HAL open science

A spatial measure-valued model for chemical reaction networks in heterogeneous systems

Lea Popovic, Amandine Véber

To cite this version:

Lea Popovic, Amandine Véber. A spatial measure-valued model for chemical reaction networks in heterogeneous systems. The Annals of Applied Probability, 2023, 33 (5), pp.3706-3754. 10.1214/22-AAP1904 . hal-03905299

HAL Id: hal-03905299

https://hal.science/hal-03905299v1

Submitted on 18 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A spatial measure-valued model for chemical reaction networks in heterogeneous systems

Lea Popovic[1], Amandine Véber[2]

December 18, 2022

Abstract

We propose a novel measure valued process which models the behaviour of chemical reaction networks in spatially heterogeneous systems. It models reaction dynamics between different molecular species and continuous movement of molecules in space. Reactions rates at a spatial location are proportional to the mass of different species present locally and to a location specific chemical rate, which may be a function of the local or global species mass as well. We obtain asymptotic limits for the process, with appropriate rescaling depending on the abundance of different molecular types. In particular, when the mass of some species in the scaling limit is discrete while the mass of the others is continuous, we obtain a new type of spatial random evolution process. This process can be shown, in some situations, to correspond to a measure-valued piecewise deterministic Markov process in which the discrete mass of the process evolves stochastically, and the continuous mass evolves in a deterministic way between consecutive jump times of the discrete part.

Contents

1 Introduction **2**

2 A spatial measure-valued reaction process **5**
2.1 The model . . . 5
2.2 Construction of the process $(M_t)_{t\geq 0}$ . . . 11
2.3 Multi-Scale Reaction Networks and Convergence to Measure-Valued PDMP . . . 14

3 Proof of the multi-scale limit (Theorem 2.11) **23**
3.1 Limiting martingale problem . . . 23
3.2 Tightness of $(M^N)_{N\geq 1}$ . . . 32
3.3 Conclusion of the proof of Theorem 2.11 . . . 34

4 Properties of the reaction-diffusion scaling limit **35**

A Proof of Theorem 2.6 **48**

[1]Department of Mathematics and Statistics Concordia University, Montreal QC H3G 1M8, Canada, e-mail: lea.popovic@concordia.ca

[2]Université Paris Cité, CNRS, MAP5, F-75006 Paris, France, e-mail: amandine.veber@parisdescartes.fr

AMS 2000 subject classification. 60J27, 60F17, 92C45, 80A30

Keywords and phrases. Measure valued process, chemical reaction network, multi-scale process, scaling limits, reaction-diffusion process, piecewise deterministic Markov process

1 Introduction

The goal of our work is to establish a mathematical framework for the dynamics of molecules of different types interacting and moving in continuous heterogeneous space. The model represents interactions that depend on both spatial location and amounts of other species, and the framework allows us to obtain results for scaling limits in scenarii of different species abundances. Our work is motivated by models for intracellular mechanisms in terms of biochemical reaction networks, but can easily be used in many other applications.

Spatial location of biochemical species within a cell plays a pivotal role in the dynamics of many key intracellular mechanisms. For example, protein movement between the nucleus and cytoplasm affects cellular responses (proteins must be present in the nucleus to regulate their target genes). Signalling proteins need to shuttle from the plasma membrane to cytoplasm and the nucleus to turn genes on or off and ultimately induce a response (spatial movement and organization is paramount to signal transduction processes).

Modelling frameworks combining spatial dynamics and biochemical reaction networks have been made at different levels of detail using deterministic and stochastic objects. The reaction-diffusion framework uses partial differential equations to model concentrations of species where reactions and movement produce deterministic changes continuously in both time and space. This framework is not appropriate for reaction dynamics which rely on changes due to molecular types in low abundances which may be localized, since it assumes the same rate of motion for all species and the same concentration scaling for all species.

The compartment framework (also referred to as the reaction-diffusion master equation, RDME) counts the number of molecules of different species in subdivided partitions of space, and in each subdivision (compartment) reactions and movement produce stochastic changes based on a Markov chain whose rates depend on the species counts in the compartment. Reactions are allowed only between species within the same compartment, and movement can occur only between different compartments. This framework is useful for speeding up simulations of stochastic dynamics (see [2, 26] for examples of variants of algorithms using RDME). The problem with this framework is that it assumes homogeneity of reaction dynamics within each compartment hence their size cannot be too large, and since reactions are allowed only within a compartment their size can also not be too small (isolating molecules). This makes the choice of partition size a challenge in many situations, in particular if the orders of abundances of different species vary.

The Brownian dynamics framework keeps track of individual molecules where each moves by an independent Brownian diffusion and participates in reactions if molecules of other source types are at close enough binding distance from it. This framework gives a more detailed account of the stochastic behaviour. Consequently, simulations in this framework are more intensive (see [3, 36] for examples of algorithms using Brownian dynamics (BD), in reaction models), and often an RDME approximation is used to simulate it (see [19] for references on multi-scale approximation simulation approaches). The BD framework can also be used to analyze the effects on the motion of molecules due to macromolecular crowding or confined geometries [9]. See the surveys [24, 49] for a discussion of appropriateness and shortcomings of different reaction modelling frameworks, as well as for examples of specific intracellular mechanisms whose size and abundance scalings are varied and for which a careful choice of framework is essential. See also the survey [18] of stochastic algorithms for reaction network dynamics with a spatial component.

This individual based model of BD motion and reaction dynamics is naturally related to the deterministic models described by reaction-diffusion partial integro-differential and reaction-

diffusion partial differential equations (in the same way that the non-spatial Markov chain model for reaction dynamics is related to the deterministic reaction dynamics described by ordinary differential equations). Some very recent works ([27, 28, 34]) have shown that for systems involving only bimolecular reactions one can obtain rigorous approximation limits connecting the BD and RD-PIDEs/RD-PDEs, under the assumptions of a single uniform scaling of all molecules and of all reaction rates. Although one can capture the behaviour of many interesting examples (see [28] for one on the trade-off between the effects of diffusion speed and interaction range), these recent results are still limited by the number of molecules involved in each reaction (and by the restriction of conservation of mass in the entropy approach of [34]). In addition to allowing for more general reactions, our representation of the Brownian dynamics also allows one to establish rigorous approximation results for varying scenarii of abundance and reaction rate scalings. The feature of flexibility to different scalings is particularly important in intracellular settings.

Current experimental evidence shows that stochasticity of reaction mechanisms in a cellular environment plays a key role in many intracellular systems (see [17, 39, 43] for gene expression experiments and a discussion of other examples). The modelling of chemical reaction networks for cellular processes requires different scalings than usual, both in terms of orders of magnitude for the abundances of different species, and in terms of orders of magnitude for reaction rates. For this reason, general multi-scale models have been developed for rigorous mathematical analysis of multi-scale stochastic reaction networks, firstly in the context of non-spatial and well-mixed (spatially homogeneous) systems (see [5, 29, 30, 12, 42, 46] for functional law of large numbers, central limit theorem and large deviation results for such models, and [10, 38] for additional scenarii with time-scale separations). Many applications use a model reduction based on some form of these results.

Spatial heterogeneity and movement of different species within it may also require different orders of magnitude for their speeds (for example small secondary messengers are faster than large complex proteins). The movement of some molecules can even be zero if a species is localized in a part of the space (for example, genes in nucleus, members of signaling pathways via anchor or scaffolding, see [8, 32]). Analytic consequences of these have been investigated within the context of compartment models (see [40, 41] for reduction limits in multi-scale compartment models), in particular reduction or increase in nonlinearity in responses can be the result of spatial heterogeneity. Many important features in applications, such as cellular adaptive ability and signal processing, rely on shaping nonlinearity in this way (see e.g. [25] for the role that movement and localization can play in generating different cellular states).

Since taking into account stochasticity may lead to qualitatively different physiological predictions, our task is to provide a representation within which one can model movement and reactions at the individual molecule level in a stochastic and heterogeneous manner. The main criterion for it is to enable macroscopic approximations of its behaviour under a diverse set of abundance and rate scalings. Using the most appropriate scalings of different molecular types, on a case by case basis, such limits will capture all of the essential reaction dynamics as well as the essential stochasticity, and provide a rigorous model reduction technique.

The applications of such models can extend beyond the molecular intracellular dynamics to processes at the organism and population levels. In recent years, tractable continuous space stochastic models of movement and interaction of different types of individuals in a heterogeneous environment have been developed in evolution and ecology (see [6, 7, 13, 22, 33] and references therein). In these models, the distribution of the population in type and space is represented by a random measure evolving by birth and death processes whose rates are linear or at most

pairwise (due to competition for resources, or fecundity selection with two types of individuals). Our framework includes such models, and extends them to more general rates as we allow any finite number of source and product species. Technically, in our framework Gronwall-type arguments are no longer sufficient to control the total interaction rates, and obtain bounds on appropriate moments of the total or local mass at any given time. Consequently, we need to assume appropriate moment conditions are satisfied (and discuss how they can be proved by other kinds of methods, such as coupling or comparison arguments).

Given our representation of the individual-based interaction and movement framework, we first show the set of assumptions on its dynamics under which this measure-valued Markov process is well defined (Theorem 2.6). We then turn to a multi-scale analysis of the process in the particular case where some species diffuse in space and others are localized (*i.e.*, can only be found at one spatial location), assuming that there are two different orders of species abundances. Letting our scaling parameter N tend to infinity, we derive the asymptotic limit in terms of the solution to a given martingale problem (Theorem 2.11). Under additional assumptions, we argue that this process is in fact a measure-valued piecewise deterministic process (PDMP), in which the unscaled mass (*i.e.*, counts) of species in lower abundances follows a Markov jump process, while the scaled mass (*i.e.*, concentration) of species in higher abundances in between these jumps evolves deterministically according to a flow that can be described as the weak solution to a system of partial integro-differential equations (Proposition 2.13). In the special case where there are no low abundance species, we recover the expected law of large number result in which the limiting process is deterministic and only given in terms of a system of nonlinear partial integro-differential equations (similar to reaction-diffusion equation results in [27, 34], see Section 4). In this case, we also focus on the qualitative effects of the localization of some molecular species, and investigate the assumptions needed to ensure regularity of limits for the subset of moving species (Proposition 4.4).

The original class of PDMPs (as defined by [16]) encompasses strong Markov processes with two types of coordinates, taking values in a subset of $\mathbb{N}^d \times \mathbb{R}^c$, in which the deterministic dynamics of the continuous coordinates and the stochastic Markov dynamics of the discrete coordinates of the processes are fully coupled. The stochastic dynamics is prescribed by jump rates and the deterministic dynamics is prescribed by a continuous flow. The appearance of infinite-dimensional PDMPs in the literature has been recent, and mostly motivated by biological applications (predominantly neuroscience models), with the continuous component taking values in a separable Hilbert space (see [4, 11, 14, 23, 45] for a possibly complete set of results to date). Since our approach derives the infinite-dimensional PDMP as a limit of a sequence of measure-valued Markov processes, some useful properties of the process are inherited from the pre-limiting sequence, and could potentially be used in devising simulation algorithms.

The rest of the paper is organised as follows. We introduce our model for chemical reaction networks in heterogeneous space in Section 2.1. We provide an algorithmic construction of the corresponding stochastic measure-valued process which we show is well-defined in Theorem 2.6 of Section 2.2, and prove in Appendix A. We then turn to a multi-scale analysis of this process in the case when some molecular species diffuse in space while others are localized, and among the set of localized species, some remain in low abundances (as our scaling parameter N tends to infinity) while the others have abundances of order $\mathcal{O}(N)$. The multi-scale limit is Theorem 2.11 of Section 2.3, which we prove in Section 3. We complement our convergence results by showing regularity properties of the limit in Proposition 4.4 of Section 4, when there are no low abundance species (and therefore the limit is deterministic and characterized by a set of partial integro-differential equations). For convenience, the notation for the different types of reactions and of

molecular species is summarised in Table 1 in Section 2.3.

2 A Spatial Measure-Valued Reaction Process – Definition and Multi-scale Analysis

2.1 The model

Let $\mathcal{T}$ be a finite set of molecular species, and let E be a compact subset of $\mathbf{R}^d$ of possible spatial locations common for all species, with a smooth boundary and a non-empty interior (the specific assumptions on E will depend on the type of motion that will model the movement of molecules). We define the underlying space for the process to be

$$\mathcal{P} = \mathcal{T} \times E,$$

each molecule being described by its characteristics $p = (x, y) \in \mathcal{P}$. The measure-valued process describing the composition of molecular species in space is given, at any time $t \geq 0$, by the counting measure (with a finite indexing set $\mathcal{I}_t$)

$$M_t = \sum_{i \in \mathcal{I}_t} \delta_{(x_i, y_i)} \tag{2.1}$$

in the space $\mathcal{M}_p$ of all finite point measures on $\mathcal{P}$. For any $M \in \mathcal{M}_p$ and any function $f : \mathcal{P} \to \mathbf{R}$, we set $\langle M, f \rangle := \sum_i f(x_i, y_i)$. Also, for a counting measure M and $n \in \mathbf{N}$, we write

$$M^{\otimes \downarrow n}(dp_1, \ldots, dp_n) := M(dp_1)(M - \delta_{p_1})(dp_2) \cdots (M - \delta_{p_1} - \cdots - \delta_{p_{n-1}})(dp_n) \tag{2.2}$$

for the measure describing the sampling of n points without replacement according to M. By convention, $M^{\otimes \downarrow n}$ is the null measure when $n = 0$ or $n > \langle M, 1 \rangle$.

Let the reaction network consist of a finite set of reactions R, where (for convenience of notation in our model) a reaction $r \in R$ is of the form

$$A_1^r + \cdots + A_{k_r}^r \mapsto B_1^r + \cdots + B_{k'_r}^r$$

with source reactants of types A_i^r and product reactants of type B_i^r, both of which are allowed to repeat. The usual notation for reaction networks by stoichiometric vectors and a stoichiometric matrix can be determined as

$$\nu_{r,x} = \sum_{j=1}^{k_r} \mathbf{1}_x(A_j^r), \quad \nu'_{r,x} = \sum_{j=1}^{k'_r} \mathbf{1}_x(B_j^r), \quad S = [\nu'_{r,x} - \nu_{r,x}]. \tag{2.3}$$

In our model, we shall consider two kinds of reactions. In a *non-localized* reaction r, close-by molecules react and are consumed or created continuously in space (with a rate dependent on the region in space where the reaction takes place, this rate being potentially equal to zero in some parts of E – see below). In contrast, a *localized* reaction takes place at a given point in space, usually where some of the species involved are attached (*e.g.*, the nucleus membrane, or the extra-cellular membrane of a cell). We shall denote the set of non-localized reactions by R_{NL} and the set of localized reactions by R_L, so that $R = R_{NL} \cup R_L$ (a disjoint union). For simplicity we shall not formally consider reactions happening continuously in some part of E and in a localized way in other parts of E, since they may be encoded as the sum of several purely

continuous or purely localized reactions in our framework (note however that this generalisation would simply consist in taking a measure $\varrho_r(d\bar{y})$ of a mixed form in (2.7)).

Non-localized reactions. The dynamics of a non-localized reaction $r \in R_{NL}$ at any location $\bar{y} \in E$ is specified by a spatially-dependent chemical reaction factor $\bar{h}_r(\bar{y}, M)$ and a mass-action-kinetics reaction function that is based on availability of source reactants in a neighbourhood of $\bar{y}$ as determined by a proximity (probability) kernel Γ_ϵ centred at $\bar{y}$ and with support in the ball $B(0,\epsilon) \subset \mathbf{R}^d$ for some $\epsilon > 0$ (for simplicity, we take the same kernel Γ_ϵ for all reactions, but this can be easily generalised to reaction-dependent kernels). Specifically, given a set of k_r source reactants $(p_i = (x_i, y_i))_{i=1,\dots,k_r}$ of the appropriate types, the rate of the reaction r which produces molecules of type $B^r_1, \dots, B^r_{k'_r}$ at location $\bar{y} \in E$ is

$$\lambda_r(\bar{y}, M; p_1, \dots, p_{k_r}) := \bar{h}_r(\bar{y}, M) \left(\prod_{i=1}^{k_r} \mathbf{1}_{A^r_i}(x_i) \Gamma_\epsilon(y_i - \bar{y}) \right). \tag{2.4}$$

The overall rate of reaction r at location $\bar{y} \in E$ is then obtained by sampling the k_r source reactants, without replacement, from the current state M_t of the measure-valued process:

$$\begin{aligned} &\int_{\mathcal{P}^{k_r}} M_t^{\otimes\downarrow k_r}(dp_1, \dots, dp_{k_r})\, \lambda_r(\bar{y}, M_t; p_1, \dots, p_{k_r}) \\ &\quad = \bar{h}_r(\bar{y}, M_t) \int_{\mathcal{P}^{k_r}} M_t^{\otimes\downarrow k_r}(dp_1, \dots, dp_{k_r}) \left(\prod_{i=1}^{k_r} \left(\mathbf{1}_{A^r_i}(x_i) \Gamma_\epsilon(y_i - \bar{y}) \right) \right). \end{aligned} \tag{2.5}$$

The dependence of the chemical reaction factor $\bar{h}_r(\bar{y}, M_t)$ on the current state of the global species composition is included in order to allow the presence of chemical reaction rates that are not simply of "mass-action" form (see [12] for examples of where elimination of extremely fast intermediate subnetworks appear in the reaction factors of the reduced network reactions). Such reactions allow the mass of some species, which is unchanged by that particular reaction, to affect the reaction rate indirectly through a role of a promoter or inhibitor. This dependence on M is of the form

$$\bar{h}_r(\bar{y}, M) = h_r(\bar{y}, \langle M, \Psi_{r,\bar{y}} \rangle), \tag{2.6}$$

for some nonnegative functions h_r, $\Psi_{r,\bar{y}}$. For instance, we may want to take $\Psi_{r,\bar{y}}(x, y) = \mathbf{1}_{B(\bar{y},\epsilon)}(y)$ when the reaction rate is affected only by nearby mass. The assumptions we make on h_r and $\Psi_{r,\bar{y}}$ are detailed in Assumption (A1) below (in particular, for technical reasons we shall need to replace the indicator function $\mathbf{1}_{B(\bar{y},\epsilon)}(y)$ by a continuous approximation to it in the above example).

Localized reactions. Suppose that reaction $r \in R_L$ occurs at a single location $\bar{y}_r \in E$. We use the same kernel Γ_ϵ to check the availability of source reactants in the neighbourhood of $\bar{y}_r$. Hence, equation (2.4) with $\bar{y} = \bar{y}_r$ still describes the rate at which a given set of k_r molecules $(p_i = (x_i, y_i))_{i=1,\dots,k_r}$ of the appropriate types react at $\bar{y}_r$ and equation (2.5) now describes the total rate at which reaction r occurs (in contrast with non-localized reactions, whose local rates have to be integrated over $\bar{y} \in E$ to obtain their global reaction rates).

Remark 2.1. Taking Γ_ϵ and h_r to be constant over the range of possible values for their arguments, we recover the mass-action form for reaction rates. Molecules still have spatial locations but the latter play no role in the dynamics of the reactions. We can also recover compartment dynamics by making all reactions localized at centres of the compartments and by taking Γ_ϵ and h_r to be constant (with $\epsilon =$ half the compartment size).

To unify the notation, let us write ℓ_E for Lebesgue measure on E and let us define the measures

$$\varrho_r(d\bar{y}) = \ell_E(d\bar{y}) \quad \text{if } r \in R_{NL} \qquad \text{and} \qquad \varrho_r(d\bar{y}) = \delta_{\bar{y}_r}(d\bar{y}) \quad \text{if } r \in R_L. \tag{2.7}$$

Hence, in both the localized and non-localized case, the total rate at which reaction r occurs when the current state of the system is described by the counting measure M is given by

$$\Lambda_r(M) := \int_{E\times\mathcal{P}^{k_r}} \varrho_r(d\bar{y}) M^{\otimes\downarrow k_r}(dp_1,\dots,dp_{k_r})\, \lambda_r(\bar{y}, M; p_1,\dots,p_{k_r}). \tag{2.8}$$

We shall rigorously construct the measure-valued process $(M_t)_{t\geq 0}$ in an algorithmic way in Section 2.2. For now we just keep introducing the key ingredients to describe its dynamics. The weak topology on the space $\mathcal{M}$ of all finite measures on the compact space $\mathcal{P}$ is determined by $\langle M, f\rangle := \int f dM$ over a sufficiently large class of functions f in $\mathcal{C}(\mathcal{P})$, the space of all continuous (hence bounded) functions on $\mathcal{P}$. Note that for a point measure $M \in \mathcal{M}_p$, we recover $\langle M, f\rangle = \sum_i f(x_i, y_i)$.

Most of our analysis of the process $(M_t)_{t\geq 0}$ will rely on the martingale problem that we now introduce. First, the operator describing the change due to reactions in R, acting on test functions of the form

$$F_f = F(\langle \cdot, f\rangle), \tag{2.9}$$

with $F \in \mathcal{C}_b(\mathbf{R})$ (*i.e.*, continuous and bounded on $\mathbf{R}$) and $f \in \mathcal{C}(\mathcal{P})$, will be given by $\sum_{r\in R} G_r$, where for all $r \in R$ we have

$$\begin{aligned} G_r F_f(M) = \int_{E\times\mathcal{P}^{k_r}} & \varrho_r(d\bar{y}) M^{\otimes\downarrow k_r}(dp_1,\dots,dp_{k_r}) \lambda_r(\bar{y}, M; p_1,\dots,p_{k_r}) \\ & \left[F\left(\langle M, f\rangle - \sum_{i=1}^{k_r} f(x_i, y_i) + \sum_{i=1}^{k_r'} f(B_i^r, \bar{y}) \right) - F_f(M) \right]. \end{aligned} \tag{2.10}$$

Recalling the definition of λ_r given in (2.4), we see that a reaction r at location $\bar{y}$ occurs if the needed types A_i^r of reactants sampled from the measure M exist in sufficient numbers in the neighbourhood of $\bar{y}$ described by the kernel Γ_ϵ. When it occurs, it removes the source reactants from M and produces molecules of types B_i^r, all at location $\bar{y}$. Note that in case $k_r = 0$ (creation of product molecules from an external source), the rate is determined by $\bar{h}_r$ (which may still be a function of the measure M).

Example 2.2. Suppose the network consists of one ($r = 1$) localized and three ($r = 2, 3, 4$) non-localized reactions on two molecular species (S, S'):

$$\emptyset \overset{\bar{h}_1(S')}{\underset{\mathbf{1}_{\bar{y}_1=0}}{\mapsto}} S, \quad S \overset{\bar{h}_2}{\mapsto} S + S', \quad S' \overset{\bar{h}_3}{\mapsto} \emptyset, \quad S \overset{\bar{h}_4}{\mapsto} \emptyset \tag{2.11}$$

This is a simplified version of the transcription-translation mechanism of a protein: here S is the mRNA and S' is the protein, the creation of mRNA occurs only in the nucleus at $\bar{y}_1 = 0 \in E$ and is given by the transcription rate $\bar{h}_1(\bar{y}_1, M) = h_1(\bar{y}_1, \langle M, \Psi_{S',\epsilon}\rangle)$, where $\Psi_{S',\epsilon}$ is a continuous approximation to $\mathbf{1}_{\{S'\}\times B(0,\epsilon)}$. In the unregulated case the function $\bar{h}_1 = h_1(0)$ is constant, while in the self-regulated case $\bar{h}_1(\bar{y}_1, M) = h_1(0, a)$ is a function of the mass $a = \langle M, \Psi_{S',\epsilon}\rangle$ of produced protein that diffuses back to the neighbourhood of the nucleus. To ensure $\bar{h}_1$ satisfies

our Assumption (A1) (see below) we ask that h_1 should be Lipschitz in a and uniformly bounded on compact sets for a. For example, we can take $h_1(\bar{y}_1, a) = c_1/(1 + (c_2 a)^k)$ if the mechanism is repressed by S', or $h_1(\bar{y}_1, a) = (1 + c_1 a^k)/(c_2^k + a^k)$ if the mechanism is activated by S' (see [37]), for some $k \geq 1$ (also referred to as Hill function coefficient). For reactions $r = 2, 3, 4$, the reaction rate factor $\bar{h}_r$ is taken to depend on the spatial coordinate but not on the mass coordinate and we ask that each should be uniformly bounded over $\bar{y} \in E$. The operators encoding these reactions are

$$
\begin{aligned}
G_1 F_f(M) &= h_1(\bar{y}_1, \langle M, \Psi_{S',\epsilon}\rangle)\Big[F\big(\langle M, f\rangle + f(S, \bar{y}_1)\big) - F_f(M)\Big] \qquad (2.12)\\
G_2 F_f(M) &= \int_{E\times\mathcal{P}} \ell_E(d\bar{y}) M(dp)\, \mathbf{1}_S(x)\Gamma_\epsilon(y - \bar{y}) h_2(\bar{y})\Big[F\big(\langle M, f\rangle + f(S', \bar{y})\big) - F_f(M)\Big]\\
G_3 F_f(M) &= \int_{E\times\mathcal{P}} \ell_E(d\bar{y}) M(dp)\, \mathbf{1}_{S'}(x)\Gamma_\epsilon(y - \bar{y}) h_3(\bar{y})\Big[F\big(\langle M, f\rangle - f(S', y)\big) - F_f(M)\Big]\\
G_4 F_f(M) &= \int_{E\times\mathcal{P}} \ell_E(d\bar{y}) M(dp)\, \mathbf{1}_S(x)\Gamma_\epsilon(y - \bar{y}) h_4(\bar{y})\Big[F\big(\langle M, f\rangle - f(S, y)\big) - F_f(M)\Big].
\end{aligned}
$$

Note that in the second reaction, for simplicity we chose to consider that the species S involved was not at all modified by the reaction. Another option, in line with the description of the model given above, would have been to consider that the source reactant (S, y) appearing in the expression for $G_2 F_f(M)$ should be withdrawn and replaced by some new molecule $(S, \bar{y})$ at location $\bar{y}$. This choice of formulation is left to the modeller. This simple example will be used later to illustrate the effects of reaction localization, and multi-scaling of abundance of molecular types, see Examples 2.14 and 2.15 in Section 2.3.

