mathematics

MDPI

Article

Kinetic Theory and Markov Chains with Stochastically Varying Transition Matrices

Bruno Carbonaro [1] and Marco Menale [2,*]

[1] Dipartimento di Matematica e Fisica, Università degli Studi della Campania "L. Vanvitelli", Viale Lincoln 5, 81100 Caserta, Italy; bruno.carbonaro@unicampania.it
[2] Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Università degli Studi di Napoli "Federico II", Via Cintia, Monte S. Angelo, 80126 Naples, Italy
* Correspondence: marco.menale@unina.it or marco.menale@outlook.com

Abstract: As is well-known, the Kinetic Theory for Active Particles is a scheme of mathematical models based on a generalization of the Boltzmann equation. It must be nowadays acknowledged as one of the most versatile and effective tools to describe in mathematical terms the behavior of any system consisting of a large number of mutually interacting objects, no matter whether they also interact with the external world. In both cases, the description is stochastic, i.e., it aims to provide at each instant the probability distribution (or density) function on the set of possible states of the particles of the system. In other words, it describes the evolution of the system as a stochastic process. In a previous paper, we pointed out that such a process can be described in turn in terms of a special kind of vector time-continuous Markov Chain. These stochastic processes share important properties with many natural processes. The present paper aims to develop the discussion presented in that paper, in particular by considering and analyzing the case in which the transition matrices of the chain are neither constant (stationary Markov Chains) nor assigned functions of time (nonstationary Markov Chains). It is shown that this case expresses interactions of the system with the external world, with particular reference to random external events.

Keywords: stochastic processes; Markov chains; probability; kinetic theory; mathematical models; complex systems

MSC: 60J10; 60G07; 82C40; 70-10

check for updates

Academic Editors: Yong Xie and Michael Voskoglou

Received: 6 October 2024
Revised: 31 October 2024
Accepted: 2 March 2025
Published: 17 June 2025

Citation: Carbonaro, B.; Menale, M. Kinetic Theory and Markov Chains with Stochastically Varying Transition Matrices. *Mathematics* **2025**, *13*, 1994. https://doi.org/10.3390/math13121994

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The evolution of many-particle systems (where the word "particle" must be read in the widest sense, as any individual belonging to a gas, or to a herd of predators or prey, or to any human collectivity, like a nation, social class, or condominium), in view of its wide range of interpretations in different scientific fields, is one of the most interesting and widely studied topics of that branch of applied mathematics dealing with models of natural, political, social, and economic phenomena. The mathematical formalization of such an evolution has been tackled from many different viewpoints and with many different languages and techniques, e.g., that based on Bayesian Networks, see [1–3], or that based on the use of Fokker–Planck equation [4–6]. Among these, kinetic-theoretic models (see Section 6 for more details) seem to be the most versatile, expressive, and effective tools to describe many-particle systems and their evolution by means of suitable Equations (we can quote [7–18] as just as a small sample of the research addressed to this field). Since the so-called transition matrices (see Section 2) characterizing Markov Chains play

a fundamental rôle in kinetic-theoretic models, as they describe in stochastic terms the results of mutual interactions between the particles of any system. It should be clearly understood that, as pointed out in formal terms in [19,20], the equations of the Kinetic Theory for Active Particles are firmly based on the use of Markov Chains; however, rather surprisingly, in the literature about the kinetic-theoretic models, Markov Chains, at least to the best of our knowledge, have never been explicitly related to the evolution equations before [19,20]. In [20], we tried to show that the explicit reference to Markov Chains actually represents a good perspective to describe the behavior of a many-particle system when it interacts with the external world (in this connection, see [21] where suitable «forcing» terms are introduced in the equations). In [20], we tried to show that, when we consider the evolution of a many-particle system as an n-tuple of joined stochastic processes, one of which is a vector Markov Chain [19], then the interactions with the external world can be described by assuming that the interaction rates and the transition matrices undergo stochastic variations at each step [19].

The main aim of [20] was, however, the interpretation rather than the formal and technical analysis of the evolution equation in the presence of random varying interaction rates and transition matrices, As a consequence, the accurate description of the equations obtained in such a case is simply outlined in the final sections. The aim of the present paper is to develop that discussion and analyze the technical differences it introduces in the description of the evolution of the system. We also try to give some interpretations of the meaning of such differences.

The rest of the paper is organized as follows. Section 6 is devoted to recalling some basic features of kinetic-theoretic models. Definitions and notions about Markov Chains as stochastic processes, with particular regard to discrete time-discrete Markov Chains are given in Section 2, while in Section 3 we introduce (two-dimensional) vector Markov Chains and describe some of their basic properties in the discrete time-discrete case. Section 4 is devoted to introducing what we have called "continuous semi-Markov coupled random processes" with particular concern with time-continuous processes, and to conclude that the equations of the Kinetic Theory, in their most general form, can be seen as a special form of the basic vector equation connecting the absolute probability distributions at different steps of a Markov Chain with each other and with the transition probabilities. Section 5 is devoted to introducing the Markov Chains with stochastically varying transition matrices in the time-discrete case, in order to give a heuristic basis to their properties, while in Section 7 we show how this kind of Markov Chain modifies the evolution equations for a many-particle system interacting with the external world. Finally, in Section 8, we discuss some problems left to be tackled in the modified scheme and outline some possible research perspectives.

2. Some Background on Markov Chains

We now turn to giving the basic definitions of Markov Chains and a short account of their basic properties.

To start with, we recall that a random process [22] is a sequence $\{X_\eta\}_{\eta\in\mathbb{H}}$ of random variables (readers not acquainted with the notion of "random variable" can usefully consult [23], p. 37) with a common range S of possible values (the *states*), called the *state space* of the process. Both the set $\mathbb{H}$ of indexes and the state space S can be either a discrete or a continuous set. If $\mathbb{H}$ is discrete ($\mathbb{H} = \mathbb{N} = \{0, 1, 2, \ldots, n, \ldots\}$), then the process is said to be *time-discrete*; if $\mathbb{H}$ is continuous ($\mathbb{H} = [0, +\infty)$), then the process is said to be *time-continuous*. Analogously, if S is discrete ($S = \{z_h\}_{h\in\mathbb{I}} \subseteq \mathbb{R}$, with $\mathbb{I} \subseteq \mathbb{Z}$), then the process is said to be *discrete*, while, if S is continuous (i.e., any real interval), then the process is said to be *continuous*. From now to the end of this section, we shall first consider

discrete (either time-discrete or time-continuous) processes. In both cases, for the sake of simplicity and without loss of generality (as regards Markov Chains,, this simply amounts to avoiding matrices with infinitely many rows and columns), the state space S will be assumed to be finite.

Now, consider a discrete and time-discrete random process $\{X_h\}_{h\in\mathbb{N}}$ with $S = \{z_1, z_2, \ldots, z_n\}$. As usual, we denote by a vector $\mathbf{p}_h$ (the *state vector of the process at time* h) the probability distribution on S according to the random variable X_h. In other words, $\mathbf{p}_h \equiv (p_{h,1}, p_{h,2}, \ldots, p_{h,n})$ with $p_{h,k} = \mathbf{P}(X_h = z_k)$. A discrete time-discrete Markov Chain [22,24,25] $\{X_h\}_{h\in\mathbb{N}}$, with $R(X_h) = S \subset \mathbb{Z}$, is a random process such that the Markov condition

$$\begin{aligned}\mathbf{P}(X_h = z_{i_h} \mid X_1 = z_{i_1}, X_2 = z_{i_2}, \ldots, X_{h-1} = z_{i_{h-1}}) = \\ = \mathbf{P}(X_h = z_{i_h} \mid X_{h-1} = z_{i_{h-1}}), \qquad \forall\, h \in \mathbb{N}\backslash\{0\}\end{aligned} \tag{1}$$

holds (where, as customary, we have used the symbol $\mathbf{P}(A \mid B)$ to denote the conditional probability of an event A under the assumption that an event B has taken place [23]). This means that, for any $h \in \mathbb{N}$, the values of the h-th random variable of the process depend only on the values of the $(h-1)$-th, not on the values of any previous variable. Now, setting $P_{ij}(h) = \mathbf{P}(X_h = z_j \mid X_{h-1} = z_i)$ (with $(i,j) \in \{1, 2, \ldots, n\}^2$), $P_{ij}(h)$ is the *transition probability* from state z_i to state z_j at the h-th step, and the matrix

$$\mathbf{P}_h \equiv (P_{ij}(h))_{1\le i,j\le n} \equiv \begin{pmatrix} P_{11}(h) & P_{12}(h) & \ldots & P_{1n}(h) \\ P_{21}(h) & P_{22}(h) & \ldots & P_{2n}(h) \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1}(h) & P_{n2}(h) & \ldots & P_{nn}(h) \end{pmatrix}, \qquad \forall\, h \in \mathbb{N}\backslash\{0\}. \tag{2}$$

is called the *transition matrix* at h-th step. It is obvious and well known that

$$\sum_{j=1}^{n} P_{ij}(h) = 1, \qquad \forall\, h \in \mathbb{N}\backslash\{0\} \tag{3}$$

and that, in virtue of the law of alternatives [23],

$$\mathbf{p}_h = \mathbf{p}_{h-1}\mathbf{P}_h, \qquad \forall\, h \in \mathbb{N}\backslash\{0\}. \tag{4}$$

For the sake of completeness, we also recall that, for any couple (r, s) of non-negative integers,

$$\mathbf{p}_{r+s} = \mathbf{p}_r\mathbf{P}_{r+1}\mathbf{P}_{r+2}\ldots\mathbf{P}_{r+s}, \tag{5}$$

and in particular, when the Markov Chain is stationary, i.e., a transition matrix $\mathbf{P}$ exists such that $\mathbf{P}_h = \mathbf{P}$ for any $h \in \mathbb{N}\backslash\{0\}$,

$$\mathbf{p}_{r+s} = \mathbf{p}_r\mathbf{P}^s, \tag{6}$$

where the power at right-hand side must be interpreted in the sense of the row-by-column product of matrices (for further details, see, e.g., [22,24–26]).

