OBSERVE THEN CONTROL: PID-INSPIRED DISCRETE LANGEVIN SEARCH FOR COMBINATORIAL OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Discrete Langevin proposals provide a training-free approach to combinatorial optimization by converting local energy information into parallel bit flips. We study how trajectory feedback can regulate such proposals under a prescribed expected flip budget. We introduce **PID-LD**, a hierarchical closed-loop discrete Langevin search framework for finite-horizon binary optimization. Bounded stochastic momentum provides a stateful dynamical backbone, while PID-inspired feedback regulates this backbone and the resulting proposal distribution: the proportional channel forms the instantaneous oriented drive, the derivative channel adjusts damping and velocity coupling using gradient-reversal observations, and the integral channel adjusts temperature and noise using accumulated stagnation and past flip-rate information. Calibrated Bernoulli centering then maps the controlled score to bit-flip probabilities while preserving the prescribed expected flip budget. We characterize this calibration as entropy-regularized score allocation and derive a conditional one-step energy bound, an oracle comparison, and mechanism-level results under explicit assumptions. Experiments on maximum independent set, maximum clique, and maximum cut show instance-dependent quality–runtime trade-offs. Ablations further distinguish the role of the stateful momentum backbone from the incremental role of closed-loop feedback on top of that backbone. Together, the formulation and analysis provide a structured view of how stateful dynamics, trajectory feedback, and budget-preserving proposal calibration interact in adaptive discrete Langevin search.

1 INTRODUCTION

Combinatorial optimization (CO) seeks high-quality feasible solutions over discrete spaces and underlies applications in logistics, scheduling, and graph optimization [16; 20; 7]. Sampling-based methods explore these spaces through an energy function, as in simulated annealing [14]. Locally informed proposals, Gibbs-with-Gradients, and discrete Langevin proposals use local density or gradient information to guide exploration [25; 11; 27]. Recent training-free solvers combine such proposals with parallel computation and explicit control of update magnitudes [21; 10; 23]. A prescribed move budget does not determine how that budget should be allocated throughout a search trajectory. Coordinate-scale differences can concentrate proposal mass on a subset of variables; momentum can interact with repeated reversals; and a peaked proposal can revisit similar moves after progress stalls. These observations motivate our research question: **how can trajectory feedback regulate a discrete Langevin-style proposal under a fixed expected flip budget, and what can be established about the resulting transition?** We address this question through a PID-inspired controller acting on the proposal construction, rather than a continuous position update.

PID control in optimization and Langevin sampling is established prior work [2; 9; 6; 5], and expected move-size regularization is also present in RLD/RLSA [10]. Our contribution is therefore not the invention of feedback control or expected move-size regulation in isolation. Instead, we organize them into a hierarchical closed-loop search architecture for finite-horizon binary optimization. PID-LD treats discrete Langevin optimization as a controlled dynamical search process. Bounded stochastic momentum provides a stateful backbone that carries search history across iterations. A PID-inspired control layer observes the evolving trajectory and regulates this backbone and its proposal: the proportional (P) channel supplies the instantaneous gradient-informed drive, the derivative

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

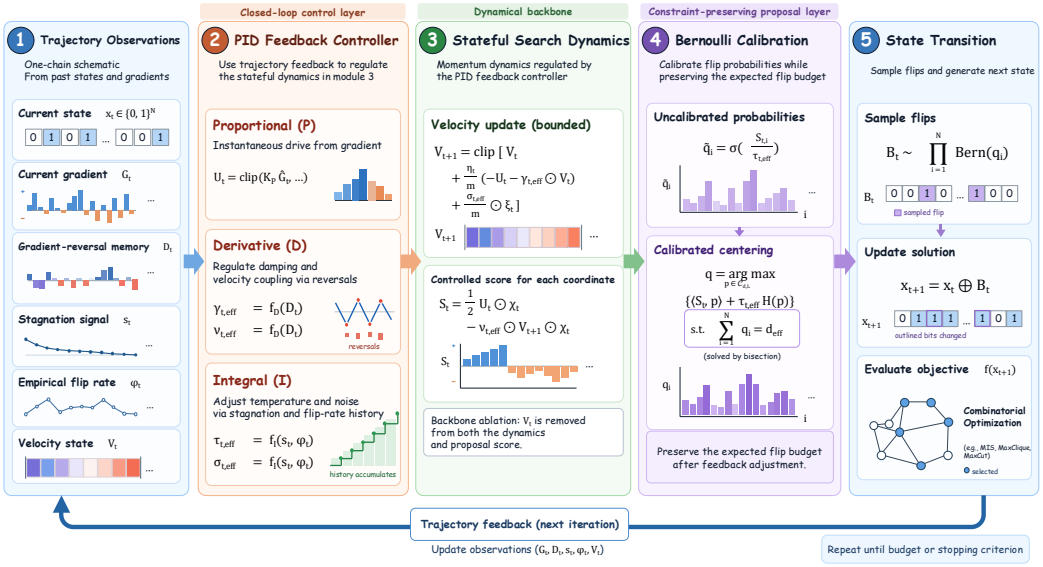


Figure 1: Overview of PID-LD as a hierarchical closed-loop search architecture.

(D) channel responds to gradient reversals by adapting damping and velocity coupling, and the integral (I) channel responds to accumulated stagnation and past empirical flip-rate observations by adapting temperature and noise. Finally, calibrated Bernoulli centering converts the controlled score into factorized bit-flip probabilities while preserving the prescribed expected flip budget. Hence, feedback changes how probability mass is allocated and evolves over time without directly changing the expected number of proposed flips.

Our contributions are threefold. **1)** We formulate PID-LD as a hierarchical closed-loop search architecture in which bounded stochastic momentum provides the stateful dynamical backbone, PID-inspired trajectory feedback forms the control layer, and calibrated Bernoulli centering serves as a constraint-preserving proposal layer. This organization separates the roles of search memory, adaptive regulation, and expected move-budget control. **2)** We characterize calibrated centering as an entropy-regularized score-allocation problem and analyze the resulting finite-budget transition. Our results include conditional one-step progress bounds, an oracle comparison, interaction-aware analysis of multi-bit moves, and mechanism-level properties of the D- and I-channels under explicit assumptions. **3)** We evaluate PID-LD on MIS, MaxCl, and MaxCut, reporting solution quality, runtime, and controlled ablations. In particular, we distinguish removal of the momentum backbone from controller ablations that retain the same stateful dynamics, thereby measuring the incremental effect of trajectory feedback without interpreting structurally different removals as an additive ranking of component importance. The experiments show setting-dependent quality–runtime trade-offs, and matched-iteration comparisons are not interpreted as matched-time superiority.

Figure 1 summarizes this organization. A key distinction is that momentum and trajectory feedback are not treated as parallel, interchangeable components. Momentum defines the stateful search dynamics, whereas the PID-inspired controller observes and regulates these dynamics and the resulting proposal allocation. Bernoulli calibration then preserves the expected move budget after the control action.

2 Methodology

2.1 Problem setup and architectural view

We consider binary combinatorial optimization problems of the form

$$\min_{x \in \mathcal{X}=\{0,1\}^N} a(x), \text{ s.t. } b(x) = 0, \tag{1}$$

and optimize the penalized energy

$$H(x) = a(x) + \lambda \Phi(b(x)), \tag{2}$$

where $\lambda > 0$, $\Phi(b) \geq 0$, and $\Phi(b) = 0$ if and only if the constraint is satisfied. The corresponding Gibbs distribution

$$\pi_\tau(x) = \frac{\exp(-H(x)/\tau)}{\sum_{x' \in \mathcal{X}} \exp(-H(x')/\tau)} \tag{3}$$

motivates a gradient-informed discrete proposal. PID-LD is nevertheless a finite-horizon optimizer with a best-so-far archive, rather than an exact sampler from π_τ. In particular, improvement in penalized energy does not by itself certify feasibility of the returned task solution.

Architectural view. PID-LD is organized hierarchically rather than as an additive collection of independent components (Figure 1). Stateful stochastic momentum provides the dynamical backbone that carries information across iterations. A PID-inspired control layer closes the feedback loop around this search process: the P-channel supplies the instantaneous gradient-informed drive, the D-channel regulates the momentum pathway through damping and velocity coupling, and the I-channel regulates exploration through temperature and noise. Calibrated Bernoulli centering then maps the controlled score to a binary transition while preserving a prescribed expected flip budget. Thus, feedback changes the allocation and temporal evolution of proposal probabilities, whereas calibration prevents these adaptations from unintentionally changing the expected move size.

PID-LD retains the factorized flip geometry of discrete Langevin proposals (DLPs) [27], but augments it with this stateful closed-loop architecture. The DLP derivation and the complete controller state are given in Appendices C.7 and C, respectively.

2.2 PID-CONTROLLED MOMENTUM DYNAMICS

For a binary state x, define

$$\chi_i(x) = 2x_i - 1, \; F_i x = x + (1 - 2x_i)e_i, \tag{4}$$

where $F_i x$ flips coordinate i. The exact one-bit energy gain is

$$g_i(x) = H(x) - H(F_i x). \tag{5}$$

Let $\widetilde{H} : [0,1]^N \to \mathbb{R}$ denote the multilinear extension of H. At every binary vertex,

$$g_i(x) = \chi_i(x)\partial_i \widetilde{H}(x). \tag{6}$$

Hence the oriented gradient $\chi_i(x)G_i(x)$, with $G(x) \approx \nabla \widetilde{H}(x)$, provides a natural local surrogate for one-bit improvement. The complete discrete-difference derivation is deferred to Appendix C.7.

PID-LD evolves K parallel binary chains together with auxiliary velocity states,

$$X_t = \left(x_t^{(1)}, \cdots, x_t^{(K)}\right) \in \{0,1\}^{N \times K}, \; V_t = \left(v_t^{(1)}, \cdots, v_t^{(K)}\right) \in \mathbb{R}^{N \times K}, \tag{7}$$

and we write

$$\chi_t = 2X_t - 1. \tag{8}$$

At iteration t, the energy-gradient oracle returns

$$(E_t, G_t) = \mathcal{EG}(X_t), \; E_t^{(k)} = H(x_t^{(k)}). \tag{9}$$

A diagonal geometry operator $\mathcal{P}_t$ optionally normalizes coordinate scales,

$$\widehat{G}_t = \mathcal{P}_t(G_t), \tag{10}$$

with the identity, RMSProp-style, and structural choices specified in Appendix C.2.

P-channel: instantaneous drive. The proportional channel converts the current normalized gradient into a bounded drive,

$$U_t = \text{clip}\left(K_P \widehat{G}_t, -U_{\max}, U_{\max}\right). \tag{11}$$

D-channel: reversal-aware regulation. For each chain, the D-channel summarizes sign reversals and changes in normalized, filtered gradient observations into a bounded statistic $D_t^{(k)}$. Its complete construction, including filtering and the reversal gate, is given in Appendix C. The statistic acts on the momentum pathway through

$$\begin{aligned}\gamma_{t,\mathrm{eff}}^{(k)} &= \mathrm{clip}\left(\gamma_0\left[1 + K_D a_\gamma D_t^{(k)}\right], \gamma_{\min}, \gamma_{\max}\right),\\ \nu_{t,\mathrm{eff}}^{(k)} &= \frac{\nu_0}{\max\left\{1 + K_D a_\nu D_t^{(k)}, \varepsilon_\nu\right\}}.\end{aligned} \tag{12}$$

For nonnegative gains, a larger reversal statistic therefore increases friction and reduces velocity–score coupling. This actuator-level direction does not by itself imply a smaller realized velocity or lower undo probability; the corresponding mechanism result requires the score-ordering condition in Appendix D.

I-channel: accumulated trajectory response. The I-channel integrates stagnation and the past empirical flip-rate history into a bounded state $I_t^{(k)}$. Importantly, it uses only information available before the current proposal is sampled. The detailed stagnation counter, flip-rate deviation, leakage, and burn-in definitions are given in Appendix C. Its actuators are

$$\begin{aligned}\tau_{t,\mathrm{eff}}^{(k)} &= \tau_t \exp\left(\mathrm{clip}(K_I a_\tau I_t^{(k)}, -c_\tau, c_\tau)\right),\\ \sigma_{t,\mathrm{eff}}^{(k)} &= \sigma_0 + K_I a_\sigma \left[I_t^{(k)}\right]_+.\end{aligned} \tag{13}$$

Thus the I-channel regulates exploration through temperature and stochastic forcing without directly changing the prescribed expected flip budget.

Given a proposal perturbation

$$\xi_t \sim \mu_t(\cdot \mid \mathcal{Y}_t), \tag{14}$$

the bounded stochastic momentum backbone evolves as

$$V_{t+1} = \mathrm{clip}\left[V_t + \frac{\eta_t}{m}\left(-U_t - \gamma_{t,\mathrm{eff}} \odot V_t\right) + \frac{\sigma_{t,\mathrm{eff}}}{m} \odot \xi_t, -V_{\max}, V_{\max}\right], \tag{15}$$

where $m > 0$ is the mass parameter and η_t is the integration step.

Because X_t is binary, V_{t+1} is not used as a continuous position update. Instead, it contributes to the score assigned to feasible bit-flip directions:

$$S_t = \mathrm{clip}\left[\frac{1}{2}U_t \odot \chi_t - \nu_{t,\mathrm{eff}} \odot V_{t+1} \odot \chi_t, -S_{\max}, S_{\max}\right]. \tag{16}$$

The negative velocity term follows because the feasible flip direction for coordinate i is $-\chi_i(x)e_i$.

