Data-driven entropic spatially inhomogeneous evolutionary games

MAURO BONAFINI[1], MASSIMO FORNASIER[2], and BERNHARD SCHMITZER[1]

[1]*Campus Institute Data Science, University of Göttingen, Göttingen, Germany*
emails: bonafini@cs.uni-goettingen.de, schmitzer@cs.uni-goettingen.de
[2]*Department of Mathematics, Technical University of Munich, Garching, Germany*
email: massimo.fornasier@ma.tum.de

We introduce novel multi-agent interaction models of entropic spatially inhomogeneous evolutionary undisclosed games and their quasi-static limits. These evolutions vastly generalise first- and second-order dynamics. Besides the well-posedness of these novel forms of multi-agent interactions, we are concerned with the learnability of individual payoff functions from observation data. We formulate the payoff learning as a variational problem, minimising the discrepancy between the observations and the predictions by the payoff function. The inferred payoff function can then be used to simulate further evolutions, which are fully data-driven. We prove convergence of minimising solutions obtained from a finite number of observations to a mean-field limit, and the minimal value provides a quantitative error bound on the data-driven evolutions. The abstract framework is fully constructive and numerically implementable. We illustrate this on computational examples where a ground truth payoff function is known and on examples where this is not the case, including a model for pedestrian movement.

1 Introduction

1.1 Multi-agent interaction models in biological, social and economical dynamics

In the course of the past two decades, there has been an explosion of research on models of multi-agent interactions [28, 29, 31, 36, 40, 42, 74], to describe phenomena beyond physics, for example, in biology, such as cell aggregation and motility [11, 43, 45, 61], coordinated animal motion [8, 16, 18, 21, 22, 24, 31, 54, 57, 58, 63, 71, 77], coordinated human [25, 30, 68], and synthetic agent behaviour and interactions, such as cooperative robots [19, 49, 60, 70]. We refer to [13, 15, 17, 75] for recent surveys.

Two main mechanisms are considered in such models to define the dynamics. The first takes inspiration from physical laws of motion and is based on pairwise forces encoding observed 'first principles' of biological, social, or economical interactions, for example, repulsion-attraction, alignment, self-propulsion/friction et cetera. Typically, this leads to Newton-like first- or second-order equations with 'social interaction' forces, see (1.1) below. In the second mechanism, the dynamics are driven by an evolutive game where players simultaneously optimise their cost.

Examples are game theoretic models of evolution [38] or mean-field games, introduced in [48] and independently under the name Nash Certainty Equivalence (NCE) in [39], later greatly popularised, for example, within consensus problems, for instance in [55, 56].

More recently, the notion of spatially inhomogeneous evolutionary games has been proposed [5, 6, 53] where the transport field for the agent population is directed by an evolutionary game on their available strategies. Unlike mean-field games, the optimal dynamics are not chosen via an underlying non-local optimal control problem but by the agents' local (in time and space) decisions, see (1.3) below.

One fundamental goal of these studies is to reveal the relationship between the simple pairwise forces or incentives acting at the individual level and the emergence of global patterns in the behaviour of a given population.

Newtonian models. A common simple class of models for interacting multi-agent dynamics with pairwise interactions is inspired by Newtonian mechanics. The evolution of N agents with time-dependent locations $x_1(t), \ldots, x_N(t)$ in $\mathbb{R}^d$ is described by the ODE system

$$\partial_t x_i(t) = \frac{1}{N} \sum_{j=1}^{N} f\big(x_i(t), x_j(t)\big) \qquad \text{for } i = 1, \ldots, N, \tag{1.1}$$

where f is a pre-determined pairwise interaction force between pairs of agents. The system is well-defined for sufficiently regular f (e.g. for f Lipschitz continuous). In this article, we will refer to such models as *Newtonian models*.

First-order models of the form (1.1) are ubiquitous in the literature and have, for instance, been used to model multi-agent interactions in opinion formation [37, 46], vehicular traffic flow [35], pedestrian motion [26], and synchronisation of chemical and biological oscillators in neuroscience [47].

Often one is interested in studying the behaviour of a very large number of agents. We may think of the agents at time t to be distributed according to a probability measure $\mu(t)$ over $\mathbb{R}^d$. The limit-dynamics of (1.1) as $N \to \infty$ can under suitable conditions be expressed directly in terms of the evolution of the distribution $\mu(t)$ according to a *mean-field equation*. The mean-field limit of (1.1) is formally given by

$$\partial_t \mu(t) + \text{div}\big(v(\mu(t)) \cdot \mu(t)\big) = 0 \quad \text{where} \quad v(\mu(t))(x) := \int_{\mathbb{R}^d} f(x, x') \, d\mu(t)(x'). \tag{1.2}$$

Here $v(\mu(t))$ is a velocity field and intuitively $v(\mu(t))(x)$ gives the velocity of infinitesimal mass particles of μ at time t and location x [12, 14].

While strikingly simple, such models only exhibit limited modelling capabilities. For instance, the resulting velocity $\partial_t x_i(t)$ is simply a linear combination of the influences of all the other agents. 'Importance' and 'magnitude' of these influences cannot be specified separately: agent j cannot tell agent i to remain motionless, regardless of what other agents suggest. Generally, agent i has no sophisticated mechanism of finding a 'compromise' between various potentially conflicting influences and merely uses their linear average. The applicability of such models to economics or sociology raises concerns as it is questionable whether a set of static interaction forces can describe the behaviour of rational agents who are able to anticipate and counteract undesirable situations.

Spatially inhomogeneous evolutionary games. In a game dynamic, the vector field $v(\mu(t))$ is not induced by a rigid force law but by the optimisation of an individual payoff by each agent. In mean-field games, this optimisation is global in time, each agent plans their whole trajectory in advance, optimisation can be thought of as taking place over repetitions of the same game (e.g. the daily commute). Alternatively, in spatially inhomogeneous evolutionary games [6] the agents continuously update their mixed strategies in a process that is local in time, which may be more realistic in some scenarios. This is implemented by the well-known replicator dynamics [38].

As above, agents may move in $\mathbb{R}^d$. Let U be the set of pure strategies. The map $e : \mathbb{R}^d \times U \to \mathbb{R}^d$ describes the spatial velocity $e(x, u)$ for an agent at position $x \in \mathbb{R}^d$ with a pure strategy $u \in U$. For example, one can pick $U \subset \mathbb{R}^d$ to be a set of admissible velocity vectors and simply set $e(x, u) = u$. e acts as dictionary between strategies and velocities and can therefore assumed to be known. $e(x, u)$ may be more complicated, for instance, when certain velocities are inadmissible at certain locations due to obstacles. A function $J : (\mathbb{R}^d \times U)^2 \to \mathbb{R}$ describes the individual benefit $J(x, u, x', u')$ for an agent at x for choosing strategy u when another agent sits at x' and intends to choose strategy u'. For example, $J(x, u, x', u')$ may be high when $e(x, u)$ points in the direction that x wants to move, but could be lowered when the other agent's velocity $e(x', u')$ suggests an impending collision. The state of each agent is given by their spatial position $x \in \mathbb{R}^d$ and a mixed strategy $\sigma \in \mathcal{M}_1(U)$ (where $\mathcal{M}_1(U)$ denotes probability measures over U). As hinted at, the evolution of σ is driven by a replicator dynamic involving the payoff function J, averaged over the benefit obtained from all other agents, the spatial velocity $\partial_t x$ is obtained by averaging $\int_U e(x, u) \, d\sigma(u)$. The equations of motion are given by

$$\partial_t x_i(t) = \int_U e(x_i(t), u) \, d\sigma_i(t)(u), \tag{1.3a}$$

$$\partial_t \sigma_i(t) = \frac{1}{N} \sum_{j=1}^{N} f^J\big(x_i(t), \sigma_i(t), x_j(t), \sigma_j(t)\big) \tag{1.3b}$$

where for $x, x' \in \mathbb{R}^d$ and $\sigma, \sigma' \in \mathcal{M}_1(U)$

$$f^J\big(x, \sigma, x', \sigma'\big) := \left[\int_U J(x, \cdot, x', u') \, d\sigma'\left(u'\right) - \int_U \int_U J(x, v, x', u') d\sigma'\left(u'\right) d\sigma(v) \right] \cdot \sigma. \tag{1.4}$$

Intuitively, the strategy σ_i tends to concentrate on the strategies with the highest benefit for agent i via (1.4), which will then determine their movement via (1.3a).

Equation (1.3a) describes the motion of agents in $\mathbb{R}^d$ and resembles (1.1). The main difference is that the vector field is determined by the resulting solution of the replicator equation (1.3b), which promotes strategies that perform best with respect to the individual payoff J with rate (1.4). Despite the different nature of the variables (x_i, σ_i) of the model, the system (1.3) can also be re-interpreted as a Newtonian-like dynamics on the space $\mathbb{R}^d \times \mathcal{M}_1(U)$ where each agent $y_i(t) = (x_i(t), \sigma_i(t))$ is characterised by its spatial location $x_i(t) \in \mathbb{R}^d$ and its mixed strategy $\sigma_i(t) \in \mathcal{M}_1(U)$. Accordingly, equations (1.3) can be more concisely expressed as

$$\partial_t y_i(t) = \frac{1}{N} \sum_{j=1}^{N} f\big(y_i(t), y_j(t)\big) \quad \text{where} \quad f\big((x, \sigma), (x', \sigma')\big) := \left(\int_U e(x, u) \, d\sigma(u), f^J\big(x, \sigma, x', \sigma'\big) \right).$$

$$\tag{1.5}$$

Similarly to mean-field limit results in [12, 14], the main contribution of [6] is to show that the large particle limit for $N \to \infty$ of solutions of (1.3) converges in the sense of probability measures to $\Sigma(t) \in \mathcal{M}_1(\mathbb{R}^d \times \mathcal{M}_1(U))$, which is, in analogy to (1.2), solution of a nonlinear transport equation of the type

$$\partial_t \Sigma(t) + \mathrm{Div}\big(v(\Sigma(t)) \cdot \Sigma(t)\big) = 0 \quad \text{where} \quad v(\Sigma(t))(y) := \int f\big(y, y'\big) \mathrm{d}\Sigma(t)(y'). \qquad (1.6)$$

Note that (1.6) is a partial differential equation having as domain a (possibly infinite-dimensional) Banach space containing $\mathbb{R}^d \times \mathcal{M}_1(U)$ and particular care must be applied to the use of an appropriate calculus needed to define differentials, in particular for the divergence operator. In [6], Lagrangian and Eulerian notions of solutions to (1.6) are introduced. As the technical details underlying the well-posedness of (1.6) do not play a central role in this article, we refer to [6] for more insights. Nevertheless, we apply some of the general statements in [6] to establish well-posedness of the models presented in this paper.

1.2 Learning or inference of multi-agent interaction models

An important challenge is to determine the precise form of the interaction force or payoff functions. While physical laws can often be determined with high precision through experiments, such experiments are often either not possible, or the models are not as precisely determined in more complex systems from biology or social sciences. Therefore, very often the governing rules are simply chosen ad hoc to reproduce at some major qualitative effects observed in reality.

Alternatively, one can employ model selection and parameter estimation methods to determine the form of the governing functions. Data-driven estimations have been applied in continuum mechanics [20, 44] computational sociology [10, 50, 51, 78] or economics [3, 23, 33]. However, even the problem of determining whether time shots of a linear dynamical system do fulfil physically meaningful models, in particular have Markovian dynamics, is computationally intractable [27]. For nonlinear models, the intractability of learning the system corresponds to the complexity of determining the set of appropriate candidate functions to fit the data. In order to break the curse of dimensionality, one requires prior knowledge on the system and the potential structure of the governing equations. For instance, in the sequence of recent papers [64, 65, 72] the authors assume that the governing equations are of first order and can be written as sparse polynomials, that is, linear combinations of few monomial terms.

In this work, we present an approach for estimating the payoff function for spatially inhomogeneous evolutionary games from observed data. It is inspired by the papers [10, 50, 51, 78], in particular by the groundbreaking paper [10] which is dedicated to data-driven evolutions of Newtonian models. In these references, the curse of dimensionality is remedied by assuming that the interaction function f in (1.1) is parametrised by a lower-dimensional function a, the identification of which is more computationally tractable. A typical example for such a parametrisation is given by functions $f = f^a$ of the type

$$f^a\big(x, x'\big) = a\big(|x - x'|\big) \cdot \big(x' - x\big) \qquad \text{for some} \qquad a : \mathbb{R} \to \mathbb{R}.$$

The corresponding model (1.1) is used, for instance, in opinion dynamics [37, 46], vehicular traffic [35], or pedestrian motion [26]. The learning or inference problem is then about the determination of a, hence of $f = f^a$, from observations of real-life realisations of the model (1.1).

Clearly, the problem can be formulated as an optimal control problem. However, as pointed out in [10], in view of the nonlinear nature of the function mapping a in the corresponding solution $x^{[a]}(t)$ of (1.1), the control cost would be nonconvex and the optimal control problem rather difficult to solve. Instead, in [10] a convex formulation is obtained by considering an empirical risk minimisation in least squares form: given an observed realisation $(x_1^N(t), \ldots, x_N^N(t))$ of the dynamics between N agents, generated by (1.1) governed by f^a, we aim at identifying a by minimising

$$\mathcal{E}^N(\hat{a}) := \frac{1}{T} \int_0^T \left[\frac{1}{N} \sum_{i=1}^N \left\| \partial_t x_i^N(t) - \frac{1}{N} \sum_{j=1}^N f^{\hat{a}} \left(x_i^N(t), x_j^N(t) \right) \right\|^2 \right] \mathrm{d}t. \tag{1.7}$$

That is, along the observed trajectories we aim to minimise the discrepancy between observed velocities and those predicted by the model. The functional $\mathcal{E}^N$ plays the role of a *loss function* in the learning or inference task. A time-discrete formulation in the framework of statistical learning theory has been proposed also in [50, 51, 78]. Under the assumption that a belongs to a suitable compact class X of smooth functions, the main results in [10, 50, 51, 78] establish that minimisers

$$\hat{a}_N \in \arg \min_{\hat{a} \in V_N} \mathcal{E}^N(\hat{a}),$$

converge to a in suitable senses for $N \to \infty$, where the ansatz spaces $V_N \subset X$ are suitable finite-dimensional spaces of smooth functions (such as finite element spaces).

1.3 Contribution and outline

The main scope of this paper is to provide a theoretical and practical framework to perform learning/inference of spatially inhomogeneous evolutionary games, so that these models could be used in real-life data-driven applications. First, we discuss some potential issues with model (1.3) and provide some adaptations. We further propose and study in Section 3 several learning functionals for inferring the payoff function J from the observation of the dynamics, that is, extending the approach of [10] to our modified version of [6]. Let us detail our contributions.

The proposed changes to the model (1.3) are as follows:

- In Section 2.1, we add entropic regularisation to the dynamics for the mixed strategies of the agents. This avoids degeneracy of the strategies and allows faster reactions to changes in the environment, see Figure 1. Entropic regularisation of games was also considered in [34]. We show that the adapted model is well-posed and has a consistent mean-field limit (Theorem 2.6).

- For interacting agents, the assumption that the mixed strategies of other agents are fully known may often be unrealistic. Therefore, we will focus our analysis on the case where $J(x, u, x', u')$ does not explicitly depend on u', the 'other' agent's pure strategy. We refer to this as the *undisclosed* setting. (The general case is still considered to some extent.)

For this undisclosed setting, we study the quasi-static *fast-reaction* limit of (1.3) where the dynamics of mixed strategies $(\sigma_i)_{i=1}^N$ run at much faster time scale than dynamics of locations $(x_i)_{i=1}^N$, that is, agents quickly adjust their strategies to changes in the environment. This will also
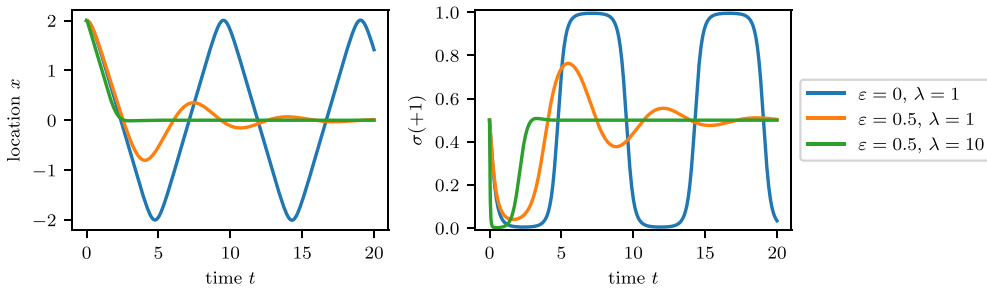
FIGURE 1. Comparison of original game dynamics and entropic regularisation (possibly with accelerated time scale for the strategy dynamics). For details, see Examples 2.1 and 2.2.

be important for designing practical inference functionals (see below). The *undisclosed fast-reaction* model is introduced in Section 2.2. Well-posedness and consistent mean-field limit are proved by Theorem 2.13, convergence to the fast-reaction limit or quasi-static evolution as the strategy time scale becomes faster is given by Theorem 2.14.

We claim that the resulting undisclosed fast-reaction entropic model is a useful alternative of the Newtonian model (1.1) and we support these considerations with theoretical and numerical examples. In particular, we show that any Newtonian model can be described (approximately) as an undisclosed fast-reaction entropic model, whereas the converse it not true (Examples 2.7 and 2.8).

We then discuss several inference functionals for the modified game model in Section 3. We start with a rather direct analogue of (1.7) (Section 3.2), which would require not only the observation of the spatial locations $(x_i)_{i=1}^N$ and velocities $(\partial_t x_i)_{i=1}^N$ but also of the mixed strategies $(\sigma_i)_{i=1}^N$, and their temporal derivatives $(\partial_t \sigma_i)_{i=1}^N$. Whether the latter two can be observed in practice is doubtful, in particular with sufficient accuracy. Therefore, as an alternative, we propose two inference functionals for the undisclosed fast-reaction setting. A functional based on observed spatial velocities is given in Section 3.3.1. A functional based on mixed strategies (but not their derivatives) is proposed in Section 3.3.2, and we discuss some options how the required data could be obtained in practice. In Section 3.3.3, some properties of these functionals are established, such as existence of minimisers and consistency of their mean-field versions.

Numerical examples are given in Section 4. These include the inference on examples where observations were generated with a true underlying undisclosed fast-reaction entropic model, as well as inference of Newtonian models and of a model for pedestrian motion adapted from [7, 32]. We close with a brief discussion in Section 5. Some longer technical proofs have been moved to Appendix A. Before presenting the new model, we collect some notation in the next subsection.

1.4 Setting and notation

General setting. Let (Y, d_Y) be a complete and separable metric space. We denote by $\mathcal{M}(Y)$ the space of signed Borel measures with finite total variation and by $\mathcal{M}_1(Y)$ the set of probability measures with bounded first moment, that is,

$$\mathcal{M}_1(Y) = \left\{ \mu \in \mathcal{M}(Y) \mid \mu \geq 0, \, \mu(Y) = 1, \, \int_Y d_Y(y, \bar{y}) \mathrm{d}\mu(y) < \infty \text{ for some } \bar{y} \in Y \right\}.$$

For a continuous function $\varphi \in C(Y)$, we denote by

$$\text{Lip}(\varphi) := \sup_{\substack{x,y \in Y \\ x \neq y}} \frac{|\varphi(x) - \varphi(y)|}{d_Y(x,y)}$$

its Lipschitz constant. The set of bounded Lipschitz functions is then denoted by $\text{Lip}_b(Y) = \{\varphi \in C(Y) \mid \|\varphi\|_\infty + \text{Lip}(\varphi) < \infty\}$, with $\|\cdot\|_\infty$ the classical sup-norm.

For $\mu_1, \mu_2 \in \mathcal{M}_1(Y)$, the 1-Wasserstein distance $W_1(\mu_1, \mu_2)$ is defined by

$$W_1(\mu_1, \mu_2) := \inf\left\{ \int_{Y \times Y} d_Y(y_1, y_2) \mathrm{d}\gamma(y_1, y_2) \mid \gamma \in \Gamma(\mu_1, \mu_2) \right\}, \qquad (1.8)$$

where $\Gamma(\mu_1, \mu_2)$ is the set of admissible coupling between μ_1 and μ_2, that is, $\Gamma(\mu_1, \mu_2) = \{\gamma \in \mathcal{M}_1(Y \times Y) \mid \gamma(A \times Y) = \mu_1(A)$ and $\gamma(Y \times B) = \mu_2(B)\}$. Due to Kantorovitch duality, one can also consider the equivalent definition

$$W_1(\mu_1, \mu_2) := \sup\left\{ \int_Y \varphi \, \mathrm{d}(\mu_1 - \mu_2) \mid \varphi \in \text{Lip}_b(Y), \text{Lip}(\varphi) \leq 1 \right\}. \qquad (1.9)$$

The metric space $(\mathcal{M}_1(Y), W_1)$ is complete because (Y, d_Y) is complete, and a sequence $(\mu_n)_{n \in \mathbb{N}} \subset \mathcal{M}_1(Y)$ converges to $\mu \in \mathcal{M}_1(Y)$ with respect to the distance W_1 if and only if μ_n converges to μ in duality with bounded Lipschitz functions and the first moments converge, that is, $\int_Y d_Y(\cdot, \bar{y}) \, \mathrm{d}\mu_n \to \int_Y d_Y(\cdot, \bar{y}) \, \mathrm{d}\mu$ for all $\bar{y} \in Y$.

Interaction setting. We fix the space of pure strategies U to be a compact metric space with distance d_U. Each agent's mixed strategy is then described by $\sigma \in \mathcal{M}_1(U)$. Agents move in $\mathbb{R}^d$, and we denote with $\|\cdot\|$ the usual Euclidean norm. For $R > 0$, $B^d(R)$ is the open ball of radius R and centre the origin in $\mathbb{R}^d$. For $N \in \mathbb{N}$ and $\mathbf{x} = (x_1, \ldots, x_N) \in [\mathbb{R}^d]^N$, we introduce the scaled norm $\|\cdot\|_N$ defined as

$$\|\mathbf{x}\|_N = \frac{1}{N} \sum_{i=1}^N \|x_i\|.$$

For $\sigma_1, \sigma_2 \in \mathcal{M}_1(U)$, we set the Kullback–Leibler divergence to be

$$\text{KL}(\sigma_1|\sigma_2) := \begin{cases} \int_U \log\left(\frac{\mathrm{d}\sigma_1}{\mathrm{d}\sigma_2}\right) \mathrm{d}\sigma_1 & \text{if } \sigma_1 \ll \sigma_2, \\ +\infty & \text{else,} \end{cases}$$

where $\mathrm{d}\sigma_1/\mathrm{d}\sigma_2$ is the Radon–Nikodym derivative of σ_1 with respect to σ_2.

Throughout the paper, we assume $e \in \text{Lip}_b(\mathbb{R}^d \times U; \mathbb{R}^d)$ to be fixed and known. Such a function maps each pure strategy $u \in U$ into a given velocity $e(\cdot, u) \in \mathbb{R}^d$. We may think of e as a label or dictionary for strategies. The sets of admissible interaction rules J are described by

$$\mathcal{X} = \text{Lip}_b\left(\mathbb{R}^d \times U \times \mathbb{R}^d \times U\right) \quad \text{and} \quad X = \text{Lip}_b\left(\mathbb{R}^d \times U \times \mathbb{R}^d\right).$$

Here, $\mathcal{X}$ consists of all possible payoff functions, modelling the case where each agent has a complete knowledge of both positions and mixed strategies of the other agents. On the other hand, in what we call the *undisclosed* setting, each agent only knows the physical locations of the others, and thus in this context, the payoff function J will no longer depend on the other's strategy, making it possible to restrict the analysis to X.

2 From spatially inhomogeneous evolutionary games to undisclosed fast-reaction dynamics

We describe in this section how we adapt the spatially inhomogeneous evolutionary games model (1.3) of [6] to make it more suitable in practice for the task of modelling and inferring interaction rules between rational agents. In Section 2.1, we add entropic regularisation to avoid degeneracy of the agents' mixed strategies. In Section 2.2, we focus on the more realistic undisclosed setting, where agents are not aware of other agents' strategy choices and derive a fast-reaction limit, describing the regime where choice of the strategies happens at a much faster time scale than physical movement of the agents.

2.1 Entropic regularisation for spatially inhomogeneous evolutionary games

In the model (1.3), mixed strategies σ_i are prone to converging exponentially to very singular distributions so that agents cannot react quickly to changes in the environment and exhibit strong inertia, contradicting the notion of rational agents. This behaviour can be illustrated with a simple example.

Example 2.1 (Slow reaction in spatially inhomogeneous evolutionary games). Let $d = 1$, $U = \{-1, +1\}$, $e(x, u) = u$, so pure strategies correspond to moving left or right with unit speed. Let now $J(x, u, x', u') = -x \cdot u$, that is, agents prefer moving towards the origin, independently of the other agents' locations and actions. For simplicity, we can set $N = 1$ and choose $x(0) = 2$, $\sigma(0) = (0.5, 0.5)$, where we identify measures on $U = \{-1, +1\}$ with vectors of the simplex in $\mathbb{R}^2$. At $t = 0$, the agent perceives strategy -1 as more attractive. The mixed strategy σ starts pivoting towards this pure strategy and the agent starts moving towards the origin. The numerically simulated trajectory for the original model () is shown in Figure 1 (corresponding to line $\varepsilon = 0$, $\lambda = 1$). We can see how σ rapidly converges to the state (0,1). Thus when the agent reaches the origin, it cannot react quickly and there is considerable overshoot.

