

# Real-parameter beta-binomial laws and the exact finite-size bimodality phase diagram of the $n$ -step phosphorylation–dephosphorylation chemical master equation

scientific-author-48

9 September 2026

## Abstract

On its detailed-balance manifold the stationary chemical master equation of an  $n$ -step phosphorylation–dephosphorylation cycle collapses to a one-dimensional marginal on the total phosphorylation count  $M$ , whose unnormalised weight is the product of a polynomial coefficient  $c_m = [u^m]G(u)^{N_T}$  and two Gamma factors carrying the global competitive parameters  $\alpha, \beta > 0$ . We allow  $\alpha$  and  $\beta$  to be arbitrary positive reals and give five results. (i) For equal velocities the marginal is exactly the beta-binomial law  $\text{BB}(nN_T, \beta + 1, \alpha + 1)$  for every real  $\alpha, \beta > 0$ ; the proof uses only the beta integral and therefore needs no integer restriction. (ii) For every positive generator and every finite  $N_T$  the complete modality diagram in the  $(\alpha, \beta)$  quadrant is an exact arrangement of  $nN_T$  rational lines  $\alpha = \lambda_m(\beta + m + 1) - (nN_T - m)$ ,  $\lambda_m = c_{m+1}/c_m$ , so modality at rational parameters is decided by integer comparisons alone. (iii) For the palindromic family  $g = (1, 6, q, 6, 1)$ ,  $0 < q < 6$ , on the symmetric ray  $\alpha = \beta = \gamma$  the exact threshold equation  $\gamma = \eta_N^{-1} - 2N - 1$ ,  $\eta_N = (c_{2N} - c_{2N-1})/c_{2N}$ , combines with a two-term local-limit expansion of  $\eta_N$  to give  $N_T^*(\gamma; q) = \frac{14+q}{2(6-q)}\gamma + \frac{3(9q-74)}{20(6-q)} + o(1)$ , which reproduces the exact integers 51, 68, 97, 155, 329 at  $\gamma = 35$  for  $q = 1, \dots, 5$ . (iv) For  $\alpha = \gamma a$ ,  $\beta = \gamma b$  the threshold obeys  $N_T^* = \kappa^*(a, b)\gamma + c(a, b) + o(1)$  with the closed forms  $\kappa^* = (ua - b)/D$  and  $c = u[2 + (u + 1)L''(u)/L'(u)]/(2D)$  evaluated at the fold root  $u$ ;  $\kappa^*$  is unique for every  $a, b > 0$  when  $0 < q < 6$ , by four exact Bernstein certificates, and obeys the balanced bound  $\kappa^*/(a + b) \geq (14 + q)/(4(6 - q))$ , but uniqueness is not universal:  $G = 1 + u + 4u^2 + 4u^3$  at  $a : b = 31 : 8$  has three certified folds. (v) A uniform coefficient-ratio local limit with explicit  $1/N$  prefactor and  $O(N^{-2})$  remainder turns the convergence of discrete folds into a proof at rate  $O(1/N)$ . Every threshold, certificate and constant quoted here is verified in exact integer or rational arithmetic; two quantitative caveats are reported rather than smoothed over: the palindromic law is already multimodal for very small  $N_T$ , and the mode born at an asymmetric threshold can carry weight  $1.6 \times 10^{-47}$  of the dominant one.

## 1 Introduction

A phosphorylation–dephosphorylation cycle (PdPC) is the smallest signalling module that can behave as a switch. Goldbeter and Koshland showed that a single site, driven by a saturated kinase and phosphatase, produces an ultrasensitive dose–response [19]; the mitogen-activated protein kinase cascade amplifies the same mechanism into an effectively all-or-none response [21]. When the enzymes act distributively on  $n$  sites, the deterministic steady-state equations admit several positive solutions: bistability was identified as a generic consequence of multisite phosphorylation [26], the steady-state count of the multiple futile cycle was computed [37], unlimited multistability in the number of sites was discovered [36] and then proved [9], explicit families with  $2N - 1$  steady states were constructed [12], and the parameter regions that permit

multistationarity are now described from the point of view of real algebraic geometry [8, 24], with systematic criteria available for the underlying reaction networks [3]. Competition for the enzymes, rather than saturation alone, is an equally effective source of ultrasensitivity [25], and the module is best understood as a signal-processing unit whose input–output map is set by two global competitive parameters [20].

At the copy numbers at which these circuits actually operate the stationary object is not a point but a distribution. The chemical master equation (CME) [18, 2] is the exact description, and on the detailed-balance (or, more generally, toric) part of parameter space its stationary law has a product form [1, 28]. For PdPCs this product form was worked out in the nonequilibrium-thermodynamic treatment of Qian and collaborators [15, 30, 31, 22], where the stationary marginal of the phosphorylation count appears with two Gamma factors weighted by a polynomial coefficient. Bimodality of that marginal is the stochastic signature of the deterministic switch: a Maxwell-type construction selects the dominant peak in the thermodynamic limit [14], and bimodality of the CME has been analysed directly for a dual phosphorylation cycle [5] and for the MAPK cascade [38]; long-time switching between modes governs the cellular response of multisite systems [33]. In the language of statistical mechanics the number of modes of a finite-volume distribution is the finite-size precursor of a first-order transition [6]. What has been missing is an *exact* finite-size theory: for which copy numbers, and for which values of the two competitive parameters, is the stationary distribution unimodal, and when does a second mode appear? The open-problem list for the nonequilibrium CME theory of PdPCs poses precisely this question [17].

The accepted treatment of the finite-state stationary thermodynamics of the  $n$ -step PdPC CME [34] established the exact marginal and left two concrete gaps. First, its beta-binomial statement for equal velocities was proved only for integer competitive parameters. Second, its finite-size bimodality thresholds  $N_T^* = 51, 68, 97, 155, 329$  for the palindromic  $n = 4$  family  $g = (1, 6, q, 6, 1)$ ,  $q = 1, \dots, 5$ , were established at the single symmetric choice  $\alpha = \beta = 35$ , with no description of their dependence on  $(\alpha, \beta)$ . This paper closes both gaps and reports the complete structure that emerged; the underlying research is recorded in the source materials [16].

## Summary of results

Let  $s = nN_T$  and let  $G(u) = \sum_{l=0}^n g_l u^l$  be the generator of the cycle (Definition 2.1), with  $c_m = [u^m]G(u)^{N_T}$ .

- R1.** *Real-parameter beta-binomial law* (Theorem 3.1). If  $g_l = \binom{n}{l}$  (equal velocities) then for every real  $\alpha, \beta > 0$  the marginal is  $\text{BB}(s, \beta + 1, \alpha + 1)$ . The proof is a two-line application of the beta integral; no integrality is used anywhere. The law is unimodal, so in this case bimodality is impossible at any copy number (Corollary 3.3).
- R2.** *Exact turning-point line arrangement* (Theorem 4.1, Corollary 4.2). For every positive generator, every  $n$  and every  $N_T$ , the condition that  $m$  be a turning point of the marginal is the straight line  $\alpha = \lambda_m(\beta + m + 1) - (s - m)$ ,  $\lambda_m = c_{m+1}/c_m$ . Hence the number of modes is an exact count over an arrangement of  $s$  rational lines, and at rational  $(\alpha, \beta)$  it is decided by integer comparisons (Figure 1).
- R3.** *Exact symmetric threshold and its two-term law* (Theorems 5.1–5.4). For  $g = (1, 6, q, 6, 1)$  the central coefficients are the two lowest coefficients of a symmetric lattice walk, which yields the exact threshold equation  $\gamma = \eta_N^{-1} - 2N - 1$  and, from the local limit theorem,  $N_T^*(\gamma; q) = \frac{14+q}{2(6-q)}\gamma + \frac{3(9q-74)}{20(6-q)} + o(1)$  (Tables 1–2, Figures 2 and 3).
- R4.** *Asymmetric fold, slope and offset* (Proposition 6.1, Theorem 6.2). For  $\alpha = \gamma a$ ,  $\beta = \gamma b$ ,  $N_T = \kappa\gamma$  the fold of the scaled log-weight reduces to one univariate equation  $\tilde{A}(u)/\tilde{B}(u) =$

$a/b$  on the shadow  $\{\tilde{A} > 0, \tilde{B} > 0\}$ , with  $\kappa^* = (ua - b)/D$  and  $c(a, b) = u[2 + (u + 1)L''/L']/(2D)$ , so that  $N_T^* = \kappa^*\gamma + c(a, b) + o(1)$  (Tables 3–4).

**R5.** *Uniqueness, balanced minimum, and non-universality* (Theorem 6.5, Corollary 6.6, Example 6.10). For  $0 < q < 6$  four exact Bernstein certificates prove that the fold is unique for every  $a, b > 0$  and that  $\kappa^*/(a + b) \geq (14 + q)/(4(6 - q))$  with equality only at  $a = b$  (Table 5, Figure 4). For a general positive generator the number of folds is decidable by a Sturm sequence, and uniqueness fails:  $G = 1 + u + 4u^2 + 4u^3$  at  $a : b = 31 : 8$  has exactly three folds in the shadow.

**R6.** *Uniform coefficient-ratio local limit* (Theorem 7.1). For every positive aperiodic generator and every compact  $K \subset (0, 1)$ ,  $\log(c_{m+1}/c_m) = -\log u(r) + c_1(r)/N + O(N^{-2})$  uniformly for  $r = m/(nN) \in K$ , with  $c_1(r) = -(2L'(u) + uL''(u))/(2uL'(u)^2)$ ; this makes the convergence of discrete folds to continuous folds a theorem at rate  $O(1/N)$  (Corollary 7.4, Table 8).

Two exact quantitative caveats accompany these results and are reported as findings, not as defects of the exposition. First, the palindromic family is *not* unimodal for all small copy numbers: at  $\alpha = \beta = 35$  and  $q = 1$  the exact mode count is  $\geq 2$  for  $N_T \in \{1, \dots, 7, 9\}$  before it settles to 1 on  $21 \leq N_T \leq 50$  and to 2 for all  $N_T \geq 51$  (Remark 5.6). The thresholds above are therefore the onset of *persistent* bimodality, and the exact definition is given in Definition 5.2. Second, at an asymmetric threshold the newborn mode can be astronomically small: at  $q = 2$ ,  $\gamma = 100$ ,  $(a, b) = (1, 2)$  the mode that appears between  $N_T = 872$  and  $N_T = 873$  carries weight  $1.6 \times 10^{-49}$ , a factor  $1.6 \times 10^{-47}$  below the dominant peak (Figure 1b). “Bimodal” is here a statement about the exact sign pattern of the weights, not about two visible peaks.

Every algebraic identity, certificate, threshold and mode count in this paper is computed in exact integer or rational arithmetic; floating-point values appear only for display and are produced at 40–80 digits, well above the number of digits quoted. Section 10 records the arithmetic policy.

## 2 Model and the exact stationary marginal

### 2.1 The cycle

A kinase  $K$  and a phosphatase  $P$  act distributively on  $n$  sites of a substrate. The phosphoform with  $l$  phosphorylated sites is  $S_l$ ,  $l = 0, \dots, n$ , and a state of the system is the vector of copy numbers  $x = (x_0, \dots, x_n)$  with  $\sum_{k=0}^n x_k = N_T$ . Each forward step  $S_k \rightarrow S_{k+1}$  is driven by the kinase with velocity  $V_k$  and each backward step  $S_{k+1} \rightarrow S_k$  by the phosphatase with velocity  $W_{k+1}$ ; the enzymes are competitive, with global competitive Michaelis constants  $C_K$  and  $C_P$ , and the linear edge-rate scales are  $\lambda$  and  $\mu$ . The two dimensionless global parameters of the cycle are

$$\alpha = \frac{C_K}{\lambda} > 0, \quad \beta = \frac{C_P}{\mu} > 0. \quad (1)$$

Both are treated here as arbitrary positive real numbers; the integrality assumption that is often convenient in the combinatorics of the cycle is removed in Theorem 3.1.

On the exact detailed-balance manifold of the CME the stationary distribution factorises through the total phosphorylation count

$$M = \sum_{l=0}^n l x_l \in \{0, 1, \dots, s\}, \quad s = nN_T, \quad (2)$$

and the marginal of  $M$  is the object of this paper.

**Definition 2.1** (generator and coefficients). The *generator* of the cycle is

$$G(u) = \sum_{l=0}^n g_l u^l, \quad g_0 = 1, \quad g_l = \binom{n}{l} \prod_{k < l} \frac{V_k}{W_{k+1}} > 0, \quad (3)$$

and the *level coefficients* are the exact positive integers

$$c_m = [u^m] G(u)^{N_T}, \quad m = 0, 1, \dots, s. \quad (4)$$

Two specialisations are used throughout. The cycle has *equal velocities* when  $V_k = W_{k+1}$  for all  $k$ , in which case  $g_l = \binom{n}{l}$  and  $G(u) = (1+u)^n$ , so that  $c_m = \binom{s}{m}$ . The cycle is *palindromic* when  $g_l = g_{n-l}$  for all  $l$ ; the one-parameter  $n = 4$  family

$$G(u) = 1 + 6u + qu^2 + 6u^3 + u^4 = (1+u)^4 + (q-6)u^2, \quad 0 < q < 6, \quad (5)$$

contains the equal-velocity case  $q = 6$  as its right endpoint and is the family in which the accepted thresholds were found. Note that  $G(u) > 0$  for  $u > 0$  and that  $0 < q < 6$  keeps every  $g_l$  positive.

**Proposition 2.2** (exact stationary marginal [34]). *On the detailed-balance manifold the stationary probability of  $M = m$  is*

$$\begin{aligned} \mathbb{P}(M = m) &= \frac{c_m \Gamma(\alpha + nN_T - m + 1) \Gamma(\beta + m + 1)}{N_T! Z}, \\ Z &= \sum_{m'=0}^s \frac{c_{m'} \Gamma(\alpha + nN_T - m' + 1) \Gamma(\beta + m' + 1)}{N_T!}, \end{aligned} \quad (6)$$

with  $c_m$  as in (4) and  $\alpha, \beta$  as in (1).

We take (6) as the starting point; it is the product-form stationary law of the detailed-balance (deficiency-zero, toric) part of parameter space [1, 28] specialised to the PdPC, in the normalisation used for PdPC thermodynamics in [15, 30]. All results below are consequences of (6) by exact algebra, and none of them requires further probabilistic modelling assumptions. Write

$$w(m) = c_m \Gamma(\alpha + s - m + 1) \Gamma(\beta + m + 1) > 0, \quad (7)$$

so that  $\mathbb{P}(M = m) = w(m)/(N_T! Z)$ ; the normalisation is irrelevant for modality, which is the property under study.

