Component — Optimizer (KL-bounded Mirror Descent)

September 26, 2025

Objective (Example)

Increase variance while preferring RTP/Hit near targets:

$$\mathcal{L}(p) = \lambda_\mu\big(\mu(p) - \mu^\star\big)^2 + \lambda_h\big(h(p) - h^\star\big)^2 + \lambda_v\big(\mathrm{Var}(p) - v^\star\big)^2. \tag{1}$$

Mirror-Descent Update on the Simplex

With entropic mirror map $\psi(p) = \sum_i p_i \log p_i$, the update is

$$p_i^+ \;\propto\; p_i \exp\big(-\eta\, g_i\big), \qquad g_i = \frac{\partial \mathcal{L}}{\partial p_i}, \tag{2}$$

followed by normalization $\sum_i p_i^+ = 1$.

KL Trust-Region

Enforce small, auditable steps:

$$D_{\mathrm{KL}}(p^+ \,\|\, p) = \sum_i p_i^+ \log \frac{p_i^+}{p_i} \le \delta. \tag{3}$$

Projection to Guardrails

If bands are violated, solve the Bregman projection

$$\min_{x \in \Delta^k} \quad D_{\mathrm{KL}}(x \| p^+) \tag{4}$$

$$\text{s.t.} \quad L_\mu \le \sum_i r_i x_i \le U_\mu, \qquad L_h \le \sum_i \mathbf{1}[r_i > 0]\, x_i \le U_h, \tag{5}$$

yielding the exponential-family form

$$x_i \;\propto\; p_i^+ \exp\big(-\lambda_0 - \lambda_\mu r_i - \lambda_h \mathbf{1}[r_i > 0]\big). \tag{6}$$

Stopping Criteria

Small KL; diminishing Var gains over M steps; guardrails satisfied in the last M steps.