

# A Controlled Mean Field Model for Chiplet Population Dynamics

Iman Nodozi, Abhishek Halder, Ion Matei

**Abstract**—In micro-assembly applications, ensemble of chiplets immersed in a dielectric fluid are steered using dielectrophoretic forces induced by an array of electrode population. Generalizing the finite population deterministic models proposed in prior works for individual chiplet position dynamics, we derive a controlled mean field model for a continuum of chiplet population in the form of a nonlocal, nonlinear partial differential equation. The proposed model accounts for the stochastic forces as well as two different types of nonlocal interactions, viz. chiplet-to-chiplet and chiplet-to-electrode interactions. Both of these interactions are nonlinear functions of the electrode voltage input. We prove that the deduced mean field evolution can be expressed as the Wasserstein gradient flow of a Lyapunov-like energy functional. With respect to this functional, the resulting dynamics is a gradient descent on the manifold of joint population density functions with finite second moments that are supported on the position coordinates.

## I. INTRODUCTION

This work is motivated by micro-assembly applications, such as printer systems [1], [2] and manufacturing of photovoltaic solar cells, where an array of electrodes can be used to generate spatio-temporally non-homogeneous electric potential landscapes for dynamically assembling the “chiplets”—micron sized particles immersed in dielectric fluid—into desired patterns. In such applications, the electric potentials generated by the array of electrodes induce non-uniform dielectrophoretic forces on the chiplets, thereby resulting in a population-level chiplet dynamics. The purpose of the present work is to propose a controlled mean field model for the same.

There have been several works [3]–[7] on the modeling and dielectrophoretic control of chiplet population. However, a continuum limit macroscopic dynamics that accounts for both chiplet-to-chiplet and chiplet-to-electrode nonlocal interactions, as considered herein, has not appeared before.

The mean field limit pursued here involves considering the number of chiplets and electrodes as infinity, i.e., to think both of them as continuum population. There are two reasons why this could be of interest. *First*, the continuum limit helps approximate and better understand the dynamics for large but finitely many chiplets and electrodes, which is indeed the situation in the engineering applications mentioned before. *Second*, the distributed control synthesis problem for large but finite population becomes computationally intractable,

as noted in recent works [6], [8], [9]. A controlled mean field model opens up the possibility of designing a controller in the continuum limit with optimality guarantees. Such a controller can then be applied to a large but finite population with sub-optimality bounds. We clarify here that in this work, we only present the mean field model and its properties. We leave the control synthesis problem for our follow up work.

As in prior works such as [6], we consider the chiplet dynamics in two dimensional position coordinate. Specifically, let  $\mathbf{x}(t) \in \mathbb{R}^2$  denote the position vector of a chiplet at any fixed time  $t \in [0, \infty)$ , and let

$$u : \mathbb{R}^2 \times [0, \infty) \mapsto [u_{\min}, u_{\max}] \subset \mathbb{R}$$

denote a causal deterministic control policy, i.e.,  $u = u(\mathbf{x}, t)$ . The control  $u$  represents the electrode voltage input, and in practice, the typical voltage range  $[u_{\min}, u_{\max}] = [-400, 400]$  Volt. We denote the collection of admissible control policies as  $\mathcal{U}$ . For a typical experimental set up detailing the sensing-control architecture, see [6, Sec. II].

A viscous drag force balances the controlled force vector field  $\mathbf{f}^u$  induced by the joint effect of the chiplet-to-chiplet and chiplet-to-electrode interactions. At the low Reynolds number context relevant here, the viscous drag force is proportional to  $\dot{\mathbf{x}}$ , where the proportionality constant  $\mu$  denotes the viscous coefficient of the dielectric fluid. Ignoring the acceleration due to negligible mass of a chiplet, the dynamics then takes a form

$$\underbrace{\mu \dot{\mathbf{x}}}_{\text{viscous drag force}} = \underbrace{\mathbf{f}^u}_{\text{controlled interaction force}} + \text{noise} \quad (1)$$

where the noise may result from stochastic forcing due to environmental fluctuations (e.g., dielectric fluid impurities) and/or unmodeled dynamics.

*Contributions:* In this paper, we make the following two specific contributions.

- We derive a controlled mean field dynamics (Sec. III) for the macroscopic motion of the chiplet population. The derived model is non-affine in control, and rather non-standard compared to the existing nonlocal dynamics models available in the literature.
- We establish that the derived mean field dynamics model can be understood as the Wasserstein gradient flow (Sec. IV) of a free energy functional over the manifold of chiplet population density functions.

## II. NOTATIONS AND PRELIMINARIES

**Wasserstein distance.** The Wasserstein distance  $W$  between a pair of probability density functions  $\rho_1(\mathbf{x}), \rho_2(\mathbf{y})$  (or between corresponding probability measures in general) with

Iman Nodozi is with the Department of Electrical and Computer Engineering, University of California, Santa Cruz, CA 95064, USA, inodozi@ucsc.edu.

Abhishek Halder is with the Department of Applied Mathematics, University of California, Santa Cruz, CA 95064, USA, ahalder@ucsc.edu.

Ion Matei is with the Palo Alto Research Center, Inc., Palo Alto, CA 94304, USA, imatei@parc.com.

This work was partially supported by NSF grant 2112755.

finite second moments, respectively supported on  $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{R}^d$ , is defined as

$$W(\rho_1, \rho_2) := \left( \inf_{\rho \in \Pi_2(\rho_1, \rho_2)} \int_{\mathcal{X} \times \mathcal{Y}} \|\mathbf{x} - \mathbf{y}\|_2^2 \rho(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} \right)^{\frac{1}{2}} \quad (2)$$

where  $\Pi_2(\rho_1, \rho_2)$  is the collection of all joint probability density functions  $\rho(\mathbf{x}, \mathbf{y})$  supported on the product space  $\mathcal{X} \times \mathcal{Y}$  having finite second moments,  $\mathbf{x}$  marginal  $\rho_1$ , and  $\mathbf{y}$  marginal  $\rho_2$ . As such, (2) involves an infinite dimensional linear program that goes back to the work of Kantorovich [10]. It is well-known [11, p. 208] that  $W$  is a metric on the space of probability density functions (more generally, on the space of probability measures). Under mild assumptions, the minimizing measure  $\rho^{\text{opt}}(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y}$  is supported on the graph of the optimal transport map  $T^{\text{opt}} : \mathcal{X} \mapsto \mathcal{Y}$  pushing the measure  $\rho_1(\mathbf{x}) d\mathbf{x}$  forward to  $\rho_2(\mathbf{y}) d\mathbf{y}$ . For many connections between the Wasserstein metric and theory of optimal mass transport, we refer the readers to [11], [12].

