

# A MULTI-FACTORIAL INNOVATION MODEL WITH MEAN-FIELD INTERACTION

GIACOMO ALETTI, IRENE CRIMALDI, AND ANDREA GHIGLIETTI

**ABSTRACT.** We introduce an Indian-buffet-type model for multi-factorial innovation in which each arriving agent may exhibit both previously observed and new features. The number of new features follows a power-law behavior, while the probability of selecting an old feature combines self-reinforcement, depending on the feature-specific popularity, with a mean-field interaction term depending on the average popularity of all observed features. The model is governed by the usual innovation parameters (mass, discount and concentration), together with two additional parameters: one controlling the strength of reinforcement against a forcing input toward zero, and one regulating the intensity of the mean-field interaction.

Although the growth of the total number of distinct observed features has the same behavior as in the three-parameter Indian buffet process, the mean-field interaction mechanism produces new asymptotic regimes. For aggregate quantities, including the predictive mean, the averaged number of features per agent, the mean inclusion probability, and the mean feature popularity, the phase transition is determined by the comparison between the discount parameter and the weight of the forcing input. For feature-specific quantities, a further transition appears according to the comparison between the interaction level and a critical threshold. In particular, high interaction leads to an asymptotic synchronization of feature-specific inclusion probabilities.

We establish strong laws and second-order asymptotic results, including central limit theorems in regimes where martingale fluctuations compete with deterministic or random terms. The analysis relies on novel general results for recursive stochastic dynamics, which may be useful beyond the present framework.

## 1. INTRODUCTION

Understanding the mechanisms by which novelties emerge and propagate is of crucial interest across various disciplines, including biology, linguistics, and social sciences. In probabilistic terms, a novelty or innovation is defined as the first occurrence of an event of interest. Widely employed mathematical models for innovation processes are urn models with infinitely many colors and species sampling sequences [5, 31, 32, 39, 43]. However, in these models, at each timestep, each observed agent/item shows only one feature, that may be an old one (that is, already observed in the past) or a new one (that is, appeared for the first time). Instead in the present work, we deal with *multi-factorial innovation processes*, where at each timestep an agent/item is observed and it can exhibit a certain (random) number of features (factors), some old and some new. The total number of features is not specified in advance but is allowed to grow as new data points (agents/items) are observed and it is inferred from the data. Our aim is not to propose a new Bayesian nonparametric prior for posterior inference, but rather to study a tractable stochastic dynamics for multi-factorial innovation with mean-field interaction, so that an old feature may be selected not only because it is itself popular, but also because the overall feature environment is highly active.

The *Indian Buffet Process* (IBP) is a classical model in Bayesian nonparametric feature allocation, designed to describe observations endowed with an unbounded number of latent features. Introduced in [25] and developed in [26, 47], it defines a probability distribution over infinite binary matrices, where each row corresponds to an agent/item and each column to a feature. This allows each agent/item to simultaneously possess multiple characteristics, distinguishing the IBP from species sampling sequences and clustering models, which typically assign each data point to a single class. In this sense, the IBP provides a useful benchmark for probabilistic models of

*Date:* July 15, 2026.

*2010 Mathematics Subject Classification.* Primary: 60F15, 60F05, 60G42. Secondary: 62F15, 91B55, 91D15, 91D30.

*Key words and phrases.* Indian buffet process; feature allocation; innovation process; mean-field interaction; reinforcement; phase transition; strong law; central limit theorem.

G. Aletti is a member of “Gruppo Nazionale per il Calcolo Scientifico (GNCS)” of the Italian Institute “Istituto Nazionale di Alta Matematica (INdAM)”. I. Crimaldi and A. Ghiglietti are members of “Gruppo Nazionale per l’Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA)” of the Italian Institute “Istituto Nazionale di Alta Matematica (INdAM)”. 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).

feature allocation, especially because it combines an infinite-dimensional feature space with a relatively explicit predictive structure.

The generative metaphor of the Indian buffet provides an intuitive interpretation of the process: a sequence of customers (representing agents/items) enters an Indian restaurant with an infinite buffet of dishes (representing features). The first customer samples a  $\text{Poisson}(\alpha)$  number of dishes, while each subsequent customer chooses previously sampled dishes with probability proportional to their popularity, and also tries new dishes according to a  $\text{Poisson}(\alpha/t)$  distribution. The resulting binary feature-allocation matrix is exchangeable, meaning that the order of the observations does not affect their joint distribution. This property ensures analytical tractability and enables efficient posterior inference, making the IBP a natural reference model for exchangeable Bayesian nonparametric feature-allocation priors. At the same time, exchangeability and feature-wise independence limit the ability of the classical IBP to describe innovation mechanisms in which the probability of adopting an old feature depends not only on its own past popularity, but also on the global activity level of the whole system.

Over the years, several generalizations of the IBP have been proposed to extend its applicability, either by preserving exchangeability while enriching the prior structure, or by relaxing exchangeability and independence assumptions. A fundamental theoretical reformulation of the IBP was introduced in [47], who demonstrated that the IBP can be viewed as the marginal distribution of a Beta–Bernoulli process. In this framework, a Beta process defines a random measure generating feature probabilities, while a Bernoulli process specifies feature allocation for each observation. This connection provides a rigorous measure-theoretic foundation, linking the IBP to other Bayesian nonparametric priors such as the Dirichlet Process and enabling efficient inference through conjugacy properties. To capture the heavy-tailed distributions often observed in empirical data, power-law variants of the IBP have been developed. Precisely, the IBP can be extended to capture power-law behavior in the total number of observed features through the three-parameter IBP [9, 46], which introduces a discount parameter (also known as stability exponent), that regulates the asymptotic behavior of the overall number of different observed features, a mass parameter, that controls the total number of new features exhibited by an agent/item and a concentration parameter, that tunes the number of agents/items per feature. Such a model better fits real-world data, but one fundamental assumption persists: the inclusion probabilities of features are independent across features and determined solely by past feature-specific counts.

Since the assumption of independence among features limits its expressiveness in settings where correlations or dependencies among features are crucial, some (non-exchangeable) variants of the IBP have been proposed. Among the most influential ones is the *dependent Indian buffet process* [50], which introduces correlations among feature allocations by making the inclusion probability of each feature a function of covariates or of latent variables modeled through Gaussian processes. This allows the IBP to represent smooth temporal or spatial dependencies, although it increases the computational burden of inference. Another notable development is the *hierarchical Indian buffet process* and related [40, 18], which introduces a hierarchical structure where groups of observations share a common global pool of features while maintaining group-specific feature usage. This model has proven particularly effective in multi-task and transfer learning contexts, enabling information sharing across related datasets but at the cost of higher model complexity and more demanding inference procedures. [19] developed the *restricted Indian buffet process*, which allows control over the distribution of the number of features per observation. This generalization enables more flexible modeling of feature sparsity patterns and partially introduces interactions through restriction mechanisms. However, these restrictions are imposed at the level of the number of active features rather than their pairwise relationships. The *Indian buffet Hawkes process* proposed in [45] models evolving features over time by combining the IBP with Hawkes processes. The interactions in their model occur in the temporal domain, where previous feature activations influence future events via self-exciting processes. While this captures a dynamic form of interaction, the focus is on temporal excitation rather than structural co-dependency among features. [29] introduced a class of *Gibbs-type Indian buffet processes*, where the underlying random measure is generalized to a Gibbs-type measure. This allows for more flexible prior specifications and includes models that exhibit power-law behavior. Although their framework can model complex feature allocation structures, the interactions among features are implicitly encoded through the prior, rather than through an explicit interaction matrix. Another direction of advancement has been to incorporate structural or spatial dependencies among agents/items. For example, the *distance-dependent Indian buffet process* [23] modifies the IBP by making the probability of feature sharing a function of the distance between data points (agents/items), allowing it to model local correlations. Similarly, the *phylogenetic Indian buffet process*

[33] introduces a tree-structured dependency among agents/items, capturing shared evolutionary histories in a Bayesian framework. More recently, the *attraction Indian buffet distribution* [48] has been developed to allow pairwise attraction between agents, which increases the likelihood of feature sharing. Another extension of the IBP is the *Indian buffet process with random weights*, introduced in [6]. In this formulation, each agent is associated with a random weight, that represents the relevance of the agent and enters the feature probabilities. This additional stochastic layer generalizes the Beta–Bernoulli construction and connects the IBP with the theory of randomly reinforced stochastic processes. The random-weight IBP maintains some of the analytical properties of the classical model while offering greater flexibility in capturing heterogeneity among the agents. Finally, recent Bayesian nonparametric works have also considered feature-allocation models in connection with prediction of unseen features. For instance, [10] introduced scaled-process priors for Bayesian nonparametric estimation of unseen genetic variation. Their stable-Beta scaled-process prior enriches the posterior distribution of the number of unseen features, replacing the Poisson posterior structure induced by completely random measure priors with a tractable negative-binomial posterior depending on both the sample size and the number of observed distinct features. This line of work is close to ours at the level of feature-allocation modeling and multi-feature data, but it has a different purpose: it concerns exchangeable Bayesian nonparametric priors and posterior prediction, whereas the present paper studies a non-exchangeable recursive innovation dynamics with explicit mean-field interaction among features.

IBP-type and related feature-allocation models have been used in several domains where observations may exhibit multiple latent or observed features. As examples, we mention bioinformatics [30], network theory [8, 34, 44], causal inference [51], modeling of choices [27] and similarity judgements [37]. Feature-allocation ideas have also been used in Bayesian nonparametric methods for unseen-feature prediction, with particular emphasis on genetic and genomic variation (see the recent paper [10] and the references therein).

Summing up, while the IBP assumes exchangeability and inclusion probabilities independent across features, several recent models have sought to extend the framework by incorporating dependencies. Indeed, ignoring such interactions can lead to oversimplified models that fail to capture the observed structure in real data. Most of these models focus on dependencies among the *rows* of the binary matrix - i.e., the data points (the agents/items) themselves - rather than among the *columns* - i.e., the features. The few models, mentioned above, that deal with feature interaction, offer a more realistic representation of evolving phenomena, but they lose the elegant mathematical formulation of the original IBP and make a detailed theoretical analysis difficult. Moreover they require more complex inference techniques, such as sequential Monte Carlo or dynamic variational methods, and often come at the cost of increased computational complexity.

The present paper takes a different route. We introduce an IBP-inspired, non-exchangeable stochastic process in which feature inclusion probabilities evolve through a combination of feature-specific self-reinforcement and a mean-field interaction term. The mean-field term represents the global level of activity of the observed feature system: a feature may be selected not only because it is itself popular, but also because the overall feature environment is highly active. This mechanism gives a simple and interpretable way to model cross-feature dependence while preserving enough recursive structure to allow a sharp asymptotic analysis. Our proposed model directly introduces interactions in the inclusion probabilities, which permits clearer interpretability and a more faithful representation of domains where features are influenced both by their own popularity and by the global level of activity of the feature system, such as in biological networks, cultural evolution, or recommendation systems. Additionally, our model, characterized by an explicit dependence among features, allows us to conduct a detailed theoretical study, proving various asymptotic results, such as strong laws and central limit theorems for several key quantities, such as the total number of observed features ( $D_t$ ), the averaged number of features exhibited by the agents/items ( $\bar{T}_t$ ), the averaged feature inclusion probability ( $\bar{P}_t$ ), the averaged number of agents/items per feature ( $\bar{K}_t$ ) and also, for feature-specific quantities, such as the inclusion probability ( $P_{t,j}$ ) and the popularity ( $K_{t,j}$ ) for an observed feature  $j$ . While the asymptotic behavior of  $D_t$  is the same as in the classical IBP with three parameters [46], we obtain new phenomena such as the power-law behavior of the averaged quantities and of the dish-specific quantities, not seen in the classical IBP, where, except in the extreme cases  $\beta = 1$ , the empirical mean  $\bar{T}_t$  converges almost surely to a strictly positive real random variable and the feature-popularity grows linearly. The mean-field interaction mechanism also produces a second level of phase transition for feature-specific quantities. In particular, when the interaction parameter is sufficiently large, the feature-specific inclusion probabilities synchronize asymptotically, whereas in the low-interaction regime the long-time behavior remains feature

dependent.

We also highlight the novel Theorem B.3 and Theorem B.4 in the supplementary, which are stated and proven in a general setting and so they could be also applied in other contexts. In a framework of recursive stochastic dynamics, the first one provides conditions to guarantee that the limit random variable of the suitable rescaled process is non-zero with probability one, while the second one deals with the related second-order asymptotic. These results are not merely technical tools for the present model. They address a class of recursive dynamics in which a polynomially rescaled process converges almost surely to a random limit, and where the second-order behavior is determined by the competition between martingale fluctuations and a slowly decaying deterministic forcing term. This setting is different from the standard stochastic-approximation framework centered around a deterministic attracting equilibrium, and it naturally appears in reinforced and innovation-type processes.

The main contribution of the paper is therefore twofold. First, we introduce a tractable interacting multi-factorial innovation process and identify the phase transitions governing both averaged and feature-specific quantities. Second, as a consequence, we develop recursive-dynamics tools that yield strong laws, exact rates, positivity of random limits and central limit theorems for the relevant observables.

The rest of the paper is organized as follows. In Section 2, we illustrate some adopted notation, present the new model, which is the object of this work, and provide an overview of the main results of our study. In Section 3, we collect several preliminary properties of the model, that will be repeatedly used throughout the paper, and we describe the asymptotic behavior of the total number  $D_t$  of the observed features. Section 4 is devoted to the asymptotic behavior of the average quantities  $Z_t$ ,  $\bar{T}_t$ ,  $\bar{K}_t$  and  $\bar{P}_t$ . In Section 5, we derive the asymptotic behavior of the quantities  $K_{t,j}$  and  $P_{t,j}$  associated with each individual feature  $j$ . Section 6 concludes the paper with some final remarks and a discussion of possible directions for future research. Finally, the paper is enriched by a wide supplementary, collecting, among other things, novel general results for recursive dynamics (we refer to Section B), and non-trivial technical results for the considered model (we refer to Section D). Sections, theorems and remarks reported in the supplementary document are referenced with capital letters, while those presented in the main document are referenced with numbers.

## 2. NOTATION, MODEL, AND OVERVIEW OF THE RESULTS

In this section, we introduce the basic notation used throughout the paper, present the interacting multi-factorial innovation model, and conclude with an overview of the main results established in this work.

**2.1. Notation.** In the sequel, given two possibly random sequences  $(\rho_t)$  and  $(V_t)$ , we use:

- the expression  $\rho_t = O(V_t)$  with  $V_t > 0$  in order to indicate that  $\limsup_t |\rho_t|/V_t < +\infty$  with probability one;
- the expression  $\rho_t = o(V_t)$  in order to indicate that we have  $\rho_t/V_t \xrightarrow{a.s.} 0$ ;
- the symbols  $O(V_t)$  and  $o(V_t)$  to denote generic, possibly random, sequences with the above properties.

Note that, if  $V_t = v_t$  is deterministic, the symbols  $O(v_t)$  and  $o(v_t)$  can indicate both random or deterministic sequences. This will usually be clear from the context, but where it is not and it is important to highlight it, we'll write it explicitly. We adopt the notation  $o_P(\cdot)$  when the above limit is in the sense of the convergence in probability, instead of a.s. convergence.

**2.2. The model.** Fix  $\alpha > 0$ ,  $0 \leq \beta \leq 1$ ,  $\theta > 0$ ,  $0 < w \leq 1$  and  $0 \leq \iota \leq 1$ . The first three parameters (i.e.  $\alpha$ ,  $\beta$  and  $\theta$ ) correspond to the three parameters (mass, discount and concentration) that characterize the Indian buffet process introduced in [46]; while the two last parameters (i.e.  $w$  and  $\iota$ ) are the new ones, introduced in order to tune the *forcing input* toward zero in the inclusion probabilities and the *interaction among the dishes*, respectively.

The dynamics is as follows. Time is discrete and indexed by customer arrivals. Customer 1 tries  $N_1 \stackrel{d}{=} \text{Poi}(\lambda_0)$  dishes with  $\lambda_0 = \alpha$ . For each time-step  $t \geq 1$ , let  $\mathcal{O}_t$  be the collection of dishes experimented by the first  $t$  customers, that we call the “old” dishes at time-step  $t$ . Denoting by  $\mathcal{F} = (\mathcal{F}_t)_t$  the (complete) natural filtration associated to the model, then:

- Customer  $t + 1$  selects a subset  $\mathcal{O}_{t+1}^* \subseteq \mathcal{O}_t$ . Each  $j \in \mathcal{O}_t$  is included or not into  $\mathcal{O}_{t+1}^*$  independently of the other members of  $\mathcal{O}_t$  with  $\mathcal{F}_t$ -conditional probability (called *inclusion probability*)

$$(1) \quad P_{t,j} = P(X_{t+1,j} = 1 | \mathcal{F}_t) = w \left[ (1 - \iota) \frac{K_{t,j}}{\theta + t} + \iota \frac{1}{D_t} \sum_{i \in \mathcal{O}_t} \left( \frac{K_{t,i}}{\theta + t} \right) \right].$$

where  $X_{n,i}$  is the indicator of the event {customer  $n$  selects dish  $i$ },  $K_{t,i} = \sum_{n=1}^t X_{n,i}$  is the number of customers who tried dish  $i$  until time-step  $t$  and  $D_t = \text{card}(\mathcal{O}_t)$ .

- In addition to  $\mathcal{O}_{t+1}^*$ , customer  $t + 1$  also tries (independently of the past  $\mathcal{F}_t$  and of the choice of  $\mathcal{O}_{t+1}^*$ ) a random number  $N_{t+1} \stackrel{d}{=} \text{Poi}(\lambda_t)$  of new dishes, where  $\text{Poi}(\lambda_t)$  denotes the Poisson distribution with mean  $\lambda_t = \alpha/(t + 1)^{1-\beta}$ .

Note that in (1) a mean-field interaction term is present: feature  $j$  interacts with each other features through the empirical average popularity of all currently observed features. Thus the model is based on a global interaction field. We can say that the case when  $w = 1$  and  $\iota = 0$  is related to the standard IBP of [46], although in the standard model we have  $P_{t,j} = (K_{t,j} - \beta)/(\theta + t)$  and  $\lambda_t = \alpha\Gamma(\theta + 1)\Gamma(\theta + \beta + t)/[\Gamma(\theta + \beta)\Gamma(\theta + 1 + t)]$ . These definitions for  $P_{t,j}$  and  $\lambda_t$  make the standard IBP to be exchangeable. However, despite these differences, it holds true (as in our model with  $w = 1$  and  $\iota = 0$ ) that  $P_{t,j} \sim K_{t,j}/(\theta + t)$  and  $\lambda_t \sim C/(t + 1)^{1-\beta}$ , for some constant  $C > 0$ , as  $t \rightarrow +\infty$ , and so the asymptotic properties of our model with  $w = 1$  and  $\iota = 0$  are similar to the ones of the standard IBP. Instead, as we will see, the second-order asymptotic behavior presents significant differences. Note that, from an applicative point of view, our simplified way to define the inclusion probabilities  $P_{t,j}$  and the parameter  $\lambda_t$  allows for a better interpretation and a clearer identification of the role played by each single parameter. This is the reason behind our choice to define them as above. Hence, the introduced model should be viewed not as an exchangeable IBP variant intended for Bayesian nonparametric inference, but as an IBP-inspired stochastic model suitable for innovation processes with possibly more than one feature (factor) per agent.

Moreover, in the above introduced model, as in standard IBP, for each fixed time-step  $t + 1$ , the random variables  $\{X_{t+1,j} : j \in \mathcal{O}_t\}$  are conditionally independent given the past; while, differently with respect to the standard model, for each fixed  $j \in \mathcal{O}_t$ , the random variable  $X_{t+1,j}$  is not independent of the random variables  $\{X_{n,i} : i \neq j, n \leq t\}$ . Finally, as in the standard model, the random variable  $N_{t+1}$  is independent of  $N_1, \dots, N_t$  and of all the choices done for the old dishes, i.e. of  $\{X_{n,j} : j \in \mathcal{O}_n, n \leq t + 1\}$ .

The parameters  $\alpha$  (known as the mass parameter) and  $\beta$  (known as the discount parameter or the stability exponent) drive the growth of the number of tested dishes (see Subsection 3.1), while the parameter  $\theta$  (known as the concentration parameter) regulate the initial condition and the behavior of the model in the initial phase (i.e. for a finite number of time-steps). Regarding the parameter  $w$ , we observe that the inclusion probability can be seen as a convex combination of the reinforcement probability with weight  $w \in (0, 1]$  and the null probability with weight  $(1 - w) \in [0, 1)$ . Hence, the parameter  $w$  tunes the weight of the two mechanisms: the *reinforcement* and the *forcing input toward zero*. The first one can be explained by the popularity principle: the more popular is a feature (dish), the higher is the probability that a future agent (customer) selects it. The last mechanism can be explained thinking about the fact that the features (dishes) could be subject to a fad and so their probability to be selected is "forced" to vanish. Finally, the reinforcement term is itself a convex combination of a term that depends on the averaged popularity of the appeared features (tested dishes) with weight  $\iota \in [0, 1]$  and a term that depends only on the specific popularity of the feature (dish)  $j$  with weight  $(1 - \iota) \in [0, 1]$ . Hence, the *interaction* present in the model dynamics is of the *mean-field type*, whose intensity is ruled by the parameter  $\iota$  and in the inclusion probability  $P_{t,j}$  the total weight of the *self-reinforcement* is  $w[(1 - \iota) + \iota/D_t]$ , while the weight of each *cross-reinforcement* (that is the reinforcement coming from a feature  $i \in \mathcal{O}_t \setminus \{j\}$ ) is  $w\iota/D_t$ .

Generally speaking, a mean-field interaction is used when each agent perceives the overall average context, not just the number of similar agents, so that the agent's behavior is influenced by an aggregate perception of the system, rather than by direct contacts. In other words, the system of interest is simplified by replacing the complex pairwise interactions with an effective average field that represents the collective effect of all the others. Hence, in the IBP framework, instead of calculating how every feature (dish) affects every other one (which is extremely difficult when the features are in a large number), we assume that each feature is affected by only an average field generated by all the others appeared in the system. Just to give an example in order to facilitate the interpretation and visualize the framework, consider the following setting. An agent (customer) enters a community and adheres to some existing proposals (old dishes), while also introducing new ones (new dishes).



Each agent decides to join an existing proposal not only based on how many agents have already joined that proposal, but also on the overall average adoption across all proposals - in other words, on a global perception of participation across all the existing proposals. This can be seen as a mean-field interaction process in an “idea space”. Each agent supports existing ideas not only based on how many others support them, but also based on the overall average level of support across all ideas.

**2.3. Overview of the Results.** Regarding the results we are going to prove, we observe that the behavior of the number  $D_t$  of the tested dishes along the time-steps is not affected by the interaction among the dishes and hence the related results are standard. We collect them in the following Section 3 for the reader’s convenience; while the main results of this work concern the asymptotic behaviors of the dish(feature)-specific inclusion probability  $P_{t,j}$ , the dish(feature)-specific popularity  $K_{t,j}$ , the number  $T_t$  of dishes tested by customer  $t$ , that is  $T_t = \sum_{j \in \mathcal{O}_t} X_{t,j} + N_t$  and the corresponding averaged quantities

$$(2) \quad \bar{P}_t = \frac{1}{D_t} \sum_{j \in \mathcal{O}_t} P_{t,j}, \quad \bar{K}_t = \frac{1}{D_t} \sum_{j \in \mathcal{O}_t} K_{t,j} \quad \text{and} \quad \bar{T}_t = \frac{1}{t} \sum_{n=1}^t T_n,$$

that corresponds to the mean dish(feature)-inclusion probability, the mean dish(feature)-popularity and the mean number of dishes (features) per customer (agent/item), respectively, until time-step  $t$ . We also analyze the predictive mean  $Z_t = E[T_{t+1} | \mathcal{F}_t]$  of the number of dishes tested (features exhibited) by the future customer (agent/item)  $t+1$  given the past  $\mathcal{F}_t$ . These results are presented in Section 4 (for the averaged quantities) and in Section 5 (for the specific-dish quantities). In particular, we highlight that the asymptotic behaviors of the averaged quantities only depend on the relationship between the parameters  $\beta$  and  $w$  (see Table 1), while the interaction intensity  $\iota$  plays a role in the results describing the asymptotics for the specific-dish quantities (see Table 2 and Table 3). This is a consequence of the specific form of the interaction considered in this work: the mean-field interaction influences the processes at the level of individual dishes but vanishes when considering the averaged quantities. This represents an advantage from the applicative point of view because this means that we can use the parameters  $\alpha$  and  $\beta$  to fit the growth of the observable number  $D_t$  of distinct dishes (features), we can fit the behavior of the observable averaged quantities  $\bar{T}_t$  by the parameter  $w$ , while we can use the parameter  $\iota$  to fit the observable  $j$ -specific quantities  $K_{t,j}$ .

For the first-order results in Table 1, the main considerations are the following. The predictive mean  $Z_t$  of the number of dishes chosen at the future time  $t+1$  converges to zero when both  $\beta$  and  $w$  are less than 1. The decay is sub-linear (specifically, a power-law or a power-law with a logarithmic correction), with a power-law exponent depending on the weight  $(1-w)$  of the forcing input toward zero or on the power-law exponent  $(1-\beta)$  governing the decrease of the expected number  $\lambda_t$  of new features. When  $\max\{\beta, w\} = 1$ , it converges to a finite strictly positive quantity if  $\beta \neq w$ , whereas it diverges to infinity if  $\beta = w$ . The averaged quantity  $\bar{T}_t$  exhibits exactly the same asymptotic behaviors, but the limit random variables coincide only when  $w = 1$  and so, in that case, we can affirm that the empirical mean  $\bar{T}_t$  can be used as a strongly consistent estimator of the predictive mean  $Z_t$ . These results for  $w = 1$  are consistent with the ones known for the standard IBP of [46] (see [6]), while the others are a consequence of the introduction in the model of the forcing input toward zero. The mean inclusion probability  $\bar{P}_t$  always converges to zero. This result is of particular interest, as it explains the sparsity of the observed feature-matrix. Although  $\bar{P}_t$  always converges to zero, the mean dish-popularity  $\bar{K}_t$  always diverges to  $+\infty$ , except when the  $w < \beta$ , i.e. when the weight  $(1-w)$  of the forcing input toward zero is higher than the power-law exponent  $(1-\beta)$  of the expected number of new dishes. Specifically, when  $w > \beta$ , we have a power-law growth (with or without a logarithmic correction) with an exponent related to the difference between  $w$  and  $\beta$ , a logarithmic growth when  $w = \beta$  and a convergence to a suitable constant when  $w < \beta$ .

Regarding the first-order results in Table 2 and Table 3, we can make the following remarks. After its appearance, the inclusion probability  $P_{t,j}$  of dish  $j$  tends to zero in all cases except when  $\iota = 0$  and  $w = 1$  (that, as already observed, essentially corresponds to the standard IBP of [46]). All the dish-specific inclusion probabilities  $P_{t,j}$  converge at the same rate, meaning that the rate does not depend on  $j$ . The level  $\iota$  of interaction influences the convergence rate at which each process  $(P_{t,j})_t$  converges towards zero. Specifically, the higher the interaction parameter  $\iota$ , the faster the convergence to zero (recall that when  $\iota = 0$  and  $w = 1$ ,  $P_{t,j}$  does not even tend to zero). Additionally, the level of interaction determines how closely the processes  $P_{t,j}$  are asymptotically related to each other. Moreover, comparing Table 1 and Table 3, we can see that, in the most of cases, the averaged inclusion probability  $\bar{P}_t$  is asymptotically lower than the specific inclusion probability  $P_{t,j}$  of any dish  $j$ . The level  $\iota$  of interaction quantifies the difference between  $\bar{P}_t$  and the single  $P_{t,j}$ . Specifically:

- *High interaction* ( $\iota = 1$  or  $\beta/w < \iota < 1$ ): The processes  $P_{t,j}$ , once rescaled, converge towards the same random variable, with the limit not depending on  $j$ . In other words, the processes  $(P_{t,j})_t$  synchronize. The process  $\bar{P}_t$  and the processes  $P_{t,j}$  converge at the same rate, but when  $0 < \beta/w < \iota < 1$  their limits differ.
- *Critical regime* ( $0 < \iota = \beta/w < 1$ ): As above, the processes  $(P_{t,j})_t$  synchronize. The process  $\bar{P}_t$  and the processes  $P_{t,j}$  converge at different rates.
- *Low interaction* ( $\iota = 0$  or  $0 < \iota < \min\{\beta/w, 1\}$ ): The processes  $P_{t,j}$ , once rescaled, converge towards possibly different random variables depending on  $j$ . Moreover, the averaged quantity  $\bar{P}_t$  and the processes  $P_{t,j}$  converge at different rates.

It may seem surprising that, when all the processes  $P_{t,j}$  synchronize, their common behavior still differs from the one of  $\bar{P}_t$ , which is their average. Specifically: in the case of high interaction,  $\bar{P}_t$  converges at the same rate as each  $P_{t,j}$  does, but, when rescaled, its limit is strictly smaller than the common limit toward which all the rescaled  $P_{t,j}$  converge, and, in the critical case, they have even different rates. The reason for this apparent paradox lies in the fact that the total number of dishes involved in the average diverges to infinity, while the rescaled processes  $(P_{t,j})_t$  do not converge uniformly to their common limit. Consequently, although older dishes approach the common limit, there will always be a large number of more recent dishes whose values remain significantly below it.

The  $j$ -specific popularity  $K_{t,j}$  always increases to  $+\infty$ . The growth is logarithmic or according to a power-law and, always sub-linear, except when  $w = 1$  and  $\iota = 0$  (consistently with the behavior in the standard IBP of [46]). We can observe that the smaller the interaction intensity  $\iota$  is, the faster is the growth to  $+\infty$ . This is coherent with the previous observation on  $P_{t,j}$ .

We complete the analysis of the model with some asymptotic results of the second-order. Specifically, for the case  $w > \beta$ , we establish some central limit theorems (see Theorem 4.3 and Corollary 4.4), that could be useful for the estimation of the random variable  $Z_\infty^*$ , present in Tables 1-3, via confidence intervals based on the observable quantity  $\bar{T}_t$ . Similarly, in the case of low-interaction, we establish a central limit theorem (see Theorem 5.4), that could be useful for the estimation of the random variable  $K_{\infty,j}^*$ , present in Table 2 and Table 3, via confidence intervals. In particular, in this result, we identify three regimes according to the relationships between the interaction intensity  $\iota$  and the ratio  $\beta/w$ : the larger  $\iota$  is, the slower the convergence to zero of the difference between the rescaled process  $K_{t,j}$  and its limit  $K_{\infty,j}^*$  is.

