Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Computational Validation of a Kinetic PDE for Reinforcement Learning in Market-Entry Games

Esther Bou Dagher

esther.bou-dagher@dauphine.psl.eu

Paris Dauphine University https://orcid.org/0000-0003-3638-6593

Research Article

Keywords: computational validation, stochastic reinforcement learning, market-entry games, kinetic equation, Fokker-Planck equation, finite-volume methods, particle simulation, Kramers-Moyal approximation, molecular-chaos closure, finite-player effects, aggregate learning, sorting dynamics

Posted Date: September 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11202562/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Computational Validation of a Kinetic PDE for Reinforcement Learning in Market-Entry Games

Esther Bou Dagher *

Abstract. One-particle kinetic equations provide efficient descriptions of large adaptive populations, but their reliability for finite, globally coupled learning systems is not automatic. We computationally assess the nonlinear transport–diffusion equation derived in [4] from a stochastic reinforcement-learning model of repeated market entry through a second-order Kramers–Moyal approximation and a molecular chaos-type closure. Rather than only quantifying particle–continuum discrepancies, we ask which observables are preserved, how far the reduced model remains informative beyond the assumptions of the analytical theory, and where finite-player information is lost. To probe the role of a positive probability floor, we compare a floored response map with a smooth vanishing-tail map outside that setting. Independently audited particle and finite-volume computations show that the PDE reproduces aggregate entry, within-agent Bernoulli uncertainty, and return towards market capacity more reliably than the complete distribution of individual entry probabilities. These aggregate conclusions persist for both maps, although the vanishing-tail case retains a detectable law-level discrepancy. Mechanism-resolving diagnostics show that finite-player dependence weakens as the population grows. Within an exact one-step generator audit, replacing focal finite-player information by global moments produces the largest local discrepancy among the tested approximation steps, whereas the second-order truncation remainder is much smaller. For the floored map, long-time computations further reveal a pronounced separation between rapid aggregate learning and slow, fluctuation-driven sorting. Taken together, the study provides a computational map of when the kinetic PDE can serve as an efficient surrogate for repeated particle simulations, which observables it preserves, and where targeted corrections could broaden its use in finite-player parameter studies and otherwise costly long-time calculations.

Key words. computational validation, stochastic reinforcement learning, market-entry games, kinetic equation, Fokker–Planck equation, finite-volume methods, particle simulation, Kramers–Moyal approximation, molecular-chaos closure, finite-player effects, aggregate learning, sorting dynamics

AMS subject classifications. 65M08, 37M05, 68T05, 35Q91, 35Q84, 91A26, 35B40

1. Introduction.

Some opportunities lose their value precisely because too many people pursue them. Consider a crowded opportunity: a trading strategy whose edge disappears when it becomes popular, a product market with only a limited pool of profitable demand, or a venue with finite room. Entry can be rewarding while only a few others make the same choice. If too many enter, competition or congestion can erase the reward and turn entry into a loss; staying out avoids that loss, but also gives up the possible gain. Market-entry games isolate this feedback in a deliberately simple form. Each participant chooses privately between entering and staying out; entry is rewarded below a known capacity and discouraged when realised attendance exceeds it [16].

The El Farol Bar problem made this coordination problem a canonical example of inductive decision making under limited information: agents choose before current attendance is known, even though that attendance, and therefore the reward from entry, is created by their

*CEREMADE (CNRS UMR 7534), PSL University, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France (esther.bou-dagher@dauphine.psl.eu).

simultaneous decisions [1]. The central question is therefore dynamic: how can decentralised agents, learning only from noisy outcomes, produce a reliable collective response? Laboratory market-entry experiments reveal a striking divide between collective and individual behaviour: aggregate entry approaches the equilibrium capacity benchmark, while persistent differences remain across participants [16]. The population can therefore learn how many should enter without yet settling who will do so.

Reinforcement learning supplies a natural mechanism for this adaptation. Rather than solve the game afresh in every round, agents update their action propensities from realised payoffs, becoming more likely to repeat rewarded actions and less likely to repeat disappointing ones. This idea has roots in adaptive-behaviour models and underlies influential models and experiments on learning in games [11, 17, 9, 10]. Studies of repeated market-entry games distinguish two outcomes: *aggregate learning*, in which the average number of entrants moves towards capacity, and *sorting*, in which individuals gradually specialise into near-certain entrants and non-entrants in proportions compatible with capacity [8, 7]. Aggregate learning can occur well before sorting is complete; matching total attendance therefore does not ensure that a model reproduces the distribution of individual behaviour [7].

Two mathematical viewpoints illuminate these dynamics. Stochastic approximation retains the finite population but relates its noisy updates to a deterministic dynamical system, whose limiting ODEs reveal the stability of candidate equilibria [2]. This framework has been used to study convergence and sorting in market-entry and El Farol games [7, 23]. The complementary kinetic viewpoint follows the distribution of propensities rather than every agent's trajectory. Such equations have been developed for adaptive learning and strategic interactions in large populations [14, 15], alongside mean-field models driven by local game-theoretic equilibria [6].

The kinetic reduction is attractive analytically and computationally: it replaces M evolving propensities, and their M-dimensional joint law, by a single density on the one-dimensional propensity space. This one-dimensional reduction permits broad parameter studies and long-time calculations at much lower cost than repeated particle simulations. In classical mean-field theory, such a reduction can sometimes be justified by *propagation of chaos*: as M grows, any fixed set of tagged particles becomes asymptotically independent, with each governed by the same one-particle law [13, 19, 5].

Against this background, Bou Dagher, Perepelitsa, and Zatorska derived a one-particle kinetic PDE for stochastic reinforcement learning in a market-entry game [4]. Their derivation contains *two distinct approximation steps*. First, the *second-order Kramers–Moyal approximation* replaces exact finite propensity shifts by the drift and diffusion determined by their first two conditional moments. Second, a factorisation *analogous to molecular chaos* replaces the two- and three-particle marginal laws by products of the one-particle density. Under the stated hypotheses, the resulting nonlinear transport–diffusion equation admits a unique strong solution and exhibits long-time sorting; under additional positive-floor and transport-dominance conditions, the theory also bounds the long-time capacity error. It further predicts that aggregate learning occurs much earlier than sorting.

That theory creates a computational question with practical consequences. The aim is not simply to label the reduced PDE as reliable or unreliable, but to determine what it reproduces at finite population sizes and jump scales, what information it loses, and why. If it preserves

the observables relevant to a given task, the one-dimensional PDE can screen broad parameter families and reach sorting horizons that are costly to explore by repeatedly simulating every agent. If a remaining discrepancy can be traced to a particular step of the reduction, that step becomes a focused target for sharper analysis or correction. The goal is therefore a map of the PDE's accuracy that can turn it into a computational backbone for finite-player problems.

The two approximation steps in [4] remove different information from the finite game. The *Kramers–Moyal approximation* replaces an agent's actual one-round shift by its first two conditional moments; its accuracy must therefore be tested against the realised size of those finite shifts. The *one-particle factorisation* removes dependence between agents. Its algebra resembles molecular chaos, but the interaction here is not the usual kinetic picture of locally colliding particles: every entrant responds to the same aggregate attendance. In our case, however, even though the entry decisions are conditionally independent, a crowded or sparse round can move many propensities in the same direction and create dependence in later states. For this learning process, the one-particle factorisation is therefore a modelling closure, not the consequence of an established propagation-of-chaos limit. Whether its loss of information becomes negligible as M grows is a question to be investigated in this paper.

A second axis of this study is the probability map that converts propensity into entry probability. This map determines how strongly a discouraged agent can withdraw from the market and enters directly into the assumptions of the long-time theory. We therefore study two branches. Branch A is a floored logistic map, so an agent's entry probability never falls below $p_{\min} = 0.05$; at the probability-map level, it is aligned with the hypotheses of the analytical paper [4]. Branch B is smooth and increasing but tends to zero in its lower tail, allowing sufficiently discouraged agents to become arbitrarily unlikely to enter. It lies outside the positive-floor hypothesis in [4] used to control aggregate learning. Such a vanishing-tail response is natural in settings where repeated losses can make future participation negligible. Comparing the branches therefore asks whether particle–PDE agreement and the learning mechanisms depend essentially on the analytical floor, or remain informative for a broader class of behavioural response rules. The results show that several key aggregate conclusions survive without the floor, broadening the computational scope of the PDE and identifying a concrete direction for extending the theory. We organise the results around four questions:

1. *When is the PDE an accurate finite-time surrogate, and what does it preserve?* A discrepancy is scientifically meaningful only after errors in the deterministic solver and stochastic simulator have been separated from errors in the model reduction. We therefore certify conservation, positivity, refinement, calibration, Bernoulli randomisation, update identities, and fixed-seed reproducibility before making matched comparisons. We then move from the least to the most demanding targets: mean entry, within-agent Bernoulli uncertainty, probability-interior mass, and the full law measured in probability-space Wasserstein distance. Across $30 \le M \le 30000$, the aggregate quantities are consistently easier to reproduce than the complete law. Branch A remains within every declared finite-time behavioural and law-level threshold. Branch B retains a small but resolvable law discrepancy even where its aggregate moments are accurate. The design and certification are given in Section 3, and the matched results in Section 4.
2. *Which approximation matters when law-level agreement is imperfect?* The two steps

leave different empirical signatures. The one-particle factorisation discards finite-player dependence, so any associated signal should weaken as M grows. After removing the multinomial sampling baseline, excess dependence decays approximately as M^{-1}: doubling the population roughly halves it. This is consistent with, but does not prove, increasing finite-population independence. The Kramers–Moyal approximation instead neglects effects beyond the first two conditional moments of each finite shift. Realised one-round probability shifts scale almost one-for-one with the prescribed jump parameter ρ and show negligible residual dependence on M.
Two further tests sharpen the diagnosis. Supplying the PDE with aggregate coefficient paths measured from the particles does not systematically improve law-level agreement, so aggregate feedback alone does not explain the discrepancy. An exact one-step weak-generator hierarchy then isolates attendance heterogeneity, focal leave-one-out information, and finite-shift truncation in sequence. For the tested states and smooth observables, the largest local change occurs when a focal agent's environment is replaced by a single global mean; the remainder beyond the second-order Kramers–Moyal approximation is much smaller. This identifies the loss of focal finite-player information as the leading local target for refinement in the audited regime. Section 5 develops this diagnosis.

3. *Does the restoring mechanism work away from the target, rather than only beside it?* The baseline comparisons begin with the mean entry probability calibrated to capacity. That is useful for detecting distributional error, but by itself it could make a model with weak attraction appear successful. We therefore displace the initial mean both above and below capacity, vary the magnitude of the displacement, and rearrange the initial density while holding its mean fixed. Every one of the 46 tested deterministic PDE cases removes more than 99% of its initial capacity error over the prescribed finite horizon. Replicated particle experiments show the same restoring direction. Because an individual finite population continues to fluctuate after its systematic error has contracted, independent zero-offset controls define the noise level at which "return" can honestly be resolved. These off-capacity tests and the finite-M bridge are reported in Section 6.
4. *How much slower is sorting than aggregate learning, and what keeps it moving?* Once expected attendance is correct, the game can look settled even though most agents remain far from pure behaviour. We therefore follow the mass left in fixed propensity windows rather than inferring sorting from the mean. This creates a numerical difficulty of its own: the analytical PDE lives on the whole real line, whereas a computation has artificial walls. The long-time study uses a greatly enlarged domain, monitors mass near its edges, and accepts a sorting milestone only before a predeclared boundary warning. For Branch A, aggregate return is visible on a scaled-time horizon of order 10, while direct evacuation of the central window requires times of order 10^5–10^6, a separation of at least four orders of magnitude. At those milestones, diffusion generated by the residual capacity error is negligible relative to the Bernoulli-fluctuation term: the mean has settled, but randomness in individual entry decisions continues to rearrange the population. The near-collapse of scaled crossing times across population sizes is robust, although the absolute times retain visible spatial-resolution

sensitivity. A comparable boundary-safe milestone for Branch B is unresolved on the available horizon and is not interpreted as failed sorting. The boundary-safe long-time analysis is given in Section 7.

The remainder of the paper is organised as follows: Section 2 introduces the finite-player process, the two probability maps, the closed kinetic PDE, and the analytical targets, while Section 3 sets out the frozen computational design, independent numerical implementations, comparison metrics, and mechanism-resolving diagnostics. Section 4 reports the numerical audits and establishes observable-dependent finite-time agreement, Section 5 investigates the origins of the remaining law discrepancy, Section 6 tests deterministic and finite-player recovery from off-capacity states, and Section 7 quantifies the separation between aggregate learning and sorting; the supplement provides the complete audits, diagnostic atlases, robustness analyses, and supporting long-time calculations.

2. Microscopic model, kinetic PDE, and analytical targets.

We use the stochastic market-entry model and kinetic PDE of Bou Dagher, Perepelitsa, and Zatorska [4], introducing only the notation and analytical targets needed for the computational study.

2.1. Finite-player process and capacity target.

There are M agents and an integer market capacity M_c. At round n, agent i has propensity $X_n^i \in \mathbb{R}$, which is mapped to an entry probability by an increasing function $p : \mathbb{R} \to (0, 1]$. Conditional on the current propensities,

$$\delta_n^i \sim \text{Bernoulli}\big(p(X_n^i)\big)\,, \qquad m_n = \sum_{j=1}^{M} \delta_n^j, \tag{2.1}$$

where δ_n^i is the entry indicator and m_n the realised attendance. The propensity evolves according to

$$X_{n+1}^i = X_n^i + h\big(M_c - m_n\big)\delta_n^i, \tag{2.2}$$

so only entrants update, being reinforced below capacity and discouraged above it; this is model (M1) of [4]. The symmetric mixed-equilibrium entry probability is

$$\kappa(M) = \frac{M_c - 1}{M - 1}, \tag{2.3}$$

while Nash-equilibrium expected attendance lies between $M_c - 1$ and M_c, corresponding to the mean-entry interval

$$\mathcal{I}_M = \left[\frac{M_c - 1}{M}, \frac{M_c}{M}\right], \qquad \kappa(M) \in \mathcal{I}_M. \tag{2.4}$$

Aggregate learning means that the expected entry rate approaches this capacity interval, whereas sorting means that individual propensities eventually leave every fixed bounded interval [4]; these behaviours are examined in Sections 6 and 7, respectively. For the particle system, the corresponding mean entry probability is

$$\alpha_n^M = \frac{1}{M}\sum_{i=1}^{M} p(X_n^i). \tag{2.5}$$

Computationally, M_c is chosen so that $\kappa(M)$ remains close to the nominal target $\kappa_\star = 0.30$, and the baseline initial law is calibrated to satisfy

$$\int_{\mathbb{R}} p(x) f_0(x)\, dx = \kappa(M) \tag{2.6}$$

up to numerical tolerance. The initial-law construction and calibration are given in Section 3.1, while attraction from deliberately off-capacity states is tested in Section 6.

