Entropy-Charged Soft Spectral Decompositions with Applications to Unique Games

Shiva Kintali*

September 5, 2026

Abstract

We construct a soft spectral decomposition whose loss is charged to the information revealed about one root endpoint pair. For a half-lazy weighted graph, the leaf multipliers form a squared partition of unity, every leaf has low base threshold rank, and the total deleted edge mass is bounded by the component-conditional endpoint entropy. A rational late-time heat selection supplies the local energy-to-information bound; the entire tree has finite implicit access.

Combining this decomposition with the Guruswami–Sinop marginal-decoding framework gives an end-to-end Unique Games guarantee governed by the smaller of a vertexwise-information envelope and a component-entropy envelope. In the general worst case, an instance of optimum at least $1-\kappa$ admits a labeling of error $O(\sqrt{\kappa \log(2/\alpha)/\alpha})$ in deterministic time $q^{O(n^{\alpha})}$ times an input-bit polynomial. The dependence on κ remains a square root at fixed α.

1 Introduction

1.1 Overview and motivation

A Unique Games instance assigns a permutation constraint to each record of a weighted graph. If a directed record $e : u \to v$ carries the permutation π_e, a labeling ℓ satisfies it when $\ell(v) = \pi_e(\ell(u))$. For n vertices and an alphabet of size q, the near-perfect recovery problem asks for a labeling of small normalized violated weight under the promise that some labeling violates at most a κ fraction of the total weight. This is the algorithmic regime relevant to Khot's Unique Games Conjecture [17]. Our focus is the quantitative relationship between the promised error κ, the recovered error, and the running time.

Spectral algorithms exploit the fact that a highly satisfiable labeling has small constraint energy. On a graph with suitable expansion or low-dimensional spectral structure, this observation can be converted into an efficient recovery procedure. On an arbitrary graph, however, one must first find useful restrictions and account for everything lost between them. Arora–Barak–Steurer [2] established the subexponential-time decomposition approach, with a detailed near-perfect tradeoff in Steurer's thesis [21]. This paper concentrates on the decomposition cost within that approach.

The guiding idea is to charge a split for the information it reveals, not to charge each recursion level independently. We use *soft* restrictions: a vertex can occur in several leaves, with nonnegative multipliers whose squares sum to one. These multipliers specify an actual retained mass for every edge. A single latent ordered endpoint pair then couples all levels of one decomposition tree, so the chain rule for mutual information bounds the total information spent. Routing the root connected

*Email: shiva.kintali@gmail.com Homepage: shivakintali.github.io.

components costs no edge mass, and its information contribution can be subtracted exactly. This yields the principal structural contribution—an entropy-charged soft spectral decomposition—from which the algorithmic results follow.

1.2 Principal results and their interpretation

A graph theorem independent of Unique Games. An endpoint law is a probability distribution a_{xy} on ordered vertex pairs. Here $a_{xy} = a_{yx}$, its marginal is $\mu(x) = \sum_y a_{xy} > 0$, and half-laziness means $a_{xx} \geq \mu(x)/2$. Its walk and scalar Dirichlet energy are $P(x,y) = a_{xy}/\mu(x)$ and $\mathcal{E}_P(f) = \frac{1}{2}\sum_{x,y} a_{xy}(f(x)-f(y))^2$. All logarithms and entropies use natural logarithms. Theorem 3.1 takes a rational symmetric half-lazy endpoint law a on at most n vertices, an integer $256 \leq R \leq n$, and rational parameters $0 < \nu \leq 1/n$ and $0 < \eta \leq 1/4$ for the minimum-mass threshold and spectral cutoff. Write $L = \log(1/\nu)$ and $L_R = L/\log R$. Under $3072\eta L_R \leq 1/34$, it constructs nonnegative rational leaf multipliers Φ_b with

$$\sum_b \Phi_b(x)^2 = 1, \qquad Z_b = \sum_{x,y} a_{xy}\Phi_b(x)\Phi_b(y), \qquad \operatorname{rank} \mathbf{1}_{[0,\eta]}(I - P_b) \leq R^{5/8},$$

where P_b is the normalized walk of the corresponding weighted restriction. The rank counts, with multiplicity, its normalized Laplacian eigenvalues in the inclusive interval $[0,\eta]$. Its total deleted edge mass, equivalently the total scalar multiplier energy, satisfies

$$D = 1 - \sum_b Z_b = \sum_b \mathcal{E}_P(\Phi_b) \leq 191488\frac{\eta\log(2L_R)}{\log R}H(E \mid C) + 2n\nu.$$

Here E is the root ordered endpoint pair and C its connected component. Thus the entropy term concerns uncertainty *within* components, not the number or relative weights of the components alone. The tree has depth at most $n-1$. Any prescribed path is computable in $\mathrm{poly}(n, \text{input length}, 1/\eta)$ bit operations. The associated sampler reaches leaf b with unconditional probability Z_b and dies with total probability D. Exact rational sampling has expected polynomial bit cost; a bounded implementation permits an explicitly controlled additional failure probability. The construction does not list a polynomial-size family of leaves, and no such bound is claimed. Its overlapping restrictions are not a hard partition of the kind produced by the earlier ABS decomposition.

A complete-labeling consequence. For $n \geq 256$, $q \geq 2$, an integer $256 \leq R \leq n$, and rational $0 < \kappa < 1$, let OPT denote the maximum satisfied fraction of constraint weight. Under the promise $\mathrm{OPT} \geq 1 - \kappa$, Theorem 4.2 combines the graph construction with a finite version of the Guruswami–Sinop conditional-marginal framework [15, 14]. Set

$$\nu = \frac{\kappa}{100n}, \qquad L_R = \frac{\log(100n/\kappa)}{\log R}, \qquad \alpha_H = \frac{\log(2L_R)}{2\log R}.$$

The recovered error is $O(\sqrt{\kappa(L_R + \alpha_H H_*)})$. The static quantity H_*, defined in Section 4, is the minimum of two static information budgets obtained by integrating residual entropy envelopes. It is not a pointwise minimum taken before integration. One envelope uses the vertexwise distribution of endpoint surprisal, and therefore records where information is concentrated; the other uses the original component-conditional entropy. In particular, $0 \leq H_* \leq 2\log n$. This is not merely a guarantee for one favorable leaf: it accounts for successive residual instances, interstage edges, and the final unprocessed mass. The randomized implementation has time $q^{O(R^{5/8})}\,\mathrm{poly}(n, q, b_{\mathrm{in}}, 1/\kappa, \log(1/\xi))$ and success probability at least $1 - \xi$, for supplied rational $0 < \xi < 1$. The deterministic implementation has an additional explicit entropy-dependent path-search exponent. Here and below b_{in} denotes the binary input length, including supplied numerical parameters.

The uniform tradeoff and a promise-free implementation. Theorem 4.4 composes the preceding algorithm with the Gupta–Talwar logarithmic approximation [13]. This removes the numerical $1/\kappa$ cost and absorbs the deterministic path-search exponent in the general worst-case guarantee. In a common time-exponent parameter, Corollary 4.5 gives

$$\begin{aligned} &\mathrm{err}(\ell) \le \min\left\{1, 600\kappa\log(2n), C_{\mathrm{tr}}\sqrt{\frac{\kappa}{\alpha}\log\frac{2}{\alpha}}\right\}, \\ &\text{time } q^{O(n^{\alpha})}\,\mathrm{poly}(n, q, b_{\mathrm{in}}), \qquad 0 < \alpha < 5/8. \end{aligned} \tag{1}$$

Here α is supplied as a rational parameter. The constants are absolute, including for parameters varying with n. Corollary 4.6, stated in the main text and proved in Appendix B.1, removes the need to supply κ: its uniform bound is expressed in terms of the unknown error $1 - \mathrm{OPT}$. It makes only $O(1 + \log\log n)$ bounded runs of the uniform algorithm; each run has its own finite sequence of localization invocations.

1.3 Contributions

The new contribution is the loss–rank–information tradeoff of Theorem 3.1 together with its finite soft-restriction construction. Its proof combines a late-time heat selection that controls energy per unit of support information (Lemma 3.2), a rational squared partition of unity with quadratic loss (Lemma 3.3), and an exact posterior coupling that charges all smooth losses to $H(E \mid C)$ through the chain rule, without any branchwise entropy-monotonicity assumption. Spectral localization, heat evolution, threshold rank, and the chain rule themselves are classical [18, 2, 21, 9].

The algorithmic theorems are consequences of that graph theorem and of the Guruswami–Sinop tools [15, 14], with two application-specific arguments: transporting the conditional-marginal ensemble back through the soft multiplier (Appendix A), and controlling every residual instance from the immutable input through the vertexwise information envelope of Lemma 4.1. Column selection, the singleton embedding, and alphabet-efficient low-rank recovery are due to Guruswami–Sinop [15]; their local hierarchy framework [14] supplies the optimization architecture, and finite-precision ellipsoid optimization is classical [12]. The logarithmic fallback and its tree-metric method are those of [13, 10]. Appendix A makes the finite ensemble interface explicit and sharpens the transport by separating label disagreement from amplitude variation: for a good leaf with cutoff rank at most $R^{5/8}$ at $0 < \eta = 1024\bar{\kappa} \le 1/4$, a normalized-leaf labeling of error at most $8\bar{\kappa}$, and $\mathcal{E}_P(\Phi_b)/Z_b \le 8D$, Corollary A.6 produces a nonempty partial assignment with

$$\Gamma^2 \le \frac{65}{432}\eta + \frac{2048}{27}D,$$

where D is the loss of that invocation's decomposition tree and Γ is the partial assignment's internal-violation-plus-twice-boundary mass divided by its assigned vertex mass, measured at the invocation root. Appendix B supplies the weighted rational fallback, and Appendix C groups optional whole-tree loss estimates, guarded policy comparisons, and exactly certified flat partitions; the principal theorems do not depend on Appendix C.

1.4 Related work and quantitative comparison

Polynomial-time recovery and structured graphs. Following Khot's formulation [17], Trevisan [22] gave polynomial-time recovery of a constant fraction of constraints from sufficiently near-perfect instances with completeness error of order $1/\log n$. Gupta–Talwar [13] obtained

Result	Recovered-error bound	Running time and scope
Gupta–Talwar [13]	$O(\kappa \log n)$	Polynomial; general graphs.
Charikar–Makarychev–Makarychev [8]	$O(\sqrt{\kappa \log q})$	Polynomial; nontrivial for small $\kappa \log q$.
Steurer [21, Theorem 5.2]; ABS approach [2]	$O(\sqrt{\kappa/\alpha^3})$	$\exp(O(qn^\alpha))$; general graphs.
Guruswami–Sinop [15, 14]	$\big(1+(2+\theta)/\lambda_{r+1}\big)\kappa$	$q^{O(r/\theta)}\,\mathrm{poly}(n)$; spectral guarantee.
This paper, Corollary 4.5	$O\big(\sqrt{\kappa \log(2/\alpha)/\alpha}\big)$	$q^{O(n^\alpha)}\,\mathrm{poly}(n, q, b_{\mathrm{in}})$; general graphs.

Table 1: Relevant recovery guarantees, with errors capped by one. The displayed thesis benchmark has $\alpha \geq \log\log n/\log n$; our range is $0 < \alpha < 5/8$. Comparison of their error coefficients is on the common range. Polynomial factors are suppressed in the historical exponential-time entry. Randomized historical rounding bounds are stated in expectation; constant-success bounds follow with an absolute constant-factor loss. Alphabet dependence and spectral hypotheses differ.

the logarithmic error guarantee used here, through probabilistic tree embeddings [10]. Charikar–Makarychev–Makarychev [8] gave the alphabet-sensitive $O(\sqrt{\kappa \log q})$ bound, which is nontrivial when $\kappa \log q$ is sufficiently small. For weighted instances, the same per-edge estimate gives this expected-error bound by weighted Jensen's inequality. These polynomial-time algorithms can be preferable to our general tradeoff, particularly for small alphabets.

Expansion gives a different source of favorable structure. Arora–Khot–Kolla–Steurer–Tulsiani–Vishnoi [4] established efficient recovery on expanding constraint graphs. Makarychev–Makarychev [19] sharpened this to error $O(\kappa/h_G)$ when $\kappa < c\lambda_G$, where h_G is edge expansion and λ_G the spectral gap in their normalization. Kolla [18] developed purely spectral recovery whose time depends on the relevant label-extended eigenspace. These hypotheses are not comparable with the arbitrary base graphs treated here.

General decomposition and low-rank hierarchies. ABS [2] is the foundational general subexponential algorithm; Steurer's thesis [21] supplies the explicit two-parameter near-perfect benchmark used below. The use of iterative partial assignments is already present there. Related divide-and-conquer algorithms were developed by Arora–Impagliazzo–Matthews–Steurer [3]. Barak–Raghavendra–Steurer [7] developed hierarchy rounding through global correlation, obtaining guarantees that recover the ABS constant-satisfaction subexponential regime and exploit low threshold rank on suitable instances. Guruswami–Sinop [15, Theorem 6.7 and Remark 6.8] give expected error at most

$$\left(1+\frac{2+\theta}{\lambda_{r+1}}\right)\kappa \quad \text{in time } q^{O(r/\theta)}\,\mathrm{poly}(n),$$

where $0 < \theta < 1$, the integer r satisfies $1 \leq r < n$, and $\lambda_{r+1} > 0$ is the $(r+1)$-st smallest eigenvalue of the base normalized Laplacian; a zero denominator gives no finite guarantee. Their Theorems 5.2 and 6.9 provide the column-selection and embedding ingredients used here. The local solver [14] is important: constructing a full hierarchy with an $n^{O(r)}$ cost would not give our stated alphabet-efficient interface. The error coefficient in our general tradeoff is attributed to the new decomposition analysis; alphabet-efficient decoding is explicitly inherited.

Table 1 separates those hypotheses and time costs. On the common parameter range, the comparison with [21, Theorem 5.2] is at fixed completeness error κ and the same time exponent α. The proved error coefficient changes from $\alpha^{-3/2}$ there to $\sqrt{\log(2/\alpha)/\alpha}$ here. Their ratio,

$1/(\alpha\sqrt{\log(2/\alpha)})$, tends to infinity as $\alpha \to 0$. Equivalently, in the asymptotic regime $x := \kappa/\delta^2 \to 0$, the new upper bound permits $\alpha = \Theta(x \log(1/x))$ to target recovered error $O(\delta)$, whereas the displayed thesis formula is balanced at $\alpha = \Theta(x^{1/3})$. These statements are subject to the respective parameter ranges. Three qualifications are important. First, this compares proved upper bounds; it is not a lower bound on an optimized version of the earlier algorithm. Second, neither bound improves the $\sqrt{\kappa}$ completeness power at fixed α. Third, the alphabet costs differ: $\exp(O(qn^\alpha))$ and $q^{O(n^\alpha)}$ agree up to constants in the exponent for fixed q, but not when q grows.

Other relevant algorithms and analytic tools. Hopkins–Schramm–Trevisan [16] obtain a subexponential Unique Games gap-distinguishing guarantee using Sherali–Adams linear programs, without SDP rounding or subspace enumeration. This is an alternative algorithmic framework; its constant-gap distinction does not by itself give the two-parameter near-perfect recovery guarantee in (1). Alev–Lau [1] use low-diameter decompositions for minor-free and bounded-genus constraint graphs. Bafna–Barak–Kothari–Schramm–Steurer [5] develop rounding for general Unique Games on graphs with certified small-set expansion. Bafna–Minzer [6] treat affine Unique Games on the broader class of globally hypercontractive graphs. These graph-class and, in the latter case, constraint-type assumptions make their guarantees incomparable with the general tradeoff here. The conditioned shift potential in [5, Definition 3.4] concerns weighted collision probabilities of relative labels in a sum-of-squares pseudodistribution. Our budget is instead the Shannon entropy of an actual root endpoint pair, conditioned on its component, and controls recursive deletion.

The analytic ingredients also have established antecedents. Heat evolution and small-support Rayleigh quotients are central to spectral profile methods [11]; the entropy chain rule and relative-entropy identities are standard [9]; and the cube analysis in Section 5 uses classical hypercontractivity [20].

1.5 Overview of our proof strategy

1. Find an informative low-energy split. At a node with all vertex masses at least ν and base cutoff rank exceeding $R^{5/8}$, the lazy walk $Q = (I + P)/2$ has $\mathrm{Tr}(Q^{2T}) > \sqrt{R}$ at the time used in Lemma 3.2. Select one mean-one density column whose heat evolution g_t has $N_T = \|g_T\|_\mu^2 > \sqrt{R}$. For $F_t = \log \|g_t\|_\mu^2 > 0$ and $e_t = \mathcal{E}_P(g_t)/\|g_t\|_\mu^2$, heat contraction gives $e_t \le 2(F_t - F_{t+1})$. Monotonicity of F_t then yields

$$\sum_{t<T} \frac{e_t}{F_t} \le 2\log(F_0/F_T) \le 2\log(2L_R).$$

Selecting a late time and truncating the density turns this estimate into an energy-to-support-information bound containing $\log(2L_R)$. A stereographic parametrization gives a squared partition of unity with quadratic energy loss. A fixed dyadic mesh preserves the bound while discarding the heat-vector denominators before recursion; the encoding length of accumulated multipliers still grows linearly with path depth. These are the local estimates in Lemmas 3.2 and 3.3.

2. Charge the entire tree to one information budget. Draw one root endpoint pair $E = (x, y)$. At a split, select child i with probability $\chi_i(x)\chi_i(y)$ and otherwise kill the path. Nonnegativity and the squared partition of unity make these probabilities valid by Cauchy–Schwarz. Conditional on reaching a node, Bayes' rule gives exactly that node's weighted endpoint law. Thus the chain rule controls the reach-weighted sum of local information by the root entropy, even though entropy need not decrease along an individual surviving branch. Routing the root connected components first makes the routing transcript a deterministic code for C with zero deletion cost,

so its contribution $H(C)$ can be subtracted from the chain rule, leaving the budget $H(E \mid C)$. Combining this budget with the local loss-per-information estimate gives Theorem 3.1; cleanup contributes the explicit $2n\nu$.

3. Decode a good leaf without reconciling overlapping leaves. At an admitted recovery invocation, the unconditional path law gives a constant probability of reaching a leaf with both a good restricted witness and small multiplier energy *relative to retained leaf mass*, as stated in Lemma 3.5. The finite conditional-marginal construction uses the inherited Guruswami–Sinop ingredients [15, 14], together with the finite local-fragment implementation proved in Appendix A. Winner-gap rounding creates an amplitude field with at most one nonzero label coordinate at each vertex. Multiplication by the leaf multiplier and ensemble coarea then produce a nonempty partial assignment, selected by its actual rational internal-violation-plus-twice-boundary charge. Keeping label disagreement separate from amplitude variation sharpens this interface. Only one accepted partial assignment is committed at a time, so no agreement between separately rounded overlapping leaves is required.

4. Preserve the information gain through complete recovery. After each accepted partial assignment, rebuild the residual from the immutable original instance and replace outgoing mass by identity loops. Lemma 4.1 bounds the resulting entropies through the original vertexwise profile and component information. Integrating the partial-assignment charges over remaining mass proves Theorem 4.2, including its stopping and terminal charges. The classical fallback handles parameter ranges in which localization would have excessive numerical or deterministic search cost. Exact satisfiability testing and a bounded geometric guessing schedule then give the promise-free corollary, including termination on false guesses.

5. Delimit what the current analysis does not show. Section 5 computes the full child-or-kill information and shows that high heat trace alone cannot imply a uniform loss-to-information ratio $K/I = O(\eta/\log R)$ without a logarithmic loss. Its cube witness already satisfies the low-rank stopping criterion, however, so the algorithm does not split it. That witness is therefore not a lower bound for the nodes on which the algorithm performs a smooth split. The hypercontractive estimate is classical [20]; the channel calculation and the rank-screen qualification specify the limitation of this argument. Neither a log-free global loss bound nor a better general completeness power follows from this analysis.

1.6 Organization of the paper

Section 2 fixes the weighted input, normalization, constraint-energy, and soft-restriction conventions. Section 3 proves the standalone spectral decomposition and its finite access guarantees. Section 4 establishes the residual information envelopes, complete recovery, uniform tradeoff, and promise-free statement. Section 5 gives the actual-channel analysis and precisely scoped limitation. Appendix A proves the required finite decoder and its sharper transport, coarea, and direct-multiplier refinements. Appendix B gives the classical fallback and the proof of the promise-free wrapper. Appendix C contains the optional actual-loss-sensitive extension. It can be omitted on a first reading without interrupting the principal decomposition–recovery–limitation narrative.

2 Notation and normalization

Definition 2.1 (Unique Games input)**.** An instance consists of a finite vertex set V with $|V| = n \geq 1$, an alphabet $[q] = \{1, \ldots, q\}$ with integer $q \geq 1$, and a finite multiset $\mathcal{R}$ of records $e = (u_e, v_e, \pi_e, w_e)$. Here $u_e, v_e \in V$, π_e is a permutation of $[q]$, and $w_e \geq 0$ is rational. A labeling $\ell : V \to [q]$ satisfies e when $\ell(v_e) = \pi_e(\ell(u_e))$. If $W = \sum_{e \in \mathcal{R}} w_e > 0$, define

$$\operatorname{err}(\ell) = \frac{1}{W} \sum_{e \in \mathcal{R}} w_e \mathbf{1}_{\ell(v_e) \neq \pi_e(\ell(u_e))}, \qquad \mathrm{OPT} = 1 - \min_{\ell} \operatorname{err}(\ell).$$

If $W = 0$, set every labeling's error to zero and $\mathrm{OPT} = 1$. Loops and parallel record occurrences are allowed. The input length includes the explicit permutations and all rational weights and parameters. Running times below count bit operations.

Preprocessing and orientation. Discard every zero-weight record before constructing any law or operator. If no positive-weight record remains, return arbitrary labels and perform no normalization. Thus all normalized-law statements below concern the case $W > 0$. A vertex incident to no positive-weight record is likewise omitted from every law and operator and receives an arbitrary output label; n continues to denote the original vertex-count bound. Unless a later transform is stated explicitly, "record" henceforth refers to this preprocessed multiset.

For each remaining record put $p_e = w_e/W$. Replace its occurrence by the two occurrences

$$(u_e, v_e, \pi_e, p_e/2) \quad \text{and} \quad (v_e, u_e, \pi_e^{-1}, p_e/2).$$

Both occurrences are retained even when $u_e = v_e$. Their satisfaction conditions are equivalent, so this symmetrization preserves the error of every labeling. Let $\mathring{a}_{xy}$ be the total mass of the resulting occurrences directed from x to y, and put $\mu^0(x) = \sum_y \mathring{a}_{xy} > 0$. Then $\mathring{a}$ is symmetric, has total mass one, and has marginal μ^0.

Half-lazification and immutable residuals. Multiply the mass of every symmetrized record occurrence by $1/2$ and add an identity-permutation loop of mass $\mu^0(x)/2$ at each vertex x. The resulting scalar endpoint law is

$$a^0_{xy} = \tfrac{1}{2}\mathring{a}_{xy} + \tfrac{1}{2}\mathbf{1}_{x=y}\mu^0(x).$$

It has total mass one, is symmetric, has marginal μ^0, and is half-lazy: $a^0_{xx} \geq \mu^0(x)/2$. The new loops are always satisfied, so the error in this normalized record law is exactly $\operatorname{err}(\ell)/2$. We retain constraint records, including nonidentity loops, for decoding; only their scalar endpoint masses are aggregated for entropy and base-graph spectral calculations.

For a residual vertex set U with $t = \mu^0(U) > 0$, start again from these immutable half-lazified records. Keep every record whose two endpoints lie in U; replace each record from $x \in U$ to $y \notin U$ by an identity loop at x of the same mass; and discard every record whose source lies outside U. The mass before renormalization is t, so divide every retained mass by t. The resulting scalar endpoint law is

$$a^U_{xy} = \frac{a^0_{xy} + \mathbf{1}_{x=y} \sum_{z \notin U} a^0_{xz}}{t} \qquad (x, y \in U).$$

For $x \in U$, summing row x gives $\mu_U(x) = \mu^0(x)/t$; consequently a^U has total mass one. Its off-diagonal entries are symmetric, while every new boundary loop is its own reverse, so the completed record law is symmetric. The retained identity-loop mass introduced during half-lazification is by itself at least $\mu_U(x)/2$. If an input labeling ℓ has $\operatorname{err}(\ell) \leq \kappa$, then its restriction to U has error at

most $\kappa/(2t)$ in this completed residual: the completion loops are satisfied, and the total violated mass in the immutable half-lazified law is at most $\kappa/2$. Later arguments sometimes use the weaker upper bound κ/t.