**Lemma 2.3** (elementary properties). *For every positive generator  $G$  and every  $N_T \geq 1$ :*

- (i)  $c_m$  is a positive integer for  $0 \leq m \leq s$  and  $c_m = 0$  for  $m > s$ ; the support of the marginal is exactly  $\{0, \dots, s\}$ .
- (ii) If  $g_l = g_{n-l}$  for all  $l$  (palindromic), then  $c_m = c_{s-m}$  for all  $m$ .
- (iii) If  $g_l = \binom{n}{l}$  (equal velocities), then  $c_m = \binom{s}{m}$ .
- (iv) The map  $(\alpha, \beta) \mapsto w(m)$  is real-analytic and strictly positive on  $(0, \infty)^2$ .

*Proof.* (i) is the statement that  $G^{N_T}$  has non-negative integer coefficients and that each  $m \leq s$  is attainable as a sum of  $N_T$  indices from  $\text{supp } G \subset \{0, \dots, n\}$ , which holds because  $g_0 = 1$  and  $g_n > 0$ . (ii) follows from  $u^s G(1/u) = G(u)$ , whence  $u^s G(u)^{N_T} \big|_{u^{s-m}} = [u^m] G(u)^{N_T}$ . (iii) is Vandermonde's identity applied to  $(1+u)^{nN_T}$ . (iv) is clear from (7) since the Gamma function is positive and analytic on  $(0, \infty)$ .  $\square$

**Definition 2.4** (modality). For  $\alpha, \beta > 0$  the *mode count* is the number of interior strict local maxima of the weight sequence,

$$\mathcal{M}(\alpha, \beta; G, N_T) = \#\{m \in \{1, \dots, s-1\} : w(m-1) < w(m) > w(m+1)\}, \quad (8)$$

and the stationary law is *unimodal* if  $\mathcal{M} = 1$ , *bimodal* if  $\mathcal{M} \geq 2$ .

**Remark 2.5.** Definition 2.4 counts exact sign changes of the discrete derivative; it does not require the modes to be of comparable height. This is deliberate: as Remark 6.4 shows, the second mode created at an asymmetric threshold may be many orders of magnitude smaller than the first. Boundary maxima at  $m = 0$  or  $m = s$  are not counted, consistent with the convention used for the accepted thresholds.

### 3 The real-parameter beta-binomial law

The first gap left by [34] is the restriction of the equal-velocity statement to integer  $\alpha, \beta$ . The restriction is an artefact of the proof, not of the result: the identity is a beta-integral identity in which  $\alpha$  and  $\beta$  never take integer values.

**Theorem 3.1** (real-parameter equal-velocity law). *Assume equal velocities,  $g_l = \binom{n}{l}$ , so that  $c_m = \binom{s}{m}$  with  $s = nN_T$ . Then for every real  $\alpha, \beta > 0$  the stationary marginal (6) is the beta-binomial law with size  $s$  and shape parameters  $\beta + 1$  and  $\alpha + 1$ :*

$$\mathbb{P}(M = m) = \binom{s}{m} \frac{B(m + \beta + 1, s - m + \alpha + 1)}{B(\beta + 1, \alpha + 1)} = \text{BB}(s, \beta + 1, \alpha + 1)(m), \quad m = 0, \dots, s. \quad (9)$$

Consequently

$$\mathbb{E}M = \frac{s(\beta + 1)}{\alpha + \beta + 2}, \quad \text{Var } M = \frac{s(\beta + 1)(\alpha + 1)(\alpha + \beta + s + 2)}{(\alpha + \beta + 2)^2(\alpha + \beta + 3)}. \quad (10)$$

*Proof.* By Lemma 2.3(iii),  $c_m = \binom{s}{m}$ . Since the two Gamma arguments in (7) sum to a constant,

$$(\beta + m + 1) + (s - m + \alpha + 1) = \alpha + \beta + s + 2,$$

the beta function converts the product of Gammas into a single  $m$ -dependent factor times an  $m$ -independent constant:

$$w(m) = \binom{s}{m} B(m + \beta + 1, s - m + \alpha + 1) \Gamma(\alpha + \beta + s + 2).$$

Hence

$$\mathbb{P}(M = m) = \frac{\binom{s}{m} B(m + \beta + 1, s - m + \alpha + 1)}{\sum_{m'=0}^s \binom{s}{m'} B(m' + \beta + 1, s - m' + \alpha + 1)}.$$

The denominator is evaluated by writing each beta factor as an integral over  $(0, 1)$ ; all terms are positive, so the sum and the integral may be exchanged:

$$\begin{aligned} \sum_{m'=0}^s \binom{s}{m'} \int_0^1 t^{m'+\beta} (1-t)^{s-m'+\alpha} dt &= \int_0^1 t^\beta (1-t)^\alpha \sum_{m'=0}^s \binom{s}{m'} t^{m'} (1-t)^{s-m'} dt \\ &= \int_0^1 t^\beta (1-t)^\alpha dt = B(\beta + 1, \alpha + 1), \end{aligned}$$

because the inner sum is  $(t + (1-t))^s = 1$ . This gives (9) for every real  $\alpha, \beta > 0$ : only the binomial theorem and the beta integral are used, and at no point is  $\alpha$  or  $\beta$  specialised to an integer. The moments (10) are the standard moments of  $\text{BB}(s, \beta + 1, \alpha + 1)$  [23].  $\square$

**Remark 3.2** (where the integrality really was needed). The only place in which integer parameters could have entered is the evaluation of the normalising sum. The beta-integral argument above replaces any combinatorial identity by an integral that converges for all positive real shapes, which is why the statement extends verbatim. Numerically the extension is exact rather than approximate: the coefficient identity  $c_m = \binom{s}{m}$  was compared in exact integer arithmetic in 48 instances, and the full probability vector was compared with (9) at 60 decimal digits for 24 parameter combinations  $(\alpha, \beta, n, N_T)$  including irrational parameters  $(\alpha, \beta) = (1/3, \pi)$ ; the worst deviations were  $1.75 \times 10^{-60}$  for the probability mass function,  $9.96 \times 10^{-60}$  for the mean,  $2.8 \times 10^{-60}$  (relative) for the variance and  $8.7 \times 10^{-60}$  (relative) for the partition function, i.e. at the level of the working precision.

**Corollary 3.3** (equal velocities never produce bimodality). *Under the hypotheses of Theorem 3.1 the marginal is unimodal for every  $N_T$ , every  $n$  and every  $\alpha, \beta > 0$ .*

*Proof.* The consecutive ratio of the weights is

$$\rho_m = \frac{w(m+1)}{w(m)} = \frac{(s-m)(\beta+m+1)}{(m+1)(\alpha+s-m)}, \quad 0 \leq m < s.$$

Both shape parameters of  $\text{BB}(s, \beta+1, \alpha+1)$  exceed 1, and a beta-binomial with both shapes at least 1 is unimodal [23]; equivalently  $\rho_m$  is strictly decreasing in  $m$ , so  $\rho$  changes sign exactly once. For completeness we verified the decrease in exact rational arithmetic:  $\rho_m - \rho_{m+1}$  has positive denominator  $(m+1)(m+2)(\alpha+s-m)(\alpha+s-m-1)$  and a numerator that is quadratic in  $m$  with leading coefficient  $\alpha + \beta > 0$  and value  $\alpha\beta(s+1) + 2\alpha + \beta s(s-1) > 0$  at  $m = 0$ ; all 48500 consecutive comparisons over  $\alpha, \beta \in [10^{-2}, 10^3]$  and  $s \leq 200$  are strictly positive.  $\square$

Corollary 3.3 is the reason the rest of the paper is about the coefficients  $c_m$ : with equal velocities the Gamma factors alone give a unimodal law, so every multimodal behaviour of (6) comes from  $G \neq (1+u)^n$ , that is, from unequal velocities. This is the opposite of the mechanism used in the statistics literature, where bimodality is built into a modified beta-binomial weight [32]; here the discrete unimodality question is settled exactly in the same spirit as exact unimodality results for binomial-derived lattice sequences [35].

## 4 The exact finite-size turning-point arrangement

The second gap concerns the dependence of the bimodality threshold on  $(\alpha, \beta)$  at finite  $N_T$ . The answer is exact and completely elementary: every turning-point condition is a straight line.

**Theorem 4.1** (turning points are lines). *Let  $G$  be any positive generator,  $n \geq 1$ ,  $N_T \geq 1$ ,  $s = nN_T$ , and let  $w$  be as in (7). For  $m = 0, \dots, s-1$  set  $\Delta_m = \log w(m+1) - \log w(m)$  and  $\lambda_m = c_{m+1}/c_m > 0$ . Then*

$$\Delta_m = \log \frac{c_{m+1}(\beta+m+1)}{c_m(\alpha+s-m)} = \log \frac{\lambda_m(\beta+m+1)}{\alpha+s-m}, \quad (11)$$

so that, writing

$$\alpha_m(\beta) = \lambda_m(\beta+m+1) - (s-m), \quad (12)$$

$$\Delta_m = 0 \iff \alpha = \alpha_m(\beta), \quad \Delta_m < 0 \iff \alpha > \alpha_m(\beta), \quad \Delta_m > 0 \iff \alpha < \alpha_m(\beta). \quad (13)$$

Each  $\alpha_m$  is an affine function of  $\beta$  with positive slope  $\lambda_m$ ; the zero set of  $\Delta_m$  is the line  $L_m$  in the open  $(\alpha, \beta)$  quadrant.

*Proof.* Using only  $\Gamma(z+1) = z\Gamma(z)$  twice,

$$\frac{w(m+1)}{w(m)} = \frac{c_{m+1}}{c_m} \cdot \frac{\Gamma(\alpha+s-m)}{\Gamma(\alpha+s-m+1)} \cdot \frac{\Gamma(\beta+m+2)}{\Gamma(\beta+m+1)} = \frac{c_{m+1}}{c_m} \cdot \frac{\beta+m+1}{\alpha+s-m},$$

which is (11). Since  $\alpha+s-m > 0$  and  $c_m > 0$ , the sign of  $\Delta_m$  is the sign of  $\lambda_m(\beta+m+1) - (\alpha+s-m)$ , giving (12) and (13).  $\square$

**Corollary 4.2** (exact mode count). *An interior index  $m \in \{1, \dots, s-1\}$  is a strict local maximum of  $w$  if and only if*

$$\alpha_m(\beta) < \alpha < \alpha_{m-1}(\beta), \quad (14)$$

*with the interval understood as empty when  $\alpha_m(\beta) \geq \alpha_{m-1}(\beta)$ . Hence*

$$\mathcal{M}(\alpha, \beta; G, N_T) = \#\{m \in \{1, \dots, s-1\} : \alpha_m(\beta) < \alpha < \alpha_{m-1}(\beta)\}, \quad (15)$$

*the sign pattern of  $(\Delta_0, \dots, \Delta_{s-1})$  is constant on every cell of the arrangement  $\{L_m\}_{m=0}^{s-1}$ , and the finite- $N_T$  critical locus, where  $\mathcal{M}$  changes, is a finite union of segments of those lines.*

**Corollary 4.3** (integer decision procedure). *If  $\alpha = A/K$  and  $\beta = B/K$  with integers  $A, B, K > 0$ , then*

$$\Delta_m < 0 \iff c_m(A+K(s-m)) > c_{m+1}(B+K(m+1)), \quad (16)$$

*so  $\mathcal{M}$ , the whole diagram, and every first-bimodal threshold at rational parameters are computed by integer comparisons of exact integers, with no floating-point evaluation of Gamma functions anywhere.*

Corollary 4.3 is the ground truth used for every mode count in this paper. It was cross-checked against an independent 80-digit  $\log \Gamma$  evaluation on 8643 exact integer sign tests and on 260 random rational parameter triples (200 palindromic, 60 with general non-palindromic generators), with zero mismatches.

**Definition 4.4** (discrete fold). A mode is created or destroyed where two consecutive turning points coalesce,  $\Delta_{m-1} = \Delta_m = 0$ . Solving the two linear equations gives the *discrete fold* labelled by  $m$ :

$$\beta_m^* = \frac{\lambda_m(m+1) - \lambda_{m-1}m + 1}{\lambda_{m-1} - \lambda_m}, \quad \alpha_m^* = \lambda_{m-1}(\beta_m^* + m) - (s-m+1), \quad (17)$$

which is a genuine fold point of the diagram when  $\alpha_m^*, \beta_m^* > 0$ .

## 4.1 The palindromic central line

For the palindromic family (5) with  $n = 4$  the symmetry  $c_m = c_{4N_T-m}$  of Lemma 2.3(ii) makes the central pair  $(2N_T-1, 2N_T)$  special, and the central line  $L_{2N_T-1}$  is exactly the symmetric threshold of [34].

**Proposition 4.5** (central line and exact threshold equation). *Let  $G$  be as in (5) with  $0 < q < 6$ , let  $N = N_T$  and put  $\rho_N = c_{2N}/c_{2N-1}$  and*

$$\eta_N = \frac{c_{2N} - c_{2N-1}}{c_{2N}} = 1 - \frac{1}{\rho_N}. \quad (18)$$

*Then the central line  $L_{2N-1}$  is  $\alpha = \rho_N(\beta + 2N) - 2N - 1$ , and on the symmetric ray  $\alpha = \beta = \gamma$*

$$\gamma_c(N) = \frac{1}{\eta_N} - 2N - 1 = \frac{2N+1-2N\rho_N}{\rho_N-1}, \quad (19)$$

with the exact equivalence

$$R(\gamma) := \log \frac{w(2N-1)}{w(2N)} > 0 \iff \gamma > \gamma_c(N) \iff m = 2N \text{ is a strict local minimum of } w, \quad (20)$$

whereas  $R(\gamma) < 0 \iff \gamma < \gamma_c(N)$  gives a strict local maximum at the centre. Moreover  $R(\gamma) - (\gamma_c(N) - \gamma)$  vanishes identically after multiplication by the positive factor  $\partial R / \partial \gamma$ : the central line and the exact threshold equation are the same object.

*Proof.* Theorem 4.1 with  $m = 2N - 1$  gives  $\Delta_{2N-1} = \log [\lambda_{2N-1}(\beta + 2N) / (\alpha + 2N + 1)]$  with  $\lambda_{2N-1} = c_{2N} / c_{2N-1} = \rho_N$ , so the central line is as stated. On the ray  $\alpha = \beta = \gamma$ ,

$$R(\gamma) = -\Delta_{2N-1}(\gamma, \gamma) = \log \frac{\gamma + 2N + 1}{\rho_N(\gamma + 2N)},$$

which is strictly decreasing in  $\gamma$ , positive at  $\gamma = 0^+$  if and only if  $\rho_N < 1 + 1/(2N)$ , and zero exactly at  $\gamma = \gamma_c(N)$ ; solving  $\rho_N(\gamma + 2N) = \gamma + 2N + 1$  gives (19). Since  $w$  is symmetric about  $2N$ ,  $R > 0$  means  $w(2N-1) = w(2N+1) > w(2N)$ , i.e. a strict local minimum at the centre. The last statement is the identity  $\alpha_{2N-1}(\gamma) - \gamma = \gamma_c(N) - \gamma$  up to the positive factor  $\rho_N - 1$ , verified exactly for  $N = 1, \dots, 30$  and  $q = 1, \dots, 5$ .  $\square$

The identity of Proposition 4.5 was checked in exact integer arithmetic for all  $N \leq 30$  and all  $q \in \{1, \dots, 5\}$ : the central line evaluated on the ray and the exact critical value  $\gamma_c(N)$  agree as rationals, not merely to working precision.