2.2. Two probability maps. The computations use two increasing maps $p : \mathbb{R} \to (0, 1]$ with deliberately different lower-tail behaviour. Branch A is the floored logistic

$$p_A(x) = p_{\min} + (1 - p_{\min})\sigma(\lambda(x - s)), \qquad \sigma(z) = \frac{1}{1 + e^{-z}}, \tag{2.7}$$

with $p_{\min} = 0.05$, $\lambda = 0.9$, and $s = 0$. This is the theorem-facing branch: its positive whole-line floor and derivative structure are compatible with the probability-map assumptions used in the analytical existence, uniqueness, and long-time results of [4]. Branch B is the softplus ratio

$$q_\ell(x) = \ell \log\Big(1 + e^{(x-s)/\ell}\Big), \qquad p_B(x) = \frac{q_\ell(x)}{q_\ell(x) + x_0}, \tag{2.8}$$

with $\ell = 0.6$, $x_0 = 1$, and $s = 0$. It retains the ratio structure used to motivate entry probabilities in the analytical model while extending it smoothly to all $x \in \mathbb{R}$. Unlike Branch A, $\inf_{\mathbb{R}} p_B = 0$, so it lies outside the positive-floor assumption used for the aggregate-learning part of the long-time theorem. The two branches therefore serve distinct purposes. Branch A tests the PDE in probability-map geometry aligned with the analytical theory [4], whereas Branch B tests how robust the finite-player/PDE agreement and the learning and sorting mechanisms are when the positive lower tail is removed. We compare the two branches throughout the numerical study. The derivative checks, precise theorem assumptions, and probability-map geometry are given in Supplementary Section SM1.1.

2.3. From the stochastic process to the closed PDE. The analytical paper passes from the finite-player learning rule to a one-particle kinetic PDE through two approximations. First, the Kolmogorov increment of the full M-particle law is expanded to second order in the update size h. Keeping the h/τ and h^2/τ terms, and discarding the $O(h^3)/\tau$ remainder under the small-step regime $h, \tau \ll 1$ with $h^2/\tau \sim 1$, gives a Kramers–Moyal, or Fokker–Planck, approximation; see Section 3 of [4]. Second, the resulting two- and three-particle marginals are factorised into products of the one-particle density, as in the molecular-chaos closure introduced in Section 3 of [4]. This closure is a modelling assumption rather than a proved propagation-of-chaos theorem for the learning process. A central aim of the present numerical study is therefore to test the closed PDE at finite M and to separate, where possible, the effects of finite jumps and finite-population dependence; see Sections 4 and 5.

Let $f(x, t)$ denote the one-particle propensity density. We use the aggregate observables

$$\alpha(t) = \int_{\mathbb{R}} p(x) f(x, t)\, dx, \qquad a(t) = \kappa(M) - \alpha(t), \qquad b(t) = \alpha(t)\big(1 - \alpha(t)\big), \tag{2.9}$$

where $\alpha(t)$ is the mean entry probability, $a(t)$ is its deviation from the symmetric capacity target, and $b(t)$ is the Bernoulli variance of the population mean. We also track

$$\beta(t) = \int_{\mathbb{R}} p(x)\big(1 - p(x)\big) f(x,t)\, dx, \tag{2.10}$$

the within-agent uncertainty quantity introduced in the analytical estimates of [4]. Since

$$b(t) - \beta(t) = \mathrm{Var}_{f(t)}\big(p(X)\big) \geq 0, \tag{2.11}$$

$\beta(t)$ measures average within-agent Bernoulli uncertainty, while $b(t) - \beta(t)$ measures heterogeneity of entry probabilities across agents. After the second-order truncation and one-particle closure, the density satisfies

$$\partial_t f + \frac{M-1}{\sqrt{\tau}}\, a(t)\, \partial_x\big(p(x) f\big) - \frac{(M-1)^2}{2}\left(a(t)^2 + \frac{b(t)}{M-1}\right) \partial_{xx}\big(p(x) f\big) = 0, \qquad f(\cdot,0) = f_0. \tag{2.12}$$

This is the closed one-particle equation derived in Section 3 of [4]. The drift is controlled by the capacity error $a(t)$, whereas the diffusion contains both the squared capacity error and the population uncertainty $b(t)$. This structure is important computationally: aggregate learning is tied to the decay of $a(t)$, while sorting may continue through the remaining diffusive dynamics.

2.4. Analytical targets for the computations. The analytical theory motivates two long-time numerical targets: aggregate learning and sorting [4]. Aggregate learning asks whether the mean entry probability approaches the finite-game equilibrium interval $\mathcal{I}_M$ in (2.4). Sorting is stronger: the density should leave every fixed bounded propensity window. We measure this directly by

$$S_R(t) = \int_{-R}^{R} f(x,t)\, dx, \tag{2.13}$$

with sorting corresponding to $S_R(t) \to 0$ for every fixed $R > 0$. Aggregate return is tested in Section 6, and direct long-time sorting diagnostics are reported in Section 7. The main long-time theorem separates these questions. Theorem 5.1(i) proves sorting under the stated regularity and curvature assumptions, while Theorem 5.1(ii) adds a positive probability floor and the sufficient transport-dominance condition

$$\widehat{c}_p \sqrt{\tau}(M-1) < p_{\min}^4. \tag{2.14}$$

This condition is used to control the long-time capacity error. Branch A is therefore theorem-oriented at the probability-map level, whereas Branch B is a robustness branch outside the positive-floor hypothesis. Finally, the analytical time-scale discussion predicts faster aggregate adjustment than sorting [4]. The computational companion tests this separation at finite M: finite-time PDE accuracy is assessed in Section 4, approximation mechanisms in Section 5, and aggregate return versus long-time redistribution in Sections 6–7.

3. Computational and statistical methodology. The computations follow a frozen staged workflow. The grids, initial laws, scaling, seeds, report times, discrepancy metrics, and operational thresholds were fixed before particle–PDE discrepancies were inspected. The PDE solver and microscopic simulator were audited independently before matched comparisons, mechanism diagnostics, and long-time runs were analysed.

3.1. Frozen design and calibrated initial laws. The population registry contains 100 distinct values of M, obtained by rounding a logarithmic grid from 30 to 30000, removing duplicates, and retaining both endpoints. With nominal capacity fraction $\kappa_\star = 0.30$, each game uses

$$M_c(M) = \mathrm{round}(1 + \kappa_\star(M-1)), \qquad \kappa(M) = \frac{M_c(M)-1}{M-1}. \tag{3.1}$$

Thus the exact finite-player target $\kappa(M)$, rather than the nominal value $\kappa_\star$, is used in both the particle system and the PDE. This keeps the integer-capacity rounding error separate from the numerical particle–PDE discrepancy. For the baseline comparisons, the initial density is a fixed smooth three-bump law,

$$\begin{aligned}\widetilde{f}_0(x;c) = 0.80\exp\left[-\frac{(x-c)^2}{2(0.85)^2}\right] + 0.10\exp\left[-\frac{(x-c+1.20)^2}{2(0.45)^2}\right] \\ + 0.10\exp\left[-\frac{(x-c-1.20)^2}{2(0.45)^2}\right],\end{aligned} \tag{3.2}$$

followed by numerical normalisation. The three-bump shape is used to make the law-level comparison nontrivial: the PDE must track internal structure, not only a single unimodal aggregate. See the supplement SM2 for the baseline-law details. For each branch and each population size, only the centre c is adjusted. A bracketed root solve selects $c = c_\star$ so that

$$\int_{\mathbb{R}} p(x) f_0(x; c_\star)\, dx = \kappa(M). \tag{3.3}$$

Equivalently, the baseline laws lie, up to numerical tolerance, in the capacity-matched set

$$\mathcal{M}_{\kappa(M)} = \left\{ f \geq 0 : \int_{\mathbb{R}} f = 1, \quad \int_{\mathbb{R}} p(x) f(x)\, dx = \kappa(M) \right\}. \tag{3.4}$$

These matched initial laws test preservation near the capacity manifold and local law evolution. Off-capacity perturbations are introduced in Section 3.6 and studied in Section 6. The calibration audit records mass, $|a(0)|$, $\beta(0)$, and the probability-space pushforward $\nu_0 = p_\# f_0$.[1] The largest residual in (3.3) over the 2×100 baseline cases was 8.31×10^{-14}, showing that later aggregate discrepancies are not caused by imperfect initial calibration. Details of the component interpretation, calibration plots, and full case registry are in the supplement SM2.

3.2. Independent deterministic and microscopic implementations.

[1] For measurable $A \subset [0,1]$, $(p_\# f_0)(A) = \int_{p^{-1}(A)} f_0(x)\, dx$; equivalently, $p_\# f_0$ is the distribution of $P = p(X)$ when $X \sim f_0$.

Finite-volume PDE solver. The closed equation (2.12) is solved on a uniform cell-centred grid over $\Omega = [-14, 14]$. This box is wide enough for the calibrated initial laws to start well inside the computational domain, but small enough for repeated refinement and enlarged-domain checks. Boundary influence is monitored by the mass in the outer 10% of the box. Writing $g_i = p(x_i)f_i$, the spatial discretisation is conservative:

$$\frac{df_i}{dt} = -\frac{F_{i+1/2} - F_{i-1/2}}{\Delta x}, \qquad F = c(t)g - d(t)\partial_x g, \tag{3.5}$$

where $c(t)$ and $d(t)$ are the drift and diffusion coefficients in (2.12). The same interface flux is shared by neighbouring cells, and zero flux is imposed at the endpoints. Advection is treated by MUSCL reconstruction with minmod limiting [20, 18], diffusion is centred, and no positivity projection is applied; any negative density is retained as a numerical diagnostic. The flux details are given in Supplementary Section SM3.1. After this spatial discretisation, the cell values satisfy a finite system of ODEs in time. This system is advanced with SciPy's variable-order BDF solver [21]. The independent PDE audit uses $N_x \in \{101, 201, 401, 801\}$ over the common scaled-time horizon $\theta = 2$, with tighter tolerances and smaller maximum steps as the grid is refined. The 801-cell solution is the finest certified solution and is used as the production reference. The refinement audit records moment errors, common-grid L^1 law errors, native-grid window masses, conservation, negative mass, boundary mass, and the empirical order $q = \log(E_h/E_{h/2})/\log 2$, where E_h and $E_{h/2}$ are successive-grid differences for the diagnostic being tested. Thus $q \simeq 2$ means that halving the mesh width reduces the observed error by about a factor of four. The finest-grid differences set the solver scale used later in the operational thresholds: about 3.5×10^{-4} for common-grid law diagnostics and at most a few 10^{-6} for smooth moments. Because probability-space W_1 was not a direct column in this deterministic refinement audit, the later 5×10^{-4} W_1 allowance is used only as a conservative numerical allowance, not as a measured W_1 refinement floor.

Microscopic simulator. The stochastic simulator implements the finite-player update (2.2) directly. Initial particles are chosen as shuffled stratified quantiles of the calibrated density:

$$u_i = \frac{i + 1/2}{M}, \qquad X_0^{\pi(i)} = F_0^{-1}(u_i), \qquad i = 0, \dots, M-1, \tag{3.6}$$

where F_0 is the CDF of f_0 and π is a random permutation. This gives an even initial empirical law without tying quantile order to agent labels. The simulator audit uses both probability branches, all 100 population sizes, four seeds, 100 microscopic rounds, and $\rho \in \{0.00125, 0.00175, 0.0025, 0.0035, 0.005, 0.0075, 0.01, 0.015, 0.02\}$, with reports at rounds $0, 25, 50, 75, 100$. This gives 1800 parameter cases and 7200 seeded trajectories. Since the number of rounds is fixed, the physical horizon is $100\rho^2/M$; the audit checks implementation and jump scaling before the matched particle–PDE study, rather than serving as a common-time convergence test. At each round the code verifies the attendance and update identities from (2.1)–(2.2), $m_n = \sum_i \delta_n^i$, and $X_{n+1}^i - X_n^i = h(M_c - m_n)\delta_n^i$, together with state, action, probability, and attendance bounds. These checks catch errors in attendance counting, update sign, update scale, or the entrant mask before any scientific comparison is made. The Bernoulli

randomisation is checked conditionally on the current propensities through

$$z_n = \frac{m_n - \sum_i P_n^i}{\sqrt{\sum_i P_n^i(1-P_n^i)}}, \qquad P_n^i = p(X_n^i),$$
$$Z_{\text{global}} = \frac{\sum_n m_n - \sum_{n,i} P_n^i}{\sqrt{\sum_{n,i} P_n^i(1-P_n^i)}}. \tag{3.7}$$

The predeclared screens $|\overline{z}| \le 0.20$, $|s_z - 1| \le 0.25$, and $|Z_{\text{global}}| \le 4$ are broad quality-control checks for systematic bias or variance distortion, not formal hypothesis tests. Fixed-seed sentinel cases must also reproduce identical final-state fingerprints and scalar summaries. The numerical outcomes of these checks are reported in Section 4.1.

3.3. Matched clocks, observables, and discrepancies. For particle–PDE comparisons, the microscopic update size, PDE time scale, and reported clock are fixed by

$$h = \frac{\rho}{\sqrt{M}}, \qquad \tau = h^2 = \frac{\rho^2}{M}, \qquad \theta = (M-1)t. \tag{3.8}$$

ρ sets the microscopic jump scale and θ is the scaled time used in the figures and tables. Let

$$\theta_{\max}(\rho) = 0.40 \left(\frac{\rho}{0.02}\right)^2 = 1000\rho^2. \tag{3.9}$$

Here $\theta = 0.40$ is the observation horizon used at the largest tested jump scale, $\rho = 0.02$. Since $\tau = \rho^2/M$, this choice corresponds to approximately $1000M/(M-1)$ microscopic rounds per case. Let S be the number of microscopic seeds. The seed-mean empirical probability law is

$$\widehat{\nu}_M(\theta) = \frac{1}{S}\sum_{s=1}^{S}\left[\frac{1}{M}\sum_{i=1}^{M}\delta_{p(X_s^i(\theta))}\right], \tag{3.10}$$

where δ_y is a unit point mass at y. The empirical moments $\widehat{\alpha}_M$ and $\widehat{\beta}_M$ are averaged over seeds in the same way, and the PDE probability law is $\nu_{\text{PDE}} = p_{\#} f_{\text{PDE}}$. Over a fixed comparison window $\mathcal{W}$, the main discrepancies are

$$D_\alpha = \sup_{\theta\in\mathcal{W}} |\widehat{\alpha}_M(\theta) - \alpha_{\text{PDE}}(\theta)|, \tag{3.11}$$

$$D_\beta = \sup_{\theta\in\mathcal{W}} |\widehat{\beta}_M(\theta) - \beta_{\text{PDE}}(\theta)|, \tag{3.12}$$

$$D_{W_1,p} = \sup_{\theta\in\mathcal{W}} W_1(\widehat{\nu}_M(\theta), \nu_{\text{PDE}}(\theta)). \tag{3.13}$$

The Wasserstein metric is the principal full-law discrepancy because it compares probability laws without choosing histogram bins.[2]

For the finite-time sorting-related comparison, define the probability-interior mass

$$S_\varepsilon^p(\theta) = \int_{\mathbb{R}} \mathbf{1}_{\{\varepsilon \le p(x) \le 1-\varepsilon\}} f(x,\theta)\, dx. \tag{3.14}$$

[2]For one-dimensional laws, we compute $W_1(\mu,\nu) = \int_0^1 |F_\mu^{-1}(u) - F_\nu^{-1}(u)|\, du$.

