TYPED AMPLITUDE COMPOSITION FOR REUSABLE INFERENCE: SEMANTICS, ERROR BOUNDS, AND EXACT CORRECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Probabilistic representations are often built before the eventual query or constraint is known. We study inference-time composition: compositional inference that modifies an already represented law while preserving its intended semantics and reusing prior computation. Algebraically natural operations can encode the wrong law even in exact arithmetic, and approximation, target mismatch, and Monte Carlo error belong to different layers. We develop one contract that keeps these layers explicit. Typed semantics fix the intended law for products of complete laws, factors relative to a shared base, labeled mixtures, event projectors, and integer tempering; explicit matrix-product-state (MPS/TT) constructions realize it; deterministic bounds propagate compression residuals, and a dimension-free amplitude-to-law inequality converts them into total-variation control of the represented law; exact density-ratio correction targets a separately declared law under support assumptions, and one scored pool answers later conditions without new target evaluations. On held-out known-answer families, composition recovers the requested law to total variation 7.4×10^{-16} and propagated bounds cover every observed error. In RNA design, where later sequence constraints are imposed on candidates compatible with a target structure, one compiled proposal per puzzle is reused for later conditions, composed with exact conditions, and repays all measured setup by 20,000 samples on a workload with 2.6×10^{17} compatible states. At 256 requested ViennaRNA scores it estimates a prespecified conditional endpoint more accurately than tested annealed SMC (equal-puzzle difference $+0.142$, interval $[0.040, 0.244]$), and its favorable pool quality transfers to EternaFold.

1 INTRODUCTION

Many inference systems operate across an amortization boundary: recurring structure is compiled once, while task-specific factors and constraints remain open until the problem is known. This reflects the broader view of generalization by construction, in which inference recombines reusable components rather than requiring a separately trained model for every task (Du, 2026). We study the probabilistic contract required at this boundary, which we call *inference-time composition*: what law does a composition define, how does representation error propagate, and when can exact correction and prior computation be reused?

The first difficulty is semantic rather than computational. Algebraically natural operations can define a different law from the one intended. Multiplying separately normalized laws that share a nonuniform base counts the base twice; adding unnormalized component energies changes requested mixture weights through their partition functions (Du et al., 2023); and renormalizing conditioned mixture branches while keeping their prior weights ignores the evidence-dependent selector posterior. These failures occur even in exact arithmetic, so a more accurate representation, such as a higher tensor rank, cannot repair them: it only reproduces the wrong law more faithfully.

The second difficulty is error attribution. A compression residual does not certify a scientific target, omitted support, importance-weight overlap, or finite Monte Carlo error. When a compact representation is used as a proposal, exact Metropolis–Hastings or importance ratios target a separately

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.

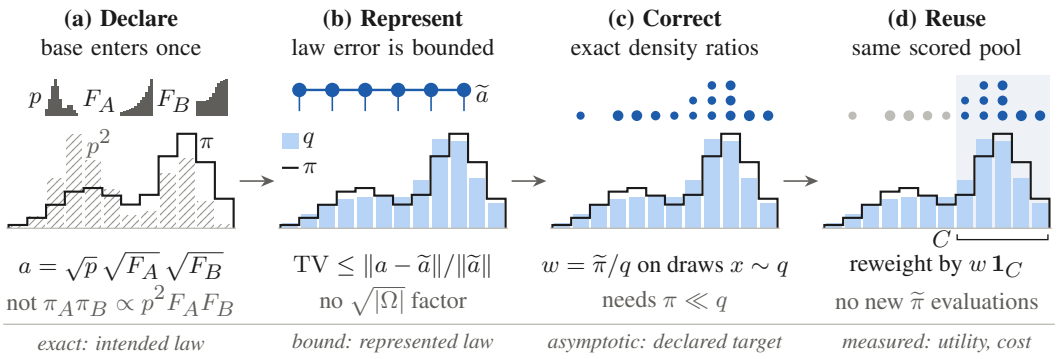

Figure 1: **One contract carries a declared law from composition to later reuse.** (a) The typed amplitude $\sqrt{p}\sqrt{F_A}\sqrt{F_B}$ squares to the declared law $\pi \propto pF_AF_B$ (outline); multiplying separately normalized controlled laws counts the shared base twice (hatched). (b) A rank-truncated MPS/TT, one possible backend, represents $q = \rho[\widetilde{a}]$; the Fuchs–van de Graaf bridge bounds its error without a $\sqrt{|\Omega|}$ factor. (c) Exact ratios $\widetilde{\pi}/q$ (dot area) recover the declared target asymptotically when $\pi \ll q$. (d) A later condition C reweights the same scored pool without new target evaluations. The bottom row states what each layer guarantees; overlap, finite accuracy, and cost are measured. Illustrative twelve-state example.

evaluable density, but finite performance remains governed by overlap (Hastings, 1970; Liu, 1996; Mengersen & Tweedie, 1996).

We address both difficulties with an integrated contract (Figure 1). Typed semantics specify the intended probability law; deterministic error bounds control changes to the represented object; exact density-ratio correction recovers a separately declared target under support assumptions; and held-out measurements separately establish finite-sample utility and cost. Our computational object is the square root of a nonnegative target, an *amplitude*, stored as a matrix product state (MPS), also called a tensor train (TT) (Oseledets, 2011; Han et al., 2018), and squared to obtain probabilities. Amplitudes are useful here not because the computation is quantum—it is entirely classical—but because squaring converts an unconstrained low-rank representation into a nonnegative law, while the geometry of normalized amplitudes converts Euclidean compression error into dimension-free probability error (Fuchs & van de Graaf, 1999).

Our contributions are a typed semantic interface for complete-law products, shared-base factors, labeled mixtures with retained selectors, event projectors, and integer tempering, with explicit MPS and weighted-automaton constructions (Section 3.1); a represented-object error account comprising deterministic stability bounds, the dimension-free amplitude-to-law bridge built on the Fuchs–van de Graaf inequality, and a sharper orthogonal case for one exact TT-SVD sweep (Section 3.2); exact correction and reuse results, in which conditioning a scored pool changes the density ratio only by a cancelling constant and exact projection reproduces rejection with an explicit amortization condition (Section 3.3); and experiments that measure what the theory leaves open: a held-out known-answer study validating typed semantics, normalization, and bound coverage (Section 4), and RNA design, where later sequence constraints are imposed on candidates compatible with a target structure (Anderson-Lee et al., 2016; Lorenz et al., 2011), demonstrating reusable conditioning, exact event composition, correction, and amortization (Section 5). Each ingredient is prior art; the contribution is their integration into one auditable framework.

2 Related Work

Tensor networks and tractable probabilistic models. MPS/TT decompositions, rounding, and canonical sampling are established (Oseledets, 2011; Han et al., 2018; Ferris & Vidal, 2012), as are nonnegative TT density models (Novikov et al., 2021) and tensor contraction for weighted model counting (Dudek et al., 2020). Tractable probabilistic circuits support broad algebras of sums, products, powers, and conditioning (Vergari et al., 2021). Squared circuits contain Born MPS as a chain-structured case and can be exponentially more compact than structured monotone circuits on specific

families (Loconte et al., 2024), and positive semidefinite models give another nonnegative density family with closed-form products and marginals (Rudi & Ciliberto, 2021). We claim no new representation family: MPS/TT is one backend, chosen because contraction, exact proposal evaluation, and TT rounding make every layer computable.

Compositional generation. Composable diffusion and energy-based models combine component scores or energies to generate unseen combinations of concepts (Liu et al., 2022; Du et al., 2023), and Du et al. (2023) note that mixing energy-based models requires the ratio of component normalizers. We focus on typed declarations with contracted normalizers and compression error tracked through composition.

Compiled proposals and exact correction. Compiling an affordable approximation and correcting it exactly is an established continuum: partial knowledge compilation with importance sampling (Friedman & Van den Broeck, 2018; Lai, 2018), tensor-train importance proposals (Cui et al., 2024), and quantum-enhanced or quantum-inspired proposals followed by Metropolis acceptance (Layden et al., 2023; Christmann et al., 2025) all rely on exact acceptance or weighting (Hastings, 1970; Liu, 1996). We add represented-object error bounds and reuse of one scored pool across later conditions.

Constrained sequential generation. Tractable models can impose lexical constraints on autoregressive generation (Zhang et al., 2023), and twisted or programmable SMC steers sequence models toward constraint- or reward-defined targets (Zhao et al., 2024; Loula et al., 2025). These methods condition on a target known before sampling; our setting compiles a proposal before later conditions are declared and then reuses or composes it. SMC samplers (Del Moral et al., 2006) and parallel tempering (Earl & Deem, 2005) serve here as baselines and references.

RNA design. RNA inverse folding, including ViennaRNA's optimizer, seeks sequences that fold into a target (Lorenz et al., 2011; Anderson-Lee et al., 2016); we instead estimate conditional expectations under a declared distribution, which support reusable design-campaign statistics rather than a single designed sequence. EternaFold provides an independently parameterized folding model for cross-model evaluation (Wayment-Steele et al., 2022).

3 Typed composition and its error contract

Let $\Omega = \Sigma_1 \times \cdots \times \Sigma_n$ be finite. A nonzero amplitude $a : \Omega \to \mathbb{R}$ represents the law $\rho[a](x) = a(x)^2/\|a\|_2^2$, so a nonnegative unnormalized target $\bar{\pi}$ is represented by $a = \sqrt{\bar{\pi}}$. An MPS stores $a(x) = A_1^{x_1} \cdots A_n^{x_n}$ with matrices $A_k^{x_k} \in \mathbb{R}^{\chi_{k-1} \times \chi_k}$, $\chi_0 = \chi_n = 1$, whose bond dimensions χ_k govern storage and contraction; TT-SVD rounding and exact sequential sampling from a canonical MPS are standard (Oseledets, 2011; Han et al., 2018; Ferris & Vidal, 2012).

