MULTIMODE TIME-MARKOV SYSTEMS: RECURSIVE TENSOR-BASED ANALYSIS, CHAOTIC GENERATION, LOCALLY LOOPING PROCESSES

G. MAZZINI, G. SETTI∗ and R. ROVATTI
Dipartimento di Ingegneria, University of Ferrara, Via Saragat 1, 44100, Ferrara, Italy
∗gsetti@ing.unife.it

Received September 29, 2004; Revised May 24, 2005

The paper proposes a systematic solution to the problem of mixing different stochastic processes, each implied by a certain mode of operation of the system at hand and with a random duration whose distribution depends on the previous and present modes. We do so by widening the scope of an existing framework for the statistical characterization of finite valued processes with memory-one properties. The point of view is that of stochastic dynamics and the state space of the process is partitioned into regions (that we identify with modes) such that, if sojourn in a mode can be assumed, the statistical characterization is fully understood. The process is also allowed to stochastically move from one mode to another and the number of time steps for which it remains in each mode is a random variable whose distribution is a function only of the mode visited before. A general theoretical framework is developed here for the computation of any-order joint probabilities. The framework is then exemplified for the case of locally looping systems that are random sequences of modes comprising the cyclical execution of given atomic actions. They are the model of choice for complex appliances that operate following the steps of a communication protocol, and/or the various phases of a bus cycle, and/or the load-compute-store mechanism of a microprocessor, etc. Exploiting the theory put forward by the paper, we highlight how these processes could be generated by suitably designed 2-d chaotic maps and how their second- and third-order spectra may be obtained and interpreted when exponentially or polynomially decaying distributions are assumed for mode sojourn times.

Keywords: Markov process; chaos; tunable statistical process; z-transform; hierarchical model of complexity.

1. Introduction

Progress in the implementation and deployment of large scale dynamic ICT systems in various areas of technology has resulted in larger and more interdependent systems whose behavior is ever more difficult to predict, test, control, and maintain as the information flow between system components grows exponentially with system size. In a societal and business context, new means of communication that have dramatically enhanced information flow have allowed new forms of interaction and have led to macroscopic behavior that often contradicts the intentions of the individual actors. Everywhere, system complexity — the high level of interdependence between often very heterogeneous system components — becomes an obstacle to further progress: while computational performance is growing, mastering system complexity remains a major hurdle that threatens to substantially slow down the information revolution.

Assume, for example, to consider a simple node in a network that features at least basic computing
and communication abilities and must obey constraints on the availability of certain resources. In performing its tasks, the node responds both to external stimuli, internal programming and previously produced data. The complexity of these interactions forces the modeling of computing and communication activities in terms of stochastic processes.

Such a node may be either a unit in a distributed sensor network (see e.g. [Akyildiz et al., 2002]), or a device in a modern computing appliance [Intille, 2002], or a piece of on-board electronics in an advanced automotive system [Leen et al., 2002], or the control unit in a networked domotic application [Want et al., 2002].

As an example, assume that we are interested in modeling and/or optimizing energy consumption and that the node may enter a standby mode where only the basic ability of responding to interrupts and passing to other modes is retained.

Assume that, for the other two modes of operation (communication or computation) the stochastic process of power consumption is known.

The amount of time spent in each operating mode is also a random variable. Since the abstract functionality of the node surely requires a coordinated sequence of basic activities, time spent in each of them depends, at least, on the task performed immediately before the current one. As an example, the time spent in standby mode after a communication activity may in general be different from the one after a computation activity.

To ascertain the stochastic process of energy consumption due to the overall node operation, we have to mix the processes of energy consumption in each of the three working modes taking into account the amount of time spent in each of them, that, in turn, depends on the sequence with which they are performed. To further complicate things, the complexity of the overall functionality of the node may prevent deterministic modeling of the sequence of activities that, instead, may be captured by Markov chain (see e.g. [Ash & Doleans-Dade, 2000]).

Should we need a broader view of the system, the process of energy consumption could be considered at network level. At this level, individual contribution of single units must be merged considering that network-level task may force ensembles of nodes to work in different “meta-modes” at different times.

This mixing has the same structural features and problems we had to tackle at node level. Thus, any procedure to model the process of energy consumption at node level can be recursively applied to all higher levels providing it is suitably systematized.

The aim of this paper is to try a systematic solution to the problem of mixing different stochastic processes, each emitted in correspondence to a certain mode of operation and for a time duration that randomly depends on the previous mode.

We will do so limited to finite-state stochastic processes and based on some past work of which we here give a brief historical review. In some previous papers, stimulated by the fertile applicative ground that the statistical approach to the study of chaotic systems has been finding in recent years, we have tried to contribute to the general theoretical framework initiated by works like [Kohda & Tsuneda, 1994, 1997a, 1997b; Götz & Schwarz, 1997] and [Götz et al., 1998].

In particular, we concentrated on the set up of a general method based on tensor algebra [Rovatti et al., 2000; Setti et al., 2002] allowing the closed form expression of higher-order statistics of processes generated by quantizing a chaotic dynamics.

The key idea on which the use of multilinear algebra relies, is that quantized observables project the multidimensional version [Rovatti et al., 2000] of the classical Perron–Frobenius operator [Lasota & Mackay, 1994] on a finite dimensional space in which tracking the statistical evolution is equivalent to the construction of a multi-index quantity by means of suitably defined tensor products of few basic building blocks.

An extremely positive side-effect of the use of tensor algebra is that the construction of the complicated expressions related to higher-order probabilities can be made systematic employing only two product operators. This prevents multiple sums over many indexes to make the computation of these quantities extremely cumbersome and error-prone. Moreover, the structure of the product chains strongly reflects the stochastic evolution of the underlying process and helps understanding it.

This “algebra” of stochastic evolution was first applied to Piecewise Affine Markov Maps that, when properly quantized, are well known to be equivalent to finite-state Markov chains. In such an elementary form, the tools have been exploited to characterize [Mazzini et al., 2000] and optimize [Rovatti et al., 2001] the performance of some chaos-based spread-spectrum communication systems.
To cope with more complex phenomena like self-similarity, 1/f noise, and intermittency, the same theoretical setting has been widened [Rovatti & Mazzini, 2002a, 2002b; Rovatti & Mazzini, 2002] by allowing the granularity of the quantization, and thus the dimensionality of projection of the Perron–Frobenius operator, to grow to countable infinity.

The strategy was to assume that the countable parts of the behavior are actually “contained” in a single macro-state and to model the transitions between these macro-states.

With this, pseudo-Markov systems are conceptually linked to denumerable Markov chains (see e.g. [Kemeny et al., 1976]) for the countability of their fine-grain structure, and with Hidden Markov Models (see e.g. [Baum & Petrie, 1966; Ferguson, 1980]) for the fact that macro-states possibly hide a complex internal behavior. They are also a slight generalization of semi-Markov processes (see e.g. [Cinlar, 1975]). In fact, they allow the sojourn time in each mode to be distributed with laws different from the exponential one and, in addition, allow such a distribution to be linked not only to the mode itself but also with the mode visited previously.

From this point of view, the memory-one property of Markov systems is brought from state-transitions to the higher level of mode sojourn-times. This is why we indicate these systems as “time Markov”.

In this form, these mathematical tools have been exploited to address problems in modeling of network traffic and design of traffic emulator based on chaotic systems [Mondragon et al., 2000; Giovanardi et al., 2001].

What we aim in this paper is a step further in widening the applicability of this kind of theoretical machinery.

In particular, we will assume that the state space of the stochastic process we want to model can be partitioned into regions (that we identify with modes) and that, if we know that the process is confined in that mode, its statistical characterization is fully understood.

The process is also allowed to stochastically move from one mode to another and the number of time steps for which it remains in each mode is a random variable whose distribution is a function of the current mode and of the mode visited before.

With this, we aim at obtaining a two-fold generalization.

- First, we relax the assumption that the behavior of the process is seen as a sequence of jumps between single states in which the sojourn-time is arbitrarily distributed. We accept jumps between modes in which the evolution can be much more variegated than simply “emitting the same value”. As we will see in the final example, this can be used, for example, to construct and model process characterized by long range dependency not only in the low-pass domain (the classical 1/fα power spectrum trend) but also in more articulated behaviors such as local cycles that are followed for a random amount of time. As an example, this may be the model for a transmission scheme employing packets with variable length fields each of them modulated with a different number of levels to provide adjustable data robustness as it is quite common in modern communication standards (see e.g. [IEEE80211g]).

- Second, this very property implies that the methodology developed in the paper can be taken as the elementary step of a recursive modeling procedure. In other words, either the stochastic characterization of each mode can be further pursued in terms of the behaviors in smaller sub-modes, or, we may model a higher level behavior stemming from a sequence of super-modes, that are nothing than whole systems already characterized. We hope that, once developed beyond the preliminary stage of this paper, such a recursive partitioning may be of help in managing the trade off between mathematical exactness and intuitive feeling that is often well addressed by hierarchical methods.

As a coarse view on the paper, please note that the following Sec. 2 gives details on its organization and a list of the main results. Moreover, at the end of the development of the general theory, Sec. 7 offers a list of the symbols associated to the main quantities entailed in the derivations and a graph of their dependencies.

2. Paper Organization and Main Results

The paper is organized as follows. In Sec. 3, the concept of time-Markov systems is formally defined as well as the key tensor quantity \( \mathcal{L} \) accounting for the inter-mode behavior. In Sec. 4, the elementary step of a possible recursive statistical analysis is developed by defining the interaction between \( \mathcal{L} \) and the tensor quantities \( \mathcal{H} \), which accounts for the
intra-mode dynamics. In this section we define the building blocks needed to give a closed form expression to joint probabilities of any order. In Sec. 5, these building blocks are composed in the time domain to yield the generic expression for the third-order joint probability. In Sec. 6, the regularity of the procedure followed in the previous section in constructing the third-order joint-probability is generalized to any-order joint probabilities by means of extensive use of multidimensional z-transformation. Section 7 is devoted to an operative summary of the theory developed so far. The main quantities are listed and briefly explained. A graph is also given showing the path to follow to obtain the complete statistical characterization starting from a separate inter- and intra-mode characterization. In Sec. 8, the link between a quite wide subclass of time-Markov systems and two-dimensional chaotic maps is formally established. In Sec. 9 the concept and model of local looping systems are introduced and analytically developed following the above theory. The special case of a two-mode local looping system is addressed in Sec. 10 and its chaotic “implementation” is given in Sec. 11. Section 12 finally reports some numerical examples of the computation of second- and third-order statistical features of the resulting processes. Some conclusions are finally drawn.

The main theoretical results of the paper are related to the characterization and generation of multimode time-Markov systems as defined in Sec. 3. They are

- The expressions for the average sojourn time in every mode \( T_p \) (7), for the probability of having a mode transition \( \psi \) (9), and for the asymptotic probabilities of observing the system in any mode-state pair \( \pi_{r,s} \) (18) (Sec. 4).
- A systematic procedure to construct a tensor accounting for the joint-probability that the systems are observed in a sequence of mode-state pairs at instants that are separated by certain time lags (Sec. 6).
- A systematic procedure to construct a 2-D chaotic map whose iteration produces a stochastic process that is indistinguishable from that of a given multimode time-Markov process (28)/(29) (Sec. 8).

3. Time-Markov Systems

A discrete-time finite-state time-Markov system is a stochastic process with values \( x_k \) in the finite set \( S \). The state set is partitioned into \( n \) modes \( M_1, \ldots, M_n \) such that \( M_r \) contains \( p_r \) states which will be named \( S_{r,0}, \ldots, S_{r,p_r-1} \).

The system is defined by the property

\[
\Pr\{S_{r_k,s_k} = x_k, x_{k+1}, \ldots, x_{k+\tau-1} \in M_r\}
\]

which, since we assume that the process is stationary, gives information that are independent of \( k \) providing that it is a transition instant.

Note that, in general, quantities related to transitions between states depend on pairs of indexes, and that, to ease reading of the corresponding expressions we separate subsequent mode-state pairs with semicolons.

\[
\mathcal{L}_{r_1,s_1,r_2,s_2}(\tau) = \begin{cases} \Pr\{S_{r_2,s_2} = x_k, x_{k+1,k+\tau-1} \in M_r \} & \text{if } r_1 \neq r_2 \\ 0 & \text{otherwise} \end{cases}
\]

Another aid in the comprehension of the dynamic modeled by various tensor quantities should be given by the proper use of , and – signs at the bottom of the symbol. In general

- a quantity marked as \( \mathcal{A} \) will refer to an immediate transition from a mode to another mode, a
possible chain of further mode transitions and an
indeterminate sojourn in the last mode;
• a quantity marked as $A$ will refer to an initial
observation in a mode, a sequence of possible
mode transition up to the last;
• a quantity marked as $\hat{A}$ will refer to both the
phenomena above.

4. Recursive Statistical Analysis
We now wish to indicate and handle joint proba-
bilities of time-Markov systems. The general set-
ing implies that $m - 1$ non-negative time lags
$\tau_1, \ldots, \tau_{m-1}$ are given and that the system
is observed at time steps $0, \tau_1, \tau_1 + \tau_2, \ldots, \sum_{i=1}^{m-1} \tau_i$.

A joint probability function is associated with
these observations, namely
\[
H_{r_1, s_1; \ldots; r_m, s_m}(\tau_1, \ldots, \tau_{m-1})
\triangleq \Pr\{x_0 = S_{r_1, s_1}, x_{\tau_1} = S_{r_2, s_2}, x_{\tau_1 + \tau_2} = S_{r_3, s_3}, \ldots, x_{\sum_{i=1}^{m-1} \tau_i} = S_{r_m, s_m}\} \quad (2)
\]

Note that, since the system is finite-state, the
quantity $H_{r_1, s_1; \ldots; r_m, s_m}(\tau_1, \ldots, \tau_{m-1})$ has the nat-
ural structure of an array with the $2m$ indexes
$r_1, s_1; \ldots; r_m, s_m$. We will call these multi-index
quantities tensors.

