, Z)$ be a Gaussian vector in $\mathbb{R}^{3d}$ with an exchangeable and non-degenerate joint distribution. Denote by $p(u, v)$ the non-degenerate joint density of $(X-Y, X-Z)$ and by $\rho$ the distribution (with a density $q$) of $X - Y$. Then, the integral operator $Q$ with kernel*

$$K(u, v) := \frac{1}{q(u)q(v)}\frac{\langle u, v\rangle}{|u||v|}p(u, v) \tag{3.9}$$

*is strictly positive on $L^2(\rho)$.*

**Proof.** For any $h, g \in L^2(\rho)$, from the definition of $Q$ and the Cauchy–Schwarz inequality, we have

$$|\langle Qh, g\rangle_{L^2(\rho)}| = \left|\mathbb{E}[h(X-Z)g(X-Z)\frac{\langle X-Y, X-Z\rangle}{|X-Y||X-Z|}]\right| \leqslant \|h\|_{L^2(\rho)}\|g\|_{L^2(\rho)}.$$

Thus, $Q$ is a bounded operator on $L^2(\rho)$. We show that $Q$ is strictly positive by

(i) $\langle Qh, h\rangle \geqslant 0$ for any $h \in L^2(\rho)$, and
(ii) $\langle Qh, h\rangle = 0 \Rightarrow h = 0$ in $L^2(\rho)$.

To prove (i), note that

$$\langle Qh, h\rangle_{L^2(\rho)} = \iint h(u)h(v)K(u,v)\overset{q}{\cancel{p}}(u)\cancel{p}(v)dudv = \iint h(u)h(v)\frac{\langle u, v\rangle}{|u||v|}p(u,v)dudv. \tag{3.10}$$

Then, by Theorem A.2, it suffices to show $p(u, v)$ is positive-definite.

Since the joint distribution of $(X, Y, Z)$ is exchangeable non-degenerate Gaussian, there exist a vector $\mu \in \mathbb{R}^d$ and an invertible matrix $\Sigma$ such that the distribution of each of $X, Y, Z$ is $N(\mu, \Sigma)$. Decompose $\Sigma^{-1} = LL^T$, and let $\tilde{X} = L(X-\mu)$, $\tilde{Y} = L(Y-\mu)$, $\tilde{Z} = L(Z-\mu)$. Then, the distribution of each of $\tilde{X}, \tilde{Y}, \tilde{Z}$ is $N(0, I_d)$ and their joint distribution is exchangeable and non-degenerate. It follows that $\mathrm{cov}(\tilde{X}, \tilde{Y}) = \mathrm{cov}(\tilde{X}, \tilde{Z}) = \mathrm{cov}(\tilde{Y}, \tilde{Z})$. Let $\Theta = \mathrm{cov}(\tilde{X}, \tilde{Y}) = \mathbb{E}[\tilde{X}\tilde{Y}^T]$. Since $\Theta$ is real symmetric, it is diagonalizable and there is a real orthonormal matrix $P$ such that $P\Theta P^T = \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$, where $\{\lambda_1, \ldots, \lambda_d\}$ are eigenvalues of $\Theta$. Note that $-1 \leqslant \lambda_i < 1$ because the joint distribution of $(\tilde{X}, \tilde{Y}, \tilde{Z})$ is non-degenerate and $\mathrm{cov}(\tilde{X}) = \mathrm{cov}(\tilde{Y}) = I_d$.

Let $X' = P\tilde{X} = LP(X-\mu)$, $Y' = P\tilde{Y} = LP(Y-\mu)$, $Z' = P\tilde{Z} = LP(Z-\mu)$. By Theorem A.2, to prove that $p(u, v)$ is positive-definite, it suffices to show that the density $q(u, v)$ of $(X'-Y', X'-Z')$ is positive-definite. Note that the covariance matrix $\mathrm{cov}(X'-Y', X'-Z')$ is invertible:

$$\mathrm{cov}(X'-Y', X'-Z')^{-1} = \begin{bmatrix} 2I-2\Theta & I-\Theta \\ I-\Theta & 2I-2\Theta \end{bmatrix}^{-1} = \frac{1}{3}\begin{bmatrix} 2\Lambda & -\Lambda \\ -\Lambda & 2\Lambda \end{bmatrix},$$

where $\Lambda := \mathrm{diag}(a_1, \ldots, a_d)$ with $a_i = \frac{1}{1-\lambda_i} \geqslant \frac{1}{2}$. Thus, with a normalizing constant $C_d > 0$, we have

$$\begin{aligned} q(u,v) &= C_d \exp\left(-\tfrac{1}{2}(u,v)\mathrm{cov}(X'-Y', X'-Z')^{-1}(u,v)^T\right) \\ &= C_d \exp\left(-\tfrac{1}{3}\sum_{i=1}^d a_i(u_i^2+v_i^2) + \tfrac{1}{3}\sum_{i=1}^d a_i u_i v_i\right). \end{aligned}$$

By Theorem A.2, $q(u, v)$ is positive-definite.

To prove (ii), let $h \in L^2(\rho)$ satisfy $\langle Qh, h\rangle = 0$. We need to prove that $h = 0$. Let

$$\begin{aligned} g_j(u) &:= \exp\left(-\frac{1}{12}\sum_{i=1}^d a_i u_i^2\right) h((PL)^{-1}u)\frac{((PL)^{-1}u)_j}{|(PL)^{-1}u|} \\ q(u) &:= \int q(u,v)dv = C_d' \exp\left(-\frac{1}{4}\sum_{i=1}^d a_i u_i^2\right) \\ f(u,v) &:= \exp\left(\frac{1}{3}\sum_{i=1}^d a_i u_i v_i\right). \end{aligned}$$

By the linear transform $X \mapsto PL(X-\mu)$, we can rewrite (3.10) as

$$\langle Qh, h\rangle = C_d(C_d')^{-2}(\det L)^2 \sum_{j=1}^d \iint g_j(u)q(u)g_j(v)q(v)f(u,v)dudv.$$

Note that the Taylor series of $f(u,v)$ is

$$f(u,v)=\sum_{k=0}^{\infty}\frac{1}{k!3^k}(\sum_{i=1}^{d}a_iu_iv_i)^k=\sum_{k=0}^{\infty}\sum_{i_1+\cdots+i_d=k}C_{k,i_1,\ldots,i_d}(u_1v_1)^{i_1}\cdots(u_dv_d)^{i_d},$$

where all coefficients are positive. By Fubini's theorem,

$$\langle Qh,h\rangle=C_d(\det L)^2\sum_j\sum_{k,i_1,\ldots,i_d}C_{k,i_1,\ldots,i_d}\left(\int g_j(u)u_1^{i_1}\cdots u_d^{i_d}q(u)du\right)^2. \tag{3.11}$$

Thus, $\langle Qh,h\rangle=0$ implies that each term must be zero

$$\int g_j(u)u_1^{i_1}\cdots u_n^{i_n}q(u)du=0,\ \text{for any integers } i_1,\ldots,i_d\geqslant 0.$$

Note that the measure $\mu$ defined by $d\mu(u):=q(u)du$ satisfies the condition of Lemma 3.7. Then the polynomials are dense in $L^2(\mathbb{R}^d,q)$. Also, note that $g_j\in L^2(\mathbb{R}^d,q)$ because $|g_j(u)|\leqslant|h((PL)^{-1}u)|$ and

$$\begin{aligned}\int|h((PL)^{-1}u)|^2q(u)du&=\mathbb{E}[|h((PL)^{-1}(X'-Y'))|^2]=\mathbb{E}[|h(X-Y)|^2]\\&=\|h\|^2_{L^2(\rho)}<\infty,\end{aligned}$$

where in the first equality we used the fact that $q(u)$ is the density of $X'-Y'$. Hence, $g_j=0$ in $L^2(\mathbb{R}^d,q)$ for all $j$, and we conclude that $h=0$ in $L^2(\rho)$. ■

**Proof of Theorem 3.6.** Similar to the proof of Theorem 3.3, we only need to consider the process $\boldsymbol{Y}^t=\boldsymbol{X}^t-\boldsymbol{X}^t_c$ because the proof only involves $\boldsymbol{r}^t_{ij}=\boldsymbol{X}^t_i-\boldsymbol{X}^t_j=\boldsymbol{Y}^t_i-\boldsymbol{Y}^t_j$.

