Analysis of the chemical diffusion master equation for creation and mutual annihilation reactions

Analysis of the chemical diffusion master equation for creation and mutual annihilation reactions

Alberto Lanconelli[1] and Berk Tan Perçin[1]

Dipartimento di Scienze Statistiche Paolo Fortunati, Università di Bologna, Bologna, Italy.

(*Electronic mail: berktan.percin2@unibo.it)

(*Electronic mail: alberto.lanconelli2@unibo.it)

(Dated: March 13, 2024)

We propose an infinite dimensional generating function method for finding the analytical solution of the so-called chemical diffusion master equation (CDME) for creation and mutual annihilation chemical reactions. CDMEs model by means of an infinite system of coupled Fokker-Planck equations the probabilistic evolution of chemical reaction kinetics associated with spatial diffusion of individual particles; here, we focus an creation and mutual annihilation chemical reactions combined with Brownian diffusion of the single particles. Using our method we are able to link certain finite dimensional projections of the solution of the CDME to the solution of a single linear fourth order partial differential equation containing as many variables as the dimension of the aforementioned projection space. Our technique extends the one presented in[22] and[23] which allowed for an explicit representation for the solution of birth-death type CDMEs.

Key words and phrases: Particle-based reaction-diffusion models, Brownian motion, Malliavin calculus.
AMS 2020 classification: 60H07; 60H30; 92E20.

1. INTRODUCTION

We consider a system of indistinguishable molecules of a chemical species S which undergo

- drift-less isotropic *diffusion* in the interval $[0,1]$;
- *creation* and *mutual annihilation* chemical reactions

$$\text{(I)} \quad \varnothing \xrightarrow{\lambda_c(x)} S \qquad\qquad \text{(II)} \quad S+S \xrightarrow{\lambda_d(x,y)} \varnothing.$$

Here, the function $[0,1] \ni x \mapsto \lambda_c(x)$ represents the stochastic rate function for reaction (I); it can be thought of being of the form $\lambda_c(x) = \gamma\pi_c(x)$ with γ being a positive constant representing the probability per unit of time for a new particle to be created while π_c is a probability density on $[0,1]$ which describes the random location for the birth of the new particle. Similarly, the function $[0,1]^2 \ni (x,y) \mapsto \lambda_d(x,y)$ is the stochastic rate function for reaction (II) to occur between two particles located at (x,y); for instance, when λ_d is constant then the location of the two particles is not relevant for reaction (II) to take place; on the contrary, if $\lambda_d(x,y) = \delta(x-y)$ (here δ stands for the Dirac delta function with mass at zero) then reaction (II) occurs (with rate one) only for particles having the same location.
To analyze the probabilistic evolution of such system the authors in[5] (see also[7] for a further discussion of the model) proposed a set of equations which describe how the number of molecules and their positions change with time. Namely, for $t \geq 0$, $n \geq 1$ and $A \in \mathcal{B}([0,1]^n)$ they set

$$\begin{aligned} \mathcal{N}(t) &:= \text{ number of molecules at time } t, \\ \rho_0(t) &:= \mathbb{P}(\mathcal{N}(t) = 0), \\ \int_A \rho_n(t,x_1,...,x_n)dx_1\cdots dx_n &:= \mathbb{P}\left(\{\mathcal{N}(t)=n\} \cap \{(X_1(t),...,X_n(t)) \in A\}\right); \end{aligned}$$

here $(X_1(t),...,X_n(t))$ is the vector collecting the positions at time t of the n particles constituting the system (we are also assuming that the stochastic processes under investigation are defined on a common probability space with

reference measure $\mathbb{P}$). Then, they write the following infinite system of equations:

$$\begin{cases} \partial_t \rho_n(t,x_1,...,x_n) = \displaystyle\sum_{i=1}^{n} \partial^2_{x_i}\rho_n(t,x_1,...,x_n) \\ \qquad + \dfrac{(n+2)(n+1)}{2}\displaystyle\int_{[0,1]^2} \lambda_d(x,y)\rho_{n+2}(t,x_1,...,x_n,x,y)dxdy \\ \qquad - \displaystyle\sum_{i<j} \lambda_d(x_i,x_j)\cdot\rho_n(t,x_1,...,x_n) \\ \qquad + \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \lambda_c(x_i)\rho_{n-1}(t,x_1,...,x_{i-1},x_{i+1},...,x_n) \\ \qquad - \displaystyle\int_{[0,1]} \lambda_c(y)dy\cdot\rho_n(t,x_1,...,x_n), \qquad n\geq 0, t>0, (x_1,...,x_n)\in[0,1]^n, \end{cases} \tag{1.1}$$

where we agree on assigning value zero to the three sums above when $n = 0$. The term

$$\sum_{i=1}^{n} \partial^2_{x_i}\rho_n(t,x_1,...,x_n)$$

in (1.1) refers to spatial diffusion of the particles; the terms

$$\frac{(n+2)(n+1)}{2}\int_{[0,1]^2} \lambda_d(x,y)\rho_{n+2}(t,x_1,...,x_n,x,y)dxdy$$

and

$$\sum_{i<j} \lambda_d(x_i,x_j)\cdot\rho_n(t,x_1,...,x_n)$$

formalize gain and loss, respectively, due to reaction (II), while

$$\frac{1}{n}\sum_{i=1}^{n} \lambda_c(x_i)\rho_{n-1}(t,x_1,...,x_{i-1},x_{i+1},...,x_n)$$

and

$$-\int_{[0,1]} \lambda_c(y)dy\cdot\rho_n(t,x_1,...,x_n)$$

represent gain and loss, respectively, associated to reaction (I). System (1.1) is combined with initial and Neumann boundary conditions

$$\begin{cases} \rho_0(0) = 1; \\ \rho_n(0,x_1,...,x_n) = 0, \quad n\geq 1, (x_1,...,x_n)\in[0,1]^n; \\ \partial_\nu\rho_n(t,x_1,...,x_n) = 0, \quad n\geq 1, t\geq 0, (x_1,...,x_n)\in\partial[0,1]^n. \end{cases} \tag{1.2}$$

The initial condition (first two equations in (1.2)) states that there are no molecules in the system at time zero while the Neumann condition prevents flux through the boundary of $[0,1]$, thus forcing the diffusion of the molecules inside $[0,1]$. The symbol ∂_ν in (1.2) stands for the directional derivative along the outer normal vector at the boundary of $[0,1]^n$.

A. Literature review

The dynamics of biochemical processes in living cells are commonly understood as an interplay between the spatial transport (diffusion) of molecules and their chemical kinetics (reaction), both of which are inherently stochastic at the molecular scale. In the case of systems with small molecule numbers in spatially well-mixed settings, the diffusion is averaged out and the probabilistic dynamics are governed by the well-known chemical master equation (CME)[14,28,29]. The CME can be seldom solved analytically[17]. However, solving a few simple cases analytically can bring valuable

insight to the solutions of more complex cases. Alternatively, one can solve it by integrating stochastic trajectories with the Gillespie or tau-leap algorithms[1,14], by approximation methods[8,11,26,30] or even by deep learning approaches[15,19]. In the case of spatially inhomogeneous systems, where diffusion is not averaged out, one would expect to obtain a similar master equation. However, obtaining such an equation is plagued with mathematical difficulties, and although it was hinted in previous work[10] and formulated for some specific systems[31], it was not until recently that this was formalized into the so-called chemical diffusion master equation (CDME)[5,7]. The CDME changes a few paradigms that have not yet been explored thoroughly in stochastic chemical kinetics models. It combines continuous and discrete degrees of freedom, and it models reaction and diffusion as a joint stochastic process. It consists of an infinite sorted family of Fokker-Planck equations, where each level of the sorted family corresponds to a certain number of particles/molecules. The equations at each level describe the spatial diffusion of the corresponding set of particles, and they are coupled to each other via reaction operators, which change the number of particles in the system. The CDME is the theoretical backbone of reaction-diffusion processes, and thus, it is fundamental to model and understand biochemical processes in living cells, as well as to develop multiscale numerical methods[6,13,21,32] and hybrid algorithms[3,4,9]. The stochastic trajectories of the CDME can be often integrated using particle–based reaction–diffusion simulations[2,16].
The problem of finding analytical solutions to some CDMEs has been recently addressed in the papers[22] and[23]. In[22] the author proposed an infinite dimensional version of the classical generating function method, which is commonly utilized to find analytical solution to some CMEs[24],[12],[25] (see also[33]). In[22] the method is employed to solve a CDME of birth-death type; this approach has been further explored in[23] and an explicit solution for the general birth-death CDME is presented.

B. Statement of the main result and structure of the paper

Our main result links certain finite dimensional projections of the solution of (1.1) to a single linear partial differential equation thus providing a tool for finding analytical solutions to (1.1). Our techniques applies to the case where the function $\lambda_d : [0,1]^2 \to [0,+\infty[$ is constant. This means that the CDME in (1.1) simplifies to

$$\begin{cases} \partial_t \rho_n(t,x_1,...,x_n) = \sum_{i=1}^n \partial_{x_i}^2 \rho_n(t,x_1,...,x_n) \\ \qquad + \lambda_d \dfrac{(n+2)(n+1)}{2} \displaystyle\int_{[0,1]^2} \rho_{n+2}(t,x_1,...,x_n,x,y)dxdy \\ \qquad - \lambda_d \dfrac{n(n-1)}{2} \rho_n(t,x_1,...,x_n) \\ \qquad + \dfrac{1}{n} \displaystyle\sum_{i=1}^n \lambda_c(x_i) \rho_{n-1}(t,x_1,...,x_{i-1},x_{i+1},...,x_n) \\ \qquad - \gamma \rho_n(t,x_1,...,x_n), \qquad n \geq 0, t > 0, (x_1,...,x_n) \in [0,1]^n, \end{cases} \tag{1.3}$$

We are now going to state our main result and we refer the reader to the next sections for a detailed discussion of our assumptions and for the proof.

