

Spatial Fluid Limits for Stochastic Mobile Networks

Max Tschaikowski, Mirco Tribastone

IMT School for Advanced Studies Lucca, Italy

Abstract

We consider Markov models of large-scale networks where nodes are characterized by their local behavior and by a mobility model over a two-dimensional lattice. By assuming random walk, we prove convergence to a system of partial differential equations (PDEs) whose size depends neither on the lattice size nor on the population of nodes. This provides a macroscopic view of the model which approximates discrete stochastic movements with continuous deterministic diffusions. This result may be used in practice to justify a space coarsening induced by the numerical PDE solver. We illustrate this by modeling a network of mobile nodes with on/off behavior performing file transfers with connectivity to 802.11 access points, empirically validated against discrete-event simulation. We show high quality of the PDE approximation in the case of moderate lattice granularity and population sizes.

Keywords: Stochastic networks, mobility models, fluid limits, reaction-diffusion.

1. Introduction

This paper is concerned with the modeling and analysis of large-scale computer and communication networks with mobile nodes. We consider an expressive model of a node which is assumed to evolve over a set of *local states*. These may represent, for instance, distinct phases of a node's behavior that distinguish between periods of activity and inactivity in a wireless network or, with a richer granularity, they may be related to different states in a communication protocol. We are interested in understanding how the interplay between the node's local behavior and its mobility pattern have an impact on the overall system's performance.

In this paper, we consider *Markov population processes* as the starting point of our analysis. These represent a general class of continuous-time Markov chains (CTMCs), characterized by a state descriptor which is given by a vector where each element gives the population of nodes in a specific local state. In general, such stochastic models do not scale well with increasing system sizes because they are based on an explicit discrete-state representation. This makes the analysis very difficult computationally and poses a detriment to model-assisted parameter-space exploration and capacity planning. However, under mild assumptions, a Markov population process may yield a *fluid approximation* [1, 2, 3, 4, 5, 6, 7] as a system of ordinary differential equations (ODEs). This can be shown to be the deterministic limit behavior that holds when the number of nodes goes to infinity; see, e.g., [8]. Fluid limits have been developed for a wide range of models of distributed and networked systems, including, for instance, load balancing [9, 10], optical switches [11], virtualized environments [12], and peer-to-peer networks [13]. In [14] Benăim and Le Boudec offer a general framework for a Markov model of *interacting agents* evolving through a set of *local states*. In all cases, while the Markov process has a state space with cardinality that is exponential in the number of agents (in the worst case), the fluid limit only depends linearly on the number of local states. This procedure is very attractive in practice because it enables computationally efficient solutions, when the local state space is small, that are typically very accurate for large populations.

*Corresponding author

Email addresses: max.tschaikowski@imtlucca.it (Max Tschaikowski), mirco.tribastone@imtlucca.it (Mirco Tribastone)

In the aforementioned models, however, spatial effects are not explicitly present because the system is inherently static, or because it is a convenient simplification to abstract away from them for the purposes of model tractability [15]. Unfortunately, explicitly taking into account mobility and space in a fluid model generally leads to a rapid growth of the local state space. In this paper, for instance, we deal with population processes defined by a CTMC with a random walk (RW) model over a two-dimensional topology. We consider a typical partitioning of the spatial domain into *regions* (or *patches*, as usually called in theoretical biology [16]). Within each region an assumption of homogeneity is made, and spatial effects are incorporated as transitions across regions [17]. If the dynamics of each region admits a fluid limit, then the overall system behavior is given by $L(K + 1)^2$ ODEs, where, according to the notation of this paper, L is the local state space size and $(K + 1)^2$ is the total number of regions. Such dependence is simply due to the fact that the model must keep track of the total population of nodes in each of the L local states, *at each region*.

We consider this problem in the context of a generic framework for *mobile reaction networks*, a concise model to describe Markov population processes. This is inspired by stochastic reaction networks, which are frequently used for the analysis of systems of (bio-)chemical reactions (e.g., [18]). We assume that nodes may interact with each other within the same region, and undergo unbiased RW with a parameter, the *migration rate*, which can be dependent on their local state. We also consider *absorbing* and *reflective* boundary conditions for the spatial domain. In the presence of the former, each node that hits the boundary exits the domain without the possibility of ever re-entering it. Instead, if the latter is considered, the boundaries build a barrier of the domain that cannot be crossed by nodes. Those assumptions do not exclude the possibility of having exogenous arrivals into the system, e.g., by means of a Poisson process in each region.

The purpose of this paper is to develop a technique to effectively analyze models when K and the total population of nodes are large. It is natural to consider an approach with two dimensions of scaling: the first one is the celebrated *density-dependent* form that allows us to obtain a continuous deterministic limit for the concentrations (or densities) of the node populations [8], while keeping the regions discrete. This is achieved by constructing a family of population processes indexed by a parameter, hereafter denoted by N , such that the initial state of each element of the family is a population of nodes that grows linearly with N . Then, the result of Kurtz [8] ensures that the sequence converges, as $N \rightarrow \infty$, to the solution of an ODE system of size $L(K + 1)^2$. The second scaling occurs on the spatial dimension. We define a suitable sequence of migration rates such that increasing K means considering regions which are closer and closer to each other on a regular mesh in the unit square in \mathbb{R}^2 (i.e., the regions are at a distance $1/K$ from each other). After setting the scene, we show that the sequence of ODE systems converges, as $K \rightarrow \infty$, to a system of L partial differential equations (PDEs) of reaction-diffusion type. In these PDEs, the diffusive terms model the continuous migration across regions, whereas the reactive terms describe the local interactions between nodes. To the best of our knowledge, such a convergence result has not been proven before. We also argue that our limit may allow for a more efficient numerical solution. This may seem surprising because analytical solutions of PDE systems are scarce and numerical solutions of PDE systems rely on a discretization of space. However while in the case of the limit ODE system of size $L(K + 1)^2$ the discretization is dictated by the stochastic model, in our limit PDE system the coarseness of the discrete mesh depends on the PDE solver. In other words, the PDE solver may in effect give rise to a coarsening of the original spatial domain.

A typical situation of practical interest to which our framework can be applied is that of large-scale mobile networks such as personal communication services (e.g., [19]): there are many base stations (e.g., in a wide-area cellular network) and the area served by a base station can be modeled as a region, which can contain potentially many mobile nodes that may migrate across regions. Here, the modeler may wish to predict how nodes distribute across the network over a given time horizon [20]. In this paper, we demonstrate the applicability of our method by studying a network of mobile nodes with connectivity provided by an 802.11 access point located in each region. First, we successfully validate our CTMC model against discrete-event simulation using the JiST/SWANS discrete-event simulation framework [21]. Then, we show that the PDE solution provides an excellent estimate of the network's performance for relatively small population sizes and moderate lattice granularity.

Paper outline. The remainder of the paper is organized as follows. Section 2 overviews related work. In order to fix notation and provide intuition on the scaling limits considered in the paper, Section 3 introduces *stationary reaction networks*, which describe *static* networks with no mobility, whilst allowing nodes to be described by a local state space. Here we discuss how such models admit a classic fluid limit as a system of ODEs. Section 4 presents mobile

reaction networks, which are defined as a conservative extension of stationary networks with an explicit mobility model. We introduce a straightforward spatial ODE limit that depends upon the lattice granularity. In Section 5 we discuss the main contribution of this paper, namely the convergence of the spatial ODE limit to a system of reaction-diffusion PDEs by assuming that nodes undergo unbiased random walk during their evolution. Section 6 discusses the numerical tests on our validation model. Finally, Section 7 concludes the paper.

2. Related Work

Reaction-diffusion PDEs. PDEs of reaction-diffusion type are very well understood in many disciplines, such as biology [22], ecology [23], and chemistry [24]. It is beyond the scope of this paper to provide a general overview of the literature. Instead, we focus on related work that, similarly to ours, considers PDEs as the macroscopic deterministic behavior of a stochastic process.

In physical chemistry, one such approach is to consider the so-called *reaction-diffusion master equation*, which corresponds to the forward equations of a CTMC that models a network of biochemical reactions occurring at discrete sites, and molecular transitions across sites [25]. Although the stochastic model corresponds to ours, a result of convergence is not provided as done in this paper, for two reasons. First, the underlying ODE system is obtained by means of a moment-closure technique, meaning that [25] *makes the approximation* that for a given stochastic process $X(t)$ and a function f it holds that $\mathbb{E}[f(X(t))] \approx f(\mathbb{E}[X(t)])$, although this does not hold in general for nonlinear f . Second, [25] *assumes* implicitly that the underlying ODE systems of size $\mathcal{O}(K^2)$ converge, as $K \rightarrow \infty$, to a PDE system. Instead, we prove both facts formally by applying the following strategy. First, we use Kurtz's result to show that the underlying rescaled stochastic process converges in probability to an ODE system of size $L(K+1)^2$. In this respect, we justify approximation $\mathbb{E}[f(X(t))] \approx f(\mathbb{E}[X(t)])$ as being asymptotically true in the limit of infinite populations. Afterwards, we use numerical analysis to show that the ODE systems converge, as $K \rightarrow \infty$, to a PDE system of size L . By combining both statements, we are able to prove that the underlying rescaled stochastic process *converges in probability* to a PDE system of size L .

In the recent work [26] the authors study a spatial process algebra for the modeling of mobile systems by means of reaction-diffusion PDEs. However, similarly to [25], no formal proof of convergence is provided, meaning that the ODE system is *assumed to converge* to the reaction-diffusion PDE system. Moreover, [26] is not as expressive as our framework, because it considers a specific communication pattern and supports only reflective boundaries. Instead, [27] provides a fluid limit result for the stochastic behavior of a single agent by combining the present approach with the concept of Fast Simulation [28, 29].

The closest approach to ours is found in [30], which was subsequently worked upon in a series of papers [31, 32, 33, 34]. In the case of absorbing boundaries, their stochastic models are different in that the local reaction rates model only birth and death, i.e. interactions between multiple types of agents are not possible. In the case of reflective boundary conditions, instead, we generalize [31] because we can treat more than one type of node.

Mobility models. There is a substantial amount of work on mobility models, both at the analytical level and experimentally through traces. In this paragraph we overview the literature that is most closely related to our approach based on the RW model; for an extensive discussion, we refer to the survey [35].

Owing to its analytical tractability, the RW model has been extensively studied in networking research. A discrete-time Markovian model was developed in [36] for the comparative evaluation of update strategies in cellular networks. The paper considers a mobility model where a node is characterized by states that determine the direction of movement. In this respect our approach is analogous, since agent types may have different movement rates. However, the results of their analysis are presented for a one-dimensional topology (over a ring) and cannot be lifted to more general local interactions; on the other hand, their mobility model is anisotropic. Unbiased RWs are instead proposed in [20] and [19] to study movements across cellular networks, and to study routing protocols [37, 38] and performance characteristics in ad-hoc networks [39]. In all cases, mobility is not coupled with a model of interaction between nodes. RW is also used in [40] as the basic mobility model by which the authors arrive at a reaction-diffusion type equation for information propagation in ad-hoc networks; their analysis is carried out at the deterministic/macroscopic level without considering a limiting regime of a counterpart microscopic/stochastic process. Instead, the PDEs used in [41] are interpreted as the deterministic limit of the empirical measure of node concentrations, by appealing to the strong law of large numbers.

Summary. In conclusion, deterministic characterizations of large-scale models of interacting agents are well established in the spatially homogeneous case; on the other hand, diffusion-reaction type equations are well understood when only the spatial/mobility model is of interest, and the results available for cases that also include certain kinds of local interactions do not generalize. In this paper, we bridge the gap between these two perspectives by providing a generic framework for modeling a richer variety of local interactions coupled with a (possibly state-dependent) RW model.

3. Stationary Reaction Networks

In this section we provide a definition of a classic stochastic reaction network which does not consider an explicit mobility model. For this reason, this is called a *stationary reaction network*. For illustrative purposes, we also show how this notation allows us to recover two different fluid models of networked computing systems already studied in the literature. This sets the stage for Section 4, which, instead, discusses the novel contribution of this paper, consisting of extending a stationary reaction network with space and mobility by means of a RW.

We introduce a CTMC population process where nodes are characterized by L local states, denoted by A_1, \dots, A_L . Our CTMC has state descriptor

$$\vec{A} := (A_1, \dots, A_L),$$

which gives the node populations in each local state. Nodes engage in J interactions, labeled $1, \dots, J$. These are given in the *reaction notation*

$$j : \sum_{1 \leq l \leq L} c_{jl} A_l \xrightarrow{F_j} \sum_{1 \leq l \leq L} d_{jl} A_l, \quad 1 \leq j \leq J. \quad (1)$$

The arrow label F_j denotes the function $F_j : \mathbb{R}^{L+P} \rightarrow \mathbb{R}$ that describes the rate at which the interaction occurs. It is dependent on the population of nodes in each of the L local states, and on $P \geq 0$ rate parameters. Additionally, F_j may explicitly depend on the *scaling parameter* $N \in \mathbb{N}$; this will be used later in this section to construct a sequence of CTMCs, indexed by N , that will enjoy convergence to the ODE fluid limit. The parameters c_{jl} and d_{jl} , with $1 \leq j \leq J$, $1 \leq l \leq L$ are instead nonnegative integers that describe the number of nodes in each local state that are simultaneously affected by the j -th transition. Without loss of generality, we assume that $c_{jl} > 0$ and $0 \leq A_l < c_{jl}$ imply $F_j(A_1, \dots, A_L, \beta_1, \dots, \beta_P) = 0$ for all $1 \leq j \leq J$ and $1 \leq l \leq L$. This condition ensures that population do not become negative after a transition occurs, thus maintaining the model meaningful from a physical point of view. Let a transition rate of the CTMC between states \vec{A} and \vec{A}' be denoted by $q(\vec{A}; \vec{A}')$. In our stationary reaction network, the rates are given by

$$q(\vec{A}; (A_1 + d_{j1} - c_{j1}, \dots, A_L + d_{jL} - c_{jL})) = \sum_{j' \in S_j} F_{j'}(A_1, \dots, A_L, \beta_1, \dots, \beta_P),$$

where S_j is the set of all actions which lead to the same state change as the j -th action, i.e., $S_j = \{1 \leq j' \leq J \mid \forall 1 \leq l \leq L (d_{j'l} - c_{j'l} = d_{jl} - c_{jl})\}$.

Let us now show two examples of how models that are already available in the literature can be captured with this definition.

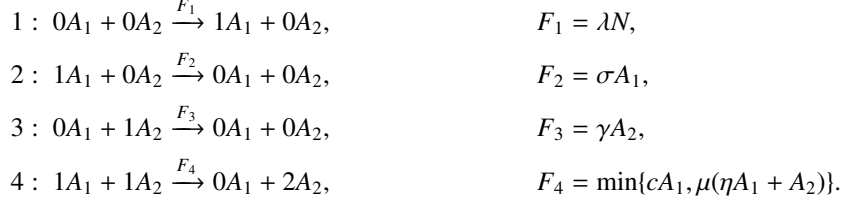
Example 1. We consider one of the models of epidemic routing studied in [15]. Its basic version has a single local state, A_1 , which denotes the number of susceptible nodes within an overall population of N nodes (thus, $N - A_1$ is the population of infected nodes). The single reaction is

$$1 : \quad 1A_1 \xrightarrow{F_1} 0A_1, \quad F_1 = \frac{1}{N} \beta A_1 (N - A_1),$$

where $\beta > 0$ is the contact rate for the infection. For instance, assuming $N > 5$, the state $\vec{A} = (5)$ has a transition to $\vec{A}' = (5 + 0 - 1) = (4)$ with rate $5\beta(1 - 5/N)$.

Example 2. The deterministic model of peer-to-peer file sharing presented in [42] can be shown to be the fluid limit of the following CTMC population process. Let A_1 denote the number of downloaders, i.e., peers that have a partial

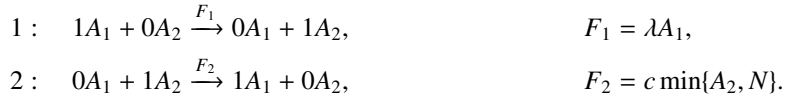
copy of a shared file, and let A_2 denote the number of seeds, i.e., peers that have already completed the download. The reactions are:



Interaction 1 gives the rate of arrivals of new downloaders into the network, parametrized by $\lambda > 0$ and dependent on the network size via N . Interactions 2 and 3 give the rate of exit of downloaders and seeds, respectively, with $\sigma, \gamma > 0$ being their individual exit rates. Finally interaction 4 gives the rate at which a downloader completes the file, thus becoming a seed. This is based on a bandwidth-sharing argument between the total requested capacity by downloaders, cA_1 , with $c > 0$, and the total upload capacity of the system. The latter is given by the total upload seed capacity (where $\mu > 0$ is the individual upload capacity) and that of the downloaders who are sharing their partial copy; $0 \leq \eta \leq 1$ is the probability that a requested portion of the file is available at a peer.

The following model represents the stationary version of the case study used in Section 6 for the numerical validation.

Example 3 (Nodes with on/off behavior). *Let us consider a node with two local states, A_1 , and A_2 , denoting an off state where the node does not use network capacity, and an on state where it downloads a file, respectively. Let us assume that all nodes share the same capacity cN and that $1/\lambda$ is the average time spent by the node in the off state. The model is given by the following reactions.*



The interpretation of the interaction function F_2 is similar to F_4 in Example 2 and is based on the same argument of bandwidth sharing. Specifically, $1/c$ is the average time to download a file. The network offers a capacity equal to cN , whereas the downloading users request an overall capacity cA_2 . Thus, the actual number of files downloaded per unit time is $\min\{cA_2, cN\} = F_2$. Notice that this does introduce resource contention, hence delay experienced by the downloading nodes. For instance, for $N = 1$ and $A_2 > 1$ the capacity is c , which will be shared amongst all nodes.

In order to obtain a fluid limit for a stationary reaction network, we define a CTMC sequence, denoted by $(\vec{A}_N(t))_{t \geq 0}$ and indexed by the scaling parameter N such that the initial state of the N -th CTMC is given by $A_i(0) = \lfloor N\alpha_i^0 \rfloor$, where $\alpha_i^0 \in \mathbb{R}_{\geq 0}$. Therefore the initial population of nodes increases linearly with N . If one considers the rescaled CTMC sequence $(\vec{A}_N(t)/N)_{t \geq 0}$ in the examples above, the larger N the faster the rates and, at the same time, the smaller the jumps; these are of order $O(1/N)$ because the non-normalized CTMC has unitary decreases in the populations. Informally, this behavior suggests a trend that is continuous in the limit.

