Exact representation of the asymptotic drift speed and diffusion matrix for a class of velocity-jump processes

Corrado Mascia

Dipartimento di Matematica "G. Castelnuovo"
Sapienza – Università di Roma
P.le Aldo Moro, 2 - 00185 Roma (ITALY)

Abstract

This paper examines a class of linear hyperbolic systems which generalizes the Goldstein–Kac model to an arbitrary finite number of speeds $\mathbf{v}^i$ with transition rates μ_{ij}. Under the basic assumptions that the transition matrix is symmetric and irreducible, and the differences $\mathbf{v}^i - \mathbf{v}^j$ generate all the space, the system exhibits a large-time behavior described by a parabolic advection-diffusion equation. The main contribution is to determine explicit formulas for the asymptotic drift speed and diffusion matrix in term of the kinetic parameters $\mathbf{v}^i$ and μ_{ij}, estabilishing a complete connection between microscopic and macroscopic coefficients. It is shown that the drift speed is the arithmetic mean of the velocities $\mathbf{v}^i$. The diffusion matrix has a more complicate representation, based on the graph with vertices the velocities $\mathbf{v}^i$ and arcs weighted by the transition rates μ_{ij}. The approach is based on an exhaustive analysis of the dispersion relation and on the application of a variant of the Kirchoff's matrix tree Theorem from graph theory.

Key words: Velocity-jump processes, Hyperbolic systems with relaxation, Graph theory

2000 MSC: 35L45 Secondary 05C05, 82C21

1. Introduction

The *Goldstein–Kac model* for correlated random walk ([7], [15]) is a differential description of a *velocity-jump process* consisting in a first order linear hyperbolic system for the couple $(f, g) = (f, g)(x, t)$ given by

$$\frac{\partial f}{\partial t} - \nu \frac{\partial f}{\partial x} = -\mu f + \mu g, \qquad \frac{\partial g}{\partial t} + \nu \frac{\partial g}{\partial x} = \mu f - \mu g$$

where $x, t \in \mathbb{R}$ and $\nu, \mu > 0$. Variables f and g represent "densities" of individuals moving, respectively, toward the left and toward the right of a one-dimensional line with velocity ν. The linear term at the right-hand side describes the fact that reversal of speed is possible with a transition rate μ. The Goldstein–Kac model was originally motivated by G.I. Taylor [23] in what M. Kac describes as "an abortive, or at least not very successful, attempt to treat turbulent diffusion.", see [15]. Even if perhaps unproductive in its original intent, the model gained a lot of attention because of its quality of being located at the crossroads of amenability and significance. "Amenability" emerges from its linear hyperbolic structure, giving raise to well-posedness and preservation of smoothness, and its quasi-monotonicity, guaranteeing the validity of a comparison principle, see [20]. "Significance" stems mainly from the fact that the

Goldstein–Kac is a prototype for a differential description of transport mechanisms. In addition, as a consequence of the fact that the sum $u = f + g$ satisfies the *telegraph equation*

$$2\mu \frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial t^2} - \nu^2 \frac{\partial^2 u}{\partial x^2} = 0,$$

such model is considered a paradigm for *hyperbolic diffusion*, hence alternative to the traditional parabolic heat equation (among many others, see [16]). In the long-run, solutions of the telegraph equation behaves as the solution to the heat equation

$$\frac{\partial u}{\partial t} - \varepsilon \frac{\partial^2 u}{\partial x^2} = 0,$$

where the coefficient ε is given by $\nu^2/2\mu$ (see [18]). The exact relation between the parameters ε, μ, ν is crucial to obtain a sharp description of the parabolic asymptotic behavior of the solution to the Goldstein–Kac model.

Here, attention is drawn to a class of first order linear hyperbolic systems in several space dimensions, describing a process where individuals may change velocity of propagation at a given rate. The key question is the relation between large-time macroscopic coefficients with the microscopic kinetic parameters: *how to provide a* <u>*quantitative*</u> *description of the asymptotically parabolic nature of the class of hyperbolic equations under observation?* The aim is to determine a generalization of the formula relation $\varepsilon = \nu^2/2\mu$, which connect velocity ν and transition rate μ with the diffusion coefficient ε in the case of the Goldstein–Kac model.

To enter the details, let $\mathbb{R}^d$ be the ambient space for the space variable $\mathbf{x}$ and let a fixed finite set of speeds $\{\mathbf{v}^1, \dots, \mathbf{v}^n\} \subset \mathbb{R}^d$ be given. The changes from one velocity to another are described by kinetic parameters $\mu_{ij} \geq 0$, with $i, j \in \{1, \dots, n\}$ and $i \neq j$, which measure the rate of transition from speed $\mathbf{v}^i$ to speed $\mathbf{v}^j$. Given such data, we consider the system for the unknown $\mathbf{f} = (f_1, \dots, f_n)$

$$\frac{\partial f_i}{\partial t} + \mathbf{v}^i \cdot \nabla_{\mathbf{x}} f_i + \sum_{j \neq i} (\mu_{ij} f_i - \mu_{ji} f_j) = 0, \qquad i = 1, \dots, n. \tag{1}$$

The Goldstein–Kac model corresponds to the one-dimensional case with two speeds, namely $d = 1$, $n = 2$, $\{\mathbf{v}^1 = -\nu, \mathbf{v}^2 = +\nu\} \subset \mathbb{R}^1$, and with $\mu_{12} = \mu_{21} = \mu > 0$. Alternatively, model (1) can also be regarded as a discrete velocity version of the model treated in [12] or as a linearization of a discrete velocity Boltzmann system.

Similarly to the destiny of the Goldstein–Kac model, system (1) can be considered as a backbone for more complicated models taking into account the dependencies on external signals, as in the case of chemotaxis (see [13, 8]), or the presence of reaction/reproduction mechanisms (see [10, 17]). Endorsing the interest of such kind of generalizations, specific one-dimensional versions of (1) has already been considered in [4, 5] in the description of transport along axons.

In this paper, model (1) is examined under the assumption that the transition rates are symmetric, $\mu_{ij} = \mu_{ji}$. The non-symmetric case is natural in the modeling of phenomena where the transport is guided by the gradient of some substance and it is left for future investigations. No specific restriction is made on the number and on the choice of the propagation velocities $\mathbf{v}^i$.

The main contribution consists in the determination of an advection-diffusion equation that describes the large time behavior of the cumulative unknown $u := \sum f_i$. Unsurpringly, such equation has the form

$$\frac{\partial w}{\partial t} + \mathbf{v}_{\text{drift}} \nabla_x w = \operatorname{div}\left(\mathbb{D}\nabla_{\mathbf{x}} w\right) \tag{2}$$

with $\mathbf{v}_{\text{drift}} \in \mathbb{R}^d$ and $\mathbb{D} \in \mathbb{R}^d \times \mathbb{R}^d$. What is much more significant is the fact that *an explicit expression of the macroscopic terms* $\mathbf{v}_{drift}$ *and* $\mathbb{D}$ *in term of the kinetic parameters* $\mathbf{v}^i$ *and* $\mu_{ij} \geq 0$ *is rigorously deduced.* The drift speed $\mathbf{v}_{\text{drift}}$ is the arithmetic average of the speeds $\mathbf{v}^i$, viz.

$$\mathbf{v}_{\text{drift}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{v}^i$$

(see Theorem 3.7); the *diffusion matrix* $\mathbb{D}$ is represented by introducing an associated (undirected) graph Γ, whose vertices are the velocities $\mathbf{v}^i$ and whose arcs are weighted by the transition rates μ_{ij}. The representation formula for the matrix $\mathbb{D}$ (see Theorem 3.11), anticipated here for the reader's convenience, is

$$\mathbb{D} = \frac{1}{I_1(\mathbb{B})}\sum_{\mathcal{F}_2} \mu_{T^1}\mu_{T^2}\left(\mathbf{w}(T_1)\otimes\mathbf{w}(T_1)\right),$$

where $I_1(\mathbb{B})$ is an invariant of the matrix $\mathbb{B} = (\mu_{ij})$, the sum is on the set $\mathcal{F}_2$ composed by the forests partitioning the verteces of the graph Γ in two (disjoint) trees T_1 and T_2, μ_T is the products of the transition rate along a tree T and $\mathbf{w}(T)$ the sum of the velocities $\mathbf{v}^i$ in a given tree T (for a more detailed description, see Section 3). The proof of the representation formula is based on a variant of a well-known result in graph theory: the *Kirchoff's matrix tree Theorem* (see [3] and references therein). For both the statements, the symmetry assumption of transition rates is crucial.

The paper is divided into three more Section (Introduction excluded). In Section 2, the basic properties of system (1) are discussed. After a brief description on how the model can be formally derived starting from a correlated random walk, it is showed that any convex function from $\mathbb{R}$ to $\mathbb{R}$ can be used to define a Lyapunov functional for the system. Such structure implies that L^p−norm are preserved and that a form of comparison principle holds. Moreover, under appropriate assumptions, system (1) fits into the class considered in [22], thus dictating a specific form of dissipation mechanism. Section 3 is the hub of the paper. Applying the Laplace–Fourier transform, the hyperbolic system is converted into its *dispersion relation*, which encodes the different modes supported by the model. The analysis of the dispersion relation close to the origin describes the large time behavior of the solutions, and thus furnishes the expressions for the drift speed $\mathbf{v}_{\text{drift}}$ and the diffusion matrix $\mathbb{D}$. Finally, in the sake of completeness, Section 4 is devoted to state and prove a result supporting the fact that the parabolic equation (2) gives the asymptotic description of (1): chosen an initial datum $\mathbf{f}_0 \in [L^1 \cap L^2(\mathbb{R}^d)]^n$ for (1), for $u := \sum_i f_i$ and u_{par} solution to (2), $u(x,0) = \sum_i f_{0,i}(x)$, there holds

$$|u - u_{\text{par}}|_{L^2}(t) \leq C\, t^{-\frac{1}{4}d-\frac{1}{2}} |\mathbf{f}_0|_{L^1\cap L^2}$$

(see Theorem 4.1). The rate of the L^2−decay for (2) with data in $L^1 \cap L^2$ is $t^{-d/4}$, thus the above estimate shows that the hyperbolic variable u and its parabolic counterpart u_{par} get closer one to the other in a time-scale shorter than the one of their ultimate decay to zero. This kind of estimate fits into a wide research stream exploring asymptotically parabolic nature of hyperbolic equations, which dates back at least to J. Hadamard, [9]. Additional bibliographical description on the subject is given in Section 4.

2. Velocity-jump processes with a finite number of speeds

Given $d \geq 1$, let us consider a family of n velocities $\mathbf{v}^1, \dots, \mathbf{v}^n \in \mathbb{R}^d$, with components $\mathbf{v}^i = (v^i_j)$, together with parameters $\mu_{ij} \geq 0$ for $i, j = 1, \dots, n$, $i \neq j$, describing the transition rate from speed $\mathbf{v}^i$ to speed $\mathbf{v}^j$. Given a population with total density $u = u(x,t)$, all the individuals are allowed to move with one of the speeds $\mathbf{v}^1, \dots, \mathbf{v}^n \in \mathbb{R}^d$. Denoted by $f_i = f_i(x,t)$ the density for the portion of the total population proceeding with velocity $\mathbf{v}^i$ and assuming that the speed change is described by the rates μ_{ij}, the dynamics is dictated by the first order linear hyperbolic system

$$\frac{\partial f_i}{\partial t} + \mathbf{v}^i \cdot \nabla_{\mathbf{x}} f_i + \sum_{j \neq i} (\mu_{ij} f_i - \mu_{ji} f_j) = 0, \qquad i = 1, \dots, n, \tag{3}$$

or, in vectorial form, with $\mathbf{f} = (f_1, \dots, f_n)$,

$$\frac{\partial \mathbf{f}}{\partial t} + \sum_{j=1}^{d} \mathbb{A}_j \frac{\partial \mathbf{f}}{\partial x_j} + \mathbb{B}\mathbf{f} = 0$$

where $\mathbb{A}_j = \mathrm{diag}\,(v^i_j)$ and $\mathbb{B} = (-\mu_{ij})$ with $\mu_{ii} := -\sum\limits_{j \neq i} \mu_{ji}$. If the coefficient μ_{ij} is zero for some i, j, there is no direct transition from the speed $\mathbf{v}^i$ to speed $\mathbf{v}^j$.