Theorem 1.1. *For $k \geq 1$ let*

$$\xi_k(x) := \sqrt{2}\cos((k-1)\pi x), x \in [0,1] \quad \text{and} \quad \alpha_k := (k-1)^2\pi^2,$$

i.e. the eigenfunctions with corresponding eigenvalues of the differential operator $-\mathcal{A} = \partial_x^2$ with homogeneous Neumann boundary conditions (as prescribed in (1.2)). We also write Π_N for the orthogonal projection onto the linear span of $\{\xi_1,...,\xi_N\}$.
If $\{\rho_n\}_{n\geq 0}$ solves (1.3)-(1.2)*, then we have the representation*

$$\Pi_N^{\otimes n} \rho_n(t,x_1,...,x_n) = \frac{1}{n!} \sum_{j_1,...,j_n=1}^N \left(\int_{\mathbb{R}^N} (\partial_{z_{j_1}} \cdots \partial_{z_{j_n}} u_N)(t,z)(2\pi)^{-N/2} e^{-\frac{|z|^2}{2}} dz \right) \xi_{j_1}(x_1) \cdots \xi_{j_n}(x_n), \tag{1.4}$$

where u_N is the solution to the following fourth order Cauchy problem

$$\begin{cases} \partial_t u_N(t,z) = -\sum_{k=1}^N \alpha_k (z_k - \partial_{z_k}) \partial_{z_k} u_N(t,z) + \frac{\lambda_d}{2} \partial_{z_1}^2 u_N(t,z) \\ \qquad\qquad -\frac{\lambda_d}{2} \sum_{j,k=1}^N (z_j - \partial_{z_j})(z_k - \partial_{z_k}) \partial_{z_j} \partial_{z_k} u_N(t,z) \\ \qquad\qquad + \sum_{k=1}^N \mathsf{c}_k (z_k - \partial_{z_k}) u_N(t,z) - \gamma u_N(t,z); \quad t > 0, z \in \mathbb{R}^N; \\ u_N(0,z) = 1, \quad z \in \mathbb{R}^N. \end{cases}$$

and $\mathsf{c}_k := \langle \lambda_c, \xi_k \rangle_{L^2([0,1])}$ *for* $k \geq 1$.

The paper is organized as follows: in Section 2 we propose an alternative derivation of equation (1.1) than the one presented in[5] and[7]. We mimic the classical approach utilized to obtain the chemical master equation through an adaptation that includes diffusion of the single particles. Even though the computation is pretty standard, we believe that such derivation helps for a better understanding of the ingredients that describe the problem under investigation. To this aim we will fix a set of transition probabilities (see Assumption 2.3 below) and derive through a limit argument the desired equation. We remark that our approach can be readily generalized to include higher order chemical reactions and more complex descriptions of the diffusive motion of the particles (i.e. anisotropic diffusion with drift). In Section 3 we employ the general method proposed in[22] to analytically solve equation (1.3)-(1.2). This requires the use of Gaussian Malliavin calculus's techniques (summarized in the Appendix below) and provides a link between the solution to (1.3)-(1.2) and the solution of a single fourth order linear PDE which describes certain finite dimensional projections of the solution to the original problem. At the end, some comments on the Gaussian features introduced in our problem by the proposed approach are also discussed.

2. ALTERNATIVE DERIVATION OF EQUATION (1.1)

In this section we present a derivation of equation (1.1) (here we do not need to confine ourselves to the case of a constant λ_d function as prescribed in the Theorem 1.1). We recall that the particles of the system under investigation are subject to the chemical reactions

$$\text{(I)} \quad \varnothing \xrightarrow{\lambda_c(x)} S \qquad \text{(II)} \quad S + S \xrightarrow{\lambda_d(x,y)} \varnothing, \tag{2.5}$$

and diffuse in space, between successive reactions, as independent Brownian motions on the interval $[0,1]$ with reflecting boundary conditions (compare with (1.2)). In the sequel we will be dealing with probabilities of the form $\mathbb{P}(\mathcal{N}(t) = n, X(t) \in A)$: this represents the probability that the system at time t is made of n many particles and that such particles are located in the region $A \subseteq [0,1]^n$. We are not going to use an extra index in $X(t)$ to stress that it is an n-dimensional vector; this vector will always come with an event of the type $\{\mathcal{N}(t) = n\}$ and hence the number of components of $X(t)$ will be uniquely determined. We mention that an analogous derivation for the chemical master equation with mutual annihilation and creation (without diffusion) can be found in[12], Section 1.4.
Now we list a couple of technical assumptions which are necessary for our derivation:

Assumption 2.1. *For any* $n \geq 1$, $A \in \mathcal{B}([0,1]^n)$ *(the Borel sets of* $[0,1]^n$*) and* $t > 0$ *there exists a symmetric function* $\rho_n(t, x_1, ..., x_n)$ *such that*

$$\mathbb{P}(\mathcal{N}(t) = n, X(t) \in A) = \int_A \rho_n(t, x_1, ..., x_n) dx_1 \cdots dx_n;$$

we also set

$$\rho_0(t) := \mathbb{P}(\mathcal{N}(t) = 0).$$

Notice that the symmetry of the functions ρ_n models the indistinguishability of the particles in the system; moreover, by construction the sequence $\{\rho_n\}_{n\geq 0}$ fulfils the constraint

$$\sum_{n\geq 0} \int_{[0,1]^n} \rho_n(t, x_1, ..., x_n) dx_1 \cdots dx_n = 1. \tag{2.6}$$

Assumption 2.2. *The functions* λ_c *and* λ_d *appearing in (2.5) are non negative, bounded and continuous. Moreover,* $\lambda_d(x,y) = \lambda_d(y,x)$ *for all* $x, y \in [0,1]$.

We are now ready to describe the probabilistic structure to be imposed on our system for the formal derivation of equation (1.1).

Assumption 2.3. *The system under investigation possesses the following properties:*

- Diffusion of particles*: in absence of chemical reactions, particles diffuse in* $[0,1]$ *like independent Brownian motions with variance* $2t$ *and reflecting boundary conditions; more precisely, the transition density* $\{p_t(x|y)\}_{t\geq 0, x,y\in[0,1]^n}$ *for the motion of* n *many particles solves*

$$\begin{cases} \partial_t p_t(x|y) = \sum_{j=1}^n \partial^2_{x_i} p_t(t,x|y), & t>0, x,y\in[0,1]^n; \\ p_0(x|y) = \delta_y(x), & x,y\in[0,1]^n; \\ \partial_\nu p_t(x|y) = 0, & t\geq 0, x\in\partial[0,1]^n, y\in[0,1]^n. \end{cases}$$

- Reaction (I) + diffusion of particles*: for any* $n\geq 1$*,* $A\in\mathcal{B}([0,1]^n)$ *and* $t,h>0$ *we have*

$$\begin{aligned} &\mathbb{P}(\mathcal{N}(t+h)=n, X(t+h)\in A|\mathcal{N}(t)=n-1, X(t)=y) \\ &\quad = h\frac{1}{n}\sum_{j=1}^n \int_A\int_{[0,1]} \mathtt{p}_h(x|y\cup_j z)\lambda_c(z)dzdx + \mathcal{O}(h^2), \end{aligned} \tag{2.7}$$

with $y\cup_j z := (y_1,...,y_{j-1},z,y_{j+1},...,y_n)\in[0,1]^n$ *and* $\mathtt{p}_h(x|y\cup_j z)$ *being the transition density between* $y\cup_j z$ *and* x *during the time interval* h*. To explain the contribution of each single term of the identity above, we imagine to split the function* λ_c *as* $\gamma\cdot\pi_c$ *where* $\gamma := \int_{[0,1]}\lambda_c(z)dz$ *while* π_c *is a probability density function supported on* $[0,1]$*. The chemical reaction (I) adds a new particle, here denoted with* z*, to the system: the rate at which this happens is* γ *while the location for the birth of the particle is distributed according to* π_c*. Moreover, once the creation takes place the outer integral* $\int_A ...dx$ *in (2.7) describes the diffusion of the* n *particles of the system from the location* $y\cup_j z$ *to the set* A *during the time frame* h*. Lastly, to make particles indistinguishable we symmetrize over the possible positions of* z *in the vector* $y\cup_j z$ *with the term* $\frac{1}{n}\sum_{j=1}^n$*.*

- Reaction (II) + diffusion of particles*: for any* $n\geq 0$*,* $A\in\mathcal{B}([0,1]^n)$ *and* $t,h>0$ *we have*

$$\begin{aligned} &\mathbb{P}(\mathcal{N}(t+h)=n, X(t+h)\in A|\mathcal{N}(t)=n+2, X(t)=y) \\ &\quad = h\sum_{j<k}\lambda_d(y_j,y_k)\int_A \mathtt{p}_h(x|\hat{y}_{j,k})dx + \mathcal{O}(h^2), \end{aligned} \tag{2.8}$$

with $\hat{y}_{j,k} := (y_1,...,y_{j-1},y_{j+1},...,y_{k-1},y_{k+1},...,y_n)\in[0,1]^n$*. The chemical reaction (II) removes two particles from the system while the others diffuse: this is the contribution of* $\int_A \mathtt{p}_h(x|\hat{y}_{j,k})dx$ *where the particles labelled* j *and* k *are those undergoing the chemical reaction through the term* $h\sum_{j<k}\lambda_d(y_j,y_k)$*. This term is mediated over all the possible couples of particles in the system: the weights of this average, represented by the sum above, are provided by* λ_d *which measures the likelihood for two particles to react depending on their locations. We also mention that for* $n=0$ *the right hand side of (2.8) simplifies to* $h\lambda_d(x_1,x_2)+\mathcal{O}(h^2)$*.*

- No reactions + diffusion of particles*: for any* $n\geq 1$*,* $A\in\mathcal{B}([0,1]^n)$ *and* $t,h>0$ *we have*

$$\begin{aligned} &\mathbb{P}(\mathcal{N}(t+h)=n, X(t+h)\in A|\mathcal{N}(t)=n, X(t)=y) \\ &\quad = \left(1-h\int_{[0,1]}\lambda_c(z)dz - h\sum_{j<k}\lambda_d(y_j,y_k)\right)\int_A \mathtt{p}_h(x|y)dx + \mathcal{O}(h^2). \end{aligned} \tag{2.9}$$

The term inside parenthesis reflects the probability of no reaction happening while the integral formalizes the diffusion of particles.

