Reward Calibration Beyond the Convex Hull: Depth-Based Feasibility and Regularized Exponential Tilting for Generative Models

Manoj Saravanan **Rohit Kumar Salla**
Virginia Tech, Department of Electrical and Computer Engineering Blacksburg, VA, USA

Abstract

We study *constraint calibration* of a base generative distribution P_0 via KL-projection onto expectation constraints. Recent work proposes a *reward*-style surrogate that approximates the maximum-entropy (exponential-tilting) solution by replacing expectations under P_0 with Monte Carlo averages (Smith et al., 2025). However, the resulting empirical maximum-entropy problem is only well-defined when the target moment vector lies in the interior of the convex hull of sampled statistics, an event that can fail with high probability in high-dimensional or rare-event regimes (Smith et al., 2025). We quantify this phenomenon by reducing reward feasibility to convex-hull membership probabilities and leveraging sharp depth-based inequalities (Hayakawa et al., 2023) together with Wendel-type phase transitions (Tikhomirov, 2023). Motivated by these limits, we propose a ridge-regularized exponential-tilting estimator that is always defined and satisfies an exact residual identity controlling constraint mismatch. We prove finite-sample bounds on parameter error, moment violation, and KL deviation, and validate the predicted feasibility transitions and bias–variance tradeoffs in synthetic experiments.

1 INTRODUCTION

Modern generative models are increasingly deployed in settings where *distribution-level constraints* are first-class requirements: class proportions in conditional generation, satisfaction rates of safety filters, composition constraints in molecular design, or matching of summary statistics in scientific simulators. A principled formalization is *information projection*: given a base distribution P_0 on $\mathcal{X}$, a statistic $T : \mathcal{X} \to \mathbb{R}^m$, and a target moment $\mu \in \mathbb{R}^m$, seek the closest distribution satisfying the constraints

$$Q^\star \in \operatorname*{arg\,min}_{Q \ll P_0} \ \mathrm{KL}(Q \| P_0) \quad \text{s.t.} \quad \mathbb{E}_Q[T(X)] = \mu. \quad (1)$$

When feasible under standard interiority conditions, $Q^\star$ is an exponential tilt of P_0 and can be characterized by a concave dual in an exponential-family parameter η (Dudík et al., 2007; Smith et al., 2025).

A popular practical surrogate is *reward calibration*: approximate the dual using Monte Carlo samples from P_0 and solve an empirical max-entropy problem (Smith et al., 2025). A key failure mode, however, is *geometric*: the unregularized empirical problem is well-defined only when the target μ lies in the convex hull of the sampled statistics. In high-dimensional or rare-event regimes, this convex-hull event can fail with high probability, leading to non-attainment/divergence of the empirical dual.

Workshop “Towards Trustworthy Predictions: Theory and Applications of Calibration for Modern AI” at AISTATS 2026, Tangier, Morocco.

Contributions. We give a learning-theoretic account of this phenomenon and a principled fix: (i) we reduce empirical reward feasibility to a convex-hull membership event and quantify its sample complexity via half-space depth (Hayakawa et al., 2023); (ii) we state a sharp Wendel-type phase transition under symmetry (Tikhomirov, 2023); (iii) we propose a ridge-regularized exponential-tilting estimator that is always well-defined and obeys an exact residual identity; (iv) we prove finite-sample bounds on parameter stability, population moment error, and KL deviation under bounded statistics; and (v) we validate the predicted transitions and bias–variance tradeoff in synthetic experiments.

2 SETUP: CONSTRAINT CALIBRATION AS KL PROJECTION

Let P_0 be a base distribution on $\mathcal{X}$ and $T : \mathcal{X} \to \mathbb{R}^m$ a constraint statistic. Given $\mu \in \mathbb{R}^m$, the calibration problem is the KL projection

$$Q^\star \in \operatorname*{arg\,min}_{Q \ll P_0} \mathrm{KL}(Q \| P_0) \quad \text{s.t.} \quad \mathbb{E}_Q[T(X)] = \mu. \tag{2}$$

Define the log-partition

$$A(\eta) \;:=\; \log \mathbb{E}_{P_0}\big[\exp(\eta^\top T(X))\big], \qquad \eta \in \mathbb{R}^m, \tag{3}$$

and the exponential tilt (when $A(\eta) < \infty$)

$$\frac{dQ_\eta}{dP_0}(x) \;=\; \exp\big(\eta^\top T(x) - A(\eta)\big). \tag{4}$$

Under standard regularity (e.g. μ in a suitable interior region), the solution to (2) satisfies $Q^\star = Q_{\eta^\star}$ where

$$\begin{aligned} \eta^\star &\in \operatorname*{arg\,max}_{\eta \in \mathbb{R}^m} \; \eta^\top \mu - A(\eta), \\ \nabla A(\eta^\star) &= \mathbb{E}_{Q_{\eta^\star}}[T(X)] = \mu. \end{aligned} \tag{5}$$

For modern generative models, evaluating $A(\eta)$ and $\nabla A(\eta)$ is typically intractable, motivating empirical surrogates.

3 EMPIRICAL REWARD CALIBRATION AND FEASIBILITY

Let $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} P_0$ and $t_i := T(X_i) \in \mathbb{R}^m$. Write $P_n := \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$. Any $Q \ll P_n$ is of the form $Q_w = \sum_i w_i \delta_{X_i}$ for some $w \in \Delta_n$. The empirical KL projection is

$$\min_{w \in \Delta_n} \; \mathrm{KL}(Q_w \| P_n) \quad \text{s.t.} \quad \sum_{i=1}^n w_i t_i = \mu, \tag{6}$$

with $\mathrm{KL}(Q_w \| P_n) = \sum_{i=1}^n w_i \log(n w_i)$. Define the empirical log-partition

$$A_n(\eta) \;:=\; \log\Big(\frac{1}{n}\sum_{i=1}^n \exp(\eta^\top t_i)\Big). \tag{7}$$

When (6) is strictly feasible, the associated dual is

$$\hat\eta \in \operatorname*{arg\,max}_{\eta \in \mathbb{R}^m} \; \eta^\top \mu - A_n(\eta), \tag{8}$$

with induced weights

$$\hat w_i(\eta) \;:=\; \frac{\exp(\eta^\top t_i)}{\sum_{j=1}^n \exp(\eta^\top t_j)}, \qquad \hat Q_\eta \;:=\; \sum_{i=1}^n \hat w_i(\eta)\, \delta_{X_i}. \tag{9}$$

We focus on when (8) is well-defined.

Lemma 1 (Feasibility ⇔ convex hull)**.** *The empirical problem* (6) *is feasible if and only if*

$$\mu \in \mathrm{conv}\{t_1, \dots, t_n\}. \tag{10}$$

Proofs and well-definedness conditions (relative interior, attainment, uniqueness) are in Appendix B following standard max-entropy duality (Dudík et al., 2007; Smith et al., 2025).

Define the feasibility probability

$$\begin{aligned} p_n(\mu) &:= \mathbb{P}\Big(\mu \in \mathrm{conv}\{Z_1, \dots, Z_n\}\Big), \\ Z_i &:= T(X_i), \quad X_i \sim P_0. \end{aligned} \tag{11}$$

4 FEASIBILITY THEORY: DEPTH AND A PHASE TRANSITION

We quantify feasibility via the halfspace (Tukey) depth of μ under $Z = T(X)$:

$$\alpha(\mu) \;:=\; \inf_{u \in \mathbb{S}^{m-1}} \mathbb{P}\big(u^\top Z \ge u^\top \mu\big)\,. \tag{12}$$

Let $N(\mu) := \min\{n : \; p_n(\mu) \ge 1/2\}$ denote the median sample size for feasibility.

Theorem 1 (Depth controls median feasibility (Hayakawa et al., 2023))**.** *For any m-dimensional random vector Z and any $\mu \in \mathbb{R}^m$ with $\alpha(\mu) > 0$,*

$$\frac{1}{2} \;\le\; \alpha(\mu)\, N(\mu) \;\le\; 3m + 1. \tag{13}$$

Thus, even when $\alpha(\mu) = \Theta(1)$, feasibility generically requires $n = \Omega(m)$ samples; for rare targets with small depth, the requirement scales as $m/\alpha(\mu)$.

Under symmetry, feasibility exhibits a sharp threshold:

Corollary 1 (Wendel formula under symmetry)**.** *Assume $Z - \mu$ is centrally symmetric and in general position (e.g. continuous density). Then*

$$p_n(\mu) \;=\; 1 - 2^{-n+1} \sum_{k=0}^{m-1} \binom{n-1}{k}, \tag{14}$$

which transitions sharply around $n \approx 2m$ (Tikhomirov, 2023).

Proofs are in Appendix C.

5 AN ALWAYS-FEASIBLE REGULARIZED REWARD ESTIMATOR

Sections 3–4 show that the unregularized empirical dual (8) can be ill-posed when $\mu \notin \mathrm{conv}\{t_1, \dots, t_n\}$ (Lemma 1). We remove this convex-hull barrier by adding a ridge penalty to the empirical maximum-entropy dual.

Ridge-regularized empirical dual. Given samples $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} P_0$ and $t_i := T(X_i)$, define A_n as in (7) and solve

$$\hat{\eta}_\lambda \in \arg\max_{\eta \in \mathbb{R}^m} \left\{ \eta^\top \mu - A_n(\eta) - \frac{\lambda}{2} \|\eta\|_2^2 \right\}, \qquad \lambda > 0. \tag{15}$$

As in (9), any η induces softmax weights $\hat{w}(\eta) \in \Delta_n$ and the empirical tilt $\hat{Q}_\eta = \sum_{i=1}^n \hat{w}_i(\eta) \delta_{X_i}$; we write $\hat{Q}_\lambda := \hat{Q}_{\hat{\eta}_\lambda}$.

Theorem 2 (Always-feasible reward calibration and residual identity)**.** *For any $\lambda > 0$, the objective in* (15) *is λ-strongly concave and has a unique maximizer $\hat{\eta}_\lambda$. Moreover,*

$$\mu - \mathbb{E}_{\hat{Q}_\lambda}[T(X)] \; = \; \lambda \, \hat{\eta}_\lambda. \tag{16}$$

Penalized primal view. Ridge regularization in the dual is equivalent to replacing the hard empirical moment constraint by a squared penalty, a standard device in generalized max-entropy estimation (Dudík et al., 2007).

Proposition 1 (Penalized empirical KL projection)**.** *Let $\hat{\eta}_\lambda$ solve* (15) *and let $\hat{w}_\lambda := \hat{w}(\hat{\eta}_\lambda)$ be the corresponding softmax weights* (9)*. Then $\hat{w}_\lambda$ is the unique minimizer of*

$$\min_{w \in \Delta_n} \left\{ \sum_{i=1}^n w_i \log(n w_i) \; + \; \frac{1}{2\lambda} \left\| \sum_{i=1}^n w_i t_i - \mu \right\|_2^2 \right\}. \tag{17}$$

Theorem 2 and Proposition 1 are proved in Appendix D. Identity (16) makes the bias–feasibility tradeoff explicit: increasing λ shrinks $\hat{\eta}_\lambda$ (and hence the distributional shift) while permitting controlled moment mismatch.