More formally, convergence of $(\vec{A}_N(t)/N)_{t \geq 0}$ to such a fluid limit needs two assumptions. First, we require that for each $1 \leq j \leq J$ there exist continuous functions $f_j : \mathbb{R}^{L+P} \rightarrow \mathbb{R}$ and $g_j : \mathbb{R}^{L+P} \rightarrow \mathbb{R}_{\geq 0}$ such that

$$(\mathbf{A1}): \quad \frac{1}{N} F_j(Na_1, \dots, Na_L, b_1, \dots, b_P) = f_j(a_1, \dots, a_L, b_1, \dots, b_P) + \mathcal{O}\left(\frac{g_j(a_1, \dots, a_L, b_1, \dots, b_P)}{N}\right).$$

Second, we require that for all $\vec{z}_0 \in \mathbb{R}^{L+P}$ there exist an open neighborhood \mathcal{D} of \vec{z}_0 and a $C \in \mathbb{R}_{\geq 0}$ such that

$$(\mathbf{A2}): \quad |f_j(\vec{z}_2) - f_j(\vec{z}_1)| \leq C \|\vec{z}_2 - \vec{z}_1\|, \quad \forall \vec{z}_1, \vec{z}_2 \in \mathcal{D}.$$

In essence, **(A1)** asserts that the underlying CTMC is in the density-dependent form, while **(A2)** requires that each f_j be locally Lipschitz continuous. Both conditions will allow us to formally relate the CTMC sequence to a system

of L ODEs. For instance, with regard to Example 3 we have that

$$f_1(a_1, a_2) = \lambda a_1 \quad \text{and} \quad f_2(a_1, a_2) = c \min\{a_2, 1\}$$

satisfy **(A1)** for F_1 and F_2 , respectively, and enjoy locally Lipschitz continuity. These functions constitute the vector field of the ODE limit. More precisely, the following holds.

Theorem 1 (ODE Fluid Limit of Stationary Networks). *Let us define*

$$\mathcal{L}_l(\vec{a}) = \sum_{1 \leq j \leq J} (d_{jl} - c_{jl}) f_j(a_1, \dots, a_L, \beta_1, \dots, \beta_P). \quad (2)$$

Then, the ODE system

$$\frac{d}{dt} a_l(t) = \mathcal{L}_l(\vec{a}(t))$$

subjected to the initial condition $\vec{a}^0 = (\alpha_1^0, \dots, \alpha_L^0)$ has a unique solution \vec{a} in \mathbb{R}^L . Moreover, under the assumption that for an arbitrary but fixed $T > 0$ the time domain of \vec{a} contains $[0; T]$, it holds that

$$\lim_{N \rightarrow \infty} \mathbb{P} \left\{ \sup_{0 \leq t \leq T} \left\| \frac{1}{N} A_N(t) - \vec{a}(t) \right\|_{\infty} > \varepsilon \right\} = 0, \quad \forall \varepsilon > 0,$$

where the supremum norm $\|\cdot\|_{\infty}$ on \mathbb{R}^L is given by $\|\vec{a}\|_{\infty} := \max_{1 \leq l \leq L} |a_l|$.

We do not give the proof of this theorem explicitly because it can be seen as a special case of the *spatial ODE fluid limit*, see Theorem 2 in Section 4.

For instance, the ODEs for Example 3 are

$$\begin{aligned} \frac{d}{dt} a_1(t) &= -\lambda a_1(t) + c \min\{a_2, 1\}, \\ \frac{d}{dt} a_2(t) &= +\lambda a_1(t) - c \min\{a_2, 1\}. \end{aligned} \quad (3)$$

Let us consider $\vec{a}^0 = (\alpha_1^0, \alpha_2^0) = (1/4, 3/4)$, i.e., we study the model under the conditions where there are initially 25% of nodes in local state A_1 and 75% of nodes in local state A_2 . This leads to setting $a_1(0) = 1/4$ and $a_2(0) = 3/4$. The initial states in the CTMC sequence $(\vec{A}_N(t))_{t \geq 0}$ are given by

$$A_1(0) = \left\lfloor \frac{N}{4} \right\rfloor \quad \text{and} \quad A_2(0) = \left\lfloor \frac{3N}{4} \right\rfloor, \quad N \geq 1.$$

Let us note that $\frac{1}{N} \vec{A}_N(0) \rightarrow (1/4, 3/4)$, as $N \rightarrow \infty$, i.e., the initial populations of the rescaled Markov population processes converge to the densities given as the ODE initial condition.

4. Mobile Reaction Networks

A mobile reaction network is a Markov population process with an explicit notion of locality and mobility. Space is partitioned in a number of regions. Nodes within the same region may communicate with each other using the interaction functions (1). Additionally, nodes may move to neighboring regions by performing an unbiased RW. The purpose of this section is to show that a straightforward spatial ODE limit result for such a CTMC where the explicit location of nodes must be kept track of leads to a number of ODEs that grows with the number of regions, hindering the practical applicability of the analysis for large-scale mobile systems. This is an intermediate deterministic model that will be used in Section 5 to study a suitable continuous approximation of space by means of a system of L PDEs.

Space Model. Our space model consists of a lattice in the unit square with $(K + 1)^2$ regions, denoted by $\mathcal{R}_K := \{(i\Delta s, j\Delta s) \mid 0 \leq i, j \leq K\}$, where $\Delta s := 1/K$. To enhance readability, the boundary of the lattice will be denoted by

$\Omega_K := \{(x, y) \mid x \in \{0, 1\} \vee y \in \{0, 1\}\}$. A node in the interior $\mathcal{R}_K \setminus \Omega_K$ can travel to one of its neighboring regions $(x - \Delta s, y)$, $(x + \Delta s, y)$, $(x, y - \Delta s)$ and $(x, y + \Delta s)$. In the case where absorbing boundary conditions are assumed, nodes that migrate to the boundary Ω_K disappear. Instead, if reflective boundary conditions are in place, nodes at the boundary Ω_K can travel only to two or three regions. In the following, $\mathcal{N}(x, y)$ denotes the neighboring regions of $(x, y) \in \mathcal{R}_K$, e.g. $\mathcal{N}(0, 0) = \{(\Delta s, 0), (0, \Delta s)\}$.

4.1. Absorbing Boundary Conditions

Stochastic Model. The state descriptor of our population process is

$$\vec{A} := (A_1^{(x,y)}, \dots, A_L^{(x,y)})_{(x,y) \in \mathcal{R}_K},$$

which gives the agent populations in each local state at each region. This conservatively extends the stationary model in Section 3 by keeping track of the population of nodes in each of the (discrete) regions in the space domain.

Let us now proceed with the discussion of the transition rates for the RW and for the local interactions separately.

For the RW, we let $\mu_l^K \geq 0$ denote the migration rate for nodes of type- l in a region contained in \mathcal{R}_K . The transition rates for an unbiased RW are defined as

$$q(\vec{A}; (\dots, A_l^{(x,y)} - 1, A_l^{(\tilde{x}, \tilde{y})} + \mathbb{1}(\tilde{x}, \tilde{y}))) = \begin{cases} \mu_l^K A_l^{(x,y)} & , (\tilde{x}, \tilde{y}) \in \mathcal{R}_K \setminus \Omega_K, \\ |\mathcal{N}(x, y) \cap \Omega_K| \mu_l^K A_l^{(x,y)} & , \text{otherwise,} \end{cases}$$

where $(x, y) \in \mathcal{R}_K \setminus \Omega_K$, $(\tilde{x}, \tilde{y}) \in \mathcal{N}(x, y)$. The ellipsis indicates that, apart from the elements of the state vector that are explicitly written, no other elements are affected by the transition. The function $\mathbb{1}(\cdot) := \mathbb{1}_{\mathcal{R}_K \setminus \Omega_K}(\cdot)$ abbreviates the indicator of the set of inner regions $\mathcal{R}_K \setminus \Omega_K$. This formally describes an unbiased RW since all the allowed neighboring regions are visited with equal probabilities. The boundary regions, instead, absorb any node that visits them.

The transition rates due to the interaction functions are similar to (1). However, in our mobile reaction network we allow the rate parameters—but not the scaling parameter $N \in \mathbb{N}$ —to be possibly dependent on the region. This might be useful, for instance, to model inhomogeneous server capacities across the spatial domain. To capture this, for any $1 \leq j \leq J$, F_j is a function $F_j : \mathbb{R}^{L+P} \rightarrow \mathbb{R}$ where the P parameters are given by continuous functions

$$\beta_p : [0; 1]^2 \rightarrow \mathbb{R}_{\geq 0}, \quad \text{with } 1 \leq p \leq P.$$

The transition rates are given by

$$q(\vec{A}; (\dots, A_1^{(x,y)} + d_{j1} - c_{j1}, \dots, A_L^{(x,y)} + d_{jL} - c_{jL})) = \sum_{j \in \mathcal{S}_j} F_j(A_1^{(x,y)}, \dots, A_L^{(x,y)}, \beta_1(x, y), \dots, \beta_P(x, y)),$$

where $(x, y) \in \mathcal{R}_K \setminus \Omega_K$. We notice that, indeed, this definition induces *local interactions*, i.e., only the population levels related to the same region (x, y) are affected by the transition. Thus, overall, our model consists of local communication between nodes in a region, together with a mobility model which can be dependent on the local state of the node.

Spatial ODE Fluid Limit. We now consider a CTMC sequence $(\vec{A}_N(t))_{t \geq 0}$ indexed by N . The initial state of the N -th CTMC is given by $A_l^{(x,y)}(0) = [N \alpha_l^0(x, y)]$, where $\alpha_1^0, \dots, \alpha_L^0 : [0; 1]^2 \rightarrow \mathbb{R}_{\geq 0}$ are now continuously differentiable functions which are zero at the boundary of $[0; 1]^2$. The assumption on the scaling with N is analogous to that of a stationary reaction network. Instead, requiring that the functions are zero at the boundary is for consistency with the fact that the boundary Ω_K of \mathcal{R}_K does not contain any nodes. Formally, this means that α_l^0 have to satisfy the Dirichlet Boundary Conditions (DBC) s

$$\alpha_l^0(x, 1) = 0, \quad \alpha_l^0(x, 0) = 0, \quad \alpha_l^0(1, y) = 0, \quad \alpha_l^0(0, y) = 0, \quad (4)$$

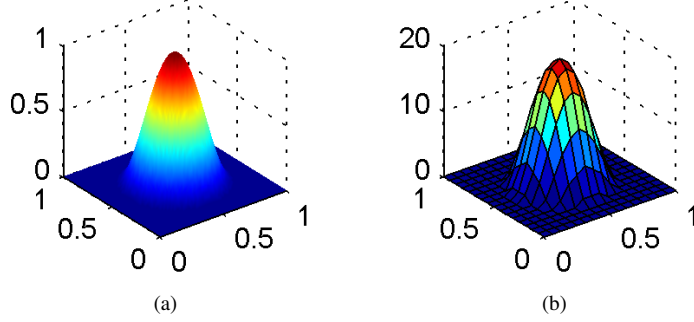


Figure 1: (a) Plot of the function $\theta(x, y)$; (b) Initial CTMC conditions from $\theta(x, y)$ for $K = 16$ and $N = 20$.

for all $1 \leq l \leq L$, $x \in [0; 1]$ and $y \in [0; 1]$. For example, in the validation study in Section 6 we will consider initial conditions based on a function, denoted by θ , which satisfies the DBCs and is defined as

$$\theta(x, y) = \begin{cases} 0 & , \|(x, y) - (\frac{1}{2}, \frac{1}{2})\|^2 \geq \frac{1}{4} \\ \frac{\exp(4)}{\exp(1/(\frac{1}{4} - \|(x, y) - (\frac{1}{2}, \frac{1}{2})\|^2))} & , \text{otherwise.} \end{cases} \quad (5)$$

This function is plotted in Figure 1a. To show how to obtain suitable initial conditions for the CTMC that yield a spatial ODE fluid limit, let us consider the case $K = 16$ and $N = 20$ and assume that the initial CTMC conditions for some l are given by $A_l^{(x,y)}(0) = \lfloor N\theta(x, y) \rfloor$, for all $(x, y) \in \mathcal{R}_K$. Then, these are represented by the values taken at the mesh visualized in Figure 1b. In the limit $K, N \rightarrow \infty$ the normalized initial conditions $\lfloor N\theta(x, y) \rfloor / N$ converge to $\theta(x, y)$.

The fluid limit for a mobile reaction network is obtained similarly to the stationary case. Let us denote $\vec{a} \in \mathbb{R}^{\mathcal{R}_K \times \{1, \dots, L\}}$ by $(a_1^{(x,y)}, \dots, a_L^{(x,y)})_{(x,y) \in \mathcal{R}_K}$. The corresponding system of ODEs is given by

$$\frac{d}{dt} a_l^{(x,y)}(t) = \mathbb{1}(x, y) \left(\mathcal{M}_l^{(x,y)}(\vec{a}(t)) + \mathcal{L}_l^{(x,y)}(\vec{a}(t)) \right), \quad (6)$$

where $(x, y) \in \mathcal{R}_K$, and, for all $1 \leq l \leq L$,

$$\begin{aligned} \mathcal{M}_l^{(x,y)}(\vec{a}) &= \mu_l^K \sum_{(\tilde{x}, \tilde{y}) \in \mathcal{N}(x,y)} (a_l^{(\tilde{x}, \tilde{y})} - a_l^{(x,y)}), \\ \mathcal{L}_l^{(x,y)}(\vec{a}) &= \sum_{1 \leq j \leq J} (d_{jl} - c_{jl}) f_j(a_1^{(x,y)}, \dots, a_L^{(x,y)}, \beta_1(x, y), \dots, \beta_P(x, y)). \end{aligned} \quad (7)$$

Essentially, $\mathcal{M}_l^{(x,y)}(\cdot)$ gives the *diffusive* dynamics of the system, i.e., the contribution to the vector field due to the movement of type- l nodes in each region (x, y) ; $\mathcal{L}_l^{(x,y)}(\cdot)$, instead, encodes the *reactive* dynamics, by considering the effects on each variable of the local interactions f_j . The latter essentially corresponds to the vector field of the corresponding stationary reaction network, i.e., compare $\mathcal{L}_l^{(x,y)}(\cdot)$ with $\mathcal{L}_l(\cdot)$ in (2). For instance, the ODEs of Example 3 in \mathcal{R}_K are

$$\begin{aligned} \frac{d}{dt} a_1^{(x,y)}(t) &= \mathbb{1}(x, y) \left(-\lambda a_1^{(x,y)}(t) + c \min\{a_2^{(x,y)}, 1\} + \mu_1^K \sum_{(\tilde{x}, \tilde{y}) \in \mathcal{N}(x,y)} (a_1^{(\tilde{x}, \tilde{y})} - a_1^{(x,y)}) \right), \\ \frac{d}{dt} a_2^{(x,y)}(t) &= \mathbb{1}(x, y) \left(+\lambda a_1^{(x,y)}(t) - c \min\{a_2^{(x,y)}, 1\} + \mu_2^K \sum_{(\tilde{x}, \tilde{y}) \in \mathcal{N}(x,y)} (a_2^{(\tilde{x}, \tilde{y})} - a_2^{(x,y)}) \right). \end{aligned} \quad (8)$$

In each ODE, the first two terms in the right-hand side correspond to those in (3). The second lines in each equation,

instead, describe the spatial behavior of the nodes.

The following result formalizes the convergence to the spatial ODE fluid limit.

Theorem 2 (ODE Fluid Limit of Spatial Networks). *The ODE system (6) subjected to the initial condition*

$$a_l^{(x,y)}(0) = \alpha_l^0(x,y), \quad (x,y) \in \mathcal{R}_K, \quad 1 \leq l \leq L,$$

has a unique solution \vec{a} in $\mathbb{R}^{\mathcal{R}_K \times \{1, \dots, L\}}$. Moreover, under the assumption that for an arbitrary but fixed $T > 0$ the time domain of \vec{a} contains $[0; T]$, it holds that

$$\lim_{N \rightarrow \infty} \mathbb{P} \left\{ \sup_{0 \leq t \leq T} \left\| \frac{1}{N} A_N(t) - \vec{a}(t) \right\|_{\infty} > \varepsilon \right\} = 0, \quad \forall \varepsilon > 0,$$

where the supremum norm $\|\cdot\|_{\infty}$ on $\mathbb{R}^{\mathcal{R}_K \times \{1, \dots, L\}}$ is given by $\|\vec{a}\|_{\infty} := \max_{(x,y) \in \mathcal{R}_K, 1 \leq l \leq L} |a_l^{(x,y)}|$.

This result can be recovered from Theorem 2.11 in [8], by a rewrite according to the notation of reaction networks adopted in this paper. However, this version represents a slight modification of that theorem in that we relax the assumption of *global* Lipschitz continuity to *local* Lipschitz continuity given by **(A2)**, by assuming that the time domain of the unique ODE solution contains $[0; T]$. It is restated in this way because it is this version that will be used in the proof of our main result, Theorem 5.

4.2. Reflective Boundary Conditions

Stochastic Model. The case of reflective boundaries differs from that of absorbing boundaries only on the boundary Ω_K . In particular, the migration rates are given by $q(\vec{A}; (\dots, A_l^{(x,y)} - 1, A_l^{(\tilde{x}, \tilde{y})} + 1)) = \mu_l^K A_l^{(x,y)}$, where $(x,y) \in \mathcal{R}_K$, $(\tilde{x}, \tilde{y}) \in \mathcal{N}(x,y)$ and $1 \leq l \leq L$. The reaction rates, instead, are

$$q(\vec{A}; (\dots, A_1^{(x,y)} + d_{j1} - c_{j1}, \dots, A_L^{(x,y)} + d_{jL} - c_{jL})) = \sum_{j' \in \mathcal{S}_j} F_{j'}(A_1^{(x,y)}, \dots, A_L^{(x,y)}, \beta_1(x,y), \dots, \beta_P(x,y)),$$

where $(x,y) \in \mathcal{R}_K$.

