Check for
updates

Exact Solution to Two-Body Financial Dealer Model: Revisited from the Viewpoint of Kinetic Theory

Kiyoshi Kanazawa[1,2] · **Hideki Takayasu[3,4]** · **Misako Takayasu[4]**

Abstract

The two-body stochastic dealer model is revisited to provide an exact solution to the average order-book profile using the kinetic approach. The dealer model is a microscopic financial model where individual traders make decisions on limit-order prices stochastically and then reach agreements on transactions. In the literature, this model was solved for several cases: an exact solution for two-body traders $N = 2$ and a mean-field solution for many traders $N \gg 1$. Remarkably, while kinetic theory plays a significant role in the mean-field analysis for $N \gg 1$, its role is still elusive for the case of $N = 2$. In this paper, we revisit the two-body dealer model $N = 2$ to clarify the utility of the kinetic theory. We first derive the exact master-Liouville equations for the two-body dealer model. We next illustrate the physical picture of the master-Liouville equation from the viewpoint of the probability currents. The master-Liouville equations are then solved exactly to derive the order-book profile and the average transaction interval. Furthermore, we introduce a generalised two-body dealer model by incorporating interaction between traders via the market midprice and exactly solve the model within the kinetic framework. We finally confirm our exact solution by numerical simulations. This work provides a systematic mathematical basis for the econophysics model by developing better mathematical intuition.

Communicated by Francesco Zamponi.

✉ Kiyoshi Kanazawa
  kiyoshi@sk.tsukuba.ac.jp

[1]   Faculty of Engineering, Information and Systems, University of Tsukuba, Tennodai, Tsukuba, Ibaraki 305-8573, Japan

[2]   JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

[3]   Sony Computer Science Laboratories, 3-14-13 Higashi-Gotanda, Shinagawa-ku, Tokyo 141-0022, Japan

[4]   Institute of Innovative Research, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8502, Japan

---

🌀 Springer

1 Introduction

Statistical physics is a powerful tool to understand the macroscopic behaviour of physical systems from their microscopic dynamics [1], and one of the historical landmarks is the kinetic theory for the Brownian motions [2]. For example, let us consider the dynamics of a small particle in water. This particle experimentally exhibits random motions due to molecular collisions, and such a physical picture is mathematically summarised within the kinetic theory. Indeed, the Langevin equation can be derived from the microscopic physical dynamics via the Liouville equation, Bogoliubov–Born–Green–Kirkwood–Yvon (BBGKY) hierarchy [3], Boltzmann equation, and the diffusive limit [4, 5].

Remarkably, the Brownian motions are ubiquitously observed in broad areas, such as in financial markets [6]. For example, the price timeseries of stock and foreign exchange rate exhibits random motions similar to the Brownian motion, which was historically pointed out by Bachelier [7] earlier than Einstein's kinetic theory [8]. Physicists curious about this similarity have studied the financial-market microstructure in terms of statistical physics, hoping to understand its macroscopic behaviour from its microscopic dynamics as a research activity of econophysics [9–12]. While there have been various econophysics models proposed, in this report, we focus on a microscopic financial-market model called the dealer model [13–17], which depicts the decision-making processes dynamics on the level of individual traders.

The dealer model is one of the earliest microscopic models that describe individual traders' decision-making process in econophysics [13]. It was first introduced as a deterministic dynamical model [13, 14] and was later extended for a stochastic model [15–17] for mathematical simplicity. The stochastic dealer model can be mathematically solved for several cases: the two-body case (see Ref. [15] for $N = 2$) and the mean-field case (see Refs. [16, 17] for $N \gg 1$) with the total number of traders N. The two-body case $N = 2$ exactly was solved in Ref. [15], where the dynamics are finally mapped into the first-hitting time problem to obtain the exact transaction-interval statistics. The mean-field case $N \gg 1$ was solved in Refs. [16, 17] by using the kinetic theory: i.e., the financial Boltzmann and Langevin equations are derived by starting from the financial Liouville equation (i.e., the master equation for the many-body joint probability density function (PDF)) and the BBGKY hierarchy. This method finally deduces the average order-book profile and transaction-interval statistics within the mean-field approximation, which exhibits excellent numerical agreements. In this sense, the dealer model is well-tractable in terms of statistical-physics theories.

However, the relationship between the previous two-body solution and the mean-field solution is still elusive since their methodologies are formally different. While the formal BBGKY hierarchy for the two-body case was briefly derived in the supplementary material of Ref. [16] in a rather incomplete form, its solution was not thoroughly analyzed in Refs. [16, 17] to understand the mathematical characters specific to the two-body case $N = 2$, because the main focus of Refs. [16, 17] was to analyze the mean-field case $N \gg 1$.

In this report, we revisit the two-body stochastic dealer model to clarify the technical and mathematical roles of the kinetic framework therein. By applying Novikov's theorem for the coloured noise, we first derive two exact master-Liouville (ML) equations (i.e., one is a reduced form and the other is the complete form) for the two-body dealer model by fully incorporating the "collision" mechanism. Before the technical derivation, we provide a technical review of Novikov's theorem [18, 19] and the Liouville equation for collisional dynamics because they are advanced topics for non-Markovian stochastic processes and

kinetic theory.[1] We also illustrate the physical meaning of the ML equations from the viewpoint of the probability current. The probability-current picture will help readers develop better intuition and catch the sense of the specific mathematical forms of the ML equations. Furthermore, the ML equation is exactly solved to obtain the average order-book profile and the average transaction interval.To demonstrate the power of our kinetic framework, we generalise the dealer model to incorporate market-midprice interaction, and solve the model exactly. Finally, we have numerically confirmed the exact solution.

This report is organised as follows. We explain the mathematical model of the stochastic dealer model for $N = 2$ in Sect. 2. In Sect. 3, we provide a technical review on Novikov's theorem and a manipulation technique of the δ-functions for collisional dynamics. In Sect. 4, we derive a reduced ML equation exactly. We exactly solve the reduced ML equation in Sect. 5 to obtain the average order-book profile and the average transaction interval. In Sect. 6, we derive the full ML equation and examine its physical meaning from the viewpoint of the probability current. In addition, we examine the consistency between the reduced and full ML equations. In Sect. 7, we consider a generalised dealer model by incorporating interaction between traders via the market midprice and exactly solve the generalised model in the kinetic framework again. Finally, we show the numerical simulations to confirm the validity of the exact solutions in Sect. 8. This report is concluded in Sect. 9 with some remarks. Eleven appendices complement these sections.

2 Model

We formulate the stochastic dealer model based on the Markovian stochastic processes after explaining our mathematical notation.

2.1 Mathematical Notation

Here we briefly explain our mathematical notation. In this report, any stochastic variable accompanies the hat symbol, such as $\hat{A}$ instead of A. In addition, the ensemble average of any stochastic variable is denoted by $\langle \hat{A}(t) \rangle$ at the continuous physical time t. All the models are based on the continuous-time Markovian dynamics. Using this notation and the δ function,[2] the PDF of $\hat{A}(t)$ is given by $P_t(A) := \langle \delta(A - \hat{A}(t)) \rangle$ by stressing the difference between the stochastic variable $\hat{A}(t)$ and the non-stochastic real number A. Inversely, the ensemble average $\langle \hat{A}(t) \rangle$ can be rewritten as

$$\langle \hat{A}(t) \rangle = \int A P_t(A) dA. \tag{1}$$

We use the square brackets for any functional argument, such as $f[\{x(s)\}_{s \in \mathbf{R}}]$, to stress that f is a functional but not a function, where $\mathbf{R} := (-\infty, \infty)$ is the set of the real numbers and $s \in \mathbf{R}$ is a time point. We sometimes abbreviate the argument, such that $f[x] := f[\{x(s)\}_s]$ if its meaning is obvious from the context. For any functional, the functional derivative is written by $(\delta f[x])/(\delta x(s))$.

---

[1] Experts on these topics can skip this review section because the main-results sections are self-contained. However, we believe such an introductory review section will be helpful for readers unfamiliar with mathematical technicalities since our kinetic theory for financial Brownian motion is very interdisciplinary, requiring the background of econophysics, non-Markovian stochastic processes, and kinetic theory.

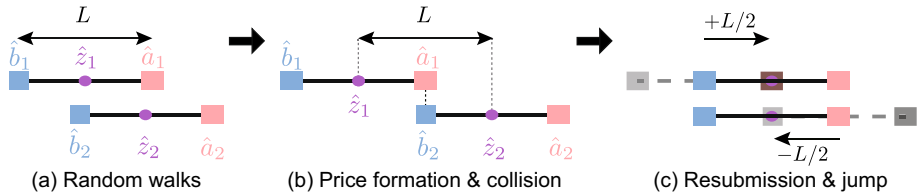[2] Readers unfamiliar with the δ functions are referred to Appendix A

Fig. 1 Schematic dynamics of the two-body stochastic dealer model. **a** In the absence of transactions $|\hat{z}_1 - \hat{z}_2| < L$, both $\hat{z}_1$ and $\hat{z}_2$ exhibits random walks. **b** When the condition $|\hat{z}_1 - \hat{z}_2| = L$ is met, a transaction happens and thus, the transacted price is updated (i.e., the price formation). These dynamics are essentially similar to the collisions in kinetic theory. **c** Just after the transaction, both traders resubmit their prices far from the transacted price. This resubmission process is mathematically implemented as jumps of $\hat{z}_1$ and $\hat{z}_2$

In this report, we regard the derivative symbols $\partial/\partial x$ and $\delta/\delta x(t)$ as linear operators acting on all subsequent terms. For example, the derivative $\partial/\partial x$ in the following formula is interpreted as

$$\frac{\partial}{\partial x}\alpha(x)\frac{\partial}{\partial x}\beta(x)P_t(x) := \frac{\partial}{\partial x}\left[\alpha(x)\left\{\frac{\partial}{\partial x}(\beta(x)P_t(x))\right\}\right]. \tag{2}$$

2.2 Model Dynamics

We consider a market composed of two traders, always quoting both bid and ask prices simultaneously. In this report, $\hat{b}_i(t)$ and $\hat{a}_i(t)$ denotes the bid and ask prices of the ith trader for $i = 1, 2$ at the continuous physical time $t \in \mathbf{R}$. The difference between the ask and bid prices $\hat{L}_i := \hat{a}_i - \hat{b}_i > 0$ is called the spread of the ith trader. We assume that the traders' spreads are the same constant, such that $\hat{L}_i = L = \text{const.} > 0$ for $i = 1, 2$ (see Fig. 1a). We also define the midprices of the trader i as $\hat{z}_i := (\hat{a}_i + \hat{b}_i)/2$. The transaction condition (see Fig. 1b) is given by

$$a_1 = b_2 \text{ or } a_2 = b_1 \implies |z_1 - z_2| = L. \tag{3}$$

In this report, we analyze the simplest case based on the model 1 in Ref. [15] and its generalisation in Sect. 7. The dynamics of the midprices $\{\hat{z}_i\}_{i=1,2}$ obey the simple Brownian motion in the absence of transactions $|\hat{z}_1 - \hat{z}_2| < L$ (see Fig. 1a):

$$\frac{d\hat{z}_i}{dt} = \sigma\hat{\xi}_i^{\mathrm{G}} \tag{4a}$$

with the white Gaussian noise $\hat{\xi}_i^{\mathrm{G}}(t)$ and a positive constant $\sigma > 0$. The white Gaussian noise is formally defined as the derivative of the Wiener process $\hat{W}_i(t)$ as $\hat{\xi}_i^{\mathrm{G}}(t) := d\hat{W}_i(t)/dt$ and satisfies

$$\langle\hat{\xi}_i^{\mathrm{G}}(t)\rangle = 0, \quad \langle\hat{\xi}_i^{\mathrm{G}}(t_1)\hat{\xi}_j^{\mathrm{G}}(t_2)\rangle = \delta(t_1 - t_2). \tag{4b}$$

We note that the higher-order cumulants are absent due to the Gaussian nature.[3] In the presence of transactions, we assume that both traders requote their prices far from the transaction

---

[3] Formally, the nth cumulant of any stochastic variable $\hat{A}(t)$ is defined by (see Ref. [1] for example)

$$\langle\hat{A}(t_1)\dots\hat{A}(t_n)\rangle_{\mathrm{c}} := \frac{\delta^n}{\delta\zeta(t_1)\dots\delta\zeta(t_n)}\log\Psi[\zeta], \quad \Psi[\zeta] := \left\langle\exp\left[i\int_0^t ds\zeta(s)\hat{A}s\right]\right\rangle.$$

For any Gaussian noise, the higher-order cumulants are absent: $\langle\hat{A}(t_1)\dots\hat{A}(t_n)\rangle_{\mathrm{c}} = 0$ for $n \geq 3$.
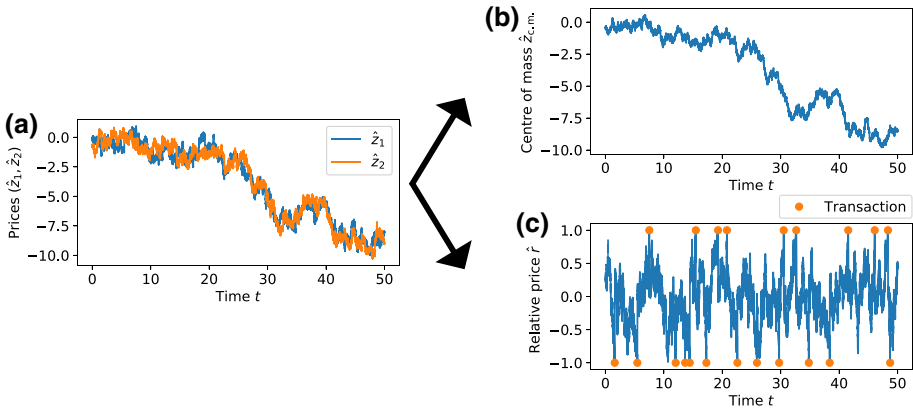
Fig. 2 Useful representation based on the centre of mass $\hat{z}_{cm}$ and the relative price $\hat{r}$. **a** The original dynamics based on $(\hat{z}_1, \hat{z}_2)$ is equivalent to those based on $(\hat{z}_{cm}, \hat{r})$. **b** The centre of mass $\hat{z}_{cm}$ exhibits the simple random walks (6a), irrelevantly to transactions (**b**). **c** The relative price $\hat{r}$ shows random walks confined in the regime $\hat{r} \in (-L/2, L/2)$. At the boundaries $r = \pm L/2$, the relative price goes back to the origin as described by Eq. (6b) due to transactions. Here we assume $L = 2$

price to avoid immediate transactions, according to the following equation (see Fig. 1b and c)

$$a_i(t) = b_j(t) \implies a_i(t + dt) = a_i(t) + \frac{L}{2}, \quad b_j(t + dt) = b_j(t) - \frac{L}{2} \quad (4c)$$

for an infinitesimal positive $dt > 0$, or equivalently

$$|z_1(t) - z_2(t)| = L \implies z_1(t + dt) = z_2(t + dt) = z_1(t) - \frac{L}{2} \operatorname{sgn}(z_1(t) - z_2(t)) \quad (4d)$$

with the sign function (i.e., $\operatorname{sgn}(x) = +1$ for $x > 0$, $\operatorname{sgn}(x) = 0$ for $x = 0$, and $\operatorname{sgn}(x) = -1$ for $x < 0$). Remarkably, this transaction rule does not change the centre of mass.

We note that this model is the most basic model; we can introduce additional elements, such as trend-following strategies (see Refs. [15–17] for examples). However, since our motivation is to reveal the mathematical characters of the dealer model from the kinetic viewpoint, we start the model as simple as possible and then consider its generalisation in Sect. 7.

2.3 Review of the Previous Solution Based on the First-Hitting Time Problem in Ref. [15]

Here we review the previous solution to this model based on Ref. [15] to clarify the difference to our approach. In Ref. [15], this model was mapped into the first-hitting time problem of the one-dimensional Brownian motion. Specifically, by introducing the centre of mass $\hat{z}_{cm}$, the relative price $\hat{r}$ (see Fig. 2), and the noise variance for the centre of mass σ_{cm}, defined by

$$\hat{z}_{cm} := \frac{\hat{z}_1 + \hat{z}_2}{2}, \quad \hat{r} := \hat{z}_1 - \hat{z}_{cm} = \frac{\hat{z}_1 - \hat{z}_2}{2}, \quad \sigma_{cm} := \frac{\sigma}{\sqrt{2}}, \quad (5)$$

we obtain

$$\hat{z}_{\text{cm}}(t + dt) = \hat{z}_{\text{cm}}(t) + dt\sigma_{\text{cm}}\hat{\chi}(t), \tag{6a}$$

$$\hat{r}(t + dt) = \begin{cases} \hat{r}(t) + dt\sigma_{\text{cm}}\hat{\eta}(t) & (|\hat{r}(t)| < L/2) \\ 0 & (|\hat{r}(t)| = L/2) \end{cases} \tag{6b}$$

for an infinitesimal positive timestep $dt > 0$. Here the two white noises are introduced as

$$\hat{\chi}(t) := \frac{1}{\sqrt{2}}\left(\hat{\xi}_1(t) + \hat{\xi}_2(t)\right), \quad \hat{\eta}(t) := \frac{1}{\sqrt{2}}\left(\hat{\xi}_1(t) - \hat{\xi}_2(t)\right), \tag{6c}$$

which satisfy

$$\langle\hat{\chi}(t)\rangle = 0, \quad \langle\hat{\eta}(t)\rangle = 0 \tag{6d}$$

$$\langle\hat{\chi}(t_1)\hat{\chi}(t_2)\rangle = \delta(t_1 - t_2), \quad \langle\hat{\eta}(t_1)\hat{\eta}(t_2)\rangle = \delta(t_1 - t_2), \quad \langle\hat{\chi}(t_1)\hat{\eta}(t_2)\rangle = 0. \tag{6e}$$

We thus find that

1. the centre of mass $\hat{z}_{\text{cm}}$ obeys the random walk irrelevantly to the transactions because the centre of mass does not move during transactions;
2. the relative price $\hat{r}$ obeys the random walk and returns back to the origin if it hits the boundaries at $\hat{r} = \pm L/2$.

These facts imply that the dynamics of the two-body dealer model is equivalent to the first-hitting time problem at the boundaries $\hat{r} = \pm L/2$. Based on these characters, the authors of Ref. [15] solved this first-hitting time problem to obtain the transaction-interval statistics. For example, the mean transaction interval is given by

$$\langle\hat{\tau}\rangle = \frac{L^2}{2\sigma^2} \tag{7}$$

with the transaction interval $\hat{\tau}$. Based on such analytic formulas on the transaction-interval statistics, the authors of Ref. [15] derived the statistical properties of various versions of the dealer model.

In addition, the diffusion constant of the centre of mass is exactly given by

$$D(N = 2) := \frac{\sigma_{\text{cm}}^2(N = 2)}{2}, \quad \sigma_{\text{cm}}^2(N = 2) := \frac{\sigma^2}{2}. \tag{8}$$

This means that the diffusive speed of the centre of mass is slower than that of a single trader. According to the mean-field theory [17], in general, the diffusion constant of the centre of mass is given by the scaling

$$D(N) := \frac{\sigma_{\text{cm}}^2(N)}{2}, \quad \sigma_{\text{cm}}^2(N) \propto \frac{\sigma^2}{N}, \tag{9}$$

suggesting that the centre of mass is "heavier" for large N. In this sense, $\sigma_{\text{cm}}(N = 2)$ can be regarded as the renormalized diffusion constants as the result of the many-body interaction.

2.4 Goal of This Report: Reformulation Based on Kinetic Theory

While the previous derivation is powerful in understanding the two-body dealer model, its relationship is not clear to the mean-field solution proposed in Refs. [16, 17]. The mean-field

solution for $N \gg 1$ in Refs. [16, 17] is based on kinetic theory: the microscopic dynamics described by the high-dimensional stochastic differential equations are mapped onto the financial Liouville equation (i.e., an ML equation for the joint PDF), which is reduced to the Boltzmann and Langevin equations via the BBGKY hierarchy. This methodology is very different from the previous first-passage-problem approach, at least superficially.

This report aims to fill this technical gap mathematically: we derive the Liouville equation for the two-body dealer model exactly. Furthermore, the ML equation is solved exactly to obtain the average order-book profile and the mean transaction interval. In addition, we examine the mathematical consistency between these different approaches to develop our better intuition.

3 Technical Review

Here we review Novikov's theorem [18, 19] and the collision problem together with related manipulation technique of the δ functions. These technical methods are relevant to the derivation of the ML equations.

This section aims to provide a technical background for the general audience unfamiliar with non-Markovian stochastic processes and kinetic theories as an elementary review. Since the main contents are self-contained without needing to refer to this section, professional readers can skip this section and proceed with the main results from Sect. 4. However, we believe that such a preliminary review will be helpful for the general audience because the kinetic framework for the dealer model is based on multidisciplinary mathematical backgrounds.

3.1 Stochastic Calculus for Coloured Noise: Novikov's Theorem

3.1.1 Setup

We first review Novikov's theorem [18, 19], a useful theorem for stochastic calculus for coloured noise. This technique will be used for the stochastic dealer model in Sect. 4.1, and we provide its elementary introduction for a simple case here.

Ornstein-Uhlenbeck coloured noise. Let us first consider an Ornstein-Uhlenbeck (OU) process with a non-zero correlation time $\epsilon > 0$:

$$\frac{d\hat{\eta}_\epsilon}{dt} = -\frac{\hat{\eta}_\epsilon}{\epsilon} + \hat{\xi}^{\mathrm{G}} \tag{10}$$

with the standard white Gaussian noise $\hat{\xi}^{\mathrm{G}}$, satisfying $\langle \hat{\xi}^{\mathrm{G}}(t) \rangle = 0$ and $\langle \hat{\xi}^{\mathrm{G}}(t_1)\hat{\xi}^{\mathrm{G}}(t_2) \rangle = \delta(t_1 - t_2)$. This OU process can be regarded as a specific implementation of the *coloured noise* because it satisfies

$$\langle \hat{\eta}_\epsilon(t) \rangle = 0, \quad \langle \hat{\eta}_\epsilon(t_1)\hat{\eta}_\epsilon(t_2) \rangle = \frac{1}{2\epsilon}e^{-|t_1-t_2|/\epsilon}, \tag{11}$$

showing the exponential decay of the correlation with the short characteristic timescale ϵ. We therefore call $\hat{\eta}_\epsilon$ the *OU coloured noise* and we use the OU coloured noise for a technical calculation of kinetic theory. We note that the OU coloured noise converges to the white noise for the small ϵ limit, such that

$$\lim_{\epsilon \downarrow 0} \langle \hat{\eta}_\epsilon(t_1)\hat{\eta}_\epsilon(t_2) \rangle = \delta(t_1 - t_2). \tag{12}$$

In addition, the OU coloured noise belongs to the Gaussian noise, which can be proved as follows. It is known that any superposition of Gaussian random variables also obeys the Gaussian statistics [1]. Because the OU coloured noise can be represented as the linear superposition of the white Gaussian noise,

$$\hat{\eta}_\epsilon(t) = \hat{\eta}_\epsilon(0) + \int_0^t e^{-(t-s)/\epsilon} \hat{\xi}^{\mathrm{G}}(s) ds, \tag{13}$$

the OU coloured noise also belongs to the Gaussian noise.