**Wasserstein gradient of a functional.** Let  $\mathcal{P}(\mathbb{R}^d)$  denote the space of all probability density functions supported over the subsets of  $\mathbb{R}^d$ , and denote the collection of probability density functions with finite second moments as  $\mathcal{P}_2(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$ . The *Wasserstein gradient* of a functional  $\Phi : \mathcal{P}_2(\mathbb{R}^d) \mapsto \mathbb{R}$ , denoted as  $\nabla^W \Phi$ , evaluated at  $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ , is given by [13, Ch. 8]

$$\nabla^W \Phi(\rho) := -\nabla \cdot \left( \rho \nabla \frac{\delta \Phi}{\delta \rho} \right) \quad (3)$$

where  $\nabla$  denotes the standard Euclidean gradient, and  $\frac{\delta}{\delta \rho}$  denotes the functional derivative w.r.t.  $\rho$ .

To exemplify the definition (3), consider the functional  $\Phi(\rho) = \int \rho \log \rho$  (negative entropy) for  $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ . Then  $\frac{\delta \Phi}{\delta \rho} = 1 + \log \rho$ ,  $\nabla(1 + \log \rho) = \nabla \rho / \rho$ , and we get  $\nabla^W \Phi(\rho) = -\nabla \cdot \nabla \rho = -\Delta \rho$ , where  $\Delta := \nabla \cdot \nabla$  denotes the Euclidean Laplacian operator.

**Other notations.** The notation  $\langle \cdot, \cdot \rangle$  is used to denote either the standard Euclidean inner product of vectors, or the  $L^2$  inner product of functions, as evident from the context. For any natural number  $n$ , we use the finite set notation  $\llbracket n \rrbracket := \{1, 2, \dots, n\}$ . The symbols  $\text{ess sup}$ ,  $\mathbb{E}$ ,  $\mathbb{P}$ ,  $\mathbf{I}_2$ ,  $\|\cdot\|_2$  and  $\|\cdot\|_\infty$  denote the essential supremum, the expectation, the probability measure, the  $2 \times 2$  identity matrix, the vector 2 and  $\infty$  norms, respectively. The symbol  $\sim$  is used as a shorthand for “follows the statistical distribution density”.

Given probability measures  $\mu_0, \mu_1$  on  $\mathbb{R}^d$ , the *total variation distance*  $\text{dist}_{\text{TV}}(\mu_0, \mu_1) := \frac{1}{2} \sup_f \left| \int f d(\mu_0 - \mu_1) \right|$  where the supremum is over all measurable  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $\|f\|_\infty \leq 1$ . For  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , we define its *Lipschitz constant*  $\|f\|_{\text{Lip}} := \sup_{\mathbf{x} \neq \mathbf{y}} \frac{|f(\mathbf{x}) - f(\mathbf{y})|}{\|\mathbf{x} - \mathbf{y}\|_2}$ , and its *bounded Lipschitz constant*  $\|f\|_{\text{BL}} := \max\{\|f\|_\infty, \|f\|_{\text{Lip}}\}$ . The *bounded Lipschitz distance* [14, Ch. 11.3] between probability measures  $\mu_0, \mu_1$  is  $\text{dist}_{\text{BL}}(\mu_0, \mu_1) := \sup_{\|f\|_{\text{BL}} \leq 1} \left| \int f d(\mu_0 - \mu_1) \right|$ . Notice that  $\text{dist}_{\text{BL}}(\mu_0, \mu_1) \leq 2 \text{dist}_{\text{TV}}(\mu_0, \mu_1)$ .

For  $\mathcal{X} \subseteq \mathbb{R}^d$ , we use  $C_b(\mathcal{X})$  to denote the space of all bounded continuous functions  $\varphi : \mathcal{X} \mapsto \mathbb{R}$ , and  $C_b^k(\mathcal{X})$  comprises those which are also  $k$  times continuously differentiable (in the sense of mixed partial derivatives of

order  $k$ ). We say that a function sequence  $\{g_n\}_{n \in \mathbb{N}}$  where  $g_n \in L^1(\mathcal{X})$ , converges weakly to a function  $g \in L^1(\mathcal{X})$ , if  $\lim_{n \rightarrow \infty} \int_{\mathcal{X}} (g_n - g) \psi = 0$  for all  $\psi \in C_b(\mathcal{X})$ . We symbolically denote the weak convergence as  $g_n \rightharpoonup g$ .

### III. CONTROLLED MEAN FIELD MODEL

In this Section, we introduce the chiplet population dynamics. Such model has its origin in the physical processes enabling silicon microchips to be manipulated by both electrophoretic and dielectrophoretic forces when they are placed in dielectric carriers such as Isopar-M [15]. These carriers have low conductivity which allows long-range Coulomb interactions. In general, the dielectrophoretic forces dominate, and they are induced by the potential energy generated by electrostatic potentials created in electrodes. The electrodes are formed by depositing nm-scale Molybdenum-Chromium (MoCr) onto a glass substrate via vapor deposition and then directly patterning them with a laser ablation tool. The electrodes are then insulated from the chiplets and dielectric fluid by thermally laminating a micrometer-scale thick perfluoroalkoxy (PFA) film. The dielectric forces act on the chiplets, while viscous drag forces proportional to their velocities oppose their motion. Due to the negligible mass of the chiplets, their acceleration can be ignored.

Let us denote the *normalized chiplet population density function* (PDF) at time  $t$  as  $\rho(\mathbf{x}, t)$ . By definition,  $\rho \geq 0$  and  $\int_{\mathbb{R}^2} \rho d\mathbf{x} = 1$  for all  $t$ .

We make the following assumptions.

- A1.** Under an admissible control policy  $u \in \mathcal{U}$ , the chiplet normalized population distribution over the two dimensional Euclidean configuration space remains absolutely continuous w.r.t. the Lebesgue measure  $d\mathbf{x}$  for all  $t \in [0, \infty)$ . In other words, the corresponding PDFs  $\rho(\mathbf{x}, t)$  exist for all  $t \in [0, \infty)$ .
- A2.** Under an admissible control policy  $u \in \mathcal{U}$ , we have  $\rho \in \mathcal{P}_2(\mathbb{R}^2)$  for all  $t$ .