Equations 15–16 make the architectural distinction explicit: momentum supplies the stateful search dynamics, whereas the controller regulates the forces, damping, coupling, temperature, and noise governing those dynamics and their conversion into proposal scores. In particular, if $\nu_0 = 0$, the score no longer depends on velocity, so velocity-side D/I actuation cannot influence the binary proposal through this pathway.

2.3 Budget-preserving Bernoulli transition

The controlled score is converted into independent bit-flip probabilities by calibrated centering. For each chain k, PID-LD chooses a scalar $\kappa_t^{(k)}$ such that

$$q_{i,t}^{(k)} = \sigma\left(\mathrm{clip}\left(\frac{S_{i,t}^{(k)} - \kappa_t^{(k)}}{\tau_{t,\mathrm{eff}}^{(k)}}, -L_{\max}, L_{\max}\right)\right), \ \sum_{i=1}^{N} q_{i,t}^{(k)} = d_{\mathrm{eff}}. \tag{17}$$

Let $\epsilon_L = \sigma(-L_{\max})$. An exact center exists whenever

$$d_{\mathrm{eff}} \in \left[N\epsilon_L,\ N(1-\epsilon_L)\right]. \tag{18}$$

Algorithm 1 PID-LD: observe, control, calibrate, and sample

Require: Energy-gradient oracle $\mathcal{EG}$, initial population X_0, horizon T, expected flip budget d_{eff}.
Ensure: Best-so-far penalized-energy state x_{best}.
1: Initialize momentum, controller memories, flip-rate memory, and archives.
2: **for** $t = 0, \dots, T-1$ **do**
3: **Observe:** $(E_t, G_t) \leftarrow \mathcal{EG}(X_t)$; form $\widehat{G}_t$ and proportional drive U_t.
4: Update D/I trajectory memories using the previous archive and past flip-rate observations; then update the archive with (X_t, E_t).
5: **Control:** form $\gamma_{t,\text{eff}}, \nu_{t,\text{eff}}, \tau_{t,\text{eff}}, \sigma_{t,\text{eff}}$.
6: **Momentum:** sample ξ_t and update V_{t+1} by Eq. 15; form the controlled score S_t.
7: **Calibrate:** solve for κ_t and obtain q_t such that $\sum_i q_{i,t} = d_{\text{eff}}$.
8: **Sample:** draw $B_{i,t} \sim \text{Bernoulli}(q_{i,t})$ and set $X_{t+1} = X_t \oplus B_t$.
9: Update only the flip-rate memory from the sampled mask B_t.
10: **end for**
11: Evaluate X_T and return x_{best}.

If the requested budget lies outside this interval, it is projected onto the feasible range. In implementation, $\kappa_t^{(k)}$ is obtained by bisection; the numerical solver and tolerance are specified in Appendix C.5.

The resulting constraint

$$\sum_{i=1}^{N} q_{i,t}^{(k)} = d_{\text{eff}} \tag{19}$$

is central to the separation between control and move size. Changes in temperature, noise, damping, or velocity alter how the probability mass is distributed across coordinates, but recalibration restores the same expected flip budget.

Conditional on the pre-proposal state $\mathcal{Y}_t$ and proposal noise ξ_t, coordinates are sampled independently:

$$B_{i,t}^{(k)} \mid \mathcal{Y}_t, \xi_t \sim \text{Bernoulli}\left(q_{i,t}^{(k)}\right), \; X_{i,t+1}^{(k)} = X_{i,t}^{(k)} \oplus B_{i,t}^{(k)}. \tag{20}$$

Only after sampling is the empirical flip-rate memory updated:

$$\widehat{\phi}_t^{(k)} = \frac{1}{N}\sum_{i=1}^{N} B_{i,t}^{(k)}, \; \overline{\phi}_{t+1}^{(k)} = \beta_\phi \overline{\phi}_t^{(k)} + (1-\beta_\phi)\widehat{\phi}_t^{(k)}. \tag{21}$$

Under exact feasible calibration,

$$\mathbb{E}\left[\widehat{\phi}_t^{(k)} \mid \mathcal{Y}_t, \xi_t\right] = \frac{1}{N}\sum_i q_{i,t}^{(k)} = \frac{d_{\text{eff}}}{N}. \tag{22}$$

Hence d_{eff} controls the expected, not realized, number of flips. The past flip-rate input to the I-channel is therefore a trajectory observation rather than an independent actuator for the current expected move size.

The resulting causal order is

observe $\rightarrow$ control $\rightarrow$ momentum update $\rightarrow$ score $\rightarrow$ calibrate $\rightarrow$ sample $\rightarrow$ feedback.

Detailed controller memories and causal ordering are given in Appendix C, while Appendix C.8 gives the conditional and noise-marginalized transition kernels.

2.4 Finite-budget analysis

We next characterize the implemented Bernoulli transition. The analysis separates two issues: interactions among simultaneously flipped coordinates, and the allocation of a fixed probability budget across coordinates.

For any $S \subseteq [N]$, let x^S denote the state obtained by flipping all coordinates in S.

Theorem 1 (One-bit residual bound for independent multi-bit moves)**.** *Let $g_i(x) = H(x) - H(F_i x)$. Suppose there exist nonnegative symmetric interaction constants $\epsilon_{ij}(x)$ such that, for every i and every $R \subseteq [N] \setminus \{i\}$,*

$$\left|g_i(x^R) - g_i(x)\right| \le \sum_{j \in R} \epsilon_{ij}(x). \tag{23}$$

If $B_i \sim \mathrm{Bernoulli}(q_i)$ independently, then

$$\mathbb{E}\left[H(x) - H(x^B)\right] \ge \sum_{i=1}^{N} q_i g_i(x) - \sum_{1 \le i < j \le N} \epsilon_{ij}(x) q_i q_j. \tag{24}$$

The proof is given in Appendix D.1. The result separates the contribution of favorable one-bit gains from a conservative penalty for simultaneous-flip interactions. For the pairwise pseudo-Boolean energies used in our experiments, these interaction terms are explicit and an exact signed expression is available in Appendix D.8.

To characterize calibration, define

$$\mathcal{C}_{d,L} = \left\{ p \in [\epsilon_L, 1 - \epsilon_L]^N : \sum_{i=1}^{N} p_i = d \right\}, \tag{25}$$

and the Bernoulli entropy

$$\mathcal{H}(p) = \sum_{i=1}^{N} \left[-p_i \log p_i - (1 - p_i) \log(1 - p_i) \right]. \tag{26}$$

Lemma 1 (Variational form of calibrated Bernoulli centering)**.** *For a fixed score $S \in \mathbb{R}^N$, temperature $\tau > 0$, and feasible budget d, the calibrated probability vector in Eq. 17 is the unique solution of*

$$q = \arg\max_{p \in \mathcal{C}_{d,L}} \left\{ \langle S, p \rangle + \tau \mathcal{H}(p) \right\}. \tag{27}$$

Thus calibration is not merely a numerical normalization step. It solves a fixed-budget score-allocation problem in which the current controlled score favors promising coordinates and Bernoulli entropy regularizes how sharply the budget is concentrated. The proof is given in Appendix D.

Define the best clipped fixed-budget one-bit gain

$$G_{d,L}(x) = \max_{p \in \mathcal{C}_{d,L}} \langle g(x), p \rangle. \tag{28}$$

Theorem 2 (A posteriori finite-budget descent certificate)**.** *Assume the residual condition in Theorem 1. Let $q \in \mathcal{C}_{d,L}$ be the calibrated PID-LD probability vector generated from score S and temperature τ. Then*

$$\mathbb{E}\left[H(x) - H(x^B) \mid \mathcal{Y}_t, \xi_t\right] \ge \langle g(x), q \rangle - \sum_{1 \le i < j \le N} \epsilon_{ij}(x) q_i q_j. \tag{29}$$

Moreover, let

$$p^\star \in \arg\max_{p \in \mathcal{C}_{d,L}} \langle g(x), p \rangle, \; G_{d,L}(x) = \langle g(x), p^\star \rangle. \tag{30}$$

For any $\alpha > 0$, define

$$\Gamma_\alpha(x, S; q, p^\star) = \langle S - \alpha g(x), q - p^\star \rangle + \tau \left(\mathcal{H}(q) - \mathcal{H}(p^\star) \right). \tag{31}$$

Then

$$\mathbb{E}\left[H(x) - H(x^B) \mid \mathcal{Y}_t, \xi_t\right] \ge G_{d,L}(x) - \frac{\Gamma_\alpha(x, S; q, p^\star)}{\alpha} - \sum_{1 \le i < j \le N} \epsilon_{ij}(x) q_i q_j. \tag{32}$$

Table 1: MIS objectives and reported total testing times. Bold objectives denote the best listed non-OR result; no equal-time or optimality claim is implied.

MIS		RB-[200–300]		RB-[800–1200]		ER-[700–800]		ER-[9000–11000]	
Method	**Type**	**Size ↑**	**Time ↓**	**Size ↑**	**Time ↓**	**Size ↑**	**Time ↓**	**Size ↑**	**Time ↓**
Gurobi	OR	19.98	47.57 m	40.90	2.17 h	41.38	50.00 m	—	—
KaMIS	OR	20.10	1.40 h	43.15	2.05 h	44.87	52.13 m	381.31	7.60 h
DGL [4]	SL	17.36	12.78 m	34.50	23.90 m	37.26	22.71 m	—	—
INTEL [15]	SL	18.47	13.07 m	34.47	20.28 m	34.86	6.06 m	284.63	5.02 m
DIFUSCO [22]	SL	18.52	16.05 m	—	—	41.12	26.67 m	—	—
LTFT [26]	UL	19.18	32 s	37.48	4.37 m	—	—	—	—
DiffUCO [18]	UL	19.24	54 s	38.87	4.95 m	—	—	—	—
SDDS [19]	UL	19.62	20 s	39.99	6.35 m	—	—	—	—
PPO [1]	RL	19.01	1.28 m	32.32	7.55 m	—	—	—	—
DIMES [17]	RL	—	—	—	—	42.06	12.01 m	332.80	12.72 m
RLNN [10]	PRL	19.52	1.64 m	38.46	6.24 m	43.34	1.37 m	363.34	11.76 m
FLD-IG [23]	PRL	19.72	1.08 m	39.56	6.31 m	43.50	1.35 m	365.03	11.41 m
iSCO [21]	H	19.29	2.71 m	36.96	11.26 m	42.18	1.45 m	365.37	1.10 h
RLSA [10]	H	19.97	35 s	40.19	1.85 m	44.10	20 s	375.31	1.66 m
FLD-EG [23]	H	20.02	38 s	40.25	1.93 m	**44.37**	19 s	377.50	1.12 m
PID-LD (ours)	H	**20.08**	41 s	**40.78**	2.05 m	44.32	22 s	**380.50**	1.32 m

Table 2: MaxCl and MaxCut objectives and reported total testing times. Bold objectives denote the best listed non-OR result.

MaxCl		RB-[200–300]		RB-[800–1200]		MaxCut		BA-[200–300]		BA-[800–1200]	
Method	**Type**	**Size ↑**	**Time ↓**	**Size ↑**	**Time ↓**	**Method**	**Type**	**Size ↑**	**Time ↓**	**Size ↑**	**Time ↓**
Gurobi	OR	19.05	1.92 m	33.89	19.67 m	Gurobi	OR	730.87	8.50 m	2944.38	1.28 h
ERDOES	UL	12.02	41 s	25.43	2.27 m	ERDOES	UL	693.45	46 s	2870.34	2.82 m
LTFT	UL	16.24	42 s	31.42	4.83 m	LTFT	UL	704.30	2.95 m	2864.61	21.33 m
DiffUCO	UL	16.22	1.00 m	—	—	DiffUCO	UL	727.32	1.00 m	2947.53	3.78 m
SDDS	UL	18.90	38 s	—	—	SDDS	UL	731.93	14 s	**2971.62**	1.08 m
RLNN	PRL	18.13	1.36 m	35.23	7.83 m	RLNN	PRL	729.00	1.58 m	2907.18	3.67 m
FLD-IG	PRL	18.52	1.40 m	37.40	6.89 m	FLD-IG	PRL	733.48	1.57 m	2922.54	3.07 m
Greedy	H	13.53	25 s	26.71	25 s	Greedy	H	688.31	13 s	2786.00	3.12 m
MFA	H	14.82	27 s	27.94	2.32 m	MFA	H	704.03	1.60 m	2833.86	7.27 m
iSCO	H	18.96	54 s	40.35	11.37 m	iSCO	H	728.24	1.67 m	2919.97	4.18 m
RLSA	H	18.97	23 s	40.53	1.27 m	RLSA	H	733.54	27 s	2955.81	1.45 m
FLD-EG	H	18.97	20 s	40.63	1.91 m	FLD-EG	H	734.18	25 s	2960.13	1.70 m
PID-LD (ours)	H	**18.98**	33 s	**40.66**	2.03 m	**PID-LD (ours)**	H	**734.35**	31 s	2961.60	2.35 m

Equation 29 is an a posteriori certificate for the implemented Bernoulli transition: positive expected descent is guaranteed only when its right-hand side is positive. Equation 32 further compares the calibrated allocation with a budgeted one-bit oracle. The comparison depends on score–gain alignment, entropy, and multi-bit interactions; it is not a comparison with another implemented solver and does not establish global convergence, faster hitting times, or unconditional superiority of PID feedback. The remaining mechanism results are deferred to Appendix D. In particular, the D-channel analysis establishes calibrated undo-mass monotonicity only under an explicit score-ordering condition, while the I-channel analysis gives temperature-dependent anti-concentration and escape-set coverage bounds at a fixed score vector. These results characterize individual control mechanisms rather than claiming monotonic improvement of the complete adaptive controller.