Technical advantage of the OU coloured noise. One of the technical advantages to introduce the coloured noise is to remove the singularity of the white noise. Indeed, the white noise is an unbounded and discontinuous function of time, which causes various delicate mathematical issues, such as the Itô-Stratonovich dilemma [20]. On the other hand, the OU coloured noise is a bounded and continuous function of time, and various delicate issues disappear for nonzero $\epsilon > 0$. Therefore, in our formulation, we first consider the coloured noise to avoid delicate issues, proceed with calculations for nonzero $\epsilon > 0$, and, finally, take the white-noise limit $\epsilon \downarrow 0$.

Stochastic dynamics under coloured noise. Next, let us consider the stochastic dynamics of a stochastic variable $\hat{x}(t)$ driven by the OU coloured noise:

$$\frac{d\hat{x}}{dt} = -\alpha(\hat{x}) + \beta(\hat{x})\hat{\eta}_\epsilon \tag{14}$$

with smooth functions $\alpha(\hat{x})$ and $\beta(\hat{x})$. Assuming that the state space is one-dimensional, this system can be regarded as non-Markovian since the OU coloured noise has a finite correlation time. In this sense, the standard mathematical method for the Markovian stochastic processes is not directly available for analytical calculations.

3.1.2 Prescription 1: Markov Embedding

Because many natural phenomena obey non-Markovian dynamics, various technical methods have been developed in statistical physics for such non-Markovian systems. One established method is based on the Markov embedding technique. This technique is based on converting the original non-Markovian onto an auxiliary higher-dimensional Markovian system by sufficiently adding many auxiliary variables. For example, let us reconsider the non-Markovian dynamics (14) from the viewpoint of the Markov embedding. While the dynamics (14) is non-Markovian in the one-dimensional space $\hat{x}(t)$, it can be interpreted as a two-dimensional Markovian dynamics specified by the two-dimensional state vector $\hat{\Gamma}(t) := (\hat{x}(t), \hat{\eta}_\epsilon(t))$. Indeed, $\hat{\Gamma}(t)$ obeys the two-dimensional simultaneous SDEs (10) and (14) that are Markovian, and thus we can derive the Fokker-Planck equation for the two-dimensional joint PDF $P_t(x, \eta_\epsilon)$. In this sense, some non-Markovian processes can be converted onto high-dimensional Markovian dynamics by the Markov embedding (e.g., see also Refs. [21–23] for the generalised Langevin equations and Refs. [23–25] for the Hawkes processes).

3.1.3 Prescription 2: Functional Calculus and Novikov's Theorem

One of the other famous methods is based on the functional calculus for non-Markovian stochastic paths, to which Novikov's theorem [18] belongs. Novikov's theorem is available for coloured Gaussian noises to evaluate ensemble averages (see Ref. [19] for its generalisation for non-Gaussian coloured noises). This theorem states that an identity
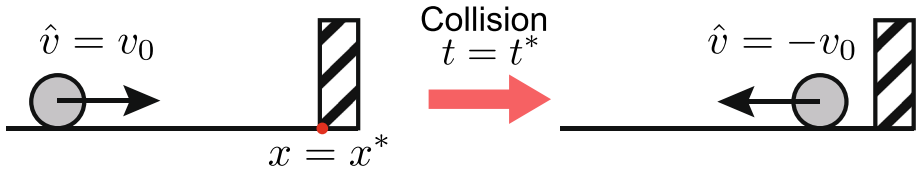
Fig. 3 Schematic of a single-particle collision. The rigid particle collides against the rigid wall placed at $x = x^*$. Here the particle is regarded as a mass point by ignoring the particle size. After the collision, the particle velocity is inverted as $\hat{v} = -v_0$. The initial condition is given by $(\hat{x}(0), \hat{v}(0)) = (x_0, v_0)$ satisfying $x_0 < x^*$ and $v_0 > 0$

$$\langle \hat{\eta}_\epsilon(t) g[\hat{\eta}_\epsilon; t] \rangle = \int_0^t ds \langle \hat{\eta}_\epsilon(t) \hat{\eta}_\epsilon(s) \rangle \left\langle \frac{\delta g[\hat{\eta}_\epsilon; t]}{\delta \hat{\eta}_\epsilon(s)} \right\rangle \tag{15a}$$

holds for any functional $g[\hat{\eta}_\epsilon; t] := g[\{\hat{\eta}_\epsilon(s)\}_{0 \leq s \leq t}]$ dependent on the path of the coloured Gaussian noise $\{\hat{\eta}_\epsilon(s)\}_{0 \leq s \leq t}$ (see Appendix B for the derivation). Particularly, for the special case where g is a function of $\hat{x}(t)$, we obtain

$$\langle \hat{\eta}_\epsilon(t) g(\hat{x}(t)) \rangle = \int_0^t ds \langle \hat{\eta}_\epsilon(t) \hat{\eta}_\epsilon(s) \rangle \left\langle \frac{\delta \hat{x}(t)}{\delta \hat{\eta}_\epsilon(s)} \frac{\partial g(\hat{x}(t))}{\partial \hat{x}} \right\rangle. \tag{15b}$$

Readers unfamilar with the functional derivative $\delta/\delta \hat{\eta}_\epsilon(s)$ are referred to Appendix C.

Novikov's theorem is helpful to analyse non-Markovian stochastic processes, in particular for the white-noise limit $\epsilon \downarrow 0$. Assuming the SDE driven by the OU coloured noise (14), let us derive the Fokker-Planck equation for the white-noise limit [19]. We finally derive the Stratonovich-type Fokker-Planck equation

$$\frac{\partial}{\partial t} P_t(x) = \left[\frac{\partial}{\partial x} \alpha(x) + \frac{1}{2} \frac{\partial}{\partial x} \beta(x) \frac{\partial}{\partial x} \beta(x) \right] P_t(x) \tag{16}$$

for the white-noise limit $\epsilon \downarrow 0$ (see Appendix D for the derivation).

The resulting Stratonovich-type Fokker-Planck equation (16) is consistent with the fact that any SDE driven by a coloured Gaussian noise reduces to the Stratonovich-type SDE in the white-noise limit [26, 27]. We also note that this result can be confirmed via the Markov embedding approach (i.e., by directly addressing the Fokker-Planck equation for the two-dimensional joint PDF $P_t(x, \eta_\epsilon)$) using the projection operators (see Chapter 8 in Ref. [6]).

3.2 Collision Problem and Manipulation Technique of δ-Functions

We next review a collision problem and a related technique for the δ functions, which will be used for the ML equation for the dealer model in Sect. 4. Interested readers in the technicality are referred to textbooks on kinetic theory, such as Chapter 13 of Ref. [28] and Chapter 3 of Ref. [29].

3.2.1 Equation of Motion with a Deterministic Collision

This review subsection aims to provide an illustrative example of a deterministic collision based on Ref. [29] and then derive the Liouville equation. Let us consider a particle of motion

characterised by its position $\hat{x}(t)$ and velocity $\hat{v}(t)$. We place a rigid wall at $x = x^*$ at which the particle's velocity is inverted due to collision (see Fig. 3). The dynamics are given by

$$\frac{d\hat{x}}{dt} = \hat{v} \tag{17a}$$

$$\frac{d\hat{v}}{dt} = -2\hat{v}(t)\delta(t - t^*) \tag{17b}$$

with the collision time t^*. The collision time t^* is determined by

$$\hat{x}(t^*) = x_0 + t^* v_0 = x^* \implies t^* := \frac{x^* - x_0}{v^*}, \tag{18}$$

where the initial condition is given by $\hat{x}(0) = x_0$ and $\hat{v}(0) = v_0$ satisfying $x_0 < x^*$ and $v_0 > 0$. Here the multiplication between $\hat{v}(t)$ and the δ function is interpreted in the Itô sense: $\hat{v}(t)\delta(t - t^*) = \lim_{h \downarrow 0} \hat{v}(t^* - h)\delta(t - t^*)$. The δ function represents the impulsive force due to the collision consistently with the invertion. Indeed, this time-evolution equation consistently deduces

$$\hat{v}(t^* + h) = \hat{v}(t^* - h) + \int_{t^*-h}^{t^*+h} dt \frac{d\hat{v}(t)}{dt} = \hat{v}(t^* - h) + \int_{t^*-h}^{t^*+h} dt \left\{ -2\hat{v}(t)\delta(t - t^*) \right\} = -\hat{v}(t^* - h) \tag{19}$$

for infinitesimal positive $h > 0$.

3.2.2 Liouville Equation

From Eq. (17), we derive the Liouville equation,[4] a time-evolution equation for the two-dimensional joint PDF $P_t(x, v) := \langle \delta(x - \hat{x}(t))\delta(v - \hat{v}(t)) \rangle$ (see Appendix E for the derivation):

$$\frac{\partial P_t(x, v)}{\partial t} = -v \frac{\partial P_t(x, v)}{\partial x} + [\Theta(v) P_t(x, v) - \Theta(-v) P_t(x, -v)] |v| \delta(x - x^*) \tag{20}$$

with the Heaviside function $\Theta(x)$, defined by

$$\Theta(x) := \begin{cases} 1 & (x > 0) \\ 1/2 & (x = 0) \\ 0 & (x < 0) \end{cases} . \tag{21}$$

It is straightforward to generalise this formulation for multi-body particle dynamics in principle by assuming the rigid spheres without contact friction, finally leading to the Liouville equation for many-body dynamics. In the standard program of the kinetic theory, the Boltzmann equation is finally deduced by integrating the Liouville equation with the assumption of molecular chaos.

The key technique for the derivation is the decomposition of the δ function. In general, the following identity holds for the δ-functions for an arbitrary function $g(t)$:

---

[4] Technically, this is called the *pseudo-Liouville equation* [28, 29]. While the conventional Liouville equation is local (i.e., a partial differential equation without jump terms) in classical mechanics, the pseudo-Liouville equation involves non-local jumps due to collisions.

$$\delta(g(t)) = \sum_i \frac{1}{|g'(t_i)|} \delta(t - t_i) \tag{22}$$

with the ith zero point t_i, by assuming $g'(t_i) \neq 0$. This technique is straightforwardly applied for the derivation of the ML equation for the dealer model in Sect. 4.1.2. Readers unfamiliar with the δ function are referred to Appendix A.

4 Result 1: A Reduced Master-Liouville Equation

We have provided enough preliminary review on Novikov's theorem and kinetic theory in Sect. 3. This section derives the master-Liouville equations for the two-body dealer model (4) or equivalently (6) using kinetic theory as the main results.

In this section, in particular, we focus on the ML equation for the relative price $\hat{r}$ obeying Eq. (6b). Finally, we will obtain a reduced ML equation as follows:

The time-evolution equation of the reduced PDF $P_t(r)$ is given by

$$\frac{\partial P_t(r)}{\partial t} = \frac{\sigma_{\text{cm}}^2}{2} \frac{\partial^2}{\partial r^2} P_t(r) + \sum_{s=\pm 1} [J_{t;s}(r + sL/2) - J_{t;s}(r)], \tag{23a}$$

$$J_{t;s}(r) := -s \frac{\sigma_{\text{cm}}^2}{2} \partial_{-s} P_t(sL/2) \delta(r - sL/2) = \frac{\sigma_{\text{cm}}^2}{2} |\partial_{-s} P_t(sL/2)| \, \delta(r - sL/2) \geq 0 \tag{23b}$$

for the one-dimensional PDF $P_t(r) := \langle \delta(r - \hat{r}(t)) \rangle$. The positive term $J_{t;s}(r)$ represents the *probability current* due to collisions, as will be discussed in detail in Sect. 4.2.

Here we have introduced the left ($s = -1$) and right ($s = +1$) derivatives, defined by

$$\partial_s f(r) := \lim_{h \downarrow 0} \frac{f(r + sh) - f(r)}{sh}. \tag{24}$$

Since the original dynamics (4) is two-dimensional specified by the full PDF $P_t(z_1, z_2) := \langle \delta(z_1 - \hat{z}_1(t)) \delta(z_2 - \hat{z}_2(t)) \rangle$, the one-dimensional PDF $P_t(r) := \langle \delta(r - \hat{r}(t)) \rangle$ can be called a *reduced PDF*. Correspondingly, we call Eq. (22) a *reduced master-Liouville equation* or just a reduced ML equation.[5]

Here we provide two independent derivations on the reduced ML equation (22): one is based on (i) Novikov's theorem for coloured noise [18, 19] and (ii) continuous limit from one-dimensional random walks on a regular lattice. While all methods deduce the same result, we believe that presenting various derivation methods will help the reader better understand our methodology.

---

[5] The time-evolution equation of PDFs is regularly called the *master equation*, the *differential form of the Chapman-Kolmogorov equation* or *Kolmogorov's forward equation*. Here we call such an equation the *master-Liouville equation* because it corresponds to the Liouville equation in the standard program of kinetic theory.
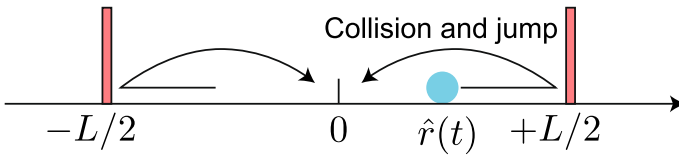
Fig. 4 The dynamics of the relative price $\hat{r}(t)$ can be regarded as a mechanical system where the particle diffuses and goes back to the origin after colliding against the boundaries at $r = \pm L/2$

4.1 Derivation Based on Novikov's Theorem for Coloured Noise

Here we derive the ML equation (22) using Novikov's theorem [18, 19] for the SDE (6b) for the relative price $\hat{r}(t)$. Interestingly, the SDE (6b) can be regarded as a Brownian motion confined in the boundaries at $r = \pm L/2$, at which the particle goes back to the origin (see Fig. 4). By regarding the particle arrivals at the boudaries $r = \pm L/2$ as "collisions", this system can be regarded as a mechanical system with collisions and jumps at the boundaries. This is the basic idea to apply the kinetic formulation to this dealer model.

This physical picture based on collisional jumps is very similar to the conventional kinetic formulation, and derivation of the ML equation should work well in the parallel method presented in Sect. 3.2. However, if we naively follow the same formal calculation as that in Sect. 3.2, we will encounter delicate mathematical issues originating from the δ singularity of the white noise. To solve this problem via kinetic theory, we reformulate the SDE (6b) by introducing three tricks:

1. replacement of the white noise $\hat{\eta}(t)$ with the OU coloured noise $\hat{\eta}_\epsilon(t)$ with nonzero correlation time $\epsilon > 0$;
2. manipulation of the δ functions to represent the boundary condition;
3. the white-noise limit $\epsilon \downarrow 0$ to keep the consistency with the original model.

Here the OU coloured noise is introduced to avoid technical delicate issues on the white noise. The OU coloured noise has the advantages that it is a bounded and continuous function of time and that the ordinary calculus is available for formal calculations by assuming non-zero $\epsilon > 0$. In the final stage of the calculation, we take the white-noise limit $\epsilon \downarrow 0$. Let us explain this procedure one by one as follow.

4.1.1 Reformulation of the SDE for the Relative Price $\hat{r}(t)$

To follow to these recipes, let us first replace the white noise $\hat{\eta}$ in Eq. (6b) with the OU coloured noise $\hat{\eta}_\epsilon$ defined by

$$\frac{d\hat{\eta}_\epsilon(t)}{dt} = -\frac{\hat{\eta}_\epsilon(t)}{\epsilon} + \hat{\xi}^G(t) \tag{25}$$

with a nonzero positive constant $\epsilon > 0$ and a white Gaussian noise $\hat{\xi}^G(t)$. This OU noise satisfies

$$\langle \hat{\eta}_\epsilon(t) \rangle = 0, \quad \langle \hat{\eta}_\epsilon(t_1)\hat{\eta}_\epsilon(t_2) \rangle = \frac{1}{2\epsilon}e^{-|t_1-t_2|/\epsilon}, \tag{26}$$

which reduces to the white noise for $\epsilon \downarrow 0$:

$$\lim_{\epsilon \downarrow 0}\langle \hat{\eta}_\epsilon(t_1)\hat{\eta}_\epsilon(t_2) \rangle = \delta(t_1 - t_2). \tag{27}$$

We then rewrite Eq. (6b) as

$$
d\hat{r}(t) = \begin{cases} dt\sigma_{\mathrm{cm}}\hat{\eta}_\epsilon(t) & (\text{without collisions: } \hat{\tau}_{s;i} \notin [t, t+dt)) \\ -sL/2 & (\text{with a collision: } \hat{\tau}_{s;i} \in [t, t+dt)) \end{cases},
\tag{28}
$$

with $d\hat{r}(t) := \hat{r}(t+dt) - \hat{r}(t)$ for an infinitesimal positive $dt > 0$. Here $\hat{\tau}_{s;i}$ is the ith arrival time of the particle at $\hat{r}(\tau_{s;i}) = sL/2$ for $s = \pm 1$. Using the δ function, this is equivalent to

$$
\frac{d\hat{r}(t)}{dt} = \sigma_{\mathrm{cm}}\hat{\eta}_\epsilon(t) + \sum_{s=\pm 1}\sum_{i=1} \frac{-sL}{2}\delta(t - \hat{\tau}_{s;i}).
\tag{29}
$$

Considering the collison directions, we note that the velocity must be positive (negative) at the arrival time at $r = +L/2$ ($r = -L/2$):

$$
\lim_{h\downarrow 0}\frac{d\hat{r}(\hat{\tau}_{+1;i} - h)}{dt} = \sigma_{\mathrm{cm}}\lim_{h\downarrow 0}\hat{\eta}_\epsilon(\hat{\tau}_{+1;i} - h) > 0, \quad \lim_{h\downarrow 0}\frac{d\hat{r}(\hat{\tau}_{-1;i} - h)}{dt} = \sigma_{\mathrm{cm}}\lim_{h\downarrow 0}\hat{\eta}_\epsilon(\hat{\tau}_{-1;i} - h) < 0.
\tag{30}
$$

In other words,

$$
s\lim_{h\downarrow 0}\frac{d\hat{r}(\hat{\tau}_{s;i} - h)}{dt} = s\sigma_{\mathrm{cm}}\lim_{h\downarrow 0}\hat{\eta}_\epsilon(\hat{\tau}_{s;i} - h) > 0.
\tag{31}
$$

for both $s = \pm 1$. Remarkably, this mathematical structure is essentially similar to the conventional kinetic theory for collision (see Eq. (153) in Appendix E, which is based on Eq. (22) in Sect. 3.2). This technical issue is important in removing the absolute operators of the OU coloured noises as shown in Sect. 4.1.2.

4.1.2 Dynamics of an Arbitrary Function $f(\hat{r}(t))$

We also consider the dynamics of an arbitrary function $f(r)$

$$
df(\hat{r}) = \begin{cases} \sigma_{\mathrm{cm}}\dfrac{df(\hat{r})}{dr}\hat{\eta}_\epsilon dt & (\text{without collisions: } \hat{\tau}_{s;i} \notin [t, t+dt)) \\ f(\hat{r} - sL/2) - f(\hat{r}) & (\text{with a collision: } \hat{\tau}_{s;i} \in [t, t+dt)) \end{cases}
\tag{32}
$$

with $df(r(t)) := f(r(t+dt)) - f(r(t))$. Using the δ-function, this relation can be rewritten as

$$
\frac{df(\hat{r})}{dt} = \underbrace{\sigma_{\mathrm{cm}}\frac{df(\hat{r})}{dr}\hat{\eta}_\epsilon}_{\text{without collisions during }[t, t+dt)} + \underbrace{\sum_{s=\pm 1}\sum_{i=1}\left[f(\hat{r} - sL/2) - f(\hat{r})\right]\delta(t - \hat{\tau}_{s;i})}_{\text{with collisions}}.
\tag{33}
$$

Here we use a decomposition formula for the δ functions:

$$
\delta(g(t)) = \sum_i \frac{1}{|g'(t)|}\delta(t - t_i) \implies h(t)|g'(t)|\delta(g(t_i)) = \sum_i h(t)\delta(t - t_i)
\tag{34}
$$

for an arbitrary functions $g(t)$ and $h(t)$, where t_i is the ith zero point of $g(t)$, defined by $g(t_i) = 0$ and $t_i < t_{i+1}$. Using this formula, we obtain

$$
\frac{df(\hat{r})}{dt} = \sigma_{\mathrm{cm}}\frac{df(\hat{r})}{dr}\hat{\eta}_\epsilon + \sigma_{\mathrm{cm}}\sum_{s=\pm 1}\left[f(\hat{r} - sL/2) - f(\hat{r})\right]\left|\frac{d\hat{r}}{dt}\right|\delta(\hat{r} - sL/2)
$$

$$
= \sigma_{\mathrm{cm}}\frac{df(\hat{r})}{dr}\hat{\eta}_\epsilon + \sigma_{\mathrm{cm}}\sum_{s=\pm 1}\left[f(\hat{r} - sL/2) - f(\hat{r})\right]|\hat{\eta}_\epsilon|\delta(\hat{r} - sL/2).
\tag{35}
$$

By considering the physical picture that the velocity $d\hat{r}/dt$ must be positive (negative) when the particle hits the boundary $\hat{r} = L/2$ ($\hat{r} = -L/2$) as summarised in Eq. (31), we can remove the absolue operator[6] in Eq. (35) such that $|\hat{\eta}_\epsilon|\delta(\hat{r} - sL/2) = s\hat{\eta}_\epsilon\delta(\hat{r} - sL/2)$, leading to

$$\frac{df(\hat{r})}{dt} = \sigma_{cm}\frac{df(\hat{r})}{dr}\hat{\eta}_\epsilon + \sigma_{cm}\sum_{s=\pm 1} s\left[f(\hat{r} - sL/2) - f(\hat{r})\right]\hat{\eta}_\epsilon\delta(\hat{r} - sL/2). \qquad (36)$$

4.1.3 Ensemble Average and Novikov's Theorem for the White-Noise Limit

Let us take the ensemble average:

$$\left\langle\frac{df(\hat{r})}{dt}\right\rangle = \left\langle\sigma_{cm}\frac{df(\hat{r})}{d\hat{r}}\hat{\eta}_\epsilon + \sigma_{cm}\sum_{s=\pm 1} s\left[f(0) - f(sL/2)\right]\hat{\eta}_\epsilon\delta(\hat{r} - sL/2)\right\rangle. \qquad (37)$$

Let us evaluate the diffusive term $\langle(df(\hat{r})/d\hat{r})\hat{\eta}_\epsilon\rangle$ and the collision term $\langle\hat{\eta}_\epsilon\delta(\hat{r} - sL/2)\rangle$ one by one.

