

Mixing trichotomy for an Ehrenfest urn with impurities

Matteo Quattropani*

Abstract

We consider a version of the classical Ehrenfest urn model with two urns and two types of balls: regular and heavy. Each ball is selected independently according to a Poisson process having rate 1 for regular balls and rate $\alpha \in (0, 1)$ for heavy balls, and once a ball is selected, is placed in a urn uniformly at random. We study the asymptotic behavior when the total number of balls, N , goes to infinity, and the number of heavy ball is set to $m_N \in \{1, \dots, N - 1\}$. We focus on the observable given by the total number of balls in the left urn, which converges to a binomial distribution of parameter $1/2$, regardless of the choice of the two parameters, α and m_N . We study the speed of convergence and show that this can exhibit three different phenomenologies depending on the choice of the two parameters of the model.

Keywords: mixing time of Markov chains; cutoff; Ehrenfest urn; negative dependence; mixing trichotomy.

MSC2020 subject classifications: Primary 60K35, Secondary 60J27; 82C26.

Submitted to ECP on December 29, 2023, final version accepted on July 11, 2024.

1 Introduction

The *cutoff phenomenon*, first denoted with this name by Aldous and Diaconis [1], is one of the most intriguing and well studied topics in the probabilistic literature of the last 30 years. Indeed, despite the topic received much attention, and the list of Markov processes for which it is possible to show such a behavior is now quite long, a general sufficient condition ensuring the presence of cutoff is still missing. To this aim, it could be beneficial to reach a good understanding of the mixing time of Markov chains that are obtained by perturbing an original chain that is known to exhibit cutoff. How does the mixing behavior of the modified chain depend on the perturbation? To which extent is the cutoff phenomenon robust? In the last few years, similar questions have been investigated in [2, 3, 5, 6]. Despite the models in the aforementioned papers are quite different, they can all be framed into a setting in which the perturbed chain is affected by the competition of two different mixing mechanisms: on the one hand, the abrupt convergence to equilibrium of the original chain, while, on the other hand, a smooth decay to equilibrium arising as an effect of the perturbation. Moreover, the models in [2, 3, 6] share the common feature that the mixing stochastic process under investigation is not a Markov process itself, but rather a non-Markovian observable of a Markov process. A similar investigation, namely the analysis of mixing behavior of

*Dipartimento di Matematica “Guido Castelnuovo”, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185 Roma, Italy. E-mail: matteo.quattropani@uniroma1.it

observables (or *features* or *statistics*) of Markov chains, has been recently performed in [17, 18] in the context of card shuffling routines and related models.

From the phenomenological point of view, the moral of the results in [2, 3, 5, 6] is that there exists a threshold for the perturbation parameter below which the mixing behavior of the process is unaffected. When the parameter exceeds the threshold, then the perturbation becomes dominant and the convergence to equilibrium takes place in a smooth fashion. Finally, and more interestingly, a non trivial mixing behavior is shown to take place at the interphase between these two regimes.

In the same spirit, this short note aims at analyzing the convergence to equilibrium of a two-parameter perturbation of one of the most classical examples of Markov chains exhibiting cutoff: the Ehrenfest urn. Beyond the classical Ehrenfest urn model, [9], the cutoff phenomenon is particularly well understood in the more general setting of birth-and-death chains [7, 8, 10, 11]. Therefore, it seems natural to consider perturbed versions of this classical model and check for the robustness of the cutoff phenomenon with respect to the perturbation. Contrarily to the works [2, 3, 5, 6], the proof of our result does not require a strong technical machinery. Indeed, as it will be explained in Section 3.1, the proof relies on a recently introduced approach to the convergence to equilibrium of spin systems, based on negative dependence, which has been developed by Salez [16] to prove cutoff for the (reversible) Simple Exclusion Process with reservoirs on arbitrary graphs. Such a technique is purely probabilistic, astonishingly simple, and perfectly suited for the analysis of our model.

2 Model and results

We consider the following generalization of the classical Ehrenfest urn model. There are N balls divided in two urns. A number $m = m_N \in \{1, \dots, N - 1\}$ of balls are *heavy*, while the remaining $n = n_N = N - m_N$ balls are *regular*. Each ball is selected at the arrival times of an independent Poisson process having rate 1 if the ball is regular, and rate $\alpha \in (0, 1)$ if the ball is heavy. When a ball is selected, it decides to which urn to move by tossing a fair coin, independently of the rest. Assuming that one can distinguish between regular and heavy balls, but that within the same class the balls are indistinguishable, for any given N, m and α as above, our model can be described as a Markov chain $(R_t, H_t)_{t \geq 0}$ with state space $\mathcal{X} := \{0, 1, \dots, n\} \times \{0, 1, \dots, m\}$, where the state $(a, b) \in \mathcal{X}$ stands for the configuration in which in the left urn there are a regular balls and b heavy balls. It is immediate to check that such a Markov chain is reversible with respect to the stationary distribution:

$$\chi_\star(a, b) := \frac{\binom{n}{a} \binom{m}{b}}{2^n 2^m}, \quad (a, b) \in \mathcal{X}.$$

In what follows, when the the initial configuration is $(r, h) \in \mathcal{X}$ i.e., $\mathbb{P}(R_0 = r, H_0 = h) = 1$, we will denote by $\chi_t^{r,h}$ the distribution of (R_t, H_t) . We are interested in the limit $N \rightarrow \infty$, and we allow α and m to depend on N . In what follows, it will turn out to be useful to define

$$\beta = \beta_N = \frac{\log(m_N)}{\log(N)}. \tag{2.1}$$

As a first step, it is worth pointing out which is the mixing behavior of this Markov chain. To this scope, call

$$\tilde{D}_{\alpha, m}^{N, \text{tv}}(t) := \max_{(r, h) \in \mathcal{X}} \|\chi_t^{r, h} - \chi_\star\|_{\text{tv}} = \frac{1}{2} \max_{(r, h) \in \mathcal{X}} \sum_{(a, b) \in \mathcal{X}} \left| \chi_t^{r, h}(a, b) - \chi_\star(a, b) \right|,$$

and define

$$t_{\text{mix}} := \inf \left\{ t \geq 0 \mid \tilde{\mathcal{D}}_{\alpha, m}^{N, \text{tv}}(t) \leq \frac{1}{4} \right\}, \quad (2.2)$$

where the dependence of t_{mix} on the parameters of the model is suppressed to ease the reading. The next result is not hard to derive, and we postpone its proof to Section 5.

