Optimal Ensemble Control of Neural Populations: Numerical Experiments*

Roman Chertovskih[1], Nikolay Pogodaev[2], Maxim Staritsyn[1] and Antonio Pedro Aguiar[1]

***Abstract*— We investigate the challenge of designing robust external excitations to control and synchronize a population of non-interacting homotypic harmonic oscillators, specifically, theta neurons. The Theta model emulates the bursting behavior observed in spiking cells, characterized by periodic oscillations in their membrane electric potential.**

Our approach involves formulating this optimization task as an optimal mean-field control problem for the linear continuity/Fokker-Planck equation in the space of probability measures. To address this problem numerically, we employ an indirect deterministic descent method, leveraging an exact representation of the increment of the objective functional.

As a main contribution, we delve into practical aspects in the implementation of the proposed method and expose several results of numerical experiments.

I. INTRODUCTION

The synchronization of oscillatory processes is a widespread phenomenon observed in various physical and natural systems composed of structurally similar interacting entities. This behavior, often emerging through different forms of time-periodic patterns, manifests in a wide range of applications; see, e.g., [1]–[6] and the references therein.

The mathematical foundation for understanding oscillatory ensembles can be traced back to the pioneering work of Y. Kuramoto and H. Araki [7], which proposed approximating the behavior of a population by its "averaged representative." This approach led to the important concept of the *mean-field limit*, providing a manageable and elegant mathematical framework for describing ensemble dynamics.

A central challenge in this context is the design of artificial signals to drive open systems towards or away from synchronized oscillations and frequency entrainment. This problem naturally aligns with the principles of ensemble control in control engineering, where the focus is on designing "simultaneous" and robust control signals for multi-agent systems [8], [9].

*The authors acknowledge the financial support of the Foundation for Science and Technology (FCT, Portugal) in the framework of ARISE (DOI 10.54499/LA/P/0112/2020) and R&D Unit SYSTEC (base UIDB/00147/2020 and programmatic UIDP/00147/2020 funds) supported by national funds through FCT/MCTES (PIDDAC). A part of the computations was carried out on the OBLIVION Supercomputer (Évora University) under FCT/RNCA computational projects 2023.10674.CPCA and 2024.07885.CPCA.A2.

[1]Roman Chertovskih, Maxim Staritsyn, and A. Pedro Aguiar are with Universidade do Porto, Faculdade de Engenharia, Rua Dr. Roberto Frias, s/n 4200-465, Porto, Portugal roman@fe.up.pt, staritsyn@fe.up.pt, pedro.aguiar@fe.up.pt

[2]Nikolay Pogodaev is with University of Padova, Dipartimento di Matematica "Tullio Levi-Civita" (DM), Via Trieste, 63 - 35121 Padova, Italy nickpogo@gmail.com

A. Background

This paper presents an ongoing study building on our previous works [10], [11], and promotes a novel approach for solving optimal ensemble control problems numerically.

Our method is based on two key ideas:

1) A mean-field approximation of the population dynamics that transforms the original nonlinear dynamic optimization problem in a finite-dimensional space into a semi-linear problem in the space of probability measures, and
2) An exact (explicit and nonlocal) representation of the increment of the objective functional, which leads to a fully deterministic numerical algorithm characterized by a remarkably fast cost-monotonic descent.

As in [10], we apply our optimization method to a specific in-phase synchronization problem involving an "infinite population" of excitable neurons undergoing periodic oscillations in their membrane potential — a simple yet significant example in mathematical neuroscience.

B. Contribution. Organization of the Paper.

The contribution of this study is twofold. First, we introduce our conceptual approach as a unified framework for both deterministic and stochastic problems, shedding new light on the role of "parabolic regularization" within the context of ensemble control. Second, we explore its practical implementation and experimental validation, addressing important issues that have previously been overlooked.

The rest of the paper is organized as follows: In Section II, we formulate the optimization problem, followed by an outline of the mathematical framework in Section III, which leads to an effective numerical algorithm for optimal ensemble control, applicable to both deterministic and stochastic settings.

Section IV focuses on the key computational steps involved in the practical implementation of our method, including the numerical integration of the resulting partial differential equations using a pseudo-spectral method. This analysis is supported by new results from numerical experiments.

The paper concludes with a brief summary of the key findings in Section V.

II. PROBLEM STATEMENT

We begin by formulating a specific optimal control problem in the space of probability measures, where the dynamics describe a population of homotypic oscillatory units known as "theta neurons" [12], [13].