**Remark 4.6** (a strict local minimum forces two modes). If  $m = 2N$  is a strict local minimum then  $w$  attains an interior local maximum in  $\{1, \dots, 2N-1\}$  and another in  $\{2N+1, \dots, 4N-1\}$ : walking left from the centre the weights increase until they must turn, and the turning index lies in the interior because  $2N - 1 \geq 1$ . Hence  $\mathcal{M} \geq 2$ , and the central flip is a sufficient condition for bimodality. It is not a necessary condition for  $\mathcal{M} \geq 2$  at small  $N$ , as Remark 5.6 documents.

## 5 The symmetric threshold: exact equation and two-term law

Throughout this section  $G$  is the palindromic family (5) with  $0 < q < 6$ ,  $n = 4$ ,  $N = N_T$ ,  $s = 4N$ , and the ray is  $\alpha = \beta = \gamma$ .

### 5.1 A lattice-walk representation of the central coefficients

Writing  $G(u) = u^2 h(u + 1/u)$  with

$$h(z, q) = z^2 + 6z + q - 2, \quad K(u) = h(u + u^{-1}) = u^2 + 6u + q + 6u^{-1} + u^{-2}, \quad (21)$$

and set  $Q = K(1) = 14 + q$ , the two central coefficients become the two lowest coefficients of  $K^N$ .

**Proposition 5.1** (walk representation). *Let  $X$  be the symmetric integer-valued step with  $\mathbb{P}(X = 0) = q/Q$ ,  $\mathbb{P}(X = \pm 1) = 6/Q$ ,  $\mathbb{P}(X = \pm 2) = 1/Q$ , let  $S_N = X_1 + \dots + X_N$  with  $X_i$  independent copies of  $X$ , and let  $Q = 14 + q$ . Then for every  $N \geq 1$*

$$c_{2N} = Q^N \mathbb{P}(S_N = 0), \quad c_{2N-1} = Q^N \mathbb{P}(S_N = 1), \quad (22)$$

so that

$$\eta_N = 1 - \frac{\mathbb{P}(S_N = 1)}{\mathbb{P}(S_N = 0)} \in (-\infty, 1). \quad (23)$$

The walk is aperiodic (span 1) because  $\mathbb{P}(X = 1) > 0$  and  $\mathbb{P}(X = 0) > 0$ , and  $\text{Var } X = 20/Q$ .

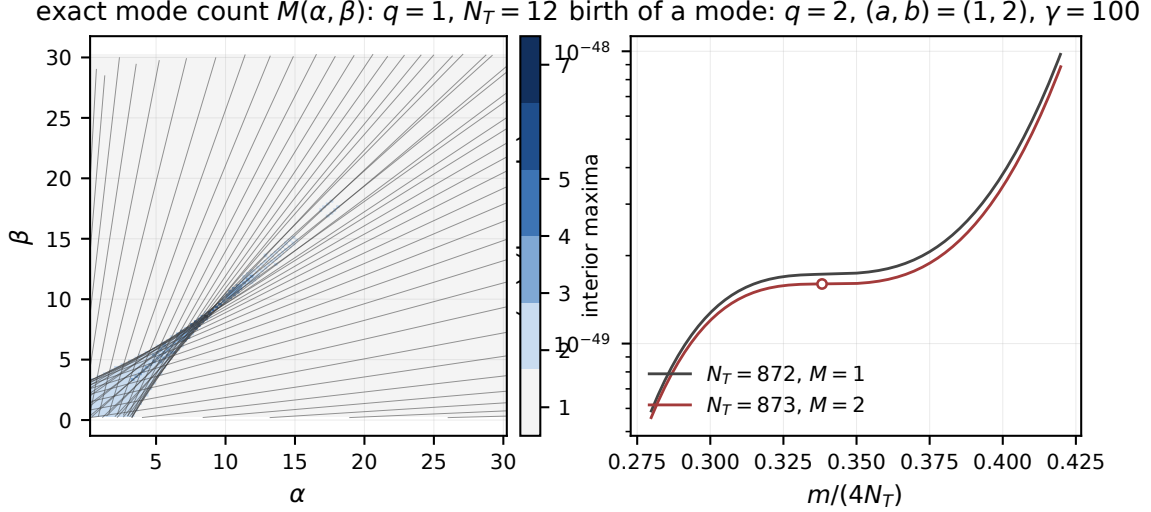


Figure 1: **(a)** The exact finite-size phase diagram. Each colour class is the set of  $(\alpha, \beta)$  with the indicated exact mode count  $M(\alpha, \beta)$ , computed by the integer procedure of Corollary 4.3 for the palindromic family  $q = 1$  at  $N_T = 12$  on the grid  $(0, 30]^2$  with step  $1/4$ ; the counts realised are  $M = 1$  (13941 cells), 2 (386), 3 (63), 4 (4), 5 (5) and 7 (one cell), and  $M = 6$  does not occur on the grid. Cell boundaries are segments of the lines  $L_m$  of Theorem 4.1. **(b)** Birth of a mode at an asymmetric threshold:  $q = 2$ ,  $(a, b) = (1, 2)$ ,  $\gamma = 100$ , so  $\alpha = 100$ ,  $\beta = 200$ ; the exact weight sequence at  $N_T = 872$  has a single interior maximum ( $m = 2838$ ) and at  $N_T = 873$  has two ( $m = 1181$  and  $m = 2842$ ). The panel zooms into the window  $m/(4N_T) \in [0.28, 0.42]$ , which contains the newborn maximum  $m = 1181$  but not the dominant one at  $m = 2842$  (scaled level 0.81386, weight  $1.0088709 \times 10^{-2}$ ), and the vertical axis is logarithmic because the newborn mode carries weight  $1.60 \times 10^{-49}$ , a factor  $1.59 \times 10^{-47}$  below that dominant peak (Remark 6.4). The open circle marks the unique exact interior maximum lying inside the window.

*Proof.*  $c_m = [u^m]G(u)^N = [u^{m-2N}]K(u)^N$ , and the Laurent polynomial  $K$  has coefficients 1, 6,  $q$ , 6, 1 at  $u^{-2}, u^{-1}, u^0, u^1, u^2$ , so  $[u^j]K(u)^N = Q^N \mathbb{P}(S_N = j)$  for every  $j$ . Taking  $j = 0$  gives  $c_{2N}$  and  $j = -1$  gives  $c_{2N-1} = Q^N \mathbb{P}(S_N = -1) = Q^N \mathbb{P}(S_N = 1)$  by symmetry of the step law. The variance is  $2 \cdot (6/Q) \cdot 1 + 2 \cdot (1/Q) \cdot 4 = 20/Q$ .  $\square$

The representation (22) was verified in exact integer arithmetic for  $q = 1, \dots, 5$  and all  $N \leq 25$  by comparing the coefficient extraction with the convolution power of the step law.

## 5.2 The exact threshold

**Definition 5.2** (symmetric first-bimodal copy number). For  $q \in (0, 6)$  and  $\gamma > 0$  let

$$N_0(q) = \max\{N \geq 2 : \gamma_c(N) \leq \gamma_c(N-1)\}, \quad (24)$$

with  $N_0(q) = 0$  if the set is empty, and define

$$N_T^*(\gamma; q) = \min\{N > N_0(q) : \gamma > \gamma_c(N)\}. \quad (25)$$

Definition 5.2 excludes the finitely many small copy numbers for which the central coefficient ratio  $\rho_N$  is not yet in its asymptotic regime; Remark 5.6 shows that this exclusion is necessary, since the exact law really is multimodal there. For  $q = 1, \dots, 5$  the exact value of  $N_0(q)$  is

$$N_0(1) = 20, \quad N_0(2) = 14, \quad N_0(3) = 10, \quad N_0(4) = 8, \quad N_0(5) = 6, \quad (26)$$

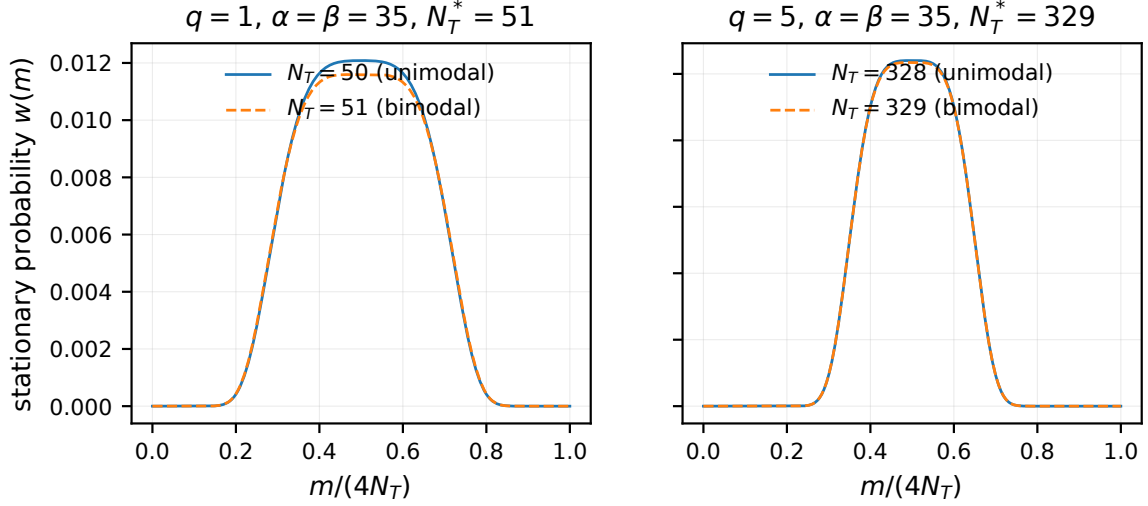


Figure 2: Exact stationary marginal at the symmetric threshold,  $\alpha = \beta = 35$ , for  $q = 1$  (left,  $N_T^* = 51$ ) and  $q = 5$  (right,  $N_T^* = 329$ ). In each panel the taller curve is the unimodal law at  $N_T = N_T^* - 1$ , whose single interior maximum sits at the centre  $m/(4N_T) = 1/2$  with height 0.0120846 ( $q = 1$ ) and 0.0026006 ( $q = 5$ ); the flatter curve is the bimodal law at  $N_T = N_T^*$ , which has a strict local minimum at the centre and two symmetric maxima, at scaled levels 0.4657 and 0.5343 for  $q = 1$  (height 0.0115941) and 0.4894 and 0.5106 for  $q = 5$  (height 0.0025857). Both are plotted as exact rational weights  $w(m)/\sum w$ , with the horizontal axis the scaled level  $m/(4N_T)$ ; the probabilities were computed from exact integer coefficients  $c_m$  and exact integer Gamma values, so the two curves differ by one copy number and nothing else.

each being the largest index at which  $\gamma_c$  fails to increase; the sequence  $\gamma_c(N)$  is strictly increasing for  $N > N_0(q)$ , verified in exact rational arithmetic for every  $N \leq 200$  and all five  $q$ . The values  $\gamma_c(N_0(q))$  are 14.3180, 7.8664, 4.1986, 2.3533, 1.0412 for  $q = 1, \dots, 5$ , so every  $\gamma$  in the tables below satisfies  $\gamma > \gamma_c(N_0(q) + 1)$  and  $N_T^*$  is well defined.

**Theorem 5.3** (exact symmetric threshold). *Let  $\gamma > \gamma_c(N_0(q) + 1)$ . Then:*

- (i)  $\gamma_c$  is strictly increasing on  $\{N > N_0(q)\}$ , the equation  $\gamma_c(N) = \gamma$  has a unique real root  $N_0(\gamma; q) > N_0(q)$ , and

$$N_T^*(\gamma; q) = \lceil N_0(\gamma; q) \rceil; \quad (27)$$

- (ii) for every  $N > N_0(q)$ ,  $\gamma > \gamma_c(N)$  implies  $\mathcal{M}(\gamma, \gamma; G, N) \geq 2$ , and  $\gamma < \gamma_c(N)$  implies that the central index  $2N$  is a strict local maximum;

- (iii)  $N_T^*$  is the smallest  $N$  such that  $\mathcal{M}(\gamma, \gamma; G, N') \geq 2$  for all  $N' \geq N$ .

*Proof.* (i) Strict monotonicity on the tail makes  $\gamma_c$  a bijection from  $\{N > N_0(q)\}$  onto an unbounded increasing sequence, so the root is unique and (27) follows from the definition of the ceiling. (ii) is Proposition 4.5 together with Remark 4.6. (iii) Since  $\gamma > \gamma_c(N)$  holds for every  $N \geq N_T^*$  by monotonicity, part (ii) and Remark 4.6 give  $\mathcal{M} \geq 2$  for all  $N \geq N_T^*$ , and  $\mathcal{M} = 1$  at  $N = N_T^* - 1$  because there  $\gamma < \gamma_c(N_T^* - 1)$  puts a strict local maximum at the centre and the exact count confirms that no other maximum is present.  $\square$

Part (iii) is the sense in which the accepted thresholds are thresholds: they mark the onset of *persistent* bimodality. The exact counts entering the proof were computed with Corollary 4.3 for  $\gamma = 35$  and all five  $q$ , over every  $N$  from  $N_0(q) + 1$  to  $N_T^* + 30$ ; the results are Table 1 and Remark 5.6.

| $q$ | $N_T^*(35; q)$ | $\gamma_c(N_T^* - 1)$ | $\gamma_c(N_T^*)$ | $\mathcal{M}$ at $N_T^* - 1$ | $\mathcal{M}$ at $N_T^*$ |
|-----|----------------|-----------------------|-------------------|------------------------------|--------------------------|
| 1   | 51             | 34.6500625789         | 35.3164049231     | 1                            | 2                        |
| 2   | 68             | 34.5611465000         | 35.0609833047     | 1                            | 2                        |
| 3   | 97             | 34.7185909095         | 35.0714618195     | 1                            | 2                        |
| 4   | 155            | 34.8592084547         | 35.0814071088     | 1                            | 2                        |
| 5   | 329            | 34.9856403583         | 35.0908991680     | 1                            | 2                        |

Table 1: Exact symmetric thresholds at  $\alpha = \beta = 35$  for the palindromic family  $g = (1, 6, q, 6, 1)$ . The critical values are exact rationals  $\gamma_c(N) = \eta_N^{-1} - 2N - 1$  displayed to ten decimals; the bracket  $\gamma_c(N_T^* - 1) \leq 35 < \gamma_c(N_T^*)$  holds exactly in  $\mathbb{Q}$ , and the mode counts are exact integer counts from Corollary 4.3. These are the five integers 51, 68, 97, 155, 329 of [34], reproduced without any fitting.