Its microscopic analogue is the seed-mean fraction of particles whose entry probabilities lie in the same interval. The corresponding discrepancy is

$$D_{S^p_\varepsilon} = \sup_{\theta\in\mathcal{W}} \left| \widehat{S}^p_{\varepsilon,M}(\theta) - S^p_{\varepsilon,\mathrm{PDE}}(\theta) \right|. \tag{3.15}$$

We record $\varepsilon \in \{0.05, 0.10, 0.20\}$ and use $\varepsilon = 0.10$ in the primary behavioural screen. Since p is increasing, this diagnostic measures mass whose entry probabilities remain away from their endpoint behaviour and is therefore sorting-related. It is not identical to the whole-line sorting diagnostic S_R in (2.13), which remains the direct long-time measure used in Section 7. As an auxiliary fixed-bin shape diagnostic, for a partition $B_1, \dots, B_K$ of $[0,1]$ we also record $H_K = \frac{1}{2}\sum_{k=1}^{K} |\widehat{m}_k - m_k|$, where $\widehat{m}_k$ and m_k are the microscopic and PDE masses in B_k. This is a histogram half-L^1 diagnostic, so it is representation-dependent. All comparison windows are fixed before inspecting the discrepancies. The finite-time atlas uses $\theta_{\mathrm{burn}} = \min(0.05, \theta_{\max}/2)$, and the mechanism study uses the final 95% of each realised horizon. These burn-in rules remove a short startup transient without choosing windows after seeing the results.

3.4. Operational adequacy, uncertainty, and censoring. Raw discrepancies are reported before any pass–fail label. For a diagnostic Q, the operational threshold is

$$T^{\mathrm{op}}_Q(M) = \max\left\{ T^{\mathrm{sampling}}_{Q,0.95}(M) + T^{\mathrm{PDE}}_Q,\ T^{\mathrm{practical}}_Q \right\}. \tag{3.16}$$

Here $T^{\mathrm{sampling}}_{Q,0.95}$ is the 95th percentile from 400 matched samples of the reference law, T^{PDE}_Q is the numerical-resolution allowance, and $T^{\mathrm{practical}}_Q$ is the predeclared accuracy margin. The practical margins are design choices, not results:

$$T^{\mathrm{practical}}_\alpha = T^{\mathrm{practical}}_\beta = 0.01,\ \ T^{\mathrm{practical}}_{W_1,p} = 0.02,$$

$$T^{\mathrm{practical}}_{H_{40}} = 0.10,\ \ T^{\mathrm{practical}}_{H_{100}} = 0.15,\ \ T^{\mathrm{practical}}_{S^p_{0.10}} = 0.05.$$

The corresponding PDE allowances are $10^{-4}, 10^{-4}, 5\times10^{-4}, 5\times10^{-4}, 5\times10^{-4}$, and 10^{-3}. We display $R_Q = D_Q / T^{\mathrm{op}}_Q(M)$ with the raw discrepancy D_Q. The ratio $R_Q > 1$ means that the observed discrepancy exceeds the declared operational resolution at that population size. Since the sampling part of the threshold decreases with M, R_Q can increase even when D_Q decreases; reporting both keeps detectability separate from absolute error. Uncertainty is handled at the seed and case levels. Two-seed atlas ranges are descriptive only. Cross-case regressions treat each parameter case as one observation and use HC3 heteroskedasticity-robust standard errors [12]; fits based on only three population sizes are labelled exploratory. Off-capacity experiments start with $a(0) \neq 0$. Persistent entry into a tolerance ε is defined by

$$T_\varepsilon = \inf\left\{ \theta : |a(s)| \le \varepsilon \text{ for every stored } s \in [\theta, \theta_{\max}] \right\}. \tag{3.17}$$

A trajectory must enter the tube and remain there. If entry is not observed by $\theta_{\max}$, the event time is right-censored rather than set to the terminal horizon. We use $\varepsilon = 10^{-3}$ and a case-matched noise-aware tolerance estimated from independent zero-offset microscopic controls.

3.5. Mechanism-resolving diagnostics. The study separates sources of finite-time law discrepancy: aggregate feedback through $a(t)$ and $b(t)$, finite-player dependence beyond independent sampling, realised one-round changes in $p(X)$, and weak-generator approximations.

Oracle intervention and corrected dependence. The oracle calculation keeps the finite volume operator fixed but replaces the PDE-generated paths $a(t)$ and $b(t)$ by piecewise-linear interpolants of the microscopic seed means. For discrepancy D_Q, define

$$I_Q^{\mathrm{abs}} = D_Q^{\mathrm{closed}} - D_Q^{\mathrm{oracle}}, \qquad I_Q^{\mathrm{rel}} = \frac{I_Q^{\mathrm{abs}}}{D_Q^{\mathrm{closed}}}, \quad D_Q^{\mathrm{closed}} \neq 0. \tag{3.18}$$

Positive I_Q^{rel} means that supplying the microscopic aggregate paths improves the comparison. To measure dependence, probability space is partitioned into bins $B_1, \dots, B_K$. At a fixed report time in seed s, define

$$N_{sk} = \sum_{i=1}^{M} \mathbf{1}_{\{p(X_s^i) \in B_k\}}, \qquad \widehat{\nu}_s = \left(\frac{N_{s1}}{M}, \dots, \frac{N_{sK}}{M} \right). \tag{3.19}$$

The independent multinomial baseline and corrected excess covariance are

$$B_M(\overline{\nu}) = \frac{\mathrm{diag}(\overline{\nu}) - \overline{\nu}\overline{\nu}^{\mathsf{T}}}{M}, \qquad C_{\mathrm{ex}} = \widehat{C}_{\mathrm{seed}} - B_M(\overline{\nu}), \qquad \chi_K^+ = \sum_{j,k} \max\{(C_{\mathrm{ex}})_{jk}, 0\}, \tag{3.20}$$

with seed mean $\overline{\nu}$ and covariance $\widehat{C}_{\mathrm{seed}}$. χ_K^+ records positive bin co-fluctuation beyond what an independent finite sample would already produce. The primary resolution is $K = 20$, with $K = 10, 40$ used for resolution checks; the reliability screen is at least 6 seeds and $M \geq 5K$.

Direct jumps and weak generators. The realised probability jump is $\Delta p_n^i = p(X_{n+1}^i) - p(X_n^i)$, with $J_{90} = Q_{0.90}(|\Delta p_n^i|)$ as the primary upper-tail coordinate. Since non-entrants do not update, move frequencies and conditional-nonzero summaries are also reported. Large jump streams are stored by uniform one-pass reservoir sampling [22], with sentinel cases accumulated exactly. The weak-generator audit is a local one-step diagnostic. For a test function ψ, it compares successive approximations to $G\psi = \tau^{-1}\mathbb{E}[\psi(X_{n+1}) - \psi(X_n) \mid X_n]$. The hierarchy is

$$G_{\mathrm{micro}} \longrightarrow G_{\mathrm{LOO}} \longrightarrow G_{\mathrm{nonlocal}} \longrightarrow G_{\mathrm{KM2}} \longrightarrow G_{\mathrm{PDE}}. \tag{3.21}$$

Here G_{micro} is the exact finite-player one-step generator; G_{LOO} replaces heterogeneous leave-one-out attendance by a binomial law with the same leave-one-out mean; G_{nonlocal} replaces focal leave-one-out information by global closed-model moments while retaining the finite shift; G_{KM2} applies the second-order Kramers–Moyal expansion; and G_{PDE} evaluates the coded PDE weak form. The corresponding telescoping identity is

$$\begin{aligned} G_{\mathrm{micro}} - G_{\mathrm{PDE}} = {} & \underbrace{G_{\mathrm{micro}} - G_{\mathrm{LOO}}}_{\text{attendance heterogeneity}} + \underbrace{G_{\mathrm{LOO}} - G_{\mathrm{nonlocal}}}_{\text{leave-one-out/global replacement}} \\ & + \underbrace{G_{\mathrm{nonlocal}} - G_{\mathrm{KM2}}}_{\text{finite-jump truncation}} + \underbrace{G_{\mathrm{KM2}} - G_{\mathrm{PDE}}}_{\text{implementation identity}}. \end{aligned} \tag{3.22}$$

This decomposition is local to one smooth observable and one stored state; it is not an additive decomposition of a global metric such as W_1. Exact Poisson–binomial quadrature is used for $M \leq 300$, with independent Monte Carlo one-step estimates as implementation checks. Further details of the generator construction are given in Supplementary Section SM3.2.

3.6. Off-capacity attraction and finite-population noise. The capacity-matched laws in (3.4) have $a(0) \approx 0$, so they test preservation near the capacity manifold. To test attraction, separate frozen experiments start from prescribed aggregate displacements, $\alpha_0^{\text{target}} = \kappa(M) + \delta_\alpha^{\text{req}}$, where $\delta_\alpha^{\text{real}} = \int_{\mathbb{R}} p(x) f_0(x)\, dx - \kappa(M) = -a(0)$. The realised displacement is recorded because finite-grid and finite-particle initialisation can slightly change the requested value. Return to capacity is measured by $C_a = \frac{|a(\theta_{\max})|}{|a(0)|}$, $a(0) \neq 0$. Thus $C_a < 1$ indicates contraction toward capacity, and $C_a \ll 1$ indicates strong return. We also record sign-correct return and persistent entry into a tolerance tube around $a = 0$, using (3.17). Zero-offset controls are handled separately because C_a is undefined when $a(0) \approx 0$. The finite-population bridge uses replicated microscopic cases at $M \in \{30, 100, 300\}$, with case-matched PDE trajectories. These runs compare aggregate paths, W_1, and S_R, while independent zero-offset controls estimate the finite-population noise floor. Persistent entry is reported both for $\varepsilon = 10^{-3}$ and for a case-matched noise-aware tolerance. The results are analysed in Section 6.

3.7. Long-time diagnostics and theorem connection. Long-time runs store the observables needed for the aggregate–sorting separation discussed in Section 2.4: $|a(\theta)|$, $\beta(\theta)$, and $S_R(\theta)$. We also store the proof-facing quantity $\phi(\theta) = \beta(\theta) g_2(\theta)^3$, where $g_2(\theta) = \int_{\mathbb{R}} (p(x) f(x, \theta))^2\, dx$, for comparison with the analytical sorting argument [4]. Domain safety is monitored by the edge mass in the outer 10% of the computational box, with headline trajectories repeated on enlarged domains. The long-time results are reported in Section 7.

3.8. Sensitivity and reproducibility. Sensitivity checks vary grid/domain, report times, bins, seeds, reservoir size, derivative step, Monte Carlo repetitions, and analysis window. Failed, skipped, censored, and excluded cases remain in audit tables; invalid regressions are omitted. Run provenance records configurations, seeds, outputs, checksums, and scripts, and fixed-seed repeats must reproduce identical results.

4. Certified finite-time validity is observable dependent. Finite-time agreement depends on the observable: the PDE may reproduce mean entry and within-agent uncertainty without matching bounded-window mass or the complete probability law. We therefore audit the PDE solver and microscopic simulator independently before their matched comparison.

4.1. Independent numerical verification. Throughout this section, a numerical result is described as "certified" when the corresponding implementation passed the predeclared convergence, conservation, boundedness, reproducibility, and calibration checks in Section 3.2. These preliminary audits ensure that a later particle–PDE discrepancy is not simply an artefact of a nonconverged PDE solver or an incorrectly implemented particle update. The four-grid refinement procedure and empirical order q are defined in Section 3.2. Across the primary moment and law diagnostics in fourteen branch–population cases, $1.9785 \leq q \leq 2.0132$. Thus halving the mesh width reduced successive-grid differences by approximately a factor of four. This near-second-order self-convergence shows that the deterministic numerical error contracts predictably under the production space–time refinement protocol and supports

using the 801-cell solution as the finite-time reference. It is an empirical result for the coupled refinement protocol. The remaining PDE checks concern the computed density f. The mass defect $\left|\int_\Omega f(x,t)\,dx - 1\right|$ was at most 2.00×10^{-15}, while the integrated negative part $\int_\Omega \max\{-f(x,t), 0\}\,dx$ remained zero and therefore below tolerance. The mass in the monitored outer part of the finest computational domain was at most 5.68×10^{-26}. These checks show that the numerical density remains normalised, does not develop resolved negative values, and is not being distorted by the artificial boundary; see Supplementary Table SM3. The microscopic simulator was audited independently over 1800 parameter cases and 7200 seeded trajectories. For the update identity, the residual $r_{\mathrm{upd}} = \max_{n,i} \left|(X^i_{n+1} - X^i_n) - h(M_c - m_n)\delta^i_n\right|$ should vanish algebraically. Its observed maximum, 4.44×10^{-16}, is consistent with floating-point roundoff. This is a code-integrity check: an error in indexing, broadcasting, update order, or the entrant mask would make the residual appreciably nonzero. The attendance identity was satisfied exactly. The Bernoulli residuals defined in (3.7) also satisfied $\max|Z_{\mathrm{global}}| = 3.06 < 4$, where 4 is the predeclared calibration screen. Because the states evolve adaptively, this is a quality-control diagnostic for systematic simulation bias rather than an independent-sample hypothesis test. The complete microscopic audit is reported in Supplementary Table SM4. The dependence of realised one-round probability changes on M and ρ is a mechanism question rather than a code-integrity check and is analysed in Section 5.3.

4.2. Aggregate and sorting-related accuracy. The matched finite-time atlas contains 2 branches $\times$ 9 ρ-values $\times$ 21 M-values $=$ 378 particle–PDE cases, with $30 \le M \le 30000$. Because $\theta_{\max} = 1000\rho^2$ in (3.9), larger ρ means both stronger jumps and a longer scaled observation window. Comparisons across ρ therefore measure their combined effect, not jump size alone. The observables α and β are defined in (2.9)–(2.10), and their discrepancies in (3.11)–(3.12). Figure 1(a)–(b) shows both errors against M for representative ρ values. They broadly decrease with population size and are largest under the high-ρ protocol. The light bands are two-seed minimum–maximum ranges, not confidence intervals. Supplementary Figure SM4 gives the complementary view against ρ: its influence is strongest at small M, whereas at large M the aggregate errors are smaller and $D_{W_1,p}$ is nearly flat in ρ. For each branch and ρ, Table 1 first takes the median over the 21 population sizes and then reports the range over the nine ρ rows. Even the upper endpoints of the median D_α and D_β ranges remain below the predeclared practical margins $T_\alpha^{\mathrm{practical}} = T_\beta^{\mathrm{practical}} = 0.01$ from Section 3.4. Supplementary Figure SM3(a) confirms the small-M/large-ρ concentration over the complete atlas. A case is behaviourally adequate when D_α, D_β, and $D_{S^p_{0.10}}$ from (3.15) all satisfy their operational thresholds (3.16). This tests finite-time particle–PDE agreement, not attainment of the long-time sorting limit. Branch A passes all 189 cases, whereas Branch B passes 179 of 189. The α and β errors pass throughout [Supplementary Figure SM3(b)], so the only failing component of the Branch B behavioural screen is the probability-interior mass (3.15). In those ten cases, the microscopic and PDE fractions with $0.10 \le p(X) \le 0.90$ differ beyond the threshold. Figure 1(d) reports the number of passing populations out of 21 at each ρ.