A. Individual Dynamics

The idealized neural cell is characterized by its membrane potential,

$$\theta(t) \in \mathbb{S}^1 \doteq \mathbb{R}/2\pi\mathbb{Z},$$

which evolves over time t and depends on an intrinsic parameter η representing the membrane baseline current.

Originally, the time evolution of the *phase* $\theta(t)$ on a given time interval $[0, T]$ is governed by the following (essentially nonlinear) ordinary differential equation (ODE):

$$\dot{\theta} = (1 - \cos\theta) + (1 + \cos\theta)\,(u + \eta)\,, \tag{1}$$

where the function

$$u\colon t \mapsto u(t)$$

represents an external stimulus applied to the neuron, acting as a *control input*.

For convenience, we will treat the parameter η as a fictitious state variable $\eta(t) \equiv \eta$, satisfying the trivial ODE

$$\dot{\eta} = 0.$$

Defining

$$x(t) \doteq [\theta(t)\ \eta(t)]^{\mathrm{T}} \text{ and } \mathbb{X} \doteq \mathbb{S}^1 \times \mathbb{R},$$

we abbreviate the right-hand side of (1) as $v_u(x)$, and recognize that v_u acts as a vector field

$$\mathbb{X} \to \mathbb{R}^2 \simeq \mathcal{T}_x(\mathbb{X})$$

to the tangent space of $\mathbb{X}$ at any point x.

B. Uncertainty and Noise

It is often practical to assume that the initial data for the model (1) are uncertain, with the parameter η subject to estimation errors. To address this, we let

$$x(0) \doteq (\theta(0), \eta)$$

be a random variable defined on a complete probability space (p.s.) $(\Omega, \mathcal{F}, \mathbb{P})$, such that η almost surely (a.s.) belongs to a given compact interval $\mathbb{I} \doteq [a, b]$. This assumption transforms a trajectory $x = x[u]$ of (1) into a random process

$$[0, T] \times \Omega \to \mathbb{X}.$$

Another natural way to account for the "imperfection" of the model is to assume that the dynamics in (1) are affected by noise, representing intrinsic system variability, environmental fluctuations, or experimental conditions. The simplest implementation of this idea is to transition to Îto's stochastic differential equation (SDE):

$$\mathrm{d}x = \begin{bmatrix} v_u(x) \\ 0 \end{bmatrix} \mathrm{d}t + \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \mathrm{d}W_t, \tag{2}$$

where the vector $c \doteq [c_1\ c_2]^{\mathrm{T}}$ of diffusion rates $c_{1,2} \geq 0$ is a parameter, and W is the standard Wiener process defined on a sufficiently rich p.s. (which we continue to denote in the same way) with the natural filtration $t \mapsto \mathcal{F}_t$ of the sigma-algebra $\mathcal{F}$ generated by the pair $(W, x(0))$. We still assume that $x(0)$ is a random vector on the underlying p.s., W is independent of $x(0)$, and the second component of $x(0)$ belong a.s. to $\mathbb{I}$.

Standard results from SDE theory [14] imply that (2), accompanied by the initial condition $x(0)$, admits a unique strong solution for any measurable input u. In this context, we denote this solution by the same letter $x = x[u]$.

The sample paths of the random/stochastic process x can be interpreted as trajectories of distinct homotypic neurons, characterized by varying initial conditions and/or dynamics, and evolving independently of one another.

Our next goal is to describe the behavior of the entire population x in terms of its distribution law.

C. Population (Mean-Field) Dynamics

In what follows, we shall consider two cases separately: $c = 0$ (Case I) and $c > 0$ (Case II).

Denote by

$$\vartheta \doteq x(0)_\sharp \mathbb{P} \doteq \mathbb{P} \circ x(0)^{-1}$$

the distribution law of $x(0)$, where $F_\sharp$ stands for the push-forward $\mathcal{P}(\Omega) \to \mathcal{P}(F(\Omega))$ of a measure by the map F. Similarly, we define

$$\mu_t \doteq x(t)_\sharp \mathbb{P},$$

where x is either a solution to the ODE (1) or the SDE (2) with a random initial condition $x(0)$, and consider the curve $\mu\colon t \mapsto \mu_t$ in the space $\mathcal{P}_c(\mathbb{X})$ of compactly supported probability measures on $\mathbb{X}$.

