Article

From Meso- to Microscopic Stochastic Modeling in Physico-Chemical Processes – the Paradigm of Chemical Reacting Systems

Chiara Pezzotti and Massimiliano Giona *

Article

From Meso- to Microscopic Stochastic Modeling in Physico-Chemical Processes—The Paradigm of Chemical Reacting Systems

Chiara Pezzotti [1,†,‡] (ORCID) **and Massimiliano Giona** [2,‡]

1   chiara.pezzotti@uniroma1.it
2   massimiliano.giona@uniroma1.it
*   Correspondence: massimiliano.giona@uniroma1.it
†   Current address: Dipartimento di Ingegneria Chimica, Materiali, Ambiente La Sapienza Università di Roma Via Eudossiana 18, 00184 Roma, Italy
‡   These authors contributed equally to this work.

Abstract: The stochastic modelling of chemical physical processes is essentially based on a coarse-grained formulation of fluctuations, usually described by means of Wiener processes. By taking the paradigmatic example of chemical reaction kinetics, we propose a simple representation of microscopic stochastic dynamics grounded on the use of distributional derivatives of counting processes that account for elementary reactive events. This approach is consistent with the statistical analysis based on the Chemical Master Equation, and provides the formal setting for the existing algorithmic approaches (Gillespie algorithm). Some practical advantages of this formulation are addressed, and several other examples are discussed in connection with cluster formation, velocity fluctuations in liquid systems, and quantum transitions (radiative interactions) at the molecular level.

Keywords: stochastic modeling; chemical reactions; microscopic models; coalescence; radiative processes

---

1. Introduction

All the chemical physical processes involve, in an atomistic perspective, a stochastic description of the events, be them reactive, associated with a change of phase (for instance adsorption) [1], or simply with momentum transfer, such as in collisions. Nonetheless, in the overwhelming majority of the cases of practical and theoretical interest, the number of molecules involved is so large to justify a mean field approach, essentially based on the Boltzmannian hypothesis of molecular chaos (the"stosszahlansatz") [2]. The mean field formulation represents the backbone of the classical theory of chemical reaction kinetics [3,4]. In between the mean-field formulation - expressed in terms of continuous fields evolving in space and time (representing the core of transport theory and of the theory of irreversible processes) [5–7] - and molecular simulations, involving force fields and quantum effects [8,9], stochastic models constitute an alternative capable of accounting for thermal and statistical fluctuations [10] and, at the same time, providing an interpretation for the transport equations of continuous theories as evolution equation for the associated probability densities [11,12].

In most of the cases, the stochastic fluctuations controlling the evolution of chemical-physical processes correspond to the superposition of a variety of elementary events at molecular level, subsumed within the evolution equation by simplified stochastic models, the basic property of which lies in their coarse-grainedness (mathematically corresponding to the application of Central Limit Theorem), independently of the specific phenomenology occurring at the molecular level [13]. It is therefore not surprising that these stochastic approaches are essentially based on a Gaussian assumption for the probability densities of the fluctuating perturbations, leading to parabolic models for the transport equations of the classical theory of irreversible processes [6].

However, it is possible to provide alternatives to this approach, defining the stochastic dynamics at a microscopic level, and accounting for the elementary chemical-physical processes involved at this

scale. The latter approach, beyond its intrinsic relevance, offers also the way to determine under what conditions the Gaussian approximation in coarse-grained models can be assumed to be valid, and if this is not the case, how specific microscopic interactions may lead to macroscopic transport anomalies.

In this article we address the mathematical formulation of microscopic stochastic modelling using chemical reaction kinetics as a prototypical case study. The rationale in these models is the description of the elementary events (reactive molecular collisions) as distributional derivatives of counting processes [10]. This type of formulation is not only important in the study of chemical reaction kinetics, but it also provides a very flexible conceptual framework for addressing any chemical-physical process in terms of elementary events involving individual molecular entities. Another typical example of this approach is represented by the thermalization dynamics in a molecular mixture controlled by radiative effects and related to the emission and absorption of energy quanta (see Section 6).

The article is organized as follows. Section 2 discusses in detail the multiscale modeling of irreversible processes, the mesoscopic nature of the existing stochastic formulations and the need for a microscopic event-based approach. Section 3 provides a succinct overview on the state of the art in the stochastic modelling of chemical kinetics and on its algorithmic representation. Section 4 addresses the mathematical formulation of stochastic chemical reactions as a paradigm for any event-based microscopic description. Section 5 applies this approach to the clustering of liquid molecules and its influence on the velocity distribution function at thermal equilibrium. Finally, Section 6 analyzes the role of radiative interactions (absorption and emission of photons) in thermalization processes, and the occurrence of deviations from the Gaussian distribution in equilibrium velocity densities.

2. Models of irreversible processes and their multiscale nature

Chemical physical processes can be described at different length scales, depending on the focus/attention on the microscopic interactions. A rough classification of possible representations, customarily adopted, involves three major classes: macroscopic, mesoscopic, and microscopic (figure 1). The two extrema in this categorization are rather clear. Macroscopic modeling involves the evolution of smooth space-time fields, such, e.g., the solute concentration $c(\mathbf{x}, t)$ in a fluid flow, and the resulting dynamics for the continuous fields is expressed in the form of partial differential equations, in most of the cases of parabolic nature (first-order in time, second-order with respect to the spatial variables). At the opposite edge, microscopic models assume the atomistic constitution of matter (molecules) and, considering the interaction between molecules, eventually including quantum effects as regards charge and dipolar moment distributions, describe the evolution of an ensemble of molecular entities via Newtonian equations of motion, complemented with suitable algorithmic/analytic solutions to account for specific thermodynamic conditions (thermostats). The goal of the molecular simulations is essentially to determine the values of transport coefficients and thermodynamic parameters, or the occurrence of phase transitions.

A particular role is played by non-equilibrium kinetic models, such as the Boltzmann equation. Depending on the way they are regarded and utilized, they are able to spam both a microscopic description (as binary collision are explicitly considered, and the collisional integral accounts for the Markovian transitions in the particle velocities controlled by elementary collisional events) [14,15] or the macroscopic transport theory (whenever the lower-order moments are considered to derive the transport equations for the collisional invariants).
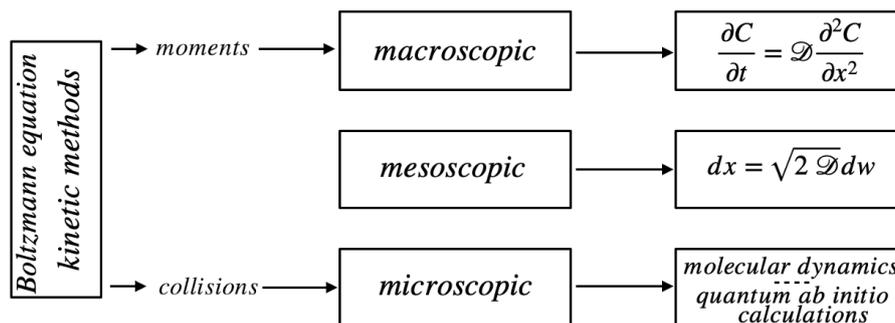
Figure 1. Level-based categorization of models in the chemical-physical theories.

In between the two aforementioned cases, there is a sea of intermediate-scale approaches that consistently use stochastic processes to model the effects of fluctuations (be them thermal, deriving from the granularity of matter, or of quantum nature) in a coarse-grained way. These models stem from the pioneering work of Einstein, Langevin, Smoluchowski on Brownian motion [16,17], i.e. the highly erratic kinematics of micrometric particles in a still fluid driven by thermal fluctuations. The equations of motion for the particles/molecules are derived taking into account the presence of hydrodynamics and/or of external potentials, and describing in a coarse-grained way the effects of thermal fluctuations. The statistical description of the stochastic motion of the particles in terms of probability densities corresponds to the macroscopic thermodynamic level involving the spatio-temporal evolution of the concentration fields.

The prototypical example of mean-field stochastic modelling refers to the motion of micrometric particles in a fluid, the hydrodynamic velocity field of which is $\mathbf{u}_f(\mathbf{x}) = (u_{f,1}(\mathbf{x}), u_{f,2}(\mathbf{x}), u_{f,3}(\mathbf{x}))$, $\mathbf{x} \in \mathbb{R}^3$. Assuming the size of the particle small compared to the length scale of the velocity variations, and isothermal conditions (constant temperature T), the equation of motion takes the form of a Wiener-driven Langevin equation

$$d\mathbf{x} = \mathbf{u}_f(\mathbf{x})\, dt + \sqrt{2\, D}\, d\mathbf{w}(t) \tag{1}$$

where $d\mathbf{w}(t) = (dw_1(t), dw_2(t), dw_3(t))$ is the increment of a vector-valued Wiener process in the time interval $(t, t + dt)$ (and thus the entries $dw_i(t)$, $i = 1, 2, 3$ are independent of each other), and D the diffusion coefficient. From eq. (2) it follows that the associated probability density $p(\mathbf{x}, t)$ satisfies the Fokker-Planck equation

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = -\nabla \cdot \left[\mathbf{u}_f(\mathbf{x})\, p(\mathbf{x}, t) \right] + D\, \nabla^2 p(\mathbf{x}, t) \tag{2}$$

which corresponds to the classical macroscopic advection-diffusion equation of mass transport theory for the solute concentration that, modulo a multiplicative constant, equals the density $p(\mathbf{x}, t)$.

Eq. (1) is a mesoscopic description of motion because the transport coefficient controlling the process, i.e. the diffusivity D, is already included into the model without a derivation of it from a microscopic interaction mechanism. Of course one could invoke a more fundamental stochastic description involving the momentum transfer mechanism, as originally formulated by Langevin: if $\mathbf{v}$ is the particle velocity and m its mass then

$$m\, d\mathbf{v} = -\eta \left[\mathbf{v} - \mathbf{u}_f(\mathbf{x}) \right] dt + \sqrt{2\, k_B\, T\, \eta}\, d\mathbf{w}(t) \tag{3}$$

where η is the friction factor, stemming from the hydrodynamic interactions with the solvent fluid, and k_B the Boltzmann constant. Eq. (2) follows from eq. (3) in the limit of vanishing inertia (overdamped approximation) where $m\, d\mathbf{v} = 0$, enforcing

$$d\mathbf{x} = \mathbf{v}\, dt \tag{4}$$

and the Stokes-Einstein fluctuation-dissipation relation

$$D = \frac{k_B\, T}{\eta} \tag{5}$$

connecting the diffusivity D to the hydromechanical properties, expressed by the friction factor η, and to the thermodynamic conditions of the systems, expressed by the temperature T.

But even eq. (3) is a mesoscopic description of the thermal fluctuations: the use of the Wiener process $\mathbf{w}(t)$, by its nature, corresponds to a long-term asymptotic characterization, emerging from the superposition of small and independent random pertubations so that the Central Limit Theorem (CLT) can be invoked [18,19].

