OPTIMAL RECURSIVE COMPOSITION AND DYADIC PHASE LAWS FOR GRADIENT DESCENT WITH PREDETERMINED STEPSIZES

Anonymous authors
Paper under double-blind review

ABSTRACT

Predetermined stepsize schedules featuring carefully chosen long steps have recently been shown to accelerate gradient descent (GD) on smooth convex functions. A prominent class of such schedules is built through recursive composition. In this paper, we characterize the convergence of these optimized recursive schedules, revealing a nonconstant log-periodic modulation across prescribed horizons. Specifically, we prove that for every $N \geq 1$, the optimized ConPP/OBS-S schedules (Zhang & Jiang, 2026; Grimmer et al., 2025b) satisfy

$$f(x_{N-1}) - f_* \leq \frac{1}{2N^p \Phi(\log_2 N) - 1} \frac{L}{2} \|x_0 - x_*\|^2, \qquad p = \log_2(1+\sqrt{2}),$$

where Φ is a positive, Lipschitz, nonconstant 1-periodic function. We derive this by proving that balanced splitting is optimal at every horizon for these constructions, resolving a conjecture of Zhang & Jiang (2026). Furthermore, for their asymmetric counterparts (ConPD/OBS-F schedules), we show that although optimal splits are not necessarily balanced, the same Silver exponent asymptotically persists alongside a distinct log-periodic modulation.

1 INTRODUCTION

Let $f : \mathbb{R}^d \to \mathbb{R}$ be convex with L-Lipschitz gradient and a nonempty minimizer set. Given a normalized schedule $h = (h_0, \ldots, h_{n-1})^\top \in \mathbb{R}^n$, gradient descent takes the form

$$x_{t+1} = x_t - \frac{h_t}{L} \nabla f(x_t), \quad t = 0, \ldots, n-1. \tag{1}$$

The schedule is fixed before optimization begins, so all algorithmic design lies in the predetermined scalar sequence h. This leads to a basic question: *how much can gradient descent be accelerated by stepsize design alone?*

For $h_t \in (0, 2)$, every nonstationary iteration decreases the objective. Tight worst-case analysis and carefully chosen varying steps can improve constants within this descent regime, while the best achievable order remains n^{-1} (Drori & Teboulle, 2014; Teboulle & Vaisbourd, 2023). One-dimensional Huber instances, in particular, rule out a uniform $o(n^{-1})$ guarantee for any schedule confined to this regime. Faster convergence therefore requires giving up monotonicity at the level of individual iterations. Long-step schedules trade this local monotonicity for progress over a larger block: some iterations may increase the objective, while their combined effect yields a stronger guarantee.

Recent work has made such long-step blocks systematic. Performance-estimation methods revealed numerically promising schedules (Taylor et al., 2017; Das Gupta et al., 2024), while subsequent constructions established first constant-factor and then polynomial improvements (Grimmer, 2024; Grimmer et al., 2023). A key structural device behind these advances is recursive composition: two shorter certified schedules are placed on either side of a specially chosen joining step, producing a longer certified schedule. Iterating this operation yields the Silver schedule of Altschuler & Parrilo (2025a;b), with objective-gap rate $O(n^{-p})$ at horizons $n = 2^k - 1$, where $\rho := 1 + \sqrt{2}$ and $p := \log_2 \rho \approx 1.27155$. Later constructions sharpened its objective-gap constant and obtained

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

the same exponent for the squared gradient norm (Grimmer et al., 2025a). At the dyadic horizons $n = 2^k - 1$, this recursion is exactly self-similar. The natural structural question is what replaces this symmetry at the horizons in between.

It is convenient to write $N = n+1$, so that the Silver horizons correspond exactly to $N = 2^k$. Away from these dyadic horizons, recursive composition becomes a split-selection problem: for each N, one must choose child sizes m and $N-m$. Two concurrent frameworks formalize this optimization. Zhang & Jiang (2026) introduced the ConPP and ConPD recursive constructions, while Grimmer et al. (2025b) independently developed the corresponding OBS-S and OBS-F families through composable certificates. The symmetric ConPP/OBS-S family extends the Silver construction to every horizon, whereas the asymmetric ConPD/OBS-F family is tailored to the final objective gap. Under matched split choices, the corresponding constructions generate the same numerical schedules (Zhang & Jiang, 2026, Remark 5.9).

For the symmetric ConPP/OBS-S family, the joining rules reduce the optimization to choosing a split $m + (N - m)$. Exhaustive dynamic programming finds an optimal split for each N, while Zhang & Jiang (2026, Conjecture 4.17) conjectured a uniform structural rule: the balanced split is always optimal, with certain unbalanced splits attaining the same value. Establishing such a uniform rule is nontrivial because the join is symmetric but nonassociative: symmetry allows the two children to be exchanged, but gives no comparison between different hierarchical pairings.

We resolve this conjecture, including all unbalanced ties. The resulting balanced recurrence exposes an exact dyadic self-similarity: doubling N multiplies the optimized scalar value by $\rho = 1 + \sqrt{2}$. This scaling links the optimal composition structure to the behavior of the convergence bound across horizons. The Silver exponent $p = \log_2 \rho$ determines the overall power-law decay, but leaves open whether the corresponding normalized coefficient approaches a single constant. We show instead that a nontrivial dependence on the fractional part of $\log_2 N$ remains. We call this quantity the *dyadic phase*: it records the position of N between consecutive powers of two. For the symmetric recursion, this phase dependence is exact at every finite horizon. For the asymmetric recursion, whose optimal splits need not be balanced, the normalized coefficients converge uniformly over dyadic blocks to a distinct nonconstant phase profile. Thus the Silver exponent together with the dyadic phase describes the leading-order horizon dependence of both optimized recursive families.

Contributions. We establish three results on the structure and convergence of these optimized recursive schedules.

Optimal recursive structure. We prove that balanced splitting is optimal at every horizon, resolving the ConPP conjecture of Zhang & Jiang (2026), and characterize all unbalanced splits that attain the same value. The key mechanism is a sharp four-point rearrangement inequality that compares different hierarchical pairings and propagates through the composition tree. As a consequence, we characterize all optimal composition trees and obtain an $O(\log N)$-time, $O(1)$-workspace evaluator for a single optimized scalar value U_N (Section 3).

Exact symmetric phase. The optimized values satisfy $U_N = N^p \Phi(\log_2 N)$, where Φ is positive, Lipschitz, 1-periodic, and nonconstant. This yields an exact formula for the tight objective coefficient and the complete nondegenerate interval of normalized subsequential limits (Section 4).

Asymmetric phase beyond balanced composition. For ConPD/OBS-F, optimal splits can be unbalanced, yet the normalized tight objective coefficients still converge uniformly over dyadic blocks to a distinct nonconstant phase profile. We further prove that this limiting profile stays strictly above the sharp homogeneous support bound obtained from the join map (Section 5).

Tightness. The finite-horizon coefficients underlying both phase laws are exact worst-case objective coefficients. For ConPP/OBS-S, equality is attained by one-dimensional Huber instances; for ConPD/OBS-F, matching one-dimensional quadratic and Huber instances attain the objective bound.

Positioning. Recent lower bounds identify the Silver exponent as the optimal polynomial exponent for predetermined nonnegative stepsizes (Ye & Liu, 2026). A complementary line of work studies anytime acceleration, where a single infinite schedule is required to provide guarantees across all stopping times (Kornowski & Shamir, 2024; Zhang et al., 2025; Tsai et al., 2026). Our results refine the finite-horizon picture for the ConPP/OBS-S and ConPD/OBS-F recursive families by characterizing their optimal composition structure and phase-dependent leading coefficients.

2 From Recursive Schedules to a Scalar Bellman Problem

A common scalar representation. The symmetric ConPP and OBS-S families admit the same scalar description. For a schedule h, define

$$u(h) := 1 + \sum_i h_i.$$

The empty schedule has value $u = 1$. If two child schedules a and b have scalar values u and v, respectively, both recursive constructions are governed by the same composition kernel

$$\mathcal{K}(u, v) := \frac{u + v + \sqrt{u^2 + 6uv + v^2}}{2}, \qquad u, v \geq 0. \tag{2}$$

They insert the joining step $\mathcal{K}(u, v) - u - v + 1$ between the two child schedules, producing

$$\left(a^\top,\ \mathcal{K}(u, v) - u - v + 1,\ b^\top\right)^\top, \tag{3}$$

whose scalar value is $\mathcal{K}(u, v)$. Thus, on any fixed composition tree, ConPP and OBS-S generate the same numerical schedule. Appendix A gives the underlying certificate formulations and the change of variables establishing this correspondence.

Each internal node corresponds to one join and therefore inserts exactly one gradient step. Since a full binary tree with N leaves has $N - 1$ internal nodes, an ordered N-leaf composition tree represents an $(N - 1)$-step schedule, matching the convention $N = n + 1$ introduced in Section 1. Starting from value 1 at every leaf, recursively applying $\mathcal{K}$ determines the scalar value at each internal node; together with equation 3, these values determine the inserted steps and hence the resulting numerical schedule. For example, joining two leaves gives the one-step schedule $(\sqrt{2})$, while joining two copies of this block gives $(\sqrt{2}, 2, \sqrt{2})$. Maximizing the common scalar value within the symmetric recursive families is therefore equivalent to maximizing the root value over N-leaf composition trees.

The Bellman problem. Let U_N be the largest root value over all composition trees with N leaves. Since $\mathcal{K}$ is strictly increasing in each argument, the two subtrees of an optimal tree may themselves be chosen optimally. Hence

$$U_1 = 1, \qquad U_N = \max_{1 \leq m < N} \mathcal{K}(U_m, U_{N-m}) \quad (N \geq 2). \tag{4}$$

Here m and $N - m$ are the leaf counts of the two child blocks, while the root contributes their joining step. Equation 4 is the scalar optimization problem underlying both symmetric recursive families.

Exact objective interpretation. Let H_N denote the maximum total stepsize in the $(N - 1)$-step ConPP family, and let s_N denote the optimal OBS-S rate at the same horizon. Their relation to the common scalar value is

$$U_N = H_N + 1 = \frac{1}{s_N}. \tag{5}$$

Since the tight objective coefficients for ConPP and OBS-S are $1/(2H_N + 1)$ and $s_N/(2 - s_N)$, respectively, equation 5 shows that both families have the same tight objective coefficient

$$B_N := \frac{1}{2U_N - 1},$$

and their optimized $(N - 1)$-step schedules satisfy

$$f(x_{N-1}) - f_* \leq B_N \frac{L}{2} \|x_0 - x_*\|^2. \tag{6}$$

A one-dimensional Huber instance attains equality in this bound for each family; the corresponding certificate derivations are collected in Appendix A. Thus B_N is their exact worst-case objective coefficient, and maximizing U_N is exactly equivalent to minimizing the worst-case objective gap.

It remains, therefore, to understand the maximization defining U_N in equation 4. For this structural question, the key properties of $\mathcal{K}$ are symmetry, positive homogeneity, and strict coordinatewise monotonicity, together with

$$\mathcal{K}(0, u) = u, \qquad \mathcal{K}(u, u) = \rho u, \qquad \rho = 1 + \sqrt{2}.$$

The central question is whether the balanced split always attains the maximum in equation 4, and when other splits can tie.

3 BALANCED OPTIMALITY AND ALL OPTIMAL COMPOSITIONS TREES

The Bellman recursion 4 optimizes over every root split, whereas the Silver construction uses equal children at dyadic horizons. The next theorem shows that the same balancing principle governs every horizon: splitting the leaf count as evenly as possible always attains the optimum. It further characterizes exactly when an unbalanced split attains the same value.

Theorem 3.1. *Extend the sequence by $U_0 = 0$. For every $N \geq 2$,*

$$U_N = \mathcal{K}\big(U_{\lfloor N/2 \rfloor}, U_{\lceil N/2 \rceil}\big). \tag{7}$$

Moreover, $(U_N)_{N \geq 0}$ is strictly increasing and satisfies the supercomposition inequality

$$\mathcal{K}(U_m, U_n) \leq U_{m+n}, \qquad m, n \geq 0.$$

For positive integers m, n, equality holds exactly when

$$\mathcal{K}(U_m, U_n) = U_{m+n} \quad \Longleftrightarrow \quad \begin{cases} m = n, & \text{or} \\ \{m, n\} = \{2^k r, 2^k(r+1)\} \text{ for some } k \geq 0,\ r \geq 1. \end{cases} \tag{8}$$

This resolves Conjecture 4.17 of Zhang & Jiang (2026) and completely characterizes the maximizing root splits. In particular, every unbalanced tie is a dyadic rescaling of two consecutive leaf counts. For example, at $N = 6$, the balanced split $3 + 3$ ties with the two ordered splits $2 + 4$ and $4 + 2$. If N is a power of two, the equal split is the unique maximizing root split; recursively, the completely balanced ordered tree is therefore the unique optimal composition tree. More generally, an ordered composition tree is optimal if and only if every internal split satisfies one of the two patterns in equation 8.

Appendix C gives the complete root-split classification, translates it to the original step-count notation of Zhang & Jiang (2026), and counts the resulting optimal ordered trees; distinct optimal trees may still generate the same numerical schedule.

3.1 A FOUR-POINT REARRANGEMENT PRINCIPLE

The obstacle is to compare compositions with different tree shapes, something symmetry alone does not do. We overcome it by examining two levels of a tree at once: four subtrees can be paired in three ways before their values are combined at the parent. The following inequality orders these three possibilities and, crucially, retains the equality information. For $a, b, c, d \geq 0$, write $Z_{ab|cd} = \mathcal{K}(\mathcal{K}(a, b), \mathcal{K}(c, d))$.

Theorem 3.2. *For $0 \leq a \leq b \leq c \leq d$,*

$$Z_{ab|cd} \leq Z_{ac|bd} \leq Z_{ad|bc}. \tag{9}$$

The equality cases are

$$Z_{ab|cd} = Z_{ac|bd} \iff b = c \text{ or } ad = bc, \tag{10}$$

$$Z_{ac|bd} = Z_{ad|bc} \iff a = b \text{ or } c = d. \tag{11}$$

The proof exploits the quadratic identity

$$\mathcal{K}(x, y)^2 - (x + y)\mathcal{K}(x, y) - xy = 0.$$

For $u = \mathcal{K}(a, b)$ and $v = \mathcal{K}(c, d)$, eliminating the inner roots reduces the first comparison to the sign of

$$abv^2 + cdu^2 - (ad + bc)uv = (av - cu)(bv - du).$$

Homogeneity gives

$$av = \mathcal{K}(ac, ad), \qquad cu = \mathcal{K}(ac, bc),$$

and similarly

$$bv = \mathcal{K}(bc, bd), \qquad du = \mathcal{K}(ad, bd).$$

Monotonicity therefore makes the two factors have opposite signs. The same factorization identifies the equality case $ad = bc$, while $b = c$ corresponds to coincident pairings. The second comparison follows analogously. Appendix B gives the complete elimination and root-sign argument, including boundary cases. This turns a comparison of nested compositions into a local exchange rule that can be propagated through the tree.

From a local comparison to global optimality. To propagate the four-point rearrangement through arbitrary trees, define the balanced candidate sequence

$$V_0 = 0, \qquad V_1 = 1, \qquad V_N = \mathcal{K}(V_{\lfloor N/2 \rfloor}, V_{\lceil N/2 \rceil}).$$

Monotonicity of $\mathcal{K}$ makes (V_N) strictly increasing. The key inductive invariant is the supercomposition inequality $\mathcal{K}(V_m, V_n) \le V_{m+n}$. Once this inequality holds, every Bellman split is bounded by V_N, while the balanced split attains V_N; hence $U_N = V_N$.

For example, in the even–odd case $1 \le a \le b$,

$$\begin{aligned} \mathcal{K}(V_{2a}, V_{2b+1}) &= \mathcal{K}\big(\mathcal{K}(V_a, V_a), \mathcal{K}(V_b, V_{b+1})\big) \\ &\le \mathcal{K}\big(\mathcal{K}(V_a, V_b), \mathcal{K}(V_a, V_{b+1})\big) && \text{(rearrangement)} \\ &\le \mathcal{K}(V_{a+b}, V_{a+b+1}) && \text{(induction)} \\ &= V_{2a+2b+1}. \end{aligned}$$

The remaining parity cases use the same two ingredients: homogeneity handles even–even indices, while the full rearrangement chain handles unequal odd indices. Thus the supercomposition inequality holds for all m, n, and the preceding argument gives $U_N = V_N$ for every N.

Tracking equality through the same parity recursion yields exactly the dyadic-consecutive pattern in equation 8: common powers of two peel off, mixed parity permits equality only for consecutive indices, and unequal odd indices are strict. Strict monotonicity then forces both children of every optimal node to be optimal for their respective sizes, giving the complete tree characterization in Theorem 3.1. Appendix B.5 gives the abstract balanced-composition induction, and Appendix C completes the equality and optimal-tree classification.

The balanced recurrence therefore replaces the exhaustive Bellman search by a deterministic binary structure, which we exploit next to evaluate U_N without tabulating all smaller horizons.

3.2 Binary recursion and single-horizon evaluation

Theorem 3.1 reduces the exhaustive Bellman search to the balanced recurrence, so the table $U_1, \dots, U_N$ can be computed in $O(N)$ scalar operations. For a single prescribed horizon, the recurrence has a stronger binary structure. Indeed,

$$(U_{2r}, U_{2r+1}) = (\rho U_r, \mathcal{K}(U_r, U_{r+1})), \quad (U_{2r+1}, U_{2r+2}) = (\mathcal{K}(U_r, U_{r+1}), \rho U_{r+1}).$$

Thus an adjacent pair (U_r, U_{r+1}) determines the pair associated with either binary extension $r \mapsto 2r$ or $r \mapsto 2r+1$. Starting from $(U_0, U_1) = (0, 1)$, read the binary digits of N from left to right. After a prefix representing r has been read, maintain $(a, b) = (U_r, U_{r+1})$. A next bit 0 sends this pair to (U_{2r}, U_{2r+1}), while a next bit 1 sends it to (U_{2r+1}, U_{2r+2}). This gives the following evaluator.

Algorithm 1: Binary evaluation of U_N

Input: An integer $N \ge 1$, read in binary from the most significant bit.

$(a, b) \leftarrow (0, 1)$
for *each binary digit* ξ *of* N **do**
  $c \leftarrow \mathcal{K}(a, b)$
  if $\xi = 0$ **then** $(a, b) \leftarrow (\rho a, c)$
  else $(a, b) \leftarrow (c, \rho b)$
end
return a

Each binary digit requires one evaluation of $\mathcal{K}$ and one multiplication by ρ. Hence Algorithm 1 evaluates a single optimized scalar value U_N in $O(\log N)$ scalar operations using $O(1)$ scalar working storage. Constructing the full $(N-1)$-step schedule requires linear output work. Appendix C.3 proves the adjacent-pair invariant and the operation count.

The same binary recursion also reveals the exact scaling $U_{2r} = \rho U_r$. Its iteration is the dyadic self-similarity underlying the phase law developed next.

4 EXACT DYADIC PHASE LAW FOR THE SYMMETRIC FAMILY

The exact scaling from Section 3.2 iterates to

$$U_{2^k N} = \rho^k U_N, \qquad s_{2^k N} = \rho^{-k} s_N, \qquad N \geq 1,\ k \geq 0. \tag{12}$$

Since $\rho = 2^p$, the normalized value U_N/N^p is invariant under dyadic dilation. Hence its remaining dependence on the horizon is determined entirely by the position of N within its dyadic block. Write $x_N := N/2^{\lfloor \log_2 N \rfloor} \in [1, 2)$, $\theta_N := \log_2 x_N = \{\log_2 N\} \in [0, 1)$, and identify θ_N with its class in $\mathbb{T} := \mathbb{R}/\mathbb{Z}$. Here x_N is the dyadic mantissa, and θ_N is the corresponding logarithmic phase. We write $\mathrm{Clust}(a_N)$ for the set of subsequential limits of a sequence $\{a_N\}$ along $N \to \infty$. The next theorem shows that all normalized optimal values lie exactly on a single continuous profile of the dyadic phase.

Theorem 4.1. *There exists a unique positive Lipschitz function $\Phi : \mathbb{T} \to (0, \infty)$ such that, for every $N \geq 1$,*

$$U_N = N^p \Phi(\theta_N). \tag{13}$$

Moreover, Φ is nonconstant and satisfies

$$0 < \phi_* := \min_{t \in \mathbb{T}} \Phi(t) < 1, \qquad \max_{t \in \mathbb{T}} \Phi(t) = \Phi(0) = 1. \tag{14}$$

Through the identity $B_N = (2U_N - 1)^{-1}$, the exact phase law for U_N induces an exact phase law for the tight objective coefficient.

Corollary 4.2. *The optimized symmetric schedules satisfy*

$$B_N = \frac{1}{2N^p \Phi(\theta_N) - 1}. \tag{15}$$

Consequently,

$$B_N = \frac{N^{-p}}{2\Phi(\theta_N)} + O(N^{-2p}), \tag{16}$$

where the remainder is uniform over the dyadic phase, and

$$\mathrm{Clust}(N^p B_N) = \left[\frac{1}{2}, \frac{1}{2\phi_*}\right]. \tag{17}$$

Thus the dyadic phase enters the leading N^{-p} coefficient itself. In particular, the normalized tight objective bound has no single asymptotic constant. Two fixed dyadic rays already exhibit distinct limits:

$$\lim_{k \to \infty} (2^k)^p B_{2^k} = \frac{1}{2}, \qquad \lim_{k \to \infty} (3 \cdot 2^k)^p B_{3 \cdot 2^k} = \frac{3^p}{2U_3} > \frac{1}{2}.$$

Continuity of Φ then fills the entire interval in equation 17.