Proposition 2.1. *For all $N \in \mathbb{N}$ fix $\alpha = \alpha_N \in (0, 1)$, $m = m_N \in \{1, \dots, N - 1\}$, and recall the definition of $\beta = \beta_N$ in (2.1). Consider the sequence*

$$\tilde{\gamma} = \tilde{\gamma}_N = \beta_N - \alpha_N, \quad (2.3)$$

and assume that

$$\exists \tilde{\gamma}_\infty = \lim_{N \rightarrow \infty} \tilde{\gamma}_N \in [-1, +1].$$

Depending of the asymptotic behavior of the parameters $\alpha = \alpha_N$ and $\beta = \beta_N$ three different mixing behaviors can take place:

- Insensitivity: if $\tilde{\gamma}_\infty < 0$ there exists $C = C(\varepsilon) > 0$ such that, called

$$t^R = \frac{1}{2} \log(N), \quad (2.4)$$

we have

$$\liminf_{N \rightarrow \infty} \tilde{\mathcal{D}}_{\alpha, m}^{N, \text{tv}}(t^R - C) \geq 1 - \varepsilon, \quad \limsup_{N \rightarrow \infty} \tilde{\mathcal{D}}_{\alpha, m}^{N, \text{tv}}(t^R + C) \leq \varepsilon. \quad (2.5)$$

- Delayed cutoff: if $\tilde{\gamma}_\infty \geq 0$ and $m \rightarrow \infty$ then, for every $\varepsilon > 0$ there exists $C = C(\varepsilon) > 0$ such that, called

$$t^H = \frac{\beta}{2\alpha} \log(N), \quad (2.6)$$

we have

$$\liminf_{N \rightarrow \infty} \tilde{\mathcal{D}}_{\alpha, m}^{N, \text{tv}}(t^H - C\alpha^{-1}) \geq 1 - \varepsilon, \quad \limsup_{N \rightarrow \infty} \tilde{\mathcal{D}}_{\alpha, m}^{N, \text{tv}}(t^H + C\alpha^{-1}) \leq \varepsilon. \quad (2.7)$$

- No cutoff: if $\tilde{\gamma}_\infty \geq 0$ and $\sup_N m_N < \infty$, then for every constant $C > 0$ sufficiently large there exists $0 < \varepsilon_1(C) < \varepsilon_2(C) < 1$ such that

$$\liminf_{N \rightarrow \infty} \tilde{\mathcal{D}}_{\alpha, m}^{N, \text{tv}}(C\alpha^{-1}) \geq \varepsilon_1(C), \quad \limsup_{N \rightarrow \infty} \tilde{\mathcal{D}}_{\alpha, m}^{N, \text{tv}}(C\alpha^{-1}) \leq \varepsilon_2(C). \quad (2.8)$$

As a first observation, notice that, for a single class of balls, namely when considering the classical Ehrenfest urn, it is well-known that the process exhibits cutoff at $\frac{t_{\text{rel}}}{2} \log(N) + \Theta(t_{\text{rel}})$, see, e.g., [9], when the number of balls goes to infinity. Therefore, Proposition 2.1 can be summarized by saying that the process exhibits cutoff at a time that is the largest among the mixing times of its two coordinates, and to disrupt this phenomenology it is necessary and sufficient to require that $m \asymp 1$ (so that the “heavy” coordinate cannot exhibit cutoff) and that at the same time $\beta \geq \alpha$.

It is worth pointing out that, given that (R_t, H_t) is a product chain, its relaxation time is $t_{\text{rel}} = \alpha^{-1}$. In light of this, it is not difficult to deduce the following result from Proposition 2.1.

Corollary 2.2. *In the setup of Proposition 2.1 there is cutoff if and only if the product condition*

$$\lim_{N \rightarrow \infty} \frac{t_{\text{mix}}}{t_{\text{rel}}} = \infty, \quad (2.9)$$

holds.

In fact, in the framework of the mixing time of Markov chains, it is known that the *product condition* in (2.9) is necessary for cutoff [12, Prop. 18.4], and it has been shown to be sufficient in specific setups, such as birth-and-death chains [10], random walks on trees [4] and the reversible simple exclusion process with reservoirs [16].

As mentioned in the Introduction, in this note we aim at understanding, rather than the mixing behavior of the chain, the mixing time of a specific statistic. More precisely, we interpret the *heavy balls* as *impurities*, and we assume that the observer is not aware of their existence. In other words, the observer can only count the number of balls in the two containers, but cannot distinguish between heavy and regular balls. In absence of impurities, by looking at such a statistic for the process at a time slightly larger than t^R for several independent experiments, the observer should see configurations “statistically indistinguishable from equilibrium”. The question we are going to address is: how long does it take for the observer to see configurations “statistically indistinguishable from equilibrium” in presence of impurities? More precisely, for every initial configuration $(r, h) \in \mathcal{X}$ and $t \geq 0$ we consider the quantity $W_t := R_t + H_t$ and call $\nu_t^{r,h} := \text{Law}(W_t)$. We aim at investigating how this convergence takes place as a function of the two parameters (α, m) in the limit in which the number of balls goes to infinity. To this aim, for every fixed N and choice of the parameters $\alpha = \alpha_N$ and $m = m_N$, we will consider the function

$$t \mapsto \mathcal{D}_{\alpha,m}^{N,\text{tv}}(t) := \max_{(r,h) \in \mathcal{X}} \left\| \nu_t^{r,h} - \nu_\star \right\|_{\text{tv}} = \frac{1}{2} \max_{(r,h) \in \mathcal{X}} \sum_{k=0}^N |\nu_t^{r,h}(k) - \nu_\star(k)|,$$

where ν_\star is the law of a Binomial random variable of parameters N and $\frac{1}{2}$.

Theorem 2.3. *Consider the sequence*

$$\gamma = \gamma_N = \frac{2\beta - 1}{\alpha} - 1, \tag{2.10}$$

and assume that

$$\exists \gamma_\infty = \lim_{N \rightarrow \infty} \gamma_N \in [-\infty, +\infty].$$

Depending of the asymptotic behavior of the parameters $\alpha = \alpha_N$ and $\beta = \beta_N$ three different mixing behaviors can take place:

- **Insensitivity:** if $\gamma_\infty < 0$ then, for every $\varepsilon > 0$ there exists $C = C(\varepsilon) > 0$ such that,

$$\liminf_{N \rightarrow \infty} \mathcal{D}_{\alpha,m}^{N,\text{tv}}(t^R - C) \geq 1 - \varepsilon, \quad \limsup_{N \rightarrow \infty} \mathcal{D}_{\alpha,m}^{N,\text{tv}}(t^R + C) \leq \varepsilon, \tag{2.11}$$

where t^R is as in (2.4).