3 Experiments

We evaluate PID-LD on MIS, MaxCl, and MaxCut using the penalized energies in Appendix B. We report complete-method comparisons separately from component-level attribution.

Datasets. We use RB and ER instances for MIS, RB instances for MaxCl, and BA instances for MaxCut [10; 23; 24; 8; 3]. Vertex weights are one. The benchmark training/validation sizes are 1000/500 graphs except for ER-[9000–11000]. Test sets contain 500 RB or BA graphs, 128 ER-[700–800] graphs, and 16 ER-[9000–11000] graphs. PID-LD itself requires no training.

Baselines and resources. The compared methods and their categories are listed in Tables 1–2. The general-purpose solver is Gurobi, with KaMIS additionally included for MIS [13; 12]. Sampling baselines include iSCO, RLSA, and FLD-EG [21; 10; 23]. Learning-based methods include INTEL, DGL, DIFUSCO, LTFT, DiffUCO, SDDS, PPO, DIMES, RLNN, and FLD-

Table 3: Paired objective comparisons on MIS RB-[200–300]. Methods are evaluated on matched graph instances under the same five-seed protocol. Positive mean differences favor PID-LD. Confidence intervals are 95%, and p-values are Holm-adjusted.

Baseline	Mean paired difference δ	95% CI	Holm-adjusted p
RLSA	$+0.106$	[+0.082,+0.110]	3×10^{-5}
FLD-EG	$+0.056$	[+0.051,+0.084]	1.5×10^{-4}

IG [15; 4; 22; 26; 18; 19; 1; 17; 10; 23]. The heuristic comparisons use the same iteration budget, not the same wall-clock budget. The reported learning-based entries use their longest evaluated runtime settings. Hardware is described in Appendix F; training-free status does not mean the absence of hyperparameter selection. The complete-method comparisons establish the overall quality–runtime behavior of PID-LD, but they do not identify which architectural layer is responsible for the observed differences. We therefore next separate the dynamical backbone from the feedback controller through targeted ablations.

Main Results. Tables 1–2 report objective values and total sequential testing times. PID-LD improves over FLD-EG on three of the four MIS configurations: the reported gains are 0.06, 0.53, and 3.00 on RB-[200–300], RB-[800–1200], and ER-[9000–11000], respectively. It is lower by 0.05 on ER-[700–800]. The MaxCl gains over FLD-EG are 0.01 and 0.03, and the MaxCut gains are 0.17 and 1.47. These improvements come with higher reported runtimes. KaMIS attains larger MIS objectives in these tables, and SDDS attains a larger objective on BA-[800–1200] MaxCut. The aggregate results therefore show instance-dependent quality–runtime trade-offs.

Paired objective comparisons. Table 3 reports paired comparisons of PID-LD against RLSA and FLD-EG on MIS RB-[200–300]. All methods are evaluated on matched graph instances under the same five-seed experimental protocol, and the reported differences are computed from the corresponding paired observations. The mean paired differences are $+0.1060$ against RLSA and $+0.0560$ against FLD-EG, with positive values favoring PID-LD. We report 95% confidence intervals together with Holm-adjusted p-values. The complete pairing and multiple-comparison protocol, together with additional within-PID-LD comparisons, is provided in Appendix G.1.

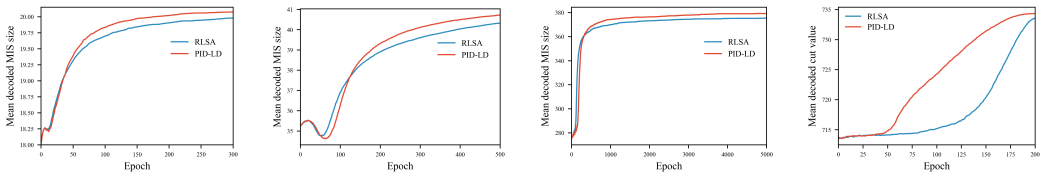
Ablation Study. Table 4 examines several structurally different modifications of PID-LD on MIS RB-[200–300]. We define $\Delta J = J_{\text{full}} - J_{\text{variant}}$, so a positive value denotes a decrease relative to the full method. These variants play different architectural roles and should therefore not be interpreted as an additive ranking of independent component importance. The "w/o momentum" variant is a backbone ablation. It removes the auxiliary velocity state from the search dynamics: the velocity recursion is disabled, V_t is not propagated across iterations, and all velocity-dependent contributions to the proposal score are removed. This ablation produces the largest objective decrease, 0.101, indicating that the stateful momentum dynamics are an important part of the underlying search backbone. By contrast, removing the D- or I-channel is a **controller ablation**. These variants retain the same momentum-based stateful dynamics while disabling one trajectory-dependent regulation mechanism. Under this common backbone, removing the D- and I-channels decreases the objective by 0.029 and 0.036, respectively. Thus, these comparisons measure the incremental effect of closed-loop feedback conditional on retaining the stateful momentum dynamics. The remaining variants separately test actuator adaptation, geometry, randomized control, and the noise family. Taken together, the ablations support complementary rather than competing architectural roles: momentum provides the stateful dynamical backbone, while PID-inspired feedback regulates this backbone and the resulting proposal. All variants retain the calibrated expected flip budget, so this table does not isolate the contribution of calibration itself. Appendix G.1 reports the available paired comparisons and their statistical scope.

Figure 2 compares reported mean decoded objectives over iterations. The plotted PID-LD endpoints are higher than those of RLSA on the displayed configurations, but these curves alone do not isolate D/I effects or establish faster wall-clock optimization. A decoded objective is not automatically the best-so-far feasible objective of an archive; the metric and archive policy must be kept distinct when interpreting nonmonotone segments.

Trajectory-level behavior. On MIS RB-[200–300], the corresponding channel comparisons show a decrease of 0.075 in the immediate undo rate and an increase of 0.58 percentage points in the

Table 4: Ablation study of PID-LD on MIS RB-[200–300]. All variants are evaluated with the same iteration budget, population size, and calibrated expected flip budget. The "w/o momentum" variant removes the velocity state update and all velocity-dependent proposal pathways.

Variant	Ablation role	Avg. obj. ↑	ΔJ
PID-LD	Full hierarchical system	**20.076**	0.000
PID-LD-$\mathcal{S}\alpha\mathcal{S}$	Alternative noise family	20.050	0.026
w/o D-channel	Controller ablation: reversal feedback	20.047	0.029
w/o I-channel	Controller ablation: history feedback	20.040	0.036
w/o momentum	Backbone ablation: remove stateful momentum dynamics	19.975	0.101
Fixed temperature/noise	I-channel actuator ablation	20.069	0.007
Fixed friction/velocity coupling	D-channel actuator ablation	20.067	0.009
Raw gradient, no geometry	Geometry ablation	20.065	0.011
Random damping	Randomized control	20.061	0.015
Random temperature boost	Randomized control	20.064	0.012
Random damping + temperature boost	Randomized control	20.065	0.011



(a) RB-[200-300]. (b) RB-[800-1200]. (c) ER-[9000-11000]. (d) BA-[200-300].

Figure 2: Reported mean decoded objective versus sampling iteration for PID-LD and RLSA: (a)–(c) MIS; (d) MaxCut. This is an iteration-based trajectory comparison, not an isolated channel ablation or a wall-clock comparison.

frequency of archive improvement within 25 iterations after a plateau of at least 25 iterations. These measurements complement the objective-level ablations by characterizing reversal behavior and progress after stagnation. They concern channel-level behavior, rather than direct comparisons with RLSA or FLD-EG. Complete diagnostics, including proposal entropy and calibration residuals, are given in Appendix G.3.

Hyperparameter sensitivity. The five-seed, one-factor analysis on MIS RB-[200–300] examines eleven parameters (Appendix G.2). Within the tested ranges, d, λ_v, and γ exhibit the largest objective changes. This analysis characterizes sensitivity on this benchmark, rather than robustness to joint changes or transfer across graph distributions.

4 Conclusion

PID-LD formulates finite-horizon binary optimization as a hierarchical closed-loop discrete Langevin search process. Stateful stochastic momentum provides the dynamical backbone, PID-inspired trajectory feedback regulates this backbone and the resulting proposal, and calibrated Bernoulli centering preserves a prescribed expected flip budget when translating controlled scores into stochastic bit flips. Our analysis characterizes the calibration objective, establishes conditional finite-budget one-step guarantees, and provides mechanism-level results for the feedback channels under explicit assumptions. These results explain the respective roles of the architectural layers without implying global convergence or unconditional superiority over alternative solvers.

Empirically, PID-LD achieves competitive solution quality across several MIS, MaxCl, and MaxCut configurations, with instance-dependent runtime trade-offs. The ablations support complementary rather than competing roles for the proposed mechanisms: removing momentum removes the underlying stateful search dynamics, whereas D- and I-channel ablations retain this backbone and measure the incremental effect of closed-loop trajectory regulation. The current evidence remains limited by matched-iteration rather than matched-time comparisons, controller ablations concentrated on one benchmark, and local theoretical results that do not imply faster mixing or uniformly faster escape. Improvement in penalized energy also does not by itself certify improvement in the feasible task objective, making the penalty, decoding, archive, and timing protocols important for reproducibility. Future work could extend the same closed-loop architecture to learned score models, integer or categorical variables, and feasibility-aware hybrid solvers.

AI USE STATEMENT

In this work, we used generative AI tools solely for language-related assistance, specifically grammar checking and sentence-level polishing of the text. We did not use generative AI tools for research ideation, algorithm or certificate design, theorem formulation or proofs, experimental design, implementation, or data analysis; all technical and scientific contributions were conceived and carried out by the authors. We have reviewed all AI-assisted edits to ensure that they did not alter the intended technical meaning. We take responsibility for the final content of this work, including all text, claims, and artifacts.

ETHICS STATEMENT

This work studies a general-purpose algorithm for combinatorial optimization and evaluates it on standard graph optimization benchmarks, including maximum independent set, maximum clique, and maximum cut. The study does not involve human subjects, personally identifiable information, sensitive personal data, or user-generated content. The experiments operate only on graph instances and their associated optimization objectives. We are not aware of direct privacy, fairness, or human-subject concerns arising from the reported experiments. As with general-purpose optimization methods, the proposed approach could in principle be applied in a broad range of downstream domains; however, this work does not study or deploy the method in safety-critical or socially sensitive applications.

REPRODUCIBILITY STATEMENT

We provide the information needed to reproduce the formulation, implementation, theoretical results, and experimental protocol of PID-LD throughout the paper and appendices. Appendix B specifies the optimization formulations, penalized energy functions, and gradients used for MIS, MaxCl, and MaxCut. Appendix C gives the complete controller state, adaptive geometry, D- and I-channel updates, calibrated-centering solver, causal ordering, DLP derivation, and conditional transition kernel. Appendix D contains the complete proofs and additional theoretical results. Appendix E reports the benchmark-specific hyperparameter configurations, and Appendix F specifies the hardware and software environment. Finally, Appendix G provides the additional experimental analyses, including paired objective comparisons (Appendix G.1), one-factor hyperparameter sensitivity (Appendix G.2), and trajectory-level diagnostics (Appendix G.3).

REFERENCES

[1] Sungsoo Ahn, Younggyo Seo, and Jinwoo Shin. Learning what to defer for maximum independent sets. In *International conference on machine learning*, pp. 134–144. PMLR, 2020.

[2] Wangpeng An, Haoqian Wang, Qingyun Sun, Jun Xu, Qionghai Dai, and Lei Zhang. A pid controller approach for stochastic optimization of deep networks. In *2018 IEEE/CVF conference on computer vision and pattern recognition*, pp. 8522–8531. IEEE, 2018.

[3] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[4] Maximilian Böther, Otto Kißig, Martin Taraz, Sarel Cohen, Karen Seidel, and Tobias Friedrich. What's wrong with deep learning in tree search for combinatorial optimization. *arXiv preprint arXiv:2201.10494*, 2022.

[5] Hongyi Chen, Jianhai Shu, Jingtao Ding, Yong Li, and Xiao-Ping Steven Zhang. Pid-controlled langevin dynamics for faster sampling of generative models. volume 38, pp. 77373–77400, 2026.

[6] Song Chen, Jiaxu Liu, Pengkai Wang, Chao Xu, Shengze Cai, and Jian Chu. Accelerated optimization in deep learning with a proportional-integral-derivative controller. *Nature Communications*, 15(1):10263, 2024.

[7] Yves Crama. Combinatorial optimization models for production scheduling in automated manufacturing systems. *European Journal of Operational Research*, 99(1):136–153, 1997.

[8] Paul Erdos, Alfréd Rényi, et al. On the evolution of random graphs. *Publ. math. inst. hung. acad. sci*, 5(1):17–60, 1960.

[9] Amir-Massoud Farahmand and Mohammad Ghavamzadeh. Pid accelerated value iteration algorithm. In *International Conference on Machine Learning*, pp. 3143–3153. PMLR, 2021.

[10] Shengyu Feng and Yiming Yang. Regularized langevin dynamics for combinatorial optimization. *arXiv preprint arXiv:2502.00277*, 2025.

[11] Will Grathwohl, Kevin Swersky, Milad Hashemi, David Duvenaud, and Chris Maddison. Oops i took a gradient: Scalable sampling for discrete distributions. In *International Conference on Machine Learning*, pp. 3831–3841. PMLR, 2021.

[12] Ernestine Großmann, Sebastian Lamm, Christian Schulz, and Darren Strash. Finding near-optimal weight independent sets at scale. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 293–302, 2023.