By Lemma 3.1, $(\boldsymbol{Y}^t_1,\ldots,\boldsymbol{Y}^t_N)$ is exchangeable Gaussian. In particular, the vector $(\boldsymbol{Y}^t_1,\boldsymbol{Y}^t_2,\boldsymbol{Y}^t_3)$ is exchangeable Gaussian on $\mathbb{R}^{3d}$. Denote by $p_t(u,v)$ the joint distribution of $(\boldsymbol{Y}^t_1-\boldsymbol{Y}^t_2,\boldsymbol{Y}^t_1-\boldsymbol{Y}^t_3)$ and recall in (3.7), we denote by $\rho_t$ the distribution (with $q_t$ denoting its density) of $\boldsymbol{Y}^t_1-\boldsymbol{Y}^t_2$ and denote by $\overline{\rho}_T$ the average of $\rho_t$ on $[0,T]$ (with $\overline{q}_T$ denoting its density).

By Proposition 3.8, the integral operator $Q_t$ with kernel

$$K_t(u,v):=\frac{1}{q_t(u)q_t(v)}\frac{\langle u,v\rangle}{|u||v|}p_t(u,v) \tag{3.12}$$

is strictly positive on $L^2(\rho_t)$. Then it follows from the non-radial version of Proposition 2.7 that the coercivity condition holds true for each time $t$.

Similar to the proof in Theorem 3.3, the coercivity on $[0,T]$ follows from the non-radial version of Propositions 2.5 and 2.8. More precisely, we would like to show that the integral operator $\bar{Q}_T$ with kernel

$$\bar{K}_T(u,v):=\frac{1}{\overline{q}_T(u)\overline{q}_T(v)}\frac{\langle u,v\rangle}{|u||v|}\int_0^T p_t(u,v)dt$$

is strictly positive on $L^2(\bar{\rho}_T)$. Note that $\langle\overline{Q}_Th,h\rangle_{L^2(\overline{\rho}_T)}=\frac{1}{T}\int_0^T\langle Q_th,h\rangle_{L^2(\rho_t)}dt$ for any $h\in L^2(\overline{\rho}_T)$, and we just showed above that $\langle Q_th,h\rangle_{L^2(\rho_t)}>0$ for each $t\in[0,T]$. Thus, $\overline{Q}_T$ is strictly positive. ■

## 4. Nonlinear systems with three particles

We consider the class of nonlinear systems with interaction functions in Theorem 1.3(c), starting from an invariant measure. Since the diffusion process $(\boldsymbol{X}^t,t\in[0,T])$ has a stationary

distribution, the coercivity condition on a time instance is the same as on the time interval $[0, T]$, because $\rho_T = \rho_t$. Thus, we simply say that coercivity condition holds true without specifying it being at a single time or on a time interval when the process is stationary.

Global solutions exist for the gradient system (2.1) with a potential $\Phi(r)$ in (1.4), because the potential leads to a locally Lipschitz drift term and the total potential $J_\phi$ in (2.2) leads to a Lyapunov function for the system.

The following theorem covers Theorem 1.3(c).

**Theorem 4.1.** *The coercivity condition holds true for the system (2.1) with $N = 3$ starting from an initial condition $\boldsymbol{X}^0$ such that the joint density of $(\boldsymbol{X}_1^0 - \boldsymbol{X}_2^0, \boldsymbol{X}_1^0 - \boldsymbol{X}_3^0)$ is $p(u, v)$ in (1.5), and with $\Phi$ in (1.4).*

Note that we only need the joint density of $(\boldsymbol{X}_1^t - \boldsymbol{X}_2^t, \boldsymbol{X}_1^t - \boldsymbol{X}_3^t)$ to study the coercivity condition. We will show first that $p(u, v)$ in (1.5) is an invariant density for $(\boldsymbol{X}_1^t - \boldsymbol{X}_2^t, \boldsymbol{X}_1^t - \boldsymbol{X}_3^t)$ when $N = 3$ (see Section 4.1). Then based on the analytical expression of $p(u, v)$, we introduce a "comparison to Gaussian kernels" technique, which makes extensive use of positive-definite kernels. We will prove the theorem and develop the technique in two steps: first when $\Phi(r) = r^{2\beta}$ in Section 4.2 and then general $\Phi$ in Section 4.3. Due to the need of the above analytical form of $p(u, v)$, our result is limited to the case when $N = 3$ and when the initial distribution is the invariant measure (see Remark 4.4).

### 4.1. Stationary distribution for pairwise differences

We show first that the process of pairwise differences $(\boldsymbol{X}_1^t - \boldsymbol{X}_2^t, \boldsymbol{X}_1^t - \boldsymbol{X}_3^t)$ admits a stationary distribution.

**Proposition 4.2.** *Suppose that $\Phi \in C^2(\mathbb{R}^+, \mathbb{R})$ and $Z = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{-H(u,v)} du dv < \infty$ for*

$$H(u, v) = \frac{1}{3}[\Phi(|u|) + \Phi(|v|) + \Phi(|u - v|)].$$

*Then the process $(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t) = (\boldsymbol{X}_1^t - \boldsymbol{X}_2^t, \boldsymbol{X}_1^t - \boldsymbol{X}_3^t)$ admits $p(u, v)$ in (1.5) as an invariant density.*

**Proof.** Note that

$$\begin{cases} d\boldsymbol{r}_{12}^t &= F(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t)dt + (d\boldsymbol{B}_1^t - d\boldsymbol{B}_2^t), \\ d\boldsymbol{r}_{13}^t &= F(\boldsymbol{r}_{13}^t, \boldsymbol{r}_{12}^t)dt + (d\boldsymbol{B}_1^t - d\boldsymbol{B}_3^t), \end{cases} \tag{4.1}$$

where the function $F : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ is given by

$$F(u, v) = -\frac{1}{3}[2\phi(|u|)u + \phi(|v|)v + \phi(|u - v|)(u - v)],$$

where $\phi(r) = \Phi'(r)$.

