

Technical Supplement: Signed Pair Jumps and Numerical Recovery for Raw Level-2 NMSC Networks

Alec Kriebel

ORCID: 0009-0001-9320-500X DOI: [10.5281/zenodo.22729545](https://doi.org/10.5281/zenodo.22729545)

September 5, 2026

Abstract

This supplement supplies detailed probability and graph arguments for the pointwise raw strong-tree-child level-at-most-two NMSC quotient. The proof uses sampling consistency, scalar co-occupancy extracted from frozen local genealogy germs, signed biological-event jumps, chronological descendant clusters, and explicit two- and three-lineage numerical formulas. It also preserves the exact hidden-clock and rank-one counterexamples, derives the five-core/twelve-repair catalog, and distinguishes universal proofs from bounded exact verification. All times and rates are normalized as in the main paper. Rates are positive and constant between demographic vertices, finite durations are positive, inheritance weights are interior, and distinct internal biological events have distinct absolute ages. The known taxon sampling map is surjective. No population occupancy or historical route tag is observed.

1 Sampled-lineage convention and genealogy cells

Let X be the terminal taxon set, S the sampled-lineage set, and $\mathfrak{s} : S \rightarrow X$ the known sampling map. At time zero there are $m_x = |\mathfrak{s}^{-1}(x)|$ labelled blocks in the terminal population x . A metric genealogy has leaves S .

The rooted presentations in this supplement satisfy the strong semi-directed condition used by the five-core theorem: after standard semi-directed reduction, every tail of a retained reticulation arrow has two incident ordinary edges. This is stronger than requiring only the supplied rooting to be tree-child.

As in the main paper, surviving blocks enter a single ancestral population above the root, with a finite rate word followed by a constant positive pair rate on the final unbounded epoch. Complete coalescence therefore occurs almost surely.

For $T \subseteq S$, a *marked T -cell* is specified by:

1. an ordered list of mergers of blocks of the partition of T ;
2. an open simplex of allowed merger times lying inside temporal chambers;
3. the condition that no other merger among the retained T -blocks occurs between listed events.

No route or population occupancy is included in the observed cell.

Lemma 1.1 (Positive cell formula). *For a fixed compatible hidden routing and occupancy history H , the density of a marked cell is*

$$c_H(\mathbf{t}) = \prod_v \gamma_v^{a_v} (1 - \gamma_v)^{b_v} \prod_{j=1}^r \lambda_{e_j} \exp \left[- \sum_e \int q_e(s) \lambda_e(s) ds \right].$$

It is strictly positive at every strict parameter point. The observed cell density is a finite sum of such terms, and hence is positive exactly when at least one hidden history is compatible.

Proof. Each route choice has positive Bernoulli probability. Conditional on a finite occupancy schedule, each listed pair merger contributes the positive pair rate of its population and each no-merger interval contributes a positive Kingman survival factor. There are finitely many route assignments for finitely many labels. Nonnegative terms cannot cancel. $\square$

Corollary 1.2 (Support survives resonance). *Equal population rates, repeated exponential exponents, and $\gamma = 1/2$ may combine analytic terms, but cannot remove a feasible labelled cell or create an infeasible one.*

2 Sampling consistency in a structured coalescent

The proof of the multilineage theorem uses a pathwise restriction construction rather than an informal appeal to exchangeability.

Lemma 2.1 (Kingman restriction on one population). *Run a Kingman coalescent with pair rate $\lambda(t)$ on a finite set of labelled blocks and restrict its partition to a subset T of initial labels. After deleting mergers involving only deleted labels and suppressing repeated restricted partitions, the restricted process is a Kingman coalescent on T with the same pair rate.*

Proof. Use independent exponential pair clocks on each constant-rate interval. For any two current restricted blocks, the clock for their current full ancestral blocks has rate λ . Clocks involving a deleted-only ancestral block do not change the restricted partition. The memoryless property restarts the same clock system after each event. Piecewise rates follow by concatenation. $\square$

Theorem 2.2 (NMSC sampling consistency). *Restricting an NMSC metric genealogy on S to $T \subseteq S$ gives exactly the NMSC law obtained from the same temporal network with only the initial blocks in T .*