3. Joint and Marginal Transition Probabilities

As already carried out in [20], and almost tracing the steps of the treatment presented there, our next step will be to now consider a special kind of Markov Chain, never explicitly considered before in the literature about random processes (as far as we are aware, a state space endowed with a structure is considered only in [27]), namely, vector Markov Chains

(see [19]). Like in [20], for the sake of simplicity, we examine the case in which $S = \Omega^2$ (where $\Omega = \{x_1, x_2, \ldots, x_m\}$, so that $|\Omega^2| = m^2$; as usual in the literature about discrete Markov Chains, the states in S are assumed to be arbitrarily ordered, so that we can set $S = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_{m^2}\}$, and the order chosen is quite immaterial. Since $\mathbf{z}_i$ is a couple of elements of Ω, we write $\mathbf{z}_i \equiv (x_{i_1}, x_{i_2})$ for any $i \in \{1, 2, \ldots, m^2\}$, with $(i_1, i_2) \in \{1, 2, \ldots, m\}^2$, and the process we want to consider is a sequence $\{(X_{h,1}, X_{h,2})\}_{h\in\mathbb{N}}$, which will be written in vector notation as $\{\mathbf{X}_h\}_{h\in\mathbb{N}}$. So, the transition probabilities are

$$\begin{aligned}\mathbf{P}(\mathbf{X}_h = \mathbf{z}_j \,|\, \mathbf{X}_{h-1} = \mathbf{z}_i) = \\ = \mathbf{P}(X_{h,1} = x_{j_1}, X_{h,2} = x_{j_2} \,|X_{h-1,1} = x_{i_1}, X_{h-1,2} = x_{i_2})\end{aligned} \tag{7}$$

(where we have set $\mathbf{z}_j \equiv (x_{j_1}, x_{j_2})$ and $\mathbf{z}_i \equiv (x_{i_1}, x_{i_2})$). Accordingly, the transition probabilities are labeled by 4 indexes, and we can set

$$\mathbf{P}(X_{h,1} = x_{j_1}, X_{h,2} = x_{j_2} \,|X_{h-1,1} = x_{i_1}, X_{h-1,2} = x_{i_2}) \equiv P_{i_1,i_2;j_1,j_2}(h)\,. \tag{8}$$

For any $h \in \mathbb{N}$, the four-dimensional transition matrix $\mathbf{P}_h \equiv (P_{i_1,i_2;j_1,j_2}(h))$ will be called the *joint transition matrix*. Each of its elements expresses the probability that the chain passes from a state (x_{i_1}, x_{i_2}) to a state (x_{j_1}, x_{j_2}) at h-th time. These are the *joint transition probabilities*. Obviously, together with these joint probabilities, we can consider also a number (in this case, $2m$) of different *marginal probabilities*. More precisely, for any $h \in \mathbb{N}$, we can consider the probabilities $\mathbf{P}(X_{h,1} = x_{j_1} \,|\, X_{h-1,1} = x_{i_1}, X_{h-1,2} = x_{i_2}) \equiv P_{i_1,i_2;j_1}(h)$ and $\mathbf{P}(X_{h,2} = x_{j_2} \,|\, X_{h-1,1} = x_{i_1}, X_{h-1,2} = x_{i_2}) \equiv P_{i_1,i_2;j_2}(h)$. The former is the probability that, if $\mathbf{X}_{h-1} = \mathbf{z}_i$, then $\mathbf{X}_h$ takes one of the values (x_{j_1}, y) with $y \in \Omega$; the latter is the probability that, if $\mathbf{X}_{h-1} = \mathbf{z}_i$, then $\mathbf{X}_h$ takes one of the values (y, x_{j_2}) with $y \in \Omega$. Thus, we have two three-dimensional matrices, each with m^3 entries. We shall denote by $P_{i_1,i_2;j_k}(h)$ $(k = 1, 2)$ the entries of each matrix. By definition, we have

$$\begin{aligned}P_{i_1,i_2;j_1}(h) = \sum_{j_2=1}^{m} P_{i_1,i_2;j_1,j_2}(h)\,, \\ P_{i_1,i_2;j_2}(h) = \sum_{j_1=1}^{m} P_{i_1,i_2;j_1,j_2}(h)\,,\end{aligned} \tag{9}$$

Recall now that $S = \Omega^2 \equiv \{(x_i, y_j)\}_{1\le i,j\le m}$, so we can agree to set $\mathbf{z}_{ij} \equiv (x_i, y_j) \in S$ and write the first of relations (9) in the form

$$P_{\mathbf{z}_{kl},x_r}(h) = \sum_{s=1}^{m} P_{\mathbf{z}_{kl},\mathbf{z}_{rs}}(h) = \sum_{s=1}^{m} P_{\mathbf{z}_{kl},(x_r,y_s)}(h)\,, \tag{10}$$

(where we have replaced the indexes i_1 and i_2 by h and k, respectively, and the indexes j_1 and j_2 by r and s, respectively). Next, we introduce the joint state vector $\mathbf{p}(h) \equiv (p_{ij}(h))$, where $p_{ij}(h) \equiv p_h(\mathbf{z}_{ij}) \equiv p_h(x_i, y_j)$ is the probability of the state $\mathbf{z}_{ij}$ at time h, and the marginal state vector $\mathbf{p}_x(h) \equiv (p_{x,h}(x_i))_{1\le i\le m}$ is the probability of the set of states $\{(x_i, y_j)\}_{y_j\in\Omega}$, so that

$$p_{x,h}(x_i) = \sum_{j=1}^{m} p_h(\mathbf{z}_{ij}) = \sum_{j=1}^{m} p_h(x_i, y_j) \tag{11}$$

and, in virtue of the law of alternatives

$$p_h(\mathbf{z}_{ij}) = \sum_{k=1}^{m}\sum_{l=1}^{m} p_{h-1}(\mathbf{z}_{kl}) P_{\mathbf{z}_{kl},\mathbf{z}_{ij}}(h)\,.$$

Hence, replacing this last relation in relation (11), and taking into account relation (10), we obtain

$$\begin{aligned} p_{x,h}(x_i) &= \sum_{j=1}^{m}\sum_{k=1}^{m}\sum_{l=1}^{m} p_{h-1}(x_k,y_l)P_{\mathbf{z}_{kl},(x_i,y_j)}(h) = \\ &= \sum_{k=1}^{m}\sum_{l=1}^{m} p_{h-1}(x_k,y_l)\left(\sum_{j=1}^{m} P_{\mathbf{z}_{kl},(x_i,y_j)}(h)\right) = \\ &= \sum_{k=1}^{m}\sum_{l=1}^{m} p_{h-1}(x_k,y_l)P_{\mathbf{z}_{kl},x_i}(h)\,. \end{aligned}$$

This last equation, by subtracting the probability $p_{x,h-1}(x_i)$ at both sides, yields

$$p_{x,h}(x_i) - p_{x,h-1}(x_i) = \sum_{k=1}^{m}\sum_{l=1}^{m} p_{h-1}(\mathbf{z}_{kl})P_{\mathbf{z}_{kl},x_i}(h) - p_{x,h-1}(x_i)\,. \tag{12}$$