3.1 Typed semantics and constructions

Composition must declare whether each input is a complete law, a control relative to a shared base, a mixture component, or an event. The same informal operator otherwise denotes different targets.

Proposition 1 (Typed composition semantics)**.** *Let $\pi_A = \rho[\psi_A]$ and $\pi_B = \rho[\psi_B]$ be laws on Ω.*

1. Complete laws. *If $\sum_y \pi_A(y)\pi_B(y) > 0$, then $\rho[\psi_A\psi_B] = \pi_A\pi_B / \sum_y \pi_A(y)\pi_B(y)$.*
2. Shared base. *For a normalized base p and factors $F_A, F_B \geq 0$ with $\sum_x pF_AF_B > 0$, the amplitude $\sqrt{p}\sqrt{F_A}\sqrt{F_B}$ represents $\propto pF_AF_B$, whereas the product of the separately normalized laws $\propto pF_A$ and $\propto pF_B$ is $\propto p^2F_AF_B$.*
3. Labeled mixtures. *For weights $w_i \geq 0$ with $\sum_i w_i = 1$ and nonzero ψ_i, the amplitude $\Psi(i,x) = \sqrt{w_i}\,\psi_i(x)/\|\psi_i\|_2$ with an orthogonal label i has x-marginal $\sum_i w_i\rho[\psi_i]$. After multiplication by $\sqrt{H}$ for a factor $H \geq 0$ with $\sum_j w_j Z_{j,H} > 0$, where $Z_{i,H} = \sum_x \rho[\psi_i](x)H(x)$, the label posterior is $w_i Z_{i,H} / \sum_j w_j Z_{j,H}$.*
4. Events and tempering. *If $\psi\mathbf{1}_C \neq 0$, then $\rho[\mathbf{1}_C\psi] = \rho[\psi](\cdot \mid C)$; for integer $m \geq 1$, $\rho[\psi^m] \propto \rho[\psi]^m$.*

Each identity depends only on amplitude magnitudes, so it is unaffected by the sign changes that truncation can introduce.

Component norms $\|\psi_i\|_2$ are semantic and must be recovered by contraction, not inferred from supplied scales. For energies with $e^{-\beta E_i} = Z_i \pi_i$, the unnormalized sum $w_A e^{-\beta E_A} + w_B e^{-\beta E_B}$ induces weight $w_A Z_A/(w_A Z_A + w_B Z_B)$ rather than w_A (Proposition 9); this concerns the displayed operation only, and reusable compositional energy models already recognize the normalizer ratio (Du et al., 2023). The shared-base interface assumes access to the evaluable base and bare factors. Fractional powers of signed amplitudes are outside the claim.

Proposition 2 (Explicit MPS constructions)**.** *In a fixed common variable order, pointwise products, labeled mixtures, event projectors, and integer powers have exact MPS representations with* $\chi(\psi_A \psi_B) \le \chi_A \chi_B$, $\chi(\Psi_{A \vee B}) \le \chi_A + \chi_B$, $\chi(\mathbf{1}_C \psi) \le \chi_C\, \chi(\psi)$, *and* $\chi(\psi^m) \le \chi(\psi)^m$, *where* χ_C *is the bond of an exact indicator MPS for* C*. A condition recognized by a weighted finite-state automaton with* R *states has an exact indicator MPS with* $\chi_C \le R$*.*

The constructions use Kronecker-product and block-direct-sum cores; tractable circuits support related operation algebras (Vergari et al., 2021; Loconte et al., 2024). The automaton form matters for global conditions: a GC-content window is recognized by a deterministic automaton whose state is the running G/C count, which Section 5 composes exactly on a non-enumerable RNA space (Appendix B). These are construction upper bounds, not tightness or affordability guarantees. The identities remove errors of meaning, but they do not control the numerical error introduced when the resulting object is compressed.

3.2 Deterministic stability and the dimension-free amplitude-to-law bridge

The represented object is built by a sequence of exact factor multiplications, each followed by rounding. Let D_f denote pointwise multiplication by an exact factor f, and write $\psi_k = D_{f_k} \psi_{k-1}$ and $\widetilde{\psi}_k = R_k(D_{f_k} \widetilde{\psi}_{k-1})$ for rounding maps R_k with certified local residuals $\|D_{f_k} \widetilde{\psi}_{k-1} - \widetilde{\psi}_k\|_2 \le e_k$.

Theorem 3 (Deterministic stability)**.** *For finite amplitudes with* $\|a - \widetilde{a}\|_2 \le E_a$ *and* $\|b - \widetilde{b}\|_2 \le E_b$, $\|ab - \widetilde{a}\widetilde{b}\|_2 \le \min\{\|b\|_\infty E_a + \|\widetilde{a}\|_\infty E_b,\ \|a\|_\infty E_b + \|\widetilde{b}\|_\infty E_a\}$*, and* $\|fa - f\widetilde{a}\|_2 \le \|f\|_\infty E_a$ *for an exact factor* f*. For the sequential construction with* $L_k = \|f_k\|_\infty$ *and* $\|\psi_0 - \widetilde{\psi}_0\|_2 \le E_0$*,*

$$\|\psi_m - \widetilde{\psi}_m\|_2 \le E_m := E_0 \prod_{j=1}^{m} L_j + \sum_{k=1}^{m} e_k \prod_{j=k+1}^{m} L_j. \tag{1}$$

Later factors amplify earlier residuals, so the order of composition enters the account. The relevant approximation is Euclidean at the amplitude level, while the downstream object is a probability law; the next result connects the two without paying for the size of the state space.

Proposition 4 (Dimension-free amplitude-to-law bridge)**.** *For nonzero finite amplitudes* $a, \widetilde{a}$ *on the same encoded support, with* $u = a/\|a\|_2$ *and* $v = \widetilde{a}/\|\widetilde{a}\|_2$*,*

$$\|\rho[a] - \rho[\widetilde{a}]\|_{\mathrm{TV}} \le \sqrt{1 - |\langle u, v \rangle|^2} \le \min\Big\{1, \frac{\|a - \widetilde{a}\|_2}{\|\widetilde{a}\|_2}\Big\}. \tag{2}$$

The normalized amplitudes are treated as pure states. Their trace distance is the middle term of Eq. (2) (unsquared overlap convention); the Fuchs–van de Graaf inequality relates fidelity and trace distance, and a fixed-basis measurement cannot increase trace distance (Fuchs & van de Graaf, 1999). The second inequality compares $\widetilde{a}$ with the one-dimensional linear subspace spanned by a. The cardinality $|\Omega|$ never appears, whereas converting an ℓ_2 error between density tables to ℓ_1 by Cauchy–Schwarz incurs a $\sqrt{|\Omega|}$ factor. This is the central geometric benefit of the amplitude lift: the norm used by tensor compression yields dimension-free control of the represented distribution. Squaring also makes probability factors enter once: multiplying a target by $F \ge 0$ multiplies its amplitude by $\sqrt{F}$, and one copy of $\sqrt{p}$ yields one copy of p in shared-base composition. Combining Eq. (1) with Eq. (2) bounds the final represented law by $\min\{1, E_m/\|\widetilde{\psi}_m\|_2\}$. Because $|\mathbb{E}_p g - \mathbb{E}_{\widetilde{p}} g| \le \|p - \widetilde{p}\|_{\mathrm{TV}}$ for every $g : \Omega \to [0, 1]$ (Corollary 12), the bound also controls every bounded observable of the same represented law. It does not control χ^2, the maximum density ratio, or

importance ESS against any law: Proposition 13 gives full-support pairs with arbitrarily small TV and arbitrarily large χ^2.

The preceding calculus applies to general local residuals; an exact sequential TT-SVD sweep has additional orthogonality and therefore admits a sharper account.

Theorem 5 (Exact-sweep TT-SVD). *Let $\|\psi\|_2 = 1$. In one exact left-to-right TT-SVD sweep, let r_b^2 be the squared singular values discarded from the current residual at bond b, and let $\phi \neq 0$ be the retained tensor. With $\delta^2 = \sum_b r_b^2$, $\|\psi - \phi\|_2^2 = \delta^2$, $\langle \psi - \phi, \phi \rangle = 0$, and $\|\rho[\psi] - \rho[\phi]\|_{\mathrm{TV}} \leq \min\{1, \delta\}$.*

Discarded blocks lie in orthogonal complements of later retained subspaces, so tails add in quadrature rather than by the triangle inequality; this is the standard TT-SVD geometry (Oseledets, 2011), specialized through Proposition 4. Interleaved factor multiplication, repeated recompression, and cross-approximation fits do not share the nesting argument and use Eq. (1) instead. Appendix C proves all four results.

3.3 Exact correction and reusable conditioning

Closeness to the represented pre-truncation object still does not imply closeness to a separately declared target π, so target correctness requires a different mechanism. If the represented law q is only a proposal but $\pi \propto \widetilde{\pi}$ is evaluable, draws $x_i \sim q$ with weights $w_i = \widetilde{\pi}(x_i)/q(x_i)$ give the self-normalized estimator $\widehat{\mu} = \sum_i w_i g(x_i) / \sum_i w_i$, which is consistent for $\mathbb{E}_\pi g$ when $\pi \ll q$; independence Metropolis–Hastings with the same ratios has π as its limiting law (Hastings, 1970; Liu, 1996). Truncation does not preserve support entrywise, so we mix in a support-covering base.

Corollary 6 (Defensive support). *If $\pi \ll q_0$ and $q_\alpha = (1-\alpha)q + \alpha q_0$ with $0 < \alpha \leq 1$, then $\pi \ll q_\alpha$ and $\|\pi - q_\alpha\|_{\mathrm{TV}} \leq (1-\alpha)\|\pi - q\|_{\mathrm{TV}} + \alpha$.*

Support is an eligibility condition, not overlap: the population ESS fraction $(\mathbb{E}_q w)^2 / \mathbb{E}_q[w^2] = 1/(1+\chi^2(\pi\|q))$ depends on relative tails that neither TV nor defensive mixing controls. Once exact ratios and support are available, the same scored proposal can answer later conditional questions without reevaluating the target.