- **Delayed cutoff:** if $\gamma_\infty \geq 0$ and $(2\beta - 1) \log(N) \rightarrow \infty$, then, for every $\varepsilon > 0$ there exists $C = C(\varepsilon) > 0$ such that, called

$$t^{\text{dc}} = \frac{1 + \gamma}{2} \log(N),$$

we have

$$\liminf_{N \rightarrow \infty} \mathcal{D}_{\alpha,m}^{N,\text{tv}}(t^{\text{dc}} - C\alpha^{-1}) \geq 1 - \varepsilon, \quad \limsup_{N \rightarrow \infty} \mathcal{D}_{\alpha,m}^{N,\text{tv}}(t^{\text{dc}} + C\alpha^{-1}) \leq \varepsilon. \tag{2.12}$$

- **No cutoff:** if $\gamma_\infty \geq 0$ and $(2\beta - 1) \log(N) \rightarrow \ell \in [0, \infty)$, then for every constant $C = C(\ell) > 0$ sufficiently large there exists $0 < \varepsilon_1(C) < \varepsilon_2(C) < 1$ such that

$$\liminf_{N \rightarrow \infty} \mathcal{D}_{\alpha,m}^{N,\text{tv}}(C\alpha^{-1}) \geq \varepsilon_1(C), \quad \limsup_{N \rightarrow \infty} \mathcal{D}_{\alpha,m}^{N,\text{tv}}(C\alpha^{-1}) \leq \varepsilon_2(C). \tag{2.13}$$

Let us now briefly comment on the result. We might define the phase $\gamma_\infty < 0$ as *subcritical*: the strength of the perturbation is not enough to affect the mixing behavior of the process. Notice that in this case, letting $\alpha_N \rightarrow 0$ arbitrarily fast, essentially constraining the heavy particles on the right urn for an arbitrarily long time, Theorem 2.3 states that at least order \sqrt{N} impurities are necessary to affect the mixing mechanism. This is not at all surprising. Indeed, the positions of $o(\sqrt{N})$ balls is irrelevant due to the size of the fluctuations of the equilibrium measure.

At the *critical line* $\gamma_\infty = 0$ a smooth transition takes place: indeed, one should notice that, if $\alpha = \Theta(1)$ and $\gamma^{-1} \gg \log N$, then we get again the same phenomenology as in the *subcritical phase*. Nevertheless, as γ grows, the cutoff time is delayed by the amount $\frac{\gamma}{2} \log N$ and, if $\alpha \rightarrow 0$, the cutoff window is enlarged by a factor α^{-1} . It is enlightening to look at such a behavior from the following perspective: let $t_{\text{rel}} = \alpha^{-1}$ be the relaxation time of the (whole) Markov process, so that it is natural to measure time on the timescale of t_{rel} . Doing so, if $\gamma_\infty \geq 0$, Theorem 2.3 tells us that

$$t_{\text{mix}}^W(\varepsilon) := \inf\{t \geq 0 \mid \mathcal{D}_{\alpha,m}^{N,\text{tv}}(t) \leq \varepsilon\} = t_{\text{rel}}[(\beta - \frac{1}{2}) \log(N) + \Theta(1)], \quad \varepsilon \in (0, 1). \quad (2.14)$$

The latter equation makes clear the transition between the *delayed cutoff phase* to the one in which the cutoff phenomenon is disrupted, as the finiteness of the limit of $(\beta - \frac{1}{2}) \log N$ as $N \rightarrow \infty$ shows. In other words, a sort of *product condition* is necessary and sufficient for the presence of cutoff.

Remark 2.4. It is worth making an explicit comparison between the regimes in Proposition 2.1 and those in Theorem 2.3. If $\tilde{\gamma}_\infty < 0$ then $\gamma_\infty < 0$, so that not only the statistics of interest, but actually the whole chain, mixes at time t^R . On the other hand, if $\gamma_\infty \geq 0$ then $\tilde{\gamma}_\infty \geq 0$ and $m \rightarrow \infty$, therefore in this case we have that the whole chain exhibits cutoff at t^H while our statistic either mixes without cutoff on the timescale $\alpha^{-1} \ll t^H$, if $(2\beta - 1) \log N \not\rightarrow \infty$, or it has cutoff at time $t^{\text{dc}} \leq t^H$ where the difference grows to infinity if $\lim_{N \rightarrow \infty} 2\beta - 1 > 0$. The remaining case is the one in which $\tilde{\gamma}_\infty \geq 0$ but $\gamma_\infty < 0$, i.e., $\alpha \leq \beta < \frac{1+\alpha}{2}$. In this case, if $m \rightarrow \infty$ then both the whole chain and our statistic exhibit cutoff at times t^H and t^R , respectively, where the former is always larger than the latter and their ratio goes to infinity if $\beta \gg \alpha$ (in particular, we must have $\alpha \ll 1$). Finally, if $\tilde{\gamma}_\infty \geq 0$, $\gamma_\infty < 0$ but $m \asymp 1$, then the whole chain and our statistic reach equilibrium on scale α^{-1} (smoothly) and t^R (with cutoff), respectively, and the ratio between the two scales goes to infinity if and only if $\alpha \ll (\log N)^{-1}$.

2.1 The coupling

Despite the fact that the stochastic process we aim at investigating is not a Markov process, one may try to adapt the techniques used to analyze the classical Ehrenfest urn model. Like in [12, Sec. 6.5.2], a natural approach to bound from above the total variation distance at time t consists in finding a coupling of the laws ν_t and ν_* and control the probability that the coupling fails. It is well-known that in the classical Ehrenfest urn model the “natural coupling” (see [12, Sec. 6.5.2]) does not suffice to obtain the exact constant of the mixing time, but it provides an estimate that is off by a factor 2. We will show that also for our model an approach in the same spirit does not provide the right prefactor. With this aim in mind we define, for an arbitrary $t \geq 0$ the following jointly independent sequences of random variables

$$\begin{aligned} (X_i^*)_{1 \leq i \leq N} &\stackrel{d}{=} \bigotimes_{i=1}^n \text{Bern}(\frac{1}{2}) \\ (\tilde{Z}_i(t))_{1 \leq i \leq N-m} &\stackrel{d}{=} \bigotimes_{i=1}^{N-m} \text{Bern}(e^{-\alpha t}), \quad (\tilde{Z}_i(t))_{N-m < i \leq N} \stackrel{d}{=} \bigotimes_{i=N-m+1}^N \text{Bern}(e^{-t}), \end{aligned} \quad (2.15)$$

and for any prescribed initial position $(r_0, h_0) \in \{0, \dots, N - m\} \times \{0, \dots, m\}$ set $y = (y_1, \dots, y_N)$ where

$$y_i = \begin{cases} 1 & \text{if } i \leq r_0 \text{ or } i > N - h_0 \\ 0 & \text{otherwise} \end{cases}. \quad (2.16)$$