[13] Gurobi Optimization, LLC. *Gurobi Optimizer Reference Manual*. Gurobi Optimization, LLC, 2023. URL `https://www.gurobi.com`.

[14] Scott Kirkpatrick, C Daniel Gelatt Jr, and Mario P Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.

[15] Zhuwen Li, Qifeng Chen, and Vladlen Koltun. Combinatorial optimization with graph convolutional networks and guided tree search. *Advances in neural information processing systems*, 31, 2018.

[16] Christos H Papadimitriou and Kenneth Steiglitz. *Combinatorial optimization: algorithms and complexity*. Courier Corporation, 1998.

[17] Ruizhong Qiu, Zhiqing Sun, and Yiming Yang. Dimes: A differentiable meta solver for combinatorial optimization problems. *Advances in Neural Information Processing Systems*, 35: 25531–25546, 2022.

[18] Sebastian Sanokowski, Sepp Hochreiter, and Sebastian Lehner. A diffusion model framework for unsupervised neural combinatorial optimization. *arXiv preprint arXiv:2406.01661*, 2024.

[19] Sebastian Sanokowski, Wilhelm Berghammer, Martin Ennemoser, Haoyu Peter Wang, Sepp Hochreiter, and Sebastian Lehner. Scalable discrete diffusion samplers: Combinatorial optimization and statistical physics. *arXiv preprint arXiv:2502.08696*, 2025.

[20] Abdelkader Sbihi and Richard W Eglese. Combinatorial optimization and green logistics. *Annals of Operations Research*, 175:159–175, 2010.

[21] Haoran Sun, Katayoon Goshvadi, Azade Nova, Dale Schuurmans, and Hanjun Dai. Revisiting sampling for combinatorial optimization. In *International Conference on Machine Learning*, pp. 32859–32874. PMLR, 2023.

[22] Zhiqing Sun and Yiming Yang. Difusco: Graph-based diffusion solvers for combinatorial optimization. *Advances in neural information processing systems*, 36:3706–3731, 2023.

[23] Shiyue Wang, Ziao Guo, Changhong Lu, and Junchi Yan. Fractional langevin dynamics for combinatorial optimization via polynomial-time escape. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

[24] Ke Xu and Wei Li. Exact phase transitions in random constraint satisfaction problems. *Journal of Artificial Intelligence Research*, 12:93–103, 2000.

[25] Giacomo Zanella. Informed proposals for local mcmc in discrete spaces. *Journal of the American Statistical Association*, 115(530):852–865, 2020.

[26] Dinghuai Zhang, Hanjun Dai, Nikolay Malkin, Aaron C Courville, Yoshua Bengio, and Ling Pan. Let the flows tell: Solving graph combinatorial problems with gflownets. *Advances in neural information processing systems*, 36:11952–11969, 2023.

[27] Ruqi Zhang, Xingchao Liu, and Qiang Liu. A langevin-like sampler for discrete distributions. In *International Conference on Machine Learning*, pp. 26375–26396. PMLR, 2022.

A Related Work

Discrete proposals and sampling-based CO. Locally balanced proposals, Gibbs-with-Gradients, and the discrete Langevin proposal (DLP) use local density or gradient information to construct discrete transitions [25; 11; 27]. Training-free CO solvers include iSCO, RLSA, and FLD-EG [21; 10; 23]. RLD explicitly regularizes the expected distance between successive solutions; accordingly, expected Hamming-move control is an inherited design principle, not a new problem formulation introduced here. Learned alternatives include GFlowNet-based solvers, DIFUSCO, DiffUCO, and scalable discrete diffusion samplers [26; 22; 18; 19]. PID-LD remains training-free and studies trajectory-dependent regulation of a calibrated Bernoulli proposal.

PID control in optimization and sampling. An et al. [2] use past, current, and changing gradients for neural-network optimization. Farahmand & Ghavamzadeh [9] apply PID-inspired control to value iteration, and Chen et al. [6] develop a PID-accelerated optimization framework. PIDLD [5] already closes a feedback loop around Langevin sampling, using gradient history and trends, and evaluates finite-step image generation and reasoning. PID-LD therefore does not claim that feedback, training-free control, or finite-step evaluation is absent from previous work. Its particular setting is binary search with a factorized flip kernel, a prescribed expected Hamming budget, and local analysis of interacting flips. Continuous drift regulation also changes a transition law; the distinction lies in the state space, observations, and constrained proposal construction, not in whether a transition law is controlled at all.

B Case Study on Energy Functions

This paper leverages formal problem formulations for three widely studied canonical cases. Each binary energy has an explicitly specified differentiable extension used by the energy-gradient oracle. For a simple undirected, unweighted graph $G = (V, E)$ with $V = \{1, \dots, N\}$, sums over edges and non-edges below count each unordered pair once and exclude self-pairs. The symmetric adjacency matrix has zero diagonal. We specify each energy and its gradient.

Case Study 1: Maximum Independent Set (MIS). An independent set S within a graph G is a subset of vertices such that no two vertices in S are adjacent. A maximum independent set S^* is an independent set of the largest possible cardinality, formally defined as $S^* \in \arg\max\{|S| : S \subseteq V,\ S \text{ is independent}\}$. This problem can be formulated as the following constrained binary optimization:

$$\max_{x\in\{0,1\}^N} \sum_{i\in V} c_i x_i, \text{ s.t. } \sum_{e(i,j)\in E} x_i x_j = 0, \tag{33}$$

where c_i represents a weight associated with vertex i. Using the penalty method, the constraint can be incorporated into an unconstrained energy function of the form given in Eq. 2:

$$H(x) = -\sum_{i\in V} c_i x_i + \lambda \sum_{e(i,j)\in E} x_i x_j = -\mathbf{c}^\top \mathbf{x} + \lambda \frac{\mathbf{x}^\top \mathbf{A}\mathbf{x}}{2}, \tag{34}$$

where $\lambda > 0$ is a penalty coefficient, and $\mathbf{A}$ denotes the adjacency matrix of G. The final expression provides a compact matrix representation of the energy function. The gradient of $H(\mathbf{x})$ follows directly from its quadratic form and is given by:

$$\nabla H(x) = -\mathbf{c} + \lambda \mathbf{A}\mathbf{x}. \tag{35}$$

Case Study 2: Maximum Clique. A clique in a graph G is a subset of vertices $C \subseteq V$ such that every pair of distinct vertices in C is connected by an edge. The maximum clique problem seeks a clique C^* of the largest possible size, formally defined as $C^* \in \arg\max\{|C| : C \subseteq V,\ C \text{ is a clique}\}$. This can be formulated as a binary optimization with a structural constraint:

$$\max_{\mathbf{x}\in\{0,1\}^N} \sum_{i\in V} c_i x_i \quad \text{s.t.} \quad \sum_{e(i,j)\notin E} x_i x_j = 0, \tag{36}$$

where c_i is a weight for vertex i, and the constraint ensures that no two selected vertices are missing an edge between them (i.e., all selected pairs must be adjacent). Employing a penalty method, the

constrained problem is transformed into an unconstrained energy minimization. The corresponding energy function, following the form of Eq. 2, is:

$$H(\mathbf{x}) = -\sum_{i \in V} c_i x_i + \lambda \sum_{e(i,j) \notin E} x_i x_j. \tag{37}$$

This expression can be rewritten in a more compact matrix form. Noting that the sum over non-edges is related to the total number of vertex pairs selected minus those connected by edges, we derive:

$$H(\mathbf{x}) = -\mathbf{c}^\top \mathbf{x} + \frac{\lambda}{2} \left((\mathbf{1}^\top \mathbf{x})^2 - \mathbf{x}^\top \mathbf{x} - \mathbf{x}^\top \mathbf{A} \mathbf{x} \right), \tag{38}$$

where $\lambda > 0$ is a penalty parameter, $\mathbf{A}$ is the adjacency matrix, and $\mathbf{1}$ denotes the vector of all ones. The term $(\mathbf{1}^\top \mathbf{x})^2$ counts all ordered pairs of selected vertices, from which we subtract the self-pairs ($\mathbf{x}^\top \mathbf{x}$) and the existing edges ($\mathbf{x}^\top \mathbf{A} \mathbf{x}$), leaving the penalized non-edges. The gradient of this energy function is readily obtained as:

$$\nabla H(\mathbf{x}) = -\mathbf{c} + \lambda \left((\mathbf{1}^\top \mathbf{x}) \mathbf{1} - \mathbf{x} - \mathbf{A} \mathbf{x} \right). \tag{39}$$

Case Study 3: Max Cut. The Maximum Cut problem is a fundamental combinatorial optimization problem on a graph. Given an undirected graph $G(V, E)$, the objective is to partition the vertex set V into two disjoint subsets, S and its complement $\bar{S} = V \setminus S$, such that the number of edges crossing the partition (i.e., edges with one endpoint in S and the other in $\bar{S}$) is maximized. Formally, the goal is to find:

$$\max_{S \subseteq V} |\{e(i,j) \in E \mid i \in S,\ j \in \bar{S}\}|. \tag{40}$$

This can be formulated using binary assignment variables. Let $x_i \in \{0, 1\}$ indicate whether vertex i belongs to set S ($x_i = 1$) or $\bar{S}$ ($x_i = 0$). An edge $e(i, j)$ is counted in the cut if its endpoints are assigned to different sets. This condition, $x_i \neq x_j$, can be encoded as $x_i + x_j - 2x_i x_j = 1$. A more common and symmetric formulation maps the binary variables to the domain $\{-1, +1\}$ via the transformation $s_i = 2x_i - 1$. This yields the standard quadratic unconstrained binary optimization form:

$$\max_{\mathbf{x} \in \{0,1\}^N} \sum_{e(i,j) \in E} \frac{1 - (2x_i - 1)(2x_j - 1)}{2}. \tag{41}$$

Transforming this objective into a minimization problem for consistency with the energy-based framework leads to the energy function:

$$H(\mathbf{x}) = -\sum_{e(i,j) \in E} \frac{1 - (2x_i - 1)(2x_j - 1)}{2}. \tag{42}$$

The constants cancel when the binary expression is expanded. Counting each undirected edge once in the sum gives the exact matrix identity:

$$H(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x} - \mathbf{1}^\top \mathbf{A} \mathbf{x}, \tag{43}$$

where $\mathbf{A}$ is the adjacency matrix of G. The quadratic term counts selected–selected edges twice, while $\mathbf{1}^\top \mathbf{A} \mathbf{x}$ sums the degrees of selected vertices. Their difference is exactly the negative cut size. The gradient of this quadratic energy function is derived as:

$$\nabla H(\mathbf{x}) = \mathbf{A}(2\mathbf{x} - \mathbf{1}). \tag{44}$$

C Implementation Details of PID-LD

This appendix records the controller memories, numerical calibration, geometry options, and proposal derivation. The D/I observations are part of the defined controller; they are not all optional stabilizers. Geometry variants and optional adaptive stepping are identified separately. Throughout this section, all clipping operations are applied elementwise unless otherwise specified. We write

$$\mathrm{clip}(z, a, b) = \min\{\max\{z, a\}, b\}. \tag{45}$$

C.1 Extended state and causal ordering

We index $\mathcal{Y}_t$ immediately before the current energy-gradient evaluation. An explicit state sufficient for the displayed update is

$$\begin{aligned} \mathcal{Y}_t &= \big(t, X_t, V_t, M^P_{t-1}, \overline{O}_{t-1}, D_{t-1}, I_{t-1}, r_{t-1}, \overline{\phi}_t, \mathcal{A}_{t-1}\big), \\ \mathcal{A}_{t-1} &= \big(X^\star_{t-1}, E^\star_{t-1}, x_{\mathrm{best},t-1}, E_{\mathrm{best},t-1}\big). \end{aligned} \tag{46}$$

The second-moment memory M^P_{t-1} is included when adaptive geometry is used. The iteration index supplies deterministic schedules; any additional adaptive structural state must likewise be included. $\overline{\phi}_t$ contains only previously sampled flips. The improvement indicator compares E_t with $E^\star_{t-1}$ before forming $\mathcal{A}_t$. The current mask updates $\overline{\phi}_{t+1}$ only after the proposal has been sampled. This makes the mathematical ordering explicit without using the current random mask to choose its own distribution.

C.2 Optional adaptive geometry

The diagonal geometry operator $\mathcal{P}_t$ normalizes coordinate scales before constructing the proportional score. The simplest choice is the identity:

$$\widehat{G}_t = \mathcal{P}_t(G_t) = G_t. \tag{47}$$

A scale-normalized implementation uses a second-moment memory

$$M^P_t = \begin{cases} G_t^{\odot 2}, & t = 0, \\ \beta_P M^P_{t-1} + (1-\beta_P) G_t^{\odot 2}, & t \geq 1, \end{cases} \tag{48}$$

and the diagonal preconditioner

$$P_t = \big(M^P_t + \varepsilon_P\big)^{-1/2}. \tag{49}$$

The normalized proposal gradient is

$$\widehat{G}_t = \mathcal{P}_t(G_t) = \Lambda_t \odot \left[\rho_P (P_t \odot G_t) + (1-\rho_P) G_t\right],\ \rho_P \in \{0,1\}. \tag{50}$$

Here $\Lambda_t \in \mathbb{R}_+^{N\times K}$ is an optional structural scaling, for example a degree-based normalization on graph instances. Setting $\rho_P = 0$ and $\Lambda_t = \mathbf{1}$ recovers the identity geometry.