- Multiple reactions*: for any* $n\geq 1$*,* $A\in\mathcal{B}([0,1]^n)$ *and* $t,h>0$ *we have*

$$\mathbb{P}(\mathcal{N}(t+h)=n, X(t+h)\in A|\mathcal{N}(t)=k, X(t)=y) = \mathcal{O}(h^2), \tag{2.10}$$

whenever $k\notin\{n-1,n,n+2\}$*.*

We now show how to use Assumption 2.1-2.2-2.3 to get the CDME (1.1). Let $n \geq 1$ and $A \in \mathcal{B}([0,1]^n)$; then, according to the law of total probability we can write

$$\begin{aligned}
\int_A \rho_n(t+h,x)dx =& \mathbb{P}(\mathcal{N}(t+h)=n, X(t+h)\in A)\\
=& \sum_{k\geq 0}\int_{[0,1]^k} \mathbb{P}(\mathcal{N}(t+h)=n, X(t+h)\in A|\mathcal{N}(t)=k, X(t)=y)\mathbb{P}(\mathcal{N}(t)=k, X(t)\in dy)\\
=& \sum_{k\geq 0}\int_{[0,1]^k} \mathbb{P}(\mathcal{N}(t+h)=n, X(t+h)\in A|\mathcal{N}(t)=k, X(t)=y)\rho_k(t,y)dy. \qquad (2.11)
\end{aligned}$$

Notice that for $k=0$ the corresponding term in the sum above should be interpreted as

$$\mathbb{P}(\mathcal{N}(t+h)=n, X(t+h)\in A|\mathcal{N}(t)=0)\rho_0(t).$$

Now, in view of Assumption 2.3 the only transitions of order one in h are those with $k=n+2$, $k=n-1$ and $k=n$ while the others are of order at least two; therefore, we can rewrite (2.11) as

$$\begin{aligned}
\int_A \rho_n(t+h,x)dx =& \sum_{k\geq 0}\int_{[0,1]^k} \mathbb{P}(\mathcal{N}(t+h)=n, X(t+h)\in A|\mathcal{N}(t)=k, X(t)=y)\rho_k(t,y)dy\\
=& \int_{[0,1]^{n+2}} \mathbb{P}(\mathcal{N}(t+h)=n, X(t+h)\in A|\mathcal{N}(t)=n+2, X(t)=y)\rho_{n+2}(t,y)dy\\
&+\int_{[0,1]^{n-1}} \mathbb{P}(\mathcal{N}(t+h)=n, X(t+h)\in A|\mathcal{N}(t)=n-1, X(t)=y)\rho_{n-1}(t,y)dy\\
&+\int_{[0,1]^{n}} \mathbb{P}(\mathcal{N}(t+h)=n, X(t+h)\in A|\mathcal{N}(t)=n, X(t)=y)\rho_{n}(t,y)dy\\
&+\mathcal{O}(h^2)\\
=& \int_{[0,1]^{n+2}} \left(\int_A h\sum_{j<k}\lambda_d(y_j,y_k)\mathtt{p}_h(x|\hat{y}_{j,k})dx\right)\rho_{n+2}(t,y)dy\\
&+\int_{[0,1]^{n-1}} \left(\int_A \frac{h}{n}\sum_{j=1}^n\int_{[0,1]}\lambda_c(z)\mathtt{p}_h(x|y\cup_j z)dzdx\right)\rho_{n-1}(t,y)dy\\
&+\int_{[0,1]^{n}} \left(\int_A \left(1-h\int_{[0,1]}\lambda_c(z)dz-h\sum_{j<k}\lambda_d(y_j,y_k)\right)\mathtt{p}_h(x|y)dx\right)\rho_{n}(t,y)dy\\
&+\mathcal{O}(h^2)\\
=& h\int_A\left(\sum_{j<k}\int_{[0,1]^{n+2}}\lambda_d(y_j,y_k)\mathtt{p}_h(x|\hat{y}_{j,k})\rho_{n+2}(t,y)dy\right)dx\\
&+h\int_A\left(\frac{1}{n}\sum_{j=1}^n\int_{[0,1]^{n-1}}\int_{[0,1]}\lambda_c(z)\mathtt{p}_h(x|y\cup_j z)\rho_{n-1}(t,y)dzdy\right)dx\\
&+\int_A\left(\int_{[0,1]^n}\left(1-h\int_{[0,1]}\lambda_c(z)dz-h\sum_{j<k}\lambda_d(y_j,y_k)\right)\mathtt{p}_h(x|y)\rho_n(t,y)dy\right)dx\\
&+\mathcal{O}(h^2)\\
=& h\int_A\left(\sum_{j<k}\int_{[0,1]^2}\lambda_d(y_j,y_k)\left(\int_{[0,1]^n}\mathtt{p}_h(x|\hat{y}_{j,k})\rho_{n+2}(t,y)d\hat{y}_{j,k}\right)dy_jdy_k\right)dx\\
&+h\int_A\left(\frac{1}{n}\sum_{j=1}^n\int_{[0,1]^{n-1}}\int_{[0,1]}\lambda_c(z)\mathtt{p}_h(x|y\cup_j z)\rho_{n-1}(t,y)dzdy\right)dx \qquad (2.12)
\end{aligned}$$

$$+\int_A\left(\int_{[0,1]^n}\left(1-h\int_{[0,1]}\lambda_c(z)dz-h\sum_{j<k}\lambda_d(y_j,y_k)\right)\mathtt{p}_h(x|y)\rho_n(t,y)dy\right)dx$$
$$+\mathcal{O}(h^2).$$

To ease the notation we now introduce the following:

$$\mathtt{T}_hf(x):=\int_{[0,1]^n}\mathtt{p}_h(x|y)f(y)dy,\quad f\in C_0([0,1]^n),$$

and recall that for suitably regular f we have

$$\lim_{h\to0}\mathtt{T}_hf(x)=f(x),\quad x\in[0,1]^n,\tag{2.13}$$

and

$$\lim_{h\to0}\frac{\mathtt{T}_hf(x)-f(x)}{h}=\sum_{i=1}^n\partial^2_{x_i}f(x)\quad x\in[0,1]^n;\tag{2.14}$$

we refer to[20] for a precise formulation of those statements. With this notation at hand we can rewrite (2.12) as

$$\begin{aligned}\int_A\rho_n(t+h,x)dx=&h\int_A\left(\sum_{j<k}\int_{[0,1]^2}\lambda_d(y_j,y_k)(\mathtt{T}_h\rho_{n+2}(t,\cdot,y_j,y_k))(x)dy_jdy_k\right)dx\\&+h\int_A\frac{1}{n}\sum_{j=1}^n\mathtt{T}_h(\lambda_c\otimes_j\rho_{n-1}(t,\cdot))(x)dx\\&+\int_A(\mathtt{T}_h\rho_n(t,\cdot))(x)dx-h\int_{[0,1]}\lambda_c(z)dz\int_A(\mathtt{T}_h\rho_n(t,\cdot))(x)dx\\&-h\sum_{j<k}(\mathtt{T}_h\lambda_d(\cdot_j,\cdot_k)\rho_n(t,\cdot))(x)dx+\mathcal{O}(h^2).\end{aligned}$$

Here, the symbol $\otimes_j$ denotes the tensor product that locates the variable of λ_c in the j-th position. We now subtract the quantity $\int_A\rho_n(t,x)dx$ from both sides of the last equality, divide by h and take the limit as h tends to zero. This gives

$$\begin{aligned}\int_A\partial_t\rho_n(t,x)dx=&\lim_{h\to0}\int_A\frac{\rho_n(t+h,x)-\rho_n(t,x)}{h}dx\\=&\lim_{h\to0}\int_A\left(\sum_{j<k}\int_{[0,1]^2}\lambda_d(y_j,y_k)(\mathtt{T}_h\rho_{n+2}(t,\cdot,y_j,y_k))(x)dy_jdy_k\right)dx\\&+\lim_{h\to0}\int_A\frac{1}{n}\sum_{j=1}^n\mathtt{T}_h(\lambda_c\otimes_j\rho_{n-1}(t,\cdot))(x)dx\\&+\lim_{h\to0}\int_A\frac{(\mathtt{T}_h\rho_n(t,\cdot))(x)-\rho_n(t,x)}{h}dx-\lim_{h\to0}\int_{[0,1]}\lambda_c(z)dz\int_A(\mathtt{T}_h\rho_n(t,\cdot))(x)dx\\&-\lim_{h\to0}\int_A\sum_{j<k}(\mathtt{T}_h\lambda_d(\cdot_j,\cdot_k)\rho_n(t,\cdot))(x)dx.\end{aligned}$$

Now, using (2.13) we get

$$\begin{aligned}&\lim_{h\to0}\int_A\left(\sum_{j<k}\int_{[0,1]^2}\lambda_d(y_j,y_k)(\mathtt{T}_h\rho_{n+2}(t,\cdot,y_j,y_k))(x)dy_jdy_k\right)dx\\&=\int_A\left(\sum_{j<k}\int_{[0,1]^2}\lambda_d(y_j,y_k)\rho_{n+2}(t,x,y_j,y_k)dy_jdy_k\right)dx\end{aligned}$$

$$= \int_A \left(\frac{(n+2)(n+1)}{2} \int_{[0,1]^2} \lambda_d(y_j, y_k) \rho_{n+2}(t, x, y_j, y_k) dy_j dy_k \right) dx,$$

and

$$\lim_{h\to 0} \int_A \frac{1}{n} \sum_{j=1}^{n} \mathtt{T}_h(\lambda_c \otimes_j \rho_{n-1}(t,\cdot))(x) dx = \int_A \left(\frac{1}{n} \sum_{j=1}^{n} \lambda_c(x_j) \rho_{n-1}(t, x_1, ..., x_{j-1}, x_{j+1}, ..., x_n) \right) dx.$$

Moreover, formula (2.14) yields

$$\lim_{h\to 0} \int_A \frac{(\mathtt{T}_h \rho_n(t,\cdot))(x) - \rho_n(t,x)}{h} dx = \int_A \left(\sum_{i=1}^{n} \partial^2_{x_i} \rho_n(t,x) \right) dx,$$

while formula (2.13) gives

$$\lim_{h\to 0} \int_{[0,1]} \lambda_c(z) dz \int_A (\mathtt{T}_h \rho_n(t,\cdot))(x) dx = \int_{[0,1]} \lambda_c(z) dz \int_A \rho_n(t,x) dx$$

and

$$\lim_{h\to 0} \int_A \sum_{j<k} (\mathtt{T}_h \lambda_d(\cdot_j, \cdot_k) \rho_n(t,\cdot))(x) dx = \int_A \sum_{j<k} \lambda_d(x_j, x_k) \rho_n(t,x) dx.$$

If we combine all the preceding equalities we can conclude that

$$\begin{aligned}
\int_A \partial_t \rho_n(t,x) dx = & \int_A \left(\frac{(n+2)(n+1)}{2} \int_{[0,1]^2} \lambda_d(y_j, y_k) \rho_{n+2}(t, x, y_j, y_k) dy_j dy_k \right) dx \\
& + \int_A \left(\frac{1}{n} \sum_{j=1}^{n} \lambda_c(x_j) \rho_{n-1}(t, x_1, ..., x_{j-1}, x_{j+1}, ..., x_n) \right) dx \\
& + \int_A \left(\sum_{i=1}^{n} \partial^2_{x_i} \rho_n(t,x) \right) dx \\
& - \int_{[0,1]} \lambda_c(z) dz \int_A \rho_n(t,x) dx \\
& - \int_A \sum_{j<k} \lambda_d(x_j, x_k) \rho_n(t,x) dx.
\end{aligned}$$

Since $A \in \mathcal{B}([0,1]^n)$ is arbitrary, this is equivalent to (1.1).