From dyadic self-similarity to a phase profile. Equation 12 collapses all horizons related by powers of two onto the single reference block $[1, 2]$. For example, the horizons $3, 6, 12, \ldots$ all have dyadic mantissa $3/2$ and satisfy $\rho^{-1}U_3 = \rho^{-2}U_6 = \rho^{-3}U_{12} = \cdots$. More generally, for a dyadic point $x = j/2^k \in [1, 2]$, define $F(x) := \rho^{-k}U_j$. The scaling $U_{2j} = \rho U_j$ makes this definition independent of the representation of x.

The balanced recurrence is compatible with dyadic refinement. If $x = j/2^k$ and $y = (j+1)/2^k$ are adjacent grid points, their midpoint corresponds to the odd index $2j + 1$, and

$$F\left(\frac{x+y}{2}\right) = \rho^{-1}\mathcal{K}(F(x), F(y)).$$

Thus the same recurrence recursively fills in the values of F on finer dyadic grids.

From the dyadic grid to a continuous phase. The adjacent ratios $R_N := U_{N+1}/U_N$ satisfy a binary recursion contracting toward 1, and in particular $R_N - 1 = O(N^{-1})$ (Lemma D.3). Lemma D.4 converts this estimate into uniform upper and lower bounds on the slopes of the dyadic piecewise-linear interpolants. Hence F extends uniquely to a strictly increasing bi-Lipschitz map $F : [1, 2] \to [1, \rho]$.

Writing $x = 2^t$, we have $x^p = \rho^t$, and therefore $U_N/N^p = F(x_N)/x_N^p$. This gives the phase profile

$$\Phi(t) := \rho^{-t} F(2^t), \qquad 0 \le t \le 1.$$

Since $F(1) = 1$ and $F(2) = \rho$, we have $\Phi(0) = \Phi(1) = 1$, so Φ extends continuously and periodically to $\mathbb{T}$.

Why the phase is nonconstant. Let $q = 1/p$. The sharp kernel inequality $\mathcal{K}(x, y) \le (x^q + y^q)^{1/q}$ implies $U_N \le N^p$, with equality exactly at powers of two. Hence $\Phi(0) = 1$, whereas $\Phi(\log_2(3/2)) = U_3/3^p < 1$. Thus Φ is nonconstant. Uniqueness follows from the density of the integer phases in $\mathbb{T}$.

This proves Theorem 4.1. Corollary 4.2 follows from $B_N = (2U_N - 1)^{-1}$; Appendix D.5 gives the uniform expansion and the complete cluster interval.

5 Limiting Dyadic phase of the asymmetric Recursion

The ConPD/OBS-F recursion has a different geometry. Each join combines an optimized symmetric prefix with an objective-specific continuation, and the optimal split may be unbalanced. This breaks the exact doubling identity that underpins the symmetric phase law in Section 4. Nevertheless, after normalization by N^p, the tight objective coefficients converge uniformly across dyadic blocks to a nonconstant phase profile. The proof proceeds by comparing the ordered recursion with its symmetric counterpart and then analyzing the equality structure of a sharp homogeneous support bound.

5.1 Ordered Bellman Recursion and Symmetric Comparison

Let η_N^F denote the tight objective-gap coefficient of the optimized ConPD/OBS-F schedule with $N - 1$ steps, and set $W_N := (\eta_N^F)^{-1}$. The composition rules underlying ConPD and OBS-F yield the ordered Bellman recursion

$$W_1 = 1, \qquad W_N = \max_{1 \le m < N} \mathcal{A}(U_m, W_{N-m}), \tag{18}$$

where $\mathcal{A}(u, w) = (4u + w + \sqrt{w^2 + 8uw})/2$. Here U_m represents an optimized symmetric prefix with m leaves, whereas W_{N-m} represents the optimized objective-specific continuation. The two arguments therefore play different roles: the recursion is ordered rather than symmetric, placing it outside the four-point rearrangement mechanism of Section 3. Appendix A.4 gives the certificate derivation and the correspondence between the ConPD and OBS-F constructions.

The corresponding optimized schedule satisfies the tight objective guarantee

$$f(x_{N-1}) - f_* \le \eta_N^{\mathrm{F}} \frac{L}{2} \|x_0 - x_*\|^2 . \tag{19}$$

Matching one-dimensional quadratic and Huber instances attain equality, so η_N^F is the exact worst-case objective coefficient for the optimized recursive schedule; see Appendix A.4.

Despite their different kernels, the symmetric and asymmetric recursions remain quantitatively coupled. Local comparisons between $\mathcal{A}$ and $\mathcal{K}$, propagated through the Bellman recursions, give

$$2U_N - 1 \le W_N \le \rho U_N. \tag{20}$$

Equivalently, $1/(\rho U_N) \le \eta_N^F \le 1/(2U_N - 1) = B_N$. Hence the asymmetric objective coefficient is at most the symmetric one at every horizon. Together with the exact symmetric phase law $U_N = N^p \Phi(\theta_N)$, equation 20 gives $\eta_N^F = \Theta(N^{-p})$. The exponent is therefore fixed; what remains is the normalized coefficient $C_N := N^p \eta_N^F$. We next show that C_N retains a nonconstant dyadic-phase dependence, converging uniformly over dyadic blocks to a limiting profile.

5.2 UNIFORM PHASE LIMIT AND STRICT SUPPORT SEPARATION

For each dyadic block $2^k \le N \le 2^{k+1}$, let $\Psi_k : [0,1] \to \mathbb{R}$ be the piecewise-affine interpolation of the points $(\log_2(N/2^k), C_N)$. Thus every dyadic block is represented by a curve on the same phase domain. In the symmetric setting, exact doubling makes the corresponding grid values consistent across scales. Here the interpolants vary with k, and the central question is whether they converge to a common profile.

A second ingredient is the sharp homogeneous support bound for the asymmetric kernel. With $q = 1/p$, define

$$B_{\sup} \coloneqq \sup_{u>0} \frac{\mathcal{A}(u,1)^q - 1}{u^q}, \qquad c_* \coloneqq B_{\sup}^{-p}.$$

By homogeneity, $\mathcal{A}(u,w)^q \le B_{\sup} u^q + w^q$, and combining this with the symmetric power bound $U_m^q \le m$ gives $C_N \ge c_*$. The key question is whether an optimizing recursion can asymptotically attain this local barrier.

Theorem 5.1. *There is a unique positive Lipschitz profile $\Psi_F : \mathbb{T} \to (0,\infty)$ such that $\|\Psi_k - \Psi_F\|_\infty \to 0$. Equivalently,*

$$\sup_{2^k \le N \le 2^{k+1}} |N^p \eta_N^{\mathrm{F}} - \Psi_F(\theta_N)| \longrightarrow 0. \tag{21}$$

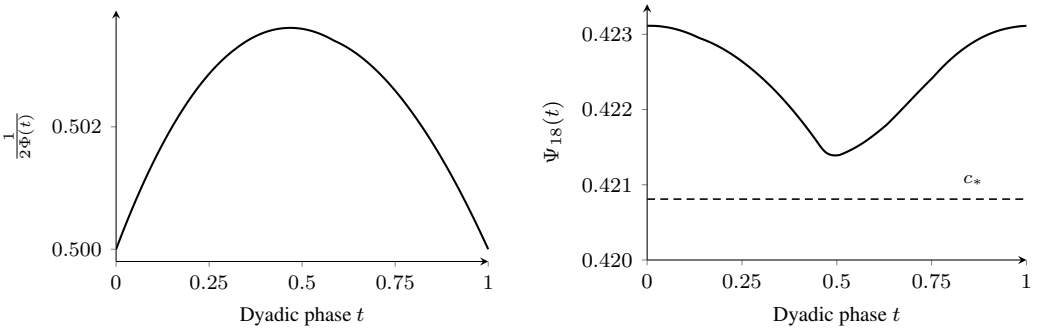
Hence $\eta_N^{\mathrm{F}} = N^{-p}\Psi_F(\theta_N) + o(N^{-p})$, with this dyadic-block uniformity, and

$$\mathrm{Clust}(N^p \eta_N^{\mathrm{F}}) = [\min \Psi_F, \max \Psi_F]. \tag{22}$$

The profile is nonconstant and satisfies

$$c_* < \min \Psi_F < \max \Psi_F \le c_*/\phi_*. \tag{23}$$

Equation 21 is the asymmetric counterpart of the exact phase law in equation 13: every dyadic block approaches the same limiting profile uniformly, although the optimizing splits may vary with the horizon. Thus both optimized recursive families share the Silver exponent p, while their leading objective coefficients retain nonconstant dyadic-phase dependence. Moreover, the strict inequality $c_* < \min \Psi_F$ shows that global recursive compatibility enforces a positive gap above the sharp local support barrier. Figure 1 visualizes the two objective-coefficient phases.



(a) Symmetric: exact phase law

(b) Asymmetric: finite-scale interpolant

Figure 1: Dyadic phase dependence of the objective-gap leading coefficients. Left: the symmetric leading-coefficient phase $1/(2\Phi(t))$, given exactly by Corollary 4.2. Right: the finite-scale asymmetric interpolant $\Psi_{18}(t)$; Theorem 5.1 establishes the uniform limit $\Psi_k \to \Psi_F$. The dashed line marks the sharp support barrier c_*.

Proof sketch. *Uniform phase limit.* For each fixed m, the inequality $\eta_{2N}^F \le \rho^{-1}\eta_N^F$ makes $C_{m2^k} = (m2^k)^p \eta_{m2^k}^F$ nonincreasing in k, while equation 20 and $U_N \le N^p$ give $C_{m2^k} \ge 1/\rho$. Hence every fixed dyadic ray has a positive limit. Moreover, $|C_{N+1} - C_N| = O(N^{-1})$, while adjacent logarithmic phases are separated by $\Theta(N^{-1})$, so the dyadic-block interpolants $\{\Psi_k\}$ are equi-Lipschitz. Thus every subsequence has a uniformly convergent subsequence. Any such limit

is already determined on the dense set of dyadic phases by the raywise limits, so all subsequential limits coincide. Therefore $\Psi_k \to \Psi_F$ uniformly.

Strict separation from c_*. We next show that Ψ_F stays strictly above the support barrier $c_* = B_{\sup}^{-p}$. Suppose instead that $\Psi_F(t_0) = c_*$. Choose $N_k \to \infty$ with $\{\log_2 N_k\} \to t_0$. Uniform phase convergence gives $N_k^p \eta_{N_k}^F \to c_*$, equivalently $E_{N_k} := B_{\sup} N_k - W_{N_k}^q = o(N_k)$.

For every optimal split $N = m + r$,

$$E_N = B_{\sup}(m - U_m^q) + E_r + \Big(B_{\sup} U_m^q + W_r^q - \mathcal{A}(U_m, W_r)^q \Big),$$

and all three terms are nonnegative. Thus near-contact with c_* requires the symmetric bound, the continuation bound, and the sharp support inequality to become asymptotically tight together.

A concentration argument based on this deficit identity shows that near-contact forces an optimal prefix m_k to satisfy $m_k/N_k \to \alpha_*$, where $\alpha_* \in (1/2, 3/4)$ is the unique fraction selected by equality in the sharp support inequality. The same near-contact also forces the symmetric bound to be asymptotically tight: $U_{m_k}^q / m_k \to 1$. By the exact symmetric phase law,

$$\frac{U_{m_k}^q}{m_k} = \Phi(\{\log_2 m_k\})^q,$$

and $\Phi(t) = 1$ only at phase 0. Hence $\{\log_2 m_k\} \to 0$. Since $\log_2 m_k = \log_2 N_k + \log_2(m_k/N_k)$, while $\{\log_2 N_k\} \to t_0$ and $m_k/N_k \to \alpha_*$, we obtain

$$t_0 + \log_2 \alpha_* = 0 \pmod 1.$$

The continuation satisfies $r_k/N_k \to 1-\alpha_*$ and still has vanishing relative deficit. Its phase therefore tends to $t_1 = t_0 + \log_2(1-\alpha_*) \pmod 1$. Applying the same argument to this continuation gives

$$t_1 + \log_2 \alpha_* = 0 \pmod 1.$$

Subtracting the two phase identities yields

$$\log_2(1-\alpha_*) \in \mathbb{Z}, \qquad 1/4 < 1 - \alpha_* < 1/2,$$

a contradiction. Therefore $\Psi_F(t) > c_*$ for every $t \in \mathbb{T}$.

Nonconstancy and upper enclosure. Finally, if Ψ_F were constant, the limiting Bellman inequality would become scale-invariant, and the sharp support geometry would force the only possible constant to be c_*, contradicting the strict separation above. The upper enclosure $\Psi_F \le c_*/\phi_*$ follows from the same limiting Bellman inequality together with the symmetric lower bound $U_m/m^p \ge \phi_*$.

Thus the asymmetric recursion changes the leading objective coefficient while retaining a nonconstant dyadic dependence.

6 Discussion and conclusion

Optimized recursive long-step schedules therefore admit a sharper description than the usual $O(N^{-p})$ rate. For the symmetric ConPP/OBS-S family, the four-point rearrangement principle solves the Bellman split problem at every horizon: balanced splitting is optimal, all optimal ties are characterized, and the resulting recursion yields exact dyadic self-similarity together with an $O(\log N)$ evaluator. This self-similarity induces an exact log-periodic law for the tight objective coefficient, in which the Silver exponent $p = \log_2(1+\sqrt{2})$ determines the power-law envelope while a nonconstant phase function captures the leading dependence on the dyadic position of N. For the asymmetric ConPD/OBS-F family, horizon-dependent unbalanced optimizers produce a distinct limiting phase profile, while the same Silver exponent persists and the leading coefficient remains strictly above the sharp local support barrier.

Taken together, these results show that the recursive mechanism underlying accelerated long-step schedules also imprints a persistent dyadic structure on their convergence guarantees: the exponent determines the scale, while the phase determines the leading finite-horizon coefficient. This exponent–phase description sharpens the usual rate-based view of predetermined stepsize acceleration and suggests a broader question of whether analogous phase phenomena arise beyond recursive constructions. Extending this viewpoint to minimax-optimal predetermined schedules, and resolving the exact geometry of the asymmetric phase together with the asymptotic structure of its optimizing splits, are natural directions for future work.

AI USE STATEMENT

The authors identified the phase-dependent variation of the leading coefficients that motivated this work, formulated the underlying mathematical questions and structural ideas, and led the development and presentation of the results. The authors substantially revised, simplified, reorganized, and verified the proofs, and carefully refined the manuscript. Generative AI tools, including ChatGPT, Codex, and AI-assisted features in Cursor, were used under the authors' direction to propose candidate proof strategies and proof arguments, assist with Lean 4 tactics and declarations, and help implement numerical checking code. AI-generated mathematical arguments were critically examined, modified where needed, and independently verified by the authors before inclusion. All submitted Lean and numerical code was reviewed and executed by the authors. The authors take responsibility for the final text, mathematical claims, and submitted artifacts.

REPRODUCIBILITY STATEMENT

The paper is primarily theoretical. The appendix proof roadmap locates the complete mathematical proofs, and Appendix A states the imported certificate interfaces and prior results. The supplementary materials provide the scalar evaluator, reference Bellman recursions, figure data, and reproduction commands. `NUMERICAL_REPRODUCIBILITY.md` distinguishes exploratory floating-point computations from rigorous interval-arithmetic certificates. The phase-minimum and dyadic-corner certificates are specified in Appendices D.8 and D.7, respectively, and the reported level-17 phase-minimum enclosure is verified by interval arithmetic. The supplementary materials also include Lean 4 proofs of the scalar and combinatorial results identified in `lean/docs/THEOREMS.md`. `LEAN_FORMALIZATION.md` documents the pinned environment, verification commands, and formalization scope. The smooth-convex gradient-descent certificate interfaces and tightness results imported from prior work are documented in Appendix A and remain outside the Lean formalization.

REFERENCES

J. M. Altschuler and P. A. Parrilo. Acceleration by stepsize hedging: Multi-step descent and the Silver stepsize schedule. *Journal of the ACM*, 72(2):Article 12, 2025a. doi: 10.1145/3708502.

J. M. Altschuler and P. A. Parrilo. Acceleration by stepsize hedging: Silver stepsize schedule for smooth convex optimization. *Mathematical Programming*, 213:1105–1118, 2025b. doi: 10.1007/s10107-024-02164-2.

S. Das Gupta, B. P. G. Van Parys, and E. K. Ryu. Branch-and-bound performance estimation programming: A unified methodology for constructing optimal optimization methods. *Mathematical Programming*, 204:567–639, 2024. doi: 10.1007/s10107-023-01973-1.

Y. Drori and M. Teboulle. Performance of first-order methods for smooth convex minimization: A novel approach. *Mathematical Programming*, 145:451–482, 2014. doi: 10.1007/s10107-013-0653-0.

B. Grimmer. Provably faster gradient descent via long steps. *SIAM Journal on Optimization*, 34(3): 2588–2608, 2024. doi: 10.1137/23M1588408.

B. Grimmer, K. Shu, and A. L. Wang. Accelerated gradient descent via long steps, 2023. URL `https://arxiv.org/abs/2309.09961`. arXiv preprint arXiv:2309.09961.

B. Grimmer, K. Shu, and A. L. Wang. Accelerated objective gap and gradient norm convergence for gradient descent via long steps. *INFORMS Journal on Optimization*, 7:156–169, 2025a. doi: 10.1287/ijoo.2024.0057.

B. Grimmer, K. Shu, and A. L. Wang. Composing optimized stepsize schedules for gradient descent. *Mathematics of Operations Research*, 2025b. doi: 10.1287/moor.2024.0764. Published online 4 November 2025.

G. Kornowski and O. Shamir. Open problem: Anytime convergence rate of gradient descent. In *Proceedings of the 37th Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pp. 5335–5339. PMLR, 2024. URL https://proceedings.mlr.press/v247/kornowski24a.html.

A. B. Taylor, J. M. Hendrickx, and F. Glineur. Smooth strongly convex interpolation and exact worst-case performance of first-order methods. *Mathematical Programming*, 161:307–345, 2017. doi: 10.1007/s10107-016-1009-3.

M. Teboulle and Y. Vaisbourd. An elementary approach to tight worst case complexity analysis of gradient based methods. *Mathematical Programming*, 201:63–96, 2023. doi: 10.1007/s10107-022-01899-0.

C.-E. Tsai, I. Fatkhullin, L. Zhang, and N. He. Lower bounds for anytime acceleration of gradient descent, 2026. URL https://arxiv.org/abs/2607.02053.

Yuhan Ye and Kaizhao Liu. Silver rate is (almost) optimal for gradient descent acceleration, 2026. URL https://arxiv.org/abs/2609.09152.

Z. Zhang and R. Jiang. Accelerated gradient descent by concatenation of stepsize schedules. *SIAM Journal on Optimization*, 36(2):1182–1210, 2026. doi: 10.1137/25M173898X.

Z. Zhang, J. D. Lee, S. S. Du, and Y. Chen. Anytime acceleration of gradient descent. In *Proceedings of COLT 2025*, volume 291 of *Proceedings of Machine Learning Research*, pp. 5991–6013. PMLR, 2025. URL https://proceedings.mlr.press/v291/zhang25a.html.

APPENDIX: PROOF ROADMAP

The main text states the structural and phase results and explains their proof mechanisms. The appendices retain all supporting certificate definitions, intermediate results, and complete proofs.

Main-text material	Supporting material and proof ingredients
Section 2: schedules and scalar values	Appendix A: primitive and s-composable certificates, joins, and objective guarantees.
Theorem 3.1: optimal compositions	Appendices B.5 and C: propagation, equality, and optimal trees.
Theorem 3.2: rearrangement	Appendix B: sign identities and algebraic comparison.
Algorithm 1: one-horizon evaluation	Appendix C.3: adjacent-pair invariant and operation count.
Theorem 4.1: exact symmetric phase	Appendices D.1–D.4: power envelope, ratio contraction, and canonical interpolation.
Corollary 4.2: tight objective coefficient	Appendix D.5: exact conversion, uniform expansion, and cluster set.
Theorem 5.1: asymmetric phase	Appendix E: sandwich, compactness, support deficit, and phase rigidity.

A CERTIFICATE FRAMEWORKS AND SCALAR RECURSIONS

The composition theorems used in the main text are due to Zhang & Jiang (2026) and Grimmer et al. (2025b). We record their certificate interfaces, the changes of variables, and the resulting connection to the gradient-descent guarantees. Throughout this appendix, $L = 1$; applying the dimensionless schedules through equation 1 restores the factor L.

A.1 CERTIFICATE INTERFACES AND RECURSIVE SCHEDULE CLASSES

Let $\mathcal{F}_1$ be the class of differentiable convex functions with 1-Lipschitz gradient and a nonempty minimizer set. For a schedule $h = (h_0, \dots, h_{n-1})^\top$ and the associated gradient-descent trajectory,

write $f_i = f(x_i)$ and $g_i = \nabla f(x_i)$, and fix a minimizer x_* with $f_* = f(x_*)$. Smooth convex interpolation (Taylor et al., 2017) gives

$$Q_{i,*} \coloneqq f_i - f_* - \langle g_i, x_i - x_* \rangle + \frac{1}{2} \|g_i\|^2 \leq 0.$$

The two frameworks below use stronger inequalities involving the final iterate to certify recursive composition.

A.1.1 PRIMITIVE SCHEDULES AND CONPP

We first recall the definition of a primitive schedule from Zhang & Jiang (2026).

Definition A.1 (Definition 2.7 in Zhang & Jiang (2026)). *A schedule $h \in \mathbb{R}_+^n$ is primitive, with $H \coloneqq \mathbf{1}^\top h$, every $f \in \mathcal{F}_1$ and every associated trajectory satisfy*

$$\frac{1}{2} \|x_0 - x_*\|^2 + \sum_{i=0}^{n-1} h_i Q_{i,*} - H(f_n - f_*) \geq \frac{1}{2} \|x_n - x_*\|^2 + \frac{H(H+1)}{2} \|g_n\|^2 .$$

The empty schedule is primitive with $H = 0$.

This certificate also yields the tight objective bound

$$f_n - f_* \leq \frac{1}{2H+1} \cdot \frac{1}{2} \|x_0 - x_*\|^2 , \tag{24}$$

with equality on a suitable one-dimensional Huber instance (Zhang & Jiang, 2026, Theorem 2.6, Proposition 2.8, and Corollary 2.9). Thus, within this certificate class, maximizing the total stepsize H is exactly equivalent to minimizing the certified objective coefficient.