The diffusion $\begin{pmatrix} d\boldsymbol{B}_1^t - d\boldsymbol{B}_2^t \\ d\boldsymbol{B}_1^t - d\boldsymbol{B}_3^t \end{pmatrix}$ has a non-degenerate covariance $\begin{pmatrix} 2I_d & I_d \\ I_d & 2I_d \end{pmatrix}$. One can then verify directly that the distribution $p(u, v)$ is a stationary solution to the Kolmogorov forward equation of (4.1). Alternatively, for ease of computation, we show that the system (4.1) is a linear transformation of a gradient system with homogeneous diffusion, which shares the same

invariant measure. Let $A = \sqrt{2}\begin{pmatrix} I_d & 0 \\ 1/2 I_d & \sqrt{3}/2 I_d \end{pmatrix}$, which satisfies $AA^T = \begin{pmatrix} 2I_d & I_d \\ I_d & 2I_d \end{pmatrix}$. Then the process

$$\begin{pmatrix} \boldsymbol{Y}_1^t \\ \boldsymbol{Y}_2^t \end{pmatrix} = A^{-1}\begin{pmatrix} \boldsymbol{r}_{12}^t \\ \boldsymbol{r}_{13}^t \end{pmatrix}$$

is a weak solution to the system

$$d\begin{pmatrix} \boldsymbol{Y}_1^t \\ \boldsymbol{Y}_2^t \end{pmatrix} = A^{-1}\begin{pmatrix} F(\boldsymbol{Y}_1^t, \frac{\sqrt{2}}{2}\boldsymbol{Y}_1^t + \frac{\sqrt{6}}{2}\boldsymbol{Y}_2^t) \\ F(\frac{\sqrt{2}}{2}\boldsymbol{Y}_1^t + \frac{\sqrt{6}}{2}\boldsymbol{Y}_2^t, \boldsymbol{Y}_1^t) \end{pmatrix} dt + \begin{pmatrix} d\widetilde{\boldsymbol{B}}_1^t \\ d\widetilde{\boldsymbol{B}}_2^t \end{pmatrix}, \tag{4.2}$$

where $(\widetilde{\boldsymbol{B}}_1^t, \widetilde{\boldsymbol{B}}_2^t)$ is a standard Brownian motion on $\mathbb{R}^{2d}$. Notice that $H(u, v) = H(v, u)$ and

$$A^{-1}\begin{pmatrix} F(\sqrt{2}u, \frac{\sqrt{2}}{2}u + \frac{\sqrt{6}}{2}v) \\ F(\frac{\sqrt{2}}{2}u + \frac{\sqrt{6}}{2}v, \sqrt{2}u) \end{pmatrix} = \begin{pmatrix} \nabla_u[H(\sqrt{2}u, \frac{\sqrt{2}}{2}u + \frac{\sqrt{6}}{2}v)] \\ \nabla_v[H(\frac{\sqrt{2}}{2}u + \frac{\sqrt{6}}{2}v, \sqrt{2}u)] \end{pmatrix}.$$

Then, it follows from Lemma A.10 that $p_Y(\boldsymbol{y}_1, \boldsymbol{y}_2) \propto e^{-2H(\boldsymbol{y}_1, \frac{\sqrt{2}}{2}\boldsymbol{y}_1 + \frac{\sqrt{6}}{2}\boldsymbol{y}_2)}$ is an invariant density for the system (4.2). Therefore, the process $(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t)$ admits $p(u, v)$ as an invariant density. ■

**Remark 4.3.** Similarly, one can prove that the process $(\boldsymbol{X}_1^t - \boldsymbol{X}_2^t, \boldsymbol{X}_1^t - \boldsymbol{X}_3^t, \ldots, \boldsymbol{X}_1^t - \boldsymbol{X}_N^t)$ admits a stationary density on $\mathbb{R}^{(N-1)d}$. In essence, we decompose the system into a reference particle and the relative positions of other particle to the reference particle. This is similar to the macro–micro decomposition of the system in [9,10,16,29]. But the above transformation has the following advantage: it leads to a gradient system with an additive white noise, and this simplifies the derivation of the stationary distribution.

**Remark 4.4.** Our current proof for the coercivity condition makes use of the analytical expression of the invariant density $p(u, v)$ of $(\boldsymbol{X}_1^t - \boldsymbol{X}_2^t, \boldsymbol{X}_1^t - \boldsymbol{X}_3^t)$ and of the Müntz-type theorem on $\mathbb{R}^+$. When $N > 3$, such an explicit form is no longer available due to the need of marginalizing the joint distribution of $(\boldsymbol{X}_1^t - \boldsymbol{X}_2^t, \boldsymbol{X}_1^t - \boldsymbol{X}_3^t, \ldots, \boldsymbol{X}_1^t - \boldsymbol{X}_N^t)$:

$$p(u, v) = \frac{1}{Z} f(u, v) e^{-\frac{2}{N}[\Phi(|u|) + \Phi(|v|) + \Phi(|u-v|)]},$$

where $Z$ is a normalizing constant and

$$f(u, v) = \int_{\mathbb{R}^{d(N-3)}} e^{-\frac{2}{N}\left[\sum_{4 \leqslant i<j}^{N} \Phi(|\mathbf{r}_{1i} - \mathbf{r}_{1j}|) + \sum_{l=4}^{N}\left[\Phi(|\mathbf{r}_{1l}|) + \Phi(|u - \mathbf{r}_{1l}|) + \Phi(|v - \mathbf{r}_{1l}|)\right]\right]} d\mathbf{r}_{14} \ldots \mathbf{r}_{1N}.$$

We expect to remove the constraint $N = 3$ in future research. Also, when the system starts from initial distributions other than the invariant measure, the density of $(\boldsymbol{X}_1^t - \boldsymbol{X}_2^t, \boldsymbol{X}_1^t - \boldsymbol{X}_3^t)$ will no longer be $p(u, v)$. Perturbation type arguments may help to address such cases.

### 4.2. *Interaction potentials in form of $\Phi(r) = r^{2\beta}$*

We develop in this section a "comparison to Gaussian kernels" technique to prove that the coercivity condition holds true for systems with interaction potential $\Phi(r) = r^{2\beta}$ for $1/2 \leqslant \beta \leqslant 1$. We first prove that $p(u, v)$ in (1.5) is positive-definite, then prove that the integral kernel $K_t$ is strictly positive-definite by comparing it with the Gaussian kernel.

**Lemma 4.5.** *Assume $\Phi(r) = r^{2\beta}$. Let $p(u, v)$ be the density function defined in (1.5).*

*(1) If $0 < \beta \leqslant 1$, all the three kernels, $\Phi(|u - v|)$, $e^{-\Phi(|u-v|)}$, and $p(u, v)$, are positive-definite.*
*(2) If $\beta > 1$, then $p(u, v)$ is not positive-definite.*

**Proof.** This is a generalization of Corollary 3.3.3 of [2] to the higher-dimensional case. Note that $p(u, v)$ is positive-definite iff $e^{-\Phi(|u-v|)}$ is positive-definite. The kernel $|u - v|^2$ for $u, v \in \mathbb{R}^d$ is a negative definite kernel, because for any $c_1, \ldots, c_n \in \mathbb{R}$, and $\sum_{i=1}^n c_i = 0$,

$$\sum_{i,j=1}^n c_i c_j |u_i - u_j|^2 = \left[\sum_{i=1}^n c_i\right]\left[\sum_{j=1}^n c_j |u_j|^2\right] + \left[\sum_{j=1}^n c_j\right]\left[\sum_{i=1}^n c_j |u_i|^2\right] - \left|\sum_{i=1}^n c_i u_i\right|^2$$
$$= -\left|\sum_{i=1}^n c_i u_i\right|^2 \leqslant 0.$$

By Theorem A.6, $|u - v|^{2\beta}$ is also a negative definite kernel for any $0 < \beta \leqslant 1$. By Theorem A.5, we obtain that $e^{-|u-v|^{2\beta}}$ is positive-definite, then Part (1) follows.