6 FINITE-SAMPLE GENERALIZATION GUARANTEES

We compare $\hat{\eta}_\lambda$ to its population analogue. Define the population regularized objective

$$g_\lambda(\eta) \; := \; \eta^\top \mu - A(\eta) - \frac{\lambda}{2} \|\eta\|_2^2, \qquad \eta_\lambda := \arg\max_{\eta \in \mathbb{R}^m} g_\lambda(\eta), \tag{18}$$

and let Q_η denote the population tilt (4). Assume bounded statistics:

$$\|T(X)\|_2 \le B \quad \text{a.s. under } P_0. \tag{19}$$

Then $\|\nabla^2 A(\eta)\|_{\text{op}} \le B^2$ for all η (Appendix A), and both η_λ and $\hat{\eta}_\lambda$ lie in $B_2(R)$ with $R = 2B/\lambda$ (Appendix E.2).

Theorem 3 (Finite-sample stability and calibration error)**.** *Assume* (19) *and $\|\mu\|_2 \le B$. Fix $\lambda > 0$ and $\delta \in (0,1)$ and set $R := 2B/\lambda$. With probability at least $1 - \delta$,*

$$\|\hat{\eta}_\lambda - \eta_\lambda\|_2 \; \le \; 2\sqrt{\frac{\varepsilon_n(R,\delta)}{\lambda}}, \tag{20}$$

where one may take (e.g. Appendix E.3, Corollary 3)

$$\varepsilon_n(R,\delta) := \frac{2BR}{n} + \left(e^{2BR} - 1\right)\sqrt{\frac{m \log(1+2n) + \log(2/\delta)}{2n}}. \tag{21}$$

Moreover,

$$\left\| \mathbb{E}_{Q_{\hat{\eta}_\lambda}}[T(X)] - \mu \right\|_2 \; \le \; \lambda \|\eta_\lambda\|_2 \; + \; B^2 \, \|\hat{\eta}_\lambda - \eta_\lambda\|_2, \tag{22}$$

and

$$\mathrm{KL}(Q_{\hat{\eta}_\lambda} \,\|\, Q_{\eta_\lambda}) \; \le \; \frac{B^2}{2} \, \|\hat{\eta}_\lambda - \eta_\lambda\|_2^2. \tag{23}$$

Full proofs are in Appendix E. Theorem 3 highlights the usual bias–variance structure: the term $\lambda \|\eta_\lambda\|_2$ is a population regularization bias, while $\|\hat{\eta}_\lambda - \eta_\lambda\|_2$ is an estimation term.

7 EXPERIMENTS

We report two synthetic experiments validating (i) feasibility phase transitions and (ii) the stabilizing role of ridge regularization. All experimental details (solvers, tolerances, seeds) are in Appendix F.

7.1 Wendel sanity check: feasibility phase transition

Let $Z \sim \mathcal{N}(0, I_m)$ and target $\mu = 0$. Figure 1 compares Monte Carlo estimates of $p_n(0) = \mathbb{P}(0 \in \mathrm{conv}\{Z_i\}_{i=1}^n)$ to the Wendel closed form (14), showing a sharp transition near $n/(2m) \approx 1$.

7.2 Regularized reward: bias–variance tradeoff (bounded statistics)

We consider bounded constraints $T_j(X) = \tanh(a_j^\top X)$ with $X \sim \mathcal{N}(0, I_d)$ and construct μ from a fixed-norm ground-truth tilt (Appendix F.3). Figure 2 reports the population moment error $\|\mathbb{E}_{Q_{\hat{\eta}_\lambda}}[T] - \mu\|_2$ versus λ at $(m,n) = (40, 80)$, exhibiting the U-shaped bias–variance behavior predicted by Theorem 3. The unregularized feasibility-rate curve as a function of n/m is reported in Appendix Fig. 3.

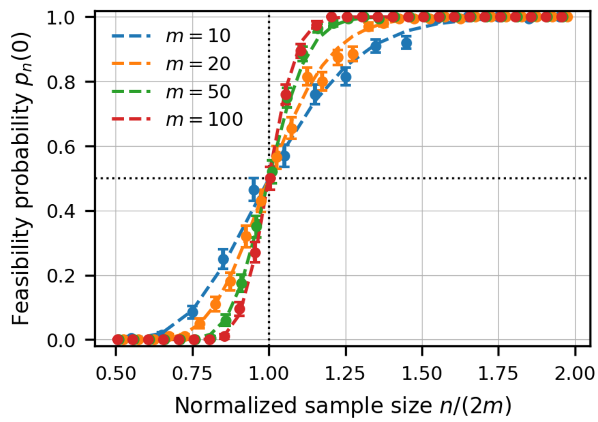

Figure 1: **Wendel phase transition for reward feasibility.** Feasibility probability $p_n(0)$ for $Z_i \sim \mathcal{N}(0, I_m)$. Markers: Monte Carlo ($\pm 1$ s.e.). Dashed: Wendel closed form (14).

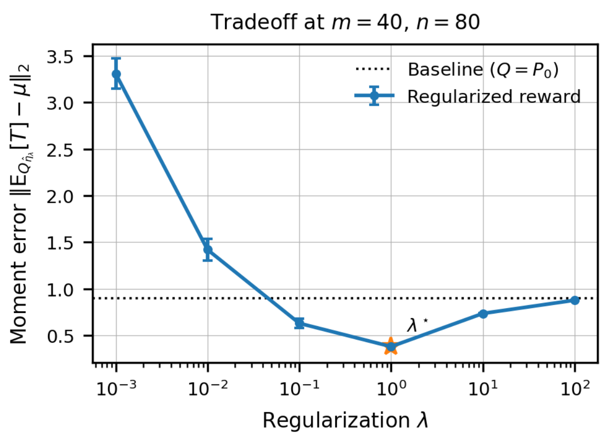

Figure 2: **Regularized reward exhibits a bias–variance tradeoff.** Population moment error $\|\mathbb{E}_{Q_{\hat{\eta}_\lambda}}[T] - \mu\|_2$ versus ridge parameter λ (log scale). Dotted line: baseline $\|\mathbb{E}_{P_0}[T] - \mu\|_2$ ($\eta = 0$).

8 DISCUSSION, LIMITATIONS, AND CONCLUSION

Unregularized empirical reward calibration is fundamentally constrained by convex-hull geometry (Lemma 1): feasibility is not guaranteed and can undergo sharp dimension-dependent transitions (Theorem 1, Corollary 1) (Hayakawa et al., 2023; Tikhomirov, 2023). A simple ridge regularization yields an always-feasible exponential-tilting estimator with an exact residual identity (Theorem 2) and finite-sample guarantees under bounded statistics (Theorem 3), aligning with the classical role of regularization in maximum-entropy estimation (Dudík et al., 2007) and with recent reward-calibration perspectives in generative modeling (Smith et al., 2025).

Practical guidance: diagnosing convex-hull failures. Our feasibility characterization (Lemma 1) suggests a concrete diagnostic for practitioners: given samples $\{X_i\}_{i=1}^n \sim P_0$ and statistics $t_i = T(X_i)$, test whether $\mu \in \text{conv}\{t_i\}_{i=1}^n$ by solving the LP feasibility system. If the LP is infeasible (or numerically ill-conditioned), this is not merely an optimization artifact: it certifies that *no* reweighting of the finite sample can exactly satisfy the constraints. Theorem 1 further suggests how to interpret such failures: feasibility requires $n = \Theta(m/\alpha(\mu))$, where $\alpha(\mu)$ is the Tukey depth of the target moment under $Z = T(X)$ (Hayakawa et al., 2023). Thus, even for moderate m, targets with small depth (rare-event constraints under P_0) can make unregularized reward calibration unreliable without very large Monte Carlo budgets.

Choosing λ and reporting constraint mismatch. The residual identity

$$\mu - \mathbb{E}_{\hat{Q}_\lambda}[T(X)] = \lambda\, \hat{\eta}_\lambda$$

makes the calibration bias operational: for a chosen λ, one can directly report the achieved mismatch $\|\mu - \mathbb{E}_{\hat{Q}_\lambda}[T]\|_2$ (or coordinatewise violations) alongside $\|\hat{\eta}_\lambda\|_2$. From a tuning perspective, Figure 2 illustrates the expected U-shaped behavior: small λ attempts aggressive matching and can be high-variance, while large λ shrinks toward the baseline $Q = P_0$ (no tilt). In applications, a principled workflow is to select λ by minimizing an estimated population moment error on held-out samples, or by targeting a specified tolerance level for constraint violation using the residual identity as a certificate.

Limitations and open directions. Our finite-sample guarantees rely on bounded statistics (Assumption (19)), which ensures globally finite log-partitions and uniform smoothness. Extending the generalization theory to *unbounded* statistics (e.g., subexponential features) would require tail-adapted concentration and localization arguments beyond the global covering bound used here. Second, while our feasibility theory is distribution-free, deriving sharp, model-specific depth estimates for structured critics used in modern generative pipelines (e.g., learned safety or preference models) remains open and would sharpen the practical sample-size predictions (Smith et al., 2025). Finally, the regularized estimator is a natural starting point for algorithmic variants that improve effective depth by adaptive sampling or proposal mixtures, thereby reducing the need for large n when $\alpha(\mu)$ is small.

Conclusion. Convex-hull feasibility is the correct geometric lens for understanding when unregularized reward calibration is well-defined. A ridge-regularized exponential tilt eliminates feasibility pathologies while retaining a transparent calibration certificate and provable finite-sample control.

References

Dudík, M., Phillips, S. J., and Schapire, R. E. (2007). Maximum entropy density estimation with generalized regularization and an application to species distribution modeling. *Journal of Machine Learning Research*, 8(44):1217–1260.

Hayakawa, S., Lyons, T., and Oberhauser, H. (2023). Estimating the probability that a given vector is in the convex hull of a random sample. *Probability Theory and Related Fields*, 185:705–746. Also available as arXiv:2101.04250.

Smith, H. D., Diamant, N. L., and Trippe, B. L. (2025). Calibrating generative models to distributional constraints. *arXiv preprint arXiv:2510.10020*.

Tikhomirov, K. (2023). On the probability that convex hull of random points contains the origin. *arXiv preprint arXiv:2304.13133*.

A Additional notation and preliminaries

This appendix collects notation and standard facts used throughout the paper. Unless stated otherwise, $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^m$, $\langle a,b\rangle = a^\top b$ the Euclidean inner product, and $\mathbb{S}^{m-1} := \{u \in \mathbb{R}^m : \|u\| = 1\}$ the unit sphere.

A.1 Measures, expectations, and KL divergence

Let $(\mathcal{X},\mathcal{F})$ be a measurable space. For probability measures P, Q on $(\mathcal{X},\mathcal{F})$, we write $Q \ll P$ if Q is absolutely continuous with respect to P, and $\frac{dQ}{dP}$ for the Radon–Nikodým derivative. For a measurable function $f : \mathcal{X} \to \mathbb{R}$, we write $\mathbb{E}_P[f] := \int f\,dP$ when the integral is well-defined. For a random variable $X \sim P$ we also write $\mathbb{E}[f(X)]$ when P is clear.

The Kullback–Leibler divergence is

$$\mathrm{KL}(Q\|P) \;:=\; \begin{cases} \displaystyle\int \log\Big(\frac{dQ}{dP}\Big)\,dQ, & Q \ll P,\\ +\infty, & \text{otherwise.}\end{cases}$$

Given n samples $X_1,\dots,X_n \in \mathcal{X}$, we write the empirical measure as

$$P_n \;:=\; \frac{1}{n}\sum_{i=1}^n \delta_{X_i}.$$

Any $Q \ll P_n$ is supported on $\{X_i\}_{i=1}^n$ and corresponds to a weight vector $w = (w_1,\dots,w_n)$ in the simplex

$$\Delta_n \;:=\; \Big\{w \in \mathbb{R}^n_{\ge 0} : \sum_{i=1}^n w_i = 1\Big\}, \qquad Q_w \;:=\; \sum_{i=1}^n w_i\,\delta_{X_i}.$$

A direct calculation gives

$$\mathrm{KL}(Q_w\|P_n) \;=\; \sum_{i=1}^n w_i \log(n w_i),$$

with the convention $0\log 0 := 0$.

