CALIBRATED α-BARYCENTERS FOR MULTIOBJECTIVE f-DIVERGENCE OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Matching one distribution to many targets can require repeated evaluations. CAL-α-MODM aggregates a nonlinear predictor into one barycenter, calibrates its residual derivatives, and checks steps against the original f-divergence objective. Residual calibration controls step-local model error under smoothness assumptions without requiring targets to be close to the model. Under stated exact-oracle assumptions, we establish convergence to stationarity. A local expansion shows that canonical geometry cancels the leading quadratic iteration error at common matching, yielding cubic convergence with second-order calibration and sufficiently accurate inner solves. High-precision experiments support this local prediction. Sensitivity prediction accelerates preference continuation along regular stationary branches. On 300-target Jensen–Shannon problems, CAL-p2 achieves median paired speedups of $3.15\times$ over L-BFGS and $1.10\times$ over second-order Direct-TR, across 23/25 and 24/25 objective-comparable stationary pairs, respectively. Gains arise when predictor savings outweigh calibration overhead.

1 INTRODUCTION

Multiple distributional targets arise in ensemble compression and generative modeling (Zhang et al., 2025). Given $P_1, \dots, P_m$ and $\mathcal{M} = \{Q_\theta : \theta \in \Theta \subseteq \mathbb{R}^d\}$, fix an orientation $\sigma \in \{-, +\}$ and consider

$$\min_{\theta\in\Theta}\big(F_1^\sigma(\theta), \dots, F_m^\sigma(\theta)\big), \qquad F_i^\sigma(\theta) = \begin{cases} D_f(P_i\|Q_\theta), & \sigma = -, \\ D_f(Q_\theta\|P_i), & \sigma = +. \end{cases} \tag{1}$$

All objectives share f and σ; positive preferences specify the trade-off. We study fixed scalarizations with exact oracles, whose repeated evaluations can require revisiting every target.

Our key idea is to reuse one barycenter for predictor queries and restore the true objective's local derivatives by residual calibration. Exact α-divergence scalarizations aggregate, but generic f need not: likelihood ratios need not approach one as parameter steps shrink, so local divergence geometry alone is insufficient.

We Taylor-correct the true objective minus the predictor and check steps against the true scalarization, following trust-region model management (Alexandrov et al., 1998; Booker et al., 1999). Aggregation lowers predictor-query cost; its nonlinear remainder can also reduce local iteration error beyond a matched quadratic Taylor model.

Contributions.

1. **Compressed models with controlled accuracy.** Exact α-aggregation and residual calibration give parameter-step error bounds without target proximity, with stationarity and certified-iteration complexity guarantees.
2. **Geometry beyond the matched Taylor jet.** The residual third derivative determines the leading local error coefficient for second-order calibration. Canonical geometry cancels it at common matching, giving a conditional cubic rate without third-order calibration; E10 audits this.
3. **Measured many-target crossover.** For 300 JS targets, CAL-p2 gives median paired speedups of $3.15\times$ over L-BFGS and $1.10\times$ over second-order Direct-TR, with 23/25 and 24/25 objective-comparable stationary pairs. Latent targets demonstrate preference-controlled generation.

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.

2 Background and Related Work

Let the distributions have positive densities relative to a common measure μ. For convex f, normalize $f(1) = f'(1) = 0$ and set $a_f = f''(1) > 0$. Define (Csiszár, 1967; Amari, 2016)

$$D_f(P\|Q) = \int q(x) f\left(\frac{p(x)}{q(x)}\right) \mathrm{d}\mu(x). \tag{2}$$

A preference $\lambda \in \Delta_m := \{\lambda \in \mathbb{R}^m : \lambda_i \geq 0,\ \sum_i \lambda_i = 1\}$ gives

$$F_\lambda^\sigma(\theta) = \sum_i \lambda_i F_i^\sigma(\theta). \tag{3}$$

Write $D_g^-(P,Q) = D_g(P\|Q)$ and $D_g^+(P,Q) = D_g(Q\|P)$ for any generator g; λ and σ remain fixed unless explicitly varied.

Distribution matching and aggregation. Zhang et al. (2025) characterize KL and inverse-KL Pareto sets in full exponential families and develop multiobjective variational autoencoders for distribution interpolation. Bregman centroid identities reduce weighted objectives to single-target problems (Banerjee et al., 2005; Nielsen & Nock, 2009); power-mean annealing paths admit α-divergence characterizations (Masrani et al., 2021). Transformed Bregman divergences characterize quasi-arithmetic means (Brekelmans & Nielsen, 2024), which also generate broader divergence families (Nielsen, 2022). Jeffreys proxy centers provide efficient centroid approximations (Nielsen, 2024). We address general smooth f and restricted model families: the barycenter supplies a compressible predictor even when it cannot itself be represented by Q_θ.

Calibration and model management. Trust-region model management provides convergence guarantees for approximation-based optimization (Alexandrov et al., 1998; Booker et al., 1999); evaluation-complexity theory (Cartis et al., 2022) and higher-order local analyses (Doikov & Nesterov, 2022) supply complementary global and local perspectives. Our analysis specializes these foundations to a nonlinear aggregated predictor with calibrated residual derivatives. Appendix A covers complementary learning applications and preference-based methods.

3 Calibrated Barycenter Predictors

3.1 Canonical Local α Geometry

For $\alpha \in \mathbb{R}$, define

$$D_\alpha(P\|Q) = \frac{4}{1-\alpha^2}\left(1 - \int p^{(1-\alpha)/2} q^{(1+\alpha)/2}\, \mathrm{d}\mu\right), \tag{4}$$

with continuous limits $D_{-1}(P\|Q) = \mathrm{KL}(P\|Q)$ and $D_1(P\|Q) = \mathrm{KL}(Q\|P)$. For $\alpha \neq \pm 1$, an affine-normalized generator is

$$f_\alpha(u) = \frac{4}{1-\alpha^2}(1 - u^{(1-\alpha)/2}) + \frac{2}{1+\alpha}(u-1), \tag{5}$$

with endpoint limits $f_{-1}(u) = u\log u - u + 1$ and $f_1(u) = -\log u + u - 1$. It satisfies $f_\alpha''(1) = 1$ and $f_\alpha'''(1) = -(3+\alpha)/2$.

Theorem 3.1 (Fourth-order local normal form)**.** *Let $0 < \delta < 1$, $f \in C^4([1-\delta, 1+\delta])$, and define*

$$\alpha_f := -3 - 2\frac{f'''(1)}{f''(1)}, \qquad h_f := f - a_f f_{\alpha_f}. \tag{6}$$

Then $h_f^{(j)}(1) = 0$ for $j = 0,1,2,3$. If $p/q \in [1-\delta, 1+\delta]$ almost everywhere, then

$$\left|D_f(P\|Q) - a_f D_{\alpha_f}(P\|Q)\right| \leq \frac{M_{f,\delta}}{24} \int q\,|p/q - 1|^4\, \mathrm{d}\mu, \tag{7}$$

where $M_{f,\delta} = \sup_{u\in[1-\delta,1+\delta]} \left|h_f^{(4)}(u)\right|$.

The bound controls likelihood-ratio displacement and does not, by itself, imply model accuracy as a parameter step tends to zero.

3.2 EXACT AGGREGATION

Let $r_\alpha = (1-\alpha)/2$ and $s_\alpha = (1+\alpha)/2$. For nonzero exponents define

$$c^-_{\lambda,\alpha} = \frac{(\sum_i \lambda_i p_i^{r_\alpha})^{1/r_\alpha}}{Z^-_{\lambda,\alpha}}, \qquad Z^-_{\lambda,\alpha} = \int (\sum_i \lambda_i p_i^{r_\alpha})^{1/r_\alpha} \, \mathrm{d}\mu, \tag{8}$$

$$c^+_{\lambda,\alpha} = \frac{(\sum_i \lambda_i p_i^{s_\alpha})^{1/s_\alpha}}{Z^+_{\lambda,\alpha}}, \qquad Z^+_{\lambda,\alpha} = \int (\sum_i \lambda_i p_i^{s_\alpha})^{1/s_\alpha} \, \mathrm{d}\mu. \tag{9}$$

For a zero exponent, normalize the geometric mean $\prod_i p_i^{\lambda_i}$. We assume finite, strictly positive normalizers and finite divergences wherever the optimization model is evaluated.

Theorem 3.2 (Exact one-target aggregation)**.** *For $\alpha \neq \pm 1$ and $0 < Z^\sigma_{\lambda,\alpha} < \infty$,*

$$\sum_i \lambda_i D_\alpha(P_i \| Q) = C^-_{\lambda,\alpha} + \omega^-_{\lambda,\alpha} D_\alpha(c^-_{\lambda,\alpha} \| Q), \qquad \omega^-_{\lambda,\alpha} = (Z^-_{\lambda,\alpha})^{r_\alpha}, \tag{10}$$

$$\sum_i \lambda_i D_\alpha(Q \| P_i) = C^+_{\lambda,\alpha} + \omega^+_{\lambda,\alpha} D_\alpha(Q \| c^+_{\lambda,\alpha}), \qquad \omega^+_{\lambda,\alpha} = (Z^+_{\lambda,\alpha})^{s_\alpha}. \tag{11}$$

The constants $C^\sigma_{\lambda,\alpha}$ do not depend on Q. The identities extend to $\alpha = \pm 1$ whenever all displayed terms admit finite continuous limits.

The positive factor ω^σ does not change the predictor's minimizers over a fixed model family, but is required to reproduce the multi-target predicted reductions. Omitting it generally changes the acceptance ratio.

3.3 DERIVATIVE CALIBRATION

Fix λ and σ and write $F = F^\sigma_\lambda$. Define the structured predictor

$$\mathcal{A}(\theta) := a_f \sum_i \lambda_i D^\sigma_{\alpha_f}(P_i, Q_\theta), \tag{12}$$

which equals a constant plus one weighted divergence to $c^\sigma_{\lambda,\alpha_f}$ by Theorem 3.2. This identity compresses $\mathcal{A}$, while F remains the objective to be optimized. We analyze the Euclidean unconstrained case; at iteration t, define $\widehat{F}_t(v) = F(\theta_t + v)$, $\widehat{\mathcal{A}}_t(v) = \mathcal{A}(\theta_t + v)$, and $\widehat{\mathcal{E}}_t = \widehat{F}_t - \widehat{\mathcal{A}}_t$.

Definition 3.3 (Order-p calibrated model)**.** For $p \in \{1, 2, 3\}$,

$$m_t^{(p)}(v) = \widehat{\mathcal{A}}_t(v) + \sum_{j=0}^{p} \frac{1}{j!} \nabla^j \widehat{\mathcal{E}}_t(0)[v]^j. \tag{13}$$

Let $T_p G(v) := \sum_{j=0}^p \nabla^j G(0)[v]^j / j!$ denote the Taylor polynomial at zero. Equivalently,

$$m_t^{(p)}(v) = T_p \widehat{F}_t(v) + \big(\widehat{\mathcal{A}}_t(v) - T_p \widehat{\mathcal{A}}_t(v)\big). \tag{14}$$

Calibration matches the true order-p jet (value and derivatives); the predictor supplies the higher-order nonlinear remainder. Thus α changes this remainder, not the matched Taylor jet. E7 tests finite-step accuracy, E10 tests the resulting local iteration error, and E9 tests whether compressed queries reduce total runtime.

Theorem 3.4 (Model order)**.** *Suppose $\widehat{F}_t, \widehat{\mathcal{A}}_t \in C^{p+1}$ on a neighborhood of the trial ball and $\sup_{\|v\| \le \Delta_t} \left\| \nabla^{p+1} \widehat{\mathcal{E}}_t(v) \right\|_{\mathrm{op}} \le L_{p+1}$. For $\|v\| \le \Delta_t$,*

$$\nabla^j m_t^{(p)}(0) = \nabla^j \widehat{F}_t(0), \quad 0 \le j \le p, \tag{15}$$

$$\left| \widehat{F}_t(v) - m_t^{(p)}(v) \right| \le \frac{L_{p+1}}{(p+1)!} \|v\|^{p+1}, \tag{16}$$

$$\left\| \nabla \widehat{F}_t(v) - \nabla m_t^{(p)}(v) \right\| \le \frac{L_{p+1}}{p!} \|v\|^p. \tag{17}$$

Target proximity is unnecessary in Theorem 3.4. Near matching, bounded scores and log-density Hessians give residual gradient and Hessian norm bounds in terms of third and second absolute moments of the likelihood-ratio displacement, respectively (Appendix B.2).

4 CAL-α-MODM

At a certified iteration, approximately minimize $m_t^{(p)}$ over $\|v\| \le \Delta_t$ and form

$$\mathrm{pred}_t = m_t^{(p)}(0) - m_t^{(p)}(s_t), \quad \mathrm{ared}_t = F(\theta_t) - F(\theta_t + s_t), \quad \rho_t = \frac{\mathrm{ared}_t}{\mathrm{pred}_t}. \tag{18}$$

Let $g_t := \nabla F(\theta_t) = \nabla m_t^{(p)}(0)$. We require the approximate trust-region solution to achieve a fixed fraction of Cauchy decrease:

$$\mathrm{pred}_t \ge \kappa_c \|g_t\| \min\left\{\Delta_t, \frac{\|g_t\|}{\kappa_H}\right\}, \qquad \kappa_c \in (0, 1], \tag{19}$$

where $\kappa_H > 0$ uniformly bounds $\left\|\nabla^2 m_t^{(p)}(v)\right\|_{\mathrm{op}}$ on the trust region. A safeguarded gradient step gives $\kappa_c = 1/2$ (Appendix C.2); larger values require verified decrease.

For $0 < \eta_1 < \eta_2 < 1$, $0 < \gamma_1 < 1 < \gamma_2$, use the explicit radius update

$$\Delta_{t+1} = \begin{cases} \gamma_1 \Delta_t, & \rho_t < \eta_1, \\ \min\{\gamma_2 \Delta_t, \Delta_{\max}\}, & \rho_t \ge \eta_2, \\ \Delta_t, & \text{otherwise.} \end{cases} \tag{20}$$

Algorithm 1: CAL-α-MODM

1. Given θ_0 and $0 < \Delta_0 \le \Delta_{\max}$, normalize f, compute (a_f, α_f), and construct $(c^{\sigma}_{\lambda,\alpha_f}, \omega^{\sigma}_{\lambda,\alpha_f})$.
2. At the certified point θ_t, evaluate F and the residual jet through order p. Stop only if $\|\nabla F(\theta_t)\| \le \varepsilon$.
3. Build (13) and compute $\|s_t\| \le \Delta_t$ satisfying (19).
4. Evaluate the true scalarized objective at $\theta_t + s_t$. Set $\theta_{t+1} = \theta_t + s_t$ if $\rho_t \ge \eta_1$ and $\theta_{t+1} = \theta_t$ otherwise.
5. Update Δ_{t+1} by (20) and repeat.

Choice of order. The order p specifies the residual derivatives: $p = 1$ requires values and gradients; $p = 2$ adds Hessians and enables local quadratic convergence under Theorem 5.3. A p1 inner solver may still use predictor Hessians. The $p = 3$ variant provides local cubic convergence for small problems. Appendix C.6 gives sufficient safeguards for deferred calibration.

Inner solves and reuse. E8/E9 use regularized Newton iterations with trust-region projection, Armijo backtracking, and a steepest-descent comparator. The center Hessian initializes its step size; it is not a bound over the ball. Appendix C.7 relates backtracking to sufficient Cauchy decrease and details the finite implementation caps. At a fixed anchor, the residual jet is reused across inner queries and rejected trials. Shrinking the radius tightens the model-error bound; accepted steps require fresh calibration.

What is actually compressed? Cached aggregate values on a reusable support or grid remove the target sum from predictor queries, which return only the requested derivative order. New nodes may require revisiting all targets. True evaluations remain necessary for calibration and acceptance; savings must offset this overhead. The aggregation identity itself is independent of quadrature. For fixed preferences the aggregate is reusable; a preference change requires constructing a new aggregate.

Higher-order information without higher-order calibration. With $p = 2$, the model preserves the true Newton jet but retains the nonlinear predictor beyond that jet. Inner evaluations exploit this geometry without forming a true third-derivative tensor. Proposition 5.4 explains how it changes the leading local iteration error; E10 tests the effect at identical calibration order.

5 Convergence and Pareto guarantees

We state the result for $\Theta = \mathbb{R}^d$. Assume a compact initial level set, a lower-bounded F with Lipschitz gradient, uniformly bounded residual derivatives and model Hessians, and the Cauchy decrease (19). Appendix C.2 gives the precise assumptions and proofs.

Theorem 5.1 (Global convergence and complexity). *For fixed λ and orientation, the zero-tolerance version of Algorithm 1 either stops at an exact stationary point or satisfies*

$$\lim_{t\to\infty} \|\nabla F_\lambda^\sigma(\theta_t)\| = 0. \tag{21}$$

For stopping tolerance $\varepsilon > 0$, there are problem-dependent constants $C_0, C_1 > 0$ such that an iterate with $\|\nabla F_\lambda^\sigma\| \le \varepsilon$ is obtained within

$$N(\varepsilon) \le C_0 + C_1 \varepsilon^{-2} \tag{22}$$

certified iterations and true scalarized trial evaluations. With explicit summation, each scalarized value or order-p jet call requires m component calls; the constants in (22) may depend on m.

This is the standard smooth nonconvex trust-region exponent (Conn et al., 2000; Cartis et al., 2010); higher p improves local accuracy, not the global exponent or certified component-oracle factor m.

Corollary 5.2 (Pareto stationarity). *For C^1 components and $\lambda \in \mathrm{ri}(\Delta_m)$, every accumulation point in Theorem 5.1 is Pareto stationary with a strictly positive multiplier. If each component is convex, that point is globally Pareto optimal.*

Theorem 5.3 (Local order). *Let $p \in \{2, 3\}$ and let $\theta^\star$ be a nondegenerate local minimizer: $\nabla F(\theta^\star) = 0$ and $\nabla^2 F(\theta^\star) \succeq \mu I$ for some $\mu > 0$. Suppose F and $\mathcal{A}$ are C^{p+1}. Let $s_p(\theta)$ be the unique local stationary branch of $m_\theta^{(p)}$ furnished by Lemma C.2. Assume the subproblem solver returns steps in the local neighborhood of Lemma C.2 satisfying*

$$\left\|\nabla m_t^{(p)}(s_t)\right\| \le \kappa_s \|\nabla F(\theta_t)\|^p . \tag{23}$$

For Algorithm 1 with zero stopping tolerance and fixed $\Delta_0 \in (0, \Delta_{\max}]$, there exists a neighborhood U_{Δ_0} of $\theta^\star$ such that, for $\theta_0 \in U_{\Delta_0}$, the local steps are eventually untruncated and accepted, and

$$\|\theta_{t+1} - \theta^\star\| \le C_p \|\theta_t - \theta^\star\|^p . \tag{24}$$

The guaranteed local order is at least quadratic for $p = 2$ and cubic for $p = 3$.

The forcing condition (23) provides a sufficient inner-accuracy requirement for our local-rate guarantee, in the spirit of inexact Newton methods (Dennis & Moré, 1974; Dembo et al., 1982). Lazy safeguards appear in Appendix C.6; common-generator Bregman objectives are zero-residual special cases (Appendix E).

Proposition 5.4 (Leading local error and canonical cancellation). *Let $F, \mathcal{A} \in C^4$ near $\theta^\star$, with $\nabla F(\theta^\star) = 0$ and $H_\star = \nabla^2 F(\theta^\star) \succ 0$. Set $\mathcal{E} = F - \mathcal{A}$. For order-two calibrated steps in the local neighborhood of Theorem 5.3, impose the stronger forcing condition $\left\|\nabla m_\theta^{(2)}(s)\right\| = O(\|\nabla F(\theta)\|^3)$. Once untruncated and accepted, with $e = \theta - \theta^\star$,*

$$e_+ := \theta + s - \theta^\star = \tfrac{1}{2} H_\star^{-1} \nabla^3 \mathcal{E}(\theta^\star)[e, e] + O(\|e\|^3). \tag{25}$$

If every positive-weight target equals $Q_{\theta^\star}$, canonical α_f gives $\nabla^3 \mathcal{E}(\theta^\star) = 0$ in either orientation, provided the divergence integrals can be differentiated through order four. The local rate is then at least cubic with $p = 2$.

Here $\nabla^3 \mathcal{E}[e, e]$ is the vector whose inner product with v is $\nabla^3 \mathcal{E}[v, e, e]$. For a quadratic Taylor model the corresponding tensor is $\nabla^3 F$; the nonlinear predictor subtracts $\nabla^3 \mathcal{A}$ from this coefficient. Common matching is a sufficient cancellation condition, not a claim of cubic convergence for heterogeneous targets. The stronger inner accuracy uses only model gradients and Hessians. Appendix C.5 proves the expansion; E10 tests both cancellation and its departure under target heterogeneity.

6 NUMERICAL EXPERIMENTS

We test model fidelity (E1–E5, E7), the geometry-dependent local error (E10), and total cost (E9), followed by learned targets (E8) and repeated preferences (E6). Table 1 links controlled tests to claims; Appendix F gives all protocols. Forward/reverse mean $\sigma = -/+$; JS uses the unhalved normalization.

6.1 WHY COMPRESSION NEEDS CALIBRATION

Geometry and step accuracy. E1 recovers residual orders $(4, 3, 2)$ near matched distributions, supporting the canonical local geometry. E2 verifies 15,680 aggregation identities with maximum scaled error 3.77×10^{-15}. E3 recovers value/directional-derivative orders $(p + 1, p)$ even with median maximum absolute log-likelihood ratio 3.94. Thus the predictor is exactly compressible, and calibration controls parameter-step errors beyond nearby targets. E5 verifies the local rates under the stated inner-solve conditions (Table 1).

Why residual derivatives matter. In E4, each calibrated method meets the true-gradient tolerance in 1,800/1,800 runs, versus 180/1,800 for the uncalibrated predictor. In E7, the corresponding counts are 480/480 and 80/480. Direct methods also succeed throughout. These comparisons show why a compressible approximation needs calibration to serve the specified divergence.

What the nonlinear remainder contributes (E10). E10 compares canonical CAL-p2, shifts $\alpha_f \pm 1$, and Taylor-p2, which share the true value, gradient, and Hessian at each anchor. Five seeds, forward JS and KL+quartic, and four heterogeneity levels $\eta \in \{0, .01, .05, .2\}$ yield 1,920 accepted, audited local maps. All methods use the same cubic inner forcing and 120-digit arithmetic; third derivatives are evaluated only offline. At common matching ($\eta = 0$), canonical CAL-p2 has median fitted order 3.000 for both generators, versus 2.000 for all controls. Subtracting the predicted quadratic error in (25) leaves a cubic remainder for every method, directly testing the coefficient formula.

For $\eta > 0$, fitted orders return to approximately two, but canonical CAL-p2 has smaller one-step errors than Taylor-p2 in all 360 paired probes. Thus equal Hessians do not imply equal local steps: the retained nonlinear geometry improves the local error in these cases without third-order calibration. This is a controlled local-accuracy result, not a timing result or a universal ordering of α choices (Appendix F.10). E7 provides complementary finite-step model-error tests; E9 measures whether compression pays off in runtime.

Table 1: Controlled evidence. Local rates require the stated neighborhood and inner-accuracy conditions; the E10 cubic result is at common matching.

Claim	Observation	Experiment
Exact aggregation	Max. scaled identity error 3.77×10^{-15}	E2
Step-based accuracy	Value/directional-derivative exponents $(p + 1, p)$	E3
Calibration fidelity	Higher true-gradient success rates than uncalibrated models	E4, E7
Canonical cancellation	CAL-p2 order 3.000; Taylor-p2 order 2.000	E10
Quadratic local iteration	Mean fitted orders 1.993–2.014	E5, $p = 2$
Cubic local iteration	Mean fitted orders 2.994–2.996	E5, $p = 3$

6.2 WHEN COMPRESSION REDUCES TOTAL COST (E9)

Design and fair accounting. E9 perturbs three fitted latent targets to form nested banks of $m \in \{3, 10, 30, 100, 300\}$ Gaussian mixtures, using five seeds, five positive preferences, and forward JS or KL. The model is a two-component Gaussian mixture with five free parameters. We compare CAL-p1 and CAL-p2 with direct L-BFGS and a quadratic true-objective trust-region solver (Direct-TR). Methods share target caches, vectorized evaluations, initial points, and stationarity thresholds. For KL, the weighted target mixture aggregates the true objective exactly up to a constant; direct baselines use this same reduction. No artificial delays are introduced.