As a remedy, we propose an adaptation of the model (1.3) by adding an entropic term to the dynamics of the mixed strategies σ_i. The update rule (1.3b) modifies σ_i towards maximising the average benefit

$$\frac{1}{N} \sum_{j=1}^{N} \int_{U \times U} J(x_i, u, x_j, u') \, \mathrm{d}\sigma_i(u) \, \mathrm{d}\sigma_j(u') \, .$$

To this objective, we will now add the entropy $\varepsilon \cdot \int_U \log\left(\frac{\mathrm{d}\sigma_i}{\mathrm{d}\eta}\right) \mathrm{d}\sigma_i$ where $\eta \in \mathcal{M}_1(U)$ is a suitable 'uniform' reference measure and $\varepsilon > 0$ is a weighting parameter. This will pull σ_i towards η and thus prevent exponential degeneration.

Entropic regularisation. The entropic regularisation implies that mixed strategies are no longer general probability measures over U, but they become densities with respect to a reference measure $\eta \in \mathcal{M}_1(U)$, which we can assume without loss of generality to have full support, that is, $\mathrm{spt}(\eta) = U$. Thus, for a fixed $\varepsilon > 0$, set

$$\mathcal{S}_{a,b} := \left\{ \sigma : U \to \mathbb{R}_+, \ \sigma \text{ measurable}, \sigma(u) \in [a, b] \ \eta\text{-a.e.}, \int_U \sigma \, \mathrm{d}\eta = 1 \right\}, \quad (2.1)$$

where $0 < a < 1 < b < \infty$ (the bounds a and b are required for technical reasons, see below). For $\lambda > 0$, the modified entropic dynamics is then given by

$$\partial_t x_i(t) = \int_U e(x_i(t), u)\, \sigma_i(t)(u)\, \mathrm{d}\eta(u), \tag{2.2a}$$

$$\partial_t \sigma_i(t) = \lambda \cdot \left[\frac{1}{N} \sum_{j=1}^N f^J\big(x_i(t), \sigma_i(t), x_j(t), \sigma_j(t)\big) + f^\varepsilon\big(\sigma_i(t)\big) \right], \tag{2.2b}$$

where the function f^J now formally needs to be given in terms of densities, instead of measures:

$$f^J(x, \sigma, x', \sigma') :=$$
$$\left[\int_U J(x, \cdot, x', u')\, \sigma'(u')\mathrm{d}\eta(u') - \int_U \int_U J(x, v, x', u')\, \sigma'(u')\, \sigma(v)\, \mathrm{d}\eta(u')\mathrm{d}\eta(v) \right] \cdot \sigma. \tag{2.3}$$

The additional term f^ε, corresponding to entropy regularisation, is given by:

$$f^\varepsilon(\sigma) := \varepsilon \cdot \left[-\log(\sigma(\cdot)) + \int_U \log(\sigma(v))\, \sigma(v)\, \mathrm{d}\eta(v) \right] \cdot \sigma. \tag{2.4}$$

In (2.2b), we have explicitly added the factor $\lambda \in (0, \infty)$ to control the relative time scale of the dynamics for mixed strategies and locations.

Example 2.2 (Faster reactions with entropic regularisation). We repeat here Example 2.1 with added entropy. We set $\eta = (0.5, 0.5)$, keeping the symmetry between strategies $+1$ and -1 in the regularised system. Numerically simulated trajectories for this setup with different choices of λ and ε are shown in Figure 1. For $(\lambda = 1, \varepsilon = 0.5)$, the strategy does not get as close to $(0,1)$ as for $(\lambda = 1, \varepsilon = 0)$. Therefore, upon crossing the origin the agent can react quicker, there is less overshoot and the trajectory eventually converges to 0. Finally, for $(\lambda = 10, \varepsilon = 0.5)$ the agent moves quickly and without overshoot to the origin, quickly adapting the mixed strategy.

Remark 2.3 (Entropic gradient flow). *Formally, the contribution of f^ε to the dynamics (2.2b) corresponds to a gradient flow in the Hellinger–Kakutani metric over $\mathcal{S}_{a,b}$ of the (negative) entropy*

$$H(\sigma) := \int_U \log(\sigma)\, \sigma\, \mathrm{d}\eta \tag{2.5}$$

of the density σ with respect to the reference measure η.

Remark 2.4 (Multiple agent species). *In the interaction models mentioned so far ((1.1), (1.3), (2.2)), all agents are of the same species (i.e. their movement is specified by the same functions). Often one wishes to model the interaction between agents of multiple species (e.g. predators and prey). The latter case can formally be subsumed into the former by expanding the physical space from $\mathbb{R}^d$ to $\mathbb{R}^{1+d}$ and using the first spatial dimension as 'species label', that is, an agent $\hat{x}_i = (n_i, x_i) \in \mathbb{R}^{1+d}$ describes an agent of species $n_i \in \mathbb{Z}$ at position $x_i \in \mathbb{R}^d$. The interaction force $\hat{f} : \mathbb{R}^{1+d} \times \mathbb{R}^{1+d} \to \mathbb{R}^{1+d}$ in the expanded version of (1.1) is then given by $\hat{f}\big((n, x), (n', x')\big) :=$*

$(0, f_{n,n'}(x, x'))$ where $f_{n,n'} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ is the interaction function for species (n,n') and we set the first component of $\hat{f}$ to zero, such that the species of each agent remains unchanged. In an entirely analogous fashion, functions J and e in models (1.3) and (2.2) can be extended.

Well-posedness and mean-field limit of the regularised system. Following [6] (see (1.5)), we analyse (2.2) as an interaction system in the Banach space $Y = \mathbb{R}^d \times L^p_\eta(U)$, $1 \le p < \infty$. For $y = (x, \sigma) \in Y$ we set $\|y\|_Y = \|x\| + \|\sigma\|_{L^p_\eta(U)}$. Define

$$\mathcal{C}_{a,b} := \mathbb{R}^d \times \mathcal{S}_{a,b}, \tag{2.6}$$

and for $y, y' \in \mathcal{C}_{a,b}$ let us set

$$f : \mathcal{C}_{a,b} \times \mathcal{C}_{a,b} \to Y, \quad f(y, y') = \left[f^e(x, \sigma), \lambda f^{J,\varepsilon}(x, \sigma, x', \sigma') \right], \tag{2.7}$$

where

$$f^e(x, \sigma) = \int_U e(x, u)\, \sigma(u)\, d\eta(u) \quad \text{and} \quad f^{J,\varepsilon}(x, \sigma, x', \sigma') = f^J(x, \sigma, x', \sigma') + f^\varepsilon(\sigma), \tag{2.8}$$

so that (2.2) takes the equivalent form

$$\partial_t y_i(t) = \frac{1}{N} \sum_{j=1}^N f(y_i(t), y_j(t)) \quad \text{for } i = 1, \ldots, N. \tag{2.9}$$

Similar to [6], well-posedness of (2.9) relies on the Lipschitz continuity of f and on the following compatibility condition (for the corresponding proofs see Appendix A, Lemmas A.1, A.2 and A.4).

Lemma 2.5 (Compatibility condition). *For $J \in \mathcal{X}$ and $\varepsilon > 0$, let f^J and f^ε be defined as in (2.3) and (2.4). Then, there exist a,b with $0 < a < 1 < b < \infty$ such that for any $(x, \sigma), (x', \sigma') \in \mathbb{R}^d \times \mathcal{S}_{a,b}$ there exists some $\theta > 0$ such that*

$$\sigma + \theta\lambda\left[f^J(x, \sigma, x', \sigma') + f^\varepsilon(\sigma) \right] \in \mathcal{S}_{a,b}. \tag{2.10}$$

Intuitively, for specific choices of the bounds a and b, (2.10) states that moving from $\sigma \in \mathcal{S}_{a,b}$ into the direction generated by f^J and f^ε, we will always remain within $\mathcal{S}_{a,b}$ for some finite time. Eventually, similar to (1.6), a mean-field limit description of (2.9) is formally given by

$$\partial_t \Sigma(t) + \mathrm{Div}\big(b(\Sigma(t)) \cdot \Sigma(t)\big) = 0 \quad \text{with} \quad b(\Sigma(t))(y) := \int_{\mathcal{C}_{a,b}} f(y, y')\, d\Sigma(t)(y'). \tag{2.11}$$

Like (1.6), this is a PDE whose domain is a Banach space and we refer to [6] for the technical details. We can then summarise the main result in the following theorem.

Theorem 2.6 (Well-posedness and mean-field limit of entropic model). *Let $J \in \mathcal{X}$, $\lambda, \varepsilon > 0$, $T < +\infty$. Let $0 < a < 1 < b < \infty$ in accordance with Lemma 2.5. Then:*

1. *Given $\bar{\mathbf{y}}^N = (\bar{y}_1^N, \ldots, \bar{y}_N^N) \in \mathcal{C}_{a,b}^N$ there exists a unique trajectory $\mathbf{y}^N = (y_1^N, \ldots, y_N^N) : [0, T] \to \mathcal{C}_{a,b}^N$ of class C^1 solving (2.9) with $\mathbf{y}^N(0) = \bar{\mathbf{y}}^N$. In particular, $\Sigma^N(t) := \frac{1}{N}\sum_{i=1}^N \delta_{y_i^N(t)}$ provides a solution of (2.11) for initial condition $\bar{\Sigma}^N := \frac{1}{N}\sum_{i=1}^N \delta_{\bar{y}_i^N}$.*

2. *Given* $\bar{\Sigma} \in \mathcal{M}_1(\mathcal{C}_{a,b})$, *there exists a unique* $\Sigma \in C([0,T]; (\mathcal{M}_1(\mathcal{C}_{a,b}), W_1))$ *satisfying in the weak sense the continuity equation (2.11) with initial condition* $\Sigma(0) = \bar{\Sigma}$.

3. *For initial conditions* $\bar{\Sigma}_1, \bar{\Sigma}_2 \in \mathcal{M}_1(\mathcal{C}_{a,b})$ *and the respective solutions* Σ_1 *and* Σ_2 *of (2.11) one has the stability estimate*

$$W_1(\Sigma_1(t), \Sigma_2(t)) \le \exp\left(2 Lip(f)(t-s)\right) \cdot W_1(\Sigma_1(s), \Sigma_2(s)) \tag{2.12}$$

for every $0 \le s \le t \le T$.

Proof. The theorem follows by invoking Theorem 4.1 from [6]. On a more technical level, Theorem 5.3 of [6] provides the uniqueness of the solution to (2.11) in a Eulerian sense. We now show that the respective requirements are met.

First, the set $\mathcal{C}_{a,b}$, (2.6), is a closed convex subset of Y with respect to $\|\cdot\|_Y$ for any $0 < a < b < +\infty$. Likewise, for any $0 < a < b < +\infty$ the map f driving the evolution (2.9) is Lipschitz continuous: indeed, one can prove that f^J and f^ε are Lipschitz continuous (see Lemmas A.1 and A.2 in Appendix A) and Lipschitz continuity of f^e follows from the fact that e is Lipschitz continuous and bounded. Furthermore, as a consequence of Lemma 2.5, one can choose a and b so that the extended compatibility condition

$$\forall R > 0 \, \exists \theta > 0 : \quad y, y' \in \mathcal{C}_{a,b} \cap B_R(0) \quad \Rightarrow \quad y + \theta f(y, y') \in \mathcal{C}_{a,b}$$

holds, cf. [6, Theorem B.1]. Therefore, we may invoke [6, Theorem 4.1]. Combining this with [6, Theorem 5.3], which is applicable since $L^p_\eta(U)$ is separable and thus so is Y, the result follows. $\qquad\square$

2.2 Undisclosed setting and fast reactions

In the model (2.2), the decision process of each agent potentially involves knowledge of the mixed strategies of the other agents. Often it is reasonable to assume that there is no such knowledge, which can be reflected in the model by assuming that the payoff $J(x, u, x', u')$ does not actually depend on u', the other agent's strategy. Note that $J(x, u, x', u')$ may still depend on x', the other agent's location. We call this the *undisclosed* setting.

Additionally, often it is plausible to assume that the (regularised) dynamic that governs the mixed strategies $(\sigma_i(t))_{i=1}^N$ of the agents runs at a much faster time scale compared to the physical motion of the spatial locations $(x_i(t))_{i=1}^N$. This corresponds to a large value of λ in (2.2). Therefore, for the undisclosed setting we study the *fast-reaction* limit $\lambda \to \infty$.

The main results of this section are Theorems 2.13 and 2.14 which establish that the undisclosed fast-reaction limit is in itself a well-defined model (with a consistent mean-field limit as $N \to \infty$) and that the undisclosed model converges to this limit as $\lambda \to \infty$.

Undisclosed setting. In the undisclosed setting, the general formulas (2.2) for the dynamics can be simplified as follows:

$$\partial_t x_i(t) = \int_U e(x_i(t), u)\, \sigma_i(t)(u)\, \mathrm{d}\eta(u), \tag{2.13a}$$

$$\partial_t \sigma_i(t) = \lambda \cdot \left[\frac{1}{N} \sum_{j=1}^N f^J\big(x_i(t), \sigma_i(t), x_j(t)\big) + f^\varepsilon\big(\sigma_i(t)\big) \right] \tag{2.13b}$$

where f^ε is as given in (2.4) and f^J simplifies to

$$f^J\big(x, \sigma, x'\big) := \left[J(x, \cdot, x') - \int_U J(x, v, x')\, \sigma(v)\, \mathrm{d}\eta(v) \right] \cdot \sigma. \tag{2.14}$$

For finite $\lambda < +\infty$, the dynamics of this model are still covered by Theorem 2.6.

Fast reactions of the agents. Intuitively, as $\lambda \to \infty$ in (2.13), at any given time t the mixed strategies $(\sigma_i(t))_{i=1}^N$ will be in the unique steady state of the dynamics (2.13b) for fixed spatial locations $(x_i(t))_{i=1}^N$. For given locations $\mathbf{x} = (x_1, \dots, x_N) \in [\mathbb{R}^d]^N$ this steady state is given by

$$\sigma_i^J(\mathbf{x}) \equiv \sigma_i^J(x_1, \dots, x_N) := \frac{\exp\left(\dfrac{1}{\varepsilon N} \sum_{j=1}^N J\big(x_i, \cdot, x_j\big) \right)}{\int_U \exp\left(\dfrac{1}{\varepsilon N} \sum_{j=1}^N J\big(x_i, v, x_j\big) \right) \mathrm{d}\eta(v)}. \tag{2.15a}$$

(This computation is explicitly shown in the proof of Theorem 2.14.) The spatial agent velocities associated with this steady state are given by

$$v_i^J(\mathbf{x}) \equiv v_i^J(x_1, \dots, x_N) := \int_U e(x_i, u)\, \sigma_i^J(x_1, \dots, x_N)(u)\, \mathrm{d}\eta(u) \tag{2.15b}$$

and system (2.13) turns into a purely spatial ODE in the form

$$\partial_t x_i(t) = v_i^J(x_1(t), \dots, x_N(t)) \quad \text{for } i = 1, \dots, N. \tag{2.15c}$$

Unlike in Newtonian models over $\mathbb{R}^d$, (1.1), here the driving velocity field $v_i^J(x_1(t), \dots, x_N(t))$ is not a linear superposition of the contributions by each $x_j(t)$. This nonlinearity is a consequence of the fast-reaction limit and allows for additional descriptive power of the model that cannot be captured by the Newton-type model (1.1). This is illustrated in the two subsequent Examples 2.7 and 2.8.

Example 2.7 (Describing Newtonian models as undisclosed fast-reaction models). Newtonian models can be approximated by undisclosed fast-reaction entropic game models. We give a sketch for this approximation procedure. For a model as in (1.1), choose $U \subset \mathbb{R}^d$ such that it contains the range of f (for simplicity we assume that it is compact), let $e(x, u) := u$ and then set

$$J\big(x, u, x'\big) := -\|u - f\big(x, x'\big)\|^2.$$

Accordingly, the stationary mixed strategy of agent i is given by

$$\sigma_i^J(\mathbf{x})(u) = \mathcal{N}\left(\exp\left(-\frac{1}{\varepsilon N} \sum_{j=1}^N \|u - f\big(x_i, x_j\big)\|^2 \right) \right),$$

where $\mathcal{N}(\cdot)$ denotes the normalisation operator for a density with respect to η. Observe now that

$$\frac{1}{N} \sum_{j=1}^N \|u - f\big(x_i, x_j\big)\|^2 = \|u\|^2 - \frac{2}{N} \sum_{j=1}^N u \cdot f\big(x_i, x_j\big) + \frac{1}{N} \sum_{j=1}^N \|f\big(x_i, x_j\big)\|^2$$
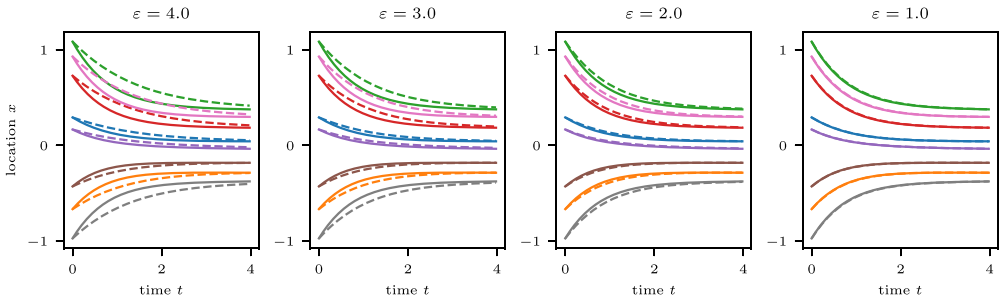
FIGURE 2. Approximation of a Newtonian model by an undisclosed fast-reaction entropic game model. Solid lines: original model, dashed lines: approximation. The Newtonian model is driven by (2.16), the approximation procedure is described in Example 2.7. The approximation becomes more accurate as ε decreases.

while

$$\left\| u - \frac{1}{N} \sum_{j=1}^{N} f\left(x_i, x_j\right) \right\|^2 = \|u\|^2 - \frac{2}{N} \sum_{j=1}^{N} u \cdot f\left(x_i, x_j\right) + \frac{1}{N^2} \left\| \sum_{j=1}^{N} f\left(x_i, x_j\right) \right\|^2.$$

Hence, since the terms not depending on u are cancelled by the normalisation, we also have

$$\sigma_i^J(\mathbf{x})(u) = \mathcal{N}\left(\exp\left(-\frac{1}{\varepsilon} \left\| u - \frac{1}{N} \sum_{j=1}^{N} f\left(x_i, x_j\right) \right\|^2 \right) \right).$$

So the maximum of the density $\sigma_i^J(\mathbf{x})$ is at $\frac{1}{N}\sum_{j=1}^{N} f(x_i, x_j)$, which is the velocity imposed by the Newtonian model. For small ε, $\sigma_i^J(\mathbf{x})$ will be increasingly concentrated around this point, and hence, the game model will impose a similar dynamic. Numerically, this is demonstrated in Figure 2 for a 1-dimensional Newtonian model as in (1.1) driven by

$$f(x, x') = -x - \frac{\tanh\left(5\left(x' - x \right) \right)}{\left(1 + \|x' - x\| \right)^2}. \tag{2.16}$$

Example 2.8 (Undisclosed fast-reaction models are strictly richer than Newtonian models). In the previous example, each agent j tried to persuade agent i to move with velocity $f(x_i, x_j)$. Deviations from this velocity were penalised in the payoff function with a quadratic function. By minimising the sum of these functions, the compromise of the linear average of all $f(x_i, x_j)$ then yields the best payoff. By picking different penalties we may obtain other interesting behaviour that cannot be described by Newtonian systems.

As above, let U be a (sufficiently large) compact subset of $\mathbb{R}^d$, $e(x, u) := u$. Let $g : \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz non-negative 'bump function' with compact support, that is, g is maximal at 0 with $g(0) = 1$, and $g(u) = 0$ for $\|u\| \geq \delta$ for some small $\delta > 0$. Then we set

$$J\left(x, u, x'\right) := \left(C - \|x - x'\| \right) \cdot g\left(u - \frac{x' - x}{\|x' - x\|} \right) - C \cdot g(u),$$

where $C > 0$ is a suitable, sufficiently large constant. Then approximately, the function $u \mapsto \frac{1}{N} \sum_j J(x_i, u, x_j)$ is maximal at $u = \frac{x_j - x_i}{\|x_j - x_i\|}$ where j is the agent that is closest to i (if it is closer than C and not closer than δ, in particular the last term in J is added to avoid that $u = 0$ is the maximiser). Therefore, approximately, we have created a model, where agents are attracted with unit speed by their nearest neighbours, but not by other agents. Such a nonlinear trade-off between the influences of the other agents is not obtainable by a Newtonian model.

Remark 2.9 (Fast-reaction limit in the fully general case). *For the fully general setting (i.e. J being also a function of u'), the study of the fast-reaction regime is a much harder problem as the stationary state of mixed strategies (for fixed spatial locations) often depends on the initial mixed strategies $(\sigma_i)_{i=1}^N$, and thus, the fast-reaction limit is a 'genuine' quasi-static evolution problem. We refer to [2, 66, 67] for derivation and well-posedness results of quasi-static evolutions of critical points of nonconvex energies on the Euclidean space by means of vanishing-viscosity methods. A similar analysis is in the course of development for nonconvex energies on Hilbert spaces [1]. Data-driven evolutions of critical points have been considered in [4].*

Well-posedness and mean-field limit of the undisclosed fast-reaction system. Similar to the full model (2.2), we are also interested in a mean-field limit of the undisclosed fast-reaction limit as the number of agents tends to infinity, $N \to \infty$. A formal limiting procedure leads to the equation

$$\partial_t \mu(t) + \text{div} \left(v^J(\mu(t)) \cdot \mu(t) \right) = 0 \tag{2.17}$$

where for $x \in \mathbb{R}^d$ and $\nu \in \mathcal{M}_1(\mathbb{R}^d)$ we set

$$v^J(\nu)(x) := \int_U e(x, u) \sigma^J(\nu)(x, u) \, d\eta(u) \tag{2.18}$$

$$\sigma^J(\nu)(x, \cdot) := \frac{\exp \left(\frac{1}{\varepsilon} \int_{\mathbb{R}^d} J(x, \cdot, x') \, d\nu(x') \right)}{\int_U \exp \left(\frac{1}{\varepsilon} \int_{\mathbb{R}^d} J(x, v, x') \, d\nu(x') \right) d\eta(v)}. \tag{2.19}$$

Given $\mathbf{x} = (x_1, \dots, x_N) \in [\mathbb{R}^d]^N$ and setting $\mu^N = \frac{1}{N} \sum_{j=1}^N \delta_{x_j}$, (2.15), (2.18) and (2.19) are related through

$$\sigma_i^J(\mathbf{x}) = \sigma_i^J(x_1, \dots, x_N) = \sigma^J(\mu^N)(x_i, \cdot) \quad \text{and} \quad v_i^J(\mathbf{x}) = v_i^J(x_1, \dots, x_N) = v^J(\mu^N)(x_i).$$

The key ingredient for the study of (2.15) and its limiting behaviour is to establish Lipschitz continuity of the fast-reaction equilibrium strategies and velocities with respect to payoff function and particle locations.

Lemma 2.10. *Let $J, J' \in X$ and consider $M > 0$ such that $M \geq \|J\|_\infty + \text{Lip}(J)$ and $M \geq \|J'\|_\infty + \text{Lip}(J')$. Let $\mu, \mu' \in \mathcal{M}_1(\mathbb{R}^d)$ and $x, x' \in \mathbb{R}^d$. Then:*

1. There exists $C = C(M, \varepsilon)$ such that, for every $u \in U$,

$$\left| \sigma^J(\mu)(x, u) - \sigma^{J'}(\mu')(x', u) \right| \leq C \left(W_1(\mu, \mu') + \|J - J'\|_\infty + \|x - x'\| \right) \tag{2.20}$$

and $1/C < \sigma^J(\mu)(x, u) < C$.

2. *There exists $C = C(M, e, \varepsilon)$ such that*

$$\left\| v^J(\mu)(x) - v^{J'}(\mu')(x') \right\| \le C\left(W_1(\mu, \mu') + \|J - J'\|_\infty + \|x - x'\| \right). \qquad (2.21)$$

Proof. Let us define two continuous functions $g, g' : U \to \mathbb{R}$ as

$$g(u) = \int_{\mathbb{R}^d} J(x, u, y) \, d\mu(y) \quad \text{and} \quad g'(u) = \int_{\mathbb{R}^d} J'(x', u, y) \, d\mu'(y)$$

For every $u \in U$, using $M \ge \|J\|_\infty$ and $M \ge \|J'\|_\infty$, we immediately obtain the global bounds $-M \le g(x) \le M$ and $-M \le g'(x) \le M$. Using the triangle inequality and the dual definition for W_1 in (1.9), we also estimate

$$|g(u) - g'(u)| \le C\left(W_1(\mu, \mu') + \|J - J'\|_\infty + \|x - x'\| \right) \qquad (2.22)$$

for $C = C(M)$. Recall now that

$$\sigma^J(\mu)(x, u) = \frac{\exp(g(u)/\varepsilon)}{\int_U \exp\left(g(v)/\varepsilon \right) \, d\eta(v)} \quad \text{and} \quad \sigma^{J'}(\mu')(x', u) = \frac{\exp(g'(u)/\varepsilon)}{\int_U \exp\left(g'(v)/\varepsilon \right) \, d\eta(v)}.$$

Thus, (2.20) follows combining global boundedness of g, g' with (2.22), the uniform bounds $1/C < \sigma^J(\mu)(x, u) < C$ are obtained in the same way, while (2.21) follows from (2.20) combined with $e \in \mathrm{Lip}_b(\mathbb{R}^d \times U; \mathbb{R}^d)$. $\qquad \qquad \square$

For the discrete fast-reaction system, the following Lemma adapts the estimate (2.21) as one in terms of discrete particle locations and their velocities.

Lemma 2.11 (Map to undisclosed fast-reaction agent velocity is Lipschitz continuous). *For any $N > 0$, consider the map*

$$\mathbf{v}^J : \left[\mathbb{R}^d \right]^N \to \left[\mathbb{R}^d \right]^N, \quad (x_1, \dots, x_N) \mapsto \left(v_1^J(x_1, \dots, x_N), \dots, v_N^J(x_1, \dots, x_N) \right)$$

associated with the fast-reaction ODE (2.15). Then, $\mathbf{v}^J$ is Lipschitz continuous under the distance induced by $\| \cdot \|_N$, with Lipschitz constant $L = L(J, e, \varepsilon)$.

