



Mean Field Games

What Does a Nonlinear Fokker–Planck Equation Buy a Mean Field Game?

Frontiers, Tails, Capacity, Congestion and Herding in Crowds Whose Agents Touch

T. Zamrik

08-OCT-2026

Abstract. Mean field games almost always pair the agents’ Hamilton–Jacobi–Bellman equation with a Fokker–Planck equation linear in the density: the agents’ noises are independent, every interaction sits in the cost, and the agents never touch. This discussion paper asks what each kind of nonlinearity in the forward equation buys. In a stationary game in one dimension the crowd is an explicit statistics of the value function; the linear equation forces Boltzmann statistics, and without a crowd cost the game decouples. Noise, capacity and congestion each select another statistics and couple the game through the dynamics alone. Noise that grows with the density gives Tsallis statistics with edges and, past an exponent of 1.146, an empty gap; noise that falls with it gathers the crowd into one cluster until the branch folds at 0.219; a capacity gives a Fermi–Dirac plateau once the linear crowd would overflow the cap 1.7-fold; congestion gives the noise family in another game, whose optimum, 0.5 to 3 per cent better, needs a fee set by the crowd’s flow. A gradient pull buys nothing: it is a crowd cost in disguise. The open questions are where the computation sees what the theory cannot yet prove.

1. Introduction

A mean field game describes many agents, each solving a control problem whose data depend on the distribution of the others, in the limit where one agent is negligible [1], [2]; the theory is now in textbooks [3], [4], [5]. Almost all of it pairs a Hamilton–Jacobi–Bellman equation, solved backward by each agent, with a Fokker–Planck equation, solved forward by the crowd, that is linear in the density. Each agent’s noise is its own and its drift is its control, so the density enters its problem only through what it pays. The agents never touch: they can make a place expensive for one another, but cannot jostle, block, slow or pull one another, as the quadratic reference case of Guéant makes explicit [6]. This paper asks what that choice gives up.

Outside game theory the forward equation has long depended on the density. The linear Fokker–Planck equation is the Wasserstein gradient flow of potential plus entropy [7], the porous medium equation that of a power of the density [8], and the free energy names the



ways particles touch: an internal energy makes the noise depend on the density [14], an interaction energy adds a pull [9], and a mobility such as $m(1 - m)$ changes the metric so that particles exclude one another [10]. Santambrogio’s book gives the framework [19], and Frank’s collects the stationary states, generalised Boltzmann distributions of the potential [11]. In a game the potential becomes the value function. Of these mechanisms only congestion and nonlinear mobility have reached games, studied one at a time for well-posedness.

The paper makes six contributions. It shows that the stationary crowd is an explicit statistics of the value function, one per way of touching (Proposition 1, Table 1), and that only touching couples a game without a crowd cost (Corollary 1). It computes the signatures on one baseline, each checked independently: edges and an empty gap (Proposition 2, Figure 1), heavy tails and a fold (Figure 2), a plateau (Figure 3), and edges made by cost (Figure 4). It shows that noise and congestion share a statistics but are two games (Corollary 2), and that the congestion optimum is reached by a fee set by the crowd’s flow (Proposition 3). It shows that a gradient pull is a crowd cost in disguise (Propositions 4 and 5, Figure 5). And it orders the open questions by payoff.

2. The Stationary Game

Definition 1 (The stationary game and its five dials). On $\Omega = [-L, L]$, $L = 2.5$, with reflecting walls, the running cost is $V(x) = h(x^2 - 1)^2$ with $h = 1$, the noise level is $\nu = 0.1$ and the mass is $M = \int_{\Omega} m \, dx = 1$. An agent moves by $dX_t = (v_t + b(X_t)) \, dt + \sqrt{2a(X_t)} \, dW_t$ and minimises the long-run average of $\ell(v_t; m(X_t)) + V(X_t)$, whose minimum is the ergodic constant λ . A stationary game is (u, λ, m) with

$$\lambda - a u'' + H(x, u', m) = V, \quad (a m)'' - ((v^* + b) m)' = 0,$$

$m \geq 0$ and $\int_{\Omega} m \, dx = M$, where $H(x, p, m) = \sup_v \{-vp - \ell(v; m)\} - bp$ and v^* attains the supremum at $p = u'$; the constant μ of Proposition 1 is fixed by the mass. The linear game D0 has $a = \nu$, $\ell = v^2/2$, $b = 0$. Each dial is D0 at its neutral value: (D1) noise exponent $\gamma > 0$, $a = \nu m^{\gamma-1}$; (D2) capacity K , $\ell = v^2/(2(1 - m/K))$, set by the load $y = (\text{linear peak})/K \in [0, 3]$; (D3) congestion exponent $c \geq 0$, $\ell = (\varepsilon + m)^c v^2/2$ with $\varepsilon = 10^{-3}$; (D4) pull θ , the drift $b = -\theta(x - \bar{x})$ toward the mean $\bar{x} = M^{-1} \int_{\Omega} x m \, dx$; (D5) conformity cost κ , D0 with running cost $V + \kappa(x - \bar{x})^2/2$.

The wells are at $x = \pm 1$ and the barrier, of height h , at $x = 0$. In D2 and D3 the dial changes the forward equation through the velocity, $v^* = -(1 - m/K) u'$ and $v^* = -u' / (\varepsilon + m)^c$; D5 is a cost, not a way of touching, and is there for Proposition 5.

Assumption 1 (Standing assumptions). (i) One dimension: at the reflecting walls the flux $(a m)' - (v^* + b) m$ and u' vanish. (ii) u is Lipschitz, twice continuously differentiable where $a > 0$, and a viscosity solution where $a = 0$. (iii) $m \geq 0$ is continuous, $a m$ is continuously differentiable, the forward equation holds weakly, and $m < K$ in D2. (iv) The equilibrium is symmetric, m and u even, unless the crowd herds.

In one dimension a stationary flux is constant; in higher dimensions a crowd can circulate.



Proposition 1 (The crowd is a statistics of the value function). Under Assumption 1, on each connected component of the support of m , $m = G(u; \mu)$ for a constant μ , with

$$\text{D0: } m = e^{(\mu-u)/\nu}, \quad \text{D1: } m = \left[1 + \frac{(\gamma-1)(\mu-u)}{\nu\gamma}\right]_+^{1/(\gamma-1)}, \quad \text{D2: } m = \frac{K}{1 + e^{(u-\mu)/\nu}},$$

$$\text{D3: } \int_1^m \frac{(\varepsilon+s)^c}{s} ds = \frac{\mu-u}{\nu}, \quad \text{D4: } m = e^{(\mu-u-\theta(x-\bar{x})^2/2)/\nu},$$

and D5 as D0. D0 is Boltzmann; D1 is the Tsallis q -exponential, $q = 2 - \gamma$ [13], [12], which tends to D0 as $\gamma \rightarrow 1$; D2 is Fermi–Dirac; D3 at $\varepsilon = 0$ is $m = [1 + c(\mu - u)/\nu]_+^{1/c}$. Under Assumption 1(iv), μ is the same on every component.