We aim at writing the tensor $H$ relying on the
knowledge of what happens (statistically) within
each mode.

\[
H_{\rho, \sigma; \rho_0, \sigma_0; \ldots; \rho_l, \sigma_l+1}(\theta_{0, l}) \triangleq \sum_{\sigma_0} L_{\rho, \sigma; \rho_0, \sigma_0} \left( \sum_{i=0}^{l} \theta_i \right) \hat{H}_{\rho_0, \sigma_0; \ldots; \rho_l, \sigma_l+1}(\theta_0 - 1, \theta_{1, l}) \quad \text{if } \theta_0 > 0 \text{ and } \theta_{1, l} \geq 0 \\
0 \quad \text{otherwise} \quad (3)
\]

4.1. Local dynamics
We will assume to know the tensor $H$ when the
state is known to belong to $S_{r_1, s_1}$ at a certain time
step and to remain within the $r_1$th mode for all the
subsequent time steps, namely
\[
\hat{H}_{r_1, s_1; \ldots; r_m, s_m}(\tau_{1, m-1}) \triangleq \hat{H}_{r_1, s_1; \ldots; r_m, s_m}(\tau_{1, m-1}) \bigg| x_k = S_{r_1, s_1}, x_{k+1+\sum_{i=0}^{m-1} \tau_i} \in \mathbb{M}_{r_1}
\]

Note that this implicitly defines a $2m$-index
quantity $\hat{H}_{r_1, s_1; \ldots; r_m, s_m}(\tau_{1, m-1})$ which is assumed to
be null whenever the mode indexes $r_1, \ldots, r_m$
are not equal.

We will also assume that the intra-mode dynami-
cs modeled by $\hat{H}$ is truly local in the sense that
• it is independent of the modes and states visited
previously;
• it is independent of the sojourn time which is
accounted for only by $L$.

4.2. Merging local and global
dynamics: A building block
Once that global and local behaviors of the system
are defined, their composition may be schematized
as in Fig. 1 where we show the coarse view of a
process whose state space is partitioned into $n = 4$
modes.

The “pineal gland” merging the local and global
statistical features of the process into an overall
characterization is the following quantity
\[
\hat{H}_{\rho, \sigma; \rho_0, \sigma_0; \ldots; \rho_l, \sigma_l+1}(\theta_{0, l}) \triangleq \sum_{\sigma_0} L_{\rho, \sigma; \rho_0, \sigma_0} \left( \sum_{i=0}^{l} \theta_i \right) \hat{H}_{\rho_0, \sigma_0; \ldots; \rho_l, \sigma_l+1}(\theta_0 - 1, \theta_{1, l}) \quad \text{if } \theta_0 > 0 \text{ and } \theta_{1, l} \geq 0 \\
0 \quad \text{otherwise} \quad (3)
\]

Note that this implicitly defines a $2(l + 2)$-index
quantity $\hat{H}_{\rho, \sigma; \rho_0, \sigma_0; \ldots; \rho_l, \sigma_l+1}$ which is assumed to
be null whenever the mode indexes $\rho_0, \ldots, \rho_l$ are not equal.

To interpret the definition in (3) note that, from our assumption on the local dynamics, we get
\[
\hat{H}_{\rho_0, \sigma_0; \ldots; \rho_l, \sigma_l+1}(\theta_0 - 1, \theta_{1, l}) = \hat{H}_{\rho_0, \sigma_0; \ldots; \rho_l+1, \sigma_l+1}(\theta_0 - 1, \theta_{1, l})
\]
\[
= \hat{H}_{\rho_0, \sigma_0; \ldots; \rho_l+1, \sigma_l+1}(\theta_0 - 1, \theta_{1, l})
\]
\[
\bigg| x_k = S_{\rho_0, \sigma_0}, x_{k+1+\sum_{i=0}^{l} \theta_i} \in \mathbb{M}_{\rho_0}
\]
\[
\bigg| x_k = S_{\rho_0, \sigma_0}, x_{k+1+\sum_{i=0}^{l} \theta_i} \in \mathbb{M}_{\rho_0}
\]
\[
\bigg| x_k = S_{\rho_0, \sigma_0}, x_{k+1+\sum_{i=0}^{l} \theta_i} \in \mathbb{M}_{\rho_0}
\]
\[
\bigg| x_k = S_{\rho_0, \sigma_0}, x_{k+1+\sum_{i=0}^{l} \theta_i} \in \mathbb{M}_{\rho_0}
\]
that yields for \( \rho \) steps from such a transition, then in \( H \) certain time, being observed in \( \theta \) (1) implies that \( H \) should be read as the probability of passing from the state from which the mode \( \rho \) was reached after a further \( \theta_0 = 1, \theta_{(1.1)} \) to obtain that the numerator in the expression of the first cancels the denominator in the second. What remains is a ratio of probabilities that can still be interpreted as a conditioned probability, i.e.

\[
H_{\rho, \sigma_0, \sigma_1; \ldots; \rho_0, \sigma_{i+1}}(\theta_{(0, l)1}) = \Pr\{x_{k, \theta_0} = 1, x_{k} \notin M_\rho\} \\
\times \Pr\{x_{k, \theta_0} = 1, x_{k} \notin M_\rho\} \\
\times \Pr\{x_{k, \theta_0} = 1, x_{k} \notin M_\rho\}
\]

that yields for \( \rho \neq \rho_0 \) while for \( \rho = \rho_0 \), definition (1) implies that \( H_{\rho, \sigma_0, \sigma_1; \ldots; \rho_0, \sigma_{i+1}}(\theta_{(0, l)1}) \neq 0 \).

In more intuitive terms, the above quantity can be read as the probability of passing from the state \( S_{\rho, \sigma} \) to a state within the \( \rho_0 \)th mode \( (\rho \neq \rho_0) \) at a certain time, being observed in \( S_{\rho_0, \sigma_1} \) after \( \theta_0 \) time steps from such a transition, then in \( S_{\rho_0, \sigma_1} \) after a further \( \theta_{i-1} \) time steps, and so on up to \( S_{\rho_0, \sigma_{i+1}} \) which is reached after a further \( \theta_l \) time steps and from which the mode \( \rho_0 \) is left, given that there has been a mode transition and that the state was previously observed in \( S_{\rho, \sigma} \).

Note that the assumption that \( \theta_0 \) counts starting from the last state \( (S_{\rho, \sigma}) \) of the previous mode implies the constraint \( \theta_0 > 0 \).

### 4.3. Merging local and global dynamics: Asymptotic probabilities

We are here concerned with the definition of three steady-state probabilities, namely

\[
\pi_{\rho, \sigma} \triangleq \Pr\{x_{k-1} = S_{\rho, \sigma}, x_k \notin M_\rho\} \\
\pi_{\rho, \sigma} \triangleq \Pr\{x_k = S_{\rho, \sigma}\} \\
\pi_{\rho} \triangleq \Pr\{x_k \in M_\rho\} = \sum_{\sigma} \pi_{\rho, \sigma}
\]

that are nothing but the probability of being observed in \( S_{\rho, \sigma} \) at a certain time and leave the \( \rho \)th mode immediately after \( (\pi_{\rho, \sigma}) \), the probability of being observed in \( S_{\rho, \sigma} \) at a certain time \( (\pi_{\rho, \sigma}) \), and the probability of being observed in \( M_\rho \) at a certain time \( (\pi_{\rho}) \).

We will begin by showing that \( \pi_{\rho, \sigma} \) satisfies an eigenrelation involving only the basic building block \( H_{\rho, \sigma_0, \sigma_1}(\theta_0) \). Under certain conditions deeply related to the ergodicity of the overall system, this eigenrelation makes \( \pi_{\rho, \sigma} \) defined up to a scaling factor \( \gamma \). To determine such a scaling factor, we can state some considerations on the average sojourn times. These considerations will be paired with an independent relation in the next subsection to derive \( \pi_{\rho, \sigma} \) and \( \pi_{\rho} \) from \( \pi_{\rho, \sigma} \).

Let us begin by considering the tensor

\[
N_{\rho, \sigma; \rho_0, \sigma_1} \triangleq \sum_{\theta_0=1}^{\infty} H_{\rho, \sigma; \rho_0, \sigma_1}(\theta_0) \\
= \Pr\{x_{k, \theta_0-1} = S_{\rho_0, \sigma_1}, x_{k} \notin M_\rho\} \\
= \Pr\{x_{k} \notin M_\rho\}
\]

that accounts for the probability of exiting mode \( M_\rho \) from \( S_{\rho_0, \sigma_1} \) given that the same mode was entered some time before from \( S_{\rho, \sigma} \).

Hence, \( N_{\rho, \sigma; \rho_0, \sigma_1} \) plays the role of a “transition matrix” for the probabilities \( \pi_{\rho, \sigma} \). We may then resort to de-conditioning and marginalization over \( \rho \) and \( \sigma \) to obtain the eigenrelation

\[
\pi_{\rho_0, \sigma_1} = \sum_{\rho, \sigma} \pi_{\rho, \sigma} N_{\rho, \sigma; \rho_0, \sigma_1}
\]
i.e. \( \pi_{\rho,\sigma} \) must be the eigenmatrix corresponding to the unit eigenvalue of \( N_{\rho,\sigma,\rho_0,\sigma_1} \).

The existence of \( \pi_{\rho,\sigma} \) is a constraint on the \( \mathcal{H} \) and \( \mathcal{L} \) pair. In particular, since \( N_{\rho,\sigma,\rho_0,\sigma_1} \) is a non-negative tensor, it features a unit eigenvalue and an unique corresponding eigenmatrix \( \pi_{\rho,\sigma} \) if

- The whole system is ergodic, i.e. if the tensor is irreducible [Minc, 1987]. Note that this tensor compounds the effect of both local (\( \mathcal{H} \)) and global (\( \mathcal{L} \)) dynamics.
- The row sums are unitary, i.e.

\[
\sum_{\rho_0,\sigma_1} N_{\rho,\sigma,\rho_0,\sigma_1} = 1
\]

for every \( \rho \) and \( \sigma \).

In the following we will assume that both conditions are met.

Note now that the entries of the matrix \( \pi_{\rho,\sigma} \) may be summed up to yield the probability that at any given time a mode exists from which the state exits at any given i.e.

\[
\sum_{\rho,\sigma} \pi_{\rho,\sigma} = \Pr\{\exists \rho \text{ s.t. } x_{k-1} \in M_\rho, x_k \notin M_\rho\} = \psi
\]

where the probability \( \psi \) of having a transition between two modes remains implicitly defined.

Since there is one mode transition at the end of every sojourn time, \( \psi \) is the inverse of the average sojourn time.

To compute such an average note that

\[
\sum_{\rho,\sigma,\sigma_1} \pi_{\rho,\sigma} \mathcal{H}_{\rho,\sigma;\rho_0,\sigma_1}(\theta_0) = \Pr\{x_{(k+1+k+\theta_0)} \in M_{\rho_0}, x_{k+\theta_0} \notin M_{\rho_0}, x_{k-1} \notin M_{\rho_0}\}
\]

To make the above expression yield the probability that \( \theta_0 \) is a sojourn time in \( M_{\rho_0} \), we must condition it with the probability of the initial transition needed to arrive at \( M_{\rho_0} \). Such a probability can be obtained by marginalizing the above with respect to \( \theta_0 \), i.e. computing

\[
\sum_{\theta_0=1}^{\infty} \sum_{\rho,\sigma,\sigma_1} \pi_{\rho,\sigma} \mathcal{H}_{\rho,\sigma;\rho_0,\sigma_1}(\theta_0) = \sum_{\rho,\sigma,\sigma_1} \pi_{\rho,\sigma} N_{\rho,\sigma,\rho_0,\sigma_1}
\]

where we have used the fact that \( \pi_{\rho,\sigma} \) is an eigenmatrix of \( N_{\rho,\sigma,\rho_0,\sigma_1} \).

With this we have that

\[
\sum_{\rho,\sigma,\sigma_1} \pi_{\rho,\sigma} \mathcal{H}_{\rho,\sigma;\rho_0,\sigma_1}(\theta_0)
\]

\[
= \Pr\{x_{(k+1+k+\theta_0)} \in M_{\rho_0}, x_{k+\theta_0} \notin M_{\rho_0}, x_{k-1} \notin M_{\rho_0}\}
\]

is the probability that \( \theta_0 \) is a sojourn time in \( M_{\rho_0} \).

If we agree to indicate by \( T_{\rho_0} \) the average of such a sojourn time we have from (6)

\[
T_{\rho_0} = \sum_{\theta_0=1}^{\infty} \sum_{\rho,\sigma,\sigma_1} \pi_{\rho,\sigma} \mathcal{H}_{\rho,\sigma;\rho_0,\sigma_1}(\theta_0)
\]

The global average sojourn time in any mode is a weighted sum of the average sojourn times in each mode. In such a sum, \( T_{\rho_0} \) must be weighted by the relative frequency of \( M_{\rho_0} \) with respect to all the possible modes.

Let us indicate by \( \xi_{\rho_0} \) such a relative frequency and consider the matrix

\[
\mathcal{M}_{\rho,\rho_0} \triangleq \sum_{\theta_0=1}^{\infty} \sum_{\rho,\sigma,\sigma_1} \pi_{\rho,\sigma} \mathcal{L}_{\rho,\sigma;\rho_0,\sigma_1}(\theta_0)
\]

\[
= \Pr\{\exists \theta_0 \text{ s.t. } x_{(k+1+k+\theta_0)} \in M_{\rho_0}, x_{k+\theta_0} \notin M_{\rho}, x_{k-1} \notin M_{\rho}\}
\]

where we have exploited the definition of \( \pi_{\rho,\sigma} \) to strip the conditional part of \( \mathcal{L}_{\rho,\sigma;\rho_0,\sigma_1} \), marginalize on \( \sigma \) and \( \sigma_1 \) and then revert to a conditional probability dividing by \( \Pr\{x_k \notin M_{\rho}, x_{k-1} \in M_{\rho}\} = \sum_{\rho,\sigma,\sigma_1} \pi_{\rho,\sigma} \).