Proof. Fix any $\mathbf{x}, \mathbf{x}' \in \left[\mathbb{R}^d \right]^N$ and define $\mu = \frac{1}{N} \sum_{j=1}^N \delta_{x_i}$ and $\mu' = \frac{1}{N} \sum_{j=1}^N \delta_{x_i'}$. Then, by Lemma 2.10, with $J' = J$, $x = x_i$ and $x' = x_i'$, we have

$$\left\| v^J(\mu)(x_i) - v^J(\mu')\left(x_i'\right) \right\| \le C\left(W_1(\mu, \mu') + \|x_i - x_i'\| \right)$$

for every $i = 1, \dots, N$, with $C = C(J, e, \varepsilon)$. Using the definition of W_1 in (1.8), we observe

$$W_1(\mu, \mu') \le \frac{1}{N} \sum_{j=1}^N \|x_i - x_i'\| = \|x - x'\|_N$$

so that

$$\|\mathbf{v}^J(\mathbf{x}) - \mathbf{v}^J(\mathbf{x}')\|_N = \frac{1}{N} \sum_{i=1}^{N} \|v_i^J(x_1, \dots, x_N) - v_i^J(x'_1, \dots, x'_N)\|$$

$$= \frac{1}{N} \sum_{i=1}^{N} \|v^J(\mu)(x_i) - v^J(\mu')(x'_i)\| \le \frac{C}{N} \sum_{i=1}^{N} (\|x - x'\|_N + \|x_i - x'_i\|) = 2C\|\mathbf{x} - \mathbf{x}'\|_N$$

which is the sought-after estimate. □

After clarifying what we mean by a solution of (2.17), we summarise the mean-field result in Theorem 2.13, whose proof then builds on Lipschitz continuity of the velocity field in the discrete system and follows by fairly standard arguments (see, e.g. [12, 14]).

Definition 2.12. We say a curve $\mu \in C([0, T]; (\mathcal{M}_1(\mathbb{R}^d), W_1))$ solves (2.17) if $\mu(t)$ has uniformly compact support for $t \in [0, T]$ and

$$\frac{d}{dt} \int_{\mathbb{R}^d} \phi(x) \mathrm{d}\mu(t)(x) = \int_{\mathbb{R}^d} \nabla\phi(x) \cdot v^J(\mu(t))(x) \, \mathrm{d}\mu(t)(x) \quad \text{for every } \phi \in C_c^\infty(\mathbb{R}^d).$$

Theorem 2.13 (Well-posedness and mean-field limit of undisclosed fast-reaction model). *Let $J \in X$, $\varepsilon > 0$, $0 < \bar{R} < +\infty$, and $0 < T < +\infty$. Define $R = \bar{R} + T \cdot \|e\|_\infty$. Then:*

1. *Given $\bar{\mathbf{x}}^N = (\bar{x}_1^N, \dots, \bar{x}_N^N) \in [B^d(\bar{R})]^N$ there exists a unique trajectory $\mathbf{x}^N = (x_1^N, \dots, x_N^N) : [0, T] \to [B^d(R)]^N$ of class C^1 solving (2.15c) with $\mathbf{x}^N(0) = \bar{\mathbf{x}}^N$. In particular, $\mu^N(t) := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i^N(t)}$ provides a solution of (2.17) for initial condition $\bar{\mu}^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{\bar{x}_i^N}$.*

2. *Given $\bar{\mu} \in \mathcal{M}_1(B^d(\bar{R}))$ there exists a unique $\mu \in C([0, T]; (\mathcal{M}_1(B^d(R)), W_1))$ satisfying in the weak sense the continuity equation (2.17) with initial condition $\mu(0) = \bar{\mu}$.*

3. *For initial conditions $\bar{\mu}_1, \bar{\mu}_2 \in \mathcal{M}_1(B^d(\bar{R}))$ and the respective solutions μ_1 and μ_2 of (2.17) one has the stability estimate*

$$W_1(\mu_1(t), \mu_2(t)) \le \exp(C(t - s)) \cdot W_1(\mu_1(s), \mu_2(s)) \tag{2.23}$$

for every $0 \le s \le t \le T$, with $C = C(J, e, \varepsilon, \bar{R}, T)$.

Proof outline. The proof is rather standard and builds on the Lipschitz continuity of the finite agents system (2.15). We highlight the main steps without in-depth details and refer to Proposition 2.1 and Theorem 2.4 of [10], and references therein, for further information.

Part 1: finite agents setting. For given $\bar{\mathbf{x}}^N = (\bar{x}_1^N, \dots, \bar{x}_N^N) \in [B^d(\bar{R})]^N$, using the Lipschitz continuity provided in Lemma 2.11, there exists a unique curve $\mathbf{x}^N : [0, T] \to [\mathbb{R}^d]^N$ of class C^1 solving (2.15) with $\mathbf{x}^N(0) = \bar{\mathbf{x}}^N$. A direct estimate provides $\|\partial x_i^N(t)\| \le \|e\|_\infty$, so that $\|x_i^N(t)\| \le \bar{R} + T \cdot \|e\|_\infty$ for every $t \in [0, T]$, $i = 1, \dots, N$. Furthermore, setting $\mu^N(t) := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i^N(t)}$ one can explicitly verify that μ^N solves (2.17) with initial condition $\bar{\mu}^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{\bar{x}_i^N}$. This establishes point 1.
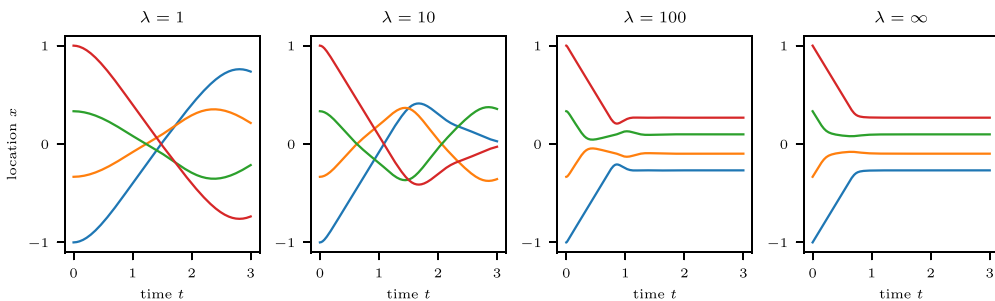
FIGURE 3. Convergence of the undisclosed model to the fast-reaction limit as $\lambda \to \infty$. The payoff function encourages agents to move to the origin, but penalises small pairwise distances (see Example 4.1 for a detailed description). With λ small, agents cannot adjust their strategies fast enough. They overshoot the origin and completely fail to avoid each other. The situation improves as λ increases. For $\lambda = 100$, the model closely resembles the fast-reaction limit, in accordance with (2.24).

Part 2: mean-field solution as limit of finite agents solutions. Fix $\bar{\mu} \in \mathcal{M}_1(B^d(\bar{R}))$. For $N > 0$, let $(\bar{\mu}^N)_N \subset \mathcal{M}_1(B^d(\bar{R}))$ be a sequence of empirical measures such that $W_1(\bar{\mu}^N, \bar{\mu}) \to 0$ as $N \to \infty$ and let $(\mu^N)_N$ be the respective sequence of solutions of (2.17) with initial conditions $\bar{\mu}^N$. This sequence of curves $(\mu^N)_{N \in \mathbb{N}} \subset C([0, T]; (\mathcal{M}_1(B^d(R)), W_1))$ is equicontinuous and equibounded. Therefore, an application of the Ascoli–Arzelà theorem provides a cluster point $\mu \in C([0, T]; (\mathcal{M}_1(B^d(R)), W_1))$ such that, up to a subsequence, we have

$$\lim_{N \to \infty} W_1\big(\mu^N(t), \mu(t)\big) = 0 \quad \text{uniformly for } t \in [0, T].$$

Invoking Lemma 2.10 for $x = x'$ and $J = J'$ this implies $\|v^J\big(\mu^N(t)\big)(x) - v^J(\mu(t))(x)\| \to 0$ uniformly in $t \in [0, T], x \in B^d(R)$ as $N \to \infty$. Consequently, the cluster point μ is a solution to (2.17) with $\bar{\mu}$ as initial condition. This establishes the existence part of point 2.

Part 3: stability estimates. For fixed N, a stability estimate of the form $\|\mathbf{x}_1^N(t) - \mathbf{x}_2^N(t)\|_N \leq \exp(C(t-s)) \cdot \|\mathbf{x}_1^N(s) - \mathbf{x}_2^N(s)\|_N$ for solutions to the discrete system (2.15) follows quickly from Grönwall's lemma. Since the $\|\cdot\|_N$-norm between two point clouds provides an upper bound for the W_1 distance between the respective empirical measures, this would provide point 3 for empirical measures. The extension to arbitrary measures can be done as in the proof of Theorem 2.4 in [10]. This then provides uniqueness of the solution μ to (2.17) for initial condition $\bar{\mu} \in \mathcal{M}_1(B^d(\bar{R}))$, completing point 2.

Finally, we establish convergence of (2.13) to (2.15) (and their respective mean-field versions) as $\lambda \to \infty$. The proof is given in Section A.2 of the Appendix, a simple numerical example is shown in Figure 3.

Theorem 2.14 (Convergence to fast-reaction limit in undisclosed setting as $\lambda \to \infty$).

1. ***Discrete setting:*** *For initial positions* $\mathbf{x}(0) = (x_1(0), \ldots, x_N(0)) \in [B^d(\bar{R})]^N$ *and initial mixed strategies* $\boldsymbol{\sigma}(0) = (\sigma_1(0), \ldots, \sigma_N(0)) \in \mathcal{S}_{a,b}^N$ *let* $\mathbf{x}(t) = (x_1(t), \ldots, x_N(t))$ *and* $\boldsymbol{\sigma}(t) = (\sigma_1(t), \ldots, \sigma_N(t))$ *be the solution to the undisclosed model (2.13). For the same initial positions, let* $\mathbf{x}^{**}(t) = (x_1^{**}(t), \ldots, x_N^{**}(t))$ *be the solution to the undisclosed fast-reaction model (2.15) and let* $\boldsymbol{\sigma}^{**}(t) = (\sigma_1^{**}(t), \ldots, \sigma_N^{**}(t))$ *with* $\sigma_i^{**}(t) = \sigma_i^J(\mathbf{x}^{**}(t))$ *be the corresponding fast-reaction strategies.*

   Then, there exists $C = C(a, b, \bar{R}, J, e)$ *(independent of N and i) such that for all* $t \in [0, \infty)$,

$$\|\mathbf{x}(t) - \mathbf{x}^{**}(t)\|_N \leq \frac{C}{\sqrt{\lambda}} \cdot \exp(t \cdot C) \tag{2.24}$$

and

$$\|\sigma_i(t) - \sigma_i^{**}(t)\|_{L^2_\eta(U)} \leq \frac{C}{\sqrt{\lambda}} \cdot \exp(t \cdot C) + \left[C \left[\frac{1}{\lambda} + \exp\left(-\frac{\lambda\,t}{C} \right) \right] \right]^{1/2}. \tag{2.25}$$

So on any compact time horizon [0, T], we find that $\|\mathbf{x}(t) - \mathbf{x}^{**}(t)\|_N \to 0$ *uniformly in time as* $\lambda \to \infty$. $\boldsymbol{\sigma}$ *converges to* $\boldsymbol{\sigma}^{**}$ *uniformly in time on the interval* $[\tau, T]$ *for any* $\tau > 0$. *Near* $t = 0$ *we cannot expect uniform convergence of* $\boldsymbol{\sigma}$, *since by initialisation it can start at a finite distance from* $\boldsymbol{\sigma}^{**}(0)$, *and thus, it takes a brief moment to relax to the vicinity of the fast-reaction state.*

2. **Mean-field setting:** *For an initial configuration* $\bar{\Sigma} \in \mathcal{M}_1(\mathcal{C}_{a,b})$, *let* Σ *be the solution to the entropic mean-field model (2.11) for a undisclosed* J *with* $\lambda \in (0, \infty)$. *Let* $P : \mathcal{C}_{a,b} \to \mathbb{R}^d$, $(x, \sigma) \mapsto x$ *be the projection from* $\mathcal{C}_{a,b}$ *to the spatial component. Set* $\bar{\mu} := P_\sharp \bar{\Sigma}$ *and let* μ *be the solution to the fast-reaction mean-field model (2.17) with initial condition* $\mu(0) = \bar{\mu}$.

   Then for the same C *as in (2.24), one has*

$$W_1(\mu(t), P_\sharp \Sigma(t)) \leq \frac{C}{\sqrt{\lambda}} \cdot \exp(t \cdot C). \tag{2.26}$$

Remark 2.15. *The statement about the mean-field equations can be expanded further: From the fast-reaction solution* μ *on* $\mathcal{M}_1(\mathbb{R}^d)$, *one can construct a 'lifted trajectory'* $\hat{\Sigma}$ *on* $\mathcal{M}_1(\mathcal{C}_{a,b})$, *intuitively by attaching to each mass particle in* μ *at* x *its corresponding fast-reaction mixed strategy* $\sigma^J(\mu)(x, \cdot)$, (2.19). *Using the bounds (2.24) and (2.25), the continuity properties of* $\sigma^J(\mu)(x, \cdot)$ *(see Lemma 2.10) and arguing as in the proof of (2.26) one can then establish a* W_1-bound *between* Σ *and* $\hat{\Sigma}$.

3 Inference for entropic evolutionary games

After having introduced a new class of models for interacting agents, we now turn to the question of how the payoff function J, that parametrises a model, can be inferred from data. Some remarks are in order before we proceed.

3.1 Discussion

Our motivation is as follows: we observe a set of rational agents, for example, pedestrians in a confined space, and want to learn about their interaction rules. As argued above (e.g. in Examples 2.7 and 2.8), undisclosed fast-reaction entropic games can be used to parametrise a rich class of interaction behaviours, in particular subsuming Newtonian models. We may therefore hope that, when feeding our observations into an inference functional for J, that by analysing and interpreting the resulting payoff function we can learn something about the interaction rules of the agents. We focus on the inference of J and assume that U, e and ε are known. We will demonstrate during the numerical examples (Section 4) that it is usually possible to make plausible choices for U, e and that qualitatively equivalent choices yield qualitatively equivalent results for the inferred J. One may also assume that $\varepsilon = 1$, since re-scaling ε can be compensated by re-scaling J in the same way, see (2.15).

In Section 3.2, we discuss a differential inference functional, proposed in analogy to (1.7), for the full entropic game model, that is, without the undisclosed assumption and not in the fast-reaction limit. For this we assume that we are able to observe the full dynamics, that is, physical locations and velocities of the agents, as well as their mixed strategies and temporal derivatives, $(x_i, \partial_t x_i, \sigma_i, \partial_t \sigma_i)_{i=1}^N$ and propose a functional that seeks to infer J by comparing its predictions with the observed $\partial_t \sigma_i$.

In many cases, it may not be possible to observe the mixed strategies σ_i, let alone their temporal derivatives $\partial_t \sigma_i$, since these may correspond to internal states of the agents that we, as external observers, cannot perceive. In Section 3.3, we turn to the undisclosed fast-reaction setting, where physical velocities and mixed strategies are functions of current physical locations, and show that there one can still perform inference with limited observations. In Section 3.3.1, we propose an inference functional, again in analogy to (1.7), that only requires the observation of the physical locations x_i and velocities $\partial_t x_i$. Although it is nonconvex, we observe in our numerical examples that it provides plausible payoff functions J, indicating that nonconvexity does not seem to be a practical issue.

In Section 3.3.2, we propose a convex inference functional that requires knowledge of the mixed strategies σ_i, but not of their temporal derivatives. This may be practical in two cases: In Section 4.1, we propose a scheme to provide such data from observed physical velocities, intuitively by 'inverting' the approximation scheme of Example 2.7. Also, mixed strategies could represent a stochastic movement of agents and the 'smooth' physical velocity $\partial_t x_i$ is merely the result of an averaging over a smaller time scale. Should we indeed be able to observe the fluctuating agent dynamics at small time scales, this could be used to approximate the mixed strategies. This was one of our motivations for providing Example 4.3.

For both functionals, there are natural candidates for the mean-field limit. We analyse these two functionals in more detail in Section 3.3.3, providing an estimate of the approximation quality (Theorem 3.3), as well as existence of minimisers and consistency in the mean-field limit (Theorem 3.6).

Now, we formalise our notion of admissible observations for inference and a notion of consistency of observations in the mean-field limit.

Assumption 3.1 (Admissible observations).

1. *Discrete observation: For a fixed number of agents N, a time horizon $T \in (0, \infty)$, and some radius $R \in (0, \infty)$ we observe the agents' physical paths $\mathbf{x}^N := (x_1^N, \ldots, x_N^N) \in C^1([0, T], B^d(R))^N$ with velocities $\mathbf{v}^N := (v_1^N, \ldots, v_N^N)$, $v_i^N = \partial_t x_i^N$.*

   Optionally, in addition we may also observe the agents' mixed strategies $\boldsymbol{\sigma}^N := (\sigma_1^N, \ldots, \sigma_N^N) \in C([0, T], \mathcal{S}_{a,b})^N$ for some bounds $0 < a \le b < \infty$ in the definition of $\mathcal{S}_{a,b}$, see (2.1). The mixed strategies are consistent with the observed velocities, that is,

$$v_i^N(t) = \int_U e\left(x_i^N(t), u\right) \sigma_i^N(t)(u) \, d\eta(u). \tag{3.1}$$

   Note: The assumption that the agent's velocity is exactly consistent with their mixed strategy may seem unrealistic, as the velocity may be subject to other external influences and noise. However, we will subsequently often assume that the mixed strategies are not observed

directly, but are only inferred from the velocities. In this case, satisfying assumption (3.1) is quite natural. In other cases, assumption (3.1) will be used in Theorem 3.3 to bound the error on trajectories based on the error on strategies. If (3.1) only holds approximately, an additional corresponding error term will appear there.

2. *Consistent mean-field behaviour:* For fixed T, R, for an increasing sequence of N we make observations as specified in part 1, and there is a limit observation $t \mapsto \mu^\infty(t) \in \mathcal{M}_1(B^d(R))$ with a velocity field $t \mapsto v^\infty(t)$ such that

$$W_1\big(\mu^N(t), \mu^\infty(t)\big) \to 0 \quad \text{uniformly in } t \in [0, T] \text{ where} \quad \mu^N(t) := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i^N(t)}, \quad (3.2)$$

$$\int_0^T \frac{1}{N} \sum_{i=1}^{N} \big\langle \phi\left(t, x_i^N(t)\right), v_i^N(t) \big\rangle \, \mathrm{d}t \to \int_0^T \int_{\mathbb{R}^d} \big\langle \phi(t, x), v^\infty(t)(x) \big\rangle \, \mathrm{d}\mu^\infty(t)(x) \, \mathrm{d}t \quad (3.3)$$

for all $\phi \in \big[C\big([0, T] \times \mathbb{R}^d\big)\big]^d$, and

$$\int_0^T \frac{1}{N} \sum_{i=1}^{N} \|v_i^N(t)\|^2 \, \mathrm{d}t \to \int_0^T \int_{\mathbb{R}^d} \|v^\infty(t)(x)\|^2 \, \mathrm{d}\mu^\infty(t)(x) \, \mathrm{d}t < \infty. \quad (3.4)$$

(3.2) implies that physical locations are consistent, (3.3) implies that observed velocities converge in a weak sense, and (3.4) implies that they are consistent. Intuitively, for (3.4) to hold, mass particles that converge to the same limit location x as $N \to \infty$ need to converge to the same velocity, otherwise Jensen's strict inequality contradicts (3.4).

If we also observe mixed strategies in part 1, then the bounds $0 < a \le b < \infty$ are uniform in N and there also is a mixed strategy mean-field $(t, x) \mapsto \sigma^\infty(t)(x) \in \mathcal{S}_{a,b}$ such that

$$\int_0^T \frac{1}{N} \sum_{i=1}^{N} \int_U \phi\left(t, x_i^N(t), u\right) \cdot \sigma_i^N(t)(u) \, \mathrm{d}\eta(u) \, \mathrm{d}t$$

$$\to \int_0^T \int_{\mathbb{R}^d} \int_U \phi(t, x, u) \cdot \sigma^\infty(t)(x)(u) \mathrm{d}\eta(u) \, \mathrm{d}\mu^\infty(t)(x) \, \mathrm{d}t \quad (3.5)$$

for all $\phi \in \big[C\big([0, T] \times \mathbb{R}^d \times U\big)\big]^d$ and

$$\int_0^T \frac{1}{N} \sum_{i=1}^{N} \int_U \sigma_i^N(t)(u) \log\big(\sigma_i^N(t)(u)\big) \, \mathrm{d}\eta(u) \, \mathrm{d}t$$

$$\to \int_0^T \int_{\mathbb{R}^d} \int_U \sigma^\infty(t)(x)(u) \log(\sigma^\infty(t)(x)(u)) \mathrm{d}\eta(u) \, \mathrm{d}\mu^\infty(t)(x) \, \mathrm{d}t. \quad (3.6)$$

These are direct equivalents of (3.3) and (3.4).

In particular, observations are admissible when they were generated by an entropic game model with some ground truth payoff J.

Lemma 3.2. *Let $J \in X$, $T > 0$, $\bar{R} > 0$, $N \in \mathbb{N}$.*

1. *For initial locations $\bar{\boldsymbol{x}}^N = (\bar{x}_1^N, \ldots, \bar{x}_N^N) \in [B^d(\bar{R})]^N$ the induced solution to (2.15) with corresponding velocities and mixed strategies provides an admissible discrete observation in the sense of Assumption 3.1, part 1, for $R = \bar{R} + \|e\|_\infty \cdot T$.*

2. *Let $\bar{\mu} \in \mathcal{M}_1(B^d(\bar{R}))$ and consider a sequence of empirical measures $(\bar{\mu}^N)_N$ in $\mathcal{M}_1(B^d(\bar{R}))$ of the form*

$$\bar{\mu}^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\bar{x}_i^N}, \quad \bar{x}_i^N \in B^d(\bar{R}), \quad \text{such that} \quad \lim_{N \to \infty} W_1(\bar{\mu}^N, \bar{\mu}) = 0.$$

Then, the solutions to (2.15) with initial positions $(\bar{x}_1^N, \ldots, \bar{x}_N^N)$ are a suitable sequence of discrete observations in the sense of Assumption 3.1, part 2, and the solution to (2.17) provides a corresponding limit observation.

The proof is quite straightforward and provided in Section A.3.

In the following, we assume that observations are admissible in the sense of Assumption 3.1.

3.2 Differential inference functional

In this section, we discuss an inference functional for payoff function J, in close analogy to (1.7). For now, we assume that in addition to $(\boldsymbol{x}^N, \boldsymbol{v}^N, \boldsymbol{\sigma}^N)$ we can even observe $\partial_t \sigma_i^N$ for all $i \in \{1, \ldots, N\}$. Our differential inference functional is therefore aimed at recovering J, by comparing the observed $\partial_t \sigma_i^N$ with the predictions by the model (2.2b). In (1.7), the discrepancy between observed v_i^N and predicted velocities is penalised by the squared Euclidean distance. As metric to compare the observed $\partial_t \sigma_i^N$ and the prediction by (2.2b), we choose the (weak) Riemannian tensor of the Hellinger–Kakutani and Fisher–Rao metrics. That is, we set for a base point $\sigma \in \mathcal{S}_{a,b}$, and two tangent directions $\delta\mu, \delta\nu \in L_\eta^p(U)$ attached to it,

$$d_\sigma(\delta\mu, \delta\nu)^2 := \frac{1}{4} \int_U \frac{(\delta\mu - \delta\nu)^2}{\sigma} \, d\eta. \tag{3.7}$$

A potential inference functional for J could then be:

$$\mathcal{E}_\sigma^N(J) :=$$
$$\frac{1}{T} \int_0^T \left[\frac{1}{N} \sum_{i=1}^{N} d_{\sigma_i^N(t)} \left(\partial_t \sigma_i^N(t), \frac{1}{N} \sum_{j=1}^{N} f^J \left(x_i^N(t), \sigma_i^N(t), x_j^N(t), \sigma_j^N(t) \right) + f^\varepsilon \left(\sigma_i^N(t) \right) \right)^2 \right] dt. \tag{3.8}$$

Minimisation should be restricted to a sufficiently regular and compact class of J. Since $d_\sigma(\cdot, \cdot)^2$ is quadratic in its arguments and f^J is linear in J, the mismatch term is again quadratic in J and thus restricted to a suitable class, this is a convex optimisation problem.

In Section 4.3, we provide a simple example where we simulate trajectories $(\boldsymbol{x}^N, \boldsymbol{\sigma}^N)$ according to the dynamics (2.2), based on a given J, and then obtain a corresponding minimiser $\hat{J}$ of (3.8). We then show that trajectories simulated with J and $\hat{J}$ are close even when the initial configurations are not drawn from training data used to infer $\hat{J}$. Indeed, upon selecting a suitable

functional framework, one can in principle extend the analysis in [10] (and the analysis presented here below) to prove existence of minimisers, limiting behaviour as $N \to \infty$ and ability of the inferred model to generalise.

However, application to real data seems challenging, as one usually only observes (x^N, v^N) but not σ^N, let alone variations $\partial_t \sigma_i^N$. To address this, we propose in the next section two new inference functionals for the undisclosed fast-reaction setting.

3.3 Undisclosed fast-reaction inference functionals

3.3.1 *Penalty on velocities*

The inference situation is simplified in the undisclosed fast-reaction setting since now, for fixed payoff function J, mixed strategies and physical velocities are a direct function of physical locations. We can therefore attempt to minimise the discrepancy between observed physical velocities and those predicted by J.