Proof. The forward equation says that the flux $J = (am)' - (v^* + b)m$ has $J' = 0$, and the walls make $J = 0$. On a component of $\{m > 0\}$ divide by m and integrate. D0: $\nu m' = -u'm$. D1: $(\nu m^\gamma)' = -u'm$, that is $\frac{\nu\gamma}{\gamma-1}(m^{\gamma-1})' = -u'$. D2: $v^* = -(1 - m/K)u'$ gives $(\ln \frac{m}{K-m})' = -u'/\nu$. D3: $v^* = -u'/(\varepsilon + m)^c$ gives $(\varepsilon + m)^c m'/m = -u'/\nu$. D4: $\nu m' = -(u' + \theta(x - \bar{x}))m$. One constant per component, equal under symmetry [11]. $\square$

Each statistics minimises $\int u m dx$ plus an internal energy: the free-energy picture with the value function as the potential [11]. Substituted into the Hamilton–Jacobi–Bellman equation it leaves one equation for u , which Algorithm 1 solves. In D1 the noise is linear in the value: $a = \nu + (\gamma - 1)(\mu - u)/\gamma$ on the support.

Source. Proposition 1, with signatures from the computed equilibria of the following sections.

Columns. The dial; what an agent feels; the statistics; the signature in the crowd.

Table 1: Five ways for agents to touch, the statistics each selects, and its signature.

dial	the agent feels	statistics	signature
D0, never touch	its own noise	Boltzmann	positive everywhere; decoupled
D1, $\gamma > 1$	noise rising with density	Tsallis, $q < 1$	edges; empty gap past $\gamma \approx 1.146$
D1, $\gamma < 1$	noise falling with density	Tsallis, $q > 1$	heavy tails; one cluster; fold near $\gamma \approx 0.219$
D2, capacity	motion dear in a full place	Fermi–Dirac	plateau from load $y \approx 1.71$
D3, congestion	motion dear in company	Tsallis at $\varepsilon = 0$	edges from cost; gap from $c \approx 0.290$
D4, pull	drift toward the mean	Boltzmann, shifted	herding past $\theta_c \approx 0.0139$, as D5 with $\kappa = \theta^2$

D5, a conformity cost, has the statistics of D0 and is not a way of touching.



Corollary 1 (Without a crowd cost, only touching couples the game). (i) In D0 the Hamilton–Jacobi–Bellman equation contains no m : u and λ do not depend on M , and m is M times a fixed profile. (ii) In D1 to D3 it contains m , so u and λ depend on M in general; doubling the crowd is, in D1, multiplying ν by $2^{\gamma-1}$ and, in D2, halving K . (iii) In D4 and D5, m enters only through $\bar{x}$, so (u, λ, sm) is an equilibrium of mass sM whenever (u, λ, m) is one of mass M .

Proof. The substitution $u = -2\nu \ln \psi$ gives $-2\nu^2 \psi'' + V\psi = \lambda\psi$ with $\psi' = 0$ at the walls, whose principal eigenpair is unique; μ is then set by M . (ii) In D1, $a(sm) = s^{\gamma-1}a(m)$, so both equations for sm are those for m with $s^{\gamma-1}\nu$ in place of ν ; in D2 both see m only through m/K . (iii) $\bar{x}[sm] = \bar{x}[m]$, and Proposition 1 absorbs s into μ . $\square$

In each nonlinear dial the crowd’s size reaches the agent through the dynamics; the pull of D4 sees only where the mean is, and by Proposition 5 that is a crowd cost in disguise.

3. Computing the Crowds

Algorithm 1 — A stationary crowd for any dial

Inputs: a dial and its value; V, ν, M, L ; N cells; tolerances 10^{-10} (HJB residual) and 10^{-12} (mass). *Output:* $u, \lambda, \mu, m = G(u; \mu)$, the velocity $v^* + b$, and $\bar{x}$ in D4 and D5.

1. Unknowns: u at the nodes, spacing $\Delta x = 2L/N$; $\lambda; \mu; \bar{x}$ in D4 and D5. Fix $u(1) = 0$.
2. Set $m = G(u; \mu)$ by Proposition 1; for D3 invert at each node by a monotone Newton iteration on $\ln m$.
3. Discretise the HJB equation: central differences where the cell Peclet number $|\partial_p H| \Delta x / (2a)$ is at most $1/2$, Godunov upwinding where it is at least 1, a linear blend between. The scheme is monotone, second order where diffusion dominates, and upwind where $a = 0$.
4. Append the mass equation, the normalisation and, in D4 and D5, the equation for $\bar{x}$.
5. Newton on the whole system: a tridiagonal Jacobian with two or three dense borders, solved by a bordered (Schur-complement) banded solve, with backtracking.
6. Continue in the dial from D0, halving the step whenever Newton fails.

Cost. $O(N)$ per Newton step. The maps use $N = 1600$ at 61 dial values, with $N = 400$ and 800 for refinement; all maps and checks ran in 32 s on one CPU. The differences follow [31]; pure upwinding was first order and missed the Hopf–Cole check, built on the change of variables of [6], by about 30 times.

The linear crowd against its Hopf–Cole ground state. *Method.* Algorithm 1 for D0, against the ground state of $-2\nu^2 \psi'' + V\psi = \lambda\psi$, with $m \propto \psi^2$, from a spectral solver on 400 cosine modes; and $V = x^2/2$ on $L = 6$, where $\lambda = \nu$ exactly.

Parameters. Definition 1; $N = 400, 800, 1600$.



Result. $\lambda = 0.271685$, far below h ; the crowd peaks at 1.017 at $|x| \approx 0.96$. The tunnelling gap to the first excited level is 3.42×10^{-4} .

Error. The error in λ is 1.5×10^{-6} , 3.8×10^{-7} , 9.5×10^{-8} and the relative error in m is 1.1×10^{-4} , 2.6×10^{-5} , 6.6×10^{-6} : second order. The harmonic case agrees to rounding. This verifies Proposition 1 for D0.

The forward equation solved without the formula. *Method.* The velocity and noise of a computed equilibrium are fed to a finite-volume solve of the forward equation, with zero flux and mass M but without G , and the result is compared with $G(u; \mu)$.

Parameters. Twelve cases across D0 to D5, herded crowds included; $N = 400, 800, 1600$.

Result. The two crowds agree in every case, to rounding for D1 at $\gamma = 2$. Over all 1477 solves the mass is exact to 5.7×10^{-11} , and the capacity crowd never exceeds $0.99987 K$.

Error. The relative difference falls about fourfold per halving of the cell, to 1.6×10^{-5} for D0 and 6.5×10^{-5} in the worst case, D1 at $\gamma = 1.5$: discretisation error, not a failure of the formula. This verifies Proposition 1 for all six dials.

Doubling the crowd. *Method.* Algorithm 1 at $M = 1$ and $M = 2$, the dial fixed.

Parameters. D0; D1 at $\gamma = 1.75$ and 0.6 ; D2 at $y = 1.5$; D3 at $c = 0.5$; D4 and D5, symmetric and herded; $N = 1600$.

Result. In D1, λ rises from 0.2168 to 0.3096 at $\gamma = 1.75$ and falls from 0.3572 to 0.2585 at $\gamma = 0.6$, as the factor $2^{\gamma-1}$ on ν requires. In D2 it rises from 0.3852 to 0.5973, the value at $y = 3$; in D3, from 0.2072 to 0.2401. The herded means, 0.8417 in D4 and 0.7003 in D5, do not move.