A.2 Convex geometry notation

For a set $S \subseteq \mathbb{R}^m$, $\mathrm{conv}(S)$ denotes its convex hull and $\mathrm{ri}(S)$ its relative interior. For points $t_1,\dots,t_n \in \mathbb{R}^m$,

$$\mathrm{conv}\{t_1,\dots,t_n\} = \Big\{\sum_{i=1}^n w_i t_i :\; w \in \Delta_n\Big\}.$$

For $u \in \mathbb{S}^{m-1}$ and $a \in \mathbb{R}$, we write the (closed) halfspace

$$H(u,a) \;:=\; \{z \in \mathbb{R}^m :\; u^\top z \ge a\}.$$

Given a random vector $Z \in \mathbb{R}^m$ and a point $\mu \in \mathbb{R}^m$, we recall the Tukey (halfspace) depth

$$\alpha(\mu) \;:=\; \inf_{u\in\mathbb{S}^{m-1}} \mathbb{P}\big(u^\top Z \ge u^\top \mu\big)\,. \tag{24}$$

We use (24) as a quantitative proxy for how "central" the target moment μ is under the base model. Depth-based bounds on convex-hull membership probabilities (used in Section 4) are developed by Hayakawa et al. (2023).

A.3 Exponential tilting and log-partition identities

Let $T : \mathcal{X} \to \mathbb{R}^m$ be a measurable statistic and let P_0 be a base distribution on $\mathcal{X}$. Define the (population) log-partition function

$$A(\eta) \;:=\; \log \mathbb{E}_{P_0}\big[\exp(\eta^\top T(X))\big]\,, \qquad \eta \in \mathbb{R}^m, \tag{25}$$

whenever the expectation is finite. The exponentially tilted distribution is

$$\frac{dQ_\eta}{dP_0}(x) \;=\; \exp\big(\eta^\top T(x) - A(\eta)\big). \tag{26}$$

This is the standard maximum-entropy / exponential-family form that underlies constraint calibration (Smith et al., 2025; Dudík et al., 2007).

Gradient and Hessian. When differentiation under the integral is justified,

$$\nabla A(\eta) \;=\; \mathbb{E}_{Q_\eta}[T(X)], \qquad \nabla^2 A(\eta) \;=\; \mathrm{Cov}_{Q_\eta}(T(X)) \;\succeq\; 0. \tag{27}$$

Similarly, given samples $X_1,\dots,X_n \sim P_0$ and $t_i := T(X_i)$, the empirical log-partition function

$$A_n(\eta) \;:=\; \log\Big(\frac{1}{n}\sum_{i=1}^n e^{\eta^\top t_i}\Big)$$

satisfies

$$\begin{aligned} \nabla A_n(\eta) &= \sum_{i=1}^n \hat{w}_i(\eta)\, t_i,\\ \nabla^2 A_n(\eta) &= \sum_{i=1}^n \hat{w}_i(\eta)\,(t_i - \bar{t}_\eta)(t_i - \bar{t}_\eta)^\top, \end{aligned} \tag{28}$$

where $\hat{w}_i(\eta) \propto e^{\eta^\top t_i}$ and $\bar{t}_\eta := \sum_i \hat{w}_i(\eta) t_i$.

Bregman/KL identity for exponential families. Define the Bregman divergence induced by A:

$$D_A(\eta',\eta) \;:=\; A(\eta') - A(\eta) - \langle \nabla A(\eta), \eta' - \eta\rangle.$$

A standard calculation using (26) shows that for any η, η' with $A(\eta), A(\eta') < \infty$,

$$\mathrm{KL}(Q_\eta\|Q_{\eta'}) \;=\; A(\eta') - A(\eta) - (\eta' - \eta)^\top \nabla A(\eta) \;=\; D_A(\eta',\eta). \tag{29}$$

We use (29) in Section 6 to convert parameter error bounds into KL bounds; see, e.g., Dudík et al. (2007) and the calibration treatment in Smith et al. (2025).

Variational representation of log-sum-exp. We will also use the standard entropy duality: for any $t_1, \dots, t_n \in \mathbb{R}^m$ and any $\eta \in \mathbb{R}^m$,

$$A_n(\eta) = \sup_{w \in \Delta_n} \left\{ \eta^\top \sum_{i=1}^n w_i t_i - \sum_{i=1}^n w_i \log(n w_i) \right\}. \quad (30)$$

Identity (30) is the finite-support analogue of the classical maximum-entropy duality and is commonly used in regularized max-entropy estimation (Dudík et al., 2007) as well as in generative constraint calibration (Smith et al., 2025).

A.4 Bounded statistics imply global smoothness

A recurring assumption in Section 6 is that T is bounded:

$$\|T(X)\| \le B \quad \text{almost surely under } P_0. \quad (31)$$

Under (31), the exponential moment in (25) is finite for all $\eta \in \mathbb{R}^m$, and (27) implies global bounds on the gradient and curvature.

Lemma 2 (Gradient and Hessian bounds under bounded T)**.** *Assume* (31)*. Then for all $\eta \in \mathbb{R}^m$,*

$$\|\nabla A(\eta)\| \le B, \qquad \|\nabla^2 A(\eta)\|_{\mathrm{op}} \le B^2, \quad (32)$$

where $\|\cdot\|_{\mathrm{op}}$ denotes the spectral/operator norm. Consequently, ∇A is B^2-Lipschitz:

$$\|\nabla A(\eta) - \nabla A(\eta')\| \le B^2 \|\eta - \eta'\|, \qquad \forall \eta, \eta' \in \mathbb{R}^m.$$

Proof. By (27), $\nabla A(\eta) = \mathbb{E}_{Q_\eta}[T(X)]$, so Jensen yields $\|\nabla A(\eta)\| \le \mathbb{E}_{Q_\eta}[\|T(X)\|] \le B$. Also $\nabla^2 A(\eta) = \mathrm{Cov}_{Q_\eta}(T)$ is positive semidefinite and satisfies $\|\mathrm{Cov}_{Q_\eta}(T)\|_{\mathrm{op}} \le \mathbb{E}_{Q_\eta}[\|T\|^2] \le B^2$. The Lipschitz bound follows by integrating the Hessian along the line segment between η and η'. ☐

Analogous bounds hold for A_n when $\|t_i\| \le B$ for all i, since $\nabla^2 A_n(\eta)$ is the covariance of $\{t_i\}$ under the softmax weights (28).

A.5 Concentration and covering preliminaries

We will use standard concentration and covering arguments to control $A_n(\eta) - A(\eta)$ uniformly over $\|\eta\| \le R$. We collect two basic tools here.

Hoeffding inequality (bounded variables). If $Y_1, \dots, Y_n$ are i.i.d. with $Y_i \in [a, b]$ almost surely, then for all $t > 0$,

$$\mathbb{P}\left(\left| \frac{1}{n} \sum_{i=1}^n Y_i - \mathbb{E}[Y_1] \right| \ge t \right) \le 2 \exp\left(-\frac{2nt^2}{(b-a)^2} \right). \quad (33)$$

Covering numbers of Euclidean balls. Let $B_2(R) := \{\eta \in \mathbb{R}^m : \|\eta\| \le R\}$ be the Euclidean ball. For any $\varepsilon > 0$, there exists an ε-net $\mathcal{N} \subseteq B_2(R)$ with cardinality bounded by

$$|\mathcal{N}| \le \left(1 + \frac{2R}{\varepsilon}\right)^m. \quad (34)$$

We combine (33)–(34) with Lipschitz properties of A and A_n (implied by bounded T) to obtain the uniform deviation bound in Lemma 11.

B Proofs for Section 3 (Setup and empirical reward calibration)

This appendix provides full proofs and derivations omitted from the main text for brevity. Throughout, fix samples $X_1, \dots, X_n \in \mathcal{X}$ and define $t_i := T(X_i) \in \mathbb{R}^m$. Recall the empirical measure $P_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ and that any $Q \ll P_n$ can be written as $Q_w = \sum_{i=1}^n w_i \delta_{X_i}$ for some $w \in \Delta_n$ (Appendix A).

B.1 Proof of Lemma 1 (Feasibility ⇔ convex hull)

Lemma 3 (restatement of Lemma 1)**.** *The empirical primal problem* (6) *is feasible if and only if $\mu \in \mathrm{conv}\{t_1, \dots, t_n\}$.*

Proof. Recall the empirical primal (6):

$$\min_{w \in \Delta_n} \sum_{i=1}^n w_i \log(n w_i) \quad \text{s.t.} \quad \sum_{i=1}^n w_i t_i = \mu.$$

Feasibility means there exists some $w \in \Delta_n$ satisfying the linear constraint $\sum_i w_i t_i = \mu$. But by definition of convex hull,

$$\mathrm{conv}\{t_1, \dots, t_n\} = \Big\{ \sum_{i=1}^n w_i t_i : \ w \in \Delta_n \Big\}.$$

Thus, $\exists\, w \in \Delta_n$ with $\sum_i w_i t_i = \mu$ if and only if $\mu \in \mathrm{conv}\{t_1, \dots, t_n\}$. ☐

B.2 Dual derivation for empirical maximum entropy

We now derive the empirical dual (8) and the exponential-tilt form (9), following standard maximum-entropy duality (see, e.g., Dudík et al. (2007)) and the generative calibration treatment in Smith et al. (2025).

Primal in weights. Define the convex objective

$$\Phi(w) \;:=\; \mathrm{KL}(Q_w \| P_n) \;=\; \sum_{i=1}^{n} w_i \log(n w_i), \qquad w \in \Delta_n,$$

and the feasible set

$$\mathcal{W}(\mu) \;:=\; \Big\{ w \in \Delta_n : \; \sum_{i=1}^{n} w_i t_i = \mu \Big\}.$$

Then (6) is $\min_{w \in \mathcal{W}(\mu)} \Phi(w)$.

Lagrangian and stationary conditions. Consider the Lagrangian with multipliers $\eta \in \mathbb{R}^m$ for the moment constraint and $\nu \in \mathbb{R}$ for the simplex constraint:

$$\mathcal{L}(w, \eta, \nu) = \sum_{i=1}^{n} w_i \log(n w_i) + \; \eta^\top \Big(\mu - \sum_{i=1}^{n} w_i t_i \Big) + \; \nu \Big(1 - \sum_{i=1}^{n} w_i \Big). \tag{35}$$

Formally differentiating with respect to w_i in the interior of the simplex yields

$$\frac{\partial}{\partial w_i} \mathcal{L}(w, \eta, \nu) = \log(n w_i) + 1 - \eta^\top t_i - \nu.$$

Setting this to zero gives

$$w_i = \frac{1}{n} \exp(\eta^\top t_i + \nu - 1). \tag{36}$$

Imposing $\sum_i w_i = 1$ determines ν and yields the normalized exponential form

$$w_i(\eta) = \frac{\exp(\eta^\top t_i)}{\sum_{j=1}^{n} \exp(\eta^\top t_j)} \;=:\; \hat{w}_i(\eta), \qquad i = 1, \ldots, n, \tag{37}$$

which coincides with (9). In particular, $w_i(\eta) > 0$ for all i and all η.