For a discrete admissible observation (x^N, v^N), as in Assumption 3.1, we define the inference functional on hypothetical payoff functions $J \in X$

$$\mathcal{E}_v^N(J) := \frac{1}{T} \int_0^T \left[\frac{1}{N} \sum_{i=1}^N \left\| v_i^N(t) - v_i^J \left(x_1^N(t), \ldots, x_N^N(t) \right) \right\|^2 \right] \mathrm{d}t. \tag{3.9a}$$

The natural candidate for the limit functional is

$$\mathcal{E}_v(J) := \frac{1}{T} \int_0^T \int_{\mathbb{R}^d} \left\| v^\infty(t)(x) - v^J \left(\mu^\infty(t) \right)(x) \right\|^2 \mathrm{d}\mu^\infty(t)(x) \mathrm{d}t, \tag{3.9b}$$

where the couple (μ^∞, v^∞) is the corresponding limit of $(x^N, v^N)_N$ as introduced in Assumption 3.1, point 2.

This functional is intuitive and requires only the observations of (x^N, v^N) and no observations about the mixed strategies. However, it is not convex. This could potentially lead to poor local minima, but we did not observe such problems in our numerical examples.

The motivation for studying mean-field inference functionals is twofold. First, it establishes asymptotic consistency of inference in the limit of many agents. Second, the mean-field equation yields an approximation for the expected behaviour of the finite-agent system under many repetitions with random initial conditions. While we do not prove this approximation property, the existence of a consistent mean-field inference functional is still an encouraging indicator that inference over the collection of many finite-agent samples will yield a reasonable result.

3.3.2 *Penalty on mixed strategies*

In the case where information about the mixed strategies of the agents is available (see Section 3.1 for a discussion), the discrepancy between observed mixed strategies and those predicted by J can be used for inference. We choose to measure this discrepancy with the Kullback–Leibler (KL) divergence.

Similar to above, for a discrete admissible observation $(\mathbf{x}^N, \mathbf{v}^N, \boldsymbol{\sigma}^N)$, we define the inference functional for hypothetical payoff functions $J \in X$

$$\mathcal{E}_\sigma^N(J) := \frac{1}{T} \int_0^T \left[\frac{1}{N} \sum_{i=1}^N \mathrm{KL}\left(\sigma_i^N(t) | \sigma_i^J\left(x_1^N(t), \ldots, x_N^N(t)\right) \right) \right] \mathrm{d}t. \tag{3.10a}$$

The natural candidate for the limit functional is

$$\begin{aligned}
\mathcal{E}_\sigma(J) &:= \frac{1}{T} \int_0^T \int_{\mathbb{R}^d} \mathrm{KL}\left(\sigma^\infty(t)(x, \cdot) | \sigma^J(\mu^\infty(t))(x, \cdot) \right) \mathrm{d}\mu^\infty(t)(x) \, \mathrm{d}t \\
&= \frac{1}{T} \int_0^T \int_{\mathbb{R}^d} \int_U \log\left(\frac{\sigma^\infty(t)(x, u)}{\sigma^J(\mu^\infty(t))(x, u)} \right) \sigma^\infty(t)(x, u) \mathrm{d}\eta(u) \mathrm{d}\mu^\infty(t)(x) \, \mathrm{d}t, \tag{3.10b}
\end{aligned}$$

where $(\mu^\infty, \sigma^\infty)$ are as introduced in Assumption 3.1, point 2. Due to the particular structure of stationary mixed strategies, (2.15), and the KL divergence, functionals (3.10) are convex on X (Proposition A.5), which is a potential advantage over (3.9).

3.3.3 *Analysis of the inference functionals*

We now provide some theoretical results about the inference functionals of Sections 3.3.1 and 3.3.2. The first result establishes that the obtained minimal inference functional value provides an upper bound on the accuracy of the trajectories that are simulated with the inferred $\hat{J}$ (the proof builds upon the proof of [10, Proposition 1.1] and is given in Section A.3).

Theorem 3.3. *Let $\hat{J} \in X$ and $(\mathbf{x}^N, \mathbf{v}^N, \boldsymbol{\sigma}^N)$ be an admissible discrete observation for N agents (cf. Assumption 3.1, point 2). Let $\hat{\mathbf{x}}^N$ be the solution of (2.15) induced by $\hat{J}$ for the initial condition $\hat{\mathbf{x}}^N(0) = \mathbf{x}^N(0) = (\bar{x}_1^N, \ldots, \bar{x}_N^N) \in [\mathbb{R}^d]^N$. Then,*

$$\|\mathbf{x}^N(t) - \hat{\mathbf{x}}^N(t)\|_N \le C\sqrt{\mathcal{E}_v^N(\hat{J})} \quad \text{and} \quad \|\mathbf{x}^N(t) - \hat{\mathbf{x}}^N(t)\|_N \le C\sqrt{\mathcal{E}_\sigma^N(\hat{J})}$$

for all $t \in [0, T]$, with $C = C(T, \hat{J}, e, \varepsilon)$. Analogously, let $(\mu^\infty, \sigma^\infty)$ as introduced in Assumption 3.1, point 1 and let $\hat{\mu}$ be the solution of (2.17) induced by $\hat{J}$ for the initial condition $\bar{\mu} = \mu^\infty(0)$. Then,

$$W_1(\mu^\infty(t), \hat{\mu}(t)) \le C\sqrt{\mathcal{E}_v(\hat{J})} \quad \text{and} \quad W_1(\mu^\infty(t), \hat{\mu}(t)) \le C\sqrt{\mathcal{E}_\sigma(\hat{J})}$$

for all $t \in [0, T]$ and the same constant C as above.

Next, we address the existence of minimising payoff functions, both in theory and numerical approximation. At the theoretical level, we need to ensure compactness of minimising sequences, which we obtain here by restriction to a suitable compact space. At the numerical level, we are interested in finite-dimensional approximations of this space that asymptotically are dense as the discretisation is refined.

Remark 3.4 (Compactness and finite-dimensional approximation). *In order to obtain compactness, we restrict the variational problems to suitable subspaces of X. In particular, for $R, M > 0$, let us define*

$$X_{R,M} = \left\{ J \in \mathrm{Lip}_b\big(B^d(R) \times U \times B^d(R)\big) \mid \|J\|_\infty + \mathrm{Lip}(J) \leq M \right\}.$$

The parameter R bounds the learning domain in space: inference where we have no data available is simply meaningless. The parameter M will ensure compactness with respect to uniform convergence.

For each $R, M > 0$, we consider a family of closed convex subsets $\big(X_{R,M}^N\big)_{N \in \mathbb{N}} \subset X_{R,M}$ satisfying the uniform approximation property: for every $J \in X_{R,M}$ there exists a sequence $(J^N)_{N \in \mathbb{N}}$ uniformly converging to J with $J^N \in X_{R,M}^N$ for each $N \in \mathbb{N}$. These approximating spaces can be selected to be finite dimensional.

An alternative way to obtain compactness would be via the introduction of additional regularising functionals. This is discussed in Section 4.1. Conceptually, both approaches serve the same purpose and we find that the former is more convenient for analysis while the latter is more attractive numerically.

The following key lemma ensures convergence of the error functionals under uniform convergence of the payoff functions and/or W_1 convergence of measures.

Lemma 3.5. *Let $\big(\mathbf{x}^N, \mathbf{v}^N, \boldsymbol{\sigma}^N\big)_N$ be a family of admissible observations and $\mathcal{E}_v^N$, $\mathcal{E}_v$, $\mathcal{E}_\sigma^N$ and $\mathcal{E}_\sigma$ be the corresponding discrete and continuous inference functionals (3.9) and (3.10). For $M > 0$, let $\big(J^N\big)_N$ be a sequence of payoff functions in $X_{R,M}$, converging uniformly to some $J \in X_{R,M}$ as $N \to \infty$. Then,*

$$\lim_{N \to \infty} \mathcal{E}_v^N(J^N) = \mathcal{E}_v(J) \quad \text{and} \quad \lim_{N \to \infty} \mathcal{E}_\sigma^N(J^N) = \mathcal{E}_\sigma(J). \tag{3.11}$$

Proof. Thanks to Assumption 3.1, point 2, we have $W_1\big(\mu^N(t), \mu^\infty(t)\big) \to 0$ uniformly in $t \in [0, T]$ as $N \to \infty$. Lemma 2.10 provides

$$\left\| v^{J^N}\big(\mu^N(t)\big)(x) - v^J(\mu^\infty(t))(x) \right\| \leq C \cdot \left(W_1\big(\mu^N(t), \mu^\infty(t)\big) + \|\hat{J}^N - \hat{J}\| \right)$$

uniformly in $x \in B^d(R), t \in [0, T]$ and N. Hence, $v^{J^N}\big(\mu^N(t)\big)(x)$ converges uniformly to $v^J(\mu^\infty(t))(x)$ in $C([0, T] \times B^d(R); \mathbb{R}^d)$. Since $\|v^J(\mu^\infty(t))(x)\| \leq \|e\|_\infty$ for every $x \in B^d(R)$ and $t \in [0, T]$, we also have $\|v^{J^N}\big(\mu^N(t)\big)(x)\|^2 \to \|v^J(\mu^\infty(t))(x)\|^2$ uniformly in $C([0, T] \times B^d(R))$. Recalling that $v_i^{J^N}\big(x_1^N(t), \ldots, x_N^N(t)\big)(x) = v^{J^N}\big(\mu^N(t)\big)\big(x_i^N(t)\big)$ by definition and using the convergence of velocities provided by Assumption 3.1, point 2, the result follows by passing to the limit in the definition:

$$\lim_{N \to \infty} \mathcal{E}_v^N\big(J^N\big) = \lim_{N \to \infty} \frac{1}{T} \int_0^T \left[\frac{1}{N} \sum_{i=1}^N \left\| v_i^N(t) - v_i^{J^N}\big(x_1^N(t), \ldots, x_N^N(t)\big) \right\|^2 \right] \mathrm{d}t$$

$$= \frac{1}{T} \int_0^T \int_{\mathbb{R}^d} \left\| v^\infty\big(\mu^\infty(t)\big)(x) - v^J\big(\mu^\infty(t)\big)(x) \right\|^2 \mathrm{d}\mu^\infty(t)(x)\mathrm{d}t = \mathcal{E}_v(J).$$

An analogous derivation applies to the σ-energies (3.10) where we use the Lipschitz estimate (2.20) and the σ-part of Assumption 3.1, point 2. $\qquad\square$

We are now ready to state the main result concerning the inference functionals and convergence of minimisers for $\mathcal{E}_v^N$ and $\mathcal{E}_\sigma^N$.

Theorem 3.6. *Let* $\left(\mathbf{x}^N, \mathbf{v}^N, \boldsymbol{\sigma}^N\right)_N$ *be a family of admissible observations and fix* $M > 0$. *Then:*

1. *The functionals (3.9a) and (3.10a) have minimisers over* $X_{R,M}^N$ *and (3.9b) and (3.10b) have minimisers over* $X_{R,M}$.

2. *For* $N > 0$, *let*

$$J_v^N \in argmin_{J \in X_{R,M}^N} \mathcal{E}_v^N(J) \quad and \quad J_\sigma^N \in argmin_{J \in X_{R,M}^N} \mathcal{E}_\sigma^N(J).$$

   The sequence $\left(J_v^N\right)_N$, *resp.* $\left(J_\sigma^N\right)_N$, *has a subsequence converging uniformly to some continuous function* $J_v \in X_{R,M}$, *resp.* $J_\sigma \in X_{R,M}$, *such that*

$$\lim_{N \to \infty} \mathcal{E}_v^N\left(J_v^N\right) = \mathcal{E}_v(J_v) \quad and \quad \lim_{N \to \infty} \mathcal{E}_\sigma^N\left(J_\sigma^N\right) = \mathcal{E}_\sigma(J_\sigma). \tag{3.12}$$

   In particular, J_v *(resp.* J_σ*) is a minimiser of* $\mathcal{E}_v^J$ *(resp.* $\mathcal{E}_\sigma^J$*) over* $X_{R,M}$.

Proof.
Part 1: existence of minimisers. All functionals (3.9) and (3.10) are continuous in J with respect to uniform convergence which follows from Lemma 3.5 setting the initial measure $\bar{\mu}$ to be the empirical one (i.e. just looking at converging J for fixed trajectories). The sets $X_{R,M}$ and $X_{R,M}^N$ are compact with respect to uniform convergence. Therefore, minimisers exist.

Part 2: convergence of minimisers. The proof is inspired by [10, Section 4.2]. Since $X_{R,M}^N \subset X_{R,M}$ and the latter is compact, the sequence of minimisers $\left(J_v^N\right)_N$ with $J_v^N \in X_{R,M}^N$ must have a cluster point J_v in $X_{R,M}$. For simplicity, denote the converging subsequence again by $\left(J_v^N\right)_N$. Let $\tilde{J} \in X_{M,R}$. By the uniform approximation property of $X_{R,M}^N$ let $\left(\tilde{J}^N\right)_N$ be a sequence converging uniformly to $\tilde{J}$ and $\tilde{J}^N \in X_{R,M}^N$. By Lemma 3.5 and optimality of each J_v^N,

$$\mathcal{E}_v(J_v) = \lim_{N \to \infty} \mathcal{E}_v^N\left(J_v^N\right) \le \lim_{N \to \infty} \mathcal{E}_v^N(\tilde{J}^N) = \mathcal{E}_v(\tilde{J}).$$

Therefore, J_v minimises $\mathcal{E}_v$ and the optimal values converge. The same argument covers the σ-functionals (3.10) and the corresponding sequence of minimisers $\left(J_\sigma^N\right)_N$. $\qquad\square$

We conclude this section with a couple of comments and remarks on the interpretation of results in Theorem 3.6.

Remark 3.7 (Realisability of data). *In general, observed data may or may not be realisable by the model. This distinguishes two regimes.*

- *Assume there is a true regular J that generated the observed data (as for example in Lemma 3.2). If we pick M large enough in Theorem 3.6, so that $J \in X_{R,M}$, one has* $\mathcal{E}_v(J_v) \le \mathcal{E}_v(J) = 0$ *and analogously* $\mathcal{E}_\sigma(J_\sigma) \le \mathcal{E}_\sigma(J) = 0$. *Hence, minimisers of* $\mathcal{E}_v$ *and* $\mathcal{E}_\sigma$ *over* $X_{R,M}$ *reproduce exactly the trajectories of the system thanks to Theorem 3.3.*

- *Assume there is no true regular J that generated the observed data. This means that, no matter how large M is taken in Theorem 3.6, the minimal limit energies* $\mathcal{E}_\sigma(J_\sigma)$ *and* $\mathcal{E}_v(J_v)$ *may not be equal to 0. However, the bound on the trajectories generated by the inferred*

model, Theorem 3.3, still holds and so the remaining residual values $\mathcal{E}_\sigma(J_\sigma)$ and $\mathcal{E}_v(J_v)$ are then an indication for how well our parametrised family of interaction models was able to capture the structure in the observed data.

Remark 3.8 (Non-uniqueness of minimisers). *We point out that the parametrisation of an entropic game model by a payoff function J is not unique. Replacing J by $\hat{J}(x, u, x', u') := J(x, u, x', u') + g(x, x', u')$ yields the same agent trajectories for arbitrary $g : \Omega \times \Omega \times U$. This can be seen from the fact that $f^{J+g} = f^J$ for f^J and f^{J+g} as given in (2.3). Intuitively, given some x, x' and u', which strategy u is most attractive does not change, if we add a constant benefit for all potential choices. The analogous observation holds for the undisclosed case. This implies that minimisers of (3.8), (3.9) and (3.10) are not unique. This has to be taken into account when we try to interpret an inferred J. In our numerical examples with the undisclosed fast-reaction model, we usually made additional structural assumptions on J (see Section 4.1) and as a consequence did not observe issues with non-unique minimisers. The phenomenon is encountered in a proof-of-concept example on the differential inference functional, Section 4.3.*

4 Numerical experiments

In this section, we present some numerical examples for the entropic game models of Section 2 and their inference with the functionals of Section 3 to provide a basic proof-of-concept and some intuition about their behaviour.

4.1 Numerical setup

Preliminaries and discretisation. We assume for simplicity that the strategy space U is finite, that is, $U = \{u_1, \ldots, u_K\}$ for some $K > 0$ and fix the reference measure $\eta = \sum_{k=1}^{K} \frac{1}{K} \delta_{u_k}$ as the normalised uniform measure over U.

For a system of N agents starting at positions $x_1(0), \ldots, x_N(0)$, we observe S snapshots of the evolution at times $t_s = s \cdot \frac{T}{S}$ for $s = 1, \ldots, S$, so that we are given locations $\{x_i(t_s)\}_{s,i}$, velocities $\{\partial_t x_i(t_s)\}_{s,i}$ and (sometimes) mixed strategies $\{\sigma_i(t_s)\}_{s,i}$ where indices s and i run from 1 to S and 1 to N respectively. Here, $\sigma_i(t_s) = (\sigma_{i,1}(t_s), \ldots, \sigma_{i,K}(t_s))$ is a discrete probability density with respect to η.

Let $\Omega = \prod_{i=1}^{d}[\ell_i, r_i]$, $\ell_i, r_i \in \mathbb{R}$, be the smallest hypercube containing all observed locations. We discretise Ω by a regular Cartesian grid and describe the unknown payoff J by its values $J(\cdot, u_k, \cdot)$ at grid points for $u_k \in U$ (or $J(\cdot, u_k, \cdot, u_{k'})$ in the fully general setting). Between grid points, J is extended by coordinate-wise linear interpolation, that is, we consider piecewise d-linear finite elements over hypercubes.

Within this setting, the inference functional (3.10) reduces to the discrete error functional

$$\mathcal{E}_\sigma^N(J) = \frac{1}{SNK} \sum_{s=0}^{S} \sum_{i=1}^{N} \sum_{k=1}^{K} \log\left(\frac{\sigma_{i,k}(t_s)}{\sigma_{i,k}^J(t_s)}\right) \sigma_{i,k}(t_s), \qquad (4.1)$$

where for $s \in \{0, \ldots, S\}$, $i \in \{1, \ldots, N\}$ and $k \in \{1, \ldots, K\}$, the optimal mixed strategy at time t_s is given by (2.15) as

$$\sigma_{i,k}^J(t_s) = \sigma_i^J(x_1(t_s), \ldots, x_N(t_s))(u_k) = \frac{\exp\left(\frac{1}{\varepsilon N}\sum_{j=1}^N J(x_i(t_s), u_k, x_j(t_s))\right)}{\frac{1}{K}\sum_{k'=1}^K \exp\left(\frac{1}{\varepsilon N}\sum_{j=1}^N J(x_i(t_s), u_{k'}, x_j(t_s))\right)}.$$

Similarly, the inference functional (3.9) reduces to the discrete functional

$$\mathcal{E}_v^N(J) = \frac{1}{SN}\sum_{s=0}^S \sum_{i=1}^N \left\| v_i^J(t_s) - \partial_t x_i(t_s) \right\|^2, \tag{4.2}$$

where

$$v_i^J(t_s) = v_i^J(x_1(t_s), \ldots, x_N(t_s)) = \frac{1}{K}\sum_{k=1}^K e(x_i(t_s), u_k)\, \sigma_{i,k}^J(t_s).$$

Reducing dimensionality. When $d \geq 2$, describing a general function $J : \Omega \times U \times \Omega \to \mathbb{R}$ where $\Omega \subset \mathbb{R}^d$ (or even $\Omega \times U \times \Omega \times U \to \mathbb{R}$ in the fully general setting) requires many degrees of freedom (one value for each node on the grid) and inferring them from data in a reliable way would require a large set of observations. After inference, interpreting such a general function to understand the interactions between agents can be a daunting task. For these two reasons, we may wish to incorporate some prior knowledge on the structure of J via a suitable ansatz. A prototypical example is

$$J(x, u, x') = J_1(x, u) + J_2(x' - x, u), \tag{4.3}$$

where the first term models how a single agent prefers to move by itself depending on its absolute location, and the second term models a translation invariant pairwise interaction with other agents. This reduces the dimensionality from one function over $\mathbb{R}^d \times U \times \mathbb{R}^d$ to two functions over $\mathbb{R}^d \times U$. Further simplifications can be made, for instance, by assuming that the interaction is also rotation invariant. Such parametrisations are very common in Newtonian models, see for instance [10], where radially symmetric interaction functions are assumed. Unless noted otherwise, we will in the following use the ansatz (4.3) in our experiments.

Regularisation and dropping the Lipschitz constraints. Even after the reduction (4.3), inferring the coefficients of J is a high-dimensional nonlinear inverse problem. We may observe no (or very few) agents near some grid points and thus have no reliable data to set the corresponding coefficients of J directly. A common approach to avoiding ill-posedness is to add regularisation to the minimisation problem. Intuitively, this will diffuse information about observations over the grid to some extent. For regularisation in Ω, we use the squared L^2-norm of the gradient of J, which boils down to weighted sums of the finite differences between grid points. For the strategy space U, one can in principle do the same, although this was not necessary in our examples. When J is split as in (4.3), a regulariser can be applied separately to each term. The regularised version of (4.1) then becomes

$$\mathcal{E}_\sigma^N(J) + \lambda_1 \cdot \mathcal{R}_1(J_1) + \lambda_2 \cdot \mathcal{R}_2(J_2). \tag{4.4}$$

Of course, over-regularisation will lead to loss of contrast and high-frequency features of J. The effect of under-regularisation is illustrated in Example 4.5.

While in Theorem 3.6 we consider increasingly finer discretisations $X_{R,M}^N$, $N \to \infty$, of the payoff function space $X_{R,M}$, in our numerical examples a fixed resolution is entirely sufficient (approximately 30 grid points along any spatial dimension). On a fixed grid, the Lipschitz and L^2-gradient semi-norms are equivalent. Consequently, for simplicity, in our numerical experiments we may drop the explicit constraint on the Lipschitz constant of J and only use the regularisation term to impose regularity. This has the advantage that our finite-dimensional discrete problem is then unconstrained, and we may approach it with quasi-Newton methods such as L-BFGS, for which we use the Julia package Optim [52].

Inferring $\sigma_i(t_s)$ from observations. In many situations, we are not able to observe the mixed strategies of agents, but merely their locations and velocities. We can then decide to use the nonconvex velocity-based error functional (4.2). Alternatively, in case we want to optimise the convex inference functional $\mathcal{E}_\sigma^N$ (4.1) instead, we can try to infer some meaningful $\sigma_i(t_s)$ from the observed velocities. In both cases, we need to make suitable choices for U and e. Unless stated otherwise, we pick U as a discrete subset of $\mathbb{R}^d$ such that observed velocities lie in the (interior of the) convex hull of U and pick $e(x, u) := u$. This means that every observed velocity can be reproduced by some mixed strategy. As a heuristic method for inferring a viable mixed strategy σ from a velocity v, we propose to solve the following minimisation problem:

$$\arg\min \left\{ \int_U \left[\varepsilon \cdot \log(\sigma(u)) + \|u - v\|^2 \right] \sigma(u) \, \mathrm{d}\eta(u) \, \middle| \, \sigma \in \mathcal{S}_{a,b}, \int_U u \cdot \sigma(u) \, \mathrm{d}\eta(u) = v \right\} \quad (4.5)$$

The constraint enforces that σ reproduces the observed velocity v. The term $\|u - v\|^2 \, \sigma(u)$ encourages σ to be concentrated in the vicinity of v in U, whereas the entropic term $\varepsilon \cdot \log(\sigma(u)) \, \sigma(u)$ keeps σ from collapsing onto a single point. It is easy to see that minimisers of (4.5) are of the form $\sigma(u) = A \cdot \exp(-\|u - \tilde{v}\|^2/\varepsilon)$ where A is a normalisation factor and $\tilde{v}$ is a velocity close to v, potentially slightly shifted to make sure that the constraint $\int_U u \cdot \sigma(u) \, \mathrm{d}\eta(u) = v$ is satisfied. (4.5) is readily solved by standard methods. We consider this approach a heuristic 'inverse' of the approximation of Newtonian models by game models, as outlined in Example 2.7.

Observations from forward models: Subsampling and multiple realisations. In all the examples here below, observations are obtained by means of simulating a forward model with explicit Euler-stepping with time-step size $\Delta t = 0.02$. Since the numerical complexity of our inference functionals grows linearly with the number of observations, we usually choose to sub-sample the trajectories in time, keeping only every δ-th step, since intuitively the information contained in subsequent steps is largely redundant as agents are confronted with a very similar configuration. Unless noted otherwise, we set $\delta = 2$.

The quality of the reconstructed model is highly related to the data quality: more observations in a given region imply a better reconstruction of the payoff function in that specific region. Generally, if we consider as inputs only data coming from a single realisation of the system we have highly correlated observations, which tend to concentrate only on a small subset of the reconstruction domain. Vaster regions of the domain are then explored taking $N \to \infty$. However, another option to enhance the quality of the reconstructed model is to combine data from multiple realisations with different initial conditions, that is, we fix the number of agents N and simulate the system for multiple different initial conditions. This provides a good coverage of the domain. Extending the functionals in Section 3 to multiple observations is straightforward.

4.2 Learning from undisclosed fast-reaction game models

We now begin with relatively simple numerical examples: an explicit game model is fixed (including U, e, J), simulations according to this model are performed and subsequently the payoff function is inferred from the observations. By possibly re-scaling J, we may w.l.o.g. set $\varepsilon = 1$ for all numerical experiments, see (2.15). In this section we focus on the undisclosed fast-reaction setting. A proof of concept of inference for the full entropic game models via the differential functional of Section 3.2 is discussed in Section 4.3.

Example 4.1 (Game model, $d = 1$). Let $d = 1$, $U = \{-1, +1\}$ and $e(x, u) = u$. We set

$$J(x, u, x') := J_1(x, u) + J_2(x' - x, u), \qquad J_1(x, u) := -u \cdot x,$$
$$J_2(\Delta x, u) := -u \cdot \tanh(5\Delta x) \cdot \left(\max\left\{1 - |\Delta x|^2, 0\right\}\right)^2. \tag{4.6}$$

J_1 encourages agents to move towards the origin $x = 0$; for example, for $x > 0$ one has $J_1(x, -1) > J_1(x, +1)$, thus preferring strategy (and velocity) $u = -1$ over $u = +1$. J_2 implements repulsion between nearby agents: when $\Delta x \in (0, 1)$, that is, the 'other agent' is to the right of 'me' and relatively close, then $J_2(\Delta x, -1) > J_2(\Delta x, +1)$, and I am encouraged to move left.