It is therefore not surprising that the resulting transport equation is of parabolic nature as, by definition of the (one-dimensional) Wiener process $w(t)$, its increment $y = w(t_2) - w(t_1)$ over the time interval $t_2 - t_1 > 0$ is a Gaussian random variable with zero mean and square variance equal to $t_2 - t_1$ [13]. Moreover, the increments in two disjoint time intervals are independent of each other. Setting $y = w(t) - w(0) = w(t)$, since $w(0) = 0$ by definition [10], the probability density for y is Gaussian

$$p_y(y) = p_w(w,t)|_{w=y} = \left. \frac{1}{\sqrt{2\,\pi\, t}}\, e^{-w^2/2\,t} \right|_{w=y} \tag{6}$$

and $p(w,t)$ so defined is the solution of the diffusion equation

$$\frac{\partial p_w(w,t)}{\partial t} = \frac{1}{2}\, \frac{\partial^2 p_w(w,t)}{\partial w^2} \tag{7}$$

Therefore, the choice of Wiener perturbations as the global descriptor of the effects of thermal fluctuations automatically implies that the resulting macroscopic transport equations for the densities are partial differential equations of parabolic nature.

The mesoscopic formulation based on Wiener-Langevin equations of the form eq. (1) represents one of the basic strongholds of contemporary statistical physics. At the same time, this setting shows intrinsic physical limitations in the broader perspective of improving the formulation of macroscopic thermodynamic models. The parabolic nature of the resulting transport equations brings with it inherent inconsistencies such as the paradox of infinite velocity of propagation [7], and this limits its application to generic (and extreme) thermodynamic conditions (such as at very low temperatures), and to generic fluids (more complex that purely Newtonian fluids, such as viscoelastic fluids) [20–22]. Moreover, there is no relation between a microscopic theory of mass transport and the properties of the stochastic force, other than the appeal to CLT. As in the absence of the external velocity field ($\mathbf{u}_f = 0$), the statistics of the velocity fluctuations deriving from eq. (3) is Gaussian, there is no indication in this form of mesoscopic stochastic theories of the conditions under which the Gaussian paradigm may fail, and new emergent phenomenologies may arise.

In order to answer these questions, and to provide a framework for describing thermal fluctuations beyond the pure CLT limit, a microscopic stochastic theory of motion is required, where the elementary mechanisms controlling momentum transfer and motion are singled out and accounted for. This may represent a general microscopic approach for handling a huge variety of chemical physical phenomena. The paradigmatic problem of such a microscopic event-based theory is represented by chemical reactions, as addressed in the next two Sections 3 and 4. In the remainder of the article, we consider several extensions of this approach to the analysis of a model for molecular coalescence in liquids and its implications as regards transport (Section 5) and to the velocity distribution in molecular systems driven by radiative effects (emission and absorptions of photons) (Section 6).

3. Stochastic modelling of chemical kinetics

Consider a typical bimolecular reaction mechanism

$$A + B \rightarrow C \tag{8}$$

where A, B and C indicate the reacting species. The stoichiometry defined by eq. (8) implies that one molecule of A and one molecule of B should collide, and that their kinetic energy is sufficiently high to overcome the activation energy so that the chemical transformation to C may take place.

In a mean-field framework, letting c_A, c_B and c_C be the concentrations of the reacting species, and assuming spatial uniformity in a closed reacting vessel, so that these concentrations depend uniquely on time, we have

$$\frac{dc_A}{dt} = \frac{dc_B}{dt} = -\frac{dc_C}{dt} = -kc_A\,c_B \tag{9}$$

where k is the rate constant, accounting for the frequency of occurrence of reactive collisions. It is immediate to recognize that the mean field approach to chemical kinetics corresponds exactly to the "stosszahlansatz" (the hypothesis of molecular chaos) enforced by Boltzmann in deriving his kinetic equation for non-equilibrium processes in diluted gaseous systems.

The latter analogy is indeed highlighting as it indicates the possibility of gathering together momentum-transfer processes, controlling molecular motion, with chemical/quantum events associated with chemical reactions and quantum transitions in a unitary stochastic event-based theory.

The stochastic theory of chemical reactions dates back to the formulation of the chemical master equation developed in order to handle reactive systems in which the granularity (limited and small number of molecules) plays a leading role and controls the statistical fluctuations in small reacting systems.

It is well known that, in all the cases where the number of molecules is small (and this occurs in subcellular biochemical reactions, in nanoscale systems, or in the growth kinetics of microorganisms [23–25]), the effects of fluctuations become significant, motivating a stochastic description of chemical kinetic processes, involving the number of molecules present in the system, thus explicitly accounting for their finite number [26–29]. The statistical theory of chemical kinetics in these conditions is grounded on the Chemical Master Equation (CME) [30,31], expressing the time evolution for the probabilities $p(\mathbf{N}, t)$ of all the possible number-configurations $\mathbf{N}(t) = (N_1(t), \ldots, N_s(t))$, where $N_h(t)$ is the number of molecules of the h-th reacting species at time t, $h = 1, \ldots, s$. However, apart from a handful of simple cases, for which the CME can be solved analytically [32], numerical methods should be applied to it in order to compute mean values and higher-order moments. But also this choice reveals itself to be unfeasible in most of the situations of practical and theoretical interest, due to the extremely large number of configurations involved, making the multi-index matrix $p(\mathbf{N}, t)$ so huge to exceed reasonable computational facilities.

In order to solve this problem, Gillespie proposed an algorithmic solution to the numerical simulation of stochastic reacting systems, based on the Markovian nature of the reactive events [33,34]. The original Gillespie algorithm has been extended and improved over time, providing a variety of slightly different computational alternatives. A common denominator of the first family of the Gillespie algorithms (namely those based on the direct method, the first reaction method or their derivates [35–37]) is to associate to every time step the occurrence of just one reaction. This formulation comes directly from the assumption that, if the time step is small enough, the probability that more than one reaction will occur is negligible. While correct, this choice brings to significant computational costs for complex reaction schemes. This problem has been highlighted several times, by the Gillespie group itself, as *stiffness* in stochastic chemical reacting systems [38]. A brilliant way to overcome this limit has been the famous tau-leaping method, which, unfortunately, requires to check that the propensity functions (i.e., the probabilities of occurrence of the reactive steps) remain *almost constant* at each iteration and can be applied just if this condition is verified [39,40]. The algorithmic solution

associated with the formalism here introduced combines the accuracy of the first stochastic simulation algorithms with the computational advantages of the tau-leaping method.

There is, moreover, a missing link between CME theory and the Gillespie algorithm, consisting in the straight mathematical formulation of the stochastic differential equations associated with a chemical reacting system, the statistical description of which would correspond to the CME. To clarify this issue, consider the conceptually analogous problem of particle diffusion over the real line, the statistical description of which is expressed by the parabolic equation $\partial p(x,t)/\partial t = D\,\partial^2 p(x,t)/\partial x^2$, for the probability density $p(x,t)$ of finding a particle at position x at time t. Setting $x_n = x(n\,\Delta t)$, an algorithm describing this process can be simply expressed by the discrete evolution equation $x_{n+1} = x_n + \sqrt{2\,D\,\Delta t}\,r_{n+1}$, where r_h, $h = 1, 2, \ldots$ represent independent random variables sampled from a normal distribution (with zero mean, and unit variance) [41]. This represents an efficient algorithmic solution of the problem, whenever the time resolution Δt is small enough. Nevertheless, the mere algorithmic approach cannot be considered physically satisfactory, in a comprehensive formulation of transport theory embedded in a continuous space-time (in which both position x and time t are real valued). In point of fact, only with the mathematical formulation due to K. Ito of stochastic differential equations driven by the increments $dw(t)$ of a Wiener process (Langevin equations) [42], namely $dx(t) = \sqrt{2\,D}\,dw(t)$ the theory of diffusive motion has found a proper mathematical physical setting.

A similar situation applies to the case of stochastic models of chemical reaction kinetics, for which up to now, the fundamental representation in terms of stochastic differential equations is missing (see e.g. the review [43]).

4. Stochastic models for chemical reactions

As mentioned above, this article is aimed at overcoming the purely algorithmic description of stochastic chemical reactions, providing the formulation of these processes as stochastic differential equations evolving continuously with time.

The basic idea is that any reactive process corresponds to a system of elementary events (the single reaction) possessing a Markovian transitional structure, and consequently, amenable to a description by means of the increments of counting processes (Poisson processes, in the Markovian case). This topic has been also pointed out in [44] in terms of Poisson measures, although the latter formulation is much less simple and physically intuitive than the approach proposed in the present article.

To begin with, consider the simple case of a first-order chemical reaction $A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} B$ (for instance, an isomerization). This model is perfectly analogous to the radiative transition of a molecule possessing two energy states, due to emission and adsorption of an energy quantum (figure 2). Let $N_A(0) + N_B(0) = N_g$ the total number of molecules at time $t = 0$. The state of the system is characterized by the state functions $\sigma_h(t)$, $h = 1, \ldots, N_g$ for each molecule, attaining values $\{0, 1\}$, and such that $\sigma_h(t) = 0$ if the energy state at time t is E_0 (or equivalently if the molecule finds itself in the state A), and $\sigma_h(t) = 1$ in the opposite case (energy state E_1, or isomeric state B).
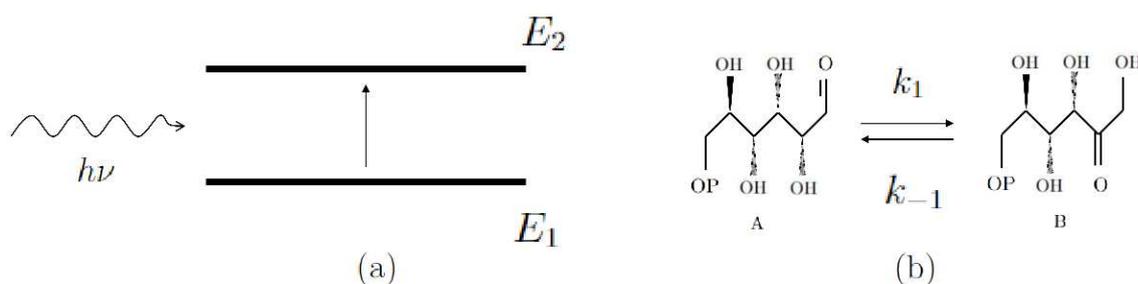

Figure 2. Schematic representation of the analogy between a two-level quantum system and a first-order chemical kinetics, such as an isomerization.