Since any mode will be eventually left, we have \( \xi_{\rho_0} = \sum_{\rho} \mathcal{M}_{\rho,\rho_0} \xi_{\rho}, \) i.e. \( \xi_{\rho_0} \) is the eigenvector of \( \mathcal{M} \) corresponding to the unit eigenvalue. The eigenvector must be obviously normalized to feature \( \sum_{\rho_0} \xi_{\rho_0} = 1 \). Once that such an eigenvector could be computed we set

\[
\psi = \frac{1}{\sum_{\rho_0} \xi_{\rho_0} T_{\rho_0}}
\]
With this, we may assume that a generic eigen-
matrix \( \pi'_{\rho, \sigma} \) of \( N_{\rho, \sigma; \rho_0, \sigma_0} \) is known and that \( \pi'_{\rho, \sigma} = \gamma \pi'_{\rho, \sigma} \) for some \( \gamma \).

Note now that (7) and (8) indicate that \( T_{\rho_0} \) and \( M \) are independent of \( \gamma \) and can therefore be computed based only on \( \pi'_{\rho, \sigma} \).

With this we know that \( \gamma \) appears only once in the substitution of (9) into (5), i.e.
\[
\gamma \sum_{\rho_0} \pi'_{\rho, \sigma} = \frac{1}{\sum_{\rho_0} \xi_{\rho_0} T_{\rho_0}} \tag{10}
\]
that can be straightforwardly solved to yield its value.

4.4. Merging local and global dynamics: A coarser view

To proceed further assume that we are interested
in the quantity \( \mathcal{H}_{\rho_0, \sigma_0; \rho_1: \ldots: \rho_{l-1}, \sigma_{l+1}}(\theta_{\{1,l\}}) \) that accounts for the probability of being observed in \( S_{\rho_0, \sigma_1} \) at a certain time and then in \( S_{\rho_0, \sigma_1} \) after a further \( \theta_{l-1} \) times steps up to \( S_{\rho_0, \sigma_{l+1}} \) from which the mode \( \rho_0 \) is left, i.e.
\[
\mathcal{H}_{\rho_0, \sigma_0; \rho_1: \ldots: \rho_{l-1}, \sigma_{l+1}}(\theta_{\{1,l\}}) = \begin{cases} \text{Pr}\{x_{k+k-1+i} = S_{\rho_0, \sigma_{j+1}} \mid j = 0, \ldots, l, \theta_{l-1} \} & \text{if } \theta_{l-1} \geq 0 \tag{11} \\
\end{cases}
\]
that implicitly defines a \( 2(l + 1) \)-index quantity \( \mathcal{H}_{\rho_0, \sigma_0; \rho_1: \ldots: \rho_{l-1}, \sigma_{l+1}}(\theta_{\{1,l\}}) \) which is assumed to be null whenever the mode indexes \( \rho_0, \ldots, \rho_l \) are not equal.

To express the above quantity, we begin considering the time step at which the mode \( \rho_0 \) is entered
\[
\mathcal{H}_{\rho_0, \sigma_1: \ldots: \rho_{l-1}, \sigma_{l+1}}(\theta_{\{1,l\}}) = \begin{cases} \sum_{\rho_0 \sigma} \pi_{\rho, \sigma} \sum_{\theta_{l-1}} \mathcal{H}_{\rho_0, \sigma_0; \rho_1: \ldots: \rho_{l-1}, \sigma_{l+1}}(\theta_{\{0,l\}}) & \text{if } \theta_{l-1} \geq 0 \tag{13} \\
\end{cases}
\]
otherwise

Let us now consider the quantity \( \mathcal{H}_{\rho_0, \sigma_0; \rho_1: \ldots: \rho_{l-1}, \sigma_{l+1}}(\theta_{\{0,l-1\}}) \) that accounts for the probability of passing from the state \( S_{\rho, \sigma} \) to a state
within the \( \rho_0 \)th mode \( (\rho \neq \rho_0) \) at a certain time, being observed in \( S_{\rho_0, \sigma_1} \) after \( \theta_0 \) time steps, and finally in \( S_{\rho_0, \sigma_{l+1}} \) after a further \( \theta_{l-1} \) time steps, given that there has been a mode transition and
that the state was previously observed in \( S_{\rho_0, \sigma} \), i.e.
\[
\mathcal{H}_{\rho_0, \sigma_0; \rho_1: \ldots: \rho_{l-1}, \sigma_{l+1}}(\theta_{\{0,l-1\}}) = \begin{cases} \text{Pr}\{x_{k+k-1+i} = S_{\rho_0, \sigma_{j+1}} \mid j = 0, \ldots, l-1, \theta_{l-1} \} & \text{if } \theta_{l-1} \geq 0 \tag{14} \\
\end{cases}
\]
that implicitly defines a \( 2(l + 1) \)-index quantity \( \mathcal{H}_{\rho_0, \sigma_0; \rho_1: \ldots: \rho_{l-1}, \sigma_{l+1}}(\theta_{\{0,l-1\}}) \) that is assumed to be
null whenever the mode indexes $\rho_0, \ldots, \rho_{l-1}$ are not equal or when $\rho = \rho_0$.

To compute it we begin considering the state $S_{\rho_0, \sigma_{l+1}}$ from which the mode $\rho_0$ is exited $\theta_l$ time steps after visiting the state $S_{\rho_0, \sigma_l}$.

With this we have that

$$
\sum_{\sigma_{l+1}} \mathcal{H}_{\rho_0, \sigma_0; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_{l+1}}(\theta_0, \theta_l)
$$

is the probability of entering $M_{\rho_0}$ at a certain time step and being observed in $S_{\rho_0, \sigma_l}$ after $\theta_0$, in $S_{\rho_0, \sigma_2}$ after $\theta_1$ and so on up to $S_{\rho_0, \sigma_l}$ from which the mode has not necessarily left.

To match the probabilistic definition of $\mathcal{H}_{\rho_0, \sigma_0; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_{l+1}}(\theta_0, \theta_l)$ it is enough that a $\theta_l \geq 0$ exists such that all of the above happens.

Hence, we may get rid of the dependency on $\theta_l$ by summing over all its possible values and get

$$
\mathcal{H}_{\rho_0, \sigma_0; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_l}(\theta_0, \theta_{l-1}) = \left\{ \begin{array}{ll}
\sum_{\sigma_{l+1}} \mathcal{H}_{\rho_0, \sigma_0; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_{l+1}}(\theta_0, \theta_l) & \text{if } \theta_0 > 0 \text{ and } \theta_{(l-1)} \geq 0 \\
0 & \text{otherwise}
\end{array} \right. 
$$

As a final quantity, let us consider $\mathcal{H}_{\rho_0, \sigma_0; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_l}(\theta_0, \theta_{l-1})$ defined to account for the probability of being observed in $S_{\rho_0, \sigma_l}$ at a certain time and then in $S_{\rho_0, \sigma_l}$ after a further $\theta_{l-1}$ steps up to $S_{\rho_0, \sigma_l}$ from which the mode has not necessarily left, i.e.

$$
\mathcal{H}_{\rho_0, \sigma_0; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_l}(\theta_0, \theta_{l-1}) = \text{Pr}\{x_{k+\sum_{i=1}^{l-1} \theta_i} \in M_{\rho_0}; x_k = S_{\rho_0, \sigma_j+1}, j = 0, \ldots, l-1\} 
$$

that implicitly defines a 2l-index quantity $\mathcal{H}_{\rho_0, \sigma_0; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_l}(\theta_0, \theta_{l-1})$ that is assumed to be null whenever the mode indexes $\rho_0, \ldots, \rho_{l-1}$ are not equal.

To express such a quantity recall (12) and its significance to recognize that what is needed is a marginalization over $\rho$, $\sigma$, $\sigma_{l+1}$ and a sum over $\theta_{l+1}$ to leave the time of exit from $M_{\rho_0}$ undefined. With this

$$
\mathcal{H}_{\rho_0, \sigma_0; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_l}(\theta_0, \theta_{l-1}) = \left\{ \begin{array}{ll}
\sum_{\theta_0=1}^{\infty} \sum_{\theta_1=0}^{\infty} \sum_{\rho, \sigma, \sigma_{l+1}} \pi_{\rho_0, \sigma_0} \mathcal{H}_{\rho, \sigma; \rho_0; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_{l+1}}(\theta_0, \theta_l) & \text{if } \theta_{(l-1)} \geq 0 \\
0 & \text{otherwise}
\end{array} \right. 
$$

As a general remark, note that from the above definitions we conclude that whenever a quantity has a $\bullet$ as the leftmost subscript, it vanishes if its first argument is 0 or less and if the other arguments are negative, while quantities with a $-$ as the leftmost subscript vanish when any of its arguments is negative.

Note finally that the definition of $\mathcal{H}_{\rho_0, \sigma_0; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_l}(\theta_0, \theta_{l-1})$ degenerates for $l = 1$ to the probability of revealing the system in the generic state $S_{\rho_0, \sigma_0}$. If we indicate with $\pi_{\rho_0, \sigma_0}$ such an asymptotic probability we have

$$
\pi_{\rho_0, \sigma_0} = \mathcal{H}_{\rho_0, \sigma_0}
= \sum_{\theta_0=1}^{\infty} \sum_{\theta_1=0}^{\infty} \sum_{\rho, \sigma, \sigma_2} \pi_{\rho, \sigma} \mathcal{H}_{\rho, \sigma; \rho_0; \rho_0, \sigma_1; \rho_0, \sigma_2}(\theta_0, \theta_1)
$$

as well as

$$
\pi_{\rho_0} = \sum_{\theta_0=1}^{\infty} \sum_{\theta_1=0}^{\infty} \sum_{\rho, \sigma, \sigma_1} \pi_{\rho, \sigma} \mathcal{H}_{\rho, \sigma; \rho_0; \rho_0, \sigma_1; \rho_0, \sigma_2}(\theta_0, \theta_1)
$$

5. Third-Order Joint Probability

When we want to observe the system at three time instants we come up with a third-order joint-probability depending on two time lags $\tau_1$ and $\tau_2$, i.e. $\mathcal{H}_{\tau_1, \tau_2; \tau_2, \tau_3, \tau_2}(\tau_1, \tau_2)$. Hence, we should take into account four different cases. In fact we may have mode transitions during both time lags, only during the first time lag but not during the second, only during the second time lag but not during the first and no mode transition at all. Each of these cases deserves a separate treatment.

We start by writing the expression of the corresponding joint-probability reasoning in the time domain where the transition mechanisms can be followed more easily.

5.1. Time-domain expression

To develop the calculations we will need to indicate a variable number of modes and time durations. In
particular, we will indicate with $\Delta_i$ the number of mode transitions between the $(i-1)$th and the $i$th observation instants (that are $\tau_i$ time steps apart). Assuming that $\Delta_i > 0$ we will indicate

- with $t'_i \geq 0$ the amount of time spent in the mode of the $(i-1)$th observation after the observation itself;
- with $t''_i > 0$ the time spent in the $j$th mode visited during the mode transitions happening between the two observations;
- with $a_i^{j}, b^{j}_i$ the state in which the trajectory was before the $j$th transition happening between the two observations;
- with $t''_i > 0$ the amount of time elapsed between the entrance in the mode of the next observation and the observation itself.

Note that, with these symbols $t'_i + \sum_{j=1}^{\Delta_i-1} t''_i + t''_i = \tau_i$. We will also make extensive use of the discrete delta symbol $\delta_a$ which is nonvanishing and equal to 1 only when $a = 0$.

The cases mentioned above can now be distinguished by means of the vanishing of $\Delta_1$ and $\Delta_2$.

- **No mode transitions ($\Delta_1 = \Delta_2 = 0$)**

  The probability of observing the system in $S_{r_1,s_1}$ at time 0, in $S_{r_2,s_2}$ at time $\tau_1$ and in $S_{r_3,s_3}$ at time $\tau_1 + \tau_2$ is nonvanishing only when $r_1 = r_2 = r_3$ and thus this case contributes to the overall joint-probability $\mathcal{H}_{r_1,s_1;r_2,s_2;r_3,s_3}((\tau_1,\tau_2)$ with the term

\[
\mathcal{H}_{r_1,s_1;r_2,s_2;r_3,s_3}((\tau_1,\tau_2)
\]

- **Mode transitions during $\tau_1$ but not during $\tau_2$ ($\Delta_1 > 0$, $\Delta_2 = 0$)**

  A different term must be considered for any possible number of mode transition during $\tau_1$. Yet, the joint probability in $\mathcal{H}((\tau_1,\tau_2)$ does not vanish only if the system remains in the mode it was in at time step $\tau_1$ for all the subsequent time lag $\tau_2$. Hence, for any given $\Delta_1$ we have

\[
\sum_{t'_1 \geq 0, t''_1 > 0} \sum_{a_1^1, a_1^2, b_1^1, b_1^2} \left\{ \mathcal{H}_{r_1,s_1;r_2,s_2;\Delta_1} (t'_1) \times \mathcal{L}_{a_1^1,a_1^2,b_1^1,b_1^2} (t''_1) \times \mathcal{H}_{r_2,s_2;r_3,s_3} (t''_1) \right\}
\]

Note that this symbol $t'_i + \sum_{j=1}^{\Delta_i-1} t''_i + t''_i = \tau_i$.

Each product in the above multiple sum is made of three recognizable blocks, namely

1. $\mathcal{H}_{r_1,s_1;a_1^1,b_1^1} (t'_1)$ which accounts for the observation in mode $r_1$ and the sojourn in the same mode up to the first mode transition that must happen during $\tau_1$.

2. $\mathcal{L}_{a_1^1,a_1^2,b_1^1,b_1^2} (t''_1) \times \cdots \times \mathcal{L}_{a_1^{\Delta_1-1},b_1^{\Delta_1-1};a_1^1,b_1^1} (t^{\Delta_1-1})$

   which accounts for the first $\Delta_1 - 1$ mode transitions during $\tau_1$.