Let $E_U = (X_U, Y_U)$ be a random ordered pair with law a^U, and let C_U be its connected-component variable in the scalar graph with edges of positive mass. Put $H_U = H(E_U)$ and $h_U = H(E_U \mid C_U)$. The original pair and its component variable are denoted $E^0 \sim a^0$ and C_0, respectively, and we write $H_0 = H(E^0)$ and $h_0 = H(E^0 \mid C_0)$. Equivalently, to sample E_U, sample an endpoint pair from a^0 conditional on its first endpoint lying in U, then move the second endpoint to the first when it lies outside U. These completed residuals are the recovery invocation roots; soft leaf weights are never carried over to the next invocation.

Scalar and constraint energies. Let $a = (a_{xy})$ be a nonnegative symmetric scalar law of total mass one, restricted to its positive-degree vertices. Write $\mu(x) = \sum_y a_{xy} > 0$, $P_{xy} = a_{xy}/\mu(x)$, and, for real scalar fields f, g,

$$\langle f, g\rangle_\mu = \sum_x \mu(x) f(x) g(x), \qquad \mathcal{E}_P(f) = \frac{1}{2}\sum_{x,y} a_{xy}(f(x) - f(y))^2 = \langle f, (I - P)f\rangle_\mu .$$

The identity follows by expansion and symmetry. In particular P is a self-adjoint Markov operator in $L^2(\mu)$. If $\mathsf{D} = \mathrm{diag}(\mu)$, its symmetric normalized Laplacian is

$$\mathcal{L} = I - \mathsf{D}^{-1/2} a \mathsf{D}^{-1/2} = \mathsf{D}^{1/2}(I - P)\mathsf{D}^{-1/2}.$$

Thus $\mathcal{L}$ in Euclidean coordinates and $I - P$ in $L^2(\mu)$ have the same cutoff eigenvalues and multiplicities. All cutoff ranks in this paper count the inclusive interval $[0, \eta]$ unless stated otherwise.

When the law carries constraints, a_e denotes the mass of a directed record $e : u \to v$, so $\sum_e a_e = 1$ and $a_{uv} = \sum_{e:u\to v} a_e$. The record multiset has a mass-preserving involution $e \mapsto \bar{e}$ that reverses the endpoints and replaces π_e by π_e^{-1}; an identity loop may be fixed by this involution. For a vector $z \in \mathbb{R}^q$, define the pullback by $(T_e z)_i = z_{\pi_e(i)}$. Fiber fields $z_u \in \mathbb{R}^q$ use the degree-weighted norm $\|z\|_\mu^2 = \sum_u \mu(u)\,\|z_u\|_2^2$ and energy

$$\widehat{\mathcal{E}}_P(z) = \frac{1}{2}\sum_e a_e \,\|z_u - T_e z_v\|_2^2 .$$

We omit the subscript when the law is specified. A one-hot field $z_u = e_{\ell(u)}$, with e_i the ith standard basis vector, has norm one. Its squared difference at a record is zero when that record is satisfied and two otherwise; hence its energy is exactly the labeling error in the current normalized record law. At the initial half-lazified root this is half the original input error.

The permutation lift has a directed edge of mass a_e from (u, i) to $(v, \pi_e(i))$ for every e and $i \in [q]$. Reverse records make this lift symmetric. Its degree measure is $\widehat{\mu}(u, i) = \mu(u)$, of total mass q, not the probability measure $\mu(u)/q$. Under the identification $f(u, i) = (z_u)_i$, its scalar Dirichlet form, with this unnormalized degree measure, is exactly $\widehat{\mathcal{E}}_P$; no factor of q is introduced. A loop at u with a nonidentity permutation can join different lift coordinates, so it can have positive constraint energy even though its scalar base-graph energy is always zero.

Soft restrictions and their degree measure. Here a is the half-lazy scalar law at the fixed root of one decomposition invocation, with positive marginal μ. For a nonzero nonnegative multiplier Φ on this vertex set, define

$$M_\Phi = \sum_u \mu(u)\Phi(u)^2, \qquad Z_\Phi = \sum_{u,v} a_{uv}\Phi(u)\Phi(v), \qquad M_\Phi - Z_\Phi = \mathcal{E}_P(\Phi).$$

The energy identity again follows by expansion. The raw restricted degree and its normalized marginal are

$$d_\Phi(x) = \sum_y a_{xy}\Phi(x)\Phi(y), \qquad \mu_\Phi(x) = d_\Phi(x)/Z_\Phi.$$

Here $\operatorname{supp}\Phi = \{x : \Phi(x) > 0\}$. These normalizations are well-defined on this support because half-laziness and symmetry give

$$d_\Phi(x) \geq \tfrac{1}{2}\mu(x)\Phi(x)^2, \qquad 0 < M_\Phi/2 \leq Z_\Phi \leq M_\Phi.$$

For the upper bound, sum $2\Phi(x)\Phi(y) \leq \Phi(x)^2 + \Phi(y)^2$ against a_{xy}. The restricted scalar law is $a^\Phi_{xy} = a_{xy}\Phi(x)\Phi(y)/Z_\Phi$, and its constraint-record masses, when present, are $a_e\Phi(u)\Phi(v)/Z_\Phi$. Vertices with zero multiplier are deleted; on the remaining vertices, μ_Φ is exactly the marginal of a^Φ. In general μ_Φ is not $\mu\Phi^2/M_\Phi$.

For example, on two vertices take $a_{xy} = 1/4$ for every pair and $\Phi = (1/2, 1)$. Then $M_\Phi = 5/8$, $Z_\Phi = 9/16$, and $d_\Phi = (3/16, 3/8)$, giving $\mu_\Phi = (1/3, 2/3)$, whereas $\mu\Phi^2/M_\Phi = (1/5, 4/5)$. Indeed

$$a^\Phi = \begin{pmatrix} 1/9 & 2/9 \\ 2/9 & 4/9 \end{pmatrix}, \qquad P^\Phi = \begin{pmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{pmatrix}.$$

Thus $P^\Phi_{11} = 1/3 < 1/2$, although the root was half-lazy.

All accumulated multipliers and their masses above refer to the fixed invocation root; local heat and split computations instead use the current normalized node law. A Unique Games recovery root has identity-loop mass at least half its degree. Descendants need not retain that pointwise fraction relative to their new degrees, as the example illustrates even for scalar diagonal mass. At every node $Q = (I + P)/2$ is used for analytic laziness: reversibility and the fact that a self-adjoint Markov operator has spectrum in $[-1, 1]$ give $0 \preceq Q \preceq I$ on the node's $L^2(\mu)$ space. Its scalar Dirichlet form satisfies $\mathcal{E}_Q(f) = \mathcal{E}_P(f)/2$. Using this heat operator does not apply a second half-lazification to the constraint objective.

Entropy conventions. All logarithms are natural unless a base is displayed. For a finite random variable with probabilities p_z, entropy is $-\sum_{z:p_z>0} p_z \log p_z$; zero-probability summands are omitted also in conditional entropy and mutual information. Endpoint entropy concerns the scalar ordered pair, not the record identity, permutation, or labels. Since a positive-mass endpoint pair never joins distinct connected components, its component variable is a deterministic function of either endpoint. In particular, $H(E_U \mid C_U) = H(E_U) - H(C_U)$, and similarly at the root.

Global notation. Table 2 lists the symbols that are used across sections; symbols introduced inside a single statement or proof are local to it. Vertices are always u, v, x, y; the scalar ratios $L_R = L/\log R$ and $\Lambda_R = \log n/\log R$ carry the subscript R so that they cannot be confused with vertices.

$n,\ q,\ b_{\mathrm{in}}$	number of vertices, alphabet size, binary input length
$u, v, x, y;\ E = (X, Y),\ C$	vertices; the random ordered endpoint pair and its component
$a,\ \mu,\ P,\ \mathcal{E}_P$	endpoint law, marginal, walk, scalar Dirichlet energy
a^0, μ^0, E^0, C_0	half-lazified input law with its pair and component variables
$a^U, \mu_U, E_U, C_U,\ t = \mu^0(U)$	completed residual on U and its mass
$H_U,\ h_U,\ H_0,\ h_0,\ H_{\max}$	$H(E_U)$, $H(E_U \mid C_U)$, their root values, and $2\log n$
$\widehat{\mathcal{E}}_P,\ T_e$	constraint energy and permutation pullback
$\Phi_b, \chi_i,\ M_\Phi, Z_\Phi, \mu_\Phi$	leaf and local multipliers; mass, retained mass, degree law
$R,\ \eta,\ \nu$	rank parameter, spectral cutoff, minimum-mass threshold
$L,\ L_R,\ J$	$\log(1/\nu)$, $L/\log R$, $\log(2L_R)$
$D,\ K_t,\ Z_t,\ N_{\mathrm{sm}}$	deleted mass; kill and reach probabilities of node t; smooth-child count
$\kappa,\ \bar{\kappa},\ \kappa_0$	input, invocation, and decoder error parameters
$\alpha_H;\ \mathcal{B}_U, \overline{\mathcal{B}}_U;\ H_*$	entropy coefficient; exact/rational stage budgets; static budget
Γ	internal-violation-plus-twice-boundary charge of a partial assignment
$\Lambda_R,\ \mathcal{B}_n$	$\log n/\log R$ and $\Lambda_R(1+\log(2\Lambda_R))$

Table 2: Symbols used across sections.

3 Entropy-charged spectral decomposition

The construction below is a graph theorem: it does not use a labeling promise or an alphabet. Its output is an implicitly accessible family of overlapping weighted restrictions, not a hard partition.

Theorem 3.1 (Soft spectral decomposition)**.** *Let a be a rational symmetric endpoint law on at most n vertices, with positive marginal μ and $a_{xx} \geq \mu(x)/2$. Fix an integer $256 \leq R \leq n$ and rational $0 < \nu \leq 1/n$, $0 < \eta \leq 1/4$. Put*

$$L = \log(1/\nu), \qquad L_R = L/\log R, \qquad J = \log(2L_R),$$

Assume $3072\eta L_R \leq 1/34$. There is a deterministic finite binary tree whose internal nodes carry nonnegative rational local multipliers χ_0, χ_1 satisfying $\chi_0^2 + \chi_1^2 = 1$, with each child support a nonempty proper subset of its parent support. Its leaf products Φ_b, extended by zero to the root, satisfy

$$\sum_b \Phi_b(x)^2 = 1, \qquad M_b = \sum_x \mu(x)\Phi_b(x)^2, \quad Z_b = \sum_{x,y} a_{xy}\Phi_b(x)\Phi_b(y), \quad M_b/2 \leq Z_b \leq M_b.$$

Each normalized restriction $a^b_{xy} = a_{xy}\Phi_b(x)\Phi_b(y)/Z_b$ has positive degrees and

$$\operatorname{rank} \mathbf{1}_{[0,\eta]}(I - P_b) \leq R^{5/8}. \tag{2}$$

The cutoff is inclusive and the projector is taken in the weighted Hilbert space. If $E \sim a$, and C is its root connected-component variable, write $h = H(E \mid C)$. Then

$$D := \sum_b \mathcal{E}_P(\Phi_b) = 1 - \sum_b Z_b \leq \min\left\{1 - \sum_x a_{xx},\ 191488\frac{\eta J}{\log R}h + 2n\nu\right\}. \tag{3}$$

In particular $D \leq 1/2$, $\sum_b M_b = 1$, and $h \leq 2\log n$.

The tree has depth at most $n-1$. Any specified path can be computed in $\mathrm{poly}(n, b_{\mathrm{in}}, 1/\eta)$ bit operations, with absolute polynomial degree; b_{in} includes the encoding of ν and η. An exact sampler, with expected polynomial bit cost, reaches leaf b with unconditional probability Z_b and otherwise dies.

For any integer $s \geq 1$, a capped version has at most 2^{-s} additional failure probability and bit cost polynomial also in s. For the exact sampler, let N_{sm} count smooth child-1 choices before the sampler reaches a leaf or dies; then $\mathbb{E}N_{\mathrm{sm}} \leq 4h/\log R$. Write $N_{\mathrm{sm}}(b)$ for this count along the path to leaf b. For every integer $k \geq 0$, every leaf with $N_{\mathrm{sm}}(b) \leq k$ can be found deterministically in at most

$$n \sum_{j=0}^{\min\{k,n-1\}} \binom{n-1}{j} \tag{4}$$

path simulations. Full-tree enumeration is not claimed to be polynomial.

Here and below, H and mutual information use natural logarithms. We prove the theorem through two local estimates and one global information calculation.

3.1 A rational choice of heat time

Lemma 3.2 (Rational heat localization)**.** *Suppose a current node has a rational normalized symmetric endpoint law, all marginal masses at least ν, and $\operatorname{rank} \mathbf{1}_{[0,\eta]}(I-P) > R^{5/8}$. With the parameters of Theorem 3.1, a rational computation finds a nonzero rational vector $g^\sharp \geq 0$ and rational λ such that*

$$\varsigma := \mu(\operatorname{supp} g^\sharp) \leq R^{-1/4}, \qquad \mathcal{E}_P(g^\sharp) \leq \lambda \left\|g^\sharp\right\|_\mu^2, \qquad \eta \leq \lambda \leq 3072\eta L_R, \qquad \frac{\lambda}{\log(1/\varsigma)} \leq \frac{44J}{T},$$

where $T = \lfloor \lfloor \log_2 R \rfloor / (32\eta) \rfloor$ and $\log R/(128\eta) \leq T \leq \log R/(16\eta)$.

Proof. Let $Q = (I+P)/2$. It is positive semidefinite in $L^2(\mu)$ even when the current node is not half-lazy. Here $L_R \geq 1$, so $\eta \leq 1/104448$. Writing $\ell = \lfloor \log_2 R \rfloor \geq 8$, we have $\ell \geq \frac{1}{2}\log_2 R$ and $\ell/(64\eta) \leq T \leq \ell/(32\eta)$; these imply $T \geq 2$ and the stated bounds, using $1/2 < \log 2 < 1$ to convert between binary and natural logarithms. Every eigenvalue of $I-P$ at most η contributes at least $(1-\eta/2)^{2T} \geq \exp(-2\eta T) \geq R^{-1/8}$ to $\operatorname{Tr}(Q^{2T})$, where $\log(1-\eta/2) \geq -\eta$ for $0 \leq \eta \leq 1/4$. Thus rank failure implies $\operatorname{Tr}(Q^{2T}) > \sqrt{R}$.

For density columns $g_{t,i} = Q^t \mathbf{1}_i/\mu(i)$, reversibility gives $\sum_i \mu(i) \left\|g_{T,i}\right\|_\mu^2 = \operatorname{Tr}(Q^{2T})$. Choose a column maximizing $N_T = \left\|g_T\right\|_\mu^2$. Every g_t is nonnegative with mean one, $N_0 = 1/\mu(i) \leq 1/\nu$, and $N_T > \sqrt{R}$. The spectral expansion writes N_t as a nonnegative combination of ρ^{2t}, $0 \leq \rho \leq 1$, so N_t is nonincreasing and log-convex by Cauchy–Schwarz (with $\rho^0 = 1$ also for $\rho = 0$). Set $F_t = \log N_t$ and $e_t = \mathcal{E}_P(g_t)/N_t$. Since $I - Q^2 \succeq I - Q$,

$$F_0 \leq L, \quad F_t \geq F_T > \tfrac{1}{2}\log R \geq \log 16, \qquad e_t \leq 2(1 - N_{t+1}/N_t) \leq 2(F_t - F_{t+1}).$$

For positive $F_t \geq F_{t+1}$, the last difference divided by F_t is at most $\log(F_t/F_{t+1})$. Consequently

$$\sum_{t=0}^{T-1} e_t/F_t \leq 2\log(F_0/F_T) \leq 2J. \tag{5}$$

There are at least $T/2$ integers in the range $\lfloor T/2 \rfloor \leq t < T$, so some such time has $e_t/F_t \leq 4J/T$. At each candidate time form $g_t^\sharp = (g_t - N_t/4)_+$. Markov's inequality gives $\varsigma_t \leq 4/N_t \leq R^{-1/4}$, while $\left\|\min(g_t, N_t/4)\right\|_\mu^2 \leq N_t/4$ follows from $\min(g_t, N_t/4)^2 \leq (N_t/4)g_t$ pointwise. The reverse triangle inequality now gives $\left\|g_t^\sharp\right\|_\mu \geq \sqrt{N_t} - \sqrt{N_t}/2$, so $g_t^\sharp \neq 0$ and $\left\|g_t^\sharp\right\|_\mu^2 \geq N_t/4$. The scalar positive-part map is 1-Lipschitz and therefore decreases energy. Hence

$$\frac{\mathcal{E}_P(g_t^\sharp)}{\left\|g_t^\sharp\right\|_\mu^2} \leq 4e_t, \qquad \log(1/\varsigma_t) \geq F_t/2.$$

For clarity, log-convexity makes $d_t = F_t - F_{t+1}$ nonnegative and nonincreasing. Thus $td_t \le \sum_{s<t} d_s = F_0 - F_t$, and every candidate $t \ge T/3$ satisfies $e_t \le 2(F_0 - F_t)/t \le 6L/T$.

No transcendental comparison is used to select a time. Compute $\lambda_t = \max\{\eta, \mathcal{E}_P(g_t^\sharp)/\left\|g_t^\sharp\right\|_\mu^2\}$ and $m_t = \lfloor \log_2 N_t \rfloor$ by rational power comparisons, and minimize λ_t/m_t. Here $m_t \ge 4$ and $m_t \log 2 \le F_t < \frac{5}{4} m_t \log 2$. Also $8\eta T \le \frac{1}{2}\log R < F_t$, so $\eta/F_t \le 1/(8T) \le J/(4T)$, where $J \ge \log 2 > 1/2$. The good time from (5) therefore has $\lambda_t/F_t \le 17J/T$; rational selection loses at most $5/4$. Indeed, for a minimizing time s and the good time t, $\lambda_s/F_s \le \lambda_s/(m_s \log 2) \le \lambda_t/(m_t \log 2) < (5/4)\lambda_t/F_t$. The chosen pair obeys $\lambda/F_s < 22J/T$ and hence the claimed ratio because $\log(1/\varsigma_s) \ge F_s/2$. Finally $\eta T \le \log R/16 \le L/16$, so at every candidate the bounds $\mathcal{E}_P(g_t^\sharp)/\left\|g_t^\sharp\right\|_\mu^2 \le 4e_t \le 24L/T$ and $\eta \le L/(16T)$ give $\lambda_t \le 24L/T \le 3072\eta L_R$. All vectors are rational matrix powers; their bit bounds are supplied in the construction below. □

3.2 A squared partition with quadratic loss

Lemma 3.3 (Rational squared split)**.** *Let the current rational node and $g^\sharp, \lambda, \varsigma$ satisfy the hypotheses and conclusion of Lemma 3.2; in particular, every current marginal mass is at least ν. One can compute nonnegative rational multipliers with proper nonempty supports and $\chi_0^2 + \chi_1^2 = 1$ such that, in the current normalized law,*

$$p_i = \sum_{x,y} a_{xy}\chi_i(x)\chi_i(y), \quad K = 1 - p_0 - p_1 \le 34\lambda p_1, \quad \mu(\operatorname{supp}\chi_1) \le \varsigma.$$

Every multiplier coordinate has $O(\log(1/\eta) + \log(1/\nu))$ bits, independently of the heat vector's denominator.

Proof. For $\vartheta = g^\sharp/\max g^\sharp$, first use $\chi_1 = 2\vartheta/(1+\vartheta^2)$ and $\chi_0 = (1-\vartheta^2)/(1+\vartheta^2)$. Writing $\mathsf{S}(\vartheta) = ((1-\vartheta^2)/(1+\vartheta^2), 2\vartheta/(1+\vartheta^2))$, we have $\|\mathsf{S}(\vartheta) - \mathsf{S}(s)\|_2^2 = 4(\vartheta - s)^2/((1+\vartheta^2)(1+s^2))$ on $[0,1]$. In particular S is 2-Lipschitz, and $\chi_1 \ge \vartheta$. Thus, with $M_1 = \|\chi_1\|_\mu^2$,

$$K = \mathcal{E}_P(\chi_0) + \mathcal{E}_P(\chi_1) \le 4\mathcal{E}_P(\vartheta) \le 4\lambda M_1.$$

Choose the smallest integer $j \ge 0$ for which the dyadic mesh $\Delta = 2^{-j}$ satisfies $\Delta^2 \le \eta\nu/10000$, and round ϑ down to this mesh, and apply the same map, denoting the results by primes. The vector perturbation $v = \mathsf{S}(\vartheta') - \mathsf{S}(\vartheta)$ has pointwise Euclidean norm at most 2Δ. Reversibility gives $\mathcal{E}_P(v) \le 2\|v\|_\mu^2 \le 8\Delta^2$, with vector energy defined as the sum of its coordinate energies. The squared triangle inequality in edge space and vertex space gives

$$K' \le 2K + 16\Delta^2, \qquad M_1 \le 2M_1' + 8\Delta^2.$$

A coordinate at which $\vartheta = 1$ remains equal to one after dyadic rounding, so the minimum-mass hypothesis gives $M_1' \ge \nu$. As $\eta \le \lambda \le 1/34$,

$$K' \le 16\lambda M_1' + (64\lambda + 16)\Delta^2 \le 17\lambda M_1'.$$

For the last inequality use $(64\lambda + 16)\Delta^2 \le 18\eta\nu/10000 \le \lambda M_1'$. Since $p_1' = M_1' - \mathcal{E}_P(\chi_1') \ge (1 - 17\lambda)M_1'$, we have $p_1' \ge M_1'/2$, giving $K' \le 34\lambda p_1'$. The nonzero vector $g^\sharp$ has proper support because $\varsigma \le R^{-1/4} \le 1/4$. Hence ϑ attains both zero and one. Rounding retains these values and cannot enlarge the small support; both child supports therefore remain nonempty and proper. Minimality of j gives $j \le 1 + \frac{1}{2}\log_2(10000/(\eta\nu))$. Thus the dyadic numerators have $O(\log(1/\eta) + \log(1/\nu))$ bits, and the displayed rational map has the same bound. Drop the primes. □

3.3 Construction, entropy accounting, and finite access

Proof of Theorem 3.1. Fix vertex order and every tie-breaking rule. Apply the first applicable rule at each current node:

1. At a singleton, stop.
2. If disconnected, split off the first connected component using its indicator and complementary indicator, and recurse on both children.
3. Otherwise compute the inclusive rank r, and stop if $r^8 \le R^5$.
4. If the rank test fails and a current marginal mass is below ν, isolate its first such vertex by an indicator split ("cleanup"), and recurse on both children.
5. Otherwise use Lemmas 3.2–3.3 and recurse on both children. Call this a smooth split.

Every tree edge strictly reduces the nonempty support, so the depth is at most $n-1$. Each leaf satisfies the rank bound: a singleton has rank one, and every other leaf passes the exact rank test. In particular, cleanup ensures the minimum-mass hypothesis whenever a smooth split is performed.

For any accumulated multiplier Φ, persistent root diagonal mass gives

$$d_\Phi(x) = \sum_y a_{xy}\Phi(x)\Phi(y) \ge \tfrac{1}{2}\mu(x)\Phi(x)^2.$$

Here Φ is the product along a node's path, extended by zero outside its support; the root product is 1. The upper bound $Z_\Phi \le M_\Phi$ follows by applying $2\Phi(x)\Phi(y) \le \Phi(x)^2 + \Phi(y)^2$ and summing. Thus $Z_\Phi > 0$, $M_\Phi/2 \le Z_\Phi \le M_\Phi$, and all retained degrees are positive. The normalized node marginal is d_Φ/Z_Φ, not $\mu\Phi^2/M_\Phi$. Induction on the finite tree gives the squared partition of unity. Index internal nodes by their binary prefixes t. If $Z_t = Z_{\Phi_t}$ is a node's root mass, define, in its normalized law,

$$p_{t,i} = \sum_{x,y} a^t_{xy}\chi_{t,i}(x)\chi_{t,i}(y), \qquad K_t = 1 - p_{t,0} - p_{t,1}.$$

Then $Z_{ti} = Z_t p_{t,i}$ and $Z_t - Z_{t0} - Z_{t1} = Z_t K_t$. Therefore

$$\sum_b M_b = 1, \qquad D = 1 - \sum_b Z_b = \sum_t Z_t K_t = \sum_b \mathcal{E}_P(\Phi_b). \tag{6}$$

Here is the information calculation with its exact posterior law. Draw one latent endpoint pair $E = (x, y)$ from the root. At a reached internal node t choose child i with conditional probability $\chi_{t,i}(x)\chi_{t,i}(y)$ and die with the remaining probability. These probabilities are nonnegative and sum to at most one by Cauchy–Schwarz, because $\sum_i \chi_{t,i}(x)^2 = \sum_i \chi_{t,i}(y)^2 = 1$. Let $\Xi = (\Xi_0, \Xi_1, \dots, \Xi_{n-2})$ be the transcript of outcomes, padded by a fixed symbol after death or arrival at a leaf. Three facts drive the calculation.