Now define

$$\tilde{X}_i(t) = (1 - \tilde{Z}_i(t))X_i^* + \tilde{Z}_i(t)y_i, \quad i \in [N], \quad (2.17)$$

and notice that

$$\left(\sum_{i=1}^{N-m} \tilde{X}_i(t), \sum_{i=N-m+1}^N \tilde{X}_i(t) \right) \stackrel{d}{=} (R_t, H_t), \quad (2.18)$$

and, as a consequence,

$$\sum_{i=1}^N \tilde{X}_i(t) \sim \nu_t^{r_0, h_0}. \quad (2.19)$$

Indeed, the random variable $\tilde{Z}_i(t)$ stands for the event that the i -ball has never been selected up to time t and, as modeled by (2.17), if $\tilde{Z}_i(t) = 0$ the i -th ball is uniformly distributed over the two urns (regardless of its initial position y_i). Under such an explicit coupled construction, we consider the event

$$\mathcal{W}_t := \left\{ \sum_{i=1}^N \tilde{Z}_i(t) = 0 \right\}.$$

By the properties of the total variation distance and by Markov inequality,

$$\mathcal{D}_{\alpha, m}^{N, \text{tv}}(t) \leq \mathbb{P}(\mathcal{W}_t^c) = \mathbb{P}\left(\sum_{i=1}^N \tilde{Z}_i(t) > 0 \right) \leq \mathbb{E}\left[\sum_{i=1}^N \tilde{Z}_i(t) \right]. \quad (2.20)$$

Since,

$$\mathbb{E}\left[\sum_{i=1}^N \tilde{Z}_i(t) \right] = me^{-\alpha t} + (N - m)e^{-t}, \quad (2.21)$$

by plugging in the values of $t = t^R + O(1)$ or $t = t^{\text{dc}} + O(\alpha^{-1})$ we see that the upper bound in (2.20) is too weak to provide the sharp estimates in Theorem 2.3.

To get a better upper bound, the key observation lies in defining another coupling as follows. Let σ be a uniformly random permutation on $[N]$, and call

$$(Z_i, Y_i) = (\tilde{Z}_{\sigma(i)}, y_{\sigma(i)}), \quad i \in [N], \quad (2.22)$$

and

$$X_i(t) = (1 - Z_i(t))X_i^* + Z_i(t)Y_i, \quad i \in [N]. \quad (2.23)$$

Since $X_i^* \stackrel{d}{=} X_j^*$ for all $i, j \in \{1, \dots, N\}$ and the random variables $(X_i^*)_{1 \leq i \leq N}$, $(\tilde{Z}_i(t))_{1 \leq i \leq N}$ and σ are independent, we deduce that

$$(X_i(t))_{1 \leq i \leq N} \stackrel{d}{=} ((1 - \tilde{Z}_{\sigma(i)}(t))X_{\sigma(i)}^* + \tilde{Z}_{\sigma(i)}(t)y_{\sigma(i)})_{1 \leq i \leq N} = (\tilde{X}_{\sigma(i)}(t))_{1 \leq i \leq N}, \quad (2.24)$$

and, by (2.19) and (2.24), we conclude that

$$\sum_{i=1}^N X_i(t) \sim \nu_t^{r_0, h_0}. \quad (2.25)$$

In what follows it will be convenient to work with a more direct representation of the sequence of random variables $(Z_i)_{1 \leq i \leq N}$. Notice indeed that, defined the jointly independent sequences of random variables

$$\begin{aligned} (P_i(t))_{1 \leq i \leq N} &\stackrel{d}{=} \bigotimes_{i=1}^N \text{Bern}(e^{-\alpha t}), & (Q_i(t))_{1 \leq i \leq N} &\stackrel{d}{=} \bigotimes_{i=1}^N \text{Bern}(e^{-t}), \\ (V_i)_{1 \leq i \leq N} &\stackrel{d}{=} \text{Unif}(\{v \in \{0, 1\}^N \mid \sum_i v_i = m\}), \end{aligned} \tag{2.26}$$

we have

$$(Z_i(t))_{1 \leq i \leq N} \stackrel{d}{=} (V_i P_i(t) + (1 - V_i) Q_i(t))_{1 \leq i \leq N}. \tag{2.27}$$

3 Strategy of proof

The strategy of proof is quite simple, and crucially relies on the recent result presented in [16], which introduces a control on the $L^2(\pi)$ distance of a negatively dependent perturbation of a product measure. More precisely, the result we are going to exploit reads as follows:

Lemma 3.1 ([16, Lemma 1]). *Let $(X_i^*)_{1 \leq i \leq N}$ be as in (2.15), $(Z_i)_{1 \leq i \leq N}$ arbitrarily distributed on $\{0, 1\}^N$ but negatively dependent, i.e.,*

$$\mathbb{E} \left[\prod_{i \in A} Z_i(t) \right] \leq \prod_{i \in A} \mathbb{E}[Z_i(t)], \quad \forall A \subseteq [N], \tag{3.1}$$

and independent of $(X_i^*)_{1 \leq i \leq N}$. Then, called π the law of $(X_i^*)_{1 \leq i \leq N}$ and μ the law of $(X_i)_{1 \leq i \leq N}$, defined as

$$X_i = (1 - Z_i) X_i^* + Z_i Y_i, \quad i = 1, \dots, N,$$

for an arbitrary $(Y_i)_{1 \leq i \leq N} \in \{0, 1\}^N$, we have

$$\left\| \frac{\mu}{\pi} - 1 \right\|_{L^2(\pi)}^2 \leq \prod_{i=1}^N (1 + \mathbb{E}[Z_i]^2) - 1. \tag{3.2}$$

Moreover, if the one in (3.1) is an equality, so is the one in (3.2).

Lemma 3.1 finds an ancestor in the exponential moment bound presented in [14, Proposition 3.2] (see also [13, Lemma 3.1]), and essentially shows that *negative dependence* can be translated into a bound for the L^2 distance as the one in (3.2), in a sharp sense. As we will show in Section 3.1, for every $t \geq 0$ the vector $(Z_i(t))_{1 \leq i \leq N}$ in (2.27) is negatively dependent, hence providing a control on the total variation distance between $(X_i^*)_{1 \leq i \leq N}$ and $(X_i(t))_{1 \leq i \leq N}$, so that the same control on the marginals given by the respective sums follows immediately by projection. As we will see in Section 3.2, the proof of the lower bound is even simpler and relies on controlling the probability that S^* and $S(t)$ lie in the subinterval $\{0, 1, \dots, \frac{N}{2} - k\}$, for certain values of $k \in \{0, \dots, \frac{N}{2}\}$. This is a particularly simple example of what is usually referred to as *distinguishing statistic*, where the function considered is the indicator of the above mentioned intervals. It is well-known, see e.g. [12, Sec. 7.3.1], that such a simple technique provides a sharp estimate for the mixing time in the classical Ehrenfest urn model. As we will show, this technique is effective also in our modified setting.