The sample path dynamics of a chiplet position is governed by a controlled nonlocal vector field

$$\mathbf{f}^u : \mathbb{R}^2 \times [0, \infty) \times \mathcal{U} \times \mathcal{P}_2(\mathbb{R}^2) \mapsto \mathbb{R}^2$$

induced by a controlled *interaction potential*  $\phi^u : \mathbb{R}^2 \times \mathbb{R}^2 \times [0, \infty) \mapsto \mathbb{R}$ , i.e.,

$$\mathbf{f}^u(\mathbf{x}, t, u, \rho) := -\nabla(\rho * \phi^u), \quad (4)$$

where  $*$  denotes *generalized convolution* in the sense

$$(\rho * \phi^u)(\mathbf{x}, t) := \int_{\mathbb{R}^2} \phi^u(\mathbf{x}, \mathbf{y}, t) \rho(\mathbf{y}, t) d\mathbf{y}.$$

The superscript  $u$  in  $\phi^u$  emphasizes that the potential depends on the choice of control policy. In particular,

$$\phi^u(\mathbf{x}, \mathbf{y}, t) := \phi_{\text{cc}}^u(\mathbf{x}, \mathbf{y}, t) + \phi_{\text{ce}}^u(\mathbf{x}, \mathbf{y}, t), \quad (5a)$$

$$\phi_{\text{cc}}^u(\mathbf{x}, \mathbf{y}, t) := \frac{1}{2} C_{\text{cc}}(\|\mathbf{x} - \mathbf{y}\|_2) (\bar{u}(\mathbf{y}, t) - \bar{u}(\mathbf{x}, t))^2, \quad (5b)$$

$$\phi_{\text{ce}}^u(\mathbf{x}, \mathbf{y}, t) := \frac{1}{2} C_{\text{ce}}(\|\mathbf{x} - \mathbf{y}\|_2) (u(\mathbf{y}, t) - \bar{u}(\mathbf{x}, t))^2, \quad (5c)$$

for  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$  and

$$\bar{u}(\mathbf{x}, t) := \frac{\int_{\mathbb{R}^2} C_{ce}(\|\mathbf{x} - \mathbf{y}\|_2) u(\mathbf{y}, t) \rho(\mathbf{y}, t) d\mathbf{y}}{\int_{\mathbb{R}^2} C_{cc}(\|\mathbf{x} - \mathbf{y}\|_2) \rho(\mathbf{y}, t) d\mathbf{y}}. \quad (6)$$

The subscripts cc and ce denote the chiplet-to-chiplet and chiplet-to-electrode interactions, respectively. As before, the superscript  $u$  highlights the dependence on the choice of control policy. In (5b)-(5c),  $C_{cc}$  and  $C_{ce}$  respectively denote the chiplet-to-chiplet and chiplet-to-electrode capacitances. These capacitances can be determined using two dimensional electrostatic COMSOL<sup>®</sup> [16] simulations for a symmetric chiplet geometry. Such simulation model comprises two metal plates with dimensions defined by the chiplet and electrode geometry, surrounded by a dielectric with properties identical to those of the Isopar-M solution. The capacitances are computed from the charges that result on each conductor when an electric potential is applied to one and the other is grounded. Once the capacitance among chiplets and electrodes at different distances are computed, differentiable parameterized capacitance function approximations (e.g., linear combination of error functions) can be fitted to that data.

In words, (5a) says that the total controlled interaction potential  $\phi^u$  is a sum of the chiplet-to-chiplet interaction potential  $\phi_{cc}^u$  given by (5b), and the chiplet-to-electrode interaction potential  $\phi_{ce}^u$  given by (5c).

The expressions for (5b), (5c), (6) arise from capacitive electrical circuit abstraction that lumps the interaction between the electrodes and the chiplets. In [6, Sec. III], such an abstraction was detailed for a finite population of  $n$  chiplets and  $m$  electrodes. The expressions (5b), (5c), (6) generalize those in the limit  $n, m \rightarrow \infty$ . On the other hand, specializing (5b), (5c), (6) for a finite population  $\{\mathbf{x}_i\}_{i \in \llbracket n \rrbracket}$  with  $\rho \equiv \frac{1}{n} \sum_{i=1}^n \delta_{\mathbf{x}_i}$  where  $\delta_{\mathbf{x}_i}$  denotes the Dirac delta at  $\mathbf{x}_i \in \mathbb{R}^2$ , indeed recovers the development in [6, Sec. III].

**Remark 1.** *An immediate observation from (5) is that even though the potential  $\phi_{cc}^u$  is symmetric in  $\mathbf{x}, \mathbf{y}$ , the potential  $\phi_{ce}^u$  is not. Therefore, the overall controlled interaction potential  $\phi^u$  is not symmetric in  $\mathbf{x}, \mathbf{y}$ .*

Without loss of generality, we assume unity viscous coefficient in (1), i.e.,  $\mu = 1$  (since otherwise we can re-scale the  $\mathbf{f}^u$ ). In addition, assuming the chiplet velocity is affected by additive standard Gaussian White noise, the sample path dynamics of the  $i$ th chiplet position  $\mathbf{x}_i(t)$  then evolves as per a controlled interacting diffusion, i.e., as a Itô stochastic differential equation (SDE) with *nonlocal* nonlinear drift:

$$d\mathbf{x}_i = \mathbf{f}^u(\mathbf{x}_i, t, u, \rho) dt + \sqrt{2\beta^{-1}} d\mathbf{w}_i(t), \quad i \in \llbracket n \rrbracket, \quad (7)$$

where  $\mathbf{f}^u$  is given by (4),  $\beta > 0$  denotes inverse temperature, and  $\mathbf{w}_i(t) \in \mathbb{R}^2$  denote i.i.d. realizations of a standard Wiener process that is  $\mathcal{F}_t$ -adapted on a complete filtered probability space with sigma-algebra  $\mathcal{F}$  and associated filtration  $(\mathcal{F}_t)_{t \geq 0}$ . In particular,  $\mathcal{F}_0$  contains all  $\mathbb{P}$ -null sets and  $\mathcal{F}_t$  is right continuous.

The study of SDEs with nonlocal nonlinear drift originated in [17], and has grown into a substantial literature, see e.g., [18], [19]. In statistical physics, such models are often referred to as ‘‘propagation of chaos’’—a terminology due to

Kac [20]. A novel aspect of the model (7) w.r.t. the existing literature is that the interaction potential  $\phi^u$  has a nonlinear dependence on the control policy  $u(\mathbf{x}, t)$  as evident from (5).