Dual objective. Substituting $w(\eta)$ from (37) into $\inf_{w \in \Delta_n} \mathcal{L}(w, \eta, \nu)$ yields the dual function

$$g_n(\eta) = \eta^\top \mu - A_n(\eta), \qquad A_n(\eta) := \log \Big(\frac{1}{n} \sum_{i=1}^{n} \exp(\eta^\top t_i) \Big), \tag{38}$$

and the dual problem $\max_{\eta \in \mathbb{R}^m} g_n(\eta)$, which is (8). (Equivalently, one may derive (38) from the variational representation (30) in Appendix A.)

Moment matching as a first-order condition. Since A_n is differentiable and convex,

$$\nabla A_n(\eta) = \sum_{i=1}^{n} \hat{w}_i(\eta)\, t_i = \mathbb{E}_{\hat{Q}_\eta}[T(X)],$$

and thus any maximizer $\hat{\eta}$ of $g_n(\eta) = \eta^\top \mu - A_n(\eta)$ satisfies the stationarity condition

$$\nabla A_n(\hat{\eta}) = \mu \quad \Longleftrightarrow \quad \sum_{i=1}^{n} \hat{w}_i(\hat{\eta})\, t_i = \mu, \tag{39}$$

i.e., exact satisfaction of the empirical constraint (when the maximizer exists and is finite).

B.3 Well-definedness: Slater condition, strong duality, and uniqueness

The main text emphasizes that the unregularized empirical reward formulation can be *ill-defined* when μ lies outside (or on the boundary of) the empirical convex hull. We formalize the standard sufficient conditions ensuring (i) strong duality and (ii) existence/uniqueness of a finite dual optimizer, closely aligned with the "well-definedness" discussion in Smith et al. (2025).

A Slater-type condition from relative interior. For the discrete simplex domain, strict feasibility corresponds to a feasible point in the relative interior of Δ_n, namely a weight vector with strictly positive entries.

Lemma 4 (Relative interior implies a strictly feasible weight vector)**.** *Assume $\mu \in \mathrm{ri}(\mathrm{conv}\{t_1, \ldots, t_n\})$. Then there exists $w \in \Delta_n$ such that $w_i > 0$ for all i and $\sum_{i=1}^{n} w_i t_i = \mu$.*

Proof. Let $P := \mathrm{conv}\{t_1, \ldots, t_n\}$ and let $V \subseteq \{t_i\}_{i=1}^{n}$ be the set of vertices (extreme points) of P. A standard characterization of relative interior for polytopes implies that $\mu \in \mathrm{ri}(P)$ admits a convex representation over vertices with strictly positive coefficients: there exist $\{\lambda_v\}_{v \in V}$ with $\lambda_v > 0$, $\sum_{v \in V} \lambda_v = 1$, and $\mu = \sum_{v \in V} \lambda_v v$.

Now consider any index j such that t_j is *not* a vertex of P. Then t_j can be expressed as a convex combination of vertices: $t_j = \sum_{v \in V} \alpha_{jv} v$ with $\alpha_{jv} \geq 0$ and $\sum_{v \in V} \alpha_{jv} = 1$. Fix a small $\varepsilon > 0$ (to be chosen) and define new weights

$$w_j := \varepsilon \quad \text{for } j \notin V, \qquad w_v := \lambda_v - \varepsilon \sum_{j \notin V} \alpha_{jv} \quad \text{for } v \in V. \tag{40}$$

By construction, $\sum_i w_i = 1$ and

$$\sum_i w_i t_i = \sum_{v \in V} w_v v + \sum_{j \notin V} w_j t_j = \sum_{v \in V} \lambda_v v = \mu.$$

Choosing $\varepsilon > 0$ small enough ensures $w_v > 0$ for all $v \in V$, and clearly $w_j = \varepsilon > 0$ for all $j \notin V$. Thus $w \in \Delta_n$ is strictly positive and feasible. □

Strong duality and attainment. We now state a standard consequence: under strict feasibility, the empirical primal and dual have the same optimal value, and the dual optimum is attained.

Proposition 2 (Strong duality and attainment under interior feasibility)**.** *Assume $\mu \in \mathrm{ri}(\mathrm{conv}\{t_1, \dots, t_n\})$. Then the primal* (6) *has a unique minimizer $\hat{w} \in \Delta_n$ with $\hat{w}_i > 0$ for all i, and the dual* (8) *attains its maximum at some finite $\hat{\eta} \in \mathbb{R}^m$. Moreover, $(\hat{w}, \hat{\eta})$ satisfies the KKT conditions, and $\hat{w} = \hat{w}(\hat{\eta})$ is given by* (37)*.*

Proof. By Lemma 4, strict feasibility holds, so Slater's condition applies to the convex program (6). Hence strong duality holds and the dual optimum is attained. Since $\Phi(w) = \sum_i w_i \log(nw_i)$ is strictly convex on $\mathrm{ri}(\Delta_n)$ and the feasible set is convex, the primal minimizer is unique. The KKT conditions yield the exponential-form weights (37). □

Uniqueness and nondegeneracy. While $\hat{w}$ is unique under the conditions above, the dual maximizer $\hat{\eta}$ is unique only when the dual objective $g_n(\eta) = \eta^\top \mu - A_n(\eta)$ is strictly concave, which is equivalent to $\nabla^2 A_n(\hat{\eta}) \succ 0$. Using (28), we can state a convenient sufficient condition.

Proposition 3 (A sufficient condition for uniqueness of $\hat{\eta}$)**.** *Assume $\mu \in \mathrm{ri}(\mathrm{conv}\{t_1, \dots, t_n\})$ and that the affine hull of $\{t_i\}_{i=1}^n$ is $\mathbb{R}^m$ (equivalently, $\{t_i - \bar{t}\}$ spans $\mathbb{R}^m$ for some $\bar{t}$). Then $\nabla^2 A_n(\hat{\eta}) \succ 0$ and the dual maximizer $\hat{\eta}$ is unique.*

Proof. From Proposition 2, $\hat{w}_i > 0$ for all i. For any nonzero $v \in \mathbb{R}^m$,

$$v^\top \nabla^2 A_n(\hat{\eta})\, v = \mathrm{Var}_{i \sim \hat{w}}\big(v^\top t_i\big).$$

If this variance were zero, then $v^\top t_i$ would be constant over all i with $\hat{w}_i > 0$, hence over all i. This would imply $\{t_i\}$ lies in a hyperplane orthogonal to v, contradicting the full-dimensional affine hull assumption. Thus $v^\top \nabla^2 A_n(\hat{\eta}) v > 0$ for all $v \neq 0$, i.e., $\nabla^2 A_n(\hat{\eta}) \succ 0$. Therefore A_n is strictly convex in a neighborhood of $\hat{\eta}$, so g_n is strictly concave and $\hat{\eta}$ is unique. □

Failure modes on the boundary or outside the hull. If $\mu \notin \mathrm{conv}\{t_i\}$, the primal is infeasible (Lemma 1). If μ lies on the boundary of $\mathrm{conv}\{t_i\}$, strict feasibility fails and the dual may not admit a finite maximizer: intuitively, the entropy/KL objective pushes weights away from the boundary of the simplex, but matching a boundary moment may require concentrating on a lower-dimensional face. This is precisely the geometric ill-posedness highlighted in Smith et al. (2025) and motivates the ridge-regularized estimator introduced in Section 5.

C Proofs for Section 4 (Feasibility theory: depth and phase transition)

This appendix provides full proofs and supporting lemmas for the feasibility results in Section 4. Throughout, let $Z \in \mathbb{R}^m$ denote a random vector (the constraint vector $T(X)$ under P_0), and let $Z_1, \dots, Z_n \overset{\text{i.i.d.}}{\sim} Z$. For a target $\mu \in \mathbb{R}^m$, recall

$$\begin{aligned} p_n(\mu) &:= \mathbb{P}\big(\mu \in \mathrm{conv}\{Z_1, \dots, Z_n\}\big), \\ N(\mu) &:= \min\{n \geq 1 :\ p_n(\mu) \geq 1/2\}, \end{aligned} \tag{41}$$

and the (Tukey/halfspace) depth

$$\alpha(\mu) \ := \ \inf_{u \in \mathbb{S}^{m-1}} \mathbb{P}\big(u^\top Z \geq u^\top \mu\big) \in [0, 1/2].$$

Depth-based bounds on $N(\mu)$ are due to Hayakawa et al. (2023). Phase-transition refinements in the Wendel setting are discussed by Tikhomirov (2023).

C.1 Convex-hull membership and separating halfspaces

We first record the standard separation characterization of convex-hull failure. This is the geometric core behind expressing infeasibility as a halfspace event.

Lemma 5 (Separation for convex hull)**.** *Let $z_1, \dots, z_n \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^m$. Then $\mu \notin \mathrm{conv}\{z_1, \dots, z_n\}$ if and only if there exists $u \in \mathbb{S}^{m-1}$ such that*

$$u^\top z_i < u^\top \mu \qquad \text{for all } i = 1, \dots, n. \tag{42}$$

Equivalently, there exists a (closed) halfspace $H = \{z :\ u^\top z \leq u^\top \mu\}$ containing $\mathrm{conv}\{z_i\}$ but with μ lying strictly outside its interior-facing side.

Proof. ($\Rightarrow$) If $\mu \notin \mathrm{conv}\{z_i\}$, then $\mathrm{conv}\{z_i\}$ is a compact convex set not containing μ. By the strong separation theorem, there exists a nonzero vector $u \in \mathbb{R}^m$ and a scalar $b \in \mathbb{R}$ such that $u^\top \mu > b \geq \sup_{z \in \mathrm{conv}\{z_i\}} u^\top z$. In particular, $u^\top \mu > u^\top z_i$ for each i. Rescaling u to unit norm yields (42).

($\Leftarrow$) If (42) holds for some u, then $\sup_{z \in \mathrm{conv}\{z_i\}} u^\top z = \max_i u^\top z_i < u^\top \mu$. Hence μ cannot belong to $\mathrm{conv}\{z_i\}$. □

As an immediate probabilistic corollary, convex-hull failure is equivalent to a *uniform* halfspace-exclusion event.

Corollary 2 (Convex-hull failure as a halfspace event)**.** *For i.i.d.* $Z_1, \dots, Z_n \sim Z$ *and any* $\mu \in \mathbb{R}^m$,

$$1 - p_n(\mu) = \mathbb{P}\big(\exists u \in \mathbb{S}^{m-1} : \ u^\top Z_i < u^\top \mu \ \ \forall i = 1, \dots, n\big). \tag{43}$$

Proof. Apply Lemma 5 pointwise to the realization $(Z_1, \dots, Z_n)$ and take probabilities. □

C.2 Depth bounds: proof of Theorem 1

We now prove the depth-based lower bound and state the sharp upper bound from Hayakawa et al. (2023) in our notation.

Lower bound ($\alpha(\mu)N(\mu) \geq \frac{1}{2}$). This direction admits a short self-contained proof.

