

Component — Constraints (Guardrails Projection)

September 26, 2025

Linear Bands (Invariants)

$$L_\mu \leq \mu(p) = \sum_i r_i p_i \leq U_\mu, \quad L_h \leq h(p) = \sum_i \mathbf{1}[r_i > 0] p_i \leq U_h. \quad (1)$$

Combined with $p \in \Delta^k$ and caps encoded by r .

Feasibility

The feasible set

$$\mathcal{C} = \left\{ p \in \Delta^k \mid L_\mu \leq \sum_i r_i p_i \leq U_\mu, \quad L_h \leq \sum_i \mathbf{1}[r_i > 0] p_i \leq U_h \right\} \quad (2)$$

is nonempty if the bands intersect the convex hull of $\{(r_i, w_i)\}$ under convex combinations; check via a small LP in practice.

KL/Bregman Projection

For a given $q \in \Delta^k$,

$$\min_{x \in \mathcal{C}} D_{\text{KL}}(x \| q) \quad \Rightarrow \quad x_i \propto q_i \exp(-\lambda_0 - \lambda_\mu r_i - \lambda_h w_i). \quad (3)$$

Priority

When both bands cannot be achieved within a δ -small change, prioritize hard legal bounds (e.g., RTP lower bound) and choose the nearest feasible hit-rate.