If a and b are primitive, then their primitive–primitive concatenation

$$\mathrm{ConPP}(a, b) \coloneqq \left[a^\top, \varphi\big(\mathbf{1}^\top a, \mathbf{1}^\top b\big), b^\top\right]^\top$$

is primitive, where

$$\varphi(x, y) \coloneqq \frac{-x - y + \sqrt{(x+y+2)^2 + 4(x+1)(y+1)}}{2}.$$

This is the ConPP theorem of Zhang & Jiang (2026, Theorem 3.1). In particular, both the inserted step and the total stepsize of the joined schedule depend on each input schedule only through its total stepsize. We now record the optimized family in the notation of Zhang & Jiang (2026, Definition 4.1).

Definition A.2 (Definition 4.1 in Zhang & Jiang (2026)). *Set $h_\circ^{(0)} = [\,]^\top$. For $n \geq 1$, choose*

$$h_\circ^{(n)} \in \operatorname*{arg\,max}_{h \in \mathcal{H}_\circ^{(n)}} \mathbf{1}^\top h, \quad \mathcal{H}_\circ^{(n)} \coloneqq \left\{ \mathrm{ConPP}\Big(h_\circ^{(k)}, h_\circ^{(n-k-1)}\Big) : 0 \leq k < n \right\}.$$

If the maximizer is not unique, any maximizing schedule may be selected.

Since $x + y + \varphi(x, y)$ is strictly increasing in each argument, this recursion implements the principle of optimality: $h_\circ^{(n)}$ has maximum total stepsize among all n-step schedules generated recursively from the empty schedule by ConPP.

A.1.2 S-COMPOSABLE SCHEDULES AND OBS-S

The s-composable framework of Grimmer et al. (2025b) uses a rate parameter η whose defining certificate is attained by canonical quadratic and Huber instances. Its certificate is symmetric in the sense that it controls the final objective gap, distance, and gradient norm through the same parameter.

Definition A.3 (Definition 3 in Grimmer et al. (2025b)). *A positive schedule $h = (h_0, \ldots, h_{n-1})^\top$ is s-composable with rate $\eta > 0$ if for every $f \in \mathcal{F}_1$ and every associated trajectory, it holds that*

$$\frac{1-\eta}{2} \|g_n\|^2 + \frac{\eta^2}{2} \|x_n - x_*\|^2 + (\eta - \eta^2)(f_n - f_*) \leq \frac{\eta^2}{2} \|x_0 - x_*\|^2 ,$$

and

$$\eta = \frac{1}{1 + \sum_{i=0}^{n-1} h_i} = \prod_{i=0}^{n-1} (h_i - 1).$$

The sum–product identity determines η from the total stepsize. The corresponding tight objective bound is

$$f_n - f_* \leq \frac{\eta}{2 - \eta} \cdot \frac{1}{2} \|x_0 - x_*\|^2, \tag{25}$$

with equality attained by a one-dimensional Huber instance (Grimmer et al., 2025b, Proposition 1). Hence minimizing η strengthens the certified objective guarantee. Grimmer et al. (2025b, Proposition 1) also gives a stronger inequality, attained by the quadratic instance, that contains an additional negative squared-norm term.

If a and b are s-composable with rates α and β, respectively, their s-join is

$$a \bowtie b \coloneqq \left[a^\top, \mu(\alpha, \beta), b^\top\right]^\top, \quad \mu(\alpha, \beta) \coloneqq 1 + \frac{\sqrt{\alpha^2 + 6\alpha\beta + \beta^2} - (\alpha + \beta)}{2\alpha\beta}.$$

The joined schedule is s-composable with rate

$$\mathcal{J}(\alpha, \beta) \coloneqq \frac{2\alpha\beta}{\alpha + \beta + \sqrt{\alpha^2 + 6\alpha\beta + \beta^2}}$$

by Grimmer et al. (2025b, Theorem 3). Repeated s-joins from the empty schedule generate the basic s-composable family.

Definition A.4 (Optimized Basic s-composable Schedule (OBS-S) in Grimmer et al. (2025b))**.** *Set* $h^{\text{OBS-S}}(0) = [\,]^\top$ *and* $\eta^{\text{OBS-S}}(0) = 1$*. For* $N \geq 2$*, define*

$$\eta^{\text{OBS-S}}(N-1) \coloneqq \min_{1 \leq m < N} \mathcal{J}\left(\eta^{\text{OBS-S}}(m-1), \eta^{\text{OBS-S}}(N-m-1)\right).$$

If m_N *is any minimizing index, set*

$$h^{\text{OBS-S}}(N-1) \coloneqq h^{\text{OBS-S}}(m_N - 1) \bowtie h^{\text{OBS-S}}(N - m_N - 1).$$

Thus $\eta^{\text{OBS-S}}(N-1)$ is the minimum rate among all basic s-composable schedules with $N-1$ steps. Different minimizing splits may yield different schedules, but the optimized rate is unambiguous. We next derive the common scalar Bellman recursion underlying the ConPP/OBS-S correspondence.

A.2 The common scalar Bellman program

We use the leaf-count convention to describe the structure of schedules. A schedule containing n gradient steps can be represented by a full binary composition tree with $N = n + 1$ leaves. Each leaf represents the empty schedule, and every internal node inserts one joining step between the schedules represented by its two children.

Define

$$H_N \coloneqq \mathbf{1}^\top h_\circ^{(N-1)}, \quad s_N \coloneqq \eta^{\text{OBS-S}}(N-1).$$

The empty schedule gives $H_1 = 0$ and $s_1 = 1$. In this notation, the two recursions become

$$H_N = \max_{1 \leq m < N} \left(H_m + H_{N-m} + \varphi(H_m, H_{N-m})\right), \quad s_N = \min_{1 \leq m < N} \mathcal{J}(s_m, s_{N-m}).$$

We now introduce the symmetric composition law

$$\mathcal{K}(x, y) \coloneqq \frac{x + y + \sqrt{x^2 + 6xy + y^2}}{2}, \quad x, y \geq 0.$$

Direct substitution gives the reciprocal-affine conjugacies

$$\frac{1}{\mathcal{J}(1/x,1/y)} = \mathcal{K}(x,y), \quad x+y+\varphi(x,y)+1 = \mathcal{K}(x+1,y+1), \tag{26}$$

and

$$\mu\left(\frac{1}{x+1},\frac{1}{y+1}\right) = \varphi(x,y).$$

Here the first identity is stated for $x,y>0$, while the remaining two hold for $x,y\geq 0$. The last identity shows that corresponding ConPP and s-join nodes insert the same numerical step, not merely that their optimized scalar values agree.

Proposition A.5. *For every $N\geq 1$,*

$$U_N \coloneqq H_N+1 = \frac{1}{s_N}.$$

The common value satisfies

$$U_1=1, \quad U_N = \max_{1\leq m<N} \mathcal{K}(U_m,U_{N-m}) \quad (N\geq 2).$$

Proof. The identities in equation 26 show that both $1/s_N$ and H_N+1 satisfy the same recursion and initial condition. Induction then gives $H_N+1=1/s_N$ for all integers $N\geq 1$.

□

Equivalently, assign value one to each leaf of an ordered full binary tree and apply $\mathcal{K}$ at every internal node. Then U_N is the largest root value among all such trees with N leaves. The objective-bound coefficients also coincide:

$$\frac{1}{2H_N+1} = \frac{s_N}{2-s_N} = \frac{1}{2U_N-1}.$$

Thus maximizing U_N simultaneously optimizes the tight primitive and OBS-S objective coefficients over their respective recursively generated classes. Throughout the paper, this correspondence is used for the symmetric families generated from the empty schedule by ConPP and the s-join.

A.3 Properties of the composition law

We record the properties of $\mathcal{K}$ needed below.

Proposition A.6. *The map $\mathcal{K}:[0,+\infty)^2\to[0,+\infty)$ is continuous, symmetric, positively homogeneous, and strictly increasing in each coordinate. It satisfies*

$$\mathcal{K}(0,x)=x, \quad \mathcal{K}(x,x)=\rho x, \quad \mathcal{K}(x,y)>x+y>\max\{x,y\} \quad (x,y>0). \tag{27}$$

Moreover, for $x,y,x',y'>0$,

$$\min\left\{\frac{x'}{x},\frac{y'}{y}\right\} \leq \frac{\mathcal{K}(x',y')}{\mathcal{K}(x,y)} \leq \max\left\{\frac{x'}{x},\frac{y'}{y}\right\}, \tag{28}$$

and both inequalities are strict if $x'/x\neq y'/y$.

Proof. The first assertions follow directly from the definition of $\mathcal{K}$ in equation 2. Strict monotonicity follows by differentiation in the positive quadrant and continuity at the boundary. The identities in equation 27 can be verified directly from equation 2. Let λ and Λ be the minimum and maximum of x'/x and y'/y. Then

$$(\lambda x,\lambda y)\leq(x',y')\leq(\Lambda x,\Lambda y)$$

coordinatewise. Monotonicity and homogeneity give equation 28, with strictness when $\lambda<\Lambda$. □

A.4 THE ASYMMETRIC CERTIFICATE AND RATE RECURSION

Definition A.7 (Definition 1 in Grimmer et al. (2025b))**.** *A positive n-step schedule $h = (h_0, \dots, h_{n-1})^\top$ is* f-composable with rate $\eta > 0$ *if, for every $f \in \mathcal{F}_1$, every minimizer x_*, and every associated gradient-descent trajectory,*

$$f(x_n) - f_* \le \frac{\eta}{2} \|x_0 - x_*\|^2 ,$$

and its rate satisfies the sum–product identities

$$\eta = \frac{1}{1 + 2\sum_{i=0}^{n-1} h_i} = \prod_{i=0}^{n-1} (h_i - 1)^2. \tag{29}$$

The empty schedule is f-composable with rate one.

The identities in equation 29 balance the matching quadratic and Huber instances in one dimension, both of which attain the bound with equality (Grimmer et al., 2025b, Definition 1 and Lemma 1). Thus the rate is the exact worst-case objective coefficient for an f-composable schedule. For a general L-smooth convex objective, applying the dimensionless schedule through $x_{i+1} = x_i - (h_i/L)\nabla f(x_i)$ restores a factor L in Definition A.7. The f-join theorem of Grimmer et al. (2025b, Theorem 1) combines an s-composable prefix of rate α with an f-composable continuation of rate β into an f-composable schedule of rate

$$\mathcal{T}(\alpha, \beta) \coloneqq \frac{2\alpha\beta}{\alpha + 4\beta + \sqrt{\alpha^2 + 8\alpha\beta}}.$$

Starting from the empty schedule and iterating these joins generates the basic f-composable family. For $N \ge 1$, let $h^{\text{OBS-F}}(N-1)$ be an $(N-1)$-step member with minimum rate, and write the rate $\eta^{\text{OBS-F}}(N-1)$ as η_N^{F}.

With $n = N - 1$ and matched optimal-split choices, the optimized ConPD schedule $h_\bullet^{(n)}$ coincides with $h^{\text{OBS-F}}(N-1)$ (Zhang & Jiang, 2026, Remark 5.9). Their common rate satisfies

$$\eta_N^{\text{F}} = \frac{1}{2\mathbf{1}^\top h_\bullet^{(N-1)} + 1}. \tag{30}$$

All results below therefore apply to both constructions.

The principle of optimality gives the ConPD/OBS-F dynamic program

$$\eta_1^{\text{F}} = 1, \quad \eta_N^{\text{F}} = \min_{1 \le m < N} \mathcal{T}(s_m, \eta_{N-m}^{\text{F}}), \tag{31}$$

where a minimizing split selects an OBS-S prefix $h^{\text{OBS-S}}(m-1)$ and an OBS-F continuation $h^{\text{OBS-F}}(N-m-1)$. By Definition A.7, after restoring a general smoothness constant L, this schedule satisfies the tight guarantee

$$f(x_{N-1}) - f_* \le \eta_N^{\text{F}} \frac{L}{2} \|x_0 - x_*\|^2 .$$

Hence all subsequent asymptotic statements about η_N^{F} are statements about the exact worst-case coefficient in equation 19 for both optimized ConPD and OBS-F schedules. In particular, minimizing equation 31 directly optimizes the certified final objective gap within the basic f-composable class, equivalently within the recursively generated ConPD family described above.

B REARRANGEMENT AND BALANCED OPTIMALITY

The proof of Theorem 3.2 is algebraic. We first establish three sign identities and then use them to compare the pairings. The final subsection propagates the four-point inequality to arbitrary trees.

B.1 SIGN BEYOND THE INNER ROOTS

The proof is based on an algebraic sign certificate. For $x, y \geq 0$, define the two conjugate roots

$$r_{\pm}(x, y) \coloneqq \frac{x + y \pm \sqrt{x^2 + 6xy + y^2}}{2}.$$

They are the roots of

$$r^2 - (x + y)r - xy = 0,$$

with $r_+(x, y) = \mathcal{K}(x, y)$ and $r_-(x, y) \leq 0$. For the pairing $(a, b) \mid (c, d)$, where $a, b, c, d \geq 0$, define

$$P_{ab|cd}(z) \coloneqq \prod_{\sigma,\tau\in\{+,-\}} \Big[z^2 - \big(r_\sigma(a, b) + r_\tau(c, d)\big)z - r_\sigma(a, b)r_\tau(c, d)\Big]. \tag{32}$$

The four quadratic factors account for every choice of the two inner conjugate roots. In particular, the distinguished nonnegative root associated with this pairing is $Z_{ab|cd}$. The first lemma explains how the sign of the degree-eight polynomial $P_{ab|cd}(z)$ recovers the position of this root once $z > \max\{\mathcal{K}(a, b), \mathcal{K}(c, d)\}$.

Lemma B.1. *Let $a, b, c, d \geq 0$ and let $P_{ab|cd}$ be defined in equation 32. Define*

$$M_{ab|cd} \coloneqq \max\{\mathcal{K}(a, b), \mathcal{K}(c, d)\}.$$

Then, for every $z > M_{ab|cd}$,

$$\operatorname{sgn} P_{ab|cd}(z) = \operatorname{sgn}\big(z - Z_{ab|cd}\big).$$

Proof. Write

$$\begin{aligned} u &= r_+(a, b), \quad r_-(a, b) = -s, \\ v &= r_+(c, d), \quad r_-(c, d) = -t, \end{aligned}$$

where $u, v, s, t \geq 0$. Vieta's relations give

$$\begin{aligned} u - s &= a + b, \quad us = ab, \\ v - t &= c + d, \quad vt = cd. \end{aligned}$$

Consequently, $0 \leq s \leq u$ and $0 \leq t \leq v$.

The $(+, +)$ factor in equation 32 is

$$q_{++}(z) = z^2 - (u + v)z - uv.$$

Its two roots are the nonnegative root $Z_{ab|cd}$ and a nonpositive root, so its sign for $z > 0$ is $\operatorname{sgn}(z - Z_{ab|cd})$. If $z > \max\{u, v\}$, the remaining three factors satisfy

$$\begin{aligned} q_{+-}(z) &= z^2 - (u - t)z + ut = z(z - u) + t(z + u) > 0, \\ q_{-+}(z) &= z^2 - (-s + v)z + sv = z(z - v) + s(z + v) > 0, \\ q_{--}(z) &= z^2 + (s + t)z - st \geq z^2 - st \geq z^2 - uv > 0. \end{aligned}$$

Thus the sign of the product in equation 32 is exactly the sign of its $(+, +)$ factor, which finishes the proof. □

B.2 PAIRING-DIFFERENCE IDENTITY

Let $e_j = e_j(a, b, c, d)$ denote the elementary symmetric polynomial of degree j in the four inputs. For a scalar E, define the auxiliary quartic

$$Q_E(z) \coloneqq z^4 - e_1 z^3 - (e_2 + 2E)z^2 - e_3 z + e_4. \tag{33}$$

The next lemma expresses the differences between pairings in terms of Q_E. A complete coefficient-level derivation is given in Appendix B.2.

Lemma B.2. *For all $a, b, c, d \geq 0$,*

$$P_{ac|bd}(z) - P_{ab|cd}(z) = 2z^2(d-a)(c-b)Q_{ad+bc}(z),$$
$$P_{ad|bc}(z) - P_{ac|bd}(z) = 2z^2(b-a)(d-c)Q_{ab+cd}(z).$$

Proof. We first derive the conjugate polynomial for a generic pairing. Let

$$A = a + b, \quad B = ab, \quad C = c + d, \quad D = cd,$$

and let $u_\pm$ and $v_\pm$ be the respective roots of

$$u^2 - Au - B = 0, \quad v^2 - Cv - D = 0.$$

For fixed u, multiplication over the two choices of v gives

$$\prod_{\tau\in\{+,-\}} \left[z^2 - (u + v_\tau)z - uv_\tau\right] = (z^2 - uz)^2 - C(z + u)(z^2 - uz) - D(z + u)^2$$
$$= L(z)u^2 + M(z)u + N(z),$$

where

$$L(z) = z^2 + Cz - D,$$
$$M(z) = -2z(z^2 + D),$$
$$N(z) = z^2(z^2 - Cz - D).$$

Since $u_+ + u_- = A$ and $u_+u_- = -B$, multiplication over u_+ and u_- then yields

$$P_{A,B,C,D}(z) \coloneqq P_{ab|cd}(z) = B^2L^2 - BM^2 + N^2 - ABLM + (A^2 + 2B)LN + AMN. \quad (34)$$

Expanding equation 34 gives the monic degree-eight polynomial

$$\begin{aligned} P_{A,B,C,D}(z) = {} & z^8 - 2(A + C)z^7 + (A^2 + 2AC + C^2 - 2B - 2D)z^6 \\ & + 2(AB + CD)z^5 \\ & + \left(-A^2C^2 - 2A^2D + 2ABC + 2ACD + B^2 - 2BC^2 - 12BD + D^2\right)z^4 \\ & + 2(AD^2 + B^2C)z^3 \\ & + \left(A^2D^2 + 2ABCD + B^2C^2 - 2B^2D - 2BD^2\right)z^2 \\ & - 2BD(AD + BC)z + B^2D^2. \end{aligned} \quad (35)$$

We now expose the dependence of this polynomial on the chosen pairing. Let e_j be the elementary symmetric polynomials in the four underlying inputs and set

$$S \coloneqq B + D.$$

For every pairing of the same four inputs,

$$e_1 = A + C, \quad e_2 = B + D + AC, \quad e_3 = AD + BC, \quad e_4 = BD.$$

The coefficients other than that of z^4 simplify immediately from

$$AB + CD = e_1S - e_3,$$
$$AD^2 + B^2C = e_3S - e_1e_4,$$
$$A^2D^2 + 2ABCD + B^2C^2 = e_3^2,$$
$$B^2D + BD^2 = e_4S.$$

For the z^4 coefficient, put $T \coloneqq AC = e_2 - S$. Since

$$A^2D + BC^2 = e_1e_3 - TS,$$
$$ABC + ACD = TS,$$
$$B^2 + D^2 - 12BD = S^2 - 14e_4,$$

we obtain

$$\begin{aligned} & -A^2C^2 - 2A^2D + 2ABC + 2ACD + B^2 - 2BC^2 - 12BD + D^2 \\ = \; & -T^2 - 2(e_1e_3 - TS) + 2TS + S^2 - 14e_4 \\ = \; & -e_2^2 - 2e_1e_3 - 14e_4 + 6e_2S - 4S^2. \end{aligned}$$

Therefore equation 35 can be written solely in terms of e_1, e_2, e_3, e_4 and the pairing-dependent scalar S as

$$\begin{aligned} P_S(z) = z^8 - 2e_1z^7 + (e_1^2 - 2S)z^6 + 2(e_1S - e_3)z^5 \\ + \big(-e_2^2 - 2e_1e_3 - 14e_4 + 6e_2S - 4S^2\big)z^4 \\ + 2(e_3S - e_1e_4)z^3 + (e_3^2 - 2e_4S)z^2 \\ - 2e_3e_4z + e_4^2. \end{aligned} \tag{36}$$

This representation makes the two desired differences transparent. Define

$$S_0 \coloneqq ab + cd, \quad S_1 \coloneqq ac + bd, \quad S_2 \coloneqq ad + bc.$$

Then

$$e_2 = S_0 + S_1 + S_2,$$

and

$$\begin{aligned} S_1 - S_0 &= -(d - a)(c - b), \\ S_2 - S_1 &= -(b - a)(d - c). \end{aligned}$$

Let $F = (d - a)(c - b)$. Subtracting equation 36 at S_0 from the same formula at S_1, and using $S_0 + S_1 = e_2 - S_2$, gives

$$\begin{aligned} P_{S_1}(z) - P_{S_0}(z) &= 2Fz^6 - 2e_1Fz^5 - 2(e_2 + 2S_2)Fz^4 - 2e_3Fz^3 + 2e_4Fz^2 \\ &= 2Fz^2\big[z^4 - e_1z^3 - (e_2 + 2S_2)z^2 - e_3z + e_4\big] \\ &= 2z^2(d - a)(c - b)Q_{ad+bc}(z). \end{aligned}$$

This is the first equality.

Likewise, let $G = (b - a)(d - c)$. Since $S_1 + S_2 = e_2 - S_0$, subtraction at S_1 from the formula at S_2 yields

$$\begin{aligned} P_{S_2}(z) - P_{S_1}(z) &= 2Gz^6 - 2e_1Gz^5 - 2(e_2 + 2S_0)Gz^4 - 2e_3Gz^3 + 2e_4Gz^2 \\ &= 2Gz^2\big[z^4 - e_1z^3 - (e_2 + 2S_0)z^2 - e_3z + e_4\big] \\ &= 2z^2(b - a)(d - c)Q_{ab+cd}(z), \end{aligned}$$

which is the second equality. □

B.3 Quartic evaluation

Lemma B.3. *Fix a pairing $(a, b) \mid (c, d)$, where $a, b, c, d \geq 0$, and set*

$$u = \mathcal{K}(a, b), \quad v = \mathcal{K}(c, d), \quad z = \mathcal{K}(u, v).$$

If $u, v > 0$, then, for every scalar E,

$$Q_E(z) = \frac{2\big(abv^2 + cdu^2 - Euv\big)(uv + uz + vz)}{uv},$$

where Q_E is formed from the four inputs a, b, c, d.