3. ANALYSIS OF EQUATION (1.3)-(1.2) THROUGH AN INFINITE DIMENSIONAL GENERATING FUNCTION METHOD

In this section we employ the general method proposed in[22] to solve analytically the CDME (1.1)-(1.2). We mention that this method has lead to an explicit representation for the solution to the general birth-death CDME[23]. For the application of that approach in the current framework we need to impose the following technical condition.

Assumption 3.1. *The function $\lambda_d : [0,1]^2 \to [0, +\infty[$ is constant; this means that equation (1.1) simplifies to*

$$\left\{ \begin{aligned}
\partial_t \rho_n(t, x_1, ..., x_n) = & \sum_{i=1}^{n} \partial^2_{x_i} \rho_n(t, x_1, ..., x_n) \\
& + \lambda_d \frac{(n+2)(n+1)}{2} \int_{[0,1]^2} \rho_{n+2}(t, x_1, ..., x_n, x, y) dx dy \\
& - \lambda_d \frac{n(n-1)}{2} \rho_n(t, x_1, ..., x_n) \\
& + \frac{1}{n} \sum_{i=1}^{n} \lambda_c(x_i) \rho_{n-1}(t, x_1, ..., x_{i-1}, x_{i+1}, ..., x_n) \\
& - \gamma \rho_n(t, x_1, ..., x_n), \qquad n \geq 0, t > 0, (x_1, ..., x_n) \in [0,1]^n,
\end{aligned} \right. \tag{3.15}$$

The need for such assumption is related to some technical features of the method used to solve (3.15). We refer to Remark 3.6 below for details.

Remark 3.2. *As pointed out in*[5]*, if we fix the reaction rates to be positive constants so* $\gamma, \lambda_d \in \mathbb{R}^+$ *and integrate equation (3.15)-(1.2) with respect to all degrees of freedom, then this reduces to the classical chemical master equation for the reactions*

$$(I) \quad \varnothing \xrightarrow{\gamma} S \qquad (II) \quad S + S \xrightarrow{\lambda_d} \varnothing.$$

In fact, identity

$$\mathbb{P}(\mathcal{N}(t) = n) = \int_{[0,1]^n} \rho_n(t, x_1, ..., x_n) dx_1 \cdots dx_n,$$

(compare with Assumption 2.1) together with the boundary conditions in (1.2) yield

$$\begin{aligned} \partial_t \mathbb{P}(\mathcal{N}(t) = n) =& \lambda_d \frac{(n+2)(n+1)}{2} \mathbb{P}(\mathcal{N}(t) = n+2) - \lambda_d \frac{n(n-1)}{2} \mathbb{P}(\mathcal{N}(t) = n) \\ &+ \gamma \mathbb{P}(\mathcal{N}(t) = n-1) - \gamma \mathbb{P}(\mathcal{N}(t) = n), \end{aligned} \tag{3.16}$$

which is indeed the desired chemical master equation (see equation (1.28) in[12]*: here, the authors employ the standard generating function method for solving equation (3.16)). We also mention that such computation, and hence the link between chemical diffusion master equations and their corresponding chemical master equations, is far from being obvious.*

The scheme for solving (3.15), as presented in[22], is made of several steps that we now discuss in the following preparatory results. Before doing that, we introduce the notation $-\mathcal{A} := \partial_x^2$ and we recall that in the Appendix below one can find a quick review of the Malliavin calculus's tools utilized in the sequel.

Lemma 3.3. *If* $\{\rho_n\}_{n \geq 0}$ *is a classical solution to equation (3.15)-(1.2), then*

$$\Phi(t) := \sum_{n \geq 0} I_n(\rho_n(t, \cdot)) \tag{3.17}$$

solves

$$\begin{cases} \partial_t \Phi = d\Gamma(-\mathcal{A})\Phi + \frac{\lambda_d}{2} D_1^2 \Phi - \frac{\lambda_d}{2} \mathtt{N}(\mathtt{N} - \mathtt{I})\Phi + D^\star_{\lambda_c} \Phi - \gamma \Phi; \\ \Phi(0) = 1, \end{cases} \tag{3.18}$$

in $\mathtt{F}^\star$.

Proof. Let $\{\rho_n\}_{n \geq 0}$ be a classical solution to equation (3.15)-(1.2); this means in particular that $\rho_0 \in C^1([0, +\infty[)$ and $\rho_n \in C^{1,2}([0, +\infty[\times [0,1]^n)$ for all $n \geq 1$. Recall also that according to Assumption 2.1, for any $n \geq 2$ and $t \geq 0$ the function $\rho_n(t, \cdot)$ is symmetric in its arguments. This allows us to consider the multiple Itô integrals $I_n(\rho_n(t, \cdot))$, formally defined in appendix (3.31), and to interchange the partial derivative ∂_t with the iterated integrals. Furthermore, employing the operators $d\Gamma(-\mathcal{A})$, D_1, $\mathtt{N}$ and $D^\star_{\lambda_c}$, whose definitions and properties can be found in the Appendix below, and equation (3.15), we can write for all $n \geq 1$ and $t \geq 0$ that

$$\begin{aligned} \partial_t I_n(\rho_n(t, \cdot)) =& I_n(\partial_t \rho_n(t, \cdot)) \\ =& d\Gamma(-\mathcal{A}) I_n(\rho_n(t, \cdot)) + \frac{\lambda_d}{2} D_1^2 I_{n+2}(\rho_{n+2}(t, \cdot)) - \frac{\lambda_d}{2} \mathtt{N}(\mathtt{N} - \mathtt{I}) I_n(\rho_n(t, \cdot)) \\ &+ D^\star_{\lambda_c} I_{n-1}(\rho_{n-1}(t, \cdot)) - \gamma I_n(\rho_n(t, \cdot)). \end{aligned}$$

If we now sum over $n \geq 0$ and recall that $\rho_{-1} \equiv 0$ while D_1^2 maps to zero any multiple Itô integral of order less than two, we obtain equation (3.18) for the stochastic process defined in (3.17). □

Remark 3.4. *It is worth to point out that condition (2.6) is already encoded in equation (3.18). In fact, if* $\{\Phi(t)\}_{t \geq 0}$ *solves (3.18), then*

$$\partial_t \langle\langle \Phi, \mathcal{E}(1) \rangle\rangle = \langle\langle \partial_t \Phi, \mathcal{E}(1) \rangle\rangle$$

$$
\begin{aligned}
=&\langle\langle d\Gamma(-\mathcal{A})\Phi, \mathcal{E}(1)\rangle\rangle \\
&+ \frac{\lambda_d}{2}\langle\langle D_1^2\Phi, \mathcal{E}(1)\rangle\rangle - \frac{\lambda_d}{2}\langle\langle \mathtt{N}(\mathtt{N}-\mathtt{I})\Phi, \mathcal{E}(1)\rangle\rangle \\
&+ \langle\langle D^\star_{\lambda_c}\Phi, \mathcal{E}(1)\rangle\rangle - \gamma\langle\langle \Phi, \mathcal{E}(1)\rangle\rangle.
\end{aligned}
$$

Now,

$$
\langle\langle d\Gamma(-\mathcal{A})\Phi, \mathcal{E}(1)\rangle\rangle = \langle\langle \Phi, d\Gamma(-\mathcal{A})\mathcal{E}(1)\rangle\rangle = \langle\langle \Phi, \mathcal{E}(1)\diamond\delta(-\mathcal{A}1)\rangle\rangle = 0,
$$

while

$$
\begin{aligned}
\langle\langle D_1^2\Phi, \mathcal{E}(1)\rangle\rangle - \langle\langle \mathtt{N}(\mathtt{N}-\mathtt{I})\Phi, \mathcal{E}(1)\rangle\rangle =&\langle\langle \Phi, D_1^\star D_1^\star \mathcal{E}(1)\rangle\rangle - \langle\langle \Phi, \mathtt{N}(\mathtt{N}-\mathtt{I})\mathcal{E}(1)\rangle\rangle \\
=&\langle\langle \Phi, \mathcal{E}(1)\diamond\delta(1)\diamond\delta(1)\rangle\rangle - \langle\langle \Phi, \mathtt{N}^2\mathcal{E}(1) - \mathtt{N}\mathcal{E}(1)\rangle\rangle \\
=&\langle\langle \Phi, \mathcal{E}(1)\diamond\delta(1)\diamond\delta(1)\rangle\rangle - \langle\langle \Phi, \mathcal{E}(1)\diamond\delta(1)\diamond\delta(1)\rangle\rangle \\
=&0.
\end{aligned}
$$

Here, in the last equality we utilized the identities

$$
\mathtt{N}\mathcal{E}(1) = \mathcal{E}(1)\diamond\delta(1) \quad \textit{and} \quad \mathtt{N}^2\mathcal{E}(1) = \mathcal{E}(1)\diamond\delta(1)\diamond\delta(1) + \mathcal{E}(1)\diamond\delta(1).
$$

Lastly,

$$
\begin{aligned}
\langle\langle D^\star_{\lambda_c}\Phi, \mathcal{E}(1)\rangle\rangle - \gamma\langle\langle \Phi, \mathcal{E}(1)\rangle\rangle =&\langle\langle \Phi, D_{\lambda_c}\mathcal{E}(1)\rangle\rangle - \gamma\langle\langle \Phi, \mathcal{E}(1)\rangle\rangle \\
=&\int_{[0,1]} \lambda_c(x)dx\cdot\langle\langle \Phi, \mathcal{E}(1)\rangle\rangle - \gamma\langle\langle \Phi, \mathcal{E}(1)\rangle\rangle \\
=&0.
\end{aligned}
$$