Spatial ODE Fluid Limit. The initial states of the underlying CTMC sequence $(\vec{A}_N(t))_{t \geq 0}$ are given by $A_l^{(x,y)}(0) = \lfloor N \alpha_l^0(x,y) \rfloor$, where $\alpha_1^0, \dots, \alpha_L^0 : [0; 1]^2 \rightarrow \mathbb{R}_{\geq 0}$ are now continuously differentiable functions whose derivatives are zero at the boundary, that is

$$0 = \partial_x \alpha_l^0(1, y) \quad 0 = \partial_x \alpha_l^0(0, y) \quad 0 = \partial_y \alpha_l^0(x, 1) \quad 0 = \partial_y \alpha_l^0(x, 0) \quad (9)$$

for all $x, y \in [0; 1]$ and $1 \leq l \leq L$. The conditions (9) are formally known as (zero) Neumann boundary conditions, short NBCs. It can be shown that (5) satisfies, apart from the DBCs (4), also the NBCs (9).

The fluid limit for a mobile reaction network is given by

$$\frac{d}{dt} a_l^{(x,y)}(t) = \mathcal{M}_l^{(x,y)}(\vec{a}(t)) + \mathcal{L}_l^{(x,y)}(\vec{a}(t)), \quad (10)$$

where $(x,y) \in \mathcal{R}_K$ and $\mathcal{M}_l^{(x,y)}, \mathcal{L}_l^{(x,y)}$ are as in (7). In particular, in the case of reflective boundary conditions, the ODEs of Example 3 in \mathcal{R}_K are

$$\begin{aligned} \frac{d}{dt} a_1^{(x,y)}(t) &= -\lambda a_1^{(x,y)}(t) + c \min\{a_2^{(x,y)}, 1\} + \mu_1^K \sum_{(\tilde{x}, \tilde{y}) \in \mathcal{N}(x,y)} (a_1^{(\tilde{x}, \tilde{y})} - a_1^{(x,y)}), \\ \frac{d}{dt} a_2^{(x,y)}(t) &= +\lambda a_1^{(x,y)}(t) - c \min\{a_2^{(x,y)}, 1\} + \mu_2^K \sum_{(\tilde{x}, \tilde{y}) \in \mathcal{N}(x,y)} (a_2^{(\tilde{x}, \tilde{y})} - a_2^{(x,y)}). \end{aligned}$$

The following result formalizes the convergence to the spatial ODE fluid limit.

Theorem 3 (ODE Fluid Limit of Spatial Networks). *The ODE system (10) subjected to the initial condition*

$$a_l^{(x,y)}(0) = \alpha_l^0(x,y), \quad (x,y) \in \mathcal{R}_K, \quad 1 \leq l \leq L,$$

has a unique solution \vec{a} in $\mathbb{R}^{\mathcal{R}_K \times \{1, \dots, L\}}$. Moreover, under the assumption that for an arbitrary but fixed $T > 0$ the time domain of \vec{a} contains $[0; T]$, it holds that

$$\lim_{N \rightarrow \infty} \mathbb{P} \left\{ \sup_{0 \leq t \leq T} \left\| \frac{1}{N} A_N(t) - \vec{a}(t) \right\|_{\infty} > \varepsilon \right\} = 0, \quad \forall \varepsilon > 0.$$

5. Spatial Reaction-Diffusion Limit

The spatial ODE fluid limit of Theorem 2 holds for any arbitrary but fixed K . However, for large K , the analysis may become infeasible because the ODE system size grows with $L(K+1)^2$. The purpose of this section is to consider a limit behavior that allows the analysis to be independent from K . To do so, we study a limit PDE of reaction-diffusion type.

Space Scaling. In addition to a suitable scaling of the population sizes with N , we require a proper scaling of the migration rates with K . Specifically, for all $1 \leq l \leq L$, it must hold that

$$\text{(A3): } \mu_l^K = \mu_l K^2.$$

Let us motivate this scaling of space by analyzing a continuous-time unbiased RW, denoted by $(W_k(t))_{t \geq 0}$, in two dimensions $\frac{1}{k}\mathbb{Z} \times \frac{1}{k}\mathbb{Z}$, with step size $1/k$ and sojourn time at each state that is exponentially distributed with mean $1/r_k$. The following result holds.

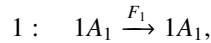
Proposition 1. *Let us assume that $W_k(0) = (0, 0)$ and denote by $d_k(t) := \|W_k(t)\|$ the Euclidian distance of $W_k(t)$ from the origin after time t . Then, it holds that $\mathbb{E}(d_k(t)^2) = \left(\frac{r_k}{k^2}\right)t$ for all $k \geq 1$ and $t \geq 0$.*

We notice that setting $r_k = k^2 r_1$ for all $k \geq 1$ implies that $\mathbb{E}(d_k(t)^2) = \mathbb{E}(d_1(t)^2)$ for all $k \geq 1$ and for all $t \geq 0$. This relation states that if the step size of the agent is $1/k$, then its migration rate should be equal to $k^2 r_1$, in order for the RW to always cover *the same distance* on average independently of k . Therefore, (A3) ensures that the average distance covered by a node across the unit square is invariant with respect to the lattice granularity.

5.1. Absorbing Boundary Conditions

Illustrating Example. The spatial PDE limit is obtained by a suitable manipulation of the diffusive dynamics of the ODE system. In order to give some intuition on this procedure, let us now consider an example of a purely diffusive population of nodes of the same type. Its CTMC population process is shown to converge to the celebrated *heat equation* (e.g, [43]).

Example 4. *A purely diffusive process can be modeled with a mobile reaction network with $L = 1$ and $J = 1$, where the interaction function is given by:*



for any $F_1 \equiv 0$.

The above example indeed gives a purely diffusive process, essentially because there are no stationary, local transitions. Let us point out that this somewhat degenerate mobile reaction network only serves an illustrative purpose.

Using the results presented in the previous section, for any K , the rescaled CTMC *density process* $(\frac{1}{N} \vec{A}_N(t))_{t \geq 0}$ for Example 4 converges in probability, as $N \rightarrow \infty$, to the solution of the system of $(K+1)^2$ ODEs

$$\frac{d}{dt} a_1^{(x,y)}(t) = \mathbf{1}(x,y) \mu_1^K \sum_{(\tilde{x}, \tilde{y}) \in \mathcal{N}(x,y)} (a_1^{(\tilde{x}, \tilde{y})}(t) - a_1^{(x,y)}(t)),$$

where $a_1^{(x,y)}(0) = \alpha_1^0(x,y)$ for all $(x,y) \in \mathcal{R}_K$, where $\alpha_1^0(x,y)$ is any function that satisfies the DBCs (4).

We continue by observing that

$$\frac{d}{dt} a_1^{(x,y)}(t) = \mathbb{1}(x,y) \frac{\mu_1^K}{K^2} \Delta^d a_1^{(x,y)}(t), \quad \Delta^d a_1^{(x,y)}(t) := \sum_{(\tilde{x}, \tilde{y}) \in \mathcal{N}(x,y)} \frac{a_1^{(\tilde{x}, \tilde{y})}(t) - a_1^{(x,y)}(t)}{(1/K)^2}, \quad (11)$$

where $(x,y) \in \mathcal{R}_K$ and $\Delta^d a_1^{(x,y)}$ refers to the discrete Laplace operator in a two dimensional lattice with neighboring distance $\Delta s = 1/K$, if $(x,y) \in \mathcal{R}_K \setminus \Omega_K$. In the case where $(x,y) \in \Omega_K$, instead, $\Delta^d a_1^{(x,y)}$ is not the discrete Laplace operator, since $|\mathcal{N}(x,y)| < 4$. Using **(A3)** by setting $\mu_1^K := \mu_1 K^2$, one can simplify (11) to the ODE system

$$\frac{d}{dt} a_1^{(x,y)}(t) = \mathbb{1}(x,y) \mu_1 \Delta^d a_1^{(x,y)}(t), \quad (x,y) \in \mathcal{R}_K. \quad (12)$$

The general result presented in the next section allows us to conclude that the solution to this ODE system converges, as $K \rightarrow \infty$, to that of the heat equation

$$\partial_t \alpha_1 = \mu_1 \Delta \alpha_1, \quad (x,y,t) \in [0;1]^2 \times \mathbb{R}_{\geq 0}, \quad (13)$$

where Δ denotes the continuous Laplace operator, subjected to the initial condition α_1^0 and the DBCs

$$\alpha_1(1,y,t) = 0, \quad \alpha_1(0,y,t) = 0, \quad \alpha_1(x,1,t) = 0, \quad \alpha_1(x,0,t) = 0$$

for all $t \geq 0$, $x \in [0;1]$ and $y \in [0;1]$.

Thus, we are dividing the question of establishing the convergence of the CTMCs $(\frac{1}{N} \vec{A}_N(t))_{t \geq 0}$ to the solution of the PDE (13) as $K, N \rightarrow \infty$ in two sub-problems where the two limits are studied separately: the first sub-problem involves convergence to a fluid limit with discrete regions, in the classical sense of Kurtz by sending N to infinity; after having established such a limit behavior, the second sub-problem studies the convergence of the spatial ODE limit (12) to the PDE (13), by sending K to infinity.

We wish to provide an insight into the proof technique by noticing that the latter issue only involves deterministic quantities. Hence, it is possible to use classical results from numerical analysis in order to study the convergence. To illustrate our strategy, we start by writing an approximating sequence, denoted by $(\vec{a}(m))_{m \geq 0}$, to solve the spatial ODE limit (12) with the Euler method using a fixed step size Δt . The iterations are given by

$$\begin{aligned} a_1^{(x,y)}(0) &:= \alpha_1^0(x,y), \\ a_1^{(x,y)}(m+1) &:= a_1^{(x,y)}(m) + \mathbb{1}(x,y) \Delta t (\mu_1 \Delta^d a_1^{(x,y)}(m)), \end{aligned} \quad (14)$$

with $(x,y) \in \mathcal{R}_K$ and $m \geq 0$. It is well-known that the sequence converges to the solution of the ODE system if $\Delta t \rightarrow 0$ [44].

The crucial observation is that this very sequence can be interpreted as a finite difference scheme (e.g., [45, 46]) for solving the PDE (13). This can be seen by discretizing the domain $[0;1]^2 \times \mathbb{R}_{\geq 0}$ into $\mathcal{R}_K \times \{m\Delta t \mid m \geq 0\}$ and observing that a sufficiently smooth solution of (13) satisfies

$$\frac{\alpha_1(x,y,(m+1)\Delta t) - \alpha_1(x,y,m\Delta t)}{\Delta t} = \mu_1 \sum_{(\tilde{x}, \tilde{y}) \in \mathcal{N}(x,y)} \frac{\alpha_1(\tilde{x}, \tilde{y}, m\Delta t) - \alpha_1(x,y,m\Delta t)}{\Delta s^2} + \mathcal{O}(\Delta t + \Delta s^2),$$

for all $(x,y) \in \mathcal{R}_K \setminus \Omega_K$, cf. [45]. As the DBCs enforce $\alpha_1(x,y,(m+1)\Delta t) = 0$ for all $(x,y) \in \Omega_K$, the PDE solution of (13) can be approximated by the finite difference scheme that computes the sequence $(\vec{a}(m))_{m \geq 0}$, where

$$\alpha_1^{(x,y)}(m+1) = \alpha_1^{(x,y)}(m) + \mathbb{1}(x,y) \Delta t \left(\mu_1 \sum_{(\tilde{x}, \tilde{y}) \in \mathcal{N}(x,y)} \frac{\alpha_1^{(\tilde{x}, \tilde{y})}(m) - \alpha_1^{(x,y)}(m)}{\Delta s^2} \right) = \alpha_1^{(x,y)}(m) + \mathbb{1}(x,y) \Delta t (\mu_1 \Delta^d \alpha_1^{(x,y)}(m)),$$

which corresponds to (14) as required.

Therefore, the desired result of convergence of the stochastic process to the PDE limit amounts to proving that such a finite difference scheme does converge to the solution of (13) as $\Delta t, \Delta s \rightarrow 0$. The main result of this paper, presented next, generalizes this convergence to a class of models where nodes also feature local interactions.

General PDE Limit. Let us now consider the ODE limit (6) for a generic mobile reaction network in \mathcal{R}_K . Observe that **(A3)** yields

$$\mathcal{M}_l^{(x,y)}(\vec{a}) = \mu_l \left(\sum_{(\bar{x}, \bar{y}) \in \mathcal{N}(x,y)} \frac{a_l^{(\bar{x}, \bar{y})} - a_l^{(x,y)}}{(1/K)^2} \right) = \mu_l \Delta^d a_l^{(x,y)}.$$

This allows us to rewrite the system of ODEs (6) as

$$\frac{d}{dt} a_l^{(x,y)}(t) = \mathbb{1}(x,y) \left(\mu_l \Delta^d a_l^{(x,y)}(t) + \mathcal{L}_l^{(x,y)}(\vec{a}(t)) \right),$$

where $a_l^{(x,y)}(0) = a_l^0(x,y)$ for $(x,y) \in \mathcal{R}_K$ and $1 \leq l \leq L$.

Similarly to (14), the approximating sequence $(\vec{a}(m))_{m \geq 0}$ arising from the Euler method with a fixed Δt to solve (6) is given by:

$$a_l^{(x,y)}(m+1) := a_l^{(x,y)}(m) + \mathbb{1}(x,y) \cdot \Delta t \cdot \left(\mu_l \Delta^d a_l^{(x,y)}(m) + \mathcal{L}_l^{(x,y)}(\vec{a}(m)) \right), \quad m \geq 0. \quad (15)$$

We now observe that this sequence corresponds to a finite difference scheme for the *reaction-diffusion* PDE system

$$\partial_t \alpha_l = \mu_l \Delta \alpha_l + \sum_{1 \leq j \leq J} (d_{jl} - c_{jl}) f_j(\alpha_1, \dots, \alpha_L, \beta_1, \dots, \beta_P), \quad (16)$$

where $(x,y,t) \in [0;1]^2 \times \mathbb{R}_{\geq 0}$ and $1 \leq l \leq L$, subjected to the DBCs

$$\alpha_l(1,y,t) = 0, \quad \alpha_l(0,y,t) = 0, \quad \alpha_l(x,1,t) = 0, \quad \alpha_l(x,0,t) = 0 \quad (17)$$

for all $1 \leq l \leq L$, $0 \leq t \leq T$, $x \in [0;1]$ and $y \in [0;1]$.

Therefore, in order to prove that the ODEs (6) converge to the PDEs (16), it remains to show that the approximating sequence (15) interpreted as a finite difference scheme for (16) is convergent with respect to the supremum norm. This is done in the following.

Theorem 4. *Assume that for a fixed $T > 0$, the family of functions $(\alpha_1, \dots, \alpha_L)$ on $[0;1]^2 \times [0;T]$ describes for all $0 < T' \leq T$ the unique solution of (16) subjected to the DBCs (17) on $[0;1]^2 \times [0;T']$. Further, assume that*

$$\partial_{txx} \alpha_l, \partial_{xxxx} \alpha_l, \partial_{yxx} \alpha_l, \partial_{tyy} \alpha_l, \partial_{xyy} \alpha_l, \partial_{yyyy} \alpha_l,$$

where $1 \leq l \leq L$, exist and are continuous. For $M \geq 1$, define $\Delta t := T/M$, $r_l := \mu_l \Delta t / \Delta s^2$ and assume that M is large enough such that $r_l \leq 1/4$ for all $1 \leq l \leq L$.

Then, there exist $K_0, M_0 \geq 1$ and $C > 0$ such that the Euler method (15) interpreted as a finite difference scheme of (16) subjected to the DBCs (17) satisfies

$$\max_{0 \leq m \leq M} \max_{(x,y) \in \mathcal{R}_K, 1 \leq l \leq L} |\alpha_l(x,y, m\Delta t) - a_l^{(x,y)}(m)| \leq C(\Delta s^2 + \Delta t)$$

for all $K \geq K_0$, $M \geq M_0$.

Theorem 2 and Theorem 4 can now be used to show that the CTMC sequence converges in probability to the solution of the PDE system (16) subjected to the DBCs (17) on $[0;1]^2 \times [0;T]$. This is stated in the following, which is the most important result of this paper.

Theorem 5 (Spatial PDE Limit). *Assume that for a fixed $T > 0$, the family of functions $(\alpha_1, \dots, \alpha_L)$ on $[0;1]^2 \times [0;T]$ describes for all $0 < T' \leq T$ the unique solution of (16) subjected to the DBCs (17) on $[0;1]^2 \times [0;T']$. Further,*

assume that

$$\partial_{lxx}\alpha_l, \partial_{xxxx}\alpha_l, \partial_{yxx}\alpha_l, \partial_{lyy}\alpha_l, \partial_{xyy}\alpha_l, \partial_{yyyy}\alpha_l,$$

where $1 \leq l \leq L$, exist and are continuous. Then, the time domain of the unique solution \vec{a} of (6) contains $[0; T]$ and for each $\varepsilon > 0$ it holds that

$$\lim_{K \rightarrow \infty} \lim_{N \rightarrow \infty} \mathbb{P} \left\{ \sup_{0 \leq t \leq T} \max_{(x,y) \in \mathcal{R}_K, 1 \leq l \leq L} \left| \frac{1}{N} A_l^{(x,y)}(t) - \alpha_l(x, y, t) \right| > \varepsilon \right\} = 0.$$

For instance, using **(A3)** the system (8) representing the spatial ODE limit for Example 3 can be simplified to

$$\begin{aligned} \frac{d}{dt} a_1^{(x,y)}(t) &= \mathbb{1}(x, y) (\mu_1 \Delta^d a_1^{(x,y)}(t) - \lambda a_1^{(x,y)}(t) + c \min\{a_2^{(x,y)}, 1\}), \\ \frac{d}{dt} a_2^{(x,y)}(t) &= \mathbb{1}(x, y) (\mu_2 \Delta^d a_2^{(x,y)}(t) + \lambda a_1^{(x,y)}(t) - c \min\{a_2^{(x,y)}, 1\}). \end{aligned} \quad (18)$$

Then, given that the assumptions of Theorem 4 are fulfilled for suitable initial concentrations a_1^0 and a_2^0 , the above ODE system converges, as $K \rightarrow \infty$, to the solution of the PDE system

$$\begin{aligned} \partial_t \alpha_1 &= -\lambda \alpha_1 + c \min\{\alpha_2, 1\} + \mu_1 \Delta \alpha_1, \\ \partial_t \alpha_2 &= +\lambda \alpha_1 - c \min\{\alpha_2, 1\} + \mu_2 \Delta \alpha_2. \end{aligned} \quad (19)$$

This implies, as stated by Theorem 5, that the sequence of normalized CTMCs converges to the solution of the PDE system when $K, N \rightarrow \infty$.

5.2. Reflective Boundary Conditions

Illustrating Example. Similarly to absorbing boundary conditions, we first convey our approach on the heat equation. As we will see, the presence of NBCs will ask for a different numerical treatment.

