

Article

Entropy Production Beyond the Diffusion Limit: A Phase Transition in the Two-Allele Wright–Fisher Model

Igor Gornyi ^{1,*} , Ruben Poghosyan ²  and David B. Saakian ³ ¹ Institute for Quantum Materials and Technologies and Institut für Theorie der Kondensierten Materie, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany² School of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; ruben.poghosyan@sjtu.edu.cn³ A.I. Alikhanyan National Science Laboratory (Yerevan Physics Institute) Foundation, 2 Alikhanian Brothers St., Yerevan 375036, Armenia; saakian@yerphi.am

* Correspondence: igor.gornyi@kit.edu

Abstract

We study stationary entropy production as a measure of irreversibility in a two-allele Wright–Fisher population, comparing the finite-population chain with its diffusion, fixed-point Gaussian, and sequential weak-mutation reductions. Whole-generation sampling generates probability-current cycles among population compositions despite the acyclic mutation graph, making the finite chain generically irreversible, unlike its three reversible reductions. For a population of size N , with selection $s = \alpha/N$ and mutation $\mu = \beta/N$, we establish a boundary-localization phase transition in entropy production at $\beta_c = 1/2$. Below it, entropy production is localized near fixation, where the minority allele has $O(1)$ copies; above it, the polymorphic bulk dominates; at criticality, contributions are approximately uniform over logarithmic intervals of distance from fixation. In the $N \rightarrow \infty$ limit, the localization order parameter jumps from one to zero at β_c , while the boundary width needed to capture a fixed fraction diverges as $\beta \rightarrow \beta_c^-$, identifying a mixed-order transition. The entropy-production rate per unit diffusion time, $N\sigma_N$, diverges as $N^{1-2\beta}$ below criticality and as $(\alpha \coth \alpha/6) \ln N$ at criticality, and converges to $\mathcal{C}(\alpha, \beta)$ above it. For $\beta > 1$, we also derive the first regular finite- N correction. The finite chain and its reversible diffusion share the leading drift and variance but differ in their third central moments. Since the diffusion is reversible for every β , the diffusion approximation conceals both finite-population irreversibility and its boundary-localization transition.

Keywords: evolutionary dynamics; entropy production; phase transitions**MSC:** 92D25

Academic Editor: Davide Valenti

Received: 14 September 2026

Revised: 18 October 2026

Accepted: 22 October 2026

Published: 27 October 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Connections between evolutionary dynamics and statistical physics have been explored for several decades, including approaches based on thermodynamic analogies, Hamilton–Jacobi methods, and quasispecies dynamics [1–15]. In deterministic settings, mean fitness and related Lyapunov quantities play roles analogous to thermodynamic potentials. For stochastic evolutionary dynamics, however, the natural question is not only whether a Lyapunov function exists [16,17], but also whether the stationary dynamics is reversible. In a stationary Markov process, this irreversibility is quantified by the entropy-production rate (EPR), which measures the asymmetry between forward and

time-reversed trajectories and vanishes when detailed balance holds [18–21]. Evolutionary time-reversal asymmetry has also been studied through fitness flux [22], which differs from the Schnakenberg EPR considered here.

Evolutionary theory also uses entropy-related quantities to characterize population states: the Shannon entropy of genotype frequencies measures genetic diversity, while relative entropies between population-state distributions and fitness-based Lyapunov functionals characterize states and their relaxation. Neither determines whether stationary evolutionary histories are statistically distinguishable from their time reversals. The EPR is a relative entropy, but between the stationary forward and time-reversed two-time distributions, rather than between population-state distributions. At stationarity, the EPR equals the average logarithmic asymmetry of forward and backward transition probabilities. The stochastic EPR also does not automatically represent physical heat dissipation or metabolic entropy production. Interpreting this formal quantity as entropy transferred to physical reservoirs requires a microscopic thermodynamic description of the transitions, for example, through local detailed balance. Evolutionary transition probabilities and sampling rules specify how population composition changes, but do not by themselves determine the heat or chemical work associated with reproduction and mutation. The resulting EPR thus quantifies the irreversibility of population histories without assigning a thermodynamic heat to mutation or selection.

A deterministic mutation–selection equation specifies the evolution of genotype frequencies but does not determine a unique stochastic process realizing that evolution. Different stochastic transition rules can produce the same deterministic drift but different stationary probability currents. Markov representations of deterministic quasispecies dynamics have been constructed [17,23,24], but the stochastic transitions induced by these embeddings need not coincide with those generated by reproduction and sampling in a finite population. Thus, the deterministic drift alone neither determines the finite-population EPR nor guarantees that a stochastic reduction preserves it. Entropy production depends on the level at which a stochastic process is resolved. Coarse-graining can hide probability currents and lower the EPR inferred from the reduced dynamics, while going from a master equation to a Fokker–Planck equation can discard entropy production carried by microscopic transitions [25–28]. The reductions considered below are related to this general issue but involve different limiting procedures: the diffusion approximation involves a singular large-population limit ($N \rightarrow \infty$, where N is the population size), a rescaling of time, and a Kramers–Moyal truncation to a Fokker–Planck equation.

In this paper, we address the dependence of evolutionary EPR on the stochastic description by comparing the exact finite- N two-allele Wright–Fisher chain [29,30] with its diffusion approximation, a fixed-point Gaussian approximation, and the sequential weak-mutation origin–fixation process. The finite Wright–Fisher model provides the reference process: we determine which features of its time-reversal asymmetry survive the three reductions. It can be irreversible because whole-generation sampling generates cycles in the population-state graph that can carry nonzero stationary probability currents. By contrast, the stationary fixed-point Gaussian approximation, the one-dimensional Wright–Fisher diffusion, and the two-state sequential origin–fixation process are reversible and have zero stationary EPR. Thus, the irreversibility of the finite Wright–Fisher process is not a property of the underlying two-allele mutation graph alone, but of the population-level transition structure retained by the stochastic description.

Our main analytical result is the large- N behavior of the finite-population EPR in the standard Wright–Fisher diffusion scaling, with selection coefficient $s = \alpha/N$ and mutation probability $\mu = \beta/N$, where α and β are fixed. For $\beta > 1/2$, the finite-population EPR obeys $\sigma_N = \mathcal{C}(\alpha, \beta)/N + o(N^{-1})$, so that the rescaled entropy-production rate on the diffusion

time scale approaches the positive constant $\mathcal{C}(\alpha, \beta)$; for $\beta > 1$, we also obtain the first correction $\mathcal{C}_1(\alpha, \beta)/N^2$ analytically. The limiting one-dimensional diffusion nevertheless has zero stationary EPR for every $\beta > 0$: as shown in Section 3.4, the diffusion discards the finite-generation information responsible for the irreversibility of the Wright–Fisher chain. The sequential weak-mutation limit also yields a reversible process, but through the elimination of polymorphic states, leaving only the two monomorphic genotypes.

The restriction $\beta > 1/2$ in the bulk EPR formula reflects a boundary-localization phase transition (Section 4) at $\beta_c = 1/2$. In the large-population limit $N \rightarrow \infty$, the dominant irreversibility changes from boundary-localized below the transition to bulk-distributed above it, with a marginal regime at criticality. The transition has mixed-order character: the localization order parameter jumps from one to zero at β_c , while the boundary width needed to capture a fixed fraction of the EPR diverges as β approaches β_c from below.

2. Materials and Methods

2.1. Entropy Production and Time Reversal

For a stationary discrete-time Markov chain with transition probabilities W_{ij} and state probabilities π_i , the probability of observing two successive states in the order $i \rightarrow j$ is

$$P_{ij}^F = \pi_i W_{ij}. \quad (1)$$

The probability assigned to the reversed ordered pair is

$$P_{ij}^R = \pi_j W_{ji}. \quad (2)$$

We refer to P^F and P^R as the stationary forward and reversed two-time distributions. The entropy production per step is their relative entropy [21],

$$\sigma = \sum_{i,j} P_{ij}^F \ln \frac{P_{ij}^F}{P_{ij}^R} = \sum_{i,j} \pi_i W_{ij} \ln \frac{\pi_i W_{ij}}{\pi_j W_{ji}} = \frac{1}{2} \sum_{i,j} (\pi_i W_{ij} - \pi_j W_{ji}) \ln \frac{\pi_i W_{ij}}{\pi_j W_{ji}}. \quad (3)$$

It is non-negative and vanishes if and only if detailed balance holds. For any function $\mathcal{F}_{ij}$ of two successive states, we denote its stationary forward average by

$$\langle \mathcal{F} \rangle_F = \sum_{i,j} P_{ij}^F \mathcal{F}_{ij} = \sum_{i,j} \pi_i W_{ij} \mathcal{F}_{ij}. \quad (4)$$

These definitions apply directly to the finite- N Wright–Fisher chain. Throughout the paper, the EPR is used as a dynamical measure of time-reversal asymmetry of the specified stochastic population process: a nonzero value means that stationary population trajectories are statistically distinguishable from their time reversals. For a continuous-time Markov process, such as the sequential origin–fixation dynamics considered below, the analogous EPR expression is a sum over transitions $i \neq j$, with transition probabilities replaced by transition rates.

2.2. The Wright–Fisher Model

Consider a Wright–Fisher population of fixed size N with two alleles ($K = 2$, where K denotes the number of allelic types) [29–32]. Let i be the number of individuals of type 1 in the current generation. The finite-population states are labeled by

$$i = 0, 1, \dots, N. \quad (5)$$

We introduce

$$X_i = \frac{i}{N} \quad (6)$$

as the frequency of type 1 in the population composition $(i, N - i)$. We write X without a subscript when treating the allele frequency as a variable, taking the discrete values X_i in the finite chain and varying continuously over $[0, 1]$ in the diffusion limit. To illustrate the state space on which Wright–Fisher transitions act, consider $N = 4$: although there are only two alleles, the population has five distinct composition states,

$$(0, 4), (1, 3), (2, 2), (3, 1), (4, 0). \quad (7)$$

The states $(0, 4)$ and $(4, 0)$ represent fixation of types 2 and 1, respectively, when the entire population carries the same allele. The neighboring states $(1, 3)$ and $(3, 1)$ contain just one copy of the minority allele and illustrate the states near fixation [$i = O(1)$ or $N - i = O(1)$ for general N]. Mutation allows the population to leave a fixation state. The type-1 allele frequencies $X_i = i/4$ should not be confused with the stationary probabilities π_i assigned to these states below. The biological mutation graph contains only the two vertices $1 \leftrightarrow 2$, whereas the stochastic population-state graph has $N + 1$ vertices.

The offspring sampling probability can be obtained from the standard Wright–Fisher reproduction scheme with selection and mutation. Suppose that type 1 and type 2 have relative reproductive fitnesses $1 + s$ and 1, respectively (positive s favors type 1). Fitness-weighted reproduction changes the type-1 frequency to

$$X_{\text{sel}} = \frac{(1 + s)X}{1 + sX}. \quad (8)$$

If mutation between the two types is symmetric, with probability μ in either direction, the probability that a sampled offspring is of type 1 is

$$w_{\text{fit}}(X) = \mu + (1 - 2\mu) \frac{(1 + s)X}{1 + sX}. \quad (9)$$

For small s and μ ,

$$w_{\text{fit}}(X) = X + sX(1 - X) + \mu(1 - 2X) + O(s^2, s\mu). \quad (10)$$

The term $sX(1 - X)$ is the leading selective change in the allele frequency, while $\mu(1 - 2X)$ is the leading change generated by symmetric mutation.

For the analytical expansions and the finite- N tables, we adopt the leading small- s , small- μ polynomial expression,

$$w(X) = X + sX(1 - X) + \mu(1 - 2X), \quad (11)$$

as the microscopic offspring sampling probability. It is the leading expansion of Equation (9) and permits explicit evaluation of the subleading EPR coefficients. Here, small μ refers only to the expansion of the one-generation sampling probability. It does not impose the sequential weak-mutation condition of Section 3.2.3, which requires mutation arrivals to be rare on the time scale of fixation or extinction of a mutant lineage. Given $w(X)$, the N individuals in the next generation are sampled independently from the common offspring pool. The number j of type-1 individuals in the next generation is binomially distributed,

$$W_{ij} = \binom{N}{j} [w(X)]^j [1 - w(X)]^{N-j}, \quad X = \frac{i}{N}. \quad (12)$$

Equations (11) and (12) define the finite-population Wright–Fisher chain studied below.

In the diffusion scaling,

$$s = \frac{\alpha}{N}, \quad \mu = \frac{\beta}{N}. \quad (13)$$

Equations (9) and (11) have the same $O(N^{-1})$ drift and differ first at $O(N^{-2})$. When discussing the robustness of the leading asymptotics, we more generally consider a binomial Wright–Fisher rule of the form

$$w_N(X) = X + \frac{b(X)}{N} + R_N(X), \quad R_N(X) = O(N^{-2}) \quad (14)$$

with the remainder uniform on $0 \leq X \leq 1$. For Equations (11) with (13), the leading drift is

$$b(X) = \alpha g(X) + \beta g'(X), \quad (15)$$

where

$$g(X) = X(1 - X), \quad g'(X) = 1 - 2X, \quad (16)$$

and $g(X)/N$ is the leading variance in one-generation sampling (throughout the article, the prime denotes the derivative with respect to X). The consequences of the sampling class (14) for the leading EPR and the boundary-localization transition are analyzed in Section 4. During scaling (13), the per-individual mutation probability vanishes as N^{-1} , but the total mutation-attempt rate is of order $N\mu = \beta$. Therefore, fixed positive β does not, in general, place the population in the sequential weak-mutation regime.

We restrict the parameters to the admissible regime:

$$0 < w(X) < 1, \quad 0 \leq X \leq 1. \quad (17)$$

A simple sufficient condition is

$$0 < \mu < \frac{1}{2}, \quad |s| \leq 1 - 2\mu. \quad (18)$$

For fixed finite α and $\beta > 0$, admissibility is automatically satisfied for sufficiently large N in Equation (13). In this regime,

$$W_{ij} > 0, \quad i, j = 0, \dots, N, \quad (19)$$

so that the chain is irreducible and has a unique stationary distribution. The population-state graph is therefore complete: every population count can be reached from every other in one generation. The stationary distribution satisfies

$$\pi_j = \sum_{i=0}^N \pi_i W_{ij}, \quad \sum_{i=0}^N \pi_i = 1. \quad (20)$$

The stationary EPR per generation is provided directly by the discrete-time expression (3).

2.3. Lattice Representation

For the large- N analysis, it is convenient to rewrite the exact lattice probabilities in a form that approaches continuum densities. We define the rescaled lattice densities as

$$p_N(X_i) = N\pi_i, \quad \mathcal{T}_N(X_j|X_i) = NW_{ij}. \quad (21)$$

These quantities are defined only at the lattice points. Their exact normalizations are

$$\frac{1}{N} \sum_{i=0}^N p_N(X_i) = 1, \quad \frac{1}{N} \sum_{j=0}^N \mathcal{T}_N(X_j|X_i) = 1. \quad (22)$$

The stationary equation becomes

$$p_N(X_j) = \frac{1}{N} \sum_{i=0}^N p_N(X_i) \mathcal{T}_N(X_j|X_i). \quad (23)$$

The exact (discrete) forward–backward affinity is then

$$\mathcal{A}_N(X_i, X_j) = \ln \frac{p_N(X_i) \mathcal{T}_N(X_j|X_i)}{p_N(X_j) \mathcal{T}_N(X_i|X_j)}, \quad (24)$$

because the factors introduced by the rescaling cancel. Accordingly, the EPR reads as

$$\sigma_N = \frac{1}{N^2} \sum_{i,j=0}^N p_N(X_i) \mathcal{T}_N(X_j|X_i) \ln \frac{p_N(X_i) \mathcal{T}_N(X_j|X_i)}{p_N(X_j) \mathcal{T}_N(X_i|X_j)}. \quad (25)$$

Ordinary integrals enter only in the bulk large- N approximation, where

$$\frac{1}{N^2} \sum_{i,j} \longrightarrow \int_0^1 dX \int_0^1 dY. \quad (26)$$

2.4. Asymptotic and Boundary-Layer Analysis

We analyze the finite Wright–Fisher chain in the diffusion scaling $s = \alpha/N$ and $\mu = \beta/N$, with α and β fixed as $N \rightarrow \infty$. Writing

$$\epsilon = \frac{1}{N}, \quad (27)$$

the sampling rule (14) reads $w_N(X) = X + \epsilon b(X) + O(\epsilon^2)$, with $b(X)$ and $g(X)$ from Equations (15) and (16). We analyze separately the bulk, away from the fixation boundaries, and the boundary layers containing $O(1)$ individuals of the minority type.

In the bulk, one Wright–Fisher generation is expanded about its local Gaussian limit. For two successive frequencies X and Y , we introduce the standardized increment:

$$z = \frac{Y - X - \epsilon b(X)}{\sqrt{\epsilon g(X)}}. \quad (28)$$

An Edgeworth expansion (see, e.g., Refs. [33,34]) systematically refines the local Gaussian approximation by adding corrections determined by higher standardized cumulants, expressed here through Hermite polynomials $H_n(z)$. For the binomial Wright–Fisher kernel, the first non-Gaussian correction is controlled by the skewness and is proportional to the third Hermite polynomial, defined in the probabilists' convention by

$$H_3(z) = z^3 - 3z. \quad (29)$$

The transition density and stationary density are expanded consistently in powers of $\epsilon^{1/2}$. We expand the exact affinity (24), written as $\mathcal{A}_N(X, Y)$ for lattice arguments X and Y , order by order in the bulk. The EPR is evaluated using the exact fluctuation identity

$$\sigma_N = \langle \mathcal{A}_N \rangle_F, \quad \left\langle e^{-\mathcal{A}_N} \right\rangle_F = 1. \quad (30)$$

The fluctuation identity follows from the normalization of the stationary reversed two-time distribution. Appendix A derives the leading EPR independently from the exact finite-chain identity and the conditional jump moments.

Near the left fixation boundary, we set $X = k/N$ with fixed $k = O(1)$. For the sampling class (14),

$$w_N(k/N) = \frac{k + \beta}{N} + O(N^{-2}), \quad (31)$$

so that, for fixed k and j , the binomial kernel has the Poisson limit

$$W_{kj} \longrightarrow \mathcal{K}_{kj}^{(L)} = e^{-(k+\beta)} \frac{(k+\beta)^j}{j!}. \quad (32)$$

The right boundary is treated analogously. The boundary stationary solution is matched to the bulk stationary density in the overlap region:

$$1 \ll k \ll N. \quad (33)$$

The consequences of this matching for the EPR are derived in Section 4, with technical details in Appendix A.3.