This proves that $\partial_t\langle\langle \Phi(t), \mathcal{E}(1)\rangle\rangle = 0$; since $\Phi(0) = 1$ (recall the initial condition in (3.18) which in turn follows from (1.2)) we deduce that $\langle\langle \Phi(0), \mathcal{E}(1)\rangle\rangle = 1$ and hence

$$
\langle\langle \Phi(t), \mathcal{E}(1)\rangle\rangle = 1, \quad \textit{for all } t \geq 0. \tag{3.19}
$$

On the other hand, by definition of dual pairing we can write

$$
\langle\langle \Phi(t), \mathcal{E}(1)\rangle\rangle = \sum_{n\geq 0}\int_{[0,1]^n} \rho_n(t, x_1, ..., x_n)dx_1\cdots dx_n,
$$

which together with (3.19) *implies* (2.6)*. Notice that a similar calculation can be carried also when λ_d is not constant; in this case the operator $\Phi \mapsto \mathtt{N}(\mathtt{N}-\mathtt{I})\Phi$ should be replaced with $\Phi \mapsto \delta^2(\lambda(\cdot,\cdot)D^2\Phi)$. This shows that condition* (2.6) *is part of equation (3.15) also in the absence of Assumption 3.1.*

The usefulness of transforming the CDME (3.15)-(1.2) into the abstract problem (3.18) becomes apparent when we consider suitable finite dimensional projections of the stochastic process $\{\Phi(t)\}_{t\geq 0}$. To this aim, we set for $k \geq 1$

$$
\xi_k(x) = \sqrt{2}\cos((k-1)\pi x), x \in [0,1] \quad \text{and} \quad \alpha_k = (k-1)^2\pi^2 \tag{3.20}
$$

to be the eigenfunctions with corresponding eigenvalues of the differential operator $-\mathcal{A} = \partial_x^2$ with homogeneous Neumann boundary conditions (as prescribed in (1.2)). We also write Π_N for the orthogonal projection onto the linear span of $\{\xi_1, ..., \xi_N\}$.

Lemma 3.5. *If $\{\Phi(t)\}_{t\geq 0}$ solves (3.18) in $\mathtt{F}^\star$, then*

$$
\Phi_N(t) := \Gamma(\Pi_N)\Phi(t), \quad t \geq 0 \tag{3.21}
$$

solves

$$
\begin{cases}
\partial_t\Phi_N = d\Gamma(-\mathcal{A})\Phi_N + \frac{\lambda_d}{2}D_1^2\Phi_N - \frac{\lambda_d}{2}\mathtt{N}(\mathtt{N}-\mathtt{I})\Phi_N + D^\star_{\Pi_N\lambda_c}\Phi_N - \gamma\Phi_N; \\
\Phi_N(0) = 1,
\end{cases} \tag{3.22}
$$

in $\mathtt{F}^\star$.

Proof. Using (3.21) and (3.18) we can write

$$\begin{aligned}\partial_t \Phi_N &= \partial_t \Gamma(\Pi_N)\Phi = \Gamma(\Pi_N)\partial_t \Phi \\ &= \Gamma(\Pi_N)\left(d\Gamma(-\mathcal{A})\Phi + \frac{\lambda_d}{2} D_1^2 \Phi - \frac{\lambda_d}{2}\mathtt{N}(\mathtt{N}-\mathtt{I})\Phi + D^\star_{\lambda_c}\Phi - \gamma\Phi \right). \end{aligned} \tag{3.23}$$

The proof consists in showing that our assumptions allow for the commutation between the operator $\Gamma(\Pi_N)$ and each of the following: $d\Gamma(-\mathcal{A})$, D_1^2, $\mathtt{N}(\mathtt{N}-\mathtt{I})$ and $D^\star_{\lambda_d}$. Let us start with the commutation between $\Gamma(\Pi_N)$ and $d\Gamma(-\mathcal{A})$: for any smooth $h \in L^2([0,1])$ we have

$$\begin{aligned}\langle\langle \Gamma(\Pi_N) d\Gamma(-\mathcal{A})\Phi, \mathcal{E}(h)\rangle\rangle =& \langle\langle d\Gamma(-\mathcal{A})\Phi, \mathcal{E}(\Pi_N h)\rangle\rangle \\ =& \langle\langle \Phi, d\Gamma(-\mathcal{A})\mathcal{E}(\Pi_N h)\rangle\rangle \\ =& \langle\langle \Phi, \mathcal{E}(\Pi_N h)\diamond \delta(-\mathcal{A}\Pi_N h)\rangle\rangle \\ =& \langle\langle \Phi, \mathcal{E}(\Pi_N h)\diamond \delta(\Pi_N(-\mathcal{A})h)\rangle\rangle \\ =& \langle\langle \Phi, \Gamma(\Pi_N)(\mathcal{E}(h)\diamond \delta(-\mathcal{A}h))\rangle\rangle \\ =& \langle\langle \Gamma(\Pi_N)\Phi, \mathcal{E}(h)\diamond \delta(-\mathcal{A}h)\rangle\rangle \\ =& \langle\langle d\Gamma(-\mathcal{A})\Gamma(\Pi_N)\Phi, \mathcal{E}(h)\rangle\rangle \\ =& \langle\langle d\Gamma(-\mathcal{A})\Phi_N, \mathcal{E}(h)\rangle\rangle. \end{aligned}$$

Comparing the first and last members of this chain of equalities we deduce that

$$\Gamma(\Pi_N) d\Gamma(-\mathcal{A})\Phi = d\Gamma(-\mathcal{A})\Phi_N, \quad \text{in } \mathtt{F}^\star.$$

It is important to observe how in the fourth equality above the commutation between $-\mathcal{A}$ and Π_N is made possible by having chosen to project onto the space generated by the eigenfunctions of $-\mathcal{A}$.
We now study the commutation between $\Gamma(\Pi_N)$ and D_1^2:

$$\begin{aligned}\langle\langle \Gamma(\Pi_N) D_1^2\Phi, \mathcal{E}(h)\rangle\rangle =& \langle\langle D_1^2\Phi, \mathcal{E}(\Pi_N h)\rangle\rangle \\ =& \langle\langle D_1 D_1\Phi, \mathcal{E}(\Pi_N h)\rangle\rangle \\ =& \langle\langle \Phi, D_1^\star D_1^\star \mathcal{E}(\Pi_N h)\rangle\rangle \\ =& \langle\langle \Phi, \mathcal{E}(\Pi_N h)\diamond \delta(1)\diamond\delta(1)\rangle\rangle \\ =& \langle\langle \Phi, \Gamma(\Pi_N)(\mathcal{E}(h)\diamond \delta(1)\diamond\delta(1))\rangle\rangle \\ =& \langle\langle \Gamma(\Pi_N)\Phi, \mathcal{E}(h)\diamond \delta(1)\diamond\delta(1)\rangle\rangle \\ =& \langle\langle D_1 D_1\Gamma(\Pi_N)\Phi, \mathcal{E}(h)\rangle\rangle \\ =& \langle\langle D_1^2\Gamma(\Pi_N)\Phi, \mathcal{E}(h)\rangle\rangle. \end{aligned}$$

In the fifth equality above we employed the identity $\Pi_N 1 = \Pi_1 1 = 1$ since the first eigenfunction of $-\mathcal{A}$ is precisely 1. We therefore can conclude that

$$\Gamma(\Pi_N) D_1^2\Phi = D_1^2\Phi_N, \quad \text{in } \mathtt{F}^\star.$$

We proceed with the commutation between $\Gamma(\Pi_N)$ and $\mathtt{N}(\mathtt{N}-\mathtt{I})$:

$$\begin{aligned}\langle\langle \Gamma(\Pi_N)\mathtt{N}(\mathtt{N}-\mathtt{I})\Phi, \mathcal{E}(h)\rangle\rangle &= \langle\langle \mathtt{N}(\mathtt{N}-\mathtt{I})\Phi, \mathcal{E}(\Pi_N h)\rangle\rangle \\ &= \langle\langle \Phi, \mathcal{E}(\Pi_N h)\diamond\delta(\Pi_N h)\diamond\delta(\Pi_N h)\rangle\rangle \\ &= \langle\langle \Phi, \Gamma(\Pi_N)(\mathcal{E}(h)\diamond\delta(h)\diamond\delta(h))\rangle\rangle \\ &= \langle\langle \Phi_N, \mathcal{E}(h)\diamond\delta(h)\diamond\delta(h)\rangle\rangle \\ &= \langle\langle \Phi_N, \mathtt{N}(\mathtt{N}-\mathtt{I})\mathcal{E}(h)\rangle\rangle \\ &= \langle\langle \mathtt{N}(\mathtt{N}-\mathtt{I})\Phi_N, \mathcal{E}(h)\rangle\rangle. \end{aligned}$$

This yields

$$\Gamma(\Pi_N)\mathtt{N}(\mathtt{N}-\mathtt{I})\Phi = \mathtt{N}(\mathtt{N}-\mathtt{I})\Phi_N, \quad \text{in } \mathtt{F}^\star.$$

Lastly,

$$\langle\langle \Gamma(\Pi_N) D^\star_{\lambda_c}\Phi, \mathcal{E}(h)\rangle\rangle = \langle\langle D^\star_{\lambda_c}\Phi, \mathcal{E}(\Pi_N h)\rangle\rangle$$

$$
\begin{aligned}
&=\langle\langle \Phi, \mathcal{E}(\Pi_N h)\rangle\rangle \int_0^1 \Pi_N h(x)\lambda_c(x)dx\\
&=\langle\langle \Gamma(\Pi_N)\Phi, \mathcal{E}(h)\rangle\rangle \int_0^1 h(x)\Pi_N\lambda_c(x)dx\\
&=\langle\langle \Phi_N, D_{\Pi_N\lambda_c}\mathcal{E}(h)\rangle\rangle\\
&=\langle\langle D^\star_{\Pi_N\lambda_c}\Phi_N, \mathcal{E}(h)\rangle\rangle,
\end{aligned}
$$

and hence

$$\Gamma(\Pi_N)D^\star_{\lambda_c}\Phi = D^\star_{\Pi_N\lambda_c}\Phi_N, \quad \text{in } \mathtt{F}^\star.$$

An implementation of all derived identities in (3.23) leads directly to (3.22). □

