Stability of a cross-diffusion system and approximation by repulsive random walks: a duality approach

Vincent Bansaye Ayman Moussa Felipe Muñoz-Hernández

September 14, 2021

Abstract

We consider conservative cross-diffusion systems for two species where individual motion rates depend linearly on the local density of the other species. We develop duality estimates and obtain stability and approximation results. We first control the time evolution of the gap between two bounded solutions by means of its initial value. As a by product, we obtain a uniqueness result for bounded solutions valid for any space dimension, under a smallness assumption. Using a discrete counterpart of our duality estimates, we prove the convergence of random walks with local repulsion in one dimensional discrete space to cross-diffusion systems. More precisely, we prove sharp quantitative estimates for the gap between the stochastic process and the cross-diffusion system. We complete this study with a rough but general estimate and convergence results, when the population and the number of sites become large.

Key words and phrases: *Cross-diffusion, duality, stability, scaling limits, repulsive random walks.*

1 Introduction and notation

Approximations of interacting large populations is motivated by physics, chemistry, biology and ecology. A famous macroscopic model was introduced by Shigesada, Kawasaki and Teramoto in [27] to describe competing species which diffuse *with* local repulsion. In the case of two species, it writes

$$\begin{cases} \partial_t u - \Delta\big(d_1 u + a_{11} u^2 + a_{12} uv\big) = u(r_1 - s_{11} u - s_{12} v), \\ \partial_t v - \Delta\big(d_2 v + a_{21} uv + a_{22} v^2\big) = v(r_2 - s_{21} u - s_{22} v), \end{cases}$$

where u and v are the densities of the two species and d_i, r_i, a_{ij} and s_{ij} are non-negative real numbers. Completed by initial and boundary conditions, this system (that we simply refer to as the *SKT system*) offers a model for the spreading of two interacting species which mutually influence their propensity to diffuse, through the cross-diffusion terms a_{ij}. The other coefficients represent either natural diffusion (d_i coefficients), reproduction (r_i coefficients) or competition (s_{ij} coefficients). The main motivation of [27] was to propose a population dynamics model able to detect segregation, that is the existence of non-constant steady states $\overline{u}$ and $\overline{v}$ having disjoint superlevel sets of low threshold value. As a consequence of this motivation, the first mathematical results dealing with this system focused on sufficient conditions for the coefficients to ensure existence of non-constant steady states, with a careful study of the stability of the latter. This study of possible segregation states is still active and we refer to the introduction of [3] for a nice state of the art. It is a striking fact that during its first years of existence within the mathematical community, the SKT system has not been studied through the prism of its Cauchy problem. As a matter of fact, existence of solutions has been tackled only a few years later: the first paper dealing with this issue is [20] and explores the system under very restrictive conditions. Several attempt followed, but only with partial results. A substantial progress was achieved by Amann [1, 2], who proposed a rather abstract approach to study generic quasilinear parabolic systems. The scope of this technology goes far beyond the sole case of cross-diffusion systems. In the specific case of the SKT system, it offers existence of local (regular) solutions, together with a criteria of explosion to decide if the existence is global or not. This fundamental result of Amann has been then used by several authors to establish existence of global solutions for particular forms of the SKT system. This is done, in general, under a strong constraint on the coefficients. For instance, [21] treats the case of equal diffusion rates in low dimension and [16], settles the one of triangular systems (that is, for two species, when $a_{12}a_{21} = 0$). However, the general question of existence of global solution for the complete system remains open, even in low dimension.

Another way to produce a global solution is to sacrifice the regularity of the solutions, and deal with only weak ones. This strategy relies on the so-called *entropic structure* of the system: SKT systems as the one previously introduced, admit Lyapunov functionals which decay along time and whose dissipation allows to control the gradient of the solution. This method has been used successfully in [4] to prove, for the first time, existence of global weak solutions for the SKT system, without restrictive assumptions on its coefficients. After it first discovery in [15], this entropic structure has been explored and generalized to several systems, allowing for the construction of global weak solutions for variants of the original SKT system (see [19, 11] and the references therein). With this low level of regularity for the solutions, uniqueness becomes an issue in itself. It has been studied either under simplifying assumptions on the system like in [24, 7] or in the weak-strong setting thanks to the use of a relative entropy (see [8]).

1.1 Objectives and state of the art

This work is initially motivated by yet another mathematical challenge offered by the SKT system: its rigorous derivation. The diffusion operator used in the system in SKT system is specific. We focus in this paper on the main difficulty raised by this operator, which is the non-linearity of diffusion term. The initial goal of the work is to approximate the conservative SKT system, without self-diffusion, that is the following one

$$\begin{cases} \partial_t u - \Delta(d_1 u + a_{12} uv) = 0, \\ \partial_t v - \Delta(d_2 v + a_{21} uv) = 0, \end{cases} \tag{1}$$

where all the coefficients d_i and a_{ij} are assumed positive. Whereas (possibly heterogeneous) diffusion of lifeless matter (*e.g.* ink or any type of chemical substance) uses the Fick diffusion operator $-\mathrm{div}(\mu\nabla\cdot)$ to express the spread, SKT systems rely on the (more singular) operator $-\Delta(\mu\,\cdot)$. As it was already explained in [27], this choice of diffusion operator is at the core of the repulsive mechanism allowing the segregation to appear. However, the justification proposed in [27] was rather formal, leaving open the question of the rigorous justification of SKT systems. As far as our knowledge goes, there exist mainly three approaches for the derivation of SKT systems

(i) The first path was proposed in [17], where an SKT model is obtained as an asymptotic limit of a family of reaction-diffusion systems. In this approach the idea is that one of the two species exists in two states (stressed or not), and switch from one to the other with a reaction rate which diverges. This was used in [17] to obtain formally a triangular cross diffusion system. This strategy has been followed with a rigorous analysis, mainly to produce triangular systems (see [28] and references therein) and more recently for a family of "full" systems in [10] which, however, do not include the SKT one.

(ii) Another strategy was proposed by Fontbona and Méléard in [14]. The idea is to start from a stochastic population model in continuous space where the individuals' displacements depend on the presence of concurrents. Then, the large population limit (under adequate scaling) leads to a non-local cross-diffusion model. In comparison with the system (1), the limit model rigorously derived in [14] is a lot less singular, because of several convolution kernels. It was explicitly asked in [14], whether letting the convolution kernels vanish to the Dirac mass was handable limit or not. A first partial answer was given in [23], but applied for only specific triangular systems. More recently, it was discovered in [13] that even for the non-local systems, it is possible to ensure the persistence of the entropy structure, allowing to answer fully to the question of Fontbona and Méléard, at least for the standard SKT system.
A little bit before [13] appeared, Chen *et. al.* proposed another strategy in [5] (see also [6] which deals with a slightly different family of systems). It also starts from a stochastic model and makes use of an intermediate non-local one. The main difference with [14, 23, 13] is that in [5] the two asymptotics are done simultaneously (size of population to infinity and parameter of regularization to 0). This direct approach amounts to "commute" the asymptotic diagram from the stochastic model to the final PDE; this is a common feature with the current work that we will comment later on.

(iii) The third path was proposed in [9] and justifies the SKT model through a semi-discrete one. The latter is itself derived from a stochastic population model in discrete space where individuals are assumed to move by pair, in order to ensure reversibility of the process and the existence of an entropy for the limit model. In [9] the link to the stochastic was done formally whereas the asymptotic analysis linking the semi-discrete model to the SKT system was proved rigorously, relying on a compactness argument which is allowed thanks to the existence of the Lyapunov functional for the semi-discrete system.

In this paper, we are interested in connections between microscopic random individual-based models (or particle system) and such macroscopic deterministic dynamics, in the spirit of strategies (ii) and (iii) described above. We do not use any non-local approximated system as in [14, 5], being inspired instead by the semi-discrete approach proposed in [9]. We consider also a discrete space and that each species moves randomly and is only sensitive to the local size of the other species. Let us comment the main differences and novelties of this work compared to [9]. First, we prove rigorously that the suitably scaled stochastic process converges in law in Skorokhod space to SKT system (1) and we perform this space and time scaling limit at once. Besides, individuals of each species move independently with a rate proportional to the number of individuals of the other species, on the same site. We do not need to make them move by pair, which may be hard to justify regarding phenomenon at stake. Indeed, we do not need a reversibility property and do not use the entropic structure. The main difficulty to prove convergence of the stochastic process at once lies in the control of the cumulative quadratic rates due to local interactions when the number of sites becomes large. As far as we have seen, entropy structure does not provide the suitable control of these non-linear terms and a way to get tightness and identification in general. We use a different approach based on generalized duality. This provides quantitative estimates in terms of space discretization and size of population. Moreover, at the level of the PDE system, it implies a local uniqueness result for bounded solutions of the SKT system. The duality approach allows to compare locally the stochastic process with its semi-discrete deterministic approximation. It is optimal in the sense that it provides the good time space scaling for such an approximation.

Let us describe now the stochastic individual-based model. The population is spatially distributed among M sites. The process under consideration is a continuous time Markov chain $(\boldsymbol{U}(t), \boldsymbol{V}(t))_{t\geq 0}$ taking values in $\mathbb{N}^M \times \mathbb{N}^M$. The two coordinates count the number of individuals of each species at each site, for each time $t \geq 0$. Each individual of each species follows a random walk and its jumps rate increases linearly with respect to the number of individuals of the other species. The dynamic is defined by the jump rates as follows. For any vector of configurations $(\boldsymbol{u}, \boldsymbol{v}) \in \mathbb{N}^M \times \mathbb{N}^M$, the transitions are

$$\begin{aligned} \boldsymbol{u} &\mapsto \boldsymbol{u} + \left(\mathbf{e}_{i+\theta} - \mathbf{e}_i\right) \quad \text{at rate} \quad 2u_i(d_1 + a_{12}v_i), \\ \boldsymbol{v} &\mapsto \boldsymbol{v} + \left(\mathbf{e}_{i+\theta} - \mathbf{e}_i\right) \quad \text{at rate} \quad 2v_i(d_2 + a_{21}u_i), \end{aligned}$$

where $(\mathbf{e}_j)_{1\leq j\leq M}$ is the canonical basis of $\mathbb{R}^M$, $\mathbf{e}_0 = \mathbf{e}_M$, $\mathbf{e}_{M+1} = \mathbf{e}_1$ and $\theta \in \{-1, 1\}$ with both values equally likely. Let us mention that hydrodynamic limits of other stochastic models with repulsive species have been considered, in particular in the context of exclusion processes, see e.g. [26]. In that case, local densities are bounded so difficulties and limits are different. In an other direction, stochastic versions of the limiting SKT systems have been considered, see e.g. [12].

This work contains two main results which at first sight can appear unrelated in their formulation. The first result is a quantitative stability estimate on the SKT system which bounds the distance between two solutions in terms of their initial distance. This result is based on a new duality lemma and applies for bounded solutions, only if one of them is small enough. As a by-product of this stability estimate, we prove uniqueness of (small) bounded solutions of the conservative SKT system. This result is valid in arbitrary dimension and is, as far as our knowledge goes, new. Uniqueness theorems for (only) bounded solutions of the full SKT system are missing in the current literature [7, 8, 24].

The second main result is the convergence of the properly scaled sequence of processes $(\boldsymbol{U}^{M,N}, \boldsymbol{V}^{M,N})_{M,N\in\mathbb{N}}$ to the SKT system. We obtain quantitative estimates of the gap between the trajectories of this process extended to the continuous space and the solution of SKT system, in a large population and diffusive regime. This analysis is performed in a one dimensional setting for the space variable. The strategy is to insert the semi-discrete model proposed in [9] and estimate separately the gap between our stochastic process and this semi-discrete system and then, estimate (with enough uniformity) the distance between the semi-discrete system and the continuous SKT limit. Following this plan, we first propose a general estimate, which rely on naive bounds of the quadratic diffusion term. Roughly, we simply bound locally the size of the population by the (constant) total number of individuals. These bounds allow for convergence with a fixed number of sites but lead to an unreasonable assumption of a superexponential number of individuals per site when the number of sites increases. When we faced this difficulty, we tried to obtain an estimate as sharp as possible to capture the good scales and compare the semi-discrete system and the continuous one. It's during this step that we discovered the stability estimate described above, which is interesting for its own sake. A nice feature of this stability estimate is that we can transfer it onto the semi-discrete and stochastic setting. We obtain then the convergence of the stochastic model towards the SKT system, with sharp estimates and relevant size scales. This asymptotic study shares a similar limitation as the previous paragraph: it holds only under the assumption of small regular solution of the SKT system, which is ensured by Amann's theorem [1, 2].

The paper is organized as follows. In the end of this section, we collect several notations which will be used throughout the paper. In Section 2 we define the (sequence of) stochastic processes we consider, we recover the semi-discrete system introduced in [9] and state our two main results. In Section 3 we show the convergence in law in path space of the stochastic process towards the semi-discrete system when the number of individuals goes to infinity but the number of sites remains fixed. We provide a quantification of this convergence. It implies the general (no restriction on the limiting SKT system) but naive (in terms of scales) convergence discussed above. Then, Section 4 is dedicated to the duality estimates with source terms and their consequences. These duality estimates account for the interacting system when one of the population is seen as an exogenous environment, which amounts to decouple the two species. In a first short paragraph (Subsection 4.1) we state and prove the generalized duality lemma and its application to the stability estimate of the SKT system in the continuous setting. This paragraph is the only one of the study in which we work in arbitrary dimension for the space variable. Then, the rest of Section 4 focuses on the translation of these estimates in the semi-discrete setting. This includes the definition of reconstruction operators, the study of the discrete laplacian matrix and the translation of classical function spaces into the discrete setting. Eventually in Section 5, we apply the previous machinery to the difference between the stochastic process and the approximated system that solutions of (1) solve when looked at a semi-discrete level. We then deduce our main asymptotic theorem by controlling some martingales and approximation errors. In a short appendix, we also give a dictionary which gives the correspondence of different objects in the discrete and continuous settings.

1.2 Notation

Finite-dimensional vectors

Throughout the article, vectors will always be written in bold letters and if not stated otherwise, the components of the vector $\boldsymbol{u} \in \mathbb{R}^M$ are $(u_i)_{1\leq i\leq M}$. The canonical basis of $\mathbb{R}^M$ will be denoted $(\mathbf{e}_j)_{1\leq j\leq M}$. Due to the periodic boundary condition that we will use, we will frequently use the convention $\mathbf{e}_0 = \mathbf{e}_M$ and $\mathbf{e}_{M+1} = \mathbf{e}_1$.

Given $M \in \mathbb{N}$ and $p \in [1,\infty]$ we denote by $\|\cdot\|_p = \left(\sum_{i=1}^M |x_i|^p\right)^{1/p}$ the usual ℓ^p norm on $\mathbb{R}^M$ and $\|\cdot\|_{p,M}$ the rescaled norm defined for $\boldsymbol{x} \in \mathbb{R}^M$ by

$$\|\boldsymbol{x}\|_{p,M} \coloneqq \left(\frac{1}{M}\sum_{i=1}^M |x_i|^p\right)^{1/p} \quad \text{for } p<\infty, \text{and} \quad \|\boldsymbol{x}\|_\infty \coloneqq \max_{1\leq i\leq M} |x_i|.$$

Similarly, the corresponding (rescaled) euclidean inner-product of $\mathbb{R}^M$ is denoted $(\cdot|\cdot)_M$:

$$(\boldsymbol{x}|\boldsymbol{y})_M = \frac{1}{M}\sum_{i=1}^M x_i y_i,$$

so that $\|\cdot\|_{2,M}^2 = (\cdot|\cdot)_M$.

The symbol $\odot$ is the internal Hadamard product on $\mathbb{R}^M$, that is $(\boldsymbol{x}\odot\boldsymbol{y})_i = x_i y_i$. We will also often use (when it makes sense) the operator $\boldsymbol{x} \oslash \boldsymbol{y}$ defined by $(\boldsymbol{x}\oslash\boldsymbol{y})_i = x_i/y_i$ and the "vectorial" square-root $\boldsymbol{x}^{1/2}$ whose components are $(\sqrt{x_i})_{1\leq i\leq M}$.

The arithmetic average of all the components of a vector $\boldsymbol{x}$ will be denoted

$$[\boldsymbol{x}]_M \coloneqq \frac{1}{M}\sum_{i=1}^M x_i.$$

The vector of $\mathbb{R}^M$ for which every component equals 1 is denoted $\mathbf{1}_M$. The orthogonal projection onto $\mathrm{Span}_{\mathbb{R}}(\mathbf{1}_M)^\perp$ is denoted with a tilde, that is: $\widetilde{\boldsymbol{x}} = \boldsymbol{x} - [\boldsymbol{x}]_M\mathbf{1}_M$.

For $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^M$ we write $\boldsymbol{x} \geq \boldsymbol{y}$ whenever $\boldsymbol{x}-\boldsymbol{y} \in \mathbb{R}_+^M$.

Functions

We will manipulate random and deterministic functions which may depend on the time variable $t \in \mathbb{R}_+$ and the space variable $x \in \mathbb{T}^d$, where $\mathbb{T} \coloneqq \mathbb{R}/\mathbb{Z}$ is the flat periodic torus. We will rely on the following convention for functions: uppercase letters will be reserved for random elements whereas lowercase letters will represent deterministic functions. Accordingly to the previous paragraph, vector valued functions will be denoted in bold whereas scalar valued functions will be denoted with the normal font.

Quite often results will be stated on a fixed time interval $[0,T]$. For this reason, we introduce the periodic cylinder $Q_T \coloneqq [0,T]\times\mathbb{T}^d$. For any function space E defined on $\mathbb{T}^d$ or Q_T, the corresponding norm will be denoted $\|\cdot\|_E$, *e.g.* $\|\cdot\|_{L^2(\mathbb{T}^d)}$. In case of a Hilbert structure, the inner-product will be denoted by $(\cdot|\cdot)_E$, *e.g.* $(\cdot|\cdot)_{L^2(\mathbb{T}^d)}$. We will use frequently two Sobolev spaces on $\mathbb{T}^d$, the definition of which we briefly recall for the reader's convenience.

Any distribution $\varphi \in \mathscr{D}'(\mathbb{T}^d)$ decomposes

$$\varphi = \sum_{k\in\mathbb{Z}^d} c_k(\varphi) e_k,$$

where $e_k(x) := e^{2i\pi k\cdot x}$, and $c_k(\varphi) := \langle \varphi, e_k\rangle$. For $s \in \mathbb{R}$ we define $H^s(\mathbb{T}^d)$ as the subspace of $\mathscr{D}'(\mathbb{T}^d)$ whose elements φ satisfy

$$\sum_{k\in\mathbb{Z}^d} |c_k(\varphi)|^2 (1+|k|^2)^s < +\infty,$$

equipped with the norm

$$\|\varphi\|_{H^s(\mathbb{T}^d)} = \left\{ \sum_{k\in\mathbb{Z}^d} |c_k(\varphi)|^2 (1+|k|^2)^s \right\}^{1/2}.$$

By analogy with the average notation of the previous paragraph, for any integrable function φ defined on $\mathbb{T}^d$, we denote

$$[\varphi]_{\mathbb{T}^d} := \int_{\mathbb{T}^d} \varphi,$$

which is in general $[\varphi]_{\mathbb{T}^d} = c_0(\varphi)$ if φ is merely a distribution. The expression

$$\|\varphi\|_{\dot{H}^s(\mathbb{T}^d)} := \left\{ \sum_{k\in\mathbb{Z}^d} |c_k(\varphi)|^2 |k|^{2s} \right\}^{1/2},$$

is only a semi-norm on $H^s(\mathbb{T}^d)$ and is a norm on the homogeneous Sobolev space $\dot{H}^s(\mathbb{T}^d)$ constituted of those elements φ belonging to $H^s(\mathbb{T}^d)$ and having a vanishing mean, *i.e.* for which $[\varphi]_{\mathbb{T}} = c_0(\varphi) = 0$. We use mainly these spaces for $s = 1$ and $s = -1$.