The integration step and the base annealing temperature may be chosen adaptively:

$$g^{\mathrm{rms}}_t = \frac{\|G_t\|_F}{\sqrt{NK}},\ \eta_t = \mathrm{clip}\left(\frac{\eta_0}{1 + c_\eta g^{\mathrm{rms}}_t}, \eta_{\min}, \eta_{\max}\right), \tag{51}$$

$$\tau_t = \tau_{\min} + (\tau_0 - \tau_{\min})\alpha_\tau^t,\ 0 < \alpha_\tau \leq 1. \tag{52}$$

If adaptive stepping is disabled, we set $c_\eta = 0$.

C.3 Derivative-channel oscillation statistic

The D-channel uses sign reversals in normalized proposal-gradient observations as an oscillation proxy. This statistic need not coincide with realized immediate bit reversals. For each chain k, define the scale

$$m^{(k)}_t = \frac{1}{N}\sum_{i=1}^{N} \left|\widehat{G}^{(k)}_{i,t}\right| + \varepsilon_m. \tag{53}$$

The normalized observation is

$$O^{(k)}_{i,t} = \frac{\widehat{G}^{(k)}_{i,t}}{m^{(k)}_t}. \tag{54}$$

We maintain an exponential moving average:

$$\overline{O}_t = \beta_G \overline{O}_{t-1} + (1-\beta_G) O_t. \tag{55}$$

The clipped filtered increment is

$$\Delta O_t = \mathrm{clip}\left(\overline{O}_t - \overline{O}_{t-1}, -R_{\max}, R_{\max}\right). \tag{56}$$

A soft sign-reversal gate is

$$\Omega_t = \sigma\left(\alpha_{\mathrm{osc}}\left(-\overline{O}_t \odot \overline{O}_{t-1}\right)\right). \tag{57}$$

When a coordinate reverses its filtered gradient direction, the product $\overline{O}_{i,t}^{(k)}\overline{O}_{i,t-1}^{(k)}$ becomes negative, and the gate becomes large.

The local oscillation magnitude is

$$D_{i,t}^{\mathrm{loc},(k)} = \Omega_{i,t}^{(k)}\left|\Delta O_{i,t}^{(k)}\right|. \tag{58}$$

The chain-level oscillation statistic is the empirical 0.7-quantile:

$$d_t^{(k)} = Q_{0.7}\left(D_{1,t}^{\mathrm{loc},(k)}, \ldots, D_{N,t}^{\mathrm{loc},(k)}\right). \tag{59}$$

The derivative-channel memory is updated by

$$D_t^{(k)} = \beta_D D_{t-1}^{(k)} + (1-\beta_D)d_t^{(k)}. \tag{60}$$

The effective friction and velocity-score coefficient are then

$$\gamma_{t,\mathrm{eff}}^{(k)} = \mathrm{clip}\left(\gamma_0\left[1 + K_D a_\gamma D_t^{(k)}\right], \gamma_{\min}, \gamma_{\max}\right), \tag{61}$$

$$\nu_{t,\mathrm{eff}}^{(k)} = \frac{\nu_0}{\max\left\{1 + K_D a_\nu D_t^{(k)}, \varepsilon_\nu\right\}}. \tag{62}$$

For nonnegative gains, a larger statistic increases the friction coefficient and reduces the velocity-coupling coefficient. This alone does not guarantee a smaller realized velocity or lower undo probability after recalibration.

C.4 Integral-channel stagnation and flip-rate statistic

The I-channel integrates stagnation and past empirical flip-rate observations rather than energy gradients. The transition retains both inputs as specified below. Under exact calibration, however, the conditional expected flip rate is fixed by Eq. 22; temperature and noise cannot independently move this mean toward a different target. The flip-rate input must therefore be interpreted as a historical trajectory statistic, not as evidence of an uncorrected expected-budget deficit.

First define the chain-wise improvement indicator using the previous archive:

$$\iota_t^{(k)} = \mathbf{1}\left\{E_t^{(k)} < E_{t-1}^{\star,(k)} - \varepsilon_{\mathrm{imp}}\right\}. \tag{63}$$

For $t = 0$, take $E_{-1}^{\star,(k)} = +\infty$. The chain-wise archive energy is updated by

$$E_t^{\star,(k)} = \min\left\{E_{t-1}^{\star,(k)}, E_t^{(k)}\right\}. \tag{64}$$

Let

$$T_b = \lfloor \rho_I T \rfloor \tag{65}$$

be the burn-in length. The stagnation counter is

$$r_t^{(k)} = \begin{cases} 0, & t < T_b, \\ \left(1 - \iota_t^{(k)}\right)\left(r_{t-1}^{(k)} + 1\right), & t \geq T_b. \end{cases} \tag{66}$$

The normalized stagnation signal is

$$s_t^{(k)} = \mathrm{clip}\left(\frac{r_t^{(k)} - p_I}{\max\{p_I, \varepsilon_m\}}, 0, 1\right). \tag{67}$$

The flip-rate signal uses only the previous flip-rate memory $\overline{\phi}_t^{(k)}$, not the current random mask. Define

$$e_{\phi,t}^{(k)} = \mathrm{clip}\left(\frac{\phi^\star - \overline{\phi}_t^{(k)}}{\max\{\phi^\star, \varepsilon_m\}}, -1, 1\right), \tag{68}$$

where $\phi^\star$ is the target empirical flip rate. The integral observation is

$$o_{I,t}^{(k)} = a_s s_t^{(k)} + a_\phi e_{\phi,t}^{(k)}. \tag{69}$$

The leaky integral state is

$$I_t^{(k)} = \mathrm{clip}\left(\max\{0, 1 - \lambda_I \eta_t\} I_{t-1}^{(k)} + \eta_t o_{I,t}^{(k)}, -I_{\max}, I_{\max}\right). \tag{70}$$

The I-channel acts through the effective temperature and noise scale:

$$\tau_{t,\mathrm{eff}}^{(k)} = \tau_t \exp\left(\mathrm{clip}\left(K_I a_\tau I_t^{(k)}, -c_\tau, c_\tau\right)\right), \tag{71}$$

$$\sigma_{t,\mathrm{eff}}^{(k)} = \sigma_0 + K_I a_\sigma \left[I_t^{(k)}\right]_+, \ [z]_+ := \max\{z, 0\}. \tag{72}$$

C.5 Calibrated centering solver

For each chain k, calibrated centering solves

$$\Psi_t^{(k)}(\kappa) = d_{\mathrm{eff}}, \tag{73}$$

where

$$\Psi_t^{(k)}(\kappa) = \sum_{i=1}^{N} \sigma\left(\mathrm{clip}\left(\frac{S_{i,t}^{(k)} - \kappa}{\tau_{t,\mathrm{eff}}^{(k)}}, -L_{\max}, L_{\max}\right)\right). \tag{74}$$

The function $\Psi_t^{(k)}(\kappa)$ is continuous and nonincreasing in κ. Because of logit clipping,

$$\lim_{\kappa \to -\infty} \Psi_t^{(k)}(\kappa) = N(1 - \varepsilon_L), \ \lim_{\kappa \to +\infty} \Psi_t^{(k)}(\kappa) = N\varepsilon_L, \tag{75}$$

where

$$\varepsilon_L = \sigma(-L_{\max}). \tag{76}$$

Thus a solution exists whenever

$$d_{\mathrm{eff}} \in [N\varepsilon_L, N(1 - \varepsilon_L)]. \tag{77}$$

If d_{eff} falls outside this range, we use the clipped target

$$d_{\mathrm{eff}}^{\mathrm{clip}} = \mathrm{clip}\left(d_{\mathrm{eff}}, N\varepsilon_L, N(1 - \varepsilon_L)\right). \tag{78}$$

The center $\kappa_t^{(k)}$ can be found by bisection. A practical initialization is

$$\kappa_{\min}^{(k)} = \min_i S_{i,t}^{(k)} - \tau_{t,\mathrm{eff}}^{(k)}(L_{\max} + c_{\mathrm{bis}}), \ \kappa_{\max}^{(k)} = \max_i S_{i,t}^{(k)} + \tau_{t,\mathrm{eff}}^{(k)}(L_{\max} + c_{\mathrm{bis}}), \tag{79}$$

with $c_{\mathrm{bis}} > 0$. Since $\Psi_t^{(k)}$ is nonincreasing, bisection updates the interval until

$$\left|\Psi_t^{(k)}(\kappa_t^{(k)}) - d_{\mathrm{eff}}^{\mathrm{clip}}\right| \le \varepsilon_{\mathrm{cal}}. \tag{80}$$

The resulting probabilities satisfy

$$\sum_{i=1}^{N} q_{i,t}^{(k)} = d_{\mathrm{eff}}^{\mathrm{clip}} \pm \varepsilon_{\mathrm{cal}}. \tag{81}$$

C.6 ALGORITHM ORDERING

Algorithm 1 is presented in the main text. The improvement indicator is evaluated against the previous archive before that archive is updated. The current flip mask is sampled only after the controller has been evaluated and the proposal calibrated. The final population X_T is evaluated before returning the archive state.

C.7 DLP DERIVATION

Let $\widetilde{H} : [0,1]^N \to \mathbb{R}$ be a differentiable continuous extension of H, and define the continuous log-density up to an additive constant by

$$\ell_\tau(u) = -\frac{\widetilde{H}(u)}{\tau},\ u \in [0,1]^N. \tag{82}$$

A discrete Langevin proposal on a product domain can be motivated by restricting a continuous Langevin proposal to feasible discrete states. For $x, y \in \{0,1\}^N$, one common form is

$$q_\alpha^{\mathrm{DLP}}(y \mid x) \propto \exp\left\{-\frac{1}{2\alpha}\left\|y - x - \frac{\alpha}{2}\nabla\ell_\tau(x)\right\|_2^2\right\}. \tag{83}$$

For a binary flip of coordinate i, the displacement is

$$y_i - x_i = 1 - 2x_i = -\chi_i(x),\ \chi_i(x) = 2x_i - 1. \tag{84}$$

Expanding Eq. 83 shows that the binary DLP flip logit contains the oriented first-order term

$$\frac{1}{2}\nabla_i\ell_\tau(x)(y_i - x_i) = \frac{1}{2\tau}\chi_i(x)\partial_i\widetilde{H}(x). \tag{85}$$

The exact one-bit gain and its multilinear gradient identity are defined in the main text. These identities motivate the oriented score; they do not establish stationary correctness for the adaptive search kernel.

C.8 CONDITIONAL BERNOULLI KERNEL

For a proposed next population $X' \in \{0,1\}^{N\times K}$, define the relative flip mask

$$M_i^{(k)}(X', X_t) = \mathbf{1}\left\{X_i'^{(k)} \neq X_{i,t}^{(k)}\right\}. \tag{86}$$

Conditional on $\mathcal{Y}_t$ and ξ_t, the proposal kernel is the factorized Bernoulli kernel

$$Q_t^\xi(X' \mid \mathcal{Y}_t) = \prod_{k=1}^{K}\prod_{i=1}^{N}\left(q_{i,t}^{(k)}\right)^{M_i^{(k)}(X',X_t)}\left(1 - q_{i,t}^{(k)}\right)^{1-M_i^{(k)}(X',X_t)}. \tag{87}$$

The marginal adaptive kernel is

$$Q_t(X' \mid \mathcal{Y}_t) = \int Q_t^\xi(X' \mid \mathcal{Y}_t)\mu_t(d\xi \mid \mathcal{Y}_t). \tag{88}$$

D PROOFS AND ADDITIONAL THEORETICAL RESULTS

For a subset $S \subseteq [N]$, x^S denotes the binary vector obtained from x by flipping all coordinates in S. For a Bernoulli flip mask B, we write x^B for the corresponding multi-bit move. The one-bit gain is

$$g_i(x) = H(x) - H(F_i x). \tag{89}$$

D.1 Proof of Theorem 1

Proof. We prove the one-bit residual bound for independent multi-bit moves. Fix a deterministic flip set $S \subseteq [N]$. Let

$$S = \{i_1, \dots, i_m\} \tag{90}$$

be any ordering of S, and define

$$T_0 = \emptyset,\ T_r = \{i_1, \dots, i_r\},\ r = 1, \dots, m. \tag{91}$$

By telescoping,

$$H(x) - H(x^S) = \sum_{r=1}^{m} \left[H(x^{T_{r-1}}) - H(x^{T_r}) \right]. \tag{92}$$

Since $T_r = T_{r-1} \cup \{i_r\}$, the r-th term is the one-bit gain of coordinate i_r evaluated after the previous flips:

$$H(x^{T_{r-1}}) - H(x^{T_r}) = g_{i_r}(x^{T_{r-1}}). \tag{93}$$

Therefore,

$$H(x) - H(x^S) = \sum_{r=1}^{m} g_{i_r}(x^{T_{r-1}}). \tag{94}$$

By the residual assumption in Theorem 1, because $T_{r-1} \subseteq [N] \setminus \{i_r\}$,

$$g_{i_r}(x^{T_{r-1}}) \geq g_{i_r}(x) - \sum_{j \in T_{r-1}} \epsilon_{i_r j}(x). \tag{95}$$

Thus

$$H(x) - H(x^S) \geq \sum_{r=1}^{m} g_{i_r}(x) - \sum_{r=1}^{m} \sum_{j \in T_{r-1}} \epsilon_{i_r j}(x). \tag{96}$$

The double sum counts each unordered pair contained in S at most once. Using the symmetry $\epsilon_{ij}(x) = \epsilon_{ji}(x)$, we obtain

$$H(x) - H(x^S) \geq \sum_{i \in S} g_i(x) - \sum_{\{i,j\} \subseteq S} \epsilon_{ij}(x). \tag{97}$$

Now let $B_i \sim \text{Bernoulli}(q_i)$ be independent. Applying the previous deterministic bound to $S = \{i : B_i = 1\}$ gives

$$H(x) - H(x^B) \geq \sum_{i=1}^{N} B_i g_i(x) - \sum_{1 \leq i < j \leq N} B_i B_j \epsilon_{ij}(x). \tag{98}$$

Taking expectation and using independence,

$$\mathbb{E}[B_i] = q_i,\ \mathbb{E}[B_i B_j] = q_i q_j,\ i \neq j, \tag{99}$$

yields

$$\mathbb{E}\left[H(x) - H(x^B) \right] \geq \sum_{i=1}^{N} q_i g_i(x) - \sum_{1 \leq i < j \leq N} \epsilon_{ij}(x) q_i q_j. \tag{100}$$

□

D.2 Proof of Lemma 1

Proof. Recall

$$\epsilon_L = \sigma(-L_{\max}) \tag{101}$$

and

$$\mathcal{C}_{d,L} = \left\{ p \in [\epsilon_L, 1-\epsilon_L]^N : \sum_{i=1}^N p_i = d \right\}. \tag{102}$$

Consider the optimization problem

$$\max_{p \in \mathcal{C}_{d,L}} \left\{ \langle S, p \rangle + \tau \mathcal{H}(p) \right\}, \tag{103}$$

where

$$\mathcal{H}(p) = \sum_{i=1}^N \left[-p_i \log p_i - (1-p_i) \log(1-p_i) \right]. \tag{104}$$

The feasible set $\mathcal{C}_{d,L}$ is compact and convex. The objective is strictly concave on $(0,1)^N$, because the binary entropy is strictly concave. Therefore the maximizer is unique.