Remark 3.6. *The previous lemma is a key ingredient of our method since it allows for finite dimensional projections of the solution to equation (3.18). In particular, it is the possibility of commuting $d\Gamma(-\mathcal{A})$ and $\mathtt{N}(\mathtt{N}-\mathtt{I})$ with $\Gamma(\Pi_N)$ that implies the desired result. It is worth to mention that such possibility exists because of Assumption 3.1, thus motivating this strong simplification. In fact, without such assumption we would not be able to commute $\Gamma(\Pi_N)$ with the operator $\Phi \mapsto \delta^2(\lambda(\cdot,\cdot)D^2\Phi)$ which is what one should work with, in the place of $\mathtt{N}(\mathtt{N}-\mathtt{I})$, for non constant λ_d. One may also wonder whether changing the projection space related to Π_N could solve this issue (maybe defining a finite dimensional space described by the function λ_d): however, this modification would imply the loss of commutativity between $\Gamma(\Pi_N)$ and $d\Gamma(-\mathcal{A})$.*

Lemma 3.7. *For any $N \geq 1$ there exists a function $u_N : [0+\infty[\times\mathbb{R}^N$ such that*

$$\Phi_N(t) = u_N(t, I_1(\xi_1), ..., I_1(\xi_N)), \quad \mathbb{P}\text{-a.s.}$$

Furthermore, the function u_N solves (weakly) the following fourth order linear problem:

$$
\begin{cases}
\partial_t u_N(t,z) = -\sum_{k=1}^N \alpha_k \partial_k^\star \partial_k u_N(t,z) + \frac{\lambda_d}{2}\partial_1^2 u_N(t,z)\\
\qquad\qquad -\frac{\lambda_d}{2}\sum_{j,k=1}^N \partial_j^\star\partial_k^\star\partial_j\partial_k u_N(t,z)\\
\qquad\qquad +\sum_{k=1}^N \mathtt{c}_k\partial_k^\star u_N(t,z) - \gamma u_N(t,z); \quad t>0, z\in\mathbb{R}^N;\\
u_N(0,z) = 1, \quad z\in\mathbb{R}^N.
\end{cases}
\tag{3.24}
$$

Here, for any $k \in \{1, ..., N\}$ the symbol ∂_k is a shorthand notation for ∂_{z_k} while $\partial_k^\star$ stands for the differential operator $-\partial_k + z_k$ (which is nothing else that the Gaussian divergence). Moreover, $\mathtt{c}_k := \langle\lambda_c, \xi_k\rangle_{L^2([0,1])}$.

Proof. It is well known (see for instance Theorem 4.9 in[18]) that the second quantization operator $\Gamma(\Pi_N)$ corresponds to the conditional expectation with respect to the sigma-algebra generated by the random variables $I_1(\xi_1)$,..., $I_1(\xi_N)$; therefore, according to (3.21) we can write

$$\Phi_N(t) = \Gamma(\Pi_N)\Phi(t) = \mathbb{E}[\Phi(t)|\sigma(I_1(\xi_1), ..., I_1(\xi_N))] = u_N(t, I_1(\xi_1), ..., I_1(\xi_N)).$$

Here, the function $u_N : [0+\infty[\times\mathbb{R}^N$ is measurable and its existence in guaranteed by Doob's lemma (see Lemma 1.13 in[20]). We now replace $\Phi_N(t)$ with $u_N(t, I_1(\xi_1), ..., I_1(\xi_N))$ in (3.22) and decompose the Malliavin calculus's operators along the orthonormal bases $\{\xi_k\}_{k\geq 1}$. More precisely:

$$
\begin{aligned}
d\Gamma(-\mathcal{A})[u_N(t, I_1(\xi_1), ..., I_1(\xi_N))] =&\delta\left(-\mathcal{A}D[u_N(t, I_1(\xi_1), ..., I_1(\xi_N))]\right)\\
=&\delta\left(-\mathcal{A}\sum_{k\geq 1} D_{\xi_k}u_N(t, I_1(\xi_1), ..., I_1(\xi_N))\xi_k\right)\\
=&\delta\left(-\mathcal{A}\sum_{k=1}^N \partial_k u_N(t, I_1(\xi_1), ..., I_1(\xi_N))\xi_k\right)\\
=&\delta\left(-\sum_{k=1}^N \alpha_k\partial_k u_N(t, I_1(\xi_1), ..., I_1(\xi_N))\xi_k\right)\\
=&-\sum_{k=1}^N \alpha_k\partial_k u_N(t, I_1(\xi_1), ..., I_1(\xi_N))I_1(\xi_k)
\end{aligned}
$$

$$+\sum_{k=1}^{N}\alpha_k\partial_k^2 u_N(t,I_1(\xi_1),...,I_1(\xi_N))$$
$$=-\sum_{k=1}^{N}\alpha_k\partial_k^\star\partial_k u_N(t,I_1(\xi_1),...,I_1(\xi_N));$$

Here, in the second-to-last equality we employed identity (1.56) from[27]. We proceed now with

$$D_1^2 u_N(t,I_1(\xi_1),...,I_1(\xi_N))=\partial_1^2 u_N(t,I_1(\xi_1),...,I_1(\xi_N)),$$

and

$$\begin{aligned}\mathtt{N}(\mathtt{N}-\mathtt{I})u_N(t,I_1(\xi_1),...,I_1(\xi_N))=&(\mathtt{N}^2-\mathtt{N})u_N(t,I_1(\xi_1),...,I_1(\xi_N))\\ =&\sum_{j=1}^{N}\partial_j^\star\partial_j\left(\sum_{k=1}^{N}\partial_k^\star\partial_k u_N(t,I_1(\xi_1),...,I_1(\xi_N))\right)\\ &-\sum_{k=1}^{N}\partial_k^\star\partial_k u_N(t,I_1(\xi_1),...,I_1(\xi_N))\\ =&\sum_{j,k=1}^{N}\partial_j^\star\partial_k^\star\partial_j\partial_k u_N(t,I_1(\xi_1),...,I_1(\xi_N)).\end{aligned}$$

In the last equality we employed the commutation relation $\partial_j\partial_k^\star=\partial_k^\star\partial_j+\delta_{jk}\mathtt{I}$ where δ_{jk} stands for the Kronecker symbol. Lastly,

$$\begin{aligned}D^\star_{\Pi_N\lambda_c}u_N(t,I_1(\xi_1),...,I_1(\xi_N))=&\sum_{k\geq 1}D^\star_{\xi_k}u_N(t,I_1(\xi_1),...,I_1(\xi_N))\langle\Pi_N\lambda_c,\xi_k\rangle_{L^2([0,1])}\\ =&\sum_{k=1}^{N}\partial_k^\star u_N(t,I_1(\xi_1),...,I_1(\xi_N))\langle\Pi_N\lambda_c,\xi_k\rangle_{L^2([0,1])}\\ =&\sum_{k=1}^{N}\partial_k^\star u_N(t,I_1(\xi_1),...,I_1(\xi_N))\mathtt{c}_k.\end{aligned}$$

Collecting all identities derived above we see how equation (3.22) is equivalent to (3.24). □

We are now ready to state the main result of the present section.

Theorem 3.8. *Let $\{\rho_n\}_{n\geq 0}$ be a classical solution to the CDME (3.15)-(1.2). Then, for any $N\geq 1$ we have the representation*

$$\Pi_N^{\otimes n}\rho_n(t,x_1,...,x_n)=\frac{1}{n!}\sum_{j_1,...,j_n=1}^{N}\mathbb{E}[(\partial_{j_1}\cdots\partial_{j_n}u_N)(t,I_1(\xi_1),...,I_1(\xi_n))]\xi_{j_1}(x_1)\cdots\xi_{j_n}(x_n),\tag{3.25}$$

for all $n\geq 1$, $t\geq 0$, $(x_1,...,x_n)\in[0,1]^n$ and with u_N solution to the Cauchy problem (3.24).

Proof. If $\{\rho_n\}_{n\geq 0}$ is a classical solution to the CDME (3.15)-(1.2), then according to Lemma 3.3 the stochastic process $\{\Phi(t)\}_{t\geq 0}$ defined in (3.17) solves equation (3.18) in $\mathtt{F}^\star$. Moreover, Lemma 3.5 shows that the finite dimensional projection of $\{\Phi(t)\}_{t\geq 0}$ introduced in (3.21) solves the auxiliary problem (3.22). Notice that by construction the kernels of the Wiener Itô chaos expansion of $\{\Phi_N(t)\}_{t\geq 0}$ are $\{\Pi_N^{\otimes}\rho_n\}_{n\geq 0}$ (since this is the action of $\Gamma(\Pi_N)$ on $\{\Phi(t)\}_{t\geq 0}$).
On the other hand, according to the Stroock-Taylor formula (see Exercise 1.2.6 in[27]) the Wiener Itô chaos expansion of $\{\Phi_N(t)\}_{t\geq 0}$ can also be represented as

$$\Pi_N^{\otimes}\rho_n(t,x_1,...,x_n)=\frac{1}{n!}\mathbb{E}[D_{x_1}\cdots D_{x_n}\Phi_N(t)].\tag{3.26}$$

From Lemma 3.7 the process $\{\Phi_N(t)\}_{t\geq 0}$ can be written as $\{u_N(t, I_1(\xi_1), ..., I_1(\xi_n))\}_{t\geq 0}$ where u_N solution to the Cauchy problem (3.24). Therefore, substituting this into (3.26) and computing the Malliavin derivatives yields

$$\begin{aligned}\Pi_N^{\otimes}\rho_n(t, x_1, ..., x_n) =& \frac{1}{n!}\mathbb{E}[D_{x_1}\cdots D_{x_n}u_N(t, I_1(\xi_1), ..., I_1(\xi_n))] \\ =& \frac{1}{n!}\sum_{j_1,...,j_n=1}^{N}\mathbb{E}[(\partial_{j_1}\cdots\partial_{j_n}u_N)(t, I_1(\xi_1), ..., I_1(\xi_n))]\xi_{j_1}(x_1)\cdots\xi_{j_n}(x_n),\end{aligned}$$

which is the formula we wanted to prove. □

A. Some comments on the PDE (3.24)

The generalization of the generating function method utilized in this section has introduced some new Gaussian features to the original problem (3.15)-(1.2). At a formal level, the infinite dimensional nature of the system of Fokker-Planck equations under investigation combined with the Fock space structure of the sequence $\{\rho_n\}_{n\geq 0}$ leads naturally to the use of Gaussian stochastic analysis's techniques.
We now try to rewrite the representation formula (3.25) in a Gaussian-free manner. To this aim, we present the following technical result.