Proposition 7 (Reusable conditional correction). *Let $\pi = \widetilde{\pi}/Z$ with $\pi \ll q$, draw $x_1, \ldots, x_N$ independently from q, and set $w_i = \widetilde{\pi}(x_i)/q(x_i)$. Let C be an event with $\pi(C) > 0$. With $\pi_C = \pi \mathbf{1}_C / \pi(C)$ and $q_C = q\mathbf{1}_C/q(C)$, on C*

$$\frac{\pi_C(x)}{q_C(x)} = \frac{q(C)}{\pi(C)} \frac{\pi(x)}{q(x)}.$$

Hence, for any fixed finite collection of events $C_1, \ldots, C_K$ with $\pi(C_k) > 0$ and observables g_k evaluable from stored samples with $\mathbb{E}_\pi|g_k| < \infty$, the estimators $\widehat{\mu}_{k,N} = \sum_i w_i \mathbf{1}_{C_k}(x_i) g_k(x_i) / \sum_i w_i \mathbf{1}_{C_k}(x_i)$, each defined when its denominator is positive, are computed from one scored pool without further target evaluations and are jointly consistent for $\mathbb{E}_\pi[g_k \mid C_k]$ as $N \to \infty$.

The constant $q(C)/\pi(C)$ cancels under self-normalization. The statement is asymptotic: a finite pool may contain no particle in C, and rare conditions or concentrated weights degrade conditional ESS. These finite-pool quantities are what Section 5 measures. When a condition is exactly representable, it can instead be composed into the proposal.

Corollary 8 (Exact projection versus rejection). *If $q(C) > 0$, sampling $\rho[\mathbf{1}_C\sqrt{q}]$ and rejection sampling from q produce the same conditioned law q_C. Rejection needs $N/q(C)$ base draws in expectation for N accepted samples. With composition setup cost S_C, per-sample composed cost c_C, and base-draw cost c_q, composition is cheaper exactly when $S_C + Nc_C < Nc_q/q(C)$, that is, when $c_q/q(C) > c_C$ and $N > S_C/(c_q/q(C) - c_C)$.*

This concerns the represented law; rounding error in the composed object remains governed by Theorem 3 and Proposition 4. Appendix D proves the correction results and the separation between TV and χ^2. Table 1 summarizes the contract: different mechanisms control different claims.

The theory fixes semantics, error scope, and correction eligibility; the experiments determine whether overlap, finite accuracy, reuse, transfer, and cost are favorable on tested workloads.

Table 1: Scope of guarantees and empirical evidence.

Layer	Controlled by	Establishes	Does not establish
Target semantics	Typed composition, contracted normalizers	Intended encoded law	Scientific validity
Representation	Sequential or exact-sweep TT-SVD bound	Error of the represented object	Fit, omitted support, target overlap
Exact correction	Support and evaluable density ratios	Declared target, asymptotically	Finite-sample accuracy or efficiency
Utility and cost	Held-out experiments	Performance on tested instances and budgets	Universal advantage or scaling

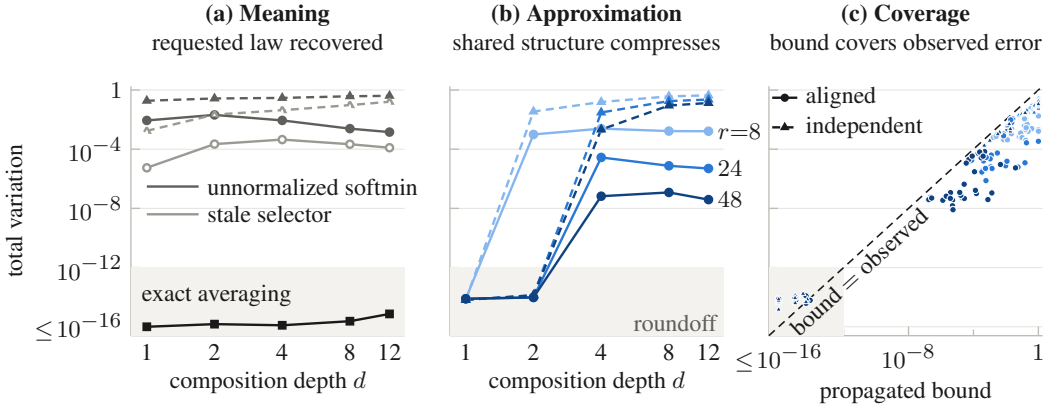

Figure 2: **Held-out closure separates errors of meaning from errors of approximation.** Three held-out seeds over all 4^6 length-six sequences; lines are medians over seeds and three queries (exact line: per-depth maximum). Solid, circles: aligned branches; dashed, triangles: independent branches. Gray: misspecified laws (dark: unnormalized softmin; light: stale selector); black: exact reference; blue: ranks $r = 8, 24, 48$, light to dark; shading: roundoff. (a) Normalizer-aware averaging recovers the requested law (TV $\leq 7.4 \times 10^{-16}$), while the misspecified laws drift. (b) Shared structure keeps finite-rank error small; independent branches do not. (c) All 270 observed errors lie at or below their propagated bounds up to roundoff; the bound is conservative. The experiment tests semantic correctness and represented-error coverage, not runtime advantage (exact min-fill is faster on every tested case of this enumerable control).

4 Held-out composition closure

This experiment asks whether the implementation constructs the requested normalized law after repeated composition, and whether shared structure changes finite-rank approximation while the propagated represented-law bound remains valid.

Data model and controls. The sample space is all $4^6 = 4096$ length-six sequences over four symbols, so every law and error is exactly enumerable; this is a synthetic known-answer setting, not an application dataset. Each mixture branch is a strictly positive nearest-neighbor chain law whose normalizer is recovered by exact contraction and whose square-root amplitude is an explicit chain MPS. A depth-d target multiplies d normalized binary mixture modules and a query factor F_q, $T_{d,q} \propto F_q \prod_{j=1}^{d} [w_j \pi_{j,0} + (1 - w_j)\pi_{j,1}]$, for $d \in \{1, 2, 4, 8, 12\}$; each branch is supplied at an arbitrary positive scale, so only contracted normalizers recover the requested weights. *Aligned* branches are small perturbations of one shared random backbone; *independent* branches are drawn independently and matched to their aligned counterparts in log-potential range and normalizer, so the families differ in shared structure rather than scale. Exact normalizer-aware averaging is the semantic reference and exact min-fill elimination the exact-inference baseline; an unnormalized softmin of branch energies (Proposition 9) and a stale selector are deliberately misspecified diagnostics. Rank-capped MPS (ranks 8, 24, 48) is compiled once per cell and reused across three

queries; three held-out seeds, separate from the development seeds that prespecified the analysis, supply 30 depth-by-family cells (Appendix E).

Results. *Composition recovers the requested law, while misspecified semantics drift.* Exact averaging agrees with the declared law to maximum TV 7.38×10^{-16} and is unchanged by positive rescaling of the supplied branch scales, whereas the misspecified laws drift, most strongly for independent branches (unnormalized-softmin TV 0.368–0.423 at depth 8 for the manufacturability query; Figure 2a). *Shared structure governs finite-rank error.* At the same ranks, aligned compositions remain accurate while independent compositions stay far from the composed law at every tested rank (Figure 2b); this is structure-dependent approximation on the tested families, not a universal compactness result. *Propagated bounds cover every observed error,* with excess at most 3.79×10^{-14}, at roundoff (Figure 2c), validating the propagated use of Proposition 4; the bound is conservative and certifies only the represented object. The experiment tests semantic correctness and bound coverage, not runtime advantage: exact min-fill is faster on every tested case.

5 RNA: ONE COMPILED PROPOSAL, REUSED AND COMPOSED

Target, proposal, and estimator. Eterna100 supplies target secondary structures S (Anderson-Lee et al., 2016). Sequences are encoded on a paired-variable space in which each base pair takes one of six canonical or wobble pairs and each unpaired site one of four nucleotides. The declared computational target is $\pi_S(x) \propto q_0(x) P_{\mathrm{V}}(S \mid x)^2$, where q_0 is uniform and $P_{\mathrm{V}}(S \mid x)$ is ViennaRNA's ensemble probability of S (Lorenz et al., 2011); it is a computational objective. The square is a deliberate design choice: the exact amplitude $\sqrt{\pi_S} \propto \sqrt{q_0}\, P_{\mathrm{V}}(S \mid x)$ makes the folding probability itself the amplitude score, and squaring favors sequences that assign high ensemble probability to S more strongly than a target proportional to $q_0 P_{\mathrm{V}}$, whose amplitude would be $\sqrt{q_0 P_{\mathrm{V}}}$. Turner nearest-neighbor loop energies of the fixed structure factor over local loops (Turner & Mathews, 2010), so the proposal amplitude $a_s(x) = \sqrt{q_0(x)} \prod_j \exp[-sE_j(x)/(2RT)]$ is an exact product of local factors. It is compiled once per puzzle to rank 8 and mixed with a fraction $\alpha = 0.05$ of q_0 (Corollary 6). The factors omit the sequence-dependent ensemble partition function inside P_{V}, so the MPS is a proposal: exact ratios $q_0 P_{\mathrm{V}}^2/q$ correct it, and its compression bound concerns only the factor product. Strength $s = 0.25$ and rank 8 were selected using Puzzle-8 calibration and development puzzles, then fixed for subsequent evaluations. For an event C and $g(x) = P_{\mathrm{V}}(S \mid x)$, the estimand is the conditional fold quality $\mu_C = \mathbb{E}_{\pi_S}[g \mid C]$, estimated as in Proposition 7 from a pool of 256 scored draws; a cell with no selected particle is unsupported, not imputed. Direct importance sampling from q_0 uses the same scoring, and annealed SMC along $q_0 P_{\mathrm{V}}^{\beta}$ with β from 0 to 2 (Del Moral et al., 2006) is the adaptive comparator at the same 256 requested target scores (Appendix F).