The change due to the movement of molecules in the spatial domain E will be given by an $\mathcal{M}_p$-valued Markov process $\widetilde{M}$ that we can take to be fairly general, provided that it can be described by means of an operator $\mathcal{D}$ well-defined on a large enough class of functions of the form (2.9), over which it is bounded by some power of the total mass function (see Assumption (A0') and Remark 2.5 below). In particular, we may be interested in situations where the local concentration in molecules of some species influences the propensity of other species to visit or avoid the corresponding region of space. A much simpler example, on which we shall concentrate in the multi-scale analysis expounded in Section 2.3, is to suppose that molecules of type x move in E independently of each other and of molecules of the other types, following a diffusion with locally bounded Lipschitz drift coefficient $b_x : E \to \mathbf{R}^d$ and locally bounded Lipschitz dispersion matrix $\Sigma_x : E \to \mathbf{R}^{d\times d}$ such that the diffusion matrix $\Sigma_x^2 := \Sigma_x(\Sigma_x)^{\mathbf{t}}$ is uniformly elliptic, this diffusion being normally reflected at the boundary of E. When the interior $\mathring{E}$ of E is non-empty, bounded and either convex, or smooth ($\mathcal{C}^3$), both with a piecewise-smooth boundary and only a finite number of convex corners, this reflected movement is well posed ([35], [50] give solutions to stochastic differential equations with reflection; [31] equates them to solutions of submartingle problems). In this particular case, for every $x \in \mathcal{T}$ and every sufficiently regular function f, we

define the functions $b_x \cdot \nabla_y f$, $\Sigma_x^2 \circ \Delta_y f$, and $\Sigma_x^2 \circ \big((\nabla_y f)(\nabla_y f)^{\mathbf{t}}\big)$ by

$$b_x \cdot \nabla_y f : (x', y') \mapsto \mathbf{1}_x(x') \sum_{i=1}^{d} b_x(y')_i \frac{\partial f}{y_i}(x', y') \tag{2.13}$$

$$\Sigma_x^2 \circ \Delta_y f : (x', y') \mapsto \frac{\mathbf{1}_x(x')}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \Sigma_x^2(y')_{ij} \frac{\partial^2 f}{\partial y_i \partial y_j}(x', y')$$

$$\Sigma_x^2 \circ \big((\nabla_y f)(\nabla_y f)^{\mathbf{t}}\big) : (x', y') \mapsto \frac{\mathbf{1}_x(x')}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \Sigma_x^2(y')_{ij} \frac{\partial f}{\partial y_i}(x', y') \frac{\partial f}{\partial y_j}(x', y'),$$

where z_i denotes the i-th coordinate of the vector z and Z_{ij} denotes the (i, j)-coordinate of the matrix Z (and the notation $\circ$ is inspired by the Hadamard product of matrices). Still in this particular example, we consider test functions of the form $F_f = F(\langle \cdot, f \rangle)$ with

(*a*) $F \in \mathcal{C}_b^2(\mathbf{R})$, *i.e.*, bounded and of classe $\mathcal{C}^2$ on $\mathbf{R}$, and

(*b*) $f \in \mathcal{C}^{0,2}(\mathcal{T} \times E)$ (*i.e.*, measurable in the first coordinate, and of class $\mathcal{C}^2$ in the second coordinate) satisfying $\nabla_y f(x', y') \cdot n(y') = 0$ for all $(x', y') \in \mathcal{T} \times \partial E$, where $n(y')$ denotes the outward normal to the boundary of E at $y' \in \partial E$ and $\cdot$ denotes scalar product in $\mathbf{R}^d$,

and the operator $\mathcal{D}$ applied to such a test function can be written (see Theorem 3.1 in [48])

$$\begin{aligned} \mathcal{D}F_f(M) = {} & F'(\langle M, f \rangle) \sum_{x \in \mathcal{T}} \langle M, b_x \cdot \nabla_y f + \Sigma_x^2 \circ \Delta_y f \rangle \\ & + \; F''(\langle M, f \rangle) \sum_{x \in \mathcal{T}} \big\langle M, \Sigma_x^2 \circ \big((\nabla_y f)(\nabla_y f)^{\mathbf{t}}\big) \big\rangle. \end{aligned} \tag{2.14}$$

Coming back to the general case for the movement of species and summing up the above, the operator that will serve as a basis for the martingale problem describing the overall dynamics of the measure-valued Markov process $(M_t)_{t \geq 0}$ is the following:

$$LF_f(M) = \sum_{r \in R} G_r F_f(M) + \mathcal{D}F_f(M). \tag{2.15}$$

For ease of reference, let us give a name to the martingale problem associated to L.

Definition 2.3. Let $\mathbf{F}$ be the set of test functions defined in Assumption (A0') below. We say that an $\mathcal{M}_p$-valued process $(M_t)_{t \geq 0}$ satisfies the martingale problem MP(L) if for every function $F_f \in \mathbf{F}$, the process

$$\left(F_f(M_t) - F_f(M_0) - \int_0^t ds \, LF_f(M_s) \right)_{t \geq 0}$$

is a martingale (for the natural filtration associated to $(M_t)_{t \geq 0}$).

Let us introduce the following **Assumptions** on the Markov process $\widetilde{M}$ describing the motion of molecules between two jumps due to reactions in R and on the operators G_r. By convention, for every localized reaction $r \in R_L$ we set $h_r(y, \cdot) \equiv 0$ for all $y \neq \bar{y}_r$. We write $D_{\mathcal{M}_p}[0, \infty)$ for the space of all càdlàg paths with values in $\mathcal{M}_p$.

(A0) The $\mathcal{M}_p$-valued Markov process $(\widetilde{M}_t)_{t\geq 0}$ is well-defined for any distribution of the initial value $\widetilde{M}_0$ belonging to the set of probability measures on $\mathcal{M}_p$ and it has càdlàg path a.s. Furthermore, it satisfies that the total number of atoms and their first coordinates $x_i \in \mathcal{T}$ are left unchanged by the dynamics (in other words, only the spatial locations in E of the atoms evolve in time).

This assumption is the minimal requirement on $\widetilde{M}$ for the process constructed in Section 2.2 to be well-defined (possibly until explosion). In order to derive the semi-martingale decompositions of the processes of interest, we shall need to impose some additional conditions on $\widetilde{M}$:

(A0') There exist an operator $\mathcal{D}$ and a set $\mathbf{F} = \{F_f : F \in \Phi_1,\, f \in \Phi_2\}$ of functions of the form (2.9), where Φ_1 is dense in $\mathcal{C}(\mathbf{R})$ for the topology of uniform convergence over compact sets, and Φ_2 is dense in $\mathcal{C}(\mathcal{P})$ for the topology of uniform convergence, satisfying

- (i) $\mathcal{D}$ is well-defined over the set $\mathbf{F}$ and for every $F_f \in \mathbf{F}$, there exists a constant $c_{F,f} > 0$ such that
$$|\mathcal{D}F_f(M)| \leq c_{F,f}\langle M, 1\rangle \qquad \text{for every } M \in \mathcal{M}_p.$$
- (ii) For every $F_f \in \mathbf{F}$,
$$\left(F_f\big(\widetilde{M}_t\big) - F_f\big(\widetilde{M}_0\big) - \int_0^t ds\, \mathcal{D}F_f\big(\widetilde{M}_s\big)\right)_{t\geq 0}$$
is a martingale with respect to the natural filtration associated to $\widetilde{M}$.

As concerns the operators describing the changes due to reactions, we shall need the following assumptions.

(A1) For each $r \in R$, the reaction factor $h_r : E \times \mathbf{R}_+ \to \mathbf{R}_+$ is uniformly bounded over compact subsets of $E \times \mathbf{R}_+$: for every $\ell \geq 0$,
$$\sup_{\bar{y}\in E}\sup_{a\in[0,\ell]} h_r(\bar{y}, a) = \|h_r\|_{\infty,\ell} < \infty.$$
It is also Lipschitz in the second coordinate, with Lipschitz constant L_r independent of the first coordinate $\bar{y}$. Finally, for every $\bar{y} \in E$ the function $\Psi_{r,\bar{y}} : \mathcal{P} \to \mathbf{R}_+$ is continuous and
$$\sup_{\bar{y}\in E}\sup_{p\in\mathcal{P}} \Psi_{r,\bar{y}}(p) = \|\Psi_r\|_\infty < \infty.$$

(A2) For some fixed $\epsilon > 0$, the function $\Gamma_\epsilon \geq 0$ is a continuous probability density with support contained in the closed ball $B(0,\epsilon) \subset \mathbf{R}^d$: in particular,
$$\sup_{y\in\mathbf{R}^d} \Gamma_\epsilon(y) = \|\Gamma_\epsilon\|_\infty < \infty, \qquad \int_{\mathbf{R}^d} \Gamma_\epsilon(y)dy = 1.$$

Remark 2.4. Assumption (A0') is satisfied in our previous example of independent inhomogeneous diffusions if we restrict our attention to $F \in \mathcal{C}_b^2(\mathbf{R})$ with bounded first and second derivatives, since the set of all $f \in \mathcal{C}^{0,2}(\mathcal{T} \times E)$ with vanishing normal derivative on ∂E is dense in $\mathcal{C}(\mathcal{P})$ for the supremum norm – see Remark 1.1 in [13].

Remark 2.5. (a) We may generalise the bound stated in Assumption (A0')-(i) into the existence of $K \in \mathbf{N}$ such that

$$|\mathcal{D}F_f(M)| \leq c_{F,f}\langle M, 1\rangle^K \qquad \text{for every } M \in \mathcal{M}_p, \tag{2.16}$$

for instance if we wanted to include some density-dependence in the movement of species. For our existence result, Theorem 2.6, to hold true, we would then have to replace Assumption (A3) stated in the theorem by the stronger condition that we can control the supremum over any finite time interval of the $(K \vee (1 + \max_r k_r))$-th moment of the total mass of the process. See Remark A.3 at the end of Appendix A.
(b) We may also relax the assumption that $\widetilde{M}$ should have càdlàg paths (at the expense of $(M_t)_{t\geq 0}$ itself not having càdlàg paths), but since it is a natural assumption in view of the applications and since our multi-scale analysis of the particular case of diffusing molecules in Section 2.3 will rely on it, we keep the simpler framework of continuously moving particles with mass evolution as càdlàg processes.

2.2 Construction of the process $(M_t)_{t\geq 0}$

In this section, we suppose that Assumptions (A0), (A1) and (A2) hold true and we construct a process that satisfies the desired dynamics. It is this particular process (appropriately rescaled) that we shall use later in our multi-scale analysis. The construction relies on the fact that the rate at which each reaction occurs is bounded from above by a polynomial in the total mass of the system. In general, we cannot rule out the possibility of an explosion in finite time of the total mass, which yields the first part of Theorem 2.6. However, assuming that the supremum of $\langle M_t, 1\rangle$ over any fixed time horizon is a.s. finite obviously prevents this pathology from appearing, and additional controls of appropriate moments of the total mass allow us to describe the process in terms of a martingale problem (see the second part of Theorem 2.6, using the stronger Assumption (A0') and Assumption (A3) introduced below).

Let M_0 be a random finite counting measure, with law $\mathcal{L}(M_0)$. All the random objects used in this section are supposed to be defined on a common probability space $(\Omega, \mathcal{F}, \mathbf{P})$.

- Write $M_0 = \sum_{i\in I_0} \delta_{(x_i,y_i)}$. Set $\tau^0 = 0$.
- Let $\widetilde{M}^0_t = \sum_{i\in I_0} \delta_{(x_i,Y^0_i(t))}$ denote the value at time t of the $\mathcal{M}_p$-valued Markov process $\widetilde{M}$ starting from M_0 (since by assumption this process only makes particles move in E, only the second coordinate $Y^0_i(t)$ evolves for every i).
- For every reaction $r \in R$, define $\mathcal{E}^1_r$ as an exponential random variable with parameter 1, independent of all other variables, and the random time τ^1_r as (recall the definition of $\Lambda_r(M)$ given in (2.8)):

$$\tau^1_r := \inf\left\{ s > 0 : \int_0^s dt\, \Lambda_r\big(\widetilde{M}^0_t\big) \geq \mathcal{E}^1_r \right\}.$$

  Using the bound on λ_r that we shall establish in Lemma A.1 together with the fact that the total mass of $\widetilde{M}^0$ is constant equal to $|I_0|$, we obtain that for each r, $\Lambda_r(\widetilde{M}^0_t)$ is bounded independently of t and so all τ^1_r are positive a.s. We can therefore set $\tau^1 := \min_{r\in R} \tau^1_r$ and let r^1 be the index of the unique reaction satisfying $\tau^1_{r^1} := \min_{r\in R} \tau^1_r$. In words, τ^1 is the random time at which the first reaction occurs when we let the $|I_0|$ particles move in space

and interact, and r^1 is the index of the reaction that takes place at time τ^1. The outcome of this reaction is given by the following procedure: Sample $(\bar{y}^1, p_1^1, \dots, p_{k_{r^1}}^1) \in E \times \mathcal{P}^{k_{r^1}}$ according to the probability measure

$$\frac{\lambda_{r^1}(\bar{y}, \widetilde{M}^0_{(\tau^1)-}; p_1, \dots, p_{k_{r^1}})}{\Lambda_{r^1}(\widetilde{M}^0_{(\tau^1)-})} \varrho_r(d\bar{y}) \big(\widetilde{M}^0_{(\tau^1)-}\big)^{\otimes\downarrow k_{r^1}}(dp_1, \dots, dp_{k_{r^1}}).$$

(Observe that the denominator is necessarily nonzero, otherwise the probability that the reaction occurring at time τ^1 is the one labelled by r^1 would be 0.) Define the new value of the measure describing the system just after the reaction by

$$\widetilde{M}^1_{\tau^1} := \sum_{i \in I_0 \setminus \bar{I}^1_s} \delta_{(x_i, Y^0_i(\tau^1))} + \sum_{i=1}^{k'_{r^1}} \delta_{(B^{r^1}_i, \bar{y}^1)},$$

where $\bar{I}^1_s$ is the index set of the particles chosen in the previous step. That is, we remove the k_{r^1} source reactants, and add the k'_{r^1} product reactants all at location $\bar{y}^1$.

- Write I_1 for the index set of $\widetilde{M}^1_{\tau^1}$, and (abusing notation[1]) let the collection of particle locations $((Y^1_i(t),\, i \in I_1))_{t \geq \tau^1}$ evolve according to the random motion in E^{I_1} generated by $\widetilde{M}$ and started at time τ^1 from the current locations $(y_i)_{i \in I_1}$ of the particles. For every time $t \geq \tau^1$, define $\widetilde{M}^1_t = \sum_{i \in I_1} \delta_{(x_i, Y^1_i(t))}$.

For every $j \geq 2$, proceed recursively following the same steps as above:

- For every $r \in R$, let $\mathcal{E}^j_r$ be an independent exponential r.v. with parameter 1. Define

$$\tau^j_r := \inf\left\{ s > \tau^{j-1} : \int_{\tau^{j-1}}^s dt\, \Lambda_r\big(\widetilde{M}^{j-1}_t\big) \geq \mathcal{E}^j_r \right\}. \tag{2.17}$$

  Let then $\tau^j := \min_{r \in R} \tau^j_r$ and let us denote the index of the unique reaction that realizes the minimum at time τ^j_r by r^j. Sample $(\bar{y}^j, p^j_1, \dots, p^j_{k_{r^j}}) \in E \times \mathcal{P}^{k_{r^j}}$ according to the probability measure

$$\frac{\lambda_{r^j}(\bar{y}, \widetilde{M}^{j-1}_{(\tau^j)-}; p_1, \dots, p_{k_{r^j}})}{\Lambda_{r^j}(\widetilde{M}^{j-1}_{(\tau^j)-})} \varrho_r(d\bar{y}) \big(\widetilde{M}^{j-1}_{(\tau^j)-}\big)^{\otimes\downarrow k_{r^j}}(dp_1, \dots, dp_{k_{r^j}}).$$

  The vector $(\bar{y}^j, p^j_1, \dots, p^j_{k_{r^j}})$ indicates the location of the j-th reaction and the k_{r^j} particles chosen to react (which will then be removed). Next, set

$$\widetilde{M}^j_{\tau^j} := \sum_{i \in I_{j-1} \setminus \bar{I}^j_s} \delta_{(x_i, Y^{j-1}_i(\tau^j))} + \sum_{i=1}^{k'_{r^j}} \delta_{(B^{r^j}_i, \bar{y}^j)},$$

  where $\bar{I}^j_s$ is the index set of the particles chosen to react during step j.

[1] More rigorously, at time τ^1 we paste an independent copy of $\widetilde{M}$ starting from $\widetilde{M}^1_{\tau^1}$ to the trajectory of $\widetilde{M}^0$ stopped at τ^1.

- Write I_j for the index set of $\widetilde{M}^j_{\tau^j}$ and, with the same abuse of notation as earlier, let $((Y^j_i(t),\, i \in I_j))_{t\geq\tau^j}$ evolve according to the random motion in E^{I_j} generated by $\widetilde{M}$, started at time τ^j from the collection $(y_i)_{i\in I_j}$ of particle locations at time τ^j. For every $t \geq \tau^j$, define $\widetilde{M}^j_t = \sum_{i\in I_j} \delta_{(x_i, Y^j_i(t))}$.

Finally, let $\tau^\infty = \sup_{j\in\mathbf{N}} \tau^j$ and define the process $(M_t)_{0\leq t<\tau^\infty}$ by

$$\forall t \in [0, \tau^\infty), \quad M_t = \widetilde{M}^{j(t)}_t, \quad \text{with } j(t) \text{ such that } t \in \left[\tau^{j(t)}, \tau^{j(t)+1}\right). \tag{2.18}$$

The main result of this section is the following theorem.

Theorem 2.6. *Suppose that Assumptions (A0), (A1) and (A2) are satisfied. Then:*

(i) For any $\mathcal{M}_p$-valued random variable M_0, the process $(M_t)_{0\leq t<\tau^\infty}$ constructed above is well-defined, it has the Markov property and has càdlàg sample paths with probability 1.

(ii) Suppose in addition that Assumption (A0') holds, together with

(A3) For every $T > 0$,

$$\sup_{t\in[0,T]} \langle M_t, 1\rangle < \infty \quad \textit{a.s., and} \quad \sup_{t\in[0,T]} \mathbf{E}\Big[\langle M_t, 1\rangle^{1+\max_r k_r}\Big] < \infty,$$

where the max in the exponent is taken over all $r \in R$.

Then $\tau^\infty = +\infty$ a.s. and the process $(M_t)_{t\geq 0}$ is solution to MP(L).

Remark 2.7. Assumption (A3) has to be checked case by case, as it may hold true for many different reasons. In Example 2.2 above, assuming that $\bar{h}_1$ and $\bar{h}_2$ are uniformly bounded in both coordinates (*e.g.*, the first reaction saturates when the concentration in species S' is high) and that Assumption (A2) is satisfied too, then the global creation rate of molecules of type S is bounded by a constant and the rate of creation of molecules of type S' is at most linear in the current number of molecules S. Hence, the total number of particles in the system is stochastically bounded by a binary branching process (with branching rate $\|\bar{h}_2\|_\infty$) with immigration at constant rate $\|\bar{h}_1\|_\infty$, for which it is straightforward to check that the two conditions stated in Assumption (A3) are satisfied.

More generally, these two properties may be proved by stochastically bounding the total mass process by an appropriate birth and death process. They will also be satisfied whenever the reactions involving more than one source reactants do not make the number of particles increase (that is, $k'_r \leq k_r$) and the reaction factors $\bar{h}_r$ are all uniformly bounded in both coordinates.

Remark 2.8. Depending on the nature of the movement of particles, the process $(M_t)_{t\geq 0}$ of Theorem 2.6 (*i*) can be characterised in different ways: for instance through a martingale problem, or through a system of stochastic differential equations combined with a Poisson point measure for the jumps (in line with the construction in [13]). In Theorem 2.6 (*ii*) we focus on the martingale problem that will allow us to obtain the semi-martingale decomposition of the processes $(\langle M_t, f\rangle)_{t\geq 0}$ that will be needed in the multiscale analysis of Section 2.3, on the principle that such a martingale problem is a rather universal approach.

The proof of Theorem 2.6 is classical and is therefore deferred until Appendix A.

2.3 Multi-Scale Reaction Networks and Convergence to Measure-Valued PDMP

In this section, we suppose that Assumptions (A1), (A2) and (A3) are satisfied. We also suppose that the set $\mathcal{T}$ of molecular species can be partitioned into a set $\mathcal{T}_L$ of *localized* species, for which we assume that all molecules of type $x \in \mathcal{T}_L$ sit at some fixed point $\bar{y}_x \in E$, and a set $\mathcal{T}_{NL}$ of *diffusive* species. Molecules of type $x \in \mathcal{T}_{NL}$ move in E independently of each other and of molecules of the other types, following a diffusion with locally bounded Lipschitz drift coefficient $b_x : E \to \mathbf{R}^d$ and locally bounded Lipschitz dispersion matrix $\Sigma_x : E \to \mathbf{R}^{d\times d}$ such that the diffusion matrix $\Sigma_x^2 := \Sigma_x(\Sigma_x)^{\mathbf{t}}$ is uniformly elliptic, this diffusion being normally reflected at the boundary of E. Again, we assume that the interior of the compact set E is nonempty and it has a piecewise-smooth $\mathcal{C}^3$ boundary with only a finite number of convex corners. Recall the notation introduced in (2.13), the set of test functions introduced just below (2.13) (that we take to be the set $\mathbf{F}$ defined in Assumption (A0')) and the operator $\mathcal{D}$ introduced in (2.14) (where we set $b_x \equiv 0$ and $\Sigma_x^2 \equiv \mathbf{0}$ for $x \in \mathcal{T}_L$). According to Remark 2.4, Assumption (A0') is satisfied and the process $(M_t)_{t\geq 0}$ constructed in Section 2.2 is well-defined by Theorem 2.6.

For consistency, in our model we can allow localized species to be source reactants of any reaction, but we have to make the following **Assumption**:

(B0) A localized species x can only be produced by a reaction r that is localized and taking place at location $\bar{y}_r = \bar{y}_x$.

Indeed, otherwise molecules of type x may pile up at locations different from $\bar{y}_x$, which would contradict our definition of localized species. Note however that this constraint is mainly a consequence of the way we define the mathematical model, and the biology behind localized species and reactions is obviously more complicated.

Let us assume that abundances of different species may scale differently. More precisely, let us suppose that all diffusive species are present in large numbers, of the same order of magnitude $N \in \mathbf{N}$. Localized species, on the other hand, can either be abundant (with the same order of magnitude as diffusive species), or may appear in small numbers bound to remain of order $\mathcal{O}(1)$. This may happen for instance if some of the localized species are made of very big molecules compared to diffusing species. Hence, the set $\mathcal{T}_L$ of localized species is further partitioned into (a disjoint union)

$$\mathcal{T}_L = \mathcal{T}_{L,s} \cup \mathcal{T}_{L,b}, \tag{2.19}$$

where $\mathcal{T}_{L,s}$ (*resp.*, $\mathcal{T}_{L,b}$) denotes the set of localized species with $\mathcal{O}(1)$ (*resp.*, $\mathcal{O}(N)$) abundance.

Now, the philosophy is as follows: for the species with $\mathcal{O}(N)$ abundance, we scale the corresponding part of each M_t by N and, as we let N tend to infinity, a law of large number-type of result will show the convergence of this part to a deterministic flow Φ. As concerns the species with $\mathcal{O}(1)$ abundance (*i.e.*, species in $\mathcal{T}_{L,s}$), since they are localized at given points in space, we only have to count how many of them sit at these locations at any time and to use descending factorials of these counts to describe each reaction rate involving at least one such molecule. Furthermore, under appropriate conditions, the rate at which molecules from low abundance species are consumed or created will remain of order $\mathcal{O}(1)$ when we let N tend to infinity, and therefore their dynamics will remain stochastic in the limit. This is the content of Theorem 2.11 below. In some cases (see Proposition 2.13), the limiting process can be shown to correspond to a measure-valued PDMP in which the spatial distribution in abundant species changes continuously according to the flow Φ (that depends on the current state of the species in $\mathcal{T}_{L,s}$) and the spatial distribution in low abundance species changes only by jumps at random discrete times whose intensity depends on Φ. In the particular case where $\mathcal{T}_{L,s} = \emptyset$, we recover

a rather classical large-population deterministic limit in which, when a density for the spatial distribution of species exists, this density satisfies a system of coupled partial integro-differential equations (for the diffusive species) and integro-differential equations (for the local amounts of localized species where they sit). This particular case will be developped in Section 4.

Remark 2.9. The assumption that *only* localized species may occur in small numbers is crucial to the results expounded in this section. Indeed, if some low abundance species were allowed to diffuse in space, then the reaction rates in the "deterministic flow" part of the limiting dynamics (see Theorem 2.11) would constantly change in a stochastic way and the limiting process would no longer be a PDMP. In this case, the construction and properties of the limit are more involved, and are left for future work.

Let us now formalise the above intuition. Let $N \in \mathbf{N}$. For every $t \geq 0$, recall the notation $\mathcal{I}_t$ for the index set of the counting measure M_t and define $M_t^N \in \mathcal{M}$ as follows:

$$M_t^N := \frac{1}{N} \sum_{\substack{i \in \mathcal{I}_t: \\ x_i \in \mathcal{T}_{L,s}^c}} \delta_{(x_i,y_i)} + \sum_{\substack{i \in \mathcal{I}_t: \\ x_i \in \mathcal{T}_{L,s}}} \delta_{(x_i,y_i)}. \tag{2.20}$$

Since $M_t^N = (1/N)M_t$ on $(\mathcal{T}_{NL} \cup \mathcal{T}_{L,b}) \times E$ and $M_t^N = M_t$ on $\mathcal{T}_{L,s} \times E$, it is easy to see that the process $(M_t^N)_{t\geq 0}$ is still Markovian and takes its values in $D_{\mathcal{M}}[0,\infty)$. As in (2.2), there is a natural notion of sampling without replacement of k particles from a measure M of the form (2.20), given by :

$$\begin{aligned} M^{\otimes\downarrow k}(dp_1,\dots,dp_k) := M(dp_1)\Big(M - \frac{1}{N^{\mathfrak{s}(x_1)}}\delta_{p_1}\Big)(dp_2)\cdots \\ \Big(M - \frac{1}{N^{\mathfrak{s}(x_1)}}\delta_{p_1} - \cdots - \frac{1}{N^{\mathfrak{s}(x_{k-1})}}\delta_{p_{k-1}}\Big)(dp_k), \end{aligned} \tag{2.21}$$

where $\mathfrak{s}(x)$ is equal to 0 if $x \in \mathcal{T}_{L,s}$ and 1 otherwise.

Let us suppose that the space-dependent chemical reaction factor h_r depends on N, and so do the functions $\Psi_{r,\bar{y}}$ which we assume (by slight abuse of notation) can be written as

$$\Psi_{r,\bar{y}}^N = \frac{1}{N}\Psi_{r,\bar{y}}\,\mathbf{1}_{\mathcal{T}_{L,s}^c \times E} + \Psi_{r,\bar{y}}\,\mathbf{1}_{\mathcal{T}_{L,s}\times E}, \tag{2.22}$$

where the $\Psi_{r,\bar{y}}$ are independent of N and satisfy the properties stated in Assumption (A1) (with $\|\Psi_r\|_\infty$ thus independent of N). This assumption is natural since the contribution of the abundant species should globally be of the same order $\mathcal{O}(1)$ as the contribution of the low abundance species.