It is well-known [15], [16] that, in both cases, μ is the unique *distributional solution* to the initial-value problem for the linear partial differential equation (PDE) (with $c = 0$ formally corresponding to Case I):

$$\partial_t \mu_t + \partial_\theta\big(v_u\,\mu_t\big) = \big\{c_1 \partial^2_{\theta\theta} + c_2 \partial^2_{\eta\eta}\big\}\,\mu_t, \quad \mu_0 = \vartheta. \tag{3}$$

By a distributional solution to (3), we mean that the following relations hold for all $\varphi \in C^\infty(\mathbb{X})$:

$$\begin{cases} \dfrac{d}{dt}\displaystyle\int \varphi\,\mathrm{d}\mu_t = \int \big\{\nabla_x \varphi \cdot v_u + c^{\mathrm{T}} E \nabla^2_{xx} \varphi\big\}\,\mathrm{d}\mu_t, \\ \displaystyle\lim_{t\to 0} \int \varphi\,\mathrm{d}\mu_t = \int \varphi\,\mathrm{d}\vartheta, \end{cases}$$

where the integrals are over $\mathbb{X}$, the first relation holds a.e. on $[0, T]$, and E stands for the 2×2 unit matrix.

The PDE (3) can be viewed as a measure-theoretic extension of the linear *Fokker-Planck-Kolmogorov (FPK) equation*, which formally reduces to the *continuity (or Liouville) equation* when $c = 0$.

D. Optimal Control

Given $\vartheta \in \mathcal{P}(\mathbb{X})$, a continuous bounded function $\eta \mapsto \check{\theta}(\eta)$, and $\alpha > 0$, we formulate the following optimal ensemble control problem in the space of probability measures:

$$(P) \quad \begin{cases} \min I[u] = \displaystyle\int F\big(\theta, \check{\theta}(\eta)\big)\,\mathrm{d}\mu_T(\theta, \eta) + \frac{\alpha}{2}\|u\|^2_{L^2}, \\ \text{subject to (3)}, \quad u \in \mathcal{U}, \end{cases}$$

where the integrand

$$F(\theta,\omega) = 1-\cos(\theta-\omega)$$

measures the "distance" between the points θ and ω on the circle $\mathbb{S}^1$.

As a natural class $\mathcal{U}$ of admissible control inputs, we choose

$$\mathcal{U} \doteq L_2([0,T];\mathbb{R}).$$

The problem (P) seeks to best match the target profile $\check{\theta}$ at the prescribed moment T while minimizing the total energy spent by the controller. This process starts from the given initial distribution ϑ and relies on a "robust" stimulus that remains invariant under uncertainty and/or noise.

Note that the first term in I can equivalently be interpreted as the expectation $\mathbb{E}\left[F(\theta_T,\check{\theta}(\eta_T))\right]$ of the value of F at the random terminal state

$$x(T) \doteq [\theta(T)\,\eta(T)]^{\mathrm{T}}$$

w.r.t. the canonical probability $\mathbb{P}$.

The addressed problem is totally deterministic, with the state variable being the probability measure μ_t. This variable enters the dynamics and the cost functional linearly, despite the non-linearity of the characteristic ODE (1). However, due to the presence of a product of μ_t and $u(t)$ in (3), the problem remains *non-convex*. This feature adds complexity to the variational analysis of (P). For instance, here, the articulations of Pontryagin's maximum principle proposed by [17] and [18] fail to provide a sufficient condition for optimality.

III. COST-VARIATION FORMULA AND NUMERICAL METHOD

We proceed by recalling the line of arguments yielding the announced optimization algorithm. This reasoning, which applies equally to both cases I and II, follows [11], [19], [20] and relies on the linearity of the problem (P) w.r.t. the state variable μ, enabling the standard "duality argument".

The first step is to establish a particular *exact* representation for the increment

$$\Delta I \doteq I[u]-I[\bar{u}]$$

in the functional I w.r.t. an *arbitrary* pair of feasible control signals $\bar{u},u\in\mathcal{U}$. In subsequent optimization contexts, the control $\bar{u}$ will be interpreted as a given (reference) input, and u as an unknown (target) one. For brevity, we denote by $t\mapsto\bar{\mu}_t\doteq\mu_t[\bar{u}]$ and $t\mapsto\mu_t\doteq\mu_t[u]$ the corresponding distributional solutions to (3).