Finally, for any metric space X, $D([0,T], X)$ denotes the space of càdlàg functions from $[0,T]$ to X endowed with the Skorokhod topology.

2 Main objects and results

Before stating our main results, we need to define precisely the objects that we aim at considering.

2.1 Repulsive random walks and scaling

Let us define the stochastic process by means of a trajectorial representation using Poisson point measures. We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying the usual conditions. We introduce a family of independent Poisson point measure $(\mathcal{N}^j)_{j\in\mathbb{N}}$ on $\mathbb{R}_+ \times \mathbb{R}_+ \times \{-1,1\}$ with common intensity $\mathrm{d}s \otimes \mathrm{d}\rho \otimes \beta(\mathrm{d}\theta)$, where β is the law of a Bernoulli$\left(\frac{1}{2}\right)$ random variable. Almost surely the initial data $(\boldsymbol{U}(0), \boldsymbol{V}(0))$ belongs to $\mathbb{N}^M \times \mathbb{N}^M$, and the corresponding process $(\boldsymbol{U}(t), \boldsymbol{V}(t))_{t\geq 0}$ is then defined as the unique strong solution in $D([0,\infty), \mathbb{N}^{2M})$ of the following system of stochastic differential equations (SDEs) driven by the aforementioned measures

$$\begin{cases} \boldsymbol{U}(t) = \boldsymbol{U}(0) + \displaystyle\int_0^t \int_{\mathbb{R}_+\times\{-1,1\}} \sum_{j=1}^M \mathbf{1}_{\rho \leq 2U_j(s^-)(d_1 + a_{12}V_j(s^-))} \big(\mathbf{e}_{j+\theta} - \mathbf{e}_j\big)\, \mathcal{N}^j(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\theta), \\ \boldsymbol{V}(t) = \boldsymbol{V}(0) + \displaystyle\int_0^t \int_{\mathbb{R}_+\times\{-1,1\}} \sum_{j=1}^M \mathbf{1}_{\rho \leq 2V_j(s^-)(d_2 + a_{21}U_j(s^-))} \big(\mathbf{e}_{j+\theta} - \mathbf{e}_j\big)\, \mathcal{N}^j(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\theta), \end{cases}$$

where the jump rates d_1, d_2, a_{12} and a_{21} are the one of (1). Uniqueness and existence for the previous system of SDEs are obtained easily from a classical inductive construction. Indeed, the total population size of each species is constant along time: $\|\boldsymbol{U}(t)\|_{1,M} = \|\boldsymbol{U}(0)\|_{1,M}$, $\|\boldsymbol{V}(t)\|_{1,M} = \|\boldsymbol{V}(0)\|_{1,M}$. Therefore, conditionally on the initial value $(\boldsymbol{U}(0), \boldsymbol{V}(0))$, the process $(\boldsymbol{U}(t), \boldsymbol{V}(t))_{t\geq 0}$ is a pure jump Markov process on a finite state space with bounded rates.

We are interested in the approximation (hydrodynamic limit) when the population size and the number of sites tend to infinity. Informally, we consider $(\boldsymbol{U}(M^2t)/N, \boldsymbol{V}(M^2t)/N)_{t\geq 0}$ and interaction now occurs through the local density of individuals. The scaling parameter $N \in \mathbb{N}^*$ yields the normalization of the population per site and provides a limiting density when N goes to infinity. The initial population per site is of order of magnitude N and each species' motion rate is an affine function of the density of the other species on the same site. The motion of each individual is centered and we consider the diffusive regime. As a consequence, we accelerate the time by the factor of M^2, which amounts to multiply the transition rates by M^2.

We denote the renormalized process by $(\boldsymbol{U}^{M,N}(t), \boldsymbol{V}^{M,N}(t))_{t\geq 0}$. Moreover, for $u, v \in \mathbb{R}$ and $i, j = 1, 2$ we set

$$\begin{aligned} \eta_{1,j}^{M,N}(t) &:= 2M^2 N U_j^{M,N}(t)\big(d_1 + a_{12}V_j^{M,N}(t)\big), \\ \eta_{2,j}^{M,N}(t) &:= 2M^2 N V_j^{M,N}(t)\big(d_2 + a_{21}U_j^{M,N}(t)\big). \end{aligned}$$

For a given initial condition $(\boldsymbol{U}^{M,N}(0), \boldsymbol{V}^{M,N}(0))$, the process $(\boldsymbol{U}^{M,N}(t), \boldsymbol{V}^{M,N}(t))_{t\geq 0}$ is the unique solution in $D([0,\infty), \mathbb{R}_+^{2M})$ of the following system of SDEs

$$\begin{cases} \boldsymbol{U}^{M,N}(t) = \boldsymbol{U}^{M,N}(0) + \displaystyle\int_0^t \int_{\mathbb{R}_+\times\{-1,1\}} \sum_{j=1}^M \mathbf{1}_{\rho \leq \eta_{1,j}^{M,N}(s^-)} \frac{\mathbf{e}_{j+\theta} - \mathbf{e}_j}{N}\, \mathcal{N}^j(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\theta), \\ \boldsymbol{V}^{M,N}(t) = \boldsymbol{V}^{M,N}(0) + \displaystyle\int_0^t \int_{\mathbb{R}_+\times\{-1,1\}} \sum_{j=1}^M \mathbf{1}_{\rho \leq \eta_{2,j}^{M,N}(s^-)} \frac{\mathbf{e}_{j+\theta} - \mathbf{e}_j}{N}\, \mathcal{N}^j(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\theta). \end{cases} \tag{2}$$

2.2 The intermediate (semi-discrete) system

To estimate the gap between the discrete stochastic process (2) and the SKT system (1), we are going to use a third system on which our asymptotic analysis will pivot

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{u}^M(t) - \Delta_M\big(d_1\boldsymbol{u}^M(t) + a_{12}\boldsymbol{u}^M(t)\odot\boldsymbol{v}^M(t)\big) = 0, \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{v}^M(t) - \Delta_M\big(d_2\boldsymbol{v}^M(t) + a_{21}\boldsymbol{u}^M(t)\odot\boldsymbol{v}^M(t)\big) = 0, \end{cases} \tag{3}$$

where the unknowns are the vector valued curves $\boldsymbol{u}^M, \boldsymbol{v}^M : \mathbb{R}_+ \to \mathbb{R}^M$, and the matrix Δ_M is the periodic laplacian matrix, that is

$$\Delta_M := M^2 \begin{pmatrix} -2 & 1 & 0 & \cdots & 1 \\ 1 & -2 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & -2 & 1 \\ 1 & \cdots & 0 & 1 & -2 \end{pmatrix} \in \mathrm{M}_M(\mathbb{R}). \tag{4}$$

This semi-discrete system corresponds to a large population approximation but fixed number of sites M. Existence and uniqueness for (3) can be proven using the standard Picard-Lindelöf theorem, as this is done in [9] where this semi-discrete system has been introduced.

2.3 Formal insight

Before stating our main results, let us give an informal argument to see how the stochastic process (2) can be linked with the SKT system (1), through the semi-discrete system (3).

We first introduce the infinitesimal generator $L^{M,N}$ of the process (2). For this purpose, we define the translation operator τ_a for any vector $a \in \mathbb{R}^M$. It acts on any function $G\colon \mathbb{R}^M \to \mathbb{R}$ by the formula $\tau_a G(\cdot) := G(\cdot + a)$. Then, for $1 \le j \le M$, we define the operator

$$\mathcal{L}_j^N = \tau_{N^{-1}(\boldsymbol{e}_{j+1}-\boldsymbol{e}_j)} + \tau_{N^{-1}(\boldsymbol{e}_{j-1}-\boldsymbol{e}_j)} - 2\mathrm{Id},$$

for $G\colon \mathbb{R}^M \to \mathbb{R}$. We recall here the periodic convention: $\mathbf{e}_0 = \mathbf{e}_M$ and $\mathbf{e}_{M+1} = \mathbf{e}_1$. Then, for any measurable and bounded function $F\colon \mathbb{R}_+^{2M} \to \mathbb{R}$, we define for $(\boldsymbol{u}, \boldsymbol{v}) \in \mathbb{R}_+^{2M}$

$$L^{M,N}F\big(\boldsymbol{u},\boldsymbol{v}\big) = \sum_{j=1}^M \frac{1}{2}\left\{\eta_{1,j}^{M,N}(u_j,v_j)\mathcal{L}_j^N[F(\cdot,\boldsymbol{v})](\boldsymbol{u}) + \eta_{2,j}^{M,N}(u_j,v_j)\mathcal{L}_j^N[F\big(\boldsymbol{u},\cdot)](\boldsymbol{v})\right\}.$$

For N going to infinity and F differentiable, Taylor's approximation ensures that $L^{M,N}F$ converges to

$$L^M F(\boldsymbol{u},\boldsymbol{v}) = \big(\Delta_M(d_1\boldsymbol{u} + a_{12}\boldsymbol{u}\odot\boldsymbol{v})\,\big|\,\nabla_{\boldsymbol{u}}F(\boldsymbol{u},\boldsymbol{v})\big) + \big(\Delta_M(d_2\boldsymbol{v} + a_{21}\boldsymbol{v}\odot\boldsymbol{u})\,\big|\,\nabla_{\boldsymbol{v}}F(\boldsymbol{u},\boldsymbol{v})\big),$$

where $(\cdot|\cdot)$ is the inner product on $\mathbb{R}^M$ and Δ_M is the discrete laplacian matrix defined in (4). Roughly, this ensures that for a fixed number of sites, the stochastic model can be approximated in large population by the semi-discrete system (3). Then, as M goes to infinity, the discrete laplacian represented by Δ_M is expected to be formally replaced by the laplacian, thus the components of $\boldsymbol{u}^M$ and $\boldsymbol{v}^M$ are expected to approach the values of u and v on a uniform grid of step $\frac{1}{M}$, yielding the cross-diffusion system (1).

2.4 Statements

Our first main result is a stability estimate for the conservative SKT system (1). As far as our knowledge goes, this result is new in the context of weak solutions for the SKT system. To measure the distance between two solutions on a time interval $[0,T]$, we introduce the following norm

$$|||\cdot|||_T := \left(\|\cdot\|^2_{L^\infty([0,T];H^{-1}(\mathbb{T}^d))} + \|\cdot\|^2_{L^2(Q_T)}\right)^{1/2}. \tag{5}$$

We define also the affine functions $\mu_i\colon \mathbb{R} \to \mathbb{R}$ for $i = 1,2$, by $\mu_i(x) := d_i + a_{ij}x$ with $\{i,j\} = \{1,2\}$.

Theorem 1. *Let $T > 0$ and consider a couple $(u,v) \in L^\infty(Q_T)^2$ and $(\overline{u},\overline{v}) \in L^\infty(Q_T)^2$ of non-negative bounded weak solutions of the SKT system (1), respectively initialized by $(u_0,v_0) \in L^\infty(\mathbb{T}^d)^2$ and $(\overline{u}_0,\overline{v}_0) \in L^\infty(\mathbb{T}^d)^2$. If the following smallness condition*

$$\|\overline{u}\|_{L^\infty(Q_T)}\|\overline{v}\|_{L^\infty(Q_T)} < \frac{d_1 d_2}{a_{12}a_{21}}, \tag{6}$$

is satisfied, then we have the stability estimate

$$\begin{aligned}|||u-\overline{u}|||^2_T + |||v-\overline{v}|||^2_T \lesssim{}& \|u_0-\overline{u}_0\|^2_{H^{-1}(\mathbb{T}^d)} + \|v_0-\overline{v}_0\|^2_{H^{-1}(\mathbb{T}^d)} \\ &+ T\Big([u_0-\overline{u}_0]^2_{\mathbb{T}^d}\|\mu_1(v_0)\|_{L^1(\mathbb{T}^d)} + [v_0-\overline{v}_0]^2_{\mathbb{T}^d}\|\mu_2(u_0)\|_{L^1(\mathbb{T}^d)}\Big),\end{aligned}$$

where the constant behind $\lesssim$ depends only on $a_{ij}, d_i, \|\overline{u}\|_{L^\infty(Q_T)}, \|\overline{v}\|_{L^\infty(Q_T)}$, and $|||\cdot|||_T$ is defined by (5). In particular, if a bounded non-negative solution satisfies (6) then, there is no other bounded non-negative solution sharing the same initial data.

Remark 1. *In case of equality in the smallness condition (6), uniqueness remains but the stability estimate controls only the H^{-1} part of the $|||\cdot|||_T$ norm.*

The proof of Theorem 1 relies on a generalized duality lemma presented in Subsection 4.1 and on the concept of *dual solutions* developed in [23], for the Kolmogorov equation. The uniqueness result contained in Theorem 1 is conditional: *if* there exists a bounded (non-negative) solution $(\overline{u},\overline{v})$ satisfying (6), then it is unique in the class of bounded weak solutions. The existence of *global* bounded solutions for the SKT system is a long standing challenge in the context of cross-diffusion systems. Partial results are known, in the wake of the quest of even more regular solutions (which are in particular bounded), like [16] or [21] that we already cited. In the weak solutions setting, the paper [18] gives sufficient –yet restrictive– conditions on the coefficients of the SKT system to ensure boundedness. Since the previous results are rather constraining on the coefficients, we prefer to rely on Amann's theory [1, 2] and understand Theorem 1 as a *local* result which holds for sufficiently small initial data. Indeed, Amann's theory proves existence of regular solutions, which exist at least in a neighborhood of the origin. Starting from an initial data satisfying (6), we recover in this way a small interval on which the estimates remains valid. As the proof of Theorem 1 (which is done in Subsection 4.1) is totally insensitive to the dimension d, it is here stated in full generality. However, the remaining part of the paper (which deals with the approximation of the SKT system by stochastic processes) will focus on the case $d = 1$.

Before stating our second main result, let us comment briefly the Section 3 in which we propose a first approach to estimate the gap between the stochastic process defined by (2) and the semi-discrete system (3) on a fixed interval $[0, T]$. The methodology at stake in this paragraph, which is quite rough, allows for asymptotic quadratic closeness between these two objects, *provided that*, as $N, M \to +\infty$, we have the following

$$N \gg M^4 \exp(cM^4 T), \tag{7}$$

where c is some constant which will become more explicit in the next section. Combining this fact with the compactness result [9, Theorem 8], we obtain convergence (up to a subsequence) of our stochastic process towards a weak solution of the SKT system. The result is general in terms of parameters and form of the solution. However, the drawbacks of this approach are twofold. First, this necessitates a self-diffusion term in the system (which tends indeed to regularize the solution) in order to use the compactness result of [9]. Second, and most importantly, the scaling condition (7) involves a superexponential and time dependent number of individuals per site in order to make the law of large numbers to hold on each site and to be able to sum local estimates. As we will see, and as we can guess from the form of quadratic variations, it is too restrictive.

We propose instead a different approach, based on the discrete translation of Theorem 1. This alternative method does not rely on [9], so that self-diffusion is not needed in the system. The convergence result is obtained by means of a quantitative estimate which bounds the expectation of the $|||\cdot|||_T$-norm of the gap between the stochastic processes and the solution of the SKT system. In particular, there are no compactness tools used and the entropy of the system is not needed. Convergence is then guaranteed only with a quadratic number of individuals per site. This corresponds to the expected scaling for having local control of the stochastic process by its semi-discrete approximation, since beyond this scaling quadratic variations do not vanish. The main disadvantage of this new method is that, like for Theorem 1, it works only in a perturbative setting: it needs the existence of a small regular solution.

In order to state the following result, we need to introduce, for any integer $M \geq 1$, the discretization of the flat (one dimensional) torus $\mathbb{T}$

$$\mathbb{T}_M := \{x_1, x_2, \cdots, x_M\}, \quad \text{with } x_k = \frac{k}{M}, \text{ for } 1 \leq k \leq M. \tag{8}$$

Given a vector $\boldsymbol{u} \in \mathbb{R}^M$, classically there exists exactly one piecewise continuous function defined on $\mathbb{T}$ for which its value on each point x_k of $\mathbb{T}_M$ is given by u_k; we denote this function

$\pi_M(\boldsymbol{u})$. We adapt the same notation if instead of $\boldsymbol{u}$ one considers a vector valued map $\boldsymbol{U}$ (which could depend on the event ω or the time t for instance), so that $\pi_M(\boldsymbol{U})$ becomes a real-valued map.

Theorem 2. *In the one dimensional case $d = 1$, assume the existence of a non-negative solution $(\overline{u}, \overline{v})$ of $\mathscr{C}^1$ regularity in time and $\mathscr{C}^4$ regularity in space of the system (1), initialized by $(\overline{u}_0, \overline{v}_0) \in \mathscr{C}^4(\mathbb{T})$ and satisfying the smallness assumption (6). Consider the stochastic processes $(\boldsymbol{U}^{M,N}, \boldsymbol{V}^{M,N})$ defined by (2) and assume the existence of C_0 such that for all $M, N \in \mathbb{N}$,*

$$\|\boldsymbol{U}^{M,N}(0)\|_{1,M} + \|\boldsymbol{V}^{M,N}(0)\|_{1,M} \leq \mathrm{C}_0, \quad \textit{almost surely.} \tag{9}$$

Then, for any $(M, N) \in \mathbb{N}^2$ such that N/M^2 is large enough, for any $T > 0$,

$$\begin{aligned}&\mathbb{E}\Big[|||\pi_M\big(\boldsymbol{U}^{M,N}\big) - \overline{u}|||_T^2 + |||\pi_M\big(\boldsymbol{V}^{M,N}\big) - \overline{v}|||_T^2\Big] \\ &\lesssim \mathbb{E}\Big[\|\pi_M\big(\boldsymbol{U}^{M,N}(0)\big) - \overline{u}_0\|^2_{H^{-1}(\mathbb{T})} + \|\pi_M\big(\boldsymbol{V}^{M,N}(0)\big) - \overline{v}_0\|^2_{H^{-1}(\mathbb{T})}\Big] + M^{-4} + \frac{M^2}{N},\end{aligned} \tag{10}$$

where $|||\cdot|||_T$ is defined (5) and the symbol $\lesssim$ depends on $\mathrm{C}, T, d_i, a_{ij}, \|\overline{u}\|_{L^\infty(Q_T)}, \|\overline{v}\|_{L^\infty(Q_T)}$.

This immediately implies the following convergence for the $|||\cdot|||_T$-norm.