Before we state the main results of this section, we define a few more pieces of notation. For every reaction $r \in R$, let us write

$$k_{r,b} := \sum_{i=1}^{k_r} \mathbf{1}_{\mathcal{T}_{L,s}^c}(A_i^r) \qquad \text{and} \qquad k'_{r,b} := \sum_{i=1}^{k'_r} \mathbf{1}_{\mathcal{T}_{L,s}^c}(B_i^r) \tag{2.23}$$

for the respective numbers of source and product molecules in reaction r that are of an abundant type (and the numbers of source and product molecules from a species in low abundance are therefore $k_r - k_{r,b}$ and $k'_r - k'_{r,b}$, respectively). Without loss of generality, we shall assume that

the source molecules of abundant types are labelled by $\{1,\dots,k_{r,b}\}$ and the ones in $\mathcal{T}_{L,s}$ by $\{k_{r,b}+1,\dots,k_r\}$. Likewise, the product molecules from an abundant species are labelled by $\{1,\dots,k'_{r,b}\}$ and those from a low-abundance species by $\{k'_{r,b}+1,\dots,k'_r\}$. Recall from (2.3) that $\nu_{r,x}$ and $\nu'_{r,x}$ stand for the stoichiometric coefficients of species x in reaction r, and define the sets

$$R^{n\ell} := \Big\{ r \in R \ : \ \sum_{x\in\mathcal{T}_{L,s}} \big|\nu_{r,x} - \nu'_{r,x}\big| = 0 \Big\}, \qquad R^{\ell} := R \setminus R^{n\ell}. \tag{2.24}$$

In words, a reaction r belongs to $R^{n\ell}$ if and only if it does not modify the number of molecules of any species in $\mathcal{T}_{L,s}$.

For every reaction $r \in R$, let the function $\widetilde{h}_r^N$ be defined by

$$\widetilde{h}_r^N := \begin{cases} N^{k_{r,b}-1} h_r^N & \text{if } r \in R^{n\ell}, \\ N^{k_{r,b}} h_r^N & \text{if } r \in R^{\ell}. \end{cases} \tag{2.25}$$

That is, the scaling property required for h_r^N depends on whether reaction r modifies the number of species of some localized $\mathcal{O}(1)$-reactant types (the second line) or not (the first line). In the case $r \in R^{\ell}$, the choice of the exponent $k_{r,b}$ instead of $k_{r,b}-1$ is imposed by the fact that the net change in the abundance of the species in $\mathcal{T}_{L,s}$ involved leads to a macroscopic jump for M^N, instead of a change of the order of $\mathcal{O}(1/N)$. In other words, $r \in R^{n\ell}$ need to be sped up by a factor of N to see $\mathcal{O}(1)$ change while $r \in R^{\ell}$ can stay at their regular speed. See (3.12) in the proof of Theorem 2.11. The **Assumptions** on the regularity and boundedness of the different functions and processes that will be needed for the convergence of $(M^N)_{N\geq 1}$ are the following.

(B1) For every $r \in R$, there exists $\widetilde{h}_r : E \times \mathbf{R}_+ \to \mathbf{R}_+$ which is Lipschitz in the second coordinate, with Lipschitz constant L_r independent of the first coordinate, and such that

$$\mathcal{S}_r := \int_E \varrho_r(d\bar{y}) \sup_{a\geq 0} \widetilde{h}_r(\bar{y},a) < \infty, \tag{2.26}$$

and

$$\lim_{N\to\infty} \int_E \varrho_r(d\bar{y}) \sup_{a\geq 0} \Big| \widetilde{h}_r^N(\bar{y},a) - \widetilde{h}_r(\bar{y},a) \Big| = 0. \tag{2.27}$$

(B2) Let $k^* := 1 + (\max_{r\in R} k_r)$. For every $T > 0$, we have

$$\sup_{N\in\mathbf{N}} \mathbf{E}\Big[\sup_{t\in[0,T]} \langle M_t^N, 1\rangle^{k^*\vee 2} \Big] < \infty.$$

Assumption (B1) is the analogue of the first part of Assumption (A1), while Assumption (B2) has a role similar to that of Assumption (A3). Both B-conditions are stronger than the corresponding A-conditions, since the latter are required for the process to be well-defined for every given N, while the former will guarantee the uniform integrability of the different terms appearing in the sequence of martingale problems. Note that Assumption (B2) is not a simple consequence of our assumptions on the reaction coefficients, since nothing guarantees that the limiting dynamics obtained by replacing $\widetilde{h}_r^N(\bar{y},a)$ with $\widetilde{h}_r(\bar{y},a)$ is non-explosive. Like Assumption (A3), it has to be proven by appropriate comparison or coupling arguments (for instance).

	Molecular species
$\mathcal{T}$	finite set of molecular species (or types) – $\mathcal{T} = \mathcal{T}_{NL} \cup \mathcal{T}_L$
$\mathcal{T}_{NL}$	subset of non-localized species
$\mathcal{T}_L$	subset of localized species – $\mathcal{T}_L = \mathcal{T}_{L,b} \cup \mathcal{T}_{L,s}$
$\bar{y}_x$	(unique) spatial location of localized species $x \in \mathcal{T}_L$
$\mathcal{T}_{L,b}$	subset of localized species of $\mathcal{O}(N)$ abundance
$\mathcal{T}_{L,s}$	subset of localized species of $\mathcal{O}(1)$ abundance
	Reaction types
R	finite set of reactions – $R = R_{NL} \cup R_L$ and $R = R^{n\ell} \cup R^{\ell}$
R_{NL}	subset of non-localized reactions
R_L	subset of localized reactions
$\bar{y}_r$	(unique) spatial location of localized reaction $r \in R_L$
$R^{n\ell}$	subset of reactions modifying the counts of none of the species in $\mathcal{T}_{L,s}$ – see (2.24)
R^{ℓ}	subset of reactions modifying the counts of at least one species in $\mathcal{T}_{L,s}$
$k_{r,b}$	number of source reactants in reaction r that are of an abundant type – see (2.23)
$k'_{r,b}$	number of product reactants in reaction r that are of an abundant type

Table 1: Main notation for the different types of molecular species and reactions.

Remark 2.10. Notice that unless there are no reactions involving at least one source species, that is, unless all reactions happen from a source external to the system, we have $k^* \vee 2 = k^* = 1 + \max_r k_r$. The fact that the exponent should be at least 2 is used in the control of the tail distribution of the total mass of M_t^N, see (3.27), and in the extension of the martingale problem to functions F_f with $F(a) = a^2$, see Remark 3.1.

Because we want to consider only spatial distributions of species in which molecules from a localized species $x \in \mathcal{T}_L$ can only be found at location $\bar{y}_x$, we shall restrict our attention to the closed subset $\mathcal{M}_{\mathcal{T}_L}$ defined by

$$\mathcal{M}_{\mathcal{T}_L} := \Big\{ M \in \mathcal{M} : \sum_{x \in \mathcal{T}_L} \langle M, \mathbf{1}_{(x,\cdot)} - \mathbf{1}_{(x,\bar{y}_x)} \rangle = 0 \Big\}. \tag{2.28}$$

Finally, recall the descending factorial notation $(n)_j = n(n-1)\cdots(n-j+1)$ (with the convention that $(n)_0 = 1$) and write $\mathcal{C}^{2,\perp}(\mathcal{P})$ for the space of all functions on $\mathcal{P} = \mathcal{T} \times E$ that are measurable in the first coordinate, of class $\mathcal{C}^2$ in the second and such that $f(x', y') \cdot n(y') = 0$ for all $(x', y') \in \mathcal{T} \times \partial E$ (where we recall that $n(y')$ is the outward normal to the boundary of E at $y' \in \partial E$ – this vector is well-defined for all but a finite number of points in ∂E by assumption).

Theorem 2.11. (i) *Suppose that Assumptions (B0), (B1) and (B2) are satisfied, that for every $N \geq 1$ we have $M_0^N \in \mathcal{M}_{\mathcal{T}_L}$ a.s., and that the sequence of random variables $(M_0^N)_{N\geq 1}$ converges in distribution to a random variable $M_0 \in \mathcal{M}_{\mathcal{T}_L}$ as N tends to infinity. Suppose also that there exists at most one solution $(M_t^\infty)_{t\geq 0}$ to the $D_{\mathcal{M}_{\mathcal{T}_L}}[0,\infty)$-martingale problem: $M_0^\infty \overset{(d)}{=} M_0$ and for every $F \in \mathcal{C}_b^1(\mathbf{R})$ with bounded first derivative and every $f \in \mathcal{C}^{2,\perp}(\mathcal{P})$,*

$$
\begin{aligned}
&\Bigg(F(\langle M_t^\infty, f\rangle) - F(\langle M_0^\infty, f\rangle) \qquad (2.29)\\
&\quad - \sum_{r\in R^{n\ell}} \int_0^t ds \int_{E\times\mathcal{P}^{k_{r,b}}} \varrho_r(d\bar{y}) M_s^\infty(dp_1)\cdots M_s^\infty(dp_{k_{r,b}})\, \widetilde{h}_r\big(\bar{y}, \langle M_s^\infty, \Psi_{r,\bar{y}}\rangle\big)\\
&\qquad\qquad \times \Big(\prod_{i=1}^{k_{r,b}} \mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\Big)\Big(\prod_{x\in\mathcal{T}_{L,s}} \Big\{(\langle M_s^\infty, \mathbf{1}_{(x,\cdot)}\rangle)_{\nu_{r,x}} \Gamma_\epsilon(\bar{y}_x-\bar{y})^{\nu_{r,x}}\Big\}\Big)\\
&\qquad\qquad \times \Big(\sum_{i=1}^{k'_r} f(B_i^r,\bar{y}) - \sum_{i=1}^{k_r} f(x_i,y_i)\Big) F'(\langle M_s^\infty, f\rangle)\\
&\quad - \sum_{r\in R^{\ell}} \int_0^t ds \int_{E\times\mathcal{P}^{k_{r,b}}} \varrho_r(d\bar{y}) M_s^\infty(dp_1)\cdots M_s^\infty(dp_{k_{r,b}})\, \widetilde{h}_r\big(\bar{y}, \langle M_s^\infty, \Psi_{r,\bar{y}}\rangle\big)\\
&\qquad\qquad \times \Big(\prod_{i=1}^{k_{r,b}} \mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\Big)\Big(\prod_{x\in\mathcal{T}_{L,s}} \Big\{(\langle M_s^\infty, \mathbf{1}_{(x,\cdot)}\rangle)_{\nu_{r,x}} \Gamma_\epsilon(\bar{y}_x-\bar{y})^{\nu_{r,x}}\Big\}\Big)\\
&\qquad\qquad \times \Big(F\Big(\langle M_s^\infty, f\rangle - \sum_{i=k_{r,b}+1}^{k_r} f(x_i,\bar{y}_{x_i}) + \sum_{i=k'_{r,b}+1}^{k'_r} f(B_i^r,\bar{y})\Big) - F\big(\langle M_s^\infty, f\rangle\big)\Big)\\
&\quad - \int_0^t ds \Big\{\sum_{x\in\mathcal{T}_{NL}} \langle M_s^\infty, b_x\cdot\nabla_y f + \Sigma_x^2 \circ \Delta_y f\rangle\Big\} F'(\langle M_s^\infty, f\rangle)\Bigg)_{t\geq 0}
\end{aligned}
$$

is a martingale.

Then, this solution exists and as N tends to infinity, M^N converges in distribution to M^∞ in $D_{\mathcal{M}_{\mathcal{T}_L}}[0,\infty)$. In addition, M^∞ also satisfies the martingale problem (2.29) *with $F = \mathrm{Id}$ and $F : a \mapsto a^2$.*

(ii) If, furthermore, for every $m_0 \in \mathcal{M}_{\mathcal{T}_L}$ the martingale problem (2.29) *with $M_0^\infty = m_0$ admits at most one solution, then the limiting process is Markovian.*

Remark 2.12. Using Lemma 3.2 and the calculations expounded in Section 3.1 with $F = \mathrm{Id}$ and $F : a \mapsto a^2$, we obtain that the limiting process may also be described through the semi-martingale decomposition of $(\langle M_t^\infty, f\rangle)_{t\geq 0}$ for every $f \in \mathcal{C}^{2,\perp}(\mathcal{P})$. We leave the straightforward computation of the finite variation part and of the predictable quadratic variation of the martingale part of this decomposition to the interested reader.

The proof of Theorem 2.11 is given in Section 3. Observe that in the term of (2.29) corresponding to $r \in R^\ell$, only the changes in the low abundance species are visible. Indeed, the simultaneous changes in abundant species are of order $\mathcal{O}(1/N)$ before taking the limit, and therefore they vanish as $N \to \infty$.

The limiting measure-valued process in Theorem 2.11 is characterised by a well-posed martingale problem, which does not provide an explicit construction of its dynamics. Based on the form of this martingale problem, a natural description of the limit should be in terms of a measure-valued piecewise deterministic Markov process with the "flow part" being the continuous change in abundant species due to reactions in $R^{n\ell}$ and to the diffusion of molecules, while

the composition in localized low abundance species would change by discrete jumps occurring at a rate given by the terms in (2.29) corresponding to reactions in R^ℓ. To our knowledge, there is no general result providing such a correspondence between solutions to martingale problems of the form (2.29) and compositions of flow and jump dynamics for infinite-dimensional stochastic processes, and so the desired relation has to be proven by hand. Although we shall not be able to do it in the full generality of our approach, let us push this direction a bit more. To parallel the general framework for non-spatial reaction networks developed in [15] (in which the PDMP take their values in $\mathbf{R}^c \times \mathbf{N}^d$), let us decompose every M_t^∞ into its "continuous" part $M_t^{\infty,c}$ seen as an element of the set $\mathcal{M}(\mathcal{T}_{L,s}^c \times E)$ of all finite measures on $\mathcal{T}_{L,s}^c \times E$, and its "discrete" part $M_t^{\infty,d}$ seen as an element of the set $\mathcal{M}_p(\mathcal{T}_{L,s} \times E)$ of all finite point measures on $\mathcal{T}_{L,s} \times E$, in such a way that $M_t^\infty = M_t^{\infty,c} \otimes M_t^{\infty,d}$. That is, $M_t^{\infty,c}$ (*resp.*, $M_t^{\infty,d}$) is the image measure of M_t^∞ under the projection map to $\mathcal{T}_{L,s}^c \times E$ (*resp.*, $\mathcal{T}_{L,s} \times E$). Let us now introduce an $\mathcal{M}(\mathcal{T}_{L,s}^c \times E) \times \mathcal{M}_p(\mathcal{T}_{L,s} \times E)$-valued process $(\mathcal{X}_t)_{t\geq 0}$, explicitly constructed in terms of a deterministic flow driving the evolution of the first coordinate and a sequence of discrete jump times for the second coordinate, which we would like to prove to be identical in distribution to $(M_t^{\infty,c}, M_t^{\infty,d})_{t\geq 0}$.

To this end, let us first suppose that for every $(m^c, m^d) \in \mathcal{M}(\mathcal{T}_{L,s}^c \times E) \times \mathcal{M}_p(\mathcal{T}_{L,s} \times E)$, there exists a unique solution $(\Phi_{m^d}(t, m^c))_{t\geq 0}$ to the following deterministic system: for every $f \in \mathcal{C}^{2,\perp}(\mathcal{P})$, we have for every $t \geq 0$

$$
\begin{aligned}
&\langle \Phi_{m^d}(t, m^c), f\rangle - \langle m^c, f\rangle \\
&= \sum_{r\in R^{n\ell}} \int_0^t ds \int_{E\times\mathcal{P}^{k_{r,b}}} \varrho_r(d\bar{y}) \Phi_{m^d}(s, m^c)^{\otimes k_{r,b}}(dp_1, \ldots, dp_{k_{r,b}})\, \widetilde{h}_r\big(\bar{y}, \langle \Phi_{m^d}(s, m^c) \otimes m^d, \Psi_{r,\bar{y}}\rangle\big) \\
&\qquad \times \Big(\prod_{i=1}^{k_{r,b}} \mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i - \bar{y})\Big)\Big(\prod_{x\in\mathcal{T}_{L,s}} \Big\{\big(m^d(\{(x, \bar{y}_x)\})\big)_{\nu_{r,x}} \Gamma_\epsilon(\bar{y}_x - \bar{y})^{\nu_{r,x}}\Big\}\Big) \\
&\qquad \times \Big(\sum_{i=1}^{k'_r} f(B_i^r, \bar{y}) - \sum_{i=1}^{k_r} f(x_i, y_i)\Big) \\
&\quad + \sum_{x\in\mathcal{T}_{NL}} \int_0^t ds\, \langle \Phi_{m^d}(s, m^c), b_x \cdot \nabla_y f + \Sigma_x^2 \circ \Delta_y f\rangle.
\end{aligned}
$$

In the above and later, integrals over $E \times (\mathcal{T}_{L,s}^c \times E)^{k_{r,b}}$ are replaced by integrals over $E \times \mathcal{P}^{k_{r,b}}$ to alleviate the notation. Notice that in particular, $\Phi_{m^d}(0, m^c) = m^c$. Using the regularity of the functions $\widetilde{h}_r$, $\Psi_{r,\bar{y}}$, Γ_ϵ, b_x and Σ_x^2, it is then easy to show that each trajectory $(\Phi_{m^d}(t, m^c))_{t\geq 0}$ is continuous for the weak topology on $\mathcal{M}(\mathcal{T}_{L,s}^c \times E)$ and that the mapping $(m^c, m^d) \mapsto (\Phi_{m^d}(t, m^c))_{t\geq 0}$ is measurable.

Let us now construct $(\mathcal{X}_t)_{t\geq 0} = (\mathcal{X}_t^c, \mathcal{X}_t^d)_{t\geq 0}$ starting at a pair $(m^c, m^d) \in \mathcal{M}(\mathcal{T}_{L,s}^c \times E) \times \mathcal{M}_p(\mathcal{T}_{L,s} \times E)$ (which we allow to be random). For every $r \in R^\ell$, define $\mathcal{E}_r^1$ as an exponential random variable with parameter 1 (independent of all other variables), and the random time τ_r^1

as

$$\tau_r^1 := \inf\Big\{t>0 : \int_0^t ds \int_{E\times\mathcal{P}^{k_{r,b}}} \varrho_r(d\bar{y})\Phi_{m^d}(s,m^c)^{\otimes k_{r,b}}(dp_1,\dots,dp_{k_{r,b}})\Big(\prod_{i=1}^{k_{r,b}} \mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\Big)$$
$$\widetilde{h}_r\big(\bar{y},\langle\Phi_{m^d}(s,m^c)\otimes m^d,\Psi_{r,\bar{y}}\rangle\big)\Big(\prod_{x\in\mathcal{T}_{L,s}}\Big\{\big(m^d(\{(x,\bar{y}_x)\})\big)_{\nu_{r,x}}\Gamma_\epsilon(\bar{y}_x-\bar{y})^{\nu_{r,x}}\Big\}\Big)\geq \mathcal{E}_r^1\Big\}.$$

Set $\tau^1 := \min_{r\in R^\ell}\tau_r^1$ and let r^1 be the index of the unique reaction satisfying $\tau_{r^1}^1 := \min_r \tau_r^1$. For every $t\in[0,\tau^1)$, set

$$\mathcal{X}_t := \big(\Phi_{m^d}(t,m^c),m^d\big).$$

In words, until the first reaction $r^1\in R^\ell$ occurs at time τ^1, only the continuous part of $\mathcal{X}$ evolves, according to the flow $\Phi_{m^d}(\cdot,m^c)$. The outcome of reaction r^1 is the jump of $\mathcal{X}$ to

$$\big(\mathcal{X}_{\tau^1}^c,\mathcal{X}_{\tau^1}^d\big) := \Big(\Phi_{m^d}(\tau^1,m^c)\,,\, m^d - \sum_{i=k_{r^1,b}+1}^{k_{r^1}} \delta_{(A_i^{r^1},\bar{y}_{A_i^{r^1}})} + \sum_{i=k'_{r^1,b}+1}^{k'_{r^1}} \delta_{(B_i^{r^1},\bar{y}_{r^1})}\Big).$$

Then for every $j\geq 2$, we proceed recursively following the same ideas:

- For every $r\in R^\ell$, define $\mathcal{E}_r^j$ as an exponential random variable with parameter 1 (independent of all other variables), and the random time τ_r^j as

$$\tau_r^j := \inf\Big\{t>\tau^{j-1} : \int_{\tau^{j-1}}^t ds \int_{E\times\mathcal{P}^{k_{r,b}}} \varrho_r(d\bar{y})\Phi_{\mathcal{X}_{\tau^{j-1}}^d}(s-\tau^{j-1},\mathcal{X}_{\tau^{j-1}}^c)^{\otimes k_{r,b}}(dp_1,\dots,dp_{k_{r,b}})$$
$$\Big(\prod_{i=1}^{k_{r,b}}\mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\Big)\widetilde{h}_r\big(\bar{y},\langle\Phi_{\mathcal{X}_{\tau^{j-1}}^d}(s-\tau^{j-1},\mathcal{X}_{\tau^{j-1}}^c)\otimes\mathcal{X}_{\tau^{j-1}}^d,\Psi_{r,\bar{y}}\rangle\big)$$
$$\Big(\prod_{x\in\mathcal{T}_{L,s}}\Big\{\big(\mathcal{X}_{\tau^{j-1}}^d(\{(x,\bar{y}_x)\})\big)_{\nu_{r,x}}\Gamma_\epsilon(\bar{y}_x-\bar{y})^{\nu_{r,x}}\Big\}\Big)\geq\mathcal{E}_r^j\Big\}. \tag{2.30}$$

  Set $\tau^j := \min_{r\in R^\ell}\tau_r^j$ and let r^j be the index of the unique reaction satisfying $\tau_{r^j}^j := \min_r\tau_r^j$. For every $t\in[\tau^{j-1},\tau^j)$, set

$$\mathcal{X}_t := \big(\Phi_{\mathcal{X}_{\tau^{j-1}}^d}(t-\tau^{j-1},\mathcal{X}_{\tau^{j-1}}^c),\mathcal{X}_{\tau^{j-1}}^d\big). \tag{2.31}$$

  Again, what is encoded here is the fact that after the jump of $\mathcal{X}$ at time τ^{j-1} and until the next jump time τ^j, the discrete part of $\mathcal{X}$ remains constant equal to its value $\mathcal{X}_{\tau^{j-1}}^d$ at time τ^{j-1} while its continuous part evolves in a deterministic way according to the flow dictated by the value of the discrete part and starting at $\mathcal{X}_{\tau^{j-1}}^c$.

- At time τ^j, a jump due to reaction r^j occurs to the discrete part of $\mathcal{X}$:

$$\big(\mathcal{X}_{\tau^j}^c,\mathcal{X}_{\tau^j}^d\big) := \Big(\Phi_{\mathcal{X}_{\tau^{j-1}}^d}(\tau^j-\tau^{j-1},\mathcal{X}_{\tau^{j-1}}^c)\,,\,\mathcal{X}_{\tau^{j-1}}^d - \sum_{i=k_{r^j,b}+1}^{k_{r^j}}\delta_{(A_i^{r^j},\bar{y}_{A_i^{r^j}})} + \sum_{i=k'_{r^j,b}+1}^{k'_{r^j}}\delta_{(B_i^{r^j},\bar{y}_{r^j})}\Big). \tag{2.32}$$

It is difficult to give general conditions under which

$$\lim_{j\to\infty} \tau^j = +\infty \qquad \text{a.s.}$$

and $(\mathcal{X}_t)_{t\geq 0}$ is a well-defined Markov process. Indeed, using Assumption (A2) and the first part of Assumption (B1), similar bounds as the ones used in the proof of Theorem 2.6 in Appendix A show that the instantaneous jump rate of the second coordinate of $\mathcal{X}$ is bounded by a constant times the total mass of $\mathcal{X}_t^c \otimes \mathcal{X}_t^d$ to the power (at most) $1 + \max_r k_r$. As long as we have not proved that $(\mathcal{X}_t^c \otimes \mathcal{X}_t^d)_{t\geq 0}$ and M^∞ are equal in law, we cannot use Assumption (B2) to control the moments of the total mass of $\mathcal{X}_t^c \otimes \mathcal{X}_t^d$ and therefore we have no a priori control over the jump times τ^j, which may accumulate in finite time (of course, the existence of $(M_t^\infty)_{t\geq 0}$ is a good indication that such an accumulation should not happen). Hence, our final result is the following.

Proposition 2.13. *Suppose that the assumptions of Theorem 2.11(i) are satisfied, and write $(M_0^c, M_0^d) \in \mathcal{M}(\mathcal{T}_{L,s}^c \times E) \times \mathcal{M}_p(\mathcal{T}_{L,s} \times E)$ for the decomposition of the initial value M_0 of the limiting process $(M_t^\infty)_{t\geq 0}$ into its "continuous" and its "discrete" parts. Suppose also that $(\mathcal{X}_t)_{t\geq 0}$ introduced above is a well-defined Markov process with $\mathcal{X}_0 = (M_0^c, M_0^d)$ and such that for every $T > 0$,*

$$\sup_{t\in[0,T]} \mathbf{E}\Big[\big(\langle \mathcal{X}_t^c \otimes \mathcal{X}_t^d, 1\rangle\big)^{1+\max_{r\in R} k_r}\Big] < \infty. \tag{2.33}$$

Then

$$(\mathcal{X}_t)_{t\geq 0} \overset{(d)}{=} \big(M_t^{\infty,c}, M_t^{\infty,d}\big)_{t\geq 0}. \tag{2.34}$$

The result stated in Proposition 2.13 relies on the facts that the solution to the martingale problem (2.29) is supposed to be unique and that the decomposition of M_t^∞ into its continuous and discrete parts is unique (so that it suffices to show that $(\mathcal{X}_t^c \otimes \mathcal{X}_t^d)_{t\geq 0}$ satisfies the martingale problem characterising the law of $(M_t^\infty)_{t\geq 0}$ to conclude). Since its proof is identical to that of Theorem 2.6, we omit it.

Example 2.14. (Diffusive species only).
Consider the simple reaction network given by the set of four reactions (2.11). Suppose both molecular types S, S' diffuse in space with constant coefficients $b_S \equiv 0, \Sigma_S^2 = \sigma_S^2 \mathrm{Id}$ and $b_{S'} \equiv 0, \Sigma_{S'}^2 = \sigma_{S'}^2 \mathrm{Id}$ respectively (where σ_S^2 and $\sigma_{S'}^2$ are positive constants), and that both have abundances of order $\mathcal{O}(N)$. Suppose the reaction factors for all non-localized (linear) reactions $h_r^N, r = 2, 3, 4$ are constant in space and have size of order $\mathcal{O}(1)$. That is, $h_r^N(\bar{y}) \equiv h_r > 0$. Recall that in Example 2.2, in the unregulated case the reaction factor $h_1(0, a) = h_1(0)$ of the localized reaction 1 was simply a constant, while in the self-regulated case we used $h_1(0, a)$ with $a = \langle M, \Psi_{S',\epsilon}\rangle$ (where $\Psi_{S',\epsilon}$ is a continuous approximation to $\mathbf{1}_{\{S'\}\times B(0,\epsilon)}$), which depended on the local mass a of S' in the form $h_1(0, a) = c_1/(1 + (c_2 a)^k)$ for self-repression, and $h_1(0, a) = c_1 a^k/(c_2^k + a^k)$ for self-activation. To match the rate of the other reactions (which happen at a rate proportional to the current counts $\mathcal{O}(N)$ of S or S' molecules), suppose that the mass function $\Psi_{1,0}^N$ is equal to $(1/N)\Psi_{S',\epsilon}$ and that the reaction factor h_1^N has size of order $\mathcal{O}(N)$: $h_1^N(0, a) = N h_1(0, a)$ with h_1 given above.