We conclude this overview by highlighting the two novel general results, Theorem B.3 and Theorem B.4 in supplementary, that could also be applied in other frameworks. In a context of a recursive dynamics, Theorem B.3 provides conditions to guarantee that the limit random variable of a suitable rescaled process is almost surely non-zero. This fact is important since it allows to conclude that the scaling factor provides the exact rate of convergence. Note that the standard arguments in literature serve to prove only the convergence of the rescaled process to a real random variable, while we succeed to prove that the limit random variable is non-zero with probability one by means of a non-standard technique. Theorem B.4 deals with the related second-order asymptotic and it faces within a framework excluded from the usual central limit theorems applied in the literature when dealing with recursive dynamics, such as the ones in [53].

### 3. PRELIMINARIES

Recall that  $K_{t,j} = \sum_{n=1}^t X_{n,j}$  is the number of customers who test dish  $j$  until time-step  $t$  and let  $T_t$  be the number of dishes tested by customer  $t$ . We have

$$\sum_{n=1}^t T_n = \sum_{n=1}^t \left( \sum_{j \in \mathcal{O}_{n-1}} X_{n,j} + N_n \right) = \sum_{j \in \mathcal{O}_t} K_{t,j}.$$

Hence, setting  $\bar{T}_t = \sum_{n=1}^t T_n / t$ , we can write

$$(3) \quad \bar{K}_t = \frac{\sum_{j \in \mathcal{O}_t} K_{t,j}}{D_t} = \frac{t}{D_t} \bar{T}_t$$

TABLE 1. Almost sure asymptotic behavior of:

- (i) the predictive mean  $Z_t$ ,
  - (ii) the empirical mean  $\bar{T}_t$ ,
  - (iii) the averaged inclusion probability  $\bar{P}_t$ ,
  - (iv) the averaged number of customers per dish  $\bar{K}_t$ .
- The real random variable  $Z_\infty^*$  is a.s. strictly positive

|                 | $Z_t = E[T_{t+1}   \mathcal{F}_t]$              | $\bar{T}_t = \sum_{n=1}^t T_n / t$        | $\bar{P}_t = \sum_{j \in \mathcal{O}_t} P_{t,j} / D_t$ | $\bar{K}_t = \sum_{j \in \mathcal{O}_t} K_{t,j} / D_t$ |
|-----------------|-------------------------------------------------|-------------------------------------------|--------------------------------------------------------|--------------------------------------------------------|
| $0 = \beta < w$ | $t^{-(1-w)} Z_\infty^*$                         | $t^{-(1-w)} \frac{Z_\infty^*}{w}$         | $\frac{t^{-(1-w)} Z_\infty^*}{\ln(t) \alpha}$          | $\frac{t^w Z_\infty^*}{\ln(t) \alpha w}$               |
| $0 < \beta < w$ | $t^{-(1-w)} Z_\infty^*$                         | $t^{-(1-w)} \frac{Z_\infty^*}{w}$         | $t^{-(1-w+\beta)} \frac{Z_\infty^* \beta}{\alpha}$     | $t^{w-\beta} \frac{Z_\infty^* \beta}{\alpha w}$        |
| $0 < \beta = w$ | $t^{-(1-w)} \ln(t) \alpha w$                    | $t^{-(1-w)} \ln(t) \alpha$                | $t^{-1} \ln(t) w^2$                                    | $\ln(t) w$                                             |
| $0 < w < \beta$ | $t^{-(1-\beta)} \frac{\alpha \beta}{\beta - w}$ | $t^{-(1-\beta)} \frac{\alpha}{\beta - w}$ | $t^{-1} \frac{w \beta}{\beta - w}$                     | $\frac{\beta}{\beta - w}$                              |

TABLE 2. Almost sure asymptotic behavior of the number of customers per dish. The real random variables  $Z_\infty^*$  and  $K_{\infty,j}^*$  are a.s. strictly positive.

| $K_{t,j}$       | $\iota = 1$                                            | $\frac{\beta}{w} < \iota < 1$                                                 | $0 < \iota = \frac{\beta}{w} < 1$                          | $\iota = 0$ or $0 < \iota < \min\{\frac{\beta}{w}, 1\}$ |
|-----------------|--------------------------------------------------------|-------------------------------------------------------------------------------|------------------------------------------------------------|---------------------------------------------------------|
| $0 = \beta < w$ | $\frac{t^w Z_\infty^*}{\ln(t) \alpha w}$               | $\frac{t^w}{\ln(t)} \frac{1}{\alpha w} Z_\infty^*$                            | not possible                                               | $t^w K_{\infty,j}^*$                                    |
| $0 < \beta < w$ | $t^{w-\beta} \frac{Z_\infty^* \beta}{\alpha(w-\beta)}$ | $t^{w-\beta} \frac{\iota}{(\iota w - \beta)} \frac{\beta}{\alpha} Z_\infty^*$ | $t^{w-\beta} \ln(t) \iota \frac{\beta}{\alpha} Z_\infty^*$ | $t^{w(1-\iota)} K_{\infty,j}^*$                         |
| $0 < \beta = w$ | $\ln^2(t) \frac{w^2}{2}$                               | not possible                                                                  | not possible                                               | $t^{w(1-\iota)} K_{\infty,j}^*$                         |
| $0 < w < \beta$ | $\ln(t) \frac{w \beta}{\beta - w}$                     | not possible                                                                  | not possible                                               | $t^{w(1-\iota)} K_{\infty,j}^*$                         |

TABLE 3. Almost sure asymptotic behavior of the inclusion probability. The real random variables  $Z_\infty^*$  and  $K_{\infty,j}^*$  are a.s. strictly positive.

| $P_{t,j}$       | $\iota = 1$                                                 | $\frac{\beta}{w} < \iota < 1$                                                                        | $0 < \iota = \frac{\beta}{w} < 1$                                                  | $\iota = 0$ or $0 < \iota < \min\{\frac{\beta}{w}, 1\}$   |
|-----------------|-------------------------------------------------------------|------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------|-----------------------------------------------------------|
| $0 = \beta < w$ | $\frac{t^{-(1-w)} Z_\infty^*}{\ln(t) \alpha}$               | $\frac{t^{-(1-w)}}{\ln(t)} \frac{1}{\alpha} Z_\infty^*$                                              | not possible                                                                       | $\frac{1}{t^{(1-w)}} w K_{\infty,j}^*$                    |
| $0 < \beta < w$ | $\frac{1}{t^{(1-w+\beta)}} \frac{Z_\infty^* \beta}{\alpha}$ | $\frac{1}{t^{(1-w+\beta)}} \frac{\iota(w-\beta)}{(\iota w - \beta)} \frac{\beta}{\alpha} Z_\infty^*$ | $\frac{\ln(t)}{t^{(1-w+\beta)}} (1-\iota) w \iota \frac{\beta}{\alpha} Z_\infty^*$ | $\frac{1}{t^{(1-w(1-\iota))}} (1-\iota) w K_{\infty,j}^*$ |
| $0 < \beta = w$ | $\frac{\ln(t)}{t} w^2$                                      | not possible                                                                                         | not possible                                                                       | $\frac{1}{t^{(1-w(1-\iota))}} (1-\iota) w K_{\infty,j}^*$ |
| $0 < w < \beta$ | $\frac{1}{t} \frac{w \beta}{\beta - w}$                     | not possible                                                                                         | not possible                                                                       | $\frac{1}{t^{(1-w(1-\iota))}} (1-\iota) w K_{\infty,j}^*$ |

and, setting  $S_0 = 0$  and  $S_t = \sum_{j \in \mathcal{O}_t} P_{t,j}$  for  $t \geq 1$ , we have

$$\begin{aligned}
 (4) \quad S_t &= \sum_{j \in \mathcal{O}_t} P_{t,j} = \frac{w(1-\iota) \sum_{j \in \mathcal{O}_t} K_{t,j} + w \iota \sum_{i \in \mathcal{O}_t} K_{t,i}}{\theta + t} \\
 &= \frac{w \sum_{j \in \mathcal{O}_t} K_{t,j}}{\theta + t} = \frac{w \sum_{n=1}^t T_n}{\theta + t} = w \frac{t}{\theta + t} \bar{T}_t.
 \end{aligned}$$

From (4) we also obtain

$$(5) \quad \bar{P}_t = \frac{\sum_{j \in \mathcal{O}_t} P_{t,j}}{D_t} = \frac{w \sum_{j \in \mathcal{O}_t} K_{t,j}}{(\theta + t) D_t} = w \frac{\bar{K}_t}{\theta + t}$$

and so we can rewrite the inclusion probability  $P_{t,j}$  as

$$(6) \quad P_{t,j} = (1-\iota) w \frac{K_{t,j}}{\theta + t} + \iota \bar{P}_t.$$



Moreover, the predictive mean  $Z_t = E[T_{t+1} | \mathcal{F}_t]$  is the sum of two terms, one related to the old dishes and the other related to the rate at which new dishes appear. Precisely, since  $T_{t+1} = \sum_{j \in \mathcal{O}_t} X_{t+1,j} + N_{t+1}$ , we have

$$(7) \quad Z_t = E[T_{t+1} | \mathcal{F}_t] = \sum_{j \in \mathcal{O}_t} P_{t,j} + \lambda_t = S_t + \lambda_t,$$

where  $\lambda_t = \alpha/(t+1)^{1-\beta}$  and  $S_t$  is defined in (4). A simple calculation yields

$$(8) \quad Z_{t+1} - Z_t = -\frac{(1-w)Z_t}{\theta+t+1} + \frac{w(T_{t+1} - Z_t)}{\theta+t+1} + \lambda_{t+1} - \lambda_t + \frac{\lambda_t}{\theta+t+1}$$

so that we obtain

$$E[Z_{t+1} | \mathcal{F}_t] - Z_t = -\frac{(1-w)Z_t}{\theta+t+1} + \lambda_{t+1} - \lambda_t + \frac{\lambda_t}{\theta+t+1}.$$

Note also that we have

$$(9) \quad E[T_{t+1}^2 | \mathcal{F}_t] = Z_t + Z_t^2 - R_t, \quad \text{where} \quad R_t = \sum_{j \in \mathcal{O}_t} P_{t,j}^2,$$

(and so  $0 \leq R_t \leq S_t \leq Z_t$ ), because

$$\begin{aligned} E[T_{t+1}^2 | \mathcal{F}_t] &= E[(\sum_{j \in \mathcal{O}_t} X_{t+1,j} + N_{t+1})^2 | \mathcal{F}_t] \\ &= \sum_{j \in \mathcal{O}_t} P_{t,j} + 2 \sum_{j < j'} P_{t,j} P_{t,j'} + \lambda_t + \lambda_t^2 + 2\lambda_t \sum_{j \in \mathcal{O}_t} P_{t,j} \\ &= \sum_{j \in \mathcal{O}_t} P_{t,j} + \lambda_t + (\sum_{j \in \mathcal{O}_t} P_{t,j} + \lambda_t)^2 - \sum_{j \in \mathcal{O}_t} P_{t,j}^2 \\ &= Z_t + Z_t^2 - R_t. \end{aligned}$$

Finally, we recall that  $D_t$  denotes the number of dishes tested until time-step  $t$  and its asymptotic behavior is described in the following subsection.

**3.1. Asymptotic behavior of  $D_t$ .** We recall that  $D_t = \text{card}(\mathcal{O}_t) = \sum_{n=1}^t N_n$ , where the random variables  $N_n$  are independent and Poisson-distributed with parameter  $\lambda_{n-1}$  so that  $D_t$  has Poisson distribution with parameter  $\Lambda_t = \sum_{n=1}^t \lambda_{n-1} = \sum_{n=1}^t \alpha/n^{1-\beta} \sim \alpha \ln(t)$  when  $\beta = 0$  and  $\sim (\alpha/\beta)t^\beta$  when  $\beta \in (0, 1]$ . (Remember that these are also the asymptotic behaviors of the mean and the variance of  $D_t$ , since they coincide with the parameter  $\Lambda_t$ .)

By standard martingale arguments and Kronecker's lemma, we have the following result regarding the first-order asymptotic behavior of  $D_t$ :

$$(10) \quad \begin{cases} \frac{D_t}{t^\beta} \xrightarrow{a.s.} \frac{\alpha}{\beta} & \text{when } 0 < \beta \leq 1; \\ \frac{D_t}{\ln(t)} \xrightarrow{a.s.} \alpha & \text{when } \beta = 0. \end{cases}$$

**Remark 3.1** (Central limit theorem). Setting

$$\begin{aligned} \lambda(\beta) &= \alpha \text{ if } \beta = 0 \quad \text{and} \quad \lambda(\beta) = \frac{\alpha}{\beta} \text{ if } \beta \in (0, 1], \\ a_t(\beta) &= \ln(t) \text{ if } \beta = 0 \quad \text{and} \quad a_t(\beta) = t^\beta \text{ if } \beta \in (0, 1], \end{aligned}$$

and recalling that

$$\sqrt{a_t(\beta)} \left\{ \frac{\Lambda_t}{a_t(\beta)} - \lambda(\beta) \right\} = \frac{\alpha \sum_{n=1}^t n^{\beta-1} - \lambda(\beta) a_t(\beta)}{\sqrt{a_t(\beta)}} = \frac{O(1)}{\sqrt{a_t(\beta)}} \longrightarrow 0,$$

by a martingale central limit theorem (see [28, Theorem 3.2, page 58]) applied to the martingale difference array  $Y_{t,n} = (N_n - \lambda_{n-1})/\sqrt{a_t(\beta)}$ , with  $t \geq 1$  and  $1 \leq n \leq t$ , we can easily obtain that

$$\sqrt{a_t(\beta)} \left( \frac{D_t}{a_t(\beta)} - \lambda(\beta) \right) \xrightarrow{\text{stably}} \mathcal{N}(0, \lambda(\beta)).$$

Moreover, by a functional central limit theorem for martingales (see [20, Theorem 2.5]) applied to the same martingale difference array  $(Y_{t,n})$ , we can get the functional version of the above limit result.

Finally, the following remark provides an asymptotic bound for the difference  $|D_t - \Lambda_t|$ .

**Remark 3.2** (Law of the iterated logarithm). For each  $\beta \in [0, 1]$ , we have with probability one that

$$\limsup_{t \rightarrow +\infty} \frac{|D_t - \Lambda_t|}{\sqrt{2\Lambda_t \ln(\ln(\Lambda_t))}} = 1.$$

Its proof is based on the classical Kolmogorov's law of the iterated logarithm and standard arguments and it is collected in Section D.1 of the supplementary.

#### 4. ASYMPTOTIC BEHAVIOR OF $\bar{T}_t$ , $\bar{P}_t$ AND $\bar{K}_t$

This section is devoted to deriving the asymptotic properties of the average quantities defined in (2), namely

$$\bar{P}_t = \frac{1}{D_t} \sum_{j \in \mathcal{O}_t} P_{t,j}, \quad \bar{K}_t = \frac{1}{D_t} \sum_{j \in \mathcal{O}_t} K_{t,j}, \quad \text{and} \quad \bar{T}_t = \frac{1}{t} \sum_{n=1}^t T_n.$$

These processes are closely related to the process of predictive means

$$Z_t = E[T_{t+1} | \mathcal{F}_t] = S_t + \lambda_t,$$

as shown by the preliminary relations (3), (4) and (5) in Section 3. Therefore, the theorems in this section focus on the asymptotic behavior of the process  $(Z_t)_t$ , from which the corresponding results for the averaged quantities can be directly derived (see Table 1 and Corollary 4.4). Firstly, we provide the first-order asymptotic properties in Subsection 4.1, and then we turn to the second-order asymptotics in Subsection 4.2.

**4.1. First-order asymptotic results.** From the recursive dynamics of  $(Z_t)_t$  in (8), one might expect that its first-order asymptotic behavior can always be derived via standard stochastic approximation techniques. However, the situation is more subtle. Indeed, the dynamics is essentially governed by two competing contributions:

$$(11) \quad -\frac{(1-w)}{\theta+t+1} Z_t \quad \text{and} \quad (\lambda_{t+1} - \lambda_t) + \frac{\lambda_t}{\theta+t+1}.$$

Heuristically, the first term, driven by the parameter  $(1-w)$  (that is the weight of the forcing input toward zero), acts as a linear drift toward zero (for  $w < 1$ ), suggesting a decay rate of order  $t^{-(1-w)}$ . The second term, induced by the decay (for  $\beta < 1$ ) of the expected number  $\lambda_t$  of new dishes, suggests instead a rate of order  $t^{-(1-\beta)}$ . The interplay between these two effects determines the asymptotic behavior of  $Z_t$ : when the two rates differ, the slower decay dominates; while, in the critical case  $1-w = 1-\beta$ , the two contributions asymptotically balance each other, leading to a cancellation of leading-order terms and resulting in a logarithmic correction, so that the decay becomes slightly slower than  $t^{-(1-w)}$  (equivalently,  $t^{-(1-\beta)}$ ). When  $\max\{w, \beta\} = 1$ , the process  $(Z_t)_t$  does not converge to zero: it converges to a random variable when  $\beta < 1$ , to a constant when  $w < 1$  and it diverges to  $+\infty$  in the case  $w = \beta = 1$ . These facts in the case  $w = 1$  are consistent with the results known for the standard IBP of [46] (see [6]).

It is worth emphasizing that, once the processes are rescaled according to the correct rate, depending on which term in (11) dominates, the proofs become substantially different in the two regimes. Determining the dominant term depends on the values of  $w$  and  $\beta$ , which requires splitting the analysis into two separate theorems: Theorem 4.1 for the case  $w > \beta$ , and Theorem 4.2 for the case  $w \leq \beta$ .

In the first case, when  $w > \beta$ , the second term in (11) behaves essentially as a remainder, and thus  $(Z_t)_t$  becomes a non-negative almost super-martingale and, besides proving the convergence of  $t^{1-w} Z_t$  (which follows from a classical convergence result for non-negative almost super-martingales), we will prove (by a non classical technique) that the corresponding limit random variable is strictly positive, concluding that  $t^{1-w}$  is the exact scaling factor.

In the second case, when  $w \leq \beta$ , the second term in (11) can not be neglected and we obtain a deterministic limit for the suitable rescaled process  $(Z_t)_t$  via stochastic approximation. In particular, if  $\beta = 1$ , the sequence  $(\lambda_t)_t$  does not decay, causing the failure of the convergence to zero for  $Z_t$ . Finally, when  $w = \beta = 1$ , the additional absence of drift leads to the divergence  $Z_t \rightarrow +\infty$ .

The proofs of all the results presented in this section are collected in Subsection C.1 of the supplementary and are essentially based on Theorem B.1, Theorem B.3 and Lemma D.1. In particular, the last one is a non immediate technical lemma needed to control the moments of the unbounded random variables  $Z_t$ . Also relation (9)

plays an important role.

We first present the convergence result in the case  $w > \beta$ .

**Theorem 4.1** (Weak forcing input toward zero). *Let  $\beta < w$  (i.e.  $(1 - w) < (1 - \beta)$ ). Then we have*

$$t^{1-w} Z_t \xrightarrow{a.s.} Z_\infty^*,$$

where  $Z_\infty^*$  is a finite and strictly positive random variable.

From Theorem 4.1 and using the preliminary relations (3), (4) and (5), we can easily derive that

$$t^{1-w} \bar{T}_t \xrightarrow{a.s.} \frac{1}{w} Z_\infty^*, \quad t^{1-w} \sum_{j \in \mathcal{O}_t} P_{t,j} \xrightarrow{a.s.} Z_\infty^*, \quad \frac{1}{t^w} \sum_{j \in \mathcal{O}_t} K_{t,j} \xrightarrow{a.s.} \frac{1}{w} Z_\infty^*,$$

that along with the behavior of  $D_t$  described in (10) leads to the results for  $\bar{P}_t$  and  $\bar{K}_t$  reported in Table 1.

We now state the result in the case  $w \leq \beta$ .

**Theorem 4.2** (Strong forcing input toward zero and Critical case). *Let  $w \leq \beta$  (i.e.  $(1 - w) \geq (1 - \beta)$ ). Then we have*

$$\begin{cases} t^{1-\beta} Z_t \xrightarrow{a.s.} \frac{\alpha\beta}{\beta-w} & \text{when } w < \beta; \\ \frac{t^{1-\beta}}{\ln(t)} Z_t \xrightarrow{a.s.} \alpha\beta & \text{when } w = \beta. \end{cases}$$

From Theorem 4.2 and using the preliminary relations (3), (4) and (5), we can easily derive that: when  $w < \beta$

$$t^{1-\beta} \bar{T}_t \xrightarrow{a.s.} \frac{\alpha}{\beta-w}, \quad t^{1-\beta} \sum_{j \in \mathcal{O}_t} P_{t,j} \xrightarrow{a.s.} \frac{\alpha w}{\beta-w}, \quad \frac{1}{t^\beta} \sum_{j \in \mathcal{O}_t} K_{t,j} \xrightarrow{a.s.} \frac{\alpha}{\beta-w},$$

while when  $w = \beta$

$$\frac{t^{1-\beta}}{\ln(t)} \bar{T}_t \xrightarrow{a.s.} \alpha, \quad \frac{t^{1-\beta}}{\ln(t)} \sum_{j \in \mathcal{O}_t} P_{t,j} \xrightarrow{a.s.} \alpha w, \quad \frac{1}{t^\beta \ln(t)} \sum_{j \in \mathcal{O}_t} K_{t,j} \xrightarrow{a.s.} \alpha,$$

that along with the behavior of  $D_t$  described in (10) leads to the results of  $\bar{P}_t$  and  $\bar{K}_t$  reported in Table 1.

**Remark 4.1** (Convergence in  $L^p$ ). By Lemma D.1, we obtain that the a.s. convergence of the suitably rescaled sequence  $Z_t$  (and the suitably rescaled sequence  $\bar{T}_t$ ) stated in Theorem 4.1 and Theorem 4.2 is also in  $L^p$  for each  $p \in [1, +\infty)$ . Moreover, by Fatou's lemma, the limit random variable  $Z_\infty^*$  that arises in the case  $w > \beta$  (see Theorem 4.1) satisfies  $E[e^{zZ_\infty^*}] < +\infty$  for some  $z > 0$ .

**4.2. Second-order asymptotic in the case when the forcing input toward zero is weak ( $\beta < w$ ).** In this section, we establish some Central Limit Theorems (CLTs) for the random variable  $Z_\infty^* > 0$ , introduced in Theorem 4.1 and related to the a.s. limits of the rescaled predictive mean  $Z_t$  and of the rescaled averaged quantities  $\bar{T}_t$ ,  $\bar{P}_t$  and  $\bar{K}_t$  in the regime  $w > \beta$ . These results are of particular interest when  $w = 1$ , as they enable the estimation of the limit predictive mean  $Z_\infty^*$  via the construction of confidence intervals based on the observable empirical means  $\bar{T}_t$ .

As discussed in Section 4.1, the dynamics of  $(Z_t)_t$  in (8) is essentially governed by the two competing contributions in (11), one driven by the weight  $(1 - w)$  of the forcing input toward zero and the other by  $(1 - \beta)$ . In the case  $w > \beta$  considered here, the first term determines the decay rate  $t^{-(1-w)}$  of  $Z_t$ , while the second one plays no role in the first-order behavior, acting essentially as a remainder term. By contrast, in the second-order asymptotic behavior (see Theorem 4.3), the parameter  $\beta$  plays an important role even in the case  $w > \beta$ . Indeed, by (8) we can observe that the convergence of  $t^{1-w} Z_t$  towards  $Z_\infty^*$  is governed by two competing contributions:

$$t^{1-w} \frac{\Delta M_{t+1}}{\theta + t + 1} \quad \text{and} \quad t^{1-w} \frac{\lambda_t}{\theta + t + 1},$$

where  $\Delta M_{t+1} = w(T_{t+1} - Z_t)$  is an unbounded martingale difference. Heuristically, the first term suggests a martingale CLT with scaling  $t^{-w/2}$ , leading to a zero-mean Gaussian limit kernel. The second term is deterministic and exhibits a decay of order  $t^{-(w-\beta)}$ . The interplay between these two contributions determines the second-order asymptotic behavior of  $t^{1-w} Z_t$ . Specifically, taking into account (12): if  $(w - \beta) > w/2$  (i.e.,  $\beta < w/2$ ), the second term is negligible also at the second order, and the asymptotics are entirely driven by the martingale term which leads to the convergence of  $t^{w/2}(t^{1-w} Z_t - Z_\infty^*)$  to a Gaussian kernel with zero mean; if  $(w - \beta) = w/2$

(i.e.,  $\beta = w/2$ ), we still obtain that  $t^{w/2}(t^{1-w}Z_t - Z_\infty^*)$  converges toward a Gaussian kernel, but the second term contributes with a constant shift in the mean equal to  $-\alpha$ ; if  $0 < w - \beta < w/2$  (i.e.,  $w/2 < \beta < w$ ), the deterministic term dominates so that we obtain  $(t^{1-w}Z_t - Z_\infty^*) \xrightarrow{P} -Ct^{-(w-\beta)}$ , with a suitable constant  $C > 0$ , and the Gaussian fluctuations arise if we consider the quantity  $(t^{1-w}Z_t - Z_\infty^* + Ct^{-(w-\beta)})$  multiplied by the scaling factor  $t^{w/2}$ . It may appear surprising that the asymptotic contribution of the second term is strictly negative in the last two cases. However, this is fully consistent with the fact that, in the dynamics (8), the second term represents a strictly positive contribution. Accordingly, it affects the second-order asymptotics by increasing the value of  $Z_\infty^*$ , and hence by reducing the deviation  $(t^{1-w}Z_t - Z_\infty^*)$ , relative to a dynamics with a random zero-mean remainder term, as typically encountered in similar settings. Moreover, it is worth emphasizing that the variance of the Gaussian limit kernel is always random. Moreover, its random component coincides with  $Z_\infty^*$ , except in the degenerate case of no-interaction ( $\iota = 0$ ) and no-forcing input toward zero ( $w = 1$ ). As we will see in Section 5, this is also the only regime in which the dish-specific inclusion probabilities  $P_{t,j}$  do not converge a.s. to zero (see Table 3).

Finally, note that in Theorem 4.3 the convergence holds true also in the sense of the a.s. conditional convergence, which is a form of convergence introduced in [12] and, subsequently, employed also by others in the urn model and Bayesian literature, e.g. [22, 52]. Specifically, in [22], it is shown how to use a CLT with this form of convergence in order to construct a credible interval in a Bayesian framework. In [52], instead, it is illustrated how to use a CLT with this form of convergence in order to prove that the limit random variable has no point masses.

The proofs of all the results presented in this section are collected in Subsection C.2 of the supplementary and are essentially based on the general Theorem B.4 and on the technical Lemma D.1 and Lemma D.2.

**Theorem 4.3** (CLT for the case of weak forcing input toward zero). *Assume  $\beta < w$ , i.e.  $(1 - w) < (1 - \beta)$ , and let*

$$\Sigma_\infty = \begin{cases} Z_\infty^* - \tilde{R}_\infty > 0 & \text{if } w = 1 \text{ and } \iota = 0; \\ Z_\infty^* > 0 & \text{otherwise.} \end{cases}$$

where  $Z_\infty^*$  is the a.s. limit of  $t^{1-w}Z_t$  introduced in Theorem 4.1 and  $\tilde{R}_\infty$  is the a.s. limit of  $R_t$  as provided in Lemma D.2. Then, we have

$$(12) \quad t^{w/2}(t^{1-w}Z_t + \frac{\alpha\beta}{w-\beta}t^{-(w-\beta)} - Z_\infty^*) \xrightarrow{\text{stably}} \mathcal{N}(0, w\Sigma_\infty),$$

where this convergence also holds true in the sense of the a.s. conditional convergence with respect to the filtration  $\mathcal{F}$ . Moreover, as a consequence, when  $(w - \beta) < w/2$  (i.e.  $w/2 < \beta < w$ ), we have

$$t^{w-\beta}(t^{1-w}Z_t - Z_\infty^*) \xrightarrow{P} -\frac{\alpha\beta}{w-\beta}.$$

**Remark 4.2** (No point-masses for the limit random variable). In the case  $w < 1$  or  $\iota > 0$ , with a similar argument used in [52] (see also the proof of [16, Theorem 2.5]) based on the a.s. conditional convergence toward a Gaussian kernel, we can obtain from (12) that, for any  $\beta < w$ ,  $P(Z_\infty^* = z) = 0$  for each  $z > 0$ .

In the next result, we provide the second order asymptotics for the averaged quantities: the one for  $\bar{T}_t$  is a direct consequence of (12), while the ones for  $\bar{P}_t$  and  $\bar{K}_t$  are also related to the second-order asymptotic of the number  $D_t$  of distinct tested dishes (observed features).

**Corollary 4.4.** *If  $\beta < w$ , then we have*

$$(13) \quad t^{w/2}\left(t^{1-w}\bar{T}_t + \frac{\alpha}{(w-\beta)}t^{-(w-\beta)} - \frac{1}{w}Z_\infty^*\right) \xrightarrow{\text{stably}} \mathcal{N}(0, \Sigma_\infty/w)$$

(where the convergence also holds true in the sense of the a.s. conditional convergence with respect to  $\mathcal{F}$ ) and, as a consequence, when  $w - \beta < w/2$  (i.e.  $w/2 < \beta < w$ ), we have

$$t^{w-\beta}\left(t^{1-w}\bar{T}_t - \frac{1}{w}Z_\infty^*\right) \xrightarrow{P} -\frac{\alpha}{(w-\beta)}.$$

Moreover, using the notation introduced in Remark 3.1, we have

$$(14) \quad \sqrt{a_t(\beta)}\left(t^{1-w}a_t(\beta)\bar{P}_t + \frac{\alpha w}{\lambda(\beta)(w-\beta)}t^{-(w-\beta)} - \frac{1}{\lambda(\beta)}Z_\infty^*\right) \xrightarrow{\text{stably}} \mathcal{N}(0, \frac{(Z_\infty^*)^2}{\lambda(\beta)^3})$$

and, as a consequence, when  $(w - \beta) < \beta/2$  (i.e.  $(2/3)w < \beta < w$ ), we have

$$t^{w-\beta}\left(t^{1-w+\beta}\bar{P}_t - \frac{\beta}{\alpha}Z_\infty^*\right) \xrightarrow{P} -\frac{\beta w}{w-\beta}.$$

Finally, we have

$$(15) \quad \sqrt{a_t(\beta)} \left( \frac{a_t(\beta)}{t^w} \bar{K}_t + \frac{\alpha}{\lambda(\beta)(w-\beta)} t^{-(w-\beta)} - \frac{1}{\lambda(\beta)w} Z_\infty^* \right) \xrightarrow{\text{stably}} \mathcal{N} \left( 0, \frac{(Z_\infty^*)^2}{\lambda(\beta)^3 w^2} \right)$$

and, as a consequence, when  $(w - \beta) < \beta/2$  (i.e.  $(2/3)w < \beta < w$ ), we have

$$t^{w-\beta} \left( \frac{1}{t^{w-\beta}} \bar{K}_t - \frac{\beta}{\alpha w} Z_\infty^* \right) \xrightarrow{P} -\frac{\beta}{w-\beta}.$$

Note that the above limit random variance  $\Sigma_\infty/w$  can be estimated by the observable process  $(T_t)_t$ . Indeed, when  $w < 1$  or  $\iota > 0$ , it can be seen as the a.s. limit of  $t^{1-w} \bar{T}_t$ ; while, when  $w = 1$  and  $\iota = 0$ , by (9) and Lemma E.5, it coincides with the a.s. limit of  $\sum_{n=1}^t T_n^2/t - (\bar{T}_t)^2$ .