Proof. Apply lemma 2.1 on each population epoch. At a tree boundary, deterministic transport plainly commutes with restriction. At a reticulation, retain the independent Bernoulli route of each selected block and sum over routes of deleted blocks; their total probability is one. Induction through the finite temporal event list proves equality of the restricted partition-valued process, including its merger times and labels. $\square$

Corollary 2.3 (Restriction of an m -block routing boundary). *Let m blocks reach one reticulation. If $A \subseteq [m]$ is the set choosing parent 1, its weight is*

$$w_A(\gamma) = \gamma^{|A|}(1 - \gamma)^{m - |A|}.$$

For a selected set T of r labels, marginalization gives

$$\sum_{B \subseteq T} \gamma^{|B|}(1 - \gamma)^{r - |B|} F_B^{(T)}.$$

Proof. For fixed $B \subseteq T$, sum over $C \subseteq [m] \setminus T$:

$$\sum_C \gamma^{|B| + |C|}(1 - \gamma)^{m - |B| - |C|} = \gamma^{|B|}(1 - \gamma)^{r - |B|}(\gamma + 1 - \gamma)^{m - r}.$$

Coalescences with deleted labels disappear by theorem 2.2. $\square$

3 Frozen local genealogy germs

Fix a time t inside a boundary-free interval. An auxiliary frozen process keeps the current populations and their current positive rates forever, with no further population-network boundaries. It may finish as a forest rather than a complete tree. It is used only to identify coefficients of the actual local analytic germ; it is not a replacement evolutionary model.

Lemma 3.1 (Pair co-occupancy from the observed law). *Let T_{ab} be the merger time in the pair restriction and $S_{ab}(t) = \Pr(T_{ab} > t)$. The unnormalized quantity*

$$q_{ab}(t) = \Pr(T_{ab} > t, a, b \text{ in one population at } t)$$

is determined by the complete pair law, including both one-sided limits at every boundary.

Proof. On a sufficiently short interval after t , the observed survival germ is

$$S_{ab}(t+u) = p_0(t) + \sum_e p_e(t) e^{-\lambda_e u}.$$

Here p_0 is the unmerged separated mass and p_e is the unmerged mass with both lineages in population e . The actual germ uniquely determines its entire exponential-polynomial extension G_t . Since every $\lambda_e > 0$,

$$q_{ab}(t) = S_{ab}(t) - \lim_{u \rightarrow \infty} G_t(u).$$

Equal positive rates combine terms but never combine with exponent zero. The finite presentation gives piecewise analytic functions and their one-sided limits. For two supplied realizations, compare at cuts outside the union of their finite boundary sets, where the same law has the same local germ; then take one-sided limits. Thus the argument does not require the raw boundaries to be supplied as observed input. Distinct finite positive exponential sums cannot agree on an open interval and have different constant terms, by analytic uniqueness (or Vandermonde independence). All masses here are unnormalized; dividing instead by $S_{ab}(t)$ would change amplitudes. $\square$

Lemma 3.2 (A three-label current grouping). *For three distinct sampled labels a, b, z , the actual three-label law determines*

$$q_{ab|z}(t) = \Pr(\text{all three unmerged at } t, a, b \text{ together and } z \text{ elsewhere}).$$

Proof. Let $F_{ab,t}(v, w)$ be the observed density of no merger before t , first merger ab at $t+v$, and no second merger during the following w units of time. Its germ for $v, w > 0$ sufficiently small lies in the same boundary-free interval and has a unique entire extension in both variables. In the frozen process it is

$$F_{ab,t}(v, w) = \sum_e p_{ab|z,e}(t) \lambda_e e^{-\lambda_e v} + \sum_e p_{abz,e}(t) \lambda_e e^{-3\lambda_e v} e^{-\lambda_e w}.$$