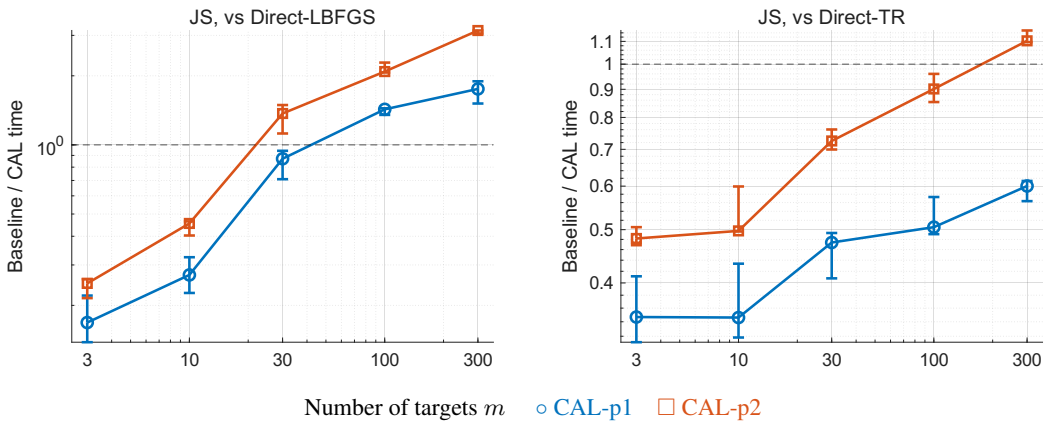

Figure 1: E9: measured JS acceleration relative to direct L-BFGS (left) and Direct-TR (right). Values above one favor CAL; points and bars are seed medians and IQRs of objective-comparable successful pairs. These are the JS panels of the complete comparison in Figure 34; Table 2 reports CAL-p2 pair coverage.

Total solve time includes scalarization setup, calibration, inner optimization, and true-objective acceptance; common data preparation and independent audits are excluded. Three unprofiled timing repeats are summarized within each case. Speedups pair successful endpoints whose scalarized objectives differ by at most 2×10^{-6}, then summarize preferences within seeds and report the median and interquartile range (IQR) across five seeds. All 1,000 primary runs attain audited stationarity; differing local solutions explain incomplete timing-pair coverage.

Lower total cost as the target bank grows. CAL-p2 crosses over against L-BFGS at $m = 30$. At $m = 300$, its median paired speedup is $3.15\times$ over L-BFGS on 23/25 objective-comparable stationary pairs and $1.10\times$ over Direct-TR on 24/25 (Figure 1, Table 2). The excluded pairs reach different local solutions; all attempts remain in the audits and reference-gap plots. At $m = 300$, CAL-p2 uses median counts of five true value/gradient/Hessian evaluations and four additional true-value evaluations, versus six and five for Direct-TR. Repeated inner queries are instead handled by the aggregated predictor.

Table 2: E9: CAL-p2 paired speedups on JS (eligible pairs out of 25). All runs pass the stationarity audit; values below one favor the baseline.

m	vs L-BFGS	vs Direct-TR
3	0.25 (22)	0.48 (24)
10	0.45 (22)	0.50 (21)
30	1.37 (25)	0.72 (24)
100	2.09 (24)	0.90 (24)
300	3.15 (23)	1.10 (24)

Why the gain appears. After aggregation, JS predictor cost is essentially independent of m, while the implemented true JS oracle processes all targets. At $m = 300$, measured true-to-predictor cost ratios are 188.8 for values, 68.1 for value/gradient pairs, and 8.79 for value/gradient/Hessian evaluations (Figure 2). This separation offsets calibration and inner-solver overhead. An implementation ablation keeps the optimization rules fixed and verifies agreement of acceptance decisions and endpoints while requesting only the needed derivative order, confirming the cost of unnecessary template Hessians (Appendix F.9).

Direct-TR supplies the matched true-derivative-order comparison; the larger L-BFGS speedup also reflects different use of curvature. CAL-p2 is slower than Direct-TR through $m = 100$, where nonlinear inner solves and calibration outweigh predictor savings. At $m = 300$, the JS speedup IQRs are $[3.06, 3.15]$ against L-BFGS and $[1.10, 1.15]$ against Direct-TR. Target count alone is insufficient: E4's direct solvers remain faster at $m = 128$ (Table 6). These results establish a cost crossover at comparable attained objectives, not uniformly better minima or faster exact-KL optimization.

At $m = 300$, CAL-p2 also makes median counts of 56 predictor-value calls and 14 predictor-Hessian calls, all charged to the reported solve time. The aggregate supports this repeated nonlinear exploration while five true value/gradient/Hessian calls provide the calibration information.

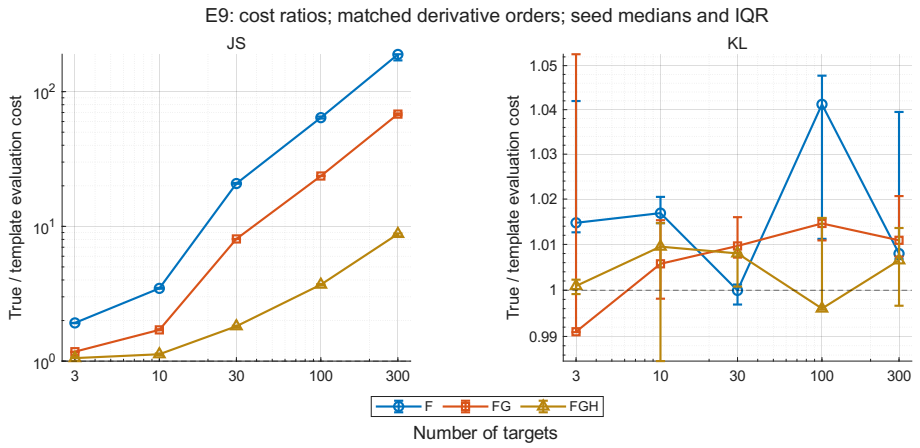

Figure 2: E9: true-objective cost divided by aggregated-predictor cost, at matched derivative orders. F, FG, and FGH include values, values/gradients, and values/gradients/Hessians. The JS cost separation grows with the target count; KL stays near one because both implementations aggregate exactly. Points and bars are seed medians and IQRs across five seeds.

6.3 Learned targets and decoded trade-offs (E8)

E8 uses MNIST digits 0, 1, and 2 with a trained autoencoder whose two-dimensional encoder and decoder are then frozen. Classwise Gaussian mixtures fitted to training encodings define R_0, R_1, R_2; three targets mix them with weights $(.7, .2, .1)$, $(.1, .7, .2)$, and $(.2, .1, .7)$. A five-parameter, two-component latent mixture Q_θ matches these targets under forward JS or KL, and its samples pass through the fixed decoder. We distinguish optimization fidelity from the observable consequences of changing preferences.

Optimization fidelity. Success requires the refined-grid true-gradient norm to be below 5×10^{-6} and passing objective, gradient, and mass checks. Across five seeds and ten preferences, all 200 calibrated runs and all 200 direct-baseline runs pass. The uncalibrated predictor passes 0/50 JS cases, despite plausible decoded images, and 50/50 exact-KL controls. Maximum objective and gradient discrepancies under grid refinement are below 1.86×10^{-8} and 1.14×10^{-7}, respectively.

Preference response. Changing λ changes which learned target the stationary latent distribution prioritizes, as shown by the objective vectors in Figure 25. Preferences weight divergence objectives; they do not prescribe generated class proportions directly.

Decoded consequences. Figure 3 shows the resulting changes in decoded digits. Independent class proportions and image-space MMD^2 describe their distributional consequences (Figure 27). These image metrics do not enter the optimized objective or certify stationarity. E8 is a controlled low-dimensional generative validation; it does not establish better image quality or faster three-target optimization.

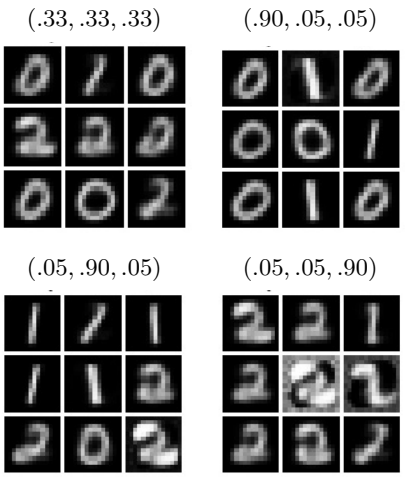

Figure 3: E8: saved decoded mean images for CAL-p2, JS, seed 1, with a fixed decoder. Each preference λ has nine latent draws, without reranking. All-method samples and held-out image diagnostics: Appendix F.8.

Each target contains six Gaussian components, while the model has only two with a fixed shared covariance. The barycenter need not belong to this family, so aggregation supplies an efficient predictor for a restricted fit rather than an already representable solution.

6.4 Complementary acceleration across preferences (E6)

Standard implicit differentiation provides a complementary initializer for repeated preferences, independently of α-aggregation (Appendix D). E6 compares Cold, Warm, and Tangent initializations with the same CAL-p2 corrector on categorical JS and KL+quartic problems, both orientations, and ten seeds. All 20,880 point solves meet the true-gradient tolerance 10^{-8}, with maximum reference objective discrepancy 9.65×10^{-15}. At 49 preferences, Tangent has median paired speedups 1.26–$1.59\times$ over Warm and 2.03–$2.10\times$ over Cold, including aggregation, tangent solves, and correction. Reusing regular stationary branches thus reduces the work of the calibrated corrector (Table 3).

Table 3: E6: Tangent costs at 49 preferences; paired seed medians. Timings include aggregation, initialization, and correction. F/R: forward/reverse; KL4: KL+quartic.

	Time speedup		True-jet ratio
Case	vs Cold	vs Warm	vs Cold
JS, F	2.099	1.590	0.581
JS, R	2.044	1.474	0.581
KL4, F	2.087	1.376	0.562
KL4, R	2.026	1.263	0.516

Along a regular branch, Tangent has $O(\|\delta\lambda\|^2)$ initialization error, while Warm generally has first-order error. Independent E6 probes recover these orders. The same corrector and stopping test are used throughout (Appendix F.6).

7 Discussion and conclusion

CAL-α-MODM retains a compressible nonlinear remainder beyond a matched Taylor jet. Calibration controls parameter-step error, and true-objective tests certify progress. The stationarity and order-p local guarantees hold under the stated assumptions without target proximity. Canonical cancellation further improves the order-two model at common matching under the stronger inner-accuracy condition.

Local accuracy at the same calibration order. In E10 at $\eta = 0.05$, canonical CAL-p2 and Taylor-p2 both have fitted order two, yet the seed-median Taylor-to-canonical ratios of the leading directional error norm are 2.37×10^3 for JS and 31.8 for KL4 (Table 9). These quantify a smaller local error coefficient along the sampled directions, with identical center values, gradients, and Hessians. The retained nonlinear geometry improves local accuracy without third-order calibration; runtime is measured separately in E9.

When to use the structure. E9 shows when cheaper predictor queries offset calibration overhead; matched-order oracle costs help identify this regime. E6 adds initialization savings for repeated preferences. If the true objective already aggregates exactly, as in the KL control, residual calibration is unnecessary.

Why latent matching is meaningful. Let K be a fixed decoding map or a common stochastic decoder, and write P_i^Z, Q_θ^Z for latent distributions, with KP denoting the decoded law of P. Data processing for f-divergences (Csiszár, 1967) gives, in either orientation,

$$D_f^\sigma(KP_i^Z, KQ_\theta^Z) \le D_f^\sigma(P_i^Z, Q_\theta^Z).$$

Thus the exact latent scalarization upper-bounds the weighted decoded divergences; this bound does not transfer stationarity to the decoded objective. In E8, the decoded target laws come from fitted latent mixtures, not empirical MNIST distributions. Independent image diagnostics assess agreement with held-out image targets.

The demonstrated scaling is in target count m, not parameter dimension d or latent dimension D: E8/E9 use $d = 5, D = 2$. Dense factorization and tensor-grid integration can cost $O(d^3)$ and $O(k^D)$, respectively, with k nodes per coordinate. Matrix-free variants and stochastic oracles remain untested extensions (Appendix C.7); the guarantees concern stationarity, not full Pareto coverage.

Reproducibility Statement

Appendices B–E contain assumptions and proofs; Appendix F details all ten experiments, seeds, tolerances, and timing. Retained histories, paired timings, and audits support verification; E10 uses 120-digit arithmetic, independent derivative checks, and a fixed fitting window. Supplementary material includes anonymized MATLAB code and a reproduction README.

AI Use Statement

Generative AI was used to aid or polish writing and to assist in drafting portions of the manuscript, including improvements to language, readability, clarity, organization, and LaTeX formatting. It also assisted with generating and debugging MATLAB code for the numerical experiments. All AI-assisted text and code were carefully reviewed, revised, and verified by the authors, who take full responsibility for the entire submission.

References

David Acuna, Guojun Zhang, Marc T. Law, and Sanja Fidler. f-domain adversarial learning: Theory and algorithms. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 66–75, 2021. URL https://proceedings.mlr.press/v139/acuna21a.html.

Rishabh Agarwal, Nino Vieillard, Yongchao Zhou, Piotr Stanczyk, Sabela Ramos Garea, Matthieu Geist, and Olivier Bachem. On-policy distillation of language models: Learning from self-generated mistakes. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=3zKtaqxLhW.

Natalia M. Alexandrov, John E. Dennis, Jr., Robert M. Lewis, and Virginia Torczon. A trust-region framework for managing the use of approximation models in optimization. *Structural Optimization*, 15(1):16–23, 1998. doi: 10.1007/BF01197433.

Shun-ichi Amari. *Information Geometry and Its Applications*. Springer, 2016. doi: 10.1007/978-4-431-55978-8.

Arindam Banerjee, Srujana Merugu, Inderjit S. Dhillon, and Joydeep Ghosh. Clustering with Bregman divergences. *Journal of Machine Learning Research*, 6:1705–1749, 2005.

Stefania Bellavia, Nataša Krejić, Benedetta Morini, and Simone Rebegoldi. A stochastic first-order trust-region method with inexact restoration for finite-sum minimization. *Computational Optimization and Applications*, 84:53–84, 2023. doi: 10.1007/s10589-022-00430-7.

Mathieu Blondel, Quentin Berthet, Marco Cuturi, Roy Frostig, Stephan Hoyer, Felipe Llinares-López, Fabian Pedregosa, and Jean-Philippe Vert. Efficient and modular implicit differentiation. In *Advances in Neural Information Processing Systems*, volume 35, pp. 5230–5242. Curran Associates, Inc., 2022. doi: 10.52202/068431-0378. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/228b9279ecf9bbafe582406850c57115-Paper-Conference.pdf.

Andrew J. Booker, John E. Dennis, Jr., Paul D. Frank, David B. Serafini, Virginia Torczon, and Michael W. Trosset. A rigorous framework for optimization of expensive functions by surrogates. *Structural Optimization*, 17(1):1–13, 1999. doi: 10.1007/BF01197708.

Rob Brekelmans and Frank Nielsen. Variational representations of annealing paths: Bregman information under monotonic embedding. *Information Geometry*, 7:193–228, 2024. doi: 10.1007/s41884-023-00129-6.

Coralia Cartis, Nicholas I. M. Gould, and Philippe L. Toint. On the complexity of steepest descent, Newton's and regularized Newton's methods for nonconvex unconstrained optimization problems. *SIAM Journal on Optimization*, 20(6):2833–2852, 2010. doi: 10.1137/090774100.

Coralia Cartis, Nicholas I. M. Gould, and Philippe L. Toint. *Evaluation Complexity of Algorithms for Nonconvex Optimization: Theory, Computation and Perspectives*, volume 30 of *MOS-SIAM Series on Optimization*. SIAM, 2022. doi: 10.1137/1.9781611976991.

Lisha Chen, AFM Saif, Yanning Shen, and Tianyi Chen. FERERO: A flexible framework for preference-guided multi-objective learning. In *Advances in Neural Information Processing Systems*, volume 37, pp. 18758–18805. Curran Associates, Inc., 2024. doi: 10.52202/079017-0595. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/21a33eba893ca3890e395651b38810df-Paper-Conference.pdf.

Andrew R. Conn, Nicholas I. M. Gould, and Philippe L. Toint. *Trust-Region Methods*. SIAM, 2000. doi: 10.1137/1.9780898719857.

Imre Csiszár. Information-type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica*, 2:299–318, 1967.

Ron S. Dembo, Stanley C. Eisenstat, and Trond Steihaug. Inexact Newton methods. *SIAM Journal on Numerical Analysis*, 19(2):400–408, 1982. doi: 10.1137/0719025.

John E. Dennis, Jr. and Jorge J. Moré. A characterization of superlinear convergence and its application to quasi-Newton methods. *Mathematics of Computation*, 28(126):549–560, 1974. doi: 10.1090/S0025-5718-1974-0343581-1.

Nikita Doikov and Yurii Nesterov. Local convergence of tensor methods. *Mathematical Programming*, 193:315–336, 2022. doi: 10.1007/s10107-020-01606-x.

Dongyoung Go, Tomasz Korbak, Germàn Kruszewski, Jos Rozen, Nahyeon Ryu, and Marc Dymetman. Aligning language models with preferences through f-divergence minimization. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 11546–11583, 2023. URL https://proceedings.mlr.press/v202/go23a.html.

Yuxian Gu, Li Dong, Furu Wei, and Minlie Huang. MiniLLM: Knowledge distillation of large language models. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=5h0qf7IBZZ.

Jiaqi Han, Mingjian Jiang, Yuxuan Song, Stefano Ermon, and Minkai Xu. f-PO: Generalizing preference optimization with f-divergence minimization. In *Proceedings of the 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 1144–1152, 2025. URL https://proceedings.mlr.press/v258/han25a.html.

Shao-Lun Huang, Anuran Makur, Gregory W. Wornell, and Lizhong Zheng. Universal features for high-dimensional learning and inference. *Foundations and Trends in Communications and Information Theory*, 21(1–2):1–299, 2024. doi: 10.1561/0100000107.

Xi Lin, Zhiyuan Yang, Xiaoyuan Zhang, and Qingfu Zhang. Pareto set learning for expensive multi-objective optimization. In *Advances in Neural Information Processing Systems*, volume 35, pp. 19231–19247. Curran Associates, Inc., 2022. doi: 10.52202/068431-1398. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/7a583691ccfcf8945ab714b677ccbf0b-Paper-Conference.pdf.

Xi Lin, Xiaoyuan Zhang, Zhiyuan Yang, Fei Liu, Zhenkun Wang, and Qingfu Zhang. Smooth Tchebycheff scalarization for multi-objective optimization. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 30479–30509. PMLR, 2024. URL https://proceedings.mlr.press/v235/lin24y.html.

Bo Liu, Xingchao Liu, Xiaojie Jin, Peter Stone, and Qiang Liu. Conflict-averse gradient descent for multi-task learning. In *Advances in Neural Information Processing Systems*, volume 34, pp. 18878–18890, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/9d27fdf2477ffbff837d73ef7ae23db9-Abstract.html.

Bo Liu, Yihao Feng, Peter Stone, and Qiang Liu. FAMO: Fast adaptive multitask optimization. In *Advances in Neural Information Processing Systems*, volume 36, pp. 57226–57243. Curran Associates, Inc., 2023. doi: 10.52202/075280-2500. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/b2fe1ee8d936ac08dd26f2ff58986c8f-Paper-Conference.pdf.

Vaden Masrani, Rob Brekelmans, Thang Bui, Frank Nielsen, Aram Galstyan, Greg Ver Steeg, and Frank Wood. q-paths: Generalizing the geometric annealing path using power means. In *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, volume 161 of *Proceedings of Machine Learning Research*, pp. 1938–1947, 2021. URL https://proceedings.mlr.press/v161/masrani21a.html.

Aviv Navon, Aviv Shamsian, Gal Chechik, and Ethan Fetaya. Learning the Pareto front with hypernetworks. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=NjF772F4ZZR.

Aviv Navon, Aviv Shamsian, Idan Achituve, Haggai Maron, Kenji Kawaguchi, Gal Chechik, and Ethan Fetaya. Multi-task learning as a bargaining game. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 16428–16446. PMLR, 2022. URL https://proceedings.mlr.press/v162/navon22a.html.

Frank Nielsen. Generalizing the alpha-divergences and the oriented Kullback–Leibler divergences with quasi-arithmetic means. *Algorithms*, 15(11):435, 2022. doi: 10.3390/a15110435.

Frank Nielsen. Fast proxy centers for the Jeffreys centroid: The Jeffreys–Fisher–Rao center and the Gauss–Bregman inductive center. *Entropy*, 26(12):1008, 2024. doi: 10.3390/e26121008.

Frank Nielsen and Richard Nock. Sided and symmetrized Bregman centroids. *IEEE Transactions on Information Theory*, 55(6):2882–2904, 2009. doi: 10.1109/TIT.2009.2018176.

Nicola Novello and Andrea M. Tonello. f-divergence based classification: Beyond the use of cross-entropy. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 38448–38473, 2024. URL https://proceedings.mlr.press/v235/novello24a.html.

Vincent Roulet, Tianlin Liu, Nino Vieillard, Michael Eli Sander, and Mathieu Blondel. Loss functions and operators generated by f-divergences. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 52110–52138, 2025. URL https://proceedings.mlr.press/v267/roulet25a.html.

Cheng Tang, Zhishuai Liu, and Pan Xu. Robust offline reinforcement learning with linearly structured f-divergence regularization. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 58842–58882, 2025. URL https://proceedings.mlr.press/v267/tang25f.html.

Dier Tang and Guangyue Han. Universal feature selection with noisy observations and weak symmetry conditions, 2026. URL https://arxiv.org/abs/2605.09396. arXiv preprint arXiv:2605.09396.

Yuqiao Wen, Zichao Li, Wenyu Du, and Lili Mou. f-divergence minimization for sequence-level knowledge distillation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 10817–10834. Association for Computational Linguistics, 2023. doi: 10.18653/v1/2023.acl-long.605. URL https://aclanthology.org/2023.acl-long.605/.

Xiangxiang Xu and Lizhong Zheng. Neural feature learning in function space. *Journal of Machine Learning Research*, 25(142):1–76, 2024. URL https://jmlr.org/papers/v25/23-1202.html.

Xiaoyuan Zhang, Peijie Li, Yingying Yu, Yichi Zhang, Han Zhao, and Qingfu Zhang. Multiobjective distribution matching. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 75413–75426, 2025. URL https://proceedings.mlr.press/v267/zhang25as.html.

APPENDIX CONTENTS

A Extended Related Work **14**

A.1 Divergence objectives in learning 14

A.2 Multiobjective learning and repeated preferences 14

A.3 Sampled objectives and model accuracy 14

B Proofs for f-divergence geometry **14**

B.1 Proof of Theorem 3.1 14

B.2 Size of the calibration tensors 15

B.3 Proof of Theorem 3.2 15

C Calibrated models and trust-region analysis **16**

C.1 Proof of Theorem 3.4 16

C.2 Precise trust-region assumptions 16

C.3 Proof of Theorem 5.1 16

C.4 Local stationary branch and proof of Theorem 5.3 17

C.5 Leading local error and canonical cancellation 18

C.6 Safeguarded lazy calibration 18

C.7 Implementing model decrease and interpreting scalability 19

D Preference continuation **20**

E Bregman transfer results **20**

F Experimental protocols and additional results **21**

F.1 Local normal-form scaling 21

F.2 Exact aggregation and amortized cost 22

F.3 Calibrated-model order diagnostics 22

F.4 Full optimization: robustness and cost 23

F.5 Local high-order convergence and condition ablations 26

F.6 Preference continuation and sensitivity prediction 29

F.7 Template choice after second-order calibration 32

F.8 E8: learned latent distribution matching and generation 37

F.9 E9: many-target cost crossover and implementation controls 41

F.10 E10: canonical geometry and the leading local error 46

A Extended Related Work

A.1 Divergence objectives in learning

Loss design and preference alignment. Variational f-divergence representations yield classification objectives (Novello & Tonello, 2024), while divergence-generated convex losses and prediction operators extend familiar cross-entropy constructions (Roulet et al., 2025). Go et al. (2023) study alignment–diversity trade-offs through target-distribution approximation, and f-PO generalizes preference optimization using divergence objectives (Han et al., 2025).

Distillation, transfer, and robustness. Sequence-level distillation uses f-divergence objectives with tractable token-level approximations (Wen et al., 2023). MiniLLM uses reverse KL (Gu et al., 2024), while generalized knowledge distillation combines student-generated sequences with flexible divergence losses (Agarwal et al., 2024).

Divergences also define discrepancies in adversarial domain adaptation (Acuna et al., 2021) and regularize transition models in robust offline reinforcement learning (Tang et al., 2025). These adversarial and robust formulations extend beyond fixed-target scalarization.

Feature learning and statistical geometry. Universal feature analysis connects local inference with low-dimensional representations of statistical dependence (Huang et al., 2024). Xu & Zheng (2024) approximate dependence components in function space, where the squared $L^2(P_X P_Y)$ norm of $P_{XY}/(P_X P_Y) - 1$ equals the χ^2 divergence between the joint distribution and the product of its marginals. Tang & Han (2026) study universal feature selection in χ^2-divergence neighborhoods and establish weaker conditions sufficient for feature universality.

