

Detecting Phase Transitions through Non-Equilibrium Work Fluctuations

*Original*

Detecting Phase Transitions through Non-Equilibrium Work Fluctuations / Colangeli, Matteo; Di Francesco, Antonio; Rondoni, Lamberto. - In: SYMMETRY. - ISSN 2073-8994. - ELETTRONICO. - 16:125(2024), pp. 1-16.  
[10.3390/sym16010125]

*Availability:*

This version is available at: 11583/2987614 since: 2024-04-07T16:56:16Z

*Publisher:*

MDPI

*Published*

DOI:10.3390/sym16010125

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

Article

# Detecting Phase Transitions through Non-Equilibrium Work Fluctuations

Matteo Colangeli <sup>1,\*</sup> , Antonio Di Francesco <sup>1</sup>  and Lamberto Rondoni <sup>2,3</sup> 

<sup>1</sup> Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila, Via Vetoio, 67100 L'Aquila, Italy; antonio.difrancesco5@graduate.univaq.it

<sup>2</sup> Department of Mathematical Sciences "Giuseppe Luigi Lagrange", Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy; lamberto.rondoni@polito.it

<sup>3</sup> National Institute for Nuclear Physics, Sezione di Torino, Via Pietro Giuria 1, 10125 Turin, Italy

\* Correspondence: matteo.colangeli1@univaq.it

**Abstract:** We show how averages of exponential functions of path-dependent quantities, such as those of Work Fluctuation Theorems, detect phase transitions in deterministic and stochastic systems. State space truncation—the restriction of the observations to a subset of state space with prescribed probability—is introduced to obtain that result. Two stochastic processes undergoing first-order phase transitions are analyzed both analytically and numerically: a variant of the Ehrenfest urn model and the 2D Ising model subject to a magnetic field. In the presence of phase transitions, we prove that even minimal state space truncation makes averages of exponentials of path-dependent variables sensibly deviate from full state space values. Specifically, in the case of discontinuous phase transitions, this approach is strikingly effective in locating the transition value of the control parameter. As this approach works even with variables different from those of fluctuation theorems, it provides a new recipe to identify order parameters in the study of non-equilibrium phase transitions, profiting from the often incomplete statistics that are available.

**Keywords:** truncated distributions; absolute irreversibility; Markov chains; Ising model



**Citation:** Colangeli, M.; Di Francesco, A.; Rondoni, L. Detecting Phase Transitions through Non-Equilibrium Work Fluctuations. *Symmetry* **2024**, *16*, 125. <https://doi.org/10.3390/sym16010125>

Academic Editor: Fernando Haas

Received: 19 December 2023

Revised: 13 January 2024

Accepted: 17 January 2024

Published: 20 January 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Exponentials of microscopically expressed variables have been used for a long time in statistical physics and molecular dynamics. Bennett's formulae for the free energy [1], Widom's relation [2], and Zwanzig's relation [3], are among them. More recently, they have become popular in non-equilibrium statistical mechanics in the form of Fluctuation Relations [4–6], which constitute a step in the direction of generalizing linear response theory and the Fluctuation Dissipation Relation, and in the form of Work Fluctuation Theorems (WFT) [7–10], which focus on equilibrium properties derived from non-equilibrium processes. See [11,12] for comprehensive reviews and [13–25] for further discussion.

Exponential observables, such as  $e^{-\beta\sigma}$ , with  $\sigma$  denoting a kind of dissipation functional of the system trajectories, may provide useful insight into the properties of non-equilibrium systems but need to be handled with care. At the very least, the amount of data for accurate calculations may be prohibitively large because substantial contributions to the statistical properties of such observables pertain to the (low probability) tails of the relevant distributions, e.g., [7]. The problem is often enhanced by high energy barriers that keep trajectories from accessing specific regions. The problem is well known. For instance, in the context of free energy differences estimation [1,26], numerous ad hoc techniques and procedures have been developed to deal with it [27–29]. In general, many challenges remain. Moreover, the recent work [30], concerning the canonical ensemble of classical Hamiltonian systems shows how exponential variables can be strongly affected by finite-size effects or lack of ergodicity [30]. This must be considered, in particular, when dealing with small systems because ensembles were originally conceived to treat specific observables of large

systems that (a) are not affected by finite-size effects and (b) their time averages equal their phase space averages.

In this paper, we will show how the above difficulties can actually be leveraged to recover important information about the system behavior. This way, often unavoidable hindrances in the investigation of the objects of modern interest may turn profitable under some conditions.

We start by concisely recalling the theoretical background of WFT for Markov processes [31–33] and the notion of absolute irreversibility [34–37]. Then, we introduce the concepts of state space truncation, which can be interpreted both as a result of incomplete information on the system of interest, or as the representation of a system with finitely many degrees of freedom, interacting with only a portion of its environment. This is of interest, for instance, when the system is subjected to rapid transformations, which is common in present-day bio- and nano-technologies. We find that the corresponding ensemble averages of  $e^{-\beta W_d}$ ,  $W_d$  being the dissipated work, may detect phase transitions and locate them in parameter space.

This is shown by analyzing a variant of the celebrated Ehrenfest urn model in detail, derived from a deterministic 2D model of particle transport [38], and a classical 2D Ising model with periodic boundary conditions in a magnetic field. For both, it is found that the ensemble average of  $e^{-\beta W_d}$  as a function of the driving control parameter, computed over trajectories starting in the lowest energy states (thus, with highest stationary probability), manifests a discontinuity in the thermodynamic limit whenever the system undergoes a first-order phase transition. A smooth dependence is instead found in the case of the Ising model at supercritical temperatures.

This paper is structured as follows: in Section 2 we introduce the theoretical background and the ensemble averages over relevant subsets of the state space, along with the notion of absolute irreversibility. Section 3 analyzes in detail the systems undergoing phase transitions mentioned above, using a suitably amended version of the Jarzynski Equality. Conclusions are drawn in Section 4.

## 2. Theory and Methods

### 2.1. Work and Free Energy for Markov Chains

Here, we concisely recall some basic concepts referring to discrete time Markov chains [39,40]. A Markov chain is a finite set of random variables  $\mathbf{X} = (X_i)_{i=0}^\tau$ , with  $\tau \in \mathbb{N}$ , taking values on a finite state space  $S$ , such that

$$P(X_i = x_i | X_0 = x_0, \dots, X_{i-1} = x_{i-1}) = P(X_i = x_i | X_{i-1} = x_{i-1})$$

for every  $0 < i \leq \tau$  and every sequence  $x_0, \dots, x_{i-1}$  for which  $P(X_0 = x_0, \dots, X_{i-1} = x_{i-1}) > 0$ . The elements of  $S$  correspond to the possible *states* of the system, with  $X_i$  representing the state at time  $i$ . The transition matrix  $Q = (Q_{xy})$  has entries  $Q_{xy} = P(X_i = y | X_{i-1} = x)$ , for all  $x, y \in S$  and  $0 < i \leq \tau$ , representing the transition probabilities from the state  $x$  to the state  $y$ , which obey  $\sum_{y \in S} Q_{xy} = 1$ , for all  $x \in S$ . A specific realization of the sequence  $X_0, \dots, X_\tau$  is characterized by the probability of the path  $\mathcal{P}_{\mu_0}(\mathbf{x}) = \mu_0(x_0) \mathcal{P}(\mathbf{x} | x_0) = \mu_0(x_0) \prod_{i=0}^{\tau-1} Q_{x_i x_{i+1}}$ , where  $\mu_0(x) = P(X_0 = x)$  is an assigned discrete probability distribution and  $\mathbf{x} = (x_0, x_1, \dots, x_\tau) \in S^{\tau+1}$ . We recall that a Markov chain admits a unique invariant distribution  $\mu$  on  $S$ , such that  $\mu Q = \mu$  (where  $\mu$  is conceived as a row vector) if it is irreducible and its states are ergodic, i.e., they are aperiodic and persistent with a finite mean recurrence time [41]. An explicit time dependence is sometimes introduced in the definition of the transition matrices to mimic the effect of an external drive affecting the evolution of the system. The action of the drive is expressed by the *protocol*  $\lambda(i) = \lambda_i$ ,  $i = 0, 1, \dots, \tau$ . This way, the Markov chain turns inhomogeneous, namely  $(Q_i)_{xy} = P(X_i = y | X_{i-1} = x)$  depends on time. One may then introduce the following path-dependent quantity, called *work* [42]:

$$W = \sum_{i=0}^{\tau-1} [E_{i+1}(x_i) - E_i(x_i)]$$

where the function  $E_i(x_i)$  is related to a distribution  $\mu_i(x)$  supported on  $S$  via:

$$\mu_i(x) = \frac{1}{Z_i} e^{-\beta E_i(x)}, \quad x \in S, \quad i = 0, 1, \dots, \tau, \quad (1)$$

with  $Z_i = \sum_{x \in S} \exp\{-\beta E_i(x)\}$ , for some  $\beta > 0$ . Correspondingly, the quantity  $F = -\beta^{-1} \log Z$  is called the *free energy* of the system. Therefore, letting  $\Delta F \equiv F_\tau - F_0 = -\beta^{-1} \log Z_\tau / Z_0$  be the free energy difference between the states  $x_\tau$  and  $x_0$  [31] and introducing  $W_d = W - \Delta F$  as the *dissipated work*, the Jarzynski Equality (JE) [7] attains the form:

$$\langle e^{-\beta W_d(x)} \rangle = 1, \quad (2)$$

where, for a generic path-dependent observable  $\mathcal{A} : S^{\tau+1} \rightarrow \mathbb{R}$ , the angular brackets represent the following average with respect to the initial distribution:

$$\langle \mathcal{A}(x) \rangle = \sum_{x \in S^{\tau+1}} \mathcal{A}(x) \mu_0(x_0) \mathcal{P}(x|x_0).$$