Corollary 1. *Under the assumptions of Theorem 2, consider an extraction function $\phi\colon \mathbb{N} \to \mathbb{N}$ such that $M^2 = \mathrm{o}(\phi(M))$. If the initial positions of the individuals are well-prepared in the sense that*

$$\mathbb{E}\Big[\|\pi_M\big(\boldsymbol{U}^{M,\phi(M)}(0)\big) - \overline{u}_0\|^2_{H^{-1}(\mathbb{T})} + \|\pi_M\big(\boldsymbol{V}^{M,\phi(M)}(0)\big) - \overline{v}_0\|^2_{H^{-1}(\mathbb{T})}\Big] \underset{M\to+\infty}{\longrightarrow} 0,$$

then for any $T > 0$, we have

$$\lim_{M\to\infty} \mathbb{E}\Big[|||\pi_M\big(\boldsymbol{U}^{M,\phi(M)}\big) - \overline{u}|||_T^2 + |||\pi_M\big(\boldsymbol{V}^{M,\phi(M)}\big) - \overline{v}|||_T^2\Big] = 0.$$

Similarly to Theorem 1, we still have a smallness condition (6) on the target solution. In some sense, this restriction is not so surprising. Even though it is a bit more hidden in this asymptotic context, the estimate (10) already contains a kind of uniqueness property for the target solution $(\overline{u}, \overline{v})$, just as the quantitative estimate of Theorem 1. At the very least, (10) states that among all possible weak solutions, $(\overline{u}, \overline{v})$ is the one who “attracts” such stochastic processes. And then, a natural way to select such a solution is to ensure uniqueness by means of sufficient regularity. These two differences come from the fact that, contrary to the previous result, Theorem 2 estimates the distance between a vector-valued stochastic process and a deterministic function which is defined on the whole torus $\mathbb{T}$. This obliges to consider corrector terms. The first one consists in the martingale term which measures locally the gap between the stochastic process and the semi-discrete deterministic approximation. Here, we observe that the estimates are sharp and the scales obtained for convergence are optimal: when $N = \phi(M)$ is of order M^2, the local behavior of the size of the population in the individual based model will remain stochastic at the limit. This limiting stochastic regime should be interesting for future works. The second correction term consists in replacing $\overline{u}$ by a piecewise continuous function in order to be able to compare it to the semi-discrete system and thus with $\pi_M(\boldsymbol{U}^{M,N})$. As a matter of fact, the proof of Theorem 2 relies on a careful translation of the (idealized) functional setting of Theorem 1 to the discrete level, together with the treatment of those corrective terms. This analysis necessitates, among other things, discrete duality lemmas including potential singular error terms. These are stated and proved in Subsection 4.4. Let us end up with a remark and perspectives. Another approach for future works would be to prove ℓ^∞ estimates for the semi-discrete system such that it is independent of M. With this one could show that the semi-discrete system is not far from verifying the limiting equation, and from here evoke the continuous version of the duality estimates in order to quantify the convergence. Also, the results obtained can be extended to the case in where the system (1) presents self-diffusion and a source term (which would correspond to adding births and deaths in the stochastic process). Last but not least, it is natural to ask to what extent the asymptotic analysis that we proposed can be generalized to higher dimension. An upper limit is fixed by the avatar of the Bramble-Hilbert lemma (which is Lemma 2). This latter remains true in higher dimension but demands a Sobolev embedding $H^2(\mathbb{T}^d) \hookrightarrow \mathscr{C}^0(\mathbb{T}^d)$, which holds only for $d = 1, 2, 3$. On the other hand, keeping in mind that solutions of the system of PDEs represent a population density in an environment, the exploration of such system in dimensions greater than 4 loses some interest. We expect that the analysis that we develop should be adaptable to dimensions 2 and 3, but this would imply a technical cost that we have preferred to avoid for now.

3 A general and rough estimate

The trajectorial representation (2) yields for each coordinate of $\boldsymbol{U}^{M,N}$

$$\begin{aligned} U_i^{M,N}(t) = U_i^{M,N}(0) - \frac{1}{N}\int_0^t \int_{\mathbb{R}_+\times\{-1,1\}} \mathbf{1}_{\rho \le \eta_{1,i}^{M,N}(s^-)} \mathcal{N}^i(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\theta) \\ + \frac{1}{N}\int_0^t \int_{\mathbb{R}_+\times\{-1,1\}} \mathbf{1}_{\rho \le \eta_{1,i-1}^{M,N}(s^-)} \mathbf{1}_{\theta=1} \mathcal{N}^{i-1}(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\theta) \\ + \frac{1}{N}\int_0^t \int_{\mathbb{R}_+\times\{-1,1\}} \mathbf{1}_{\rho \le \eta_{1,i+1}^{M,N}(s^-)} \mathbf{1}_{\theta=-1} \mathcal{N}^{i+1}(\mathrm{d}s, \mathrm{d}\rho, \mathrm{d}\theta). \end{aligned} \tag{11}$$

By compensating the Poisson point measure, we obtain the semimartingale decomposition

$$\boldsymbol{U}^{M,N}(t) = \boldsymbol{A}^{M,N}(t) + \boldsymbol{\mathcal{M}}^{M,N}(t), \tag{12}$$

where $\boldsymbol{A}^{M,N} = (A_i^{M,N})_{1\le i\le M}$ is a continuous process defined by

$$\boldsymbol{A}^{M,N}(t) = \boldsymbol{U}^{M,N}(0) + \int_0^t d_1 \Delta_M \boldsymbol{U}^{M,N}(s)\,\mathrm{d}s + \int_0^t a_{12}\Delta_M\big(\boldsymbol{U}^{M,N}(s) \odot \boldsymbol{V}^{M,N}(s)\big)\,\mathrm{d}s,$$

with Δ_M as defined in (4), and $\mathcal{M}_i^{M,N}$ is a square integrable martingale whose predictable quadratic variation is given by

$$\langle \mathcal{M}_i^{M,N} \rangle(t) = \frac{M^2}{N} \int_0^t d_1 \Big(2U_i^{M,N}(s) + U_{i+1}^{M,N}(s) + U_{i-1}^{M,N}(s) \Big) \mathrm{d}s \qquad (13)$$
$$+ \frac{M^2}{N} \int_0^t a_{12} \Big(2U_i^{M,N}(s) V_i^{M,N}(s) + U_{i+1}^{M,N}(s) V_{i+1}^{M,N}(s) + U_{i-1}^{M,N}(s) V_{i-1}^{M,N}(s) \Big) \mathrm{d}s.$$

The analogous decomposition holds for the coordinates of $(\boldsymbol{V}^{M,N}(t))_{t\geq 0}$, the second species.

Let us give first estimates of the gap between the stochastic process and its approximation in large population for a fixed number of sites. Let

$$\boldsymbol{\mathcal{U}}^{M,N}(t) = \boldsymbol{U}^{M,N}(t) - \boldsymbol{u}^M(t), \quad \boldsymbol{\mathcal{V}}^{M,N}(t) = \boldsymbol{V}^{M,N}(t) - \boldsymbol{v}^M(t).$$

Proposition 1. *We assume that there exists $C_0 > 0$ such that almost surely, for any $M, N \geq 1$,*

$$\max(\|\boldsymbol{U}^{M,N}(0)\|_{1,M}, \|\boldsymbol{V}^{M,N}(0)\|_{1,M}, \|\boldsymbol{u}^M(0)\|_{1,M}, \|\boldsymbol{v}^M(0)\|_{1,M}) \leq \mathrm{C}_0$$

and that for any $T \geq 0$, there exist $c_1, c_2 > 0$ such that for any $M, N \geq 1$,

$$\mathbb{E}\Bigg(\sup_{t\in[0,T]} \big\|\boldsymbol{\mathcal{U}}^{M,N}(t)\big\|_{2,M}^2 + \sup_{t\in[0,T]} \big\|\boldsymbol{\mathcal{V}}^{M,N}(t)\big\|_{2,M}^2 \Bigg)$$
$$\leq \Bigg(\mathbb{E}\Big(\big\|\boldsymbol{\mathcal{U}}^{M,N}(0)\big\|_{2,M}^2 + \big\|\boldsymbol{\mathcal{V}}^{M,N}(0)\big\|_{2,M}^2 \Big) + c_1 \bigg(\frac{M^2}{\sqrt{N}} + T\frac{M^3}{N} \bigg) \Bigg) e^{c_2\left(M^4 + \frac{M^2}{\sqrt{N}}\right)T},$$

where c_1 only depends on the diffusion parameters and the initial bounds and c_2 only depends on the diffusion parameters.

In particular, this estimate guarantees that the normalized stochastic process converges to the semi-discrete SKT system when the population size becomes large and the number of sites is fixed. As evoked in the introduction, this is a first step for obtaining convergence to the continuous SKT system, when the semi-discrete system itself converges to the expected continuous limit. Moreover, provided of an estimate for this last convergence, combining both of them will enable to prove convergence of the stochastic processes towards the cross-diffusion system with simultaneously the size of the population and the number of sites going to infinity. This constitutes an alternative approach for the rigorous derivation of the SKT system of [5], starting from discrete space. Both results seem to involve the same scales, with a number of individuals exponentially large compared to the inverse of the range of interaction. Our approach, in where the interaction is restricted to the same site, seems to relax the condition of small cross-diffusion parameters in [5]. Nevertheless, our main motivation in the rest of the paper is to go beyond this exponential scale and provide sharper estimates.

Proof. First, using the fact that the total number of individuals is constant along time, we observe that under our assumptions

$$\max(\|\boldsymbol{U}^{M,N}(t)\|_{1,M}, \|\boldsymbol{V}^{M,N}(t)\|_{1,M}) = \max(\|\boldsymbol{U}^{M,N}(0)\|_{1,M}, \|\boldsymbol{V}^{M,N}(0)\|_{1,M}) \leq \mathrm{C}_0, \quad (14)$$

almost surely for any $M, N \geq 1$, and

$$\max(\|\boldsymbol{u}^M(t)\|_{1,M}, \|\boldsymbol{v}^M(t)\|_{1,M}) = \max(\|\boldsymbol{u}^M(0)\|_{1,M}, \|\boldsymbol{v}^M(0)\|_{1,M}) \leq \mathrm{C}_0, \tag{15}$$

for any $M \geq 1$. Combining (12) and (3), we notice that the process $\boldsymbol{\mathcal{U}}^{M,N}(t) = \boldsymbol{U}^{M,N}(t) - \boldsymbol{u}^M(t)$ has finite variations and satisfies

$$\begin{aligned}\boldsymbol{\mathcal{U}}^{M,N}(t) = \boldsymbol{\mathcal{U}}^{M,N}(0) &+ \int_0^t d_1 \Delta_M \boldsymbol{\mathcal{U}}^{M,N}(s)\,\mathrm{d}s \\ &+ \int_0^t a_{12} \Delta_M \big(\boldsymbol{U}^{M,N}(s) \odot \boldsymbol{V}^{M,N}(s) - \boldsymbol{u}^M(s) \odot \boldsymbol{v}^M(s)\big)\,\mathrm{d}s + \boldsymbol{\mathcal{M}}^{M,N}(t).\end{aligned}$$

Consider now the square of its coordinates

$$\mathcal{U}_i^{M,N}(t)^2 = \mathcal{U}_i^{M,N}(0)^2 + \int_0^t 2\mathcal{U}_i^{M,N}(s^-)\,\mathrm{d}\mathcal{U}_i^{M,N}(s) + R_i^{M,N}(t),$$

for $i = 1, \dots, M$, where

$$\begin{aligned}R_i^{M,N}(t) &= \sum_{0<s\leq t} \Big\{\mathcal{U}_i^{M,N}(s)^2 - \mathcal{U}_i^{M,N}(s^-)^2 - 2\mathcal{U}_i^{M,N}(s^-)\big(\mathcal{U}_i^{M,N}(s) - \mathcal{U}_i^{M,N}(s^-)\big)\Big\} \\ &= \Big(\frac{1}{N}\Big)^2 \sum_{0<s\leq t} \mathbf{1}_{U_i^{M,N}(s) \neq U_i^{M,N}(s-)},\end{aligned}$$

since the jumps of $\mathcal{U}_i^{M,N}$ and $U_i^{M,N}$ are of size $1/N$. Putting the two expressions together yields

$$\begin{aligned}\mathcal{U}_i^{M,N}(t)^2 = \mathcal{U}_i^{M,N}(0)^2 &+ 2d_1 \int_0^t \mathcal{U}_i^{M,N}(s)\big(\Delta_M \boldsymbol{\mathcal{U}}^{M,N}(s)\big)_i\,\mathrm{d}s \\ &+ 2a_{12} \int_0^t \mathcal{U}_i^{M,N}(s)\big(\Delta_M \big(\boldsymbol{U}^{M,N}(s) \odot \boldsymbol{V}^{M,N}(s) - \boldsymbol{u}^M(s) \odot \boldsymbol{v}^M(s)\big)\big)_i\,\mathrm{d}s \\ &+ 2\int_0^t \mathcal{U}_i^{M,N}(s^-)\,\mathrm{d}\mathcal{M}_i^{M,N}(s) + R_i^{M,N}(t).\end{aligned}$$

Given $\boldsymbol{u} \in \mathbb{R}^M$ let us introduce the discrete gradient vector $\nabla_M^+ \boldsymbol{u} = (M(u_{i+1} - u_i))_{1\leq i\leq M}$ (recalling the periodic convention). Summing over all the sites $i \in \{1, \dots, M\}$ and using discrete integration by parts in the second and third terms of the right hand side yields

$$\begin{aligned}\big\|\boldsymbol{\mathcal{U}}^{M,N}(t)\big\|_2^2 &= \big\|\boldsymbol{\mathcal{U}}^{M,N}(0)\big\|_2^2 - 2d_1 \int_0^t \big\|\nabla_M^+ \boldsymbol{\mathcal{U}}^{M,N}(s)\big\|_2^2\,\mathrm{d}s \\ &- 2a_{12} \int_0^t \sum_{i=1}^M \big(\nabla_M^+ \boldsymbol{\mathcal{U}}^{M,N}(s)\big)_i \big(\nabla_M^+ \big(\boldsymbol{U}^{M,N}(s) \odot \boldsymbol{V}^{M,N}(s) - \boldsymbol{u}^M(s) \odot \boldsymbol{v}^M(s)\big)\big)_i\,\mathrm{d}s \\ &+ 2\sum_{i=1}^M \int_0^t \mathcal{U}_i^{M,N}(s^-)\,\mathrm{d}\mathcal{M}_i^{M,N}(s) + \big\|\boldsymbol{R}^{M,N}(t)\big\|_1.\end{aligned}$$

Dropping the second term which is negative, taking absolute value in the third term and using $2|ab| \leq |a|^2 + |b|^2$ ensures that

$$\begin{aligned}\left\|\boldsymbol{\mathcal{U}}^{M,N}(t)\right\|_2^2 \leq \left\|\boldsymbol{\mathcal{U}}^{M,N}(0)\right\|_2^2 + a_{12}\int_0^t \left\|\nabla_M^+ \boldsymbol{\mathcal{U}}^{M,N}(s)\right\|_2^2 \mathrm{d}s \\ + a_{12}\int_0^t \left\|\nabla_M^+\big(\boldsymbol{U}^{M,N}(s)\odot \boldsymbol{V}^{M,N}(s) - \boldsymbol{u}^M(s)\odot \boldsymbol{v}^M(s)\big)\right\|_2^2 \mathrm{d}s \\ + 2\sum_{i=1}^M \int_0^t \mathcal{U}_i^{M,N}(s^-)\,\mathrm{d}\mathcal{M}_i^{M,N}(s) + \left\|\boldsymbol{R}^{M,N}(t)\right\|_1.\end{aligned}$$

Let us observe that $\left\|\boldsymbol{R}^{M,N}(t)\right\|_1$ is given by the number of jumps before time t

$$\mathbb{E}\left(\left\|\boldsymbol{R}^{M,N}(t)\right\|_1\right) = 2N^{-2}\mathbb{E}(\#\{t \geq 0 : \boldsymbol{U}^{M,N}(s) \neq \boldsymbol{U}^{M,N}(s^-)\}).$$

Moreover, the total jump rate in the scaled process $\boldsymbol{\mathcal{U}}^{M,N}$, when the number of individuals of each species in site i is equal to (u_i, v_i), is

$$2M^2\sum_{i=1}^M u_i\Big(d_1 + a_{12}\frac{v_i}{N}\Big) \leq 2M^2\|\boldsymbol{u}\|_1\Big(d_1 + a_{12}\frac{\|\boldsymbol{v}\|_1}{N}\Big) \leq C_0' M^3 N(1+M),$$

where $C_0' = 2(d_1 + a_{12})C_0$, by (14). Then we get

$$\mathbb{E}\left(\left\|\boldsymbol{R}^{M,N}(t)\right\|_1\right) \leq 2C_0'\, t\,\frac{M^3}{N}(1+M).$$

Lets us now deal with the third and fourth terms. We notice that

$$\left(\nabla_M^+ \boldsymbol{\mathcal{U}}^{M,N}(s)\right)_i^2 = M^2\Big(\mathcal{U}_{i+1}^{M,N}(s) - \mathcal{U}_i^{M,N}(s)\Big)^2 \leq 2M^2\left(\mathcal{U}_{i+1}^{M,N}(s)^2 + \mathcal{U}_i^{M,N}(s)^2\right),$$

Similarly, using also $|ab - cd| \leq |a-c|b + c|b-d|$ to deal with the difference of products of positive terms, we get

$$\begin{aligned}&\left(\nabla_M^+(\boldsymbol{U}^{M,N}(s)\odot \boldsymbol{V}^{M,N}(s) - \boldsymbol{u}^M(s)\odot \boldsymbol{v}^M(s))\right)_i^2 \\ &\qquad \leq 4M^2\Big(\|\boldsymbol{u}^M(0)\|_1^2\,\mathcal{V}_{i+1}^{M,N}(s)^2 + \|\boldsymbol{u}^M(0)\|_1^2\,\mathcal{V}_i^{M,N}(s)^2 \\ &\qquad\qquad + \|\boldsymbol{V}^{M,N}(0)\|_1^2\,\mathcal{U}_{i+1}^{M,N}(s)^2 + \|\boldsymbol{V}^{M,N}(0)\|_1^2\,\mathcal{U}_i^{M,N}(s)^2\Big) \\ &\qquad \leq 4\mathrm{C}_0^2 M^4\Big(\mathcal{V}_{i+1}^{M,N}(s)^2 + \mathcal{V}_i^{M,N}(s)^2 + \mathcal{U}_{i+1}^{M,N}(s)^2 + \mathcal{U}_i^{M,N}(s)^2\Big),\end{aligned}$$

using (14) and (15). Gathering these bounds, taking supremum and then expectation gives us

$$\begin{aligned}\mathbb{E}\Big(\sup_{s\in[0,t]}\|\boldsymbol{\mathcal{U}}^{M,N}(s)\|_2^2\Big)&\\ \le \mathbb{E}\big(\|\boldsymbol{\mathcal{U}}^{M,N}(0)\|_2^2\big) &+ 4a_{12}M^2\int_0^t \mathbb{E}\big(\|\boldsymbol{\mathcal{U}}^{M,N}(s)\|_2^2\big)\,\mathrm{d}s\\ &+ 8\mathrm{C}_0^2 a_{12}M^4\left(\int_0^t \mathbb{E}\big(\|\boldsymbol{\mathcal{V}}^{M,N}(s)\|_2^2\big)\,\mathrm{d}s + \int_0^t \mathbb{E}\big(\|\boldsymbol{\mathcal{U}}^{M,N}(s)\|_2^2\big)\,\mathrm{d}s\right)\\ &+ 2\sum_{i=1}^{M}\mathbb{E}\Big(\sup_{s\in[0,t]}\int_0^s \mathcal{U}_i^{M,N}(r^-)\,\mathrm{d}\mathcal{M}_i^{M,N}(r)\Big) + 2C_0' T\,\frac{M^3}{N}(1+M).\end{aligned}$$

For the martingale part, we use Cauchy-Schwarz and Burkholder-Davis-Gundy inequalities which together with (13) and (14) yield

$$\begin{aligned}\mathbb{E}\Big(\sup_{s\in[0,t]}\int_0^s \mathcal{U}_i^{M,N}(r^-)\,\mathrm{d}\mathcal{M}_i^{M,N}(r)\Big)^2&\\ &\le \mathbb{E}\Big(\sup_{s\in[0,t]}\Big|\int_0^s \mathcal{U}_i^{M,N}(r^-)\,\mathrm{d}\mathcal{M}_i^{M,N}(r)\Big|^2\Big)\\ &\le \mathbb{E}\Big(\int_0^t \mathcal{U}_i^{M,N}(r^-)^2\,\mathrm{d}\langle \mathcal{M}_i^{M,N}\rangle(r)\Big)\\ &\le 2\frac{M^2}{N}\mathbb{E}\Big(\big\|\boldsymbol{U}^{M,N}(0)\big\|_1\Big(d_1 + a_{12}\,\big\|\boldsymbol{V}^{M,N}(0)\big\|_1\Big)\int_0^t \mathcal{U}_i^{M,N}(s)^2\,\mathrm{d}s\Big)\\ &\le C_0'\frac{M^3}{N}(1+M)\int_0^t \mathbb{E}\Big(\mathcal{U}_i^{M,N}(s)^2\Big)\,\mathrm{d}s.\end{aligned}$$

Using that $\sqrt{1+x}\le 1+x$ for all $x\ge 0$, we obtain

$$\mathbb{E}\Big(\sup_{s\in[0,t]}\int_0^s \mathcal{U}_i^{M,N}(r^-)\,\mathrm{d}\mathcal{M}_i^{M,N}(r)\Big) \le \sqrt{2C_0''}\,\frac{M^2}{\sqrt{N}}\left(1+\int_0^t \mathbb{E}\Big(\mathcal{U}_i^{M,N}(s)^2\Big)\,\mathrm{d}s\right).$$

Putting everything together and using again (14) yields

$$\begin{aligned}\mathbb{E}\Big(\sup_{s\in[0,t]}\|\boldsymbol{\mathcal{U}}^{M,N}(s)\|_2^2\Big) &\le \mathbb{E}\big(\|\boldsymbol{\mathcal{U}}^{M,N}(0)\|_2^2\big) + 2\sqrt{2C_0''}\,\frac{M^3}{\sqrt{N}} + 2C_0' T\frac{M^4}{N}\\ &\quad+\left(8\mathrm{C}_0 a_{12}M^4 + 2\sqrt{2C_0''}\,\frac{M^2}{\sqrt{N}}\right)\int_0^t \mathbb{E}\Big(\sup_{r\in[0,s]}\|\boldsymbol{\mathcal{U}}^{M,N}(r)\|_2^2\Big)\,\mathrm{d}s\\ &\quad+ 8\mathrm{C}_0 a_{12}M^4\int_0^t \mathbb{E}\Big(\sup_{r\in[0,s]}\|\boldsymbol{\mathcal{V}}^{M,N}(r)\|_2^2\Big)\,\mathrm{d}s,\end{aligned}$$

for some $C_0''>0$. In a similar way we can obtain analogous bounds for $\boldsymbol{V}^{M,N}$. Adding the two inequalities and then applying Gronwall's lemma leads us to the desired conclusion. □

The proof above is general in the sense that we have no conditions on the limiting SKT system. But as explained in the previous sections, convergence with a large number of sites requires a superexponential number of individuals per site. The bounds in the previous proof are indeed not sharp at several steps. In particular, we have controlled the quadratic terms by bounding the local size of one species by the total number of individuals, which is fixed and thus controlled quantity. Similarly, the gradient term has been dominated by brute force since we have summed the components. To go beyond these estimates and deal with the quadratic term, we develop a duality approach. This will bring stability property and allow us to compare the terms involved in the stochastic process to those of the targeted SKT limit. The stochastic process will then appear as a stable perturbation of this SKT system.