If all three labels start together, the total initial rate is $3\lambda_e$ while the specified ab merger has rate λ_e ; afterward the remaining pair has rate λ_e . Every other grouping has zero density for first merger ab . Therefore

$$q_{ab|z}(t) = \lim_{w \rightarrow \infty} \int_0^\infty F_{ab,t}(v, w) dv.$$

The finite positive sum justifies both operations and gives the formula also at repeated rates. One-sided limits give the value just before a reticulation. This extracts one current grouping mass; it does not disintegrate arbitrary historical routes or their full suffix laws. $\square$

4 Signed biological jumps and the graph

At a tree vertex of age h with child sample clusters A, B ,

$$\Delta q_{ij}(h) = \Pr(T_{ij} > h, (i \in e_A, j \in e_B) \text{ or } (i \in e_B, j \in e_A)) \geq 0.$$

Its strict-positive support is the cross-pair graph $A \times B$, with repeated labels excluded. At a reticulation v with child population e , age h and inheritance γ ,

$$\Delta q_{ij}(h) = -2\gamma(1 - \gamma) \Pr(T_{ij} > h, i, j \in e) \leq 0.$$

The strict-negative support is the full pair clique of the descendant sample cluster $C(v)$. The incoming reticulation arcs are new at h , so no exterior lineage is already in either of them immediately above the boundary. No merger has an atom at a fixed biological time. Demographic changes do not move a lineage and hence do not jump q .

Every permitted finite routing choice has positive probability, and finite no-merger survival is positive. Thus the displayed support equalities hold pointwise, including equal rates and inheritance one-half, even when A, B overlap. Temporal separation prevents cancellation of distinct biological events.

Lemma 4.1 (Tree-only witnesses). *Every vertex of a rooted tree-child network has a descendant taxon reached by a path whose vertices after the start have indegree one. Incomparable vertices have mutually noncontained descendant clusters. Every tree vertex has at least two descendant taxa.*

Proof. Repeatedly choose an ordinary child. The finite acyclic graph eventually reaches a leaf. An alternative path to that leaf must enter the indegree-one suffix through its start, or start on the suffix. If incomparable vertices shared one of these witness leaves, a root-to-leaf path would contain both and make them comparable. For a tree vertex, its ordinary-child witness cannot be reached through the other child: entering the same suffix would give a directed cycle. A descendant of the other child is therefore a different taxon. $\square$

Consequently every tree vertex is visible and its cluster is the union of the endpoints of its positive jump graph. Every reticulation with at least two sampled descendants is visible with its full cluster. A reticulation with only one descendant sampled label can have neither a tree child (which has two descendant taxa) nor a reticulation child. After degree-two suppression its child is a singly sampled taxon leaf. Contract this reticulation into that leaf, retaining the two incoming parent incidences. Call the graph so obtained H ; multiply sampled taxon leaves remain single graph vertices with their known sample clusters.

For retained internal vertices, ancestry is exactly cluster containment together with decreasing forward age: u precedes v if and only if $C(v) \subseteq C(u)$ and $\tau(u) > \tau(v)$. Containment forces comparability by [lemma 4.1](#); equal clusters are ordered by their distinct observed times. Leaf incidences follow from the known taxon sample clusters. Thus the signed event book determines the complete reachability poset and its transitive reduction.

Lemma 4.2 (Reticulation clusters at level two). *Distinct reticulations have distinct clusters. Incomparable reticulations in a level-at-most-two network have disjoint clusters.*

Proof. For comparable reticulations r preceding r' , a tree-only witness below r cannot be entered from r' without revisiting r ; it belongs to $C(r) \setminus C(r')$. Incomparability precludes equal clusters by [lemma 4.1](#).

If incomparable reticulations had a common descendant, choose root paths with last common vertex b and descendant paths with first common vertex c . Prefix interiors are disjoint, as are suffix interiors. An intersection between one prefix and the opposite suffix would make the two reticulations comparable. The four paths therefore form a simple undirected cycle through the two reticulations and c . The two distinct incoming arcs at c make it a third reticulation on that same cycle. A cycle lies in one blob, contradicting level at most two. This argument uses no observed blob assignment and permits arbitrary nested decorations. $\square$

Lemma 4.3 (Unique shortcut completion). *The reachability poset and signed jump supports determine H , except for the topological reversal of a pendant bi-unary sister triangle.*

Proof. An arc into an indegree-one vertex is a cover: any alternate path must use the same last arc and return to its source. Reticulations have one outgoing cover. A tree vertex retains its ordinary-child cover and loses at most one outgoing shortcut in the transitive reduction. A missing arc targets a retained reticulation or a contracted singleton leaf. The other parent is lower, and its incoming arc is a cover.