A.2 Multiobjective learning and repeated preferences

Adaptive task balancing. CAGrad regulates average-loss updates using the worst local task improvement (Liu et al., 2021). Nash-MTL derives gradient aggregation from bargaining (Navon et al., 2022), and FAMO adapts task weights without storing all task gradients (Liu et al., 2023). These methods adapt task influence; our preference weights are inputs.

Preference specification. Smooth Tchebycheff scalarization provides a differentiable nonlinear alternative to weighted sums (Lin et al., 2024). FERERO supports cone-defined relative preferences and absolute constraints on objective values (Chen et al., 2024). These methods broaden preference specification; our aggregation identity applies to fixed linear scalarizations.

Repeated preference queries. Hypernetworks learn preference-to-parameter maps (Navon et al., 2021), and Pareto set learning models solution sets for expensive multiobjective Bayesian optimization (Lin et al., 2022). Implicit differentiation obtains solution sensitivities from optimality conditions (Blondel et al., 2022). E6 tests sensitivity initialization of a common corrector along regular stationary branches, without learning a solution map.

A.3 Sampled objectives and model accuracy

Bellavia et al. (2023) combine stochastic trust-region models with inexact restoration for finite-sum minimization. Sampled divergences require derivative and acceptance-test error controls beyond our deterministic bounds.

B Proofs for f-divergence geometry

B.1 Proof of Theorem 3.1

The normalization gives $f(1) = f'(1) = 0$. From the derivatives of f_α and (6),

$$a_f f''_{\alpha_f}(1) = f''(1), \qquad a_f f'''_{\alpha_f}(1) = -a_f \frac{3+\alpha_f}{2} = f'''(1). \tag{26}$$

Therefore $h_f = f - a_f f_{\alpha_f}$ has a zero of order at least four at $u = 1$. Taylor's theorem gives $h_f(u) = h_f^{(4)}(\xi_u)(u-1)^4/24$ for some ξ_u between 1 and u. Substitute $u = p/q$, multiply by q, integrate, and apply the triangle inequality to obtain (7). □

B.2 Size of the calibration tensors

Proposition B.1. *Under the normalization and smoothness assumptions of Theorem 3.1, let* $R^-(\theta) = D_f(P\|Q_\theta) - a_f D_{\alpha_f}(P\|Q_\theta)$. *Suppose differentiation under the integral is valid,* $p/q_\theta \in [1-\delta, 1+\delta]$, *and the score and log-density Hessian are uniformly bounded. Then*

$$\left\|\nabla R^-(\theta)\right\| \le C_1 \int q_\theta \left|p/q_\theta - 1\right|^3 \,\mathrm{d}\mu, \tag{27}$$

$$\left\|\nabla^2 R^-(\theta)\right\|_{\mathrm{op}} \le C_2 \int q_\theta \left|p/q_\theta - 1\right|^2 \,\mathrm{d}\mu. \tag{28}$$

Analogous bounds hold for $R^+(\theta) = D_f(Q_\theta\|P) - a_f D_{\alpha_f}(Q_\theta\|P)$ *when* $q_\theta/p \in [1-\delta, 1+\delta]$. *If* $\nabla^3 \log q_\theta$ *is also uniformly bounded and third-order differentiation under the integral is valid, then*

$$\left\|\nabla^3 R^\sigma(\theta)\right\|_{\mathrm{op}} \le C_3 \int q_\theta \left|u_\sigma - 1\right| \,\mathrm{d}\mu, \qquad u_- = p/q_\theta, \quad u_+ = q_\theta/p. \tag{29}$$

The constants depend on the density-derivative bounds, δ, *and* $M_{f,\delta}$, *not on the parameter step.*

Proof. Write $h = h_f$, $u = p/q_\theta$, $z = \nabla \log q_\theta$, and $J = \nabla^2 \log q_\theta$. For the forward orientation,

$$\nabla R^-(\theta) = \int q_\theta \big(h(u) - uh'(u)\big) z \,\mathrm{d}\mu, \tag{30}$$

$$\nabla^2 R^-(\theta) = \int q_\theta \left[\big(h - uh' + u^2 h''\big) zz^\top + \big(h - uh'\big) J\right] \mathrm{d}\mu. \tag{31}$$

Because $h^{(j)}(1) = 0$ for $j = 0, 1, 2, 3$, we have $h(u) = O(|u-1|^4)$, $h'(u) = O(|u-1|^3)$, and $h''(u) = O(|u-1|^2)$. The claimed bounds follow. For the reverse orientation, with $u = q_\theta/p$,

$$\nabla R^+(\theta) = \int q_\theta h'(u) z \,\mathrm{d}\mu, \tag{32}$$

$$\nabla^2 R^+(\theta) = \int q_\theta [(h'(u) + uh''(u)) zz^\top + h'(u) J] \,\mathrm{d}\mu, \tag{33}$$

and the same Taylor bounds apply. For the third derivative, put $L = \nabla^3 \log q_\theta$ and let $\mathrm{Sym}_3(z \otimes J)$ denote the sum over its three distinct placements. Differentiating once more gives, in the forward orientation,

$$\nabla^3 R^- = \int q_\theta \left[(h - uh' - u^3 h''') z^{\otimes 3} + (h - uh' + u^2 h'') \,\mathrm{Sym}_3(z \otimes J) + (h - uh') L\right] \mathrm{d}\mu.$$

For the reverse orientation the three coefficients are respectively $h' + 3uh'' + u^2 h'''$, $h' + uh''$, and h'. Using $h'''(u) = O(|u-1|)$, the preceding Taylor bounds, and the bounded tensors z, J, L proves (29). Weighted sums give the corresponding scalarized residual bound. □

B.3 Proof of Theorem 3.2

Let $K_\alpha = 4/(1-\alpha^2)$ and $H^-(x) = \sum_i \lambda_i p_i(x)^{r_\alpha}$. Then

$$\sum_i \lambda_i D_\alpha(P_i\|Q) = K_\alpha \left(1 - \int H^-(x) q(x)^{s_\alpha} \,\mathrm{d}\mu(x)\right). \tag{34}$$

By (8), $H^- = (Z^-_{\lambda,\alpha})^{r_\alpha} (c^-_{\lambda,\alpha})^{r_\alpha}$, and hence

$$\sum_i \lambda_i D_\alpha(P_i\|Q) = K_\alpha(1 - \omega^-_{\lambda,\alpha}) + \omega^-_{\lambda,\alpha} D_\alpha(c^-_{\lambda,\alpha}\|Q). \tag{35}$$

The reverse identity follows by taking $H^+(x) = \sum_i \lambda_i p_i(x)^{s_\alpha}$ and exchanging the arguments. At the KL endpoints, write $\bar{p} = \sum_i \lambda_i p_i$ and $g = Z_g^{-1} \prod_i p_i^{\lambda_i}$, with $Z_g = \int \prod_i p_i^{\lambda_i} \,\mathrm{d}\mu$. Direct expansion gives

$$\sum_i \lambda_i \,\mathrm{KL}(P_i\|Q) = \sum_i \lambda_i \,\mathrm{KL}(P_i\|\bar{P}) + \mathrm{KL}(\bar{P}\|Q), \tag{36}$$

$$\sum_i \lambda_i \,\mathrm{KL}(Q\|P_i) = \mathrm{KL}(Q\|G) - \log Z_g. \tag{37}$$

Here $\bar{P}$ and G have densities $\bar{p}$ and g; finiteness is assumed for the displayed terms. These are the $\alpha = -1$ identities in the two orientations; $\alpha = 1$ exchanges them. Their scale factors are one, as also obtained from the stated finite continuous limits. □

C CALIBRATED MODELS AND TRUST-REGION ANALYSIS

C.1 PROOF OF THEOREM 3.4

Because $m_t^{(p)} = \widehat{\mathcal{A}}_t + T_p\widehat{\mathcal{E}}_t$ and $\widehat{F}_t = \widehat{\mathcal{A}}_t + \widehat{\mathcal{E}}_t$, derivative matching follows from the Taylor polynomial. Taylor's theorem with integral remainder gives

$$\left|\widehat{\mathcal{E}}_t(v) - T_p\widehat{\mathcal{E}}_t(v)\right| \le \frac{L_{p+1}}{(p+1)!}\|v\|^{p+1}. \tag{38}$$

Differentiation commutes with Taylor truncation: for $0 \le j \le p$, $\nabla^j T_p\widehat{\mathcal{E}}_t$ is the degree-$(p-j)$ Taylor polynomial of $\nabla^j\widehat{\mathcal{E}}_t$. Applying the integral remainder to this tensor-valued function bounds its error by $L_{p+1}\|v\|^{p+1-j}/(p+1-j)!$. The case $j=1$ gives (17). □

C.2 PRECISE TRUST-REGION ASSUMPTIONS

Assumption C.1 (Trust-region regularity). *Fix λ and σ, set $\Theta = \mathbb{R}^d$, use exact objective and derivative oracles, and assume:*

(A1) The level set $\{\theta : F(\theta) \le F(\theta_0)\}$ is compact, $F \ge F_{\inf}$ there, and ∇F is L_g-Lipschitz, with $L_g > 0$, on a neighborhood containing every trial segment.

(A2) The objective and predictor satisfy Theorem 3.4 on every trial ball, with a uniform residual-derivative bound for $\Delta_t \le \Delta_{\max}$.

(A3) $\sup_{t,\|v\|\le\Delta_t}\left\|\nabla^2 m_t^{(p)}(v)\right\|_{\mathrm{op}} \le \kappa_H$.

(A4) The computed step has $\|s_t\| \le \Delta_t$ and satisfies (19) for fixed $\kappa_c \in (0,1]$.

(A5) The acceptance decision uses (18), with $\mathrm{pred}_t > 0$, and the radius follows (20) for fixed $0 < \eta_1 < \eta_2 < 1$ and $0 < \gamma_1 < 1 < \gamma_2$.

Condition (19) is attainable for the nonlinear model itself. For $g_t \ne 0$, let $\tau_t = \min\{\Delta_t/\|g_t\|, 1/\kappa_H\}$ and $s_t^{\mathrm{C}} = -\tau_t g_t$. The Hessian bound along this segment gives

$$m_t^{(p)}(0) - m_t^{(p)}(s_t^{\mathrm{C}}) \ge \tau_t\|g_t\|^2 - \tfrac{1}{2}\kappa_H\tau_t^2\|g_t\|^2 \ge \tfrac{1}{2}\|g_t\|\min\{\Delta_t, \|g_t\|/\kappa_H\}.$$

Thus any feasible step with at least this model decrease satisfies (19) with $\kappa_c = 1/2$; Appendix C.7 describes a backtracking safeguard when κ_H is not known.

For every $p \ge 1$, Theorem 3.4 implies fully linear bounds

$$\left|F(\theta_t + v) - m_t^{(p)}(v)\right| \le \kappa_f\Delta_t^2, \tag{39}$$

$$\left\|\nabla F(\theta_t + v) - \nabla m_t^{(p)}(v)\right\| \le \kappa_g\Delta_t \tag{40}$$

for $\|v\| \le \Delta_t$, choosing positive constants κ_f, κ_g. For $p = 2$ the sharper orders are $(3,2)$, and for $p = 3$ they are $(4,3)$. An exact stationary endpoint is extended constantly when discussing asymptotic rates. The global result concerns the unconstrained Euclidean iteration; a bounded affine probability chart is covered by the local result only while its boundary safeguard is inactive.

C.3 PROOF OF THEOREM 5.1

Since $m_t^{(p)}(0) = F(\theta_t)$,

$$|\mathrm{ared}_t - \mathrm{pred}_t| \le \kappa_f\Delta_t^2. \tag{41}$$

Define

$$c_\rho := \min\left\{\frac{1}{\kappa_H}, \frac{(1-\eta_2)\kappa_c}{\kappa_f}\right\}. \tag{42}$$

If $g_t \ne 0$ and

$$\Delta_t \le c_\rho\|g_t\|, \tag{43}$$

then (19)–(41) give $|1-\rho_t| \le 1-\eta_2$. Thus a sufficiently small radius relative to the gradient produces a very successful step and cannot be repeatedly contracted.

Fix $0 < \varepsilon \le \varepsilon_0 := \Delta_0/(\gamma_1 c_\rho)$. Before the first ε-stationary point,

$$\Delta_t \ge \underline{\Delta}_\varepsilon := \gamma_1 c_\rho \varepsilon. \tag{44}$$

Indeed, an unsuccessful iteration must have $\Delta_t > c_\rho \|g_t\| > c_\rho \varepsilon$, so its contracted successor exceeds (44); a radius below $c_\rho \varepsilon$ instead gives $\rho_t \ge \eta_2$ and is not contracted. The claim follows by induction from the definition of ε_0.

Every successful iteration before termination therefore decreases the true objective by at least

$$\eta_1 \kappa_c \varepsilon \min\{\gamma_1 c_\rho \varepsilon, \varepsilon/\kappa_H\} = c_s \varepsilon^2. \tag{45}$$

Boundedness below limits the successful iterations to $(F(\theta_0) - F_{\inf})/(c_s \varepsilon^2)$. If N_u and N_v denote unsuccessful and very successful iterations, respectively, the radius recursion and (44) imply

$$N_u |\log \gamma_1| \le N_v \log \gamma_2 + \log \frac{\Delta_{\max}}{\gamma_1 c_\rho \varepsilon}. \tag{46}$$

Since N_v is no larger than the successful count and $|\log \varepsilon| = O(\varepsilon^{-2})$ on $(0, 1)$, this proves (22); larger ε are absorbed into C_0. The constants need not be uniform in m or target heterogeneity.

For the zero-tolerance run, either an exact stationary point is reached or the preceding argument gives $\liminf_t \|g_t\| = 0$. To prove convergence of the entire gradient sequence, fix $\varepsilon > 0$ and set

$$b_\varepsilon = \eta_1 \kappa_c \varepsilon \min\{1, \varepsilon/(\kappa_H \Delta_{\max})\} > 0. \tag{47}$$

At each successful iteration with $\|g_t\| \ge \varepsilon$, Cauchy decrease implies

$$F(\theta_t) - F(\theta_{t+1}) \ge b_\varepsilon \Delta_t \ge b_\varepsilon \|s_t\|. \tag{48}$$

If $\limsup_t \|g_t\| > 0$, choose ε with $2\varepsilon < \limsup_t \|g_t\|$. Since the liminf is zero, there are infinitely many disjoint intervals starting with $\|g_t\| \ge 2\varepsilon$ and ending at its first subsequent value below ε. Within each interval, Lipschitz continuity requires total accepted-step length at least ε/L_g; rejected steps do not move the iterate. Equation (48) therefore gives objective decrease at least $b_\varepsilon \varepsilon/L_g$ per interval, contradicting the lower bound on F. Thus $\|g_t\| \to 0$. □

Proof of Corollary 5.2. Continuity at any accumulation point $\bar\theta$ gives $\sum_i \lambda_i \nabla F_i^\sigma(\bar\theta) = 0$. A direction that strictly decreases every component to first order would have negative inner product with this weighted sum, a contradiction. If the component objectives are convex, the first-order convexity inequality shows that $\bar\theta$ globally minimizes their weighted sum. Because every $\lambda_i > 0$, any point dominating $\bar\theta$ would strictly lower that sum. Hence $\bar\theta$ is globally Pareto optimal. □

C.4 Local stationary branch and proof of Theorem 5.3

For a generic anchor θ, write $m_\theta^{(p)}$ for (13) and define $G_p(\theta, s) = \nabla_s m_\theta^{(p)}(s)$.

Lemma C.2 (Local model stationary branch)**.** *Under the hypotheses of Theorem 5.3, there are open balls around $\theta^\star$ and 0 and a unique continuously differentiable map $s_p(\theta)$ such that $G_p(\theta, s_p(\theta)) = 0$, $s_p(\theta^\star) = 0$, and*

$$\|s_p(\theta)\| \le C_s \|\theta - \theta^\star\|, \qquad \nabla_s^2 m_\theta^{(p)}(s) \succeq \frac{\mu}{2} I \tag{49}$$

throughout a smaller product neighborhood. If s lies in this neighborhood, then $\|s - s_p(\theta)\| \le (2/\mu)\|G_p(\theta, s)\|$.

Proof. Derivative matching gives $G_p(\theta^\star, 0) = \nabla F(\theta^\star) = 0$ and $\nabla_s G_p(\theta^\star, 0) = \nabla^2 F(\theta^\star) \succ 0$. The implicit function theorem gives the unique branch. Its derivative is locally bounded, which yields the first bound in (49); continuity gives the Hessian bound. Strong monotonicity of $G_p(\theta, \cdot)$ then gives the final inequality. □

Proof of Theorem 5.3. Let $e_t = \|\theta_t - \theta^\star\|$ and $s_p^t = s_p(\theta_t)$. Strong regularity gives constants $c_1, c_2 > 0$ such that locally

$$c_1 e_t \le \|\nabla F(\theta_t)\| \le c_2 e_t. \tag{50}$$

Lemma C.2 and (23) give $\|s_t - s_p^t\| = O(e_t^p)$ and hence $\|s_t\| \le c_3 e_t$. At the trial point, (17) and (23) give

$$\|\nabla F(\theta_t + s_t)\| \le \frac{L_{p+1}}{p!} \|s_t\|^p + \kappa_s \|\nabla F(\theta_t)\|^p \le c_4 e_t^p. \tag{51}$$

Equation (50) at the trial point yields the order-p error bound whenever the step is accepted. To verify acceptance, let M bound the local model Hessians. Since $g_t = -\int_0^1 \nabla_s^2 m_t^{(p)}(us_p^t)s_p^t\,\mathrm{d}u$, we have $\|s_p^t\| \geq \|g_t\|/M$. Strong convexity gives $m_t^{(p)}(0) - m_t^{(p)}(s_p^t) \geq \mu\|s_p^t\|^2/4$, while $m_t^{(p)}(s_t) - m_t^{(p)}(s_p^t) = O(e_t^{2p})$. Together with the local upper bounds, this proves $\mathrm{pred}_t = \Theta(e_t^2)$ for nonstationary anchors. The value-model error is $O(e_t^{p+1})$, so $|1-\rho_t| = O(e_t^{p-1}) \to 0$. Fix $\Delta_0 \in (0, \Delta_{\max}]$. Choose the initial neighborhood sufficiently small, depending on Δ_0, so that the returned steps satisfy $\|s_t\| < \Delta_0$, the local error bound is contractive, and $\rho_t \geq \eta_2$. The accepted iterates then remain in this neighborhood, while the radius update never decreases the radius. Thus the local steps remain untruncated and accepted, and (24) applies. □

C.5 Leading local error and canonical cancellation

Proof of Proposition 5.4. Write $x = \theta$, $y = x + s$, $e = x - \theta^\star$, and $r = \nabla m_x^{(2)}(s)$. On a fixed sufficiently small neighborhood, C^4 regularity bounds the fourth derivatives. Taylor expansion of the residual gradient gives

$$\nabla F(y) = r + \tfrac{1}{2}\nabla^3\mathcal{E}(x)[s,s] + O(\|s\|^3). \tag{52}$$

The local stationary model branch from Lemma C.2 is $O(\|e\|)$, and local strong convexity bounds the distance of an inexact step from that branch by $C\|r\|$. Since $\|\nabla F(x)\| = O(\|e\|)$, cubic forcing gives $r = O(\|e\|^3)$ and $s = O(\|e\|)$. Equation (52), together with local invertibility of ∇F, then yields $y - \theta^\star = O(\|e\|^2)$. Consequently $s = -e + O(\|e\|^2)$ and

$$\nabla^3\mathcal{E}(x)[s,s] = \nabla^3\mathcal{E}(\theta^\star)[e,e] + O(\|e\|^3).$$

Expanding $\nabla F(y) = H_\star(y-\theta^\star) + O(\|y-\theta^\star\|^2)$ in (52) proves (25). These same local steps satisfy the $p=2$ forcing condition and are eventually accepted as in the proof of Theorem 5.3.

For canonical α_f, each component of $\mathcal{E}$ is an integral with residual generator $h_f = f - a_f f_{\alpha_f}$. At $P_i = Q_{\theta^\star}$, its density ratio equals one. Every term in the first three parameter derivatives of either $q_\theta h_f(p_i/q_\theta)$ or $p_i h_f(q_\theta/p_i)$ contains $h_f^{(j)}(1)$ for some $j \leq 3$, which vanishes by Theorem 3.1. Differentiating under the integral and summing over positive-weight targets gives $\nabla^3\mathcal{E}(\theta^\star) = 0$ and hence $\|e_+\| = O(\|e\|^3)$. For a finite positive categorical model, the interchange is automatic for a C^4 parameterization near an interior point. No third derivative is needed by the order-two optimizer: it appears only in the analysis and offline audit. □

Scope of the refinement. For the quadratic Taylor model, take $\mathcal{A} = 0$ so that the leading tensor is $\nabla^3 F(\theta^\star)$, recovering the usual Newton error coefficient. A calibrated nonlinear predictor replaces this by the residual tensor. A small tensor can improve the error constant without changing the asymptotic order. If this tensor is nonzero, a transient near-cubic slope does not establish cubic convergence; it may reflect a crossover before the quadratic term dominates. The bounds in Appendix B.2 explain why canonical residual tensors can be small near matching under bounded density derivatives, but do not imply that canonical α minimizes the error coefficient for arbitrary heterogeneous targets.

C.6 Safeguarded lazy calibration

The following sufficient safeguards make the extension's convergence argument explicit. After a successful certified step s_t, set $z_t = \theta_t + s_t$ and retain $g_t = \nabla F(\theta_t)$. Before the next calibration, allow at most K_{lazy} extra trials using the fixed-origin, value-matched predictor

$$m_{z_t}^{(0)}(v) = \mathcal{A}(z_t + v) + F(z_t) - \mathcal{A}(z_t).$$

A trial from $z_{t,j} = z_t + v_j$ queries this model at $v_j + u_j$; it is not recentered without recalibration. Accept a lazy increment only if, for fixed positive constants,

$$F(z_{t,j}) - F(z_{t,j} + u_j) \geq c_{\mathrm{lazy}}\|u_j\|^2,$$
$$\sum_j \|u_j\| \leq \min\{\kappa_{\mathrm{path}}\|g_t\|, \kappa_{\mathrm{move}}\|s_t\|\}. \tag{53}$$

The sum is over accepted increments within this block. Its endpoint becomes θ_{t+1} and receives a fresh residual jet and true-gradient stopping test. No lazy block follows a rejected certified step, and lazy trials do not alter the radius update (20). The added step-length budget links uncalibrated motion to certified progress; a bound involving only the anchor gradient does not supply that link in the proof below.

Corollary C.3 (Lazy-search complexity)**.** *Under Assumption C.1, including Lipschitz continuity on the lazy segments, the preceding block protocol and (53) give vanishing gradients along the complete accepted sequence of a zero-tolerance run, or finite exact stationarity. With tolerance $\varepsilon > 0$, the certified jet complexity is $O(\varepsilon^{-2})$ and the true scalarized value complexity is $O((K_{\mathrm{lazy}} + 1)\varepsilon^{-2})$.*

Proof. Lazy steps only decrease F. Thus each successful certified iteration still contributes its original true decrease, and the radius recursion is unchanged. The radius lower bound and successful-iteration count in the proof of Theorem 5.1 therefore apply before the first ε-stationary certified anchor.

For the full gradient limit, the path from θ_t to θ_{t+1} has length at most $(1+\kappa_{\mathrm{move}})\,\|s_t\|$ at a successful iteration and zero at a rejected one. Equation (48), with its constant divided by $1 + \kappa_{\mathrm{move}}$, controls this entire path length whenever $\|g_t\| \geq \varepsilon$. The same disjoint-excursion argument proves $\|g_t\| \to 0$ at certified anchors. Every intermediate lazy point is within $\kappa_{\mathrm{path}}\,\|g_t\|$ of the block endpoint; Lipschitz continuity transfers the limit to those points as well. At most K_{lazy} extra true-value calls accompany each certified iteration, proving the stated evaluation count. These safeguards establish global stationarity, not the local order of the augmented sequence. □

The E4 lazy implementation is an empirical variant using a local curvature estimate and its original path budget; the additional step-length budget above was not part of those runs. Its observed stationarity is reported directly, rather than inferred from this sufficient-condition corollary.