Moreover, for any $l \in \{1,2,\ldots,q\}$,

$$\sum_{i'=1}^{m}\sum_{l'=1}^{m} P_{(x_i,y_l),(x_{i'},y_{l'})}(h) = 1\,,$$

so that

$$\begin{aligned} p_{x,h-1}(x_i) &= p_{x,h-1}(x_i)\sum_{i'=1}^{m}\sum_{l'=1}^{m} P_{(x_i,y_l),(x_{i'},y_{l'})}(h) = \\ &= \sum_{l=1}^{m} p_{h-1}(\mathbf{z}_{il})\sum_{i'=1}^{m}\sum_{l'=1}^{m} P_{(x_i,y_l),(x_{i'},y_{l'})}(h) = \\ &= \sum_{l=1}^{m} p_{h-1}(\mathbf{z}_{il})\sum_{i'=1}^{m} P_{(x_i,y_l),x_{i'}}(h) \end{aligned}$$

and we can rewrite Equation (12) in the final form

$$\begin{aligned} p_{x,h}(x_i) - p_{x,h-1}(x_i) = \sum_{l=1}^{m}\Bigg\{ &\sum_{k\neq i}^{1\ldots m} p_{h-1}(\mathbf{z}_{kl})P_{\mathbf{z}_{kl},x_i}(h) + \\ &- p_{h-1}(\mathbf{z}_{il})\sum_{i'\neq i}^{1\ldots m} P_{(x_i,y_l),x_{i'}}(h)\Bigg\}\,. \end{aligned} \tag{13}$$

Finally, if the random variables $X_{h,1}$ and $X_{h,2}$ are independent for any $h \in \mathbb{N}$, and the chain is stationary, then we find

$$\begin{aligned} p_{x,h}(x_i) - p_{x,h-1}(x_i) = \sum_{l=1}^{q}\Bigg\{ &\sum_{k\neq i}^{1\ldots m} p_{x,h-1}(x_k)p_{y,h-1}(y_l)P_{\mathbf{z}_{kl},x_i}(h) + \\ &- p_{x,h-1}(x_i)\sum_{k\neq i}^{1\ldots m} p_{y,h-1}(y_l)P_{\mathbf{z}_{il},x_k}(h)\Bigg\}\,. \end{aligned} \tag{14}$$

The first term at right-hand side will be called the *gain term* of the subset of states $S_i = \{\mathbf{z}_{il}\}_{1\leq l\leq m}$, since it accounts for all the possible transitions from other states to a state of S_i, while the last term is called the *loss term* of S_i, since it accounts for all the possible transitions from a state of S_i to any state *outside* S_i.

Therefore, we are on the way to show that the statistical equations governing the evolution of many-particle systems (see Section 6) are simply equations describing the

evolution of suitably defined Markov Chains. In the next Section this conclusion will be completely achieved.

4. Continuous Semi-Markov Coupled Random Processes

We now introduce what we shall call a *continuous semi-Markov coupled random process*, that is a random process $\{(\mathbf{X}_t, \boldsymbol{\chi}_t)\}$, where $t \in [0, T] \subset \mathbb{R}$ is a continuous parameter interpreted as the time variable and $T > 0$ can be either finite or infinite. Here, $\{\mathbf{X}_t\}$ is a two-dimensional Markov Chain like the one described in Section 3, and $\{\boldsymbol{\chi}_t\}$ is a m^2-dimensional vector Bernoullian process. More precisely,

1. $\boldsymbol{\chi}_t \equiv (\chi_{ij,t})$, with $(i, j) \in \{1, 2, \ldots, m\}^2$;
2. For any $(i, j) \in \{1, 2, \ldots, m\}^2$ and any $t \in [0, T]$, $\chi_{ij,t}$ is the classical Bernoulli variable, with range $R(\chi_{ij,t}) = \{0, 1\}$;
3. For any pair of triples (i_1, j_1, t_1) and $(i_2, j_2, t_2) \neq (i_1, j_1, t_1)$ in $\{1, 2, \ldots, m\}^2 \times [0, T]$, the random variables $\chi_{i_1 j_1, t_1}$ and χ_{i_2, j_2, t_2} are independent.

The meaning of the last condition requires at least a short discussion. When interpreted in the framework of the description of the behavior of many particle systems, it means that the probability that two particles interact at a given time t_1, when they are in two given states x_{i_1} and y_{j_1}, is not influenced by the assumption that they interacted at a different time t_1 nor on the assumption that an interaction between them has occurred when they were in a different couple of states. Of course, at first glance such a condition seems to introduce an undue restriction, but it is classically acknowledged to be quite plausible, since the probability of interaction is linked to *states*, not to *particles*, and the condition simply expresses the almost obvious fact that the way in which the states x_{i_1} and y_{j_1} have been achieved should not be taken into account.

As already observed in [20], the state vectors obviously depend on time, but the transition probabilities and the probabilities associated with the Bernoulli variables $\chi_{ij,t}$ may or may not depend on time. Obviously, on a strictly empirical ground, a positive probability distribution $u_{ij}(t)$ on $[0, T]$ such that $\mathbf{P}(\chi_{ij,t} = 1) = u_{ij}(t)$ for any t cannot be assigned; all what can be done is to assign a continuous "probability density" $\tau_{ij}(t)$ such that, for any $s \in [0, T]$ and any sufficiently small Δt, $\tau_{ij}(s)\Delta t$ is (approximately) the probability to find in any interval $[s, s + \Delta t]$ some points such that $\chi_{ij,t} = 1$. Roughly speaking, and referring to the interpretation of probability as relative frequency, we can state that $\tau_{ij}(s)\Delta t$ is the length of the set $\{t \in [s, s + \Delta t] \,|\, \chi_{ij,t} = 1\}$, so that $\tau_{ij}(s)$ is the ratio of that length to the length Δt. So, in terms of relative frequency, each $\tau_{ij}(s)$ must be interpreted as the *instantaneous* interaction rate in the states x_i and x_j at time t. We preliminarily consider the case in which the interaction rate $\tau_{ij}(s)$ is constant with respect to time for any $(i, j) \in \{1, 2, \ldots, m\}^2$, and write $\tau_{ij}(s) = \tau_{ij}$. So, for any assigned couple (i, j), the Bernoulli variables χ_{ij,t_1} and χ_{ij,t_2} are identically distributed for any $(t_1, t_2) \in [0, T]^2$ and, for simplicity, we may write $\chi_{ij,t} \equiv \chi_{ij}$ for any t.

The continuous Markov Chain $\{\mathbf{X}_t\}_{t\in[0,T)}$ and the Bernoulli random process $\{\boldsymbol{\chi}_t\}_{t\in[0,T)}$ are coupled in the sense that

1. A function
$$\chi : \mathbf{z}_{ij} \in \Omega^2 \longrightarrow \chi_{ij} = \chi(\mathbf{z}_{ij}) \tag{15}$$
is given;
2. For any $t \in [0, T]$, the transition matrix $\mathbf{P}_t$ is assumed to depend on $\boldsymbol{\chi}_t$ (and only on $\boldsymbol{\chi}_t$), in such a way that
$$P_{\mathbf{z}_{kl}, \mathbf{z}_{ij}} = P_{\mathbf{z}_{kl}, \mathbf{z}_{ij} \,|\, \chi_{kl}=1} \tau_{kl} \Delta t + P_{\mathbf{z}_{kl}, \mathbf{z}_{ij} \,|\, \chi_{kl}=0} (1 - \tau_{kl} \Delta t)\,, \tag{16}$$

where $P_{\mathbf{z}_{kl},\mathbf{z}_{ij}\,|\,\chi_{kl}=1}$ (independent of time) is the transition probability from the state $\mathbf{z}_{kl}$ to the state $\mathbf{z}_{ij}$ under the assumption $\chi_{kl}=1$, and $P_{\mathbf{z}_{kl},\mathbf{z}_{ij}\,|\,\chi_{kl}=0}$ is at any time the transition probability from the state $\mathbf{z}_{kl}$ to the state $\mathbf{z}_{ij}$ under the assumption $\chi_{kl}=0$;

3. We assume that

$$P_{\mathbf{z}_{kl},\mathbf{z}_{ij}\,|\,\chi_{kl}=0}(t)=\delta_k^i\delta_l^j\,,\qquad \forall\, t\in[0,T]\,, \tag{17}$$

where δ_k^i and δ_l^j are the well-known Kronecker symbols, that is, $\delta_k^i=0$ when $i\neq k$ and $\delta_k^i=1$ when $i=k$, and $\delta_l^j=0$ when $j\neq l$ and $\delta_l^j=1$ when $j=l$. This means that the Chain remains in each state $\mathbf{z}_{ij}$ when $\chi_{ij}=0$ (notice that the Markov Chain is stationary if and only if the matrix $P_{\mathbf{z}_{ij},\mathbf{z}_{kl}\,|\,\chi_{ij}=1}$ is constant with respect to t. This remark will prove meaningful in the final sections.)