Let $\{\chi_h^{(1)}(t, k_1), \chi_h^{(2)}(t, k_{-1})\}_{h=1}^{N_g}$ be two systems of independent Poisson processes, characterized by the transition rates k_1, and k_{-1}, respectively. The evolution of $\sigma_h(t)$ can be expressed via the stochastic differential equation

$$\frac{d\sigma_h(t)}{dt} = (1 - \sigma_h(t)) \frac{d\chi_h^{(1)}(t, k_1)}{dt} - \sigma_h(t) \frac{d\chi_h^{(2)}(t, k_{-1})}{dt} \tag{10}$$

$h = 1, \ldots, N_g$, where $d\chi(t, \lambda)/dt$ is the distributional derivative of the Poisson process $\chi(t, \lambda)$, corresponding to a sequence of unit impulsive functions at the transition instants t_i^*, $i = 1, 2, \ldots$, $0 < t_i^* < t_{i+1}^*$, where for $\varepsilon > 0$, $\lim_{\varepsilon \to 0} \int_{t_i^* - \varepsilon}^{t_i^* + \varepsilon} d\chi(t, \lambda) = 1$. Summing over $h = 1, \ldots N_g$, and observing that $N_A(t) = \sum_{h=1}^{N_g} (1 - \sigma_h(t))$, $N_B(t) = \sum_{h=1}^{N_g} \sigma_h(t)$, we have

$$\frac{dN_B(t)}{dt} = \sum_{h=1}^{N_A(t)} \frac{d\chi_h^{(1)}(t, k_1)}{dt} - \sum_{h=1}^{N_B(t)} \frac{d\chi_h^{(2)}(t, k_{-1})}{dt} \tag{11}$$

and $dN_A(t)/dt = -dN_B(t)/dt$, representing the evolution equation for $N_A(t)$ and $N_B(t)$, attaining integer values. The stochastic evolution of the number of molecules $N_A(t)$, $N_B(t)$ is thus expressed as a differential equation with respect to the continuous physical time $t \in \mathbb{R}^+$, over the increments of a Poisson process. Interpreted in a mean-field way, if c_{tot} is the overall concentration of the reactants at time $t = 0$, then the concentrations $c_\alpha(t)$, $\alpha = A, B$, at time t can be recovered from eq. (11) as

$$c_\alpha(t) = c_{\text{tot}} \frac{N_\alpha(t)}{N_g}, \qquad \alpha = A, B \tag{12}$$

representing the calibration relation connecting the stochastic description, in terms of the number of molecules $N_\alpha(t)$, and the concentrations $c_\alpha(t)$, $\alpha = A, B$ entering the mean-field description.

The analytical formulation of a stochastic differential equation for chemical kinetics, expressed in terms of the number of molecules of the chemical species involved, rather than an algorithm defined for discretized times, allows the development of a variety of different numerical strategies, that naturally perform a modified tau-leaping procedure, as the occurrence of several distinct reactive events in any elementary time step Δt is intrinsically accounted for. This can be easily seen by considering the simple reaction defined by the evolution equation (11). In terms of increments, eq. (11) can be written as $dN_B(t) = \sum_{h=1}^{N_A(t)} d\chi^{(1)}(t, k_1) - \sum_{h=1}^{N_B(t)} d\chi^{(2)}(t, k_{-1})$. If Δt is the chosen time step, it follows from this formulation, a simple numerical approximation for eq. (11), namely,

$$\Delta N_B(t) = N_B(t + \Delta t) - N_B(t) = \sum_{h=1}^{N_A(t)} \xi_h^{(1)}(k_1 \Delta t) - \sum_{h=1}^{N_B(t)} \xi_h^{(2)}(k_{-1} \Delta t) \tag{13}$$

where $\xi^{(1)}(k_1 \Delta t)$, $\xi_h^{(2)}(k_{-1} \Delta t)$ $h = 1, 2, \ldots$, are two families of independent binary random variables, where

$$\xi_h^{(\alpha)}(p) = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

$\alpha = 1, 2$, $h = 1, 2, \ldots$. The time step Δt, can be chosen in eq. (13) from the condition

$$K \Delta t < 1, \qquad K = \max\{k_1, k_{-1}\} \tag{15}$$

In practice, we choose $\Delta t = 0.1/K$. As can be observed, the choice of Δt is limited only by the intrinsic rates of the process. The advantage of deriving different algorithmic schemes for solving numerically the stochastic kinetic equations becomes more evident in dealing with bimolecular reactions (addressed below) [45].

The same approach can be extended to include amongst the elementary events not only the reactive steps, but also feeding conditions, thus representing the evolution of chemically reacting systems with a finite number of molecules in a perfectly stirred open reactor. This is the case of the tank-loading problem, in which a tracer is injected in an open vessel, assumed perfectly mixed, for which, in the absence of chemical reactions, the mean field equation for the concentration of the tracer reads

$$\frac{dc(t)}{dt} = D\ (c_0 - c(t)) \tag{16}$$

where c_0 is the inlet concentration and D the dilution rate (reciprocal of the mean retention time), and $c(0) = 0$. Fixing N_g so that $c(t) = c_0\, N(t)/N_g$, the corresponding stochastic differential equation for the integer $N(t)$ involves, also in this case, two families of counting processes, one for the loading at constant concentration c_0, and the other for tracer discharge in the outlet stream, characterized by the same transition rate D,

$$\frac{dN(t)}{dt} = \sum_{h=1}^{N_g} \frac{d\chi_h^{(1)}(t,D)}{dt} - \sum_{k=1}^{N(t)} \frac{d\chi_h^{(2)}(t,D)}{dt} \tag{17}$$

starting from $N(0) = 0$. Figure 3 depicts several realizations of the tank-loading process, obtained by discretizing eq. (17) with a time step $\Delta t = 10^{-3}$.
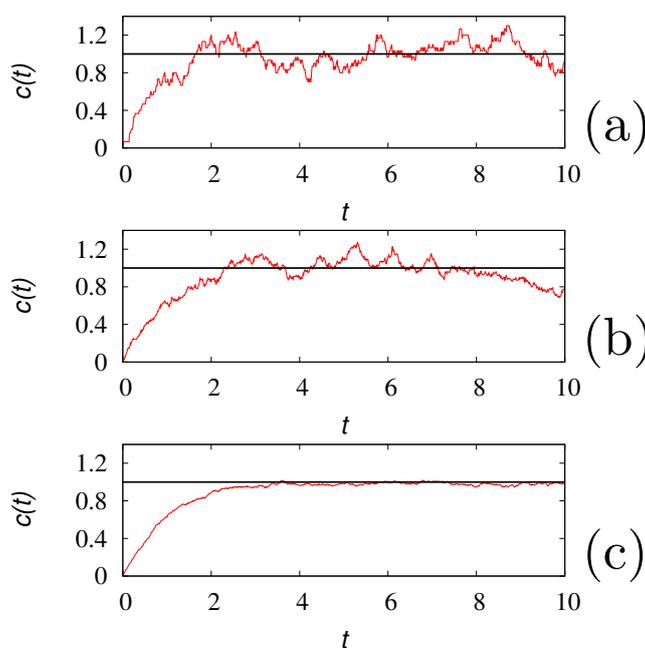

Figure 3. $c(t) = N(t)/N_g$ vs t from a single realization of the tank-loading process eq. (17) with $D = 1$, $c_0 = 1$ a.u.. Panel (a): $N_g = 30$, panel (b) $N_g = 100$, panel (c) $N_g = 1000$. The solid horizontal lines represent the steady-state value $c^* = 1$.

This approach can be extended to the analysis of chemical kinetics in flow systems (specifically microfluidic devices) for aggregation and gelation processes, that possess persistent chemical fluctuations, see e.g. [46].

Despite the simplicity of eqs. (16) or (17), this process permits to highlight the role of N_g, that can be referred to as the *granularity number*, and the way stochastic models of chemical reactions can be fruitfully applied. Indeed, there is a two-fold use of the stochastic formulation of chemical kinetic schemes. The first refers to a chemical reacting system involving a small number of molecules, and in this case N_g represents the effective number of molecules present in the system. The other is to use stochastic algorithms for simulating reacting systems as an alternative way, sometimes more efficient, with respect to the solution of the corresponding mean-field equations. In the latter case,

the granularity number N_g represents essentially a computational parameter, tuning the intensity of the fluctuations. Two choices are then possible: (i) N_g can be chosen large enough, in order to obtain from a single realization of the process an accurate approximation of the mean-field behavior, or (ii) it can be chosen small enough in order, to deal with extremely fast simulations of a single realization of the process, that could be averaged over a statistically significant number of realizations in due time. These two choices are depicted in figure 3 (panel c), choosing $N_g = 10^3$, and in figure 4 panel (a) obtained for $N_g = 30$. Of course, the latter approach is valid as long as the low-granularity (low values of N_g) does not influence the qualitative properties of the kinetics.
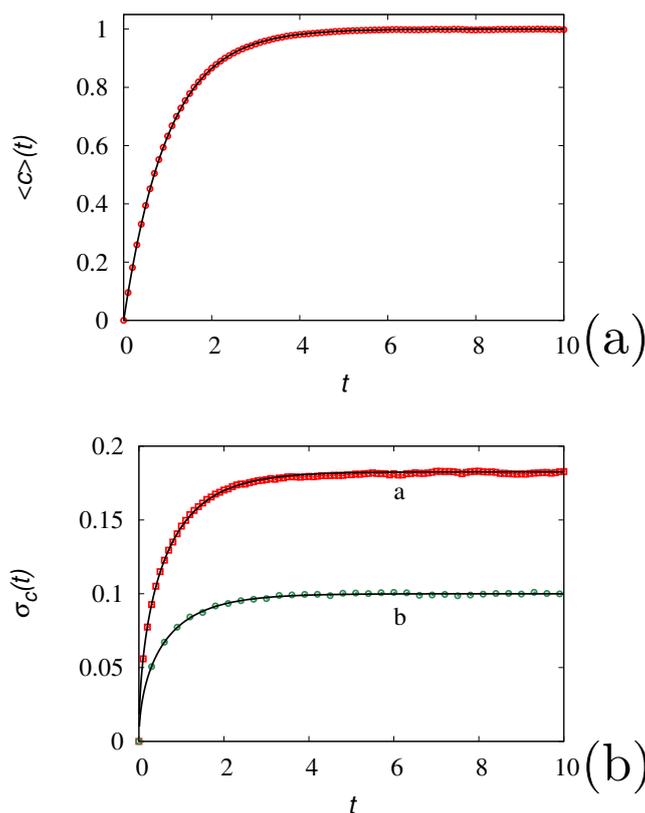

Figure 4. Panel (a): $\langle c \rangle (t)$ vs t at $N_g = 30$ (symbols) averaged over $[10^6/N_g]$ realizations of the tank-loading process with $D = 1$, $c_0 = 1$ a.u. Here, $[\cdot]$ indicates the integer part of its argument. The solid line represents the mean-field result $\langle c \rangle (t) = 1 - e^{-t}$. Panel (b): Variance $\sigma_c(t)$ vs t for the tank-loading process. Symbols are the results of stochastic simulations of eq. (17) averaged over $[10^6/N_g]$ realizations, lines the solutions of eq. (19). Line (a) refers to $N_g = 30$, line (b) to $N_g = 100$.