#### A. Existence-Uniqueness of Solution for (7)

For a given causal control policy  $u \in \mathcal{U}$ , it is known [21, Thm. 2.4] that an interacting diffusion of the form (7) with initial condition  $\mathbf{x}_{i0} \sim \rho_0$  admits unique weak solution provided the following four conditions hold:

- (i) the drift  $\mathbf{f}^u$  is jointly Borel measurable w.r.t.  $\mathbb{R}^2 \times [0, \infty) \times \mathcal{P}(\mathbb{R}^2)$ ,
- (ii) the diffusion coefficient  $\sqrt{2\beta^{-1}} \mathbf{I}_2$  is invertible, and the driftless SDE  $d\mathbf{z}(t) = \sqrt{2\beta^{-1}} d\mathbf{w}(t)$  admits unique strong solution,
- (iii) the drift  $\mathbf{f}^u$  is uniformly bounded,
- (iv) there exists  $\kappa > 0$  such that

$$\begin{aligned} & \|\mathbf{f}^u(\mathbf{x}, t, u(\mathbf{x}, t), \rho) - \mathbf{f}^u(\mathbf{x}, t, u(\mathbf{x}, t), \tilde{\rho})\|_2 \\ & \leq \kappa \text{dist}_{\text{TV}}(\rho, \tilde{\rho}) \quad \text{uniformly in } (\mathbf{x}, t) \in \mathbb{R}^2 \times [0, \infty). \end{aligned}$$

We assume that the capacitances  $C_{cc}, C_{ce}$  in (5)-(6) are sufficiently smooth, and the control  $u$  can be parameterized to ensure smoothness for guaranteeing that  $\nabla_{\mathbf{x}} \phi_{cc}^u, \nabla_{\mathbf{x}} \phi_{ce}^u$  (and thus  $\nabla_{\mathbf{x}} \phi^u$ ) are  $\|\cdot\|_2$  Lipschitz and uniformly bounded.

As  $\nabla_{\mathbf{x}} \phi^u$  is bounded,  $\mathbf{f}^u = \int_{\mathbb{R}^2} \nabla_{\mathbf{x}} \phi^u(\mathbf{x}, \mathbf{y}, t) \rho(\mathbf{y}) d\mathbf{y}$ , which being an average of Lipschitz, is itself Lipschitz and thus continuous. Since  $\mathbf{f}^u$  is continuous, the preimage of any Borel set in  $\mathbb{R}^2$  under  $\mathbf{f}^u$  is a measurable set in  $\mathbb{R}^2 \times [0, \infty) \times \mathcal{U} \times \mathcal{P}_2(\mathbb{R}^2)$ . Thus, condition (i) holds.

Condition (ii) holds for any  $\beta > 0$  since  $\mathbf{z}(t)$  is a Wiener process with variance  $2\beta^{-1}$ .

For (iii), we find  $\text{ess sup}_{(\mathbf{x}, t) \in \mathbb{R}^2 \times [0, \infty)} \|\mathbf{f}^u(\mathbf{x}, t, u(\mathbf{x}, t), \rho)\|_{\infty}$

$$\begin{aligned} & = \text{ess sup}_{(\mathbf{x}, t) \in \mathbb{R}^2 \times [0, \infty)} \left\| \int_{\mathbb{R}^2} \nabla_{\mathbf{x}} \phi^u(\mathbf{x}, \mathbf{y}, t) \rho(\mathbf{y}) d\mathbf{y} \right\|_{\infty} \\ & \leq \text{ess sup}_{(\mathbf{x}, t) \in \mathbb{R}^2 \times [0, \infty)} \int_{\mathbb{R}^2} \|\nabla_{\mathbf{x}} \phi^u(\mathbf{x}, \mathbf{y}, t) \rho(\mathbf{y})\|_{\infty} d\mathbf{y} \\ & \leq \int_{\mathbb{R}^2} \text{ess sup}_{(\mathbf{x}, t) \in \mathbb{R}^2 \times [0, \infty)} \|\nabla_{\mathbf{x}} \phi^u(\mathbf{x}, \mathbf{y}, t) \rho(\mathbf{y})\|_{\infty} d\mathbf{y} \\ & = \int_{\mathbb{R}^2} \text{ess sup}_{(\mathbf{x}, t) \in \mathbb{R}^2 \times [0, \infty)} \|\nabla_{\mathbf{x}} \phi^u(\mathbf{x}, \mathbf{y}, t)\|_{\infty} \rho(\mathbf{y}) d\mathbf{y} \quad (8) \end{aligned}$$

where we used the Leibniz rule, triangle inequality, and that  $\rho \geq 0$ . Per assumption,  $\nabla_{\mathbf{x}} \phi^u$  is uniformly bounded, and we have: (8)  $\leq M \int_{\mathbb{R}^2} \rho(\mathbf{y}) d\mathbf{y} = M$  for some constant  $M > 0$ .

Condition (iv) holds because

$$\begin{aligned} & \|\mathbf{f}^u(\mathbf{x}, t, u(\mathbf{x}, t), \rho) - \mathbf{f}^u(\mathbf{x}, t, u(\mathbf{x}, t), \tilde{\rho})\|_2 \\ & = \|\nabla_{\mathbf{x}} \int_{\mathbb{R}^2} \phi^u(\mathbf{x}, \mathbf{y}, t) (\rho(\mathbf{y}) - \tilde{\rho}(\mathbf{y})) d\mathbf{y}\|_2 \\ & = \left\| \int_{\mathbb{R}^2} (\nabla_{\mathbf{x}} \phi^u(\mathbf{x}, \mathbf{y}, t)) (\rho(\mathbf{y}) - \tilde{\rho}(\mathbf{y})) d\mathbf{y} \right\|_2 \\ & \leq c \text{dist}_{\text{BL}}(\rho, \tilde{\rho}) \leq \kappa \text{dist}_{\text{TV}}(\rho, \tilde{\rho}) \quad \forall (\mathbf{x}, t) \in \mathbb{R}^2 \times [0, \infty) \end{aligned}$$

for some constant  $c > 0$ ,  $\kappa := 2c$ , and the second to last inequality follows from  $\nabla_{\mathbf{x}} \phi^u$  being bounded and Lipschitz.

Thus, we can guarantee the existence-uniqueness of sample path  $\mathbf{x}_i(t)$ ,  $i \in \llbracket n \rrbracket$ , solving the interacting diffusion (7).

## B. Derivation of the Controlled Mean Field Model

Our next result (Theorem 1) derives the macroscopic mean field dynamics as a *nonlinear* Fokker-Planck-Kolmogorov partial differential equation (PDE), and establishes the consistency of the mean field dynamics in the continuum limit vis-à-vis the finite population dynamics.