Novikov's theorem for short-time interval. To evaluate the ensemble averages, we use Novikov's theorem that is valid for an arbitrary coloured Gaussian noise [18, 19]:

$$\langle\hat{\eta}_\epsilon(t)g(\hat{r}(t))\rangle = \int_0^t dt'\langle\hat{\eta}_\epsilon(t)\hat{\eta}_\epsilon(t')\rangle\left\langle\frac{\delta g(\hat{r}(t))}{\delta\hat{\eta}_\epsilon(t')}\right\rangle. \qquad (38)$$

Since the correlation time ϵ is finally set infinitesimal, we evaluate this ensemble average by dropping minor correction terms disappearing for the white noise limit $\epsilon \downarrow 0$. Here we remark a useful identity identity for integrals with short memory:

$$\lim_{\epsilon\downarrow 0}\int_0^t dt'\frac{e^{-(t-t')/\epsilon}}{2\epsilon}f(t') = \lim_{\epsilon\downarrow 0}\int_{t_{ini}(\epsilon)}^t dt'\frac{e^{-(t-t')/\epsilon}}{2\epsilon}f(t') = f(t),\ t_{ini}(\epsilon) := t-\epsilon^{1/2} \quad (39)$$

for any function $f(t)$ that decays sufficiently rapidly for $t \to \infty$. This relation holds because

$$\lim_{\epsilon\downarrow 0}\int_0^t dt' f(t')\frac{e^{-(t-t')/\epsilon}}{2\epsilon} = \lim_{\epsilon\downarrow 0}\int_0^\infty d\tilde{t}\, f(t - \epsilon\tilde{t})e^{-\tilde{t}}/2 = f(t) \qquad (40)$$

and

$$\lim_{\epsilon\downarrow 0}\int_{t_{ini}(\epsilon)}^t dt' f(t')\frac{e^{-(t-t')/\epsilon}}{2\epsilon} = \lim_{\epsilon\downarrow 0}\int_0^{\epsilon^{-1/2}} d\tilde{t}\, f(t - \epsilon\tilde{t})e^{-\tilde{t}}/2 = f(t) \qquad (41)$$

with the variable transformation $\tilde{t} := (t - t')/\epsilon$. This result is intuitively reasonable because the integral interval $t - t_{ini}(\epsilon) = \epsilon^{1/2}$ is much larger than the correlation time ϵ: $\epsilon^{1/2} \gg \epsilon$ for $\epsilon \downarrow 0$.

Considering this relationship, it is sufficient to consider the contribution of the path $\{\hat{r}(t'')\}_{t''\in[t_{ini}(\epsilon),t)}$ with sufficiently short-time interval $t - t_{ini}(\epsilon) = \epsilon^{1/2}$:

$$\lim_{\epsilon\downarrow 0}\langle\hat{\eta}_\epsilon(t)g(\hat{r}(t))\rangle = \lim_{\epsilon\downarrow 0}\int_{t_{ini}(\epsilon)}^t dt'\langle\hat{\eta}_\epsilon(t)\hat{\eta}_\epsilon(t')\rangle\left\langle\frac{\delta g(\hat{r}(t))}{\delta\hat{\eta}_\epsilon(t')}\right\rangle. \qquad (42)$$

---

[6] The introduction of the OU coloured noise $\hat{\eta}_\epsilon$ with nonzero $\epsilon > 0$ plays a technically important role here. If we assumed the white noise from the beginning, there appeared the absolute values of the white noise $|\hat{\xi}^G|$, which is mathematically ill-defined. In addition, the collisional velocities (31) should have diverged for the white noise. To avoid these delicate issues, we proceed with the calculation for nonzero $\epsilon > 0$ and finally take the white-noise $\epsilon \downarrow 0$.
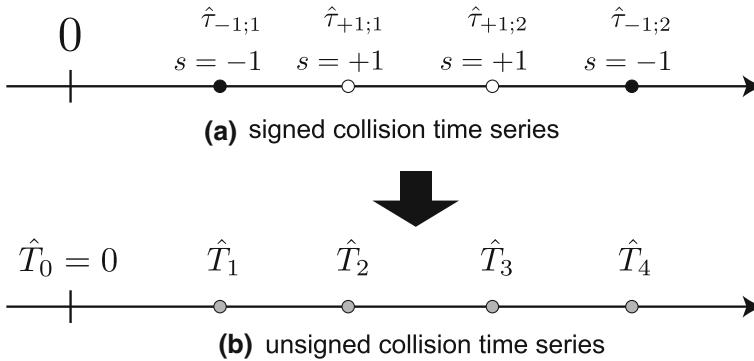
(a) signed collision time series

(b) unsigned collision time series

Fig. 5 The unsigned timeseries $\{\hat{T}_k\}_k$ is introduced as the rearrangement of the signed timeseries $\{\hat{\tau}_{s;k}\}_{s=\pm 1;k}$

This formula implies that only short-time information on the path $\{\hat{r}(t'')\}_{t''\in[t_{\text{ini}}(\epsilon),t)}$ is necessary in evaluating the ensemble average for the white-noise limit.

Rearranged collision timeseries.

Here we reformulate the mathematical notation of the collision times. We have already introduced the signed collision timeseries $\{\hat{\tau}_{s;i}\}_i$ for both $s = \pm 1$. To define the unsigned collision timeseries $\{\hat{T}_i\}_i$, let us rearrange these collision timeseries in the ascending order without considering the sign $s = \pm 1$ (see Fig. 5):

for any $i = 1, 2, ...$, there exists k and s, such that $\hat{T}_i = \hat{\tau}_{s;k}$, satisfying $\hat{T}_i < \hat{T}_{i+1}$.   (43)

We also assume $\hat{T}_0 = 0$. Using this notation, the presence of collision at time t can be written as the condition $t = \hat{T}_i$ for some i, whereas the absence of collision can be written as $t \neq \hat{T}_i$ for any i.

Notably, if there is a collision at time t (i.e., $t = \hat{T}_i$ for some i), the collision term $[f(0) - f(sL/2)]\delta(\hat{r} - sL/2)$ is much more dominant than the diffusive term $\sigma_{\text{cm}}(df/d\hat{r})\hat{\eta}_\epsilon$ in Eq. (37), such that $|\sigma_{\text{cm}}(df/d\hat{r})\hat{\eta}_\epsilon| \ll |[f(0) - f(sL/2)]\delta(\hat{r} - sL/2)|$, due to the δ singularity. Therefore, at time t, we can assume

– the absense of collisions (i.e., $t \neq \hat{T}_i$ for any i) in evaluating $\langle \sigma_{\text{cm}}(df/d\hat{r})\hat{\eta}_\epsilon \rangle$, and
– the presense of a collision (i.e., $t = \hat{T}_i$ for some i) in evaluating $\langle [f(0) - f(sL/2)]\delta(\hat{r} - sL/2) \rangle$.

This means that the dominant term is switched whether there is a collision at the time t or not.

We also estimate the probability of a collision during a infinitesimal-time interval $[t, t + dt)$. The probability $p_{\text{col}}(t, t + dt)$ should be proportional to dt, such that

$$p_{\text{col}}(t, t + dt) = \lambda_{\text{col}}(t)dt + o(dt) \qquad (44)$$

with intensity $\lambda_{\text{col}}(t)$. Indeed, if $p_{\text{col}}(t, t + dt) \propto dt^m$ with $m < 1$, the expected number of collisions $\hat{N}(T)$ during $[0, T)$ is estimated as

$$\langle \hat{N}(T) \rangle \propto \frac{T}{dt}dt^m = Tdt^{m-1}, \qquad (45)$$

which diverges to infinity for $dt \downarrow 0$.

Evaluation of the diffusive term. Based on Novikov's theorem (42) for short-time interval $t - t_{\text{ini}}(\epsilon)$, we evaluate the diffusive term $\langle \sigma_{\text{cm}}(df(\hat{r})/d\hat{r})\hat{\eta}_\epsilon \rangle$ as

$$\lim_{\epsilon \downarrow 0} \left\langle \sigma_{\text{cm}} \frac{df(\hat{r}(t))}{d\hat{r}} \hat{\eta}_\epsilon \right\rangle = \lim_{\epsilon \downarrow 0} \int_{t_{\text{ini}}(\epsilon)}^{t} dt' \frac{e^{-(t-t')/\epsilon}}{2\epsilon} \left\langle \sigma_{\text{cm}} \frac{\delta \hat{r}(t)}{\delta \hat{\eta}_\epsilon(t')} \frac{\partial}{\partial \hat{r}(t)} \left(\frac{df(\hat{r}(t))}{d\hat{r}(t)} \right) \right\rangle. \quad (46)$$

We then evaluate the functional derivative of the path $\delta \hat{r}(t)/\delta \hat{\eta}_\epsilon(t')$, by assuming $t \neq \hat{T}_i$ for any i. Since the integral interval $t - t_{\text{ini}}(\epsilon) = \epsilon^{1/2}$ is to be set infinitesimal, we can assume the absense of collisions during $[t_{\text{ini}}(\epsilon), t)$, such that $\hat{T}_{i-1} < t_{\text{ini}}(\epsilon) < t < \hat{T}_i$.

This assumption can be quantitatively discussed by considering the statistical number of collisions during $[t_{\text{ini}}(\epsilon), t)$. Let us introduce $p_k(t_{\text{ini}}(\epsilon), t)$ and $\langle \ldots \rangle_{k;[t_{\text{ini}}(\epsilon),t)}$ as the probability of k-times collisions during $[t_{\text{ini}}(\epsilon), t)$ and the ensemble average conditional on k-times collisions, respectively. The ensemble average can be written as

$$\left\langle \sigma_{\text{cm}} \frac{\delta \hat{r}(t)}{\delta \hat{\eta}_\epsilon(t')} \frac{\partial}{\partial \hat{r}(t)} \left(\frac{df(\hat{r}(t))}{d\hat{r}(t)} \right) \right\rangle = \sum_{k=0}^{\infty} p_k(t_{\text{ini}}(\epsilon), t) \left\langle \sigma_{\text{cm}} \frac{\delta \hat{r}(t)}{\delta \hat{\eta}_\epsilon(t')} \frac{\partial}{\partial \hat{r}(t)} \left(\frac{df(\hat{r}(t))}{d\hat{r}(t)} \right) \right\rangle_{k;[t_{\text{ini}}(\epsilon),t)}. \quad (47)$$

Since the interval is proportional short $t - t_{\text{ini}}(\epsilon) = \epsilon^{1/2}$, the probability of k-times collisions during $[t_{\text{ini}}(\epsilon), t)$ is estimated to be $[\lambda_{\text{col}}(t)(t - t_{\text{ini}}(\epsilon))]^k \propto \epsilon^{k/2}$. For $\epsilon \downarrow 0$, it is sufficient to consider the case $k = 0$ as the leading-order contribution.

Therefore, by assuming the absense of collisions (i.e., $k = 0$), the path is given by

$$\hat{r}(t) = \hat{r}(t_{\text{ini}}(\epsilon)) + \int_{t_{\text{ini}}(\epsilon)}^{t} dt' \sigma_{\text{cm}} \hat{\eta}_\epsilon(t'), \quad (48)$$

as the formal solution of the SDE (29). This implies the functional-derivative relationship

$$\frac{\delta \hat{r}(t)}{\delta \hat{\eta}_\epsilon(t')} = \sigma_{\text{cm}}, \quad (49)$$

leading to

$$\lim_{\epsilon \downarrow 0} \left\langle \sigma_{\text{cm}} \frac{df(\hat{r}(t))}{d\hat{r}} \hat{\eta}_\epsilon(t) \right\rangle = \lim_{\epsilon \downarrow 0} \int_{t_{\text{ini}}(\epsilon)}^{t} dt' \frac{e^{-(t-t')/\epsilon}}{2\epsilon} \left\langle \sigma_{\text{cm}}^2 \frac{\partial}{\partial \hat{r}} \left(\frac{df(\hat{r})}{d\hat{r}} \right) \right\rangle = \frac{\sigma_{\text{cm}}^2}{2} \left\langle \frac{d^2 f(\hat{r})}{d\hat{r}^2} \right\rangle. \quad (50)$$

Evaluation of the collision term. We next evaluate the collisional term $\langle \hat{\eta}_\epsilon(t)\delta(\hat{r}(t) - sL/2) \rangle$, by assuming $t = \hat{T}_i$ for some i. Using Novikov's theorem (42) for short-time interval, we obtain

$$\lim_{\epsilon \downarrow 0} \langle \hat{\eta}_\epsilon(t)\delta(\hat{r}(t) - sL/2) \rangle = \lim_{\epsilon \downarrow 0} \int_{t_{\text{ini}}(\epsilon)}^{t} dt' \frac{e^{-(t-t')/\epsilon}}{2\epsilon} \left\langle \frac{\delta \hat{r}(t)}{\delta \hat{\eta}_\epsilon(t')} \frac{\partial}{\partial \hat{r}(t)} \delta(\hat{r}(t) - sL/2) \right\rangle. \quad (51)$$

Since the integral interval $t - t_{\text{ini}}(\epsilon) = \epsilon^{1/2}$ is to be set infinitesimal, we can assume the absense of collisions during $[t_{\text{ini}}(\epsilon), t)$, such that $\hat{T}_{i-1} < t_{\text{ini}}(\epsilon) < t = \hat{T}_i$, similarly to the diffusive term. Therefore, the formal solution of the SDE (29) and its functional derivative are given by

$$\hat{r}(t) = \hat{r}(t_{\text{ini}}(\epsilon)) + \int_{t_{\text{ini}}(\epsilon)}^{t} dt' \sigma_{\text{cm}} \hat{\eta}_\epsilon(t') \implies \frac{\delta \hat{r}(t)}{\delta \hat{\eta}_\epsilon(t')} = \sigma_{\text{cm}}. \quad (52)$$

We then obtain

$$\lim_{\epsilon \downarrow 0} \langle \hat{\eta}_\epsilon(t)\delta(\hat{r}(t) - sL/2) \rangle = \lim_{\epsilon \downarrow 0} \int_{t_{\text{ini}}(\epsilon)}^{t} dt' \frac{e^{-(t-t')/\epsilon}}{2\epsilon} \left\langle \sigma_{\text{cm}} \frac{\partial}{\partial \hat{r}(t)} \delta(\hat{r}(t) - sL/2) \right\rangle$$
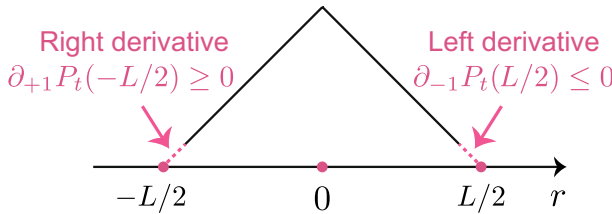
Fig. 6 Technically, the derivatives $\partial P_t(L/2)/\partial r$ and $\partial P_t(-L/2)/\partial r$ should be regarded as left and right derivatives, respectively. Indeed, considering the obvious character $P_t(r) = 0$ for $r > L/2$, $\partial_{+1} P_t(L/2) = 0$ and $\partial_{-1} P_t(-L/2) = 0$. In addition, the signs of the derivatives are given by $\partial_{-1} P_t(L/2) \leq 0$ and $\partial_{+1} P_t(-L/2) \geq 0$

$$
= \frac{\sigma_{\text{cm}}}{2} \left\langle \frac{\partial}{\partial \hat{r}(t)} \delta(\hat{r}(t) - sL/2) \right\rangle
$$

$$
= -\frac{\sigma_{\text{cm}}}{2} \frac{\partial}{\partial r} P_t(sL/2). \tag{53}
$$

Obtaining the ML equation. In summary, we obtain

$$
\int_{-\infty}^{\infty} dr' \frac{\partial}{\partial t} P_t(r') f(r') = \int_{-\infty}^{\infty} dr' \left\{ P_t(r') \frac{\sigma_{\text{cm}}^2}{2} \frac{d^2}{dr'^2} f(r') \right\} - \frac{\sigma_{\text{cm}}^2}{2} \sum_{s=\pm 1} s \left[f(0) - f(sL/2) \right] \frac{\partial}{\partial r} P_t(sL/2). \tag{54}
$$

By substituting $f(r') = \delta(r' - r)$ and performing the partial integration, we obtain

$$
\frac{\partial}{\partial t} P_t(r) = \frac{\sigma_{\text{cm}}^2}{2} \frac{\partial^2}{\partial r^2} P_t(r) - \frac{\sigma_{\text{cm}}^2}{2} \sum_{s=\pm 1} s \left[\delta(r) - \delta(r - sL/2) \right] \frac{\partial}{\partial r} P_t(sL/2)
$$

$$
= \frac{\sigma_{\text{cm}}^2}{2} \frac{\partial^2}{\partial r^2} P_t(r) + \sum_{s=\pm 1} \left[J_{t;s}(r + sL/2) - J_{t;s}(r) \right] \tag{55a}
$$

with the probability current due to collisions

$$
J_{t;s}(r) := -s \frac{\sigma_{\text{cm}}^2}{2} \frac{\partial P_t(sL/2)}{\partial r} \delta(r - sL/2). \tag{55b}
$$

While this is a valid time-evolution equation of the PDF $P_t(r)$, we next rewrite Eq. (55a) to a more intuitve form by examining several technical issues.

4.1.4 Technical Issue on the Left and Right Derivatives

Technically, the derivative $\partial P_t(sL/2)/\partial r$ in the reduced ML equation (55a) should be interpreted as the left (right) derivatives for positive (negative) s (see Fig. 6), such that

$$
\frac{\partial P_t(sL/2)}{\partial r} \rightarrow \partial_{-s} P_t(sL/2), \quad \partial_s f(r) := \lim_{h \downarrow 0} \frac{f(r + sh) - f(r)}{sh}, \tag{56}
$$

because the probability must be exactly zero beyond the boundary

$$
P_t(r) = 0 \quad (\text{for } |r| \geq L/2). \tag{57}
$$

In other words, the particle must come from left (right) when it collides against the right (left) boundary $r = L/2$ ($r = -L/2$), which is reflected for the selection of the derivative direction. This technical issue is more evident in another derivation based on a lattice model in the continuous limit (see Appendix F).
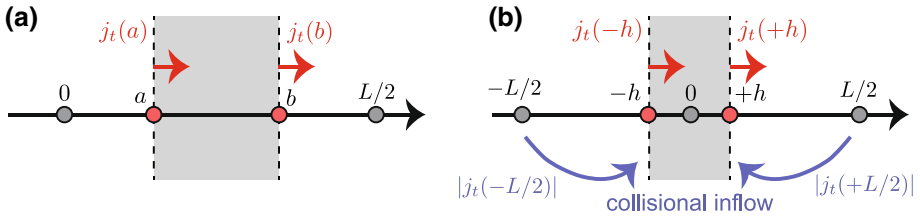
Fig. 7 Interpretation of the ML equation (61b) based on the probability current $j_t(r)$ defined by Eq. (61a). **a** $j_t(r)$ represents the probability current because the probability conservation relation is written as $(d/dt) \int_a^b dr\, P_t(r) = -j_t(b) + j_t(a)$ for $a, b \in (0, L/2)$. **b** The δ contributions from the term $J_{t;s}(r)$ represents the collisional inflow at $r = \pm L/2$ because the probability conservation relation is written as $(d/dt) \int_{-h}^h dr\, P_t(r) = -j_t(h) + j_t(-h) + |j_t(-L/2)| + |j_t(L/2)|$. These schematic illustrates the intuitive meaning of the δ contributions in $J_{t;s}(r)$

4.1.5 Sign of Derivatives

We further examine the sign of the derivatives in rewriting Eq. (55a). Since $P_t(r) = 0$ for $|r| \geq L/2$, we obtain the signs of the left and right derivative (see Fig. 6) as

$$\partial_{-1} P_t(+L/2) \leq 0, \quad \partial_{+1} P_t(-L/2) \geq 0 \quad \Longleftrightarrow \quad s\partial_{-s} P_t(sL/2) = -|\partial_{-s} P_t(sL/2)|. \quad (58)$$

Indeed, since $P_t(r) = 0$ for $|r| \geq L/2$ and $P_t(r) \geq 0$ for $|r| < L/2$, we can show

$$\partial_{-1} P_t(L/2) = \lim_{h \downarrow 0} \frac{P_t(L/2 - h) - P_t(L/2)}{-h} = -\lim_{h \downarrow 0} \frac{P_t(L/2 - h)}{h} \leq 0, \qquad (59a)$$

$$\partial_{+1} P_t(L/2) = \lim_{h \downarrow 0} \frac{P_t(L/2 + h) - P_t(L/2)}{h} = +\lim_{h \downarrow 0} \frac{P_t(L/2 + h)}{h} \geq 0. \qquad (59b)$$

This means that the probability current $J_{t;s}(r)$ defined by Eq. (55b) can be rewritten as an explicitly positive form

$$J_{t;s}(r) := \frac{\sigma_{cm}^2}{2} |\partial_{-s} P_t(sL/2)| \delta(r - sL/2) \geq 0. \qquad (60)$$

This is equivalent to the reduced ML equation (22), which is a more intuitive form than Eq. (55a) because the direction of the probability current is clear as discussed in Sect. 4.2.

4.2 Intuitive Interpretation of the Reduced Master-Liouville Equation (22)

Here we provide an intuitive interpretation of the master-Louville equation (22) from the viewpoint of the probability inflow and outflow relevant to the probability conservation.