The second (computational) use of stochastic simulations of chemical kinetics requires a further discussion. At a first sight, it may appear that any stochastic simulation would be computationally less efficient than the solution of the corresponding mean-field equations. This is certainly true for classical chemical reaction schemes in a perfectly mixed system, for which the mean-field model reduces to a system of ordinary differential equations for the concentrations of the reactants. But there are kinetic problems e.g., associated with the growth of microorganisms and eukaryotic cell lines in bioreactors (these growth phenomena, are indeed amenable to a description in terms of equivalent chemical reactions), the mean-field model of which is expressed in the form of higher-dimensional nonlinear integro-differential equations . For this class of problems, the use of stochastic simulations is the most efficient, if not the only way to achieve a quantitative description of the process, in those cases where the number n_p of internal parameters describing the physiological state of an eukaryotic cell becomes large enough [45]. This case is altogether similar to some transport problems, such as Taylor-Aris dispersion for high Péclet numbers or the analysis of microfluidic separation processes (DLD devices)

for which the stochastic simulation of particle motion is far more efficient that the corresponding solution of the corresponding mean-field model expressed in the form of advection-diffusion equations [47,48].

To complete the analysis of the tank-loading problem, the associated CME reads

$$\frac{dp(n,t)}{dt} = D\,N_g\,[p(n-1,t)\,\eta_{n-1} - p(n,t)] + D\,[(n+1)\,p(n+1,t) - n\,p(n,t)] \tag{18}$$

where $\eta_h = 1$ for $h \geq 0$ and $\eta_h = 0$ otherwise. It follows that $\langle c \rangle(t) = c_0 \sum_{n=1}^{\infty} n\,p(n,t)/N_g$ satisfies identically the mean-field equation (due to the linearity of the problem), while the variance $\sigma_c(t)$, with $\sigma_c^2(t) = c_0^2 \sum_{n=1}^{\infty} n^2\,p(n,t)/N_g^2 - \left(c_0 \sum_{n=1}^{\infty} n\,p(n,t)/N_g\right)^2$, satisfies the equation

$$\frac{d\sigma_c^2}{dt} = -2\,D\,\sigma_c^2 + D\left(\frac{1}{N_g} + \frac{\langle c \rangle}{N_g}\right) \tag{19}$$

Figure 4 panel (b) compares the results of stochastic simulations against the solutions of eq. (19) for two values of N_g.

The above approach can be extended to any system of nonlinear reaction schemes involving unimolecular and bimolecular reaction, and in the presence of slow/fast kinetics. The structure of the reaction mechanism can be arbitrarily complicated without adding any further complexity (other than purely notational) in the formulation of the stochastic evolution expressed in terms of number of molecules. The only practical issue, is that the number of different families of stochastic processes grows with the number of elementary reactive processes considered. For instance, in the case of the substrate-inhibited Michaelis-Menten kinetics

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES$$
$$ES \overset{k_2}{\rightarrow} E + P \tag{20}$$
$$ES + S \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} ESS$$

there are five reactive processes (five channels in the language of the Gillespie algorithm) and consequently five families of counting processes $\{\chi_{i_h}^{(h)}(t,\cdot)\}$, $h = 1,\ldots,5$, should be introduced, so that the formulation of the discrete stochastic dynamics reads

$$\frac{dN_S(t)}{dt} = -\sum_{i=1}^{N_S(t)} \frac{d\chi_i^{(1)}(t,\widetilde{k}_1\,N_E(t))}{dt} + \sum_{j=1}^{N_{ES}(t)} \frac{d\chi_j^{(2)}(t,k_{-1})}{dt}$$

$$\frac{dN_E(t)}{dt} = -\sum_{i=1}^{N_S(t)} \frac{d\chi_i^{(1)}(t,\widetilde{k}_1\,N_E(t))}{dt} + \sum_{j=1}^{N_{ES}(t)} \frac{d\chi_j^{(2)}(t,k_{-1})}{dt} + \sum_{h=1}^{N_{ES}(t)} \frac{d\chi_h^{(3)}(t,k_2)}{dt}$$

$$\frac{dN_{ES}(t)}{dt} = \sum_{i=1}^{N_S(t)} \frac{d\chi_i^{(1)}(t,\widetilde{k}_1\,N_E(t))}{dt} - \sum_{j=1}^{N_{ES}(t)} \frac{d\chi_j^{(2)}(t,k_{-1})}{dt} - \sum_{h=1}^{N_{ES}(t)} \frac{d\chi_h^{(3)}(t,k_2)}{dt}$$
$$\quad - \sum_{k=1}^{N_S(t)} \frac{d\chi_k^{(4)}(t,\widetilde{k}_3\,N_{ES}(t))}{dt} + \sum_{l=1}^{N_{ESS}(t)} \frac{d\chi_l^{(5)}(t,k_{-3})}{dt} \tag{21}$$

$$\frac{dN_{ESS}(t)}{dt} = \sum_{k=1}^{N_S(t)} \frac{d\chi_k^{(4)}(t,\widetilde{k}_3\,N_{ES}(t))}{dt} - \sum_{l=1}^{N_{ESS}(t)} \frac{d\chi_l^{(5)}(t,k_{-3})}{dt}$$

$$\frac{dN_P(t)}{dt} = \sum_{h=1}^{N_{ES}(t)} \frac{d\chi_h^{(3)}(t,k_2)}{dt}$$

equipped with the initial conditions $c_S(0) = c_{S,0}, c_E(0) = c_{E,0}, c_{ES}(0) = c_{ESS}(0) = c_P(0) = 0$. Observe that for the bimolecular steps we have used a number-dependent rate coefficient.

The granularity number N_g can be fixed, so that

$$N_S(0) = [c_{S,0} N_g], \qquad N_{E,0} = [c_{E,0} N_g] \tag{22}$$

where $[\xi]$ indicates the integer part of ξ, thus defining the relation between $N_\alpha(t)$ and $c_\alpha(t)$, namely $c_\alpha(t) = N_\alpha(t)/N_g$, $\alpha = S, E, ES, ESS, P$. This implies also that the effective rate parameters entering the discrete stochastic evolution equation (21), and associated with the two bimolecular reactive steps, are given by $\widetilde{k}_1 = k_1/N_g$, and $\widetilde{k}_3 = k_3/N_g$.

Consider the case $k_{-1} = k_2 = k_3 = k_{-3} = 1$, $c_{S,0} = 4$, $c_{E,0} = 0.1$. In this case the quasi steady-state approximation of the c_{ES}-c_S curve (representing the slow manifold of the kinetics) takes the expression

$$c_{ES} = \frac{c_{E,0} \, c_S}{K_M + c_S + \beta \, c_S^2}, \qquad K_M = \frac{k_{-1} + k_2}{k_1}, \qquad \beta = \frac{k_{-3}}{k_3} \tag{23}$$

Figure 5 depicts the c_{ES}-c_S graph obtained from a single realization of the stochastic process eq. (21) at several values of k_1 so as to modify the Michaelis-Menten constant K_M for a value $N_g = 10^6$ of the granularity number.

Apart from the initial transient giving rise to an overshot in the values of c_{ES} near $c_S \simeq c_{S,0}$, the dynamics rapidly collapses towards the slow manifold and the stochastic simulations at high N_g-value provide a reliable description of the mean-field behavior starting from a single stochastic realization.
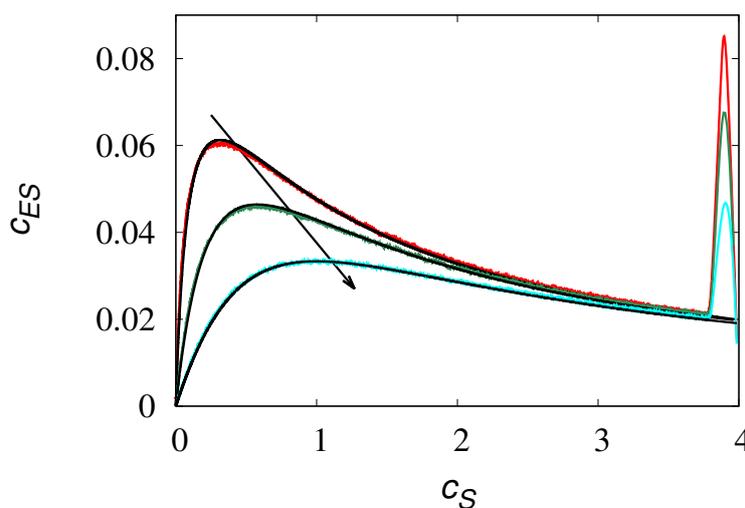

Figure 5. c_{ES} vs c_S plot of the substrate-inhibited enzymatic kinetics discussed in the main text. Symbols (in color) are the results of stochastic simulations of a single realization of the process eq. (21), (black) solid lines the graph of the quasi steady-state approximation. The arrow indicates increasing values of K_M, i.e. decreasing values of $k_1 = 20, 6, 2$.

5. Extensions: clustering of liquid molecules and velocity fluctuations

The mechanism controlling fluctuations in chemical kinetics can be extended to any kind of chemical physical process in which some properties change discontinuously as a consequence of elementary events. This is the case of aggregation/fragmentation phenomena, adsorption and quantum transitions. In this Section we consider a particular case of aggregation, related to the clustering of molecules in liquids such as water due to weak interactions [49–51]. The dynamics of creation and destruction of clusters is controlled by weak forces and hydrogen-bond formation. This determines that the effective mean square velocity of water molecules at constant temperature T is not $k_B T/m$ (where m is the mass of a water molecule), but rather $\langle v^2 \rangle = k_B T/m_{\text{eff}}$, with $m_{\text{eff}} > m$, due to

clustering. Therefore, cluster kinetics has a relevant effect not only at a molecular level but also on the thermodynamic and transport properties of the solution.