4.3. Finite-population excess and residual law discrepancy. Section 4.2 established close agreement for the aggregate observables. We now ask the stronger question of whether the PDE reproduces the complete distribution of entry probabilities. The microscopic empirical law, PDE pushforward, and their worst-over-window Wasserstein discrepancy $D_{W_1,p}$ are

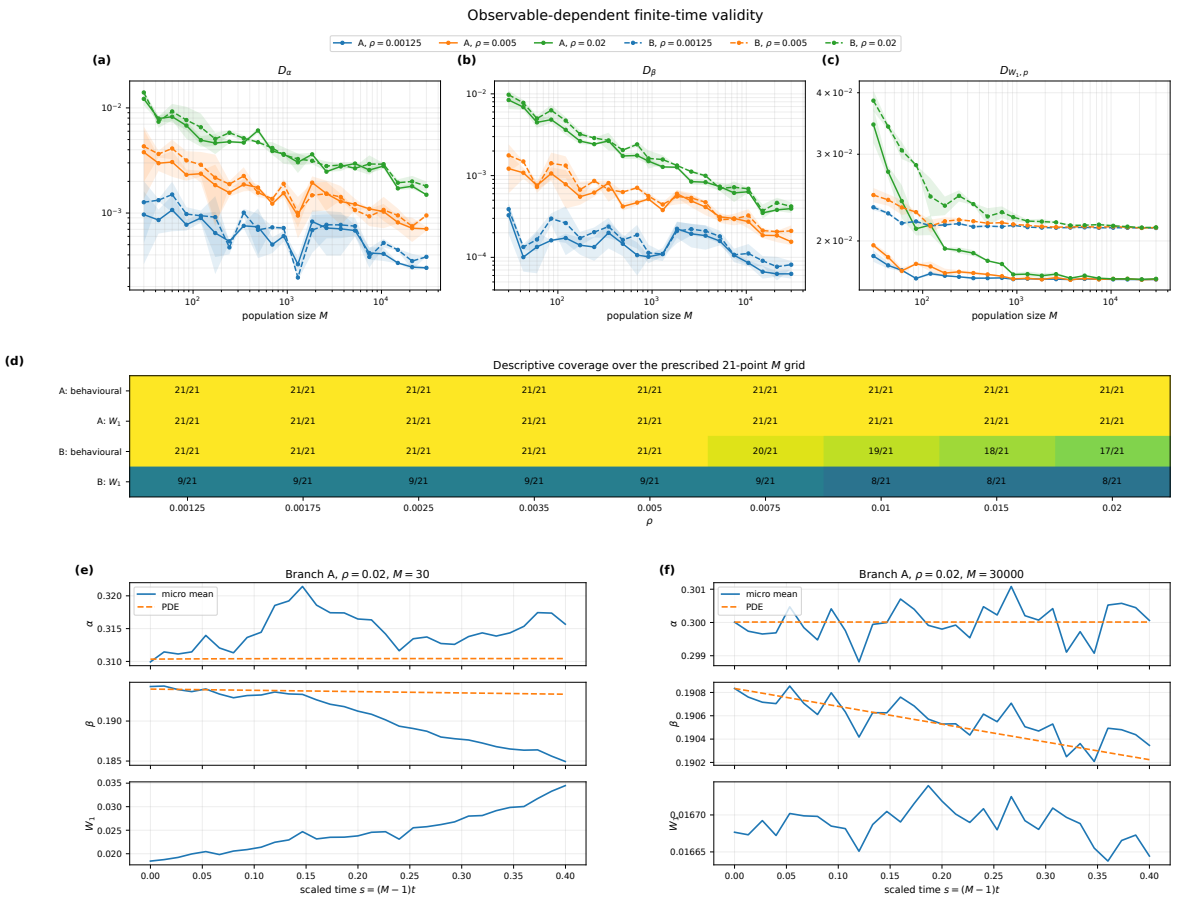


Figure 1. *Observable-dependent finite-time validity. Panels (a)–(c) show D_α, D_β, and $D_{W_1,p}$ against population size for representative ρ values. Colours encode ρ, line style distinguishes Branches A and B, and the light bands are two-seed minimum–maximum ranges. Panel (d) reports coverage over the prescribed M grid: each row is one branch–criterion pair, each column fixes ρ, and each cell gives the number of passing populations out of* 21*. Panels (e)–(f) compare Branch A at $\rho = 0.02$ for $M = 30$ and $M = 30000$.*

Table 1

Finite-time summary over the 9 ρ rows. Each discrepancy range is the minimum and maximum, over ρ, of the median across the 21 *population sizes. Coverage gives the total number of passing cases, followed in parentheses by the least favourable ρ-row coverage. The W_1 criterion additionally requires behavioural adequacy.*

Branch	median D_α range	median D_β range	median $D_{W_1,p}$ range	behavioural coverage	W_1 coverage
A	6.82×10^{-4}–3.641×10^{-3}	1.34×10^{-4}–1.504×10^{-3}	0.016751–0.017186	189/189 (21/21)	189/189 (21/21)
B	7.31×10^{-4}–3.581×10^{-3}	1.72×10^{-4}–1.624×10^{-3}	0.021375–0.022356	179/189 (17/21)	78/189 (8/21)

defined in (3.10)–(3.13). Panels 1(e)–(f) illustrate the distinction. At $M = 30$, discrepancies are visible in both the aggregate trajectories and $W_{1,p}$. At $M = 30000$, α and β follow the PDE much more closely, while $W_{1,p}$ remains nonzero. This is possible because α and β are only two functionals of the probability law and do not determine it uniquely: $\alpha_{\text{micro}} \approx \alpha_{\text{PDE}}$, $\beta_{\text{micro}} \approx \beta_{\text{PDE}}$, $W_1\left(\nu^{\text{micro}}, \nu^{\text{PDE}}\right) > 0$. Since both systems use the same

$\kappa(M)$, agreement in α also implies agreement in the capacity error $a = \kappa - \alpha$. Thus accurate capacity-restoring dynamics do not require exact agreement of the complete probability laws. Figure 1(c) shows how the law discrepancy changes with population size. At small M, $D_{W_1,p}$ is elevated, especially under the high-ρ protocol. It then decreases towards an approximately flat branch-dependent level. We call the small-M elevation above this level the *finite-population excess*. The complete atlas in Supplementary Figure SM3(a) confirms this pattern, and Table 1 reports the branchwise median ranges. Increasing M therefore reduces the finite-population excess, but not the entire measured law discrepancy. A nonzero raw W_1 value is not automatically an operational failure. The threshold (3.16) includes an allowance for sampling variability in the finite microscopic law. This allowance is largest at small M and decreases as M grows. Branch A remains below the nested W_1 threshold in all 189 cases, so its nonzero law discrepancy is operationally acceptable throughout the atlas. Branch B satisfies the criterion in only 78 of 189 cases. Its ρ-specific median discrepancies, 0.021375–0.022356, lie slightly above the practical margin 0.02 (Table 1); as the sampling allowance contracts, this level becomes detectable. The coverage boundary in Figure 1(d) therefore reflects increasing resolution of the test, rather than a sudden growth of the raw discrepancy. For completeness, the fixed-bin diagnostic compares the masses assigned to the same probability bins: $D_{\mathrm{hist}} = \frac{1}{2}\sum_j \left| H_j^{\mathrm{micro}} - H_j^{\mathrm{PDE}} \right|$. Because a small displacement across a bin edge can change this quantity substantially, it depends strongly on the chosen binning. We therefore retain it only as a representation-sensitive shape check; see Supplementary Figure SM3(c). These results establish a law-level discrepancy but leave its origin unresolved. Section 5 goes beyond pass–fail validation to a mechanism-resolving diagnosis: by isolating aggregate-feedback closure, finite-player dependence, and finite-jump and local-generator approximations, it tests which approximation steps materially create or amplify the discrepancy, and which do not.

5. Mechanisms of the finite-time law discrepancy. Section 4.3 showed that close agreement in α and β does not imply equality of the microscopic and PDE probability laws. Throughout this section, $D_{W_1,p}^{\mathrm{closed}}$ denotes the discrepancy (3.13) for the unmodified closed PDE, evaluated over the final 95% of each realised horizon as specified in Section 3.3. It measures the law-level difference but does not explain its origin. As reviewed in Section 2.3, two approximation steps are central. First, the full M-particle Kolmogorov equation is expanded to second order in h: contributions of orders h/τ and h^2/τ are retained, while $O(h^3)/\tau$ is discarded under $h, \tau \ll 1$ and $h^2/\tau \sim 1$; see Section 3 of [4]. Second, the marginal equation is closed through $g(x,y,t) = f(x,t)f(y,t)$, $l(x,y,z,t) = f(x,t)f(y,t)f(z,t)$, as in the molecular-chaos closure of Section 3 of [4]. These are, respectively, the second-order Kramers–Moyal truncation and the molecular-chaos closure. The measured W_1 difference may also include finite empirical sampling, probability-map geometry, and accumulation over many rounds. We therefore test specific mechanisms by asking four questions. *First, is there finite-player dependence beyond sampling noise?* The molecular-chaos factorisation treats particles as independent, but finite bin counts fluctuate even under independence. Section 5.1 subtracts this multinomial sampling covariance from the observed seed-to-seed covariance. Any excess that remains is evidence, within this binned diagnostic, of dependence not represented by the closure. *Second, do inaccurate aggregate paths explain the law discrepancy?* The oracle in Section 5.2 changes only the PDE-generated coefficient paths $a(t)$ and $b(t)$, replacing them by their microscopic

seed means. A reduction in $D_{W_1,p}$ would show that aggregate-feedback error explains part of the law discrepancy; otherwise, those paths are not sufficient to explain it. *Third, how large are the finite microscopic jumps?* The update produces a propensity shift and a corresponding probability change, $\Delta x_n^i = h(M_c - m_n)\delta_n^i$; $\Delta p_n^i = p(X_n^i + \Delta x_n^i) - p(X_n^i)$; $\psi(x+\Delta x) - \psi(x) = \psi'(x)\Delta x + \frac{1}{2}\psi''(x)(\Delta x)^2 + O\big(|\Delta x|^3\big)$. The Kramers–Moyal approximation expands the finite shift Δx, whereas Δp records how that same shift appears in the probability space where W_1 is measured. Thus $|\Delta p|$ indicates the realised jump scale relevant to the law comparison. Section 5.3 measures these jumps, and Section 5.4 compares the jump and dependence diagnostics with each other and with $D_{W_1,p}^{\mathrm{closed}}$. *Fourth, which approximation changes the local dynamics most?* At a fixed microscopic state and for a fixed smooth observable, the weak-generator audit in Section 5.5 changes one step of the hierarchy (3.21) at a time to separate the effects of attendance heterogeneity, the closure-related replacement of focal leave-one-out information by global moments, and the truncation of the exact finite shift at second order.

5.1. Corrected finite-player dependence. At a fixed report time, each microscopic seed produces the probability-bin fractions $\widehat{\nu}_s$ defined in (3.19). Their covariance $\widehat{C}_{\mathrm{seed}}$ measures how the bin occupancies vary across seeds. Even if the particles were perfectly independent, two finite samples of M particles would not place exactly the same numbers of particles in every bin. Thus $\widehat{C}_{\mathrm{seed}}$ contains ordinary sampling variability as well as any dependence between particles. The matrix $B_M(\overline{\nu})$ in (3.20) is the theoretical covariance predicted for M independent draws from the common bin law $\overline{\nu}$. It is therefore the finite-bin independence benchmark corresponding to the molecular-chaos factorisation in equation (3.7) of [4]. The correction $C_{\mathrm{ex}} = \widehat{C}_{\mathrm{seed}} - B_M(\overline{\nu})$,(3.20), removes this predicted sampling covariance from the observed covariance. Under perfect independence, C_{ex} should be close to zero, apart from finite-seed estimation noise. A systematic residual instead identifies binned dependence not represented by the closure. The score χ_K^+ in (3.20) sums the positive entries of C_{ex}, preventing positive and negative residuals from cancelling. A large value therefore means that some bin masses co-fluctuate more strongly than independent sampling predicts. For the primary resolution $K = 20$, each (branch, ρ, M) case is summarised over the post-startup window by

$$\chi_{\mathrm{ex}}^+ = Q_{0.90}\big(\{\chi_{20}^+(\theta) : \theta \in \mathcal{W}\}\big)\,. \tag{5.1}$$

The 90th percentile represents a repeatedly reached upper-tail level without allowing one isolated maximum to determine the case. The main question is whether this excess over the molecular-chaos benchmark decreases as M grows. We estimate the trend with

$$\chi_{\mathrm{ex}}^+ \approx C_\chi M^{\gamma_M} \rho^{\gamma_\rho} \exp\big(\gamma_B \mathbf{1}_{\{B\}}\big)\,, \tag{5.2}$$

where C_χ is a scale constant, γ_M measures the population trend, γ_ρ measures the stress dependence, and $\exp(\gamma_B)$ is the Branch B-to-A multiplier. A common γ_M tests one population trend across both branches, while γ_B allows their fitted levels to differ. This is a descriptive fit to the data: after taking logarithms, we fit it by ordinary least squares and report HC3 robust uncertainty [12]. Across the 104 strictly positive cases, $\widehat{\gamma}_M = -1.001017$, with HC3 standard error 0.003876 and 95% interval $[-1.008708, -0.993326]$ (Table SM5). The interval contains -1 and excludes zero, supporting inverse population scaling rather than a flat trend.

At fixed ρ and branch, doubling M multiplies the fitted score by $2^{\widehat{\gamma}_M} = 2^{-1.001017} = 0.4996$. Thus doubling the population approximately halves the corrected dependence excess. By comparison, doubling ρ multiplies the score by $2^{0.129224} \simeq 1.094$, an increase of 9.4%, while Branch B is $\exp(0.049603) \simeq 1.051$ times Branch A at the same M and ρ (Table SM5). Population size therefore has the largest fitted effect on these matched comparisons. The nearly linear log–log decline in Figure 2(a) has $R^2 = 0.998718$. Robustness checks using the occupancy restriction, alternative bin resolutions $K = 10, 40$, and leave-one-ρ-out fits all preserve the near-M^{-1} trend; Figure SM5. Hence the result is not driven by sparse bins, one stress level, or the primary bin width. This is empirical evidence that the specific binned finite-player dependence omitted by the molecular-chaos closure weakens over the audited population range.

5.2. Does supplying microscopic aggregate paths repair the law?. The closed PDE computes its coefficient paths from its own evolving law, (2.9). The oracle intervention of Section 3.5 reruns the same PDE from the same calibrated initial law, on the same grid, with the same finite-volume discretisation and time integrator, changing only the coefficient paths:

$$\begin{aligned} a_{\mathrm{PDE}}(t) = \kappa(M) - \alpha_{\mathrm{PDE}}(t) \quad &\longrightarrow \quad a_{\mathrm{oracle}}(t) = \kappa(M) - \overline{\alpha}_{\mathrm{micro}}(t), \\ b_{\mathrm{PDE}}(t) = \alpha_{\mathrm{PDE}}(t)\big(1 - \alpha_{\mathrm{PDE}}(t)\big) \quad &\longrightarrow \quad b_{\mathrm{oracle}}(t) = \overline{b}_{\mathrm{micro}}(t). \end{aligned}$$

Here the bars denote means over the same eight microscopic seeds; both paths are interpolated linearly. These realised paths would be unavailable to a stand-alone PDE prediction, hence the term *oracle*. The test deliberately favours the aggregate-path explanation: if errors in $a(t)$ and $b(t)$ drive the law discrepancy, supplying their observed microscopic paths should reduce it. We measure this using $I^{\mathrm{rel}}_{W_1}$, (3.18), for which positive values indicate improvement. Because the same eight seeds supply both the paths and the microscopic comparison law, this is an in-sample mechanism test rather than a predictive validation. Across the 104 paired cases, the median response is -2.314×10^{-3}, a 0.23% deterioration. The upper quartile is -1.95×10^{-4}, so at least 75% of cases do not improve; even the 90th percentile is only 2.52×10^{-4}, or 0.025% improvement (Table SM5). Accordingly, the points in Figure 2(b) lie mostly on or above the equality line. Supplying the observed aggregate paths therefore does not systematically repair the probability law: aggregate-path error alone is insufficient to explain the discrepancy.