By introducing the *probability current* [6] defined by

$$j_t(r) := -\frac{\sigma_{\text{cm}}^2}{2} \frac{\partial P_t(r)}{\partial r}, \tag{61a}$$

the reduced ML equation (22) can be rewritten as

$$\frac{\partial P_t(r)}{\partial t} = -\frac{\partial}{\partial r} j_t(r) + \sum_{s=\pm 1} [J_{t;s}(r+sL/2) - J_{t;s}(r)], \quad J_{t;s}(r) := |j_t(r)|\delta(r-sL/2) \geq 0. \tag{61b}$$

Here we have abbreviated the technical minor symbol on the left and right derivatives. The term $j_t(r)$ is called the probability current because the probability-conservation relation

$$\frac{d}{dt} \int_a^b dr\, P_t(r) = \int_a^b \left[-\frac{\partial}{\partial r} j_t(r)\right] = -\underbrace{j_t(b)}_{\text{outflow}} + \underbrace{j_t(a)}_{\text{inflow}} \tag{62}$$

implies that the total probability within the interval (a, b) is determined by the balance of the probability outflow $j_t(b)$ and inflow $j_t(a)$ (see Fig. 7a for a schematic), where (a, b) is any interval which does not include the singular points $r = 0$ and $r = \pm L/2$, such that $0 < a < b < L/2$ or $-L/2 < a < b < 0$.

Considering the meaning of the probability current $j_t(r)$, the δ-type contribution can be interpreted as follows: let us consider the probability conservation near $r = 0$ by integrating the ML equation (61b) over $(-h, h)$ as

$$\frac{d}{dt} \int_{-h}^h dr\, P_t(r) = \underbrace{-j_t(h) + j_t(-h)}_{\text{diffusive flow}} + \underbrace{|j_t(+L/2)| + |j_t(-L/2)|}_{\text{collisional inflow}} \tag{63}$$

with infinitesimal positive $h > 0$. This means that the δ-contribution from $J_{t;s}(r)$ is to transfer the collisional probability current at $r = \pm L/2$ to the origin $r = 0$ (see Fig. 7b for a schematic). This picture illustrates the balance of the probability currents described by the ML equation (61b), consistently with the physical dynamics where the particle returns back to the origin after collision.

In addition, let us consider the probability conservation near the wall at $r = L/2$, by integrating the ML equation (61b) over $(L/2 - h, L/2 + h)$ as

$$\frac{d}{dt} \int_{L/2-h}^{L/2+h} dr\, P_t(r) = -j_t(L/2 + h) + j_t(L/2 - h) - |j_t(L/2)| = 0$$
$$+ \underbrace{j_t(L/2 - h)}_{\text{diffusive inflow}} - \underbrace{j_t(L/2)}_{\text{collisional outflow}}, \tag{64}$$

where we have used $j_t(L/2+h) = 0$ and $j_t(L/2) > 0$. We thus find that the total probability within the interval $(L/2 - h, L/2 + h)$ is determined by the balance between the diffusive inflow $j_t(L/2 - h)$ and the collisional outflow $j_t(L/2)$.
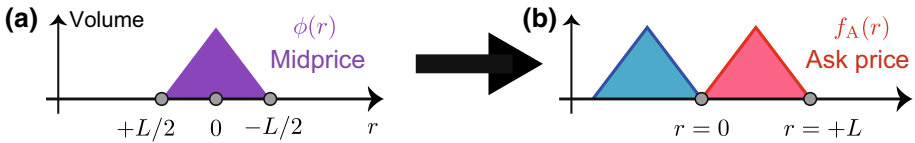
(a) Volume

$\phi(r)$
Midprice

$+L/2 \quad 0 \quad -L/2 \qquad r$

(b)

$f_A(r)$
Ask price

$r = 0 \qquad r = +L$

Fig. 8 **a** The average PDF $\phi(r) := \lim_{t\to\infty} P_t(r)$ in the steady state is given by the tent function (65). **b** $\phi(r)$ is directly related to the average order-book profile (66), where the depth r is measured from the centre of mass $\hat{z}_{cm}$

5 Result 2: Exact Solution to the Reduced Master-Liouville Equation (22)

The ML equation (22) can be solved exactly in the steady state. We finally obtain the tent function as the exact steady solution (see Fig. 8a and Appendix G.1)

$$\phi(r) = \max\left\{0, \frac{L/2 - |r|}{L^2/4}\right\}. \tag{65}$$

5.1 Average Order-Book Profile

The solution (65) is directly related to the normalised average order-book profile $f_A(r)$ (see Fig. 8b) as

$$f_A(r) := \left\langle \frac{1}{2}\sum_{i=1,2}\delta\left(\hat{a}_i - \hat{z}_{cm} - r\right)\right\rangle = \phi(r - L/2) = \max\left\{0, \frac{L/2 - |r - L/2|}{L^2/4}\right\}, \tag{66}$$

where the depth r is measured from the centre of mass $\hat{z}_{cm}$, or equivalently from the market midprice for the special case of $N = 2$.

5.2 Average Transaction Interval

We also discuss the average transaction interval. The average transaction interval can be evaluated by considering the physical meaning of the steady probability current $j_{ss}(r) := \lim_{t\to\infty} j_t(r)$. Indeed, since the absolute value of the steady probability current $|j_{ss}(r)|$ at the walls $r = \pm L/2$ represents the transaction probability per unit time, the average transaction interval $\langle\tau\rangle$ is given by

$$\langle\tau\rangle = \lim_{h\downarrow 0}\frac{1}{|j_{ss}(-L/2 + h)| + |j_{ss}(L/2 - h)|} = \frac{L^2}{4\sigma_{cm}^2} = \frac{L^2}{2\sigma^2}. \tag{67}$$

This is exactly equal to the formula (7) derived in Ref. [15]. We thus rederive the statistics of the transaction interval via the kinetic theory.

6 Result 3: The Full Master-Liouville Equation

We have derived the reduced ML equation to exactly obtain the average order-book profile and the average transaction interval. In this section, we derive the full ML equation for the two-body dealer model (4) via Novikov's theorem in the parallel calculation to Sect. 4.1:

The full PDF $P_t(z_1, z_2) := \langle \delta(z_1 - \hat{z}_1)\delta(z_2 - \hat{z}_2)\rangle$ obeys the *full ML equation* defined by

$$\frac{\partial P_t(z_1, z_2)}{\partial t} = \sum_{i=1,2} \frac{\sigma^2}{2} \frac{\partial^2 P_t(z_1, z_2)}{\partial z_i^2} + \sum_{s=\pm 1} \left[J_{t;s}(z_1 + sL/2, z_2 - sL/2) - J_{t;s}(z_1, z_2) \right]$$
(68a)

with the *probability current* $J_{t;s}(z_1, z_2)$ defined by

$$J_{t;s}(z_1, z_2) := \frac{-s\sigma^2}{2}\delta(z_1 - z_2 - sL)\tilde{\partial}_{12;s} P_t(z_1, z_2)$$
(68b)

$$= \frac{\sigma^2}{2}\delta(z_1 - z_2 - sL)\left|\tilde{\partial}_{12;s}\right| P_t(z_1, z_2) \geq 0.$$

Here we have introduced the left ($s = -1$) and right ($s = +1$) derivatives defined by

$$\partial_{1;s} f(z_1, z_2) := \lim_{h\downarrow 0} \frac{f(z_1 + sh, z_2) - f(z_1, z_2)}{sh}, \quad \partial_{2;s} f(z_1, z_2) := \lim_{h\downarrow 0} \frac{f(z_1, z_2 + sh) - f(z_1, z_2)}{sh},$$
(69)

and

$$\tilde{\partial}_{12;s} f(z_1, z_2) := \partial_{1;-s} f(z_1, z_2) - \partial_{2;s} f(z_1, z_2),$$
(70)

$$\left|\tilde{\partial}_{12;s}\right| f(z_1, z_2) := \left|\partial_{1;-s} f(z_1, z_2)\right| + \left|\partial_{2;s} f(z_1, z_2)\right| \geq 0.$$
(71)

The consistency between the full and reduced ML equations (68) and (22) can be confirmed as shown in Appendix H.

6.1 Reformulation Based on the OU Coloured Noise

First, we reformulate the SDE (4) using the OU coloured noise:

$$\hat{z}_1(t + dt) = \hat{z}_1(t) + \begin{cases} \sigma \hat{\eta}_{1;\epsilon}(t)dt & \text{if } |\hat{z}_1(t) - \hat{z}_2(t)| < L \\ -L/2 & \text{if } \hat{z}_1(t) - \hat{z}_2(t) = +L \\ +L/2 & \text{if } \hat{z}_1(t) - \hat{z}_2(t) = -L \end{cases}$$
(72a)

$$\hat{z}_2(t + dt) = \hat{z}_2(t) + \begin{cases} \sigma \hat{\eta}_{2;\epsilon}(t)dt & \text{if } |\hat{z}_1(t) - \hat{z}_2(t)| < L \\ +L/2 & \text{if } \hat{z}_1(t) - \hat{z}_2(t) = +L , \\ -L/2 & \text{if } \hat{z}_1(t) - \hat{z}_2(t) = -L \end{cases}$$
(72b)

where the OU coloured noises $\hat{\eta}_{1;\epsilon}$ and $\hat{\eta}_{2;\epsilon}$ are defined by

$$\frac{d\hat{\eta}_{1;\epsilon}}{dt} = -\frac{1}{\epsilon}\hat{\eta}_{1;\epsilon} + \hat{\xi}_1^G, \quad \frac{d\hat{\eta}_{2;\epsilon}}{dt} = -\frac{1}{\epsilon}\hat{\eta}_{2;\epsilon} + \hat{\xi}_2^G,$$
(72c)

where $\hat{\xi}_1^G$ and $\hat{\xi}_2^G$ are the standard independent white Gaussian noises (i.e., $\langle \hat{\xi}_i^G \rangle = 0$ and $\langle \hat{\xi}_i^G(t_k)\hat{\xi}_j^G(t_l)\rangle = \delta_{i,j}\delta(t_k - t_l)$). We finally take the white-noise limit $\epsilon \downarrow 0$ to keep the consistency with the original model (4).

Equations (72) can be rewritten by the δ functions as follows: let us introduce the transaction time $\hat{\tau}_{s;i}$ as the ith transaction time with sign $s = \pm 1$ satisfying

$$\hat{z}_1(\hat{\tau}_{s;i}) - \hat{z}_2(\hat{\tau}_{s;i}) = sL, \quad \hat{\tau}_{s;i} < \hat{\tau}_{s;i+1}, \quad s \in \{-1, +1\}.$$
(73a)

Using the δ functions, we can rewrite Eq. (72) as

$$\frac{d\hat{z}_1}{dt} = \sigma\hat{\eta}_{1;\epsilon}(t)dt - \sum_{s=\pm 1}\sum_i \frac{sL}{2}\delta(t-\hat{\tau}_{s;i}), \tag{73b}$$

$$\frac{d\hat{z}_2}{dt} = \sigma\hat{\eta}_{2;\epsilon}(t)dt + \sum_{s=\pm 1}\sum_i \frac{sL}{2}\delta(t-\hat{\tau}_{s;i}). \tag{73c}$$

We note that, for the collision $\hat{z}_1(\hat{\tau}_{s;i}) - \hat{z}_2(\hat{\tau}_{s;i}) = sL$, the relative velocity $d(\hat{z}_1 - \hat{z}_2)/dt$ must be positive for $s = 1$ and negative for $s = -1$, respectively. In other words, the following relation holds,

$$s\lim_{h\downarrow 0}\frac{d}{dt}\left\{\hat{z}_1(\hat{\tau}_{s;i}-h) - \hat{z}_2(\hat{\tau}_{s;i}-h)\right\} = s\sigma\lim_{h\downarrow 0}\left\{\hat{\eta}_{1;\epsilon}(\hat{\tau}_{s;i}-h) - \hat{\eta}_{2;\epsilon}(\hat{\tau}_{s;i}-h)\right\} > 0. \tag{74}$$

6.2 Dynamics of an Arbitrary Function $f(\hat{z}_1, \hat{z}_2)$

We next consider the dynamics of an arbitrary function $f(\hat{z}_1, \hat{z}_2)$:

$$\begin{aligned}\frac{df(\hat{z}_1, \hat{z}_2)}{dt} =&\sigma\hat{\eta}_{1;\epsilon}\frac{\partial f(\hat{z}_1, \hat{z}_2)}{\partial\hat{z}_1} + \sigma\hat{\eta}_{2;\epsilon}\frac{\partial f(\hat{z}_1, \hat{z}_2)}{\partial\hat{z}_2}\\ &+ \sum_{s=\pm 1}\sum_i\left[f(\hat{z}_1 - sL/2, \hat{z}_2 + sL/2) - f(\hat{z}_1, \hat{z}_2)\right]\delta(t-\hat{\tau}_{s;i}).\end{aligned} \tag{75}$$

By the way, since $\hat{\tau}_{s;i}$ is the solution of $\hat{z}_1(\hat{\tau}_{s;i}) - \hat{z}_2(\hat{\tau}_{s;i}) = sL$, the δ function $\delta(\hat{z}_1 - \hat{z}_2 - sL)$ can be decomposed as follows:

$$\delta(\hat{z}_1 - \hat{z}_2 - sL) = \sum_i\left|\frac{d}{dt}(\hat{z}_1 - \hat{z}_2 - sL)\right|^{-1}\delta(t-\hat{\tau}_{s;i}) = \sum_i\sigma^{-1}\left|\hat{\eta}_{1;\epsilon} - \hat{\eta}_{2;\epsilon}\right|^{-1}\delta(t-\hat{\tau}_{s;i}). \tag{76}$$

By considering the sign of the relative velocity (equivalently, the direction of collisions) given by Eq. (74), we obtain

$$s\sigma g(\hat{z}_1, \hat{z}_2)(\hat{\eta}_{1;\epsilon} - \hat{\eta}_{2;\epsilon})\delta(\hat{z}_1 - \hat{z}_2 - sL) = \sum_i sg(\hat{z}_1, \hat{z}_2)\delta(t-\hat{\tau}_{s;i}) \tag{77}$$

for an arbitrary function $g(\hat{z}_1, \hat{z}_2)$. Equation (75) then can be rewritten as

$$\begin{aligned}\frac{df(\hat{z}_1, \hat{z}_2)}{dt} =&\sigma\hat{\eta}_{1;\epsilon}\frac{\partial f(\hat{z}_1, \hat{z}_2)}{\partial\hat{z}_1} + \sigma\hat{\eta}_{2;\epsilon}\frac{\partial f(\hat{z}_1, \hat{z}_2)}{\partial\hat{z}_2}\\ &+ \sum_{s=\pm 1}s\sigma(\hat{\eta}_{1;\epsilon} - \hat{\eta}_{2;\epsilon})\left[f(\hat{z}_1 - sL/2, \hat{z}_2 + sL/2) - f(\hat{z}_1, \hat{z}_2)\right]\delta(\hat{z}_1 - \hat{z}_2 - sL).\end{aligned} \tag{78}$$

We then take the ensemble average of both sides as

$$\begin{aligned}\left\langle\frac{df(\hat{z}_1, \hat{z}_2)}{dt}\right\rangle =&\left\langle\sigma\hat{\eta}_{1;\epsilon}\frac{\partial f(\hat{z}_1, \hat{z}_2)}{\partial\hat{z}_1} + \sigma\hat{\eta}_{2;\epsilon}\frac{\partial f(\hat{z}_1, \hat{z}_2)}{\partial\hat{z}_2}\right\rangle\\ &+ \sum_{s=\pm 1}s\sigma\left\langle(\hat{\eta}_{1;\epsilon} - \hat{\eta}_{2;\epsilon})\left[f(\hat{z}_1 - sL/2, \hat{z}_2 + sL/2) - f(\hat{z}_1, \hat{z}_2)\right]\delta(\hat{z}_1 - \hat{z}_2 - sL)\right\rangle.\end{aligned} \tag{79}$$

6.3 Application of Novikov's Theorem

To proceed the calculation further, we have to calculate the correlation terms of the form $\langle \hat{\eta}_{i;\epsilon}(t) g(\hat{z}_1, \hat{z}_2) \rangle$. Such a correlation term can be calculated via Novikov's theorem,

$$
\begin{aligned}
\langle \hat{\eta}_{i;\epsilon}(t) g(\hat{z}_1, \hat{z}_2) \rangle &= \int_0^t dt' \langle \hat{\eta}_{i;\epsilon}(t) \hat{\eta}_{i;\epsilon}(t') \rangle \left\langle \frac{\delta g(\hat{z}_1, \hat{z}_2)}{\delta \hat{\eta}_{i;\epsilon}} \right\rangle \\
&= \int_0^t dt' \frac{e^{-(t-t')/\epsilon}}{2\epsilon} \left\langle \left\{ \frac{\delta \hat{z}_1(t)}{\delta \hat{\eta}_{i;\epsilon}(t')} \frac{\partial}{\partial \hat{z}_1} + \frac{\delta \hat{z}_2(t)}{\delta \hat{\eta}_{i;\epsilon}(t')} \frac{\partial}{\partial \hat{z}_2} \right\} g(\hat{z}_1, \hat{z}_2) \right\rangle, \quad (80)
\end{aligned}
$$

which implies

$$
\lim_{\epsilon \downarrow 0} \langle \hat{\eta}_{i;\epsilon}(t) g(\hat{z}_1, \hat{z}_2) \rangle = \lim_{t' \uparrow t} \frac{1}{2} \left\langle \left\{ \frac{\delta \hat{z}_1(t)}{\delta \hat{\eta}_{i;\epsilon}(t')} \frac{\partial}{\partial \hat{z}_1} + \frac{\delta \hat{z}_2(t)}{\delta \hat{\eta}_{i;\epsilon}(t')} \frac{\partial}{\partial \hat{z}_2} \right\} g(\hat{z}_1, \hat{z}_2) \right\rangle \quad (81)
$$

for the white-noise limit. Since the formal solution of the SDE (73) is given by

$$
\hat{z}_1(t) = \hat{z}_1(t_{\text{ini}}) + \sigma \int_{t_{\text{ini}}}^t dt' \hat{\eta}_{1;\epsilon}(t'), \quad (82)
$$

$$
\hat{z}_2(t) = \hat{z}_2(t_{\text{ini}}) + \sigma \int_{t_{\text{ini}}}^t dt' \hat{\eta}_{2;\epsilon}(t'), \quad (83)
$$

assuming the absense of transactions during $[t_{\text{ini}}, t)$, we obtain

$$
\lim_{t' \uparrow t} \frac{\delta \hat{z}_i(t)}{\delta \hat{\eta}_j(t')} = \sigma. \quad (84)
$$

We thus obtain

$$
\lim_{\epsilon \downarrow 0} \langle \hat{\eta}_{i;\epsilon}(t) g(\hat{z}_1, \hat{z}_2) \rangle = \frac{\sigma}{2} \left\langle \frac{\partial}{\partial \hat{z}_i} g(\hat{z}_1, \hat{z}_2) \right\rangle. \quad (85)
$$

The formula (85) is useful in deriving the full ML equation. Indeed, for the white-noise limit $\epsilon \downarrow 0$, we obtain

$$
\left\langle \frac{df(\hat{z}_1, \hat{z}_2)}{dt} \right\rangle = \int_{-\infty}^\infty dz_1 dz_2 \frac{\partial P_t(z_1, z_2)}{\partial t} f(z_1, z_2), \quad (86)
$$

$$
\begin{aligned}
\left\langle \sigma \hat{\eta}_{i;\epsilon} \frac{\partial f(\hat{z}_1, \hat{z}_2)}{\partial \hat{z}_i} \right\rangle &= \int_{-\infty}^\infty dz_1 dz_2 P_t(z_1, z_2) \frac{\sigma^2}{2} \frac{\partial^2 f(\hat{z}_1, \hat{z}_2)}{\partial z_i^2} \\
&= \int_{-\infty}^\infty dz_1 dz_2 f(\hat{z}_1, \hat{z}_2) \frac{\sigma^2}{2} \frac{\partial^2 P_t(z_1, z_2)}{\partial z_i^2}, \quad (87)
\end{aligned}
$$

$$
\begin{aligned}
&\langle (\hat{\eta}_{1;\epsilon} - \hat{\eta}_{2;\epsilon}) f(\hat{z}_1 - sL/2, \hat{z}_2 + sL/2) \delta(\hat{z}_1 - \hat{z}_2 - sL) \rangle \\
&= \frac{\sigma}{2} \int_{-\infty}^\infty dz_1 dz_2 P_t(z_1, z_2) \left(\frac{\partial}{\partial z_1} - \frac{\partial}{\partial z_2} \right) f(z_1 - sL/2, z_2 + sL/2) \delta(z_1 - z_2 - sL) \\
&= -\frac{\sigma}{2} \int_{-\infty}^\infty dz_1 dz_2 f(z_1 - sL/2, z_2 + sL/2) \delta(z_1 - z_2 - sL) \left(\frac{\partial}{\partial z_1} - \frac{\partial}{\partial z_2} \right) P_t(z_1, z_2), \quad (88)
\end{aligned}
$$

and

$$
\begin{aligned}
&\langle (\hat{\eta}_{1;\epsilon} - \hat{\eta}_{2;\epsilon}) f(\hat{z}_1, \hat{z}_2) \delta(\hat{z}_1 - \hat{z}_2 - sL) \rangle \\
&= \frac{\sigma}{2} \int_{-\infty}^\infty dz_1 dz_2 P_t(z_1, z_2) \left(\frac{\partial}{\partial z_1} - \frac{\partial}{\partial z_2} \right) f(z_1, z_2) \delta(z_1 - z_2 - sL)
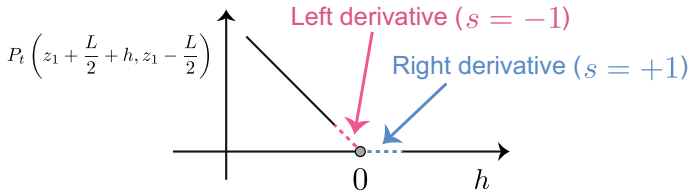\end{aligned}
$$

Fig. 9 Schematic of the left ($\partial_{1;-1}$) and right ($\partial_{1;+1}$) derivatives of $P_t(z_1 + L/2, z_2 - L/2)$ on the condition $z_2 = z_1$. The right derivative is zero, while the left is nonpositive

$$= -\frac{\sigma}{2} \int_{-\infty}^{\infty} dz_1 dz_2 f(z_1, z_2) \delta(z_1 - z_2 - sL) \left(\frac{\partial}{\partial z_1} - \frac{\partial}{\partial z_2} \right) P_t(z_1, z_2), \quad (89)$$

where we have performed the partial integration. By introducing a symbol

$$\tilde{\partial}_{12} := \frac{\partial}{\partial z_1} - \frac{\partial}{\partial z_2} \tag{90}$$

and by substituting $f(\hat{z}_1, \hat{z}_2) = \delta(\hat{z}_1 - z_1)\delta(\hat{z}_2 - z_2)$, we obtain the full ML equation

$$\frac{\partial P_t(z_1, z_2)}{\partial t} = \sum_{i=1,2} \frac{\sigma^2}{2} \frac{\partial^2 P_t(z_1, z_2)}{\partial z_i^2} + \sum_{s=\pm 1} \frac{-s\sigma^2}{2} \delta(z_1 - z_2) \tilde{\partial}_{12} P_t \left(z_1 + \frac{sL}{2}, z_2 - \frac{sL}{2} \right)$$

$$- \sum_{s=\pm 1} \frac{-s\sigma^2}{2} \delta(z_1 - z_2 - sL) \tilde{\partial}_{12} P_t(z_1, z_2). \tag{91}$$

We will further rewrite this ML equation by considering several technical issues.