Calibration on an enumerable target. Factor extraction is exact: five local Turner tables, from 516 local energy evaluations and no full-target calls, reproduce ViennaRNA's fixed-structure energy on all 331,776 compatible sequences of puzzle 8. The rank-8 factor proposal has population ESS fraction 0.649 against the declared target, versus 0.330 for q_0. A cross-approximation proposal is more target-faithful but needs thousands of target calls to build, and a fixed-energy proposal at the target's exponent fails, so the omitted partition function is load-bearing (Appendix G).

Later conditions reuse the scored pool. Nine later conditions—GC content 45–55%, no homopolymer run of four, local diversity, their conjunctions, a leftmost exterior `CG` pair, and all four together—were evaluated on the immutable scored pools of nine completed puzzles from a prospectively selected twelve-puzzle cohort, with no additional ViennaRNA evaluations, no refitting, and no new particles. The factor and direct pools jointly supported 281 of 324 condition–replicate cells, and the equal-puzzle Factor-to-Direct conditional-ESS ratio was 1.141 with puzzle-cluster interval $[1.032, 1.284]$ (Figure 3a). Support and gain decline with conjunction depth, from 1.23 for single conditions to 0.95 for the four-factor condition, which both pools support in only 19 of 36 cells. This is the finite-pool behavior that Proposition 7 leaves open: the evidence supports selected later conditions, not arbitrary-query reuse.

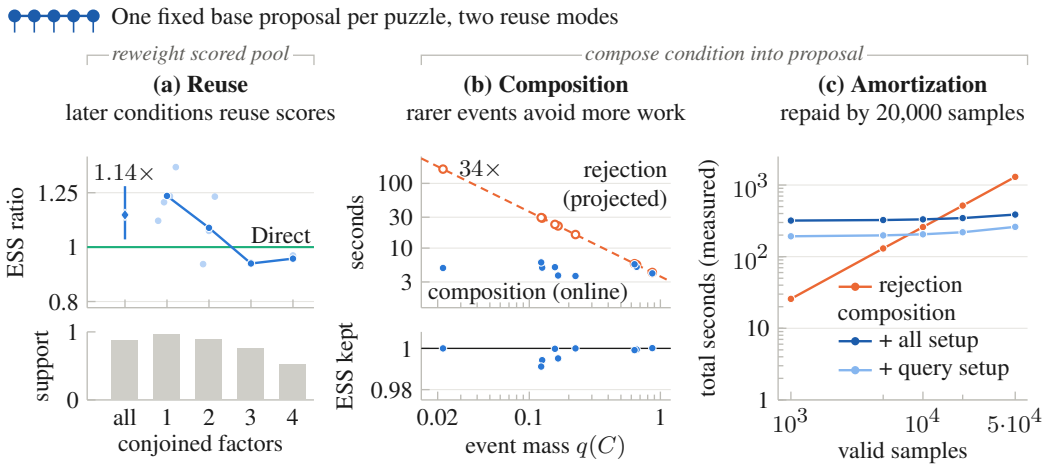

Figure 3: **One fixed base proposal per puzzle, two reuse modes.** (a) Later conditions reuse scores: nine later conditions reweight the immutable scored pools of nine puzzles without new ViennaRNA evaluations. Top: Factor-to-Direct conditional-ESS ratio (bold: per depth; faint: per condition; "all": interval $[1.03, 1.28]$; green line: parity with Direct). Bottom: fraction of cells supported by both pools. Gains hold for shallow conditions and fade for deeper conjunctions. (b, c) Composition multiplies an exact condition projector into the fixed base proposal and recompresses, creating a condition-specific tensor; the base is not refit. (b) Nine exact Puzzle-8 conditions: composition (measured online time, filled) keeps 99.1–100% of exact-rejection population ESS, while projected rejection time (hollow) grows as $1/q(C)$. (c) Five measured Puzzle-3 workloads over about 2.6×10^{17} states: with all setup charged, composition overtakes rejection by $20{,}000$ samples and is $3.36\times$ faster at $50{,}000$.

Composing exact conditions avoids rejection work. On puzzle 8, each of the nine conditions was composed into the compiled proposal as an exact projector and rounded. Composition keeps 0.991–1.000 of the population ESS of exact rejection conditioning, and every rounded conditioned proposal lies within TV 0.0355 of its exact conditioned law (Figure 3b). Rejection time, projected from the exact represented event mass and the measured base-draw time via Corollary 8, grows as $1/q(C)$; for the rarest condition ($q(C) = 0.022$) it is $34\times$ the measured online composition time, while for common conditions the two are comparable. On the non-enumerable puzzle 3, a condition conjoining the GC-content automaton of Proposition 2 with two local constraints acts on about 2.60×10^{17} compatible states and is composed exactly (Appendix B). Over five measured workloads, composition sampled about $18\times$ faster than rejection; including condition-specific setup, its total time fell below rejection by 10,000 samples, and including all setup by 20,000 samples, reaching $3.36\times$ at 50,000 (Figure 3c; Appendix G).

Conditional quality at a fixed target-score budget. At 256 requested ViennaRNA scores per replicate, the fixed factor recipe estimates the prespecified conditional endpoint more accurately than the tested annealed SMC. The endpoint is the fraction of cells whose estimate of μ_C lies within 0.05 of a reference, scored only where independent high-budget SMC and parallel-tempering references (Earl & Deem, 2005) agree. In a prospective evaluation on six prespecified puzzles, five completed; one resource noncompletion was not replaced. On the 152 reference-resolved cells, factor and direct were each accurate on 112 and SMC on 96; the equal-puzzle Factor$-$SMC difference was $+0.142$ with puzzle-cluster interval $[0.040, 0.244]$, favoring factor on four of five puzzles (Figure 4a); an earlier nine-puzzle benchmark ($+0.120$ $[0.063, 0.179]$) and a matched-256 robustness check with a stronger SMC schedule on the same cohort ($+0.139$ $[0.040, 0.229]$; not a fresh replication) agree. With identical requested scores, this also reads as better use of the tested target-evaluation budget—one result read two ways, not a total-compute or wall-time advantage.

Two counterpoints qualify the comparison. Direct importance ties factor on the binary endpoint (Figure 4b), so for single-target estimation the factor proposal's distinct value is reuse and composition rather than accuracy over direct sampling. The endpoint difference partly reflects finite-pool

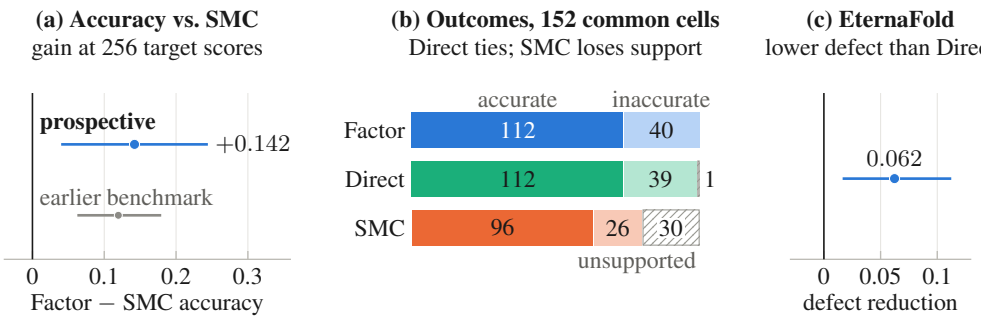

Figure 4: **At 256 requested ViennaRNA scores, the factor proposal estimates the conditional endpoint more accurately than tested SMC.** (a) Equal-puzzle Factor−SMC difference in the fraction of conditional fold-quality estimates within 0.05 of the reference, with puzzle-cluster intervals: prospective evaluation (five of six prespecified puzzles completed) and the earlier nine-puzzle benchmark. (b) Outcomes on the 152 cells with an agreed reference: Direct ties Factor, while SMC leaves 30 cells unsupported. (c) On outcome-independently selected particles, the corrected pool lowers EternaFold ensemble defect relative to Direct (interval $[0.017, 0.112]$): cross-model computational transfer.

support: SMC produced no usable conditional estimate in 30 of 152 reference-resolved condition–replicate cells, which the prespecified endpoint counts as failures, whereas Factor supported all 152. Accuracy conditional on method-specific support is not directly comparable because the supported cell sets differ and missing estimates concentrate on rarer conditions. In the earlier matched-score cohort, SMC also had $13.38\times$ higher terminal weight ESS than the factor pool while supporting fewer condition cells (218 versus 288 of 324): terminal weight balance and downstream conditional accuracy are different quantities, and neither observation is a general statement about SMC.

Favorable pool quality transfers to an independently parameterized folding model. On an outcome-independent, method-balanced selection of 1,728 existing sequences, frozen before any EternaFold scoring (Wayment-Steele et al., 2022), the corrected pool reduced Vienna-weighted EternaFold ensemble defect relative to Direct by 0.0623 with puzzle-cluster interval $[0.0165, 0.1123]$ over eight posterior-complete puzzles (Figure 4c). Both corrected pools estimate the same Vienna-defined law asymptotically, so this compares realized finite pools; EternaFold was trained partly on Eterna-derived data, and this is cross-model computational transfer. With 16 selected rows per method and replicate, finite-subset variance remains material.

6 Conclusion

The central lesson is that reusable probabilistic inference requires a contract, not merely an algebra. One must first declare the law intended by each composition, then distinguish approximation of its representation from correction to a separately specified target, and finally measure the support, overlap, finite-sample error, and cost that neither algebra nor compression bounds determine. The amplitude representation makes this contract especially natural: probability factors enter once through their square roots, ordinary low-rank compression still yields a nonnegative law after squaring, and Euclidean residuals translate into dimension-free probability control.