Let us assume that in the clustering process only one molecule at a time is either incorporated in a cluster or eliminated from it, as depicted in figure 6, and that clustering dynamics can be regarded as linear, i.e. with rates independent of the concentration of the molecules. The process is defined by two parameters: the rate λ at which a cluster undergoes a transition, and the probability α that this transition corresponds to the acquisition of another molecule. Moreover, let us assume that there is a maximum number of molecules N above which no cluster can be formed. Therefore, for each molecule, we can define a state variable $\sigma_h(t)$, attaining values $1, 2, \ldots, N$, specifying the cluster-state of the molecule: $\sigma_h(t)$ corresponds to the size of the cluster to which the molecule belongs to. The dynamics of the cluster formation is thus described by the stochastic evolution equation for $\sigma_h(t)$,

$$\frac{d\sigma_h(t)}{dt} = \frac{d\chi^{(1)}(t, \lambda^{(1)}_{\sigma_h(t)})}{dt} - \frac{d\chi^{(2)}(t, \lambda^{(2)}_{\sigma_h(t)})}{dt} \tag{24}$$

where

$$\begin{aligned}
\lambda^{(1)}_1 &= \lambda, \qquad \lambda^{(1)}_h = \lambda\,(1-\alpha)\,(1-\delta_{h,N}), \quad h = 2, \ldots, N \\
\lambda^{(2)}_h &= \lambda\,\alpha\,(1-\delta_{h,1}), \quad h = 1, \ldots, N-1, \qquad \lambda^{(2)}_N = \lambda
\end{aligned} \tag{25}$$

Indicating with $P_h(t)$, $h = 1, \ldots, N$ the probability that a molecule belongs at time t to a cluster of h elements, we have

$$\begin{aligned}
\frac{dP_1}{dt} &= -\lambda\,P_1 + \alpha\,\lambda\,P_2 \\
\frac{dP_2}{dt} &= \lambda\,P_1 - \lambda\,P_2 + \alpha\,\lambda\,P_3 \\
\frac{dP_k}{dt} &= (1-\alpha)\,\lambda\,P_{k-1} - \lambda\,P_k + \alpha\,\lambda\,P_{k+1}, \qquad k = 3, \ldots, N-2 \\
\frac{dP_{N-1}}{dt} &= (1-\alpha)\,\lambda\,P_{N-1} - \lambda\,P_{N-1} + \lambda\,P_N \\
\frac{dP_N}{dt} &= (1-\alpha)\,\lambda P_{N-1} - \lambda\,P_N
\end{aligned} \tag{26}$$

Equations (26) represent a particular case of the Becker-Döring nucleation kinetics [52–54], the matrix representation of which attains the form

$$\frac{dP_k}{dt} = \sum_{j=1}^{N} \Lambda_{k,j}\,P_j, \qquad k = 1, \ldots, N \tag{27}$$

The stationary equilibrium distribution P_k^* is given by

$$P_k^* = C, \qquad P_k^* = C\,\frac{(1-\alpha)^{k-2}}{\alpha^{k-1}}, \; k = 2, \ldots, N_1, \qquad P_N^* = C\,\frac{(1-\alpha)^{N-2}}{\alpha^{N-2}} \tag{28}$$

where C is the normalization constant such that $\sum_{k=1}^{N} P_k^* = 1$. Figure 7 depicts the comparison between stochastic simulations using an ensemble of $N_p = 10^5$ realizations of eq. (24) and the steady-sate solution eq. (28).
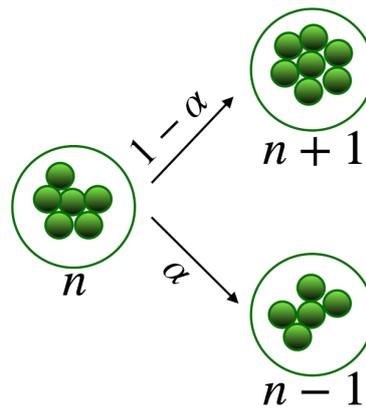
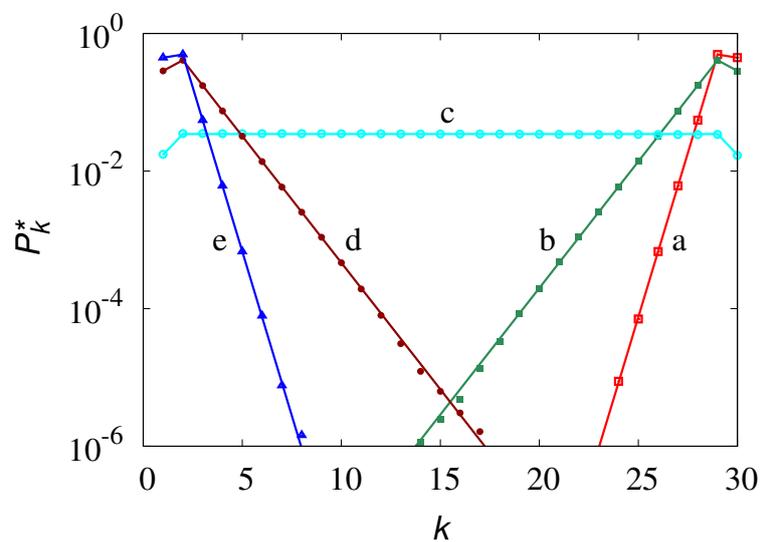**Figure 6.** Schematic representation of the clustering model.

Figure 7. Equilibrium P_k^* vs k obtained from stochastic simulations of eq. (24) at different values of α, $N = 30$. Symbols refer to stochastic simulations, lines to eq. (28): (a) $\alpha = 0.1$, (b) $\alpha = 0.3$, (c) $\alpha = 0.5$, (d) $\alpha = 0.7$, (e) $\alpha = 0.9$.

The microscopic phenomenon of coalescence impacts on macroscopic transport properties, as the motion of clusters is controlled by its size, both as regards inertial (mass) and hydrodynamic interactions (friction factor). If we indicate with m_k and η_k the mass and the friction factor of a cluster of k molecules, the velocity v_k of a k-cluster satisfies the Langevin equation

$$m_k \, dv_k = -\eta_k \, v_k \, dt + \sqrt{2 \, k_B \, T \, \eta_k} \, dw_k(t) \tag{29}$$

where $w_k(t)$, $k = 1, \ldots, N$ are a family of Wiener processes independent of each other. We use here a scalar representation for the velocity as the vector representation does not add any new physics, making the notation more cumbersome. The masses m_k and the friction factors η_k depend on k. Since $m_k = m \, k \sim R_k^3$, where R_k is the gyration radius of the k-cluster, and $\eta_k \sim R_k$, we can set

$$\eta_k = \eta_0 \, k^{1/3} \tag{30}$$

where $\eta_0 > 0$ is a constant. Let $P_{v,k}(v, t)$ be the joint density function for v and $\sigma_i(t)$ for the generic i-th molecule

$$P_{v,k}(v', t)dv = \text{Prob} \left[v(t) \in (v', v' + dv), \sigma_i(t) = k \right] \tag{31}$$

The system of partial densities $P_{v,k}(v,t)$ satisfies the family of Fokker-Planck equations

$$\frac{\partial P_{v,k}(v,t)}{\partial t} = \frac{\partial}{\partial v}\left(\frac{\eta_k\, v}{m_k} P_{v,k}(v,t)\right) + \frac{k_B\, T\, \eta_k}{m_k^2}\frac{\partial^2 P_{v,k}(v,t)}{\partial v^2} + \sum_{j=1}^{N} \Lambda_{k,j}\, P_j(v,t) \tag{32}$$

$k = 1,\ldots,N$. The average value of any function $f(v)$ of the velocity at time t is thus given by

$$\langle f(v(t))\rangle = \sum_{k=1}^{N} \langle f(v_k(t))\rangle = \sum_{k=1}^{N}\int_{-\infty}^{\infty} f(v)\, P_{v,k}(v,t)\, dv \tag{33}$$

Of particular interest is the value of $\langle v^2(t)\rangle$. From eq. (32) it follows that $\langle v_k^2(t)\rangle$, $k = 1,\ldots,N$, satisfy the equations

$$\frac{d\langle v_k^2\rangle}{dt} = -\frac{2\,\eta_k}{m_k}\langle v_k^2\rangle + \frac{2\,k_B\,T\,\eta_k}{m_k^2} P_k + \sum_{j=1}^{N}\Lambda_{k,j}\langle v_j^2\rangle \tag{34}$$

$k = 1,\ldots,N$, where $P_k(t)$ are the marginal densities with respect to σ_i, solution of eq. (27). Figure 8 depicts the evolution for $\langle v^2(t)\rangle$ obtained from stochastic simulations with $N_p = 10^5$ elements, initially in the form of single molecules $\sigma_i(0) = 1$, $i = 1,\ldots,N_p$, with velocities sampled from the equilibrium distribution. Throughout these simulations we set $N = 30$, $k_B\,T=1$ a.u., $m = 1$ a.u. and $\eta_0 = 1$ a.u., as the values of these parameters can be rescaled out and they do not influence the qualitative properties of the system. Correspondingly, the initial velocities for the molecules are sampled from a normal distribution with zero mean and unit variance. The data depicted in figure 8 refers to two values of the parameter α.
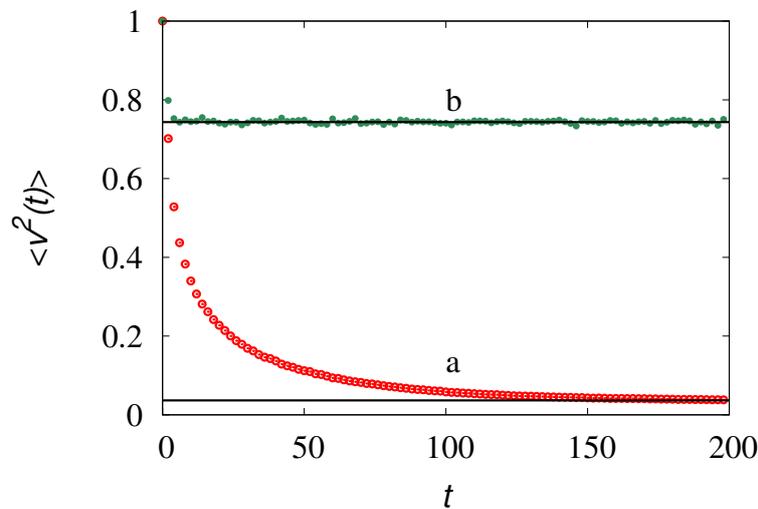

Figure 8. $\langle v^2(t)\rangle$ vs t at $\lambda = 1$ for two values of α, $N = 30$. Line (a) refers to $\alpha = 0.4$, line (b) to $\alpha = 0.9$. Symbols correspond to the results of stochastic simulations, solid horizontal lines to the steady values obtained from the solution of eq. (34) at steady state.