The Cauchy problem for (3) determined by the initial condition

$$f_i(\mathbf{x}, 0) = f_{0,i}(\mathbf{x}) \qquad \mathbf{x} \in \mathbb{R}^d, \qquad i = 1, \dots, n \tag{4}$$

has a unique (mild) solution continuously dependent on the initial data whenever the initial datum $\mathbf{f}_0 = (f_{0,i})$ is chosen in an appropriate functional space. Later on, we will concentrate on the case $f_0 \in [L^1 \cap L^2(\mathbb{R}^d)]^n$; for the moment, we continue the discussion with choices of initial data depending case by case.

The *transition matrix* $\mathbb{B}$ is singular since the sum of its columns is zero. Moreover, the total density $u = \sum f_i$ satisfies the homogeneous transport equation

$$\frac{\partial u}{\partial t} + \mathrm{div}\,\mathbf{j} = 0 \qquad \text{where} \quad \mathbf{j} := \sum_{i=1}^{n} f_i \mathbf{v}^i.$$

Since 0 has algebraic multiplicity one and the remaining part of the spectrum of $\mathbb{B}$ has negative real part, as time goes to $+\infty$, the dynamics of the solution $\mathbf{f}$ to (3)–(4) is expected to be described by the single scalar quantity $u = \sum f_i$.

Derivation from a correlated random walk

System (3) can be heuristically derived from a correlated random walk, in the same spirit of what is usually done for the Goldstein–Kac model. Given the velocities $\{\mathbf{v}^1, \dots, \mathbf{v}^n\} \subset \mathbb{R}^d$ and the transition rates μ_{ij}, with $i, j \in \{1, \dots, n\}$, $i \neq j$, let $\mathrm{dt} > 0$ be such that

$$p_i := 1 - \sum_{j \neq i} \mu_{ij} \mathrm{dt} \geq 0 \qquad i = 1, \dots, n.$$

Let X be defined by

$$X := \left\{ \mathbf{x} \in \mathbb{R}^d \ : \ \mathbf{x} = \sum_{i=1}^{n} c_i \, \mathbf{v}^i \mathrm{dt} \quad \text{for some} \quad c_1, \dots, c_n \in \mathbb{N} \right\}.$$

Then, assume that each particle of a given finite set is located at time $t = 0$ at some $\mathbf{x} \in X$ and it has a given state $i \in \{1, \dots, n\}$, corresponding to a "preferential" speed. The set of particles with state i will be denoted by F_i. At each time interval dt, the displacement of every particle in F_i amounts to $\mathrm{dx} = \mathbf{v}^i \mathrm{dt}$ with a probability p_i and to $\mathrm{dx} = \mathbf{v}^j \mathrm{dt}$ with a probability $\mu_{ij}\mathrm{dt}$. In the latter case, the particle changes state from i to j.

Denoting by $f_i(\mathbf{x}, n\,\mathrm{dt})$ the fraction of particles with state i at time $t = n\,\mathrm{dt}$ and at position $\mathbf{x}$, the relation between the values f_i at step n and $n+1$ is

$$f_i(\mathbf{x}, (n+1)\mathrm{dt}) = p_i f_i(\mathbf{x} - \mathbf{v}^i \mathrm{dt}, n\,dt) + \sum_{j \neq i} \mu_{ji} f_j(\mathbf{x} - \mathbf{v}^i \mathrm{dt}, n\,\mathrm{dt})\mathrm{dt}.$$

Adding and subtracting the term $f_i(\mathbf{x} + \mathbf{v}^i \mathrm{dt}, n)$, we obtain

$$\begin{aligned} f_i(\mathbf{x} + \mathbf{v}^i \mathrm{dt}, (n+1)\mathrm{dt}) - f_i(\mathbf{x} + \mathbf{v}^i \mathrm{dt}, n\,\mathrm{dt}) + f_i(\mathbf{x} + \mathbf{v}^i \mathrm{dt}, n\,\mathrm{dt}) - f_i(\mathbf{x}, n\,\mathrm{dt}) \\ = \Big(\sum_{j \neq i} \mu_{ij} \Big) f_i(\mathbf{x}, n\,\mathrm{dt})\mathrm{dt} + \sum_{j \neq i} \mu_{ji} f_j(\mathbf{x}, n\,\mathrm{dt})\mathrm{dt} \end{aligned}$$

For time interval dt small, assuming f_i to be smooth with respect to its first argument, we may approximate the difference $f_i(\mathbf{x} + \mathbf{v}^i \mathrm{dt}, n\,\mathrm{dt}) - f_i(\mathbf{x}, n\,\mathrm{dt})$ with the scalar product between the gradient of f with respect to $\mathbf{x}$ and the increment $\mathbf{v}^i \mathrm{dt}$, getting the relation

$$\frac{1}{\mathrm{dt}} \{ f_i(\mathbf{x} + \mathbf{v}^i \mathrm{dt}, (n+1)\mathrm{dt}) - f_i(\mathbf{x} + \mathbf{v}^i \mathrm{dt}, n\,\mathrm{dt}) \} + \mathbf{v}^i \cdot \nabla_{\mathbf{x}} f_i + \sum_{j \neq i} (\mu_{ij} f_i - \mu_{ji} f_j) \approx 0.$$

Passing to the limit $\mathrm{dt} \to 0$, we formally obtain (3).

Properties of the first order linear system.

If the *transition rates* μ_{ij} are assumed to be symmetric, i.e.

$$\mu_{ij} = \mu_{ji} \qquad \forall i \neq j, \tag{5}$$

the structure of the zero-th order term in (3) and the symmetry of the matrix $\mathbb{B}$ trigger a number of additional properties.

Proposition 2.1. *Assume hypothesis* (5). *Let* $\eta : \mathbb{R} \to \mathbb{R}$ *be a convex Lipschitz continuous function. Then, each solution* $\mathbf{f}$ *to* (3) *is such that, whenever the right-hand side is finite,*

$$\sum_{i=1}^{n} \int_{\mathbb{R}^d} \eta(f_i)(\mathbf{x}, t_2)\, d\mathbf{x} \leq \sum_{i=1}^{n} \int_{\mathbb{R}^d} \eta(f_i)(\mathbf{x}, t_1)\, d\mathbf{x} \qquad\qquad t_1 < t_2. \tag{6}$$

Proof. Let us consider the case of a smooth initial datum $\mathbf{f}_0$ so that the solution $\mathbf{f}$ is also smooth. The general case can be obtained by applying a density argument.

Given a Lipschitz continuous function η, multiplying (3) by $\eta'(f_i)$ and summing up with respect to i, we infer

$$\frac{\partial}{\partial t}\sum_{i=1}^{n}\eta(f_i)+\sum_{\ell=1}^{d}\frac{\partial}{\partial x_\ell}\left(\sum_{i=1}^{n}v_i^\ell\,\eta(f_i)\right)+\sum_{i=1}^{n}\sum_{j\neq i}\eta'(f_i)(\mu_{ji}\,f_i-\mu_{ij}\,f_j)=0.$$

Given $t_1 < t_2$, integrating with respect to (x,t) in $\mathbb{R}^d\times[t_1,t_2]$, we get

$$\sum_{i=1}^{n}\int_{\mathbb{R}^d}\eta(f_i)(\mathbf{x},t_2)\,d\mathbf{x}+\int_{\mathbb{R}^d}\mathcal{I}[\mathbf{f}]\,d\mathbf{x}\,dt=\sum_{i=1}^{n}\int_{\mathbb{R}^d}\eta(f_i)(\mathbf{x},t_1)\,d\mathbf{x},$$

where

$$\mathcal{I}[\mathbf{f}]:=\sum_{i=1}^{n}\sum_{j\neq i}\eta'(f_i)(\mu_{ji}\,f_i-\mu_{ij}\,f_j).$$

Since $\mu_{ij}=\mu_{ji}$, there holds

$$\mathcal{I}[\mathbf{f}]=\sum_{i,j=1}^{n}\eta'(f_i)(\mu_{ji}\,f_i-\mu_{ij}\,f_j)=\sum_{i,j=1}^{n}\mu_{ji}\left(\eta'(f_i)-\eta'(f_j)\right)f_i=\frac{1}{2}\sum_{i,j=1}^{n}\mu_{ij}\left(\eta'(f_i)-\eta'(f_j)\right)(f_i-f_j)\geq 0, \tag{7}$$

for η' increasing and $\mu_{ij}\geq 0$. The conclusion follows. □

In particular, the solution semigroup of (3) is such that the L^p−norms are (weakly) decreasing in time for any $p\geq 1$. Additionally, a comparison principle holds.

Corollary 2.2. *Let* $\mathbf{f}$ *and* $\mathbf{g}$ *be two solutions to* (3) *corresponding to the initial conditions* $\mathbf{f}(\mathbf{x},0)=\mathbf{f}_0(\mathbf{x})$ *and* $\mathbf{g}(\mathbf{x},0)=\mathbf{g}_0(\mathbf{x})$*, respectively. If* $\mathbf{f}_0$ *and* $\mathbf{g}_0$ *are such that* $f_{0,i}\leq g_{0,i}$ *for any* $i=1,\dots,n$*, then the same ordering relation holds for any positive time, i.e.* $f_i(\mathbf{x},t)\leq g_i(\mathbf{x},t)$ *for any* $i=1,\dots,n$ *and for any* $t>0$*.*

Proof. Linearity permits to restrict the attention to the case $g_0=0$. Choosing $\eta(s)=[s]_+$ in (6), we deduce

$$\sum_{i=1}^{n}\int_{\mathbb{R}^d}[f_i(\mathbf{x},t)]_+\,d\mathbf{x}\leq\sum_{i=1}^{n}\int_{\mathbb{R}^d}[f_i(\mathbf{x},0)]_+\,d\mathbf{x}.$$

In particular, if $f_i\leq 0$ for any i at time $t=0$, then $f_i\leq 0$ for any $t\geq 0$ for any i. □

If the transition matrix $\mathbb{B}$ is irreducible (see [1], Chap.2, Sect.2) and an appropriate additional assumption on the velocities $\mathbf{v}^i$ holds, the *Shizuta–Kawashima dissipativity condition* is valid (see [22] and, for a discussion on its limits of validity, see [19]).

Proposition 2.3. *Let the matrix* $\mathbb{B}$ *be symmetric and irreducible. If*

$$\operatorname{span}\{\mathbf{v}^i-\mathbf{v}^j\ :\ i,j=1,\dots,n\}=\mathbb{R}^d, \tag{8}$$

then there holds

$$\lambda F+\sum_{j=1}^{n}k_j\mathbb{A}_jF\neq 0\qquad\qquad\forall\,F\in\ker\mathbb{B}$$

for any $\lambda\in\mathbb{R}$ *and any* $\mathbf{k}=(k_1,\dots,k_d)$ *with* $\mathbf{k}\neq 0$*.*

Proof. The specific structure of the matrix $\mathbb{B}$ together with its irreducibility imply that $\ker \mathbb{B}$ is generated by the vector $\mathbf{1} = (1, \dots, 1)$. Since $(\mathbb{A}_j \mathbf{1})_i = v^i_j$, we infer

$$\lambda + \sum_{j=1}^{n} k_j (\mathbb{A}_j \mathbf{1})_i = \lambda + \sum_{j=1}^{n} k_j v^i_j = \lambda + \mathbf{k} \cdot \mathbf{v}^i$$

with $i = 1, \dots, n$. If $\lambda + \mathbf{k} \cdot \mathbf{v}^i = 0$ for any i, then $\mathbf{k} \cdot (\mathbf{v}^i - \mathbf{v}^j) = 0$ for any i, j. Thus, as a consequence of assumption (8), the vector $\mathbf{k}$ is zero. □

The role of the two hypotheses in Proposition 2.3 can be clarified showing that under these assumptions, invoking the *LaSalle Invariance Principle*, the solution to the Cauchy problem (3)–(4) decayes to zero as $t \to +\infty$ for integrable initial datum in some L^p. Indeed, the ω−limit of a given trajectory is contained in the largest invariant region for which the functional $\sum \int \eta(f_i)$ is constant. For η' strictly increasing, the final representation of $\mathcal{I}[\mathbf{f}]$ in (7) together with the requirement of the irreducibility of the matrix $\mathbb{B}$, shows that

$$\mathcal{I}[\mathbf{f}] = 0 \qquad \Longleftrightarrow \qquad f_i = f_j \quad \forall\, i, j.$$

Hence, solutions preserving the value of $\sum \int \eta(f_i)$ are such that $f_i = g$ for some scalar function g satisfying the homogeneous system

$$\frac{\partial g}{\partial t} + \mathbf{v}^i \cdot \nabla_x g = 0, \qquad i = 1, \dots, n.$$

which gives

$$(\mathbf{v}^i - \mathbf{v}^j) \cdot \nabla_{\mathbf{x}} g = 0, \qquad i, j = 1, \dots, n.$$

Thus, as a consequence of (8), g is constant.