4 Duality estimates

4.1 The continuous setting

The duality lemma is a tool first introduced by Martin, Pierre and Schmitt [22, 25], in the context of reaction-diffusion systems. It consists in an *a priori* estimate for solutions of the Kolmogorov equation. The strength of the estimate is that it requires very low regularity on the diffusivity (merely integrability), which allows its use when dealing with rather weak solutions. We propose below a small generalization of the duality lemma, which was suggested in [23, Remark 7]. As a matter of fact, we will not directly use the duality lemma presented in this paragraph, but rather translate it in a discrete setting (see Subsection 4.4 below). The purpose of this paragraph is then twofold. First, prove Theorem 1. Second, explain, avoiding several technicalities inherent to the discrete setting, the core ideas that will be used in Subsection 4.4. Below, we call a *weak* solution to a solution in the distributional sense. During (and only in) this whole paragraph, we work in arbitrary dimension d.

Lemma 1. *Consider $\mu \in L^\infty(Q_T)$ such that $\alpha := \inf_{Q_T} \mu > 0$, $z_0 \in H^{-1}(\mathbb{T}^d)$ and $f \in L^2(Q_T)$. Then, there exists a unique $z \in L^2(Q_T)$ that solves weakly the Kolmogorov equation*

$$\begin{cases} \partial_t z - \Delta(\mu z) = \Delta f, \\ z(0,\cdot) = z_0. \end{cases} \tag{16}$$

Furthermore, this solution z belongs to $\mathscr{C}([0,T]; H^{-1}(\mathbb{T}^d))$ and satisfies the duality estimate

$$\|z(T)\|^2_{H^{-1}(\mathbb{T}^d)} + \int_{Q_T} \mu z^2 \le \|z_0\|^2_{H^{-1}(\mathbb{T}^d)} + [z_0]^2_{\mathbb{T}^d} \int_{Q_T} \mu + \frac{1}{\alpha} \int_{Q_T} f^2. \tag{17}$$

Remark 2. *This duality estimate is stronger than the one stated in [23]: it contains a (singular) source term and allows a uniform-in-time control of the $H^{-1}(\mathbb{T}^d)$ norm.*

Proof. The proof of existence and uniqueness is exactly the same as [23, Theorem 3]: following the naming of this article, z is the unique dual solution of (16). For this z, the regularity $\mathscr{C}([0,T]; H^{-1}(\mathbb{T}^d))$ is obtained classically from the belongings $z \in L^2(Q_T)$ and $\partial_t z \in$

$L^2([0,T]; H^{-2}(\mathbb{T}^d))$, which come from the equation itself. We can thus focus here on the duality estimate which needs to be proven only in the case when every function involved in (17) is smooth, in the sense that they are $\mathscr{C}^\infty$. Indeed, the assumptions on the data give us a smooth sequence $(\mu^n, z_0^n, f_n)_{n\in\mathbb{N}}$ converging to (μ, z_0, f) in $L^1(Q_T) \times H^{-1}(\mathbb{T}^d) \times L^2(Q_T)$, with a uniform bound for the first component. Let's call $(z^n)_{n\in\mathbb{N}}$ the corresponding sequence of solutions. Note that, by parabolic regularity, the z^n's are also smooth. Then, if the duality estimate (17) is proved in the smooth setting, we get (up to some subsequence) weak(-$\star$) convergence of $(z^n)_{n\in\mathbb{N}}$, in $L^\infty([0,T]; H^{-1}(\mathbb{T}^d)) \cap L^2(Q_T)$. But, by uniqueness of the target equation, the only possible limit point is precisely z, the solution of (16). The whole sequence $(z^n)_n$ converges therefore weakly(-$\star$) towards z, and (17) is recovered by the usual semi-continuity argument for weak convergence.

So, without loss of generality, we assume now that μ, z_0, f and z are smooth. This allows to justify rigorously the following computations. For any function w defined on $\mathbb{T}^d$ and having zero average there exists a unique function ϕ of zero average satisfying $\Delta\phi = w$ (which is easily seen *via* the Fourier coefficients). In particular, for any $t \in [0,T]$ there exists a unique $\phi(t)$ of vanishing mean such that $-\Delta\phi(t) = z(t) - [z(t)]_{\mathbb{T}^d}$. By integrating the Kolmogorov equation we get

$$\frac{\mathrm{d}}{\mathrm{d}t}[z(t)]_{\mathbb{T}^d} = 0,$$

so that $[z(t)]_{\mathbb{T}^d} = [z_0]_{\mathbb{T}^d}$ and $-\partial_t\Delta\phi = \partial_t z$. In particular, we have by integration by parts

$$\int_{\mathbb{T}^d} \phi(t)\,\partial_t z(t) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{T}^d} |\nabla\phi(t)|^2.$$

Therefore, multiplying equation (16) by ϕ and using integration by parts

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{T}^d} |\nabla\phi(t)|^2 + \int_{\mathbb{T}^d} \mu z(z - [z_0]_{\mathbb{T}^d}) = -\int_{\mathbb{T}^d} (z - [z_0]_{\mathbb{T}^d}) f.$$

Integrating in time and using Young's inequality for the right hand side, we get

$$\begin{aligned} \frac{1}{2}\int_{\mathbb{T}^d} |\nabla\phi(T)|^2 + \int_{Q_T} \mu z^2 \le \int_{Q_T} \mu z [z_0]_{\mathbb{T}^d} + \frac{1}{2}\int_{\mathbb{T}^d} |\nabla\phi(0)|^2 \\ + \frac{1}{2}\int_{Q_T} (z - [z_0]_{\mathbb{T}^d})^2\mu + \frac{1}{2}\int_{Q_T} \frac{f^2}{\mu}, \end{aligned}$$

and thus, using $\mu \ge \alpha > 0$,

$$\int_{\mathbb{T}^d} |\nabla\phi(T)|^2 + \int_{Q_T} \mu z^2 \le \int_{\mathbb{T}^d} |\nabla\phi(0)|^2 + [z_0]_{\mathbb{T}^d}^2 \int_{Q_T} \mu + \frac{1}{\alpha}\int_{Q_T} f^2.$$

Noticing that $\|z(t)\|_{\dot H^{-1}(\mathbb{T}^d)} = \|z(t) - [z_0]_{\mathbb{T}^d}\|_{H^{-1}(\mathbb{T}^d)} = \|\nabla\phi(t)\|_2$, once we add $[z_0]_{\mathbb{T}^d}$ to each side of the inequality to get the full $H^{-1}(\mathbb{T}^d)$ norms, the proof is over. □

In Subsection 4.4, we will give (in the discrete setting) variants of the previous duality lemma which include in the r.h.s. some error term, which is possibly singular in the time variable. Being able to take into account those error terms will be crucial in the final asymptotic limit studied in Section 5. However, already in its current form, the previous duality lemma is a valuable piece of information. We highlight this with an application of this lemma: the proof of Theorem 1, which applies to the conservative SKT system (1) that we consider here with (u_0, v_0) as initial data. We recall the definition of the affine functions $\mu_i(x) := d_i + a_{ij}x$ for $i, j = 1, 2$, so that (1) rewrites

$$\begin{cases} \partial_t u - \Delta(\mu_1(v)u) = 0, \\ \partial_t v - \Delta(\mu_2(u)v) = 0. \end{cases}$$

In particular, we recover the framework of Lemma 1, as soon as v and u are bounded and non-negative.

Proof of Theorem 1. Let's introduce $z := \overline{u} - u$ and $w := \overline{v} - v$, so that, by subtraction

$$\partial_t z - \Delta(\mu_1(v)z) = \Delta f,$$
$$\partial_t w - \Delta(\mu_2(u)w) = \Delta g,$$

where $f := a_{12}\overline{u}(\overline{v} - v)$ and $g := a_{21}\overline{v}(\overline{u} - u)$. Since u and v are bounded and non-negative, we recover the structure of Lemma 1 and we get

$$\|z(T)\|^2_{H^{-1}(\mathbb{T}^d)} + d_1 \int_{Q_T} z^2 \leq \|z_0\|^2_{H^{-1}(\mathbb{T}^d)} + [z_0]^2_{\mathbb{T}^d} \int_{Q_T} \mu_1(v) + \frac{a_{12}^2}{d_1} \|\overline{u}\|^2_{L^\infty(Q_T)} \int_{Q_T} w^2,$$
$$\|w(T)\|^2_{H^{-1}(\mathbb{T}^d)} + d_2 \int_{Q_T} w^2 \leq \|w_0\|^2_{H^{-1}(\mathbb{T}^d)} + [w^0]^2_{\mathbb{T}^d} \int_{Q_T} \mu_2(u) + \frac{a_{21}^2}{d_2} \|\overline{v}\|^2_{L^\infty(Q_T)} \int_{Q_T} z^2,$$

since $\inf_{Q_T} \mu_i \geq d_i$, $|f| \leq a_{12}|w|\|\overline{u}\|_{L^\infty(Q_T)}$ and $|g| \leq a_{21}|z|\|\overline{v}\|_{L^\infty(Q_T)}$. By combining the two inequalities we infer

$$\begin{aligned} \|z(T)\|^2_{H^{-1}(\mathbb{T}^d)} + d_1 \int_{Q_T} z^2 \leq \|z_0\|^2_{H^{-1}(\mathbb{T}^d)} + [z_0]^2_{\mathbb{T}^d} \int_{Q_T} \mu_1(v) \\ + \frac{a_{12}^2}{d_1 d_2} \|\overline{u}\|^2_{L^\infty(Q_T)} \Big(\|w_0\|^2_{H^{-1}(\mathbb{T}^d)} + [w_0]^2_{\mathbb{T}^d} \int_{Q_T} \mu_2(u) \Big) \\ + d_1 \left(\frac{a_{12}a_{21}}{d_1 d_2} \right)^2 \|\overline{u}\|^2_{L^\infty(Q_T)} \|\overline{v}\|^2_{L^\infty(Q_T)} \int_{Q_T} z^2. \end{aligned}$$

In particular, if we want to absorb the last term of the r.h.s. in the l.h.s. the inequality that we need is exactly the smallness condition (6). If the later is satisfied, and if we allow the symbol $\lesssim$ to depend on $d_i, a_{ij}, \|\overline{u}\|_{L^\infty(Q_T)}$ and $\|\overline{v}\|_{L^\infty(Q_T)}$, we have actually established

$$\begin{aligned} \|z(T)\|^2_{H^{-1}(\mathbb{T}^d)} + \int_{Q_T} z^2 \lesssim \|z_0\|^2_{H^{-1}(\mathbb{T}^d)} + \|w^0\|^2_{H^{-1}(\mathbb{T}^d)} \\ + [z_0]^2_{\mathbb{T}^d} \int_{Q_T} \mu_1(v) + [w_0]^2_{\mathbb{T}^d} \int_{Q_T} \mu_2(u). \end{aligned}$$

Since the previous computation is still valid replacing T by any $t \in [0,T]$, we have in fact

$$|||z|||_T^2 \lesssim \|z_0\|^2_{H^{-1}(\mathbb{T}^d)} + \|w_0\|^2_{H^{-1}(\mathbb{T}^d)} + [z_0]^2_{\mathbb{T}^d} \int_{Q_T} \mu_1(v) + [w_0]^2_{\mathbb{T}^d} \int_{Q_T} \mu_2(u).$$

Exchanging the roles $(z,\overline{u},\overline{v},u,v) \leftrightarrow (w,\overline{v},\overline{u},v,u)$, the previous right hand side remains unchanged: we have exactly the same estimate for $\|w\|_T^2$ on the left hand side. The proof is over once we notice that $\int_{Q_T} \mu_1(v) = T\int_{\mathbb{T}^d} \mu_1(v_0)$ and $\int_{Q_T} \mu_2(v) = T\int_{\mathbb{T}^d} \mu_2(u_0)$, since the space integrals of u and v are conserved through time. □

4.2 Reconstruction operators

As explained in the previous paragraph, we plan now to transfer the previous duality and stability estimates into a discrete setting. The purpose is to be able to use these results on the semi-discrete system (3). We will have to manipulate several norms on $\mathbb{R}^M$, reminiscent of classical function spaces of the continuous variable. As the number of points M of the discretization will be sent to infinity, it will be crucial to have estimates which do not depend on this parameter. In particular, the following notion of uniform equivalence will be relevant.

Definition 1. *Given norms $P_{1,M}$ and $P_{2,M}$ on $\mathbb{R}^M$, we say that $P_{1,M}$ and $P_{2,M}$ are uniformly equivalent if there exists $\alpha,\beta > 0$ such that*

$$\forall M \in \mathbb{N}, \quad \forall \boldsymbol{u} \in \mathbb{R}^M, \quad \alpha P_{1,M}(\boldsymbol{u}) \leq P_{2,M}(\boldsymbol{u}) \leq \beta P_{1,M}(\boldsymbol{u}).$$

If this is satisfied, we write $P_{1,M} \sim P_{2,M}$.

Given a discretization like (8), there are several ways to build a function defined on the whole torus $\mathbb{T}$. The generic approach is to fix a profile θ (generally compactly supported) and consider

$$x \mapsto \sum_{k=1}^{M} \theta\left(M(x-x_k)\right) u_k. \tag{18}$$

Definition 2. *For $\boldsymbol{u} \in \mathbb{R}^M$ and $\theta := \mathbf{1}_{[-1,0]}$, the function defined by (18) is a step function that we denote $\sigma_M(\boldsymbol{u})$. For $\boldsymbol{u} \in \mathbb{R}^M$ and $\theta(z) := (1-|z|)^+$, the function defined by (18) is a piecewise linear function that we denote $\pi_M(\boldsymbol{u})$. The corresponding vector space of functions (step and continuous piecewise linear functions respectively) are denoted*

$$\mathfrak{s}_M := \left\{\sigma_M(\boldsymbol{u}) : \boldsymbol{u} \in \mathbb{R}^M\right\} \quad \text{and} \quad \mathfrak{p}_M := \left\{\pi_M(\boldsymbol{u}) : \boldsymbol{u} \in \mathbb{R}^M\right\}.$$

If $t \mapsto \boldsymbol{u}(t)$ is a map from $[0,T]$ to $\mathbb{R}^M$, we simply denote by $\sigma_M(\boldsymbol{u})$ and $\pi_M(\boldsymbol{u})$ the respective maps from $[0,T]$ to $\mathfrak{s}_M$ and $\mathfrak{p}_M$ respectively.

Proposition 2. *For $\boldsymbol{u} \in \mathbb{R}^M$ we have $\|\boldsymbol{u}\|_\infty = \|\sigma_M(\boldsymbol{u})\|_{L^\infty(\mathbb{T})} = \|\pi_M(\boldsymbol{u})\|_{L^\infty(\mathbb{T})}$ and for $p < \infty$ we have $\|\boldsymbol{u}\|_{p,M} = \|\sigma_M(\boldsymbol{u})\|_{L^p(\mathbb{T})} \geq \|\pi_M(\boldsymbol{u})\|_{L^p(\mathbb{T})}$. Furthermore, the equivalence $\|\sigma_M(\cdot)\|_{L^p(\mathbb{T})} \sim \|\pi_M(\cdot)\|_{L^p(\mathbb{T})}$ holds on the positive cone $\mathbb{R}^M_+$.*

Proof. We first notice $\mathbf{1}_{[-1,0]} \star \mathbf{1}_{[0,1]}(x) = \int_{-1}^{0} \mathbf{1}_{[0,1]}(x-y)\,\mathrm{d}y = (1-|x|)^{+}$. In particular, we infer for $\varphi(x) = (1-|x|)^{+}$

$$
\begin{aligned}
\varphi_{k,M}(x) \coloneqq \varphi\left(M(x-x_k)\right) &= \int \mathbf{1}_{[-1,0]}\left(M(x-x_k)-y\right)\mathbf{1}_{[0,1]}(y)\,\mathrm{d}y \\
&= N \int \mathbf{1}_{[-1,0]}\left(M(x-z-x_k)\right)\mathbf{1}_{[0,1]}(Mz)\,\mathrm{d}z \\
&= \theta_{k,M} \star \rho_M(x).
\end{aligned}
$$

We have thus established $\pi_M(\boldsymbol{u}) = \sigma_M(\boldsymbol{u}) \star \rho_M$ where $(\rho_M)_M$ is an approximation of the identity. Therefore, we have $\|\pi_M(\boldsymbol{u})\|_{L^p(\mathbb{T})} \le \|\sigma_M(\boldsymbol{u})\|_{L^p(\mathbb{T})}$.

Conversely, assume $\boldsymbol{u} \ge 0$. By definition we have

$$
\pi_M(\boldsymbol{u}) = \sum_{k=1}^{M} u_k \varphi_{k,M},
$$

with $\varphi_{k,M}(x) = \varphi(M(x-x_k))$ and $\varphi(x) = (1-|x|)^{+}$. Recall that for any vector $\boldsymbol{w} \in \mathbb{R}^M$, one has $M^{1/p}\|\boldsymbol{w}\|_{p,M} \le M\|\boldsymbol{w}\|_{1,M}$. In particular, using $u_k \ge 0$, we infer at any point $x \in \mathbb{T}$

$$
\pi_M(\boldsymbol{u})(x) = \sum_{k=1}^{M} u_k \varphi_{k,M} \ge \left(\sum_{k=1}^{M} u_k^p \varphi_{k,M}^p(x)\right)^{1/p},
$$

from where we conclude

$$
\|\pi_M(\boldsymbol{u})\|_{L^p(\mathbb{T})} \ge M\|\boldsymbol{u}\|_{p,M}^p \|\varphi\|_{L^p(\mathbb{T})}^p = \|\sigma_M(\boldsymbol{u})\|_{L^p(\mathbb{T})}^p \frac{2}{p+1},
$$

using that $\|\varphi_{k,M}\|_{L^p(\mathbb{T})}^p = \frac{1}{M}\|\varphi\|_{L^p(\mathbb{T})}^p = \frac{1}{M}\frac{2}{p+1}$. □

We end this paragraph with an estimate that belongs to the folklore of the finite element method and omit the proof. It is usually proved using the Bramble-Hilbert lemma, but since here we focus here on the one dimensional case, it is also possible to give a direct, elementary proof.