Of major interest are the long-term equilibrium properties, corresponding to the steady solution of eq. (34). The squared velocity variance $\langle v^2\rangle$ is given by the sum of the marginal squared variances

$$\langle v^2\rangle = \sum_{k=1}^{N}\langle v_k^2\rangle \tag{35}$$

Set $\Lambda_{k,j} = \lambda \, \widetilde{\Lambda}_{k,j}$, where $\widetilde{\Lambda}_{k,j} \sim O(1)$. Two limit regimes occur. If $\lambda \ll 1$, the recombination term (last term at the r.h.s. of eq. (34)) can be neglected and thus the equilibrium values for the squared velocity variances are given by

$$\langle v_k^2 \rangle = \frac{k_B \, T}{m_k} \, P_k^* \,, \qquad k = 1, \ldots, N \tag{36}$$

where P_k^* are the equilibrium fractions eq. (28). Conversely, if $\lambda \gg 1$, the recombination term prevails, and the squared partial variances approach the solution of the equation

$$\sum_{j=1}^N \widehat{\Lambda}_{k,j} \, \langle v_j^2 \rangle = 0 \tag{37}$$

This means that the $\langle v_k^2 \rangle$ should be proportional to the equilibrium fractions

$$\langle v_k^2 \rangle = c \, P_k^* \tag{38}$$

where the proportionality constant $c > 0$ can be determined by substituting this expression into eq. (34) at steady-state, from which we obtain

$$c = k_B \, T \, \frac{\sum_{k=1}^N \frac{\eta_k}{m_k^2} \, P_k^*}{\sum_{k=1}^N \frac{\eta_k}{m_k} \, P_k^*} \tag{39}$$

Therefore, $\langle v^2 \rangle$ attains the two limit values

$$\lim_{\lambda \to 0} \langle v^2 \rangle = \langle v^2 \rangle_0 \,, \qquad \lim_{\lambda \to \infty} \langle v^2 \rangle = \langle v^2 \rangle_\infty \tag{40}$$

where

$$\langle v^2 \rangle_0 = k_B \, T \, \sum_{k=1}^N \frac{P_k^*}{m_k} \,, \qquad \langle v^2 \rangle_\infty = k_B \, T \, \frac{\sum_{k=1}^N \frac{\eta_k \, P_k^*}{m_k^2}}{\sum_{k=1}^N \frac{\eta_k \, P_k^*}{m_k}} \tag{41}$$

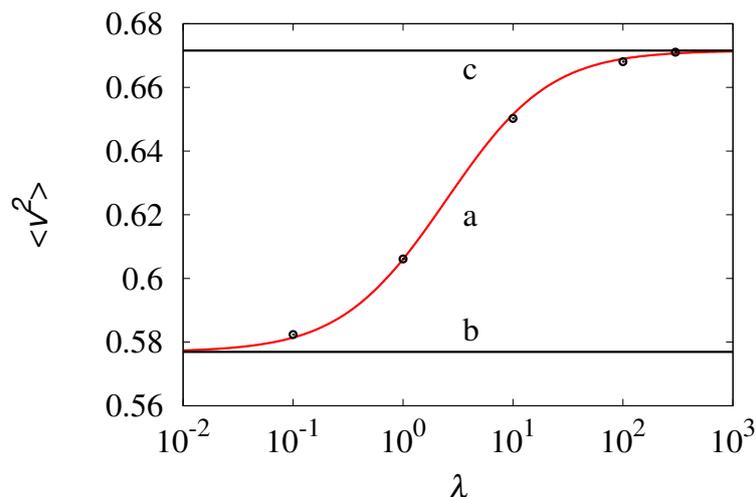The occurrence of these two limit regimes is depicted in Figures 9 and 10.

Figure 9. Equilibrium value $\langle v^2 \rangle$ vs λ at $\alpha = 0.7$ (line (a) and symbols). Symbols represent the results of stochastic simulations, line (a) stems from the steady solution of eq. (28). Horizontal lines (b) and (c) correspond to the limit values $\langle v^2 \rangle_0$, $\langle v^2 \rangle_\infty$, respectively, defined by eq. (41).
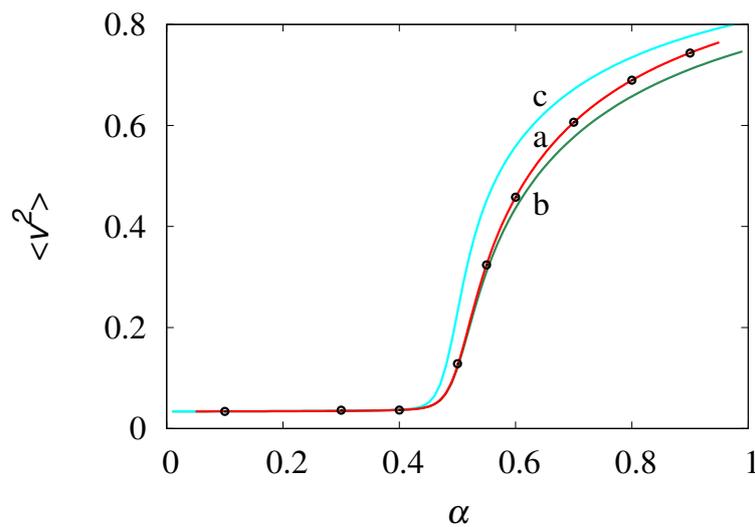
Figure 10. Equilibrium value $\langle v^2 \rangle$ vs α. Line (a) and symbols refer to $\lambda = 1$, and corresponds to the results of stochastic simulations, and of the solution of eq. (28), respectively. Lines (b) and (c) represent the behavior of $\langle v^2 \rangle_0$ and $\langle v^2 \rangle_\infty$ vs α, respectively.

Observe that for low values of α ($\alpha \leq 0.4$) the influence of the rate λ on the intensity of the velocity fluctuations is practically negligible, and $\langle v^2 \rangle_\infty \simeq \langle v^2 \rangle_0$. This condition corresponds to the formation of very large clusters, having the size of the order of N. Much higher variability occurs for values of $\alpha > 0.5$ approaching $\alpha = 1$ as this corresponds to the most active kinetic regime with the formation of smaller clusters of the size of few molecules, in which the value of the rate λ plays a leading role in determining the statistical properties of the velocity fluctuations.

It is also interesting to analyze the velocity density functions $P_v^*(v)$ at equilibrium. Figure 11 depicts the equilibrium densities in these two limit regimes, for values of α below and above the threshold $\alpha_c = 1/2$.
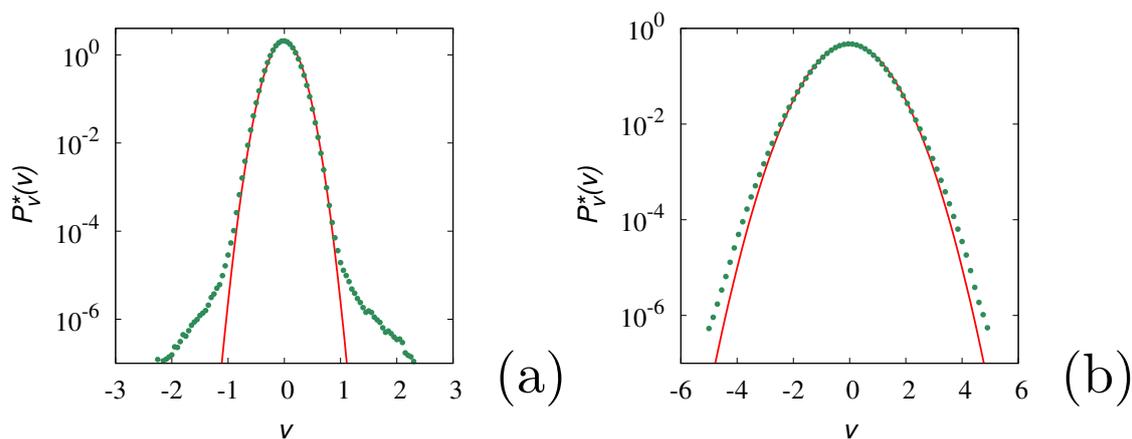

Figure 11. Equilibrium velocity distribution $P_v^*(v)$ vs v at $\lambda = 1$, $N = 30$, obtained from stochastic simulations (symbols). Panel (a) refers to $\alpha = 0.4$, panel (b) to $\alpha = 0.9$. The solid lines represent the Gaussian densities eq. (42) with zero mean and squared variance equal to that of the stochastic simulations.

The results of the stochastic simulations (symbols in figure 11) refer to a sample of $N_p = 10^6$ realizations. Once the stationary equilibrium conditions have been achieved, the densities are evaluated over a sufficiently wide time window to ensure the estimate of the velocity densities from a sample of

about 10^{11} realizations. This ensures that, at least for the range of P_v^*-values considered, the velocity statistics is accurate. The solid lines in this figure refer to the Maxwellian (Gaussian) distribution

$$g(v,\sigma) = \frac{1}{2\,\pi\,\sigma^2}\,e^{-v^2/2\,\sigma^2} \tag{42}$$

with zero mean and variance $\sigma^2 = \langle v^2 \rangle_{\text{stoch.simul.}}$ equal to that obtained from the stochastic simulations. It can be observed that for small $\alpha < 1/2$ (panel a), a main Gaussian core occurrs at low velocities, but significant deviations from Gaussianity appear in the high-velocity tails. This is non-surprising as the velocity fluctuations are mainly controlled by the formation of large molecular clusters, and the velocity tails account for the relatively small contribution of these clusters. Conversely, for $\alpha > 1/3$ (as in panel (b) at $\alpha = 0.9$) a slight, but uniform deviation from the Gaussian profile is observed throughout the whole range of v-values.

Next, consider the particle motion defined by the kinematics

$$dx_i = v_i\,dt \tag{43}$$

While in the analysis developed above the values of m and η_0 were immaterial, so that we set them equal to unity, in the description of particle motion they play a relevant role. Therefore, in the remainder let us assume that m and η_0 attain physical values, which implies for the characteristic dissipation time $t_{\text{diss}} = m/\eta_0 \ll 1$. Consequently, we can assume in eq. (29), the overdamped approximation even for the larger clusters. This means that eq. (43) becomes

$$dx_i = \sqrt{\frac{2\,k_B\,T}{\eta_i}}\,dw_i(t) \tag{44}$$

Assuming without loss of generality $x_i(0) = 0$, we have for the mean square displacement

$$\langle x^2(t) \rangle = 2\,k_B\,T\,\sum_{k=1}^{N}\frac{P_k^*}{\eta_k}\,t \tag{45}$$

providing, for the effective diffusion coefficient D the expression

$$D = k_B\,T\,\sum_{k=1}^{N}\frac{P_k^*}{\eta_k} = k_B\,T\,\langle \eta^{-1} \rangle \tag{46}$$

The latter expression for D should be compared with the effective friction factor $\langle \eta \rangle$,

$$\langle \eta \rangle = \sum_{k=1}^{N}\eta_k\,P_k^* \tag{47}$$

From the comparison of eqs. (46)-(47) it is natural to expect that the classical fluctuation-dissipation relation may be violated [55]. The order parameter χ for this phenomenon is the product

$$\chi = \langle \eta \rangle\,\langle \eta^{-1} \rangle \tag{48}$$

as $D\,\langle \eta \rangle = k_B\,T\,\chi$, and it should be strictly equal to 1 if the classical fluctuation-dissipation relation is verified. Observe that the averages entering eqs. (46)-(47) are evaluated for the equilibrium cluster distribution P_k^*.