(a) *Exact posterior.* Conditional on $E = (x, y)$, the probability of reaching the node with binary prefix t is the product $\Phi_t(x)\Phi_t(y)$ of the local probabilities along its path. Hence the unconditional reach probability is $Z_t = \sum_{x,y} a_{xy}\Phi_t(x)\Phi_t(y)$, and by Bayes' rule the conditional law of E given that t is reached is $a_{xy}\Phi_t(x)\Phi_t(y)/Z_t$, which is exactly the node law a^t.

(b) *Padding carries no information.* Conditional on a prefix that has died or reached a leaf, the next outcome is constant.

(c) *Routing is a code for C.* Let $\mathcal{R}_0$ be the set of indicator nodes used before a path has been restricted to one root connected component; it is empty if the root is connected. On a realized path these routing decisions form an initial segment. Let Ξ^{rt} be their outcome sequence, padded by a fixed symbol after routing ends. Routing has no death because a positive-mass endpoint pair never straddles two components. Consequently Ξ^{rt} is a deterministic prefix code for C, and the map from C to its codeword is injective. After routing ends, every subsequent node lies inside one root component. Thus Ξ^{rt} is a function of Ξ, and Ξ^{rt} and C determine one another, even if the path later dies.

Let I_t be the mutual information, under the node law a^t, between the endpoint pair and the child-or-death outcome at t. The chain rule for mutual information, applied depth by depth and then decomposed over the prefixes of each depth, gives

$$I(E;\Xi) = \sum_{k\ge 0} I(E;\Xi_k \mid \Xi_{<k}) = \sum_{k\ge 0}\sum_{|t|=k} \Pr[\text{reach } t]\, I(E;\Xi_k \mid \text{reach } t) = \sum_t Z_t I_t, \tag{7}$$

where the last sum runs over internal nodes only: by (a) the conditional law given a reached prefix t is a^t, so the corresponding term equals $Z_t I_t$, and by (b) dead and terminal prefixes contribute zero. Applying the chain rule to the routing decision tree and using (c) gives $\sum_{t\in\mathcal{R}_0} Z_t I_t = I(E;\Xi^{\mathrm{rt}}) = H(C)$. Because Ξ^{rt} is a function of Ξ and generates the same sigma-algebra as C,

$$I(E;\Xi) - I(E;\Xi^{\mathrm{rt}}) = I(E;\Xi \mid \Xi^{\mathrm{rt}}) = I(E;\Xi \mid C).$$

Subtracting the routing contributions therefore yields

$$\sum_{t\notin\mathcal{R}_0} Z_t I_t = I(E;\Xi \mid C) \le H(E \mid C) = h. \tag{8}$$

The initial routing contribution $H(C)$ is not charged to h; later component cuts may consume conditional information but cost no mass. At a smooth node, outcome 1 forces both endpoints into $S_t := \operatorname{supp}\chi_{t,1}$. Put $s_t = \mu_t(S_t)$, and denote the support mass returned by the heat lemma at this node by ς_t. Then Lemma 3.3 gives $s_t \le \varsigma_t \le R^{-1/4}$. The prior probability $z_t = a^t(S_t \times S_t)$ is at most s_t. The relative-entropy support bound yields $I_t \ge p_{t,1}\log(1/s_t)$: indeed, for a distribution supported on an event of prior probability z, its divergence from the prior equals its divergence from the conditioned prior plus $\log(1/z)$, and here $z_t \le s_t$. The other outcome contributions are nonnegative. Lemma 3.3 gives $K_t \le 34\lambda_t p_{t,1}$ and Lemma 3.2 gives $\lambda_t \le (44J/T)\log(1/\varsigma_t) \le (44J/T)\log(1/s_t)$, so $K_t \le (1496J/T)\,p_{t,1}\log(1/s_t) \le (1496J/T)I_t$ at every smooth node. With $T \ge \log R/(128\eta)$ this gives

$$\sum_{t\ \text{smooth}} Z_t K_t \le \frac{1496J}{T}\sum_{t\ \text{smooth}} Z_t I_t \le 191488\frac{\eta J}{\log R}h, \qquad \mathbb{E}N_{\mathrm{sm}} = \sum_{t\ \text{smooth}} Z_t p_{t,1} \le \frac{4h}{\log R}.$$

The equality for $\mathbb{E}N_{\mathrm{sm}}$ counts a smooth child-1 transition even on a path that subsequently dies. Its upper bound uses $\log(1/s_t) \ge \frac{1}{4}\log R$ and $\sum_{t\ \text{smooth}} Z_t I_t \le h$. Cleanup at x costs $K_t = 2\sum_{y\ne x} a^t_{xy} < 2\nu$. Moreover, $\sum_{t\ \text{cleanup}} Z_t$ is the expected number of cleanup nodes reached by the sampler, which is at most the maximum depth $n-1$. Their total contribution is therefore less than $2n\nu$. This proves the entropy bound; $h \le H(E) \le \log(n^2)$ because there are at most n^2 ordered endpoint pairs. Finally, the vectors $(\Phi_b(x))_b$ are nonnegative and have unit Euclidean norm, so $k(x,y) = \sum_b \Phi_b(x)\Phi_b(y)$ lies in $[0,1]$ by Cauchy–Schwarz and has $k(x,x) = 1$. Hence $D = \sum_{x,y} a_{xy}(1-k(x,y))$ is at most $1 - \sum_x a_{xx} \le 1/2$.

For completeness, the construction is rational in the bit model. At a node with rational symmetric raw matrix A and positive degree matrix D, count the nonpositive eigenvalues of $(1-\eta)\mathsf{D}-\mathsf{A}$. Congruence by $\mathsf{D}^{-1/2}$ identifies this count with the inclusive normalized cutoff rank. Rational symmetric elimination uses a nonzero diagonal as a one-dimensional pivot; if all remaining diagonals vanish, a nonzero off-diagonal supplies an invertible two-dimensional pivot of negative determinant. An all-zero remainder contributes only zero eigenvalues. The signs of the one-dimensional pivots and the inertia $(1,1)$ of each two-dimensional pivot determine the count, including all zeros. To bound bit length, clear the input denominators once. Each Schur complement entry is a quotient of determinants of submatrices of the permuted original integer matrix, whose pivot block is nonsingular. Hadamard's determinant bound gives polynomial bit lengths for these integers. Reducing rational entries after each operation therefore keeps all pivot computations polynomial in the node's encoding length. Thus neither an exact irrational eigenbasis nor a numerical equality test is needed.

All accumulated multipliers have $O(nj)$ bits per coordinate because Lemma 3.3 replaces every heat vector by a dyadically rounded local multiplier; $j = O(\log(1/\eta) + \log(1/\nu))$ is fixed by the root parameters. Compute raw node weights directly from the root law and these products. Raw weights, their sums, and normalized degrees then have polynomial encoding length, regardless of how small Z_Φ is numerically. If b_Q is a common integer denominator for the current rational Q, each entry of Q^t has denominator dividing b_Q^t; numerators have $O(t \log b_Q)$ bits up to polynomial dimension terms. Since $T = O(\log R/\eta)$, the heat computation and its rational score comparisons have the promised cost. Only the rounded multipliers are propagated; the heat denominators are discarded. The support-depth bound gives the same polynomial cost for a path.

Computational sampling uses only the rational probabilities p_0, p_1, K; the latent pair is not sampled or needed. Express these probabilities with a common integer denominator B. For $B > 1$, draw $\lceil \log_2 B \rceil$ fair bits, accept the resulting integer if it is below B, and otherwise repeat; $B = 1$ is deterministic. The accepted integer is uniform on $\{0, \dots, B-1\}$, and the acceptance probability is greater than $1/2$. Assigning consecutive intervals of the appropriate integer lengths gives an exact draw in expected polynomial bit cost. There are at most $n-1$ draws on a path. For the confidence parameter s in the theorem, cap each draw at $r = s + \lceil \log_2 n \rceil$ attempts. This adds failure probability at most $n2^{-r} \le 2^{-s}$ and gives a polynomial bit budget. This is an additional failure event coupled to the exact sampler, not a renormalization of the surviving leaf law. To obtain (4), fix any vertex x in a leaf support and record the absolute depths at which its path takes smooth child 1. Depths are numbered $0, \dots, n-2$. At every indicator split, including cleanup, follow the unique child containing x; at a smooth split follow the recorded choice. Abort if x leaves the selected support. The pair consisting of x and the recorded depth subset reconstructs its leaf exactly. Enumerating vertices and depth subsets gives the asserted count and covers every required leaf. □

Definition 3.4 (good leaf)**.** Fix an invocation root carrying a Unique Games witness of error at most $\bar{\kappa} \ge 0$, and let D be the total loss of its decomposition tree. A leaf b is *good* if the witness has error at most $8\bar{\kappa}$ in the normalized leaf law a^b and $\mathcal{E}_P(\Phi_b)/Z_b \le 8D$. Both conditions are exactly the hypotheses of the leaf decoder, Corollary A.6.

Lemma 3.5 (A good leaf under the unconditional sampling law)**.** *If the root carries a Unique Games witness of error at most $\bar{\kappa} \ge 0$ and $D \le 1/4$, the exact sampler of Theorem 3.1 reaches a good leaf with probability at least $1/2$.*

Proof. Let v_b be the root-law mass of violated constraint records weighted by $\Phi_b(x)\Phi_b(y)$. The witness error in the normalized leaf is v_b/Z_b. Since $k(x,y) \le 1$ recordwise, $\sum_b v_b \le \bar{\kappa}$. Also $\sum_b Z_b(\mathcal{E}_P(\Phi_b)/Z_b) = D$ by (6). Under the unconditional leaf law each bad event has probability

at most 1/8 by Markov's inequality. The two bad-leaf events and death are the only excluded outcomes, so

$$\Pr[\text{reach a good leaf}] \geq 1 - D - \tfrac{1}{8} - \tfrac{1}{8} \geq \tfrac{1}{2}.$$

When either bound is zero its nonnegative summands vanish, so the same conclusion holds without division by that bound. □

The theorem also gives an exact deleted-edge interpretation. Put $a^{\mathrm{loss}}_{xy} = a_{xy}(1 - k(x,y))$ and $d^{\mathrm{loss}}(x) = \sum_y a^{\mathrm{loss}}_{xy}$. This is a nonnegative symmetric measure of total mass D. For any such measure m, use the unnormalized quadratic form

$$\mathcal{E}_m(f) := \frac{1}{2}\sum_{x,y} m_{xy}(f(x) - f(y))^2.$$

In particular $\mathcal{E}_{a^{\mathrm{loss}}}$ involves no division by D and is zero when $D = 0$. For every real-valued f on the root vertices, restricting f to a leaf support in terms involving P_b, we have

$$\begin{aligned}
\mathcal{E}_P(f) &= \sum_b Z_b \mathcal{E}_{P_b}(f) + \mathcal{E}_{a^{\mathrm{loss}}}(f), \\
\sum_b \mathcal{E}_P(\Phi_b f) &= \sum_b Z_b \mathcal{E}_{P_b}(f) + \sum_x d^{\mathrm{loss}}(x) f(x)^2 \leq \mathcal{E}_P(f) + D\,\|f\|^2_\infty .
\end{aligned}$$

The first identity follows from $a_{xy} = a_{xy}k(x,y) + a^{\mathrm{loss}}_{xy}$. For the second, expanding its left side gives $\sum_x \mu(x)f(x)^2 - \sum_{x,y} a_{xy}k(x,y)f(x)f(y)$; the sum of leaf energies has the same cross term and diagonal coefficient $\sum_y a_{xy}k(x,y) = \mu(x) - d^{\mathrm{loss}}(x)$. Finally $\mathcal{E}_{a^{\mathrm{loss}}}(f) \geq 0$ and $\sum_x d^{\mathrm{loss}}(x) = D$ give the inequality. Thus D measures actual omitted edge mass, not an auxiliary entropy potential alone.

4 Entropy-sensitive recovery of a complete labeling

4.1 The immutable information envelope

Throughout this section $n \geq 256$ and $q \geq 2$. As in Section 2, zero-total-weight inputs are returned immediately with arbitrary labels; the normalized-law definitions and profile-dependent guarantees below concern positive-total-weight inputs. Put $E^0 \sim a^0$, $H_{\max} = 2\log n$, $H_0 = H(E^0)$ and $h_0 = H(E^0 \mid C_0)$. For every positive-degree vertex define

$$r_x = \frac{1}{\mu^0(x)} \sum_{y : a^0_{xy} > 0} a^0_{xy} \log \frac{1}{a^0_{xy}}, \qquad \mathcal{T}(t) = \max_{\substack{0 \leq z_x \leq 1 \\ \sum_x \mu^0(x) z_x = t}} \sum_x \mu^0(x) r_x z_x \quad (0 \leq t \leq 1).$$

The fractional relaxation defines the profile at every mass, not only at masses attained by actual vertex subsets. For $0 < s, t \leq 1$, with $h_{\mathrm{bin}}(s) = -s\log s - (1-s)\log(1-s)$ and zero terms omitted, define

$$\begin{aligned}
\gamma(s) &= \min\left\{H_{\max}, \max\left\{0, \min\left\{\mathcal{T}(s)/s + \log s, (H_0 - h_{\mathrm{bin}}(s))/s\right\}\right\}\right\}, \\
G(t) &= \sup_{t \leq s \leq 1} \gamma(s), \qquad H_G = \frac{1}{4}\left(\int_0^1 \sqrt{G(t)/t}\, dt\right)^2, \\
\psi(\theta) &= \theta\big(1 + \tfrac{1}{2}\log(1/\theta)\big)^2 \quad (0 < \theta \leq 1), \qquad \psi(0) = 0, \\
H_* &= \min\{H_G, H_{\max}\psi(h_0/H_{\max})\}.
\end{aligned}$$

The quantities r_x include vertex surprisal, not merely transition-row entropy. The supremum defining G is needed because γ need not be decreasing. The integral defining H_G is improper at zero; it converges because $0 \le G(t) \le H_{\max}$ and $t^{-1/2}$ is integrable. No value of $\gamma(0)$ is needed. These functions describe the guarantee and are not oracles used by the algorithm.

Lemma 4.1 (Residual information bounds)**.** *The profile $\mathcal{T}$ is continuous, nondecreasing and concave, with endpoints $0, H_0$, and $\mathcal{T}(t)/t$ is nonincreasing on $(0,1]$. Every nonempty completed residual of mass t satisfies*

$$tH_U \le \mathcal{T}(t) + t\log t, \qquad tH_U \le H_0 - h_{\mathrm{bin}}(t), \qquad th_U \le h_0.$$

Consequently

$$h_U \le H_U \le G(t), \qquad H_0 \le G(t) \le \min\{H_{\max}, \mathcal{T}(t)/t\}, \qquad h_U \le \min\{H_{\max}, h_0/t\}.$$

In particular $H_0 \le H_G \le H_{\max}$ and $h_0 \le H_{\max}\psi(h_0/H_{\max}) \le H_{\max}$.

Proof. Ordering vertices by decreasing r_x fills the maximizing fractional set in that order: moving mass from a lower value to an unsaturated higher value cannot decrease its objective. Thus $\mathcal{T}$ is piecewise linear with nonnegative decreasing slopes. Concavity and $\mathcal{T}(0) = 0$ imply that $\mathcal{T}(t)/t$ is nonincreasing. Moreover, $\mathcal{T}(1) = \sum_x \mu^0(x) r_x = H(E^0) = H_0$, so this ratio is at least H_0 on $(0,1]$. This also proves the asserted endpoints and continuity.

Condition $E^0 = (X, Y)$ on $X \in U$ and map (X, Y) to (X, X) when $Y \notin U$. This deterministic map has output law E_U. Its input entropy is

$$H(E^0 \mid X \in U) = t^{-1} \sum_{x \in U} \mu^0(x) r_x + \log t.$$

Entropy decreases under a deterministic map, and the indicator of U is feasible in the fractional maximization, proving the first inequality. For $0 < t < 1$, the chain rule gives

$$H_0 = h_{\mathrm{bin}}(t) + tH(E^0 \mid X \in U) + (1-t)H(E^0 \mid X \notin U).$$

Discard the last nonnegative term to prove the second inequality; $t = 1$ follows directly from the deterministic map.

Let original components have masses m_j^0, and let $t_j = \mu^0(U \cap C_j)$. Applying the same argument within component j gives, when $t_j > 0$,

$$t_j H(E_U \mid C_0 = j) \le m_j^0 H(E^0 \mid C_0 = j).$$

Here the event has conditional probability t_j/m_j^0 in component j, so the same chain rule justifies the factor m_j^0. In the residual, component j has probability t_j/t. Summing the displayed inequality gives $tH(E_U \mid C_0) \le h_0$. Boundary completion creates no edges between original components, so C_U refines C_0 and further conditioning only decreases entropy. This proves $th_U \le h_0$.

At a mass attained by a residual, both expressions inside the inner minimum defining γ are upper bounds on $H_U \ge 0$; also $H_U \le H_{\max}$. Hence $H_U \le \gamma(t) \le G(t)$. At every mass, $\gamma(s) \le \min\{H_{\max}, \mathcal{T}(s)/s\}$, so taking the supremum over $s \ge t$ proves the asserted upper bound for G. Nested supremum domains make G nonincreasing, while $\gamma(1) = H_0$ gives its lower bound. Integrating $\sqrt{H_0/t}$ and $\sqrt{H_{\max}/t}$ proves the bounds on H_G. Finally $\theta \le \psi(\theta) \le 1$: $\psi(1) = 1$, its limit at zero is zero, and $\psi'(\theta) = \frac{1}{2}\log(1/\theta) + \frac{1}{4}(\log(1/\theta))^2 \ge 0$. □

4.2 The end-to-end theorem

Theorem 4.2 (Entropy-sensitive complete recovery)**.** *Let $n \geq 256$, $q \geq 2$, and $256 \leq R \leq n$ be integers, and let $0 < \kappa < 1$ be rational. Given an instance with $\mathrm{OPT} \geq 1 - \kappa$, put*

$$L = \log(100n/\kappa), \quad L_R = L/\log R, \quad \alpha_H = \frac{\log(2L_R)}{2\log R}, \quad \mathcal{B}_* = L_R + \alpha_H H_*.$$

There is an absolute C_{rec} and a component-first algorithm producing, on success, a complete labeling of error at most

$$\min\{1, C_{\mathrm{rec}}\sqrt{\kappa \mathcal{B}_*}\}.$$

For rational $0 < \xi < 1$, it has a bounded randomized implementation with success probability at least $1 - \xi$ and time

$$q^{O(R^{5/8})}\,\mathrm{poly}(n, q, \textit{input length}, 1/\kappa, \log(1/\xi)).$$

It also has a deterministic implementation of time

$$\exp\left(O\left(R^{5/8}\log q + \frac{\mathcal{H}_* \log n}{\log R}\right)\right)\mathrm{poly}(n, q, \textit{input length}, 1/\kappa),$$

where

$$\mathcal{H}_* = \begin{cases} \min\{G(\sqrt{\kappa L_R}), h_0/\sqrt{\kappa L_R}\}, & \kappa L_R < 1, \\ 0, & \kappa L_R \geq 1. \end{cases}$$

One may take $C_{\mathrm{rec}} = 2^{23}$. No decision requires evaluating G, H_G or an improper integral.

We first give the finite decision quantities used in its proof. For a rational probability vector p put

$$\overline{H}(p) = \sum_{p_z > 0} p_z \left(\frac{7}{10}k_z + \frac{7}{5}(1 - 2^{k_z} p_z)\right), \qquad k_z = \lfloor \log_2(1/p_z) \rfloor.$$

For the current residual, let $m_j = \Pr[C_U = j]$ and let $a^{(j)}$ be the conditional law of E_U given $C_U = j$. Set

$$\overline{h}_U = \min\{\overline{H}(E_U), \sum_j m_j \overline{H}(a^{(j)})\}.$$

Define the global rational numbers

$$L_2 = \lceil \log_2(100n/\kappa) \rceil, \quad s_2 = \lfloor \log_2 R \rfloor, \qquad \overline{L_R} = L_2/s_2,$$

$$J_2 = \lceil \log_2(2\overline{L_R}) \rceil, \qquad \overline{J} = 7J_2/10, \qquad \overline{d} = 100/(69 s_2),$$

and the stage budgets

$$\mathcal{B}_U = L_R + \alpha_H h_U, \qquad \overline{\mathcal{B}}_U = \overline{L_R} + \tfrac{1}{2}\overline{d}\,\overline{J}\,\overline{h}_U.$$

Lemma 4.3 (Exact rational budget)**.** *The quantities with overbars have polynomial bit complexity and*

$$h_U \leq \overline{h}_U \leq \tfrac{7}{5} h_U, \qquad L_R \leq \mathcal{B}_U \leq \overline{\mathcal{B}}_U < 4\mathcal{B}_U, \qquad L_R \leq \overline{L_R} \leq \tfrac{81}{64} L_R.$$

The entropy estimates preserve exact zero.

Proof. For $x \in [1/2, 1]$, the tangent and chord bounds give $1 - x \le -\log x \le 2\log 2(1-x)$. Since $69/100 < \log 2 < 7/10$, put $x_z = 2^{k_z} p_z \in (1/2, 1]$ and write $-\log p_z = k_z \log 2 - \log x_z$. Applying the preceding bounds to each summand yields $H(p) \le \overline{H}(p) \le (7/5)H(p)$. Both terms defining $\overline{h}_U$ majorize h_U, and its second term is at most $(7/5)h_U$. Binary comparisons compute every integer exponent; their magnitudes and all rational encodings are polynomially bounded. Moreover, the displayed approximation is zero exactly when the original entropy is zero, which proves the final assertion.

Put $\tilde{L} = \log_2(100n/\kappa)$ and $\tilde{s} = \log_2 R$. Then $\tilde{L} \ge \tilde{s} \ge 8$, $\tilde{L} \le L_2 < \tilde{L} + 1$, and $s_2 \le \tilde{s} < s_2 + 1$. Thus $\overline{L_R}/L_R = (L_2/\tilde{L})(\tilde{s}/s_2)$ lies in $[1, (9/8)^2]$. These bounds and $69/100 < \log 2 < 7/10$ also give

$$\frac{1}{\log R} \le \overline{d} < \frac{105}{92 \log R}, \qquad \log(2L_R) \le \overline{J} < \tfrac{5}{2}\log(2L_R).$$

Indeed $\overline{d} \log R < 100(9/8)(7/10)/69 = 105/92$. For the last upper bound use $\log(81/64) < 1/4$ and $J_2 < \log_2(2\overline{L_R}) + 1$ to obtain $\overline{J} < (70/69)\log(2L_R) + 1$, and then $\log(2L_R) > 69/100$. The entropy coefficient inflates by less than $(105/92)(7/5)(5/2) = 735/184 < 4$; the L_R coefficient inflates by at most $81/64$. This proves the budget comparison. □

Proof of Theorem 4.2. Table 3 collects the quantities that drive the algorithm's decisions. All of them are exact rationals computed from the current residual, and none requires G, H_G, or an improper integral.

quantity	definition	role
t	$\mu^0(U)$	mass of the current residual U
t_{tail}	least power of two $\ge \sqrt{\kappa \overline{L_R}}$, capped at 1	tail stop when $t \le t_{\text{tail}}$
$\bar{\kappa},\ \eta,\ \nu$	$\kappa/t,\ 1024\bar{\kappa},\ \kappa/(100n)$	parameters of the invocation tree
$\overline{h}_U,\ \overline{\mathcal{B}}_U$	Lemma 4.3	rational stand-ins for h_U and $\mathcal{B}_U = L_R + \alpha_H h_U$
budget stop	$(\kappa/t)\overline{\mathcal{B}}_U > 2^{-34}$	stop and label the residual arbitrarily
acceptance	$\Gamma^2 \le 2^{38}\bar{\kappa}\,\overline{\mathcal{B}}_U$	test applied to every decoded candidate
k_U	$\lceil 16\overline{d}\,\overline{h}_U \rceil$	cap on N_{sm} in the deterministic search

Table 3: Decision quantities in the proof of Theorem 4.2.