For the quantity of interest in Equation (2), the following can be established:

$$\begin{aligned} e^{-\beta W_d} &= e^{-\beta(W - \Delta F)} = \frac{Z_0}{Z_\tau} e^{-\beta W} = \\ &= \frac{Z_0}{Z_\tau} \frac{e^{-\beta E_1(x_0)}}{e^{-\beta E_0(x_0)}} \frac{e^{-\beta E_2(x_1)}}{e^{-\beta E_1(x_1)}} \cdots \frac{e^{-\beta E_\tau(x_{\tau-1})}}{e^{-\beta E_{\tau-1}(x_{\tau-1})}}. \end{aligned} \quad (3)$$

which, together with Equation (1), yields:

$$e^{-\beta W_d} = \frac{Z_0}{Z_\tau} \frac{Z_1 \mu_1(x_0)}{Z_0 \mu_0(x_0)} \frac{Z_2 \mu_2(x_1)}{Z_1 \mu_1(x_1)} \cdots \frac{Z_\tau \mu_\tau(x_{\tau-1})}{Z_{\tau-1} \mu_{\tau-1}(x_{\tau-1})} = \frac{\mu_1(x_0)}{\mu_0(x_0)} \frac{\mu_2(x_1)}{\mu_1(x_1)} \cdots \frac{\mu_\tau(x_{\tau-1})}{\mu_{\tau-1}(x_{\tau-1})} \quad (4)$$

This result applies as long as the distribution  $\mu_i(x)$  is stationary for the process at fixed  $\lambda = \lambda_i$  and is supported on  $S$ , for all  $i = 0, \dots, \tau$ . For  $N$ -step transformations within this framework, with  $N > 1$ , the validity of the JE was established in [31].

Let us now take one-step evolutions going from a state  $x_0$  to a state  $x_1$ , which means  $\tau = 1$  in Equation (4). Under the above assumption that the support of  $\mu_0$  and  $\mu_1$  is the same  $S$ , we obtain the identity:

$$\langle e^{-\beta W_d(x)} \rangle = \sum_{x \in S} \frac{\mu_1(x)}{\mu_0(x)} \mu_0(x) = \sum_{x \in S} \mu_1(x) = 1. \quad (5)$$

Note that the sum in Equation (5) is performed over states  $x$  that belong to the state space  $S$  of the Markov chain associated with the initial value  $\lambda(0) = \lambda_0$  of the protocol, and that such states are eventually weighted with the distribution  $\mu_1(\cdot)$ , which is invariant for the final value  $\lambda(1) = \lambda_1$  of the protocol. In other words, the result follows from sampling the states  $x$  belonging to the initial state space  $S$ , with the probability corresponding to final distribution  $\mu_1(\cdot)$ . It is important to remark that the last sum in Equation (5) would not equal 1, in general, if the initial and final state spaces do not coincide. If they do, Equation (5) shows that the JE holds regardless of the specific protocol. In particular, it holds even if the stationary states undergo a phase transition between the initial and final values of  $\lambda$ . In the next subsection, we will investigate variations of the ensemble averages considered above, obtained by restricting the set of observed states to a subset of the whole state space  $S$ .

## 2.2. State Space Truncation and Absolute Irreversibility

We now show how state space truncation, which may be due to insufficient statistic or to finite-size effects, affects the ensemble averages of the exponential of path-dependent observables such as  $W_d$ . We define the *important state space* as the restriction of the state space to the subset of states  $\eta \subset \mathcal{S}$  that can be accessed with probability larger than a certain threshold  $\bar{\mu}$  [43]:

$$\eta = \{x \in \mathcal{S} \mid \mu(x) \geq \bar{\mu}\} \quad (6)$$

One consequently has

$$\sum_{x \in \eta} \mu(x) = 1 - \delta, \quad \text{where } \delta = \sum_{x \in \mathcal{S} \setminus \eta} \mu(x) \quad (7)$$

and a probability measure with support on  $\eta$  is obtained from  $\mu$  dividing it by  $1 - \delta$ :

$$\mu_\delta(x) = \frac{\mu(x)}{\sum_{x \in \eta} \mu(x)} = \frac{\mu(x)}{1 - \delta} \quad (8)$$

Let us now study how the average in (5) is affected by this restriction of the state space to the subset  $\eta$ . Following [43], we denote  $\eta_0 = \{x \in \mathcal{S} \mid \mu_0(x) \geq \bar{\mu}\}$  and  $\eta_1 = \{x \in \mathcal{S} \mid \mu_1(x) \geq \bar{\mu}\}$ , with the same cutoff probability  $\bar{\mu}$  for both  $\eta_0$  and  $\eta_1$ . Correspondingly, we introduce the symbols  $\delta_0 \equiv 1 - \sum_{x \in \eta_0} \mu_0(x)$  and  $\delta_1 \equiv 1 - \sum_{x \in \eta_1} \mu_1(x)$ . Next, we call  $\mathcal{I} = \eta_0 \cap \eta_1$ , so that the sets  $\eta_0 = \mathcal{I} \cup (\eta_0 \setminus \mathcal{I})$  and  $\eta_1 = \mathcal{I} \cup (\eta_1 \setminus \mathcal{I})$  are conveniently written as the union of disjoint subsets. This way, we can write

$$\left\langle e^{-\beta W_d(x)} \right\rangle_{\eta_0} = \sum_{x \in \eta_0} \frac{\mu_1(x)}{\mu_0(x)} \mu_{0,\delta_0}(x) = \frac{1}{1 - \delta_0} \sum_{x \in \eta_0} \mu_1(x). \quad (9)$$

A careful rewriting of Equation (9) yields

$$\left\langle e^{-\beta W_d(x)} \right\rangle_{\eta_0} = \frac{1 - \delta_1}{1 - \delta_0} + \frac{1}{1 - \delta_0} \left( \sum_{x \in \eta_0 \setminus \mathcal{I}} \mu_1(x) - \sum_{x \in \eta_1 \setminus \mathcal{I}} \mu_1(x) \right), \quad (10)$$

where we used the identity  $\mathcal{I} = \eta_1 \setminus (\eta_1 \setminus \mathcal{I})$ . The second term in (10) vanishes, yielding  $\langle \exp(-\beta W_d) \rangle_{\eta_0} = 1$ , even in the presence of a reduction in the state space, if  $\mu_0$  and  $\mu_1$  coincide, so that  $\eta_0 = \eta_1$  and  $\delta_0 = \delta_1$ . The latter equality does not hold, in general, when  $\mu_0 \neq \mu_1$ . Deviations from unity are small, as long as  $\mu_0$  and  $\mu_1$  differ little from one another. This is the case, for instance, when the initial and the final stationary states, corresponding to the initial and final values of the protocol, fall within the same macroscopic phase, as will be evidenced in the next section.

Two different cases are worth investigating when  $\mu_0 \neq \mu_1$ . Consider first the case with  $\eta_1 \subseteq \eta_0$ , which implies  $\mathcal{I} = \eta_1$ ,  $\eta_1 \setminus \mathcal{I} = \emptyset$  and  $\eta_0 \setminus \mathcal{I} = \eta_0 \setminus \eta_1$ . Then, Equation (10) yields:

$$\left\langle e^{-\beta W_d(x)} \right\rangle_{\eta_0} = \frac{1 - \delta_1}{1 - \delta_0} + \frac{1}{1 - \delta_0} \left( \sum_{x \in \eta_0 \setminus \eta_1} \mu_1(x) \right) = \frac{1 - (1 - A)\delta_1}{1 - \delta_0} \quad (11)$$

where  $A = \sum_{x \in \eta_0 \setminus \eta_1} \mu_1(x) / \delta_1$  obeys  $A \in [0, 1]$ . It appears that for  $A$  close to 1, the value  $\langle e^{-\beta W_d} \rangle_{\eta_0}$  can exceed 1, depending on the values of  $\mu_0(x)$  and  $\mu_1(x)$ , with  $x \in \eta_0$ . For simplicity, we illustrate this fact with hypothetical Gaussian distributions with standard deviations  $\sigma_0$  and  $\sigma_1$ , centered at  $x = \bar{x}$ , for which analytical results are readily obtained (Figure 1). While such Gaussians are not directly related to our finite state Markov chains, they may arise as limit cases of finite state processes, which will lead to analogous results. In the case of Gaussians, the state space is  $\mathcal{S} = \mathbb{R}$ , and the initial truncated space we consider is one interval,  $\eta_0 = [\bar{x} - \Delta, \bar{x} + \Delta]$ , of width  $|\eta_0| = 2\Delta$ . The top right panel of the

figure shows that, for different values of  $\bar{\mu}$  (which is the same for both the initial and the final distributions, as generally assumed throughout this work), values of  $\sigma_1/\sigma_0 < 1$  yield  $\langle e^{-\beta W_d} \rangle_{\eta_0} > 1$ .