Lemma 2. *For $\varphi \in H^2(\mathbb{T})$ and $M \in \mathbb{N}^*$ there exists a unique $\iota_M(\varphi) \in \mathfrak{p}_M$ matching the values of φ on the grid $(x_k)_{1\le k\le M}$. It satisfies*

$$
\begin{aligned}
\|\varphi - \iota_M(\varphi)\|_{\dot{H}^{-1}(\mathbb{T})} &\lesssim M^{-2}\|\varphi\|_{\dot{H}^2(\mathbb{T})}, \\
\|\varphi - \iota_M(\varphi)\|_{L^2(\mathbb{T})} &\lesssim M^{-2}\|\varphi\|_{\dot{H}^2(\mathbb{T})}, \\
\|\varphi - \iota_M(\varphi)\|_{\dot{H}^1(\mathbb{T})} &\lesssim M^{-1}\|\varphi\|_{\dot{H}^2(\mathbb{T})},
\end{aligned}
$$

where the symbol $\lesssim$ means that the inequality holds up to a constant independent of φ and M.

4.3 Prerequisites on the discrete laplacian matrix

We give in this paragraph several useful properties linked to the discrete periodic laplacian matrix introduced in (4). This matrix Δ_M is not invertible: we have $\mathrm{Ker}(\Delta_M) = \mathrm{Span}_{\mathbb{R}}(\mathbf{1}_M)$ and $\mathrm{Ran}(\Delta_M) = \mathrm{Ker}(\Delta_M)^\perp = \{\boldsymbol{u} \in \mathbb{R}^M : [\boldsymbol{u}]_M = 0\}$. We refer to Subsection 1.2 for the definition of $\mathbf{1}_M$ and $[\cdot]_M$.

Definition 3. *For $\boldsymbol{u} \in \mathrm{Ran}(\Delta_M)$ there exists a unique $\Phi \in \mathrm{Ran}(\Delta_M)$ such that $\boldsymbol{u} = \Delta_M \Phi$. By a small abuse of notation we write then $\Phi = \Delta_M^{-1}\boldsymbol{u}$.*

Proposition 3. *The matrix $-\Delta_M$ is symmetric non-negative and admits therefore a unique symmetric non-negative square root that we denote $\sqrt{-\Delta_M}$.*

Proof. The proof is standard and we simply note that the spectrum of $-\Delta_M$ is given by

$$\left\{M^2\left(2-2\cos\left(\frac{2\pi k}{M}\right)\right) : 0 \le k \le M-1\right\} = \left\{4M^2\sin^2\left(\frac{\pi k}{M}\right) : 0 \le k \le M-1\right\} \subset \mathbb{R}_+,$$

which establishes the non-negativeness. □

Proposition 4. *For any $\mathbf{\Phi} \in \mathbb{R}^M$ we have the estimate $\|\mathbf{\Phi} - [\mathbf{\Phi}]_M\|_{2,M} \le \|\Delta_M \mathbf{\Phi}\|_{2,M}$.*

Remark 3. *This is the discrete counterpart of the following consequence of the Poincaré-Wirtinger inequality $\|\varphi - [\varphi]_{\mathbb{T}}\|_{L^2(\mathbb{T})} \lesssim \|\Delta\varphi\|_{L^2(\mathbb{T})}$, for $\varphi \in H^2(\mathbb{T})$.*

Proof. Using the identity $\sin(\pi k/M) = \sin(\pi(M-k)/M)$, the spectrum $-\Delta_M$ that we identified in the proof of Proposition 3 rewrites

$$\left\{4M^2\sin^2\left(\frac{\pi k}{M}\right) : 0 \le k \le \frac{M-1}{2}\right\}.$$

In particular, using the inequality $\sin(x) \ge \frac{2}{\pi}x$ valid on $[0,\pi/2]$ we see that apart from 0 all the eigenvalues of $-\Delta_M$ are lower-bounded by 16. $-\Delta_M$ being symmetric, its diagonalization can be written in an orthonormal basis of $\mathbb{R}^M$ that we denote $(\boldsymbol{w}_k)_{0\le k\le M-1}$, with $\boldsymbol{w}_0$ being the (only) element of this set belonging to $\mathrm{Ker}(\Delta_M)$. We have therefore

$$\|\mathbf{\Phi} - [\mathbf{\Phi}]_M\|_{2,M}^2 = \frac{1}{M}\sum_{k=1}^{M-1} |(\mathbf{\Phi}|\boldsymbol{w}_k)|^2 \le \frac{1}{M}\frac{1}{16^2}\sum_{k=1}^{M-1} \lambda_k^2 |(\mathbf{\Phi}|\boldsymbol{w}_k)|^2 = \frac{1}{16^2}\|\Delta_M \mathbf{\Phi}\|_{2,M}^2. \quad \square$$

Before introducing an analog of the negative Sobolev norm, we recall a standard computation linked with the Lagrange finite elements method for which we need to introduce the following matrix

$$B_M := \begin{pmatrix} \frac{2}{3} & \frac{1}{6} & 0 & \cdots & \frac{1}{6} \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ \frac{1}{6} & \cdots & 0 & \frac{1}{6} & \frac{2}{3} \end{pmatrix}. \tag{19}$$

Proposition 5. *For $\boldsymbol{w} \in \mathbb{R}^M$ we have*

$$- (\boldsymbol{w}|\Delta_M \boldsymbol{w})_M = \int_{\mathbb{T}} |\nabla \pi_M(\boldsymbol{w})(x)|^2 \, \mathrm{d}x, \tag{20}$$

where we recall that $(\cdot|\cdot)_M$ denotes the rescaled inner product on $\mathbb{R}^M$ (see Subsection 1.2). Furthermore, for any $\boldsymbol{u} \in \mathbb{R}^M$ we have

$$B_M \boldsymbol{u} = -\Delta_M \boldsymbol{w} \iff \forall \psi \in \mathfrak{p}_M, \int_{\mathbb{T}} \psi(x) \, \pi_M(\boldsymbol{u})(x) \, \mathrm{d}x = \int_{\mathbb{T}} \nabla\psi(x) \cdot \nabla \pi_M(\boldsymbol{w})(x) \, \mathrm{d}x. \tag{21}$$

Proof. $\mathfrak{p}_M$ is the vector space spanned by the functions $\varphi_{k,M}(x) := \varphi(M(x - x_k))$ where $\varphi(x) := (1 - |x|)^+$, so the r.h.s. of the equivalence (21) boils down to

$$\int_{\mathbb{T}} \varphi_{k,M}(x) \pi_M(\boldsymbol{u})(x) \, \mathrm{d}x = \int_{\mathbb{T}} \nabla \varphi_{k,M}(x) \cdot \nabla \pi_M(\boldsymbol{w})(x) \, \mathrm{d}x,$$

for $k \in \{1, \ldots, M\}$, and one checks that

$$\int_{\mathbf{T}} \varphi_{i,M}(x) \varphi_{j,M}(x) \, \mathrm{d}x = \frac{1}{M} \left(\frac{2}{3} \mathbf{1}_{i=j} + \frac{1}{6} \mathbf{1}_{|i-j|=1} \right),$$
$$\int_{\mathbf{T}} \nabla \varphi_{i,M}(x) \cdot \nabla \varphi_{j,M}(x) \, \mathrm{d}x = M(2\mathbf{1}_{i=j} - \mathbf{1}_{|i-j|=1}),$$

where the equality $|i-j| = 1$ has to be understood modulo M. Expanding $\pi_M(\boldsymbol{u})$ and $\pi_M(\boldsymbol{w})$ on the basis $(\varphi_{k,M})_{1 \le k \le M}$, we get the equivalence (21). Formula (20) is obtained in the same fashion, expanding $\pi_M(\boldsymbol{w})$ on the basis. □

We observe that $\boldsymbol{u} \mapsto -(\boldsymbol{u}|\Delta_M^{-1} \boldsymbol{u})_M$ is non-negative, due to the symmetry and non-negativity of $-\Delta_M$ (see Proposition 3). For $\boldsymbol{u} \in \mathbb{R}^M$, recalling that $\widetilde{\boldsymbol{u}} = \boldsymbol{u} - [\boldsymbol{u}]_M \mathbf{1}_M$, we have then $-(\widetilde{\boldsymbol{u}}|\Delta_M^{-1} \widetilde{\boldsymbol{u}})_M \ge 0$. This enables us to introduce the following norm $\| \cdot \|_{-1,M}$, which is a discrete counterpart of the $H^{-1}(\mathbb{T})$ norm.

Definition 4. *For $\boldsymbol{u} \in \mathbb{R}^M$, we define*

$$\|\boldsymbol{u}\|_{-1,M} := \sqrt{-(\widetilde{\boldsymbol{u}}|\Delta_M^{-1} \widetilde{\boldsymbol{u}})_M + [\boldsymbol{u}]_M^2}.$$

This is a norm on $\mathbb{R}^M$.

Proposition 6. *We have the equivalence*

$$M \|\pi_M(\cdot)\|_{H^{-1}(\mathbb{T})} + \|\pi_M(\cdot)\|_{L^2(\mathbb{T})} \sim M \| \cdot \|_{-1,M} + \|\pi_M(\cdot)\|_{L^2(\mathbb{T}^d)}. \tag{22}$$

Moreover for any $\boldsymbol{u} \in \mathbb{R}^M$,

$$\|\boldsymbol{u}\|_{-1,M} \le \|\boldsymbol{u}\|_{2,M}. \tag{23}$$

Remark 4. *The above definition is reminiscent of the equality*

$$\|\varphi - [\varphi]_{\mathbb{T}}\|^2_{H^{-1}(\mathbb{T})} = -\int_{\mathbb{T}} (\varphi - [\varphi]_{\mathbb{T}})\psi,$$

where ψ is the unique solution of $-\Delta\psi = \varphi - [\varphi]_{\mathbb{T}}$.

Proof. We first observe the uniform equivalences

$$\begin{aligned} \|\pi_M(\boldsymbol{u})\|_{L^2(\mathbb{T})} &\sim \|\pi_M(\widetilde{\boldsymbol{u}})\|_{L^2(\mathbb{T})} + |[\boldsymbol{u}]_M|, \\ \|\pi_M(\boldsymbol{u})\|_{H^{-1}(\mathbb{T})} &\sim \|\pi_M(\widetilde{\boldsymbol{u}})\|_{H^{-1}(\mathbb{T})} + |[\boldsymbol{u}]_M|, \\ \|\boldsymbol{u}\|_{-1,M} &\sim \|\widetilde{\boldsymbol{u}}\|_{-1,M} + |[\boldsymbol{u}]_M|. \end{aligned}$$

Without loss of generality we can therefore establish the uniform equivalence (22) under the assumption $[\boldsymbol{u}]_M = 0$.

We have $\|\boldsymbol{u}\|^2_{-1,M} = -(\boldsymbol{u}|\Delta_M^{-1}\boldsymbol{u})_M = -(\Delta_M\boldsymbol{\Phi}, \boldsymbol{\Phi})_M$ where $\boldsymbol{\Phi} := -\Delta_M^{-1}\boldsymbol{u}$. Thanks to Proposition 5 we have therefore

$$\|\boldsymbol{u}\|^2_{-1,M} = \|\nabla\pi_M(\boldsymbol{\Phi})\|^2_{L^2(\mathbb{T})}. \tag{24}$$

The matrix B_M defined by (19) satisfies $6B_M = M^{-2}\Delta_M + 6\mathrm{I}_M$, so it commutes with Δ_M. In particular, the equation $\boldsymbol{u} = -\Delta_M\boldsymbol{\Phi}$ is strictly equivalent to

$$B_M\boldsymbol{u} = -\Delta_M\boldsymbol{w},$$

where $\boldsymbol{w} := B_M\boldsymbol{\Phi}$. We obtain from Proposition 5 that this last equation is exactly equivalent to

$$\forall\psi \in \mathfrak{p}_M, \quad \int_{\mathbb{T}} \psi(x)\,\pi_M(\boldsymbol{u})(x)\,\mathrm{d}x = \int_{\mathbb{T}} \nabla\psi(x)\cdot\nabla\pi_M(\boldsymbol{w})(x)\,\mathrm{d}x.$$

Since we assumed $[\boldsymbol{u}]_M = 0$, we have also $[\pi_M(\boldsymbol{u})]_{\mathbb{T}} = 0$ and we can therefore solve $-\Delta\varphi_M = \pi_M(\boldsymbol{u})$, for a unique $\varphi_M \in \dot{H}^2(\mathbb{T})$. We have then, by integration by parts,

$$\forall\psi \in \mathfrak{p}_M, \quad \int_{\mathbb{T}} \psi(x)\,\pi_M(\boldsymbol{u})(x)\,\mathrm{d}x = \int_{\mathbb{T}} \nabla\psi(x)\cdot\nabla\varphi_M(x)\,\mathrm{d}x.$$

In particular, we have established

$$\forall\psi \in \mathfrak{p}_M, \quad \int_{\mathbb{T}} \nabla\psi(x)\cdot(\nabla\pi_M(\boldsymbol{w})(x) - \nabla\varphi_M(x))\,\mathrm{d}x = 0,$$

and this equality holds in particular for $\psi = \pi_M(\boldsymbol{w})$. We deduce that for each $\psi \in \mathfrak{p}_M$

$$\begin{aligned} &\int_{\mathbb{T}} |\nabla\pi_M(\boldsymbol{w})(x) - \nabla\varphi_M(x)|^2\,\mathrm{d}x \\ &\quad= \int_{\mathbb{T}} (\nabla\pi_M(\boldsymbol{w})(x) - \nabla\varphi_M(x) + \nabla\psi(x) - \nabla\pi_M(\boldsymbol{w})(x))\cdot(\nabla\pi_M(\boldsymbol{w})(x) - \nabla\varphi_M(x))\,\mathrm{d}x \\ &\quad= \int_{\mathbb{T}} (\nabla\psi(x) - \nabla\varphi_M(x))\cdot(\nabla\pi_M(\boldsymbol{w})(x) - \nabla\varphi_M(x))\,\mathrm{d}x, \end{aligned}$$

and we get by the Cauchy-Schwarz inequality

$$\|\nabla \pi_M(\boldsymbol{w}) - \nabla \varphi_M\|_{L^2(\mathbb{T})} \le \inf_{\psi \in \mathfrak{p}_M} \|\nabla \psi - \nabla \varphi_M\|_{L^2(\mathbb{T})}.$$

Taking $\psi = \iota_M(\varphi)$ and using successively $\|\nabla f\|_{L^2(\mathbb{T})} = 2\pi \|f\|_{\dot{H}^1(\mathbb{T})}$ for any $f \in \dot{H}^1(\mathbb{T})$ and the third estimate of Lemma 2, we get

$$\begin{aligned}\|\nabla \pi_M(\boldsymbol{w}) - \nabla \varphi_M\|_{L^2(\mathbb{T})} &\lesssim \|\nabla \iota_M(\varphi) - \nabla \varphi_M\|_{L^2(\mathbb{T})} \\ &\lesssim \|\iota_M(\varphi) - \varphi_M\|_{\dot{H}^1(\mathbb{T})} \lesssim \frac{1}{M}\|\varphi_M\|_{\dot{H}^2(\mathbb{T})},\end{aligned}$$

where we refer to Subsection 1.2 for the definition of the homogeneous norms $\|\cdot\|_{\dot{H}^s(\mathbb{T})}$. Recalling that $-\Delta\varphi_M = \pi_M(\boldsymbol{u})$ we have $\|\pi_M(\boldsymbol{u})\|_{\dot{H}^{-1}(\mathbb{T})} = \|\nabla\varphi_M\|_{L^2(\mathbb{T})}$ and $\|\varphi_M\|_{\dot{H}^2(\mathbb{T})} = \|\Delta\varphi_M\|_{L^2(\mathbb{T})} = \|\pi_M(\boldsymbol{u})\|_{L^2(\mathbb{T})}$. All in all, using the reversed triangular inequality we have established

$$\left| \|\nabla \pi_M(\boldsymbol{w})\|_{L^2(\mathbb{T})} - \|\pi_M(\boldsymbol{u})\|_{\dot{H}^{-1}(\mathbb{T})} \right| \lesssim \frac{1}{M}\|\pi_M(\boldsymbol{u})\|_{L^2(\mathbb{T})}.$$

To conclude, due to (24), it is thus sufficient to prove that $\|\nabla\pi_M(\boldsymbol{w})\|_{L^2(\mathbb{T})} \sim \|\nabla\pi_M(\Phi)\|_{L^2(\mathbb{T})}$, where we recall $\boldsymbol{w} = B_M\Phi$. This last equality implies in particular

$$\pi_M(\boldsymbol{w}) = \frac{2}{3}\pi_M(\Phi) + \frac{1}{6}\tau_{\frac{1}{M}}\pi_M(\Phi) + \frac{1}{6}\tau_{-\frac{1}{M}}\pi_M(\Phi),$$

where we recall for $a \in \mathbb{R}$ the translation operator τ_a defined by $\tau_a f(x) = f(x+a)$. We have therefore

$$\nabla\pi_M(\boldsymbol{w}) = \frac{2}{3}\nabla\pi_M(\Phi) + \frac{1}{6}\tau_{\frac{1}{M}}\nabla\pi_M(\Phi) + \frac{1}{6}\tau_{-\frac{1}{M}}\nabla\pi_M(\Phi). \tag{25}$$

Both $\nabla\pi_M(\boldsymbol{w})$ and $\nabla\pi_M(\Phi)$ belong to $\mathfrak{s}_M(\mathbb{T})$ *i.e.* are respectively equal to some functions $\sigma_M(\boldsymbol{\lambda})$ and $\sigma_M(\boldsymbol{\gamma})$, for some $\boldsymbol{\lambda}, \boldsymbol{\gamma} \in \mathbb{R}^M$.

Note that B_M is uniformly well-conditioned: the spectral radii of B_M and B_M^{-1} are bounded independently of M. This can be seen writing $B_M = \frac{2}{3}I_M + \frac{1}{6}J_M$, where J_M is the matrix

$$J_M := \begin{pmatrix} 0 & 1 & 0 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & 1 \\ 1 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

The eigenvalues of J_M are $\left\{2\cos\left(\frac{2\pi k}{M}\right) \,:\, k \in \{0,\dots,M-1\}\right\}$, so the spectrum of B_M lies within $[1/3, 1]$.

The identity (25) shows that $\boldsymbol{\lambda} = B_M\boldsymbol{\gamma}$ and we have just controlled the euclidean subordinate norms of B_M and B_M^{-1}: we have $\|\boldsymbol{\gamma}\|_{2,M} \sim \|B_M\boldsymbol{\gamma}\|_{2,M}$, and therefore $\|\nabla\pi_M(\boldsymbol{w})\|_{L^2(\mathbb{T})} \sim \|\nabla\pi_M(\Phi)\|_{L^2(\mathbb{T})}$, thanks to Proposition 2, concluding the proof of (22).