Lemma 3.9. *Let $f \in C^1(\mathbb{R}^N)$ be, together with all its first order partial derivatives, polynomially bounded at infinity. Then, setting*

$$\tilde{f}(z) := \int_{\mathbb{R}^N} f(y)(2\pi)^{-N/2}e^{-|z-y|^2/2}dy, \quad z \in \mathbb{R}^N,$$

we have for all $k \in \{1, ..., N\}$ and $z \in \mathbb{R}^N$ that

$$\widetilde{\partial_k f}(z) = \partial_k \tilde{f}(z) \quad \text{and} \quad \widetilde{\partial_k^\star f}(z) = z_k \tilde{f}(z) \tag{3.27}$$

Proof. It is a direct verification. □

Proposition 3.10. *If u_N solves the PDE (3.24), then $v_N := \tilde{u}_N$ solves*

$$\begin{cases}\partial_t v_N(t,z) = -\sum_{k=1}^N \alpha_k z_k \partial_k v_N(t,z) + \frac{\lambda_d}{2}\partial_1^2 v_N(t,z) \\ \qquad\qquad -\frac{\lambda_d}{2}\sum_{j,k=1}^N z_j z_k \partial_j\partial_k v_N(t,z) \\ \qquad\qquad +\sum_{k=1}^N \mathsf{c}_k z_k v_N(t,z) - \gamma v_N(t,z); \quad t>0, z\in\mathbb{R}^N; \\ v_N(0,z) = 1, \quad z \in \mathbb{R}^N.\end{cases} \tag{3.28}$$

Proof. Follows immediately from (3.27). □

The PDE (3.28) represents a version of (3.24) in which the Gaussian features inherited from our approach have been removed. Equation (3.28) has certainly the advantage over (3.24) of being of second order (contrary to the fourth order of the latter); moreover, if we consider the case $N = 1$ and remember that $\xi_1 \equiv 1$ and $\alpha_1 = 0$ we obtain

$$\begin{cases}\partial_t v_1(t,z) = \frac{\lambda_d}{2}(1-z_1^2)\partial_1^2 v_1(t,z) + \gamma(z_1-1)v_1(t,z); \quad t>0, z\in\mathbb{R}; \\ v_1(0,z) = 1, \quad z\in\mathbb{R}.\end{cases} \tag{3.29}$$

This is exactly the equation you obtain via the classical generating function method applied to the CME (3.16), which is the diffusion-free analogue of our system (3.15)-(1.2). See[12] for a detailed study of (3.29). Therefore, from this point of view equation (3.28) is the natural extension of (3.29) to a model that includes diffusion of the particles.
Even though equation (3.28) possesses some desirable properties, its investigation from both analytical and numerical points of view presents some important obstacles. First of all, if we use the function v_N solution to (3.28) in the place of u_N solution to (3.24), then the representation formula (3.25) takes the form

$$\Pi_N^{\otimes n}\rho_n(t, x_1, ..., x_n) = \frac{1}{n!}\sum_{j_1,...,j_n=1}^{N}(\partial_{j_1}\cdots\partial_{j_n}v_N)(t, 0, ..., 0)\xi_{j_1}(x_1)\cdots\xi_{j_n}(x_n), \tag{3.30}$$

as it follows immediately by the definition of v_N and Lemma 3.9. This means that the natural domain for solving (3.28) would be a neighborhood of the origin instead of the whole space; this can be seen already in the case $N = 1$, i.e. equation (3.29), where z_1 should be taken in $[-1, 1]$ in order to avoid a sign change in the leading second order term. However, it is very hard to find a reasonable argument for assigning a boundary value to the problem (3.28) (this issue is also discussed in[25]).
A second main difficulty in analyzing equation (3.28) is due to its intrinsic ill-posedness. In fact, if for simplicity we take $N = 2$ and focus on the second order (i.e. leading) term of the differential operator appearing in the right hand side of (3.28), we see that the matrix describing its coefficients is a multiple of

$$A(z_1, z_2) = \begin{bmatrix} 1 - z_1^2 & -z_1 z_2 \\ -z_1 z_2 & -z_2^2 \end{bmatrix}.$$

Checking the positive semi-definiteness of the matrix A (recall that where have an initial condition for solving equation (3.28)) we see that

$$\begin{aligned} \langle A(z_1, z_2)\theta, \theta\rangle =& (1 - z_1^2)\theta_1^2 - 2z_1 z_2 \theta_1 \theta_2 - z_2^2 \theta_2^2 \\ =& \theta_1^2 - (z_1\theta_1 + z_2\theta_2)^2 \end{aligned}$$

and the last quantity cannot be non negative for any choice of $(\theta_1, \theta_2) \in \mathbb{R}^2$ unless $z_2 = 0$ (to see this take $\theta_1 = 0$). Therefore, there is no open neighborhood of the origin for the space variable z where the matrix A is positive semi-definiteness. This entails the ill-posedness of the PDE (3.28).
The discussion presented above highlights some potential advantages in embedding the CDME (3.15)-(1.2) into the Gaussian framework utilized in this section for deriving the representation formula (3.25).

APPENDIX

In this section we collect some definitions and formulas utilized in the proofs of Section 3. For more details on the subject we refer the reader to one the books[18] and[27].

Wiener chaos and spaces of random variables

Let $(\Omega, \mathcal{B}, \mathbb{P})$ be the classical Wiener space over the interval $[0, 1]$. We denote by

$$\begin{aligned} B_x : \Omega &\to \mathbb{R} \\ \omega &\mapsto B_x(\omega) := \omega(x), \quad x \in [0, 1], \end{aligned}$$

the coordinate process which by construction is a one dimensional Brownian motion under $\mathbb{P}$. According to the Wiener-Itô chaos expansion theorem, any random variable Φ in $\mathbb{L}^2(\Omega)$ can be uniquely represented as

$$\Phi = \sum_{n\geq 0} I_n(h_n),$$

where for $n \geq 1$, $I_n(h_n)$ stands for the n-th order multiple Itô integral defined as

$$I_n(h_n) := n! \int_0^1 \int_0^{x_1} \cdots \int_0^{x_{n-1}} h_n(x_1, ..., x_n) dB_{x_n} \cdots dB_{x_2} dB_{x_1}. \tag{3.31}$$

Two notable dense subset of $\mathbb{L}^2(\Omega)$ are

$$\mathtt{F} := \left\{ \sum_{n=0}^{M} I_n(h_n), \text{ for some } M \in \mathbb{N} \cup \{0\}, \, h_0 \in \mathbb{R} \text{ and } h_n \in L_s^2([0, 1]^n), \, n = 1, ..., M \right\},$$

which collects the random variables with a finite order chaos expansion, and

$$\mathtt{E} := \left\{ \mathcal{E}(f) := \sum_{n\geq 0} I_n\left(\frac{f^{\otimes n}}{n!}\right), \text{ for some } f \in L^2([0, 1]) \right\},$$

which is the family of the so-called *stochastic exponentials*.

Malliavin derivative and its adjoint

The *Malliavin derivative* of $\Phi = \sum_{n=0}^{M} I_n(h_n) \in \mathtt{F}$, denoted $\{D_x\Phi\}_{x\in[0,1]}$, is the element of $L^2([0,1];\mathtt{F})$ defined by

$$D_x\Phi := \sum_{n=0}^{M-1} (n+1) I_n(h_{n+1}(\cdot, x)), \quad x \in [0,1].$$

For $l \in L^2([0,1])$ and $\Phi = \sum_{n=0}^{M} I_n(h_n) \in \mathtt{F}$, we also write

$$\begin{aligned} D_l\Phi := \langle D\Phi, l\rangle_{L^2([0,1])} &= \sum_{n=0}^{M-1} (n+1) I_n\left(\int_0^1 h_{n+1}(\cdot, y) l(y) dy\right) \\ &= \sum_{n=0}^{M-1} (n+1) I_n\left(h_{n+1} \otimes_1 l\right) \end{aligned}$$

for the *directional Malliavin derivative* of Φ along l. Here, we denote the *r-th order contraction* of h_n and h_m by $h_n \otimes_r h_m$, i.e.

$$\begin{aligned} &(h_n \otimes_r h_m)(x_1, ..., x_{n+m-2r}) \\ &:= \int_{[0,1]^r} h_n(x_1, ..., x_{n-r}, y_1, ..., y_r) h_m(y_1, ..., y_r, x_{n-r+1}, ..., x_{n+m-2r}) dy_1 \cdots dy_r. \end{aligned}$$

We have:

$$D_x\mathcal{E}(f) = f(x)\mathcal{E}(f), x \in [0,1] \quad \text{and} \quad D_l\mathcal{E}(f) = \langle f, l\rangle_{L^2([0,1])}\mathcal{E}(f).$$

If we now take $l \in L^2([0,1])$, $\Phi = \sum_{n=0}^{M} I_n(h_n) \in \mathtt{F}$ and $\Psi = \sum_{n=0}^{K} I_n(g_n) \in \mathtt{F}$, we can write

$$\mathbb{E}[D_l\Phi \cdot \Psi] = \mathbb{E}[\Phi \cdot D_l^\star\Psi],$$

where

$$D_l^\star\Psi := \sum_{n=1}^{K+1} I_n(l\hat{\otimes} g_{n-1})$$

and

$$(l\hat{\otimes} g_{n-1})(x_1, ..., x_n) := \frac{1}{n}\sum_{i=1}^{n} f(x_i) g_{n-1}(x_1, ..., x_{i-1}, x_{i+1}, ..., x_n).$$

The following identity holds:

$$D_l^\star\Psi + D_l\Psi = \Psi \cdot I_1(l).$$

One can also introduce the adjoint of D_x, denoted δ:

$$\delta(\Phi(\cdot)) := \sum_{n=0}^{M} I_{n+1}(\tilde{h}_n) \in \mathtt{F},$$

where $\tilde{h}_n$ stands for the symmetrization of h_n with respect to the $n+1$ variables $x_1, ..., x_n, x$. We mentioned that $D_l^\star\Psi$ it is sometimes written as $\Phi \diamond \delta(l)$.