Accordingly, it is readily seen that, if for any $t\in[0,T]$, the random variables $X_{1,t}$ and $X_{2,t}$ are assumed to be independent, then system (14) may be written as follows:

$$\begin{aligned} p_{x,t+\Delta t}(x_i)-p_{x,t}(x_i)&= \\ &=\Delta t\sum_{l=1}^{m}\Bigg\{\sum_{k\neq i}^{1\ldots m}\tau_{kl}\,p_{x,t}(x_k)p_{y,t}(y_l)P_{z_{kl},x_i\,|\,\chi_{kl}=1}+ \\ &-\tau_{il}\,p_{x,t}(x_i)\sum_{k\neq i}^{1\ldots m}p_{y,t}(y_l)P_{z_{il},x_k\,|\,\chi_{il}=1}\Bigg\}\,. \end{aligned}\qquad (i=1,2,\ldots,p)\,, \tag{18}$$

Next, by dividing both sides by Δt and letting $\Delta t\to 0$, we obtain

$$\begin{aligned} \frac{\mathrm{d}\,p_{x,t}}{\mathrm{d}\,t}(x_i)&=\sum_{l=1}^{m}\Bigg\{\sum_{k\neq i}^{1\ldots m}\tau_{kl}\,p_{x,t}(x_k)p_{y,t}(y_l)P_{\mathbf{z}_{kl},x_i\,|\,\chi_{kl}=1}+ \\ &-\tau_{il}\,p_{x,t}(x_i)\sum_{k\neq i}^{1\ldots m}p_{y,t}(y_l)P_{\mathbf{z}_{il},x_k\,|\,\chi_{il}=1}\Bigg\} \end{aligned}\qquad (i=1,2,\ldots,p)\,, \tag{19}$$

if the transition matrix $\mathbf{P}_{\boldsymbol{\chi}=\mathbf{1}}\equiv(P_{\mathbf{z}_{kl},\mathbf{z}_{ij}\,|\,\chi_{kl}=1})$ is stationary.

A comparison between system (19) and system (23), which will be made in Section 6, now shows at once that the latter can be viewed simply as a particular interpretation of the basic Chapman–Kolmogorov vector equation for the marginal state vectors of a two-dimensional vector Markov Chain. Furthermore, the random variables of the chain are pairs of states of the particles of the system. It could also be easily proved that the dimension of the vector Markov Chain describing the evolution of the system is strictly connected to the choice to allow and consider only pairwise interactions between the particles of the system. If one would decide to allow and to consider multiple interactions (e.g., simultaneous interactions between $p>2$ particles), then the Markov Chain describing the evolution should be p-dimensional.

This will also allow us to state that all the conditions leading to the deduction of system (23) in the mathematical physical framework can be reduced to the sole following axiom:

The random process of p-tuples of states of the particles of the system is a p-dimensional vector Markov Chain conditioned by simultaneous interactions of p particles.

This conclusion allows to acknowledge that all assumptions about the influence of the external world on the evolution of the system can and must involve only the form of the transition matrices and their possible dependence on time or external parameters. In the remaining part of the paper, we shall treat the case in which the transition matrices are in turn random variables depending on other random variables describing relevant features of the external world.

5. Markov Chains with Stochastic Transition Matrices: A Heuristic Description

In this Section, we leave aside for the moment any reference to Kinetic Theory and to the mathematical description of the evolution of a many-particle system, in order to discuss in quite general terms the Markov Chains for which the transition matrices are never assigned but simply treated as random variables, depending on the outcomes of some additional experiments. In particular, we want to offer an example of a sequence of extractions from an urn for which this is the case.

Confining ourselves to the case of stationary Markov Chains, consider a sequence of extractions from an urn U containing b black balls and w white balls, and set $n = b + w$. The extractions are performed according to the following rule: for $h \geq 3$, before extracting the h-th ball from the urn, we replace in the urn the ball we have drawn at the $(h-2)$-th extraction. So, it is readily seen that the transition matrix at h-th step is

$$\mathbf{P}_h \equiv \begin{pmatrix} \dfrac{b-1}{n-1} & \dfrac{w}{n-1} \\ \dfrac{b}{n-1} & \dfrac{w-1}{n-1} \end{pmatrix} \tag{20}$$

for any $h \geq 2$, at least if nothing modifies the state space (i.e., the urn) at any step. However, now, at each step, we actually intervene on the urn U, and our intervention will be decided by an extraction from another urn U' containing only three balls, labeled in any way we want (for instance, one can be white, one can be black, and the third red). If we draw from U' the white ball, then we add only a white ball in U; if we draw the black ball, then we add only a black ball; if we draw the red ball, then we leave U unchanged; in any case, we replace the ball in the urn U'. Accordingly, the transition matrix from the outcomes of the first extraction from U to those of the second one will be:

$$\mathbf{P}_1^1 = \begin{pmatrix} \dfrac{b-1}{n} & \dfrac{w+1}{n} \\ \dfrac{b}{n} & \dfrac{w}{n} \end{pmatrix} \tag{21}$$

if we have drawn from U' the white ball;

$$\mathbf{P}_1^2 = \begin{pmatrix} \dfrac{b}{n} & \dfrac{w}{n} \\ \dfrac{b+1}{n} & \dfrac{w-1}{n} \end{pmatrix} \tag{22}$$

if the ball drawn from U' is black; finally, $\mathbf{P}_1^3$ will simply be the matrix (20) if we have drawn from U' the red ball. In the considered case, the matrices $\mathbf{P}_1^1$, $\mathbf{P}_1^2$ and $\mathbf{P}_1^3$ have the same probability $1/3$, but if U' contained for instance three white balls, four black balls, and only two red balls, then the probability of $\mathbf{P}_1^1$ was $3/9$, that of $\mathbf{P}_1^2$ was $4/9$ and that of $\mathbf{P}_1^3$ was $2/9$.

This example aims to introduce in as general terms as possible the notion of Markov Chains *coupled* with another random process. As a matter of fact, we considered a random process $\{X_i\}_{i\in\mathbb{N}}$, which is a time-discrete Markov Chain with state space $S \equiv \{0,1\}$ (where "0" stands for "black" and "1" stands for "white") and another time-discrete *independent* random process $\{E_j\}_{j\in\mathbb{N}}$ with state space $S' \equiv \{0,1,2\}$ (where "0" stands for "black", "1" stands for "white", and "2" stands for "red"). The form of the transition matrix of

process $\{X_i\}_{i\in\mathbb{N}}$ at each step i is ruled by the outcomes of process $\{E_i\}_{i\in\mathbb{N}}$ at all the steps from the first one to the i-th one, so that we shall give the following two definitions:

1. The processes $\{X_i\}_{i\in\mathbb{N}}$ and $\{E_i\}_{i\in\mathbb{N}}$ are said to be *coupled*;
2. $\{E_i\}_{i\in\mathbb{N}}$ is called *the ruling process* of the Markov Chain $\{X_i\}_{i\in\mathbb{N}}$.

Notice, however, that the variables of the ruling process are assumed to be not only independent but also identically distributed, while independence is assumed only for the sake of simplicity, and does not affect the generality of our description, the identical distribution of the random variables E_i ($i \in \mathbb{N}$) assures the independence from h of all conditional transition probabilities, at least when each transition matrix $\mathbf{P}_i$ is ruled only by E_i and is independent of E_h for any $h < i$. Of course, the above elementary example does not satisfy such condition, but it is possible to find different examples—though less simple—for which it is actually met.

In order to go a little deeper into the behavior of a Markov Chain in the presence of a ruling process, let us denote by $\mathbf{X} \equiv \{X_i\}_{i\in\mathbb{N}}$ the given Markov Chain and by $\mathbf{E} \equiv \{E_i\}_{i\in\mathbb{N}}$, by $D \equiv \{u_1, \ldots, u_m\}$ the state space of $\mathbf{X}$ and by $S \equiv \{\eta_1, \ldots, \eta_n\}$ that of $\mathbf{E}$, by $\mathbf{p}_i \equiv (p_i(u_1), p_i(u_2), \ldots, p_i(u_m))$ the absolute probability distribution on D at i-th step (the i-th state vector of the chain) and by $\boldsymbol{\pi}_i \equiv (\pi_i(\eta_1), \pi_i(\eta_2), \ldots, \pi_i(\eta_n))$ the i-th state vector of the ruling process $\mathbf{E}$. Moreover, assuming once and for all that the transition matrix of $\mathbf{X}$ at each step is ruled only by the value of $\mathbf{E}$ at the same step, we shall use the symbols $\mathbf{p}_{k\,|\,\eta_h} \equiv (p_k(u_1\,|\,\eta_h), p_k(u_2\,|\,\eta_h), \ldots, p_k(u_m\,|\,\eta_h))$ and $\mathbf{P}_k(\cdot|\eta_h) \equiv (P_{k,ij}(\eta_h))_{1\le i,j\le m}$ to denote the state vector at k-th step and the transition matrix of $\mathbf{X}$ at the same step conditional to the assumption that $E_k = \eta_h$..