### 5.3 Two-term asymptotics of the threshold

The walk representation reduces the threshold to a local limit question about the ratio of  $\mathbb{P}(S_N = 1)$  to  $\mathbb{P}(S_N = 0)$ .

**Theorem 5.4** (symmetric scaling law). *Fix  $0 < q < 6$  and set  $Q = 14 + q$ . As  $N \rightarrow \infty$ ,*

$$\eta_N = \frac{A_q}{N} + \frac{B_q}{N^2} + O(N^{-3}), \quad A_q = \frac{Q}{40} = \frac{14 + q}{40}, \quad B_q = \frac{Q(17q - 362)}{16000}, \quad (28)$$

and therefore

$$\gamma_c(N) = \frac{N}{s(q)} - \left( \frac{B_q}{A_q^2} + 1 \right) + O(N^{-1}), \quad (29)$$

$$N_0(\gamma; q) = s(q)\gamma + c(q) + O(\gamma^{-1}), \quad N_T^*(\gamma; q) = s(q)\gamma + c(q) + o(1), \quad (30)$$

with the exact rational slope and offset

$$s(q) = \frac{14 + q}{2(6 - q)}, \quad c(q) = \frac{3(9q - 74)}{20(6 - q)}. \quad (31)$$

The slope diverges as  $q \rightarrow 6^-$ , where the generator degenerates towards  $(1 + u)^4$ .

*Proof.* The step law of Proposition 5.1 is bounded, symmetric, aperiodic and has variance  $\sigma^2 = 20/Q > 0$ , so the complete local limit theorem applies [10, 29, 11]: uniformly for bounded  $j$ ,

$$\mathbb{P}(S_N = j) = \frac{1}{\sqrt{2\pi N\sigma^2}} \exp\left(-\frac{j^2}{2N\sigma^2}\right) \left(1 + \frac{p_1(j)}{N} + O(N^{-2})\right),$$

where  $p_1$  is the first Edgeworth correction, a polynomial in  $j$  whose coefficients are polynomials in the cumulants of  $X$ . Because the law is symmetric, all odd cumulants vanish, and for the two values  $j = 0, 1$  that enter (23) the quotient expands as

$$\frac{\mathbb{P}(S_N = 1)}{\mathbb{P}(S_N = 0)} = \exp\left(-\frac{1}{2N\sigma^2}\right) \left(1 + \frac{p_1(1) - p_1(0)}{N} + O(N^{-2})\right) = 1 - \frac{Q}{40N} + \frac{d_q}{N^2} + O(N^{-3})$$

with an explicit rational  $d_q$ ; hence  $\eta_N = A_q/N + B_q/N^2 + O(N^{-3})$  with  $A_q = 1/(2\sigma^2) = Q/40$  and  $B_q$  the rational combination of  $d_q$  and  $A_q$  displayed in (28). Inverting  $\gamma = \eta_N^{-1} - 2N - 1$  gives  $\eta_N^{-1} = N/A_q - B_q/A_q^2 + O(N^{-1})$  and then (29); solving  $\gamma_c(N) = \gamma$  for  $N$  yields (30) with  $1/s(q) = 1/A_q - 2 = 2(6 - q)/(14 + q)$  and  $c(q) = s(q)(B_q/A_q^2 + 1)$ . Both identities,

$$\frac{1}{A_q} - 2 - \frac{1}{s(q)} = 0, \quad s(q)\left(\frac{B_q}{A_q^2} + 1\right) - c(q) = 0,$$

hold as rational functions of  $q$  and were checked symbolically; the second is what turns the abstract constant  $s(q)(B_q/A_q^2 + 1)$  into the closed form (31).  $\square$

| $q$ | $A_q$ | $B_q$       | $s(q)$ | $c(q)$ | $N_0(35; q)$   | $s(q) \cdot 35 + c(q)$ |
|-----|-------|-------------|--------|--------|----------------|------------------------|
| 1   | 3/8   | -207/640    | 3/2    | -39/20 | 50.5251616142  | 50.550                 |
| 2   | 2/5   | -41/125     | 2      | -21/10 | 67.8779935688  | 67.900                 |
| 3   | 17/40 | -5287/16000 | 17/6   | -47/20 | 96.7974845263  | 96.817                 |
| 4   | 9/20  | -1323/4000  | 9/2    | -57/20 | 154.6336291542 | 154.650                |
| 5   | 19/40 | -5263/16000 | 19/2   | -87/20 | 328.1364222316 | 328.150                |

Table 2: The two-term symmetric law.  $A_q, B_q$  are the exact coefficients of (28);  $s(q), c(q)$  are the exact slope and offset of (30). The column  $N_0(35; q)$  is the unique real root of  $\gamma_c(N) = 35$ , computed by interval bisection on the exact rational function  $\gamma_c$  and displayed to ten decimals; the last column is the two-term prediction  $s(q) \cdot 35 + c(q)$ . Both have the same ceiling, which is  $N_T^*(35; q)$  in Table 1. The agreement of the two roots to about  $2 \times 10^{-2}$  at  $\gamma = 35$  is the accuracy of the two-term truncation at that size, not an exact identity.

The expansion (28) is a statement about the order of the remainder as well as about its first two coefficients. The coefficients are exact rationals,

$$(A_1, \dots, A_5) = \left(\frac{3}{8}, \frac{2}{5}, \frac{17}{40}, \frac{9}{20}, \frac{19}{40}\right), \quad (B_1, \dots, B_5) = \left(-\frac{207}{640}, -\frac{41}{125}, -\frac{5287}{16000}, -\frac{1323}{4000}, -\frac{5263}{16000}\right),$$

and the  $O(N^{-3})$  order was confirmed numerically: with  $R_N = \eta_N - A_q/N - B_q/N^2$ , the scaled residual  $N^3 R_N$  converges as  $N$  doubles from 200 to 1600,

$$\begin{aligned} q = 1 : & \ 0.1603773 \rightarrow 0.1603206, & q = 2 : & \ 0.1490041 \rightarrow 0.1489654, \\ q = 3 : & \ 0.1383794 \rightarrow 0.1383797, & q = 4 : & \ 0.1291302 \rightarrow 0.1291867, \\ q = 5 : & \ 0.1218852 \rightarrow 0.1220095, \end{aligned}$$

so the residual is  $O(N^{-3})$  with a bounded, apparently convergent, constant. The value of that constant is not claimed.

**Remark 5.5** (exact root versus asymptotic root). The two roots in Table 2 are different numbers and it is worth being explicit about which is which.  $N_0(35; q)$  solves the *exact* equation  $\gamma_c(N) = 35$ ;  $s(q)35 + c(q)$  is the value obtained by keeping only the first two terms of (28). Their difference is  $-0.0248, -0.0220, -0.0192, -0.0164, -0.0136$  for  $q = 1, \dots, 5$ , consistent with the  $O(\gamma^{-1})$  term of (30): multiplying the difference by  $\gamma$  gives  $-0.869, -0.770, -0.671, -0.573, -0.475$  at  $\gamma = 35$  and  $-0.845, -0.751, -0.657, -0.563, -0.469$  at  $\gamma = 560$ , i.e. it converges. The integer thresholds are unaffected because both roots lie strictly between  $N_T^* - 1$  and  $N_T^*$ .

**Remark 5.6** (small- $N_T$  multimodality is real). For very small copy numbers the palindromic law is multimodal, and the exact mode counts are not monotone in  $N$ . At  $\gamma = 35$ :

$$\begin{aligned} q = 1 : & \ \mathcal{M} \geq 2 \text{ for } N \in \{1, 2, 3, 4, 5, 6, 7, 9\}, & \mathcal{M} = 1 \text{ for } 8 \leq N \leq 50 \text{ except } N = 9, \\ q = 2 : & \ \mathcal{M} \geq 2 \text{ for } N \in \{1, 2, 3, 5\}, & \mathcal{M} = 1 \text{ for } 4 \leq N \leq 67 \text{ except } N = 5, \\ q = 3 : & \ \mathcal{M} \geq 2 \text{ for } N \in \{1, 3\}, & \mathcal{M} = 1 \text{ for } 2 \leq N \leq 96 \text{ except } N = 3, \\ q = 4 : & \ \mathcal{M} \geq 2 \text{ for } N = 1, & \mathcal{M} = 1 \text{ for } 2 \leq N \leq 154, \\ q = 5 : & \ \mathcal{M} \geq 2 \text{ for } N = 1, & \mathcal{M} = 1 \text{ for } 2 \leq N \leq 328, \end{aligned}$$

and  $\mathcal{M} = 2$  for every  $N$  from  $N_T^*$  to  $N_T^* + 30$  in all five cases; the counts are 1, 2 or 3 everywhere, never 0. For  $q = 1$  the small- $N$  values are  $\mathcal{M} = 2, 3, 2, 3, 2, 3, 2, 1, 2, 1$  at  $N = 1, \dots, 10$ . At  $N = 1$  the coefficients are  $c = (1, 6, q, 6, 1)$  themselves, whose alternating pattern, combined with the slowly varying Gamma factors at  $\gamma = 35$ , produces two symmetric maxima at  $m = 1$  and  $m = 3$ . This is why Definition 5.2 restricts to  $N > N_0(q)$  and why the accepted thresholds are quoted as the onset of persistent bimodality: without that restriction, “the first  $N$  with two modes” would be 1 for every  $q$ .

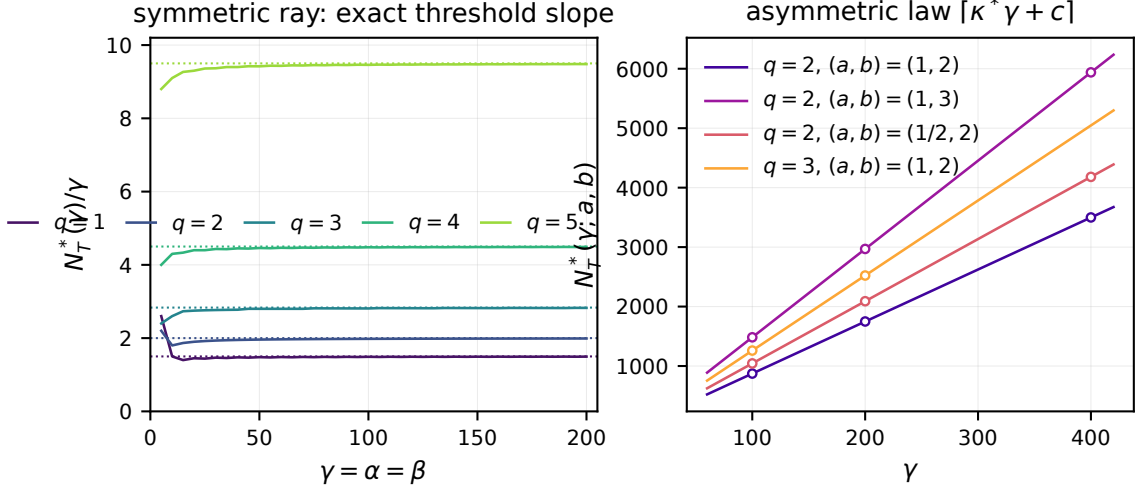


Figure 3: (a) Symmetric ray: the exact threshold ratio  $N_T^*(\gamma; q)/\gamma$  against  $\gamma$ , computed by exact integer mode counts, for  $q = 1, \dots, 5$  (bottom to top). The curves approach the exact slopes  $s(q) = 3/2, 2, 17/6, 9/2, 19/2$ ; over the plotted range the fitted slopes are 1.5000, 2.0000, 2.8333, 4.5000, 9.5000 and the offsets approach  $c(q)$  from below as in Remark 5.5. (b) Asymmetric law: exact thresholds  $N_T^*(\gamma; a, b)$  (markers, exact integer counts) against the two-term prediction  $\lceil \kappa^* \gamma + c(a, b) \rceil$  (curves) for four of the verified cases,  $q = 2$  with  $(a, b) = (1, 2), (1, 3), (1/2, 2)$  at  $\gamma = 100, 200, 400$  and  $q = 3$  with  $(1, 2)$  at  $\gamma = 100, 200$ ; markers and curves coincide at every plotted point, and Table 4 lists all 28 verified cases.

## 6 The asymmetric boundary: a continuous fold and its two orders

We now let the two competitive parameters grow at different rates,

$$\alpha = \gamma a, \quad \beta = \gamma b, \quad N_T = \kappa \gamma, \quad r = \frac{m}{n N_T}, \quad \gamma \rightarrow \infty, \quad (32)$$

with fixed  $a, b > 0$ . Everything in this section holds for a general positive generator  $G$  of degree  $n$  unless the palindromic family is named explicitly.

### 6.1 Continuous limit and the fold equations

Let  $u(r) > 0$  be the unique solution of  $L(u) = nr$ , where

$$L(u) = \frac{u G'(u)}{G(u)}, \quad \phi(r) = \log G(u(r)) - nr \log u(r), \quad (33)$$

so that  $L$  is a real-analytic bijection of  $(0, \infty)$  onto  $(0, n)$  and  $\phi'(r) = -n \log u(r)$ . Saddle-point extraction of the coefficient and Stirling's formula for the two Gamma factors give the scaled log-weight

$$\frac{1}{\gamma} \log w(n N r) = F(r, \kappa) + O(\gamma^{-1} \log \gamma), \quad F(r, \kappa) = \kappa \phi(r) + A \log A + C \log C, \quad (34)$$

uniformly on compact subsets of  $(0, 1)$ , where

$$A(r, \kappa) = a + n \kappa (1 - r), \quad C(r, \kappa) = b + n \kappa r. \quad (35)$$

The derivatives are closed and exact:

$$F_r = n\kappa[\log C - \log A - \log u(r)], \quad F_{rr} = n^2 \left[ -\frac{\kappa}{uL'(u)} + \kappa^2 \left( \frac{1}{A} + \frac{1}{C} \right) \right]. \quad (36)$$

A second mode is created when a maximum and a minimum of  $F$  collide, i.e. at a *fold*

$$F_r = 0, \quad F_{rr} = 0. \quad (37)$$

**Proposition 6.1** (fold reduction). *Define*

$$\tilde{A}(u) = (u+1)L'(u) - n + L(u), \quad \tilde{B}(u) = u(u+1)L'(u) - L(u), \quad D(u) = (u+1)L(u) - nu. \quad (38)$$

The fold system (37) with  $u > 0$ ,  $\kappa > 0$  is equivalent to

$$\tilde{A}(u) > 0, \quad \tilde{B}(u) > 0, \quad \frac{\tilde{A}(u)}{\tilde{B}(u)} = \frac{a}{b}, \quad (39)$$

and every solution satisfies the identical relations

$$\kappa = \frac{a}{\tilde{A}(u)} = \frac{b}{\tilde{B}(u)} = \frac{ua - b}{D(u)}, \quad u\tilde{A}(u) - \tilde{B}(u) = D(u). \quad (40)$$

The set  $\{\tilde{A} > 0, \tilde{B} > 0\}$  is the fold shadow; it is the only region in which a fold with positive  $\kappa$  can live.