A rigorous application of the principle requires some compactness of the trajectories, guaranteed by estimates in some appropriate Sobolev space, at the price of some additional regularity requirements on the initial data. Indeed, taking again advantage of the dissipation described by Proposition 2.1 and the linear structure of the equation, one obtains

$$\sum_{i=1}^{n} |f_i(\cdot, t)|^p_{W^{k,p}} \le \sum_{i=1}^{n} |f_{0,i}(\cdot, t)|^p_{W^{k,p}} \qquad \text{for any } k \ge 0, p \ge 1.$$

All the above arguments concur in showing that the system (3) shares a number of properties with scalar linear diffusion equations. Rigorous statement on such a asymptotic diffusive behavior can be obtained applying results already present in the literature, such as the general theory in [2] (which is applicable thanks to Proposition 2.3).

Our next aim is to show that *it is possible to identify a specific diffusion equation of parabolic type describing the large-time behavior of the solutions to the hyperbolic system* (3). Such identification is achieved by means of explicit expressions for both the asymptotic drift speed and diffusion matrix, which give a complete connection between macroscopic and microscopic parameters.

3. Drift velocity and diffusion matrix

Given a square matrix $\mathbb{A}$ of dimension $n \times n$ and k indeces $i_1, \dots, i_k$, let $\mathbb{A}(i_1, \dots, i_k)$ be the principal minor that results from deleting sets of k rows and columns with indeces $i_1, \dots, i_k$. By definition, we set $\mathbb{A}(1, \dots, n) := 1$. Moreover, given n column vectors $w_1, \dots, w_n$ in $\mathbb{R}^n$, let

$$w_1 \wedge \cdots \wedge w_n := \det(w_1 \dots w_n).$$

Applying Laplace–Fourier transform, that is converting time derivatives with multiplication by $\lambda \in \mathbb{R}$ and space derivatives with scalar multiplication by $\mathbf{k} \in i\mathbb{R}^d$, the partial differential equation (3) is turned into

$$(\lambda + \mathbf{v}^i \cdot \mathbf{k})\hat{f}_i + \sum_{j=1}^{n} \mu_{ij} \hat{f}_j = 0 \qquad i = 1, \dots, n,$$

where $\hat{f}_i = \hat{f}_i(\lambda, \mathbf{k})$ is the transform of f. The *dispersion relation* of system (3) is the polynomial relation in λ and $\mathbf{k}$

$$p(\lambda, \mathbf{k}) := \det(\lambda\,\mathbb{I} + \mathrm{diag}\,(\mathbf{v}^i \cdot \mathbf{k}) + \mathbb{B}) = 0.$$

Denoting by B_i the columns of the matrix $\mathbb{B}$, the dispersion relation can be written in compact form as

$$p(\lambda, \mathbf{k}) = V_1(\lambda, \mathbf{k}) \wedge \cdots \wedge V_n(\lambda, \mathbf{k}) = 0.$$

where $V_i(\lambda, \mathbf{k}) := (\lambda + \mathbf{v}^i \cdot \mathbf{k})E_i + B_i$ and $E_i = (\delta_{ij})_{j=1,\dots,n}$ is the i-th element of the canonical basis of $\mathbb{R}^n$.

Given $\mathbf{k}$, the main target of the analysis is to locate the values $\lambda = \lambda(\mathbf{k})$ such that the dispersion relation $p(\lambda, \mathbf{k}) = 0$ is satisfied. In particular, being interested in the large-time behavior of the solutions, the attention is mainly directed to the region of $\mathbf{k}$ with $|\mathbf{k}|$ small.

Proposition 3.1. *Let $\mathbb{B}$ be a real $n \times n$ singular matrix such that*

$$I_1(\mathbb{B}) := \sum_{i=1}^{n} \det \mathbb{B}(i) \neq 0.$$

Then, there is a smooth function $\mathbf{k} \mapsto \lambda(\mathbf{k})$ defined in a neighborhood of $\mathbf{0}$ such that $p(\lambda, \mathbf{k}) = 0$ if and only if $\lambda = \lambda(\mathbf{k})$. Moreover, there hold

$$1^{st} \text{ order:} \qquad \frac{\partial \lambda}{\partial k_\ell}(\mathbf{0}) = -\frac{1}{I_1(\mathbb{B})} \sum_{i=1}^{n} v^i_\ell \det \mathbb{B}(i),$$

$$2^{nd} \text{ order:} \qquad \frac{\partial^2 \lambda}{\partial k_j \partial k_\ell}(\mathbf{0}) = -\frac{1}{I_1(\mathbb{B})} \sum_{\{h,k\,:\,h \neq k\}} (v^j_h\, v^\ell_k)\, \mathbb{B}(h, k) \qquad \text{if } \nabla_{\mathbf{k}} \lambda(\mathbf{0}) = \mathbf{0}.$$

Proof. Since $\det \mathbb{B} = 0$, the couple $(0, \mathbf{0})$ satisfies the dispersion relation. Moreover, at $(\lambda, \mathbf{k})$

$$\begin{aligned}\frac{\partial p}{\partial \lambda} &= \frac{\partial V_1}{\partial \lambda} \wedge V_2 \wedge \cdots \wedge V_n + \cdots + V_1 \wedge \cdots \wedge V_{n-1} \wedge \frac{\partial V_n}{\partial \lambda} \\ &= E_1 \wedge V_2 \wedge \cdots \wedge V_n + \cdots + V_1 \wedge \cdots \wedge V_{n-1} \wedge E_n,\end{aligned}$$

there holds

$$\frac{\partial p}{\partial \lambda}(0,\mathbf{0}) = E_1 \wedge \cdots \wedge B_n + \cdots + B_1 \wedge \cdots \wedge E_n = I_1(\mathbb{B}),$$

and the existence of the function λ follows from the Implicit function Theorem.

Moreover, differentiating with respect to k_ℓ for $\ell \in \{1,\ldots,n\}$, we obtain at $(\lambda,\mathbf{k})$

$$\begin{aligned}\frac{\partial p}{\partial k_\ell} &= \frac{\partial V_1}{\partial k_\ell} \wedge \cdots \wedge V_n + \cdots + V_1 \wedge \cdots \wedge \frac{\partial V_n}{\partial k_\ell} \\ &= v_\ell^1 E_1 \wedge \cdots \wedge V_n + \cdots + V_1 \wedge \cdots \wedge v_\ell^n E_n;\end{aligned} \tag{9}$$

and thus, calculating at $(0,\mathbf{0})$, we infer

$$\frac{\partial p}{\partial k_\ell}(0,\mathbf{0}) = v_\ell^1 E_1 \wedge \cdots \wedge B_n + \cdots + B_1 \wedge \cdots \wedge v_\ell^n E_n = \sum_{i=1}^n v_\ell^i \det \mathbb{B}(i).$$

which gives the first order expansion.

Differentating with respect to k_ℓ and then with respect to k_j the relation $p(\lambda(k),k) = 0$, we get the equality

$$\frac{\partial^2 p}{\partial \lambda^2}\frac{\partial \lambda}{\partial k_\ell}\frac{\partial \lambda}{\partial k_j} + \frac{\partial^2 p}{\partial \lambda \partial k_j}\frac{\partial \lambda}{\partial k_\ell} + \frac{\partial p}{\partial \lambda}\frac{\partial^2 \lambda}{\partial k_j \partial k_\ell} + \frac{\partial^2 p}{\partial \lambda \partial k_\ell}\frac{\partial \lambda}{\partial k_j} + \frac{\partial^2 p}{\partial k_j \partial k_\ell} = 0.$$

Calculating at $(0,\mathbf{0})$, since $\partial_\lambda p(0,\mathbf{0}) = I_1(\mathbb{B})$ and $\nabla_{\mathbf{k}}\lambda(\mathbf{0}) = \mathbf{0}$, we obtain

$$\frac{\partial^2 \lambda}{\partial \kappa_j \partial \kappa_\ell}(\mathbf{0}) = -\frac{1}{I_1(\mathbb{B})}\frac{\partial^2 p}{\partial \kappa_j \partial \kappa_\ell}(0,\mathbf{0}).$$

Upon differentiation of (9), we deduce at $(\lambda,\mathbf{k})$

$$\begin{aligned}\frac{\partial^2 p}{\partial k_j\,\partial k_\ell} &= \frac{\partial}{\partial k_j}\left(v_1^\ell E_1 \wedge \cdots \wedge V_n\right) + \cdots + \frac{\partial}{\partial k_j}\left(V_1 \wedge \cdots \wedge v_n^\ell E_n\right) \\ &= \sum_{\{h,k\,:\,h\neq k\}} (v_h^j\, v_k^\ell)\, V_1 \wedge \cdots \wedge E_h \wedge \cdots \wedge E_k \wedge \cdots \wedge V_n.\end{aligned}$$

Therefore, calculating at $(0,\mathbf{0})$, we end up with

$$\frac{\partial^2 p}{\partial \kappa_j\,\partial \kappa_\ell}(0,\mathbf{0}) = \sum_{\{h,k\,:\,h\neq k\}} (v_h^j\, v_k^\ell)\, B_1 \wedge \cdots \wedge E_h \wedge \cdots \wedge E_k \wedge \cdots \wedge B_n = \sum_{\{h,k\,:\,h\neq k\}} (v_h^j\, v_k^\ell)\, \mathbb{B}(h,k),$$

that gives the second order derivatives of $\lambda = \lambda(\mathbf{k})$ in the origin. □

Condition $I_1(\mathbb{B}) \neq 0$ is satisfied if the matrix $\mathbb{B}$ is irreducible. From now on, we assume such hypothesis.

The expressions for the first and second order term in the expansion of the function λ at $\mathbf{k} = 0$ can be restyled. The gradient of λ can be rewritten as

$$\nabla_{\mathbf{k}}\lambda(\mathbf{0}) = -\frac{1}{I_1(\mathbb{B})}\left((\mathbb{B}(1),\ldots,\mathbb{B}(n)) \cdot (v_j^1,\ldots,v_j^n)\right)_{j=1,\ldots,d}$$

Also, when the gradient of λ at $\mathbf{k} = \mathbf{0}$ is null, the hessian matrix of λ at the same point is

$$D^2\lambda(\mathbf{0}) = -\frac{2}{I_1(\mathbb{B})}\sum_{h<k} \mathbb{B}(h,k)\,(\mathbf{v}^h \otimes \mathbf{v}^k)^*$$

where $\otimes$ is the tensor product of vectors and * denotes the symmetric part of a matrix.