Now, given the initial state vector $\mathbf{p}_0$ of $\mathbf{X}$, we obviously have

$$\begin{aligned}
&\mathbf{p}_{1\,|\,\eta_h} = \mathbf{p}_0\mathbf{P}_0(\cdot|\eta_h)\,, \qquad \forall\,\eta_h \in S\,,\\
&\mathbf{p}_{2\,|\,\eta_h} = \mathbf{p}_{1|\eta_h}\mathbf{P}_1(\cdot|\eta_h)\,, \qquad \forall\,\eta_h \in S\,,\\
&\vdots\\
&\mathbf{p}_{k\,|\,\eta_h} = \mathbf{p}_{k-1|\eta_h}\mathbf{P}_{k-1}(\cdot|\eta_h)\,, \qquad \forall\,\eta_h \in S\,,
\end{aligned}$$

so that

$$\mathbf{p}_{k\,|\,\eta_h} = \mathbf{p}_0\mathbf{P}_1(\cdot|\eta_h)\mathbf{P}_2(\cdot|\eta_h)\ldots\mathbf{P}_{k-1}(\cdot|\eta_h)\,, \qquad \forall\,k \in \mathbb{N}\,,\ \forall\,\eta_h \in S\,.$$

This shows that—under the assumption that the random variables E_i are independent and identically distributed—a chain conditioned to the repetition of the same "external" event is *stationary*. More generally, for any fixed k, consider any element $\boldsymbol{\eta}^{(k)} \equiv (\eta_{i_1}, \eta_{i_2}, \ldots, \eta_{i_k})$ of S^k. One has

$$\begin{aligned}
&\mathbf{p}_{1\,|\,\eta_{i_1}} = \mathbf{p}_0\mathbf{P}_0(\cdot|\eta_{i_1})\,, \qquad \forall\,\eta_{i_1} \in S\,,\\
&\mathbf{p}_{2\,|\,(\eta_{i_1},\eta_{i_2})} = \mathbf{p}_{1\,|\,\eta_{i_1}}\mathbf{P}_1(\cdot|\eta_{i_2})\,, \qquad \forall\,(\eta_{i_1},\eta_{i_2}) \in S^2\,,\\
&\vdots\\
&\mathbf{p}_{k\,|\,\boldsymbol{\eta}^{(k)}} = \mathbf{p}_{k-1\,|\,(\eta_{i_1},\ldots\eta_{i_{k-1}})}\mathbf{P}_{k-1}(\cdot|\eta_{i_k})\,, \qquad \forall\,\boldsymbol{\eta}^{(k)} \in S^k\,,
\end{aligned}$$

so that

$$\mathbf{p}_{k\,|\,\boldsymbol{\eta}^{(k)}} = \mathbf{p}_0\mathbf{P}_0(\cdot|\eta_{i_1})\mathbf{P}_1(\cdot|\eta_{i_2})\,\ldots\,\mathbf{P}_{k-1}(\cdot|\eta_{i_k})\,, \qquad \forall\,k \in \mathbb{N}\,,\ \forall\,\boldsymbol{\eta}^{(k)} \in S^k$$

(note that the symbol $\mathbf{p}_{k\,|\,\boldsymbol{\eta}^{(k)}}$ denotes the state vector at k-th step depending on all the external events from the first step to "just before" the k-th step).

These relations are of the greatest importance for applications to the behavior of many-particle systems subjected to random external influences, not only to understand its dependence on the sequence of all past and present external events but also to see that a complete description of such behavior cannot ignore a comparison between different predictions corresponding to different sequences of such events.

6. Kinetic-Theoretic Models: A Recall

As laid out in Section 4, the equations expressing the link between the state vectors of a coupled semi-Markov continuous process via the transition matrices, when written in the form of differential equations, simply become the equations describing in statistical terms the evolution of a many-particle system according to the Kinetic Theory for Active Particles. In order to clarify this point, we devote this Section to recalling some basic features of the kinetic-theoretic scheme, so that, in the subsequent Section, once the rôle of Markov Chains in the equations of the model is pointed out, we shall be able to study the way in which the dependence on random events modify the equations themselves.

A *complex system* is a set S of a very large number of objects called *individuals* or *active particles*, that, as already laid out in the Introduction and shown in the papers cited therein, can represent not only particles of a mechanical systems but also cells, living individuals, or human beings. The *state* of each individual in S is defined, according to the context in which we aim to develop our description, as a scalar variable u or a vector variable $\mathbf{u} \equiv (u^1, u^2, \ldots, u^m)$ which, much more for historical reasons than for reasons suggested by the context, is called *activity* or—more simply—*a state variable*. This variable may describe the kinetic state of a material particle, or the activity of a cell in its interactions with other cells of a different tissue, or the health state of a living individual, or the level of wealth and the social condition of a human being in a collectivity. According to the context, the activity may be assumed to take its values in a *discrete* domain (typically, a finite or countable subset of $\mathbb{Z}$ or $\mathbb{Z}^m$) or in a *continuous* domain (typically, a bounded or unbounded real interval or a bounded or unbounded domain of $\mathbb{R}^m$). In any case, the domain D_u (or $D_{\mathbf{u}}$) in which the state variable is allowed to take its values is called the *state space* of the system [21].

As in Boltzmann's Kinetic Theory of Gases [7], the viewpoint from which the mathematical framework is developed is statistical, i.e., we are interested to describe the evolution in time of the system as a whole rather than the way in which the states of single particles vary with time. Since we are aware that any precise description of the state of each particle of S must be given up, for both theoretical and technical reasons, we decide to study the relative frequency (or probability) distribution over the state space at each instant. In other words, the state variable is conceived as a *random variable*, and the goal of the study becomes the forecast of the evolution of its probability density function as well as of some of its parameters (expected value, moment of order two, standard deviation, and so on) [21].

In many cases of interest, S is considered to be split into a family $\{S_1, S_2, \ldots, S_k\}$ of subsystems, called its *functional subsystems*. The introduction of such subsystems aims to describe the interactions between different subclasses of individuals of S, characterized by different reactions to interactions, and offers several applications (1) in biology, for instance to describe the fight between tumor cells and immune system, or the competition between different species to model Darwinian selection; (2) in social sciences, to model interactions between social classes; and (3) in economics, to model the fluxes of wealth, e.g., between the class of financial managers, that of entrepreneurs, and that of salaried workers.

In some cases, the same variable can be used to identify the state of the members of all the functional subsystems, while in other cases, different state variables must be used for different subsystems.

In view of the aims of the present paper, we shall, however, disregard any possible decomposition of the system into different subsystems. Such a decomposition would simply result in a quite useless complication of notation and multiplication of similar equations. Accordingly, in this Section, we shall describe in some details only the case of a system S consisting in one subsystem equal to the whole S. The state space associated with the system will be denoted by the symbol

$$D = \{u_1, u_2, \ldots, u_m\}.$$

For any t in a time-interval $\mathbb{I} \subseteq \mathbb{R}$, the state of S at time t will be identified by a probability distribution (also called *a state vector*) $\mathbf{f}_t \equiv (f_t(u_1), f_t(u_2), \ldots, f_t(u_m))$ on D, and, for any $k \in \{1, 2, \ldots, m\}$, we can define the function $f_k : t \in \mathbb{I} \longrightarrow f_k(t) \equiv f_t(u_k) \in [0, 1]$. According to this definition, we have

$$\sum_{k=1}^{m} f_k(t) = \sum_{k=1}^{m} f_t(u_k) = 1,.$$

Now, the evolution of the system S is viewed as a time-continuous stochastic process, and the time derivative of each probability function f_k is expressed, according to the law of alternatives, in terms of transition probabilities. More precisely, for any $(r, s, j) \in \{1, 2, \ldots m\}^3$, the symbol ${F^r}_{sj} \equiv F_h(u_s, u_j; u_r)$ will denote the probability that a particle of S *falls* from the state u_s to the state u_r after an interaction with another particle of S which is in the state u_j: accordingly,

$$\sum_{r=1}^{m} {F^r}_{sj} = 1, \qquad \forall\, (s, j) \in \{1, 2, \ldots, m\}^2.$$

Thus, using the law of alternatives, we write the following system of differential equations

$$\frac{\mathrm{d}\, f_h}{\mathrm{d}\, t}(t) = \sum_{i,j}^{1\ldots m} \tau_{ij} {F^h}_{ij} f_i(t) f_j(t) - f_h(t) \sum_{i,j}^{1\ldots m_1} \tau_{hj} {F^i}_{hj} f_j(t) \tag{23}$$

where, for any $(r, s) \in \{1, 2, \ldots, m\}^2$, $\tau_{rs} \equiv \tau(u_r, u_s)$ is the so-called *encounter rate* of particles of S in the states u_r and u_s, that is the number of pairwise interactions between particles of S, that are in the states u_r and u_s, per time unit, so that for any sufficiently small Δt, the product $\tau_{rs}\Delta t$ is the probability that one such interaction occur in the time interval Δt *provided* the individuals involved are in the states u_r and u_s (accordingly, $\tau_{rs}\Delta t$ is in turn a *conditional* probability and τ_{rs} is the ratio of a conditional probability to time).