Using similar arguments as in the case of absorbing boundary conditions, we infer that the CTMC sequence $(\frac{1}{N} \vec{A}_N(t))_{t \geq 0}$ underlying Example 4 converges in probability, as $N \rightarrow \infty$, to the solution of the system of $(K + 1)^2$ ODEs

$$\frac{d}{dt} a_1^{(x,y)}(t) = \mu_1 \Delta^d a_1^{(x,y)}(t), \quad (x, y) \in \mathcal{R}_K, \quad (20)$$

where $a_1^{(x,y)}(0) = a_1^0(x, y)$ for all $(x, y) \in \mathcal{R}_K$. Crucially, $a_1^0(x, y)$ is now a function that satisfies the NBCs (9).

Further, we observe that the spatial ODE system (20) can be solved by means of the Euler method using a fixed step size Δt . The iterations are given by

$$\begin{aligned} a_1^{(x,y)}(0) &:= a_1^0(x, y), \\ a_1^{(x,y)}(m+1) &:= a_1^{(x,y)}(m) + \Delta t (\mu_1 \Delta^d a_1^{(x,y)}(m)), \end{aligned} \quad (21)$$

with $(x, y) \in \mathcal{R}_K$ and $m \geq 0$. As before, the crucial observation is that this very sequence can be interpreted as a finite difference scheme for solving the heat equation (13) which is now subject to the NBCs

$$0 = \partial_x \alpha_1(1, y, t) \quad 0 = \partial_x \alpha_1(0, y, t) \quad 0 = \partial_y \alpha_1(x, 1, t) \quad 0 = \partial_y \alpha_1(x, 0, t), \quad (22)$$

where $x, y \in [0; 1]$ and $0 \leq t \leq T$. To see this, we discretize the domain $[0; 1]^2 \times \mathbb{R}_{\geq 0}$ into $\mathcal{R}_K \times \{m\Delta t \mid m \geq 0\}$ and observe that a sufficiently smooth solution of the heat equation (13) satisfies

$$\frac{\alpha_1(x, y, (m+1)\Delta t) - \alpha_1(x, y, m\Delta t)}{\Delta t} = \mu_1 \sum_{(\tilde{x}, \tilde{y}) \in \mathcal{N}(x, y)} \frac{\alpha_1(\tilde{x}, \tilde{y}, m\Delta t) - \alpha_1(x, y, m\Delta t)}{\Delta s^2} + \mathcal{O}(\Delta t + \Delta s^2),$$

for all $(x, y) \in \mathcal{R}_K \setminus \Omega_K$. The treatment of the boundary Ω_K , instead, is done by means of so-called *ghost regions*

$(-\Delta s, j\Delta s), (1 + \Delta s, \Delta s), (i\Delta s, -\Delta s), (i\Delta s, 1 + \Delta s)$ where $0 \leq i, j \leq K$, cf. [45, Section 1.4]. These allow one to define the discrete Laplace operator also at Ω_K . More importantly, for a given ghost region (x_0, y_0) , the NBCs allow us to express $a_1^{(x_0, y_0)}(m)$ in terms of $\{a_1^{(x, y)}(m) \mid (x, y) \in \mathcal{R}_K\}$. For instance, the assumptions $\partial_x \alpha_1(0, 0, m\Delta t) = 0$ and $\partial_y \alpha_1(0, 0, m\Delta t) = 0$ yield the approximations

$$\frac{a_1^{(0,0)}(m) - a_1^{(-\Delta s, 0)}(m)}{\Delta s} = 0 \qquad \frac{a_1^{(0,0)}(m) - a_1^{(0, -\Delta s)}(m)}{\Delta s} = 0,$$

which imply, in turn, that

$$\begin{aligned} a_1^{(0,0)}(m+1) &= a_1^{(0,0)}(m) + \Delta t \mu_1 \left(\frac{a_1^{(-\Delta s, 0)}(m) - 2a_1^{(0,0)}(m) + a_1^{(\Delta s, 0)}(m)}{\Delta s^2} + \frac{a_1^{(0, -\Delta s)}(m) - 2a_1^{(0,0)}(m) + a_1^{(0, \Delta s)}(m)}{\Delta s^2} \right) \\ &= a_1^{(0,0)}(m) + \Delta t \left(\mu_1 \sum_{(\bar{x}, \bar{y}) \in ((\Delta s, 0), (0, \Delta s))} \frac{a_1^{(\bar{x}, \bar{y})}(m) - a_1^{(0,0)}(m)}{\Delta s^2} \right) \\ &= a_1^{(0,0)}(m) + \Delta t \mu_1 \Delta^d a_1^{(0,0)}. \end{aligned}$$

Thus, we recover the exact correspondence with (21) also at the corner $(0, 0)$. As a similar calculation can be shown to apply to all regions in Ω_K , we conclude that the Euler sequence (21) can be interpreted as a finite difference scheme of the heat equation (13) that is subject to NBCs (22).

General PDE Limit. We now consider the ODE limit (6) for a generic mobile reaction network in \mathcal{R}_K . Similarly to the case of absorbing boundary conditions, (6) can be rewritten into

$$\frac{d}{dt} a_l^{(x, y)}(t) = \mu_l \Delta^d a_l^{(x, y)}(t) + \mathcal{L}_l^{(x, y)}(\vec{a}(t)),$$

where $a_l^{(x, y)}(0) = \alpha_l^0(x, y)$ for $(x, y) \in \mathcal{R}_K$ and $1 \leq l \leq L$. The approximating sequence $(\vec{a}(m))_{m \geq 0}$ arising from the Euler method with a fixed Δt to solve (6) is given by

$$a_l^{(x, y)}(m+1) := a_l^{(x, y)}(m) + \Delta t \left(\mu_l \Delta^d a_l^{(x, y)}(m) + \mathcal{L}_l^{(x, y)}(\vec{a}(m)) \right), \quad m \geq 0. \quad (23)$$

Using the concept of ghost regions, this sequence corresponds to a finite difference scheme of the *reaction-diffusion* PDE system (16) that is subject to the NBCs

$$0 = \partial_x \alpha_l(1, y, t) \qquad 0 = \partial_x \alpha_l(0, y, t) \qquad 0 = \partial_y \alpha_l(x, 1, t) \qquad 0 = \partial_y \alpha_l(x, 0, t), \quad (24)$$

for all $1 \leq l \leq L$, $0 \leq t \leq T$, $x \in [0; 1]$ and $y \in [0; 1]$.

Unfortunately, the proof of Theorem 4 does not carry over if reflective boundary conditions are assumed, meaning that we cannot establish the convergence of the finite difference scheme (23). However, we are able to prove that (23) is stable and consistent [45, Section 2.5 and 6.2] with respect to a certain norm. Informally, this means that the scheme arises from a Taylor approximation (consistency) and gives rise to an error that does not explode (stability).

Theorem 6. *Assume that for a fixed $T > 0$, the family of functions $(\alpha_1, \dots, \alpha_L)$ on $[0; 1]^2 \times [0; T]$ describes for all $0 < T' \leq T$ the unique solution of (16) subjected to the NBCs (24) on $[0; 1]^2 \times [0; T']$. Further, assume that*

$$\partial_{txx} \alpha_l, \partial_{xxxx} \alpha_l, \partial_{yxx} \alpha_l, \partial_{lyy} \alpha_l, \partial_{xyy} \alpha_l, \partial_{yyyy} \alpha_l,$$

where $1 \leq l \leq L$, exist and are continuous. For $M \geq 1$, define $\Delta t := T/M$, $r_l := \mu_l \Delta t / \Delta s^2$ and assume that M is large enough such that $r_l \leq 1/4$ for all $1 \leq l \leq L$. Then, the Euler method (23) interpreted as a finite difference scheme of (16) is stable and consistent with respect to the ‘‘averaging’’ norm $\|\vec{a}\|_{1, \Delta s} := \sum_{(x, y) \in \mathcal{R}_K, 1 \leq l \leq L} |a_l^{(x, y)}| \Delta s^2$.

For linear PDEs, that is, if the reactive terms may be written in the form $f_j(z_1, \dots, z_L) = \sum_{l=1}^L \xi_{j,l} z_l$, for $\xi_{j,l} \in \mathbb{R}$, Theorem 6 and Lax’s Theorem [45, Section 2.5 and 6.2] assert convergence.

Corollary 1. *If the PDE system (16) is linear, then it is also well-posed [46, Section 6.2]. Let us further assume that the unique solution on $[0; 1]^2 \times [0; T]$ is such that*

$$\partial_{lxx}\alpha_l, \partial_{lxxx}\alpha_l, \partial_{lyxx}\alpha_l, \partial_{lyy}\alpha_l, \partial_{xyy}\alpha_l, \partial_{yyyy}\alpha_l,$$

exist and are continuous for all $1 \leq l \leq L$.

Then, there exist $K_0, M_0 \geq 0$ and $C > 0$ such that the Euler method (23) interpreted as a finite difference scheme of (16) subjected to the NBCs (24) satisfies

$$\max_{0 \leq m \leq M} \frac{1}{K^2} \sum_{\substack{(x,y) \in \mathcal{R}_K \\ 1 \leq l \leq L}} |\alpha_l(x, y, m\Delta t) - a_l^{(x,y)}(m)| \leq C(\Delta s + \Delta t)$$

for all $K \geq K_0$ and $M \geq M_0$, where $\Delta t := T/M$.

Theorem 2 and Corollary 1 can now be used to show that the CTMC sequence converges in probability to the solution of the PDE system (16) subjected to the NBCs (24) on $[0; 1]^2 \times [0; T]$. Note that $\|\cdot\|_{1,\Delta s} \leq 4L\|\cdot\|_\infty$, meaning that convergence in the supremum norm implies convergence in the averaging norm.

Theorem 7 (Spatial PDE Limit). *If the PDE system (16) is linear, then it is also well-posed [46, Section 6.2] and the finite difference scheme (23) satisfies*

$$\lim_{K \rightarrow \infty} \lim_{N \rightarrow \infty} \mathbb{P} \left\{ \sup_{0 \leq t \leq T} \frac{1}{K^2} \sum_{\substack{(x,y) \in \mathcal{R}_K \\ 1 \leq l \leq L}} \left| \frac{1}{N} A_l^{(x,y)}(t) - \alpha_l(x, y, t) \right| > \varepsilon \right\} = 0$$

provided that $\partial_{lxx}\alpha_l, \partial_{lxxx}\alpha_l, \partial_{lyxx}\alpha_l, \partial_{lyy}\alpha_l, \partial_{xyy}\alpha_l, \partial_{yyyy}\alpha_l$, where $1 \leq l \leq L$, exist and are continuous on $[0; 1]^2 \times [0; T]$.

6. Case Study

The purpose of this section is to show how to use our PDE limit result for the performance modeling and analysis of a communication network with mobile nodes. In the following, we study in detail the mobile version of Example 3 with its PDE limit (19) using absorbing boundary conditions.

System Description. Our mobile network consists of a lattice of $(K + 1)^2$ regions where each region represents an area offering Internet connectivity to all nodes therein located by means of an 802.11 access point. We assume that nodes feature an *on/off* behavior whereby they interpose an exponentially distributed think time between successive connections. Whilst pausing, nodes may decide to move across one of the neighboring regions according to an unbiased RW. When connected, a node downloads a file with an exponentially distributed size, and does not move until the download has finished. We assume that the overall network's bottleneck is represented by the link between the node and the access point, therefore the performance depends on the current load at each region and on how the capacity is shared amongst the nodes in a region.

This system was implemented in the JiST framework, using the SWANS module for the modeling of wireless and ad-hoc networks [21]. The physical layer was configured using default values for all parameters. For simplicity, we used a model of transmission with no packet loss. Because of the scenario considered in this study, nodes were not equipped with a dynamically updating routing protocol because there was no direct communication between nodes, and the destination of a packet could be determined based on the node location, by maintaining a simple mapping between the current location and the MAC address of the access point. The application layer was implemented as a lightweight file transfer protocol on top of UDP. The file size to download was exponentially distributed with mean 40 KB, and the file was transmitted in 2 KB packets. We assume that the think time and the RW behavior during the *off* state of the nodes are given by independent distributions. Let λ be the think rate and μK^2 be the movement rate to each neighboring region. Thus, the holding time in the *off* state is exponentially distributed with mean $1/(\lambda + 4\mu K^2)$.

V	$Nodes$	$\mu = 0.001$			$\mu = 0.010$		
		$JiST$	$CTMC$ (Error)	PDE (Error)	$JiST$	$CTMC$ (Error)	PDE (Error)
10	56	0.8852	0.8638 (0.0214)	0.9007 (0.0155)	0.5161	0.5368 (0.0207)	0.2716 (0.2445)
25	156	0.6968	0.6939 (0.0029)	0.7186 (0.0218)	0.5207	0.5480 (0.0273)	0.3807 (0.1400)
50	320	0.4734	0.4578 (0.0156)	0.4646 (0.0088)	0.3926	0.3742 (0.0184)	0.2840 (0.1086)

Table 1: Validation results: Average fraction of nodes in the *off* state after 10 time units in an area divided in 64 regions ($K = 7$), for $N = 1$. The errors are given as the absolute difference with respect to the JiST estimated average.

N	K	$V = 10$			$V = 25$			$V = 50$		
		$Nodes$	$CTMC$	$Err. PDE$	$Nodes$	$CTMC$	$Err. PDE$	$Nodes$	$CTMC$	$Err. PDE$
1	7	56	0.5368	0.2652	156	0.5480	0.1673	320	0.3742	0.0902
2	7	124	0.5073	0.2350	320	0.5500	0.1693	660	0.4172	0.1332
1	15	268	0.3897	0.1181	716	0.4611	0.0804	1496	0.3397	0.0557
2	15	546	0.3897	0.1181	1496	0.4658	0.0851	3028	0.3415	0.0575
1	31	1132	0.3424	0.0708	3100	0.4139	0.0332	6338	0.3102	0.0262
2	31	2248	0.3341	0.0625	6388	0.4256	0.0449	12948	0.3061	0.0221
1	47	2592	0.3238	0.0522	7108	0.3984	0.0177	14628	0.3035	0.0195
2	47	5572	0.3096	0.0380	14628	0.4046	0.0239	29736	0.3011	0.0171
1	63	4676	0.3049	0.0335	12736	0.3893	0.0009	26296	0.2928	0.0088
2	63	10084	0.3016	0.0300	29296	0.3943	0.0136	53480	0.2973	0.0133
1	79	7268	0.3072	0.0356	19996	0.3813	0.0006	41392	0.2907	0.0067
2	79	15736	0.2943	0.0227	41312	0.3877	0.0070	84084	0.2910	0.0070

Table 2: Scaling behavior for $\mu = 0.010$. The PDE approximation error is the absolute difference with respect to the CTMC estimate.

Stochastic Model. To model this system, we use the mobile reaction network with interaction functions given in Example 3, where we let $\mu_1^K = \mu K^2$, for some $\mu > 0$ which will be varied in our numerical experiments, and $\mu_2^K = 0$. The scaling of μ_1^K satisfies assumption (A3), while the choice of μ_2^K models the fact that the nodes stand still until the download has finished. With this, we wish to highlight that the unbiased RW can indeed be made dependent upon the local state of the node; analogous results to those presented in this section would be obtained with $\mu_2^K > 0$, i.e., allowing the node to transfer to another region even during the download process. By varying μ in our experiments, instead, we wish to show the impact of the mobility model on the network performance.

Remarks. Let us observe that, in essence, our stochastic model is a *spatial* version of a queueing network with exponentially distributed service times. In each region (x, y) , $A_1^{(x,y)}$ represents the queue length at a delay station whereas $A_2^{(x,y)}$ is the queue length at a multi-server station with N servers with individual rate c . A job serviced at this station goes to the delay station in the same region with probability 1, whereas a job at the delay station may go into service in the same region, or into a neighboring delay station according to the movement rate μ_k^1 . The knowledge of the queue lengths in each region, given by the PDE solution, therefore provides a complete characterization of the system's performance.

Due to the assumption on the boundary conditions and to the forms of F_1 and F_2 , all nodes will eventually leave the spatial domain with probability 1. Although we make use of this scenario in the present paper, we wish to remark that our framework also supports exogenous arrivals. For this, however, the arrival rate must scale linearly with N . For instance, interaction 1 in Example 2 describes the arrival of A_1 -type nodes into each region with rate $N\lambda$. This models Poisson arrivals with intensity proportional to some parameter function $\lambda(x, y)$ that must be consistent with the DBCs.

Model Validation. We tested the validity of our stochastic model against the reference behavior as given by discrete-event simulation of the JiST implementation. The performance metric of interest was chosen to be the ratio between the expected total number of nodes in state A_1 and the initial population of nodes after 10 time units. The CTMC was solved by simulation, since closed-form expressions for the transient behavior are not available and numerical CTMC solution was made infeasible by the excessively large state spaces. In both cases, simulations were conducted using the method of independent replicas; the stopping criterion was the convergence of the 95% confidence interval within 5% of the mean. Throughout the remainder of this section, we report such means. Finally, the PDEs were solved using Matlab and its *Partial Differential Equation Toolbox* with the function `parabolic`, which implements the numerical solution of systems of reaction-diffusion PDE with both NBCs and DBCs.

The JiST implementation and the CTMC model were parametrized as follows. Throughout all tests, we kept λ fixed at 0.250. We estimated the parameter $c = 2.857$ by measuring the download times in a network with a single node, that is, in the absence of contention. The initial population of nodes across the spatial domain was generated according to the function θ given in (5). The initial concentrations of nodes were set as

$$\alpha_1^0(x, y) = 0 \quad \text{and} \quad \alpha_2^0 = V\theta(x, y),$$

where V is a parameter that was varied in our tests to observe the network under different node densities. Thus, the N -th element of the CTMC sequence had an initial number of downloading nodes in region (x, y) equal to $\lfloor NV\theta(x, y) \rfloor$.

Results. Table 1 shows the results of the comparison between the average fraction of nodes in the *off* state after 10 time units as computed by the JiST discrete-event simulation, CTMC simulation, and PDE analysis, for different values of V and μ , and for $N = 1$ and $K = 7$ (corresponding to an area with 64 regions). As can be expected, our parametrizations exercise the network under a range of different operating regimes. This can be noticed by the fact that our performance index varies sensibly with V in Table 1. For any fixed V , the estimates for the different values of μ numerically show the impact of the mobility model on the network performance. We remark that the observation window of interest, 10 time units, is in all cases sufficient to study the network under interesting conditions where the fraction of nodes in the *off* state is significantly different than 0 (which corresponds to the initial condition).