With a terminology that will be fully motivated in the subsequent Section, we define the *drift velocity* of the system (3) to be the vector

$$\mathbf{v}_{\text{drift}} := -\nabla_{\mathbf{k}}\lambda(\mathbf{0}) = \frac{1}{I_1(\mathbb{B})}\big((\det\mathbb{B}(1),\dots,\det\mathbb{B}(n))\cdot(v_j^1,\dots,v_j^n)\big) \tag{10}$$

and the *diffusion matrix* of the system (3) as

$$\mathbb{D} := \frac{1}{2}\,D^2\lambda(\mathbf{0}) = -\frac{1}{I_1(\mathbb{B})}\sum_{i<j}\mathbb{B}(i,j)\,(\mathbf{v}^i\otimes\mathbf{v}^j)^*. \tag{11}$$

In a frame moving with speed $\mathbf{v}_{\text{drift}}$, the velocities $\mathbf{v}^i$ in the system are modified in $\mathbf{v}^i - \mathbf{v}_{\text{drift}}$ and coherently, the gradient of λ at $\mathbf{0}$ is null. Thus, without loss of generality, we may assume that the conditions

$$(\det\mathbb{B}(1),\dots,\det\mathbb{B}(n))\cdot(v_j^1,\dots,v_j^n) = \mathbf{0}$$

henceforth holds for $j = 1,\dots,d$.

Example 3.2. [Modified Goldstein–Kac model] The simplest example needs two velocities $\mathbf{v}^1,\mathbf{v}^2$ and a transition matrix of the form

$$\mathbb{B} = \begin{pmatrix}\mu_{12} & -\mu_{21}\\ -\mu_{12} & \mu_{21}\end{pmatrix}.$$

for some positive values μ_{12},μ_{21}. Since $\mathbb{B}(1) = \mu_{21}$ and $\mathbb{B}(2) = \mu_{12}$, the drift velocity of the system is

$$\mathbf{v}_{\text{drift}} = \frac{\mu_{21}\mathbf{v}^1 + \mu_{12}\mathbf{v}^2}{\mu_{21}+\mu_{12}}.$$

Since $\det\mathbb{B}(1,2) = 1$, if $\mu_{21}\mathbf{v}^1 + \mu_{12}\mathbf{v}^2 = 0$, there holds

$$\mathbb{D} = -\frac{1}{\mu_{21}+\mu_{12}}(\mathbf{v}^1\otimes\mathbf{v}^2)^* = \frac{\mu_{21}/\mu_{12}}{\mu_{21}+\mu_{12}}(\mathbf{v}^1\otimes\mathbf{v}^1)$$

When the symmetry condition $\mu_{21} = \mu_{12}$ holds, $\mathbf{v}_{\text{drift}}$ is the algebraic mean of $\mathbf{v}^1$ and $\mathbf{v}^2$ and the diffusion matrix $\mathbb{D}$ reduces to $(\mathbf{v}^1\otimes\mathbf{v}^1)/2\mu_{12}$.

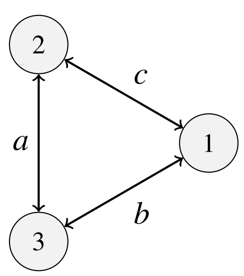

Figure 1: Schematic representation of the transitions in Examples 3.3.

Example 3.3. Let us consider the case with $n = 3$ given by the symmetric transition matrix

$$\mathbb{B} = \begin{pmatrix} b+c & -c & -b \\ -c & a+c & -a \\ -b & -a & a+b \end{pmatrix}.$$

so that, for $i = 1, 2, 3$,

$$\det \mathbb{B}(i) = ab + ac + bc, \qquad \det \mathbb{B}(1,2) = a + b, \qquad \det \mathbb{B}(1,3) = a + c, \qquad \det \mathbb{B}(2,3) = b + c.$$

Therefore, the drift velocity $\mathbf{v}_{\text{drift}}$, is the arithmetic average $\mathbf{v}^1, \mathbf{v}^2, \mathbf{v}^3$. Moreover, if $\mathbf{v}_{\text{drift}} = 0$, the diffusion matrix is

$$\begin{aligned}
\mathbb{D} &= -\frac{1}{3(ab+bc+ca)}\{(a+b)(\mathbf{v}^1 \otimes \mathbf{v}^2)^* + (a+c)(\mathbf{v}^1 \otimes \mathbf{v}^3)^* + (b+c)(\mathbf{v}^2 \otimes \mathbf{v}^3)^*\} \\
&= -\frac{1}{3(ab+bc+ca)}\{a[(\mathbf{v}^1 \otimes \mathbf{v}^2)^* + (\mathbf{v}^1 \otimes \mathbf{v}^3)^*] + b[(\mathbf{v}^1 \otimes \mathbf{v}^2)^* + (\mathbf{v}^2 \otimes \mathbf{v}^3)^*] + c[(\mathbf{v}^1 \otimes \mathbf{v}^3)^* + (\mathbf{v}^2 \otimes \mathbf{v}^3)^*]\} \\
&= \frac{1}{3(ab+bc+ca)}\left\{a\,\mathbf{v}^1 \otimes \mathbf{v}^1 + b\,\mathbf{v}^2 \otimes \mathbf{v}^2 + c\,\mathbf{v}^3 \otimes \mathbf{v}^3\right\}
\end{aligned}$$

Example 3.4. Next, consider four velocities with symmetric transition rates

$$\mathbb{B} = \begin{pmatrix} * & -\mu_{12} & 0 & -\mu_{14} \\ -\mu_{12} & * & -\mu_{23} & 0 \\ 0 & -\mu_{23} & * & -\mu_{34} \\ -\mu_{14} & 0 & -\mu_{34} & * \end{pmatrix},$$

where the diagonal entries are the sums of the column element changed by sign. Then, there hold for $i \in \{1, 2, 3, 4\}$

$$\det \mathbb{B}(i) = \mu_{12}\mu_{14}\mu_{23} + \mu_{12}\mu_{14}\mu_{34} + \mu_{12}\mu_{23}\mu_{34} + \mu_{14}\mu_{23}\mu_{34}$$

and, for the second order principal minors,

$$\begin{aligned}
&\det \mathbb{B}(1,2) = \mu_{14}\mu_{23} + \mu_{14}\mu_{34} + \mu_{23}\mu_{34} && \det \mathbb{B}(1,3) = \mu_{12}\mu_{14} + \mu_{12}\mu_{34} + \mu_{14}\mu_{23} + \mu_{23}\mu_{34} \\
&\det \mathbb{B}(1,4) = \mu_{12}\mu_{23} + \mu_{12}\mu_{34} + \mu_{23}\mu_{34} && \det \mathbb{B}(2,3) = \mu_{12}\mu_{14} + \mu_{12}\mu_{34} + \mu_{14}\mu_{34} \\
&\det \mathbb{B}(2,4) = \mu_{12}\mu_{23} + \mu_{12}\mu_{34} + \mu_{14}\mu_{23} + \mu_{14}\mu_{34} && \det \mathbb{B}(3,4) = \mu_{12}\mu_{14} + \mu_{12}\mu_{23} + \mu_{14}\mu_{23}.
\end{aligned}$$

Since $\det \mathbb{B}(i)$ is independent of i, as in the previous cases, the drift velocity is the average of the speeds. Then, assuming $\mathbf{v}_{\text{drift}} = 0$, the expression for the diffusion matrix can be rearranged by collecting the term multiplied by the same product of transition rates, obtaining

$$\begin{aligned}
\mathbb{D} = -\frac{1}{\det \mathbb{B}(i)}\Big\{&\mu_{23}\mu_{34}\big((\mathbf{v}^2 + \mathbf{v}^3 + \mathbf{v}^4) \otimes \mathbf{v}^1\big)^* + \mu_{14}\mu_{34}\big((\mathbf{v}^1 + \mathbf{v}^3 + \mathbf{v}^4) \otimes \mathbf{v}^2\big)^* \\
&+ \mu_{12}\mu_{14}\big((\mathbf{v}^1 + \mathbf{v}^2 + \mathbf{v}^4) \otimes \mathbf{v}^3\big)^* + \mu_{12}\mu_{23}\big((\mathbf{v}^1 + \mathbf{v}^2 + \mathbf{v}^3) \otimes \mathbf{v}^4\big)^* \\
&+ \mu_{12}\mu_{34}\big((\mathbf{v}^1 + \mathbf{v}^2) \otimes (\mathbf{v}^3 + \mathbf{v}^4)\big)^* + \mu_{14}\mu_{23}\big((\mathbf{v}^1 + \mathbf{v}^4) \otimes (\mathbf{v}^2 + \mathbf{v}^3)\big)^*\Big\}.
\end{aligned}$$

Since $\mathbf{v}^1 + \mathbf{v}^2 + \mathbf{v}^3 + \mathbf{v}^4 = 0$, the matrix $\mathbb{D}$ is given by

$$\mathbb{D} = \frac{1}{\det \mathbb{B}(i)} \Big\{ \mu_{23}\mu_{34}(\mathbf{v}^1 \otimes \mathbf{v}^1) + \mu_{14}\mu_{34}(\mathbf{v}^2 \otimes \mathbf{v}^2) + \mu_{12}\mu_{14}(\mathbf{v}^3 \otimes \mathbf{v}^3) + \mu_{12}\mu_{23}(\mathbf{v}^4 \otimes \mathbf{v}^4)$$
$$+ \mu_{12}\mu_{34}\big((\mathbf{v}^1 + \mathbf{v}^2) \otimes (\mathbf{v}^1 + \mathbf{v}^2)\big) \; + \mu_{14}\mu_{23}\big((\mathbf{v}^1 + \mathbf{v}^4) \otimes (\mathbf{v}^1 + \mathbf{v}^4)\big) \Big\}.$$

showing, in particular, that the diffusion matrix is non-negative definite.

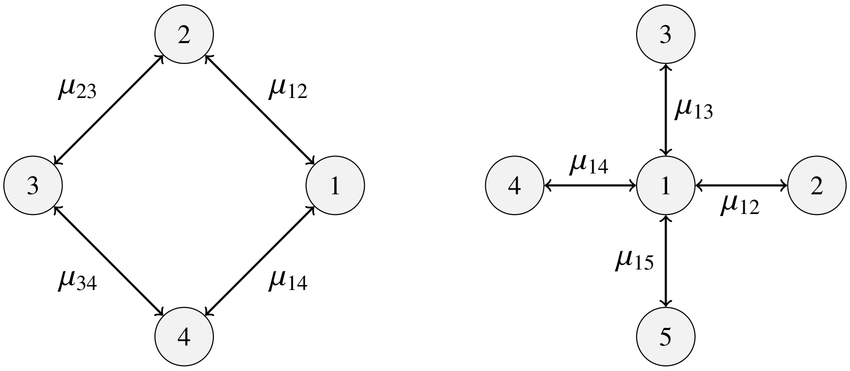

Figure 2: The graphs relative to the Examples 3.4–3.5.