3. $\mathcal{H}_{r_2,s_2;r_3,s_3} (t''_1)$ which accounts for the $\Delta_2$th mode transitions during $\tau_1$ and for the observations in the new and last mode $r_2 = r_3$.

- **Mode transitions during $\tau_2$ but not during $\tau_1$ ($\Delta_1 = 0$, $\Delta_2 > 0$)**

  This is the opposite situation. In order to make the joint probability nonvanishing, $r_1$ and $r_2$ must coincide and the system must stay in the initial state for all $\tau_1$. For any given $\Delta_2$, this case contributes to the final $\mathcal{H}((\tau_1,\tau_2)$ with

\[
\sum_{t''_1 > 0} \sum_{a_1^1,a_1^2,b_1^1,b_1^2} \left\{ \mathcal{H}_{r_1,s_1;r_2,s_2;\Delta_2} (\tau_1,t'_2) \times \mathcal{L}_{a_1^1,a_1^2,b_1^1,b_1^2} (t'_2) \times \cdots \times \mathcal{L}_{a_1^{\Delta_2-1},b_1^{\Delta_2-1};a_1^1,b_1^1} (t^{\Delta_2-1}) \times \mathcal{H}_{r_2,s_2;r_3,s_3} (t'_2) \right\}
\]

That must be summed up for all the values of $\Delta_2 = 1, 2, \ldots$. 

\[
\sum_{t''_1 > 0} \sum_{a_1^1,a_1^2,b_1^1,b_1^2} \left\{ \mathcal{H}_{r_1,s_1;r_2,s_2;\Delta_2} (\tau_1,t'_2) \times \mathcal{L}_{a_1^1,a_1^2,b_1^1,b_1^2} (t'_2) \times \cdots \times \mathcal{L}_{a_1^{\Delta_2-1},b_1^{\Delta_2-1};a_1^1,b_1^1} (t^{\Delta_2-1}) \times \mathcal{H}_{r_2,s_2;r_3,s_3} (t'_2) \right\}
\]
Each product in the above multiple sum is made of three recognizable blocks, namely

1. \( \mathcal{H}_{r_1,s_1:t_2} (t'_1) \) which accounts for the observation in mode \( r_1 \) and the sojourn in the same mode up to the first mode transition that must happen during \( r_2 \).

2. \( \mathcal{L}_{a_2} (t'_2) \) which accounts for the first \( \Delta_2 - 1 \) mode transitions during \( r_2 \).

3. \( \mathcal{H}_{a_2} (t''_2) \) which accounts for the first of the \( \Delta_2 \) mode transitions during \( r_2 \) and for the observation in the new and last mode \( r_3 \).

- **Mode transitions during both \( r_1 \) and \( r_2 \)**

For any given \( \Delta_1 \), \( \Delta_2 \) we obtain a contribution of the kind

\[
\begin{align*}
&\sum_{t'_1, t''_1 > 0} \sum_{t''_1 > 0} \mathcal{L}_{a_1} (t'_1) \times \cdots \times \mathcal{L}_{a_1} (t'_1) \times \mathcal{H}_{a_1} (t'_1, t''_2) \\
&\times \mathcal{L}_{a_2} (t''_2) \times \cdots \times \mathcal{L}_{a_2} (t''_2) \times \mathcal{H}_{a_2} (t''_2) \times \mathcal{H}_{a_r} (t''_3, \ldots, t''_{r_3}) \\
&\quad \times \mathcal{H}_{a_{r_3}} (t''_3, \ldots, t''_{r_3}) \\
&\quad \times \mathcal{H}_{a_{r_{r_3}}} (t''_{r_3}) \\
&\quad \times \mathcal{H}_{a_{r_{r_{r_3}}}} (t''_{r_{r_{r_3}}}) 
\end{align*}
\]

which accounts for the first \( \Delta_2 - 1 \) mode transitions during \( r_2 \).

Note how each product in the above cumbersome multiple sum is made of five easily recognizable blocks, namely

1. \( \mathcal{H}_{r_1,s_1:t_2} (t'_1) \) which accounts for the observation in mode \( r_1 \) and the sojourn in the same mode up to the first mode transition that must happen during \( r_1 \).

2. \( \mathcal{L}_{a_2} (t'_2) \) which accounts for the first \( \Delta_2 - 1 \) mode transitions during \( r_2 \).

3. \( \mathcal{H}_{a_2} (t''_2) \) which accounts for the first of the \( \Delta_2 \) mode transitions during \( r_2 \) and for the observation in the new and last mode \( r_3 \).

The complexity of writing the above expressions mainly derives from the proper arrangement of indexes and corresponding sums. Hence, the definition of suitable operators that work on multi-index quantities may greatly help to focus on the fundamental structure and significance.

We will use the identity tensor \( I \) as a shortcut for the constraint of having no mode transition by defining \( I_d = \delta_{d,c} \delta_{b,d} \).

The tensor product we need can be defined as follows.

**Definition 1.** Given the \( 2m^\prime \)-rank tensor of multivariate integer or complex functions \( A \) and the \( 2m^\prime \)-rank tensor of multivariate integer or complex functions \( B \) such that the range of the last two indexes of \( A \) is the same as the first two indexes of \( B \), their **product** is an \( 2(m^\prime + m^\prime - 2) \)-rank tensor of multivariate integer or complex functions \( C = A \cdot B \) such that

\[
C_{\rho,\sigma;\ldots;\rho_{m^\prime}+m^\prime-2;\sigma_{m^\prime}+m^\prime-2} = \sum_{\rho,\sigma} A_{\rho_1,\sigma_1;\ldots;\rho_{m^\prime-1};\sigma_{m^\prime-1};\rho;\sigma} \\
\times B_{\rho,\sigma;\ldots;\rho_{m^\prime}+m^\prime-2;\sigma_{m^\prime}+m^\prime-2}
\]

where \( \rho \) and \( \sigma \) are annihilated sweeping all their common ranges.

The product is a noncommutative, associative and bilinear operator. In the following we will sometimes indicate the product between two tensors neglecting the central, whenever the significance cannot be mistaken.
Definition 2. Given the $2m'$-rank tensor of functions of $v'$ integer variables $A$ and the $2m''$-rank tensor of functions of $v''$ integer variables $B$ such that the range of the last two indexes of $A$ is the same of the first two indexes of $B$, their star product is an $2(m'+m''-2)$-rank tensor of functions of $(v'+v''-1)$ integer variables $C = A \ast B$ such that

$$C_{\rho_1,\sigma_1;\ldots;\rho_{m'+m''-2},\sigma_{m'+m''-2}}(l'_1,v'_1), l''_{2,v''})$$

$$= \sum_{\rho,\sigma} \sum_{l=-\infty}^{\infty} A_{\rho_1,\sigma_1;\ldots;\rho_{m'-1},\sigma_{m'-1};\rho,\sigma}(l'_1,v'_1-1), l'_v - l)$$

$$\times B_{\rho,\sigma;\rho_{m'+1},\sigma_{m'+1};\ldots;\rho_{m'+m''-2},\sigma_{m'+m''-2}}(l,l''_{2,v''})$$

The star product is noncommutative, associative and bilinear operator.

For a square tensor of functions of one integer variable we also define $A^{*p} = {A} \ast {A} \ast \cdots \ast {A}$ and assume that $A^{*0}$ is such that $B \ast A^{*0} = A^{*0} \ast B = B$ for any compatible tensor $B$.

Note that the star product is nothing but the conventional row-by-column product of Definition 1 in which the product between entries is substituted by convolution along the last argument of the left-hand entry and the first argument of the right-hand entry.

With these definitions we have four contributions that can be rewritten as

- **No mode transitions**
  $$\mathcal{H}(\tau_1, \tau_2)$$

- **Mode transition during $\tau_1$ but not during $\tau_2$**
  $$\mathcal{H}(\tau_1) \ast \left[ \sum_{\Delta_1=0}^{\infty} L^{*\Delta_1}(\tau_1) \right] \ast \mathcal{H}(\tau_1, \tau_2)$$

- **Mode transitions during $\tau_2$ but not during $\tau_1$**
  $$\mathcal{H}(\tau_1, \tau_2) \ast \left[ \sum_{\Delta_2=0}^{\infty} L^{*\Delta_2}(\tau_2) \right] \ast \mathcal{H}(\tau_2)$$

- **Mode transitions during both $\tau_1$ and $\tau_2$**
  $$\mathcal{H}(\tau_1) \ast \left[ \sum_{\Delta_1=0}^{\infty} L^{*\Delta_1}(\tau_1) \right] \ast \mathcal{H}(\tau_1, \tau_2)$$
  $$\ast \left[ \sum_{\Delta_2=0}^{\infty} L^{*\Delta_2}(\tau_2) \right] \ast \mathcal{H}(\tau_2)$$

5.2. **$z$-Domain expression**

It is clear that in passing from third-order to fourth- and higher-order characterization the complexity of the resulting time-domain expressions rapidly increases to unmanageability due to multiple convolutions.

Yet, such a complexity is greatly reduced if we consider the multidimensional $z$-transform ($Z$) that can be defined for a generic multi-index as

$$\tilde{D}(z_{1,v}) = Z[D(l_{1,v})](z_{1,v})$$

$$= \sum_{l_{1,v}=0}^{\infty} D(l_{1,v}) \prod_{i=1}^{v} z_i^{-l_i}$$

As an example, multidimensional $z$-transforms have been already exploited for the statistical characterization of processes in [Krummenauer, 1998].

With this, we have that if $\tilde{C}(l'_1,v'_1), l''_{2,v''}) = A(l'_1,v'_1) \ast B(l''_{2,v''})$ then

$$\tilde{C}(z'_1,v'_1), z''_{2,v''}) = \tilde{A}(z'_1,v'_1) \cdot \tilde{B}(z''_{2,v''})$$

Note that the product in Definition 1 does not entail operations in the time-axis and thus simply propagates from the time-domain to the $z$-domain.

In the light of the above considerations we may, for example, write the complete two-dimensional $z$-transform $\tilde{H}(z_1, z_2)$ of the third-order joint-probability $\mathcal{H}(\tau_1, \tau_2)$ as

$$\tilde{H}(z_1, z_2)$$

$$= \tilde{H}(z_1, z_2) + \tilde{H}(z_1) \tilde{J}(z_1) \tilde{H}(z_2)$$

$$+ \tilde{H}(z_1, z_2) \tilde{J}(z_2) \tilde{H}(z_2)$$

$$+ \tilde{H}(z_1) \tilde{J}(z_1) \tilde{H}(z_1, z_2) \tilde{J}(z_2) \tilde{H}(z_2)$$

where the $z$-transform of the generic $\sum_{\Delta=0}^{\infty} L^{*\Delta}(\tau)$ is

$$\tilde{J}(z) = \sum_{\Delta=0}^{\infty} \tilde{L}^{\Delta}(z) = [I - \tilde{L}(z)]^{-1}$$

To interpret tensor inversion in the above formula note that each $L_{a,b,c,d}$ is the representation of the linear transformation that associates to each two-index quantity $A_{c,d}$ the two-index quantity $B_{a,b} = \sum_{c,d} L_{a,b,c,d} A_{c,d}$ that belongs to the same vector space. Hence, the inverse of the tensor is the representation of the inverse of the corresponding linear transformation (providing it is a bijection).
6. Any Order Joint-Probability in the $z$-Domain

Equation (24) has such a strong regularity that we may now leap forward and devise a general procedure to write the $z$-domain expression of an $m$th order joint-probability.

Since there are $m$ time instants at which the mode is observed, there are $m - 1$ time lags $\tau_1, \ldots, \tau_{m-1}$ so that the expression we aim at is the $(m - 1)$-dimensional $z$-transform $\mathcal{H}(z_1, \ldots, z_{m-1})$.

To proceed we define a string of binary flags $\mathbf{v}$ such as $v_i = 0$ if $\Delta_i = 0$ and $v_i = 1$ if $\Delta_i > 0$. The binary flag string $\mathbf{v}$ has $m - 1$ entries for a total of $2^{m-1}$ terms in the expression of the $m$th order $\mathcal{H}$.

Given a particular instance of $\mathbf{v}$ the expression corresponding to it can be constructed by following few simple rules that start from the decomposition of the augmented flag string $1\mathbf{v}$ into runs [Rovatti & Mazzini, 2002b].

We say that a run is a subsequence of the kind $10\ldots01$ with $l \geq 0$ being the length of the run. We say that $1\mathbf{v}$ is decomposed into runs if it is the concatenation of runs, the last entry of each run being the first entry of the following run. As an example, if $\mathbf{v} = 110$ we have $1\mathbf{v} = 11011$ which can be decomposed into the three runs 11, 11 and 101 while, if $\mathbf{v} = 0010$ we have that $1\mathbf{v} = 100101$ is decomposed into the two runs 1001 and 101.

For any possible instance of $\mathbf{v}$ we have $\Lambda \in \{1, \ldots, m - 1\}$ runs whose lengths will indicate by $l_1, \ldots, l_\Lambda$ and that are linked by the equality $\sum_{i=1}^\Lambda l_i = \Lambda = m + 1$. These numbers, along with the accumulations $q_i = \sum_{j=1}^{i-1} (l_j + 1)$ are all we need to build the term corresponding to that $\mathbf{v}$, provided we follow some simple rules.

- If $\Lambda = 1$ then we must have $l_1 = m$ and thus $\mathbf{v} = 0\ldots0$. The corresponding term is then $\mathcal{H}(z_1, m-1))$.

- If $\Lambda > 1$ then we may distinguish the first and the last run from the other possible intermediate runs. Contribution relative to each kind of run can be written depending on the following quantity.

  — The first run gives raise to the term $\mathcal{H}(z_1, l_1+1)$.