Now we prove Part (2). Suppose now that for some $\beta > 1$, $p(u, v)$ is a positive-definite kernel. Then for any $t > 0$, $x_1, \ldots, x_n \in \mathbb{R}^d$ and $c_1, \ldots, c_n \in \mathbb{R}$, we have

$$\sum_{j,k=1}^n c_j c_k e^{-t|x_j - x_k|^{2\beta}} = \sum_{j,k=1}^n c_j c_k e^{-\left|t^{\frac{1}{2\beta}} x_j - t^{\frac{1}{2\beta}} x_k\right|^{2\beta}} \geqslant 0$$

By Theorem A.5, the kernel $|u - v|^{2\beta}$ is negative definite, and by Theorem A.7, $|u - v|^{\beta}$ is a metric on $\mathbb{R}$. Let $\mathbf{0} = (0, \ldots, 0) \in \mathbb{R}^d$, $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^d$ and $\mathbf{2} = (2, \ldots, 2) \in \mathbb{R}^d$. Note that

$$|\mathbf{0} - \mathbf{1}|^{\beta} = d^{\frac{\beta}{2}}, \quad |\mathbf{0} - \mathbf{2}|^{\beta} = 2^{\beta} d^{\frac{\beta}{2}} > 2|\mathbf{0} - \mathbf{1}|^{\beta}$$

when $\beta > 1$. The contradiction to the triangle inequality implies Part (2). ■

Recall that the coercivity condition depends only on the distribution of the process $(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t)$. When the process $(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t)$ is stationary, the coercivity condition at a time instance in Definition 2.4 is equivalent to that of Definition 1.2. Following Proposition 2.5, the coercivity condition is equivalent to the positivity of the integral operator $Q$ on $L^2(\rho)$ with kernel $K(r, s) : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ defined by

$$K(r, s) := \frac{1}{q(r)q(s)}(rs)^{d-1} \int_{S^{d-1}} \int_{S^{d-1}} \langle \xi, \eta \rangle p(r\xi, s\eta) d\xi d\eta, \tag{4.3}$$

where $p(u, v)$ is the stationary density defined in (1.5), and $\rho$ denotes the distribution of $|\boldsymbol{r}_{12}^t|$ with $q$ being its density. For the case $\beta = 1$ in the previous section, we witnessed that the Gaussian distribution neatly ensures strict positivity of the integral operator through Taylor expansion of $\langle u, v \rangle e^{c\langle u, v \rangle}$. However, when $\beta \neq 1$, such a "quadratic structure" of the Gaussian kernel is no longer available. Using the positive-definiteness of $-|u - v|^{2\beta}$, we uncover a quadratic structure by bounding the kernel $K$ by another positive-definite kernel from below and by a Gamma integral representation of the power function. We call this procedure a "comparison to a Gaussian kernel" technique.

We start with two inequalities using positive-definiteness of integral kernels.

**Lemma 4.6.** *Let $\Phi_i : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be positive-definite kernels for $i = 1, 2$. Then,*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(u)h(v)\Phi_1(u,v)e^{\Phi_2(u,v)}dudv \geqslant \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(u)h(v)\Phi_1(u,v)dudv,$$
$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(u)h(v)\Phi_1(u,v)e^{\Phi_2(u,v)}dudv \geqslant \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(u)h(v)\Phi_1(u,v)\Phi_2(u,v)dudv$$

*for any $h \neq 0$, as long as the integrals exist.*

**Proof.** By Theorem A.2, $\Phi_2(u,v)^n\Phi_1(u,v)$ is positive-definite for each integer $n \geqslant 0$, so for any $h \neq 0$ and $n \geqslant 0$, we have $\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(u)h(v)\Phi_1(u,v)\Phi_2(u,v)^n dudv > 0$. Then the inequalities follow from the Taylor expansion of $e^{\Phi_2(u,v)}$, with the first inequality keeping only the term with $n = 0$ and the second inequality keeping only the term with $n = 1$. ■

The following proposition is a counterpart of Proposition 3.8.

**Proposition 4.7.** *Let $\beta \in (0, 1]$ and $p(u, v)$ be a density function in (1.5) with $\Phi(r) = r^{2\beta}$, i.e., $p(u,v) = \frac{1}{Z}e^{-\frac{2}{3}(|u|^{2\beta}+|v|^{2\beta}+|u-v|^{2\beta})}$. Let $\rho(r)$ be the distribution of $|U|$ with $(U, V)$ having a joint distribution $p(u, v)$. Then, for any $0 \neq h \in L^2(\rho)$,*

$$I = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(|u|)h(|v|)\frac{\langle u,v\rangle}{|u||v|}p(u,v)dudv > 0.$$

**Proof.** The factor $\frac{2}{3}$ and the normalizing constant $Z$ do not play a role in the above inequality, so we omit them in the following proof. We only consider the case $\beta < 1$, since when $\beta = 1$, the Gaussian distribution neatly ensures strict positivity of the integral operator through Taylor expansion of $\langle u, v\rangle e^{c\langle u,v\rangle}$. Note that

$$\begin{aligned} I &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(|u|)e^{-2|u|^{2\beta}}h(|v|)e^{-2|v|^{2\beta}}\frac{\langle u,v\rangle}{|u||v|}e^{|u|^{2\beta}+|v|^{2\beta}-|u-v|^{2\beta}}dudv \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \widetilde{h}(|u|)\widetilde{h}(|v|)\frac{\langle u,v\rangle}{|u||v|}e^{\Phi_2(u,v)}dudv, \end{aligned}$$

where $\widetilde{h}(r) := h(r)e^{-2r^{2\beta}}$ and

$$\Phi_2(u,v) := |u|^{2\beta} + |v|^{2\beta} - |u-v|^{2\beta}.$$

By Lemma 4.5, $|u-v|^{2\beta}$ is negative definite. Then, by Theorem A.4, $\Phi_2(u,v)$ is positive-definite. Thus, by Lemma 4.6 with $\Phi_1(u,v) = \langle u,v\rangle$ and $\Phi_2(u,v)$ as above, we have

$$\begin{aligned} I &\geqslant \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \widetilde{h}(|u|)\widetilde{h}(|v|)\frac{\langle u,v\rangle}{|u||v|}\left(|u|^{2\beta}+|v|^{2\beta}-|u-v|^{2\beta}\right)dudv \\ &= -\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \widetilde{h}(|u|)\widetilde{h}(|v|)\frac{\langle u,v\rangle}{|u||v|}|u-v|^{2\beta}dudv =: \widetilde{I}, \end{aligned} \tag{4.4}$$

where in the equality we dropped the term $|u|^\beta + |v|^\beta$, due to symmetry of $\langle u, v\rangle$,

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g_1(|u|)g_2(|v|)\frac{\langle u,v\rangle}{|u||v|} = 0$$

for any $g_1, g_2 \in L^2(\rho)$. We shall use this property several times in the following.