A deficient tail v has one ordinary cover child w , with $C = C(v) = C(w)$, and a shortcut head with nonempty proper cluster $D \subset C$. The positive jump graph at v is the complete graph on C with the clique on $C \setminus D$ removed. If $|C \setminus D| \geq 2$, its universal vertices identify D . If the graph is complete, then $D = C \setminus \{a\}$. For $|C| \geq 3$, two distinct candidate heads would both be retained reticulations in the same observed poset, with overlapping equal-size clusters. This contradicts lemma 4.2. The head is unique.

For $|C| = 2$ the two samples belong to different singly sampled taxa. All tree vertices with that cluster form a chain. Every nonbottom same-cluster tree vertex has a singleton-reticulation shortcut child; tree-childness therefore forces its sole cover child to be ordinary. An exposed reticulation may occur above the whole chain, but cannot interrupt it. The bottom tree vertex has the two terminal heads after unary-reticulation contraction. Two upper shortcut tails cannot target one head without giving indegree three, or target different heads without giving the bottom tree vertex two reticulation children. Thus at most one upper deficient tail exists. In that ambiguous case exactly one terminal head is reticulate, with parents the upper and bottom tree vertices; the other terminal is ordinary. Neither admits an additional outside parent. The subgraph is a pendant sister triangle, possibly the whole graph when its top is the root. The two head choices give exactly its bi-unary Q6 reversal. If there is no upper deficient tail, only absence of shortcut ambiguity follows: a terminal can still have an outside parent, whose incidence is already a cover. This distinction is essential. $\square$

These arguments recover all decorated attachment positions directly as vertices, clusters and times. They neither infer an arbitrary larger graph from deleted-port catalogs nor glue unrelated choices from separate marginal carriers.

5 Detailed scalar numerical inversion

Assume the reduced graph has been recovered as above, with a canonical choice in each bi-unary Q6 class. All observations below are on actual sampled subsets of size at most three.

5.1 Ordinary words and oriented weights

For an ordinary arc e (its lower endpoint is a tree vertex or taxon leaf), choose a tree-only descendant witness a . In a pair restriction its unmerged lineage occupies e deterministically throughout the

arc. If another sampled label b can reach e , then the only possible shared population is e , and

$$\lambda_e(t) = f_{ab}(t)/q_{ab}(t), \quad q_{ab}(t) > 0.$$

This recovers the entire finite exposed word, including every unequal-rate boundary, and applies also above the root. A singleton terminal arc has no coalescent-rate contribution.

For a reticulation v , let its parent ages be $a_1 < a_2$. Choose a tree-only witness a below v and a tree-only witness z in the ordinary side of the earlier parent. The two paths cannot meet below a_1 . If γ is the route weight toward that parent,

$$\gamma = q_{az}(a_1+) - q_{az}(a_1-) = q_{az}(a_1+).$$

The other reticulation arm is isolated until a_2 . This formula also recovers the typed Q8 weight when the reticulation has a single sampled descendant and its age is hidden.

5.2 Two incoming arms before the earlier parent

For an exposed reticulation of age h , choose a different descendant b . Put $q = q_{ab}(h-) > 0$ and $q_3 = q_{ab|z}(h-) > 0$, with no prior merger in the respective restrictions. Since a is deterministically in the reticulation child and z is on its disjoint ordinary side, the grouping mass q_3 is exactly the child-pair mass needed below; [lemma 3.2](#) recovers it without route labels.