3.1 Upper bound via negative dependence

For notational simplicity, let us denote

$$\mu_t = \text{Law}(X_i(t), i \leq N), \quad \pi = \text{Law}(X_i^*, i \leq N) = \bigotimes_{i=1}^N \text{Bern}(\frac{1}{2}).$$

Since both $S(t)$ and S^* are statistics of the laws μ_t and π , respectively, it is possible to bound from above

$$\mathcal{D}_{\alpha, m}^{N, \text{tv}}(t) \leq \|\mu_t - \pi\|_{\text{tv}}. \tag{3.3}$$

We will show that for every given $t \geq 0$ the collection of random variables $(Z_i(t))_{1 \leq i \leq N}$ defined in (2.27) is *negatively dependent*, i.e., (3.1) holds. Thanks to this fact, using the L^2 bound we get

$$\begin{aligned} \|\mu_t - \pi\|_{\text{tv}} &\leq \frac{1}{2} \left\| \frac{\mu_t}{\pi} - 1 \right\|_{L^2(\pi)} \leq \frac{1}{2} \sqrt{\left[1 + \frac{1}{4} \left(\frac{m}{N} e^{-\alpha t} + \frac{N-m}{N} e^{-t} \right)^2 \right]^N - 1} \\ &\leq \frac{1}{2} \sqrt{\left[1 + \frac{1}{4} (N^{\beta-1} e^{-\alpha t} + N e^{-t})^2 \right]^N - 1}, \end{aligned} \tag{3.4}$$

where the last inequality follows by Lemma 3.1 and (3.11).

Lemma 3.2. *For every $N \in \mathbb{N}$, $t \geq 0$ and $\alpha_N, \beta_N \in [0, 1]$ the random vector $(Z_i(t))_{1 \leq i \leq N}$ in (2.27) satisfies (3.1).*

Proof. Fix $A \subseteq [N]$. Called $u_t = e^{(1-\alpha)t} \geq 1$, we have

$$\begin{aligned} \mathbb{E}[Z_i(t)]^{|A|} &= \left(\frac{m}{N} e^{-\alpha t} + \frac{N-m}{N} e^{-t} \right)^{|A|} \\ &= \frac{1}{N^{|A|}} \sum_{a=0}^{|A|} \binom{|A|}{a} m^a (N-m)^{|A|-a} (e^{-\alpha t})^a (e^{-t})^{|A|-a} \\ &= (e^{-t})^{|A|} \mathbb{E}[u_t^B], \quad B \sim \text{Bin}(|A|, \frac{m}{N}). \end{aligned}$$

For the expectation on the left-hand-side of (3.1) we have

$$\begin{aligned} \mathbb{E} \left[\prod_{i \in A} Z_i(t) \right] &= \frac{1}{(N)^{|A|}} \sum_{a=0}^{|A|} \binom{|A|}{a} (m)_a (N-m)_{|A|-a} (e^{-\alpha t})^a (e^{-t})^{|A|-a} \\ &= \sum_{a=0}^{|A|} \frac{\binom{m}{a} \binom{N-m}{|A|-a}}{\binom{N}{|A|}} (e^{-\alpha t})^a (e^{-t})^{|A|-a} \\ &= (e^{-t})^{|A|} \mathbb{E}[u_t^H], \quad H \sim \text{Hypergeom}(N, m, |A|), \end{aligned} \tag{3.5}$$

where $(x)_k$ denotes the falling factorial $x(x-1) \cdots (x-k+1)$. Indeed, the expectation on the left-hand side of (3.5) can be computed by conditioning on the number of elements in $i \in A$ such that $V_i = 1$. For every $a \in \{0, \dots, |A|\}$ the probability that $|\{i \in A \mid V_i = 1\}| = a$ is given by

$$\frac{m \dots (m-a+1) \cdot (N-m) \dots (N-m-|A|+a+1)}{N(N-1) \cdots (N-|A|+1)} = \frac{(m)_a (N-m)_{|A|-a}}{(N)_A},$$

and, conditionally on this event, for each $\bar{i} \in \{i \in A \mid V_i = 1\}$ the expectation of $Z_{\bar{i}}(t)$ is $e^{-\alpha t}$ (and similarly for $\bar{i} \in \{i \in A \mid V_i = 0\}$). Therefore, to conclude the proof it is sufficient to show that for every $u \geq 1$ we have

$$\mathbb{E}[u^B] \geq \mathbb{E}[u^H]. \tag{3.6}$$

Notice that for every $k \geq 1$ and $D = B, H$

$$\frac{d^k}{du^k} \mathbb{E}[u^D] \Big|_{u=1} = \mathbb{E}[(D)_k u^{D-k}] \Big|_{u=1} = \mathbb{E}[(D)_k]. \tag{3.7}$$

Moreover, it is known that the factorial moments of B and H are given by (see, e.g., [15])

$$\mathbb{E}[(B)_k] = (|A|)_k \frac{m^k}{N^k} \geq (|A|)_k \frac{\binom{m}{k}}{\binom{N}{k}} = \mathbb{E}[(H)_k]. \tag{3.8}$$

Therefore (3.6) follows by a finite (since for $k > |A|$ the coefficients are zero) Taylor expansion of $\mathbb{E}[u^B]$ and $\mathbb{E}[u^H]$ around $u = 1$ noticing that, thanks to (3.7) and (3.8), the coefficients of the former polynomials are larger than or equal to those of the latter. \square

3.2 TV lower bound via distinguishing statistics

In this section we will focus on the initial condition $(r, h) = (0, 0)$, and we will call

$$S(t) := \sum_{i=1}^t X_i(t), \quad S^* := \sum_{i=1}^t X_i^*, \tag{3.9}$$

where the sequences $(X_i^*)_{1 \leq i \leq N}$ and $(X_i(t))_{1 \leq i \leq N}$ are defined in (2.15) and (2.23), respectively. Notice that, with this initial condition and introducing the convenient notation

$$p_t := \frac{1}{2}(1 - e^{-\alpha t}), \quad q_t := \frac{1}{2}(1 - e^{-t}), \tag{3.10}$$

we get

$$\begin{aligned} \mathbb{E}[S(t)] &= mp_t + (N - m)q_t, \\ \text{Var}(S(t)) &= mp_t(1 - p_t) + (N - m)q_t(1 - q_t) \leq \frac{N}{4}, \\ \sum_{i=1}^N \mathbb{E}[Z_i(t)] &= N\mathbb{E}[Z_1(t)] = N - 2\mathbb{E}[S(t)]. \end{aligned} \tag{3.11}$$