Error. Against the predicted zero, D0, D4 and D5 change by at most 6.7×10^{-16} in λ and 6.4×10^{-13} in u . This verifies Corollary 1.

4. Noise That Grows With the Crowd

In D1 with $\gamma > 1$ an agent's noise $\nu m^{\gamma-1}$ grows with the density around it and vanishes where nobody is. The forward equation is a porous medium equation with a drift [14], the gradient flow of $\int m^\gamma/(\gamma-1) dx$ [8], whose signature is an edge beyond which the density is exactly zero: $m^{\gamma-1}$ is linear in u and vanishes where $u = \mu + \nu\gamma/(\gamma-1)$. An agent at the edge has no noise; the crowd is jostled inside and still at its edge. The linear crowd is positive everywhere, the barrier included. This one can withdraw from the walls and the barrier, and split into two islands.

Proposition 2 (Where the crowd can be absent). Under Assumption 1, let $\gamma > 1$ in D1. If $m = 0$ on an open interval I , then $V \geq \lambda$ on I .

Proof. On I , $a = 0$, so u is a Lipschitz viscosity solution of $\lambda + \frac{1}{2}(u')^2 = V$. Where it is differentiable, almost everywhere, its derivative lies in both semidifferentials and the equation holds classically, so $V - \lambda = \frac{1}{2}(u')^2 \geq 0$ almost everywhere on I , and everywhere by continuity. $\square$



The condition is necessary, not sufficient: $V(0) = h$ exceeds λ for every γ computed, yet the barrier empties only past $\gamma \approx 1.146$.

Script. `fig_frontier.py`: the D1 map, Algorithm 1 at 61 values of γ in $[1, 2.5]$.

Shows. Proposition 1 for D1 with $\gamma \geq 1$: edges and an empty gap.

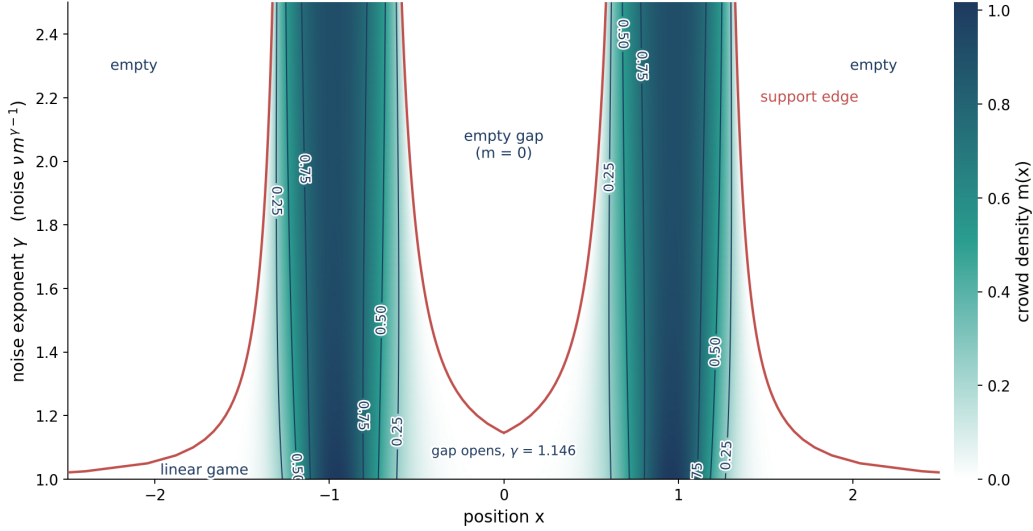


Figure 1: The D1 crowd density over position and noise exponent γ from 1 to 2.5, contours at 0.25, 0.5 and 0.75, with the support edges in red: they leave the walls at $\gamma \approx 1.021$, and an empty gap opens on the barrier at $\gamma \approx 1.146$.

The crowd is one island spanning both wells until the barrier empties, then two islands that narrow as γ rises, with peak density 0.917 at $\gamma = 2$ against 1.017 in the linear game.

The edges and the empty gap. *Method.* Algorithm 1 for D1; edges where m reaches zero, by local cubic interpolation; onsets interpolated between rows 0.025 apart.

Parameters. Definition 1; $N = 1600$; $N = 400, 800, 1600$ at $\gamma = 2$.

Result. The support leaves the walls at $\gamma \approx 1.021$ and the gap opens at $\gamma \approx 1.146$. The edges lie at $|x| = 0.422$ and 1.435 for $\gamma = 1.5$, 0.538 and 1.359 for $\gamma = 2$, 0.585 and 1.326 for $\gamma = 2.5$, while λ falls from 0.272 to 0.230, 0.206 and 0.191. Each empty region lies where $V > \lambda$.

Error. At $\gamma = 2$ the inner edge is 0.53801, 0.53776, 0.53769 on the three grids, order 1.85, stable to three figures; the forward-equation check gives 6.5×10^{-5} at $\gamma = 1.5$. This checks Proposition 2.

Remark 1 (Open: how foresight moves a frontier). In a game played in time the edge would move, and how is open. The porous medium spreads at finite speed [14], but outside the support the agents' equation is first order: a deterministic agent can plan to arrive exactly at the edge. How the backward equation moves a forward edge, and whether an edge can jump, are not known; the nearest theory is that of first-order games with density constraints, where the region is given and pressure is a price [17]. A smaller question is why λ falls as γ rises. A reading, not a proof: the density is near or below one, so the

noise is at most about ν ; doubling the crowd raises λ .

5. Noise That Falls With the Crowd

With $\gamma < 1$ lone agents are the noisiest. The forward equation is a fast diffusion equation, and Proposition 1 gives Tsallis’s q -exponential with $q > 1$ [13], the stationary state Plastino and Plastino found for a nonlinear Fokker–Planck equation [12], whose tails are powers of the value function, $m \sim u^{-1/(1-\gamma)}$; the linear game has a heavy tail only if V does. A second signature is less expected. Agents leave sparse places fast and dense ones slowly, so mass collects where it already is, and as γ falls the crowd gives up the wells for one cluster near the barrier, until the symmetric branch folds near $\gamma \approx 0.219$.

Script. `fig_tails.py`: the D1 map for $\gamma \leq 1$, 61 values from 0.225 to 1, as $\log_{10} m$.

Shows. Proposition 1 for D1 with $\gamma < 1$: power tails, and the crowd moving onto the barrier.