**Remark 4.3** (Comparison with the standard IBP of [46]). From [6, Theorem 8], using our notation, we get that the standard IBP of [46] satisfies for every  $\beta < 1$

$$\sqrt{t} (Z_t - Z_\infty^*) \xrightarrow{\text{stably}} \mathcal{N}(0, \Sigma_\infty) \quad \text{and} \quad \sqrt{t} (\bar{T}_t - Z_\infty^*) \xrightarrow{\text{stably}} \mathcal{N}(0, \Sigma_\infty),$$

with  $\Sigma_\infty > 0$ . This is obviously a difference with our model in the case  $w < 1$ , but, importantly, it is a difference with our model even when  $w = 1$  and  $\iota = 0$ . Indeed, as illustrated above, we have a non-centred Gaussian kernel in the case  $\beta = 1/2$  and a convergence in probability with a rate depending on  $(1 - \beta)$  to a suitable constant when  $1/2 < \beta < 1$ . This difference is due to the fact that, also in the case  $w = 1$  and  $\iota = 0$ , the inclusion probabilities  $P_{t,j}$  and the parameter  $\lambda_t$  in our model differ from the ones in the standard IBP, where the inclusion probabilities and the parameter  $\lambda_t$  are defined so that the sequence  $(Z_t)$  of the predictive means is a martingale.

## 5. ASYMPTOTIC BEHAVIOR OF $P_{t,j}$ AND $K_{t,j}$

In this section, we aim to derive the asymptotic behavior of the quantities  $K_{t,j}$  and  $P_{t,j}$  related to each individual observed dish  $j$ . To this end, let  $\tau_j$  be the first time at which a dish  $j$  appears. The observed dishes are the ones with  $\tau_j < +\infty$ . Then, define the process  $B_{t,j} = K_{t,j}/(\theta + t)$  (which is equal to zero for  $t < \tau_j$  and  $> 0$  for  $t \geq \tau_j$ ) whose asymptotic behavior will be essential to establish the one of  $P_{t,j}$  and  $K_{t,j}$ . For any observed dish  $j$ , for  $t \geq \tau_j$ , we have

$$\begin{aligned} B_{t+1,j} - B_{t,j} &= \frac{(K_{t,j} + \Delta K_{t+1,j})}{(\theta + t + 1)} - \frac{K_{t,j}}{(\theta + t)} \\ &= -\frac{1}{\theta + t + 1} B_{t,j} + \frac{1}{\theta + t + 1} \Delta K_{t+1,j}, \end{aligned}$$

that is

$$(16) \quad B_{t+1,j} = \left( 1 - \frac{1}{\theta + t + 1} \right) B_{t,j} + \frac{1}{\theta + t + 1} \Delta K_{t+1,j},$$

where

$$E[\Delta K_{t+1,j} \mid \mathcal{F}_t] = P_{t,j} = (1 - \iota)w B_{t,j} + \iota \bar{P}_t.$$

A simple calculation yields to

$$(17) \quad B_{t+1,j} - B_{t,j} = -\frac{(1 - (1 - \iota)w)}{\theta + t + 1} B_{t,j} + \frac{\iota \bar{P}_t}{\theta + t + 1} + \frac{\Delta M_{t+1,j}}{\theta + t + 1},$$

where  $\Delta M_{t+1,j} = (\Delta K_{t+1,j} - P_{t,j})$  is a martingale difference.

Firstly, we establish the first-order asymptotic properties in Subsection 5.1, and then we turn to the second-order asymptotics in Subsection 5.2.



**5.1. First-order asymptotic results.** The dynamics of  $B_{t,j}$  in (17) is essentially governed by two competing contributions:

$$(18) \quad -\frac{(1 - (1 - \iota)w)}{\theta + t + 1} B_{t,j} \quad \text{and} \quad \frac{\iota \bar{P}_t}{\theta + t + 1}.$$

Heuristically, the first term, driven by the parameter  $(1 - (1 - \iota)w)$ , acts as a linear drift toward zero (for  $(1 - \iota)w < 1$ ), suggesting a decay rate of order  $t^{-(1 - (1 - \iota)w)}$ . The second term, proportional to the interaction intensity  $\iota$  and to the random averaged quantity  $\bar{P}_t$ , suggests instead a rate as derived by means of Theorems 4.1 and Theorem 4.2, and described in Table 1. The interplay between these two effects determines the asymptotic behavior of  $B_{t,j}$  (and hence of  $K_{t,j}$  and  $P_{t,j}$ ): when the two rates differ, the slower decay dominates; while, in the critical case, when the two contributions asymptotically balance each other, the leading-order terms disappear and a logarithmic correction appears. Once the processes are rescaled according to the right rate, depending on which term in (18) dominates, the proofs become substantially different in the various regimes. The dominant term depends on the values of  $\iota$  with respect to the ratio  $\beta/w$  and, to better highlight how these results change with the level of interaction  $\iota$ , we will categorize the results into three distinct cases:

- (a) *Low interaction*: where the interaction parameter  $\iota$  is either zero (no interaction) or smaller than the threshold  $\beta/w$ ;
- (b) *High interaction*: where the interaction parameter  $\iota$  equals one (maximum interaction) or it is smaller than one but exceeds the threshold  $\beta/w$ ;
- (c) *Critical regime*: where the interaction parameter  $\iota$  exactly equals the threshold  $\beta/w$ .

In the first case (a), the second term in (18) behaves essentially as a remainder, and thus  $(B_{t,j})_t$  becomes a non-negative almost super-martingale and we will prove (by the non classical general result Theorem B.3) that its right scaling factor is  $t^{(1 - (1 - \iota)w)}$ . In the other cases (b)-(c), the second term in (18) is of the same order as the first one, and we obtain the desired results via a general stochastic approximation theorem with a random “attractor” (see Theorem B.1).

The fact that the presence of interaction ( $\iota > 0$ ) causes each  $P_{t,j}$  to approach zero (even when the weight  $(1 - w)$  of the forcing input toward zero is null) can be understood by recalling that, according to (6),  $P_{t,j}$  is a linear combination involving the term  $\iota \bar{P}_t$ , which always tends to zero (see Table 1). Hence, a higher value of  $\iota$  amplifies the effect of  $\bar{P}_t$ , accelerating the convergence of  $P_{t,j}$  to zero.

**5.1.1. Statement of the results.** In this section we are going to provide the first-order asymptotic results about the dish-specific quantities  $K_{t,j}$  and  $P_{t,j}$ . The first theorem holds in the case when the interaction is low.

**Theorem 5.1** (Case (a): Low interaction). *When*

- (1)  $\iota = 0$  (no interaction at all) or
- (2)  $0 < \iota < \min\{\beta/w, 1\}$ ,

for each observed dish  $j$ , we have

$$\frac{K_{t,j}}{t^{(1 - \iota)w}} \xrightarrow{a.s.} K_{\infty,j}^*$$

and

$$t^{1 - (1 - \iota)w} P_{t,j} \xrightarrow{a.s.} (1 - \iota)w K_{\infty,j}^*,$$

where  $K_{\infty,j}^*$  is a finite and strictly positive random variable.

This result for  $w = 1$  and  $\iota = 0$  is coherent with the one for the standard IBP of [46].

The next theorem holds in the case when the interaction is high.

**Theorem 5.2** (Case (b): High interaction). *When*

- (1)  $\iota = 1$ , for each observed dish  $j$ , we have

$$P_{t,j} = \bar{P}_t \quad \forall t \geq 1 \quad \text{and} \quad \frac{K_{t,j}}{\sum_{n=1}^t \bar{P}_{n-1}} \xrightarrow{a.s.} 1.$$

- (2)  $0 < \beta/w < \iota < 1$ , for each observed dish  $j$ , we have

$$\frac{K_{t,j}}{t^{w - \beta}} \xrightarrow{a.s.} K_{\infty} = \frac{\iota}{(\iota w - \beta)} \frac{\beta}{\alpha} Z_{\infty}^*$$

and

$$t^{1-w+\beta} P_{t,j} \xrightarrow{a.s.} \frac{\iota(w-\beta)}{(\iota w - \beta)} \frac{\beta}{\alpha} Z_{\infty}^*,$$

where  $Z_{\infty}^* > 0$  a.s. is defined in Theorem 4.1;

(3)  $0 = \beta/w < \iota < 1$  (i.e.  $\beta = 0$  and  $0 < \iota < 1$ ), for each observed dish  $j$ , we have

$$\frac{K_{t,j}}{t^w / \ln(t)} \xrightarrow{a.s.} K_{\infty} = \frac{1}{\alpha w} Z_{\infty}^*$$

and

$$t^{1-w} \ln(t) P_{t,j} \xrightarrow{a.s.} \frac{1}{\alpha} Z_{\infty}^*.$$

Finally, the last theorem holds in the critical case.

**Theorem 5.3** (Case (c): Critical regime). *When  $0 < \iota = \beta/w < 1$ , for each observed dish  $j$ , we have*

$$\frac{K_{t,j}}{t^{w-\beta} \ln(t)} \xrightarrow{a.s.} K_{\infty} = \iota \frac{\beta}{\alpha} Z_{\infty}^*$$

and

$$\frac{t^{1-w+\beta}}{\ln(t)} P_{t,j} \xrightarrow{a.s.} (1-\iota)w\iota \frac{\beta}{\alpha} Z_{\infty}^*.$$

### 5.1.2. Proofs.

*Proof of Theorem 5.1 (Low interaction).* Considering the process  $B_{t,j} = K_{t,j}/(\theta + t)$  and denoting  $B_{t,j}^* = \zeta_t(1 - (1-\iota)w)B_{t,j}$ , from the dynamics (17) we obtain

$$(19) \quad B_{t+1,j}^* = \left(1 + O\left(\frac{1}{t^2}\right)\right) B_{t,j}^* + \frac{\zeta_{t+1}(1-(1-\iota)w)}{\theta+t+1} (\Delta M_{t+1,j} + \iota \bar{P}_t)$$

and so, by the asymptotics for  $\bar{P}_t$  (see Table 1) and the classical Theorem E.4, the process  $(B_{t,j}^*)_t$  is a non-negative almost super-martingale that converges a.s. to a finite r.v.  $B_{\infty}^*$ . However, this is not enough for our scopes. Indeed, in that way, we do not know that such random limit  $B_{\infty}^*$  is a.s. strictly positive, and this is crucial for proving the right rate of convergence. For this reason, we are going to apply the more sophisticated Theorem B.3 (case i) together with Remark B.1 with  $\tau = \tau_j$  to dynamics (16) with  $X_t = B_{t,j}$ ,  $Y_{t+1} = \Delta K_{t+1,j} \in \{0, 1\}$ ,  $\delta = (1-\iota)w \in (0, 1]$ . To this purpose, we note that  $\rho_t = 0$  when  $\iota = 0$ ; while  $\rho_t = \iota \bar{P}_t > 0$  when  $\iota > 0$ . To handle the case  $\iota > 0$ , from Table 1 we get  $\bar{P}_t = O(1/t^{1-w+\beta})$  when  $0 < \beta < w \leq 1$ ,  $\bar{P}_t = O(\ln(t)/t)$  when  $0 < \beta = w \leq 1$  and  $\bar{P}_t = O(1/t)$  when  $0 < w < \beta \leq 1$  (the case  $\beta = 0$  is not possible with  $\iota > 0$  because  $\beta > \iota w$  in the case of low interaction). Therefore, setting  $\delta = (1-\iota)w$ , we have

$$\sum_t \frac{\rho_t}{(\theta+t)^{\delta-\epsilon}} < +\infty \text{ a.s. for some } 0 < \epsilon < \delta,$$

provided that  $\iota = 0$  or  $0 < \iota < \beta/w$ . Moreover, we trivially (since  $Y_{t+1} \in \{0, 1\}$ ) have

$$E[Y_{t+1}^2 \mid \mathcal{F}_t] \leq E[Y_{t+1} \mid \mathcal{F}_t].$$

Hence, under our conditions, we find

$$t^{1-(1-\iota)w} B_{t,j} \xrightarrow{a.s.} B_{\infty,j}^* \in (0, +\infty)$$

and so the first statement of the theorem follows with  $K_{\infty,j}^* = B_{\infty,j}^*$ . Finally, it follows, by (6), that

$$t^{1-w(1-\iota)} P_{t,j} \xrightarrow{a.s.} (1-\iota)w B_{\infty,j}^* = (1-\iota)w K_{\infty,j}^* \in (0, +\infty),$$

which is the second statement of the theorem.  $\square$

*Proof of Theorem 5.2 (High interaction).* We will prove the two cases separately.

*Proof of case (1):* From (6) we immediately get  $P_{t,j} = \bar{P}_t$  for each  $j \in \mathcal{O}_t$  and so the asymptotic behavior of  $P_{t,j}$  is described in Table 1, while, by Lemma E.2, for  $K_{t,j}$  we have  $K_{t,j}/\sum_{n=1}^t \bar{P}_{n-1} \xrightarrow{a.s.} 1$  because  $K_{t,j} = \sum_{n=1}^t X_{n,j}$  with  $E[X_{n+1,j} \mid \mathcal{F}_n] = P_{n,j} = \bar{P}_n$  and  $\sum_n \bar{P}_n = +\infty$  a.s..

*Proof of case (2):* Recall that we have  $B_{t,j} = K_{t,j}/(\theta + t)$  with dynamics (17):

$$B_{t+1,j} = B_{t,j} - \frac{(1-(1-\iota)w)B_{t,j} - \iota \bar{P}_t}{\theta+t+1} + \frac{\Delta M_{t+1,j}}{\theta+t+1},$$

where

$$\Delta M_{t+1,j} = \Delta K_{t+1,j} - P_{t,j} = \Delta K_{t+1,j} - (1-\iota)wB_{t,j} - \iota\bar{P}_t.$$

When  $\iota w > \beta > 0$ , the term  $\iota\bar{P}_t$  is not negligible as in Theorem 5.1 and so we have to study the dynamics of  $B_{t,j}$  paired with the one of  $\bar{P}_t$ . From Table 1 we know that, since  $w > \beta$ , we have  $t^{1-w+\beta}\bar{P}_t \xrightarrow{a.s.} P_\infty^* > 0$ . Now, let  $B_{t,j}^* = \zeta_t(1-w+\beta)B_{t,j}$ ,  $\bar{P}_t^* = \zeta_t(1-w+\beta)\bar{P}_t$  and  $\Delta M_{t+1,j}^* = \zeta_{t+1}(1-w+\beta)\Delta M_{t+1,j}$ . Then, using (25), we obtain

$$\begin{aligned} B_{t+1,j}^* &= \left( B_{t,j}^* - \frac{(1-(1-\iota)w)B_{t,j}^* - \iota\bar{P}_t^*}{\theta+t+1} \right) \frac{\zeta_{t+1}(1-w+\beta)}{\zeta_t(1-w+\beta)} + \frac{\Delta M_{t+1,j}^*}{\theta+t+1} \\ &= \left( B_{t,j}^* - \frac{(1-(1-\iota)w)B_{t,j}^* - \iota\bar{P}_t^*}{\theta+t+1} \right) \left( 1 + \frac{1-w+\beta}{\theta+t+1} \right) + \frac{\Delta M_{t+1,j}^*}{\theta+t+1} \\ &= B_{t,j}^* - \frac{(\iota w - \beta)B_{t,j}^* - \iota\bar{P}_t^*}{\theta+t+1} + \frac{\Delta M_{t+1,j}^*}{\theta+t+1} + \rho_t, \end{aligned}$$

with  $\rho_t = O\left(\frac{B_{t,j} + \bar{P}_t}{t^{1+w-\beta}}\right)$  being  $\mathcal{F}_t$ -measurable. This suggests that  $(B_t^*)_t$  is not a non-negative almost supermartingale, as in Theorem 5.1, but instead it evolves according to a stochastic approximation dynamics, which should converge to  $\iota P_\infty^*/(\iota w - \beta)$ . For proving this convergence, we are going to apply Theorem B.1 with  $X_t = B_t^*$ , together with Lemma B.2. To this end, we observe that, by (9) and since  $0 \leq P_{t,j} \leq 1$ ,

$$\begin{aligned} E[(\Delta M_{t+1,j}^*)^2 | \mathcal{F}_t] &= \zeta(1-w+\beta)^2 P_{t,j}(1-P_{t,j}) \\ &\leq \zeta(1-w+\beta)^2 P_{t,j} \\ &= \zeta(1-w+\beta)[(1-\iota)wB_{t,j}^* + \iota\bar{P}_t^*] \\ &= O(t^{1-w+\beta+\epsilon}) \quad \forall \epsilon > 0, \end{aligned}$$

where we have used  $B_t^* = o(t^\epsilon)$  that follows by Lemma B.2 applied to  $X_t = B_t^*$  with  $A_t = \bar{P}_t^*$  and  $b = (\iota w - \beta) > 0$ . (Indeed,  $A_t = \bar{P}_t^*$  is a.s. convergent by Table 1 and, since both  $B_{t,j}$  and  $\bar{P}_t$  are bounded by one and  $w > \beta$ , we a.s. have  $\sum_t X_t/t^2 = \sum_t O(B_{t,j}/t^{1-w+\beta}) < +\infty$  and  $\sum_t |\rho_t| = \sum_t O\left(\frac{B_{t,j} + \bar{P}_t}{t^{1+w-\beta}}\right) < +\infty$ .) Hence, choosing a suitable  $\epsilon$ , we get

$$\sum_t \frac{1}{(\theta+t+1)^2} E[(\Delta M_{t+1,j}^*)^2 | \mathcal{F}_t] < +\infty \quad \text{a.s.}$$

and so, by Theorem B.1, we have

$$B_{t,j}^* = \zeta_t(1-w+\beta)B_{t,j} \xrightarrow{a.s.} \frac{\iota}{(\iota w - \beta)} P_\infty^*.$$

From Table 1 (case  $w > \beta > 0$ ) we get  $P_\infty^* = \frac{\beta}{\alpha} Z_\infty^* > 0$  a.s. and so the first statement of case (2) of the theorem follows. Finally, by (6), we get

$$t^{1-w+\beta} P_{t,j} \xrightarrow{a.s.} \frac{(1-\iota)w\iota}{(w\iota-\beta)} P_\infty^* + \iota P_\infty^* = \frac{\iota(w-\beta)}{(\iota w - \beta)} P_\infty^* = \frac{\iota(w-\beta)}{(\iota w - \beta)} \frac{\beta}{\alpha} Z_\infty^*,$$

which is the second statement of the case (2) of the theorem.

*Proof of case (3):* When  $\iota w > \beta = 0$ , analogously to case (2), the term  $\iota\bar{P}_t$  is not negligible in the dynamics (17) of  $B_{t,j}$  and so we have to study the dynamics of  $B_{t,j}$  paired with the one of  $\bar{P}_t$ . From Table 1 we know that, since  $w > \beta = 0$ , we have  $t^{1-w} \ln(t)\bar{P}_t \xrightarrow{a.s.} P_\infty > 0$ , so the dynamics we need to consider is

$$B_{t+1,j} = B_{t,j} - \frac{(1-(1-\iota)w)B_{t,j} - \iota\bar{P}_t}{\theta+t+1} + \frac{\Delta M_{t+1,j}}{\theta+t+1},$$

where  $\Delta M_{t+1,j} = \Delta K_{t+1,j} - P_{t,j} = \Delta K_{t+1,j} - (1-\iota)wB_{t,j} - \iota\bar{P}_t$ . Now, let  $B_{t,j}^* = \zeta_t(1-w)\ln(\theta+t)B_{t,j}$ ,  $\bar{P}_t^* = \zeta_t(1-w)\ln(\theta+t)\bar{P}_t$  and  $\Delta M_{t+1,j}^* = \zeta_{t+1}(1-w)\ln(\theta+t+1)\Delta M_{t+1,j}$ . Then, using (27) from Section A of the supplementary, we obtain

$$\begin{aligned} B_{t+1,j}^* &= \left( B_{t,j}^* - \frac{(1-(1-\iota)w)B_{t,j}^* - \iota\bar{P}_t^*}{\theta+t+1} \right) \frac{\zeta_{t+1}(1-w)}{\zeta_t(1-w)} \frac{\ln(\theta+t+1)}{\ln(\theta+t)} + \frac{\Delta M_{t+1,j}^*}{\theta+t+1} \\ &= \left( B_{t,j}^* - \frac{(1-(1-\iota)w)B_{t,j}^* - \iota\bar{P}_t^*}{\theta+t+1} \right) \left( 1 + \frac{1-w}{\theta+t+1} - \frac{1}{(\theta+t+1)\ln(\theta+t+1)} + O\left(\frac{1}{t^2 \ln(t)}\right) \right) \\ &\quad + \frac{\Delta M_{t+1,j}^*}{\theta+t+1} \end{aligned}$$

$$\begin{aligned}
&= B_{t,j}^* - \iota \frac{wB_{t,j}^* - \bar{P}_t^*}{\theta+t+1} - \frac{B_{t,j}^*}{(\theta+t+1)\ln(\theta+t+1)} + \frac{\Delta M_{t+1,j}^*}{\theta+t+1} + \rho_t \\
&= B_{t,j}^* + \frac{1}{\theta+t+1} \left( \iota \bar{P}_t^* - \left( \iota w + \frac{1}{\ln(\theta+t+1)} \right) B_{t,j}^* \right) + \frac{\Delta M_{t+1,j}^*}{\theta+t+1} + \rho_t,
\end{aligned}$$

with  $\rho_t = O\left(\frac{B_{t,j}^* + \bar{P}_t^*}{t^{1+w}}\right)$  being  $\mathcal{F}_t$ -measurable. We want to apply Theorem B.1 with  $X_t = B_t^*$ , together with Lemma B.2. To this end, analogously to the proof of case (2), we have that

$$\begin{aligned}
E[(\Delta M_{t+1,j}^*)^2 \mid \mathcal{F}_t] &= \zeta(1-w)^2 \ln^2(\theta+t+1) P_{t,j}(1-P_{t,j}) \\
&\leq \zeta(1-w)^2 \ln^2(\theta+t+1) P_{t,j} \\
&= \zeta(1-w) \ln(\theta+t+1) [(1-\iota)wB_{t,j}^* + \iota \bar{P}_t^*] \\
&= O(t^{1-w+\epsilon} \ln(t)) \quad \forall \epsilon > 0,
\end{aligned}$$

where we have used  $B_t^* = o(t^\epsilon)$  that follows by Lemma B.2 applied to  $X_t = B_t^*$ , with  $A_t = \iota \bar{P}_t^*$  and  $b_t = \iota w + 1/\ln(\theta+t+1)$  (note that we a.s. have  $\sum_t X_t/t^2 = \sum_t O(B_{t,j} \ln(t)/t^{1+w}) < +\infty$  and  $\sum_t |\rho_t| = \sum_t O\left(\frac{B_{t,j} + \bar{P}_t}{t^{1+w}}\right) < +\infty$  as  $w > 0$ ). Hence, choosing a suitable  $\epsilon$ , we get

$$\sum_t \frac{1}{(\theta+t+1)^2} E[(\Delta M_{t+1,j}^*)^2 \mid \mathcal{F}_t] < +\infty \quad \text{a.s.}$$

Then, by Theorem B.1, we have

$$B_{t,j}^* = \zeta_t(1-w) \ln(\theta+t) B_{t,j} \xrightarrow{\text{a.s.}} \frac{1}{w} P_\infty^*.$$

From Table 1 (case  $w > \beta = 0$ ) we get  $P_\infty^* = \frac{1}{\alpha} Z_\infty^* > 0$  a.s. and so the first statement of the case (3) of the theorem follows. Finally, by (6), we get

$$t^{1-w} \ln(t) P_{t,j} \xrightarrow{\text{a.s.}} (1-\iota)w \frac{1}{w} P_\infty^* + \iota P_\infty^* = P_\infty^* = \frac{1}{\alpha} Z_\infty^*,$$

which is the second statement of the case (3) of the theorem.  $\square$

*Proof of Theorem 5.3 (Critical regime).* The structure of this proof is analogous to the proof of Theorem 5.2 as, also in the case  $\iota w = \beta > 0$ , we have to study the dynamics (17) of  $B_{t,j} = K_{t,j}/(\theta+t)$  paired with the one of  $\bar{P}_t$ . From Table 1 we know that, since  $w > \beta$  (because  $0 < \iota < 1$  and  $\iota w = \beta$ ), we have  $t^{1-w+\beta} \bar{P}_t \xrightarrow{\text{a.s.}} P_\infty^* > 0$ . Now, let  $B_{t,j}^* = \frac{\zeta_t(1-w+\beta)}{\ln(\theta+t)} B_{t,j}$ ,  $\bar{P}_t^* = \zeta_t(1-w+\beta) \bar{P}_t$  and  $\Delta M_{t+1,j}^* = \zeta_{t+1}(1-w+\beta) \Delta M_{t+1,j}$ . Then, using (25) and (27) from Section A of the supplementary, we obtain

$$\begin{aligned}
B_{t+1,j}^* &= \left( 1 - \frac{(1-(1-\iota)w)}{\theta+t+1} \right) B_{t,j}^* \frac{\zeta_{t+1}(1-w+\beta)}{\ln(\theta+t+1)} \frac{\ln(\theta+t)}{\zeta_t(1-w+\beta)} \\
&\quad + \frac{\iota \bar{P}_t^*}{(\theta+t+1)\ln(\theta+t+1)} \frac{\zeta_{t+1}(1-w+\beta)}{\zeta_t(1-w+\beta)} + \frac{\Delta M_{t+1,j}^*}{(\theta+t+1)\ln(\theta+t+1)} \\
&= \left( 1 - \frac{(1-(1-\iota)w)}{\theta+t+1} \right) B_{t,j}^* \left( 1 + \frac{1-w+\beta}{\theta+t+1} - \frac{1}{(\theta+t+1)\ln(\theta+t+1)} + O\left(\frac{1}{t^2 \ln(t)}\right) \right) \\
&\quad + \frac{\iota \bar{P}_t^*}{(\theta+t+1)\ln(\theta+t+1)} \left( 1 + \frac{1-w+\beta}{\theta+t+1} \right) + \frac{\Delta M_{t+1,j}^*}{(\theta+t+1)\ln(\theta+t+1)} \\
&= B_{t,j}^* - \frac{B_{t,j}^* - \iota \bar{P}_t^*}{(\theta+t+1)\ln(\theta+t+1)} + \frac{\Delta M_{t+1,j}^*}{(\theta+t+1)\ln(\theta+t+1)} + \rho_t,
\end{aligned}$$

with  $\rho_t = O\left(\frac{B_{t,j}^* + \bar{P}_t^*}{t^{1+w-\beta} \ln(t)}\right) = O\left(\frac{B_{t,j} + \bar{P}_t}{t^{1+w-\beta}}\right)$  being  $\mathcal{F}_t$ -measurable. We want to apply Theorem B.1 with  $X_t = B_t^*$ , together with Lemma B.2. To this end, analogously to the proof of Theorem 5.2, we have that

$$E[(\Delta M_{t+1,j}^*)^2 \mid \mathcal{F}_t] \leq \zeta(1-w+\beta) [(1-\iota)wB_{t,j}^* + \iota \bar{P}_t^*] = O(t^{1-w+\beta+\epsilon}),$$

for any  $\epsilon > 0$ , where we have used  $B_t^* = o(t^\epsilon)$  that follows by Lemma B.2 applied to  $X_t = B_{t,j}^*$ , with  $A_t = \bar{P}_t^*$  and  $b = 0$  (note that we a.s. have  $\sum_t X_t/t^2 = \sum_t O(B_{t,j}/t^{1+w-\beta}) < +\infty$  and  $\sum_t |\rho_t| = \sum_t O\left(\frac{B_{t,j} + \bar{P}_t}{t^{1+w-\beta}}\right) < +\infty$  as  $w > \beta$ ). Hence, choosing a suitable  $\epsilon$ , we get

$$\sum_t \frac{1}{(\theta+t+1)^2 \ln^2(\theta+t+1)} E[(\Delta M_{t+1,j}^*)^2 \mid \mathcal{F}_t] < +\infty \quad \text{a.s.}$$

Then, by Theorem B.1, we have

$$B_{t,j}^* = \frac{\zeta_t(1-w+\beta)}{\ln(\theta+t)} B_{t,j} \xrightarrow{a.s.} \iota P_\infty^*.$$

From Table 1 (case  $w > \beta > 0$ ) we get  $P_\infty^* = \frac{\beta}{\alpha} Z_\infty^* > 0$  a.s. and so the first statement of the theorem follows. Finally, by (6), we get

$$\frac{t^{1-w+\beta}}{\ln(\theta+t)} P_{t,j} \xrightarrow{a.s.} (1-\iota)w\iota P_\infty^* = (1-\iota)w\iota \frac{\beta}{\alpha} Z_\infty^*,$$

which is the second statement of the theorem.  $\square$

**5.2. Second-order asymptotics in the case of low interaction.** We focus here on the low-interaction regime, when each dish-specific process  $K_{t,j}$  converges to a limit random variable, proportional to  $K_{\infty,j}^* > 0$  (see Theorem 5.1). Therefore, it could result useful to establish a CLT for these random variables  $K_{\infty,j}^*$ , in order to construct confidence intervals and estimate them from the observable quantities  $K_{t,j}$  and  $\bar{T}_t$ . This result is presented in Theorem 5.4. Its proof relies on the auxiliary process  $B_{t,j} = K_{t,j}/(\theta+t)$ , introduced in Section 5.1 and on the results proven for the averaged quantity  $\bar{P}_t$ . By (19) from the proof of Theorem 5.1, we can observe that the convergence of  $t^{1-(1-\iota)w} B_{t,j}$  toward  $K_{\infty,j}^*$  is governed by the two competing contributions:

$$(20) \quad t^{1-(1-\iota)w} \frac{\Delta M_{t+1,j}}{\theta+t+1} \quad \text{and} \quad t^{1-(1-\iota)w} \frac{\iota \bar{P}_t}{\theta+t+1},$$

where  $\Delta M_{t+1,j} = (\Delta K_{t+1,j} - P_{t,j})$ . Heuristically, the first term suggests a martingale CLT with scaling  $t^{-(1-\iota)w/2}$ , leading to a zero-mean Gaussian limit kernel. The second term, which is absent in the case of no interaction ( $\iota = 0$ ), depends on the random averaged quantity  $\bar{P}_t$  and so it exhibits a different order of decay according to the values of  $\beta$  and  $w$  (see Table 1 for the first-order asymptotics derived by means of Theorems 4.1 and Theorem 4.2, and see Corollary 4.4 for the second-order asymptotics obtained from Theorem 4.3). In presence of interaction (we here mean low interaction, i.e.  $0 < \iota < \min\{\beta/w, 1\}$ ), the interplay between these two contributions determines the second-order asymptotic behavior of  $t^{1-(1-\iota)w} B_{t,j}$ . Specifically: if  $\iota = 0$  OR  $0 < \iota < \min\{2\beta/w - 1, 1\}$  (note that the case  $\beta \geq w$  and  $0 \leq \iota < 1$  is contained in the union of these two cases), the second term is negligible also at the second order and  $t^{(1-\iota)w/2}(K_{t,j}/t^{(1-\iota)w} - K_{\infty,j}^*)$  converges to a Gaussian kernel with zero mean; if  $0 < \iota = 2\beta/w - 1 < 1$  (that implies  $\beta < w$ ), we still obtain that  $t^{(1-\iota)w/2}(K_{t,j}/t^{(1-\iota)w} - K_{\infty,j}^*)$  converges to a Gaussian kernel, but the second term contributes with a strictly negative random shift in the mean; if  $\max\{0, (2\beta/w - 1)\} < \iota < \beta/w < 1$ , the second term dominates so that  $(K_{t,j}/t^{(1-\iota)w} - K_{\infty,j}^*) \stackrel{P}{\sim} -C Z_\infty^* t^{-(\beta-w\iota)}$ , with a suitable constant  $C > 0$ , and, if we also have  $\iota < 3\beta/w - 1$ , then the Gaussian fluctuations arise for  $(K_{t,j}/t^{(1-\iota)w} - K_{\infty,j}^* + C Z_\infty^* t^{-(\beta-w\iota)})$  multiplied by the scaling factor  $t^{(1-\iota)w/2}$ . Analogously to the discussion in Section 4.2, also here the asymptotic contribution of the second term is strictly negative (when it is not negligible), because, in the dynamics (19), that term represents a strictly positive contribution. This affects the second-order asymptotics by increasing the value of  $K_{\infty,j}^*$  and, consequently, by reducing the difference  $(K_{t,j}/t^{w(1-\iota)} - K_{\infty,j}^*)$ . Moreover, the variance of the Gaussian limit kernel is always random and equal to the dish-specific limit  $K_{\infty,j}^*$ . Finally, it is interesting to discuss how the convergence rate in the CLT changes according to the level  $\iota$  of interaction. Since the second term in (20) is proportional to  $\iota \bar{P}_t$ , it is natural to expect that the larger the parameter  $\iota$  is, the slower the convergence of the CLT will be. Indeed, we can observe from Theorem 5.4 that the decay rate of the difference  $(K_{t,j}/t^{w(1-\iota)} - K_{\infty,j}^*)$  is of the form  $1/t^{f(\iota)}$ , where  $f(\iota)$  is a decreasing function in the parameter  $\iota$ . In particular, we will see that  $f(\iota)$  tends to zero as  $\iota$  approaches  $\min\{\beta/w, 1\}$  (that is its maximum possible value in the case of low-interaction), while, for  $0 < \iota \downarrow 0$ , we have that  $f(\iota)$  converges to  $\min\{\beta, w/2\} > 0$  and, finally,  $f(0) = w/2$ . This means that a discontinuity in the convergence rate of the CLT arises when  $\beta < w/2$ .

