Stochastic Processes and their Applications 201 (2026) 105039

Contents lists available at ScienceDirect

Stochastic Processes and their Applications

journal homepage: www.elsevier.com/locate/spa

Central limit theorems for interacting innovation processes, related statistical tools and general results

Giacomo Aletti [a,1], Irene Crimaldi [b,2,3], Andrea Ghiglietti [c,2,*]

[a] Department of Environmental Science and Policy, Universitá degli Studi di Milano, Italy
[b] IMT School for Advanced Studies Lucca, Italy
[c] Department of Statistics and Quantitative Methods, Universit degli Studi di Milano-Bicocca, Italy

ARTICLE INFO

2000 MSC:
Primary
60F15
60K35
Secondary
91D30

Keywords:
Reinforcement
Interaction
Urn model
Poisson-Dirichlet process
Innovation process
Central limit theorem
Stable convergence

ABSTRACT

We study a networked system of innovation processes, where each process is modeled as an urn with infinitely many colors-a classical framework for capturing the emergence of novelties. Extending this paradigm, we analyze a model of interacting urns, where the probability of generating or reusing elements in one process is influenced by the histories of others. This interaction is governed by two matrices that control innovation triggering and reinforcement dynamics across the system. The core contribution of this work is a detailed analysis of the second-order asymptotic behavior of the model. Building on these theoretical results, we develop statistical tools to infer the structure and strength of inter-process influence. The methodology is framed in a general setting, making it broadly applicable. We validate our approach with applications to two real-world datasets from Reddit discussions and Gutenberg text corpora.

1. Introduction

Understanding how novelties emerge, propagate, and catalyze further innovations is crucial in several disciplines, including biology, linguistics, and social sciences [1–6]. In probabilistic terms, a *novelty* (or *innovation*) is the first occurrence of an event of interest.

Widely adopted mathematical models for such *innovation processes* include *urn models with infinitely many colors* [7–10] and *species sampling sequences* [11–15]. In the urn metaphor, for each $t \in \mathbb{N} \setminus \{0\}$, the random variable C_t denotes the observed color at extraction

* Corresponding author.
E-mail addresses: giacomo.aletti@unimi.it (G. Aletti), irene.crimaldi@imtlucca.it (I. Crimaldi), andrea.ghiglietti@unimib.it (A. Ghiglietti).
[1] Member of "Gruppo Nazionale per il Calcolo Scientifico (GNCS)" of the Italian Institute "Istituto Nazionale di Alta Matematica (INdAM)"
[2] Member of "Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA)" of the Italian Institute "Istituto Nazionale di Alta Matematica (INdAM)"
[3] Irene Crimaldi thanks the "Resilienza Economica e Digitale" (RED) project (CUP D67G23000060001) funded by the Italian Ministry of University and Research (MUR) as "Department of Excellence" (Dipartimenti di Eccellenza 2023–2027, Ministerial Decree no. 230/2022). Partially supported by the European Union - NextGenerationEU through the Italian Ministry of University and Research under the National Recovery and Resilience Plan (PNRR) - M4C2 - Investment 1.3, title [Public Sector Indicators for Sustainability and Wellbeing (PUWELL)] - Program [Growing Resilient, INclusive and Sustainable (GRINS)] - PE18 - CUP J33C22002910001.

https://doi.org/10.1016/j.spa.2026.105039
Received 31 July 2025; Received in revised form 1 July 2026; Accepted 1 July 2026
Available online 3 July 2026
0304-4149/

t. The probability that the $(t+1)$-th draw yields a "new" color (i.e., one not previously observed) is Z_t^*, a function of $C_1, \dots, C_t$, while the probability of drawing an "old" (i.e., already observed) color c is $P_{c,t} = \sum_{n=1}^{t} Q_{n,t} \mathbb{1}_{\{C_n = c\}}$, where $Q_{n,t}$ depends on $C_1, \dots, C_t$. The quantities Z_t^* and $Q_{n,t}$ specify the model. In this framework, the set of possible colors is unbounded, so new colors may continuously enter the system: the urn represents the space of possibilities, and the extracted sequence the realized history.

A canonical example is the *Poisson-Dirichlet (PD) model* [16,17]. In this model,

$$Z_t^* = \frac{\theta + \gamma D_t}{\theta + t}, \qquad Q_{n,t} = \frac{1 - \gamma / K_{C_n,t}}{\theta + t}, \qquad \text{and so} \quad P_{c,t} = \frac{K_{c,t} - \gamma}{\theta + t}, \tag{1}$$

where $0 \le \gamma < 1$, $\theta > -\gamma$, D_t is the number of distinct colors observed up to time-step t, and $K_{c,t}$ is the count of color c up to time-step t. This model embodies the "preferential attachment" (or "popularity") principle, since the probability of observing a color is proportional to its frequency. The PD process effectively captures empirical laws such as *Heaps' law* [18], describing the sublinear growth of distinct elements with the number of observations, and *Zipf's law* [19], characterizing the sublinear decrease of ranked frequencies. In particular, in the PD model $D_t \propto t^{\gamma}$ for $0 < \gamma < 1$, and $D_t \propto \ln(t)$ when $\gamma = 0$ [20,21].

An intuitive representation of the PD process is the *urn with triggering* model [22,23] (or the equivalent formulation in [24]). Initially, the urn contains $N_0 > 0$ distinct balls of different colors. At each time-step $t+1$, a ball is drawn at random with replacement and:

- if the drawn color is *new*, we add $(\hat{\rho} - 1)$ balls of the same color and $(\nu + 1)$ balls of entirely new (distinct) colors;
- if the drawn color is *old*, we add ρ balls of the same color.

When the *balance condition* $\hat{\rho} + \nu = \rho$ holds (so that the number of added balls is constant), the model reproduces (1) by taking $\rho > \nu \ge 0$ and setting $\theta = N_0/\rho$ and $\gamma = \nu/\rho$.

This formulation highlights both *reinforcement* and *Kauffman's principle of the adjacent possible* [25]: when a novelty occurs, it expands the space of potential future novelties by introducing new colors.

To study how distinct innovation processes influence each other, in [26] a finite network of urns with triggering was introduced and analyzed (see Section 2). Each node has its own urn, and the probability of observing a new or old item at a node depends on both the node's history and the histories of the other nodes. Interactions are governed by two matrices: Γ and W. Matrix Γ modulates how novelties produced in the system affect each node's probability of generating new novelties, extending the adjacent possible to the system level. Matrix W regulates reinforcement across nodes, making the probability of extracting already observed items depend on their frequencies in all nodes.

When Γ and W are irreducible, the model captures synchronization: as shown in [26, Theorem 3.1 & Theorem 3.2] all nodes produce innovations at the same asymptotic rate, and the distribution of observed items becomes uniform across the network. These results are summarized in the following Theorems 2.1, and 2.2.

This work aims to:

- derive second-order asymptotic behaviour of the model by providing Central Limit Theorems (CLTs) for the quantities of interest;
- develop statistical tools for inference on the interaction matrices Γ and W, as well as on the limiting probabilities of observing an extracted item c;
- apply the theoretical findings to real-world datasets, assessing interaction intensities in empirical contexts.

We also emphasize that the results developed in the appendix are of interest beyond the present model. They concern central limit theorems for a pair of multidimensional stochastic processes following a suitable recursive dynamics, and involve random centering, random limiting covariance matrices, and non-standard scaling. Since the asymptotic covariance matrices in these results are random, we do not rely on standard convergence in distribution; instead, we employ the notion of *stable convergence*.

1.1. Related literature

Although interacting urn models with a finite set of colors are widely studied (see, e.g., [27,28] and references therein), works on urns with infinitely many colors are comparatively scarce. Notable exceptions are [29,30], which investigate Dirichlet processes with random reinforcement and urns with triggering, respectively. These approaches differ from ours in the definition of novelty and in the interaction structure, in particular concerning the system-level adjacent possible and the dual reinforcement mechanism.

In [29] (see Example 3.8 therein), a finite collection of Dirichlet processes with random reinforcement is considered. A random weight $W_{t,h}$ is associated with the extraction at time-step t from urn h, and the probability of extracting an old color (where "old" is defined at the level of urn h only) is proportional to the weight of that color. Interactions across urns are introduced through stochastic dependence among the weights: for instance, the weights may be shared across urns, depend on outcomes of other urns, or depend on common observable or latent variables.

The present work and [30] both study interacting urns with triggering, where interaction acts when a new color is extracted, thus extending Kauffman's principle from a single agent to a network. In contrast with [30], in our model the extraction of an "old" item is also interactive: the probability of extracting an old item depends on its total number of occurrences across all nodes. Moreover, [30] focuses on node-level novelties ("discoveries"), i.e. first appearances in a given node, possibly already observed in other nodes, whereas we study the sequence of novelties for the whole network produced by each node.

1.2. Structure of the work

The remainder of the paper is organized as follows. In Section 2 we describe the model and recall the first-order convergence results. In Section 3 we present the main results, consisting of CLTs describing the second-order asymptotic behaviour of the interacting innovation processes. Then, in Section 4 we illustrate statistical tools derived from the CLTs. The inferential methodology is presented for a suitable bi-dimensional recursive process in a separate appendix, including theoretical results and proofs; sections, theorems, and remarks in the appendix are referenced with capital letters. Finally, in Section 5 we apply the proposed methods to two real datasets, evaluating interaction intensities.

2. Model and first-order convergence results

Suppose we have N urns, labeled from 1 to N. At time-step 0, the colors inside each urn are different from those in the other urns. Let $N_{0,h} > 0$ be the number of distinct balls with distinct colors inside the urn h. Then, at each time-step $t \geq 1$, one ball is randomly drawn with replacement from each urn and, for any $h = 1, \dots N$, urn h is so updated according to the colors extracted from urn h itself and from all the other urns $j \neq h$:

- if the color $C_{t,h}$ of the ball extracted from urn h is “new” (i.e., it appears for the first time in the system), then into urn h we add $(\widehat{\rho}_{h,h} - 1)$, with $\widehat{\rho}_{h,h} > 0$, balls of the same color as the extracted one plus $(\nu_{h,h} + 1)$, with $\nu_{h,h} \geq 0$, balls of distinct “new” colors (i.e. not already present in the system);
- if the color $C_{t,h}$ of the ball extracted from urn h is “old” (i.e., it has been already extracted in the system), we add $\rho_{h,h} > 0$ balls of the same color into urn h;
- for each $j \neq h$, if the color $C_{t,j}$ of the ball extracted from urn j is “new” (i.e., it appears for the first time in the system), then into urn h we add $\widehat{\rho}_{j,h} \geq 0$ balls of the same color as the one extracted from urn j plus $\nu_{j,h} \geq 0$ balls of distinct “new” colors (i.e. not already present in the system);
- for each $j \neq h$, if the color $C_{t,j}$ of the ball extracted from urn j is “old” (i.e., it has been already extracted in the system), then into urn h we add $\rho_{j,h} \geq 0$ balls of the same color as the one extracted from urn j.

The terms “new” and “old” refer to the entire system. We assume that the “new” colors added to a given urn are always different from those added to the other urns (at the same time-step or in the past). Together with the initial condition, this implies that the same novelty cannot be observed simultaneously in different urns. Hence, for each observed novelty c, there exists a unique urn, denoted $j^*(c)$, that produced it. After its first extraction, c may subsequently appear in other urns due to interactions among urns. Furthermore, as in the standard Poisson-Dirichlet model, we assume the *balance condition*

$$\widehat{\rho}_{j,h} + \nu_{j,h} = \rho_{j,h}, \quad \text{i.e. } \widehat{\rho}_{j,h} = \rho_{j,h} - \nu_{j,h}\,, \tag{2}$$

so that, at each time-step, each urn j contributes to increase the number of balls inside urn h by $\rho_{j,h} \geq 0$, with $\rho_{h,h} > 0$. Moreover, letting $\rho_h = \sum_{j=1}^N \rho_{j,h}$ and setting (without loss of generality)

$$\theta_h = N_{0,h}/\rho_h, \quad \gamma_{j,h} = \nu_{j,h}/\rho_h, \quad \lambda_{j,h} = \widehat{\rho}_{j,h}/\rho_h \quad \text{and} \quad w_{j,h} = \rho_{j,h}/\rho_h\,, \tag{3}$$

we obtain

$$Z^*_{t,h} = P(C_{t+1,h} = \text{“new”} \mid \mathbf{C}_1, \dots, \mathbf{C}_t) = \frac{\theta_h + \sum_{j=1}^N \gamma_{j,h} D^*_{t,j}}{\theta_h + t}\,, \tag{4}$$

where $\mathbf{C}_n = (C_{n,1}, \dots, C_{n,N})^\top$ with $n = 1, \dots, t$, $D^*_{t,j}$ denotes the number, until time-step t, of distinct observed colors extracted for their first time from urn j, that is the number of distinct novelties for the whole system “produced” by urn (process) j until time-step t, and, for each “old” color c,

$$\begin{aligned} P_t(h,c) &= P(C_{t+1,h} = c \mid \mathbf{C}_1, \dots, \mathbf{C}_t) \\ &= \frac{\sum_{j \neq j^*(c)} w_{j,h} K_t(j,c) + w_{j^*(c),h}(K_t(j^*(c),c) - 1) + \lambda_{j^*(c),h}}{\theta_h + t} \\ &= \frac{\sum_{j=1}^N w_{j,h} K_t(j,c) - \gamma_{j^*(c),h}}{\theta_h + t}\,, \end{aligned} \tag{5}$$

where $K_t(j,c)$ denotes the number of times the color c has been extracted from urn j until time-step t and $j^*(c)$ denotes the urn from which the color c has been extracted for the first time. Note that the probability that urn h produces at time-step $t+1$ a novelty for the entire system depends increasingly on the number $D^*_{t,j}$ of novelties produced by urn j up to time t, and $\gamma_{j,h}$ regulates this dependence. Thus, Kauffman’s principle of the adjacent possible holds at the system level: for each pair (j,h), $\gamma_{j,h}$ quantifies how the production of a novelty by urn j induces potential novelties in urn h. Similarly, the probability that urn h extracts an old color c at time-step $t+1$ depends increasingly on the counts $K_t(j,c)$, with $w_{j,h}$ quantifying the influence of urn j on urn h through reinforcement.

In [26] (see also [31, Appendix B.1] for a general theory) the first-order asymptotic properties of the model was proved:

Theorem 2.1 ([26, Theorem 3.1]). *Suppose that the matrix $\Gamma = (\gamma_{j,h})_{j,h=1,\dots,N}$ is irreducible. Denote by $\gamma^* \in (0,1)$ the Perron-Frobenius eigenvalue of Γ and by $\mathbf{u}$ its left eigenvector. Then, we have*

$$t^{-\gamma^*}\mathbf{D}^*_t \xrightarrow{a.s.} D^{**}_\infty \mathbf{u}\,,$$

where $\mathbf{D}^*_t = (D^*_{t,1}, \dots, D^*_{t,N})^\top$ *and* D^{**}_∞ *is an integrable strictly positive random variable.*

To connect this first-order result with the fluctuations studied later in Theorem 3.2, note that the above stated vector convergence implies that,

$$\frac{D^*_{t,h}}{D^*_{t,j}} \longrightarrow \frac{u_h}{u_j} \quad \text{a.s. for all } h, j, \qquad \text{and} \qquad \frac{D^*_{t,h}}{\sum_{\ell=1}^N D^*_{t,\ell}} \longrightarrow \frac{u_h}{\sum_{\ell=1}^N u_\ell} \quad \text{a.s.}$$

Theorem 2.2 ([26, Theorem 3.2]). *Suppose that the matrix* $W = (w_{j,h})_{j,h=1,\dots,N}$ *is irreducible. Then, for each observed color* c *in the system, we have*

$$\mathbf{P}_t(c) \xrightarrow{a.s.} \widetilde{P}_\infty(c)\mathbf{1}$$

$$\frac{1}{t}\mathbf{K}_t(c) \xrightarrow{a.s.} \widetilde{P}_\infty(c)\mathbf{1} \quad \textit{and so} \quad \frac{K_t(h,c)}{\sum_{j=1}^N K_t(j,c)} \xrightarrow{a.s.} \frac{1}{N},$$

where $\mathbf{P}_t(c) = (P_t(1,c), \dots, P_t(N,c))^\top$, $\mathbf{K}_t(c) = (K_t(1,c), \dots, K_t(N,c))^\top$, $\mathbf{1}$ *denotes the vector with all the components equal to* 1 *and* $\widetilde{P}_\infty(c)$ *is a random variable that takes values in* $(0,1)$. [1]

The first result states that, when Γ is irreducible, equivalently when the graph whose nodes are the innovation processes and whose adjacency matrix is Γ is strongly connected, all processes exhibit the same asymptotic Heaps' exponent. More precisely, $D^*_{t,h} \propto t^{\gamma^*}$ a.s. for every $h = 1, \dots, N$, with $\gamma^* \in (0,1)$. Thus, in the long run, all nodes in the network generate innovations for the system at the same rate. Moreover, the ratio $D^*_{t,h}/D^*_{t,j}$ is a strongly consistent estimator of the ratio u_h/u_j of the relative eigenvector centrality scores, computed with respect to $\Gamma^\top$. Similarly, the fraction of system-wide innovations produced by process h converges almost surely to its absolute eigenvector centrality score.

The second result concerns the reinforcement matrix W. If W is irreducible, or equivalently if the graph with adjacency matrix W is strongly connected, then, for each observed item c, the number of times c is observed in each process h grows linearly with time. Furthermore, the occurrences of c become asymptotically uniformly distributed across the processes. At equilibrium, the limiting probability $\widetilde{P}_\infty(c)$ of observing item c is the same for every process.

In the next section we complete the analysis of the asymptotics of the model providing central limit theorems for the fluctuations of the number of system-level novelties produced by each process and for the fluctuations of the empirical frequencies of a fixed observed color across the urns.

3. Central limit theorems

Assume $N \geq 2$. Let Γ, W, Λ be the *non-negative* $N \times N$ square matrices with elements equal to the model parameters $\gamma_{j,h}$, $w_{j,h}$ (with $w_{h,h} > 0$) and $\lambda_{j,h}$ (with $\lambda_{h,h} > 0$), respectively. We recall that, by the balance condition (2) and the reparametrization (3), we have

$$W = \Gamma + \Lambda, \qquad \mathbf{1}^\top\Gamma < \mathbf{1}^\top \qquad \text{and} \qquad \mathbf{1}^\top W = \mathbf{1}^\top, \tag{6}$$

where $\mathbf{1}$ denotes the vector with all the components equal to 1. As observed above, the matrix Γ rules the production of potential novelties and, in particular, its elements out of the diagonal regulate the interaction among the innovation processes with respect to this issue; while, the matrix W rules the interaction among the innovation processes with respect to the choice of an old item. Moreover, recall that $\mathbf{D}^*_t = (D^*_{t,1}, \dots, D^*_{t,N})^\top$, $\mathbf{P}_t(c) = (P_t(1,c), \dots, P_t(N,c))^\top$ and $\mathbf{K}_t(c) = (K_t(1,c), \dots, K_t(N,c))^\top$. Finally, the symbol $\mathbf{0}$ will denote the vector with all the components equal to 0.

In order to simplifying the statements of the following theorems, we summarize the required notation and conditions in the following two sets of assumptions:

Assumption 3.1. Assume that the non-negative matrix $\Gamma = (\gamma_{j,h})_{j,h=1,\dots,N}$ is irreducible and diagonalizable. Denote by $\gamma^* \in (0,1)$ the Perron-Frobenius eigenvalue of Γ, by $\mathbf{v}$ the corresponding right eigenvector with strictly positive entries and such that $\mathbf{v}^\top\mathbf{1} = 1$ and denote by $\mathbf{u}$ the corresponding left eigenvector with strictly positive entries and such that $\mathbf{v}^\top\mathbf{u} = 1$. Finally, let γ^*_2 be an eigenvalue of Γ different from γ^* and with highest real part and assume that $\mathcal{R}e(\gamma^*_2)/\gamma^* < 1/2$.

Assumption 3.2. Assume that the non-negative matrix $W = (w_{j,h})_{j,h=1,\dots,N}$ is irreducible and diagonalizable. Denote by $\mathbf{v}_W$ the right eigenvector associated to the leading eigenvalue $w^* = 1$, with strictly positive entries and such that $\mathbf{v}_W{}^\top\mathbf{1} = 1$. Finally, let w^*_2 be an eigenvalue of W different from $w^* = 1$ and with highest real part and assume that $\mathcal{R}e(w^*_2) < \frac{1}{2}$.

Irreducibility means that the directed interaction graph has a single strongly connected component: every process can influence every other process, possibly through intermediate processes. In empirical terms, this corresponds to a genuinely connected system in which no subpopulation evolves asymptotically independently from the rest. If irreducibility fails, the model may decompose into

[1] Note that in [26, Theorem 3.2] this random variable takes values in $(0,1]$, but, applying [26, Theorem S1.3] also replacing $P_t(h,c)$ by $1 - P_t(h,c)$, one immediately obtains $(1 - \widetilde{P}_\infty(c)) > 0$ a.s., that is $\widetilde{P}_\infty(c) < 1$ a.s.

communicating classes, and different blocks may display different limiting innovation rates or different reinforcement behavior. This requires a separate block-wise analysis.

Diagonalizability is mainly a technical condition used to obtain transparent formulas for the covariance matrices and to keep the proof readable. In applications it is not a restrictive structural requirement in most continuous parametric families, since non-diagonalizable matrices form a lower-dimensional exceptional set.

The quantities $\mathcal{R}e(\gamma_2^*)/\gamma^*$ and $\mathcal{R}e(w_2^*)$ measure the persistence of non-leading modes, namely the modes describing deviations from the Perron–Frobenius synchronization direction. The conditions $\mathcal{R}e(\gamma_2^*)/\gamma^* < 1/2$ and $\mathcal{R}e(w_2^*) < 1/2$ require these disagreement modes to decay sufficiently fast for martingale fluctuations to dominate, yielding the Gaussian limits with the stated scaling. From an applied perspective, they correspond to systems in which the synchronization induced by the network is strong enough that long-lived community or block effects do not dominate the second-order fluctuations.

We are now ready to provide the central limit theorems for the considered model, that describe the second-order behaviour of the quantities in Theorems 2.1 and 2.2 under the above assumptions. A discussion of possible extensions of these results is postponed to Remark 3.1, after the proofs, at the end of this section. For completeness, we also point out that when $N = 1$, the model reduces to the usual Poisson–Dirichlet innovation process and the corresponding CLT recovers the known one-dimensional results, as investigated for instance by [32,33].

Theorem 3.1. *Under Assumption 3.1, we have*

$$t^{\gamma^*/2}\left(\frac{\mathbf{D}_t^*}{t^{\gamma^*}} - D_\infty^{**}\mathbf{u}\right) \xrightarrow{stably} \mathcal{N}(\mathbf{0}, D_\infty^{**}\gamma^* C_{det,\Gamma}),$$

where D_∞^{**} is defined in Theorem 2.1 and $C_{det,\Gamma}$ is a deterministic matrix depending on the eigen-structure of Γ, which is possible to write explicitly.