2.5. Exact Finite-Population Numerics

The asymptotic results are tested against direct calculations for the full finite- N Wright-Fisher chain. For each N , α , and β , we construct the complete $(N+1) \times (N+1)$ transition matrix in the binomial form (12), with $s = \alpha/N$ and $\mu = \beta/N$, and determine its stationary distribution numerically. The figures of Section 4 use the full sampling rule (9), and the tables use Equation (11); the two rules differ at $O(N^{-2})$ and give the same exponents (Appendix B.3). Total-EPR scans use iteration of the exact transition matrix, while pair-resolved calculations use a direct solution of the stationarity equations. Numerical implementation and convergence tests are described in Appendix B.

For each unordered pair $i < j$, we define the nonnegative EPR contribution

$$e_{ij}^{(N)} = (\pi_i W_{ij} - \pi_j W_{ji}) \ln \frac{\pi_i W_{ij}}{\pi_j W_{ji}}, \quad i < j, \quad (34)$$

so that

$$\sigma_N = \sum_{i < j} e_{ij}^{(N)}. \quad (35)$$

To resolve where the EPR is produced in population-composition space, let $d_i = \min(i, N-i)$ and assign each pair the distance

$$D_{ij} = \max(d_i, d_j). \quad (36)$$

The cumulative fraction

$$\Phi_N(M) = \frac{1}{\sigma_N} \sum_{i < j; D_{ij} \leq M} e_{ij}^{(N)} \quad (37)$$

is the fraction of the EPR carried by pairs whose two endpoints both lie within distance M of a fixation boundary. The normalized EPR weight per logarithmic interval of distance is

$$G_N(M) = M [\Phi_N(M) - \Phi_N(M-1)], \quad M \geq 1. \quad (38)$$

For $M \gg 1$, $\sigma_N G_N(M)$ represents the EPR per unit logarithmic distance from fixation.

The asymptotic boundary and bulk sectors are also measured separately by assigning each pair the midpoint distance

$$d_{ij}^{\text{mid}} = \frac{1}{2} \min(i + j, 2N - i - j). \quad (39)$$

We obtain the inner and outer EPRs, $\sigma_{\text{in}}(N)$ and $\sigma_{\text{out}}(N)$, by summing $e_{ij}^{(N)}$ over pairs with $d_{ij}^{\text{mid}} \leq 10$ and $d_{ij}^{\text{mid}} \geq 0.1N$, respectively. The intervening region is omitted: σ_{in} and σ_{out} are not intended to partition σ_N ; they isolate the boundary and bulk scaling sectors. Using the midpoint rather than the endpoint farther from fixation avoids a leading $O(N^{-1/2})$ boundary correction, since typical Wright–Fisher jumps in the bulk span $O(\sqrt{N})$ individuals, while the cutoff lies $O(N)$ individuals from fixation.

For the total EPR and its separately resolved inner and outer parts, we define the finite-size effective exponents:

$$\theta_a(N) = -\frac{\ln[\sigma_a(2N)/\sigma_a(N)]}{\ln 2}, \quad a \in \{\text{in}, \text{out}, \text{eff}\}, \quad (40)$$

where $\sigma_{\text{eff}}(N) \equiv \sigma_N$ denotes the total EPR, while $\sigma_{\text{in}}(N)$ and $\sigma_{\text{out}}(N)$ are the EPR contributions assigned to the two spatial sectors defined above.

3. Results

Finite-population irreversibility is encoded in the full one-generation transition kernel and is not determined by its leading drift and variance alone. The reduced descriptions considered below retain different parts of this information. The fixed-point Gaussian approximation freezes the state dependence of the sampling fluctuations and discards non-Gaussian finite-step corrections; the diffusion retains state-dependent noise but replaces the microscopic one-generation kernel by continuous local dynamics; the sequential weak-mutation description eliminates polymorphic population states. Thus, their reversibility has different immediate origins.

3.1. Finite-Population Irreversibility

The difference between the genotype graph and the population-state graph can be seen explicitly already for $N = 3$. Consider neutral reproduction with symmetric mutation probability u , where $0 < u < 1$. For the four population states $i = 0, 1, 2, 3$, corresponding to $X_i = i/3$, the type-1 sampling probability is

$$w(X) = u + (1 - 2u)X. \quad (41)$$

For the population-state cycle

$$0 \rightarrow 1 \rightarrow 3 \rightarrow 0, \quad (42)$$

define the Kolmogorov cycle ratio as

$$\Gamma_C = \frac{W_{01}W_{13}W_{30}}{W_{03}W_{31}W_{10}}. \quad (43)$$

Using the binomial transition probabilities yields

$$\Gamma_C = \frac{1-u}{u} \left(\frac{1+u}{2-u} \right)^3. \quad (44)$$

Since

$$\Gamma_C - 1 = \frac{(1 - 2u)^3}{u(2 - u)^3}, \quad (45)$$

the Kolmogorov cycle condition $\Gamma_C = 1$ is satisfied only for $u = 1/2$. For $u \neq 1/2$, the forward and reverse products around this cycle differ, and the finite-population chain is not reversible [18]. A two-allele Wright–Fisher population can thus be irreversible at finite N (see also, e.g., Refs. [35–37]).

The role of the transition structure is particularly clear by comparison with the Moran process [38]. For the same $K = 2$ system with symmetric mutation considered here, the Moran model has the same $N + 1$ allele-count states as the Wright–Fisher chain, but a single update changes the population counts only by one. Its state graph is a birth–death chain, i.e., a line with no nontrivial closed cycles, and every finite irreducible birth–death chain is reversible. Thus, in this two-allele setting, the difference between the Wright–Fisher and Moran processes does not arise from the number of population states. It arises from whole-generation Wright–Fisher sampling, which permits nonlocal transitions and thereby creates population-state cycles.

3.2. Reversible Descriptions of Two-Allele Evolution

In this section, we address three reversible reductions of the finite Wright–Fisher chain: the fixed-point Gaussian approximation, the one-dimensional diffusion, and the sequential weak-mutation process. Each retains certain features of mutation–selection dynamics while losing the time-reversal asymmetry of whole-generation sampling. Comparing them reveals how different approximations can describe evolutionary behavior without preserving the irreversibility of the original process.

3.2.1. Fixed-Point Gaussian Approximation

Before deriving the leading nonzero large- N behavior of the finite-population Wright–Fisher EPR, it is instructive to examine whether the simplest local approximation near a deterministic fixed point retains time-reversal asymmetry. The fixed-point Gaussian approximation linearizes the conditional mean and removes two features of the sampling fluctuations that can carry time-reversal asymmetry: their state dependence and non-Gaussian finite-step structure. The resulting linear Gaussian process is reversible because, at stationarity, the allele frequencies in two consecutive generations have equal means and variances, making their joint Gaussian distribution symmetric under exchange. For $\sqrt{\alpha^2 + 4\beta^2} \gg 1$, the Gaussian approximation accurately describes the stationary fluctuations around the fixed point while missing the leading nonzero EPR. Otherwise, in the diffusion scaling $s, \mu \sim N^{-1}$, its stationary distribution does not become localized near the fixed point as $N \rightarrow \infty$.

The argument does not depend on the polynomial form of the sampling probability (11). Consider any smooth Wright–Fisher sampling $w(X)$ with an interior fixed point:

$$w(x_0) = x_0, \quad 0 < x_0 < 1, \quad (46)$$

and define

$$a = w'(x_0). \quad (47)$$

Writing

$$x = X - x_0, \quad y = Y - x_0 \quad (48)$$

for the deviations of two successive allele frequencies from the fixed point, the conditional mean has the local expansion

$$\langle y \rangle_x = ax + O(x^2). \quad (49)$$

The local Gaussian approximation reproduces the leading sampling fluctuations of the finite Wright–Fisher step. Since the next generation is drawn binomially with success probability $w(X)$, the conditional variance of the allele frequency Y is

$$\text{Var}(Y|X) = \frac{w(X)[1 - w(X)]}{N}. \quad (50)$$

At the fixed point x_0 , the conditional variance becomes

$$\text{Var}(Y|X = x_0) = \frac{g_0}{N}, \quad g_0 \equiv x_0(1 - x_0). \quad (51)$$

The approximation retains the linear term in the conditional mean and freezes the conditional variance at this fixed-point value. The resulting process can be written as

$$y = ax + \xi, \quad \langle \xi \rangle = 0, \quad \langle \xi^2 \rangle = \frac{g_0}{N}, \quad (52)$$

with ξ Gaussian and independent of x . Its transition kernel is

$$\mathcal{T}_G(y|x) = \sqrt{\frac{N}{2\pi g_0}} \exp\left[-\frac{N(y - ax)^2}{2g_0}\right]. \quad (53)$$

The stationary variance $V \equiv \text{Var}(x)$ satisfies

$$V = a^2V + \frac{g_0}{N}, \quad (54)$$

as follows from Equation (52) and the independence of x and ξ . A finite positive solution exists precisely when

$$|a| < 1, \quad (55)$$

and is

$$V = \frac{g_0}{N(1 - a^2)}. \quad (56)$$

Since a Gaussian distribution is mapped by Equation (53) to another Gaussian distribution, the zero-mean Gaussian with the variance in Equation (56) is stationary:

$$p_G(x) = \sqrt{\frac{N(1 - a^2)}{2\pi g_0}} \exp\left[-\frac{N(1 - a^2)x^2}{2g_0}\right]. \quad (57)$$

Detailed balance now follows directly. In the stationary process, x and y are jointly Gaussian, with equal variances V and covariance aV . Their joint distribution is therefore symmetric under exchange of x and y . Yet, neither factor in the product $p_G(x)\mathcal{T}_G(y|x)$ is separately symmetric under $x \leftrightarrow y$. Remarkably, their asymmetries cancel in the product. Combining Equations (53) and (57), we obtain

$$p_G(x)\mathcal{T}_G(y|x) \propto \exp\left\{-\frac{N}{2g_0}[(1 - a^2)x^2 + (y - ax)^2]\right\} = \exp\left[-\frac{N}{2g_0}(x^2 + y^2 - 2axy)\right]. \quad (58)$$

The stationary variance fixes the coefficient $1 - a^2$, which converts the two asymmetric quadratic terms into the symmetric expression. Hence,

$$p_G(x)\mathcal{T}_G(y|x) = p_G(y)\mathcal{T}_G(x|y), \quad (59)$$

and

$$\sigma_{\text{Gauss}} = 0. \quad (60)$$

Reversibility is thus a general property of the fixed-point linear-Gaussian approximation. Any smooth Wright–Fisher sampling rule with the same local construction produces a reversible Gaussian surrogate whenever $|a| < 1$.

For the sampling rule (11), the fixed point and the slope in the scaling (13) are

$$x_0 = \frac{1}{2} \left(1 + \frac{\alpha}{\sqrt{\alpha^2 + 4\beta^2 + 2\beta}} \right), \quad (61)$$

and

$$a = 1 - \frac{\sqrt{\alpha^2 + 4\beta^2}}{N}, \quad (62)$$

respectively. For sufficiently large N , $|a| < 1$, while Equation (56) yields

$$V \longrightarrow \frac{x_0(1-x_0)}{2\sqrt{\alpha^2 + 4\beta^2}} \quad (N \rightarrow \infty). \quad (63)$$

The stationary width remains finite at fixed α and β . A Gaussian approximation to the Wright–Fisher stationary distribution on $0 \leq X \leq 1$ requires

$$\sqrt{V} \ll \min(x_0, 1 - x_0). \quad (64)$$

Increasing N does not by itself ensure this condition. Thus, the fixed-point Gaussian approximation is not generally asymptotically controlled in the diffusion scaling.

The vanishing Gaussian EPR does not establish which features of the exact Wright–Fisher sampling process are responsible for its finite- N irreversibility: the approximation simultaneously linearizes the conditional mean, freezes the state-dependent variance, and discards higher cumulants of the binomial distribution. The large- N analysis of Section 3.3 determines how the finite-population EPR arises from the full sampling kernel and establishes its asymptotic scaling and localization in population space.

3.2.2. Diffusion Approximation for the Two-Allele Model

The second reduction is the Wright–Fisher diffusion approximation. Unlike the Gaussian approximation, it retains the state dependence of the sampling variance: as the allele frequency changes, so does the strength of the fluctuations. Although its generator contains only first and second derivatives, the diffusion has generally non-Gaussian transition probabilities. Indeed, successive applications of the Fokker–Planck operator generate derivatives of the state-dependent diffusion coefficient, producing nonvanishing higher conditional cumulants. Nevertheless, the diffusion for $K = 2$ is reversible: in one dimension, stationarity makes the probability current constant, while the absence of flux through the boundaries forces it to vanish. For transitions over one Wright–Fisher generation, corresponding to a diffusion-time interval $\epsilon = 1/N$, the reversibility of the limiting diffusion manifests itself as a cancellation between the forward–backward asymmetry of the state-dependent Gaussian factor and the leading non-Gaussian correction. This cancellation is incomplete for the finite Wright–Fisher chain because its third central moment has a different coefficient, as demonstrated in Section 3.4.

With $X = i/N$ and $\epsilon = 1/N$, the first two conditional moments of the one-generation increment in the scaling of Equation (13) are

$$\langle \Delta X \rangle_X = \epsilon b(X) + O(\epsilon^2), \quad (65)$$

$$\langle (\Delta X)^2 \rangle_X = \epsilon g(X) + O(\epsilon^2), \quad (66)$$

where $b(X)$ and $g(X)$ are defined in Equations (15) and (16). Let $P(X, \tau)$ denote the probability density of the allele frequency X at time $\tau = t/N$. The Fokker–Planck equation is

$$\frac{\partial P}{\partial \tau} = -\frac{\partial J}{\partial X} = -\frac{\partial}{\partial X}[b(X)P] + \frac{1}{2} \frac{\partial^2}{\partial X^2}[g(X)P], \quad (67)$$

where the probability current is

$$J(X, \tau) = b(X)P(X, \tau) - \frac{1}{2} \frac{\partial}{\partial X}[g(X)P(X, \tau)]. \quad (68)$$

At stationarity, $\partial_\tau P = 0$, and Equation (67) yields

$$\frac{dJ_{\text{st}}}{dX} = 0, \quad (69)$$

so that the stationary current is independent of X . For $\beta \geq 1/2$, the diffusion cannot reach fixation, whereas for $0 < \beta < 1/2$ it can, but mutation prevents fixation from being absorbing. In both regimes, there is no probability flux through the endpoints. Since J_{st} is constant, it follows that

$$J_{\text{st}}(X) = 0. \quad (70)$$

Solving Equation (68) with $J_{\text{st}} = 0$, the stationary probability density is

$$p_0(X) = \frac{1}{Z} e^{2\alpha X} [X(1-X)]^{2\beta-1}, \quad (71)$$

where the normalization is provided by

$$Z = \int_0^1 dX e^{2\alpha X} [X(1-X)]^{2\beta-1}. \quad (72)$$

For $0 < \beta < 1/2$, it is U-shaped with integrable divergences at both endpoints; at $\beta = 1/2$, it remains finite at the endpoints; for $\beta > 1/2$, it vanishes there. In the scaling (13), $\beta = 1/2$ corresponds to Wright's threshold $2N\mu = 1$ [29].

For a one-dimensional diffusion, the stationary EPR on the diffusion time scale is [19]

$$\sigma_{\text{diff}} = 2 \int_0^1 dX \frac{J_{\text{st}}(X)^2}{g(X)p_0(X)}. \quad (73)$$

Since the stationary current vanishes, Equation (70) yields

$$\sigma_{\text{diff}} = 0. \quad (74)$$

Thus, the two-allele diffusion remains reversible for every $\beta > 0$, regardless of whether the fixation boundaries are accessible.

3.2.3. Sequential Weak-Mutation Limit for Two Genotypes

We finally consider the sequential weak-mutation or origin–fixation limit of the two-genotype finite-population dynamics [39–41]. This limit requires a mutation scaling differ-

ent from the fixed- β scaling used in the finite-population and diffusion analyses above and eliminates polymorphic population states, leaving only the two monomorphic endpoints. The reduction thus removes the population-state cycles that can support stationary probability currents in the finite Wright–Fisher chain. The resulting two-state Markov process is necessarily reversible: at stationarity, the probability fluxes in the two directions must balance because there is no alternative path along which a current could circulate. Its zero EPR thus holds independently of the fixation probabilities and does not rely on a diffusion approximation.

Let μ_{ij} be the per-individual mutation probability from genotype i to genotype j . In a population monomorphic for genotype i , the rate at which single j mutants are generated is, in the rare-mutation limit,

$$v_{ij} \simeq N\mu_{ij}. \quad (75)$$

For the sequential description to apply, a mutant lineage must typically reach extinction or fixation before another mutation initiates a competing lineage. Let T_{seg}^{ij} denote the mean time, with mutation switched off, for a population initiated by one j mutant in an i background to reach either extinction or fixation. The sequential approximation requires

$$v_{ij}T_{\text{seg}}^{ij} \ll 1. \quad (76)$$

A simple way to enforce this separation of time scales is to take $\mu_{ij} \rightarrow 0$ at fixed N . By contrast, the diffusion scaling $\mu = \beta/N$ with fixed positive β leaves the total mutation-attempt rate of order $N\mu = \beta$ and does not generally satisfy Equation (76). In particular, the numerical regimes $\beta = 2$ and $\beta = 30$ used in Section 3.3 are not sequential weak-mutation regimes. The selection scaling $s = \alpha/N$, however, is compatible with the sequential limit. Under Equation (76), the population is monomorphic for most of the time. Let f_{ij} be the probability that a single mutant of genotype j fixes in a population resident at genotype i . The effective substitution rate is

$$\mathcal{R}_{ij} = v_{ij}f_{ij}, \quad i \neq j. \quad (77)$$

The fixation probabilities, and hence the substitution rates $\mathcal{R}_{ij}$, depend on the population model. Their values, however, do not affect the reversibility of the reduced process. To demonstrate this explicitly, let $P_i(t)$ denote the probability that the population is in monomorphic state i . The two-state master equation is

$$\begin{aligned} \frac{dP_1}{dt} &= -\mathcal{R}_{12}P_1 + \mathcal{R}_{21}P_2, \\ \frac{dP_2}{dt} &= \mathcal{R}_{12}P_1 - \mathcal{R}_{21}P_2. \end{aligned} \quad (78)$$

The two equations contain the same probability current with opposite signs. At stationarity, this current must vanish: there is no third state through which probability could return and sustain a circulating flux.