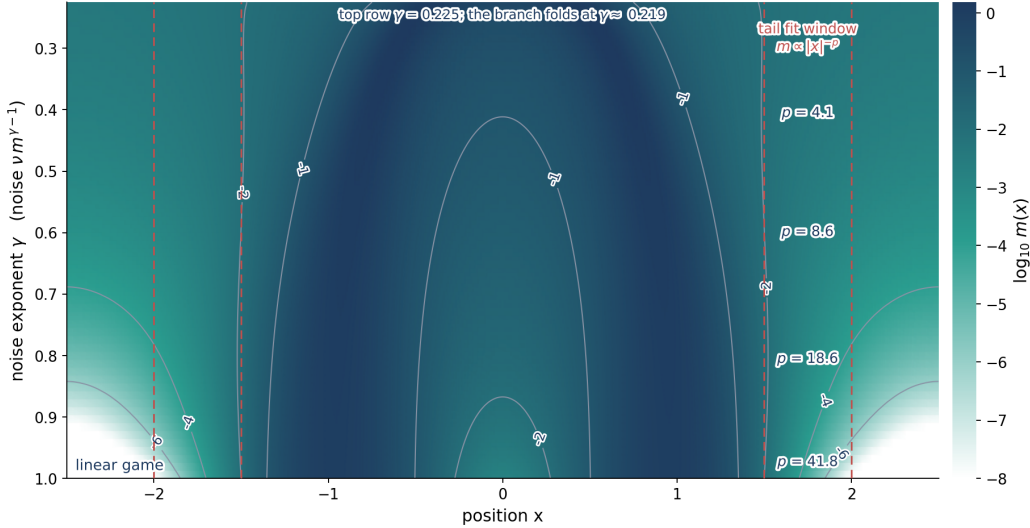


Figure 2: The base-10 logarithm of the D1 crowd density over position and noise exponent γ , from the linear game at the bottom to 0.225 at the top, with the tail-fit window $1.5 \leq |x| \leq 2$ dashed and the fitted effective exponents p .

The linear crowd falls to $10^{-20.6}$ at the wall. As γ falls the wall density rises to $10^{-2.7}$ at $\gamma = 0.4$, and the barrier density from 0.0014 to 0.43 at $\gamma = 0.225$, where the two wells’ crowds have merged.

How fast the tails decay. *Method.* Algorithm 1 for D1 on $[0.225, 1]$; a least-squares fit of $\ln m = \text{const} - p \ln x$ on $1.5 \leq x \leq 2$, an effective exponent since $L = 2.5$ allows no clean power-law range; continuation in λ ($N = 800$) past the lower end.

Parameters. Definition 1; $N = 400, 800, 1600$.

Result. $p = 52.2, 18.6, 8.56, 4.14$ and 2.18 at $\gamma = 1, 0.8, 0.6, 0.4, 0.225$, and λ rises from 0.272 to 0.867. Continuation in γ stops at 0.2194 on all three grids: a fold. The branch is S-shaped, turning at $\gamma \approx 0.2194$ and 0.2213; beyond, with $\lambda > h$, the crowd sits on the



barrier top and its density there grows to 21.8 at $\gamma = 0.140$, where the continuation stops.

Error. The worst fit has $R^2 = 0.980$; at $\gamma = 0.6$, $p = 8.56$ against the formal far-field 7.5 of Remark 2; the forward-equation check at $\gamma = 0.8$ gives 2.1×10^{-5} . This verifies Proposition 1 for $\gamma < 1$.

Remark 2 (Open: the fold, and where the crowd goes past it). Only symmetric states were followed. Whether the end of the branch is a loss of existence or a switch to one cluster in one well, not looked for here, is open, as is which of the three states between the folds an evolving crowd selects. So are the tails on the whole line: a formal balance of the far-field terms with $u \approx Ax^3$, using $a = \nu + (1 - \gamma)(u - \mu)/\gamma$, requires

$$A^2 \left(\frac{9}{2} - \frac{6(1 - \gamma)}{\gamma} \right) = h,$$

solvable only for $\gamma > 4/7$, where $m \sim x^{-3/(1-\gamma)}$. Below $4/7$ it fails while the box computation converges. The balance is formal and unproved, and Frank’s account of tails [11] assumes a given potential, not a value function the tails shape.

6. A Crowd With a Capacity

In D2 a place holds at most K agents per unit length, and moving through it costs $v^2/(2(1 - m/K))$. The best speed is scaled by the free room, $v^* = -(1 - m/K)u'$, and the forward equation acquires the mobility $m(1 - m/K)$ of the exclusion process [15], brought into pedestrian games by Burger, Di Francesco, Markowich and Wolfram [16], a change of metric in the gradient-flow picture [10]. By Proposition 1 the crowd is exactly Fermi–Dirac in the value function. The linear game can cap a density only by charging for presence; this crowd fills its best places to the brim and overflows.

Script. `fig_capacity.py`: the D2 map, Algorithm 1 at 61 loads y in $[0, 3]$.

Shows. Proposition 1 for D2: a Fermi–Dirac crowd saturating at the cap.

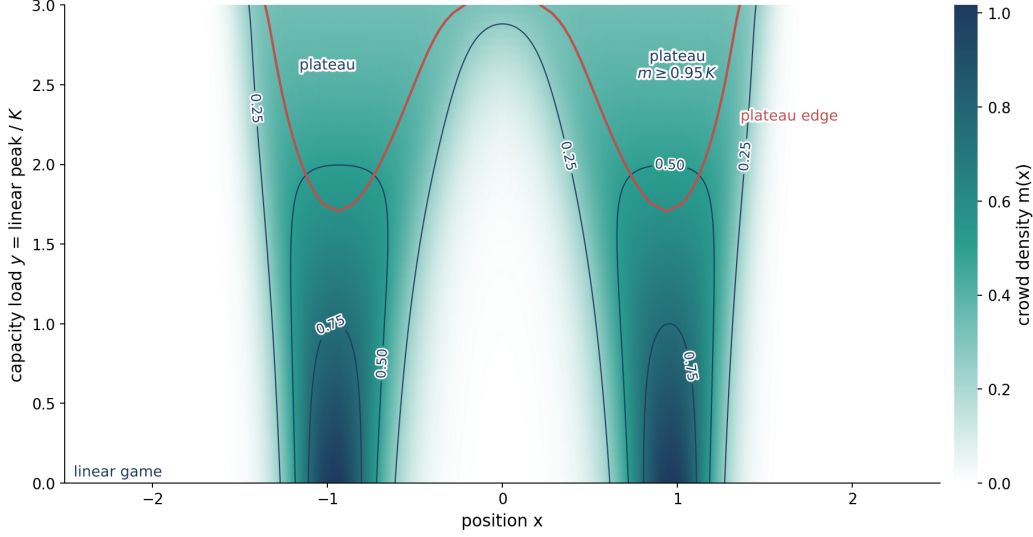


Figure 3: The D2 crowd density over position and capacity load y from 0 to 3, contours at 0.25, 0.5 and 0.75, with the plateau edges in red where the density reaches 0.95 of the capacity K , from load 1.71 upward.

The peaks first flatten and widen; from $y \approx 1.706$ a plateau forms in each well, and the overflow moves outward and onto the barrier. The agents pay for the lost room: λ rises from 0.272 to 0.597.

How far the linear crowd overflows. *Method.* Algorithm 1 for D2 with K the linear peak, 1.0173, over y ; the plateau is the length where $m > 0.95 K$.

Parameters. Definition 1; $N = 1600$; $N = 400, 800, 1600$ at $y = 2$.

Result. The plateau appears at $y \approx 1.706$: the linear crowd must overflow the cap 1.7-fold. Its length is 0.890, 1.528 and 2.280 at $y = 2, 2.5$ and 3, and λ is 0.335, 0.453 and 0.597 at $y = 1, 2, 3$.