Proof. Let

$$A = a + b, \quad B = ab, \quad C = c + d, \quad D = cd.$$

The two inner quadratic equations and the assumption $u, v > 0$ give

$$A = u - \frac{B}{u}, \quad C = v - \frac{D}{v}. \tag{37}$$

The outer quadratic equation gives

$$z^2 = (u + v)z + uv. \tag{38}$$

Set $S = u + v$ and $T = uv$. Repeated use of equation 38 yields

$$\begin{aligned} z^3 &= (S^2 + T)z + ST, \\ z^4 &= (S^3 + 2ST)z + S^2T + T^2. \end{aligned}$$

For the four inputs a, b, c, d, their elementary symmetric polynomials satisfy

$$e_1 = A + C, \quad e_2 = B + D + AC, \quad e_3 = AD + BC, \quad e_4 = BD.$$

Substitution into equation 33 reduces $Q_E(z)$ to a linear polynomial in z:

$$Q_E(z) = R_1 z + R_0,$$

where

$$\begin{aligned} R_0 &= S^2T + T^2 - e_1ST - (e_2 + 2E)T + e_4, \\ R_1 &= S^3 + 2ST - e_1(S^2 + T) - (e_2 + 2E)S - e_3. \end{aligned}$$

Using equation 37, we first compute

$$S^2 + T - e_1S - (e_2 + 2E) = \frac{2Bv}{u} + \frac{2Du}{v} - \frac{BD}{uv} - 2E.$$

Consequently,

$$\begin{aligned} R_0 &= T\left(\frac{2Bv}{u} + \frac{2Du}{v} - \frac{BD}{uv} - 2E\right) + BD \\ &= 2\big(Bv^2 + Du^2 - Euv\big). \end{aligned} \tag{39}$$

It remains to relate the coefficient R_1 to R_0. Directly from their definitions,

$$R_1 - \frac{S}{T}R_0 = ST - e_1T - e_3 - \frac{Se_4}{T}. \tag{40}$$

On the other hand, equation 37 gives

$$\begin{aligned} T(S - e_1) &= Bv + Du, \\ e_3 &= \left(u - \frac{B}{u}\right)D + B\left(v - \frac{D}{v}\right) \\ &= uD + Bv - \frac{BD}{u} - \frac{BD}{v}, \\ \frac{Se_4}{T} &= \frac{BD}{u} + \frac{BD}{v}. \end{aligned}$$

Substitution into equation 40 shows that $R_1 = (S/T)R_0$. Combining this identity with equation 39 and $T + Sz = uv + uz + vz$ gives

$$Q_E(z) = R_0\left(1 + \frac{S}{T}z\right) = \frac{2\big(Bv^2 + Du^2 - Euv\big)(uv + uz + vz)}{uv},$$

which finishes the proof. □

B.4 PROOF OF THE FOUR-POINT REARRANGEMENT THEOREM

Proof of Theorem 3.2. For brevity, write

$$Z_0 \coloneqq Z_{ab|cd} = \mathcal{K}\big(\mathcal{K}(a,b),\mathcal{K}(c,d)\big),$$
$$Z_1 \coloneqq Z_{ac|bd} = \mathcal{K}\big(\mathcal{K}(a,c),\mathcal{K}(b,d)\big),$$
$$Z_2 \coloneqq Z_{ad|bc} = \mathcal{K}\big(\mathcal{K}(a,d),\mathcal{K}(b,c)\big).$$

If $b = 0$, then $a = 0$, and $\mathcal{K}(0,x) = x$ gives $Z_0 = Z_1 = Z_2 = \mathcal{K}(c,d)$. Both inequalities and their stated equality conditions therefore hold. Henceforth assume $b > 0$, so all inner quantities appearing below are positive.

For the first comparison, set

$$u = \mathcal{K}(a,b), \quad v = \mathcal{K}(c,d), \quad z_0 = \mathcal{K}(u,v) = Z_0.$$

The quadratic form in Lemma B.3 factors as

$$abv^2 + cdu^2 - (ad+bc)uv = (av - cu)(bv - du). \tag{41}$$

By homogeneity of $\mathcal{K}$,

$$av = \mathcal{K}(ac,ad), \quad cu = \mathcal{K}(ac,bc), \quad bv = \mathcal{K}(bc,bd), \quad du = \mathcal{K}(ad,bd).$$

Thus, by strict coordinatewise monotonicity of $\mathcal{K}$, $av - cu$ has the sign of $ad - bc$, whereas $bv - du$ has the opposite sign. Consequently,

$$abv^2 + cdu^2 - (ad+bc)uv \le 0,$$

with equality if and only if $ad = bc$. Applying Lemma B.3 with $E = ad+bc$ yields $Q_{ad+bc}(z_0) \le 0$. Since $P_{ab|cd}(z_0) = 0$, Lemma B.2 gives

$$P_{ac|bd}(z_0) = 2z_0^2(d-a)(c-b)Q_{ad+bc}(z_0) \le 0.$$

Moreover,

$$z_0 > v = \mathcal{K}(c,d) \ge \max\{\mathcal{K}(a,c),\mathcal{K}(b,d)\} = M_{ac|bd}.$$

The sign certificate in Lemma B.1 therefore implies $Z_0 \le Z_1$.

If $b = c$, the two pairings defining Z_0 and Z_1 coincide. If $ad = bc$, equality holds in equation 41, hence $Q_{ad+bc}(z_0) = 0$ and the same sign certificate gives $Z_0 = Z_1$. Conversely, if $b < c$ and $ad \ne bc$, then equation 41 is strict and $(d-a)(c-b) > 0$. Thus $P_{ac|bd}(z_0) < 0$, which gives $Z_0 < Z_1$. This proves equation 10.

For the second comparison, set

$$u' = \mathcal{K}(a,c), \quad v' = \mathcal{K}(b,d), \quad z_1 = \mathcal{K}(u',v') = Z_1.$$

The corresponding quadratic form factors as

$$ac(v')^2 + bd(u')^2 - (ab+cd)u'v' = (cv' - bu')(av' - du'). \tag{42}$$

Homogeneity and coordinatewise monotonicity give

$$cv' = \mathcal{K}(bc,cd) \ge \mathcal{K}(ab,bc) = bu',$$
$$av' = \mathcal{K}(ab,ad) \le \mathcal{K}(ad,cd) = du'.$$

Hence the right-hand side of equation 42 is nonpositive. By Lemma B.3 with $E = ab+cd$, we have $Q_{ab+cd}(z_1) \le 0$. Since $P_{ac|bd}(z_1) = 0$, Lemma B.2 yields

$$P_{ad|bc}(z_1) = 2z_1^2(b-a)(d-c)Q_{ab+cd}(z_1) \le 0.$$

Furthermore,

$$z_1 > v' = \mathcal{K}(b,d) \ge \max\{\mathcal{K}(a,d),\mathcal{K}(b,c)\} = M_{ad|bc}.$$

Another application of Lemma B.1 gives $Z_1 \le Z_2$.

If $a = b$ or $c = d$, the pairings defining Z_1 and Z_2 agree up to symmetry. Conversely, if $a < b$ and $c < d$, strict monotonicity gives $cv' > bu'$ and $av' < du'$. Thus $Q_{ab+cd}(z_1) < 0$. Since $(b-a)(d-c) > 0$, we have $P_{ad|bc}(z_1) < 0$. The sign certificate then gives $Z_1 < Z_2$. This proves equation 11 and completes the proof. □

B.5 AN ABSTRACT BALANCED-COMPOSITION PRINCIPLE

The propagation argument uses the order, homogeneity, and rearrangement properties of the kernel, not its radical formula.

Theorem B.4. *Let $\star : [0,\infty)^2 \to [0,\infty)$ be symmetric, positively homogeneous, and strictly increasing in each coordinate. Assume that, for some $\lambda > 1$,*

$$0 \star x = x, \quad x \star x = \lambda x,$$

hold for all $x \geq 0$, and that the four-point rearrangement inequalities

$$(a \star b) \star (c \star d) \leq (a \star c) \star (b \star d) \leq (a \star d) \star (b \star c)$$

hold whenever $0 \leq a \leq b \leq c \leq d$. Define the balanced sequence $\{V_N\}_{N\geq 0}$ by

$$V_0 = 0, \quad V_1 = 1, \quad V_N = V_{\lfloor N/2 \rfloor} \star V_{\lceil N/2 \rceil} \ (N \geq 2), \tag{43}$$

and define the Bellman sequence $\{W_N\}_{N\geq 1}$ by

$$W_1 = 1, \quad W_N = \max_{1\leq m<N} W_m \star W_{N-m} \ (N \geq 2). \tag{44}$$

Then $\{V_N\}_{N\geq 0}$ is strictly increasing and satisfies the supercomposition inequality

$$V_m \star V_n \leq V_{m+n} \quad (m, n \geq 0).$$

Consequently, $W_N = V_N$ for every $N \geq 1$.

Proof. Since $V_0 < V_1 < V_2$, and since $V_r < V_{r+1}$ implies

$$V_{2r} = V_r \star V_r < V_r \star V_{r+1} = V_{2r+1} < V_{r+1} \star V_{r+1} = V_{2r+2},$$

strict monotonicity follows by induction. We prove the supercomposition inequality by strong induction on $m + n$, assuming $0 < m \leq n$; the case $m = 0$ is immediate. For even–even indices, homogeneity and the induction hypothesis give

$$V_{2a} \star V_{2b} = \lambda(V_a \star V_b) \leq \lambda V_{a+b} = V_{2a+2b}.$$

The even–odd case follows from the first rearrangement inequality and the induction hypothesis:

$$V_{2a} \star V_{2b+1} = (V_a \star V_a) \star (V_b \star V_{b+1}) \leq (V_a \star V_b) \star (V_a \star V_{b+1}) \leq V_{a+b} \star V_{a+b+1} = V_{2a+2b+1}.$$

For odd–even indices, $a + 1 \leq b$, and the same inequality with induction hypothesis gives

$$V_{2a+1} \star V_{2b} = (V_a \star V_{a+1}) \star (V_b \star V_b) \leq (V_a \star V_b) \star (V_{a+1} \star V_b) \leq V_{a+b} \star V_{a+b+1} = V_{2a+2b+1}.$$

For two odd indices, $a = b$ gives equality. If $a < b$, the full rearrangement chain and the induction hypothesis give

$$\begin{aligned} V_{2a+1} \star V_{2b+1} &= (V_a \star V_{a+1}) \star (V_b \star V_{b+1}) \leq (V_a \star V_{b+1}) \star (V_{a+1} \star V_b) \\ &\leq V_{a+b+1} \star V_{a+b+1} = V_{2a+2b+2}. \end{aligned}$$

This proves supercomposition. Finally, if $W_j = V_j$ for $j < N$, every root split has value at most V_N, whereas the balanced split attains V_N by equation 43. Induction gives $W_N = V_N$. □

Corollary B.5. *Extend the common scalar sequence by $U_0 := 0$. Then $\{U_N\}_{N\geq 0}$ is strictly increasing, and for all $m, n \geq 0$,*

$$\mathcal{K}(U_m, U_n) \leq U_{m+n}.$$

Moreover, every $N \geq 2$ satisfies the balanced recurrence

$$U_N = \mathcal{K}\big(U_{\lfloor N/2 \rfloor}, U_{\lceil N/2 \rceil}\big).$$

In particular, the balanced root split solves the Bellman program equation 4 at every leaf count.

Proof. Note that equation 44 is exactly the Bellman sequence defined by equation 4. Apply Theorem B.4 with $\star = \mathcal{K}$ and $\lambda = \rho$ to obtain that $U_N = V_N$ for all $N \geq 0$, which finishes the proof. □

These statements give the balanced recurrence and supercomposition part of Theorem 3.1. Equality and complete-tree optimality are proved next.

C Optimal splits, composition trees, and fast evaluation

The equality conditions below complete the proof of Theorem 3.1. All counts refer to ordered composition trees; distinct certificates need not produce distinct numerical stepsize vectors.

C.1 Equality in supercomposition and all maximizing root splits

The sharp rearrangement equality conditions determine whether an unbalanced split can tie. The next lemma identifies exactly when the inequality in Corollary B.5 is tight. The exceptional pairs are those that become consecutive after removal of a common power of two; this is the source of every nonbalanced tie in the Bellman recursion.

Lemma C.1. *For positive integers m, n,*

$$\mathcal{K}(U_m, U_n) = U_{m+n}$$

if and only if either

(i) $m = n$; or

(ii) there exist integers $k \geq 0$ and $r \geq 1$ such that

$$\{m, n\} = \{2^k r, 2^k(r+1)\}.$$

Proof. We use strong induction on $m + n$ and assume $m \leq n$ by symmetry. The base case $m = n = 1$ follows from equation 7. If $m = 2a$ and $n = 2b$, then

$$\mathcal{K}(U_m, U_n) = \rho\mathcal{K}(U_a, U_b), \quad U_{m+n} = \rho U_{a+b}.$$

Equality for (m, n) is therefore equivalent to equality for (a, b). By the induction hypothesis, either $a = b$, which gives $m = n$, or $\{a, b\} = \{2^k r, 2^k(r+1)\}$. Conversely, either lower-level equality propagates to (m, n) by the same identities.

Next let $m = 2a$ and $n = 2b + 1$. Here $a \geq 1$ and $a \leq b$. The supercomposition proof begins with

$$\mathcal{K}\big(\mathcal{K}(U_a, U_a), \mathcal{K}(U_b, U_{b+1})\big) \leq \mathcal{K}\big(\mathcal{K}(U_a, U_b), \mathcal{K}(U_a, U_{b+1})\big).$$

For the ordered quadruple $U_a \leq U_a \leq U_b \leq U_{b+1}$, the first equality condition in Theorem 3.2 requires either $U_a = U_b$ or $U_a U_{b+1} = U_a U_b$. The second alternative is impossible because $U_a > 0$ and $U_{b+1} > U_b$, while strict monotonicity of $\{U_j\}$ makes the first equivalent to $a = b$. Hence the supercomposition inequality is strict unless $(m, n) = (2a, 2a + 1)$. In that remaining case the two indices are consecutive, and equality follows directly from equation 7.

Similarly, for $(m, n) = (2a + 1, 2b)$, it forces $b = a + 1$, again giving consecutive indices. These equalities also follow directly from the balanced recurrence.

Finally, for $(m, n) = (2a + 1, 2b + 1)$, equality is immediate when $a = b$. If $a < b$, then

$$\begin{aligned}\mathcal{K}\big(\mathcal{K}(U_a, U_{a+1}), \mathcal{K}(U_b, U_{b+1})\big) &\leq \mathcal{K}\big(\mathcal{K}(U_a, U_b), \mathcal{K}(U_{a+1}, U_{b+1})\big) \\ &< \mathcal{K}\big(\mathcal{K}(U_a, U_{b+1}), \mathcal{K}(U_{a+1}, U_b)\big) \\ &\leq \mathcal{K}(U_{a+b+1}, U_{a+b+1}) = U_{2a+2b+2}.\end{aligned}$$

The middle inequality is strict by Theorem 3.2. Thus unequal odd indices never tie, completing the induction in both directions. □

The optimizer set now follows from the 2-adic factorization of the total leaf count.

Corollary C.2. *For $N \geq 2$, define the set of maximizing left-child leaf counts by*

$$\mathcal{M}_N \coloneqq \operatorname*{arg\,max}_{1 \leq m < N} \mathcal{K}(U_m, U_{N-m}).$$

Write $N = \nu \cdot 2^\ell$, where ν is odd. Then

(i) If $\nu = 1$, then

$$\mathcal{M}_N = \{N/2\}.$$

(ii) If $\nu > 1$ and $\ell = 0$, then

$$\mathcal{M}_N = \{(N-1)/2, (N+1)/2\}.$$

(iii) If $\nu > 1$ and $\ell \geq 1$, then

$$\mathcal{M}_N = \left\{(\nu-1)2^{\ell-1}, \nu \cdot 2^{\ell-1}, (\nu+1)2^{\ell-1}\right\}. \tag{45}$$

Proof. By the Bellman recursion and Corollary B.5, a split is maximizing exactly when

$$\mathcal{K}(U_m, U_{N-m}) = U_N = U_{m+(N-m)}.$$

The equality characterization in Lemma C.1 therefore applies. The equal-child alternative gives the unique balanced split $m = N/2$ whenever N is even. Under the scaled-consecutive alternative,

$$N = 2^k r + 2^k(r+1) = 2^k(2r+1).$$

Since $2r+1$ is odd, uniqueness of the factorization of N into an odd part and a power of two forces $k = \ell$ and $2r+1 = \nu$. Thus $r = (\nu-1)/2$, and the two child counts are

$$2^\ell r = 2^{\ell-1}(\nu-1), \quad 2^\ell(r+1) = 2^{\ell-1}(\nu+1).$$

If $\nu = 1$, this alternative is excluded by $r \geq 1$, leaving only the equal split. If $\nu > 1$ and $\ell = 0$, N is odd and only the two orientations of the consecutive split remain. If $\nu > 1$ and $\ell \geq 1$, the equal split contributes $2^{\ell-1}\nu$ in addition to the two orientations of the scaled-consecutive split. This gives all three cases. □

We finish by translating leaf counts into the step-count notation used by Zhang & Jiang (2026). This makes explicit both the balanced assertion and all the additional ties in their conjecture.

Theorem C.3 (Conjecture 4.17 in Zhang & Jiang (2026))**.** *For every step count $n \geq 1$,*

$$\mathbf{1}^\top h_\circ^{(n)} = \mathbf{1}^\top \operatorname{ConPP}\left(h_\circ^{(\lfloor (n-1)/2 \rfloor)}, h_\circ^{(\lceil (n-1)/2 \rceil)}\right).$$

Moreover, if $n = \nu \cdot 2^\ell - 1$ with $\nu > 1$ odd and $\ell \geq 1$, then, for every $i \in \{-1, 0, 1\}$,

$$\mathbf{1}^\top h_\circ^{(n)} = \mathbf{1}^\top \operatorname{ConPP}\left(h_\circ^{((\nu+i)2^{\ell-1}-1)}, h_\circ^{((\nu-i)2^{\ell-1}-1)}\right),$$

and no other ConPP root split between two recursively optimized primitive schedules attains the same total stepsize.

Proof. Note that $N = n + 1$. The balanced leaf counts $\lfloor N/2 \rfloor$ and $\lceil N/2 \rceil$ therefore become the child step counts $\lfloor (n-1)/2 \rfloor$ and $\lceil (n-1)/2 \rceil$, which proves the first assertion by Proposition A.5 and Corollary B.5. If $N = \nu \cdot 2^\ell$ with $\nu > 1$ odd and $\ell \geq 1$, by Corollary C.2, the maximizing leaf counts are $(\nu + i)2^{\ell-1}$ for $i \in \{-1, 0, 1\}$. Completeness follows again from Corollary C.2. □

By the treewise conjugacy established in Proposition A.5, the same split rules characterize all optimal basic s-join trees: replace ConPP by the s-join and interpret maximization of U_N as minimization of $s_N = U_N^{-1}$. It remains only to pass from the root optimizer set to complete trees.

Corollary C.4. *Let T be an ordered full binary composition tree with N leaves, assign value 1 to every leaf, and evaluate each internal node using $\mathcal{K}$. Then T attains the optimal root value U_N if and only if, at every internal node with j descendant leaves, its two child subtrees are optimal for their respective leaf counts and its left-child leaf count belongs to $\mathcal{M}_j$. Equivalently, T is optimal if and only if every internal split satisfies one of the two equality patterns in Lemma C.1.*

Proof. If the root subtrees have m and $N - m$ leaves and values X and Y, then $X \leq U_m$ and $Y \leq U_{N-m}$. Strict monotonicity of $\mathcal{K}$ shows that the root can attain U_N only if $X = U_m$, $Y = U_{N-m}$, and $m \in \mathcal{M}_N$; these conditions are also sufficient by Corollary C.2. Recursion gives necessity at every node, while upward induction gives sufficiency. □

The resulting multiplicity of optimal plane composition certificates is quantified in Appendix C.2.

C.2 OPTIMAL-TREE MULTIPLICITY

Although each internal node has at most three maximizing splits, independent choices across the hierarchy can produce many optimal composition certificates. Let T_N be the number of *plane* full binary trees with N unit-valued leaves whose evaluation under $\mathcal{K}$ equals U_N; plane means that left and right children are ordered. These numbers count recursive certificates, which need not yield distinct numerical schedule vectors.

Proposition C.5. *The sequence $\{T_N\}_{N\geq 1}$ satisfies*

$$T_1 = 1, \quad T_N = \sum_{m\in\mathcal{M}_N} T_m T_{N-m} \quad (N \geq 2). \tag{46}$$

More explicitly, write $N = \nu 2^\ell$, where ν is odd. For $N \geq 2$,

(i) if $\nu = 1$, then $T_N = T_{N/2}^2 = 1$;

(ii) if $\nu > 1$ and $\ell = 0$, then

$$T_N = 2T_{(N-1)/2}T_{(N+1)/2};$$

(iii) if $\nu > 1$ and $\ell \geq 1$, setting

$$a = (\nu-1)2^{\ell-1}, \quad b = \nu 2^{\ell-1}, \quad c = (\nu+1)2^{\ell-1},$$

gives

$$T_N = 2T_aT_c + T_b^2. \tag{47}$$

Proof. There is only one full binary tree with one leaf, so $T_1 = 1$. Fix $N \geq 2$ and consider an optimal plane tree with N leaves. If its left subtree has m leaves, then its right subtree has $N - m$ leaves. By Corollary C.4, optimality of the whole tree is equivalent to the conjunction of the following three conditions: $m \in \mathcal{M}_N$, the left subtree is optimal for m leaves, and the right subtree is optimal for $N - m$ leaves.