---

### Table 1. Construction of the $z$-transform of the fifth order joint-probability.

<table>
<thead>
<tr>
<th>$1\mathbf{v}$</th>
<th>$l_i$</th>
<th>$\mathcal{H}(z_1, z_2, z_3, z_4) =$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10001</td>
<td>4</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>11001</td>
<td>0, 3</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>11011</td>
<td>1, 2</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>11111</td>
<td>0, 0, 2</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>10111</td>
<td>2, 1</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>11011</td>
<td>0, 1, 1</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>10111</td>
<td>1, 0, 1</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>11111</td>
<td>0, 0, 0, 1</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>10011</td>
<td>3, 0</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>11011</td>
<td>0, 2, 0</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>10111</td>
<td>1, 1, 0</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>11111</td>
<td>0, 0, 1, 0</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>10011</td>
<td>2, 0, 0</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>11011</td>
<td>0, 1, 0, 0</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>10111</td>
<td>1, 0, 0, 0</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
<tr>
<td>11111</td>
<td>0, 0, 0, 0, 0</td>
<td>$\mathcal{H}(z_1, z_2, z_3, z_4) +$</td>
</tr>
</tbody>
</table>
— The intermediate runs, of lengths \( l_i \) for \( i = 2, \ldots, \Lambda - 1 \), exist only if \( \Lambda > 2 \) and give raise to terms of the kind
\[
\tilde{J}(z_{q_i}) \tilde{H}(z_{q_i + l_i + 1})
\]
— The last run gives raise to the term
\[
\tilde{J}(z_{q_\Lambda}) \tilde{H}(z_{q_\Lambda + m - 1})
\]
What we have to do is to generate all the \( 2^m \) possible strings \( \mathbf{v} \) and add up all the product chains deriving from each of them. As an example Table 1 shows the \( z \)-transform of a fifth order joint-probability

By means of (13)–(17) we finally have that \( \tilde{H} \), \( \tilde{H} \), and \( \tilde{H} \) can be expressed in terms of \( \tilde{H} \). More formally we have
\[
\tilde{H}_{\rho_0, \sigma_1; \ldots; \rho_0, \sigma_{l+1}}(z_{1l})
\]
\[
= \lim_{z_0^{-1}} \sum_{\rho, \sigma} \pi_{\rho, \sigma} \tilde{H}_{\rho, \sigma; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_{l+1}}(z_{0l})
\]
\[
\tilde{H}_{\rho, \sigma; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_{l+1}}(z_{0l-1})
\]
\[
= \lim_{z_l^{-1}} \sum_{\sigma_{l+1}} \tilde{H}_{\rho, \sigma; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_{l+1}}(z_{0l})
\]
\[
\tilde{H}_{\rho_0, \sigma_1; \ldots; \rho_0, \sigma_{l+1}}(z_{1l-1})
\]
\[
= \lim_{z_0^{-1}} \sum_{\rho, \sigma, \sigma_{l+1}} \rho, \sigma, \tilde{H}_{\rho, \sigma; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_{l+1}}(z_{0l})
\]
\[
(27)
\]
As usual the above expressions concentrate on those tensor components that are non null (check definitions (11), (14), (16) and the following discussion).

The expressions follow directly from the linearity of the \( z \)-transform and from the trivial property \( \lim_{z \to 1} A(z) = \sum_{k=0}^{\infty} A(k) \).

Note finally how it is here evident that \( \tilde{H} \), \( \tilde{H} \), and \( \tilde{H} \) (\( \tilde{H} \), \( \tilde{H} \), and \( \tilde{H} \)) are nothing but “marginalizations” of the probability information contained in \( \tilde{H} \). In the \( z \)-domain, in fact, \( \lim_{z_0^{-1}} \sum_{\rho, \sigma} \pi_{\rho, \sigma} \) is the marginalization with respect to the first observation while \( \lim_{z_l^{-1}} \sum_{\sigma_{l+1}} \) is the marginalization with respect to the last observation.

7. Summary of Theory

As a recapitulation of the above derivations recall that the main aim is to compute the tensor accounting for all possible joint probabilities \( \mathcal{H}_{\rho_1, \sigma_1; \ldots; \rho_m, \sigma_m} \), starting from the knowledge of the local stochastic dynamics taking place in each mode, i.e., the tensors \( \mathcal{H}_{\rho_1, \sigma_1; \ldots; \rho_m, \sigma_m} \) and from the knowledge of inter-mode dynamics through \( \mathcal{L}_{\rho_1, \sigma_1; \rho_2, \sigma_2} \).

The path from \( \tilde{H} \) and \( \tilde{H} \) to \( \mathcal{H} \), is, in general term, quite complex. Such a complexity is partly tamed if we agree to compute the \((m-1)\)-dimensional \( z \)-transform of the joint-probability that is given in the time domain in (2).

Even in the \( z \)-domain, the task of obtaining \( \mathcal{H} \) still involves the computation of many intermediate quantities with different probabilistic interpretation.

Figure 2 shows these quantities and their relative dependencies, and distinguishes between time-domain computations and \( z \)-domain computations. The following list recalls the meaning of the main quantities and the mechanism for passing from one to another.

- The quantity \( \mathcal{H}_{\rho, \sigma; \rho_0, \sigma_1; \ldots; \rho_0, \sigma_{l+1}}(\theta_{0l}) \) can be read as the probability of passing from the state \( S_{\rho, \sigma} \) to a state within the \( \rho_0 \)th mode \( (\rho \neq \rho_0) \) at a certain time, being observed in \( S_{\rho_0, \sigma_1} \) after \( \theta_0 \) time steps from such a transition, then in \( S_{\rho_0, \sigma_1} \) after a further \( \theta_{l-1} \) time steps up to \( S_{\rho_0, \sigma_{l+1}} \) which is reached after a further \( \theta_l \) time steps and from which the mode \( \rho_0 \) is left, given that there has been a mode transition and that the state was previously observed in \( S_{\rho, \sigma} \).
Note that $\mathcal{H}$ is null by definition when all the mode indexes but the first are not identical. This leaves $l + 4$ degrees of freedom in the choice of the tensor indexes.

- The four-indexes quantity $N_{\rho,\sigma;i_0,j_0}$ accounts for the probability of exiting mode $M_{i_0}$ from $S_{j_0}$, given that the same mode was entered some time ago from $S_{\rho,\sigma}$.

- The two-indexes quantity $\mathcal{I}_{\rho,\sigma} = \Pr\{x_{k-1} = S_{\rho,\sigma}, x_k \neq M_{\rho}\}$ accounts for the asymptotic probability of being revealed in $S_{\rho,\sigma}$ and leave $M_{\rho}$ immediately afterwards. Such a matrix must be proportional to any other eigenmatrix $\tilde{\pi}_{\rho,\sigma}$ of $N_{\rho,\sigma;i_0,j_0}$ corresponding to the unit eigenvalue.

- The scalar $\psi$ is the probability of observing a mode transition at any time. It is such that $\sum_{\rho,\sigma} \pi_{\rho,\sigma} = \psi$. It is computed as the inverse of the average sojourn time in the generic mode.

- The one-indexes quantity $I_{\rho}$ contains the average sojourn times in the modes $M_{\rho}$.

- The one-indexes quantity $\xi_{\rho}$ contains the relative frequencies of visits to each modes with respect to all other modes.

- The two-indexes quantity $M_{\rho_1,\rho_2}$ accounts for the probability of jumping from $M_{\rho_1}$ to $M_{\rho_2}$ when leaving $M_{\rho_1}$. Hence, $\xi$ is the unique eigenvector of $M$ corresponding to the unit eigenvalue whose components are normalized to have unit sum.

- The four-indexes quantity $\tilde{\mathcal{J}}(z) = [I - \mathcal{L}(z)]^{-1}$ is defined in the $z$-domain to take into account the convolution between probability distributions of sojourn times in a possibly infinite sequences of visits to different modes.

- The quantity $\mathcal{H}_{\rho_0,\sigma_{i_1},...,\rho_0,\sigma_{i_1}}(\theta_{1,1})$ accounts for the probability of being observed in $S_{\rho_0,\sigma_i}$ at a certain time and then in $S_{\rho_0,\sigma_i}$ after a further $\theta_{i-1}$ times steps up to $S_{\rho_0,\sigma_{i+1}}$ from which the mode $\rho_0$ has left.

  Note that $\mathcal{H}$ is null by definition when all the mode indexes are not identical. This leaves $l + 2$ degrees of freedom in the choice of the tensor indexes.

- The quantity $\mathcal{H}_{\rho_0,\sigma_{i_1},...,\rho_0,\sigma_{i_1}}(\theta_{1,1})$ is defined to account for the probability of being observed in $S_{\rho_0,\sigma_1}$ at a certain time and then in $S_{\rho_0,\sigma_i}$ after a further $\theta_{i-1}$ times steps up to $S_{\rho_0,\sigma_{i+1}}$ from which the mode has not been necessarily left.

  Note that $\mathcal{H}$ is null by definition when all the mode indexes are not identical. This leaves $l + 1$ degrees of freedom in the choice of the tensor indexes.

- The quantity $\mathcal{H}_{\rho,\sigma;i_1,...,i_l}(\theta_{l-1,0})$ accounts for the probability of passing from the state $S_{\rho,\sigma}$ to a state within the $\rho_0$th mode ($\rho \neq \rho_0$) at a certain time, being observed in $S_{\rho_0,\sigma_i}$ after $\theta_0$ time steps, and finally in $S_{\rho_0,\sigma_i}$ after a further $\theta_{i-1}$ time steps, given that there has been a mode transition and that the state was previously observed in $S_{\rho,\sigma}$.

  Note that $\mathcal{H}$ is null by definition when all the mode indexes but the first are not identical. This leaves $l + 3$ degrees of freedom in the choice of the tensor indexes.

Once that $\mathcal{H}$, $\tilde{\mathcal{H}}$, $\bar{\mathcal{H}}$, $\tilde{\mathcal{H}}$, and $\tilde{\mathcal{J}}$ can be obtained, a systematic recipe is given in Sec. 6 to obtain the multidimensional $z$-transform of any desired joint-probability in (2).

8. Multimode Time-Markov System and Chaotic Maps

In this section we aim at highlighting the link between time-Markov system and chaotic maps by showing that, under suitable assumptions, a two-dimensional piecewise affine chaotic map can be designed to produce quantized trajectories that are statistically indistinguishable from those produced by the machinery discussed above.

More formally, the evolution of the two state variables $\alpha_k, \beta_k \in [0, 1]$ is controlled by two maps $M, N : [0, 1]^2 \mapsto [0, 1]$ as in

$$\alpha_{k+1} = M(\alpha_k, \beta_k)$$

$$\beta_{k+1} = N(\alpha_k, \beta_k)$$

The $\alpha_k$ interval is partitioned into $\sum_{r=1}^{n} I_{r,s}$ adjacent intervals $I_{r,s}$ such that $\cup_{s=0}^{r-1} I_{r,s} = [(r - 1)/n, r/n]$. The state of the process is obtained indicating that $x_k = S_{r,s}$ when $\alpha_k \in I_{r,s}$, regardless of the value of $\beta_k$.

To make the above general theory fit within this model we must additionally assume that the local dynamics accounted for by $\mathcal{H}_{\rho_0,\sigma_{i_1},...,\rho_0,\sigma_{i_1}}(\theta_{1,1})$ is actually generated by a piecewise affine chaotic map $M_r : ([r - 1]/n, r/n] \mapsto ([r - 1]/n, r/n]$.

The two components $M$ and $N$ can be straightforwardly defined once we provide a basic “building block”.

Such a building block is a function $A_{[a,b],(P_i,Y_i)} : [a, b] \mapsto [0, 1]$ depending on its domain $[a, b]$ and on a possibly countable collection of probability-interval pairs $\{(P_i,Y_i)\}$. It takes a random variable $x$ uniformly distributed in $[a, b]$ and
produces a random variable \( y = A_{[a,b], (P, Y_i)}(x) \) that is uniformly distributed in each \( Y_i \) and such that \( \Pr \{ y \in Y_i \} = P_i \). 

Such a function can be constructed by following a procedure originally devised in [Kalman, 1956] and further discussed and optimized in [Kohda & Fujisaki, 1999] and [Kohda, 2002], once that it is properly adjusted to comply with the constraint that not only the probabilities \( P_i \) but also the intervals \( Y_i \) are given. The resulting procedure is formalized in the following few points (Adapted from [Kohda, 2002]).

1. Define the sequence of points \( \{d_i\}_i \) such that \( b = d_1 \geq d_2 \geq \cdots \geq a \) and \( d_j - d_{j+1} = P_j(b - a) \).

2. Define the function \( A_{[a,b], (P, Y_i)} \), as one of the functions that is affine in each interval \([d_{j+1}, d_j] \)

and such that

\[
[d_{j+1}, d_j] = A_{[a,b], (P, Y_i)}^{-1}(Y_j)
\]

for the sake of simplicity of future developments we assume also that \( A_{[a,b], (P, Y_i)}(x) = 0 \) for every \( x \notin [a, b] \).

To informally justify such a construction note that the counterimages of the \( Y_j \) form a partition of \([a, b]\) and that, since \( x \) is uniformly distributed, it is enough to ensure that each counterimage is an interval whose length is proportional to the probability that \( A(x) \) lands in the corresponding interval.

As an example, Fig. 3 reports the function \( A(x) \) for a certain interval \([a, b]\), some intervals \( Y_1, Y_2, Y_3, Y_4 \) and \( P_1 = 1/6, P_2 = 1/3, P_3 = 1/4 \) and \( P_4 = 1/3 \).