Proof. Denote by $X_{t,h}^*$ the random variable that takes value 1 when the ball extracted from urn h at time-step t has a new (for all the system) color and is equal to 0 otherwise. Then $Z_{t,h}^*$ coincides with $P(X_{t+1,h}^* = 1 \mid \mathbf{C}_1, \dots, \mathbf{C}_t) = E[X_{t+1,h}^* \mid \mathbf{C}_1, \dots, \mathbf{C}_t]$ and $D_{t,h}^*$ can be written as $\sum_{n=1}^t X_{n,h}^*$. Since we have

$$Z_{t,h}^* = \frac{\theta_h + \sum_{n=1}^t \sum_{j=1}^N \gamma_{j,h} X_{n,j}^*}{\theta_h + t},$$

we obtain the following dynamics for $Z_{t,h}^*$:

$$Z_{0,h}^* = 1, \qquad Z_{t+1,h}^* = (1 - r_{t,h})Z_{t,h}^* + r_{t,h}\sum_{j=1}^N \gamma_{j,h} X_{t+1,j}^* \quad \text{for } t \geq 0,$$

where $r_{t,h} = 1/(\theta_h + t + 1) = 1/(t+1) + O(1/t^2)$ [2] The corresponding vectorial dynamics for $\mathbf{Z}_t^* = (Z_{t,1}^*, \dots, Z_{t,N}^*)^\top$ is

$$\begin{aligned} \mathbf{Z}_0^* &= \mathbf{1} \\ \mathbf{Z}_{t+1}^* &= \left(1 - \frac{1}{t+1}\right)\mathbf{Z}_t^* + \frac{1}{t+1}\Gamma^\top \mathbf{X}_{t+1}^* + O(1/t^2)\mathbf{1} \\ &= \left[1 - \frac{1}{t+1}(I - \Gamma^\top)\right]\mathbf{Z}_t^* + \frac{1}{t+1}\Gamma^\top \Delta\mathbf{M}_{t+1}^* + O(1/t^2)\mathbf{1} \quad \text{for } t \geq 0, \end{aligned} \tag{7}$$

where $\Delta\mathbf{M}_{t+1}^* = \mathbf{X}_{t+1}^* - \mathbf{Z}_t^*$. Now, we fix $x > 0$ and set

$$\zeta_0(x) = 1, \qquad \zeta_t(x) = \frac{\Gamma(t+x)}{\Gamma(t)} \sim t^x \uparrow +\infty.$$

More precisely, from [33, Lemma 4.1] we have

$$\zeta_t(x) = t^x + O(t^{x-1}) \quad \text{and} \quad \frac{1}{\zeta_t(x)} = \frac{1}{t^x} + O(1/t^{x+1}),$$

and so

$$\begin{aligned} \frac{1}{\zeta_{t+1}(x)\zeta_{t+1}(1-x)} &= \left(\frac{1}{(t+1)^x} + O(1/t^{x+1})\right)\left(\frac{1}{(t+1)^{1-x}} + O(1/t^{2-x})\right) \\ &= \frac{1}{t+1} + O(1/t^2). \end{aligned} \tag{8}$$

Hence, multiplying (7) by $\zeta_{t+1}(1-\gamma^*)$ and using the relation

$$\frac{\zeta_{t+1}(x)}{\zeta_t(x)} = \frac{\Gamma(t+x+1)}{\Gamma(t+1)}\frac{\Gamma(t)}{\Gamma(t+x)} = 1 + \frac{x}{t} = 1 + \frac{x}{t+1} + O(1/t^2),$$

[2] In all the paper, the notation $O(s_t)$ denotes a generic (possibly random) remainder term R_t such that $|R_t| \leq C s_t$ for a suitable deterministic constant C and for t large enough.

with $x = 1-\gamma^*$, we get the following dynamics for $\mathbf{Z}^{**}_t = \zeta_t(1-\gamma^*)\mathbf{Z}^*_t$, where we set $\Delta\mathbf{M}^{**}_{t+1} = \zeta_t(1-\gamma^*)\Delta\mathbf{M}^*_{t+1}$:

$$\begin{aligned}
\mathbf{Z}^{**}_0 &= \mathbf{1} \\
\mathbf{Z}^{**}_{t+1} &= \left[1 - \frac{1}{t+1}(I-\Gamma^\top)\right]\frac{\zeta_{t+1}(1-\gamma^*)}{\zeta_t(1-\gamma^*)}\mathbf{Z}^{**}_t + \frac{1}{t+1}\Gamma^\top\Delta\mathbf{M}^{**}_{t+1} + O\left(\frac{\zeta_{t+1}(1-\gamma^*)}{t^2}\right)\mathbf{1} \\
&= \left[1 - \frac{1}{t+1}(I-\Gamma^\top)\right]\left(1 + \frac{1-\gamma^*}{t+1} + O(1/t^2)\right)\mathbf{Z}^{**}_t + \frac{1}{t+1}\Gamma^\top\Delta\mathbf{M}^{**}_{t+1} + O\left(\frac{\zeta_{t+1}(1-\gamma^*)}{t^2}\right)\mathbf{1} \\
&= \mathbf{Z}^{**}_t - \frac{1}{t+1}(\gamma^* I - \Gamma^\top)\mathbf{Z}^{**}_t + \frac{1}{t+1}\Gamma^\top\Delta\mathbf{M}^{**}_{t+1} + O\left(\frac{\zeta_{t+1}(1-\gamma^*)}{t^2}\right)\mathbf{1}.
\end{aligned}$$

Moreover, recalling that $D^*_{t,h} = \sum_{n=1}^t X^*_{n,h}$, we find

$$\begin{aligned}
\frac{\mathbf{D}^*_0}{\zeta_0(\gamma^*)} &= \mathbf{0} \\
\frac{\mathbf{D}^*_{t+1}}{\zeta_{t+1}(\gamma^*)} &= \frac{\zeta_t(\gamma^*)}{\zeta_{t+1}(\gamma^*)}\frac{\mathbf{D}^*_t}{\zeta_t(\gamma^*)} + \frac{1}{\zeta_{t+1}(\gamma^*)}\mathbf{X}^*_{t+1} \\
&= \left(1 - \frac{\gamma^*}{t+1} + O(t^{-2})\right)\frac{\mathbf{D}^*_t}{\zeta_t(\gamma^*)} + \frac{1}{\zeta_{t+1}(\gamma^*)}\Delta\mathbf{M}^*_{t+1} + \frac{1}{\zeta_{t+1}(\gamma^*)}\mathbf{Z}^*_t \\
&= \left(1 - \frac{\gamma^*}{t+1} + O(t^{-2})\right)\frac{\mathbf{D}^*_t}{\zeta_t(\gamma^*)} + \frac{1}{\zeta_{t+1}(\gamma^*)\zeta_{t+1}(1-\gamma^*)}\Delta\mathbf{M}^{**}_{t+1} \\
&\quad + \frac{1}{\zeta_{t+1}(\gamma^*)\zeta_{t+1}(1-\gamma^*)}\frac{\zeta_{t+1}(1-\gamma^*)}{\zeta_t(1-\gamma^*)}\mathbf{Z}^{**}_t.
\end{aligned}$$

Using (8) and the relation $\zeta_{t+1}(x)/\zeta_t(x) = 1 + O(1/t)$, and combining together with the previous dynamics of $(\mathbf{Z}^{**}_t)_t$ and $(\mathbf{D}^*_{t+1}/\zeta_{t+1}(\gamma^*))_t$, we obtain

$$\begin{aligned}
\mathbf{Z}^{**}_{t+1} &= \mathbf{Z}^{**}_t - \frac{1}{t+1}(\gamma^* I - \Gamma^\top)\mathbf{Z}^{**}_t + \frac{1}{t+1}\Gamma^\top\Delta\mathbf{M}^{**}_{t+1} + O\big(\zeta_{t+1}(1-\gamma^*)/t^2\big)\mathbf{1} \\
\frac{\mathbf{D}^*_{t+1}}{\zeta_{t+1}(\gamma^*)} &= \frac{\mathbf{D}^*_t}{\zeta_t(\gamma^*)} - \frac{1}{t+1}\left(\gamma^*\frac{\mathbf{D}^*_t}{\zeta_t(\gamma^*)} - \mathbf{Z}^{**}_t\right) + \frac{1}{t+1}\Delta\mathbf{M}^{**}_{t+1} + O(\zeta_t(1-\gamma^*)/t^2)\mathbf{1},
\end{aligned}$$

where $O(\zeta_t(1-\gamma^*)/t^2) = O(1/t^{1+\gamma^*})$. Moreover, we recall that in [26, Theorem S1.1 and Lemma S1.2] (see also [8, Appendix B.1] for a general theory) it was proved that

$$\widetilde{Z}^{**}_t = \zeta_t(1-\gamma^*)\widetilde{Z}^*_t = \zeta_t(1-\gamma^*)\mathbf{v}^\top\mathbf{Z}^*_t \sim t^{1-\gamma^*}\mathbf{v}^\top\mathbf{Z}^*_t = t^{1-\gamma^*}\widetilde{Z}^*_t \xrightarrow{a.s./mean} \widetilde{Z}^{**}_\infty,$$

where $\widetilde{Z}^{**}_\infty$ is an integrable strictly positive random variable and $\mathbf{Z}^{**}_t \xrightarrow{a.s.} \widetilde{Z}^{**}_\infty\mathbf{u}$ so that $D^{**}_\infty = \widetilde{Z}^{**}_\infty/\gamma^*$. In addition, we have

$$\begin{aligned}
t^{-(1-\gamma^*)}E[\Delta\mathbf{M}^{**}_{t+1}(\Delta\mathbf{M}^{**}_{t+1})^\top \mid \mathcal{F}_t] &= \frac{\zeta_t(1-\gamma^*)^2}{t^{1-\gamma^*}}E[\Delta\mathbf{M}^*_{t+1}(\Delta\mathbf{M}^*_{t+1})^\top \mid \mathcal{F}_t] \\
&= \frac{\zeta_t(1-\gamma^*)^2}{t^{1-\gamma^*}}diag(\mathbf{Z}^*_t)(I - diag(\mathbf{Z}^*_t)) \\
&= \frac{\zeta_t(1-\gamma^*)}{t^{1-\gamma^*}}diag(\mathbf{Z}^{**}_t)(I - diag(\mathbf{Z}^*_t)) \xrightarrow{a.s.} \Sigma_\infty = \widetilde{Z}^{**}_\infty diag(\mathbf{u}).
\end{aligned}$$

Therefore, we can apply Theorem A.4 with $\mathbf{A}_t = \mathbf{Z}^{**}_t = \zeta_t(1-\gamma^*)\mathbf{Z}^*_t$, $\mathbf{B}_t = \frac{\mathbf{D}^*_t}{\zeta_t(\gamma^*)}$, $\Phi = \Gamma$, $\phi^* = \gamma^*$, $\Sigma_\infty = \widetilde{Z}^{**}_\infty diag(\mathbf{u})$, $\Delta\mathbf{M}_{t+1} = \Delta\mathbf{M}^{**}_{t+1} = \zeta_t(1-\gamma^*)\Delta\mathbf{M}^*_{t+1} = \zeta_t(1-\gamma^*)(\mathbf{X}^*_{t+1} - \mathbf{Z}^*_t)$. Finally, we can conclude by observing that, since $\gamma^* < 1$, we have

$$t^{\gamma^*/2}\left(\frac{\mathbf{D}^*_t}{t^{\gamma^*}} - \frac{\mathbf{D}^*_t}{\zeta_t(\gamma^*)}\right) = t^{\gamma^*/2}\mathbf{D}^*_t O\left(1/t^{\gamma^*+1}\right) = \frac{\mathbf{D}^*_t}{t^{\gamma^*}}O\left(1/t^{1-\gamma^*/2}\right) \xrightarrow{a.s.} \mathbf{0}.$$

Note that in Theorem A.4 we have the explicit formulas for the elements of the matrix $C_{det,\Gamma}$: that is, $C_{det,\Gamma} = 1/\gamma^*$ when $N = 1$ and

$$C_{det,\Gamma} = \frac{1}{\gamma^*}(\mathbf{v}^\top diag(\mathbf{u})\mathbf{v})\mathbf{u}\mathbf{u}^\top + \frac{1}{\widetilde{Z}^{**}_\infty}\mathcal{M}^{33}_\infty,$$

where $\mathcal{M}^{33}_\infty$ is the matrix defined in Theorem A.3 when $N \geq 2$ with the matrix Φ equal to Γ. Note that, since the expression for $\mathcal{M}^{33}_\infty$ contains the matrix Σ_∞, and so the term $\widetilde{Z}^{**}_\infty$, we have that the matrix $\frac{1}{\widetilde{Z}^{**}_\infty}\mathcal{M}^{33}_\infty$ is deterministic. □

Theorem 3.2. *Under Assumption 3.1, we have*

$$t^{\gamma^*/2}\left(\frac{\mathbf{D}^*_t}{\sum_{h=1}^N D^*_{t,h}} - \frac{\mathbf{u}}{\sum_{h=1}^N u_h}\right) \xrightarrow{stably} \mathcal{N}\left(\mathbf{0}, \frac{1}{D^{**}_\infty}\frac{\gamma^*}{(\sum_{j=1}^N u_j)^2}Q_\Gamma C_{det,\Gamma}Q_\Gamma^\top\right),$$

and

$$\sqrt{\sum_{h=1}^{N} D^*_{t,h}}\left(\frac{\mathbf{D}^*_t}{\sum_{h=1}^{N} D^*_{t,h}} - \frac{\mathbf{u}}{\sum_{h=1}^{N} u_h}\right) \xrightarrow{stably} \mathcal{N}\left(\mathbf{0}, \frac{\gamma^*}{\sum_{j=1}^{N} u_j} Q_\Gamma C_{det,\Gamma} Q_\Gamma^\top\right),$$

where $Q_\Gamma = I - \mathbf{u}\mathbf{1}^\top/\mathbf{1}^\top\mathbf{u}$ *and* $C_{det,\Gamma}$ *is defined in Theorem 3.1. Moreover, for each* j *we have*

$$t^{\gamma^*/2}\left(\frac{\mathbf{D}^*_t}{D^*_{t,j}} - \frac{\mathbf{u}}{u_j}\right) \xrightarrow{stably} \mathcal{N}\left(\mathbf{0}, \frac{1}{D^{**}_\infty}\frac{\gamma^*}{u_j^2} Q_j C_{det,\Gamma} Q_j^\top\right),$$

and

$$\sqrt{D^*_{t,j}}\left(\frac{\mathbf{D}^*_t}{D^*_{t,j}} - \frac{\mathbf{u}}{u_j}\right) \xrightarrow{stably} \mathcal{N}\left(\mathbf{0}, \frac{\gamma^*}{u_j} Q_j C_{det,\Gamma} Q_j^\top\right),$$

where $Q_j = I - \mathbf{u}\mathbf{e}_j^\top/u_j$ *with* $\mathbf{e}_j$ *the* j*th vector of the canonical base and* $C_{det,\Gamma}$ *is defined in Theorem 3.1.*

Proof. We have to apply Corollary A.6 reported in the appendix with $\mathbf{A}_t = \mathbf{Z}^{**}_t = \zeta_t(1-\gamma^*)\mathbf{Z}^*_t$, $\mathbf{B}_t = \frac{\mathbf{D}^*_t}{\zeta_t(\gamma^*)}$, $\mathbf{x} = \mathbf{1}$, $\widetilde{A}_\infty = \widetilde{Z}^{**}_\infty = \gamma^* D^{**}_\infty$ and $\phi^* = \gamma^*$. Note that, as specified in the proof of this corollary, we have $\widetilde{Z}^{**}_\infty Q_\Gamma C_{det,\Gamma} Q_\Gamma^\top = Q_\Gamma \mathcal{M}^{33}_\infty Q_\Gamma^\top$, where $\mathcal{M}^{33}_\infty$ is the matrix defined in Theorem A.3 when $N \geq 2$ with the matrix Φ equal to Γ. Similarly, we obtain the second part of the statement by means of Corollary A.6 with $\mathbf{x} = \mathbf{e}_j$. □

Theorem 3.3. *Under Assumption 3.2, we have for a given observed color* c *in the system*

$$\sqrt{t}\left(\frac{\mathbf{K}_t(c)}{t} - \widetilde{P}_\infty(c)\mathbf{1}\right) \xrightarrow{stably} \mathcal{N}(\mathbf{0}, \widetilde{P}_\infty(c)(1-\widetilde{P}_\infty(c))C_{det,W}),$$

where $\widetilde{P}_\infty(c) = a.s. - \lim_t P_t(h,c) = a.s. - \lim_t K_t(h,c)/t$ *for each* $h = 1,\dots,N$, $\widetilde{P}_\infty(c) \in (0,1)$ *a.s. and* $C_{det,W}$ *is a deterministic matrix depending on the eigen-structure of* W*, which is possible to write explicitly.*

Proof. For a given color c, let $t^*(c)$ be the random time-step of the first extraction of c in the system. Conditionally on $\{t^*(c) < +\infty\}$, for $t \geq t^*(c)$ we can write

$$P_t(h,c) = \frac{\sum_{n=1}^{t}\sum_{j=1}^{N} w_{j,h}\Delta K_n(j,c)}{\theta_h + t} - \frac{\gamma_{j^*(c),h}}{\theta_h + t},$$

where $\Delta K_n(j,c) = K_n(j,c) - K_{n-1}(j,c)$ takes values in $\{0,1\}$ and $E[\Delta K_{n+1}(j,c)|\mathbf{C}_1,\dots,\mathbf{C}_n] = P_n(j,c)$. Then, we obtain the following dynamics for $P_t(h,c)$:

$$P_{t+1}(h,c) = (1-r_{t,h})P_t(h,c) + r_{t,h}\sum_{j=1}^{N} w_{j,h}\Delta K_{t+1}(j,c),$$

where $r_{t,h} = 1/(\theta_h + t + 1) = 1/(t+1) + O(1/t^2)$. Thus the corresponding vectorial dynamics for $\mathbf{P}_t(c) = (P_t(1,c),\dots,P_t(N,c))^\top$ is

$$\begin{aligned}
&\mathbf{P}_{t^*(c)}(c) \neq \mathbf{0},\\
&\mathbf{P}_{t+1}(c) = \left(1 - \frac{1}{t+1}\right)\mathbf{P}_t(c) + \frac{1}{t+1}W^T\Delta\mathbf{K}_{t+1}(c) + O(1/t^2)\mathbf{1}\\
&\qquad = \mathbf{P}_t(c) - \frac{1}{t+1}(I - W^\top)\mathbf{P}_t(c) + \frac{1}{t+1}W^\top\Delta\mathbf{M}_{t+1}(c) + O(1/t^2)\mathbf{1},
\end{aligned}$$

where $\Delta\mathbf{K}_t(c) = (\Delta K_t(1,c),\dots,\Delta K_t(N,c))^\top$ and $\Delta\mathbf{M}_{t+1}(c) = \Delta\mathbf{K}_{t+1}(c) - \mathbf{P}_t(c)$.

Considering also the dynamics of $\mathbf{K}_t(c)/t$, we obtain

$$\begin{aligned}
&\mathbf{P}_{t+1}(c) = \mathbf{P}_t(c) - \frac{1}{t+1}(I - W^\top)\mathbf{P}_t(c) + \frac{1}{t+1}W^\top\Delta\mathbf{M}_{t+1}(c) + O(1/t^2)\mathbf{1}\\
&\frac{\mathbf{K}_{t+1}(c)}{t+1} = \frac{\mathbf{K}_t(c)}{t} - \frac{1}{t+1}\left(\frac{\mathbf{K}_t(c)}{t} - \mathbf{P}_t(c)\right) + \frac{1}{t+1}\Delta\mathbf{M}_{t+1}(c).
\end{aligned}$$

Moreover, from Theorem 2.2, we have $\widetilde{P}_t(c) = \mathbf{v}_W^\top\mathbf{P}_t(c) \xrightarrow{a.s.} \widetilde{P}_\infty(c)$, and this convergence is also in mean because $(\widetilde{P}_t(c))_t$ is uniformly bounded. In addition, we have

$$E[\Delta\mathbf{M}_{t+1}(\Delta\mathbf{M}_{t+1})^\top \mid \mathcal{F}_t] = diag(\mathbf{P}_t)(I - diag(\mathbf{P}_t)) \xrightarrow{a.s.} \Sigma_\infty = \widetilde{P}_\infty(c)(1-\widetilde{P}_\infty(c))I.$$

Therefore, we can apply Theorem A.4 with $\mathbf{A}_t = \mathbf{P}_t(c)$, $\mathbf{B}_t = \mathbf{K}_t(c)/t$, $\Phi = W$, $\phi^* = w^* = 1$, $\mathbf{u} = \mathbf{u}_W = \mathbf{1}$, $\mathbf{v} = \mathbf{v}_W$, $\Sigma_\infty = \widetilde{P}_\infty(c)(1-\widetilde{P}_\infty(c))I$, $\Delta\mathbf{M}_{t+1} = \Delta\mathbf{M}_{t+1}(c) = \Delta\mathbf{K}_{t+1}(c) - \mathbf{P}_t(c)$. Note that in Theorem A.4 we have the explicit formulas for the elements of the matrix $C_{det,W}$: that is, $C_{det,W} = 1$ when $N = 1$ and

$$C_{det,W} = (\mathbf{v}_W^\top\mathbf{v}_W)\mathbf{1}\mathbf{1}^\top + \frac{1}{\widetilde{P}_\infty(c)(1-\widetilde{P}_\infty(c))}\mathcal{M}^{33}_\infty,$$

where $\mathcal{M}^{33}_\infty$ is the matrix defined in Theorem A.3 when $N \geq 2$ with the matrix Φ equal to W. Note that, since the expression for $\mathcal{M}^{33}_\infty$ contains the matrix Σ_∞, and so the factor $\widetilde{P}_\infty(c)(1-\widetilde{P}_\infty(c))$, we have that the matrix $\frac{1}{\widetilde{P}_\infty(c)(1-\widetilde{P}_\infty(c))}\mathcal{M}^{33}_\infty$ is deterministic. □

Theorem 3.4. *Under Assumption 3.2, we have for a given observed color c in the system*

$$\sqrt{t}\left(\frac{\mathbf{K}_t(c)}{\sum_{h=1}^N K_t(h,c)}-\frac{1}{N}\mathbf{1}\right)\xrightarrow{stably}\mathcal{N}\left(\mathbf{0},\frac{(1-\widetilde{P}_\infty(c))}{N^2\widetilde{P}_\infty(c)}Q_W C_{det,W}Q_W^\top\right),$$

and

$$\sqrt{\frac{\sum_{h=1}^N K_t(h,c)}{1-(Nt)^{-1}\sum_{h=1}^N K_t(h,c)}}\left(\frac{\mathbf{K}_t(c)}{\sum_{h=1}^N K_t(h,c)}-\frac{1}{N}\mathbf{1}\right)\xrightarrow{stably}\mathcal{N}\left(\mathbf{0},\frac{1}{N}Q_W C_{det,W}Q_W^\top\right).$$

where $Q_W=(I-N^{-1}\mathbf{1}\mathbf{1}^\top)$ and $C_{det,W}$ is defined in Theorem 3.3.