Next, we use Gamma function to bound $\widetilde{I}$ in (4.4) from below by a Gaussian kernel as in the previous section. Note that for any $x > 0$ and $\beta < 1$,

$$|u-v|^{2\beta} = \frac{\beta}{\Gamma(1-\beta)}\int_0^\infty (1-e^{-\lambda|u-v|^2})\frac{d\lambda}{\lambda^{\beta+1}}.$$

Plugging this into the integral in (4.4), and using the symmetry of $\langle u, v\rangle$ again, we obtain

$$\begin{aligned}\widetilde{I} &= \frac{\beta}{\Gamma(1-\beta)}\int_0^\infty\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\widetilde{h}(|u|)\widetilde{h}(|v|)\frac{\langle u, v\rangle}{|u||v|}(e^{-\lambda|u-v|^2}-1)dudv\frac{d\lambda}{\lambda^{\beta+1}}\\ &= \frac{\beta}{\Gamma(1-\beta)}\int_0^\infty\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\widetilde{h}(|u|)\widetilde{h}(|v|)\frac{\langle u, v\rangle}{|u||v|}e^{-\lambda|u-v|^2}dudv\frac{d\lambda}{\lambda^{\beta+1}}\\ &= \frac{\beta}{\Gamma(1-\beta)}\int_0^\infty\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\widetilde{h}(|u|)\widetilde{h}(|v|)e^{-\lambda(|u|^2+|v|^2)}\frac{\langle u, v\rangle}{|u||v|}e^{2\lambda\langle u,v\rangle}dudv\frac{d\lambda}{\lambda^{\beta+1}}.\end{aligned}$$

By the symmetry of $\langle u, v\rangle$ and Taylor expansion of $e^{2\lambda\langle u,v\rangle}$, we have

$$\begin{aligned}\widetilde{I} &= \frac{\beta}{\Gamma(1-\beta)}\int_0^\infty\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\widetilde{h}(|u|)\widetilde{h}(|v|)e^{-\lambda(|u|^2+|v|^2)}\frac{\langle u, v\rangle}{|u||v|}\\ &\qquad\times\sum_{n=0}^\infty\frac{\langle u, v\rangle^{2n+1}}{(2n+1)!}dudv2^{2n+1}\lambda^{2n-\beta}d\lambda\\ &= \sum_{n=0}^\infty\frac{\beta}{\Gamma(1-\beta)}\int_0^\infty\int_0^\infty\int_0^\infty\widetilde{h}(r)\widetilde{h}(s)e^{-\lambda(r^2+s^2)}(rs)^{d+2n+1}C_n drds\lambda^{2n-\beta}d\lambda,\end{aligned}$$

where we denote $C_n := \int_{\xi\in S^{d-1}}\int_{\eta\in S^{d-1}}\frac{2^{2n+1}\langle\xi,\eta\rangle^{2n+2}}{(2n+1)!}d\xi d\eta$, which is positive. Thus,

$$\begin{aligned}\widetilde{I} &= \sum_{n=0}^\infty\frac{C_n\beta}{\Gamma(1-\beta)}\int_0^\infty\lambda^{2n-\beta}\left[\int_0^\infty\widetilde{h}(r)e^{-\lambda r^2}r^{d+2n+1}dr\right]^2d\lambda\\ &\geqslant \sum_{n=0}^\infty\frac{C_n\beta}{\Gamma(1-\beta)}\int_1^2\lambda^{2n-\beta}d\lambda\left[\int_0^\infty\widetilde{h}(r)e^{-\lambda r^2}r^{d+2n+1}dr\right]^2.\end{aligned}$$

Note that $\tilde{C}_{n,\beta} = \frac{C_n\beta}{\Gamma(1-\beta)}\int_1^2\lambda^{2n-\beta}d\lambda > 0$ for each $n \geqslant 0$. Combining all the above, we have

$$I \geqslant \sum_{n=0}^\infty\tilde{C}_{n,\beta}\left[\int_0^\infty h(r)e^{-2r^{2\beta}-2r^2}r^{d+2n+1}dr\right]^2,$$

which is positive if $h \neq 0 \in L^2(f)$ with $f(r) := r^{d+1}e^{-2r^{2\beta}-2r^2}$, because by Lemma A.9, the set of functions span$\{1, r^2, r^4, \ldots\}$ is complete in $L^2(\mathbb{R}^+, f)$. Note that supp $f$ = supp $\rho = \mathbb{R}^+$, so $h \neq 0 \in L^2(f)$ when $h \neq 0 \in L^2(\rho)$. ■

**Proof of Theorem 4.1 with $\Phi(r) = r^{2\beta}$.** Since $\beta \in [1/2, 1]$ and the density of $(\boldsymbol{r}_{12}^0, \boldsymbol{r}_{13}^0)$ is $p(u, v)$, it follows from Proposition 4.2 that the process $(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t)$ is stationary with distribution $p(u, v)$. Then, the coercivity condition is equivalent to that

$$I := \mathbb{E}[h(|\boldsymbol{r}_{12}^t|)h(|\boldsymbol{r}_{13}^t|)\frac{\langle\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t\rangle}{|\boldsymbol{r}_{12}^t||\boldsymbol{r}_{13}^t|}] > 0$$

for any $h \neq 0 \in L^2(\rho)$, where $\rho$ is the stationary probability density of $|\boldsymbol{r}_{12}^t|$. Note that

$$I = \frac{1}{Z}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(|u|)h(|v|)\frac{\langle u, v\rangle}{|u||v|}e^{-\frac{2}{3}(|u|^{2\beta}+|v|^{2\beta}+|u-v|^{2\beta})}dudv.$$

Then we can conclude the theorem by Proposition 4.7. ■

**Remark 4.8.** We point out that the requirement $\beta \in (0, 1]$ is to ensure that the stationary density $p(u, v)$ is a positive-definite kernel. When $\beta > 1$, the above method does no longer work, because $p(u, v)$ is not positive-definite as shown in Lemma 4.5. The requirement $\beta \geqslant \frac{1}{2}$ is to ensure that the drift term is continuous, so that a strong solution exists. When $\beta < \frac{1}{2}$, the drift is moderately singular, the existence of a solution is open [1,34].

### *4.3. General interaction potentials*

The "comparison to a Gaussian kernel" technique in Lemma 4.6 and Proposition 4.7 can be generalized to prove the coercivity condition for a large class of interaction functions. The following lemma provides the key element in such a generalization.

**Lemma 4.9.** *Let $\Phi$ be a potential in (1.4) and let $\rho(r)$ be the distribution of $|U|$ with $(U, V)$ having a joint distribution $p(u, v) = \frac{1}{Z}e^{-\frac{2}{3}[\Phi(|u|)+\Phi(|v|)+\Phi(|u-v|)]}$. Then,*

$$I = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(|u|)h(|v|)\frac{\langle u, v\rangle}{|u||v|}e^{-[\Phi(|u|)+\Phi(|v|)+\Phi(|u-v|)]}dudv > 0$$

*for any $0 \neq h \in L^2(\rho)$.*

**Proof.** Rewrite the integral as

$$\begin{aligned} I &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} [h(|u|)e^{-\frac{2}{3}\Phi(|u|)}][h(|v|)e^{-\frac{2}{3}\Phi(|v|)}]\frac{\langle u, v\rangle}{|u||v|}e^{-\frac{2}{3}a|u-v|^{2\beta}-\frac{2}{3}\Phi_0(|u-v|)}dudv \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \widetilde{h}(|u|)\widetilde{h}(|v|)\frac{\langle u, v\rangle}{|u||v|}e^{-\frac{2}{3}a|u-v|^{2\beta}+\frac{2}{3}\widetilde{\Phi}(|u-v|)}dudv, \end{aligned}$$

where $\widetilde{h}(r) = h(r)e^{-\frac{2}{3}\Phi(r)-\frac{2}{3}\Phi_0(r)}$ and

$$\widetilde{\Phi}(u, v) := \Phi_0(|u|) + \Phi_0(|v|) - \Phi_0(|u - v|).$$

Since $\Phi_0(|u - v|)$ is negative definite, by Theorem A.3, $\widetilde{\Phi}(u, v)$ is positive-definite. Also, by Lemma 4.5, $\langle u, v\rangle e^{-\frac{2}{3}a|u-v|^{2\beta}}$ is positive-definite. Hence, by Lemma 4.6, we have

$$I \geqslant \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \widetilde{h}(|u|)\widetilde{h}(|v|)\frac{\langle u, v\rangle}{|u||v|}e^{-\frac{2}{3}a|u-v|^{2\beta}}dudv.$$

Then the strictly positive-definiteness follows from Proposition 4.7. ■

We can now conclude the proof of Theorem 4.1.