Error. $\max m/K = 0.99987$, so the cap is kept by the statistics, not by clipping. The length at $y = 2$ converges at order 2.02 (0.890334, 0.890273, 0.890258), and the forward-equation check gives 2.2×10^{-5} . This verifies Proposition 1 for D2.

Remark 3 (Open: is a full room expensive or impassable). In D2 a full place is expensive but never full. In first-order games with density constraints the cap is hard and its pressure is a price [17]; in crowd motion it projects velocities onto admissible ones [18], [19]. As $\nu \rightarrow 0$ the Fermi–Dirac crowd formally becomes a saturated block, K where $u < \mu$ and empty elsewhere. Whether that is the pressure-equals-price equilibrium, how the vanishing mobility becomes a pressure, and whether a stationary crowd can tell the two apart, are open.

7. A Crowd That Slows Its Movers

In D3 moving through company costs $(\varepsilon + m)^c v^2/2$, so agents in dense places move slowly and the forward equation is nonlinear through the velocity alone. Games have studied this



form of touching most: existence and uniqueness for time-dependent congestion games, with a Hamiltonian undefined where the density vanishes [20], existence for stationary ones [21], and trade crowding, where liquidating costs more when others liquidate [35]. Proposition 1 gives D3 the statistics of noise, but its edge is made by cost. Two things are new: one statistics, two games (Corollary 2), and an externality corrected by a fee set by the crowd's flow (Proposition 3).

Corollary 2 (Two mechanisms, one statistics, two games). At $\varepsilon = 0$, D1 with $\gamma > 1$ and D3 with $c = \gamma - 1$ have the same q -exponential statistics, D1's argument scaled by $1/\gamma$:

$$m = \left[1 + \frac{(\gamma - 1)(\mu - u_{D1})}{\nu\gamma} \right]_+^{1/(\gamma-1)} = \left[1 + \frac{(\gamma - 1)(\mu - u_{D3})}{\nu} \right]_+^{1/(\gamma-1)}$$

exactly when $\mu - u_{D3} = (\mu - u_{D1})/\gamma$; the velocities are then $-u'_{D1}$ and $-u'_{D1}/(\gamma m^{\gamma-1})$. A density does not tell noise from congestion; the value function and the velocities do.

Proof. Proposition 1 for D1 and for D3; the relation gives $u'_{D3} = u'_{D1}/\gamma$, and the D3 velocity is $-u'_{D3}/m^c$. $\square$

D1 puts m in the diffusion and D3 in the Hamiltonian, so for one running cost they select different members of the family.

Proposition 3 (The planner's fee is set by the crowd's flow). In D3 with $\varepsilon > 0$, let (m^*, v^*) , smooth with $m^* > 0$, minimise $J = M^{-1} \int_{\Omega} (\ell(v, m) + V) m dx$ over velocity fields v, m being the stationary density of v , and let $\tau = m^* \partial_m \ell(v^*, m^*) = \frac{c}{2} m^* (\varepsilon + m^*)^{c-1} (v^*)^2$. (i) There are u and λ with $\partial_v \ell(v^*, m^*) = -u'$, $u' = 0$ at the walls and $\lambda - \nu u'' + H(u', m^*) = V + \tau$, the multiplier of the mass being $-\lambda$: the optimum is the equilibrium of the game in which each agent pays at x the fee $\tau(x)$, set by the crowd's velocity there. (ii) For $c > 0$, the same formula charged on the payer's own speed does not make (m^*, v^*) an equilibrium.

Proof. Zero flux gives $v = \nu m'/m$, so J depends on m alone. With $\delta v = \nu(\delta m/m)'$ and a multiplier η , the first variation of $MJ + \eta(\int_{\Omega} m dx - M)$ is

$$\int_{\Omega} (\ell + V + m \partial_m \ell + \eta) \delta m dx + \nu [\partial_v \ell \delta m]_{-L}^L - \nu \int_{\Omega} (m \partial_v \ell)' \frac{\delta m}{m} dx.$$

Set $u' = -\partial_v \ell(v^*, m^*)$. The boundary term gives $u' = v^* = 0$ at the walls, and

$$-\nu (m \partial_v \ell)' / m = \nu u'' + v^* u';$$

as $\ell + v^* u' = -H(u', m^*)$, stationarity is the stated equation with $\lambda = -\eta$, the optimality system of mean field type control [22], [23]. In the fee game τ is a function of x , so the agent's speed rule and equation are these: (i). For (ii) the payer's cost is $k v^2/2$, $k = (\varepsilon + m)^c + c m(\varepsilon + m)^{c-1}$, with Hamiltonian $\ell + \tau$ at $\tilde{u}' = -k v^*$. Subtracting the equations, with $g = \tilde{u}' - u' = -c m^* (\varepsilon + m^*)^{c-1} v^*$, so that $2\tau = -v^* g$, gives $\nu g' + v^* g = \tilde{\lambda} - \lambda$, that is $\nu(m^* g)' = (\tilde{\lambda} - \lambda) m^*$. Integrating, with $g = 0$ at the walls, gives $\tilde{\lambda} = \lambda$, then $g \equiv 0$, so $v^* \equiv 0$, $u' \equiv 0$ and $\lambda = V$, impossible for a non-constant V . $\square$



The fee charges an agent for what its presence adds to the cost of those passing at speed $v^*(x)$; at $c = 0$ it vanishes.

Script. `fig_congestion.py`: the D3 map, Algorithm 1 at 61 values of c in $[0, 1]$, $\varepsilon = 10^{-3}$.

Shows. Proposition 1 for D3: islands whose edges are made by cost.

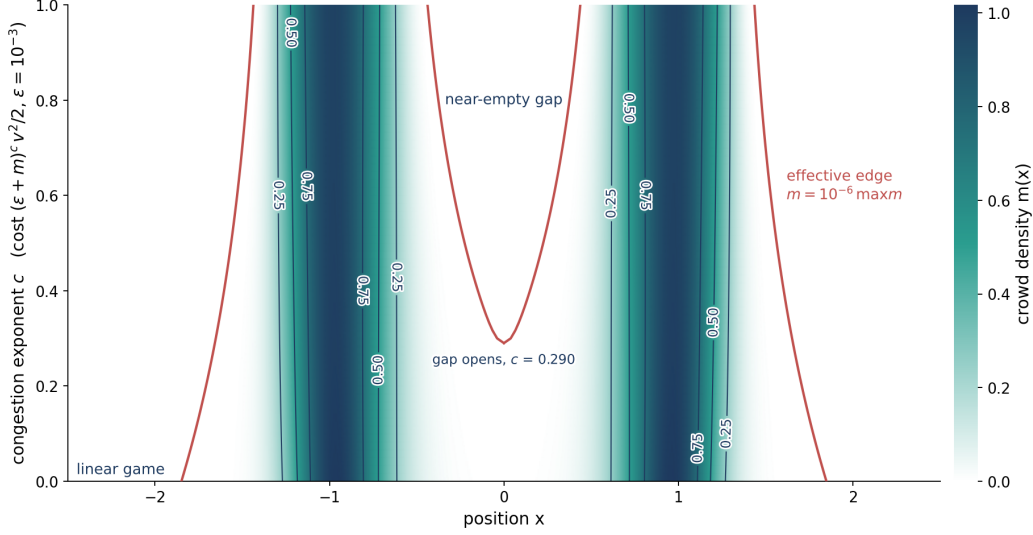


Figure 4: The D3 crowd density over position and congestion exponent c from 0 to 1 at $\varepsilon = 10^{-3}$, with effective edges in red where the density falls to 10^{-6} of its maximum; the edges of the near-empty gap depend on ε .