Let us turn to the proof of (23). Using Proposition 4, $\|\Delta_M^{-1}\widetilde{\boldsymbol{u}}\|_{2,M} \le \|\widetilde{\boldsymbol{u}}\|_{2,M}$ and Cauchy-Schwarz inequality entails that $-(\widetilde{\boldsymbol{u}}|\Delta_M^{-1}\widetilde{\boldsymbol{u}})_M \le \|\widetilde{\boldsymbol{u}}\|_{2,M}^2$. By Pythagore's identity, we obtain (23), since $\boldsymbol{u} = \widetilde{\boldsymbol{u}} + [\boldsymbol{u}]_M \mathbf{1}_M$ and $\|[\boldsymbol{u}]_M \mathbf{1}_M\|_{2,M}^2 = [\boldsymbol{u}]_M^2$. □

Proposition 7. *For $\boldsymbol{w} \in \mathscr{C}^1([0,T]; \mathrm{Ran}(\Delta_M))$, we have*

$$-(\Delta_M^{-1}\boldsymbol{w}(t)|\boldsymbol{w}'(t))_M = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{w}(t)\|_{-1,M}^2,$$

where as usual $(\cdot|\cdot)_M$ denotes the rescaled inner-product on $\mathbb{R}^M$.

Proof. If $\boldsymbol{v}(t) := -\Delta_M^{-1}\boldsymbol{w}(t)$, we have $\Delta_M \boldsymbol{v}(t) = -\boldsymbol{w}(t)$ and therefore $\Delta_M \boldsymbol{v}'(t) = -\boldsymbol{w}'(t)$, with still $[\boldsymbol{v}'(t)]_M = 0$. We then have $\boldsymbol{v}'(t) = -\Delta_M^{-1}\boldsymbol{w}'(t)$. We infer, by symmetry of $\sqrt{-\Delta_M}$,

$$\begin{aligned}
-(\Delta_M^{-1}\boldsymbol{w}(t)|\boldsymbol{w}'(t))_M &= -\big(\boldsymbol{v}(t)|\Delta_M \boldsymbol{v}'(t)\big)_M \\
&= \Big(\sqrt{-\Delta_M}\boldsymbol{v}(t)|\sqrt{-\Delta_M}\boldsymbol{v}'(t)\Big)_M \\
&= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\sqrt{-\Delta_M}\boldsymbol{v}(t)|\sqrt{-\Delta_M}\boldsymbol{v}(t)\Big)_M \\
&= -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{v}(t)|\Delta_M \boldsymbol{v}(t))_M = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{w}(t)\|_{-1,M}^2.
\end{aligned}$$

□

4.4 The discrete duality lemma

We are now all set to state and prove the discrete duality lemmas. These estimates will apply to linear differential equations with source terms. We first consider the case when the source term is continuous and then the case when it is not regular, respectively Lemma 3 and 4. We need to combine them to deal with the approximation of the stochastic process and this is achieved in Proposition 8.

Lemma 3. *Consider $\boldsymbol{\mu} \in \mathscr{C}([0,T]; \mathbb{R}_{>0}^M)$ so that each component is uniformly (w.r.t. to time and index) lower bounded by a positive constant $\alpha > 0$. Assume that $\boldsymbol{z} \in \mathscr{C}^1([0,T]; \mathbb{R}^M)$ solves*

$$\boldsymbol{z}'(t) = \Delta_M\Big[\boldsymbol{z}(t) \odot \boldsymbol{\mu}(t) + \boldsymbol{f}(t)\Big] + \boldsymbol{r}(t),$$

where $\boldsymbol{f}$ and $\boldsymbol{r}$ are two functions in $\mathscr{C}([0,T]; \mathbb{R}^M)$. Then we have the following estimate, for any parameter $a > 0$

$$\begin{aligned}
&\sup_{t\in[0,T]} \|\boldsymbol{z}(t)\|_{-1,M}^2 + \int_{Q_T} \sigma_M(\boldsymbol{z} \odot \boldsymbol{\mu}^{1/2})(s,x)^2 \,\mathrm{d}s\,\mathrm{d}x \\
&\quad\le (1+a)\Big[\|\boldsymbol{z}(0)\|_{-1,M}^2 + [\boldsymbol{z}(0)]_M^2 \int_0^T [\boldsymbol{\mu}(s)]_M \,\mathrm{d}s + \frac{1}{\alpha}\int_{Q_T} \sigma_M(\boldsymbol{f})(s,x)^2 \,\mathrm{d}s\,\mathrm{d}x\Big] \\
&\qquad + (1+a^{-1})\left(T + T\int_0^T [\boldsymbol{\mu}(s)]_M \,\mathrm{d}s + \frac{1}{\alpha}\right)\int_{Q_T} \sigma_M(\boldsymbol{r})(s,x)^2 \,\mathrm{d}s\,\mathrm{d}x, \quad (26)
\end{aligned}$$

where the Hadamard product $\odot$ and the square-root $\boldsymbol{\mu}^{1/2}$ are defined in Subsection 1.2.

This is a counterpart of Lemma 1. In the case, $\boldsymbol{r} = 0$, one can indeed get rid of a.

Proof. We follow the proof of the continuous case, Lemma 1. Since $\mathrm{Ran}(\Delta_M) \subseteq \mathbf{1}_M^\perp$, we claim

$$[\boldsymbol{z}(t)]_M' = \frac{1}{M}(\boldsymbol{z}'(t), \mathbf{1}_M) = [\boldsymbol{r}(t)]_M,$$

and therefore

$$[\boldsymbol{z}(t)]_M = [\boldsymbol{z}(0)]_M + \int_0^t [\boldsymbol{r}(s)]_M \,\mathrm{d}s. \tag{27}$$

Recalling the definition $\widetilde{\boldsymbol{z}}(t) := \boldsymbol{z}(t) - [\boldsymbol{z}(t)]_M$ we also have

$$\boldsymbol{z}'(t) = \widetilde{\boldsymbol{z}}'(t) + [\boldsymbol{r}(t)]_M.$$

Now, taking the inner-product with the vector $\Delta_M^{-1}\widetilde{\boldsymbol{z}}(t)$ in the differential equation solved by $\boldsymbol{z}$, we get, using the symmetry of Δ_M and the fact $\Delta_M^{-1}\widetilde{\boldsymbol{z}}(t) \in \mathrm{Span}_{\mathbb{R}}(\mathbf{1}_M)^\perp$ (see Subsection 4.3),

$$-\Big(\Delta_M^{-1}\widetilde{\boldsymbol{z}}(t)\big|\widetilde{\boldsymbol{z}}'(t)\Big)_M + \Big(\widetilde{\boldsymbol{z}}(t)\big|\boldsymbol{z}(t) \odot \boldsymbol{\mu}(t)\Big)_M = -\Big(\widetilde{\boldsymbol{z}}(t)\big|\boldsymbol{f}(t) + \Delta_M^{-1}\widetilde{\boldsymbol{r}}(t)\Big)_M.$$

We use Proposition 7 to identify the first term of the l.h.s. and get

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\widetilde{\boldsymbol{z}}(t)\|_{-1,M}^2 + \Big(\widetilde{\boldsymbol{z}}(t)\big|\boldsymbol{z}(t) \odot \boldsymbol{\mu}(t)\Big)_M = -\Big(\widetilde{\boldsymbol{z}}(t)\big|\boldsymbol{f}(t) + \Delta_M^{-1}\widetilde{\boldsymbol{r}}(t)\Big)_M. \tag{28}$$

Using that the entries of $\boldsymbol{\mu}(t)$ are all lower-bounded by $\alpha > 0$ we have the following inequality (see Subsection 1.2 for the notation $\oslash$), for any vector $\boldsymbol{g} \in \mathbb{R}^M$

$$\begin{aligned}
\left|\Big(\widetilde{\boldsymbol{z}}(t)|\boldsymbol{g}\Big)_M\right| &= \left|\Big(\widetilde{\boldsymbol{z}}(t) \odot \boldsymbol{\mu}(t)^{1/2}|\boldsymbol{g} \oslash \boldsymbol{\mu}(t)^{1/2}\Big)_M\right| \\
&\le \|\widetilde{\boldsymbol{z}}(t) \odot \boldsymbol{\mu}(t)^{1/2}\|_{2,M}\|\boldsymbol{g} \oslash \boldsymbol{\mu}(t)^{1/2}\|_{2,M} \\
&\le \frac{1}{\sqrt{\alpha}}\|\widetilde{\boldsymbol{z}}(t) \odot \boldsymbol{\mu}(t)^{1/2}\|_{2,M}\|\boldsymbol{g}\|_{2,M} \\
&\le \frac{1}{2}\|\widetilde{\boldsymbol{z}}(t) \odot \boldsymbol{\mu}(t)^{1/2}\|_{2,M}^2 + \frac{1}{2\alpha}\|\boldsymbol{g}\|_{2,M}^2 \\
&= \frac{1}{2}\Big(\widetilde{\boldsymbol{z}}(t)|\widetilde{\boldsymbol{z}}(t) \odot \boldsymbol{\mu}(t)\Big)_M + \frac{1}{2\alpha}\|\boldsymbol{g}\|_{2,M}^2,
\end{aligned}$$

where we used Young's inequality. We use this estimate in (28) with $\boldsymbol{g} := \boldsymbol{f}(t) + \Delta_M^{-1}\widetilde{\boldsymbol{r}}(t)$

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\widetilde{\boldsymbol{z}}(t)\|_{-1,M}^2 + \Big(\widetilde{\boldsymbol{z}}(t)\Big|\boldsymbol{z}(t) \odot \boldsymbol{\mu}(t)\Big)_M \\
&\qquad\qquad \le \frac{1}{2}\Big(\widetilde{\boldsymbol{z}}(t)\Big|\widetilde{\boldsymbol{z}}(t) \odot \boldsymbol{\mu}(t)\Big)_M + \frac{1}{2\alpha}\|\boldsymbol{f}(t) + \Delta_M^{-1}\widetilde{\boldsymbol{r}}(t)\|_{2,M}^2,
\end{aligned}$$

which, after expanding the definition $\widetilde{\boldsymbol{z}}(t) := \boldsymbol{z}(t) - [\boldsymbol{z}(t)]_M$, becomes

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\widetilde{\boldsymbol{z}}(t)\|_{-1,M}^2 + \frac{1}{2}\Big(\boldsymbol{z}(t)\Big|\boldsymbol{z}(t) \odot \boldsymbol{\mu}(t)\Big)_M \\
&\qquad\qquad \le \frac{1}{2}[\boldsymbol{z}(t)]_M^2[\boldsymbol{\mu}(t)]_M + \frac{1}{2\alpha}\|\boldsymbol{f}(t) + \Delta_M^{-1}\widetilde{\boldsymbol{r}}(t)\|_{2,M}^2.
\end{aligned}$$

Using Proposition 4 to infer $\|\Delta_M^{-1}\widetilde{\boldsymbol{r}}(t)\|_{2,M} \leq \|\widetilde{\boldsymbol{r}}(t)\|_{2,M} \leq \|\boldsymbol{r}(t)\|_{2,M}$ and the convex inequality $(x+y)^2 \leq (1+a)x^2 + (1+a^{-1})y^2$ we eventually get, after integration in time

$$\|\widetilde{\boldsymbol{z}}(t)\|_{-1,M}^2 + \int_0^t \|\boldsymbol{z}(s) \odot \boldsymbol{\mu}(s)^{1/2}\|_{2,M}^2 \,\mathrm{d}s \\ \leq \|\widetilde{\boldsymbol{z}}(0)\|_{-1,M}^2 + \int_0^t [\boldsymbol{z}(s)]_M^2[\boldsymbol{\mu}(s)]_M \,\mathrm{d}s \\ + \frac{1+a}{\alpha}\int_0^t \|\boldsymbol{f}(s)\|_{2,M}^2 \,\mathrm{d}s + \frac{1+a^{-1}}{\alpha}\int_0^t \|\boldsymbol{r}(s)\|_{2,M}^2 \,\mathrm{d}s. \quad (29)$$

On the other hand, using once more the above convex inequality, we claim from (27) and Cauchy-Schwarz inequality that

$$[\boldsymbol{z}(t)]_M^2 \leq (1+a)[\boldsymbol{z}(0)]_M^2 + (1+a^{-1})T\int_0^T [\boldsymbol{r}(s)]_M^2 \,\mathrm{d}s.$$

Summing the two last inequalities we obtain (26) since for any vector $\boldsymbol{u} \in \mathbb{R}^M$, $\|\boldsymbol{u}\|_{2,M} = \|\sigma_M(\boldsymbol{u})\|_{L^2(\mathbb{T})}$. □

Lemma 4. *Consider $\boldsymbol{\mu} \in \mathscr{C}([0,T];\mathbb{R}^M_{>0})$ so that each component is uniformly (w.r.t. to time and index) lower bounded by a positive constant $\alpha > 0$. Assume that $\boldsymbol{z}_d \colon [0,T] \to \mathbb{R}^M$ solves*

$$\boldsymbol{z}_d(t) = \int_0^t \Delta_M\Big[\boldsymbol{z}_d(s) \odot \boldsymbol{\mu}(s)\Big]\,\mathrm{d}s + \boldsymbol{x}_d(t), \quad (30)$$

where $\boldsymbol{x}_d$ is any càdlàg $\mathbb{R}^M$ valued function over $[0,T]$. Then we have the following estimate

$$\sup_{t\in[0,T]} \|\boldsymbol{z}_d(t)\|_{-1,M}^2 + \int_{Q_T} \sigma_M(\boldsymbol{z}_d \odot \boldsymbol{\mu}^{1/2})(s,x)^2 \,\mathrm{d}s\,\mathrm{d}x \\ \lesssim \sup_{t\in[0,T]} \|\boldsymbol{x}_d(t)\|_{-1,M}^2 + \int_0^T [\boldsymbol{\mu}(s)]_M[\boldsymbol{x}_d(s)]_M^2 \,\mathrm{d}s, \quad (31)$$

where the constant behind $\lesssim$ is universal and $\boldsymbol{\mu}^{1/2}$ denotes the vector map whose entries are the square-roots of the ones of $\boldsymbol{\mu}$.

Remark 5. *In this lemma we consider the (discrete) Kolmogorov equation with a singular source term $\boldsymbol{x}_d$. The mere integrability of this term forbids to differentiate in time this equation, so we cannot proceed as we have done in the proof Lemma 3.*

Proof. Using (30), we first remark that $[\boldsymbol{z}_d]_M = [\boldsymbol{x}_d]_M$ and therefore

$$\widetilde{\boldsymbol{z}}_d(t) = \int_0^t \Delta_M\Big[\boldsymbol{z}_d(s) \odot \boldsymbol{\mu}(s)\Big]\,\mathrm{d}s + \widetilde{\boldsymbol{x}}_d(t).$$

We take as usual the inner product with $-\Delta_M^{-1}\widetilde{\boldsymbol{z}}_d(t)$ and use symmetry to write

$$-\Big(\Delta_M^{-1}\widetilde{\boldsymbol{z}}_d(t)\big|\widetilde{\boldsymbol{z}}_d(t)\Big)_M + \int_0^t \Big(\widetilde{\boldsymbol{z}}_d(s)\big|\boldsymbol{z}_d(s) \odot \boldsymbol{\mu}(s)\Big)_M \,\mathrm{d}s = -\Big(\Delta_M^{-1}\widetilde{\boldsymbol{z}}_d(t), \widetilde{\boldsymbol{x}}_d(t)\Big)_M.$$

We use the definition of the $\|\cdot\|_{-1,M}$ norm (see Proposition 6) and the equality $\widetilde{\boldsymbol{z}}_d = \boldsymbol{z}_d - [\boldsymbol{z}_d]_M$ to infer

$$\|\widetilde{\boldsymbol{z}}_d(t)\|_{-1,M}^2 + \int_0^t \|\boldsymbol{z}_d(s) \odot \boldsymbol{\mu}(s)^{1/2}\|_{2,M}^2 \,\mathrm{d}s \\ = \int_0^t [\boldsymbol{z}_d(s)]_M \Big(\mathbf{1}_M | \boldsymbol{z}_d(s) \odot \boldsymbol{\mu}(s)\Big)_M \mathrm{d}s - \Big(\Delta_M^{-1} \widetilde{\boldsymbol{z}}_d(t), \widetilde{\boldsymbol{x}}_d(t)\Big)_M.$$

The first term of the r.h.s. can be handled using Young's inequality to absorb a part of it in the l.h.s. and get

$$\|\widetilde{\boldsymbol{z}}_d(t)\|_{-1,M}^2 + \int_0^t \|\boldsymbol{z}_d(s) \odot \boldsymbol{\mu}(s)^{1/2}\|_{2,M}^2 \,\mathrm{d}s \\ \lesssim \int_0^t [\boldsymbol{z}_d(s)]_M^2 \|\boldsymbol{\mu}(s)^{1/2}\|_{2,M}^2 \,\mathrm{d}s - \Big(\Delta_M^{-1} \widetilde{\boldsymbol{z}}_d(t), \widetilde{\boldsymbol{x}}_d(t)\Big)_M. \tag{32}$$

Now, defining $\boldsymbol{\Phi}_M := \Delta_M^{-1} \widetilde{\boldsymbol{z}}_d$ and $\boldsymbol{\Psi}_M := \Delta_M^{-1} \widetilde{\boldsymbol{x}}_d$ we have that using Cauchy-Schwarz's inequality, the definition of the $\|\cdot\|_{-1,M}$ norm and the symmetry of the discrete laplacian matrix

$$\begin{aligned} -\Big(\Delta_M^{-1} \widetilde{\boldsymbol{z}}_d(t) \mid \widetilde{\boldsymbol{x}}_d(t)\Big)_M &= -\Big(\boldsymbol{\Phi}_M(t) \mid \Delta_M \boldsymbol{\Psi}_M(t)\Big)_M \\ &= \Big(\sqrt{-\Delta_M}\boldsymbol{\Phi}_M(t) \mid \sqrt{-\Delta_M}\boldsymbol{\Psi}_M(t)\Big)_M \\ &\leq \|\sqrt{-\Delta_M}\boldsymbol{\Phi}_M(t)\|_{2,M} \|\sqrt{-\Delta_M}\boldsymbol{\Psi}_M(t)\|_{2,M} \\ &= \|\widetilde{\boldsymbol{z}}_d(t)\|_{-1,M} \|\widetilde{\boldsymbol{x}}_d(t)\|_{-1,M}. \end{aligned}$$

Plugging this estimate in (32), we have

$$\|\widetilde{\boldsymbol{z}}_d(t)\|_{-1,M}^2 + \int_0^t \|\boldsymbol{z}_d(s) \odot \boldsymbol{\mu}(s)^{1/2}\|_{2,M}^2 \,\mathrm{d}s \lesssim \int_0^t [\boldsymbol{z}_d(s)]_M^2 \|\boldsymbol{\mu}(s)^{1/2}\|_{2,M}^2 \,\mathrm{d}s + \|\widetilde{\boldsymbol{x}}_d(t)\|_{-1,M}^2.$$

Recalling that $[\boldsymbol{z}_d]_M^2 = [\boldsymbol{x}_d]_M^2$ and adding this term to the inequality, we get (31). □

Proposition 8. *Consider $\boldsymbol{\mu} \in \mathscr{C}([0,T]; \mathbb{R}^M_{>0})$ so that each component is uniformly (w.r.t. to time and index) lower bounded by a positive constant $\alpha > 0$. Assume that $\boldsymbol{z}\colon [0,T] \to \mathbb{R}^M$ solves*

$$\boldsymbol{z}(t) = \boldsymbol{z}(0) + \int_0^t \Delta_M \Big[\boldsymbol{z}(s) \odot \boldsymbol{\mu}(s) + \boldsymbol{f}(s)\Big] \,\mathrm{d}s + \boldsymbol{x}(t),$$

where $\boldsymbol{f}$ is a function in $\mathscr{C}([0,T]; \mathbb{R}^M)$ and $\boldsymbol{x} = \boldsymbol{x}_r + \boldsymbol{x}_d$, with the regular component $\boldsymbol{x}_r \in \mathscr{C}^1([0,T], \mathbb{R}^M)$ and the singular component $\boldsymbol{x}_d$ is any càdlàg $\mathbb{R}^M$ valued function over $[0,T]$.