We set $N = 8$, simulated 100 instances over a time span of 0.2, collecting 5 data points per instance, that is, a total of 500 observed configurations (consisting of locations and mixed strategies). In each instance, initial locations are sampled uniformly from $[-1, 1]$. Observed relative locations between two agents are thus distributed over $[-2, 2]$. Describing the discrete J required 178 coefficients. Since observed mixed strategies are available in this example, we use the energy (4.1) for inference augmented with a regulariser as discussed in Section 4.1, with $\lambda_1 = \lambda_2 = 10^{-6}$, see (4.4). The results are illustrated in Figure 4. We find that, as intended, the model describes agents moving towards the origin while avoiding getting too close to each other. The functions J_1 and J_2 are accurately recovered from the data. For newly generated initial conditions, that is, not taken from the training data, the trajectories simulated with the inferred J are in excellent agreement with the underlying true model, demonstrating that the inferred model is able to generalise.

Example 4.2 (Towards the mean-field regime). We can also explore the mean-field regime numerically. Figure 5 shows trajectories for the same model as in the previous example for an increasing number of particles where the initial locations are sampled from some underlying distribution. As anticipated, the behaviour approaches a consistent limit.

Inference can also be performed for larger numbers of particles. We applied the same approach as in the previous example with $N = 100$ agents. Since we are dealing with a large number of agents, already a small number of configurations carries enough information for the inference. Thus, we simulate 10 instances over a time span of 0.02 with time-step size $\Delta t = 0.002$, collecting 2 data points per instance; that is, a total of 20 observed configurations (for locations and mixed strategies). In each instance, initial locations are sampled uniformly from $[-1, 1]$. The inferred payoff function is essentially identical to the one obtained in Figure 4.

Example 4.3 (Game model, $d = 1$, noisy data). We repeat the previous example with some added noise. We consider the same set of observations, but the observed mixed strategies are corrupted
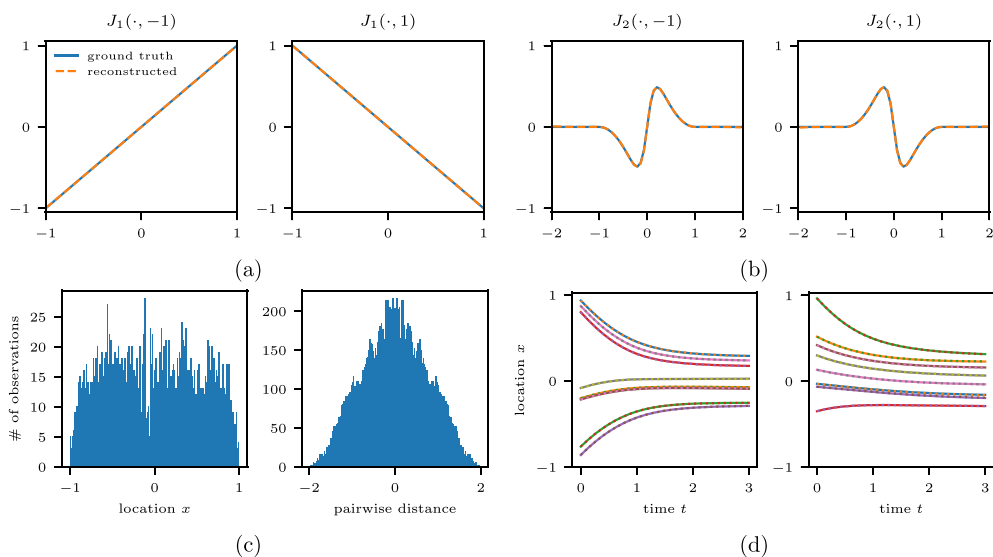
FIGURE 4. Inference from 1d observations of an undisclosed fast-reaction system. (a,b): reconstruction of the payoff function. (c): distribution of the data in the training set. (d): comparison between true trajectories (solid lines) and trajectories generated by the inferred model (dashed lines) for two new realisations (not part of the training data).
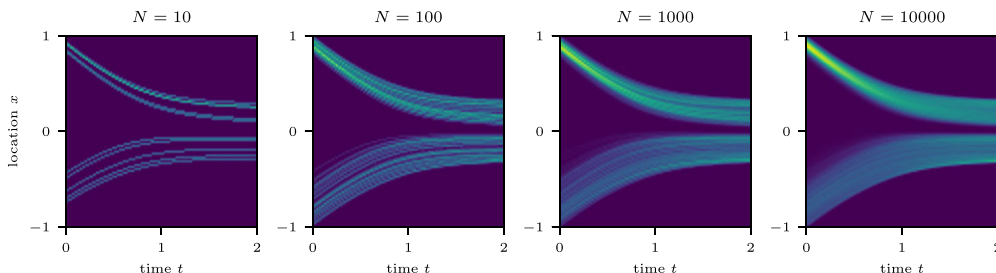

FIGURE 5. Numerical approximation of the mean-field limit. The forward model of Example 4.1 is simulated for an increasing number of agents. Each panel shows a discrete histogram of the particle distribution over time. As N increases, the histograms approach a consistent limit.

by re-sampling: For a given 'true' simulated σ, we draw 20 times from U, according to σ and now use the resulting empirical distribution for inference. These new corrupted σ provide corrupted velocities predictions: in our example, the overall standard deviation between exact and corrupted velocities is ≈ 0.2 (with the range of admissible velocities being $[-1, 1]$). This could model the situation where observation of the mixed strategies is imperfect or by observing very noisy (stochastic) agent movement and using empirical distributions of agent velocities over short time intervals as substitute for mixed strategies. The results are shown in Figure 6, for parameters $\lambda_1 = \lambda_2 = 10^{-5}$ to enforce more regularity due to noisy inputs. The inferred payoff functions are similar to the previous example with just a few spurious fluctuations. The trajectories simulated with the inferred J are close to the (unperturbed) true trajectories, the error is consistently smaller than the one caused by the re-sampling noise, indicating that inference also works well in noisy settings.
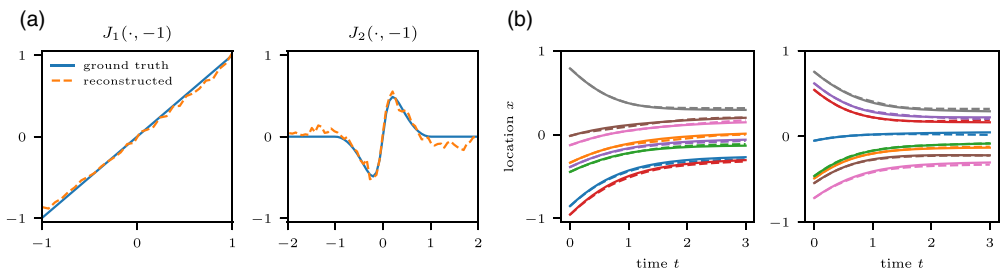
FIGURE 6. Inference from noisy 1d observations of an undisclosed fast-reaction system. (a): reconstruction of the payoff function for $u = -1$. (b): comparison between true trajectories (solid lines) and trajectories generated by the inferred model (dashed lines) for multiple new realisations (not part of the training data).

Example 4.4 (Game model, $d = 2$). Let $d = 2$, $U = \{(1,1), (-1,1), (-1,-1), (1,-1)\} \subset \mathbb{R}^2$ and $e(x, u) = u$. We use the same J defined in (4.6), thus combining a driving force towards the origin $x = (0,0)$ and repulsion between nearby agents (we assume tanh to be applied componentwise to vectors).

Data collection follows the same approach as in Example 4.1: for $N = 8$ agents we simulate 100 instances over a time span of 0.2, collecting 5 data points per instance, for a total of 500 observed configurations. In each instance, initial locations are sampled uniformly from $[-0.75, 0.75]^2$. Observed relative locations between two agents are thus distributed over $[-1.5, 1.5]^2$, as displayed in Figure 7(b). Describing the discrete J required 10,656 coefficients, resulting from a 30×30 spatial grid for J_1 and a 42×42 spatial grid for J_2. We use again the energy (4.1) for inference with a regulariser on derivatives, with $\lambda_1 = \lambda_2 = 10^{-5}$. The results are illustrated in Figure 7. The recovered functions J_1 and J_2 reproduce the same structural features of the exact ones and the trajectories simulated with the inferred J follow closely the underlying true model (also in this case newly generated initial conditions are considered).

Example 4.5 (Under-regularisation). As is common in inverse problems, if the regularisation parameters λ_1, λ_2 in (4.4) are chosen too high, one obtains an over-regularised result, where the large penalty for spatial variation leads to a loss of contrast and high-frequency features of J. Conversely, we may experience under-regularisation when the parameters are small and the observed data does not cover the learning domain sufficiently, for example when we only observe few realisations of the system. Such a case is illustrated in Figure 8(c) where the observations are concentrated on a few one-dimensional trajectories. For weak regularisation, the reconstructed J then tends to be concentrated on these trajectories as well, see Figure 8(a). And while the inferred J may predict the agent behaviour with high accuracy on these trajectories, it will generalise very poorly to different data. With stronger regularisation, J is extended beyond the trajectories in a meaningful way, see Figure 8(b).

4.3 An example for differential inference

The 'differential inference' functional $\mathcal{E}_{\hat{\sigma}}^{J,N}$ introduced in Section 3.2 was constructed in close analogy to (1.7). We have argued that it may be challenging to apply to real data, as it would require the observation of mixed strategies $\{\sigma_i(t_s)\}_{s,i}$ and their temporal derivatives $\{\partial_t \sigma_i(t_s)\}_{s,i}$.
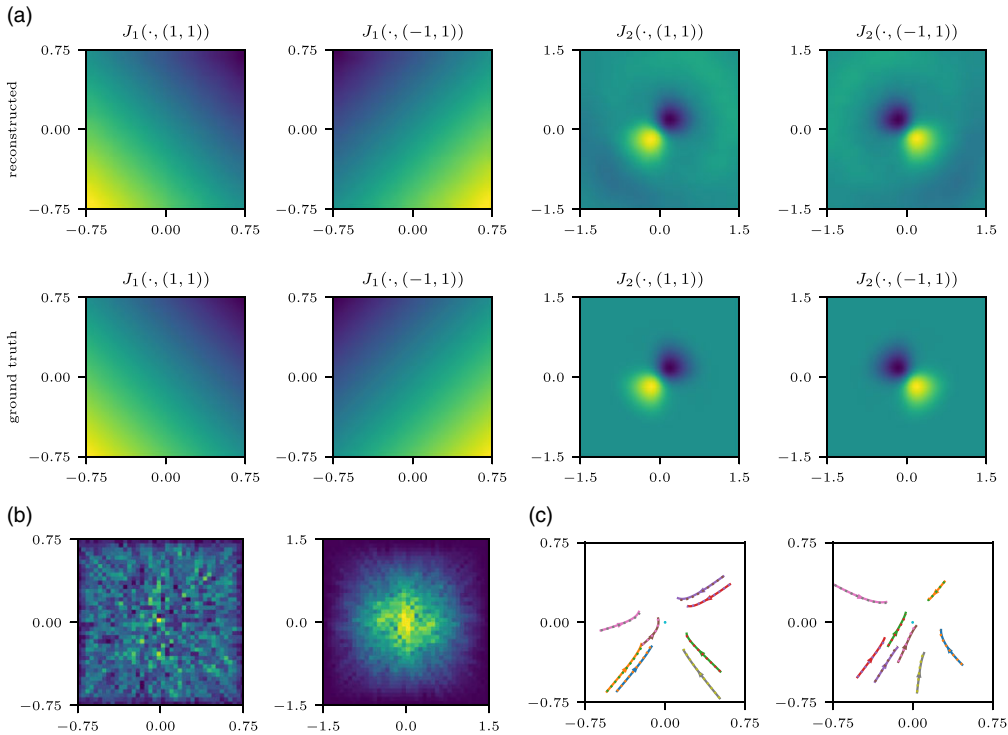
FIGURE 7. Inference from 2d observations of an undisclosed fast-reaction system. (a): inferred payoff function and ground truth for pure strategies (1,1) and (−1, 1). (b): distribution of the data in the training set (agents locations and relative locations), dark blue : low, yellow: high. (c): comparison between exact trajectories (solid lines) and trajectories generated by the inferred model (dashed lines) for two new realisations (not part of the training data).

As a proof of concept, we now demonstrate that if these observations are indeed available, the numerical setup described up to now can be easily adapted to $\mathcal{E}_{\hat{\sigma}}^{J,N}$. After discretisation, $\mathcal{E}_{\hat{\sigma}}^{J,N}$ reduces to a finite-dimensional quadratic objective. For a quadratic regulariser, such as the squared L^2-norm of the gradient of J, minimisation then corresponds to solving a linear system. Alternatively, we can implement the Lipschitz constraint on J which amounts to adding linear constraints to the quadratic problem. We solve the resulting finite-dimensional optimisation problem using the operator splitting approach provided by the OSQP solver [69].

As an example, we simulate the entropic regularised model in (2.2) for $U = \{-1, 1\}$ and $J: \mathbb{R} \times \{-1, 1\} \times \mathbb{R} \times \{-1, 1\}$ defined by

$$J(x, u, x', u') = -\frac{1}{2} (u + u') \left((u + 1)x^5 + (u - 1)(x + x')^3 \right)$$

Our choice of J is not motivated by providing a meaningful model for anything, but simply to illustrate that inference works in principle. We collect data using the same sub-sampling approach presented above: 8 agents, 100 realisations for a time span of 0.2, initial conditions uniformly sampled from $[-1, 1]$, only every other observation is kept for a total of 500 input configurations. Observed positions (x, x') with $x \neq x'$ are displayed in Figure 9(a). No particular structure is assumed for J.
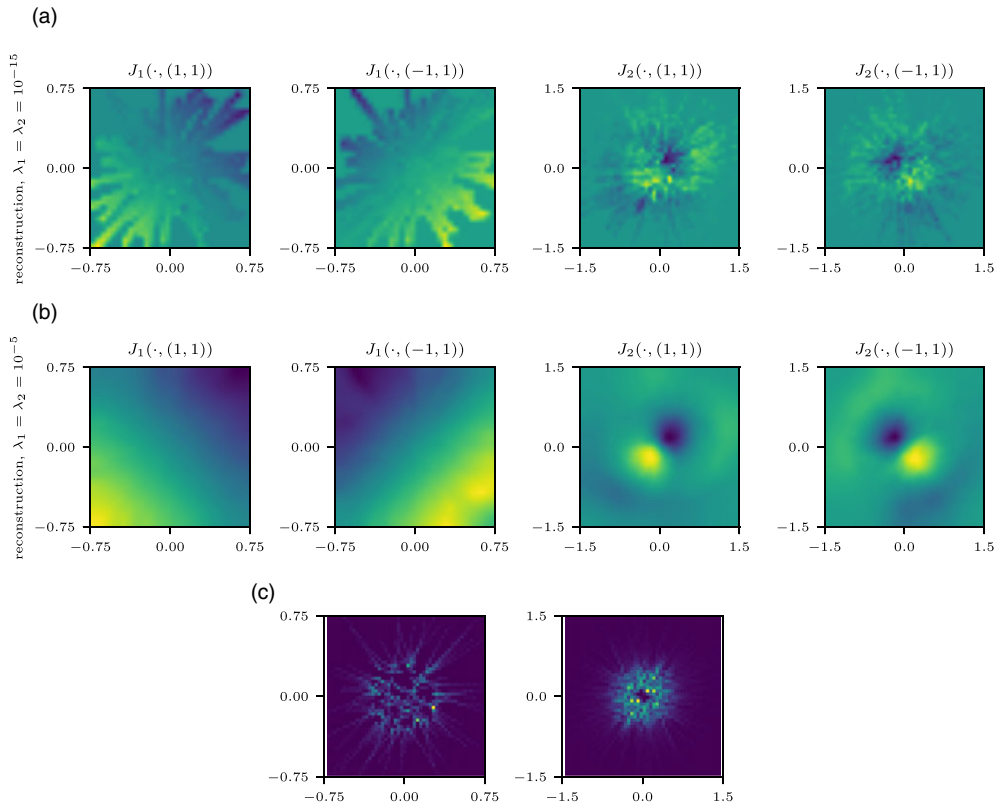
FIGURE 8. Influence of the regularisation parameter. Analogous to Figure 7 but with fewer observations, as displayed in (c). (a): reconstruction for low regularisation parameters, $\lambda_1 = \lambda_2 = 10^{-15}$. (a): reconstruction for higher regularisation parameters, $\lambda_1 = \lambda_2 = 10^{-5}$. See Example 4.5 for more details.
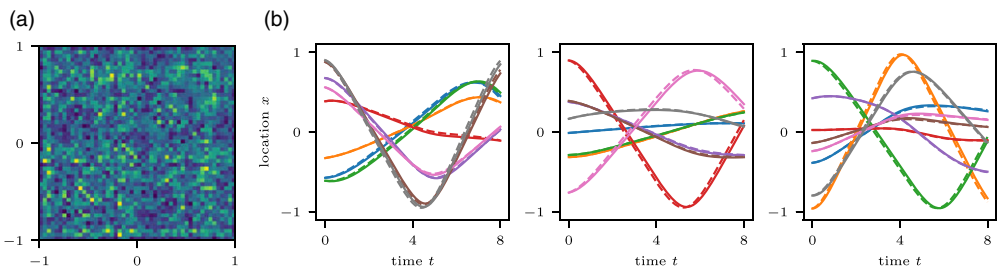

FIGURE 9. Inference from 1d observations generated with (2.2). (a): distribution of observed pairs (x, x'), $x \neq x'$. (b): comparison between exact trajectories (solid lines) and inferred model (dashed lines) on new realisations (not part of the training data).

As discussed in Remark 3.8, the optimal inferred J is not unique. As a remedy, we normalise the reconstructed J^r such that $J^r(x, 1, x', -1) = J^r(x, -1, x', 1) = 0$. Comparisons for trajectories are reported in Figure 9(b) for new randomly selected initial conditions with $T = 8$: trajectories with the reconstructed J^r follow closely the trajectories generated with the true J.
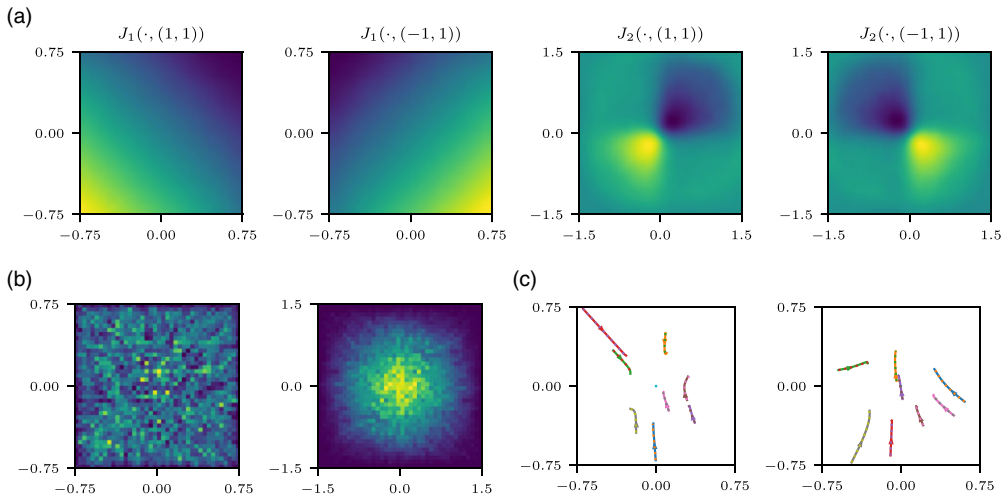
FIGURE 10. Inference from 2d observations of a Newtonian system. Top rows: inferred payoff function for pure strategies (1,1) and (−1, 1), bottom line: distribution of the data in the training set (agents locations and pairwise distances) and comparison between exact Newtonian trajectories (solid lines) and trajectories generated by the inferred game model (dashed lines).

4.4 Learning for Newtonian models

In Example 2.7, we have shown that entropic game models with fast reaction can also capture Newtonian interaction models. In this section, we perform inference with an entropic game model ansatz on data generated by a Newtonian model. As an example, we pick

$$\partial_t x_i(t) = \frac{1}{N} \sum_{j=1}^{N} f\big(x_i(t), x_j(t)\big) \qquad \text{with} \qquad f(x, x') = -x - \frac{\tanh\big(5\left(x' - x\right)\big)}{(1 + \|x' - x\|)^2}, \qquad (4.7)$$

where the function tanh is assumed to be applied componentwise. The first term in f encourages agents to move towards the origin, the second term implements a pairwise repulsion at close range. Trajectory data are collected for a two dimensional system: for $N = 8$ agents, we simulate 100 instances of (4.7) over a time span of 0.2, collecting 5 data points per instance to total 500 observed configurations. In each instance, initial locations are sampled uniformly from $[-0.75, 0.75]^2$. Observed relative locations between two agents are thus distributed over $[-1.5, 1.5]^2$, as displayed in Figure 10(b).

We use the parametrisation (4.3) for the payoff function. Unlike in the previous examples, the original model does not involve a strategy set U or a map e. Thus, prior to inference, we have to make informed choices for these. As in earlier examples, we choose $U \subset \mathbb{R}^2$ and $e(x, u) = u$ (i.e. pure strategies correspond directly to velocities) and in particular we pick $U = \{u_1, \ldots, u_K\}$ such that the convex hull of U in $\mathbb{R}^2$ contains (almost) all observed velocities, meaning that they can be reproduced by suitable mixed strategies. So for (almost) every i and s, we can write $\partial_t x_i(t_s) = \sum_k u_k \sigma_k$ for some mixed strategy σ over U.
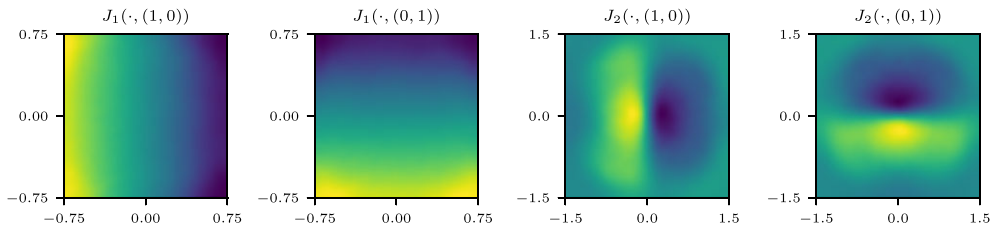
FIGURE 11. Inferred payoff function via $\mathcal{E}_v^N$-based energy for pure strategies (1,0) and (0,1) (same Newtonian input data as in Figure 10(c)).
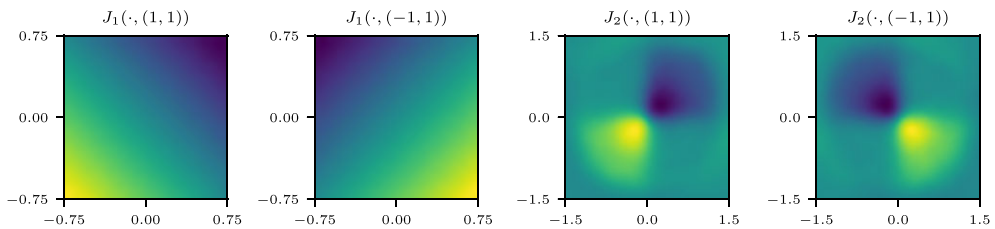

FIGURE 12. Inferred payoff function via $\mathcal{E}_\sigma^N$-based energy for pure strategies (1,1) and $(-1,1)$ (same Newtonian input data as in Figure 10(c)). To compare with $\mathcal{E}_v^N$-based reconstruction in Figure 10(a).

Example 4.6. We pick $U = \{(1,1), (-1,1), (-1,-1), (1,-1)\} \subset \mathbb{R}^2$, optimise the velocity-based energy (4.2) because no mixed strategies are provided by observations and consider a regulariser on derivatives as described in Section 4.1, with $\lambda_1 = \lambda_2 = 10^{-5}$. Using the same setup as Example 4.4 we describe the discrete J using 10656 coefficients (30×30 grid for J_1 and 42×42 grid for J_2). The results are illustrated in Figure 10. The recovered functions J_1 and J_2 have the same structure as the ones in Example 4.4: the qualitative behaviour of the two systems is indeed the same. In Figure 10(c), we simulate a couple of dynamics with newly generated initial conditions and observe again how Newtonian trajectories and game trajectories computed through the reconstructed J are essentially the same.

Example 4.7. In most cases, there will not be a unique preferred choice for U and various options could be reasonable. We now demonstrate robustness of the inference results between different choices that are essentially equivalent. Thus, we repeat Example 4.6 changing U into $U = \{(1,0), (0,1), (-1,0), (0,-1)\}$. The results are illustrated in Figure 11. Essentially, a rotation of U caused a corresponding rotation in the reconstructed functions, without changing the underlying behaviour that is being described.

Example 4.8. In the previous two examples, we used the velocity-based functional $\mathcal{E}_v^N$ (4.2) for inference. As discussed in Section 4.1, we can also optimise the error functional $\mathcal{E}_\sigma^N$ (4.1) upon providing some suitable mixed strategies via (4.5). We repeat Example 4.6 with this approach. The results are illustrated in Figure 12: the reconstructed functions closely resemble the previous reconstruction provided in Figure 10(a).