C.7 Implementing model decrease and interpreting scalability

The computational regime. Grouping oracle queries by returned derivative order r (values, values/gradients, or higher jets) gives

$$T_{\mathrm{CAL}} = T_{\mathrm{setup}} + \sum_r n_r^F \bar{c}_r^F + \sum_r n_r^{\mathcal{A}} \bar{c}_r^{\mathcal{A}} + T_{\mathrm{other}}, \tag{54}$$

where $n_r^F, n_r^{\mathcal{A}}$ count true-objective and predictor calls, and $\bar{c}_r^F, \bar{c}_r^{\mathcal{A}}$ are their respective mean costs. T_{other} includes subproblem algebra and control overhead. Setup is charged once per preference. Cached aggregate densities remove the target sum from predictor queries; the true generic-f oracle can still depend on m.

A net gain requires true-oracle savings to exceed predictor setup, inner-query, and model-management costs. E9 measures matched-order oracle costs and total solve time at comparable objectives as m increases.

Nonlinear inner solves used in E8/E9. The calibrated model is generally nonquadratic. The implementation starts at $v = 0$, shifts its current Hessian by $\max\{0, 10^{-7} - \lambda_{\min}(\nabla^2 m(v))\}I$, and computes a Newton direction, falling back to the negative model gradient if needed. Trial iterates are projected onto the trust-region ball; Armijo backtracking uses coefficient 10^{-4} and at most 22 trials. The best model decrease is retained together with a steepest-descent comparator. E8/E9 cap the inner loop at 12 iterations and use relative gradient tolerance 0.01 with absolute floor 10^{-9}. These settings target practical stationarity; E5 separately enforces the stronger local forcing condition (23).

A Cauchy comparator without a supplied uniform bound. Write $g = \nabla m(0) \neq 0$. The E8/E9 comparator initializes

$$\alpha_0 = \min\left\{\frac{\Delta}{\|g\|}, \frac{1}{\max\{1, \|\nabla^2 m(0)\|_{\mathrm{op}}\}}\right\}$$

and halves α while seeking

$$m(-\alpha g) - m(0) \leq -\tfrac{1}{4}\alpha\,\|g\|^2. \tag{55}$$

The center Hessian supplies an initial scale, not a certified bound over the ball. To connect this test to the theory, suppose $\left\|\nabla^2 m(v)\right\|_{\mathrm{op}} \leq \kappa_H$ on the ball, taking $\kappa_H \geq 1$ without loss of generality.

Table 4: Distinct notions of order used in the paper.

Notion	Order	Meaning
Likelihood-ratio normal form	4	Error as $p/q \to 1$
Calibrated model p	$(p+1, p)$	Value/gradient error in the step
Global complexity	$O(\varepsilon^{-2})$	Certified iterations to first-order stationarity
Local iteration	At least Q-p	Error recursion for $p = 2, 3$

Taylor's inequality ensures (55) whenever $\alpha \le 3/(2\kappa_H)$. In exact arithmetic, uncapped halving therefore terminates with

$$\alpha \ge \tfrac{1}{2}\min\{\Delta/\|g\|, 1/\kappa_H\}, \qquad m(0) - m(-\alpha g) \ge \tfrac{1}{8}\|g\|\min\{\Delta, \|g\|/\kappa_H\}.$$

Any returned step with at least this model decrease satisfies (19) with $\kappa_c = 1/8$; the algorithm need not know the numerical value of κ_H. The stored E8/E9 implementations cap this comparator search at 40 trials and retain a decreasing candidate; finite caps and arithmetic floors are implementation safeguards, not independent verification of the uniform-bound assumption. E4 uses its stated conservative constant, while E5 uses an analytic ballwise bound.

Parameter dimension and matrix-free extensions. Target count m, parameter dimension d, and sample-space dimension D impose distinct costs. Dense residual Hessians need $O(d^2)$ storage and dense Newton factorization $O(d^3)$ work; $p = 3$ also requires third derivatives. Tensor storage is not inherent: for $p = 2$,

$$\nabla_v^2 m_\theta^{(2)}(v)u = \nabla^2\mathcal{A}(\theta + v)u + [\nabla^2 F(\theta) - \nabla^2\mathcal{A}(\theta)]u. \tag{56}$$

Exact Hessian–vector products permit matrix-free inner solves under the same decrease and forcing conditions, but repeated true residual products can revisit all targets and negate caching savings. CAL-p1 permits a safeguarded first-order inner solver using a cached residual gradient. Such large-scale implementations are untested; approximate derivatives need additional error controls.

Integration dimension and sampled oracles. Aggregation holds for integrable densities independently of discretization. For $\alpha \ne \pm 1$, the forward predictor integral can use $\sum_i \lambda_i p_i^{r_\alpha}$ without normalizing the pseudo-target, but efficiency requires a reusable representation. Tensor grids have k^D nodes for k nodes per coordinate; fresh Monte Carlo nodes may require reevaluating all targets. E8/E9 test two-dimensional grids with refinement audits, not high-dimensional integration scalability or the theorems' exact continuous oracles. Target minibatching also requires error controls for noisy derivatives and acceptance tests; stochastic rates are not established here.

D Preference continuation

This extension uses standard implicit differentiation for smooth scalarizations (Blondel et al., 2022); it does not depend on α-aggregation. CAL supplies the common certified corrector.

Fix an orientation σ. Assume that each F_i^σ is C^3 near $\theta^\star$, that $\lambda^\star \in \mathrm{ri}(\Delta_m)$, and that

$$\nabla F_{\lambda^\star}^\sigma(\theta^\star) = 0, \qquad \nabla^2 F_{\lambda^\star}^\sigma(\theta^\star) \text{ is nonsingular.}$$

Parameterize the simplex by $\lambda_m = 1 - \sum_{j<m}\lambda_j$. The implicit function theorem yields a locally unique C^2 stationary branch $\theta^\star(\lambda)$. Writing $H_\lambda = \nabla^2 F_\lambda^\sigma(\theta^\star(\lambda))$, its derivatives satisfy

$$\frac{\partial\theta^\star}{\partial\lambda_j} = -H_\lambda^{-1}\big(\nabla F_j^\sigma(\theta^\star) - \nabla F_m^\sigma(\theta^\star)\big), \quad j < m. \tag{57}$$

The corresponding first-order predictor has parameter error $O(\|\delta\lambda\|^2)$ for nearby feasible preferences and initializes the corrector used in E6. The expansion assumes exact branch points and sensitivities; numerical errors and singular-Hessian crossings require separate analysis.

E Bregman transfer results

Let ϕ be differentiable and strictly convex on an open convex domain Ω, with $x_1, \dots, x_m, y \in \Omega$ and $\lambda \in \Delta_m$. For the dual-centroid identity, assume that $\sum_i \lambda_i \nabla\phi(x_i) \in \nabla\phi(\Omega)$, so that $\widetilde{x}_\lambda$ exists and is unique.

For $B_\phi(x,y) = \phi(x) - \phi(y) - \langle \nabla\phi(y), x-y\rangle$,

$$\sum_i \lambda_i B_\phi(x_i, y) = C_\lambda^- + B_\phi(\bar{x}_\lambda, y), \qquad \bar{x}_\lambda = \sum_i \lambda_i x_i, \tag{58}$$

$$\sum_i \lambda_i B_\phi(y, x_i) = C_\lambda^+ + B_\phi(y, \widetilde{x}_\lambda), \qquad \nabla\phi(\widetilde{x}_\lambda) = \sum_i \lambda_i \nabla\phi(x_i). \tag{59}$$

Expanding the definition proves (58); collecting the linear terms in y proves (59). For a common-generator objective, choosing the exact aggregate, including its constant, as the predictor gives $\mathcal{A} = F$, $\mathcal{E} = 0$, and vanishing calibration tensors. Theorems 5.1 and 5.3 then transfer whenever their Euclidean assumptions hold.

For the variable-first heterogeneous objective $\sum_i \lambda_i B_{\phi_i}(v, a_i)$, define

$$\Phi_\lambda = \sum_i \lambda_i \phi_i, \qquad b_\lambda = \sum_i \lambda_i \nabla\phi_i(a_i). \tag{60}$$

If Φ_λ is Legendre and $a_\lambda = (\nabla\Phi_\lambda)^{-1}(b_\lambda)$ exists, then

$$\sum_i \lambda_i B_{\phi_i}(v, a_i) = C_\lambda + B_{\Phi_\lambda}(v, a_\lambda). \tag{61}$$

Indeed, both sides equal $\Phi_\lambda(v) - \langle b_\lambda, v\rangle$ up to a constant. A derivative-matching Bregman template can therefore replace the α_f module under the same regularity, model-decrease, and acceptance assumptions. For heterogeneous generators, this algebraic identity alone does not guarantee that evaluating Φ_λ is cheaper than evaluating its component sum.

F Experimental protocols and additional results

Table 5: Structural numerical checks. "Max. dev." is the largest absolute deviation of an aggregate mean fitted exponent from theory.

Test	Design	Main observation
E1: normal form	20 seeds, three f, two orientations	Orders $(4,3,2)$; max. dev. 7.23×10^{-4}.
E2: aggregation	15,680 query checks	Max. scaled identity error 3.77×10^{-15}.
E2: amortization	$K = 2048$, 30 paired repeats	At $m = 256$, end-to-end speedup 2.46–3.14×.
E3: model order	720 raywise trials	Orders $(p+1, p)$; max. dev. 1.67×10^{-2}.

Generator conventions. The labels JS, KLChi, and KL4 denote

$$f_{\mathrm{JS}}(u) = u\log u - (u+1)\log\left(\frac{u+1}{2}\right),$$

$$f_{\mathrm{KLChi}}(u) = u\log u - u + 1 + \tfrac{1}{2}(u-1)^2,$$

$$f_{\mathrm{KL4}}(u) = u\log u - u + 1 + (u-1)^4.$$

Their (a_f, α_f) pairs are $(1/2, 0)$, $(2, -2)$, and $(1, -1)$, respectively. JS uses the unhalved normalization: $D_{f_{\mathrm{JS}}}(P\|Q) = \mathrm{KL}(P\|M) + \mathrm{KL}(Q\|M)$, where $M = (P+Q)/2$.

F.1 Local normal-form scaling

For E1, each of 20 seeds uses $K = 128$, draws $z_k \sim \mathcal{N}(0, 0.55^2)$ and $v_k \sim \mathcal{N}(0,1)$, sets $P = \mathrm{softmax}(z)$, centers v under P, and rescales it to $\|v\|_\infty = 1$. Along

$$Q_t(k) = \frac{P_k \exp(t v_k)}{\sum_\ell P_\ell \exp(t v_\ell)}, \qquad Q_0 = P, \tag{62}$$

we evaluate $R_f^-(t) = D_f(P\|Q_t) - a_f D_{\alpha_f}(P\|Q_t)$ and $R_f^+(t) = D_f(Q_t\|P) - a_f D_{\alpha_f}(Q_t\|P)$ and their first two analytic derivatives. The 21 step sizes span $[10^{-3.2}, 10^{-0.7}]$ and slopes are fitted on $[2\times10^{-3}, 3\times10^{-2}]$. We test JS, KLChi, and KL4 under the conventions above, together with the exact canonical control.

The fitted value/first-/second-derivative orders are $(4,3,2)$. Across the 18 aggregate cells, the largest mean-exponent deviation is 7.23×10^{-4}, the largest across-seed standard deviation is 5.55×10^{-3}, and all exact-control residuals are below 10^{-12}.

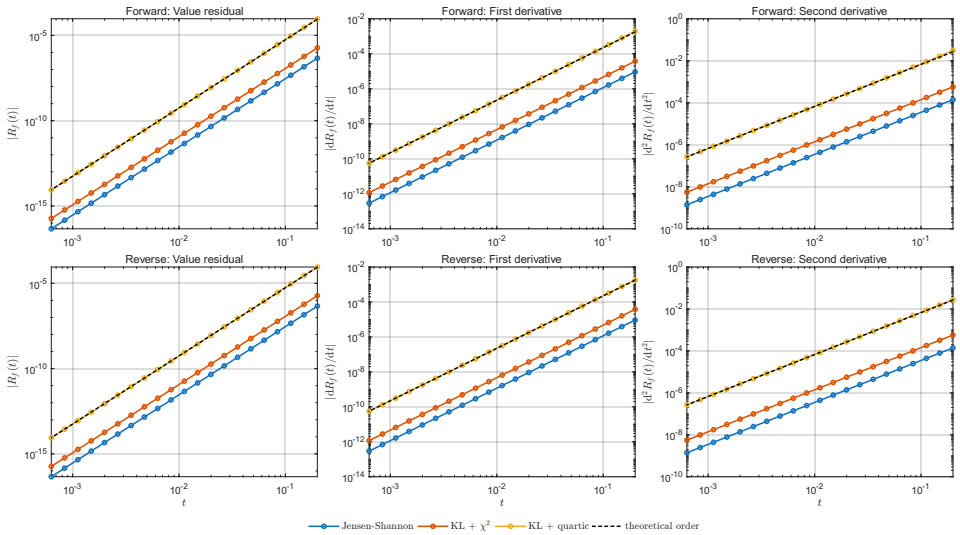

Figure 4: Local normal-form scaling in the forward and reverse orientations. Curves show medians over 20 seeds; dashed references have slopes 4, 3, and 2. Both axes are logarithmic.

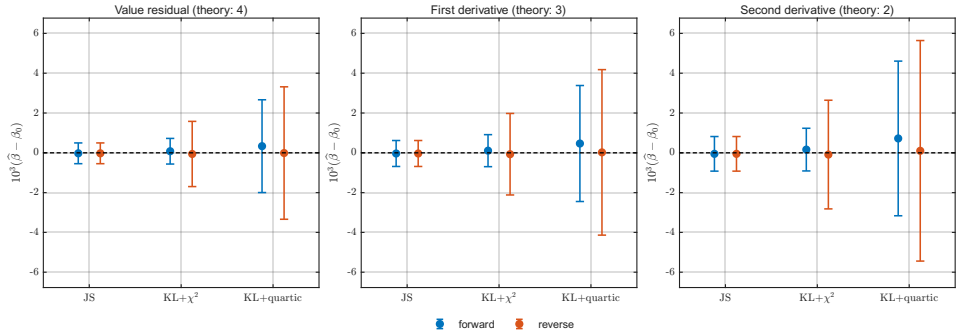

Figure 5: Fitted-exponent deviations for E1, scaled by 10^3. Markers show means over 20 seeds, error bars one standard deviation, and the dashed line exact agreement.

F.2 Exact aggregation and amortized cost

For the identity check, E2 uses $K = 128$, $m \in \{2, 4, 8, 16, 32, 64, 128\}$, $\alpha \in \{-3, -2, -1, 0, 1, 2, 3\}$, both orientations, 20 seeds, and eight query distributions per cell. All distributions are softmaxes of Gaussian logits with scale 0.65 and all preferences are positive. The direct scalarization is compared with (10)–(11), including $\omega^{\sigma}_{\lambda,\alpha}$. All 15,680 query comparisons pass the 10^{-10} tolerance; the maximum scaled identity and normalization errors are 3.77×10^{-15} and 6.66×10^{-16}.

The timing benchmark uses $K = 2048$, $m \in \{2, 4, 8, 16, 32, 64, 128, 256\}$, 256 sequential queries, and 30 paired repetitions with alternating order. A break-even count is reported only when the central empirical 95% interval of paired per-query savings is positive. At $m = 256$, median target-kernel speedups are 12.00×, 10.39×, and 12.08× for forward KL, reverse KL, and forward $\alpha = 0$; end-to-end speedups are 2.79×, 2.46×, and 3.14×, with break-even counts 45, 5, and 100.

F.3 Calibrated-model order diagnostics

E3 uses $K = 64$ categories, 16 targets, and 20 random seeds. We test the same three generators as in Appendix F.1, in both divergence orientations. For each seed, we sample calibration logits $z_c \sim \mathcal{N}(0, 0.45^2 I_K)$ and independent target logits $z_i \sim \mathcal{N}(0, I_K)$, and define

$$Q_0 = \mathrm{softmax}(z_c), \qquad P_i = \mathrm{softmax}(z_i), \qquad i = 1, \ldots, 16.$$

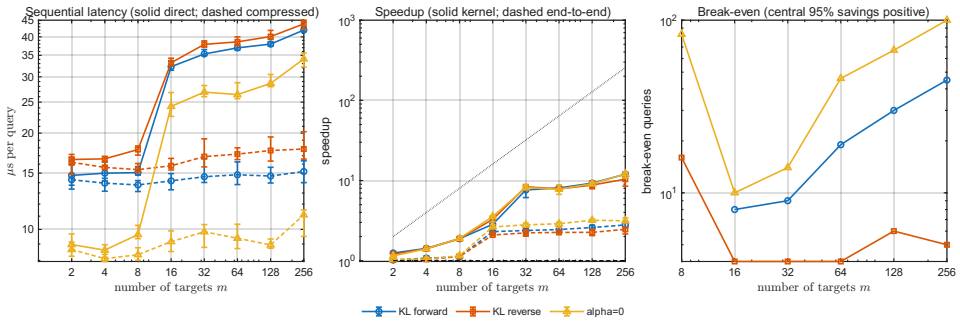

Figure 6: E2 amortized cost, $K = 2048$, 30 paired repetitions. Left: direct (solid) and aggregated (dashed) sequential time per query, with IQRs. Middle: median target-kernel (solid) and end-to-end (dashed, including query-feature formation) speedups; dotted diagonal and dashed horizontal references indicate m-fold and unit speedups. Right: break-even query count; missing segments indicate a nonpositive empirical 2.5th percentile of paired per-query savings.

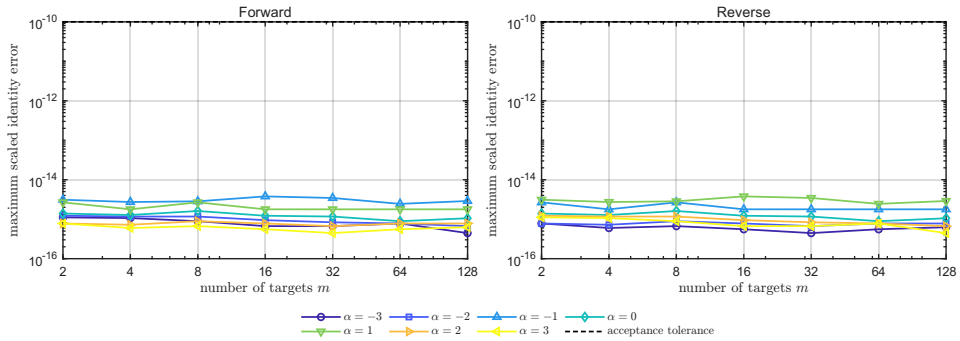

Figure 7: E2 aggregation identity errors in both orientations: maximum scaled discrepancy over 20 seeds and eight queries per seed, including the aggregate's constant and scale. The dashed line marks tolerance 10^{-10}; errors remain at double-precision roundoff.

The preference weights satisfy $\lambda_i \propto 0.25 + U_i$, where $U_i \sim \mathrm{Unif}(0, 1)$, followed by normalization onto the simplex. A Gaussian logit direction is centered so that $\sum_k Q_0(k)d_k = 0$ and rescaled to satisfy $\|d\|_\infty = 1$. Both signed rays are evaluated.

We use 25 logarithmically spaced step magnitudes in $[10^{-3}, 2 \times 10^{-1}]$. Slopes are fitted over the prespecified window $[2 \times 10^{-3}, 3 \times 10^{-2}]$ after excluding errors below 10^{-13}. If fewer than six eligible points remain, the upper endpoint is extended to 6×10^{-2}. The structured predictor includes the aggregation constant and normalization coefficient from Theorem 3.2. Residual Taylor coefficients through degree three are obtained from fourth-order truncated-series arithmetic, rather than numerical differentiation.

Along each signed ray, we fit the log–log slopes of $|F(\theta_c + td) - m_c^{(p)}(td)|$ and the corresponding directional-derivative error for $p \in \{1, 2, 3\}$. Figure 8 shows the predicted $(p + 1, p)$ scaling. The targets are not artificially close: the median maximum absolute log-likelihood ratio is 3.94 and the median target-to-center KL is 0.540.

Each generator–orientation–order cell contains 40 trials (20 seeds, two signs). Across 18 cells, the largest mean-order deviation is 1.67×10^{-2} and the largest standard deviation 0.148 (Figure 9); all 720 trials yield valid value and directional-derivative fits.

At $p = 2$, one value fit and two directional-derivative fits have $R^2 < 0.95$; the latter residuals change sign within the window. All fits are retained.

As an algebraic control, we additionally use an exact α-divergence with $\alpha = 0.4$. All 80 control configurations pass, with maximum value, directional-derivative, and residual-jet discrepancies 1.94×10^{-15}, 1.25×10^{-16}, and 1.78×10^{-15}, respectively.

F.4 Full optimization: robustness and cost

Problem construction. We use $Q_\theta = \mathrm{softmax}(b + B\theta)$ with K categories and $d < K$ parameters. Let $\Pi = I_K - \mathbf{1}\mathbf{1}^\top/K$. We obtain $U \in \mathbb{R}^{K \times d}$ by QR factorization of a column-centered

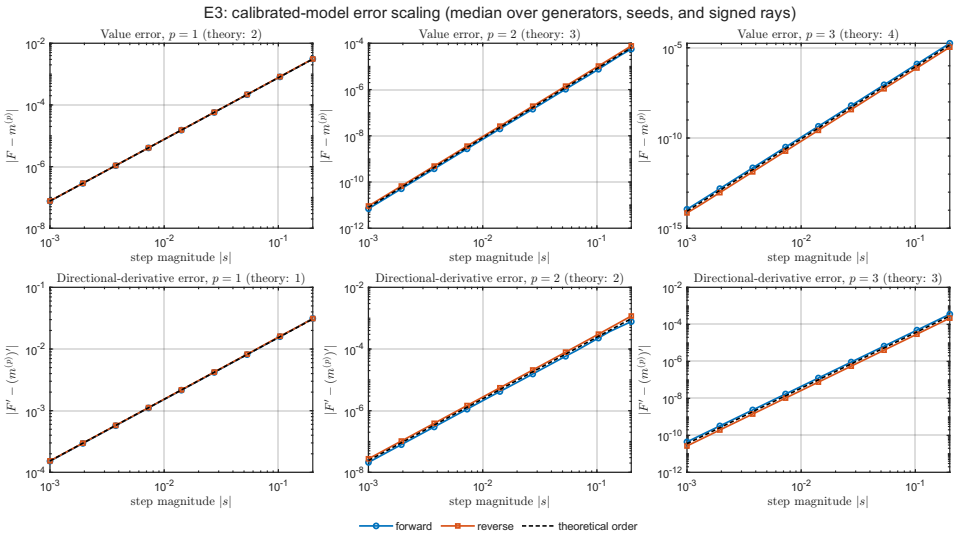

Figure 8: Step-local calibrated-model error scaling. Columns correspond to $p = 1, 2, 3$; rows show value and directional-derivative errors. Curves are medians over three generators, 20 seeds, and two signed rays; dashed references have slopes $p + 1$ and p.

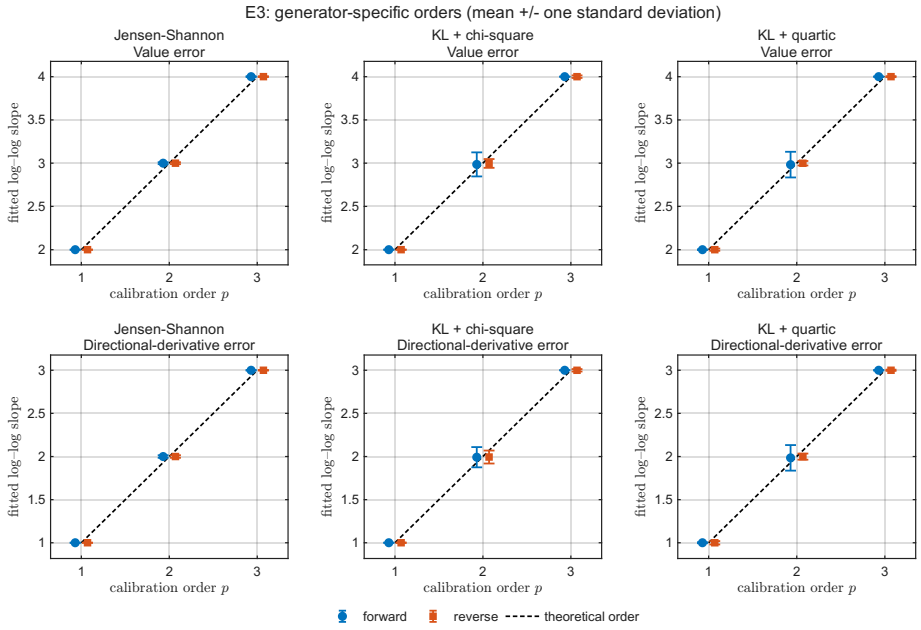

Figure 9: E3 fitted model-error orders by generator (columns): value errors (top, order $p + 1$) and directional-derivative errors (bottom, order p). Markers and bars show means and one standard deviation over 20 seeds and two signed rays, not confidence intervals. Blue circles/orange squares denote forward/reverse orientations, offset for visibility; dashed lines mark theoretical orders.