Example 3.5. As a final example, let us consider a case with 5 velocities and admissible transitions only between $\mathbf{v}^1$ and $\mathbf{v}^j$ for $j = 2, 3, 4, 5$. The transition matrix is

$$\mathbb{B} = \begin{pmatrix} * & -\mu_{12} & -\mu_{13} & -\mu_{14} & -\mu_{15} \\ -\mu_{12} & * & 0 & 0 & 0 \\ -\mu_{13} & 0 & * & 0 & 0 \\ -\mu_{14} & 0 & 0 & * & 0 \\ -\mu_{15} & 0 & 0 & 0 & * \end{pmatrix}.$$

where the diagonal entries are the sums of the column element changed by sign. A direct computation shows that $\det \mathbb{B}(i) = \mu_{12}\mu_{13}\mu_{14}\mu_{15}$ for any $i \in \{1, 2, 3, 4, 5\}$. The second order principal minors are

$$\begin{aligned}
&\det \mathbb{B}(1,2) = \mu_{13}\mu_{14}\mu_{15} && \det \mathbb{B}(1,3) = \mu_{12}\mu_{14}\mu_{15} \\
&\det \mathbb{B}(1,4) = \mu_{12}\mu_{13}\mu_{15} && \det \mathbb{B}(1,5) = \mu_{12}\mu_{13}\mu_{14} \\
&\det \mathbb{B}(2,3) = \mu_{12}\mu_{14}\mu_{15} + \mu_{13}\mu_{14}\mu_{15} && \det \mathbb{B}(2,4) = \mu_{12}\mu_{13}\mu_{15} + \mu_{13}\mu_{14}\mu_{15} \\
&\det \mathbb{B}(2,5) = \mu_{12}\mu_{13}\mu_{14} + \mu_{13}\mu_{14}\mu_{15} && \det \mathbb{B}(3,4) = \mu_{12}\mu_{13}\mu_{15} + \mu_{12}\mu_{14}\mu_{15} \\
&\det \mathbb{B}(3,5) = \mu_{12}\mu_{13}\mu_{14} + \mu_{12}\mu_{14}\mu_{15} && \det \mathbb{B}(4,5) = \mu_{12}\mu_{13}\mu_{14} + \mu_{12}\mu_{13}\mu_{15}.
\end{aligned}$$

As in the previous cases, being $\det \mathbb{B}(i)$ is independent of i, the drift velocity is the average of the speeds. Then, setting $\mathbf{v}_{\text{drift}} = 0$, the diffusion matrix turns to be

$$\mathbb{D} = \frac{1}{\mu_{12}}(\mathbf{v}^2 \otimes \mathbf{v}^2) + \frac{1}{\mu_{13}}(\mathbf{v}^3 \otimes \mathbf{v}^3) + \frac{1}{\mu_{14}}(\mathbf{v}^4 \otimes \mathbf{v}^4) + \frac{1}{\mu_{15}}(\mathbf{v}^5 \otimes \mathbf{v}^5).$$

Note, in this case, the absence of the dependency of $\mathbb{D}$ on the velocity $\mathbf{v}^1$.

From all the examples, in the case of symmetric transition matrix $\mathbb{B}$, two specific features come evident: firstly, the drift speed is the average of the elements in the speed set $\{\mathbf{v}^1, \dots \mathbf{v}^n\}$; secondly, the diffusion matrix is described as a linear combination with non-negative coefficients of tensor of the form $w \otimes w$. This latter property, in particular, implies that the matrix $\mathbb{D}$ is non-negative definite.

In what follows, we show that the same hallmark is shared by any system (3) if the matrix $\mathbb{B}$ is symmetric as a consequence of the special structure of its principal minors.

Proposition 3.6. *If the transition matrix* $\mathbb{B}$ *is symmetric, then* $\det \mathbb{B}(i) = \det \mathbb{B}(j)$ *for any* $i, j \in \{1, \dots, n\}$.

Proof. Substituting the first row with the sum of all the rows, using symmetry and, then, the first column with the sum of all the columns

$$\det \mathbb{B}(1) = \det \begin{pmatrix} \sum\limits_{j\neq 2} \mu_{j2} & -\mu_{23} & \cdots \\ -\mu_{32} & \sum\limits_{j\neq 3} \mu_{j3} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \det \begin{pmatrix} \mu_{12} & \mu_{13} & \cdots \\ -\mu_{23} & \sum\limits_{j\neq 3} \mu_{j3} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \det \begin{pmatrix} \sum\limits_{j\neq 1} \mu_{1j} & \mu_{13} & \cdots \\ \mu_{13} & \sum\limits_{j\neq 3} \mu_{j3} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$= -\det \begin{pmatrix} -\sum\limits_{j\neq 1} \mu_{1j} & -\mu_{13} & \cdots \\ \mu_{13} & \sum\limits_{j\neq 3} \mu_{j3} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \det \begin{pmatrix} \sum\limits_{j\neq 1} \mu_{1j} & -\mu_{13} & \cdots \\ -\mu_{13} & \sum\limits_{j\neq 3} \mu_{j3} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \det \mathbb{B}(2).$$

The proof is complete. □

As a consequence of the independency of the principal minors of order 1 of the matrix $\mathbb{B}$, as suggested by the previous examples, the form of the drift velocity becomes particularly simple.

Theorem 3.7. *If the transition matrix* $\mathbb{B}$ *is symmetric, there holds*

$$\mathbf{v}_{drift} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{v}^i. \tag{12}$$

Proof. The proof is straightforward. Thanks to Proposition 3.6, $\det \mathbb{B}(i)$ is independent on i and thus $I_1(\mathbb{B}) = n \det \mathbb{B}(i)$. Then, from (10) follows

$$(\mathbf{v}_{\mathrm{drift}})_j = \frac{\det \mathbb{B}(i)}{n \det \mathbb{B}(i)} (\mathbf{1} \cdot (v_j^1, \dots, v_j^n)) = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{v}^i)_j$$

where $\mathbf{1} = (1, \dots, 1)$. □

In order to get a deeper knowledge on the structure on the diffusion matrix $\mathbb{D}$ relative to (3), it is needed a more precise understanding of the form of the principal minors of the transition matrix $\mathbb{B}$. In this direction, a fundamental tools is provided by a generalization, proved in [3], of the *Kirchoff's matrix tree Theorem*, a well-known result in graph theory.

In its original version, this result affirms that the number of spanning trees of a given graph coincides with the determinant of an appropriate matrix associated to the graph. Here, we use this fascinating connection the other way round: given the transition matrix $\mathbb{B}$, we consider a corresponding graph and determine the values of its minors by means of trees contained in the graph. Indeed, the velocity changes, dictated by the transition matrix $\mathbb{B}$, can be equivalently represented by means of a directed graph, shortly called a *digraph*, whose vertices are the speeds $\mathbf{v}^i$, and with arcs from the i-th to the j-th node weighted by the transition rate μ_{ij}. The symmetry assumption on $\mathbb{B}$ translates into the fact that the graph is actually undirected with weights given by the common values of the transition rates. In this respect, Example 3.3 correspond to the graph depicted in Figure 1 and Examples 3.4–3.5 to the graphs in Figure 2. Also, irreducibility of the matrix $\mathbb{B}$ is equivalent to the connectedness of the associated graph.

The collection of velocities $\mathbf{v}^i$ with arcs weighted by the rate μ_{ij} will be referred to as the *graph associated to* (3). Let us stress that the graph representation of the vertex is not directly related with the effective value of the velocity $\mathbf{v}^i$ as vector in $\mathbb{R}^d$. The content of the graph is only illustrative on the admissibile speed transitions.

To proceed, let us also recall that a *tree* is an undirected graph in which any two vertices are connected by exactly one simple path. Equivalently, a tree is a connected graph without simple cycles. A *forest* is a disjoint union of trees. Finally, given a forest F in the graph Γ, we denote by μ_F the product of the weights of all the arcs in F and we call it *weight of the forest*. By definition, if F is composed by a single vertex, its weight is equal to 1.

Next, we state a (facilitated) version of the *(All minors) matrix tree theorem* (see [3]).

Theorem 3.8. *Let $\Gamma = (\mathbf{v}^i, \mu_{ij})$ be the graph associated to* (3)*. Let the transition matrix $\mathbb{B}$ be symmetric and let $\mathcal{I} = \{i_1 < \cdots < i_k\}$ be a set of indeces in $\{1, \ldots, n\}$. Then*

$$\det \mathbb{B}(\mathcal{I}) = \sum_{F \in \mathcal{F}} \mu_F$$

where $\mathcal{F}$ is the set of forests F in Γ such that

i. *all the verteces in Γ are contained in F;*

ii. *F contains exactly k trees;*

iii. *each tree in F contains exactly one vertex $\mathbf{v}^i$ with $i \in \mathcal{I}$.*

Formulas (10) and (11) can be now re-interpreted taking advantage of Theorem 3.8 in the cases $k = 1$ and $k = 2$. Indeed, to compute the value $\det \mathbb{B}(i)$ (which is independent on i thanks to Proposition 3.6) it is sufficient to consider all the trees composed by all the verteces $\mathbf{v}^i$ and contained in the graph associated to (3) and to compute the sum of the weights of such trees.

Example 3.9. As an explicative example, let us consider the case of n velocities $\mathbf{v}^i$ with rate μ_{ij} symmetric and positive if and only if, for $i < j$, either $j = i + 1$ or $i = 1$ and $j = n$ (see Figure 3, left). The graph associated to such choice is a cycle and all of its trees containing all of the n vertices are obtained by removing a single arc. Thus

$$\det \mathbb{B}(i) = \sum_{j=1}^{n} \frac{\mu_{12}\mu_{23} \cdots \mu_{n-1,n}\mu_{n1}}{\mu_{j,j+1}}$$

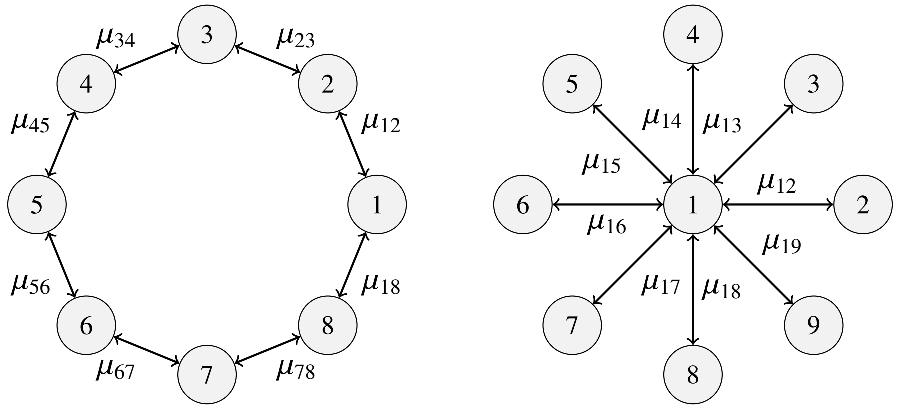

Figure 3: The graphs relative to the Example 3.9 in the case $n = 8$ and to the Example 3.10 in the case $n = 9$.

where $\mu_{n,n+1} = \mu_{n1}$.

Example 3.10. Somewhat oppositely with respect to Example 3.9, we can consider a matrix $\mathbb{B}$ such that transitions are possibile only from and towards a given specific speed, say $\mathbf{v}^1$, i.e. $\mu_{ij} > 0$ for $i < j$ if and only if $i = 1$ (see Figure 3, right). In this case, the graph is a tree. Hence, $\det \mathbb{B}(i)$ is the product of all the rates μ_{1j} with $j \in \{2, \ldots, n\}$.

Next, we take advantage of Theorem 3.8 to give a different representation of the diffusion matrix $\mathbb{D}$ defined in (11). Given the graph Γ associated to (3), we set

$$\mathcal{F}_2 := \{\text{forests } F \text{ partitioning verteces of } \Gamma \text{ in two trees}\}.$$

Any element F in $\mathcal{F}_2$ is composed by two trees, T^1 and T^2, ordered according to the size of the smallest node in each tree. Additionaly, given indeces i, j, with $i < j$, we use the notation

$$\mathcal{F}_2(i, j) := \{F = \{T^1, T^2\} \in \mathcal{F}_2 \,:\, i \text{ and } j \text{ are verteces of } T^1 \text{ and } T^2, \text{ respectively}\}.$$

Then, given indeces i and j, there holds

$$\det \mathbb{B}(i, j) = \sum_{\mathcal{F}_2(i,j)} \mu_{T^1} \mu_{T^2}.$$

and thus, for $\mathbf{v}_{\text{drift}} = 0$,

$$\mathbb{D} = -\frac{1}{I_1(\mathbb{B})} \sum_{h<k} \sum_{\mathcal{F}_2(h,k)} \mu_{T^1} \mu_{T^2} \, (\mathbf{v}^h \otimes \mathbf{v}^k)^*.$$

Reversing the order of the sums, we infer

$$\mathbb{D} = -\frac{1}{I_1(\mathbb{B})} \sum_{\mathcal{F}_2} \mu_{T^1} \mu_{T^2} \sum_{h \in T^1} \sum_{k \in T^2} (\mathbf{v}^h \otimes \mathbf{v}^k)^* = -\frac{1}{I_1(\mathbb{B})} \sum_{\mathcal{F}_2} \mu_{T^1} \mu_{T^2} \left(\sum_{h \in T^1} \mathbf{v}^h \otimes \sum_{k \in T^2} \mathbf{v}^k \right)^*$$

Thanks to Proposition 3.6, $\mathbf{v}_{\text{drift}} = 0$ if and only if the sum of the velocities $\mathbf{v}^i$ is null. As a consequence, we obtain the proof of the following statement, the more intriguing contribution of the present paper, giving an explicit formula for the diffusion matrix in the symmetric setting.