**Theorem 1.** *Supposing A1, consider a population of  $n$  interacting chiplets, where the  $i$ th chiplet position  $\mathbf{x}_i \in \mathbb{R}^2$ ,  $i \in \llbracket n \rrbracket$ , evolves via (7). Denote the Dirac measure concentrated at  $\mathbf{x}_i$  as  $\delta_{\mathbf{x}_i}$  and let the random empirical measure  $\rho^n := \frac{1}{n} \sum_{i=1}^n \delta_{\mathbf{x}_i}$ . Consider the empirical version of the dynamics (7) given by*

$$d\mathbf{x}_i = \mathbf{f}^u(\mathbf{x}_i, t, u, \rho^n) dt + \sqrt{2\beta^{-1}} d\mathbf{w}_i(t),$$

with respective initial conditions  $\mathbf{x}_{0i} \in \mathbb{R}^2$ ,  $i \in \llbracket n \rrbracket$ , which are independently sampled from a given PDF  $\rho_0$  supported on a subset of  $\mathbb{R}^2$ . Then, as  $n \rightarrow \infty$ , almost surely  $\rho^n \rightarrow \rho$  where the deterministic function  $\rho$  is a PDF that evolves as per the macroscopic dynamics

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= -\nabla \cdot (\rho \mathbf{f}^u) + \beta^{-1} \Delta \rho \\ &= \nabla \cdot (\rho \nabla (\rho * \phi^u + \beta^{-1}(1 + \log \rho))), \end{aligned} \quad (9)$$

with the initial condition

$$\rho(\cdot, t=0) = \rho_0 \in \mathcal{P}(\mathbb{R}^2) \text{ (given)}. \quad (10)$$

*Proof.* To describe the dynamics of  $\rho^n$  as  $n \rightarrow \infty$ , we start with investigating the time evolution of the quantity

$$\langle \varphi, \rho^n \rangle := \frac{1}{n} \sum_{i=1}^n \varphi(\mathbf{x}_i) \quad (11)$$

for any compactly supported test function  $\varphi \in C_b^2(\mathbb{R}^2)$ .

Using Ito's rule, we have

$$d\varphi(\mathbf{x}_i) = L_{\rho^n} \varphi(\mathbf{x}_i) dt + \nabla \varphi^\top(\mathbf{x}_i) \sqrt{2\beta^{-1}} d\mathbf{w}_i \quad (12)$$

wherein the infinitesimal generator

$$L_{\rho^n} \varphi(\mathbf{x}) := \langle \mathbf{f}^u(\mathbf{x}, t, u, \rho^n), \nabla_{\mathbf{x}} \varphi(\mathbf{x}) \rangle + \beta^{-1} \Delta \varphi. \quad (13)$$

Thus,

$$\begin{aligned} d\langle \varphi, \rho^n \rangle &= \frac{1}{n} \sum_{i=1}^n d\varphi(\mathbf{x}_i) \\ &= \langle L_{\rho^n} \varphi, \rho^n \rangle dt + \frac{1}{n} \sum_{i=1}^n \sqrt{2\beta^{-1}} \nabla \varphi^\top(\mathbf{x}_i) d\mathbf{w}_i \\ &:= \langle L_{\rho^n} \varphi, \rho^n \rangle dt + dM_t^n \end{aligned} \quad (14)$$

where  $M_t^n$  is a local martingale.

Because  $\varphi \in C_b^2(\mathbb{R}^2)$ , we have  $|\sqrt{2\beta^{-1}} \nabla \varphi^\top(\mathbf{x}_i)| \leq C$  uniformly for some  $C > 0$ . Notice that the quadratic variation of the noise term in (14) is

$$[M_t^n] := \frac{1}{n^2} \sum_{i=1}^n \int_0^t \left| \sqrt{2\beta^{-1}} \nabla \varphi^\top(\mathbf{x}_i(s)) \right|^2 ds \leq \frac{tC^2}{n},$$

and using Doob's martingale inequality [22, Ch. 14.11],

$$\mathbb{E} \left( \sup_{t \leq T} M_t^n \right)^2 \leq \mathbb{E} \left( \sup_{t \leq T} (M_t^n)^2 \right) \leq 4\mathbb{E} \left( (M_t^n)^2 \right)$$

$$\leq 4\mathbb{E}([M_t^n]) \leq \frac{4tC^2}{n}.$$

Hence in the limit  $n \rightarrow \infty$ , the noise term in (14) vanishes, resulting in a deterministic evolution equation.

For any  $t > 0$ , we take  $\{\rho^n\}_{n=1}^\infty$  to be the (random) elements of  $\Omega = C([0, \infty), \mathcal{P}(\mathbb{R}^2))$ , the set of continuous functions from  $[0, \infty)$  into  $\mathcal{P}(\mathbb{R}^2)$  endowed with the topology of weak convergence. Following the argument of Oelschläger [23, Proposition 3.1], the sequence  $\mathbb{P}_n$  of joint PDFs on  $\Omega$  induced by the processes  $\{\rho^n\}_{n=1}^\infty$ , is relatively compact in  $\mathcal{P}(\Omega)$ , which is the space of probability measures on  $\Omega$ . Oelschläger's proof makes use of the Prohorov's theorem [24, Ch. 5]. The relative compactness implies that the sequence  $\mathbb{P}_n$  weakly converges (along a subsequence) to some  $\mathbb{P}$ , where  $\mathbb{P}$  is the joint PDFs induced by the limiting process  $\rho$ . By Skorohod representation theorem [24, Theorem 6.7], the sequence  $\{\rho^n\}_{n=1}^\infty$  converges  $\mathbb{P}$ -almost surely to  $\rho$ . Since the martingale term in (14) vanishes as  $n \rightarrow \infty$ , we obtain

$$d\langle \varphi, \rho \rangle = \langle L_\rho \varphi, \rho \rangle dt = \langle \varphi, L_\rho^* \rho \rangle dt \quad (15)$$

where  $L^*$  is the adjoint (see e.g., [25, Ch. 2.3, 2.5], [26, p. 278]) of the generator  $L$  given by (13), and is defined as

$$\begin{aligned} L_m^* \rho(\mathbf{x}, t) &:= -\nabla \cdot (\rho \mathbf{f}^u(\mathbf{x}, t, u, m)) + \beta^{-1} \Delta \rho \\ &= \nabla \cdot (\rho \nabla (m * \phi^u + \beta^{-1}(1 + \log \rho))) \end{aligned}$$

where  $m \in \mathcal{P}(\mathbb{R}^2)$ . For any test function  $\varphi \in C_b^2(\mathbb{R}^2)$ , (15) is valid almost everywhere, and therefore,  $\rho$  is almost surely a weak solution to the nonlinear Fokker-Planck-Kolmogorov PDE initial value problem (9)-(10). ■

Notice that the Cauchy problem (9)-(10) involves a *nonlinear nonlocal PDE* which in turn depends on control policy  $u$ .