The lower bounds are proved by considering the events $\mathcal{E}_k^- := \{0, 1, \dots, N/2 - k\}$, and, for appropriate choices of t and k , we use the lower bound

$$\mathcal{D}_{\alpha, m}^{N, tv}(t) \geq \mathbb{P}(S(t) \in \mathcal{E}_k^-) - \mathbb{P}(S^* \in \mathcal{E}_k^-) = 1 - \mathbb{P}(S(t) \notin \mathcal{E}_k^-) - \mathbb{P}(S^* \in \mathcal{E}_k^-). \tag{3.12}$$

On the one hand, by Chebyshev inequality we have

$$\mathbb{P}(S^* \in \mathcal{E}_k^-) \leq \mathbb{P}(|S^* - \mathbb{E}[S^*]| \geq k) \leq \frac{N}{4k^2}, \tag{3.13}$$

so that, choosing $k = \lceil \frac{c}{2} \sqrt{N} \rceil$, we obtain

$$\mathbb{P}(S^* \in \mathcal{E}_k^-) \leq \frac{1}{c^2}. \tag{3.14}$$

On the other hand, by (3.11) and Chebyshev inequality

$$\begin{aligned} \mathbb{P}(S(t) \notin \mathcal{E}_k^-) &= \mathbb{P}\left(S(t) - \mathbb{E}[S(t)] \geq \frac{N}{2} - \mathbb{E}[S(t)] - k + 1\right) \\ &\leq \frac{N}{4} \left(\frac{N}{2} - \mathbb{E}[S(t)] - k + 1\right)^{-2}, \end{aligned} \tag{3.15}$$

hence, if for some $c > 0$,

$$\frac{N}{2} - \mathbb{E}[S(t)] \geq c\sqrt{N} \tag{3.16}$$

choosing $k = \lceil \frac{c}{2} \sqrt{N} \rceil$ we obtain

$$\mathbb{P}(S(t) \notin \mathcal{E}_k^-) \leq \frac{1}{c^2}. \tag{3.17}$$

3.2.1 Refined lower bounds via CLT

For the *no cutoff* case in Theorem 2.3 we will need a refined estimate which consists in bounding the probabilities on the right-hand side of (3.12) using the CLT rather than Chebyshev inequality. To do so, notice that

$$\mathbb{P}(S^* \in \mathcal{E}_k^-) = \mathbb{P}\left(2\sqrt{N} \left(\frac{1}{N}S^* - \frac{1}{2}\right) \leq -\frac{2k}{\sqrt{N}}\right), \tag{3.18}$$

so that, choosing $k = \lceil \frac{c}{2}\sqrt{N} \rceil$ for some $c > 0$, by the Central Limit Theorem we have

$$\mathbb{P}(S^* \in \mathcal{E}_k^-) \rightarrow \Phi(-c), \quad \text{as } N \rightarrow \infty. \tag{3.19}$$

Similarly,

$$\begin{aligned} \mathbb{P}(S(t) \notin \mathcal{E}_k^-) = \\ \mathbb{P}\left(\frac{N}{\sqrt{\text{Var}(S(t))}} \left(\frac{S(t)}{N} - \frac{\mathbb{E}[S(t)]}{N}\right) > \frac{1}{\sqrt{\text{Var}(S(t))}} \left(\frac{N}{2} - \mathbb{E}[S(t)] - k\right)\right), \end{aligned} \tag{3.20}$$

hence, if for some $c > 0$

$$\frac{N}{2} - \mathbb{E}[S(t)] \geq c\sqrt{N} \tag{3.21}$$

choosing $k = \lceil \frac{c}{2}\sqrt{N} \rceil$ we obtain, for $t = t_N \rightarrow \infty$,

$$\begin{aligned} \mathbb{P}(S(t) \notin \mathcal{E}_k^-) \leq \mathbb{P}\left(\frac{N}{\sqrt{\text{Var}(S(t))}} \left(\frac{1}{N}S(t) - \mathbb{E}[S(t)]\right) > \frac{c}{2} \frac{\sqrt{N}}{\sqrt{\text{Var}(S(t))}}\right) \\ \rightarrow 1 - \Phi(c), \end{aligned} \tag{3.22}$$

where we used that $t = t_N \rightarrow \infty$ implies $\sqrt{\text{Var}(S(t))N^{-1}} \rightarrow \frac{1}{2}$, and in the last inequality that N is sufficiently large.

4 Proof of Theorem 2.3

We start by considering the case in which $\gamma_\infty < 0$ and therefore, for every N large enough, $\beta < \frac{1}{2}(1 + \alpha)$.

Insensitivity. We start by proving the upper bound. Let $t^+ = \frac{1}{2} \log N + C$, for some $C > 0$ to be determined later. Then, by (3.4) we have, for all N sufficiently large

$$\begin{aligned} \mathcal{D}_{\alpha,m}^{N,tv}(t^+) &\leq \frac{1}{2} \sqrt{\left[1 + \frac{1}{4} \left(N^{\beta-1-\frac{\alpha}{2}} e^{-C\alpha} + N^{-\frac{1}{2}} e^{-C}\right)^2\right]^N - 1} \\ &\leq \frac{1}{2} \sqrt{\left[1 + \frac{e^{-2C}}{N}\right]^N - 1} \leq \sqrt{e^{e^{-2C}} - 1}, \end{aligned}$$

and the latter can be made arbitrarily small by taking $C \rightarrow \infty$. We now prove the lower bound and set $t^- = \frac{1}{2} \log N - C$. By (3.11) we have

$$\frac{N}{2} - \mathbb{E}[S(t^-)] = \frac{1}{2} \left(mN^{-\frac{\alpha}{2}} e^{\alpha C} + (N - m)N^{-\frac{1}{2}} e^C\right) \geq \frac{1}{4} \sqrt{N} e^C, \tag{4.1}$$

where for the inequality we only used that if $\gamma_\infty < 0$ then $\beta < 1$, thus $m \leq N/2$ for all N large enough. Hence, (3.16) holds with $c = \frac{1}{4} e^C$. Choosing $k = \lceil \frac{c}{2}\sqrt{N} \rceil$, thanks to (3.17) and (3.14) we have

$$\mathbb{P}(S(t^-) \notin \mathcal{E}_k^-) \leq 16e^{-2C}, \quad \mathbb{P}(S^* \in \mathcal{E}_k^-) \leq 16e^{-2C}, \tag{4.2}$$

and the conclusion follows by (3.12) and (4.2) by taking $C \rightarrow \infty$. □

We now prove the result for the phase in which $\gamma_\infty \geq 0$ and we first focus on the case in which $(2\beta - 1) \log N \rightarrow \infty$.