Theorem 3.11. *Let the transition matrix $\mathbb{B}$ be symmetric. Then, denoting by $\mathbf{w}(T)$ the sum of the velocities $\mathbf{v}^i$ in a given tree T of the graph associated to* (3)*, the diffusion matrix relative to* (3) *is*

$$\mathbb{D} = \frac{1}{I_1(\mathbb{B})} \sum_{\mathcal{F}_2} \mu_{_{T^1}} \mu_{_{T^2}} \left(\mathbf{w}(T_1) \otimes \mathbf{w}(T_1) \right). \tag{13}$$

In particular, the matrix $\mathbb{D}$ is non-negative definite.

When compared with the representation for the asymptotic parabolic equation provided in [2] (formula (5.80), Section 5.5), formula (13) exhibits a number of strengths, the more evident being its capability of pinpointing the different role played by the choice of the microscopic speeds and transition rates. Such advantages are consequence of the specific class of system considered in this paper and of its meaningful microscopic interpretation. No statement in the spirit of Theorem 3.11 can be expected for the very general type of equations considered in [2].

To get acquainted with formula (13), let us consider some of the previous Examples.

Example 3.12. For the Goldstein–Kac model, the graph Γ is composed by a single arc connecting the speeds $\mathbf{v}^1 = -\nu$ and $\mathbf{v}^2 = \nu$ with weight μ. Then, the set $\mathcal{F}_2$ has a single element with trees $T_1 = \{\mathbf{v}^1\}$ and $T_2 = \{\mathbf{v}^2\}$. Hence, the diffusion coefficient is $\mathbb{D} = \nu^2/2\mu$, since, by definition, $I_1(\mathbb{B}) = 2\mu$ and $\mu_{_{T^1}} = 1$.

Example 3.13. For what concerns the system of Example 3.4, the set $\mathcal{F}_2$ is composed by six elements, obtained by removing two arcs of the corresponding graph (see Figure 5). Each of the element $F = \{T^1, T^2\}$ in $\mathcal{F}_2$ gives the

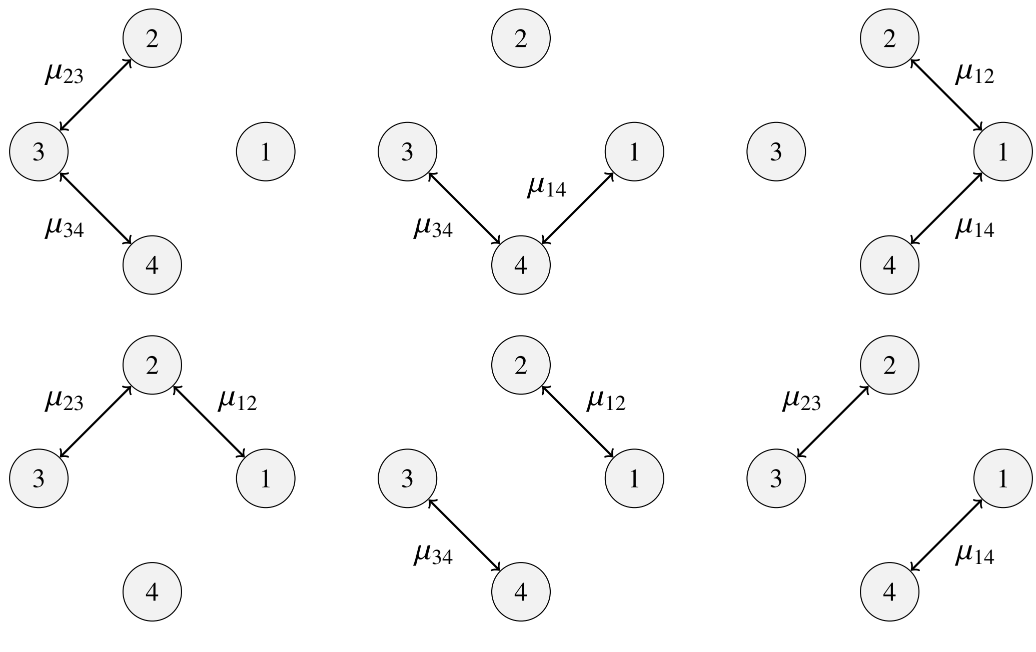

Figure 4: The six elements of $\mathcal{F}_2$ for the graph of the Example 3.4.

coefficient of an appropriate matrix of the form $\mathbf{w} \otimes \mathbf{w}$ in the formula for the diffusion matrix $\mathbb{B}$. The vector $\mathbf{w}$ is obtained by summing the verteces of the tree T^1, or of the tree T^2. These two sum of verteces coincides because of the requirement $\mathbf{v}_{\text{drift}} = 0$. The final expression of $\mathbb{D}$ has been given in Example 3.4.

Example 3.14. Example 3.9 generalizes the previous one. The family $\mathcal{F}_2$ is completely described by the trees T^1 containing a given vertex, say $\mathbf{v}^1$. Since such threes are composed by path of k consecutive arcs, for $k = 1, \dots, n-1$, there are exactly k of such threes containing the vertex. Therefore, the family $\mathcal{F}_2$ is composed by $n(n-1)/2$ elements. The diffusion matrix $\mathbb{D}$ is the sum of weighted matrix $\mathbf{w} \otimes \mathbf{w}$ where $\mathbf{w}$ is the sum of the verteces in the tree T^1 and the weights is the product of the arcs in T^1 and in T^2 (thus the product of all the coefficients μ_{ij} with the exception of the one corresponding to the two arcs not in T^1 and T^2).

Example 3.15. The situation for Example 3.10 is particularly simple. Indeed, the set $\mathcal{F}_2$ is composed by $n-1$ element each of which correspond to the forest of a tree composed by a single vertex $\mathbf{v}^i$, $i \in \{2, \dots, n\}$ and the tree, denoted by $T(i)$ composed by the remaining part of the graph excluding both the vertex $\mathbf{v}^i$ and the arc from $\mathbf{v}^1$ to $\mathbf{v}^i$. Thus, the diffusion matrix takes the form

$$\mathbb{D} = \frac{1}{\mu_{12} \dots \mu_{1n}} \sum_{i=2}^{n} \mu_{T(i)} \left(\mathbf{v}^i \otimes \mathbf{v}^i \right) = \sum_{i=2}^{n} \frac{1}{\mu_{1i}} \left(\mathbf{v}^i \otimes \mathbf{v}^i \right).$$

Apparently, the formula for $\mathbb{D}$ does not depend on the choice of $\mathbf{v}^1$, but it should be always kept in mind that formula (13) gives the diffusion matrix when the drift term $\mathbf{v}_{\text{drift}}$ has been set to zero by applying a change in the reference frame. Thus, the speed $\mathbf{v}^1$ cannot be chosen independently on the other speeds.

To conclude the Section let us consider a specific case for which the expression of the diffusion matrix simplifies further, involving only term of the form $\mathbf{v}^i \otimes \mathbf{v}^i$.

Proposition 3.16. *Let $\mathbb{B}$ be a symmetric matrix and assume:*

i. *either $n = 2k$ for some k and $\mathbf{v}^{2\ell} = -\mathbf{v}^{2\ell-1}$ for any $\ell \in \{1, \dots, k\}$;*

ii. *or $n = 2k+1$ for some k, $\mathbf{v}^{2\ell} = -\mathbf{v}^{2\ell-1}$ for any $\ell \in \{1, \dots, k\}$ and $\mathbf{v}^{2k+1} = 0$.*

If the transition matrix $\mathbb{B}$ is such that

$$\mathbb{B}(2i-1, j) = \mathbb{B}(2i, j) \tag{14}$$

for any $i = 1, \dots, k$ and $j = 1, \dots, n$, then there holds

$$\mathbb{D} = -\frac{1}{I_1(\mathbb{B})} \sum_{i=1}^{m} \sum_{\mathcal{F}_2(2i-1,2i)} \mu_{T^1} \mu_{T^2} (\mathbf{v}^{2i} \otimes \mathbf{v}^{2i}) \tag{15}$$

Proof. Let us deal with case **i.**, the other one being similar. The proof consists in showing that formula (11), under assumption (14), can be rearranged as

$$\mathbb{D} = -\frac{1}{I_1(\mathbb{B})} \sum_{i=1}^{k} \mathbb{B}(2i-1, 2i) \, (\mathbf{v}^{2i} \otimes \mathbf{v}^{2i}).$$

Let us proceed by induction. For $k = 1$, there holds

$$I_1(\mathbb{B}) \, \mathbb{D} = -\mathbb{B}(1,2) \, (\mathbf{v}^1 \otimes \mathbf{v}^2)^* = \mathbb{B}(1,2)(\mathbf{v}^2 \otimes \mathbf{v}^2)$$

Then, assuming that the thesis holds for $k-1$, we infer

$$\begin{aligned} I_1(\mathbb{B})\,\mathbb{D} &= -\sum_{j>1}\mathbb{B}(1,j)\,(\mathbf{v}^1\otimes\mathbf{v}^j)^* - \sum_{j>2}\mathbb{B}(2,j)\,(\mathbf{v}^2\otimes\mathbf{v}^j)^* - \sum_{2<i<j}\mathbb{B}(i,j)\,(\mathbf{v}^i\otimes\mathbf{v}^j)^* \\ &= -\mathbb{B}(1,2)\,(\mathbf{v}^1\otimes\mathbf{v}^2)^* + \sum_{j>2}(\mathbb{B}(1,j)-\mathbb{B}(2,j))(\mathbf{v}^2\otimes\mathbf{v}^j)^* - \sum_{2<i<j}\mathbb{B}(i,j)\,(\mathbf{v}^i\otimes\mathbf{v}^j)^* \\ &= \mathbb{B}(1,2)\,(\mathbf{v}^2\otimes\mathbf{v}^2)^* + \sum_{i=2}^{k}\mathbb{B}(2i-1,2i)\,(\mathbf{v}^{2i}\otimes\mathbf{v}^{2i}) = \sum_{i=1}^{k}\mathbb{B}(2i-1,2i)\,(\mathbf{v}^{2i}\otimes\mathbf{v}^{2i}), \end{aligned}$$

that gives the conclusion. □

In term of transitions, hypothesis (14) asserts that the probability to jump on/from velocity $\mathbf{v}^{2j}$ or $-\mathbf{v}^{2j}$ to a different given speed $\mathbf{v}^\ell$ is the same. Such assumption is meaningful either when each couple of speeds $\mathbf{v}^{2j-1}, \mathbf{v}^{2j}$ corresponds to a cartesian direction or also in the case of "undirected" motion in the sense that the two directions on the same line have same probability of success after transition. As an example, the latter situation is of interest in the modeling of undirected tissues as considered in [11] in the mathematical description of mesenchymal motion.

4. Asymptotically parabolic behavior

To corroborate the analysis of the previous Section, we now show that the representation of the drift speed $\mathbf{v}_{\text{drift}}$ and of the diffusion matrix $\mathbb{D}$ is indeed significant for the description of the large-time behavior of solutions to (3)–(4). As mentioned in the Introduction, the main result of this Section fits into a well-estabilished research strand, whose main target is to quantify the large-time parabolic behavior of solutions of a class of dissipative hyperbolic equations. Specifically, the main point of the analysis is to show that the distance of the hyperbolic solution and some solution to a corresponding parabolic problem, decayes to zero faster than the decay of each separate term, showing that the dissipation mechanism is asymptotically of the same type. Being the parabolic behavior preferred for regularity reasons, it is usually stated that the hyperbolic equation has an *asymptotically parabolic nature*.