Under these specifications, all reaction factors $\tilde{h}_r = N^{k_r-1} h_r^N(\bar{y}, a)$ are uniformly bounded in $\bar{y}$ and for $r = 1$, it is uniformly bounded in a as well. Hence the Assumptions of Theorem 2.11, as well as those of Lemma 4.2 and Proposition 4.4 stated in Section 4 and dealing with the existence of a density for the limiting process, can be shown to hold using a multi-type birth and death

process as an upper envelope (see Remark 2.7). The density μ_t of the limiting process solves the following set of deterministic equations: for all $x \in \{S, S'\}$ and $y \in \partial E$, $\nabla_y \mu_t(x,y) \cdot n(y) = 0$, and for every $t > 0$ and $y \in E$,

$$\begin{aligned}
\partial_t \mu_t(S,y) &= \sigma_S^2 \Delta_y \mu_t(S,y) + h_1\Big(0, \int_E \ell_E(d\bar{y}) \Psi_{S',\epsilon}(S',\bar{y}) \mu_t(S',\bar{y})\Big) \mathbf{1}_{\{y=0\}} \\
&\quad - h_4\Big(\int_E \ell_E(d\bar{y}) \Gamma_\epsilon(y-\bar{y})\Big) \mu_t(S,y), \\
\partial_t \mu_t(S',y) &= \sigma_{S'}^2 \Delta_y \mu_t(S',y) + h_2 \int_E \ell_E(d\bar{y}) \Gamma_\epsilon(\bar{y}-y) \mu_t(S,\bar{y}) \\
&\quad - h_3\Big(\int_E \ell_E(d\bar{y}) \Gamma_\epsilon(y-\bar{y})\Big) \mu_t(S',y),
\end{aligned} \tag{2.35}$$

with for instance $\mu_0(S,y) = \mu^S \mathbf{1}_{\{y=0\}}$ and $\mu_0(S',y) \equiv 0$ (recall that ℓ_E stands for Lebesgue measure on E).

Note that in the unregulated case $h_1(0,a) = h_1(0)$, the partial differential equation for $\mu_t(S,y)$ is autonomous (with the integral multiplying h_4 being a function of y in the ϵ-boundary of E and constant equal to 1 elsewhere), while in the self-regulated case it is fully coupled with the equation for $\mu_t(S',y)$. When it exists, the steady state solution $y \mapsto (\pi(S,y), \pi(S',y))$ is solution to the system (2.35) with the left-hand side set to 0.

Example 2.15. (Diffusive and localized low abundance species).
Using again the reaction network given by (2.11), suppose now that only molecules of species S' diffuse in space (in a symmetric and homogeneous way, *i.e.*, $b_{S'} \equiv 0$ and $\Sigma_{S'}^2 = \sigma_{S'}^2 \mathrm{Id}$ with $\sigma_{S'}^2 > 0$) while species S is localized at $\bar{y} = 0$. For consistency, assume that reaction 4 is now localized at $\bar{y} = 0$ too. Suppose also that the abundance of species S' is of order $\mathcal{O}(N)$, while the abundance of species S is of order $\mathcal{O}(1)$. This is a simplification of the mechanism where only few molecules of mRNA are transcribed and they stay close to the nucleus, while a larger number of proteins are generated and they diffuse throughout the cell. By definition (see Table 1), we thus have $k_{1,b} = 0 = k_{2,b} = k_{4,b}$ and $k_{3,b} = 1$ and reactions 2 and 3 belong to $R^{n\ell}$ (*i.e.*, do not modify the counts of the low abundance species S) while reactions 1 and 4 belong to R^ℓ. Coming back to (2.25), we obtain

$$k_{1,b} - \mathbf{1}_{\{1 \in R^{n\ell}\}} = 0 \ ; \ k_{2,b} - \mathbf{1}_{\{2 \in R^{n\ell}\}} = -1 \ ; \ k_{3,b} - \mathbf{1}_{\{3 \in R^{n\ell}\}} = 0 \ ; \ k_{4,b} - \mathbf{1}_{\{4 \in R^{n\ell}\}} = 0.$$

Hence, let us suppose that the reaction factor h_2^N is constant through space and of order $\mathcal{O}(N)$ (that is, $h_2^N(\bar{y}) \equiv N h_2$ with $h_2 > 0$), that the reaction factor h_3^N is also constant but of order $\mathcal{O}(1)$ (that is, $h_3^N(\bar{y}) \equiv h_3 > 0$), that the reaction factor h_4^N for the localized reaction 4 is equal to some constant $h_4 > 0$ independent of N, that the reaction factor h_1^N for the localized reaction 1 is of order $\mathcal{O}(1)$ too (that is, $h_1^N(0,a) = h_1(0,a)$ with h_1 defined as in Example 2.14) and finally that the mass function $\Psi_{1,0}^N$ at the location of reaction 1 is again equal to $(1/N)\Psi_{S',\epsilon}$.

Under these scaling specifications, the assumptions of Theorem 2.11 and of Proposition 2.13 can be shown to hold, and the limiting process is a measure-valued piecewise deterministic Markov process described as follows.

(i) The mass of the discrete coordinate $M_t^{\infty,d}$ (number of S molecules at $\bar{y} = 0$) is a Markov jump process started at $M_0(\{(S,0)\}) = m_0 \in \mathbf{N}$ with birth rate $h_1(0,\cdot)$ and linear death rate $h_4 \Gamma_\epsilon(0) M_t^\infty(\{(S,0)\})$; as before, in the unregulated case $h_1(0,a) = h_1(0)$, and in the self-regulated case $h_1(0,a) = c_1/(1+(c_2 a)^k)$ or $h_1(0,a) = c_1 a^k/(c_2^k + a^k)$, where $a = \langle M_t^\infty, \Psi_{S',\epsilon}\rangle$.

(ii) The continuous coordinate $M_t^{\infty,c}$ (concentration of S' molecules in E) evolves in a deterministic way between the jump times $(\tau^j, j \geq 1)$ of $M_t^\infty(\{(S,0)\})$. In each random time interval $[\tau^j, \tau^{j+1})$, when $M_t^\infty(\{(S,0)\})$ is constant, the density $\mu_t(S',y) = dM_t^\infty(\{(S',y)\})/dy$ satisfies the equation: $\forall y \in E, \forall t \in [\tau^j, \tau^{j+1})$,

$$\partial_t \mu_t(S',y) = \sigma_{S'}^2 \Delta_y \mu_t(S',y) + h_2 \Gamma_\epsilon(-y) M_t^\infty(\{(S,0)\}) - h_3 \bigg(\int_E \ell_E(d\bar{y}) \Gamma_\epsilon(y-\bar{y}) \bigg) \mu_t(S',y)$$

with (for instance) $\mu_0(S',y) \equiv 0$ and with subsequent initial values given by $\mu_{\tau^j}(S',y) = \lim_{t \uparrow \tau^j} \mu_t(S',y)$.

In the unregulated case, the autonomous dynamics for $M(\{(S,0)\})$ has

$$\pi(S,0) \sim \text{Poisson}\bigg(\frac{h_1(0)}{h_4 \Gamma_\epsilon(0)} \bigg)$$

as its stationary distribution; and conditional on the value of $\pi(S,0)$, the steady state for $\mu(S',y)$ is determined by solving

$$0 = \sigma_{S'}^2 \Delta_y \pi(S',y) + h_2 \Gamma_\epsilon(-y) \pi(S,0) - h_3 \bigg(\int_E \ell_E(d\bar{y}) \Gamma_\epsilon(y-\bar{y}) \bigg) \pi(S',y), \ \forall y \in E.$$

3 Proof of the multi-scale limit (Theorem 2.11)

We proceed as usual, by first showing that any limit of a converging subsequence of $(M^N)_{N\geq 1}$ necessarily satisfies the martingale problem stated in Theorem 2.11 (but with trajectories that are *a priori* $\mathcal{M}$-valued). This is done in Section 3.1, where we also argue that this limiting martingale problem holds with $F = \text{Id}$ and $F : a \mapsto a^2$ under our assumptions (see Remark 3.1(a)), and that indeed all limiting trajectories a.s. take their values in the subset $\mathcal{M}_{\mathcal{T}_L}$ (see Remark 3.1(b)). In Section 3.2 we show that the sequence $(M^N)_{N\geq 1}$ is tight in $D_{\mathcal{M}}[0,\infty)$. We put all these bricks together and conclude in Section 3.3.

3.1 Limiting martingale problem

For $f \in \mathcal{C}(\mathcal{P})$, let us set

$$f^N := \frac{1}{N} f \, \mathbf{1}_{\mathcal{T}_{L,s}^c \times E} + f \, \mathbf{1}_{\mathcal{T}_{L,s} \times E}, \qquad \text{so that} \quad \langle M_t, f^N \rangle = \langle M_t^N, f \rangle, \ \forall t \geq 0. \tag{3.1}$$

Writing the operator G_r in (2.10) for functions of the form $F(\langle \cdot, f^N \rangle)$ (with f^N defined as above), we obtain that the operator describing the dynamics of $(M_t^N)_{t\geq 0}$ due to reaction r, applied to functions of the form $F_f : M \mapsto F(\langle M, f \rangle)$ with $F \in \mathcal{C}_b(\mathbf{R})$ and $f \in \mathcal{C}(\mathcal{P})$, is

$$\begin{aligned} & G_r^N F_f\big(M^N\big) \qquad (3.2) \\ & := \int_{E \times \mathcal{P}^{k_r}} \varrho_r(d\bar{y}) N^{k_{r,b}} (M^N)^{\otimes \downarrow k_r}(dp_1, \cdots dp_{k_r}) h_r^N(\bar{y}, \langle M^N, \Psi_{r,\bar{y}} \rangle) \bigg(\prod_{i=1}^{k_r} \big(\mathbf{1}_{A_i^r}(x_i) \Gamma_\epsilon(y_i - \bar{y}) \big) \bigg) \\ & \quad \times \bigg[F\bigg(\langle M^N, f \rangle - \frac{1}{N} \sum_{i=1}^{k_{r,b}} f(x_i,y_i) - \sum_{i=k_{r,b}+1}^{k_r} f(x_i,y_i) + \frac{1}{N} \sum_{i=1}^{k'_{r,b}} f(B_i^r, \bar{y}) + \sum_{i=k'_{r,b}+1}^{k'_r} f(B_i^r, \bar{y}) \bigg) \\ & \qquad - F_f\big(M^N\big) \bigg], \end{aligned}$$

where we recall that the measure $(M^N)^{\otimes\downarrow k}$ corresponding to sampling k particles from M^N without replacement was defined in (2.21).

Likewise, recalling that only the molecules from abundant species diffuse in space, the operator describing the (independent) motion all non-localized molecules can be derived from (2.14) in the same way as above and is given by

$$\begin{aligned}\mathcal{D}^N F_f(M^N) := F'(\langle M^N, f\rangle) \sum_{x\in\mathcal{T}_{NL}} \langle M^N, b_x\cdot\nabla_y f + \Sigma_x^2\circ\Delta_y f\rangle \\ + \frac{1}{N}F''(\langle M^N, f\rangle)\sum_{x\in\mathcal{T}_{NL}}\langle M^N, \Sigma_x^2\circ\big((\nabla_y f)(\nabla_y f)^{\mathbf{t}}\big)\rangle\end{aligned} \tag{3.3}$$

for every $F\in\mathcal{C}_b^2(\mathbf{R})$, and every $f\in\mathcal{C}^{2,\perp}(\mathcal{P})$. We then have that for each given N, and every function of the form F_f satisfying the above conditions,

$$\left(F_f\big(M_t^N\big) - F_f\big(M_0^N\big) - \int_0^t ds\left\{\sum_{r\in R} G_r^N F_f\big(M_s^N\big) + \mathcal{D}^N F_f(M_s^N)\right\}\right)_{t\geq 0} \tag{3.4}$$

is a martingale.

Let us first consider the terms in the martingale problem corresponding to reactions $r\in R$. We shall treat the cases $r\in R^{n\ell}$ and R^ℓ separately, since the former requires a Taylor expansion to identify the leading term while the latter does not.

Let thus $r\in R^{n\ell}$. In this case, the sums over $i\in\{k_{r,b}+1,\cdots,k_r\}$ and $i\in\{k'_{r,b}+1,\cdots,k'_r\}$ in (3.2) cancel out since reaction r does not modify the counts of species in $\mathcal{T}_{L,s}$ by assumption, meaning that each species $x\in\mathcal{T}_{L,s}$ is neither source nor product of the reaction, or it is both source and product reactant, with $\nu_{r,x}=\nu'_{r,x}$ and (by Assumption (B0)) reaction r is necessarily localized at $\bar y=\bar y_x$. Hence, only the sums over abundant reactants remain. The key idea below is that since the number of molecules of each abundant species tends to infinity when we let N tend to infinity, in a region where the appropriate molecular types are present, sampling without replacement from abundant species is nearly the same as sampling with replacement, up to an error of order $\mathcal{O}(1/N)$ that will vanish as $N\to\infty$.

Using Assumption (A2) on the uniform boundedness of Γ_ϵ and a Taylor expansion of F, we have for every $M\in\mathcal{M}$ of the form (2.20)

$$\begin{aligned}G_r^N F_f(M) = &\int_{E\times\mathcal{P}^{k_r}} \varrho_r(d\bar y) N^{k_{r,b}} M^{\otimes\downarrow k_r}(dp_1,\cdots dp_{k_r}) h_r^N(\bar y,\langle M,\Psi_{r,\bar y}\rangle) \\ &\times\left(\prod_{i=1}^{k_r}\big(\mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar y)\big)\right)\frac{1}{N}F'\big(\langle M,f\rangle\big)\left(\sum_{i=1}^{k'_r} f(B_i^r,\bar y) - \sum_{i=1}^{k_r} f(x_i,y_i)\right) \\ &+\varepsilon_{r,1}^N(M),\end{aligned} \tag{3.5}$$

where the error term $\varepsilon_{r,1}^N(M)$ satisfies

$$|\varepsilon_{r,1}^N(M)| \leq \frac{\|F''\|_\infty[(k_r+k'_r)\|f\|_\infty]^2}{2N}\|\Gamma_\epsilon\|_\infty^{k_r}\left(\int_E \varrho_r(d\bar y)\,\widetilde{h}_r^N(\bar y,\langle M,\Psi_{r,\bar y}\rangle)\right)\langle M,1\rangle^{k_r} \tag{3.6}$$

and $\widetilde{h}_r^N$ was defined in (2.25). Next, if $k_{r,b}<k_r$, for convenience let us suppose that the low-abundance (localized) reactants $A_{k_{r,b}+1}^r,\ldots,A_{k_r}^r$ are listed in such a way that molecules of the same species are grouped together and can be written $a_1^r,\ldots,a_1^r,a_2^r,\ldots,a_2^r$, *etc.*, where a_ι^r

appears with multiplicity ν_{r,a_ι^r}. When $k_{r,b} > 1$ (otherwise sampling of abundant species with or without replacement is the same), we can then replace $M^{\otimes\downarrow k_r}$ by

$$M^{\otimes k_{r,b}} \otimes \left(\bigotimes_\iota \left(M(\{(a_\iota^r, \bar{y}_{a_\iota^r})\})\right)_{\nu_{r,a_\iota^r}} \delta^{\otimes \nu_{r,a_\iota^r}}_{(a_\iota^r, \bar{y}_{a_\iota^r})} \right) \tag{3.7}$$

in the integral on the r.h.s. of (3.5), up to a combinatorial error term

$$\begin{aligned} &\varepsilon^N_{r,2}(M) \\ &:= \int_{E\times\mathcal{P}^{k_r}} \varrho_r(d\bar{y})\big(M^{\otimes\downarrow k_{r,b}} - M^{\otimes k_{r,b}}\big)(dp_1,\dots,dp_{k_{r,b}}) M^{\otimes\downarrow(k_r-k_{r,b})}(dp_{k_{r,b}+1},\dots,dp_{k_r}) \\ &\quad\times \left(\prod_{i=1}^{k_r}\big(\mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\big)\right)\left(\sum_{i=1}^{k_r'} f(B_i^r,\bar{y}) - \sum_{i=1}^{k_r} f(x_i,y_i)\right)\widetilde{h}_r^N(\bar{y},\langle M,\Psi_{r,\bar{y}}\rangle)F'(\langle M,f\rangle). \end{aligned} \tag{3.8}$$

Indeed, the second part of the measure in (3.7) encodes sampling without replacement of the low-abundance localized species involved in reaction r, and so is identical to the measure $M^{\otimes\downarrow(k_r-k_{r,b})}\mathbf{1}_{A^r_{k_{r,b}+1}}\cdots\mathbf{1}_{A^r_{k_r}}$ appearing on the r.h.s. of (3.8). By the relation between $M^{\otimes\downarrow k_{r,b}}$ and $M^{\otimes k_{r,b}}$, this error term satisfies :

$$\begin{aligned} |\varepsilon^N_{r,2}(M)| \le (k_r+k_r')\|f\|_\infty\|F'\|_\infty\|\Gamma_\epsilon\|_\infty^{k_r}\left(\int_E \varrho_r(d\bar{y})\,\widetilde{h}_r^N(\bar{y},\langle M,\Psi_{r,\bar{y}}\rangle)\right)\prod_\iota M(\{(a_\iota^r,y_{a_\iota^r})\})^{\nu_{r,a_\iota^r}} \\ \times\left\langle M^{\otimes k_{r,b}} - M^{\otimes\downarrow k_{r,b}}, \mathbf{1}_{(\mathcal{T}^c_{L,s}\times E)^{k_{r,b}}}\right\rangle. \end{aligned} \tag{3.9}$$

Now,

$$\begin{aligned} &\left\langle M^{\otimes k_{r,b}} - M^{\otimes\downarrow k_{r,b}}, \mathbf{1}_{(\mathcal{T}^c_{L,s}\times E)^{k_{r,b}}}\right\rangle \\ &= \langle M,\mathbf{1}_{\mathcal{T}^c_{L,s}\times E}\rangle^{k_{r,b}} - \frac{1}{N^{k_{r,b}}}\mathrm{Card}\Big(\big\{(i_1,i_2,\dots,i_{k_{r,b}})\in\{1,\dots,N\langle M,\mathbf{1}_{\mathcal{T}^c_{L,s}\times E}\rangle\}^{k_{r,b}} : i_k\neq i_l\ \forall k,l\big\}\Big) \\ &= \langle M,\mathbf{1}_{\mathcal{T}^c_{L,s}\times E}\rangle^{k_{r,b}}\left(1-\prod_{j=0}^{k_{r,b}-1}\left(1-\frac{j}{N\langle M,\mathbf{1}_{\mathcal{T}^c_{L,s}\times E}\rangle}\right)\right), \end{aligned}$$

with the convention that the product on the r.h.s. is 0 if $N\langle M,\mathbf{1}_{\mathcal{T}^c_{L,s}\times E}\rangle < k_{r,b}$. When $k_{r,b} > 1$, this quantity is of order $\mathcal{O}(\langle M,\mathbf{1}_{\mathcal{T}^c_{L,s}\times E}\rangle^{k_{r,b}-1}/N)$ for a fixed M of total mass of order $\mathcal{O}(1)$, which, together with (3.9), gives us :

$$|\varepsilon^N_{r,2}(M)| = \mathcal{O}\left(\frac{\langle M,\mathbf{1}_{\mathcal{T}^c_{L,s}\times E}\rangle^{k_{r,b}-1}}{N}\langle M,1\rangle^{k_r-k_{r,b}}\right)\left(\int_E \varrho_r(d\bar{y})\widetilde{h}_r^N(\bar{y},\langle M,\Psi_{r,\bar{y}}\rangle)\right). \tag{3.10}$$

Assumption (B1) ensures that the integral over $\bar{y}$ in the above is uniformly bounded in N and M.

Let us now consider $r \in R^\ell$. For every $M \in \mathcal{M}$ of the form (2.20), we can first write that

$$\begin{aligned} &F\Big(\langle M, f\rangle - \frac{1}{N}\sum_{i=1}^{k_{r,b}} f(x_i, y_i) - \sum_{i=k_{r,b}+1}^{k_r} f(x_i, y_i) + \frac{1}{N}\sum_{i=1}^{k'_{r,b}} f(B_i^r, \bar{y}) + \sum_{i=k'_{r,b}+1}^{k'_r} f(B_i^r, \bar{y})\Big) \\ &\qquad - F\big(\langle M, f\rangle\big) \\ &= F\Big(\langle M, f\rangle - \sum_{i=k_{r,b}+1}^{k_r} f(x_i, y_i) + \sum_{i=k'_{r,b}+1}^{k'_r} f(B_i^r, \bar{y})\Big) - F\big(\langle M, f\rangle\big) + \eta_r^N \\ &\quad + \frac{1}{N} F'\Big(\langle M, f\rangle - \sum_{i=k_{r,b}+1}^{k_r} f(x_i, y_i) + \sum_{i=k'_{r,b}+1}^{k'_r} f(B_i^r, \bar{y})\Big)\Big(\sum_{i=1}^{k'_{r,b}} f(B_i^r, \bar{y}) - \sum_{i=1}^{k_{r,b}} f(x_i, y_i)\Big), \end{aligned} \tag{3.11}$$

where $|\eta_r^N| \leq (2N^2)^{-1}\|F''\|_\infty (k_r + k'_r)^2 \|f\|_\infty^2$. Plugging this equation in (3.2), we obtain that

$$\begin{aligned} G_r^N F_f(M) = &\int_{E\times\mathcal{P}^{k_r}} \varrho_r(d\bar{y}) N^{k_{r,b}} M^{\otimes\downarrow k_r}(dp_1, \cdots dp_{k_r}) h_r^N(\bar{y}, \langle M, \Psi_{r,\bar{y}}\rangle)\Big(\prod_{i=1}^{k_r}\big(\mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i - \bar{y})\big)\Big) \\ &\times \Big[F\Big(\langle M, f\rangle - \sum_{i=k_{r,b}+1}^{k_r} f(x_i, y_i) + \sum_{i=k'_{r,b}+1}^{k'_r} f(B_i^r, \bar{y})\Big) - F\big(\langle M, f\rangle\big)\Big] + \hat{\varepsilon}_{r,1}^N(M), \end{aligned} \tag{3.12}$$

where

$$\begin{aligned} |\hat{\varepsilon}_{r,1}^N(M)| \leq \|\Gamma_\epsilon\|_\infty^{k_r}\Big(\int_E \varrho_r(d\bar{y})\widetilde{h}_r^N(\bar{y}, \langle M, \Psi_{r,\bar{y}}\rangle)\Big)\langle M, 1\rangle^{k_r}\Big(&\frac{\|F'\|(k_r + k'_r)\|f\|_\infty}{N} \\ &+ \frac{\|F''\|_\infty (k_r + k'_r)^2 \|f\|_\infty^2}{2N^2}\Big). \end{aligned} \tag{3.13}$$

Next, using the same arguments as in the case $r \in R^{n\ell}$, we can replace sampling without replacement of the abundant species in (3.12) by sampling with replacement up to a combinatorial error $\hat{\varepsilon}_{r,2}^N(M)$ of a similar form as (3.8).

Combining the above, we obtain that for $M \in \mathcal{M}$ of the form (2.20),

$$\begin{aligned}
\sum_{r\in R} G_r^N F_f(M) = F'\big(\langle M, f\rangle\big) \sum_{r\in R^{n\ell}} \int_{E\times\mathcal{P}^{k_{r,b}}} \varrho_r(d\bar{y}) M^{\otimes k_{r,b}}(dp_1,\cdots dp_{k_{r,b}}) \widetilde{h}_r^N(\bar{y}, \langle M, \Psi_{r,\bar{y}}\rangle) \\
\times \Big(\prod_{i=1}^{k_{r,b}} \big(\mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\big)\Big)\Big(\prod_{x\in\mathcal{T}_{L,s}} \Big\{\big(M(\{(x,\bar{y}_x)\})\big)_{\nu_{r,x}} \Gamma_\epsilon(\bar{y}_x-\bar{y})^{\nu_{r,x}}\Big\}\Big) \\
\times \Big(\sum_{i=1}^{k'_r} f(B_i^r,\bar{y}) - \sum_{i=1}^{k_r} f(x_i,y_i)\Big) \\
+ \sum_{r\in R^\ell} \int_{E\times\mathcal{P}^{k_{r,b}}} \varrho_r(d\bar{y}) M^{\otimes k_{r,b}}(dp_1,\cdots dp_{k_{r,b}}) \widetilde{h}_r^N(\bar{y}, \langle M, \Psi_{r,\bar{y}}\rangle) \\
\times \Big(\prod_{i=1}^{k_{r,b}} \big(\mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\big)\Big)\Big(\prod_{x\in\mathcal{T}_{L,s}} \Big\{\big(M(\{(x,\bar{y}_x)\})\big)_{\nu_{r,x}} \Gamma_\epsilon(\bar{y}_x-\bar{y})^{\nu_{r,x}}\Big\}\Big) \\
\times \Big(F\Big(\langle M, f\rangle + \sum_{i=k'_{r,b}+1}^{k'_r} f(B_i^r,\bar{y}) - \sum_{i=k_{r,b}+1}^{k_r} f(x_i,y_i)\Big) - F_f(M)\Big) \\
+ \sum_{r\in R^{n\ell}} \big(\varepsilon_{r,1}^N(M) + \varepsilon_{r,2}^N(M)\big) + \sum_{r\in R^\ell} \big(\hat{\varepsilon}_{r,1}^N(M) + \hat{\varepsilon}_{r,2}^N(M)\big). \qquad (3.14)
\end{aligned}$$

Let us use these results to show that any limit point of a subsequence of $(M^N)_{N\geq 1}$ satisfies the $D_{\mathcal{M}}[0,\infty)$-martingale problem stated in Theorem 2.11, for test functions of the form (2.9) with $F \in \mathcal{C}_b^2(\mathbf{R})$ and $f \in \mathcal{C}^{2,\perp}(\mathcal{P})$. The fact that this property holds true when $F \in \mathcal{C}_b^1(\mathbf{R})$ can then be obtained by a simple density argument, and we shall argue that the trajectories of the limit point(s) of $(M^N)_{N\geq 1}$ remain in $D_{\mathcal{M}_{\mathcal{T}_L}}[0,\infty)$ in Remark 3.1(b) at the end of this Section.