For $h < t < a_1$ and sufficiently small $u > 0$, let $D(t, u)$ be the actual joint density of first merger ab at t and second merger $(ab)z$ at $a_1 + u$. Then

$$D(t, u) = q_3 \gamma^2 \lambda_1(t) e^{-\int_h^t \lambda_1(s) ds} \lambda_o e^{-\lambda_o u},$$

where λ_o is the known ordinary rate just above the earlier parent. The other arm gives zero density for this second merger before a_2 ; neither incoming arm admits a side entry before its parent. The earlier arbitrary history is integrated once into q_3 .

Thus

$$g_1(t) = \frac{\lim_{u \downarrow 0} D(t, u)}{q_3 \gamma^2 \lambda_o}, \quad A_1(t) = 1 - \int_h^t g_1(s) ds, \quad \lambda_1(t) = g_1(t)/A_1(t).$$

The denominators are strictly positive, with $A_1(t) = e^{-\int_h^t \lambda_1}$. The pair density before a_1 then gives

$$g_2(t) = \frac{f_{ab}(t)/q - \gamma^2 g_1(t)}{(1 - \gamma)^2}, \quad A_2(t) = 1 - \int_h^t g_2(s) ds, \quad \lambda_2(t) = g_2(t)/A_2(t).$$

These formulas link entire words to physical parent incidences even at equal rates and inheritance one-half.

5.3 The longer arm and the order of inversion

Recover all ordinary words and all inheritance weights first. Process reticulations from ancestors to descendants within a blob. For a reticulation with no preceding reticulation in its blob, the earlier-parent lineage follows a known deterministic chain of ordinary populations until age a_2 . Any reticulation in a preceding blob is above the current blob's entry, whose age is at least a_2 .

Write $q_2(a_1) = q(1 - \gamma)^2 A_2(a_1)$ for the known both-later-arm mass. The route-one co-occupied mass starts at

$$q_1(a_1+) = q_{ab}(a_1+) - q_2(a_1).$$

This includes histories in which b bypassed the reticulation child and entered at the earlier parent. On each ordinary chamber, $q'_1 = -\lambda_{\text{path}} q_1$. At every tree join on that path it receives the whole observed positive jump of q_{ab} , including later bypass entries. The isolated longer arm receives none. An off-path reticulation cannot move a lineage into the already existing ordinary arc: its parent arcs are new at its own boundary. Thus finite scalar propagation determines q_1 , and

$$q_2(t) = q_{ab}(t) - q_1(t) > 0, \quad \lambda_2(t) = -q'_2(t)/q_2(t) \quad (a_1 < t < a_2).$$

If another reticulation precedes v in its level-two blob, v is the lower sink of the comparable-reticulation theta cases. Its child arc is a bridge, as follows directly. If its sole outgoing edge $v \rightarrow w$ remained in that blob B , biconnectivity would give an undirected path in B from w to a parent of v , avoiding v . Let D be the vertices reachable from v . This path starts in D and ends outside D . Its first exit edge must point from outside D into D , since a directed edge leaving D contradicts reachability. It enters $c \in D$, distinct from v ; a directed path from v to c supplies another incoming edge from inside D . Thus c is a third reticulation in B , besides v and its predecessor, a contradiction. Therefore the child edge is a bridge, and every root-to-descendant path crosses it and v . Every descendant lineage reaches the child before h in every surviving pair history. The both-route-one mass at a_1 is exactly $q\gamma^2 A_1(a_1)$. Starting from that known mass in the ordinary population above the earlier parent, propagate the known two-lineage process through the already recovered preceding reticulation and ordinary words until a_2 . This computes its co-occupied contribution $q_1(t)$. Split assignments still cannot meet while one lineage is isolated on the longer arm. The same subtraction recovers q_2 and its full rate word. This is forward propagation after parameter recovery, not inversion of unknown hidden columns. No circular dependence or additional observed labels is introduced.