RNA design demonstrates the resulting workflow rather than defining its scope. A puzzle-specific proposal can be constructed once, corrected by exact target ratios, reused to answer later conditional questions without rescoring, and composed with new exact conditions when direct conditioned sampling is preferable. The same pattern applies whenever an expensive structured representation precedes a changing collection of evaluable factors or events. Our results do not establish universal tensor tractability or superiority to specialized inference methods. They instead establish a backend-independent foundation for deciding what law has been constructed, what approximation is certified, what exact correction recovers, and which benefits of reuse must still be demonstrated empirically.

REPRODUCIBILITY STATEMENT

Section 3 states every assumption used by the formal results, and Appendices A–D give complete proofs. Section 4 and Appendix E specify the controlled data generator, queries, ranks, and seed-selection protocol. Section 5 and Appendices F–G specify the RNA target, encoding, proposal construction, configuration selection, baselines, reference-agreement rule, budgets, prospective cohort, and complete result tables, including unfavorable and unsupported outcomes. Multi-puzzle intervals resample puzzles rather than particles.

AI USE STATEMENT

AUTHOR_AI_USE_STATEMENT_REQUIRED

This mandatory ICLR 2027 section must be written by the authors before submission.
Submission is blocked while this placeholder remains.

REFERENCES

Jeff Anderson-Lee, Eli Fisker, Vineet Kosaraju, Michelle Wu, Justin Kong, Jeehyung Lee, Minjae Lee, Mathew Zada, Adrien Treuille, and Rhiju Das. Principles for predicting RNA secondary structure design difficulty. *Journal of Molecular Biology*, 428(5):748–757, 2016. doi: 10.1016/j.jmb.2015.11.013.

Johannes Christmann, Petr Ivashkov, Mattia Chiurco, and Guglielmo Mazzola. From quantum-enhanced to quantum-inspired Monte Carlo. *Physical Review A*, 111:042615, 2025. doi: 10.1103/PhysRevA.111.042615.

Tiangang Cui, Sergey Dolgov, and Robert Scheichl. Deep importance sampling using tensor trains with application to a priori and a posteriori rare events. *SIAM Journal on Scientific Computing*, 46(1):C1–C29, 2024. doi: 10.1137/23M1546981.

Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):411–436, 2006. doi: 10.1111/j.1467-9868.2006.00553.x.

Yilun Du. Generalization by construction. Research perspective, `https://yilundu.github.io/perspective/generalization_by_construction.html`, 2026. Accessed 2026-09-26.

Yilun Du, Conor Durkan, Robin Strudel, Joshua B. Tenenbaum, Sander Dieleman, Rob Fergus, Jascha Sohl-Dickstein, Arnaud Doucet, and Will Sussman Grathwohl. Reduce, reuse, recycle: Compositional generation with energy-based diffusion models and MCMC. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 8489–8510, 2023.

Jeffrey M. Dudek, Leonardo Dueñas-Osorio, and Moshe Y. Vardi. Efficient contraction of large tensor networks for weighted model counting through graph decompositions, 2020. arXiv:1908.04381v2.

David J. Earl and Michael W. Deem. Parallel tempering: Theory, applications, and new perspectives. *Physical Chemistry Chemical Physics*, 7(23):3910, 2005. doi: 10.1039/B509983H.

Andrew J. Ferris and Guifré Vidal. Perfect sampling with unitary tensor networks. *Physical Review B*, 85:165146, 2012. doi: 10.1103/PhysRevB.85.165146.

Tal Friedman and Guy Van den Broeck. Approximate knowledge compilation by online collapsed importance sampling, 2018. arXiv:1805.12565.

Christopher A. Fuchs and Jeroen van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, 1999. doi: 10.1109/18.761271.

Zhao-Yu Han, Jun Wang, Heng Fan, Lei Wang, and Pan Zhang. Unsupervised generative modeling using matrix product states. *Physical Review X*, 8:031012, 2018. doi: 10.1103/PhysRevX.8.031012.

W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. doi: 10.1093/biomet/57.1.97.

Yong Lai. Approximate model counting by partial knowledge compilation, 2018. arXiv:1805.07180.

David Layden, Guglielmo Mazzola, Ryan V. Mishmash, Mario Motta, Pawel Wocjan, Jin-Sung Kim, and Sarah Sheldon. Quantum-enhanced Markov chain Monte Carlo. *Nature*, 619:282–287, 2023. doi: 10.1038/s41586-023-06095-4.

Jun S. Liu. Metropolized independent sampling with comparisons to rejection sampling and importance sampling. *Statistics and Computing*, 6(2):113–119, 1996. doi: 10.1007/BF00162521.

Nan Liu, Shuang Li, Yilun Du, Antonio Torralba, and Joshua B. Tenenbaum. Compositional visual generation with composable diffusion models. In *Computer Vision – ECCV 2022*, Lecture Notes in Computer Science, pp. 423–439, 2022. doi: 10.1007/978-3-031-19790-1_26.

Lorenzo Loconte, Aleksanteri M. Sladek, Stefan Mengel, Martin Trapp, Arno Solin, Nicolas Gillis, and Antonio Vergari. Subtractive mixture models via squaring: Representation and learning, 2024. arXiv:2310.00724v3.

Ronny Lorenz, Stephan H. Bernhart, Christian Höner zu Siederdissen, Hakim Tafer, Christoph Flamm, Peter F. Stadler, and Ivo L. Hofacker. ViennaRNA Package 2.0. *Algorithms for Molecular Biology*, 6:26, 2011. doi: 10.1186/1748-7188-6-26.

João Loula, Benjamin LeBrun, Li Du, Ben Lipkin, Clemente Pasti, Gabriel Grand, Tianyu Liu, Yahya Emara, Marjorie Freedman, Jason Eisner, Ryan Cotterell, Vikash Mansinghka, Alexander K. Lew, Tim Vieira, and Timothy J. O'Donnell. Syntactic and semantic control of large language models via sequential Monte Carlo. In *International Conference on Learning Representations*, 2025. arXiv:2504.13139.

K. L. Mengersen and R. L. Tweedie. Rates of convergence of the Hastings and Metropolis algorithms. *The Annals of Statistics*, 24(1), 1996. doi: 10.1214/aos/1033066201.

Georgii S. Novikov, Maxim E. Panov, and Ivan V. Oseledets. Tensor-train density estimation. In *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, volume 161 of *Proceedings of Machine Learning Research*, pp. 1321–1331, 2021.

Ivan V. Oseledets. Tensor-train decomposition. *SIAM Journal on Scientific Computing*, 33(5):2295–2317, 2011. doi: 10.1137/090752286.

Alessandro Rudi and Carlo Ciliberto. PSD representations for effective probability models. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Douglas H. Turner and David H. Mathews. NNDB: The nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucleic Acids Research*, 38(suppl. 1): D280–D282, 2010. doi: 10.1093/nar/gkp892.

Antonio Vergari, YooJung Choi, Anji Liu, Stefano Teso, and Guy Van den Broeck. A compositional atlas of tractable circuit operations for probabilistic inference. In *Advances in Neural Information Processing Systems*, volume 34, pp. 13189–13201, 2021.

Hannah K. Wayment-Steele, Wipapat Kladwang, Alexandra I. Strom, Jeehyung Lee, Adrien Treuille, Alex Becka, Eterna Participants, and Rhiju Das. RNA secondary structure packages evaluated and improved by high-throughput experiments. *Nature Methods*, 19(10):1234–1242, 2022. doi: 10.1038/s41592-022-01605-0.

Honghua Zhang, Meihua Dang, Nanyun Peng, and Guy Van den Broeck. Tractable control for autoregressive language generation. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 40932–40945, 2023.

Stephen Zhao, Rob Brekelmans, Alireza Makhzani, and Roger Baker Grosse. Probabilistic inference in language models via twisted sequential Monte Carlo. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 60704–60748, 2024.

A TYPED SEMANTICS: PROOFS

Proof of Proposition 1. Throughout, $\pi_i(x) = |\psi_i(x)|^2/\|\psi_i\|_2^2$.

Item 1. $|\psi_A(x)\psi_B(x)|^2 = \|\psi_A\|_2^2\|\psi_B\|_2^2\,\pi_A(x)\pi_B(x)$. The norm factors do not depend on x and cancel on division by the positive total mass.

Item 2. Squaring $\sqrt{p}\sqrt{F_A}\sqrt{F_B}$ gives pF_AF_B pointwise. The separately normalized laws are pF_A/Z_A and pF_B/Z_B; their product is $p^2F_AF_B/(Z_AZ_B)$, proportional to $p^2F_AF_B$. Normalization cannot remove the extra state-dependent factor p, so the two constructions agree only in special cases such as a base that is uniform on the retained support.

Item 3. Because label states are orthogonal, $\sum_{i,x}|\Psi(i,x)|^2 = \sum_i w_i \sum_x |\psi_i(x)|^2/\|\psi_i\|_2^2 = 1$ and no cross terms between labels arise. Marginalizing the label gives $\sum_i |\Psi(i,x)|^2 = \sum_i w_i\rho[\psi_i](x)$. After multiplication by $\sqrt{H}$, label i carries unnormalized mass $\sum_x |\Psi(i,x)|^2 H(x) = w_i Z_{i,H}$; normalizing over i, which requires $\sum_j w_j Z_{j,H} > 0$, gives the stated posterior. A construction that renormalizes each conditioned branch but keeps the prior weights w_i instead encodes $\sum_i w_i\rho[\psi_i\sqrt{H}]$, which differs whenever the $Z_{i,H}$ differ (the stale-selector law of Section 4).

Item 4. $|\mathbf{1}_C(x)\psi(x)|^2 = \mathbf{1}_C(x)|\psi(x)|^2$; dividing by the positive retained mass gives the conditional law. For integer $m \geq 1$, $|\psi^m|^2 = |\psi|^{2m}$, so $\rho[\psi^m] \propto \rho[\psi]^m$.