Let thus $F \in \mathcal{C}_b^2(\mathbf{R})$ and $f \in \mathcal{C}^{2,\perp}(\mathcal{P})$. By Theorem 2.6($ii$) (see also (3.4)), we know that for any given $N \in \mathbf{N}$,

$$\Big(F_f\big(M_t^N\big) - F_f\big(M_0^N\big) - \int_0^t ds \Big\{\sum_{r\in R} G_r^N F_f\big(M_s^N\big) + \mathcal{D}^N F_f\big(M_s^N\big)\Big\}\Big)_{t\geq 0}$$

is a martingale. Let us prove that for every $t, t' \geq 0$, $k \in \mathbf{N}$, $0 \leq t_1 < \ldots < t_k \leq t$ and $\beta_1, \ldots, \beta_k \in \mathcal{C}_b(\mathcal{M})$, we have along any subsequence (which we also denote by $(M^N)_{N\geq 1}$ for simplicity) converging to a limit $(\bar{M}_t)_{t\geq 0}$

$$\begin{aligned}
\lim_{N\to\infty} \mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(M_{t_j}^N\big)\Big)\Big(F_f\big(M_{t+t'}^N\big) - F_f\big(M_t^N\big) \qquad (3.15) \\
- \int_t^{t+t'} ds \Big\{\sum_{r\in R} G_r^N F_f\big(M_s^N\big) + \mathcal{D}^N F_f\big(M_s^N\big)\Big\}\Big)\Big] \\
= \mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(\bar{M}_{t_j}\big)\Big)\Big(F_f\big(\bar{M}_{t+t'}\big) - F_f\big(\bar{M}_t\big) \\
- \int_t^{t+t'} ds \Big\{\sum_{r\in R} G_r^\infty F_f\big(\bar{M}_s\big) + \mathcal{D}^\infty F_f\big(\bar{M}_s\big)\Big\}\Big)\Big],
\end{aligned}$$

where

$$G_r^\infty F_f(M) := F'\big(\langle M, f\rangle\big) \int_{E\times\mathcal{P}^{k_{r,b}}} \varrho_r(d\bar{y}) M^{\otimes k_{r,b}}(dp_1, \cdots dp_{k_{r,b}}) \widetilde{h}_r(\bar{y}, \langle M, \Psi_{r,\bar{y}}\rangle) \tag{3.16}$$
$$\times \Big(\prod_{i=1}^{k_{r,b}} \big(\mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\big)\Big)\Big(\prod_{x\in\mathcal{T}_{L,s}} \Big\{\big(M(\{(x,\bar{y}_x)\})\big)_{\nu_{r,x}} \Gamma_\epsilon(\bar{y}_x-\bar{y})^{\nu_{r,x}}\Big\}\Big)$$
$$\times \Big(\sum_{i=1}^{k_r'} f(B_i^r,\bar{y}) - \sum_{i=1}^{k_r} f(x_i,y_i)\Big)$$

for $r \in R^{n\ell}$,

$$G_r^\infty F_f(M) := \int_{E\times\mathcal{P}^{k_{r,b}}} \varrho_r(d\bar{y}) M^{\otimes k_{r,b}}(dp_1, \cdots dp_{k_{r,b}}) \widetilde{h}_r(\bar{y}, \langle M, \Psi_{r,\bar{y}}\rangle) \tag{3.17}$$
$$\times \Big(\prod_{i=1}^{k_{r,b}} \big(\mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\big)\Big)\Big(\prod_{x\in\mathcal{T}_{L,s}} \Big\{\big(M(\{(x,\bar{y}_x)\})\big)_{\nu_{r,x}} \Gamma_\epsilon(\bar{y}_x-\bar{y})^{\nu_{r,x}}\Big\}\Big)$$
$$\times \Big(F\Big(\langle M, f\rangle + \sum_{i=k_{r,b}'+1}^{k_r'} f(B_i^r,\bar{y}) - \sum_{i=k_{r,b}+1}^{k_r} f(x_i,y_i)\Big) - F_f(M)\Big)$$

for $r \in R^\ell$, and

$$\mathcal{D}^\infty F_f(M) = F'(\langle M, f\rangle) \sum_{x\in\mathcal{T}_{NL}} \langle M, b_x\cdot\nabla_y f + \Sigma_x^2 \circ \Delta_y f\rangle. \tag{3.18}$$

Since the expectation on the l.h.s. of (3.15) is zero for every N, (3.15) will show that the limit $\bar{M}$ solves the desired martingale problem.

We proceed in three steps.

Step 1. By assumption, $(M_{t_1}^N, \ldots, M_{t_k}^N, M_t^N, M_{t+t'}^N)$ converges in distribution to $(\bar{M}_{t_1}, \ldots, \bar{M}_{t+t'})$ as $N \to \infty$ (along the subsequence considered). Since $F_f, \beta_1, \ldots, \beta_k$ are bounded continuous functions on $\mathcal{M}$, we obtain that

$$\lim_{N\to\infty} \mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(M_{t_j}^N\big)\Big)\Big(F_f\big(M_{t+t'}^N\big) - F_f\big(M_t^N\big)\Big)\Big] = \mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(\bar{M}_{t_j}\big)\Big)\Big(F_f\big(\bar{M}_{t+t'}\big) - F_f\big(\bar{M}_t\big)\Big)\Big].$$

Step 2. Let us show that

$$\lim_{N\to\infty} \mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(M_{t_j}^N\big)\Big)\int_t^{t+t'} ds\, \big|G_r^N F_f\big(M_s^N\big) - G_r^\infty F_f\big(M_s^N\big)\big|\Big] = 0 \quad \forall r \in R, \text{ and} \tag{3.19}$$

$$\lim_{N\to\infty} \mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(M_{t_j}^N\big)\Big)\int_t^{t+t'} ds\, \big|\mathcal{D}^N F_f\big(M_s^N\big) - \mathcal{D}^\infty F_f\big(M_s^N\big)\big|\Big] = 0. \tag{3.20}$$

Starting with (3.19) for $r \in R^{n\ell}$ and using (3.14) and then (3.6), we can write that for every $M \in \mathcal{M}$ of the form (2.20),

$$\begin{aligned}
&\big|G_r^N F_f(M) - G_r^\infty F_f(M)\big| \\
&\le (k_r + k_r')\|F'\|_\infty \|f\|_\infty \|\Gamma_\epsilon\|_\infty^{k_r} \langle M, 1\rangle^{k_r} \int_E \varrho_r(d\bar{y}) \Big|\widetilde{h}_r^N\big(\bar{y}, \langle M, \Psi_{r,\bar{y}}\rangle\big) - \widetilde{h}_r\big(\bar{y}, \langle M, \Psi_{r,\bar{y}}\rangle\big)\Big| \\
&\qquad + |\varepsilon_{r,1}^N(M)| + |\varepsilon_{r,2}^N(M)| \\
&\le C_1 \langle M, 1\rangle^{k_r} \int_E \varrho_r(d\bar{y}) \Big|\widetilde{h}_r^N\big(\bar{y}, \langle M, \Psi_{r,\bar{y}}\rangle\big) - \widetilde{h}_r\big(\bar{y}, \langle M, \Psi_{r,\bar{y}}\rangle\big)\Big| \\
&\qquad + \frac{C_2}{N} \langle M, 1\rangle^{k_r} \int_E \varrho_r(d\bar{y})\, \widetilde{h}_r^N(\bar{y}, \langle M, \Psi_{r,\bar{y}}\rangle) + |\varepsilon_{r,2}^N(M)|.
\end{aligned}$$

Let us focus on the first two terms on the r.h.s. We have

$$\begin{aligned}
&\mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(M_{t_j}^N\big)\Big) \int_t^{t+t'} ds\, \langle M_s^N, 1\rangle^{k_r} \Big\{C_1 \int_E \varrho_r(d\bar{y}) \Big|\widetilde{h}_r^N\big(\bar{y}, \langle M_s^N, \Psi_{r,\bar{y}}\rangle\big) - \widetilde{h}_r\big(\bar{y}, \langle M_s^N, \Psi_{r,\bar{y}}\rangle\big)\Big| \\
&\qquad\qquad + \frac{C_2}{N} \int_E \varrho_r(d\bar{y}) \widetilde{h}_r^N(\bar{y}, \langle M, \Psi_{r,\bar{y}}\rangle)\Big\}\Big] \\
&\le \Big(\prod_{j=1}^k \|\beta_j\|_\infty\Big)\Big(C_1 \int_E \varrho_r(d\bar{y}) \sup_{a\ge 0} \Big|\widetilde{h}_r^N(\bar{y}, a) - \widetilde{h}_r(\bar{y}, a)\Big| + \frac{C_2}{N} \int_E \varrho_r(d\bar{y}) \sup_{a\ge 0} \widetilde{h}_r^N(\bar{y}, a)\Big) \\
&\qquad \times \int_t^{t+t'} ds\, \mathbf{E}\Big[\langle M_s^N, 1\rangle^{k_r}\Big].
\end{aligned}$$

Assumption (B1) guarantees that the second bracketed expression goes to 0 as $N \to \infty$. In addition, Assumption (B2) implies that the last integral is bounded by some constant depending on $T := t + t'$ but not on N, and so the whole quantity on the above r.h.s. converges to 0 as $N \to \infty$.

Next, let us control the term involving $\varepsilon_{r,2}^N$, which is nonzero only if $k_{r,b} \ge 2$. Recall (3.10), where the $\mathcal{O}(1/N)$ asymptotics are valid only if $\langle M, \mathbf{1}_{\mathcal{T}_{L,s}^c \times E}\rangle$ is not too close to 0. Hence, let us fix $\delta > 0$ and let us decompose the bound on $\varepsilon_{r,2}^N(M_s^N)$ according to whether $\langle M_s^N, \mathbf{1}_{\mathcal{T}_{L,s}^c \times E}\rangle < \delta$ or not. We obtain, writing $C_\delta > 0$ for the constant independent of N entering the first estimate and using (in this case only) that $\langle M_s^N, \mathbf{1}_{\mathcal{T}_{L,s}^c \times E}\rangle^{k_{r,b}-1} \le \langle M_s^N, 1\rangle^{k_{r,b}-1}$,

$$\begin{aligned}
&\mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(M_{t_j}^N\big)\Big) \int_t^{t+t'} ds\, \big|\varepsilon_{r,2}^N\big(M_s^N\big)\big|\Big] \\
&\le C_r \Big(\prod_{j=1}^k \|\beta_j\|_\infty\Big)\Big(\int_E \varrho_r(d\bar{y}) \sup_{a\ge 0} \widetilde{h}_r^N(\bar{y}, a)\Big) \int_t^{t+t'} ds\, \mathbf{E}\Big[\mathbf{1}_{\{\langle M_s^N, \mathbf{1}_{\mathcal{T}_{L,s}^c \times E}\rangle \ge \delta\}} \frac{C_\delta \langle M_s^N, 1\rangle^{k_r - 1}}{N} \\
&\qquad\qquad + \mathbf{1}_{\{\langle M_s^N, \mathbf{1}_{\mathcal{T}_{L,s}^c \times E}\rangle < \delta\}} \langle M_s^N, \mathbf{1}_{\mathcal{T}_{L,s}^c \times E}\rangle^{k_{r,b}} \langle M_s^N, 1\rangle^{k_r - k_{r,b}}\Big].
\end{aligned}$$

By Assumption (B2), $\mathbf{E}[\langle M_s^N, 1\rangle^{k_r-1}]$ is bounded uniformly in $N \ge 1$ and $s \in [0, t+t']$, and so the first term in the expectation is of the order of $\mathcal{O}(1/N)$. Furthermore, the second term is bounded by $\delta^{k_{r,b}} \mathbf{E}[\langle M_s^N, 1\rangle^{k_r - k_{r,b}}]$. Consequently, using the uniform bound on the integral of

$\widetilde{h}_r^N$ derived from Assumption (B1), the fact that $k_r - k_{r,b} < k_r - 1$ and Assumption (B2), and letting N tend to infinity, we obtain that

$$\limsup_{N\to\infty} \mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(M_{t_j}^N\big)\Big)\int_t^{t+t'} ds\,\big|\varepsilon_2\big(M_s^N\big)\big|\Big] \le Ct'\delta^{k_{r,b}}$$

for a constant C independent of δ. Since this is true for any $\delta > 0$ and since we consider only the case $k_{r,b} \ge 2$, we can conclude that this limit is actually 0. Combining the above, we obtain (3.19) for $r \in R^{n\ell}$. Exactly the same reasoning applies and yields (3.19) for $r \in R^\ell$.

The convergence result (3.20) is obtained by noticing that

$$\Big|\frac{1}{N}\,F''(\langle M,f\rangle)\langle M,\Sigma_x^2\circ\big((\nabla_y f)(\nabla_y f)^{\mathbf{t}}\big)\rangle\Big| \le \frac{1}{N}\|F''\|_\infty\|\Sigma_x^2\circ\big((\nabla_y f)(\nabla_y f)^{\mathbf{t}}\big)\|_\infty\,\langle M,1\rangle,$$

and using the same arguments as before.

Step 3. Finally, let us show that

$$\begin{aligned}&\lim_{N\to\infty}\mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(M_{t_j}^N\big)\Big)\Big(\int_t^{t+t'} ds\Big\{\sum_{r\in R} G_r^\infty F_f\big(M_s^N\big)+\mathcal{D}^\infty F_f\big(M_s^N\big)\Big\}\Big)\Big]\\ &=\mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(\bar{M}_{t_j}\big)\Big)\Big(\int_t^{t+t'} ds\Big\{\sum_{r\in R} G_r^\infty F_f\big(\bar{M}_s\big)+\mathcal{D}^\infty F_f\big(\bar{M}_s\big)\Big\}\Big)\Big].\end{aligned} \tag{3.21}$$

By construction and Assumption (B1), we have for every $M \in \mathcal{M}$

$$\big|G_r^\infty F_f(M)\big| \le \langle M,1\rangle^{k_r}\mathcal{S}_r\|\Gamma_\epsilon\|_\infty^{k_r}\|F'\|_\infty(k_r+k_r')\|f\|_\infty \tag{3.22}$$

when $r \in R^{n\ell}$,

$$\big|G_r^\infty F_f(M)\big| \le \langle M,1\rangle^{k_r}\mathcal{S}_r\|\Gamma_\epsilon\|_\infty^{k_r}2\|F\|_\infty \tag{3.23}$$

when $r \in R^\ell$, and

$$\big|\mathcal{D}^\infty F_f(M)\big| \le \langle M,1\rangle\|F'\|_\infty\sum_{x\in\mathcal{T}_{NL}}\|b_x\cdot\nabla_y f+\Sigma_x^2\circ\Delta_y f\|_\infty. \tag{3.24}$$

Thanks to these bounds, together with Assumption (B2) which ensures that each of these quantities is integrable, we can use Fubini's theorem to exchange order between integration and summation in the above and treat each term separately.

Starting with the terms corresponding to reactions $r \in R$, if $k_r = 0$, a simple dominated convergence argument together with the continuity of $G_r^\infty F_f$ suffice to conclude. If $k_r \ge 1$, we use (3.22) to observe that the function $G_r^\infty F_f$ is bounded over any subset of $\mathcal{M}$ of measures with total mass less than a given quantity. Hence, let $\delta > 0$ be small, g be a continuous function with values in $[0,1]$, such that $g(a) = 1$ if $a \le -\delta$ and $g(a) = 0$ if $a > \delta$, and let $A > 0$ be a large constant. We use $g(\cdot - A)$ as a continuous approximation to $\mathbf{1}_{\{\cdot\le A\}}$. We have

$$\begin{aligned}&\int_t^{t+t'} ds\,\mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(M_{t_j}^N\big)\Big)G_r^\infty F_f\big(M_s^N\big)\Big]\\ &\quad=\int_t^{t+t'} ds\,\mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(M_{t_j}^N\big)\Big)G_r^\infty F_f\big(M_s^N\big)g\big(\langle M_s^N,1\rangle - A\big)\Big]\\ &\qquad+\int_t^{t+t'} ds\,\mathbf{E}\Big[\Big(\prod_{j=1}^k \beta_j\big(M_{t_j}^N\big)\Big)G_r^\infty F_f\big(M_s^N\big)\big(1-g\big(\langle M_s^N,1\rangle - A\big)\big)\Big].\end{aligned} \tag{3.25}$$

Since $G_r^\infty F_f(\cdot) g\big(\langle \cdot, 1\rangle - A\big)$ is bounded and continuous (as well as each β_j), the dominated convergence theorem guarantees that the first term on the r.h.s. of (3.25) converges to

$$\int_t^{t+t'} ds\, \mathbf{E}\Bigg[\Bigg(\prod_{j=1}^k \beta_j\big(\bar{M}_{t_j}\big)\Bigg) G_r^\infty F_f\big(\bar{M}_s\big) g\big(\langle \bar{M}_s, 1\rangle - A\big)\Bigg] \tag{3.26}$$

as $N \to \infty$ (along the converging subsequence). On the other hand, the second term on the r.h.s. of (3.25) is bounded by

$$\begin{aligned} &C \int_t^{t+t'} ds\, \mathbf{E}\Big[\langle M_s^N, 1\rangle^{k_r} \big(1 - g\big(\langle M_s^N, 1\rangle - A\big)\big)\Big] \\ &\le C \int_t^{t+t'} ds\, \mathbf{E}\Big[\langle M_s^N, 1\rangle^{pk_r}\Big]^{1/p} \mathbf{E}\Big[\big(1 - g\big(\langle M_s^N, 1\rangle - A\big)\big)^q\Big]^{1/q} \end{aligned} \tag{3.27}$$

for some constant C independent of N, A, and any pair (p, q) such that $p > 1$, $1/p + 1/q = 1$ and $pk_r \le k^* \vee 2$ (so that we may use Hölder's inequality to pass from the first to the second line). By Assumption (B2), the first expectation on the r.h.s. is bounded uniformly in N and $s \in [0, t + t']$. Because $1 - g$ is bounded, continuous and satisfies $1 - g(\cdot - A) \le \mathbf{1}_{\{\cdot > A - \delta\}}$, we have

$$\begin{aligned} \lim_{N\to\infty} \int_t^{t+t'} ds\, \mathbf{E}\Big[\big(1 - g\big(\langle M_s^N, 1\rangle - A\big)\big)^q\Big]^{1/q} &= \int_t^{t+t'} ds\, \mathbf{E}\Big[\big(1 - g\big(\langle \bar{M}_s, 1\rangle - A\big)\big)^q\Big]^{1/q} \\ &\le \int_t^{t+t'} ds\, \mathbf{P}\big[\langle \bar{M}_s, 1\rangle > A - \delta\big]^{1/q}. \end{aligned}$$

Consequently, we have

$$\begin{aligned} &\limsup_{N\to\infty} \int_t^{t+t'} ds\, \mathbf{E}\Bigg[\Bigg(\prod_{j=1}^k \beta_j\big(M_{t_j}^N\big)\Bigg) G_r^\infty F_f\big(M_s^N\big)\big(1 - g\big(\langle M_s^N, 1\rangle - A\big)\big)\Bigg] \\ &\qquad \le C' \int_t^{t+t'} ds\, \mathbf{P}\big[\langle \bar{M}_s, 1\rangle > A - \delta\big]^{1/q}, \end{aligned}$$

where the constant C' is independent of A. Since the integrand is bounded by 1, and tends to 0 as A tends to infinity provided that we know that $\langle \bar{M}_s, 1\rangle < \infty$ a.s. for every $s \in [t, t + t']$ (see below for a proof), we can use the dominated convergence theorem to conclude that the quantity on the last line tends to 0 as $A \to \infty$. It only remains to show that as $A \to \infty$, the quantity in (3.26) tends to

$$\int_t^{t+t'} ds\, \mathbf{E}\Bigg[\Bigg(\prod_{j=1}^k \beta_j\big(\bar{M}_{t_j}\big)\Bigg) G_r^\infty F_f\big(\bar{M}_s\big)\Bigg].$$

But using Fatou's Lemma and Assumption (B2), we see that

$$\sup_{s\in[0,T]} \mathbf{E}\big[\langle \bar{M}_s, 1\rangle^{k^*\vee 2}\big] < \infty, \qquad \forall T > 0. \tag{3.28}$$

The first consequence is that, as supposed in the previous paragraph, $\langle \bar{M}_s, 1\rangle$ is a.s. finite for every s. Second, by (3.22), $G_r^\infty F_f(\bar{M}_s)$ is also integrable (and so is $\mathcal{D}^\infty F_f(\bar{M}_s)$, by the same

arguments). Since g is bounded by 1 and $g(\cdot - A)$ converges pointwise to 1 as A tends to infinity, the dominated convergence theorem gives us the desired convergence.

Exactly the same chain of arguments gives us the convergence of the term in (3.21) involving $\mathcal{D}^\infty F_f$, and so we do not repeat it.

As desired, we have proved that any limit point of a convergent subsequence of $(M^N)_{N\in\mathbf{N}}$ satisfies the martingale problem stated in Theorem 2.11 (with trajectories that are $\mathcal{M}$-valued).

Remark 3.1. (a) By taking two sequences of functions $(F^{1,n})_{n\geq 1}$ and $(F^{2,n})_{n\geq 1}$ in $\mathcal{C}_b^2(\mathbf{R})$, each with bounded first and second derivatives, and converging respectively to the identity function and to $a \mapsto a^2$ uniformly over compact intervals and such that their first derivatives converge respectively to 1 and $a \mapsto 2a$ uniformly over compact intervals, we can show (using again Assumptions (B1-B2) and the different bounds of $G_r^N F_f^{i,n}$ and $\mathcal{D}^N F_f^{i,n}$ obtained in this section) that the pre-limiting and limiting martingale problems hold also with $F^1(a) = a$ and $F^2(a) = a^2$. For F^2, observe that the component $(F^2)'(\langle M, f\rangle)$ of some of the terms in the (approximations to the) operators involved adds another factor $\langle M, 1\rangle$ in all the bounds. However, Assumption (B2) controls the $(k^* \vee 2)$-th moment of the total mass of the processes, where $k^* = 1 + \max_r k_r$, which is precisely the moment appearing in those bounds.

(b) Likewise, by taking a sequence $(f_n)_{n\geq 1}$ in $\mathcal{C}^{2,\perp}(\mathcal{P})$ approximating $f = \mathbf{1}_{\{(x,\cdot)\}} - \mathbf{1}_{\{(x,\bar{y}_x)\}}$ uniformly over $\mathcal{P}$, for any $x \in \mathcal{T}_L$, we can show that any limit point of $(M^N)_{N\geq 1}$ has trajectories taking their values in $\mathcal{M}_{\mathcal{T}_L}$ a.s. (provided the sequence of initial values $(M_0^N)_{N\geq 1}$ converges in $\mathcal{M}_{\mathcal{T}_L}$).

3.2 Tightness of $(M^N)_{N\geq 1}$

Using the tightness criterion for measure-valued processes given in Theorem 2.1 of [47] together with the fact that $\mathcal{C}^{2,\perp}(\mathcal{P})$ is dense in $\mathcal{C}(\mathcal{P})$ for the topology of uniform convergence (and contains the constant function equal to 1), tightness of the sequence of laws of M^N will be proved if we show that the sequence of laws of $\langle M^N, f\rangle$ is tight for every $f \in \mathcal{C}^{2,\perp}(\mathcal{P})$.

To start with, let us show the following *compact containment condition*, which will give us the tightness of $(\langle M_t^N, f\rangle)_{N\geq 1}$ for each $t \geq 0$ and each $f \in \mathcal{C}^{2,\perp}(\mathcal{P})$. Since $\mathcal{P}$ is compact, for every $a > 0$ the set

$$E_a := \{M \in \mathcal{M} : \langle M, 1\rangle \leq a\}$$

is a compact subset of $\mathcal{M}$. By Assumption (B2) and the Markov inequality, for every $T > 0$ and every $\eta \in (0,1)$ there exists $a_{T,\eta} > 0$ such that

$$\inf_{N\geq 1} \mathbf{P}\big(M_t^N \in E_{a_{T,\eta}},\ \forall t \in [0,T]\big) \geq 1 - \eta,$$

and so the compact containment condition holds indeed.

We now call on the criterion from Aldous [1] and Rebolledo [44] and use the semi-martingale decomposition of $(\langle M_t^N, f\rangle)_{t\geq 0}$ to show that this sequence of processes is tight. More precisely, we use the following lemma.