Going back to the aim of the section and relying on the above definition we may now choose any number \( \epsilon \in ]0, 1[, \) set \( l_{r_1, s_1, r_2, s_2} = \sum_{\tau=1}^{\infty} \mathcal{L}_{r_1, s_1, r_2, s_2}(\tau) \) and define

\[
M(\alpha, \beta) = \begin{cases} 
  \sum_{r_1, s_1} 1_{I_{r_1, s_1}}(\alpha)A_{I_{r_1, s_1}, \{l_{r_1, s_1, r_2, s_2}\}}(\alpha) & \text{if } \beta \geq \epsilon \\
  M_{[n\alpha]+1}(\alpha) & \text{if } \beta < \epsilon 
\end{cases}
\]

and then set \( h_{r_1, s_1, r_2, s_2} = (\mathcal{L}_{r_1, s_1, r_2, s_2}(\tau)/\mathcal{H}_{r_1, s_1, r_2, s_2}(1)) \) and define

\[
N(\alpha, \beta) = \begin{cases} 
  \sum_{r_1, s_1} \sum_{r_2, s_2} 1_{I_{r_1, s_1} \cap M^{-1}(I_{r_2, s_2})}(\alpha)A_{[\epsilon, 1], \{h_{r_1, s_1, r_2, s_2}\}[\epsilon, \epsilon+1]}(\beta) & \text{if } \beta \geq \epsilon \\
  \beta/\epsilon & \text{if } \beta < \epsilon 
\end{cases}
\]

where \( 1_S(\cdot) \) evaluates to 1 when its argument belongs to the set \( S \) and to zero otherwise.

To see how the above definition works assume \( \beta_k < \epsilon \) for \( k' \leq k \leq k'' \). With this, the evolution of the map \( M \) is globally “nonergodic” since if \( \beta_k \in [(r-1)/n, r/n] \) for a certain \( r \) and some \( k < k'' \), then \( \beta_{k+1} \in [(r-1)/n, r/n] \).

Hence, from the point of view of the discrete-valued process, we have \( x_k \in \mathbb{N}_r \) for \( k' \leq k \leq k'' \) and the interval of times for which \( \beta_k < \epsilon \) is the sojourn time in a mode.

In the same interval of time the real-valued state \( \beta_k \) increases at the constant rate \( r^{-1} \). If \( \beta_k' < \epsilon \) then \( \beta_k \) will not be smaller than \( r \) when \( k - k' = \lfloor \log_r \beta_k' \rfloor \). Sojourn time is then controlled by the distribution of \( \beta_k \) in the interval \([0, k]\).

More formally, the above argument indicates that if \( k' \) is the first time step spent in a certain
mode and \( \beta_{k'} \in [e^{r+1}, e^r] \) then the sojourn time in that mode is exactly \( \tau \). The statistics of the sojourn times is therefore controlled by the assignment of proper probabilities to the intervals \([e^{r+1}, e^r]\).

This is exactly the task of the branch of the map \( N \) corresponding to \( \beta_k \geq e \) that is activated at the end of a sojourn time, i.e. in the transition from a mode to another.

In such a branch, by means of properly defined functions \( A \), every possible pair of starting state \((r_1, s_1)\) and final state \((r_2, s_2)\) is considered to remap \( \beta_k \) over the whole interval \([0, 1]\) assigning to \([e^{r}, e^{r-1}]\) the probability indicated by

\[
\hat{h}_{r_1, s_1; r_2, s_2} = \frac{\mathcal{L}_{r_1, s_1; r_2, s_2}(\tau)}{\mathcal{H}_{r_1, s_1; r_2, s_2}(1)} = \frac{\Pr\{ \mathcal{S}_{r_2, s_2} = x_k, x_{k+1:k+r-1} \in \mathcal{M}_{r_2} \neq x_{k+r}| x_{k-1} = S_{r_1, s_1}, x_k \notin \mathcal{M}_{r_1} \}}{\Pr\{ x_k = S_{r_2, s_2}| x_{k-1} = S_{r_1, s_1}, x_k \notin \mathcal{M}_{r_1} \}}
\]

that is assumed to be vanishing for \( r_1 = r_2 \).

As a corollary to this construction, note that when the sojourn time is exponentially distributed as \( e^{r+1} \) we have that

\[
A_{[e, 1], [e^{r+1}, [e^r, e^{r-1}]_r]}(\beta) = \frac{x}{e} - 1
\]

At the same time step at which \( N \) set the sojourn time, the map \( M \) takes care of the mode transition by exploiting properly defined functions \( A \) so that, each possible state \((r_1, s_1)\) from which we may exit is mapped into the corresponding entering states \((r_2, s_2)\) with probability dictated by \( l_{r_1, s_1; r_2, s_2} = \sum_{\tau=1}^\infty \mathcal{L}_{r_1, s_1; r_2, s_2}(\tau) \).

As a final remark, note that the same result could be obtained by a one-dimensional chaotic map. This would imply merging of the intra-mode dynamics due to \( M \) with the “counting” procedure contained in the evolution of the state \( \beta \).

To obtain such a merging, the intervals \( [(k-1)/n, k/n] \) should be partitioned into a possibly countable number of subintervals each of them corresponding to a different \( I_{k,j} \). On each of these subintervals a replica of the \( \hat{M}_k \) has to be accommodated.

Hence, the resulting map would feature points at which replica of maps \( \hat{M}_k \) “accumulate” and the overall structure would be much more complex than the proposed two-dimensional one.

9. Local Looping Processes

Following the theory, we are able to compute any-order joint-probability (at least in the \( z \)-domain) starting from the knowledge of \( \mathcal{L} \) and \( \mathcal{H} \). Since the \( \mathcal{H} \) are themselves joint-probabilities this may imply a recursive step if the dynamics can be seen, under a greater magnification, as the combination of a time-Markov jumping mechanism between smaller modes whose local dynamics is, again, fully characterized by a smaller-scale joint-probability. The process terminates when this recursive decomposition cannot be performed either because there is no time-Markov property or because the finest granularity has been reached.

In our example we assume to be at the last possible step since the dynamics within each of the modes is a simple cycle between all the corresponding states.

More formally, we will assume that all the \( n \) modes contain the same number of states \( p_1 = p_2 = \cdots = p_n = P \) and that

\[
\hat{\mathcal{H}}_{r_1, s_1; \ldots; r_m, s_m}(\tau_1, \ldots, \tau_{m-1}) = \prod_{i=1}^{m-1} \delta_{[s_{i+1} - s_i - \tau_i]_p}
\]

where we indicate with \([x]_p\) the remainder of the integer division of \( x \) by \( P \).

Note that the assumption that all the periods of the local cycles are equal causes no loss of generality. In fact, if they were different we could think of an augmented system in which each cycle has \( P = \text{lcm}(p_1, \ldots, p_n) \) and so that the states in the \( r \)th mode replicate \( P/p_r \) times the original cycle. With this, if we agree to let each original state be indistinguishable from all its replica (e.g. by making all the observable functions assume the same value in all the replica) the original and the augmented systems must feature the same statistical properties.
Though the theory is more general, we will here assume that $L_{0,0,0,0}$ is independent of $s_2$, i.e. the sojourn time in a mode is distributed independently of the entrance state. Note that this choice along with (30) matches the constraint (4) and guarantees the validity of the above theory. In fact, expliciting all the sums,

$$
\sum_{\rho_0,\sigma_0,\theta_0=0}^{\infty} \mathcal{H}_{\rho,\sigma,\rho_0,\sigma_1}(\theta_0) = \sum_{\rho_0,\sigma_0,\sigma_1}^{\infty} \sum_{\theta_0=0}^{\infty} L_{\rho,\sigma,\rho_0,\sigma_0}(\theta_0) \delta_{[\sigma_1 - \sigma_0 - \theta_0 + 1]} = \sum_{\rho_0,\sigma_0,\sigma_1,\theta_0=0}^{\infty} L_{\rho,\sigma,\rho_0,\sigma_0}(\theta_0) = P \sum_{\rho_0,\sigma_0,\theta_0=0}^{\infty} L_{\rho,\sigma,\rho_0,\sigma_0}(\theta_0) = 1
$$

where we have noted that being $L_{\rho,\sigma,\rho_0,\sigma_0}$ independent of $\sigma_0$ the sum on $\sigma_0$ annihilates the $\delta$ term, implying that $\sum_{\rho_0}^{\infty} \sum_{\theta_0=0}^{\infty} L_{\rho,\sigma,\rho_0,0}(\theta_0) = 1/P$.

Exploiting the same invariance we may begin the computation of the $z$-transform of the key quantity

$$
\mathcal{H}_{\rho,\sigma,\rho_0,\sigma_1,\ldots,\rho_0,\sigma_1,\ldots}(\theta_{0\cdot}) = \begin{cases} A_{\rho,\sigma,\rho_0}(\theta_{0\cdot}) & \text{if } \sigma_{i+1} = \sigma_i \text{ for } i = 1, \ldots, l \text{ and } \theta_0 > 0 \\
0 & \text{otherwise} \end{cases}
$$

From now on we will define the quantities $\Delta_i = [\sigma_{i+1} - \sigma_i]_P \geq 0$ and $\Delta_0 = 1$.

With this we have

$$
\mathcal{H}_{\rho,\sigma,\rho_0,\sigma_1,\ldots,\rho_0,\sigma_1,\ldots}(z_{0\cdot}) = \sum_{k_{0\cdot}=0}^{\infty} \mathcal{A}_{\rho,\sigma,\rho_0}(k_0 + \Delta_0, P k_{1\cdot} + \Delta_{1\cdot}) z_0^{-k_0} z_i^{-P k_i - \Delta_i} = \prod_{i=0}^{l} z_{i}^{-\Delta_i} \sum_{k_{0\cdot}=0}^{\infty} \mathcal{B}_{\rho,\sigma,\rho_0}(k_0, P k_{1\cdot}) z_0^{-k_0} \prod_{i=1}^{l} z_i^{-P k_i}
$$

(32)

where we have defined $\mathcal{B}_{\rho,\sigma,\rho_0}(\theta_{0\cdot}) = A_{\rho,\sigma,\rho_0}(\theta_{0\cdot}) + \Delta_{0\cdot}$.

The subexpression (*) can be expressed as the linear combination of the $z$-transforms of the whole $\mathcal{B}_{\rho,\sigma,\rho_0}(\theta_{0\cdot})$ computed for various powers of the $P$th root of unity $\xi = e^{2\pi i/P}$.

In fact

$$
\prod_{i=1}^{l} \left[ \frac{1}{P} \sum_{j_i=0}^{P-1} \xi^{k_{i} j_i} \right] = \prod_{i=1}^{l} \delta_{\{k_i\}}_P
$$

and we may recast subexpression (*) as

$$
\sum_{k_{0\cdot}=0}^{\infty} \mathcal{B}(k_0, P k_{1\cdot}) z_0^{-k_0} \prod_{i=1}^{l} z_i^{-P k_i} = \sum_{k_{0\cdot}=0}^{\infty} \mathcal{B}(k_{0\cdot}) \prod_{i=1}^{l} \left[ \frac{1}{P} \sum_{j_i=0}^{P-1} \xi^{k_{i} j_i} \right] \prod_{i=0}^{l} z_i^{-k_i}
$$

(33)
What we need now is to compute the z-transform in the subexpression (***) of \( B \) in terms of the z-transform of \( L \), exploiting the intermediate definition of \( A \).

\[
\hat{B}(z_{(0,l)}) = \sum_{\theta_{(0,l)}=0}^{\infty} B(\theta_{(0,l)}) \prod_{i=0}^{l} z_i^{-\theta_i}
\]

\[
= \sum_{\theta_{(0,l)}=0}^{\infty} A(\theta_{(0,l)}) + \Delta_{(0,l)} \prod_{i=0}^{l} z_i^{-\theta_i}
\]

\[
= \sum_{\theta_{(0,l)}=0}^{\infty} L(\sum_{i=0}^{l} \theta_i + \Delta_i) \prod_{i=0}^{l} z_i^{-\theta_i}
\]

Finally, we substitute (34) into (33) and (33) into (32) to arrive at

\[
\hat{H}_{\rho,\sigma;\rho_0,\sigma_1;\ldots;\rho_l,\sigma_{l+1}}(z_{(0,l)})
\]

\[
= (-1)^l \prod_{i=0}^{l} z_i^{1-\Delta_i} \sum_{j_{(1,l)}=0}^{P-1} \xi - \sum_{i=1}^{l} j_i \left\{ \frac{z_0^{D-1}}{\prod_{k=1}^{l} (z_0 - z_k \xi^{-j_k})} \left( \hat{\mathcal{L}}_{\rho,\sigma;\rho_0,0}(z_0) - \sum_{s=1}^{D-1} \mathcal{L}_{\rho,\sigma;\rho_0,0}(s) z_0^{-s} \right) \right\}
\]

\[
+ \sum_{i=1}^{l} \left[ \frac{(z_i \xi^{-j_i})^{D-1}}{(z_i \xi^{-j_i} - z_0) \prod_{k=1}^{i} (z_i \xi^{-j_k} - z_k \xi^{-j_k})} \left( \hat{\mathcal{L}}_{\rho,\sigma;\rho_0,0}(z_i \xi^{-j_i}) - \sum_{s=1}^{D-1} \mathcal{L}_{\rho,\sigma;\rho_0,0}(s) (z_i \xi^{-j_i})^{-s} \right) \right]
\]

Recall that the 2\((l + 2)\)-index quantity \( \hat{H}_{\rho,\sigma;\rho_0,\sigma_1;\ldots;\rho_l,\sigma_{l+1}} \) is assumed to be null whenever the mode indexes \( \rho_0, \ldots, \rho_l \) are not equal.

Once that \( L \) is known, we may plug the above expression into (25)–(27) to obtain the non-null components of all the building blocks needed to specify the z-transform of any possible joint-probability.

### 10. A Particular Case

We may now show how the general framework developed above for locally looping system can be applied once the inter-mode dynamics is also assigned.