The solution  $\rho(\mathbf{x}, t)$ ,  $\mathbf{x} \in \mathbb{R}^2$ ,  $t \in [0, \infty)$ , for the Cauchy problem (9)-(10) is understood in weak sense. In other words, for all compactly supported smooth test functions  $\theta \in C_c^\infty(\mathbb{R}^2, [0, \infty))$ , the solution  $\rho(\mathbf{x}, t)$  satisfies

$$\int_0^\infty \int_{\mathbb{R}^2} \left( \frac{\partial \theta}{\partial t} + L_\rho \theta \right) \rho d\mathbf{x} dt + \int_{\mathbb{R}^2} \rho_0(\mathbf{x}) \theta(\mathbf{x}, 0) d\mathbf{x} = 0 \quad (16)$$

where  $L_\rho$  is defined as in (13). The reason why  $\rho$  satisfying (16) for all  $\theta \in C_c^\infty(\mathbb{R}^2, [0, \infty))$  is called a "weak solution" of (9)-(10) is because such  $\rho$  may not be sufficiently smooth to satisfy (9). In the next Section, we provide a variational interpretation of the solution for problem (9)-(10).

## IV. CHIPLET POPULATION DYNAMICS AS WASSERSTEIN GRADIENT FLOW

The structure of the PDE in (9) motivates defining an *energy functional*

$$\begin{aligned} \Phi(\rho) &:= \Phi_{cc}(\rho) + \Phi_{ce}(\rho) + \mathbb{E}_\rho [\beta^{-1} \log \rho] \\ &= \mathbb{E}_\rho [\rho * \phi^u + \beta^{-1} \log \rho] \end{aligned} \quad (17)$$

where  $\mathbb{E}_\rho$  denotes the expectation w.r.t. the PDF  $\rho$ , and

$$\Phi_{cc}(\rho) := \int_{\mathbb{R}^2 \times \mathbb{R}^2} \phi_{cc}^u(\mathbf{x}, \mathbf{y}) \rho(\mathbf{x}) \rho(\mathbf{y}) d\mathbf{x} d\mathbf{y}, \quad (18a)$$

$$\Phi_{ce}(\rho) := \int_{\mathbb{R}^2 \times \mathbb{R}^2} \phi_{ce}^u(\mathbf{x}, \mathbf{y}) \rho(\mathbf{x}) \rho(\mathbf{y}) d\mathbf{x} d\mathbf{y}. \quad (18b)$$

In (17), the term  $\mathbb{E}_\rho[\rho * \phi^u]$  quantifies the *interaction energy* while the term  $\beta^{-1}\mathbb{E}_\rho[\log \rho]$  (scaled negative entropy) quantifies the *internal energy*. We have the following result.

**Theorem 2.** Let  $\Phi : \mathcal{P}_2(\mathbb{R}^2) \mapsto \mathbb{R}$  be the energy functional given in (17). Then,

(i) the chiplet population dynamics given by (4), (5), (9) is Wasserstein gradient flow of the functional  $\Phi$ , i.e.,

$$\frac{\partial \rho}{\partial t} = -\nabla^W \Phi(\rho). \quad (19)$$

(ii)  $\Phi$  is a Lyapunov functional that is decreasing along the flow generated by (9), i.e.,  $\frac{d}{dt}\Phi \leq 0$ .

*Proof.* (i) We start by noticing that the functional derivative

$$\frac{\delta \Phi}{\delta \rho} = \rho * \phi^u + \beta^{-1}(1 + \log \rho). \quad (20)$$

Next, we rewrite (9) as

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left( \rho \nabla \frac{\delta \Phi}{\delta \rho} \right), \quad (21)$$

which by definition (3), yields (19).

(ii) To show that  $\Phi$  is decreasing along the flow generated by (9), we find

$$\begin{aligned} \frac{d}{dt}\Phi &= \int \frac{\delta \Phi}{\delta \rho} \frac{\partial \rho}{\partial t} dx \\ &\stackrel{(21)}{=} \int \frac{\delta \Phi}{\delta \rho} \nabla \cdot \left( \rho \nabla \frac{\delta \Phi}{\delta \rho} \right) dx \\ &= - \int \left\langle \nabla \frac{\delta \Phi}{\delta \rho}, \rho \nabla \frac{\delta \Phi}{\delta \rho} \right\rangle dx \\ &= - \int \left\langle \nabla \frac{\delta \Phi}{\delta \rho}, \nabla \frac{\delta \Phi}{\delta \rho} \right\rangle \rho dx \\ &= -\mathbb{E}_\rho \left[ \left\| \nabla \frac{\delta \Phi}{\delta \rho} \right\|_2^2 \right] \leq 0. \end{aligned} \quad (22)$$

In order to get from the second line to the third line of (22), we used the duality<sup>1</sup> between the gradient and divergence operators, namely the fact that for differentiable scalar field  $s(\mathbf{x})$  and vector field  $\mathbf{v}(\mathbf{x})$ , we have

$$\langle \nabla s, \mathbf{v} \rangle_{L_2} + \langle s, \nabla \cdot \mathbf{v} \rangle_{L_2} = 0, \quad (23)$$

where  $\langle \mathbf{p}, \mathbf{q} \rangle_{L_2} := \int \langle \mathbf{p}, \mathbf{q} \rangle dx$ . Specifically, in (22),  $s \equiv \frac{\delta \Phi}{\delta \rho}$  and  $\mathbf{v} \equiv \rho \nabla \frac{\delta \Phi}{\delta \rho}$ . ■

**Remark 2.** Theorem 2 shows that for an admissible control policy  $u \in \mathcal{U}$ , the chiplet population dynamics (9)-(10) can be seen as gradient descent of the energy functional  $\Phi$  on the manifold  $\mathcal{P}_2(\mathbb{R}^2)$  w.r.t. the Wasserstein metric. We point out that the statement of Theorem 2 remains valid in the deterministic limit, i.e., when the noise strength  $\sqrt{2\beta^{-1}} \downarrow 0$ . In that case, the functional  $\Phi$  in (17) comprises of only the interaction energy term  $\mathbb{E}_\rho[\rho * \phi^u]$ , and  $\frac{\delta \Phi}{\delta \rho} = \rho * \phi^u$ . Other than this, the proof of Theorem 2 remains unchanged.