Introduce the function

$$\bar{p}\doteq p[\bar{u}]\colon\ (t,x\doteq(\theta,\eta))\mapsto\bar{p}_t(x),\quad [0,T]\times\mathbb{X}\to\mathbb{R},$$

by

$$\bar{p}_t(x)\doteq -F\big(\Phi_{t,T}(x),\check{\theta}(\eta)\big).$$

It can be shown (see, e.g. []) that $\bar{p}$ is a distributional solution of the backward (Kolmogorov/transport) equation

$$\begin{cases}\partial_t p_t+\partial_\theta p_t\cdot v_{\bar{u}(t)} = -\left\{c_1\partial^2_{\theta\theta}+c_2\partial^2_{\eta\eta}\right\}p_t,\\ p_T=-F\big(\theta,\check{\theta}(\eta)\big).\end{cases}\tag{4}$$

Furthermore, $\bar{p}$ is precisely the adjoint trajectory in the standard formulations [18], [21] of Pontryagin's maximum principle for the problem (P).

By definition of $\bar{p}$, the duality pairing

$$t\mapsto\int\bar{p}_t\,d\bar{\mu}_t$$

is a constant function on $[0,T]$. Using this fact, the difference ΔI is represented (refer to [19], [20] for further details) as

$$-\Delta I=\int_0^T\left(\bar{H}_t\left(\mu_t,u(t)\right)-\bar{H}_t\left(\mu_t,\bar{u}(t)\right)\right)dt,\tag{5}$$

where

$$\bar{H}_t(\mu,u)\doteq H\left(\mu,\partial_\theta\bar{p}_t,u\right),$$

and

$$H(\mu,\zeta,u)\doteq u\int\zeta(\theta,\eta)(1+\cos\theta)d\mu(\theta,\eta)-\frac{\alpha}{2}u^2$$

is the Hamilton-Pontryagin functional.

Expression (5) provides the desired exact cost-increment formula, which we now recruit to design an iterative descent method for problem (P).

Let $\bar{u}$ be given/computed. Denote $\bar{\xi}\doteq\partial_\theta\bar{p}$ and define $w=w_t[\mu]$ as the unique maximizer

$$w_t[\mu]=\frac{1}{\alpha}\int\bar{\xi}_t(\theta,\eta)\cdot(1+\cos\theta)d\mu(\theta,\eta),\tag{6}$$

in the mathematical programming problem:

$$H\left(\mu,\bar{\xi}_t,u\right)\to\max,\quad u\in\mathbb{R}.$$

Assume we find a solution $t\mapsto\check{\mu}_t$ of the PDE (3) with the nonlocal time-dependent vector field

$$(t,\mu)\mapsto v_{w_t[\mu]},$$

i.e.,

$$\partial_t\mu_t+\partial_\theta\big(v_{u[\mu_t]}\mu_t\big)=\left\{c_1\partial^2_{\theta\theta}+c_2\partial^2_{\eta\eta}\right\}\mu_t,\quad \mu_0=\vartheta,\tag{7}$$

and define:

$$u(t)=u_t[\check{\mu}_t],\tag{8}$$

it follows immediately from (5) that

$$\Delta I\le 0,$$

i.e., the control (8) guarantees a decrease in the cost I from the reference point $\bar{u}$.

By repeating this reasoning, we obtain an iterative Algorithm 1, which generates a cost-monotone sequence $\{u^k\}_{k\ge0}\subset\mathcal{U}$ of controls:

$$I^{k+1}\doteq I[u^{k+1}]<I[u^k]\doteq I^k.$$

Since the sequence $(I^k)_{k\ge0}$ is bounded below by $\min(P)$, it converges.

We stress that the proposed method does not involve any hidden free parameters to be optimized. In contrast, common indirect methods for optimal control, such as algorithms based on Pontryagin's maximum principle proposed in [22]–[24], typically incorporate intrinsic "modules" of parametric

Algorithm 1: Optimal ensemble control

Data: $\bar{u} \in \mathcal{U}$ (initial guess), $\varepsilon > 0$ (tolerance)
Result: $\{u^k\}_{k\geq 0} \subset \mathcal{U}$ such that $I[u^{k+1}] < I[u^k]$
$k \leftarrow 0$;
$u^0 \leftarrow \bar{u}$;
repeat
  $\mu^k \leftarrow \check{\mu}[u^k]$;
  $u^{k+1} \leftarrow w[\mu^k]$;
  $k \leftarrow k + 1$;
until $I[u^{k-1}] - I[u^k] < \varepsilon$;

line search as part of each iteration. Such parameters control, for example, the intensity of the employed (needle-shaped or weak) variations of the nominal control, thereby modulating the step size along the descent direction.