From $c \approx 0.290$ a near-empty gap opens. The density stays positive, and the gap edges depend on the regularisation at the level of 0.03: at $c = 0.5$ the inner edge is 0.140, 0.225 and 0.259 for $\varepsilon = 10^{-2}, 10^{-3}, 10^{-4}$. The ergodic constant falls to 0.173 at $c = 1$ because the density stays below one, where $(\varepsilon + m)^c < 1$.

Noise and congestion in the same family. *Method.* Algorithm 1 for D1 at $\gamma = 1.5$ and D3 at $c = 0.5$, both with exponent 2: $[1 + (\mu - u)/(3\nu)]_+^2$ and, at $\varepsilon = 0$, $[1 + (\mu - u)/(2\nu)]_+^2$.

Parameters. Definition 1; $\varepsilon = 10^{-3}$; $N = 1600$.

Result. The densities differ by at most 4.6 per cent of the peak, and λ is 0.230 under noise and 0.207 under congestion. The velocities differ by up to 1.95 on the noise crowd's support, where its largest speed is 1.60, and by 5.6 near the edges, where congested agents move fast.

Error. The forward-equation checks give 6.5×10^{-5} and 7.7×10^{-6} , and ε from 10^{-2} to 10^{-4} moves the congested λ only from 0.2090 to 0.2070: the differences are the games'. This verifies Corollary 2.

The planner found two ways. *Method.* J minimised directly over velocity fields, with no HJB equation (L-BFGS, then Newton on the same discrete J), against the fee game solved by Algorithm 1; 100 random perturbations of the optimum.

Parameters. D3, $\varepsilon = 10^{-3}$, $c = 0$ to 1; $N = 25\,600$.



Result. The fee game reproduces the optimum at every c . Without the fee the equilibrium costs 0.46, 1.31, 2.18 and 2.95 per cent more at $c = 0.25, 0.5, 0.75, 1$, and nothing at $c = 0$. At $c = 0.5$ the planner's multiplier, -0.248062 , is minus the fee game's λ .

Error. J agrees to 2.0×10^{-11} relative or better, the velocities to 5.0×10^{-4} , and no perturbation lowered J . This verifies Proposition 3(i).

Remark 4 (Open: fees out of equilibrium, and uniqueness). The fee of Proposition 3 is set by the flow at the optimum; recomputed from a moving flow it becomes a feedback, and whether that reaches the optimum is open. So is uniqueness: at $\varepsilon = 0$ the Hamiltonian $p^2/(2m^c)$ is singular where $m = 0$, handled for time-dependent games by weak solutions [20] but not for a stationary crowd with a near-empty gap, whose inner edge moves from 0.225 to 0.259 when ε falls tenfold. For trade crowding [35] the result suggests a fee set by the flow, not the trader's own rate.

8. A Crowd That Pulls

In D4 the crowd pulls each agent toward its mean, whatever the agent chooses. This McKean–Vlasov pull is the classical source of a phase transition: particles in a double well lose their symmetric state past a critical strength [24], and such transitions are well understood on the torus [25]; in finance the same drift couples banks' reserves in the systemic-risk game of Carmona, Fouque and Sun [36]. A linear game can make agents herd only through a conformity cost, as in D5. A pull looks different, a force and not a price, but in a stationary game it is not: Proposition 4 locates the branch point, and Proposition 5 shows that a pull θ is a conformity cost $\kappa = \theta^2$.

Proposition 4 (Where a herded crowd can branch off). Under Assumption 1(i)–(iii), in D4, and in D5 with κ for θ , $\bar{x} = 0$ is an equilibrium for every θ . Let $m[\bar{x}]$ solve the game with $\bar{x}$ held fixed, $\Phi(\bar{x}) = M^{-1} \int_{\Omega} x m[\bar{x}] dx$, whose fixed points are the equilibria, and suppose Φ is continuously differentiable in $(\bar{x}, \theta)$, with Φ' its derivative in $\bar{x}$. If $\Phi'(0) \neq 1$ at θ_0 , the symmetric equilibrium is the only one near $(0, \theta_0)$: a herded branch can leave it only where $\Phi'(0) = 1$.

Proof. With $\bar{x}$ fixed the equation for u contains no m , and $u = -2\nu \ln \psi$ makes it a Sturm–Liouville problem with a unique principal eigenpair. At $\bar{x} = 0$ the drift is odd and V even, so u is even and $\Phi(0) = 0$. The zeros of $F(\bar{x}, \theta) = \Phi(\bar{x}) - \bar{x}$ include $\bar{x} = 0$, and $\partial_{\bar{x}} F(0, \theta_0) = \Phi'(0) - 1 \neq 0$, so by the implicit function theorem they are only $\bar{x} = 0$ near $(0, \theta_0)$. $\square$

The threshold θ_c is the root of $\Phi'(0) = 1$; the branch found there is supercritical.

Proposition 5 (A gradient pull is a crowd cost in disguise). Let $\ell = v^2/2$, $a = \nu$ and $b = -\Psi'$ with $\Psi = \Psi[m](x)$, and put $w = u + \Psi$. Then $\lambda - \nu u'' + \frac{1}{2}(u')^2 + \Psi' u' = V$ is

$$\lambda - \nu w'' + \frac{1}{2}(w')^2 = V + F[m], \quad F[m] = \frac{1}{2}(\Psi')^2 - \nu \Psi'',$$

and $m \propto e^{-(u+\Psi)/\nu}$ is $m \propto e^{-w/\nu}$: the pull game is the linear game with crowd cost $F[m]$, with the same density, λ and total velocity $-w'$, the wall condition becoming $w' = \Psi'$. For D4, $F = \theta^2(x - \bar{x})^2/2 - \nu\theta$: D4 with pull θ is D5 with $\kappa = \theta^2$, λ shifted by $\nu\theta$.

Proof. With $u = w - \Psi$, $\frac{1}{2}(u')^2 + \Psi'u' = \frac{1}{2}(w')^2 - \frac{1}{2}(\Psi')^2$ and $-\nu u'' = -\nu w'' + \nu\Psi''$; zero flux reads $\nu m' = -(u' + \Psi')m = -w'm$. For D4, $\Psi' = \theta(x - \bar{x})$ and $\Psi'' = \theta$. $\square$

This is the shift that makes the quadratic game tractable [6]. In one dimension every drift is a gradient, so every physical force in a stationary quadratic game is a crowd cost in disguise; in higher dimensions only gradient forces are. What differs is the split of the velocity into chosen and suffered parts, and the wall condition, invisible here: the linear crowd's density at the walls is about $10^{-20.6}$.