Sign invariance. Each displayed law is a function of the magnitudes only, so changing the sign of $\psi_i(x)$ at any state leaves every magnitude, every norm, and hence every law unchanged. □

Proposition 9 (Normalizer-induced softmin weights)**.** *Let $w_A, w_B \geq 0$ with $w_A + w_B = 1$, let $e^{-\beta E_i} = Z_i\pi_i$ with $Z_A, Z_B > 0$, and hence $w_AZ_A + w_BZ_B > 0$. The unnormalized sum $w_Ae^{-\beta E_A} + w_Be^{-\beta E_B}$ induces $\pi_\vee = \omega_A\pi_A + (1-\omega_A)\pi_B$ with $\omega_A = w_AZ_A/(w_AZ_A + w_BZ_B)$, and $\|\pi_\vee - (w_A\pi_A + w_B\pi_B)\|_{\mathrm{TV}} = |\omega_A - w_A|\,\|\pi_A - \pi_B\|_{\mathrm{TV}}$.*

Proof. Substituting $e^{-\beta E_i} = Z_i\pi_i$ and normalizing by the total mass $w_AZ_A + w_BZ_B$ gives $\pi_\vee$. Its difference from the requested mixture is $(\omega_A - w_A)(\pi_A - \pi_B)$; taking half the ℓ_1 norm gives the identity. □

The proposition diagnoses only the displayed operation and is the source of the unnormalized-softmin diagnostic in Section 4; it is not a claim about any published composition method.

Proof of Proposition 2. Let $A_k^{x_k}$ and $B_k^{x_k}$ be cores in a common order. The Kronecker cores $A_k^{x_k} \otimes B_k^{x_k}$ contract to $\psi_A(x)\psi_B(x)$ because the mixed-product property factorizes the chain product, and their bonds are products of the input bonds. For a labeled mixture, place normalized component cores in block-diagonal direct sums, put the coefficients $\sqrt{w_i}/\|\psi_i\|_2$ at one boundary, and terminate block i in its own label state; bonds add. Applying the product construction to an exact indicator MPS of bond χ_C gives the projector bound, and $m-1$ repeated products give the power bound. The automaton statement is Lemma 10. All are exact construction upper bounds; no tightness follows. □

B WEIGHTED AUTOMATA AND THE EXACT PUZZLE-3 CONDITION

Lemma 10 (Weighted automaton to MPS)**.** *Let $u, v \in \mathbb{R}^R$ and let position-dependent transfer matrices $T_{x_k}^{(k)} \in \mathbb{R}^{R\times R}$ satisfy $\psi(x) = u^\top T_{x_1}^{(1)} \cdots T_{x_n}^{(n)} v$ for all x. Then ψ has an exact MPS with every bond at most R.*

Proof. Set $A_1^{x_1} = u^\top T_{x_1}^{(1)}$, $A_k^{x_k} = T_{x_k}^{(k)}$ for $1 < k < n$, and $A_n^{x_n} = T_{x_n}^{(n)} v$. The MPS contraction is exactly the automaton path sum, and every virtual index ranges over at most R states. Position-dependent state sets embed into a common set of the largest size. □

Proposition 11 (Exact global-count conditions)**.** *Let each site k emit a nucleotide string $s_k(x_k)$, let $c_k(x)$ be the number of G or C symbols emitted by sites $1, \ldots, k$, and let $C = \{x : \ell \leq c_n(x) \leq u\}$. The deterministic automaton whose state after site k is c_k, restricted to counts from which $[\ell, u]$*

remains reachable and at most u, has transfer entries in $\{0,1\}$ and accepts exactly C. Hence $\mathbf{1}_C$ has an exact MPS whose bond at each cut is at most the number of such reachable counts. Conjoining conditions that depend on a bounded window of recently emitted symbols multiplies the state by the window contents that remain active across the cut, and conjoining single-site restrictions leaves the bond unchanged.

Proof. The automaton is deterministic: from state c_{k-1} and symbol x_k it moves to $c_{k-1} + |s_k(x_k)|_{\mathrm{GC}}$, or to a rejecting sink (omitted, contributing zero) when the count exceeds u or can no longer reach ℓ with the remaining sites. Each accepted path has weight one and each rejected path weight zero, so the path sum equals $\mathbf{1}_C$, and Lemma 10 gives the MPS. A window condition is checked when its last symbol is read; storing the active window symbols in the state makes each check a deterministic transition, and a single-site restriction only zeroes transfer entries. □

The Puzzle-3 condition of Section 5 conjoins the GC-content window $[0.45L, 0.55L]$ on the full sequence of length L with two single-site restrictions: the paired variable with the smallest left position is a `CG` or `GC` pair, and the unpaired variable with the smallest position is `C` or `G`. Its exact projector realized maximum bond 20. Multiplying it into the reusable rank-8 Turner-factor MPS (Proposition 2) and rounding with a cap of 256 realized rank 137 and represented-object bound 2.75×10^{-13} (Theorem 5); every one of 1,024 checked composed draws satisfied the condition.

C STABILITY PROOFS

Proof of Theorem 3. Write $ab - \widetilde{a}\widetilde{b} = (a - \widetilde{a})b + \widetilde{a}(b - \widetilde{b})$. Since $\|uv\|_2 \le \|u\|_2\|v\|_\infty$, the triangle inequality gives $\|ab - \widetilde{a}\widetilde{b}\|_2 \le \|b\|_\infty E_a + \|\widetilde{a}\|_\infty E_b$; inserting $a\widetilde{b}$ instead gives the other branch. For an exact factor, the diagonal multiplication operator has norm $\|D_f\|_{2\to 2} = \|f\|_\infty$. At sequential step k,

$$\|\psi_k - \widetilde{\psi}_k\|_2 \le \|D_{f_k}(\psi_{k-1} - \widetilde{\psi}_{k-1})\|_2 + \|D_{f_k}\widetilde{\psi}_{k-1} - \widetilde{\psi}_k\|_2 \le L_k E_{k-1} + e_k,$$

and induction gives Eq. (1): E_0 crosses every later factor, while e_k crosses only factors $k+1, \dots, m$. The statement is in exact arithmetic with exact factors; approximate factor values require separate multiplication-error terms. □

Proof of Proposition 4. Let $u = a/\|a\|_2$ and $v = \widetilde{a}/\|\widetilde{a}\|_2$. For pure states, the trace distance is $D(u,v) = \sqrt{1 - |\langle u, v\rangle|^2}$ (unsquared overlap). Measuring in the fixed computational basis produces the classical laws $\rho[a]$ and $\rho[\widetilde{a}]$, and trace distance equals the maximum over measurements of the total variation between outcome laws (Fuchs & van de Graaf, 1999); hence $\|\rho[a] - \rho[\widetilde{a}]\|_{\mathrm{TV}} \le D(u,v)$. Next, $D(u,v)\|\widetilde{a}\|_2$ is the Euclidean distance from $\widetilde{a}$ to the one-dimensional linear subspace spanned by u, which is at most $\|\widetilde{a} - a\|_2$ because a lies in that subspace. Dividing by $\|\widetilde{a}\|_2$ and capping at one, the maximum possible TV, proves Eq. (2). Both amplitudes must be nonzero; the bound becomes uninformative when $\|\widetilde{a}\|_2$ is small, and it compares laws on the same encoded support, so omitted states are not detected. □

Corollary 12 (Bounded observables)**.** *If $\|p - \widetilde{p}\|_{\mathrm{TV}} \le \delta$ and $g : \Omega \to [0,1]$, then $|\mathbb{E}_p g - \mathbb{E}_{\widetilde{p}} g| \le \delta$.*

Proof. $\mathbb{E}_p g - \mathbb{E}_{\widetilde{p}} g = \sum_x (p(x) - \widetilde{p}(x))(g(x) - \frac{1}{2})$, and $|g - \frac{1}{2}| \le \frac{1}{2}$, so the magnitude is at most $\frac{1}{2}\|p - \widetilde{p}\|_1 = \delta$. □

The corollary transfers a represented-law bound to bounded observables of that same represented law. It does not transfer the bound to a different declared target: in Section 5, the factor-product bound does not bound the ViennaRNA estimand.

Proof of Theorem 5. At bond b, the exact SVD of the current residual splits it into a retained block K_b and a discarded block D_b with $\|D_b\|_F^2 = r_b^2$. The left cores produced so far are isometries, so embedding D_b in the full tensor preserves its norm. Every later SVD acts inside the retained subspace, so every later discarded block and the final tensor ϕ lie in that subspace and are orthogonal to D_b. Induction makes the embedded blocks pairwise orthogonal and orthogonal to ϕ, so $\psi =$

$\phi + \sum_b D_b$, $\|\psi - \phi\|_2^2 = \sum_b r_b^2 = \delta^2$, and $\langle \psi - \phi, \phi \rangle = 0$. Because $\|\psi\|_2 = 1$, $\|\phi\|_2^2 = 1 - \delta^2$ and, with $\eta = \phi/\|\phi\|_2$, $|\langle \psi, \eta \rangle| = \sqrt{1 - \delta^2}$. The pure-state trace distance is therefore δ, and fixed-basis measurement cannot increase it. For an unnormalized input T, applying the result to $T/\|T\|_2$ gives $\|\rho[T] - \rho[\phi]\|_{\mathrm{TV}} \le \delta/\|T\|_2 = \delta/\sqrt{\|\phi\|_2^2 + \delta^2}$. □

The argument requires one exact left-to-right sweep whose tails are measured on successive current residuals, with the canonical form that sweep assumes. It does not cover interleaved factor multiplication, repeated recompression, fitting error, or floating-point arithmetic.