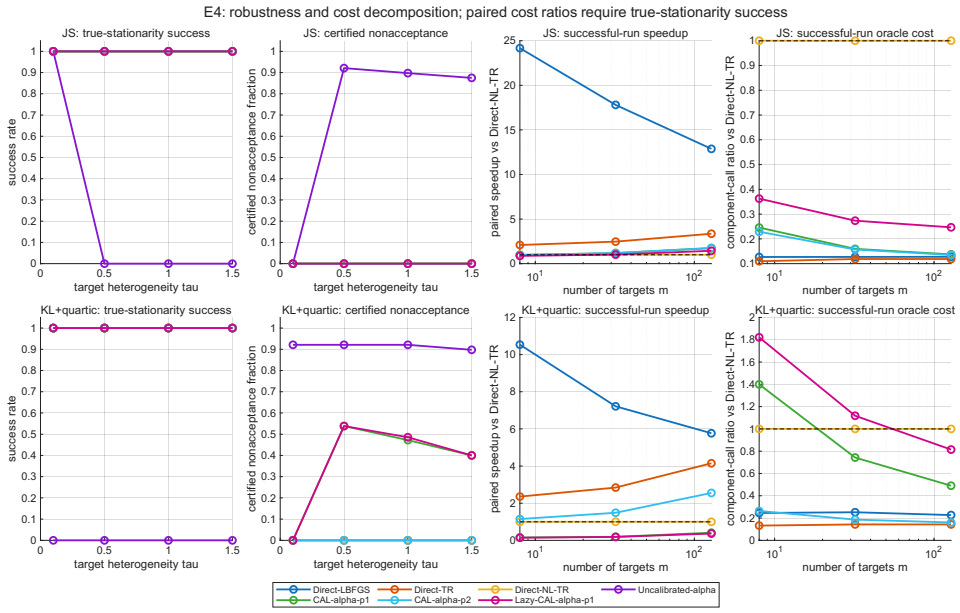

Figure 10: E4 robustness and cost, 30 paired seeds. Rows: JS and KL+quartic. Left: success rates and median certified nonacceptance fractions, including failed subproblems and rejected trials (not applicable to L-BFGS). Right: median paired runtime speedups and component-call ratios against Direct-NL-TR in the target-count sweep; both runs must succeed.

Gaussian matrix and set $B = 1.2U$. Independently draw $b \sim \mathcal{N}(0, 0.25^2 I_K)$, $\xi_i \sim \mathcal{N}(0, I_d)$, $z_i \sim \mathcal{N}(0, I_K)$, and $\zeta \sim \mathcal{N}(0, I_d)$. The targets and initialization are

$$P_i = \mathrm{softmax}(b + \tau B \xi_i + 0.1\tau \Pi z_i)\,, \tag{63}$$

$$\theta_0 = -\frac{0.75\tau}{m} \sum_{i=1}^{m} \xi_i + 0.8 \max\{\tau, 0.25\}\zeta. \tag{64}$$

The perturbation Πz_i allows targets outside the model family. A deterministic seed mapping fixes each problem, and all methods use the same targets, preference, and initialization for a paired comparison.

We use JS and KL4 with the generator conventions stated at the beginning of Appendix F, and test both divergence orientations.

Experimental configurations. The robustness experiment fixes $(K, d, m) = (128, 16, 32)$ and varies $\tau \in \{0.1, 0.5, 1.0, 1.5\}$. Preferences are uniform, drawn from $\mathrm{Dirichlet}(1, \ldots, 1)$, or biased with $\lambda_1 = 0.6$ and $\lambda_i = 0.4/(m-1)$ for $i > 1$. The scaling experiment uses $(K, d) = (512, 64)$, $m \in \{8, 32, 128\}$, $\tau = 1$, and Dirichlet preferences. Both experiments use seeds $1, \ldots, 30$, two generators, two orientations, and seven methods. They contain 10,080 and 2,520 method runs, respectively, giving 1,800 main-experiment runs per method.

Methods and numerical settings. The seven methods are Direct-LBFGS, Direct-TR, Direct-NL-TR, Uncalibrated-α, CAL-α-$p1$, CAL-α-$p2$, and Lazy-CAL-α-$p1$. The uncalibrated method uses the aggregated predictor without residual-derivative calibration, but its success is assessed using the true objective gradient.

The stopping tolerance is 10^{-6}, with a budget of 120 certified iterations and 360 trust-region trials. Trust-region parameters are $\Delta_0 = 0.5$, $\Delta_{\max} = 2$, $\Delta_{\min} = 10^{-10}$, $(\eta_1, \eta_2) = (0.1, 0.75)$, and $(\gamma_1, \gamma_2) = (0.5, 2)$. The Cauchy-decrease check uses $\kappa_c = 0.1$ and the conservative curvature bound 10^8. The common nonlinear subproblem routine permits at most 12 inner iterations and uses gradient tolerance 10^{-6}. Direct-LBFGS uses memory 10 and Armijo backtracking.

The lazy variant permits at most three trials per block and uses $c_{\mathrm{lazy}} = 10^{-4}$. Its path coefficient is selected as $0.25/L_{\mathrm{est}}$, where

$$L_{\mathrm{est}} = 2 \max\{1, \left\|\nabla^2 F(\theta^c)\right\|_{\mathrm{op}}\}.$$

Table 6: E4 at $m = 128$: median paired ratios against Direct-NL-TR. Each displayed method succeeds in all 60 runs per generator. Uncalibrated fails in all 60, so its paired costs are unavailable. Higher speedups and lower call ratios are better.

	Time speedup		Call ratio	
Method	JS	KL+quartic	JS	KL+quartic
Direct-LBFGS	12.87	5.77	0.128	0.227
Direct-TR	3.36	4.15	0.119	0.143
Direct-NL-TR	1.00	1.00	1.000	1.000
CAL-α-$p1$	1.74	0.41	0.137	0.490
CAL-α-$p2$	1.76	2.55	0.135	0.160
Lazy-CAL-α-$p1$	1.44	0.36	0.247	0.815

This is a numerical local-curvature estimate, not a verified global Lipschitz bound. The implementation retains the stale model's certified coordinate origin, enforces the path budget and true-objective sufficient decrease, and leaves the certified radius unchanged during lazy trials. The additional step-length budget in Appendix C.6 was not enforced here, so these measurements evaluate the implemented heuristic, rather than verify every hypothesis of Corollary C.3.

Metrics and paired comparisons. Success is defined exclusively by $\|\nabla F_\lambda^\sigma(\theta_{\text{final}})\| \le 10^{-6}$. End-to-end time includes per-method pseudo-target construction, residual calibration, inner solves, and true-objective acceptance checks; data generation, derivative audits, plotting, and export are excluded.

The component-equivalent ledger charges m units per full scalarization query or pseudo-target construction and one per predictor query. Derivative orders are recorded but equally weighted: the ledger measures query structure, not floating-point work.

For method a and a paired Direct-NL-TR run, define

$$S_a = \frac{T_{\text{Direct-NL-TR}}}{T_a}, \qquad R_a = \frac{C_a}{C_{\text{Direct-NL-TR}}},$$

where T is recorded end-to-end time and C is the component-equivalent count. We report medians of per-problem ratios only when both runs pass the true-gradient test.

Robustness and cost results. All six direct and calibrated methods succeed in every main-experiment run, whereas the uncalibrated method succeeds in 180/1,800 runs. Table 6 summarizes the scaling results at $m = 128$, with 30 seeds and two orientations per generator.

For CAL-α-$p2$, percentile bootstrap 95% intervals are $[1.72, 1.79]$ (JS) and $[2.45, 2.66]$ (KL+quartic), using 10,000 resamples of 30 seed clusters with both orientations retained. They measure between-instance variation, not repeated-timing noise.

Figure 11 contrasts calibrated convergence with uncalibrated stalling. Direct-LBFGS and Direct-TR remain competitive or faster; the reported gain is against Direct-NL-TR's repeated full-objective queries.

Implementation controls. Four generator–orientation derivative checks pass, with maximum relative gradient and Hessian discrepancies 3.06×10^{-12} and 3.71×10^{-12}. An additional exact-α experiment uses $\alpha = 0.4$, $(K, d, m) = (64, 8, 16)$, ten seeds, both orientations, and five methods, giving 100 runs. The 80 comparisons against Direct-NL-TR agree in their success decisions, with maximum final-objective discrepancy 2.12×10^{-11}.

A separate small-dimensional control uses $(K, d, m) = (64, 4, 16)$, ten seeds, two generators, both orientations, and five methods, giving 200 runs. All 40 CAL-α-$p3$ runs reach the true-gradient tolerance. In this control, the third residual tensor is approximated by centered differences of residual Hessians with step 2×10^{-4} and then symmetrized. This tests attainable optimization accuracy, not the Q-cubic convergence rate. Together with the main experiments, E4 contains 12,900 runs.

F.5 Local high-order convergence and condition ablations

Problem construction. E5 tests Theorem 5.3 on the affine chart $q_\theta = q_0 + B\theta > 0$, with $(K, d, m) = (6, 2, 4)$. For seed $s \in \{1, \dots, 10\}$, the MATLAB random stream is

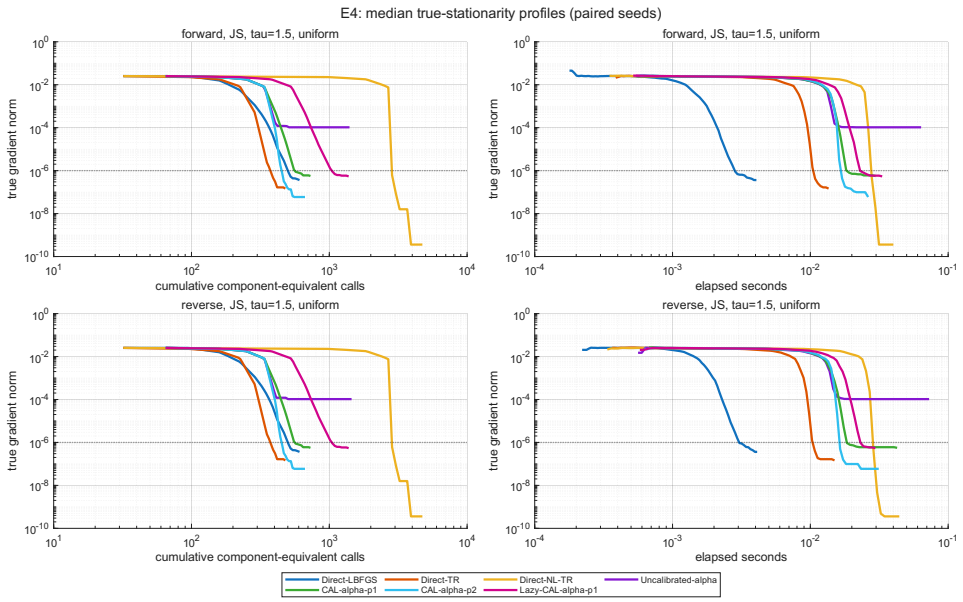

Figure 11: E4 true-gradient profiles: JS, $(K, d, m) = (128, 16, 32)$, $\tau = 1.5$, uniform preferences; rows separate orientations. Budgets are component-equivalent calls (left) and time (right). Curves are pointwise medians over 30 seeds after log-coordinate interpolation; final gradients are carried forward, and budgets before the first record are unavailable. The dotted line marks 10^{-6}.

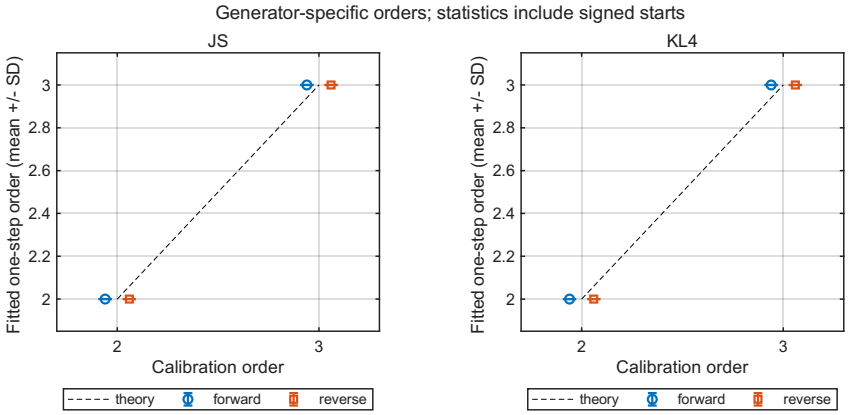

Figure 12: E5 one-step orders: means and one standard deviation over ten seeds and two paired signs (bars smaller than markers; not confidence intervals). Dashed lines mark order p; actual-iterate orders appear in Table 7.

`rng(27100+s,'twister')`. Draw $z \sim \mathcal{N}(0, 0.35^2 I_K)$ and a standard Gaussian matrix $G \in \mathbb{R}^{K\times d}$, and set

$$q_0 = \mathrm{softmax}(z), \qquad \widetilde{G} = G - \mathbf{1}(q_0^\top G), \qquad B = \frac{0.4\ \mathrm{diag}(q_0)\widetilde{G}}{\max_k \left\|\widetilde{G}_{k,:}\right\|_2}. \tag{65}$$

Thus $\mathbf{1}^\top B = 0$. Targets are $P_i = \mathrm{softmax}(\log q_0 + \xi_i)$ with independent $\xi_i \sim \mathcal{N}(0, 0.7^2 I_K)$, and preferences are $\lambda_i = w_i / \sum_j w_j$, where w_i are independent $\mathrm{Unif}(0.25, 1.25)$ draws. We use the JS and KL+quartic generators specified in Appendix F.1, in both orientations. The same primitive random draws are used across generators, orientations, and calibration orders within each seed.

Precision and local solves. All reported E5 results use 120-digit variable-precision arithmetic. A separate damped Newton solve obtains a reference $\theta^\star$ with $\|\nabla F(\theta^\star)\| \le 10^{-90}$. The 40 nondegenerate references have minimum Hessian eigenvalues between 2.206×10^{-3} and 0.665. Reference points are used only for initialization and error measurement; acceptance and stopping use the true objective and its gradient. The first seed's four generator–orientation references are recomputed at

160 digits; the largest parameter discrepancy is 4.88×10^{-108}. The error-resolution threshold is 10^{-80} in all cases; this is a numerical diagnostic, not an interval certificate.

Primary runs use analytic residual derivatives through order $p \in \{2, 3\}$. Damped Newton iterations from zero seek the local model-stationary branch while maintaining positive model Hessians, with at most 35 inner iterations. The inner condition is (23) with $\kappa_s = 0.01$; an arithmetic floor of 10^{-105} is recorded separately, and an untruncated primary step is accepted only if the original forcing condition is met. The initial and maximum radii are 1 and 4, $(\eta_1, \eta_2) = (0.1, 0.75)$, and radii are halved after rejection and doubled after very successful steps. Each applied ball is restricted to 0.45 times the distance to the boundary of the positive affine chart. An analytic model-curvature bound over that ball is used in the Cauchy test with $\kappa_c = 0.1$. Runs stop at $\|\nabla F(\theta)\| \le 10^{-30}$ or after 60 trials.

Independent one-step maps. Let v be a normalized Gaussian direction and define

$$r_\star = \min_{k:\|B_{k,:}\|_2>0} \frac{q_{\theta^\star}(k)}{\|B_{k,:}\|_2}, \qquad c_\star = \min\{1, 0.2r_\star\}, \qquad \theta_{\epsilon,h} = \theta^\star + \epsilon h c_\star v, \quad \epsilon \in \{-1, 1\}. \tag{66}$$

For each problem, order, and sign, one step is initialized separately at each $h \in \{10^{-2}, 10^{-3}, \ldots, 10^{-8}\}$, giving 1,120 probes. We regress the logarithm of the resulting parameter error against the logarithm of the initial parameter error, using the four prespecified radii $h \in \{10^{-3}, \ldots, 10^{-6}\}$. A fit requires all four points to be accepted, untruncated, forcing-compliant, and above the error-resolution threshold. All 1,120 probes and all 160 fits are valid; no slope-based filtering or adaptive fitting window is used. The largest individual slope deviation from p is 5.79×10^{-4}, the largest within-group standard deviation is 2.21×10^{-4}, and the minimum R^2 is 0.999999960. Figures 13 and 12 report scaling and orders. Signed starts share a seed; JS orientations coincide and are not independent replications.

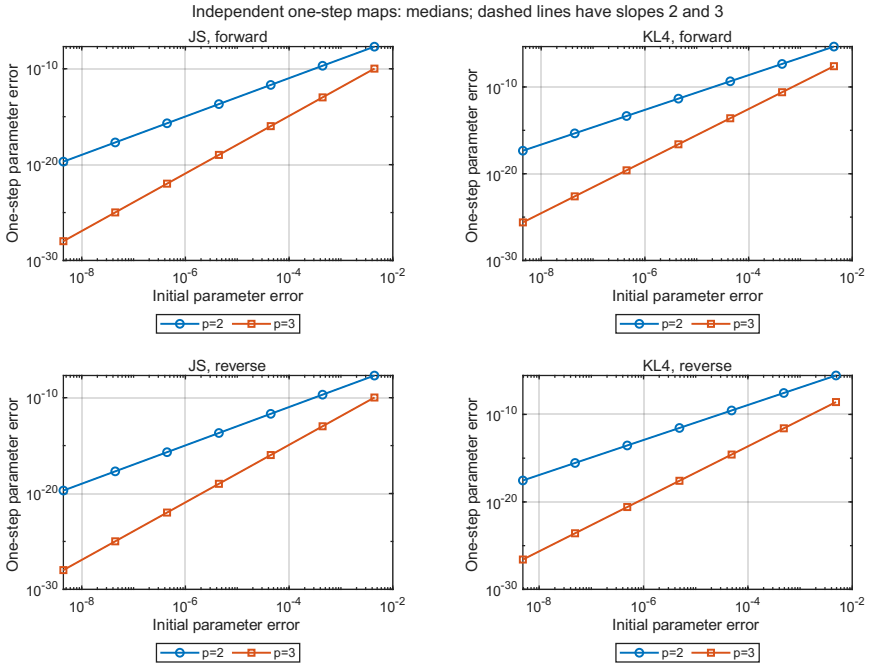

Figure 13: E5 independent one-step errors by generator (columns) and orientation (rows). Coordinates are medians over ten seeds and two signs; dashed slopes two and three largely overlap the curves. Fits use each seed/sign's prespecified window, not the median curves.

Actual iterates. Primary trajectories start at $\theta_{\epsilon,0.1}$, giving $10 \times 2 \times 2 \times 2 \times 2 = 160$ runs. All reach the true-gradient tolerance; the largest final gradient norm is 3.43×10^{-31}. All 542 primary steps are accepted and untruncated and pass the Cauchy test. The largest ratio $\left\|\nabla m_t^{(p)}(s_t)\right\| / \|\nabla F(\theta_t)\|^p$ is $0.009790 < 0.01$. For $e_t = \|\theta_t - \theta^\star\|$, define

$$\widehat{q}_t = \frac{\log(e_{t+1}/e_t)}{\log(e_t/e_{t-1})}. \tag{67}$$

Table 7: E5 actual-iterate orders: mean $\pm$ one standard deviation of the last eligible $\widehat{q}_t$ from each run. Each cell contains ten seeds and two paired signs. JS orientations coincide mathematically.

Generator	Orientation	$p = 2$	$p = 3$
JS	Forward	1.993 ± 0.045	2.996 ± 0.012
JS	Reverse	1.993 ± 0.045	2.996 ± 0.012
KL+quartic	Forward	2.014 ± 0.029	2.996 ± 0.012
KL+quartic	Reverse	1.995 ± 0.015	2.994 ± 0.013

An eligible observation requires three decreasing errors above the resolution threshold and two consecutive accepted, untruncated steps satisfying the local checks. Table 7 reports the last eligible observation from all 160 runs as finite-trajectory evidence.

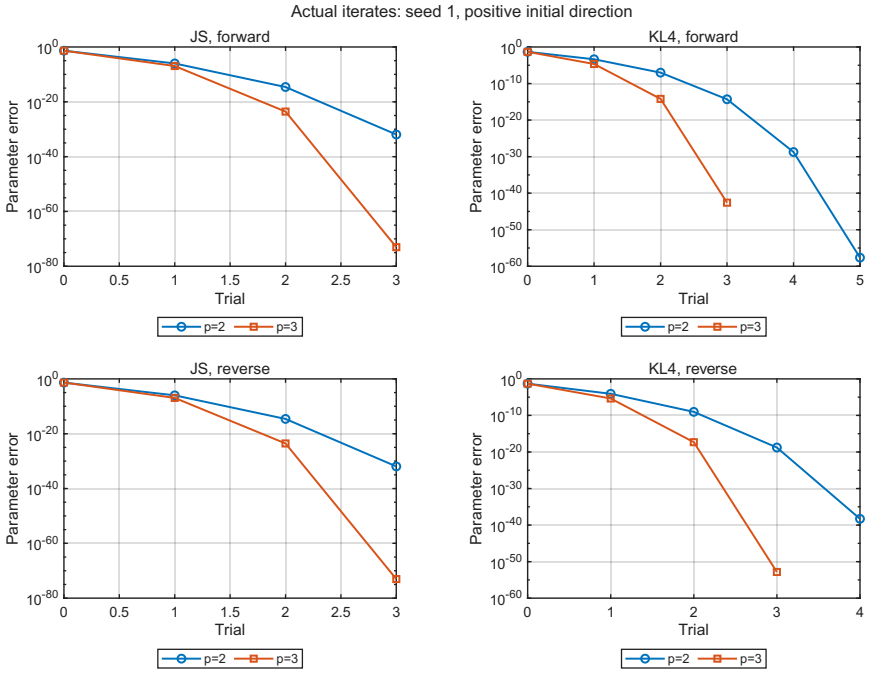

Figure 14: E5 parameter-error trajectories, seed 1, positive initial direction, 120-digit arithmetic. Panels separate generators and orientations; Table 7 summarizes all runs.

Condition ablations and singular control. Four additional $p = 3$ runs modify one condition on the first KL+quartic forward problem with the positive initial direction (Figure 15). A fixed inner tolerance of 10^{-8} stalls at a true-gradient norm of 6.23×10^{-11}. Restricting every applied step to at most one quarter of the computed branch-step norm produces a late error ratio of 0.75 and observed order 1.00; the run reaches its 60-trial budget before the outer tolerance. Centered differences of residual Hessians with fixed step 2×10^{-4} still attain stationarity, but the last observed order is 2.257 rather than close to three. This short trajectory illustrates third-jet sensitivity, not a definitive finite-difference asymptotic order. An initial radius of 10^{-3} causes five truncated steps, followed by recovery of an observed order of 2.982.

Finally, setting the second column of B to zero creates a flat parameter direction. Starting from $\theta^\star + (0.02, 0.1)^\top$, the singular control reaches a gradient norm of 3.64×10^{-58} while its distance to the selected reference parameter remains 0.1. Thus stationarity need not identify a unique parameter without nondegeneracy. These five illustrative controls bring the total to 165 trajectories.

F.6 Preference continuation and sensitivity prediction

Problem construction. E6 evaluates the nonsingular preference sensitivity formula in Appendix D. We use the affine probability chart $q_\theta = q_0 + B\theta > 0$, with $(K, d, m) = (24, 3, 3)$. For seed $s \in \{1, \ldots, 10\}$, the MATLAB random stream is `rng(61000+s,'twister')`. Draw

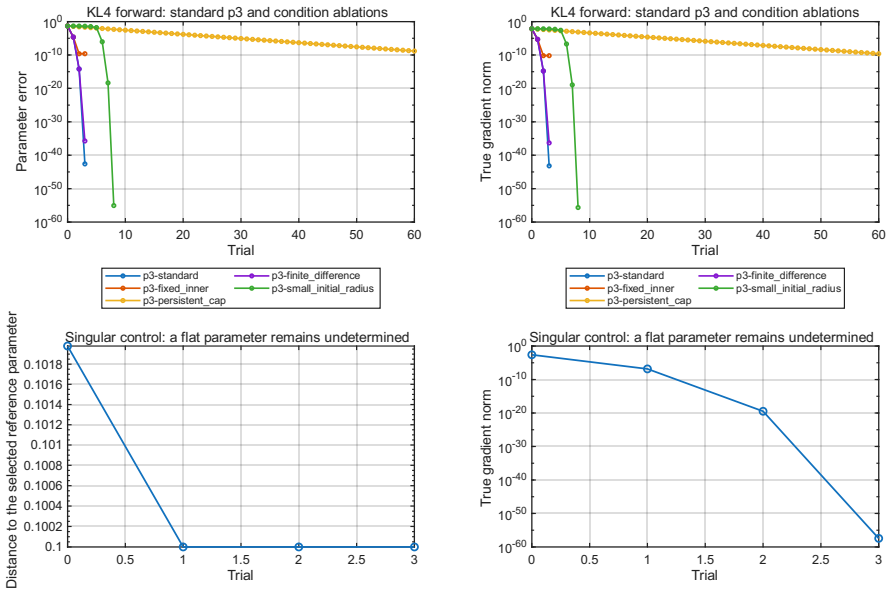

Figure 15: E5 individual condition ablations on the first forward KL+quartic problem. Top: parameter errors and true gradients for standard $p = 3$ and four controls. Bottom: a flat parameter direction permits vanishing gradient without convergence to the selected reference. The text reports each control's outcome.