Decisions and existence of accepted candidates. Let t_{tail} be the lesser of one and the least power of two that is at least $\sqrt{\kappa \overline{L_R}}$. Thus, analytically, $t_{\text{tail}} = \min\{1, 2^{\lceil \log_2 \sqrt{\kappa \overline{L_R}} \rceil}\}$. Comparisons of rational squares implement this rule without computing the square root or logarithm. Power-of-two rounding loses less than a factor two, and $\sqrt{81/64} = 9/8$, giving

$$t_{\text{tail}} < 3\sqrt{\kappa L_R}, \qquad t_{\text{tail}} \ge \sqrt{\kappa L_R} \quad \text{if } t_{\text{tail}} < 1.$$

At residual mass t, stop and label the residual arbitrarily if $t \le t_{\text{tail}}$, or if $(\kappa/t)\overline{\mathcal{B}}_U > 2^{-34}$. Otherwise invoke the component-first tree with

$$\bar{\kappa} = \kappa/t, \qquad \eta = 1024\bar{\kappa}, \qquad \nu = \kappa/(100n).$$

The restricted witness has error at most $\bar{\kappa}$; this is a statement about the complete mixture, not each component separately. Admission and $\overline{\mathcal{B}}_U \ge L_R \ge 1$, $t \le 1$ give $\kappa \le \bar{\kappa} \le 2^{-34}$ and $\eta \mathcal{B}_U \le 2^{-24}$, in particular $\eta \le 2^{-24} < 1/4$. Hence $3072\eta L_R \le 3 \cdot 2^{-14} < 1/34$. Theorem 3.1 therefore applies and gives

$$D \le 3 \cdot 2^{17} \eta \mathcal{B}_U + \kappa/50 \le 3/128 + 2^{-34}/50 < 1/32.$$

Here $191488\eta \log(2L_R) h_U / \log R = 382976\eta\alpha_H h_U \le 3 \cdot 2^{17}\eta\mathcal{B}_U$, and $2n\nu = \kappa/50$. By the good-leaf and finite-decoder interfaces (Lemma 3.5 and Corollary A.6), a sampled path reaches a good leaf with unconditional probability at least $1/2$. On every such leaf the bounded decoder, supplied with the invocation-root records, the leaf product Φ_b, and $\kappa_0 = 8\bar{\kappa}$, enumerates a nonempty partial labeling $\mathcal{S}$ whose actual invocation-root charge obeys

$$\Gamma = \frac{B + 2C_{\mathrm{out}}}{\mu_U(\mathcal{S})}, \qquad \Gamma^2 \le \frac{65}{432}\eta + \frac{2048}{27}D < 2^{38}\bar{\kappa}\mathcal{B}_U.$$

Here B is violated internal directed mass and C_{out} is outgoing directed mass, both in the normalized invocation root. For the strict final inequality, substitute $\eta = 1024\bar{\kappa}$ and the bound on D to obtain

$$\Gamma^2 \le \frac{65 \cdot 1024}{432}\bar{\kappa} + \frac{2^{38}}{9}\bar{\kappa}\mathcal{B}_U + \frac{2048}{1350}\kappa < 2^{38}\bar{\kappa}\mathcal{B}_U,$$

where $\kappa \le \bar{\kappa}$ and $\mathcal{B}_U \ge 1$. The decoder hypothesis $128\kappa_0 \le \eta$ holds with equality; the good-leaf event supplies its witness-error and loss bounds. Test the exactly rational condition $\Gamma^2 \le 2^{38}\bar{\kappa}\overline{\mathcal{B}}_U$ on every enumerated candidate; accept one that passes, install it and remove its vertices. Decoder failure at an inspected leaf simply continues the search. At least one candidate passes on every good leaf. There are at most n successful invocations, and every accepted score satisfies $\Gamma \le 2^{20}\sqrt{\kappa\overline{\mathcal{B}}_U/t}$, because $\overline{\mathcal{B}}_U < 4\mathcal{B}_U$.

One proof for both entropy envelopes. Apply the following argument separately to

$$\mathcal{W}(t) = G(t), \quad H_{\mathcal{W}} = H_G, \qquad \text{and} \qquad \mathcal{W}(t) = \min\{H_{\max}, h_0/t\}, \quad H_{\mathcal{W}} = H_{\max}\psi(h_0/H_{\max}).$$

In either case $\mathcal{W}$ is nonnegative and nonincreasing, $h_U \le \mathcal{W}(t)$, and

$$\int_0^1 \sqrt{\mathcal{W}(t)/t}\, dt = 2\sqrt{H_{\mathcal{W}}}.$$

For the second case this follows by splitting at $h_0/H_{\max}$: the integral is $\sqrt{h_0}(2 + \log(H_{\max}/h_0))$, with value zero when $h_0 = 0$. Let $\mathcal{B}_{\mathcal{W}}(t) = L_R + \alpha_H\mathcal{W}(t)$. If accepted stage i removes immutable mass $\Delta_i = t_i - t_{i+1}$, then $\mathcal{B}_{U_i} \le \mathcal{B}_{\mathcal{W}}(t_i)$ and $\sqrt{\mathcal{B}_{\mathcal{W}}(t)/t}$ is nonincreasing in t. Consequently $\Delta_i\sqrt{\mathcal{B}_{\mathcal{W}}(t_i)/t_i} \le \int_{t_{i+1}}^{t_i} \sqrt{\mathcal{B}_{\mathcal{W}}(t)/t}\, dt$. Summing disjoint stage intervals and using $\sqrt{L_R + \alpha_H\mathcal{W}} \le \sqrt{L_R} + \sqrt{\alpha_H\mathcal{W}}$ gives

$$\sum_i \Delta_i\Gamma_i \le 2^{20}\sqrt{\kappa}\int_0^1 \sqrt{\mathcal{B}_{\mathcal{W}}(t)/t}\, dt \le 2^{21}\sqrt{2\kappa(L_R + \alpha_H H_{\mathcal{W}})}.$$

The last inequality also uses $\sqrt{L_R} + \sqrt{\alpha_H H_{\mathcal{W}}} \le \sqrt{2(L_R + \alpha_H H_{\mathcal{W}})}$. This holds for every accepted trajectory, not merely in expectation.

The terminal mass t_f must also be charged; if all vertices were removed, put $t_f = 0$. A tail stop gives $t_f < 3\sqrt{\kappa L_R}$. At a budget stop, $t_f < 2^{36}\kappa\mathcal{B}_{U_f}$: indeed, the stopping inequality and $\overline{\mathcal{B}}_{U_f} < 4\mathcal{B}_{U_f}$ give $2^{-34} < \kappa\overline{\mathcal{B}}_{U_f}/t_f < 4\kappa\mathcal{B}_{U_f}/t_f$. Therefore

$$t_f^2 < 2^{36}\kappa(L_R t_f + \alpha_H t_f h_{U_f}) \le 2^{36}\kappa(L_R + \alpha_H H_{\mathcal{W}}).$$

For the component envelope use $t_f h_{U_f} \le h_0 \le H_{\mathcal{W}}$; for G use $t_f h_{U_f} \le t_f H_{U_f} \le H_0 \le H_G$. For the charging identity, if stage i removes $\mathcal{S}_i$, then $\mu_{U_i}(\mathcal{S}_i) = \Delta_i/t_i$ and $\Delta_i\Gamma_i = t_i(B_i + 2C_{\mathrm{out},i})$. This pays original half-lazified internal violations at that stage and both orientations of every edge to a later stage. Completion loops are identities and contribute no violation; edges to earlier stages have

already been paid. Only terminal internal records remain unpaid, with total mass at most t_f. Thus, undoing half-lazification,

$$\mathrm{err} * \mathrm{original} \le \min\{1, 2[\sum_i \Delta_i \Gamma_i + t_f]\} \le \min\{1, 2^{23}\sqrt{\kappa(L_R + \alpha_H H * \mathcal{W})}\}.$$

The terminal bounds give $t_f \le 2^{18}\sqrt{\kappa(L_R + \alpha_H H_{\mathcal{W}})}$, and $2(2^{21}\sqrt{2} + 2^{18}) < 2^{23}$ verifies the displayed constant. This includes zero-entropy and trivial-bound cases. Both envelopes analyze the same decisions and success event, so their minimum gives H_*.

Bounded random search. At each invocation use $N_{\mathrm{path}} = \lceil \log_2(2n/\xi) \rceil$ fresh killed paths (Lemma 3.5). Conditional on any earlier successful history, failure to find a good leaf has probability at most $2^{-N_{\mathrm{path}}}$. A union bound over at most n invocations costs at most $\xi/2$. There are at most $n^2 N_{\mathrm{path}}$ rational branch draws. To sample a rational categorical law, draw a uniform integer with the required denominator by binary rejection, whose acceptance probability is at least $1/2$. Cap each draw at $\lceil \log_2(2n^2 N_{\mathrm{path}}/\xi) \rceil$ trials. Coupling with uncapped draws bounds the additional failure probability by $\xi/2$. A cap failure, or exhaustion of the path batch without an accepted candidate, is reported as failure and the remaining labels are arbitrary. No exact conditional leaf law is claimed after capping.

Deterministic search. At an invoked root let $k_U = \lceil 16 \bar{d}\, \bar{h}_U \rceil$. The sampler satisfies $\mathbb{E}N_{\mathrm{sm}} \le 4h_U/\log R$, where N_{sm} counts smooth small-child choices. Markov's inequality gives $\Pr(N_{\mathrm{sm}} > k_U) \le 1/4$; when $h_U = 0$, both k_U and N_{sm} vanish. Since a good leaf has probability at least $1/2$, the probability of reaching a good leaf with at most k_U counted choices is at least $1/2 - 1/4 > 0$. Thus such a leaf exists.

The support-vertex encoding in (4) covers every such leaf with $\exp(O((k_U + 1)\log n))$ path simulations; it does not enumerate hard component and cleanup decisions independently. If no candidate passes after this finite enumeration, report failure and label the remaining vertices arbitrarily. This exit cannot occur under the promised witness; specifying it also bounds runs on false guesses. The rational estimates give $k_U < 26h_U/\log R + 2$. Every invoked root has $t > t_{\mathrm{tail}} \ge \sqrt{\kappa L_R}$, so Lemma 4.1 gives $h_U \le \mathcal{H}_*$. If $\kappa L_R \ge 1$, the tail test invokes no root and the zero cap is valid.

Finally every visited tree node has polynomial bit cost in n, the input length and $1/\kappa$, by finite access in Theorem 3.1; its dyadic multiplier and bit-length bounds apply equally after zero–one component restrictions. The decoder costs $q^{O(R^{5/8})}$ times that polynomial. Rational residual rebuilding, entropy estimates and charge comparisons are polynomial. Multiplying by the stated path counts and at most n invocations proves the two running times. □

4.3 The general accuracy–time tradeoff

The entropy-sensitive implementation retains numerical $1/\kappa$ work and a deterministic path-search term. The next composition is a separate worst-case guarantee; it does not remove those costs from Theorem 4.2 with its sharper instance-specific envelope.

Theorem 4.4 (Completeness-uniform worst-case guarantee)**.** *Let $n \ge 256$, $q \ge 2$, and $256 \le R \le n$ be integers, and let $0 < \kappa < 1$ be rational. For an instance with* $\mathrm{OPT} \ge 1 - \kappa$*, put*

$$\Lambda_R = \frac{\log n}{\log R}, \qquad \mathcal{B}_n = \Lambda_R(1 + \log(2\Lambda_R)).$$

A deterministic algorithm produces a labeling with

$$\mathrm{err}(\ell) \le \min\{1, C_{\mathrm{GT}}\kappa \log(2n), C_{\mathrm{unif}}\sqrt{\kappa \mathcal{B}_n}\}$$

in time $q^{O(R^{5/8})}\,\mathrm{poly}(n, q, \textit{input length})$*, where* C_{GT} *and* C_{unif} *are absolute. The polynomial has no numerical* $1/\kappa$ *factor.*

Proof. Always run the deterministic Gupta–Talwar fallback (Lemma B.2), whose error is at most $C_{\mathrm{GT}}\kappa\log(2n)$ and whose time is polynomial in the binary input length. Set $l = \lceil \log_2 n \rceil$. Also run the deterministic algorithm of Theorem 4.2 precisely when $R > l^4$ and $n\kappa \geq 1$. Return the labeling with least actual rational error. All branch tests are rational or integer comparisons.

When that invocation is made, $1/\kappa \leq n$ and $n \geq 256 > 100$, so $L \leq \log(100n^2) \leq 3\log n$, whence $L_R \leq 3\Lambda_R$. Since $H_* \leq 2\log n$,

$$\mathcal{B}_* \leq L_R + \Lambda_R \log(2L_R) \leq 3\Lambda_R + \Lambda_R \log(6\Lambda_R) \leq 3\Lambda_R(1 + \log(2\Lambda_R)) = 3\mathcal{B}_n.$$

The last inequality uses $\log 3 \leq 2\log(2\Lambda_R)$ for $\Lambda_R \geq 1$. Thus the entropy-sensitive candidate has error at most $\sqrt{3}C_{\mathrm{rec}}\sqrt{\kappa\mathcal{B}_n}$. Also $R^{5/8} > l^{5/2} \geq (\log n)^2$, which absorbs the additional deterministic path exponent, since $\mathcal{H}_* \leq 2\log n$ and $\log q \geq \log 2$. The numerical cost is polynomial in n because $1/\kappa \leq n$ on this branch.

If $\kappa < 1/n$, the fallback alone suffices:

$$\kappa\log(2n) \leq \sqrt{\kappa}\,\frac{\log(2n)}{\sqrt{n}} \leq \sqrt{\kappa} \leq \sqrt{\kappa\mathcal{B}_n}.$$

Here $\log(2n) \leq \sqrt{n}$ for $n \geq 256$, as follows at 256 and thereafter by differentiating $\sqrt{n} - \log(2n)$.

It remains to handle $R \leq l^4$. Write $\mathsf{x} = \log n$, $\mathsf{y} = \log R$. Since $l \leq \mathsf{x}/\log 2 + 1 \leq 2\mathsf{x}$ for $n \geq 256$, we have $\mathsf{y} \leq 4\log(2\mathsf{x})$ and $\log\mathsf{y} \leq 3\mathsf{y}/16 + 1$: the maximum of $\log\mathsf{y} - 3\mathsf{y}/16$ is $\log(16/3) - 1 < 1$. Hence

$$\mathcal{B}_n = \frac{\mathsf{x}}{\mathsf{y}}(1 + \log(2\mathsf{x}) - \log\mathsf{y}) \geq \frac{\mathsf{x}}{\mathsf{y}}(\log(2\mathsf{x}) - 3\mathsf{y}/16) \geq \mathsf{x}/16.$$

The last step uses $\mathsf{y} \leq 4\log(2\mathsf{x})$. Every labeling has error at most one, and $\min\{1, z\} \leq \sqrt{z}$ for $z \geq 0$. The fallback therefore has error at most

$$\sqrt{C_{\mathrm{GT}}\kappa\log(2n)} \leq \sqrt{32C_{\mathrm{GT}}\kappa\mathcal{B}_n}.$$

Here $\log(2n) \leq 2\mathsf{x} \leq 32\mathcal{B}_n$. Taking $C_{\mathrm{unif}} = \max\{\sqrt{3}C_{\mathrm{rec}}, C_{\mathrm{GT}}, \sqrt{32C_{\mathrm{GT}}}\}$ handles all branches. Its polynomial fallback cost and the preceding large-R calculation prove the claimed uniform time. □

Corollary 4.5 (Near-perfect accuracy–time tradeoff)**.** *For integers* $n \geq 256$*,* $q \geq 2$ *and rational parameters* $0 < \kappa < 1$*,* $0 < \alpha < 5/8$*, an instance with* $\mathrm{OPT} \geq 1 - \kappa$ *admits a deterministically computed labeling with*

$$\mathrm{err}(\ell) \leq \min\left\{1, C_{\mathrm{GT}}\kappa\log(2n), C_{\mathrm{tr}}\sqrt{\frac{\kappa}{\alpha}\log\frac{2}{\alpha}}\right\}$$

in time $q^{O(n^\alpha)}\,\mathrm{poly}(n, q, \textit{input length})$*. All constants are absolute, also for parameters varying with* n*.*

Proof. Put $\beta = 8\alpha/5$, compute $l = \lceil \log_2 n \rceil$ and $k = \lceil \beta l \rceil$, and choose $R = \min\{n, \max\{256, 2^k\}\}$. Since $0 < \beta < 1$ and $\log_2 n \leq l < \log_2 n + 1$, we have $n^\beta \leq 2^k < 2^{1+\beta}n^\beta < 4n^\beta$. The cap at n preserves the lower bound, and the floor at 256 gives $n^\beta \leq R \leq 256n^\beta$. Thus $1 \leq \Lambda_R \leq 1/\beta$. Consequently

$$\mathcal{B}_n \leq \frac{1 + \log(2/\beta)}{\beta} \leq \frac{3\log(2/\beta)}{\beta} \leq \frac{3\log(2/\alpha)}{\alpha}.$$

The middle inequality uses $\log(2/\beta) > \log 2 > 1/2$; the last uses that $x \mapsto \log(2/x)/x$ is decreasing on $(0, 1]$ and $\beta = 8\alpha/5 > \alpha$. Theorem 4.4 applies, and $R^{5/8} \le 256^{5/8} n^{\alpha}$. All parameter choices are finite integer or rational operations, with their encodings included in the input length. The first two error terms are unchanged, while the third is at most $\sqrt{3} C_{\mathrm{unif}} \sqrt{(\kappa/\alpha)\log(2/\alpha)}$. Thus one may take $C_{\mathrm{tr}} = \sqrt{3} C_{\mathrm{unif}}$. □

Corollary 4.6 (Recovery without a supplied completeness promise)**.** *Given only a rational weighted Unique Games instance with $n \ge 256$, $q \ge 2$, and an integer $256 \le R \le n$, a deterministic algorithm returns a labeling satisfying*

$$\mathrm{err}(\ell) \le \min\{1, C_{\mathrm{GT}} \varepsilon_* \log(2n), C_{\mathrm{pf}} \sqrt{\varepsilon_* \mathcal{B}_n}\}, \qquad \varepsilon_* = 1 - \mathrm{OPT},$$

where $\mathcal{B}_n = (\log n/\log R)(1 + \log(2\log n/\log R))$, $C_{\mathrm{GT}} = 600$, and C_{pf} is absolute. Its time is $q^{O(R^{5/8})}\,\mathrm{poly}(n, q, \textit{input length})$, with at most $O(1 + \log\log n)$ bounded runs of the algorithm of Theorem 4.4. Neither an upper bound on ε_* nor its value is an input.

Proof. Appendix B.1 first tests satisfiability exactly and otherwise obtains, from the Gupta–Talwar fallback, a rational error ϵ_0 satisfying $\varepsilon_* \le \epsilon_0 \le C_{\mathrm{GT}} \varepsilon_* \log(2n)$. It runs Theorem 4.4 on the retained guesses $\epsilon_0/2^j$ for $0 \le j \le \lceil \log_2(C_{\mathrm{GT}} \lceil \log_2(2n) \rceil) \rceil$. When $\varepsilon_* \ge 1/n$, one retained guess lies in $[\varepsilon_*, 2\varepsilon_*)$; when $\varepsilon_* < 1/n$, the fallback already gives the stated square-root bound. The appendix also specifies a finite exhaustion exit for every false guess, proving both the claimed number of bounded runs and the uniform running time. □

5 Actual-channel information and its limitations

The decomposition proof charges killing to a lower bound on the information of a split. We compute the full information and show that even this quantity cannot give a uniform bound $K/I = O(\eta/\log R)$ over the entire high-trace class. The same example passes the algorithm's low-rank stopping test: this is a limitation of an unrestricted local argument, not a lower bound on the algorithm's performance.

For a probability law a and $f \ge 0$, write $\mathrm{Ent}_a(f) = \mathbb{E}_a[f \log f] - (\mathbb{E}_a f)\log(\mathbb{E}_a f)$, with $0 \log 0 = 0$. Let $a_{xy} = \mu(x) P(x, y)$ be a normalized symmetric endpoint law on a finite vertex set with positive marginal μ. For a finite nonnegative squared partition of unity $\sum_{i=1}^{k} \chi_i(x)^2 = 1$, define

$$w_i(x, y) = \chi_i(x)\chi_i(y), \qquad w_\dagger(x, y) = 1 - \sum_i w_i(x, y), \qquad p_z = \mathbb{E}_a w_z.$$

Cauchy–Schwarz gives $0 \le \sum_i w_i(x, y) \le 1$, so these functions define a probability channel from $E = (x, y)$ to the child or killed outcome $\mathcal{O} \in \{1, \dots, k, \dagger\}$. Set $K = p_\dagger$ and $I = I(E; \mathcal{O})$. In the binary case we index the children by $0, 1$. At a node of the construction, use the actual rounded multipliers passed to the children; the identities below do not require rationality. We write $D_{\mathrm{KL}}(p\|q) = \sum_{e: p(e) > 0} p(e) \log(p(e)/q(e))$ when q is positive on the support of p.

Lemma 5.1 (The exact split channel)**.** *For every positive-probability outcome put $\mathcal{A}_z = \{(x, y) : a_{xy} w_z(x, y) > 0\}$, $\mathrm{pr}_z = a(\mathcal{A}_z)$, and $a^{(z)}(x, y) = a_{xy} w_z(x, y)/p_z$. Then*

$$K = \sum_i \mathcal{E}_P(\chi_i), \tag{9}$$

$$I = \sum_z \mathrm{Ent} * a(w_z) = \sum * z : p_z > 0 p_z \left[\log(1/\mathrm{pr} * z) + D * \mathrm{KL}(a^{(z)} \| a(\cdot \mid \mathcal{A}_z))\right]. \tag{10}$$

For the stereographic split $\chi_0 = (1-\vartheta^2)/(1+\vartheta^2)$, $\chi_1 = 2\vartheta/(1+\vartheta^2)$, $0 \le \vartheta \le 1$, *writing* $N_{xy} = (1+\vartheta_x^2)(1+\vartheta_y^2)$ *gives*

$$(w_0, w_1, w_\dagger) = \frac{((1-\vartheta_x^2)(1-\vartheta_y^2),\ 4\vartheta_x\vartheta_y,\ 2(\vartheta_x-\vartheta_y)^2)}{N_{xy}}, \qquad \mathcal{E}_P(\vartheta) \le K \le 4\mathcal{E}_P(\vartheta).$$

Proof. The two marginals of a are μ, and therefore $\mathcal{E}_P(\chi_i) = \mathbb{E}_\mu \chi_i^2 - \mathbb{E}_a w_i$. Summing over i, and using $\sum_i \chi_i^2 = 1$, proves the first identity. The conditional endpoint law given $\mathcal{O} = z$ is $a^{(z)}$, so mutual information equals $\sum_{z:p_z>0} p_z D_{\mathrm{KL}}(a^{(z)}\|a) = \sum_z \mathrm{Ent}_a(w_z)$. On $\mathcal{A}_z$, the law $a(\cdot \mid \mathcal{A}_z)$ is a/pr_z. Hence

$$D_{\mathrm{KL}}(a^{(z)}\|a) = \log(1/\mathrm{pr} * z) + D * \mathrm{KL}(a^{(z)}\|a(\cdot \mid \mathcal{A}_z)),$$

which proves the second identity. Direct substitution of the stereographic multipliers gives the displayed probabilities. In particular,

$$K = \mathbb{E}_a \frac{2(\vartheta_x - \vartheta_y)^2}{N_{xy}}.$$

The bounds $1 \le N_{xy} \le 4$, together with $\mathcal{E}_P(\vartheta) = \frac{1}{2}\mathbb{E}_a(\vartheta_x - \vartheta_y)^2$, now give $\mathcal{E}_P(\vartheta) \le K \le 4\mathcal{E}_P(\vartheta)$. □

For each positive-mass outcome, its relative-entropy remainder vanishes if and only if $w_z = p_z/\mathrm{pr}_z$ a-almost surely on $\mathcal{A}_z$. Thus all remainders vanish precisely when each such likelihood is constant on its support. This holds for indicator multipliers of a hard vertex partition: the child likelihoods indicate within-part pairs and the killed likelihood indicates between-part pairs. The remainders are therefore not uniformly positive.