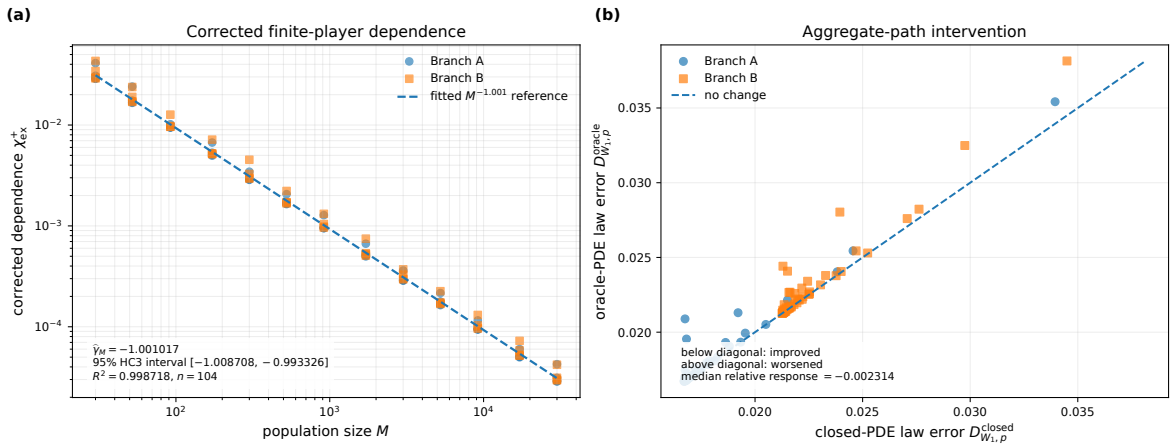


Figure 2. *Corrected finite-player dependence and aggregate-path intervention.* ***(a)*** *The corrected dependence score χ^{+}_{ex} versus population size M, with one point per* $(\mathrm{branch}, \rho, M)$ *case.* ***(b)*** *Closed-PDE law error versus oracle law error. The dashed diagonal denotes no change; points below it improve.*

5.3. Realised one-round probability changes. For the realised probability jump $\Delta p^i_n = p(X^i_{n+1}) - p(X^i_n)$, the primary case statistic over the post-startup window, defined as the final 95% of each realised horizon in Section 3.3, is the mean of the six seed-level 90th percentiles,

$$J_{90} = \frac{1}{6}\sum_{s=1}^{6} Q^{(s)}_{0.90}(|\Delta p|)\,. \tag{5.3}$$

For large streams, each seed is represented by a uniform one-pass reservoir of at most 10^5 observations [22]. Supplementary Figure SM6(d) shows that the relative seed-to-seed dispersion does not increase systematically with source-stream size. By the microscopic update (2.2), non-entrants have $\Delta p^i_n = 0$; an entrant also has zero change when $m_n = M_c$. Consequently, the unconditional jump law has a substantial atom at zero: approximately 70%–75% of the observations in the representative cases of Supplementary Figure SM6(a) are zero. The primary statistic J_{90} retains them and therefore reflects both update frequency and move size. As a sensitivity analysis, J^{+}_{90} removes the zeros before the seed-level 90th percentiles are computed, thereby isolating the size of nonzero moves; the two definitions are compared in Supplementary Figure SM6(b). The scaling $h = \rho/\sqrt{M}$ in (3.8) is designed to offset attendance fluctuations of order $\sqrt{M}$ near capacity. In that regime, $\Delta x = h(M_c - m_n)\delta^i_n$ is of order ρ. The question is whether this cancellation survives the nonlinear map p: if it does, $|\Delta p|$ should grow nearly proportionally to ρ while showing little residual dependence on M. The first two conditional moments of the same finite increment generate the drift and diffusion retained in the Kramers–Moyal derivation [4]; measuring $|\Delta p|$ records its realised size in the probability space where W_1 is evaluated. For $J \in \{J_{90}, J^{+}_{90}\}$, we fit $J \approx C_J M^{\eta_M} \rho^{\eta_\rho} \exp\big(\eta_B \mathbf{1}_{\{B\}}\big)$, where C_J is a scale constant, η_M and η_ρ are the population and jump-stress exponents, and $\exp(\eta_B)$ is the fitted Branch B-to-A multiplier. The log–linear regressions use HC3 robust uncertainty. For the primary unconditional statistic, Table SM6 gives $\widehat{\eta}_\rho = 0.992954$, with HC3 95% confidence interval $[0.984416, 1.001493]$. Thus doubling ρ multiplies the fitted J_{90} by $2^{0.992954} = 1.990$, an

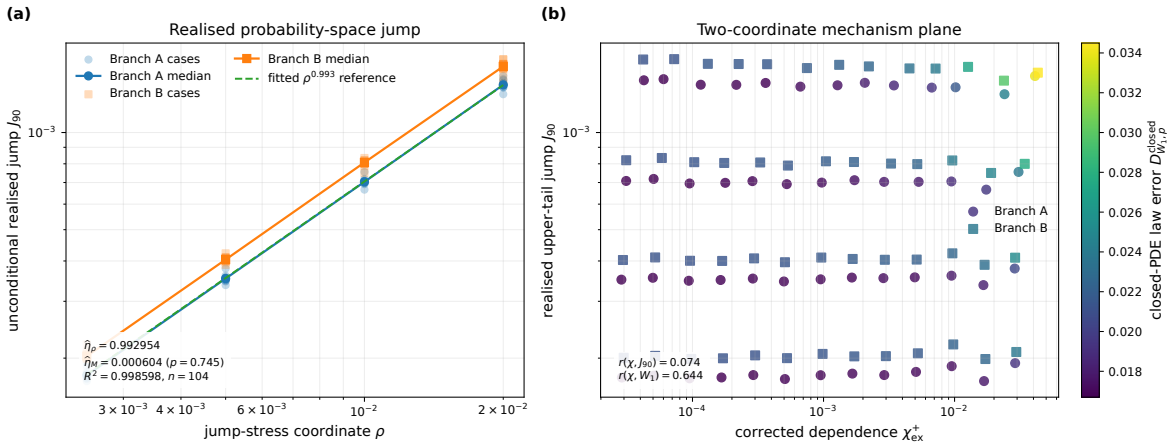


Figure 3. ***(a)*** *Primary unconditional J_{90} versus ρ. Faint points show all* 104 *cases and curves show branch medians. Doubling ρ nearly doubles J_{90}, while the fitted population effect is negligible (Table SM6).* ***(b)*** *J_{90} versus χ^+_{ex}, coloured by $D^{\mathrm{closed}}_{W_1,p}$; symbols distinguish the branches. Larger M tends to move cases left and larger ρ upward. The two axes have correlation* 0.0737*, confirming that they order the cases differently; the colour overlay shows association rather than an additive error decomposition (Supplementary Figure SM7).*

approximately 99% increase. By contrast, $\widehat{\eta}_M = 0.000604$, with interval $[-0.003072, 0.004280]$ and $p = 0.745$; doubling M changes the fitted statistic by only $2^{0.000604} = 1.0004$, or 0.04%. This negligible population effect is the intended consequence of the factor $M^{-1/2}$ in h. Finally, $\exp(0.133054) = 1.142$, so at fixed M and ρ, Branch B has a 14.2% larger fitted upper-tail jump than Branch A. These relations are displayed in Figure 3(a). Removing the zeros gives the same leading picture. The sensitivity fit has $\widehat{\eta}^+_\rho = 0.994065$, with HC3 95% confidence interval $[0.987576, 1.000554]$, so doubling ρ again nearly doubles the fitted jump. Its population exponent is $\widehat{\eta}^+_M = -0.004653$, with interval $[-0.007420, -0.001887]$ and $p = 0.00119$: doubling M multiplies J^+_{90} by 0.9968, a 0.32% reduction; although the confidence interval excludes zero, the population effect is practically small. The Branch B multiplier is $\exp(0.119307) = 1.127$, corresponding to a 12.7% increase (Table SM6). Hence the treatment of zero updates does not change the conclusion: ρ controls the realised upper-tail jump almost one-for-one, Branch B produces moderately larger jumps, and the frozen scaling removes nearly all direct population dependence. This result identifies the scale of the finite shift that the PDE must represent, but not the truncation error itself. J_{90} measures the total realised probability change, including the first- and second-order effects retained by the drift–diffusion approximation. The generator component $G_{\mathrm{nonlocal}} - G_{\mathrm{KM2}}$ in Section 5.5 instead measures the omitted third- and higher-order contribution for smooth observables. The direct-jump result therefore establishes that ρ sets the realised probability-space shift; the generator audit determines whether approximating that shift to second order is an important local error.

5.4. Dependence, jump size, and law error. Figure 3(b) plots the corrected dependence score χ^+_{ex} against the unconditional jump statistic J_{90}, with colour recording $D^{\mathrm{closed}}_{W_1,p}$. The fitted scalings organise the plane: because $\chi^+_{\mathrm{ex}} \sim M^{-1}$, larger populations tend to move cases left, whereas $J_{90} \sim \rho$ moves higher-stress cases upward. Across the 104 cases, $\mathrm{Corr}(\chi^+_{\mathrm{ex}}, J_{90}) =$

0.0737 and $\mathrm{Corr}(\chi^{+}_{\mathrm{ex}}, J^{+}_{90}) = 0.0872$. These low correlations, Supplementary Figure SM7(a), show that the dependence and jump diagnostics order the cases differently. This is consistent with what they measure: χ^{+}_{ex} probes finite-player dependence omitted by the molecular-chaos closure, whereas J_{90} records the finite shift approximated by the Kramers–Moyal expansion [4]. Their weak co-movement therefore suggests distinct empirical signatures rather than duplicate measurements; it does not establish independence or causal shares of the law error. Both diagnostics are nevertheless positively associated with the law error: $\mathrm{Corr}(\chi^{+}_{\mathrm{ex}}, D^{\mathrm{closed}}_{W_1,p}) = 0.6444$, $\mathrm{Corr}(J_{90}, D^{\mathrm{closed}}_{W_1,p}) = 0.3137$, and $\mathrm{Corr}(J^{+}_{90}, D^{\mathrm{closed}}_{W_1,p}) = 0.3183$. Thus stronger corrected dependence and larger jumps both tend to accompany a larger discrepancy, Supplementary Figure SM7(a). To test what each diagnostic adds beyond the design variables, we fit

$$D^{\mathrm{closed}}_{W_1,p} = \beta_0 + \beta_M \log M + \beta_\rho \log \rho + \beta_B \mathbf{1}_{\{B\}} + \beta_\chi \chi^{+}_{\mathrm{ex}} + \beta_J J_{90} + \varepsilon. \tag{5.4}$$

with one observation per case and Branch A as the reference. Here β_0 is the intercept, β_M and β_ρ control for population and stress, β_B is the Branch B-to-A shift, and β_χ and β_J are the adjusted associations with corrected dependence and jump size. The residual ε is the unexplained case variation. The design-only baseline $(\log M, \log \rho, \mathbf{1}_{\{B\}})$ has $R^2 = 0.697708$. Adding χ^{+}_{ex} raises R^2 to 0.885224, an increase of 18.75 percentage points; adding J_{90} raises it only to 0.887941, a further 0.27 percentage points (Supplementary Figure SM7(b) and Supplementary Table SM7). Corrected dependence therefore adds substantial case-to-case information beyond branch, M, and ρ. The small final increase does not make jumps dynamically irrelevant: because J_{90} is nearly proportional to ρ, it adds little new case ordering once $\log \rho$ is present. The full coefficients and HC3 intervals are reported in Supplementary Table SM7. The branch term answers a different question. At the same M, ρ, χ^{+}_{ex}, and J_{90}, the fitted Branch B law error exceeds the Branch A error by $\widehat{\beta}_B = 0.004413$, with HC3 95% confidence interval $[0.003800, 0.005027]$ (Supplementary Table SM7). Thus an otherwise matched Branch B case is predicted to have approximately 0.0044 more absolute probability-space W_1 error. For scale, 0.0044 is about 26% of the representative unadjusted Branch A level near 0.017 in Section 4.3. This shows that the remaining branch gap is substantial. Corrected dependence and jump size therefore do not exhaust why Branch B is harder for the closed PDE to match. The design cannot isolate one cause because changing branch simultaneously changes the probability floor, derivatives and curvature of p, visited states, and pushforward geometry; doing so would require probability maps that vary these features separately.

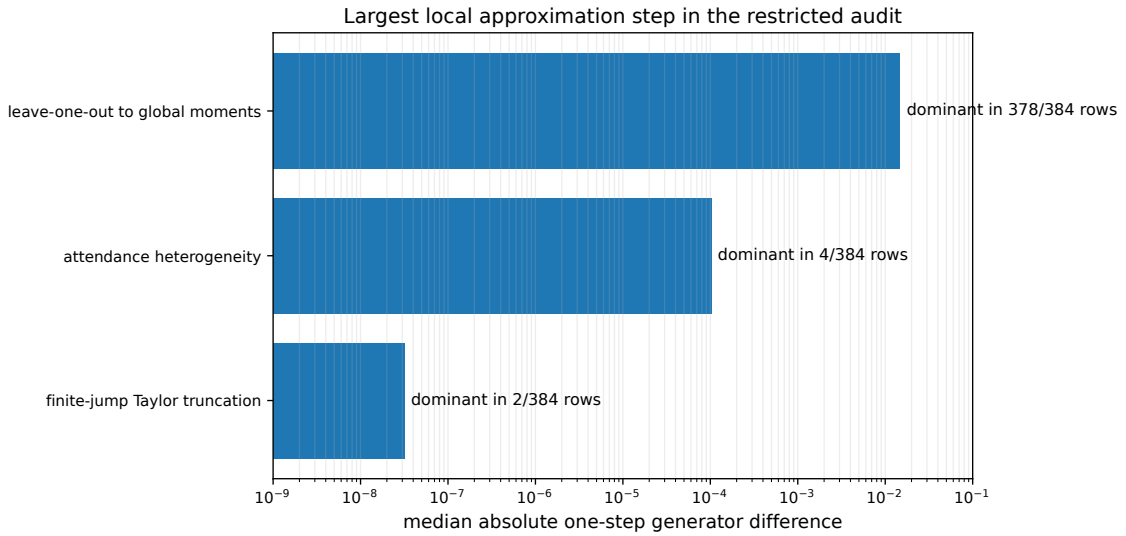


Figure 4. *Exact local weak-generator hierarchy. The audit contains* $384 = 16 \times 4 \times 6$ *case–snapshot–observable rows from* 16 *selected cases, four snapshots, and six smooth observables across both branches and* $M \in \{30, 100, 300\}$. *Bars show the median absolute differences for Steps 1–3 of* (3.22)*; annotations give the number of rows in which each is largest. The focal-to-global replacement* $G_{\mathrm{LOO}} - G_{\mathrm{nonlocal}}$ *dominates* 378/384 *rows, compared with four for attendance heterogeneity and two for finite-jump truncation.*

5.5. Exact one-step weak-generator audit. The dependence score χ^{+}_{ex} and J_{90} identify two case-level signatures, but neither measures the local error produced by its associated approximation. The weak-generator audit supplies that missing comparison. It freezes one microscopic state and one smooth observable, then changes one ingredient of the expected next-round evolution at a time. Closure-related and Kramers–Moyal effects are therefore compared before subsequent trajectory dynamics mix them. The operator hierarchy (3.21) and its exact decomposition (3.22) are defined in Section 3.5; further construction details are given in Supplementary Section SM3.2. At a stored state $\mathbf{X}_n = (X_n^1, \dots, X_n^M)$, the conditional weak generator for focal agent i and smooth test function ψ is

$$G\psi(\mathbf{X}_n, i) = \frac{1}{\tau}\mathbb{E}\left[\psi(X_{n+1}^i) - \psi(X_n^i) \mid \mathbf{X}_n\right], \qquad \overline{G}\psi(\mathbf{X}_n) = \frac{1}{M}\sum_{i=1}^{M} G\psi(\mathbf{X}_n, i). \tag{5.5}$$

This is the expected one-round change of ψ, averaged uniformly over the focal agents. Conditional on $\mathbf{X}_n$, only the next Bernoulli actions are random. For agent i, write

$$K_{-i} = \sum_{j \neq i} \delta_n^j, \qquad \overline{p}_{-i} = \frac{1}{M-1}\sum_{j \neq i} p(X_n^j).$$

The exact law of K_{-i} is Poisson–binomial. Dynamic programming evaluates its complete probability mass function, after which (5.5) is averaged over i. The primary generator values therefore contain no Monte-Carlo sampling noise.