Lemma 3.2. *Let $f \in \mathcal{C}^{2,\perp}(\mathcal{P})$ and let $N \in \mathbf{N}$. For every $t \geq 0$, let us define V_t^N and Z_t^N by:*

$$V_t^N := \int_0^t ds \left\{ \sum_{r\in R} G_r^N \mathrm{Id}_f(M_s^N) + \mathcal{D}^N \mathrm{Id}_f(M_s^N) \right\},$$

where $\mathrm{Id}_f(M) = \langle M, f\rangle$, *and*

$$Z_t^N := \langle M_t^N, f\rangle - \langle M_0^N, f\rangle - V_t^N.$$

Then the process $(Z_t^N)_{t\geq 0}$ *is a square integrable martingale, with predictable quadratic variation*

$$\begin{aligned}
\big\langle Z^N\big\rangle_t = \frac{1}{N}\int_0^t ds \Big\{ &\sum_{r\in R^{n\ell}} \int_{E\times\mathcal{P}^{k_r}} \varrho_r(d\bar{y})\big(M_s^N\big)^{\otimes\downarrow k_r}(dp_1,\dots,dp_{k_r})\widetilde{h}_r^N\big(\bar{y},\langle M_s^N,\Psi_{r,\bar{y}}\rangle\big)\\
&\times\Big(\prod_{i=1}^{k_r}\big(\mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\big)\Big)\Big(\sum_{i=1}^{k_r'} f\big(B_i^r,\bar{y}\big)-\sum_{i=1}^{k_r} f(x_i,y_i)\Big)^2\Big\}\\
+\int_0^t ds\Big\{ &\sum_{r\in R^{\ell}} \int_{E\times\mathcal{P}^{k_r}} \varrho_r(d\bar{y})\big(M_s^N\big)^{\otimes\downarrow k_r}(dp_1,\dots,dp_{k_r})\widetilde{h}_r^N\big(\bar{y},\langle M_s^N,\Psi_{r,\bar{y}}\rangle\big)\\
&\times\Big(\prod_{i=1}^{k_r}\big(\mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\big)\Big)\\
&\times\Big(\frac{1}{N}\sum_{i=1}^{k_{r,b}'} f\big(B_i^r,\bar{y}\big)+\sum_{i=k_{r,b}'+1}^{k_r'} f\big(B_i^r,\bar{y}\big)-\frac{1}{N}\sum_{i=1}^{k_{r,b}} f(x_i,y_i)-\sum_{i=k_{r,b}+1}^{k_r} f(x_i,y_i)\Big)^2\Big\}\\
+\frac{2}{N}\int_0^t ds\Big\{ &\sum_{x\in\mathcal{T}_{NL}}\big\langle M_s^N,\Sigma_x^2\circ\big((\nabla_y f)(\nabla_y f)^{\mathbf{t}}\big)\big\rangle\Big\}.
\end{aligned}$$

Proof. (Proof of Lemma 3.2.) By Remark 3.1(a), the processes $(\langle M_t^N, f\rangle)_{t\geq 0}$ and $(\langle M_t^N, f\rangle^2)_{t\geq 0}$ satisfy the martingale problem (3.4) written with $F = \mathrm{Id}$ and $F : a \mapsto a^2$ respectively. Based on this, the identification of the predictable finite variation and quadratic variation terms in the semi-martingale decomposition of $\langle M^N, f\rangle$ is standard (see, *e.g.*, the proof of Theorem 3.3 in [13]). Note that the factor $1/N$ in the first term of the expression for $\langle Z^N\rangle_t$ comes from the fact that for $r \in R^{n\ell}$, $(1/N^2)\times N^{k_{r,b}}h_r^N(\bar{y},\langle M,\Psi_{r,\bar{y}}\rangle) = (1/N)\widetilde{h}_r^N(\bar{y},\langle M,\Psi_{r,\bar{y}}\rangle)$ for all $\bar{y}, M$ by definition of $\widetilde{h}_r^N$. □

Let us now fix $f \in \mathcal{C}^{2,\perp}(\mathcal{P})$. Let $T > 0$ and $(\tau_N)_{N\geq 1}$ be a sequence of stopping times bounded by T. Using Lemma 3.2 and (3.2–3.3) with $F = \mathrm{Id}$, we can write for every $t \in [0, 1]$

$$\begin{aligned}
\big|V_{\tau^N+t}^N - V_{\tau^N}^N\big| &= \Big|\int_{\tau^N}^{\tau^N+t} ds\Big\{\sum_{r\in R} G_r^N\mathrm{Id}_f(M_s^N)+\mathcal{D}^N\mathrm{Id}_f(M_s^N)\Big\}\Big|\\
&\leq \int_{\tau^N}^{\tau^N+t} ds\Big\{\sum_{r\in R}\big|G_r^N\mathrm{Id}_f(M_s^N)\big|+\big|\mathcal{D}^N\mathrm{Id}_f(M_s^N)\big|\Big\}\\
&\leq \sum_{r\in R}\Big[(k_r+k_r')\|f\|_\infty\|\Gamma_\epsilon\|_\infty^{k_r}\Big(\int_E \varrho_r(d\bar{y})\sup_{a\geq 0}\widetilde{h}_r^N(\bar{y},a)\Big)\int_{\tau^N}^{\tau^N+t} ds\,\langle M_s^N,1\rangle^{k_r}\Big]\\
&\quad+\sum_{x\in\mathcal{T}_{NL}}\Big[\|b_x\cdot\nabla_y f+\Sigma_x^2\circ\Delta_y f\|_\infty\int_{\tau^N}^{\tau^N+t} ds\,\langle M_s^N,1\rangle\Big].
\end{aligned}$$

Since $\tau^N + t \leq T + 1$, Assumption (B2) guarantees that

$$\mathbf{E}\Big[\sup_{s\in[\tau^N,\tau^N+t]}\langle M_s^N,1\rangle^k\Big]\leq\mathbf{E}\Big[\sup_{s\in[0,T+1]}\langle M_s^N,1\rangle^k\Big]$$

is bounded uniformly in N for every $k \leq k^* \vee 2$. Therefore, using Fubini's theorem as well as Assumption (B1) to control the terms involving the $\widetilde{h}_r^N$, we obtain the existence of a constant $C_T > 0$ independent of N such that for every $t \in [0,1]$,

$$\mathbf{E}\Big[\big|V^N_{\tau^N+t} - V^N_{\tau^N}\big|\Big] \leq C_T t.$$

Using the Markov inequality, we can therefore conclude that for any $\eta > 0$, there exists $\delta = \delta(\eta, T) > 0$ such that

$$\sup_{N\geq 1}\sup_{t\in[0,\delta]} \mathbf{P}\Big[\big|V^N_{\tau^N+t} - V^N_{\tau^N}\big| > \eta\Big] \leq \eta,$$

and the first part of the Aldous-Rebolledo criterion is satisfied. Likewise, there exists $C'_{T,n\ell}, C'_{T,\ell}$ independent of N such that

$$\mathbf{E}\Big[\big|\langle Z^N\rangle_{\tau^N+t} - \langle Z^N\rangle_{\tau^N}\big|\Big] \leq \frac{C'_{T,n\ell}\mathrm{Card}(R^{n\ell})}{N}t + C'_{T,\ell}\mathrm{Card}(R^\ell)t, \tag{3.29}$$

and so the same conclusion as above holds for the quadratic variation process $(\langle Z^N\rangle_t)_{t\geq 0}$. The second part of the Aldous-Rebolledo criterion is satisfied too, and hence the sequence $(\langle M^N, f\rangle)_{N\geq 1}$ is tight. This completes the proof of the tightness of $(M^N)_{N\geq 1}$ in $D_{\mathcal{M}}[0,\infty)$.

3.3 Conclusion of the proof of Theorem 2.11

To conclude, let us write L^∞ for the operator on which the martingale problem (2.29) is based :

$$L^\infty := \sum_{r\in R} G_r^\infty + \mathcal{D}^\infty, \tag{3.30}$$

where G_r^∞ and $\mathcal{D}^\infty$ were defined respectively in (3.16–3.17) and (3.18). We use Theorem 4.8.10 in [20], in which the condition that the operator L^∞ should take its values in $\mathcal{C}_b(\mathcal{M})$ is replaced by the bounds (3.22), (3.23) and (3.24) which, together with (3.28), ensure that the limiting local martingales

$$\Big(F_f(M_t^\infty) - F_f(M_0^\infty) - \int_0^t ds\, L^\infty F_f(M_s^\infty)\Big)_{t\geq 0}$$

are integrable and are therefore true martingales for every limit M^∞ of a converging subsequence of $(M^N)_{N\geq 1}$. By Remark 3.1(b), such a limit takes its values in $\mathcal{M}_{\mathcal{T}_L}$ a.s. Equation (3.15) shows that Condition (b') of Theorem 4.8.10 is satisfied and since we assumed that there was at most one solution to the $D_{\mathcal{M}_{\mathcal{T}_L}}[0,\infty)$-martingale problem (2.29), we thus obtain that the unique solution M^∞ to the $D_{\mathcal{M}_{\mathcal{T}_L}}[0,\infty)$-martingale problem indeed exists and that M^N converges to it in distribution as $N \to \infty$, in $D_{\mathcal{M}}[0,\infty)$. Using again Remark 3.1(b) and Corollary 3.3.2 in [20], we can conclude that the convergence also holds in $D_{\mathcal{M}_{\mathcal{T}_L}}[0,\infty)$. Remark 3.1($a$) shows that the limiting martingale problem also holds for $F = \mathrm{Id}$ and $F : a \mapsto a^2$, and so Theorem 2.11(i) is proved.

The fact that if uniqueness of the solution to the limiting martingale problem holds for every initial condition $m_0 \in \mathcal{M}_{\mathcal{T}_L}$, then M^∞ has the Markov property, is a consequence of Theorem 4.4.2(a) in [20]. The proof of Theorem 2.11 is now complete.

4 Properties of the reaction-diffusion scaling limit (abundant species only)

In this section, we further explore the properties of the limiting process obtained in Theorem 2.11 in the particular case where all species are abundant, that is $\mathcal{T}_{L,s} = \emptyset$. We also suppose that the initial value M_0^∞ is deterministic and equal to some $M_0 \in \mathcal{M}_{\mathcal{T}_L}$. In particular, we give a sufficient condition for the assumption on uniqueness of the solution to (2.29) to be satisfied (see Lemma 4.2), then we give some conditions under which the measures M_t^∞ admits a density (see Proposition 4.4 for a more precise statement). The results obtained in this section and their proofs are largely inspired by the results and methods developed in [13], in which only linear birth of particles and pairwise interactions are considered (but these "reactions" are allowed to have spatially inhomogeneous rates, as in our framework).

Remark 4.1. Notice that, in what follows, we allow the presence of localized species, but assume that their abundances are of order $\mathcal{O}(N)$ before taking the limit.

Since $\mathcal{T}_{L,s} = \emptyset$, we obviously have $\mathrm{Card}(R^\ell) = 0$ and so by (3.29) and Doob's maximal inequality, for every $f \in \mathcal{C}^{2,\perp}(\mathcal{P})$, $T \geq 0$, $\eta > 0$ and every $N \geq 1$ we have

$$\mathbf{P}\left(\sup_{0\leq t\leq T} \left|Z_t^N\right| \geq \eta\right) \leq \frac{1}{\eta^2}\mathbf{E}\big[(Z_T^N)^2\big] = \frac{1}{\eta^2}\mathbf{E}\Big[\big\langle Z^N\big\rangle_T\Big] \leq \frac{C_T' T}{\eta^2 N},$$

where $Z^N = Z^N(f)$ is the martingale defined in Lemma 3.2. This probability goes to 0 as $N \to \infty$, and so the limit M^∞ of $(M^N)_{N\geq 1}$ is the deterministic solution to : $M_0^\infty = M_0$ and for every $f \in \mathcal{C}^{2,\perp}(\mathcal{P})$,

$$\langle M_t^\infty, f\rangle = \langle M_0^\infty, f\rangle + \int_0^t ds \left\{\sum_{r\in R} G_r^\infty \mathrm{Id}_f(M_s^\infty) + \mathcal{D}^\infty \mathrm{Id}_f(M_s^\infty)\right\}, \qquad \forall t \geq 0. \tag{4.1}$$

As in the definition of the pre-limiting stochastic model (see Remark 2.7), the assumption on uniqueness of the solution M^∞ has to be checked case by case, as it may hold for very different reasons. Note however the following result.

Lemma 4.2. *Suppose that the conditions on the functions $\Psi_{r,\bar{y}}$ stated in Assumption (A1) are satisfied, together with Assumptions (A2) and (B1). Suppose also that for any solution $(m_t)_{t\geq 0}$ to the set of equations* (4.1) *and any $T > 0$, we have*

$$m_T^* := \sup_{t\in[0,T]} \langle m_t, 1\rangle < \infty. \tag{4.2}$$

Then for any $m \in \mathcal{M}_{\mathcal{T}_L}$, there is at most one solution to (4.1) *with initial condition m.*

Proof. (Proof of Lemma 4.2.) For each $x \in \mathcal{T}_{NL}$, let $(P_x^t,\, t \geq 0)$ denote the semigroup on $\mathcal{C}(E)$ of the diffusion process with drift coefficient b_x and dispersion matrix Σ_x, normally reflected at the boundary of E. By extension, for $x \in \mathcal{T}_L$ we also write $(P_x^t,\, t \geq 0)$ for the semigroup of the trivial process following which particles do not move, *i.e.*, $P_x^t = \mathrm{Id}$ for every $t \geq 0$. When necessary, we shall abuse notation and write $P_x^t\varphi(x, y)$ for the function $P_x^t\varphi(x, \cdot)$ applied at $y \in E$. The

following is a mild formulation of the equations satisfied by a solution $(m_t)_{t\geq 0}$ to (4.1) (the proof follows exactly the same lines as that of Lemma 4.5 in [13]). For each $\varphi \in \mathcal{C}^{2,\perp}(\mathcal{P})$, we have

$$\begin{aligned}\langle m_t, \varphi\rangle - \langle m_0, P^t_\cdot \varphi\rangle \; &= \sum_{r\in R} \int_0^t ds \int_{E\times\mathcal{P}^{k_r}} \varrho_r(d\bar{y}) m_s(dp_1)\cdots m_s(dp_{k_r})\, \widetilde{h}_r\big(\bar{y}, \langle m_s, \Psi_{r,\bar{y}}\rangle\big) \\ &\quad \times \Big(\prod_{i=1}^{k_r} \mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\Big)\Big(\sum_{i=1}^{k_r'} P^{t-s}_{B_i^r}\varphi(B_i^r,\bar{y}) - \sum_{i=1}^{k_r} P^{t-s}_{x_i}\varphi(x_i,y_i)\Big),\end{aligned} \tag{4.3}$$

for every $t \geq 0$. By Remark 2.4 and the dominated convergence theorem, the equality (4.3) also holds for any $\varphi \in \mathcal{C}(\mathcal{P})$.

Let m', m'' be two solutions to (4.1), and let us consider $\varphi \in \mathcal{C}(\mathcal{P})$ such that $\|\varphi\|_\infty \leq 1$. Then

$$\begin{aligned}&\big|\langle m_t' - m_t'', \varphi\rangle\big| - |\langle m_0' - m_0'', P^t_\cdot\varphi\rangle| \\ &\leq \sum_{r\in R}\int_0^t ds \Bigg| \int_{E\times\mathcal{P}^{k_r}} \varrho_r(d\bar{y})\Big(\prod_{i=1}^{k_r}\mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i-\bar{y})\Big)\Big(\sum_{i=1}^{k_r'} P^{t-s}_{B_i^r}\varphi(B_i^r,\bar{y}) - \sum_{i=1}^{k_r} P^{t-s}_{x_i}\varphi(x_i,y_i)\Big) \\ &\qquad \times \Big(\widetilde{h}_r\big(\bar{y},\langle m_s', \Psi_{r,\bar{y}}\rangle\big) m_s'(dp_1)\cdots m_s'(dp_{k_r}) - \widetilde{h}_r\big(\bar{y},\langle m_s'', \Psi_{r,\bar{y}}\rangle\big) m_s''(dp_1)\cdots m_s''(dp_{k_r})\Big)\Bigg|.\end{aligned} \tag{4.4}$$

Using the facts that $\|\varphi\|_\infty \leq 1$ and that the semigroups $(P^t_x)_{t\geq 0}$ preserve the supremum norm, we obtain that the r.h.s. of (4.4) is bounded by

$$\begin{aligned}\sum_{r\in R}\|\Gamma_\epsilon\|_\infty^{k_r}(k_r'+k_r)\int_0^t ds \Bigg| \int_{E\times\mathcal{P}^{k_r}} \varrho_r(d\bar{y})\Big(\prod_{i=1}^{k_r}\psi_i^r(x_i,y_i,\bar{y})\Big)\Big(\widetilde{h}_r\big(\bar{y},\langle m_s', \Psi_{r,\bar{y}}\rangle\big) {m_s'}^{\otimes k_r} \\ - \widetilde{h}_r\big(\bar{y},\langle m_s'', \Psi_{r,\bar{y}}\rangle\big) {m_s''}^{\otimes k_r}\Big)(dp_1,\ldots,dp_{k_r})\Bigg|,\end{aligned} \tag{4.5}$$

where we have defined for every $r \in R$, $i \in \{1,\ldots,k_r\}$, and $x_i, y_i, \bar{y} \in E$,

$$\psi_i^r(x_i,y_i,\bar{y}) := \mathbf{1}_{A_i^r}(x_i)\,\frac{\Gamma_\epsilon(y_i-\bar{y})}{\|\Gamma_\epsilon\|_\infty}.$$

Since $\|\psi_i^r\|_\infty \le 1$, if $k_r \ge 1$ we can write that for every $\bar{y} \in E$ and every $s \ge 0$,

$$\begin{aligned}
&\Big| \int_{\mathcal{P}^{k_r}} \Big(\prod_{i=1}^{k_r} \psi_i^r(x_i, y_i, \bar{y}) \Big) m_s'(dp_1) \cdots m_s'(dp_{k_r}) - \int_{\mathcal{P}^{k_r}} \Big(\prod_{i=1}^{k_r} \psi_i^r(x_i, y_i, \bar{y}) \Big) m_s''(dp_1) \cdots m_s''(dp_{k_r}) \Big| \\
&= \Big| \prod_{i=1}^{k_r} \langle m_s', \psi_i^r(\cdot,\cdot,\bar{y}) \rangle - \prod_{i=1}^{k_r} \langle m_s'', \psi_i^r(\cdot,\cdot,\bar{y}) \rangle \Big| \\
&= \Big| \big(\langle m_s', \psi_1^r(\cdot,\cdot,\bar{y}) \rangle - \langle m_s'', \psi_1^r(\cdot,\cdot,\bar{y}) \rangle \big) \prod_{i=2}^{k_r} \langle m_s', \psi_i^r(\cdot,\cdot,\bar{y}) \rangle \\
&\qquad + \langle m_s'', \psi_1^r(\cdot,\cdot,\bar{y}) \rangle \Big(\prod_{i=2}^{k_r} \langle m_s', \psi_i^r(\cdot,\cdot,\bar{y}) \rangle - \prod_{i=2}^{k_r} \langle m_s'', \psi_i^r(\cdot,\cdot,\bar{y}) \rangle \Big) \Big| \\
&\le \langle m_s', 1 \rangle^{k_r - 1} \big| \langle m_s', \psi_1^r(\cdot,\cdot,\bar{y}) \rangle - \langle m_s'', \psi_1^r(\cdot,\cdot,\bar{y}) \rangle \big| \\
&\qquad + \langle m_s'', 1 \rangle \Big| \prod_{i=2}^{k_r} \langle m_s', \psi_i^r(\cdot,\cdot,\bar{y}) \rangle - \prod_{i=2}^{k_r} \langle m_s'', \psi_i^r(\cdot,\cdot,\bar{y}) \rangle \Big|.
\end{aligned}$$

By an easy recursion, we thus obtain that

$$\int_{\mathcal{P}^{k_r}} \Big(\prod_{i=1}^{k_r} \psi_i^r(x_i, y_i, \bar{y}) \Big) \Big({m_s'}^{\otimes k_r} - {m_s''}^{\otimes k_r} \Big)(d\mathbf{p}) \le k_r \langle m_s' + m_s'', 1 \rangle^{k_r - 1} \sup_{\psi : \|\psi\|_\infty \le 1} \big| \langle m_s' - m_s'', \psi \rangle \big|,$$

where the supremum is taken over all $\psi \in \mathcal{C}(\mathcal{P})$ such that $\|\psi\|_\infty \le 1$. In addition, by Assumption (B1) we have for $r \in R$

$$\begin{aligned}
\int_E \varrho_r(d\bar{y}) \big| \widetilde{h}_r\big(\bar{y}, \langle m_s', \Psi_{r,\bar{y}} \rangle\big) - \widetilde{h}_r\big(\bar{y}, \langle m_s'', \Psi_{r,\bar{y}} \rangle\big) \big| &\le L_r \int_E \varrho_r(d\bar{y}) \big| \langle m_s' - m_s'', \Psi_{r,\bar{y}} \rangle \big| \\
&\le L_r \|\Psi_r\|_\infty (\mathrm{Vol}(E) + 1) \sup_{\psi : \|\psi\|_\infty \le 1} \big| \langle m_s' - m_s'', \psi \rangle \big|.
\end{aligned}$$

Combining the above and using the assumption on the total mass of $(m_t')_{t \ge 0}$ and $(m_t'')_{t \ge 0}$ stated in Lemma 4.2, we obtain that for every time horizon $T > 0$ and every $s \in [0, T]$,

$$\begin{aligned}
&\Big| \int_E \varrho_r(d\bar{y}) \int_{\mathcal{P}^{k_r}} \Big(\prod_{i=1}^{k_r} \psi_i^r(x_i, y_i, \bar{y}) \Big) \Big(\widetilde{h}_r(\bar{y}, \langle m_s', \Psi_{r,\bar{y}} \rangle) {m_s'}^{\otimes k_r} - \widetilde{h}_r(\bar{y}, \langle m_s'', \Psi_{r,\bar{y}} \rangle) {m_s''}^{\otimes k_r} \Big)(d\mathbf{p}) \Big| \\
&\le \Big| \int_E \varrho_r(d\bar{y}) \int_{\mathcal{P}^{k_r}} \Big(\prod_{i=1}^{k_r} \psi_i^r(x_i, y_i, \bar{y}) \Big) \Big(\widetilde{h}_r(\bar{y}, \langle m_s', \Psi_{r,\bar{y}} \rangle) - \widetilde{h}_r(\bar{y}, \langle m_s'', \Psi_{r,\bar{y}} \rangle) \Big) {m_s'}^{\otimes k_r}(d\mathbf{p}) \Big| \\
&\qquad + \Big| \int_E \varrho_r(d\bar{y}) \, \widetilde{h}_r(\bar{y}, \langle m_s'', \Psi_{r,\bar{y}} \rangle) \int_{\mathcal{P}^{k_r}} \Big(\prod_{i=1}^{k_r} \psi_i^r(x_i, y_i, \bar{y}) \Big) \big({m_s'}^{\otimes k_r} - {m_s''}^{\otimes k_r} \big)(d\mathbf{p}) \Big| \\
&\le L_r \|\Psi_r\|_\infty (\mathrm{Vol}(E) + 1) \langle m_s', 1 \rangle^{k_r} \sup_{\psi : \|\psi\|_\infty \le 1} \big| \langle m_s' - m_s'', \psi \rangle \big| \\
&\qquad + \mathcal{S}_r k_r \langle m_s' + m_s'', 1 \rangle^{k_r - 1} \sup_{\psi : \|\psi\|_\infty \le 1} \big| \langle m_s' - m_s'', \psi \rangle \big| \\
&\le \Big[L_r \|\Psi_r\|_\infty (\mathrm{Vol}(E) + 1) \big((m')_T^* + (m'')_T^* \big) + \mathcal{S}_r k_r \Big] \\
&\qquad \times \big((m')_T^* + (m'')_T^* \big)^{k_r - 1} \sup_{\psi : \|\psi\|_\infty \le 1} \big| \langle m_s' - m_s'', \psi \rangle \big|,
\end{aligned} \tag{4.6}$$

where $\mathcal{S}_r$ was defined in (2.26). Observe that for any reaction r with $k_r = 0$, the rate of reaction is independent of mass so its contribution to $\langle m'_t - m''_t, \varphi\rangle$ is zero. This implies that for every $T > 0$, there exists a constant $C > 0$ such that for every $\varphi \in \mathcal{C}(\mathcal{P})$ with $\|\varphi\|_\infty \leq 1$, and every $t \leq T$, we have

$$\left|\langle m'_t - m''_t, \varphi\rangle\right| \leq \left|\langle m'_0 - m''_0, P^t_\cdot \varphi\rangle\right| + C\int_0^t ds \sup_{\psi:\|\psi\|_\infty\leq 1} \left|\langle m'_s - m''_s, \psi\rangle\right|,$$

and Gronwall's inequality yields for all $t \in [0, T]$

$$\sup_{\varphi:\|\varphi\|_\infty\leq 1} \left|\langle m'_t - m''_t, \varphi\rangle\right| \leq \sup_{\varphi:\|\varphi\|_\infty\leq 1} \left|\langle m'_0 - m''_0, P^t_\cdot \varphi\rangle\right| e^{Ct}. \tag{4.7}$$

Hence, whenever $m'_0 = m''_0$ we have $m'_t = m''_t$ for all $t \leq T$ and uniqueness holds. □

For our last result, Proposition 4.4, which gives conditions under which at any time the spatial distributions of non-localized species have a density with respect to Lebesgue measure on E while the localized species have an evolving mass at the locations where they sit, we restrict our attention to a particular case which is already rich (and notationally heavy): for every $x \in \mathcal{T}_{NL}$, we suppose that the diffusion matrix Σ^2_x is of the form $\sigma^2_x \mathrm{Id}$, where $\sigma^2_x : E \mapsto \mathbf{R}_+$ is Lipschitz.

We shall use the following set of functions $\mathbf{D}$ as the set of possible densities. Recall the notation ℓ_E for Lebesgue measure on E.

Definition 4.3. We say that a function $\varphi : \mathcal{P} \to \mathbf{R}$ is in $\mathbb{L}^1(\mathcal{P})$ if

$$\|\varphi\|_1 := \sum_{x\in\mathcal{T}_{NL}} \int_E \ell_E(dy)|\varphi(x, y)| + \sum_{x\in\mathcal{T}_L} |\varphi(x, \bar{y}_x)| < \infty. \tag{4.8}$$

We call $\mathbf{D}$ the set of nonnegative functions φ such that $\varphi \in \mathbb{L}^1(\mathcal{P})$ and $\varphi(x, y) = 0$ whenever $x \in \mathcal{T}_L$ and $y \neq \bar{y}_x$.

The additional constraints on the drift and variance coefficients for the diffusion of non-localized species are summarised in the following **Assumption**:

(B3) There exists $\sigma^2_* > 0$ such that $\sigma^2_x(y) \geq \sigma^2_*$ for all $(x, y) \in \mathcal{T}_{NL} \times E$. Furthermore, for every $x \in \mathcal{T}_{NL}$, σ^2_x is of class C^2 on E and its second derivatives with respect to the d spatial coordinates $y_1, \ldots, y_d$ are α-Hölderian for some $\alpha > 0$.

In addition, for every $x \in \mathcal{T}_{NL}$, the drift coefficient b_x is of class C^1 and its derivatives with respect to the d spatial coordinates are α-Hölderian for some $\alpha > 0$.

Assumption (B3) is analogous to Assumption (H2) in [13], where it is used to guarantee the existence of a spatial density for the semigroups corresponding to the motions of particles (which are all assumed to diffuse in [13]). We shall also need a last assumption, that can only be stated in a rigourous way once we have introduced the appropriate sequence $\{\mu^n, n \geq 0\}$ of approximations to the density. This assumption is used in the proof of Proposition 4.4 to have a uniform control on the total mass of the approximate densities over any compact time interval.