The table shows good agreement, in all cases, between the JiST estimates and the CTMC model, according to the notion of error defined as the absolute difference with respect to the JiST average. The PDE approximation, instead, yielded different quality depending on the choice of μ . Indeed, it already provides good accuracy for $\mu = 0.001$, whereas it suffered more severe errors for $\mu = 0.010$. This is, however, unsurprising because the PDE limit holds for large N and K . To numerically show convergence to the PDE limit, we considered all cases of V with $\mu = 0.010$ and applied the space scaling by simulating the CTMCs for different values of K until 79, thus dividing the space in at most 6400 regions. The numerical results are presented in Table 2, where the PDE approximation error is measured as the absolute difference with respect to the corresponding average estimate computed by CTMC simulation. It is interesting to note that, for a fixed K , the PDE approximation error may tend to increase with N , see, for instance, $K = 31$ for $V = 25$. However, this is not in contradiction with our limit result which holds when both N and K are large. Indeed, in all cases, increasing K leads to sensibly increased accuracy, with errors that are negligible for all practical purposes for $K = 79$, corresponding to a situation with at most ca 84,000 nodes in the network. There, we also verified that the means for $N = 1$ and $N = 2$ were statistically equal at the 0.05 level.

ODE/PDE Comparison. As discussed, the main advantage in using the PDE limits instead of the the ODE limits is that the mesh discretization becomes a parameter of the solver, instead of being a parameter of the model (i.e., the actual number of regions in the spatial domain). An ODE/PDE comparison is proposed in Figure 2, where we consider the absolute error between the numerical ODE and PDE solutions of the fraction of agents in the *off* state after 10.0 time units, together with their runtimes. This was done by fixing $V = 10$ and varying $\lambda = 0.25, 1.25, 2.25$. To study the quality of the numerical PDE limit, we applied Matlab's PDE solver on a conservative mesh consisting of 5408 triangles and an excessively fine PDE mesh containing 338944 triangles. Our findings can be summarized as follows.

- The distance between the numerical ODE and PDE solutions decreases, for bigger values of K , if the PDE mesh is refined. Similarly, with a PDE mesh being fixed, the error stabilizes or decreases with increasing K values. This confirms our main result: both, the numerical PDE and numerical ODE solutions converge to a common object, namely the analytical PDE solution.

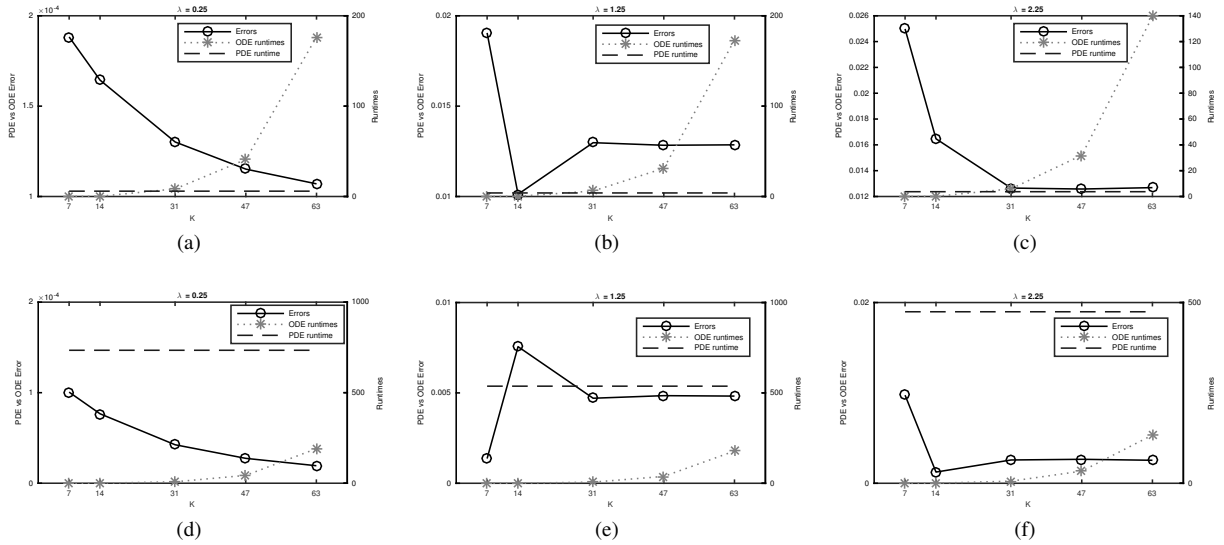


Figure 2: Comparison of ODE/PDE estimation errors and runtimes for different PDE meshes and varying values of K and λ . Subfigures (a)-(c) under a common PDE mesh, while (d)-(f) arise from an excessively fine PDE mesh (5408 versus 338944 triangles, respectively); the left y axis shows the absolute error between the ODE solution and the PDE solution of the fraction of agents in the *off* state after 10.0 time units; the right y axis shows the runtimes (in seconds) of the ODE solution with varying K relative to the PDE runtime (dashed line).

- The ODE limit and the PDE limit are not comparable for small values of K , as can be seen by comparing Subfigures (b) and (e) of Figure 2. Hence, ODE limits should be used whenever the spatial model has a coarse granularity, while the PDE limit is appropriate whenever the spatial domain enjoys a fine granularity or discretizes a continuous domain.
- In the case of the excessively refined PDE mesh, the PDE running times exceed the ones of ODEs. However, by noting that even an excessive refinement of the PDE mesh does not affect the numerical PDE solution substantially, we conclude that in practice it suffices to consider PDE meshes of moderate size. This, in turn, allows for efficient computation.

We wish to stress that in the case of the conservative mesh, all experiments considered in this paper were analyzed within 8 s on an ordinary machine (consistently with [26], where a similar analysis was conducted on epidemiological models). This turned out to be significantly cheaper from a computational point of view than directly solving the system of $2(K + 1)^2$ ODEs (with Matlab's `ode15s` function). In particular, for $K = 63$ we registered ODE runtimes two orders of magnitude larger than the PDE solution times, whilst yielding accurate results in all cases (up to at most ca 1%). Clearly, instead, CTMC simulations required a sheer amount of computational power — for instance, we did measure runtimes of ca 32 hours for the CTMC simulation for the cases $N = 2$ and $K = 79$, with independent replicas being run in parallel on eight cores.

7. Conclusion

In this paper we have proposed an approach to analyze stochastic models of mobile networks by means of a deterministic approximation represented by a system of partial differential equations. These offer a macroscopic, continuous view of the network dynamics which holds in the limit of infinite populations of nodes in the network and infinite number of regions in a regular two-dimensional spatial lattice. In practice, in our numerical tests we observed good accuracy in the case of moderate lattice granularity and population size by carrying out a validation study against Markov-chain simulations and detailed discrete-event simulations at the network-protocol level. Furthermore, we found good speed of convergence with increasing system sizes, yielding high accuracy for networks consisting of

thousands of nodes. These observations allow us to conclude that partial differential equations are appropriate for the analysis of large-scale mobile systems.

Acknowledgement

This work was partially supported by the EU project QUANTICOL, 600708.

- [1] J. Hillston, Fluid flow approximation of PEPA models, in: QEST, 2005, pp. 33–43. doi:10.1109/QEST.2005.12. URL <http://dx.doi.org/10.1109/QEST.2005.12>
- [2] B. V. Houdt, A mean field model for a class of garbage collection algorithms in flash-based solid state drives, in: SIGMETRICS, 2013, pp. 191–202. doi:10.1145/2465529.2465543.
- [3] W. Minnebo, B. V. Houdt, A fair comparison of pull and push strategies in large distributed networks, *IEEE/ACM Trans. Netw.* 22 (3) (2014) 996–1006. doi:10.1109/TNET.2013.2270445.
- [4] M. Kowal, M. Tschaikowski, M. Tribastone, I. Schaefer, Scaling Size and Parameter Spaces in Variability-Aware Software Performance Models, in: 30th IEEE/ACM International Conference on Automated Software Engineering, ASE 2015, 2015, pp. 407–417. doi:10.1109/ASE.2015.16.
- [5] L. Cardelli, M. Tribastone, M. Tschaikowski, A. Vandin, Forward and backward bisimulations for chemical reaction networks, in: CONCUR, 2015.
- [6] M. Tschaikowski, M. Tribastone, Approximate reduction of heterogenous nonlinear models with differential hulls, *IEEE Trans. Automat. Contr.* 61 (4) (2016) 1099–1104. doi:10.1109/TAC.2015.2457172.
- [7] L. Cardelli, M. Tribastone, M. Tschaikowski, A. Vandin, Symbolic computation of differential equivalences, in: POPL, 2016.
- [8] T. G. Kurtz, Solutions of ordinary differential equations as limits of pure Markov processes, *J. Appl. Prob.* 7 (1) (1970) 49–58.
- [9] G. Sharma, A. Ganesh, P. Key, Performance analysis of contention based medium access control protocols, *IEEE Trans. Inf. Theory* 55 (4) (2009) 1665–1682.
- [10] N. Gast, B. Gaujal, A Mean Field Model of Work Stealing in Large-Scale Systems, in: SIGMETRICS, ACM, 2010.
- [11] B. Van Houdt, L. Bortolussi, Fluid limit of an asynchronous optical packet switch with shared per link full range wavelength conversion, in: SIGMETRICS, 2012.
- [12] J. Anselmi, I. Verloop, Energy-aware capacity scaling in virtualized environments with performance guarantees, *Performance Evaluation* 68 (11) (2011) 1207–1221.
- [13] X. Zhou, S. Ioannidis, L. Massoulie, On the stability and optimality of universal swarms, in: SIGMETRICS, ACM, 2011.
- [14] M. Benaïm, J.-Y. Le Boudec, A class of mean field interaction models for computer and communication systems, *Performance Evaluation* 65 (11-12) (2008) 823–838.
- [15] X. Zhang, G. Neglia, J. Kurose, D. Towsley, Performance modeling of epidemic routing, *Computer Networks* 51 (10) (2007) 2867–2891. doi:DOI: 10.1016/j.comnet.2006.11.028.
- [16] J. Wu, O. L. Loucks, From balance of Nature to hierarchical patch dynamics: A paradigm shift in ecology, *The Quarterly Review of Biology* 70 (4) (1995) 4390–466.
- [17] A. Chaintreau, J.-Y. Le Boudec, N. Ristanovic, The age of gossip: spatial mean field regime, in: SIGMETRICS/Performance, 2009.
- [18] D. Gillespie, Exact stochastic simulation of coupled chemical reactions, *Journal of Physical Chemistry* 81 (25) (1977) 2340–2361.
- [19] I. Akyildiz, Y.-B. Lin, W.-R. Lai, R.-J. Chen, A new random walk model for PCS networks, *IEEE J. Sel. Areas Commun.* 18 (7) (2000) 1254–1260.
- [20] I. F. Akyildiz, J. S. M. Ho, Y.-B. Lin, Movement-based location update and selective paging for PCS networks, *IEEE/ACM Trans. Netw.* 4 (4) (1996) 629–638.
- [21] R. Barr, Z. J. Haas, R. van Renesse, JiST: an efficient approach to simulation using virtual machines, *Softw. Pract. Exper.* 35 (6) (2005) 539–576.
- [22] J. D. Murray, *Mathematical Biology I: An Introduction*, 3rd Edition, Springer, 2002.
- [23] A. Okubo, S. A. Levin, *Diffusion and ecological problems: modern perspectives*, Springer, 2001.
- [24] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd Edition, Elsevier, 2007.
- [25] S. Isaacson, The reaction-diffusion master equation as an asymptotic approximation of diffusion to a small target, *SIAM Journal on Applied Mathematics* 70 (1) (2009) 77–111.
- [26] M. Tschaikowski, M. Tribastone, A Partial-differential Approximation for Spatial Stochastic Process Algebra, in: VALUETOOLS, 2014.
- [27] L. Bortolussi, M. Tschaikowski, Fluid Analysis of Spatio-Temporal Properties of Agents in a Population Model, in: ASMTA 2016, 2016, pp. 92–106.
- [28] R. Darling, J. Norris, Differential equation approximations for markov chains., *Probability Surveys [electronic only]* 5 (2008) 37–79.
- [29] L. Bortolussi, J. Hillston, Fluid model checking, *CONCUR’12*, Springer, 2012.
- [30] L. Arnold, M. Theodosopulu, Deterministic limit of the stochastic model of chemical reactions with diffusion, *Advances in Applied Probability* 12 (2) (1980) 367–379.
- [31] P. Kotelenez, Law of large numbers and central limit theorem for linear chemical reactions with diffusion, *The Annals of Probability* 14 (1) (1986) 173–193.
- [32] D. Blount, Comparison of stochastic and deterministic models of a linear chemical reaction with diffusion, *The Annals of Probability* 19 (4) (1991) 1440–1462.
- [33] D. Blount, Limit theorems for a sequence of nonlinear reaction-diffusion systems, *Stochastic Processes and their Applications* 45 (2) (1993) 193–207.
- [34] D. Blount, Diffusion limits for a nonlinear density dependent space-time population model, *The Annals of Probability* 24 (2) (1996) 639–659.

- [35] T. Camp, J. Boleng, V. Davies, A survey of mobility models for ad hoc network research, *Wireless Communications and Mobile Computing* 2 (5) (2002) 483–502.
- [36] A. Bar-Noy, I. Kessler, M. Sidi, Mobile users: To update or not to update?, in: *INFOCOM*, 1994, pp. 570–576.
- [37] S. Ioannidis, P. Marbach, A brownian motion model for last encounter routing, in: *INFOCOM*, 2006.
- [38] M. Grossglauser, M. Vetterli, Locating mobile nodes with EASE: learning efficient routes from encounter histories alone, *IEEE/ACM Trans. Netw.* 14 (3) (2006) 457–469.
- [39] A. Gamal, J. Mammen, B. Prabhakar, D. Shah, Throughput-delay trade-off in wireless networks, in: *INFOCOM*, 2004.
- [40] M. Durvy, P. Thiran, Reaction-diffusion based transmission patterns for ad hoc networks, in: *INFOCOM*, Vol. 3, 2005, pp. 2195–2205.
- [41] M. Garetto, E. Leonardi, Analysis of random mobility models with partial differential equations, *IEEE Transactions on Mobile Computing* 6 (11) (2007) 1204–1217.
- [42] D. Qiu, R. Srikant, Modeling and performance analysis of BitTorrent-like peer-to-peer networks, in: *SIGCOMM*, 2004.
- [43] A. Okubo, *Diffusion and Ecological Problems: Mathematical Models*, Springer-Verlag, 1980.
- [44] C. W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice Hall PTR, 1971.
- [45] J. W. Thomas, *Numerical Partial Differential Equations: Finite Difference Methods*, Springer-Verlag, 1995.
- [46] J. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, Wadsworth & Brooks, Pacific Grove, California, 1989.
- [47] G. H. Weiss, *Aspects and Applications of the Random Walk*, North-Holland Publishing Co., Amsterdam, 1994.
- [48] P. Brémaud, *Markov chains, Gibbs fields, Monte Carlo simulation, and queues*, Springer-Verlag, 1999.

Appendix A. Proofs

Proof of Theorem 2. Let us denote by E_N the state space of $(\frac{1}{N}\vec{A}_N(t))_{t \geq 0}$ subjected to $\frac{1}{N}A_l^{(x,y)}(0) = \lfloor N\alpha_l^0(x,y) \rfloor / N$. For an $\vec{a} \in E_N$, let $\lambda_N(\vec{a})$ denote the waiting time distribution in state \vec{a} , meaning that

$$\mathbb{P}\{\tau(\vec{a}) > t \mid \frac{1}{N}\vec{A}_N(0) = \vec{a}\} = e^{-\lambda_N(\vec{a})t}, \quad \text{where } \tau(\vec{a}) := \inf\{t \geq 0 \mid \frac{1}{N}\vec{A}_N(t) \neq \vec{a}\}.$$

For every set Γ in the Borel σ -algebra of E_N we further define the jump distribution in \vec{a} as

$$\mu_N(\vec{a}, \Gamma) := \mathbb{P}\left\{\frac{1}{N}\vec{A}_N(\tau(\vec{a})) \in \Gamma \mid \frac{1}{N}\vec{A}_N(0) = \vec{a}\right\}.$$

Then, the drift $D_N(\vec{a})$ of $(\frac{1}{N}\vec{A}_N(t))_{t \geq 0}$ in the state $\vec{a} \in E_N$ can be written in the form

$$\begin{aligned} D_N(\vec{a}) &= \lambda_N(\vec{a}) \int_{\vec{b} \in E_N} (\vec{b} - \vec{a}) \mu_N(\vec{a}, d\vec{b}) = \int_{\vec{b} \in E_N} (\vec{b} - \vec{a}) \lambda_N(\vec{a}) \mu_N(\vec{a}, d\vec{b}) \\ &= \sum_{(x,y)} \left[\sum_{l, (\tilde{x}, \tilde{y})} \frac{w_l^{(x,y), (\tilde{x}, \tilde{y})}}{N} \mu_l^K(N\vec{a}_l^{(x,y)}) + \sum_j \frac{w_j^{(x,y)}}{N} \cdot F_j(N\vec{a}_1^{(x,y)}, \dots, N\vec{a}_L^{(x,y)}, \beta_1(x,y), \dots, \beta_P(x,y)) \right], \end{aligned}$$

where $w_j^{(x,y)}$ and $w_l^{(x,y), (\tilde{x}, \tilde{y})}$ denote the corresponding jump vectors of $(\vec{A}_N(t))_{t \geq 0}$. For instance, $w_l^{(\Delta s, \Delta s), (2\Delta s, \Delta s)}$ decrements $A_l^{(\Delta s, \Delta s)}$ and increments $A_l^{(2\Delta s, \Delta s)}$, whereas $w_l^{(\Delta s, \Delta s), (0, \Delta s)}$ just decrements $A_l^{(\Delta s, \Delta s)}$. Also, $w_j^{(x,y)} = 0 = w_l^{(x,y), (\tilde{x}, \tilde{y})}$ whenever $(x,y) \in \Omega$. Note that

$$\left(\sum_{(x,y)} \left[\sum_{l, (\tilde{x}, \tilde{y})} w_l^{(x,y), (\tilde{x}, \tilde{y})} \mu_l^K(\vec{a}_l^{(x,y)}) + \sum_j w_j^{(x,y)} f_j(\vec{a}_1^{(x,y)}, \dots, \vec{a}_L^{(x,y)}, \beta_1(x,y), \dots, \beta_P(x,y)) \right] \right)_l^{(x,y)} = \mathbf{1}(x,y) \cdot (\mathcal{M}_l^{(x,y)}(\vec{a}) + \mathcal{L}_l^{(x,y)}(\vec{a}))$$

for all $\vec{a} \in E_N$, $1 \leq l \leq L$ and $(x,y) \in \mathcal{R}$. Moreover, since the theorem of Picard-Lindelöf and **(A2)** assert that (6) has a unique solution \vec{a} in $\mathbb{R}^{\mathcal{R} \times \{1, \dots, L\}}$ and the time domain of \vec{a} contains $[0; T]$ by assumption, it holds that $c := \sup_{0 \leq t \leq T} \|\vec{a}(t)\| + 1 < \infty$. Thanks to the compactness of $E := [-c; c]^{\mathcal{R} \times \{1, \dots, L\}}$ and **(A1)**, the above implies

$$\lim_{N \rightarrow \infty} \sup_{\vec{a} \in \Gamma_N} \left\| D_N(\vec{a}) - \sum_{(x,y)} \left[\sum_{l, (\tilde{x}, \tilde{y})} w_l^{(x,y), (\tilde{x}, \tilde{y})} \mu_l^K(\vec{a}_l^{(x,y)}) + \sum_j w_j^{(x,y)} f_j(\vec{a}_1^{(x,y)}, \dots, \vec{a}_L^{(x,y)}, \beta_1(x,y), \dots, \beta_P(x,y)) \right] \right\| = 0,$$

where $\Gamma_N := E_N \cap E$. Moreover, it holds that

$$E_N \cap \left\{ \vec{b} \in \mathbb{R}^{\mathcal{R} \times (1, \dots, L)} \mid \inf_{0 \leq t \leq T} \|\vec{b} - \vec{a}(t)\| \leq \frac{1}{2} \right\} \subseteq E \quad \text{and} \quad \sup_{N \geq 1} \sup_{\vec{a} \in \Gamma_N} \lambda_N(\vec{a}) \int_{\vec{b} \in E_N} \|\vec{b} - \vec{a}\| \mu_N(\vec{a}, d\vec{b}) < \infty.$$