For a fixed ordered root index $m \in \mathcal{M}_N$, the two subtrees can be chosen independently. There are T_m possibilities for the left subtree and T_{N-m} possibilities for the right subtree, hence T_mT_{N-m} optimal trees with that root index. Different values of m specify different ordered root splits and therefore give disjoint classes of plane trees. Summing over all admissible root indices proves equation 46.

We next substitute the explicit optimizer sets from Corollary C.2. If $\nu = 1$, then N is a power of two and $\mathcal{M}_N = \{N/2\}$, so

$$T_N = T_{N/2}^2.$$

Repeatedly halving N eventually reaches 1; since $T_1 = 1$, induction on ℓ yields $T_N = 1$. Thus the completely balanced plane tree is the unique optimal certificate at dyadic leaf counts.

If $\nu > 1$ and $\ell = 0$, then N is odd and

$$\mathcal{M}_N = \left\{\frac{N-1}{2}, \frac{N+1}{2}\right\}.$$

The two terms in equation 46 are equal after interchanging the ordered children, which gives

$$\begin{aligned} T_N &= T_{(N-1)/2}T_{(N+1)/2} + T_{(N+1)/2}T_{(N-1)/2} \\ &= 2T_{(N-1)/2}T_{(N+1)/2}. \end{aligned}$$

Finally, suppose $\nu > 1$ and $\ell \geq 1$. In the notation of the statement, Corollary C.2 gives

$$\mathcal{M}_N = \{a, b, c\}.$$

Since $a + c = N$ and $2b = N$, the three contributions to equation 46 are

$$T_aT_c, \quad T_b^2, \quad T_cT_a,$$

and their sum is exactly equation 47. □

In particular, if $A_k \coloneqq T_{3\cdot 2^k}$, then

$$A_0 = 2, \quad A_{k+1} = A_k^2 + 2,$$

and consequently

$$T_{3\cdot 2^k} = \exp\big(\Theta(2^k)\big) = \exp\big(\Theta(N)\big).$$

Thus dyadic leaf counts have a unique optimal plane certificate, whereas the subsequence $N = 3\cdot 2^k$ has exponentially many. For reference,

$$(T_1, \dots, T_{12}) = (1, 1, 2, 1, 4, 6, 4, 1, 8, 28, 48, 38),$$
$$(T_{3\cdot 2^k})_{k=0}^{5} = (2, 6, 38, 1446, 2090918, 4371938082726).$$

C.3 Correctness and complexity of the binary evaluator

For every $r \geq 0$, the balanced recurrence and $\mathcal{K}(x, x) = \rho x$ give

$$(U_{2r}, U_{2r+1}) = (\rho U_r, \mathcal{K}(U_r, U_{r+1})),$$
$$(U_{2r+1}, U_{2r+2}) = (\mathcal{K}(U_r, U_{r+1}), \rho U_{r+1}).$$

Indeed, after a binary prefix of value r has been processed, the invariant is $(a, b) = (U_r, U_{r+1})$. The two identities above map this pair to (U_{2r}, U_{2r+1}) or (U_{2r+1}, U_{2r+2}) according to the next digit. Hence, at termination, $a = U_N$. Since $k = \lfloor \log_2 N \rfloor + 1$, Algorithm 1 uses exactly k evaluations of $\mathcal{K}$, k multiplications by ρ, and $\Theta(1)$ scalar memory. Thus U_N is computed in $\Theta(\log N)$ scalar time for $N \geq 2$, without forming $U_1, \dots, U_{N-1}$. These complexity statements use the unit-cost scalar-arithmetic model, with evaluations of $\mathcal{K}$ counted as unit-cost operations and input storage excluded.

Under this model, computing the whole prefix table costs $O(N)$ scalar operations, and explicitly returning an $(N-1)$-step schedule requires $\Omega(N)$ output operations.

D Symmetric phase: proofs and further properties

This appendix proves the exact symmetric phase law in Theorem 4.1 and derives the tight objective consequences in Corollary 4.2. It also records subdivision derivatives, a dyadic corner, certified minimum bounds, and marginal increments. Throughout, $q = 1/p = \log_\rho 2$.

D.1 Sharp power envelope

The balanced split of $2N$ is (N, N). By Corollary B.5 and $\mathcal{K}(x, x) = \rho x$, we have $U_{2N} = \rho U_N$. Iteration, together with $s_N = U_N^{-1}$, yields

$$U_{2^k N} = \rho^k U_N, \quad s_{2^k N} = \rho^{-k} s_N, \quad N \geq 1,\ k \geq 0.$$

Lemma D.1. *For all $x, y \geq 0$,*

$$\mathcal{K}(x, y) \leq (x^q + y^q)^{1/q}.$$

Equality holds if and only if $x = y$ or $xy = 0$.

Proof. The equality is immediate if $xy = 0$, so assume $x, y > 0$. Write

$$x = \mathrm{e}^{m-s}, \quad y = \mathrm{e}^{m+s}, \quad s \geq 0,$$

where symmetry allows the last restriction. Direct calculation gives

$$\mathcal{K}(x, y) = \mathrm{e}^m \exp\big(\operatorname{arsinh}(\cosh s)\big),$$
$$(x^q + y^q)^{1/q} = \mathrm{e}^m \big(2\cosh(qs)\big)^{1/q}.$$

It therefore suffices to show

$$G(s) \coloneqq \frac{1}{q}\log\big(2\cosh(qs)\big) - \operatorname{arsinh}(\cosh s) \geq 0.$$

Since $2^{1/q} = \rho$, one has $G(0) = 0$, while the large-s asymptotics give $G(s) \to 0$ as $s \to \infty$. Moreover,

$$G'(s) = \tanh(qs) - \frac{\sinh s}{\sqrt{1 + \cosh^2 s}}.$$

Define

$$\alpha(s) \coloneqq \operatorname{artanh}\left(\frac{\sinh s}{\sqrt{1 + \cosh^2 s}}\right).$$

Then

$$\alpha'(s) = \frac{\cosh s}{\sqrt{1 + \cosh^2 s}},$$

which increases strictly from $1/\sqrt{2}$ to 1. Also $1/\sqrt{2} < q < 1$: the upper bound follows from $\rho > 2$, and the lower bound from $\rho^3 = 7 + 5\sqrt{2} < 16$. Hence $\alpha(s) - qs$ is strictly convex, initially decreasing, tends to $+\infty$, and has a unique positive zero s_0. Since $\tanh$ is strictly increasing, G' is positive on $(0, s_0)$ and negative on (s_0, ∞). Together with the endpoint values, this proves $G(s) > 0$ for $s > 0$. Equality therefore occurs only when $s = 0$, equivalently $x = y$. □

Proposition D.2. *For every $N \geq 1$,*

$$U_N \leq N^p.$$

Equivalently, $s_N \geq N^{-p}$. Equality holds if and only if N is a power of two.

Proof. We argue by induction on N. For $N = 1$, we directly have $U_1 = 1 = 1^p$. For $N \geq 2$, let $a = \lfloor N/2 \rfloor$ and $b = \lceil N/2 \rceil$. Monotonicity of $\mathcal{K}$, the induction hypothesis, Lemma D.1, and $pq = 1$ give

$$U_N = \mathcal{K}(U_a, U_b) \leq \mathcal{K}(a^p, b^p) \leq (a^{pq} + b^{pq})^{1/q} = (a + b)^p = N^p.$$

If N is odd, then $a \neq b$, so the second inequality is strict. If $N = 2a$, equality holds precisely when $U_a = a^p$. Recursing proves that equality occurs exactly at powers of two. □

D.2 Adjacent-ratio contraction

The regularity of the phase profile is controlled by a one-dimensional dynamical system. For $N \geq 1$, define

$$R_N \coloneqq \frac{U_{N+1}}{U_N}, \quad k(r) \coloneqq \mathcal{K}(1, r), \quad T_0(r) \coloneqq \frac{k(r)}{\rho}, \quad T_1(r) \coloneqq \frac{\rho r}{k(r)}.$$

Lemma D.3. *For every $N \geq 1$,*

$$R_{2N} = T_0(R_N), \quad R_{2N+1} = T_1(R_N), \tag{48}$$

and

$$1 < R_N \leq \rho, \quad 0 < R_N - 1 \leq \frac{\rho - 1}{2^{\lfloor \log_2 N \rfloor}} \leq \frac{2(\rho - 1)}{N}. \tag{49}$$

Proof. The adjacent-pair identities in Section 3 and the homogeneity of $\mathcal{K}$ give equation 48. Put

$$\Delta(r) \coloneqq \sqrt{r^2 + 6r + 1}.$$

Then

$$T_0(r) = \frac{1 + r + \Delta(r)}{2\rho}, \quad T_1(r) = \frac{\rho}{2}\big(\Delta(r) - r - 1\big), \tag{50}$$

where the second identity follows from $r/k(r) = k(r) - (1+r)$. Both maps fix 1. Their derivatives are

$$T_0'(r) = \frac{1}{2\rho}\left(1 + \frac{r+3}{\Delta(r)}\right), \tag{51}$$
$$T_1'(r) = \frac{\rho}{2}\left(\frac{r+3}{\Delta(r)} - 1\right).$$

Since

$$\left(\frac{r+3}{\Delta(r)}\right)' = -\frac{8}{\Delta(r)^3} < 0,$$

both derivatives are positive and at most $1/2$ on $[1, \rho]$, with equality at $r = 1$. Furthermore,

$$\rho = \mathcal{K}(0, \rho) < k(\rho) < \mathcal{K}(\rho, \rho) = \rho^2,$$

so $1 < T_0(\rho) < \rho$ and $1 < T_1(\rho) < \rho$. Hence T_0 and T_1 map $[1, \rho]$ into itself and are $1/2$-Lipschitz on this interval, with common fixed point 1.

For $N = 1$, the claimed bounds follow immediately from $R_1 = \rho$. Hence assume $N \geq 2$ and write

$$N = (1\epsilon_1\epsilon_2 \cdots \epsilon_m)_2, \quad \epsilon_j \in \{0, 1\}, \quad m = \lfloor \log_2 N \rfloor,$$

and let $n_j = (1\epsilon_1 \cdots \epsilon_j)_2$ be the integer represented by the first $j+1$ digits, with $n_0 = 1$. Then

$$n_j = 2n_{j-1} + \epsilon_j, \quad 1 \leq j \leq m, \quad n_m = N.$$

The two identities in equation 48 give, successively,

$$R_{n_j} = T_{\epsilon_j}(R_{n_{j-1}}), \quad R_N = T_{\epsilon_m} \circ T_{\epsilon_{m-1}} \circ \cdots \circ T_{\epsilon_1}(R_1).$$

Thus each digit after the leading digit selects the corresponding branch map T_{ϵ_j}. For either $T \in \{T_0, T_1\}$, the fixed-point identity $T(1) = 1$ and the derivative bound $0 < T'(r) \leq 1/2$ on $[1, \rho]$ imply

$$0 \leq T(r) - 1 = T(r) - T(1) \leq \frac{1}{2}(r - 1), \quad 1 \leq r \leq \rho.$$

Because $R_1 = U_2/U_1 = \rho$ and both maps preserve $[1, \rho]$, induction along the binary prefixes yields

$$0 \leq R_{n_j} - 1 \leq 2^{-j}(R_1 - 1) = 2^{-j}(\rho - 1), \quad 0 \leq j \leq m.$$

Taking $j = m = \lfloor \log_2 N \rfloor$ gives $R_N - 1 \leq (\rho - 1)2^{-m}$ and, in particular, $R_N \leq \rho$. Moreover, $2^m \leq N < 2^{m+1}$ implies

$$2^{-m} \leq \frac{2}{N},$$

which proves both upper bounds in equation 49. Finally, the strict increase of $\{U_N\}$ established in the preceding section gives $U_{N+1} > U_N > 0$, and hence

$$R_N = \frac{U_{N+1}}{U_N} > 1.$$

This proves the strict lower bound and completes the proof. □

D.3 Canonical interpolation

For each $k \geq 0$, define the level-k dyadic grid and its intervals by

$$\mathcal{D}_k := \left\{ \frac{j}{2^k} : 2^k \leq j \leq 2^{k+1} \right\}, \quad I_{k,j} := \left[\frac{j}{2^k}, \frac{j+1}{2^k} \right], \ 2^k \leq j < 2^{k+1}.$$

The elements of $\mathcal{D}_k$ are called level-k nodes, the sets $I_{k,j}$ are called level-k intervals, and $\mathcal{D} := \bigcup_{k \geq 0} \mathcal{D}_k$ is the set of dyadic rationals in $[1, 2]$. For $x = j/2^k \in \mathcal{D}_k$, define

$$F(x) := \rho^{-k} U_j. \tag{52}$$

This is well defined since the definition is independent of the chosen level by equation 12. Let F_k be the continuous function that agrees with F on $\mathcal{D}_k$ and is affine on every $I_{k,j}$. For the endpoints $x < y$ of any level-k interval, the balanced recurrence gives the nonlinear midpoint rule

$$F\left(\frac{x+y}{2}\right) = \frac{1}{\rho}\mathcal{K}\big(F(x), F(y)\big). \tag{53}$$

Lemma D.4. *The function in equation 52 extends uniquely to a strictly increasing bi-Lipschitz map*

$$F : [1, 2] \to [1, \rho]$$

that satisfies $F(1) = 1$, $F(2) = \rho$, and equation 53 on every dyadic interval. More precisely,

$$c_-(y - x) \le F(y) - F(x) \le c_+(y - x), \quad 1 \le x < y \le 2, \tag{54}$$

where

$$c_- \coloneqq (\rho - 1)\left(1 - \frac{1}{2\rho}\right), \quad c_+ \coloneqq (\rho - 1)\exp\left(\frac{1}{2\rho}\right).$$

In particular, one may take $L_F \coloneqq c_+$ as a Lipschitz constant for F.

For $k \ge 0$ and $2^k \le j < 2^{k+1}$, recall that

$$I_{k,j} = \left[\frac{j}{2^k}, \frac{j+1}{2^k}\right].$$

For $r > 1$, define the subdivision weights

$$w_0(r) \coloneqq \frac{2(T_0(r) - 1)}{r - 1}, \quad w_1(r) \coloneqq \frac{2(r - T_0(r))}{r - 1}, \tag{55}$$

and set $w_0(1) = w_1(1) = 1$. Thus $w_0 + w_1 = 2$.

Proof. Let $S_{k,j}$ be the slope of F_k on $I_{k,j}$. Write $a \coloneqq F(j/2^k)$. Since the right endpoint value is aR_j, the midpoint rule, the homogeneity of $\mathcal{K}$, and the definition of T_0 give

$$F\left(\frac{2j+1}{2^{k+1}}\right) = \frac{1}{\rho}\mathcal{K}(a, aR_j) = \frac{a}{\rho}\mathcal{K}(1, R_j) = aT_0(R_j).$$

The left and right children of $I_{k,j}$ are $I_{k+1,2j}$ and $I_{k+1,2j+1}$. Their increments are $a(T_0(R_j) - 1)$ and $a(R_j - T_0(R_j))$, respectively, whereas the increment over $I_{k,j}$ is $a(R_j - 1)$. Therefore

$$\begin{aligned} S_{k,j} &= 2^k a(R_j - 1), \\ S_{k+1,2j} &= 2^{k+1}a(T_0(R_j) - 1) = \frac{2(T_0(R_j) - 1)}{R_j - 1} S_{k,j}, \\ S_{k+1,2j+1} &= 2^{k+1}a(R_j - T_0(R_j)) = \frac{2(R_j - T_0(R_j))}{R_j - 1} S_{k,j}. \end{aligned}$$

Thus

$$S_{k+1,2j} = w_0(R_j)S_{k,j}, \quad S_{k+1,2j+1} = w_1(R_j)S_{k,j}. \tag{56}$$

Differentiating equation 51 once more gives

$$T_0''(r) = -\frac{4}{\rho(r^2 + 6r + 1)^{3/2}}.$$

Thus T_0 is strictly concave and $T_0'(1) = 1/2$. On $[1, \rho]$,

$$0 \le -T_0''(r) \le M, \quad M \coloneqq \frac{1}{4\rho(\rho - 1)}, \tag{57}$$

because $-T_0''$ is decreasing on $[1, \rho]$ and $-T_0''(1) = 1/[4\rho(\rho - 1)]$. For $r > 1$, the strict increase of T_0 gives $w_0(r) > 0$, while concavity and $T_0(1) = 1$, $T_0'(1) = 1/2$ give

$$T_0(r) - 1 \le T_0'(1)(r - 1) = \frac{1}{2}(r - 1),$$

so $w_0(r) \le 1$. Moreover, Taylor's theorem yields some $\xi_r \in (1, r)$ such that

$$T_0(r) = T_0(1) + T_0'(1)(r - 1) + \frac{1}{2}T_0''(\xi_r)(r - 1)^2 = 1 + \frac{1}{2}(r - 1) + \frac{1}{2}T_0''(\xi_r)(r - 1)^2.$$

Hence, using $w_0 + w_1 = 2$ and equation 57,

$$0 \le 1 - w_0(r) = w_1(r) - 1 = -T_0''(\xi_r)(r-1) \le M(r-1).$$

The definitions at $r = 1$ give the same bounds there. Thus, for every $r \in [1, \rho]$,

$$0 < w_0(r) \le 1 \le w_1(r), \quad 0 \le 1 - w_0(r) = w_1(r) - 1 \le M(r-1). \tag{58}$$

Fix $I_{k,J}$ and, for $0 \le \ell \le k$, set

$$j_\ell := \left\lfloor \frac{J}{2^{k-\ell}} \right\rfloor.$$

Then I_{ℓ,j_ℓ} is the unique level-ℓ interval containing $I_{k,J}$. The nested sequence $I_{0,j_0} \supset I_{1,j_1} \supset \cdots \supset I_{k,j_k}$ is called the ancestor chain of $I_{k,J}$. Here $j_0 = 1$, $j_k = J$, and for each $0 \le \ell < k$ there is a unique $\delta_\ell \in \{0, 1\}$ such that $j_{\ell+1} = 2j_\ell + \delta_\ell$; the values 0 and 1 select the left and right child, respectively. Set $r_\ell := R_{j_\ell}$ and

$$e_\ell := 1 - w_0(r_\ell) = w_1(r_\ell) - 1, \quad \sigma_\ell := 2\delta_\ell - 1 \in \{-1, 1\}.$$

Then $w_{\delta_\ell}(r_\ell) = 1 + \sigma_\ell e_\ell$. Since $I_{0,1} = [1, 2]$ has slope $\rho - 1$, iteration of equation 56 along this chain gives

$$S_{k,J} = (\rho - 1) \prod_{\ell=0}^{k-1} (1 + \sigma_\ell e_\ell).$$

Since $2^\ell \le j_\ell < 2^{\ell+1}$, equation 49 and equation 58 imply

$$0 \le e_\ell \le M(r_\ell - 1) \le M(\rho - 1)2^{-\ell}, \quad 0 \le \ell < k.$$

Consequently,

$$E := \sum_{\ell=0}^{\infty} M(\rho - 1)2^{-\ell} = 2M(\rho - 1) = \frac{1}{2\rho} < 1.$$

In particular, $0 \le e_\ell < 1$. Since $1 - e_\ell \le 1 + \sigma_\ell e_\ell \le 1 + e_\ell$, the elementary inequalities $\prod_{\ell=0}^{k-1}(1 - e_\ell) \ge 1 - \sum_{\ell=0}^{k-1} e_\ell$, proved by induction, and $1 + t \le \mathrm{e}^t$ for $t \ge 0$ yield

$$S_{k,J} \ge (\rho - 1) \prod_{\ell=0}^{k-1} (1 - e_\ell) \ge (\rho - 1) \left(1 - \sum_{\ell=0}^{k-1} e_\ell \right) \ge (\rho - 1)(1 - E) = c_-,$$

$$S_{k,J} \le (\rho - 1) \prod_{\ell=0}^{k-1} (1 + e_\ell) \le (\rho - 1) \exp \left(\sum_{\ell=0}^{k-1} e_\ell \right) \le (\rho - 1)\mathrm{e}^E = c_+.$$

Thus every slope of every F_k lies in $[c_-, c_+]$.

If $x = m/2^k < y = n/2^k$ are in the same grid $\mathcal{D}_k$, then

$$F(y) - F(x) = 2^{-k} \sum_{j=m}^{n-1} S_{k,j}.$$

The slope bounds therefore give equation 54 for $x, y \in \mathcal{D}$, after passing to a common level when necessary. The upper bound makes F uniformly continuous on the dense set $\mathcal{D}$, so F has a unique continuous extension to $[1, 2]$. Taking dyadic limits gives equation 54 for all $1 \le x < y \le 2$. The lower bound makes the extension strictly increasing, and $F(1) = U_1 = 1$, $F(2) = U_2 = \rho$ show that its range is $[1, \rho]$. Finally, equation 53 remains valid on every dyadic interval because its two endpoints and midpoint belong to $\mathcal{D}$. □

D.4 Exact scalar phase and cluster sets

Theorem D.5. *For $0 \le t \le 1$, define*

$$\Phi(t) \coloneqq \rho^{-t} F(2^t),$$

and extend Φ as a 1-periodic function. Let $\mathbb{T} \coloneqq \mathbb{R}/\mathbb{Z}$. Then $\Phi : \mathbb{T} \to \mathbb{R}$ is positive, Lipschitz, and nonconstant, and

$$U_N = N^p \Phi\big(\{\log_2 N\}\big) \quad (N \ge 1),$$

where $\{\cdot\}$ takes the fractional part of a real. It is the unique continuous 1-periodic function with this property. Furthermore,

$$0 < \phi_* \coloneqq \min_{t\in[0,1]} \Phi(t) < 1, \quad \max_{t\in[0,1]} \Phi(t) = 1.$$

One valid Lipschitz constant on the phase circle is $L_\Phi \coloneqq 2(\log 2)c_+ + \rho \log \rho$.