Lemma 6 (Depth lower bound on median feasibility)**.** *For any* $\mu \in \mathbb{R}^m$ *with depth* $\alpha(\mu) \in (0, 1/2]$,

$$N(\mu) \ \geq \ \frac{1}{2\,\alpha(\mu)} \qquad \Longleftrightarrow \qquad \alpha(\mu)\,N(\mu) \ \geq \ \frac{1}{2}. \tag{44}$$

Proof. Fix μ and let $\alpha = \alpha(\mu)$. By definition of the infimum, for any $\varepsilon > 0$ there exists $u \in \mathbb{S}^{m-1}$ such that

$$\mathbb{P}(u^\top Z \geq u^\top \mu) \leq \alpha + \varepsilon.$$

Define the closed halfspace $H = \{z : \ u^\top z \geq u^\top \mu\}$ containing μ on its boundary. If all samples fall in the complement $H^c = \{z : \ u^\top z < u^\top \mu\}$, then by Lemma 5 we have $\mu \notin \operatorname{conv}\{Z_1, \dots, Z_n\}$. Therefore,

$$\begin{aligned} 1 - p_n(\mu) &\geq \mathbb{P}(Z_1, \dots, Z_n \in H^c) \\ &= \mathbb{P}(Z \in H^c)^n \\ &\geq \big(1 - (\alpha + \varepsilon)\big)^n. \end{aligned} \tag{45}$$

Letting $\varepsilon \downarrow 0$ yields $1 - p_n(\mu) \geq (1-\alpha)^n$, hence $p_n(\mu) \leq 1 - (1-\alpha)^n$.

Now suppose $n \leq \frac{1}{2\alpha}$. Since $\alpha \leq 1/2$, the function $f(\alpha) := (1-\alpha)^{1/(2\alpha)}$ is minimized at $\alpha = 1/2$ with $f(1/2) = 1/2$, and $f(\alpha) \geq 1/2$ for all $\alpha \in (0, 1/2]$. Consequently,

$$(1-\alpha)^n \ \geq \ (1-\alpha)^{1/(2\alpha)} \ \geq \ \frac{1}{2},$$

and thus $p_n(\mu) \leq 1 - (1-\alpha)^n \leq 1/2$. Therefore, the smallest n such that $p_n(\mu) \geq 1/2$ must satisfy $N(\mu) > 1/(2\alpha)$, which proves (44). □

Upper bound ($\alpha(\mu)N(\mu) \leq 3m + 1$). The converse inequality is nontrivial and was proved by Hayakawa et al. (2023). We restate it in our notation.

Theorem 4 (Hayakawa–Lyons–Oberhauser median-sample bound)**.** *Let* $Z \in \mathbb{R}^m$ *be any random vector and* $\mu \in \mathbb{R}^m$ *any point with depth* $\alpha(\mu) > 0$*. Let* $N(\mu)$ *be the minimal* n *such that* $p_n(\mu) \geq 1/2$*. Then*

$$\frac{1}{2} \ \leq \ \alpha(\mu)\,N(\mu) \ \leq \ 3m + 1. \tag{46}$$

Proof. The lower bound is Lemma 6. The upper bound is precisely the main inequality of Hayakawa et al. (2023, Theorem 1.2) (stated there for a general d-dimensional random vector), specialized to dimension m and translated into our notation. □

In particular, (46) implies the sample-complexity corollary used in the main text: if $n \geq (3m+1)/\alpha(\mu)$, then $p_n(\mu) \geq 1/2$.

C.3 Wendel formula and a sharp phase transition (proof of Corollary 1)

We now prove the Wendel expression used in Section 4 and highlight the resulting threshold at $n \approx 2m$. We follow the standard random-sign/arrangement argument; see Tikhomirov (2023) for a modern discussion and extensions beyond the continuous setting.

General position. A collection $x_1, \dots, x_n \in \mathbb{R}^m$ is in *general position* if no subset of size m is linearly dependent (equivalently, the hyperplanes $\{u : \ u^\top x_i = 0\}$ are in general position as a central arrangement). If X has a continuous distribution, then i.i.d. samples are in general position with probability one.

Lemma 7 (Counting sign patterns of a central hyperplane arrangement)**.** *Let* $x_1, \dots, x_n \in \mathbb{R}^m$ *be in general position, and consider the hyperplanes* $H_i := \{u \in \mathbb{R}^m : u^\top x_i = 0\}$*. Then the arrangement* $\{H_i\}_{i=1}^n$ *partitions* $\mathbb{R}^m$ *into*

$$R(n, m) \ = \ 2 \sum_{k=0}^{m-1} \binom{n-1}{k} \tag{47}$$

open polyhedral cones (regions). Equivalently, the number of distinct sign patterns $\operatorname{sign}(u^\top x_1), \dots, \operatorname{sign}(u^\top x_n) \in \{\pm 1\}^n$ *attained by some* $u \in \mathbb{R}^m$ *is* $R(n, m)$*.*

Proof. This is a classical fact (often attributed to Schläfli/Cover) for central arrangements. A short proof proceeds by the standard recursion: let $R(n, m)$ be the number of regions for n hyperplanes in $\mathbb{R}^m$ in general position. Adding the n-th hyperplane H_n splits precisely those existing regions whose intersection with H_n is nonempty. The induced arrangement on $H_n \simeq \mathbb{R}^{m-1}$ consists of $n-1$ hyperplanes in general position, hence has $R(n-1, m-1)$ regions. Therefore,

$$R(n, m) = R(n-1, m) + R(n-1, m-1),$$

with base cases $R(1,m)=2$ and $R(n,1)=2$. Solving this recursion yields (47); one verifies by induction that the closed form satisfies the same recursion and base cases. □

Theorem 5 (Wendel formula for continuous centrally symmetric distributions)**.** *Let $X\in\mathbb{R}^m$ have a continuous distribution satisfying central symmetry: $X\overset{d}{=}-X$. Let $X_1,\ldots,X_n\overset{\text{i.i.d.}}{\sim}X$ and assume $n>m$. Then*

$$\mathbb{P}(0\in\operatorname{conv}\{X_1,\ldots,X_n\}) \;=\; 1-2^{-n+1}\sum_{k=0}^{m-1}\binom{n-1}{k}. \tag{48}$$

Proof. By Lemma 5 with $\mu=0$, the event $0\notin\operatorname{conv}\{x_1,\ldots,x_n\}$ is equivalent to the existence of $u\in\mathbb{S}^{m-1}$ such that $u^\top x_i>0$ for all i (or equivalently $u^\top x_i<0$ for all i). Now condition on a realization $(x_1,\ldots,x_n)$ in general position (which holds a.s. by continuity).

Let $\varepsilon_1,\ldots,\varepsilon_n$ be i.i.d. Rademacher signs, independent of X_i. By symmetry, $(\varepsilon_1X_1,\ldots,\varepsilon_nX_n)\overset{d}{=}(X_1,\ldots,X_n)$, hence

$$\mathbb{P}\big(0\notin\operatorname{conv}\{X_i\}_{i=1}^n\big)=\mathbb{E}\Big[\mathbb{P}\big(0\notin\operatorname{conv}\{\varepsilon_iX_i\}_{i=1}^n \;\big|\; X_1,\ldots,X_n\big)\Big]. \tag{49}$$

Fix $(x_1,\ldots,x_n)$. For a given sign vector $\varepsilon\in\{\pm1\}^n$, we have

$$\begin{aligned}0\notin\operatorname{conv}\{\varepsilon_ix_i\}_{i=1}^n &\iff \exists u\in\mathbb{R}^m:\ u^\top(\varepsilon_ix_i)>0\ \forall i\\ &\iff \exists u\in\mathbb{R}^m:\ \operatorname{sign}(u^\top x_i)=\varepsilon_i\ \forall i.\end{aligned} \tag{50}$$

Thus, the number of sign vectors ε for which the event holds equals the number of realizable sign patterns $\operatorname{sign}(u^\top x_i)$ as u ranges over $\mathbb{R}^m\setminus\cup_iH_i$, where $H_i=\{u:\ u^\top x_i=0\}$. By Lemma 7, this number is $R(n,m)=2\sum_{k=0}^{m-1}\binom{n-1}{k}$. Since ε is uniform over $\{\pm1\}^n$, we obtain

$$\begin{aligned}\mathbb{P}\big(0\notin\operatorname{conv}\{\varepsilon_ix_i\}_{i=1}^n \mid x_1,\ldots,x_n\big)&=\frac{R(n,m)}{2^n}\\ &=2^{-n+1}\sum_{k=0}^{m-1}\binom{n-1}{k}.\end{aligned} \tag{51}$$

Taking expectations over $(X_1,\ldots,X_n)$ yields

$$\mathbb{P}(0\notin\operatorname{conv}\{X_1,\ldots,X_n\})=2^{-n+1}\sum_{k=0}^{m-1}\binom{n-1}{k},$$

which is equivalent to (48). □

Corollary 1 in the main text is Theorem 5 with $Z=X$ and $\mu=0$. The same expression is denoted $p_{n,m}$ in Tikhomirov (2023), who further shows that (48) remains essentially correct (up to exponentially small error) for broad classes of symmetric subgaussian distributions.

A sharp threshold around $n\approx 2m$. The Wendel probability (48) can be written as a binomial tail: if $B\sim\operatorname{Bin}(n-1,1/2)$ then

$$\mathbb{P}(0\in\operatorname{conv}\{X_1,\ldots,X_n\})=1-\mathbb{P}(B\le m-1). \tag{52}$$

A direct Hoeffding bound yields a quantitative transition.

Lemma 8 (Hoeffding window for the Wendel transition)**.** *Let $B\sim\operatorname{Bin}(n-1,1/2)$ and define $p_{n,m}:=1-\mathbb{P}(B\le m-1)$. Then for any $n>m$,*

$$\begin{aligned}p_{n,m}&\ge 1-\exp\left(-\frac{(n-2m+1)^2}{2(n-1)}\right),\\ p_{n,m}&\le \exp\left(-\frac{(2m-n-1)^2}{2(n-1)}\right)\quad \textit{if } n<2m+1.\end{aligned} \tag{53}$$

Consequently, if $n=2m+\omega(\sqrt{m})$ then $p_{n,m}\to1$, while if $n=2m-\omega(\sqrt{m})$ then $p_{n,m}\to0$.

Proof. Let $\mu_B:=\mathbb{E}[B]=(n-1)/2$. If $n\ge 2m+1$ then $\mu_B-(m-1)=(n-2m+1)/2=:t\ge0$, so

$$\begin{aligned}\mathbb{P}(B\le m-1)=\mathbb{P}(B-\mu_B\le -t)&\le\exp\left(-\frac{2t^2}{n-1}\right)\\ &=\exp\left(-\frac{(n-2m+1)^2}{2(n-1)}\right).\end{aligned} \tag{54}$$

where we used Hoeffding's inequality for binomials. Plugging into $p_{n,m}=1-\mathbb{P}(B\le m-1)$ gives the first bound.

If $n<2m+1$, then $(m-1)-\mu_B=(2m-n-1)/2=:t>0$, hence

$$\begin{aligned}p_{n,m}=\mathbb{P}(B\ge m)=\mathbb{P}(B-\mu_B\ge t)&\le\exp\left(-\frac{2t^2}{n-1}\right)\\ &=\exp\left(-\frac{(2m-n-1)^2}{2(n-1)}\right).\end{aligned} \tag{55}$$

which is the second bound. The asymptotic statements follow by setting $n=2m\pm\omega(\sqrt{m})$ and observing that the exponents diverge to $-\infty$. □

D Proofs for Section 5 (Regularized reward estimator)

This appendix provides full proofs for the results in Section 5. We follow standard maximum-entropy /

exponential-family duality arguments (e.g., Dudík et al. (2007)) and the recent generative calibration framing of Smith et al. (2025).

Throughout, let $X_1, \dots, X_n \sim P_0$, write $t_i := T(X_i) \in \mathbb{R}^m$, and recall the empirical log-partition

$$A_n(\eta) \;:=\; \log\Big(\frac{1}{n}\sum_{i=1}^{n}\exp(\eta^\top t_i)\Big), \qquad \eta \in \mathbb{R}^m.$$

For $\eta \in \mathbb{R}^m$, define the induced softmax weights and empirical tilt as in (9).