Killing also carries information. Let a be a half-lazy root law and let Φ be a nonzero soft restriction. When the split channel is applied to the descendant endpoint law a^Φ, the notation of Section 2 gives

$$a^\Phi\{x = y\} = \frac{\sum_x a_{xx}\Phi(x)^2}{Z_\Phi} \ge \frac{M_\Phi}{2Z_\Phi} \ge \frac{1}{2}.$$

Since $w_\dagger(x,x) = 0$, the support of the killed outcome contains no diagonal pair, and therefore $\mathrm{pr}_\dagger \le 1/2$ whenever $K > 0$. Hence (10), now with base law a^Φ, gives $I \ge p_1 \log(1/\mathrm{pr}_1) + K \log 2$ for a binary split, with the first term omitted if $p_1 = 0$ and the second interpreted as zero if $K = 0$. This is an aggregate diagonal-mass assertion, not pointwise half-laziness of the descendant walk.

Proposition 5.2 (A full-channel obstruction, and its rank screen)**.** *Let* $d \ge 2$ *be an integer, let* Q *independently flip every coordinate of* $\{0,1\}^d$ *with probability* $1/(4d)$*, and put* $\varrho = 1 - 1/(2d)$ *and* $P = 2Q - I$*. This is a rational half-lazy walk with uniform stationary law. For a real number* $U \ge 2$*, let*

$$T_U = \min\{t \in \mathbb{Z}_{\ge 0} : (1+\varrho^{2t})^d \le 2^{d/U}\}.$$

Every nonnegative binary squared partition of unity with both child supports nonempty and proper satisfies

$$0 < K, \qquad 0 < I \le 6dK, \qquad \frac{K}{I} \ge \frac{1}{6d} \ge \frac{\log U}{12T_U}. \tag{11}$$

For integer $U \ge 2$ *and* $d \ge 16U$*, put*

$$n = 2^d, \quad \nu = 1/n, \quad s = \lfloor d/U \rfloor, \quad R = 2^s, \quad T = T_U, \quad \eta = s/(32T).$$

Then

$$\mathrm{Tr}(Q^{2T}) > \sqrt{R}, \qquad \frac{\log R}{\eta}\frac{K}{I} \ge \frac{8}{3}(\log 2)\log U, \qquad \mathrm{rank}\,\mathbf{1}_{[0,\eta]}(I-P) < R^{5/8}. \tag{12}$$

These are rational algorithm parameters, with $256 \le R \le n$, $0 < \eta < 1/16$, *and* $\lfloor \lfloor \log_2 R \rfloor / (32\eta) \rfloor = T_U$, *the heat time prescribed in Lemma 3.2. Moreover, with* $L_R^{\mathrm{h}} = \log(1/\nu)/\log R$, $U \le L_R^{\mathrm{h}} \le 16U/15$ *and* $3072\eta L_R^{\mathrm{h}} \le 192/\log U$.

Proof. We give the entropy estimate, full-channel bound, and spectral computation separately. The cube entropy–energy inequality is classical [20, Exercises 10.23 and 10.26]; we include its proof in our normalization rather than use it as a black box.

Entropy and energy. For the uniform two-point law and $\varphi = (a, b) \ge 0$, $\mathrm{Ent}(\varphi^2) \le (a-b)^2/2$. The case $a = b = 0$ is immediate. Otherwise homogeneity permits the normalization $(a^2+b^2)/2 = 1$; after interchanging a, b if necessary, write $a^2 = 1+t$, $b^2 = 1-t$, where $0 \le t \le 1$. The derivative of the right side minus the left side is $t/\sqrt{1-t^2} - \operatorname{arctanh} t \ge 0$: its derivative is $(1-t^2)^{-3/2} - (1-t^2)^{-1} \ge 0$, and both differences vanish at zero. Continuity covers $t = 1$. For independent finite random variables X, Y and $f \ge 0$, entropy tensorizes because

$$\mathrm{Ent} * \mathsf{X}, \mathsf{Y} f = \mathbb{E} * \mathsf{Y}\, \mathrm{Ent} * \mathsf{X} f + \mathrm{Ent} * \mathsf{Y}(\mathbb{E}_{\mathsf{X}} f) \le \mathbb{E}_{\mathsf{Y}}\, \mathrm{Ent} * \mathsf{X} f + \mathbb{E} * \mathsf{X}\, \mathrm{Ent} * \mathsf{Y} f.$$

The inequality uses convexity of entropy: for positive f, its second directional derivative is $\mathbb{E}(k^2/f) - (\mathbb{E}k)^2/\mathbb{E}f \ge 0$ by Cauchy–Schwarz; continuity covers zeros. Iteration and the two-point estimate give

$$\mathrm{Ent} * \mu(\varphi^2) \le \frac{1}{2} \sum_{j=1}^{d} \mathbb{E}_\mu (\varphi(x) - \varphi(x^{(j)}))^2 = d\, \mathcal{E}_{P_{\mathrm{cube}}}(\varphi),$$

where $x^{(j)}$ flips coordinate j of x and P_{cube} flips one uniformly selected coordinate. Bernoulli's inequality gives $P(x, x^{(j)}) = (2d)^{-1}(1 - 1/(4d))^{d-1} \ge 3/(8d)$. Keeping just these nonnegative contributions in the Dirichlet form gives $\mathcal{E}_P(\varphi) \ge (3/8)\mathcal{E}_{P_{\mathrm{cube}}}(\varphi)$. Consequently

$$\mathrm{Ent}_\mu(\varphi^2) \le (8d/3)\mathcal{E}_P(\varphi). \tag{13}$$

Full information. For every finite nonnegative squared partition of unity on any normalized symmetric endpoint law (not just the cube),

$$I \le \sum_i \mathrm{Ent} * \mu(\chi_i^2) + K(1 + 2\log(1/K)), \tag{14}$$

with the final term zero at $K = 0$. To prove it, for $0 \le q \le v$, $m_v = \mathbb{E}v$ and $d' = m_v - \mathbb{E}q$, write $h_{\mathrm{bin}}(z) = -z\log z - (1-z)\log(1-z)$. When $m_v > 0$, the identity

$$\mathrm{Ent}(q) + \mathrm{Ent}(v-q) - \mathrm{Ent}(v) = m_v h * \mathrm{bin}(d'/m_v) - \mathbb{E}[v h_{\mathrm{bin}}(q/v)]$$

(with zero summands omitted), nonnegativity of entropy, and $h_{\mathrm{bin}}(z) \le z\log(e/z)$ give $\mathrm{Ent}(q) \le \mathrm{Ent}(v) + d'\log(em_v/d')$. If $m_v = 0$, both q and v vanish almost surely and the bound is immediate; if $d' = 0$, interpret its last term as zero. Apply this with $q_i = \chi_i(x)\chi_i(y)$ and $v_i = (\chi_i(x)^2 + \chi_i(y)^2)/2$. The pointwise inequality $q_i \le v_i$ is the arithmetic–geometric mean inequality. Convexity of entropy and the fact that both endpoint marginals are μ give $\mathrm{Ent}_a(v_i) \le \mathrm{Ent}_\mu(\chi_i^2)$, while $M_i = \mathbb{E}_a v_i = \mathbb{E}_\mu \chi_i^2$, $d_i = M_i - \mathbb{E}_a q_i = \mathcal{E}_P(\chi_i)$, $\sum_i d_i = K$, and $\sum_i M_i = 1$. For $K > 0$, Jensen's inequality with weights d_i/K gives

$$\sum_i d_i \log(eM_i/d_i) \le K \log\left(\sum_{i : d_i > 0} \frac{d_i}{K} \frac{eM_i}{d_i} \right) \le K\log(e/K),$$

where the last inequality uses $\sum_i M_i = 1$. The assertion is immediate when $K = 0$. Finally $0 \le w_\dagger \le 1$ implies $\mathrm{Ent}_a(w_\dagger) \le K\log(1/K)$, proving (14). Terms of zero mass are interpreted by continuity.

Return now to the cube, with uniform law μ and endpoint law $a_{xy} = 2^{-d}P(x,y)$. Walsh characters of degree j have Q-eigenvalue ϱ^j, so $I - P$ has gap $2(1-\varrho) = 1/d$. Properness of the support of child 1 gives a vertex x_0 with $\chi_1(x_0) = 0$, and properness of the support of child 0 gives a vertex x_1 with $\chi_0(x_1) = 0$. Thus $z = (\chi_0, \chi_1)$ equals $e_0 = (1,0)$ and $e_1 = (0,1)$, respectively. They are distinct because $\|z(x)\|_2 = 1$ everywhere. For any vector v, $\|e_0 - v\|_2^2 + \|e_1 - v\|_2^2 = 1 + 2\,\|v - (e_0+e_1)/2\|_2^2 \geq 1$. The spectral-gap inequality, applied to each coordinate of z, therefore gives

$$K \geq \frac{1}{d}\mathbb{E}_\mu \|z - \mathbb{E}_\mu z\|^2 \geq \frac{1}{d2^d},$$

where the last bound keeps just x_0, x_1, each of mass 2^{-d}. Combining (13) and (14) gives

$$I \leq K\left(\frac{8d}{3} + 1 + 2d\log 2 + 2\log d\right) \leq 6dK.$$

For the last inequality, $(1 + 2\log d)/d$ decreases for $d \geq 2$, and $19/6 + 3\log 2 < 6$. Because the support of child 1 is nonempty, some diagonal pair with positive a-mass has $w_1 > 0$, so $p_1 > 0$. Because that support is proper, another positive-mass diagonal pair lies outside $\mathcal{A}_1$, so $\mathrm{pr}_1 < 1$. Thus (10) implies $I > 0$.

Time, trace, and rank. Bernoulli's inequality gives $Q(x,x) = (1 - 1/(4d))^d \geq 3/4$, so $P(x,x) \geq 1/2$. Off-diagonal entries of P are positive, its row sums are one, and it is symmetric and rational. This verifies the asserted half-lazy walk and uniform stationary law. Walsh expansion gives $\mathrm{Tr}(Q^{2t}) = (1 + \varrho^{2t})^d$ and

$$T_U = \left\lceil \frac{\log(1/(2^{1/U} - 1))}{-2\log\varrho} \right\rceil, \qquad \frac{d}{2}\log U \leq T_U \leq 4d\log U.$$

Here $T_U \geq 1$ because the defining inequality fails at $t = 0$. For the time bounds use $(\log 2)/U \leq 2^{1/U} - 1 \leq 1/U$ and $1/d \leq -2\log\varrho \leq 2/d$. In particular the lower bound follows before taking the ceiling, while the upper bound follows from $T_U \leq d\log(U/\log 2) + 1 \leq 4d\log U$ for $d, U \geq 2$. The lower time bound gives $1/(6d) \geq (\log U)/(12T_U)$, completing (11). Moreover, $\log R/\eta = 32T_U \log 2$, so

$$\frac{\log R}{\eta}\frac{K}{I} \geq 32T_U \log 2 \frac{\log U}{12T_U} = \frac{8}{3}(\log 2)\log U,$$

which is the middle inequality in (12). Put $F_t = d\log(1 + \varrho^{2t})$. Concavity and $\log(1+0) = 0$ give $F_T \geq \varrho^2 F_{T-1}$. The minimality of T gives $F_{T-1} > d\log 2/U$, hence $F_T > \varrho^2 d\log 2/U > \frac{1}{2}s\log 2$, proving the trace assertion. Here $s \geq 16$ and $\varrho^2 \geq 9/16$. Moreover, $R = 2^s \geq 2^{16} > 256$, while $R \leq 2^d = n$, and $U \leq d/s \leq 16U/15$. Also

$$\eta L_R^{\mathrm{h}} = \frac{d}{32T} \leq \frac{1}{16\log U}.$$

In particular, $0 < \eta \leq 1/(16U\log U) < 1/16$, and multiplying the displayed bound by 3072 gives $3072\eta L_R^{\mathrm{h}} \leq 192/\log U$. For integer U, the test defining T_U is equivalently $(1 + \varrho^{2t})^U \leq 2$, an exact rational comparison, so T_U can be found by finite search. Thus R and η are rational algorithm parameters. Since $\log_2 R = s$ and $32\eta = s/T$, the heat-time formula in Lemma 3.2 equals T exactly.

Finally let $m \in \{0, \ldots, d\}$ be the largest degree with $2(1 - \varrho^m) \leq \eta$. Since $\eta \leq 1/(16U\log U) < 1/16$, the defining inequality for m and $-\log(1-z) \leq z/(1-z)$ give

$$m \leq \frac{-\log(1 - \eta/2)}{-\log\varrho} \leq \frac{d\eta}{1 - \eta/2} \leq \frac{2s}{31\log U} < \frac{s}{15\log U} =: a < d/2.$$

If $m = 0$, the rank is one, which is already less than $R^{5/8}$. Otherwise the inclusive cutoff rank r obeys

$$r = \sum_{j=0}^{m} \binom{d}{j} \le \exp(d h_{\mathrm{bin}}(m/d)) \le \exp(m \log(ed/m)) \le \exp\left(\frac{s \log(16eU \log U)}{15 \log U}\right).$$

The first inequality follows by setting $z = m/(d-m) \le 1$ in $\sum_{j \le m} \binom{d}{j} \le z^{-m}(1+z)^d$; the last inequality uses monotonicity of $x \log(ed/x)$ for $x < d$ and $d/s \le 16U/15$. The function $[\log(16e) + y + \log y]/(15y)$ decreases for $y \ge \log 2$, and its value at $\log 2$ is below $2/5$, since $e \log 2 < 2$. Hence $\log r < 2s/5 < (5/8)s \log 2$, proving the rank bound. □

Taking integer U arbitrarily large and an integer $d \ge 16U$ makes $3072\eta L_R^{\mathrm{h}}$ arbitrarily small, while the normalized full-channel ratio is bounded below by a positive constant times $\log U$, uniformly over the proper binary splits in the proposition. For example, $\log U \ge 6528$ suffices for the hypothesis $3072\eta L_R^{\mathrm{h}} \le 1/34$ of Theorem 3.1; the displayed spectral and channel inequalities themselves hold for every integer $U \ge 2$, $d \ge 16U$. Thus even within the smallness regime no absolute constant C can guarantee the existence of a proper binary split with $K \le C(\eta/\log R)I$ on every high-trace node.

Nevertheless, these roots are connected and their inclusive cutoff ranks satisfy the stopping test $r^8 \le R^5$ of Theorem 3.1, before any split. Stationary masses equal ν, so no mass lies below the cleanup threshold either. In particular, the construction has the single leaf $\Phi = 1$ and total loss $D = 0$ on these roots. The example therefore does not rule out a log-free local estimate *under rank-screen failure*, or a better charge restricted to reached nodes. It gives no lower bound on whole-tree loss, labeling error, or runtime. The Fourier and entropy–energy ingredients are classical; the limitation is their application to this particular channel and parameter regime.

A From a low-rank restriction to a partial assignment

We use one conditional-marginal construction and select among its roundings by their actual partial-assignment charge. Its ingredients are the singleton embedding and column-selection theorems of Guruswami–Sinop [15, Theorems 5.2 and 6.9]. Their local-fragment approach [14] avoids constructing an entire moment hierarchy. We record the fragment and its finite construction because a theorem about the final labeling alone does not supply the interface needed to transport a soft restriction back to its invocation root.

A.1 The finite conditional-marginal ensemble

The constructions below use four finite-precision facts as black boxes. We state them once, in the form in which they are used.

Lemma A.1 (Finite-precision toolkit)**.** *Each of the following runs in time polynomial in the binary length of its input.*

(i) Rational inertia. *For a rational symmetric matrix* A*, symmetric elimination with* 1×1 *and* 2×2 *pivots computes the numbers of positive, negative, and zero eigenvalues exactly, and when a negative eigenvalue exists it returns a rational vector* v *with* $v^{\top}\mathsf{A}v < 0$*. Bit lengths stay polynomial because every Schur complement entry is a quotient of minors of the integer matrix obtained by clearing denominators; the details are given in the proof of Theorem 3.1.*

(ii) Shallow-cut ellipsoid method. *Let $C \subset \mathbb{R}^N$ be a convex body containing a ball of radius ρ and contained in a given rational ellipsoid. Suppose that for every rational center x an oracle either certifies $x \in C$ or returns a rational half-space containing C whose bounding hyperplane lies within relative depth $1/(N+1)$ of x. Then the method with rounded rational updates certifies a point of C after $O(N^2 \log(N \cdot \operatorname{diam}/\rho))$ oracle calls [12, Section 3.3].*

(iii) Convex minimization with an inradius. *A convex quadratic with rational data can be minimized over a rational polytope of known inradius, to a prescribed rational additive accuracy, by the ellipsoid method and binary search [12, Section 3.3 and Chapter 4].*

(iv) Exact rational linear programming *[12, Theorem 6.4.12].*

Lemma A.2 (A conditional-marginal ensemble at low base rank)**.** *Let a normalized rational weighted Unique Games instance have positive degree measure μ, alphabet $q \geq 2$, and symmetric normalized Laplacian H. Suppose*

$$r = \operatorname{rank} \mathbf{1}_{[0,\eta]}(H), \qquad 0 < 128\kappa_0 \leq \eta \leq 1/4.$$

If its optimum is at least $1 - \kappa_0$, a deterministic algorithm returns at most q^{2r} rational simplex fields p^σ with positive rational weights w_σ summing to one, such that

$$\mathbb{E}_\sigma \widehat{\mathcal{E}}(p^\sigma) \leq \frac{65}{64}\kappa_0, \qquad \mathbb{E}_\sigma \sum_u \mu(u)(1 - \|p_u^\sigma\|^2) \leq \frac{5\kappa_0}{\eta}. \tag{15}$$

Its running time is $q^{O(r)} \operatorname{poly}(n, q, b_{\mathrm{in}})$, where b_{in} is the binary input length, including η and κ_0. Without the optimum promise, it returns an ensemble passing the displayed rational checks or declares failure in the same time. Parallel records and arbitrary permutation loops are allowed.

Proof. If $2r \geq n$, enumerate every labeling and retain a minimum-error one. This costs at most q^{2r} times polynomial work. If its error exceeds $65\kappa_0/64$, declare failure; otherwise return its one-hot field as a singleton ensemble of zero impurity. Hence assume $2r < n$.

The required moment fragment. Its singleton indicator vectors $X_{u,a}$ satisfy $\sum_a X_{u,a} = X_\varnothing$, $\|X_\varnothing\|^2 = 1$, and orthogonality within every fiber. Its objective is

$$F = \frac{1}{2} \sum_e a_e \sum_a \left\| X_{u,a} - X_{v,\pi_e(a)} \right\|^2 \leq F_{\max} := 65\kappa_0/64.$$

We construct a seed set S, $|S| \leq 2r$, and compatible joint indicators for $S \cup \{u\}$, for every u. If Π_S^X is projection onto the seeded singleton span, the construction ensures

$$\sum_u \mu(u) \sum_a \left\| (I - \Pi_S^X) X_{u,a} \right\|^2 \leq 5\kappa_0/\eta. \tag{16}$$

Seed selection from singleton inner products. The embedding of [15, Theorem 6.9] is

$$Y_u = \sum_{a:\, X_{u,a} \neq 0} \frac{X_{u,a} \otimes X_{u,a}}{\|X_{u,a}\|}, \qquad \|Y_u\|^2 = \sum_a \|X_{u,a}\|^2 = 1,$$

where the norm identity uses within-fiber orthogonality. It supplies unit vectors Y_u with

$$\|Y_u - Y_v\|^2 \leq 2 \sum_a \left\| X_{u,a} - X_{v,\pi(a)} \right\|^2, \qquad \sum_a \left\| (I - \Pi_S^X) X_{u,a} \right\|^2 \leq \left\| (I - \Pi_S^Y) Y_u \right\|^2,$$

where Π_S^Y projects onto the span of $(Y_v)_{v\in S}$. This embedding uses within-fiber orthogonality, not entrywise positivity of cross-fiber inner products. Indeed, for $A = \|x\| > 0$, $B = \|y\| > 0$ and $c = \langle x, y\rangle$, the matched-pair estimate follows from

$$A^2 + B^2 - 4c + 2c^2/(AB) = (A-B)^2 + 2AB(1 - c/(AB))^2 \geq 0.$$

For the Gram matrix $\mathcal{G}$ of the columns $\sqrt{\mu(u)}Y_u$, $\mathrm{Tr}(\mathcal{G}H) \leq 2F$, and therefore $\sum_{i>r} \lambda_i^{\downarrow}(\mathcal{G}) \leq 2F/\eta$. The column-selection theorem [15, Theorem 5.2], with $2r$ columns, gives residual at most twice this spectral tail.

Here is a rational implementation retaining slack. Put $\delta = \kappa_0/(16n\eta)$ and compute a rational positive definite $\mathcal{G}_+$ with $\mathcal{G} \preceq \mathcal{G}_+ \preceq \mathcal{G} + 2\delta I$. Its entries use

$$\mathcal{G}_{uv} = \sqrt{\mu(u)\mu(v)} \sum_{a,b} \frac{\langle X_{u,a}, X_{v,b}\rangle^2}{\sqrt{\|X_{u,a}\|^2 \|X_{v,b}\|^2}},$$

where a zero-denominator term is zero. Terms with denominator below $\tau = \delta/(16nq^2)$ can be omitted: positive semidefiniteness bounds each such term by τ. Their sum in any entry is at most $q^2\tau = \delta/(16n)$. Rational square-root bisection on the remaining terms produces a symmetric rational $\widetilde{\mathcal{G}}$ with entrywise error at most $\delta/(4n)$. Set $\mathcal{G}_+ = \widetilde{\mathcal{G}} + \delta I$. The operator-norm error of $\widetilde{\mathcal{G}}$ is at most $\delta/4$, so $\mathcal{G} + 3\delta I/4 \preceq \mathcal{G}_+ \preceq \mathcal{G} + 5\delta I/4$. All required precision lengths are polynomial in the rational input length and $\log(1/\delta)$.

Here are the exact conditional-expectation computations for selecting $2r$ columns. For a positive definite matrix $\mathcal{D}$, put $e_j(\mathcal{D}) = \sum_{|T|=j} \det \mathcal{D}[T,T]$, with $e_0(\mathcal{D}) = 1$; these are the coefficients of $\det(I + t\mathcal{D})$. Choosing k columns with weights proportional to their principal determinants gives expected Schur residual $(k+1)e_{k+1}(\mathcal{D})/e_k(\mathcal{D})$. Indeed each residual diagonal is a bordered determinant divided by the selected principal determinant, so each determinant of order $k+1$ is counted $k+1$ times. If $\mathcal{D}$ is the Schur complement after the already fixed choices and k more columns are needed, write $\mathcal{D}/i$ for the Schur complement of its diagonal pivot at i. The conditional expected final residual and its averaging weights are

$$E_i = \frac{k e_k(\mathcal{D}/i)}{e_{k-1}(\mathcal{D}/i)}, \qquad p_i = \frac{\mathcal{D}_{ii} e_{k-1}(\mathcal{D}/i)}{k e_k(\mathcal{D})}.$$

They satisfy $\sum_i p_i = 1$ and $\sum_i p_i E_i = (k+1)e_{k+1}(\mathcal{D})/e_k(\mathcal{D})$ by the same determinant counting. Choose an index minimizing E_i, replace $\mathcal{D}$ by $\mathcal{D}/i$, and decrement k until no choices remain. Characteristic-polynomial coefficients and Schur complements have polynomial rational bit cost; no family of n^{2r} subsets is enumerated. For eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n > 0$, each product of $k+1$ eigenvalues contains at least $k+1-r$ tail indices. Thus $(k+1-r)e_{k+1} \leq (\sum_{i>r} \lambda_i)e_k$. At $k = 2r$, the initial expected residual is less than twice the tail beyond rank r. Loewner monotonicity of Schur residuals now gives

$$\sum_u \mu(u) \sum_a \left\|(I - \Pi_S^X)X_{u,a}\right\|^2 \leq 4F/\eta + 4n\delta \leq \frac{69}{16}\frac{\kappa_0}{\eta} < 5\kappa_0/\eta.$$

Indeed, a Schur-complement diagonal is the least squared residual obtained by regressing that column on the selected columns. Hence $\mathcal{G} \preceq \mathcal{G}_+$ implies that the residual trace for $\mathcal{G}$ is at most the one just bounded for $\mathcal{G}_+$. This seed procedure reads only the singleton inner products.