6.4 Technical Issues on the Left and Right Derivatives

Here we consider the technical issues on the left and right derivatives in Eq. (91). Let us first consider the meaning of the derivatives in $\tilde{\partial}_{12}$ for the contribution of $s = 1$. For the nonnegativity of the probability,

$$P_t(z_1, z_2) \geq 0 \quad \text{for all } z_1, z_2. \tag{92a}$$

At the same time, the transaction rule imposes an obvious restriction,

$$P_t(z_1, z_2) = 0 \quad \text{for } z_1 - z_2 \geq L. \tag{92b}$$

As illustrated in Fig. 9, this implies that

$$\partial_{1;+1} P_t \left(z_1 + \frac{L}{2}, z_1 - \frac{L}{2} \right) := \lim_{h \downarrow 0} \frac{P_t \left(z_1 + \frac{L}{2} + h, z_1 - \frac{L}{2} \right) - P_t \left(z_1 + \frac{L}{2}, z_1 - \frac{L}{2} \right)}{h}$$

$$= \lim_{h \downarrow 0} \frac{0 - 0}{h} = 0, \tag{93a}$$

$$\partial_{2;-1} P_t \left(z_1 + \frac{L}{2}, z_1 - \frac{L}{2} \right) := \lim_{h \downarrow 0} \frac{P_t \left(z_1 + \frac{L}{2}, z_1 - \frac{L}{2} - h \right) - P_t \left(z_1 + \frac{L}{2}, z_1 - \frac{L}{2} \right)}{-h}$$

$$= \lim_{h \downarrow 0} \frac{0 - 0}{-h} = 0, \tag{93b}$$

where we have introduced the left ($s = -1$) and right ($s = +1$) derivatives as

$$\partial_{1;s} f(z_1, z_2) := \lim_{h \downarrow 0} \frac{f(z_1 + sh, z_2) - f(z_1, z_2)}{sh}, \quad \partial_{2;s} f(z_1, z_2) := \lim_{h \downarrow 0} \frac{f(z_1, z_2 + sh) - f(z_1, z_2)}{sh}. \tag{94}$$

We can also show

$$\partial_{1;+1} P_t(z_1, z_1 - L) = \partial_{2;-1} P_t(z_1, z_1 - L) = 0. \tag{95}$$

Considering the relations (93) and (95), the derivatives $\tilde{\partial}_{12}$ for $s = 1$ should be understood as

$$\tilde{\partial}_{12} \to \tilde{\partial}_{12;+1} := \partial_{1;-1} - \partial_{2;+1}. \tag{96}$$

Similarly, by considering the apparent restriction imposed by the transaction rule

$$P_t(z_1, z_2) = 0 \quad \text{for } z_1 - z_2 \le -L, \tag{97}$$

the derivative $\tilde{\partial}_{12}$ for $s = -1$ should be understood as

$$\tilde{\partial}_{12} \to \tilde{\partial}_{12;-1} := \partial_{1;+1} - \partial_{2;-1}. \tag{98}$$

In summary, the ML equation (91) should be technically interpreted as

$$\frac{\partial P_t(z_1, z_2)}{\partial t} = \sum_{i=1,2} \frac{\sigma^2}{2} \frac{\partial^2 P_t(z_1, z_2)}{\partial z_i^2} + \sum_{s=\pm 1} \frac{-s\sigma^2}{2} \delta(z_1 - z_2) \tilde{\partial}_{12;s} P_t\left(z_1 + \frac{sL}{2}, z_2 - \frac{sL}{2}\right)$$

$$- \sum_{s=\pm 1} \frac{-s\sigma^2}{2} \delta(z_1 - z_2 - sL) \tilde{\partial}_{12;s} P_t(z_1, z_2) \tag{99}$$

in terms of the left and right derivatives.

6.5 Sign of Derivatives

Next, we consider the sign of the derivatives. As illustrated in Fig. 9, the relation (92) implies

$$\partial_{1;-1} P_t\left(z_1 + \frac{L}{2}, z_1 - \frac{L}{2}\right) := \lim_{h \downarrow 0} \frac{P_t\left(z_1 + \frac{L}{2} - h, z_1 - \frac{L}{2}\right) - P_t\left(z_1 + \frac{L}{2}, z_1 - \frac{L}{2}\right)}{-h}$$

$$= \lim_{h \downarrow 0} \frac{P_t\left(z_1 + \frac{L}{2} - h, z_1 - \frac{L}{2}\right) - 0}{-h} \le 0, \tag{100a}$$

$$\partial_{2;+1} P_t\left(z_1 + \frac{L}{2}, z_1 - \frac{L}{2}\right) := \lim_{h \downarrow 0} \frac{P_t\left(z_1 + \frac{L}{2}, z_1 - \frac{L}{2} + h\right) - P_t\left(z_1 + \frac{L}{2}, z_1 - \frac{L}{2}\right)}{h}$$

$$= \lim_{h \downarrow 0} \frac{P_t\left(z_1 + \frac{L}{2}, z_1 - \frac{L}{2} + h\right) - 0}{h} \ge 0. \tag{100b}$$

We can also show

$$\partial_{1;-1} P_t(z_1, z_1 - L) \le 0, \quad \partial_{2;+1} P_t(z_1, z_1 - L) \ge 0. \tag{101}$$

This implies that

$$\frac{-s\sigma^2}{2} \delta(z_1 - z_2) \tilde{\partial}_{12;s} P_t\left(z_1 + \frac{sL}{2}, z_2 - \frac{sL}{2}\right) \ge 0, \tag{102a}$$
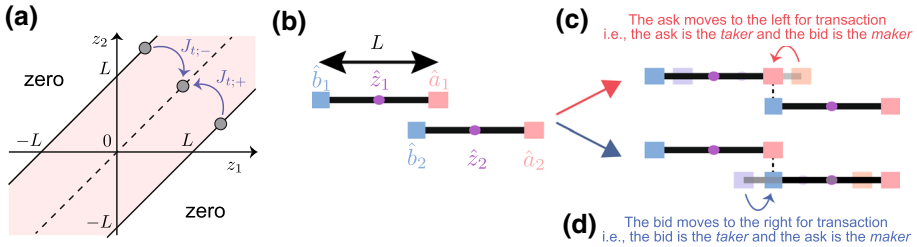
Fig. 10 a The transaction rule imposes the condition $P_t(z_1, z_2) = 0$ for $|z_1 - z_2| \geq L$ and the transaction occurs on the line $|z_1 - z_2| = L$. In the light-red regime $|z_1 - z_2| < L$, the PDF can be non-zero such that $P_t(z_1, z_2) > 0$. The full ML Eq. (68) implies that the probability current $J_{t;s}$ on the line $z_1 - z_2 = sL$ is transferred to the line $z_1 = z_2$ for both $s = \pm 1$. **b** Let us consider the time just before a transaction. There are two scenarios: **c** The ask price a_1 moves to the left, leading to a transaction. This type of transactions implies that the ask is the taker, and the bid is the maker. **d** The bid price b_2 moves to the right, leading to a transaction. This type of transactions implies that the bid is the taker and the ask is the maker

$$\frac{-s\sigma^2}{2}\delta(z_1 - z_2 - sL)\tilde{\partial}_{12;s}P_t(z_1, z_2) \geq 0 \tag{102b}$$

for $s = +1$. Similarly, we can show the inequality (102) even for $s = -1$. By considering the inequality (102), it is useful to introduce the nonnegative probability current:

$$J_{t;s}(z_1, z_2) := \frac{-s\sigma^2}{2}\delta(z_1 - z_2 - sL)\tilde{\partial}_{12;s}P_t(z_1, z_2) \geq 0. \tag{103}$$

Using the nonnegative probability current $J_{t;s}(z_1, z_2)$, we can rewrite the ML equation (99) as

$$\frac{\partial P_t(z_1, z_2)}{\partial t} = \sum_{i=1,2}\frac{\sigma^2}{2}\frac{\partial^2 P_t(z_1, z_2)}{\partial z_i^2} + \sum_{s=\pm 1}\left[J_{t;s}(z_1 + sL/2, z_2 - sL/2) - J_{t;s}(z_1, z_2)\right]. \tag{104}$$

Considering the nonnegativity of the probability current $J_{t;s}(z_1, z_2) \geq 0$, it would be useful to introduce a notation where the nonnegativity is apparent. We thus introduce the symbol $|\tilde{\partial}_{12;s}|$ defined by

$$\left|\tilde{\partial}_{12;s}\right|f(z_1, z_2) := \left|\partial_{1;-s}f(z_1, z_2)\right| + \left|\partial_{2;s}f(z_1, z_2)\right| \geq 0, \tag{105}$$

whose nonnegativity is apparent by its definition. We can thus rewrite the probability current $J_{t;s}(z_1, z_2)$ as

$$J_{t;s}(z_1, z_2) := \frac{\sigma^2}{2}\delta(z_1 - z_2 - sL)\left|\tilde{\partial}_{12;s}\right|P_t(z_1, z_2) \geq 0. \tag{106}$$

We thus obtain the full ML equation (68).

Notably, the selection of the left and right derivatives in the ML equations were not apparent in our previous publications [16, 17]. In this sense, the full ML equation (68) that we have derived in this report is the complete form in terms of the mathematical interpretation.

6.6 Intuitive Interpretation of the Full Master-Liouville Equation (68)

Here we provide an intuitive interpretation of the full ML Eq. (68) from the viewpoint of the probability current. In this subsection, we abbreviate the technical symbol of the left and right derivatives for simplicity. The transaction rule apparently imposes the rule

$$P_t(z_1, z_2) = 0 \quad \text{for } |z_1 - z_2| \geq L \tag{107}$$

and the transaction occurs on the lines $|z_1 - z_2| = L$. The full ML equation (68) means that the probability current $J_{t;s}$ on the line $z_1 - z_2 = sL$ is transferred to the line $z_1 = z_2$ for $s = \{+1, -1\}$, due to the transaction rule (see Fig. 10a).

In addition, we can intuitive decompose the probability current as

$$J_{t;s}(z_1, z_2) = \delta(z_1 - z_2 - sL) \sum_{i=1,2} \left| j_{t;s}^{(i)}(z_1, z_2) \right|, \quad j_{t;s}^{(i)}(z_1, z_2) := -\frac{\sigma^2}{2} \frac{\partial}{\partial z_i} P_t(z_1, z_2). \tag{108}$$

The decomposed probability current $j_{t;s}^{(i)}$ represents the transaction where the ith trader is the *taker* and the other is the *maker*. In addition, similarly to Eqs. (61), we can rewrite the full ML equation as

$$\frac{\partial P_t(z_1, z_2)}{\partial t} = -\sum_{i=1,2} \frac{\partial}{\partial z_i} j_{t;s}^{(i)}(z_1, z_2) + \sum_{s=\pm 1} \left[J_{t;s}(z_1 + sL/2, z_2 - sL/2) - J_{t;s}(z_1, z_2) \right]. \tag{109}$$

Apparently, this formula is a natural extension of the probability-current representation (61) of the reduced ML equation.

Let us explain the necessary background knowledge on takers and makers in financial markets. In the double-auction systems, the trader leading to the decision via the market orders is called a *taker*, whereas the trader waiting for transactions is called *maker*. This distinction is practically vital because the takers and makers regarded as liquidity consumers and providers, respectively. Many market regulators offer a financial incentive to makers because sufficient liquidity provision will stabilise the market.

From the viewpoints of the taker and maker, any transaction can be classified whether the first trader is a taker or maker. For example, let us consider the timing just before a transaction $\hat{a}_1 = \hat{b}_2$ (see Fig. 10b). There are two classifications for this transaction: the first case is that the ask price $\hat{a}_1$ moves to the left and then leads to the transaction (Fig. 10c). In this case, the first trader is the taker and the second trader is the maker. Another case is that the bid price moves to the right and then leads to the transaction. In this case, the second trader is the taker, and the first trader is the maker (Fig. 10d). Considering that the probability current $j_{t;s}^{(i)}$ originates from the diffusion of the ith trader, $j_{t;s}^{(i)}$ represents the contribution where the ith trader leads a transaction as the taker.
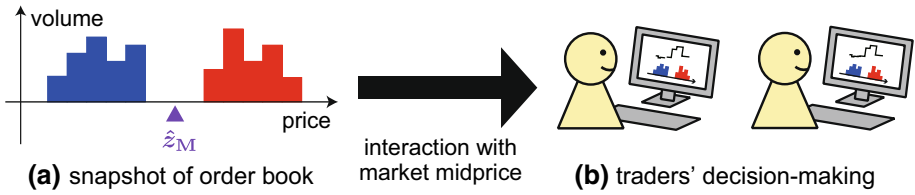
Fig. 11 Traders make their decisions on the basis of the current order book. One of the key feature variables is the market midprice $\hat{z}_M$. As an advanced modelling, interaction between traders and market midprice is incorporated in Sect. 7

7 Advanced Modelling: Interaction Between Traders via the Market Midprice

Here we introduce a generalised dealer model by incorporating interaction between traders and then exactly solve the model straightforwardly on the basis of the kinetic approach.

7.1 Dealer Model with Interaction via the Market Midprice

We have shown the exact solution to the simple two-body dealer model by the kinetic theory. At the same time, it is realistic to introduce interaction between the traders via the order book. For example, it is a reasonable assumption that traders avoid immediate transactions by submitting orders far from the market midprice $\hat{z}_M = \hat{z}_{cm} = (z_1 + z_2)/2$. In the absence of transactions, we thus consider the following generalised dealer model (see Fig. 11):

$$\frac{d\hat{z}_1}{dt} = -\frac{d}{d\hat{r}_1} U\left(\hat{r}_1\right) + \sigma \hat{\xi}_1^G \tag{110a}$$

$$\frac{d\hat{z}_2}{dt} = -\frac{d}{d\hat{r}_2} U\left(\hat{r}_2\right) + \sigma \hat{\xi}_2^G \tag{110b}$$

with the interaction via the market midprice U, the independent white Gaussian noises $\hat{\xi}_i^G$, and relative prices $\hat{r}_i = \hat{z}_i - \hat{z}_M$ from the midprice for $i = 1, 2$. We assume that the potential is a symmetric function with minimum at $r = 0$:

$$U(0) = 0, \quad U(r) = U(-r) \geq U(0) \quad \text{for any } r. \tag{110c}$$

This potential has the effect to keep the distance between the market midprice and the best bid (ask) price. The market spread tends to be kept wide due to this potential and, thus, immediate transactions are unlikely if the potential strength is strong (see Fig. 12).

In the presence of transactions, the jump rule is given by the same rule as the conventional dealer model:

$$|\hat{z}_1(t) - \hat{z}_2(t)| = L \implies \hat{z}_1(t + dt) = \hat{z}_2(t + dt) = \hat{z}_1(t) - \frac{L}{2} \operatorname{sgn}(\hat{z}_1(t) - \hat{z}_2(t)). \tag{110d}$$

We then describe this generalised dealer model in terms of the relative coordinate to the centre of mass $\hat{r} := \hat{z}_1 - \hat{z}_{cm} = (\hat{z}_1 - \hat{z}_2)/2$:
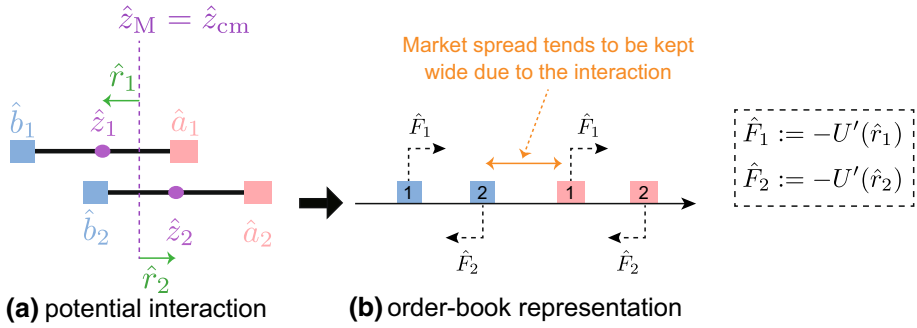
(a) potential interaction   **(b)** order-book representation

Fig. 12 Schematic of the potential interaction via the market midprice $\hat{z}_{\mathrm{M}} = \hat{z}_{\mathrm{cm}}$. **a** The relative coordinate $\hat{r}_i$ is introduced from the market midprice for $i = 1, 2$. We can define the "force" $\hat{F}_i := -U'(\hat{r}_i)$ which has the effect to keep the market spread wide and thus to prevent immediate transactions

$$r(t + dt) = \begin{cases} r(t) + \left(-U'(\hat{r}) + \sigma_{\mathrm{cm}}\hat{\eta}(t)\right) dt & (|r(t)| < L/2) \\ 0 & (|r(t)| = L/2) \end{cases} \tag{111}$$

with $\hat{\eta}(t) := (\hat{\xi}_1^{\mathrm{G}} - \hat{\xi}_2^{\mathrm{G}})/\sqrt{2}$, $U'(r) := dU(r)/dr$, and $\sigma_{\mathrm{cm}} := \sigma/\sqrt{2}$.

7.2 ML Equation in the Presence of $U(r)$

We derive the ML equation corresponding to Eq. (111). As a natural extension of the reduced ML Eq. (61) for $U(r) = 0$, we obtain the following reduced ML equation in the presence of the potential:

$$\frac{\partial P_t(r)}{\partial t} = -\frac{\partial}{\partial r}\left(j_t^{\mathrm{D}}(r) + j_t^{\mathrm{P}}(r)\right) + \sum_{s=\pm 1}\left[J_{t;s}(r + sL/2) - J_{t;s}(r)\right], \tag{112a}$$

where $j_t^{\mathrm{D}}(r)$, $j_t^{\mathrm{P}}(r)$, and $J_{t;s}(r)$ are the probability currents due to diffusion, potential, and jump, respectively, defined by

$$j_t^{\mathrm{D}}(r) := -\frac{\sigma_{\mathrm{cm}}^2}{2}\frac{\partial P_t(r)}{\partial r}, \quad j_t^{\mathrm{P}}(r) := -U'(r)P_t(r), \quad J_{t;s}(r) := |j_t^{\mathrm{D}}(r)|\delta(r - sL/2) \geq 0, \tag{112b}$$

where we have ignored the minor technical issues on the left and right derivatives.

The derivation of Eq. (112) is essentially parallel to that of Eq. (61) (see Appendix I for the detail).

7.3 Exact Steady Solution for General Avoiding Potential

We next show the exact solution for the order-book profile for a symmetric general avoiding potential (see Appendix G.2 for the derivation):

For a symmetric potential

$$U(r) = U(-r), \tag{113a}$$

we obtain the exact solution

$$\phi(r) = \begin{cases} \dfrac{1}{Z}\exp\left[-\dfrac{2U(r)}{\sigma_{\text{cm}}^2}\right]\{\mathcal{G}(L/2) - \mathcal{G}(|r|)\} & (|r| \le L/2) \\ 0 & (|r| > 0) \end{cases} \tag{113b}$$

with

$$\mathcal{G}(r) := \int_0^r dx\, \exp\left[\dfrac{2U(x)}{\sigma_{\text{cm}}^2}\right], \quad Z := \int_{-L/2}^{L/2} dr\, \exp\left[-\dfrac{2U(r)}{\sigma_{\text{cm}}^2}\right]\{\mathcal{G}(L/2) - \mathcal{G}(|r|)\}. \tag{113c}$$

7.4 Exact Steady Solution for Harmonic Avoidng Potential

We next consider the specific case where the avoiding potential is harmonic:

For the harmonic avoiding potential

$$U(r) = \dfrac{u^2 r^2}{2}, \quad u > 0, \tag{114}$$

the exact steady solution is given by

$$\phi(r) = \begin{cases} \dfrac{1}{Z}e^{-\frac{u^2 r^2}{\sigma_{\text{cm}}^2}}\left[\operatorname{erfi}\left(\dfrac{uL}{2\sigma_{\text{cm}}}\right) - \operatorname{erfi}\left(\dfrac{u|r|}{\sigma_{\text{cm}}}\right)\right] & (|r| \le L/2) \\ 0 & (|r| > L/2) \end{cases} \tag{115}$$

with

$$Z = \dfrac{1}{2u\sigma_{\text{cm}}\sqrt{\pi}}\left[2\pi\sigma_{\text{cm}}^2\operatorname{erf}\left(\dfrac{uL}{2\sigma_{\text{cm}}}\right)\operatorname{erfi}\left(\dfrac{uL}{2\sigma_{\text{cm}}}\right) - u^2 L^2 {}_2F_2\left(\genfrac{}{}{0pt}{}{1,1}{3/2,2}\Big| - \dfrac{u^2 L^2}{4\sigma_{\text{cm}}^2}\right)\right]. \tag{116}$$

Here we have introduced the following special functions:

$$\operatorname{erf}(x) \equiv (2/\sqrt{\pi})\int_0^x e^{-t^2}\,dt \tag{117a}$$

$$\operatorname{erfi}(x) \equiv -i\operatorname{erf}(iz) \tag{117b}$$

$${}_2F_2\left(\genfrac{}{}{0pt}{}{a_1,a_2}{b_1,b_2}\Big|z\right) \equiv \sum_{n=0}^{\infty} \dfrac{(a_1)_n (a_2)_n}{(b_1)_n (b_2)_n}\dfrac{z^n}{n!} \tag{117c}$$

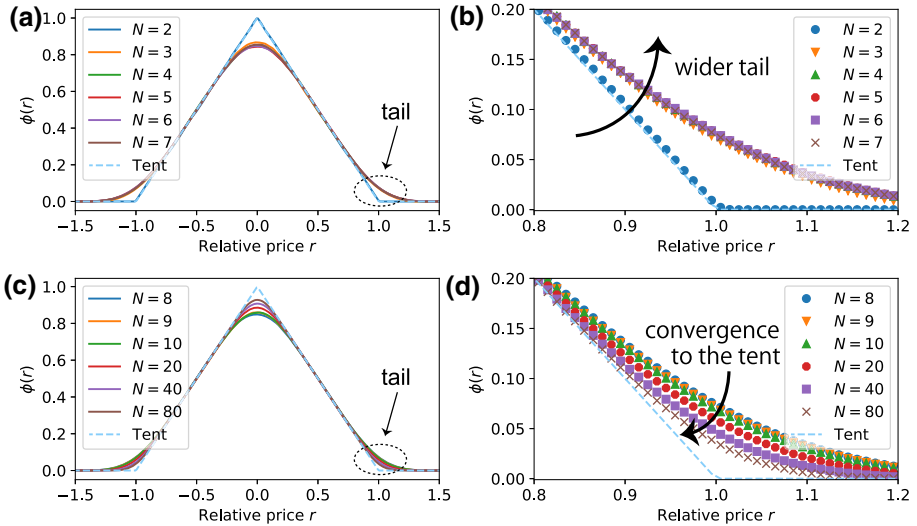with the Pochhammer symbol $(a)_n \equiv a(a+1)(a+2)\dots(a+n-1)$ for $n \ge 1$ and $(a)_0 = 1$.