Delayed cutoff. For the upper bound, letting $t^+ = \frac{1+\gamma}{2} \log N + C\alpha^{-1} = \frac{2\beta-1}{2\alpha} \log N + C\alpha^{-1}$ and using (3.4) we get, for all N sufficiently large,

$$\begin{aligned} \mathcal{D}_{\alpha,m}^{N,tv}(t^+) &\leq \frac{1}{2} \sqrt{\left[1 + \frac{1}{4} \left(N^{\beta-1-\beta+\frac{1}{2}} e^{-C} + N^{-\frac{1}{2}-\frac{\gamma}{2}} e^{-C\alpha^{-1}}\right)^2\right]^N - 1} \\ &\leq \frac{1}{2} \sqrt{\left[1 + \frac{e^{-2C}}{N}\right]^N - 1} \leq \sqrt{e^{e^{-2C}} - 1}. \end{aligned}$$

For the lower bound, set $t^- = \frac{2\beta-1}{2\alpha} \log N - C\alpha^{-1}$ and notice that, similarly to (4.1),

$$\frac{N}{2} - \mathbb{E}[S(t^-)] \geq \frac{1}{2} m N^{\frac{1}{2}-\beta} e^C \geq \frac{1}{4} \sqrt{N} e^C. \tag{4.3}$$

Therefore, the conclusion follows again by (4.2). □

Finally, we consider the case in which $\gamma_\infty \geq 0$ and $(2\beta - 1) \log N \rightarrow \ell \in [0, \infty)$.

No cutoff. Let $C > 0$ be a constant to be tuned later. Since $(2\beta - 1) \log N \rightarrow \ell < \infty$, for all $\varepsilon > 0$ we have

$$2\beta - 1 < \frac{\ell + \varepsilon}{\log(N)}, \tag{4.4}$$

for all N sufficiently large. On the other hand, since $\gamma_\infty \geq 0$, we have $\lim_{N \rightarrow \infty} \frac{2\beta-1}{\alpha} \geq 1$ and by (4.4) we deduce that

$$\alpha < \frac{\ell + \varepsilon}{\log(N)}, \tag{4.5}$$

for all N large enough. By (3.4), and exploiting (4.4) and (4.5) we have: for all $\varepsilon > 0$ and N large enough

$$\begin{aligned} \mathcal{D}_{\alpha,m}^{N,tv}(C\alpha^{-1}) &\leq \frac{1}{2} \sqrt{\left[1 + \frac{1}{4} \left(N^{\beta-1} e^{-C} + N e^{-C\alpha^{-1}}\right)^2\right]^N - 1} \\ &\leq \frac{1}{2} \sqrt{\left[1 + \frac{1}{4} \left(N^{-\beta} e^{-C+\ell+\varepsilon} + N^{1-\frac{C}{\ell+\varepsilon}}\right)^2\right]^N - 1}. \end{aligned} \tag{4.6}$$

Choosing now $C > 2(\ell + \varepsilon)$ we have

$$N^{-\beta} e^{-C+\ell+\varepsilon} > N^{1-\frac{C}{\ell+\varepsilon}},$$

for all N large enough, and therefore

$$\begin{aligned} \mathcal{D}_{\alpha,m}^{N,tv}(C\alpha^{-1}) &\leq \frac{1}{2} \sqrt{\left[1 + N^{-2\beta} e^{-2(C-\ell-\varepsilon)}\right]^N - 1} \\ &\leq \frac{1}{2} \sqrt{\exp(N^{-2\beta+1} e^{-2(C-\ell-\varepsilon)}) - 1} \leq \frac{1}{2} \sqrt{e^{e^{-2(C-\ell-\varepsilon)-\ell+\varepsilon}} - 1}, \end{aligned} \tag{4.7}$$

where in the last step we used that: for all $\varepsilon > 0$

$$-(2\beta - 1) < \frac{-\ell + \varepsilon}{\log(N)}, \tag{4.8}$$

for all N sufficiently large. In conclusion, for all $\varepsilon > 0$ and N large enough

$$\mathcal{D}_{\alpha,m}^{N,tv}(C\alpha^{-1}) \leq \frac{1}{2} \sqrt{e^{e^{-2C+\ell+3\varepsilon}} - 1}, \tag{4.9}$$

and the latter is strictly smaller than 1 as soon as $C = C(\ell)$ is sufficiently large.

For the lower bound we use the estimates in Section 3.2.1. Indeed, for all $\varepsilon > 0$ and N large enough

$$\frac{N}{2} - \mathbb{E}[S(C\alpha^{-1})] \geq \frac{1}{2}N^\beta e^{-C} \geq \frac{1}{2}\sqrt{N}e^{-C-\varepsilon},$$

where we used that

$$2\beta - 1 > \frac{\ell - \varepsilon}{\log(N)} \implies \beta > \frac{1}{2} - \frac{\varepsilon}{\log(N)}, \tag{4.10}$$

for all N sufficiently large. Therefore, by (3.19) and (3.22) with $c = \frac{1}{2}e^{-C-\varepsilon}$ and $k = \lceil \frac{\varepsilon}{2}\sqrt{N} \rceil$ we have: for all $\varepsilon, \delta > 0$

$$\mathbb{P}(S(C\alpha^{-1}) \notin \mathcal{E}_k^-) \leq 1 - \Phi(\frac{1}{2}e^{-C-\varepsilon}) + \delta, \quad \mathbb{P}(S^* \in \mathcal{E}_k^-) \leq \Phi(-\frac{1}{2}e^{-C-\varepsilon}) + \delta,$$

so that, for all $C > 0$ and every N sufficiently large, thanks to (3.12),

$$\mathcal{D}_{\alpha, m}^{N, \text{tv}}(C\alpha^{-1}) \geq \Phi(\frac{1}{2}e^{-C-\varepsilon}) - \Phi(-\frac{1}{2}e^{-C-\varepsilon}) - 2\delta, \tag{4.11}$$

and, for any $C > 0$, taking $\delta = \varepsilon = \varepsilon(C)$ sufficiently small, the right-hand side of (4.11) is strictly larger than 0 for all N large enough. \square

This concludes the proof of Theorem 2.3.