Let P_i^* denote the stationary probability of monomorphic state i . Setting the left-hand side of Equation (78) to zero and using $P_1^* + P_2^* = 1$ yields, for arbitrary positive substitution rates,

$$P_1^* = \frac{\mathcal{R}_{21}}{\mathcal{R}_{12} + \mathcal{R}_{21}}, \quad P_2^* = \frac{\mathcal{R}_{12}}{\mathcal{R}_{12} + \mathcal{R}_{21}}, \quad (79)$$

so that

$$P_1^*\mathcal{R}_{12} = P_2^*\mathcal{R}_{21}. \quad (80)$$

Detailed balance therefore holds identically, and

$$\sigma_{\text{seq}} = 0. \quad (81)$$

The fixation probabilities determine the relative stationary weights:

$$\frac{P_j^*}{P_i^*} = \frac{\mathcal{R}_{ij}}{\mathcal{R}_{ji}} = \frac{v_{ij} f_{ij}}{v_{ji} f_{ji}}, \quad i \neq j. \quad (82)$$

Although the sequential process is reversible, its substitution rates retain information about the polymorphic dynamics through the fixation probabilities. We evaluate the stationary weights for Wright–Fisher and Moran dynamics to identify which properties of the underlying population models survive this reduction.

For the finite Wright–Fisher chain, the fixation probabilities can be defined without a diffusion approximation. With mutation switched off during the segregation event, let h_i be the probability that type 1 ultimately fixes when the population initially contains i copies of type 1. Then,

$$h_i = \sum_{j=0}^N W_{ij}^{(\mu=0)} h_j, \quad h_0 = 0, \quad h_N = 1. \quad (83)$$

Thus, $f_{21} = h_1$ and $f_{12} = 1 - h_{N-1}$. For symmetric mutation-attempt rates, the exact stationary ratio of the corresponding sequential process is

$$\frac{P_1^*}{P_2^*} = \frac{h_1}{1 - h_{N-1}}. \quad (84)$$

In general, these exact finite- N fixation probabilities do not obey a simple exponential ratio.

A closed expression is obtained in the Wright–Fisher diffusion approximation. If s_i denotes the selection parameter associated with genotype i ,

$$f_{ij}^{\text{diff}} = \frac{1 - \exp[-2(s_j - s_i)]}{1 - \exp[-2N(s_j - s_i)]}, \quad (85)$$

with neutral limit $1/N$. This is an approximation to the fixation probability of the finite binomial chain. It yields

$$\frac{f_{ij}^{\text{diff}}}{f_{ji}^{\text{diff}}} = \exp[2(N-1)(s_j - s_i)]. \quad (86)$$

For symmetric mutation-attempt rates, Equation (82) then yields

$$P_i^* \propto \exp[2(N-1)s_i]. \quad (87)$$

These exponential weights follow from the diffusion fixation formula and are not exact finite- N results for the binomial Wright–Fisher chain. Consistency with Section 2.2 requires $s_1 - s_2 = s = \alpha/N$, so positive α favors type 1.

For the Moran process, the fixation probability is known exactly. If F_i is the reproductive fitness of genotype i and

$$\Delta_{ij} = \ln \frac{F_j}{F_i}, \quad (88)$$

then

$$f_{ij}^{\text{M}} = \frac{1 - \exp(-\Delta_{ij})}{1 - \exp(-N\Delta_{ij})}, \quad (89)$$

and

$$\frac{f_{ij}^M}{f_{ji}^M} = \exp[(N-1)\Delta_{ij}]. \quad (90)$$

For symmetric mutation-attempt rates, this exact fixation ratio yields

$$P_i^* \propto F_i^{N-1} = \exp[(N-1)\ell_i], \quad \ell_i = \ln F_i, \quad (91)$$

the Boltzmann-like stationary weights discussed in Ref. [40]. If $F_i = 1 + s_i$, then $\Delta_{ij} = \ln[(1 + s_j)/(1 + s_i)]$; replacing the logarithm with $s_j - s_i$ is a weak-selection approximation.

The fixation formulas show that the sequential reduction retains model-dependent stationary selection biases, but this information does not determine irreversibility. The full Moran chain is reversible, whereas the full Wright–Fisher chain remains irreversible at finite N even for very weak mutation. For example, a transition from fixation to two mutant copies requires two mutations, while the reverse loss of both copies can occur through ordinary sampling. At fixed N and small μ , the corresponding stationary fluxes are of order μ^2 and μ , respectively, since the stationary probability of a polymorphic state is of order μ . Such irreversible transitions are absent from the sequential model, which replaces polymorphic excursions by direct substitutions. Although the EPR of the full Wright–Fisher chain also vanishes as $\mu \rightarrow 0$ at fixed N , the sequential reduction gives exactly zero EPR by construction. It captures the stationary selection bias, but not the irreversibility generated during rare polymorphic excursions. This shows that a reduction can correctly describe stationary evolutionary outcomes without reproducing the entropy production of the full process. In particular, the fixed- N , $\mu \rightarrow 0$ result does not determine how the EPR of the full Wright–Fisher chain scales when $N \rightarrow \infty$ and $\mu \rightarrow 0$ jointly.

3.3. Large- N Entropy Production in the Two-Allele Wright–Fisher Model

The finite Wright–Fisher chain remains irreversible on the diffusion time scale, even though its limiting diffusion is reversible. We show that, for $\beta > 1/2$, the EPR per generation decreases as $1/N$, with a strictly positive coefficient even in the absence of selection. The origin of this residual irreversibility lies in the one-generation transition kernel: the leading non-Gaussian correction cancels only part of the forward–backward asymmetry generated by the state dependence of the Gaussian factor. The remaining affinity is of order $N^{-1/2}$ and produces an EPR of order N^{-1} . For the limiting diffusion, the corresponding cancellation is complete, as demonstrated in Section 3.4.

To derive this result, we return to the diffusion scaling (13), with α and β fixed as $N \rightarrow \infty$. The stationary density of the finite chain approaches the diffusion density $p_0(X)$ of Equation (71) in the bulk, but the forward–backward affinity must be calculated from the original discrete one-generation kernel rather than the transition kernel of the diffusion. For the polynomial sampling rule (11), the sampling probability is exactly

$$w(X) = X + \frac{b(X)}{N}, \quad (92)$$

with the drift $b(X)$ and the variance function $g(X)$ of Equations (15) and (16).

We use the rescaled lattice kernel $\mathcal{T}_N(X_j|X_i) = NW_{ij}$ introduced in Equation (21). In the bulk, where the lattice spacing is small compared with the width of the one-generation transition kernel, it admits the continuous local approximation

$$\mathcal{T}_N(Y|X) = \frac{1}{\sqrt{2\pi\epsilon g(X)}} e^{-z^2/2} \left[1 + \sqrt{\epsilon} \frac{g'(X)}{6\sqrt{g(X)}} H_3(z) + O(\epsilon) \right], \quad (93)$$

where z is the standardized increment defined in Equation (28) and $H_3(z)$ is defined in Equation (29). The corresponding local two-time affinity defined in Section 2.4 is

$$\mathcal{A}_N(X, Y) = -\frac{1}{6\sqrt{N}} \frac{g'(X)}{\sqrt{g(X)}} H_3(z) + O(N^{-1}). \quad (94)$$

The affinity in Equation (94) follows from Equation (93) and the expansion of the stationary density; Appendix A confirms the resulting leading EPR by an independent calculation from two binomial jump moments.

Equation (94) is a bulk expansion. It requires $NX \gg 1$ and $N(1-X) \gg 1$ and is not uniform in the endpoint layers, where $g(X)$ vanishes. Appendix A.3 treats those regions separately. For $\beta > 1/2$, their contribution is $o(N^{-1})$, so that they do not change the leading EPR. For $\beta > 1$, they are also $o(N^{-2})$ and do not change the first bulk correction. Using the exact fluctuation identity (30), the EPR can be expanded in moments of the affinity. In the bulk, $\mathcal{A}_N = O(N^{-1/2})$; hence, all terms beyond the quadratic one are $o(N^{-1})$. Therefore,

$$\sigma_N = \frac{1}{2} \langle \mathcal{A}_N^2 \rangle_F + o(N^{-1}). \quad (95)$$

The leading term in Equation (94) is proportional to $H_3(z)$ and has zero Gaussian average, so that the leading EPR arises at second order in the $O(N^{-1/2})$ forward-backward affinity. Using $\langle H_3^2 \rangle = 6$, we obtain

$$\sigma_N = \frac{\mathcal{C}(\alpha, \beta)}{N} + o(N^{-1}), \quad \beta > \frac{1}{2}, \quad (96)$$

with

$$\mathcal{C}(\alpha, \beta) = \frac{1}{12} \left\langle \frac{[g'(X)]^2}{g(X)} \right\rangle_{p_0} = \frac{1}{12} \left\langle \frac{(1-2X)^2}{X(1-X)} \right\rangle_{p_0}. \quad (97)$$

The condition $\beta > 1/2$ follows from the endpoint behavior of this average. An explicit evaluation is provided in Appendix A.5. The drift $b(X)$ enters the leading coefficient only through p_0 . In particular, $\mathcal{C}(0, \beta) = 1/[6(2\beta - 1)] > 0$; thus, selection is not required for irreversibility.

For $\beta > 1$, Appendix A.4 shows

$$N\sigma_N = \mathcal{C}(\alpha, \beta) + \frac{\mathcal{C}_1(\alpha, \beta)}{N} + o(N^{-1}). \quad (98)$$

Table 1 confirms this prediction by comparing the exact numerical values with the two-term results for $N\sigma_N$. For the parameters

$$\alpha = \beta = 30, \quad (99)$$

the analytical coefficients are

$$\begin{aligned} \mathcal{C}(30, 30) &= 0.0226976982\dots, \\ \mathcal{C}_1(30, 30) &= -6.7247278006\dots \end{aligned} \quad (100)$$

The negative $\mathcal{C}_1$ predicts an approach to the asymptotic value from below. The convergence rate strongly depends on the scaled parameters. For $\alpha = \beta = 30$, $|\mathcal{C}_1/\mathcal{C}|$ is close to 300. Therefore, rather large N is needed before the asymptotic regime is reached. As a second test, $\alpha = 5$ and $\beta = 2$ give $\mathcal{C} \approx 0.1822$ and $\mathcal{C}_1 \approx -6.6406$; at $N = 1600$,

$(N\sigma_N - \mathcal{C})N \approx -6.59$. The divergence of $\mathcal{C}(\alpha, \beta)$ as $\beta \rightarrow 1/2$ signals the breakdown of the bulk expansion near fixation, not a divergence of the finite- N EPR.

Table 1. Entropy production rate of the two-allele Wright–Fisher chain for $s = 30/N$ and $\mu = s$. The first row shows the direct finite- N result from the full transition matrix and stationary distribution. The second row shows the two-term large- N asymptotics $N\sigma_N = \mathcal{C} + \mathcal{C}_1/N$ evaluated analytically. The last column shows the analytical limit rounded to six decimal places.

N	1000	2000	3000	5000	9000	∞
exact $N\sigma_N$	0.016386	0.019437	0.020501	0.021369	0.021955	–
two-term $N\sigma_N$	0.015973	0.019335	0.020456	0.021353	0.021951	0.022698

3.4. Comparison with Diffusion

The finite Wright–Fisher chain and its limiting diffusion have the same leading drift and state-dependent variance, yet different entropy production. The reason is not that the diffusion eliminates non-Gaussian fluctuations. As discussed in Section 3.2.2, even in a stationary diffusion, the conditional distribution is generally non-Gaussian: the noise amplitude depends on the allele frequency, and the resulting evolution generates nonzero higher cumulants.

The essential difference lies in the third central moment of the transition kernel. One Wright–Fisher generation is a single binomial sampling step, whereas the diffusion over the corresponding time interval accumulates infinitesimal fluctuations whose variance changes with X . The diffusion value produces a non-Gaussian correction that cancels the leading forward–backward asymmetry of the state-dependent Gaussian factor. For the binomial Wright–Fisher kernel, this cancellation is incomplete, leaving a nonzero affinity and EPR. Below, we derive this difference and show that the leading EPR of a broader class of one-step kernels is determined by the deviation of their third central moment from its diffusion value.

For the Wright–Fisher kernel, the third central moment is

$$\left\langle (Y - \langle Y \rangle_X)^3 \right\rangle_X^{\text{WF}} = \epsilon^2 g(X) g'(X) + O(\epsilon^3). \quad (101)$$

For the diffusion generated by Equation (67), the corresponding short-time moment is instead

$$\left\langle (Y - \langle Y \rangle_X)^3 \right\rangle_X^{\text{diff}} = \frac{3}{2} \epsilon^2 g(X) g'(X) + O(\epsilon^3). \quad (102)$$

A derivation of Equation (102) is presented in Appendix A.2. Both kernels have conditional mean $X + \epsilon b(X)$ and variance $\epsilon g(X)$. The Wright–Fisher expansion is Equation (93), with coefficient $1/6$ of the $H_3(z)$ correction. Equation (102) instead gives coefficient $1/4$:

$$\mathcal{T}_{\text{diff}}(Y|X) = \frac{e^{-z^2/2}}{\sqrt{2\pi\epsilon g}} \left[1 + \sqrt{\epsilon} \frac{g'}{4\sqrt{g}} H_3(z) + O(\epsilon) \right]. \quad (103)$$

The coefficients $1/6$ and $1/4$ of the $H_3(z)$ correction determine whether the leading forward–backward affinity cancels. The leading factors in the Wright–Fisher and diffusion Edgeworth expansions share the state-dependent variance $\epsilon g(X)$. Combined with the stationary density $p_0(X)$, they give the same forward–backward affinity:

$$\mathcal{A}^{(0)} = -\frac{\sqrt{\epsilon}}{2} \frac{g'}{\sqrt{g}} H_3(z) + O(\epsilon). \quad (104)$$

If the variance was frozen at $g(x_0)$, as in the fixed-point Gaussian approximation of Section 3.2.1, this contribution to the affinity would vanish. The non-Gaussian correction for the two cases is provided by

$$\mathcal{A}_{\text{WF}}^{(E)} = +\frac{\sqrt{\epsilon}}{3} \frac{g'}{\sqrt{g}} H_3(z) + O(\epsilon), \quad (105)$$

$$\mathcal{A}_{\text{diff}}^{(E)} = +\frac{\sqrt{\epsilon}}{2} \frac{g'}{\sqrt{g}} H_3(z) + O(\epsilon). \quad (106)$$

For the Wright–Fisher kernel, the coefficients combine as $-1/2 + 1/3 = -1/6$, reproducing the nonzero affinity in Equation (94). For the diffusion, $-1/2 + 1/2 = 0$, so the leading affinity vanishes. This cancellation is consistent with the exact reversibility of the stationary diffusion established by Equation (70).

The mechanism is not restricted to binomial sampling. Consider one-step kernels whose bulk transition densities admit a regular local Gaussian–Edgeworth expansion, with conditional mean $X + \epsilon b(X) + O(\epsilon^2)$, variance $\epsilon g(X) + O(\epsilon^2)$, and the third central moment

$$\langle (Y - \langle Y \rangle_X)^3 \rangle_X = \kappa_3 \epsilon^2 g(X) g'(X) + O(\epsilon^3), \quad (107)$$

where κ_3 is constant. We assume that higher cumulants introduce no additional terms at order $\sqrt{\epsilon}$, that the bulk stationary density approaches $p_0(X)$, and that boundary contributions to the EPR are $o(\epsilon)$ for $\beta > 1/2$. These conditions hold for the Wright–Fisher kernel and the diffusion considered above. The skewness is then $\kappa_3 \sqrt{\epsilon} g' / \sqrt{g}$, the coefficient of $H_3(z)$ in the local kernel is $\kappa_3/6$, and the non-Gaussian affinity is $(1/3)\kappa_3 \sqrt{\epsilon} (g' / \sqrt{g}) H_3(z)$. Adding Equation (104) yields

$$\mathcal{A} = \frac{1}{3} (\kappa_3 - 3/2) \sqrt{\epsilon} \frac{g'}{\sqrt{g}} H_3(z) + O(\epsilon). \quad (108)$$

For $\beta > 1/2$, Equation (95) and $\langle H_3^2 \rangle = 6$ then yield

$$\lim_{N \rightarrow \infty} N \sigma_N = \frac{1}{3} (\kappa_3 - 3/2)^2 \left\langle \frac{[g'(X)]^2}{g(X)} \right\rangle_{p_0}. \quad (109)$$

The binomial Wright–Fisher kernel has $\kappa_3 = 1$ and reproduces Equation (97), whereas the diffusion has $\kappa_3 = 3/2$ and gives zero. Equation (109) shows that, for one-step kernels with a regular Gaussian–Edgeworth expansion specified above, the leading EPR on the diffusion time scale depends on the microscopic kernel through the deviation $\kappa_3 - 3/2$. For $\kappa_3 \neq 3/2$, this EPR has a strictly positive large- N limit, although the limiting diffusion is reversible. The value $\kappa_3 = 3/2$ is exceptional because it cancels the leading term; this condition alone does not establish detailed balance for the finite-step kernel. For $\kappa_3 = 0$, corresponding to a Gaussian one-step kernel with state-dependent variance, Equation (109) predicts $9\mathcal{C}(\alpha, \beta)$, provided the stated boundary conditions hold. Thus, eliminating the Wright–Fisher skewness increases the leading EPR ninefold rather than restoring reversibility.

The Fokker–Planck truncation of the Kramers–Moyal expansion retains the leading drift $b(X)$ and variance $g(X)$ of the one-generation kernel, but not the coefficient κ_3 of its third central moment. The general-kernel result applies to the above Gaussian–Edgeworth class, not to arbitrary population dynamics. In particular, the two-allele Moran process has nearest-neighbor elementary transitions and forms a reversible birth–death chain at every finite N . Its transition structure provides a separate reason for zero EPR.

The diffusion limit is singular with respect to entropy production. For $\beta > 1/2$, the Wright–Fisher EPR per generation satisfies $\sigma_N \rightarrow 0$ as $N \rightarrow \infty$, whereas the EPR per unit diffusion time, $N\sigma_N$, has a positive limit. In contrast, the limiting diffusion has zero stationary EPR:

$$\lim_{N \rightarrow \infty} N\sigma_N = \mathcal{C}(\alpha, \beta) \neq \sigma_{\text{diff}} = 0, \quad \beta > \frac{1}{2}. \quad (110)$$

Thus, the diffusion approximation preserves the leading evolution of the allele-frequency distribution but not the entropy production associated with whole-generation transitions. Computing the EPR and taking the diffusion limit do not commute.