Then we have the following estimate, for any $a > 0$,

$$\sup_{t\in[0,T]} \|\boldsymbol{z}(t)\|_{-1,M}^2 + \int_{Q_T} \sigma_M(\boldsymbol{z}\odot\boldsymbol{\mu}^{1/2})(s,x)^2\,\mathrm{d}s\,\mathrm{d}x$$
$$\leq (1+a)^2\left[\|\boldsymbol{z}(0)\|_{-1,M}^2 + [\boldsymbol{z}(0)]_M^2\int_0^T[\boldsymbol{\mu}(s)]_M\mathrm{d}s + \frac{1}{\alpha}\int_{Q_T}\sigma_M(\boldsymbol{f})(s,x)^2\,\mathrm{d}s\,\mathrm{d}x\right]$$
$$+ (1+a)(1+a^{-1})\left(T + T\int_0^T[\boldsymbol{\mu}(s)]_M\mathrm{d}s + \frac{1}{\alpha}\right)\int_{Q_T}\sigma_M(\boldsymbol{x}_r')(s,x)^2\,\mathrm{d}s\,\mathrm{d}x$$
$$+ (1+a^{-1})\Big[\sup_{t\in[0,T]}\|\boldsymbol{x}_d(t)\|_{-1,M}^2 + \int_0^T[\boldsymbol{\mu}(s)]_M[\boldsymbol{x}_d(s)]_M^2\,\mathrm{d}s\Big]. \quad (33)$$

Proof. Let's define $\boldsymbol{z}_r \in \mathscr{C}^1([0,T];\mathbb{R}^M)$ as the unique solution of

$$\boldsymbol{z}_r(t) = \boldsymbol{z}(0) + \int_0^t \Delta_M\Big[\boldsymbol{z}_r(s)\odot\boldsymbol{\mu}(s) + \boldsymbol{f}(s)\Big]\,\mathrm{d}s + \boldsymbol{x}_r(t),$$

which, since $\boldsymbol{x}_r$ is continuously differentiable, is equivalent to the Cauchy problem

$$\boldsymbol{z}_r'(t) = \Delta_M\Big[\boldsymbol{z}_r(t)\odot\boldsymbol{\mu}(t) + \boldsymbol{f}(t)\Big]\,\mathrm{d}s + \boldsymbol{x}_r'(t), \quad (34)$$
$$\boldsymbol{z}_r(0) = \boldsymbol{z}(0). \quad (35)$$

Now, defining $\boldsymbol{z}_d \coloneqq \boldsymbol{z} - \boldsymbol{z}_r$, one readily checks that it solves

$$\boldsymbol{z}_d(t) = \int_0^t \Delta_M\Big[\boldsymbol{z}_d(s)\odot\boldsymbol{\mu}(s)\Big]\,\mathrm{d}s + \boldsymbol{x}_d(t).$$

The Cauchy problem (34) – (35) is exactly the one of Lemma 3, with $\boldsymbol{r}(t) \coloneqq \boldsymbol{x}_r'(t)$, we therefore infer from this very lemma, for any $a > 0$

$$\sup_{t\in[0,T]} \|\boldsymbol{z}_r(t)\|_{-1,M}^2 + \int_{Q_T} \sigma_M(\boldsymbol{z}_r\odot\boldsymbol{\mu}^{1/2})(s,x)^2\,\mathrm{d}s\,\mathrm{d}x$$
$$\leq (1+a)\left[\|\boldsymbol{z}(0)\|_{-1,M}^2 + [\boldsymbol{z}(0)]_M^2\int_0^T[\boldsymbol{\mu}(s)]_M\,\mathrm{d}s + \frac{1}{\alpha}\int_{Q_T}\sigma_M(\boldsymbol{f})(s,x)^2\,\mathrm{d}s\,\mathrm{d}x\right]$$
$$+ (1+a^{-1})\left(T + T\int_0^T[\boldsymbol{\mu}(s)]_M\,\mathrm{d}s + \frac{1}{\alpha}\right)\int_{Q_T}\sigma_M(\boldsymbol{x}_r')(s,x)^2\,\mathrm{d}s\,\mathrm{d}x.$$

Now, since $\boldsymbol{z} = \boldsymbol{z}_d + \boldsymbol{z}_r$, combining the triangular inequality and the convex inequality $(x+y)^2 \leq (1+a)x^2 + (1+a^{-1})y^2$ implies

$$\sup_{t\in[0,T]} \|\boldsymbol{z}(t)\|_{-1,M}^2 + \int_{Q_T} \sigma_M(\boldsymbol{z}\odot\boldsymbol{\mu}^{1/2})(s,x)^2\,\mathrm{d}s\,\mathrm{d}x$$
$$\leq (1+a)\left[\sup_{t\in[0,T]}\|\boldsymbol{z}_r(t)\|_{-1,M}^2 + \int_{Q_T}\sigma_M(\boldsymbol{z}_r\odot\boldsymbol{\mu}^{1/2})(s,x)^2\,\mathrm{d}s\,\mathrm{d}x\right]$$
$$+ (1+a^{-1})\left[\sup_{t\in[0,T]}\|\boldsymbol{z}_d(t)\|_{-1,M}^2 + \int_{Q_T}\sigma_M(\boldsymbol{z}_d\odot\boldsymbol{\mu}^{1/2})(s,x)^2\,\mathrm{d}s\,\mathrm{d}x\right],$$

so that the proof follows from Lemma 4, which focuses on the non-regular component. $\square$

5 Quantitative estimates and proof of Theorem 2

Let u, v be two functions defined on Q_T of $\mathscr{C}^1$ regularity in time and $\mathscr{C}^4$ regularity in space solution of the system (1). We have, by Taylor expansion, for any $h > 0$ and $\mathscr{C}^4(\mathbb{T})$ function f

$$\tau_h f = f + hf' + \frac{h^2}{2!} f'' + \frac{h^3}{3!} f''' + \mathrm{O}_{h\to 0}(h^4),$$
$$\tau_{-h} f = f - hf' + \frac{h^2}{2!} f'' - \frac{h^3}{3!} f''' + \mathrm{O}_{h\to 0}(h^4),$$

where $\mathrm{O}_{h\to 0}$ refers to the $L^\infty(Q_T)$ norm. We have therefore

$$\frac{\tau_h f + \tau_{-h} f - 2f}{h^2} = f'' + \mathrm{O}_{h\to 0}(h^2).$$

In particular, denoting by $\widehat{\boldsymbol{u}}^M(t)$ and $\widehat{\boldsymbol{v}}^M(t)$ the respectives values of u and v at the points (t, x_k) for $k = 1, \dots, M$, we have the following discrete system:

$$\partial_t \widehat{\boldsymbol{u}}^M(t) = \Delta_M \big[d_1 \widehat{\boldsymbol{u}}^M(t) + a_{12} \widehat{\boldsymbol{u}}^M(t) \odot \widehat{\boldsymbol{v}}^M(t) \big] + \boldsymbol{r}^M(t),$$
$$\partial_t \widehat{\boldsymbol{v}}^M(t) = \Delta_M \big[d_1 \widehat{\boldsymbol{v}}^M(t) + a_{21} \widehat{\boldsymbol{v}}^M(t) \odot \widehat{\boldsymbol{u}}^M(t) \big] + \boldsymbol{s}^M(t),$$

with

$$\|\boldsymbol{r}^M(t)\|_\infty + \|\boldsymbol{s}^M(t)\|_\infty \lesssim M^{-2}, \tag{36}$$

uniformly for t on compact intervals.

On the other hand, we recall that our stochastic process satisfies

$$\boldsymbol{U}^{M,N}(t) = \boldsymbol{U}^{M,N}(0) + \int_0^t \Delta_M \Big(d_1 \boldsymbol{U}^{M,N}(s) + a_{12} \boldsymbol{U}^{M,N}(s) \odot \boldsymbol{V}^{M,N}(s) \Big) \, \mathrm{d}s + \boldsymbol{\mathcal{M}}^{M,N}(t),$$
$$\boldsymbol{V}^{M,N}(t) = \boldsymbol{V}^{M,N}(0) + \int_0^t \Delta_M \Big(d_2 \boldsymbol{V}^{M,N}(s) + a_{21} \boldsymbol{U}^{M,N}(s) \odot \boldsymbol{V}^{M,N}(s) \Big) \, \mathrm{d}s + \boldsymbol{\mathcal{N}}^{M,N}(t),$$

where $\boldsymbol{\mathcal{M}}^{M,N}$ is square integrable martingale whose quadratic variation is given by (13) and $\boldsymbol{\mathcal{N}}^{M,N}$ satisfies similar properties. By symmetry, we can focus on the first species $\boldsymbol{U}^{M,N}$. Denoting

$$\boldsymbol{Z}^{M,N}(t) = \widehat{\boldsymbol{u}}^M(t) - \boldsymbol{U}^{M,N}(t), \quad \boldsymbol{X}^{M,N}(t) = \int_0^t \boldsymbol{r}^M(s) \, \mathrm{d}s - \boldsymbol{\mathcal{M}}^{M,N}(t),$$

we have yet another system satisfied by these quantities

$$\boldsymbol{Z}^{M,N}(t) = \boldsymbol{Z}^{M,N}(0) + \int_0^t \Delta_M \Big(\boldsymbol{Z}^{M,N}(s) \odot \boldsymbol{\Lambda}^{M,N}(s) + \boldsymbol{F}^{M,N}(s) \Big) \, \mathrm{d}s + \boldsymbol{X}^{M,N}(t), \tag{37}$$

where

$$\begin{aligned}\boldsymbol{\Lambda}^{M,N}(t) &= d_1\mathbf{1}_M + a_{12}\boldsymbol{V}^{M,N}(t),\\ \boldsymbol{W}^{M,N}(t) &= \widehat{\boldsymbol{v}}^M(t) - \boldsymbol{V}^{M,N}(t),\\ \boldsymbol{F}^{M,N}(t) &= a_{12}\widehat{\boldsymbol{u}}^M \odot \boldsymbol{W}^{M,N}(t).\end{aligned}$$

We can now apply the discrete duality lemma developed in the previous section to control the gap $\boldsymbol{Z}^{M,N}$. This is the core of the next result, which yields Theorem 2. For $\boldsymbol{z}\colon [0,T] \to \mathbb{R}^M$, let

$$|||\boldsymbol{z}|||_{T,M} \coloneqq \left(\sup_{t\in[0,T]} \|\boldsymbol{z}(t)\|_{-1,M}^2 + \|\sigma_M(\boldsymbol{z})\|_{L^2(Q_T)}^2\right)^{1/2}.$$

Proposition 9. *Let u, v be a $\mathscr{C}^4(Q_T)$ solution of the system (1), if*

$$\frac{a_{12}a_{21}}{d_1 d_2}\|u\|_{L^\infty(Q_T)}\|v\|_{L^\infty(Q_T)} < 1,$$

then for any $(M,N) \in \mathbb{N}^2$ such that N/M^2 is large enough

$$\begin{aligned}&\mathbb{E}\big(|||\boldsymbol{Z}^{M,N}|||_{T,M}^2 + |||\boldsymbol{W}^{M,N}|||_{T,M}^2\big)\\ &\quad\lesssim |||\boldsymbol{Z}^{M,N}(0)|||_{T,M}^2(1+[\boldsymbol{V}^{M,N}(0)]_M) + |||\boldsymbol{W}^{M,N}(0)|||_{T,M}^2(1+[\boldsymbol{U}^{M,N}(0)]_M)\\ &\qquad\qquad + \big(1+T^2+T^2[\boldsymbol{U}^{M,N}(0)+\boldsymbol{V}^{M,N}(0)]_M\big)\,M^{-4} + TM^2N^{-1}.\end{aligned} \tag{38}$$

Proof. We first observe that $t \mapsto [\boldsymbol{\Lambda}^{M,N}(t)]_M$ is constant and we set

$$\begin{aligned}\boldsymbol{\lambda}_T^{M,N} &= T + T\int_0^T [\boldsymbol{\Lambda}^{M,N}(s)]_M \mathrm{d}s + \frac{1}{d_1}\\ &= T + T^2(d_1 + a_{12}[\boldsymbol{V}^{M,N}(0)]_M) + \frac{1}{d_1}.\end{aligned}$$

By applying Proposition 8 with $\boldsymbol{x}_d \coloneqq -\boldsymbol{\mathcal{M}}^{M,N}$ and

$$\boldsymbol{x}_r\colon t \mapsto \int_0^t \boldsymbol{r}^M(\sigma)\,\mathrm{d}\sigma,$$

we obtain that

$$\begin{aligned}&\sup_{t\in[0,T]}\|\boldsymbol{Z}^{M,N}(t)\|_{-1,M}^2 + \int_{Q_T}\sigma_M\big(\boldsymbol{Z}^{M,N}\odot(\boldsymbol{\Lambda}^{M,N})^{1/2}\big)(s,x)^2\,\mathrm{d}s\mathrm{d}x\\ &\quad\lesssim \|\boldsymbol{Z}^{M,N}(0)\|_{-1,M}^2 + [\boldsymbol{Z}^{M,N}(0)]_M^2\int_0^T[\boldsymbol{\Lambda}^{M,N}(s)]_M\,\mathrm{d}s\\ &\qquad + \frac{1}{d_1}\int_{Q_T}\sigma_M(\boldsymbol{F}^{M,N})(s,x)^2\,\mathrm{d}s\mathrm{d}x + \boldsymbol{\lambda}_T^{M,N}\int_{Q_T}\sigma_M(\boldsymbol{r}^M)(s,x)^2\,\mathrm{d}s\mathrm{d}x\\ &\qquad\qquad + \sup_{t\in[0,T]}\|\boldsymbol{\mathcal{M}}^{M,N}(t)\|_{-1,M}^2 + \int_0^T[\boldsymbol{\Lambda}^{M,N}(s)]_M[\boldsymbol{\mathcal{M}}^{M,N}(s)]_M^2\,\mathrm{d}s.\end{aligned} \tag{39}$$

Since $\Lambda_i^{M,N} \geq d_1$ and $|\sigma_M(\boldsymbol{F}^{M,N})(s,x)| \leq a_{12}\|u\|_{L^\infty(Q_T)}|\sigma_M(\boldsymbol{W}^{M,N})(s,x)|$, we obtain

$$\begin{aligned}
&|||\boldsymbol{Z}^{M,N}(t)|||_{T,M}^2 \\
&\quad \lesssim \|\boldsymbol{Z}^{M,N}(0)\|_{-1,M}^2 + T[\boldsymbol{Z}^{M,N}(0)]_M^2[\boldsymbol{\Lambda}^{M,N}(0)]_M \\
&\qquad + \frac{a_{12}\|u\|_{L^\infty(Q_T)}}{d_1}\int_{Q_T}\sigma_M(\boldsymbol{W}^{M,N})(s,x)^2\,\mathrm{d}s\mathrm{d}x + \boldsymbol{\lambda}_T^{M,N}\int_{Q_T}\sigma_M(\boldsymbol{r}^M)(s,x)^2\,\mathrm{d}s\mathrm{d}x \\
&\qquad\qquad + \sup_{t\in[0,T]}\|\boldsymbol{\mathcal{M}}^{M,N}(t)\|_{-1,M}^2 + \int_0^T[\boldsymbol{\Lambda}^{M,N}(s)]_M[\boldsymbol{\mathcal{M}}^{M,N}(s)]_M^2\,\mathrm{d}s.
\end{aligned}$$

As the roles of $\boldsymbol{Z}^{M,N}$ and $\boldsymbol{W}^{M,N}$ are symmetric in the previous inequality, we have a similar estimate for $\boldsymbol{W}^{M,N}$. Thus, by setting

$$\boldsymbol{\Gamma}^{M,N}(t) = d_2 + a_{21}\boldsymbol{U}^{M,N}(t),$$

and

$$\begin{aligned}
\boldsymbol{\gamma}_T^{M,N} &= T + T\int_0^T[\boldsymbol{\Gamma}^{M,N}(s)]_M\mathrm{d}s + \frac{1}{d_2} \\
&= T + T^2(d_2 + a_{21}[\boldsymbol{U}^{M,N}(0)]_M) + \frac{1}{d_2},
\end{aligned}$$

we get

$$\begin{aligned}
&|||\boldsymbol{W}^{M,N}|||_{T,M}^2 \\
&\quad \lesssim \|\boldsymbol{W}^{M,N}(0)\|_{-1,M}^2 + T[\boldsymbol{W}^{M,N}(0)]_M^2[\boldsymbol{\Gamma}^{M,N}(0)]_M \\
&\qquad + \frac{a_{21}\|\overline{v}\|_{L^\infty(Q_T)}}{d_2}\int_{Q_T}\sigma_M(\boldsymbol{Z}^{M,N})(s,x)^2\,\mathrm{d}s\mathrm{d}x + \boldsymbol{\gamma}_T^{M,N}\int_{Q_T}\sigma_M(\boldsymbol{s}^M)(s,x)^2\,\mathrm{d}s\mathrm{d}x \\
&\qquad\qquad + \sup_{t\in[0,T]}\|\boldsymbol{\mathcal{N}}^{M,N}(t)\|_{-1,M}^2 + \int_0^T[\boldsymbol{\Gamma}^{M,N}(s)]_M[\boldsymbol{\mathcal{N}}^{M,N}(s)]_M^2\,\mathrm{d}s.
\end{aligned}$$

Using now our assumption on the bound of $\|\overline{u}\|_{L^\infty(Q_T)}\|\overline{v}\|_{L^\infty(Q_T)}$ and letting $\lesssim$ to depend on those parameters, we get that

$$\begin{aligned}
|||\boldsymbol{Z}^{M,N}|||_{T,M}^2 \lesssim\ & \|\boldsymbol{Z}^{M,N}(0)\|_{-1,M}^2 + T[\boldsymbol{Z}^{M,N}(0)]_M^2[\boldsymbol{\Lambda}^{M,N}(0)]_M \\
&+ \|\boldsymbol{W}^{M,N}(0)\|_{-1,M}^2 + T[\boldsymbol{W}^{M,N}(0)]_M^2[\boldsymbol{\Gamma}^{M,N}(0)]_M \\
&+ \left(\boldsymbol{\lambda}_T^{M,N} + \boldsymbol{\gamma}_T^{M,N}\right)\int_{Q_T}\left(\sigma_M(\boldsymbol{r}^M)(s,x)^2 + \sigma_M(\boldsymbol{s}^M)(s,x)^2\right)\mathrm{d}s\mathrm{d}x \\
&+ \sup_{t\in[0,T]}\|\boldsymbol{\mathcal{M}}^{M,N}(t)\|_{-1,M}^2 + \int_0^T[\boldsymbol{\Lambda}^{M,N}(s)]_M[\boldsymbol{\mathcal{M}}^{M,N}(s)]_M^2\,\mathrm{d}s \\
&+ \sup_{t\in[0,T]}\|\boldsymbol{\mathcal{N}}^{M,N}(t)\|_{-1,M}^2 + \int_0^T[\boldsymbol{\Gamma}^{M,N}(s)]_M[\boldsymbol{\mathcal{N}}^{M,N}(s)]_M^2\,\mathrm{d}s.
\end{aligned}$$