5 Proof of Proposition 2.1 and Corollary 2.2

Proof of Proposition 2.1. If $\tilde{\gamma}_\infty < 0$ then $N - m \rightarrow \infty$, hence the “regular” coordinate exhibits cutoff at t^R with window of size $\Theta(1)$. Therefore, the lower bound in (2.5) follows by projection. Moreover, if $m \rightarrow \infty$ the coordinate $(H_t)_{t \geq 0}$ exhibits cutoff at t^H with a window of size $\Theta(\alpha^{-1})$. Hence, also the lower bound (2.7) follows by projection. To show the lower bound in (2.8), recall that $t_{\text{rel}} = \alpha^{-1}$ and that in this case we have $\alpha \lesssim (\log N)^{-1}$. Then, by [12, Lemma 20.12], at time $t = C\alpha^{-1}$ the “heavy” coordinate is at total-variation distance at least $\frac{1}{2}e^{-\frac{t}{t_{\text{rel}}}} = \frac{1}{2}e^{-C}$ from its stationary distribution, and therefore the desired lower bound follows again by projection. To get the upper bounds in (2.5) and (2.7) it is enough to argue as in the proof of [12, Theorem 20.7]: thanks to the product structure of the chain we have

$$\max_{(r, h) \in \mathcal{X}} \|\chi_t^{r, h} - \chi\|_{\text{tv}}^2 \leq 2N^\beta e^{-2\alpha t} + 2Ne^{-2t}. \tag{5.1}$$

Hence, the upper bounds in (2.5), (2.7) and (2.8) follow by choosing $t = t^R + C$, $t = t^H + C\alpha^{-1}$ and $t = C\alpha^{-1}$, respectively, for some $C = C(\varepsilon)$ large enough. \square

Proof of Corollary 2.2. Regardless of the choice of the parameters we have

$$t_{\text{mix}} \asymp t^R + t^H \asymp \log(N) + t^H, \tag{5.2}$$

since the mixing time of the whole chain is always larger or equal than the mixing time of one of the two coordinates and smaller or equal than their sum. Using $t_{\text{rel}} = \alpha^{-1}$ and (5.2), simple algebra shows that $\liminf t_{\text{mix}}/t_{\text{rel}} < \infty$ if and only if $\alpha + \beta \lesssim (\log N)^{-1}$ (at least along subsequences), and the latter implies $m \asymp 1$ and $\tilde{\gamma}_\infty = 0$. Conversely, in the *insensitivity* regime the condition $\tilde{\gamma}_\infty$ ensures that $\alpha \asymp 1$ and therefore $t_{\text{mix}}/t_{\text{rel}} \gg 1$; and, similarly, in the *delayed cutoff* regime the assumption $m \rightarrow \infty$ ensures that $\beta \gg (\log N)^{-1}$ and therefore $t_{\text{mix}}/t_{\text{rel}} \gg 1$. \square

References

- [1] David Aldous and Persi Diaconis. Shuffling cards and stopping times. *Amer. Math. Monthly*, 93(5):333–348, 1986. MR0841111
- [2] Luca Avena, Hakan Guldás, Remco van der Hofstad and Frank den Hollander. Random walks on dynamic configuration models: A trichotomy. *Stochastic Process. Appl.*, 129(9):3360–3375, 2019. MR3985565
- [3] Luca Avena, Hakan Guldás, Remco van der Hofstad, Frank den Hollander and Oliver Nagy. Linking the mixing times of random walks on static and dynamic random graphs. *Stochastic Process. Appl.*, 153:145–182, 2022. MR4474678
- [4] Riddhipratim Basu, Jonathan Hermon and Yuval Peres. Characterization of cutoff for reversible Markov chains. *Ann. Probab.*, 45:1448–1487, 2017. MR3650406
- [5] Pietro Caputo and Matteo Quattropiani. Mixing Time of PageRank Surfers on Sparse Random Digraphs *Random Structures Algorithms*, 59(3):376–406, 2021. MR4295568
- [6] Pietro Caputo and Matteo Quattropiani. Mixing Time Trichotomy in Regenerating Dynamic Digraphs *Stochastic Process. Appl.*, 137:222–251, 2021. MR4244192
- [7] Guan-Yu Chen and Laurent Saloff-Coste. On the mixing time and spectral gap for birth and death chains. *ALEA Lat. Am. J. Probab. Math. Stat.*, 10(1):293–321, 2013. MR3083928
- [8] Guan-Yu Chen and Laurent Saloff-Coste. Computing cutoff times of birth and death chains. *Electron. J. Probab.*, 20:1–47, 2015. MR3371435
- [9] Persi Diaconis. The cutoff phenomenon in finite Markov chains. *PNAS*, 93(4):1659–1664, 1996. MR1374011
- [10] Jian Ding, Eyal Lubetzky, and Yuval Peres. Total variation cutoff in birth-and-death chains. *Probab. Theory Related Fields*, 146:61, 2010. MR2550359
- [11] Persi Diaconis and Laurent Saloff-Coste. Separation cut-offs for birth and death chains. *Ann. Appl. Probab.*, 16(4):2098–2122, 2006. MR2288715
- [12] David A. Levin and Yuval Peres. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2017. Second edition of [MR2466937], with contributions by Elizabeth L. Wilmer, with a chapter on “Coupling from the past” by James G. Propp and David B. Wilson.
- [13] Eyal Lubetzky and Allan Sly. Universality of cutoff for the Ising model. *Ann. Probab.*, 45(6):3664–3696, 2017. MR3729612
- [14] Jason Miller and Yuval Peres. Uniformity of the uncover set of random walk and cutoff for lamplighter chains. *Ann. Probab.*, 40(2):535–577, 2012. MR2952084
- [15] Renfrey B. Potts. Note on the factorial moments of standard distributions. *Aust. J. Phys.*, 6(4):498–499, 1953.
- [16] Justin Salez. Universality of cutoff for exclusion with reservoirs. *Ann. Probab.*, 51(2):478–494, 2023. MR4546624
- [17] Graham White. Strong stationary times for features of random walks. *arXiv preprint*, arXiv:1910.03837, 2019.
- [18] Graham White. Coupling for features of random walks. *arXiv preprint*, arXiv:1910.03265, 2019.

Acknowledgments. The author is a member of GNAMPA-INdAM, and he thanks the German Research Foundation (project number 444084038, priority program SPP2265) for financial support. Moreover, the author wishes to thank Pietro Caputo and Federico Sau for helpful discussions on the topic, and the anonymous referee for having suggested several key improvements on the first draft of this paper.

Electronic Journal of Probability

Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS¹)
- Easy interface (EJMS²)

Economical model of EJP-ECP

- Non profit, sponsored by IMS³, BS⁴, ProjectEuclid⁵
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund⁶ (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

¹LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

²EJMS: Electronic Journal Management System: <https://vtex.lt/services/ejms-peer-review/>

³IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

⁴BS: Bernoulli Society <http://www.bernoulli-society.org/>

⁵Project Euclid: <https://projecteuclid.org/>

⁶IMS Open Access Fund: <https://imstat.org/shop/donation/>