Fig. 13 Numerical steady PDF $\phi_N(r)$ for the N-body dealer model for various N, showing the non-monotonic convergence to the tent function (118). **a, b** For $N \in [N, N^*]$, the PDF tail becomes wider with $N^* \simeq 7$, at least numerically. We note that the figure b enlarges the tail near $r = L$. **c, d** The numerical tail seems to exhibit the convergence to the tent function (118) for $N \in (N^*, \infty)$

8 Numerical Confirmation and Discussion

8.1 Numerical Confirmation Without Avoiding Potential

8.1.1 Exact Solution for $N = 2$

Here we numerically confirm the validity of the tent-function formula (65) for $\phi(r)$ (see Appendix J for the detailed numerical scheme). We have plotted the numerical PDF for $\phi(r)$ as shown in Fig. 13a and b, where the tent function is precisely consistent with the numerical result for $N = 2$.

8.1.2 Non-monotonic Convergence for $N \to \infty$

Let us discuss the relationship of the two-body exact solution (65) and the numerical solutions for the N-body dealer model. According to Refs. [16, 17], remarkably, the mean-field solution for $N \to \infty$ is also given by the tent function

$$\lim_{N \to \infty} \phi_N(r) = \max \left\{ 0, \frac{L/2 - |r|}{L^2/4} \right\}, \tag{118}$$

where the steady PDF $\phi_N(r) := \langle \delta(\hat{z}_i - \hat{z}_{cm} - r) \rangle$ is defined for the N-body dealer model, by making assumptions (see Appendix J and Refs. [16, 17] for the model assumptions) that

– all the traders share the same value of their buy-sell spread: $\hat{L}_i := \hat{a}_i - \hat{b}_i = L = \text{const.}$
– the dynamics is given by the straightforward generalisation of Eq. (4); i.e., random walks with "collisions" when bid and ask prices coincide with each other.
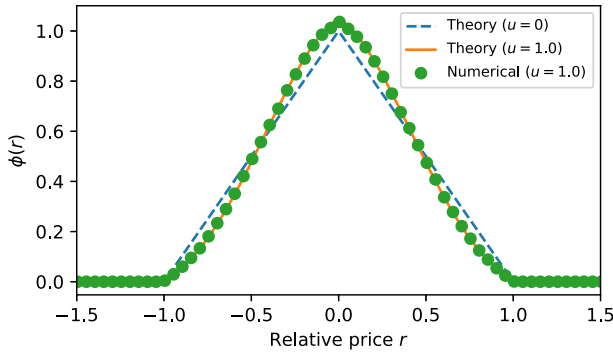
Fig. 14 Numerical steady PDF $\phi(r)$ for the two-body dealer model under the harmonic avoiding potential $U(r) = u^2 r^2/2$ with $u = 1$. The numerical PDF $\phi(r)$ excellently fits the theoretical line (115) for $u \neq 0$ but shows discrepancy with the tent function for $u = 0$

This implies that the exact solution for $N = 2$ is equal to the mean-field solution for $N \to \infty$:

$$\phi_{N=2}(r) = \lim_{N \to \infty} \phi_N(r). \tag{119}$$

On the other hand, the solution $\phi_N(r)$ is different from the tent function (65) for general $N \neq 2, \infty$: $\phi_N(r) :\neq \max\{0, (L/2 - |r|)/L^2/4\}$ for $N \neq 2, \infty$. Indeed, the next-leading-order (NLO) mean-field solution for large $N \gg 1$ is given by

$$\phi_N(r) = \frac{4\varepsilon}{L^2} \left[\mathcal{F}\left(\frac{|r| - L/2}{\varepsilon} \right) - 2\mathcal{F}\left(\frac{|r|}{\varepsilon} \right) \right] \tag{120}$$

with the thickness of the boundary layer ε and the tail function $\mathcal{F}(r)$ defined by

$$\varepsilon := \frac{L}{2\sqrt{N}}, \quad \mathcal{F}(r) := \frac{1}{\sqrt{2\pi}} e^{-r^2/2} - \frac{r}{2} \mathrm{erfc}\left(\frac{r}{\sqrt{2}} \right). \tag{121}$$

These relations (119) and (120) suggest that the convergence behaviour of the steady PDF $\phi_N(r)$ is not monotonic in terms of the tail; the tail becomes wider from $N = 2$ to $N = N^*$ with some fixed value $N^* > 0$ and then it finally converges to the tent function (65).

To confirm this picture, we have performed the numerical simulations of the N-body dealer model for various N as shown in Fig. 13 (see Appendix J for the detailed numerical scheme). Figure 13a and b shows that the tail becomes wider up to $N^* \simeq 7$, numerically. On the other hand, the tail monotonically converges to the tent function (118) for $N > N^*$, as suggested by the numerical Fig. 13c and d. This non-monotonic convergence suggests that one of the approximate criteria to apply the mean-field solution (120) might be to satisfy the condition $N > N^*$ since it is a threshold whereby the solution exhibits qualitatively different behaviours. It might be interesting to investigate the reason behind this non-monotonic convergence numerically and theoretically.

8.2 Numerical Confirmation Under Harmonic Avoiding Potential for $N = 2$

We next confirm our exact solution (115) in the presence of the harmonic avoiding potential for $N = 2$ (see Appendix J for the detailed numerical scheme). The numerical plot nicely agrees with the theoretical line (115) in Fig. 14. In addition, the PDF $\phi(r)$ shrinks around

$r = \pm L/2$ in the presence of $U(r)$ and, thus, immediate transactions are unlikely for large u.

8.3 Discussion to Other Order-Book Models

The dealer model is a simple microscopic model that can replicate various empirical findings, as shown in Refs. [16, 17]. In this report, we have discussed the analytical characters of the *revealed* order book in terms of explicit exact solutions. The revealed order book is defined by the visible order book gathering all the limit orders, in contrast to the latent order book (see the related discussion below). While several order-book models have been proposed in the literature, such as the zero-intelligence order-book model [30, 31], their explicit analytical characters have not been well-documented even in terms of the revealed order book due to their high-dimensional characters.[7] The dealer model is currently exceptional in the sense that the kinetic theory can thoroughly scrutinise its analytical characters for both mean-field limit $N \to \infty$ and two-body case $N = 2$. It might be interesting to extend the kinetic financial framework beyond the dealer model, such as for the zero-intelligence models.

Another research direction would be to address the *latent* order book instead of the revealed order book. While the revealed order book is visible and is mainly contributed by high-frequency market makers, the latent order book is invisible and is considered to be managed by low-frequency actors. In Refs. [33], a latent order book model based on the reaction-diffusion process was proposed to explain the nonlinear market impact after meta-order splitting. This model's book profile is locally linear near the market midprice, which is directly related to the square-market-impact law.

The dealer model was validated on the empirical observation of the high-frequency market makers in Refs. [16, 17]. It mainly focuses on the revealed order book, and studying the latent order book is out of scope in this paper. However, we think it would be possible to extend this model to analyse the latent order book. For example, by introducing high-frequency and low-frequency actors, the revealed order book can be decomposed into the high-frequency market-maker order book and the other (i.e, the latent order book). The book shape of such a generalised dealer model is expected to be also locally linear near the midprice (as implied from the tent-function profile (118) in the mean-field limit $N \to \infty$), consistently with the latent order book model. Such extensions will be very interesting for a deeper understanding of the relationship to other models and are left for future studies.

9 Conclusion

We have exactly scrutinised the stochastic dealer model by focusing on the specific case $N = 2$ from the viewpoint of kinetic theory. We first derive a reduced form of the master-Liouville (ML) equation based on Novikov's theorem for coloured noise. We also examine the physical meaning of the reduced ML equation from the probability current viewpoint, intuitively discerning why the reduced ML equation takes its form as it is. The reduced ML equation is exactly solved to obtain the average-order book profile and the transaction

---

[7] One remarkable formula was derived for the revealed book profile by a phenomenological theory based on the method of images under the diffusive approximation [11, 32]. While this theory includes a fitting parameter that cannot be fixed within the theory, it predicts a scaling behaviour of the book profile qualitatively consistent with empirical results (such as hump shapes). To the best of our knowledge, this was the only analytical formula for the book shape before our kinetic works.

interval. Remarkably, the average transaction interval coincides with that in the previous literature [15], showing the consistency between the different approaches. We next derive the full ML equation to examine its physical meaning and consistency with the reduced ML equation. To demonstrate the power of this approach, we generalise the dealer model in terms of the interaction between traders via the order book and again exactly solve the generalised model within the kinetic approach. Finally, we provide the numerical simulations to test our exact solution's validity.

Since the MLEs are derived in this paper, various traditional tools for the master equations will be available for the mathematical analysis of the dealer model. For example, while we have only focused on the steady solution $\phi(r)$, it is possible to consider the time-dependent solutions (since the steady solution corresponds to the ML operator's zero eigenfunction, time-dependent solution corresponds to non-zero eigenfunctions). In addition, it might be interesting to apply the full counting-statistics framework for the MLE to study the complete transaction interval statistics from a different angle.

In this paper, we have attempted to thoroughly investigate the mathematical structure of the kinetic theory for financial Brownian motion by focusing on the simplest case of $N = 2$. We have shown that various theoretical methods finally produce the same results, which guarantees the mathematical soundness of our approach. While our previous long paper [17] has meticulously revealed the mean-field mathematical structure of the N-body dealer model with $N \gg 1$, this report supplements our previous Letter [16] from the viewpoint of the exact solution of the simplest case $N = 2$, by the detailed description of the ML equations in the complete form. Also, the utility of this mathematical formulation is demonstrated by solving an advanced and realistic dealer model. It would be interesting to observe such a potential interaction from microscopic data analysis directly. In addition, we believe that the thick technical review section would help non-expert readers understand our mathematical theory without hurdles.

A Brief Review on the δ Function

We briefly review the δ function. The δ function is formally defined by the following relation:

$$\delta(x) = \begin{cases} 0 & (x \neq 0) \\ \infty & (x = 0) \end{cases}, \quad \int_{-\infty}^{\infty} dx f(x) \delta(x) = f(0) \tag{122}$$

for an arbitrary function $f(x)$.

Using this relation, we can derive the variable-transformation formulas. For example, we can derive

$$\delta(ax) = \frac{1}{|a|}\delta(x) \tag{123}$$

with constant $a \neq 0$. This relation can be derived as follows: for an arbitrary function $f(x)$ and positive $a > 0$, we obtain

$$\int_{-\infty}^{\infty} dx f(x)\delta(ax) = \frac{1}{a}\int_{-\infty}^{\infty} dy f\left(\frac{y}{a}\right)\delta(y) = \frac{1}{a}f(0) \tag{124}$$

by the variable transformation $y := ax$. The same calculation can be performed for $a < 0$ likewise.

The variable-transformation formula (123) can be generalised

$$\delta(g(x)) = \sum_i \frac{1}{|g'(x_i)|}\delta(x - x_i) \tag{125}$$

for an arbitary function $g(x)$, where x_i is the ith zero point of $g(x)$ such that $g(x_i) = 0$ and $x_i < x_{i+1}$ by assuming $g'(x_i) \neq 0$. This can be derived as follows: for x sufficiently near x_i, $g(x)$ can be expanded as

$$g(x) \simeq g(x_i) + g'(x_i)(x - x_i) + o((x - x_i)). \tag{126}$$

This means that

$$\delta(g(x)) \simeq \delta\left(g'(x_i)(x - x_i) + o((x - x_i))\right) \simeq \frac{1}{|g'(x_i)|}\delta(x - x_i) \quad \text{for } x \text{ near } x_i. \tag{127}$$

By considering the contributions for all x near $\{x_i\}_i$, we obtain Eq. (125).

B Derivation of Novikov's Theorem

We derive Novikov's theorem (15a) as follows: let us apply the functional Taylor expansion (see Appendix C),

$$g[\hat{\eta}_\epsilon; t] = g[0; t] + \sum_{n=1}^{\infty} \frac{1}{n!}\int_0^t ds_n \frac{\delta^n g[\eta_\epsilon; t]}{\delta\eta_\epsilon(s_1)\dots\eta_\epsilon(s_n)}\bigg|_{\eta_\epsilon=0} \hat{\eta}_\epsilon(s_1)\dots\hat{\eta}_\epsilon(s_n), \; ds_n := \prod_{k=1}^{n} ds_k. \tag{128}$$

By introducing $s_n := (s_1, \dots, s_n)$, this implies

$$\langle\hat{\eta}_\epsilon(t)g[\hat{\eta}_\epsilon; t]\rangle = \sum_{n=1}^{\infty} \frac{1}{n!}\int_0^t ds_n R^{(n)}(s_n)\langle\hat{\eta}_\epsilon(t)\hat{\eta}_\epsilon(s_1)\dots\hat{\eta}_\epsilon(s_n)\rangle, \tag{129a}$$

$$R^{(n)}(s_n) := \frac{\delta^n g[\eta_\epsilon; t]}{\delta\eta_\epsilon(s_1)\dots\eta_\epsilon(s_n)}\bigg|_{\eta_\epsilon=0}. \tag{129b}$$

Here we use a mathematical fact on the Gaussian random numbers:

$$\langle\hat{\eta}_\epsilon(t)\hat{\eta}_\epsilon(s_1)\dots\hat{\eta}_\epsilon(s_n)\rangle = \sum_{i=1}^{n}\langle\hat{\eta}_\epsilon(t)\hat{\eta}_\epsilon(s_i)\rangle\left\langle\prod_{j=1|j\neq i}^{n}\hat{\eta}_\epsilon(s_j)\right\rangle. \tag{130}$$

By considering a symmetry of $R^{(n)}(\boldsymbol{s}_n)$

$$R^{(n)}(\boldsymbol{s}_n) = R^{(n)}(\boldsymbol{s}'_n), \quad \text{where } \boldsymbol{s}'_n \text{ is any permutation of } \boldsymbol{s}_n, \tag{131}$$

we obtain

$$
\begin{aligned}
\langle \hat{\eta}_\epsilon(t) g[\hat{\eta}_\epsilon; t] \rangle &= \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^t d\boldsymbol{s}_n R^{(n)}(\boldsymbol{s}_n) \sum_{i=1}^{n} \langle \hat{\eta}_\epsilon(t)\hat{\eta}_\epsilon(s_i) \rangle \left\langle \prod_{j=1|j\neq i}^{n} \hat{\eta}_\epsilon(s_j) \right\rangle \\
&= \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{i=1}^{n} \int_0^t ds_i \langle \hat{\eta}_\epsilon(t)\hat{\eta}_\epsilon(s_i) \rangle \int_0^t \left(\prod_{k=1|k\neq i}^{n} ds_k \right) R^{(n)}(\boldsymbol{s}_n) \left\langle \prod_{j=1|j\neq i}^{n} \hat{\eta}_\epsilon(s_j) \right\rangle \\
&= \int_0^t ds_1 \langle \hat{\eta}_\epsilon(t)\hat{\eta}_\epsilon(s_1) \rangle \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int_0^t ds_2 \dots ds_n R^{(n)}(\boldsymbol{s}_n) \langle \hat{\eta}_\epsilon(s_2) \dots \hat{\eta}_\epsilon(s_n) \rangle.
\end{aligned}
\tag{132}
$$

Eq. (128) implies

$$
\begin{aligned}
\frac{\delta g[\hat{\eta}_\epsilon; t]}{\delta \hat{\eta}_\epsilon(s)} &= \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^t d\boldsymbol{s}_n R^{(n)}(\boldsymbol{s}_n) \frac{\delta}{\delta \hat{\eta}_\epsilon(s)} \{\hat{\eta}_\epsilon(s_1) \dots \hat{\eta}_\epsilon(s_n)\} \\
&= \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int_0^t ds_2 \dots ds_n R^{(n)}(s, s_2, \dots, s_n) \hat{\eta}_\epsilon(s_2) \dots \hat{\eta}_\epsilon(s_n),
\end{aligned}
\tag{133}
$$

where we have used the symmetry (131) again. We thus obtain Eq. (15a) by substituing Eq. (133) into Eq. (132).

C Brief Review on the Functional Taylor Expansion

We briefly reivew the functional Taylor expansion in this Appendix. Before the review of the functional Taylor expansion, we first review the Taylor expansion for a n-dimensional vector $\boldsymbol{x} := (x_1, \dots, x_n)$ with a positive integer n. For an arbitrary function $f(\boldsymbol{x})$, the Taylor expansion implies

$$f(\boldsymbol{x}) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{i=1}^{n} x_i \frac{\partial}{\partial y_i} \right)^k f(\boldsymbol{y}) \bigg|_{\boldsymbol{y}=\boldsymbol{0}}. \tag{134}$$

The functional Taylor expansion is a generalisation of the relation (134). In other words, the functional Taylor expansion implies

$$f[x] = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\int_{-\infty}^{\infty} dt x(t) \frac{\delta}{\delta y(t)} \right)^k f[y] \bigg|_{y=0} \tag{135}$$

with an arbitrary functional $f[x] := f[\{x(t)\}_t]$ for a function $\{x(t)\}_t$. Here, $\delta/\delta x(t)$ is the functional derivative.

The functional derivative $\delta/\delta x(t)$ is related to the δ function as

$$\frac{\delta x(t)}{\delta x(t')} = \delta(t - t'), \tag{136}$$

which is a generalisation of the following relation for a finite-dimensional vector $\boldsymbol{x}$:

$$\frac{\partial x_i}{\partial x_j} = \delta_{i,j}. \tag{137}$$

D Derivation of the Fokker–Planck Equation (16) in the White Noise Limit

Let us derive the Fokker-Planck equation (16) corresponding to the SDE (14) in the white-noise limit. The derivative of an arbitary function $f(\hat{x}(t))$ is given by the chain rule[8]:

$$\frac{df(\hat{x})}{dt} = \left(-\alpha(\hat{x}) + \beta(\hat{x})\hat{\eta}_\epsilon\right) \frac{df(\hat{x})}{d\hat{x}}. \tag{138}$$

Next, let us take the ensemble average of both sides. We first obtain two identities:

$$\left\langle \frac{df(\hat{x})}{dt} \right\rangle = \lim_{dt \downarrow 0} \left\langle \frac{f(\hat{x}(t+dt)) - f(\hat{x}(t))}{dt} \right\rangle = \int dx f(x) \lim_{dt \downarrow 0} \frac{P_{t+dt}(x) - P_t(x)}{dt}$$
$$= \int dx f(x) \frac{\partial}{\partial t} P_t(x), \tag{139}$$

and

$$\left\langle \alpha(\hat{x}) \frac{df(\hat{x})}{d\hat{x}} \right\rangle = \int dx \, P_t(x)\alpha(x) \frac{df(x)}{dx} = -\int dx f(x) \frac{\partial}{\partial x} \alpha(x) P_t(x), \tag{140}$$

where we have used the partial integration by assuming the sufficiently-rapid decay of the PDF:

$$\lim_{x \to \pm\infty} P_t(x) = 0, \quad \lim_{x \to \pm\infty} \frac{\partial^n P_t(x)}{\partial x^n} = 0, \tag{141}$$

with any positive integer n. Using Novikov's theorem (15b), we obtain an identity for the white-noise limit $\epsilon \downarrow 0$:

$$\lim_{\epsilon \downarrow 0} \left\langle \hat{\eta}_\epsilon(t)\beta(\hat{x}(t)) \frac{df(\hat{x}(t))}{d\hat{x}} \right\rangle = \lim_{\epsilon \downarrow 0} \int_0^t ds \frac{1}{2\epsilon} e^{-(t-s)/\epsilon} \left\langle \frac{d}{d\hat{x}} \left\{ \beta(\hat{x}) \frac{df(\hat{x})}{d\hat{x}} \right\} \frac{\delta \hat{x}(t)}{\delta \hat{\eta}_\epsilon(s)} \right\rangle$$
$$= \frac{1}{2} \lim_{s \uparrow t} \left\langle \frac{d}{d\hat{x}} \left\{ \beta(\hat{x}) \frac{df(\hat{x})}{d\hat{x}} \right\} \frac{\delta \hat{x}(t)}{\delta \hat{\eta}_\epsilon(s)} \right\rangle. \tag{142}$$

The formal solution of the SDE (14) is given by

$$x(t) = x(t_{\text{ini}}) + \int_{t_{\text{ini}}}^t dt' \left\{ -\alpha(\hat{x}(t')) + \beta(\hat{x}(t'))\hat{\eta}_\epsilon(t') \right\} \tag{143}$$

for any $t_{\text{ini}} \in [0, t]$. Notably, considering the causality, the effect of $\hat{\eta}_\epsilon(s)$ is only related to the later value of $\hat{x}(t')$ with $t' > s$. In other words, we obtain

$$\frac{\delta \hat{x}(t')}{\delta \hat{\eta}_\epsilon(s)} = 0 \quad \text{for } t' < s. \tag{144}$$

Using these relations, we obtain the formal functional derivative

$$\frac{\delta \hat{x}(t)}{\delta \hat{\eta}_\epsilon(s)} = \int_{t_{\text{ini}}}^t dt' \left\{ -\frac{\partial \alpha(\hat{x}(t'))}{\partial \hat{x}} \frac{\delta \hat{x}(t')}{\delta \hat{\eta}_\epsilon(s)} + \frac{\partial \beta(\hat{x}(t'))}{\partial \hat{x}} \frac{\delta \hat{x}(t')}{\delta \hat{\eta}_\epsilon(s)}\hat{\eta}_\epsilon(t') + \beta(\hat{x}(t')) \frac{\delta \hat{\eta}_\epsilon(t')}{\delta \hat{\eta}_\epsilon(s)} \right\}$$
$$= \int_{t_{\text{ini}}}^t dt' \left\{ -\frac{\partial \alpha(\hat{x}(t'))}{\partial \hat{x}} + \frac{\partial \beta(\hat{x}(t'))}{\partial \hat{x}}\hat{\eta}_\epsilon(t') \right\} \frac{\delta \hat{x}(t')}{\delta \hat{\eta}_\epsilon(s)} + \int_{t_{\text{ini}}}^t dt' \left\{ \beta(\hat{x}(t'))\delta(t' - s) \right\}$$
$$= \beta(\hat{x}(s)) + \int_s^t dt' \left\{ -\frac{\partial \alpha(\hat{x}(t'))}{\partial \hat{x}} + \frac{\partial \beta(\hat{x}(t'))}{\partial \hat{x}}\hat{\eta}_\epsilon(t') \right\} \frac{\delta \hat{x}(t')}{\delta \hat{\eta}_\epsilon(s)} \tag{145}$$

---

[8] The ordinary chain rule is available because the OU coloured noise is a continuous and bounded function of time. If we take the white-noise limit first, the ordinary chain rule must be replaced with the Itô formula.

for any $s \in (t_{\text{ini}}, t)$, where we have used $\delta \hat{\eta}_\epsilon(t')/\delta \hat{\eta}_\epsilon(s) = \delta(t' - s)$ and the causality relation (144). We thus obtain

$$\lim_{s \uparrow t} \frac{\delta \hat{x}(t)}{\delta \hat{\eta}_\epsilon(s)} = \beta(\hat{x}(t)), \tag{146}$$

deducing an identity

$$\lim_{\epsilon \downarrow 0} \left\langle \hat{\eta}_\epsilon(t) \beta(\hat{x}(t)) \frac{df(\hat{x}(t))}{d\hat{x}} \right\rangle = \frac{1}{2} \left\langle \frac{d}{d\hat{x}} \left\{ \beta(\hat{x}) \frac{df(\hat{x})}{d\hat{x}} \right\} \beta(\hat{x}) \right\rangle = \frac{1}{2} \int dx f(x) \frac{\partial}{\partial x} \beta(x) \frac{\partial}{\partial x} \beta(x) P_t(x), \tag{147}$$

where we have used the partial integration implicitly. Since an integral identity

$$\int dx f(x) \frac{\partial P_t(x)}{\partial t} = \int dx f(x) \left[\frac{\partial}{\partial x} \alpha(x) + \frac{1}{2} \frac{\partial}{\partial x} \beta(x) \frac{\partial}{\partial x} \beta(x) \right] P_t(x) \tag{148}$$

holds for any function $f(x)$, we obtain the Stratonovich-type Fokker-Planck equation (16).