4. Boundary-Localization Phase Transition

The bulk coefficient $\mathcal{C}(\alpha, \beta)$ in Equation (96) is finite only for $\beta > 1/2$. Indeed, as β approaches $1/2$ from above, the integral defining this coefficient receives increasingly large contributions from populations close to fixation, where one allele is rare. At $\beta = 1/2$, the integral develops a logarithmic divergence. This does not mean that the EPR of a finite population diverges. Rather, it tells us that the bulk calculation alone is no longer sufficient to determine how the EPR scales with population size.

The reason is that one-generation sampling behaves differently in the population bulk and near fixation. In the bulk, where both alleles have many copies, the EPR scales as N^{-1} . Near fixation, the minority allele has only $O(1)$ copies, and its offspring count follows a Poisson rather than a Gaussian sampling law. The mutation parameter β controls how much stationary probability is concentrated near fixation: the probability of each state with a fixed number of minority individuals scales as $N^{-2\beta}$. Since the transition affinities in this region remain finite, its EPR contribution has the same scaling. For $\beta < 1/2$, this contribution decays more slowly than the bulk contribution and thus dominates. For $\beta > 1/2$, the bulk dominates instead. At $\beta = 1/2$, comparable contributions come from successive logarithmic ranges of distance from fixation. Their sum produces the $(\ln N)/N$ scaling. This change in where the EPR is produced is the *boundary-localization phase transition*. The transition does not require different continuum limits on the two sides of $\beta_c = 1/2$. Throughout the analysis, we use the same scaling $s = \alpha/N$, $\mu = \beta/N$. What changes is which part of the finite population contributes most to the total EPR.

4.1. Boundary Scaling and Localization Order Parameter

The bulk EPR coefficient in Equation (97) reveals the spatial distribution of entropy production: the contribution at allele frequency X is proportional to $p_0(X)[g'(X)]^2/g(X)$, where the factor $[g'(X)]^2/g(X)$ diverges near fixation because $g(X) = X(1-X)$ vanishes. At the same time, the stationary probability density behaves as $p_0(X) \propto X^{2\beta-1}$ near $X = 0$. Combining these two factors yields

$$p_0(X) \frac{[g'(X)]^2}{g(X)} \sim \frac{1}{Z} X^{2\beta-2}, \quad (111)$$

with an analogous expression near $X = 1$.

To determine the scaling of the boundary contribution to EPR, we use the variables and Poisson kernel introduced in Section 2.4, Equations (31) and (32). Matching to the interior stationary density in the overlap region $1 \ll k \ll N$ yields $\pi_k \sim Z^{-1} N^{-2\beta} k^{2\beta-1}$. The stationary solution of the boundary Poisson kernel, matched to this asymptotic form, then yields $\pi_k = O(N^{-2\beta})$ for fixed k , as derived in Appendix A.3. Since the boundary

transition probabilities and forward–backward affinities remain $O(1)$, the EPR carried by a fixed part of either boundary layer consequently scales as

$$\sigma_{N,M}^{\text{bdry}} = O(N^{-2\beta}), \quad M \text{ fixed.} \quad (112)$$

Comparison with the interior contribution $O(N^{-1})$ provides three large-population regimes of the EPR:

$$\sigma_N \sim \begin{cases} C_{\text{bdry}}(\alpha, \beta) N^{-2\beta}, & 0 < \beta < \frac{1}{2}, \\ C_{\log}(\alpha) \frac{\ln N}{N}, & \beta = \frac{1}{2}, \\ \frac{C(\alpha, \beta)}{N}, & \beta > \frac{1}{2}. \end{cases} \quad (113)$$

The power $N^{-2\beta}$ follows from Equation (112), whereas the amplitude $C_{\text{bdry}}(\alpha, \beta)$ requires the stationary solution of the boundary Poisson kernel (32); this kernel is independent of α , but the amplitude depends on α through the matching to the interior solution.

At the critical point $\beta = 1/2$,

$$p_0(X) = \frac{e^{2\alpha X}}{Z_c}, \quad Z_c = \int_0^1 dX e^{2\alpha X} = \frac{e^{2\alpha} - 1}{2\alpha}. \quad (114)$$

Here, $Z_c = 1$ at $\alpha = 0$ by continuity. The logarithmic singularities from the two endpoints yield

$$C_{\log}(\alpha) = \frac{1 + e^{2\alpha}}{12Z_c} = \frac{\alpha \coth \alpha}{6}, \quad \sigma_N \sim C_{\log}(\alpha) \frac{\ln N}{N}. \quad (115)$$

The continuous neutral limit is $C_{\log}(0) = 1/6$. Table A1 in Appendix B provides direct finite- N checks of these two regimes. At $\beta = 1/2$, the logarithmic slope of $N\sigma_N$ approaches the predicted coefficient $C_{\log}(\alpha)$; at $\beta = 0.3$, the rescaled EPR $N^{2\beta}\sigma_N$ approaches a finite boundary amplitude. These tests use the polynomial sampling rule (11). On the diffusion time scale, Equation (113) implies that $N\sigma_N$ diverges as $N^{1-2\beta}$ for $0 < \beta < 1/2$, grows as $C_{\log}(\alpha) \ln N$ at $\beta = 1/2$, and approaches $C(\alpha, \beta)$ for $\beta > 1/2$.

The large- N singularity can be characterized in two complementary ways. First, the EPR decay exponent

$$\theta(\beta) = - \lim_{N \rightarrow \infty} \frac{\ln \sigma_N}{\ln N} = \min(1, 2\beta) \quad (116)$$

is continuous but has a discontinuous derivative at $\beta_c = 1/2$. This asymptotic exponent is the large- N limit of the finite-size exponent $\theta_{\text{eff}}(N, \beta)$ defined in Equation (40). Second, the transition is visible in the spatial localization of the EPR. The fraction localized within a finite distance from fixation is characterized by taking first the large- N and then the large- M limit of the cumulative fraction defined in Equation (37):

$$\Phi(\alpha, \beta) = \lim_{M \rightarrow \infty} \lim_{N \rightarrow \infty} \Phi_N(M) = \begin{cases} 1, & 0 < \beta < \frac{1}{2}, \\ 0, & \beta \geq \frac{1}{2}. \end{cases} \quad (117)$$

To obtain the EPR contribution outside the boundary regions of width M , we consider allele-count distances $k > M$ from fixation. For $1 \ll M < k \ll N$, the bulk EPR density in Equation (111) can be matched to $X = k/N$. Since $dX = dk/N$, the EPR carried per unit distance k from the left fixation boundary behaves as

$$\frac{d\sigma_N}{dk} \sim \frac{1}{12Z} N^{-2\beta} k^{2\beta-2}, \quad 1 \ll k \ll N, \quad (118)$$

with an analogous contribution from the right boundary. This is the spatial EPR density after summing over transitions, not the stationary probability π_k or an individual pairwise term $e_{ij}^{(N)}$. For $\beta < 1/2$, the tail is summable: its contribution to the EPR beyond a boundary width M , relative to the leading $N^{-2\beta}$ scaling, decreases as $M^{2\beta-1}$. Increasing M after taking $N \rightarrow \infty$ thus captures the full leading boundary EPR, yielding $\Phi = 1$. For $\beta > 1/2$, every fixed-width boundary region contributes only $O(N^{-2\beta}) = o(N^{-1})$. At $\beta = 1/2$, a fixed-width region contributes $O(N^{-1})$, whereas the total EPR scales as $(\ln N)/N$, yielding $\Phi = 0$. The critical EPR is distributed over the hierarchy $1 \ll k \ll N$, rather than confined to a fixed boundary width or dominated by distances $k = O(N)$. $\times$

To display the spatial distribution of the EPR, we resolve it by the distance M from fixation using Equations (34) and (38). The quantity $\sigma_N G_N(M)$ gives the EPR per logarithmic interval of M ; Figure 1 shows it multiplied by the boundary scaling factor $N^{2\beta}$. In Figure 1, we see that the microscopic boundary structure persists on both sides of β_c . The four panels display the redistribution of irreversibility from the fixation boundary to the bulk at $\beta_c = 1/2$. The transition occurs because its share of the total EPR changes with population size. At criticality, the logarithmic enhancement comes from the increasing range of distances carrying comparable EPR, not from an increase in the height of the microscopic boundary peak. The order parameter Φ in Equation (117) measures the fraction that remains localized near fixation in the large-population limit.

4.2. Mixed-Order Transition and Critical Crossover

The jump of the localization order parameter tells us whether the EPR remains concentrated near fixation as $N \rightarrow \infty$. It does not tell us how wide a boundary region must be to capture a given fraction of the EPR. As β approaches $1/2$ from below, the boundary tail becomes increasingly broad. We now determine how rapidly the required width grows. For $0 < \beta < 1/2$, define the large-population cumulative boundary fraction $\Phi_\infty(M; \alpha, \beta) = \lim_{N \rightarrow \infty} \Phi_N(M)$, where Φ_N is provided by Equation (37). For $0 < \eta < 1$, let M_η be the smallest boundary width satisfying $\Phi_\infty(M_\eta; \alpha, \beta) \geq \eta$. Then, the large-population limit is taken before approaching β_c .

The boundary tail in Equation (118) decays with exponent $2\beta - 2 = -1 - (1 - 2\beta)$, approaching the marginal $1/M$ decay as $\beta \rightarrow 1/2$ from below. Summing the EPR contributions from minority-allele counts beyond M at both fixation boundaries yields

$$1 - \Phi_\infty(M; \alpha, \beta) \sim \frac{1 + e^{2\alpha}}{12Z(\alpha, \beta)(1 - 2\beta)\mathcal{C}_{\text{bdry}}(\alpha, \beta)} M^{-(1-2\beta)}, \quad M \gg 1. \quad (119)$$

As $\beta \rightarrow (1/2)^-$, the contribution of this large- M tail to the boundary amplitude $\mathcal{C}_{\text{bdry}}$ diverges, while the contribution from any fixed number of states near fixation remains finite. Since $(1 + e^{2\alpha})/[12Z(\alpha, \beta)]$ approaches $\mathcal{C}_{\text{log}}(\alpha)$,

$$\mathcal{C}_{\text{bdry}}(\alpha, \beta) \sim \frac{\mathcal{C}_{\text{log}}(\alpha)}{1 - 2\beta}. \quad (120)$$

The normalized tail approaches $M^{-(1-2\beta)}$, yielding

$$\ln M_\eta \sim -\frac{\ln(1 - \eta)}{1 - 2\beta}, \quad \beta \rightarrow \frac{1}{2}^-. \quad (121)$$

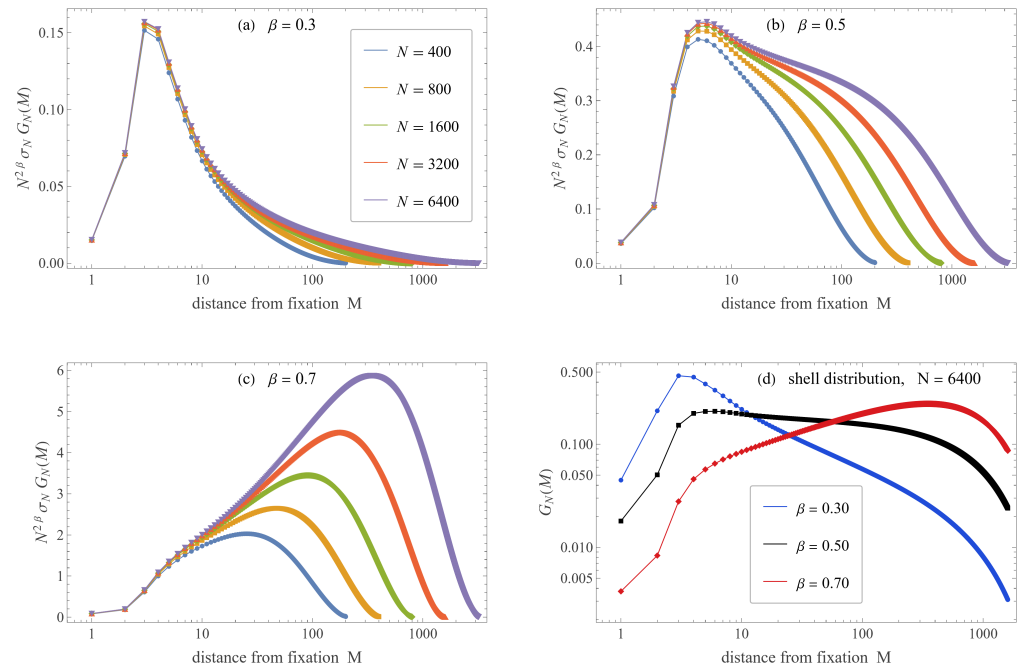


Figure 1. Spatial distribution of the stationary entropy production for $\alpha = 2$. Panels (a–c) show $N^{2\beta} \sigma_N G_N(M)$ for $N = 400, 800, 1600, 3200, 6400$, with the colors of the curves indicated in the legend of panel (a). Here, M is the distance from fixation and $G_N(M)$ is the dimensionless shell weight defined so that $\sigma_N G_N(M)$ gives the EPR carried per logarithmic interval of M . (a) For $\beta = 0.3 < 1/2$, the near-boundary profile approaches an N -independent form, and increasing N only extends a decaying tail whose integrated weight converges. (b) At $\beta = 1/2$, the shell-resolved EPR retains a fixed near-boundary maximum at $M \lesssim 10$, while a broad shoulder extends to progressively larger M as N increases. In the large- N limit, this shoulder approaches an approximately flat plateau in $N \sigma_N G_N(M)$ over the range $1 \ll M \ll N$, while its outer cutoff moves to distances $M = O(N)$. The growing range of logarithmic distances carrying appreciable EPR produces the $(\ln N)/N$ scaling of the total EPR. (c) For $\beta = 0.7 > 1/2$, a remnant of the microscopic boundary structure remains visible as a shoulder around $M \lesssim 10$, while at any fixed M , the left wing of the rescaled shell weight $N^{2\beta} \sigma_N G_N(M)$ approaches an N -independent profile. A pronounced maximum develops at large M and shifts outward proportionally to N , $M_{\max} = O(N)$. The EPR carried by this bulk region grows relative to the boundary wing and dominates the total EPR at large N . (d) Unscaled shell distribution $G_N(M)$ at $N = 6400$ for $\beta = 0.3, 0.5, 0.7$, showing the change from a boundary-localized profile to a bulk-dominated one.

The divergence of M_η measures the increasing range of distances needed to capture a fixed fraction of the EPR, rather than the broadening of the microscopic boundary peak. At finite N , the tail extends to distances of order N . Its accumulated weight changes appreciably when $|2\beta - 1| \ln N$ is of order one. This provides the finite-size crossover variable

$$v = (2\beta - 1) \ln N, \quad (122)$$

which describes the gradual change between boundary- and bulk-dominated EPR near β_c . (For the finite-size data collapse, the logarithm is refined to $L_N = \ln(N/k_0)$, with the microscopic scale k_0 determined from the critical EPR; see Appendix B.6.)

The jump of Φ together with the divergence of M_η identifies a mixed-order transition [42,43]. The structure here parallels that of one-dimensional mixed-order models, where a power-law distribution crosses a summability threshold. A classic example is the one-dimensional Ising model with $1/r^2$ interactions, for which the magnetization is discontinuous at the transition [44,45]. In the present problem, the corresponding power-

law object is not a domain-size distribution but the distribution of stationary entropy production over the distance from fixation in population space.

Unlike dynamical critical phenomena characterized by critical slowing down [46], the present transition does not involve a closing relaxation gap of the neutral limiting diffusion. For $\alpha = 0$, its relaxation rates on the diffusion clock are $(n/2)(n - 1 + 4\beta)$ and remain analytic across β_c . The order parameter is constructed from stationary forward-backward probability fluxes and probes trajectory-level time-reversal asymmetry rather than the stationary allele-frequency density alone. The limiting Wright–Fisher diffusion remains reversible on both sides of β_c , even though the finite Wright–Fisher chain has nonzero EPR. The transition concerns how finite-population irreversibility is distributed and survives the large-population limit, rather than a change in the reversibility of the limiting diffusion.

The approach to the transition from the bulk side fixes the amplitude of the critical logarithm. As $\beta \rightarrow (1/2)^+$, the endpoint singularities of Equation (111) provide

$$\mathcal{C}(\alpha, \beta) \simeq \frac{C_{\log}(\alpha)}{2\beta - 1}, \quad \beta \rightarrow \frac{1}{2}^+, \quad (123)$$

with the same $C_{\log}(\alpha)$ that multiplies the EPR logarithm at criticality. For $\alpha = 0$, this relation is exact for all $\beta > 1/2$, since $\mathcal{C}(0, \beta) = 1/[6(2\beta - 1)]$. The pole of the bulk amplitude as $\beta \rightarrow \beta_c$ and the logarithmic growth $C_{\log}(\alpha) \ln N$ at β_c arise from the same boundary singularity. The same threshold $\beta_c = 1/2$ governs the endpoint structure of the limiting diffusion, discussed after Equation (71).

The finite-size crossover near criticality follows directly from the singular endpoint integral, with cutoffs of order N^{-1} and unity:

$$\frac{1}{N} \int_{1/N}^1 X^{2\beta-2} dX = \frac{1}{N} \frac{1 - N^{-(2\beta-1)}}{2\beta - 1}. \quad (124)$$

The integration limits here are schematic but sufficient for the scaling argument. The lower cutoff marks where the bulk description must be matched to the boundary Poisson kernel; the upper cutoff marks where the endpoint power law must be replaced by the full bulk expression. Their precise values affect finite-size corrections but not the leading near-critical scaling function. With the scaling variable v from Equation (122), in the joint limit $N \rightarrow \infty$, $\beta \rightarrow 1/2$ at fixed v , the singular part takes the finite-size scaling form

$$\sigma_N^{\text{sing}} \simeq C_{\log}(\alpha) \frac{\ln N}{N} \frac{1 - e^{-v}}{v}. \quad (125)$$

Thus, at every finite N , the transition matrix and its stationary state vary smoothly with β near β_c , and the change of regime is a crossover controlled by v . The nonanalyticity appears only after the large-population limit is taken, as in an ordinary finite-size scaling problem.