The terms on the right-hand side of Equation (23) express, respectively, the increase in probability of state u_h as the probability that some «candidate» particles of S in the state u_i interact with some *field* particles of S in the state u_j and, as a consequence, fall in the state u_h with a positive probability, and the decrease in probability of state u_h as the probability that some «test» particles of S in the state u_h interact with some *field* particles of S in the state u_j and, as a consequence, *leave* the state u_h with a positive probability.

7. The Continuous Case and KTAP Systems with Stochastic Transition Matrices

Now, in order to deepen the reasoning presented in [20] (Section 7), we consider again a stochastic process $(\mathbf{X}, \boldsymbol{\chi}, \mathbf{E})$, where $\mathbf{X} \equiv \{\mathbf{X}_t\}_{t\in[0,T)}$, $\boldsymbol{\chi} \equiv \{\boldsymbol{\chi}_t\}_{t\in[0,T)}$ and $\mathbf{E} \equiv \{E_t\}_{t\in[0,T)}$; just for the sake of simplicity, $\mathbf{X}_t \equiv (X_{1,t}, X_{2,t})$ is assumed to be two-dimensional and E_t is assumed to be scalar for any $t \in [0, T)$; $\chi_t \equiv (\chi_{ij,t})_{1\leq i,j\leq m}$ is instead a $n \times n$ matrix-valued

random variable for any $t \in [0, T)$. For the sake of completeness and clarity, we collect again our assumptions on these processes:

1. We denote by $D^2 \equiv \{u_1, u_2, \ldots, u_m\}^2$ and $S \equiv \{e_1, e_2, \ldots, e_n\}$, respectively, the ranges of $\mathbf{X}_t$ and E_t for any $t \in [0, T)$. The range of $\chi_{ij,t}$ is simply the set $\{0, 1\}$ for any $(i, j) \in \{1, 2, \ldots, m\}^2$ and any $t \in [0, T)$;
2. $\mathbf{X}$ is a Markov Chain, and for any $t \in [0, T)$ the variable $\mathbf{X}_t$ depends on all the variables $\chi_{ij,t}$ but is independent on any variable $\chi_{ij,t'}$ for $t' \neq t$;
3. For any $t \in [0, T)$ each variable $X_{h,t}$ $(h = 1, 2)$ depends on all the variables E_s $(s \leq t)$ only via the dependence of its transition matrices $\mathbf{P}_s$ $(s < t)$ on the corresponding variables E_s;
4. For any $t \in [0, T)$ the variables $X_{1,t}$ and $X_{2,t}$ are independent;
5. For any pair of triples (i, j, t) and (i', j', t') in $D^2 \times [0, T)$, with $t \neq t'$, the random variables $\chi_{ij,t}$ and $\chi_{i'j',t'}$ are independent;
6. For any pair (t, t') of times in $[0, T)$, the random variables $\chi_{ij,t}$ and $\chi_{ij,t'}$ are identically distributed;
7. Each variable $\chi_{ij,t}$ is independent of all variables E_s, with $s \in [0, T)$;
8. The variables E_t are independent and identically distributed.

This stated, before going on to the technical development of the consequences of our definitions and assumptions, we need to discuss and explain their meaning in some details. To this end, we start by outlining an example of a stochastic process of the type under examination, and try to show how it suggests the above conditions to be quite reasonable.

It is to be noted that a complete description of any example of a time-continuous process of such kind would require a long and rather complicated treatment, which would go far beyond the limits of this paper and requires a separate paper. So, we shall confine ourselves to a concise outline. Furthermore, in the same way as for the case of time-discrete processes, we shall not refer to the case of interacting subjects, because this would force us to consider at least two-dimensional vector variables, and we want to avoid formal complications.

This stated, consider the evolution in time of the bank account of any subject we may choose. Such evolution is readily seen to be well described as a *continuous* "random walk", once the probability distribution on the set of possible amounts of payments and incomes and the rates have been assigned; the probability of any possible amount of money in the account at each instant t does not depend on its history, but—roughly speaking—only on the amount at instant $t - \mathrm{d}t$. Now, as the "ruling" urn, we choose a set of possible abrupt external events producing unpredictable new expenses or incomes (climate changes requiring unusual heating or cooling expenses, unexpected taxes, thefts, fires, earthquakes, unexpected inheritances, lottery winnings, unexpected salary advancements), all labeled with the amounts of money they could produce, so that the range D and the range S are the same set (with different probability distribution functions, of course). These events are random, with rates (probabilities) that are usually very low, and must be *estimated* by means of a statistical analysis of the history of similar events, and are very likely to be independent of each other, as it is difficult to assume that any one of the above events take into account the preceding ones to decide whether they will take place or not. This gives a full explanation of assumption 8. In the same way, the assumptions on the variables $\chi_{ij,t}$ seem to be very plausible, as these variables simply express accounting information, which is essentially linked to the life and habits of the subject under examination, and we could say that they *define* the system.

We are well aware that, while the condition on $\mathbf{E}$, as expressing the abruptness and the randomness of external events, can be assumed in almost any application of the model, the conditions on the interaction rates would require to be modified according to the

meaning of the word "interaction" in each context, and could turn out to be quite unlikely for more complicated examples, but this point cannot be examined here, and must be postponed to future analysis.

Now, turning again to the formal development proposed above, in order to study the form of System (19) when the transition matrices of the Markov Chain $\mathbf{X}$ depend on the values of random variables E_t, we shall first agree to set $\mathbf{u}_{hk} \equiv (u_h, u_k)$ and to denote by the symbol $P_{\mathbf{u}_{il},u_k \,|\, \chi_{il}=1, E_t=ej}$ the probability that the couple of states (u_i, u_l) falls in any state whose first component is u_k when not only a particle in the state u_i interacts with a particle in the state u_l but also the random variable E_t takes the value e_j. Furthermore, we agree to denote by $\rho_{ilj}\Delta t$ the joint probability that an interaction between a particle in the state u_i and a particle in the state u_l occurs *and* $E_t = e_j$ in any time interval of length Δt, so that ρ_{ilj} will be the *instantaneous* probability density of these joint events (which we have implicitly assumed to be constant).

In addition, we must also introduce—for any $r \in \mathbb{N}$—the symbol $\mathbf{e}^{(r)} \equiv (e_{h_1}, e_{h_2}, \ldots, e_{h_r})$ to denote any unspecified element of S^r. For reasons that will become clear later, we also agree to denote by $\sigma_{h_1 h_2 \ldots h_r}$ the joint probability of events $e_{h_1}, e_{h_2}, \ldots, e_{h_r}$. Note that, in view of the independence of $\chi_{ij,t}$ from the variables of the process $\mathbf{E}$, the product $\tau_{ij}\sigma_{h_1 h_2 \ldots h_r}$ is the joint probability density of events $e_{h_1}, e_{h_2}, \ldots, e_{h_r}$ *and* interactions between a particle in the state u_i with a particle in the state u_j. According to the classical treatment of the kinetic model, we will assume that τ_{ij} and $\sigma_{h_1 h_2 \ldots h_r}$ are independent of time for any i, j, h_1, h_2, $\ldots$, h_r.

This stated, in the time-continuous model with transition matrices that vary randomly over time depending on the values of the variables E_t, System (19) can be rewritten in many ways, according to the information we want to take into account and the information we want to obtain.

First of all, if we want to link the instantaneous variation of the distribution of the states over the particles of the system to assumptions about possible past r-tuples of external events (where r is assigned arbitrarily in advance), we may simply consider the system

$$\begin{aligned}
&\frac{\mathrm{d}\, p_t}{\mathrm{d}\, t}(u_i \,|\, \mathbf{e}^{(r)}, E_t = e_h) = \\
&\quad = \sum_{l=1}^{m} \left\{ \sum_{k \neq i}^{1 \ldots m} \tau_{kl}\, p_t(u_k \,|\, \mathbf{e}^{(r)}) p_t(u_l \,|\, \mathbf{e}^{(r)}) P_{\mathbf{u}_{kl},u_i \,|\, \chi_{kl}=1, E_t=e_h} + \right. \qquad (i = 1, 2, \ldots, m)\,, \qquad (24) \\
&\quad \left. - \tau_{il}\, p_t(u_i \,|\, \mathbf{e}^{(r)}) \sum_{k \neq i}^{1 \ldots m} p_t(u_l \,|\, \mathbf{e}^{(r)}) P_{\mathbf{u}_{il},u_k \,|\, \chi_{il}=1, E_t=e_h} \right\},
\end{aligned}$$

where (a) we have dropped out the index of the component of the state vectors because the variable is always the same; (b) $\mathbf{u}_{kl} \equiv (u_k, u_l)$ and $\mathbf{u}_{il} \equiv (u_i, u_l)$; (c) the conditional symbol "$|\, \mathbf{e}^{(r)}$" should be read "under the assumption that r variables of the process $\mathbf{E}$ in the time interval $[0, t)$ have taken the values $e_{h_1}, e_{h_2}, \ldots, e_{h_r}$"; (d) we have set

$$P_{\mathbf{u}_{il},u_k \,|\, \chi_{il}=1, E_t=e_h} \equiv P_{|\, \chi_{il}=1}(\mathbf{u}_{il}, u_k \,|\, E_t = e_h) \equiv P(\mathbf{u}_{il}, u_k \,|\, \chi_{il} = 1, E_t = e_h)\,.$$