Figure 12 depicts the behavior of $\chi(\alpha)$ vs α for a maximum cluster size $N = 30$. The order parameter is practically equal to 1 for small values of α, and the main departure from the fluctuation-dissipation relation (albeit of relative small entity, $|\chi - 1| < 0.1$), occurs near the critical point $\alpha_c = 1/2$. This result indicates how microscopic phenomena, such as coalescence, may impact

on macroscopic thermodynamic and transport properties, and that it is possible to provide simple and efficient stochastic modelling of them, accounting for the elementary transitions occurring at the molecular level.
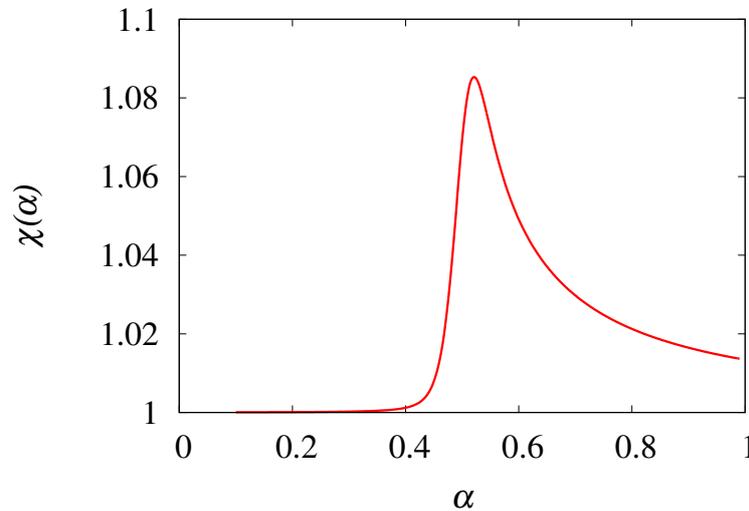

Figure 12. Order parameter $\chi(\alpha)$ eq. (48) vs α for $N = 30$.

6. Radiative interactions and thermalization

Chemical reaction kinetics and quantum transitions can be treated stochastically with the same Markovian setting, once the transition rates are defined starting from the microscopic physics of the elementary events. This is analyzed in detail in [56] to which the reader is referred for further details.

It is important to stress in this context, similarly to the case of coalescence dynamics analyzed in the previous Section, the influence of radiative processes on transport properties. It is well known that the electromagnetic field carries momentum, and consequently it is able to interact "at a mechanical level" with massive bodies, i.e. influencing their momentum dynamics (equations of motion). This problem has been analyzed in detail by Einstein [57], (the English translation can be found in [58], see also the discussion [59,60]), showing that, if a molecular system composed of identical molecules of mass m, and possessing an internal quantized energy structure, amenable to radiative transitions (i.e. absorption and emission of energy quanta), interacts with equilibrium thermal radiation at temperature T, the resulting momentum transfer due to photon kicks and molecular recoil effects determines statistically a squared velocity variance equal to $\langle v^2 \rangle = k_B T / m$ (componentwise), consistently with the kinetic theory.

This result is of the highest physical importance as it indicates that radiative interactions at the molecular level determine the thermalization of the system independently of other mechanical interactions (such as binary molecular collisions), which possibly may speed up the thermalization process, but not modify its outcome (for the simple reason that binary elastic collisions preserve the sum of the squared norm of velocities of the colliding particles).

The stochastic dynamics of this process has been analyzed in [56]. Due to the impulsive nature of the radiative events, at least in a semiclassical description, momentum-transfer dynamics can be represented by means of an impulsive Langevin equation

$$m \frac{d\mathbf{v}}{dt} = (-\eta \, \mathbf{v} + \mathbf{b}) \, \frac{d\chi(t,\lambda)}{dt} \tag{49}$$

where $\chi(t,\lambda)$ is a Poisson counting process with transition rate λ, accounting for the stochasticity and for the frequency of radiative events, $\mathbf{b} = b \, \mathbf{r}$, where $b = h\nu/c$ the momentum of the photon ($\nu = (E_2 - E_1)/h$ is the resonant frequency of the transition between the energy levels E_1 and $E_2 > E_1$,

h the Planck constant, and *c* the light speed *in vacuo*), **r** a random unit vector, uniformly distributed on the surface of the unit sphere (as thermal radiation is here considered), and η the radiative friction factor, possessing the physical dimension of a mass.

Integrating eq. (49) between the occurrence of a radiative event, we have

$$\mathbf{v}' = \alpha\,\mathbf{v} + \frac{b\,(1-\alpha)\,\mathbf{r}}{\eta} \tag{50}$$

where **v** and **v'** are the velocities before and after the event, $\alpha = e^{-\eta/m}$. The parameter α and η are related to each other by a radiative fluctuation-dissipation relation [56]

$$3\,\frac{k_B\,T}{m}\,(1-\alpha^2) = \frac{b^2\,(1-\alpha)^2}{\eta^2} \tag{51}$$

Also in this case, it is interesting to analyze the effects of this microscopic process on the transport properties, using the parameter $\alpha \in (0,1)$ as the quantity controlling radiative dissipation. If α is close to 1, and this physically occurs at "high" temperatures, above 10^{-5}-10^{-6} K, the resulting velocity density functions are of Gaussian shape, consistently with the Maxwellian theory. Conversely, at very low temperatures, say below 10^{-6} K, α is definitely smaller than 1, and deviations from Gaussian profiles become evident. This is depicted in figure 13 obtained by simulating an ensemble of 10^8 particles, undergoing the dynamics (50)-(51) with $k_B T/m$ set equal to 1 a.u., for two values of α, starting from an initially uniform velocity density function, and considering the equilibrium velocity profile $P_v^*(v)$ for a generic component of **v**.
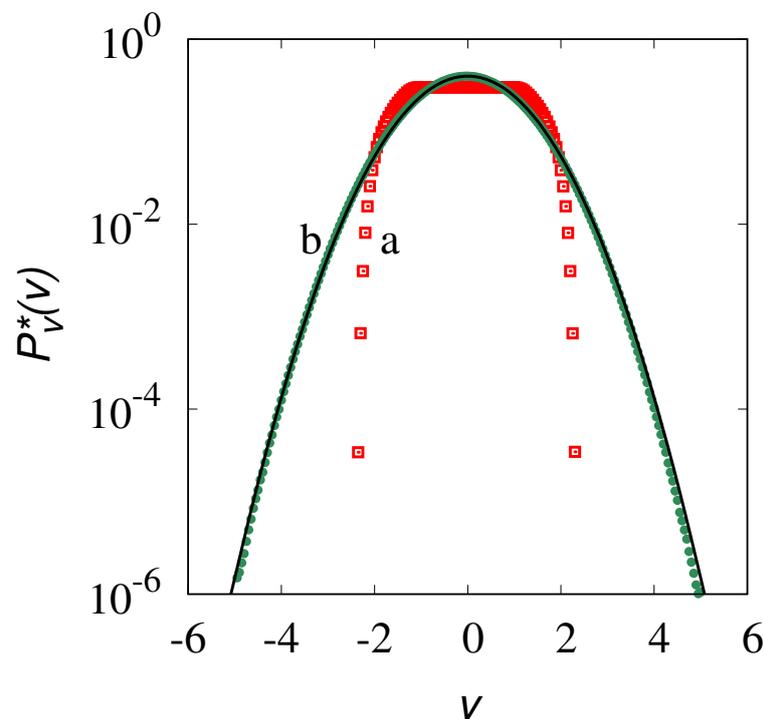

Figure 13. $P_v^*(v)$ vs v at equilibrium obtained from the radiative dynamics eqs. (50)-(51) at two values of α. Symbols represent the results of stochastic simulations: (a) refers to $\alpha = 0.4$, (b) to $\alpha = 0.9$. The solid line represents the normal distribution with zero mean and unit variance.

The analysis of the mechanical effects of radiative interactions is important at least for the following main reasons:

- it provides a simple physical mechanism, beyond the realm of mechanical interactions (collisions) for the thermalization process leading to thermal equilibrium conditions in a molecular system at constant T, where the temperature T is defined by the statistical properties of the incoming radiation;
- it provides a way to model momentum transfer in the presence of a generic out-of-equilibrium radiation, i.e., in the case the photons are not "thermal photons", but may possess other statistical properties;
- it indicates that Gaussianity in statistical physics is by no mean a fundamental constitutive principle (see e.g. the analysis of the Central Limit Theory addressed in [68]) as regards kinetic variables (velocities), but it results as a consequence of the physical conditions at which commonly we perform experiments. In the case of the radiative processes, this corresponds to temperatures above 10^{-5}-10^{-6} K;
- consequently, it provides a way to address the thermodynamic and transport features of cold atoms that, as discussed in [61], are characterized by highly anomalous transport and thermodynamic properties.

7. Concluding remarks

The stochastic modelling of chemical reaction kinetics is paradigmatic for a huge variety of chemical physical processes. In this article we have outlined the transition from a purely algorithmic description of these processes to a formal and simple stochastic formulation driven by distributional derivatives of counting processes. The approach proposed:

- shows a direct analogy between chemical reaction kinetics, radiative processes and stochastic formulation of open quantum systems, thus, paving the way for a unified treatment of the interplay between these phenomena, that is particularly important in the field of photochemistry, and in the foundation of statistical physics [56,62];
- can be easily extended to semi-Markov transitions. This is indeed the case of the growth kinetics of eukaryotic microorganisms, the physiological state of which can be parametrized with respect to internal (hidden) parameters such as the age, the cytoplasmatic content, etc.;
- can be easily extended to include transport phenomena. In point of fact, the occurrence of Markovian or semi-Markovian transitions in modeling chemical kinetics is analogous to the transitions occurring in the direction of motion (Poisson-Kac processes, Lévy flights, Extended Poisson-Kac processes) or in the velocity (linearized Boltzmannian schemes) [63–65].
- it is closely related to the formulation of stochastic differential equations for the thermalization of athermal system [66], in which the classical mesoscopic description of thermal fluctuations, using the increments of a Wiener process, is replaced by a dynamic model involving the increments of a counting process.

In a broader perspective, we have shown that it is possible to develop a stochastic microscopic event-based theory of chemical physical processes, and derive out of it macroscopic and emergent features. When microscopic processes are considered, Gaussianity for velocity density functions may be violated, as shown in connection with coalescence dynamics and radiative transitions. Although this violation refers to equilibrium property, it is easy to foresee that that this may impact non-equilibrium phenomena as well, indicating the necessity of modifying the classical parabolic paradigm of transport phenomena Gaussian density profiles stem from, by including and accounting for microscopic event-based physical transitions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Krapivsky, P. L. Redner, S. Ben-Naim, E. *A Kinetic View to Statistical Physics*; Cambridge University Press, Cambridge, 2010.