A fundamental stimulus on the topic has been provided by [14], where the asymptotic equivalence between Euler equations in Lagrangian coordinates with friction and the porous media equation is analyzed. To quote some more recent contributions, without any intention of completeness, let us mention the approach in [6] based on the used of (parabolic) scaling variables for scalar damped wave equations with general nonlinear lower order terms, the direction that explores the form of the Green function in [26] (nonequilibrium gasdynamics) and in [2] (general class of hyperbolic relaxation systems) is explored in details, the L^∞-bound in [25] and the L^p-L^q estimates in [18, 21] (and descendants) relative to the prototypical case of the relation between heat and telegraph equation, the analysis L^2-asymptotic expansions of the solutions for the heat and the damped wave equation proposed in [24], which clearly shows how the diffusive behavior is the effect of the cancellation of leading order terms.

A rigorous statement relative to the connection between the solutions to (3) and the ones to the corresponding parabolic equation, given by the asymptotic drift velocity and diffusion matrix can be obtained by observing that the model considered fits into the class analyzed in [2] (thanks to the validity of Proposition 2.3) at the price of facing all the complicated machinery needed to deal with a very general class of dissipative nonlinear hyperbolic system. Hence, for reader convenience, this final Section intend to provide a self-contained statement and proof on the diffusive behavior of system (3), taylored on the specific structure of the model and, mainly, on the scalar nature of the corresponding asymptotic parabolic equation.

Theorem 4.1. *Assume the transition matrix $\mathbb{B}$ to be symmetric with strictly positive diagonal elements and the diffusion matrix $\mathbb{D}$ to be positive definite. Let* $\mathbf{f}$ *be the solution to* (3)–(4)*, u be the sum of the components f_i, i.e. $u := \mathbf{1}\cdot\mathbf{f}$, and u_{par} be the solution to*

$$\frac{\partial w}{\partial t} = \mathrm{div}\,(\mathbb{D}\nabla_{\mathbf{x}} w) \qquad w(x,0) = \mathbf{1}\cdot\mathbf{f}_0.$$

Then, if $\mathbf{f}_0 \in [L^1 \cap L^2(\mathbb{R}^d)]^n$, there holds

$$|u - u_{par}|_{L^2}(t) \le C\, t^{-\frac{1}{4}d-\frac{1}{2}} |\mathbf{f}_0|_{L^1\cap L^2} \tag{16}$$

for some $C > 0$ (independent of t and $\mathbf{f}_0$).

Hypotheses on the matrices $\mathbb{B}$ and $\mathbb{D}$ are satisfied when the conditions required in Proposition 2.3 hold, namely

$$\mathbb{B} \quad \text{irreducible} \qquad \text{and} \qquad \mathrm{span}\,\{\mathbf{v}^i - \mathbf{v}^j \;:\; i,j = 1,\dots,n\} = \mathbb{R}^d.$$

Indeed, if the matrix $\mathbb{B}$ is irreducible (or, equivalently, if the graph associated to the system (3) is connected), the diagonal elements are strictly positive. Moreover, since in [22] it is proved that the property in Proposition 2.3 is equivalent to the bound

$$\mathrm{Re}\lambda \le -\frac{c_0|\mathbf{k}|^2}{1+|\mathbf{k}|^2}$$

for some $c_0 > 0$ and for any $(\lambda,\mathbf{k})$ satisfying the dispersion relation, the matrix $\mathbb{D}$ is forced to be positive definite, whenever the differences $\mathbf{v}^i - \mathbf{v}^j$ generates all of $\mathbb{R}^d$.

To prove Theorem 4.1, we apply Fourier transform to (3). Denoting by $\hat{u}$ the Fourier transform of function u, we obtain a system of ordinary differential equations for the frequency variables $\hat{f} = \hat{f}(\mathbf{k},t)$

$$\frac{\partial \hat{f}_i}{\partial t} + (\mathbf{v}^i\cdot\mathbf{k})\hat{f}_i + \sum_{j\neq i}(\mu_{ij}\,\hat{f}_i - \mu_{ji}\,\hat{f}_j) = 0 \qquad\qquad i = 1,\dots,n, \tag{17}$$

with initial conditions

$$\hat{f}_i(\mathbf{k},0) = \hat{f}_{0,i}(\mathbf{k}) \qquad \mathbf{k}\in\mathbb{R}^d \qquad\qquad i = 1,\dots,n. \tag{18}$$

Denoting by $\hat{\mathbf{f}}$ and $\hat{\mathbf{f}}_0$ the vectors of components $\hat{f}_i$ and $\hat{f}_{0,i}$, respectively, the solution to (17)–(18) is given by

$$\hat{\mathbf{f}}(\mathbf{k},t) = \exp\left\{-(\mathrm{diag}\,(\mathbf{v}^i\cdot\mathbf{k}) + \mathbb{B})t\right\}\hat{\mathbf{f}}_0(\mathbf{k}).$$

Let us consider the functions

$$\hat{u}(\mathbf{k},t) = \mathbf{1}\cdot\exp\left\{-(\mathrm{diag}\,(\mathbf{v}^i\cdot\mathbf{k}) + \mathbb{B})t\right\}\hat{\mathbf{f}}_0(\mathbf{k}) \qquad \text{and} \qquad \hat{u}_{\mathrm{par}}(\mathbf{k},t) = e^{(\mathbf{k}\cdot\mathbb{D}\mathbf{k})t}\hat{u}_0(\mathbf{k})$$

where $\hat{u}_0 = \mathbf{1}\cdot\hat{\mathbf{f}}_0$. We are interested in estimating the L^2-norm of the difference $\hat{u} - \hat{u}_{\text{par}}$.

Let $\varepsilon_0 > 0$ to be chosen. Then, there holds

$$|\hat{u} - \hat{u}_{\text{par}}|^2_{L^2}(t) \leq I_1(t) + I_2(t) \tag{19}$$

where

$$I_1(t) := \int_{|\mathbf{k}|<\varepsilon_0} \Big|\mathbf{1}\cdot\exp\Big\{-(\text{diag}\,(\mathbf{v}^i\cdot\mathbf{k}) + \mathbb{B})t\Big\}\,\hat{\mathbf{f}}_0(\mathbf{k}) - e^{(\mathbf{k}\cdot\mathbb{D}\mathbf{k})t}(\mathbf{1}\cdot\hat{\mathbf{f}}_0)(\mathbf{k})\Big|^2\,d\mathbf{k}$$

$$I_2(t) := \int_{|\mathbf{k}|\geq\varepsilon_0} \Big\{|\hat{u}|^2 + |\hat{u}_{\text{par}}|^2\Big\}(\mathbf{k},t)\,d\mathbf{k}$$

The dispersion relation of (3) in the limit $\mathbf{k}\to 0$ has been explored in Section 3. In particular, in this regime, it is possible to identify the branch of solution of $p(\lambda,\mathbf{k}) = 0$ passing through the origin $(0,\mathbf{0})$ and to describe its behavior by means of a scalar-valued function $\lambda = \lambda(\mathbf{k})$ for which the following second-order expansion holds

$$\lambda(\mathbf{k}) = \mathbf{k}\cdot\mathbb{D}\mathbf{k} + o(|\mathbf{k}|^2) \qquad \text{as} \quad \mathbf{k}\to 0.$$

Taking advantage of this relation, we are able to state a result concerning the bound of I_1.

Lemma 4.2. *Let the vectors $\mathbf{v}^i$ be such that $\sum_i \mathbf{v}^i = 0$ and let the matrix $\mathbb{D}$ be positive definite. If $\mathbf{f}_0 \in L^1\cap L^2(\mathbb{R}^d)$, then there exist $\varepsilon_0, C, c > 0$ such that*

$$I_1(t) \leq C(t^{-\frac{1}{2}d-1}|\mathbf{f}_0|^2_{L^1} + e^{-ct}|\mathbf{f}_0|^2_{L^2}). \tag{20}$$

for any $t > 0$.

Proof. Denoting by $\mathbb{P}(\mathbf{k})$ the spectral projector relative to $\lambda = \lambda(\mathbf{k})$, we deduce that, for ε_0 small, there holds

$$\exp\Big\{-(\text{diag}\,(\mathbf{v}^i\cdot\mathbf{k}) + \mathbb{B})t\Big\} = e^{\lambda(\mathbf{k})t}\mathbb{P}(\mathbf{k}) + O(e^{-\theta t})$$

for some $\theta > 0$ uniform with respect to $\mathbf{k}$ such that $|\mathbf{k}| \leq \varepsilon_0$. Then, the term I_1 can be estimated by

$$I_1(t) \leq \int_{|\mathbf{k}|<\varepsilon_0} \Big|\Big(e^{\lambda(\mathbf{k})t}\mathbf{1}\,\mathbb{P}(\mathbf{k}) - e^{(\mathbf{k}\cdot\mathbb{D}\mathbf{k})t}\mathbf{1}\Big)\cdot\hat{\mathbf{f}}_0(\mathbf{k})\Big|^2\,d\mathbf{k} + O(e^{-2\theta t})\int_{|\mathbf{k}|<\varepsilon_0}\Big|\hat{\mathbf{f}}_0(\mathbf{k})\Big|^2\,d\mathbf{k}.$$

The projection $\mathbb{P}(\mathbf{k})$ has the form $r(\mathbf{k})\otimes\ell(\mathbf{k})$ where ℓ and r are, respectively, left and right eigenvectors of $\text{diag}\,(\mathbf{v}^i\cdot\mathbf{k}) + \mathbb{B}$ relative to the eigenvalue $\lambda(\mathbf{k})$, normalized so that the condition $\ell(\mathbf{k})\cdot r(\mathbf{k}) = 1$ holds. Since the sum of columns and rows of $\mathbb{B}$ is zero, by assumption, we have $r(\mathbf{0}) = \ell(\mathbf{0}) = \mathbf{1}/\sqrt{n}$. In particular, the zero-th order expansion for $\mathbb{P}(\mathbf{k})$ is

$$\mathbb{P}(\mathbf{k}) = \frac{1}{n}(\mathbf{1}\otimes\mathbf{1}) + O(\mathbf{k}) \qquad \text{as} \quad \mathbf{k}\to\mathbf{0}.$$

Therefore, the term I_1 is bounded by

$$I_1(t) \leq I_{11}(t) + I_{12}(t) + O(e^{-2\theta t})|\hat{\mathbf{f}}_0|^2_{L^2},$$

with

$$I_{11}(t) := o(1)\int_{|\mathbf{k}|<\varepsilon_0} e^{2(\mathbf{k}\cdot\mathbb{D}\mathbf{k})t}|\mathbf{k}|^4\Big|\hat{u}_0(\mathbf{k})\Big|^2\,d\mathbf{k}, \qquad \text{and} \qquad I_{12}(t) := O(1)\int_{|\mathbf{k}|<\varepsilon_0} e^{2\text{Re}\lambda(\mathbf{k})t}|\mathbf{k}|^2|\hat{\mathbf{f}}_0(\mathbf{k})|^2\,d\mathbf{k}.$$

For $v \in L^1(\mathbb{R}^d)$, there holds

$$\int_{|\mathbf{k}|<\varepsilon_0} e^{-2\theta|\mathbf{k}|^2 t}|\mathbf{k}|^{2\ell}|\hat{v}(\mathbf{k})|^2\,d\mathbf{k} \le |v|^2_{L^1}\int_{|\mathbf{k}|<\varepsilon_0} e^{-2\theta|\mathbf{k}|^2 t}|\mathbf{k}|^{2\ell}\,d\mathbf{k}$$
$$= |v|^2_{L^1}\, t^{-\frac{1}{2}d-\ell}\int_{|\mathbf{y}|<\varepsilon_0\sqrt{\theta t}} e^{-2|\mathbf{y}|^2}|\mathbf{y}|^{2\ell}\,d\mathbf{k} \le C|v|^2_{L^1}\, t^{-\frac{1}{2}d-\ell}.$$

Thus, collecting, we end up with

$$I_1(t) \le C(\,t^{-\frac{1}{2}d-2}|u_0|^2_{L^1} + C\,t^{-\frac{1}{2}d-1}|\mathbf{f}_0|^2_{L^1} + e^{-2\theta t}|\mathbf{f}_0|^2_{L^2}),$$

which gives (20). □

Concerning the term I_2 for value of $\mathbf{k}$ at positive distance from the origin $\mathbf{0}$, the following estimate holds true.