$z \sim \mathcal{N}(0, 0.35^2 I_K)$ and a standard Gaussian matrix $A \in \mathbb{R}^{K\times d}$, and define

$$q_0 = \mathrm{softmax}(z), \qquad \widetilde{A} = A - \mathbf{1}(q_0^\top A), \qquad B = \frac{0.4\ \mathrm{diag}(q_0)\widetilde{A}}{\max_k \left\|\widetilde{A}_{k,:}\right\|_2}. \tag{68}$$

Then $\mathbf{1}^\top B = 0$. For independent $u_i \sim \mathcal{N}(0, I_d)$, set

$$\vartheta_i = 1.2\,\frac{u_i}{\|u_i\|_2}, \qquad P_i = q_0 + B\vartheta_i, \qquad i = 1, 2, 3. \tag{69}$$

The construction gives $P_i(k) \geq 0.52 q_0(k) > 0$; the implementation normalizes the target columns to remove floating-point normalization error. The same primitive random draws are used across generators and divergence orientations within each seed. We use the JS and KL+quartic generators from Appendix F.1; KL4 in the figures denotes KL+quartic.

Preference loops and initializers. The preference path is the closed triangle with vertices

$$(0.90, 0.05, 0.05)^\top, \qquad (0.05, 0.90, 0.05)^\top, \qquad (0.05, 0.05, 0.90)^\top.$$

Each edge is subdivided into $L \in \{4, 8, 16\}$ equal intervals, giving $3L + 1 \in \{13, 25, 49\}$ sampled points including the repeated starting vertex. Both traversal directions are tested independently of the divergence orientation.

Fix a divergence orientation σ. Let θ_{j-1} denote the preceding corrected solution, $\delta\lambda_j = \lambda_j - \lambda_{j-1}$, and

$$G_{j-1} = \left[\nabla F_1^\sigma(\theta_{j-1}),\ \nabla F_2^\sigma(\theta_{j-1}),\ \nabla F_3^\sigma(\theta_{j-1})\right], \qquad H_{j-1} = \nabla^2 F^\sigma_{\lambda_{j-1}}(\theta_{j-1}),$$

where the component objectives are defined in (1).

At subsequent preferences, the three initializers are

$$\begin{aligned} \theta_{j,0}^{\mathrm{Cold}} &= 0, \\ \theta_{j,0}^{\mathrm{Warm}} &= \theta_{j-1}, \\ \theta_{j,0}^{\mathrm{Tangent}} &= \theta_{j-1} - H_{j-1}^{-1} G_{j-1}\delta\lambda_j. \end{aligned} \tag{70}$$

Since $\mathbf{1}^\top \delta\lambda_j = 0$, Tangent equals (57). Each method uses its own preceding corrected endpoint, with a common first initialization $\theta = 0$. Tangent reuses retained component jets; independent references never initialize timed paths.

Common corrector and safeguards. All three strategies use the same order-two nonlinear calibrated model, Cauchy decrease safeguard, and true-objective acceptance test. The initial trust-region radius is reset to 1 at every preference, with maximum radius 8. We use $(\eta_1, \eta_2) = (0.1, 0.75)$ and Cauchy constant $\kappa = 0.1$. The radius is halved after rejection and doubled after a very successful step, subject to the maximum radius. Each applied ball is restricted to 0.45 times the distance to the affine-chart boundary.

The nonlinear model solve allows at most 25 inner Newton iterations and uses the relative gradient target

$$\max\{10^{-12},\, 0.01\,\|\nabla F_\lambda^\sigma(\theta)\|\}.$$

The solver also evaluates the model's center Newton candidate, retains the larger predicted reduction, and falls back to a Cauchy step if needed. Each preference solve permits at most 200 trials and stops when $\|\nabla F_\lambda^\sigma\| \le 10^{-8}$. These settings assess continuation efficiency; they are not the strict high-order inner forcing protocol of E5.

Tangent falls back to Warm after a failed Hessian definiteness/conditioning check and may shorten predictions to preserve positivity; neither occurs here. All methods freshly evaluate true component jets at each initializer.

Reference accuracy and validation. Independent damped Newton solves, each initialized at $\theta = 0$, obtain reference solutions with target gradient norm 10^{-13}. The maximum recorded reference gradient norm is 9.88×10^{-14}, and the minimum reference Hessian eigenvalue is 1.96×10^{-3}. Derivative, aggregation, center-calibration, and stable objective-difference checks all pass. The maximum normalized finite-difference discrepancy for the preference derivative is 8.24×10^{-9}.

The first timing repetition retains 20,880 primary preference-point records:

$$10 \times 2 \times 2 \times 2 \times 3 \times (13 + 25 + 49) = 20{,}880.$$

The factors correspond to seeds, generators, divergence orientations, traversal directions, and initializers. Including three timing repetitions gives 62,640 point solves. All timed runs satisfy the stationarity and reference objective-agreement checks, and deterministic work counts agree across repetitions. The largest primary final gradient norm is 9.96×10^{-9} and the largest absolute scalarization gap to a reference is 9.65×10^{-15}.

Independent local prediction maps. To isolate initializer accuracy, use $\lambda_0 = (0.2, 0.3, 0.5)^\top$ and a random unit vector $v \in \mathbb{R}^3$ satisfying $\mathbf{1}^\top v = 0$. For both signs and

$$h = 0.04\,2^{-r}, \qquad r = 0, \ldots, 5,$$

compute independent reference solutions at λ_0 and $\lambda_0 + \delta\lambda$, where $\delta\lambda = \pm h v$. The two errors are

$$e_{\mathrm{W}}(h) = \|\theta^\star(\lambda_0) - \theta^\star(\lambda_0 + \delta\lambda)\|\,, \tag{71}$$

$$e_{\mathrm{T}}(h) = \left\|\theta^\star(\lambda_0) - H_0^{-1} G_0 \delta\lambda - \theta^\star(\lambda_0 + \delta\lambda)\right\|. \tag{72}$$

Here H_0 and G_0 are evaluated at the base reference. On a smooth nonsingular solution branch, the expected orders are $e_{\mathrm{W}}(h) = O(h)$ and $e_{\mathrm{T}}(h) = O(h^2)$.

Fits exclude errors below the diagnostic resolution floor

$$\max\left\{10^{-12}, 20\left(\frac{\|g_{\mathrm{ref},0}\|}{\mu_{\mathrm{ref},0}} + \frac{\|g_{\mathrm{ref},h}\|}{\mu_{\mathrm{ref},h}}\right)\right\},$$

where μ_{ref} is the smallest reference Hessian eigenvalue. Tangent fits additionally require a feasible prediction. All 160 signed fits are valid and use all six step sizes. We average the two signed slopes within each seed before computing the mean and standard deviation across ten seeds.

Timing and work accounting. After untimed warm-up, each method/traversal is timed three times in randomized method order. For each seed, generator, divergence orientation, and grid, paired speedups use median whole-path times over repetitions, then the median ratio over traversals, followed by the median and IQR across seeds. Repetitions and traversals are not independent instances; a failed traversal invalidates its seed summary.

Whole-path times include the common first solve, aggregation, initialization, tangent algebra, model construction, inner solves, and acceptance checks; references, diagnostics, tables, plots, and export

Table 8: E6 local prediction orders. Entries are mean ± one standard deviation across ten seeds, after averaging the two signed slopes within each seed. The reference orders are one for Warm and two for Tangent.

Generator	Orientation	Warm	Tangent
JS	Forward	0.999989 ± 0.000005	2.000089 ± 0.000283
JS	Reverse	0.999989 ± 0.000005	2.000089 ± 0.000283
KL+quartic	Forward	1.000184 ± 0.000122	2.000039 ± 0.000690
KL+quartic	Reverse	1.000238 ± 0.000082	2.000341 ± 0.000808

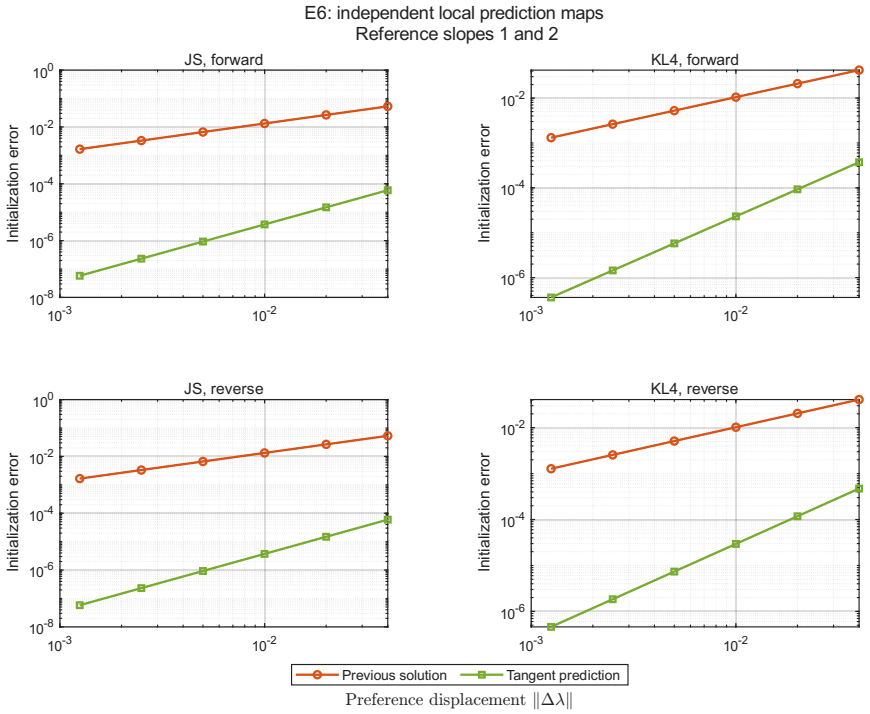

Figure 16: E6 independent prediction errors: medians over signs within each seed, then over ten seeds. Warm uses the base reference; Tangent adds its sensitivity prediction. Dashed slopes one and two nearly overlap the curves. The horizontal axis measures preference displacement, not corrector iteration.

are excluded. A true-jet pass evaluates all component values, gradients, and Hessians. Aggregation, predictor queries, objective differences, and linear solves are separately counted but included in runtime; jet counts are not total operation counts.

Table 3 in the main text reports the finest-grid costs.

Pathwise initializer quality and objective tradeoffs. On the finest grid, Tangent reduces pre-correction gradients after the common first point; changes at triangle vertices reflect the new preference direction (Figure 18).

All corrected paths agree numerically with independent references (Figure 19). This loop samples supported tradeoffs, not the full Pareto surface; singular-Hessian crossings, branch switching, and real-data performance are untested. JS orientations are algebraically identical, not independent replications.

F.7 Template choice after second-order calibration

Purpose and compared models. E7 fixes calibration order and varies the nonlinear template. Write $F = F_\lambda^\sigma$ and define

$$\mathcal{A}_\alpha(\theta) = a_f \sum_{i=1}^{m} \lambda_i D_\alpha^\sigma(P_i, Q_\theta).$$

Let

$$T_{2,\theta} G(s) = G(\theta) + \langle \nabla G(\theta), s \rangle + \frac{1}{2} s^\top \nabla^2 G(\theta) s.$$

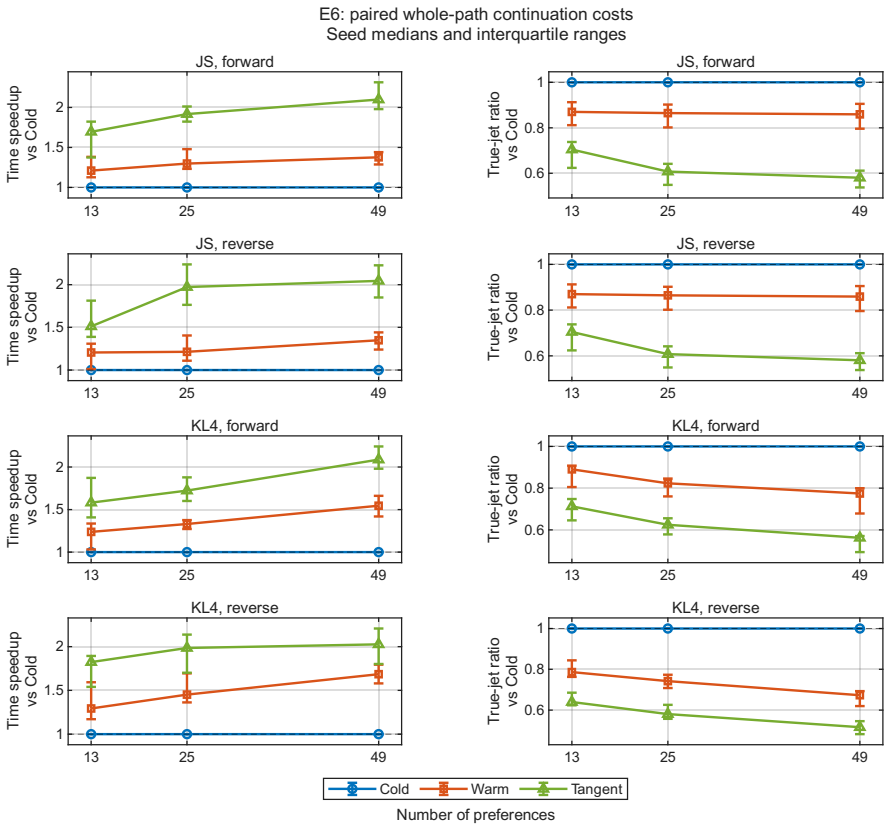

Figure 17: E6 whole-path costs by generator/orientation: runtime speedup against Cold (left) and true-jet pass ratio (right). Paired ratios use median times over three repetitions, then medians over traversals within seeds and medians/IQRs over ten seeds. Initialization and the common first solve are included.

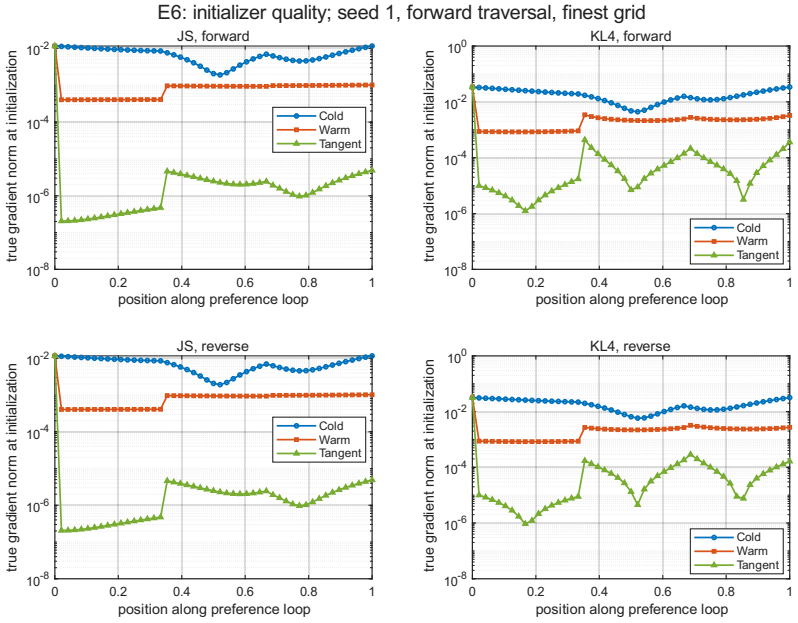

Figure 18: E6 pre-correction gradient norms, seed 1, forward traversal of the 49-point loop, both divergence orientations. All methods share the first cold initialization; the reference line is the corrector tolerance 10^{-8}.

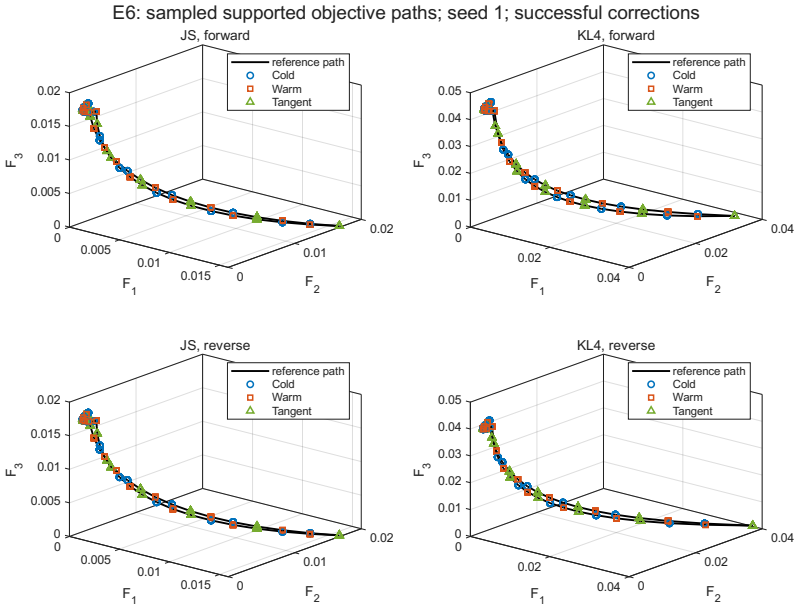

Figure 19: E6 corrected objective paths, seed 1, finest grid, forward traversal. The black line connects independent references; staggered markers show Cold, Warm, and Tangent solutions. All sampled points enter validation. The loop does not cover the full Pareto surface.

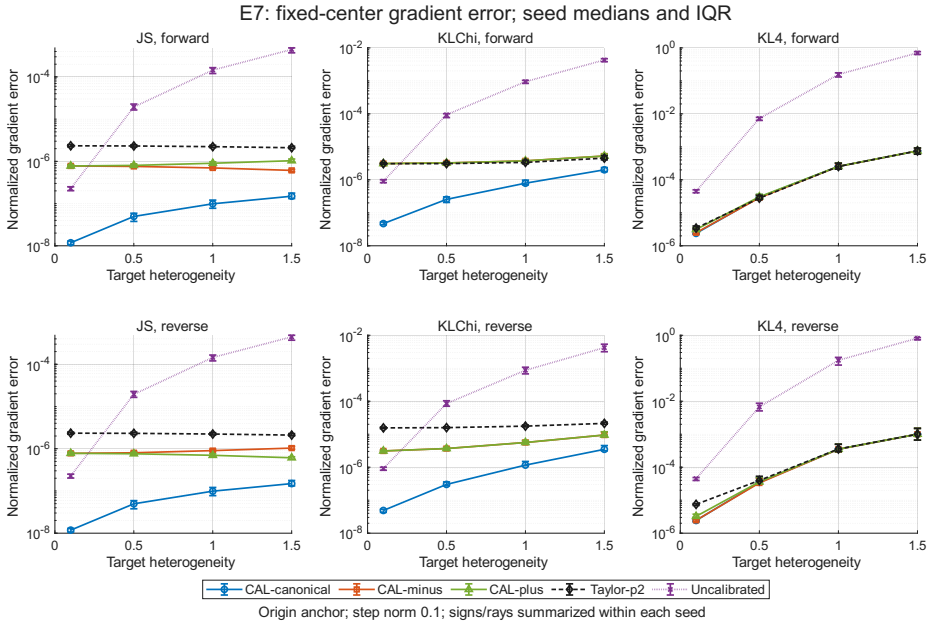

Figure 20: E7 normalized gradient errors (74), $\theta = 0$, $\|s\| = 0.1$: medians over two directions and both signs within seeds, then medians/IQRs across 20 seeds. CAL-minus/plus use $\alpha_f \mp 1$; all calibrated models and Taylor-p2 match through order two. JS orientations coincide.

The $p = 2$ specialization of (14) is

$$m^{(2)}_{\alpha,\theta}(s) = T_{2,\theta}F(s) + \mathcal{A}_\alpha(\theta + s) - T_{2,\theta}\mathcal{A}_\alpha(s). \tag{73}$$

The templates share F's quadratic expansion and differ in higher-order terms. Each satisfies the model-order bounds under residual smoothness; canonical matching does not guarantee minimal finite-step error.

We compare CAL-canonical with $\alpha = \alpha_f$, CAL-minus with $\alpha = \alpha_f - 1$, CAL-plus with $\alpha = \alpha_f + 1$, and Taylor-p2 with model $T_{2,\theta}F$. The fifth method uses the value-aligned uncalibrated model

$$m^{\mathrm{uncal}}_\theta(s) = F(\theta) + \mathcal{A}_{\alpha_f}(\theta + s) - \mathcal{A}_{\alpha_f}(\theta).$$

This control omits residual derivative corrections. All α predictors use exact aggregation with the required scale.

Problem construction. We use $Q_\theta = \mathrm{softmax}(b + B\theta)$ with $(K, d, m) = (128, 16, 32)$. The entries of b are independent $\mathcal{N}(0, 0.25^2)$ variables. We set $B = 1.2U$, where U is obtained by a thin QR factorization of a column-centered Gaussian matrix. Targets are

$$P_i(\tau) = \mathrm{softmax}(b + \tau B\xi_i + 0.1\tau z_i), \qquad \tau \in \{0.1, 0.5, 1, 1.5\},$$

where ξ_i are independent standard Gaussian vectors and each z_i is a coordinate-centered standard Gaussian vector. Preferences are sampled from $\mathrm{Dirichlet}(1, \ldots, 1)$. The common initial parameter is

$$\theta_{\mathrm{init}} = -0.75\tau\bar{\xi} + 0.8\max\{\tau, 0.25\}\zeta, \qquad \bar{\xi} = \frac{1}{m}\sum_i \xi_i, \quad \zeta \sim \mathcal{N}(0, I_d).$$

Twenty deterministic seeds are used, with MATLAB's Twister generator initialized by $71000 + \mathrm{seed}$. Primitive random draws are shared across heterogeneities, generators, orientations, and methods.

We evaluate JS, KLChi, and KL4 in both orientations, using the generator conventions stated at the beginning of Appendix F. The factorial design contains 480 cases and 2,400 primary method runs. JS forward and reverse coincide mathematically and are not independent replications of the underlying objective.

Common optimization protocol. All computations use double precision. Success requires $\|\nabla F(\theta)\| \le 10^{-6}$. Each run permits at most 360 trials and 120 accepted steps. The initial, maximum, and minimum trust-region radii are 0.5, 2, and 10^{-12}. The acceptance thresholds are $\eta_1 = 0.1$ and $\eta_2 = 0.75$. Rejected trials halve the radius; accepted trials with $\rho \ge \eta_2$ double it, subject to the maximum radius. Predicted reductions must satisfy the implemented Cauchy-decrease check with coefficient 0.1.

Every method uses the same generic inner routine: at most 15 regularized Newton iterations with projection onto the trust-region ball and an Armijo line search. Its model-gradient stopping threshold is

$$\max\{10^{-10},\, 0.01\,\|\nabla m(0)\|\}.$$

A Cauchy candidate safeguards model decrease. True-objective changes determine step acceptance, and true gradients determine success. The uncalibrated method is recorded as unsuccessful if it terminates at surrogate stationarity without meeting the true-gradient tolerance. Taylor-p2 uses this same inner routine; it is a controlled template baseline rather than a separately optimized quadratic trust-region implementation.

Independent reference runs use Taylor-p2 from $\theta = 0$, a true-gradient tolerance of 10^{-10}, and larger iteration budgets. All 480 references converge, with maximum gradient norm 9.71×10^{-11}. Their smallest observed Hessian eigenvalue is 3.95×10^{-3}. These are numerical reference solutions, not certificates of global optimality.

Fixed-center model diagnostics. Each case is probed at three centers: $\theta = 0$, the common initialization, and the independent reference parameter. At every center, we use two random unit directions, both signs, and step norms

$$h \in \{0.2, 0.1, 0.05, 0.02, 0.01, 0.005, 0.002, 0.001\}.$$

The normalized errors are

$$e_v(\theta, s) = \frac{|F(\theta + s) - m_\theta(s)|}{1 + |F(\theta)|},$$
$$e_g(\theta, s) = \frac{\|\nabla F(\theta + s) - \nabla_s m_\theta(s)\|}{1 + \|\nabla F(\theta + s)\|}. \tag{74}$$

Stable objective differences are used when evaluating small value errors. All 230,400 probe records are finite.