D CORRECTION AND REUSE PROOFS

Self-normalized importance sampling and independence Metropolis–Hastings. Let $\pi = \widetilde{\pi}/Z$, $\pi \ll q$, $w = \widetilde{\pi}/q$ on $\{q > 0\}$, and $x_1, x_2, \ldots$ i.i.d. from q. Then $\mathbb{E}_q[wg] = Z\,\mathbb{E}_\pi g$ and $\mathbb{E}_q w = Z$, so if $\mathbb{E}_\pi|g| < \infty$ the strong law of large numbers applied to numerator and denominator gives $\sum_i w_i g(x_i)/\sum_i w_i \to \mathbb{E}_\pi g$ almost surely. The estimator is consistent but not finite-sample unbiased. For independence Metropolis–Hastings restricted to $\Omega_\pi = \{\pi > 0\}$, rejecting proposals outside it, the accepted flow between $x \neq y$ is $\min\{\pi(x)q(y), \pi(y)q(x)\}$, which is symmetric; $q > 0$ on Ω_π gives irreducibility and a positive self-loop, so the finite chain converges to π from every start in Ω_π (Hastings, 1970; Liu, 1996). Finite-run accuracy is not implied.

Proof of Corollary 6. If $\pi(x) > 0$ then $q_0(x) > 0$ and $q_\alpha(x) \ge \alpha q_0(x) > 0$, so $\pi \ll q_\alpha$. Also $\pi - q_\alpha = (1 - \alpha)(\pi - q) + \alpha(\pi - q_0)$, so the triangle inequality and $\|\pi - q_0\|_{\mathrm{TV}} \le 1$ give the bound. If q and q_0 are evaluable, so is q_α. No lower bound on relative tails follows. □

Population ESS. For normalized $\pi \ll q$ with $w = \pi/q$, $\mathbb{E}_q w = 1$ and $\mathbb{E}_q w^2 - 1 = \sum_x (\pi(x) - q(x))^2/q(x) = \chi^2(\pi\|q)$, so $(\mathbb{E}_q w)^2/\mathbb{E}_q[w^2] = 1/(1 + \chi^2(\pi\|q))$. This is direct algebra.

Proposition 13 (Small TV does not control overlap)**.** *For every $\delta \in (0, 1/2)$, the full-support laws $\pi = (1-\delta, \delta)$ and $q = (1-\delta^3, \delta^3)$ satisfy $\|\pi - q\|_{\mathrm{TV}} \le \delta$, $\chi^2(\pi\|q) \ge 9/(16\delta)$, $\max_x \pi(x)/q(x) = \delta^{-2}$, and population ESS fraction at most $16\delta/9$.*

Proof. $\|\pi - q\|_{\mathrm{TV}} = \delta - \delta^3 \le \delta$. Also $\chi^2(\pi\|q) \ge (\delta - \delta^3)^2/\delta^3 = (1 - \delta^2)^2/\delta \ge (3/4)^2/\delta$, the ratio at the second state is δ/δ^3, and the ESS fraction is $1/(1 + \chi^2) \le 1/\chi^2$. Both laws have full support on the two states. □

Thus a represented-law bound, however small, does not bound importance weights or ESS against a different law; the RNA experiments report these quantities separately.

Proof of Proposition 7. On C, $\pi_C(x)/q_C(x) = [\pi(x)/\pi(C)]/[q(x)/q(C)]$, which is the displayed identity; $q(C) > 0$ because $\pi(C) > 0$ and $\pi \ll q$. With $w = \widetilde{\pi}/q$ and i.i.d. draws from q, $\mathbb{E}_q[w\mathbf{1}_C g] = Z\,\mathbb{E}_\pi[\mathbf{1}_C g]$ and $\mathbb{E}_q[w\mathbf{1}_C] = Z\,\pi(C) > 0$. The strong law applied to both averages gives almost-sure convergence of the ratio to $\mathbb{E}_\pi[\mathbf{1}_C g]/\pi(C) = \mathbb{E}_\pi[g \mid C]$ for each k; a finite intersection of probability-one events has probability one, so the K estimators converge jointly. They use only the stored x_i, w_i, $\mathbf{1}_{C_k}(x_i)$, and $g_k(x_i)$, hence no new target evaluation. Equivalently, by the identity each estimator is self-normalized importance sampling for π_{C_k} with proposal q_{C_k}, whose unknown constant cancels. □

The standard asymptotic variance of each estimator, $\mathbb{E}_q[w^2\mathbf{1}_C(g - \mu_C)^2]/\pi(C)^2$ for normalized w, grows as $\pi(C)$ shrinks or weights concentrate, which is why the experiments report support and conditional ESS for rare and deep conditions rather than only aggregate accuracy.

Proof of Corollary 8. By Proposition 1(4), $\rho[\mathbf{1}_C\sqrt{q}] = q\mathbf{1}_C/q(C) = q_C$. Rejection sampling accepts a draw from q with probability $q(C)$ and returns accepted draws distributed as q_C; the number of base draws per acceptance is geometric with mean $1/q(C)$, so N accepted samples need $N/q(C)$ base draws in expectation. Composition is cheaper exactly when $S_C + Nc_C < Nc_q/q(C)$, i.e. $N(c_q/q(C) - c_C) > S_C$, which is impossible if $c_q/q(C) \le c_C$ and otherwise holds exactly for $N > S_C/(c_q/q(C) - c_C)$. □

E CONTROLLED-CLOSURE DETAILS

Generator. Symbols are $\{\mathtt{A}, \mathtt{C}, \mathtt{G}, \mathtt{U}\}$ and sequences have length six. A shared backbone draws first-site and pair log-potentials independently from $\mathcal{N}(0, 0.35^2)$; four perturbation directions are drawn and normalized. An aligned branch adds $0.08 \sum_{r=1}^{4} \gamma_r v_r$ with $\gamma \sim \mathcal{N}(0, I_4)$ to the backbone log-potentials. An independent branch draws its log-potentials from $\mathcal{N}(0, 1)$, recenters them and rescales them to the peak-to-peak range of the corresponding aligned branch, and then rescales its first-site potential so that its normalizer equals the aligned branch's. Mixture weights are uniform on $[0.25, 0.75]$, and each branch amplitude is supplied with a log-uniform positive scale in $[0.2, 5]$; two further independent rescalings per cell test invariance. The manufacturability query requires two to four G/C symbols, no run of four identical symbols, and at least three distinct symbols in the length-six window; the left-anchor query requires $x_1 x_2 = \mathtt{CG}$.

Protocol. Candidate held-out seeds 2026091501–2026091599 were ordered by SHA-256 of a fixed namespace string and the first three were taken before any outcome was generated; they do not overlap the six development seeds used to prespecify the held-out analysis. All computations used float64 on one CPU thread with per-task limits of 120 s and 2 GiB. Exact min-fill completed every cell. An exact untruncated MPS completed 12 of 30 cells within the memory limit (maximum TV 4.18×10^{-14} where completed), which is why rank-capped construction carries the represented-object account. Table 2 reports median observed error and median propagated bound. Base compilation took 0.0138–0.0793 s, 0.00125–0.679 s, and 0.00130–4.612 s at ranks 8, 24, and 48, and three query updates took 0.0175–0.0333 s, 0.0669–0.396 s, and 0.133–3.769 s; exact min-fill took 0.00019–0.00506 s per query.

Table 2: Held-out controlled closure: median observed represented-law TV and median propagated bound over three held-out seeds and three queries, by family, rank, and depth. Entries below 10^{-12} are floating-point roundoff.

Family	Rank	Median	$d = 1$	$d = 2$	$d = 4$	$d = 8$	$d = 12$
Aligned	8	observed	7.8×10^{-15}	9.9×10^{-4}	0.00245	0.00167	0.00164
		bound	1.0×10^{-17}	0.0123	0.158	0.135	0.205
	24	observed	7.8×10^{-15}	9.6×10^{-15}	2.8×10^{-5}	7.5×10^{-6}	5.0×10^{-6}
		bound	1.0×10^{-17}	4.9×10^{-15}	6.1×10^{-4}	0.00107	0.00116
	48	observed	7.8×10^{-15}	9.6×10^{-15}	6.7×10^{-8}	1.2×10^{-7}	3.9×10^{-8}
		bound	1.0×10^{-17}	4.9×10^{-15}	2.1×10^{-6}	2.7×10^{-5}	2.2×10^{-5}
Independent	8	observed	6.6×10^{-15}	0.0351	0.154	0.372	0.436
		bound	2.3×10^{-17}	0.426	1	1	1
	24	observed	6.6×10^{-15}	1.3×10^{-14}	0.0303	0.178	0.227
		bound	2.3×10^{-17}	4.8×10^{-15}	0.521	1	1
	48	observed	6.6×10^{-15}	1.3×10^{-14}	0.00214	0.0891	0.133
		bound	2.3×10^{-17}	4.8×10^{-15}	0.0188	1	1

F RNA PROTOCOLS

Target and proposal. Paired variables take the six pairs AU, UA, GC, CG, GU, UG; unpaired variables take four nucleotides; q_0 is uniform on this space. ViennaRNA 2.7.2 supplies $P_{\mathrm{V}}(S \mid x)$ and the Turner-2004 loop energies. Factor extraction evaluates each local loop energy of the fixed structure on its variable scope and stores it as an exact table; the tempered amplitude is multiplied in a fixed order and rounded to rank 8, and the proposal is $q = 0.95\,\rho[\widetilde{a}_s] + 0.05\,q_0$. For composed proposals, mixture weights between the factor and uniform components are recomputed by contraction after conditioning (Proposition 1).