2. Boltzmann, L. Weitere Studien *ü*ber das W*ä*rmeglichgenicht unter Gas-molek*ü*len, *Sitzungsberichte Akademie der Wissenschaften*, 1872, 66, 275-370.

3. Marin, G. B. Yablonsky, G. S. Constales, D. *Kinetics of chemical reactions: decoding complexity*, John Wiley & Sons, New York, 2019.

4. Levenspiel, O. *Chemical Reaction Engineering*, J. Wiley & Sons, New York 1998.

5. Bird, R. B. Stewart, W. F. and Lightfoot, E. N. *Transport Phenomena*, J. Wiley, New York 2002.

6. de Groot S. R. and Mazur, P. *Non-equilibrium Thermodynamics* Dover Publications, New York, 1984.

7. Jou, D. Casa-Vazquez J. and Lebon G., *Extended Irreversible Thermodynamics*, Springer-Verlag, Berlin, 2001.

8. Frenkel D. and Smit, B. *Understanding Molecular Simulations*, Academic Press, New York, 2002.

9. Ollitrault, P. J. Miessen A. and Tavernelli, I. Molecular Quantum Dynamics: A Quantum Computing Perspective, *Acc. Chem. Res.* 2021, 54, 4229-4238.

10. Gardiner, C. *Stochastic Methods*, Springer-Verlag, Berlin, 2009.

11. Venerus D. C. and Öttinger, H. C. *A Modern Course in Transport Phenomena*, Cambridge University Press, Cambridge, 2018.

12. Giona, M. Brasiello A. and Crescitelli, S. Stochastic foundations of undulatory transport phenomena: generalized Poisson–Kac processes—part I basic theory, *J. Phys. A* 2017, 50, 335002.

13. van Kampen, N. G. *Stochastic Processes in Physics and Chemistry*, Elsevier, Amsterdam, 2007.

14. Harris, S. *An Introduction to the Theory of the Boltzmann Equation*, Dover Publishing, New York, 2004.

15. Giona, M. Brasiello A. and Crescitelli, S. Stochastic foundations of undulatory transport phenomena: generalized Poisson–Kac processes—part III extensions and applications to kinetic theory and transport. Stochastic foundations of undulatory transport phenomena: generalized Poisson–Kac processes—part III extensions and applications to kinetic theory and transport, *J. Phys. A*, 2017, 50 335004.

16. Einstein, A. *Investigations on the theory of the Brownian movement*, Dover Publishing, New York, 1956.

17. Langevin, P. Sur la théorie du mouvement brownien, *C. R. Acad. Sci.* Paris, 1908, 146 530–533.

18. Gnedenko B. V. and Kolmogorov, A. N. *Limit distributions for sums of independent random variables* Addison-Wesley, Cambridge MA 1954.

19. Fischer, H. *A History of the Central Limit Theorem*, Springer, New York, 2010.

20. Procopio, G. and Giona, M. Modal Representation of Inertial Effects in Fluid–Particle Interactions and the Regularity of the Memory Kernels *Fluids* 2023, 8 84.

21. Procopio G. and Giona, M. Thermodynamics of irreversible processes: fundamental constraints, representations, and formulation of boundary conditions, submitted to Physics, 2023.

22. Jou, D. Casas-Vazquez J. and Criado- Sancho, M. *Thermodynamics of Fluids under flow*, Springer-Verlag, Berlin, 2010.

23. Wang, Z. Hou, Z. Xin, H. Internal noise stochastic resonance of synthetic gene network, *Chemical Physics Letters*, 2005, 401 (1-3), 307-311.

24. Perc, M. Gosak, M. and Marhl, M. From stochasticity to determinism in the collective dynamics of diffusively coupled cells, *Chemical Physics Letters*, 2006, 421 (1-3) 106–110.

25. Lente, G. A binomial stochastic kinetic approach to the Michaelis–Menten mechanism, *Chemical Physics Letters*, 2013, 568 167–169.

26. McQuarrie, D. A. Stochastic approach to chemical kinetics, *Journal of Applied Probability*, 1967, 4 (3) 413-478.

27. Gillespie, D. T. Stochastic simulation of chemical kinetics, *Annual Review of Physical Chemistry*, 2007, 58 (1) 35-55.

28. Delbrück, M. Statistical fluctuations in autocatalytic reactions, *The Journal of Chemical Phsics*, 1940, 8 (1) 120-124.

29. Bartholomay, A. F. A stochastic approach to statistical kinetics with application to enzyme kinetics, *Biochemistry*, 1962, 1 (2) 223-230.

30. Gillespie, D. T. A rigorous derivation of the chemical master equation, *Physica A*, 1992, 188 (1-3) 404-425.

31. Keizer, J. On the necessity of using the master equation to describe the chemical reaction $X + A \rightleftharpoons B + X$, *Chemical Physics Letters*, 1971, 10 (4) 371–374.

32. Gaynor, B. J. Gilbert, R. G. King, K. D. Solution of the master equation for unimolecular reactions, *Chemical Physics Letters*, 1978, 55 (1) 40-43.

33. Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *Journal of Computational Physics*, 1976, 22 (4) 403-434.

34. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions, *The Journal of Physical Chemistry*, 1977, 81 (25) 2340-2361.

35. Gibson, M. A. Bruck, J. Efficient exact stochastic simulation of chemical systems with many species and many channels, *The Journal of Physical Chemistry A*, 2000, 104 (9) 1876-1889.

36. Lok , L. Brent, R. Automatic generation of cellular reaction networks with Moleculizer, *Nature Biotechnology*, 2005, 23 131–36

37. Cao, Y. Li, H. Petzold, L. R. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems, *The Journal of Chemical Physics*, 2004, 121 4059–67.

38. Rathinam, M. . Petzold, L. R Cao, Y. Gillespie, D. T. Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method, *The Journal of Chemical Physics*, 2003, 119 (24) 12784-12794.

39. Yang, C. Gillespie, D. T. Petzold, L. R. Efficient step size selection for the tau-leaping simulation method, *The Journal of Chemical Physics*, 2006, 124 (4) 044109.

40. C. Yang, D. T. Gillespie, L. R. Petzold, Adaptive explicit-implicit tau-leaping method with automatic tau selection, *The Journal of Chemical Physics*, 2007, 126 (22) 224101.

41. Venerus D. C. and Öttinger, H. C. *A modern Course in Transport Phenomena*, Cambridge University Press, Cambridge, 2018.

42. Ito, K. and McKean Jr., H. P. *Diffusion Processes and their Sample Paths*, Springer, Berlin, 1974.

43. Lecca, P. Stochastic chemical kinetics - A review of the modelling and simulation approaches, *Biophysical Review*, 2013, 5 323-345.

44. Campillo, F. Chebbi, M. Toumi, S. Stochastic modeling for biotechnologies Anaerobic model AM2b, Revue Africaine de la de la Recherche en Informatique et Mathématiques Appliqués, *Mathematics for Biology and the Environment*, INRIA, 2018-2019, 28 13-23.

45. Pezzotti C. Stochastic modelling of chemical reactions and transport-madiated chemical kinetics with applications to thermalization and biology, Master Thesis, University of Rome La Sapienza (2022).

46. Smith, B. T. and Hashmi, S. M. In situ gelation in confined flow controls intermittent dynamics, *chemrxiv-2023-mq8s2*, 2023 .

47. Aris, R. On the dispersion of a solute in a fluid flowing through a tube, *Proceedings of the Royal Society of London A*, 1956, **235**, 67-77.

48. Cerbelli, S., Giona, M., Garofalo, F. Quantifying dispersion of finite-sized particles in deterministic lateral displacement microflow separators through Brenner's macrotransport paradigm, *Microfluidics and Nanofluidics*, **15**, 431-449, 2013.

49. Egelstaff, P. A. *An Introduction to the Liquid State*, Clarendon Press, Oxford, 1994.

50. Zandveld, P., Andriesse, C. D., Bregman, J. D., Hasman, A., & Van Loef, J. J. Temperature dependence of the atomic self-motion in liquid argon, *Physica*, 1970, **50**, 511-523.

51. Endo, H. Density dependence of the effective mass for model liquid with twelfth inverse power potential, *Progress in Theoretical Physics*, 1977, **57**, 1457-1473.

52. Becker, R., Döring, W. Kinetische behandlung der keimbildung in übersättigten dämpfen. *Annalen der Physik*, 1935, **416**(8), 719-752.

53. Bressloff, P. C. *Stochastic processes in cell biology* (Vol. 41, pp. 608-614), Berlin, Springer, 2014.

54. Hingant, E., Yvinec, R. Deterministic and stochastic Becker–Döring equations: Past and recent mathematical developments. *Stochastic Processes, Multiscale Modeling, and Numerical Methods for Computational Cellular Biology*, 2017, 175-204.

55. Kubo, R., Toda, M., Hashitsume, N. *Statistical Physics II*, Springer Verlag, Berlin, 1991.

56. Pezzotti, C., Giona, M. Particle-photon radiative interactions and thermalization, *Physical Review E*, 2023, **108**, 024147.

57. Einstein, A. Zur Quantentheorie der Strahlung, *Physik. Z.*, 1917, **18**, 121-128.

58. Van der Waerden, B. L. *Sources of Quantum Mechanics*, Dover Publications, New York, 1968.

59. Milonni, P. W. *The Quantum Vacuum*, Academic Press, San Diego, 1994.

60. Fowler, R.H. *Statistical Mechanics, the theory of the properties of matter in equilibrium*, Cambridge University Press, Cambridge, 1929.

61. Bardou, F., Bouchaud, J.-F., Aspect, A., Coen-Tannoudji, C. *Lévy Statistics and Laser Cooling*, Cambridge University Press, Cambridge, 2002.

62. Breuer, H.-P., Petruccione, F. *The Theory of Open Quantum Systems*, Clarendon Press, Oxford, 2002.

63. Giona, M., Brasiello, A., Crescitelli, S. Stochastic foundations of undulatory transport phenomena: Generalized Poisson–Kac processes—Part I basic theory, *Journal of Physics A*, 2017, **50**, 335002.

64. Giona, M., Cairoli, A., Klages, R. Extended Poisson-Kac theory: A unifying framework for stochastic processes with finite propagation velocity, *Physical Review X*, 2022, **12**, 021004.

65. Sato, K. I. *Lévy processes and infinitely divisible distributions*, Cambridge University Press, Cambridge, 1999.

66. Kanazawa, K. *Statistical Mechanics for Athermal Fluctuations*, Springer Nature, Singapore, 2017.

67. Cocco, D., Giona, M. Generalized Counting Processes in a Stochastic Environment, *Mathematics*, 2021, **9**, 25-73.

68. Giona, M., Pezzotti, C., Procopio, G. The fourfold way to Gaussianity: physical interactions, distributional models and monadic transformations, *Axioms*, 2023 **12**, 278.