The boundary layers also change the corrections to scaling. For $1/2 < \beta < 1$, the leading term remains $\mathcal{C}/N$, but the first correction is $O(N^{-2\beta})$ rather than $O(N^{-2})$. At $\beta = 1$, the correction becomes marginal, $O[(\ln N)/N^2]$. Only for $\beta > 1$ does the regular two-term expansion (98) apply. Thus, $\beta = 1$ is a correction-to-scaling threshold, not a second transition of the leading EPR. The endpoint power counting suggests possible further correction-to-scaling thresholds at $\beta = n/2$, with integer $n \geq 3$. Establishing this hierarchy requires higher-order bulk expansions and boundary matching, which we leave for future work.

4.3. Scaling Signatures of the Transition

A fixed-width boundary contribution to the EPR scales as $N^{-2\beta}$ on both sides of the transition, whereas an outer contribution taken at a fixed fraction of the population scales

as N^{-1} . Their exponents cross at $2\beta = 1$. The numerical resolution of these two sectors is presented in Section 4.4, and the shell distribution of the EPR in Section 4.1; details and finite-size diagnostics are presented in Appendix B.

Figure 2 provides two complementary tests of the boundary-localization transition. Panel (a) tests the spatial power law of Equation (126) below. Panel (b) tests the finite-size crossover of the total EPR near $\beta_c = 1/2$. After accounting for the microscopic scale through the effective logarithm L_N , the numerical data approach the predicted scaling function $f(v) = (1 - e^{-v})/v$, without fitted crossover parameters. For the spatial test in panel (a), we divide the shell profile by the smooth outer envelope $\mathcal{E}_{\alpha,\beta}(M/N)$ defined in Appendix B. The resulting profile obeys

$$\tilde{G}_N(M) \propto M^{2\beta-1}, \quad 1 \ll M \ll N. \quad (126)$$

Normalizing $\tilde{G}_N(M)$ at a fixed reference scale removes the boundary-core amplitude without changing the exponent in Equation (126). The dashed guides in Figure 2a use the predicted exponents $2\beta - 1$, with only their vertical offsets fitted. The comparison is performed without fitting its exponent: the profile decreases for $\beta < 1/2$, increases for $\beta > 1/2$, and is approximately scale-independent in the overlap region at criticality.

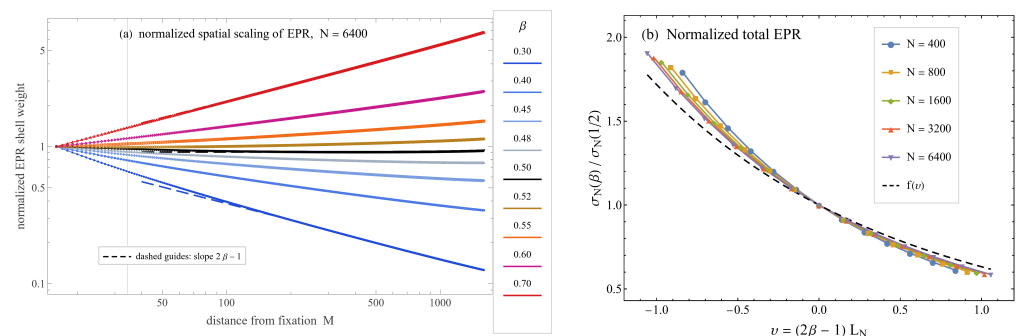


Figure 2. Spatial scaling and finite-size crossover of entropy production for $\alpha = 2$. **(a)** Normalized spatial scaling of the EPR at $N = 6400$. The smooth outer envelope $\mathcal{E}_{\alpha,\beta}(x) = (1 - x)^{2\beta-2}(1 - 2x)^2[\exp(2\alpha x) + \exp(2\alpha(1 - x))]$, with $x = M/N$, is divided out, and the compensated shell weight is normalized at $M = 16$. The thin vertical line indicates the approximate crossover from the microscopic boundary region to the power-law regime, where the compensated weight scales as $M^{2\beta-1}$. Dashed guides for $\beta = 0.3, 0.5, 0.7$ have slopes $2\beta - 1$; only their vertical offsets are fitted over $64 \leq M \leq N/4$, with weights $1/M$, and are extended to $M = 40$ so that they remain visible where the boundary-core correction lifts the data above them. **(b)** Finite-size collapse of the total EPR for $N = 400$ to 6400 , using the full sampling rule. The EPR is normalized by its critical value and plotted against $v = (2\beta - 1)L_N$, where $L_N = N\sigma_N(\beta = 1/2)/C_{\log}(\alpha)$ is determined from the critical EPR, Equation (A64). The dashed curve is the predicted scaling function $f(v) = (1 - e^{-v})/v$ without fitted crossover parameters.

The transition at $\beta_c = 1/2$ is not tied to the polynomial approximation (11) used for the offspring sampling probability. Consider the binomial Wright–Fisher class (14), with the same leading mutation–selection drift $b(X)$ and a remainder uniform up to the fixation boundaries. Such rules have the same limiting stationary density, with $p_0(X) \sim Z^{-1}X^{2\beta-1}$ near $X = 0$. Moreover, Equation (31) gives the same leading boundary scaling, which yields the same Poisson kernel (32). Importantly, $O(N^{-2})$ modifications of the microscopic sampling rule do not alter the competition between the bulk EPR, which scales as N^{-1} , and the boundary EPR, which scales as $N^{-2\beta}$. Hence, the transition point $\beta_c = 1/2$ and the three leading EPR scaling regimes are unchanged. Such modifications can affect subleading quantities, including the coefficient C_1 .

A second test of the finite-size crossover uses the effective exponent of the total EPR. The rescaled quantity $(1 - \theta_{\text{eff}})L_{\text{mid}}$ approaches $\gamma(v) = v/(e^v - 1)$, where $v = (2\beta - 1)L_{\text{mid}}$ and $L_{\text{mid}} = (L_N + L_{2N})/2$. Figure A2b in Appendix B.6 shows this collapse, with smaller finite-size deviations than in Figure 2b. Neither collapse requires fitted crossover parameters, unlike the interpolation used for Figure 3a.

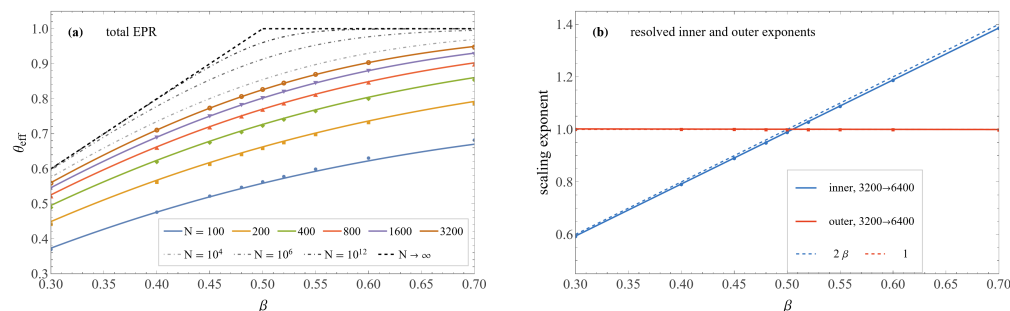


Figure 3. Finite-size crossover and competing scaling exponents for $\alpha = 2$. (a) Effective exponent $\theta_{\text{eff}}(N, \beta) = -\log_2[\sigma_{2N}(\beta)/\sigma_N(\beta)]$. Symbols show exact finite-population results for $N = 100, 200, 400, 800, 1600, 3200$. The solid curves show the crossover interpolation of Appendix B.5, constrained by the endpoint asymptotics, with two constants fitted at $N = 3200$ and kept fixed for the other population sizes. The dash-dotted curves, extrapolated with the same two constants for $N = 10^4, 10^6, 10^{12}$, illustrate the slow sharpening of the crossover near $\beta = 1/2$; the dashed limiting curve is $\theta_{\infty}(\beta) = \min(2\beta, 1)$. (b) Exponents of the separately resolved inner and outer contributions, obtained from the size pair $N = 3200 \rightarrow 6400$. The inner exponent follows 2β , whereas the outer exponent follows 1. The crossing of the limiting curves at $2\beta = 1$ identifies β_c .

4.4. Finite-Population Resolution of the Transition

The asymptotic exponent of the total EPR, $\theta(\beta) = \min(1, 2\beta)$, has a kink at $\beta_c = 1/2$, but the finite- N exponent of the unresolved total is a poor way to locate it. Near β_c , the finite chain contains simultaneously an inner contribution with scale $N^{-2\beta}$ and an outer contribution with scale N^{-1} , together with the crossover region between them. Their relative weight varies on the scale $(2\beta - 1) \ln N$, so that any quantity formed from the total EPR changes smoothly over $|\beta - 1/2| = O(1/\ln N)$. The transition becomes clear at accessible N when the two spatial sectors are resolved separately. Figure 3a shows the finite-size rounding of the asymptotic kink. With two constants fitted at $N = 3200$, the crossover interpolation reproduces the exact effective exponents over the full displayed interval of β within 0.003 for $N \geq 1600$ and within 0.008 for N from 200 to 800; at $N = 100$, the deviations reach 0.012. Extrapolating the interpolation to larger N illustrates the slow sharpening expected from the crossover variable $v = (2\beta - 1) \ln N$. Even population sizes far beyond those accessible to the exact calculation approach the limiting kink only gradually.

To separate the two contributions responsible for the finite-size rounding, we measure the EPR in boundary and bulk sectors. The inner and outer sectors defined in Section 2.5 have independent asymptotic scalings:

$$\sigma_{\text{in}} \propto N^{-2\beta}, \quad \sigma_{\text{out}} \propto N^{-1}, \quad (127)$$

so that

$$\frac{\sigma_{\text{out}}}{\sigma_{\text{in}}} \propto N^{2\beta-1}. \quad (128)$$

Figure A2a tests this prediction using the exact finite-population EPR assigned to the two sectors. The ratio decreases with N below β_c , increases above β_c , and becomes approximately independent of population size at criticality. Equation (127) provides the limiting sector exponents against which the finite- N measurements in Figure 3b are compared.

In the large- N limit, the inner and outer sector exponents cross at the transition:

$$\theta_{\text{in}} = 2\beta, \quad \theta_{\text{out}} = 1, \quad \theta_{\text{in}} = \theta_{\text{out}} \implies \beta_c = \frac{1}{2}. \quad (129)$$

At finite N , the two measured exponents coincide at $\beta = 0.537$, 0.519 , and 0.510 for the size pairs starting at $N = 400$, 800 , and 1600 . The crossing displacement decreases approximately twofold at each doubling of N . Logarithmically enhanced inner-sector corrections can produce a displacement of order $(\ln N)/N$, so that the observed decrease should not be interpreted as evidence for a pure $1/N$ correction. At $\beta = 1/2$ both sectors have the same power N^{-1} , but the total EPR is not a pure N^{-1} law: the logarithmic enhancement comes from the growing number of logarithmic distance scales that carry comparable EPR. This connects the exponent crossing in Figure 3b to the critical shoulder in Figure 1b and the nearly horizontal critical curve in Figure 2a.

5. Discussion

The finite Wright–Fisher chain, its diffusion approximation, the fixed-point Gaussian approximation, and the sequential origin–fixation process describe two-allele mutation–selection dynamics under different approximations and limiting procedures. The chain retains all $N + 1$ population compositions and whole-generation transitions between them. The diffusion replaces these finite jumps by infinitesimal motion on a continuous allele-frequency interval, with time rescaled by N . The fixed-point Gaussian approximation linearizes the dynamics around an equilibrium and replaces state-dependent sampling noise with additive Gaussian noise. The sequential origin–fixation process instead assumes a much lower mutation rate, retaining only the two monomorphic states.

The Moran comparison shows that the number of population states alone does not determine irreversibility. The Moran process has the same $N + 1$ allele-count states as the Wright–Fisher process, but its nearest-neighbor transitions form a birth–death chain, which is reversible [36,37,47]. Whole-generation Wright–Fisher resampling, by contrast, connects distant population compositions [48], creating cycles that can support stationary probability currents. Such currents can occur even under neutral reproduction with symmetric mutation. Thus, the finite-population EPR is not determined by the biological mutation graph or by state-space size alone, but by the transition probabilities on the graph of population compositions.

Earlier population-genetic work connected time reversal to allele ages, extinction and fixation times, genealogies, and backward simulation [36,37,49–51]. Watterson identified two-allele models whose discrete dynamics can be nonreversible even though their diffusion approximations are reversible [35]. The vanishing stationary current of the limiting two-allele diffusion explains its reversibility [52], while weak-mutation substitution models provide a distinct reduction [40,41]. These works did not determine the stationary entropy production of the finite Wright–Fisher population-count chain, its large- N scaling, or its distribution over population compositions.

The analysis in this paper has shown that a large-population approximation can describe the leading allele-frequency dynamics while missing its statistical arrow of time. At stationarity, the distribution of population compositions is constant, but a sequence of compositions and the same sequence in reverse need not be equally probable. For the finite Wright–Fisher process, the entropy production per generation vanishes as the population grows. Its accumulated value over N generations, however, remains nonzero or diverges, depending on the mutation parameter, whereas the limiting diffusion has zero EPR [see Equation (110)].

The limiting Wright–Fisher diffusion remains reversible for every $\beta > 0$. Its zero EPR is not obtained by removing state-dependent noise or finite-time non-Gaussianity, since the diffusion retains both. Rather, its higher short-time moments have precisely the coefficients required to cancel the local forward–backward affinity generated by the state dependence of the Gaussian kernel. For finite-generation Wright–Fisher sampling, this cancellation is incomplete. Over the time corresponding to one generation, the leading third central moments of the two processes have the same allele-frequency dependence, but different numerical coefficients: $\kappa_3 = 1$ for binomial sampling and $\kappa_3 = 3/2$ for the diffusion. For $\beta > 1/2$, Equation (109) shows that the accumulated EPR is proportional to $(\kappa_3 - 3/2)^2$ within the regular class of kernels specified in Section 3.4. Matching the diffusion value cancels the leading EPR, but does not establish exact detailed balance. The fixed-point Gaussian and sequential reductions are also reversible, for different reasons. The former yields a reversible stationary Gaussian process whenever its stationary state exists, while the latter gives a reversible two-state chain for arbitrary positive substitution rates.

We have demonstrated that, in the large-population limit, the entropy production of the finite Wright–Fisher chain exhibits a mutation-controlled boundary-localization transition at $\beta_c = 1/2$. At fixed positive β , states with only a finite number of minority-allele copies become increasingly rare as N grows. Transitions among these states can nevertheless retain a finite forward–backward asymmetry, while the corresponding asymmetry of transitions in the polymorphic bulk decreases with N . The mutation parameter decides which of the two effects prevails. Below β_c , near-fixation populations dominate the EPR; above it, the polymorphic bulk dominates. At criticality, successive logarithmic ranges of minority-allele number contribute comparably, producing the logarithmic enhancement of the total EPR. The value $\beta_c = 1/2$ coincides with the known mutation–drift threshold [29] in the Wright–Fisher diffusion. Here, however, it marks a transition in the distribution of time-reversal asymmetry, not a change in the reversibility of the diffusion.

The transition is of mixed order: the localization order parameter jumps from one to zero at β_c , while the boundary width required to capture a prescribed fraction of the EPR diverges as $\beta \rightarrow \beta_c^-$. This divergence comes from the increasingly extended power-law tail, not from a broadening of the microscopic boundary peak. The finite-size crossover narrows only as $1/\ln N$, so the transition can appear smooth even in large populations. The transition is robust within the binomial Wright–Fisher class (14): $O(N^{-2})$ changes of the offspring probability can modify subleading coefficients such as $\mathcal{C}_1$, but not the leading bulk–boundary competition.

Observing the localization of entropy production in biological systems requires records of transitions between population compositions; stationary allele-frequency distributions alone do not reveal the temporal asymmetry of these transitions. Below β_c , the dominant contribution to the EPR comes from rare near-fixation populations, making resolution of small minority-allele counts particularly important. Near criticality, the contributing transitions span a broad range of minority-allele numbers. Coarse frequency binning or sparse sampling near fixation can make the transition from boundary- to bulk-dominated EPR difficult to detect. High-resolution lineage-tracking methods, such as those demonstrated in Ref. [53], may provide a route toward such measurements, provided that population compositions and their transitions can be resolved.

The phase transition identified in the present paper opens several directions for further research. For three or more alleles, the limiting diffusion can itself support circulating probability currents. How boundary and bulk contributions compete when the diffusion is itself irreversible remains unknown. Another question concerns the alternative large- N mutation scaling behind the sequential weak-mutation reduction. Although this reduction has zero EPR, it excludes rare polymorphic excursions and therefore cannot determine the EPR of the

full Wright–Fisher chain under the same scaling. Whether the full chain exhibits a different EPR scaling or localization transition remains an open question. Further directions include offspring-sampling laws beyond the class studied here, the EPR that can be recovered when population compositions are recorded only every few generations, and the feasibility of testing the boundary-localization transition in controlled microbial populations.

In conclusion, the large-population limit is singular with respect to irreversibility. Although the entropy production per generation vanishes as $N \rightarrow \infty$, its accumulation over N generations does not approach the zero entropy production of the reversible limiting diffusion. Instead, the scaling and spatial distribution of finite-population irreversibility undergo a genuine phase transition at $\beta_c = 1/2$. A reversible continuum limit can thus conceal a nontrivial phase structure of the irreversible stochastic dynamics from which it emerges. In this sense, the boundary-localization transition is not a transition of the diffusion, but a transition in the finite-population irreversibility that the diffusion limit ultimately erases.

Author Contributions: Conceptualization, I.G., R.P., and D.B.S.; Methodology, I.G., R.P., and D.B.S.; Software, I.G., R.P., and D.B.S.; Validation, I.G., R.P., and D.B.S.; Formal analysis, I.G., R.P., and D.B.S.; Investigation, I.G., R.P., and D.B.S.; Writing—original draft, I.G., R.P., and D.B.S.; Writing—review and editing, I.G., R.P., and D.B.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the HESC of Armenia, grants No. 24IRF/2-1C001 and 10-27/24FP-1D009.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: We thank A. Allahverdyan, S. Ito, J. Wang, and D. Wolpert for discussions.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviation is used in this manuscript:

EPR Entropy Production Rate

Appendix A. Large- N EPR in the Two-Allele Wright–Fisher Chain

We derive the leading EPR and its first regular finite-population correction directly from the exact stationary EPR identity. Throughout this appendix, $\epsilon = 1/N$, $s = \alpha\epsilon$, and $\mu = \beta\epsilon$. We use the polynomial sampling rule, the drift and variance functions, and the limiting stationary density $p_0(X)$ defined in Equation (71). The leading bulk EPR exists for $\beta > 1/2$, while the regular two-term expansion requires $\beta > 1$.