4.5 Pedestrian dynamics

The examples presented in Sections 4.2 and 4.3 were based on observations generated with the underlying entropic game model with known U and e. The examples in Section 4.4 were based on observations generated by Newtonian models, which by virtue of Example 2.7 are known to be covered by the entropic game models, if suitable U and e are chosen. Here we consider a more challenging scenario, trying to learn interaction rules between pedestrians in a model adapted from [7, 32]. This model attempts to describe 'rational' agents (pedestrians) that avoid collisions with other agents not merely via a repulsion term involving the other agents' current locations, but by explicit anticipation of their trajectories. Related datasets and models in the computer vision community can be found, for instance, in [59, 62, 76]. It will be interesting future work to see if our model can be trained on such data. Note that this differs from the setting where pedestrians in high-density and high-stress situations are modelled as a non-sentient fluid that is solely driven by external gradients and internal pressure, see for instance [9, 41] for reviews on the literature of micro- and macroscopic crowd motion models.

Forward model. We now describe the model for pedestrian movement that we use to generate observations. It is slightly adapted from [7, 32], to which we refer for more details and motivation of the quantities we re-introduce here. We consider N pedestrians in $\mathbb{R}^2$, each of which with desired target velocity $\bar{v}_i \in \mathbb{R}^2$, $i = 1, \ldots, N$. The dynamic is described by the coupled system of second-order equations

$$\begin{cases} \partial_t x_i(t) = v_i(t), \\ \partial_t v_i(t) = -\nabla_{v_i} \Phi_i(x_1(t), \ldots, x_N(t), v_1(t), \ldots, v_N(t)), \\ \|v_i(t)\| = 1, \end{cases} \tag{4.8}$$

where the third line means that we confine the gradient movement of the velocities in the second line to the unit sphere. This reduces the dimension of the strategy space and thus simplifies inference (see below). The potential Φ_i is designed to steer the velocity towards the target velocity $\bar{v}_i$ while at the same time trying to avoid as much as possible close encounters with other pedestrians. The potential is computed from pairwise interaction terms that involve anticipated (near-)collision events. A key assumption is that agents are able to evaluate at each time the current locations and velocities of the other agents in their cone of vision; that is, they fully know the relevant status of the system. Thus, for locations $x_1, \ldots, x_N$ and velocities $v_1, \ldots, v_N$, each agent i assumes the others' velocities to be fixed and can evaluate 'optimality' of a new velocity v through the following quantities:

$$D_{i,j}(v) = -\frac{(x_j - x_i) \cdot (v_j - v)}{\|v_j - v\|^2} \|v\| \tag{4.9}$$

$$C_{i,j}(v) = \left(\|x_j - x_i\|^2 - \frac{((x_j - x_i) \cdot (v_j - v))^2}{\|v_j - v\|^2} \right)^{1/2} \tag{4.10}$$

$$S_{i,j}(v) = \begin{cases} 1 & \text{if } \dfrac{(x_j - x_i) \cdot v}{\|x_j - x_i\| \|v\|} > \cos\left(\dfrac{7}{12}\pi\right) \text{ and } D_{i,j}(v) > 0 \\ 0 & \text{else.} \end{cases} \tag{4.11}$$

$$0 \;\leq\; D_{i,j} \;\leq\; L$$
distance to closest approach

$x_i \quad v_i$

$$0 \;\leq\; C_{i,j} \;\leq\; R$$
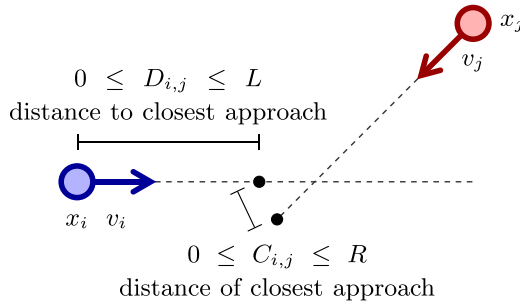distance of closest approach

x_j

v_j

FIGURE 13. Adapted from [7].

These are illustrated in Figure 13. $D_{i,j}$ is the travelling distance of i to the closest encounter with j, based on velocities v and v_j. It is negative if the closest encounter would have happened in the past. $C_{i,j}$ is their distance realised at that closest encounter and $S_{i,j}$ counts whether agent j is visible to i and getting closer. Note that $D_{i,j}$, $C_{i,j}$ and S_j are functions of the other agent's location and velocity. From the perspective of agent i, these are considered parameters, and thus, we drop them in the explicit notation of arguments. For a parameter $a > 0$ define the quadratic penalty

$$\phi_a(x) = \begin{cases} \left(\dfrac{x}{a} - 1\right)^2 & \text{if } x \leq a, \\ 0 & \text{else,} \end{cases}$$

and then set the i-th potential to be

$$\Phi_i(v) = k_1 \|v - \bar{v}_i\|^2 + \frac{k_2}{\sum_{j \neq i} S_{i,j}(v)} \sum_{j \neq i} S_{i,j}(v) \cdot \phi_R\left(C_{i,j}(v)\right) \cdot \phi_L\left(D_{i,j}(v)\right) \tag{4.12}$$

for some parameters $k_1, k_2, L, R > 0$. Agent i will then evolve its velocity based on the gradient of Φ_i, which balances the desire to stay close to the target velocity $\bar{v}_i$ and the avoidance of close encounters with other agents. The second term penalises anticipated minimal distances $C_{i,j}$ below R, that happen on a distance horizon of L, with a stronger penalty if the anticipated encounter is closer.

State space and strategy space. System (4.8) is simulated with potentials Φ_i defined as in (4.12). By the third line of (4.8) we can write $v_i(t) = (\cos(\theta_i(t)), \sin(\theta_i(t)))$ at each t for a suitable time depended angle θ_i. Observations consist of locations $\{x_i(t_s)\}_{s,i}$ and velocities $\{v_i(t_s)\}_{s,i}$ for $s \in \{0, \ldots, S\}$ and $i \in \{1, \ldots, N\}$. In particular, observations on velocities can be recast into observations of the angles $\{\theta_i(t_s)\}_{s,i}$, $\theta_i(t_s) \in S^1$, so that $v_i(t_s) = (\cos(\theta_i(t_s)), \sin(\theta_i(t_s)))$. Likewise, the desired velocity $\bar{v}_i$ can be encoded by an orientation $\bar{\theta}_i \in S^1$.

We want to model the system as an undisclosed fast-reaction entropic game. Consequently, we must choose a suitable 'location' space, strategy space and map e for the agents. In the undisclosed setting, the choice of strategy of an agent does not depend on the choices of strategies of the other agents. However, for pedestrians, the movement choices of one agent clearly depend on the current velocities of other agents. This is explicit in the above model, and of course also true in reality, where we can learn about other persons' current movement by their orientation (people walk forward). Consequently, the orientations θ_i cannot be interpreted as strategies and should be considered as part of the agents' locations. Instead, the second line of (4.8) suggests

that strategies describe the agents' choice to change their orientations. Finally, different agents may have different desired orientations $\bar{\theta}_i$; that is, we may have different 'species' of agents. As discussed in Remark 2.4, this can be formally incorporated into our model by enriching the physical space.

Consequently, as physical space we use $\mathbb{R}^2 \times S^1 \times S^1$ where the first component describes the physical location of the agent, the second component their current orientation and the third component their desired orientation.

As strategy space, we pick $U \subset \{-C, +C\}$ for some $C > 0$, where a pure strategy $u \in U$ corresponds to the desire to change the orientation with rate u. C should be picked sufficiently large to capture the maximally observed changes of orientation of the agents.

In summary, in this particular instance of (2.15) we choose

$$
e\left((x, \theta, \bar{\theta}), u\right) = \left(\cos(\theta), \sin(\theta), u, 0\right),
$$

and consequently obtain the full system

$$
\begin{cases}
\partial_t x_i(t) = (\cos(\theta_i(t)), \sin(\theta_i(t))) \\
\partial_t \theta_i(t) = \vartheta_i^J\left((x_1(t), \theta_1(t), \bar{\theta}_1(t)), \ldots, (x_N(t), \theta_N(t), \bar{\theta}_N(t))\right) \\
\qquad := \dfrac{1}{K} \sum_{k=1}^{K} u_k \cdot \sigma_i^J\left((x_1(t), \theta_1(t), \bar{\theta}_1(t)), \ldots, (x_N(t), \theta_N(t), \bar{\theta}_N(t))\right)(u_k) \\
\partial_t \bar{\theta}_i(t) = 0
\end{cases}
\tag{4.13}
$$

Ansatz for payoff function. For the payoff function, in analogy to (4.3), we make the following ansatz:

$$
J\left((x, \theta, \bar{\theta}), u, (x', \theta', \bar{\theta}')\right) = J_1\left(\theta, u, \bar{\theta}\right) + J_2\left(R_{-\theta} \cdot (x' - x), \theta' - \theta, u\right),
\tag{4.14}
$$

where $R_{-\theta}$ denotes the rotation matrix by angle $-\theta$. J_1 is intended to recover the drive towards the desired orientation $\bar{\theta}_i$ and J_2 the pairwise interaction between agents. We have added the natural assumption that J_2 is invariant under translations and rotations. For each agent, we may choose a coordinate system such that they are currently in the origin ($x = 0$), heading right ($\theta = 0$), whereas the other agent is located at $y = R_{-\theta}(x' - x)$ with orientation $\Delta\theta = \theta' - \theta$. Further dimensionality reduction can be obtained by setting J_2 to zero, when the other agent is not within the agent's field of vision. Alternatively, we can attempt to infer the field of vision from the observations during learning.

Discrete inference functional. We approximate the missing velocities $\{\partial_t \theta_i(t_s)\}_{s,i}$ by finite differences,

$$
\partial_t \theta_i(t_s) \approx \frac{\theta_i(t_s) - \theta_i(t_{s-1})}{t_s - t_{s-1}} \quad \text{for all } s = 1, \ldots, S, \text{ and } i = 1, \ldots, N,
$$

and the corresponding velocity-based inference functional (3.9) is given by

$$
\mathcal{E}_v^N(J) = \frac{1}{SN} \sum_{s=1}^{S} \sum_{i=1}^{N} \left|\vartheta_i^J(t_s) - \partial_t \theta_i(t_s)\right|^2,
\tag{4.15}
$$

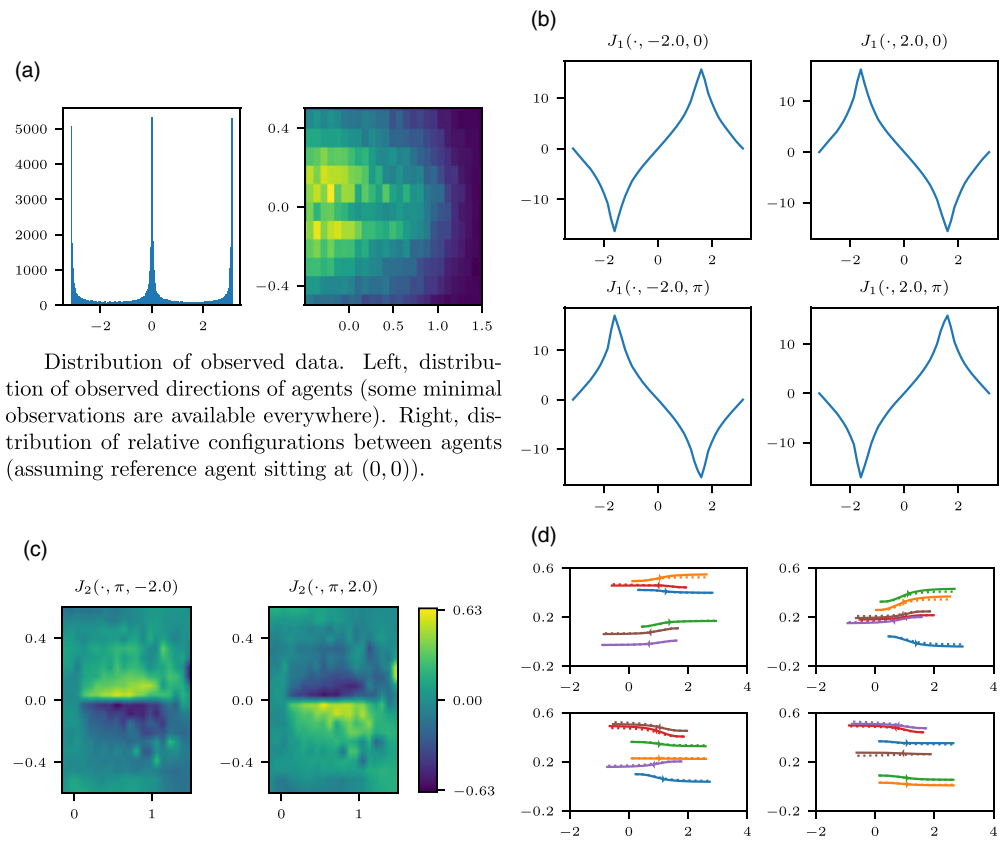where ϑ_i^J is introduced in (4.13).

FIGURE 14. Inference from 2d observations of a pedestrian model. (a): distribution of the data in the training set. (b,c): reconstruction of the payoff function (section for J_2 at $\Delta\theta = \pi$). (d): comparison between exact trajectories (solid lines) and trajectories generated by the inferred model (dashed lines) for four new realisations (not part of the training data).

Example 4.9. Analogously to previous examples, we simulate multiple realisation of (4.8): for each instance, we consider 6 agents, 3 starting at random locations inside $[0, 0.5] \times [0, 0.5]$ with desired direction $\bar{\theta} = 0$, while the other 3 start somewhere within $[1.0, 1.5] \times [0, 0.5]$ with target direction $\bar{\theta} = \pi$. We simulate 400 instances of (4.8), assuming for 300 of them a slight perturbation of the target direction as initial direction, and sampling randomly in $(-\pi, \pi)$ for the initial directions of the remaining 100 runs (to observe the behaviour at different directions). Each run we perform 500 steps of an explicit Euler-stepping with $\Delta t = 0.005$. Observations are then subsampled every 20-th step. The final distribution of directions within the training data is reported in Figure 14(a).

In the reconstruction, we consider $U = \{-2, 2\}$. Energy (4.15) is then optimised coupled with regularisers in the same fashion as in (4.4), using a 30 nodes regular grid over $[-\pi, \pi]$ for J_1 and a $20 \times 20 \times 20$ grid over $[-0.15, 1.5] \times [-0.6, 0.6] \times [-\pi, \pi]$ for J_2 (we assume a priori that agents outside my cone of vision go unseen). Results are reported in Figure 14. The self-interaction function J_1, Figure 14(b), correctly dictates the steering of the direction towards the

desired direction. Observe how the self-driving force to go back to the desired direction is maximal around $\pm\pi/2$, as expected when looking at potential (4.12). The interaction function J_2, Figure 14(c), encodes the expected deviation for avoiding a close interaction: when approaching an agent to my left I favour deviations to the right and vice versa. Eventually, we compare in Figure 14(d) simulated trajectories obtained with the ground truth model and the inferred one, starting at locations outside the training set. The two models reproduce the same behaviour.

5 Conclusion and outlook

In this article, we proposed several modifications to the model of spatially inhomogeneous evolutionary games as introduced by [6]. We added entropic regularisation to obtain more robust trajectories, considered the undisclosed setting and then derived the quasi-static fast-reaction limit. Relying on results from [6], we established well-posedness and a mean-field limit for all considered model variants. The resulting new family of multi-agent interaction models vastly generalises first- and second-order Newtonian models.

We then posed the inverse problem of inferring the agent payoff function from observed interactions. We gave several variational formulations that seek to minimise the discrepancy between the observed particle velocities and those predicted by the payoff function. The mismatch functionals provide quantitative control on the ability of the reconstructed payoff function to replicate the observed trajectories, and the functionals are (under suitable assumptions) consistent in the limit of increasing observed data.

Finally, we demonstrated that the inference scheme is not merely limited to abstract analysis but can be implemented numerically. Our computational examples included the inference from models for which a true payoff function was available (and known), examples where this was not the case (e.g. data generated by a Newtonian model or a simple 'rational' model for pedestrian movement), and examples with simulated noise. These showed that our scheme can infer meaningful payoff functions from a broad variety of input data.

A logical next step would be the numerical application to real data. Here one will need to be particularly careful about choosing the family of models to match (i.e. the complexity of the space of pure strategies and the generality of the form of the payoff function). On one hand, a relatively simple class of models will be numerically robust and inexpensive but involve considerable modelling bias due to a potential lack of flexibility. On the other hand, a rather general class of models provides more descriptive power but will require many observations for meaningful inference and be numerically more challenging. On top of that, real data will be corrupted by noise, and thus, a more detailed analysis of learning from noisy data will become crucial (in the spirit of Example 4.3).

Acknowledgement

References

[1] AGOSTINIANI, V., ROSSI, R. & SAVARÉ, G. (2018) Singular vanishing-viscosity limits of gradient flows in Hilbert spaces. Personal communication: in preparation.

[2] AGOSTINIANI, V. & ROSSI, R. (2017) Singular vanishing-viscosity limits of gradient flows: the finite-dimensional case. *J. Differ. Equations* **263**(11), 7815–7855.

[3] ALBANI, V., ASCHER, U. M., YANG, X. & ZUBELLI, J. P. (2017) Data driven recovery of local volatility surfaces. *Inverse Problems Imaging* **11**(5), 799–823.

[4] ALMI, S., FORNASIER, M. & HUBER, R. (2020) Data-driven evolutions of critical points. *Found. Data Sci.* **2**(3), 207–255.

[5] ALMI, S., MORANDOTTI, M. & SOLOMBRINO, F. (2021) A multi-step Lagrangian scheme for spatially inhomogeneous evolutionary games. *J. Evol. Equ.* **21**, 2691–2733.

[6] AMBROSIO, L., FORNASIER, M., MORANDOTTI, M. & SAVARÉ, G. (2018) Spatially inhomogeneous evolutionary games. arXiv:1805.04027.

[7] BAILO, R., CARRILLO, J. A. & DEGOND, P. (2018) Pedestrian models based on rational behaviour. In: *Crowd Dynamics*, Vol. **1**, Springer, pp. 259–292.

[8] BALLERINI, M., CABIBBO, N., CANDELIER, R., CAVAGNA, A., CISBANI, E., GIARDINA, L., LECOMTE, L., ORLANDI, A., PARISI, G., PROCACCINI, A., VIALE, M. & ZDRAVKOVIC, V. (2008) Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study. *PNAS* **105**(4), 1232–1237.

[9] BELLOMO, N. & DOGBE, C. (2011) On the modeling of traffic and crowds: a survey of models, speculations, and perspectives. *SIAM Rev.* **53**(3), 409–463.

[10] BONGINI, M., FORNASIER, M., HANSEN, M. & MAGGIONI, M. (2017) Inferring interaction rules from observations of evolutive systems I: the variational approach. *Math. Models Methods Appl. Sci.* **27**(5), 909–951.

[11] CAMAZINE, S., DENEUBOURG, J. L., FRANKS, N. R., SNEYD, J., THERAULAZ, G. & BONABEAU, E. (2003) Self-organization in Biological Systems, *Princeton University Press*, Princeton.

[12] CAÑIZO, J. A., CARRILLO, J. A. & ROSADO, J. (2011) A well-posedness theory in measures for some kinetic models of collective motion. *Math. Models Methods Appl. Sci.* **21**(03), 515–539.

[13] CARRILLO, J. A., CHOI, Y.-P. & HAURAY, M. (2014) The derivation of swarming models: mean-field limit and Wasserstein distances. In: A. Muntean and F. Toschi (editors), *Collective Dynamics from Bacteria to Crowds*, CISM International Centre for Mechanical Sciences, Springer, pp. 1–46.

[14] CARRILLO, J. A., CHOI, Y.-P. & HAURAY, M. (2014) *The Derivation of Swarming Models: Mean-Field Limit and Wasserstein Distances*, Springer Vienna, Vienna, pp. 1–46.

[15] CARRILLO, J. A., CHOI, Y.-P. & PEREZ, S. P. (2017) *A Review on Attractive–Repulsive Hydrodynamics for Consensus in Collective Behavior*, Springer International Publishing, Cham, pp. 259–298.

[16] CARRILLO, J. A., D'ORSOGNA, M. R. & PANFEROV, V. (2009) Double milling in self-propelled swarms from kinetic theory. *Kinet. Relat. Models* **2**(2), 363–378.

[17] CARRILLO, J. A., FORNASIER, M., TOSCANI, G. & VECIL, F. (2010) Particle, kinetic, and hydrodynamic models of swarming. In: G. Naldi, L. Pareschi, G. Toscani and N. Bellomo (editors), *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, *Modeling and Simulation in Science*, Engineering and Technology, Birkhäuser, Boston, pp. 297–336.

[18] CHUANG, Y. L., D'ORSOGNA, M., MARTHALER, D., BERTOZZI, A. L. & CHAYES, L. S. (2007) State transition and the continuum limit for the 2D interacting, self-propelled particle system. *Physica D* (232), 33–47.

[19] CHUANG, Y. L., HUANG, Y. R., D'ORSOGNA, M. R. & BERTOZZI, A. L. (2007) Multi-vehicle flocking: scalability of cooperative control algorithms using pairwise potentials. In: *IEEE International Conference on Robotics and Automation*, pp. 2292–2299.

[20] CONTI, S., MÜLLER, S. & ORTIZ, M. (2018) Data-driven problems in elasticity. *Arch. Ration. Mech. Anal.* **229**(1), 79–123.

[21] COUZIN, I. D. & FRANKS, N. R. (2002) Self-organized lane formation and optimized traffic flow in army ants. *Proc. R. Soc. Lond.*, B 270, 139–146.

[22] COUZIN, I. D., KRAUSE, J., FRANKS, N. R. & LEVIN, S. A. (2005) Effective leadership and decision making in animal groups on the move. *Nature* **433**, 513–516.

[23] CRÉPEY, S. (2003) Calibration of the local volatility in a generalized Black-Scholes model using Tikhonov regularization. *SIAM J. Math. Anal.* **34**(5), 1183–1206.

[24] CRISTIANI, E., PICCOLI, B. & TOSIN, A. (2010) Modeling self-organization in pedestrians and animal groups from macroscopic and microscopic viewpoints. In: G. Naldi, L. Pareschi, G. Toscani and N. Bellomo (editors), *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, Modeling and Simulation in Science, Engineering and Technology, Birkhäuser, Boston.

[25] CRISTIANI, E., PICCOLI, B. & TOSIN, A. (2011) Multiscale modeling of granular flows with application to crowd dynamics. *Multiscale Model. Simul.* **9**(1), 155–182.

[26] CRISTIANI, E., PICCOLI, B. & TOSIN, A. (2014) *Multiscale Modeling of Pedestrian Dynamics*, Springer, Cham.

[27] CUBITT, T. S., EISERT, J. & WOLF, M. M. (2012) Extracting dynamical equations from experimental data is np hard. *Phys. Rev. Lett.* **108**, 120503.

[28] CUCKER, F. & DONG, J.-G. (2011) A general collision-avoiding flocking framework. *IEEE Trans. Autom. Control* **56**(5), 1124–1129.

[29] CUCKER, F. & MORDECKI, E. (2008) Flocking in noisy environments. *J. Math. Pures Appl. (9)* **89**(3), 278–296.

[30] CUCKER, F., SMALE, S. & ZHOU, D. (2004) Modeling language evolution. *Found. Comput. Math.* **4**(5), 315–343.

[31] CUCKER, F. & SMALE, S. (2007) Emergent behavior in flocks. *IEEE Trans. Automat. Control* **52**(5), 852–862.

[32] DEGOND, P., APPERT-ROLLAND, C., MOUSSAÏD, M., PETTRÉ, J. & THERAULAZ, G. (2013) A hierarchy of heuristic-based models of crowd dynamics. *J. Stat. Phys.* **152**(6), 1033–1068.

[33] EGGER, H. & ENGL, H. W. (2005) Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates. *Inverse Problems* **21**(3), 1027–1045.

[34] FLÅM, S. D. & CAVAZZUTI, E. (2005) Entropic penalties in finite games. *Ann. Operations Res.* **137**, 331–348.

[35] GARAVELLO, M. & PICCOLI, B. (2006) *Traffic Flow on Networks: Conservation Laws Model*. AIMS Series on Applied Mathematics, American Institute of Mathematical Sciences.

[36] GRÉGOIRE, G. & CHATÉ, H. (2004) Onset of collective and cohesive motion. *Phys. Rev. Lett.* **92**, 025702.

[37] HEGSELMANN, R. & KRAUSE, U. (2002) Opinion dynamics and bounded confidence: models, analysis and simulation. *J. Artif. Soc. Soc. Simul.* **5**(3).

[38] HOFBAUER, J. & SIGMUND, K. (1998) *Evolutionary Games and Population Dynamics*, Cambridge University Press, Cambridge.

[39] HUANG, M., CAINES, P. E. & MALHAMÉ, R. P. (2003) Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions. In: *Proceedings of the 42nd IEEE Conference on Decision and Control Maui, Hawaii USA, December 2003*, pp. 98–103.

[40] JADBABAIE, A., LIN, J. & STEPHEN MORSE, A. (2003) Correction to: "Coordination of groups of mobile autonomous agents using nearest neighbor rules" [IEEE Trans. Automat. Control 48 (2003), no. 6, 988–1001; MR 1986266]. *IEEE Trans. Autom. Control* **48**(9), 1675.

[41] S. Martinand Marie-Therese Wolfram Jose A. Carrillo (2016) An improved version of the Hughes model for pedestrian flow. *Math. Models Methods Appl. Sci.* **26**(4), 671–697.

[42] KE, J., MINETT, J., AU, C.-P. & WANG, W.-Y. (2002) Self-organization and selection in the emergence of vocabulary. *Complexity* **7**, 41–54.

[43] KELLER, E. F. & SEGEL, L. A. (1970) Initiation of slime mold aggregation viewed as an instability. *J. Theor. Biol.* **26**(3), 399–415.

[44] KIRCHDOERFER, T. & ORTIZ, M. (2016) Data-driven computational mechanics. *Comput. Methods Appl. Mech. Engrg.* **304**, 81–101.

