Regime-Adaptive Bayesian Optimization

via Dirichlet Process Mixtures of Gaussian Processes

RAMBO

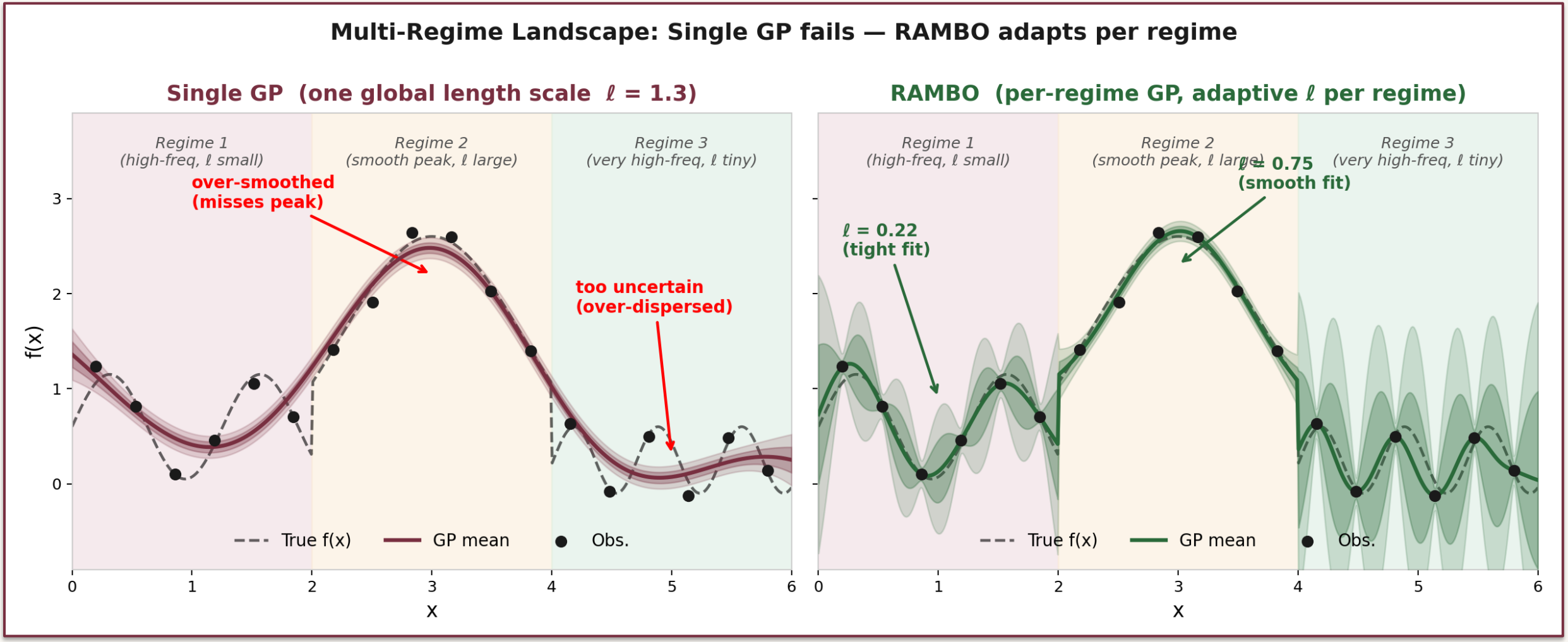
ICML 2026 Poster Recording

Presenter: Yan Zhang

Problem: Standard GP Fails on Multi-Regime Landscapes

Real-world objectives have abrupt transitions — one global length scale is not enough

- ☒ Single GP uses one global length scale $\ell \rightarrow$ either over-smooths sharp transitions or hallucinates noise in smooth regions
- ☑ RAMBO assigns each regime its own GP with locally-optimized $\ell_k, \sigma^2_{f,k}, \sigma^2_{n,k}$



Method: RAMBO

Regime-Adaptive Mixture Bayesian Optimization

Key Insight

Replace the single monolithic GP surrogate with a mixture of GPs, where each component captures a distinct local regime.

DPMM-GP Surrogate

DPMM-GP with a CRP prior automatically discovers K regimes — no need to pre-specify K!

$$z_i \sim CRP(\alpha), f_k \mid \theta_k \sim GP(0, k_{\theta_k})$$

Collapsed Gibbs Sampling

Resamples regime assignments via CRP prior × GP predictive likelihood.

$$p(z_i = k \mid \cdot) \propto p_{CRP}(k \mid z_{-i}, \alpha) \cdot p_{GP}(y_i \mid x_i, D_{k,-i}, \theta_k)$$

Adaptive α Scheduling

Controls regime creation rate. Sparse early → detailed later. Matches data availability.

$$\alpha_t = \alpha_0 \cdot \sqrt{t} / \log(t + e)$$

RAMBO = DPMM-GP surrogate + Collapsed Gibbs inference + Mixture EI + Adaptive α-scheduling

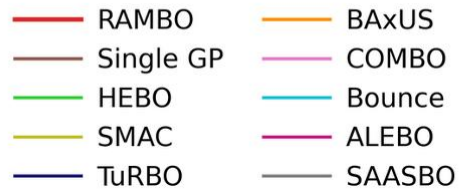
Mixture EI: $EI_{mix}(x) = \Sigma_k w_k(x) \cdot EI_k(x)$