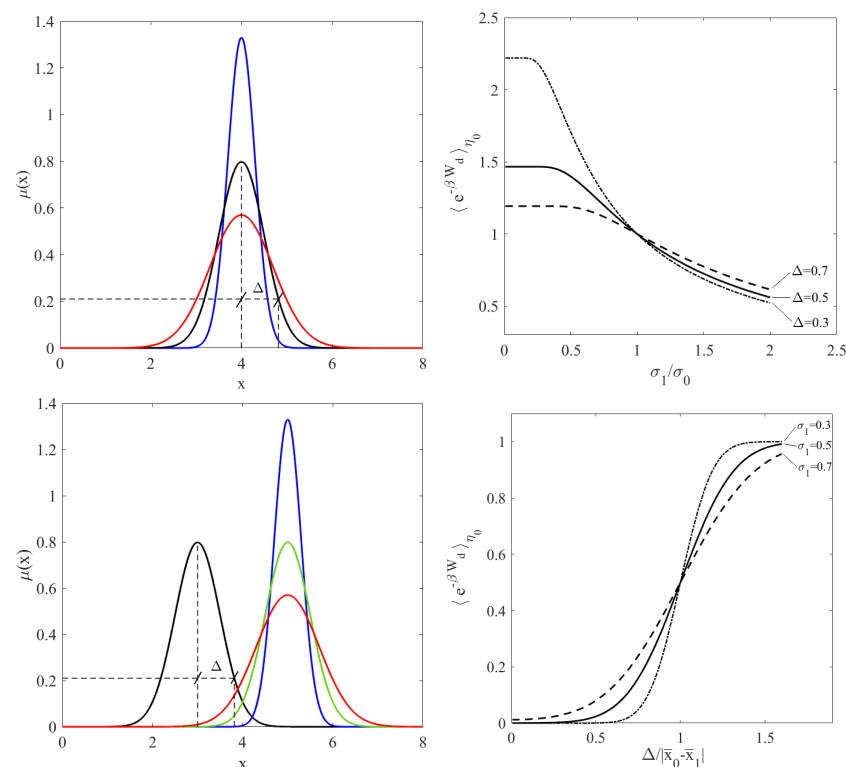
Another scenario concerns distributions  $\mu_0$  and  $\mu_1$  for which  $\eta_1 \neq \eta_0$  and  $\mathcal{I} = \emptyset$ . In this case, the following bound is found:

$$\langle e^{-\beta W_d(x)} \rangle_{\eta_0} = \frac{1 - \delta_1}{1 - \delta_0} + \frac{1}{1 - \delta_0} \left( \sum_{x \in \eta_0} \mu_1(x) - \sum_{x \in \eta_1} \mu_1(x) \right) = \frac{1}{1 - \delta_0} \sum_{x \in \eta_0} \mu_1(x) < 1 \quad (12)$$

The inequality follows from the fact that  $\mu_1 < \bar{\mu}$  on  $\eta_0$ , because  $\mathcal{I} = \emptyset$ , hence,

$$\sum_{x \in \eta_0} \mu_1(x) < |\eta_0| \bar{\mu} \leq \sum_{x \in \eta_0} \mu_0(x) = 1 - \delta_0$$

where  $|\eta_0|$  now represents the cardinality of  $\eta_0$ . It is useful to comment on the behavior of the quantity  $\langle e^{-\beta W_d} \rangle_{\eta_0}$  as a function of  $|\eta_0|$ , whose value can be tuned by varying the cutoff probability  $\bar{\mu}$ . For simplicity, we refer again to Gaussian distributions, taking  $\eta_0 = [\bar{x} - \Delta, \bar{x} + \Delta]$ , and  $\bar{\mu}$  and  $\Delta$  in a range that allows  $\mathcal{I} = \emptyset$ , so that Equation (12) applies. In the bottom-left panel of Figure 1, the distributions  $\mu_0$  and  $\mu_1$  are Gaussian densities centered at  $\bar{x}_1$  and  $\bar{x}_1 \neq \bar{x}_0$ , with standard deviations  $\sigma_0$  and  $\sigma_1$ , respectively. The bottom right panel of Figure 1 shows that, for values of  $\Delta$  guaranteeing  $\mathcal{I} = \emptyset$ , a decrease of  $\Delta$  leads to a decrease of  $\langle e^{-\beta W_d} \rangle_{\eta_0}$  to values substantially smaller than 1.



**Figure 1.** Top row: Gaussian distributions centered at  $\bar{x}_0 = 4$  (left panel):  $\mu_0(x)$  has standard deviation  $\sigma_0 = 0.5$  (black), while the standard deviation of  $\mu_1(x)$  is  $\sigma_1 = 0.3$  (red) and  $0.7$  (blue). The horizontal dashed line denotes the value  $\bar{\mu}$ , which determines  $\eta_0 = [\bar{x}_0 - \Delta, \bar{x}_0 + \Delta]$ . (right panel): behavior of  $\langle e^{-\beta W_d} \rangle_{\eta_0}$  as a function of  $\sigma_1/\sigma_0$  for different values of  $\Delta$ , computed using Equation (12). Bottom row, (left panel):  $\mu_0(x)$  is a Gaussian centered at  $\bar{x}_0 = 3$  with standard deviation  $\sigma_0 = 0.5$  (black curve), while  $\mu_1(x)$  is a Gaussian function centered at  $\bar{x}_1 = 5$  with standard deviations  $\sigma_1 = 0.3$  (blue),  $0.5$  (green),  $0.7$  (red). (right panel): graph of  $\langle e^{-\beta W_d} \rangle_{\eta_0}$  as a function of  $\Delta/|\bar{x}_0 - \bar{x}_1|$  for different values of  $\sigma_1$ , computed using Equation (12).

One possible mechanism leading to a reduction in the available state space is known as *absolute irreversibility* [34], which is also investigated within the realm of Ising models [43]. Essentially, absolute irreversibility occurs if the probability of a certain “forward” path in the state space is zero, whereas the probability of the corresponding “reverse” path (tracing back in opposite order all the states visited along the forward bath) is non-zero. This phenomenon is experimentally realized when an external field keeps the system first trapped inside a sub-region of the state space, and then the constraint is released and the system can explore the whole state space. This is also the case of, e.g., the free expansion of a gas, in which particles are initially confined by a wall inside a fraction of the volume of the box and then, after the swift removal of the wall, they can eventually fill up the entire volume. We denote by  $\mu_1(x)$  and  $\mu_0(x)$  the probability distributions describing the final and initial states, respectively, which are supported on  $\mathcal{S}_1$  and  $\mathcal{S}_0 \subset \mathcal{S}_1$ . In such cases, Equation (5) yields

$$\langle e^{-\beta W_d(x)} \rangle = \sum_{x \in \mathcal{S}_0} \mu_1(x) < 1. \quad (13)$$

The key aspect of these processes—which encodes much of the physics of the model—is that the state space in the initial state is intrinsically smaller than that in the final state. At times, absolute irreversibility and the state space truncation appear to be interwoven. In Section 3.1 we shall discuss a model in which the two phenomena in fact mingle, while  $\mathcal{I} = \emptyset$ . In this case, one finds that

$$\langle e^{-\beta W_d(x)} \rangle_{\eta_0} < \langle e^{-\beta W_d(x)} \rangle < 1, \quad (14)$$

where the first inequality holds provided that

$$\sum_{x \in \mathcal{S}_0 \setminus \eta_0} \mu_0(x) \sum_{x \in \eta_0} \mu_1(x) < \sum_{x \in \eta_0} \mu_0(x) \sum_{x \in \mathcal{S}_0 \setminus \eta_0} \mu_1(x)$$

Note, indeed, that the latter inequality can be rewritten as

$$0 < -\delta_0 \sum_{x \in \eta_0} \mu_1(x) + (1 - \delta_0) \sum_{x \in \mathcal{S}_0 \setminus \eta_0} \mu_1(x)$$

Next, adding  $\sum_{x \in \eta_0} \mu_1(x)$  on both sides, one obtains

$$\sum_{x \in \eta_0} \mu_1(x) < (1 - \delta_0) \left( \sum_{x \in \eta_0} \mu_1(x) + \sum_{x \in \mathcal{S}_0 \setminus \eta_0} \mu_1(x) \right)$$

hence

$$\frac{1}{1 - \delta_0} \sum_{x \in \eta_0} \mu_1(x) < \sum_{x \in \mathcal{S}_0} \mu_1(x)$$

which yields the result.

### 3. The Models

In this section, we investigate different models for which an explicit analytical solution can be provided. The Jarzynski relation, revisited with the notions of state space reduction and absolute irreversibility introduced in Section 2.2, is used to detect possible discontinuous behaviors of an order parameter as a function of an external control parameter. First, a probabilistic urn model, described by a Markov chain, is considered, in which a threshold parameter allows to tune between different steady states. The second model is the classical 2D Ising model subject to an external magnetic field.