Appendix A.1. Leading EPR from Stationary Jump Moments

Let $X = i/N$, $Y = j/N$, $\Delta = Y - X$, $q(X) = \ln[w(X)/(1 - w(X))]$, and $P_{ij}^F = \pi_i W_{ij}$. The relation

$$\left\langle e^{-\mathcal{A}_N} \right\rangle_F = \sum_{ij} P_{ij}^F P_{ji}^F / P_{ij}^F = \sum_{ij} P_{ji}^F = 1 \quad (\text{A1})$$

provides Equation (30). For the binomial kernel (12),

$$\ln \frac{P_{ij}^F}{P_{ji}^F} = \phi(i) - \phi(j) + jq(X) - iq(Y), \quad \phi(i) = \ln \pi_i - \ln \binom{N}{i} + N \ln[1 - w(X)]. \quad (\text{A2})$$

Since $\sum_j P_{ij}^F = \pi_i$ and $\sum_i P_{ij}^F = \pi_j$, the stationary average of $\phi(i) - \phi(j)$ vanishes, which yields the exact identity

$$\sigma_N = N \langle Yq(X) - Xq(Y) \rangle_F. \quad (\text{A3})$$

For any sufficiently smooth function Ψ , stationarity yields $\langle \Psi(Y) - \Psi(X) \rangle_F = 0$. Choose $\Psi'(X) = q(X) - Xq'(X)$. Subtracting $\Psi(Y) - \Psi(X)$ from the integrand in Equation (A3) cancels the terms linear and quadratic in Δ . Taylor expansion at a fixed interior X yields

$$\sigma_N = N \left\langle \frac{q''}{6} \Delta^3 + \frac{q'''}{12} \Delta^4 + \frac{q''''}{40} \Delta^5 + \frac{q'''''}{180} \Delta^6 + \dots \right\rangle_F. \quad (\text{A4})$$

The cancellation follows from stationarity and is exact before the Taylor expansion.

For the binomial Wright–Fisher kernel with $w = X + \epsilon b$, the first moments needed at leading order are

$$\langle \Delta^3 \rangle_X = \epsilon^2 (gg' + 3bg) + O(\epsilon^3), \quad \langle \Delta^4 \rangle_X = 3\epsilon^2 g^2 + O(\epsilon^3). \quad (\text{A5})$$

Writing $q = q_0 + O(\epsilon)$ with $q_0 = \ln[X/(1-X)]$, and denoting the backward generator by

$$\mathcal{L}_0 = b\partial_X + \frac{1}{2}g\partial_X^2,$$

the local coefficient is

$$\frac{gg' + 3bg}{6} q_0'' + \frac{g^2}{4} q_0''' = \frac{(g')^2}{12g} - \frac{1}{2} \mathcal{L}_0 \ln g. \quad (\text{A6})$$

The zero-current identity $(gp_0)' = 2bp_0$ implies $\langle \mathcal{L}_0 \ln g \rangle_{p_0} = 0$ for $\beta > 1/2$; the boundary term vanishes because $gp_0(\ln g)' = O(X^{2\beta-1})$ near $X = 0$, and similarly at $X = 1$. Thus,

$$\sigma_N = \frac{\mathcal{C}(\alpha, \beta)}{N} + o(N^{-1}), \quad \mathcal{C}(\alpha, \beta) = \frac{1}{12} \left\langle \frac{(g')^2}{g} \right\rangle_{p_0}, \quad \beta > \frac{1}{2}. \quad (\text{A7})$$

Appendix A.2. Third Moment of the Wright–Fisher Diffusion

We derive Equation (102) for the Wright–Fisher diffusion with generator [cf. Equation (67)]

$$\mathcal{L} = b(X)\partial_X + [g(X)/2]\partial_X^2. \quad (\text{A8})$$

One Wright–Fisher generation corresponds to the diffusion-time interval $\epsilon = 1/N$. The conditional expectation over this interval follows from the diffusion semigroup,

$$\langle f(Y) \rangle_X^{\text{diff}} = [e^{\epsilon \mathcal{L}} f](X) = \left[1 + \epsilon \mathcal{L} + \frac{\epsilon^2}{2} \mathcal{L}^2 + O(\epsilon^3) \right] f(X). \quad (\text{A9})$$

For the third raw increment moment, set $f_X(Y) = (Y - X)^3$, treating the initial point X as fixed when differentiating with respect to Y . The first action of the generator yields

$$\mathcal{L}f_X(Y) = 3b(Y)(Y - X)^2 + 3g(Y)(Y - X). \quad (\text{A10})$$

At $Y = X$, this expression vanishes, while its first and second derivatives are

$$\partial_Y(\mathcal{L}f_X)|_{Y=X} = 3g(X), \quad \partial_Y^2(\mathcal{L}f_X)|_{Y=X} = 6b(X) + 6g'(X). \quad (\text{A11})$$

Applying the generator once more, we obtain

$$[\mathcal{L}^2 f_X](X) = 3bg + \frac{g}{2}(6b + 6g') = 6bg + 3gg'. \quad (\text{A12})$$

Since $f_X(X) = 0$ and $[\mathcal{L}f_X](X) = 0$, Equation (A9) yields

$$\langle (\Delta X)^3 \rangle_X^{\text{diff}} = \epsilon^2 \left(3bg + \frac{3}{2}gg' \right) + O(\epsilon^3). \quad (\text{A13})$$

The first two raw increment moments are

$$\langle \Delta X \rangle_X^{\text{diff}} = \epsilon b + O(\epsilon^2) \quad \text{and} \quad \langle (\Delta X)^2 \rangle_X^{\text{diff}} = \epsilon g + O(\epsilon^2). \quad (\text{A14})$$

Subtracting the drift-dependent terms from the third raw moment,

$$\langle (\Delta X - \langle \Delta X \rangle_X)^3 \rangle_X = \langle (\Delta X)^3 \rangle_X - 3\langle \Delta X \rangle_X \langle (\Delta X)^2 \rangle_X + 2\langle \Delta X \rangle_X^3, \quad (\text{A15})$$

we find

$$\langle (Y - \langle Y \rangle_X)^3 \rangle_X^{\text{diff}} = \frac{3}{2}\epsilon^2 gg' + O(\epsilon^3), \quad (\text{A16})$$

which proves Equation (102).

For the binomial Wright–Fisher step, the third central moment is instead

$$\frac{w(X)[1 - w(X)][1 - 2w(X)]}{N^2} = \epsilon^2 gg' + O(\epsilon^3). \quad (\text{A17})$$

The coefficient 3/2 in the diffusion result arises from the second-order term of the diffusion semigroup, not from a third-order derivative in its infinitesimal generator. The two processes have different third moments over the time interval corresponding to one Wright–Fisher generation, despite having the same leading drift and variance.

Appendix A.3. Boundary Layers and the Range of the Bulk Expansion

The Taylor expansion above requires $NX \gg 1$ and $N(1 - X) \gg 1$. At the left endpoint, put $i = k = O(1)$, so that $X = k/N$. According to Equation (31), the transition probabilities then approach, for fixed k and j , the boundary Poisson kernel $\mathcal{K}_{kj}^{(L)}$ defined in Equation (32). The fixed- k boundary dynamics is not described by the Gaussian bulk expansion. Matching the stationary probabilities to $p_0(X) \sim Z^{-1}X^{2\beta-1}$ in $1 \ll k \ll N$ yields

$$\pi_k \sim \frac{1}{N}p_0(k/N) \sim \frac{N^{-2\beta}}{Z}k^{2\beta-1}, \quad (\text{A18})$$

$$\pi_k = N^{-2\beta}\rho_k^{(L)}[1 + o(1)] \quad (k \text{ fixed}). \quad (\text{A19})$$

The boundary profile solves the stationary equation for the limiting Poisson kernel,

$$\rho_j^{(L)} = \sum_{k=0}^{\infty} \rho_k^{(L)} \mathcal{K}_{kj}^{(L)}, \quad \rho_k^{(L)} \sim Z^{-1}k^{2\beta-1} \quad (k \rightarrow \infty). \quad (\text{A20})$$

The asymptotic condition fixes its normalization by matching to the bulk; the profile $\rho_k^{(L)}$ is not summable on the half-line for $\beta > 0$. The stationary sum is well defined: for each fixed j , its large- k summand is proportional to $e^{-k}k^{2\beta+j-1}$. The local boundary kernel is independent of α , but the matched normalization Z^{-1} is not.

For fixed k, j , the forward–backward affinity tends to

$$\ln \frac{\pi_k W_{kj}}{\pi_j W_{jk}} = \ln \frac{\rho_k^{(L)} \mathcal{K}_{kj}^{(L)}}{\rho_j^{(L)} \mathcal{K}_{jk}^{(L)}} + o(1). \quad (\text{A21})$$

Using the pairwise EPR $e_{kj}^{(N)}$ defined in the main text, the contribution from the left boundary region $0 \leq k < j \leq M$, with M fixed, becomes

$$\sigma_{N,M}^{(L)} = N^{-2\beta} \mathcal{C}_M^{(L)} + o(N^{-2\beta}), \quad (\text{A22})$$

$$\mathcal{C}_M^{(L)} = \sum_{0 \leq k < j \leq M} (\rho_k^{(L)} \mathcal{K}_{kj}^{(L)} - \rho_j^{(L)} \mathcal{K}_{jk}^{(L)}) \ln \frac{\rho_k^{(L)} \mathcal{K}_{kj}^{(L)}}{\rho_j^{(L)} \mathcal{K}_{jk}^{(L)}}. \quad (\text{A23})$$

The coefficient $\mathcal{C}_M^{(L)}$ is independent of N and describes the EPR carried by a fixed number of states near the left fixation boundary. At the right endpoint, setting $i = N - k$ gives $1 - w(1 - k/N) = (k + \beta)/N + O(N^{-2})$ and the same leading Poisson kernel for the minority allele. The contribution of the right boundary contribution also scales as $N^{-2\beta}$; its normalization can differ because selection weights the two endpoints differently.

The endpoint behavior of the bulk integrand provides an independent power count:

$$p_0(X) \frac{[g'(X)]^2}{g(X)} \sim Z^{-1} X^{2\beta-2} \quad (X \rightarrow 0), \quad (\text{A24})$$

with the corresponding factor $e^{2\alpha}/Z$ near $X = 1$. For $\beta > 1/2$, the part omitted by cutting the bulk integral off at $X \sim N^{-1}$ is $O(N^{-2\beta}) = o(N^{-1})$. At $\beta = 1/2$, integration from $X \sim N^{-1}$ produces $(\ln N)/N$, and for $0 < \beta < 1/2$ the boundary term $N^{-2\beta}$ dominates the nominal bulk N^{-1} term. The critical logarithmic amplitude is given in Equation (115); the boundary-dominated amplitude requires the Poisson stationary solution rather than a divergent bulk integral.

The integrands of the regular N^{-2} correction behave as $X^{2\beta-3}$ near an endpoint. Its boundary omission is $O(N^{-2\beta})$ for $\beta > 1$. Thus, the regular two-term expansion is valid for $\beta > 1$; for $1/2 < \beta < 1$ its first correction is instead $O(N^{-2\beta})$, and for $\beta = 1$ it is $O[(\ln N)/N^2]$. The threshold $\beta = 1$ concerns subleading terms, not the leading boundary-localization transition.

Appendix A.4. First Finite-Population Correction

For $\beta > 1$, write the interior stationary density as

$$p_N(X) = p_0(X)[1 + \epsilon r(X) + O(\epsilon^2)], \quad \langle r \rangle_{p_0} = 0. \quad (\text{A25})$$

Let $(P_N f)(X) = \langle f(Y) \rangle_X$ be the one-generation backward operator. Its Kramers–Moyal expansion is

$$\epsilon^{-1}(P_N - I) = \mathcal{L}_0 + \epsilon \mathcal{L}_1 + O(\epsilon^2),$$

with

$$\mathcal{L}_0 f = b f' + \frac{g}{2} f'', \quad \mathcal{L}_1 f = \frac{b^2 + b g'}{2} f'' + \frac{g g' + 3 b g}{6} f''' + \frac{g^2}{8} f'''. \quad (\text{A26})$$

The stationary equation at order ϵ is

$$\mathcal{L}_0^*(p_0 r) = -\mathcal{L}_1^* p_0. \quad (\text{A27})$$

Using $(gp_0)' = 2bp_0$ and integrating Equation (A27) once, with the integration constant fixed by zero stationary current, yields

$$\frac{gp_0}{2}r_0' = -\frac{1}{2}[(b^2 + bg')p_0]' + \frac{1}{6}[(gg' + 3bg)p_0]'' - \frac{1}{8}(g^2p_0)''' \quad (\text{A28})$$

The normalized solution is

$$r(X) = r_0(X) - \langle r_0 \rangle_{p_0}, \quad (\text{A29})$$

with

$$r_0(X) = -\frac{2\alpha^2}{3}g(X) + \frac{\alpha(22\beta + 1)}{3}X + \frac{5\beta(2\beta - 1)}{6g(X)} + \frac{\beta(14\beta - 1)}{3}\ln g(X) - \frac{7\alpha\beta}{3}\ln \frac{X}{1-X}. \quad (\text{A30})$$

To expand the EPR identity (A4), set $q_0 = \ln[X/(1-X)]$ and $q_1 = b/g$, so that $q = q_0 + \epsilon q_1 + O(\epsilon^2)$. The required conditional jump moments are

$$\begin{aligned} \langle \Delta^3 \rangle_X &= \epsilon^2 m_3 + \epsilon^3 d_3 + O(\epsilon^4), & \langle \Delta^4 \rangle_X &= \epsilon^2 m_4 + \epsilon^3 d_4 + O(\epsilon^4), \\ \langle \Delta^5 \rangle_X &= \epsilon^3 d_5 + O(\epsilon^4), & \langle \Delta^6 \rangle_X &= \epsilon^3 d_6 + O(\epsilon^4), \end{aligned} \quad (\text{A31})$$

where

$$\begin{aligned} m_3 &= gg' + 3bg, & m_4 &= 3g^2, & d_3 &= b[(g')^2 - 2g] + 3b^2g' + b^3, \\ d_4 &= g(1 - 6g) + 10bgg' + 6b^2g, & d_5 &= 10g^2g' + 15bg^2, & d_6 &= 15g^3. \end{aligned} \quad (\text{A32})$$

Substitution into Equation (A4) yields the local N^{-2} coefficient

$$B(X) = \frac{d_3q_0'' + m_3q_1''}{6} + \frac{d_4q_0''' + m_4q_1'''}{12} + \frac{d_5q_0''''}{40} + \frac{d_6q_0''''}{180}, \quad (\text{A33})$$

and hence,

$$\sigma_N = \frac{C(\alpha, \beta)}{N} + \frac{C_1(\alpha, \beta)}{N^2} + o(N^{-2}), \quad (\text{A34})$$

$$C_1 = \langle B + rA \rangle_{p_0}, \quad \beta > 1. \quad (\text{A35})$$

Stationarity at order ϵ yields $\langle r\mathcal{L}_0 \ln g \rangle_{p_0} = -\langle \mathcal{L}_1 \ln g \rangle_{p_0}$. Using Equation (A6), the coefficient becomes

$$C_1 = \left\langle B + \frac{1}{2}\mathcal{L}_1 \ln g \right\rangle_{p_0} + \frac{1}{12} \left\langle \frac{(g')^2}{g} r \right\rangle_{p_0}. \quad (\text{A36})$$

Using $(gp_0)' = 2bp_0$ and integrating by parts, with all boundary terms vanishing for $\beta > 1$, we obtain

$$\left\langle B + \frac{1}{2}\mathcal{L}_1 \ln g \right\rangle_{p_0} = \langle \chi_{\alpha, \beta} \rangle_{p_0},$$

where

$$\chi_{\alpha, \beta}(X) = \frac{1}{24g^2} [9 - 38g + 22g^2 - 2\alpha gg'(5 - 6g) - 2\beta(g')^2(9 - 10g) + 6(g')^2b^2]. \quad (\text{A37})$$

The reduction is an equality of stationary averages, not a pointwise identity for $\chi_{\alpha, \beta}$ and

$$B + \frac{1}{2}\mathcal{L}_1 \ln g.$$

All boundary terms vanish for $\beta > 1$. With $(g')^2/g = 1/X + 1/(1-X) - 4$, Equations (A29) and (A36) yield

$$C_1(\alpha, \beta) = \langle \chi_{\alpha, \beta} \rangle_{p_0} + \frac{1}{12} \left[\left\langle \frac{r_0}{X} \right\rangle_{p_0} + \left\langle \frac{r_0}{1-X} \right\rangle_{p_0} \right] - \frac{1}{12} \left(\left\langle \frac{1}{X} \right\rangle_{p_0} + \left\langle \frac{1}{1-X} \right\rangle_{p_0} \right) \langle r_0 \rangle_{p_0}. \quad (\text{A38})$$

Appendix A.5. Analytical Evaluation with Beta–Kummer Moments

Define the normalized moments

$$\mathcal{I}_{mn}(\alpha, \beta) = \langle X^m (1-X)^n \rangle_{p_0} = \frac{B(2\beta+m, 2\beta+n)}{B(2\beta, 2\beta)} \frac{{}_1F_1(2\beta+m; 4\beta+m+n; 2\alpha)}{{}_1F_1(2\beta; 4\beta; 2\alpha)}. \quad (\text{A39})$$

Here, $B(\lambda, \tilde{\lambda})$ denotes Euler's beta function and ${}_1F_1(\lambda; \Lambda; \zeta)$ Kummer's confluent hypergeometric function. The leading EPR coefficient $C = \langle (g')^2/g \rangle_{p_0}/12$ becomes

$$C(\alpha, \beta) = \frac{1}{12} [\mathcal{I}_{-1,-1}(\alpha, \beta) - 4], \quad \beta > 1/2. \quad (\text{A40})$$

The rational function $\chi_{\alpha, \beta}(X)$ defined in Equation (A37) can be expressed in beta–Kummer moments without further integration:

$$\begin{aligned} \langle \chi_{\alpha, \beta} \rangle_{p_0} = & \alpha^2 \mathcal{I}_{20} + (-\alpha^2 + 4\alpha\beta - \alpha) \mathcal{I}_{10} + \left(\frac{\alpha^2}{4} - 2\alpha\beta + \frac{\alpha}{2} + 4\beta^2 - \frac{10\beta}{3} + \frac{11}{12} \right) \\ & + \frac{6\alpha\beta - 5\alpha - 18\beta^2 + 28\beta - 10}{12} \mathcal{I}_{-1,0} + \frac{-6\alpha\beta + 5\alpha - 18\beta^2 + 28\beta - 10}{12} \mathcal{I}_{0,-1} \\ & + \frac{2\beta^2 - 6\beta + 3}{8} (\mathcal{I}_{-2,0} + \mathcal{I}_{0,-2}). \end{aligned} \quad (\text{A41})$$

The three averages containing r_0 follow from $R_{mn} = \langle X^m (1-X)^n r_0(X) \rangle_{p_0}$:

$$\begin{aligned} R_{mn} = & \frac{2\alpha^2}{3} \mathcal{I}_{m+2,n} + \frac{\alpha(-2\alpha + 22\beta + 1)}{3} \mathcal{I}_{m+1,n} + \frac{5\beta(2\beta - 1)}{6} (\mathcal{I}_{m-1,n} + \mathcal{I}_{m,n-1}) \\ & - \frac{\beta(7\alpha - 14\beta + 1)}{3} \frac{\partial \mathcal{I}_{mn}}{\partial m} + \frac{\beta(7\alpha + 14\beta - 1)}{3} \frac{\partial \mathcal{I}_{mn}}{\partial n}. \end{aligned} \quad (\text{A42})$$

Only R_{00} , $R_{-1,0}$, and $R_{0,-1}$ are needed. Equation (A38) then becomes

$$C_1(\alpha, \beta) = \langle \chi_{\alpha, \beta} \rangle_{p_0} + \frac{R_{-1,0} + R_{0,-1} - (\mathcal{I}_{-1,0} + \mathcal{I}_{0,-1})R_{00}}{12}, \quad \beta > 1. \quad (\text{A43})$$

Equations (A41) and (A42) provide all terms on the right-hand side in terms of $\mathcal{I}_{mn}$ and their derivatives.