Proof. We have to apply Corollary A.6 reported in the appendix with $\mathbf{A}_t=\mathbf{P}_t(c)$, $\mathbf{B}_t=\mathbf{K}_t(c)/t$, $\mathbf{x}=\mathbf{1}$, $\widetilde{A}_\infty=\widetilde{P}_\infty(c)$, $\phi^*=w^*=1$ and $\mathbf{u}=\mathbf{u}_W=\mathbf{1}$. Note that, as specified in the proof of this corollary, we have $\widetilde{P}_\infty(c)(1-\widetilde{P}_\infty(c))Q_W C_{det,W}Q_W^\top=Q_W\mathcal{M}_\infty^{33}Q_W^\top$, where $\mathcal{M}_\infty^{33}$ is the matrix defined in Theorem A.3 with the matrix Φ equal to W. $\square$

Remark 3.1. (Discussion of the Assumption 3.1 and Assumption 3.2)
As observed in Remark A.1 in the appendix, removing the diagonalizability of the matrices Γ and W leads to analogous results (cf. [34] for block upper-triangular matrices). Moreover, in order to deal with reducible matrices, the directed graph must be decomposed into communicating classes. The leading classes, their Perron-Frobenius eigenvalues, and the accessibility relations between blocks determine the first-order rates (we refer to [26,28,35]). A second-order theory would require additional block-wise fluctuation analysis (e.g., [34] for block upper-triangular matrices). Finally, when $\mathcal{R}e(\gamma_2^*)/\gamma^*\geq 1/2$, as highlighted in Remark A.2 in appendix, the rate of convergence is slower than $t^{\gamma^*/2}$. Indeed, similarly to the stochastic approximation literature (e.g. [36]), for the case $\mathcal{R}e(\gamma_2^*)/\gamma^*=1/2$ we expect a CLT with a logarithmic correction in the scaling; while in the case $\mathcal{R}e(\gamma_2^*)/\gamma^*>1/2$, a power-law rate with an exponent smaller than $\gamma^*/2$ could hold true, but not a Gaussian limit. Analogous considerations can be done for the cases $\mathcal{R}e(w_2^*)=1/2$ and $\mathcal{R}e(w_2^*)>1/2$.

4. Statistical inference

We here present some statistical tools based on the results of Section 3. More precisely, we provide a statistical test for each of the two interaction matrices, Γ and W, based on the observable processes $\mathbf{D}_t^*$ and $\mathbf{K}_t(c)$. Furthermore, we use the latter process to provide also a confidence interval for the limit probability $\widetilde{P}_\infty(c)$. The general procedure to be followed in order to obtain these statistical instruments is described in Section B and Section C. In particular, in our framework, we have to apply it with $\mathbf{B}_t=\frac{\mathbf{D}_t^*}{\zeta_t(\gamma^*)}$ (or, equivalently $\mathbf{B}_t=\frac{\mathbf{D}_t^*}{t^{\gamma^*}}$) and with $\mathbf{B}_t=\mathbf{K}_t(c)/t$. To this end, we observe that Assumption B.1 is true in both cases with $\Sigma_{det}=\Sigma_{det,\Gamma}=diag(\mathbf{u})$ in the first one and $\Sigma_{det}=\Sigma_{det,W}=I$ in the second one.

We firstly focus on the case $N=2$ and, then, we consider a particular type of interaction (mean-field interaction) in the general case $N\geq 2$. It is worthwhile to underline that the statistical tools described in Section B are very general and so it is also possible to explicit the formulas for any family of interaction matrices satisfying the required assumptions.

4.1. The case $N=2$

4.1.1. Hypothesis test on Γ

In what follows we assume the parameters $\gamma^*\in(0,1)$ and $r=u_1/u_2>0$ to be known.

We use the re-parametrization of the matrix Γ given by the constrains

$$r\gamma_{1,1}+\gamma_{2,1}=r\gamma^* \quad\text{and}\quad r\gamma_{1,2}+\gamma_{2,2}=\gamma^*.$$

and make inference on the other two parameters $\gamma_{1,2}$ and $\gamma_{2,1}$ (that have to be such that the assumptions on the matrix Γ are verified). If we use the parametrization $\iota=\gamma_{1,2}+\gamma_{2,1}$ and $\eta=\frac{\gamma_{1,2}}{\gamma_{1,2}+\gamma_{2,1}}$ so that ι rules the intensity of the whole interaction between the two processes due to matrix Γ, while η (resp. $(1-\eta)$) denotes the percentage of this intensity due to the influence of process 1 on process 2 (resp., of process 2 on process 1), the expressions for the elements of the matrix Γ become

$$\gamma_{1,1}=\gamma^*-\frac{(1-\eta)\iota}{r},\quad \gamma_{1,2}=\eta\iota,\quad \gamma_{2,1}=(1-\eta)\iota\quad\text{and}\quad\gamma_{2,2}=\gamma^*-r\eta\iota, \tag{9}$$

where (by the required conditions on Γ)

$$\eta\in(0,1)\quad\text{and}\quad\iota\in J_\eta\quad\text{with,}$$
$$J_\eta=\left(0,\min\left(\tfrac{\gamma^* r}{1-\eta},\tfrac{\gamma^*}{r\eta}\right)\right]\cap\left(0,\tfrac{1-\gamma^*}{\eta(1-r)}I_{(r<1)}+\tfrac{1-\gamma^*}{(1-\eta)(1-1/r)}I_{(r>1)}+\infty I_{(r=1)}\right). \tag{10}$$

Two-sided test. If we take the null hypothesis $H_0:\ \iota=\iota_0,\ \eta=\eta_0$, where ι_0 and η_0 satisfy (10), i.e. $\iota_0\in J_{\eta_0}$, and the required condition in Assumption 3.1 for the second eigenvalue of Γ, that is $\iota_0\frac{(1-\eta_0)+r^2\eta_0}{r}>\gamma^*/2$, then we can use Theorem 3.1 and obtain the test statistic (see (C.1) from Remark C.1 with $\phi = \gamma^*$, $\phi_2^* = \gamma_2^* = \gamma^* - [(1-\eta_0)\iota_0 + r^2\eta_0\iota_0]/r$, $g(x) = x$, and $\mathbf{B}_t = \frac{\mathbf{D}_t^*}{\zeta_t(\gamma^*)}$ or, equivalently $\mathbf{B}_t = \frac{\mathbf{D}_t^*}{t\gamma^*}$)

$$q_{r,\eta_0}\Delta_0 \frac{(D^*_{1,t} - rD^*_{2,t})^2}{\widetilde{D}^*_t} \xrightarrow{d} \chi^2(1) \quad \text{under } H_0\,, \tag{11}$$

where $\widetilde{D}^*_t = \mathbf{v}^\top \mathbf{D}^*_t = \frac{\eta_0 r D^*_{1,t} + (1-\eta_0)D^*_{2,t}}{\eta_0 r + (1-\eta_0)}$, $q_{r,\eta_0} = \frac{2[(1-\eta_0)+\eta_0 r^2]}{r(1+r)[(1-\eta_0)+\eta_0 r]}$ and

$$\Delta_0 = \frac{1}{2} - \frac{\gamma_2^*}{\gamma^*} = \frac{\iota_0}{\gamma^*}\frac{(1-\eta_0)+\eta_0 r^2}{r} - \frac{1}{2}\,.$$

It is immediate to see that the test statistic (11) has an increasing dependence on the parameter ι_0, while its behaviour with respect the parameter η_0 is not clear. However, we can observe that it can be rewritten as

$$\frac{2}{r(1+r)}\frac{(1-\eta_0)+\eta_0 r^2}{(1-\eta_0)+\eta_0 r D^*_{1,t}/D^*_{2,t}}\Delta_0 D^*_{2,t}(D^*_{1,t}/D^*_{2,t} - r)^2\,,$$

where, for any value of η_0, we have

$$\frac{(1-\eta_0)+\eta_0 r^2}{(1-\eta_0)+\eta_0 r D^*_{1,t}/D^*_{2,t}} \xrightarrow{a.s.} 1\,.$$

Hence, we can simplify (11) as

$$\tfrac{2}{r(1+r)}\Delta_0 D^*_{2,t}(D^*_{1,t}/D^*_{2,t} - r)^2 \xrightarrow{d} \chi^2(1) \quad \text{under } H_0\,, \tag{12}$$

where the dependence on ι_0 and η_0 is only in the factor Δ_0. Notice also that in the present framework the two eigenvalues γ^* and γ_2^* are both real and so the quantity Δ_0 represents the distance of the ratio between the second and the leading eigenvalue from the critical value $1/2$ (see condition $\mathcal{R}e(\gamma_2^*)/\gamma^* < 1/2$ in Assumption 3.1). Moreover, Δ_0 is the only quantity in the test statistics (12) that contains the parameters ι_0 and η_0. Moreover, it is neither random nor time-dependent. The random quantity $D^*_{2,t}(D^*_{1,t}/D^*_{2,t} - r)^2 = [\sqrt{D^*_{2,t}}(D^*_{1,t}/D^*_{2,t} - r)]^2$ describes the fluctuations of $D^*_{1,t}/D^*_{2,t}$ around r (note that (12) is coherent with the convergence result stated in Theorem 3.2).

One-sided test. Since Δ_0, and so the statistic (12), is increasing in ι_0, we can replace the condition $\iota = \iota_0$ in H_0 by $\iota \geq \iota_0$. Further, the quantity Δ_0 is increasing in η_0 when $r > 1$ and decreasing in η_0 when $r < 1$. Hence, the statistic (12) works well also if we replace the condition $\eta = \eta_0$ in H_0 by $\eta \geq \eta_0$ when $r > 1$ or by $\eta \leq \eta_0$ when $r < 1$. Therefore, if we set $\eta_0 = 1/2$, we get

$$\Delta_0 = \frac{1}{2} - \frac{\gamma_2^*}{\gamma^*} = \frac{\iota_0}{\gamma^*}\frac{1+r^2}{2r} - \frac{1}{2}$$

and we can perform a one-sided test with one of the following two null hypotheses, according to the value of r:

- Case $r > 1$: $H_0 : \iota \geq \iota_0, \eta \geq 1/2$ (the interaction intensity is at least ι_0 and the processes equally influence each other or the interaction is mostly due to the influence of process 1 on process 2), where ι_0 satisfies (10) for some η in H_0, i.e. $\iota_0 \in \cup_{\eta\geq 1/2} J_\eta$, and $\iota_0 > \gamma^* \frac{r}{1+r^2}$;
- Case $r < 1$: $H_0 : \iota \geq \iota_0, \eta \leq 1/2$ (the interaction intensity is at least ι_0 and the processes equally influence each other or the interaction is mostly due to the influence of process 2 on process 1), where ι_0 satisfies (10) for some η in H_0, i.e. $\iota_0 \in \cup_{\eta\leq 1/2} J_\eta$, and $\iota_0 > \gamma^* \frac{r}{1+r^2}$.

Finally, in the very special case when $r = 1$, the test statistic (12) does not depend on the value of η_0 and it can be used for testing the intensity ι only.

Power. Following Remark B.1, we find that, under the alternative $H_1 : \iota = \iota_1, \eta = \eta_1$, with $\iota_1 \in J_{\eta_1}$ and $\iota_1\frac{(1-\eta_1)+r^2\eta_1}{r} > \gamma^*/2$, the statistic (12) is asymptotically distributed as

$$\frac{\Delta_0}{\Delta_1}\chi^2(1) = \frac{\frac{\iota_0}{\gamma^*}\frac{(1-\eta_0)+\eta_0 r^2}{r} - \frac{1}{2}}{\frac{\iota_1}{\gamma^*}\frac{(1-\eta_1)+\eta_1 r^2}{r} - \frac{1}{2}}\chi^2(1)\,.$$

Notice that the term $\frac{\Delta_0}{\Delta_1}$ is time-invariant and this means that increasing the time t does not increase the power. Nevertheless, the power increases as η_1 decreases if $r > 1$ or as η_1 increases if $r < 1$. In addition, the power improves as the value of ι_1 decreases and it tends toward 1 as $\Delta_1 \to 0$, that is as ι_1 approaches its minimum admissible value $\frac{r}{(1-\eta_1)+\eta_1 r^2}\frac{\gamma^*}{2}$. Moreover, when ι_1 is even lower or equal to $\frac{r}{(1-\eta_1)+\eta_1 r^2}\frac{\gamma^*}{2}$, i.e. when condition $\mathcal{R}e(\gamma_2^*)/\gamma^* < 1/2$ does not hold, the power is asymptotically equal to 1 because the test statistic (12) diverges to infinity as a consequence of Remark 3.1.

4.1.2. Hypothesis test on W

As $W^\top \mathbf{1} = \mathbf{1}$ (i.e. $w^* = 1$ and $r = 1$), we have

$$w_{1,1} = 1 - w_{2,1} \quad \text{and} \quad w_{2,2} = 1 - w_{1,2}$$

and so we can make inference on the two parameters $w_{2,1}$ and $w_{1,2}$. Using the parametrization $\imath = w_{1,2} + w_{2,1}$ for the intensity of the interaction due to the matrix W, and $\eta = \frac{w_{1,2}}{w_{1,2}+w_{2,1}}$ (resp. $(1-\eta)$) for the percentage of this intensity due to the influence of process 1 on process 2 (resp., of process 2 on process 1), we have

$$w_{1,1} = 1 - (1-\eta)\imath, \qquad w_{1,2} = \eta\imath, \qquad w_{2,1} = (1-\eta)\imath \qquad \text{and} \qquad w_{2,2} = 1 - \eta\imath\,, \tag{13}$$

where (by the required conditions on W)

$$\eta \in (0,1) \qquad \text{and} \qquad \imath \in \left(0, \min(1/\eta, 1/(1-\eta)\right) \subseteq (0,2)\,. \tag{14}$$

Two-sided test. If we take the null hypothesis $H_0\,:\ \imath = \imath_0\,,\ \eta = \eta_0$, where (14) and $\imath_0 > 1/2$ (the required condition in Assumption 3.2 for the second eigenvalue of W) are satisfied, then we can use Theorem 3.3 and obtain the test statistic (see (C.1) from Remark C.1 with $\phi = w^* = 1$, $r = 1$, $\phi_2^* = 1 - \imath_0$, $g(x) = x(1-x)$ and $\mathbf{B}_t = \mathbf{K}_t(c)/t$ with c an arbitrary old color)

$$\begin{aligned}
&\Delta_0\,\frac{(K_{1,t}(c) - K_{2,t}(c))^2}{\widetilde{K}_t(c)(1-\widetilde{K}_t(c)/t)} \xrightarrow{d} \chi^2(1) \quad \text{under } H_0\,,\\
&\text{where} \qquad \Delta_0 = \frac{1}{2} - (1-\imath_0) = \imath_0 - \tfrac{1}{2},
\end{aligned} \tag{15}$$

and

$$\widetilde{K}_t(c) = \mathbf{v}^\top \mathbf{K}_t(c) = \eta_0 K_{1,t}(c) + (1-\eta_0)K_{2,t}(c). \tag{16}$$

Since $K_{j,t}(c)/t \xrightarrow{a.s.} \widetilde{P}_\infty(c)$, we obtain for any η_0 and η_1

$$\frac{\eta_1 K_{1,t}(c)/t + (1-\eta_1)K_{2,t}(c)/t}{\widetilde{K}_t(c)/t}\ \frac{(1-\eta_1 K_{1,t}(c)/t + (1-\eta_1)K_{2,t}(c)/t)}{(1-\widetilde{K}_t(c)/t)} \xrightarrow{a.s.} 1\,,$$

and so the choice of the value of η_0 in H_0 does not affect the asymptotic distribution of the test statistic. This means that we can use the test statistic (15) with an arbitrary value of η_0 in order to perform a test on the intensity $\imath$. In particular, if we choose $\eta_0 = 1/2$, we simply have $\widetilde{K}_t(c) = \frac{K_{1,t}(c)+K_{2,t}(c)}{2}$.

One-sided test. The statistic (15) is increasing in $\imath_0$ and so it works well also for a one-sided test with the condition $\imath = \imath_0$ in H_0 replaced by $\imath \geq \imath_0$, where $\imath_0 \in (\frac{1}{2}, 2)$.

Power. Following Remark B.1, we find that, under the alternative $H_1\,:\ \imath = \imath_1$, where $\imath_1 \in (\frac{1}{2}, 2)$ with $\imath_1 \neq \imath_0$, the statistic (15) is asymptotically distributed as

$$\frac{\Delta_0}{\Delta_1}\chi^2(1) = \frac{\imath_0 - \frac{1}{2}}{\imath_1 - \frac{1}{2}}\,\chi^2(1)\,.$$

Notice that the term $\frac{\Delta_0}{\Delta_1}$ is time-invariant and so, as before, increasing the time t does not increase the power. However, the power is monotone decreasing in $\imath_1$, and it tends to 1 as $\imath_1$ approaches its lower bound $1/2$, and it is equal to one for $\imath_1 \leq 1/2$ by Remark 3.1.

4.1.3. Confidence interval for $\widetilde{P}_\infty(c)$

Fix the confidence level $(1-\alpha)$ and denote by z_α the quantile of order $(1-\alpha/2)$ of the standard normal distribution, that is $\mathcal{N}(0,1)(z_\alpha, +\infty) = \alpha/2$. Assuming $\imath > 1/2$, we can use Theorem 3.3 in order to construct a confidence interval for the random limit $\widetilde{P}_\infty(c)$ associated to a given item c (see the interval (C.2) from Remark C.1 with $\phi = w^* = 1$, $r = 1$, $g(x) = x(1-x)$ and $\mathbf{B}_t = \mathbf{K}_t(c)/t$). Precisely, we obtain:

$$CI_{1-\alpha}(\widetilde{P}_\infty(c)) \ = \ \frac{\widetilde{K}_t(c)}{t} \pm \frac{z_\alpha}{t^{1/2}}\sqrt{\frac{\widetilde{K}_t(c)}{t}\left(1 - \frac{\widetilde{K}_t(c)}{t}\right)c_\eta}, \tag{17}$$

with $c_\eta = \eta^2 + (1-\eta)^2$ and $\widetilde{K}_t(c)$ defined as in (16). It is interesting to notice the absence of the dependence on the specific value of $\imath$. On the other hand, while the dependence of $\widetilde{K}_t(c)$ on η is not important, because $K_{j,t}(c) \xrightarrow{a.s.} \widetilde{P}_\infty(c)$ so that $\frac{\eta_1 K_{1,t}(c)/t + (1-\eta_1)K_{2,t}(c)/t}{\widetilde{K}_t(c)/t} \xrightarrow{a.s.} 1$, the presence of the quantity c_η implies that we need to know the value of the parameter η. However, c_η is always smaller than 1 for $\eta \in (0,1)$ and hence, if we do not know the value of η, we can use $\widetilde{K}_t(c)$ with $\eta = 1/2$, so that we have $\widetilde{K}_t(c) = \frac{K_{1,t}(c)+K_{2,t}(c)}{2}$, and take the largest interval obtained from (17) with c_η replaced by 1.

As an example, we have simulated $S = 200$ independent pair of innovation processes with Γ and W defined with $r = 0.75$, $\gamma^* = 0.75$, $\eta = 1/2$, $\imath_\Gamma = 1$ and $\imath_W = 1.25$ and then, for any simulation, we have constructed the confidence intervals (17) of level $(1-\alpha) = 95\%$

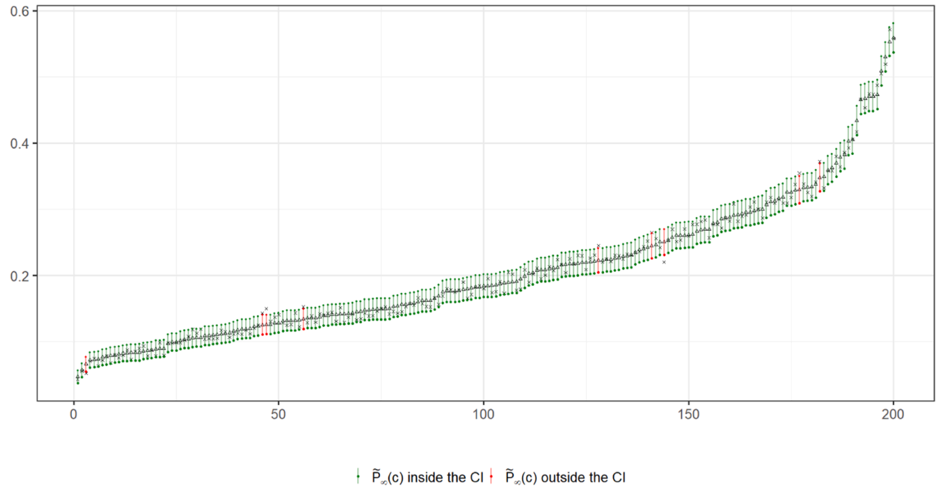

Fig. 1. Confidence interval (see (17)) for the random variable $\widetilde{P}_\infty(c)$ at time-step $t = 10^3$ and related to item c with highest $\mathbf{1}^\top \mathbf{K}_t(c)$ in the case $N = 2$, Γ defined as in (9) with $r = 0.75$, $\gamma^* = 0.75$, $\eta = 1/2$ and $\iota_\Gamma = 1$, and W defined as in (13) with $\eta = 1/2$ and $\iota_W = 1.25$. Number of simulations $S = 200$ (sorted on the x-axis according to the value of $\mathbf{1}^\top \mathbf{K}_t(c)$). The crosses represent the value (reported on the y-axis) of $\widetilde{P}_\infty(c)$ estimated by $\mathbf{1}^\top \mathbf{K}_{t_\infty}(c)/(N t_\infty)$ at time-step $t_\infty = 10^5$.

computed at time-step $t = 10^3$ associated to the item c with highest $(K_{t,1}(c) + K_{t,2}(c))$ (recall that the interval has the same expression for any given item c). In Fig. 1 we can see these $S = 200$ confidence intervals (sorted according to the value of $K_{t,1}(c) + K_{t,2}(c)$) and the estimates of the corresponding $\widetilde{P}_\infty(c)$ computed as $\frac{K_{t_\infty,1}(c)+K_{t_\infty,2}(c)}{2t_\infty}$ at time-step $t_\infty = 10^5$. The empirical coverage is represented by the proportion of times the estimates of $\widetilde{P}_\infty(c)$ belong to the corresponding confidence interval, and in this simulation this is equal to 0.955, which is very close to the nominal value of 0.95.

4.2. Mean-field interaction for $N \geq 2$

4.2.1. Hypothesis test on Γ

Suppose that Γ is of the form

$$\gamma_{j,h} = \phi\big(\iota/N + \delta_{j,h}(1-\iota)\big) \qquad \phi, \iota \in (0,1]\,. \tag{18}$$

We refer to this kind of interaction matrix as the mean-field interaction because in this case the term $\sum_{j=1}^N \gamma_{j,h} D^*_{t,j}$ in (4) becomes $\phi\Big(\iota \frac{1}{N}\sum_{j=1}^N D^*_{t,j} + (1-\iota) D^*_{t,h}\Big)$ and so the conditional probability $Z^*_{t,h}$ depends on a convex combination between $D^*_{t,h}$ and the averaged value of the all $D^*_{t,j}$ in the system. Note that, for this matrix Γ, the leading eigenvalue γ^* coincides with the parameter ϕ.

We can make inference on the parameter ι that rules the intensity of the interaction.

Two-sided test. If we take the null hypothesis $H_0 : \iota = \iota_0$, where $\iota_0 \leq 1$ and $\iota_0 > 1/2$ (the required condition in Assumption 3.1 for the second eigenvalue of Γ), then we can use Theorem 3.1 and obtain the test statistic (see (C.4) from Remark C.2 with $\phi = \gamma^*$, $\phi^*_2 = \gamma^*(1-\iota_0)$, $g(x) = x$, and $\mathbf{B}_t = \frac{\mathbf{D}^*_t}{\zeta_t(\gamma^*)}$ or, equivalently $\mathbf{B}_t = \frac{\mathbf{D}^*_t}{t\gamma^*}$)

$$\begin{aligned} &2\Delta_0 \frac{\left\|\mathbf{D}^*_t - \widetilde{D}^*_t \mathbf{1}\right\|^2}{\widetilde{D}^*_t} \xrightarrow{d} \chi^2(N-1) \quad \text{under } H_0,\\ &\text{where} \quad \Delta_0 = \frac{1}{2} - \frac{\gamma^*(1-\iota_0)}{\gamma^*} = \iota_0 - \frac{1}{2} \end{aligned} \tag{19}$$

and $\widetilde{D}^*_t = \mathbf{v}^\top \mathbf{D}^*_t = \frac{\mathbf{1}^\top \mathbf{D}^*_t}{N}$. Note that this test statistic does not depend on the parameter $\phi = \gamma^*$ and so the value of this parameter is not needed for the test on ι.