IV. NUMERICAL EXPERIMENTS

In this section, we discuss key features of Algorithm 1, highlight potential pitfalls in its practical implementation, and provide insights drawn from our experience in numerical analysis of problem (P) for both cases I and II.

A. Implementation

For computations, the initial distribution ϑ is typically assumed to be absolutely continuous w.r.t. the Lebesgue measure on $\mathbb{X}$, with a smooth density ρ_0. Consequently, all μ_t also possess densities, denoted by ρ_t.

Each algorithm iteration requires a numerical solution to two problems: the backward linear PDE (4) and the forward nonlocal equation (7). For simplicity, we always assume that $c_2 = 0$. In this case, since the drift in (3) does not involve the derivative ∂_η, it is possible to treat η as a parameter and solve η-parametric families of one-dimensional PDEs on $\mathbb{S}^1$ with $\eta \in \mathbb{I}$.

The PDEs (4) and (7) are solved by the classical spectral method [25], applicable to both continuity and FPK equations. Since the solution is 2π-periodic w.r.t. θ, it can be represented as a truncated Fourier series:

$$\rho_t = \sum_{n=-N/2}^{N/2} \widehat{\rho}_n(t,\eta)\mathrm{e}^{in\theta}. \tag{9}$$

Substituting (9) into (7) and equating coefficients with the corresponding basis functions transforms the PDEs into systems of ODEs:

$$\frac{\mathrm{d}\widehat{\rho}_n}{\mathrm{d}t} = -in\,\widehat{(v\rho)}_n - (c_1 n^2 + c_2\partial^2_{\eta\eta})\widehat{\rho}_n, \tag{10}$$

where $n = -N/2, \ldots, N/2$ and the wide hat means taking the n-th Fourier coefficient.

Typically, such equations require pseudospectral methods, where fast Fourier transforms (FFTs, see, e.g., [26]) are employed at each time step to compute the right-hand side. However, in our case, when all $\widehat{\rho}_k$ for given t and η are known, the right-hand side of (10) can be computed entirely in Fourier space without converting back to physical space. Therefore, the method used here is purely spectral rather than pseudospectral.

This computational simplification significantly reduces the algorithm's burden, as the control u does not explicitly depend on θ. Thus, operations such as multiplication by $\cos\theta$ and the integration in (6) — which represent convolutions — can be performed solely in the Fourier domain.

Since the solution is real-valued, only half of the Fourier coefficients in (9) need to be stored, i.e., $\widehat{\rho}_n$ for $n = 0, \ldots, N/2$. The remaining coefficients are determined by complex conjugation, $\widehat{\rho}_{-n} = \widehat{\rho}^*_n$.

B. Experimental Data

In numerical experiments, we took the following values of the computational parameters: $T = 4$, $\mathbb{I} = [0, 1]$ and $\alpha = 1$. The initial density, taken in the form

$$\rho_0(\theta, \eta) = \frac{1}{2} - \frac{1}{5}\sin(2\theta),$$

was represented by its Fourier coefficients in the expansion (9) with $N = 512$ Fourier harmonics.

The target function was specified in the form $\check{\theta}(\eta) = 2\pi\eta$; thus, our goal was to frame the neurons' distribution near the diagonal of the rectangle $(\theta, \eta) = [0, 2\pi] \times \mathbb{I}$ by the time moment T with the aid of fairly small resource of control.

The systems of ODEs in (10) were solved using the classical 4th-order Runge-Kutta method with a constant time step Δt.

Numeric integration was performed using a uniform grid for η with the step $\Delta\eta = 5 \cdot 10^{-3}$, and the time marching was made with a constant step $\Delta t = 10^{-3}$.

Also, other parameters of the algorithm were chosen to be $\bar{u} \equiv -1$, $\varepsilon = 10^{-2}$.

1) Case I (Continuity Equation): For the hyperbolic case ($c_1 = 0$), the algorithm converged after 10 iterations to the desired accuracy. The initial cost was $I[u^0] = I[\bar{u}] \approx 5.29$; at the final iteration $k = 10$ the cost decreased to $I[u^{10}] = 1.53$ (a reduction by a factor of approximately 3.5). The cost drops markedly during the first few iterations and then stagnates, a behavior typical of algorithms based on exact increment formulae, as observed in [11], [19], [20].