**Theorem 5.4** (CLT for the case of low interaction). *Assume*

- (1)  $\iota = 0$  (no interaction at all) or
- (2)  $0 < \iota < \min\{\beta/w, 1\}$ .

Let  $K_{\infty,j}^*$  be the a.s. limit of  $K_{t,j}/t^{w(1-\iota)}$  introduced in Theorem 5.1. Then:

- (i) when  $\iota = 0$  OR  $0 < \iota < \min\{2\beta/w - 1, 1\}$  (that implies  $(1 - \iota)w/2 < \beta - w\iota$ ) OR  $0 < \iota = 2\beta/w - 1 < 1$  (that implies  $(1 - \iota)w/2 = \beta - w\iota$ ), for each observed dish  $j$ , we have

$$(21) \quad t^{(1-\iota)w/2} \left( t^{-(1-\iota)w} K_{t,j} + I_{\{0 < \iota = (2\frac{\beta}{w} - 1) < 1\}} \frac{\iota(\beta/\alpha)}{(\beta - w\iota)} Z_{\infty}^* t^{-(\beta - w\iota)} - K_{\infty,j}^* \right) \xrightarrow{\text{stably}} \mathcal{N}(0, K_{\infty,j}^*),$$

and the convergence also holds true in the sense of the a.s. conditional convergence with respect to  $\mathcal{F}$ ,

- (ii) when  $\max\{0, (2\beta/w - 1)\} < \iota < \beta/w < 1$  (that implies  $(1 - \iota)w/2 > \beta - w\iota$ ), for each observed dish  $j$ , we have

$$t^{\beta - w\iota} (t^{-w(1-\iota)} K_{t,j} - K_{\infty,j}^*) \xrightarrow{P} -\frac{\iota\beta/\alpha}{\beta - w\iota} Z_{\infty}^*.$$

Moreover, when we also have  $(w - \beta) - (1 - \iota)w/2 < \beta/2$  (i.e.  $\iota < 3\beta/w - 1$ ), for each observed dish  $j$ , we have

$$t^{(1-\iota)w/2} \left( t^{-(1-\iota)w} K_{t,j} + \frac{\iota(\beta/\alpha)}{(\beta - w\iota)} Z_{\infty}^* t^{-(\beta - w\iota)} - K_{\infty,j}^* \right) \xrightarrow{\text{stably}} \mathcal{N}(0, K_{\infty,j}^*).$$

Note that, in both cases (i) and (ii), the random quantity  $Z_{\infty}^*$  can be seen as the a.s. limit of  $wt^{1-w}\bar{T}_t$  (for case (ii), recall (13)).

*Proof of Theorem 5.4.* We will first prove the CLT for the process  $B_{t,j} = K_{t,j}/(\theta + t)$  and then we will derive from it the one for  $K_{t,j}$ . Indeed, since the dynamics of  $B_{t,j}$  in (17), i.e.

$$B_{t+1,j} = \left( 1 - \frac{(1 - (1 - \iota)w)}{\theta + t + 1} \right) B_{t,j} + \frac{(\Delta K_{t+1,j} - P_{t,j})}{\theta + t + 1} + \frac{\iota \bar{P}_t}{\theta + t + 1}$$

is equal to the general dynamics (39) in Section B.3 of the supplementary, with  $\delta = w(1 - \iota) \in (0, 1]$ ,  $\Delta M_{t+1} = (\Delta K_{t+1,j} - P_{t,j})$ ,  $\rho_{t+1} = \iota \bar{P}_t$ ,

$$\rho = \begin{cases} w - \beta, & \text{if } w > \beta \text{ AND } 0 < \iota < 1; \\ w(1 - \iota)\varepsilon_0 & \text{if } w \leq \beta \text{ OR } \iota = 0; \end{cases}$$

for an arbitrary  $\varepsilon_0 \in (0, 1)$ , and

$$\rho_{\infty} = \begin{cases} \iota(\beta/\alpha)Z_{\infty}^*, & \text{if } w > \beta \text{ AND } 0 < \iota < 1; \\ 0 & \text{if } w \leq \beta \text{ OR } \iota = 0; \end{cases}$$

(because, by Theorem 4.1 and Theorem 4.2,  $t^{1-\rho} \iota \bar{P}_t$  converges a.s. to  $\rho_{\infty}$ ), and since we are assuming  $w\iota < \beta$ , in all cases we have  $0 < \rho < \delta$  and so the CLT for  $B_{t,j}$  can be obtained by applying Theorem B.4 with  $X_t = B_{t,j}$  and  $X_{\infty}^* = K_{\infty,j}^*$ . Indeed, by Theorem 5.1, we have  $t^{1-(1-\iota)w} B_{t,j} \xrightarrow{\text{a.s.}} t^{-(1-\iota)w} K_{t,j} \xrightarrow{\text{a.s.}} K_{\infty,j}^*$  and, in addition, condition (40) of Theorem B.4 follows by (64) in Lemma D.3, because for any  $p \geq 1$  and  $\epsilon > 0$ , we have

$$E \left[ \sum_{t=1}^{\infty} \frac{|\Delta M_{t+1}|^p}{t^{(1-\iota)w+\epsilon}} \right] = \sum_{t=1}^{\infty} \frac{1}{t^{1+\epsilon}} E[t^{1-(1-\iota)w} |\Delta K_{t+1}|^p] < +\infty.$$

Moreover, condition (41) of Theorem B.4 follows by Theorem 5.1, because

$$t^{1-(1-\iota)w} E[(\Delta M_{t+1})^2 | \mathcal{F}_t] = t^{1-(1-\iota)w} P_{t,j} (1 - P_{t,j}) \xrightarrow{\text{a.s.}} V_{\infty}^* = w(1 - \iota) K_{\infty,j}^*.$$

Then, from Theorem B.4, choosing  $\varepsilon_0 < 1/2$  (so that we always have  $\rho < \delta/2$  when  $w \leq \beta$  OR  $\iota = 0$ ), we obtain that:

- (i) For  $\rho \leq \delta/2 = (1 - \iota)w/2$ , i.e. for  $w \leq \beta$  OR  $\iota = 0$  OR  $w > \beta$  and  $0 < \iota < 2\beta/w - 1$  (that implies  $\iota < \beta/w$ ) OR  $0 < \iota = 2\beta/w - 1 < 1$  (that implies  $\iota < \beta/w < 1$ ), we have

$$t^{(1-\iota)w/2} \left( t^{1-(1-\iota)w} B_{t,j} + \frac{\rho_{\infty}}{(1-\iota)w-\rho} t^{-(1-\iota)w-\rho} - K_{\infty,j}^* \right) \xrightarrow{\text{stably}} \mathcal{N}(0, K_{\infty,j}^*),$$

where  $\rho$  and  $\rho_{\infty}$  take different values according to the different cases and so we obtain (21). Moreover, the convergence also holds true in the sense of the a.s. conditional convergence with respect to  $\mathcal{F}$ . (Note that the above four cases can be grouped as in point (i) of the statement: indeed, the case  $w \leq \beta$  is covered by the union of case  $\iota = 0$  and case  $0 < \iota < \min\{2\beta/w - 1, 1\}$ , that implies  $\iota < \min\{\beta/w, 1\}$ .)



- (ii) For  $(1 - \iota)w/2 = \delta/2 < \rho < \delta$ , i.e. for  $\max\{0, (2\beta/w - 1)\} < \iota < \beta/w < 1$  (that obviously implies  $\beta < w$  and it is possible only if  $\beta > 0$ ), we have

$$t^{\beta-w\iota}(t^{1-w(1-\iota)}B_{t,j} - K_{\infty,j}^*) \xrightarrow{P} -\frac{\iota\beta/\alpha}{\beta-w\iota}Z_{\infty}^*.$$

Moreover, since  $\rho_{t+1} = \iota\bar{P}_t$ , by the limit relation (14) in Corollary 4.4, we have

$$\begin{aligned} t^{\rho-\delta/2}(t^{1-\rho}\rho_{t+1} - \rho_{\infty}) &= \iota \frac{t^{(w-\beta)-(1-\iota)w/2}}{t^{\beta/2}} t^{\beta/2}(t^{1-(w-\beta)}\bar{P}_t + \frac{\beta w}{(w-\beta)}t^{-(w-\beta)} - \frac{\beta}{\alpha}Z_{\infty}^*) \\ &\quad - \frac{\beta w}{(w-\beta)}t^{-(1-\iota)w/2} \xrightarrow{P} 0, \end{aligned}$$

provided  $\rho - \delta/2 = (w - \beta) - (1 - \iota)w/2 < \beta/2$ , and so, by Theorem B.4, in this case we have

$$t^{(1-\iota)w/2} \left( t^{1-(1-\iota)w} B_{t,j} + \frac{\iota(\beta/\alpha)}{(\beta-w\iota)} Z_{\infty}^* t^{-(\beta-w\iota)} - K_{\infty,j}^* \right) \xrightarrow{\text{stably}} \mathcal{N}(0, K_{\infty,j}^*).$$

Finally, the thesis follows because  $B_{t,j} = \frac{K_{t,j}}{(\theta+t)} = \frac{K_{t,j}}{t}(1 + O(\frac{1}{t}))$ .  $\square$

## 6. FINAL REMARKS AND POSSIBLE EXTENSIONS FOR FUTURE RESEARCH

We have proposed an Indian-buffet-type model for multi-factorial innovation, in which the inclusion probability of each feature  $P_{t,j}$  is subjected to a forcing input toward zero, controlled by the parameter  $w$ , and a mean-field interaction among the observed features, regulated by the parameter  $\iota$ . The mean-field term represents the global level of activity of the observed feature system: an old feature may be selected not only because it is itself popular, but also because the overall feature environment is highly active. Our analysis has focused on rigorously establishing the asymptotic behavior of the model, with all results depending on these two novel parameters,  $w$  and  $\iota$ , alongside the standard innovation parameters  $\alpha$ ,  $\beta$  and  $\theta$ . Given the length and the theoretical richness of the paper, we have not addressed the statistical challenges associated with estimating these parameters. Nevertheless, the bases for constructing effective estimators is already provided by the theoretical results presented here: indeed, from the observation of the number  $D_t$  of the distinct appeared features, we can estimate the parameters  $\alpha$  and  $\beta$ ; while the power-law behavior proven for observable processes, such as the average number  $\bar{T}_t$  of features per agent and feature popularity  $K_{t,j}$ , suggests that the new parameters,  $w$  and  $\iota$ , can be estimated as slopes or intercepts via linear regressions on a log-log scale.

The numerous phase transitions and the proliferation of different regimes (see Table 1, 2 and 3) arising from comparisons between parameters might suggest that the model is over-parameterized. Nevertheless, we do not believe this to be the case. Indeed, the three parameters  $\alpha$ ,  $\beta$ , and  $\theta$  were already present in the standard IBP [46], while the new parameters  $\iota$  and  $w$  serve distinct and specific roles, as explained above. In fact, these two parameters enable the model to capture a wide range of behaviors and reproduce diverse asymptotic regimes, which would not be possible using only the standard parameters. The phenomenon under study is inherently complex, and it is therefore natural that several parameters are required to adequately characterize it. We also emphasize that, in order to establish all these different regimes, we had to employ a range of distinct proof techniques, as the dynamics of the stochastic processes of interest behave differently across the various phases. In particular, some of these proofs required the development of general theorems that may be of independent interest beyond the context considered in this paper, and non immediate technical lemmas; specifically, we refer to the results provided in Section B (for general results) and in Section D (for technical lemmas) of the supplementary.

Finally, we highlight that, although the interaction structure considered in this article is a simple mean-field average over all features, it is already sufficient to generate a rich variety of asymptotic behaviors. This interaction can be described using just two additional parameters,  $w$  and  $\iota$ , which can be analytically related to the convergence rates and to the almost-sure limits of certain observable processes, such as  $(\bar{T}_t)_t$  and  $(K_{t,j})_t$ . Consequently, as already observed above, the explicit dependence of the asymptotic rates on the parameters suggests possible routes toward statistical estimation. While one could consider more general interaction matrices, including sparse or heterogeneous interactions, this would substantially increase the number of parameters and make their estimation considerably more challenging. Therefore, the mean-field formulation represents a natural balance between complexity and tractability, enabling the model to capture diverse behaviors without compromising statistical feasibility.

In addition, based on our experience with general interacting reinforced models on finite networks, the asymptotic behavior in this context is typically governed by the eigenstructure of the interaction matrix. In particular, first-order convergence is determined by the leading eigenvalue, with the a.s. limit identified by its corresponding

eigenvector, while second-order asymptotics are influenced by the eigenvalue with the second largest real part. However, in the context of this work, the number of vertices of the network is not fixed, as in the previous literature, since the number of observed features  $D_t$  diverges to infinity. As a consequence, the increasing dimension of the interaction matrix implies that its eigenstructure may evolve over time. The mean-field interaction has the desirable property that the eigenvalues remain constant over time, and are given by  $w$  and  $w(1 - \iota)$ . Analogously, the leading eigenvector is always proportional to the uniform vector with all the entries equal to one. This stability property does not hold for a generic type of interaction.

An interesting direction for future research would be to identify an interaction matrix that grows in dimension while maintaining a fixed leading eigenvalue/eigenvector and a fixed relative second eigenvalue. This is not straightforward and would require a separate dedicated work. Nevertheless, we recognize that this is an important future development that could further enhance the significance of this topic. We believe, however, that the present work, based on the mean-field interaction, provides a first mathematically tractable step in the literature on stochastic multi-factorial innovation models with feature-interaction.

Finally, a newsworthy direction for future work is to develop inferential procedures for the proposed interacting dynamics. In particular, it would be interesting to compare or combine the present recursive approach with Bayesian nonparametric methods for unseen-feature prediction, such as scaled-process priors [10], which provide tractable posterior distributions for the number of unseen genetic variants. Such a comparison would clarify the respective roles of exchangeable feature-allocation priors and non-exchangeable interacting innovation dynamics.

**Authors' contributions.** All the authors contributed equally to the present work.

Sections, theorems and remarks reported in this supplementary document are referenced with capital letters, while those presented in the main document are referenced with numbers.

#### APPENDIX A. TECHNICAL LEMMAS ON DETERMINISTIC SEQUENCES

For  $x \geq 0$ , set

$$(22) \quad \zeta_t(x) = \frac{\Gamma(x+\theta+t+1)}{\Gamma(\theta+t+1)} \quad \text{for } t \geq 0,$$

so that we have (see [24, Lemma 4.1])

$$(23) \quad \zeta_t(x) = (\theta + t + 1)^x + O((\theta + t + 1)^{x-1}) \sim_{t \rightarrow +\infty} t^x$$

and

$$(24) \quad \frac{1}{\zeta_t(x)} = \frac{1}{(\theta+t+1)^x} + O\left(\frac{1}{(\theta+t+1)^{x+1}}\right).$$

Moreover, we have

$$(25) \quad \frac{\zeta_{t+1}(x)}{\zeta_t(x)} = \frac{\Gamma(x+\theta+t+2)}{\Gamma(x+\theta+t+1)} \frac{\Gamma(\theta+t+1)}{\Gamma(\theta+t+2)} = 1 + \frac{x}{\theta+t+1}$$

and since for  $y \geq 0$  we have  $1 - y \leq (1 + y)^{-1} \leq 1 - y + y^2$ ,

$$(26) \quad 1 - \frac{x}{\theta+t+1} \leq \frac{\zeta_t(x)}{\zeta_{t+1}(x)} \leq 1 - \frac{x}{\theta+t+1} + \left(\frac{x}{\theta+t+1}\right)^2.$$

Finally, by (25), we have for  $t \geq 1$

$$(27) \quad \begin{aligned} \frac{\frac{\zeta_{t+1}(x)}{\ln(\theta+t+1)}}{\frac{\zeta_t(x)}{\ln(\theta+t)}} &= \left(1 + \frac{x}{\theta+t+1}\right) \left(1 + \frac{\ln(1 - \frac{1}{\theta+t+1})}{\ln(\theta+t+1)}\right) \\ &= \left(1 + \frac{x}{\theta+t+1}\right) \left(1 - \frac{1}{(\theta+t+1)\ln(\theta+t+1)} + O\left(\frac{1}{t^2 \ln(t)}\right)\right) \\ &= 1 + \frac{x}{\theta+t+1} - \frac{1}{(\theta+t+1)\ln(\theta+t+1)} + O\left(\frac{1}{t^2 \ln(t)}\right). \end{aligned}$$

**Lemma A.1.** Let  $\zeta_t(x) \sim t^x$  defined as in (22),  $0 < \delta \leq 1$ ,  $0 \leq \rho \leq 1$ , and set

$$(28) \quad \delta_0 = \max\{\delta, \rho\} \in (0, 1], \quad b_k = \frac{\theta+k}{r_k} \quad \text{and} \quad a_k = \frac{(\theta+k+1)^2}{b_k^2 r_k},$$

where  $r_0 = \zeta_0(1 - \delta_0)$  and for  $k \geq 1$

$$(29) \quad r_k = \begin{cases} \zeta_k(1 - \delta_0) & \text{if } \delta \neq \rho; \\ \zeta_k(1 - \delta_0)/\ln(\theta + k) & \text{if } \delta = \rho. \end{cases}$$

Then for  $n \rightarrow +\infty$

$$b_n \sum_{k>n} \frac{r_k}{(\theta+k+1)^2} = b_n \sum_{k>n} \frac{1}{a_k b_k^2} \rightarrow \frac{1}{\delta_0}, \quad \text{and} \quad \frac{1}{b_n} \sum_{k \leq n} \frac{1}{r_k} \rightarrow \frac{1}{\delta_0}.$$

*Proof of Lemma A.1.* The proof is based on the continuous integral approximation of the series with proper integrals, namely

$$(30) \quad \left| b_n \sum_{k>n} \frac{1}{a_k b_k^2} - \frac{1}{\delta_0} \right| \leq \left| b_n \left( \sum_{k>n} \frac{1}{a_k b_k^2} - \int_n^\infty \bar{c}(t) dt \right) \right| + b_n \int_n^\infty |\bar{c}(t) - c(t)| dt + \left| b_n \int_n^\infty c(t) dt - \frac{1}{\delta_0} \right|,$$

where  $\bar{c}(t) = r_t/(\theta+t+1)^2$  is the convex function that interpolates  $c_k = (b_k^2 a_k)^{-1} = r_k/(\theta+k+1)^2$  and  $c(t) = (t^{1+\delta_0})^{-1}$  if  $\delta \neq \rho$ ;  $c(t) = (t^{1+\delta_0} \ln t)^{-1}$  if  $\delta = \rho = \delta_0$ .

For the first term in (30), the trapezoidal rule overestimates  $\int_n^{n+M} \bar{c}(t) dt < \frac{c_n}{2} + \sum_{n+1}^{N-1} c_k + \frac{c_M}{2}$  and hence

$$0 < b_n \left( \int_n^\infty \bar{c}(t) dt - \sum_{k>n} \frac{1}{b_k^2 a_k} \right) \leq b_n \frac{c_n}{2} = \frac{\theta+n}{2(\theta+n+1)^2} = O(n^{-1}).$$

For the second term in (30), by (23),  $|\bar{c}(t) - c(t)| \leq C t^{-2-w_0}$ , and hence

$$b_n \int_n^\infty |\bar{c}(t) - c(t)| dt = O(n^{-1}).$$

The second term in (30) is immediate when  $\delta \neq \rho$ , since  $\int_n^\infty c(t) dt = n^{-\delta_0}/\delta_0$ . When  $\delta = \rho = \delta_0$ ,

$$b_n \int_n^\infty c(t) dt = n^{\delta_0} \ln(n) \int_{\delta_0 \ln n}^\infty \frac{e^{-y}}{y} dy \rightarrow \frac{1}{\delta_0},$$

since, for any  $x > 0$  [1, eq.5.1.20, pag.229]

$$\frac{1}{2} \ln \left( 1 + \frac{2}{x} \right) < e^x \int_x^\infty \frac{e^{-y}}{y} dy < \ln \left( 1 + \frac{1}{x} \right).$$

The proof of  $\frac{1}{b_n} \sum_{k \leq n} \frac{1}{r_k} \rightarrow \frac{1}{\delta_0}$  is trivial for  $\delta_0 = 1$ , since  $r_k \equiv 1$  when  $\delta \neq \rho$  and  $\sum_{k \leq n} r_k^{-1} \sim n \ln(n)$  for  $\delta = \rho = 1$ . For  $\delta_0 < 1$ , the proof is again made on continuous integral approximations, here

$$(31) \quad \left| \frac{1}{b_n} \sum_{k \leq n} \frac{1}{r_k} - \frac{1}{\delta_0} \right| \leq \left| \frac{1}{b_n} \left( \sum_{k \leq n} \frac{1}{r_k} - \int_0^n \frac{1}{r_t} dt \right) \right| + \frac{1}{b_n} \int_0^n \left| \frac{1}{r_t} - c(t) \right| dt + \left| \frac{1}{b_n} \int_0^n c(t) dt - \frac{1}{\delta_0} \right|,$$

where  $c(t) = (t^{-1+w_0})$  if  $w \neq \beta$ ;  $c(t) = (t^{-1+\delta_0} \ln t)$  if  $\delta = \rho = \delta_0$ . Since  $\frac{1}{r_t}$  is eventually decreasing by (25) and (27), the first term in (31) is vanishing, since

$$0 < \frac{1}{b_n} \left( \sum_{k \leq n} \frac{1}{r_k} - \int_0^n \frac{1}{r_t} dt \right) \leq \frac{1}{b_n} \frac{2}{r_{\min}} \rightarrow 0.$$

Moreover, by (24),  $|\frac{1}{r_t} - c(t)| \leq C t^{-2+\delta_0}$ , and hence also the second term in (31) is vanishing:

$$\frac{1}{b_n} \int_0^n \left| \frac{1}{r_t} - c(t) \right| dt \rightarrow 0.$$

When  $\delta \neq \rho$ ,  $\int_0^n c(t) dt = n^{\delta_0}/\delta_0$  and the proof is done. When  $\delta = \rho = \delta_0$ , with  $e^y = t^{\delta_0}$ , we obtain

$$\int_0^n c(t) dt = \int_0^n \frac{t^{\delta_0} \ln t}{t} dt = \int_{-\infty}^{\delta_0 \ln n} \frac{e^y y}{\delta_0^2} dy = \frac{n^{\delta_0} (\delta_0 \ln(n) - 1)}{\delta_0^2}. \quad \square$$

**Lemma A.2** ([4, Supplementary material]). *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  $\lim_{t \rightarrow +\infty} y_t = 0$ .*

## APPENDIX B. GENERAL RESULTS FOR SOME RECURSIVE DYNAMICS

We start by providing in Subsection B.1 two results, probably known, but for which we were not able to find a reference in the literature and so, for the reader's convenience, we state and prove them here. In Subsection B.2 and Subsection B.3 we provide two new general results.

### B.1. Results of the first-order for a generalized stochastic approximation dynamics.

**Theorem B.1.** Let  $(X_t)_t$  and  $(A_t)_t$  be two  $\mathcal{F}$ -adapted real stochastic processes, where  $(X_t)_t$  follows the dynamics

$$X_{t+1} = X_t + \eta_t(A_t - b_t X_t) + \eta_t \Delta M_{t+1} + \rho_{t+1},$$

where  $(b_t)_t$  and  $(\eta_t)_t$  are two sequences of real numbers with  $b_t \rightarrow b > 0$ ,  $\eta_t \geq 0$  for  $t$  large enough,  $\sum_t \eta_t = +\infty$ ,  $\sum_t \eta_t^2 < +\infty$ ,  $(\Delta M_{t+1})_t$  is an  $\mathcal{F}$ -martingale difference such that

$$(32) \quad \sum_t \eta_t^2 E[(\Delta M_{t+1})^2 | \mathcal{F}_t] < +\infty \quad a.s.$$

and  $\rho_{t+1}$  is a  $\mathcal{F}_{t+1}$ -measurable remainder term such that  $\sum_t |\rho_{t+1}| < +\infty$  a.s. Then, we have

$$\frac{1}{b} \liminf_{t \rightarrow \infty} A_t \leq \liminf_{t \rightarrow \infty} X_t \leq \limsup_{t \rightarrow \infty} X_t \leq \frac{1}{b} \limsup_{t \rightarrow \infty} A_t \quad a.s..$$

As a consequence, if  $(A_t)_t$  converges a.s. to a random variable  $A_\infty$ , then we have  $X_t \xrightarrow{a.s.} A_\infty/b$ .

*Proof of Theorem B.1.* Without loss of generality, since  $b_t \rightarrow b > 0$ , we can assume  $b_t > 0$  for all  $t$ . Moreover, we can focus on proving the thesis with  $b_t$  replaced by its limit  $b > 0$ . Indeed, since  $(b_t/b) \rightarrow 1$ , the thesis with a time-dependent sequence  $(b_t)_t$  follows directly with the one with constant  $b$  by setting  $\tilde{\eta}_t = (b_t/b)\eta_t \geq 0$  for  $t$  large enough,  $\tilde{A}_t = (b/b_t)A_t$  and  $\Delta \tilde{M}_{t+1} = (b/b_t)\Delta M_{t+1}$ .

Assume  $b_t = b > 0$  for all  $t$ . Without loss of generality, since  $\eta_t \geq 0$  for  $t$  large enough and  $\eta_t \rightarrow 0$ , we can assume  $\eta_t$  such that  $\eta_t \geq 0$  and  $(1 - b\eta_t) > 0$  for all  $t$ . The rest of the proof is divided in two steps.

*Step (1):* For any  $a \in \mathbb{R}$  and  $t_0 \in \mathbb{N}$  we can define the process  $(X_t^{a,t_0})_{t \geq 0}$  as follows:  $X_0^{a,t_0} = X_0$  and  $X_t^{a,t_0} = \sum_{n=0}^{t-1} \Delta X_{n+1}^{a,t_0}$  for  $t \geq 1$ , with

$$\Delta X_{n+1}^{a,t_0} = \begin{cases} \eta_n(A_n - bX_n^{a,t_0}) + \eta_n \Delta M_{n+1} + \rho_{n+1} & \text{if } n \leq t_0; \\ \eta_n(a - bX_n^{a,t_0}) + \eta_n \Delta M_{n+1} + \rho_{n+1} & \text{if } n > t_0; \end{cases}$$

so that  $X_t^{a,t_0} \equiv X_t$  for any  $t \leq t_0$ . Moreover notice that, for any  $n > t_0$ ,

$$\Delta X_{n+1}^{a,t_0} - \Delta X_{n+1} = \eta_n((a - A_n) - b(X_n^{a,t_0} - X_n))$$

and then

$$(X_{n+1}^{a,t_0} - X_{n+1}) = \eta_n(a - A_n) + (1 - b\eta_n)(X_n^{a,t_0} - X_n).$$

This means that

- (i)  $\cap_{t \geq t_0} \{X_t^{a,t_0} \leq X_t\}$  with probability one on  $\cap_{t \geq t_0} \{a < A_t\}$ ,
- (ii)  $\cap_{t \geq t_0} \{X_t^{a,t_0} \geq X_t\}$  with probability one on  $\cap_{t \geq t_0} \{a > A_t\}$ .