E Derivation of the Liouville Equation (20) for the Collisional Dynamics (17)

In this appendix, the Liouville equation (20) is derived from the collisional microscopic dynamics (17). For an arbitrary function $f(\hat{x}, \hat{v})$, its time-evolusion during $[t, t + dt)$ is given by

$$df(\hat{x}, \hat{v}) = \begin{cases} \hat{v}(t) \dfrac{\partial f(\hat{x}(t), \hat{v}(t))}{\partial \hat{x}} dt & (t^* \notin [t, t+dt)) \\ f(\hat{x}(t), -\hat{v}(t)) - f(\hat{x}(t), \hat{v}(t)) & (t^* \in [t, t+dt)) \end{cases} \tag{149}$$

with $df(\hat{x}, \hat{v}) := f(\hat{x}(t + dt), \hat{v}(t + dt)) - f(\hat{x}(t), \hat{v}(t))$ and infinitesimal positive $dt > 0$. This relation is equivalent to

$$\frac{df(\hat{x}, \hat{v})}{dt} = \hat{v} \frac{\partial f(\hat{x}, \hat{v})}{\partial \hat{x}} + \left[f(\hat{x}, -\hat{v}) - f(\hat{x}, \hat{v}) \right] \delta(t - t^*). \tag{150}$$

Let us take the ensemble average of both sides to obtain

$$\left\langle \frac{df(\hat{x}, \hat{v})}{dt} \right\rangle = \left\langle \hat{v} \frac{\partial f(\hat{x}, \hat{v})}{\partial \hat{x}} + \left[f(\hat{x}, -\hat{v}) - f(\hat{x}, \hat{v}) \right] \delta(t - t^*) \right\rangle. \tag{151}$$

We then consider a relation,

$$\left\langle \hat{v} \frac{\partial f(\hat{x}, \hat{v})}{\partial \hat{x}} \right\rangle = \int_{-\infty}^{\infty} dx dv \, P_t(x, v) v \frac{\partial f(x, v)}{\partial x} = -\int_{-\infty}^{\infty} dx dv f(x, v) v \frac{\partial P_t(x, v)}{\partial x} \tag{152}$$

using the partial integration.

We next make a transformation of the term $\langle [f(\hat{x}, \hat{v}) - f(\hat{x}, -\hat{v})] \delta(t - t^*) \rangle$ using a useful identity

$$\delta(\hat{x}(t) - x^*) = \frac{1}{|d\hat{x}(t)/dt|} \delta(t - t^*) = \frac{1}{|\hat{v}(t)|} \delta(t - t^*), \quad \text{assuming } \hat{x}(t) < x^* \text{ and } \hat{v}(t) > 0. \tag{153}$$
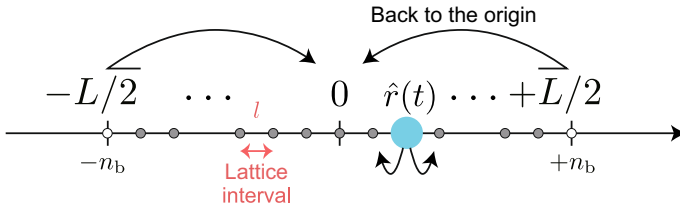
Fig. 15 Discrete random walks model (158) on a lattice with interval l. The particle $\hat{r}(t)$ hops to the nearest-neighbors $\hat{r}(t) - l$ and $\hat{r}(t) + l$ with equal probability, according to the Poisson process of intensity λ. The particle goes back to the origin at the boundary $|\hat{r}(t)| = L/2$ (equivalently, when $|\hat{r}(t)| = ln_b$, the particle goes back to the origin with the intensity $\lambda/2$)

Here the condition $\hat{v}(t) > 0$ is essential because it restricts the collision direction. This implies

$$
\delta(t-t^*) = |\hat{v}(t)|\Theta(\hat{v}(t))\delta(\hat{x}(t)-x^*) \text{with the Heaviside function } \Theta(x) := \begin{cases} 1 & (x > 0) \\ 1/2 & (x = 0) \\ 0 & (x < 0) \end{cases}.
$$
(154)

We obtain

$$
\langle [f(\hat{x}, \hat{v}) - f(\hat{x}, -\hat{v})]\delta(t - t^*) \rangle
$$
$$
= \int_{-\infty}^{\infty} dx dv \, P_t(x, v) \left[f(x, v) - f(x, -v) \right] |v|\Theta(v)\delta(x - x^*)
$$
$$
= \int_{-\infty}^{\infty} dx dv f(x, v) \left[\Theta(v)P_t(x, v) - \Theta(-v)P_t(x, -v) \right] |v|\delta(x - x^*),
$$
(155)

where we have made a variable transformation $-v \to v$ for the second term. We thus obtain an identity for any $f(x, v)$:

$$
\int_{-\infty}^{\infty} dx dv f(x, v) \frac{\partial P_t(x, v)}{\partial t}
$$
$$
= \int_{-\infty}^{\infty} dx dv f(x, v) \left\{ -v \frac{\partial P_t(x, v)}{\partial x} + \left[\Theta(v)P_t(x, v) - \Theta(-v)P_t(x, -v) \right] |v|\delta(x - x^*) \right\},
$$
(156)

leading the Liouville equation (20).

F Derivation Based on the Continuous Limit from a Lattice Model

We have derived the ML equation (22) via Novikov's theorem, which requires advanced techniques on both non-Markovian stochastic processes and kinetic theory. However, considering that the derivation is multidisciplinary, we believe that another more elementary derivation without requiring such technicalities will be helpful for non-experts. Therefore, we provide a more elementary derivation of the reduced ML equation (22) based on a random-walk model on the one-dimensional lattice. This derivation will clearly present the technical issues on the left and right derivatives.

Let us consider a lattice of interval l and the particle dynamics on the lattice (see Fig. 15). The location of the particle is denoted by $\hat{r}(t)$, which obeys the symmetric Poisson jump

process whose jump sizes are $\pm l$ in the absense of collisions against barriers, such that

$$\hat{r}(t + dt) = \hat{r}(t) + \begin{cases} +l & (\text{prob.} = \frac{\lambda}{2} dt) \\ -l & (\text{prob.} = \frac{\lambda}{2} dt) \\ 0 & (\text{prob.} = 1 - \lambda dt) \end{cases} \tag{157}$$

with an infinitesimal positive $dt > 0$ and the total Poisson intensity $\lambda > 0$. We note that the probability of occurrence of a Poisson jump during $[t, t + dt)$ is given by λdt. The hopping barriers are placed at $r = \pm L/2 := \pm l n_b$ with positive integer n_b. At the moment of collisions $\hat{r}(t) = \pm L/2$, the particle comes back to the origin $r = 0$. In summary, the stochastic dynamics is given by

$$\hat{r}(t + dt) = \begin{cases} \hat{r}(t) + l & (\text{prob.} = \frac{\lambda}{2} dt, \text{ if } \hat{r}(t) \leq l(n_b - 2)) \\ \hat{r}(t) - l & (\text{prob.} = \frac{\lambda}{2} dt, \text{ if } \hat{r}(t) \geq l(2 - n_b)) \\ 0 & (\text{prob.} = \frac{\lambda}{2} dt, \text{ if } |\hat{r}(t)| = l(n_b - 1)) \\ \hat{r}(t) & (\text{prob.} = 1 - \lambda dt) \end{cases} . \tag{158}$$

This system is a discrete version of the SDE (6b) and reduces to the SDE (6b) for the continuous limit, as will be discussed later.

F.1 Master-Liouville Equation for the Lattice Model

Let us derive the ML equation for this lattice model as follow. For an arbitrary function $f(\hat{r}(t))$, we obtain the following identity,

$$df(\hat{r}(t)) = \begin{cases} f(\hat{r}(t) + l) - f(\hat{r}(t)) & (\text{prob.} = \frac{\lambda}{2} dt, \text{ if } \hat{r}(t) \leq l(n_b - 2)) \\ f(\hat{r}(t) - l) - f(\hat{r}(t)) & (\text{prob.} = \frac{\lambda}{2} dt, \text{ if } \hat{r}(t) \geq l(2 - n_b)) \\ f(0) - f(\hat{r}(t)) & (\text{prob.} = \frac{\lambda}{2} dt, \text{ if } |\hat{r}(t)| = l(n_b - 1)) \\ 0 & (\text{prob.} = 1 - \lambda dt) \end{cases} . \tag{159}$$

with $df(\hat{r}(t)) := f(\hat{r}(t + dt)) - f(\hat{r}(t))$. By taking the ensemble average of both sides, we obtain

$$\left\langle \frac{df(\hat{r})}{dt} \right\rangle = \left\langle \frac{\lambda}{2} \left[f(\hat{r} + l) - f(\hat{r}) \right] \mathbf{1}_{D_+}(\hat{r}) + \frac{\lambda}{2} \left[f(\hat{r} - l) - f(\hat{r}) \right] \mathbf{1}_{D_-}(\hat{r}) + \frac{\lambda}{2} \left[f(0) - f(\hat{r}) \right] \mathbf{1}_{D_\partial}(\hat{r}) \right\rangle \tag{160}$$

with the indicator function $\mathbf{1}_D(x)$ for the following domain symbols

$$D_+ := (-\infty, l(n_b - 2)], \quad D_- := [-l(n_b - 2), \infty), \quad D_\partial := \{r = \pm l(n_b - 1)\}. \tag{161}$$

The indicator function $\mathbf{1}_D(x)$ is introduced for any domain D, such that

$$\mathbf{1}_D(x) = \begin{cases} 1 & (x \in D) \\ 0 & (x \notin D) \end{cases} . \tag{162}$$

Equation (160) is equivalent to

$$\sum_r f(r) \frac{\partial P_t(r)}{\partial t}$$

$$= \frac{\lambda}{2} \sum_r P_t(r) \left[\{f(r + l) - f(r)\} \mathbf{1}_{D_+}(r) + \{f(r - l) - f(r)\} \mathbf{1}_{D_-}(r) + \{f(0) - f(r)\} \mathbf{1}_{D_\partial}(r) \right]. \tag{163}$$

By substituting $f(r) = \delta_{r,r'}$ with the Kronecker delta $\delta_{r,r'}$, the right-hand side can be transformed into

$$\frac{\lambda}{2} \left[P_t(r' - l)\mathbf{1}_{D_+}(r' - l) - P_t(r')\mathbf{1}_{D_+}(r') + P_t(r' + l)\mathbf{1}_{D_-}(r' + l) - P_t(r')\mathbf{1}_{D_-}(r') \right.$$

$$\left. - P_t(r')\mathbf{1}_{D_\partial}(r') \right] + \frac{\lambda}{2}\delta_{0,r'} \sum_{r \in D_\partial} P_t(r). \tag{164}$$

After replacing $r' \to r$, we thus obtain the ML equation

$$\frac{\partial P_t(r)}{\partial t} = \frac{\lambda}{2} \left\{ P_t(r - l)\mathbf{1}_{D_+}(r - l) - P_t(r)\mathbf{1}_{D_+}(r) + P_t(r + l)\mathbf{1}_{D_-}(r + l) \right.$$

$$\left. - P_t(r)\mathbf{1}_{D_-}(r) - P_t(r)\mathbf{1}_{D_\partial}(r) \right\} + \frac{\lambda}{2}\delta_{0,r} \left\{ P_t(-ln_b + 1) + P_t(ln_b - l) \right\}. \tag{165}$$

This ML equation can be rewritten as

$$\frac{\partial P_t(r)}{\partial t} = \frac{\lambda}{2}\Delta^2 P_t(r) + \begin{cases} 0, & \text{if } l \le |r| \le l(n_b - 1) \\ \frac{\lambda}{2}\left\{ P_t(-ln_b + l) + P_t(ln_b - l) \right\}, & \text{if } r = 0 \\ -\frac{\lambda}{2}P_t(ln_b - l), & \text{if } r = +ln_b \\ +\frac{\lambda}{2}P_t(-ln_b + l), & \text{if } r = -ln_b \end{cases}$$

$$= \frac{\lambda}{2}\Delta^2 P_t(r) + \delta_{r,0}\frac{\lambda}{2}\left\{ P_t(-ln_b + l) + P_t(ln_b - l) \right\}$$

$$- \delta_{r,ln_b}\frac{\lambda}{2}P_t(ln_b - l) + \delta_{r,-ln_b}\frac{\lambda}{2}P_t(-ln_b + l), \tag{166}$$

where we have used the boundary condition[9]

$$P_t(r) = 0 \quad \text{for } |r| \ge ln_b \tag{167}$$

and the second-order difference operator Δ^2 defined by

$$\Delta^2 P_t(r) := P_t(r + l) - 2P_t(r) + P_t(r - l). \tag{168}$$

F.2 Continuous Limit

We take a continuous limit of the lattice model (158) consistently with the original continuous version (6b) according to the system size expansion [4–6]. In other words, we consider the diffusive limit:

$$l \to 0, \quad n_b \to \infty, \quad \lambda \to \infty, \quad \lambda l^2 = \sigma_{cm}^2 = \text{const.}, \quad ln_b = L/2 = \text{const.} \tag{169}$$

The discrete ML equation (166) reduces to the continuous version (22) for the diffusive limit (169). Indeed, by using the Kramers-Moyal expansion [4–6]

$$\Delta^2 P_t(r) \simeq l^2 \frac{\partial^2 P_t(r)}{\partial r^2} + O(l^3), \tag{170}$$

$$P_t(+ln_b - l) = P_t(+L/2 - l) \simeq P_t(+L/2) - l\partial_{-1}P_t(+L/2) + O(l^2) = -l\partial_{-1}P_t(+L/2) + O(l^2), \tag{171}$$

---

[9] The obvious relation $\partial P_t(\pm ln_b)/\partial t = 0$ should hold at the boundary $r = \pm ln_b$, which is consistent with the ML equation (166) under the boundary condition (167).

$$P_t(-ln_b + l) = P_t(-L/2 + l) \simeq P_t(-L/2) + l\partial_{+1}P_t(-L/2) + O(l^2) = +l\partial_{+1}P_t(-L/2) + O(l^2), \tag{172}$$

and replacing the Kronecker δ with the Dirac δ function

$$\lim_{l \downarrow 0} \frac{1}{l}\delta_{r,x} = \delta(r - x), \tag{173}$$

we obtain Eq. (22) from Eq. (166). This derivation is rather elementary without requiring either Novikov's theorem or functional calculus. We believe this derivation will develop the better intuition of the general audience, particularly about the technical issues on the left and right derivatives.

G Derivation of the Exact Steady Solution $\phi(r)$

We derive the exact steady solution $\phi(r)$ with (without) the avoiding potential $U(r)$. In this appendix, we assume the symmetry of the PDF in the steady state:

$$\phi(r) := \lim_{t \to \infty} P_t(r), \quad \phi(r) = \phi(-r), \quad \frac{\partial}{\partial r}\phi(r) = -\frac{\partial}{\partial r}\phi(-r). \tag{174}$$

This symmetry implies that it is sufficient to investigate the steady solution $\phi(r)$ only for $r \in [0, \infty)$.

G.1 In the Absence of the Avoiding Potential $U(r) = 0$

Let us first assume the absence of the avoiding potential $U(r)$. On the basis of the ML equation (22), $\phi(r)$ satisfies

$$\frac{\sigma_{cm}^2}{2}\frac{\partial^2}{\partial r^2}\phi(r) = 0, \quad r \in (0, L/2), \tag{175}$$

which implies the piecewise-linear steady solution

$$\phi(r) = C_1 + C_2 r, \quad r \in (0, L/2) \tag{176}$$

with coefficients C_1 and C_2. We will then consider the boundary condition at $r = 0, L/2$.

G.1.1 Boundary Condition at $r = 0$

The boundary condition at $r = 0$ is given by integrating Eq. (22) over $(-h, +h)$ with infinitesimal positive $h > 0$ as

$$0 = \int_{-h}^{+h} dr \left[\frac{\sigma_{cm}^2}{2}\frac{\partial^2}{\partial r^2}\phi(r) - \delta(r)\frac{\sigma_{cm}^2}{2}\partial_{-1}\phi(L/2) + \delta(r)\frac{\sigma_{cm}^2}{2}\partial_{+1}\phi(-L/2) \right]$$

$$= \frac{\sigma_{cm}^2}{2}\left(\frac{\partial}{\partial r}\phi(h) - \frac{\partial}{\partial r}\phi(-h) \right) - \frac{\sigma_{cm}^2}{2}\partial_{-1}\phi(L/2) + \frac{\sigma_{cm}^2}{2}\partial_{+1}\phi(-L/2)$$

$$= \sigma_{cm}^2\frac{\partial}{\partial r}\phi(h) - \sigma_{cm}^2\partial_{-1}\phi(L/2), \tag{177}$$

where we have used $|\partial_{-s}\phi(sL/2)| = -s\partial_{-s}\phi(sL/2)$. For $h \downarrow 0$, we obtain the boundary condition

$$\partial_{+1}\phi(0) = \partial_{-1}\phi(L/2). \tag{178}$$

G.1.2 Boundary Condition at $r = L/2$

We examine the boundary condition at $r = L/2$. Let us integrate the ML equation (22) over $(L/2 - h, L/2 + h)$ with infinitesimal positive $h > 0$ as

$$0 = \int_{L/2-h}^{L/2+h} dr \left[\frac{\sigma_{cm}^2}{2} \frac{\partial^2}{\partial r^2} \phi(r) + \delta(r - L/2) \frac{\sigma_{cm}^2}{2} |\partial_{-1}\phi(L/2)| \right]$$
$$= \frac{\sigma_{cm}^2}{2} \frac{\partial}{\partial r} \phi(L/2 - h) - \frac{\sigma_{cm}^2}{2} \frac{\partial}{\partial r} \phi(L/2 + h) + \frac{\sigma_{cm}^2}{2} |\partial_{-1}\phi(L/2)|. \tag{179}$$

Considering the fact $\partial_{-1}\phi(L/2) < 0$, we obtain

$$\frac{\partial}{\partial r} \phi(L/2 + h) = \frac{\partial}{\partial r} \phi(L/2 - h) - \partial_{-1}\phi(L/2). \tag{180}$$

By taking the limit $h \to 0$, we obtain the boundary condition

$$\lim_{h \downarrow 0} \frac{\partial}{\partial r} \phi(L/2 + h) = \partial_{+1}\phi(L/2) = 0. \tag{181}$$

G.1.3 Normalisation Condition and the Explicit Steady Solution

Considering the boundary conditions (178) and (181) together with the local solution (176) for $r \in (0, L/2)$, the global solution is given by

$$\phi(r) = \begin{cases} C_1 + C_2 r & (r \in (0, L/2)) \\ C_1 + C_2 L/2 = \text{const.} & (r \in [L/2, \infty)) \end{cases}. \tag{182}$$

The coefficients C_1 and C_2 for the steady solution (176) is determined by the normalisation condition:

$$\int_{-\infty}^{\infty} dr \phi(r) = 2 \int_0^{\infty} dr \phi(r) = 1, \tag{183}$$

which deduces the tent function (65) with $C_1 = 2/L$ and $C_2 = -4/L^2$.