D.1 Proof of Theorem 2 (existence/uniqueness and residual identity)

Recall the ridge-regularized empirical dual objective

$$F_n(\eta) \;:=\; \eta^\top \mu - A_n(\eta) - \frac{\lambda}{2}\|\eta\|^2, \qquad \lambda > 0. \tag{56}$$

Proof. **Step 1 (Strong concavity).** Since A_n is convex, $-A_n$ is concave. The map $\eta \mapsto -\frac{\lambda}{2}\|\eta\|^2$ is λ-strongly concave on $\mathbb{R}^m$. Therefore F_n is λ-strongly concave: for all $\eta, \eta' \in \mathbb{R}^m$ and $\theta \in [0,1]$,

$$\begin{aligned} &F_n(\theta\eta + (1-\theta)\eta') \\ &\quad\geq \theta F_n(\eta) + (1-\theta)F_n(\eta') + \frac{\lambda}{2}\theta(1-\theta)\|\eta - \eta'\|^2. \end{aligned} \tag{57}$$

In particular, if a maximizer exists then it is unique.

Step 2 (Coercivity and existence). Let $T_{\max} := \max_{1\leq i\leq n}\|t_i\|$. For any $\eta \in \mathbb{R}^m$ and any k,

$$A_n(\eta) = \log\Big(\frac{1}{n}\sum_{i=1}^{n} e^{\eta^\top t_i}\Big) \geq \log\Big(\frac{1}{n}e^{\eta^\top t_k}\Big) = \eta^\top t_k - \log n,$$

hence $A_n(\eta) \geq \max_i \eta^\top t_i - \log n$. Substituting into (56) yields

$$\begin{aligned} F_n(\eta) &\leq \log n + \eta^\top\mu - \max_i \eta^\top t_i - \frac{\lambda}{2}\|\eta\|^2 \\ &\leq \log n + \|\eta\|\,\|\mu\| + \|\eta\|\,T_{\max} - \frac{\lambda}{2}\|\eta\|^2, \end{aligned} \tag{58}$$

where $-\max_i \eta^\top t_i \leq \max_i |\eta^\top t_i| \leq \|\eta\| T_{\max}$. The RHS tends to $-\infty$ as $\|\eta\| \to \infty$, so F_n is coercive and continuous, hence attains its maximum. Together with (57), the maximizer $\hat{\eta}_\lambda$ exists and is unique.

Step 3 (Residual identity). Using $\nabla A_n(\eta) = \sum_{i=1}^{n}\hat{w}_i(\eta)t_i = \mathbb{E}_{\hat{Q}_\eta}[T(X)]$ (cf. (9)),

$$\nabla F_n(\eta) = \mu - \nabla A_n(\eta) - \lambda\eta.$$

At $\hat{\eta}_\lambda$, the first-order condition $\nabla F_n(\hat{\eta}_\lambda) = 0$ gives

$$\mu - \mathbb{E}_{\hat{Q}_{\hat{\eta}_\lambda}}[T(X)] = \lambda\hat{\eta}_\lambda,$$

which is (16). □

A useful norm bound (optional). If additionally $\|t_i\| \leq B$ for all i and $\|\mu\| \leq B$, then (16) implies $\lambda\|\hat{\eta}_\lambda\| \leq \|\mu\| + \|\mathbb{E}_{\hat{Q}_\lambda}[T]\| \leq 2B$, hence $\|\hat{\eta}_\lambda\| \leq 2B/\lambda$.

D.2 Proof of Proposition 1 (penalized primal equivalence)

Proof. Start from the variational representation of log-sum-exp (Appendix A, (30)):

$$A_n(\eta) = \sup_{w\in\Delta_n}\left\{\eta^\top\sum_{i=1}^{n} w_i t_i - \sum_{i=1}^{n} w_i\log(nw_i)\right\}.$$

Plugging this into the ridge dual (15) gives

$$\begin{aligned} &\max_{\eta\in\mathbb{R}^m}\left\{\eta^\top\mu - A_n(\eta) - \tfrac{\lambda}{2}\|\eta\|^2\right\} \\ &= \max_{\eta\in\mathbb{R}^m}\inf_{w\in\Delta_n}\left\{\eta^\top(\mu - T_w) + S(w) - \tfrac{\lambda}{2}\|\eta\|^2\right\}, \end{aligned} \tag{59}$$

where $T_w := \sum_{i=1}^{n} w_i t_i$ and $S(w) := \sum_{i=1}^{n} w_i\log(nw_i)$.

For each fixed w, the map $\eta \mapsto \eta^\top(\mu - T_w) - \frac{\lambda}{2}\|\eta\|^2$ is concave and continuous (indeed λ-strongly concave). For each fixed η, the map $w \mapsto S(w) - \eta^\top T_w$ is convex and lower semicontinuous on the compact convex set Δ_n. Thus Sion's minimax theorem applies, allowing us to exchange $\max_\eta$ and $\inf_w$:

$$\begin{aligned} \max_{\eta\in\mathbb{R}^m}\inf_{w\in\Delta_n}\psi(\eta,w) &= \inf_{w\in\Delta_n}\max_{\eta\in\mathbb{R}^m}\psi(\eta,w), \\ \psi(\eta,w) &:= \eta^\top(\mu - T_w) + S(w) - \tfrac{\lambda}{2}\|\eta\|^2. \end{aligned} \tag{60}$$

Fix $w \in \Delta_n$ and set $r(w) := \mu - T_w$. The inner maximization over η is

$$\max_{\eta\in\mathbb{R}^m}\left\{\eta^\top r(w) - \frac{\lambda}{2}\|\eta\|^2\right\} = \frac{1}{2\lambda}\|r(w)\|^2,$$

attained uniquely at $\eta = r(w)/\lambda$. Therefore (60) becomes

$$\inf_{w\in\Delta_n}\left\{\sum_{i=1}^{n} w_i\log(nw_i) + \frac{1}{2\lambda}\left\|\mu - \sum_{i=1}^{n} w_i t_i\right\|_2^2\right\},$$

which is exactly the penalized primal (17).

Finally, strict convexity of $w \mapsto \sum_i w_i\log(nw_i)$ on Δ_n plus convexity of the squared moment-mismatch term implies the minimizer is unique. The saddle-point/KKT conditions identify it with the softmax weights at the dual optimizer: $w = \hat{w}(\hat{\eta}_\lambda)$ where $\hat{w}(\cdot)$ is defined in (9). □

D.3 Smoothness properties of A_n under bounded statistics

For completeness, we record the empirical analogue of Appendix A (Lemma 2).

Lemma 9 (Empirical gradient/Hessian bounds). *Assume* $\|t_i\| \le B$ *for all* $i \in \{1, \dots, n\}$. *Then for all* $\eta \in \mathbb{R}^m$,

$$\|\nabla A_n(\eta)\| \le B, \qquad \|\nabla^2 A_n(\eta)\|_{\mathrm{op}} \le B^2,$$

and hence ∇A_n *is* B^2*-Lipschitz:*

$$\|\nabla A_n(\eta) - \nabla A_n(\eta')\| \le B^2 \|\eta - \eta'\| \qquad \forall \eta, \eta' \in \mathbb{R}^m.$$

Proof. Using the covariance form (28) (Appendix A),

$$\begin{aligned}\nabla A_n(\eta) &= \sum_{i=1}^{n} \hat{w}_i(\eta)\, t_i \\ &\Rightarrow \quad \|\nabla A_n(\eta)\| \le \sum_{i=1}^{n} \hat{w}_i(\eta) \|t_i\| \le B.\end{aligned} \tag{61}$$

Moreover, $\nabla^2 A_n(\eta)$ is the covariance of the discrete distribution supported on $\{t_i\}$ under weights $\hat{w}(\eta)$, so

$$\begin{aligned}\|\nabla^2 A_n(\eta)\|_{\mathrm{op}} &= \sup_{\|v\|=1} \mathrm{Var}_{i \sim \hat{w}(\eta)}(v^\top t_i) \\ &\le \sup_{\|v\|=1} \mathbb{E}_{i \sim \hat{w}(\eta)}[(v^\top t_i)^2] \\ &\le \mathbb{E}_{i \sim \hat{w}(\eta)}[\|t_i\|^2] \le B^2.\end{aligned} \tag{62}$$

The Lipschitz claim follows by integrating the Hessian along the line segment between η and η'. ☐

E Proofs for Section 6 (Finite-sample generalization guarantees)

This appendix proves the finite-sample guarantees stated in Section 6 for the ridge-regularized reward estimator. The analysis follows a standard "regularized M-estimation" template: (i) localize both the population and empirical optimizers to a norm ball using first-order optimality and boundedness of T, (ii) prove uniform concentration of the (log-)partition function on that ball via covering arguments, and (iii) invoke strong convexity to translate uniform function deviation into parameter error, then convert parameter error into moment and KL control using smoothness of A. Regularization in maximum-entropy estimation is classical; see, e.g., Dudík et al. (2007), and the calibration viewpoint is emphasized in Smith et al. (2025).

E.1 Setup and population/empirical objectives

Let P_0 be the base distribution on $\mathcal{X}$, and let $T : \mathcal{X} \to \mathbb{R}^m$ be the constraint statistic. Define the population log-partition

$$A(\eta) := \log \mathbb{E}_{P_0}\left[\exp(\eta^\top T(X))\right], \qquad \eta \in \mathbb{R}^m,$$

and the empirical log-partition (given $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} P_0$ and $t_i := T(X_i)$)

$$A_n(\eta) := \log \Big(\frac{1}{n} \sum_{i=1}^{n} \exp(\eta^\top t_i) \Big).$$

Fix a target moment $\mu \in \mathbb{R}^m$ and ridge parameter $\lambda > 0$. It will be technically convenient to work with the convex minimization form of the regularized dual:

$$\begin{aligned} J(\eta) &:= A(\eta) + \tfrac{\lambda}{2}\|\eta\|^2 - \mu^\top \eta, \\ J_n(\eta) &:= A_n(\eta) + \tfrac{\lambda}{2}\|\eta\|^2 - \mu^\top \eta.\end{aligned} \tag{63}$$

Let

$$\eta_\lambda \in \arg\min_{\eta \in \mathbb{R}^m} J(\eta), \qquad \hat{\eta}_\lambda \in \arg\min_{\eta \in \mathbb{R}^m} J_n(\eta). \tag{64}$$

Equivalently, η_λ and $\hat{\eta}_\lambda$ maximize the corresponding concave objectives $\eta^\top \mu - A(\eta) - \frac{\lambda}{2}\|\eta\|^2$ and $\eta^\top \mu - A_n(\eta) - \frac{\lambda}{2}\|\eta\|^2$ (Section 5 / Appendix D).