For direct evaluation of the derivatives, set

$$\lambda = 2\beta + m, \quad \tilde{\lambda} = 2\beta + n, \quad \Lambda = \lambda + \tilde{\lambda} = 4\beta + m + n, \quad \zeta = 2\alpha. \quad (\text{A44})$$

Differentiating Equation (A39) yields

$$\frac{\partial \mathcal{I}_{mn}}{\partial m} = \mathcal{I}_{mn} [\psi(\lambda) - \psi(\Lambda) + (\partial_\lambda + \partial_\Lambda) \ln {}_1F_1(\lambda; \Lambda; \zeta)], \quad (\text{A45})$$

$$\frac{\partial \mathcal{I}_{mn}}{\partial n} = \mathcal{I}_{mn} [\psi(\tilde{\lambda}) - \psi(\Lambda) + \partial_\Lambda \ln {}_1F_1(\lambda; \Lambda; \zeta)], \quad (\text{A46})$$

where $\psi(x) = d \ln \Gamma(x)/dx$ is the digamma function. The parameter derivatives of the Kummer function are

$$\frac{\partial}{\partial \lambda} {}_1F_1(\lambda; \Lambda; \zeta) = \sum_{k=1}^{\infty} \frac{(\lambda)_k}{(\Lambda)_k} \frac{\zeta^k}{k!} [\psi(\lambda + k) - \psi(\lambda)], \quad (\text{A47})$$

$$\frac{\partial}{\partial \Lambda} {}_1F_1(\lambda; \Lambda; \zeta) = \sum_{k=1}^{\infty} \frac{(\lambda)_k}{(\Lambda)_k} \frac{\zeta^k}{k!} [\psi(\Lambda) - \psi(\Lambda + k)], \quad (\text{A48})$$

where $(x)_k = \Gamma(x+k)/\Gamma(x)$. Equations (A39) and (A41)–(A48) express $\mathcal{C}_1(\alpha, \beta)$ in terms of beta–Kummer functions and convergent series, for all real α and $\beta > 1$.

The transformation $X \mapsto 1 - X$ sends $\alpha \mapsto -\alpha$ and implies $\mathcal{I}_{mn}(\alpha, \beta) = \mathcal{I}_{nm}(-\alpha, \beta)$. Both EPR coefficients are even in selection:

$$\mathcal{C}(\alpha, \beta) = \mathcal{C}(-\alpha, \beta), \quad \mathcal{C}_1(\alpha, \beta) = \mathcal{C}_1(-\alpha, \beta). \quad (\text{A49})$$

For $\alpha = 0$, the moments reduce to beta-function ratios, with

$$\mathcal{C}(0, \beta) = \frac{1}{6(2\beta - 1)}, \quad \mathcal{C}_1(0, \beta) = -\frac{\beta(3\beta - 1)(6\beta - 5)}{6(\beta - 1)(2\beta - 1)^2}. \quad (\text{A50})$$

The pole at $\beta = 1$ marks the correction-to-scaling threshold discussed in Appendix A.3; $\mathcal{C}_1(0, \beta) < 0$ for every $\beta > 1$. For $\alpha = \beta = 30$, Equation (A41) yields $\langle \chi_{\alpha, \beta} \rangle_{p_0} \approx -6.8323761601$. Equation (A42) yields $R_{00} \approx 3149.0904003$, $R_{-1,0} \approx 5126.1454500$, and $R_{0,-1} \approx 8329.2331744$. Their contribution to Equation (A43) is 0.1076483595, yielding $\mathcal{C}_1(30, 30) \approx -6.7247278006$, in agreement with Equation (100).

Appendix B. Numerical Checks and Robustness

The numerical results in Figures 1–3 are obtained from the full fitness-weighted Wright–Fisher sampling rule (9). The finite- N tables use the polynomial rule (11). This appendix compares the two microscopic rules and gives the definitions needed for the compensated spatial profile and the finite-size crossover curves used in the main text. The stationary distributions are calculated from the exact finite transition matrices, as described in Section 2.5.

Appendix B.1. Numerical Implementation

The finite- N transition probabilities are evaluated in logarithmic form and normalized by a log-sum-exp procedure. For the total-EPR scans, the stationary distribution is obtained by repeated application of the exact transition matrix. Once the L^1 increments enter their asymptotic geometric regime, the ratio of successive increments is used to estimate and accelerate the remaining single-mode tail. The extrapolated vector is returned to the ordinary iteration, which is continued until the estimated remaining error is below 10^{-15} . The final distribution is checked independently for normalization and against the stationarity residual of the exact finite matrix. The discretized stationary density of the limiting diffusion is used as the initial vector; for $\beta < 1/2$, the endpoint values are initialized from the nearest interior values because the density diverges at the boundaries.

For the pair-resolved calculations, the stationary equations are solved directly by LU factorization, with one redundant stationarity equation replaced by normalization. The iterative and direct stationary distributions agree on their common range. In evaluating Equation (34), the forward and reversed fluxes are retained in logarithmic form as long as possible. Differences between nearly equal fluxes are evaluated by a small-argument expansion, and the pair contributions are accumulated with compensated summation. No probability-current cutoff is introduced. As an independent check, the total EPR is also

evaluated from the exact identity (A3), using the same stationary distribution and transition matrix; the pairwise sum and the identity agree within numerical precision.

Table A1. Finite- N tests of the critical and boundary-dominated regimes for the polynomial sampling rule (11). Upper part: at $\beta = 1/2$, the logarithmic slope $S_N = (2N\sigma_{2N} - N\sigma_N) / \ln 2$, compared with the critical amplitude $\mathcal{C}_{\log}(\alpha)$ of Equation (115). Lower part: for $\beta = 0.3$, the rescaled EPR $N^{2\beta}\sigma_N$ and its extrapolated boundary amplitude $\mathcal{C}_{\text{bdry}}(\alpha, \beta)$.

$\beta = 1/2$: slope S_N for $\alpha = 0, 1, 2, 5$ (top to bottom)								
$\alpha \setminus N$	100	200	400	800	1600	3200	6400	$\mathcal{C}_{\log}(\alpha)$
0	0.1676	0.1686	0.1685	0.1681	0.1677	0.1673	0.1671	0.1667
1	0.2190	0.2216	0.2219	0.2214	0.2208	0.2201	0.2197	0.2188
2	0.3421	0.3497	0.3516	0.3511	0.3499	0.3487	0.3477	0.3458
5	0.7882	0.8289	0.8451	0.8487	0.8471	0.8439	0.8407	0.8334
$\beta = 0.3$: $N^{2\beta}\sigma_N$ for $\alpha = 0, 1, 2, 5$ (top to bottom)								
$\alpha \setminus N$	100	200	400	800	1600	3200	6400	$\mathcal{C}_{\text{bdry}}(\alpha, \beta)$
0	0.1452	0.1610	0.1732	0.1825	0.1895	0.1948	0.1989	0.212
1	0.1687	0.1911	0.2084	0.2216	0.2317	0.2393	0.2451	0.263
2	0.2151	0.2513	0.2798	0.3017	0.3184	0.3310	0.3405	0.370
5	0.3115	0.3907	0.4550	0.5054	0.5441	0.5736	0.5958	0.666

Appendix B.2. Finite- N Amplitude Checks

The tables in the article use the polynomial-rule calculation, Equation (11). Table 1 tests the regular bulk expansion, and Table A1 the critical and boundary-dominated regimes. In Table A1, the displayed results use $N = 100, 200, 400, 800, 1600, 3200, 6400$ and $\alpha = 0, 1, 2, 5$. At $\beta = 1/2$, we use

$$S_N = \frac{2N\sigma_{2N} - N\sigma_N}{\ln 2}. \quad (\text{A51})$$

For $\beta = 0.3$, the boundary amplitude is obtained by removing the leading N^{-1} bulk correction. Writing

$$Q_N = N^{2\beta}\sigma_N = \mathcal{C}_{\text{bdry}}(\alpha, \beta) + c_{\text{sub}}N^{-(1-2\beta)} + \dots, \quad (\text{A52})$$

two successive population sizes yield

$$\mathcal{C}_{\text{bdry}}(\alpha, \beta) \simeq \frac{Q_{2N} - 2^{-(1-2\beta)}Q_N}{1 - 2^{-(1-2\beta)}}. \quad (\text{A53})$$

The values quoted in Table A1 use $N = 3200$ and 6400 . The successive increments of Q_N are also consistent with the predicted ratio r . At $N = 6400$, the critical slopes differ from $\mathcal{C}_{\log}(\alpha)$ by less than about 0.9% for the four values of α shown. Repeating the boundary-amplitude extrapolation with the preceding pair $N = 1600, 3200$ changes the quoted values of $\mathcal{C}_{\text{bdry}}$ by at most 6×10^{-4} .

Appendix B.3. Robustness to the Microscopic Sampling Rule

The full fitness-weighted rule (9) and the polynomial rule (11) differ by $O(N^{-2})$ under the scaling $s = \alpha/N$ and $\mu = \beta/N$. Figure A1 compares the two rules directly. The total EPRs approach one another as N increases, while the separately resolved inner and outer contributions have the same exponents for both rules. The crossing of the boundary exponent 2β and the bulk exponent 1 remains at $\beta_c = 1/2$.

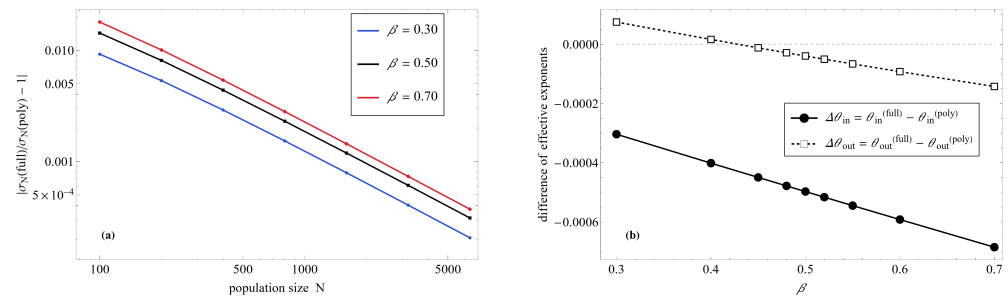


Figure A1. Robustness with respect to the microscopic Wright–Fisher sampling rule for $\alpha = 2$. The full fitness-weighted rule, Equation (9), is compared with the polynomial rule, Equation (11). (a) Relative difference $|\sigma_N^{\text{full}} / \sigma_N^{\text{poly}} - 1|$ of the total EPR for $\beta = 0.3, 0.5, 0.7$. The difference decreases as $1/N$. (b) Differences between the effective inner and outer scaling exponents obtained from the two rules, $\Delta\theta_{\text{in}} = \theta_{\text{in}}^{\text{full}} - \theta_{\text{in}}^{\text{poly}}$ and $\Delta\theta_{\text{out}} = \theta_{\text{out}}^{\text{full}} - \theta_{\text{out}}^{\text{poly}}$, evaluated from the pair $N = 3200, 6400$. Both differences remain below 7×10^{-4} over the displayed range of β , showing that the extracted boundary and bulk scaling exponents are essentially unchanged by the choice of sampling rule.

Appendix B.4. Compensated Spatial Profile

The smooth envelope used to compensate the shell distribution in Figure 2a follows from the bulk EPR density. By Equation (97), the leading EPR per unit allele frequency is

$$\frac{d\sigma_N}{dX} \simeq \frac{1}{12N} p_0(X) \frac{[g'(X)]^2}{g(X)} = \frac{e^{2\alpha X}}{12NZ} [X(1-X)]^{2\beta-2} (1-2X)^2. \quad (\text{A54})$$

This local expression applies for every $\beta > 0$. The restriction $\beta > 1/2$ arises only when it is integrated over the full allele-frequency interval to obtain the leading total EPR.

A shell at distance M from fixation receives contributions from both $X = M/N$ and $X = 1 - M/N$. Writing $x = M/N$ and multiplying their combined density by x to obtain the EPR per logarithmic interval of M , we find

$$\sigma_N G_N(M) \simeq \frac{N^{-2\beta} M^{2\beta-1}}{12Z} \mathcal{E}_{\alpha,\beta}(M/N), \quad (\text{A55})$$

where

$$\mathcal{E}_{\alpha,\beta}(x) = (1-x)^{2\beta-2} (1-2x)^2 \left[e^{2\alpha x} + e^{2\alpha(1-x)} \right]. \quad (\text{A56})$$

For $1 \ll M \ll N$, the envelope approaches $1 + e^{2\alpha}$, leaving the predicted $M^{2\beta-1}$ power law. At larger M/N , the envelope accounts for the smooth variation of the bulk EPR density across the allele-frequency interval. The compensated profile is

$$\tilde{G}_N(M) = \frac{G_N(M)}{\mathcal{E}_{\alpha,\beta}(M/N)}. \quad (\text{A57})$$

It scales as $\tilde{G}_N(M) \propto M^{2\beta-1}$ in the overlap region. Figure 2a normalizes this quantity by its value at $M = 16$, removing the nonuniversal amplitude and exposing the exponent $2\beta - 1$ directly.

Appendix B.5. Finite-Size Crossover in Figure 3a

The endpoint analysis determines the limiting EPR exponents and the near-critical crossover variable $v = (2\beta - 1) \ln N$, but not the finite- N corrections to the crossover. To compare the effective exponents at accessible population sizes, we use an interpolation motivated by the endpoint integral. Writing $\delta = 2\beta - 1$, consider the singular contribu-

tion with an effective microscopic cutoff $k_m(\delta)/N$ and a fixed upper cutoff x_* , refining Equation (124) as follows:

$$\int_{k_m(\delta)/N}^{x_*} X^{\delta-1} dX = x_*^\delta \frac{1 - \exp[-\delta \ell_N(\delta)]}{\delta}, \quad \ell_N(\delta) = \ln N + \ln \frac{x_*}{k_m(\delta)}. \quad (\text{A58})$$

At $\delta = 0$, this effective logarithm plays the same role as L_N , introduced in Appendix B.6 and determined directly from the critical EPR. The difference is that $\ell_N(\delta)$ allows the matching scale to depend on β , whereas L_N is defined at $\beta = 1/2$ and requires no fitted parameters. The cutoff represents matching between the Poisson boundary region and the bulk approximation. Its precise value is not fixed by the leading endpoint analysis. The factor x_*^δ is independent of N and cancels when two population sizes are compared.

Near $\delta = 0$, we approximate the finite part of the effective logarithm by

$$\ell_N(\delta) = \ln N + a_0 + a_1 \delta. \quad (\text{A59})$$

Here, a_0 represents the finite shift of the critical logarithm, while a_1 allows the effective matching scale to vary with β . This truncation is a phenomenological approximation to finite-size corrections, not a derivation of the two coefficients from the boundary Poisson kernel. The resulting interpolation is

$$\mathcal{F}_N(\delta) = \begin{cases} \frac{1 - \exp\{-\delta[\ln N + a_0 + a_1 \delta]\}}{\delta}, & \delta \neq 0, \\ \ln N + a_0, & \delta = 0. \end{cases} \quad (\text{A60})$$

It is continuous at $\delta = 0$ and reproduces the leading N dependence on both sides of the transition. Approximating the total EPR by $\sigma_N(\beta) \propto N^{-1} \mathcal{F}_N(\delta)$, with an N -independent proportionality factor, the definition of the two-size effective exponent yields

$$\theta_{\text{th}}(N, \beta) = 1 - \log_2 \left[\frac{\mathcal{F}_{2N}(2\beta - 1)}{\mathcal{F}_N(2\beta - 1)} \right]. \quad (\text{A61})$$

The two constants,

$$a_0 = -2.66865, \quad a_1 = -0.69131, \quad (\text{A62})$$

are fitted to the exact effective exponents at $N = 3200$ for the full sampling rule and kept fixed for all other population sizes. The fitted value of a_0 corresponds to an effective microscopic scale $e^{-a_0} \simeq 14$, consistent with the scale extracted from the critical EPR in Appendix B.6. Figure 3a compares the resulting interpolation with the exact finite- N effective exponents. Its agreement across population sizes tests the usefulness of this two-parameter description of finite-size corrections. The curves for $N = 10^4, 10^6, 10^{12}$ extrapolate the fitted interpolation beyond the range of exact calculations; they are not parameter-free analytical predictions. The limiting exponent $\theta_\infty(\beta) = \min(2\beta, 1)$ follows independently from Equation (116).