Therefore, on the set  $\{a > \limsup_{t \rightarrow \infty} A_t\}$ , with probability one we have that

$$\limsup_{t \rightarrow \infty} X_t = \inf_{t_0} \sup_{t \geq t_0} X_t \leq \inf_{t_0} \sup_{t \geq t_0} X_t^{a,t_0} = \limsup_{t \rightarrow \infty} X_t^{a,t_0}.$$

Analogously, on the set  $\{a < \liminf_{t \rightarrow \infty} A_t\}$ , with probability one we have that

$$\liminf_{t \rightarrow \infty} X_t = \sup_{t_0} \inf_{t \geq t_0} X_t \geq \sup_{t_0} \inf_{t \geq t_0} X_t^{a,t_0} = \liminf_{t \rightarrow \infty} X_t^{a,t_0}.$$

*Step (2):* We will now show that, for any fixed  $a \in \mathbb{R}$  and  $t_0 \in \mathbb{N}$ ,

$$\limsup_{t \rightarrow \infty} X_t^{a,t_0} = \liminf_{t \rightarrow \infty} X_t^{a,t_0} = \frac{a}{b} \quad a.s..$$

We define  $\tilde{\rho}_{t_0} = 0$  and  $\tilde{\rho}_{t+1} = (1 - b\eta_t)\tilde{\rho}_t + \rho_{t+1}$  for each  $t \geq t_0$ , so that we have  $|\tilde{\rho}_{t+1}| \leq (1 - b\eta_t)|\tilde{\rho}_t| + |\rho_{t+1}|$  and hence, by Lemma A.2, we get  $\tilde{\rho}_t \xrightarrow{a.s.} 0$ . Set  $Y_t = X_t^{a,t_0} - a/b - \tilde{\rho}_t$  for any  $t \geq t_0$ . After some computations, we arrive to

$$Y_{t+1} = (1 - b\eta_t)Y_t + \eta_t \Delta M_{t+1}$$

and so

$$E[Y_{t+1}^2 | \mathcal{F}_t] = (1 + b^2 \eta_t^2)Y_t^2 + \eta_t^2 E[(\Delta M_{t+1})^2 | \mathcal{F}_t] - 2b\eta_t Y_t^2.$$

Hence,  $(Y_t^2)_t$  is a non-negative almost super-martingale, which converges a.s. since  $\sum_t \eta_t^2 < +\infty$  and (32) (see Theorem E.4). We also get  $\sum_t \eta_t Y_t^2 < +\infty$  a.s. and so, since  $\sum_t \eta_t = +\infty$ , we necessarily have  $Y_t^2 \xrightarrow{a.s.} 0$ , that is

$X_t^{a,t_0} \xrightarrow{a.s.} a/b$ .

*Conclusion:* We have shown that, for any pair  $a_1 < a_2$ , on the set

$$\{a_1 < \liminf_{t \rightarrow \infty} A_t < \limsup_{t \rightarrow \infty} A_t < a_2\},$$

we have that

$$\frac{a_1}{b} \leq \liminf_{t \rightarrow \infty} X_t < \limsup_{n \rightarrow \infty} X_t \leq \frac{a_2}{b} \quad \text{a.s.}$$

Finally, since  $a_1 < a_2$  are arbitrary, we get the thesis.  $\square$

**Lemma B.2.** *Let  $(X_t)$  be a non-negative stochastic process such that  $\sum_t X_t/t^2 < +\infty$  and with dynamics*

$$X_{t+1} = X_t + \frac{1}{\theta+t+1}(A_t - b_t X_t) + \frac{\Delta M_{t+1}}{\theta+t+1} + \rho_t,$$

where  $b_t \geq 0$ ,  $(A_t)_t$  is a.s. convergent,  $(\Delta M_{t+1})_t$  is an  $\mathcal{F}$ -martingale difference and  $(\rho_t)$  is an  $\mathcal{F}$ -adapted process with  $\sum_t |\rho_t| < +\infty$  a.s. Then we have  $X_t = o(t^\epsilon)$  for any  $\epsilon > 0$ .

*Proof.* Set  $X_t^* = X_t/\zeta_t(\epsilon)$ .

$$X_{t+1}^* = \left(1 - \frac{b_t}{\theta+t+1}\right) \frac{\zeta_t(\epsilon)}{\zeta_{t+1}(\epsilon)} X_t^* + \frac{\Delta M_{t+1}}{(\theta+t+1)\zeta_{t+1}(\epsilon)} + \frac{A_t/(\theta+t+1)+\rho_t}{\zeta_{t+1}(\epsilon)}.$$

Since (26), we get

$$\begin{aligned} X_{t+1}^* &= \left(1 - \frac{b_t}{\theta+t+1}\right) \left(1 - \frac{\epsilon}{\theta+t+1} + O\left(\frac{1}{t^2}\right)\right) X_t^* + \frac{\Delta M_{t+1}}{(\theta+t+1)\zeta_{t+1}(\epsilon)} + \frac{\frac{A_t}{(\theta+t+1)} + \rho_t}{\zeta_{t+1}(\epsilon)} \\ &= X_t^* - \frac{b_t + \epsilon}{\theta+t+1} X_t^* + O(1/t^2) X_t^* + \frac{\Delta M_{t+1}}{(\theta+t+1)\zeta_{t+1}(\epsilon)} + \frac{\frac{A_t}{(\theta+t+1)} + \rho_t}{\zeta_{t+1}(\epsilon)} \\ &\leq X_t^* + \frac{\Delta M_{t+1}}{(\theta+t+1)\zeta_{t+1}(\epsilon)} + \frac{\frac{A_t}{(\theta+t+1)} + \rho_t}{\zeta_{t+1}(\epsilon)} + O(1/t^2) X_t^*. \end{aligned}$$

Hence, we have

$$E[X_{t+1}^* | \mathcal{F}_t] \leq X_t^* + \frac{|A_t|/(\theta+t+1)+|\rho_t|}{\zeta_{t+1}(\epsilon)} + O(1/t^2) X_t^*$$

and so  $(X_t^*)$  is a non-negative almost super-martingale (see Theorem E.4) and, since we have

$$\sum_t \left[ \frac{|A_t|/(\theta+t+1)+|\rho_t|}{\zeta_{t+1}(\epsilon)} + O(1/t^2) X_t^* \right] < +\infty \quad \text{a.s.},$$

we can conclude that  $X_t^*$ , and so  $X_t/t^\epsilon$ , converges a.s. to a finite random variable. Since  $\epsilon > 0$  is arbitrary, the limit random variable has to be equal to zero, and so we can conclude that  $X_t = o(t^\epsilon)$ .  $\square$

**B.2. Strict positiveness of the a.s. limit for a particular class of non-negative almost super-martingales.** Let  $X = (X_t)_t$  be a non-negative real stochastic process with discrete time, adapted to a filtration  $\mathcal{F} = (\mathcal{F}_t)_t$  and such that

$$(33) \quad X_{t+1} = \left(1 - \frac{1}{\theta+t+1}\right) X_t + \frac{1}{(\theta+t+1)} Y_{t+1},$$

where  $\theta > 0$  and  $(Y_t)_t$  is a non-negative  $\mathcal{F}$ -adapted real stochastic process. Now, let us set  $H_t = (\theta+t)X_t$ . From (33), we get

$$H_{t+1} = H_t + Y_{t+1}$$

and so  $H_t = H_0 + \sum_{n=1}^t (H_n - H_{n-1}) = H_0 + \sum_{n=1}^t Y_n$ . Hence,  $(H_t)_t$  is a non-decreasing sequence of random variables with  $\sup_t H_t = H_0 + \sum_t Y_t$ . Therefore, if the series  $\sum_t Y_t$  is a.s. convergent, then  $(H_t)$ , and so  $(tX_t)$ , is a.s. convergent. This means that  $X_t = O(1/t)$  and so it a.s. converges to zero. The following result deals with the case when  $\sum_t Y_t = +\infty$ .

(Note we use the notation  $\rho_t^+ = \max\{\rho_t, 0\}$  and  $\rho_t^- = \max\{-\rho_t, 0\}$  so that  $\rho_t = \rho_t^+ - \rho_t^-$  and  $|\rho_t| = \rho_t^+ + \rho_t^-$ .)

**Theorem B.3** (Strict positiveness of the a.s. limit). *In the framework described above, if  $\sum_t Y_t = +\infty$  a.s., then  $(tX_t)_t$  converges a.s. to  $+\infty$ .*

Moreover, if we also assume

$$(34) \quad E[Y_{t+1} | \mathcal{F}_t] = \delta X_t + \rho_t,$$

where  $0 < \delta \leq 1$  and  $(\rho_t)_t$  is an  $\mathcal{F}$ -adapted stochastic process such that  $\rho_t \xrightarrow{a.s.} 0$  with  $(*) \sum_t \rho_t^+ / (\theta+t)^\delta < +\infty$  a.s., then  $\sum_t |\rho_t| / (\theta+t)^\delta < +\infty$  a.s. and  $(t^{(1-\delta)} X_t)_t$  converges a.s. to a finite random variable  $X_\infty^*$ , that is

$X_t = O(1/t^{1-\delta})$ .

Finally,  $P(X_\infty^* > 0) = 1$ , that is  $t^{-(1-\delta)}$  is the exact convergence rate of  $(X_t)_t$  toward zero, if  
a) we replace condition (\*) by the more restrictive one:

$$(35) \quad \sum_t \frac{|\rho_t|}{(\theta+t)^{\delta-\epsilon}} < +\infty \text{ a.s. for some } 0 < \epsilon < \delta,$$

b) we also add the assumption

$$(36) \quad E[Y_{t+1}^2 \mid \mathcal{F}_t] \leq O(E[Y_{t+1} \mid \mathcal{F}_t]),$$

and c) we require that  $\rho_t$  is a.s. in one of the following conditions:

- i)  $\rho_t \geq 0$  eventually;
- ii)  $t|\rho_t| = O(1)$ ;
- iii)  $\rho_t$  is a function of  $X_t$  such that  $|\rho_t| = o(X_t)$ .

**Remark B.1.** When the random variables  $Y_t$  are uniformly bounded and  $\rho_t \geq 0$ , in order to have  $\sum_t Y_t = +\infty$  a.s., it is enough that exists a finite time  $\tau$ , possibly random, such that  $X_\tau > 0$  a.s.. Indeed we a.s. have

$$\sum_t E[Y_{t+1} \mid \mathcal{F}_t] = \delta \sum_t X_t + \sum_t \rho_t \geq \delta \sum_t \frac{H_t}{(\theta+t)} \geq \delta H_\tau \sum_{t \geq \tau} \frac{1}{\theta+t} = +\infty,$$

where we used that  $H_t = (\theta+t)X_t$  is non-decreasing (see above) and  $H_\tau = (\theta+\tau)X_\tau > 0$  a.s.. This is sufficient in order to conclude because of Lemma E.1.

*Proof of Theorem B.3.* Let us set  $H_t = (\theta+t)X_t$  as we have done above so that, from (33), we have

$$H_{t+1} = H_t + Y_{t+1}$$

and so  $H_t = H_0 + \sum_{n=1}^t (H_n - H_{n-1}) = H_0 + \sum_{n=1}^t Y_n$ . Hence,  $(H_t)_t$  is a non-decreasing sequence of random variables with  $\sup_t H_t = H_0 + \sum_t Y_t$ . Therefore, if  $\sum_t Y_t = +\infty$  a.s., then  $(H_t)$ , and so  $(tX_t)$ , diverges a.s. to  $+\infty$  (and so  $X_t > 0$  for  $t$  large enough).

Now, we set  $X_t^* = \zeta_t(1-\delta)X_t$  and  $\Delta M_{t+1} = Y_{t+1} - E[Y_{t+1} \mid \mathcal{F}_t]$  so that we have

$$X_{t+1}^* = \left(1 - \frac{(1-\delta)}{\theta+t+1}\right) \frac{\zeta_{t+1}(1-\delta)}{\zeta_t(1-\delta)} X_t^* + \frac{\zeta_{t+1}(1-\delta)\Delta M_{t+1}}{\theta+t+1} + \frac{\zeta_{t+1}(1-\delta)}{\theta+t+1} \rho_t.$$

Since (25), we get

$$(37) \quad \begin{aligned} X_{t+1}^* &= \left(1 - \frac{(1-\delta)}{\theta+t+1}\right) \left(1 + \frac{(1-\delta)}{\theta+t+1}\right) X_t^* + \frac{\zeta_{t+1}(1-\delta)\Delta M_{t+1}}{\theta+t+1} + \frac{\zeta_{t+1}(1-\delta)}{\theta+t+1} \rho_t \\ &= X_t^* - \frac{(1-\delta)^2}{(\theta+t+1)^2} X_t^* + \frac{\zeta_{t+1}(1-\delta)\Delta M_{t+1}}{\theta+t+1} + \frac{\zeta_{t+1}(1-\delta)}{\theta+t+1} \rho_t \\ &\leq X_t^* + \frac{\zeta_{t+1}(1-\delta)\Delta M_{t+1}}{\theta+t+1} + \frac{\zeta_{t+1}(1-\delta)}{\theta+t+1} \rho_t. \end{aligned}$$

Hence, we have

$$E[X_{t+1}^* \mid \mathcal{F}_t] \leq X_t^* + \frac{\zeta_{t+1}(1-\delta)}{\theta+t+1} \rho_t^+ - \frac{\zeta_{t+1}(1-\delta)}{\theta+t+1} \rho_t^-$$

and so  $(X_t^*)$  is a non-negative almost super-martingale (see Theorem E.4) and, since we have

$$\sum_t \frac{\zeta_{t+1}(1-\delta)}{(\theta+t+1)} \rho_t^+ \leq \sum_t \frac{\rho_t^+}{(\theta+t)^\delta} + \sum_t \frac{O((\theta+t)^{-\delta})}{(\theta+t)} \rho_t^+ < +\infty \text{ a.s.},$$

we can conclude that  $X_t^*$ , and so  $t^{1-\delta}X_t$ , converges a.s. to a finite random variable  $X_\infty^*$  (and so  $X_t \xrightarrow{a.s.} 0$ ) and also  $\sum_t \frac{\zeta_{t+1}(1-\delta)}{(\theta+t)} \rho_t^- < +\infty$  a.s. so that we get  $\sum_t |\rho_t|/(\theta+t)^\delta < +\infty$  a.s..

Finally, we assume conditions (36), (35) and one of the conditions i)-iii) and we are going to prove that  $X_\infty^* > 0$  a.s. To this purpose we proceed in some steps.

*Step 1 (Proof of the fact that, for each  $0 < \eta < \delta \leq 1$ , we have  $1/X_t = o(t^{1-\eta})$  or, equivalently,  $1/H_t = o(t^{-\eta})$ ):*

We set  $Q_t = (\theta+t)/H_t^\gamma$  with  $\gamma > 0$  and we observe that we have

$$\frac{Q_{t+1}}{Q_t} - 1 = \frac{(\theta+t+1)}{(\theta+t)} \frac{H_t^\gamma}{H_{t+1}^\gamma} - 1 = \left(1 + \frac{1}{\theta+t}\right) \frac{H_t^\gamma}{H_{t+1}^\gamma} - 1$$



$$\begin{aligned}
&= \left(\frac{H_t^\gamma}{H_{t+1}^\gamma}\right)^\gamma - 1 + \frac{1}{(\theta+t)} \left(\frac{H_t}{H_{t+1}}\right)^\gamma = \left(1 - \frac{Y_{t+1}}{H_{t+1}}\right)^\gamma - 1 + \frac{1}{(\theta+t)} \left(\frac{H_t}{H_{t+1}}\right)^\gamma \\
&\leq -\gamma \frac{Y_{t+1}}{H_{t+1}} + \frac{1}{(\theta+t)} = -\gamma \frac{Y_{t+1}}{H_t} + \gamma Y_{t+1} \left(\frac{1}{H_t} - \frac{1}{H_{t+1}}\right) + \frac{1}{(\theta+t)} \\
&= -\gamma \frac{Y_{t+1}}{H_t} + \gamma \frac{Y_{t+1}^2}{H_t H_{t+1}} + \frac{1}{(\theta+t)} \leq -\gamma \frac{Y_{t+1}}{H_t} + \gamma \frac{Y_{t+1}^2}{H_t^2} + \frac{1}{(\theta+t)},
\end{aligned}$$

where the first inequality follows since  $(1-x)^\gamma \leq 1 - \gamma x$  for  $0 \leq x \leq 1$  and  $H_t \leq H_{t+1}$  and this last relation implies also the second inequality. If we take the conditional expectation given the past  $\mathcal{F}_t$  and use the equality  $H_t = (\theta+t)X_t$  and  $E[Y_{t+1} | \mathcal{F}_t] = \delta X_t + \rho_t$ , we obtain

$$\begin{aligned}
(38) \quad E\left[\frac{Q_{t+1}}{Q_t} - 1 \mid \mathcal{F}_t\right] &\leq -\gamma \frac{(\delta X_t + \rho_t)}{H_t} + \gamma \frac{E[Y_{t+1}^2 | \mathcal{F}_t]}{H_t^2} + \frac{1}{(\theta+t)} \\
&\leq -\gamma \frac{(\delta X_t + \rho_t)}{H_t} + \gamma \frac{O(E[Y_{t+1} | \mathcal{F}_t])}{H_t^2} + \frac{1}{(\theta+t)} \\
&= \frac{(1-\gamma\delta)}{(\theta+t)} - \frac{\gamma}{(\theta+t)} \frac{(\theta+t)\rho_t}{H_t} + \frac{1}{(\theta+t)} O\left(\frac{1}{H_t}\right) + \frac{1}{(\theta+t)} O\left(\frac{(\theta+t)|\rho_t|}{H_t^2}\right) \\
&= \frac{1}{(\theta+t)} \left[ (1-\gamma\delta) + O\left(\frac{1}{H_t}\right) + \frac{(\theta+t)|\rho_t|}{H_t} (-\gamma \text{sign}(\rho_t) + O\left(\frac{1}{H_t}\right)) \right].
\end{aligned}$$

If we prove that the above last quantity into the square brackets is a.s. eventually non-positive for  $\gamma\delta > 1$ , then we get that for  $\gamma\delta > 1$  we a.s. have

$$E[Q_{t+1} - Q_t \mid \mathcal{F}_t] = Q_t E\left[\frac{Q_{t+1}}{Q_t} - 1 \mid \mathcal{F}_t\right] \leq 0 \quad \text{eventually.}$$

This proves that, for each  $\gamma > 1/\delta$ , the stochastic process  $(Q_t)_t$ , i.e.  $((\theta+t)/H_t^\gamma)_t$  is eventually a (non-negative) super-martingales and so, for each  $\gamma > 1/\delta$ , it converges a.s. to a real random variable. Since  $\gamma > 1/\delta$  is arbitrary, we necessarily have that  $(\theta+t)/H_t^\gamma$  converges a.s. to zero. This fact concludes the proof of Step 1 because it implies that  $(\theta+t)^{1/\gamma}/H_t$  converges a.s. to zero and we can set  $\eta = 1/\gamma < \delta \leq 1$ . Therefore the point is to have additional assumptions that entail the last term in the square brackets of (38) to be non-positive for  $\gamma\delta > 1$ . When  $\gamma\delta > 1$ , we have  $(1-\gamma\delta) < 0$  and so, since  $H_t \rightarrow +\infty$ , the first quantity  $O(1/H_t)$  can be neglected. Hence the crucial term is

$$\frac{(\theta+t)|\rho_t|}{H_t} \left( -\gamma \text{sign}(\rho_t) + O\left(\frac{1}{H_t}\right) \right).$$

In the case i), that is when  $\rho_t \geq 0$  eventually with probability one, we are done because the above quantity is equal to zero when  $\rho_t = 0$  and, when  $\rho_t > 0$ , the quantity  $-\gamma \text{sign}(\rho_t) + O(1/H_t) = -\gamma + O(1/H_t)$  is eventually non-positive with probability one, because  $-\gamma < 0$  and  $O(1/H_t) \rightarrow 0$  a.s.. Another simple cases are case ii), that is when  $\limsup_t(t|\rho_t|) < +\infty$  a.s., and case iii), that is when  $|\rho_t| = o(X_t)$ . Indeed, in these cases we a.s. have

$$\lim_t \frac{(\theta+t)|\rho_t|}{H_t} = \lim_t \frac{|\rho_t|}{X_t} = 0.$$

*Step 2 (Convergence of fundamental series):*

From Step 1 and assumption (35), we a.s. get

$$\sum_t \frac{|\rho_t|}{H_t} = \sum_t \frac{(\theta+t)^\eta}{H_t} \frac{|\rho_t|}{(\theta+t)^\eta} = \sum_t o(1) \frac{|\rho_t|}{(\theta+t)^\eta} < +\infty,$$

where  $0 < \eta = \delta - \epsilon < \delta$ . Moreover, we a.s. have

$$\sum_t \frac{1}{tH_t} = \sum_t \frac{t^\eta}{H_t} \frac{1}{t^{1+\eta}} = \sum_t o(1) \frac{1}{t^{1+\eta}} < +\infty,$$

where  $0 < \eta < \delta$ .

*Step 3 (Conclusion):* Set  $L_t = \ln(H_t/(\theta+t)^\delta)$  and  $U_t = E[L_{t+1} - L_t \mid \mathcal{F}_t]$  and  $V_t = E[(L_{t+1} - L_t)^2 \mid \mathcal{F}_t]$ . If we prove that  $\sum_t U_t$  and  $\sum_t V_t$  are a.s. convergent, then  $L_t$  converges a.s. to a real random variable (see Lemma E.3). This fact implies that  $(\theta+t)^{1-\delta}X_t = (\theta+t)X_t/(\theta+t)^\delta = H_t/(\theta+t)^\delta$  converges to a random variable with values in  $(0, +\infty)$  and so  $P(X_\infty^* > 0) = 1$ . The rest of the proof is devoted to verify that  $\sum_t |U_t| < +\infty$  and  $\sum_t V_t < +\infty$  a.s.. To this regard, we note that

$$\begin{aligned}
U_t &= E[\ln(H_{t+1}) - \ln(H_t) \mid \mathcal{F}_t] - \delta (\ln(\theta+t+1) - \ln(\theta+t)) \\
&= E[\ln(H_t + Y_{t+1}) - \ln(H_t) \mid \mathcal{F}_t] - \delta \ln(1 + 1/(\theta+t)) \\
&= E\left[\int_0^{Y_{t+1}} \frac{1}{H_t+x} dx \mid \mathcal{F}_t\right] - \delta \ln(1 + 1/(\theta+t)).
\end{aligned}$$

Since  $1/(H_t + x) \leq 1/H_t$  and  $\ln(1 + 1/(\theta + t)) \geq 1/(\theta + t) - 1/(2(\theta + t)^2)$  for each  $x \geq 0$  and each  $t$ , the last term of the above equalities is smaller than or equal to

$$\begin{aligned} \frac{E[Y_{t+1}|\mathcal{F}_t]}{H_t} - \frac{\delta}{\theta+t} + \frac{\delta}{2(\theta+t)^2} &= \frac{\delta X_t + \rho_t}{H_t} - \frac{\delta}{\theta+t} + \frac{\delta}{2(\theta+t)^2} \\ &= \frac{\delta}{\theta+t} + \frac{\rho_t}{H_t} - \frac{\delta}{\theta+t} + \frac{\delta}{2(\theta+t)^2} = \frac{\rho_t}{H_t} + O(1/t^2). \end{aligned}$$

Now, we note that  $-U_t = \delta \ln(1 + 1/(\theta + t)) - \ln(H_{t+1}) + \ln(H_t)$ . Using  $\ln(1 + 1/(\theta + t)) \leq 1/(\theta + t)$  and  $1/(H_t + x) \geq 1/H_t - x/H_t^2$  for each  $x \geq 0$  and each  $t$ , we find that  $-U_t$  is smaller than or equal to

$$\begin{aligned} \frac{\delta}{(\theta+t)} - \frac{1}{H_t} E[Y_{t+1} | \mathcal{F}_t] + \frac{1}{2H_t^2} E[Y_{t+1}^2 | \mathcal{F}_t] \\ \leq \frac{\delta}{(\theta+t)} - \frac{\delta X_t + \rho_t}{H_t} + \frac{O(\delta X_t + \rho_t)}{2H_t^2} = \frac{\delta}{\theta+t} - \frac{\delta}{\theta+t} - \frac{\rho_t}{H_t} + \frac{O(\delta X_t + \rho_t)}{2H_t^2} \\ = -\frac{\rho_t}{H_t} + O(1/(tH_t)) + o(\rho_t/H_t). \end{aligned}$$

Thus, by Step 2, we can derive that  $\sum_t |U_t| < +\infty$  a.s.. Finally, we observe we have

$$\begin{aligned} E[(\ln(H_{t+1}) - \ln(H_t) - \delta \ln(\theta + t + 1) + \delta \ln(\theta + t))^2 | \mathcal{F}_t] \\ \leq 2 \{ E[(\ln(H_{t+1}) - \ln(H_t))^2 | \mathcal{F}_t] + \delta^2 (\ln(1 + 1/(\theta + t)))^2 \} \\ \leq 2E \left[ \left( \int_0^{Y_{t+1}} \frac{1}{H_t + x} dx \right)^2 \middle| \mathcal{F}_t \right] + 2\delta^2 / (\theta + t)^2 \\ \leq 2E[(Y_{t+1}/H_t)^2 | \mathcal{F}_t] + O(1/t^2) \leq \frac{2}{H_t^2} E[Y_{t+1}^2 | \mathcal{F}_t] + O(1/t^2) \\ = \frac{1}{H_t^2} O(\delta X_t + \rho_t) + O(1/t^2) \\ = O(1/(tH_t)) + o(\rho_t/H_t) + O(1/t^2). \end{aligned}$$

Therefore we have also  $\sum_t V_t < +\infty$  a.s. and we can conclude.  $\square$

**B.3. Central limit theorem for a recursive dynamics.** Let  $X = (X_t)_t$  be a real stochastic process with discrete time, adapted to a complete filtration  $\mathcal{F} = (\mathcal{F}_t)_t$  with dynamics

$$(39) \quad X_{t+1} = \left(1 - \frac{(1-\delta)}{\theta+t+1}\right) X_t + \frac{1}{(\theta+t+1)} \Delta M_{t+1} + \frac{1}{\theta+t+1} \rho_{t+1},$$

where  $\theta > 0$ ,  $0 < \delta \leq 1$ ,  $\Delta M_{t+1}$  is a martingale difference and  $(\rho_t)_t$  is an  $\mathcal{F}$ -adapted stochastic process such that  $t^{1-\rho} \rho_{t+1} \xrightarrow{a.s.} \rho_\infty$  with  $0 \leq \rho \leq 1$  and  $\rho_\infty$  being a real ( $\mathcal{F}_\infty$ -measurable) random variable.

Note that a non-negative stochastic process  $X = (X_t)_t$  with dynamics (33), where  $Y = (Y_t)_t$  is a sequence of integrable random variable that verify condition (34) with  $t^{1-\rho} \rho_t \xrightarrow{a.s.} \rho_\infty$  for a suitable  $\rho \in [0, 1]$ , satisfy also dynamics (39) with  $\Delta M_{t+1} = Y_{t+1} - E[Y_{t+1}|\mathcal{F}_t]$  (and in this case the reminder term  $\rho_{t+1}$  in (39) coincides with  $\rho_t$  of condition (34) and so it is even  $\mathcal{F}_t$ -measurable).

When  $t^{1-\delta} X_t$  a.s. converges toward a real random variable  $X_\infty^*$  (to this regards, see Remark B.3 after the following proof), the result below provides a central limit theorem for the difference  $(t^{1-\delta} X_t - X_\infty^*)$ .

**Theorem B.4 (CLT).** *Given the dynamics (39) and the existence of  $X_\infty^* = a.s. - \lim_t t^{1-\delta} X_t$ , assume  $0 \leq \rho < \delta$ ,*

$$(40) \quad E \left[ \sum_{t=1}^{\infty} \frac{|\Delta M_{t+1}|^p}{t^{\delta+\epsilon}} \right] < +\infty$$

for  $p = 2$  and  $p = 4$  and some  $0 < \epsilon < \delta$ , and

$$(41) \quad t^{1-\delta} E[(\Delta M_{t+1})^2 | \mathcal{F}_t] \xrightarrow{a.s.} V_\infty^*$$

for some real ( $\mathcal{F}_\infty$ -measurable) random variable  $V_\infty^*$ . Then:

(i) For  $0 \leq \rho \leq \delta/2$ , we have

$$(42) \quad t^{\delta/2} (t^{1-\delta} X_t - X_\infty^*) \xrightarrow{stably} \begin{cases} \mathcal{N}(0, V_\infty^*/\delta) & \text{if } 0 \leq \rho < \delta/2; \\ \mathcal{N}(-2\rho_\infty/\delta, V_\infty^*/\delta) & \text{if } \rho = \delta/2; \end{cases}$$

or, equivalently, in a compact form

$$(43) \quad t^{\delta/2} (t^{1-\delta} X_t + \frac{\rho_\infty}{(\delta-\rho)} t^{-(\delta-\rho)} - X_\infty^*) \xrightarrow{\text{stably}} \mathcal{N}(0, V_\infty^*/\delta),$$

and the all the limit relations are also in the sense of the a.s. conditional convergence with respect to the filtration  $\mathcal{F}$ .

(ii) For  $\delta/2 < \rho < \delta$ , we have

$$t^{\delta-\rho} (t^{1-\delta} X_t - X_\infty^*) \xrightarrow{P} -\frac{\rho_\infty}{(\delta-\rho)}.$$

Moreover, adding the assumption  $(*)$   $t^{\rho-\delta/2} (t^{1-\rho} \rho_{t+1} - \rho_\infty) \xrightarrow{P} 0$ , the stable convergence (43) holds true. Finally, if the additional assumption  $(*)$  holds true in the sense of the a.s. convergence, then the limit relation (43) also holds true in the sense of the a.s. conditional convergence.

**Remark B.2** (Comparison with the CLTs by L.-X. Zhang [53]). The dynamics of the process  $X_t^* = \zeta_t(1-\delta)X_t$ , where  $\zeta_t(x) \sim t^x$  is defined in (22), is

$$X_{t+1}^* = X_t^* - \frac{(1-\delta)^2}{(\theta+t+1)^2} X_t^* + \frac{\zeta_{t+1}(1-\delta)\Delta M_{t+1}}{\theta+t+1} + \frac{\zeta_{t+1}(1-\delta)}{\theta+t+1} \rho_t$$

and, if we compare it with the dynamics (1.1) in [53], then, using the notation in that paper, we get  $h(\cdot) = 0$ , a case which is excluded in [53]. In addition, there is also another important difference. For  $\delta < 1$ , we have by (41) that  $E[(\zeta_{t+1}(1-\delta)\Delta M_{t+1})^2 | \mathcal{F}_t]$  tends a.s. to infinity and so assumption (2.4) in [53] does not hold. Therefore Theorem B.4 is not a consequence of the results in [53].

Before presenting the proof of Theorem B.4, we will report a preliminary result used in that proof. Indeed, a crucial point will be played by the following martingale:  $\widehat{M}_0 = 0$  and, for any  $t \geq 0$ ,

$$(44) \quad \widehat{M}_{t+1} = \sum_{k=0}^t \Delta \widehat{M}_{k+1}, \quad \text{with} \quad \Delta \widehat{M}_{k+1} = \frac{r_k}{\theta+k+1} \Delta M_{k+1},$$

where the sequence  $(r_k)$  is defined as in (29). The proof of the following auxiliary result will be presented after the proof of the CLT (Theorem B.4).