Step 1: attendance heterogeneity. G_{micro} retains the unequal probabilities $p(X_n^j)$ of the other agents. G_{LOO} replaces them by Binomial$(M-1, \overline{p}_{-i})$, which preserves their mean attendance. Hence $G_{\mathrm{micro}} - G_{\mathrm{LOO}}$ measures the effect of unequal attendance probabilities beyond their leave-one-out mean.

Step 2: focal environment versus global mean field. The global mean at the frozen microscopic state is

$$\alpha_n^{\mathrm{micro}} = \frac{1}{M}\sum_{j=1}^{M} p(X_n^j). \tag{5.6}$$

G_{nonlocal} replaces every focal-specific $\overline{p}_{-i}$ by $\alpha_n^{\mathrm{micro}}$, while retaining the complete finite increment $\psi(x+\Delta x)-\psi(x)$. Consequently, $G_{\mathrm{LOO}}-G_{\mathrm{nonlocal}}$ isolates the loss of focal environmental information before Taylor expansion. This is the finite-player analogue of replacing conditional multi-particle information by the one-particle mean under molecular-chaos closure [4].

Step 3: exact finite shift versus second order. Using the same global moments and attendance model, G_{KM2} replaces $\psi(x+\Delta x)-\psi(x) \quad \longrightarrow \quad \psi'(x)\Delta x + \frac{1}{2}\psi''(x)(\Delta x)^2$. Therefore $G_{\mathrm{nonlocal}} - G_{\mathrm{KM2}}$ directly measures the omitted third- and higher-order contribution for the selected state and observable [4]. This is the local truncation error that the total jump-size statistic J_{90} in Section 5.3 cannot determine.

Step 4: coded PDE weak form. G_{PDE} evaluates the same second-order expression as G_{KM2} in the coded PDE weak form. Their difference is an implementation check, not another modelling approximation.

These are the adjacent terms in (3.22). Every pair uses the same state, observable, and uniform focal-agent average, so only the declared ingredient changes. The signed terms telescope exactly, but they are local one-step differences, not additive or causal shares of $D_{W_1,p}^{\mathrm{closed}}$. The restricted audit contains 16 selected cases spanning both branches and $M \in \{30, 100, 300\}$. Four snapshots and six smooth observables per case give $16\times4\times6 = 384$ case–snapshot–observable rows (Figure 4 and Supplementary Table SM8). The observables include localised bump functions, which detect regional mass movement that aggregate moments can miss. For each row, the dominant modelling component is the largest absolute value among Steps 1–3; Step 4 is excluded because it is a code check. Within this audit, the hierarchy is clear. The focal-to-global replacement $G_{\mathrm{LOO}} - G_{\mathrm{nonlocal}}$ (Step 2) is dominant in $378/384 = 98.44\%$ of the rows, compared with $4/384 = 1.04\%$ for attendance heterogeneity (Step 1) and $2/384 = 0.52\%$ for finite-jump truncation (Step 3). Across the 60 observable-by-observable grouped summaries derived from these 384 generator rows, the corresponding median absolute generator differences are 1.494×10^{-2}, 1.039×10^{-4}, and 3.209×10^{-8} (Figure 4 and Supplementary Table SM8). The principal local change therefore occurs when the focal environment is replaced by one global mean. Because Step 2 preserves the global value $\alpha_n^{\mathrm{micro}}$ but removes which agent experiences which leave-one-out environment, it can retain the population-wide restoring signal while altering how updates are distributed across agents. This makes the aggregate–law split intuitive: agreement of the scalar observables α and β can coexist with a nonzero whole-law W_1 discrepancy, as recorded in (4.3). The generator result is a local mechanism consistent with that pattern. By contrast, Step 3 is extremely small for the selected states, smooth observables, and populations $M \le 300$. This is positive evidence that the

second-order Kramers–Moyal replacement is locally accurate in the audited regime, without extrapolating the ordering to larger populations or other observables. Nor does it require J_{90} to be small: J_{90} measures the complete realised shift, whereas Step 3 measures only the remainder after its first- and second-order contributions have been retained. The exact G_{micro} calculation was independently checked on 64 rows using 10^5–10^6 simulated next rounds. With $Z_{\mathrm{MC}} = \frac{\widehat{G}_{\mathrm{MC}} - G_{\mathrm{exact}}}{\mathrm{SE}(\widehat{G}_{\mathrm{MC}})}$, the largest observed $|Z_{\mathrm{MC}}|$ was approximately 2.38, consistent with ordinary simulation variability (Supplementary Figure SM8 and Table SM8). Because the exact and Monte-Carlo calculations estimate the same conditional expectation by different algorithms, this is an independent validation of the exact generator computation. Taken together, the local audit identifies a clear hierarchy over the selected states, observables, and populations $M \leq 300$. Replacing the focal leave-one-out environment by one global mean produces the largest tested local change, whereas replacing the exact finite shift by its second-order Kramers–Moyal expansion is much smaller.

6. Return to capacity from off-capacity initial states. Sections 4 and 5 begin from initial laws whose population-average entry probability is calibrated to the capacity target. They therefore test *preservation* and particle–PDE agreement near capacity, not recovery from an incorrect aggregate state. Here we ask: if the population initially enters too often or too rarely, do the microscopic and PDE dynamics move its average entry probability back towards capacity? With $\alpha(\theta)$ defined in (2.9) and $\kappa(M)$ in (2.3), we use the signed capacity error

$$a(\theta) = \kappa(M) - \alpha(\theta). \tag{6.1}$$

Thus $a < 0$ denotes over-entry, $a > 0$ under-entry, and $a = 0$ the symmetric capacity target. Aggregate learning in [4] concerns approach to the finite-game equilibrium interval $\mathcal{I}_M$ in (2.4). Our numerical diagnostic measures displacement from the particular point $\kappa(M) \in \mathcal{I}_M$, consistently with the calibration in (2.6) and the capacity-matched set (3.4). Hence $|a(\theta)|$ measures the distance from the chosen target, while its sign records the direction of the imbalance. This section examines the faster aggregate stage over the pre-specified intermediate horizon $\theta_{\max} = 20$. The horizon tests recovery from a finite displacement rather than the limit $\theta \to \infty$. The slower sorting problem is treated separately in Section 7, using a broad screen at $\theta = 200$ and targeted trajectories on the $O(10^5)$–$O(10^6)$ scales required to resolve direct bounded-window sorting. This separation reflects the analytical prediction that aggregate learning precedes the slower redistribution of the full propensity density [4].

6.1. A controlled deterministic return-to-capacity test.

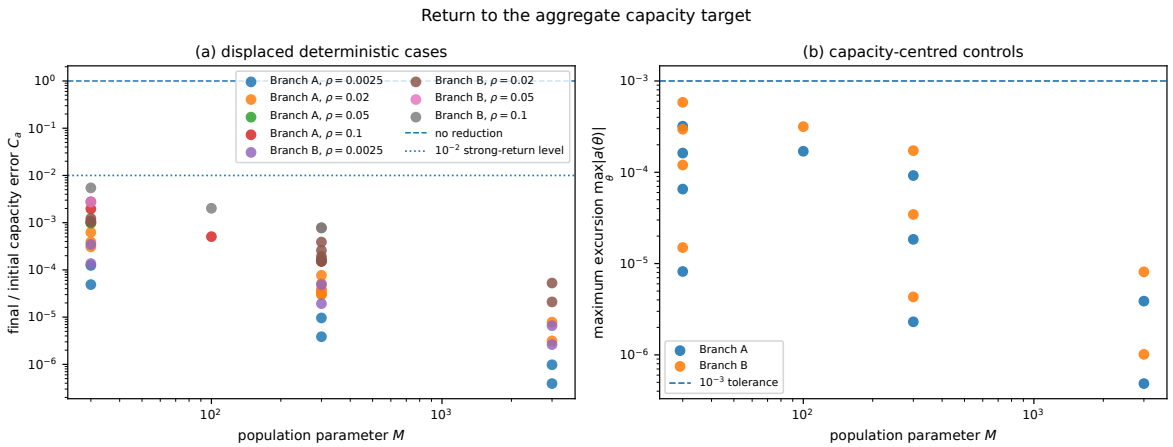


Figure 5. ***(a)*** *Final-to-initial capacity-error ratio C_a for all 46 displaced PDE cases. The lines at 1 and 10^{-2} mark no contraction and the strong-return benchmark, respectively.* ***(b)*** *Maximum excursion $E_{\max}$ for the 20 capacity-matched controls. The dashed line is the fixed 10^{-3} preservation tolerance.*

The capacity-matched three-bump law defined in (3.2) and calibrated by (3.3) has $a(0) \simeq 0$. We use it to test two complementary properties of the deterministic PDE: whether an initially displaced aggregate mean returns towards capacity, and whether an already matched mean remains there. For a displaced case, $a(0) \neq 0$, the remaining fraction of the initial capacity error is

$$C_a = \frac{|a(\theta_{\max})|}{|a(0)|}. \tag{6.2}$$

Thus $C_a = 1$ means no net contraction, whereas $C_a = 10^{-2}$ means that only 1% of the initial error remains. We use 10^{-2} as the strong-return benchmark. When $a(0) \simeq 0$, this ratio is uninformative, so preservation is measured instead by the maximum excursion

$$E_{\max} = \max_{0 \le \theta \le \theta_{\max}} |a(\theta)|. \tag{6.3}$$

The imposed displacement is $\delta_\alpha = \alpha(0) - \kappa(M) = -a(0)$. Positive and negative values represent initial over-entry and under-entry, respectively. The pre-specified deterministic family contains both signs and several magnitudes, with principal nonzero magnitudes $|\delta_\alpha| \in \{0.02, 0.05\}$. Because C_a uses absolute values, Figure 5(a) compares the amount of contraction without encoding the initial sign. To impose a prescribed δ_α without changing the probability map, we exponentially reweight a chosen base density b:

$$f_\eta(x) = \frac{b(x)e^{\eta p(x)}}{Z(\eta)}, \qquad Z(\eta) = \int_{\mathbb{R}} b(y)e^{\eta p(y)}\,\mathrm{d}y, \qquad \int_{\mathbb{R}} p(x)f_\eta(x)\,\mathrm{d}x = \kappa(M) + \delta_\alpha. \tag{6.4}$$

A one-dimensional root solve selects η. The construction preserves positivity and unit mass, and the target moment is monotone because $\frac{\mathrm{d}}{\mathrm{d}\eta}\int p(x)f_\eta(x)\,\mathrm{d}x = \mathrm{Var}_{f_\eta}(p(X)) \ge 0$. Moment calibration fixes the initial aggregate displacement but not the arrangement of mass in

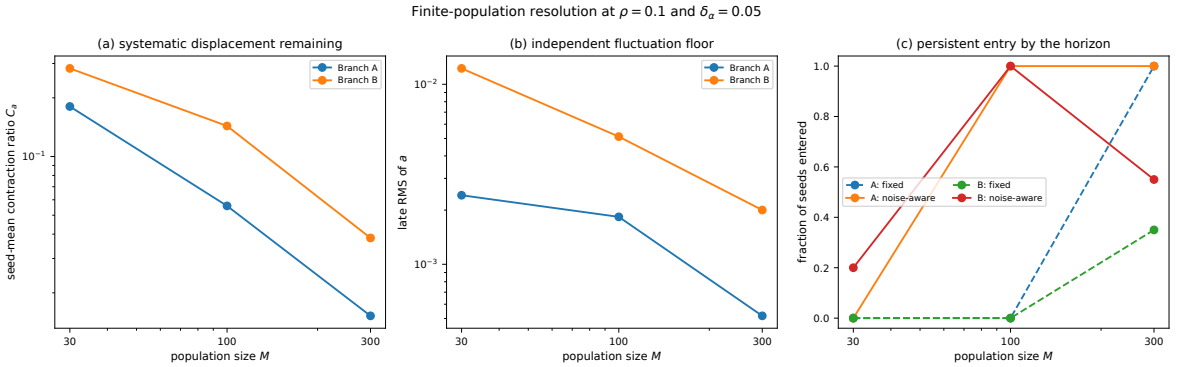


Figure 6. ***(a)*** *Final-to-initial capacity-error ratio C_a of the microscopic seed-mean path.* ***(b)*** *Late RMS capacity error in independent zero-offset controls.* ***(c)*** *Fraction of displaced realisations that persistently enter the fixed 10^{-3} tube (dashed) or the case-specific noise-aware tube (solid) by the case horizon. Exact values are reported in Supplementary Table SM10 and Supplementary Table SM11.*

propensity space. We therefore apply (6.4) to three base shapes. With $\mathcal{T}_s f(x) = f(x - s)$,

$$\begin{aligned} b_{\mathrm{tilt}}(x) &= f_0^{\mathrm{match}}(x), \\ b_{\mathrm{trans}}(x) &= \mathcal{T}_4 f_0^{\mathrm{match}}(x), \\ b_{\mathrm{bi}}(x) &= \tfrac{1}{2}\mathcal{T}_{-7} f_0^{\mathrm{match}}(x) + \tfrac{1}{2}\mathcal{T}_{+7} f_0^{\mathrm{match}}(x). \end{aligned} \tag{6.5}$$

These are the baseline law, a translated copy, and a separated two-component mixture. Within each comparison, all three are reweighted to the same chosen value $\alpha(0) = \kappa(M) + \delta_\alpha$. The shape therefore changes while the magnitude and direction of the aggregate displacement are held fixed. Figure 5 reports the complete deterministic screen. Panel (a) contains all 46 displaced PDE cases, spanning both probability maps, both displacement signs, the tested magnitudes, M and ρ values, and the three initial-law shapes. Every case satisfies $C_a < 10^{-2}$, so more than 99% of the initial capacity error is removed by the endpoint. Panel (b) contains the 20 zero-offset controls. All satisfy $E_{\max} < 10^{-3}$, and the largest observed excursion is 5.84×10^{-4}. The PDE therefore both restores a displaced aggregate mean and preserves one that is already capacity matched. Across both branches, $\rho = 0.02$, $M \in \{30, 300\}$, and all three shapes, the twelve deterministic shape cases satisfy $C_a \le 1.233 \times 10^{-3} < 10^{-2}$. Exact values are given in Supplementary Table SM9. Because C_a is only an endpoint statistic, Supplementary Figure SM9 shows the normalized trajectories over the full horizon and an enlarged early-time interval. Their curvature and return speed vary, so the evidence supports robust endpoint contraction across the tested shapes, not a universal shape-independent decay law. This finite-horizon result should be distinguished from the sufficient asymptotic statement in Theorem 5.1(ii) of [4]. At the probability-map level, Branch A satisfies the regularity and positive-floor requirements, although application of the theorem also requires the transport-dominance condition (2.14). Branch B has $\inf_{x \in \mathbb{R}} p_B(x) = 0$ and therefore lies outside the theorem's positive-floor assumption. Strong return in both branches is thus direct numerical evidence for the specified family, not a proof of global attraction for arbitrary initial laws.