We first consider coordinates that are not saturated by the box constraints. Let κ be the Lagrange multiplier for the equality constraint

$$\sum_{i=1}^N p_i = d. \tag{105}$$

For an interior coordinate, the KKT stationarity condition is

$$S_i + \tau \log \frac{1-p_i}{p_i} - \kappa = 0. \tag{106}$$

Equivalently,

$$\log \frac{p_i}{1-p_i} = \frac{S_i - \kappa}{\tau}, \tag{107}$$

and hence

$$p_i = \sigma\left(\frac{S_i - \kappa}{\tau} \right). \tag{108}$$

The box constraint $p_i \in [\epsilon_L, 1-\epsilon_L]$ is equivalent to bounding its logit by $[-L_{\max}, L_{\max}]$, because

$$\epsilon_L = \sigma(-L_{\max}),\ 1-\epsilon_L = \sigma(L_{\max}). \tag{109}$$

Thus, including the saturated coordinates, the unique maximizer has the form

$$p_i = \sigma\left(\operatorname{clip}\left(\frac{S_i - \kappa}{\tau}, -L_{\max}, L_{\max} \right) \right). \tag{110}$$

The scalar κ is chosen so that $\sum_{i=1}^N p_i = d$. This is exactly the calibrated centering rule used by PID-LD.

It remains only to justify existence of such a κ. Define

$$\Psi(\kappa) = \sum_{i=1}^N \sigma\left(\operatorname{clip}\left(\frac{S_i - \kappa}{\tau}, -L_{\max}, L_{\max} \right) \right). \tag{111}$$

The function Ψ is continuous and nonincreasing in κ, with

$$\lim_{\kappa \to -\infty} \Psi(\kappa) = N(1-\epsilon_L),\ \lim_{\kappa \to +\infty} \Psi(\kappa) = N\epsilon_L. \tag{112}$$

Therefore, for any feasible $d \in [N\epsilon_L, N(1-\epsilon_L)]$, there exists at least one κ satisfying $\Psi(\kappa) = d$. Although κ need not be unique on a saturated plateau, the probability vector q is unique because the optimization problem has a unique maximizer. This proves Lemma 1. □

D.3 Proof of Theorem 2

Proof. Condition on the extended state $\mathcal{Y}_t$ and the proposal perturbation ξ_t. Then the Bernoulli mask is independent across coordinates with

$$B_i \mid \mathcal{Y}_t, \xi_t \sim \text{Bernoulli}(q_i). \tag{113}$$

By Theorem 1, applied to the implemented probability vector q, we have

$$\mathbb{E}\left[H(x) - H(x^B) \mid \mathcal{Y}_t, \xi_t\right] \geq \sum_{i=1}^{N} q_i g_i(x) - \sum_{1\leq i<j\leq N} \epsilon_{ij}(x) q_i q_j. \tag{114}$$

This proves the a posteriori descent certificate.

For the oracle-comparison statement, let

$$p^\star \in \arg\max_{p\in\mathcal{C}_{d,L}} \langle g(x), p\rangle,\ G_{d,L}(x) = \langle g(x), p^\star\rangle. \tag{115}$$

By Lemma 1, the calibrated probability vector q satisfies

$$\langle S, q\rangle + \tau\mathcal{H}(q) \geq \langle S, p^\star\rangle + \tau\mathcal{H}(p^\star). \tag{116}$$

Write

$$e = S - \alpha g(x). \tag{117}$$

Substituting $S = \alpha g(x) + e$ into the previous inequality gives

$$\alpha\langle g(x), q\rangle + \langle e, q\rangle + \tau\mathcal{H}(q) \geq \alpha\langle g(x), p^\star\rangle + \langle e, p^\star\rangle + \tau\mathcal{H}(p^\star). \tag{118}$$

Rearranging yields

$$\langle g(x), q\rangle \geq G_{d,L}(x) - \frac{\langle S - \alpha g(x), q - p^\star\rangle + \tau\left(\mathcal{H}(q) - \mathcal{H}(p^\star)\right)}{\alpha}. \tag{119}$$

Combining this inequality with the a posteriori descent certificate proves

$$\mathbb{E}\left[H(x) - H(x^B) \mid \mathcal{Y}_t, \xi_t\right] \geq G_{d,L}(x) - \frac{\Gamma_\alpha(x, S; q, p^\star)}{\alpha} - \sum_{1\leq i<j\leq N} \epsilon_{ij}(x) q_i q_j. \tag{120}$$

□

D.4 Centered conservative form of Theorem 2

For a scale parameter $\alpha > 0$, define the centered score-alignment error

$$\rho_{\text{cen},\alpha}(x, S) = \inf_{b\in\mathbb{R}} \|S - \alpha g(x) - b\mathbf{1}\|_\infty. \tag{121}$$

The centered form is natural because calibrated Bernoulli centering is invariant to adding the same constant to all scores.

Proposition 1 (Centered conservative sufficient form)**.** *Under the assumptions of Theorem 2, let*

$$W_{d,L} = \sup_{p,r\in\mathcal{C}_{d,L}} \|p - r\|_1. \tag{122}$$

Then, for any $p^\star \in \arg\max_{p\in\mathcal{C}_{d,L}} \langle g(x), p\rangle$,

$$\begin{aligned}&\mathbb{E}\left[H(x) - H(x^B) \mid \mathcal{Y}_t, \xi_t\right]\\ &\quad\geq G_{d,L}(x) - \frac{W_{d,L}\rho_{\text{cen},\alpha}(x, S) + \tau\left[\mathcal{H}(q) - \mathcal{H}(p^\star)\right]_+}{\alpha} - \sum_{1\leq i<j\leq N} \epsilon_{ij}(x) q_i q_j.\end{aligned} \tag{123}$$

where $[z]_+ = \max\{z, 0\}$. *Moreover,*

$$W_{d,L} \leq 2\min\{d - N\varepsilon_L,\ N(1 - \varepsilon_L) - d\} \leq 2d. \tag{124}$$

Consequently, using $\mathcal{H}(q) \leq Nh(d/N)$ *and* $\epsilon_{ij}(x) \leq \epsilon_{\max}$ *gives the simplified conservative corollary*

$$\mathbb{E}\left[H(x) - H(x^B) \mid \mathcal{Y}_t, \xi_t\right] \geq G_{d,L}(x) - \frac{2d\rho_{\text{cen},\alpha}(x, S) + \tau Nh(d/N)}{\alpha} - \frac{\epsilon_{\max}}{2} d^2. \tag{125}$$

Proof. Since $q, p^\star \in \mathcal{C}_{d,L}$, they have the same total mass. Hence, for any $b \in \mathbb{R}$,

$$\langle b\mathbf{1}, q - p^\star \rangle = 0. \tag{126}$$

Therefore,

$$\langle S - \alpha g(x), q - p^\star \rangle = \langle S - \alpha g(x) - b\mathbf{1}, q - p^\star \rangle . \tag{127}$$

Taking the infimum over b gives

$$\langle S - \alpha g(x), q - p^\star \rangle \le \rho_{\mathrm{cen},\alpha}(x, S)\|q - p^\star\|_1 \le W_{d,L}\rho_{\mathrm{cen},\alpha}(x, S). \tag{128}$$

Also,

$$\tau \left(\mathcal{H}(q) - \mathcal{H}(p^\star)\right) \le \tau \left[\mathcal{H}(q) - \mathcal{H}(p^\star)\right]_+ . \tag{129}$$

Substituting these two upper bounds into the localized oracle-comparison inequality in Theorem 2 proves the first claim.

It remains to bound $W_{d,L}$. For $p, r \in \mathcal{C}_{d,L}$, define

$$y_i = p_i - \varepsilon_L,\ z_i = r_i - \varepsilon_L. \tag{130}$$

Then $y_i, z_i \in [0, 1 - 2\varepsilon_L]$ and

$$\sum_{i=1}^N y_i = \sum_{i=1}^N z_i = d - N\varepsilon_L. \tag{131}$$

Since the positive and negative parts of $y - z$ have equal mass,

$$\|p - r\|_1 = \|y - z\|_1 = 2\sum_{i=1}^N (y_i - z_i)_+. \tag{132}$$

The positive part is at most the total mass of y and also at most the unused capacity of z, so

$$\sum_{i=1}^N (y_i - z_i)_+ \le \min\{d - N\varepsilon_L,\ N(1 - \varepsilon_L) - d\}. \tag{133}$$

This proves the displayed bound on $W_{d,L}$.

Finally, Jensen's inequality gives

$$\mathcal{H}(q) = \sum_{i=1}^N h(q_i) \le Nh\left(\frac{1}{N}\sum_{i=1}^N q_i\right) = Nh(d/N), \tag{134}$$

and

$$\sum_{1 \le i < j \le N} q_i q_j = \frac{1}{2}\left[\left(\sum_{i=1}^N q_i\right)^2 - \sum_{i=1}^N q_i^2\right] \le \frac{d^2}{2}. \tag{135}$$

The simplified conservative corollary follows. □

Numerical calibration. The statements above assume exact calibration $\sum_i q_i = d$. If the bisection solver is stopped with tolerance, define the realized budget

$$d_q = \sum_{i=1}^N q_i \tag{136}$$

and replace $\mathcal{C}_{d,L}$ by

$$\mathcal{C}_{q,L} = \left\{ p \in [\varepsilon_L, 1 - \varepsilon_L]^N : \sum_{i=1}^N p_i = d_q \right\}. \tag{137}$$

The same proof holds verbatim with d replaced by d_q.

D.5 Score-alignment decomposition for the implemented PID-LD score

This subsection justifies the score-alignment estimate used after Theorem 2. Consider one chain and define the unclipped score before score-side clipping by

$$\widetilde{S} = \frac{1}{2} K_P \widehat{G}(x) \odot \chi(x) - \nu_{\text{eff}} V_{t+1} \odot \chi(x). \tag{138}$$

The velocity V_{t+1} in this expression is already the implemented, clipped velocity. We collect the remaining proportional-score and final-score clipping discrepancy in the residual

$$r_{\text{clip}} = \left\| S - \widetilde{S} \right\|_\infty. \tag{139}$$

Assume $\widetilde{H}$ is the multilinear extension of H, so that at every binary vertex

$$g(x) = \chi(x) \odot \nabla \widetilde{H}(x). \tag{140}$$

Assume $K_P > 0$ and set

$$\alpha = \frac{K_P}{2}. \tag{141}$$

Since $\rho_{\text{cen},\alpha}(x, S) \le \|S - \alpha g(x)\|_\infty$, we have

$$\begin{aligned} \rho_{\text{cen},\alpha}(x, S) &\le \left\| S - \frac{K_P}{2} \chi(x) \odot \nabla \widetilde{H}(x) \right\|_\infty \\ &\le \left\| S - \widetilde{S} \right\|_\infty + \left\| \frac{K_P}{2} \chi(x) \odot \left(\widehat{G}(x) - \nabla \widetilde{H}(x) \right) \right\|_\infty + \left\| \nu_{\text{eff}} V_{t+1} \odot \chi(x) \right\|_\infty. \end{aligned} \tag{142}$$

Because $|\chi_i(x)| = 1$ for every coordinate,

$$\left\| \chi(x) \odot \left(\widehat{G}(x) - \nabla \widetilde{H}(x) \right) \right\|_\infty = \left\| \widehat{G}(x) - \nabla \widetilde{H}(x) \right\|_\infty \tag{143}$$

and

$$\left\| V_{t+1} \odot \chi(x) \right\|_\infty = \left\| V_{t+1} \right\|_\infty. \tag{144}$$

Therefore,

$$\rho_{\text{cen},\alpha}(x, S) \le \frac{K_P}{2} \left\| \widehat{G}(x) - \nabla \widetilde{H}(x) \right\|_\infty + \nu_{\text{eff}} \|V_{t+1}\|_\infty + r_{\text{clip}}. \tag{145}$$

This is an error decomposition, not a monotonic improvement theorem. Rescaling an already exact gradient can increase the first term. Reducing ν_{eff} lowers its multiplier at fixed velocity, but the velocity also changes with damping, noise, and history. Consequently, the inequality does not establish that P normalization or D control always improves score alignment.