Configuration selection and evaluation sequence. Puzzle-8 calibration and development puzzles 1, 23, and 26 selected strength 0.25 with rank 8 for corrected estimation and reuse, and strength 0.5 for a separate yield objective. These settings were then fixed for: a two-puzzle confirmation (puzzles 30 and 11); a prospectively selected twelve-puzzle cohort; nine-condition reuse on the nine completed cohort puzzles; the EternaFold evaluation; matched-score SMC and an earlier conditional-accuracy benchmark on the same nine puzzles; and the prospective six-puzzle evaluation. Nine

of twelve prospectively selected cohort puzzles completed; three 284–380-nt targets exceeded the fixed per-step limit and were not replaced. On Puzzle 8, condition-specific rank caps were chosen from $\{8, 16, 32, 64\}$ by a rule fixed before any outcome that depends only on the represented-object bound, not on timing or utility; they are reported in Table 4.

Budgets and baselines. Each method has four replicates per puzzle with 256 requested ViennaRNA target scores each. Direct importance sampling draws from q_0. Annealed SMC follows $q_0 P_{\mathrm{V}}^{\beta}$ with $\beta \in \{0, 0.5, 1, 1.5, 2\}$, 64 particles, systematic resampling when ESS falls below half the particle count, one symmetric single-decision Metropolis move per particle at each intermediate exponent, and no final resampling; its terminal weighted pool is reused for all conditions. The matched-256 robustness schedule used 256 particles, the same resampling trigger, no intermediate Metropolis sweeps, and no final resampling.

Reference and endpoint. Four high-budget SMC and four parallel-tempering replicates per puzzle define the reference. A puzzle–condition cell is scored only if the two reference methods agree and have sufficient conditional support; otherwise it is unresolved for every method. A candidate estimate is accurate if it lies within 0.05 of the reference; a candidate cell with no selected particle is unsupported and is not accurate.

EternaFold evaluation. Rows were selected by a deterministic index rule that ignored sequence, ViennaRNA score, target-MFE status, proposal density, and weight, and the selection was frozen before any EternaFold call: nine puzzles, three methods, four replicates, and 16 rows per group (1,728 unique sequences). With posterior pair probabilities $P_{ab}^{\mathrm{EF}}(x)$, the defect is $D(x, S) = 1 - L^{-1} \sum_{a,b} P_{ab}^{\mathrm{EF}}(x) S_{ab}$, including unpaired probabilities on the diagonal. For corrected pools, means are weighted by self-normalized ViennaRNA importance weights. One invalid posterior row was excluded from posterior metrics, leaving eight posterior-complete puzzles for the corrected comparison.

G Additional RNA results

Proposal selection and overlap. On development puzzles, strength 0.25 improved corrected importance ESS over direct sampling by a geometric mean of $1.232\times$ (10 of 12 replicate pairs), with a loss on puzzle 26. On puzzles 30 and 11 it gave $2.087\times$ (7 of 8), although puzzle 11 remained weight-degenerate in absolute terms. Across the nine completed cohort puzzles the gain was $1.275\times$ with puzzle-bootstrap interval $[0.982, 1.841]$, favoring factor on seven puzzles (24 of 36 replicate pairs); puzzle 45 gained $4.81\times$, and puzzles 41 and 62 lost ($0.93\times$, $0.74\times$). Corrected ESS is therefore heterogeneous, which is why the main utility claims rest on reuse, composition, and conditional accuracy rather than on aggregate overlap.

Puzzle-8 calibration. The rank-8 factor proposal's declared-target TV is 0.299 while its factor-product compression TV is 0.0057 (bound 0.022): compression error and distance to the declared target are different layers. The selected cross-approximation proposal reached target TV 0.157 and population ESS fraction 0.803 but consumed 4,302 target calls to build; the fixed-energy proposal at the target's exponent had ESS fraction 0.022.

Reuse and composition. Table 3 gives every reused condition, Table 4 every Puzzle-8 composition, and Table 5 every measured Puzzle-3 workload. The Puzzle-8 bound and the exact TV are reported together: the bound covers the rounded represented object, is conservative, and does not certify corrected estimation under the ViennaRNA target. Puzzle-3 timings were measured on one GPU host class with peak resident memory 17.8 GiB; rejection acceptance was 0.045–0.046. Condition-specific composition setup took 191 s; all setup, adding factor extraction and base compilation, took 318 s.

Conditional accuracy. Table 6 gives the prospective result by puzzle; puzzle 40 favors SMC, and puzzle 33 contributes only four resolved cells. In the earlier nine-puzzle benchmark, factor, direct, and SMC were accurate on 166, 151, and 136 of 264 resolved cells; factor favored 8 of 9 puzzles against SMC, and factor minus direct was $+0.042$ with interval $[0.000, 0.089]$. Factor did

Table 3: Reuse of immutable scored pools for nine later conditions (no additional ViennaRNA evaluations): equal-puzzle Factor-to-Direct conditional-ESS ratio with puzzle-bootstrap 95% interval, and condition–replicate cells supported by both pools.

Condition	Factors	Ratio	Interval	Supported
GC 45–55%	1	1.20	$[0.92, 1.61]$	36/36
No 4-base run	1	1.39	$[1.06, 1.91]$	36/36
Local diversity	1	1.11	$[0.97, 1.27]$	31/36
GC + no run	2	1.23	$[1.04, 1.44]$	36/36
GC + diversity	2	0.93	$[0.83, 1.04]$	29/36
No run + diversity	2	1.07	$[0.96, 1.20]$	31/36
All three	3	0.94	$[0.85, 1.03]$	27/36
Leftmost CG pair	1	1.23	$[0.97, 1.61]$	36/36
All three + CG	4	0.97	$[0.83, 1.13]$	19/36

Table 4: Exact Puzzle-8 composition for all nine conditions. $q(C)$ is the exact represented event mass; the bound is the represented-object TT-SVD bound of the rounded conditioned proposal; TV is its exact distance to the untruncated conditioned law; ESS kept is population ESS relative to exact rejection conditioning; the last column divides projected rejection time by measured online composition time for 10^4 samples.

Condition	$q(C)$	Rank	Bound	TV	ESS kept	Rejection/composition
GC 45–55%	0.1658	16	0.0788	0.0178	0.9951	5.84
No 4-base run	0.8679	32	0.0389	0.00266	1.0001	1.03
Local diversity	0.6590	64	0.0642	0.01	0.9994	1.08
GC + no run	0.1569	64	0.0309	0.00399	0.9998	4.58
GC + diversity	0.1254	64	0.159	0.0269	0.9943	5.83
No run + diversity	0.6352	64	0.122	0.0163	0.9992	1.02
All three	0.1232	64	0.249	0.0355	0.9912	4.93
Leftmost CG pair	0.2251	8	4.2×10^{-17}	2.9×10^{-14}	1.0000	4.37
All three + CG	0.0220	64	4.2×10^{-4}	7.7×10^{-5}	1.0000	33.70

not win every estimand: for query mass within 0.03, the fractions were 0.390, 0.407, and 0.395. The matched-256 stronger-SMC check favored factor on four of five puzzles.

Yield objective and its EternaFold transfer. The strength-0.5 arm targets a different objective: unique sequences whose ViennaRNA MFE is the target. On the nine completed cohort puzzles it produced 676 such sequences versus 39 for direct sampling at matched emitted scores (286 versus 2 with $P_{\mathrm{V}}(S \mid x) \geq 0.5$), but puzzles 45 and 13 contributed 655 of the 676, and its diversity retention relative to direct sampling was 0.951 (Hamming), 0.944 (site entropy), and 0.820 (unique predicted structures). ViennaRNA's RNAinverse remains the stronger optimizer for this objective. Under EternaFold, the unweighted yield pool reduced defect by 0.1566 with interval $[0.0684, 0.2455]$ over nine puzzles; exact EternaFold target folds were sparse (0, 7, and 20 of 576 for direct, strength 0.25, and strength 0.5), all 20 on puzzles 45 and 13. For the corrected strength-0.25 comparison of the main text, six of eight puzzles were favorable and every leave-one-puzzle-out mean remained positive (0.0465–0.0753).

Attribution and sensitivity. Replacing the canonical factor multiplication order by deterministic shuffled schedules changed corrected ESS by $1.0068\times$ $[0.9916, 1.0250]$ and reused conditional ESS by $0.9971\times$ $[0.9855, 1.0086]$: the results are robust to order, and no gain is attributed to it. Permuting nucleotide labels along each factor axis, which preserves every value multiset and local singular spectrum, gave $1.3410\times$ $[0.8861, 2.2303]$ globally and $1.0879\times$ $[0.9880, 1.2111]$ for reuse; only the prespecified homopolymer-avoidance condition resolved ($1.34\times$ $[1.03, 1.79]$), so the results are not broadly attributed to detailed Turner values. On development puzzles 1, 23, and 26, ranks 4, 8, 16, and 32 were compared descriptively: higher ranks tightened the factor-product bound (puzzle 23: 0.732, 0.027, 0.0016, 0) without uniformly improving corrected ESS (puzzle 23: 4.06, 3.89, 3.85, 3.85), so no rank-selection rule is inferred.

Table 5: Five measured Puzzle-3 workloads on one host class (seconds). Composition totals add condition-specific setup or all setup (including factor extraction and base compilation) to measured composed sampling; ratios are rejection time divided by composition time.

		Composition			Rejection/composition ratio		
Samples	Rejection	Sampling	+ condition	+ all	Sampling	+ condition	+ all
1,000	25.74	1.44	192.57	318.97	17.93	0.134	0.081
5,000	130.25	7.08	198.22	324.61	18.40	0.657	0.401
10,000	258.98	14.04	205.18	331.58	18.44	1.262	0.781
20,000	516.14	28.07	219.21	345.61	18.38	2.355	1.493
50,000	1299.91	69.48	260.62	387.02	18.71	4.988	3.359

Table 6: Prospective conditional accuracy by completed puzzle: fraction of reference-resolved cells within 0.05 of the reference.

Puzzle	Resolved cells	Factor	SMC	Factor$-$SMC
88	40	1.000	0.950	$+0.050$
40	40	0.000	0.025	-0.025
66	40	1.000	0.850	$+0.150$
33	4	1.000	0.750	$+0.250$
25	28	1.000	0.714	$+0.286$