Note that, as a matter of fact, System (24) is *not* only one system, but a whole set of n^r systems of equations, one for each possible choice of assumption $\mathbf{e}^{(r)}$ in S^r. As we shall discuss in detail in the final Section of this paper, this would allow us to compare the possible behaviors of the many-particle system under examination corresponding to different possible *histories* of external influences if we could treat the distributions at right-hand side like *known* functions of $\mathbf{e}^{(r)}$. Since this is not the case, bearing in mind the independence from time of $\sigma_{h_1 h_2 \ldots h_r}$ for any choice of $(h_1, h_2, \ldots, h_r) \in \{1, 2, \ldots, n\}^n$, we

may multiply each system of the set by the corresponding $\sigma_{h_1 h_2 \dots h_r}$, then add side by side all the n^r systems of the set, to obtain the unique system

$$\begin{aligned}
&\frac{\mathrm{d}\, p_t}{\mathrm{d}\, t}(u_i \,|\, E_t = e_h) = \\
&= \sum_{l=1}^{m} \left\{ \sum_{k \neq i}^{1 \dots m} \tau_{kl} p_t(u_k) p_t(u_l) P_{\mathbf{u}_{kl}, u_i \,|\, \chi_{kl}=1, E_t = e_h} + \right. \qquad (i = 1, 2, \dots, p)\,, \\
&\left. - \tau_{il} p_t(u_i) \sum_{k \neq i}^{1 \dots m} p_t(u_l) P_{\mathbf{u}_{il}, u_k \,|\, \chi_{il}=1, E_t = e_h} \right\},
\end{aligned} \tag{25}$$

where we have taken into account the law of total probabilities

$$\begin{aligned}
&\sum_{h_1, h_2, \dots, h_r}^{1 \dots n} p_t(u_i \,|\, \mathbf{e}^{(r)}, E_t = e_h) \sigma_{h_1 h_2 \dots h_r} = p_t(u_i \,|\, E_t = e_h)\,, \\
&\sum_{h_1, h_2, \dots, h_r}^{1 \dots n} p_t(u_k \,|\, \mathbf{e}^{(r)}) p_t(u_l \,|\, \mathbf{e}^{(r)}) \sigma_{h_1 h_2 \dots h_r} = p_t(u_k) p_t(u_l)\,, \\
&\sum_{h_1, h_2, \dots, h_r}^{1 \dots n} p_t(u_i \,|\, \mathbf{e}^{(r)}) p_t(u_l \,|\, \mathbf{e}^{(r)}) \sigma_{h_1 h_2 \dots h_r} = p_t(u_i) p_t(u_l)\,.
\end{aligned}$$

It is obvious that, while the last system (25) is much simpler to manage, the set of systems expressed by (24) gives much more information in connection with the sequence of external events preceding the one identified by the value of E_t. This point requires a careful discussion, which will be proposed in the concluding Section of the paper.

8. Conclusions and Perspectives

As the reader will have undoubtedly noticed, the main aim of the present paper was to give a detailed and precise account, in the light of the basic language of Markov Chains, of the case in which the transition matrices appearing in the equations governing the evolution of a many-particle system in turn depend on stochastic events taking place in the external world and influencing the behavior of the system. This case was treated only as a hint in [20], which focused on possible applications of the Kinetic Theory for Active Particles to social and economic problems, and has been discussed in more general terms here.

In order to synthesize the main significant aspects of the discussion presented here, we want to direct the attention of the reader to the fact that the dependence of the form of transition matrices of a Markov Chain from random events at each step produces a "tree" of possible ordinary, deterministic chains, whose number of branches, in the time-discrete case, grows exponentially with the number of steps. More precisely, at the r-th step, if the number of possible events influencing the form of the transition matrix is n, the number of chains is n^r, even when the form of r-th transition matrix is modified only by the r-th random external event.

In the time-continuous case (which is the most interesting one in view of applications to physical or biological or economic or social problems), the situation is far more complicated. In principle, at any time $t > 0$, we could have any number of different sequences of transition matrices, with different probabilities, and we should formulate suitable assumptions on the sequences to be taken into account. However, since the form of the system hides the dependence of the distributions from $\mathbf{e}^{(r)}$, we are forced to use the law of alternatives, and to *ignore* any assumption on the sequence of external events preceding the one influencing the transition matrix at time t.

Conversely, as a consequence of this remark, the solution of system (25) in an arbitrary time interval $[0, T)$ is somehow misleading, as it suggests the idea that a particular external event took place at time $t = 0$ and influenced the whole behavior of the system in the whole interval $[0, T)$. So, we lose any reference to a sequence of random external events. The only correct way to apply the notion of "random variable transition matrix" is the following: we assume that a sequence of times $\{t_1, t_2, \ldots, t_r\} \subset [0, T)$ exists such that some events $e_{h_1}, e_{h_2}, \ldots, e_{h_r}$ happened, influencing the transition matrix in such a way that in the time interval $[t_j, t_{j+1})$, the system governing the evolution of the system is

$$\begin{aligned} &\frac{\mathrm{d}\, p_t}{\mathrm{d}\, t}(u_i) = \\ &\quad = \sum_{l=1}^{m} \Bigg\{ \sum_{k\neq i}^{1\ldots m} \tau_{kl}\, p_t(u_k)\, p_t(u_l)\, P_{\mathbf{u}_{kl},u_i \,|\, \chi_{kl}=1} + \qquad (i = 1, 2, \ldots, m)\,, \\ &\quad - \tau_{il}\, p_t(u_i) \sum_{k\neq i}^{1\ldots m} p_t(u_l)\, P_{\mathbf{u}_{il},u_k \,|\, \chi_{il}=1} \Bigg\}\,, \end{aligned} \tag{26}$$

for $t \in [0, t_1)$,

$$\begin{aligned} &\frac{\mathrm{d}\, p_t}{\mathrm{d}\, t}(u_i \,|\, E_t = e_{h_1}) = \\ &\quad = \sum_{l=1}^{m} \Bigg\{ \sum_{k\neq i}^{1\ldots m} \tau_{kl}\, p_t(u_k)\, p_t(u_l)\, P_{\mathbf{u}_{kl},u_i \,|\, \chi_{kl}=1, E_t=e_{h_1}} + \qquad (i = 1, 2, \ldots, m)\,, \\ &\quad - \tau_{il}\, p_t(u_i) \sum_{k\neq i}^{1\ldots m} p_t(u_l)\, P_{\mathbf{u}_{il},u_k \,|\, \chi_{il}=1, E_t=e_{h_1}} \Bigg\}\,, \end{aligned} \tag{27}$$

for $t \in [t_1, t_2)$ and

$$\begin{aligned} &\frac{\mathrm{d}\, p_t}{\mathrm{d}\, t}(u_i \,|\, \mathbf{e}^{(j-1)}, E_t = e_{h_j}) = \\ &\quad = \sum_{l=1}^{m} \Bigg\{ \sum_{k\neq i}^{1\ldots m} \tau_{kl}\, p_t(u_k \,|\, \mathbf{e}^{(j-1)})\, p_t(u_l \,|\, \mathbf{e}^{(j-1)})\, P_{\mathbf{u}_{kl},u_i \,|\, \chi_{kl}=1, E_t=e_{h_j}} + \qquad (i = 1, 2, \ldots, m)\,, \\ &\quad - \tau_{il}\, p_t(u_i \,|\, \mathbf{e}^{(j-1)}) \sum_{k\neq i}^{1\ldots m} p_t(u_l \,|\, \mathbf{e}^{(j-1)})\, P_{\mathbf{u}_{il},u_k \,|\, \chi_{il}=1, E_t=e_{h_j}} \Bigg\}\,, \end{aligned} \tag{28}$$

for $t \in [t_j, t_{j+1})$, where $2 \leq j \leq r$ and $\mathbf{e}^{(j)} = \{e_{h_1}, e_{h_2}, \ldots, e_{h_j}\}$. Since the distributions on the right-hand side of each system in the interval $[t_j, t_{j+1})$ are known, as obtained by solving the analogous system in the interval $[t_{j-1}, t_j)$, we obtain a *possible* evolution of the system in the whole interval $[0, T)$. More precisely, we obtain a solution for any assigned r-tuple $\mathbf{e}^{(r)} \equiv (e_{h_1}, e_{h_2}, \ldots, e_{h_r})$, and each solution has the same probability of the sequence $\{e_{h_1}, e_{h_2}, \ldots, e_{h_r}\}$, that is the probability $\sigma_{h_1 h_2 \ldots h_r}$. In view of identical distribution and independence of random variables E_t, one has

$$\sigma_{h_1 h_2 \ldots h_r} = \prod_{s=1}^{r} q_{h_s}\,,$$

where $q_{h_s} = \mathbf{P}(E_t = e_{h_s})$ for any t.