[45] KOCH, A. L. & WHITE, D. (1998) The social lifestyle of myxobacteria. *Bioessays* **20**, pp. 1030–1038.

[46] KRAUSE, U. (2000) A discrete nonlinear and non-autonomous model of consensus formation. In: *Communications in difference equations. Proceedings of the 4th International Conference on Difference Equations, Poznan, Poland, August 27–31, 1998*, Gordon and Breach Science Publishers, Amsterdam, pp. 227–236.

[47] KURAMOTO, Y. (2003) *Chemical Oscillations, Waves, and Turbulence*. Dover Books on Chemistry Series, Dover Publications, Mineola, New York.

[48] LASRY, J.-M. & LIONS, P.-L. (2007) Mean field games. *Japanese Journal of Mathematics (3)*, **2**(1), 229–260.

[49] LEONARD, N. E. & FIORELLI, E. (2001) Virtual leaders, artificial potentials and coordinated control of groups. In: *Proceedings of the 40th IEEE Conference on Decision and Control*, pp. 2968–2973.

[50] LU, F., MAGGIONI, M. & TANG, S. (2021) Learning interaction kernels in heterogeneous systems of agents from multiple trajectories. to appear *J. Mach. Learn. Res.* **22**(32), 1–67.

[51] LU, F., ZHONG, M., TANG, S. & MAGGIONI, M. (2019) Nonparametric inference of interaction laws in systems of agents from trajectory data. *Proc. Nat. Acad. Sci. U.S.A.* 116(29), 14424–14433.

[52] MOGENSEN, P. K. & RISETH, A. N. (2018) Optim: A mathematical optimization package for Julia. *J. Open Source Software* **3**(24), 615.

[53] MORANDOTTI, M. & SOLOMBRINO, F. (2020) Mean-field analysis of multipopulation dynamics with label switching. *SIAM J. Math. Anal.* **52**(2), 1427–1462.

[54] NIWA, H. S. (1994) Self-organizing dynamic model of fish schooling. *J. Theoret. Biol.* **171**, 123–136.

[55] NUORIAN, M., CAINES, P. E. & MALHAMÉ, R. P. (2010) Synthesis of Cucker-Smale type flocking via mean field stochastic control theory: Nash equilibria. In: *Proceedings of the 48th Allerton Conference on Communication, Control, and Computing, Monticello, Illinois, September 2010*, pp. 814–819.

[56] NUORIAN, M., CAINES, P. E. & MALHAMÉ, R. P. (2011) Mean field analysis of controlled Cucker-Smale type flocking: linear analysis and perturbation equations. In: *Proceedings of 18th IFAC World Congress Milano (Italy) August 28–September 2, 2011*, pp. 4471–4476.

[57] PARRISH, J. & EDELSTEIN-KESHET, L. (1999) Complexity, pattern, and evolutionary trade-offs in animal aggregation. *Science* **294**, 99–101.

[58] PARRISH, J. K., VISCIDO, S. V. & GRUENBAUM, D. (2002) Self-organized fish schools: an examination of emergent properties. *Biol. Bull.* **202**, 296–305.

[59] PELLEGRINI, S., ESS, A., SCHINDLER, K. & VAN GOOL, L. (2009) You'll never walk alone: modeling social behavior for multi-target tracking. In: *2009 IEEE 12th International Conference on Computer Vision*, pp. 261–268.

[60] PEREA, L., GÓMEZ, G. & ELOSEGUI, P. (2009) Extension of the Cucker-Smale control law to space flight formations. *AIAA J. Guidance Control Dyn.* **32**, 527–537.

[61] PERTHAME, B. (2007) *Transport Equations in Biology*, Birkhäuser, Basel.

[62] RAHMATI, Y., TALEBPOUR, A., MITTAL, A. & FISHELSON, J. (2020) Game theory-based framework for modeling human–vehicle interactions on the road. *Transp. Res. Rec.* **2674**(9), 701–713.

[63] ROMEY, W. L. (1996) Individual differences make a difference in the trajectories of simulated schools of fish. *Ecol. Model.* **92**, 65–77.

[64] SCHAEFFER, H., TRAN, G. & WARD, R. (2018) Extracting sparse high-dimensional dynamics from limited data. *SIAM J. Appl. Math.* **78**(6), 3279–3295.

[65] SCHAEFFER, H., TRAN, G. & WARD, R. Learning dynamical systems and bifurcation via group sparsity, arXiv:1709.01558.

[66] SCILLA, G. & SOLOMBRINO, F. (2018) Delayed loss of stability in singularly perturbed finite-dimensional gradient flows. *Asymptotic Anal.* **110**(1–2), 1–19.

[67] SCILLA, G. & SOLOMBRINO, F. (2018) Multiscale analysis of singularly perturbed finite dimensional gradient flows: the minimizing movement approach. *Nonlinearity* **31**(11), 5036–5074.

[68] SHORT, M. B., D'ORSOGNA, M. R., PASOUR, V. B., TITA, G. E., BRANTINGHAM, P. J., BERTOZZI, A. L. & CHAYES, L. B. (2008) A statistical model of criminal behavior. *Math. Models Methods Appl. Sci.* **18**(suppl.), 1249–1267.

[69] STELLATO, B., BANJAC, G., GOULART, P., BEMPORAD, A. & BOYD, S. (2020) OSQP: an operator splitting solver for quadratic programs. *Math. Program. Comput.* **12**, 637–672

[70] SUGAWARA, K. & SANO, M. (1997) Cooperative acceleration of task performance: foraging behavior of interacting multi-robots system. *Physica D* **100**, 343–354.

[71] TONER, J. & TU, Y. (1995) Long-range order in a two-dimensional dynamical xy model: how birds fly together. *Phys. Rev. Lett.* **75**, 4326–4329.

[72] TRAN, G. & WARD, R. (2017) Exact recovery of chaotic systems from highly corrupted data. *Multiscale Model. Simul.* **15**(3), 1108–1129.

[73] TSYBAKOV, A. B. (2009) *Introduction to Nonparametric Estimation*, Springer, New York.

[74] VICSEK, T., CZIROK, A., BEN-JACOB, E., COHEN, I. & SHOCHET, O. (1995) Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.* **75**, 1226–1229.

[75] VICSEK, T. & ZAFEIRIS, A. (2012) Collective motion. *Phys. Rep.* **517**, 71–140.

[76] YANG, D., LI, L., REDMILL, K. A. & ÖZGÜNER, Ü. Top-view trajectories: a pedestrian dataset of vehicle-crowd interaction from controlled experiments and crowded campus. In: 2019 *IEEE Intelligent Vehicles Symposium (IV)*, pp. 899–904, 2019.

[77] YATES, C., ERBAN, R., ESCUDERO, C., COUZIN, L., BUHL, J., KEVREKIDIS, L., MAINI, P. & SUMPTER, D. (2009) Inherent noise can facilitate coherence in collective swarm motion. *Proc. Nat. Acad. Sci.* 106, 5464–5469.

[78] ZHONG, M., MILLER, J. & MAGGIONI, M. (2020) Data-driven discovery of emergent behaviors in collective dynamics. *Physica D Nonlinear Phenomena* **411**, 132542.

A. Mathematical analysis

A.1 Well-posedness for entropic games

In this Section, we prove the key technical ingredients involved in the proof of Theorem 2.6 in Section 2.1, which establishes the well-posedness of the modified entropic dynamics given by (2.2). The two subsequent Lemmas establish that $f^{J,\varepsilon} = f^J + f^\varepsilon$, (2.8), is Lipschitz continuous in the chosen setting for strategies in $\mathcal{S}_{a,b}$.

Lemma A.1 (f^ε is L^p_η-Lipschitz continuous on $\mathcal{S}_{a,b}$). *For $\sigma_1, \sigma_2 \in \mathcal{S}_{a,b}$ and f^ε defined in (2.4), we have*

$$\|f^\varepsilon(\sigma_1) - f^\varepsilon(\sigma_2)\|_{L^p_\eta(U)} \le L \cdot \|\sigma_1 - \sigma_2\|_{L^p_\eta(U)} \tag{A1}$$

with $L = L(a, b)$.

Proof. Define $G(x) := x \log x$, let $L_{a,b}$ be its Lipschitz constant and $M_{a,b}$ its maximal absolute value over the interval $[a, b]$. For any $\sigma_1, \sigma_2 \in \mathcal{S}_{a,b}$, L^p_η-Lipschitz continuity can be proved via direct computations:

$$\|f^\varepsilon(\sigma_2) - f^\varepsilon(\sigma_1)\|_{L^p_\eta(U)}$$

$$= \left\| -G(\sigma_2(\cdot)) + G(\sigma_1(\cdot)) + \sigma_2(\cdot) \int_U G(\sigma_2(v)) \, d\eta(v) - \sigma_1(\cdot) \int_U G(\sigma_1(v)) \, d\eta(v) \right\|_{L^p_\eta(U)}$$

$$\le \|G(\sigma_2(\cdot)) - G(\sigma_1(\cdot))\|_{L^p_\eta(U)} + \left\| \sigma_1(\cdot) \int_U |G(\sigma_2(v)) - G(\sigma_1(v))| \, d\eta(v) \right\|_{L^p_\eta(U)}$$

$$+ \left\| |\sigma_2(\cdot) - \sigma_1(\cdot)| \cdot \int_U G(\sigma_2(v)) \, d\eta(v) \right\|_{L^p_\eta(U)}$$

$$\le (L_{a,b} + b \cdot L_{a,b} + M_{a,b}) \cdot \|\sigma_2 - \sigma_1\|_{L^p_\eta(U)}$$

where we used $\sigma_1(\cdot) \le b$ and Jensen's inequality in the second term. $\qquad \square$

Lemma A.2 (f^J is Lipschitz continuous). *For $J \in \mathcal{X}$, let f^J be defined as in (2.3). Then, f^J is Lipschitz continuous in all four arguments, that is, given $x_i, x_i' \in \mathbb{R}^d$ and $\sigma_i, \sigma_i' \in \mathcal{S}_{a,b}$, $i = 1, 2$, we have*

$$\|f^J(x_2, \sigma_2, x_2', \sigma_2') - f^J(x_1, \sigma_1, x_1', \sigma_1')\|_{L^p_\eta(U)} \le L \cdot \left[\|x_2 - x_1\| + \|\sigma_2 - \sigma_1\|_{L^p_\eta(U)} \right.$$
$$\left. + \|x_2' - x_1'\| + \|\sigma_2' - \sigma_1'\|_{L^p_\eta(U)} \right]$$

with $L = L(J, b)$.

Proof. Define the payoff for the single interaction $j \colon \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{S}_{a,b} \times U \to \mathbb{R}$ as

$$j(x, x', \sigma', u) := \int_U J(x, u, x', u')\, \sigma'(u') d\eta(u')$$

and observe that, for any $x, x' \in \mathbb{R}^d$, $u \in U$ and $\sigma_1', \sigma_2' \in \mathcal{S}_{a,b}$, we have

$$|j(x, x', \sigma_2', u) - j(x, x', \sigma_1', u)| = \left| \int_U J(x, u, x', u')\, (\sigma_2'(u') - \sigma_1'(u'))\, d\eta(u') \right|$$
$$\le \|J\|_\infty \cdot \|\sigma_2' - \sigma_1'\|_{L^p_\eta(U)}. \tag{A2}$$

The function f^J introduced in (2.3) can then be written as

$$f^J(x, \sigma, x', \sigma') = \left(j(x, x', \sigma', \cdot) - \int_U j(x, x', \sigma', v)\, \sigma(v)\, d\eta(v) \right) \sigma.$$

The Lipschitz continuity in x and x' is clear: it follows by Lipschitz continuity of j in x, x', descending directly from Lipschitz continuity of J. It remains to investigate the dependence on σ, σ'. For $\sigma_1, \sigma_2, \sigma' \in \mathcal{S}_{a,b}$, direct estimates provide

$$\|f^J(x, \sigma_2, x', \sigma') - f^J(x, \sigma_1, x', \sigma')\|_{L^p_\eta(U)} = \left\| j(x, x', \sigma', \cdot)\, (\sigma_2(\cdot) - \sigma_1(\cdot)) \right.$$

$$\left. - \sigma_2(\cdot) \int_U j(x, x', \sigma', v)\, \sigma_2(v)\, d\eta(v) + \sigma_1(\cdot) \int_U j(x, x', \sigma', v)\, \sigma_1(v)\, d\eta(v) \right\|_{L^p_\eta(U)}$$

$$\le \left\| |j(x, x', \sigma', \cdot)| \cdot |\sigma_2(\cdot) - \sigma_1(\cdot)| \right\|_{L^p_\eta(U)} + \left\| \sigma_1(\cdot) \cdot \int_U |j(x, x', \sigma', v)| \cdot |\sigma_2(v) - \sigma_1(v)|\, d\eta(v) \right\|_{L^p_\eta(U)}$$

$$+ \left\| |\sigma_2(\cdot) - \sigma_1(\cdot)| \cdot \int_U |j(x, x', \sigma', v)| \cdot \sigma_2(v)\, d\eta(v) \right\|_{L^p_\eta(U)}$$

$$\le (\|J\|_\infty + b \cdot \|J\|_\infty + \|J\|_\infty) \cdot \|\sigma_2 - \sigma_1\|_{L^p_\eta(U)},$$

where we used $j(\cdot, \cdot, \cdot, \cdot) \le \|J\|_\infty$, $\sigma(\cdot) \le b$ and Jensen's inequality. Similarly, for $\sigma, \sigma_1', \sigma_2' \in \mathcal{S}_{a,b}$, taking into account (A2), we have

$$\|f^J\left(x, \sigma, x', \sigma'_2\right) - f^J\left(x, \sigma, x', \sigma'_1\right)\|_{L^p_\eta(U)}$$

$$= \left\|\left[\left[j\left(x, x', \sigma'_2, \cdot\right) - j\left(x, x', \sigma'_1, \cdot\right)\right] - \int_U \left[j\left(x, x', \sigma'_2, v\right) - j\left(x, x', \sigma'_1, v\right)\right]\sigma(v)\,d\eta(v)\right]\sigma(\cdot)\right\|_{L^p_\eta(U)}$$

$$\leq \left(b \cdot \|J\|_\infty + b^2 \cdot \|J\|_\infty\right) \cdot \|\sigma'_2 - \sigma'_1\|_{L^p_\eta(U)}$$

which concludes the proof. $\qquad\square$

Remark A.3 (Entropic f^ε is not BL-Lipschitz continuous). *In [6], for the study of the original un-regularised system (1.3), the authors work in the space $Y = \mathbb{R}^d \times F(U)$, where*

$$F(U) := \overline{\mathrm{span}(\mathcal{M}_1(U))}^{\|\cdot\|_{BL}}$$

with the bounded Lipschitz norm defined as

$$\|\sigma\|_{BL} := \sup\left\{\int_U \varphi(u)\mathrm{d}\sigma(u) \mid \varphi \in C(U), \|\varphi\|_\infty + \mathrm{Lip}(\varphi) \leq 1\right\}. \tag{A3}$$

Large parts of the analysis in [6] are based on this bounded Lipschitz norm. However, the function f^ε is not Lipschitz continuous with respect to this norm. For $0 < a \ll 1 \ll b$, let $U = [0, 1]$ and η the uniform Lebesgue measure over U. For $n \geq 2$ we split $[0,1]$ into 2^n intervals and define the nodes $x_i^n = i/2^n$, for $0 \leq i \leq 2^n$. Define $s^n : (0, 1) \to (0, 2)$ as

$$s^n(x) = \begin{cases} 2 - a & \text{if } x \in \left(x_i^n, x_{i+1}^n\right), i \text{ even} \\ a & \text{if } x \in \left(x_i^n, x_{i+1}^n\right), i \text{ odd}. \end{cases}$$

This function alternates between $2 - a$ and a over each sub-interval. Set now

$$\sigma_1^n(x) = \begin{cases} 1 & \text{if } x \in (0, 1/2) \\ s^n(x) & \text{if } x \in (1/2, 1) \end{cases} \quad \text{and} \quad \sigma_2^n(x) = \begin{cases} s^n(x) & \text{if } x \in (0, 1/2) \\ 1 & \text{if } x \in (1/2, 1) \end{cases}$$

On the one hand, we find that $\|\sigma_2^n - \sigma_1^n\|_{BL} \leq W_1(\sigma_2^n, \sigma_1^n) \to 0$ as $n \to \infty$. On the other hand, observe that

$$\int_0^1 \sigma_1^n(x) \log\left(\sigma_1^n(x)\right)dx = \int_0^1 \sigma_2^n(x) \log\left(\sigma_2^n(x)\right)dx = \frac{1}{4}\left(a\log(a) + (2 - a)\log(2 - a)\right) =: M$$

so that, for $\varphi(x) = x/2$ (which is admissible in (A3)), one has

$$\frac{1}{\varepsilon}\|f^\varepsilon\left(\sigma_2^n\right) - f^\varepsilon\left(\sigma_1^n\right)\|_{BL} \geq -\int_0^{1/2} \frac{x}{2} \cdot s^n(x)\log(s^n(x))dx + \int_{1/2}^1 \frac{x}{2} \cdot s^n(x)\log\left(s^n(x)\right)dx$$

$$+ M\underbrace{\int_0^1 \frac{x}{2} \cdot \left(\sigma_2^n(x) - \sigma_1^n(x)\right)dx}_{\to 0 \text{ as } n \to \infty} = \frac{1}{4}\int_0^{1/2} s^n(x)\log\left(s^n(x)\right)dx + o(1) = \frac{M}{4} + o(1).$$

Hence, f^ε is not Lipschitz continuous with respect to the bounded Lipschitz norm.

We are left with the proof of the compatibility Lemma 2.5, which we briefly restate for the reader's convenience.

Lemma A.4 (Compatibility condition). *For $J \in \mathcal{X}$ and $\varepsilon > 0$, let f^J and f^ε be defined as in (2.3) and (2.4). Then, there exist a,b with $0 < a < 1 < b < \infty$ such that for any $(x,\sigma),(x',\sigma') \in \mathbb{R}^d \times \mathcal{S}_{a,b}$ there exists some $\theta > 0$ such that*

$$\sigma + \theta\lambda \left[f^J(x,\sigma,x',\sigma') + f^\varepsilon(\sigma) \right] \in \mathcal{S}_{a,b}. \tag{A4}$$

Proof. Fix any $(x,\sigma),(x',\sigma') \in (\mathbb{R}^d \times \mathcal{S}_{a,b})^2$. Taking into account that $\|\sigma\|_{L^1_\eta(U)} = 1$, one obtains

$$\int_U \left[\sigma + \theta\lambda \left(f^J(x,\sigma,x',\sigma') + f^\varepsilon(\sigma) \right) \right] d\eta = 1$$

for any $\theta > 0$. We are left with proving that, for suitable $0 < a \ll 1 \ll b$, there exists $\theta > 0$ such that $a \leq \sigma + \theta\lambda(f^J(x,\sigma,x',\sigma') + f^\varepsilon(\sigma)) \leq b$. On the one hand, thanks to global boundedness of J, one has

$$-2\|J\|_\infty \leq f^J(x,u,x',\sigma') \leq 2\|J\|_\infty.$$

On the other hand, writing any $\sigma \in \mathcal{S}_{a,b}$ as $\sigma(u) = \xi(u)a + (1 - \xi(u))b$, with $0 \leq \xi \leq 1$, and using convexity of $x \mapsto x \log x$, one has

$$\int_U \sigma(v) \log(\sigma(v)) \, d\eta(v) \leq \int_U \xi(u)a \log a \, d\eta(u) + \int_U (1 - \xi(u))b \log b \, d\eta(u).$$

The integral of ξ can be explicitly computed (recall $\|\sigma\|_{L^1_\eta(U)} = 1$), so that

$$0 \leq \int_U \sigma(v) \log(\sigma(v)) \, d\eta(v) \leq \frac{a(b-1)}{b-a} \log(a) + \frac{b(1-a)}{b-a} \log(b) =: K_{a,b}, \tag{A5}$$

where non-negativity follows by Jensen's inequality. The proof boils then down to find suitable a, b such that for any given $\sigma \in \mathcal{S}_{a,b}$ there exists $\theta > 0$ such that

$$a \leq \sigma(u)(1 - \theta\lambda \left[2\|J\|_\infty + \varepsilon \log(\sigma(u)) \right]) \tag{L}$$

$$\sigma(u) \left(1 + \theta\lambda \left[2\|J\|_\infty + \varepsilon(K_{a,b} - \log(\sigma(u))) \right] \right) \leq b \tag{U}$$

for every $u \in U$. Select a, b in the following way:

$$0 < a < \exp\left(-\frac{2\|J\|_\infty}{\varepsilon} \right) = L^\ell_\varepsilon < L^u_\varepsilon = \exp\left(\frac{2\|J\|_\infty}{\varepsilon} + K_{a,b} \right) < b. \tag{A6}$$

The existence of b satisfying this requirement follows by a direct one-dimensional asymptotic analysis. We distinguish three regions:

- $a \leq \sigma(u) < L^\ell_\varepsilon$: the lower bound (L) is satisfied for any $\theta > 0$ because $[2\|J\|_\infty + \varepsilon \log(\sigma(u))] < 0$ in the given region, while the upper bound (U) follows taking

$$\theta \leq \frac{b - L^\ell_\varepsilon}{\lambda L^\ell_\varepsilon(2\|J\|_\infty + \varepsilon \log(b/a))}$$

- $L_\varepsilon^\ell \le \sigma(u) \le L_\varepsilon^u$: a direct estimate on (L) and (U) leads to

$$\theta \le \min \left(\frac{L_\varepsilon^\ell - a}{\lambda L_\varepsilon^\ell (2\|J\|_\infty + \varepsilon \log b)}, \frac{b - L_\varepsilon^u}{\lambda L_\varepsilon^u (2\|J\|_\infty + \varepsilon \log(b/a))} \right)$$

- $L_\varepsilon^u < \sigma(u) \le b$: the lower bound (L) is satisfied for

$$\theta \le \frac{L_\varepsilon^u - a}{\lambda L_\varepsilon^u (2\|J\|_\infty + \varepsilon \log b)}$$

while upper bound (U) is satisfied for any $\theta > 0$ because $\left[2\|J\|_\infty + \varepsilon(K_{a,b} - \log(\sigma(u))) \right] < 0$ in the given region. □

The sought-after θ can then be selected to be the smallest between the previous upper bounds.

A.2 Convergence to fast-reaction limit in undisclosed setting

In this Section, we prove Theorem 2.14, which is concerned with the convergence to the undisclosed fast-reaction limit as $\lambda \to \infty$.

Proof of Theorem 2.14.
Part 1: $\sigma_i(t)$ close to $\sigma_i^J(x(t))$. For $i \in \{1, \ldots, N\}$ the dynamics of σ_i are described by

$$\partial_t \sigma_i(t)(u) = \lambda \cdot \left(- g_{i,u}'(t, \sigma_i(t)(u)) + \underbrace{\int_U g_{i,u}'(\sigma_i(t)(u)) \, \sigma_i(t)(u) \, d\eta(u)}_{:= m_i(t)} \right) \cdot \sigma_i(t)(u) \qquad \text{(A7)}$$

$$\text{where} \quad g_{i,u}(t, s) := -\frac{1}{N} \sum_{j=1}^N J(x_i(t), u, x_j(t)) s + \varepsilon \ [s \ \log(s) - s + 1] \qquad \text{(A8)}$$

and $g_{i,u}'$ denotes the first derivative with respect to s. This is formally the gradient flow (scaled by λ) of the (explicitly time-dependent) energy

$$G_i(t, \sigma) := \int_U g_{i,u}(t, \sigma(u)) \, d\eta(u) \qquad \text{(A9)}$$

with respect to the (spherical) Hellinger–Kakutani distance over $\mathcal{S}_{a,b}$. (This relation only serves as intuition for our proof and is never used in a mathematical argument.)

Let a, b such that $0 < a < b < \infty$. Then, there exists a $c \in (0, \infty)$ such that $g_{i,u}''(t, s) \ge c$ for $t \in [0, \infty)$, $s \in [a, b]$, $u \in U$ and $i \in \{1, \ldots, N\}$ (where analogous to above $g_{i,i}''(t, s)$ refers to the second derivative with respect to s). This implies that $G_i(t, \cdot)$ is strongly convex on $\mathcal{S}_{a,b}$.

Consider now the fast-reaction mixed strategies $\sigma^* = (\sigma_i^*)_{i=1}^N$ associated with locations $\mathbf{x} = (x_i)_{i=1}^N$, (2.15), (i.e. $\sigma_i^*(t) := \sigma_i^J(\mathbf{x}(t))$) where we introduce a normalisation constant $\mathcal{A}_i$:

$$\sigma_i^*(t)(u) = \mathcal{A}_i(t) \cdot \exp \left(\frac{1}{N\,\varepsilon} \sum_{j=1}^N J(x_i(t), u, x_j(t)) \right), \qquad \text{(A10)}$$

$$\mathcal{A}_i(t)^{-1} = \int_U \exp\left(\frac{1}{N\,\varepsilon} \sum_{j=1}^{N} J(x_i(t), u, x_j(t))\right) d\eta(u). \tag{A11}$$

For these we find

$$g'_{i,u}(t, \sigma_i^*(t)(u)) = -\frac{1}{N} \sum_{j=1}^{N} J(x_i(t), u, x_j(t)) + \varepsilon\,\log(\sigma_i^*(t)(u)) = \varepsilon\,\log(\mathcal{A}_i(t)). \tag{A12}$$

Using convexity of the functions $g_{i,u}(t, \cdot)$ we find for arbitrary $\sigma \in \mathcal{S}_{a,b}$:

$$G_i(t, \sigma) \geq \int_U \Big[g_{i,u}(t, \sigma_i^*(t)(u)) + \underbrace{g'_{i,u}(t, \sigma_i^*(t)(u))}_{=\varepsilon\,\log(\mathcal{A}_i(t))} \cdot (\sigma(u) - \sigma_i^*(t)(u)) \Big] d\eta(u) = G_i(t, \sigma_i^*(t)), \tag{A13}$$

where the second term integrates to zero, since $\varepsilon\,\log(\mathcal{A}_i(t))$ is constant with respect to $u \in U$ and both σ and $\sigma_i^*(t)$ are normalised. Therefore, $\sigma_i^*(t)$ is a minimiser of $G_i(t, \cdot)$ over $\mathcal{S}_{a,b}$, and since $G_i(t, \cdot)$ is strongly convex, it is the unique minimiser.