Proof. Let $k = \lfloor \log_2 N \rfloor$, $x = N/2^k \in [1, 2)$, and $t = \log_2 x = \{\log_2 N\}$. By equation 52, $F(x) = \rho^{-k} U_N$, while $N^p = \rho^{k+t}$. Hence

$$U_N = \rho^k F(2^t) = N^p \rho^{-t} F(2^t) = N^p \Phi(t),$$

which proves equation 13. Since $\Phi(0) = F(1) = 1$ and $\Phi(1) = \rho^{-1} F(2) = 1$, the periodic extension is continuous. For $s, t \in [0, 1]$, equation 54 gives

$$|\Phi(t) - \Phi(s)| \le c_+ |2^t - 2^s| + \rho |\rho^{-t} - \rho^{-s}| \le \big(2(\log 2)c_+ + \rho \log \rho\big)|t - s|.$$

The same estimate across the identified endpoints gives the stated Lipschitz constant. Positivity follows from $F > 0$.

For $x = j/2^m \in \mathcal{D}$, Proposition D.2 gives

$$F(x) = \rho^{-m} U_j \le \rho^{-m} j^p = x^p.$$

Density extends this bound to $[1, 2]$, so $\Phi \le 1$. At $t_* = \log_2(3/2)$, the strict part of Proposition D.2 gives

$$\Phi(t_*) = \frac{U_3}{3^p} < 1.$$

Thus equation 14 holds and Φ is nonconstant. Finally, the phases $\{\{\log_2 N\} : N \ge 1\}$ are dense: for any $\theta \in [0, 1)$, take $N_k = \big\lfloor 2^{k+\theta} \big\rfloor$. Any other continuous periodic profile satisfying equation 13 agrees with Φ on this dense set and hence everywhere. □

As in Section 4, $\mathrm{Clust}(a_N)$ denotes the set of subsequential limits of a_N as $N \to \infty$.

Corollary D.6. *The normalized scalar and rate sequences satisfy*

$$\mathrm{Clust}\left(\frac{U_N}{N^p}\right) = [\phi_*, 1], \quad \mathrm{Clust}(N^p s_N) = \left[1, \frac{1}{\phi_*}\right].$$

In particular, neither sequence converges.

Proof. By equation 13, $U_N/N^p = \Phi(\theta_N)$ with $\theta_N = \{\log_2 N\}$. The sequence $\{\theta_N\}$ is dense on the phase circle, and every phase is approached along integers tending to infinity. Continuity of Φ and equation 14 therefore give

$$\mathrm{Clust}\left(\frac{U_N}{N^p}\right) = \Phi([0, 1]) = [\phi_*, 1].$$

Since $N^p s_N = (U_N/N^p)^{-1}$ and Φ is positive, taking reciprocals proves the second identity. □

D.5 TIGHT OBJECTIVE-GAP PHASE

We now translate the scalar phase into worst-case guarantees for gradient descent with recursive primitive schedules and OBS-S. Restore a general smoothness constant L, and apply a dimensionless schedule $h = (h_0, \dots, h_{N-2})^\top$ through

$$x_{i+1} = x_i - \frac{h_i}{L}\nabla f(x_i).$$

For an optimized primitive schedule of $N-1$ steps, Proposition A.5 gives $H_N = U_N - 1$. Hence the tight Huber certificate equation 24 becomes

$$f(x_{N-1}) - f_* \le B_N \frac{L}{2}\left\|x_0 - x_*\right\|^2, \quad B_N \coloneqq \frac{1}{2U_N - 1}.$$

The same coefficient governs the objective consequence of OBS-S. Indeed, its optimized s-composable rate is $\eta = s_N = 1/U_N$, and equation 25 gives

$$\frac{\eta}{2-\eta} = \frac{1/U_N}{2 - 1/U_N} = \frac{1}{2U_N - 1} = B_N.$$

For primitive schedules, equality in equation 6 is attained on the one-dimensional Huber instance of Zhang & Jiang (2026, Theorem 2.6 and Corollary 2.9); for OBS-S, a matching one-dimensional Huber instance is supplied by Grimmer et al. (2025b, Proposition 1). Thus B_N is the exact worst-case objective coefficient for the corresponding optimized schedules, not merely an upper estimate produced by the scalar analysis.

It is useful to distinguish B_N from the s-composable rate itself. Their exact relation is

$$B_N = \frac{s_N}{2 - s_N}, \quad B_N - \frac{s_N}{2} = \frac{s_N^2}{2(2 - s_N)} > 0. \tag{59}$$

Consequently, the objective coefficient is asymptotically one half of the OBS-S rate, while the positive correction in equation 59 records the finite-horizon effect of the term -1 in the denominator $2U_N - 1$.

Corollary D.7. *Let $\theta_N = \{\log_2 N\}$. For every $N \ge 1$,*

$$B_N = \frac{1}{2N^p\Phi(\theta_N) - 1}.$$

As $N \to \infty$,

$$B_N = \frac{N^{-p}}{2\Phi(\theta_N)} + O(N^{-2p}),$$

where the $O(N^{-2p})$ remainder is uniform in the dyadic phase θ_N. Furthermore,

$$\mathrm{Clust}(N^p B_N) = \left[\frac{1}{2}, \frac{1}{2\phi_*}\right].$$

Consequently, the normalized tight objective coefficient has no single asymptotic constant.

Proof. Substituting equation 13 into the definition of B_N gives equation 15. Since $\Phi(\theta_N) \ge \phi_* > 0$,

$$B_N - \frac{N^{-p}}{2\Phi(\theta_N)} = \frac{1}{\big(2N^p\Phi(\theta_N) - 1\big)2N^p\Phi(\theta_N)}.$$

Moreover, $N^p\Phi(\theta_N) = U_N \ge 1$, so $2N^p\Phi(\theta_N) - 1 \ge N^p\Phi(\theta_N)$. The displayed remainder is therefore bounded above by $N^{-2p}/(2\phi_*^2)$, proving equation 16. After multiplication by N^p,

$$N^p B_N = \frac{1}{2\Phi(\theta_N)} + O(N^{-p}).$$

Density of the integer phases, continuity of Φ, and equation 14 now yield

$$\mathrm{Clust}(N^p B_N) = \left\{\frac{1}{2\Phi(t)} : 0 \le t \le 1\right\} = \left[\frac{1}{2}, \frac{1}{2\phi_*}\right].$$

□

The oscillation occurs in the leading term and cannot be absorbed into the $O(N^{-2p})$ remainder. This proves the exact formula, uniform expansion, and cluster-set statement in Corollary 4.2.

D.6 SUBDIVISION DERIVATIVES AND BINARY-PATH PRODUCTS

Proposition D.8. *There exists a bounded function $D : [1, 2] \to [c_-, c_+]$ such that*

$$F(x) = 1 + \int_1^x D(u)\,\mathrm{d}u. \tag{60}$$

The function D is continuous at every nondyadic point $x \in (1, 2)$, and $F'(x) = D(x)$ there. At every interior dyadic point, both one-sided derivatives of F exist.

For a binary path $\epsilon = \{\epsilon_i\}_{i\geq 1} \in \{0, 1\}^{\mathbb{N}}$, set $r_0 = \rho$ and recursively define

$$r_i = \begin{cases} T_0(r_{i-1}), & \epsilon_i = 0, \\ T_1(r_{i-1}), & \epsilon_i = 1. \end{cases} \tag{61}$$

For $k \geq 1$, define the partial product

$$P_k(\epsilon) \coloneqq (\rho - 1) \prod_{i=1}^{k} w_{\epsilon_i}(r_{i-1}).$$

Then P_k converges uniformly on the path space $\{0, 1\}^{\mathbb{N}}$ to

$$P(\epsilon) \coloneqq (\rho - 1) \prod_{i=1}^{\infty} w_{\epsilon_i}(r_{i-1}) \tag{62}$$

in the precise sense that $\sup_\epsilon |P_k(\epsilon) - P(\epsilon)| \to 0$. More precisely, with $M = 1/[4\rho(\rho - 1)]$,

$$\sup_{\epsilon\in\{0,1\}^{\mathbb{N}}} \sum_{i=k+1}^{\infty} |\log w_{\epsilon_i}(r_{i-1})| \leq 4M(\rho - 1)2^{-k}.$$

If $x \in (1, 2)$ is nondyadic and $x - 1 = 0.\epsilon_1\epsilon_2\cdots$, then $D(x) = F'(x) = P(\epsilon)$. If an interior dyadic point x_0 has the two binary expansions

$$\begin{aligned} x_0 - 1 &= 0.\epsilon_1 \cdots \epsilon_m 000 \cdots \\ &= 0.\epsilon_1 \cdots \epsilon_{m-1} 0111 \cdots, \quad \epsilon_m = 1, \end{aligned}$$

then

$$\begin{aligned} F'_+(x_0) &= P(\epsilon_1, \ldots, \epsilon_m, 0, 0, \ldots), \\ F'_-(x_0) &= P(\epsilon_1, \ldots, \epsilon_{m-1}, 0, 1, 1, \ldots). \end{aligned}$$

Proof. For step-function bookkeeping, define the half-open level-k cells

$$I^{\mathrm{h}}_{k,j} \coloneqq \left[\frac{j}{2^k}, \frac{j+1}{2^k}\right), \quad 2^k \leq j < 2^{k+1}.$$

For $x \in I^{\mathrm{h}}_{k,j}$, set $S_k(x) \coloneqq S_{k,j}$, where $S_{k,j}$ is the slope of F_k on $I_{k,j}$, and set $S_k(2) \coloneqq S_{k,2^{k+1}-1}$. Thus S_k is a step function on $[1, 2]$ whose value on each level-k cell is the corresponding slope of F_k.

Fix $x \in I^{\mathrm{h}}_{k,j}$, and let $\delta \in \{0, 1\}$ be determined by $x \in I^{\mathrm{h}}_{k+1,2j+\delta}$. By equation 56, $S_{k+1}(x) = w_\delta(R_j)S_k(x)$. Since $S_k(x) \leq c_+$, $\lfloor \log_2 j \rfloor = k$, and equation 49 and equation 58 give $|w_\delta(R_j) - 1| \leq M(R_j - 1) \leq M(\rho - 1)2^{-k}$, it follows that

$$|S_{k+1}(x) - S_k(x)| = S_k(x)|w_\delta(R_j) - 1| \leq c_+M(\rho - 1)2^{-k}.$$

The same estimate holds at $x = 2$ by using the rightmost level-k cell. Hence, for $m > k$,

$$\|S_m - S_k\|_\infty \leq \sum_{\ell=k}^{m-1} \|S_{\ell+1} - S_\ell\|_\infty \leq c_+M(\rho - 1) \sum_{\ell=k}^{m-1} 2^{-\ell} \leq 2c_+M(\rho - 1)2^{-k}.$$

Thus $\{S_k\}$ is uniformly Cauchy and converges uniformly to a function $D : [1,2] \to [c_-, c_+]$, because every S_k takes values in the closed interval $[c_-, c_+]$.

The function F_k is continuous and piecewise affine, has slope S_k away from its finitely many level-k nodes, and satisfies $F_k(1) = 1$. Therefore

$$F_k(x) = 1 + \int_1^x S_k(u)\,\mathrm{d}u, \quad 1 \le x \le 2.$$

Uniform convergence of S_k implies uniform convergence of the right-hand side to $G(x) := 1 + \int_1^x D(u)\,du$. If $x \in \mathcal{D}_m$, then $F_k(x) = F(x)$ for every $k \ge m$, so $G = F$ on the dense set $\mathcal{D}$. Since both functions are continuous, $G = F$ on $[1,2]$, proving equation 60.

Now let $x \in (1,2)$ be nondyadic and write $x - 1 = 0.\epsilon_1\epsilon_2\cdots$. Define the indices of its successive containing cells by

$$j_0 := 1, \quad j_i := 2j_{i-1} + \epsilon_i = 2^i + \left\lfloor 2^i(x-1) \right\rfloor, \quad i \ge 1.$$

Then $x \in I^{\mathrm{h}}_{i,j_i}$ and I^{h}_{i,j_i} is the child of $I^{\mathrm{h}}_{i-1,j_{i-1}}$ selected by ϵ_i. Moreover, $r_0 = \rho = R_{j_0}$, and induction using equation 48 and equation 61 gives $r_i = R_{j_i}$. Iterating equation 56 along these containing cells therefore yields

$$S_k(x) = (\rho - 1)\prod_{i=1}^{k} w_{\epsilon_i}(r_{i-1}).$$

Letting $k \to \infty$ and using $S_k(x) \to D(x)$ proves equation 62.

The same construction applies to any infinite binary path $\{\epsilon_i\}_{i\ge1}$. Since j_{i-1} is a level-$(i-1)$ index, equation 49 and equation 58 give

$$|w_{\epsilon_i}(r_{i-1}) - 1| \le M(r_{i-1} - 1) \le M(\rho - 1)2^{-(i-1)} =: a_i.$$

Here $a_i \le a_1 = M(\rho-1) = 1/(4\rho) < 1/2$. The elementary estimate $|\log(1+t)| \le 2|t|$ for $|t| \le 1/2$ consequently gives, uniformly over all binary paths,

$$\sum_{i=k+1}^{\infty} |\log w_{\epsilon_i}(r_{i-1})| \le 2M(\rho-1)\sum_{i=k+1}^{\infty} 2^{-(i-1)} = 4M(\rho-1)2^{-k}.$$

This path-independent tail bound is precisely uniform convergence of the logarithms of the partial products. Since all partial products lie in the common compact interval $[c_-, c_+]$, exponentiation gives $\sup_\epsilon |P_k(\epsilon) - P(\epsilon)| \to 0$. Thus the uniformity is with respect to the binary-path parameter indexing the family of products.

To prove continuity, use for every $y \in [1,2)$ its canonical binary expansion, namely the expansion that does not end in recurring ones. For dyadic y, this is the terminating expansion, and the preceding slope-product argument still gives $D(y)$ as the corresponding infinite product. Fix a nondyadic x and an integer k. Since x lies in the interior of its half-open level-k cell, all sufficiently close y lie in the same cell and hence share the first k binary digits with x. The preceding tail estimate then gives

$$|\log D(y) - \log D(x)| \le 8M(\rho-1)2^{-k}.$$

First choosing k large and then y sufficiently close to x proves that D is continuous at x. The integral representation equation 60 and the fundamental theorem of calculus now yield $F'(x) = D(x)$.

Finally, let $x_0 \in (1,2)$ be dyadic. Its two binary expansions have the form

$$x_0 - 1 = 0.\epsilon_1\cdots\epsilon_m000\cdots, \quad \epsilon_m = 1, \quad x_0 - 1 = 0.\epsilon_1\cdots\epsilon_{m-1}0111\cdots.$$

The uniform logarithmic tail estimate shows that the infinite products associated with these two paths are well defined; denote them by $D_+(x_0)$ and $D_-(x_0)$, respectively. For every $n \ge m$, the level-n cells selected by the first n digits of these paths are $[x_0, x_0 + 2^{-n})$ and $[x_0 - 2^{-n}, x_0)$, respectively. The common-prefix argument therefore gives

$$\lim_{x\downarrow x_0} D(x) = D_+(x_0), \quad \lim_{x\uparrow x_0} D(x) = D_-(x_0).$$

For $h > 0$, equation 60 implies

$$\frac{F(x_0+h)-F(x_0)}{h} = \frac{1}{h}\int_{x_0}^{x_0+h} D(u)\,\mathrm{d}u, \quad \frac{F(x_0)-F(x_0-h)}{h} = \frac{1}{h}\int_{x_0-h}^{x_0} D(u)\,\mathrm{d}u.$$

Passing to the limit $h \downarrow 0$ gives $F'_+(x_0) = D_+(x_0)$ and $F'_-(x_0) = D_-(x_0)$. Hence the terminating expansion gives the right derivative and the expansion ending in recurring ones gives the left derivative. □

D.7 A CERTIFIED DYADIC CORNER

Corollary D.9. *The two one-sided derivatives at $x = 3/2$ satisfy*

$$F'_-(3/2) > F'_+(3/2). \tag{63}$$

Consequently, F is not differentiable at $3/2$ and is not continuously differentiable on $[1, 2]$.

Proof. The two binary expansions

$$\frac{3}{2} - 1 = \frac{1}{2} = 0.0111\cdots = 0.1000\cdots$$

and equation 62 give

$$F'_-(3/2) = (\rho-1)w_0(\rho)\prod_{i\geq 0} w_1\big(T_1^i(T_0(\rho))\big),$$
$$F'_+(3/2) = (\rho-1)w_1(\rho)\prod_{i\geq 0} w_0\big(T_0^i(T_1(\rho))\big).$$

For $m \geq 0$, set

$$x_m \coloneqq T_1^m(T_0(\rho)), \quad P_m^- \coloneqq (\rho-1)w_0(\rho)\prod_{i=0}^{m-1} w_1(x_i),$$
$$y_m \coloneqq T_0^m(T_1(\rho)), \quad P_m^+ \coloneqq (\rho-1)w_1(\rho)\prod_{i=0}^{m-1} w_0(y_i),$$

where an empty product equals one. Thus P_6^- and P_6^+ contain the initial factor and exactly six tail factors.

We now give a rational interval certificate for these two finite products. For positive rational intervals $X = [a, b]$ and $Y = [c, d]$, with $c > 0$, use

$$X + Y = [a+c, b+d], \quad X - Y = [a-d, b-c],$$
$$XY = [ac, bd], \quad \frac{X}{Y} = \left[\frac{a}{d}, \frac{b}{c}\right].$$

An enclosure $\sqrt{X} \subseteq [s_-, s_+]$ is certified by the rational inequalities $s_-^2 \leq a \leq b \leq s_+^2$. For $G \in \{T_0, T_1, w_0, w_1\}$, let $G^{\square}(X)$ denote the interval obtained by applying these rules to the explicit formulas in equation 50 and equation 55; every square-root endpoint is chosen by the preceding squaring test. Finally, let $\langle X\rangle_{12}$ be the smallest interval with endpoints in $10^{-12}\mathbb{Z}$ that contains X.

Start from

$$\boldsymbol{\rho} \coloneqq [2.414213562373095048801688724209, 2.414213562373095048801688724210]\,,$$

which contains $\rho = 1 + \sqrt{2}$ because the squares of the two endpoints after subtracting one lie below and above 2, respectively. Define

$$\mathbf{x}_0 \coloneqq \left\langle T_0^{\square}(\boldsymbol{\rho})\right\rangle_{12}, \quad \mathbf{P}_0^- \coloneqq \left\langle (\boldsymbol{\rho}-1)w_0^{\square}(\boldsymbol{\rho})\right\rangle_{12},$$
$$\mathbf{y}_0 \coloneqq \left\langle T_1^{\square}(\boldsymbol{\rho})\right\rangle_{12}, \quad \mathbf{P}_0^+ \coloneqq \left\langle (\boldsymbol{\rho}-1)w_1^{\square}(\boldsymbol{\rho})\right\rangle_{12},$$

and, for $0 \le m < 6$, recursively set

$$\mathbf{x}_{m+1} \coloneqq \left\langle T_1^{\square}(\mathbf{x}_m) \right\rangle_{12}, \quad \mathbf{P}_{m+1}^{-} \coloneqq \left\langle \mathbf{P}_m^{-} w_1^{\square}(\mathbf{x}_m) \right\rangle_{12},$$
$$\mathbf{y}_{m+1} \coloneqq \left\langle T_0^{\square}(\mathbf{y}_m) \right\rangle_{12}, \quad \mathbf{P}_{m+1}^{+} \coloneqq \left\langle \mathbf{P}_m^{+} w_0^{\square}(\mathbf{y}_m) \right\rangle_{12}.$$

Because rational interval operations preserve inclusion, induction gives $x_m \in \mathbf{x}_m$, $y_m \in \mathbf{y}_m$, $P_m^{-} \in \mathbf{P}_m^{-}$, and $P_m^{+} \in \mathbf{P}_m^{+}$. Six exact rational interval iterations give the terminal enclosures

$$x_6 \in [1.006713216536, 1.006713216564], \qquad P_6^{-} \in [1.419687938587, 1.419687949396],$$
$$y_6 \in [1.006668059132, 1.006668059134], \qquad P_6^{+} \in [1.419604771595, 1.419604772437].$$

In particular,

$$1.4196879 < P_6^{-} < 1.4196880, \quad 1.4196047 < P_6^{+} < 1.4196048.$$

By equation 58, every omitted factor in the left product is at least one, whereas every omitted factor in the right product is at most one. Therefore

$$F'_{-}(3/2) \ge P_6^{-} > P_6^{+} \ge F'_{+}(3/2),$$

which proves equation 63. □

D.8 Nested certificates for the phase minimum

The qualitative phase theorem is established analytically. We next give rigorously convergent finite-grid certificates for its minimum. For $k \ge 0$, define

$$\underline{\phi}_k \coloneqq \min_{2^k \le j < 2^{k+1}} \frac{U_j}{(j+1)^p}, \tag{64}$$
$$\overline{\phi}_k \coloneqq \min_{2^k \le j \le 2^{k+1}} \frac{U_j}{j^p}.$$

Proposition D.10. *For every $k \ge 0$,*

$$\underline{\phi}_k \le \phi_* \le \overline{\phi}_k, \quad 0 \le \overline{\phi}_k - \underline{\phi}_k \le \frac{p}{2^k}.$$

The sequence $\{\underline{\phi}_k\}$ is nondecreasing, the sequence $\{\overline{\phi}_k\}$ is nonincreasing, and both converge to ϕ_*. All quantities through level k are computable in $O(2^k)$ scalar operations.

Proof. Writing $x = 2^t$ gives

$$\phi_* = \min_{x \in [1,2]} \frac{F(x)}{x^p}.$$

If $x \in I_{k,j}$, monotonicity of F and equation 52 imply

$$\frac{F(x)}{x^p} \ge \frac{F(j/2^k)}{((j+1)/2^k)^p} = \frac{U_j}{(j+1)^p},$$

which gives the lower certificate. Minimizing $F(x)/x^p$ over the level-k grid gives the upper certificate.