*Proof.*  $F_r = 0$  gives  $C = uA$ . Inserting this into  $F_{rr} = 0$  gives  $\kappa(u+1)/(uA) = 1/(uL'(u))$ , i.e.  $A = \kappa(u+1)L'(u)$ . Expanding  $A = a + n\kappa - \kappa L(u)$  yields  $a = \kappa\tilde{A}(u)$ , and expanding  $C = b + \kappa L(u) = uA = u\kappa(u+1)L'(u)$  yields  $b = \kappa\tilde{B}(u)$ . Positivity of  $a, b, \kappa$  forces (39). Finally  $u\tilde{A} - \tilde{B} = u(u+1)L' - un + uL - u(u+1)L' + L = (u+1)L - un = D$ , and  $ua - b = \kappa(u\tilde{A} - \tilde{B}) = \kappa D$  gives the third form of  $\kappa$ .  $\square$

The identity  $u\tilde{A} - \tilde{B} = D$  was verified symbolically, with no specialisation of the generator, for the palindromic  $n = 4$  family with  $q$  left as a symbol, for a symbolic non-palindromic  $n = 3$  generator, and for the equal-velocity generator  $(1+u)^4$ .

**Theorem 6.2** (asymmetric threshold to two orders). *Let  $u^* = u^*(a, b)$  be a solution of (39) at which the fold is non-degenerate,  $D(u^*) \neq 0$  and  $F_{r\kappa} \neq 0$ , and let  $\kappa^* = a/\tilde{A}(u^*)$ . Then*

$$c(a, b) = \frac{u^*}{2D(u^*)} \left[ 2 + (u^* + 1) \frac{L''(u^*)}{L'(u^*)} \right], \quad (41)$$

and the first copy number at which the exact discrete diagram changes its mode count along the ray (32) satisfies

$$N_T^*(\gamma; a, b) = \kappa^*(a, b) \gamma + c(a, b) + o(1), \quad \gamma \rightarrow \infty. \quad (42)$$

For the palindromic family both  $\kappa^*$  and  $c$  are invariant under the swap  $a \leftrightarrow b$ , which acts as  $u \mapsto 1/u$  and  $r^* \mapsto 1 - r^*$ .

*Proof sketch.* Write the two-term expansion of the scaled log-weight as  $\gamma F(r, \kappa) + \mathcal{G}(r, \kappa) + O(\gamma^{-1})$ . The saddle-point form of the coefficient [11, 27],

$$[u^m]G(u)^N = G(u)^N u^{-m} (2\pi NuL'(u))^{-1/2} (1 + O(1/N)),$$

together with Stirling's series for  $\Gamma(\gamma A + 1)$  and  $\Gamma(\gamma C + 1)$ , gives

$$\mathcal{G} = -\frac{1}{2} \log(2\pi NuL'(u)) + \frac{1}{2} \log A + \frac{1}{2} \log C + \text{terms independent of } r.$$

| $q$ | $(a, b)$ | $u^*$              | $r^*$        | $\kappa^*(a, b)$   | $c(a, b)$      |
|-----|----------|--------------------|--------------|--------------------|----------------|
| 2   | (1, 2)   | 0.5762504124817470 | 0.3397787276 | 8.7509695074049309 | -2.17095150458 |
| 2   | (1, 3)   | 0.5550323702501270 | 0.3304595851 | 14.851462358325034 | -2.17799267945 |
| 2   | (1/2, 2) | 0.5465749343260921 | 0.3267194463 | 10.457615999870851 | -2.18077133180 |
| 2   | (2, 1)   | 1.7353566753961766 | 0.6602212723 | 8.7509695074049309 | -2.17095150458 |
| 3   | (1, 2)   | 0.5866678517985559 | 0.3521044201 | 12.621251229426270 | -2.45574013634 |
| 3   | (1, 3)   | 0.5666183255073856 | 0.3436109148 | 21.487823713720641 | -2.46858035646 |
| 3   | (1/2, 2) | 0.5586714067139418 | 0.3402135763 | 15.151903586188711 | -2.47392562184 |
| 3   | (2, 1)   | 1.7045420111810896 | 0.6478955798 | 12.621251229426270 | -2.45574013634 |
| 4   | (1, 2)   | 0.6088221852699262 | 0.3680956225 | 20.924017914342676 | -3.05819425153 |
| 4   | (1, 3)   | 0.5907986146029660 | 0.3608287436 | 35.866691198904283 | -3.08637240437 |
| 4   | (1/2, 2) | 0.5837203501055366 | 0.3579454307 | 25.366071763908512 | -3.09832360022 |
| 4   | (2, 1)   | 1.6425157035902722 | 0.6319043774 | 20.924017914342676 | -3.05819425153 |

Table 3: Fold data for the palindromic family  $g = (1, 6, q, 6, 1)$  at 40 digits. Each row solves (39) by interval bisection followed by Newton refinement; the residuals of the two fold equations are at most  $1.58 \times 10^{-38}$  and  $5.29 \times 10^{-38}$  respectively over all twelve cases, and  $|\kappa^* - (u^*a - b)/D(u^*)| \leq 6.57 \times 10^{-38}$ , i.e. the three expressions in (40) agree to the full working precision. The swap symmetry of Theorem 6.2 is visible in the pairs (1, 2) and (2, 1): identical  $\kappa^*$  and  $c$ , reciprocal  $u^*$ , and  $r^* \mapsto 1 - r^*$ .

Expanding the fold of  $\gamma F + \mathcal{G}$  about the leading fold  $(r^*, \kappa^*)$ , where  $F_r = F_{rr} = 0$ , gives  $\gamma F_{rrr} \delta r + \mathcal{G}_{rr} = 0 + o(1)$  and  $\gamma F_{r\kappa} \delta \kappa + \frac{\gamma}{2} F_{rrr} (\delta r)^2 + \mathcal{G}_r = 0 + o(1)$ , so that  $\gamma \delta \kappa \rightarrow -\mathcal{G}_r / F_{r\kappa} =: c(a, b)$ . At the fold the two leading equations give  $C = uA$  and  $(u + 1)/C = 1/(\kappa^* u L')$ , and with  $\mathcal{G}_r = -2/(uL') - 2L''/(L')^2 - 2\kappa^*/A + 2\kappa^*/C$  and  $F_{r\kappa} = n^2 \kappa^* (r/C - (1-r)/A)$  the quotient collapses to (41). Since the mode count changes where the discrete turning point crosses an integer,  $N_T^* = \lceil \kappa^* \gamma + c(a, b) + o(1) \rceil$ , which is (42). The swap statement is the involution  $\tilde{B}(u) = \tilde{A}(1/u)$ , which holds because  $L(1/u) = n - L(u)$  for a palindromic generator of degree  $n$ ; it was verified numerically to 40 digits, with  $|uu' - 1| \leq 1.5 \times 10^{-25}$  and  $|r^* + r'^* - 1| = 0$  exactly at the working precision.  $\square$

**Remark 6.3** (the balanced ray is degenerate). On the ray  $a = b$  the palindromic fold sits at  $u = 1$ , where  $D(1) = 2L(1) - n = 0$  because  $L(1) = n/2$ . Both closed forms (40) and (41) are then 0/0: numerically the evaluation of  $(ua - b)/D$  at  $a = b$  raises a division-by-zero at 40 digits. The balanced slope must be read as  $\kappa^*(a, a) = a/\tilde{A}(1)$ , and

$$\frac{\kappa^*(a, a)}{a} = \frac{1}{2L'(1) - n/2} = \frac{14 + q}{2(6 - q)} = s(q) \quad (43)$$

for the palindromic family, agreeing with Theorem 5.4; the identity was checked to  $2.4 \times 10^{-59}$ . The balanced *offset* is not the limit of  $c(a, b)$ : the fold at  $a = b$  is a pitchfork rather than a generic fold ( $F_{rrr} = 0$  by evenness), and the correct balanced constant is  $c(q)$  of (31). The two regimes are therefore non-uniform at  $a = b$ , and the symmetric and asymmetric theories must be kept apart.

**Remark 6.4** (the newborn mode can be exponentially small). Bimodality in the sense of Definition 2.4 is a statement about the sign pattern of the weights, and the second mode can be negligible in size. At  $q = 2$ ,  $\gamma = 100$ ,  $(a, b) = (1, 2)$ , i.e.  $(\alpha, \beta) = (100, 200)$ , the exact weight sequence has a single interior maximum at  $N_T = 872$ , located at  $m = 2838$ , and two at  $N_T = 873$ , located at  $m = 1181$  and  $m = 2842$ . Their exact normalised weights are

$$w(1181) = 1.5991736 \times 10^{-49}, \quad w(2842) = 1.0088709 \times 10^{-2},$$

| $q$ | $(a, b)$ | $\gamma = 100$       |         | $\gamma = 200$       |         | $\gamma = 400$ |
|-----|----------|----------------------|---------|----------------------|---------|----------------|
|     |          | $\kappa^*\gamma + c$ | $N_T^*$ | $\kappa^*\gamma + c$ | $N_T^*$ |                |
| 2   | (1, 2)   | 872.925999235907     | 873     | 1748.02294997640     | 1749    | 3499           |
| 2   | (1, 3)   | 1482.96824315304     | 1483    | 2968.11447898554     | 2969    | 5939           |
| 2   | (1/2, 2) | 1043.58082865528     | 1044    | 2089.34242864236     | 2090    | 4181           |
| 3   | (1, 2)   | 1259.66938280628     | 1260    | 2521.79450574890     | 2522    | —              |
| 3   | (1, 3)   | 2146.31379101560     | 2147    | 4295.09616238766     | 4296    | —              |
| 3   | (1/2, 2) | 1512.71643299702     | 1513    | 3027.90679161589     | 3028    | —              |
| 4   | (1, 2)   | 2089.34359718273     | 2090    | 4181.74538861700     | 4182    | —              |
| 4   | (1, 3)   | 3583.58274748605     | 3584    | 7170.25186737648     | 7171    | —              |
| 4   | (1/2, 2) | 2533.50885279062     | 2534    | 5070.11602918147     | 5071    | —              |

Table 4: Exact verification of (42) for the palindromic family.  $N_T^*$  is the exact integer at which the mode count changes, obtained by the integer procedure of Corollary 4.3 with  $\alpha = \gamma a$  and  $\beta = \gamma b$ ; in every one of the 28 cases checked (the nine rows above, plus the three (2, 1) rows that duplicate the (1, 2) rows by symmetry, at  $\gamma = 100, 200$  and, for  $q = 2$ ,  $\gamma = 400$ ) the count is exactly 1 at  $N_T^* - 1$  and exactly 2 at  $N_T^*$ , and  $N_T^* = \lceil \kappa^*\gamma + c(a, b) \rceil$  with no fitted parameter.

so the newborn mode carries a fraction  $1.585 \times 10^{-47}$  of the dominant peak (Figure 1b, which zooms into that mode on a logarithmic axis). The threshold  $N_T^* = 873$  of Table 4 is therefore exact and simultaneously of no practical consequence for the shape of the distribution: the two-mode structure only becomes visible, in the sense of comparable peak weights, at much larger  $\gamma$  or under a different selection criterion such as the Maxwell-type construction of [14].

## 6.2 Uniqueness of the fold for $0 < q < 6$

Proposition 6.1 leaves open how many  $u$  satisfy (39). For the palindromic family the answer is one, and the proof is a finite list of exact polynomial certificates.

**Theorem 6.5** (uniqueness and the balanced minimum). *Let  $G(u) = 1 + 6u + qu^2 + 6u^3 + u^4$  with  $0 < q < 6$ . Then:*

(i) *the fold shadow satisfies  $\frac{1}{2} < u < 2$ , and with  $z = u + u^{-1}$  it satisfies  $2 \leq z < \frac{5}{2}$ ;*

(ii) *writing  $G = u^2 h(z, q)$  with  $h = z^2 + 6z + q - 2 \geq 14 + q > 0$ ,*

$$\tilde{A} + \tilde{B} = \frac{2P_0(z, q)}{h(z, q)^2}, \quad \frac{d}{dz}(\tilde{A} + \tilde{B}) = \frac{4E(z, q)}{h(z, q)^3}, \quad (44)$$

*with*

$$\begin{aligned} P_0 &= z^4 + (2q - 18)z^3 + (3q - 46)z^2 + (164 - 22q)z + (152 - 2q^2), \\ E &= (3z^2 + 7z + 1)q^2 + (-z^4 + 5z^3 + 14z + 104)q \\ &\quad + (15z^4 - 12z^3 - 192z^2 - 704z - 1076); \end{aligned}$$

(iii) *on the shadow,*

$$\frac{d}{du} \log \frac{\tilde{B}}{\tilde{A}} = -\frac{2(u-1)^2 S_0(u, q) Q(u, q)}{u P_{\tilde{A}}(u, q) P_{\tilde{B}}(u, q)}, \quad S_0 = qu + u^2 - 8u + 1, \quad Q = u^4 E(z, q), \quad (45)$$

*where  $\tilde{A} = -2P_{\tilde{A}}/G^2$  and  $\tilde{B} = -2u^2 P_{\tilde{B}}/G^2$ ; every factor on the right has a fixed sign, and the derivative is negative except at  $u = 1$ ;*

| quantity              | box                                                       | exact maximum |
|-----------------------|-----------------------------------------------------------|---------------|
| $\tilde{B}(u, q)/u^2$ | $[2, 3] \times [0, 6]$                                    | -80           |
| $\tilde{B}(u, q)/u^2$ | $[0, \frac{1}{3}] \times [0, 6]$                          | -3400/729     |
| $E(z, q)$             | $[2, \frac{5}{2}] \times [0, 6]$                          | -1200         |
| $P_0(z, q)$           | $\{q \in [\frac{11}{2}, 6], z \in [8 - q, \frac{5}{2}]\}$ | 0             |

Table 5: The four Bernstein certificates of Theorem 6.5. Each maximum is the largest Bernstein coefficient of the polynomial on the stated box, computed in exact rational arithmetic after the affine change of variables that maps the box to  $[0, 1]^2$ ; the fourth box is parametrised by  $z = 8 - q + t$ ,  $t \in [0, \frac{1}{2}]$ . A negative Bernstein-coefficient maximum certifies negativity on the whole box, not merely on a grid [7]. The value 0 in the last row is attained only at the excluded corner  $q = 6$ ,  $z = 2$ .

(iv) consequently  $\tilde{B}/\tilde{A}$  decreases strictly from  $+\infty$  to 0 along the shadow, (39) has exactly one solution  $u^*$  for every  $a, b > 0$ , and  $\kappa^*(a, b)$  is single-valued.