To see this, we note that

$$\begin{aligned} \lambda_N(\vec{a}) \int_{\vec{b} \in E_N} \|\vec{b} - \vec{a}\| \mu_N(\vec{a}, d\vec{b}) &\leq \sqrt{(K+1)^2 L} \int_{\vec{b} \in E_N} \|\vec{b} - \vec{a}\| \lambda_N(\vec{a}) \mu_N(\vec{a}, d\vec{b}) \\ &\leq (K+1) \sqrt{L} \sum_{(x,y)} \left[\sum_j \frac{C}{N} F_j(N \underline{a}_1^{(x,y)}, \dots, N \underline{a}_L^{(x,y)}, \vec{\beta}(x,y)) + \sum_l \frac{1}{N} 4\mu_l^K N \underline{a}_l^{(x,y)} \right], \end{aligned}$$

where $C := \max \{|d_{jl} - c_{jl}| \mid 1 \leq j \leq J \wedge 1 \leq l \leq L\}$. Moreover, for all $N \geq 1$, it holds that

$$\sup_{\vec{a} \in \Gamma_N} \lambda_N(\vec{a}) \int_{\|\vec{b} - \vec{a}\| > \varepsilon_N} \|\vec{b} - \vec{a}\| \mu_N(\vec{a}, d\vec{b}) = 0, \quad \text{where} \quad \varepsilon_N := \frac{(C+1)(K+1)\sqrt{L}}{N},$$

because $\varepsilon_N < \|\vec{b} - \vec{a}\| \leq \sqrt{(K+1)^2 L} \|\vec{b} - \vec{a}\|_\infty$ implies $\frac{C+1}{N} < \|\vec{b} - \vec{a}\|_\infty$. This allows us to conclude the claim using Theorem 2.11 in [8]. \square

Proof of Theorem 3. Similarly to the proof of Theorem 2. \square

Proof of Proposition 1. Using the concept of characteristic function, it is possible to establish an analytic expression of the Fourier-Laplace transform of the transient probability distribution of a continuous-time random walk (CTRW) in the case of general holding times. This, in turn, characterizes $\mathbb{E}(d_k^2(t))$ in terms of the underlying Laplace transform, see [47, Section 2.5]. However, in the case of exponentially distributed holding times, it is possible to provide a direct proof that relies on the uniformization method for CTMCs with countable state spaces [48, Chapter 8, Section 2.1]. For this, let the CTRW $(W_k(t))_{t \geq 0}$ be given by means of the standard graphical notation in Figure A.3 and let $(\widehat{W}_k(n))_{n \geq 0}$ denote the unbiased random walk in *discrete* time on $\frac{1}{k}\mathbb{Z} \times \frac{1}{k}\mathbb{Z}$ with $\widehat{W}_k(0) = (0, 0)$. Further, let $(N(t))_{t \geq 0}$ be the homogenous Poisson process with intensity r_k and P denote the transition matrix of $(\widehat{W}_k(n))_{n \geq 0}$. Then it holds that $W_k(t) = \widehat{W}_k(N(t))$ and

$$\begin{aligned} \mathbb{E}(d_k^2(t)) &= \sum_{(x,y) \in \mathbb{Z}^2} \mathbb{P}\{W_k(t) = \left(\frac{x}{k}, \frac{y}{k}\right)\} \left[\left(\frac{x}{k}\right)^2 + \left(\frac{y}{k}\right)^2 \right] = \frac{1}{k^2} \sum_{(x,y) \in \mathbb{Z}^2} \left(\sum_{n=0}^{\infty} \vec{e}_{\left(\frac{x}{k}, \frac{y}{k}\right)}^T P^n \frac{(r_k t)^n}{n!} e^{-r_k t} \right) \vec{e}_{\left(\frac{x}{k}, \frac{y}{k}\right)}(x^2 + y^2) \\ &= \frac{1}{k^2} e^{-r_k t} \sum_{n=0}^{\infty} \frac{(r_k t)^n}{n!} \left(\sum_{(x,y) \in \mathbb{Z}^2} \vec{e}_{\left(\frac{x}{k}, \frac{y}{k}\right)}^T P^n \vec{e}_{\left(\frac{x}{k}, \frac{y}{k}\right)}(x^2 + y^2) \right) = \frac{1}{k^2} e^{-r_k t} \sum_{n=0}^{\infty} \frac{(r_k t)^n}{n!} \mathbb{E}(\|\widehat{W}_k(n)\|^2) \\ &= \frac{1}{k^2} e^{-r_k t} \sum_{n=0}^{\infty} \frac{(r_k t)^n}{n!} n = \frac{1}{k^2} e^{-r_k t} r_k t \sum_{n=1}^{\infty} \frac{(r_k t)^{n-1}}{(n-1)!} = \left(\frac{r_k}{k^2}\right) t. \end{aligned}$$

\square

Proposition 2. Let us fix an $f : \mathbb{R}^n \rightarrow \mathbb{R}$ which is locally Lipschitz and some $c > 0$. Then, the function

$$\underline{f}(x) := \begin{cases} f(x) & , \|x\| \leq c \\ f\left(c \frac{x}{\|x\|}\right) & , \text{otherwise} \end{cases}$$

is globally Lipschitz on \mathbb{R}^n .

Proof. As the function f is locally Lipschitz on \mathbb{R}^n and the set $\{z \in \mathbb{R}^n \mid \|z\| \leq c\}$ is compact, there exists a $\Lambda > 0$ such that

$$|f(y) - f(x)| \leq \Lambda \|y - x\|$$

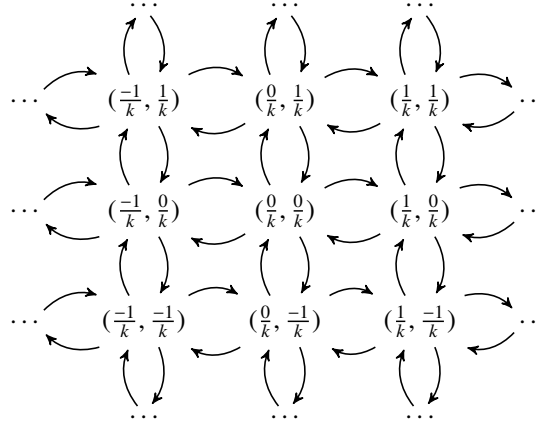


Figure A.3: Continuous-time random walk $(W_k(t))_{t \geq 0}$ on the lattice $\frac{1}{k}\mathbb{Z} \times \frac{1}{k}\mathbb{Z}$. The rate of each transition is $r_k/4$ and was suppressed to enhance readability. In particular, the average sojourn time in each state is $1/r_k$.

for all $x, y \in \{z \in \mathbb{R}^n \mid \|z\| \leq c\}$. In the following, we show that

$$|\underline{f}(y) - \underline{f}(x)| \leq \Lambda \|y - x\|$$

for all $x, y \in \mathbb{R}^n$. Let us fix for this two arbitrary $x, y \in \mathbb{R}^n$.

We note that the cases $(x = 0 \wedge y = 0)$, $(x = 0 \wedge y \neq 0)$ and $(x \neq 0 \wedge y = 0)$ are obvious. For the case $x \neq 0 \wedge y \neq 0$, we may assume without loss of generality that $\|x\| \leq \|y\|$. Then, the case can be divided into the subcases $(\|x\| < c \wedge \|y\| < c)$, $(\|x\| \geq c \wedge \|y\| \geq c)$ and $(0 < \|x\| < c \leq \|y\|)$. The first subcase is clear, as $\underline{f} \equiv f$ on $\{x \in \mathbb{R}^n \mid \|x\| \leq c\}$.

Let us consider now the case $(\|x\| \geq c \wedge \|y\| \geq c)$. Specifically, we assume first that y and x are linearly dependent, i.e. $y = \lambda x$ for some $\lambda \in \mathbb{R}$. As $\underline{f}(y) = \underline{f}(x)$ in the case of $\lambda \geq 0$, we may assume that $\lambda < 0$. Then, as

$$|\underline{f}(y) - \underline{f}(x)| = \left| f\left(\frac{c\lambda x}{|\lambda|\|x\|}\right) - f\left(\frac{cx}{\|x\|}\right) \right| \leq \Lambda \left\| -\frac{cx}{\|x\|} - \frac{cx}{\|x\|} \right\| = \Lambda \frac{c}{\|x\|} 2\|x\| = \Lambda 2c,$$

the estimation $\|y - x\| = (1 + |\lambda|)\|x\| = \|x\| + \|y\| \geq 2c$ yields the claim.

Let us assume now that x and y are linearly independent. Then, the calculation

$$|\underline{f}(y) - \underline{f}(x)| = \left| f\left(\frac{cy}{\|y\|}\right) - f\left(\frac{cx}{\|x\|}\right) \right| \leq \Lambda \left\| \frac{cy}{\|y\|} - \frac{cx}{\|x\|} \right\| = \Lambda \frac{c}{\|x\|} \left\| \frac{\|x\|}{\|y\|} y - x \right\| \leq \Lambda \left\| \frac{\|x\|}{\|y\|} y - x \right\|$$

shows that it is sufficient to prove that $\|p - x\| \leq \|y - x\|$, where $p := \frac{\|x\|}{\|y\|}y$. An informal pictorial description of the situation in the two-dimensional space is given in Figure A.4. The calculation

$$\begin{aligned} \|y - x\|^2 &= \|y - p + p - x\|^2 = \langle y - p + p - x, y - p + p - x \rangle = \\ &= \langle y - p, y - p \rangle + 2\langle y - p, p - x \rangle + \langle p - x, p - x \rangle = \|y - p\|^2 + \|x - p\|^2 - 2\langle y - p, x - p \rangle \end{aligned}$$

shows then that $\|x - p\| \leq \|y - x\|$ if $\langle y - p, x - p \rangle \leq 0$. (Note that the above calculation establishes nothing more than the general version of the cosine law.) In order to prove this, we introduce the auxiliary points $y' := \frac{\|y\|}{\|x\|}x$ and

$$p' := y + \left\langle \frac{y' - y}{\|y' - y\|}, p - y \right\rangle \frac{y' - y}{\|y' - y\|}.$$

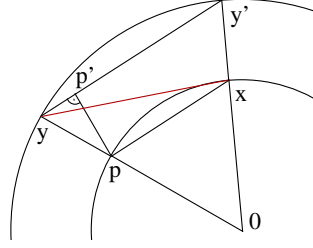


Figure A.4: Graphical illustration in the two-dimensional case for the proof idea of Proposition 2 to show the Lipschitz continuity of \underline{f} in the case of linearly independent vectors x and y .

Note that the linear independence of x and y implies $y' - y \neq 0$. Hence, p' is well-defined and we conclude

$$\langle y - p, x - p \rangle = \langle y - p' + p' - p, x - p \rangle = \langle y - p', x - p \rangle + \langle p' - p, x - p \rangle.$$

Consequently, it is sufficient to show that $\langle y - p', x - p \rangle \leq 0$ and $\langle p' - p, x - p \rangle = 0$. The latter follows from

$$\begin{aligned} \frac{\|y\|}{\|x\|} \langle p' - p, x - p \rangle &= \left\langle p' - p, \frac{\|y\|}{\|x\|} (x - p) \right\rangle = \langle p' - p, y' - y \rangle = \left\langle y + \left\langle \frac{y' - y}{\|y' - y\|}, p - y \right\rangle \frac{y' - y}{\|y' - y\|} - p, y' - y \right\rangle \\ &= \langle y - p, y' - y \rangle + \left\langle \frac{y' - y}{\|y' - y\|}, p - y \right\rangle \left\langle \frac{y' - y}{\|y' - y\|}, y' - y \right\rangle \\ &= \langle y - p, y' - y \rangle + \langle y' - y, p - y \rangle \frac{\langle y' - y, y' - y \rangle}{\|y' - y\|^2} \\ &= \langle y - p, y' - y \rangle - \langle y - p, y' - y \rangle = 0, \end{aligned}$$

whereas the calculation

$$\begin{aligned} \langle y - p', x - p \rangle &= \left\langle - \left\langle \frac{y' - y}{\|y' - y\|}, p - y \right\rangle \frac{y' - y}{\|y' - y\|}, x - p \right\rangle = \left\langle \frac{y' - y}{\|y' - y\|}, y - p \right\rangle \left\langle \frac{y' - y}{\|y' - y\|}, x - p \right\rangle \\ &= \frac{1}{\|y' - y\|^2} \langle y' - y, y - p \rangle \langle y' - y, x - p \rangle \end{aligned}$$

shows that $\langle y - p', x - p \rangle \leq 0$ in the case of $\langle y' - y, y - p \rangle \leq 0$ and $\langle y' - y, x - p \rangle \geq 0$. This can be shown using the Cauchy-Schwarz inequality, as

$$\begin{aligned} \langle y' - y, y - p \rangle &= \left\langle \frac{\|y\|}{\|x\|} x - y, y - \frac{\|x\|}{\|y\|} y \right\rangle = \frac{\|y\|}{\|x\|} \langle x, y \rangle - \langle x, y \rangle - \|y\|^2 + \frac{\|x\|}{\|y\|} \|y\|^2 = \left(\frac{\|y\|}{\|x\|} - 1 \right) \langle x, y \rangle - \|y\|^2 + \|x\| \|y\| \\ &\leq \left(\frac{\|y\|}{\|x\|} - 1 \right) |\langle x, y \rangle| - \|y\|^2 + \|x\| \|y\| \leq \left(\frac{\|y\|}{\|x\|} - 1 \right) \|x\| \|y\| - \|y\|^2 + \|x\| \|y\| = 0 \end{aligned}$$

and

$$\begin{aligned} \langle y' - y, x - p \rangle &= \left\langle \frac{\|y\|}{\|x\|} x - y, x - \frac{\|x\|}{\|y\|} y \right\rangle = \frac{\|y\|}{\|x\|} \|x\|^2 - \langle x, y \rangle - \langle y, x \rangle + \frac{\|x\|}{\|y\|} \|y\|^2 \\ &= \|y\| \|x\| - 2\langle x, y \rangle + \|x\| \|y\| = 2(\|x\| \|y\| - \langle x, y \rangle) \geq 0. \end{aligned}$$

We consider next the case ($0 < \|x\| < c \leq \|y\|$). Let us assume first that y and x are linearly dependent, i.e. $y = \lambda x$

for some $\lambda \in \mathbb{R}$. Then it holds that

$$|\underline{f}(y) - \underline{f}(x)| = \left| f\left(\frac{c\lambda x}{|\lambda||x|}\right) - f(x) \right| \leq \Lambda \left\| x\left(\frac{\lambda c}{|\lambda||x|} - 1\right) \right\| = \Lambda \left| \frac{\lambda c}{|\lambda||x|} - 1 \right| \|x\| = \begin{cases} \Lambda(c - \|x\|) & , \lambda \geq 0 \\ \Lambda(c + \|x\|) & , \lambda < 0 \end{cases}$$

Thus, if $\lambda \geq 0$, $\|x\| < \|y\|$ implies $\lambda > 1$ and we conclude that

$$|\underline{f}(y) - \underline{f}(x)| \leq \Lambda(c - \|x\|) \leq \Lambda\|y - x\|,$$

as $\|y - x\| = (\lambda - 1)\|x\| = \|y\| - \|x\| \geq c - \|x\|$. Instead, if $\lambda < 0$, it holds that $\|y - x\| = |\lambda - 1|\|x\| = (|\lambda| + 1)\|x\| = \|y\| + \|x\| \geq c + \|x\|$, which implies in turn

$$|\underline{f}(y) - \underline{f}(x)| \leq \Lambda(c + \|x\|) \leq \Lambda\|y - x\|.$$

Let us assume now that x and y are linearly independent. Then it holds that

$$|\underline{f}(y) - \underline{f}(x)| = \left| f\left(\frac{cy}{\|y\|}\right) - f(x) \right| \leq \Lambda \left\| \frac{cy}{\|y\|} - x \right\|.$$

As $g : [0; 1] \rightarrow \mathbb{R}_{\geq 0}$, $t \mapsto \|(1-t)x + ty\|$ is continuous, the set $\mathcal{O} := \{t \in [0; 1] \mid g(t) < c\} = g^{-1}([0; c])$ is open in $[0; 1]$ and $\lambda := \sup \mathcal{O} \notin \mathcal{O}$. Obviously, $\|g(t)\| \geq c$ for all $\lambda \leq t \leq 1$ and the continuity of g yields $g(\lambda) = c$. (Moreover, the convexity of $\{z \in \mathbb{R}^n \mid \|z\| < c\}$ asserts that $\|g(t)\| < c$ for all $0 \leq t < \lambda$.) Together with $x_0 := (1-\lambda)x + \lambda y$, it holds then that

$$\left\| \frac{cy}{\|y\|} - x \right\| \leq \left\| \frac{cy}{\|y\|} - x_0 \right\| + \|x_0 - x\| = \left\| \frac{\|x_0\|}{\|y\|} y - x_0 \right\| + \|x_0 - x\|$$