Compatible local moments, without an $n^{O(r)}$ hierarchy. Use linearly independent centered indicator functions at each vertex, $z_{u,a} = \mathbf{1}_{\ell(u)=a} - 1/q$, $a < q$; eliminate the last indicator by the sum-to-one identity and reduce repeated-vertex products by

$$z_a^2 = (1 - 2/q)z_a + (q-1)/q^2, \qquad z_a z_b = -(z_a + z_b)/q - 1/q^2 \quad (a \neq b).$$

The base moments are the monomials on at most two vertices; their nonconstant number is $d_{\mathrm{b}} = n(q-1) + \binom{n}{2}(q-1)^2$. For a seed set of size s, index both rows and columns of the symmetric matrix $\mathcal{M}_S$ by the distinct reduced monomials supported in $S \cup \{u\}$ for some u, including the constant monomial. Its entry at two row indices is the formal expectation of their reduced product. Use one scalar moment variable for each distinct nonconstant monomial occurring in these products; all occurrences of the same monomial share that variable. There are $m_S \le (1 + n(q-1))q^s$ rows and $N_S = O(n^2 q^{s+2})$ nonconstant moment coordinates. Let

$$\mathcal{K}_S = \{y : \mathcal{M}_S(y) \succeq 0,\ F(y) \le F_{\max}, -2 \le y_A \le 2\}.$$

The constant moment is fixed and all indicator identities have already been eliminated. Rational inertia (Lemma A.1(i)) gives a strong separation oracle for this spectrahedron in time polynomial in its explicit size: if $\mathcal{M}_S(y)$ is not positive semidefinite, a rational witness v with $v^\top \mathcal{M}_S(y) v < 0$ yields the valid linear inequality $v^\top \mathcal{M}_S(y') v \ge 0$, which is violated at y; the box and objective constraints are linear.

The following inradii justify finite local-fragment optimization. Mix a promised witness distribution with uniform independent labels, using uniform mass $\sigma_{\mathrm{mix}} = \kappa_0/256$. Its objective is at most $\kappa_0 + \sigma_{\mathrm{mix}} < F_{\max}$. For each base monomial, the product of the appropriate functions $\psi_{u,a} = q(\mathbf{1}_{\ell(u)=a} - \mathbf{1}_{\ell(u)=q})$ is its dual under uniform measure and has absolute value at most q^2. Thus a base perturbation of infinity norm at most

$$\rho_* := \frac{\kappa_0}{1024 d_{\mathrm{b}} q^2}$$

is realized by adding the corresponding dual linear combination to the mixture's density relative to uniform measure. Its absolute size is at most $d_{\mathrm{b}} q^2 \rho_* = \sigma_{\mathrm{mix}}/4$, and its mean is zero. The resulting distribution stays nonnegative and has objective less than $F_{\max}$. Consequently $C_* := \bigcap_{|S| \le 2r} \mathrm{proj}_{\mathrm{base}} \mathcal{K}_S$ contains a Euclidean ρ_*-ball. This body is used only in the proof and is never enumerated.

Each $\mathcal{K}_S$ also contains a local ball of explicitly known radius; its center need not be known to the algorithm. Use the moments of the same witness–uniform mixture as its center. Under uniform labels, different vertex supports of centered monomials are orthogonal, and the one-vertex covariance has minimum eigenvalue q^{-2}. The uniform row Gram matrix therefore has minimum eigenvalue at least $q^{-2(s+1)}$, so the mixture's Gram matrix has minimum eigenvalue at least $\sigma_{\mathrm{mix}} q^{-2(s+1)}$. A coordinate perturbation of Euclidean norm a changes $\mathcal{M}_S$ in operator norm by at most $m_S 2^{s+1} a$, because reduced products have coefficient one-norm at most 2^{s+1}. The objective has coefficient one-norm at most $10q^2$. It follows that $\mathcal{K}_S$ contains a ball of radius

$$\rho_S = \min\left\{\frac{1}{4}, \frac{\kappa_0}{4096 q^2}, \frac{\kappa_0 q^{-2(s+1)}}{1024 m_S 2^{s+1}}\right\}.$$

Indeed the last radius bound permits a Gram perturbation of at most one quarter of this minimum eigenvalue. The objective increases by at most $10q^2\rho_S \le 10\kappa_0/4096$, smaller than its slack at the mixture, and the moment box has slack at least one at the center. In particular $\log(1/\rho_S) = O(s \log q + \log n + \log(1/\kappa_0))$.

Run a shallow-cut ellipsoid in the d_{b} base coordinates, starting with a ball containing $[-2,2]^{d_{\mathrm{b}}}$. At a base point x, first test the box, objective, and base PSD conditions, the latter using $\mathcal{M}_\varnothing(x)$ with constant and centered-singleton row indices. A violated condition supplies a separating inequality; weakening it so that its boundary passes through x gives a valid central cut. If all pass, select $S(x)$ by the preceding seed procedure and apply Lemma A.3 below to $\mathcal{K}_{S(x)}$, with fixed base coordinates

x and tolerance $\epsilon = \rho_*/(100d_{\mathrm{b}})$. An extension ends the algorithm. Otherwise the returned inequality $\langle c, z - x\rangle \le \epsilon$, $\|c\|_\infty = 1$, holds for every $z \in C_*$. Every current ellipsoid contains its ρ_*-ball; its directional half-width is at least ρ_*, so this is a valid shallow cut of relative depth at most $1/(100d_{\mathrm{b}})$. After $O(d_{\mathrm{b}}^2 \log(d_{\mathrm{b}}/\rho_*))$ unsuccessful cuts the volume would be smaller than that ball, a contradiction. The algorithm therefore returns a fragment with exactly the base moments that selected its seeds. Rounded rational ellipsoid updates with fixed precision are used throughout (Lemma A.1(ii)). Keep violated box cuts exact. For all remaining calls the center x lies in the box; round each normalized cut direction to infinity error at most $\epsilon/(16d_{\mathrm{b}})$ and add the corresponding $4d_{\mathrm{b}}$ times that error to its offset; this preserves validity and changes the relative depth by less than a factor two. This also prevents successive oracle calls from inflating the outer coefficient length. All dimensions are $q^{O(r)}\,\mathrm{poly}(n, q)$, and the logarithms of the inverse radii and precisions are polynomial in the stated binary inputs. Fixed budgets and exact output checks give the asserted failure behavior without the promise.

Rounding the fragment. Let $X_S(\sigma)$ be its mutually orthogonal seed indicators, $w_\sigma = \|X_S(\sigma)\|^2$, and omit zero weights. They sum to $X_\varnothing$, so the positive weights sum to one. Set

$$p_u^\sigma(a) = \frac{\langle X_S(\sigma), X_{u,a}\rangle}{w_\sigma}.$$

The seed–singleton indicator is a local row combination; its idempotence and PSD show that its moment is nonnegative. Exact indicator identities show that each p_u^σ is a simplex vector. The projection Π_S^{seed} onto the seed-indicator span contains the seeded singleton span, and orthogonal projection gives

$$\mathbb{E}_\sigma \widehat{\mathcal{E}}(p^\sigma) \le F, \qquad \mathbb{E}_\sigma \sum_u \mu(u)(1 - \|p_u^\sigma\|^2) = \sum_u \mu(u) \sum_a \left\|(I - \Pi^{\mathrm{seed}} * S)X * u, a\right\|^2 \le 5\kappa_0/\eta.$$

These are precisely (15); retain all positive-weight seeds, rather than imposing separate seedwise energy and purity tests. The simplex, normalization, and two mean inequalities can be checked exactly. Small positive weights affect bit length, not repetition counts. Algorithmically, every weight and numerator above is an entry or a fixed linear combination of entries of the rational moment matrix; no Gram factorization, and hence no irrational vector representation, is computed. A label-changing loop contributes the nonnegative unary objective $a_e \sum_{a:\pi_e(a)\neq a} \|X_{u,a}\|^2$; hence none of the preceding arguments discard such constraints. □

Lemma A.3 (Fixed-coordinate extension or a shallow cut)**.** *Let $\mathcal{K} \subset [-2, 2]^N$ be a rational strongly separable convex body containing a ball of known rational radius $0 < \rho \le 1/4$. Given integers $1 \le d_{\mathrm{b}} \le N$, rational fixed coordinates $x \in [-2, 2]^{d_{\mathrm{b}}}$, and rational $0 < \epsilon < 1$, one can return either a rational $y \in \mathcal{K}$ with $y_{\mathrm{base}} = x$, or rational $c \in \mathbb{R}^{d_{\mathrm{b}}}$ with*

$$\|c\|_\infty = 1, \qquad \langle c, z_{\mathrm{base}} - x\rangle \le \epsilon \quad (z \in \mathcal{K}).$$

The time is polynomial in the dimensions, oracle and input lengths, $\log(1/\rho)$, and $\log(1/\epsilon)$. With these budgets imposed, failure is permitted when the inradius promise is false.

The cut alternative is allowed even if the fiber is nonempty. The lemma does not assert an exact decision procedure for feasibility of an arbitrary lower-dimensional spectrahedral slice.

Proof. This is the inradius-aware restricted-support construction underlying [14]; the short proof specifies its geometric margin. If $N = d_{\mathrm{b}}$, query the strong oracle at x. It either certifies $x \in \mathcal{K}$ or returns a strict separating inequality; after normalizing its base normal in infinity norm, that

inequality has the stated form with right side at most zero. Otherwise put $\delta = \epsilon\rho/(128\lceil\sqrt{d_{\mathrm{b}}N}\rceil)$ and $\beta = (\epsilon\delta/(64d_{\mathrm{b}}))^2$. Run a rational ellipsoid on the affine fiber, retaining all full-space cuts and the enclosing box. A feasible point ends the algorithm. A violated inequality with zero free part is already a separator in the base coordinates. Otherwise finite volume reduction returns a polytope $\mathcal{K} \subset P \subset [-2,2]^N$ whose fixed-coordinate fiber has volume less than a free-coordinate ball of radius δ.

Write $P = \{z : C_i z \le b_i\}$. Choose rational $\|C_i\|_2 \le h_i \le 2\|C_i\|_2$ and set $P' = \{z : C_i z \le b_i - 2\delta h_i\}$. It contains a $\rho/2$-ball and lies inside the radius-2δ erosion of P. Minimize $\|z_{\mathrm{base}} - x\|^2$ over P' to additive accuracy β, obtaining a rational feasible point $\widehat{y}$. This requires only ordinary bounded convex quadratic optimization with the known inradius (Lemma A.1(iii)). More explicitly, let $f(z) = \|z_{\mathrm{base}} - x\|^2$ and $0 < \gamma \le 1$. If y_* is a minimizer and $B(w, \rho/2) \subset P'$, the points $(1-t)y_* + tz$, $z \in B(w,\rho/2)$, with $t = \gamma/(32d_{\mathrm{b}})$, form a ball of radius $\gamma\rho/(64d_{\mathrm{b}})$. Convexity and $f(z) \le 16d_{\mathrm{b}}$ throughout $P' \subset [-2,2]^N$ place this entire ball in the sublevel $f \le f(y_*) + \gamma/2$. Ellipsoid feasibility of such sublevels and binary search therefore have bit-polynomial budgets in $\log(1/\rho) + \log(1/\beta)$; no unknown center is computed.

Let $v = \widehat{y}_{\mathrm{base}} - x$ and let v_* be the base displacement of a minimizer. The thin-fiber certificate and the 2δ erosion imply $\|v\| > \delta$. Indeed $\widehat{y} \in P'$ gives $B(\widehat{y}, 2\delta) \subseteq P$; if $\|v\| \le \delta$, this ball would meet the fiber $\{z : z_{\mathrm{base}} = x\}$ in a free-coordinate ball of radius at least $\sqrt{3}\,\delta > \delta$, contradicting the volume certificate, which bounds the fiber of P and hence also the fiber of $\mathcal{K} \subseteq P$. Optimality gives $\|v - v_*\| \le \sqrt{\beta}$. If w is a ρ-ball center in P, then for every $z \in P$, $z_0 = (1 - 4\delta/\rho)z + (4\delta/\rho)w$ lies in P' and $\|z - z_0\| \le 16\delta\sqrt{N}/\rho$. First-order optimality gives $\langle v_*, (z_0)_{\mathrm{base}} - x\rangle \ge \|v_*\|^2 \ge 0$. Consequently $c = -v/\|v\|_\infty$ satisfies

$$\langle c, z_{\mathrm{base}} - x\rangle \le \frac{4d_{\mathrm{b}}\sqrt{\beta}}{\delta} + \frac{16\delta\sqrt{d_{\mathrm{b}}N}}{\rho} < \epsilon.$$

Certify this inequality over the rational polytope P by rational linear programming (Lemma A.1(iv)) before returning it; since $\mathcal{K} \subseteq P$, validity over P implies validity over $\mathcal{K}$, which is the direction the lemma asserts. Check any proposed extension with the original strong oracle as well. A failed budget, failed check, or zero normal to be normalized results in failure, not an unchecked output; these exits cannot occur under the promised conditions. Thus failed inradius promises cannot install invalid cuts. Each ellipsoid call uses bounded rational updates and its predetermined volume budget. All stated inverse precisions occur logarithmically, proving the running-time assertion. □

A.2 Ensemble selection with separate disagreement and amplitude costs

All quantities in this subsection use one normalized symmetric directed invocation root. Its scalar law is half-lazy. For a nonempty partial labeling on U', write B for its directed internal violated mass and $C_{\mathrm{out}} = \sum_{u\in U', v\notin U'} a_{uv}$, and define its charge

$$\Gamma = \frac{B + 2C_{\mathrm{out}}}{\mu(U')}.$$

Let $g_u = s_u e_{\ell(u)}$ with $s_u \ge 0$; labels at zero amplitudes are arbitrary. Define the root quantities

$$M(g) = \sum_u \mu(u) s_u^2, \qquad A(g) = \mathcal{E}_P(s),$$

$$\mathsf{J}(g) = \sum_{e:u\to v} a_e s_u s_v \mathbf{1}_{\ell(v)\neq\pi_e(\ell(u))}, \qquad \Omega(g) = \sum_u \sum_{v\neq u} a_{uv} s_u^2,$$

$$\mathcal{I}(g) = \sum_{e:u\to v} a_e \min\{s_u^2, s_v^2\} \mathbf{1}_{\ell(v)\neq\pi_e(\ell(u))}, \qquad \mathcal{V}(g) = \frac{1}{2}\sum_{u,v} a_{uv}|s_u^2 - s_v^2|.$$

Lemma A.4 (Disagreement-sensitive ensemble coarea)**.** *For each such field,*

$$\widehat{\mathcal{E}}(g) = A(g) + \mathsf{J}(g), \quad 0 \le A(g) \le \Omega(g) \le M(g)/2, \quad \mathcal{I}(g) \le \mathsf{J}(g), \quad \mathcal{V}(g)^2 \le A(g)(2\Omega(g) - A(g)). \tag{17}$$

For a finite ensemble with positive weights summing to one, let bars denote weighted means and assume $\overline{M} > 0$. Threshold each field on $U_{\sigma,t} = \{u : s_u^2 > t\}$ for $t \ge 0$. Its least nonempty seed–threshold charge satisfies

$$\Gamma_* \le \frac{\overline{\mathcal{I}} + 2\overline{\mathcal{V}}}{\overline{M}} \le \frac{\overline{\mathsf{J}}}{\overline{M}} + \frac{2\sqrt{\overline{A}(2\overline{\Omega} - \overline{A})}}{\overline{M}} \le \mathsf{j} + 2\sqrt{a(1-a)}, \tag{18}$$

where $a = \overline{A}/\overline{M} \in [0, 1/2]$ and $\mathsf{j} = \overline{\mathsf{J}}/\overline{M}$. For rational data the minimum is computable exactly with at most n thresholds per nonzero field and polynomial work in the ensemble encoding.

Proof. On every directed record, with permutation pullback understood,

$$\left\| s_u e_{\ell(u)} - T_e s_v e_{\ell(v)} \right\|^2 = (s_u - s_v)^2 + 2 s_u s_v \mathbf{1}_{\ell(v) \ne \pi_e(\ell(u))}.$$

Summing proves $\widehat{\mathcal{E}} = A + \mathsf{J}$. Also $A = \Omega - \sum_{u \ne v} a_{uv} s_u s_v \le \Omega$, and half-laziness gives $\Omega \le M/2$. The inequality $\min\{s_u^2, s_v^2\} \le s_u s_v$ gives $\mathcal{I} \le \mathsf{J}$. After discarding scalar loops, Cauchy–Schwarz gives

$$\mathcal{V}^2 \le \left(\frac{1}{2} \sum_{u \ne v} a_{uv} (s_u - s_v)^2 \right) \left(\frac{1}{2} \sum_{u \ne v} a_{uv} (s_u + s_v)^2 \right) = A(2\Omega - A).$$

For one field, integrating over $t \in [0, \infty)$, coarea gives exactly $\int \mu(U_t)\, dt = M$, $\int B_t\, dt = \mathcal{I}$, and $\int C_{\mathrm{out},t}\, dt = \mathcal{V}$: internal records survive an interval of length $\min\{s_u^2, s_v^2\}$, and a directed boundary record contributes $(s_u^2 - s_v^2)_+$. Symmetry pairs opposite boundary orientations. Loops contribute only to the internal term. Average $\Gamma_{\sigma,t}$ with density $w_\sigma \mu(U_{\sigma,t})/\overline{M}$ to prove the first bound. Another Cauchy–Schwarz inequality gives $\overline{\mathcal{V}} \le \sqrt{\overline{A}(2\overline{\Omega} - \overline{A})}$. Finally, $2\overline{\Omega} - \overline{A} \le \overline{M} - \overline{A}$ because $\overline{\Omega} \le \overline{M}/2$; division by $\overline{M}$ proves the last inequality in (18). For finite selection sort the distinct positive rational s_u^2 and test every interval midpoint, including the interval from zero to the smallest level. All nonempty thresholds occur and all charges are rational. Zero fields are omitted. No square roots or random draws from possibly tiny seed masses are used in selection. □

For a nonzero root multiplier $\chi \ge 0$, use

$$M_\chi = \sum_u \mu(u) \chi(u)^2, \quad Z_\chi = \sum_e a_e \chi(u) \chi(v) > 0, \quad \Delta_\chi = M_\chi - Z_\chi, \quad r_\chi = \Delta_\chi / Z_\chi.$$

The normalized restriction has record weights $a_e \chi(u)\chi(v)/Z_\chi$ and degree law μ_χ. Write $\Delta_\chi(\cdot)$ and $\widehat{\mathcal{E}}_\chi(\cdot)$ for its scalar and constraint energies, respectively; the scalar $\Delta_\chi = M_\chi - Z_\chi$ continues to denote the root multiplier energy. Scalar half-laziness gives $Z_\chi \mu_\chi(u) \ge \mu(u)\chi(u)^2/2$ and $0 < Z_\chi \le M_\chi \le 2Z_\chi$, so $0 \le r_\chi \le 1$.

Theorem A.5 (Transported ensemble and the sharper leaf selector)**.** *Suppose a finite rational simplex ensemble on the restriction, with positive weights summing to one, has mean energy at most F and mean impurity at most ϖ. Let $W(p)_a = (p_a - \max_{b \ne a} p_b)_+$ and set $g^\sigma = \chi W(p^\sigma)$, extending by zero outside the support. If $\underline{M} = M_\chi - 4 Z_\chi \varpi > 0$, then*

$$\overline{M} \ge \underline{M}, \qquad 0 \le \overline{\mathsf{J}} \le Z_\chi F, \qquad \overline{A} \le \left(\sqrt{2 Z_\chi F - 2\overline{\mathsf{J}}} + \sqrt{\Delta_\chi} \right)^2. \tag{19}$$

Writing $\underline{m} = \underline{M}/Z_\chi$ and $z = \overline{\mathsf{J}}/Z_\chi \in [0, F]$, the least actual seed–threshold charge obeys

$$\Gamma_* \le \frac{z}{\underline{m}} + \frac{2}{\sqrt{\underline{m}}}\left(\sqrt{2(F-z)} + \sqrt{r_\chi}\right). \tag{20}$$

For the ensemble of Lemma A.2, in particular,

$$\Gamma_*^2 \le \frac{520}{27}\kappa_0 + \frac{256}{27}r_\chi. \tag{21}$$

For that constructed ensemble there are at most nq^{2r} seed–threshold pairs, with rational polynomial-time selection per pair. The selected seed need not obey separate energy or impurity thresholds.

Proof. The permutation-equivariant winner-gap map has at most one positive coordinate and is $\sqrt{2}$-Lipschitz: on each ordering region its only nonzero row is the winning coordinate minus the runner-up. Continuity and subdivision of line segments prove the global assertion. If $a \ge b$ are the two largest coordinates of a simplex vector, then $\|p\|^2 \le a^2 + b(1-a)$ and $a + b \le 1$, giving $2\|p\|^2 - (a-b)^2 \le a^2 + 2b - b^2 \le 1$. Therefore

$$\|W(p)\| \le 1, \qquad 1 - \|W(p)\|^2 \le 2(1 - \|p\|^2).$$

Degree domination consequently yields $M(\chi W(p)) \ge M_\chi - 4Z_\chi \sum_u \mu_\chi(u)(1 - \|p_u\|^2)$; average to obtain $\overline{M} \ge \underline{M}$.

We also retain how the winner map allocates its energy. If h, k are the two output gaps and i, j their winners, then

$$(h-k)^2 + 4hk\mathbf{1}_{i\ne j} \le 2\|p-q\|^2.$$

For a zero gap or coincident winners this follows from Lipschitzness. Otherwise $h + k \le (p_i - q_i) - (p_j - q_j) \le \sqrt{2}\|p-q\|$, whose square is the desired left side. Apply this recordwise, including permutation loops. With $d_u = \|W(p_u)\|$, it gives $\Delta_\chi(d) + 2\mathsf{J}_\chi(W(p)) \le 2\widehat{\mathcal{E}}_\chi(p)$. Disagreement transports exactly: $\mathsf{J}(\chi W(p)) = Z_\chi \mathsf{J}_\chi(W(p))$. Thus $\overline{\mathsf{J}} \le Z_\chi F$ and the mean leaf amplitude energy is at most $2F - 2\overline{\mathsf{J}}/Z_\chi$.

For $a \ge b \ge 0$ and $x, y \in [0,1]$, write $ax - by = b(x-y) + (a-b)x$; reverse the roles for $b > a$. The triangle inequality over weighted edge spaces gives

$$\sqrt{\mathcal{E}_P(\chi d)} \le \sqrt{Z_\chi \Delta_\chi(d)} + \sqrt{\Delta_\chi}.$$

Indeed, the squared norm of the first summand is bounded by $\frac{1}{2}\sum_{u,v} a_{uv}\chi(u)\chi(v)(d_u - d_v)^2 = Z_\chi\Delta_\chi(d)$, while $0 \le d \le 1$ bounds the squared norm of the second by Δ_χ. Apply the same triangle inequality in the weighted direct sum over seeds. It proves the amplitude estimate in (19). Lemma A.4 gives $\Gamma_* \le \overline{\mathsf{J}}/\overline{M} + 2\sqrt{\overline{A}/\overline{M}}$; substituting the preceding bounds proves (20).

For the constructed ensemble, $F = 65\kappa_0/64$, $\varpi = 5\kappa_0/\eta$, and $\underline{m} = M_\chi/Z_\chi - 20\kappa_0/\eta \ge 27/32$. Put $f = F/\underline{m}$ and $x = z/\underline{m} \in [0, f]$. Since $\kappa_0 \le 1/512$, we have $f < 2$, and

$$2\sqrt{2f} - 2\sqrt{2(f-x)} = \frac{2\sqrt{2}\,x}{\sqrt{f} + \sqrt{f-x}} \ge x.$$

The zero case follows by continuity. Hence $\Gamma_* \le 2(\sqrt{2F} + \sqrt{r_\chi})/\sqrt{\underline{m}}$. Squaring and using $(a+b)^2 \le 2a^2 + 2b^2$ gives $\Gamma_*^2 \le (16F + 8r_\chi)/\underline{m}$, proving (21). Exact finite selection is supplied by Lemma A.4; no additional moment extension is required. □

Corollary A.6 (The leaf decoder used by the outer algorithm)**.** *Suppose a visited restriction has base cutoff rank at most $R^{5/8}$ at $0 < \eta = 1024\bar{\kappa} \le 1/4$, optimum at least $1 - 8\bar{\kappa}$, and $r_\chi \le 8D$. A deterministic algorithm returns a nonempty partial assignment satisfying*

$$\Gamma^2 \le \frac{65}{432}\eta + \frac{2048}{27}D \tag{22}$$

at cost $q^{O(R^{5/8})}$ times polynomial work in the binary input.