One-sided test. Since the test statistic (19) is increasing in ι_0, it works well also for the one-sided test with $H_0 : \iota \geq \iota_0$, where $\iota_0 \in (1/2, 1]$.

Power. Following Remark B.1, we observe that, under the alternative $H_1 : \iota = \iota_1$, with $\iota_1 \in (1/2, 1]$, the statistics (19) is asymptotically distributed as

$$\frac{\Delta_0}{\Delta_1}\chi^2(N-1) = \frac{\iota_0 - \frac{1}{2}}{\iota_1 - \frac{1}{2}}\chi^2(N-1)\,.$$

Therefore, as before, increasing the time t does not increase the power. Moreover, the power is monotone decreasing in ι_1, and it tends to 1 as ι_1 approaches its lower bound $1/2$, being equal to 1 for $\iota_1 \leq 1/2$ by Remark 3.1.

4.2.2. Hypothesis test on W

Suppose that W is of the form

$$w_{j,h} = \iota/N + \delta_{j,h}(1-\iota) \qquad \iota \in (0,1]\,. \tag{20}$$

Two-sided test. If we take the null hypothesis $H_0\,:\,\iota = \iota_0$, where $\iota_0 \leq 1$ and $\iota_0 > 1/2$ (the required condition in Assumption 3.2 for the second eigenvalue of W), then we can use Theorem 3.3 and obtain the test statistic (see (C.4) from Remark C.2 with $\phi = w^* = 1$, $\phi_2^* = w^*(1-\iota_0) = (1-\iota_0)$, $g(x) = x(1-x)$, and $\mathbf{B}_t = \mathbf{K}_t(c)/t$)

$$\begin{aligned} &2\Delta_0 \frac{\left\|\mathbf{K}_t(c) - \widetilde{K}_t(c)\mathbf{1}\right\|^2}{\widetilde{K}_t(c)(1-\widetilde{K}_t(c)/t)} \xrightarrow{d} \chi^2(N-1) \quad \text{under } H_0, \\ &\text{where} \qquad \Delta_0 = \frac{1}{2} - (1-\iota_0) = \iota_0 - \frac{1}{2}, \end{aligned} \tag{21}$$

c is an arbitrary observed item and $\widetilde{K}_t(c) = \mathbf{v}^\top \mathbf{K}_t(c) = \frac{\mathbf{1}^\top \mathbf{K}_t(c)}{N}$.

One-sided test. The test statistic (21) is increasing in ι_0 and so it works well also for the one-sided test with $H_0\,:\,\iota \geq \iota_0$, where $\iota_0 \in (1/2, 1]$.

Power. Following Remark B.1, we observe that, under the alternative $H_1\,:\,\iota = \iota_1$, with $\iota_1 \in (1/2, 1]$, the statistics (21) is asymptotically distributed exactly as the previous statistics (19) and so the same considerations on the power of the test hold true.

4.2.3. Confidence interval for $\widetilde{P}_\infty(c)$

Fix the confidence level $(1-\alpha)$ and denote by z_α the quantile of order $(1-\alpha/2)$ of the standard normal distribution, that is $\mathcal{N}(0,1)(z_\alpha, +\infty) = \alpha/2$. Assuming $\iota \in (1/2, 1]$, we can use Theorem 3.3 in order to construct a confidence interval for the random limit $\widetilde{P}_\infty(c)$ associated to a given item c (see the interval (C.3) from Remark C.2 with $\phi = w^* = 1$, $g(x) = x(1-x)$ and $\mathbf{B}_t = \mathbf{K}_t(c)/t$). Precisely, we obtain:

$$CI_{1-\alpha}(\widetilde{P}_\infty(c)) \;=\; \frac{\widetilde{K}_t(c)}{t} \pm \frac{z_\alpha}{t^{1/2}} \sqrt{\frac{\widetilde{K}_t(c)}{t}\left(1 - \frac{\widetilde{K}_t(c)}{t}\right)\frac{1}{N}}, \tag{22}$$

where $\widetilde{K}_t(c) = \mathbf{v}^\top \mathbf{K}_t(c) = \frac{\mathbf{1}^\top \mathbf{K}_t(c)}{N}$. Note that this interval does not depend on the value of the parameter ι.

As an example, we have simulated $S = 200$ independent triplets of innovation processes with Γ and W of the mean-field type defined with $N = 3$, $\gamma^* = 0.75$, $\iota_\Gamma = \iota_W = 0.8$, and we show the confidence intervals (22) of level $(1-\alpha) = 95\%$ at time-step $t = 10^3$, associated to the item c with highest $\mathbf{1}^\top \mathbf{K}_t(c)$ (see Fig. 2). In this case, the empirical coverage is equal to 0.970, which is again very close to the nominal value 0.95.

5. Applications

In this section we apply the proposed inferential methodology on the two real data sets, already considered in [26], to test the minimum level of interaction presented by the matrices Γ and W in these specific contexts. In particular, the first one is taken from the social content aggregation website Reddit, collected, elaborated and made freely available on the web by the authors of [37] (see https://github.com/corradomonti/demographic-homophily), while the second one is downloaded from the on-line library Project Gutenberg, which is a collection of public domain books (see https://www.gutenberg.org).

5.1. Reddit data set

The data set includes a collection of news posts and their corresponding comments from the Reddit community r/news over the period 2016–2020, downloaded from the website Reddit [37]. Each news item is linked to the user who posted it. In addition, the data provide the topic associated with each news post and assign to every comment a sentiment score, expressed as a real number in the interval $(-1, 1)$. We restrict our analysis to comments on news classified under the topic "Politics". The sentiment variable is then discretized as follows: a comment is labeled as "positive" if its sentiment score exceeds $+0.35$, and as "negative" if it is below -0.35. Comments with original sentiment values in $(-0.35, +0.35)$ are excluded. The resulting data set contains 2 602 173 comments in each of the two categories. Consequently, we define two innovation processes: one associated with authors receiving at least one negative comment on their posts, and another associated with authors receiving at least one positive comment. More precisely, we have two urns, one associated to positive sentiment and the other to negative sentiment, the extracted balls represent the comments to the posts and the color of a ball is associated with the author who wrote the commented post. Hence, a new color is extracted when the comment refers to a post written by an author who has not previously received comments on any of her/his posts. A first analysis of

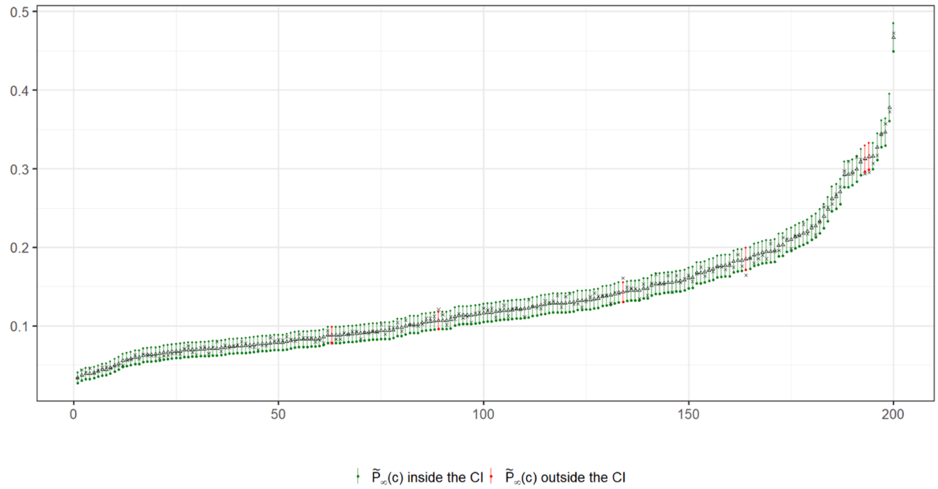

Fig. 2. Confidence interval (see (22)) for the random variable $\widetilde{P}_\infty(c)$ at time-step $t = 10^3$ and related to item c with highest $\mathbf{1}^\top \mathbf{K}_t(c)$ in the case $N = 3$, Γ defined as in (18) with $\phi = 0.75$ and $\iota_\Gamma = 0.8$, and W defined as in (20) with $\iota_W = 0.8$. Number of simulations $S = 200$ (sorted on the x-axis according to the value of $\mathbf{1}^\top \mathbf{K}_t(c)$). The crosses represent the value (reported on the y-axis) of $\widetilde{P}_\infty(c)$ estimated by $\mathbf{1}^\top \mathbf{K}_{t_\infty}(c)/(N t_\infty)$ at time-step $t_\infty = 10^5$.

Table 1
Reddit data set. p-values associated to the hypothesis tests described in Section 4.1 with $H_{\Gamma,0} : \iota_\Gamma \geq \iota_{\Gamma,0}, \eta \leq 1/2$ and $H_{W,0} : \iota_W \geq \iota_{W,0}$, respectively. The table also contains the summary statistics related to the sample of the 100 items c with highest $(K_{t,1}(c) + K_{t,2}(c))$.

Interaction Intensity $\iota_{\Gamma,0}$	p-value	Interaction Intensity $\iota_{W,0}$	p-value distribution $q_{50\%}$, $q_{75\%}$, $q_{95\%}$
0.15	0.384	0.60	<0.001, 0.007, 0.486
0.18	0.064	0.80	<0.001, <0.001, 0.229
0.21	0.014	1.00	<0.001, <0.001, 0.121
0.24	0.003	1.25	<0.001, <0.001, 0.058
0.27	0.001	1.50	<0.001, <0.001, 0.029

this data set can be found in [26], where it is shown that these two sequences exhibit empirical behaviours that are in accordance with those predicted by Theorems 2.1 and 2.2, and we provide the estimates $\widehat{\gamma}^* = 0.781$ and $\widehat{r} = 10^{-0.727} = 0.187$ (these estimates have been obtained by means of a suitable regression on $(\mathbf{D}^*_t)_t$, which simulations have shown to be a highly accurate estimation method). We here continue the study performing the two hypothesis tests on Γ (taking the above estimates as the known values of γ^* and r) and on W described in Subsection 4.1. More precisely, we consider the two one-sided tests:

$$H_{\Gamma,0} : \ \iota \geq \iota_{\Gamma,0}, \eta \leq 1/2 \qquad \text{and} \qquad H_{W,0} : \ \iota \geq \iota_{W,0},$$

where $\iota_{\Gamma,0}$ must be selected in the interval $(0.141, 0.291]$ and $\iota_{W,0}$ must be selected in the interval $(\frac{1}{2}, 2)$. The p-values associated to the null hypothesis $H_{\Gamma,0}$ and $H_{W,0}$, for different values of $\iota_{\Gamma,0}$ and $\iota_{W,0}$, respectively, are collected in Table 1. In order to better describe the empirical distribution of the p-values for testing W, instead of taking a single item c, we take the 100 items c with highest $(K_{t,1}(c) + K_{t,2}(c))$, and report in Table 1 the summary statistics of this sample (i.e. the quantiles of order 50%, 75% and 95%). We can see that, considering a test on Γ performed at level $\alpha = 0.05$, we can reject the null hypothesis with $\iota_{\Gamma,0} = 0.21$, but we cannot reject it for $\iota_{\Gamma,0} = 0.18$. Regarding the test on W instead, we observe that the proportion of items c providing a p–value that let us reject the null hypothesis that the interaction intensity ι_W is higher than $\iota_{W,0}$ is from 75% to 95% for $\iota_{W,0} \leq 1.25$, while is even higher than 95% or $\iota_{W,0} = 1.50$.

5.2. Gutenberg data set

This data set is drawn from the on-line library Project Gutenberg and concerns seven books written in English that belong to two specific literary genres: “Western” and “History”. More precisely, the corpus contains 476 948 words for each genre. Therefore, we have two urns, one associated to "Western" and the other to "History", the extracted balls represent the words and different words are associated with different colors. The two resulting word sequences constitute the innovation processes under investigation. A word is classified as new (respectively, old) for the overall system if it has not (respectively, has already) appeared in any of the considered genres. A first analysis of also this data set can be found in [26], where we show that the behaviors of these two sequences predicted by Theorems 2.1 and 2.2 match with the observed ones, and we provide the estimates $\widehat{\gamma}^* = 0.466$ and $\widehat{r} = 10^{-0.238} = 0.578$. As in the

Table 2
Gutenberg data set. p-values associated to the hypothesis tests described in Section 4.1 with $H_{\Gamma,0} : \iota_\Gamma \geq \iota_{\Gamma,0}$, $\eta \leq 1/2$ and $H_{W,0} : \iota_W \geq \iota_{W,0}$, respectively. The table also contains the summary statistics related to the sample of the 100 items c with highest $(K_{t,1}(c) + K_{t,2}(c))$.

Interaction Intensity $\iota_{\Gamma,0}$	p-value	Interaction Intensity $\iota_{W,0}$	p-value distribution $q_{50\%}$, $q_{75\%}$, $q_{95\%}$
0.21	0.034	0.60	0.010, 0.209, 0.823
0.24	<0.001	0.80	<0.001, 0.030, 0.699
0.27	<0.001	1.00	<0.001, 0.005, 0.618
0.30	<0.001	1.25	<0.001, 0.001, 0.541
0.40	<0.001	1.50	<0.001, <0.001, 0.480

previous application, we here continue the study performing the two hypothesis tests on Γ and W described in Section 4.1. More precisely, we consider the two one-sided tests:

$$H_{\Gamma,0} : \ \iota \geq \iota_0, \eta \leq 1/2 \qquad \text{and} \qquad H_{W,0} : \ \iota \geq \iota_0,$$

where $\iota_{\Gamma,0}$ must be selected in the interval $(0.202, 0.539]$ and $\iota_{W,0}$ must be selected in the interval $(\frac{1}{2}, 2)$. The p-values associated to the null hypothesis $H_{\Gamma,0}$ and $H_{W,0}$, for different values of $\iota_{\Gamma,0}$ and $\iota_{W,0}$, respectively, are collected in Table 2. Moreover, we take the 100 items c with highest $(K_{t,1}(c) + K_{t,2}(c))$ and report in the table the empirical quantiles of the corresponding sample of p-values. We can observe that, considering a test on Γ performed at level $\alpha = 0.05$, we can never reject the null hypothesis. Regarding the test on W instead, we observe that the proportion of items c providing a p-value that let us reject the null hypothesis that the interaction intensity ι_W is higher than $\iota_{W,0}$ is from 50% to 75% for $\iota_{W,0} = 0.60$, while is from 75% to 95% for $\iota_{W,0} \geq 0.80$.

Author contributions statement

All the authors contributed equally to the present work.

Acknowledgments

The authors gratefully acknowledge the anonymous referees for their constructive comments and suggestions, which greatly improved the quality of this manuscript.

Appendix

Notation. In all the appendix, the notation $R_t = O(s_t)$ means that $|R_t| \leq C s_t$ for a suitable deterministic constant C and for t large enough. Similarly, the vectorial notation $\mathbf{R}_t = O(s_t)\mathbf{1}$ means that each component of $\mathbf{R}_t$, say $R_{t,j}$, is such that $|R_{t,j}| \leq C s_t$ for an appropriate deterministic constant C and for t large enough. (Analogous results can be obtained for a suitable integrable random variable C.)

The results in appendix are given in terms of stable convergence or its variants. For more details on these concepts, we refer the reader to [38–42].

Appendix A. General results

The results in this appendix are formulated in a general setting and are intended to be applicable beyond the interacting innovation processes studied in the main text. In particular, they apply to multidimensional recursive dynamics of stochastic-approximation type and involve random centering, random limiting covariance matrices and non-standard scaling. Such dynamics arise naturally in reinforced stochastic processes and interacting urn models.

Consider two multi-dimensional real stochastic processes $(\mathbf{A}_t, \mathbf{B}_t)_t$, adapted to a filtration $(\mathcal{F}_t)_t$, with the following joint dynamics:

$$\begin{aligned} \mathbf{A}_{t+1} &= \mathbf{A}_t - \frac{1}{t+1}(\phi^* I - \Phi^\top)\mathbf{A}_t + \frac{1}{t+1}\Phi^\top \Delta\mathbf{M}_{t+1} + \mathbf{R}_{A,t+1}, \\ \mathbf{B}_{t+1} &= \mathbf{B}_t - \frac{1}{t+1}(\phi^* \mathbf{B}_t - \mathbf{A}_t) + \frac{1}{t+1}\Delta\mathbf{M}_{t+1} + \mathbf{R}_{B,t+1}\,, \end{aligned} \tag{A.1}$$

where $\mathbf{A}_0$ and $\mathbf{B}_0$ are integrable and

(i) $\Phi^\top$ is a non-negative irreducible diagonalizable matrix with leading eigenvalue $0 < \phi^* \leq 1$;
(ii) $\mathbf{R}_{A,t+1} = O(t^{-(1+\beta)})\mathbf{1}$ and $\mathbf{R}_{B,t+1} = O(t^{-(1+\beta)})\mathbf{1}$ for some $\beta > \frac{\phi^*}{2}$;
(iii) $\Delta\mathbf{M}_{t+1} = O(t^{1-\eta})\mathbf{1}$ is a martingale difference with $\eta > \phi^*/2$ and $t^{-(1-\phi^*)} E[\Delta\mathbf{M}_{t+1}\Delta\mathbf{M}_{t+1}^\top | \mathcal{F}_t] \xrightarrow{a.s.} \Sigma_\infty$ with Σ_∞ a positive semidefinite symmetric *random* matrix.

It is worthwhile to note that dynamics similar to (A.1) have already been considered in the Stochastic Approximation literature. However, the most part of those results do not apply to our situation because the covariance matrices in those results are deterministic (e.g. [43,44]), while the asymptotic covariance matrices in our CLTs are random. This is also why we do not use the simple convergence in distribution, but we employ the notion of stable convergence. On the other hand, the paper [36] also takes into account random asymptotic covariance matrices, but it is able to catch our CLT only in the case $\phi^* = 1$ (compare our assumption (iii) by assumption (2.4) in [36]). The recent work [34], again in the case $\phi^* = 1$, studies a related class of stochastic approximation dynamics where the required diagonalizability and irreducibility of the matrix Φ is replaced by a block triangularity assumption (see the following Remark A.1 for a more detailed discussion).

Let $\mathbf{u}$ and $\mathbf{v}$ be the right and the left eigenvectors of $\Phi^\top$ associated to ϕ^* and such that $\mathbf{v}^\top \mathbf{1} = 1$ and $\mathbf{v}^\top \mathbf{u} = 1$. (Recall that the components of these vectors are all strictly positive according to the Frobenius-Perron theory.)

Let $\widetilde{A}_t = \mathbf{v}^\top \mathbf{A}_t$ and get its dynamics by multypling $\mathbf{v}^\top$ to the dynamics of $\mathbf{A}_t$ as follows:

$$\widetilde{A}_{t+1} = \widetilde{A}_t + \frac{1}{t+1}\phi^* \mathbf{v}^\top \Delta \mathbf{M}_{t+1} + \mathbf{v}^\top \mathbf{R}_{A,t+1}.$$

Define

$$\widehat{\mathbf{A}}_t = \mathbf{A}_t - \widetilde{A}_t \mathbf{u} = (I - \mathbf{u}\mathbf{v}^\top)\mathbf{A}_t \qquad \text{and} \qquad \widehat{\mathbf{B}}_t = \mathbf{B}_t - \frac{1}{\phi^*}\widetilde{A}_t \mathbf{u}. \tag{A.2}$$

An intuitive reasoning of this decomposition relates to the eigen-structure of the matrix $(\phi^* I - \Phi^\top)$ in (A.1). Indeed, $\widetilde{A}_t$ is the projection along its kernel, that corresponds to the Perron Frobenius eigen-space of Φ, and $\widehat{\mathbf{A}}_t$ completes with the projections on the remaining directions. The process $(\mathbf{B}_t)_t$ in (A.1) will follow the dynamics of the attracting process $(\mathbf{A}_t)_t$, and $\widehat{\mathbf{B}}_t$ plays the role of $\widehat{\mathbf{A}}_t$.

Theorem A.1 (CLT for $(\widetilde{A}_t)$ when $N \geq 1$)**.** *In addition to assumptions (i), (ii) and (iii), also assume*

(iv) *$(\widetilde{A}_t)_t$ is uniformly integrable.*

Then $(\widetilde{A}_t)_t$ converges almost surely and in mean to an integrable random variable $\widetilde{A}_\infty$ and we have

$$t^{\phi^*/2}(\widetilde{A}_t - \widetilde{A}_\infty) \xrightarrow{stably} \mathcal{N}(0, \phi^* \mathbf{v}^\top \Sigma_\infty \mathbf{v}),$$

Moreover, the above convergence is in the sense of the almost sure conditional convergence with respect to the filtration $(\mathcal{F}_t)_t$.