**Theorem B.5.** *Given the dynamics (39), let  $\delta_0$  and  $(b_t)_t$  defined as in (28) and  $(\widehat{M}_t)_t$  as in (44). Assume that, for  $p = 2$  and  $p = 4$  and some  $0 < \epsilon < \delta_0$ , we have*

$$(45) \quad E \left[ \sum_t t^{p\delta_0} \frac{|\Delta \widehat{M}_{t+1}|^p}{t^{\delta_0+\epsilon}} \right] < +\infty$$

and

$$(46) \quad \frac{(\theta+t+1)^2}{r_t} E[(\Delta \widehat{M}_{t+1})^2 | \mathcal{F}_t] = r_t E[(\Delta M_{t+1})^2 | \mathcal{F}_t] \xrightarrow{\text{a.s.}} V_\infty^*$$

for some real ( $\mathcal{F}_\infty$ -measurable) random variable  $V_\infty^*$ . Then,  $(\widehat{M}_t)_t$  is a martingale bounded in  $L^4$  and so it converges a.s. and in  $L^4$  to a real random variable  $\widehat{M}_\infty$  with  $E[\widehat{M}_\infty^4] < +\infty$ . Moreover, we have

$$\sqrt{b_t}(\widehat{M}_t - \widehat{M}_\infty) \rightarrow \mathcal{N}(0, V_\infty^*/\delta_0),$$

where the convergence is in the sense of the a.s. conditional convergence with respect to the filtration  $\mathcal{F}$  (and so also in the sense of the stable convergence).

We can now present the proof of the CLT.

*Proof of Theorem B.4.* Let us consider the process  $X^* = (X_t^*)_t$  defined as  $X_t^* = \zeta_t(1-\delta)X_t$ , where  $\zeta_t(x) \sim t^x$  is defined in (22), so that  $X_t^* \xrightarrow{\text{a.s.}} t^{1-\delta} X_t \xrightarrow{\text{a.s.}} X_\infty^*$  by assumption. Starting from dynamics (39) and recalling that  $r_k = \zeta_k(1-\delta)$  (by assumption  $\delta > \rho$ ), we can follow the analogous calculations as in (37) to get

$$(47) \quad X_{k+1}^* - X_k^* = A_k + B_{k+1} + C_{k+1} + D_{k+1},$$

where

$$\begin{cases} A_k = -\left(\frac{1-\delta}{\theta+k+1}\right)^2 X_k^* & (\text{which is } 0 \text{ if } \delta = 1); \\ B_{k+1} = \frac{\zeta_{k+1}(1-\delta)\rho_{k+1}}{\theta+k+1}; \\ C_{k+1} = \left(\frac{\zeta_{k+1}(1-\delta)}{\zeta_k(1-\delta)} - 1\right)\Delta\widehat{M}_{k+1} & (\text{which is } 0 \text{ if } \delta = 1); \\ D_{k+1} = \Delta\widehat{M}_{k+1}. \end{cases}$$

Note that  $\sum_{k \geq t} (X_{k+1}^* - X_k^*) = (X_\infty^* - X_t^*)$ . Therefore, it is enough to study the asymptotic behavior of the four terms  $\sum_{k \geq t} A_k$ ,  $\sum_{k \geq t} B_{k+1}$ ,  $\sum_{k \geq t} C_{k+1}$  and  $\sum_{k \geq t} D_{k+1}$ .

Let us start with the last term  $\sum_{k \geq t} D_{k+1}$ . First of all we observe that, since  $\delta > \rho$  by assumption, we have  $r_t \sim t^{1-\delta}$  and so

$$E[|\Delta\widehat{M}_{t+1}|^p] = E[|\Delta M_{t+1}|^p] \frac{r_t^p}{(\theta+t+1)^p} = E[|\Delta M_{t+1}|^p] O(t^{-p\delta}).$$

Hence, relation (45) with  $\delta_0 = \delta$  follows from (40). Moreover, condition (41) equals condition (46). It follows by Theorem B.5 and the fact that  $b_t \sim t^\delta$ , that we have  $\widehat{M}_t \xrightarrow{a.s.} \widehat{M}_\infty$  for a suitable real random variable  $\widehat{M}_\infty$  and

$$t^{\delta/2} \sum_{k \geq t} D_{k+1} = t^{\delta/2} \sum_{k \geq t} \Delta\widehat{M}_{k+1} = t^{\delta/2} (\widehat{M}_\infty - \widehat{M}_t) \xrightarrow{stably} \mathcal{N}(0, V_\infty^*/\delta),$$

where the convergence is also in the sense of the a.s. conditional convergence with respect to the filtration  $\mathcal{F}$ .

Let us now focus on the term  $\sum_{k \geq t} C_{k+1}$  when  $\delta < 1$ . Let us define  $Y_k = \zeta_k(1-\delta)\Delta M_{k+1}$  and  $\mathcal{G}_k = \mathcal{F}_{k+1}$ , then  $E[Y_k | \mathcal{G}_{k-1}] = 0$ . Now, let  $\tilde{a}_k = k^{-\delta}(\theta+k+1)^2$  and  $\tilde{b}_k = k^{\delta/2}$ . Then  $\tilde{b}_k \uparrow +\infty$  and

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{E[Y_k^2]}{\tilde{a}_k^2 \tilde{b}_k^2} &= \sum_{k=1}^{\infty} \frac{\zeta_k(1-\delta)^2}{\tilde{a}_k^2 \tilde{b}_k^2} E[(\Delta M_{k+1})^2] = \sum_{k=1}^{\infty} \frac{\zeta_k(1-\delta)^2(\theta+k+1)^2}{\tilde{a}_k^2 \tilde{b}_k^2 r_k^2} E[(\Delta\widehat{M}_{k+1})^2] \\ &= \sum_{k=1}^{\infty} \frac{\zeta_k(1-\delta)^2(\theta+k+1)^2}{\tilde{a}_k^2 \tilde{b}_k^2 \zeta_k(1-\delta)^2} E[(\Delta\widehat{M}_{k+1})^2] = \sum_{k=1}^{\infty} \frac{k^\delta}{(\theta+k+1)^2} E[(\Delta\widehat{M}_{k+1})^2] \\ &\leq \sum_{k=1}^{\infty} k^{2\delta} \frac{E[(\Delta\widehat{M}_{k+1})^2]}{k^{\delta+\epsilon}}, \end{aligned}$$

which is finite, by (45) with  $p = 2$ ,  $\delta_0 = \delta$  and some  $0 < \epsilon < \delta$ . Since  $\tilde{b}_t \sum_{k \geq t} \frac{1}{\tilde{a}_k \tilde{b}_k^2} \rightarrow 0$ , then by Lemma E.5, we obtain

$$\begin{aligned} t^{\delta/2} \sum_{k \geq t} C_{k+1} &= t^{\delta/2} \sum_{k \geq t} \left( \frac{\zeta_{k+1}(1-\delta)}{\zeta_k(1-\delta)} - 1 \right) \Delta\widehat{M}_{k+1} = t^{\delta/2} \sum_{k \geq t} \frac{1-\delta}{\theta+k+1} \Delta\widehat{M}_{k+1} \\ &= (1-\delta) t^{\delta/2} \sum_{k \geq t} \frac{\zeta_k(1-\delta)\Delta M_{k+1}}{(\theta+k+1)^2} = (1-\delta) \tilde{b}_t \sum_{k \geq t} \frac{Y_k}{\tilde{a}_k \tilde{b}_k^2} \xrightarrow{a.s.} 0. \end{aligned}$$

Let us now focus on the term  $\sum_{k \geq t} A_k$  when  $\delta < 1$ . Since  $X_t^*$  converges a.s. to a real random variable  $X_\infty^*$ , we have

$$t^{\delta/2} \sum_{k \geq t} A_k = t^{\delta/2} \sum_{k \geq t} \left( \frac{1-\delta}{\theta+k+1} \right)^2 X_k^* = t^{\delta/2} \sum_{k \geq t} \frac{O(1)}{k^{1+\delta}} \xrightarrow{a.s.} 0.$$

Finally, regarding the term  $\sum_{k \geq t} B_{k+1}$ , since (23), we have

$$(48) \quad t^{\delta-\rho} \zeta_{k+1}(1-\delta)\rho_{k+1} = \rho_\infty + o(1),$$

and hence

$$\sum_{k \geq t} B_{k+1} = \sum_{k \geq t} \frac{\zeta_{k+1}(1-\delta)\rho_{k+1}}{\theta+k+1} = \sum_{k \geq t} \frac{\rho_\infty}{k^{1+\delta-\rho}} + o\left(\frac{1}{t^{\delta-\rho}}\right) \sim \frac{\rho_\infty}{(\delta-\rho)} \frac{1}{t^{\delta-\rho}},$$

so that

$$\begin{aligned} t^{\delta/2} \sum_{k \geq t} B_{k+1} &= \frac{t^{\delta/2}}{t^{\delta-\rho}} t^{\delta-\rho} \sum_{k \geq t} B_{k+1} \xrightarrow{a.s.} \begin{cases} 0 & \text{if } \rho < \delta/2; \\ \rho_\infty/\rho = 2\rho_\infty/\delta & \text{if } \rho = \delta/2; \end{cases} \\ &\quad t^{\delta-\rho} \sum_{k \geq t} B_{k+1} \xrightarrow{a.s.} \frac{\rho_\infty}{\delta-\rho} \quad \text{if } \rho > \delta/2. \end{aligned}$$

Recalling that the stable convergence combines well with convergence in probability (e.g. Theorem 2.4.2 in [13]), and so with the a.s. convergence, and the a.s. conditional convergence combines well with the a.s. convergence (see Lemma E.7), we are now ready to conclude:

- (i) if  $\rho \leq \delta/2$ , then, we have to combine the stable convergence and the a.s. conditional convergence of the suitable rescaled sequence  $\sum_{k \geq t} D_{k+1} = (\widehat{M}_\infty - \widehat{M}_t)$  with the a.s. convergences of the suitably rescaled sequences  $\sum_{k \geq t} A_k$  (only when  $\delta < 1$ ),  $\sum_{k \geq t} B_{k+1}$  and  $\sum_{k \geq t} C_{k+1}$  (only when  $\delta < 1$ ) so that we obtain (42) (with  $X_t^*$  instead of  $t^{1-\delta}X_t$ ), in the sense of the stable convergence and in the sense of the a.s. conditional convergence with respect to  $\mathcal{F}$ .
- (ii) if  $\delta/2 < \rho < \delta$ , then the leading term is  $\sum_{k \geq t} B_{k+1}$  and the a.s. convergence of  $t^{\delta-\rho} \sum_{k \geq t} B_{k+1}$  has to be combined with the a.s. limit results for the other terms involving  $A_k$  and  $C_{k+1}$  (only when  $\delta < 1$ ) and, above all, with the convergence in probability to zero of  $\frac{t^{\delta-\rho}}{t^{\delta/2}} t^{\delta/2} (\widehat{M}_\infty - \widehat{M}_t)$ , where  $\frac{t^{\delta-\rho}}{t^{\delta/2}} = t^{-(\rho-\delta/2)} \rightarrow 0$ .

Furthermore, the (43) (with  $X_t^*$  instead of  $t^{1-\delta}X_t$ ) is obvious when  $\rho \leq \delta/2$ , because the stable convergence and the a.s. conditional convergence combine well with the a.s. convergence of  $t^{-(\delta/2-\rho)}\rho_\infty/(\delta-\rho)$  to zero when  $\rho < \delta/2$  and to  $2\rho_\infty/\delta$  when  $\rho = \delta/2$ . In the case  $\delta/2 < \rho < \delta$ , if we assume  $(*)$   $t^{\rho-\delta/2}(t^{1-\rho}\rho_{t+1} - \rho_\infty) \xrightarrow{P} 0$ , then, equation (48) becomes

$$k^{\delta-\rho}\zeta_{k+1}(1-\delta)\rho_{k+1} = \rho_\infty + o_P(k^{-(\rho-\delta/2)}),$$

so that

$$t^{\delta/2} \left( \sum_{k \geq t} B_{k+1} - \frac{\rho_\infty}{(\delta-\rho)} \frac{1}{t^{\delta-\rho}} \right) = o_P(1) \xrightarrow{P} 0$$

and this convergence in probability combines well with the stable convergence of  $t^{\delta/2} \sum_{k \geq t} D_{k+1} = t^{\delta/2}(\widehat{M}_\infty - \widehat{M}_t)$ . If the additional assumed convergence  $(*)$  is a.s., then we can replace  $o_P()$  with  $o()$  and the above limit relation is in the sense of a.s. convergence and it combines well with the a.s. conditional convergence of  $t^{\delta/2} \sum_{k \geq t} D_{k+1}$ .

Finally, we can obtain the results for  $t^{1-\delta}X_t$  from those proven for  $X_t^*$ , because of (23) with  $x = 1 - \delta$ .  $\square$

**Remark B.3** (Sufficient conditions for the a.s. convergence of  $t^{1-\delta}X_t$ ). Consider  $X_t^* = \zeta_t(1-\delta)X_t$ , where  $\zeta_t(x) \sim t^x$  is defined in (22). From (47), if we can take the conditional expectation, we get

$$E[X_{t+1}^* | \mathcal{F}_t] - X_t^* = A_t + \widetilde{B}_t,$$

with  $\widetilde{B}_t = E[B_{t+1} | \mathcal{F}_t]$ . Then, if the process  $X$  (and so  $X^*$ ) is non-negative, we get that  $A_t \leq 0$  for each  $t$  and  $X^*$  results to be a non-negative almost martingale which converges a.s. when  $\sum_t |\widetilde{B}_t| < +\infty$  a.s. (as it is, for instance, when  $\rho < \delta$  and  $\rho_{t+1}$  is  $\mathcal{F}_t$ -measurable so that  $\widetilde{B}_t = B_{t+1} = O(1/t^{1+\delta-\rho})$ ).

In general, if  $\sum_t E[|A_t|] + \sum_t E[|\widetilde{B}_t|] < +\infty$ , the process  $X^*$  is a quasi-martingale and a sufficient condition for its a.s. convergence is  $\sup_t E[|X_t^*|] < +\infty$  (which implies  $\sum_t E[|A_t|] < +\infty$ ). Summing up,  $\sup_t E[|X_t^*|] < +\infty$  and  $\sum_t E[|B_{t+1}|] < +\infty$  (which implies  $\sum_t E[|\widetilde{B}_t|] < +\infty$ ) are sufficient conditions for the a.s. convergence of  $X_t^*$ .

*Proof of Theorem B.5.* First of all we observe that relation (45) with  $p = 4$  implies that  $(\widehat{M}_t)_t$  is a martingale bounded in  $L^4$  and so it converges a.s. and in  $L^4$  (in particular, in mean) to a real  $(\mathcal{F}_\infty$ -measurable) random variable  $\widehat{M}_\infty$  with  $E[\widehat{M}_\infty^4] < +\infty$ . The sequel of the proof is based on Theorem E.6. To this aim, for any  $n \geq 1$ , define  $M_{n,0} = 0$  and  $M_{n,t} = \sqrt{b_n}(\widehat{M}_n - \widehat{M}_{n+t-1})$  for  $t \geq 1$  (which is  $\mathcal{F}_{n+t-1}$ -measurable). We have  $M_{n,t} \xrightarrow{L^1} \sqrt{b_n}(\widehat{M}_n - \widehat{M}_\infty)$  as  $t \rightarrow +\infty$ . Moreover, we have  $X_{n,1} = M_{n,1} - M_{n,0} = 0$  and, for each  $t \geq 2$

$$X_{n,t} = M_{n,t} - M_{n,t-1} = \sqrt{b_n}(\widehat{M}_{n+t-2} - \widehat{M}_{n+t-1}),$$

so that

$$U_n = b_n \sum_{t \geq n} (\widehat{M}_{t+1} - \widehat{M}_t)^2, \quad X_n^* = \sqrt{b_n} \sup_{t \geq n} |\widehat{M}_{t+1} - \widehat{M}_t|.$$

*First step:*  $X_n^* \xrightarrow{a.s.} 0$ : From (45), with  $p = 4$  and some  $0 < \epsilon < \delta_0$ , we have

$$(49) \quad \begin{aligned} E \left[ \sum_t (\sqrt{b_t})^4 |\Delta \widehat{M}_{t+1}|^4 \right] &\leq E \left[ \sum_t t^{2\delta_0 + (\delta_0 - \epsilon)} |\Delta \widehat{M}_{t+1}|^4 \right] \\ &= E \left[ \sum_t t^{p\delta_0} \frac{|\Delta \widehat{M}_{t+1}|^p}{t^{\delta_0 + \epsilon}} \right] < +\infty, \end{aligned}$$

which implies  $\sum_t (\sqrt{b_t})^4 (\Delta \widehat{M}_{t+1})^4 < +\infty$  a.s.. Then we have

$$(X_n^*)^4 \leq \sup_{t \geq n} (\sqrt{b_t})^4 |\Delta \widehat{M}_{t+1}|^4 \leq \sum_{t \geq n} (\sqrt{b_t})^4 |\Delta \widehat{M}_{t+1}|^4 \xrightarrow{a.s.} 0 \quad \text{as } n \rightarrow +\infty.$$

*Second step:*  $(X_n^*)_n$  dominated in  $L^1$ : We are going to show that  $(X_n^*)_n$  is dominated in  $L^4$ . In fact, with  $p = 4$  and some  $0 < \epsilon < \delta_0$ , we have

$$\begin{aligned} E \left[ \left( \sup_n X_n^* \right)^4 \right] &\leq E \left[ \sup_n \sup_{t \geq n} (\sqrt{b_t})^4 |\Delta \widehat{M}_{t+1}|^4 \right] \\ &\leq E \left[ \sum_t (\sqrt{b_t})^4 |\Delta \widehat{M}_{t+1}|^4 \right] < +\infty \end{aligned}$$

as shown in (49) in the first step.

*Third step:*  $U_n \xrightarrow{a.s.} V_\infty^*/\delta_0$ : Let  $a_k$  as defined in (28) (so that  $a_k b_k^2 = (\theta + k + 1)^2/r_k$ ) and set  $Y_k = a_k b_k^2 (\Delta \widehat{M}_{k+1})^2 = r_k (\Delta M_{k+1})^2$ . Then, we have  $E[Y_k | \mathcal{F}_{k-1}] \rightarrow V_\infty^*$  by assumption (46). Moreover, we have  $b_k \uparrow +\infty$  and, by (45) with  $p = 4$  and some  $0 < \epsilon < \delta_0$ , we get

$$\sum_{k=1}^{\infty} \frac{E[Y_k^2]}{a_k^2 b_k^2} = \sum_{k=1}^{\infty} b_k^2 E[(\Delta \widehat{M}_{k+1})^4] \leq C \sum_{k=1}^{\infty} k^{4\delta_0} \frac{E[(\Delta \widehat{M}_{k+1})^4]}{k^{\delta_0 + \epsilon}} < +\infty.$$

Since  $b_t \sum_{k \geq t} \frac{1}{a_k b_k^2} \rightarrow \frac{1}{\delta_0}$  by the technical Lemma A.1, then  $U_t = b_t \sum_{k \geq t} \frac{Y_k}{a_k b_k^2} \xrightarrow{a.s.} V_\infty^*/\delta_0$  by Lemma E.5.  $\square$

#### APPENDIX C. PROOF OF THE RESULTS IN SECTION 4

We here collect the proofs for the results presented in Section 4 regarding the predictive mean  $Z_t$  and the averaged quantities  $\bar{T}_t$ ,  $\bar{P}_t$  and  $\bar{K}_t$ .

##### C.1. First-order asymptotic results.

*Proof of Theorem 4.1.* In the case  $w > \beta$  the term involving  $\lambda_t$  in the dynamics (8) of  $(Z_t)_t$  behaves essentially as a remainder and hence we can focus on the process  $S_t = \sum_{j \in \mathcal{O}_t} P_{t,j} = Z_t - \lambda_t$ , so that from (8) we get

$$\begin{aligned} S_{t+1} - S_t &= Z_{t+1} - Z_t - (\lambda_{t+1} - \lambda_t) \\ &= w \frac{T_{t+1} - Z_t}{\theta + t + 1} - (1 - w) \frac{Z_t}{\theta + t + 1} + \frac{\lambda_t}{\theta + t + 1} \\ &= -\frac{Z_t}{\theta + t + 1} + w \frac{T_{t+1}}{\theta + t + 1} + \frac{\lambda_t}{\theta + t + 1} = -\frac{S_t}{\theta + t + 1} + \frac{w T_{t+1}}{\theta + t + 1}. \end{aligned}$$

The above dynamics for  $S_t$  can be written as

$$(50) \quad S_{t+1} = \left(1 - \frac{1}{\theta + t + 1}\right) S_t + w \frac{T_{t+1}}{\theta + t + 1} = \left(1 - \frac{1-w}{\theta + t + 1}\right) S_t + \frac{\Delta M_{t+1}}{\theta + t + 1} + \frac{w \lambda_t}{\theta + t + 1},$$

with  $\Delta M_{t+1} = w(T_{t+1} - Z_t)$ . Denoting  $S_t^* = \zeta_t(1 - w)S_t$ , where  $\zeta_t(x) \sim t^x$  is defined in (22), we can obtain

$$S_{t+1}^* = \left(1 - \frac{(1-w)^2}{(\theta + t + 1)^2}\right) S_t^* + \frac{\zeta_{t+1}(1-w)\Delta M_{t+1}}{\theta + t + 1} + \frac{w \zeta_{t+1}(1-w)\lambda_t}{\theta + t + 1},$$

and so, by the classical Theorem E.4, the process  $(S_t^*)_t$  is a non-negative almost super-martingale that converges a.s. to a finite r.v.  $S_\infty^*$ . However, this is not enough for our scopes. Indeed, in that way, we do not know that such random limit  $S_\infty^*$  is a.s. strictly positive, and this is crucial for proving the right rate of convergence. For this reason we are going to apply the more sophisticated Theorem B.3 with  $X_t = S_t$ ,  $Y_{t+1} = w T_{t+1}$  and  $\delta = w \in (0, 1]$ . To this end, we observe that we have

$$\sum_t Y_t = w \sum_t T_t \geq w \sum_t N_t = w \sup_t \sum_{n=1}^t N_n = w \sup_t D_t = +\infty \quad \text{a.s.}$$



(by (10)) and

$$E[Y_{t+1} | \mathcal{F}_t] = E[wT_{t+1} | \mathcal{F}_t] = wZ_t = wS_t + w\lambda_t = \delta X_t + \rho_t,$$

with  $\rho_t = w\lambda_t > 0$  and

$$\sum_t \frac{\rho_t}{(\theta+t)^{\delta-\epsilon}} = O\left(\sum_t \frac{1}{(\theta+t)^{1+w-(\beta+\epsilon)}}\right) < +\infty$$

for some  $\epsilon > 0$  such that  $w - (\beta + \epsilon) > 0$  (note that this  $\epsilon$  exists because  $w > \beta$  by assumption). Therefore, by Theorem B.3, we firstly obtain that  $S_t = O(1/t^{1-w})$  and so  $Z_t = S_t + \lambda_t = O(1/t^{1-w}) + O(1/t^{1-\beta}) = O(1/t^{1-w})$  because  $\beta < w$ . Hence, by (9), we get

$$E[Y_{t+1}^2 | \mathcal{F}_t] = w^2(Z_t + Z_t^2 - R_t) = O(Z_t) = O(E[Y_{t+1} | \mathcal{F}_t]).$$

Therefore, we can apply Theorem B.3 (case *i*)) and we obtain that

$$t^{1-w}S_t \xrightarrow{a.s.} S_\infty^* \in (0, +\infty).$$

As before, since  $Z_t = S_t + \lambda_t$  and  $\beta < w$ , we get

$$t^{1-w}Z_t = t^{1-w}S_t + t^{1-w}\lambda_t \xrightarrow{a.s.} S_\infty^*$$

and so the first convergence holds true with  $Z_\infty^* = S_\infty^*$ .  $\square$

*Proof of Theorem 4.2.* In the case  $w \leq \beta$  the term involving  $\lambda_t$  in the dynamics (8) of  $(Z_t)_t$  cannot be considered as a remainder term as in the proof of Theorem 4.1. Indeed, letting

$$Z_t^* = r_t Z_t \quad \text{with} \quad r_t = \begin{cases} \zeta_t(1-\beta) & \text{when } w < \beta, \\ \frac{\zeta_t(1-\beta)}{\ln(\theta+t)} & \text{when } w = \beta, \end{cases}$$

where  $\zeta_t(x) \sim t^x$  is defined in (22). From (8), taking into account (25) and (27), we get

$$(51) \quad Z_{t+1}^* - Z_t^* = \eta_{t+1}(\alpha\beta - bZ_t^*) + \eta_{t+1}\Delta M_{t+1}^* + \rho_{t+1},$$

where  $\rho_{t+1} = Z_t^*O(\frac{1}{t^2}) + O(\frac{1}{t^{1+\beta}})$  (here  $O(\frac{1}{t^2})$  and  $O(\frac{1}{t^{1+\beta}})$  are deterministic terms and  $\beta > 0$ ),  $\Delta M_{t+1}^* = w\zeta_{t+1}(1-\beta)(T_{t+1} - Z_t)$  and

$$\begin{cases} b = \beta - w & \text{and} & \eta_{t+1} = \frac{1}{(\theta+t+1)} & \text{if } w < \beta; \\ b = 1 & \text{and} & \eta_{t+1} = \frac{1}{(\theta+t+1)\ln(\theta+t+1)} & \text{if } w = \beta. \end{cases}$$

This suggests that  $(Z_t^*)_t$  is not a non-negative almost super-martingale, as in the case  $w > \beta$ . Instead, it evolves according to the stochastic approximation dynamics in (51), which suggests convergence toward a deterministic limit given by  $\alpha\beta/b$ . For this reason, we apply Theorem B.1 to the process  $(Z_t^*)_t$  governed by (51). To this end, note that from Lemma D.1 we have that  $\sup_t E[r_t Z_t] < +\infty$  and hence, since by (9),

$$\begin{aligned} E[(\Delta M_{t+1}^*)^2 | \mathcal{F}_t] &= w^2\zeta_{t+1}(1-\beta)^2 (E[T_{t+1}^2 | \mathcal{F}_t] - Z_t^2) \\ &= w^2\zeta_{t+1}(1-\beta)^2 (Z_t - R_t) \\ &\leq w^2\zeta_{t+1}(1-\beta)^2 Z_t, \end{aligned}$$

we get, for any  $w \leq \beta$ , that

$$\sum_t \eta_{t+1}^2 E[(\Delta M_{t+1}^*)^2] \leq \sum_t \frac{w^2 r_t}{(\theta+t+1)^2} E[r_t Z_t] < +\infty,$$

which implies

$$\sum_t \eta_{t+1}^2 E[(\Delta M_{t+1}^*)^2 | \mathcal{F}_t] < +\infty \quad \text{a.s.}$$

In addition, since  $\sup_t E[Z_t^*] = \sup_t E[r_t Z_t] < +\infty$  (by Lemma D.1), the expression for  $\rho_{t+1}$  implies that  $\sum_t E[|\rho_{t+1}|] = \sum_t O(E[Z_t^*]/t^2) + O(1/t^{1+\beta})$  is convergent, and so  $\sum_t |\rho_{t+1}| < +\infty$  with probability one. Then, we can apply Theorem B.1 to the dynamics (51) so obtaining

$$Z_t^* \xrightarrow{a.s.} \frac{\alpha\beta}{b} = \begin{cases} \frac{\alpha\beta}{\beta-w} & \text{if } w < \beta; \\ \alpha\beta & \text{if } w = \beta. \end{cases}$$

$\square$

## C.2. Second-order asymptotic results.

*Proof of Theorem 4.3.* We will first prove the CLT for the process  $S_t = \sum_{j \in \mathcal{O}_t} P_{t,j}$  and then, since  $S_t = Z_t - \lambda_t$ , we will derive from it the one for  $Z_t$ . Indeed, since the dynamics of  $S_t$  in (50), i.e.

$$S_{t+1} = \left(1 - \frac{1-w}{\theta+t+1}\right) S_t + \frac{w(T_{t+1}-Z_t)}{\theta+t+1} + \frac{w\lambda_t}{\theta+t+1},$$

is equal to the general dynamics in (39) with  $\delta = w \in (0, 1]$ ,  $\Delta M_{t+1} = w(T_{t+1} - Z_t)$ ,  $\rho_{t+1} = w\lambda_t$ ,  $\rho = \beta \in [0, 1]$  and  $\rho_\infty = w\alpha$  (because  $t^{1-\beta}\lambda_t$  trivially converges to  $\alpha$ ), and since we are assuming  $\beta < w$ , that is  $\rho < \delta$ , the CLT for  $S_t$  can be obtained by applying Theorem B.4 with  $X_t = S_t$  and  $X_\infty^* = Z_\infty^*$  (recall that, by Theorem 4.1,  $t^{1-w}S_t \xrightarrow{a.s.} Z_\infty^*$ ). Indeed, condition (40) follows by (53) in Lemma D.1 as for any  $p \geq 1$  and  $\epsilon > 0$  we have

$$E\left[\sum_{t=1}^{\infty} \frac{|\Delta M_{t+1}|^p}{t^{w+\epsilon}}\right] = \sum_{t=1}^{\infty} \frac{1}{t^{1+\epsilon}} E[t^{1-w}|\Delta M_{t+1}|^p] < +\infty.$$

Moreover, condition (41) follows by (9), Theorem 4.1 and Lemma D.2 (in the case  $\beta < w$ ):

$$t^{1-w} E[(\Delta M_{t+1})^2 | \mathcal{F}_t] = w^2 t^{1-w} (Z_t - R_t) \xrightarrow{a.s.} V_\infty^* = w^2 \Sigma_\infty.$$

Then, since  $(t^{1-\beta}\lambda_t - \alpha) = \alpha[1 - 1/(t+1)]^{1-\beta} - \alpha = O(t^{-1}) = o(t^{-(\beta-w/2)})$ , also condition (\*) of Theorem B.4 is verified with a.s. convergence and so, from that theorem, we get

$$t^{w/2}(t^{1-w}S_t + \frac{w\alpha}{(w-\beta)}t^{-(w-\beta)} - Z_\infty^*) \xrightarrow{stably} \mathcal{N}(0, w\Sigma_\infty).$$

For what concerns the process  $(Z_t)_t$  recall that  $Z_t = S_t + \lambda_t$  and so  $t^{1-w}Z_t = t^{1-w}S_t + t^{1-w}\lambda_t$ , and hence the thesis follows by the fact that  $t^{1-w}\lambda_t = \alpha t^{\beta-w} + O(t^{-(1+w-\beta)}) = \alpha t^{\beta-w} + o(t^{-w/2})$ . In addition, note that, from (12), for  $w/2 < \beta < w$  we have

$$\begin{aligned} t^{w-\beta}(t^{1-w}Z_t - Z_\infty^*) + \frac{\alpha\beta}{w-\beta} &= \frac{1}{t^{\beta-w/2}} t^{w/2} t^{\beta-w} [t^{w-\beta}(t^{1-w}Z_t - Z_\infty^*) + \frac{\alpha\beta}{w-\beta}] \\ &= \frac{1}{t^{\beta-w/2}} t^{w/2} [(t^{1-w}Z_t - Z_\infty^*) + \frac{\alpha\beta}{w-\beta} t^{\beta-w}] \xrightarrow{P} 0. \end{aligned}$$

□

*Proof of Corollary 4.4.* From from (12), we can derive the corresponding result for  $\bar{T}_t$ . Indeed, by (4), we have  $\bar{T}_t = (1 + \frac{\theta}{t})S_t/w$  and, in addition, we have  $\lambda_t = \alpha/(t+1)^{1-\beta} = \alpha/t^{1-\beta} + (1-\beta)O(1/t^{2-\beta})$ , so that we obtain

$$\begin{aligned} t^{w/2}(t^{1-w}\bar{T}_t + \frac{\alpha}{(w-\beta)}t^{-(w-\beta)} - \frac{1}{w}Z_\infty^*) &= \\ t^{w/2}(t^{1-w}\bar{T}_t + \frac{\alpha\beta}{w(w-\beta)}t^{-(w-\beta)} + \frac{t^{1-w}}{w} \frac{\alpha}{t^{1-\beta}} - \frac{1}{w}Z_\infty^*) &= \\ t^{w/2}(t^{1-w}\bar{T}_t + \frac{\alpha\beta}{w(w-\beta)}t^{-(w-\beta)} + \frac{t^{1-w}}{w}\lambda_t - \frac{1}{w}Z_\infty^*) + O(1/t^{1+w/2-\beta}) &= \\ \frac{t^{w/2}}{w}(t^{1-w}Z_t + \frac{\alpha\beta}{(w-\beta)}t^{\beta-w} - Z_\infty^*) + \frac{\theta}{w} \frac{S_t}{t^{w/2}} + O(\frac{1}{t^{1+w/2-\beta}}) &\xrightarrow{stably} \mathcal{N}(0, \Sigma_\infty/w) \end{aligned}$$

and this convergence also holds true in the sense of the a.s. conditional convergence with respect to  $\mathcal{F}$ . Consequently, when  $(w-\beta) < w/2$ , the quantity

$$t^{w-\beta}(t^{1-w}\bar{T}_t - \frac{1}{w}Z_\infty^*) + \frac{\alpha}{(w-\beta)} = \frac{t^{w-\beta}}{t^{w/2}} t^{w/2}(t^{1-w}\bar{T}_t + \frac{\alpha}{(w-\beta)}t^{-(w-\beta)} - \frac{1}{w}Z_\infty^*)$$

converges in probability toward zero.