The numerical solution exhibits large gradients in certain regions of the computational domain (see Fig. 1, middle and bottom panels). Capturing such steep gradients necessitates high spatial resolution and, consequently, significant computational resources.

The computed control exhibits smooth oscillations bounded approximately within $[-0.61, 0.7]$, see Fig. 2.

2) Case II (FPK Equation): In the parabolic setup ($c_1 = 0.02, c_2 = 0$), similar computations were performed using the same initial density and baseline control signal $u^0 \equiv -1$. Here, the initial cost $I[u^0] \approx 5.3$ is reduced to $I[u^{10}] \approx 1.67$ after 10 iterations, yielding a solution that — although numerically acceptable — is less accurate than that of the hyperbolic case. The computed control and corresponding density snapshots are shown in Figs. 4 and 3.

We observe that the "optimal" control profile in Case II is nearly identical to that obtained in Case I, while

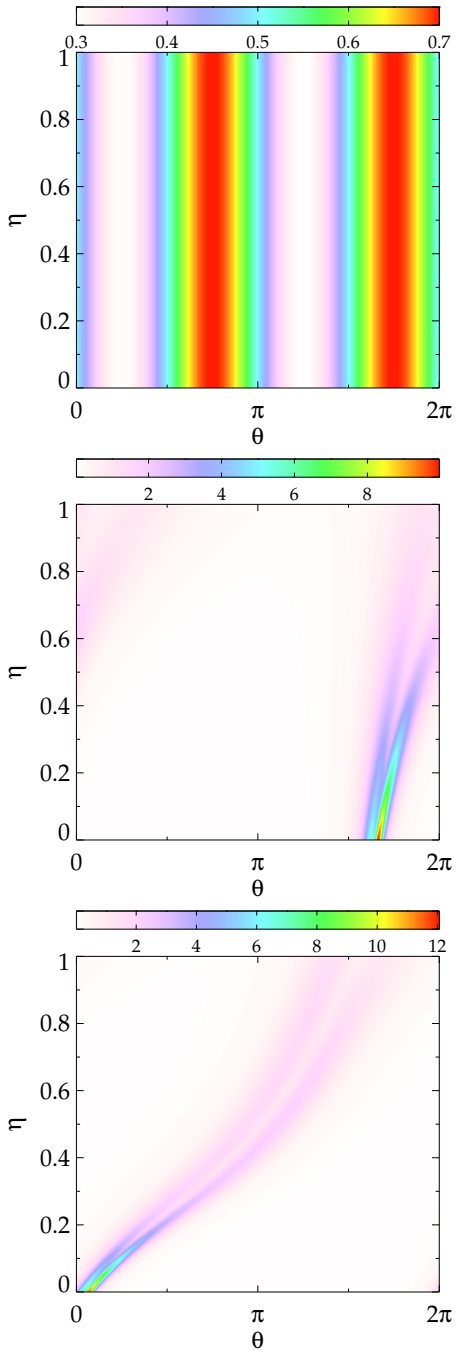

Fig. 1. Case I. Density snapshots, ρ_t, of the "optimal" solution to (3) at time moments $t = 0$ (top panel), $t = T/2$ (middle panel) and $t = T$ (bottom panel).

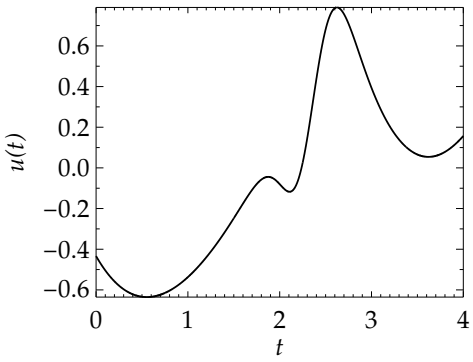

Fig. 2. Case I. The "optimal" control $u = u^{10}$.

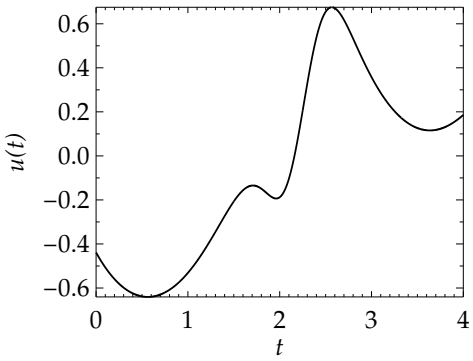

Fig. 3. Case II. The "optimal" control $u = u^{10}$.

the corresponding density heatmaps naturally exhibit greater dispersion as a result of diffusion.