**Proof of Theorem 4.1.** The case when $\Phi(r) = r^{2\beta}$ is proved in the previous section. For general potential $\Phi$, the proof is similar. In fact, since $\beta \in [1/2, 1]$ and $\Phi_0$ is smooth, the process $(\boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t)$ is stationary with invariant density $p(u, v)$. Then, $\mathbb{E}[h(|\boldsymbol{r}_{12}^t|)h(|\boldsymbol{r}_{13}^t|)\frac{\langle \boldsymbol{r}_{12}^t, \boldsymbol{r}_{13}^t\rangle}{|\boldsymbol{r}_{12}^t||\boldsymbol{r}_{13}^t|}]$ is the same as the integral $I$ in Lemma 4.9, so the coercivity condition follows. We conclude the proof. ■

We provide a few examples of negative definite radial kernels and the related positive kernels.

**Lemma 4.10.** *For $0 < \alpha \leqslant 2$, $0 < \gamma \leqslant 1$ and $a \geqslant 0$, the following kernels are negative definite:*

$$\Phi_1(|u-v|) = (a + |u-v|^\alpha)^\gamma;$$
$$\Phi_2(|u-v|) = \log[1 + (a + |u-v|^\alpha)^\gamma].$$

*For any $c > 0$ and any integer $k \geqslant 1$, the following kernels are positive-definite:*

$$e^{-c\Phi_1(|u-v|)},\ e^{-c\Phi_2(|u-v|)},\ \Phi_2(|u-v|)^{-k}.$$

**Proof.** By Lemma 4.5, if $0 \leqslant \alpha \leqslant 2$, then $|u-v|^\alpha$ is a negative definite kernel. By definition of a negative definite kernel $a + |u-v|^\alpha$ is also negative definite for any $a \in \mathbb{R}$. By Theorem A.6, $\Phi_1(|u-v|) = (a + |u-v|^\alpha)^\gamma$ is also a negative definite kernel when $0 < \gamma \leqslant 1$ and $a \geqslant 0$.

Since $\Phi_1(|u-u|) = a^\gamma \geqslant 0$, by Theorem A.6, $\log(1 + \Phi_1(|u-v|)) = \Phi_2(|u-v|)$ is negative definite.

The positive-definiteness of $e^{-c\Phi_1(|u-v|)}$ and $e^{-c\Phi_2(|u-v|)}$ follows directly from Theorem A.5. The kernel $\Phi_2(|u-v|)^{-k}$ is positive-definite because

$$\int_0^\infty e^{-s\Phi_2(|u-v|)} ds = \frac{1}{\Phi_2(|u-v|)}$$

and because that the product of positive-definite kernels are positive-definite. ■

**Proposition 4.11.** *Assume that the series*

$$\Phi_1(r) = c_0 + \sum_{j=1}^{\infty} c_j \log\left[1 + (a_j + r^{\alpha_j})^{\gamma_j}\right] - \sum_{j=-1}^{-\infty} c_j [\log(1 + (a_j + r^{\alpha_j})^{\gamma_j})]^{-k_j} \tag{4.5}$$

$$\Phi_2(r) = \sum_{i=1}^{\infty} c_i' [(a_i' + r^{\alpha_i'})^{\gamma_i'}] - \sum_{i=-1}^{-\infty} c_i' [1 + (a_i' + r^{\alpha_i'})^{\gamma_i'}]^{-\beta i} \tag{4.6}$$

*converge for every $r \in \mathbb{R}^+$, where the coefficients satisfy the following conditions*

1. *$a_j \geqslant 0, a_i' \geqslant 0, c_j \geqslant 0, c_i' \geqslant 0$ for $i, j \neq 0$ and*
2. *$0 < \gamma_i \leqslant 1$, $\alpha_j, \alpha_i' \in [1, 2]$ for $i, j \neq 0$, and*
3. *$\beta_j > 0$ and $k_j \geqslant 1$ is a positive integer for each $j$.*

*Let $K : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ be an integral kernel defined in (4.3) with $p(u, v)$ defined in (1.5) and with*

$$\Phi(r) = \Phi_1(r) + \Phi_2(r). \tag{4.7}$$

*Then $K(r, s)$ is a positive-definite kernel. Furthermore, if there exists $i_0 \geqslant 1$, such that*

$$a_{i_0}' = 0,\ \gamma_{i_0}' = 1,\ \text{and } c_{i_0}' > 0, \tag{4.8}$$

*then the coercivity condition holds true on $L^2(\rho)$ for the system (2.1) with potential $\Phi$ in (4.7), if it starts from $(\boldsymbol{r}_{12}^0, \boldsymbol{r}_{13}^0)$ with a joint density $p(u, v)$ in (1.5), where $\rho$ is the distribution of $|\boldsymbol{r}_{12}^0|$.*

**Proof.** It follows directly from Lemma 4.10 that $K$ is positive-definite. Note that with the above conditions, the drift term is smooth and dominated by a term $r^{2\beta}$ with $\beta = \alpha'_{i_0}/2 \in [1/2, 1]$, so the system leads to a stationary process $(\boldsymbol{r}^t_{12}, \boldsymbol{r}^t_{13})$. It follows from Lemma 4.9 that the coercivity condition holds true. ■

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

We thank the two anonymous reviewers whose comments helped improve this manuscript. ZL is grateful for support from National Science Foundation (NSF), USA-1608896 and Simons, USA-638143; FL and MM are grateful for partial support from NSF, USA-1913243; FL for NSF, USA-1821211; MM for NSF, USA-1837991, NSF, USA-1934979, NSF, USA-2031985 and Air Force Office of Scientific Research (AFOSR), USA-FA9550-17-1-0280 and the Simons Fellowship; ST for the American Mathematics Society Simons travel grant. FL would like to thank Professor Yaozhong Hu and Dr. Yulong Lu for helpful discussions.

## Appendix

### *A.1. Positive-definite integral kernels*

In this section, we review the definitions of positive and negative definite kernels, as well as their basic properties. The following definition is a real version of the definition in [2, p. 67].

**Definition A.1.** Let $\mathbb{X}$ be a nonempty set. A function $K : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ is called a (real) positive-definite kernel iff it is symmetric (i.e. $K(x, y) = K(y, x)$) and

$$\sum_{j,k=1}^{n} c_j c_k K(x_j, x_k) \geqslant 0 \tag{A.1}$$

for all $n \in \mathbb{N}$, $\{x_1, \ldots, x_n\} \in \mathbb{X}$ and $\{c_1, \ldots, c_n\} \in \mathbb{R}$. We call the function $\phi$ a (real) negative definite kernel iff it is symmetric and

$$\sum_{j,k=1}^{n} c_j c_k K(x_j, x_k) \leqslant 0 \tag{A.2}$$

for all $n \geqslant 2$, $\{x_1, \ldots, x_n\} \in \mathbb{X}$ and $\{c_1, \ldots, c_n\} \in \mathbb{R}$ with $\sum_{j=1}^{n} c_j = 0$.