Second quantization operators

Let $A : L^2([0,1]) \to L^2([0,1])$ be a bounded linear operator; for $\Phi = \sum_{n=0}^{M} I_n(h_n) \in \mathtt{F}$ we define the *second quantization operator* of A as

$$\Gamma(A)\Phi := \sum_{n=0}^{M} I_n\left(A^{\otimes n} h_n\right),$$

and the *differential second quantization operator* of A as

$$d\Gamma(A)\Phi := \sum_{n=1}^{M} I_n\left(\sum_{i=1}^{n} A_i h_n\right),$$

where A_i stands for the operator A acting on the i-th variable of h_n. Notice in addition that for A being the identity, we recover from $d\Gamma(A)$ the well known *number operator*:

$$\mathtt{N}\Phi = \sum_{n=1}^{M} nI_n\left(h_n\right).$$

The following identities hold true:

$$\begin{array}{l}\mathbb{E}[\Gamma(A)\Phi] = \mathbb{E}[\Phi]; \quad \mathbb{E}[d\Gamma(A)\Phi] = 0;\\ \mathbb{E}[\Gamma(A)\Phi\cdot\Psi] = \mathbb{E}[\Phi\cdot\Gamma(A^\star)\Psi]; \quad \mathbb{E}[d\Gamma(A)\Phi\cdot\Psi] = \mathbb{E}[\Phi\cdot d\Gamma(A^\star)\Psi];\\ \Gamma(A)\mathcal{E}(f) = \mathcal{E}(Af); \quad d\Gamma(A)\mathcal{E}(f) = D^\star_{Af}\mathcal{E}(f); \quad d\Gamma(A)\Phi = \delta\left(AD.\Phi\right).\end{array}$$

A space of generalized random variables

Let

$$\mathtt{F}^\star := \left\{\sum_{n\geq 0} I_n(h_n),\ \text{for some } h_0\in\mathbb{R} \text{ and } h_n\in L^2_s([0,1]^n),\ n\geq 1\right\}$$

be a family of *generalized* random variables. The action of $T=\sum_{n\geq 0} I_n(h_n)\in\mathtt{F}^\star$ on $\varphi=\sum_{n=0}^{M} I_n(g_n)\in\mathtt{F}$ is defined as

$$\langle\langle T,\varphi\rangle\rangle := \sum_{n=0}^{M} n!\langle h_n, g_n\rangle_{L^2([0,1]^n)}.$$

By construction, we have the inclusions

$$\mathtt{F}\subset\mathbb{L}^2(\Omega)\subset\mathtt{F}^\star$$

with

$$\langle\langle T,\varphi\rangle\rangle = \mathbb{E}[T\varphi],$$

whenever $T\in\mathbb{L}^2(\Omega)$. We will say that $T=U$ in $\mathtt{F}^\star$ if

$$\langle\langle T,\varphi\rangle\rangle = \langle\langle U,\varphi\rangle\rangle, \quad \text{for all } \varphi\in\mathtt{F}.$$

Let

$$\mathtt{F}^\star := \left\{\sum_{n\geq 0} I_n(h_n),\ \text{for some } h_0\in\mathbb{R} \text{ and } h_n\in L^2_s([0,1]^n),\ n\geq 1\right\}$$

be a family of *generalized* random variables. The action of $T=\sum_{n\geq 0} I_n(h_n)\in\mathtt{F}^\star$ on $\varphi=\sum_{n=0}^{M} I_n(g_n)\in\mathtt{F}$ is defined as

$$\langle\langle T,\varphi\rangle\rangle := \sum_{n=0}^{M} n!\langle h_n, g_n\rangle_{L^2([0,1]^n)}.$$

By construction, we have the inclusions

$$\mathtt{F}\subset\mathbb{L}^2(\Omega)\subset\mathtt{F}^\star$$

with

$$\langle\langle T, \varphi \rangle\rangle = \mathbb{E}[T\varphi],$$

whenever $T \in \mathbb{L}^2(\Omega)$. We will say that $T = U$ in $\mathtt{F}^\star$ if

$$\langle\langle T, \varphi \rangle\rangle = \langle\langle U, \varphi \rangle\rangle, \quad \text{for all } \varphi \in \mathtt{F}.$$

Lastly, we recall a generalized version of the so-called Stroock-Taylor formula: if $T = \sum_{n\geq 0} I_n(h_n) \in \mathtt{F}^\star$, then

$$h_n(x_1, ..., x_n) = \frac{1}{n!}\mathbb{E}[D_{x_1}...D_{x_n}T], \quad (x_1, ..., x_n) \in [0,1]^n.$$

REFERENCES

[1]Anderson, D. F. and Kurtz, T. G. (2015). *Stochastic analysis of biochemical systems*, volume 1. Springer.

[2]Andrews, S. S. (2017). Smoldyn: particle-based simulation with rule-based modeling, improved molecular interaction and a library interface. *Bioinformatics*, 33(5):710–717.

[3]Chen, W., Erban, R., and Chapman, S. J. (2014). From Brownian dynamics to Markov chain: An ion channel example. *SIAM Journal of Applied Mathematics*, 74(1):208–235.

[4]del Razo, M. J., Dibak, M., Schütte, C., and Noé, F. (2021). Multiscale molecular kinetics by coupling markov state models and reaction-diffusion dynamics. *Journal of Chemical Physics*, 155(12):124109.

[5]del Razo, M. J., Frömberg, D., Straube, A. V., Schütte, C., Höfling, F., and Winkelmann, S. (2022). A probabilistic framework for particle-based reaction-diffusion dynamics using classical Fock space representations. *Lett. Math. Phys.*, 112(3):Paper No. 49, 59.

[6]del Razo, M. J., Qian, H., and Noé, F. (2018). Grand canonical diffusion-influenced reactions: A stochastic theory with applications to multiscale reaction-diffusion simulations. *Journal of Chemical Physics*, 149(4):044102.

[7]del Razo, M. J., Winkelmann, S., Klein, R., and Höfling, F. (2023). Chemical diffusion master equation: Formulations of reaction–diffusion processes on the molecular level. *J. Math. Phys.*, 64(1):013304.

[8]Deuflhard, P., Huisinga, W., Jahnke, T., and Wulkow, M. (2008). Adaptive discrete galerkin methods applied to the chemical master equation. *SIAM Journal on Scientific Computing*, 30(6):2990–3011.

[9]Dibak, M., del Razo, M. J., De Sancho, D., Schütte, C., and Noé, F. (2018). Msm/rd: Coupling markov state models of molecular kinetics with reaction-diffusion simulations. *Journal of Chemical Physics*, 148(21):214107.

[10]Doi, M. (1976). Second quantization representation for classical many-particle system. *Journal of Physics A: Mathematical and General*, 9(9):1465.

[11]Engblom, S. (2009). Spectral approximation of solutions to the chemical master equation. *Journal of computational and applied mathematics*, 229(1):208–221.

[12]Erban, R. and Chapman, S. J. (2020). *Stochastic modelling of reaction-diffusion processes*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge.

[13]Flegg, M. B., Chapman, S. J., and Erban, R. (2012). The two-regime method for optimizing stochastic reaction–diffusion simulations. *Journal of the Royal Society Interface*, 9(70):859–868.

[14]Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361.

[15]Gupta, A., Schwab, C., and Khammash, M. (2021). Deepcme: A deep learning framework for computing solution statistics of the chemical master equation. *PLoS computational biology*, 17(12):e1009623.

[16]Hoffmann, M., Fröhner, C., and Noé, F. (2019). Readdy 2: Fast and flexible software framework for interacting-particle reaction dynamics. *PLoS computational biology*, 15(2):e1006830.

[17]Jahnke, T. and Huisinga, W. (2007). Solving the chemical master equation for monomolecular reaction systems analytically. *Journal of mathematical biology*, 54:1–26.

[18]Janson, S. (1997). *Gaussian Hilbert spaces*, volume 129 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge.

[19]Jiang, Q., Fu, X., Yan, S., Li, R., Du, W., Cao, Z., Qian, F., and Grima, R. (2021). Neural network aided approximation and parameter inference of non-markovian models of gene expression. *Nature communications*, 12(1):1–12.

[20]Kallenberg, O. ([2021] ©2021). *Foundations of modern probability*, volume 99 of *Probability Theory and Stochastic Modelling*. Springer, Cham. Third edition [of 1464694].

[21]Kostré, M., Schütte, C., Noé, F., and del Razo, M. J. (2021). Coupling particle-based reaction-diffusion simulations with reservoirs mediated by reaction-diffusion pdes. *SIAM Multiscale Modeling & Simulation*, 19(4):1659–1683.

[22]Lanconelli, A. (2023). Using Malliavin calculus to solve a chemical diffusion master equation. *J. Math. Anal. Appl.*, 526(2):Paper No. 127352.

[23]Lanconelli, A., Perçin, B. T., and del Razo, M. J. (2023). Solution formula for the general birth-death chemical diffusion master equation. *arXiv:2302.10700 [math.PR]*.

[24]Lecca, P., Ian, L., and Jordan, F. (2013). *Deterministic Versus Stochastic Modelling in Biochemistry and Systems Biology*. Oxford: Woodhead Publishing.

[25]McQuarrie, D. A. (1967). Stochastic approach to chemical kinetics. *J. Appl. Probability*, 4:413–478.

[26]Munsky, B. and Khammash, M. (2006). The finite state projection algorithm for the solution of the chemical master equation. *The Journal of chemical physics*, 124(4):044104.

[27]Nualart, D. (2006). *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition.

[28]Qian, H. and Bishop, L. M. (2010). The chemical master equation approach to nonequilibrium steady-state of open biochemical systems: Linear single-molecule etheornzyme kinetics and nonlinear biochemical reaction networks. *Int. J. Mol. Sci.*, 11(9):3472–3500.

[29]Qian, H. and Ge, H. (2021). *Stochastic Chemical Reaction Systems in Biology*. Springer.
[30]Schnoerr, D., Sanguinetti, G., and Grima, R. (2017). Approximation and inference methods for stochastic biochemical kinetics—a tutorial review. *Journal of Physics A: Mathematical and Theoretical*, 50(9):093001.
[31]Schweitzer, F. and Farmer, J. D. (2003). *Brownian agents and active particles: collective dynamics in the natural and social sciences*, volume 1. Springer.
[32]Smith, C. A. and Yates, C. A. (2018). Spatially extended hybrid methods: a review. *Journal of the Royal Society Interface*, 15(139):20170931.
[33]Zhang, X., De Cock, K., Bugallo, M. F., and Djurić, P. M. (2005). A general method for the computation of probabilities in systems of first order chemical reactions. *Journal of Chemical Physics*, 122(10).