Proof. Set $\breve{M}_t = \sum_{n=1}^t \frac{1}{n}\mathbf{v}^\top \Delta \mathbf{M}_n$ and $\widetilde{R}_{A,t+1} = \mathbf{v}^\top \mathbf{R}_{A,t+1}$. Therefore, we have

$$\begin{aligned}
\widetilde{A}_{t+1} - \widetilde{A}_0 &= \sum_{n=0}^{t} \widetilde{A}_{n+1} - \widetilde{A}_n \\
&= \phi^* \sum_{n=0}^{t} \frac{1}{n+1}\mathbf{v}^\top \Delta \mathbf{M}_{n+1} + \sum_{n=0}^{t} \widetilde{R}_{A,n+1} \\
&= \phi^* \breve{M}_{t+1} + \sum_{n=0}^{t} \widetilde{R}_{A,n+1}\,.
\end{aligned}$$

It follows from (iv) and $\sum_{n\geq 0} \widetilde{R}_{A,n+1} = \sum_{n\geq 0} O(1/n^{1+\beta}) < +\infty$ that $(\breve{M}_t)_t$ is a uniformly integrable martingale. Therefore, we obtain that $\breve{M}_t \xrightarrow{a.s./in\,mean} \breve{M}_\infty$ for a certain integrable random variable $\breve{M}_\infty$ and so $\widetilde{A}_t \xrightarrow{a.s./in\,mean} \widetilde{A}_\infty$. Moreover, we have

$$\begin{aligned}
t^{\phi^*/2}\left(\widetilde{A}_t - \widetilde{A}_\infty\right) &= t^{\phi^*/2} \sum_{n\geq t} \left(\widetilde{A}_n - \widetilde{A}_{n+1}\right) \\
&= t^{\phi^*/2}\phi^*(\breve{M}_t - \breve{M}_\infty) + t^{\phi^*/2} \sum_{n\geq t} \widetilde{R}_{A,n+1} \\
&= t^{\phi^*/2}\phi^*(\breve{M}_t - \breve{M}_\infty) + t^{\phi^*/2} \sum_{n\geq t} O(1/n^{1+\beta}) \\
&= t^{\phi^*/2}\phi^*(\breve{M}_t - \breve{M}_\infty) + t^{\phi^*/2} O(1/t^{\beta})\,,
\end{aligned}$$

where $\beta - \phi^*/2 > 0$. Therefore, it is enough to prove the convergence of $t^{\phi^*/2}(\breve{M}_t - \breve{M}_\infty)$. For this purpose, we want to apply Theorem D.1. Let us consider, for each t the filtration $(\mathcal{F}_{t,k})_k$ and the process $(L_{t,k})_k$ defined by

$$\mathcal{F}_{t,0} = \mathcal{F}_{t,1} = \mathcal{F}_t, \qquad L_{t,0} = L_{t,1} = 0$$

and, for $k \geq 2$,

$$\mathcal{F}_{t,k} = \mathcal{F}_{t+k-1}, \qquad L_{t,k} = t^{\phi^*/2}(\breve{M}_t - \breve{M}_{t+k-1}).$$

The process $(L_{t,k})_k$ is a martingale w.r.t. $(\mathcal{F}_{t,k})_k$ which converges (for $k \to +\infty$) a.s. and in mean to the random variable $L_{t,\infty} = t^{\phi^*/2}(\breve{M}_t - \breve{M}_\infty)$. In addition, the increment $Y_{t,n} = L_{t,n} - L_{t,n-1}$ is equal to zero for $n = 1$ and, for $n \geq 2$, it coincides with a random variable of the form $t^{\phi^*/2}(\check{M}_k - \check{M}_{k+1})$ with $k \geq t$. Therefore, we have

$$\begin{aligned}\sum_{n\geq 1} Y_{t,n}^2 &= t^{\phi^*} \sum_{k\geq t} (\check{M}_k - \check{M}_{k+1})^2 \\ &= t^{\phi^*} \sum_{k\geq t} \frac{1}{(k+1)^2} (\mathbf{v}^\top \Delta \mathbf{M}_{k+1})^2 \\ &= t^{\phi^*} \sum_{k\geq t} \frac{k^{1-\phi^*}}{(k+1)^2} k^{-(1-\phi^*)} (\mathbf{v}^\top \Delta \mathbf{M}_{k+1})^2 \xrightarrow{a.s.} \mathbf{v}^\top \Sigma_\infty \mathbf{v}/\phi^* ,\end{aligned}$$

where the almost sure convergence follows by applying [45, Lemma 4.1] and by the fact that

$$k^{-(1-\phi^*)} E[(\mathbf{v}^\top \Delta \mathbf{M}_{k+1})^2 | \mathcal{F}_t] = k^{-(1-\phi^*)} \mathbf{v}^\top E[\Delta \mathbf{M}_{k+1} \Delta \mathbf{M}_{k+1}^\top | \mathcal{F}_t] \mathbf{v} \xrightarrow{a.s.} \mathbf{v}^\top \Sigma_\infty \mathbf{v} .$$

Finally, we have

$$\sup_{n\geq 1} |Y_{t,n}| = t^{\phi^*/2} \sup_{k\geq t} |\check{M}_k - \check{M}_{k+1}| \leq \sup_{k\geq t} k^{\phi^*/2} \frac{|\mathbf{v}^\top \Delta \mathbf{M}_{k+1}|}{(k+1)}$$

and, since $\Delta \mathbf{M}_{k+1} = O((k+1)^{1-\eta})\mathbf{1}$ by assumption (iii) and so $k^{\phi^*/2} \frac{|\mathbf{v}^\top \Delta \mathbf{M}_{k+1}|}{(k+1)} = O(k^{-(\eta - \phi^*/2)})$, the thesis follows by Theorem D.1. $\square$

Theorem A.2 (Convergence result for $(\widehat{A}_t, \widehat{B}_t)$ when $N = 1$). *Under the assumptions* (i) *and* (ii), *we have*

$$\widehat{A}_t \equiv 0 \qquad \textit{and} \qquad \widehat{B}_t = B_t - \frac{1}{\phi^*} A_t \overset{a.s.}{=} o(t^{-\frac{\phi^*}{2}}).$$

Proof. First, when $N = 1$ it is immediate to see $\widetilde{A}_t \equiv A_t$ and hence $\widehat{A}_t = A_t - \widetilde{A}_t \equiv 0$ for any $t \geq 1$. The dynamics of $\widehat{B}_t$ can be obtained by subtracting the dynamics of $\frac{1}{\phi^*}\widetilde{A}_t = \frac{1}{\phi^*} A_t$ to the dynamics of B_t as follows:

$$\widehat{B}_{t+1} = \widehat{B}_t - \frac{1}{t+1}(\phi^* \widehat{B}_t + \phi^* \frac{1}{\phi^*} \widetilde{A}_t - A_t) + \frac{1}{t+1} \Delta M_{t+1} - \frac{1}{\phi^*}(\widetilde{A}_{t+1} - \widetilde{A}_t) + R_{B,t+1} ,$$

from which we can derive

$$\widehat{B}_{t+1} = \widehat{B}_t - \frac{1}{t+1}(\phi^* \widehat{B}_t) + \frac{1}{t+1} \Delta M_{t+1} - \frac{1}{\phi^*}\left(\frac{1}{t+1} \phi^* \Delta M_{t+1}\right) + R_{B,t+1} - \frac{1}{\phi^*} R_{A,t+1} ,$$

that is

$$\widehat{B}_{t+1} = \left(1 - \frac{\phi^*}{t+1}\right) \widehat{B}_t + O(t^{-(1+\beta)}).$$

Then, since $\sum_t \frac{\phi^*}{t+1} = +\infty$ and $\sum_t O(t^{-(1+\beta)}) < +\infty$, we can apply Lemma D.2 to obtain $\widehat{B}_t \xrightarrow{a.s.} 0$. Moreover, if we multiply the previous dynamics of $\widehat{B}_t$ by $(t+1)^{\frac{\phi^*}{2}}$ and we use the relation $(\frac{t+1}{t})^{\frac{\phi^*}{2}} = (1 + \frac{1}{t})^{\frac{\phi^*}{2}} = 1 + \frac{\phi^*/2}{t+1} + O(1/t^2)$, we get the following dynamics for the process $\widehat{B}_t^* = t^{\frac{\phi^*}{2}} \widehat{B}_t$:

$$\widehat{B}_{t+1}^* = \left(1 - \frac{\phi^*}{t+1}\right)\left(1 + \frac{\phi^*/2}{t+1}\right) \widehat{B}_t^* + O(t^{-(1+\beta-\frac{\phi^*}{2})}) + O(t^{-(2-\frac{\phi^*}{2})}),$$

which is equivalent to

$$\widehat{B}_{t+1}^* = \left(1 - \frac{\phi^*/2}{t+1}\right) \widehat{B}_t^* + O(t^{-(1+\beta-\frac{\phi^*}{2})}) + O(t^{-(2-\frac{\phi^*}{2})}).$$

Therefore, since $\sum_t t^{-(2-\frac{\phi^*}{2})} < +\infty$ because $\phi^*/2 < 1$ and $\sum_t t^{-(1+\beta-\frac{\phi^*}{2})} < +\infty$ because $\beta > \phi^*/2$ by assumption, we can apply Lemma D.2 also to the sequence $\widehat{B}_t^*$ to obtain $\widehat{B}_t^* \xrightarrow{a.s.} 0$, i.e. $\widehat{B}_t \overset{a.s.}{=} o(t^{-\frac{\phi^*}{2}})$. $\square$

We recall that the diagonalizable and irreducible matrix $\Phi^\top$ can be decomposed as

$$\Phi^\top = \phi^* \mathbf{u}\mathbf{v}^\top + UDV^\top , \tag{A.3}$$

where D is the diagonal matrix whose elements are the eigenvalues ϕ_j, $j \geq 2$ of Φ (i.e. of $\Phi^\top$) different from ϕ^* and U and V denote the matrices whose columns are the left (right) and the right (left) eigenvectors of Φ (of $\Phi^\top$, respectively) associated to these eigenvalues, such that we have

$$V^\top \mathbf{u} = U^\top \mathbf{v} = 0, \quad V^\top U = U^\top V = I, \quad I = \mathbf{u}\mathbf{v}^\top + UV^\top . \tag{A.4}$$

Let ϕ_2^* be an eigenvalue of Φ different from ϕ^* with highest real part, i.e. $\phi_2^* \in Sp(\Phi) \setminus \{\phi^*\}$ such that $\mathcal{R}e(\phi_2^*) = \max\{\mathcal{R}e(\phi) : \phi \in Sp(\Phi) \setminus \{\phi^*\}\}$.

Theorem A.3 (CLT for $(\widehat{\mathbf{A}}_t, \widehat{\mathbf{B}}_t)$ when $N \geq 2$). *In addition to assumptions* (i), (ii) *and* (iii), *assume*

(v) $\mathcal{R}e(\phi_2^*)/\phi^* < 1/2$.

Then, we have

$$t^{\phi^*/2}(\widehat{\mathbf{A}}_t,\widehat{\mathbf{B}}_t)\overset{stably}{\longrightarrow}\mathcal{N}(\mathbf{0},\mathcal{M}_\infty)$$

with

$$\mathcal{M}_\infty=\begin{pmatrix}\mathcal{M}_\infty^{11} & \mathcal{M}_\infty^{13}\\ \mathcal{M}_\infty^{13} & \mathcal{M}_\infty^{33}\end{pmatrix}=\begin{pmatrix}US_\infty^{11}U^\top & US_\infty^{13}U^\top\\ US_\infty^{13}U^\top & US_\infty^{33}U^\top\end{pmatrix},\tag{A.5}$$

where, for $h,j=2,\dots,N$

$$[S_\infty^{11}]_{h,j}=\frac{\phi_j\phi_h}{\phi^*-\phi_j-\phi_h}(\mathbf{v}_j^\top\Sigma_\infty\mathbf{v}_h),$$

$$[S_\infty^{33}]_{h,j}=\frac{1}{\phi^*-\phi_j-\phi_h}(\mathbf{v}_j^\top\Sigma_\infty\mathbf{v}_h),$$

$$[S_\infty^{13}]_{h,j}=\frac{\phi_j}{\phi^*-\phi_j-\phi_h}(\mathbf{v}_j^\top\Sigma_\infty\mathbf{v}_h).$$

Proof. We split the proofs in some steps.

Step 1 (Dynamics of $\theta_t=(\widehat{A}_t,\widehat{B}_t)$).

First of all, we observe that $\widehat{\mathbf{A}}_t=UV^\top\mathbf{A}_t$ and the dynamics of $\widehat{\mathbf{A}}_t$ can be obtained by multypling by $UV^\top$ the dynamics of $\mathbf{A}_t$ as follows:

$$\widehat{\mathbf{A}}_{t+1}=\widehat{\mathbf{A}}_t-\frac{1}{t+1}U(I\phi^*-D)V^\top\widehat{\mathbf{A}}_t+\frac{1}{t+1}UDV^\top\Delta\mathbf{M}_{t+1}+UV^\top\mathbf{R}_{A,t+1}.$$

The dynamics of $\widehat{\mathbf{B}}_t$ can be obtained by subtracting the dynamics of $\frac{1}{\phi^*}\widetilde{A}_t\mathbf{u}$ to the dynamics of $\mathbf{B}_t$ as follows:

$$\widehat{\mathbf{B}}_{t+1}=\widehat{\mathbf{B}}_t-\frac{1}{t+1}(\phi^*\widehat{\mathbf{B}}_t+\phi^*\frac{1}{\phi^*}\widetilde{A}_t\mathbf{u}-\widehat{\mathbf{A}}_t-\widetilde{A}_t\mathbf{u})+\frac{1}{t+1}\Delta\mathbf{M}_{t+1}-\frac{1}{\phi^*}\mathbf{u}(\widetilde{A}_{t+1}-\widetilde{A}_t)+\mathbf{R}_{B,t+1}\,,$$

from which we can derive

$$\widehat{\mathbf{B}}_{t+1}=\widehat{\mathbf{B}}_t-\frac{1}{t+1}(\phi^*\widehat{\mathbf{B}}_t-\widehat{\mathbf{A}}_t)+\frac{1}{t+1}(I-\mathbf{u}\mathbf{v}^\top)\Delta\mathbf{M}_{t+1}+\mathbf{R}_{B,t+1}-\frac{1}{\phi^*}\mathbf{u}\mathbf{v}^\top\mathbf{R}_{A,t+1}\,.$$

Then the dynamics of $\boldsymbol{\eta}_t=(\widehat{\mathbf{A}}_t,\widehat{\mathbf{B}}_t)$ can be expressed as

$$\boldsymbol{\eta}_{t+1}-\boldsymbol{\eta}_t=-\frac{1}{t+1}Q\boldsymbol{\eta}_t+\frac{1}{t+1}R\Delta\mathbf{M}_{\eta,t+1}+\mathbf{R}_{\eta,t+1}\,,$$

where[3]

$$Q=\begin{pmatrix}U(I\phi^*-D)V^\top & \mathcal{O}\\ -UV^\top & I\phi^*\end{pmatrix},\qquad \Delta\mathbf{M}_{\eta,t+1}=\begin{pmatrix}\Delta\mathbf{M}_{t+1}\\ \Delta\mathbf{M}_{t+1}\end{pmatrix},\qquad R=\begin{pmatrix}UDV^\top & \mathcal{O}\\ \mathcal{O} & UV^\top\end{pmatrix}.$$

Now, let us define the $(2N)\times(2N-1)$ matrices

$$U_\eta=\begin{pmatrix}U & \mathbf{0} & 0\\ 0 & \mathbf{u} & U\end{pmatrix},\qquad\text{and}\qquad V_\eta^\top=\begin{pmatrix}V^\top & 0\\ \mathbf{0}^\top & \mathbf{v}^\top\\ 0 & V^\top\end{pmatrix},$$

so that $V_\eta^\top U_\eta=I$ and

$$U_\eta V_\eta^\top=\begin{pmatrix}UV^\top & 0\\ \mathbf{0}^\top & I\end{pmatrix}.$$

Then, defining the $(2N-1)\times(2N-1)$ matrices

$$S_Q=\begin{pmatrix}I\phi^*-D & \mathbf{0} & 0\\ \mathbf{0}^\top & \phi^* & \mathbf{0}^\top\\ -I & \mathbf{0} & I\phi^*\end{pmatrix},\qquad\text{and}\qquad S_R=\begin{pmatrix}D & \mathbf{0} & 0\\ \mathbf{0}^\top & 0 & \mathbf{0}^\top\\ 0 & \mathbf{0} & I\end{pmatrix},$$

we have that $Q=U_\eta S_Q V_\eta^\top$ and $R=U_\eta S_R V_\eta^\top$. Then, from the above relations on U_η and $V_\eta^\top$ we get $U_\eta V_\eta^\top\boldsymbol{\eta}_t=\boldsymbol{\eta}_t$, and so multypling the dynamics of $\boldsymbol{\eta}_t$ by $U_\eta V_\eta^\top$ we finally get

$$\boldsymbol{\eta}_{t+1}=\frac{1}{t+1}U_\eta\Big(I-\frac{1}{t+1}S_Q\Big)V_\eta^\top\boldsymbol{\eta}_t+\frac{1}{t+1}R\Delta\mathbf{M}_{\eta,t+1}+\mathbf{R}_{\eta,t+1}.$$

Step 2 (Explicit expression for $\boldsymbol{\eta}_t=(\widehat{\mathbf{A}}_t,\widehat{\mathbf{B}}_t)$).

Recalling that the eigenvalues of $\Phi^\top$ different from ϕ^* are denoted by ϕ_j for $j=2,\dots,N$, let us set $\alpha_j=\phi^*-\phi_j$. Recall that $\mathcal{R}e(\alpha_j)>0$

[3] The symbol $\mathcal{O}$ denotes the null matrix.

since ϕ^* is the leading eigenvalue and so $\mathcal{R}e(\phi_j) < \phi^*$. Then, setting an initial time m_0 such that $\mathcal{R}e(\alpha_j)/(m_0+1) < 1$ for each j, we can write

$$\begin{aligned}\boldsymbol{\eta}_{t+1} &= C_{m_0,t}\boldsymbol{\eta}_{m_0} + \sum_{k=m_0}^{t} \frac{1}{k+1} C_{k+1,t} R \Delta \mathbf{M}_{\boldsymbol{\eta},k+1} + \sum_{k=m_0}^{t} C_{k+1,t}\mathbf{R}_{\boldsymbol{\eta},k+1} \\ &= C_{m_0,t}\boldsymbol{\eta}_{m_0} + \sum_{k=m_0}^{t} \mathbf{T}_{t,k+1} + \boldsymbol{\rho}_{t+1}\,,\end{aligned}$$

where

$$C_{k+1,t} = U_{\boldsymbol{\eta}} A_{k+1,t} V_{\boldsymbol{\eta}}^{\top}, \qquad \text{and}$$

$$A_{k+1,t} = \begin{cases} \prod_{m=k+1}^{t}\left(I - \frac{1}{m+1}S_Q\right) = \begin{pmatrix} A^{11}_{k+1,t} & \mathbf{0} & 0 \\ 0 & a^{22}_{k+1,t} & 0 \\ A^{31}_{k+1,t} & \mathbf{0} & A^{33}_{k+1,t}\end{pmatrix} & k = m_0-1,\dots,t-1, \\ I & k = t\,. \end{cases}$$

Notice that the blocks $A^{11}_{k+1,t}$, $A^{31}_{k+1,t}$ and $A^{33}_{k+1,t}$ are all diagonal $(N-1)\times(N-1)$ matrices. Moreover, setting for any $x \in \mathbb{C}$ with $\mathcal{R}e(x)/(m_0+1) < 1$, $p_{m_0-1}(x) = 1$ and $p_k(x) = \prod_{m=m_0}^{k}(1 - \frac{x}{m+1})$ for $k \geq m_0$ and $F_{k+1,t}(x) = \frac{p_t(x)}{p_k(x)}$ for $m_0 - 1 \leq k \leq t-1$, we get (see [46, Lemma A.5] for similar technical calculations)

$$\begin{aligned}[A^{11}_{k+1,t}]_{jj} &= F_{k+1,t}(\alpha_j) \\ [A^{33}_{k+1,t}]_{jj} &= a^{22}_{k+1,t} = F_{k+1,t}(\phi^*) \\ [A^{31}_{k+1,t}]_{jj} &= \begin{cases} -\frac{F_{k+1,t}(\alpha_j)-F_{k+1,t}(\phi^*)}{\alpha_j - \phi^*} & \text{for} \alpha_j \neq \phi^*, \text{i.e.} \phi_j \neq 0; \\ \frac{1-\phi^*}{\phi^*}F_{k+1,t}(\phi^*)\ln\left(\frac{t}{k}\right) + O\left(\frac{1}{t}\right) & \text{for} \alpha_j = \phi^*, \text{i.e.} \phi_j = 0. \end{cases}\end{aligned}$$

Step 3 (Study of the terms $C_{m_0,t}\boldsymbol{\eta}_{m_0}$ and $\boldsymbol{\rho}_{t+1}$).

We will prove that $t^{\frac{\phi^*}{2}}|C_{m_0,t}\boldsymbol{\eta}_{m_0}| \xrightarrow{a.s.} 0$ and $t^{\frac{\phi^*}{2}}|\boldsymbol{\rho}_{t+1}| \xrightarrow{a.s.} 0$. To this end, first notice that $O(|C_{k+1,t}|) = O(|A_{k+1,t}|)$ and, setting $a_2^* = \phi^* - \mathcal{R}e(\phi_2^*)$, we have that

$$\begin{aligned}|A_{k+1,t}| &= O\left(\frac{|p_t(\alpha_2^*)|}{|p_k(\alpha_2^*)|}\right) + O\left(\frac{|p_t(\phi^*)|}{|p_k(\phi^*)|}\right) + O\left(\frac{|p_t(\phi^*)|}{|p_k(\phi^*)|}\ln\left(\frac{t}{k}\right)\right) + O\left(\frac{1}{t}\right) \\ &= O\left(\left(\frac{k}{t}\right)^{a_2^*}\right) + O\left(\left(\frac{k}{t}\right)^{\phi^*}\right) + O\left(\left(\frac{k}{t}\right)^{\phi^*}\ln\left(\frac{t}{k}\right)\right) + O\left(\frac{1}{t}\right) \\ &= O\left(\left(\frac{k}{t}\right)^{a_2^*}\right) + O\left(\left(\frac{k}{t}\right)^{\phi^*}\ln(t)\right) \\ &= O\left(\left(\frac{k}{t}\right)^{(\phi^* - \mathcal{R}e(\phi_2^*))}\right) + O\left(\left(\frac{k}{t}\right)^{\phi^*}\ln(t)\right) \quad \text{for } k = m_0-1,\dots,t-1,\end{aligned}$$

and simply $|A_{t+1,t}| = O(1)$ for $k = t$. Therefore, since assumption (v), we have

$$t^{\frac{\phi^*}{2}}|C_{m_0,t}\boldsymbol{\eta}_{m_0}| = |\boldsymbol{\eta}_{m_0}|\,O\left(\left(\frac{1}{t}\right)^{\phi^*/2 - \mathcal{R}e(\phi_2^*)} + \left(\frac{1}{t}\right)^{\phi^*/2}\ln(t)\right) \xrightarrow{a.s.} 0\,.$$

Moreover, recalling that $\mathbf{R}_{\boldsymbol{\eta},t+1} = O(t^{-(1+\beta)})\mathbf{1}$ for some $\beta > \frac{\phi^*}{2}$, we have

$$\begin{aligned}t^{\frac{\phi^*}{2}}|\boldsymbol{\rho}_{t+1}| &= t^{\frac{\phi^*}{2}}\left|\sum_{k=m_0}^{t} C_{k+1,t}\mathbf{R}_{\boldsymbol{\eta},k+1}\right| \\ &= t^{\frac{\phi^*}{2}}O\left(\sum_{k=m_0}^{t-1}\left(\frac{k}{t}\right)^{a_2^*}\frac{1}{k^{1+\beta}} + \sum_{k=m_0}^{t-1}\left(\frac{k}{t}\right)^{\phi^*}\ln(t)\frac{1}{k^{1+\beta}}\right) + t^{\phi^*/2}O(1/t^{1+\beta}) \\ &= t^{\frac{\phi^*}{2}}O\left(\frac{1}{t^{a_2^*}}\sum_{k=m_0}^{t-1}k^{a_2^*-1-\beta} + \frac{\ln(t)}{t^{\phi^*}}\sum_{k=m_0}^{t-1}k^{\phi^*-1-\beta}\right) + O(1/t^{1+\beta-\phi^*/2}),\end{aligned}$$

which tends to zero because $\beta > \frac{\phi^*}{2}$ and $a_2^* > 0$.

Step 4 (Study of the "delta-martingale" term).

We want to apply Theorem D.3. Therefore we have to check conditions (1) and (2) in Theorem D.3.