Appendix B.6. Finite-Size Data Collapse

To incorporate the finite part of $N\sigma_N$ at criticality, we introduce an effective microscopic cutoff k_0 that absorbs contributions from the boundary core and the regular bulk. With $L_N = \ln(N/k_0)$, the refined form of Equation (125) is

$$N\sigma_N \simeq C_{\log}(\alpha) \frac{1 - e^{-v}}{2\beta - 1}, \quad v = (2\beta - 1)L_N. \quad (\text{A63})$$

At $\beta = 1/2$, this reduces to $N\sigma_N \simeq \mathcal{C}_{\log}(\alpha)L_N$, where the effective logarithm is provided by

$$L_N = \frac{N\sigma_N(\beta = 1/2)}{\mathcal{C}_{\log}(\alpha)}, \quad (\text{A64})$$

which incorporates this finite contribution without fitting a crossover parameter. For the full sampling rule at $\alpha = 2$, the finite- N estimate $\ln N - L_N$ of $\ln k_0$ varies from 2.65 to 2.59 between $N = 100$ and $N = 6400$, giving an effective cutoff $k_0 \approx 14$. The constant a_0 of Equation (A60) plays the role of $-\ln k_0$ in the asymptotic form, calibrated from the critical EPR.

The variable v of Equation (A63) is the scaling variable of Equation (122) with $\ln N$ replaced by L_N ; the two coincide when $L_N/\ln N \rightarrow 1$. At the population sizes studied here, however, this ratio increases from 0.56 at $N = 400$ to 0.70 at $N = 6400$. Evaluating v with $\ln N$ then introduces substantial size-dependent corrections to both the crossover variable and the normalization of the EPR. Using L_N , determined directly from the critical EPR, accounts for these corrections. Dividing Equation (A63) by its critical value yields $N\sigma_N/[\mathcal{C}_{\log}(\alpha)L_N] = f(v)$ with $f(v) = (1 - e^{-v})/v$, and by Equation (A64) the left-hand side equals $\sigma_N(\beta)/\sigma_N(1/2)$, the quantity plotted in Figure 2b.

The same finite-size scaling form can also be tested through the effective exponent $\theta_{\text{eff}}(N, \beta)$ from Equation (40), extracted from the EPR at population sizes N and $2N$:

$$\theta_{\text{eff}}(N, \beta) = -\log_2[\sigma_{2N}(\beta)/\sigma_N(\beta)]. \quad (\text{A65})$$

Since $L_{2N} - L_N \simeq \ln 2$, a centered finite-difference approximation to Equation (A63) yields

$$[1 - \theta_{\text{eff}}(N, \beta)]L_{\text{mid}} \simeq \gamma(v), \quad \gamma(v) = \frac{v}{e^v - 1}, \quad v = (2\beta - 1)L_{\text{mid}}, \quad (\text{A66})$$

with $L_{\text{mid}} = (L_N + L_{2N})/2$. Evaluating v at the midpoint of the size pair removes the first-order effect of the discrete difference. Any N -independent multiplicative factor in Equation (A63) cancels in σ_{2N}/σ_N . Corrections remain from the regular part of the EPR, the finite-difference approximation, and the deviation of $L_{2N} - L_N$ from $\ln 2$.

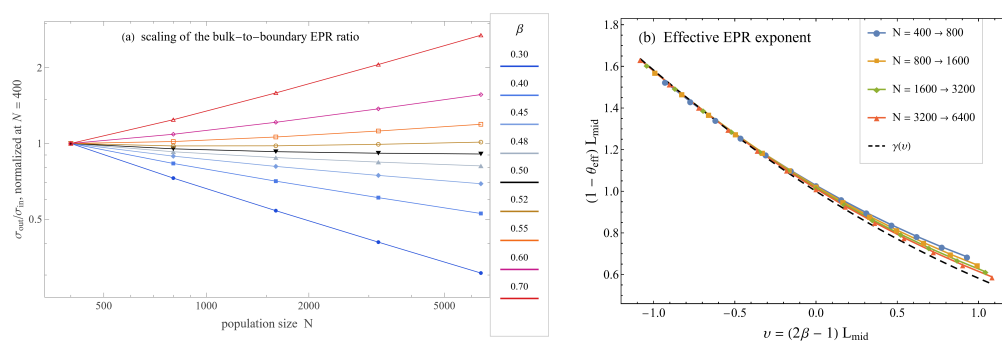


Figure A2. Complementary finite-size tests of the boundary-localization transition for $\alpha = 2$, cf. Figure 2: scaling of the ratio of the outer and inner EPR contributions and collapse of the effective exponent of the total EPR for the size pairs $N \rightarrow 2N$. **(a)** Ratio $\sigma_{\text{out}}/\sigma_{\text{in}}$, normalized at $N = 400$. The inner sector contains pairs whose midpoint lies within 10 individuals of either fixation boundary; the outer sector contains pairs whose midpoint is at least $0.1N$ from fixation. The predicted scaling is $N^{2\beta-1}$. **(b)** The rescaled exponent $[1 - \theta_{\text{eff}}]L_{\text{mid}}$ is plotted against $v = (2\beta - 1)L_{\text{mid}}$, where $L_{\text{mid}} = (L_N + L_{2N})/2$. The dashed curve is $\gamma(v) = v/(e^v - 1)$, Equation (A66), without fitted parameters.

The procedure contains no adjustable parameters and no fitted crossover scale. In Figure A2, for each N , the exact stationary EPR is computed as described in Section 2.5, at

$\beta = 1/2$ and at values of β for which $(2\beta - 1) \ln N$ runs over a uniform grid in $[-1.5, 1.5]$, and L_N follows from Equation (A64). Figure 2b displays $N\sigma_N/[C_{\log}(\alpha)L_N]$ against $v = (2\beta - 1)L_N$. For Figure A2b, σ_{2N} is computed at the same values of β , and $(1 - \theta_{\text{eff}})L_{\text{mid}}$ is plotted against $(2\beta - 1)L_{\text{mid}}$.

The data collapse well onto the predicted scaling functions, especially for the effective-exponent form in Figure A2b, with residual deviations that decrease systematically as N increases. In Figure 2b, the maximum relative deviation from $f(v)$ over the sampled β grid decreases from 14% at $N = 400$ to 7% at $N = 6400$. Its product with L_N stays between 0.45 and 0.48, consistent with a correction to scaling of order $1/\ln(N/k_0)$. In Figure A2b, the relative deviation from $\gamma(v)$ is below 3% for $v \leq 0$ for all size pairs and below 1.2% for the pair $N = 3200 \rightarrow 6400$; near $v = 1$ it decreases from 13% for $N = 400 \rightarrow 800$ to 6% for $N = 3200 \rightarrow 6400$.

References

1. Eigen, M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **1971**, *58*, 465–523. <https://doi.org/10.1007/BF00623322>.
2. Eigen, M.; McCaskill, J.S.; Schuster, P. The molecular quasi-species. In *Advances in Chemical Physics*; Prigogine, I., Rice, S.A., Eds.; John Wiley & Sons: New York, NY, USA, 1989; Volume 75, pp. 149–263. <https://doi.org/10.1002/9780470141243.ch4>.
3. Leuthäusser, I. Statistical mechanics of Eigen’s evolution model. *J. Stat. Phys.* **1987**, *48*, 343–360. <https://doi.org/10.1007/BF01010413>.
4. Tarazona, P. Error thresholds for molecular quasispecies as phase transitions: From simple landscapes to spin-glass models. *Phys. Rev. A* **1992**, *45*, 6038–6050. <https://doi.org/10.1103/PhysRevA.45.6038>.
5. Baake, E.; Baake, M.; Wagner, H. Ising quantum chain is equivalent to a model of biological evolution. *Phys. Rev. Lett.* **1997**, *78*, 559–562. <https://doi.org/10.1103/PhysRevLett.78.559>.
6. Baake, E.; Wagner, H. Mutation-selection models solved exactly with methods of statistical mechanics. *Genet. Res.* **2001**, *78*, 93–117. <https://doi.org/10.1017/S0016672301005110>.
7. Saakian, D.B.; Hu, C.K. Eigen model as a quantum spin chain: Exact dynamics. *Phys. Rev. E* **2004**, *69*, 021913. <https://doi.org/10.1103/PhysRevE.69.021913>.
8. Park, J.M.; Deem, M.W. Phase diagrams of quasispecies theory with recombination and horizontal gene transfer. *Phys. Rev. Lett.* **2007**, *98*, 058101. <https://doi.org/10.1103/PhysRevLett.98.058101>.
9. Saakian, D.B. Evolution models with base substitutions, insertions, deletions, and selection. *Phys. Rev. E* **2008**, *78*, 061920. <https://doi.org/10.1103/PhysRevE.78.061920>.
10. Saakian, D.B.; Kirakosyan, Z.; Hu, C.K. Evolution models with lethal mutations on symmetric or random fitness landscapes. *Phys. Rev. E* **2012**, *86*, 031920. <https://doi.org/10.1103/PhysRevE.86.031920>.
11. Saakian, D.B.; Ghazaryan, M.H.; Hu, C.K. Punctuated equilibrium and shock waves in molecular models of biological evolution. *Phys. Rev. E* **2014**, *90*, 022712. <https://doi.org/10.1103/PhysRevE.90.022712>.
12. Saakian, D.B.; Hu, C.K. Mathematical models of quasi-species theory and exact results for the dynamics. In *Quasispecies: From Theory to Experimental Systems*; Current Topics in Microbiology and Immunology; Domingo, E., Schuster, P., Eds.; Springer: Cham, Switzerland, 2016; Volume 392, pp. 121–139. https://doi.org/10.1007/82_2015_471.
13. Yakushkina, T.; Saakian, D.B.; Hu, C.K. Exact dynamics for a mutator gene model. *Chin. J. Phys.* **2015**, *53*, 100904. <https://doi.org/10.6122/CJP.20150910A>.
14. Ancliff, M.; Park, J.M. Maximum, minimum, and optimal mutation rates in dynamic environments. *Phys. Rev. E* **2009**, *80*, 061910. <https://doi.org/10.1103/PhysRevE.80.061910>.
15. de Vladar, H.P.; Barton, N.H. The contribution of statistical physics to evolutionary biology. *Trends Ecol. Evol.* **2011**, *26*, 424–432. <https://doi.org/10.1016/j.tree.2011.04.002>.
16. Tran, T.D.; Hofrichter, J.; Jost, J. The free energy method and the Wright–Fisher model with 2 alleles. *Theory Biosci.* **2015**, *134*, 83–92. <https://doi.org/10.1007/s12064-015-0218-2>.
17. Smerlak, M. Effective potential reveals evolutionary trajectories in complex fitness landscapes. *bioRxiv* **2019**, 869883. <https://doi.org/10.48550/arXiv.1912.05890>.
18. Schnakenberg, J. Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Mod. Phys.* **1976**, *48*, 571–585. <https://doi.org/10.1103/RevModPhys.48.571>.
19. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **2012**, *75*, 126001. <https://doi.org/10.1088/0034-4885/75/12/126001>.

20. Sughiyama, Y.; Kobayashi, T.J. Steady-state thermodynamics for population growth in fluctuating environments. *Phys. Rev. E* **2017**, *95*, 012131. <https://doi.org/10.1103/PhysRevE.95.012131>.
21. Fang, X.; Kruse, K.; Lu, T.; Wang, J. Nonequilibrium physics in biology. *Rev. Mod. Phys.* **2019**, *91*, 045004. <https://doi.org/10.1103/RevModPhys.91.045004>.
22. Mustonen, V.; Lässig, M. Fitness flux and ubiquity of adaptive evolution. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4248–4253. <https://doi.org/10.1073/pnas.0907953107>.
23. Smerlak, M. Quasi-species evolution maximizes genotypic reproductive value (not fitness or flatness). *J. Theor. Biol.* **2021**, *522*, 110699. <https://doi.org/10.1016/j.jtbi.2021.110699>.
24. Poghosyan, R.; Saakian, D.B. Dynamics of entropy in evolution models. *Phys. A* **2023**, *617*, 128652. <https://doi.org/10.1016/j.physa.2023.128652>.
25. Esposito, M. Stochastic thermodynamics under coarse graining. *Phys. Rev. E* **2012**, *85*, 041125. <https://doi.org/10.1103/PhysRevE.85.041125>.
26. Busiello, D.M.; Hidalgo, J.; Maritan, A. Entropy production for coarse-grained dynamics. *New J. Phys.* **2019**, *21*, 073004. <https://doi.org/10.1088/1367-2630/ab29c0>.
27. Roldán, E.; Parrondo, J.M.R. Entropy production and Kullback–Leibler divergence between stationary trajectories of discrete systems. *Phys. Rev. E* **2012**, *85*, 031129. <https://doi.org/10.1103/PhysRevE.85.031129>.
28. Seifert, U. Universal bounds on entropy production from fluctuating coarse-grained trajectories. *Nat. Rev. Phys.* **2026**, *8*, 493–507. <https://doi.org/10.1038/s42254-026-00954-5>.
29. Wright, S. Evolution in Mendelian populations. *Genetics* **1931**, *16*, 97–159. <https://doi.org/10.1093/genetics/16.2.97>.
30. Ewens, W.J. *Mathematical Population Genetics 1: Theoretical Introduction*, 2nd ed.; Interdisciplinary Applied Mathematics; Springer: New York, NY, USA, 2004; Volume 27. <https://doi.org/10.1007/978-0-387-21822-9>.
31. Fisher, R.A. *The Genetical Theory of Natural Selection*; Clarendon Press: Oxford, UK, 1930. <https://doi.org/10.5962/bhl.title.27468>.
32. Wright, S. The differential equation of the distribution of gene frequencies. *Proc. Natl. Acad. Sci. USA* **1945**, *31*, 382–389. <https://doi.org/10.1073/pnas.31.12.382>.
33. Feller, W. *An Introduction to Probability Theory and Its Applications*, 2nd ed.; Wiley: New York, NY, USA, 1971; Volume 2.
34. Blinnikov, S.; Moessner, R. Expansions for nearly Gaussian distributions. *Astron. Astrophys. Suppl. Ser.* **1998**, *130*, 193–205. <https://doi.org/10.1051/aas:1998221>.
35. Watterson, G.A. Reversibility and the age of an allele. II. Two-allele models, with selection and mutation. *Theor. Popul. Biol.* **1977**, *12*, 179–196. [https://doi.org/10.1016/0040-5809\(77\)90041-7](https://doi.org/10.1016/0040-5809(77)90041-7).
36. Tavaré, S. Time Reversal and Age Distributions, I. Discrete-Time Markov Chains. *J. Appl. Probab.* **1980**, *17*, 33–46. <https://doi.org/10.2307/3212922>.
37. Taylor, C.; Iwasa, Y.; Nowak, M.A. A symmetry of fixation times in evolutionary dynamics. *J. Theor. Biol.* **2006**, *243*, 245–251. <https://doi.org/10.1016/j.jtbi.2006.06.016>.
38. Moran, P.A.P. Random processes in genetics. *Math. Proc. Camb. Philos. Soc.* **1958**, *54*, 60–71. <https://doi.org/10.1017/S0305004100033193>.
39. Lynch, M. Mutation pressure, drift, and the pace of molecular coevolution. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2306741120. <https://doi.org/10.1073/pnas.2306741120>.
40. Sella, G.; Hirsh, A.E. The application of statistical physics to evolutionary biology. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 9541–9546. <https://doi.org/10.1073/pnas.0501865102>. Correction in *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 14475.
41. Manhart, M.; Haldane, A.; Morozov, A.V. A universal scaling law determines time reversibility and steady state of substitutions under selection. *Theor. Popul. Biol.* **2012**, *82*, 66–76. <https://doi.org/10.1016/j.tpb.2012.03.007>.
42. Bar, A.; Mukamel, D. Mixed-Order Phase Transition in a One-Dimensional Model. *Phys. Rev. Lett.* **2014**, *112*, 015701. <https://doi.org/10.1103/PhysRevLett.112.015701>.
43. Mukamel, D., Mixed-Order Phase Transitions. In *50 Years of the Renormalization Group: Dedicated to the Memory of Michael E. Fisher*; Aharony, A., Entin-Wohlman, O., Huse, D., Radzihovsky, L., Eds.; World Scientific: Singapore, 2024; Chapter 10, pp. 89–101. https://doi.org/10.1142/9789811282386_0010.
44. Thouless, D.J. Long-Range Order in One-Dimensional Ising Systems. *Phys. Rev.* **1969**, *187*, 732–733. <https://doi.org/10.1103/PhysRev.187.732>.
45. Aizenman, M.; Chayes, J.T.; Chayes, L.; Newman, C.M. Discontinuity of the magnetization in one-dimensional $1/|x - y|^2$ Ising and Potts models. *J. Stat. Phys.* **1988**, *50*, 1–40. <https://doi.org/10.1007/BF01022985>.
46. Hohenberg, P.C.; Halperin, B.I. Theory of dynamic critical phenomena. *Rev. Mod. Phys.* **1977**, *49*, 435–479. <https://doi.org/10.1103/RevModPhys.49.435>.
47. Muirhead, C.A.; Wakeley, J. Modeling multiallelic selection using a Moran model. *Genetics* **2009**, *182*, 1141–1157. <https://doi.org/10.1534/genetics.108.089474>.

48. Huillet, T.E. On discrete-time multiallelic evolutionary dynamics driven by selection. *J. Probab. Stat.* **2010**, 2010, 580762. <https://doi.org/10.1155/2010/580762>.
49. Watterson, G.A. Reversibility and the age of an allele. I. Moran's infinitely many neutral alleles model. *Theor. Popul. Biol.* **1976**, 10, 239–253. [https://doi.org/10.1016/0040-5809\(76\)90018-6](https://doi.org/10.1016/0040-5809(76)90018-6).
50. Slatkin, M. Simulating genealogies of selected alleles in a population of variable size. *Genet. Res.* **2001**, 78, 49–57. <https://doi.org/10.1017/S0016672301005183>.
51. Barton, N.H.; Etheridge, A.M.; Sturm, A.K. Coalescence in a random background. *Ann. Appl. Probab.* **2004**, 14, 754–785. <https://doi.org/10.1214/105051604000000099>.
52. Huillet, T.E. On Wright–Fisher diffusion and its relatives. *J. Stat. Mech. Theory Exp.* **2007**, P11006. <https://doi.org/10.1088/1742-5468/2007/11/P11006>.
53. Jasinska, W.; Manhart, M.; Lerner, J.; Gauthier, L.; Serohijos, A.W.R.; Bershtein, S. Chromosomal barcoding of *E. coli* populations reveals lineage diversity dynamics at high resolution. *Nat. Ecol. Evol.* **2020**, 4, 437–452. <https://doi.org/10.1038/s41559-020-1103-z>.

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.