Lemma 4.3. *Let $\mathbf{v}^i$ be such that $\sum_i \mathbf{v}^i = 0$ and let the matrix $\mathbb{B}$ be such that $\mu_{ii} > 0$ for any $i = 1,\dots,n$. If $\mathbf{f}_0 \in L^2(\mathbb{R}^d)$, then for any $\varepsilon_0 > 0$ there exist $C, c > 0$ such that*

$$I_2(t) \le C\,e^{-ct}|\mathbf{f}_0|^2_{L^2} \qquad \forall\, t > 0. \tag{21}$$

Proof. Since $\mathbb{D}$ is positive definite, from the definition of $\hat{u}_{\text{par}}$ it follows

$$\int_{|\mathbf{k}|\ge\varepsilon_0} |\hat{u}_{\text{par}}|^2\,d\mathbf{k} \le \int_{|\mathbf{k}|\ge\varepsilon_0} e^{-2c_0|\mathbf{k}|^2 t}|\hat{u}_0(\mathbf{k})|^2\,d\mathbf{k} \le e^{-2c_0\varepsilon_0^2 t}|\hat{u}_0|^2_{L^2} = e^{-2c_0\varepsilon_0^2 t}|\mathbf{1}\cdot\mathbf{f}_0|^2_{L^2}.$$

for some $c_0 >$. Thus, estimate (21) is proved if we show that there exists $\theta > 0$ such that

$$\text{Re}\lambda \le -\theta < 0 \tag{22}$$

for any λ such that $\det(\lambda\,\mathbb{I} + \text{diag}\,(\mathbf{v}^i\cdot\mathbf{k}) + \mathbb{B}) = 0$ for some $|\mathbf{k}| > \varepsilon_0$.

The description of the dispersion relation in Section 3 guarantees that property $\text{Re}\lambda < 0$ holds for $\mathbf{k} \ne 0$ and small. Next, we show that the system does not support pure imaginary value of λ corresponding to purely imaginary values of $\mathbf{k}$. Let F be such that

$$(\lambda + \mathbf{v}^i\cdot\mathbf{k})F_i + \sum_{j=1}^n \mu_{ij}\,F_j = 0 \qquad i = 1,\dots,n. \tag{23}$$

By multiplying for the complex conjugate $\bar{F}_i$ and summing with respect to i, we get

$$\sum_{i=1}^n(\lambda + v_i\cdot\mathbf{k})|F_i|^2 + \sum_{i,j=1}^n \mu_{ij}\,\bar{F}_i\,F_j = 0$$

Since the matrix $\mathbb{B}$ is symmetric, the last term is real; hence, for $\lambda, \mathbf{k}$ purely imaginary,

$$\sum_{i=1}^n(\lambda + \mathbf{v}^i\cdot\mathbf{k})|F_i|^2 = \sum_{i,j=1}^n \mu_{ij}\,\bar{F}_i\,F_j = 0$$

Since $F \in \ker B$, then $F = C\mathbf{1}$ for some $C \in \mathbb{R}$ and thus

$$\lambda = -\frac{1}{n}\sum_{i=1}^n \mathbf{v}^i\cdot\mathbf{k} = 0$$

At the moment, the bound (22) holds for $|\mathbf{k}| \in [\varepsilon_0, M]$ for any $0 < \varepsilon_0 < M$ for some $\theta = \theta(\varepsilon_0, M)$. The last step of the proof concerns with the high frequency regime $\mathbf{k} \in i\mathbb{R}^d$ with $|\mathbf{k}| \to \infty$.

To this aim, let us consider the eigenvalue problem (23) with $\mathbf{k} = \varepsilon^{-1}\mathbf{h}$, $|\mathbf{h}| = 1$ and $\lambda = \varepsilon^{-1}\nu$, that is

$$(\mu + \mathbf{v}^i \cdot \mathbf{h})F_i + \varepsilon \sum_{j=1}^{n} \mu_{ij} F_j = 0 \qquad i = 1, \dots, n \tag{24}$$

in the limit $\varepsilon \to 0$. Setting $\nu = \nu_0 + \varepsilon \nu_1 + o(\varepsilon)$ and $F = F_0 + \varepsilon F_1 + o(\varepsilon)$, plugging into (24) and collecting the terms with same power of ε, we get the relations

$$(\nu_0 + \mathbf{v}^i \cdot \mathbf{h})F_{0,i} = 0 \qquad \text{and} \qquad (\nu_0 + \mathbf{v}^i \cdot \mathbf{h})F_{1,i} + \nu_1 F_{0,i} + \sum_{j=1}^{n} \mu_{ij} F_{0,j} = 0$$

Hence, we infer for the 0-th order coefficients $\nu_0 = -\mathbf{v}^i \cdot \mathbf{h}$ and $F_{0,i} = E_i$ where E_i denotes the i-th element of the canonical base of $\mathbb{R}^n$. Plugging into the 1-st order relation, we get the formula for the first coefficient in the expansion of ν, that is $\nu_1 = -\mu_{ii}$. Coming back to the original variables λ and $\mathbf{k}$, the asymptotic expansions reads as

$$\lambda(\mathbf{k}) = -\mathbf{v}^i \cdot \mathbf{k} - \mu_{ii} + o(1) \qquad \text{as} \quad |\mathbf{k}| \to \infty.$$

Since the diagonal values μ_{ii} are assumed to be strictly positive, the bound (22) can be prolunged to $M \to +\infty$, changing, if necessary, the value of the constant θ. □

By means of the results of Lemmas 4.2–4.3, the completion of the proof of Theorem 4.1 is at hand. Indeed, resuming from (19) and using (20)–(21), we obtain

$$|\hat{u} - \hat{u}_{\rm par}|^2_{L^2}(t) \leq C(t^{-\frac{1}{2}d-1}|\mathbf{f}_0|^2_{L^1} + e^{-ct}|\mathbf{f}_0|^2_{L^2}) \leq C\, t^{-\frac{1}{2}d-1}|\mathbf{f}_0|^2_{L^1\cap L^2}$$

that gives (16) passing to the square roots and invoking Plancherel identity.

Acknowledgements. The author is partially supported by the Italian Project FIRB 2012 "Dispersive dynamics: Fourier Analysis and Variational Methods".

References

[1] Berman A. and Plemmons R.J.; "Nonnegative matrices in the mathematical sciences", Revised reprint of the 1979 original. Classics in Applied Mathematics, 9. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

[2] Bianchini S., Hanouzet B. and Natalini, R.; Asymptotic behavior of smooth solutions for partially dissipative hyperbolic systems with a convex entropy. *Comm. Pure Appl. Math.* **60** (2007), no. 11, 1559–1622.

[3] Chaiken S.; A combinatorial proof of the all minors matrix tree theorem. *SIAM J. Algebraic Discrete Methods* **3** (1982), no. 3, 319–329.

[4] Friedman A. and Craciun G.; Approximate traveling waves in linear reaction-hyperbolic equations. *SIAM J. Math. Anal.* **38** (2006), no. 3, 741–758.

[5] Friedman A., Hu B. and Keener J.P.; The diffusion approximation for linear nonautonomous reaction-hyperbolic equations. *SIAM J. Math. Anal.* **45** (2013), no. 4, 2285–2298.

[6] Gallay Th. and Raugel G.; Scaling variables and asymptotic expansions in damped wave equations. *J. Differential Equations* **150** (1998), no. 1, 42–97.

[7] Goldstein, S.; On diffusion by discontinuous movements, and on the telegraph equation. *Quart. J. Mech. Appl. Math.* **4** (1951) 129–156.

[8] Guarguaglini F.R., Mascia C., Natalini R. and Ribot M.; Stability of constant states of qualitative behavior of solutions to a one dimensional hyperbolic model of chemotaxis. *Discrete Contin. Dyn. Syst. Ser. B* **12** (2009), no. 1, 39–76.

[9] Hadamard, J.; "Lectures on Cauchy's problem in linear partial differential equations", Dover Publications, New York, 2003. Originally published by Yale University Press, Mrs. Hepsa Ely Silliman memorial Lectures, New York, 1923.

[10] Hadeler K. P.; Travelling fronts for correlated random walks. *Canad. Appl. Math. Quart.* **2** (1994), no. 1, 27–43.

[11] Hillen T.; M^5 mesoscopic and macroscopic models for mesenchymal motion. *J. Math. Biol.* **53** (2006), no. 4, 585–616.

[12] Hillen T. and Othmer H.G.; The diffusion limit of transport equations derived from velocity-jump processes. *SIAM J. Appl. Math.* **61** (2000), no. 3, 751–775.

[13] Hillen T. and Stevens A.; Hyperbolic models for chemotaxis in 1-D. *Nonlinear Anal. Real World Appl.* **1** (2000), no. 3, 409–433.

[14] Hsiao L. and Liu T.-P.; Convergence to nonlinear diffusion waves for solutions of a system of hyperbolic conservation laws with damping. *Comm. Math. Phys.* **143** (1992), no. 3, 599–605.

[15] Kac M.; *A stochastic model related to the telegrapher's equation*. Reprinting of an article published in 1956. Papers arising from a Conference on Stochastic Differential Equations (Univ. Alberta, Edmonton, Alta., 1972). Rocky Mountain J. Math. 4 (1974), 497–509.

[16] Keller J.B.; Diffusion at finite speed and random walks. *Proc. Natl. Acad. Sci. USA* **101** (2004), no. 5, 1120–1122.

[17] Lattanzio C., Mascia C., Plaza R.G. and Simeoni C.; Analytical and numerical investigation of traveling waves for Allen–Cahn model with relaxation. *in preparation*.

[18] Marcati P. and Nishihara K.; The $L^p - L^q$ estimates of solutions to one-dimensional damped wave equations and their application to the compressible flow through porous media. *J. Differential Equations* **191** (2003), no. 2, 445–469.

[19] Mascia C. and Natalini R.; On relaxation hyperbolic systems violating the Shizuta–Kawashima condition. *Arch. Ration. Mech. Anal.* **195** (2010), no.3, 729–762.

[20] Natalini R. and Hanouzet B.; Weakly coupled systems of quasilinear hyperbolic equations. *Differential Integral Equations* **9** (1996), no. 6, 1279–1292.

[21] Nishihara K.; $L^p - L^q$ estimates of solutions to the damped wave equation in 3-dimensional space and their application. *Math. Z.* **244** (2003), no. 3, 631–649.

[22] Shizuta Y. and Kawashima S.; Systems of equations of hyperbolic-parabolic type with applications to the discrete Boltzmann equation. *Hokkaido Math. J.* **14** (1985), no. 2, 249–275.

[23] Taylor G. I.; Diffusion by continuous movements. *Proc. London Math. Soc.* **s2-20** (1922), no. 1, 196–212.

[24] Volkmer H.; Asymptotic expansion of L^2-norms of solutions to the heat and dissipative wave equations. *Asymptot. Anal.* **67** (2010), no. 1-2, 85–100.

[25] Yang H. and Milani A.; On the diffusion phenomenon of quasilinear hyperbolic waves. *Bull. Sci. Math.* **124** (2000), no. 5, 415–433.

[26] Zeng Y.; Gas dynamics in thermal nonequilibrium and general hyperbolic systems with relaxation. *Arch. Ration. Mech. Anal.* **150** (1999), no. 3, 225–279.