D.6 Calibrated undo-mass suppression

We next prove a monotonicity property showing that reducing scores on previously flipped coordinates reduces their total calibrated Bernoulli probability mass, even after the centering parameter is recomputed.

Proposition 2 (Calibrated undo-mass monotonicity)**.** *Let $q(S)$ denote the calibrated probability vector*

$$q(S) = \arg \max_{p \in \mathcal{C}_{d,L}} \left\{ \langle S, p \rangle + \tau \mathcal{H}(p) \right\}. \tag{146}$$

For $A \subseteq [N]$, define

$$U_A(S) = \sum_{i \in A} q_i(S). \tag{147}$$

If two score vectors $S, S' \in \mathbb{R}^N$ satisfy

$$S'_i \le S_i, \quad i \in A, \; S'_j \ge S_j, \quad j \notin A, \tag{148}$$

then

$$U_A(S') \le U_A(S). \tag{149}$$

Proof. Let

$$\phi_\tau(u) = \sigma\left(\mathrm{clip}\left(u/\tau, -L_{\max}, L_{\max}\right)\right). \tag{150}$$

For a score vector S, the calibrated probabilities can be written as

$$q_i(S) = \phi_\tau(S_i - \kappa),\ \sum_{i=1}^{N} \phi_\tau(S_i - \kappa) = d, \tag{151}$$

for some center κ. The map ϕ_τ is continuous and nondecreasing.

Let κ and κ' be centers for S and S', respectively. Suppose, for contradiction, that

$$U_A(S') > U_A(S). \tag{152}$$

If $\kappa' \geq \kappa$, then for every $i \in A$ we have

$$S'_i - \kappa' \leq S_i - \kappa, \tag{153}$$

and hence $q_i(S') \leq q_i(S)$ by monotonicity of ϕ_τ. This contradicts $U_A(S') > U_A(S)$. Therefore,

$$\kappa' < \kappa. \tag{154}$$

For every $j \notin A$, the assumptions give

$$S'_j - \kappa' \geq S_j - \kappa, \tag{155}$$

and therefore $q_j(S') \geq q_j(S)$. Together with $U_A(S') > U_A(S)$, this implies

$$\sum_{i=1}^{N} q_i(S') > \sum_{i=1}^{N} q_i(S), \tag{156}$$

which contradicts the common budget constraint. Hence $U_A(S') \leq U_A(S)$. □

Connection to the D-channel. Let $A_t = \{i : B_{i,t-1} = 1\}$ be the coordinates flipped in the previous step. The expected immediate undo mass is

$$U_{\mathrm{undo}}(S_t) = \sum_{i \in A_t} q_i(S_t). \tag{157}$$

In a local oscillatory regime, the D-channel increases $\gamma_{t,\mathrm{eff}}$ and decreases $\nu_{t,\mathrm{eff}}$. Under the local score-ordering condition

$$S'_i \leq S_i, \quad i \in A_t,\ S'_j \geq S_j, \quad j \notin A_t, \tag{158}$$

Proposition 2 gives

$$U_{\mathrm{undo}}(S') \leq U_{\mathrm{undo}}(S). \tag{159}$$

This gives a conditional interpretation of undo suppression at the proposal level. The implemented D-channel does not automatically enforce the score-ordering condition, and a gradient reversal is not the same event as an immediate bit reversal. The proposition establishes neither reduced undo on every trajectory nor improved final objective whenever undo mass decreases.

D.7 Temperature-driven escape-set coverage

This subsection bounds coordinate probabilities for a fixed score vector and feasible expected budget. Increasing temperature tightens the lower and upper envelopes below without changing $\sum_i q_i = d$. A monotone envelope is not, by itself, a monotonicity theorem for every actual escape event under the full adaptive controller.

Proposition 3 (Temperature-induced anti-concentration)**.** *Fix a score vector S, temperature $\tau > 0$, and feasible clipped budget $d \in [N\varepsilon_L, N(1-\varepsilon_L)]$. Let $q_i = q_i(S, \tau, d)$ be the calibrated probabilities satisfying*

$$\sum_{i=1}^{N} q_i = d. \tag{160}$$

Let the score range be

$$R = \max_i S_i - \min_i S_i. \tag{161}$$

Then every coordinate satisfies

$$q_i \geq p_{\min}(\tau, d, R) := \frac{d}{d + (N-d)\exp(R/\tau)} \tag{162}$$

and

$$q_i \leq p_{\max}(\tau, d, R) := \frac{d\exp(R/\tau)}{N - d + d\exp(R/\tau)}. \tag{163}$$

Moreover, $p_{\min}(\tau, d, R)$ is nondecreasing in τ, and $p_{\max}(\tau, d, R)$ is nonincreasing in τ.

Proof. Let

$$\ell_i = \operatorname{clip}\left((S_i - \kappa)/\tau, -L_{\max}, L_{\max}\right),\ q_i = \sigma(\ell_i). \tag{164}$$

Since the clipping map is monotone and 1-Lipschitz,

$$\ell_j - \ell_i \leq \frac{S_j - S_i}{\tau} \leq \frac{R}{\tau} \tag{165}$$

whenever $S_j \geq S_i$; the same upper bound holds trivially when $S_j < S_i$. Hence, for all i, j,

$$\frac{q_j/(1-q_j)}{q_i/(1-q_i)} = \exp(\ell_j - \ell_i) \leq \exp(R/\tau). \tag{166}$$

Let $r = \exp(R/\tau) \geq 1$. Then

$$q_j \leq \frac{rq_i}{1 - q_i + rq_i},\ \forall j. \tag{167}$$

Using the budget constraint,

$$d = \sum_{j=1}^{N} q_j \leq N\frac{rq_i}{1 - q_i + rq_i}. \tag{168}$$

Solving for q_i gives

$$q_i \geq \frac{d}{d + (N-d)r} = \frac{d}{d + (N-d)\exp(R/\tau)}. \tag{169}$$

This proves the lower bound.

For the upper bound, apply the same lower-bound argument to the non-flip probabilities $1 - q_i$, whose total mass is $N - d$ and whose log-odds range is also bounded by R/τ. This gives

$$1 - q_i \geq \frac{N-d}{N - d + dr}. \tag{170}$$

Equivalently,

$$q_i \leq \frac{dr}{N - d + dr} = \frac{d\exp(R/\tau)}{N - d + d\exp(R/\tau)}. \tag{171}$$

Since $r = \exp(R/\tau)$ is nonincreasing in τ, $p_{\min}$ is nondecreasing in τ and $p_{\max}$ is nonincreasing in τ. □

Proposition 4 (Marginal escape-set coverage)**.** *For any nonempty set $A \subseteq [N]$, the calibrated Bernoulli proposal satisfies*

$$\mathbb{P}(B \cap A \neq \emptyset \mid \mathcal{Y}_t, \xi_t) = 1 - \prod_{i \in A}(1 - q_i) \geq 1 - (1 - p_{\min}(\tau, d, R))^{|A|}. \tag{172}$$

Consequently,

$$\mathbb{P}(B \cap A \neq \emptyset \mid \mathcal{Y}_t, \xi_t) \geq 1 - \exp\left(-|A| p_{\min}(\tau, d, R)\right). \tag{173}$$

Proof. The first identity follows from independence of the Bernoulli coordinates. By Proposition 3, $q_i \geq p_{\min}(\tau, d, R)$ for all i, so

$$\prod_{i \in A}(1 - q_i) \leq (1 - p_{\min}(\tau, d, R))^{|A|}. \tag{174}$$

The exponential bound follows from $1 - u \leq \exp(-u)$. □

Proposition 5 (Exact escape-set success)**.** *Suppose x is a one-bit local optimum,*

$$g_i(x) \le 0,\ i \in [N], \tag{175}$$

but there exists an escape set $S^\star \subseteq [N]$ with $|S^\star| = m^\star$ and margin $\Gamma > 0$ such that

$$H(x^{S^\star}) \le H(x) - \Gamma. \tag{176}$$

Let q be a calibrated Bernoulli proposal with score range R, temperature τ, and budget d. Then

$$\mathbb{P}\left(H(x^B) \le H(x) - \Gamma \mid \mathcal{Y}_t, \xi_t\right) \ge p_{\min}(\tau, d, R)^{m^\star} \left(1 - p_{\max}(\tau, d, R)\right)^{N - m^\star}. \tag{177}$$

Proof. Consider the event that the Bernoulli proposal flips exactly the coordinates in $S^\star$ and no others:

$$\mathcal{E}_{S^\star} = \{B_i = 1\ \forall i \in S^\star,\ B_j = 0\ \forall j \notin S^\star\}. \tag{178}$$

By independence,

$$\mathbb{P}(\mathcal{E}_{S^\star}) = \prod_{i \in S^\star} q_i \prod_{j \notin S^\star} (1 - q_j). \tag{179}$$

Using Proposition 3,

$$q_i \ge p_{\min}(\tau, d, R),\ 1 - q_j \ge 1 - p_{\max}(\tau, d, R). \tag{180}$$

Therefore,

$$\mathbb{P}(\mathcal{E}_{S^\star}) \ge p_{\min}(\tau, d, R)^{m^\star} \left(1 - p_{\max}(\tau, d, R)\right)^{N - m^\star}. \tag{181}$$

On $\mathcal{E}_{S^\star}$, $x^B = x^{S^\star}$, and hence

$$H(x^B) = H(x^{S^\star}) \le H(x) - \Gamma. \tag{182}$$

The claim follows. □

Connection to the I-channel. The calibrated budget constraint fixes

$$\sum_{i=1}^{N} q_i = d. \tag{183}$$

Therefore the I-channel does not increase the prescribed expected number of flips. If temperature is increased while the score vector and budget are held fixed, Proposition 3 implies that $p_{\min}$ increases and $p_{\max}$ decreases. Proposition 4 shows that the probability of touching an escape set increases in this conservative lower-bound sense, while Proposition 5 gives a sufficient exact-set success probability for an energy-improving multi-bit move. The full I update also changes noise and hence the score range, so these comparisons do not directly compare its complete on/off kernels. Touching an escape set need not improve energy; the exact-set result is a sufficient event bound and may be small in high dimensions.

D.8 Pairwise pseudo-Boolean energies

This subsection records how the residual constants in Theorem 1 specialize to the pairwise energies used in the experiments. It also gives an exact one-step expression for pairwise energies, which is sharper than the worst-case residual bound when the signs of the pairwise interactions are informative.

Consider a pairwise pseudo-Boolean energy

$$H(x) = H_0 + \sum_{i=1}^{N} h_i x_i + \sum_{1 \le i < j \le N} J_{ij} x_i x_j. \tag{184}$$

For notational convenience, set $J_{ij} = J_{ji}$ and $J_{ii} = 0$. The one-bit gain is

$$g_i(x) = \chi_i(x) \left(h_i + \sum_{j \ne i} J_{ij} x_j \right). \tag{185}$$

If $R \subseteq [N] \setminus \{i\}$, then

$$g_i(x^R) - g_i(x) = \chi_i(x) \sum_{j \in R} J_{ij} \left[(1 - x_j) - x_j \right]. \tag{186}$$

Since

$$|(1 - x_j) - x_j| = 1, \tag{187}$$

we have

$$\left| g_i(x^R) - g_i(x) \right| \leq \sum_{j \in R} |J_{ij}|. \tag{188}$$

Thus the residual condition in Theorem 1 holds with

$$\epsilon_{ij}(x) = |J_{ij}|. \tag{189}$$

Exact expected decrease. For pairwise energies, the expected one-step decrease under independent flips has the exact form

$$\mathbb{E}\left[H(x) - H(x^B) \mid \mathcal{Y}_t, \xi_t \right] = \sum_{i=1}^{N} q_i g_i(x) - \sum_{1 \leq i < j \leq N} J_{ij} \chi_i(x) \chi_j(x) q_i q_j. \tag{190}$$

Indeed, $x_i^B = x_i - \chi_i(x) B_i$, and expanding the pairwise energy gives

$$H(x) - H(x^B) = \sum_{i=1}^{N} B_i g_i(x) - \sum_{1 \leq i < j \leq N} J_{ij} \chi_i(x) \chi_j(x) B_i B_j. \tag{191}$$

Taking conditional expectation and using

$$\mathbb{E}[B_i \mid \mathcal{Y}_t, \xi_t] = q_i, \ \mathbb{E}[B_i B_j \mid \mathcal{Y}_t, \xi_t] = q_i q_j \tag{192}$$

proves the formula.

For maximum independent set with penalty energy

$$H_{\mathrm{MIS}}(x) = -\sum_i w_i x_i + \lambda \sum_{(i,j) \in E} x_i x_j, \tag{193}$$

the pairwise coupling is

$$J_{ij} = \lambda \mathbf{1}\{(i,j) \in E\}, \tag{194}$$

and hence $\epsilon_{\max} = \lambda$. Equivalently, in the matrix form $H_{\mathrm{MIS}}(x) = -w^\top x + \lambda x^\top A x / 2$, the factor $1/2$ compensates for the double counting in $x^\top A x$, so the unordered-edge coupling is still $J_{ij} = \lambda A_{ij}$.