G.2 In the Presence of the Avoiding Potential $U(r) \neq 0$

From Eq. (112) for $r \in (0, L/2)$, we obtain

$$\frac{d}{dr} \left[U'(r) + \frac{\sigma_{cm}^2}{2} \frac{d}{dr} \right] \phi(r) = 0. \tag{184}$$

The boundady condition at $r = 0$ is given by integrating the both hand sides (BHSs) of Eq. (112a) over $(-h, h)$ with $h > 0$:

$$0 = \frac{\sigma_{cm}^2}{2} \left(\frac{\partial}{\partial r} \phi(+h) - \frac{\partial}{\partial r} \phi(-h) \right) - \frac{\sigma_{cm}^2}{2} \frac{\partial}{\partial r} \phi(L/2) + \frac{\sigma_{cm}^2}{2} \frac{\partial}{\partial r} \phi(-L/2) + o(h), \tag{185}$$

which is equivalent to

$$\frac{\partial}{\partial r} \phi(+0) = \frac{\partial}{\partial r} \phi(L/2) \tag{186}$$

for $h \downarrow 0$, by considering the symmetry $(\partial/\partial r)\phi(r) = -(\partial/\partial r)\phi(-r)$. The boundary condition at $r = L/2$ is also given by integrating the BHSs of Eq. (112a) over $(L/2 - h, L/2 + h)$ with $h > 0$:

$$0 = \frac{\sigma_{\rm cm}^2}{2}\left(\frac{\partial}{\partial r}\phi(L/2 + h) - \frac{\partial}{\partial r}\phi(L/2 - h)\right) + \frac{\sigma_{\rm cm}^2}{2}\frac{\partial}{\partial r}\phi(L/2) + o(h). \qquad (187)$$

Since $(\partial/\partial r)\phi(L/2 + h) = 0$, the boundary condition (187) always holds for $h \downarrow 0$. Furthermore, the normalisation condition is given by

$$\int_0^{L/2} \phi(r)dr = \frac{1}{2}. \qquad (188)$$

By solving Eq. (184) under the conditions (186) and (188), we obtain the exact solution (113).

H Consistency Confirmation with the Reduced Master-Liouville Equation (22)

Let us confirm the consistency between the full and reduced ML equations (68) and (22). In this Appendix, the technical symbol on the left and right derivatives are abbreviated for simplicity, such as $\tilde{\partial}_{12;s} \to \tilde{\partial}_{12}$. Starting from the full ML equation (68), let us apply a variable transformation:

$$\hat{z}_{\rm cm} := \frac{\hat{z}_1 + \hat{z}_2}{2}, \quad \hat{r} := \hat{z}_1 - \hat{z}_{\rm cm} = \frac{\hat{z}_1 - \hat{z}_2}{2} \implies \hat{z}_1 = \hat{z}_{\rm cm} + \hat{r}, \quad \hat{z}_2 = \hat{z}_{\rm cm} - \hat{r}. \quad (189)$$

This implies the derivative chain rule

$$\frac{\partial}{\partial z_1} = \frac{1}{2}\left(\frac{\partial}{\partial z_{\rm cm}} + \frac{\partial}{\partial r}\right), \quad \frac{\partial}{\partial z_2} = \frac{1}{2}\left(\frac{\partial}{\partial z_{\rm cm}} - \frac{\partial}{\partial r}\right). \qquad (190)$$

We thus obtain

$$\frac{\partial P_t(z_{\rm cm}, r)}{\partial t} = \frac{\sigma_{\rm cm}^2}{2}\left[\frac{\partial^2}{\partial z_{\rm cm}^2} + \frac{\partial^2}{\partial r^2}\right] P_t(z_{\rm cm}, r) + \sum_{s=\pm 1}\left[J_{t;s}(z_{\rm cm}, r - sL/2) - J_{t;s}(z_{\rm cm}, r)\right],$$

$$\qquad (191)$$

$$J_{t;s}(z_{\rm cm}, r) := -\frac{s\sigma_{\rm cm}^2}{2}\delta(r - sL/2)\frac{\partial}{\partial r}P_t(z_{\rm cm}, r), \qquad (192)$$

where we have used $\sigma_{\rm cm}^2 = \sigma^2/2$. By introducing the reduced PDF

$$P_t(r) := \int_{-\infty}^{\infty} dz_{\rm cm} P_t(z_{\rm cm}, r) \qquad (193)$$

as the result of the marginalisation, we obtain the reduced ML equation (22). This consistency suggests that the kinetic formulation for the dealer model and the technical calculation therein are very reasonable from various viewpoints.

I Derivation of the ML Equation with Avoiding Potential $U(r)$

Equation (112) can be derived in an essentially parallel manner to Eq. (61). Let us first replace the white Gaussian noise $\hat{\eta}(t)$ with a coloured Gaussian noise $\hat{\eta}_\epsilon(t)$, satisfying $\langle\hat{\eta}_\epsilon(t)\rangle = 0$

and $\langle \hat{\eta}_\epsilon(t_1)\hat{\eta}_\epsilon(t_2)\rangle = (1/(2\epsilon))e^{-|t_1-t_2|/\epsilon}$. The time evolution of an arbitrary function $f(\hat{r})$ is given by

$$
df(\hat{r}) = \begin{cases} \dfrac{df(\hat{r})}{d\hat{r}}\left(-U'(\hat{r})+\sigma_{\mathrm{cm}}\hat{\eta}_\epsilon\right)dt & (\text{with a collision: } \hat{\tau}_{s;i} \notin [t, t+dt)) \\ f(\hat{r}-sL/2)-f(\hat{r}) & (\text{with a collision: } \hat{\tau}_{s;i} \in [t, t+dt)) \end{cases}, \quad (194)
$$

where $df(\hat{r}(t)) := f(\hat{r}(t+dt)) - f(\hat{r}(t))$ and $\hat{\tau}_{s;i}$ is the ith arrival time of the particle at $\hat{r}(\hat{\tau}_{s;i}) = sL/2$ for $s = \pm 1$. Using the δ functions, this is equivalent to

$$
\begin{aligned}
\frac{df(\hat{r})}{dt} &= \frac{df(\hat{r})}{d\hat{r}}\left(-U'(\hat{r})+\sigma_{\mathrm{cm}}\hat{\eta}_\epsilon\right) + \sum_{s=\pm1}\sum_{i=1}\left[f(\hat{r}-sL/2)-f(\hat{r})\right]\delta(t-\hat{\tau}_{s;i}) \\
&= \frac{df(\hat{r})}{dr}\left(-U'(\hat{r})+\sigma_{\mathrm{cm}}\hat{\eta}_\epsilon\right) + \sum_{s=\pm1}\left[f(\hat{r}-sL/2)-f(\hat{r})\right]\left|\frac{d\hat{r}}{dt}\right|\delta(\hat{r}-sL/2).
\end{aligned}
$$
$$(195)$$

By considering the sign of the velocity $d\hat{r}/dt$ as $d\hat{r}/dt > 0$ for $\hat{r} = L/2$ and $d\hat{r}/dt < 0$ for $\hat{r} = -L/2$, we obtain $|d\hat{r}/dt| = s(-U'(\hat{r})+\sigma_{\mathrm{cm}}\hat{\eta}_\epsilon)$ just before the collosion at $\hat{r} = sL/2$. We thus have

$$
\frac{df(\hat{r})}{dt} = \frac{df(\hat{r})}{d\hat{r}}\left(-U'(\hat{r})+\sigma_{\mathrm{cm}}\hat{\eta}_\epsilon\right) + \sum_{s=\pm1} s\left(-U'(\hat{r})+\sigma_{\mathrm{cm}}\hat{\eta}_\epsilon\right)\left[f(\hat{r}-sL/2)-f(\hat{r})\right]\delta(\hat{r}-sL/2).
$$
$$(196)$$

By taking the ensemble average of BHSs, we obtain

$$
\begin{aligned}
\left\langle\frac{df(\hat{r})}{dt}\right\rangle &= \left\langle\frac{df(\hat{r})}{d\hat{r}}\left(-U'(\hat{r})+\sigma_{\mathrm{cm}}\hat{\eta}_\epsilon\right)\right. \\
&\quad \left. + \sum_{s=\pm1} s\left(-U'(\hat{r})+\sigma_{\mathrm{cm}}\hat{\eta}_\epsilon\right)\left[f(\hat{r}-sL/2)-f(\hat{r})\right]\delta(\hat{r}-sL/2)\right\rangle. \quad (197)
\end{aligned}
$$

By substituting $f(\hat{r}) = \delta(\hat{r}-r)$, we first obtain the following relations:

$$
\left\langle\frac{df(\hat{r})}{dt}\right\rangle = \frac{d}{dt}\langle f(\hat{r})\rangle = \frac{d}{dt}\int_{-\infty}^{\infty}dr' f(r')P_t(r') = \frac{\partial}{\partial t}P_t(r), \quad (198)
$$

$$
\begin{aligned}
\left\langle\frac{df(\hat{r})}{d\hat{r}}U'(\hat{r})\right\rangle &= \int_{-\infty}^{\infty}P_t(r')\frac{df(r')}{dr'}U'(r')dr' = -\int_{-\infty}^{\infty}f(r')\frac{\partial}{\partial r'}\left[U'(r')P_t(r')\right]dr \\
&= \frac{\partial}{\partial r}\left[U'(r)P_t(r)\right], \quad (199)
\end{aligned}
$$

$$
\begin{aligned}
\left\langle U'(\hat{r})f(\hat{r}-sL/2)\delta(\hat{r}-sL/2)\right\rangle &= \int_{-\infty}^{\infty}dr' P_t(r')U'(r')f(\hat{r}'-sL/2)\delta(r'-sL/2) \\
&= P_t(sL/2)U'(sL/2)\delta(r) = 0, \quad (200)
\end{aligned}
$$

and

$$
\begin{aligned}
\left\langle U'(\hat{r})f(\hat{r})\delta(\hat{r}-sL/2)\right\rangle &= \int_{-\infty}^{\infty}dr' P_t(r')U'(r')f(\hat{r}')\delta(r'-sL/2) \\
&= P_t(sL/2)U'(sL/2)\delta(r-sL/2) = 0, \quad (201)
\end{aligned}
$$

Table 1 Parameter table for the numerical simulations

Variable	Meaning	Numerical value	Dimension
L	Spread constant	2	Price
u^2	Strength of the avoiding potential	0 or 1	Time^{-1}
σ^2	Variance of the random walks	1	Price^2/time
Δt	Discrete timestep	$10^{-4}/\sqrt{N/2}$	Time
T_{ini}	Time for the initialisation	20	Time
T_{end}	Time for the sampling	10^4	Time
Δr	Interval between the bins	10^{-2}	Price
$r_{\max}$	Maximum value of the bins	3	Price
$r_{\min}$	Minimum value of the bins	-3	Price

Here Δt depends on the total number of the traders N because the minimum characteristic length $\varepsilon :=$
$L/(2\sqrt{N})$ depends on N likewise

where we have used $P_t(sL/2) = 0$ (i.e., $P_t(r)$ must have the vanishing probability at the boundary). Using Novikov's theorem and substituting $f(\hat{r}) = \delta(\hat{r} - r)$, in addition, we obtain

$$\lim_{\epsilon \downarrow 0} \left\langle \sigma_{\text{cm}} \frac{df(\hat{r})}{d\hat{r}} \hat{\eta}_\epsilon \right\rangle = \left\langle \frac{\sigma_{\text{cm}}^2}{2} \frac{df(\hat{r})}{d\hat{r}^2} \right\rangle = \int_{-\infty}^{\infty} f(r') \frac{\sigma_{\text{cm}}^2}{2} \frac{\partial^2}{\partial r'^2} P_t(r') dr' = \frac{\sigma_{\text{cm}}^2}{2} \frac{\partial^2}{\partial r^2} P_t(r), \tag{202}$$

$$\lim_{\epsilon \downarrow 0} \left\langle \hat{\eta}_\epsilon f(\hat{r} - sL/2) \delta(\hat{r} - sL/2) \right\rangle = \left\langle \frac{\sigma_{\text{cm}}}{2} \frac{\partial}{\partial \hat{r}} f(\hat{r} - sL/2) \delta(\hat{r} - sL/2) \right\rangle$$
$$= \int_{-\infty}^{\infty} dr' P_t(r') \frac{\sigma_{\text{cm}}}{2} \frac{\partial}{\partial r'} f(r' - sL/2) \delta(r' - sL/2)$$
$$= -\frac{\sigma_{\text{cm}}}{2} \delta(r) \frac{\partial}{\partial r} P_t(sL/2), \tag{203}$$

and

$$\lim_{\epsilon \downarrow 0} \left\langle \hat{\eta}_\epsilon f(\hat{r}) \delta(\hat{r} - sL/2) \right\rangle = \left\langle \frac{\sigma_{\text{cm}}}{2} \frac{\partial}{\partial \hat{r}} f(\hat{r}) \delta(\hat{r} - sL/2) \right\rangle$$
$$= \int_{-\infty}^{\infty} dr' P_t(r') \frac{\sigma_{\text{cm}}}{2} \frac{\partial}{\partial r'} f(r') \delta(r' - sL/2)$$
$$= -\frac{\sigma_{\text{cm}}}{2} \delta(r - sL/2) \frac{\partial}{\partial r} P_t(sL/2). \tag{204}$$

In summary, we have

$$\frac{\partial}{\partial t} P_t(r) = \frac{\partial}{\partial r} \left[U'(r) + \frac{\sigma_{\text{cm}}^2}{2} \frac{\partial}{\partial r} \right] P_t(r) + \sum_{s=\pm 1} [\delta(r) - \delta(r - sL/2)] \left(-s \frac{\sigma_{\text{cm}}^2}{2} \frac{\partial}{\partial r} P_t(sL/2) \right), \tag{205}$$

which is equivalent to Eq. (112) by considering that $(\partial/\partial r) P_t(L/2) < 0$ and $(\partial/\partial r) P_t(-L/2) > 0$.

J Numerical Scheme for the *N*-Body Dealer Model

We explain the numerical scheme for the N-body dealer model. We also note that the original codes are available as "Supplementary Code S1.jl" (without potential $U(r) = 0$) and "Supplementary Code S2.jl" (with potential $U(r) = u^2 r^2/2$), which can be run on the Julia Programming Language (see Appendix K for the data availability statements).

Let us define the bid price, ask price, and midprice of the ith trader as $\hat{b}_i$, $\hat{a}_i$, and $\hat{z}_i := (\hat{b}_i + \hat{a}_i)/2$ for $i \in 1, 2, \ldots, N$. We assume that the spread is identical to all traders, such that $\hat{a}_i - \hat{b}_i = L = $ const. for all $i = 1, 2, \ldots, N$. The discrete time is introduced as $t_k = k\Delta t$ with discrete timestep $\Delta t > 0$ and integer k. We run the simulation during $[-T_{\text{ini}}, T_{\text{end}})$, where the data sample during $[-T_{\text{ini}}, 0)$ is discorded for the initialisation. The dynamical equation is basically given by

$$\hat{z}_i(t_{k+1}) = \hat{z}_i(t_k) - u^2(\hat{z}_i(t_k) - \hat{z}_{\text{M}}(t_k))\Delta t + \sigma\sqrt{\Delta t}\hat{\xi}(t_k), \quad \hat{z}_{\text{M}}(t_k) := \frac{1}{2}\left(\hat{z}_1(t_k) + \hat{z}_2(t_k)\right) \tag{206}$$

with the normal random number $\hat{\xi}(t_k)$ with the unit variance $\langle\hat{\xi}(t_k)\hat{\xi}(t_{k'})\rangle = \delta_{k,k'}$. If there is a transaction, there occurs a jump representing resubmissions:

$$|\hat{z}_i(t_k) - \hat{z}_j(t_k)| \geq L \implies \hat{z}_i(t_{k+1}) = \hat{z}_j(t_{k+1}) = \frac{\hat{z}_i(t_k) + \hat{z}_j(t_k)}{2}. \tag{207}$$

Even though it is a rare event for $dt \downarrow 0$, but if there are multiple collisions at the same time, only one pair composed of the highest bid and lowest ask is selected to make a transaction during one discrete timestep. The steady PDF $\phi_N(r)$ is calculated according to the following formula,

$$\phi_N(r) := \left\langle \frac{1}{N}\sum_{i=1}^{N}\delta(\hat{z}_i - \hat{z}_{\text{cm}} - r)\right\rangle \simeq \frac{1}{T_{\text{end}}}\sum_{t_k \in [0, T_{\text{end}})}\Delta t\left(\frac{1}{N}\sum_{i=1}^{N}\delta(\hat{z}_i - \hat{z}_{\text{cm}} - r)\right) \tag{208}$$

for sufficiently large T_{end} by assuming the ergodicity. Numerically, we introduce the bins, such as $[r_k, r_{k+1})$ where $r_k := r_0 + k\Delta r$, $r_0 = r_{\min}$, $r_{N_r} = r_{\max}$, and $N_r := (r_{\max} - r_{\min})/\Delta r$ and then make an approximation on the δ function as

$$\delta(\hat{z}_i - \hat{z}_{\text{cm}} - r) \simeq \begin{cases} 1/\Delta r & \text{if } r \in [r_k, r_{k+1}) \text{ and } \hat{z}_i - \hat{z}_{\text{cm}} \in [r_k, r_{k+1}) \text{ for some } k \in 0, \ldots, N_r - 1 \\ 0 & \text{otherwise} \end{cases}. \tag{209}$$

By applying this binning approximation to Eq. (208), we numerically evaluate the steady PDF $\phi_N(r)$ under the parameter set summarised in Table 1.

References

1. Kubo, R., Toda, M., Hashitsume, N.: Statistical Physics II: Non-equilibrium Statistical Mechanics, 2nd edn. Springer, Berlin (1991)
2. Chapman, S., Cowling, T.G.: The Mathematical Theory of Non-uniform Gases. Cambridge University Press, Cambridge (1970)
3. Hansen, J.-P., McDonald, I.R.: Theory of Simple Liquids, 3rd edn. Academic Press, London (2006)
4. van Kampen, N.G.: Stochastic Processes in Physics and Chemistry, 3rd edn. Elsevier, Amsterdam (2007)
5. van Kampen, N.G.: A power series expansion of the master equation. Can. J. Phys. **39**, 551 (1961)
6. Gardiner, C.W.: Handbook of Stochastic Methods, 4th edn. Springer, Berlin (2009)
7. Bachelier, L.: Théorie de la spéculation. Annales Scientifiques de l'École Normale Supérieure **17**, 21 (1900)

8. Einstein, A.: Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Ann. Phys. **322**, 549 (1905)

9. Mantegna, R.N., Stanley, H.E.: Introduction to Econophysics. Cambridge University Press, Cambridge (1999)

10. Slanina, F.: Essentials of Econophysics Modelling. Oxford University Press, Oxford (2014)

11. Bouchaud, J.-P., Bonart, J., Donier, J., Gould, M.: Trades, Quotes and Prices. Cambridge University Press, Cambridge (2018)

12. Jusup, M., et al.: Social physics. Phys. Rep. **948**, 1 (2022)

13. Takayasu, H., Miura, H., Hirabayashi, T., Hamada, K.: Statistical properties of deterministic threshold elements—the case of market price. Phys. A: Stat. Mech. Appl. **184**, 127 (1992)

14. Sato, A.-H., Takayasu, H.: Dynamic numerical models of stock market price: from microscopic determinism to macroscopic randomness. Phys. A: Stat. Mech. Appl. **250**, 231 (1998)

15. Yamada, K., Takayasu, H., Ito, T., Takayasu, M.: Solvable stochastic dealer models for financial markets. Phys. Rev. E **79**, 051120 (2009)

16. Kanazawa, K., Sueshige, T., Takayasu, H., Takayasu, M.: Derivation of the Boltzmann equation for financial Brownian motion: direct observation of the collective motion of high-frequency traders. Phys. Rev. Lett. **120**, 138301 (2018)

17. Kanazawa, K., Sueshige, T., Takayasu, H., Takayasu, M.: Kinetic theory for financial Brownian motion from microscopic dynamics. Phys. Rev. E **98**, 052317 (2018)

18. Novikov, E.A.: Functionals and the random-force method in turbulence theory. Sov. Phys. JETP **20**, 1290 (1965)

19. Hänggi, P.: Correlation functions and masterequations of generalized (non-Markovian) Langevin equations. Z. Phys. B **31**, 407 (1978)

20. van Kampen, N.G.: Itô versus Stratonovich. J. Stat. Phys. **24**, 175 (1981)

21. Zwanzig, R.: Nonequilibrium Statistical Mechanics. Oxford University Press, New York (2001)

22. Kupferman, R.: Fractional kinetics in Kac-Zwanzig heat bath models. J. Stat. Phys. **114**, 291 (2004)

23. Kanazawa, K., Sornette, D.: Field master equation theory of the self-excited Hawkes process. Phys. Rev. Res. **2**, 033442 (2020)

24. Kanazawa, K., Sornette, D.: Nonuniversal power law distribution of intensities of the self-excited Hawkes process: a field-theoretical approach. Phys. Rev. Lett. **125**, 138301 (2020)

25. Kanazawa, K., Sornette, D.: Ubiquitous power law scaling in nonlinear self-excited Hawkes processes. Phys. Rev. Lett. **127**, 188301 (2021)

26. Wong, E., Zakai, M.: On the relation between ordinary and stochastic differential equations. Int. J. Eng. Sci. **3**, 213 (1965)

27. Wong, E., Zakai, M.: On the convergence of ordinary integrals to stochastic integrals. Ann. Math. Stat. **36**, 1560 (1965)

28. Brilliantov, N.V., Pöschel, T.: Kinetic Theory of Granular Gases. Oxford University Press, England (2004)

29. Kanazawa, K.: Statistical Mechanics for Athermal Fluctuations (Springer Theses). Springer Nature, Singapore (2017)

30. Daniels, M.G., Farmer, J.D., Gillemot, L., Iori, G., Smith, E.: Quantitative model of price diffusion and market friction based on trading as a mechanistic random process. Phys. Rev. Lett. **90**, 108102 (2003)

31. Smith, E., Farmer, J.D., Gillemot, L., Krishnamurthy, S.: Statistical theory of the continuous double auction. Quant. Financ. **3**, 481 (2003)

32. Bouchaud, J.-P., Mézard, M., Potters, M.: Statistical properties of stock order books: empirical results and models. Quant. Financ. **2**, 251 (2002)

33. Donier, J., Bonart, J., Mastromatteo, I., Bouchaud, J.-P.: A fully consistent, minimal model for non-linear market impact. Quant. Financ. **15**, 1109 (2015)