Figures 20 and 21 use the preselected origin and $h = 0.1$: medians over directions/signs within seeds, then medians and IQRs across seeds. IQRs measure dispersion, not confidence; fixed-step errors do not estimate convergence order.

Diagnostic fits use radii $\{0.1, 0.05, 0.02, 0.01\}$ and require all four errors above recorded roundoff thresholds. Of 57,600 value/gradient fits, 57,353 qualify; 247 remain unavailable in the records. Neither slope agreement nor R^2 determines eligibility.

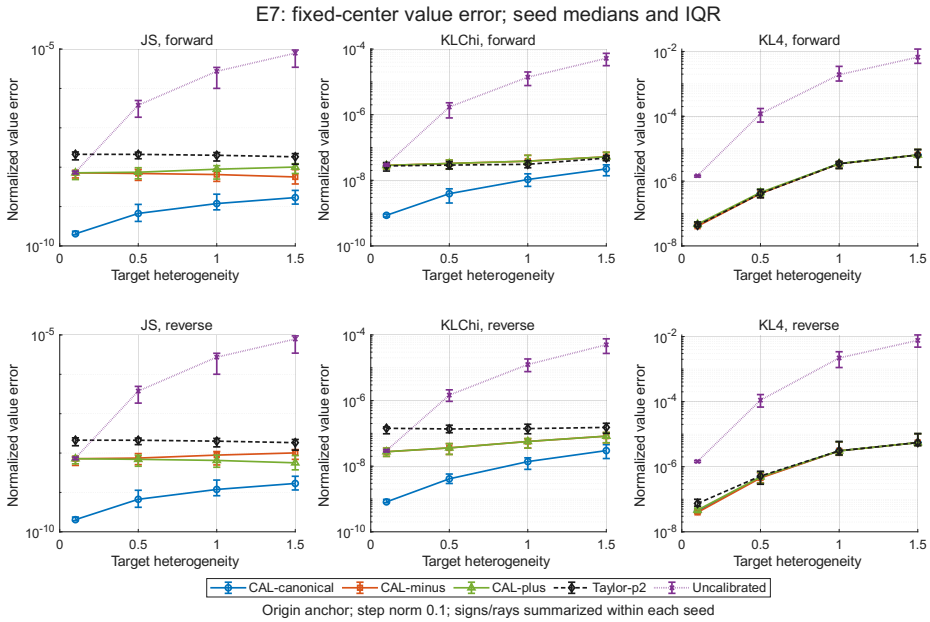

Figure 21: E7 normalized value errors with the probes and aggregation of Figure 20. Canonical median errors improve on JS and KL+chi-square; calibrated and Taylor curves nearly coincide for KL+quartic at larger heterogeneities. The horizontal axis varies targets, not step size.

Convergence and timing results. Each of CAL-canonical, CAL-minus, CAL-plus, and Taylor-p2 succeeds in all 480 cases. Their largest final gradient norm is 9.68×10^{-7}, and their largest absolute objective discrepancy from the numerical reference is 7.28×10^{-11}. None of these 1,920 runs fails the Cauchy-decrease check. Uncalibrated succeeds in 40/160 JS, 40/160 KL+chi-square, and 0/160 KL+quartic cases, only at $\tau = 0.1$. Denominators include every attempt.

After untimed warm-up, three repetitions per case/method in randomized order give 7,200 timed runs with identical numerical outcomes and work counts. For method M, the paired speedup uses median times:

$$S_{\text{time}}(M) = \frac{\text{median}_r\, t_{\text{Taylor},r}}{\text{median}_r\, t_{M,r}}.$$

A timing pair is eligible only if both methods succeed and their final objectives satisfy

$$|F_M - F_{\text{Taylor}}| \le 10^{-8}(1 + \max\{|F_M|, |F_{\text{Taylor}}|\}).$$

Eligible ratios are summarized by medians/IQRs across 20 seeds; repetitions are not independent instances, and groups without pairs are unplotted.

Timings include aggregation, true-jet evaluation, model construction and calibration, inner solves, and true-objective acceptance checks. Independent references and offline model diagnostics are

excluded. The uncalibrated control retains the common true-jet evaluation pipeline, so its timing is not a benchmark of the cheapest possible standalone predictor solver.

CAL-canonical's group-median speedups are 0.470–0.816 (Figure 22): better model accuracy does not yield lower runtime here. Figure 23 gives representative gradient histories; success statistics use all seeds.

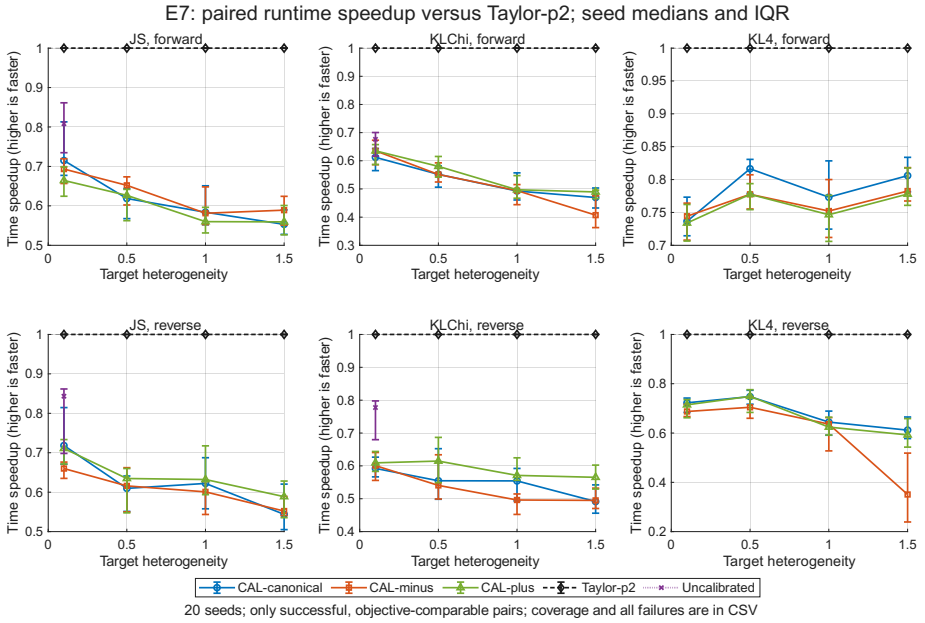

Figure 22: E7 paired speedups against Taylor-p2 (above one is faster), using median times over three repetitions and seed medians/IQRs. Each calibrated group has 20 successful, objective-comparable pairs. Uncalibrated pairs exist only for JS and KL+chi-square at $\tau = 0.1$; missing points indicate no eligible pairs.

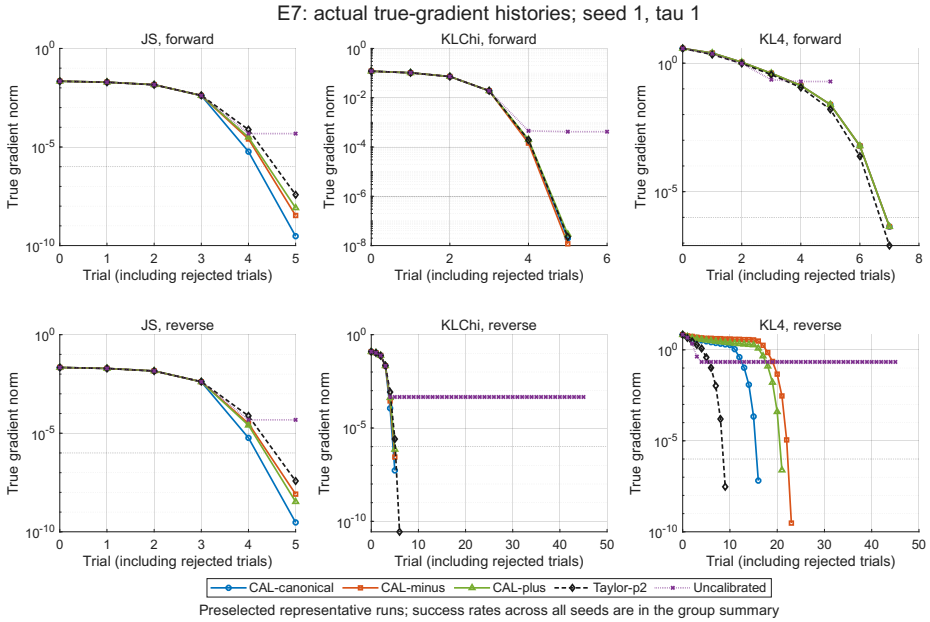

Figure 23: E7 true-gradient histories, preselected seed 1 and $\tau = 1$, including rejected trials. These representative histories do not measure asymptotic local order or replace all-seed success statistics.

F.8 E8: LEARNED LATENT DISTRIBUTION MATCHING AND GENERATION

Why the matching problem remains nontrivial. Each P_i contains six Gaussian components, whereas Q_θ has only two and a fixed shared covariance. Even for forward KL, whose unrestricted

solution is $\bar{P}_\lambda = \sum_i \lambda_i P_i$, that mixture need not be representable by Q_θ. For example, preference $(.90, .05, .05)$ produces the KL target mixture $.645R_0 + .220R_1 + .135R_2$. These are weights of fitted latent laws, not prescribed image-class frequencies. The restricted fit, divergence choice, and decoding can each affect the observed proportions. E8 tests restricted-family optimization with a fixed learned representation.

Data and representation. We use the MNIST digits 0, 1, and 2, downsampled from 28×28 to 14×14 by averaging each 2×2 block and scaled to $[0, 1]$. Each of five seeds selects 1,200 training and 180 validation images per class from the original training split, and 500 held-out images per class from the original test split. An autoencoder with widths 196–64–2–64–196 uses tanh hidden activations, a linear bottleneck, and a sigmoid output. We minimize pixelwise binary cross-entropy plus a 2×10^{-4} latent squared-norm penalty with Adam, learning rate 10^{-3} and batch size 128, for at most 120 epochs. Validation early stopping has patience 15 and restores the best checkpoint. Training-encoding statistics standardize the latent coordinates. The encoder, decoder, and standardization are then frozen for all optimization methods and preferences within that seed.

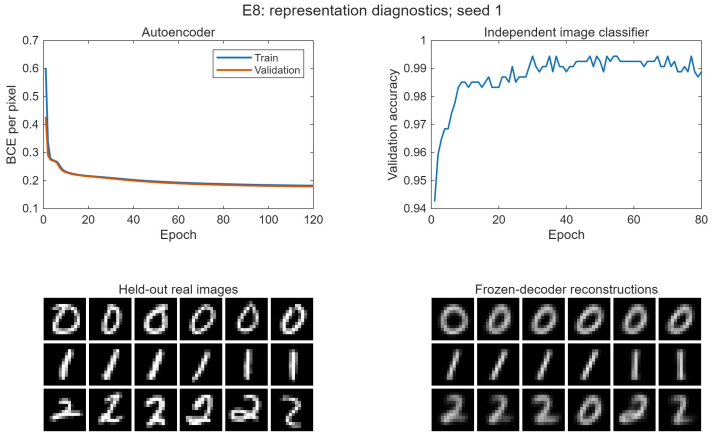

Figure 24: E8 representation diagnostics for seed 1: training and validation reconstruction loss, validation accuracy of an independent image classifier, and held-out inputs with frozen-decoder reconstructions.

Targets and optimized model. For each class, a two-component diagonal-covariance Gaussian mixture R_k is fitted to training encodings with three EM starts, at most 80 iterations, and standard-deviation floor 0.35. The targets are

$$\begin{aligned} P_1 &= .7R_0 + .2R_1 + .1R_2, \\ P_2 &= .1R_0 + .7R_1 + .2R_2, \\ P_3 &= .2R_0 + .1R_1 + .7R_2. \end{aligned} \tag{75}$$

The optimized latent distribution Q_θ has two Gaussian components with a common fixed isotropic standard deviation equal to the larger of 0.5 and the median fitted teacher standard deviation. Its five free parameters are one mixture-weight logit and four mean coordinates; means use a bounded tanh parameterization. All methods use the same initial parameter vector. There is no continuation warm start between preferences.

The ten preferences are the order-three simplex lattice transformed by $\lambda = .05\mathbf{1} + .85r$, where r ranges over nonnegative triples with sum one and coordinates in $\{0, 1/3, 2/3, 1\}$. This includes the uniform preference and the three $(.90, .05, .05)$ permutations. For JS and forward KL, we minimize $\sum_i \lambda_i D_f(P_i \| Q_\theta)$. With $a_f = 1/2$ and $\alpha_f = 0$, the normalized-grid JS template can equivalently be evaluated as

$$\mathcal{A}(\theta) = 2\left(1 - \sum_x H_\lambda(x)\sqrt{q_\theta(x)}\right), \qquad H_\lambda(x) = \sum_i \lambda_i \sqrt{p_i(x)}, \tag{76}$$

where p_i and q_θ denote probability masses on the quadrature grid. This expression retains the scale and constant in Theorem 3.2.

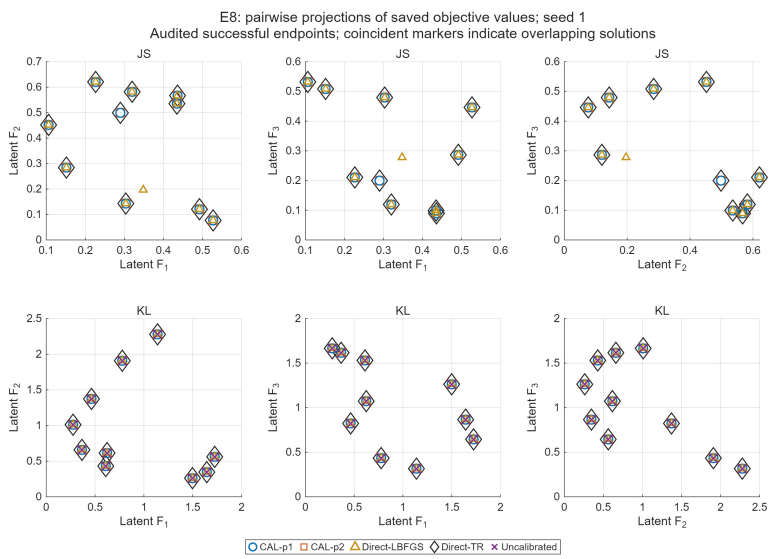

Figure 25: E8 pairwise projections of latent objectives at audited successful endpoints, seed 1. Coincident markers indicate overlapping solutions. The points are sampled stationary trade-offs, not a certified full Pareto front.

Solvers, budgets, and timing. CAL-p1 and CAL-p2 use residual derivatives through orders one and two, respectively; the p1 implementation can still use the predictor Hessian inside its subproblem. Direct-LBFGS uses memory eight and Armijo backtracking. Direct-TR uses the true quadratic model and an eigendecomposition-based trust-region step. The uncalibrated control aligns the predictor value but omits residual derivative corrections. All methods use the same true-objective stopping tolerance 10^{-6}, a budget of 220 trial iterations and 100 accepted steps, initial radius 0.5, maximum radius 2, and minimum radius 10^{-11}. CAL inner solves use at most 12 iterations, relative gradient tolerance 0.01, and absolute floor 10^{-9}. The design comprises 500 primary runs (five seeds, ten preferences, two generators, five methods), with three timing repetitions in rotated method order. Four independent reference starts per seed/preference/generator yield 400 reference solves, excluded from optimization timing and never used for initialization. References are local best-found solutions, not certified global optima.

Quadrature and endpoint audits. Optimization uses a fixed 97×97 midpoint grid and normalized probability masses. The integration box is chosen using a 10^{-10} Gaussian tail-mass threshold. Every endpoint is re-evaluated on independently refined 193×193 and 385×385 grids. Audit thresholds are 5×10^{-6} for the refined gradient norm, 10^{-7} for objective discrepancies, 5×10^{-6} for gradient discrepancies, and 10^{-7} for discrete-mass normalization error. The maximum observed objective and gradient discrepancies are 1.85594×10^{-8} and 1.13101×10^{-7}; maximum normalization error is 8.88×10^{-16}. Grid refinement checks numerical consistency, not rigorous continuous-objective error bounds; mass normalization alone does not certify quadrature accuracy.

Generation and independent image diagnostics. We sample $z \sim Q_\theta$ and display the frozen decoder's mean image, without adding Bernoulli pixel noise. Image-space diagnostics do not enter the optimized latent divergence. An independent linear softmax classifier is trained on training images for at most 80 epochs with learning rate 0.005. Its minimum held-out accuracy across seeds is 98.27%; the minimum classification accuracy of reconstructed held-out images is 97.2%. Unbiased image MMD^2 is computed in a 24-dimensional PCA representation learned from training images, using a Gaussian kernel with training-data bandwidth. Each image diagnostic uses 256 latent draws. Class proportions and MMD^2 are summarized across seeds separately for each categorical preference ID. Failed endpoints remain in these descriptive diagnostics and are identified in the saved samples.

Results and interpretation. All 200 calibrated runs and all 200 direct runs attain audited stationarity; the uncalibrated predictor succeeds on all 50 KL cases but no JS case. Comparable-pair coverage against Direct-TR is 43/50 for JS CAL-p1 and 46/50 for JS CAL-p2, and 50/50 for both

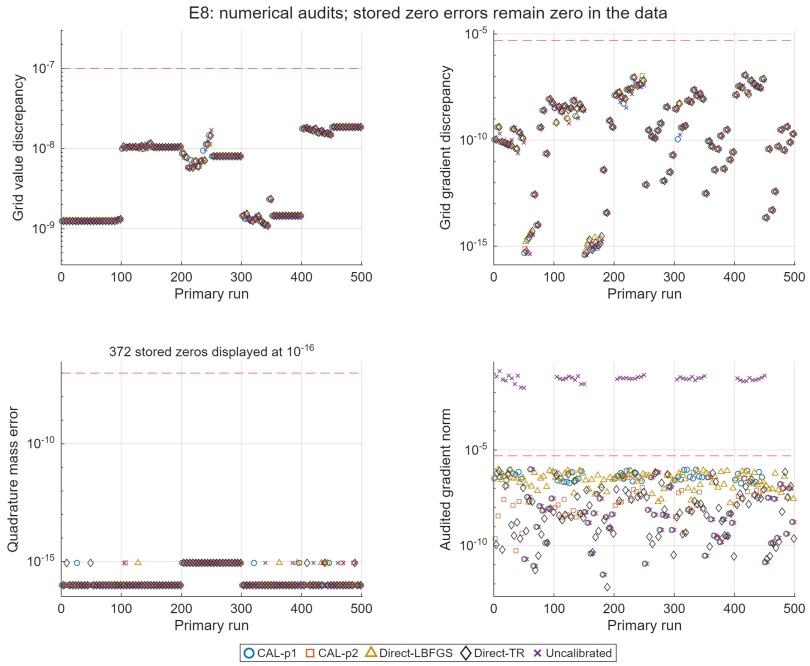

Figure 26: E8 quadrature and endpoint audits. Dashed lines mark the configured thresholds. The mass panel reports normalization error of the discrete masses; stored zeros are displayed at 10^{-16} without changing the data. Large true gradients of uncalibrated JS endpoints explain their failed audits.

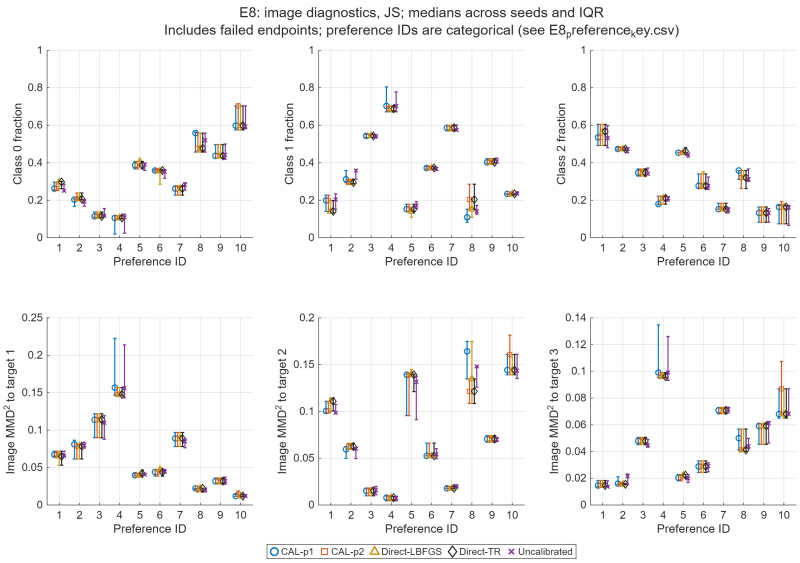

Figure 27: E8 held-out image diagnostics: class fractions and image MMD^2 under JS, summarized by seed medians and IQRs. Preference IDs are categorical and follow the saved preference key. Failed endpoints are included; these quantities describe decoded distributions rather than certify optimality.

KL variants. Median paired speedups against Direct-TR are 0.108 and 0.131 for JS CAL-p1/p2, and 0.178 and 0.155 for KL: the three-target application establishes reliable matching and preference-controlled generation, not acceleration. Signed reference gaps retain both improvements and worse local solutions.

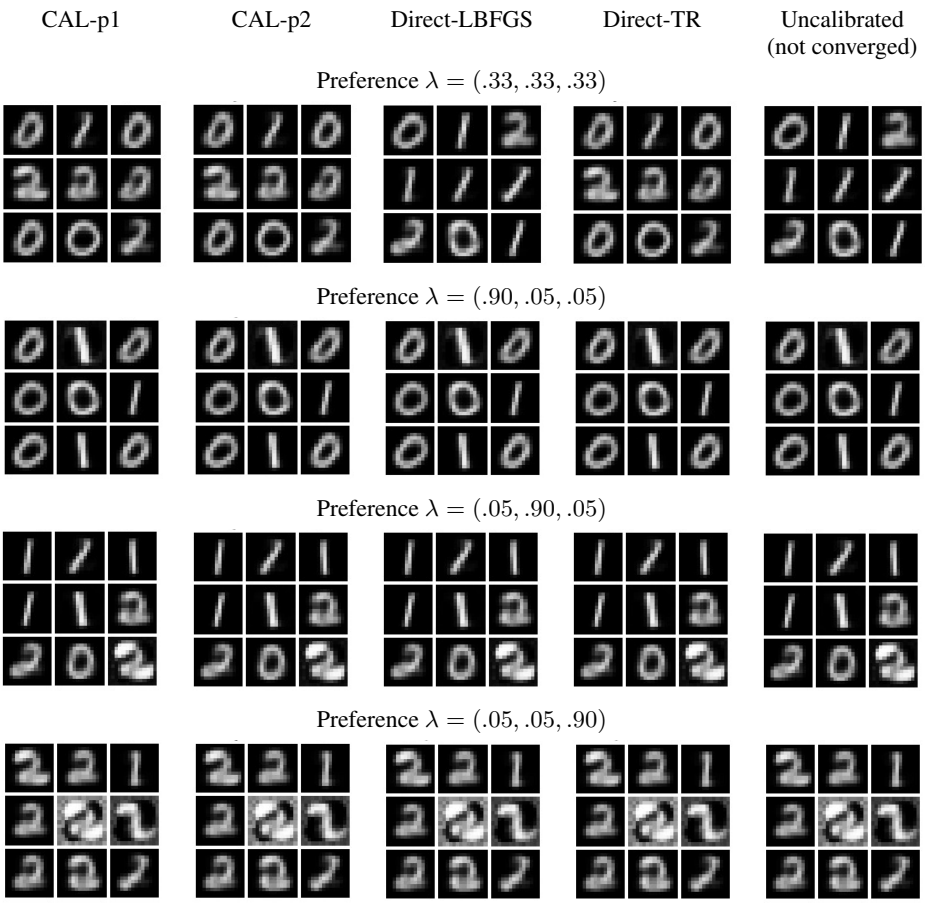

Figure 28: E8 saved decoded samples, four JS preferences, seed 1. Columns compare methods; all displayed uncalibrated endpoints failed to converge. Mosaics retain the saved samples, cropping only surrounding whitespace; CAL-p2 also appears in Figure 3.

F.9 E9: MANY-TARGET COST CROSSOVER AND IMPLEMENTATION CONTROLS

Nested target banks. E9 reuses each E8 seed's three fitted latent targets and common initial model, then creates a bank of 300 six-component Gaussian mixtures. The first three are the original targets. Additional targets cycle through the three families, perturb component means by Gaussian noise of scale 0.15 times the student standard deviation, multiply component standard deviations by $\exp(.075 \tanh Z)$, and multiply mixture weights by $\exp(.15 \tanh Z)$ followed by renormalization, with independent standard-normal Z. The bank seed is $91000+$ the E8 seed. Target counts $3, 10, 30, 100, 300$ use nested prefixes of the same bank. The common integration box is fixed using the full 300-target bank, extending to nine component standard deviations beyond the extreme means when necessary. These are fitted and perturbed latent densities, not 300 separately trained neural networks.