Since the relation $V_\eta^\top U_\eta = I$ implies $V_\eta^\top R = S_R V_\eta^\top$, we have that

$$\begin{aligned}\sum_{k=m_0}^{t} (t^{\frac{\phi^*}{2}}\mathbf{T}_{t,k+1})(t^{\frac{\phi^*}{2}}\mathbf{T}_{t,k+1}^\top) &= t^{\phi^*}\sum_{k=m_0}^{t}\frac{1}{(k+1)^2}C_{k+1,t}R\Delta\mathbf{M}_{\eta,k+1}\Delta\mathbf{M}_{\eta,k+1}^\top R^\top C_{k+1,t}^\top\\ &= U_\eta\left(t^{\phi^*}\sum_{k=m_0}^{t}\frac{1}{(k+1)^2}A_{k+1,t}V_\eta^\top R\Delta\mathbf{M}_{\eta,k+1}\Delta\mathbf{M}_{\eta,k+1}^\top R^\top V_\eta A_{k+1,t}^\top\right)U_\eta^\top\\ &= U_\eta\left(t^{\phi^*}\sum_{k=m_0}^{t}\frac{1}{(k+1)^2}A_{k+1,t}S_R V_\eta^\top \Delta\mathbf{M}_{\eta,k+1}\Delta\mathbf{M}_{\eta,k+1}^\top V_\eta S_R A_{k+1,t}^\top\right)U_\eta^\top.\end{aligned}$$

Therefore, omitting the last term of the sum because $|A_{t+1,t}| = O(1)$ and , by assumption (iii), we have $(t^{\phi^*}/(t+1)^2)|\Delta\mathbf{M}_{t+1}| = O(t^{\phi^*-2+1-\eta}) = O(t^{-(1-\phi^*+\eta)}) \to 0$, it is enough to study the convergence of

$$t^{\phi^*}\sum_{k=m_0}^{t-1}\frac{1}{(k+1)^2}A_{k+1,t}S_R V_\eta^\top\Delta\mathbf{M}_{\eta,k+1}\Delta\mathbf{M}_{\eta,k+1}^\top V_\eta S_R A_{k+1,t}^\top. \tag{A.6}$$

To this purpose, setting $B_{\eta,k+1} = V_\eta^\top\Delta\mathbf{M}_{\eta,k+1}\Delta\mathbf{M}_{\eta,k+1}^\top V_\eta$, $B_{k+1} = V^\top\Delta\mathbf{M}_{k+1}\Delta\mathbf{M}_{k+1}^\top V$, $\mathbf{b}_{k+1} = V^\top\Delta\mathbf{M}_{k+1}\Delta\mathbf{M}_{k+1}^\top\mathbf{v}$ and $b_{k+1} = \mathbf{v}^\top\Delta\mathbf{M}_{k+1}\Delta\mathbf{M}_{k+1}^\top\mathbf{v}$, we observe that

$$B_{\eta,k+1} = \begin{pmatrix} B_{k+1} & \mathbf{b}_{k+1} & B_{k+1}\\ \mathbf{b}_{k+1}^\top & b_{k+1} & \mathbf{b}_{k+1}^\top\\ B_{k+1} & \mathbf{b}_{k+1} & B_{k+1}\end{pmatrix}.$$

Since in $B_{\eta,k+1}$ the first and the third row and column of blocks are the same, the matrix $(A_{k+1,t}S_R)$ can be replaced by a diagonal matrix, with the following diagonal blocks: $A_{k+1,t}^1 = A_{k+1,t}^{11}D$, $A_{k+1,t}^3 = A_{k+1,t}^{31}D + A_{k+1,t}^{33}$ and $a_{k+1,t}^2 = 0$. Hence, the above expression (A.6) can be rewritten as

$$t^{\phi^*}\sum_{k=m_0}^{t-1}\frac{1}{(k+1)^2}\begin{pmatrix} A_{k+1,t}^1 B_{k+1}A_{k+1,t}^1 & \mathbf{0} & A_{k+1,t}^1 B_{k+1}A_{k+1,t}^3\\ \mathbf{0}^\top & 0 & \mathbf{0}^\top\\ A_{k+1,t}^3 B_{k+1}A_{k+1,t}^1 & \mathbf{0} & A_{k+1,t}^3 B_{k+1}A_{k+1,t}^3\end{pmatrix}.$$

The elements of $A_{k+1,t}^1$ and $A_{k+1,t}^3$ in the above matrix can be rewritten in terms of $F_{k+1,t}(\cdot)$ as

$$\begin{aligned}{[A_{k+1,t}^1]_{jj}} &= (\phi^*-\alpha_j)F_{k+1,t}(\alpha_j) = \phi_j F_{k+1,t}(\alpha_j)\\ {[A_{k+1,t}^3]_{jj}} &= \begin{cases} F_{k+1,t}(\alpha_j) & \text{for } \alpha_j\neq\phi^*,\ \text{i.e. } \phi_j\neq 0\\ F_{k+1,t}(\phi^*) & \text{for } \alpha_j=\phi^*,\ \text{i.e. } \phi_j = 0.\end{cases}\end{aligned}$$

Moreover, with some computations (see [47, Lemma A.3] for similar technical calculations) for complex numbers x, y such that $\mathcal{R}e(x+y)-\phi^* > 0$ we have

$$t^{\phi^*}\sum_{k=m_0}^{t-1}\frac{1}{(k+1)^{1+\phi^*}}F_{k+1,t}(x)F_{k+1,t}(y)\to\frac{1}{x+y-\phi^*}. \tag{A.7}$$

Hence, combining (see [45, Lemma 4.1]) this limit with the limit relation

$$t^{-(1-\phi^*)}E[B_{t+1}|\mathcal{F}_t] = t^{-(1-\phi^*)}V^\top E[\Delta\mathbf{M}_{t+1}\Delta\mathbf{M}_{t+1}^\top|\mathcal{F}_t]V\to V^\top\Sigma_\infty V,$$

we get

$$t^{\phi^*}\sum_{k=m_0}^{t-1}\frac{1}{(k+1)^{1+\phi^*}}\Big((k+1)^{-(1-\phi^*)}B_{k+1}\Big)F_{k+1,t}(x)F_{k+1,t}(y)\to\frac{1}{x+y-\phi^*}V^\top\Sigma_\infty V\,. \tag{A.8}$$

As a consequence, recalling that $\alpha_j = \phi^*-\phi_j$ and, by assumption (v), $\phi^*-2\mathcal{R}e(\phi_2^*) > 0$, the almost sure convergences of all the elements of the previous matrix can be obtained by applying (A.8) as follows:

$$\begin{aligned} t^{\phi^*}\sum_{k=m_0}^{t-1}\frac{1}{(k+1)^2}[A_{k+1,t}^1 B_{k+1}A_{k+1,t}^1]_{h,j} &\to [S_\infty^{11}]_{h,j} = \frac{\phi_j\phi_h}{\phi^*-\phi_j-\phi_h}(\mathbf{v}_j^\top\Sigma_\infty\mathbf{v}_h),\\ t^{\phi^*}\sum_{k=m_0}^{t-1}\frac{1}{(k+1)^2}[A_{k+1,t}^3 B_{k+1}A_{k+1,t}^3]_{h,j} &\to [S_\infty^{33}]_{h,j} = \frac{1}{\phi^*-\phi_j-\phi_h}(\mathbf{v}_j^\top\Sigma_\infty\mathbf{v}_h),\\ t^{\phi^*}\sum_{k=m_0}^{t-1}\frac{1}{(k+1)^2}[A_{k+1,t}^1 B_{k+1}A_{k+1,t}^3]_{h,j} &\to [S_\infty^{13}]_{h,j} = \frac{\phi_j}{\phi^*-\phi_j-\phi_h}(\mathbf{v}_j^\top\Sigma_\infty\mathbf{v}_h).\end{aligned}$$

Therefore condition (1) of Theorem D.3 is satisfied with the matrix $\mathcal{M}_\infty$ defined in (A.5).

Regarding condition (2) of Theorem D.3, we observe that, by assumption (iii), we have

$$|\mathbf{T}_{t,k+1}| = O\Big(\frac{1}{k+1}|A_{k+1,t}||\Delta\mathbf{M}_{\boldsymbol{\eta},k+1}|\Big) = O\big(|A_{k+1,t}|(k+1)^{-\eta}\big)$$
$$= O\left(k^{-\eta}\left(\frac{k}{t}\right)^{(\phi^*-\mathcal{R}e(\phi_2^*))}\right) + O\left(k^{-\eta}\left(\frac{k}{t}\right)^{\phi^*}\ln(t)\right) \quad \text{for } k = m_0,\dots,t-1$$

and $|\mathbf{T}_{t,t+1}| = O\Big(\frac{1}{t+1}|A_{t+1,t}||\Delta\mathbf{M}_{\boldsymbol{\eta},t+1}|\Big) = O(t^{-\eta})$ for $k = t$. Hence, for any u, we have

$$\left(\sup_{m_0\le k\le t}\left|t^{\phi^*/2}\mathbf{T}_{t,k+1}\right|\right)^{2u} \le t^{\phi^* u}\sum_{k=m_0}^{t-1}|\mathbf{T}_{t,k+1}|^{2u} + t^{\phi^* u}|\mathbf{T}_{t,t+1}|^{2u}$$
$$= O\left(t^{-(\phi^*-2\mathcal{R}e(\phi_2^*))u}\sum_{k=m_0}^{t-1}k^{-(\eta-\phi^*+2\mathcal{R}e(\phi_2^*))u}\right)$$
$$+ O\left(\frac{\ln(t)^{2u}}{t^{\phi^* u}}\sum_{k=m_0}^{t-1}k^{-(\eta-\phi^*)2u}\right) + O(t^{-(2\eta-\phi^*)u}),$$

which tends to zero for $u > \max\{1/\eta; 1/(2\eta-\phi^*)\}$ since $\phi^* - 2\mathcal{R}e(\phi_2^*) > 0$ and $\eta > \phi^*/2$. This result, in particular, implies condition (2) of Theorem D.3.

By Theorem D.3, we can finally conclude that $t^{\phi^*/2}\sum_{k=m_0}^{t}\mathbf{T}_{t,k+1}$ converges stably to the Gaussian kernel $\mathcal{N}(\mathbf{0},\mathcal{M}_\infty)$. □

Remark A.1. Replacing the diagonal spectral decomposition by a Jordan decomposition introduces nilpotent terms in the powers of the linearized recursion. These terms produce polynomial factors, and in critical cases logarithmic factors, in the normalization and in the covariance computations. Consequently, the limiting covariance matrices must be expressed through sums involving Jordan chains rather than through simple eigenmode-wise contributions. The martingale arguments remain conceptually similar, but the estimates of the products associated with the stochastic approximation recursion require a more careful treatment. Note that the fundamental ideas of the proof will remain the same, only the computations and the final formulas will become more complex and difficult to read. To get an idea, we refer to [34], where, for a different context, the assumption of a diagonalizable and irriducible matrix is replaced by the assumption of a block upper-triangular matrix. A substantial part of the analysis in [34] is devoted to understanding the contribution of the terms generated by the nilpotent matrix that appears in the block-triangular decomposition. Note that the proofs are still based on [27, Lemma A.4], [46, Lemma A.5] and its analogous [34, Lemma A.5].

For reducible matrices, the directed graph must be decomposed into communicating classes. The leading classes, their Perron–Frobenius eigenvalues, and the accessibility relations between blocks determine the first-order rates. The appropriate first-order analysis in the case of a reducible matrix may be found in [26,28,35]. A second-order theory would then have to be developed separately for the leading blocks and for the propagation of fluctuations to downstream blocks and would require additional block-wise fluctuation analysis (see, e.g., [34] for block upper-triangular matrices).

Remark A.2. A key point in the proof of Theorem A.3 is the convergence of the left-hand term of (A.7) toward a bounded limit, which is ensured by assumption (v), i.e. $\mathcal{R}e(\phi_2^*) < \phi^*/2$. On the other hand, when $\mathcal{R}e(\phi_2^*) \ge \phi^*/2$, this quantity can diverge to $+\infty$, implying the divergence of $t^{\phi^*/2}(\widehat{\mathbf{A}}_t,\widehat{\mathbf{B}}_t)$. In fact, since Φ is a real matrix, in its spectrum we have ϕ_2^* and its conjugate, so that $F_{k+1,t}(x)F_{k+1,t}(y) = |F_{k+1,t}(x)|^2$ when we take $x = \alpha_2^* = \phi^* - \phi_2^*$ and for y its conjugate (or $y = x$ if x is real). Now, we observe that $|F_{k+1,t}(\alpha_2^*)|^2$ increases with $\mathcal{R}e(\phi_2^*)$ and the corresponding limit in (A.7), that is $1/(2\mathcal{R}e(\alpha_2^*)-\phi^*) = 1/(\phi^*-2\mathcal{R}e(\phi_2^*))$, goes to $+\infty$ for $\mathcal{R}e(\phi_2^*) \uparrow \phi^*/2$. As a consequence of these two facts, we can conclude that the limit of the left-hand term of (A.7) is necessarily $+\infty$ whenever $\mathcal{R}e(\phi_2^*) \ge \phi^*/2$. This means that, in that case, the rate of convergence is slower than $t^{\phi^*/2}$.

Theorem A.4. *When $N = 1$, under assumptions (i), (ii), (iii) and (iv), we have*

$$t^{\phi^*/2}\begin{pmatrix} A_t - \widetilde{A}_\infty \\ B_t - \widetilde{A}_\infty/\phi^* \end{pmatrix} \xrightarrow{stably} \mathcal{N}\left(0, \begin{pmatrix} \phi^*\Sigma_\infty & \Sigma_\infty \\ \Sigma_\infty & \Sigma_\infty/\phi^* \end{pmatrix}\right).$$

When $N \ge 2$, under assumptions (i), (ii), (iii), (iv) and (v), we have

$$t^{\phi^*/2}\begin{pmatrix} \mathbf{A}_t - \widetilde{A}_\infty\mathbf{u} \\ \mathbf{B}_t - \widetilde{A}_\infty/\phi^*\mathbf{u} \end{pmatrix} \xrightarrow{stably} \mathcal{N}\left(\mathbf{0}, \mathbf{v}^\top\Sigma_\infty\mathbf{v}\begin{pmatrix} \phi^*\mathbf{u}\mathbf{u}^\top & \mathbf{u}\mathbf{u}^\top \\ \mathbf{u}\mathbf{u}^\top & \mathbf{u}\mathbf{u}^\top/\phi^* \end{pmatrix} + \mathcal{M}_\infty\right),$$

where $\mathcal{M}_\infty$ is given in (A.5).

Proof. When $N = 1$, we simply have $(A_t, B_t) = (\widetilde{A}_t, \widetilde{A}_t/\phi^* + o(t^{-\phi^*/2}))$ and so the statement follows from Theorem A.1 and Theorem A.2 for $\widetilde{A}_t$. When $N \ge 2$, we have

$$t^{\phi^*/2}(\mathbf{A}_t - \widetilde{A}_\infty\mathbf{u}, \mathbf{B}_t - \widetilde{A}_\infty/\phi^*\mathbf{u})$$
$$= t^{\phi^*/2}((\widetilde{A}_t - \widetilde{A}_\infty)\mathbf{u} + \widehat{\mathbf{A}}_t, (\widetilde{A}_t - \widetilde{A}_\infty)/\phi^*\mathbf{u} + \widehat{\mathbf{B}}_t)$$
$$= t^{\phi^*/2}(\widetilde{A}_t - \widetilde{A}_\infty)(\mathbf{u}, \mathbf{u}/\phi^*) + t^{\phi^*/2}(\widehat{\mathbf{A}}_t, \widehat{\mathbf{B}}_t).$$

The first term converges in the sense of the almost sure conditional convergence toward the Gaussian kernel with mean $\mathbf{0}$ and covariance matrix

$$\mathbf{v}^\top \Sigma_\infty \mathbf{v} \begin{pmatrix} \phi^* \mathbf{u}\mathbf{u}^\top & \mathbf{u}\mathbf{u}^\top \\ \mathbf{u}\mathbf{u}^\top & \mathbf{u}\mathbf{u}^\top/\phi^* \end{pmatrix}.$$

The second term converges stably toward the Gaussian kernel $\mathcal{N}(\mathbf{0}, \mathcal{M}_\infty)$. By Lemma D.4, we can combine the two convergences and obtain the desidered convergence. □

As an immediate corollary, we can get the central limit theorem for $\widetilde{B}_t = \mathbf{v}^\top \mathbf{B}_t$ when $N \geq 2$.

Corollary A.5 (CLT for $(\widetilde{B}_t)$ when $N \geq 2$). *When $N \geq 2$, under assumptions (i), (ii), (iii), (iv) and (v), we have*

$$t^{\phi^*/2}(\widetilde{B}_t - \widetilde{A}_\infty/\phi^*) \xrightarrow{stably} \mathcal{N}\left(0\,, \mathbf{v}^\top \Sigma_\infty \mathbf{v}/\phi^*\right).$$

Proof. Setting $\widetilde{B}_t = \mathbf{v}^\top \mathbf{B}_t$ and using the relations $\mathbf{v}^\top \mathbf{u} = 1$ and $\mathbf{v}^\top U = \mathbf{0}$ in Theorem A.4, we find the desidered convergence. □

We can also derive the following result that consists in a CLT for a normalized version of $(\mathbf{A}_t, \mathbf{B}_t)$.

Corollary A.6 (CLT for $(\mathbf{A}_t/(\mathbf{x}^\top \mathbf{A}_t), \mathbf{B}_t/(\mathbf{x}^\top \mathbf{B}_t))$ when $N \geq 2$). *Under assumptions (i), (ii), (iii), (iv) and (v), if $\widetilde{A}_\infty \neq 0$ a.s., then, for any given vector with $\mathbf{x}^\top \mathbf{u} \neq 0$, we have*

$$t^{\phi^*/2} \begin{pmatrix} \mathbf{A}_t/(\mathbf{x}^\top \mathbf{A}_t) - \mathbf{u}/(\mathbf{x}^\top \mathbf{u}) \\ \mathbf{B}_t/(\mathbf{x}^\top \mathbf{B}_t) - \mathbf{u}/(\mathbf{x}^\top \mathbf{u}) \end{pmatrix} \xrightarrow{stably} \mathcal{N}\left(\mathbf{0}, \frac{1}{(\mathbf{x}^\top \mathbf{u})^2 \widetilde{A}_\infty^2} \begin{pmatrix} Q_\mathbf{x} \mathcal{M}_\infty^{11} Q_\mathbf{x}^\top & \phi^* Q_\mathbf{x} \mathcal{M}_\infty^{13} Q_\mathbf{x}^\top \\ \phi^* Q_\mathbf{x} \mathcal{M}_\infty^{13} Q_\mathbf{x}^\top & (\phi^*)^2 Q_\mathbf{x} \mathcal{M}_\infty^{33} Q_\mathbf{x}^\top \end{pmatrix} \right),$$

where $Q_\mathbf{x} = I - \mathbf{u}\mathbf{x}^\top/(\mathbf{x}^\top \mathbf{u})$ and the matrices $\mathcal{M}_\infty^{11}$, $\mathcal{M}_\infty^{13}$, $\mathcal{M}_\infty^{33}$ are defined in (A.5). In particular, for $\mathbf{x} = \mathbf{v}$ we have

$$t^{\phi^*/2} \begin{pmatrix} \mathbf{A}_t/\widetilde{A}_t - \mathbf{u} \\ \mathbf{B}_t/\widetilde{B}_t - \mathbf{u} \end{pmatrix} \xrightarrow{stably} \mathcal{N}\left(\mathbf{0}, \frac{1}{\widetilde{A}_\infty^2} \begin{pmatrix} \mathcal{M}_\infty^{11} & \phi^* \mathcal{M}_\infty^{13} \\ \phi^* \mathcal{M}_\infty^{13} & (\phi^*)^2 \mathcal{M}_\infty^{33} \end{pmatrix} \right).$$

Remark A.3. Note that, by (A.2) and (A.4), we have $\widehat{\mathbf{A}}_t = \mathbf{A}_t - \widetilde{A}_t \mathbf{u} = UV^\top \mathbf{A}_t$ and, similarly, $\mathbf{B}_t - \widetilde{B}_t \mathbf{u} = UV^\top \mathbf{B}_t$. Therefore, the processes in the above result in the case $\mathbf{x} = \mathbf{v}$ coincide, up to multiplicative (random) factors, with the processes $UV^\top \mathbf{A}_t$ and $UV^\top \mathbf{B}_t$ (note also that we have $Q_\mathbf{v} = I - \mathbf{u}\mathbf{v}^\top/(\mathbf{v}^\top \mathbf{u}) = UV^\top$ by (A.4)). These processes will play a fundamental role in the following section dedicated to the statistical tools.

Proof. The result follows from the CLT of Theorem A.4 once we note that

$$\begin{aligned}
&t^{\phi^*/2} \begin{pmatrix} \mathbf{A}_t/(\mathbf{x}^\top \mathbf{A}_t) - \mathbf{u}/(\mathbf{x}^\top \mathbf{u}) \\ \mathbf{B}_t/(\mathbf{x}^\top \mathbf{B}_t) - \mathbf{u}/(\mathbf{x}^\top \mathbf{u}) \end{pmatrix} \\
&= t^{\phi^*/2} \begin{pmatrix} (\mathbf{x}^\top \mathbf{A}_t)^{-1} I & 0 \\ 0 & (\mathbf{x}^\top \mathbf{B}_t)^{-1} I \end{pmatrix} \begin{pmatrix} \mathbf{A}_t - \mathbf{u}\mathbf{x}^\top \mathbf{A}_t/(\mathbf{x}^\top \mathbf{u}) \\ \mathbf{B}_t - \mathbf{u}\mathbf{x}^\top \mathbf{B}_t/(\mathbf{x}^\top \mathbf{u}) \end{pmatrix} \\
&= t^{\phi^*/2} \begin{pmatrix} (\mathbf{x}^\top \mathbf{A}_t)^{-1} Q_\mathbf{x} & 0 \\ 0 & (\mathbf{x}^\top \mathbf{B}_t)^{-1} Q_\mathbf{x} \end{pmatrix} \begin{pmatrix} \mathbf{A}_t \\ \mathbf{B}_t \end{pmatrix} \\
&= t^{\phi^*/2} \begin{pmatrix} (\mathbf{x}^\top \mathbf{A}_t)^{-1} Q_\mathbf{x} & 0 \\ 0 & (\mathbf{x}^\top \mathbf{B}_t)^{-1} Q_\mathbf{x} \end{pmatrix} \begin{pmatrix} \mathbf{A}_t - \widetilde{A}_\infty \mathbf{u} \\ \mathbf{B}_t - \widetilde{A}_\infty/\phi^* \mathbf{u} \end{pmatrix}.
\end{aligned}$$

Indeed, we have $Q_\mathbf{x} \mathbf{u}\mathbf{u}^\top = 0$ and so, since $\mathbf{x}^\top \mathbf{A}_t \to (\mathbf{x}^\top \mathbf{u})\widetilde{A}_\infty$ and $\mathbf{x}^\top \mathbf{B}_t \to (\mathbf{x}^\top \mathbf{u})\widetilde{A}_\infty/\phi^*$, we have

$$\begin{aligned}
&\begin{pmatrix} (\mathbf{x}^\top \mathbf{A}_t)^{-1} Q_\mathbf{x} & 0 \\ 0 & (\mathbf{x}^\top \mathbf{B}_t)^{-1} Q_\mathbf{x} \end{pmatrix} \left(\mathbf{v}^\top \Sigma_\infty \mathbf{v} \begin{pmatrix} \phi^* \mathbf{u}\mathbf{u}^\top & \mathbf{u}\mathbf{u}^\top \\ \mathbf{u}\mathbf{u}^\top & \mathbf{u}\mathbf{u}^\top/\phi^* \end{pmatrix} + \mathcal{M}_\infty \right) \begin{pmatrix} (\mathbf{x}^\top \mathbf{A}_t)^{-1} Q_\mathbf{x} & 0 \\ 0 & (\mathbf{x}^\top \mathbf{B}_t)^{-1} Q_\mathbf{x} \end{pmatrix}^\top = \\
&\begin{pmatrix} (\mathbf{x}^\top \mathbf{A}_t)^{-1} Q_\mathbf{x} & 0 \\ 0 & (\mathbf{x}^\top \mathbf{B}_t)^{-1} Q_\mathbf{x} \end{pmatrix} \mathcal{M}_\infty \begin{pmatrix} (\mathbf{x}^\top \mathbf{A}_t)^{-1} Q_\mathbf{x}^\top & 0 \\ 0 & (\mathbf{x}^\top \mathbf{B}_t)^{-1} Q_\mathbf{x}^\top \end{pmatrix} \xrightarrow{a.s.} \\
&\frac{1}{(\mathbf{x}^\top \mathbf{u})^2 \widetilde{A}_\infty^2} \begin{pmatrix} Q_\mathbf{x} \mathcal{M}_\infty^{11} Q_\mathbf{x}^\top & \phi^* Q_\mathbf{x} \mathcal{M}_\infty^{13} Q_\mathbf{x}^\top \\ \phi^* Q_\mathbf{x} \mathcal{M}_\infty^{13} Q_\mathbf{x}^\top & (\phi^*)^2 Q_\mathbf{x} \mathcal{M}_\infty^{33} Q_\mathbf{x}^\top \end{pmatrix}.
\end{aligned}$$

The result for $\mathbf{x} = \mathbf{v}$ simply follows by the fact that, since (A.4), we have the equalities $Q_\mathbf{x} = Q_\mathbf{v} = (I - \mathbf{u}\mathbf{v}^\top/(\mathbf{v}^\top \mathbf{u})) = UV^\top$ and so $UV^\top \mathcal{M}_\infty^{ij} VU^\top = \mathcal{M}_\infty^{ij}$. □

Appendix B. Statistical tools

In this section we present some statistical tools, based on Theorem A.4, to provide an interval estimation of the random limit $\widetilde{A}_\infty$ and an hypothesis test on the structure of the matrix Φ. To this end, we need to assume that the asymptotic conditional covariance matrix Σ_∞ of the martingale vector $\Delta \mathbf{M}_t$ satisfies the following assumption:

Assumption B.1. Assume the matrix Σ_∞ can be factorized as

$$\Sigma_\infty = g(\widetilde{A}_\infty)\Sigma_{det}\,,$$

with g being a continuous function with $g(\widetilde{A}_\infty) > 0$ a.s. and Σ_{det} a deterministic positive definite symmetric matrix determined by the eigen-structure of Φ.