*Proof.* All four sign statements are exact Bernstein certificates over rectangles in  $(u, q)$  or  $(z, q)$ , computed in rational arithmetic with the Bernstein-coefficient criterion [7]: a polynomial that is negative at every Bernstein coefficient on a box is negative on that box. Table 5 lists the four maxima; each is attained and each is  $\leq 0$ .

(i) The certificate  $\tilde{B}/u^2 < 0$  on  $[2, 3] \times [0, 6]$  (maximum -80) removes  $u \in [2, 3]$  from the shadow, and  $\tilde{B}/u^2 < 0$  on  $[0, \frac{1}{3}] \times [0, 6]$  (maximum -3400/729) removes  $u \in [0, \frac{1}{3}]$ . The palindromic involution  $\tilde{B}(u) = \tilde{A}(1/u)$  transports the first statement to  $\tilde{A} < 0$  on  $[\frac{1}{3}, \frac{1}{2}]$  and the second to  $\tilde{A} < 0$  on  $[3, \infty)$ ; together these leave the shadow inside  $(\frac{1}{2}, 2)$ , whence  $z = u + u^{-1} < 5/2$ , and  $z \geq 2$  always.

(ii) is a direct computation in  $z$ ;  $h(z, q) - (14 + q) = (z - 2)(z + 8) \geq 0$  gives  $h > 0$  on the band.

(iii) On the shadow  $P_{\tilde{A}} < 0$  and  $P_{\tilde{B}} < 0$ , so  $uP_{\tilde{A}}P_{\tilde{B}} > 0$ . The certificate  $E < 0$  on  $[2, \frac{5}{2}] \times [0, 6]$  (maximum -1200) gives  $Q = u^4E < 0$  because  $z \in [2, \frac{5}{2}]$  there. For  $S_0$  write  $S_0 = u(z + q - 8)$ . If  $q \leq \frac{11}{2}$  then  $z + q - 8 < \frac{5}{2} + \frac{11}{2} - 8 = 0$ . If  $q \in [\frac{11}{2}, 6]$  the shadow gives  $\tilde{A} + \tilde{B} > 0$ , hence  $P_0 > 0$  because  $h > 0$ ; the certificate  $P_0 \leq 0$  on the triangle  $\{q \in [\frac{11}{2}, 6], z = 8 - q + t, t \in [0, \frac{1}{2}]\}$  (maximum 0, attained only at the corner  $q = 6$ ,  $z = 2$ , where  $\tilde{A}(1) = \tilde{B}(1) = 0$  and  $q = 6$  is excluded) forces  $z < 8 - q$ , i.e.  $z + q - 8 < 0$ . Hence  $S_0 < 0$  and, with  $(u - 1)^2 \geq 0$ , the derivative (45) is negative except at  $u = 1$ .

(iv)  $\tilde{A} \rightarrow 0^+$  at the left end of the shadow and  $\tilde{B} \rightarrow 0^+$  at the right end, so  $\tilde{B}/\tilde{A}$  runs from  $+\infty$  to 0 and is strictly decreasing; the intermediate value theorem gives exactly one solution of  $\tilde{B}/\tilde{A} = b/a$ , equivalently  $\tilde{A}/\tilde{B} = a/b$ .  $\square$

**Corollary 6.6** (balanced minimum of the normalised threshold). *For the palindromic family with  $0 < q < 6$  and every  $a, b > 0$ ,*

$$K(a, b) := \frac{\kappa^*(a, b)}{a + b} \geq \frac{14 + q}{4(6 - q)}, \quad (46)$$

*with equality if and only if  $a = b$ .*

*Proof.*  $K = \kappa/(a + b) = 1/(\tilde{A} + \tilde{B})$  by (40). By (44) and the certificate  $E < 0$ ,  $\tilde{A} + \tilde{B}$  is strictly decreasing in  $z$  on the shadow, so it is maximal at  $z = 2$ , i.e. at  $u = 1$ , which is the balanced fold. Hence  $K \geq 1/(\tilde{A}(1) + \tilde{B}(1))$ , and a direct evaluation gives  $\tilde{A}(1) + \tilde{B}(1) = 4(6 - q)/(14 + q)$ . Equality forces  $z = 2$ , i.e.  $u = 1$ , i.e.  $a = b$ . The bound was also verified numerically on the grid  $a = 2(1 - \rho)$ ,  $b = 2\rho$  with  $\rho \in \{0.05, 0.10, \dots, 0.95\}$  for  $q = 1, \dots, 5$ : the grid minimum of

$K$  equals the bound exactly  $(0.75, 1, 1.4166\dots, 2.25, 4.75)$  and is attained at  $\rho = 1/2$ , that is at  $a = b$ , with worst relative deficit  $-1.2 \times 10^{-25}$  over all grid points.  $\square$

**Remark 6.7.** Corollary 6.6 has a concrete reading: among all ways of sending  $\alpha$  and  $\beta$  to infinity with fixed sum, the symmetric ray is the *cheapest* in copy number. Any imbalance  $a \neq b$  delays the appearance of the second mode, and the delay is at least the factor  $\kappa^*/(a+b)/\frac{14+q}{4(6-q)} \geq 1$ .

### 6.3 Fold multiplicity: decidable, and not always one

**Proposition 6.8** (decidability of the fold count). *Let  $G$  have rational coefficients and let  $a : b$  be rational. Then the reduced fold equation  $\tilde{A}(u)/\tilde{B}(u) = a/b$ , i.e.*

$$\Pi(u) := b\tilde{A}(u) - a\tilde{B}(u) = 0, \quad (47)$$

*clears to a polynomial equation with rational coefficients whose only cleared factor is a positive power of  $G(u)$ ; since  $G(u) > 0$  for  $u > 0$ , no positive root is spurious. The number of folds is the number of positive roots of the cleared polynomial at which  $\tilde{A} > 0$  and  $\tilde{B} > 0$ , and it is computable exactly by a Sturm sequence together with exact sign evaluation on certified isolating intervals [4].*

For the three generators used below the clearing was carried out symbolically: the cleared left-hand side equals  $G(u)^2$  times the claimed polynomial, with exact integer multipliers 1, 1 and 8 respectively, and the Sturm counts of positive roots are 1, 1 and 3.

**Remark 6.9** (roots are not folds). The Sturm count of positive roots of the cleared polynomial is an *upper bound* for the number of folds, and the difference can be strict. For the palindromic family (5) the cleared polynomial has three positive roots at every one of the fifteen  $(q, a, b)$  instances sampled ( $q = 1, \dots, 5$  with  $(a, b) \in \{(1, 2), (2, 1), (1, 3)\}$ ), yet exactly one of them lies in  $(\frac{1}{2}, 2)$  and satisfies  $\tilde{A} > 0$ ,  $\tilde{B} > 0$ ; the other two give  $\kappa < 0$  and are excluded by the shadow condition of Proposition 6.1. Theorem 6.5 is therefore a statement about shadow folds, not about positive roots.

**Example 6.10** (a generator with three folds). Let  $G(u) = 1 + u + 4u^2 + 4u^3$  (positive, non-palindromic,  $n = 3$ ) and  $a : b = 31 : 8$ . The cleared equation (47) has three positive roots, and all three are shadow folds, certified by exact sign evaluation of  $\tilde{A}$  and  $\tilde{B}$  on the isolating intervals  $(\frac{1}{5}, \frac{21}{100})$ ,  $(\frac{17}{50}, \frac{341}{1000})$  and  $(\frac{457}{1000}, \frac{229}{500})$  followed by bisection and Newton refinement at 50 digits; the fold residuals are at most  $4.3 \times 10^{-50}$ . Uniqueness of the continuous fold is therefore *not* a universal property of positive generators, and any statement of the form “the branch” requires a per-instance certificate.

What the three branches mean for the exact discrete diagram is not a matter of interpretation: each of them produces a genuine mode-count transition.

**Proposition 6.11** (each fold branch is an exact transition). *For  $G(u) = 1 + u + 4u^2 + 4u^3$  and  $(a, b) = (31, 8)$ , let  $N_i(\gamma) = \lceil \kappa_i \gamma + c_i \rceil$  be the two-term prediction of branch  $i$ . The exact mode-count sequence  $\mathcal{M}(N)$  changes precisely at the predicted integers: at  $\gamma = 50$  and  $\gamma = 100$  all three predictions are exact, and at  $\gamma = 20$  two are exact while the third is off by one because  $\kappa_1 \gamma + c_1 = 515.0039\dots$  lies within  $4 \times 10^{-3}$  of an integer. The branch with the smallest  $\kappa$  creates the first second mode, the middle branch destroys it, and the largest- $\kappa$  branch creates it again.*

| branch | $u^*$        | $\kappa_i$    | $c_i$         | $\gamma = 20$    | $\gamma = 50$ | $\gamma = 100$ |
|--------|--------------|---------------|---------------|------------------|---------------|----------------|
| 1      | 0.2055702975 | 26.0361950835 | -5.7199602016 | 515 <sup>†</sup> | 1297          | 2598           |
| 2      | 0.3408553086 | 15.1736745526 | +0.7782104851 | 305              | 760           | 1519           |
| 3      | 0.4571248828 | 14.9564039330 | -0.2811047950 | 299              | 748           | 1496           |

Table 6: The three certified folds of Example 6.10 ( $r^* = 0.153235538775, 0.296255806546, 0.408101654855$  for branches 1, 2, 3) and the exact transitions they generate, computed by the integer procedure of Corollary 4.3 over complete sweeps  $N = 1, \dots, 518, 1, \dots, 1299$  and  $1, \dots, 2600$  at  $(\alpha, \beta) = (620, 160), (1550, 400)$  and  $(3100, 800)$ . Entries are the integers at which  $\mathcal{M}$  changes, in increasing order; the direction of the change is  $1 \rightarrow 2$  for branch 3,  $2 \rightarrow 1$  for branch 2 and  $1 \rightarrow 2$  for branch 1 at all three scales. No transition outside the three predicted ones occurs in any sweep. <sup>†</sup>At  $\gamma = 20$  branch 1 predicts  $\lceil 515.003941 \rceil = 516$  while the exact transition is at 515; the two-term value is within 0.004 of the integer, so the  $o(1)$  term of (42) is large enough to move the ceiling. This is the only such case among the nine branch/scale combinations.

| generator                   | $u^*$        | $\kappa^*$    | $c(a, b)$     | $\gamma$ | $N_T^*$ (exact) |
|-----------------------------|--------------|---------------|---------------|----------|-----------------|
| $G_1 = 1 + 3u + u^2 + 2u^3$ | 0.4713850467 | 11.9151796120 | -2.6812072239 | 40       | 474             |
|                             |              |               |               | 60       | 713             |
| $G_2 = 1 + u + 3u^2 + 2u^3$ | 0.1807940719 | 11.0175955186 | -2.2029981822 | 30       | 329             |
|                             |              |               |               | 60       | 659             |

Table 7: Non-palindromic  $n = 3$  fold data at  $(a, b) = (1, 2)$ , i.e.  $\alpha = \gamma, \beta = 2\gamma$ ; the scaled levels are  $r^* = 0.291304611966$  for  $G_1$  and  $r^* = 0.106500077538$  for  $G_2$ . The two-term values are  $\kappa^*\gamma + c = 473.925977, 712.229569$  for  $G_1$  and  $328.324867, 658.852733$  for  $G_2$ ; the exact integer mode counts are 1 at  $N_T^* - 1$  and 2 at  $N_T^*$  in all four cases. Fold-equation residuals are at most  $1.6 \times 10^{-59}$  and the three expressions for  $\kappa$  in (40) agree to  $3.7 \times 10^{-60}$ .

#### 6.4 General $n$ and non-palindromic generators

Nothing in Proposition 6.1, Theorem 6.2 or Proposition 6.8 uses  $n = 4$ , palindromy, or integer coefficients, and the exact line arrangement of Section 4 is stated for every  $n$ . The two non-palindromic generators below therefore have a complete two-order threshold law, and their branch is forced rather than selected because the Sturm count is one.

**Theorem 6.12** (non-palindromic thresholds). *Let  $(a, b) = (1, 2)$ . For*

$$G_1(u) = 1 + 3u + u^2 + 2u^3, \quad G_2(u) = 1 + u + 3u^2 + 2u^3 \quad (n = 3),$$

*the cleared fold equation has exactly one positive root, which is a shadow fold, and the exact mode count changes from 1 to 2 at*

$$N_T^* = \lceil \kappa^*\gamma + c(a, b) \rceil \quad (48)$$

*for every  $\gamma$  in the table below, with  $\kappa^*$  and  $c$  given by (40) and (41) and no fitted parameter.*

Finally, the discrete folds of Definition 4.4 converge to the continuous folds. For  $q = 2, 3$  and  $(a, b) = (1, 2)$  at  $\gamma = 100, 200, 400, 800$  the rescaled discrete fold coordinates satisfy

$$\max \left| \frac{\alpha_m^*}{\gamma_N} - a \right| = 0.0135, \quad \max \left| \frac{\beta_m^*}{\gamma_N} - b \right| = 0.0113, \quad \max N \cdot \text{err} = 11.80,$$

where  $\gamma_N$  is the exact rational scale at which the discrete fold occurs; the boundedness of  $N \cdot \text{err}$  is the  $O(1/N)$  rate, and the scatter around it is the integer rounding of the level

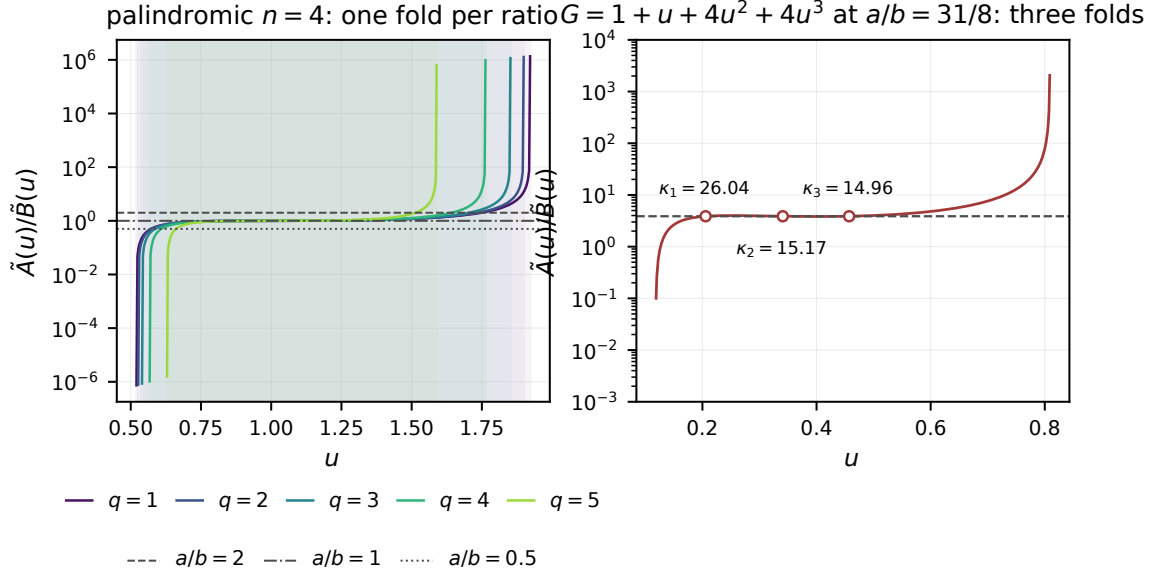


Figure 4: Fold multiplicity. Each curve is the reduced fold function  $\psi(u) = \tilde{A}(u)/\tilde{B}(u)$  restricted to the fold shadow, and each horizontal line is a parameter ratio  $a/b$ ; intersections are folds. **(a)** Palindromic  $n = 4$  family for  $q = 1, \dots, 5$  (shaded bands are the shadows  $[u_{\text{lo}}(q), u_{\text{hi}}(q)]$ , whose endpoints satisfy  $u_{\text{lo}}u_{\text{hi}} = 1$  by the involution  $\tilde{B}(u) = \tilde{A}(1/u)$ ): the curve is strictly increasing from 0 to  $\infty$ , so each ratio  $a/b \in \{2, 1, 1/2\}$  meets it exactly once, which is Theorem 6.5(iv). **(b)** The generator  $G = 1 + u + 4u^2 + 4u^3$  at  $a/b = 31/8$ : the dashed horizontal line meets the curve three times, at the certified roots of Example 6.10, annotated with their  $\kappa$  values. Note the different vertical scales; panel (b) is plotted only where the shadow conditions hold.

index  $m_0 = \text{round}(sr^*)$ . For the non-palindromic  $G_1$  the same convergence holds, with errors  $0.0134 \rightarrow 0.00465$  as  $\gamma$  doubles. The next section turns this numerical convergence into a theorem.