6.2. Finite-player return under stochastic resolution. The preceding deterministic test showed strong return towards capacity. We now ask whether the stochastic finite-player game has the same restoring tendency and, because individual populations continue to fluctuate, when a finite population can reasonably be said to have returned. We use

$$\rho = 0.1, \qquad M \in \{30, 100, 300\}, \qquad \delta_\alpha = 0.05,$$

with 20 independent realisations for each branch and population size. For each realisation we track the empirical mean entry probability defined in (2.5); averaging the 20 paths gives the *seed-mean path*. Each ensemble is compared with its case-matched PDE as specified in Section 3.6. The design and law-level comparisons are summarized in Figure 6 and Supplementary Table SM10. Additional initial-shape checks and representative aggregate trajectories are given in Supplementary Table SM9 and Supplementary Figure SM10, respectively. Figure 6 separates three questions. Panel (a) applies the contraction ratio C_a from (6.2) to the seed-mean path and therefore measures the systematic displacement remaining on average. Panel (b) measures the fluctuation present even without an imposed displacement. For an independent zero-offset control, let $\mathcal{W}_{\mathrm{late}}$ be the stored times in the final half of the trajectory and define

$$\mathrm{RMS}_{\mathrm{late}}(a) = \left[\frac{1}{|\mathcal{W}_{\mathrm{late}}|} \sum_{\theta \in \mathcal{W}_{\mathrm{late}}} a(\theta)^2 \right]^{1/2}. \tag{6.6}$$

This is the ordinary late fluctuation scale of a population that began at capacity. Panel (c) uses the persistent-entry rule in (3.17): a trajectory must enter a tube around $a = 0$ and remain there at every later stored time. The fixed tube $|a| \le 10^{-3}$ asks whether every case reaches the same strict absolute accuracy. The noise-aware tube instead uses

$$\varepsilon_{\mathrm{noise}} = Q_{0.95}\left(|a(\theta)| : \theta \in \mathcal{W}_{\mathrm{late}} \text{ of the matched zero-offset control}\right), \tag{6.7}$$

calibrated separately for each branch, M, and ρ. It asks whether the residual displacement is no larger than the fluctuations ordinarily observed in an otherwise identical population that started at capacity. The seed mean returns towards capacity in both branches, with less systematic error remaining as M increases. From $M = 30$ to $M = 300$, C_a falls from 0.1806 to 0.0152 in Branch A and from 0.2832 to 0.0382 in Branch B (Figure 6(a) and Supplementary Table SM10). Thus only about 1.5% and 3.8%, respectively, of the imposed seed-mean displacement remains at $M = 300$. Branch B retains the larger aggregate residual, but both branches show the same strengthening of return over the tested population ladder. The same supplementary table also shows that aggregate return does not imply complete law-level agreement: bounded-window and probability-space discrepancies remain visible after the mean has moved close to capacity. The background fluctuation floor narrows with population size but is larger for Branch B. Between $M = 30$ and $M = 300$, the late RMS decreases from 2.419×10^{-3} to 5.17×10^{-4} in Branch A and from 1.2256×10^{-2} to 2.002×10^{-3} in Branch B (Figure 6(b) and Supplementary Table SM11). This explains why a strict absolute tube and a noise-aware tube can give different answers. At $M = 100$, neither branch enters the fixed 10^{-3} tube

persistently, whereas all 20 realisations in each branch enter their noise-aware tubes. The imposed displacement has therefore fallen to the scale of ordinary finite-population noise, even though that noise is still wider than 10^{-3} (Figure 6(c) and the same supplementary table). Noise-aware entry need not vary monotonically with M, because its tube tightens as the control fluctuations decrease. For Branch B, for example, the threshold narrows from 0.006316 at $M = 100$ to 0.002270 at $M = 300$; the corresponding entry fraction changes from 100% to 55%, while the fixed-tube fraction at $M = 300$ is 35%. Branch A at $M = 300$ reaches 100% under both criteria (Figure 6(c) and Supplementary Table SM11). The reported case horizons are not all identical: most equal 20, whereas the Branch A, $M = 300$ case has horizon 2.9. The entry fractions are therefore finite-horizon resolution diagnostics, not a scaling law in M (same supplementary table). If persistent entry is not observed by a case's final stored time, the event is right-censored: $T_\varepsilon > \theta_{\max}$, rather than assigned to the endpoint or treated as impossible. The full entry-time curves are shown in Supplementary Figure SM11. Overall, the finite-player game reproduces the aggregate restoring direction of the PDE, while finite-M fluctuations set the resolution at which return can be recognized for an individual population. Section 7 turns to the slower distributional question: after the aggregate mean is near capacity, does the propensity density continue to reorganise towards sorting?

7. Long-time dynamics: fast aggregate learning, slow sorting. Section 6 established the fast aggregate stage: over $\theta \leq 20$, all 46 displaced deterministic PDE cases ended with $C_a < 10^{-2}$, leaving less than 1% of their initial capacity error (Figure 5). In the notation of (2.9), small $|a(\theta)|$ therefore means that the population-average entry probability is close to $\kappa(M)$. Sorting is stronger: as reviewed in Section 2.4, S_R from (2.13) is the mass remaining in a fixed propensity window $[-R, R]$, and sorting requires

$$S_R(\theta) \longrightarrow 0 \qquad \text{as } \theta \to \infty \quad \text{for every fixed } R > 0. \tag{7.1}$$

Aggregate learning concerns the mean entry rate, whereas decreasing S_R records the slower outward redistribution of the full density towards increasingly extreme propensities. This is the one-particle sorting criterion proved in Theorem 5.1(i) of [4]. Section 5 of [4] explains the expected separation of scales. In the original parameters, the transport coefficient $h(M-1)/\tau$ governs aggregate adjustment; once a is small, the remaining b-driven diffusion scales as $h^2(M-1)/(2\tau)$ and is associated with slower sorting. The long-time computations test whether this ordering is visible at finite M and finite time. Because the analytical problem is posed on $\mathbb{R}$ whereas the solver uses a finite domain, boundary interaction is monitored explicitly below.

7.1. Boundary-safe sorting after aggregate learning. The long-time calculations start from the capacity-matched baseline laws satisfying (2.6). Unlike the displaced tests in Section 6, they ask whether the propensity density continues to reorganise after the aggregate entry probability has settled near capacity. For the bounded-window mass $S_R(\theta)$ defined in (2.13), we monitor $R \in \{10, 50, 100\}$. Because these intervals are nested, $S_{10}(\theta) \leq S_{50}(\theta) \leq S_{100}(\theta)$. Reaching the same low threshold for a larger R therefore requires density to propagate farther from the centre. In particular, $S_{10} \leq 0.10$ means that at least 90% of the density lies outside $[-10, 10]$, while $S_{10} \leq 0.05$ raises this fraction to at least 95%. These are finite-time milestones for the sorting condition (7.1). The whole-line PDE is approximated by the finite-volume solver of Section 3.2 on $x \in [-L, L]$, $L = 2000$, with zero numerical flux at

$x = \pm L$. To identify central-window evacuation before the artificial walls materially influence the solution, we monitor the mass in the outer 10% of the box,

$$B_L(\theta) = \int_{\{x:\, 0.9L \le |x| \le L\}} f(x,\theta)\,\mathrm{d}x, \tag{7.2}$$

and retain a sorting event only while

$$B_L(\theta) < 5 \times 10^{-3}. \tag{7.3}$$

Thus less than 0.5% of the numerical density lies near the boundaries when an event is recorded, so the observed loss of central mass precedes appreciable interaction with the truncated domain. We count a threshold crossing only when it is sustained. For a window R and level q, $\Theta_{R,q}$ is the linearly interpolated first crossing for which

$$S_R(\theta_j) \le q \quad \text{at every later stored time } \theta_j \text{ for which (7.3) holds.} \tag{7.4}$$

The 56-case PDE screen already shows the separation. With $\alpha(\theta)$ and the capacity error $a(\theta)$ defined in (2.9), at $\theta = 200$ all 56 values of $\alpha(200)$ lie in the finite-game equilibrium interval $\mathcal{I}_M$ from (2.4), and the median $|a(200)|$ is only 5.628×10^{-7}. The aggregate entry probability is therefore effectively at capacity. At the same time, $S_{100}(200)$ remains equal to 1 to stored numerical precision, far from the 0.10 milestone: essentially all of the density is still inside $[-100, 100]$. The complete values are reported in the supplement. The direct boundary-safe test follows the eight Branch A trajectories

$$\rho \in \{0.0025, 0.02\}, \qquad M \in \{30, 300, 3000, 30000\}. \tag{7.5}$$

Whereas the aggregate moment is already settled by $\theta = 200$, the first direct sorting milestones occur only when the scaled time reaches the order of 10^5–10^6. Every trajectory reaches $S_{10} \le 0.10$ before the boundary warning, and the six trajectories with $M \ge 300$ reach $S_{10} \le 0.05$:

$$2.551 \times 10^5 \le \Theta_{10,0.10} \le 2.704 \times 10^5, \tag{7.6}$$

$$1.068 \times 10^6 \le \Theta_{10,0.05} \le 1.079 \times 10^6. \tag{7.7}$$

Figure 7(a) shows the evacuation of the central window, while Supplementary Table SM13 gives the case counts and ranges. The first range is the time required to place at least 90% of the density outside $[-10, 10]$; the second is the later, stricter event at which this fraction reaches at least 95%. The quoted ranges are the baseline results. Their dependence on spatial resolution is examined separately in Section 7.2 and Supplementary Table SM15. No boundary-safe $S_{50} \le 0.10$ or $S_{100} \le 0.10$ crossing is resolved within the available domain and horizon. This does not indicate an absence of outward motion: it reflects the longer propagation required to evacuate wider intervals and shows how slowly the later stages of sorting unfold. The nested-window trajectories in Supplementary Figure SM12 display this ordering without assigning post-warning crossing times. The contrast with Section 6 is therefore sharp. The deterministic capacity error is strongly reduced by $\theta = 20$ (Figure 5), whereas the directly resolved S_{10} milestones occur only at scaled times of order 10^5–10^6 (Figure 7 and Supplementary Table SM13). Aggregate learning precedes bounded-window sorting by several orders of magnitude; monitoring only the population mean would hide most of the subsequent redistribution of individual propensities.

7.2. A common late sorting scale across population sizes. The computational clock is the scaled time $\theta = (M-1)t$ defined in (3.8). Let $t_{10,q}$ be the physical time corresponding to the persistent $S_{10} = q$ crossing defined in (7.4). Its scaled crossing time is

$$\Theta_{10,q} = (M-1)t_{10,q}. \tag{7.8}$$

If $\Theta_{10,q}$ is approximately independent of population size, then

$$t_{10,q} = \frac{\Theta_{10,q}}{M-1} \propto (M-1)^{-1}, \tag{7.9}$$

so a log–log regression of physical crossing time against $M-1$ should have slope -1. Figure 7(b) shows this scaled-time collapse. At fixed ρ, the fitted physical-time slopes are

$$\frac{\mathrm{d}\log t_{10,0.10}}{\mathrm{d}\log(M-1)} \simeq -0.9923, \tag{7.10}$$

$$\frac{\mathrm{d}\log t_{10,0.05}}{\mathrm{d}\log(M-1)} \simeq -0.9977, \tag{7.11}$$

with $R^2 \geq 0.999996$. The slopes change negligibly when ρ increases from 0.0025 to 0.02; the complete regressions are reported in Supplementary Table SM14. Thus the tested Branch A sorting milestones occur at nearly common values of the scaled time across M, and are also stable under the tested eight-fold change in ρ. This statement concerns the late PDE clock, not an absence of population-size effects. Physical crossing times still decrease with M according to (7.9), while the finite-population and closure effects studied in Sections 4 and 5 remain M-dependent. The analytical paper motivates the characteristic scaling in its closing discussion of Section 5 [4]; the near-collapse of these finite S_{10} thresholds is an additional numerical result. The numerical value of the crossing time is more sensitive to spatial resolution. To measure this effect, the $\rho = 0.0025$ cases with $M = 3000$ and $M = 30000$ were recomputed on the same domain with $\Delta x \in \{4, 2, 1\}$. Averaging the two population sizes, refinement from $\Delta x = 2$ to $\Delta x = 1$ lowers the mean $S_{10} = 0.10$ crossing time from 2.6388×10^5 to 2.3221×10^5, and the mean $S_{10} = 0.05$ crossing time from 1.0709×10^6 to 9.4199×10^5. Relative to the fine-grid values, these are shifts of 13.6% and 13.7%, respectively (Figure 7(c) and Supplementary Table SM15). The population-size comparison is much more stable. Holding Δx fixed, the relative separation $\frac{\left|\Theta_{10,q}^{(M=3000)} - \Theta_{10,q}^{(M=30000)}\right|}{\frac{1}{2}\left(\Theta_{10,q}^{(M=3000)} + \Theta_{10,q}^{(M=30000)}\right)}$ lies between 0.0293% and 0.0301% across both thresholds and all three grids (Supplementary Table SM15). Refining the grid therefore moves the two crossing times together: it changes the absolute clock appreciably, but leaves their near-collapse across population sizes essentially unchanged.