In comparison, the hyperbolic (continuity equation) case tends to produce solutions with sharper gradients and more pronounced localization effects. The diffusion in the parabolic case acts as a regularizing mechanism, smoothing out steep features and thereby stabilizing the numerical solution. However, when employing pseudospectral methods, one must exercise caution: the regularization can mask localized phenomena, and aliasing errors may become significant if the resolution is not sufficiently high.

V. CONCLUSION

In this paper, we have numerically validated an approach based on the mean-field control paradigm, leveraging exact cost-increment formulas to tackle optimization and synchronization challenges in oscillatory processes.

Using the Theta model as an illustrative example, we demonstrated that even one of the simplest models in this context can capture essential dynamics. The proposed technique is remarkably versatile, extending to a broad class of optimal control problems with a state-linear structure, and is applicable to both finite and infinite non-interacting statistical ensembles with diverse characteristics.

Future work will extend this framework to interacting ensembles and more complex network structures, further broadening its applicability to a wider range of synchronization and control challenges.

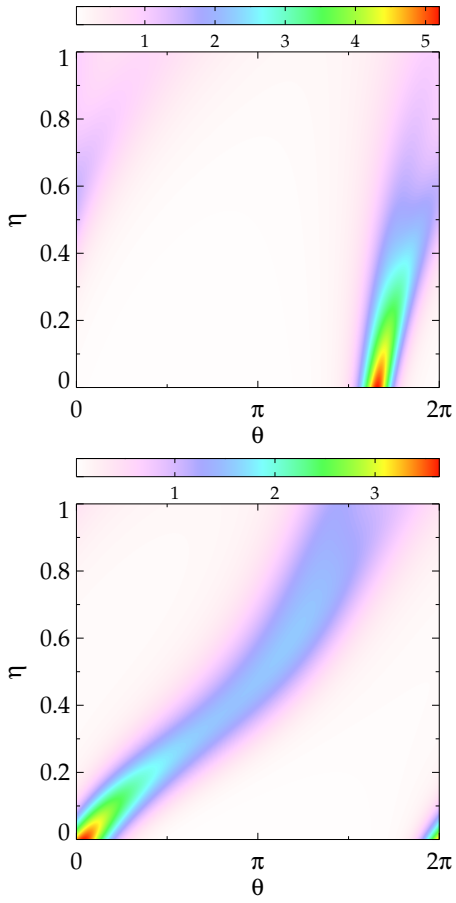

Fig. 4. Case II. Density snapshots, ρ_t, of the "optimal" solution to (3) at time moments $t = T/2$ (top panel) and $t = T$ (bottom panel).

REFERENCES

[1] I. Fischer, Y. Liu, and P. Davis, "Synchronization of chaotic semiconductor laser dynamics on subnanosecond time scales and its potential for chaos communication," *Phys. Rev. A*, vol. 62, p. 011801, Jun 2000. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.62.011801

[2] I. Blekhman, I. Blekhman, and E. Rivin, *Synchronization in Science and Technology*, ser. ASME press translations. ASME Press, 1988. [Online]. Available: https://books.google.ru/books?id=ao1QAAAAMAAJ

[3] T. Nishikawa, N. Gulbahce, and A. E. Motter, "Spontaneous reaction silencing in metabolic optimization," *PLOS Computational Biology*, vol. 4, no. 12, pp. 1–12, 12 2008. [Online]. Available: https://doi.org/10.1371/journal.pcbi.1000236

[4] P. J. Uhlhaas and W. Singer, "Neural synchrony in brain disorders: Relevance for cognitive dysfunctions and pathophysiology," *Neuron*, vol. 52, no. 1, pp. 155–168, 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0896627306007276

[5] S. J. Schiff, K. Jerger, D. H. Duong, T. Chang, M. L. Spano, and W. L. Ditto, "Controlling chaos in the brain," *Nature*, vol. 370, no. 6491, pp. 615–620, Aug 1994. [Online]. Available: https://doi.org/10.1038/370615a0

[6] L. Glass, "Cardiac arrhythmias and circle maps-a classical problem," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 1, no. 1, pp. 13–19, 1991. [Online]. Available: https://doi.org/10.1063/1.165810

[7] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*, ser. Dover books on chemistry. Dover Publications, 2003. [Online]. Available: https://books.google.ru/books?id=4ADt7smO5Q8C

[8] J.-S. Li, "Ensemble control of finite-dimensional time-varying linear systems," *IEEE Trans. Autom. Control*, vol. 56, no. 2, pp. 345–357, 2011.