For each parent index j, the lower candidates of its two children satisfy

$$\frac{U_{2j}}{(2j+1)^p} > \frac{U_j}{(j+1)^p},$$
$$\frac{U_{2j+1}}{(2j+2)^p} > \frac{\rho U_j}{\rho(j+1)^p} = \frac{U_j}{(j+1)^p},$$

where we used equation 12 and $U_{2j+1} = \mathcal{K}(U_j, U_{j+1}) > \rho U_j$. Hence $\underline{\phi}_{k+1} \ge \underline{\phi}_k$. Since the level-$k$ grid is contained in the level-$(k+1)$ grid, $\overline{\phi}_{k+1} \le \overline{\phi}_k$.

Choose j_k attaining $\underline{\phi}_k$. By Proposition D.2 and Bernoulli's inequality,

$$0 \le \overline{\phi}_k - \underline{\phi}_k \le \frac{U_{j_k}}{j_k^p}\left[1 - \left(\frac{j_k}{j_k+1}\right)^p\right] \le \frac{p}{j_k+1} \le \frac{p}{2^k}.$$

The nested bounds and vanishing gap force both endpoints to converge to ϕ_*. Finally, iterating the balanced recurrence computes all required U_j through 2^{k+1} in $O(2^k)$ scalar operations; scanning the two finite candidate sets has the same order. □

At level $k = 17$, the minimizing indices in equation 64 and directed outward intervals are

Quantity	Index	Outward interval
$\underline{\phi}_{17}$	181314	$(0.9928123698,\ 0.9928123700)$
$\overline{\phi}_{17}$	181324	$(0.9928193322,\ 0.9928193324)$

Enlarging each side by 10^{-8} gives the certified enclosure

$$0.99281236 < \phi_* < 0.99281934. \tag{65}$$

The certificate requires only the balanced recurrence through 2^{18} and the exact finite minima in equation 64. The displayed level-17 decimal enclosure is certified by interval arithmetic, with details provided in the supplementary material.

D.9 Marginal increments and fixed dyadic rays

The derivative products also determine the first-order gain from adding one leaf at a fixed dyadic mantissa. If $x \in (1,2)$ is nondyadic and $N_k = \lfloor 2^k x \rfloor$, then

$$\lim_{k\to\infty} N_k^{1-p}\big(U_{N_k+1} - U_{N_k}\big) = x^{1-p}D(x). \tag{66}$$

Indeed, the slope of the level-k cell containing x is

$$2^k \rho^{-k}\big(U_{N_k+1} - U_{N_k}\big).$$

It converges to $D(x)$ by Proposition D.8; multiplying by $(N_k/2^k)^{1-p}$ and using $N_k/2^k \to x$ proves equation 66. At a dyadic x, the same argument along the right or left adjacent cells replaces $D(x)$ by $F'_+(x)$ or $F'_-(x)$, respectively.

Exact scaling separately shows that a fixed finite seed selects a phase but cannot change the exponent. For every $m \ge 1$ and $k \ge 0$,

$$\frac{U_{m2^k}}{(m2^k)^p} = \frac{U_m}{m^p}, \quad (m2^k)^p s_{m2^k} = m^p s_m.$$

More generally, a seed with m leaves and reciprocal scalar value $W > 0$ has value $\rho^k W$ after k symmetric self-compositions, so its normalized scalar and rate constants remain W/m^p and m^p/W. For example,

$$(2^k)^p s_{2^k} = 1, \quad (3\cdot 2^k)^p s_{3\cdot 2^k} = \frac{3^p}{U_3} \approx 1.006818956.$$

E Asymmetric phase: proofs and strict support separation

The proof of Theorem 5.1 has three parts: comparison with the symmetric recursion, uniform convergence of the phase interpolants, and a support deficit argument excluding both a constant phase and contact with the support bound.

E.1 Reciprocal comparison and fixed-ray limits

The analysis of Grimmer et al. (2025b, Lemma 10 and Theorem 5) also gives, for an explicit $c_{\mathrm{low}} > 0$,

$$\eta^{\mathrm{F}}_{2N} \le \rho^{-1}\eta^{\mathrm{F}}_N, \quad \eta^{\mathrm{F}}_N \ge c_{\mathrm{low}} N^{-p}. \tag{67}$$

Reciprocating equation 31 exposes the Bellman structure. By equation 30, write $W_N \coloneqq (\eta^{\mathrm{F}}_N)^{-1} = 2\mathbf{1}^\top h^{(N-1)}_\bullet + 1$. Then

$$W_1 = 1, \quad W_N = \max_{1\le m<N} \mathcal{A}(U_m, W_{N-m}),$$

where

$$\mathcal{A}(u,w) \coloneqq \frac{4u + w + \sqrt{w^2 + 8uw}}{2}.$$

The map $\mathcal{A}$ is continuous, increasing, and positively homogeneous, but asymmetric in its two arguments: U_m represents the optimized ConPP/OBS-S prefix, whereas W_{N-m} represents the optimized ConPD/OBS-F continuation. Thus equation 18 is an ordered Bellman recursion, and the balanced rearrangement argument of Section 3 does not apply.

Before comparing the two recursions, note a simple consequence of the imported estimates. For each fixed $m \ge 1$, the sequence $\{(m2^k)^p \eta^{\mathrm{F}}_{m2^k}\}_k$ is nonincreasing and bounded below by equation 67, and hence converges to a positive limit, which is denoted by C^F_m. Shifting the ray index gives $C^F_{2m} = C^F_m$. We use these limits to identify the phase below.

E.1.1 Local comparisons for the reciprocal sandwich

Lemma E.1. *For $x, y > 0$,*

$$\mathcal{A}(x, \rho y) \le \rho \mathcal{K}(x,y), \tag{68}$$

$$\mathcal{A}(x, 2y-1) \ge 2\mathcal{K}(x,y) - 1, \quad x, y \ge 1. \tag{69}$$

Proof. Let $k = \mathcal{K}(x,y)$, so

$$k^2 = (x+y)k + xy. \tag{70}$$

For $x, y > 0$, the value $\mathcal{A}(x, \rho y)$ is the larger root of

$$Q(z) = z^2 - (4x + \rho y)z + 4x^2.$$

Using equation 70 and $\rho^2 = 2\rho + 1$ gives

$$Q(\rho k) = x\big(4x + \rho^2 y - (2\rho - 1)k\big).$$

Put $\theta = x/y$ and $s = \sqrt{2}$. After substituting the explicit formula for k, the nonnegativity of the factor in parentheses is equivalent to

$$(7 - 2s)\theta + 5 + 2s \ge (1 + 2s)\sqrt{\theta^2 + 6\theta + 1}.$$

Both sides are positive, and the left-hand side squared minus the right-hand side squared equals

$$8(6 - 4s)\left(\theta - 2 - \frac{3s}{2}\right)^2 \ge 0.$$

Hence $Q(\rho k) \ge 0$. The smaller root of Q is below $2x$, while $\rho k > 2x$; thus ρk lies to the right of the larger root. This proves equation 68.

For equation 69, $\mathcal{A}(x, 2y-1)$ is the larger root of

$$\widetilde{Q}(z) = z^2 - (4x + 2y - 1)z + 4x^2.$$

At $z = 2k - 1$, reduction by equation 70 gives the equivalent inequality

$$(2x+1)\sqrt{x^2 + 6xy + y^2} \ge 2x^2 + 2xy + 3x + y.$$

The difference of squares is $8x^2(2xy - x + y - 1) \ge 0$ for $x, y \ge 1$. Therefore $\widetilde{Q}(2k-1) \le 0$. Since $2k - 1 > 2(x+y) - 1 > 2x$ lies to the right of the smaller root, it does not exceed the larger root, proving equation 69. □

Proposition E.2. *For every $N \geq 1$,*

$$2U_N - 1 \leq W_N \leq \rho U_N.$$

Consequently, if $t = \{\log_2 m\}$, then

$$\frac{1}{\rho\Phi(t)} \leq C_m^F \leq \frac{1}{2\Phi(t)}.$$

Proof. The local comparisons established in Lemma E.1 are

$$\mathcal{A}(x, \rho y) \leq \rho\mathcal{K}(x, y), \quad x, y > 0,$$
$$\mathcal{A}(x, 2y - 1) \geq 2\mathcal{K}(x, y) - 1, \quad x, y \geq 1.$$

Assume equation 20 below N. For every split,

$$\mathcal{A}(U_m, W_{N-m}) \leq \mathcal{A}(U_m, \rho U_{N-m}) \leq \rho\mathcal{K}(U_m, U_{N-m}).$$

Maximizing and using equation 4 gives $W_N \leq \rho U_N$. For the lower bound, choose a balanced optimal split $N = a + b$. Since $U_N = \mathcal{K}(U_a, U_b)$,

$$W_N \geq \mathcal{A}(U_a, W_b) \geq \mathcal{A}(U_a, 2U_b - 1) \geq 2\mathcal{K}(U_a, U_b) - 1 = 2U_N - 1.$$

The base case is immediate. Applying the sandwich along $N = m2^k$, dividing by $(m2^k)^p$, and using the exact phase law for U_N gives the stated bounds on C_m^F. □

E.2 Discrete modulus and interpolants

The fixed-ray limits do not by themselves define a continuous phase. One also needs uniform control of neighboring normalized values. Define

$$D_N \coloneqq \frac{W_N}{N^p}, \quad C_N \coloneqq N^p \eta_N^{\mathrm{F}} = \frac{1}{D_N},$$

and recall that $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ and $\phi_* = \min_{t\in\mathbb{T}} \Phi(t) > 0$. Recall that L_F is a Lipschitz constant for the canonical interpolation F and set $L_U \coloneqq \max\{1, L_F\}$. The discrete estimates proved in Lemma E.3 give

$$0 < U_m - U_{m-1} \leq L_U m^{p-1}, \quad 0 < W_{N+1} - W_N \leq 4L_U N^{p-1}, \quad |C_{N+1} - C_N| \leq \frac{L_{\mathrm{obs}}}{N},$$

where $L_{\mathrm{obs}} \coloneqq (4L_U + \rho p)/\phi_*^2$. Since adjacent logarithmic phases are separated by $\log_2(1+1/N) \asymp N^{-1}$, the preceding estimate yields a uniform bound on the slopes of the phase interpolants.

For each $k \geq 0$ and $0 \leq j \leq 2^k$, define

$$t_{k,j} \coloneqq \log_2 \frac{2^k + j}{2^k}.$$

Let $\Psi_k : [0, 1] \to \mathbb{R}$ be the unique continuous function satisfying

$$\Psi_k(t_{k,j}) = C_{2^k+j}, \quad 0 \leq j \leq 2^k,$$

and affine on each interval $[t_{k,j}, t_{k,j+1}]$.

Lemma E.3. *Set $L_U \coloneqq \max\{1, L_F\}$, and put $\phi_* \coloneqq \min \Phi$. With $U_0 = 0$, for every $m, N \geq 1$,*

$$0 < U_m - U_{m-1} \leq L_U m^{p-1},$$
$$0 < W_{N+1} - W_N \leq 4L_U N^{p-1},$$
$$|C_{N+1} - C_N| \leq \frac{4L_U + \rho p}{\phi_*^2 N}.$$

Proof. The first estimate is immediate for $m = 1$. If $m \geq 2$, put $r = \lfloor \log_2(m - 1) \rfloor$. Exact interpolation gives

$$U_j = \rho^r F(j/2^r), \quad j = m - 1, m,$$

and hence

$$U_m - U_{m-1} \le L_F \rho^r 2^{-r} = L_F 2^{r(p-1)} \le L_U m^{p-1}.$$

Strict positivity follows from Corollary B.5.

The join satisfies $\mathcal{A}(0, w) = w$ and

$$\partial_u \mathcal{A}(u, w) = 2 + \frac{2w}{\sqrt{w^2 + 8uw}} \le 4.$$

The split $m = 1$ in equation 18 gives $W_{N+1} > W_N$. Choose a maximizing pivot $m \in \{1, \dots, N\}$ for W_{N+1}. If $m \ge 2$, the split $(m-1, N+1-m)$ is admissible for W_N; if $m = 1$, use $\mathcal{A}(U_0, W_N) = W_N$. Therefore

$$\begin{aligned} W_{N+1} - W_N &\le \mathcal{A}(U_m, W_{N+1-m}) - \mathcal{A}(U_{m-1}, W_{N+1-m}) \\ &\le 4(U_m - U_{m-1}) \le 4L_U N^{p-1}. \end{aligned}$$

The reciprocal sandwich and $\phi_* \le U_N/N^p \le 1$ give $\phi_* \le D_N \le \rho$. Thus

$$\begin{aligned} |D_{N+1} - D_N| &\le \frac{W_{N+1} - W_N}{(N+1)^p} + W_N \left(\frac{1}{N^p} - \frac{1}{(N+1)^p} \right) \\ &\le \frac{4L_U}{N} + \rho \left[1 - \left(\frac{N}{N+1} \right)^p \right] \le \frac{4L_U + \rho p}{N}. \end{aligned}$$

Since $C_N = D_N^{-1}$ and $D_N, D_{N+1} \ge \phi_*$, the final estimate follows. □

E.3 Uniform phase convergence

Theorem E.4. *There exists a unique positive Lipschitz function $\Psi_F : \mathbb{T} \to (0, \infty)$ such that the interpolation functions Ψ_k converge uniformly to Ψ_F on $[0, 1]$:*

$$\|\Psi_k - \Psi_F\|_{L^\infty([0,1])} \to 0.$$

Equivalently, because the maximal interpolation mesh tends to zero and Ψ_F is Lipschitz,

$$\sup_{2^k \le N \le 2^{k+1}} \left| N^p \eta_N^{\mathrm{F}} - \Psi_F(\{\log_2 N\}) \right| \to 0.$$

Thus

$$\eta_N^{\mathrm{F}} = N^{-p} \Psi_F(\{\log_2 N\}) + o(N^{-p})$$

as $N \to \infty$. Moreover,

$$\mathrm{Clust}(N^p \eta_N^{\mathrm{F}}) = \Psi_F(\mathbb{T}) = [\min \Psi_F, \max \Psi_F].$$

Thus equation 22 is the complete cluster interval of the normalized tight objective coefficient in equation 19. Combining this expansion with equation 19 gives the following guarantee for the optimized ConPD/OBS-F schedules

$$f(x_{N-1}) - f_* \le \left[N^{-p} \Psi_F(\{\log_2 N\}) + o(N^{-p}) \right] \frac{L}{2} \|x_0 - x_*\|^2 \tag{71}$$

where the $o(N^{-p})$ remainder has the same dyadic-block uniformity. One may take $2L_{\mathrm{obs}} \log 2$ as a Lipschitz constant for Ψ_F.

Proof. For adjacent interpolation nodes,

$$\log_2 \frac{N+1}{N} = \frac{\log(1 + 1/N)}{\log 2} \ge \frac{1}{(N+1)\log 2}.$$

The discrete modulus therefore makes $\{\Psi_k\}$ equi-Lipschitz. The reciprocal sandwich gives common positive upper and lower bounds, so Arzelà–Ascoli yields subsequences converging uniformly on $[0, 1]$.

Every subsequential limit has the same values on dyadic mantissas. Indeed, if $x = j/2^\ell \in [1, 2]$ and $k \geq \ell$, then

$$\Psi_k(\log_2 x) = C_{j2^{k-\ell}},$$

which converges to the fixed-ray limit C_j^F. Hence all uniformly convergent subsequences have limits agreeing on a dense set and therefore, by continuity, everywhere. Since the family is relatively compact and has only one possible limit, the full function sequence $\{\Psi_k\}$ converges uniformly on $[0, 1]$. The identity $C_{2m}^F = C_m^F$ matches the endpoint values and gives a function on $\mathbb{T}$, proving equation 21 and uniqueness.

Compactness of the phase circle shows that every cluster point of C_N lies in $\Psi_F(\mathbb{T})$. Conversely, $N_k = \lfloor 2^{k+t} \rfloor$ approaches any prescribed phase t, so $C_{N_k} \to \Psi_F(t)$. Since the continuous image of the circle is a compact interval, equation 22 follows. □

E.4 Sharp homogeneous support

Lemma E.5. *Let $\xi_* \in (0, 1)$ be the unique interior solution of*

$$\xi^{2q-1}(2 - \xi) = 1, \tag{72}$$

and define

$$\tau_* \coloneqq \frac{1 - \xi_*}{2\xi_*^2}, \quad B_{\sup} \coloneqq 2^q \frac{1 - \xi_*^{2q}}{(1 - \xi_*)^q}.$$

Then $B_{\sup} > 2^q$ and, for all $u, w \geq 0$,

$$\mathcal{A}(u, w)^q \leq B_{\sup} u^q + w^q. \tag{73}$$

For $u, w > 0$, equality holds if and only if $u/w = \tau_*$. Equality also holds when $u = 0$, whereas the inequality is strict when $w = 0 < u$.

Proof. By homogeneity, normalize $w = 1$ and write $t = u > 0$. Set

$$z = \frac{1 + \sqrt{1 + 8t}}{2}.$$

Then $z^2 - z - 2t = 0$, so $t = z(z - 1)/2$ and $\mathcal{A}(t, 1) = z^2$. With $x = 1/z \in (0, 1)$, the smallest admissible support coefficient is the supremum of

$$h(x) \coloneqq \frac{\mathcal{A}(t, 1)^q - 1}{t^q} = 2^q \frac{1 - x^{2q}}{(1 - x)^q}.$$

Its logarithmic derivative is

$$\frac{h'(x)}{qh(x)} = \frac{1}{1 - x} - \frac{2x^{2q-1}}{1 - x^{2q}},$$

whose sign is that of $1 - x^{2q-1}(2 - x)$. Put $a = 2q - 1 \in (0, 1)$ and $g(x) = x^a(2 - x)$. Since

$$g'(x) = x^{a-1}\big(2a - (a + 1)x\big),$$

g first increases and then decreases on $(0, 1)$; moreover, $g(0) = 0$, $g(1) = 1$, and $g'(1) = a - 1 < 0$. Thus $g(x) = 1$ has exactly one interior solution ξ_* in addition to the endpoint 1, and h has its unique maximum at ξ_*. The endpoint limits are

$$\lim_{x \downarrow 0} h(x) = 2^q, \quad \lim_{x \uparrow 1} h(x) = 0,$$

so $B_{\sup} = h(\xi_*) > 2^q$. Back-substitution gives the unique positive equality ratio

$$t = \frac{1 - \xi_*}{2\xi_*^2} = \tau_*.$$

Homogeneity yields the stated condition $u/w = \tau_*$. The boundary cases follow from $\mathcal{A}(0, w) = w$ and $\mathcal{A}(u, 0) = 2u$, together with $B_{\sup} > 2^q$. □

Set

$$d_* \coloneqq B_{\sup}^p, \quad c_* \coloneqq d_*^{-1} = B_{\sup}^{-p}, \quad r_* \coloneqq \tau_*^q, \quad \alpha_* \coloneqq \frac{B_{\sup} r_*}{1 + B_{\sup} r_*}. \tag{74}$$

Numerically, $B_{\sup} \approx 1.97530507031103$ and $c_* \approx 0.420809352823451$. The support inequality equation 73 recovers the lower barrier $C_N \geq c_* = c_{\text{low}}$ established by Grimmer et al. (2025b, Theorem 5). Below we will establish a strict separation $c_* < \min_{t \in \mathbb{T}} \Psi_F(t)$, obtained from the deficit and phase-rigidity arguments. The exact deficit identity in Proposition E.6 then shows what asymptotic contact with this barrier would require: along an optimizing right spine, every macroscopic pivot must simultaneously saturate the symmetric power bound and the support inequality. The latter selects the unique fraction α_*, while the former forces a zero symmetric phase. The technical concentration and phase-circle arguments implementing this rigidity are collected in Appendix E.6 and Appendix E.7.

E.5 Support deficit and contact geometry

Proposition E.6. *For every $N \geq 1$,*

$$W_N^q \leq B_{\sup} N, \quad C_N = N^p \eta_N^{\mathrm{F}} \geq c_*.$$

Define $E_N \coloneqq B_{\sup} N - W_N^q \geq 0$. If m is any maximizing pivot for W_N and $r = N - m$, then

$$E_N = B_{\sup}(m - U_m^q) + E_r + \sigma(U_m, W_r), \tag{75}$$

where

$$\sigma(u, w) \coloneqq B_{\sup} u^q + w^q - \mathcal{A}(u, w)^q,$$

and all three terms on the right-hand side of equation 75 are nonnegative.

Proof. The sharp symmetric power bound gives $U_m^q \leq m$. Assuming $W_j^q \leq B_{\sup} j$ below N, Lemma E.5 gives, for every split $N = m + r$,

$$\mathcal{A}(U_m, W_r)^q \leq B_{\sup} U_m^q + W_r^q \leq B_{\sup} m + B_{\sup} r = B_{\sup} N.$$

Maximizing proves $W_N^q \leq B_{\sup} N$; since $pq = 1$, this is equivalent to $C_N \geq B_{\sup}^{-p} = c_*$. For a maximizing split, adding and subtracting $B_{\sup} U_m^q + W_r^q$ gives equation 75. □

Lemma E.7. *For every $x \in (1, 2)$,*

$$F(x) < x^p.$$

Equivalently, $\Phi(t) < 1$ for every nonzero phase $t \in (0, 1)$.