Proposition 4.4. *Suppose that the regularity and boundedness assumptions on each $\widetilde{h}_r$ stated in Assumption (B1) are satisfied, together with Assumption (B3) and Assumption (B4) (see the*

proof of the Proposition for a statement). Let M^∞ be the solution to (4.1)*, and suppose that there exists $\mu_0^\infty \in \mathbf{D}$ such that*

$$M_0^\infty = \sum_{x\in\mathcal{T}_{NL}} \mu_0^\infty(x,y)\delta_x \otimes \ell_E + \sum_{x\in\mathcal{T}_L} \mu_0^\infty(x,\bar{y}_x)\delta_x \otimes \delta_{\bar{y}_x}.$$

That is, for every $f \in \mathcal{C}(\mathcal{P})$,

$$\langle M_0^\infty, f\rangle = \sum_{x\in\mathcal{T}_{NL}} \int_E \ell_E(dy)\, \mu_0^\infty(x,y) f(x,y) + \sum_{x\in\mathcal{T}_L} \mu_0^\infty(x,\bar{y}_x) f(x,\bar{y}_x).$$

Then for every $t \geq 0$, there exists a function $\mu_t \in \mathbf{D}$ such that

$$M_t^\infty = \sum_{x\in\mathcal{T}_{NL}} \mu_t(x,y)\delta_x \otimes \ell_E + \sum_{x\in\mathcal{T}_L} \mu_t(x,\bar{y}_x)\delta_x \otimes \delta_{\bar{y}_x}. \tag{4.9}$$

Moreover, for every $T > 0$

$$\sup_{t\in[0,T]} \|\mu_t\|_1 < \infty \tag{4.10}$$

and $(\mu_t)_{t\geq 0}$ is a weak solution to the following system of integro-differential and partial integro-differential equations: for every $x \in \mathcal{T}_{NL}$, and for every $y \in E$

$$\begin{aligned}
\partial_t \mu_t(x,y) &= \Delta_y\big(\sigma_x^2(y)\mu_t(x,y)\big) - \nabla_y \cdot \big(b_x(y)\mu_t(x,y)\big) \\
&\quad - \sum_{r\in R}\sum_{i=1}^{k_r} \mathbf{1}_{A_i^r}(x)\mu_t(x,y) \int_E \varrho_r(d\bar{y})\Bigg(\widetilde{h}_r(\bar{y},\mu_t)\Gamma_\epsilon(y-\bar{y}) \prod_{\substack{j=1\\ j\neq i,}}^{k_r} \Big\{\mathbf{1}_{\{A_j^r\in\mathcal{T}_L\}}\Gamma_\epsilon(\bar{y}_{A_j^r}-\bar{y})\mu_t(A_j^r,\bar{y}_{A_j^r}) \\
&\qquad + \mathbf{1}_{\{A_j^r\in\mathcal{T}_{NL}\}} \int_E \ell_E(dy_j)\Gamma_\epsilon(y_j-\bar{y})\mu_t(A_j^r,y_j)\Big\}\Bigg) \\
&\quad + \sum_{r\in R_{NL}} \Bigg(\sum_{i=1}^{k_r'} \mathbf{1}_{B_i^r}(x)\Bigg)\widetilde{h}_r(y,\mu_t) \prod_{j=1}^{k_r} \Big\{\mathbf{1}_{\{A_j^r\in\mathcal{T}_L\}}\Gamma_\epsilon(\bar{y}_{A_j^r}-y)\mu_t(A_j^r,\bar{y}_{A_j^r}) \\
&\qquad + \mathbf{1}_{\{A_j^r\in\mathcal{T}_{NL}\}} \int_E \ell_E(dy_j)\Gamma_\epsilon(y_j-y)\mu_t(A_j^r,y_j)\Big\} \\
&\quad + \sum_{r\in R_L} \mathbf{1}_{\{y=\bar{y}_r\}}\Bigg(\sum_{i=1}^{k_r'} \mathbf{1}_{B_i^r}(x)\Bigg)\widetilde{h}_r(\bar{y}_r,\mu_t) \prod_{j=1}^{k_r} \Big\{\mathbf{1}_{\{A_j^r\in\mathcal{T}_L\}}\Gamma_\epsilon(\bar{y}_{A_j^r}-\bar{y}_r)\mu_t(A_j^r,\bar{y}_{A_j^r}) \\
&\qquad + \mathbf{1}_{\{A_j^r\in\mathcal{T}_{NL}\}} \int_E \ell_E(dy_j)\Gamma_\epsilon(y_j-\bar{y}_r)\mu_t(A_j^r,y_j)\Big\}\ ; \\
\mu_0(x,\cdot) &= \mu_0^\infty(x,\cdot);\ \ \nabla_y\mu_t(x,y')\cdot n(y') = 0 \textit{ for all } t>0 \textit{ and } y' \in \partial E\ ;
\end{aligned} \tag{4.11}$$

and for every $x \in \mathcal{T}_L$,

$$
\begin{aligned}
&\partial_t \mu_t(x,\bar{y}_x) \\
&= -\sum_{r\in R}\sum_{i=1}^{k_r} \mathbf{1}_{A_i^r}(x)\mu_t(x,\bar{y}_x)\int_E \varrho_r(d\bar{y})\Bigg(\widetilde{h}_r(\bar{y},\mu_t)\Gamma_\epsilon(\bar{y}_x-\bar{y})\prod_{\substack{j=1\\ j\neq i}}^{k_r}\Bigg\{\mathbf{1}_{\{A_j^r\in\mathcal{T}_L\}}\Gamma_\epsilon(\bar{y}_{A_j^r}-\bar{y})\mu_t(A_j^r,\bar{y}_{A_j^r}) \\
&\qquad\qquad + \mathbf{1}_{\{A_j^r\in\mathcal{T}_{NL}\}}\int_E \ell_E(dy_j)\Gamma_\epsilon(y_j-\bar{y})\mu_t(A_j^r,y_j)\Bigg\}\Bigg) \\
&\quad + \sum_{r\in R_{NL}}\Bigg(\sum_{i=1}^{k_r'}\mathbf{1}_{B_i^r}(x)\Bigg)\widetilde{h}_r(\bar{y}_x,\mu_t)\prod_{j=1}^{k_r}\Bigg\{\mathbf{1}_{\{A_j^r\in\mathcal{T}_L\}}\Gamma_\epsilon(\bar{y}_{A_j^r}-\bar{y}_x)\mu_t(A_j^r,\bar{y}_{A_j^r}) \\
&\qquad\qquad + \mathbf{1}_{\{A_j^r\in\mathcal{T}_{NL}\}}\int_E \ell_E(dy_j)\Gamma_\epsilon(y_j-\bar{y}_x)\mu_t(A_j^r,y_j)\Bigg\} \\
&\quad + \sum_{r\in R_L}\mathbf{1}_{\{\bar{y}_x=\bar{y}_r\}}\Bigg(\sum_{i=1}^{k_r'}\mathbf{1}_{B_i^r}(x)\Bigg)\widetilde{h}_r(\bar{y}_r,\mu_t)\prod_{j=1}^{k_r}\Bigg\{\mathbf{1}_{\{A_j^r\in\mathcal{T}_L\}}\Gamma_\epsilon(\bar{y}_{A_j^r}-\bar{y}_x)\mu_t(A_j^r,\bar{y}_{A_j^r}) \\
&\qquad\qquad + \mathbf{1}_{\{A_j^r\in\mathcal{T}_{NL}\}}\int_E \ell_E(dy_j)\Gamma_\epsilon(y_j-\bar{y}_x)\mu_t(A_j^r,y_j)\Bigg\}\ ; \\
&\mu_0(x,\bar{y}_x) = \mu_0^\infty(x,\bar{y}_x).
\end{aligned}
\tag{4.12}
$$

In the above, we have abused notation and written $\widetilde{h}_r(\cdot,\mu_t)$ *to mean* $\widetilde{h}_r(\cdot,\langle M_t,\Psi_{r,\cdot}\rangle)$, *where* M_t *is the measure built out of the function* μ_t.

Proof. (Proof of Proposition 4.4.) We follow the lines of the proof of Theorem 4.6 in [13] closely, mainly using the same arguments as in our proof of Lemma 4.2 (which therefore we do not repeat entirely). In all that follows, we shall suppose that for each n, $\mu_t^n(x,\cdot)$ is defined for all $x\in\mathcal{T}$ and $t\geq 0$ with the understanding that $\mu_t^n(x,y)=\mu_t^n(x,\bar{y}_x)\mathbf{1}_{\bar{y}_x}(y)$ when $x\in\mathcal{T}_L$ (that is, $\mu^n\in\mathbf{D}$). To ease the notation, as in (2.7) we also define

$$
\varrho^x(d\bar{y}) = \ell_E(d\bar{y}) \quad \text{if } x\in\mathcal{T}_{NL} \qquad \text{and} \qquad \varrho^x(d\bar{y}) = \delta_{\bar{y}_x}(d\bar{y}) \quad \text{if } x\in\mathcal{T}_L. \tag{4.13}
$$

Consider the following collection of inductively constructed functions $\{\mu^n, n\geq 0\}$. First, we set $\mu_t^0=\mu_0^\infty$ for all $t\geq 0$. Suppose that for some $n\geq 0$, $(\mu_t^n)_{t\geq 0}$ is well-defined and takes its values in $\mathbf{D}$. For every $x\in\mathcal{T}_{NL}$, define $(\mu_t^{n+1}(x,\cdot))_{t\geq 0}$ as the weak solution to the partial differential equation

$$
\begin{aligned}
&\partial_t\mu_t^{n+1}(x,y) = \Delta_y\big(\sigma_x^2(y)\mu_t^{n+1}(x,y)\big) - \nabla_y\cdot\big(b_x(y)\mu_t^{n+1}(x,y)\big) \\
&\quad - \sum_{r\in R}\sum_{i=1}^{k_r}\mathbf{1}_{A_i^r}(x)\mu_t^{n+1}(x,y)\int_E\varrho_r(d\bar{y})\Bigg(\widetilde{h}_r(\bar{y},\mu_t^n)\Gamma_\epsilon(y-\bar{y})\prod_{\substack{j=1\\ j\neq i}}^{k_r}\Bigg\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y})\mu_t^n(A_j^r,y_j)\Bigg\}\Bigg) \\
&\quad + \sum_{r\in R_{NL}}\Bigg(\sum_{i=1}^{k_r'}\mathbf{1}_{B_i^r}(x)\Bigg)\widetilde{h}_r(y,\mu_t^n)\prod_{j=1}^{k_r}\Bigg\{\int_E\varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-y)\mu_t^n(A_j^r,y_j)\Bigg\} \\
&\quad + \sum_{r\in R_L}\mathbf{1}_{\{y=\bar{y}_r\}}\Bigg(\sum_{i=1}^{k_r'}\mathbf{1}_{B_i^r}(x)\Bigg)\widetilde{h}_r(\bar{y}_r,\mu_t^n)\prod_{j=1}^{k_r}\Bigg\{\int_E\varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y}_r)\mu_t^n(A_j^r,y_j)\Bigg\}\ ; \\
&\mu_0^{n+1}(x,\cdot)=\mu_0^\infty(x,\cdot);\ \ \nabla_y\mu_t^{n+1}(x,y')\cdot n(y')=0 \text{ for all } t>0 \text{ and } y'\in\partial E.
\end{aligned}
\tag{4.14}
$$

For every $x \in \mathcal{T}_L$, define $(\mu_t^{n+1}(x,\bar{y}_x))_{t\geq 0}$ as the solution to the ordinary differential equation:

$$\begin{aligned}
&\partial_t \mu_t^{n+1}(x,\bar{y}_x) \\
&= -\sum_{r\in R}\sum_{i=1}^{k_r} \mathbf{1}_{A_i^r}(x)\mu_t^{n+1}(x,\bar{y}_x)\int_E \varrho_r(d\bar{y})\Big(\widetilde{h}_r(\bar{y},\mu_t^n)\Gamma_\epsilon(\bar{y}_x-\bar{y})\prod_{\substack{j=1\\ j\neq i}}^{k_r}\Big\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y})\mu_t^n(A_j^r,y_j)\Big\}\Big) \\
&\quad + \sum_{r\in R_{NL}}\Big(\sum_{i=1}^{k_r'}\mathbf{1}_{B_i^r}(x)\Big)\widetilde{h}_r(\bar{y}_x,\mu_t^n)\prod_{j=1}^{k_r}\Big\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y}_x)\mu_t^n(A_j^r,y_j)\Big\} \\
&\quad + \sum_{r\in R_L}\mathbf{1}_{\{\bar{y}_x=\bar{y}_r\}}\Big(\sum_{i=1}^{k_r'}\mathbf{1}_{B_i^r}(x)\Big)\widetilde{h}_r(\bar{y}_r,\mu_t^n)\prod_{j=1}^{k_r}\Big\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y}_r)\mu_t^n(A_j^r,y_j)\Big\}\ ; \\
&\mu_0^{n+1}(x,\bar{y}_x) = \mu_0^\infty(x,\bar{y}_x). \qquad (4.15)
\end{aligned}$$

Existence of these functions follows from standard results on linear parabolic equations, see *e.g.* Theorem 7.3 in [21], since the lower bound on σ^2 required in Assumption (B3) yields the uniform ellipticity of the diffusion operator for the non-localized species. Nonnegativity of each $\mu^{n+1}(x,\cdot)$ follows from the nonnegativity of μ_0^∞, μ^n, Γ_ϵ and all $\widetilde{h}_r$, from the rate of removal of mass being proportional to mass itself and from standard maximal inequality arguments. Recall that for definiteness, we also set $\mu_t^{n+1}(x,y)=0$ for all $x\in\mathcal{T}_L$ and $y\neq\bar{y}_x$.

Now that we have defined the sequence μ^n, let us state the mysterious **Assumption** appearing in Proposition 4.4. Recall the definition of $\|\varphi\|_1$ given in (4.8).

(B4) For every $T>0$, there exists $C_T<\infty$ such that

$$\sup_{n\geq 0}\sup_{0\leq t\leq T}\|\mu_t^n\|_1 \leq C_T.$$

Just like Assumptions (A3) and (B2), because of the generality of our formulation, we believe that this condition can only be checked case by case.

Let us first focus on (4.14). By construction, for every $\varphi_x\in\mathcal{C}^2(E)$ with $\nabla_y\varphi_x(y)\cdot n(y)=0$ for all $y\in\partial E$, we have

$$\begin{aligned}
&\int_E \ell_E(dy)\varphi_x(y)\mu_t^{n+1}(x,y) \qquad (4.16)\\
&= \int_E \ell_E(dy)\varphi_x(y)\mu_0^\infty(x,y) + \int_0^t\int_E ds\,\ell_E(dy)\big(\sigma_x^2(y)\Delta_y\varphi_x(y) - b_x(y)\cdot\nabla_y\varphi_x(y)\big)\mu_s^{n+1}(x,y) \\
&\quad - \int_0^t ds\sum_{r\in R}\sum_{i=1}^{k_r}\mathbf{1}_{A_i^r}(x)\int_E \ell_E(dy)\varphi_x(y)\mu_s^{n+1}(x,y) \\
&\qquad\qquad \times\int_E \varrho_r(d\bar{y})\widetilde{h}_r(\bar{y},\mu_s^n)\Gamma_\epsilon(y-\bar{y})\prod_{\substack{j=1\\ j\neq i}}^{k_r}\Big\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y})\mu_s^n(A_j^r,y_j)\Big\} \\
&\quad + \int_0^t ds\sum_{r\in R_{NL}}\Big(\sum_{i=1}^{k_r'}\mathbf{1}_{B_i^r}(x)\Big)\int_E \ell_E(dy)\varphi_x(y)\widetilde{h}_r(y,\mu_s^n)\prod_{j=1}^{k_r}\Big\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-y)\mu_s^n(A_j^r,y_j)\Big\} \\
&\quad + \int_0^t ds\sum_{r\in R_L}\Big(\sum_{i=1}^{k_r'}\mathbf{1}_{B_i^r}(x)\Big)\varphi_x(\bar{y}_r)\widetilde{h}_r(\bar{y}_r,\mu_s^n)\prod_{j=1}^{k_r}\Big\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y}_r)\mu_s^n(A_j^r,y_j)\Big\}.
\end{aligned}$$

As in the proof of Lemma 4.2, the above equation has a mild form which can be written using the density $p_x^t(\cdot,\cdot)$ of $P_x^t\varphi_x(y) = \int_E \ell_E(dy')p_x^t(y,y')\varphi_x(y')$. Indeed, the regularity of the diffusion and drift coefficients stated in Assumption (B3), together with boundedness of E and smoothness of its boundary ∂E imply that a unique such density $p_x^t(\cdot,\cdot)$ exists which is continuous in t, y, y' for all $x \in \mathcal{T}_{NL}$ (see Lemma 4.5 in [13], and note that the original Sato-Ueno result allows ∂E to have finitely many piecewise smooth components). Hence, for all $x \in \mathcal{T}_{NL}$ and each continuous function φ_x on E, we have

$$\begin{aligned}
\int_E \ell_E(dy)\,\varphi_x(y)\mu_t^{n+1}(x,y) &= \int_E \ell_E(dy)\int_E \ell_E(dy')p_x^t(y,y')\varphi_x(y')\mu_0^\infty(x,y)\\
&- \int_0^t ds\sum_{r\in R}\sum_{i=1}^{k_r}\mathbf{1}_{A_i^r}(x)\int_E \ell_E(dy)\int_E \ell_E(dy')p_x^{t-s}(y,y')\varphi_x(y')\mu_s^{n+1}(x,y)\\
&\qquad\times\int_E \varrho_r(d\bar{y})\widetilde{h}_r(\bar{y},\mu_s^n)\Gamma_\epsilon(y-\bar{y})\prod_{\substack{j=1\\ j\neq i}}^{k_r}\Big\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y})\mu_s^n(A_j^r,y_j)\Big\}\\
&+ \int_0^t ds\sum_{r\in R_{NL}}\Big(\sum_{i=1}^{k_r'}\mathbf{1}_{B_i^r}(x)\Big)\int_E \ell_E(dy)\int_E \ell_E(dy')p_x^{t-s}(y,y')\varphi_x(y')\widetilde{h}_r(y,\mu_s^n)\\
&\qquad\times\prod_{j=1}^{k_r}\Big\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-y)\mu_s^n(A_j^r,y_j)\Big\}\\
&+ \int_0^t ds\sum_{r\in R_L}\Big(\sum_{i=1}^{k_r'}\mathbf{1}_{B_i^r}(x)\Big)\int_E \ell_E(dy')p_x^{t-s}(\bar{y}_r,y')\varphi_x(y')\widetilde{h}_r(\bar{y}_r,\mu_s^n)\\
&\qquad\times\prod_{j=1}^{k_r}\Big\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y}_r)\mu_s^n(A_j^r,y_j)\Big\}.
\end{aligned}$$

Fubini's theorem then implies that for all $y' \in E$,

$$\begin{aligned}
\mu_t^{n+1}(x,y') &= \int_E \ell_E(dy)p_x^t(y,y')\mu_0^\infty(x,y)dy \qquad (4.17)\\
&- \int_0^t ds\sum_{r\in R}\sum_{i=1}^{k_r}\mathbf{1}_{A_i^r}(x)\int_E \ell_E(dy)p_x^{t-s}(y,y')\mu_s^{n+1}(x,y)\int_E \varrho_r(d\bar{y})\widetilde{h}_r(\bar{y},\mu_s^n)\Gamma_\epsilon(y-\bar{y})\\
&\qquad\times\prod_{\substack{j=1\\ j\neq i}}^{k_r}\Big\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y})\mu_s^n(A_j^r,y_j)\Big\}\\
&+ \int_0^t ds\sum_{r\in R_{NL}}\Big(\sum_{i=1}^{k_r'}\mathbf{1}_{B_i^r}(x)\Big)\int_E \ell_E(dy)p_x^{t-s}(y,y')\widetilde{h}_r(y,\mu_s^n)\prod_{j=1}^{k_r}\Big\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-y)\mu_s^n(A_j^r,y_j)\Big\}\\
&+ \int_0^t ds\sum_{r\in R_L}\Big(\sum_{i=1}^{k_r'}\mathbf{1}_{B_i^r}(x)\Big)p_x^{t-s}(\bar{y}_r,y')\widetilde{h}_r(\bar{y}_r,\mu_s^n)\prod_{j=1}^{k_r}\Big\{\int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y}_r)\mu_s^n(A_j^r,y_j)\Big\}.
\end{aligned}$$

Integrating (4.15) with respect to time, for every $x \in \mathcal{T}_L$ we directly obtain an analogue of (4.17) where the measure $\ell_E(dy)\,p_x^{t-s}(y,y')$ is replaced by a Dirac mass at $\bar{y}_x$. Let us now show that the sequence $\{\mu^n,\, n \geq 0\}$ converges as $n \to \infty$ and that the limit satisfies (4.11) and (4.12).

Let $T > 0$ and recall from Assumption (B4) that we assume that we can prove the existence of $C_T < \infty$ such that

$$\sup_{n\geq 0}\sup_{t\in[0,T]} \|\mu_t^n\|_1 \leq C_T.$$

For $x \in \mathcal{T}_{NL}$ and $y' \in E$, we can write

$$\begin{aligned}
&|\mu_t^{n+1}(x,y') - \mu_t^n(x,y')| \\
&= \Bigg| \int_0^t ds \sum_{r\in R}\sum_{i=1}^{k_r} \mathbf{1}_{A_i^r}(x) \int_E \ell_E(dy) p_x^{t-s}(y,y') \\
&\quad \Bigg(\mu_s^{n+1}(x,y) \int_E \varrho_r(d\bar{y}) \widetilde{h}_r(\bar{y},\mu_s^n)\Gamma_\epsilon(y-\bar{y}) \prod_{\substack{j=1\\ j\neq i}}^{k_r} \Bigg\{ \int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y})\mu_s^n(A_j^r,y_j)\Bigg\} \\
&\quad - \mu_s^n(x,y) \int_E \varrho_r(d\bar{y}) \widetilde{h}_r(\bar{y},\mu_s^{n-1})\Gamma_\epsilon(y-\bar{y}) \prod_{\substack{j=1\\ j\neq i}}^{k_r} \Bigg\{ \int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y})\mu_s^{n-1}(A_j^r,y_j)\Bigg\}\Bigg)\Bigg| \\
&+ \Bigg| \int_0^t ds \sum_{r\in R_{NL}} \Bigg(\sum_{i=1}^{k_r'} \mathbf{1}_{B_i^r}(x)\Bigg) \int_E \ell_E(dy) p_x^{t-s}(y,y') \Bigg(\widetilde{h}_r(y,\mu_s^n) \prod_{j=1}^{k_r} \Bigg\{ \int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-y)\mu_s^n(A_j^r,y_j)\Bigg\} \\
&\qquad - \widetilde{h}_r(y,\mu_s^{n-1}) \prod_{j=1}^{k_r} \Bigg\{ \int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-y)\mu_s^{n-1}(A_j^r,y_j)\Bigg\}\Bigg)\Bigg| \\
&+ \Bigg| \int_0^t ds \sum_{r\in R_L} \Bigg(\sum_{i=1}^{k_r'} \mathbf{1}_{B_i^r}(x)\Bigg) p_x^{t-s}(\bar{y}_r,y') \Bigg(\widetilde{h}_r(\bar{y}_r,\mu_s^n) \prod_{j=1}^{k_r} \Bigg\{ \int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y}_r)\mu_s^n(A_j^r,y_j)\Bigg\} \\
&\qquad - \widetilde{h}_r(\bar{y}_r,\mu_s^{n-1}) \prod_{j=1}^{k_r} \Bigg\{ \int_E \varrho^{A_j^r}(dy_j)\Gamma_\epsilon(y_j-\bar{y}_r)\mu_s^{n-1}(A_j^r,y_j)\Bigg\}\Bigg)\Bigg|.
\end{aligned}$$

Analogous expressions trivially hold for $|\mu_t^{n+1}(x,\bar{y}_x) - \mu_t^n(x,\bar{y}_x)|$ when $x \in \mathcal{T}_L$. Considering each molecular species (non-localized and localized) separately and, for a given species, each term in the above sums over R, R_{NL} and R_L one by one, we may then proceed exactly as in the proof of Lemma 4.2 and obtain the existence of a finite constant C_T', which depends on the finitely many parameters k_r, k_r', L_r, $\|\Psi_r\|_\infty$, $\mathcal{S}_r$, C_T and $\mathrm{Vol}(E)$ but not on n, and such that for all $t \in [0,T]$ and all $n \geq 0$,

$$\sup_{s\in[0,t]} \|\mu_s^{n+1} - \mu_s^n\|_1 \leq C_T' \int_0^t ds \Bigg(\sup_{u\in[0,s]} \|\mu_u^{n+1} - \mu_u^n\|_1 + \sup_{u\in[0,s]} \|\mu_u^n - \mu_u^{n-1}\|_1 \Bigg). \tag{4.18}$$

Using Gronwall's lemma, we find that there exists another constant $C_T'' < \infty$ such that for every $t \in [0,T]$ and all $n \geq 1$,

$$\sup_{s\in[0,t]} \|\mu_t^{n+1} - \mu_t^n\|_1 \leq C_T'' \int_0^t ds \sup_{u\in[0,s]} \|\mu_u^n - \mu_u^{n-1}\|_1.$$

Picard's iteration proof gives us that

$$\sum_{n\geq 0} \sup_{t\in[0,T]} \|\mu_t^{n+1} - \mu_t^n\|_1 < \infty,$$

which implies that the sequence $\{\mu_t^n,\, n \geq 0\}$ converges uniformly over $t \in [0,T]$ to a function μ_t on $\mathcal{P}$ which furthermore satisfies

$$\sup_{t\in[0,T]} \|\mu_t\|_1 \leq C_T.$$

In particular, each μ_t belongs to the set $\mathbf{D}$ of densities. Using the uniform convergence and passing to the limit in (4.16) (and in the analogous equations for localized species), we obtain that $(\mu_t)_{t\geq 0}$ satisfies the partial and ordinary integro-differential equations (4.14) and (4.15). Since by assumption (see Theorem 2.11) the measure-valued solution to (4.1) is unique, we can finally conclude that for every $t \geq 0$, μ_t is a density for M_t^∞ in the sense of (4.9), and Proposition 4.4 holds true. □

Remark 4.5. In the more general case where species of low abundances are present, the characterization of the limiting process in Theorem 2.11 as a measure-valued PDMP suggests that under conditions similar to the assumptions of Proposition 4.4, the continuous part $M_t^{\infty,c}$ may be written as

$$M_t^{\infty,c} = \sum_{x\in\mathcal{T}_{NL}} \mu_t(x,y)\delta_x \otimes \ell_E + \sum_{x\in\mathcal{T}_{L,b}} \mu_t(x,\bar{y}_x)\delta_x \otimes \delta_{\bar{y}_x}, \tag{4.19}$$

where $(\mu_t)_{t\geq 0}$ satisfies a system of partial integro-differential equations of the same form as (4.11–4.12), whose parameters may change in a stochastic way when the discrete part $(M_t^{\infty,d})_{t\geq 0}$ jumps (recall that the terms $\widetilde{h}_r(\bar{y}, \langle M_t^\infty, \Psi_{r,\bar{y}}\rangle)$ a priori depend on both the continuous and the discrete part of M_t^∞). We leave this exercise to the reader as it is notationally very heavy, but we expect that most of the proof of Proposition 4.4 can be reused in a very straightforward way to prove such a result (at least in particular cases), since the effect of the discrete component on the continuous one is only through $\tilde{h}_r$.

Acknowledgements

This research was supported by NSERC (Natural Sciences and Engineering Research Council of Canada) and the CRM (Centre de Recherches Mathématiques) -UMI travel allowance. AV was also supported in part by the *chaire Modélisation Mathématique et Biodiversité* of Veolia Environnement-École Polytechnique-Museum National d'Histoire Naturelle-Fondation X. The authors are grateful to the two reviewers for their particularly useful and constructive comments on previous versions of the manuscript, which helped to improve readability and to simplify some technical aspects of the construction of the process.

References

[1] Aldous, D. (1978). Stopping times and tightness. *Annals of Probability*, 6, 335–340.

[2] Ander, M., Beltrao, P., Di Ventura, B., Ferkinghoff-Borg, J., Foglierini, M.A.F.M., Lemerle, C., Tomas-Oliveira, I., Serrano, L. (2004). SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks. Systems biology, 1(1), 129–138.

[3] Andrews, S.S., Bray, D. (2004) Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Physical Biology*, 1, 137–151.

[4] Austin T. (2008) The emergence of the deterministic Hodgkin-Huxley equations as a limit from the underlying stochastic ion-channel mechanism. *Annals of Applied Probability*, 18, 1279–1325.

[5] Ball, K., Kurtz, T.G., Popovic, L., Rempala, G. (2006). Asymptotic analysis of multiscale approximations to reaction networks. *Annals of Applied Probability*, 16(4), 1925–1961.

[6] Bansaye, V., Méléard, S. (2015). *Stochastic Models for Structured Populations: Scaling Limits and Long Time Behavior*. Stochastics in Biological Systems. Springer.

[7] Barton, N.H., Etheridge, A.M., Véber, A. (2010). A new model for evolution in a spatial continuum. *Electronic Journal of Probability*, 15, 162–216.