Maximum clique has the same penalty form on the complement conflict graph:

$$H_{\mathrm{Cl}}(x) = -\sum_i w_i x_i + \lambda \sum_{(i,j) \notin E} x_i x_j. \tag{195}$$

Thus

$$J_{ij} = \lambda \mathbf{1}\{(i,j) \notin E\}, \tag{196}$$

and again $\epsilon_{\max} = \lambda$.

For max cut, using the energy

$$H_{\mathrm{Cut}}(x) = -\sum_{(i,j) \in E} \left(x_i + x_j - 2 x_i x_j \right), \tag{197}$$

we have

$$J_{ij} = 2\mathbf{1}\{(i,j) \in E\}, \tag{198}$$

and therefore $\epsilon_{\max} = 2$. This is consistent with the matrix form $H_{\mathrm{Cut}}(x) = x^\top A x - \mathbf{1}^\top A x$ for a symmetric adjacency matrix, since $x^\top A x$ counts every undirected edge twice.

Consequently, the simplified residual term

$$\frac{\epsilon_{\max}}{2} d^2 \tag{199}$$

equals

$$\frac{\lambda}{2} d^2 \tag{200}$$

for the MIS and maximum clique penalty energies, and equals

$$d^2 \tag{201}$$

for max cut. When available, the exact pairwise expression above should be used for diagnostics because it keeps the sign of the interaction term instead of replacing it by the worst-case magnitude $|J_{ij}|$.

D.9 Finite-horizon archive consequence

This result is a generic consequence of uniform conditional hitting bounds, not a PID-specific convergence-rate comparison. Let $\mathcal{F}_t$ contain the complete search history before the proposals at time t, including the archives, and define

$$\mathcal{A}_h = \{x : H(x) \le h\},\ E_t^\star = \min_{0 \le s \le t} \min_{1 \le k \le K} H(x_s^{(k)}),\ \tau_h = \inf\{t : E_t^\star \le h\}. \tag{202}$$

Proposition 6 (Finite-horizon archive hitting under deterministic bounds)**.** *For each t, k, let $\underline{p}_{t,k} \in [0,1]$ be a deterministic number satisfying, almost surely on $\{\tau_h > t\}$,*

$$\mathbb{P}\Big(x_{t+1}^{(k)} \in \mathcal{A}_h \mid \mathcal{F}_t\Big) \ge \underline{p}_{t,k}. \tag{203}$$

If the K next-state proposals are conditionally independent given $\mathcal{F}_t$, then

$$\mathbb{P}(\tau_h \le T) \ge 1 - \prod_{t=0}^{T-1} \prod_{k=1}^{K} (1 - \underline{p}_{t,k}). \tag{204}$$

In particular, a deterministic uniform bound $\underline{p}_{t,k} \ge p$ yields $\mathbb{P}(\tau_h \le T) \ge 1 - (1-p)^{KT}$.

Proof. Set $a_t = \prod_k (1 - \underline{p}_{t,k})$. Conditional independence gives

$$\mathbb{P}(\tau_h > t+1 \mid \mathcal{F}_t) \le \mathbf{1}_{\{\tau_h > t\}} a_t. \tag{205}$$

Since a_t is deterministic, taking expectations gives $\mathbb{P}(\tau_h > t+1) \le a_t \mathbb{P}(\tau_h > t)$. Induction and $\mathbb{P}(\tau_h > 0) \le 1$ yield $\mathbb{P}(\tau_h > T) \le \prod_{t=0}^{T-1} a_t$. Taking complements proves the claim. □

The deterministic bounds must hold uniformly over the relevant unhit histories. Conditional independence given $(\mathcal{Y}_t, \xi_t)$ does not by itself imply independence after integrating out shared proposal noise. The proposition therefore requires its stated history-level assumption; it does not infer it solely from the factorized conditional mask distribution. History-dependent random lower bounds cannot be substituted into the product formula by the same induction.

E Hyperparameters

Table 5 reports the benchmark-specific configurations used in our experiments. Here, K is the number of parallel search chains, T is the finite search horizon, and d specifies the prescribed expected number of bit flips per proposal. The initial temperature is controlled by τ_0, while k_p, k_i, and k_d denote the gains of the proportional, integral, and derivative feedback channels, respectively. The momentum backbone is parameterized by the mass m, the base friction coefficient γ, the velocity-update scale η, and the stochastic noise scale σ; λ_v controls the contribution of the velocity state to the final proposal score. Each row gives the configuration used for the corresponding benchmark, rather than a universally optimal parameter setting. The search horizon and expected flip budget are adjusted to the problem scale, while the controller and momentum parameters determine how trajectory feedback regulates the stateful search dynamics before Bernoulli calibration. Unless otherwise stated, the listed values are kept fixed within the corresponding benchmark configuration.

Table 5: Reported PID-LD configurations. Values are retained from the experimental record.

Problem	Benchmark	K	T	d	β	τ_0	k_p	k_i	k_d	m	γ	η	σ	λ_v
MIS	RB-[200–300]	200	300	5	1.02	0.01	0.78	0.01	0.01	1.0	3.0	0.05	0.01	0.06
	RB-[800–1200]	200	500	5	1.02	0.01	1.00	0.01	0.01	1.0	3.0	0.05	0.01	0.06
	ER-[700–800]	200	500	20	1.001	0.012	1.00	0.01	0.10	0.8	3.5	0.03	0.01	0.01
	ER-[9000–11000]	200	5000	20	1.001	0.01	1.00	0.01	0.01	0.8	3.0	0.03	0.01	0.05
MaxCl	RB-[200–300]	200	100	2	1.02	4.0	1.00	0.50	0.01	1.0	3.0	0.05	0.00	0.05
	RB-[800–1200]	200	500	2	1.02	4.0	1.00	0.50	0.01	1.0	3.0	0.05	0.01	0.05
MCut	BA-[200–300]	200	200	20	1.02	5.0	1.05	0.20	0.01	0.8	3.8	0.03	0.01	0.01
	BA-[800–1200]	200	500	20	1.02	5.0	1.05	0.16	0.01	0.8	3.8	0.03	0.01	0.01

d denotes `num_d`; η denotes `pidld_vscale`; σ denotes `pidld_noise_scale`; and λ_v denotes `pidld_velocity_score_scale`. All rows use $\tau_{\min} = 10^{-3}$.

F Hardware and Software Devices

Experiments are conducted on a Linux workstation using an H100 GPU and an Intel(R) Xeon(R) Platinum 8468 CPU, with programs implemented in *PyTorch*.

G Additional Experimental Analyses

This appendix presents complementary analyses of PID-LD on MIS RB-[200–300], covering paired objective comparisons, one-factor hyperparameter sensitivity, and trajectory-level measurements. The objective is the independent set size, for which larger values indicate better solutions.

G.1 Paired objective comparisons

Table 6 reports paired objective comparisons on MIS RB-[200–300]. Each configuration is evaluated using five independent random seeds, with methods compared on matched graph instances under the same evaluation protocol.

For a comparison between methods A and B, let $\Delta_{j,s} = J_{A,j,s} - J_{B,j,s}$, where j indexes the matched graph instance and s the random seed. The reported δ denotes the mean paired objective difference under the evaluation protocol, with positive values favoring the first method. We report 95% confidence intervals and Holm-adjusted p-values. Holm correction is applied within the planned comparison family reported in this table.

Table 6: Paired objective comparisons on MIS RB-[200–300] under the five-seed evaluation protocol. For each comparison, δ is the mean paired difference between the first and second methods. Positive values favor the first method. Confidence intervals are 95%, and p-values are Holm-adjusted.

Comparison	Mean difference δ	95% confidence interval	Holm-adjusted p
PID-LD vs RLSA	$+0.1060$	$[+0.082, +0.110]$	3×10^{-5}
PID-LD vs FLD-EG	$+0.0560$	$[+0.051, +0.084]$	1.5×10^{-4}
PID-LD vs w/o I+D	$+0.0142$	$[+0.007, +0.022]$	1.4×10^{-3}
w/o D vs w/o I+D	-0.0148	$[-0.019, -0.007]$	3.2×10^{-2}
w/o D vs RLSA	$+0.0770$	$[+0.064, +0.089]$	3×10^{-5}
w/o D vs FLD-EG	$+0.0270$	$[+0.015, +0.037]$	1.2×10^{-4}

The comparisons against RLSA and FLD-EG evaluate the complete PID-LD system against external baselines, whereas comparisons involving channel removal quantify differences among matched PID-LD configurations. Accordingly, the full-system differences against external baselines should not be interpreted as the isolated contribution of PID feedback. In particular, the comparison between PID-LD and the variant with both D- and I-feedback disabled measures the joint incremental effect of trajectory-dependent regulation while retaining the same momentum backbone and calibrated proposal mechanism. The channel-removal results also indicate non-additive interactions between the D- and I-feedback mechanisms, so individual ablations should be interpreted conditionally rather than as additive component contributions.

G.2 ONE-FACTOR HYPERPARAMETER SENSITIVITY

We evaluate sensitivity to eleven parameters by varying one parameter at a time around its listed default on MIS RB-[200–300], with the remaining parameters held fixed. Results are averaged over five random seeds. Each δ in Table 7 is the mean objective under the perturbed setting minus that under the default configuration used in this sensitivity analysis. Negative values therefore indicate a decrease in solution quality. Within each row, the three objective changes correspond, in order, to the three tested values. Here λ_v controls only the contribution of the velocity state to the proposal score; changing λ_v does not disable the velocity recursion itself.

Table 7: One-factor hyperparameter sensitivity on MIS RB-[200–300]. Objective changes are measured relative to the default configuration of the sensitivity analysis and averaged over five random seeds.

Parameter	Default	Tested values	Objective changes δ
k_p	0.78	$0.50/1.00/1.50$	$-0.0104/\ \ 0.0000/-0.0052$
k_i	0.01	$0.00/0.05/0.10$	$-0.0360/-0.0156/-0.0208$
k_d	0.01	$0.00/0.05/0.10$	$-0.0290/-0.0052/-0.0156$
λ_v	0.06	$0.00/0.03/0.10$	$-0.1010/-0.0312/\ \ 0.0000$
τ_0	0.01	$0.005/0.02/0.05$	$-0.0156/-0.0156/-0.0312$
γ	3.0	$1.0/5.0/10.0$	$-0.0417/-0.0312/-0.0677$
m	1.0	$0.5/2.0/5.0$	$-0.0104/-0.0052/-0.0208$
η	0.05	$0.01/0.03/0.10$	$-0.0156/-0.0104/-0.0312$
σ	0.01	$0.005/0.02/0.05$	$-0.0208/-0.0104/-0.0104$
β	1.02	$1.01/1.05/1.10$	$-0.0156/-0.0104/-0.0156$
d	5	$3/7/10$	$-0.0104/-0.0260/-0.1094$

Within the tested ranges, the expected flip budget d, velocity–score coupling λ_v, and friction coefficient γ produce the largest objective changes. Increasing d from 5 to 10 gives the largest decrease in the table, 0.1094. Setting $\lambda_v = 0$ decreases the objective by 0.1010. This perturbation should not be confused with the "w/o momentum" ablation in Table 4. The sensitivity experiment with $\lambda_v = 0$ retains the velocity state and its stochastic dynamics but removes its direct contribution to the proposal score. In contrast, the "w/o momentum" variant removes the velocity recursion itself and disables all velocity-dependent proposal pathways. The similar objective changes therefore arise from two distinct interventions and should not be interpreted as identical configurations. Setting $k_i = 0$ and $k_d = 0$ gives decreases of 0.0360 and 0.0290, respectively, matching the corresponding I- and D-channel removals in Table 4. Increasing γ from 3.0 to 10.0 decreases the objective by 0.0677.

G.3 TRAJECTORY-LEVEL MEASUREMENTS

We examine search behavior on MIS RB-[200–300] using five random seeds. Metrics are aggregated within each trajectory before comparison across trajectories. Table 8 complements the objective comparisons with measurements of immediate reversals, progress after stagnation, proposal entropy, and calibration accuracy.

The immediate undo rate is the fraction of coordinates flipped at iteration t that are flipped back at iteration $t+1$. Plateau escape is measured as an archive improvement within 25 iterations following a plateau of at least 25 iterations. This measurement uses a patience parameter of 25 and a burn-in of $\rho_I T = 60$ iterations; the baseline event rate is 1.26%. The change in this event rate is expressed in percentage points (pp).

Proposal entropy is measured in nats during stagnation windows, with δ defined as the on-minus-off counterfactual difference. Calibration accuracy is measured by the mean absolute residual $|\sum_i q_i - d_{\text{eff}}|$, with $d_{\text{eff}} = 5$.

The trajectory measurements show a lower immediate undo rate and a higher frequency of archive improvement after stagnation in the corresponding comparisons. The entropy difference is negative for the specified on-minus-off comparison and therefore should not be interpreted as a uniform entropy increase. The calibration residual measures numerical agreement with the prescribed expected flip budget; it is not, on its own, a direct measure of coordinate-scale correction by the P-channel.

Table 8: Trajectory-level measurements on MIS RB-[200–300]. The plateau-escape change is expressed in percentage points, and the entropy change is the on-minus-off difference in nats.

Channel	Metric	Difference δ	p-value
D	Immediate undo rate	-0.075	$< 10^{-5}$
I	Plateau escape: archive improvement within the specified window	$+0.58$ pp	$< 10^{-5}$
I	Proposal entropy during stagnation windows (nats)	-0.750	$< 10^{-5}$
P	Pre-calibration expected-budget mismatch $\lvert \sum_i \widetilde{q}_i - d_{\mathrm{eff}} \rvert$	-0.102	$< 10^{-5}$

These diagnostics describe different aspects of the search process and should be distinguished from improvements in the final objective or computational efficiency.