To this aim, let us assume that the system jumps from mode \( M_j \) to mode \( M_{j+1} \) for \( j = 1, \ldots, n - 1 \) and from \( M_n \) to \( M_1 \). The sojourn time probability assignment depends only on the landing mode. With this we may set

\[
\mathcal{L}_{\rho,\sigma;\rho_0,\sigma_0}(\tau) = F_{\rho_0}(\tau) \delta_{[\rho_0-\rho-1]_n}
\]

where \( F_{\rho_0}(\tau) \) is the function giving the probability that the sojourn time in \( M_{\rho_0} \) is \( \tau \) and that the first state in such a mode is a certain \( S_{\rho_0,\sigma_0} \).
We will consider two kinds of probability assignment
\[ F_q^{(1)}(\tau) = \frac{1 - q}{Pq} q^\tau \quad 0 \leq q < 1 \]
\[ F_q^{(2)}(\tau) = \frac{1}{P\zeta(q)} \tau^{-q} \quad 2 < q \]
where \( \zeta(\cdot) \) is the Riemann’s zeta function. We have
\[ \tilde{F}_q^{(1)}(z) = \frac{1 - q}{P(z - q)} \quad \tilde{F}_q^{(2)}(z) = \frac{1}{P\zeta(q)} \text{Li}_q \left( \frac{1}{z} \right) \]
where \( \text{Li}_q(x) \) is the polylogarithmic function defined by \( \text{Li}_q(x) = \sum_{k=1}^{\infty} x^k/k^q \).

In general we may compute
\[ N_{\rho,\sigma;\rho_0,\sigma_1} = \sum_{\theta_0=1}^{\infty} L_{\rho,\sigma;\rho_0,\sigma_1}(\theta_0) H_{\rho_0,\sigma_0;\rho_1,\sigma_1}(\theta_0) \]
\[ = \sum_{\theta_0=1}^{\infty} F_{\rho_0}(\theta_0) \delta_{[\rho_0-\rho-1]_n} \delta_{[\sigma_1-\sigma_0-\theta_0]_p} \]
\[ = \sum_{\theta_0=1}^{\infty} F_{\rho_0}(\theta_0) \delta_{[\rho_0-\rho-1]_n} \]
\[ = \frac{1}{P} \delta_{[\rho_0-\rho-1]_n} \]

It can be verified that one eigenmatrix of \( N_{\rho,\sigma;\rho_0,\sigma_1} \) corresponding to the unit eigenvalue is \( \tilde{\pi}'_{\rho,\sigma} = 1 \).

Recall now that \( \pi_{\rho,\sigma} = \gamma \pi'_{\rho,\sigma} \) if \( \gamma \) satisfies (10). To compute such a scaling factor we shall substitute (31) and (35) into (7) to compute
\[ T_{\rho_0} = \sum_{\theta_0=1}^{\infty} \sum_{\rho,\sigma,\sigma_1} \tilde{\pi}'_{\rho,\sigma} H_{\rho,\sigma;\rho_0,\sigma_1}(\theta_0) \]
\[ = \frac{1}{P} \sum_{\theta_0=1}^{\infty} \theta_0 \sum_{\rho,\sigma,\sigma_1} F_{\rho_0}(\theta_0) \delta_{[\rho_0-\rho-1]_n} \]
\[ = P \sum_{\theta_0=1}^{\infty} \theta_0 F_{\rho_0}(\theta_0) \]
\[ = -P \frac{\partial F_{\rho_0}}{\partial z}(1) \]

To compute the relative frequencies \( \xi_{\rho_0} \) we may express the matrix \( M \) as (8)
\[ M_{\rho,\rho_0} = \sum_{\sigma,\sigma_1,\theta_0=1}^{\infty} \sum_{\rho,\sigma} \tilde{\pi}'_{\rho,\sigma} L_{\rho,\sigma;\rho_0,\sigma_1}(\theta_0) \]
\[ = \frac{1}{P} \sum_{\sigma,\sigma_1,\theta_0=1}^{\infty} \sum_{\rho,\sigma} \tilde{\pi}'_{\rho,\sigma} \delta_{[\rho_0-\rho-1]_n} \]
\[ = \delta_{[\rho_0-\rho-1]_n} \]

that implies \( \xi_{\rho} = 1/n \).

With this (10) yields
\[ \gamma = \frac{1}{P} \sum_{\rho_0} \sum_{\rho} \pi_{\rho_0} T_{\rho_0} = \frac{1}{P} \sum_{\rho_0} T_{\rho_0} \]

The equation giving the asymptotic state probabilities (18) can now be joined with (31) and (35) to yield
\[ \pi_{\rho_0,\sigma_1} \]
\[ = \gamma \sum_{\theta_0=1}^{\infty} \sum_{\rho,\sigma,\sigma_2} \tilde{\pi}'_{\rho,\sigma} H_{\rho,\sigma;\rho_0,\sigma_1;\rho_0,\sigma_2}(\theta_0, \theta_1) \]
\[ = \gamma \sum_{\theta_0=1}^{\infty} \sum_{\rho,\sigma,\sigma_2} F_{\rho_0}(\theta_0 + \theta_1) \delta_{[\rho_0-\rho-1]_n} \delta_{[\sigma_2-\sigma_1-\theta_1]_p} \]
\[ = \gamma T_{\rho_0} \]
\[ = \frac{T_{\rho_0}}{P} \sum_{\rho_0} T_{\rho_0} \]

Considering the two kinds of \( F_{\rho_0} \) listed above we have
\[ T_{\rho_0}^{(1)} = \frac{1}{1 - q_{\rho_0}} \]
\[ T_{\rho_0}^{(2)} = \frac{\zeta(q_{\rho_0}) - 1}{\zeta(q_{\rho_0})} \]

when, respectively, we assume \( F = F^{(1)} \) and \( F = F^{(2)} \).

11. Local Looping Systems and Chaotic Maps

As an example, we want to construct the 2-D chaotic system whose evolution mimics a Local Looping System alternating between \( n = 2 \) modes when \( F_1 = F_{q_1}^{(1)} \) and \( F_2 = F_{q_2}^{(2)} \).

We get from Sec. 8 that such an equivalent chaotic system controls the evolution of its two state variables \( \alpha_k, \beta_k \in [0,1] \) with two state update functions \( M, N : [0,1]^2 \mapsto [0,1] \) such that
\[ \alpha_{k+1} = M(\alpha_k, \beta_k) \]
\[ \beta_{k+1} = N(\alpha_k, \beta_k) \]

Recall also that the interval containing the \( \alpha_k \) interval is partitioned into \( \sum_{r=1}^{n} p_r \) adjacent intervals \( I_{r,s} \) such that \( \bigcup_{s=0}^{p_r-1} I_{r,s} = [(r-1)/n, r/n]. \) The
state of the process is obtained indicating that $x_k = S_{r,s}$ when $\alpha_k \in I_{r,s}$, regardless of the value of $\beta_k$.

From Sec. 8 we get that the definition of $M$ and $N$ depends on the family of functions $A_{[a,b],i(P_i,Y_i)}$, mapping the interval $[a,b]$ into the union of adjacent intervals $\bigcup_j Y_j$ such that, defining the points $b = d_1 \geq d_2 \geq \ldots \geq a$ with $d_j - d_{j+1} = P_j(b-a)$,

$$[d_j, d_{j+1}] = A^{-1}_{[a,b],i(P_i,Y_i)}(Y_j)$$

From (29) we also get that the probabilities $P_i$ are related to $L_{r_1,s_1;r_2,s_2}$ and

$$H_{r_1,s_1;r_2,s_2}(1) = \sum_{\theta_1=0}^{\infty} H_{r_1,s_1;r_2,s_2}(1, \theta_1)$$

$$= \sum_{\theta_1=0}^{\infty} L_{r_1,s_1;r_2,0}(1 + \theta_1) \delta_{[s_3-s_1-\theta_1]} = \sum_{\theta_1=0}^{\infty} F_{q_{r_2}}(1 + \theta_1) \delta_{[r_2-r_1-1]} = \frac{1}{P} \delta_{[r_2-r_1-1]}$$

With this and setting $\epsilon = 1/2$ we obtain

$$M(\alpha, \beta) = \begin{cases} 
\left\lfloor \frac{\alpha + \frac{1}{2P}}{2} \right\rfloor & \text{if } \beta < \frac{1}{2} \text{ and } \alpha < \frac{1}{2} \\
\frac{1}{2} + \left\lfloor \frac{\alpha + \frac{1}{2P}}{2} \right\rfloor & \text{if } \beta < \frac{1}{2} \text{ and } \alpha \geq \frac{1}{2} \\
1_{I_{1,s_1}}(\alpha) A_{I_{1,s_1}\left\{I_{2,s_2}\right\}}(\alpha) & \text{if } \beta \geq \frac{1}{2} \text{ and } \alpha < \frac{1}{2} \\
1_{I_{2,s_1}}(\alpha) A_{I_{2,s_1}\left\{I_{1,s_2}\right\}}(\alpha) & \text{if } \beta \geq \frac{1}{2} \text{ and } \alpha < \frac{1}{2} 
\end{cases}$$

Note that each of the lower branches contains a function $A$ that is in charge of mapping a random variable uniformly distributed in a certain interval into a set of $P$ identical adjacent intervals so that the probability of falling in any of them is $1/P$.

Such a function is obviously a globally affine mapping of its domain onto such a set of intervals and the two branches above are nothing but two translated replicas of a $P$-way Bernoulli shift.

---

Fig. 4. Plots of the two components of a chaotic system whose behavior is equivalent to a Local Looping System with $n = 2$, $P = 6$ and exponentially distributed sojourn times.
that is an m-linear quantity in the values \( Q(S_{r_1,s_1}) \).

In the following we will concentrate on a case with \( n = 2 \) modes of operation. As an example, this may be considered as an extremely simplified model for the repeated transmission of a packet with two fields of variable length employing different modulations, the amount of time spent in each mode being proportional to the length of the packet field.

In this simplified model, each field is thought of as the cyclical repetition of the possible transmitted symbols that, in general, are complex quantities indicating the instantaneous amplitude and phase of the modulator output.

With our assumptions we get that the breakpoints \( d_j \) defining the function \( A_{[1/2,1],(PF_{q}^{(1)}(\tau),[2^{\tau},2^{2^{\tau}}+1])}(\beta) \) are

\[
d_j = 1 - \frac{P}{2} \sum_{\tau=1}^{j-1} F_q^{(1)}(\tau) = 1 + \frac{q^{j-1}}{2}
\]

With this the slope of \( A_{[1/2,1],(PF_{q}^{(1)}(\tau),[2^{\tau},2^{2^{\tau}}+1])}(\beta) \) in the \( \tau \)th interval is \( (2q)^{-\tau}\right/(q^{-1} - 1) \) and is obviously constant when \( q = 1/2 = \epsilon \).

Figure 4 reports the plots of the two functions

\[
E \left[ \prod_{j=1}^{m} Q(x_{\tau_1-s_1}^{j-1}) \right] = \sum_{r_1,s_1,\ldots,r_m,s_m} \prod_{j=1}^{m} Q(S_{r_1,s_1})H_{r_1,s_1;\ldots;r_m,s_m}(\tau_1,\ldots,\tau_m-1)
\]

that is an \( m \)-linear quantity in the values \( Q(S_{r_1,s_1}) \).

12. Numerical Results

As a tangible example of a computation carried over along the lines reported above, we here assume that the system emits symbols depending on the states it visits. More formally, we assume that a function \( Q : S \rightarrow C \) is given and that we observe the process \( Q(x_k) \).

Under this assumption we may think of computing expectations of any order by noting that

\[
\Pi(f) = 2 \text{Re} \left\{ \sum_{r_1,s_1,r_2,s_2} Q(S_{r_1,s_1})Q(S_{r_2,s_2})\tilde{H}_{r_1,s_1;r_2,s_2}(e^{2\pi f}) - \sum_{r,s} |Q(S_{r,s})|^2 \pi_{r,s} \right\}
\]

where we have set \( z = e^{2\pi i f} \) and constrain \( f \in [-1/2,1/2] \).

In Fig. 5 we consider \( P = 6 \) as well as \( Q(S_{1,\sigma}) = e^{2\pi i \sigma/3} \) and \( Q(S_{2,\sigma}) = e^{2\pi i \sigma/6} \). If these were to

be interpreted as transmission symbols they would imply the adoption of a three-level Phase Shift Keying (PSK) for the first part of the packed and a six-level PSK for the second part.
As far as the sojourn time distribution is concerned we assume $F_1 = F^{(1)}_{q_1}$ and $F_2 = F^{(1)}_{q_2}$ and plot $\Pi(f)$ in dB for $q_1 = 0.6, 0.7, 0.8, 0.9$ and $q_2 = 0.9$, i.e. for average sojourn times $T_1 = 2.5, 3.33, 5, 10$ and $T_2 = 10$.

Note that, though the process is not periodic, the power spectrum reflects the presence of two persistent looping modes with peaks at frequencies $f_1 = 1/3$ and $f_2 = 1/6$.

The heights of the two peaks are monotonically related with the two sojourn times $T_\rho$ since the unitary total power of the process is distributed favoring the most persistent cycle.

In Fig. 6 we consider $P = 6$ as well as $Q(S_{1,\sigma}) = e^{2\pi i}\sigma/3$ and $Q(S_{2,\sigma}) = e^{2\pi i}\sigma/6$. As far as the sojourn time distribution is concerned we assume $F_1 = F^{(2)}_{q_1}$ and $F_2 = F^{(2)}_{q_2}$ and plot $\Pi(f)$ in dB for $q_1 = 2.05, 2.1, 2.2, 2.3$ and $q_2 = 2.05$, i.e. for average sojourn times $T_1 = 12.86, 6.78, 3.75, 1.94$ and $T_2 = 12.86$.

Again, peaks are detectable at frequencies $f_1 = 1/3$ and $f_2 = 1/6$. Yet, comparing Fig. 6 with Fig. 5...
in which sojourn times were distributed exponentially one notices that peaks are much narrower in the polynomial case where the trends clearly hint at vertical asymptotes. From an intuitive point of view, this may be ascribed to the heavy-tailed feature of the polynomially distributed looping time.