6 The quotient soundness and exact obstructions

Lemma 6.1 (Contextual unary-fan response). *On the theorem class a nontrivial unexposed reticulation has one singly sampled terminal descendant. Its maximal pre-exposure fan consists of the unary path into that reticulation and its two incoming arms, stopped at their two parent tree vertices. Its complete contextual response retains exactly the two typed parent outlets with weights $(\gamma, 1 - \gamma)$; the internal age and all singly occupied rate words are absent.*

Proof. The singleton-descendant assertion follows from lemma 4.1 and tree-childness. Both parent tree vertices have ordinary-side sampled witnesses, so every route first becomes capable of co-occupancy at its parent vertex. Until then there is one sampled block and hence no Kingman merger or survival loss. Summing its routes gives precisely the stated two weighted exits. Cutting a complete marked genealogy at those exits proves equality in every fixed surrounding context. Replacing one valid strict realization by another with the same response is Q8; a one-outlet unary corridor gives Q4 by the same argument with total weight one. No observable operator selecting the past route is needed. $\square$

Example 6.2 (Hidden-clock collision). *Take $R \rightarrow S, w, S \rightarrow T, X, T \rightarrow z, X, X \rightarrow x$, where X is reticulate and $0 < h = \tau(X) < a = \tau(T) < d = \tau(S) < r = \tau(R)$. This is a strict strong-tree-child triangle with a root-side sample w . If the route to T has weight γ and the exposed donor interval (a, d) has constant rate λ , then*

$$f_{xz}(a + u) = \gamma \lambda e^{-\lambda u}, \quad \Pr(T_{xz} > d) = 1 - \gamma + \gamma e^{-\lambda(d-a)}.$$

The entire contextual law is independent of $h \in (0, a)$ and of the unary words below the two exits, by [lemma 6.1](#). Distinct hidden clocks therefore refute the former Q1–Q7 quotient and explain the necessity of Q8.

Lemma 6.3 (The whole bi-unary donor word). *For a sister triangle with one sampled block on each side, first contact a , reconvergence d , joining weight γ and finite positive donor word λ_D , the local merger density and unmerged/merged transition matrix are*

$$f(a+u) = \gamma \lambda_D(a+u) e^{-\Lambda_D(u)}, \quad K = \begin{pmatrix} 1 - \gamma + \gamma e^{-\Lambda_D(d-a)} & \gamma(1 - e^{-\Lambda_D(d-a)}) \\ 0 & 1 \end{pmatrix},$$

where $\Lambda_D(u) = \int_0^u \lambda_D(a+s) ds$. Transferring this entire word and the same joining weight upon reversal preserves every marked cell in every common context.

Proof. The pair joins the donor population with probability γ , then has the stated inhomogeneous Kingman density; otherwise it remains separate until d . There are no other blocks inside the bi-unary pendant module. At d , each history has exactly the same labelled partition in the same ancestral population in the two directions. All upper continuation factors therefore agree. Moving only one constant-rate value would not preserve a multi-epoch density; every retained breakpoint and value must transfer. $\square$

Example 6.4 (Rank one on a strict nonbridge). *In a minimum strong cycle repair put two child labels x, y below the reticulation, a side label on one parent arm, inheritance $1/2$, and equal arm rates λ until the side-entry time a . For every cut $t \in (h, a)$, all prefixes with one current block give parent posterior $(1/2, 1/2)$, while the two-block partition has only its no-merger prefix. Thus the feasible partition-indexed prefix-suffix kernels have rank one throughout this nonbridge interval. Yet the pair survival germ is*

$$G_t(u) = \frac{1 + e^{-\lambda(t-h+u)}}{1 + e^{-\lambda(t-h)}},$$

which has a positive exponent-zero atom. The signed-jump proof reconstructs the reticulation from co-occupancy and does not use rank one as a bridge converse.

Example 6.5 (Why a historical route projector cannot descend). *After the two routes of the preceding hidden-clock triangle reconverge at d , a no-merger history from either route has the same current population, labelled partition and future law. Write their formal coordinates as e_0, e_1 . The observable future map sends both to the same residual function, so $e_0 - e_1$ is in its kernel. The historical route selector $Pe_0 = e_0$, $Pe_1 = 0$ sends that difference to e_0 , which is not in the kernel. Equivalently,*

$$b = (1, 1), \quad v = (1, -1)^\top, \quad P = \text{diag}(1, 0), \quad bv = 0, \quad bPv = 1.$$

The selector therefore does not descend to the minimal future quotient. Retaining formal tags cannot repair kernel preservation after they are identified by future behavior. The present proof instead uses current scalar masses, literal merger densities, and then forward propagation of a recovered graph.