**Remark 3.** In the recent systems-control literature, the Wasserstein gradient flow interpretations and related proximal algorithms [27], [28] for several linear and nonlinear

Fokker-Planck-Kolmogorov PDEs in prediction and density control have appeared. New gradient flow interpretations have also appeared [29]–[31] for well-known filtering equations. We next point out that the Wasserstein gradient flow interpretation deduced in Theorem 2 allows approximating the weak solution of (19) by recursive evaluation of a Wasserstein proximal operator on the manifold  $\mathcal{P}_2(\mathbb{R}^2)$ .

**Theorem 3.** For a given control policy  $u \in \mathcal{U}$  and potentials (5), let  $\widehat{\Phi}(\varrho, \varrho_{k-1}) := \mathbb{E}_\varrho[\varrho_{k-1} * \phi^u + \beta^{-1} \log \varrho]$ ,  $\varrho, \varrho_{k-1} \in \mathcal{P}_2(\mathbb{R}^2)$ ,  $k \in \mathbb{N}$ . Consider the Wasserstein proximal recursion:

$$\begin{aligned} \varrho_k &= \text{prox}_{\tau \widehat{\Phi}}^W(\varrho_{k-1}) \\ &:= \arg \inf_{\varrho \in \mathcal{P}_2(\mathbb{R}^2)} \left\{ \frac{1}{2} W^2(\varrho, \varrho_{k-1}) + \tau \widehat{\Phi}(\varrho, \varrho_{k-1}) \right\} \end{aligned} \quad (24)$$

over discrete time  $t_{k-1} := (k-1)\tau$  with fixed step-size  $\tau > 0$ , and with initial condition  $\varrho_0 \equiv \rho_0 \in \mathcal{P}_2(\mathbb{R}^2)$ . Let  $\rho(\mathbf{x}, t)$  be the weak solution of (19) for the same  $u \in \mathcal{U}$  and the functional  $\Phi$  given by (17)-(18). Using the sequence of functions  $\{\varrho_{k-1}\}_{k \in \mathbb{N}}$  generated by the recursion (24), define an interpolation  $\varrho_\tau : \mathbb{R}^2 \times [0, \infty) \mapsto [0, \infty)$  as

$$\varrho_\tau(\mathbf{x}, t) := \varrho_{k-1}(\mathbf{x}, \tau) \quad \forall t \in [(k-1)\tau, k\tau), \quad k \in \mathbb{N}.$$

Then  $\varrho_\tau(\mathbf{x}, t) \xrightarrow{\tau \downarrow 0} \rho(\mathbf{x}, t)$  in  $L^1(\mathbb{R}^2)$  for all  $t \in [0, \infty)$ .

*Proof.* Follows the development in [32, Sec. 12.3–12.5]. ■

**Remark 4.** For a given control policy  $u \in \mathcal{U}$ , the Wasserstein proximal recursion (24) can in turn be leveraged for numerically updating the PDFs over discrete time with a small step-size  $\tau$ . To illustrate Theorem (2), we fixed a linear control policy  $u = \langle \mathbf{k}, \mathbf{x} \rangle$  with gain  $\mathbf{k} := (8.5 \times 10^{-3}, -1 \times 10^{-2})^\top$ , and solved (24) with  $\tau = 0.1$  via [27, Algorithm 1] for  $n = 400$  uniformly spaced grid samples in the domain  $[-4 \text{ mm}, 4 \text{ mm}]^2$  starting from an initial bivariate Gaussian  $\rho_0 = \mathcal{N}((0.5, 0.5)^\top, 0.1 \mathbf{I}_2)$ . Fig. 1 shows the corresponding decay of the energy functional  $\Phi$  in (17)-(18), computed using these PDFs obtained from the Wasserstein proximal updates. As in [6, Sec. III], our simulation used capacitances  $C_{cc}(\|\mathbf{x} - \mathbf{y}\|_2)$ ,  $C_{ce}(\|\mathbf{x} - \mathbf{y}\|_2)$  in (5b)-(5c) of the form  $\sum_{i=1}^n a_i [\text{erf}((\|\mathbf{x} - \mathbf{y}\|_2 + \delta)/c_i) - \text{erf}((\|\mathbf{x} - \mathbf{y}\|_2 - \delta)/c_i)]$  where  $\text{erf}(\cdot)$  denotes the error function, the parameters  $a_i, c_i$  are sampled uniformly random in  $[0, 1]$ , and  $\delta$  (half of the electrode pitch) = 10 micrometer.

## V. CONCLUSIONS

We presented a controlled mean field model for the population dynamics of chiplets, which are tiny (micron sized or smaller) particles immersed in a dielectric liquid, and are amenable to reshape into desired concentrations for micro-assembly applications. In such applications, an array of electrodes generate a space-time varying electric potential landscape, thereby strategically inducing the collective motion of the chiplet ensemble. Our derived model quantifies how exactly the two types of nonlocal nonlinear

<sup>1</sup>In words, the gradient and the negative divergence are adjoint maps.

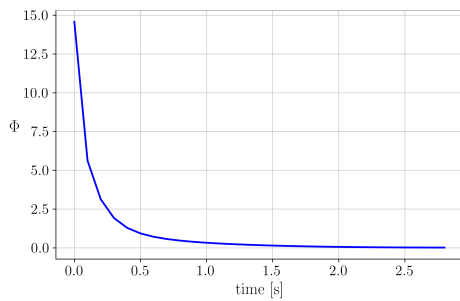


Fig. 1: The energy functional  $\Phi$  given by (17)–(18) versus time for the simulation set up summarized in Remark 4.

interactions (viz. chiplet-to-chiplet and chiplet-to-electrode) jointly induce a macroscopic dynamics in terms of the joint PDF evolution of the chiplet ensemble. Our results establish consistency of the model in a limiting sense, and demonstrate that the resulting PDF evolution can be seen as an infinite dimensional gradient descent of a Lyapunov-like energy functional w.r.t. the Wasserstein metric.

While we focused our development for the derivation of the controlled mean field model, our future work will investigate the synthesis of optimal control of the chiplet joint PDF w.r.t. suitable performance objective that allows steering an initial joint PDF to a desired terminal joint PDF. Such feedback steering problems are generalized variants of the so-called Schrödinger bridge problem [33]. We note that the feedback synthesis for density steering subject to a controlled mean field nonlocal PDE is relatively less explored but has started appearing in recent works; see e.g., [34]–[36].