[8] Batada, N.N., Shepp, L.A., Siegmund, D.O. (2004). Stochastic model of protein-protein interaction: Why signaling proteins need to be colocalized. *Proceedings of the National Academy of Sciences*, 101(17), 6445–6449.

[9] Bruna, M., Chapman, S. J. (2013). Diffusion of finite-size particles in confined geometries. *Bulletin of Mathematical Biology*, 76, 947–982.

[10] Bruna, M., Chapman, S. J., Smith, M. J. (2014). Model reduction for slow-fast stochastic systems with metastable behaviour. *Journal of Chemical Physics*, 140(17), 174107.

[11] Buckwar, E., Riedler, M.G. (2011). Exact modelling of neuronal membranes including spatio-temportal evolution. *Journal of Mathematical Biology*, 63(6), 1053–1091.

[12] Cappelletti, D., Wiuf, C. (2016). Elimination of intermediate species in multiscale stochastic reaction networks. *Annals of Appied Probability*, 26(5): 2915–2958.

[13] Champagnat, N., Méléard, S. (2007). Invasion and adaptive evolution for individual-based spatially structured populations. *Journal of Mathematical Biology*, 55(2), 147–188.

[14] Cloez, B., De Saporta, B., Joubaud, M. (2020). Optimal stopping for measure-valued piecewise deterministic Markov processes. *Journal of Applied Probability*, 57(2), 497–512.

[15] Crudu, A., Debussche, A., Muller, A., Radulescu, O. (2012). Convergence of stochastic gene networks to hybrid piecewise deterministic processes. *Annals of Applied Probability*, 22(5), 1822–1859.

[16] Davis, M.H. (1984). Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society: Series B* (Methodological), 46(3), 353–376.

[17] Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S. (2002). Stochastic gene expression in a single cell. *Science*, 297(5584), 1183–1186.

[18] Erban, R., Chapman, J., Maini, P. (2007). A practical guide to stochastic simulations of reaction-diffusion processes. 35pages, *arXiv:0704.1908*.

[19] Erban, R., Othmer, H. G. (2014). Special Issue on Stochastic Modelling of Reaction-Diffusion Processes in Biology. *Bulletin of Mathematical Biology*, 76:761–765.

[20] Ethier, S.N., Kurtz, T.G. (1986). *Markov processes: characterization and convergence*. Wiley.

[21] Evans, L.C. (1998). *Partial Differential Equations*. Graduate Studies in Mathematics Volume 19, American Mathematical Society.

[22] Forien, R. (2017). *Spatial structure of genetic diversity: the influence of natural selection and of heterogeneous environments*. PhD thesis, Paris-Saclay University.

[23] Genadot, A., Thieullen, M. (2012). Averaging for a fully coupled Piecewise Deterministic Markov Process in infinite dimension. *Advances in Applied Probability*, 44(3), 749–773.

[24] Grima, R., Schnell, S. (2008). Modelling reaction kinetics inside cells. *Essays in biochemistry*, 45, 41–56.

[25] Harrington, H. A., Feliu, E., Wiuf, C., Stumpf, M. P. (2013). Cellular compartments cause multistability and allow cells to process more information. *Biophysical journal*, 104(8), 1824–1831.

[26] Hattne, J., Fange, D., Elf, J. (2005). Stochastic reaction-diffusion simulation with MesoRD. *Bioinformatics*, 21(12), 2923–2924.

[27] Isaacson, S. A., Ma, J., Spiliopoulos, K. (2020). Mean Field Limits of Particle-Based Stochastic Reaction-Diffusion Models. arXiv preprint arXiv:2003.11868.

[28] Isaacson, S. A., Ma, J., Spiliopoulos, K. (2020). How reaction-diffusion PDEs approximate the large-population limit of stochastic particle models. *SIAM Journal on Applied Mathematics*, 81(6), 2622–2657.

[29] Kang, H.W., Kurtz, T.G. (2013). Separation of time-scales and model reduction for stochastic reaction networks. *Annals of Applied Probability*, 23(2), 529–583.

[30] Kang, H.W., Kurtz, T.G., Popovic, L. (2014). Central limit theorems and diffusion approximations for multiscale Markov chain models. *Annals of Applied Probability*, 24(2), 721–759.

[31] Kang, W., Ramanan, K. (2017). On the submartingale problem for reflected diffusions in domains with piecewise smooth boundaries. *The Annals of Probability*, 45(1), 404–468.

[32] Lajoie, P., Goetz, J.G., Dennis, J.W., Nabi, I.R. (2009). Lattices, rafts, and scaffolds: domain regulation of receptor signaling at the plasma membrane. *Journal of Cell Biology*, 185(3), 381–385.

[33] Leman, H. (2016). Convergence of an infinite dimensional stochastic process to a spatially structured trait substitution sequence. *Stochastic partial differential equations: analysis and computations*, 4(4), 791–826.

[34] Lim, T. S., Lu, Y., Nolen, J. H. (2020). Quantitative Propagation of Chaos in a Bimolecular Chemical Reaction-Diffusion Model. *SIAM Journal on Mathematical Analysis*, 52(2), 2098–2133.

[35] Lions, P. L., Sznitman, A. S. (1984). Stochastic differential equations with reflecting boundary conditions. *Communications on Pure and Applied Mathematics*, 37(4), 511–537.

[36] Loew, L.M., Schaff, J.C. (2001) The Virtual Cell: a software environment for computational cell biology. *Trends in Biotechnology*, 19, 401–406.

[37] Mackey, M. C., Santillán, M., Tyran-Kamińska, M., Zeron, E. S. (2016). *Simple mathematical models of gene regulatory dynamics*. Springer International Publishing.

[38] McSweeney, J. and Popovic, L. (2014). Stochastically-induced bistability in chemical reaction systems, *Annals of Applied Probability*, 24(3), 1226–1268.

[39] Paulsson, J. (2004). Summing up the noise in gene networks. *Nature*, 427(6973), 415–418.

[40] Pfaffelhuber, P., Popovic, L. (2015). Scaling limits of spatial compartment models for chemical reaction networks. *Annals of Applied Probability*, 25(6), 3162–3208.

[41] Pfaffelhuber, P., Popovic, L. (2015). How spatial heterogeneity shapes multiscale biochemical reaction network dynamics. *Journal of the Royal Society Interface*, 12(104), 20141106.

[42] Popovic, L. (2019). Large deviations of Markov chains with multiple time-scales. *Stochastic Processes and their Applications*, 129(9), 3319–3359.

[43] Raj, A., Van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*, 135(2), 216–226.

[44] Rebolledo, R. (1980). Sur l'existence de solutions à certains problèmes de semimartingales. *Comptes Rendus de l'Académie des Sciences Paris*, 290.

[45] Riedler, M., Thieullen, M., Wainrib, G. (2012). Limit theorems for infinite-dimensional piecewise deterministic Markov processes. Applications to stochastic excitable membrane models. *Electronic Journal of Probability*, 17.

[46] Robert, P. (2019). Mathematical models of gene expression. *Probability Surveys*, 16, 277–332.

[47] Roelly-Coppoletta, S. (1986). A criterion of convergence of measure-valued processes: application to measure-valued branching processes. *Stochastics*, 17, 43–65.

[48] Roelly, S., Rouault, A. (1990). Construction et propriétés de martingales des branchements spatiaux interactifs. *International Statistical Review/Revue Internationale de Statistique*, 173–189.

[49] Smith, S., Grima, R. (2019). Spatial stochastic intracellular kinetics: A review of modelling approaches. *Bulletin of Mathematical Biology*, 81(8), 2960–3009.

[50] Tanaka, H. (1979). Stochastic differential equations with reflecting boundary condition in convex regions. *Hiroshima Mathematical Journal*, 9(1), 163–177.

A Proof of Theorem 2.6

We first state and prove two lemmas which will be used later.

Lemma A.1. *Under Assumptions (A1) and (A2), for every $r \in R$, $M \in \mathcal{M}$, $(p_1, \dots, p_{k_r}) \in \mathcal{P}^{k_r}$ and $\bar{y} \in E$, we have*

$$\big|\lambda_r(\bar{y}, M; p_1, \dots, p_{k_r})\big| \leq \|h_r\|_{\infty,\ell(r,M)} \times \|\Gamma_\epsilon\|_\infty^{k_r},$$

where $\lambda_r(\bar{y}, M; p_1, \dots, p_{k_r})$ is the local reaction rate at $\bar{y}$ of particles $p_1, \dots, p_{k_r}$ defined in (2.4) and

$$\ell(r, M) = \|\Psi_r\|_\infty \langle M, 1\rangle.$$

In addition, the function $M \mapsto \lambda_r(\bar{y}, M; p_1, \dots, p_{k_r})$ is continuous on $\mathcal{M}$.

Proof. (Proof of Lemma A.1.) The bound is a straightforward consequence of the Assumptions and of the fact that $|\langle M, \Psi_{r,\bar{y}}\rangle| \leq \|\Psi_r\|_\infty \langle M, 1\rangle$. The continuity of the mapping $M \mapsto \lambda_r(\bar{y}, M; p_1, \dots, p_{k_r})$ comes from the facts that, according to Assumption (A1), h_r is Lipschitz in its second coordinate and $\Psi_{r,y}$ is continuous (and thus bounded) on $\mathcal{P}$. □

Lemma A.2. *Suppose Assumptions (A1) and (A2) are satisfied. Then the function Λ_r defined in* (2.8) *is continuous on $\mathcal{M}_p$.*

Proof. (Proof of Lemma A.2.) Let $r \in R$.

Let us first derive a bound on $|\Lambda_r(M_1) - \Lambda_r(M_2)|$ for every $M_1, M_2 \in \mathcal{M}_p$. We shall then use this bound to show that if $(M_n)_{n\geq 0}$ converges weakly to M in $\mathcal{M}_p$, then $\Lambda_r(M_n) \to \Lambda_r(M)$ as $n \to \infty$.

Let thus $M_1, M_2 \in \mathcal{M}_p$. Recalling the definition (2.4) of λ_r and our notation $p_i = (x_i, y_i)$, and leaving aside the integral over E for a moment, we have

$$\begin{aligned}
&\int_{\mathcal{P}^{k_r}} M_1^{\otimes\downarrow k_r}(dp_1, \dots, dp_{k_r}) \left(\prod_{i=1}^{k_r} \mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i - \bar{y})\right) \bar{h}_r(\bar{y}, M_1) \\
&= \int_{\mathcal{P}^{k_r}} M_1^{\otimes\downarrow k_r}(dp_1, \dots, dp_{k_r}) \left(\prod_{i=1}^{k_r} \mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i - \bar{y})\right) \big(\bar{h}_r(\bar{y}, M_1) - \bar{h}_r(\bar{y}, M_2)\big) \\
&\quad + \int_{\mathcal{P}^{k_r}} M_1^{\otimes\downarrow k_r}(dp_1, \dots, dp_{k_r}) \left(\prod_{i=1}^{k_r} \mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i - \bar{y})\right) \bar{h}_r(\bar{y}, M_2). \qquad \text{(A.1)}
\end{aligned}$$

By Assumption (A1), the reaction factor h_r is Lipschitz in its second coordinate, with Lipschitz constant L_r independent of the first coordinate, and so the absolute value of the first term on the r.h.s. of (A.1) is bounded by

$$\begin{aligned}
&\|\Gamma_\epsilon\|_\infty^{k_r} \langle M_1, 1\rangle^{k_r} \big|h_r\big(\bar{y}, \langle M_1, \Psi_{r,\bar{y}}\rangle\big) - h_r\big(\bar{y}, \langle M_2, \Psi_{r,\bar{y}}\rangle\big)\big| \\
&\leq \|\Gamma_\epsilon\|_\infty^{k_r} \langle M_1, 1\rangle^{k_r} L_r \big|\langle M_1, \Psi_{r,\bar{y}}\rangle - \langle M_2, \Psi_{r,\bar{y}}\rangle\big|. \qquad \text{(A.2)}
\end{aligned}$$

Integrating now with respect to $\varrho_r(d\bar{y})$, we can thus write

$$\begin{aligned}
\big|\Lambda_r(M_1) - \Lambda_r(M_2)\big| \leq &\|\Gamma_\epsilon\|_\infty^{k_r} \langle M_1, 1\rangle^{k_r} L_r \int_E \varrho_r(d\bar{y}) \big|\langle M_1, \Psi_{r,\bar{y}}\rangle - \langle M_2, \Psi_{r,\bar{y}}\rangle\big| \qquad \text{(A.3)} \\
&+ \int_E \varrho_r(d\bar{y})\, \bar{h}_r(\bar{y}, M_2) \left|\left\langle M_1^{\otimes\downarrow k_r} - M_2^{\otimes\downarrow k_r}, \prod_{i=1}^{k_r} \mathbf{1}_{A_i^r}(x_i)\Gamma_\epsilon(y_i - \bar{y})\right\rangle\right|.
\end{aligned}$$

Now suppose that $(M_n)_{n\geq 1}$ is a sequence of measures in $\mathcal{M}_p$ converging weakly to $M \in \mathcal{M}_p$, and consider the expression in (A.3) with $M_1 = M_n$ and $M_2 = M$. First, $\langle M_n, 1\rangle \to \langle M, 1\rangle$ and so there exists n_0 such that for every $n \geq n_0$, we have $\langle M_n, 1\rangle \leq 2\langle M, 1\rangle$. Thanks to the continuity and uniform boundedness of $\Psi_{r,\bar{y}}$ stated in Assumption (A1) and the fact that E has finite volume, we can use the dominated convergence theorem to conclude that

$$\lim_{n\to\infty} \|\Gamma_\epsilon\|_\infty^{k_r} \langle M_n, 1\rangle^{k_r} L_r \int_E \varrho_r(d\bar{y}) \left|\langle M_n, \Psi_{r,\bar{y}}\rangle - \langle M, \Psi_{r,\bar{y}}\rangle\right| = 0.$$

For the second term on the r.h.s. of (A.3), first observe that by Assumption (A1),

$$\forall \bar{y} \in E, \qquad \left|\bar{h}_r(\bar{y}, M)\right| \leq \|h_r\|_{\infty, \|\Psi_r\|_\infty \langle M, 1\rangle} < \infty.$$

Together with the continuity of Γ_ϵ, the fact that the set $\mathcal{T}$ of molecular types is finite and the weak convergence of the counting measure M_n to the counting measure M, this allows us to conclude that the second term on the r.h.s. of (A.3) (again, with $M_1 = M_n$ and $M_2 = M$) converges to 0 as $n \to \infty$ and consequently

$$\lim_{n\to\infty} \left|\Lambda_r(M_n) - \Lambda_r(M)\right| = 0.$$

This concludes the proof of the continuity of Λ_r on $\mathcal{M}_p$. □

We can now proceed to the proof of Theorem 2.6.

Proof. (Proof of Theorem 2.6.) (i) By construction, $(M_t)_{0\leq t<\tau^\infty}$ is an $\mathcal{M}_p$-valued càdlàg process a.s. and is Markovian. Indeed, by Assumption (A0) the movement of particles between two reaction times is given by a càdlàg Markov process. In addition, in the time interval $[\tau^j, \tau^{j+1})$ we have $\widetilde{M}_t^j = M_t$, and so the integrals appearing in the definition of the random times τ_r^{j+1} are adapted to the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ of $(M_t)_{t\geq 0}$ (for the filtration to be well-defined for all times, one may add a cemetery state ∂ and declare that $M_t = \partial$ for all $t \geq \tau^\infty$). This tells us that each τ_r^{j+1} (for $r \in R$), and hence τ^{j+1}, is a stopping time for $(\mathcal{F}_t)_{t\geq 0}$, for every $j \geq 0$; for every $t \geq 0$ and $k \in \mathbf{N}$, the event $\{j(t) = k\} = \{\tau^k \leq t\} \setminus \{\tau^{k+1} \leq t\}$ is thus $\mathcal{F}_t$-measurable. Using the lack of memory property of the exponential r.v., on the event $\{j(t) = k\}$ the distribution of τ^{k+1} conditionally on $\mathcal{F}_t$ is equal to its distribution conditionally on M_t. Finally, since the sampling rule and updating of the measure at time τ^{k+1} depends only on the state of the process at time $(\tau^{k+1})-$, we can conclude that $(M_t)_{0\leq t<\tau^\infty}$ has the Markov property.

(ii) The fact that $\tau^\infty = +\infty$ a.s. is an easy consequence of the first part of Assumption (A3), namely that

$$S_T := \sup_{t\in[0,T]} \langle M_t, 1\rangle < \infty \qquad \text{a.s.}, \tag{A.4}$$

for any fixed $T > 0$. Indeed, let us fix $T > 0$ and show that $\tau^\infty > T$ with probability one. Recalling the bound on λ_r stated in Lemma A.1, we simply have to observe that for every $r \in R$ and $t \in [0, T]$, we have

$$\begin{aligned}\int_{E\times\mathcal{P}^{k_r}} &\varrho_r(d\bar{y}) M_t^{\otimes\downarrow k_r}(dp_1, \ldots, dp_{k_r})\, \lambda_r\big(\bar{y}, M_t; p_1, \ldots, p_{k_r}\big) \\ &\leq (1 + \mathrm{Vol}(E)) S_T^{k_r} \|\Gamma_\epsilon\|_\infty^{k_r} \|h_r\|_{\infty, \|\Psi_r\|_\infty S_T} < \infty \quad \text{a.s.}\end{aligned}$$

Consequently, each τ^j (until time T) is stochastically bounded from below by the minimum of $|R|$ independent exponentially distributed random variables whose parameters are independent

of j. The number of j's such that $\tau^j \leq T$ is thus a.s. finite, which is equivalent to the property $\tau^\infty > T$ a.s. that we were seeking. The $\mathcal{M}_p$-valued process $(M_t)_{t\geq 0}$ is therefore defined for all times.

Finally, let us show that $(M_t)_{t\geq 0}$ satisfies the martingale problem MP(L) with initial distribution $\mathcal{L}(M_0)$, in the particular case where $\mathcal{L}(M_0) = \delta_{m_0}$ for some fixed $m_0 \in \mathcal{M}_p$. The general case can be obtained by integrating with respect to the law of M_0. Using Lemma A.1 and the fact that until τ^1, the (finite) total mass of the system remains unchanged, we have for every $r \in R$ and every $s < \tau^1$:

$$\begin{aligned}\Lambda_r\big(\widetilde{M}^0_s\big) &= \int_{E\times\mathcal{P}^{k_r}} \varrho_r(d\bar{y})\Big(\widetilde{M}^0_s\Big)^{\otimes\downarrow k_r}(dp_1,\dots,dp_{k_r})\,\lambda_r(\bar{y},\widetilde{M}^0_s;p_1,\dots,p_{k_r})\\ &\leq (1+\mathrm{Vol}(E))\langle M_0,1\rangle^{k_r}\|\Gamma_\epsilon\|_\infty^{k_r}\|h_r\|_{\infty,\ell(r,M_0)} < \infty \qquad \text{a.s.},\end{aligned}$$

where we recall that $\ell(r,M_0) = \|\Psi_r\|_\infty\langle M_0,1\rangle < \infty$. Consequently, for every $t \geq 0$ we can write

$$\begin{aligned}\mathbf{P}\big(\tau^1 > t\big) &= \mathbf{E}\Big[\mathbf{P}\big(\tau^1 > t \,\big|\, (\widetilde{M}^0_s)_{s\geq 0}\big)\Big] = \mathbf{E}\Big[\prod_{r\in R}\mathbf{P}\big(\tau^1_r > t \,\big|\, (\widetilde{M}^0_s)_{s\geq 0}\big)\Big]\\ &= \mathbf{E}\Big[\exp\Big\{-\sum_{r\in R}\int_0^t ds\,\Lambda_r\big(\widetilde{M}^0_s\big)\Big\}\Big]\\ &= 1 - \sum_{r\in R}\mathbf{E}\Big[\int_0^t ds\,\Lambda_r\big(\widetilde{M}^0_s\big)\Big] + \mathcal{O}\big(t^2\big). \end{aligned} \tag{A.5}$$

The same reasoning applies to $\tau^2 - \tau^1$, since the total mass of the system between times τ^1 and τ^2 is bounded by $\langle M_0,1\rangle + \sup_{r\in R}(k'_r - k_r) < \infty$. Let us now write for every function F_f of the form (2.9) belonging to the set $\mathbf{F}$ of Assumption (A0'):

$$\begin{aligned}\frac{1}{t}\,\mathbf{E}\big[F_f(M_t) - F_f(M_0)\big] &= \frac{1}{t}\,\mathbf{E}\big[\big(F_f(M_t) - F_f(M_0)\big)\mathbf{1}_{\{\tau^1>t\}}\big]\\ &\quad + \frac{1}{t}\,\mathbf{E}\big[\big(F_f(M_t) - F_f(M_0)\big)\mathbf{1}_{\{\tau^1\leq t\}}\big].\end{aligned} \tag{A.6}$$

Since $\mathbf{1}_{\{\tau^1>t\}} \to 1$ a.s. as $t \to 0$, $M_t = \widetilde{M}^0_t$ evolves only through the spatial movement of the particles initially described by M_0 as long as $t < \tau^1$ and the function F_f belongs to the domain of $\mathcal{D}$, we have that

$$\lim_{t\to 0}\frac{1}{t}\,\mathbf{E}\big[\big(F_f(M_t) - F_f(M_0)\big)\mathbf{1}_{\{\tau^1>t\}}\big] = \mathcal{D}F_f(M_0). \tag{A.7}$$

Next, we have

$$\frac{1}{t}\,\mathbf{E}\big[\big(F_f(M_t) - F_f(M_0)\big)\mathbf{1}_{\{\tau^1\leq t\}}\big] = \mathbf{E}\big[\big(F_f(M_t) - F_f(M_0)\big)\,\big|\,\tau^1 \leq t\big]\frac{\mathbf{P}(\tau^1\leq t)}{t}. \tag{A.8}$$

Using (A.5) and the continuity of Λ_r shown in Lemma A.2, we obtain that

$$\lim_{t\to 0}\frac{\mathbf{P}(\tau^1\leq t)}{t} = \sum_{r\in R}\Lambda_r(M_0).$$

Furthermore, conditionally on τ^1, the lapse of time $\tau^2 - \tau^1$ between the first and second reaction satisfies the same type of asymptotics as in (A.5) and so the probability that two reactions occur

before time t (*i.e.*, that $\tau^2 \leq t$) is of the order of $\mathcal{O}(t^2)$. As a consequence, integrating over all possible values of M_{τ^1} and disregarding the unlikely event that two reactions occur before time t yields

$$\begin{aligned}
&\lim_{t\to 0} \frac{1}{t}\,\mathbf{E}\big[\big(F_f(M_t) - F_f(M_0)\big)\mathbf{1}_{\{\tau^1\leq t\}}\big] \\
&= \Big(\sum_{r'\in R}\Lambda_{r'}(M_0)\Big)\Bigg\{\sum_{r\in R}\frac{\Lambda_r(M_0)}{\sum_{r'\in R}\Lambda_{r'}(M_0)}\int_{E\times\mathcal{P}^{k_r}}\varrho_r(d\bar{y})M_0^{\otimes\downarrow k_r}(dp_1,\dots,dp_{k_r}) \\
&\qquad \frac{\lambda_r(\bar{y},M_0;p_1,\dots,p_{k_r})}{\Lambda_r(M_0)}\Big[F\Big(\Big\langle M_0 - \sum_{i=1}^{k_r}\delta_{p_i} + \sum_{i=1}^{k_r'}\delta_{(B_i^r,\bar{y})}, f\Big\rangle\Big) - F_f(M_0)\Big] \\
&= \sum_{r\in R} G_r F_f(M_0),
\end{aligned}$$

where in the first equality, each summand is the product of the probability that reaction r is the first to occur and of the integral describing the sampling of a location and a set of source reactants as specified in our construction of $(M_t)_{t\geq 0}$. Note that the continuity of λ_r and Λ_r on $\mathcal{M}_p$, stated respectively in Lemmas A.1 and A.2, are used here to obtain that the sampling of the location and particles at time 0+ is made according to $\lambda_r(\bar{y},M_0;p_1,\dots,p_{k_r})/\Lambda_r(M_0)$. Together with (A.6) and (A.7), this gives us that

$$\lim_{t\to 0}\frac{1}{t}\,\mathbf{E}\big[F_f(M_t) - F_f(M_0)\big] = LF_f(M_0).$$

Using the bound on $\mathcal{D}F_f$ stated in Assumption (A0') and the bound on the $(1+\max_{r\in R}k_r)$-th moment of the total mass stated in Assumption (A3) together with the inequalities

$$\begin{aligned}
|G_rF_f(M_s)| &\leq 2\|F\|_\infty\|\Gamma_\epsilon\|_\infty^{k_r}\int_{E\times\mathcal{P}^{k_r}}\varrho_r(d\bar{y})M_s^{\otimes\downarrow k_r}(dp_1,\dots,dp_{k_r})\,h_r(\bar{y},\langle M_s,\Psi_{r,\bar{y}}\rangle) \\
&\leq 2\|F\|_\infty\|\Gamma_\epsilon\|_\infty^{k_r}(1+\mathrm{Vol}(E))\langle M_s,1\rangle^{k_r}\big(h_r(\bar{y},0) + L_r\langle M_s,\Psi_{r,\bar{y}}\rangle\big) \\
&\leq 2\|F\|_\infty\|\Gamma_\epsilon\|_\infty^{k_r}(1+\mathrm{Vol}(E))\langle M_s,1\rangle^{k_r}\big(h_r(\bar{y},0) + L_r\|\Psi_r\|_\infty\langle M_s,1\rangle\big),
\end{aligned}$$

where the second inequality uses the fact that h_r is Lipschitz in its second coordinate by Assumption (A1), we can then write that for every $t>0$ and F_f as above,

$$F_f(M_t) - \int_0^t ds\, LF_f(M_s)$$

is integrable. Extending the above computations to $F_f(M_{t_2}) - F_f(M_{t_1})$ for any $0\leq t_1<t_2$ and using the Markov property of $(M_t)_{t\geq 0}$ allows us to conclude that $(M_t)_{t\geq 0}$ is a solution to MP(L). Theorem 2.6 is proved. □

Remark A.3. If instead of assuming that $|\mathcal{D}F_f(M)| \leq c_{F,f}\langle M,1\rangle$ as in the statement of Assumption (A0'), we suppose that there exists $K\in\mathbf{N}$ and $c_{F,f}>0$ such that

$$|\mathcal{D}F_f(M)| \leq c_{F,f}\langle M,1\rangle^K \qquad \text{for every } M\in\mathcal{M}_p,$$

we see from the end of the proof of Theorem 2.6 that the integrability of

$$F_f(M_t) - \int_0^t ds\, LF_f(M_s)$$

will be a guaranteed by the stronger condition:

(A3') For every $T > 0$,

$$\sup_{t\in[0,T]} \langle M_t, 1\rangle < \infty \quad \text{a.s., and} \quad \sup_{t\in[0,T]} \mathbf{E}\Big[\langle M_t, 1\rangle^{K\vee(1+\max_r k_r)}\Big] < \infty.$$

This justifies the claim in Remark 2.5(a).