Proof. Apply Lemma A.2 and Theorem A.5 with $\kappa_0 = 8\bar{\kappa}$. Then $128\kappa_0 = \eta$, and substitution in (21) gives the displayed constants. □

The total tree loss D is used only for existence; selection and acceptance use the actual rational partial-label charge.

A.3 Direct multiplier thresholds and a rank-certified menu

There is a complementary candidate: round labels throughout the support of χ and threshold only χ^2. Here we use the invocation root's stronger property that its *identity-loop* mass at u is at least $\mu(u)/2$. No half-laziness of the normalized leaf is assumed. Put

$$m_e = a_e \min\{\chi(u)^2, \chi(v)^2\}, \qquad V_\chi = \tfrac{1}{2}\sum_{u,v} a_{uv}|\chi(u)^2 - \chi(v)^2|.$$

For a simplex field on the support, define the rational quantity

$$\mathsf{Q}_\chi(p) = \sum_e m_e b_e(p), \qquad b_e(p) = \begin{cases} 1 - \sum_a p_u(a)p_v(\pi_e(a)), & u \ne v, \\ \sum_{a:\pi_e(a)\ne a} p_u(a), & u = v. \end{cases} \tag{23}$$

Every min-overlap sum below omits terms with $m_e = 0$, and vertex sums involving marginals are restricted to the support of χ. Thus no labels or probabilities outside that support are required.

Theorem A.7 (Direct multiplier-threshold selection)**.** *For any finite rational simplex ensemble with positive weights summing to one, mean leaf energy at most F and mean impurity at most ϖ, deterministic selection produces a nonempty partial labeling with*

$$\Gamma_{\mathrm{dir}} \le \Theta_\chi := \frac{\min_\sigma \mathsf{Q}_\chi(p^\sigma) + 2V_\chi}{M_\chi} \le \frac{Z_\chi(F + c\varpi) + 2V_\chi}{M_\chi}, \tag{24}$$

$$\Gamma_{\mathrm{dir}} \le \frac{F + c\varpi + 2\sqrt{r_\chi}}{1 + r_\chi}, \qquad c = \begin{cases} 1/2, & \text{all root loops are identities}, \\ 1, & \text{otherwise}. \end{cases} \tag{25}$$

There is no purity-denominator condition. If $\chi = \mathbf{1}_C$, replace $2\sqrt{r_\chi}$ by $2r_\chi$. The cost is polynomial in the full ensemble encoding, with at most n thresholds per seed and no additional moment construction.

Proof. For a labeling ℓ of the support put $\mathcal{I}_\chi(\ell) = \sum_e m_e \mathbf{1}_{\ell(v)\ne\pi_e(\ell(u))}$. Independent labels at distinct vertices give $\mathbb{E}\mathcal{I}_\chi(\ell) = \mathsf{Q}_\chi(p)$. A loop uses one random label, which explains the unary, not bilinear, second line of (23). For each seed, fix vertices successively to labels of positive probability minimizing the remaining conditional expectation. The averaged value equals the previous value, so the final labeling satisfies $\mathcal{I}_\chi(\ell) \le \mathsf{Q}_\chi(p)$. Each step evaluates sums involving at most two remaining probabilities, with polynomial rational bit cost; no inverse-probability sampling is used.

We prove the impurity coefficient, including label-changing loops. Let $C_u = \mathrm{Diag}(p_u) - p_u p_u^\top \succeq 0$ and $\delta_u = \operatorname{Tr} C_u = 1 - \|p_u\|^2$. Write $c_u = \sum_{e:u\to v,\, v\ne u} m_e$ and $A_m = \frac{1}{2}\sum_e m_e \|p_u - T_e p_v\|^2 \le Z_\chi \widehat{\mathcal{E}}_\chi(p)$.

For a non-loop record, its expected violation equals its energy contribution plus $(\delta_u + \delta_v)/2$. For a loop, the difference is $\operatorname{Tr}((I - T_e)C_u)$. Thus symmetry gives the exact identity

$$\mathsf{Q}_\chi(p) = A_m + \sum_u c_u \delta_u + \sum_{e:u\to u} m_e \operatorname{Tr}((I - T_e)C_u).$$

Only the symmetric part of T_e contributes; its operator norm is at most one, so the loop trace is between zero and $2\delta_u$, and vanishes on identity loops. Denote the two root loop masses by l_u^{id} and l_u^{non}. Identity half-laziness gives $l_u^{\mathrm{id}} \ge l_u^{\mathrm{non}} + \sum_{v\neq u} a_{uv}$, whence

$$c_u + 2\chi(u)^2 l_u^{\mathrm{non}} \le \chi(u)^2 (l_u^{\mathrm{id}} + l_u^{\mathrm{non}}) \le Z_\chi \mu_\chi(u).$$

If all loops are identities, the stronger two bounds $c_u \le \chi(u)^2 l_u^{\mathrm{id}}$ and $c_u \le Z_\chi \mu_\chi(u) - \chi(u)^2 l_u^{\mathrm{id}}$ instead give $2c_u \le Z_\chi \mu_\chi(u)$. Consequently, in both cases $\mathsf{Q}_\chi(p) \le Z_\chi(\widehat{\mathcal{E}}_\chi(p) + c\sum_u \mu_\chi(u)\delta_u)$.

The coarea identities of Lemma A.4, with amplitudes $s = \chi$ and labels ℓ, give a multiplier threshold of charge at most $(\mathcal{I}_\chi(\ell) + 2V_\chi)/M_\chi$. Its scalar bound also gives $V_\chi^2 \le \Delta_\chi(M_\chi - \Delta_\chi) = \Delta_\chi Z_\chi$: here $A = \Delta_\chi$ and half-laziness gives $2\Omega - \Delta_\chi \le M_\chi - \Delta_\chi = Z_\chi$. For $\chi = \mathbf{1}_C$, direct computation instead gives $V_\chi = \Delta_\chi$. Enumerate all nonempty multiplier thresholds for every derandomized seed, including the interval below the smallest positive squared multiplier, and retain the least actual charge. This proves the first bound in (24). The minimum of Q_χ is at most its ensemble mean, proving the second. Use $M_\chi = Z_\chi(1 + r_\chi)$ for the final envelopes. All threshold and score comparisons are rational, and their number is polynomial. □

For the ensemble of Lemma A.2, substitute $F = 65\kappa_0/64$ and $\varpi = 5\kappa_0/\eta$. For an integer cap $r_{\max}$ at least the inclusive η-cutoff rank, one can additionally test the rational cutoffs $\eta, 2\eta, 4\eta, \ldots$ below $1/4$, together with $1/4$, keeping only those ζ whose inclusive leaf rank is at most $r_{\max}$. Rational inertia of the generalized matrix $L_\chi - \zeta D_\chi$, counting nonpositive eigenvalues, certifies that rank. Here $D_\chi = \operatorname{diag}(\mu_\chi)$ and $L_\chi = D_\chi - a^\chi$ are the normalized restriction's degree and unnormalized Laplacian matrices. Every retained cutoff satisfies $128\kappa_0 \le \eta \le \zeta \le 1/4$, so the same marginal construction applies with unchanged energy bound and impurity at most $5\kappa_0/\zeta$. The menu has $O(1 + \log(1/\eta))$ entries, polynomial in the binary input; its total cost is $q^{O(r_{\max})}\operatorname{poly}(n, q, b_{\mathrm{in}})$. No larger cutoff is asserted to pass on an arbitrary leaf.

Retain the winner-gap candidates at the original cutoff when adjoining these direct candidates. For use in the outer algorithm set $r_{\max} = \max\{r \in \mathbb{Z}_{\ge 0} : r^8 \le R^5\}$ and apply this cap to every retained cutoff. Keep the original acceptance thresholds and stopping rules. The old guaranteed candidate is still present, and taking the least actual charge preserves all stated approximation and search guarantees. The menu makes only polynomially many additional bounded decoder calls per inspected leaf, each taking $q^{O(R^{5/8})}\operatorname{poly}(n, q, b_{\mathrm{in}})$ time; this preserves the stated global time class, not necessarily the running time at a particular lower-rank leaf. In particular, the direct certificate may improve an observed charge but does not authorize skipping the terminal-mass charge or any invocation test.

B The classical fallback and promise-free recovery

The uniform running-time statement uses the logarithmic approximation of Gupta–Talwar [13]. We give the weighted rational implementation and its deterministic selection step; neither the approximation factor nor the metric-tree method is claimed as new. The construction below does not use the supplied completeness parameter in its computation. The final subsection uses this implementation to remove the supplied completeness promise from the uniform recovery theorem.

Lemma B.1 (Finite weighted tree embedding)**.** *Given a positive rational metric d on $n_{\mathrm{t}} \geq 2$ terminals and rational weights $w_{ab} \geq 0$ on unordered pairs, a deterministic polynomial-time algorithm constructs a rational 2-hierarchically separated tree metric $\mathbb{D}$ such that*

$$\mathbb{D}(a,b) \geq d(a,b), \qquad \sum_{a<b} w_{ab}\mathbb{D}(a,b) \leq 32\log(2n_{\mathrm{t}}) \sum_{a<b} w_{ab} d(a,b).$$

Here outgoing edges at each internal node have equal lengths, and edge lengths decrease by a factor two down the tree. Polynomial time is measured in the full rational input length.

Proof. This is the FRT construction [10, Sections 2.1–2.3 and 2.5], with its finite conditional-expectation implementation made explicit. Scale d by a power of two s so that the minimum positive distance of $\widetilde{d} = \mathsf{s}d$ lies in $(1,2]$, and put $m_{\mathrm{lev}} = \lceil \log_2 \operatorname{diam}(\widetilde{d}) \rceil \geq 1$. Both choices use rational comparisons; m_{lev} is bounded by the input bit length. For $1 < \alpha_{\mathrm{F}} < 2$ and a permutation of all terminals, start from one root cluster at level m_{lev}. At levels $i = m_{\mathrm{lev}} - 1, \ldots, 0$, partition each parent by assigning every terminal in it to the first center at distance less than $r_i = \alpha_{\mathrm{F}} 2^{i-1}$ in $\widetilde{d}$. Centers range over all terminals, also those outside the parent. A child is indexed by its parent and center. Join a level-i child to its parent by an edge of length $2^{i+1}/\mathsf{s}$, keeping a node even when its cluster equals its parent. Level zero consists of singletons, so the tree has at most $1 + n_{\mathrm{t}} m_{\mathrm{lev}}$ nodes.

If (a,b) first separates at level $i < m_{\mathrm{lev}} - 1$, its parent lies in a ball of radius $r_{i+1} < 2^{i+1}$, so $\widetilde{d}(a,b) < 2^{i+2}$, whereas its tree path contains two edges of length $2^{i+1}/\mathsf{s}$. For separation immediately below the root use $\operatorname{diam}(\widetilde{d}) \leq 2^{m_{\mathrm{lev}}}$ instead. This proves domination. Summing the geometric lengths gives $\mathbb{D}(a,b) < 8 \cdot 2^i/\mathsf{s}$. Let ι_i indicate that the first center whose radius-r_i ball contains either endpoint contains exactly one. At first separation $\iota_i = 1$, whence

$$\mathbb{D}(a,b) \leq (8/\mathsf{s}) \sum_{i=0}^{m_{\mathrm{lev}}-1} 2^i \iota_i. \tag{26}$$

For a fixed pair let $a_z = \min\{\widetilde{d}(a,z), \widetilde{d}(b,z)\}$, $b_z = \max\{\widetilde{d}(a,z), \widetilde{d}(b,z)\}$, $A(r) = \{z : a_z < r\}$, and $B(r) = \{z : a_z < r \leq b_z\}$. For a uniform permutation the probability of $\iota_i = 1$ at radius r is $|B(r)|/|A(r)|$. Order centers by increasing a_z, breaking ties arbitrarily. The jth center, when in $B(r)$, has at least j centers in $A(r)$; also $b_z - a_z \leq \widetilde{d}(a,b)$. Hence

$$\int_0^\infty \frac{|B(r)|}{|A(r)|}\, dr \leq \sum_{j=1}^{n_{\mathrm{t}}} \frac{b_j - a_j}{j} \leq H_{n_{\mathrm{t}}} \widetilde{d}(a,b), \qquad H_{n_{\mathrm{t}}} = \sum_{j=1}^{n_{\mathrm{t}}} \frac{1}{j}.$$

Averaging also over uniform $\alpha_{\mathrm{F}} \in (1,2)$ and substituting $r = \alpha_{\mathrm{F}} 2^{i-1}$ gives $\mathbb{E}\sum_i 2^i \iota_i \leq 2H_{n_{\mathrm{t}}} \widetilde{d}(a,b)$. Thus (26) bounds the expected weighted cost by $16 H_{n_{\mathrm{t}}} \sum w_{ab} d(a,b) \leq 32\log(2n_{\mathrm{t}}) \sum w_{ab} d(a,b)$.

For derandomization use the rational estimator

$$\Psi(\alpha_{\mathrm{F}}, p) = (8/\mathsf{s}) \sum_{a<b} w_{ab} \sum_i 2^i \Pr[\iota_i = 1 \mid \alpha_{\mathrm{F}}, p],$$

where p is a permutation prefix. If p has encountered $A(r_i)$, its first such center fixes ι_i; otherwise the conditional probability is exactly $|B(r_i)|/|A(r_i)|$, since none of these centers has been used. All center sets change only at rational breakpoints $\alpha_{\mathrm{F}} = 2^{1-i}\widetilde{d}(a,z)$ in $(1,2)$. Enumerate the $O(n_{\mathrm{t}}^2 m_{\mathrm{lev}})$ breakpoints, adjoin the endpoints $1, 2$, and sort and remove duplicates. Minimize $\Psi(\alpha_{\mathrm{F}}, \varnothing)$ over rational interval midpoints. Since it is constant on each open interval, the minimum is no larger than its continuous average. This includes the single interval $(1,2)$ when there are no interior breakpoints. Then choose each next center to minimize the conditional estimator. Its average

over remaining centers is the current value, so it cannot increase. The final permutation and (26) prove the claim. All enumerations are polynomial; distances, interval midpoints, and dyadic lengths have polynomial bit length. Conditional probabilities have denominators at most n_{t}, so a common denominator for every estimator is a product of input weight denominators, the denominator of $1/\mathsf{s}$, and $\operatorname{lcm}(1,\dots,n_{\mathrm{t}})$, with dyadic factors of $O(m_{\mathrm{lev}})$ bits. Thus comparisons also have polynomial bit cost. □

Lemma B.2 (Deterministic weighted logarithmic approximation)**.** *For a weighted rational Unique Games instance on $n \geq 1$ vertices, there is a deterministic algorithm, polynomial in n, q and the instance bit length, returning a labeling of error at most*

$$\min\{1, 600\kappa \log(2n)\} \quad \textit{whenever} \quad \mathrm{OPT} \geq 1 - \kappa.$$

Parallel constraints, nonidentity loops, disconnected graphs, and arbitrarily small positive rational weights are allowed.

Proof. Discard zero-weight records. If none remain, any labeling suffices; otherwise normalize their weights to sum to one. Each record $e = (u, v)$ has an orientation and permutation π_e from u to v. Use variables $x(u, a), d(e, a) \in [0, 1]$ with

$$\sum_a x(u,a) = 1, \qquad d(e,a) \geq |x(u,a) - x(v,\pi_e(a))|, \qquad L_{\mathrm{LP}} = \frac{1}{2}\sum_e w_e \sum_a d(e,a).$$

The absolute value represents two linear constraints. In the undirected label lift, the edge from (u, a) to $(v, \pi_e(a))$ has length $d(e, a)$. For each lifted root $s = (r, a)$ add potentials $t_s(z) \in [0, 1]$ satisfying

$$t_s(s) = 0, \qquad t_s(r,b) \geq x(r,a) \ (b \neq a), \qquad t_s(j) \leq t_s(i) + d(e,c)$$

in both orientations of every lifted edge i–j. This is a polynomial-size rational LP: it has $O(n^2q^2+mq)$ variables and $O(nmq^2 + n^2q^2)$ constraints for m records. Minimize L_{LP} by exact rational linear programming [12, Theorem 6.4.12]. Every conflict walk W from (r, a) to (r, b), $b \neq a$, obeys

$$x(r,a) \leq \sum_{(e,c)\in W} d(e,c). \tag{27}$$

Every integral labeling is feasible: use one-hot marginals, their absolute differences for d, and truncated shortest-path distances from each lifted root for t_s, taking distance one for unreachable states. A walk from a true label to a false label crosses a length-one edge. Its LP objective is exactly its error, including loops. Thus the rational optimum has $L_{\mathrm{LP}} \leq \kappa$ and polynomial encoding length.

If $L_{\mathrm{LP}} = 0$, every lift length is zero because all remaining weights are positive. In each non-loop connected component choose a root label with positive marginal and propagate it along a spanning tree. The marginal of each propagated label stays positive because the edge marginals agree under transport. Any violated record, including a loop, would form, with the two root paths, a conflict walk moving the positive root label. This contradicts (27). Isolated vertices use their own positive-marginal label. This returns a satisfying labeling.

Assume henceforth $L_{\mathrm{LP}} > 0$. Put $\ell_e = \sum_a d(e,a)$ and, for non-loop records, $\ell'_e = \ell_e + L_{\mathrm{LP}}/(10n)$. Shortest-path completion gives a positive rational metric on each non-loop component and $\sum_{e \text{ non-loop}} w_e\ell'_e \leq 2L_{\mathrm{LP}} + L_{\mathrm{LP}}/(10n)$. Apply Lemma B.1 componentwise, adding weights of parallel records for the embedding only. Eliminate the tree's Steiner vertices as follows. Represent each nested cluster by its least terminal in a fixed order, and contract edges with equal representatives. Every representative's fiber is connected and contains its terminal leaf, so the quotient is a

tree on terminals. Give each surviving edge the original metric distance between its representatives. For an old edge of length L from a parent to a child, each representative is a descendant terminal of the parent and is at tree distance less than $2L$ from it, by geometric decrease. Their original metric distance is thus less than $4L$, by domination. The image of any old terminal-to-terminal path is the unique quotient-tree path after contracted edges are removed. Summing the surviving edge bounds gives $\mathbb{D}_T \le 4\mathbb{D}$, while the triangle inequality for the original metric gives $\mathbb{D}_T \ge d$. Represent each new tree edge by an original shortest path for ℓ', with its composed permutation. Its unperturbed path length is at most its tree length. Singleton components need no embedding.

For a tree edge with transport σ, write $p_a = x(u,a)$, $q_b = x(v,b)$, $m_a = \min\{p_a, q_{\sigma(a)}\}$, $r_a = p_a - m_a$, $s_b = q_b - m_{\sigma^{-1}(b)}$, and $t = \sum_a r_a = \sum_b s_b$. The rational coupling

$$\Upsilon(a,b) = \mathbf{1}_{b=\sigma(a)} m_a + r_a s_b / t$$

uses only its first term at $t = 0$. It has marginals p, q and mismatch probability $t = \frac{1}{2}\sum_a |p_a - q_{\sigma(a)}|$, at most half the represented path length by telescoping. Indeed $r_a s_{\sigma(a)} = 0$. Root each tree, orient the couplings away from the root, and set $\Upsilon_{uv}(a,b) = \Upsilon(a,b)/p_a$ when $p_a > 0$, using any rational probability row when $p_a = 0$. Together with the root marginal these tables define a rational Markov law on labelings, with marginal $x(v,\cdot)$ at every vertex; take the product law across components.

For a non-loop original record $e = (u,v)$, expand the tree path into its representing paths and close it by the reverse of e. For each starting label moved by this closed walk, (27) bounds its marginal by the lifted walk length. Summing over moved labels gives at most $\mathbb{D}_T(u,v) + \ell_e$, since every prefix transport is a permutation. If the label is not moved and all tree edges on the path agree with their transports, the record is satisfied. The union bound on tree mismatches therefore gives $\Pr[e \text{ violated}] \le \ell_e + 2\mathbb{D}_T(u,v)$. For a loop, (27) gives $x(v,a) \le d(e,a)$ whenever $\pi_e(a) \ne a$, so the violation probability is at most ℓ_e. Consequently

$$\mathbb{E}[\mathrm{err}] \le 2L_{\mathrm{LP}} + 256\log(2n)(2L_{\mathrm{LP}} + L_{\mathrm{LP}}/(10n)) \le 600 L_{\mathrm{LP}} \log(2n).$$

For the last inequality, divide by $L_{\mathrm{LP}}\log(2n)$ and use $\log(2n) \ge \log 2 > 1/2$ and $n \ge 1$: the coefficient is at most $4 + 512 + 128/5 < 600$.

Finally derandomize this law exactly. For a partial assignment $\mathcal{F}$, let $\psi_v(b)$ indicate compatibility with its assignment at v. For a nonroot vertex v with parent u compute the upward message

$$M_v(a) = \sum_b \Upsilon_{uv}(a,b)\psi_v(b) \prod_{z \text{ child of } v} M_z(b).$$

For each root r, sum $\mathcal{Z}_r(\mathcal{F}) = \sum_b x(r,b)\psi_r(b)\prod_{z \text{ child of } r} M_z(b)$. Their product is $\mathcal{Z}(\mathcal{F}) = \Pr[\mathsf{X} \text{ extends } \mathcal{F}]$. Inconsistent clamps have probability zero. For $\mathcal{Z}(\mathcal{F}) > 0$ the conditional expected error is

$$\frac{\sum_e w_e \left[\mathcal{Z}(\mathcal{F}) - \sum_a \mathcal{Z}(\mathcal{F}, \mathsf{X}_u = a, \mathsf{X}_v = \pi_e(a))\right]}{\mathcal{Z}(\mathcal{F})}.$$

For each successive vertex, choose a positive-probability label minimizing this expression. The average over those labels is the current value, so it cannot increase. After all choices the actual error is at most the initial expectation, and it is trivially at most one.

Each probability evaluation uses $O(nq^2)$ rational operations, and only polynomially many evaluations are required. A product of all root and transition-table denominators is a common denominator of all complete-label probabilities, so exact messages and conditional comparisons have polynomial bit length even for tiny probabilities. The optimal LP solution, shortest paths, and metric aspect ratio have polynomial rational encoding length; Lemma B.1 therefore also runs in polynomial bit time. No numerical factor $1/\kappa$ or $1/L_{\mathrm{LP}}$ is used. □

B.1 Removing the supplied completeness promise

Proof of Corollary 4.6. We retain the corollary's hypotheses $n \geq 256$ and $q \geq 2$. Discard zero-weight records before any constraint check. First test satisfiability exactly. In every positive-weight non-loop connected component, try each label of one root, propagate it along a spanning tree, and check every record, including loops and parallel constraints. The component is satisfiable precisely when a root label passes all checks. Isolated vertices use the same test on their loops; zero-degree vertices are arbitrary. This is polynomial work. If all components pass, return the resulting satisfying labeling. Zero total weight is also handled this way. Otherwise $\varepsilon_* > 0$.

Run the algorithm of Lemma B.2. It needs no completeness parameter: its perturbation is determined by the rational LP optimum. The proof of that lemma gives an output ℓ_0 with

$$\varepsilon_* \leq \epsilon_0 := \mathrm{err}(\ell_0) \leq \min\{1, C_{\mathrm{GT}}\varepsilon_* \log(2n)\}, \qquad C_{\mathrm{GT}} = 600.$$

In particular $\epsilon_0 > 0$, and it is computed exactly from rational record weights. Set

$$l = \lceil \log_2(2n) \rceil, \qquad J_{\max} = \lceil \log_2(C_{\mathrm{GT}} l) \rceil, \qquad \kappa_j = \epsilon_0/2^j \quad (0 \leq j \leq J_{\max}).$$

Retain only guesses satisfying $1/n \leq \kappa_j < 1$. Run the deterministic algorithm of Theorem 4.4 with each retained guess, with the exhaustion convention specified below. Return the labeling of least actual rational error among these outputs and ℓ_0. No decision compares unknown optimum values or expressions involving real logarithms.