We will provide statistical tools based on the process $(\mathbf{A}_t)$ or on the process $(\mathbf{B}_t)$ so that the suitable tool can be selected in the practical applications when only one of the two processes is observable.

B.1. Confidence intervals for the random limit $\widetilde{A}_\infty$ based on $(\mathbf{A}_t)$

First of all, we can observe that by means of the central limit theorem for the stochastic process $\widetilde{A}_t = \mathbf{v}^\top \mathbf{A}_t$ described in Theorem A.1, it is possible to construct an asymptotic confidence interval for its limit $\widetilde{A}_\infty$. Specifically, an asymptotic confidence interval for $\widetilde{A}_\infty$ with approximate level $(1-\alpha)$ is the following:

$$CI_{1-\alpha}(\widetilde{A}_\infty) \;:=\; \left(\widetilde{A}_t - \frac{z_\alpha}{t^{\phi^*/2}}\sqrt{g(\widetilde{A}_t)\phi^*\mathbf{v}^\top\Sigma_{det}\mathbf{v}}\;;\; \widetilde{A}_t + \frac{z_\alpha}{t^{\phi^*/2}}\sqrt{g(\widetilde{A}_t)\phi^*\mathbf{v}^\top\Sigma_{det}\mathbf{v}}\right),$$

where z_α is such that $\mathcal{N}(0,1)(z_\alpha,+\infty) = \alpha/2$. Note that the term $g(\widetilde{A}_\infty)$ in the confidence interval has been replaced by its strongly consistent estimator $g(\widetilde{A}_t)$.

Since the convergence in Theorem A.1 is also in the sense of the almost sure conditional convergence with respect to $(\mathcal{F}_t)$, the above interval can be recovered as a credible interval (see [48]) in a Bayesian framework.

B.2. Confidence intervals for the random limit $\widetilde{A}_\infty$ based on $(\mathbf{B}_t)$

We here construct an asymptotic confidence interval for $\widetilde{A}_\infty$ based on the process $(\mathbf{B}_t)_t$. First notice that using Theorem A.4 we can obtain for $N=1$

$$CI_{1-\alpha}(\widetilde{A}_\infty) \;=\; \phi^* B_t \;\pm\; \frac{z_\alpha}{t^{\phi^*/2}}\sqrt{g(\phi^* B_t)\phi^*\Sigma_{det}},$$

where $g(\phi^* B_t)$ represents a consistent estimator of $g(\widetilde{A}_\infty)$ and z_α is such that $\mathcal{N}(0,1)(z_\alpha,+\infty) = \alpha/2$.

When $N \geq 2$, it is required the knowledge of $\mathcal{R}e(\phi_2^*)$. Indeed, when $\mathcal{R}e(\phi_2^*)/\phi^* < 1/2$, setting $\widetilde{B}_t = \mathbf{v}^\top\mathbf{B}_t$ and using Corollary A.5, we obtain

$$CI_{1-\alpha}(\widetilde{A}_\infty) \;=\; \phi^*\widetilde{B}_t \;\pm\; \frac{z_\alpha}{t^{\phi^*/2}}\sqrt{g(\phi^*\widetilde{B}_t)\phi^*\mathbf{v}^\top\Sigma_{det}\mathbf{v}},$$

where z_α is such that $\mathcal{N}(0,1)(z_\alpha,+\infty) = \alpha/2$.

B.3. Hypothesis tests on the matrix Φ based on $(\mathbf{A}_t)$

We now focus on the case $N \geq 2$ and the inferential problem of testing the null hypothesis $H_0 : \Phi = \Phi_0$, using the multi-dimensional stochastic process $(\mathbf{A}_t)_t$. Since the distribution of $\widetilde{A}_\infty$ should be unknown in practice, we propose a test statistics whose limit does not involve $\widetilde{A}_\infty$. First, we need to introduce some notation. Let us define the matrix $\mathcal{M}^{11}_{det}$ as the matrix $\mathcal{M}^{11}_\infty$ defined in Theorem A.3 with Σ_∞ replaced by Σ_{det}, i.e. $\mathcal{M}^{11}_\infty = g(\widetilde{A}_\infty)\mathcal{M}^{11}_{det}$. Then, notice that the matrix $\mathcal{M}^{11}_{det}$ has rank $(N-1)$ and it admits a spectral decomposition as follows: $\mathcal{M}^{11}_{det} = O^{11}\mathcal{D}^{11}(O^{11})^\top$, where $\mathcal{D}^{11}$ is the diagonal matrix of dimension $(N-1)$ containing the eigenvalues of $\mathcal{M}^{11}_{det}$ different from 0 and O^{11} is the $N \times (N-1)$ matrix whose columns form an orthonormal basis of $Im(\mathcal{M}^{11}_{det})$ of corresponding eigenvectors. Then, we set $L^{11} = (\mathcal{D}^{11})^{-1/2}$ (diagonal matrix of dimension $(N-1)$) and take the product $L^{11}(O^{11})^\top$ (matrix of dimension $(N-1)\times N$). Fixed the matrix assumed under H_0, i.e. Φ_0, we can compute for it the eigenvectors $\mathbf{u}$ and $\mathbf{v}$, the eigenvectors contained in the matrices U and V and the matrices L^{11} and O^{11} defined above. Hence, we can obtain under H_0 the real process $\widetilde{A}_t = \mathbf{v}^\top \mathbf{A}_t$ and the multi-dimensional process $\widehat{\mathbf{A}}_t = \mathbf{A}_t - \widetilde{A}_t\mathbf{u} = U\,V^\top \mathbf{A}_t$. Then, from Theorem A.3 we have that, under H_0 and the assumption $\mathcal{R}e(\phi_2^*)/\phi^* < 1/2$,

$$\mathbf{T}_t = t^{\phi^*/2} g(\widetilde{A}_t)^{-1/2} L^{11}(O^{11})^\top\, U\,V^\top \mathbf{A}_t = t^{\phi^*/2} g(\widetilde{A}_t)^{-1/2} L^{11}(O^{11})^\top\,\widehat{\mathbf{A}}_t \xrightarrow{d} \mathcal{N}\big(\mathbf{0}, I_{N-1}\big)$$

(in the above formulas the matrices L^{11} and O^{11} are related to $\mathcal{M}^{11}_{det}$ computed under H_0 and so for Φ_0). Hence, under H_0, the test statistics $\|\mathbf{T}_t\|^2$ is asymptotically chi-squared distributed with $(N-1)$ degrees of freedom. This result lets us construct an asymptotic critical region for testing any Φ_0.

Remark B.1. *(Power)* The performance in terms of power of this inferential procedure is strongly related to the matrix Φ_1 considered under the alternative hypothesis H_1. For instance, the leading eigenvalue of Φ under H_0 may not be equal to the one under H_1 and so the test statistics may go to infinity with t at different rates under H_0 or H_1. Moreover, the leading right eigenvector $\mathbf{v}_0$ computed under $H_0 : \Phi = \Phi_0$ may not be equal to the leading right eigenvector $\mathbf{v}_1$ of Φ_1 and so it is possible that $(\mathbf{v}_0^\top\mathbf{A}_t) \neq (\mathbf{v}_1^\top\mathbf{A}_t)$. In that case, we would have, under H_1, that $\mathbf{A}_t \xrightarrow{a.s.} \widetilde{A}_\infty\mathbf{u}_1$ but $(\mathbf{v}_0^\top\mathbf{A}_t) \xrightarrow{a.s.} (\mathbf{v}_0^\top\mathbf{u}_1)\widetilde{A}_\infty$ and hence, since we use $\widetilde{A}_t = \mathbf{v}_0^\top\mathbf{A}_t$ in the test statistics, we can be sure that $g(\widetilde{A}_t) \xrightarrow{a.s.} g(\widetilde{A}_\infty)$ remains valid under H_1 only if $(\mathbf{v}_0^\top\mathbf{u}_1) = 1$. This condition is guaranteed in a framework where the leading left eigenvector $\mathbf{u}$ of Φ is known in advanced (e.g. under the popular balance condition $\Phi^\top \mathbf{1} = \mathbf{1}$) and so in that case we would always have $\mathbf{u}_0 = \mathbf{u}_1 = \mathbf{u}$. Analogously, the columns of U_1 and V_1 computed under $H_1 : \Phi = \Phi_1$ may not be equal to the corresponding U_0 and V_0 derived under $H_0 : \Phi = \Phi_0$. However, if we assume $V_0^\top \mathbf{u}_1 = \mathbf{0}$ (that is automatically true when $\mathbf{u}_0 = \mathbf{u}_1$) we can be sure that, under H_1, we have $U_0 V_0^\top \mathbf{A}_t = U_0 V_0^\top \widehat{\mathbf{A}}_t$, since

$$U_0 V_0^\top \mathbf{A}_t = U_0 V_0^\top (\mathbf{A}_t - \widetilde{A}_t \mathbf{u}_1) = U_0 V_0^\top \widehat{\mathbf{A}}_t.$$

In summary, whenever $\mathbf{u}_0 = \mathbf{u}_1$ we have under H_1 that $\mathbf{T}_t$ is still asymptotically normally distributed with covariance matrix $L^{11}(O^{11})^\top U V^\top \mathcal{M}^{11}_{det} V U^\top O^{11} (L^{11})^\top$, where $\mathcal{M}^{11}_{det}$ is referred to the eigen-structure of Φ_1, while $L^{11}(O^{11})^\top U V^\top$ is referred to Φ_0. Thus, the distance between the identity I and the matrix $L^{11}(O^{11})^\top U V^\top \mathcal{M}^{11}_{det} V U^\top O^{11} (L^{11})^\top$ describes the relation between the asymptotic distribution of $\|\mathbf{T}_t\|^2$ under H_0 and the one under H_1, which determines the power of the test. For instance, note that $E[\|\mathbf{T}_t\|^2] = (N-1)$ under H_0, while $E[\|\mathbf{T}_t\|^2]$ is equal to the trace of $L^{11}(O^{11})^\top U V^\top \mathcal{M}^{11}_{det} V U^\top O^{11} (L^{11})^\top$ under H_1.

B.4. Hypothesis tests on the matrix Φ based on $(\mathbf{B}_t)$

We here want to build for the case $N \geq 2$ a test statistic for $H_0 : \Phi = \Phi_0$, that is based only on the process $(\mathbf{B}_t)$. To this end, analogously as in Section B.3, let us define the matrix $\mathcal{M}^{33}_{det}$ such that $\mathcal{M}^{33}_\infty = g(\widetilde{A}_\infty)\mathcal{M}^{33}_{det}$ and the corresponding matrices L^{33} and O^{33}. Fixed the matrix Φ_0 assumed under H_0, we can obtain the real process $\widetilde{B}_t = \mathbf{v}^\top \mathbf{B}_t$ and the multi-dimensional process

$$\mathbf{B}'_t = \mathbf{B}_t - \widetilde{B}_t \mathbf{u} = U\, V^\top \mathbf{B}_t.$$

By Corollary A.6 and Remark A.3, we get that, under H_0 and the assumption $\mathcal{R}e(\phi^*_2)/\phi^* < 1/2$, we have

$$\mathbf{T}_t = t^{\phi^*/2} g(\phi^* \widetilde{B}_t)^{-1/2} L^{33}(O^{33})^\top \mathbf{B}'_t = t^{\phi^*/2} g(\phi^* \widetilde{B}_t)^{-1/2} L^{33}(O^{33})^\top\, U\, V^\top \mathbf{B}_t \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, I_{N-1}\right)$$

(in the above formulas the matrices L^{33} and O^{33} are related to $\mathcal{M}^{33}_{det}$ computed under H_0 and so for Φ_0). The considerations on the power made for the test based on $\mathbf{A}_t$ in Remark B.1 remain valid as well for the test based on $\mathbf{B}_t$.

Appendix C. Statistical tools in some examples

We here specify the previous general statistical tools in two particular cases: firstly, we consider the case $N = 2$ and, then, we take into account the case when Φ is of the “mean-field” type and $N \geq 2$.

C.1. Case $N = 2$

A matrix Φ with dimension 2×2 and satisfying condition (i) can be reparametrized imposing that the number $\phi \in (0, 1]$ is an eigenvalue with left eigenvector $\mathbf{u}$ such that $r = u_1/u_2$, with $r > 0$, that is

$$r\phi_{11} + \phi_{21} = r\phi \quad \text{and} \quad r\phi_{12} + \phi_{22} = \phi\,.$$

This reparametrization has been inspired by the fact that we typically are able to estimate with high accuracy the leading eigenvalue $\phi^* \in (0, 1]$ and the ratio $r = u_1/u_2 > 0$ of the two componets of the leading left eigenvector $\mathbf{u}$. Hence, we may think that $\phi = \phi^*$ and r are known in practice and the only parameters that are object of the test are ϕ_{21} and ϕ_{12}. We can also calculate the second eigenvalue from the trace of Φ:

$$\phi^*_2 = \phi_{1,1} + \phi_{2,2} - \phi = \phi - \frac{\phi_{2,1} + r^2\phi_{1,2}}{r},$$

so that the condition $\mathcal{R}e(\phi^*_2)/\phi^* < 1/2$ required in the previous results corresponds to the condition

$$2\frac{\phi_{2,1} + r^2\phi_{1,2}}{r} > \phi.$$

Moreover, since according to our notation we have $\mathbf{v}^\top \mathbf{1} = 1$ and $\mathbf{v}^\top \mathbf{u} = 1$, we obtain the following relations:

$$v_1 = \frac{r\phi_{1,2}}{r\phi_{1,2} + \phi_{2,1}}, \qquad v_2 = 1 - v_1, \qquad u_1 = \frac{r}{1 - (1-r)v_1}, \qquad u_2 = \frac{1}{1 - (1-r)v_1},$$

which leads to

$$v_1 = \frac{r\phi_{1,2}}{r\phi_{1,2} + \phi_{2,1}}, \qquad v_2 = \frac{\phi_{2,1}}{r\phi_{1,2} + \phi_{2,1}}, \qquad u_1 = \frac{r\phi_{2,1} + r^2\phi_{1,2}}{\phi_{2,1} + r^2\phi_{1,2}}, \qquad u_2 = \frac{\phi_{2,1} + r\phi_{1,2}}{\phi_{2,1} + r^2\phi_{1,2}}.$$

If we adopt the parametrization $\iota = \phi_{1,2} + \phi_{2,1}$ and $\eta = \frac{\phi_{1,2}}{\phi_{1,2} + \phi_{2,1}}$ so that ι represents the intensity of the whole “interaction” among the two components 1 and 2 due to matrix Φ and η (resp. $(1-\eta)$) gives the percentage of this intensity due to the influence of the first component on the second one (resp., of the second component on the first one). Then, we have

$$\mathbf{u} = \frac{(1-\eta) + r\eta}{(1-\eta) + r^2\eta} \begin{pmatrix} r \\ 1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{v} = \frac{1}{r\eta + (1-\eta)} \begin{pmatrix} r\eta \\ 1-\eta \end{pmatrix},$$

and so

$$\mathbf{u}\mathbf{v}^\top = \frac{1}{(1-\eta)+r^2\eta}\begin{pmatrix} r \\ 1 \end{pmatrix}\begin{pmatrix} r\eta & 1-\eta \end{pmatrix} \qquad \text{and} \qquad UV^\top = \frac{1}{(1-\eta)+r^2\eta}\begin{pmatrix} 1-\eta \\ -r\eta \end{pmatrix}\begin{pmatrix} 1 & -r \end{pmatrix}.$$

Finally, imposing $U^\top U = 1$, we get

$$U = \frac{1}{\sqrt{(1-\eta)^2 + r^2\eta^2}}\begin{pmatrix} 1-\eta \\ -r\eta \end{pmatrix} \qquad \text{and} \qquad V = \frac{\sqrt{(1-\eta)^2+r^2\eta^2}}{(1-\eta)+r^2\eta}\begin{pmatrix} 1 \\ -r \end{pmatrix}.$$

We thus obtain

$$\phi^*\mathbf{v}^\top\Sigma_{det}\mathbf{v} = \frac{\phi}{(r\eta+(1-\eta))^2}\begin{pmatrix} r\eta & 1-\eta \end{pmatrix}\Sigma_{det}\begin{pmatrix} r\eta \\ 1-\eta \end{pmatrix}$$

and, for $2\iota\frac{(1-\eta)+r^2\eta}{r} > \phi$,

$$\mathcal{M}_{det}^{11} = US_{det}^{11}U^\top = \frac{(\iota\frac{(1-\eta)+r^2\eta}{r}-\phi)^2}{2\iota\frac{(1-\eta)+r^2\eta}{r}-\phi}\,\frac{1}{((1-\eta)+r^2\eta)^2}\begin{pmatrix} 1-\eta \\ -r\eta \end{pmatrix}\begin{pmatrix} 1 & -r \end{pmatrix}\Sigma_{det}\begin{pmatrix} 1 \\ -r \end{pmatrix}\begin{pmatrix} 1-\eta & -r\eta \end{pmatrix},$$

$$\mathcal{M}_{det}^{33} = US_{det}^{33}U^\top = \frac{1}{2\iota\frac{(1-\eta)+r^2\eta}{r}-\phi}\,\frac{1}{((1-\eta)+r^2\eta)^2}\begin{pmatrix} 1-\eta \\ -r\eta \end{pmatrix}\begin{pmatrix} 1 & -r \end{pmatrix}\Sigma_{det}\begin{pmatrix} 1 \\ -r \end{pmatrix}\begin{pmatrix} 1-\eta & -r\eta \end{pmatrix}.$$

(where S_{det}^{ii} are the matrices defined in Theorem A.3 with Σ_∞ replaced by Σ_{det}). Therefore, recalling that

$$\widetilde{A}_t = \mathbf{v}^\top\mathbf{A}_t = \frac{r\eta A_{1,t}+(1-\eta)A_{2,t}}{r\eta+(1-\eta)} \qquad \text{and} \qquad \widetilde{B}_t = \mathbf{v}^\top\mathbf{B}_t = \frac{r\eta B_{1,t}+(1-\eta)B_{2,t}}{r\eta+(1-\eta)},$$

and defining the constant

$$c_{r,\eta} = \frac{\begin{pmatrix} r\eta & (1-\eta) \end{pmatrix}\Sigma_{det}\begin{pmatrix} r\eta \\ (1-\eta) \end{pmatrix}}{(r\eta+(1-\eta))^2},$$

the confidence interval based on $\mathbf{A}_t$ and $\mathbf{B}_t$ are respectively

$$CI_{1-\alpha}(\widetilde{A}_\infty) = \widetilde{A}_t \pm \frac{z_\alpha}{t^{\phi/2}}\sqrt{\phi g(\widetilde{A}_t)c_{r,\eta}}, \qquad \text{and} \qquad CI_{1-\alpha}(\widetilde{A}_\infty) = \phi\widetilde{B}_t \pm \frac{z_\alpha}{t^{\phi/2}}\sqrt{\phi g(\phi\widetilde{B}_t)c_{r,\eta}}.$$

Analogously, the test statistics based on $\mathbf{A}_t$ is

$$T_t = t^{\phi/2}\frac{1}{(1-\eta)+r^2\eta}L^{11}(O^{11})^\top\begin{pmatrix} (1-\eta) \\ -r\eta \end{pmatrix}\frac{(A_{1,t}-rA_{2,t})}{g(\widetilde{A}_t)^{1/2}},$$

and the one based on $\mathbf{B}_t$

$$T_t = t^{\phi/2}\frac{1}{(1-\eta)+r^2\eta}L^{33}(O^{33})^\top\begin{pmatrix} (1-\eta) \\ -r\eta \end{pmatrix}\frac{(B_{1,t}-rB_{2,t})}{g(\phi\widetilde{B}_t)^{1/2}}.$$

Remark C.1. In the particular case $\Sigma_{det} = diag(\mathbf{u})$ we have

$$S_{det}^{33} = \frac{1}{\left(2\iota\frac{(1-\eta)+r^2\eta}{r}-\phi\right)}\frac{(r+r^2)((1-\eta)+r\eta)((1-\eta)^2+r^2\eta^2)}{((1-\eta)+r^2\eta)^3},$$

which leads to $L^{33}(O^{33})^\top = (S_{det}^{33})^{-1/2}U^\top$ and so the test statistics based on $\mathbf{B}_t$ becomes

$$T_t = t^{\phi/2}\sqrt{\left(2\iota\frac{(1-\eta)+r^2\eta}{r}-\phi\right)\frac{(1-\eta)+r^2\eta}{(r+r^2)((1-\eta)+r\eta)}}\,\frac{(B_{1,t}-rB_{2,t})}{g(\phi\widetilde{B}_t)^{1/2}}.$$

Then, denoting

$$\Delta = \frac{1}{2} - \frac{\phi_2^*}{\phi} = \frac{\iota}{\phi}\frac{(1-\eta)+r^2\eta}{r} - \frac{1}{2} \qquad \text{and} \qquad q_{r,\eta} = \frac{2((1-\eta)+r^2\eta)}{r(1+r)((1-\eta)+r\eta)},$$

we have

$$\|T_t\|^2 = t^\phi q_{r,\eta}\Delta\,\frac{(B_{1,t}-rB_{2,t})^2}{g(\phi\widetilde{B}_t)/\phi} \xrightarrow{d} \chi^2(N-1). \tag{C.1}$$