Now, remember that  $\bar{P}_t = S_t/D_t = (Z_t - \lambda_t)/D_t$  so that, using the notation and the CLT for  $(D_t)_t$  provided in Remark 3.1, we can write

$$\begin{aligned} \bar{P}_t &= \frac{1}{\lambda(\beta)a_t(\beta)} S_t (1 - \frac{D_t - \lambda(\beta)a_t(\beta)}{D_t}) \\ &= \frac{1}{\lambda(\beta)a_t(\beta)} S_t \left[ 1 - \frac{1}{\frac{D_t}{a_t(\beta)\lambda(\beta)}} \frac{1}{\sqrt{a_t(\beta)}} \cdot \frac{\sqrt{a_t(\beta)}}{\lambda(\beta)} \left( \frac{D_t}{a_t(\beta)} - \lambda(\beta) \right) \right] \\ &= \frac{1}{\lambda(\beta)a_t(\beta)} S_t \left[ 1 + \frac{1}{\sqrt{a_t(\beta)}} W_t \right], \end{aligned}$$

where  $W_t \xrightarrow{\text{stably}} \mathcal{N}(0, 1/\lambda(\beta))$ . Hence, since  $\lambda_t = \alpha/t^{1-\beta} + (1-\beta)O(1/t^{2-\beta})$ , we get

$$\begin{aligned} t^{1-w}\lambda(\beta)a_t(\beta)\bar{P}_t + \frac{\alpha w}{w-\beta}t^{-(w-\beta)} - Z_\infty^* &= \\ t^{1-w}\lambda(\beta)a_t(\beta)\bar{P}_t + \alpha t^{-(w-\beta)} + \frac{\alpha\beta}{w-\beta}t^{-(w-\beta)} - Z_\infty^* &= \\ \left(t^{1-w}Z_t + \frac{\alpha\beta}{w-\beta}t^{-(w-\beta)} - Z_\infty^*\right) + O\left(\frac{1}{t^{1+w-\beta}}\right) + \frac{t^{1-w}S_t}{\sqrt{a_t(\beta)}}W_t. \end{aligned}$$

Hence, taking into account (12), since  $\sqrt{a_t(\beta)} = o(t^{w/2})$ , we obtain (14). As a consequence, when  $(w-\beta) < \beta/2$ , we have  $\sqrt{a_t(\beta)} = t^{\beta/2}$  and

$$\begin{aligned} t^{w-\beta} \left( t^{1-w}\lambda(\beta)a_t(\beta)\bar{P}_t - Z_\infty^* \right) + \frac{\alpha w}{w-\beta} &= \\ \frac{t^{w-\beta}}{\sqrt{a_t(\beta)}} \sqrt{a_t(\beta)} \left( t^{1-w}\lambda(\beta)a_t(\beta)\bar{P}_t + \frac{\alpha w}{w-\beta}t^{-(w-\beta)} - Z_\infty^* \right) &\xrightarrow{P} 0. \end{aligned}$$

Finally, recalling that  $\bar{K}_t = (\theta + t)\bar{P}_t/w$  (see (5)), we have

$$\begin{aligned} w\lambda(\beta)\frac{a_t(\beta)}{t^w}\bar{K}_t + \frac{\alpha w}{w-\beta}t^{-(w-\beta)} - Z_\infty^* &= \\ (t^{1-w}\lambda(\beta)a_t(\beta)\bar{P}_t + \frac{\alpha w}{w-\beta}t^{-(w-\beta)} - Z_\infty^*) + \theta\lambda(\beta)\frac{a_t(\beta)}{t^w}\bar{P}_t \end{aligned}$$

so that, from (14), we obtain (15) and, as a consequence, when  $(w-\beta) < \beta/2$ , we have (with the same arguments used above)

$$t^{w-\beta} \left( \lambda(\beta)\frac{a_t(\beta)}{t^w}\bar{K}_t - \frac{Z_\infty^*}{w} \right) + \frac{\alpha}{w-\beta} \xrightarrow{P} 0.$$

□

#### APPENDIX D. TECHNICAL RESULTS RELATED TO THE MODEL

In this section we present some technical lemmas used in the paper. To this end, let us take the sequence  $(r_t)$  as in (29) with  $\delta = w$  e  $\rho = \beta$ .

We prove a lemma regarding a condition of uniform integrability.

**Lemma D.1.** *We have*

$$(52) \quad \sup_{t \geq 1} E[e^{z r_t \bar{T}_t}] < +\infty, \quad \sup_{t \geq 1} E[e^{z r_t Z_t}] < +\infty$$

for some  $z > 0$ . For the sequence  $(T_t)_t$  and any  $p \geq 1$ , we have

$$(53) \quad \begin{cases} \sup_t E[(r_t T_{t+1})^p] < \infty & \text{for } w = 1 \text{ or } \beta = 1; \\ \sup_t E[r_t (T_{t+1})^p] < \infty & \text{for } w < 1 \text{ and } \beta < 1. \end{cases}$$

As a consequence, we note that the sequences  $(r_t Z_t)_t$  and  $(r_t \bar{T}_t)_t$ , are bounded in  $L^p$  for any  $p \geq 1$ .

*Proof.* We split the proof in some steps.

*Step 1):* First notice that, for any  $z > 0$ ,

$$\begin{aligned} E[e^{z T_{t+1}} \mid \mathcal{F}_t] &= E[e^{z \sum_{j \in \mathcal{O}_t} X_{t+1,j} + z N_{t+1}} \mid \mathcal{F}_t] \\ &= E[e^{z N_{t+1}} \mid \mathcal{F}_t] \prod_{j \in \mathcal{O}_t} E[e^{z X_{t+1,j}} \mid \mathcal{F}_t] \\ &= e^{(e^z - 1)\lambda_t} \prod_{j \in \mathcal{O}_t} (e^z P_{t,j} + 1 - P_{t,j}). \end{aligned}$$

Then, denoting  $g(z) = e^z - 1$ , since  $g(z) \geq 0$  and  $1 + x \leq e^x$  for  $x \geq 0$  and using (4) we have

$$\begin{aligned}
 E[e^{z T_{t+1}} | \mathcal{F}_t] &= e^{g(z) \lambda_t} \prod_{j \in \mathcal{O}_t} (1 + g(z) P_{t,j}) \\
 &\leq \exp \left( g(z) \lambda_t + g(z) \sum_{j \in \mathcal{O}_t} P_{t,j} \right) \\
 &= \exp \left( g(z) \lambda_t + g(z) \frac{w \sum_{n=1}^t T_n}{\theta + t} \right) \\
 &= \exp \left( g(z) \lambda_t + g(z) w \frac{t}{\theta + t} \bar{T}_t \right) \quad \text{a.s.}
 \end{aligned}
 \tag{54}$$

*Step 2):* Define  $\bar{T}_t^* = r_t \bar{T}_t$  and note that

$$\begin{aligned}
 \bar{T}_{t+1}^* &= r_{t+1} \frac{\sum_{n=1}^t T_n + T_{t+1}}{t+1} = \frac{r_{t+1}}{r_t} \frac{t}{t+1} \bar{T}_t^* + T_{t+1} \frac{r_{t+1}}{t+1} \\
 &= (1 - v_t) \bar{T}_t^* + u_t T_{t+1},
 \end{aligned}
 \tag{55}$$

where by (25) and (27)

$$u_t = \frac{r_{t+1}}{t+1} \rightarrow 0 \quad \text{and} \quad v_t = 1 - \frac{u_t}{u_{t-1}} \geq 0.$$

Hence, by (54) and (55), we get

$$\begin{aligned}
 E[e^{z \bar{T}_{t+1}^*}] &= E \left[ \exp \left( z \bar{T}_t^* (1 - v_t) \right) E[e^{z u_t T_{t+1}} | \mathcal{F}_t] \right] \\
 &\leq \exp \left( \lambda_t g(z u_t) \right) E \left[ \exp \left( g(z u_t) w \bar{T}_t + z (1 - v_t) \bar{T}_t^* \right) \right] \\
 &= \exp \left( \lambda_t g(z u_t) \right) E \left[ \exp \left( g(z u_t) \frac{w}{r_t} \bar{T}_t^* + z (1 - v_t) \bar{T}_t^* \right) \right].
 \end{aligned}
 \tag{56}$$

Since  $v_t \geq 0$  and  $w \leq 1$ , it follows

$$E[e^{z \bar{T}_{t+1}^*}] \leq \exp \left( \lambda_t g(z u_t) \right) E \left[ \exp \left( \left( z + \frac{g(z u_t)}{r_t} \right) \bar{T}_t^* \right) \right].$$

Iterating this last inequality, we obtain

$$E[e^{z \bar{T}_{t+1}^*}] \leq a_t(z) E[e^{b_t(z) \bar{T}_1^*}]$$

for suitable deterministic quantities  $a_t(z)$  and  $b_t(z)$ . Since we have  $\bar{T}_1^* = r_1 T_1$  with  $T_1 \stackrel{d}{=} \text{Poi}(\alpha)$ , then  $E[e^{b_t(z) r_1 T_1}] < +\infty$  so that  $E[e^{z \bar{T}_t^*}] < +\infty$  for all fixed  $t \geq 1$  and  $z > 0$ . This also implies  $E[e^{z' \bar{T}_t}] < +\infty$  for all fixed  $t \geq 1$  and  $z' > 0$  (take in the above relation  $z = z'/r_t$ ), which by (54) also entails  $E[e^{z T_t}] < +\infty$  for all fixed  $t \geq 1$  and  $z > 0$ .

*Step 3):* Observe now that  $g(x) \leq 2x$  and  $g(x) - x \leq x^2$  for  $x \in [0, 1/2]$ . Moreover, since  $u_t \rightarrow 0$ , we know  $z u_t \leq u_t \leq 1/2$  for  $z \in (0, 1]$  and  $t$  large enough. These facts are used for  $z \in (0, 1]$  to write  $g(z u_t) \leq 2z u_t$  and  $g(z u_t) \leq (z u_t) + (z u_t)^2$  when  $t$  is large enough. Hence, by (56), we get for  $z \in (0, 1]$  and  $t$  large enough:

$$\begin{aligned}
 E[e^{z \bar{T}_{t+1}^*}] &\leq \exp \left( 2z \lambda_t u_t \right) E \left[ \exp \left( (z u_t)^2 \frac{w}{r_t} \bar{T}_t^* + \left( 1 - v_t + u_t \frac{w}{r_t} \right) z \bar{T}_t^* \right) \right] \\
 &\leq \exp \left( 2z \lambda_t u_t \right) E \left[ \left( 1 - v_t + (u_t + u_t^2) \frac{w}{r_t} \right) z \bar{T}_t^* \right],
 \end{aligned}$$

where we have used that  $z^2 \leq z$  as  $z \in (0, 1]$ . Recalling that  $v_t = 1 - u_t/u_{t-1}$  and  $u_{t-1} = r_t/t$ , we have

$$\begin{aligned}
 \pi_t &= 1 - v_t + (u_t + u_t^2) \frac{w}{r_t} = \frac{u_t}{u_{t-1}} + (1 + u_t) \frac{u_t}{u_{t-1}} \frac{w}{t} \\
 &= \frac{u_t}{u_{t-1}} \left( 1 + \frac{w}{t} + \frac{w}{t} u_t \right) = \frac{u_t}{u_{t-1}} \left( 1 + \frac{w}{t} \right) \left( 1 + \frac{w/t}{1 + w/t} u_t \right) \\
 &= \frac{u_t}{u_{t-1}} \frac{\tilde{\zeta}_t(w)}{\tilde{\zeta}_{t-1}(w)} \left( 1 + \frac{w}{t + w} u_t \right),
 \end{aligned}$$

where  $\tilde{\zeta}_t(x)$  is defined as  $\zeta_t(x)$  in (22) with  $\theta = 0$ . This leads for  $z \in (0, 1]$  to

$$E[e^{z \bar{T}_{t+1}^*}] \leq \exp \left( 2z \lambda_t u_t \right) E \left[ \exp \left( z \pi_t \bar{T}_t^* \right) \right]. \tag{57}$$

Now, note that for any  $n < t$

$$(58) \quad \prod_{j=n+1}^t \pi_j = \prod_{j=n+1}^t \frac{u_j}{u_{j-1}} \frac{\tilde{\zeta}_j(w)}{\tilde{\zeta}_{j-1}(w)} \left(1 + \frac{w}{j+w} u_j\right) = \frac{u_t}{u_n} \frac{\tilde{\zeta}_t(w)}{\tilde{\zeta}_n(w)} \prod_{j=n+1}^t \left(1 + \frac{w}{j+w} u_j\right)$$

and since

- (i)  $\limsup_{t \rightarrow \infty} u_t \tilde{\zeta}_t(w) \leq 1$  and
- (ii)  $C = \prod_{t \geq 1} \left(1 + \frac{w}{t+w} u_t\right) < +\infty$  as  $\sum_{t \geq 1} \frac{w}{t+w} u_t < +\infty$

we can choose  $z_0 \in (0, 1)$  and  $t_0$  large enough such that

$$\sup_{t \geq t_0} \left( \prod_{j=t_0}^t \pi_j \right) z_0 \leq 1.$$

This ensures that we can iterate the inequality (57), so obtaining

$$(59) \quad \begin{aligned} E[e^{z_0 \bar{T}_{t+1}^*}] &\leq \exp \left( 2z_0 \sum_{n=t_0}^t \lambda_n u_n \prod_{j=n+1}^t \pi_j \right) E[e^{(\prod_{j=t_0}^t \pi_j) z_0 \bar{T}_{t_0}^*}] \\ &\leq \exp \left( 2z_0 \sum_{n=t_0}^t \lambda_n u_n \prod_{j=n+1}^t \pi_j \right) E[e^{\bar{T}_{t_0}^*}] \end{aligned}$$

for each  $t \geq t_0$ . Then, the second term in the right-hand product in (59) is always bounded by Step 2) and so we can now focus on the first term, which by (58) becomes less or equal than

$$\exp \left( 2z_0 C u_t \tilde{\zeta}_t(w) \sum_{n=t_0}^t \frac{\lambda_n}{\tilde{\zeta}_n(w)} \right) \sim \exp \left( 2z_0 C \frac{r_t}{t^{1-w}} \sum_{n=t_0}^t \frac{\alpha}{n^{1+w-\beta}} \right),$$

that is finite in all cases as

- (a) in the case  $w > \beta$  we have  $\frac{r_t}{t^{1-w}} \sim 1$  and  $\sum_{n=t_0}^t \frac{\alpha}{n^{1+w-\beta}} < +\infty$ ;
- (b) in the case  $w = \beta$  we have  $\frac{r_t}{t^{1-w}} \sim 1/\ln(t)$  and  $\sum_{n=t_0}^t \frac{\alpha}{n^{1+w-\beta}} = O(\ln(t))$ ;
- (c) in the case  $w < \beta$  we have  $\frac{r_t}{t^{1-w}} \sim 1/t^{\beta-w}$  and  $\sum_{n=t_0}^t \frac{\alpha}{n^{1+w-\beta}} = O(t^{\beta-w})$ .

Therefore from (59) we can conclude that, for any  $z \leq z_0$ ,  $\sup_{t \geq 1} E[e^{z \bar{T}_t^*}] \leq \sup_{t \geq 1} E[e^{z_0 \bar{T}_t^*}] < +\infty$ .

Finally, since by (4) we have  $r_t Z_t = w \frac{t}{t+\theta} \bar{T}_t^* + r_t \lambda_t \leq \bar{T}_t^* + c$  for some suitable  $c > 0$ , we also obtain  $\sup_{t \geq 1} E[e^{z r_t Z_t}] < +\infty$  for any  $z \leq z_0$ .

The  $L^p$ -boundness of the sequences  $(r_t Z_t)_t$  and  $(r_t \bar{T}_t)_t$  follows immediately, as for non-negative random variables we have

$$\begin{aligned} \sup_t E(e^{z_0 r_t Z_t}) < \infty &\iff \sup_t \sum_{p \geq 0} \frac{z_0^p E[(r_t Z_t)^p]}{p!} < +\infty \\ &\implies \sup_t E[(r_t Z_t)^p] < +\infty \quad \forall p \geq 1. \end{aligned}$$

and the same for  $(r_t \bar{T}_t)_t$ . For what concerns  $(T_t)_t$ , from (54), (4), (7) and the fact that  $g(z) \leq 2z$  when  $z \leq 1/2$ , we obtain for any  $t \geq 0$  and  $z \leq 1/2$

$$(60) \quad E[e^{z T_{t+1}} | \mathcal{F}_t] \leq \exp(g(z) (\lambda_t + S_t)) = \exp(g(z) Z_t) \leq \exp(2z Z_t).$$

Now, note that  $r_t \equiv 1$  when  $w < \beta = 1$  or  $\beta < w = 1$ , and in this case the two expressions in (53) coincide. More generally, we can consider the following cases.

*Case  $w = 1$  or  $\beta = 1$ :* we have  $r_t \leq 1$  and so, by (60), for each  $t \geq 0$ , we have

$$E[e^{\frac{z_0}{2} r_t T_{t+1}}] \leq E[e^{z_0 r_t Z_t}].$$

Then, by (52), we get  $\sup_t E[e^{\frac{z_0}{2} r_t T_{t+1}}] < +\infty$ , that implies as above  $\sup_t E[(r_t T_{t+1})^p] < +\infty$  for all  $p \geq 1$ .

*Case  $w < 1$  or  $\beta < 1$ :* we have  $r_t \geq 1$  for  $t$  large enough, say  $t \geq \bar{t}$ . Hence, by (60), for  $t \geq \bar{t}$ , we can write

$$\begin{aligned}
r_t(E[e^{\frac{z_0}{2} T_{t+1}}] - 1) &\leq r_t(E[e^{z_0 Z_t}] - 1) = r_t(E[e^{z_0 \frac{Z_t^*}{r_t}}] - 1) \\
&= \sum_{p \geq 1} \frac{z_0^p E[(Z_t^*)^p]}{(r_t)^{p-1} p!} \leq \sum_{p \geq 1} \frac{z_0^p E[(Z_t^*)^p]}{p!} = (E[e^{z_0 Z_t^*}] - 1).
\end{aligned}$$

Hence, we can conclude that  $\sup_{t \geq \bar{t}} r_t(E[e^{\frac{z_0}{2} T_{t+1}}] - 1) < +\infty$ , that reads

$$\sup_{t \geq \bar{t}} \sum_{p \geq 1} \frac{(z_0)^p E[r_t(T_{t+1})^p]}{2^p p!} < +\infty,$$

and so we can get  $\sup_{t \geq \bar{t}} E[r_t(T_{t+1})^p] < +\infty$ , for all  $p \geq 1$ . Since, by Step 2), for all  $p \geq 1$  and each fixed  $t$  we have  $E[(T_{t+1})^p] < +\infty$ , and so  $E[r_t(T_{t+1})^p] < +\infty$ , we can also affirm that  $\sup_t E[r_t(T_{t+1})^p] < +\infty$  for all  $p \geq 1$ .  $\square$

**Lemma D.2.** *We have*

$$r_t R_t \xrightarrow{a.s.} \begin{cases} \tilde{R}_\infty & \text{if } \iota = 0, w = 1 \text{ and } \beta < 1; \\ 0 & \text{otherwise.} \end{cases}$$

where  $\tilde{R}_\infty$  is a real random variable such that  $0 < \tilde{R}_\infty < Z_\infty^*$ .

*Proof.* From (6), since  $D_t \bar{P}_t^2 = \bar{P}_t S_t$  we get

$$\begin{aligned}
R_t &= \sum_{j \in \mathcal{O}_t} P_{t,j}^2 = \sum_{j \in \mathcal{O}_t} (w(1-\iota) \frac{K_{t,j}}{(\theta+t)} + \iota \bar{P}_t)^2 \\
&= (w(1-\iota))^2 \tilde{R}_t + \iota^2 D_t \bar{P}_t^2 + 2\iota(1-\iota) \bar{P}_t S_t \\
&= (w(1-\iota))^2 \tilde{R}_t + \iota(2-\iota) \bar{P}_t S_t,
\end{aligned}$$

where, by (4),

$$0 \leq \tilde{R}_t = \frac{\sum_{j \in \mathcal{O}_t} K_{t,j}^2}{(\theta+t)^2} \leq \left( \frac{\sum_{j \in \mathcal{O}_t} K_{t,j}}{\theta+t} \right)^2 = S_t^2/w^2.$$

Since  $\bar{P}_t \rightarrow 0$  a.s. and  $r_t S_t = O(1)$  from (4) and Table 1, then  $r_t R_t \rightarrow 0$  a.s. when  $\iota = 1$ , otherwise if and only if  $r_t \tilde{R}_t \rightarrow 0$  a.s. and this is always the case when  $\max\{w, \beta\} < 1$ , because  $r_t \tilde{R}_t \leq r_t S_t^2/w^2 \rightarrow 0$  a.s..

From now on, we assume  $\iota < 1$  and  $\max\{w, \beta\} = 1$  (and so  $r_t = 1$  when  $\beta \neq w$  and  $r_t \rightarrow 0$  when  $\beta = w$ ) and we are going to prove that  $r_t \tilde{R}_t \rightarrow 0$  a.s. unless  $\iota = 0, w = 1$  and  $\beta < 1$ . We observe that

$$\tilde{R}_{t+1} = \sum_{j \in \mathcal{O}_t} \left( \frac{K_{t+1,j}}{\theta+t+1} \right)^2 + \sum_{j \in \mathcal{O}_{t+1} \setminus \mathcal{O}_t} \left( \frac{1}{\theta+t} \right)^2 = \sum_{j \in \mathcal{O}_t} \left( \frac{K_{t,j} + \Delta K_{t+1,j}}{\theta+t+1} \right)^2 + \frac{N_{t+1}}{(\theta+t+1)^2}$$

and so, since  $(\Delta K_{t+1,j})^2 = \Delta K_{t+1,j}$ , we have

$$(61) \quad E[\tilde{R}_{t+1} \mid \mathcal{F}_t] = \sum_{j \in \mathcal{O}_t} \frac{K_{t,j}^2 + P_{t,j} + 2K_{t,j} P_{t,j}}{(\theta+t+1)^2} + \frac{\lambda_t}{(\theta+t+1)^2},$$

while, the martingale difference reads

$$\begin{aligned}
\tilde{R}_{t+1} - E[\tilde{R}_{t+1} \mid \mathcal{F}_t] &= \frac{1}{\theta+t+1} \left( \sum_{j \in \mathcal{O}_t} (\Delta K_{t+1,j} - P_{t,j}) \frac{1+2K_{t,j}}{\theta+t+1} + \frac{N_{t+1} - \lambda_t}{(\theta+t+1)} \right) = \frac{\Delta M_{t+1}^{(\tilde{R})}}{\theta+t+1}.
\end{aligned}$$

Since the random variables  $(\Delta K_{t+1,j})_{j \in \mathcal{O}_t}$  are, conditionally on  $\mathcal{F}_t$ , independent among each other, and  $N_{t+1}$  is independent of them and  $\mathcal{F}_t$  (see the description of the model dynamics), we have

$$(62) \quad E[(\Delta M_{t+1}^{(\tilde{R})})^2 \mid \mathcal{F}_t] = \sum_{j \in \mathcal{O}_t} P_{t,j} (1 - P_{t,j}) \left( \frac{1+2K_{t,j}}{\theta+t+1} \right)^2 + \frac{\lambda_t}{(\theta+t+1)^2}.$$

Now, we go on with the computations for  $E[\tilde{R}_{t+1} \mid \mathcal{F}_t]$ . From (61), recalling that  $\tilde{R}_t = \sum_{j \in \mathcal{O}_t} K_{t,j}^2/(\theta+t)^2$  and using (4) and (6), we get

$$E[\tilde{R}_{t+1} \mid \mathcal{F}_t] = \tilde{R}_t + \sum_{j \in \mathcal{O}_t} K_{t,j}^2 \left( \frac{1}{(\theta+t+1)^2} - \frac{1}{(\theta+t)^2} \right) + \frac{S_t}{(\theta+t+1)^2}$$

$$\begin{aligned}
& + \frac{2}{(\theta+t+1)^2} \sum_{j \in \mathcal{O}_t} K_{t,j} \left( (1-\iota)w \frac{K_{t,j}}{\theta+t} + \iota \bar{P}_t \right) + \frac{\lambda_t}{(\theta+t+1)^2} \\
& = \tilde{R}_t + \sum_{j \in \mathcal{O}_t} K_{t,j}^2 \left( \frac{1}{(\theta+t+1)^2} - \frac{1}{(\theta+t)^2} + \frac{2(1-\iota)w}{(\theta+t)(\theta+t+1)^2} \right) \\
& \quad + \frac{S_t}{(\theta+t+1)^2} + \frac{2\iota}{(\theta+t+1)^2} \bar{P}_t \sum_{j \in \mathcal{O}_t} K_{t,j} + \frac{\lambda_t}{(\theta+t+1)^2} \\
& = \tilde{R}_t - \frac{2(\iota+1-w)}{\theta+t+1} \tilde{R}_t + \frac{2\iota}{(\theta+t+1)^2} \bar{P}_t \sum_{j \in \mathcal{O}_t} K_{t,j} + \rho_t.
\end{aligned}$$

with  $\rho_t$  being  $\mathcal{F}_t$ -measurable and such that  $\sum_t |\rho_t| < +\infty$  since  $\rho_t = O((\tilde{R}_t + S_t + 1)/t^2) = O((S_t + 1)/t^2)$ . Using (5), we finally obtain

$$E[\tilde{R}_{t+1} \mid \mathcal{F}_t] = \tilde{R}_t - \frac{2(\iota+1-w)}{\theta+t+1} R_t^* + \frac{2\iota(\theta+t)D_t}{(\theta+t+1)^2 w} \bar{P}_t^2 + \rho_t$$

and hence

$$(63) \quad \tilde{R}_{t+1} = \tilde{R}_t + \frac{1}{\theta+t+1} \left( \frac{2\iota(\theta+t)D_t \bar{P}_t^2}{(\theta+t+1)w} - 2(\iota+1-w)\tilde{R}_t \right) + \frac{\Delta M_{t+1}^{(\tilde{R})}}{\theta+t+1} + \rho_t.$$

Now, we consider two cases.

*Case  $\iota > 0$  or  $w < 1$  (and  $\beta = 1$ ):* in this case we can apply Theorem B.1 to (63) with  $X_t = \tilde{R}_t$ ,  $A_t = \frac{2\iota(\theta+t)D_t \bar{P}_t^2}{(\theta+t+1)w}$ ,  $b = 2(\iota+1-w) > 0$ ,  $\eta_t = (\theta+t+1)^{-1}$  and  $\Delta M_{t+1} = \Delta M_{t+1}^{(\tilde{R})}$  (and in the present case the reminder term  $\rho_{t+1}$  in Theorem B.1 coincides here with  $\rho_t$  and so it is  $\mathcal{F}_t$ -measurable). Indeed, by (62), we have

$$\begin{aligned}
& \sum_t \eta_t^2 E[(\Delta M_{t+1})^2 \mid \mathcal{F}_t] \\
& = \sum_t \frac{1}{(\theta+t+1)^2} \left( \sum_{j \in \mathcal{O}_t} P_{t,j} (1 - P_{t,j}) \left( \frac{1+2K_{t,j}}{\theta+t+1} \right)^2 + \frac{\lambda_t}{(\theta+t+1)^2} \right) \\
& \leq \sum_t \frac{1}{(\theta+t+1)^2} \left( \sum_{j \in \mathcal{O}_t} P_{t,j} \left( \frac{1+2t}{\theta+t+1} \right)^2 + \frac{\lambda_t}{(\theta+t+1)^2} \right) \\
& \leq \sum_t \left( \frac{4S_t}{(\theta+t+1)^2} + \frac{\lambda_t}{(\theta+t+1)^4} \right) < +\infty.
\end{aligned}$$

In addition, we have

$$A_t = \frac{2\iota}{w} \frac{(\theta+t)}{(\theta+t+1)} D_t \bar{P}_t^2 \xrightarrow{a.s.} 0.$$

Hence, by Theorem B.1, we obtain  $\tilde{R}_t \rightarrow 0$  a.s. and, since in this case  $r_t = 1$ , this means  $r_t \tilde{R}_t \rightarrow 0$  a.s..

*Case  $\iota = 0$  and  $w = 1$  (and  $\beta < 1$  or  $\beta = 1$ ):* in this case  $R_t = \tilde{R}_t$  and  $\tilde{R}_t$  is a non-negative almost super-martingale converging a.s. to a non-negative real random variable  $\tilde{R}_\infty$  (see Theorem E.4). Then, when  $w = \beta = 1$ , since  $r_t \rightarrow 0$  in this case, we trivially get  $r_t R_t = r_t \tilde{R}_t \rightarrow 0$  a.s..