### 3.1. The Urn Model

The modification of the classical Ehrenfest urn model [41,44] that we analyze was originally introduced in [38] to understand billiard dynamics comprising a Maxwell demon [45–47], and is embedded in a long history of urn models studied in statistical mechanics [48–59]. It refers to  $N$  particles initially distributed inside two urns. At each discrete time step, one of the two urns is selected with a probability proportional to the number of particles contained in it, and then one of such particles is moved to the other urn, unless a clogging phenomenon stops it. To represent this process, let  $n(t) \in \mathcal{S} = \{0, 1, \dots, N\}$  denote the number of particles in one selected urn at time  $t \in \mathbb{N}$ , and introduce a parameter  $\lambda \in [0, 1]$  to define a threshold  $T = \lfloor \lambda N \rfloor \in \{0, 1, \dots, N\}$ , which implements the clogging mechanism together with another parameter  $\epsilon \in [0, 1]$ . We then consider the discrete time Markov chain  $\{n(t) : t \geq 0\}$  that obeys the following transition rules:

$$p_{n,n+1}^{(t)}(T) = \frac{N-n}{N} \epsilon_1(T), \quad p_{n,n-1}^{(t)}(T) = \frac{n}{N} \epsilon_2(T) \quad (15)$$

where

$$\epsilon_1(T) = \begin{cases} \epsilon, & n < N - T \\ 1, & n \geq N - T \end{cases}, \quad \epsilon_2 = \begin{cases} 1, & n \leq T \\ \epsilon, & n > T \end{cases} \quad (16)$$

The stationary distributions  $\mu_\lambda(n)$  of the process take the form [38]:

$$\mu_\lambda(n) = \frac{1}{C_\lambda} \binom{N}{n} \begin{cases} \epsilon^n & \text{if } n \leq T \\ \epsilon^T & \text{if } T + 1 \leq n \leq N - T - 1 \\ \epsilon^{N-n} & \text{if } n \geq N - T \end{cases} \quad (17)$$

when  $T < N/2$ , and

$$\mu_\lambda(n) = \frac{1}{C_\lambda} \binom{N}{n} \begin{cases} \epsilon^n & \text{if } n \geq N - T + 1 \\ \epsilon^{N-T+1} & \text{if } N - T + 2 \leq n \leq T - 1 \\ \epsilon^{N-n} & \text{if } n \leq T \end{cases} \quad (18)$$

when  $T > N/2$ , where  $C_\lambda$  is the normalization coefficient associated to  $\mu_\lambda$ . We only analyze the case  $T < N/2$ , as the case  $T > N/2$  trivially follows. The dependence of the stationary state on the parameters  $\epsilon$  and  $\lambda$ , along with the stationary distribution  $\mu(n)$  evaluated in different regions of the parameter space are shown in Figure 2, showing that  $\mu$  concentrates about different states for different values  $\lambda$ . In particular, the left panel of Figure 2 reveals the presence, in parameter space, of equilibrium as well as non-equilibrium phases, portrayed as white and brown regions, respectively [38]. While in the equilibrium region the stationary distribution peaks at  $n = N/2$  (second panel from left), in the non-equilibrium ones the distribution concentrates around the two states  $n = n^* = (\epsilon N - 1)/(\epsilon + 1)$  and  $n = N - n^*$  (right panel). Moreover, in the light blue region, shown in the left panel, the states  $n = n^*$  and  $n = N - n^*$  are metastable while  $n = N/2$  is stable. The distribution corresponding to the latter case shows a global maximum at  $n = N/2$  and two symmetric local maxima at  $n = n^*$  and  $n = N - n^*$  (third panel).

Let us now calculate the normalization coefficient  $C_\lambda$  and analyze its dependence on the threshold parameter to highlight its singular behavior. First, observe that

$$\begin{aligned} C_\lambda &= \sum_{n=0}^T \epsilon^n \binom{N}{n} + \sum_{n=T+1}^{N-T-1} \epsilon^T \binom{N}{n} + \sum_{n=N-T}^N \epsilon^{N-n} \binom{N}{n} = 2 \sum_{n=0}^T \epsilon^n \binom{N}{n} + \sum_{n=T+1}^{N-T-1} \epsilon^T \binom{N}{n} \\ &\approx 2 \sum_{n=0}^N \epsilon^n \binom{N}{n} + \sum_{n=0}^N \epsilon^T \binom{N}{n} = 2(1 + \epsilon)^N + 2^N \epsilon^T \end{aligned} \quad (19)$$

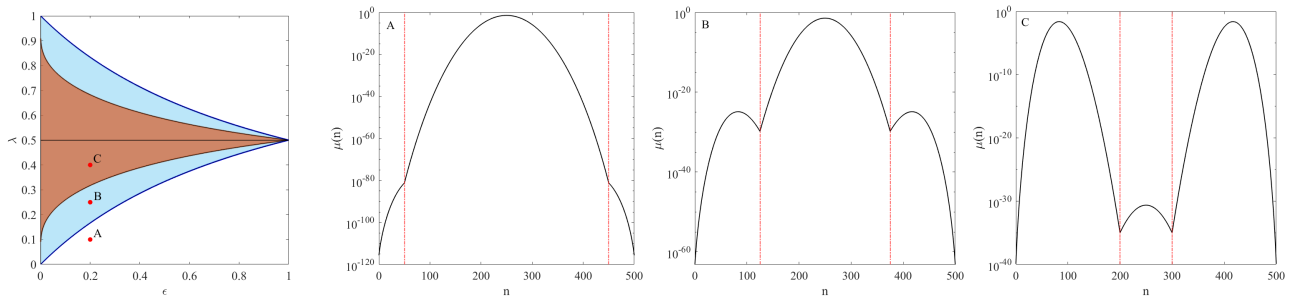


where the last equality follows from the binomial theorem. Therefore, for  $\epsilon < 1$  and  $T > 1$ ,  $C_\lambda$  is an exponentially decaying function of  $\lambda$ . Moreover, the term  $(1 + \epsilon)^N$  is negligible with respect to  $2^N \epsilon^T$  if

$$T < N \log\left(\frac{1 + \epsilon}{2}\right) (\log \epsilon)^{-1}$$

while  $2^N \epsilon^T$  is negligible with respect to  $(1 + \epsilon)^N$  if

$$T > N \log\left(\frac{1 + \epsilon}{2}\right) (\log \epsilon)^{-1}$$



**Figure 2.** Stationary distributions for  $N = 500$ . Right panel: stability regions in the  $\epsilon$ - $\lambda$  parameter space. White corresponds to stable equilibrium ( $n = N/2$ ); cyan to two metastable non-equilibrium states ( $n = n^*, N - n^*$ ) and stable equilibrium; and brown to two stable non-equilibrium states and metastable equilibrium. The remaining three panels, from left to right, plot with vertical log-scale the stationary distributions (17) referring to the points  $A = (0.2, 0.1)$ ,  $B = (0.2, 0.25)$  and  $C = (0.2, 0.4)$ .

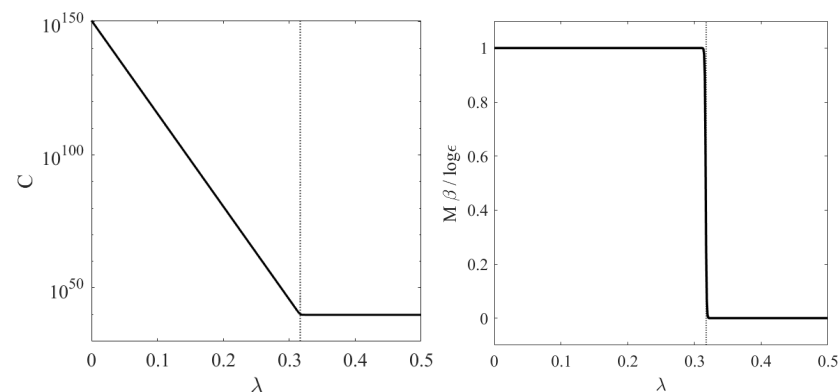
In the large  $N$  limit, the separation between these two regimes becomes sharp, and defines the critical value [38]

$$\lambda^* = \log\left(\frac{1 + \epsilon}{2}\right) (\log \epsilon)^{-1} \tag{20}$$

of a phase transition, implying

$$\begin{aligned} C_{\lambda < \lambda^*} &\approx 2^N \epsilon^T, \\ C_{\lambda > \lambda^*} &\approx 2(1 + \epsilon)^N. \end{aligned} \tag{21}$$

The behavior of  $C_\lambda$  as a function of the normalized threshold  $\lambda$  is illustrated in Figure 3.



**Figure 3.** Left panel: Normalization constant  $C_\lambda$  as a function of  $\lambda$ , Equation (19), with vertical axis in logarithmic scale. Dotted vertical line indicates the critical value  $\lambda^*$  defined in Equation (20). Right panel:  $M(\lambda)\beta / \log \epsilon$  obtained through Equation (23), with dotted vertical line indicating the critical value  $\lambda^*$ . Both plots are obtained for  $N = 500$ ,  $\beta = 1$  and  $\epsilon = 0.2$ .

Mimicking the theory of phase transitions, let us now look at the first derivative with respect to  $\lambda$  of the quantity  $\mathcal{F} = -\beta^{-1} \log C$ :

$$M := -\frac{1}{N} \frac{\partial}{\partial \lambda} \mathcal{F} = -\frac{1}{N} \frac{\partial}{\partial \lambda} \left( -\frac{1}{\beta} \log C(\lambda) \right) \quad (22)$$

which, using Equation (19), and taking large  $N$  so that  $T \approx \lambda N$ , writes

$$M = \frac{\log \epsilon}{\beta} \frac{2^N \epsilon^{\lambda N}}{(1 + \epsilon)^N + 2^N \epsilon^{\lambda N}} = \frac{\log \epsilon}{2\beta} \left[ 1 + \tanh \left( \frac{N}{2} (\lambda - \lambda^*) \log \epsilon \right) \right] \quad (23)$$

The graph of  $M(\lambda)$ , which acts as an order parameter for this model, is shown in the right panel of Figure 3 for  $N = 500$ . At such moderate  $N$ , hence even more sharply in the large  $N$  limit,  $M$  undergoes a jump discontinuity at  $\lambda = \lambda^*$ , which in statistical mechanics is typically associated with first-order phase transitions.