The previous r.h.s. is again invariant with respect to the roles of $\boldsymbol{Z}^{M,N}$ and $\boldsymbol{W}^{M,N}$. Then using the uniform bounds on $\sigma_M(\boldsymbol{r}^M)$ and $\sigma_M(\boldsymbol{s}^M)$ from (36) and taking expectation, we get

$$
\begin{aligned}
&\mathbb{E}\big(|||\boldsymbol{Z}^{M,N}|||_{T,M}^2 + |||\boldsymbol{W}^{M,N}|||_{T,M}^2\big) \\
&\quad \lesssim \|\boldsymbol{Z}^{M,N}(0)\|_{-1,M}^2 + T[\boldsymbol{Z}^{M,N}(0)]_M^2[\boldsymbol{\Lambda}^{M,N}(0)]_M \\
&\qquad + \|\boldsymbol{W}^{M,N}(0)\|_{-1,M}^2 + T[\boldsymbol{W}^{M,N}(0)]_M^2[\boldsymbol{\Gamma}^{M,N}(0)]_M + \Big(\boldsymbol{\lambda}_T^{M,N} + \boldsymbol{\gamma}_T^{M,N}\Big) M^{-4} \\
&\qquad + \mathbb{E}\Big(\sup_{t\in[0,T]} \|\boldsymbol{\mathcal{M}}^{M,N}(t)\|_{-1,M}^2\Big) + [\boldsymbol{\Lambda}^{M,N}(0)]_M \int_0^T \mathbb{E}\big([\boldsymbol{\mathcal{M}}^{M,N}(s)]_M^2\big)\,\mathrm{d}s \\
&\qquad + \mathbb{E}\Big(\sup_{t\in[0,T]} \|\boldsymbol{\mathcal{N}}^{M,N}(t)\|_{-1,M}^2\Big) + [\boldsymbol{\Gamma}^{M,N}(0)]_M \int_0^T \mathbb{E}\big([\boldsymbol{\mathcal{N}}^{M,N}(s)]_M^2\big)\,\mathrm{d}s. \qquad (40)
\end{aligned}
$$

We are left then with controlling the local martingale terms that appear at the end. Since

$$
\begin{aligned}
[\boldsymbol{\mathcal{M}}^{M,N}(s)]_M^2 &= \Big(\frac{1}{M}\sum_{i=1}^M \mathcal{M}_i^{M,N}(s)\Big)^2 \\
&\le \frac{1}{M}\sum_{i=1}^M \mathcal{M}_i^{M,N}(s)^2 \le \frac{1}{M}\sum_{i=1}^M \sup_{t\in[0,T]} \mathcal{M}_i^{M,N}(t)^2,
\end{aligned}
$$

and recalling that $[\boldsymbol{\Lambda}^{M,N}(0)]_M = [d_1 + a_{12}\boldsymbol{V}^{M,N}(0)]_M$, we have

$$
\begin{aligned}
&[\boldsymbol{\Lambda}^{M,N}(0)]_M \int_0^T \mathbb{E}\big([\boldsymbol{\mathcal{M}}^{M,N}(s)]_M^2\big)\,\mathrm{d}s \\
&\qquad\qquad \le (d_1 + a_{12}\|\boldsymbol{V}^{M,N}(0)\|_{1,M})\frac{T}{M}\sum_{i=1}^M \mathbb{E}\Big(\sup_{t\in[0,T]} \mathcal{M}_i^{M,N}(t)^2\Big).
\end{aligned}
$$

Besides, we also have using (23)

$$
\begin{aligned}
\mathbb{E}\Big(\sup_{t\in[0,T]} \|\boldsymbol{\mathcal{M}}^{M,N}(t)\|_{-1,M}^2\Big) &\le \mathbb{E}\Big(\sup_{t\in[0,T]} \|\boldsymbol{\mathcal{M}}^{M,N}(t)\|_{2,M}^2\Big) \\
&\le \mathbb{E}\Big(\sup_{t\in[0,T]} \Big(\frac{1}{M}\sum_{i=1}^M \mathcal{M}_i^{M,N}(t)^2\Big)\Big) \\
&\le \frac{1}{M}\sum_{i=1}^M \mathbb{E}\Big(\sup_{t\in[0,T]} \mathcal{M}_i^{M,N}(t)^2\Big).
\end{aligned}
$$

Now, Doob's inequality ensures that $\mathbb{E}\big(\sup_{t\in[0,T]} \mathcal{M}_i^{M,N}(t)^2\big) \lesssim \mathbb{E}(\langle \mathcal{M}_i^{M,N}\rangle(T))$, where the expression of the quadratic variation $\langle \mathcal{M}_i^{M,N} \rangle$ is found in (13). Moreover, since

$$\sum_{i=1}^{M}\Big(2U_i^{M,N}(s)V_i^{M,N}(s) + U_{i+1}^{M,N}(s)V_{i+1}^{M,N}(s) + U_{i-1}^{M,N}(s)V_{i-1}^{M,N}(s)\Big)$$
$$\lesssim \sum_{i=1}^{M}\Big(U_i^{M,N}(s)^2 + V_i^{M,N}(s)^2\Big) = \|\boldsymbol{U}^{M,N}(t)\|_2^2 + \|\boldsymbol{V}^{M,N}(t)\|_2^2,$$

we get

$$\frac{1}{M}\sum_{i=1}^{M}\mathbb{E}\Big(\sup_{t\in[0,T]} \mathcal{M}_i^{M,N}(t)^2\Big) \lesssim \frac{1}{M}\mathbb{E}\Big(\frac{M^2}{N}\int_0^T \|\boldsymbol{U}^{M,N}(s)\|_1 \,\mathrm{d}s + \frac{M^2}{N}\int_0^T \big(\|\boldsymbol{U}^{M,N}(s)\|_2^2 + \|\boldsymbol{V}^{M,N}(s)\|_2^2\big)\,\mathrm{d}s\Big),$$

Moreover $\|\boldsymbol{U}^{M,N}(s)\|_1 = \|\boldsymbol{U}^{M,N}(0)\|_1$ a.s. and we recall that $\boldsymbol{U}^{M,N}(t) = \widehat{\boldsymbol{u}}^M(t) - \boldsymbol{Z}^{M,N}(t)$ and $\boldsymbol{V}^{M,N}(t) = \widehat{\boldsymbol{v}}^M(t) - \boldsymbol{W}^{M,N}(t)$ for any $s \geq 0$. Adding that boundedness assumption on the solution of the SKT system and (14) ensure that

$$T\frac{M^2}{N}\|\boldsymbol{U}^{M,N}(0)\|_{1,M} + \frac{M^2}{N}\int_{Q_T} \sigma_M\big(\widehat{\boldsymbol{u}}^M\big)^2 + \sigma_M\big(\widehat{\boldsymbol{v}}^M\big)^2 = T\mathcal{O}\big(\frac{M^2}{N}\big),$$

we finally have

$$\begin{aligned}\frac{1}{M}\sum_{i=1}^{M}\mathbb{E}\Big(\sup_{t\in[0,T]} \mathcal{M}_i^{M,N}(t)^2\Big) &\lesssim \frac{M}{N}\int_0^T \mathbb{E}\big(\|\boldsymbol{Z}^{M,N}(s)\|_2^2 + \|\boldsymbol{W}^{M,N}(s)\|_2^2\big)\,\mathrm{d}s + T\frac{M^2}{N} \\ &\lesssim \frac{M^2}{N}\int_{Q_T} \mathbb{E}\left(\sigma_M\big(\boldsymbol{Z}^{M,N}\big)(s,x)^2 + \sigma_M\big(\boldsymbol{W}^{M,N}\big)(s,x)^2\right)\,\mathrm{d}s\mathrm{d}x + T\frac{M^2}{N} \\ &\lesssim \frac{M^2}{N}|||\boldsymbol{Z}^{M,N}|||_{T,M}^2 + \frac{M^2}{N}|||\boldsymbol{W}^{M,N}|||_{T,M}^2 + T\frac{M^2}{N}.\end{aligned}$$

By symmetry we have bounds of the same order for the terms involving $(\boldsymbol{\mathcal{N}}^{M,N}(t))_{t\geq 0}$. We plug these bounds in (40) and gather the terms $|||\boldsymbol{Z}^{M,N}|||_{T,M}^2$ and $|||\boldsymbol{W}^{M,N}|||_{T,M}^2$ in the left hand side. For N/M^2 large enough, we can control the left hand side and get

$$\begin{aligned}\mathbb{E}(|||\boldsymbol{Z}^{M,N}|||_{T,M}^2 + |||\boldsymbol{W}^{M,N}|||_{T,M}^2) &\lesssim \|\boldsymbol{Z}^{M,N}(0)\|_{-1,M}^2 + T[\boldsymbol{Z}^{M,N}(0)]_M^2[\boldsymbol{\Lambda}^{M,N}(0)]_M \\ &\quad + \|\boldsymbol{W}^{M,N}(0)\|_{-1,M}^2 + T[\boldsymbol{W}^{M,N}(0)]_M^2[\boldsymbol{\Gamma}^{M,N}(0)]_M \\ &\quad + \left(\boldsymbol{\lambda}_T^{M,N} + \boldsymbol{\gamma}_T^{M,N}\right)M^{-4} + T\Big(1 + [\boldsymbol{\Lambda}^{M,N}(0) + \boldsymbol{\Gamma}^{M,N}(0)]_M\Big)\frac{M^2}{N}.\end{aligned}$$

Using that $T[\boldsymbol{u}]_M^2 \leq \|\sigma_M(\boldsymbol{u})\|_{L^2(Q_T)}^2$ for any $\boldsymbol{u} \in \mathbb{R}^M$, by rearranging the terms we conclude the proof. □

Now we can prove the remaining main result.

Proof of Theorem 2. We have

$$\begin{aligned}\zeta^{M,N} &:= \pi_M(\boldsymbol{U}^{M,N}) - u \\ &= \pi_M(\boldsymbol{U}^{M,N} - \widehat{\boldsymbol{u}}^M) + \pi_M(\widehat{\boldsymbol{u}}^M) - u = \pi_M(\boldsymbol{Z}^{M,N}) + \iota_M(u) - u,\end{aligned}$$

where the interpolation operator ι_M is the one used in Lemma 2. Using the triangular inequality, we infer

$$\begin{aligned}&\mathbb{E}\Big[\sup_{t\in[0,T]} \|\zeta^{M,N}(t)\|_{\dot{H}^{-1}(\mathbb{T})}^2 + \|\zeta^{M,N}\|_{L^2(Q_T)}^2\Big] \\ &\qquad \leq \mathbb{E}\Big[\sup_{t\in[0,T]} \|\pi_M(\boldsymbol{Z}^{M,N})(t)\|_{\dot{H}^{-1}(\mathbb{T})}^2 + \|\pi_M(\boldsymbol{Z}^{M,N})\|_{L^2(Q_T)}^2\Big] \\ &\qquad\qquad + \sup_{t\in[0,T]} \|\iota_M(u) - u\|_{\dot{H}^{-1}(\mathbb{T})}^2 + \|\iota_M(u) - u\|_{L^2(Q_T)}^2. \qquad (41)\end{aligned}$$

Now, using Proposition 2 we have that $\|\pi_M(\boldsymbol{Z}^{M,N})\|_{L^2(Q_T)} \leq \|\sigma_M(\boldsymbol{Z}^{M,N})\|_{L^2(Q_T)}$, and using the equivalence (22) of Proposition 6 we get for all $t \in [0,T]$

$$\|\pi_M(\boldsymbol{Z}^{M,N})(t)\|_{H^{-1}(\mathbb{T})} \lesssim \|\boldsymbol{Z}^{M,N}(t)\|_{-1,M} + M^{-1}\|\pi_M(\boldsymbol{Z}^{M,N}(t)\|_{L^2(\mathbb{T})}.$$

This means that the expectation term in the r.h.s. of (41) satisfies the following bound for $M \geq 1$

$$\mathbb{E}\Big[\sup_{t\in[0,T]} \|\pi_M(\boldsymbol{Z}^{M,N})(t)\|_{\dot{H}^{-1}(\mathbb{T})}^2 + \|\pi_M(\boldsymbol{Z}^{M,N})\|_{L^2(Q_T)}^2\Big] \lesssim_T \mathbb{E}\Big[|||\boldsymbol{Z}^{M,N}|||_{T,M}^2\Big].$$

All in all, using Proposition 9, we get

$$\mathbb{E}\Big[\sup_{t\in[0,T]} \|\pi_M(\boldsymbol{Z}^{M,N})(t)\|_{\dot{H}^{-1}(\mathbb{T})}^2 + \|\pi_M(\boldsymbol{Z}^{M,N})\|_{L^2(Q_T)}^2\Big] \lesssim_T \varepsilon_{M,N},$$

where $\varepsilon_{M,N}$ is the r.h.s. of (38). Getting back to (41), we still have to control the second expectation term of its r.h.s., for which invoke Lemma 2 which allow us to write

$$\sup_{t\in[0,T]} \|\iota_M(u) - u\|_{\dot{H}^{-1}(\mathbb{T})}^2 + \|\iota_M(u) - u\|_{L^2(Q_T)}^2 \lesssim M^{-4}\|u\|_{L^\infty\cap L^2([0,T];H^2(\mathbb{T}))}^2.$$

Gathering all the terms leads to the conclusion. □

Appendix

Discrete–continuous dictionary

Discrete	Continuous
Δ_M	Δ
$\lVert\cdot\rVert_{p,M}$	$\lVert\cdot\rVert_{L^p(\mathbb{T})}$
$(\cdot\vert\cdot)_M$	$(\cdot\vert\cdot)_{L^2(\mathbb{T})}$
$\lVert\cdot\rVert_{-1,M}$	$\lVert\cdot\rVert_{H^{-1}(\mathbb{T})}$
$\vert\vert\vert\cdot\vert\vert\vert_{T,M}$	$\vert\vert\vert\cdot\vert\vert\vert_T$
$[\cdot]_M$	$[\cdot]_{\mathbb{T}}$

Acknowledgements The authors thank Joaquin Fontbona for stimulating discussions on this topic. A. M. would like to thank Ariane Trescases for a fruitful discussion which took place at the Tsinghua University in China, back in April 2015, during which the core ideas of Lemma 1 and Theorem 1 where discovered, even though the proof was not yet formalized ([23] did not exist back then!). F. M.-H. acknowledges financial support received under the Doctoral Fellowship ANID-PFCHA/Doctorado Nacional/2017-21171912 and thanks support from Millennium Nucleus Stochastic Models of Complex and Disordered Systems from Millennium Scientific Initiative. This work was partially funded by the Chair "Modélisation Mathématique et Biodiversité" of VEOLIA-Ecole Polytechnique-MNHN-F.X and ANR ABIM 16-CE40-0001.

References

[1] Herbert Amann. Dynamic theory of quasilinear parabolic equations. II. Reaction-diffusion systems. *Differential Integral Equations*, 3(1):13–75, 1990.

[2] Herbert Amann. Erratum: "Dynamic theory of quasilinear parabolic systems. III. Global existence" [Math. Z. **202** (1989), no. 2, 219–250; MR1013086 (90i:35125)]. *Math. Z.*, 205(2):331, 1990.

[3] Maxime Breden, Christian Kuehn, and Cinzia Soresina. On the influence of cross-diffusion in pattern formation. *Journal of Computational Dynamics*, 8(2):213, 2021.

[4] L. Chen and A. Jüngel. Analysis of a parabolic cross-diffusion population model without self-diffusion. *J. Differential Equations*, 224(1):39–59, 2006.

[5] Li Chen, Esther S. Daus, A. Holzinger, and A. Jüngel. Rigorous derivation of population cross-diffusion systems from moderately interacting particle systems. *arXiv:2010.12389 [math]*, 2020.

[6] Li Chen, Esther S. Daus, and Ansgar Jüngel. Rigorous mean-field limit and cross-diffusion. *Zeitschrift für angewandte Mathematik und Physik*, 70(4), July 2019.

[7] Xiuqing Chen and Ansgar Jüngel. A note on the uniqueness of weak solutions to a class of cross-diffusion systems. *Journal of Evolution Equations*, 18(2):805–820, November 2017.

[8] Xiuqing Chen and Ansgar Jüngel. Weak–strong uniqueness of renormalized solutions to reaction–cross-diffusion systems. *Mathematical Models and Methods in Applied Sciences*, 29(02):237–270, February 2019.

[9] Esther S. Daus, Laurent Desvillettes, and Helge Dietert. About the entropic structure of detailed balanced multi-species cross-diffusion equations. *Journal of Differential Equations*, 266(7):3861–3882, March 2019.

[10] Esther S. Daus, Laurent Desvillettes, and Ansgar Jüngel. Cross-diffusion systems and fast-reaction limits. *Bulletin des Sciences Mathématiques*, 159:102824, March 2020.

[11] L. Desvillettes, T. Lepoutre, A. Moussa, and A. Trescases. On the entropic structure of reaction-cross diffusion systems. *Communications in Partial Differential Equations*, 40(9):1705–1747, June 2015.

[12] Gaurav Dhariwal, Florian Huber, and Ansgar Jüngel. Global martingale solutions for a stochastic shigesada-kawasaki-teramoto population model. *arXiv:2012.12765 [math]*, 2020.

[13] Helge Dietert and Ayman Moussa. Persisting entropy structure for nonlocal cross-diffusion systems. *arXiv:2101.02893 [math]*, 2021.

[14] Joaquin Fontbona and Sylvie Méléard. Non local lotka-volterra system with cross-diffusion in an heterogeneous medium. *Journal of Mathematical Biology*, 70(4):829–854, April 2014.

[15] Gonzalo Galiano, María L. Garzón, and Ansgar Jüngel. Semi-discretization in time and numerical convergence of solutions of a nonlinear cross-diffusion population model. *Numer. Math.*, 93(4):655–673, 2003.

[16] Luan T. Hoang, Truyen V. Nguyen, and Tuoc V. Phan. Gradient estimates and global existence of smooth solutions to a cross-diffusion system. *SIAM Journal on Mathematical Analysis*, 47(3):2122–2177, 2015.

[17] M. Iida, M. Mimura, and H. Ninomiya. Diffusion, cross-diffusion and competitive interaction. *Journal of mathematical biology*, 53(4):617–641, 2006.

[18] Ansgar Jüngel and Nicola Zamponi. Qualitative behavior of solutions to cross-diffusion systems from population dynamics. *Journal of Mathematical Analysis and Applications*, 440(2):794–809, August 2016.

[19] Ansgar Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity*, 28(6):1963–2001, may 2015.

[20] Jong Uhn Kim. Smooth solutions to a quasi-linear system of diffusion equations for a certain population model. *Nonlinear analysis*, 8(10):1121–1144, 1984.

[21] Yuan Lou and Michael Winkler. Global existence and uniform boundedness of smooth solutions to a cross-diffusion system with equal diffusion rates. *Communications in Partial Differential Equations*, 40(10):1905–1941, June 2015.

[22] R. H. Martin, Jr. and M. Pierre. Nonlinear reaction-diffusion systems. In *Nonlinear equations in the applied sciences*, volume 185 of *Math. Sci. Engrg.*, pages 363–398. Academic Press, Boston, MA, 1992.

[23] Ayman Moussa. From nonlocal to classical shigesada–kawasaki–teramoto systems: Triangular case with bounded coefficients. *SIAM Journal on Mathematical Analysis*, 52(1):42–64, January 2020.

[24] Du Pham and Roger Temam. A result of uniqueness of solutions of the shigesada–kawasaki–teramoto equations. *Advances in Nonlinear Analysis*, 8(1):497–507, June 2017.

[25] Michel Pierre and Didier Schmitt. Blowup in reaction-diffusion systems with dissipation of mass. *SIAM J. Math. Anal.*, 28(2):259–269, 1997.

[26] Jeremy Quastel. Diffusion of color in the simple exclusion process. *Communications on pure and applied mathematics*, 45(6), 1992.

[27] Nanako Shigesada, Kohkichi Kawasaki, and Ei Teramoto. Spatial segregation of interacting species. *J. Theoret. Biol.*, 79(1):83–99, 1979.

[28] Ariane Trescases. On triangular reaction cross-diffusion systems with possible self-diffusion. *Bulletin des Sciences Mathématiques*, 140(7):796–829, 2016.

Vincent Bansaye
Centre de Mathématiques Appliquées (CMAP), École Polytechnique, CNRS, Palaiseau, France.
E-mail address: vincent.bansaye@polytechnique.edu

Ayman Moussa
Laboratoire Jacques-Louis Lions (LJLL), Sorbonne Université, CNRS, Paris, France.
E-mail address: ayman.moussa@sorbonne-universite.fr

Felipe Muñoz-Hernández
Departamento de Ingeniería Matemática (DIM), Universidad de Chile, Santiago, Chile.
Centre de Mathématiques Appliquées (CMAP), École Polytechnique, CNRS, Palaiseau, France.
E-mail address: fmunozh@dim.uchile.cl