Now, besides the complete irrelevance of any precise identification of times $t_1, t_2, \ldots, t_r$, we must remark that $r \in \mathbb{N}$ is arbitrary, so that the number of possible solutions in the

given time interval $[0, T)$ is potentially infinite. More precisely, we may repeat the same procedure for any $r \in \mathbb{N}$, to obtain n^r solutions

$$\mathbf{p}_r(\mathbf{e}^{(r)}, t) \equiv (p_r(\mathbf{e}^{(r)}; u_1, t), p_r(\mathbf{e}^{(r)}; u_2, t), \ldots, p_r(\mathbf{e}^{(r)}; u_m, t)), \qquad (t \in [0, T)),$$

defined by the conditions

$$p(\mathbf{e}^{(r)}; u_i, t) = \begin{cases} p(u_i, t) & \text{for } t \in [0, t_1) \\ p(u_i, t \mid E_t = e_{h_1}) \equiv p(u_i, t \mid \mathbf{e}^{(1)}) & \text{for } t \in [t_1, t_2) \\ p(u_i, t \mid \mathbf{e}^{(1)}, E_t = e_{h_2}) \equiv p(u_i, t \mid \mathbf{e}^{(2)}) & \text{for } t \in [t_2, t_3) \\ \quad \vdots & \quad \vdots \\ p(u_i, t \mid \mathbf{e}^{(r-2)}, E_t = e_{h_{r-1}}) \equiv p(u_i, t \mid \mathbf{e}^{(r-1)}) & \text{for } t \in [t_{r-1}, t_r) \\ p(u_i, t \mid \mathbf{e}^{(r-1)}, E_t = e_{h_r}) \equiv p(u_i, t \mid \mathbf{e}^{(r)}) & \text{for } t \in [t_r, T), \end{cases}$$

where $i = 1, 2, \ldots, m$.

We cannot but acknowledge that this is a purely abstract treatment of the problem of describing the behavior of a many-particle system influenced by random external events. Furthermore, we cannot hide that its abstract character, joined with the formal complication and the increased number of variables and parameters, can make the application of the model to real phenomena remarkably difficult; however, we are encouraged by the remark that the difficulty is in nature the same as in the classical kinetic model in which the transition matrices are assumed to be constant, and the difference is only technical. The subsequent steps, besides the classical existence, uniqueness, and stability problems for each of the partial solutions listed above, and the explicit computation of solutions (at least by means of numerical simulations in suitable particular cases), must be essentially of a statistical type, and mainly oriented by applications (however, this is true also in the classical model). For each application, we need to establish whether a r^* exists such that the rate of r-tuples of external events is zero for any $r > r^*$. At the same time, one should also study the *expected behavior* of the system, described by the relation

$$\overline{\mathbf{p}}(t) = \sum_{r=1}^{\infty} \rho_r \sigma_{e_{h_1} e_{h_2} \ldots e_{h_r}} \mathbf{p}_r(\mathbf{e}^{(r)}, t),$$

where ρ_r is the probability density (rate) of a sequence of r external events. It is quite clear that the terms corresponding to large values of r are smaller the larger r is. However, the application to real systems on the one hand, suggests, and on the other hand, requires that only a finite number of external events actually occur in any given time interval of assigned length. The analysis of these points is the natural development of the present proposal, and should be the object of future papers.

Author Contributions: Conceptualization, B.C. and M.M.; Formal analysis, B.C. and M.M.; Writing—original draft, B.C. and M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in the study are included in the articlel, further inquiries can be directed to the corresponding author.

Acknowledgments: The research by M.M. has been carried out under the auspices of GNFM (National Group of Mathematical-Physics) and of INdAM (National Institute of Advanced Mathematics).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Waldmann, M.R.; Martignon, L. A Bayesian network model of causal learning. In *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*; Routledge: London, UK, 2022; pp. 1102–1107.
2. Sajid, Z.; Khan, F.; Zhang, Y. Integration of interpretive structural modelling with Bayesian network for biodiesel performance analysis. *Renew. Energy* **2017**, *107*, 194–203. [CrossRef]
3. Hosseini, S.; Barker, K. A Bayesian network model for resilience-based supplier selection. *Int. J. Prod. Econ.* **2016**, *180*, 68–87. [CrossRef]
4. Toscani, G.; Sen, P.; Biswas, S. Kinetic exchange models of societies and economies. *Phil. Trans. Royal Soc. A Math. Phys. Eng. Sci.* **2022**, *380*, 2224. [CrossRef] [PubMed]
5. Dimarco, G.; Pareschi, L.; Toscani, G.; Zanella, M. Wealth distribution under the spread of infectious diseases. *Phys. Rev. E* **2020**, *102*, 022303. [CrossRef] [PubMed]
6. Bernardi, E.; Pareschi, L.; Toscani, G.; Zanella, M. Effects of vaccination efficacy on wealth distribution in kinetic epidemic models. *Entropy* **2022**, *24*, 216. [CrossRef] [PubMed]
7. Boltzmann, L. *Lectures on Gas Theory*; Dover Publications: New York, NY, USA, 2012.
8. Aylaj, B.; Bellomo, N.; Gibelli, L. *Crowd Dynamics by Kinetic Theory Modeling: Complexity, Modeling, Simulations, and Safety*; Springer: New York, NY, USA, 2020.
9. Bellomo, N.; Bellouquid, A.; Gibelli, L.; Outada, N. *A Quest Towards a Mathematical Theory of Living Systems*; Birkhäuser: Basel, Switzerland, 2017.
10. Bellomo, N.; Bertotti, M.L.; Delitala, M. From the kinetic theory of active particles to the modeling of social behaviors and politics. *Qual. Quant.* **2007**, *41*, 545–555. [CrossRef]
11. Bellomo, N.; Brezzi, F. Traffic, crowds and swarms. *Math. Model. Methods Appl. Sci.* **2008**, *18*, 1145–1148. [CrossRef]
12. Bellomo, N.; Degond, P.; Tadmor, E. (Eds.) *Active Particles, Volume 1: Advances in Theory, Models, and Applications*; Birkhäuser: Basel, Switzerland, 2017.
13. Bertotti, M.L.; Modanese, G. Economic inequality and mobility in kinetic models for social sciences. *Eur. Phys. J. Spec. Top.* **2019**, *225*, 1945–1958. [CrossRef]
14. Bianca, C.; Menale, M. Mathematical analysis of a thermostatted equation with a discrete real activity variable. *Mathematics* **2020**, *8*, 57. [CrossRef]
15. Carbonaro, B.; Menale, M. Dependence on the Initial Data for the Continuous Thermostatted Framework. *Mathematics* **2019**, *7*, 602. [CrossRef]
16. Chinesta, F.; Abisset-Chavanne, E. *A Journey Around the Different Scales Involved in the Description of Matter and Complex Systems: A Brief Overview with Special Emphasis on Kinetic Theory Approaches*; Springer: New York, NY, USA, 2017.
17. Marsan, G.A.; Bellomo, N.; Gibelli, L. Stochastic Evolving Differential Games Toward a Systems Theory of Behavioral Social Dynamics. *arXiv* **2015**, arXiv:1506.05699v2.
18. Bellomo, N.; Marsan, G.A.; Tosin, A. *Complex Systems and Society: Modeling and Simulation*; Springer: New York, NY, USA, 2006.
19. Carbonaro, B.; Vitale, F. Some remarks on vector Markov Chains and their applications to the description of many-particle systems. In *Stochastic Processes—Theoretical Advances and Applications in Complex Systems*; Kulasiri, D., Ed.; IntechOpen: London, UK, 2024.
20. Carbonaro, B.; Menale, M. Markov Chains and Kinetic Theory: A Possible Application to Socio-Economic Problems. *Mathematics* **2024**, *12*, 1571. [CrossRef]
21. Carbonaro, B. Modeling epidemics by means of the stochastic description of complex systems. *Comput. Math. Methods* **2021**, *3*, 1208–1220. [CrossRef]
22. Lawler, G.F. *Introduction to Stochastic Processes*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2006.
23. Rozanov, Y.A. *Probability Theory: A Concise Course*; Dover Publications: New York, NY, USA, 2019.
24. Gilch, L. *Markov Chains: An Introduction: Lecture Notes*; Independently Published: Chicago, IL, USA, 2022; ISBN 979-8358906143.
25. Norris, J.R. *Markov Chains*; Cambridge University Press: Cambridge, UK, 1998.
26. Crisanti, A.; Paladin, G.; Vulpiani, A. *Producta of Random Matrices in Statistical Mechanics*; Springer: Berlin, Germany, 1993.
27. Benaim, M.; Hurth, T. *Markov Chains on Metric Spaces: A Short Course*; Springer Nature: New York, NY, USA, 2022.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.