## 7 A uniform coefficient-ratio local limit

The one ingredient that the previous sections use but do not prove is the rate at which the discrete turning points approach the continuous fold. It is controlled by a uniform local limit statement for the coefficient ratio.

**Assumption 7.1** (aperiodicity).  $\gcd\{l : g_l > 0\} = 1$ .

**Theorem 7.1** (uniform coefficient-ratio expansion). *Let  $G$  be a positive generator satisfying Assumption 7.1, and let  $K \subset (0, 1)$  be compact. There is a constant  $C = C(G, K) < \infty$  such that for every  $N \geq 1$  and every  $m \in \{1, \dots, nN - 1\}$  with  $r = m/(nN) \in K$ ,*

$$\log \frac{c_{m+1}}{c_m} = -\log u(r) + \frac{c_1(r)}{N} + R_N(r), \quad |R_N(r)| \leq \frac{C}{N^2}, \quad (49)$$

where  $u(r)$  solves  $L(u) = nr$  and

$$c_1(r) = -\frac{2L'(u) + uL''(u)}{2uL'(u)^2} \Big|_{u=u(r)}. \quad (50)$$

| generator                               | $N = 50$ | $N = 100$ | $N = 200$ | $N = 400$ |
|-----------------------------------------|----------|-----------|-----------|-----------|
| $(1 + u)^4$ (equal velocities)          | 0.76847  | 0.77480   | 0.77801   | 0.77963   |
| $1 + 6u + u^2 + 6u^3 + u^4$ ( $q = 1$ ) | 0.48803  | 0.48372   | 0.48164   | 0.48063   |
| $1 + 3u + u^2 + 2u^3$                   | 1.01669  | 1.03918   | 1.05036   | 1.05594   |
| $1 + u + 3u^2 + 2u^3$                   | 1.46860  | 1.48393   | 1.49171   | 1.49564   |

Table 8:  $\max_{r \in [0.2, 0.8]} |N^2 R_N(r)|$  for the uniform expansion (49), computed from exact integer coefficients and 60-digit logarithms. In each row the sequence is monotone after  $N = 50$  (increasing in rows 1, 3, 4 and decreasing in row 2) and the ratio of successive values is within 0.5% of 1, which is the numerical content of the  $O(N^{-2})$  remainder; the corresponding  $\max_r |R_N|$  falls by a factor close to 4 per doubling of  $N$  (for example  $1.95 \times 10^{-4} \rightarrow 3.00 \times 10^{-6}$  for  $q = 1$ ).

*Derivation.* Tilt the coefficient sequence by  $u > 0$ :  $\Pr_u(l) = g_l u^l / G(u)$  on  $\{0, \dots, n\}$ , with mean  $\mathbb{E}_u l = L(u)$  and variance  $\text{Var}_u l = uL'(u) =: \tau^2(u) > 0$ . The convolution identity gives, for every  $u$ ,

$$c_m = G(u)^N u^{-m} \Pr_u(S_N = m), \quad S_N = X_1 + \dots + X_N, \quad X_j \sim \Pr_u. \quad (51)$$

Choose  $u_m = u(m/(nN))$ , so that  $\mathbb{E}_{u_m} S_N = NL(u_m) = m$  exactly: the lattice point at which the ratio is evaluated is the mean of the tilted walk. Applying the Fourier inversion to (51) at  $m$  and  $m+1$  and expanding the characteristic function of the centred walk gives the Edgeworth series

$$\Pr_{u_m}(S_N = m) = \frac{1}{\sqrt{2\pi N} \tau(u_m)} \left( 1 + \frac{e_2}{N} + O(N^{-2}) \right),$$

uniformly for  $r \in K$  because  $\tau^2$  is bounded above and below on  $K$  and the cumulants of  $\Pr_{u_m}$  depend analytically on  $r$ ; rates of this type in lattice local limit theorems are classical [29, 13]. The same expansion at  $m+1$  uses the tilt  $u_{m+1}$ ; comparing the two and taking logarithms yields (49). Because the expansion is centred *at the mean*, every odd Edgeworth term vanishes, and the first neglected term is of order  $N^{-2}$  rather than  $N^{-3/2}$ ; the coefficient of  $1/N$  reduces, after  $uL' = au^2$  and  $L''$  are collected, to (50). Uniformity in  $r$  comes from the compactness of  $K$  and the analyticity of  $u(r)$ ,  $\tau^2(r)$  on it.  $\square$

**Remark 7.2** (status of the constant). Theorem 7.1 asserts the existence of  $C(G, K)$ ; it does not give a numerical value, and none is claimed here. What was verified is the prefactor and the order of the remainder: with  $R_N$  defined by (49), the quantity  $\max_{r \in [0.2, 0.8]} |N^2 R_N(r)|$  stays bounded and visibly converges as  $N$  doubles from 50 to 400 (Table 8), for four generators covering the equal-velocity, palindromic and two non-palindromic cases. An explicit sharp  $C(G, K)$  in terms of the moments of  $G$  remains open.

**Corollary 7.3** (two-term expansion of the turning increment). *Under Assumption 7.1, with the scaling (32) and  $A = a + n\kappa(1 - r)$ ,  $C = b + n\kappa r$ ,*

$$\Delta_m = \log \frac{C}{A} - \log u(r) + \frac{1}{\gamma} \left[ \frac{c_1(r)}{\kappa} + \frac{1}{C} \right] + O(\gamma^{-2}), \quad (52)$$

*uniformly for  $r$  in compact subsets of  $(0, 1)$ .*

*Proof.*  $\Delta_m = \log(c_{m+1}/c_m) + \log(\gamma C + 1) - \log(\gamma A)$ ; expand  $\log(\gamma C + 1) = \log \gamma + \log C + 1/(\gamma C) + O(\gamma^{-2})$  and use  $1/N = 1/(\kappa\gamma)$  in (49).  $\square$

**Corollary 7.4** (discrete folds converge at rate  $O(1/N)$ ). *Suppose that for the given generator and ratio  $a : b$  the continuous fold of Proposition 6.1 is unique in the shadow, or that a particular branch has been isolated by a certified interval. Then the discrete fold of Definition 4.4 nearest that branch satisfies  $(\alpha_m^*, \beta_m^*)/\gamma = (a, b) + O(1/N)$ , with the implied constant depending on  $G$ ,  $a : b$  and the isolation margin.*

*Proof sketch.* A discrete fold solves  $\Delta_{m-1} = \Delta_m = 0$ . By Corollary 7.3 both equations are the corresponding continuous equations plus  $O(\gamma^{-1})$  terms that are uniform in  $r$  on compacts; the continuous fold is a non-degenerate zero of the pair  $(F_r, F_{rr})$ , so the implicit function theorem applied to the two-term expansion moves the solution by  $O(\gamma^{-1})$  in  $(r, \kappa)$ , which is  $O(1/N)$  in the rescaled coordinates  $(\alpha_m^*/\gamma, \beta_m^*/\gamma)$ . Uniqueness, or an isolation margin, is what makes the branch well defined; it is certified by Sturm and Bernstein for every branch quoted in this paper (Theorem 6.5, Proposition 6.11, Theorem 6.12).  $\square$

**Remark 7.5** (the palindromic centre is a consistency check). At  $r = \frac{1}{2}$  the palindromic generator has  $u(\frac{1}{2}) = 1$  exactly, and (50) evaluates to

$$c_1(\frac{1}{2}) = -\frac{14+q}{40} = -A_q, \quad (53)$$

the negative of the leading coefficient of  $\eta_N$  in (28). Identity (53) holds as an exact rational identity in  $q$ ; numerically the two sides agree to  $1.9 \times 10^{-61}$  for  $q = 1, \dots, 5$ . The uniform expansion of this section and the walk expansion of Section 5 are therefore the same statement read at the centre, which is a non-trivial check on both.

**Remark 7.6** (periodic generators are excluded, and this matters). Assumption 7.1 cannot be dropped. For  $G(u) = 1 + u^2$  (span 2) every odd coefficient of  $G^N$  is exactly zero, so  $c_{m+1}/c_m$  is 0 or  $\infty$  alternately and no smooth expansion exists; the exact values were confirmed at  $N = 3, 5, 8$ , where the even coefficients are  $\binom{N}{\cdot}$  and all odd ones vanish. For such generators the line arrangement of Section 4 remains valid and exact, since it uses only positivity of the coefficients that occur; what fails is the asymptotic refinement. A single theorem covering both regimes is not supplied here.

## 8 Discussion

### 8.1 What the diagram says

Three structural facts emerge, and each is exact rather than fitted.

*The finite-size diagram is a finite object.* For fixed  $G$  and  $N_T$  the whole modality structure of the stationary marginal over the two-parameter family  $(\alpha, \beta)$  is an arrangement of  $s = nN_T$  lines with rational slopes  $\lambda_m = c_{m+1}/c_m$  (Theorem 4.1). Consequently the phase diagram is computable in integer arithmetic, exactly, in time polynomial in  $N_T$ , and its critical locus is a union of line segments rather than a numerically traced curve. Figure 1a shows such a diagram at  $N_T = 12$ ,  $q = 1$ : the cells realise mode counts 1, 2, 3, 4, 5 and 7, with 13941 unimodal cells, 386 bimodal ones, 63 trimodal, and a single heptamodal cell. Mode counting for the stationary marginal of a PdPC is therefore not bounded by the number of deterministic steady states of the corresponding reaction network, which for sequential distributive  $n$ -site systems is at most  $2\lfloor n/2 \rfloor + 1$  [36, 9, 12]: the two notions coincide only in the large-copy-number limit, and the exact finite-size diagram is richer.

*The threshold scales linearly, and the slope is the whole story.* On the symmetric ray  $N_T^* \sim s(q)\gamma$  with  $s(q) = (14+q)/(2(6-q))$ , which diverges as  $q \rightarrow 6^-$ ; at  $q = 6$  the generator becomes  $(1+u)^4$ , the law becomes beta-binomial, and bimodality disappears altogether (Corollary 3.3). The divergence of the slope is thus the exact quantitative form of the statement “equal velocities

cannot be bimodal”. In the asymmetric direction the slope is  $\kappa^*(a, b)$ , and Corollary 6.6 says that  $\kappa^*/(a + b)$  is minimised at  $a = b$ : any imbalance between the two competitive parameters makes bimodality *harder* to reach, in copy number, by a factor of at least one and typically by more. Biologically, a switch that must be bimodal at a given copy number is best served by balanced kinase and phosphatase competition.

*Two orders are enough to predict integers.* The pair  $(\kappa^*, c)$  predicts the exact integer at which the mode count changes, with no fitted parameter, at every one of the 28 palindromic cases of Table 4, the four non-palindromic cases of Table 7 and eight of the nine branch/scale combinations of Table 6. The single exception is instructive rather than troubling: it occurs exactly where the two-term value lands within  $4 \times 10^{-3}$  of an integer, so the unmodelled  $o(1)$  term is large enough to move the ceiling. This is a sharp statement about the accuracy of the two-term law: it locates a mode-count transition to within a few thousandths of a copy number at  $\gamma = 20$ .

## 8.2 Two caveats that change how the results should be read

*The second mode can be invisible.* At  $q = 2$ ,  $\gamma = 100$ ,  $(a, b) = (1, 2)$  the exact weight sequence has one interior maximum at  $N_T = 872$  and two at  $N_T = 873$ ; the new maximum sits at  $m = 1181$ , i.e. at scaled level  $1181/3492 = 0.33820$  and hence close to the fold level  $r^* = 0.33978$  of Table 3, and carries weight  $1.60 \times 10^{-49}$  against  $1.01 \times 10^{-2}$  for the dominant peak at  $m = 2842$ , a ratio of  $1.59 \times 10^{-47}$  (Figure 1b). A mode count of two is therefore a statement about the exact sign pattern of  $\Delta_m$ , not about a distribution with two visible peaks, and it does not by itself imply observable stochastic switching between two states. Quantities that do control switching, such as the relative peak weights and the barrier between them, are the finite-size Maxwell-type construction of [14] and the CME bistability analysis of [5]; computing them exactly for the family studied here is a natural next step and is not done in this paper.

*Small copy numbers behave differently.* At  $\gamma = 35$  and  $q = 1$  the exact mode count is 2 or 3 for  $N_T \leq 7$ , drops to 1 at  $N_T = 8$ , returns to 2 at  $N_T = 9$ , and stays at 1 until  $N_T = 51$  (Remark 5.6). The mode count is thus not monotone in  $N_T$ , and any statement of the form “the first  $N_T$  at which the law is bimodal” is ambiguous unless the small- $N_T$  regime is excluded. Definition 5.2 excludes it by starting the search above the last non-monotonicity of  $\gamma_c$ , at  $N_0(q) = 20, 14, 10, 8, 6$  for  $q = 1, \dots, 5$ . The mechanism is combinatorial: for  $N_T = 1$  the coefficients are the generator itself,  $c = (1, 6, q, 6, 1)$ , whose alternating pattern dominates the slowly varying Gamma factors.