Note that the last subcase asserts that $\left\| \frac{\|x_0\|}{\|y\|} y - x_0 \right\| \leq \|y - x_0\|$ if y and x_0 are linearly independent. Let us assume towards a contradiction that y and x_0 are linearly dependent. Then there exists a $\mu \neq 0$ such that $y = \mu x_0 = \mu(1-\lambda)x + \mu\lambda y$, which implies in turn $(1-\mu\lambda)y = \mu(1-\lambda)x$. As $\lambda < 1$, this induces $x = \frac{1-\mu\lambda}{\mu(1-\lambda)}y$ and $x \neq 0$ yields the contradiction $y = \frac{\mu(1-\lambda)}{1-\mu\lambda}x$. Hence, it holds that $\left\| \frac{\|x_0\|}{\|y\|} y - x_0 \right\| \leq \|y - x_0\|$ and it is sufficient to show that $\|y - x_0\| + \|x_0 - x\| = \|y - x\|$. For this, we note that

$$x_0 - x = -\lambda x + \lambda y = \lambda(y - x) \quad \text{and} \quad y - x_0 = y - (1-\lambda)x - \lambda y = (1-\lambda)(y - x)$$

imply together with

$$\begin{aligned} \|y - x\|^2 &= \langle y - x_0 + x_0 - x, y - x_0 + x_0 - x \rangle = \|y - x_0\|^2 + \|x_0 - x\|^2 + 2\langle y - x_0, x_0 - x \rangle \\ &= \|y - x_0\|^2 + \|x_0 - x\|^2 + 2\lambda(1-\lambda)\langle y - x, y - x \rangle = \|y - x_0\|^2 + \|x_0 - x\|^2 + 2\|y - x_0\|\|x_0 - x\| \\ &= (\|y - x_0\| + \|x_0 - x\|)^2 \end{aligned}$$

the desired relation. □

Proof of Theorem 4. The proof borrows ideas from convergence proofs of the Euler method [44] and the heat equation [45], respectively. Let us define

$$c := \max_{(x,y,t) \in [0;1]^2 \times [0;T]} \max_{1 \leq l \leq L} \max_{1 \leq p \leq P} \max \left\{ |\alpha_l(x, y, t)|, |\beta_p(x, y)| \right\}, \quad c' := \sqrt{(L+P)(c+1)^2},$$

and $\underline{f}_j : \mathbb{R}^{L+P} \rightarrow \mathbb{R}$, where $1 \leq j \leq J$, by

$$\underline{f}_j(\vec{z}) := \begin{cases} f_j(\vec{z}) & , \|\vec{z}\| \leq c' \\ f_j(c' \frac{\vec{z}}{\|\vec{z}\|}) & , \text{otherwise.} \end{cases}$$

Since $\|(z_1, \dots, z_{L+P})\|_\infty \leq c + 1$ implies that $\|(z_1, \dots, z_{L+P})\| \leq c'$, we infer that $\underline{f}_j(z_1, \dots, z_{L+P}) = f_j(z_1, \dots, z_{L+P})$ for

all $\|(z_1, \dots, z_{L+P})\|_\infty \leq c + 1$. This shows that the solution of (16) on $[0; 1]^2 \times [0; T]$ is also a solution of the PDE system

$$\partial_t \alpha_l = \mu_l \Delta \alpha_l + \sum_{1 \leq j \leq J} (d_{jl} - c_{jl}) \underline{f}_j(\alpha_1, \dots, \alpha_L, \beta_1, \dots, \beta_P), \quad (\text{A.1})$$

where $1 \leq l \leq L$, subjected to the DBCs.

Consequently, as it can be shown that (A.1) has a unique solution on $[0; 1]^2 \times [0; T]$, the solution of (16) on $[0; 1]^2 \times [0; T]$ is the solution of (A.1) on $[0; 1]^2 \times [0; T]$. To see this, let us assume towards a contradiction that $(\hat{\alpha}_1, \dots, \hat{\alpha}_L)$ is a solution of (A.1) satisfying $\xi(t) > 0$ for some t , where $\xi : [0; T] \rightarrow \mathbb{R}_{\geq 0}$ is given by $\xi(t) := \max_{(x,y) \in [0;1]^2} \max_{1 \leq l \leq L} |\alpha_l(x, y, t) - \hat{\alpha}_l(x, y, t)|$. As the function $\zeta : (x, y, t) \mapsto \max_{1 \leq l \leq L} |\alpha_l(x, y, t) - \hat{\alpha}_l(x, y, t)|$ is continuous on the compactum $[0; 1]^2 \times [0; T]$, it is uniformly continuous on $[0; 1]^2 \times [0; T]$ and thus it satisfies

$$\forall \varepsilon > 0 \exists \delta > 0 \forall (x, y) \in [0; 1]^2 \forall t, t' \in [0; T] (|t - t'| < \delta \Rightarrow |\zeta(x, y, t) - \zeta(x, y, t')| < \varepsilon). \quad (\text{A.2})$$

Since $\forall (x, y) \in [0; 1]^2 (|\zeta(x, y, t)| - \varepsilon < |\zeta(x, y, t')| < |\zeta(x, y, t)| + \varepsilon)$ implies $\xi(t) - \varepsilon < \xi(t') < \xi(t) + \varepsilon$, (A.2) yields the (uniform) continuity of ξ on $[0; T]$. Consequently, the set $\mathcal{O} := \{t \in [0; T] \mid \xi(t) > 0\} = \xi^{-1}((0; \infty))$ is open in $[0; T]$ and $t_0 := \inf \mathcal{O}$, which exists because $\mathcal{O} \neq \emptyset$, is not contained in \mathcal{O} . Also, since the solutions are continuous in t , it holds that $t_0 < T$. Since ξ is continuous, there exists a $\delta > 0$ such that $|\xi(t')| = |\xi(t') - \xi(t_0)| < 1$ for all $t' \in [t_0; t_0 + \delta] \subseteq [0; T]$ and we get

$$\forall 1 \leq l \leq L \forall t_0 \leq t' \leq t_0 + \delta \forall (x, y) \in [0; 1]^2 (|\hat{\alpha}_l(x, y, t')| \leq |\hat{\alpha}_l(x, y, t') - \alpha_l(x, y, t')| + |\alpha_l(x, y, t')| < 1 + c).$$

Consequently, $(\hat{\alpha}_1, \dots, \hat{\alpha}_L)$ is a solution of (16) on $[0; 1]^2 \times [0; t_0 + \delta]$. The choice of t_0 ensures that $\xi(t') > 0$ for some $t_0 < t' \leq t_0 + \delta$. This is a contradiction to the uniqueness of the solution of (16) on $[0; 1]^2 \times [0; t_0 + \delta]$ and we infer that the solution of (16) on $[0; 1]^2 \times [0; T]$ can be recovered by solving (A.1).

We show next that the finite difference scheme of (A.1)

$$\begin{aligned} \underline{\mathcal{L}}^{(x,y)}(\underline{a}) &:= \sum_{1 \leq j \leq J} (d_{jl} - c_{jl}) \underline{f}_j(\underline{a}_1^{(x,y)}, \dots, \underline{a}_L^{(x,y)}, \beta_1(x, y), \dots, \beta_P(x, y)), \quad \underline{a} \in \mathbb{R}^{\mathcal{R} \times \{1, \dots, L\}} \\ \underline{a}_l^{(x,y)}(m+1) &:= \underline{a}_l^{(x,y)}(m) + \mathbb{1}(x, y) \cdot \Delta t \cdot (\mu_l \Delta^d \underline{a}_l^{(x,y)}(m) + \underline{\mathcal{L}}^{(x,y)}(\underline{a}(m))), \quad m \geq 0, \end{aligned} \quad (\text{A.3})$$

is convergent in the supremum norm $\|\underline{a}(m)\|_\infty := \max_{(x,y) \in \mathcal{R}, 1 \leq l \leq L} |\underline{a}_l^{(x,y)}(m)|$ on $\mathbb{R}^{\mathcal{R} \times \{1, \dots, L\}}$. For this, we observe first that each \underline{f}_j is Lipschitz with respect to $\|\cdot\|$ by Proposition 2. This and the equivalence of all norms on a finite dimensional vector space show that each \underline{f}_j is also Lipschitz with respect to $\|\cdot\|_\infty$. Let us denote the space step error and the time step error of the scheme at $(x, y, t) \in \mathcal{R} \times [0; T - \Delta t]$ by

$$\begin{aligned} \rho_l(x, y, t) &:= \mu_l \Delta \alpha_l(x, y, t) - \mu_l \Delta^d \alpha_l(x, y, t) \\ \eta_l(x, y, t) &:= \alpha_l(x, y, t + \Delta t) - \alpha_l(x, y, t) - \Delta t (\mu_l \Delta \alpha_l(x, y, t) + g_l(x, y, t)), \end{aligned}$$

respectively, where $g_l : [0; 1]^2 \times [0; T] \rightarrow \mathbb{R}$ is defined as $g_l := \sum_{1 \leq j \leq J} (d_{jl} - c_{jl}) \underline{f}_j(\alpha_1, \dots, \alpha_L, \beta_1, \dots, \beta_P)$.

We begin with bounding the space step error. For this, we fix a region $(x, y) \in \mathcal{R} \setminus \Omega$ and note that there exist $x' \in (x - \Delta s; x + \Delta s)$ and $y' \in (y - \Delta s; y + \Delta s)$ such that

$$\begin{aligned} \frac{\rho_l(x, y, t)}{\mu_l} &= \partial_{xx} \alpha_l(x, y, t) + \partial_{yy} \alpha_l(x, y, t) \\ &\quad - \left(\frac{\alpha_l(x + \Delta s, y, t) - 2\alpha_l(x, y, t) + \alpha_l(x - \Delta s, y, t)}{\Delta s^2} + \frac{\alpha_l(x, y + \Delta s, t) - 2\alpha_l(x, y, t) + \alpha_l(x, y - \Delta s, t)}{\Delta s^2} \right) \\ &= \frac{\partial_{xxx} \alpha_l(x', y, t)}{12} \Delta s^2 + \frac{\partial_{yyy} \alpha_l(x, y', t)}{12} \Delta s^2. \end{aligned}$$

Consequently, there exists a $C_1 \in \mathbb{R}_{\geq 0}$ which *does not depend* on $\Delta t, \Delta s$ and satisfies $|\rho_l(x, y, t)| \leq C_1 \Delta s^2$ for all $(x, y) \in \mathcal{R} \setminus \Omega, 0 \leq t \leq T$ and $1 \leq l \leq L$.

We turn now to the time step error. Applying the mean value theorem to $t \mapsto \alpha_l(x, y, t)$ allows us to infer the existence of a $t \leq t' \leq t + \Delta t$ such that $\alpha_l(x, y, t + \Delta t) - \alpha_l(x, y, t) = \partial_t \alpha_l(x, y, t') \Delta t$. Hence

$$\begin{aligned} \eta_l(x, y, t) &= \alpha_l(x, y, t + \Delta t) - \alpha_l(x, y, t) - \Delta t(\mu_l \Delta \alpha_l(x, y, t) + g_l(x, y, t)) = \Delta t(\partial_t \alpha_l(x, y, t') - \mu_l \Delta \alpha_l(x, y, t) - g_l(x, y, t)) \\ &= \Delta t(\mu_l \Delta \alpha_l(x, y, t') + g_l(x, y, t') - \mu_l \Delta \alpha_l(x, y, t) - g_l(x, y, t)). \end{aligned}$$

Let $\Lambda' > 0$ refer to the Lipschitz constant of $\mu_1 \Delta \alpha_1, \dots, \mu_L \Delta \alpha_L$ on $[0; 1]^2 \times [0; T]$ and $\Lambda'' > 0$ be the *global* Lipschitz constant of $\underline{f}_1, \dots, \underline{f}_J$. Then, together with $\Lambda := \max_{1 \leq l \leq L} \sum_{1 \leq j \leq J} |d_{jl} - c_{jl}| \Lambda''$, the above calculation yields

$$\begin{aligned} \frac{|\eta_l(x, y, t)|}{\Delta t} &\leq \sum_{1 \leq j \leq J} |d_{jl} - c_{jl}| \Lambda'' \|\vec{\alpha}(x, y, t') - \vec{\alpha}(x, y, t)\|_{\infty} + \Lambda' \|(x, y, t') - (x, y, t)\|_{\infty} \\ &\leq \Lambda \|\vec{\alpha}(x, y, t') - \vec{\alpha}(x, y, t)\|_{\infty} + \Lambda' \Delta t \leq \Lambda \Delta t \max_{1 \leq l' \leq L} \max_{(x, y, t) \in [0; 1]^2 \times [0; T]} |\partial_t \alpha_{l'}(x, y, t)| + \Lambda' \Delta t, \end{aligned}$$

where the last inequality follows again by means of the mean value theorem. Hence, there exists a $C_2 \in \mathbb{R}_{\geq 0}$ which *does not depend* on $\Delta t, \Delta s$ and satisfies $|\eta_l(x, y, t)| \leq C_2 \Delta t^2$ for all $(x, y) \in \mathcal{R}, 0 \leq t \leq T - \Delta t$ and $1 \leq l \leq L$.

After we were able to derive bounds for $|\eta_l|$ and $|\rho_l|$, we note that

$$\eta_l(x, y, t) = \alpha_l(x, y, t + \Delta t) - \alpha_l(x, y, t) - \Delta t(\mu_l \Delta^d \alpha_l(x, y, t) + \rho_l(x, y, t) + g_l(x, y, t))$$

is equivalent to

$$\alpha_l(x, y, t + \Delta t) = \alpha_l(x, y, t) + \Delta t(\mu_l \Delta^d \alpha_l(x, y, t) + g_l(x, y, t)) + \Delta t \rho_l(x, y, t) + \eta_l(x, y, t).$$

Consequently, together with $t_m := m \Delta t, e_l^{(x, y)}(m) := \alpha_l(x, y, t_m) - \underline{a}_l^{(x, y)}(m)$ and

$$h_l(x, y, m) := \sum_{1 \leq j \leq J} (d_{jl} - c_{jl}) \underline{f}_j(\underline{a}_1^{(x, y)}(m), \dots, \underline{a}_L^{(x, y)}(m), \beta_1(x, y), \dots, \beta_P(x, y)),$$

it holds that

$$\begin{aligned} |e_l^{(x, y)}(m+1)| &\leq \left| (1 - |\mathcal{N}(x, y)| r_l) e_l^{(x, y)}(m) + r_l \sum_{(\bar{x}, \bar{y}) \in \mathcal{N}(x, y)} e_l^{(\bar{x}, \bar{y})}(m) + \Delta t (g_l(x, y, t_m) - h_l(x, y, m)) \right| + \\ &\quad + \Delta t |\rho_l(x, y, t_m)| + |\eta_l(x, y, t_m)| \leq \|\vec{e}(m)\|_{\infty} + \Delta t \Lambda \|\vec{e}(m)\|_{\infty} + \Delta t C_1 \Delta s^2 + C_2 \Delta t^2 \end{aligned}$$

for all $0 \leq m \leq M - 1$ and $(x, y) \in \mathcal{R} \setminus \Omega$, thanks to $r_l \leq 1/4$. Noting that $e_l^{(x, y)}(m+1) = 0$ for all $(x, y) \in \Omega$ because of the DBCs, we conclude that the estimation

$$|e_l^{(x, y)}(m+1)| \leq (1 + \Delta t \Lambda) \|\vec{e}(m)\|_{\infty} + \Delta t C_1 \Delta s^2 + C_2 \Delta t^2$$

holds for all $(x, y) \in \mathcal{R}$. Together with $C_3 := \max\{C_1, C_2\}$, this yields

$$\begin{aligned} \|\vec{e}(m)\|_{\infty} &\leq (1 + \Delta t \Lambda)^m \|\vec{e}(0)\|_{\infty} + C_3 (\Delta t \Delta s^2 + \Delta t^2) \sum_{k=0}^{m-1} (1 + \Delta t \Lambda)^k \leq (1 + \Delta t \Lambda)^m \|\vec{e}(0)\|_{\infty} + C_3 (\Delta t \Delta s^2 + \Delta t^2) \frac{(1 + \Delta t \Lambda)^m - 1}{\Delta t \Lambda} \\ &\leq e^{m \Delta t \Lambda} \|\vec{e}(0)\|_{\infty} + C_3 (\Delta s^2 + \Delta t) \frac{e^{m \Delta t \Lambda} - 1}{\Lambda} \leq e^{T \Lambda} \|\vec{e}(0)\|_{\infty} + C_3 (\Delta s^2 + \Delta t) \frac{e^{T \Lambda} - 1}{\Lambda} = 0 + \frac{e^{T \Lambda} - 1}{\Lambda / C_3} (\Delta s^2 + \Delta t) \end{aligned}$$

for all $0 \leq m \leq M$. This shows that (A.3) converges in the supremum norm $\|\cdot\|_{\infty}$ to the solution of (A.1) and (16) on

$[0; 1]^2 \times [0; T]$. Moreover, it holds that

$$|\underline{a}_l^{(x,y)}(m)| \leq |\underline{a}_l^{(x,y)}(m) - \alpha_l(x, y, t_m)| + |\alpha_l(x, y, t_m)| \leq \|\vec{e}(m)\|_\infty + c \leq 1/2 + c$$

for all $0 \leq m \leq M$, $1 \leq l \leq L$ and $(x, y) \in \mathcal{R}$ if Δs and Δt are sufficiently small. Specifically, there exist K_0 and M_0 such that $\underline{\mathcal{L}}_l^{(x,y)}(\vec{d}(m)) = \underline{\mathcal{L}}_l^{(x,y)}(\vec{d}(m))$ for all $0 \leq m \leq M$, $(x, y) \in \mathcal{R}$ and $1 \leq l \leq L$, if $K \geq K_0$ and $M \geq M_0$. Consequently, (A.3) coincides with the scheme (15) in the case of $K \geq K_0$ and $M \geq M_0$ and the proof is complete. \square