7 The five directed cores and twelve repairs

Proposition 7.1 (Graph-axiom core exhaustion). *After pendant bridge components are replaced by labelled boundary ports and ordinary unported degree-two vertices are suppressed, every nontrivial*

binary rooted blob of level at most two is a directed cycle or one of four directed theta cores. Its full strong expansions contain exactly the twelve segment-indexed inclusion-minimal repair sets displayed below; these form eight orbits under the corresponding role-preserving core automorphisms.

Proof. Forget ordinary directions in one blob while retaining its distinguished incoming boundary. The resulting connected multigraph is biconnected and has minimum degree two after suppression. With cyclomatic number $\beta = |E| - |V| + 1$, the identity

$$\sum_v (\deg v - 2) = 2|E| - 2|V| = 2(\beta - 1)$$

shows that $\beta = 1$ gives a cycle. For $\beta = 2$, binarity forces exactly two degree-three vertices; removing them leaves three internally disjoint paths with the same endpoints. Biconnectedness excludes a figure eight and two cycles joined at a bridge, so the kernel is theta.

Retain the unique incoming boundary and every degree-two vertex carrying a root or reticulation role, and orient the three paths. Impose: acyclicity; one boundary source; binary tree degrees; exactly two indegree-two/outdegree-one reticulations; distinct reticulation parents in the expanded raw graph; and no directed two-cycle. Boundary-preserving isomorphism leaves the four theta orbits whose primitive arc lists are recorded below. This check is exhaustive: each path direction is fixed between consecutive retained role vertices, and every remaining assignment either contains a directed cycle, changes the number of indegree-two vertices, or violates a binary degree. The analogous one-cycle check gives its unique directed orbit.

An independent exact certificate implements this argument without importing the displayed catalog. It constructs explicit undirected **MultiGraph** theta kernels whose three path lengths range independently over 1, 2, 3, assigns every edge orientation, and filters for a directed acyclic graph with the distinguished root boundary, binary degrees, and exactly two reticulations. It then suppresses every indegree-one/outdegree-one ordinary vertex and quotients the surviving objects by exact boundary- and role-preserving **MultiDiGraph** isomorphism (not a hash or Weisfeiler–Leman surrogate). The stable output has exactly four theta orbits. To justify the path-length bound, after ordinary degree-two suppression an internal path vertex can only be the unique source (two outgoing path edges) or a reticulation sink (two incoming path edges, with its child port leaving the blob). Two internal sinks on one path would force the unique source between them. Both paths from that source would terminate at those sinks, making the theta endpoints unreachable, contrary to the unique-source DAG condition. Thus a primitive path has at most one internal sink and the one source, hence at most two retained internal vertices and length at most three. Degree-three reticulation roles are kernel endpoints. Longer representatives therefore add only suppressible subdivisions.

For repair exhaustion, a subset R of the m primitive segments is processed without using any scalar data: subdivide every segment in R by a tree vertex with one sampled-leaf child; attach every compulsory path-sink child; restore the incoming boundary; and test rooted tree-childness and the strong local-tail predicate. The latter counts as ordinary every incident edge except an edge whose head is a reticulation and requires exactly the two ordinary incidences forced by binary raw degree at every reticulation-arrow tail. Enumerating all 2^m subsets for $m = 2, 5, 5, 6, 6$ and deleting each satisfying set that contains a smaller satisfying set yields

cycle	$\{0\}, \{1\}$
θ_0	$\{2, 3\}, \{3, 4\}$
θ_1	$\{2, 3\}, \{2, 4\}$
θ_2	$\{2, 3\}, \{2, 5\}, \{3, 4\}, \{4, 5\}$
θ_3	$\{2\}, \{4\}$.