[9] J.-S. Li, I. Dasanayake, and J. Ruths, "Control and synchronization of neuron ensembles," *IEEE Transactions on Automatic Control*, vol. 58, no. 8, pp. 1919–1930, 2013.

[10] R. Chertovskih, N. Pogodaev, M. Staritsyn, J. Da Silva Sewane, and A. P. Aguiar, "Optimization of external stimuli for populations of theta neurons via mean-field feedback control," in *2023 9th International Conference on Control, Decision and Information Technologies (CoDIT)*, 2023, pp. 512–517.

[11] M. Staritsyn, N. Pogodaev, and F. L. Pereira, "Linear-quadratic problems of optimal control in the space of probabilities," *IEEE Control Systems Letters*, vol. 6, pp. 3271–3276, 2022.

[12] G. B. Ermentrout and N. Kopell, "Parabolic bursting in an excitable system coupled with a slow oscillation," *SIAM Journal on Applied Mathematics*, vol. 46, no. 2, pp. 233–253, 1986. [Online]. Available: http://www.jstor.org/stable/2101582

[13] S. H. Strogatz, *Nonlinear dynamics and chaos : with applications to physics, biology, chemistry, and engineering*. Second edition. Boulder, CO : Westview Press, a member of the Perseus Books Group, [2015], [2015], includes bibliographical references and index. [Online]. Available: https://search.library.wisc.edu/catalog/9910223127702121

[14] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, ser. Universitext. Springer Berlin Heidelberg, 2010.

[15] L. Ambrosio and G. Savaré, "Gradient flows of probability measures," in *Handbook of differential equations: evolutionary equations. Vol. III*, ser. Handb. Differ. Equ. Elsevier/North-Holland, Amsterdam, 2007, pp. 1–136.

[16] V. Bogachev, N. Krylov, M. Röckner, and S. Shaposhnikov, *Fokker-Planck-Kolmogorov Equations*, ser. Mathematical Surveys and Monographs. American Mathematical Society, 2015.

[17] N. Pogodaev, "Program strategies for a dynamic game in the space of measures," *Optimization Letters*, vol. 13, no. 8, pp. 1913–1925, Nov 2019. [Online]. Available: https://doi.org/10.1007/s11590-018-1318-y

[18] B. Bonnet, C. Cipriani, M. Fornasier, and H. Huang, "A measure theoretical approach to the mean-field maximum principle for training neurodes," *Nonlinear Analysis*, vol. 227, p. 113161, 2023.

[19] M. Staritsyn, N. Pogodaev, R. Chertovskih, and F. L. Pereira, "Feedback maximum principle for ensemble control of local continuity equations: An application to supervised machine learning," *IEEE Control Systems Letters*, vol. 6, pp. 1046–1051, 2022.

[20] R. Chertovskih, N. Pogodaev, M. Staritsyn, and A. Pedro Aguiar, "Optimal control of diffusion processes: Infinite-order variational analysis and numerical solution," *IEEE Control Systems Letters*, vol. 8, pp. 1469–1474, 2024.

[21] N. Pogodaev, "Optimal control of continuity equations," *NoDEA Nonlinear Differential Equations Appl.*, vol. 23, no. 2, pp. Art. 21, 24, 2016.

[22] A. V. Arguchintsev, V. A. Dykhta, and V. A. Srochko, "Optimal control: Nonlocal conditions, computational methods, and the variational principle of maximum," *Russian Mathematics*, vol. 53, no. 1, pp. 1–35, Jan 2009. [Online]. Available: https://doi.org/10.3103/S1066369X09010010

[23] P. Drag, K. Styczen, M. Kwiatkowska, and A. Szczurek, "A review on the direct and indirect methods for solving optimal control problems with differential-algebraic constraints," *Studies in Computational Intelligence*, vol. 610, pp. 91–105, 07 2016.

[24] A. Borzì, *The Sequential Quadratic Hamiltonian Method: Solving Optimal Control Problems*, ser. Chapman & Hall/CRC Numerical Analysis and Scientific Computing Series. CRC Press, 2023.

[25] J. P. Boyd, *Chebyshev and Fourier spectral methods.*, 2nd ed. Mineola, NY: Dover Publications, 2001.

[26] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical recipes in Fortran. The art of scientific computing.*, 2nd ed. Cambridge University Press, 1992.