**Remark.** In [2, p. 67], a function $K : \mathbb{X} \times \mathbb{X} \to \mathbb{C}$ is defined to be positive-definite iff

$$\sum_{j,k=1}^{n} c_j \bar{c}_k K(x_j, x_k) \geqslant 0 \tag{A.3}$$

for all $n \in \mathbb{N}$, $\{x_1, \ldots, x_n\} \in \mathbb{X}$ and $\{c_1, \ldots, c_n\} \in \mathbb{C}$, where $\bar{c}$ denotes the complex conjugate of a complex number $c$. It is straightforward to check that when $\phi$ is real-valued

and symmetric, the definitions (A.1) and (A.3) are equivalent. Similarly, In the definition of negative definiteness in [2, p. 67], a function $K : \mathbb{X} \times \mathbb{X} \to \mathbb{C}$ is negative definite iff it is Hermitian (i.e. $K(x, y) = \overline{K(y, x)}$) and

$$\sum_{j,k=1}^{n} c_j \bar{c}_k K(x_j, x_k) \leqslant 0 \tag{A.4}$$

for all $n \geqslant 2$, $\{x_1, \ldots, x_n\} \in \mathbb{X}$ and $\{c_1, \ldots, c_n\} \in \mathbb{C}$ with $\sum_{j=1}^{n} c_j = 0$. We can again check that when $\phi$ is real-valued, the definitions (A.2) and (A.4) are equivalent. In this paper, we only consider real-valued, symmetric kernels.

**Theorem A.2** (*Properties of Positive-definite Kernels*)**.** *Suppose that $k, k_1, k_2 : \mathbb{X} \times \mathbb{X} \subset \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ are positive-definite kernels. Then*

1. *$c_1 k_1 + c_2 k_2$ is positive-definite, for $c_1, c_2 \geqslant 0$*
2. *$k_1 k_2$ is positive-definite. ([2, p. 69])*
3. *$\exp(k)$ is positive-definite. ([2, p. 70])*
4. *$k(f(u), f(v))$ is positive-definite for any map $f : \mathbb{R}^d \to \mathbb{R}^d$*
5. *Inner product $\langle u, v\rangle = \sum_{j=1}^{d} u_j v_j$ is positive-definite ([2, p. 73])*
6. *$f(u)f(v)$ is positive-definite for any function $f : \mathbb{X} \to \mathbb{R}$ ([2, p. 69]).*
7. *If $k(u, v)$ is measurable and integrable, then $\iint k(u, v)du dv \geqslant 0$ ([32, p. 524])*

**Theorem A.3** (*[2, Theorem 3.1.17]*)**.** *Let $K : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ be symmetric. Then $K$ is positive-definite iff*

$$\det(K(x_j, x_k)_{j,k\leqslant n}) \geqslant 0$$

*for all $n \in \mathbb{N}$ and all $\{x_1, \ldots, x_n\} \subseteq \mathbb{X}$.*

**Theorem A.4** (*[2, Lemma 3.2.1]*)**.** *Let $\mathbb{X}$ be a nonempty set, $x_0 \in \mathbb{X}$ and let $\psi : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ be a symmetric kernel. Put $K(x, y) := \psi(x, x_0) + \psi(y, x_0) - \psi(x, y) - \psi(x_0, x_0)$. Then $K$ is positive-definite iff $\psi$ is negative definite.*

**Theorem A.5.** *Let $\mathbb{X}$ be a nonempty set and let $\psi : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ be a kernel. Then $\psi$ is negative definite iff $\exp(-t\psi)$ is positive-definite for all $t > 0$.*

**Proof.** The complex version of this theorem is proved in Theorem 3.2.2 of [2]. The real version can be proved in a similar way. ■

**Theorem A.6.** *If $\psi : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ is negative definite and $\psi(x, x) \geqslant 0$, then so are $\psi^\alpha$ for $0 < \alpha < 1$ and $\log(1 + \psi)$.*

**Proof.** The complex version of this theorem is proved in Theorem 3.2.10 of [2]. The real version can be proved in a similar way. ■

**Theorem A.7** (*[2, Proposition 3.3.2]*)**.** *Let $\mathbb{X}$ be nonempty and $\psi : \mathbb{X} \times \mathbb{X} \to \mathbb{C}$ be negative definite. Assume $\{(x, y) \in \mathbb{X} \times \mathbb{X}, \psi(x, y) = 0\} = \{(x, x) : x \in \mathbb{X}\}$, then $\sqrt{\psi}$ is a metric on $\mathbb{X}$.*

### A.2. Müntz-type theorems on half-line

We recall first the following theorem on the completeness of $\{t^{a_n}\}$ in weighted $L^2$ space on unbounded domain (see [4,14] and see [15,18] for recent developments).

**Theorem A.8.** *Let $a_k$ be positive numbers, such that $a_{k+1} - a_k \geqslant d > 0$, $(k = 1, 2, \ldots)$, and let*

$$\log \psi(r) = \begin{cases} 2\sum_{a_k<r} \frac{1}{a_k}, & \text{if } r > a_1 \\ \frac{2}{a_1}, & \text{if } r \leqslant a_1. \end{cases}$$

*Then $\{e^{-t}t^{a_k}\}$ is complete in $L^2(0,\infty)$ iff $\int_1^\infty \frac{\psi(r)}{r} dr = \infty$.*

**Lemma A.9.** *The set of functions $\{r^{2k}, k = 1, 2, \ldots\}$ is complete in $L^2([0,\infty), \rho)$ for any probability density $\rho$ such that $\sup_{r>0} \rho(r)e^{2r} < \infty$.*

**Proof.** Let $a_k = 2k$ for $k = 1, 2, \ldots$. We define the function $\log(\psi(r)) = 2\sum_{a_k<r} \frac{1}{a_k}$, if $r > a_1$, and $\log(\psi(r)) = \frac{2}{a_1}$ if $r \leqslant a_1$. Note that $2\sum_{a_k<r} \frac{1}{a_k} = \sum_{k=1}^{\lfloor r/2 \rfloor} \frac{1}{k} > \ln(\lfloor r/2 \rfloor)$. Then $\psi(r) \geqslant r$ and $\int_1^\infty \frac{\psi(r)}{r^2} = \infty$. We conclude that $\{e^{-t}t^{2k}, k = 1, 2, \ldots\}$ is complete in $L^2(0,\infty)$ by Theorem A.8.