Let us focus on the remaining sub-case, that is when  $\iota = 0$ ,  $w = 1$  and  $\beta < 1$  (and so  $r_t = 1$ ). First, since  $P_{t,j}^2 \leq P_{t,j}$ , we have  $\tilde{R}_t = R_t \leq S_t = (Z_t - \lambda_t) \rightarrow Z_\infty^*$  and so we obtain  $0 \leq \tilde{R}_\infty \leq Z_\infty^*$ . Now, let  $\tau_1 = \inf\{t: N_t > 0\}$  and note that in this case (1) reads

$$P_{\tau_1+n}(1) = \frac{K_{\tau_1+n}}{\tau_1+n+\theta},$$

that evolves, conditionally on the value taken by  $\tau_1$  (note that  $\tau_1 < +\infty$  a.s.), as the proportion of black balls in a Pólya urn that starts with 1 black ball and  $(\tau_1 + \theta)$  white balls. Then, conditionally on the value taken by  $\tau_1$ , we have  $P_{\tau_1+n}(1) \xrightarrow{a.s.} P_\infty(1) \stackrel{d}{=} \text{Beta}(1, \tau_1 + \theta)$ , so that we have  $P_\infty(1)(1 - P_\infty(1)) > 0$ . Note that, since  $\beta < w = 1$  and  $\iota = 0$ , we have

$$\tilde{R}_\infty = P_\infty(1)^2 + \text{a.s.-}\lim_t \sum_{j>1, j \in \mathcal{O}_t} P_{t,j}^2$$

and

$$Z_\infty^* = S_\infty^* = P_\infty(1) + \text{a.s.-}\lim_t \sum_{j>1, j \in \mathcal{O}_t} P_{t,j}$$



and hence

$$0 < P_\infty(1)^2 \leq \tilde{R}_\infty < \tilde{R}_\infty + P_\infty(1)(1 - P_\infty(1)) \leq Z_\infty^*.$$

Hence, we can conclude because in this last sub-case  $r_t = 1$  and so  $r_t \tilde{R}_t = \tilde{R}_t$ .  $\square$

**Lemma D.3.** *When  $0 \leq w\iota < \beta$  we have*

$$(64) \quad \sup_{t \geq 1} E[t^{1-w(1-\iota)} \Delta K_{t+1,j}] < +\infty.$$

*Proof.* First, recall from Section 3.1 that  $D_t = \sum_{n=1}^t N_n \stackrel{d}{=} \text{Poi}(\Lambda_t)$  with  $\Lambda_t = \sum_{n=1}^t \lambda_{n-1} \sim (\alpha/\beta)t^\beta$  when  $\beta > 0$ . Then, consider the sequence  $(r_t)$  as defined in (29) with  $\delta = w$  e  $\rho = \beta$ . Since  $E[\Delta K_{t+1,j}] = w(1-\iota)E[B_{t,j}] + \iota E[\bar{P}_t]$ , the proof will be divided in the proof of the following two facts:

- (1)  $E[t^{1-w+w\iota}\bar{P}_t] = O(t^{-(\min\{\beta,w\}-w\iota)} \ln(t)) \rightarrow 0$ ,
- (2)  $\sup_{t \geq 1} t^{1-w+w\iota} E[B_{t,j}] < +\infty$ .

*Proof of (1):* Consider

$$E[t^{1-w+w\iota}\bar{P}_t] = \frac{t^{1-w}}{r_t} E\left[\frac{r_t S_t}{t^{-w\iota} D_t}\right].$$

Since  $D_t \stackrel{d}{=} \text{Poi}(\Lambda_t)$ , by Chernoff's bound (see [36]), we have

$$P(D_t < (1-\epsilon)\Lambda_t) \leq e^{-\frac{\epsilon^2}{2}\Lambda_t}$$

and so, using also Holder's inequality, we get

$$\begin{aligned} E\left[\frac{r_t S_t}{t^{-w\iota} D_t}\right] &= E\left[\frac{r_t S_t}{t^{-w\iota} D_t} \mathbb{1}_{\{D_t \geq (1-\epsilon)\Lambda_t\}}\right] + E\left[\frac{r_t S_t}{t^{-w\iota} D_t} \mathbb{1}_{\{D_t < (1-\epsilon)\Lambda_t\}}\right] \\ &\leq E\left[\frac{r_t S_t}{t^{-w\iota}(1-\epsilon)\Lambda_t}\right] + E[(r_t S_t)^2]^{\frac{1}{2}} t^{w\iota} P(D_t < (1-\epsilon)\Lambda_t)^{\frac{1}{2}} \\ &\leq E[r_t S_t] \frac{t^{w\iota}}{(1-\epsilon)\Lambda_t} + E[(r_t S_t)^2]^{\frac{1}{2}} t^{w\iota} e^{-\frac{\epsilon^2}{4}\Lambda_t}. \end{aligned}$$

Now, we observe that, by Lemma D.1 and the fact that  $0 \leq S_t = Z_t - \lambda_t \leq Z_t$ , we have  $\sup_t E[(r_t S_t)^p] \leq \sup_t E[(r_t Z_t)^p] < +\infty$  for  $p = 1, 2$  and so the above inequality implies that

$$E[t^{1-w+w\iota}\bar{P}_t] = \frac{t^{1-w}}{r_t} O\left(\frac{t^{w\iota}}{\Lambda_t}\right) = O(t^{-(\min\{\beta,w\}-w\iota)} \ln(t)) \rightarrow 0.$$

*Proof of (2):* Setting  $m_t = E[B_t]$ , by (16) we have

$$m_{t+1} = \left(1 - \frac{1}{\theta+t+1}\right)m_t + \frac{1}{\theta+t+1} E[P_{t,j}] = \left(1 - \frac{1-w(1-\iota)}{\theta+t+1}\right)m_t + \frac{\iota}{\theta+t+1} E[\bar{P}_t].$$

Then, setting  $m_t^* = \zeta_t(1-w+w\iota)m_t$  and using (25), we get

$$\begin{aligned} m_{t+1}^* &= \left(1 - \frac{(1-w(1-\iota))^2}{(\theta+t+1)^2}\right)m_t^* + \frac{\iota}{\theta+t+1} E[t^{1-w+w\iota}\bar{P}_t] \\ &\leq m_t^* + \frac{\iota}{\theta+t+1} E[t^{1-w+w\iota}\bar{P}_t], \end{aligned}$$

which implies that  $(m_t^*)_t$  is bounded because, by fact (1), we have

$$\sum_t E[t^{1-w+w\iota}\bar{P}_t]/(\theta+t+1) = \sum_t O(1/t^{1+\min\{\beta,w\}-w\iota}) < +\infty.$$

$\square$

**D.1. Proof of the law of the iterated logarithm for  $D_t$ .** We here provide the proof of the result for  $D_t$  given in Remark 3.2.

*Proof.* Recall that  $D_t - \Lambda_t = \sum_{n=1}^t (N_n - \lambda_{n-1})$ , where  $(N_n - \lambda_{n-1})_n$  is a sequence of zero-mean independent random variables and  $\text{Var}[D_t] = \sum_{n=1}^t \text{Var}[N_n] = \sum_{n=1}^t \lambda_{n-1} = \Lambda_t$ . Set  $c_n = \sqrt{\frac{\Lambda_n}{\ln(\Lambda_n)}}$ . For  $n$  large enough, we

have  $\lambda_{n-1} < c_n$  and consequently  $P(N_n - \lambda_{n-1} \leq -c_n) = P(N_n \leq \lambda_{n-1} - c_n) = 0$ . Then, by the bounds for the tail probabilities of a Poisson random variable (e.g. [35, Theorem 5.4]), we obtain for  $n$  large enough

$$\begin{aligned} P(|N_n - \lambda_{n-1}| \geq c_n) &= P(N_n \geq \lambda_{n-1} + c_n) \leq e^{-\lambda_{n-1}} \frac{(e\lambda_{n-1})^{\lambda_{n-1}+c_n}}{(\lambda_{n-1}+c_n)^{\lambda_{n-1}+c_n}} \\ &= \left(\frac{e\lambda_{n-1}}{c_n}\right)^{c_n} \left(\frac{\lambda_{n-1}}{c_n}\right)^{\lambda_{n-1}} \frac{1}{\left(\frac{\lambda_{n-1}}{c_n}+1\right)^{c_n} \left(\frac{\lambda_{n-1}}{c_n}+1\right)^{\lambda_{n-1}}} \\ &\leq C \left(\frac{e\lambda_{n-1}}{c_n}\right)^{c_n}, \end{aligned}$$

where the last inequality is due to the fact that  $\lambda_{n-1} = o(c_n)$ . Moreover, we have  $\sum_n \left(\frac{e\lambda_{n-1}}{c_n}\right)^{c_n} < +\infty$ . In fact, for  $\beta = 0$ , we have  $\lambda_{n-1} = \alpha/n$  and hence  $\left(\frac{e\lambda_{n-1}}{c_n}\right)^{c_n} = \left(\frac{e\alpha}{nc_n}\right)^{c_n} = o(n^{-2})$  and for  $\beta > 0$ , we have  $\lambda_{n-1} = \alpha/n^{1-\beta}$  and so  $c_n \geq n^{\beta/(2+\epsilon)}$  eventually, which, together with  $\lambda_{n-1}/c_n \rightarrow 0$ , implies  $\left(\frac{e\lambda_{n-1}}{c_n}\right)^{c_n} \leq 2^{-n^{\beta/(2+\epsilon)}}$  eventually. By Borel-Cantelli Lemma, we finally obtain

$$P\left(\limsup_n \{|N_n - \lambda_{n-1}| \geq c_n\}\right) = 0$$

and so

$$|N_n - \lambda_{n-1}| = O(c_n) = o\left(\sqrt{\frac{\Lambda_n}{\ln \ln \Lambda_n}}\right) \quad \text{a.s.}$$

Therefore, (3.2) is a consequence of Kolmogorov's law of the iterated logarithm [21] applied to the sequence  $(N_n - \lambda_{n-1})_n$  (see [11, Theorem 1, p. 343]).  $\square$

#### APPENDIX E. USEFUL RECALLS

For the reader's convenience, we here recall some technical results for discrete-time real stochastic processes, used in the previous proofs.

**Lemma E.1** ([17, Theorem 46, p. 40]). *Let  $\mathcal{F} = (\mathcal{F}_t)_t$  be a filtration and  $(Y_t)$  be an  $\mathcal{F}$ -adapted non-negative stochastic process.*

*Then the set  $\{\sum_t E[Y_{t+1}|\mathcal{F}_t] < +\infty\}$  is a.s. contained in the set  $\{\sum_t Y_t < +\infty\}$ . If, in addition, the random variables  $Y_t$  are uniformly bounded by a constant, then these two sets are a.s. equal.*

**Lemma E.2** (Lévy's extension of Borel-Cantelli Lemma [49, Sec. 12.15]). *Let  $(Y_t)_t$  be a sequence of Bernoulli random variables, adapted to a filtration  $\mathcal{F} = (\mathcal{F}_t)_t$  and such that  $Z_t = P(Y_{t+1} = 1 | \mathcal{F}_t)$  and  $\sum_t Z_t = +\infty$  a.s.. Then  $\sum_{n=0}^t Y_{n+1} / \sum_{n=0}^t Z_n \xrightarrow{\text{a.s.}} 1$ .*

**Lemma E.3** ([38, Lemma 3.2]). *Let  $\mathcal{F} = (\mathcal{F}_t)_t$  be a filtration and let  $(L_t)_t$  be an  $\mathcal{F}$ -adapted stochastic process. Set  $U_t = E[L_{t+1} - L_t | \mathcal{F}_t]$  and  $V_t = E[(L_{t+1} - L_t)^2 | \mathcal{F}_t]$ . If  $\sum_t U_t$  and  $\sum_t V_t$  are a.s. convergent, then  $(L_t)_t$  converges a.s. to a real random variable.*

**Theorem E.4** ([42, Non-negative almost super-martingale]). *Let  $(Y_t)_t$  be a non-negative real almost super-martingale with respect to a filtration  $\mathcal{F} = (\mathcal{F}_t)_t$ , that is an  $\mathcal{F}$ -adapted non-negative real stochastic process satisfying*

$$E[Y_{t+1}|\mathcal{F}_t] \leq (1 + \Delta_t)Y_t + \rho_{1,t} - \rho_{2,t},$$

*where  $\Delta_t, \rho_{1,t}, \rho_{2,t}$  are all  $\mathcal{F}$ -adapted non-negative real stochastic processes. Then  $(Y_t)_t$  a.s. converges to a real random variable and  $\sum_t \rho_{2,t} < +\infty$  a.s. on the event  $\{\sum_t \Delta_t < +\infty, \sum_t \rho_{1,t} < +\infty\}$ .*

**Lemma E.5** ([14, Lemma 4.1]). *Let  $\mathcal{F}$  be a filtration and  $(Y_t)$  be an  $\mathcal{F}$ -adapted stochastic process such that each  $Y_t$  is square-integrable and  $E[Y_{t+1}|\mathcal{F}_t] \rightarrow Y_\infty$  a.s. for some real random variable  $Y_\infty$ . Moreover, let  $(a_t)$  and  $(b_t)$  be two sequences of strictly positive real numbers such that*

$$b_t \uparrow +\infty, \quad \sum_{t=1}^{\infty} \frac{E[Y_t^2]}{a_t^2 b_t^2} < +\infty.$$

*Then we have:*

- a) *if  $\frac{1}{b_t} \sum_{n=1}^t \frac{1}{a_n} \rightarrow \gamma$  for some constant  $\gamma$ , then  $\frac{1}{b_t} \sum_{n=1}^t \frac{Y_n}{a_n} \xrightarrow{\text{a.s.}} \gamma Y_\infty$ ;*
- b) *If  $b_t \sum_{n \geq t} \frac{1}{a_n b_n^2} \rightarrow \gamma$  for some constant  $\gamma$ , then  $b_t \sum_{n \geq t} \frac{Y_n}{a_n b_n^2} \xrightarrow{\text{a.s.}} \gamma Y_\infty$ .*

**Theorem E.6** ([12, Theorem A.1]). Let  $(\mathcal{F}_t)_{t \geq 0}$  be a complete filtration and, for each  $n \geq 1$ , let  $(M_{n,t})_{t \geq 0}$  be a martingale with respect to the filtration  $(\mathcal{F}_{n+(t-1)})_{t \geq 0}$ , with  $M_{n,0} = 0$ , and converging in  $L^1$  to a real  $(\mathcal{F}_\infty$ -measurable) random variable  $M_{n,\infty}$ . Moreover, set

$$X_{n,t} = M_{n,t} - M_{n,t-1} \quad \text{for } t \geq 1, \quad U_n = \sum_{t \geq 1} X_{n,t}^2, \quad X_n^* = \sup_{t \geq 1} |X_{n,t}|$$

and assume that the following conditions are satisfied:

- i)  $X_n^* \xrightarrow{a.s.} 0$ ;
- ii)  $(X_n^*)_n$  is dominated in  $L^1$ ;
- iii)  $(U_n)_n \xrightarrow{a.s.} U$  for a real  $(\mathcal{F}_\infty$ -measurable) random variable  $U$ .

Then, the sequence  $(M_{n,\infty})_n$  converges to the Gaussian kernel  $\mathcal{N}(0, U)$  in the sense of the a.s. conditional convergence with respect to  $(\mathcal{F}_n)_n$  (and so also stably), that is the a.s. convergence of the conditional distributions holds true: for almost every  $\omega$ , the conditional distribution  $\mu_n(\omega, \cdot)$  of  $M_{n,\infty}$  given  $\mathcal{F}_n$  converges weakly to the Gaussian distribution  $\mathcal{N}(0, U(\omega))$ .

The a.s. conditional convergence is a variant of the stable convergence. For more information on stable convergence and its variants, we refer the reader to [2, 12, 13, 15, 28, 41]. An useful summary can be found in the appendix of [3]. We here provide only a lemma, regarding the fact that the a.s. conditional convergence combines well with the a.s. convergence, that we need in the proof of Theorem B.4.

**Lemma E.7.** Let  $\mathcal{F} = (\mathcal{F}_t)_t$  be a filtration,  $K$  an  $\mathcal{F}_\infty$ -measurable kernel and  $Y$  an  $\mathcal{F}_\infty$ -measurable real random variable. If  $X_t \rightarrow K$  in the sense of the a.s. conditional convergence with respect to a filtration  $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$  and  $Y_t \xrightarrow{a.s.} Y$ , then

$$[X_t, Y_t] \rightarrow K \otimes \delta_Y$$

in the sense of the a.s. conditional convergence with respect to  $\mathcal{F}$ .

In particular, when  $K$  is a Gaussian kernel  $\mathcal{N}(M, U)$ , we have

$$X_t + Y_t \rightarrow \mathcal{N}(M + Y, U).$$

in the sense of the a.s. conditional convergence with respect to  $\mathcal{F}$ .

*Proof.* Since Cramér-Wold device and [12, Lemma A.3], it is enough to prove that, for each  $a_1, a_2 \in \mathbb{R}$ , we have

$$E[e^{i(a_1 X_t + a_2 Y_t)} \mid \mathcal{F}_t] \xrightarrow{a.s.} \left( \int e^{ia_1 x} K(\cdot)(dx) \right) e^{ia_2 Y}.$$

We first prove the above convergence under the additional assumption that  $Y$  is integrable and then we will show how to get the result in the general case. Suppose  $Y$  to be integrable and set  $Z_t = E[Y \mid \mathcal{F}_t]$  so that we have  $Z_t \xrightarrow{a.s.} Y$  (by the martingale convergence theorem for martingales bounded in  $L^1$ ). Moreover, we have

$$\begin{aligned} |D_t| &= |E[e^{i(a_1 X_t + a_2 Y_t)} \mid \mathcal{F}_t] - \left( \int e^{ia_1 x} K(\cdot)(dx) \right) e^{ia_2 Y}| \\ &\leq |T_{1,t}| + |T_{2,t}|, \end{aligned}$$

where

- $|T_{1,t}| = |E[e^{ia_1 X_t}(e^{ia_2 Y_t} - e^{ia_2 Z_t}) \mid \mathcal{F}_t]|$ , which is smaller than or equal to  $E[|e^{ia_2 Y_t} - e^{ia_2 Z_t}| \mid \mathcal{F}_t]$ , that converges a.s. to zero, because of [12, Lemma A.2(d)] (see also [7, Th. 2]) and the a.s. convergence of both  $Y_t$  and  $Z_t$  to the same limit  $Y$ ;
- $|T_{2,t}| = |E[e^{i(a_1 X_t + a_2 Z_t)} \mid \mathcal{F}_t] - \left( \int e^{ia_1 x} K(\cdot)(dx) \right) e^{ia_2 Y}|$ , which is equal to the difference

$$|e^{ia_2 Z_t} E[e^{ia_1 X_t} \mid \mathcal{F}_t] - \left( \int e^{ia_1 x} K(\cdot)(dx) \right) e^{ia_2 Y}|$$

and converges a.s. to zero by the assumption on  $X_t$  and the a.s. converge of  $Z_t$  to  $Y$ .

Therefore, we can conclude that  $D_t \xrightarrow{a.s.} 0$  and so the result is proven when  $Y$  is integrable.

If  $Y$  is not integrable, then, for any  $N \in \mathbb{N}$ , we can define  $Y_t^{(N)} = (-N) \vee Y_t \wedge N$  and  $Y^{(N)} = (-N) \vee Y \wedge N$  so that  $|Y_t^{(N)}| \leq N$ ,  $|Y^{(N)}| \leq N$ ,  $Y_t^{(N)} \xrightarrow{a.s.} Y^{(N)}$  as  $t \rightarrow +\infty$ ,  $Y_t^{(N)} \xrightarrow{a.s.} Y_t$  and  $Y^{(N)} \xrightarrow{a.s.} Y$  as  $N \rightarrow +\infty$ . For each fixed  $N$ , applying the result to  $X_t \rightarrow K$  and  $Y_t^{(N)} \xrightarrow{a.s.} Y^{(N)}$  (which is trivially integrable), we find for  $t \rightarrow +\infty$

$$D_t^{(N)} = E[e^{i(a_1 X_t + a_2 Y_t^{(N)})} \mid \mathcal{F}_t] - \left( \int e^{ia_1 x} K(\cdot)(dx) \right) e^{ia_2 Y^{(N)}} \xrightarrow{a.s.} 0.$$

Moreover, since [12, Lemma A.2(d)] and the a.s. convergence of the difference  $|e^{ia_2 Y_t} - e^{ia_2 Y_t^{(N)}}|$  to  $|e^{ia_2 Y} - e^{ia_2 Y^{(N)}}|$  for  $t \rightarrow +\infty$  (and each fixed  $N$ ), we have  $E[|e^{ia_2 Y_t} - e^{ia_2 Y_t^{(N)}}| | \mathcal{F}_t] \rightarrow |e^{ia_2 Y} - e^{ia_2 Y^{(N)}}|$  for  $t \rightarrow +\infty$  (and each fixed  $N$ ) and we get

$$\limsup_t |D_t - D_t^{(N)}| \leq 2|e^{ia_2 Y} - e^{ia_2 Y^{(N)}}| \xrightarrow{a.s.} 0 \quad \text{as } N \rightarrow +\infty.$$

Hence,  $D_t \xrightarrow{a.s.} 0$  and the proof is so concluded.  $\square$

## REFERENCES

- [1] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 1972.
- [2] D. J. Aldous and G. K. Eagleson. On mixing and stability of limit theorems. *Ann. Probab.*, 6(2):325–331, 1978.
- [3] G. Aletti, I. Crimaldi, and A. Ghiglietti. Synchronization of reinforced stochastic processes with a network-based interaction. *Ann. Appl. Probab.*, 27(6):3787–3844, 2017.
- [4] G. Aletti, I. Crimaldi, and A. Ghiglietti. Networks of reinforced stochastic processes: A complete description of the first-order asymptotics. *Stochastic Process. Appl.*, 176:104427, 2024.
- [5] A. Bandyopadhyay and D. Thacker. A new approach to Pólya urn schemes and its infinite color generalization. *Ann. Appl. Probab.*, 32(1):46–79, 2022.
- [6] P. Berti, I. Crimaldi, L. Pratelli, and P. Rigo. Central Limit Theorems for an Indian Buffet Model with Random Weights. *Ann. Appl. Probab.*, 25(2):523–547, 2015.
- [7] D. Blackwell and L. Dubins. Merging of opinions with increasing information. *Ann. Math. Statist.*, 33(3):882–886, 1962.
- [8] P. Boldi, I. Crimaldi, and C. Monti. A network model characterized by a latent attribute structure with competition. *Inf. Sci.*, 354(C):236–256, 2016.
- [9] T. Broderick, M. I. Jordan, and J. Pitman. Beta processes, stick-breaking, and power laws. *Bayesian Anal.*, 7(2):439–476, 2012.
- [10] Federico Camerlenghi, Stefano Favaro, Lorenzo Masoero, and Tamara Broderick. Scaled process priors for bayesian nonparametric estimation of the unseen genetic variation. *Journal of the American Statistical Association*, 119(545):320–331, 2024.
- [11] Yuan Shih Chow and Henry Teicher. *Limit Theorems for Independent Random Variables*, pages 354–403. Springer New York, New York, NY, 1997.
- [12] I. Crimaldi. An almost sure conditional convergence result and an application to a generalized Pólya urn. *Int. Math. Forum*, 4(23):1139–1156, 2009.
- [13] I. Crimaldi. *Introduzione alla nozione di convergenza stabile e sue varianti (Introduction to the notion of stable convergence and its variants)*, volume 57. Unione Matematica Italiana, Monograf s.r.l., Bologna, Italy, 2016. Book written in Italian.
- [14] I. Crimaldi, P. Dai Pra, and I. G. Minelli. Fluctuation theorems for synchronization of interacting Pólya’s urns. *Stochastic Process. Appl.*, 126(3):930–947, 2016.
- [15] I. Crimaldi, G. Letta, and L. Pratelli. *A Strong Form of Stable Convergence*, volume 1899, pages 203–225. Springer, 2007.
- [16] I. Crimaldi, P. Dai Pra, P.-Y. Louis, and I. G. Minelli. Synchronization and functional central limit theorems for interacting reinforced random walks. *Stochastic Process. Appl.*, 129(1):70 – 101, 2019.
- [17] C. Dellacherie and P.-A. Meyer. *Probabilities and potential. B*, volume 72 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1982. Theory of martingales, Translated from the French by J. P. Wilson.
- [18] F. Doshi-Velez and Z. Ghahramani. Correlated non-parametric latent feature models. In *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence (UAI)*, UAI ’09, pages 143–150, Arlington, Virginia, USA, 2009. AUAI Press.
- [19] F. Doshi-Velez and S. A. Williamson. Restricted Indian Buffet Processes. *Stat. Comput.*, 27(5):1205–1223, 2017.
- [20] R. Durrett and S. I. Resnick. Functional limit theorems for dependent variables. *Ann. Probab.*, 6(5):829–846, 1978.
- [21] E. Fisher. On the Law of the Iterated Logarithm for Martingales. *Ann. Probab.*, 20(2):675–680, 1992.
- [22] S. Fortini and S. Petrone. Quasi-bayes properties of a procedure for sequential learning in mixture models. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 82(4):1087–1114, 2020.
- [23] S. J. Gershman, P. I. Frazier, and D. M. Blei. Distance-dependent infinite latent feature models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 37(2):334–345, 2015.
- [24] R. Gouet. Martingale functional central limit theorems for a generalized Pólya urn. *Ann. Probab.*, 21(3):1624 – 1639, 1993.
- [25] T. L. Griffiths and Z. Ghahramani. Infinite latent feature models and the Indian Buffet Process. In *Advances in Neural Information Processing Systems (NIPS)*, volume 18, 2005.
- [26] T. L. Griffiths and Z. Ghahramani. The Indian Buffet Process: An introduction and review. *J. Mach. Learn. Res.*, 12:1185–1224, 2011.
- [27] D. Görür, F. Jäkel, and C. E. Rasmussen. A choice model with infinitely many latent features. In *Proceedings of the International Conference on Machine Learning (ICML)*, volume 23, pages 361–368, 2006.
- [28] P. Hall and C. C. Heyde. *Martingale Limit Theory and Its Application*. Academic Press, New York, 1980.
- [29] C. Heaukulani and D. M. Roy. Gibbs-type Indian Buffet Processes. *Bayesian Anal.*, 15(3):683–710, 2020.
- [30] D. A. Knowles and Z. Ghahramani. Nonparametric Bayesian sparse factor models with application to gene expression modeling. *Ann. Appl. Stat.*, 5(2B):1534–1552, 2011.
- [31] A. Lijoi, R. H. Mena, and I. Prünster. Bayesian nonparametric estimation of the probability of discovering new species. *Biometrika*, 94(4):769–786, 2007.
- [32] A. Lijoi and I. Prünster. *Models beyond the Dirichlet process*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2010.

- [33] K. T. Miller, T. L. Griffiths, and M. I. Jordan. The phylogenetic Indian Buffet Process: A non-exchangeable nonparametric prior for latent features. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 24, pages 403–410, 2008.
- [34] K. T. Miller, T. L. Griffiths, and M. I. Jordan. Nonparametric latent feature models for link prediction. In *Advances in Neural Information Processing Systems (NIPS)*, 2009.
- [35] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.
- [36] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Cambridge University Press, Cambridge, 2nd edition, 2017.
- [37] D. J. Navarro and T. L. Griffiths. Latent features in similarity judgment: A nonparametric bayesian approach. *Neural Comput.*, 20(11):2597–2628, 2008.
- [38] R. Pemantle and S. Volkov. Vertex-reinforced random walk on  $\mathbf{Z}$  has finite range. *Ann. Probab.*, 27(3):1368–1388, 1999.
- [39] J. Pitman. Some developments of the Blackwell-Macqueen urn scheme. *Lecture Notes Monogr. Ser.*, 30:245–267, 1996.
- [40] P. Rai and H. Daume. The infinite hierarchical factor regression model. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems*, volume 21. Curran Associates, Inc., 2008.
- [41] A. Rényi. On stable sequences of events. *Sankhyā Ser. A*, 25:293–302, 1963.
- [42] H. Robbins and D. Siegmund. A convergence theorem for non negative almost supermartingales and some applications. In J. S. Rustagi, editor, *Optimizing Methods in Statistics*, pages 233–257. Academic Press, New York, 1971.
- [43] H. Sariev, S. Fortini, and S. Petrone. Infinite-color randomly reinforced urns with dominant colors. *Bernoulli*, 29(1):132 – 152, 2023.
- [44] P. Sarkar, D. Chakrabarti, and M. I. Jordan. Nonparametric link prediction in dynamic networks. In *Proceedings of the 29th International Conference on Machine Learning (ICML)*, 2012.
- [45] X. Tan, V. Rao, and J. Neville. The Indian Buffet Hawkes Process to model evolving latent influence. In *Proceedings of the 34th Conference on Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI)*, 2018.
- [46] Y. W. Teh and D. Görür. Indian Buffet Processes with power-law behaviour. In *Advances in Neural Information Processing Systems (NIPS)*, volume 22, pages 1838–1846, 2009.
- [47] R. Thibaux and M. I. Jordan. Hierarchical beta processes and the Indian Buffet Process. In *Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 11, pages 564–571, 2007.
- [48] R. L. Warr, D. B. Dahl, J. M. Meyer, and A. Lui. The Attraction Indian Buffet Distribution. *Bayesian Anal.*, 17(3):931 – 967, 2022.
- [49] D. Williams. *Probability with Martingales*. Cambridge University Press, 1991.
- [50] S. Williamson, P. Orbanz, and Z. Ghahramani. Dependent Indian Buffet Processes. In Yee Whye Teh and Mike Titterton, editors, *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 9 of *Proceedings of Machine Learning Research*, pages 924–931. PMLR, 2010.
- [51] F. Wood, T. L. Griffiths, and Z. Ghahramani. A non-parametric bayesian method for inferring hidden causes. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 22, pages 536–543, 2006.
- [52] L.-X. Zhang. A Gaussian process approximation for two-color randomly reinforced urns. *Electron. J. Probab.*, 19:1–19, 2014.
- [53] L.-X. Zhang. Central limit theorems of a recursive stochastic algorithm with applications to adaptive designs. *Ann. Appl. Probab.*, 26(6):3630–3658, 2016.

UNIVERSITÀ DEGLI STUDI DI MILANO, MILAN, ITALY

Email address: [giacomo.aletti@unimi.it](mailto:giacomo.aletti@unimi.it)

IMT SCHOOL FOR ADVANCED STUDIES LUCCA, LUCCA, ITALY

Email address: [irene.crimaldi@imtlucca.it](mailto:irene.crimaldi@imtlucca.it)

UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA, MILAN, ITALY

Email address: [andrea.ghiglietti@unimib.it](mailto:andrea.ghiglietti@unimib.it) (Corresponding author)