Suppose now that the threshold depends on time, so that  $\lambda = \lambda(t)$ , with  $t = 0, 1$ , constitutes a protocol in accord with the theory illustrated in Section 2.1, and  $\mu_{\lambda(t)}$  is understood as the stationary distribution corresponding to a fixed value of  $t$  [31]. In particular, take the one-step transformation expressed by the protocol  $\lambda_0 \rightarrow \lambda_1$ , with  $\lambda_0 \equiv \lambda(0) = 0$  and  $\lambda_1 \equiv \lambda(1) \in [0, 0.5]$ . The associated stationary distributions can be computed using Equations (17) and (21). Specifically, denoting by  $\mu_0(n) \equiv \mu_{\lambda_0}(n)$  the distribution evaluated at  $\lambda = \lambda_0$ , we find

$$\mu_0(n) = \frac{1}{2^N} \binom{N}{n},$$

whereas  $\mu_1(n) \equiv \mu_{\lambda_1}(n)$  depends on the final value  $\lambda_1$  of the protocol. As discussed earlier in Section 2.2, a summation over all possible initial states  $n \in \mathcal{S}_0 = \{0, 1, \dots, N\}$  in Equation (10), which corresponds to no state space reduction ( $\bar{\mu} = 0$ ), yields the JE, regardless of the chosen value of  $\lambda_1$ . Specifically, for  $\lambda^* < \lambda_1 \leq 0.5$ , two symmetric steady states  $n = n^*$  and  $n = N - n^*$  appear. Indeed, we have  $\mathcal{S}_0 = \mathcal{S}_1 = \{0, 1, \dots, N\}$ , and Equation (5) applies, independently of  $\lambda_1$ . Therefore, the JE does not reveal the jump discontinuity portrayed in the right panel of Figure 3, because it is not affected by that. The situation is different if the state space is reduced to a subset  $\eta_0$  by a cutoff probability  $\bar{\mu} > 0$ , see Equation (9). The left panel of Figure 4 shows indeed that  $\langle e^{-\beta W_d} \rangle_{\eta_0} \simeq 1$  for  $\lambda_1 < \lambda^*$ , while a sudden drop to a value  $\langle e^{-\beta W_d} \rangle_{\eta_0} < 1$  occurs for  $\lambda_1 > \lambda^*$ . In fact,  $\bar{\mu}$  as small as  $10^{-20}$  suffices to get  $\langle e^{-\beta W_d} \rangle_{\eta_0} \simeq 0$  for  $\lambda_1 > \lambda^*$ . This behavior signals the sudden separation of the state space regions occupied by the sets  $\eta_0$  and  $\eta_1$ , in Equation (10), occurring at  $\lambda_1 = \lambda^*$ . While for  $\lambda_1 < \lambda^*$  the sets  $\eta_0$  and  $\eta_1$  overlap consistently, for  $\lambda_1 > \lambda^*$  one can safely consider  $\mathcal{I} = \emptyset$ , provided  $\bar{\mu}$  is large enough. Thus, following the discussion of Section 2.2, for  $\lambda_1 > \lambda^*$  one expects not only that  $\langle e^{-\beta W_d} \rangle_{\eta_0} < 1$ , but also that, this average decreases for growing  $\bar{\mu}$ . This is indeed demonstrated in the left panel of Figure 4.

Consider now the one-step transformations with  $\lambda_0 = 0.5$ , which is larger than  $\lambda^*$ , and any  $\lambda_1 \in [0, 1]$ . In this case, the translational symmetry is broken because the initial condition selects only one of the two non-equilibrium states, say, the one with  $n = n^*$ . Then, taking  $T = N/2$  in Equation (17) yields

$$\mu_{\lambda=0.5} = \frac{1}{C_{0.5}} \binom{N}{n} \epsilon^n, \quad n < N/2.$$

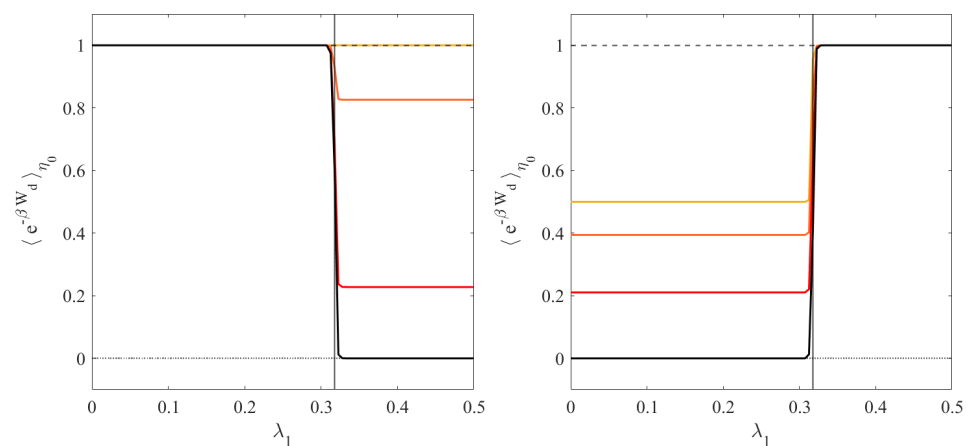
where the second row of Equation (21) turns  $C_{0.5} \approx (1 + \epsilon)^N$ , because only half of the state space  $\mathcal{S}$  is involved. One obtains

$$\mu_{\lambda(0)}(n) = \begin{cases} \frac{\epsilon^n}{(1+\epsilon)^N} \binom{N}{n} & \text{for } n < N/2 \\ 0 & \text{for } n \geq N/2 \end{cases}$$

which is supported on  $\mathcal{S}_0 = \{0, \dots, N/2 - 1\}$ , while the support of the distribution  $\mu_{\lambda(1)}$  is  $\mathcal{S}_1 = \mathcal{S}_0$  for  $\lambda^* < \lambda_1 \leq 0.5$  (non-equilibrium phase) and  $\mathcal{S}_1 = \{0, \dots, N\}$  for  $0 \leq \lambda_1 < \lambda^*$  (equilibrium phase). Consequently, Equation (5) holds for  $\lambda^* < \lambda_1 \leq 0.5$  whereas Equation (13) applies for  $0 \leq \lambda_1 < \lambda^*$ , because  $\mathcal{S}_0 \subset \mathcal{S}_1$ , which implies:

$$\begin{aligned} \langle e^{-\beta W_d(n)} \rangle &= \sum_{n \in \mathcal{S}_0} \mu_1(n) = \sum_{n=0}^{T-1} \frac{\epsilon^n}{2^N \epsilon^T} \binom{N}{n} + \sum_{n=T}^{N/2-1} \frac{1}{2^N} \binom{N}{n} \\ &\simeq \sum_{n=T}^{N/2-1} \frac{1}{2^N} \binom{N}{n} \simeq 1/2, \end{aligned} \tag{24}$$

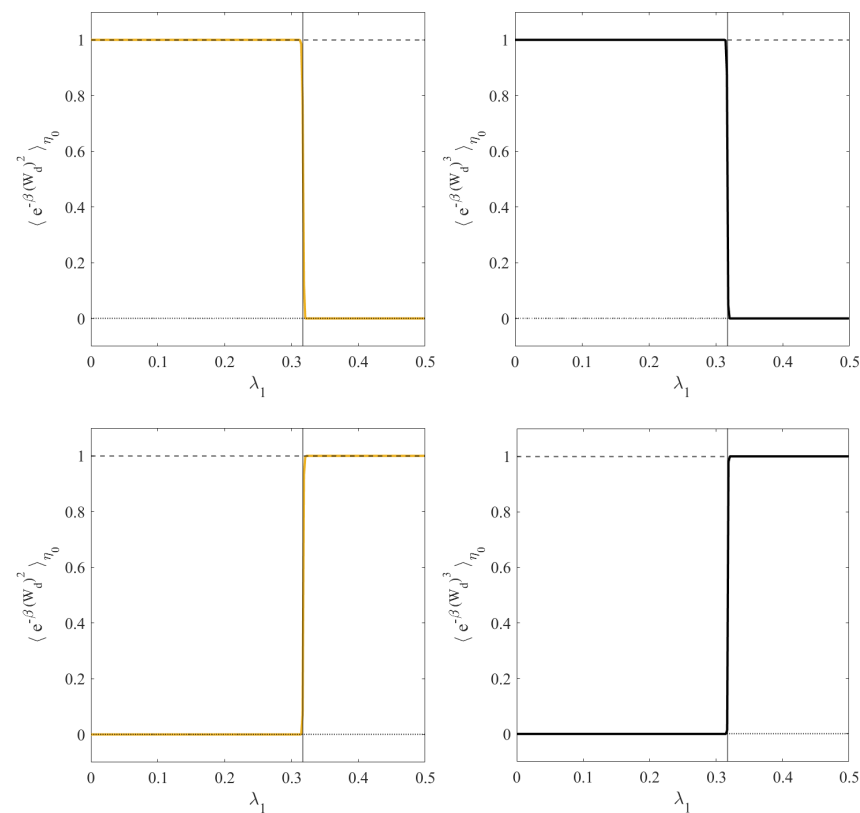
where the approximations turn exact for  $T = 0$  and  $N \rightarrow \infty$ . The violation of the JE stems from the “absolute irreversibility” of the process  $\lambda_0 \rightarrow \lambda_1$ . This is illustrated by the numerical results reported in the right panel of Figure 4, which also clarifies the effect of reducing the state space with a cutoff probability  $\bar{\mu}$ . The effect is a  $\bar{\mu}$  dependent further decrease of the average of  $e^{-\beta W_d}$  for  $0 \leq \lambda_1 < \lambda^*$ , visible in the left panel of Figure 4, cf. Equation (14).