To show that $\{r^{2k}, k = 1, 2, \ldots\}$ is complete in $L^2(\rho)$, assume that $\langle h(r), r^{2k}\rangle_{L^2(\rho)} = 0$ for all $k \geqslant 1$. Then

$$\int_0^\infty h(r)\rho(r)e^r r^{2k} e^{-r} dr = \int_0^\infty h(r) r^{2k} \rho(r) dr = 0$$

for all $k$. This implies that $h(r)\rho(r)e^r = 0$ in $L^2[0,\infty)$ (note that $h(r)\rho(r)e^r \in L^2[0,\infty)$ because $\sup_{r>0} \rho(r)e^{2r} < \infty$). Hence $h(r)\rho(r) = 0$ almost everywhere, and $h = 0$ in $L^2([0,\infty), \rho)$. ■

### A.3. Stationary measure for a gradient system

**Lemma A.10.** *Suppose $H : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz and $\lim_{|x|\to\infty} H(x) = +\infty$ so that $Z = \int_{\mathbb{R}^n} e^{-2H(x)} dx < \infty$. Then $p(x) = \frac{1}{Z} e^{-2H(x)}$ is an invariant density to gradient system*

$$dX_t = -\nabla H(X_t)dt + dB_t,$$

*where $(B_t)$ is an n-dimensional standard Brownian motion.*

**Proof.** It follows directly by showing that $p(x)$ is a stationary solution to the Kolmogorov forward equation,

$$\frac{1}{2}\Delta p + \nabla \cdot (p \nabla H) = 0. \quad ■$$

## References

[1] S. Albeverio, Y. Kondratiev, M. Röckner, Strong Feller properties for distorted Brownian motion and applications to finite particle systems with singular interactions, in: H.-H. Kuo, A.N. Sengupta (Eds.), Contemporary Mathematics, Vol. 317, American Mathematical Society, Providence, Rhode Island, 2003, pp. 15–35.

[2] C. Berg, J.P.R. Christensen, P. Ressel, Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions, Vol. 100, Springer, New York, 1984.
[3] A.J. Bernoff, C.M. Topaz, Nonlocal aggregation models: A primer of swarm equilibria, SIAM Rev. 55 (4) (2013) 709–747.
[4] R. Boas, H. Pollard, Properties equivalent to the completeness of $\{e^{-t}t^{\lambda_n}\}$, Bull. Amer. Math. Soc. 52 (4) (1946) 348–351.
[5] M. Bongini, M. Fornasier, M. Hansen, M. Maggioni, Inferring interaction rules from observations of evolutive systems I: The variational approach, Math. Models Methods Appl. Sci. 27 (05) (2017) 909–951.
[6] C. Brugna, G. Toscani, Kinetic models of opinion formation in the presence of personal conviction, Phys. Rev. E 92 (5) (2015) 052818.
[7] J.A. Carrillo, Y.-P. Choi, S.P. Perez, A review on attractive–repulsive hydrodynamics for consensus in collective behavior, in: N. Bellomo, P.T.E. Degond (Eds.), Active Particles, Vol. 1, Birkhäuser, Cham, 2017, pp. 259–298.
[8] J. Casadiego, M. Nitzan, S. Hallerberg, M. Timme, Model-free inference of direct network interactions from nonlinear collective dynamics, Nature Commun. 8 (1) (2017) 2192.
[9] P. Cattiaux, F. Delebecque, L. Pédèches, et al., Stochastic cucker–smale models: Old and new, Ann. Appl. Probab. 28 (5) (2018) 3239–3286.
[10] P. Cattiaux, A. Guillin, F. Malrieu, Probabilistic approach for granular media equations in the non-uniformly convex case, Probab. Theory Related Fields 140 (1–2) (2008) 19–40.
[11] F. Cucker, S. Smale, On the mathematical foundations of learning, Bull. Amer. Math. Soc. 39 (1) (2002) 1–49.
[12] M.R. D'Orsogna, Y.-L. Chuang, A.L. Bertozzi, L.S. Chayes, Self-propelled particles with soft-core interactions: patterns, stability, and collapse, Phys. Rev. Lett. 96 (2006) 104–302.
[13] G.E. Fasshauer, Positive definite kernels: past, present and future, Dolomite Res. Notes Approx. 4 (2011) 21–63.
[14] W. Fuchs, On the closure of $\{e^{-t}t^{a_\nu}\}$, Math. Proc. Camb. Phil. Soc. 22 (1946) 91–105.
[15] Z. Gao, G. Deng, The generalized bernstein problem on weighted lacunary polynomial approximation, J. Approx. Theory 136 (1) (2005) 108–114.
[16] S.-Y. Ha, K. Lee, D. Levy, et al., Emergence of time-asymptotic flocking in a stochastic cucker-smale system, Commun. Math. Sci. 7 (2) (2009) 453–469.
[17] J.E. Herbert-Read, A. Perna, R.P. Mann, T.M. Schaerf, D.J. Sumpter, A.J. Ward, Inferring the rules of interaction of shoaling fish, Proc. Natl. Acad. Sci. USA 108 (46) (2011) 18726–18731.
[18] A.P. Horváth, Müntz-type theorems on the half-line with weights, J. Math. Anal. Appl. 410 (2) (2014) 699–712.
[19] H. Huang, J.-G. Liu, J. Lu, Learning interacting particle systems: Diffusion parameter estimation for aggregation equations, Math. Models Methods Appl. Sci. 29 (01) (2019) 1–29.
[20] P.-E. Jabin, Z. Wang, Mean field limit for stochastic particle systems, in: Active Particles, Volume 1, Springer, 2017, pp. 379–402.
[21] I. Karatzas, S.E. Shreve, Brownian motion, in: Brownian Motion and Stochastic Calculus, second ed., Springer, 1998.
[22] R. Khasminskii, Stochastic Stability of Differential Equations, Vol. 66, second ed., Springer-Verlag, Berlin Heidelberg, 2012.
[23] Y.A. Kutoyants, Statistical Inference for Ergodic Diffusion Processes, Springer, 2004.
[24] L. Li, J.-G. Liu, P. Yu, On the mean field limit for brownian particles with coulomb interaction in 3d, J. Math. Phys. 60 (11) (2019) 111501.
[25] J.-G. Liu, R. Yang, Propagation of chaos for large brownian particle system with coulomb interaction, Res. Math. Sci. 3 (1) (2016) 40.
[26] F. Lu, M. Maggioni, S. Tang, Learning interaction kernels in heterogeneous systems of agents from multiple trajectories, arXiv:1910.04832.
[27] F. Lu, M. Maggioni, S. Tang, Learning interaction kernels in stochastic systems of interacting particles from multiple trajectories, arXiv:2007.15174.
[28] F. Lu, M. Zhong, S. Tang, M. Maggioni, Nonparametric inference of interaction laws in systems of agents from trajectory data, Proc. Natl. Acad. Sci. USA 116 (29) (2019) 14424–14433.
[29] F. Malrieu, Convergence to equilibrium for granular media equations and their euler schemes, Ann. Appl. Probab. 13 (2) (2003) 540–560.
[30] S. Méléard, Asymptotic behaviour of some interacting particle systems; mckean-vlasov and boltzmann models, in: Probabilistic Models for Nonlinear Partial Differential Equations, Springer, 1996, pp. 42–95.

[31] S. Motsch, E. Tadmor, Heterophilious dynamics enhances consensus, SIAM Rev. 56 (4) (2014) 577–621.
[32] S.T. Rachev, L. Klebanov, S.V. Stoyanov, F. Fabozzi, The Methods of Distances in the Theory of Probability and Statistics, Springer Science & Business Media, 2013.
[33] B. Schmuland, Dirichlet forms with polynomial domain, Math. Jpn. 37 (1992) 1015–1024.
[34] A.V. Skorokhod, On the regularity of many-particle dynamical systems perturbed by white noise, J. Appl. Math. Stoch. Anal. 9 (4) (1996) 427–437.
[35] S. Smale, D.-X. Zhou, Geometry on probability spaces, Constr. Approx. 30 (3) (2009) 311.
[36] T. Vicsek, A. Zafeiris, Collective motion, Phys. Rep. 517 (3–4) (2012) 71–140.