Proof. Let an interior dyadic point x first appear as the midpoint of a dyadic interval $[a, b]$, and set $H = \rho^{-1}\mathcal{K}$. The midpoint rule, endpoint power bounds, and monotonicity give

$$F(x) = H(F(a), F(b)) \leq H(a^p, b^p) < \rho^{-1}(a + b)^p = x^p,$$

where strictness follows from $a < b$ and Lemma D.1. For nondyadic x, choose a sufficiently fine dyadic interval $[a, b]$ containing x with $1 < a < b < 2$. With

$$M = \max\{F(a)/a^p, F(b)/b^p\} < 1,$$

homogeneity and the same midpoint argument propagate $F(y) \leq M y^p$ to all dyadic descendants. Density and continuity yield $F(x) \leq M x^p < x^p$. □

Lemma E.8. *Write r_* and α_* as in equation 74. Set $x_* \coloneqq 2\alpha_*$. The contact fraction satisfies*

$$\alpha_* = 1 - \xi_*^{2q}, \quad 1 - \alpha_* = \xi_*^{2q}, \tag{76}$$

and

$$\frac{1}{4} < 1 - \alpha_* < \frac{1}{2}. \tag{77}$$

In particular, $1/2 < \alpha_* < 1$ and $x_* \in (1, 2)$.

Proof. The definitions of $B_{\sup}$ and τ_* give

$$B_{\sup}\tau_*^q = 2^q \frac{1-\xi_*^{2q}}{(1-\xi_*)^q}\frac{(1-\xi_*)^q}{2^q\xi_*^{2q}} = \frac{1-\xi_*^{2q}}{\xi_*^{2q}}.$$

Since $r_* = \tau_*^q$ and $\alpha_* = B_{\sup}r_*/(1+B_{\sup}r_*)$, this proves equation 76.

We next locate ξ_*. The inequalities $q < 15/19$ and $2^{30} < 3^{19}$ imply $2^{2q} < 3$. Indeed, $q < 15/19$ is equivalent to $\rho^{15} > 2^{19}$, and

$$\rho^{15} = 275807 + 195025\sqrt{2} > 2^{19}.$$

Consequently,

$$(1/2)^{2q-1}(2-1/2) = 3/2^{2q} > 1.$$

Because ξ_* is the first interior solution of equation 72, $\xi_* < 1/2$. Since $2q > 1$,

$$1-\alpha_* = \xi_*^{2q} < \xi_* < \frac{1}{2}.$$

For the other direction, $q > 11/14$ and $q < 4/5$, equivalently $\rho^{11} < 2^{14}$ and $\rho^4 > 2^5$. At $x = 12/25$, these bounds give

$$x^{2q-1}(2-x) < x^{4/7}\frac{38}{25} < 1,$$

where the last inequality is equivalent to $12^4 38^7 < 25^{11}$. Hence $\xi_* > 12/25$. Since $2q < 8/5$,

$$\xi_*^{2q} > \xi_*^{8/5} > \left(\frac{12}{25}\right)^{8/5} > \frac{1}{4};$$

the last inequality is equivalent to $12^8 4^5 > 25^8$. Together with equation 76, this proves equation 77 and the remaining assertions. □

E.6 Deficit concentration and nonconstancy

Lemma E.9. *The dyadic-ray limit satisfies*

$$C_1^F = \lim_{k\to\infty}(2^k)^p \eta_{2^k}^{\mathrm{F}} > c_*.$$

Proof. Suppose instead that $C_1^F = c_*$. For $N = 2^k$, this is equivalent to

$$\frac{W_N}{N^p} \to d_* = B_{\sup}^p, \quad \frac{E_N}{N} = B_{\sup} - \left(\frac{W_N}{N^p}\right)^q \to 0. \tag{78}$$

Starting at $n_0 = N$, choose a maximizing pivot m_i for W_{n_i} and set $n_{i+1} = n_i - m_i$ until $n_L = 1$. Iterating equation 75 gives

$$E_N = E_1 + \sum_{i=0}^{L-1}\left[B_{\sup}(m_i - U_{m_i}^q) + \sigma(U_{m_i}, W_{n_{i+1}})\right], \tag{79}$$

with every summand nonnegative.

Small pivots incur a linear support deficit. Since $q < 1$ and $\mathcal{A}(t,1) = 1 + O(t)$ as $t \downarrow 0$,

$$\frac{\sigma(t,1)}{t^q} \to B_{\sup}.$$

By homogeneity, there is $\delta > 0$ such that

$$0 < \frac{u}{w} \le \delta \quad \Longrightarrow \quad \sigma(u,w) \ge \frac{B_{\sup}}{2}u^q.$$

The sandwich and symmetric phase bounds give

$$W_r \geq U_r \geq \phi_* r^p, \quad U_m \leq m^p, \quad U_m^q \geq \phi_*^q m.$$

Choose $\varepsilon > 0$ so that

$$\phi_*^{-1}\left(\frac{\varepsilon}{1-\varepsilon}\right)^p \leq \delta.$$

If $m_i/n_i \leq \varepsilon$, then $n_{i+1} \geq (1-\varepsilon)n_i$ and $U_{m_i}/W_{n_{i+1}} \leq \delta$. Thus

$$\sigma(U_{m_i}, W_{n_{i+1}}) \geq c_0 m_i, \quad c_0 := \frac{B_{\sup}}{2}\phi_*^q > 0. \tag{80}$$

Let j be the first index with $m_j/n_j > \varepsilon$. Such an index exists for all large N; otherwise equation 79, equation 80, and $\sum_i m_i = N-1$ would contradict equation 78. Moreover,

$$\sum_{i<j} m_i \leq \frac{E_N}{c_0} = o(N), \quad \frac{n_j}{N} \to 1.$$

Nonnegativity in equation 79 then gives

$$m_j - U_{m_j}^q = o(N), \quad E_{n_{j+1}} = o(N), \quad \sigma(U_{m_j}, W_{n_{j+1}}) = o(N). \tag{81}$$

Pass to a subsequence with $m_j/n_j \to \alpha \in [\varepsilon, 1]$. The preceding relations imply

$$\frac{U_{m_j}^q}{n_j} \to \alpha, \quad \frac{W_{n_{j+1}}^q}{n_j} \to B_{\sup}(1-\alpha).$$

Because σ is homogeneous of degree q and $pq = 1$, its vanishing limit becomes

$$B_{\sup}\alpha + B_{\sup}(1-\alpha) - \mathcal{A}\big(\alpha^p, [B_{\sup}(1-\alpha)]^p\big)^q = 0.$$

The case $\alpha = 1$ is impossible because $B_{\sup} > 2^q$. Equality in Lemma E.5 therefore forces

$$\frac{\alpha^p}{[B_{\sup}(1-\alpha)]^p} = \tau_*, \quad \text{hence} \quad \alpha = \frac{B_{\sup} r_*}{1 + B_{\sup} r_*} = \alpha_*.$$

Thus every convergent subsequence has the same limit, and $m_j/N \to \alpha_* \in (1/2, 1)$.

Since $N = 2^k$, eventually $2^{k-1} < m_j < 2^k$ and $2m_j/N \to x_* = 2\alpha_* \in (1, 2)$. Exact interpolation on $[2^{k-1}, 2^k]$ gives

$$\frac{U_{m_j}^q}{m_j} = \frac{F(2m_j/N)^q}{2m_j/N}.$$

The left-hand side tends to one by equation 81. Passage to the limit yields $F(x_*)^q = x_*$, or $F(x_*) = x_*^p$, contradicting Lemma E.7. Therefore $C_1^F > c_*$. □

Lemma E.10. *If the phase profile in Theorem E.4 were constant, then it would be identically equal to c_*.*

Proof. Suppose $\Psi_F \equiv c$ and set $d = c^{-1}$. The support bound gives $d \leq d_*$. Let $h \geq 0$ and $b \geq 1$ be arbitrary integers, and set $a = 2^h$. For each $k \geq 0$, consider the Bellman recursion for $W_{(a+b)2^k}$ and use the admissible pivot $a2^k$. Exact scaling of U and the constant-profile asymptotics give, after division by 2^{kp} and passage to the limit,

$$d(a+b)^p \geq \mathcal{A}(a^p, db^p).$$

By homogeneity, with $\alpha = a/(a+b)$,

$$d \geq \mathcal{A}(\alpha^p, d(1-\alpha)^p). \tag{82}$$

The set $\{b/2^h : h \geq 0, b \geq 1\}$ is dense in $(0, \infty)$, so equation 82 holds for every $\alpha \in [0, 1]$.

If $d^q < B_{\sup}$, choose $\alpha = d^q r_*/(1 + d^q r_*)$. The two positive arguments in equation 82 then have equality ratio τ_*. By Lemma E.5, the joined value has q-th power

$$B_{\sup}\alpha + d^q(1-\alpha) > d^q,$$

contradicting equation 82. Hence $d^q \geq B_{\sup}$, so $d \geq d_*$. Therefore $d = d_*$ and $c = c_*$. □

The preceding two results imply that Ψ_F is nonconstant: a constant profile would equal c_* by Lemma E.10, contradicting its strictly larger value at the dyadic phase from Lemma E.9.

E.7 GLOBAL STRICT SEPARATION

Lemma E.11. *For every $t \in \mathbb{T}$,*

$$\Psi_F(t) > c_*.$$

Consequently, $\min_{t\in\mathbb{T}} \Psi_F(t) > c_*$.

Proof. By Proposition E.6 and the uniform phase convergence in Theorem E.4, $\Psi_F(t) \geq c_*$ for every $t \in \mathbb{T}$. Suppose, for a contradiction, that $\Psi_F(t_0) = c_*$ for some $t_0 \in \mathbb{T}$. Choose integers $N_k \to \infty$ such that $\{\log_2 N_k\} \to t_0$ in $\mathbb{T}$. Then

$$\frac{W_{N_k}}{N_k^p} \to d_*, \quad \frac{E_{N_k}}{N_k} = B_{\sup} - \left(\frac{W_{N_k}}{N_k^p}\right)^q \to 0.$$

We first make explicit the concentration argument underlying estimate equation 80, which applies to any integer sequence with $E_{N_k}/N_k \to 0$. Along a maximizing right spine, write

$$n_{k,0} = N_k, \quad n_{k,i+1} = n_{k,i} - m_{k,i}, \quad n_{k,L_k} = 1.$$

The deficit identity of Proposition E.6 gives

$$E_{N_k} = E_1 + \sum_{i=0}^{L_k-1} \left[B_{\sup}\big(m_{k,i} - U^q_{m_{k,i}}\big) + \sigma\big(U_{m_{k,i}}, W_{n_{k,i+1}}\big)\right],$$

and every term is nonnegative. The small-pivot estimate equation 80 supplies constants $\varepsilon \in (0,1)$ and $c_0 > 0$, independent of the initial size, such that

$$\frac{m_{k,i}}{n_{k,i}} \leq \varepsilon \quad \Longrightarrow \quad \sigma\big(U_{m_{k,i}}, W_{n_{k,i+1}}\big) \geq c_0 m_{k,i}.$$

For all sufficiently large k, there must therefore be a first index j_k with $m_{k,j_k}/n_{k,j_k} > \varepsilon$. Otherwise, $E_{N_k} \geq c_0(N_k - 1)$, contradicting $E_{N_k}/N_k \to 0$. Set

$$b_k \coloneqq \sum_{i<j_k} m_{k,i}, \quad s_k \coloneqq n_{k,j_k} = N_k - b_k, \quad m_k \coloneqq m_{k,j_k}, \quad r_k \coloneqq n_{k,j_k+1} = s_k - m_k.$$

In particular, r_k is the actual continuation size after the first macroscopic pivot. Nonnegativity and the small-pivot estimate yield

$$0 \leq b_k \leq \frac{E_{N_k}}{c_0} = o(N_k), \quad \frac{s_k}{N_k} \to 1.$$

Applying the deficit identity at s_k, and observing that the deficit at any later spine node is at most E_{N_k}, also gives

$$m_k - U^q_{m_k} = o(N_k), \quad E_{r_k} = o(N_k), \quad \sigma(U_{m_k}, W_{r_k}) = o(N_k).$$

Consider any subsequence on which $m_k/s_k \to \beta \in [\varepsilon, 1]$. Since $r_k = s_k - m_k$, the preceding estimates imply

$$\frac{U^q_{m_k}}{s_k} \to \beta, \quad \frac{W^q_{r_k}}{s_k} = B_{\sup}\frac{r_k}{s_k} - \frac{E_{r_k}}{s_k} \to B_{\sup}(1-\beta).$$

Using $pq = 1$, continuity, and the degree-q homogeneity of σ, we obtain

$$0 = B_{\sup} - \mathcal{A}(\beta^p, [B_{\sup}(1-\beta)]^p)^q.$$

The case $\beta = 1$ would give $B_{\sup} = 2^q$, which is excluded by Lemma E.5. Thus both arguments of $\mathcal{A}$ are positive, and the equality characterization in that lemma forces

$$\frac{\beta^p}{[B_{\sup}(1-\beta)]^p} = \tau_*, \quad \beta = \frac{B_{\sup}\tau_*^q}{1 + B_{\sup}\tau_*^q} = \alpha_*.$$

Every convergent subsequence of m_k/s_k has this same limit. Consequently,

$$\frac{m_k}{N_k} \to \alpha_*, \quad \frac{r_k}{N_k} \to 1-\alpha_*, \quad \frac{U^q_{m_k}}{m_k} \to 1, \quad \frac{E_{r_k}}{r_k} \to 0.$$

Here the last limit uses $1-\alpha_* > 0$; in particular, $m_k, r_k \to \infty$.

The exact symmetric phase law now gives

$$\frac{U^q_{m_k}}{m_k} = \Phi(\{\log_2 m_k\})^q \to 1.$$

By Lemma E.7, $\Phi(t) = 1$ only at the zero phase. Continuity and compactness of $\mathbb{T}$ therefore imply $\{\log_2 m_k\} \to 0$ in $\mathbb{T}$. Since

$$\{\log_2 m_k\} = \{\log_2 N_k + \log_2(m_k/N_k)\},$$

we conclude that

$$t_0 + \log_2 \alpha_* = 0 \quad \text{in } \mathbb{T}.$$

On the other hand, $E_{r_k}/r_k \to 0$ implies $W_{r_k}/r_k^p \to d_*$, while the actual continuation phases satisfy

$$\{\log_2 r_k\} \to t_1 := t_0 + \log_2(1-\alpha_*) \quad \text{in } \mathbb{T}.$$

Uniform phase convergence hence gives $\Psi_F(t_1) = c_*$. The same concentration argument, now applied to the sequence r_k, yields

$$t_1 + \log_2 \alpha_* = 0 \quad \text{in } \mathbb{T}.$$

Subtracting the two phase identities shows that $\log_2(1-\alpha_*) = 0$ in $\mathbb{T}$. Thus $1-\alpha_*$ would be an integral power of two, contrary to $1/4 < 1-\alpha_* < 1/2$ from Lemma E.8. This contradiction proves $\Psi_F(t) > c_*$ for every phase. Finally, Ψ_F is continuous on the compact circle and attains its minimum, so $\min_{t\in\mathbb{T}} \Psi_F(t) > c_*$. □

Proposition E.12. *For every $t \in \mathbb{T}$,*

$$\Psi_F(t) \le \frac{c_*}{\phi_*}.$$

Equivalently, with $\Delta_F(t) = 1/\Psi_F(t)$, one has $\Delta_F(t) \ge d_*\phi_*$.

Proof. Let t_0 minimize Δ_F, and choose $N_k \to \infty$ whose phases approach t_0. Fix $\alpha \in (0,1)$ and put $m_k = \lfloor \alpha N_k \rfloor$ and $r_k = N_k - m_k$. The Bellman recursion, the minimum property of $\Delta_F(t_0)$, and $U_{m_k}/m_k^p \ge \phi_*$ give, after division by N_k^p and passage to the limit,

$$\Delta_F(t_0) \ge \mathcal{A}\big(\phi_*\alpha^p, \Delta_F(t_0)(1-\alpha)^p\big). \tag{83}$$

If $\Delta_F(t_0)^q < B_{\sup}\phi_*^q$, choose

$$\alpha = \frac{\Delta_F(t_0)^q r_*}{\phi_*^q + \Delta_F(t_0)^q r_*}.$$

The two positive arguments in equation 83 then have equality ratio τ_*. Their joined value has q-th power

$$B_{\sup}\phi_*^q\alpha + \Delta_F(t_0)^q(1-\alpha) > \Delta_F(t_0)^q,$$

contradicting equation 83. Hence $\Delta_F(t_0)^q \ge B_{\sup}\phi_*^q$, so $\Delta_F(t_0) \ge d_*\phi_*$. Since t_0 minimizes Δ_F, this bound holds at every phase; reciprocation proves the result. □

E.8 Completion of the asymmetric phase theorem

Theorem E.13. *The phase Ψ_F is nonconstant and satisfies*

$$c_* < \min_{t\in\mathbb{T}} \Psi_F(t) < \max_{t\in\mathbb{T}} \Psi_F(t) \le \frac{c_*}{\phi_*}.$$

Consequently, the cluster interval in equation 22 is nondegenerate and lies strictly above the sharp homogeneous support barrier. By equation 71, this is precisely the nondegenerate interval of asymptotic leading constants in the tight objective-gap guarantee for optimized ConPD/OBS-F schedules.

Proof. The support and deficit decomposition are established in Proposition E.6. The dyadic-ray separation and constant-phase rigidity in Lemma E.9 and Lemma E.10 imply that Ψ_F is nonconstant. The global strict lower bound and the upper phase enclosure are proved in Lemma E.11 and Proposition E.12. These statements give equation 23, while equation 22 identifies the complete cluster interval. □

Together with the uniform convergence and cluster-set statement above, this proves Theorem 5.1.

F Additional numerical phase illustrations

For the symmetric family, figure 2 displays the underlying scalar phase Φ, whereas the left panel of figure 1 displays the corresponding objective-gap leading phase $1/(2\Phi)$. The detailed plot complements the exact phase law and the rigorous enclosure of $\phi_* = \min \Phi$ established in Appendix D.8.

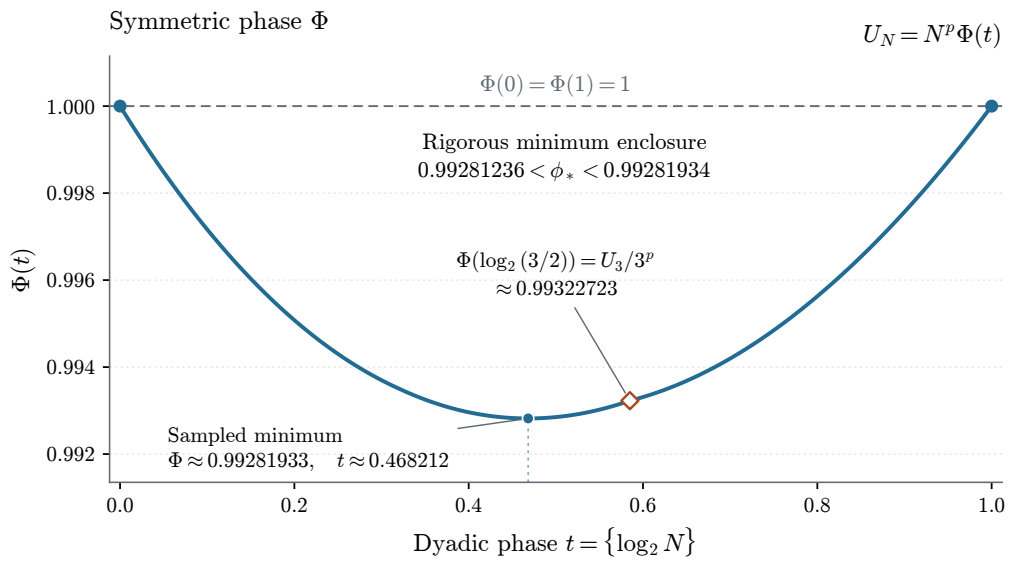


Figure 2: Detailed symmetric scalar-phase illustration. The annotations show an exact nonconstancy witness and a sampled minimum, together with the certified range of the phase minimum.

For the asymmetric family, figure 3 displays the finite-scale interpolant Ψ_{18} over one dyadic block. It provides a more detailed numerical view of the asymmetric phase than the compact panel in figure 1. Theorem 5.1 establishes the uniform convergence $\Psi_k \to \Psi_F$ and gives the limiting interpretation of these finite-scale profiles.

Implementation details, plotting data, reproduction commands, interval-arithmetic checks, and software and hardware information are provided in the supplementary material.

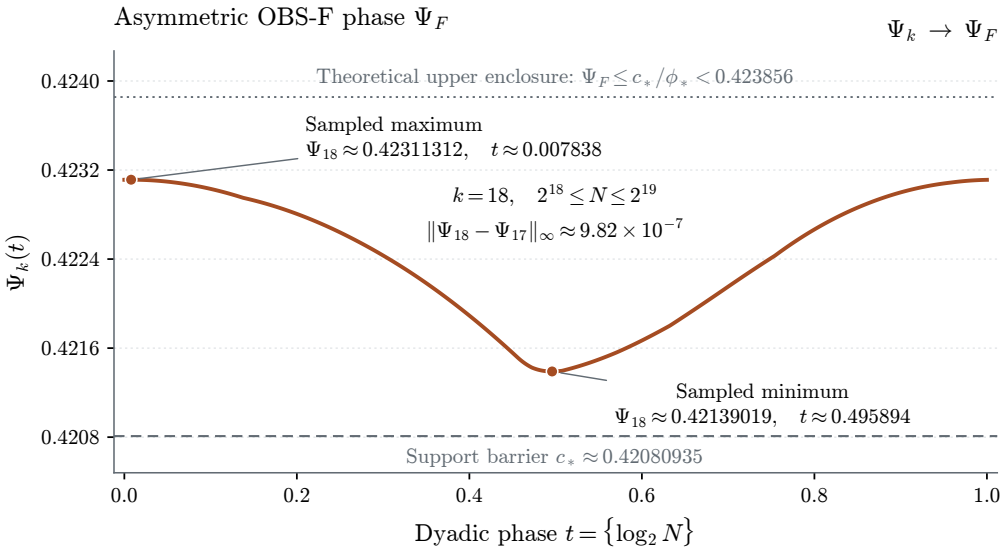


Figure 3: Detailed finite-scale asymmetric phase illustration. The curve is Ψ_{18}. The annotations show its sampled extrema, the sharp support barrier c_*, the theoretical upper enclosure for Ψ_F, and the successive-scale comparison with Ψ_{17}.