**Figure 4.** Numerical estimation of  $\langle e^{-\beta W_d} \rangle_{\eta_0}$ , as a function of  $\lambda_1$ , given by Equation (10). The yellow, orange, red, and black lines correspond to cutoffs  $\bar{\mu} = 0, 10^{-60}, 10^{-50}, 10^{-20}$ , respectively. The vertical solid black lines lie at  $\lambda = \lambda^*$ . The dashed black lines correspond to Equation (5).  $N, \beta$  and  $\epsilon$  are set to the same values considered in Figure 3. **Left panel:** The process starts in a homogeneous state with  $\lambda_0 = 0$ . The yellow line is always at 1; the phase transition is only revealed by the truncated distributions. **Right panel:** The process starts in a non-homogeneous state with  $\lambda_0 = 0.5$ . The values for  $\bar{\mu} = 0$  and  $\lambda < \lambda^*$  do not equal 1, as prescribed by Equation (24). Even the full distribution reveals the phase transition.

We conclude this Section observing that different exponential variables may be used to efficiently identify phase transitions. For instance, we considered  $\exp\{-\beta(W_d)^p\}$ , with  $p \in \mathbb{N}$ , and numerically studied the cases  $p = 2$  and  $p = 3$ . The numerical evaluation of  $\langle \exp\{-\beta(W_d)^p\} \rangle_{\eta_0}$  for the two different values of  $p$  and for different initial values of  $\lambda_0$  is represented in Figure 5. It is important to note that, unlike the standard expression, corresponding to  $p = 1$ , the quantity  $\langle e^{-\beta(W_d)^2} \rangle_{\eta_0}$  displays a sharp transition at  $\lambda = \lambda^*$  even for  $\bar{\mu} = 0$ , and not only for the truncated distributions. Moreover,  $(W_d)^2$  is a quantity substantially different from  $W_d$ , because it only takes non-negative values, while negative values are fundamental for the usual fluctuation relations. This indicates that our approach may result effective with variables of quite different nature. Let us also note that taking higher odd powers of  $W_d$ , the calculation of the full state space averages may become prohibitive, from a computational point of view. Indeed, the negative values of  $W_d$ , raised to a large odd power and exponentiated, may result in being non-computable. However,

our method based on truncated distributions may remain effective (Figure 5 right panel) for  $p = 3$ .



**Figure 5.** Numerical estimation of  $\langle e^{-\beta(W_d)^p} \rangle_{\eta_0}$  as a function of  $\lambda_1$  for  $p = 2$  and  $p = 3$  (left and right column, respectively), with the parameters  $N$ ,  $\beta$ , and  $\epsilon$  set to the same values as in Figure 3. The plots refer to a process starting in a homogeneous state with  $\lambda_0 = 0$  (top row) or in a non-homogeneous state with  $\lambda_0 = 0.5$  (bottom row). In the left panels, the yellow line refers to the cutoff  $\bar{\mu} = 0$ , whereas the curves corresponding to cutoff values  $\bar{\mu} > 0$  are not shown in the plots, as they are graphically indistinguishable from the yellow line. In the right panels, the solid black line refers to the cutoff  $\bar{\mu} = 10^{-15}$ .

### 3.2. The 2D Ising Model

The second example we tackle is the classical 2D Ising model under an external magnetic field  $h$ , defined on a 2D lattice  $\Lambda = \{1, \dots, L\}^2$ , with the state space  $\mathcal{S} = \{-1, +1\}^\Lambda$ . Here,  $\beta = (k_B T)^{-1}$  denotes the inverse temperature. Let  $\sigma \in \mathcal{S}$  denote a generic spin configuration, and introduce the Ising Hamiltonian, which reads

$$H(\sigma) = -J \sum_{\substack{\langle j,k \rangle \\ j,k \in \Lambda}} \sigma_j \sigma_k - h \sum_{j \in \Lambda} \sigma_j,$$

where  $\sigma_j = \{-1, +1\}$  for  $j \in \Lambda$ ,  $J > 0$  denotes the ferromagnetic coupling constant and the first sum is taken over pairs of adjacent spins (every pair is counted only once). Periodic boundary conditions are imposed along the horizontal and vertical directions. The model exhibits a critical value  $\beta_c = \ln(1 + \sqrt{2})/2J$ , marking the transition from a disordered (paramagnetic) to an ordered (ferromagnetic) phase. We then let the magnetic field depend on time and define a protocol  $\lambda(t) = h(t)$ . In [43], a similar problem is investigated, taking  $\lambda(t) = J(t)$ . We denote by  $H_i(\sigma)$  and  $H_f(\sigma)$  the Hamiltonians evaluated at the initial and the final values of the protocol, respectively, and we iden-

tify  $\mu_k(\sigma) = \exp(-\beta H_k(\sigma))/Z_k$  as the Gibbs measure at the inverse temperature  $\beta$ , with  $k = 0, 1$  and  $Z_k = \sum_{\sigma \in \mathcal{S}} \exp(-\beta H_k(\sigma))$ . So both  $\mu_0$  and  $\mu_1$  are supported on  $\mathcal{S}$ .

We now introduce the reduced state spaces  $\eta_0 = \{\sigma | \mu_0(\sigma) \geq \bar{\mu}\}$  and  $\eta_1 = \{\sigma | \mu_1(\sigma) \geq \bar{\mu}\}$  and we correspondingly denote  $\bar{Z}_0 = \sum_{\sigma \in \eta_0} \exp(-\beta H_0(\sigma))$  and  $\bar{Z}_1 = \sum_{\sigma \in \eta_1} \exp(-\beta H_1(\sigma))$ . With the notation of Section 2.2, we have

$$1 - \delta_k = \sum_{\sigma \in \eta_k} \mu_k(\sigma) = \frac{\sum_{\sigma \in \eta_k} \exp(-\beta H_k(\sigma))}{\sum_{\sigma \in \mathcal{S}} \exp(-\beta H_k(\sigma))} = \frac{\bar{Z}_k}{Z_k}, \quad k = 0, 1, \quad (25)$$

which yields

$$\frac{\bar{Z}_0}{\bar{Z}_1} = \frac{Z_0 (1 - \delta_0)}{Z_1 (1 - \delta_1)}.$$

Furthermore, for a function  $f(\sigma) : \mathcal{S} \rightarrow \mathbb{R}$ , we set

$$\langle f \rangle_{\eta_0} = \sum_{\sigma \in \eta_0} f(\sigma) \frac{e^{-\beta H_0(\sigma)}}{\bar{Z}_0} = \frac{1}{1 - \delta_0} \sum_{\sigma \in \eta_0} f(\sigma) \mu_0(\sigma). \quad (26)$$

Then, taking

$$f(\sigma) = e^{-\beta W_d} = e^{-\beta(W - \Delta F)} = e^{-\beta(H_1(\sigma) - H_0(\sigma))} \frac{Z_0}{Z_1}, \quad \text{where } \beta \Delta F = \log \frac{Z_0}{Z_1},$$

we find

$$f(\sigma) \mu_0(\sigma) = \frac{e^{-\beta H_1(\sigma)} Z_0}{Z_0 Z_1} = \mu_1(\sigma),$$

and

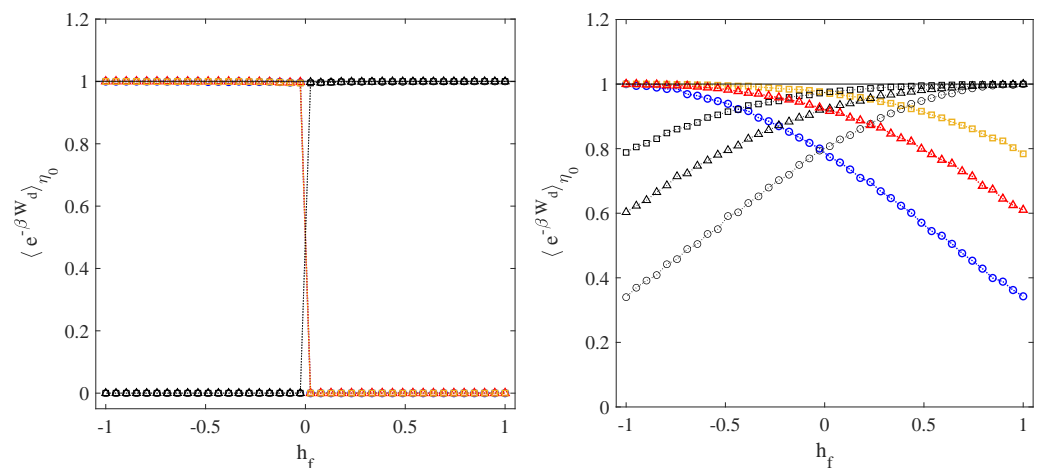
$$\langle e^{-\beta W_d} \rangle_{\eta_0} = \frac{1}{1 - \delta_0} \sum_{\sigma \in \eta_0} \mu_1(\sigma). \quad (27)$$

The latter equation coincides with Equation (10) and recovers Equation (14) [43]. We are interested in studying the application of Equation (27) to evolutions in which the magnetic field changes from  $\lambda_0 = h_0$  to  $\lambda_1 = h_1$ . Setting  $J = 1$ , which yields  $\beta_c \approx 0.4406$ , we consider different scenarios. We first choose  $\beta > \beta_c$ , so that the 2D Ising model undergoes a first-order phase transition when the magnetic field  $h$  changes sign, and we study the behavior of  $\langle e^{-\beta W_d} \rangle_{\eta_0}$  as a function of  $h_1$ . Later, we study the case with  $\beta < \beta_c$ , where no phase transition takes place. The computation of  $\langle e^{-\beta W_d} \rangle_{\eta_0}$  can be conveniently carried out by sampling configurations  $\sigma \in \mathcal{S}$ , with a prescribed magnetization density  $m = N^{-1} \sum_{j \in \Lambda} \sigma_j$ , from the distribution [43]:

$$P(m) = \sum_{\sigma \in \mathcal{S}} \mu(\sigma) \delta \left( m - \frac{1}{N} \sum_{j \in \Lambda} \sigma_j \right)$$