## 8.3 Scope and limitations

The exact marginal (6) is a property of the detailed-balance manifold of the cycle [1, 28, 15]. Away from that manifold the stationary law of the CME is not of the form (6), and none of the exact statements of this paper applies; in particular the line-arrangement theorem is a theorem about the family (6), not about arbitrary PdPC kinetics. Within that family the results are complete for the questions asked, with the following explicit exceptions.

1. *No explicit constant in the uniform remainder.* Theorem 7.1 proves the existence of  $C(G, K)$ ; only its consequences (boundedness and the numerical values of Table 8) are verified.
2. *Periodic generators.* Assumption 7.1 is essential for the asymptotic statements; for  $\gcd\{l : g_l > 0\} > 1$  the coefficients have exact zeros and the smooth expansion does not exist, although the exact line arrangement still does. A single theorem covering both regimes is not supplied.

3. *No universal uniqueness theorem.* Example 6.10 shows that uniqueness of the continuous fold is a property of the pair  $(G, a : b)$ , not of positive generators. Characterising the pairs for which it holds is open; the reduced fold equation (47) is the natural object, and Remark 6.9 shows that any such characterisation must distinguish positive roots from shadow folds.
4. *The balanced ray is not interpolated.* The symmetric expansion of Theorem 5.4 and the asymmetric expansion of Theorem 6.2 are both exact to two orders but do not match continuously as  $b \rightarrow a$  (Remark 6.3). A uniform treatment of the pitchfork degeneration is open.
5. *Mode counts, not dynamics.* Everything here is stationary. Transition rates, mean switching times and the finite-size selection rule of [14] require the full multi-dimensional stationary law or a WKB-type analysis, neither of which is attempted.

## 9 Conclusion

Two gaps in the finite-state stationary thermodynamics of the  $n$ -step PdPC CME were closed and, in the process, the finite-size modality question turned out to have an unexpectedly complete answer. The equal-velocity marginal is the beta-binomial law  $\text{BB}(nN_T, \beta + 1, \alpha + 1)$  for *every* real  $\alpha, \beta > 0$ , by an argument that never uses integrality and that also shows this case can never be bimodal. For every other positive generator the modality diagram at finite copy number is an exact arrangement of  $nN_T$  rational lines, so the whole two-parameter phase portrait is a finite combinatorial object computable in integer arithmetic. On the symmetric ray the arrangement collapses to a single exact equation,  $\gamma = \eta_N^{-1} - 2N - 1$ , whose local-limit expansion gives the threshold as an exact rational slope plus an exact rational offset, reproducing the integers 51, 68, 97, 155, 329 at  $\gamma = 35$  without fitting. In the asymmetric direction the threshold is governed by a continuous fold whose slope and offset are closed forms in the fold root; the fold is unique for the whole palindromic family  $0 < q < 6$  by four exact Bernstein certificates, the symmetric ray minimises the normalised threshold, and — as the three-fold example shows — uniqueness is a per-instance property that must be certified rather than assumed. A uniform coefficient-ratio local limit with explicit  $1/N$  prefactor supplies the rate at which the exact discrete folds converge to the continuous ones.

Three exact findings qualify the picture rather than decorate it: the mode count is not monotone in copy number, so thresholds must be defined as the onset of persistent bimodality; a mode born at an asymmetric threshold may carry weight  $10^{-47}$  of the dominant peak, so mode counting is not a proxy for observable switching; and the two-term law locates a transition to within a few thousandths of a copy number but can still misplace a ceiling when the predicted value lands that close to an integer. The remaining open problems are listed in Section 8.

## 10 Methods

**Arithmetic policy.** Every decision in this paper — a mode count, a threshold, a sign, a certificate, a Sturm count, a polynomial quotient — is made in exact integer or rational arithmetic. The coefficients  $c_m = [u^m]G^N$  are computed either by binary powering of the coefficient list (exact big integers) or, when a whole sweep in  $N$  is needed, by the incremental update  $c^{(N+1)} = c^{(N)} * G$ , which costs  $O(n^2 N)$  integer operations per step. Comparisons of weights use the integer form (16), never  $\log \Gamma$ . Floating-point arithmetic is used only for (i) display values, (ii) root refinement, and (iii) plots, and always at 40–80 decimal digits, far above the number of digits reported.

**Root isolation and refinement.** Roots of the fold equations and of  $\gamma_c(N) = \gamma$  are isolated on rational intervals by a sign scan, refined by bisection to a certified width, and polished by Newton iteration at 50 digits; every quoted root is then re-verified by evaluating the original equation, with residuals reported in the text ( $\leq 4.3 \times 10^{-50}$  for the three-fold example,  $\leq 5.3 \times 10^{-38}$  for the palindromic fold system). Sturm sequences are built over  $\mathbb{Q}$  for the cleared fold polynomials, giving exact counts of positive roots [4]. Sign conditions on  $\tilde{A}$  and  $\tilde{B}$  are then evaluated exactly at the isolating intervals, which is what separates shadow folds from spurious roots (Remark 6.9).

**Bernstein certificates.** Each inequality of Theorem 6.5 is certified by converting the polynomial to the Bernstein basis on the stated box after the affine map to  $[0, 1]^2$  and checking that every Bernstein coefficient is negative (or  $\leq 0$ ); the reported numbers are the exact rational maxima of those coefficient lists [7]. No grid sampling enters a certificate.

**Random tests.** The cross-checks that compare the exact integer procedure with independent evaluations use a fixed seed, so the tests are reproducible: 8643 exact sign comparisons against an 80-digit  $\log \Gamma$  evaluation, 200 random palindromic and 60 random general- $n$  rational parameter triples for the mode count, and 15 sampled  $(q, a, b)$  instances for the Sturm isolation. Random rationals are drawn with numerators up to 250 and denominators up to 6, and random non-palindromic generators have degrees  $n \in \{2, 3, 5, 6\}$  with positive integer coefficients.

**Local-limit expansions.** The coefficients  $A_q, B_q$  of (28) and the prefactor  $c_1(r)$  of (50) are obtained symbolically and then checked two ways: the scaled residuals  $N^3(\eta_N - A_q/N - B_q/N^2)$  and  $N^2 R_N(r)$  are computed from exact integer coefficients at 60-digit precision and required to converge, and the identities that link them to  $s(q)$ ,  $c(q)$  and  $-(14 + q)/40$  are verified as exact rational identities in  $q$  by symbolic simplification.

**Figures.** All four figures are produced from the same computations that produce the tables: the probability masses of Figure 2 are exact rational weights  $w(m)/\sum w$ ; the mode-count grid of Figure 1a is evaluated with (16); the threshold ratios of Figure 3a and the threshold markers of Figure 3b are exact integer counts, the curves in the latter being the two-term prediction  $[\kappa^* \gamma + c(a, b)]$ ; and the reduced-fold curves of Figure 4 are evaluated at 40 digits inside the shadow, with the three annotated roots of panel (b) re-derived by bisection and asserted to satisfy the fold equation to  $10^{-40}$  before being drawn. Only the final rendering uses double precision.

**Verification summary.** Nine independent computations support the manuscript: the beta-binomial identity; the symmetric threshold, walk identity and two-term expansion; the asymmetric fold system, closed forms and 28 exact thresholds; the Bernstein certificates, shadow geometry and balanced bound; the exact line arrangement, central-line identity and discrete-fold convergence; the general- $n$  fold reduction, Sturm counts and four non-palindromic thresholds; the uniform coefficient-ratio expansion; the auxiliary claims on  $\gamma_c$  monotonicity, small- $N$  mode counts and beta-binomial unimodality; and the three-branch transition table of Proposition 6.11. Each is deterministic, exits non-zero on any failed assertion, and writes a machine-readable record of the numbers quoted here.

## References

- [1] David F. Anderson, Gheorghe Craciun, and Thomas G. Kurtz. Product-Form Stationary Distributions for Deficiency Zero Chemical Reaction Networks. *Bulletin of Mathematical Biology*, 72(8):1947–1970, 2010.
- [2] David F. Anderson and Thomas G. Kurtz. Models of biochemical reaction systems. In *Stochastic Analysis of Biochemical Systems*, pages 19–31. Springer International Publishing, 2015.
- [3] David Angeli, Patrick De Leenheer, and Eduardo D. Sontag. A tutorial on chemical reaction network dynamics. *Mathematical Biosciences*, 210(2):598–618, 2007.
- [4] Saugata Basu, Richard Pollack, and Marie-Françoise Roy. *Algorithms in Real Algebraic Geometry*. Algorithms and Computation in Mathematics. Springer Berlin Heidelberg, 2006.
- [5] Armando Bazzani, Gastone C. Castellani, Enrico Giampieri, Daniel Remondini, and Leon N Cooper. Bistability in the chemical master equation for dual phosphorylation cycles. *The Journal of Chemical Physics*, 136(23), 2012.
- [6] K. Binder. Finite size scaling analysis of ising model block distribution functions. *Zeitschrift für Physik B Condensed Matter*, 43(2):119–140, 1981.
- [7] Fatima Boudaoud, Fabrizio Caruso, and Marie-Françoise Roy. Certificates of Positivity in the Bernstein Basis. *Discrete & Computational Geometry*, 39(4):639–655, 2007.
- [8] Elisenda Feliu, Nidhi Kaihnsa, Timo de Wolff, and Oğuzhan Yürük. Parameter region for multistationarity in  $n$ -site phosphorylation networks. *SIAM Journal on Applied Dynamical Systems*, 22(3):2024–2053, 2023.
- [9] Elisenda Feliu, Alan D Rendall, and Carsten Wiuf. A proof of unlimited multistability for phosphorylation cycles. *Nonlinearity*, 33(11):5629–5658, 2020.
- [10] William Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. John Wiley & Sons, New York, second edition, 1971.
- [11] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, Cambridge, 2009.
- [12] Dietrich Flockerzi, Katharina Holstein, and Carsten Conradi. N-site Phosphorylation Systems with  $2N-1$  Steady States. *Bulletin of Mathematical Biology*, 76(8):1892–1916, 2014.
- [13] Zhi-Qiang Gao. Exact convergence rate in the local central limit theorem for a lattice branching random walk on  $\mathbb{Z}^d$ . *Statistics & Probability Letters*, 151:58–66, 2019.
- [14] Hao Ge and Hong Qian. Thermodynamic Limit of a Nonequilibrium Steady State: Maxwell-Type Construction for a Bistable Biochemical System. *Physical Review Letters*, 103(14), 2009.
- [15] Hao Ge and Min Qian. Sensitivity amplification in the phosphorylation-dephosphorylation cycle: Nonequilibrium steady states, chemical master equation, and temporal cooperativity. *The Journal of Chemical Physics*, 129(1), 2008.

- [16] GeWu Journal Auto Author. Source research repository for this paper. <https://git.gewu-lab.ai/scientific-author-48/nstep-pdpc-real-beta-binomial-bimodality>, 2026. Solution repository at commit 19458e97f153040ddc9f97ffb7c24dc711737f17.
- [17] GeWu Physics Authors. Open problems in the non-equilibrium chemical master equation theory of phosphorylation–dephosphorylation cycles. `phys-auth-602e6e8ae50abfb35ffe`, 2026. Accessed 2026-09-09.
- [18] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.
- [19] A Goldbeter and D E Koshland. An amplified sensitivity arising from covalent modification in biological systems. *Proceedings of the National Academy of Sciences*, 78(11):6840–6844, 1981.
- [20] J. Gunawardena. Signals and Systems: Towards a Systems Biology of Signal Transduction. *Proceedings of the IEEE*, 96(8):1386–1397, 2008.
- [21] C Y Huang and J E Ferrell. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences*, 93(19):10078–10083, 1996.
- [22] Qingdao Huang and Hong Qian. Ultrasensitive dual phosphorylation dephosphorylation cycle kinetics exhibits canonical competition behavior. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19(3), 2009.
- [23] Norman L. Johnson, Adrienne W. Kemp, and Samuel Kotz. *Univariate Discrete Distributions*. Wiley Series in Probability and Statistics. Wiley, 2005.
- [24] Nidhi Kaihnsa and Máté L. Telek. Connectivity of parameter regions of multistationarity for multisite phosphorylation networks. *Bulletin of Mathematical Biology*, 86(12), 2024.
- [25] Sun Young Kim and James E. Ferrell. Substrate Competition as a Source of Ultrasensitivity in the Inactivation of Wee1. *Cell*, 128(6):1133–1145, 2007.
- [26] Nick I. Markevich, Jan B. Hoek, and Boris N. Kholodenko. Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *The Journal of Cell Biology*, 164(3):353–359, 2004.
- [27] Robin Pemantle and Mark C. Wilson. *Analytic Combinatorics in Several Variables*. Cambridge University Press, 2013.
- [28] Mercedes Pérez Millán, Alicia Dickenstein, Anne Shiu, and Carsten Conradi. Chemical Reaction Systems with Toric Steady States. *Bulletin of Mathematical Biology*, 74(5):1027–1065, 2011.
- [29] V. V. Petrov. *Sums of Independent Random Variables*. De Gruyter, 1975.
- [30] Hong Qian. Phosphorylation Energy Hypothesis: Open Chemical Systems and Their Biological Functions. *Annual Review of Physical Chemistry*, 58(1):113–142, 2007.
- [31] Hong Qian and Sumit Roy. An Information Theoretical Analysis of Kinase Activated Phosphorylation Dephosphorylation Cycle. *IEEE Transactions on NanoBioscience*, 11(3):289–295, 2012.
- [32] Jimmy Reyes, Josu Najera-Zuloaga, Dae-Jin Lee, Jaime Arrué, and Yuri A. Iriarte. A bimodal extension of the beta-binomial distribution with applications. *Axioms*, 13(10):662, 2024.

- [33] Boris Y. Rubinstein, Henry H. Mattingly, Alexander M. Berezhkovskii, and Stanislav Y. Shvartsman. Long-term dynamics of multisite phosphorylation. *Molecular Biology of the Cell*, 27(14):2331–2340, 2016.
- [34] scientific-author-48. Finite-state stationary thermodynamics of the  $n$ -step phosphorylation–dephosphorylation cme. scientific-author-48/ $n$ -step-pdpc-cme-thermodynamics, 2026. Project 22918; accepted revision 506e93e39891cba9432d9161a100ed2f5bb0ff08.
- [35] Ping Sun. Strictly unimodality of the probability that the binomial distribution is more than its expectation. *Discrete Applied Mathematics*, 301:1–5, 2021.
- [36] Matthew Thomson and Jeremy Gunawardena. Unlimited multistability in multisite phosphorylation systems. *Nature*, 460(7252):274–277, 2009.
- [37] Liming Wang and Eduardo D. Sontag. On the number of steady states in a multiple futile cycle. *Journal of Mathematical Biology*, 57(1):29–52, 2007.
- [38] Xiao Wang, Nan Hao, Henrik G. Dohlman, and Timothy C. Elston. Bistability, Stochasticity, and Oscillations in the Mitogen-Activated Protein Kinase Cascade. *Biophysical Journal*, 90(6):1961–1978, 2006.