Proof of Theorem 5. In the following, we will use the auxiliary results from the proof of Theorem 4. Recall that the sequence induced by the finite difference scheme (A.3) remains bounded by $c + 1/2$ in the supremum norm if $M \geq M_0$ and $K \geq K_0$ and that (A.3) can be interpreted as the Euler method of the ODE system

$$\frac{d}{dt} \underline{a}_l^{(x,y)}(t) = \mathbb{1}(x, y) \cdot \left(\mu_l \Delta^d \underline{a}_l^{(x,y)} + \underline{\mathcal{L}}_l^{(x,y)}(\vec{d}(t)) \right), \quad (\text{A.4})$$

where $(x, y) \in \mathcal{R}$ and $1 \leq l \leq L$. The global Lipschitz continuity of (A.4) and the global version of Picard-Lindelöf's theorem assert that (A.4) has a unique solution \vec{d} which is defined on the whole real line. Define $C'_1 := \frac{e^{T\Lambda} - 1}{\Lambda/C_3}$ and fix a K such that $K \geq K_0$ and $C'_1 \Delta s^2 < \varepsilon/12$. The uniform continuity of \vec{d} on $[0; T]$ implies the existence of a $\delta_1 > 0$ such that

$$\forall 1 \leq l \leq L \forall (x, y) \in \mathcal{R} \forall t, t' \in [0; T] (|t - t'| < \delta_1 \Rightarrow |\underline{a}_l^{(x,y)}(t) - \underline{a}_l^{(x,y)}(t')| < \varepsilon/12).$$

Similarly, thanks to the uniform continuity of $(\alpha_1, \dots, \alpha_L)$ on $[0; 1]^2 \times [0; T]$, there exists a $\delta_2 > 0$ such that

$$\forall 1 \leq l \leq L \forall (x, y) \in [0; 1]^2 \forall t, t' \in [0; T] (|t - t'| < \delta_2 \Rightarrow |\alpha_l(x, y, t) - \alpha_l(x, y, t')| < \varepsilon/12).$$

Moreover, by [44] there exists a $C'_2 > 0$ such that

$$\max_{0 \leq m \leq M} \max_{(x,y) \in \mathcal{R}, 1 \leq l \leq L} |\underline{a}_l^{(x,y)}(t_m) - \underline{a}_l^{(x,y)}(m)| \leq C'_2 \Delta t$$

for all $M \geq M_0$. Note that $\underline{a}_l^{(x,y)}(t_m)$ denotes the exact ODE solution at time t_m , whereas $\underline{a}_l^{(x,y)}(m)$ denotes the m -th sequence element in the method of Euler (A.3). In the following, let us fix an M such that $\Delta t = \frac{T}{M} < \min \left\{ \frac{T}{M_0}, \delta_1, \delta_2, \frac{\varepsilon}{4(C'_1 + C'_2)} \right\}$ and assume without loss of generality that $\varepsilon < 1$. Then, for an arbitrary $t \in [0; T]$ and an $0 \leq m \leq M$ such that $|t - t_m| < \Delta t$, it holds that

$$|\underline{a}_l^{(x,y)}(t)| \leq |\underline{a}_l^{(x,y)}(t) - \underline{a}_l^{(x,y)}(t_m)| + |\underline{a}_l^{(x,y)}(t_m) - \underline{a}_l^{(x,y)}(m)| + |\underline{a}_l^{(x,y)}(m)| \leq \varepsilon/12 + C'_2 \Delta t + (c + 1/2) \leq \varepsilon/12 + \varepsilon/4 + c + 1/2$$

for all $(x, y) \in \mathcal{R}$ and $1 \leq l \leq L$, meaning that the solution of (A.4) on $[0; T]$ is bounded by $c + 1$ in the supremum norm. Consequently, since $\|(z_1, \dots, z_{L+P})\|_\infty \leq c + 1$ implies $f_j(z_1, \dots, z_{L+P}) = f_j(z_1, \dots, z_{L+P})$ for all $1 \leq j \leq J$, it holds that the solution \vec{d} of the original ODE system (6) is well-defined on $[0; T]$ and satisfies $\vec{d}_{|[0; T]} = \vec{d}_{|[0; T]}$. Hence, for an arbitrary $t \in [0; T]$ and an $0 \leq m \leq M$ such that $|t - t_m| < \Delta t$, it holds that

$$\begin{aligned} |a_l^{(x,y)}(t) - \alpha_l(x, y, t)| &\leq |a_l^{(x,y)}(t) - \underline{a}_l^{(x,y)}(t_m)| + |\underline{a}_l^{(x,y)}(t_m) - \underline{a}_l^{(x,y)}(m)| + |\underline{a}_l^{(x,y)}(m) - \alpha_l(x, y, t_m)| + |\alpha_l(x, y, t_m) - \alpha_l(x, y, t)| \\ &\leq \varepsilon/12 + C'_2 \Delta t + C'_1 (\Delta s^2 + \Delta t) + \varepsilon/12 \leq \varepsilon/4 + (C'_1 + C'_2) \Delta t \leq \varepsilon/2 \end{aligned}$$

for all $(x, y) \in \mathcal{R}$ and $1 \leq l \leq L$. Together with

$$\begin{aligned}\phi_1 &:= \sup_{0 \leq t \leq T} \max_{(x,y) \in \mathcal{R}, 1 \leq l \leq L} \left| \frac{1}{N} A_l^{(x,y)}(t) - \alpha_l(x, y, t) \right|, \\ \phi_2 &:= \sup_{0 \leq t \leq T} \max_{(x,y) \in \mathcal{R}, 1 \leq l \leq L} \left| \frac{1}{N} A_l^{(x,y)}(t) - a_l^{(x,y)}(t) \right|, \\ \phi_3 &:= \sup_{0 \leq t \leq T} \max_{(x,y) \in \mathcal{R}, 1 \leq l \leq L} \left| a_l^{(x,y)}(t) - \alpha_l(x, y, t) \right|,\end{aligned}$$

we infer that $\mathbb{P}\{\phi_1 > \varepsilon\} \leq \mathbb{P}\{\phi_2 + \phi_3 > \varepsilon\} \leq \mathbb{P}\{\phi_2 > \varepsilon/2 \vee \phi_3 > \varepsilon/2\} \leq \mathbb{P}\{\phi_2 > \varepsilon/2\} + \mathbb{P}\{\phi_3 > \varepsilon/2\}$. Since the choice of K from above implies that $\phi_3 \leq \varepsilon/2$, it holds that $\mathbb{P}\{\phi_3 > \varepsilon/2\} = 0$ and Theorem 2 yields the claim. \square

Proof of Theorem 6. In the following, we will use results and variables from the proof of Theorem 4. We first discuss the space step error and start by fixing an $(x, y) \in \Omega$ where $x = 0$ and $\Delta s \leq y \leq 1 - \Delta s$. Due to Taylor's theorem there exist $x' \in (0; \Delta s)$ and $y' \in (y - \Delta s; y + \Delta s)$ such that

$$\begin{aligned}\frac{\rho_l(0, y, t)}{\mu_l} &= \partial_{xx} \alpha_l(0, y, t) + \partial_{yy} \alpha_l(0, y, t) - \left(\frac{\alpha_l(0 + \Delta s, y, t) - \alpha_l(0, y, t)}{\Delta s^2} + \frac{\alpha_l(0, y + \Delta s, t) - 2\alpha_l(0, y, t) + \alpha_l(0, y - \Delta s, t)}{\Delta s^2} \right) \\ &= \partial_{xx} \alpha_l(0, y, t) - \left(\frac{\partial_x \alpha_l(0, y, t)}{\Delta s} + \frac{\partial_{xx} \alpha_l(0, y, t)}{2!} + \frac{\partial_{xxx} \alpha_l(x', y, t)}{3!} \Delta s - \frac{\partial_{yyyy} \alpha_l(0, y', t)}{12} \Delta s^2 \right) \\ &= \frac{\partial_{xx} \alpha_l(0, y, t)}{2} + \Delta s \left(\frac{\partial_{xxx} \alpha_l(x', y, t)}{3!} - \frac{\partial_{yyyy} \alpha_l(\Delta s, y', t)}{12} \Delta s \right)\end{aligned}$$

As a similar calculation applies in the other seven boundary cases, we infer that there exists a $C'_1 \in \mathbb{R}_{\geq 0}$ which *does not depend* on $\Delta t, \Delta s$ and satisfies $|\rho_l(x, y, t)| \leq C'_1$ for all $(x, y) \in \Omega$, $0 \leq t \leq T$ and $1 \leq l \leq L$. By choosing $C_1 \geq C'_1$, we may assume without loss of generality that $C_1 \geq C'_1$.

In order to show consistency with respect to $\|\cdot\|_{1, \Delta s}$, we have to prove that

$$\left\| \frac{\vec{\eta}(x, y, t)}{\Delta t} + \vec{\rho}(x, y, t) \right\|_{1, \Delta s} \rightarrow 0, \quad \Delta t, \Delta s \rightarrow 0 \quad (\text{A.5})$$

For this, we define

$$C_{(x,y)} := \begin{cases} C'_1 & , (x, y) \in \Omega \\ C_1 \Delta s^2 & , (x, y) \in \mathcal{R} \setminus \Omega \end{cases}$$

conclude (by using results from the proof of Theorem 4)

$$\left\| \frac{\vec{\eta}(x, y, t)}{\Delta t} + \vec{\rho}(x, y, t) \right\|_{1, \Delta s} \leq \sum_{l=1}^L \sum_{(x,y) \in \mathcal{R}} (\Delta t C_2 + C_{(x,y)}) \Delta s^2 \leq L(4C_2 \Delta t + \sum_{(x,y) \in \mathcal{R} \setminus \Omega} (C_2 \Delta s^2) \Delta s^2 + \sum_{(x,y) \in \Omega} C_1 \Delta s^2)$$

Hence, as $|\Omega| = 2(K+1) + 2(K-1) = 4K$ and $|\mathcal{R} \setminus \Omega| = (K-1)^2$ we infer that $|\Omega| \Delta s^2 = |\Omega|/K^2 = 4/K = 4\Delta s$ and $|\mathcal{R} \setminus \Omega| \Delta s^2 \leq 1$. Consequently, it holds

$$L(4C_2 \Delta t + \sum_{(x,y) \in \mathcal{R} \setminus \Omega} (C_2 \Delta s^2) \Delta s^2 + \sum_{(x,y) \in \Omega} C_1 \Delta s^2) \leq L(4C_2 \Delta t + C_2 \Delta s^2 + 4C_1 \Delta s),$$

yielding the consistency (A.5).

Next, we show that the scheme is stable with respect to $\|\cdot\|_{\infty}$. (Note that this is sufficient, as $\|\cdot\|_{1, \Delta s} \leq 4L \|\cdot\|_{\infty}$.) From the proof of Theorem 4 we know that

$$\Delta t |\rho_l(x, y, t_m)| + |\eta_l(x, y, t_m)| \leq C_1 \Delta t + C_2 \Delta t^2$$

if Δs is sufficiently small. By performing a similar calculation as in the case of DBCs, we infer for $C'_3 := \max\{C_1, C_2\}$

$$|e_l^{(x,y)}(m)| \leq \|\vec{e}(m)\|_\infty + \Delta t \Lambda \|\vec{e}(m)\|_\infty + C'_3(\Delta t + \Delta t^2) \leq (1 + \Delta t \Lambda) \|\vec{e}(m)\|_\infty + C'_3(\Delta t + \Delta t^2),$$

which allows us to conclude, for all $0 \leq m \leq M$, that

$$\begin{aligned} \|\vec{e}(m)\|_\infty &\leq (1 + \Delta t \Lambda)^m \|\vec{e}(0)\|_\infty + C'_3(\Delta t + \Delta t^2) \sum_{k=0}^{m-1} (1 + \Delta t \Lambda)^k \leq (1 + \Delta t \Lambda)^m \|\vec{e}(0)\|_\infty + C'_3(\Delta t + \Delta t^2) \frac{(1 + \Delta t \Lambda)^m - 1}{\Delta t \Lambda} \\ &\leq e^{m\Delta t \Lambda} \|\vec{e}(0)\|_\infty + C'_3(1 + \Delta t) \frac{e^{m\Delta t \Lambda} - 1}{\Lambda} \leq e^{T\Lambda} \|\vec{e}(0)\|_\infty + C'_3(1 + \Delta t) \frac{e^{T\Lambda} - 1}{\Lambda} \\ &= 0 + \frac{e^{T\Lambda} - 1}{\Lambda/C'_3} (1 + \Delta t) \end{aligned}$$

Consequently, as \vec{a} is bounded on $[0; 1]^2 \times [0; T]$,

$$|a_l^{(x,y)}(m) - a_l^{(x,y)}(0)| \leq |a_l^{(x,y)}(m) - a_l(x, y, t_m)| + |\alpha_l(x, y, t_m) - a_l^{(x,y)}(0)|$$

and $\|\cdot\|_{1, \Delta s} \leq 4L\|\cdot\|_\infty$, the scheme is stable with respect to $\|\cdot\|_{1, \Delta s}$. \square

Proof of Corollary 1. From the proof of Theorem 6 it becomes apparent that the scheme (23) is accurate of order $(1, 1)$ [45, Definition 2.3.3] with respect to $\|\cdot\|_{\Delta s, 1}$. Since Theorem 6 implies also that (23) is stable with respect to $\|\cdot\|_{\Delta s, 1}$, the proof of Lax's theorem [45, Theorem 2.5.2] yields the claim. \square

Proof of Theorem 7. Theorem 6 ensures that there exist $K_0, M_0 \geq 0$ and $C > 0$ such that

$$\max_{0 \leq m \leq M} \frac{1}{K^2} \sum_{\substack{(x,y) \in \mathcal{R}_K \\ 1 \leq l \leq L}} |\alpha_l(x, y, m\Delta t) - a_l^{(x,y)}(m)| \leq C(\Delta s + \Delta t)$$

for all $K \geq K_0$ and $M \geq M_0$. Moreover, the ODE system

$$\frac{d}{dt} a_l^{(x,y)}(t) = \mu_l \Delta^d a_l^{(x,y)} + \mathcal{L}_l^{(x,y)}(\vec{a}(t)), \quad (\text{A.6})$$

where $(x, y) \in \mathcal{R}$ and $1 \leq l \leq L$, is globally Lipschitz. Thus, the global version of Picard-Lindelöf's theorem assert that (A.6) has a unique solution \vec{a} which is defined on the whole real line. The uniform continuity of \vec{a} on $[0; T]$ implies the existence of a $\delta_1 > 0$ such that

$$\forall 1 \leq l \leq L \forall (x, y) \in \mathcal{R} \forall t, t' \in [0; T] (|t - t'| < \delta_1 \Rightarrow |a_l^{(x,y)}(t) - a_l^{(x,y)}(t')| < \varepsilon/48L).$$

Similarly, thanks to the uniform continuity of $(\alpha_1, \dots, \alpha_L)$ on $[0; 1]^2 \times [0; T]$, there exists a $\delta_2 > 0$ such that

$$\forall 1 \leq l \leq L \forall (x, y) \in [0; 1]^2 \forall t, t' \in [0; T] (|t - t'| < \delta_2 \Rightarrow |\alpha_l(x, y, t) - \alpha_l(x, y, t')| < \varepsilon/48L).$$

Moreover, by [44] there exists a $C' > 0$ such that

$$\max_{0 \leq m \leq M} \max_{(x,y) \in \mathcal{R}, 1 \leq l \leq L} |a_l^{(x,y)}(t_m) - a_l^{(x,y)}(m)| \leq C' \Delta t$$

for all $M \geq M_0$. Note that $a_l^{(x,y)}(t_m)$ denotes the exact ODE solution at time t_m , whereas $a_l^{(x,y)}(m)$ denotes the m -th sequence element in the method of Euler (A.6). In the following, let us fix an $M \geq M_0$ such that $\Delta t = T/M < \min\{\delta_1, \delta_2, \varepsilon/48LC, \varepsilon/48LC'\}$ and $K \geq K_0$ such that $C\Delta s \leq \varepsilon/48L$. Then, for an arbitrary $t \in [0; T]$ and an $0 \leq m \leq M$

such that $|t - t_m| < \Delta t$, it holds that

$$\begin{aligned} |a_l^{(x,y)}(t) - \alpha_l(x, y, t)| &\leq |a_l^{(x,y)}(t) - a_l^{(x,y)}(t_m)| + |a_l^{(x,y)}(t_m) - a_l^{(x,y)}(m)| + |a_l^{(x,y)}(m) - \alpha_l(x, y, t_m)| + |\alpha_l(x, y, t_m) - \alpha_l(x, y, t)| \\ &\leq \varepsilon/48L + \varepsilon/48L + \varepsilon/24L + \varepsilon/48L \leq \varepsilon/8L \end{aligned}$$

for all $(x, y) \in \mathcal{R}$ and $1 \leq l \leq L$. Together with

$$\begin{aligned} \phi_1 &:= \sup_{0 \leq t \leq T} \frac{1}{K^2} \sum_{\substack{(x,y) \in \mathcal{R}_K \\ 1 \leq l \leq L}} \left| \frac{1}{N} A_l^{(x,y)}(t) - \alpha_l(x, y, t) \right|, \\ \phi_2 &:= \sup_{0 \leq t \leq T} \frac{1}{K^2} \sum_{\substack{(x,y) \in \mathcal{R}_K \\ 1 \leq l \leq L}} \left| \frac{1}{N} A_l^{(x,y)}(t) - a_l^{(x,y)}(t) \right|, \\ \phi_3 &:= \sup_{0 \leq t \leq T} \frac{1}{K^2} \sum_{\substack{(x,y) \in \mathcal{R}_K \\ 1 \leq l \leq L}} \left| a_l^{(x,y)}(t) - \alpha_l(x, y, t) \right|, \end{aligned}$$

we infer that $\mathbb{P}\{\phi_1 > \varepsilon\} \leq \mathbb{P}\{\phi_2 + \phi_3 > \varepsilon\} \leq \mathbb{P}\{\phi_2 > \varepsilon/2 \vee \phi_3 > \varepsilon/2\} \leq \mathbb{P}\{\phi_2 > \varepsilon/2\} + \mathbb{P}\{\phi_3 > \varepsilon/2\}$. Since the choice of M, K from above implies that $\phi_3 \leq (\varepsilon/8L)3L \leq \varepsilon/2$, it holds that $\mathbb{P}\{\phi_3 > \varepsilon/2\} = 0$. Consequently, Theorem 3 and $\|\cdot\|_{\Delta s, 1} \leq 4L\|\cdot\|_\infty$ yield the claim. \square