The independent certificate performs this second stage on the exact multidigraph representatives: it enumerates every primitive-port subset, requires both rooted tree-childness and the strong local-tail predicate, and then minimizes under inclusion before exact role-preserving isomorphism. The segment-indexed raw-row counts are 2, 2, 4, 2 for $\theta_0, \theta_1, \theta_2, \theta_3$, respectively and 2 for the cycle, hence twelve. Exact role-preserving quotienting gives orbit counts 2, 1, 2, 2 for the four theta classes and 1 for the cycle, hence eight. Every larger strong expansion contains an inclusion-minimal satisfying subset; after it is chosen, all remaining port-bearing vertices are ordered words on the primitive segments. This proves exhaustion and the count twelve directly from the graph predicates. $\square$

Core	Ordered primitive arcs (indices start at zero)
cycle	$S \rightarrow X, S \rightarrow X$
θ_0	$S \rightarrow U, S \rightarrow V, U \rightarrow X, V \rightarrow X, U \rightarrow V$
θ_1	$S \rightarrow U, S \rightarrow X, V \rightarrow X, U \rightarrow V, U \rightarrow V$
θ_2	$S \rightarrow U, S \rightarrow V, U \rightarrow X_0, V \rightarrow X_0, U \rightarrow X_1, V \rightarrow X_1$
θ_3	$S \rightarrow U, S \rightarrow X_0, V \rightarrow X_0, U \rightarrow X_1, V \rightarrow X_1, U \rightarrow V$

The arc-order table must be interpreted with the catalog’s compulsory child ports at reticulation sinks and its distinguished incoming boundary. Parallel segments occur only in the suppressed core; the full raw repairs have distinct reticulation parents. The main inverse proof also has a direct graph argument and does not depend on importing a precomputed observable catalog.

8 Small marginals and consistency across them

The graph event book uses the collection of all pair laws on the fixed labelled set S . Its clusters are subsets of this same set, and tree ages are absolute normalized times. Thus all pairs determine one common chronological poset before any numeric witness is selected. Each witness set used for ordinary rates, inheritance or arm words has size at most three and is selected from that recovered graph. All such laws belong to the supplied small-marginal collection by sampling consistency. Numerical formulas are uniquely determined on their positive domains. They therefore attach to one graph and cannot select incompatible carriers independently in different triples.

Once the complete graph-and-parameter quotient has been recovered, its entire full-sample law is given by ordinary finite-state NMSC forward propagation and the Q4/Q8 response kernels. This proves that the collection of all marginals on at most three sampled lineages determines the complete Q1–Q8 quotient and full law. Every larger indexed family also suffices, including the eight-lineage family. No minimality is claimed. The finite proof does not reconstruct each triple separately and then presume they glue. No locus-length or finite-data stability bound follows from analytic continuation of an exact metric-genealogy law.

9 Verification and limits of the evidence

The universal graph and numerical arguments above cover arbitrary finite decorations and rate words. Separate exact checks exercise signed-jump supports on all twelve minimum repairs and on specified single-port expansions, and test scalar identities on strict examples with rates in positive integer multiples of $\log 2$ and integer durations. Their manifests declare finite enumeration domains, isomorphism conventions, event-order choices and counterexample controls. Favorable finite counts are not an exhaustion of arbitrary decorations or of all positive parameters.

The documented validation contains 90 exact tests and 57 executable mutation records in 56 conceptual groups. Its fixed-(X,S,s) exact structural regression is not a universal proof. Separate

raw-event computations check the scalar inverse formulas; independent source reviews and exact-recomputation records accompany the manuscript. The route-information counterexample above illustrates why finite computations must be interpreted at their precise mathematical scope. No numerical tolerance is used as a proof of equality.

The theorem concerns actual normalized metric-genealogy laws of supplied positive models, not their Euclidean or Zariski closures. It does not address level three, simultaneous biological events, nonreversible substitution, recombination, unknown locus boundaries, arbitrary demographic mixtures, or stable practical reconstruction from finitely many noisy genealogies. The companion study ([doi:10.5281/zenodo.22729537](https://doi.org/10.5281/zenodo.22729537)) provides separate compact-semianalytic finite-determination and quotient-consistency results with exactly their stated sequence-law consequences.