Figure 6 shows that  $\langle e^{-\beta W_d} \rangle = 1$  when  $\bar{\mu} = 0$ , for both  $\beta < \beta_c$  and  $\beta > \beta_c$ , cf. Equation (5). Therefore, the phase transition in the Ising model is not detected by the JE in one-step transformations. For the reduced state spaces, corresponding to cutoffs  $\bar{\mu} > 0$ , the situation changes, as evidenced by the left panel of Figure 6. The phase transition is properly captured by  $\langle e^{-\beta W_d} \rangle_{\eta_0}$ , which drops abruptly from the value 1 to a value close to 0 when crossing  $h_1 = 0$ , for  $h_0 = -1$  and with  $h_1 \in [-1, 1]$ . An analogous behavior is also observed when  $h_0 = 1$  and  $h_1$  is varied in the same interval. It is worth recalling that, in the Curie–Weiss approximation of the Ising model, the value of the magnetization density  $m$  is obtained from the solution of the mean field equation  $m = \tanh[\beta(m + h)]$ , which, for  $\beta > \beta_c$ , gives rise to metastable branches [60]. One such branch corresponds to

positive values of  $m$  in the interval  $[-h_c, 0]$ , and another branch corresponds, symmetrically, to negative values of  $m$  in the interval  $[0, h_c]$ , with  $h_c > 0$ . This is not visible in the left panel of Figure 6, which shows that for  $h_0 = \pm 1$  the discontinuity occurs in both cases sharply at  $h_1 = 0$ . This is a consequence of our choice of initial datum for the Monte Carlo simulations, that mitigates the effects of the two metastable branches. More precisely, as initial datum we took  $\sigma_j = -1$  for  $h_1 < 0$  and  $\sigma_j = 1$  for  $h_1 > 0$ , for all  $j \in \Lambda$ , producing a distribution  $P(m)$  sharply concentrated near  $m = -1$  for  $h_1 < 0$  and near  $m = 1$  for  $h_1 > 0$ . The discontinuity at  $h_1 = 0$ , therefore, stems from Equation (12), which rules the behavior of  $\langle e^{-\beta W_d} \rangle_{\eta_0}$  for  $\beta > \beta_c$ , where one can assume  $\mathcal{I} = \emptyset$ . In the left panel, it is also interesting to note that, for  $\beta = 2.5 > \beta_c$ , the different sets of data points corresponding to different values of  $\bar{\mu}$  fully overlap with one another and display the same discontinuity at  $h_1 = 0$ .



**Figure 6.**  $\langle e^{-\beta W_d} \rangle_{\eta_0}$  as a function of  $h_f$ , Equation (27), from Monte Carlo simulations of a 2D Ising model on a square grid of linear size  $L = 10$ . Here,  $J = 1$ , hence  $\beta_c \simeq 0.4406$ , and periodic boundary conditions are applied in both horizontal and vertical directions. Simulations have been performed with  $h_i = -1$  (colored symbols) and  $h_i = 1$  (black symbols), for  $\beta = 2.5 > \beta_c$  (left panel) and  $\beta = 0.1 < \beta_c$  (right panel). Circles, triangles, and squares correspond to  $\bar{\mu} = (\max\{\mu_0\} + \max\{\mu_1\}) / (2k)$  with  $k = 5, 25, 100$ , respectively. The horizontal black line at 1 represents in both panels the case with  $\bar{\mu} = 0$ .

For  $\beta < \beta_c$  the phase transition in the Ising model is instead absent, and the magnetization density  $m$  solving the mean field equation becomes a continuous function of  $h$ , which vanishes at  $h = 0$ . As a result, the data points shown in the right panel of Figure 6 show that  $\langle e^{-\beta W_d} \rangle_{\eta_0}$  no longer presents a sharp discontinuity at  $h_1 = 0$ , as this quantity decreases in a continuous fashion while increasing the difference  $|h_0 - h_1|$ . The effect of varying the value of  $\bar{\mu}$  is clearly visible in the right panel. Again, the JE is not affected by the variation of  $h_1$ .

#### 4. Conclusions

In this study, we propose one way of benefiting from incomplete statistics or finite-size effects, which result in a reduction in the state space. As this is often an unavoidable fact, we show that it can provide valuable insight, if associated with exponential observables. In particular, we have considered the Jarzynski Equality and its instantaneous transformation variant (also known as Free Energy Perturbation), to identify the critical values of the parameters corresponding to a first-order phase transition. Smooth variations of states have also been characterized. This is due to the sensitivity of exponentials of extensive quantities, when averaged over the available subsets of the system state space, which is usually considered a hindrance. Interestingly enough, our study shows that reducing the amount of available information about the system can enhance our understanding of its critical behaviors. In fact, this is not obtained from a perfect sampling of states. The JE, in

particular, results are insensitive to parameters variations, hiding interesting features of the system at hand.

We illustrated our theoretical findings by numerically analyzing two stochastic models: a modified Ehrenfest urn model and a classical 2D Ising model subject to an external magnetic field. In both cases, we observed a jump discontinuity in the quantity  $\langle e^{-\beta W_d} \rangle_{\eta_0}$  occurring at the same value of the control parameter distinguishing different macroscopic phases in the thermodynamic limit of the model, remaining quite far from that limit. Similarly, when the phase transition is absent, as in the case of the Ising model at supercritical temperatures, the variation of state with the control parameter, is still revealed by a continuous variation of  $\langle e^{-\beta W_d} \rangle_{\eta_0}$ .

Studying the behavior of powers of  $W_d$  we also found that our approach works even with variables different from those of fluctuation theorems. Indeed, it provides a new recipe to identify order parameters in the study of non-equilibrium phase transitions, which is particularly effective with truncated distributions.

Future works are intended to develop this approach, so that incomplete information on a given system may be used beyond the problem of discovering phase transitions.

**Author Contributions:** Conceptualization, M.C., A.D.F. and L.R.; methodology, M.C., A.D.F. and L.R.; software, A.D.F.; validation, M.C., A.D.F. and L.R.; formal analysis, M.C., A.D.F. and L.R.; investigation, M.C., A.D.F. and L.R.; data curation, A.D.F.; writing—original draft preparation, M.C., A.D.F. and L.R.; writing—review and editing, M.C., A.D.F. and L.R.; visualization, A.D.F.; supervision, M.C. and L.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** L.R. gratefully acknowledges support from the Italian Ministry of University and Research (MUR) through the grant PRIN2022-PNRR project (No. P2022Z7ZAJ) “A Unitary Mathematical Framework for Modelling Muscular Dystrophies” (CUP: E53D23018070001). This work has been performed under the auspices of Italian National Group of Mathematical Physics (GNFM) of the Istituto Nazionale di Alta Matematica (INdAM).

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Bennett, C.H. Efficient estimation of free energy differences from monte carlo data. *J. Comput. Phys.* **1976**, *22*, 245–268. [[CrossRef](#)]
- Widom, B. Some topics in the theory of fluids. *J. Chem. Phys.* **1963**, *39*, 2808–2812. [[CrossRef](#)]
- Zwanzig, R.W. High-temperature equation of state by a perturbation method. ii. Polar gases. *J. Chem. Phys.* **1955**, *23*, 1915–1922. [[CrossRef](#)]
- Evans, D.J.; Searles, D.J. The fluctuation theorem. *Adv. Phys.* **2002**, *52*, 1529. [[CrossRef](#)]
- Evans, D.J.; Cohen, E.G.D.; Morriss, G.P. Probability of second law violations in shearing steady states. *Phys. Rev. Lett.* **1993**, *71*, 2401. [[CrossRef](#)] [[PubMed](#)]
- Jepps, O.G.; Rondoni, L. A dynamical-systems interpretation of the dissipation function, T-mixing and their relation to thermodynamic relaxation. *J. Phys. A Math. Theor.* **2016**, *49*, 154002. [[CrossRef](#)]
- Jarzynski, C. Non-equilibrium equality for free energy differences. *Phys. Rev. Lett.* **1997**, *78*, 2690. [[CrossRef](#)]
- Crooks, G.E. Non-equilibrium measurements of free energy differences for microscopically reversible markovian systems. *J. Stat. Phys.* **1998**, *90*, 1481–1487. [[CrossRef](#)]
- Crooks, G.E. Path-ensemble averages in systems driven far from equilibrium. *Phys. Rev. E* **2000**, *61*, 2361–2366. [[CrossRef](#)]
- Hatano, T.; Sasa, S. Steady-state thermodynamics of langevin systems. *Phys. Rev. Lett.* **2001**, *86*, 3463. [[CrossRef](#)]
- Marconi, U.M.B.; Puglisi, A.; Rondoni, L.; Vulpiani, A. Fluctuation–dissipation: Response theory in statistical physics. *Phys. Rep.* **2008**, *461*, 111–195. [[CrossRef](#)]
- Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **2012**, *75*, 126001. [[CrossRef](#)]
- Cohen, E.G.D.; Mauzerall, D. A note on the jarzynski equality. *J. Stat. Mech. Theory Exp.*, **2004**, *2004*, P07006. [[CrossRef](#)]
- Vilar, J.M.G.; Rubi, J.M. Failure of the work-hamiltonian connection for free-energy calculations. *Phys. Rev. Lett.* **2008**, *100*, 020601. [[CrossRef](#)] [[PubMed](#)]
- Talkner, P.; Hänggi, P. Open system trajectories specify fluctuating work but not heat. *Phys. Rev. E* **2016**, *94*, 022143. [[CrossRef](#)] [[PubMed](#)]
- Peliti, L. On the work—Hamiltonian connection in manipulated systems. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, P05002. [[CrossRef](#)]
- Vilar, J.M.G.; Rubi, J.M. Vilar and rubi reply. *Phys. Rev. Lett.* **2008**, *101*, 098902. [[CrossRef](#)]