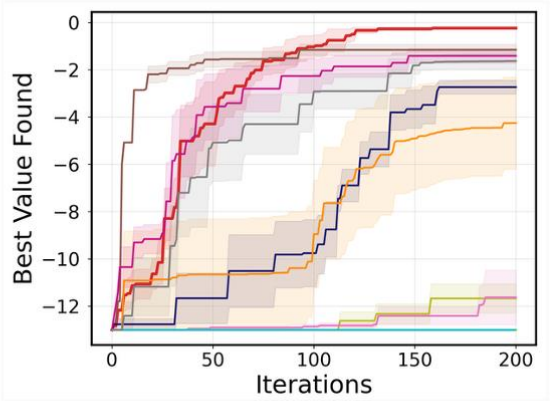
Experiments: Synthetic & Scientific Benchmarks

Levy & Schwefel functions in 6D and 10D

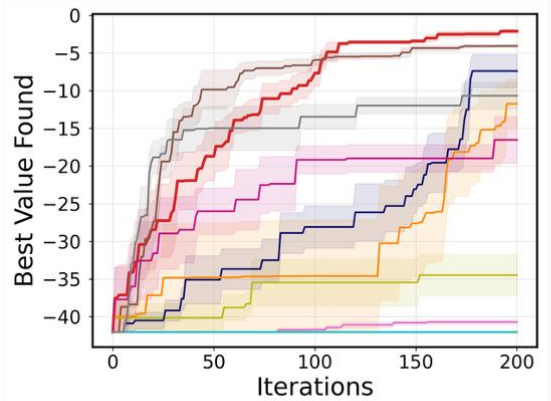
Molecular conformer (12D), drug discovery (50D), fusion reactor (81D)

RAMBO consistently matches or outperforms state-of-the-art baselines

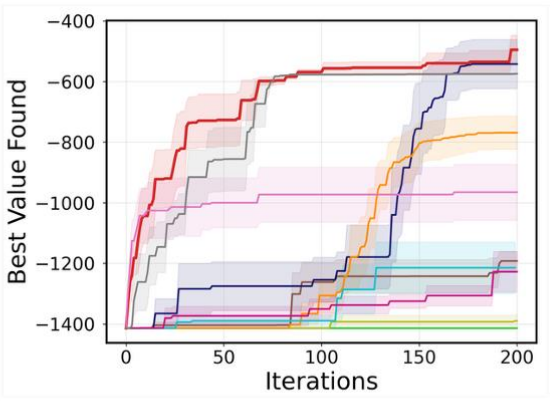




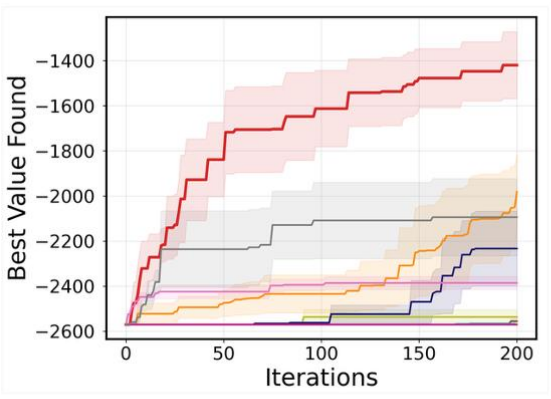
(a) Levy(6D)



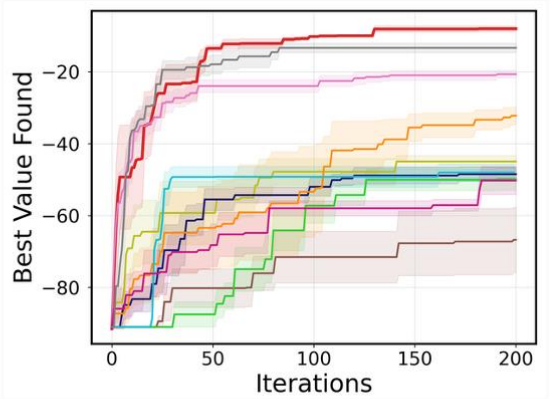
(b) Levy(10D)



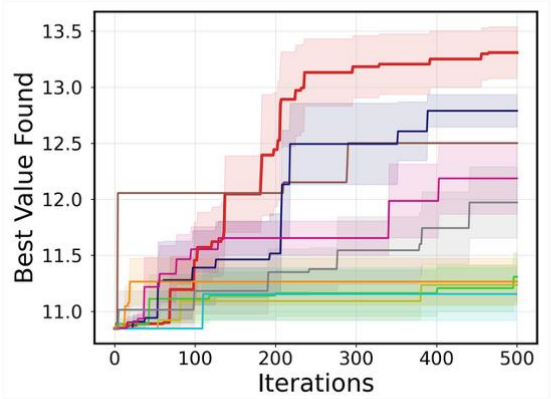
(c) Schwefel(6D)



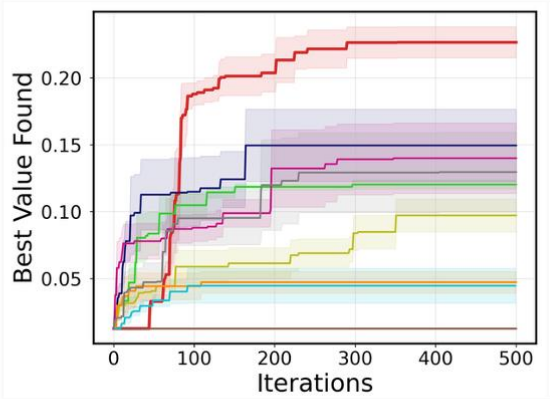
(d) Schwefel(10D)



(e) Molecular



(f) Drug Discovery



(g) ConStellaration

Thank you!

Please visit our poster for more details.