Standing boundedness assumption. We assume throughout Section 6 and this appendix that T is bounded:

$$\|T(X)\| \le B \quad \text{almost surely under } P_0. \tag{65}$$

In particular, any feasible target moment must satisfy $\|\mu\| \le B$; we assume this henceforth. Under (65), $A(\eta)$ and $A_n(\eta)$ are finite for all η, and Appendix A (Lemma 2) yields global smoothness:

$$\|\nabla A(\eta)\| \le B, \qquad \|\nabla^2 A(\eta)\|_{\mathrm{op}} \le B^2, \qquad \forall \eta \in \mathbb{R}^m. \tag{66}$$

E.2 Localization: both optimizers lie in a deterministic ball

Lemma 10 (Norm bounds for η_λ and $\hat{\eta}_\lambda$). *Assume* (65) *and* $\|\mu\| \le B$. *Then both minimizers in* (64) *satisfy*

$$\|\eta_\lambda\| \le \frac{2B}{\lambda}, \qquad \|\hat{\eta}_\lambda\| \le \frac{2B}{\lambda}. \tag{67}$$

Proof. We use first-order optimality. Since J is differentiable and strictly convex (because of the ridge term), its minimizer η_λ is unique and satisfies

$$\begin{aligned} 0 = \nabla J(\eta_\lambda) &= \nabla A(\eta_\lambda) + \lambda \eta_\lambda - \mu \\ \Longrightarrow \quad \lambda \eta_\lambda &= \mu - \nabla A(\eta_\lambda).\end{aligned} \tag{68}$$

Taking norms and using (66) gives

$$\lambda \|\eta_\lambda\| \le \|\mu\| + \|\nabla A(\eta_\lambda)\| \le B + B = 2B,$$

hence $\|\eta_\lambda\| \le 2B/\lambda$.

The same argument applies to $\hat{\eta}_\lambda$ because J_n is also differentiable and strictly convex:

$$\begin{aligned} 0 &= \nabla J_n(\hat{\eta}_\lambda) = \nabla A_n(\hat{\eta}_\lambda) + \lambda\hat{\eta}_\lambda - \mu \\ &\implies \quad \lambda\hat{\eta}_\lambda = \mu - \nabla A_n(\hat{\eta}_\lambda). \end{aligned} \tag{69}$$

Under (65) we have $\|t_i\| \le B$ almost surely for all i, and Appendix D (Lemma 9) yields $\|\nabla A_n(\eta)\| \le B$ for all η, hence $\lambda\|\hat{\eta}_\lambda\| \le \|\mu\| + B \le 2B$. □

In the remainder, set

$$R := \frac{2B}{\lambda}. \tag{70}$$

Lemma 10 shows that both η_λ and $\hat{\eta}_\lambda$ lie in the Euclidean ball $B_2(R)$.

E.3 Uniform concentration of the log-partition on $B_2(R)$

We now prove the uniform deviation lemma used in Section 6.

Lemma 11 (Uniform deviation of A_n on a norm ball)**.** *Assume* (65)*. Fix $R > 0$, $\rho \in (0, R]$, and $\delta \in (0,1)$. Then with probability at least $1-\delta$ (over $X_1, \dots, X_n \sim P_0$),*

$$\begin{aligned} &\sup_{\|\eta\|\le R} |A_n(\eta) - A(\eta)| \\ &\quad \le 2B\rho + \left(e^{2BR} - 1\right)\sqrt{\frac{m\log(1+\frac{2R}{\rho}) + \log(\frac{2}{\delta})}{2n}}. \end{aligned} \tag{71}$$

Proof. **Step 1 (an ρ-net).** Let $\mathcal{N} \subseteq B_2(R)$ be an ρ-net in Euclidean norm. By (34) in Appendix A, we may take

$$|\mathcal{N}| \le \left(1 + \frac{2R}{\rho}\right)^m. \tag{72}$$

Step 2 (pointwise concentration for each net point). Fix $\eta \in \mathcal{N}$. Define $Y := \exp(\eta^\top T(X))$ and $Y_i := \exp(\eta^\top T(X_i))$. Under (65) and $\|\eta\| \le R$, we have deterministically

$$\begin{aligned} \eta^\top T(X) \in [-BR, BR] \Rightarrow Y \in \left[e^{-BR}, e^{BR}\right], \\ Y_i \in \left[e^{-BR}, e^{BR}\right] \forall i. \end{aligned} \tag{73}$$

Hence Hoeffding's inequality (33) gives: for any $\delta_\eta \in (0,1)$,

$$\begin{aligned} &\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^n Y_i - \mathbb{E}[Y]\right| \ge \left(e^{BR} - e^{-BR}\right)\sqrt{\frac{\log(2/\delta_\eta)}{2n}}\right) \\ &\quad \le \delta_\eta. \end{aligned} \tag{74}$$

Since $A_n(\eta) = \log\left(\frac{1}{n}\sum_i Y_i\right)$ and $A(\eta) = \log(\mathbb{E}[Y])$, and both arguments lie in $[e^{-BR}, e^{BR}]$, the map $\log(\cdot)$ is e^{BR}-Lipschitz on this interval. Therefore, on the event above,

$$\begin{aligned} |A_n(\eta) - A(\eta)| &\le e^{BR}\left|\frac{1}{n}\sum_{i=1}^n Y_i - \mathbb{E}[Y]\right| \\ &\le e^{BR}\left(e^{BR} - e^{-BR}\right)\sqrt{\frac{\log(2/\delta_\eta)}{2n}} \\ &= \left(e^{2BR} - 1\right)\sqrt{\frac{\log(2/\delta_\eta)}{2n}}. \end{aligned} \tag{75}$$

Step 3 (union bound over the net). Set $\delta_\eta := \delta/|\mathcal{N}|$ and apply a union bound over $\eta \in \mathcal{N}$. With probability at least $1-\delta$,

$$\sup_{\eta\in\mathcal{N}} |A_n(\eta) - A(\eta)| \le \left(e^{2BR} - 1\right)\sqrt{\frac{\log(2|\mathcal{N}|/\delta)}{2n}}. \tag{76}$$

Using (72) and $\log(2|\mathcal{N}|/\delta) \le \log(2/\delta) + m\log(1 + 2R/\rho)$ yields

$$\begin{aligned} &\sup_{\eta\in\mathcal{N}} |A_n(\eta) - A(\eta)| \\ &\quad \le \left(e^{2BR} - 1\right)\sqrt{\frac{m\log(1+2R/\rho) + \log(2/\delta)}{2n}}. \end{aligned} \tag{77}$$

Step 4 (extend from the net to the full ball). Let $\eta \in B_2(R)$ and choose $\eta_0 \in \mathcal{N}$ with $\|\eta - \eta_0\| \le \rho$. Under (65), Appendix A (Lemma 2) and Appendix D (Lemma 9) imply that both A and A_n are B-Lipschitz:

$$\begin{aligned} |A(\eta) - A(\eta_0)| \le B\|\eta - \eta_0\| \le B\rho, \\ |A_n(\eta) - A_n(\eta_0)| \le B\|\eta - \eta_0\| \le B\rho. \end{aligned} \tag{78}$$

Hence

$$\begin{aligned} |A_n(\eta) - A(\eta)| &\le |A_n(\eta_0) - A(\eta_0)| + |A_n(\eta) - A_n(\eta_0)| \\ &\quad + |A(\eta) - A(\eta_0)| \\ &\le \sup_{\xi\in\mathcal{N}} |A_n(\xi) - A(\xi)| + 2B\rho. \end{aligned} \tag{79}$$

Combining with (77) proves (71). □

Corollary 3 (A convenient instantiation)**.** *Under the assumptions of Lemma 11, choosing $\rho = R/n$ yields: with probability at least $1-\delta$,*

$$\begin{aligned} &\sup_{\|\eta\|\le R} |A_n(\eta) - A(\eta)| \\ &\quad \le \frac{2BR}{n} + \left(e^{2BR} - 1\right)\sqrt{\frac{m\log(1+2n) + \log(2/\delta)}{2n}}. \end{aligned} \tag{80}$$

E.4 From uniform deviation to parameter, moment, and KL bounds (proof of Theorem 3)

We now prove the main generalization theorem from Section 6.

Lemma 12 (Bregman upper bound from smoothness)**.** *Assume* (65)*. Then for any* $\eta, \eta' \in \mathbb{R}^m$*,*

$$D_A(\eta', \eta) := A(\eta') - A(\eta) - \langle \nabla A(\eta), \eta' - \eta \rangle \le \frac{B^2}{2} \|\eta' - \eta\|^2. \tag{81}$$

Proof. By Taylor's theorem with integral remainder and twice differentiability of A,

$$D_A(\eta', \eta) = \int_0^1 (1-t)\,(\eta' - \eta)^\top \nabla^2 A\big(\eta + t(\eta' - \eta)\big) \times (\eta' - \eta)\, dt. \tag{82}$$

Using $\|\nabla^2 A(\cdot)\|_{\mathrm{op}} \le B^2$ from (66) gives $D_A(\eta', \eta) \le \int_0^1 (1-t) B^2 \|\eta' - \eta\|^2 dt = \frac{B^2}{2}\|\eta' - \eta\|^2$. □

Proof of Theorem 3. Let $R = 2B/\lambda$ as in (70). By Lemma 10, $\eta_\lambda, \hat\eta_\lambda \in B_2(R)$.

Step 1 (reduce estimation to uniform deviation of A_n). Define $\varepsilon_A := \sup_{\|\eta\| \le R} |A_n(\eta) - A(\eta)|$. Since $J_n(\hat\eta_\lambda) \le J_n(\eta_\lambda)$ and $J(\eta_\lambda) \le J(\hat\eta_\lambda)$ by optimality,

$$\begin{aligned} &J(\hat\eta_\lambda) - J(\eta_\lambda) \\ &= \big(J(\hat\eta_\lambda) - J_n(\hat\eta_\lambda)\big) + \big(J_n(\hat\eta_\lambda) - J_n(\eta_\lambda)\big) + \big(J_n(\eta_\lambda) - J(\eta_\lambda)\big) \\ &\le |J(\hat\eta_\lambda) - J_n(\hat\eta_\lambda)| + |J_n(\eta_\lambda) - J(\eta_\lambda)| \\ &= |A(\hat\eta_\lambda) - A_n(\hat\eta_\lambda)| + |A_n(\eta_\lambda) - A(\eta_\lambda)| \\ &\le 2\varepsilon_A. \end{aligned} \tag{83}$$

where we used that the ridge and linear terms cancel in $J_n - J$ (see (63)).

Step 2 (parameter error via strong convexity). Since A is convex, $J(\eta) = A(\eta) + \frac{\lambda}{2}\|\eta\|^2 - \mu^\top \eta$ is λ-strongly convex. Therefore, for all η,

$$J(\eta) \ge J(\eta_\lambda) + \frac{\lambda}{2}\|\eta - \eta_\lambda\|^2.$$

Applying this at $\eta = \hat\eta_\lambda$ and combining with (83) yields

$$\frac{\lambda}{2}\|\hat\eta_\lambda - \eta_\lambda\|^2 \le J(\hat\eta_\lambda) - J(\eta_\lambda) \le 2\varepsilon_A,$$

hence

$$\|\hat\eta_\lambda - \eta_\lambda\| \le 2\sqrt{\frac{\varepsilon_A}{\lambda}}. \tag{84}$$

Invoking Lemma 11 (or Corollary 3) bounds ε_A with probability $1 - \delta$, yielding the stated finite-sample parameter bound.