18. Ciliberto, S. Experiments in stochastic thermodynamics: Short history and perspectives. *Phys. Rev. X* **2017**, *7*, 021051. [[CrossRef](#)]
19. Liphardt, J.; Dumont, S.; Smith, S.B.; Tinoco, I.; Bustamante, C. Equilibrium information from non-equilibrium measurements in an experimental test of Jarzynski's equality. *Science* **2002**, *296*, 1832–1835. [[CrossRef](#)]
20. Monge, A.M.; Manosas, M.; Ritort, F. Experimental test of ensemble inequivalence and the fluctuation theorem in the force ensemble in dna pulling experiments. *Phys. Rev. E* **2018**, *98*, 032146. [[CrossRef](#)]
21. Peliti, L. Comment on "failure of the work-hamiltonian connection for free-energy calculations". *Phys. Rev. Lett.* **2008**, *101*, 098903. [[CrossRef](#)] [[PubMed](#)]
22. Bustamante, C.; Liphardt, J.; Ritort, F. The Non-equilibrium Thermodynamics of Small Systems. *Phys. Today* **2005**, *58*, 43–48. [[CrossRef](#)]
23. Lua, R.C.; Grosberg, A.Y. Practical applicability of the Jarzynski relation in statistical mechanics: A pedagogical example. *J. Phys. Chem. B* **2005**, *109*, 6805–6811. [[CrossRef](#)] [[PubMed](#)]
24. Crooks, G.E.; Jarzynski, C. Work distribution for the adiabatic compression of a dilute and interacting classical gas. *Phys. Rev. E* **2007**, *75*, 021116. [[CrossRef](#)] [[PubMed](#)]
25. Jarzynski, C. Stochastic and macroscopic thermodynamics of strongly coupled systems. *Phys. Rev. X* **2017**, *7*, 011008. [[CrossRef](#)]
26. Hansen, N.; Van Gunsteren, W. F. Practical aspects of free-energy calculations: A review. *J. Chem. Theory Comput.* **2014**, *10*, 2632–2647. [[CrossRef](#)]
27. Torrie, G.M.; Valleau, J.P. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187–199. [[CrossRef](#)]
28. Pohorille, A.; Jarzynski, C.; Chipot, C. Good practices in free-energy calculations. *J. Phys. Chem. B* **2010**, *114*, 10235–10253. [[CrossRef](#)]
29. Wu, D.; Kofke, D.A. Phase-space overlap measures. ii. Design and implementation of staging methods for free-energy calculations. *J. Chem. Phys.* **2005**, *123*, 084109. [[CrossRef](#)]
30. Colangeli, M.; Di Francesco, A.; Rondoni, L. Finite Reservoirs Corrections to Hamiltonian Systems Statistics and Time Symmetry Breaking. *Symmetry* **2023**, *15*, 1268. [[CrossRef](#)]
31. Hack, P.; Gottwald, S.; Braun, D.A. Jarzynski's Equality and Crooks' Fluctuation Theorem for General Markov Chains with Application to Decision-Making Systems. *Entropy* **2022**, *42*, 1731. [[CrossRef](#)]
32. Ge, H.; Qian, M. Generalized Jarzynski's equality in inhomogeneous Markov chains. *J. Math. Phys.* **2007**, *48*, 053302. [[CrossRef](#)]
33. Horowitz, J.M.; Vaikuntanathan, S. Non-equilibrium detailed fluctuation theorem for repeated discrete feedback. *Phys. Rev. E* **2010**, *82*, 061120. [[CrossRef](#)] [[PubMed](#)]
34. Murashita, Y.; Funo, K.; Ueda, M. Non-equilibrium equalities in absolutely irreversible processes. *Phys. Rev. E* **2014**, *90*, 042110. [[CrossRef](#)] [[PubMed](#)]
35. Murashita, Y.; Ueda, M. Gibbs paradox revisited from the fluctuation theorem with absolute irreversibility. *Phys. Rev. Lett.* **2017**, *118*, 060601. [[CrossRef](#)] [[PubMed](#)]
36. Funo, K.; Murashita, Y.; Ueda, M. Quantum non-equilibrium equalities with absolute irreversibility. *New J. Phys.* **2015**, *17*, 075005. [[CrossRef](#)]
37. Manikandan, S.K.; Elouard, C.; Jordan, A.N. Fluctuation theorems for continuous quantum measurements and absolute irreversibility. *Phys. Rev. A* **2019**, *99*, 022117. [[CrossRef](#)]
38. Cirillo, E.N.; Colangeli, M.; Muntean, A.; Richardson, O.; Rondoni, L. Deterministic reversible model of non-equilibrium phase transitions and stochastic counterpart. *J. Phys. A Math. Theor.* **2020**, *53*, 305001. [[CrossRef](#)]
39. Stroock, D.J. *An Introduction to Markov Processes*; Springer: Berlin/Heidelberg, Germany, 2005.
40. Billingsley, P. *Probability and Measures*; John Wiley & Sons: Hoboken, NJ, USA, 1995.
41. Feller, W. *An Introduction to Probability. Theory and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 1968.
42. Sekimoto, K. *Stochastic Energetics*; Springer: Berlin/Heidelberg, Germany, 2010.
43. Hoang, D.-T.; Venkatesh, B.P.; Han, S.; Jo, J.; Watanabe, G.; Choi, M.-S. Scaling law for irreversible entropy production in critical systems. *Sci. Rep.* **2016**, *6*, 27603. [[CrossRef](#)]
44. Kac, M. *Probability and Related Topics in Physical Sciences, 1 of Lectures in Applied Mathematics*; Interscience Publishers: Geneva, Switzerland, 1959.
45. Cirillo, E.N.; Colangeli, M.; Richardson, O.; Rondoni, L. Deterministic model of battery, uphill currents, and nonequilibrium phase transitions. *Phys. Rev. E* **2021**, *103*, 032119. [[CrossRef](#)]
46. Cirillo, E.N.; Colangeli, M.; Di Francesco, A.; Kröger, M.; Rondoni, L. Transport and nonequilibrium phase transitions in polygonal urn models. *Chaos Interdiscip. J. Nonlinear Sci.* **2022**, *32*, 093127. [[CrossRef](#)] [[PubMed](#)]
47. Cirillo, E.N.; Colangeli, M.; Kröger, M.; Rondoni, L. Nonequilibrium phase transitions in feedback-controlled three-dimensional particle dynamics. *Phys. Rev. Res.* **2023**, *5*, 043063. [[CrossRef](#)]
48. Godreche, C.; Luck, J.M. Non-equilibrium dynamics of urn models. *J. Phys. Condens. Mat.* **2002**, *14*, 1601–1615. [[CrossRef](#)]
49. Dette, H. On a generalization of the Ehrenfest urn model. *J. Appl. Prob.* **1994**, *31*, 930–939. [[CrossRef](#)]
50. Ohkubo, J.; Yasuda, M.; Tanaka, K. Preferential urn model and nongrowing complex networks. *Phys. Rev. E* **2005**, *72*, 065104. [[CrossRef](#)] [[PubMed](#)]
51. Siegmund, D.; Yakir, B. An urn model of diaconis. *Annal. Prob.* **2005**, *33*, 2036–2042. [[CrossRef](#)]
52. Chen, M.-R.; Hsiao, S.-R. On a generalized q-urn model. *Prob. Eng. Sci.* **2015**, *29*, 99–115. [[CrossRef](#)]

53. Idriss, S. Nonlinear unbalanced urn models via stochastic approximation. *Meth. Comput. Appl. Prob.* **2022**, *24*, 413–430. [[CrossRef](#)]
54. Zhang, L.-X.; Hu, F.; Cheung, S.H.; Chan, W.S. Asymptotic properties of multicolor randomly reinforced Pólya urns. *Adv. Appl. Prob.* **2014**, *46*, 585–602. [[CrossRef](#)]
55. Kao, Y.M.; Luan, P.G. Poincare cycle of a multibox Ehrenfest urn model with directed transport. *Phys. Rev. E* **2003**, *67*, 031101. [[CrossRef](#)]
56. Chen, M.-R.; Kuba, M. On generalized Pólya urn models. *J. Appl. Prob.* **2013**, *50*, 1169–1186. [[CrossRef](#)]
57. Bai, Z.D.; Hu, F.F.; Zhang, L.X. Gaussian approximation theorems for urn models and their applications. *Annal. Appl. Prob.* **2002**, *12*, 1149–1173. [[CrossRef](#)]
58. Klein, M.J. Entropy and the Ehrenfest urn model. *Physica* **1956**, *22*, 569–575. [[CrossRef](#)]
59. Thornblad, E. The dominating colour of an infinite Pólya urn model. *J. Appl. Prob.* **2016**, *53*, 914–924. [[CrossRef](#)]
60. Presutti, E. *Scaling Limits in Statistical Mechanics and Microstructures in Continuum Mechanics*; Springer: Berlin/Heidelberg, Germany, 2009.

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