Bounded runs under false guesses. In a deterministic localization invocation, enumerate its prescribed finite set of paths and bounded decoder calls. If none supplies a candidate passing the actual rational charge test, assign the remaining vertices arbitrarily and terminate that guessed run. Successful invocations each remove a nonempty set, so there are at most n of them before completion or this exhaustion exit. The graph construction, rank tests, support reduction, rational entropy estimates, path count, and bounded decoder failure behavior do not require a completeness witness. Only the proof that some candidate passes uses the promise. Thus an incorrect guess cannot force an unbounded search, and it is never assumed to give a good labeling.

These bounded runs have the claimed uniform time even without a witness. Theorem 4.4 invokes localization only when $R > \lceil \log_2 n \rceil^4$; on this branch the unconditional cap $h_U \leq 2 \log n$ absorbs the deterministic path exponent into $O(R^{5/8} \log q)$. Every retained guess has $1/\kappa_j \leq n$, so numerical inverse-completeness work is polynomial in n. On the other branch only the polynomial-time fallback is run. When a guess is valid, the original good-leaf proof guarantees a passing candidate before exhaustion, so the convention does not change Theorem 4.4's guarantee.

A sufficiently accurate valid guess, or the fallback. If $1/n \leq \varepsilon_* < 1/2$, then

$$1 \leq \epsilon_0/\varepsilon_* \leq C_{\mathrm{GT}} \log(2n) \leq C_{\mathrm{GT}} l.$$

The integer $j = \lfloor \log_2(\epsilon_0/\varepsilon_*) \rfloor$, used only in this existence proof, lies in $[0, J_{\max}]$ and satisfies

$$\varepsilon_* \leq \kappa_j < 2\varepsilon_* < 1.$$

It is retained and valid. Writing C_{unif} for the constant in Theorem 4.4, its output has error at most $\sqrt{2} C_{\mathrm{unif}} \sqrt{\varepsilon_* \mathcal{B}_n}$. If $0 < \varepsilon_* < 1/n$, the fallback instead satisfies

$$C_{\mathrm{GT}} \varepsilon_* \log(2n) \leq C_{\mathrm{GT}} \sqrt{\varepsilon_*} \frac{\log(2n)}{\sqrt{n}} \leq C_{\mathrm{GT}} \sqrt{\varepsilon_* \mathcal{B}_n},$$

using $\log(2n) \leq \sqrt{n}$ and $\mathcal{B}_n \geq 1$. For $\varepsilon_* \geq 1/2$, the bound follows from error at most one and $2\sqrt{\varepsilon_* \mathcal{B}_n} \geq 1$. The case $\varepsilon_* = 0$ was already handled by the exact test. Thus $C_{\mathrm{pf}} = \max\{\sqrt{2} C_{\mathrm{unif}}, C_{\mathrm{GT}}, 2\}$ suffices; retaining ℓ_0 gives the linear branch simultaneously.

Finally $J_{\max} = O(1 + \log\log n)$. All guesses have encoding length polynomial in the original input length, since ϵ_0 is an exactly evaluated rational error and division by 2^j adds only j bits. The bounded run costs, number of calls, and exact final comparisons prove the stated running time without an input completeness bound. □

C Whole-tree loss, guarded policies, and certified flat partitions

This optional refinement measures the loss of a complete invocation, not the ratio of one split. Here a policy specifies every node decision and tie rule before its validation samples are drawn. It defines a finite binary tree of nonnegative rational local multipliers with $\chi_0^2 + \chi_1^2 = 1$ and nonempty proper child supports. Each child law is the normalized soft restriction of its parent, and each branch draw uses the child-or-death probabilities of Theorem 3.1. Thus every path has depth at most $n - 1$. Admissibility additionally requires the stated leaf-rank, loss-cap, and uniform polynomial path-access guarantees; these are hypotheses to be verified for each policy, not consequences of its pilot estimates. For a fixed policy, the telescoping calculation of Theorem 3.1 gives

$$D = \sum_{v \text{ internal}} Z_v K_v = 1 - \sum_{b \text{ leaf}} Z_b = \Pr\{\text{an exact path is killed}\}. \tag{28}$$

Pilot paths stop only at death or a terminal leaf; cleanup deaths count. Conditioning on survival or counting decoder failures would estimate a different quantity. These pilots never run a leaf decoder.

C.1 Estimating and comparing complete policies

Lemma C.1 (Bounded rational loss estimation)**.** *Fix a deterministic binary policy with the access guarantees of Theorem 3.1 and a rational cap $0 \le D \le d \le 1/2$. For rational $0 < \epsilon, \delta < 1$, put*

$$b_\delta = \lceil \log_2(4/\delta) \rceil, \qquad N = \left\lceil 16(d + \epsilon/2) b_\delta / \epsilon^2 \right\rceil.$$

There is a bounded procedure using at most N paths which, with probability at least $1 - \delta$, returns rational endpoints satisfying

$$0 \le L_D \le D \le U_D \le d, \qquad U_D - L_D \le \epsilon.$$

It can instead declare failure, included in the exceptional probability. Every nonfailure output, even an exceptional one, has $0 \le L_D \le U_D \le d$ and width at most ϵ. The bit cost is $N \,\mathrm{poly}(n, b_{\mathrm{in}}, 1/\eta, \log(N/\delta))$. For $d = 0$ return $[0, 0]$ without sampling.

Proof. On independent ideal exact paths, death indicators y_j are Bernoulli with mean D. For $|\lambda| \le 1$,

$$\log \mathbb{E} e^{\lambda(\mathsf{y}_j - D)} \le D(e^\lambda - 1 - \lambda) \le D\lambda^2.$$

Here the first inequality is $\log(1 + x) \le x$ and the second follows from the exponential series. Thus either tail of $\widehat{D} = N^{-1} \sum_j \mathsf{y}_j$ at distance s is at most $\exp(-N\lambda s + ND\lambda^2)$ for $0 < \lambda \le 1$. Take $s = \epsilon/2$ and $\lambda = s/(2(d + s))$. The exponent is at most $-Ns^2/(4(d + s)) \le -b_\delta$, so $|\widehat{D} - D| \le \epsilon/2$ except with probability $2e^{-b_\delta} \le \delta/2$. Output

$$U_D = \min\{d, \widehat{D} + \epsilon/2\}, \qquad L_D = \min\{U_D, \max\{0, \widehat{D} - \epsilon/2\}\}.$$

These intervals have the asserted unconditional formatting properties and contain D on the stated deviation event.

For bounded rational sampling, clear the three branch-probability denominators to an integer B. Draw $\lceil \log_2 B \rceil$ fair bits, reject the represented integer when it is at least B, and otherwise use the three integer intervals of lengths Bp_0, Bp_1, BK. Each trial accepts with probability greater than $1/2$ (and probability one when B is a power of two). Cap each draw at $k = \lceil \log_2(2Nn/\delta) \rceil$ trials, accepting a valid final trial, and declare pilot failure if any draw exhausts its cap. At most $N(n-1)$ draws occur. Coupling to ideal draws using the same bits bounds the probability of any discrepancy by $Nn2^{-k} \le \delta/2$. We do not condition on avoiding failure. Combining the two exceptional probabilities proves the confidence claim. The path access bound, polynomial denominator lengths, and rational power comparisons prove the bit bound. □

Proposition C.2 (Guarded comparison of complete policies)**.** *At one fixed root, let $m \ge 1$ admissible deterministic policies with losses $D_1, \dots, D_m$ be fixed, with policy 1 the baseline. Each must independently have the same certified analytic loss cap, rank, partition-of-unity, and polynomial path-access guarantees. Apply Lemma C.1 with scale ϵ and allowance δ/m to each policy. Exactly known losses may instead use singleton intervals. If any pilot declares failure, keep the baseline and its analytic upper cap. Otherwise let j minimize U_j, and choose j only when $U_j \le L_1$; if this fails, retain policy 1. With probability at least $1-\delta$,*

$$D_{\mathrm{sel}} \le D_1, \qquad D_{\mathrm{sel}} \le U_{\mathrm{sel}} \le \min\{D_{\mathrm{sel}} + \epsilon, \min_k D_k + 2\epsilon\}. \tag{29}$$

The cost is the sum of the pilot costs, not an enumeration of all policies.

Proof. With probability at least $1-\delta$, no pilot declares failure and all intervals contain their losses. On this event $U_j \le \min_k D_k + \epsilon$. If the guard accepts, $D_j \le U_j \le L_1 \le D_1$. Otherwise $U_1 \le L_1 + \epsilon < U_j + \epsilon \le \min_k D_k + 2\epsilon$. In both cases the selected interval contains its loss and has width at most ϵ, so $U_{\mathrm{sel}} \le D_{\mathrm{sel}} + \epsilon$. The union bound requires only the validity of each pilot, not independence between different policies. □

C.2 Exactly certified flat partitions

At an admitted residual root of Theorem 4.2, retain its parameters and the rational budget of Lemma 4.3, and define

$$\bar\kappa = \kappa/t, \quad \eta = 1024\bar\kappa, \quad d_U = \min\{1/2, 3\cdot 2^{17}\eta\overline{\mathcal{B}} * U + \kappa/50\}, \qquad A_U = 2^{38}\bar\kappa\overline{\mathcal{B}}U. \tag{30}$$

The component-first construction has $D \le d_U$. Admission gives $\eta\overline{\mathcal{B}}_U \le 2^{-24}$ and $\kappa \le 2^{-34}$, so $d_U < 1/32$. An explicit partition $\mathcal{P} = \{C_1, \dots, C_s\}$ into nonempty sets proposes leaf laws and a complete-tree loss

$$z_j = \sum x, y \in C_j a * xy, \qquad a^j_{xy} = a_{xy}/z_j \ (x, y \in C_j), \qquad D_{\mathcal{P}} = 1 - \sum_j z_j.$$

Accept this proposal only if exact rational inertia verifies $\operatorname{rank} \mathbf{1}_{[0,\eta]}(I - P_j) \le R^{5/8}$ for every part and the rational test $D_{\mathcal{P}} \le d_U$ passes. Explicitly, for each part form $\mathsf{A}_j = (a_{xy})_{x,y\in C_j}$ and $\mathsf{D}_j = \operatorname{diag}(\sum_{y\in C_j} a_{xy})_{x\in C_j}$. The diagonal is positive by the root loops. Count the nonpositive eigenvalues of $(1-\eta)\mathsf{D}_j - \mathsf{A}_j$ by rational inertia and test $r_j^8 \le R^5$. Division by z_j does not change this count.

Proposition C.3 (Certified flat candidates)**.** *Every passing partition is admissible at this invocation. Its loss is known exactly, and its at most n leaves can all be decoded within $nq^{O(R^{5/8})}$ times polynomial bit cost. At least one resulting candidate has actual root charge*

$$\Gamma^2 \le \min\{A_U, a_{\mathrm{rd}}\eta + b_{\mathrm{rd}} D_{\mathcal{P}}\}, \qquad a_{\mathrm{rd}} = 65/432, \qquad b_{\mathrm{rd}} = 2048/27. \tag{31}$$

Proof. Successive hard cuts isolating the parts give a binary squared partition of unity with leaves $\mathbf{1}_{C_j}$ and leaf probabilities $z_j > 0$; positivity follows from the root loops. Its loss is the crossing ordered mass $D_{\mathcal{P}}$, independently of isolation order. The identities in Theorem 3.1 and the proof of Lemma 3.5 apply directly: they require these global identities and the loss cap, not the heat-selection rule. Hence some leaf is good. The checked rank and Corollary A.6 give $\Gamma^2 \le a_{\mathrm{rd}}\eta + b_{\mathrm{rd}} D_{\mathcal{P}}$. This also passes the baseline because, using $\kappa \le \eta/1024$ and $\overline{\mathcal{B}}_U \ge 1$,

$$a_{\mathrm{rd}}\eta + b_{\mathrm{rd}} d_U \le \left(2^{28}/9 + 65/432 + 1/675\right)\eta\overline{\mathcal{B}}_U < 2^{28}\eta\overline{\mathcal{B}}_U = A_U.$$

All leaf decoders are bounded, including on nongood leaves; enumerate them and minimize actual charge. All admission tests are rational. □

Such a flat candidate has a certified invocation-level loss bound; it need not obey the sharper standalone entropy coefficient. It may be included in Proposition C.2 with its exact singleton interval. No claim is made that every proposed partition passes the tests.

C.3 The end-to-end actual-loss guarantee

Theorem C.4 (An actual-loss-sensitive complete-labeling guarantee)**.** *Under Theorem 4.2, at each admitted root fix a finite portfolio containing the component-first baseline. Every other implicit policy must independently satisfy that theorem's conditional-entropy loss cap and the access and rank guarantees above; include only flat proposals passing Proposition C.3. Suppose the portfolio size is polynomially bounded, its proposals are constructed in polynomial time (or supplied as part of the input), and node access has a uniform polynomial bound. There is a bounded randomized implementation with the same running-time class as Theorem 4.2 which, with probability at least $1-\xi$, retains its H_* error guarantee and additionally satisfies*

$$\mathrm{err}(\ell) \le \min\left\{1, 2\left[\sum_i \Delta_i \sqrt{\min\{A_i, a_{\mathrm{rd}}\eta_i + b_{\mathrm{rd}} V_i\}} + t_f\right]\right\}. \tag{32}$$

Here i ranges over accepted invocations on the realized residual trajectory, $t_i = \mu^0(U_i)$, $\Delta_i = t_i - t_{i+1}$, $\eta_i = 1024\kappa/t_i$, A_i is (30), and

$$V_i = \min\left\{D_{i,\mathrm{sel}} + \eta_i, \min_k D_{i,k} + 2\eta_i, \min_{\mathcal{P}} D_{i,\mathcal{P}}\right\}.$$

The last minimum is over that root's passing flat proposals and is $+\infty$ if there are none. Moreover $D_{i,\mathrm{sel}} \le D_{i,1}$ simultaneously. Loss comparisons are at the same realized root, not between different adaptive trajectories. The original tail and budget stops, and terminal mass t_f, are retained.

Proof. At each root freeze its policies and use Proposition C.2 with $\epsilon = \eta$ and total allowance $\delta = \xi/(2n)$. On declared pilot failure use the baseline and $U = d_U$; otherwise use its selected upper endpoint U. Decode all passing flat-partition leaves and retain their least-charge baseline-passing candidate. Write D_{flat} for the least exact flat loss; if there is no flat candidate, omit its threshold from the following minimum. On fresh randomness try at most $J_{\mathrm{path}} = \lceil \log_2(4n/\xi) \rceil$ paths of the selected policy. At each leaf run the bounded decoder and retain the least actual charge among candidates passing $\Gamma^2 \le A_U$. Accept early when

$$\Gamma^2 \le \min\{A_U, a_{\mathrm{rd}}\eta + b_{\mathrm{rd}} U, a_{\mathrm{rd}}\eta + b_{\mathrm{rd}} D_{\mathrm{flat}}\}. \tag{33}$$

The same test is made on retained flat candidates. If no tightened acceptance occurs, use the best retained baseline candidate; if none exists, declare production-search failure and complete arbitrarily

as in the original bounded implementation. Killed paths and leaf-decoder failures only end their attempts.

Every accepted candidate passes A_U, regardless of pilot accuracy. Therefore the deterministic charge, residual-envelope, and terminal-stop arguments of Theorem 4.2 remain valid on every run without production-search failure, giving the same H_* bound. This uses the immutable residual inequalities, not equality of the old and new trajectories. For implicit candidates the required cap is the component-conditional cap; an arbitrary total-entropy policy is not silently assigned this smaller cap. Flat candidates instead have the separate exact certificate above.

Condition on the entire history before each validation. The root and its portfolio are then fixed, so the conditional union bound over at most n roots makes all pilots nonfailing, all intervals covering, and (29) valid except with probability $\xi/2$. For any fixed current pilot history with valid intervals, every good leaf of the selected policy gives a candidate with $\Gamma^2 \le \min\{A_U, a_{\mathrm{rd}}\eta + b_{\mathrm{rd}} D_{\mathrm{sel}}\}$, by Lemma 3.5 and Corollary A.6. If the flat threshold in (33) is smallest, its enumerated candidates already pass. Otherwise every good sampled leaf passes the tightened test, since $D_{\mathrm{sel}} \le U$. Thus J_{path} ideal attempts fail to give tightened acceptance with probability at most $2^{-J_{\mathrm{path}}} \le \xi/(4n)$. Without any pilot accuracy assumption the same argument with only A_U guarantees a baseline-passing candidate.

Cap each production branch draw at $\lceil \log_2(4n^2 J_{\mathrm{path}}/\xi) \rceil$ rejection trials, treating cap failure as a failed attempt. There are at most $n^2 J_{\mathrm{path}}$ draws. Coupling to fresh ideal paths bounds any cap discrepancy by $\xi/4$, and the ideal search failures over all roots cost at most another $\xi/4$. Together with validation this gives total exceptional probability at most ξ. This argument conditions on history before the current samples, not on successful rejection draws or an optional stopping event.

On its complement every acceptance is tightened. The two upper bounds on U in (29), and the exact flat threshold, give the expression in (32). Charge accounting yields $\mathrm{err} \le \min\{1, 2(\sum_i \Delta_i \Gamma_i + t_f)\}$; the factor two undoes half-lazification. This proves the claimed bound.

At one root, every implicit policy needs $O((d_U + \eta)\eta^{-2} \log(nm/\xi)) = O(\overline{\mathcal{B}}_U \eta^{-1} \log(nm/\xi))$ pilot paths. Here $\kappa \le \eta/1024$, $\overline{\mathcal{B}}_U \ge 1$, and $1/\eta \le 1/(1024\kappa)$. Their costs are polynomial in the parameters already allowed in Theorem 4.2; they require no leaf decoders. The at most polynomially many flat leaves and nJ_{path} production leaves contribute only polynomially many bounded $q^{O(R^{5/8})}$ calls. All counts, means, tests and probability tables have polynomial bit length. This proves the same bounded running-time class. □

For a single implicit policy without flat proposals, the new bound uses $(a_{\mathrm{rd}} + b_{\mathrm{rd}})\eta_i + b_{\mathrm{rd}} D_i$. In particular it implies

$$\mathrm{err}(\ell) \le \min\left\{1, 128\sqrt{(a_{\mathrm{rd}} + b_{\mathrm{rd}})\kappa} + 2\sqrt{b_{\mathrm{rd}} \sum_i \Delta_i D_i} + 2t_f\right\}.$$

Indeed $\sum_i \Delta_i/\sqrt{t_i} \le \int_0^1 t^{-1/2}\,dt = 2$, and Cauchy–Schwarz gives $\sum_i \Delta_i \sqrt{D_i} \le \sqrt{\sum_i \Delta_i D_i}$. This bound is trajectory-sensitive; the displayed losses need not be uniformly small. The exactly recorded actual charges always certify $2(\sum_i \Delta_i \Gamma_i + t_f)$ on successful runs; confidence is needed to relate pilot bounds to true losses and to obtain efficient tightened acceptance. Guarding complete policies does not assert that greedily decreasing a local K/I decreases whole-tree loss. Policies modified after viewing validation data require fresh validation. These pilot-enhanced guarantees are randomized; deterministic exhaustive path search alone does not make their confidence statements deterministic.

Use of AI: The author prepared a two-page draft (without using AI) on treating spectral decomposition as an information-budgeted process. The author fed this draft to multiple AI tools (Claude, ChatGPT and two open-weight models with open-source harnesses) and prompted them to apply it to improve the best known algorithms for Unique Games, work out the details of the proofs and identify missing gaps. The author fixed those missing gaps and iteratively verified and expanded the proofs using AI tools. The author is solely responsible for the correctness of the proofs.

References

[1] V. L. Alev and L. C. Lau. *Approximating Unique Games Using Low Diameter Graph Decomposition*. APPROX/RANDOM 2017, LIPIcs 81, Article 18, pp. 18:1–18:15. Publisher version. https://arxiv.org/abs/1702.06969.

[2] S. Arora, B. Barak, and D. Steurer. *Subexponential Algorithms for Unique Games and Related Problems*. FOCS 2010, pp. 563–572; journal version, *Journal of the ACM* 62(5), Article 42 (2015). doi:10.1145/2775105. Author preprint, 8 April 2010.

[3] S. Arora, R. Impagliazzo, W. Matthews, and D. Steurer. *Improved Algorithms for Unique Games via Divide and Conquer*. ECCC Report TR10-041, 2010. https://eccc.weizmann.ac.il/report/2010/041/.

[4] S. Arora, S. Khot, A. Kolla, D. Steurer, M. Tulsiani, and N. K. Vishnoi. *Unique Games on Expanding Constraint Graphs Are Easy*. STOC 2008, pp. 21–28. doi:10.1145/1374376.1374380. Author version.

[5] M. Bafna, B. Barak, P. K. Kothari, T. Schramm, and D. Steurer. *Playing Unique Games on Certified Small-Set Expanders*. STOC 2021, pp. 1629–1642; arXiv:2006.09969. Theorem 1.1 in the full version. https://arxiv.org/abs/2006.09969.

[6] M. Bafna and D. Minzer. *Solving Unique Games over Globally Hypercontractive Graphs*. CCC 2024, LIPIcs 300, Article 3, pp. 3:1–3:15. doi:10.4230/LIPIcs.CCC.2024.3. Full version: https://arxiv.org/abs/2304.07284.

[7] B. Barak, P. Raghavendra, and D. Steurer. *Rounding Semidefinite Programming Hierarchies via Global Correlation*. FOCS 2011; arXiv:1104.4680. https://arxiv.org/abs/1104.4680.

[8] M. Charikar, K. Makarychev, and Y. Makarychev. *Near-Optimal Algorithms for Unique Games*. STOC 2006, pp. 205–214. Theorem 4.5 in the author version. https://konstantin.makarychev.net/pdf/unique.pdf.

[9] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Second edition, Wiley, 2006. Chapter 2. doi:10.1002/047174882X.

[10] J. Fakcharoenphol, S. Rao, and K. Talwar. *A Tight Bound on Approximating Arbitrary Metrics by Tree Metrics*. STOC 2003, especially Section 2.5. https://www.cs.cmu.edu/~guyb/paralg/papers/FakcharoenpholRaoTalwar03.pdf.

[11] S. Goel, R. Montenegro, and P. Tetali. *Mixing Time Bounds via the Spectral Profile*. *Electronic Journal of Probability* 11(1), 2006, pp. 1–26. doi:10.1214/EJP.v11-300.

[12] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Algorithms and Combinatorics 2, Springer, 1988. Sections 3.2–3.3; Theorem 6.4.12. Author-hosted first-edition PDF.

[13] A. Gupta and K. Talwar. *Approximating Unique Games*. SODA 2006, Sections 2–3. https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/19-p99-gupta.pdf.

[14] V. Guruswami and A. K. Sinop. *Faster SDP Hierarchy Solvers for Local Rounding Algorithms*. Expanded version of FOCS 2012, arXiv:1207.4372v2, 2012. Sections 2–5. https://arxiv.org/abs/1207.4372.

[15] V. Guruswami and A. K. Sinop. *Rounding Lasserre SDPs using Column Selection and Spectrum-Based Approximation Schemes for Graph Partitioning and Quadratic IPs*. Merged full version, arXiv:1312.3024, 2013. Theorems 5.2, 6.7, and 6.9; Remark 6.8. https://arxiv.org/abs/1312.3024.

[16] S. B. Hopkins, T. Schramm, and L. Trevisan. *Subexponential LPs Approximate Max-Cut*. FOCS 2020; arXiv:1911.10304. Theorem 4.5 in the full version. https://arxiv.org/abs/1911.10304.

[17] S. Khot. *On the Power of Unique 2-Prover 1-Round Games*. STOC 2002, pp. 767–775. doi:10.1145/509907.510017.

[18] A. Kolla. *Spectral Algorithms for Unique Games*. *Computational Complexity* 20 (2011), pp. 177–206. https://arxiv.org/abs/1102.2300.

[19] K. Makarychev and Y. Makarychev. *How to Play Unique Games on Expanders*. arXiv:0903.0367, 2009. https://arxiv.org/abs/0903.0367.

[20] R. O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014; author-posted edition updated 2021. Section 10.6. https://www.cs.cmu.edu/~odonnell/papers/Analysis-of-Boolean-Functions-by-Ryan-ODonnell.pdf.

[21] D. Steurer. *On the Complexity of Unique Games and Graph Expansion*. PhD thesis, Princeton University, 2010. Chapter 5, especially Theorem 5.2, pp. 65 and 77–78. https://www.dsteurer.org/paper/thesis.pdf.

[22] L. Trevisan. *Approximation Algorithms for Unique Games*. *Theory of Computing* 4(5), 2008, pp. 111–128. doi:10.4086/toc.2008.v004a005.