Using the bound $c \leq g''_{i,u}(t, s)$ for $s \in [a, b]$, we find for $s^* \in [a, b]$

$$g'_{i,u}(t, s^*) \cdot (s - s^*) + \frac{c}{2}(s - s^*)^2 \leq g_{i,u}(t, s) - g_{i,u}(t, s^*) \tag{A14}$$

and applying this bound to $G_i(t, \cdot)$, using $\sigma_i^*(t) \in \mathcal{S}_{a,b}$ due to (A6), analogous to (A13) one obtains for any $\sigma \in \mathcal{S}_{a,b}$ that

$$\frac{c}{2}\|\sigma - \sigma_i^*(t)\|_{L^2_\eta(U)}^2 \leq G_i(t, \sigma) - G_i(t, \sigma_i^*(t)). \tag{A15}$$

Strong convexity of $G_i(t, \cdot)$ provides another useful estimate (where m_i is defined in (A7)):

$$\|g'_{i,u}(t, \sigma_i(t)) - m_i(t)\|_{L^2_\eta(U)} \cdot \|\sigma_i(t) - \sigma_i^*(t)\|_{L^2_\eta(U)}$$
$$\geq \int_U \big[g'_{i,u}(t, \sigma_i(t)(u)) - m_i(t) \big] \cdot \big[\sigma_i(t)(u) - \sigma_i^*(t)(u) \big] d\eta(u)$$
$$= \int_U g'_{i,u}(t, \sigma_i(t)(u)) \cdot \big[\sigma_i(t)(u) - \sigma_i^*(t)(u) \big] d\eta(u) \geq G_i(t, \sigma_i(t)) - G_i(t, \sigma_i^*(t)),$$

where we used the Cauchy-Schwarz inequality. With (A15), this implies

$$\|g'_{i,u}(t, \sigma_i(t)) - m_i(t)\|_{L^2_\eta(U)}^2 \geq \frac{c}{2}[G_i(t, \sigma_i(t)) - G_i(t, \sigma_i^*(t))]. \tag{A16}$$

We need two additional estimates to control the explicit time-dependency of G_i. Note that $\|\partial_t x_i(t)\| \leq \|e\|_\infty$ and thus $|\partial_t g_{i,u}(t, s)| \leq 2 \cdot \mathrm{Lip}(J) \cdot \|e\|_\infty$. ($J$ may not be differentiable, but since $t \mapsto x_i(t)$ and J are Lipschitz, so is $g_{i,u}$, and thus, this derivative exists for almost all t.) This implies $|\partial_t G_i(t, \sigma)| \leq 2 \cdot \mathrm{Lip}(J) \cdot \|e\|_\infty$. Next, an explicit computation yields

$$G_i(t, \sigma_i^*(t)) = \varepsilon \cdot \log \mathcal{A}_i(t) = -\varepsilon \log\left(\int_U \exp\left(\frac{1}{N\,\varepsilon} \sum_{j=1}^{N} J(x_i(t), u, x_j(t))\right) d\eta(u)\right)$$

from which we deduce that $\left|\dfrac{\mathrm{d}}{\mathrm{d}t} G_i(t, \sigma_i^*(t))\right| \leq 2 \cdot \mathrm{Lip}(J) \cdot \|e\|_\infty$.

Summarising now, we find for almost every $t \in [0, \infty)$:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \left[G_i(t, \sigma_i(t)) - G_i(t, \sigma_i^*(t)) \right] &= \int_U g_{i,u}'(t, \sigma_i(t)(u)) \cdot \partial_t \sigma_i(t)(u) \, \mathrm{d}\eta(u) \\
&\quad + (\partial_t G_i)(t, \sigma_i(t)) - \frac{\mathrm{d}}{\mathrm{d}t} G_i(t, \sigma_i^*(t)) \\
&= \int_U [g_{i,u}'(t, \sigma_i(t)(u)) - m_i(t)] \cdot \partial_t \sigma_i(t)(u) \, \mathrm{d}\eta(u) \\
&\quad + (\partial_t G_i)(t, \sigma_i(t)) - \frac{\mathrm{d}}{\mathrm{d}t} G_i(t, \sigma_i^*(t)) \\
&= -\lambda \int_U [g_{i,u}'(t, \sigma_i(t)(u)) - m_i(t)]^2 \cdot \underbrace{\sigma_i(t)(u)}_{\geq a} \, \mathrm{d}\eta(u) \\
&\quad + (\partial_t G_i)(t, \sigma_i(t)) - \frac{\mathrm{d}}{\mathrm{d}t} G_i(t, \sigma_i^*(t)) \\
&\leq -\frac{\lambda \, a \, c}{2} \left[G_i(t, \sigma_i(t)) - G_i(t, \sigma_i^*(t)) \right] + 4 \cdot \mathrm{Lip}(J) \cdot \|e\|_\infty. \quad \text{(A17)}
\end{aligned}
$$

Applying Grönwall's lemma to $h(t) := G_i(t, \sigma_i(t)) - G_i(t, \sigma_i^*(t)) - \frac{8 \, \mathrm{Lip}(J) \, \|e\|_\infty}{\lambda \, a \, c}$ yields

$$
\begin{aligned}
G_i(t, \sigma_i(t)) - G_i(t, \sigma_i^*(t)) &\leq \frac{8 \, \mathrm{Lip}(J) \, \|e\|_\infty}{\lambda \, a \, c} \\
&\quad + \left[(G_i(0, \sigma_i(0)) - G_i(0, \sigma_i^*(0))) - \frac{8 \, \mathrm{Lip}(J) \, \|e\|_\infty}{\lambda \, a \, c} \right] \cdot \exp\left(-\frac{\lambda \, a \, c}{2} t \right). \quad \text{(A18)}
\end{aligned}
$$

Using $\sigma_i(0) \in \mathcal{S}_{a,b}$ and $x_i(0) \in B^d(R)$ for all $i \in \{1, \dots, N\}$, there exists some $C = C(a, b, R, \mathrm{Lip}(J), \|e\|_\infty) < \infty$ (not depending on N or i) such that

$$
G_i(t, \sigma_i(t)) - G_i(t, \sigma_i^*(t)) \leq C \left[\frac{1}{\lambda} + \exp\left(-\frac{\lambda \, t}{C} \right) \right]. \quad \text{(A19)}
$$

By virtue of (A15), an analogous bound with a different C (with the same dependency structure) holds for $\|\sigma_i(t) - \sigma_i^*(t)\|_{L_\eta^2(U)}$:

$$
\|\sigma_i(t) - \sigma_i^*(t)\|_{L_\eta^2(U)} \leq C \left[\frac{1}{\lambda} + \exp\left(-\frac{\lambda \, t}{C} \right) \right] \quad \text{(A20)}
$$

Part 2: x_i and σ_i close to x_i^{**} and σ_i^{**}. Recall now that the fast-reaction dynamics (2.15) for $\mathbf{x}^{**}$ are given by

$$
\partial_t x_i^{**}(t) = \int_U e(x_i^{**}(t), u) \, \sigma_i^{**}(t)(u) \, \mathrm{d}\eta(u) = v_i^J(x_1^{**}(t), \dots, x_N^{**}(t))
$$

where $\sigma_i^{**}(t) = \sigma_i^J(\mathbf{x}^{**}(t))$, (2.15a). The dynamics for $\mathbf{x}(t)$ follow

$$
\begin{aligned}
\partial_t x_i(t) &= \int_U e(x_i(t), u) \, \sigma_i(t)(u) \, \mathrm{d}\eta(u) = v_i^J(x_1(t), \dots, x_N(t)) \\
&\quad + \int_U e(x_i(t), u) \, [\sigma_i(t)(u) - \sigma_i^*(t)(u)] \, \mathrm{d}\eta(u).
\end{aligned}
$$

Introduce now the deviation measure $\delta(t) := \|\mathbf{x}(t) - \mathbf{x}^{**}(t)\|_N$ and obtain (*t*-almost everywhere, due to δ being Lipschitz)

$$\partial_t \delta(t) \le \|\partial_t \mathbf{x}(t) - \partial_t \mathbf{x}^{**}(t)\|_N$$

$$\le \|\mathbf{v}^J(\mathbf{x}(t)) - \mathbf{v}^J(\mathbf{x}^{**}(t))\|_N + \sum_{i=1}^N \frac{1}{N} \left\| \int_U e(x_i(t), u) \left[\sigma_i(t)(u) - \sigma_i^*(t)(u) \right] d\eta(u) \right\|$$

$$\le Lip\left(\mathbf{v}^J\right) \cdot \underbrace{\|\mathbf{x}(t) - \mathbf{x}^{**}(t)\|_N}_{=\delta(t)} + \|e\|_\infty \cdot \sum_{i=1}^N \frac{1}{N} \underbrace{\|\sigma_i(t) - \sigma_i^*(t)\|}_{\le C_{t,\lambda} := \left[C \left[\frac{1}{\lambda} + \exp\left(-\frac{\lambda\, t}{C} \right) \right] \right]^{1/2}} \tag{A21}$$

For every $\tau > 0$, we get $\delta(t) \le 2\,\tau\, \|e\|_\infty$ on $t \in [0, \tau]$ (due to $\|\partial_t x_i\|, \|\partial_t x_i^{**}\| \le \|e\|_\infty$) and for $t \in [\tau, \infty)$ from (A21) via Grönwall's lemma (with a change of variables as above)

$$\delta(t) \le \left(2\,\tau\, \|e\|_\infty + \frac{C_{\tau,\lambda}}{Lip\left(\mathbf{v}^J\right)} \right) \cdot \exp(Lip\left(\mathbf{v}^J\right) \cdot (t - \tau)) - \frac{C_{\tau,\lambda}}{Lip\left(\mathbf{v}^J\right)}. \tag{A22}$$

Setting now $\tau = \sqrt{\lambda}^{-1}$, there is a suitable C (depending on a, b, R, J, e, but not on N or i) such that

$$\delta(t) \le \frac{C}{\sqrt{\lambda}} \cdot \exp(t \cdot C). \tag{A23}$$

From Lemma 2.10, it follows that (cf. Lemma 2.11)

$$\|\sigma_i^*(t) - \sigma_i^{**}(t)\|_{L^2_\eta(U)} = \|\sigma_i^J(\mathbf{x}(t)) - \sigma_i^J(\mathbf{x}^{**}(t))\|_{L^2_\eta(U)} \le L \cdot \|\mathbf{x}(t) - \mathbf{x}^{**}(t)\|_N \tag{A24}$$

for some $L < \infty$ depending on J and ε. Combining (A20), (A23) and (A24), one obtains that (for different C with same dependency structure)

$$\|\sigma_i(t) - \sigma_i^{**}(t)\|_{L^2_\eta(U)} \le \frac{C}{\sqrt{\lambda}} \cdot \exp(t \cdot C) + \left[C \left[\frac{1}{\lambda} + \exp\left(-\frac{\lambda\, t}{C} \right) \right] \right]^{1/2}. \tag{A25}$$

Part 3: Mean-field setting. For $\Sigma_1, \Sigma_2 \in \mathcal{M}_1(\mathcal{C}_{a,b})$ one has

$$W_1(P_\sharp \Sigma_1, P_\sharp \Sigma_2) \le W_1(\Sigma_1, \Sigma_2) \tag{A26}$$

due to

$$\|x_1 - x_2\| = \|P(x_1, \sigma_1) - P(x_2, \sigma_2)\| \le \|x_1 - x_2\| + \|\sigma_1 - \sigma_2\|_{L^p_\eta(U)} = \|(x_1, \sigma_1) - (x_2, \sigma_2)\|_Y.$$

Let $(\bar{\Sigma}^N)_N$ be a sequence of initial empirical measures with $W_1(\bar{\Sigma}^N, \bar{\Sigma}) \to 0$ as $N \to \infty$ and let Σ^N be the corresponding solution of (2.11), as discussed in Theorem 2.6, point 1. Set $\bar{\mu}^N := P_\sharp \bar{\Sigma}^N$, which consequently satisfies $W_1(\bar{\mu}^N, \bar{\mu}) \to 0$ as $N \to \infty$ and let μ^N be the corresponding solution of (2.17). By Theorem 2.6, point 3 and Theorem 2.13, point 3 (both for $s = 0$), for any $\alpha > 0$ there is some N such that

$$W_1(\Sigma^N(t), \Sigma(t)) \le \alpha, \qquad\qquad W_1\left(\mu^N(t), \mu(t)\right) \le \alpha \tag{A27}$$

for $t \in [0, T]$.

The points in Σ^N evolve according to the full discrete model, the points in μ^N evolve according to the fast-reaction limit. Therefore, by virtue of (A23) one has

$$W_1\big(P_\sharp\Sigma^N(t),\mu^N(t)\big) \le \frac{C}{\sqrt{\lambda}}\cdot\exp(t\cdot C) \tag{A28}$$

with the same C as in (A23). Combining (A28), (A27) and (A26) one obtains

$$W_1\big(P_\sharp\Sigma(t),\mu(t)\big) \le W_1\big(P_\sharp\Sigma(t),P_\sharp\Sigma^N(t)\big) + W_1\big(P_\sharp\Sigma^N(t),\mu^N(t)\big) + W_1\big(\mu^N(t),\mu(t)\big)$$

$$\le 2\alpha + \frac{C}{\sqrt{\lambda}}\cdot\exp(t\cdot C).$$

The fact that $\alpha > 0$ is arbitrary, means the inequality also holds in the limit $\alpha \to 0$ and thus establishes (2.26).

A.3 Undisclosed fast-reaction inference functionals

We report in this section some technical proofs for statements in Section 3.3. We start with proving that observations generated through explicit simulations of the undisclosed fast-reaction system are indeed admissible in the sense of Assumption 3.1.

Proof of Lemma 3.2.
The first statement is just point 1 in Theorem 2.13. The proof of the second statement is a direct by-product of Theorem 2.13 combined with Lemma 2.10. Let μ^N and μ be the unique solutions of (2.17) for initial conditions $\bar\mu^N$ and $\bar\mu$, as provided by Theorem 2.13, point 1. By the stability estimate in Theorem 2.13, point 3, one has for any $t \in [0, T]$

$$W_1\big(\mu^N(t),\mu(t)\big) \le C\, W_1\big(\bar\mu^N,\bar\mu\big) \to 0 \text{ as } N \to \infty. \tag{A29}$$

Hence, we claim $\mu^\infty = \mu$ and in particular $v^\infty = v^J(\mu)$ and $\sigma^\infty = \sigma^J(\mu)$ (as defined in (2.18) and (2.19)). Regarding convergence of velocities, observe that by definition

$$v_i^N(t) = v_i^J\big(x_1^N(t),\dots,x_N^N(t)\big) = v^J\big(\mu^N(t)\big)\big(x_i^N(t)\big)$$

while (2.21) in Lemma 2.10 provides the uniform bound

$$\big\|v^J\big(\mu^N(t)\big)(x) - v^J(\mu(t))(x)\big\| \le C\, W_1\big(\mu^N(t),\mu(t)\big).$$

This yields directly (3.3) while (3.4) follows thanks to the uniform bound $\|v^J(\mu(t))(x)\| \le \|e\|_\infty$. An analogous argument applies to prove (3.5) and (3.6): first recall that

$$\sigma_i^N(t) = \sigma_i^J\big(x_1^N(t),\dots,x_N^N(t)\big) = \sigma^J\big(\mu^N(t)\big)\big(x_i^N(t),\cdot\big)$$

and thanks to (2.20) in Lemma 2.10 one has the uniform bounds

$$\big|\sigma^J\big(\mu^N(t)\big)(x,u) - \sigma^J(\mu(t))(x,u)\big| \le C\, W_1\big(\mu^N(t),\mu(t)\big)$$

and $1/C < \sigma^J(\mu^N)(x,u) < C$, which immediately provides the conclusions when combined with uniform continuity of $s \mapsto s\log(s)$ over $[1/C, C]$.

Next, following the pipeline provided by [10, Proposition 1.1], we show that the obtained minimal inference functional value provides an upper bound on the accuracy of the trajectories that are simulated with the inferred $\hat{J}$.

Proof of Theorem 3.3.

Part 1: discrete velocity-error functional (3.9a). Let $\mu^N(t) = \frac{1}{N} \sum_{i=1}^N \delta_{x_i^N(t)}$. For every $t \in [0, T]$, we estimate

$$
\begin{aligned}
\|\mathbf{x}^N(t) - \hat{\mathbf{x}}^N(t)\|_N &= \left\| \int_0^t (\partial_t \mathbf{x}^N(s) - \partial_t \hat{\mathbf{x}}^N(s)) \, \mathrm{d}s \right\|_N \leq \int_0^t \left\| \partial_t \mathbf{x}^N(s) - \partial_t \hat{\mathbf{x}}^N(s) \right\|_N \, \mathrm{d}s \\
&= \int_0^t \frac{1}{N} \sum_{i=1}^N \left\| v_i^N(s) - v_i^{\hat{J}}(\hat{\mathbf{x}}^N(s)) \right\| \, \mathrm{d}s \\
&\leq \int_0^t \left[\frac{1}{N} \sum_{i=1}^N \left\| v_i^N(s) - v_i^{\hat{J}}(\mathbf{x}^N(s)) \right\| + \frac{1}{N} \sum_{i=1}^N \left\| v_i^{\hat{J}}(\mathbf{x}^N(s)) - v_i^{\hat{J}}(\hat{\mathbf{x}}^N(s)) \right\| \right] \mathrm{d}s \\
&\leq T \left(\frac{1}{TN} \int_0^T \sum_{i=1}^N \left\| v_i^N(s) - v_i^{\hat{J}}(\mathbf{x}^N(s)) \right\|^2 \, \mathrm{d}s \right)^{1/2} + L \int_0^t \|\mathbf{x}^N(s) - \hat{\mathbf{x}}^N(s)\|_N \, \mathrm{d}s \\
&\leq T \sqrt{\mathcal{E}_v^N(\hat{J})} + L \int_0^t \|\mathbf{x}^N(s) - \hat{\mathbf{x}}^N(s)\|_N \, \mathrm{d}s
\end{aligned}
$$

where we used Jensen's inequality (or Cauchy-Schwarz) from the third to the fourth line and with $L = L(\hat{J}, e, \varepsilon)$ denoting the Lipschitz constant of the velocities (cf. Lemma 2.11). An application of Grönwall's inequality yields

$$
\|\mathbf{x}^N(t) - \hat{\mathbf{x}}^N(t)\|_N \leq T e^{LT} \cdot \sqrt{\mathcal{E}_v^N(\hat{J})}. \tag{A30}
$$

Part 2: discrete σ-error functional (3.10a). To obtain the same estimate for the discrete σ-error functional (3.10a) we first recall Pinsker's inequality (cf. [73, Lemma 2.5]): for $\sigma_1, \sigma_2 \in \mathcal{S}_{a,b}$ one has

$$
\|\sigma_1 - \sigma_2\|_{L_\eta^1(U)} \leq \sqrt{2\mathrm{KL}(\sigma_1 \mid \sigma_2)}.
$$

Thus, for any $t \in [0, T]$, we estimate

$$
\begin{aligned}
\left\| v_i^N(t) - v_i^{\hat{J}}(\mathbf{x}^N(t)) \right\|^2 &= \left\| \int_U e(x_i^N(t), u) \left(\sigma_i^N(t)(u) - \sigma_i^{\hat{J}}(\mathbf{x}^N(t))(u) \right) d\eta(u) \right\|^2 \\
&\leq \left(\int_U \| e(x_i^N(t), u) \| \left| \sigma_i^N(t)(u) - \sigma_i^{\hat{J}}(\mathbf{x}^N(t))(u) \right| d\eta(u) \right)^2 \\
&\leq \|e\|_\infty^2 \| \sigma_i^N(t) - \sigma_i^{\hat{J}}(\mathbf{x}^N(t)) \|_{L_\eta^1(U)}^2 \leq 2\|e\|_\infty^2 \mathrm{KL}\left(\sigma_i^N(t) \mid \sigma_i^{\hat{J}}(\mathbf{x}^N(t)) \right)
\end{aligned}
$$

This provides $\mathcal{E}_v^N(\hat{J}) \leq 2\|e\|_\infty^2 \mathcal{E}_\sigma^N(\hat{J})$, and the conclusion follows from (A90).

Part 3: extension to continuous functionals. Let $\left(\hat{\mathbf{x}}^N \right)_N$ be the sequence of discrete solutions with initial conditions $\left(\mathbf{x}^N(0) \right)_N$ induced by $\hat{J}$ and $(\hat{\mu}^N)_N$ the corresponding sequence of empirical

measures. By Assumption 3.1, point 2, we have $W_1\left(\mu^N(t), \mu^\infty(t)\right) \to 0$ uniformly for $t \in [0, T]$, in particular $W_1\left(\mu^N(0), \mu^\infty(0)\right) \to 0$. Hence, by Theorem 2.13, point 3, $W_1(\hat{\mu}^N(t), \hat{\mu}(t)) \le CW_1\left(\mu^N(0), \mu^\infty(0)\right) \to 0$ as $N \to \infty$ uniformly for $t \in [0, T]$. Then

$$W_1(\mu^\infty(t), \hat{\mu}(t)) = \lim_{N \to \infty} W_1(\mu^N(t), \hat{\mu}^N(t)) \le \liminf_{N \to \infty} \|\mathbf{x}^N(t) - \hat{\mathbf{x}}^N(t)\|_N$$

$$\le \liminf_{N \to \infty} C\sqrt{\mathcal{E}_v^N(\hat{J})} = C\sqrt{\mathcal{E}_v(\hat{J})}$$

where we applied the discrete result in the third step and used Lemma 3.5, for the constant sequence $(\hat{J})_N$, in the fourth step. The same argument covers the σ-functionals (3.10). $\qquad\square$

We close this section by proving convexity of the σ-based inference functionals $\mathcal{E}_\sigma^N$ and $\mathcal{E}_\sigma$.

Proposition A.5. (Convexity). *$\mathcal{E}_\sigma^N$ and $\mathcal{E}_\sigma$, (3.10), are convex on X.*

Proof. We fix a time $t \in [0, T]$ and ignore the t-dependency for notational simplicity. Also recall that all mixed strategy densities in $\mathcal{S}_{a,b}$ are strictly bounded away from zero, making quotients and logarithms in the following well-defined. For $\mathbf{x} = (x_1, \ldots, x_N) \in [\mathbb{R}^d]^N$ and any $i \in \{1, \ldots, N\}$, we compute

$$\mathrm{KL}\left(\sigma_i | \sigma_i^{\hat{J}}(x_1, \ldots, x_N)\right) = \int_U \log\left(\frac{\sigma_i(u)}{\sigma_i^{\hat{J}}(x_1, \ldots, x_N)(u)}\right) \sigma_i(u) \mathrm{d}\eta(u)$$

$$= \int_U \left[\sigma_i(u) \log \sigma_i(u) - \sigma_i(u) \log\left(\sigma_i^{\hat{J}}(x_1, \ldots, x_N)(u)\right)\right] \mathrm{d}\eta(u)$$

$$= \int_U \left[\sigma_i(u) \log \sigma_i(u) - \sigma_i(u) \log\left(\frac{\exp\left(\frac{1}{\varepsilon N} \sum_{j=1}^N \hat{J}(x_i, u, x_j)\right)}{\int_U \exp\left(\frac{1}{\varepsilon N} \sum_{j=1}^N \hat{J}(x_i, v, x_j)\right) \mathrm{d}\eta(v)}\right)\right] \mathrm{d}\eta(u)$$

$$= \log\left(\int_U \exp\left(\frac{1}{\varepsilon N} \sum_{j=1}^N \hat{J}(x_i, v, x_j)\right) \mathrm{d}\eta(v)\right) + \int_U \sigma_i(u)\left(\log \sigma_i(u) - \frac{1}{\varepsilon N} \sum_{j=1}^N \hat{J}(x_i, u, x_j)\right) \mathrm{d}\eta(u).$$

Convexity of the first term follows by Hölder's inequality: let $\hat{J}_1, \hat{J}_2 \in X$ and set $S_1(v) := \frac{1}{\varepsilon N} \sum_{j=1}^N \hat{J}_1(x_i, v, x_j)$ and $S_2(v) := \frac{1}{\varepsilon N} \sum_{j=1}^N \hat{J}_2(x_i, v, x_j)$, for $\alpha \in (0, 1)$ we have

$$\log\left(\int_U \exp\left(\alpha S_1(v) + (1-\alpha)S_2(v)\right) \mathrm{d}\eta(v)\right) = \log\left(\int_U \exp(\alpha S_1(v)) \cdot \exp((1-\alpha)S_2(v)) \, \mathrm{d}\eta(v)\right)$$

$$\le \log\left(\left(\int_U \exp(\alpha S_1(v))^{1/\alpha} \, \mathrm{d}\eta(v)\right)^\alpha \cdot \left(\int_U \exp((1-\alpha)S_2(v))^{1/(1-\alpha)} \, \mathrm{d}\eta(v)\right)^{1-\alpha}\right)$$

$$= \alpha \log\left(\int_U \exp(S_1(v)) \, \mathrm{d}\eta(v)\right) + (1-\alpha) \log\left(\int_U \exp(S_2(v)) \, \mathrm{d}\eta(v)\right).$$

The second term is constant, and the last term is linear in $\hat{J}$, and so $\mathrm{KL}\left(\sigma_i | \sigma_i^{\hat{J}}(x_1, \ldots, x_N)\right)$ is convex in $\hat{J}$. Summing over i (for (3.10a)) or integrating over $\mu^\infty(t)$ (for (3.10b)) and then integrating in time preserves convexity and concludes the proof. $\qquad\square$