Script. `fig_herding.py`: the D4 map, Algorithm 1 at 61 values of θ up to about $2\theta_c$, with the herded branch continued in $\bar{x}$.

Shows. Proposition 4: the herded branch leaving the symmetric crowd at θ_c .

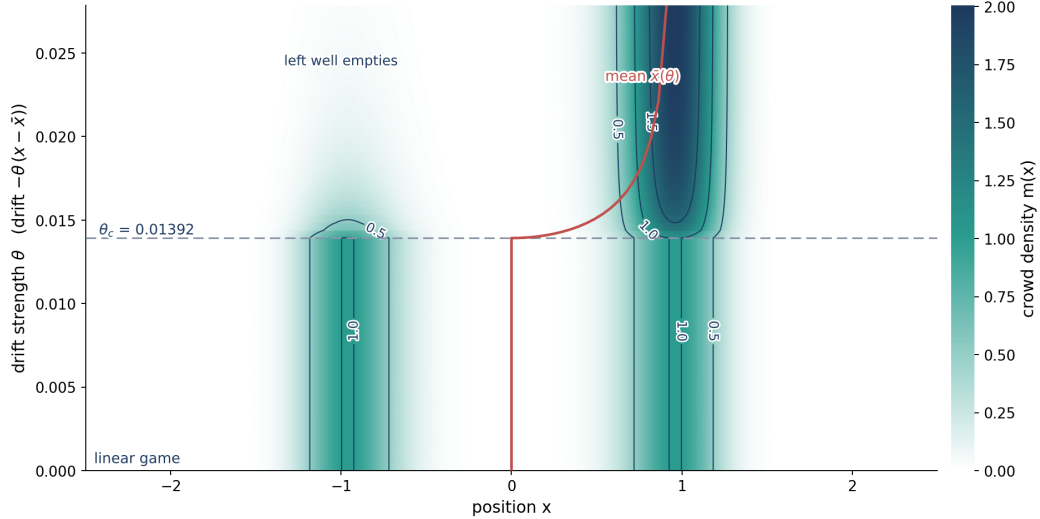


Figure 5: The D4 crowd density over position and pull θ from 0 to about twice the threshold, contours at 0.5, 1 and 1.5, with the mean $\bar{x}$ in red and $\theta_c \approx 0.0139$ dashed; past it the crowd herds into the right well.

Past θ_c the mean rises steeply, to 0.84 at $1.5\theta_c$ and 0.91 at $2\theta_c$, as the left well empties; the mirror branch and the symmetric crowd remain equilibria.

The two thresholds, each found two ways. *Method.* For D4 and D5: the root of $\Phi'(0) = 1$ from the exact derivative of the discrete map; and the herded branch, parameterised by $\bar{x}$, with the dial fitted as $\alpha_0 + \alpha_1\bar{x}^2 + \alpha_2\bar{x}^4$ on $\bar{x} \leq 0.1$ and extrapolated to α_0 .

Parameters. Definition 1; $N = 400, 800, 1600$.

Result. $\theta_c = 0.01392213$ and $\kappa_c = 1.938256 \times 10^{-4}$, each both ways, and $\sqrt{\kappa_c} = 0.01392213$; $\alpha_1 = 0.00394 > 0$, a supercritical branch. The herded mean is 0.842 at $1.5\theta_c$ and 0.700 at



$1.5 \kappa_c$, consistent since $1.5 \theta_c$ is $(1.5)^2 \kappa_c$. Half the tunnelling gap gives 1.71×10^{-4} for κ_c . *Error.* The two ways agree to 5.6×10^{-9} and 1.1×10^{-8} relative, and $\theta_c = \sqrt{\kappa_c}$ to 1.6×10^{-9} , far inside the grid error ($\theta_c = 0.0138825, 0.0139142, 0.0139221$, order 2.00). This verifies Propositions 4 and 5.

Remark 5 (Open: forces that are not gradients, and the dynamics). In one dimension every force is a gradient, so this paper cannot see what touching through a force adds. Three settings could: non-gradient forces in two dimensions, rotational or non-reciprocal; non-quadratic Hamiltonians; and the dynamics, where a force acts at once and a cost through anticipation, as synchronisation games show [26], [27], [28], against the uncontrolled transitions of [25]. The threshold is small because a long-run average forgets the one-off cost of moving between nearly degenerate wells, whose tunnelling gap is 3.4×10^{-4} ; a discounted game would be expected to herd later; by how much is open.

9. What the Insight Costs

Remark 6 (Five tools that lean on linearity). Each mechanism here is paid for with a tool the linear game supplies.

1. *Uniqueness.* The Lasry–Lions argument [1] needs a monotone crowd cost and a Hamiltonian free of m ; the three crowds between the folds, and herding, show it failing.
2. *Convex formulations.* Potential games are convex problems solved by augmented Lagrangians [33]; convexity survives a concave mobility such as $m(1 - m/K)$ [10], but congestion needs a fee to become an optimum.
3. *Many players.* Lacker’s limit [30] concerns interactions through the empirical measure; noise set by the local density needs moderate interactions [29], and for games that limit is open.
4. *Numerics.* Structure-preserving schemes rest on the adjoint structure of the linear game [31], [32]; this paper used Newton on the full system, affordable in one dimension.
5. *Time.* The master equation [4] and convergence to the ergodic game [34] are built on the linear equation; each crowd here is a long-time state only if that convergence holds.

The open questions, ordered by what an answer would buy:

1. **How foresight moves a frontier** (Proposition 2, Remark 1): how the backward equation moves an edge where the noise vanishes.
2. **What a force adds** (Proposition 5, Remark 5): non-gradient forces, non-quadratic Hamiltonians and the dynamics.



3. **Fees out of equilibrium** (Proposition 3, Remark 4): whether a fee that follows a moving flow reaches the optimum.
4. **Uniqueness under congestion** (Figure 4, Remark 4): a stationary crowd with a near-empty gap.
5. **Expensive or impassable** (Figure 3, Remark 3): the Fermi–Dirac crowd as $\nu \rightarrow 0$.
6. **Past the fold** (Figure 2, Remark 2): loss of existence or a one-well cluster, and the tails below $\gamma = 4/7$.
7. **Which tools survive** (Remark 6): uniqueness, convexity, the many-player limit and convergence in time.

10. Conclusion

A linear Fokker–Planck equation describes agents that never touch: the stationary crowd is a Boltzmann statistics of the value function, and without a crowd cost the game decouples. Each nonlinear forward equation selects another statistics and couples the game through the dynamics alone.

Growing noise gives edges and an empty gap; falling noise gives heavy tails and one cluster, until the branch folds; a capacity gives a Fermi–Dirac plateau; congestion gives the noise family in another game, whose optimum needs a fee set by the crowd’s flow. A gradient pull buys nothing: it is a conformity cost in disguise.

The price is part of the linear toolkit; the open questions mark where the computation runs ahead of the theory.