## REFERENCES

- [1] I. Matei, S. Nelaturi, J. P. Lu, J. A. Bert, L. S. Crawford, and E. Chow, “Towards printing as an electronics manufacturing method: Micro-scale chiplet position control,” in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 1549–1555.
- [2] I. Matei, S. Nelaturi, E. M. Chow, J. P. Lu, J. A. Bert, and L. S. Crawford, “Micro-scale chiplets position control,” *Journal of Microelectromechanical Systems*, vol. 28, no. 4, pp. 643–655, 2019.
- [3] J. Lu, J. Thompson, G. Whiting, D. Biegelsen, S. Raychaudhuri, R. Lujan, J. Veres, L. Lavery, A. Völkel, and E. Chow, “Open and closed loop manipulation of charged microchips in an electric field,” *Applied Physics Letters*, vol. 105, no. 5, p. 054104, 2014.
- [4] T. D. Edwards and M. A. Bevan, “Controlling colloidal particles with electric fields,” *Langmuir*, vol. 30, no. 36, pp. 10793–10803, 2014.
- [5] I. Matei, J. de Kleer, C. Somarakis, A. Plochowietz, and J. S. Baras, “Micro-scale 2d chiplet position control: a formal approach to policy design,” in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 5519–5524.
- [6] I. Matei, A. Plochowietz, J. de Kleer, and J. S. Baras, “Micro-scale chiplet assembly control with chiplet-to-chiplet potential interaction,” in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 623–628.
- [7] A. Lefevre, V. Gauthier, M. Gauthier, and A. Bolopion, “Closed-loop control of particles based on dielectrophoretic actuation,” *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 6, pp. 4764–4773, 2022.
- [8] I. Matei, J. de Kleer, and M. Zhenirovskyy, “2D density control of micro-particles using kernel density estimation,” *arXiv preprint arXiv:2209.03550*, 2022.
- [9] I. Matei, A. Plochowietz, S. Nelaturi, J. De Kleer, J. P. Lu, L. S. Crawford, and E. M. Chow, “System and method for machine-learning enabled micro-assembly control with the aid of a digital computer,” Jun. 16 2022, US Patent App. 17/121,411.
- [10] L. V. Kantorovich, “On the translocation of masses,” in *Dokl. Akad. Nauk. USSR (NS)*, vol. 37, 1942, pp. 199–201.
- [11] C. Villani, *Topics in Optimal Transportation*. American Mathematical Soc., 2003, no. 58.
- [12] C. Villani, *Optimal transport: old and new*. Springer, 2009, vol. 338.
- [13] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.
- [14] R. M. Dudley, *Real analysis and probability*. volume 74 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2002.
- [15] J. P. Lu, J. D. Thompson, G. L. Whiting, D. K. Biegelsen, S. Raychaudhuri, R. Lujan, J. Veres, L. L. Lavery, A. R. Völkel, and E. M. Chow, “Open and closed loop manipulation of charged microchips in an electric field,” *Applied Physics Letters*, vol. 105, no. 5, pp. 054104–1–054104–4, 2014. [Online]. Available: <https://scitation.aip.org/content/aip/journal/apl/105/5/10.1063/1.4891957>
- [16] COMSOL Multiphysics® v. 6.1., COMSOL AB, Stockholm, Sweden. [Online]. Available: [www.comsol.com](http://www.comsol.com)
- [17] H. P. McKean Jr, “A class of Markov processes associated with nonlinear parabolic equations,” *Proceedings of the National Academy of Sciences*, vol. 56, no. 6, pp. 1907–1911, 1966.
- [18] A.-S. Sznitman, “Topics in propagation of chaos,” in *Ecole d’été de probabilités de Saint-Flour XIX—1989*. Springer, 1991, pp. 165–251.
- [19] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications I-II*. Springer, 2018.
- [20] M. Kac, “Foundations of kinetic theory,” in *Proceedings of The third Berkeley symposium on mathematical statistics and probability*, vol. 3, 1956, pp. 171–197.
- [21] D. Lacker, “On a strong form of propagation of chaos for McKean-Vlasov equations,” *Electronic Communications in Probability*, vol. 23, no. 45, pp. 1–11, 2018.
- [22] D. Williams, *Probability with martingales*. Cambridge university press, 1991.
- [23] K. Oelschläger, “A law of large numbers for moderately interacting diffusion processes,” *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 69, no. 2, pp. 279–322, 1985.
- [24] P. Billingsley, *Convergence of probability measures*. John Wiley & Sons, 2013.
- [25] G. A. Pavliotis, *Stochastic processes and applications: diffusion processes, the Fokker-Planck and Langevin equations*. Springer, 2014, vol. 60.
- [26] V. I. Bogachev, N. V. Krylov, M. Röckner, and S. V. Shaposhnikov, *Fokker-Planck-Kolmogorov Equations*. American Mathematical Soc., 2015, vol. 207.
- [27] K. F. Caluya and A. Halder, “Gradient flow algorithms for density propagation in stochastic systems,” *IEEE Transactions on Automatic Control*, vol. 65, no. 10, pp. 3991–4004, 2019.
- [28] K. F. Caluya and A. Halder, “Wasserstein proximal algorithms for the Schrödinger bridge problem: Density control with nonlinear drift,” *IEEE Transactions on Automatic Control*, vol. 67, no. 3, pp. 1163–1178, 2021.
- [29] A. Halder and T. T. Georgiou, “Gradient flows in uncertainty propagation and filtering of linear Gaussian systems,” in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 3081–3088.
- [30] A. Halder and T. T. Georgiou, “Gradient flows in filtering and Fisher-Rao geometry,” in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 4281–4286.
- [31] A. Halder and T. T. Georgiou, “Proximal recursion for the Wonham filter,” in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 660–665.
- [32] M. Laborde, “On some nonlinear evolution systems which are perturbations of Wasserstein gradient flows,” *Topological Optimization and Optimal Transport: In the Applied Sciences*, vol. 17, pp. 304–332, 2017.
- [33] Y. Chen, T. T. Georgiou, and M. Pavon, “On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint,” *Journal of Optimization Theory and Applications*, vol. 169, pp. 671–691, 2016.
- [34] Y. Chen, “Density control of interacting agent systems,” *arXiv preprint arXiv:2108.07342*, 2021.
- [35] C. Sinigaglia, F. Braghin, and S. Berman, “Optimal control of velocity and nonlocal interactions in the mean-field Kuramoto model,” in *2022 American Control Conference (ACC)*. IEEE, 2022, pp. 290–295.
- [36] T. Zheng, Q. Han, and H. Lin, “Backstepping mean-field density control for large-scale heterogeneous nonlinear stochastic systems,” in *2022 American Control Conference (ACC)*. IEEE, 2022, pp. 4832–4837.