Moreover, the confidence interval based on $\mathbf{B}_t$ becomes

$$CI_{1-\alpha}(\widetilde{A}_\infty) = \phi\widetilde{B}_t \pm \frac{z_\alpha}{t^{\phi/2}}\sqrt{\phi g(\phi\widetilde{B}_t)c_{r,\eta}}, \qquad \text{with} \qquad c_{r,\eta} = \frac{(r^3\eta^2+(1-\eta)^2)}{(r\eta+(1-\eta))(r^2\eta+(1-\eta))}. \tag{C.2}$$

C.2. Mean-field case for $N \geq 2$

This first case we consider is the "Mean-field" case, which refers to the family of matrices Φ that can be expressed as follows: for any $1 \leq h, j \leq N$

$$\phi_{j,h} \;=\; \phi\Big(\frac{\imath}{N} \;+\; \delta_{j,h}(1-\imath)\Big) \qquad \text{with } \phi, \imath \in (0,1],$$

where $\delta_{j,h}$ is equal to 1 when $h = j$ and to 0 otherwise. Note that Φ is irreducible and doubly stochastic, we have $\mathbf{v} = N^{-1}\mathbf{1}$ and $\mathbf{u} = \mathbf{1}$ and so the random variable $\widetilde{A}_t$ coincides with the average of the processes $A_{t,h}$, i.e. $N^{-1}\mathbf{1}^\top\mathbf{A}_t$ and $\widehat{\mathbf{A}}_t = \big(I - N^{-1}\mathbf{1}\mathbf{1}^\top\big)\mathbf{A}_t$. Furthermore, in this case, the leading eigenvalue ϕ^* coincides with ϕ, while all the eigenvalues of Φ different from ϕ^* are equal to $\phi(1-\imath)$ and, consequently, the condition $\mathcal{R}e(\phi^*_2)/\phi^* < 1/2$ required in the previous results corresponds to the condition $2\imath > 1$. Finally, since Φ is also symmetric, we have $U = V$ and so $U^\top U = V^\top V = I$ and $UU^\top = VV^\top = I - N^{-1}\mathbf{1}\mathbf{1}^\top$. Notice that the mean-field matrix has two parameters, ϕ and $\imath$, but, since the leading eigenvalue ϕ^* is typically easy to be estimated with high accuracy, we may think that the only parameter that is object of inference in this case is $\imath$, which is the quantity that rules the intensity of the "interaction" among the different components due to matrix Φ. We thus obtain $\phi^*\mathbf{v}^\top\Sigma_{det}\mathbf{v} = \phi N^{-2}(\mathbf{1}^\top\Sigma_{det}\mathbf{1})$ and, for $2\imath > 1$,

$$\mathcal{M}^{11}_{det} = US^{11}_{det}U^\top = \frac{\phi(1-\imath)^2}{2\imath - 1}UV^\top\Sigma_{det}VU^\top = \frac{\phi(1-\imath)^2}{2\imath - 1}(I - N^{-1}\mathbf{1}\mathbf{1}^\top)\Sigma_{det}(I - N^{-1}\mathbf{1}\mathbf{1}^\top),$$

$$\mathcal{M}^{33}_{det} = US^{33}_{det}U^\top = \frac{1}{\phi(2\imath - 1)}UV^\top\Sigma_{det}VU^\top = \frac{1}{\phi(2\imath - 1)}(I - N^{-1}\mathbf{1}\mathbf{1}^\top)\Sigma_{det}(I - N^{-1}\mathbf{1}\mathbf{1}^\top),$$

Therefore, recalling $\widetilde{A}_t = \frac{\mathbf{1}^\top\mathbf{A}_t}{N}$, $\widetilde{B}_t = \frac{\mathbf{1}^\top\mathbf{B}_t}{N}$ and defining $c_N = \frac{\left(\mathbf{1}^\top\Sigma_{det}\mathbf{1}\right)}{N^2}$, the confidence interval based on $\mathbf{A}_t$ and $\mathbf{B}_t$ are respectively

$$CI_{1-\alpha}(\widetilde{A}_\infty) \;=\; \widetilde{A}_t \;\pm\; \frac{z_\alpha}{t^{\phi/2}}\sqrt{\phi g(\widetilde{A}_t)c_N} \qquad \text{and} \qquad CI_{1-\alpha}(\widetilde{A}_\infty) \;=\; \phi\widetilde{B}_t \;\pm\; \frac{z_\alpha}{t^{\phi/2}}\sqrt{\phi g(\phi\widetilde{B}_t)c_N}.$$

Analogously, the test statistics based on $\mathbf{A}_t$ is

$$\mathbf{T}_t = t^{\phi/2}L^{11}(O^{11})^\top UV^\top\,\frac{\mathbf{A}_t}{g(\widetilde{A}_t)^{1/2}} = t^{\phi/2}L^{11}(O^{11})^\top(I - N^{-1}\mathbf{1}\mathbf{1}^\top)\,\frac{\mathbf{A}_t}{g(\widetilde{A}_t)^{1/2}},$$

and the one based on $\mathbf{B}_t$ is

$$\mathbf{T}_t = t^{\phi/2}L^{33}(O^{33})^\top UV^\top\,\frac{\mathbf{B}_t}{g(\phi\widetilde{B}_t)^{1/2}} = t^{\phi/2}L^{33}(O^{33})^\top(I - N^{-1}\mathbf{1}\mathbf{1}^\top)\,\frac{\mathbf{B}_t}{g(\phi\widetilde{B}_t)^{1/2}}.$$

Remark C.2. In the particular case $\Sigma_{det} = I$ we have $\mathbf{1}^\top\Sigma_{det}\mathbf{1} = N$ and hence the following confidence interval based on $\mathbf{B}_t$:

$$CI_{1-\alpha}(\widetilde{A}_\infty) \;=\; \phi\widetilde{B}_t \;\pm\; \frac{z_\alpha}{t^{\phi/2}}\sqrt{\phi g(\phi\widetilde{B}_t)c_N} \qquad \text{with} \qquad c_N = \frac{1}{N}. \tag{C.3}$$

Moreover, we have $\mathcal{M}^{33}_{det} = US^{33}_{det}U^\top$, with

$$UU^\top = (I - N^{-1}\mathbf{1}\mathbf{1}^\top) \qquad \text{and} \qquad \mathcal{S}^{33}_{det} = \frac{1}{\phi(2\imath - 1)}I,$$

which leads to $L^{33}(O^{33})^\top = \sqrt{\phi(2\imath - 1)}U^\top$ and so the test statistics based on $\mathbf{B}_t$ becomes

$$\mathbf{T}_t = t^{\phi/2}\sqrt{\phi(2\imath - 1)}U^\top\,\frac{\mathbf{B}_t}{g(\phi\widetilde{B}_t)^{1/2}},$$

from which we can get

$$\begin{aligned}\|\mathbf{T}_t\|^2 = t^\phi(2\imath - 1)\frac{\mathbf{B}_t^\top(I - N^{-1}\mathbf{1}\mathbf{1}^\top)\mathbf{B}_t}{g(\phi\widetilde{B}_t)/\phi} &= t^\phi(2\imath - 1)\frac{\|\mathbf{B}_t\|^2 - \frac{(\mathbf{1}^\top\mathbf{B}_t)^2}{N}}{g(\widetilde{B}_t)/\phi} \\ &= t^\phi(2\imath - 1)\frac{\left\|\mathbf{B}_t - \left(\frac{\mathbf{1}^\top\mathbf{B}_t}{N}\right)\mathbf{1}\right\|^2}{g(\phi\widetilde{B}_t)/\phi},\end{aligned}$$

that is

$$\|T_t\|^2 = t^\phi 2\Delta\,\frac{\left\|\mathbf{B}_t - \widetilde{B}_t\mathbf{1}\right\|^2}{g(\phi\widetilde{B}_t)/\phi} \xrightarrow{\;d\;} \chi^2(N-1) \qquad \text{with} \qquad \Delta = \frac{1}{2} - \frac{\phi^*_2}{\phi} = \imath - \frac{1}{2}. \tag{C.4}$$

Appendix D. Technical results

For the reader's convenience, we here recall some results used in the above proofs. (For a synthetic review about the notion of stable convergence and its variants, see for instance the appendix of [47] and the references recalled therein.)

Theorem D.1 (A consequence of [49, Theorem A.1]). *On $(\Omega, \mathcal{A}, P)$, let $(\mathcal{F}_t)_t$ be a filtration and, for each $t \geq 1$, let $(L_{t,k})_{k\in\mathbb{N}}$ be a real martingale with respect to $(\mathcal{F}_{t+k-1})_k$, with $L_{t,0} = 0$, which converges in mean to a random variable $L_{t,\infty}$. Set*

$$Y_{t,n} = L_{t,n} - L_{t,n-1} \quad \textit{for } n \geq 1, \quad U_t = \sum_{n\geq 1} Y_{t,n}^2, \quad Y_t^* = \sup_{n\geq 1} |Y_{t,n}|.$$

Further, suppose that $(U_t)_t$ converges almost surely to a positive real ($\mathcal{F}_\infty$-measurable) random variable U and $Y_t^* = O(y_t)$ with y_t a deterministic sequence such that $\lim_{t\to+\infty} y_t = 0$.

Then, with respect to $(\mathcal{F}_t)_t$, the sequence $(L_{t,\infty})_t$ converges to the Gaussian kernel $\mathcal{N}(0, U)$ in the sense of the almost sure conditional convergence (and so also stably).

Lemma D.2 ([28, Lemma S2.2]). *If $a_t \geq 0$, $a_t \leq 1$ for t large enough, $\sum_t a_t = +\infty$, $\delta_t \geq 0$, $\sum_t \delta_t < +\infty$, $b > 0$, $y_t \geq 0$ and $y_{t+1} \leq (1 - a_t)^b y_t + \delta_t$, then $\lim_{t\to+\infty} y_t = 0$.*

Theorem D.3 (A consequence of [50, Proposition 3.1]). *Let $(\mathbf{T}_{t,k+1})_{t\geq m_0, m_0\leq k\leq t}$ be a triangular array of d-dimensional real random vectors, such that, for each fixed t, the finite sequence $(\mathbf{T}_{t,k+1})_{m_0\leq k\leq t}$ is a martingale difference array with respect to a given filtration $(\mathcal{F}_{k+1})_{k\geq m_0}$. Moreover, let $(s_t)_t$ be a sequence of positive real numbers and assume that the following conditions hold:*

(1) $\sum_{k=m_0}^t (s_t \mathbf{T}_{t,k+1})(s_t \mathbf{T}_{t,k+1})^\top = s_t^2 \sum_{k=m_0}^t \mathbf{T}_{t,k+1}\mathbf{T}_{t,k+1}^\top \xrightarrow{P} \mathcal{M}_\infty$, *where $\mathcal{M}_\infty$ is a ($\mathcal{F}_\infty$-measurable) random positive semidefinite matrix;*
(2) $\sup_{m_0\leq k\leq t} |s_t \mathbf{T}_{t,k+1}| = s_t \sup_{m_0\leq k\leq t} \sum_{i=1}^d |T_{t,k+1,i}| = O(y_t)$ *with (y_t) a deterministic sequence such that $\lim_{t\to+\infty} y_t = 0$.*

Then $(s_t \sum_{k=m_0}^t \mathbf{T}_{t,k+1})_t$ converges stably to the Gaussian kernel $\mathcal{N}(\mathbf{0}, \mathcal{M}_\infty)$.

Recall that, according to the notation used here, the above condition 2) means that $\sup_{m_0\leq k\leq t} |s_t \mathbf{T}_{t,k+1}| = s_t \sup_{m_0\leq k\leq t} \sum_{i=1}^d |T_{t,k+1,i}| \leq C y_t$ for a suitable deterministic constant C and t large enough, so that we trivially have the convergence in L^1 of $\sup_{m_0\leq k\leq t} |s_t \mathbf{T}_{t,k+1}|$ toward zero, as requested in [50, Proposition 3.1].

The following result combines together stable convergence and almost sure conditional convergence.

Lemma D.4 (A consequence of [51, Lemma 1]). *Suppose that $(\mathbf{Y}_t^{(1)})$ and $(\mathbf{Y}_t^{(2)})$ are real random vectors such that the first is adapted to a filtration $(\mathcal{F}_t)_t$ and the second is measurable with respect to $\mathcal{F}_\infty$. If $(\mathbf{Y}_t^{(1)})$ converges stably to $\mathcal{N}(0, \mathcal{M}_\infty^{(1)})$ and $(\mathbf{Y}_t^{(2)})$ converges to $\mathcal{N}(0, \mathcal{M}_\infty^{(2)})$ in the sense of the almost sure conditional convergence with respect to $(\mathcal{F}_t)_t$, then*

$$(\mathbf{Y}_t^{(1)} + \mathbf{Y}_t^{(2)}) \xrightarrow{stably} \mathcal{N}(\mathbf{0}, \mathcal{M}_\infty^{(1)} + \mathcal{M}_\infty^{(2)}).$$

References

[1] G. Armano, M.A. Javarone, The beneficial role of mobility for the emergence of innovation, Sci. Rep. 7 (2017) 1781. https://doi.org/10.1038/s41598-017-01955-2
[2] W.B. Arthur, The Nature of Technology: What It Is and How It Evolves, Free Press, 2009.
[3] T.M.A. Fink, M. Reeves, R. Palma, R.S. Farr, Serendipity and strategy in rapid innovation, Nat. Commun. 8 (2002) (2017). https://doi.org/10.1038/s41467-017-02042-w
[4] A. Puglisi, A. Baronchelli, V. Loreto, Cultural route to the emergence of linguistic categories, Proc. Natl. Acad. Sci. 105 (23) (2008) 7936–7940. https://doi.org/10.1073/pnas.0802485105
[5] E.M. Rogers, A. Singhal, M.M. Quinlan, Diffusion of innovations, Taylor and Francis, 2019. https://doi.org/10.4324/9780203710753-35
[6] A. Rzhetsky, J.G. Foster, I.T. Foster, J.A. Evans, Choosing experiments to accelerate collective discovery, Proc. Natl. Acad. Sci. 112 (47) (2015) 14569–14574. https://doi.org/10.1073/pnas.1509757112
[7] A. Bandyopadhyay, D. Thacker, A new approach to Pólya urn schemes and its inifinite color generalization, Ann. Appl. Probab. 32 (1) (2022) 46–79.
[8] C. Mailler, D. Villemonais, Stochastic approximation on noncompact measure spaces and application to measure-valued Pólya processes, Ann. Appl. Probab. 30 (5) (2020) 2393–2438. https://doi.org/10.1214/20-AAP1561
[9] H. Sariev, S. Fortini, S. Petrone, Infinite-color randomly reinforced urns with dominant colors, Bernoulli 29 (1) (2023) 132 – 152. https://doi.org/10.3150/21-BEJ1452
[10] H. Sariev, M. Savov, Characterization of exchangeable measure-valued pólya urn sequences, Electron. J. Probab. 29 (2024) Paper No. 73, 23. https://doi.org/10.1214/24-ejp1132
[11] S. Favaro, A. Lijoi, R.H. Mena, I. Prünster, Bayesian Non-Parametric inference for species variety with a two-parameter poisson-dirichlet process prior, J. R. Stat. Soc. Ser. B 71 (5) (2009) 993–1008. https://doi.org/10.1111/j.1467-9868.2009.00717.x
[12] B. Hansen, J. Pitman, Prediction rules for exchangeable sequences related to species sampling, Statist. Probab. Lett. 46 (2000) 251–256.
[13] A. Lijoi, R. Mena, I. Prünster, Bayesian nonparametric estimation of the probability of discovering new species, Biometrika 94 (2007) 769–786. https://doi.org/10.1093/biomet/asm061
[14] A. Lijoi, I. Prünster, Bayesian Nonparametrics, Models beyond the Dirichlet process, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2010.
[15] J. Pitman, Some Developments of the Blackwell-Macqueen URN scheme, Lect. Notes-Monogr. Ser. 30 (1996) 245–267.
[16] L.F. James, A. Lijoi, I. Prünster, Distributions of linear functionals of two parameter poisson-dirichlet random measures, Ann. Appl. Probab. 18 (2) (2008) 521 – 551. https://doi.org/10.1214/07-AAP462
[17] J. Pitman, M. Yor, The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator, Ann. Probab. 25 (2) (1997) 855–900.
[18] H.S. Heaps, Information Retrieval-Computational and Theoretical Aspects, Academic Press, 1978.
[19] G.K. Zipf, Human Behavior and the Principle of Least Effort, Addison-Wesley, Reading MA (USA), 1949.
[20] S. Janson, Limit theorems for triangular urn schemes, Prob. Theory Relat. Fields 134 (2005) 417–452.

[21] F. Tria, V. Loreto, V.D.P. Servedio, Zipf's, Heaps' and Taylor's laws are determined by the expansion into the adjacent possible, Entropy 20 (10) (2018). https://doi.org/{10.3390/e20100752}
[22] G. Aletti, I. Crimaldi, Twitter as an innovation process with damping effect, Sci. Rep. 11 (1) (2021). https://doi.org/10.1038/s41598-021-00378-4
[23] F. Tria, V. Loreto, V.D.P. Servedio, S.H. Strogatz, The dynamics of correlated novelties, Sci. Rep. 4 (2014). https://doi.org/10.1038/srep05890
[24] S.L. Zabell, The Continuum of Inductive Methods Revisited, Cambridge Studies in Probability, Induction and Decision Theory, Cambridge University Press, 2005, p. 243-274.
[25] S.A. Kauffman, Investigations, Oxford University Press, New York/Oxford, New York/Oxford, 2000.
[26] G. Aletti, I. Crimaldi, A. Ghiglietti, Interacting innovation processes, Sci. Rep. 13 (2023) 17187. https://doi.org/10.1038/s41598-023-43967-1
[27] G. Aletti, I. Crimaldi, A. Ghiglietti, Synchronization of reinforced stochastic processes with a network-based interaction, Ann. Appl. Probab. 27(6) (2017) 3787–3844. https://doi.org/10.1214/17-AAP1296
[28] G. Aletti, I. Crimaldi, A. Ghiglietti, Networks of reinforced stochastic processes: a complete description of the first-order asymptotics, Stochastic Process. Appl. 176 (2024) 104427. https://doi.org/10.1016/j.spa.2024.104427
[29] S. Fortini, S. Petrone, P. Sporysheva, On a notion of partially conditionally identically distributed sequences, Stoch. Process. Appl. 128 (3) (2018) 819–846. https://doi.org/10.1016/j.spa.2017.06.008
[30] I. Iacopini, G. Di Bona, E. Ubaldi, V. Loreto, V. Latora, Interacting discovery processes on complex networks, Phys. Rev. Lett. 125 (2020) 248301. https://doi.org/10.1103/PhysRevLett.125.248301
[31] G. Aletti, I. Crimaldi, A. Ghiglietti, F. Nutarelli, Modeling Innovation Ecosystem Dynamics through Interacting Reinforced Bernoulli Processes, 2025, (????). arXiv:2505.13364 [stat.AP]
[32] B. Bercu, S. Favaro, A martingale approach to gaussian fluctuations and laws of iterated logarithm for ewens-pitman model, Stoch. Process. Their Appl. 178 (2024) 104493. https://doi.org/10.1016/j.spa.2024.104493
[33] R. Gouet, Martingale functional central limit theorems for a generalized polya urn, Ann. Probab. 21 (3) (1993) 1624 – 1639. https://doi.org/10.1214/aop/1176989134
[34] L. Yang, D. Jiang, J. Hu, Z. Bai, Asymptotics for Reinforced Stochastic Processes on Hierarchical Networks, 2025, (????). arXiv:2511.04562 [math.ST]
[35] G. Aletti, A. Ghiglietti, Interacting generalized friedman's urn systems, Stoch. Process. Appl. 127 (8) (2017) 2650–2678. https://doi.org/10.1016/j.spa.2016.12.003
[36] L.-X. Zhang, Central limit theorems of a recursive stochastic algorithm with applications to adaptive designs, Ann. Appl. Probab. 26 (6) (2016) 3630–3658.
[37] C. Monti, J. D'Ignazi, M. Starnini, G. De Francisci Morales, Evidence of demographic rather than ideological segregation in news discussion on reddit, in: Proceedings of the ACM Web Conference 2023, WWW '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 2777-2786. https://doi.org/10.1145/3543507.3583468
[38] D.J. Aldous, G.K. Eagleson, On mixing and stability of limit theorems, Ann. Probab. 6 (2) (1978) 325–331.
[39] I. Crimaldi, Introduzione alla nozione di convergenza stabile e sue varianti (Introduction to the notion of stable convergence and its variants), 57, Unione Matematica Italiana, Monograf s.r.l., Bologna, Italy, Bologna, Italy, 2016. Book written in Italian.
[40] I. Crimaldi, G. Letta, L. Pratelli, A strong form of stable convergence, in: C. Donati-Martin, M. Émery, A. Rouault, C. Stricker (Eds.), Séminaire De Probabilités XL, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007, pp. 203–225. https://doi.org/10.1007/978-3-540-71189-6-9.
[41] P. Hall, C.C. Heyde, Martingale limit theory and its application, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, New York, 1980. Probability and Mathematical Statistics.
[42] A. Rényi, On stable sequences of events, Sankhyā Ser. A 25 (1963) 293 302.
[43] S. Laruelle, G. Pagès, Randomized urn models revisited using stochastic approximation, Ann. Appl. Probab. 23 (4) (2013) 1409–1436.
[44] A. Mokkadem, M. Pelletier, Convergence rate and averaging of nonlinear two-time-scale stochastic approximation algorithms, Ann. Appl. Probab. 16 (3) (2006) 1671–1702. https://doi.org/10.1214/105051606000000448
[45] I. Crimaldi, P. Dai Pra, I.G. Minelli, Fluctuation theorems for synchronization of interacting pólya's urns, Stoch. Process. Appl. 126 (3) (2016) 930–947. https://doi.org/10.1016/j.spa.2015.10.005
[46] G. Aletti, I. Crimaldi, A. Ghiglietti, Networks of reinforced stochastic processes: asymptotics for the empirical means, Bernoulli 25 (4 B) (2019) 3339–3378. https://doi.org/10.3150/18-BEJ1092
[47] G. Aletti, I. Crimaldi, A. Ghiglietti, Interacting reinforced stochastic processes: statistical inference based on the weighted empirical means, Bernoulli 26 (2) (2020) 1098–1138. https://doi.org/10.3150/19-BEJ1143
[48] S. Fortini, S. Petrone, Quasi-Bayes properties of a procedure for sequential learning in mixture models, J. R. Stat. Soc. B: Stat. Methodol. 82 (4) (2020) 1087–1114. https://doi.org/10.1111/rssb.12385
[49] I. Crimaldi, An almost sure conditional convergence result and an application to a generalized Pólya urn, Int. Math. Forum 4 (21–24) (2009) 1139–1156.
[50] I. Crimaldi, L. Pratelli, Convergence results for multivariate martingales, Stoch. Process. Appl. 115 (4) (2005) 571–577. https://doi.org/10.1016/j.spa.2004.10.004
[51] P. Berti, I. Crimaldi, L. Pratelli, P. Rigo, A central limit theorem and its applications to multicolor randomly reinforced urns, J. Appl. Probab. 48 (2) (2011) 527–546. https://doi.org/10.1239/jap/1308662642