To clarify this comparison, Fig. 7 still considers a two-modes system. In this case we set $T_1 = F_q^{(1)}$ and $F_2 = F_q^{(2)}$ and choose the two parameters $q_1$ and $q_2$ such that $T_1 = T_2 = 1.94, 3.75, 6.78, 12.86, 31.10$. The resulting power spectrum is plotted in dB. Since $Q(S_{1,σ}) = e^{2πiσ/3}$ and $Q(S_{2,σ}) = e^{-2πiσ/3}$ peaks are present at the frequencies $f_{1,2} = ±1/3$, the rightmost of the two being due to the exponential mode and the leftmost of the two being due to the polynomial mode.

Numerical evaluation of $\int_0^{1/2} Π(f) df$ indicates that the total unit power is evenly distributed between the two symmetric peaks. Yet, the shape of the right and left peaks are completely different and show that the heavy-tailed polynomial distribution of sojourn time implies much sharper trends and vertical asymptotes.

**12.2. Third-order features**

When $m = 3$, the $z$-transform expression for $\tilde{H}$ (see (24)) may help us to investigate a higher-order spectrum, i.e. the Fourier transform of a third-order correlation $C(τ_1, τ_2) = E[Q(x_0)Q^*(x_{τ_1})Q(x_{τ_1+τ_2})]$. Since (36) is valid only for $τ_1, τ_2 ≥ 0$, we shall define $C_{Q_0,Q_1}(τ_1, τ_2) = E[Q_0(x_0)Q_1(x_{τ_1})]$ and $C_{Q_0,Q_1,Q_2}(τ_1, τ_2) = E[Q_0(x_0)Q_1(x_{τ_1})Q_2(x_{τ_1+τ_2})]$ to write

\[
C(τ_1, τ_2) = \begin{cases} 
C_{Q,Q^*,Q}(τ_1, τ_2) & \text{if } τ_1 ≥ 0, τ_2 ≥ 0 \\
C_{Q,Q,Q^*}(τ_1 + τ_2, -τ_2) & \text{if } τ_1 ≥ 0, τ_2 < 0, τ_1 + τ_2 ≥ 0 \\
C_{Q,Q,Q^*}(-τ_1 - τ_2, τ_1) & \text{if } τ_1 ≥ 0, τ_2 < 0, τ_1 + τ_2 < 0 \\
C_{Q,Q^*,Q}(τ_2, -τ_2, -τ_1) & \text{if } τ_1 < 0, τ_2 < 0 \\
C_{Q^*,Q,Q}(τ_2, -τ_1 - τ_2) & \text{if } τ_1 < 0, τ_2 ≥ 0, τ_1 + τ_2 < 0 \\
C_{Q^*,Q,Q}(τ_1, τ_1 + τ_2) & \text{if } τ_1 < 0, τ_2 ≥ 0, τ_1 + τ_2 ≥ 0 
\end{cases}
\]

(38)

The third-order spectrum $Π(f_1, f_2)$ is the two-dimensional bilateral $z$-transform of $C(τ_1, τ_2)$ computed for $z_1 = e^{2πi f_1}, z_2 = e^{2πi f_2}$.

The double series in such a bilateral $z$-transform can be broken into six pieces corresponding to the six possible conditions in (38). With this we obtain

---

**Fig. 7.** Power spectrum of a local looping system with $n = 2$ modes with exponentially and polynomially distributed sojourn times; average sojourn times in the two modes are kept equal.

\[ \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} C(\tau_1, \tau_2) z_1^{-\tau_1} z_2^{-\tau_2} = \tilde{C}_Q Q^* Q(z_1, z_2) + \tilde{C}_Q Q Q^* \left( z_1 \frac{z_1}{z_2} \right) - \tilde{C}_Q Q^* Q(z_1) + \tilde{C}_Q Q Q^* \left( \frac{1}{z_2} \frac{z_1}{z_2} \right) - \tilde{C}_Q Q Q^* \left( \frac{z_1}{z_2} \right) + \tilde{C}_Q Q Q^* \left( \frac{1}{z_2} \frac{1}{z_1} \right) - \tilde{C}_Q Q^2 Q(z_1) - \tilde{C}_Q Q^2 Q^* \left( \frac{1}{z_2} \frac{1}{z_1} \right) + \tilde{C}_Q Q^2 Q^2 \left( \frac{z_2}{z_1} \right) + \tilde{C}_Q Q Q^2 Q \left( \frac{z_2}{z_1} \right) - \tilde{C}_Q Q^2 Q Q^* \left( \frac{z_2}{z_1} \right) \]

and thus

\[ \Pi(f_1, f_2) = Q(S_{r_1, s_1}) Q^*(S_{r_2, s_2}) Q(S_{r_3, s_3}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} \tilde{\mathcal{H}}_{r_3, s_3} (e^{2\pi i f_1}, e^{2\pi i f_2}) + Q(S_{r_1, s_1}) Q(S_{r_2, s_2}) Q^*(S_{r_3, s_3}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} \tilde{\mathcal{H}}_{r_3, s_3} (e^{2\pi i f_1}, e^{2\pi i f_1 - f_2}) - Q(S_{r_1, s_1}) Q(S_{r_2, s_2}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} (e^{2\pi i f_1}) + Q(S_{r_1, s_1}) Q(S_{r_2, s_2}) Q^*(S_{r_3, s_3}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} \tilde{\mathcal{H}}_{r_3, s_3} (e^{2\pi i f_1}, e^{2\pi i f_1 - f_2}) - Q^2(S_{r_1, s_1}) Q^*(S_{r_2, s_2}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} (e^{2\pi i f_1 - f_2}) + Q(S_{r_1, s_1}) Q^*(S_{r_2, s_2}) Q(S_{r_3, s_3}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} \tilde{\mathcal{H}}_{r_3, s_3} (e^{2\pi i f_2}, e^{2\pi i f_1}) - Q(S_{r_1, s_1}) Q(S_{r_2, s_2}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} (e^{2\pi i f_2}) + Q(S_{r_1, s_1}) Q(S_{r_2, s_2}) Q^2(S_{r_3, s_3}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} \tilde{\mathcal{H}}_{r_3, s_3} (e^{2\pi i f_2}, e^{2\pi i f_1}) - Q(S_{r_1, s_1}) Q^2(S_{r_2, s_2}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} \tilde{\mathcal{H}}_{r_3, s_3} (e^{2\pi i f_2}, e^{2\pi i f_2}) - Q(S_{r_1, s_1}) Q^2(S_{r_2, s_2}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} \tilde{\mathcal{H}}_{r_3, s_3} (e^{2\pi i f_2 - f_1}, e^{2\pi i f_2}) - Q^2(S_{r_1, s_1}) Q(S_{r_2, s_2}) Q^2(S_{r_3, s_3}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} \tilde{\mathcal{H}}_{r_3, s_3} (e^{2\pi i f_2 - f_1}, e^{2\pi i f_2}) - Q(S_{r_1, s_1}) Q^2(S_{r_2, s_2}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} \tilde{\mathcal{H}}_{r_3, s_3} (e^{2\pi i f_2 - f_1}, e^{2\pi i f_2}) - Q(S_{r_1, s_1}) Q^2(S_{r_2, s_2}) \tilde{\mathcal{H}}_{r_1, s_1} \tilde{\mathcal{H}}_{r_2, s_2} \tilde{\mathcal{H}}_{r_3, s_3} (e^{2\pi i f_2 - f_1}, e^{2\pi i f_2}) + (39) \]

Let us now consider a system with \( P = 6 \) as well as \( Q(S_{1,0}) = 1/2, Q(S_{1,1}) = 0, Q(S_{1,2}) = -1, Q(S_{1,3}) = 0, Q(S_{1,4}) = 1/2, Q(S_{1,5}) = 0, Q(S_{2,0}) = i, Q(S_{2,1}) = 0, Q(S_{2,2}) = -i, Q(S_{2,3}) = i, Q(S_{2,4}) = 0, Q(S_{2,5}) = -i. \) We also set \( F_1 = F_1^{(1)} \) and \( F_2 = F_2^{(1)} \). Note that the observables are such that when the process is in \( M_1 \) it follows a cycle of period 6 while \( M_2 \) is a cycle of period 3.

To ease the physical interpretation of the resulting graphics, note that a more natural way of computing a third order spectrum could be to transform a correlation function \( E(Q(x_1), Q^*(x_2), Q(x_3)) \) depending on the absolute instants \( t_1 \) and \( t_2 \) instead of the time lags \( \tau_1 \) and \( \tau_2 \). The relationship between the more intuitive spectrum \( \Pi'(f_1, f_2) \) and the one computed in (39) is that \( \Pi'(f_1, f_2) = \Pi(f_1 + f_2, f_2) \).

In Fig. 8 we report \( \Pi'(f_1, f_2) \) for \( T_1 = 5 \) and \( T_2 = 3.33 \) both as 3D and contour plots. Note how peaks are present with frequency coordinates that are \( \pm 1/6 \) and \( \pm 1/3 \).

To show that such a spectral profile signifies the super-imposition of two cycling behaviors, in Fig. 9 we report \( \Pi'(f_1, f_2) \) when, respectively, the period-3 mode and the period-6 mode emit negligible values, i.e. for \( Q(S_{2,2}) = 0 \) and for \( Q(S_{1,1}) = 0 \).

Note how in Fig. 9(a) the basic harmonics \( \pm (1/6, 1/6) \) and \( \pm (1/3, 1/3) \) are present as well as all the other cross-harmonics \( ((1/3, -1/6), (1/6, -1/3), (-1/3, 1/6), (-1/6, 1/3)) \) that do not imply cancellation of zero-mean behaviors, i.e. \( f_1 + f_2 = 0 \) and \( f_1 + f_2 = \pm 1/2 \). With this in mind, Fig. 9(b) exposes \( \Pi'(f_1, f_2) \) for a period-3 cycle in which only principal harmonics can be detected for \( (f_1, f_2) \in [-1/2, 1/2] \times [-1/2, 1/2] \). As (39) shows, the super-imposition of the two partial \( \Pi' \) is not a linear operation. Hence, the behavior of the overall systems highlights a non-symmetric interaction between the period-3 and period-6 cycles in that the peak in \( (1/3, 1/3) \), is higher than the one in \( (-1/3, -1/3) \).
13. Conclusion

The paper widens the scope of a previously proposed framework for the statistical characterization of finite-valued processes whose evolution has some memory-one properties.

We further assume that the state space of such processes can be partitioned into modes and that, if we know that the process is confined in that mode, the statistical characterization is fully understood.

The process is also allowed to stochastically move from one mode to another and the number of time steps for which it remains in each mode is a random variable whose distribution is a function only of the mode visited before as well as the current one.
Relying on multidimensional z-transformation and multilinear algebraic tools we give a systematic procedure to compute any-order joint probability of such processes.

This statistical characterization is independent of the actual values associated to different process states by the quantization function and can be reused “as is” when the values change.

We have also highlighted how, under suitable assumptions, the behavior of some processes falling in the above class is indistinguishable from that of two-dimensional piecewise affine chaotic map for which we give a systematic design procedure.

As an example, this technique is applied in the final part to processes generated by local looping systems. Local looping processes model stochastic phenomena merging cycles with different permanency times and may be the model of choice for a system switching between different modes of operation, each of them characterized by the cyclical repetition of a certain number of actions, such as the subsequent steps in a communication protocol, the various phases of a bus cycle, the load-compute-store mechanism of a microprocessor, etc.

Calculations are developed in detail to show how such processes could be generated by suitably designed 2-D chaotic maps.

Moreover, in a more specific case, loosely related to the modeling of a transmission system with a variable modulation protocol, we show how second- and third-order spectra may be obtained when exponentially or polynomially decaying distribution are assumed for cycle sojourn times.

References


Rovatti, R., Mazzini, G. & Setti, G. [2000] “A tensor approach to higher-order expectations of quantized
Appendix
The following is an auxiliary result.

Lemma 1.

\[
\sum_{\sum_{i=0}^{l} a_i = s} \prod_{i=0}^{l} z_i^{-a_i} = \begin{cases} 
(-1)^{l} \prod_{i=0}^{l} z_i^{-b_i} \sum_{i=0}^{l} \frac{z_i^{-1-s+B}}{\prod_{j=0}^{l} (z_i - z_j)} & \text{if } s \geq B \\
0 & \text{otherwise}
\end{cases}
\]

where \( B = \sum_{i=0}^{l} b_i \). That, if we also have \( B \leq l \) can be rewritten as

\[
\sum_{\sum_{i=0}^{l} a_i = s} \prod_{i=0}^{l} z_i^{-a_i} = \begin{cases} 
(-1)^{l} \prod_{i=0}^{l} z_i^{-b_i} \sum_{i=0}^{l} \frac{z_i^{-1-s+B}}{\prod_{j=0}^{l} (z_i - z_j)} & \text{if } s \geq 0 \\
0 & \text{otherwise}
\end{cases}
\]

Proof. Let us define \( g_i(a) = z_i^{-a} \) for \( a \geq b_i \) and 0 otherwise. Knowing that the \( z \)-transform of \( g_i \) expressed in \( x \) is \( \tilde{g}_i(x) = (z_i x)^{1-b_i} / (z_i x - 1) \) we may proceed as follows

\[
\sum_{\sum_{i=0}^{l} a_i = s} \prod_{i=0}^{l} z_i^{-a_i} = g_0 * g_1 * \cdots * g_l(s) = Z^{-1} \left[ \prod_{i=0}^{l} \tilde{g}_i(x) \right] (s)
\]

\[
= Z^{-1} \left[ x^{l+1-B} \prod_{i=0}^{l} z_i^{-b_i} \prod_{i=0}^{l} \frac{1}{z_i x - 1} \right] (s) = \begin{cases} 
(-1)^{l} \prod_{i=0}^{l} z_i^{-b_i} \sum_{i=0}^{l} \frac{z_i^{-1-s+B}}{\prod_{j=0}^{l} (z_i - z_j)} & \text{if } s \geq B \\
0 & \text{otherwise}
\end{cases}
\]

For \( B \leq l \) the expression in the \( (s \geq B) \)-branch of the above expression vanishes for all the non-negative integers \( s < B \) so that the corresponding constraint can be relaxed.