7.3. Late sorting is fluctuation driven.

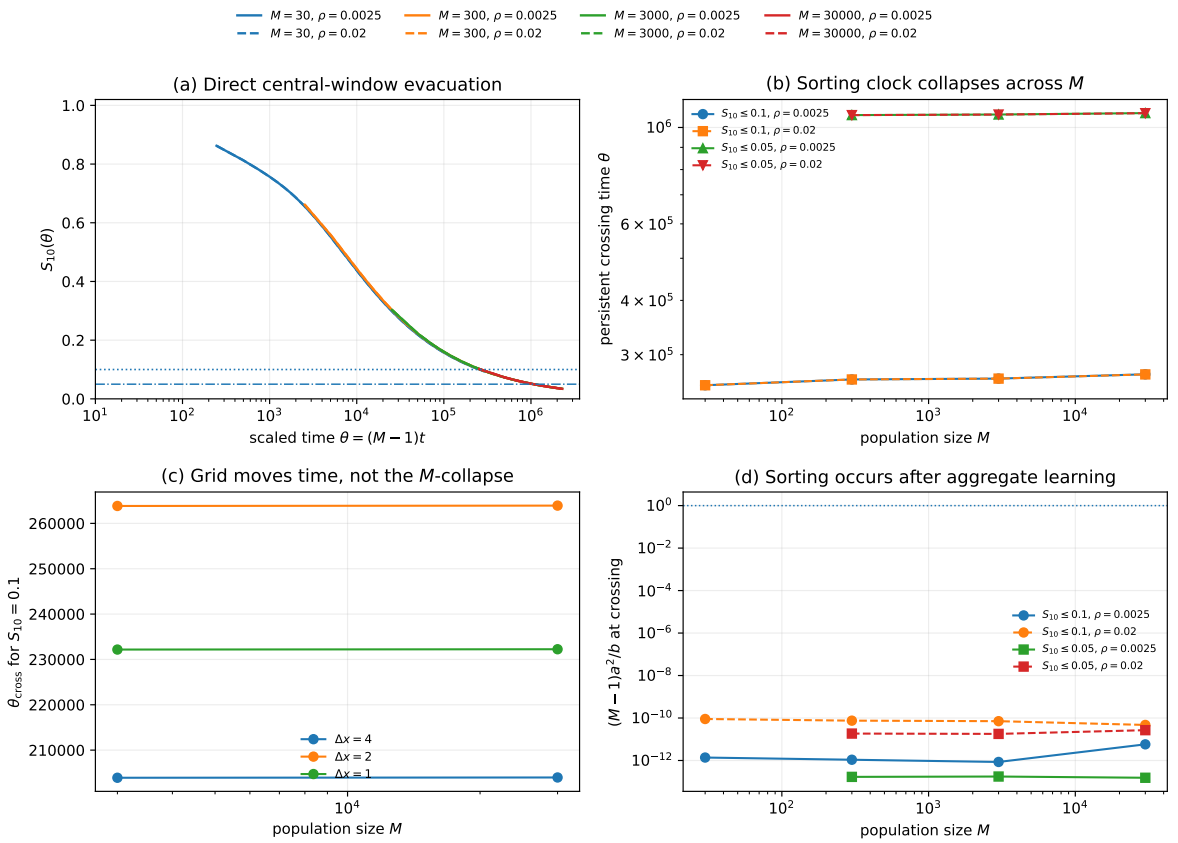


Figure 7. ***(a)*** *Central-window mass S_{10}; the levels* 0.10 *and* 0.05 *correspond to* 90% *and* 95% *of the density outside* $[-10, 10]$. ***(b)*** *Persistent crossing times on the scaled clock, showing weak residual dependence on M.* ***(c)*** *Spatial-resolution audit: changing Δx shifts the absolute crossing time but leaves the two population sizes nearly coincident.* ***(d)*** *The mechanism ratio* (7.13) *at the direct crossings. Values far below one show that the surviving b-driven diffusion, rather than the residual capacity error, governs the observed redistribution. All reported crossings precede the prescribed boundary warning.*

Rewriting the closed one-particle PDE (2.12) on the scaled clock $\theta = (M-1)t$ from (3.8) gives

$$\partial_\theta f + \frac{a}{\sqrt{\tau}}\,\partial_x\big(pf\big) - \frac{1}{2}\big[(M-1)a^2 + b\big]\partial_{xx}\big(pf\big) = 0, \qquad b = \alpha(1-\alpha). \tag{7.12}$$

This is the same PDE used in the computations, expressed on the common time scale. Equation (7.12) separates the two stages of the dynamics. When the capacity error a is appreciable, the transport term $a/\sqrt{\tau}$ drives the aggregate correction studied in Section 6; the diffusion also contains the a-dependent contribution $(M-1)a^2$. Both weaken as $a \to 0$. The second diffusion contribution survives because, near capacity, $b \simeq \kappa(M)(1-\kappa(M)) > 0$. Thus correcting the mean entry probability does not freeze the density: fluctuations in individual entry decisions continue to redistribute propensities. This transition from aggregate restoration to slower diffusion is the mechanism predicted in the analytical long-time analysis [4]. The direct

S_{10} crossings let us test which diffusion contribution is active when sorting is observed. Let

$$\mathcal{R}_{\text{diff}} = \frac{(M-1)a^2}{b}. \tag{7.13}$$

This is a mechanism diagnostic, not a second sorting criterion: $\mathcal{R}_{\text{diff}} \ll 1$ means that diffusion caused by the residual capacity error is negligible relative to the surviving b-term. At the $S_{10} = 0.10$ crossings, its median and maximum are 2.65×10^{-11} and 8.99×10^{-11}; at the $S_{10} = 0.05$ crossings, they are 9.02×10^{-12} and 2.67×10^{-11} (Figure 7(d) and Supplementary Section SM6.1). Even the largest value is below 9×10^{-11}, so the a^2 contribution is more than ten orders of magnitude smaller than the b-contribution. The maximum transport coefficient also falls from 4.40×10^{-4} at the 0.10 milestone to 1.19×10^{-4} at the 0.05 milestone, confirming that the a-driven motion continues to weaken as sorting progresses (Supplementary Section SM6.1). The leading late-time dynamics are therefore

$$\partial_\theta f \simeq \frac{b}{2}\,\partial_{xx}(pf). \tag{7.14}$$

Although b, $\kappa(M)$, the probability map, and the evolving density can still depend on the model parameters, equation (7.14) contains no explicit factor of M or ρ. This explains why the Branch A sorting times can nearly collapse on the scaled clock. The additional numerical finding is that this fluctuation-dominated regime is already established at the directly observed finite-time sorting events. As a proof-facing consistency check, we also monitor the diagnostic introduced in Section 2.4,

$$g_2(\theta) = \int_{\mathbb{R}} \big(p(x)f(x,\theta)\big)^2\,\mathrm{d}x, \qquad \phi(\theta) = \beta(\theta)g_2(\theta)^3. \tag{7.15}$$

Under the hypotheses of Lemma 5.4 and Corollary 5.5 in [4], this quantity controls the mass in any fixed bounded window through

$$S_R(\theta) \le C_R\,\phi(\theta)^{1/7}. \tag{7.16}$$

In the Branch A trajectories, ϕ continues to decrease while S_{10} passes through both direct milestones (Supplementary Figure SM12). This is consistent with the analytical route to sorting, but the reported crossing times are measured directly from S_{10}; the bound (7.16) is not used as a finite-time clock. The direct quantitative result is restricted to Branch A. For Branch B, $p_B(x) > 0$ at every finite x but $\inf_{\mathbb{R}} p_B = 0$, as discussed in Section 2.2. Because the PDE flux acts through $p_B f$, very small values of p_B in the far-left tail make outward propagation harder to resolve, and no comparable boundary-safe S_{10} milestone is observed on the available horizon (Supplementary Section SM6.1). The Branch B crossing time is therefore unresolved, not evidence of failed sorting. This remains compatible with Theorem 5.1: the uniform positive lower bound on p is an additional assumption for its aggregate-learning conclusion, not for the sorting statement [4].

8. Conclusion, limitations, and future directions. This study assessed when the one-particle kinetic PDE provides a useful finite-population surrogate for stochastic reinforcement

learning in a market-entry game. The answer is observable dependent. Across the tested range $30 \leq M \leq 30000$, the PDE reproduced the aggregate entry rate and within-agent Bernoulli uncertainty more accurately than the complete probability law. Branch A satisfied all declared finite-time behavioural and law-level criteria, whereas Branch B retained a small but detectable distributional discrepancy despite accurate aggregate moments. Thus the PDE is especially well suited to questions about aggregate learning, while law-level predictions require greater caution for probability maps with a vanishing lower tail.

The mechanism diagnostics clarify this distinction. Corrected finite-player dependence decreased approximately as M^{-1}, while realised probability jumps scaled almost linearly with ρ and showed negligible residual dependence on M. Supplying microscopic aggregate coefficient paths did not systematically improve the PDE law, and the exact one-step generator audit identified the replacement of focal leave-one-out information by a single global mean as the leading local change in the audited regime; the second-order Kramers–Moyal remainder was much smaller. The off-capacity experiments further showed a robust restoring mechanism: all tested deterministic cases removed more than 99% of their initial capacity error, and the finite-player systems displayed the same restoring direction once their intrinsic noise was taken into account. Finally, Branch A exhibited a separation of at least four orders of magnitude between aggregate adjustment and directly resolved sorting. At the sorting milestones, redistribution was governed by the surviving fluctuation-driven diffusion rather than by residual capacity error.

Limitations. These findings are numerical and apply to the specified probability maps, parameter ranges, initial-law families, observables, and finite horizons; they do not constitute a propagation-of-chaos theorem or a uniform error bound. The corrected dependence statistic is bin based, and the generator hierarchy is a local audit restricted to selected states and smooth observables. Moreover, the absolute long-time crossing times retain visible spatial-resolution sensitivity. A boundary-safe sorting milestone was obtained only for Branch A, so the corresponding Branch B time remains unresolved rather than disproved.

Further directions. The most direct next step is to develop a finite-(M) correction that retains focal or two-particle information and to test whether it improves law-level accuracy without sacrificing the efficiency of the one-dimensional PDE. Analytical work could seek quantitative propagation-of-chaos or finite-population error estimates adapted to the shared attendance and extend the long-time theory to probability maps with vanishing lower tails. A parameterised family of response maps that varies the probability floor, slope, and curvature separately would help identify why Branch B is harder to reproduce. Adaptive-domain and higher-resolution methods could resolve its long-time behaviour and sharpen the absolute Branch A sorting times, while fixed-horizon experiments could isolate jump-size effects from their accumulation over time. The focal-to-global discrepancy also provides a principled target for reduced-order or data-assisted corrections that retain the closed PDE as their computational backbone. Taken together, the results position the kinetic PDE as an efficient, observable-specific reduced model: its aggregate predictions are already strong, and its principal finite-player limitation has been sufficiently localised to guide the next analytical and computational refinements.

Acknowledgements. The author is funded by the European Union's Horizon 2020 research

and innovation programme under the Marie Skłodowska-Curie grant agreement n°101034255.

Code and data availability. The complete computational reproduction bundle is publicly available online [3].

Conflicts of interest. The author declares that there are no conflicts of interest for this work.

REFERENCES

[1] W. B. Arthur, *Inductive reasoning and bounded rationality*, American Economic Review, 84 (1994), pp. 406–411, https://www.jstor.org/stable/2117868.

[2] M. Benaïm, *Dynamics of stochastic approximation algorithms*, in Séminaire de Probabilités XXXIII, J. Azéma, M. Émery, M. Ledoux, and M. Yor, eds., vol. 1709 of Lecture Notes in Mathematics, Springer, Berlin, Heidelberg, 1999, pp. 1–68, https://doi.org/10.1007/BFb0096509.

[3] E. Bou Dagher, *Computational validation of a kinetic pde for reinforcement learning in market-entry games*, 2026, https://github.com/EstherBD/Computational-Validation-of-a-Kinetic-Equation. Public-release branch, commit 6bbda6cb247183ad8792ea989e7b15821c6bfb8e.

[4] E. Bou Dagher, M. Perepelitsa, and E. Zatorska, *On a PDE model for learning in stochastic market entry games*, 2026, https://doi.org/10.48550/arXiv.2603.06514, https://arxiv.org/abs/2603.06514, https://arxiv.org/abs/2603.06514.

[5] L.-P. Chaintron and A. Diez, *Propagation of chaos: A review of models, methods and applications. I. models and methods*, Kinetic and Related Models, 15 (2022), pp. 895–1015, https://doi.org/10.3934/krm.2022017.

[6] P. Degond, J.-G. Liu, and C. Ringhofer, *Large-scale dynamics of mean-field games driven by local Nash equilibria*, Journal of Nonlinear Science, 24 (2014), pp. 93–115, https://doi.org/10.1007/s00332-013-9185-2.

[7] J. Duffy and E. Hopkins, *Learning, information, and sorting in market entry games: Theory and evidence*, Games and Economic Behavior, 51 (2005), pp. 31–62, https://doi.org/10.1016/j.geb.2004.04.007.

[8] I. Erev and A. Rapoport, *Coordination, "magic," and reinforcement learning in a market entry game*, Games and Economic Behavior, 23 (1998), pp. 146–175, https://doi.org/10.1006/game.1997.0619.

[9] I. Erev and A. E. Roth, *Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria*, American Economic Review, 88 (1998), pp. 848–881, https://www.jstor.org/stable/117009.

[10] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*, MIT Press, Cambridge, MA, 1998.

[11] C. B. Harley, *Learning the evolutionarily stable strategy*, Journal of Theoretical Biology, 89 (1981), pp. 611–633, https://doi.org/10.1016/0022-5193(81)90032-1.

[12] J. G. MacKinnon and H. White, *Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties*, Journal of Econometrics, 29 (1985), pp. 305–325, https://doi.org/10.1016/0304-4076(85)90158-7.

[13] J. McKean, Henry P., *Propagation of chaos for a class of nonlinear parabolic equations*, in Stochastic Differential Equations, vol. 7 of Lecture Series in Differential Equations, Air Force Office of Scientific Research, Office of Aerospace Research, Arlington, VA, 1967, pp. 41–57. Proceedings of Session 7, Catholic University of America, Washington, DC.

[14] M. Perepelitsa, *Adaptive learning in large populations*, Journal of Mathematical Biology, 79 (2019), pp. 2237–2253, https://doi.org/10.1007/s00285-019-01427-3.

[15] M. Perepelitsa, *A model of cultural evolution in the context of strategic conflict*, Kinetic and Related Models, 14 (2021), pp. 523–539, https://doi.org/10.3934/krm.2021014.

[16] A. Rapoport, D. A. Seale, I. Erev, and J. A. Sundali, *Equilibrium play in large group market entry games*, Management Science, 44 (1998), pp. 119–141, https://doi.org/10.1287/mnsc.44.1.119.

[17] A. E. Roth and I. Erev, *Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term*, Games and Economic Behavior, 8 (1995), pp. 164–212, https://doi.org/10.1016/S0899-8256(05)80020-X.

[18] P. K. Sweby, *High resolution schemes using flux limiters for hyperbolic conservation laws*, SIAM Journal on Numerical Analysis, 21 (1984), pp. 995–1011, https://doi.org/10.1137/0721062.

[19] A.-S. Sznitman, *Topics in propagation of chaos*, in École d'Été de Probabilités de Saint-Flour XIX—1989, P.-L. Hennequin, ed., vol. 1464 of Lecture Notes in Mathematics, Springer, Berlin, Heidelberg, 1991, pp. 165–251, https://doi.org/10.1007/BFb0085169.

[20] B. van Leer, *Towards the ultimate conservative difference scheme. V. a second-order sequel to Godunov's method*, Journal of Computational Physics, 32 (1979), pp. 101–136, https://doi.org/10.1016/0021-9991(79)90145-1.

[21] P. Virtanen et al., *SciPy 1.0: Fundamental algorithms for scientific computing in python*, Nature Methods, 17 (2020), pp. 261–272, https://doi.org/10.1038/s41592-019-0686-2.

[22] J. S. Vitter, *Random sampling with a reservoir*, ACM Transactions on Mathematical Software, 11 (1985), pp. 37–57, https://doi.org/10.1145/3147.3165.

[23] D. Whitehead, *The El Farol bar problem revisited: Reinforcement learning in a potential game*, ESE Discussion Paper 186, Edinburgh School of Economics, University of Edinburgh, 2008, https://www.research.ed.ac.uk/en/publications/the-el-farol-bar-problem-revisited-reinforcement-learning-in-a-po/.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementComputationalPDE.pdf