Step 3 (population residual and moment error). Recall $\nabla A(\eta) = \mathbb{E}_{Q_\eta}[T(X)]$ (Appendix A, (27)). The population first-order condition at η_λ is

$$\mu - \mathbb{E}_{Q_{\eta_\lambda}}[T(X)] = \lambda \eta_\lambda. \tag{85}$$

Define the *population residual* at $\hat\eta_\lambda$:

$$r(\hat\eta_\lambda) := \mu - \mathbb{E}_{Q_{\hat\eta_\lambda}}[T(X)] - \lambda \hat\eta_\lambda.$$

Subtracting (85) gives

$$r(\hat\eta_\lambda) = \Big(\mathbb{E}_{Q_{\eta_\lambda}}[T(X)] - \mathbb{E}_{Q_{\hat\eta_\lambda}}[T(X)]\Big) + \lambda(\eta_\lambda - \hat\eta_\lambda).$$

Using Lipschitzness of ∇A from (66), $\|\mathbb{E}_{Q_\eta}[T] - \mathbb{E}_{Q_{\eta'}}[T]\| = \|\nabla A(\eta) - \nabla A(\eta')\| \le B^2 \|\eta - \eta'\|$, we obtain

$$\|r(\hat\eta_\lambda)\| \le (B^2 + \lambda)\|\hat\eta_\lambda - \eta_\lambda\|. \tag{86}$$

Similarly, the population moment error relative to μ obeys

$$\begin{aligned} &\|\mathbb{E}_{Q_{\hat\eta_\lambda}}[T] - \mu\| \\ &\le \|\mathbb{E}_{Q_{\hat\eta_\lambda}}[T] - \mathbb{E}_{Q_{\eta_\lambda}}[T]\| + \|\mathbb{E}_{Q_{\eta_\lambda}}[T] - \mu\| \\ &\le B^2 \|\hat\eta_\lambda - \eta_\lambda\| + \lambda \|\eta_\lambda\|, \end{aligned} \tag{87}$$

where the second term used (85). This decomposes the error into an *estimation term* and a *regularization bias* term.

Step 4 (KL deviation). Using the exponential-family identity (29) in Appendix A,

$$\mathrm{KL}(Q_{\hat\eta_\lambda} \| Q_{\eta_\lambda}) = D_A(\eta_\lambda, \hat\eta_\lambda).$$

Applying Lemma 12 with $\eta' = \eta_\lambda$ and $\eta = \hat\eta_\lambda$ yields

$$\mathrm{KL}(Q_{\hat\eta_\lambda} \| Q_{\eta_\lambda}) \le \frac{B^2}{2}\|\hat\eta_\lambda - \eta_\lambda\|^2. \tag{88}$$

Combining with (84) gives a KL bound in terms of ε_A (and hence in terms of (n, m, δ) via Lemma 11).

Step 5 (assemble the high-probability statement). Finally, choose ρ in Lemma 11 (e.g. Corollary 3) to obtain an explicit bound $\varepsilon_A \le \varepsilon_A(n, m, R, \delta)$ with probability at least $1-\delta$. On this event, the parameter bound (84), residual bound (86), moment bound (87), and KL bound (88) all hold simultaneously. □

F Experimental details and reproducibility

This appendix documents experimental configurations, implementation details, and reproducibility notes for Section 7. All experiments are synthetic and are designed to validate the geometric feasibility theory (Section 4) and the regularized reward estimator (Sections 5–6) in controlled settings.

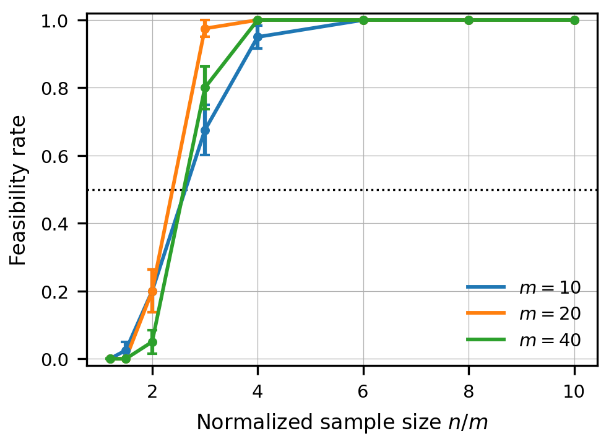

Figure 3: **(Appendix) Feasibility of unregularized reward calibration.** Empirical feasibility rate of the event $\mu \in \mathrm{conv}\{T(X_i)\}_{i=1}^n$ versus normalized sample size n/m for bounded constraints $T_j(X) = \tanh(a_j^\top X)$ and $m \in \{10, 20, 40\}$. Error bars denote ± 1 s.e.

F.1 Software and numerical solvers

All experiments were implemented in Python using `numpy` for linear algebra and random number generation, `scipy` for optimization, and `matplotlib` for plotting. Convex-hull membership events were tested via linear programming feasibility using `scipy.optimize.linprog` with the HiGHS backend (`method="highs"`), solving

$$\exists\, w \in \Delta_n \ \text{ s.t. } \ \sum_{i=1}^n w_i t_i = \mu,$$

which is equivalent to $\mu \in \mathrm{conv}\{t_1, \dots, t_n\}$ (Lemma 1). Regularized dual optimization was performed using L-BFGS-B (`scipy.optimize.minimize` with `method="L-BFGS-B"`) on the smooth, strongly convex objective (63). We used numerically stable `logsumexp` and softmax computations to avoid overflow when evaluating $A_n(\eta)$ and its gradient.

F.2 Experiment 7.1: Wendel sanity check (Figure 1)

Goal. Validate the closed-form feasibility probability in the symmetric, continuous case and empirically demonstrate the sharp transition near $n/(2m) \approx 1$ predicted by Wendel's formula (Corollary 1) (Tikhomirov, 2023).

Data model. For each dimension $m \in \{10, 20, 50, 100\}$ and sample size n on a grid ranging from $m+1$ up to $4m$, we draw $Z_1, \dots, Z_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, I_m)$ and test feasibility of $\mu = 0$:

$$\mathbb{I}\{0 \in \mathrm{conv}\{Z_1, \dots, Z_n\}\}.$$

The Monte Carlo estimate $\hat{p}_n(0)$ is computed over 200 independent repetitions, with standard error $\sqrt{\hat{p}_n(0)(1 - \hat{p}_n(0))/200}$.

Theory curve. The Wendel probability is computed in closed form as

$$\begin{aligned} p_n(0) &= 1 - 2^{-n+1} \sum_{k=0}^{m-1} \binom{n-1}{k} \\ &= 1 - \mathbb{P}\big(B \le m-1\big), \quad B \sim \mathrm{Bin}(n-1, 1/2). \end{aligned} \tag{89}$$

and plotted alongside Monte Carlo estimates (Figure 1).

Convex-hull feasibility test. We solve the linear feasibility program:

$$\text{find } w \in \mathbb{R}^n \text{ s.t. } w \ge 0,\ \mathbf{1}^\top w = 1,\ Z^\top w = 0,$$

where $Z \in \mathbb{R}^{n \times m}$ has rows Z_i. A trial is marked feasible if the HiGHS solver returns a feasible point with maximal constraint residual at most 10^{-7} in ℓ_∞ norm.

F.3 Experiment 7.2: bounded-statistics calibration and regularization tradeoff

Goal. Demonstrate (i) the nontrivial feasibility transition of unregularized empirical reward calibration as n/m varies, and (ii) that ridge-regularized reward calibration is always well-defined and exhibits a bias–variance tradeoff in population moment error (Figures 3 and 2).

Base distribution and bounded statistics. We set $P_0 = \mathcal{N}(0, I_d)$ with $d = 10$. For each constraint dimension $m \in \{10, 20, 40\}$, we draw random unit vectors $a_1, \dots, a_m \in \mathbb{R}^d$ by sampling i.i.d. standard Gaussians and normalizing each row. We define bounded constraint statistics

$$T_j(X) := \tanh(a_j^\top X) \in [-1, 1], \qquad j = 1, \dots, m,$$

so $\|T(X)\| \le \sqrt{m}$ almost surely and the bounded-statistics condition holds with $B = \sqrt{m}$.

Constructing a population-feasible target moment. We construct a target moment μ by sampling a ground-truth tilt parameter $\eta^\star \in \mathbb{R}^m$ with fixed norm $\|\eta^\star\|_2 = 1$, independent of m. Using a large Monte Carlo pool $X_1^{\text{pool}}, \dots, X_{N_\mu}^{\text{pool}} \overset{\text{i.i.d.}}{\sim} P_0$ (with

$N_\mu = 150{,}000$), we compute $t_k^{\text{pool}} := T(X_k^{\text{pool}})$ and approximate

$$\mu = \mathbb{E}_{Q_{\eta^\star}}[T(X)]$$

via self-normalized importance sampling:

$$\hat{\mu} = \sum_{k=1}^{N_\mu} \tilde{w}_k \, t_k^{\text{pool}}, \quad \tilde{w}_k \propto \exp\big((\eta^\star)^\top t_k^{\text{pool}}\big), \quad \sum_{k=1}^{N_\mu} \tilde{w}_k = 1. \tag{90}$$

We use $\hat{\mu}$ as μ in the experiments. Fixing $\|\eta^\star\|$ avoids making μ artificially extreme as m grows, which would trivially drive feasibility to zero.

Unregularized feasibility sweep (Figure 3). For each $m \in \{10, 20, 40\}$ and each multiplier $c \in \{1.2, 1.5, 2, 3, 4, 6, 8, 10\}$ we set $n = \lceil c\, m \rceil$ (and enforce $n \geq m + 1$). For each (m, n), we run 40 independent repetitions: draw $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} P_0$, set $t_i = T(X_i)$, and test whether

$$\mu \in \operatorname{conv}\{t_1, \ldots, t_n\}$$

via the LP feasibility program described above. We report the empirical feasibility rate and ± 1 standard error.

Regularized tradeoff (Figure 2). We fix $(m, n) = (40, 80)$ and solve the ridge-regularized estimator

$$\hat{\eta}_\lambda \in \arg\max_\eta \; \eta^\top \mu - A_n(\eta) - \frac{\lambda}{2}\|\eta\|^2$$

for $\lambda \in \{10^{-3}, 10^{-2}, 10^{-1}, 1, 10, 10^2\}$ using L-BFGS-B with tolerances `ftol=1e-12` and `gtol=1e-8`, and `maxiter=400`. Population moment error is estimated using an independent evaluation pool of size $N_{\text{eval}} = 50{,}000$: draw $X_k^{\text{eval}} \overset{\text{i.i.d.}}{\sim} P_0$, compute $t_k^{\text{eval}} = T(X_k^{\text{eval}})$, and approximate

$$\mathbb{E}_{Q_{\hat{\eta}_\lambda}}[T(X)] \approx \sum_{k=1}^{N_{\text{eval}}} \bar{w}_k \, t_k^{\text{eval}}, \qquad \bar{w}_k \propto \exp\big(\hat{\eta}_\lambda^\top t_k^{\text{eval}}\big).$$

We report $\|\mathbb{E}_{Q_{\hat{\eta}_\lambda}}[T] - \mu\|_2$ averaged over 30 repetitions (fresh n-sample each time) with ± 1 standard error. The plotted dotted baseline corresponds to $Q = P_0$ (i.e., $\eta = 0$), estimated as $\|\mathbb{E}_{P_0}[T] - \mu\|_2$ from the evaluation pool.

F.4 Random seeds and determinism

All random number generation uses `numpy`'s `default_rng` with a fixed seed (default 0) to ensure bitwise reproducibility of figures on a fixed software stack. The optimization routines are deterministic given fixed data. Minor numerical differences across platforms can arise due to differences in BLAS/LAPACK implementations and HiGHS versions, but the qualitative conclusions (phase transitions and bias–variance behavior) are stable.

F.5 Connection to theory

Experiment 7.1 directly visualizes Corollary 1 (Tikhomirov, 2023). Experiment 7.2 illustrates both the convex-hull feasibility phenomenon quantified via depth in Section 4 (Hayakawa et al., 2023) and the stabilizing role of ridge regularization in exponential-tilting / maximum-entropy estimation (Dudík et al., 2007; Smith et al., 2025).