References

- [1] Lasry, J.-M., & Lions, P.-L. (2007). Mean field games. *Japanese Journal of Mathematics*. <https://doi.org/10.1007/s11537-007-0657-8>
- [2] Huang, M., Malhamé, R. P., & Caines, P. E. (2006). Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information and Systems*. <https://doi.org/10.4310/cis.2006.v6.n3.a5>
- [3] Carmona, R., & Delarue, F. (2018). *Probabilistic Theory of Mean Field Games with Applications I*. Springer. <https://doi.org/10.1007/978-3-319-58920-6>
- [4] Cardaliaguet, P., Delarue, F., Lasry, J.-M., & Lions, P.-L. (2019). *The Master Equation and the Convergence Problem in Mean Field Games*. Princeton University Press. <https://doi.org/10.23943/princeton/9780691190716.001.0001>
- [5] Achdou, Y., Cardaliaguet, P., Delarue, F., Porretta, A., & Santambrogio, F. (2020). *Mean Field Games*. Springer. <https://doi.org/10.1007/978-3-030-59837-2>
- [6] Guéant, O. (2009). A reference case for mean field games models. *Journal de Mathématiques Pures et Appliquées*. <https://doi.org/10.1016/j.matpur.2009.04.008>
- [7] Jordan, R., Kinderlehrer, D., & Otto, F. (1998). The variational formulation of the Fokker-Planck equation. *SIAM Journal on Mathematical Analysis*. <https://doi.org/10.1137/s0036141096303359>
- [8] Otto, F. (2001). The geometry of dissipative evolution equations: the porous medium equation. *Communications in Partial Differential Equations*. <https://doi.org/10.1081/pde-100002243>
- [9] Carrillo, J. A., McCann, R. J., & Villani, C. (2003). Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Revista Matemática Iberoamericana*. <https://doi.org/10.4171/rmi/376>
- [10] Dolbeault, J., Nazaret, B., & Savaré, G. (2008). A new class of transport distances between measures. *Calculus of Variations and Partial Differential Equations*. <https://doi.org/10.1007/s00526-008-0182-5>
- [11] Frank, T. D. (2005). *Nonlinear Fokker-Planck Equations*. Springer. <https://doi.org/10.1007/b137680>
- [12] Plastino, A. R., & Plastino, A. (1995). Non-extensive statistical mechanics and generalized Fokker-Planck equation. *Physica A: Statistical Mechanics and its Applications*. [https://doi.org/10.1016/0378-4371\(95\)00211-1](https://doi.org/10.1016/0378-4371(95)00211-1)
- [13] Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*. <https://doi.org/10.1007/bf01016429>
- [14] Vázquez, J. L. (2006). *The Porous Medium Equation*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198569039.001.0001>



- [15] Kipnis, C., & Landim, C. (1999). *Scaling Limits of Interacting Particle Systems*. Springer. <https://doi.org/10.1007/978-3-662-03752-2>
- [16] Burger, M., Di Francesco, M., Markowich, P. A., & Wolfram, M.-T. (2014). Mean field games with nonlinear mobilities in pedestrian dynamics. *Discrete and Continuous Dynamical Systems - B*. <https://doi.org/10.3934/dcdsb.2014.19.1311>
- [17] Cardaliaguet, P., Mészáros, A. R., & Santambrogio, F. (2016). First order mean field games with density constraints: pressure equals price. *SIAM Journal on Control and Optimization*. <https://doi.org/10.1137/15m1029849>
- [18] Maury, B., Roudneff-Chupin, A., & Santambrogio, F. (2010). A macroscopic crowd motion model of gradient flow type. *Mathematical Models and Methods in Applied Sciences*. <https://doi.org/10.1142/s0218202510004799>
- [19] Santambrogio, F. (2015). *Optimal Transport for Applied Mathematicians*. Birkhäuser. <https://doi.org/10.1007/978-3-319-20828-2>
- [20] Achdou, Y., & Porretta, A. (2018). Mean field games with congestion. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*. <https://doi.org/10.1016/j.anihpc.2017.06.001>
- [21] Gomes, D., & Mitake, H. (2015). Existence for stationary mean-field games with congestion and quadratic Hamiltonians. *Nonlinear Differential Equations and Applications NoDEA*. <https://doi.org/10.1007/s00030-015-0349-7>
- [22] Achdou, Y., & Laurière, M. (2016). Mean field type control with congestion. *Applied Mathematics & Optimization*. <https://doi.org/10.1007/s00245-016-9342-8>
- [23] Bensoussan, A., Frehse, J., & Yam, P. (2013). *Mean Field Games and Mean Field Type Control Theory*. Springer. <https://doi.org/10.1007/978-1-4614-8508-7>
- [24] Dawson, D. A. (1983). Critical dynamics and fluctuations for a mean-field model of cooperative behavior. *Journal of Statistical Physics*. <https://doi.org/10.1007/bf01010922>
- [25] Carrillo, J. A., Gvalani, R. S., Pavliotis, G. A., & Schlichting, A. (2019). Long-time behaviour and phase transitions for the McKean-Vlasov equation on the torus. *Archive for Rational Mechanics and Analysis*. <https://doi.org/10.1007/s00205-019-01430-4>
- [26] Yin, H., Mehta, P. G., Meyn, S. P., & Shanbhag, U. V. (2012). Synchronization of coupled oscillators is a game. *IEEE Transactions on Automatic Control*. <https://doi.org/10.1109/tac.2011.2168082>
- [27] Carmona, R., Cormier, Q., & Sonner, H. M. (2023). Synchronization in a Kuramoto mean field game. *Communications in Partial Differential Equations*. <https://doi.org/10.1080/03605302.2023.2264611>



- [28] Cirant, M. (2019). On the existence of oscillating solutions in non-monotone mean-field games. *Journal of Differential Equations*.
<https://doi.org/10.1016/j.jde.2018.12.025>
- [29] Oelschläger, K. (1985). A law of large numbers for moderately interacting diffusion processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*.
<https://doi.org/10.1007/bf02450284>
- [30] Lacker, D. (2015). A general characterization of the mean field limit for stochastic differential games. *Probability Theory and Related Fields*.
<https://doi.org/10.1007/s00440-015-0641-9>
- [31] Achdou, Y., & Capuzzo-Dolcetta, I. (2010). Mean field games: numerical methods. *SIAM Journal on Numerical Analysis*. <https://doi.org/10.1137/090758477>
- [32] Achdou, Y., Camilli, F., & Capuzzo-Dolcetta, I. (2013). Mean field games: convergence of a finite difference method. *SIAM Journal on Numerical Analysis*.
<https://doi.org/10.1137/120882421>
- [33] Benamou, J.-D., & Carlier, G. (2015). Augmented Lagrangian methods for transport optimization, mean field games and degenerate elliptic equations. *Journal of Optimization Theory and Applications*. <https://doi.org/10.1007/s10957-015-0725-9>
- [34] Cardaliaguet, P., Lasry, J.-M., Lions, P.-L., & Porretta, A. (2012). Long time average of mean field games. *Networks and Heterogeneous Media*.
<https://doi.org/10.3934/nhm.2012.7.279>
- [35] Cardaliaguet, P., & Lehalle, C.-A. (2017). Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics*.
<https://doi.org/10.1007/s11579-017-0206-z>
- [36] Carmona, R., Fouque, J.-P., & Sun, L.-H. (2015). Mean field games and systemic risk. *Communications in Mathematical Sciences*.
<https://doi.org/10.4310/cms.2015.v13.n4.a4>