The correlated bank isolates target-count effects within three families; its speedups do not establish performance for independent, highly heterogeneous teachers. E7 tests heterogeneity effects on model accuracy, not this timing comparison.

For family $g(j) = 1 + \mathrm{mod}(j-1, 3)$ and its count n_g in the chosen prefix, target weights are $\lambda_j = \pi_{g(j)}/n_{g(j)}$. The five family preferences are uniform, the three permutations of $(.8, .1, .1)$, and $(.5, .35, .15)$. This keeps family-level preferences comparable as m grows. The student remains

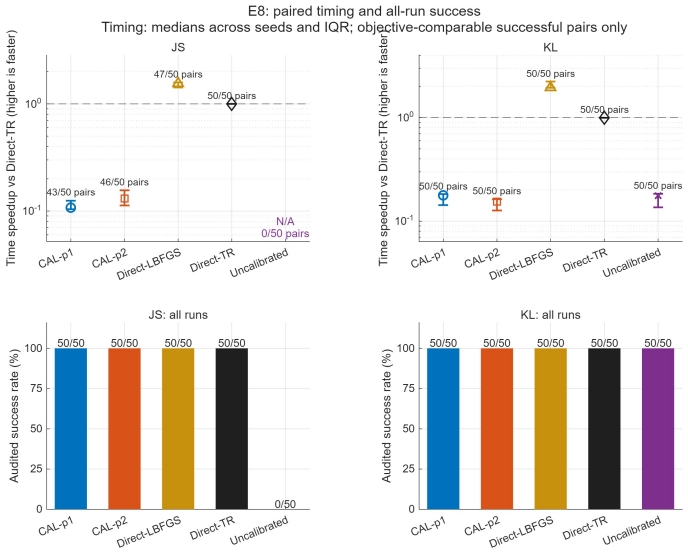

Figure 29: E8 paired runtime and all-run audited success. Timing uses only successful, objective-comparable pairs and reports pair counts. Success uses all attempted cases; N/A denotes absence of eligible timing pairs.

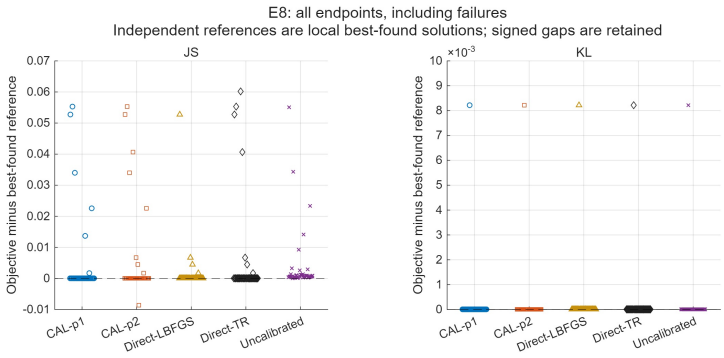

Figure 30: E8 signed objective gaps to independent local best-found references, including failed endpoints. A negative value means that the tested run improved on its reference.

the same five-parameter, two-component mixture used in E8. Both forward JS and forward KL are tested.

Shared implementations and cost boundaries. All solvers share cached teacher masses, log masses, and entropies on a 97×97 grid, block-vectorized evaluation with block size 32, and a per-solve cache for the current student density. A cached higher-order evaluation can satisfy a lower-order query. Direct KL objectives use exact aggregation, just as the KL template does. Experiments run in double-precision MATLAB R2025b Update 5 on 64-bit Windows; comparisons are paired on the same execution environment.

Total solve time includes preference-dependent setup, predictor aggregation, calibration and refreshes, inner solves, true-objective trial tests, and stopping checks. Shared bank/grid preparation, representation training, references, audits, profiling, and microbenchmarks are excluded equally; E9 measures optimization of fitted densities.

What the baselines isolate. CAL-p2 and Direct-TR use true values, gradients, and Hessians; L-BFGS uses values and gradients. All share the scalarization and stationarity test. FAMO's adaptive weighting (Liu et al., 2023) and FERERO's relative/constrained preferences (Chen et al., 2024) change the preference specification. Exact KL/inverse-KL Pareto constructions for full exponential families (Zhang et al., 2025) do not directly solve restricted-mixture JS matching; our KL baselines

already aggregate exactly. Target-subsampled stochastic solvers are untested and would require consistent sampling costs and full-objective audits.

Stage A: derivative-demand implementation ablation. For the original three-target problem, we compare eager template evaluation (computing a Hessian at every template query) with demand-driven evaluation that computes only the requested derivative order. Optimization rules, tolerances, and initialization are held fixed. The ablation uses five seeds, two representative preferences, and CAL-p1/p2. Outer acceptance decisions agree, and endpoint parameter/objective differences pass tolerances $10^{-8}/10^{-10}$. Figure 31 shows the timing and template-Hessian savings.

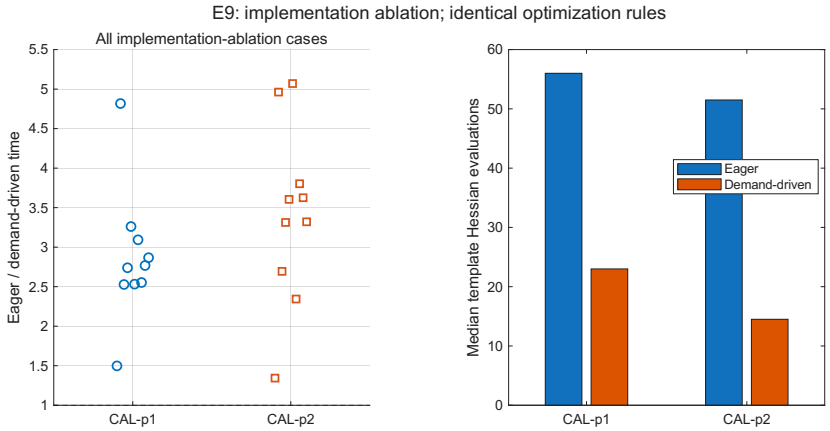

Figure 31: E9 implementation ablation with identical optimization rules. Left: eager/demand-driven time ratios for CAL-p1/p2. Right: median template-Hessian counts. This compares two implementations of the same method, not CAL against a direct optimizer.

Stage B: matched-order microbenchmarks. We separately time the true objective and aggregated predictor at three parameter points, with three repetitions. F returns a value, FG a value and gradient, and FGH a value, gradient, and Hessian. A timing batch targets at least 0.04 seconds, with a 500-call cap; evaluation batches reach the requested duration. Aggregation preparation is timed separately, and is already charged to full solves. At $m = 300$, JS true/predictor cost ratios are 188.788, 68.120, and 8.786 for F, FG, and FGH; KL ratios remain approximately one. See Figures 32 and 2.

Stage C: optimization, references, and audits. Five seeds, five target counts, five preferences, two generators, and four methods yield 1,000 primary runs and 3,000 timed repetitions. Before measured repeats, each method performs a warm-up solve limited to two trial iterations; the measured method order is rotated. The methods are CAL-p1, CAL-p2, Direct-LBFGS, and Direct-TR. CAL-p1 refreshes true FG information, CAL-p2 and Direct-TR use true FGH information, and L-BFGS uses true FG information; the CAL inner solver may request predictor Hessians for either calibration order. Independent local references use three Direct-TR starts and true-gradient tolerance 10^{-7}, without initializing any compared run.

Primary solves use the E8 limits of 220 trials and 100 accepted steps, radii $(\Delta_0, \Delta_{\max}, \Delta_{\min}) = (0.5, 2, 10^{-11})$, and at most 12 inner iterations with relative tolerance 0.01 and absolute floor 10^{-9}. They stop at true-gradient norm 10^{-6}; endpoint audits require at most 2×10^{-6}. All 1,000 pass, with maximum audited norm 9.948×10^{-7}. All 140 formula and derivative checks pass. Forty selected endpoints (seed 1, first preference, all target counts, both generators, all methods) are independently re-evaluated on a 193×193 grid; all satisfy objective-discrepancy tolerance 10^{-7} and gradient-discrepancy tolerance 5×10^{-6}. Saved validation checks also cover case uniqueness, repetition counts, independent references, and replay consistency.

Paired summaries and local-solution coverage. Let $t^M_{s,k}$ be the median of three unprofiled times for method M, seed s, and preference k at a fixed generator and target count. For a baseline B, a pair is eligible if both runs pass the endpoint audit and

$$\left|F_\lambda(\theta^M_{s,k}) - F_\lambda(\theta^B_{s,k})\right| \le 2 \times 10^{-6}. \tag{77}$$

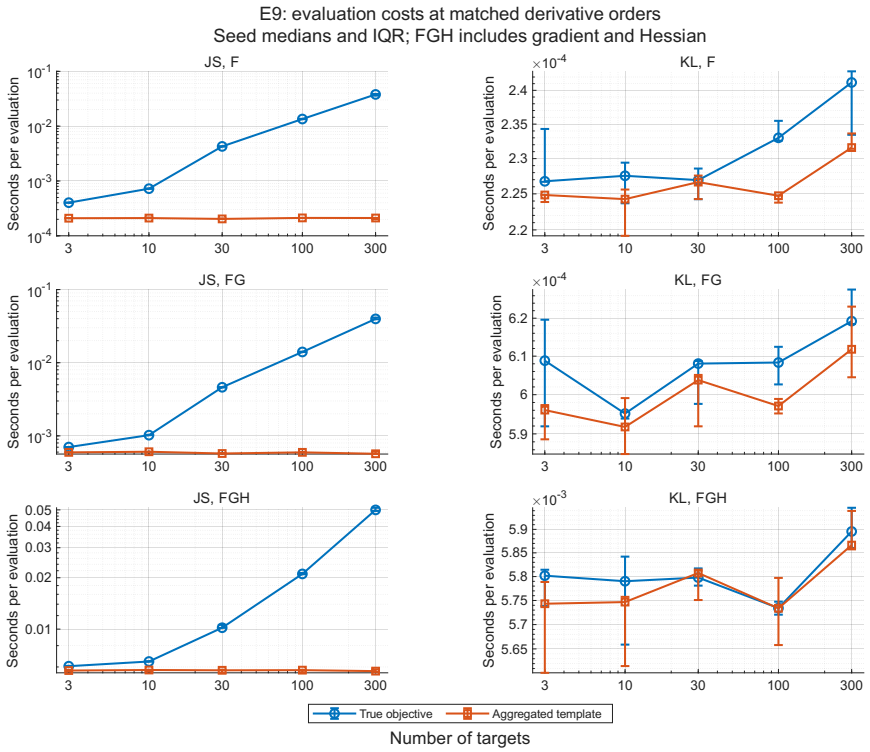

Figure 32: E9 absolute evaluation costs at matched derivative orders, summarized by seed medians and IQRs. FGH includes both gradients and Hessians; predictor preprocessing is accounted for separately and included in total solves.

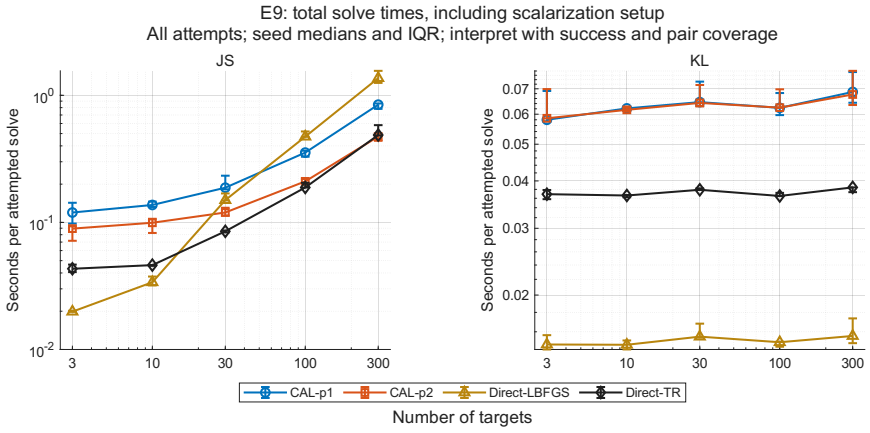

Figure 33: E9 total solve times including preference-dependent scalarization setup, over all attempts. Because different methods can reach different local solutions, these marginal medians supplement rather than replace paired speedups.

Take median paired ratios $t^B_{s,k}/t^M_{s,k}$ over eligible preferences within each seed, then the median/IQR across five seeds. IQRs measure dispersion, not confidence. Report pair coverage separately from stationarity; retain every endpoint in raw tables and signed-gap plots.

At $m = 300$, JS CAL-p2 has 23/25 eligible pairs against L-BFGS and 24/25 against Direct-TR. Its median speedups are 3.1498 (IQR 3.0587–3.1510) and 1.1026 (IQR 1.0968–1.1513). Against Direct-TR, 17/24 individual pairs and four of five seed medians are faster. The two excluded L-BFGS pairs have higher CAL-p2 objectives, with gaps approximately 0.0262 and 0.0184; they are not evidence of improved solution quality. Direct methods remain faster in the KL exact-aggregation control. These controls delimit the cost regime in which JS aggregation yields a net benefit.

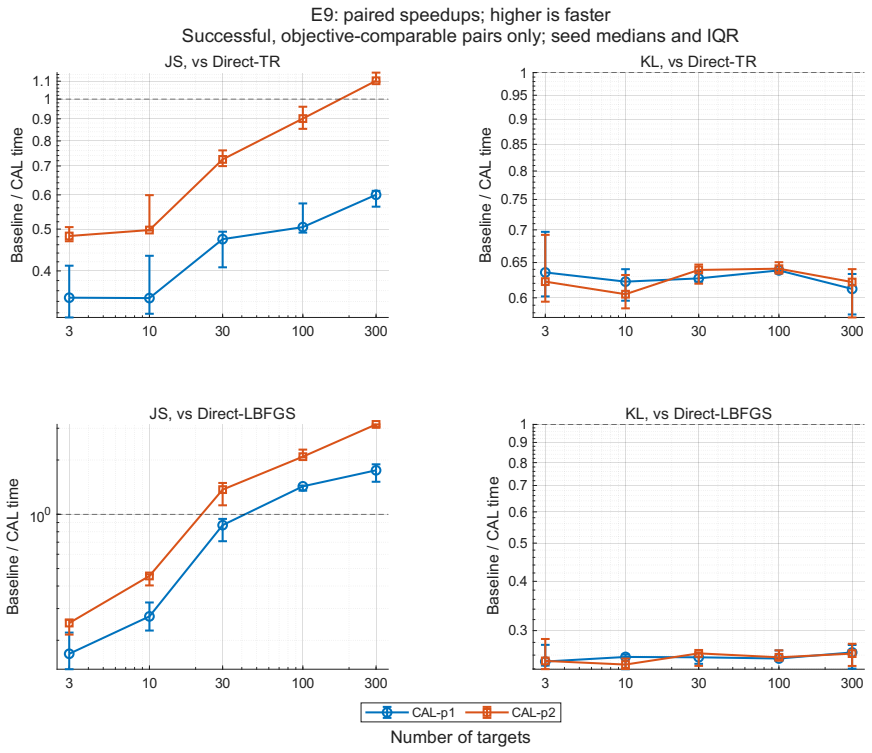

Figure 34: E9 full paired speedups for JS and the exact-aggregation KL control, against Direct-TR and Direct-LBFGS. Curves summarize successful objective-comparable pairs by seed medians and IQRs. CAL-p2 JS pair coverage is given in Table 2; raw records retain every attempted case.

Diagnostic cost breakdown. Separate instrumented replays measure preparation, true/template evaluations, linear algebra, and other work (Figure 35). Component medians need not sum to median runtime; speedups use unprofiled paired times. At $m = 300$ on JS, the median CAL-p2 counts are five true FGH calls, four additional true F calls, 56 template F calls, and 14 template Hessian calls, compared with six true FGH and five additional F calls for Direct-TR. The resulting gain combines lower true-evaluation demand with inexpensive repeated template work.

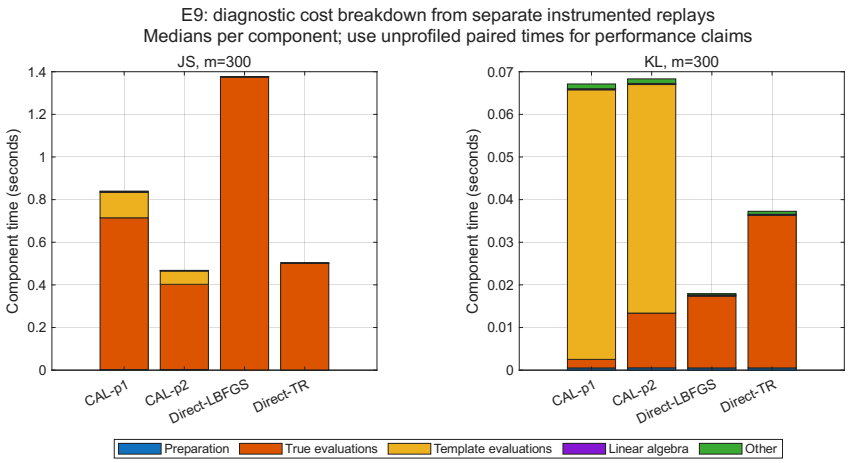

Figure 35: E9 diagnostic cost breakdown at $m = 300$ from separate instrumented replays. Each component is summarized by its median. Performance claims use the unprofiled paired timing experiment.

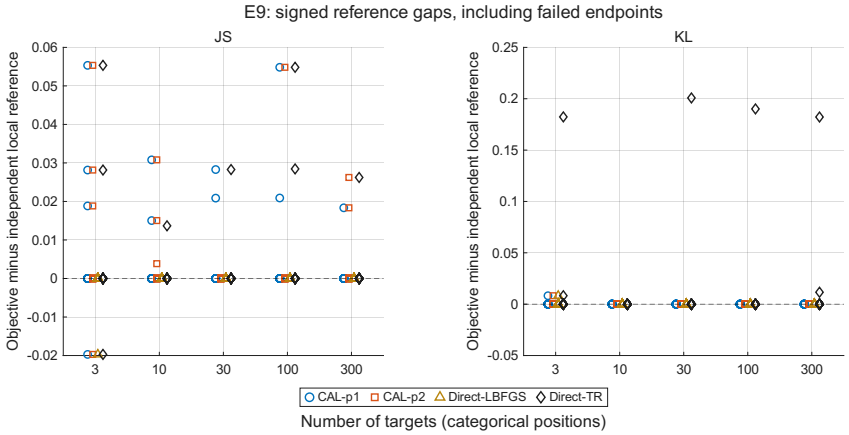

Figure 36: E9 signed objective gaps relative to independent local references, retaining all endpoints. Negative gaps are possible because references are best-found local solutions. These plots report solution variation separately from stationarity success and timing-pair coverage.

F.10 E10: CANONICAL GEOMETRY AND THE LEADING LOCAL ERROR

E10 tests the leading-error cancellation in Proposition 5.4 at fixed $p = 2$, complementing E5's generic local orders and E7's finite-step errors. Third derivatives are offline diagnostics; all optimizers use only values, gradients, and Hessians.

Problems and target heterogeneity. Use $K = 6$ categories, $d = 2$ parameters, $m = 4$ targets, seeds $s = 1, \ldots, 5$, and MATLAB's Twister generator with seed $27100 + s$. Draw $b \sim \mathcal{N}(0, 0.35^2 I)$, a standard Gaussian $K \times d$ matrix G_0, target perturbations $\xi_i \sim \mathcal{N}(0, 0.7^2 I)$, positive raw preferences $w_i = 0.25 + U_i$ with $U_i \sim U(0, 1)$, and one standard Gaussian direction. Set $q_0 = \mathrm{softmax}(b)$, center G_0 by subtracting $\mathbf{1}(q_0^\top G_0)$, and rescale it to a matrix G whose largest row norm is 0.4. The affine family and preferences are

$$q_\theta = q_0 + B\theta, \qquad B = \mathrm{diag}(q_0)G, \qquad \Theta = \{\theta : q_\theta > 0\}, \qquad \lambda_i = w_i / \sum_j w_j.$$

The laws are normalized since $\mathbf{1}^\top B = 0$. For $\eta \in \{0, 0.01, 0.05, 0.2\}$, set $P_i(\eta) = \mathrm{softmax}(\log q_0 + \eta \xi_i)$, explicitly assigning $P_i(0) = q_0$. Forward JS and KL4 use the generator conventions at the start of this appendix. Thus $\eta = 0$ is common matching; positive η perturbs the targets using shared random draws.

Comparators and common inner accuracy. CAL-canonical uses α_f; CAL-alpha-minus/plus use $\alpha_f - 1/\alpha_f + 1$, with the same scale a_f and true residual value, gradient, and Hessian correction.

Taylor-p2 uses the true quadratic Taylor model. All four models therefore agree through order two at every anchor. All methods impose

$$\|\nabla m_\theta(s)\| \le 0.01 \|\nabla F(\theta)\|^3 . \tag{78}$$

This strengthens Theorem 5.3's $p = 2$ forcing to reveal the coefficient in (25). Newton iterations on the model start from zero, with at most 30 iterations and 45 backtracking halvings, retaining feasible steps that reduce the model-gradient norm. The trial radius is $0.45 \min_k q_\theta(k)/\|B_{k,:}\|$. True-objective acceptance requires positive actual decrease, $\rho \ge 0.1$, and the Cauchy decrease test with constant 0.1 and a ballwise model-Hessian bound. No reference solution or third derivative is passed to the step solver.

Precision, references, and audits. The completed `paper` run uses MATLAB Symbolic Math Toolbox variable-precision arithmetic at 120 digits. An independent true-objective Newton solve from zero computes each local reference $\theta^\star$ to gradient tolerance 10^{-90} and checks positive definiteness of its Hessian. References are used only to position and audit local probes. Across 40 problems, the reference gradient norms are below 1.10×10^{-92} and the smallest reference Hessian eigenvalue exceeds 2.44×10^{-3}. Centered finite differences with step 10^{-20} independently check analytic gradients, Hessians, and third derivatives for the true objective and three predictors. All 160 checks pass tolerance 10^{-30}; the maximum normalized third-derivative discrepancy is 1.21×10^{-41}. Direct and aggregated predictor evaluations also agree, with maximum normalized discrepancy 7.12×10^{-128}. The inner arithmetic floor is 10^{-105}; runs needing smaller forcing tolerances are flagged rather than counted as successful.

Local maps and prespecified summaries. For each seed, use a unit direction d_s and the scale $r_s = \min\{1, 0.2 \min_k q_{\theta^\star}(k)/\|B_{k,:}\|\}$, initialize at $\theta^\star \pm r_s r d_s$, with $r \in \{10^{-2}, 10^{-3}, \dots, 10^{-7}\}$. The two signs share one random direction and are not independent seeds. The design gives $5 \times 2 \times 4 \times 4 \times 2 \times 6 = 1{,}920$ one-step maps. An eligible map must be accepted, meet (78), pass the Cauchy test, avoid precision and boundary flags, and contract its parameter error to a value above the reference error floor 10^{-80}. All 1,920 maps qualify; every attempt is retained, independently of fitted order.

Fit log error against log initial radius on the fixed relative window $10^{-7} \le r \le 10^{-3}$, using five points per seed and sign. Summarize signs within each seed and then report medians and IQRs across five seeds. The same procedure fits the expansion remainder

$$\|e_+ - b(e)\|, \qquad b(e) = \tfrac{1}{2} H_\star^{-1} \nabla^3 \mathcal{E}(\theta^\star)[e, e].$$

For Taylor-p2, use $\mathcal{E} = F$. Figures 37 and 38 show the raw scaling and remainder audit.

Cancellation, error constants, and limits. At $\eta = 0$, the median canonical order is 3.000 for both generators, whereas shifted templates and Taylor-p2 have order 2.000. The computed canonical residual third tensor has norm below 4×10^{-130}, consistent with the exact cancellation. At positive η, all group median orders are within 0.001 of two. Canonical one-step errors remain below Taylor-p2 in all 360 paired positive-heterogeneity probes. Table 9 reports directional error-constant ratios, not speedups or a uniform ordering over directions. All 320 individual expansion-remainder fits lie between 2.99995 and 3.00005, supporting the vector coefficient, not merely a fitted order (Figure 38).

Table 9: E10: leading directional error ratio $\|b_{\mathrm{Taylor}}(d_s)\| / \|b_{\mathrm{CAL}}(d_s)\|$, median over five seeds. Values above one favor canonical CAL-p2. The denominator vanishes theoretically at $\eta = 0$.

η	JS	KL4
0.01	12,793.3	177.6
0.05	2,369.0	31.8
0.2	514.5	2.77

Canonical α is not universally optimal: for KL4 at $\eta = 0.2$, the negative shift has a smaller residual tensor norm in all five seeds. Common matching is a nonconflicting cancellation limit; positive-η cases test departures from it.

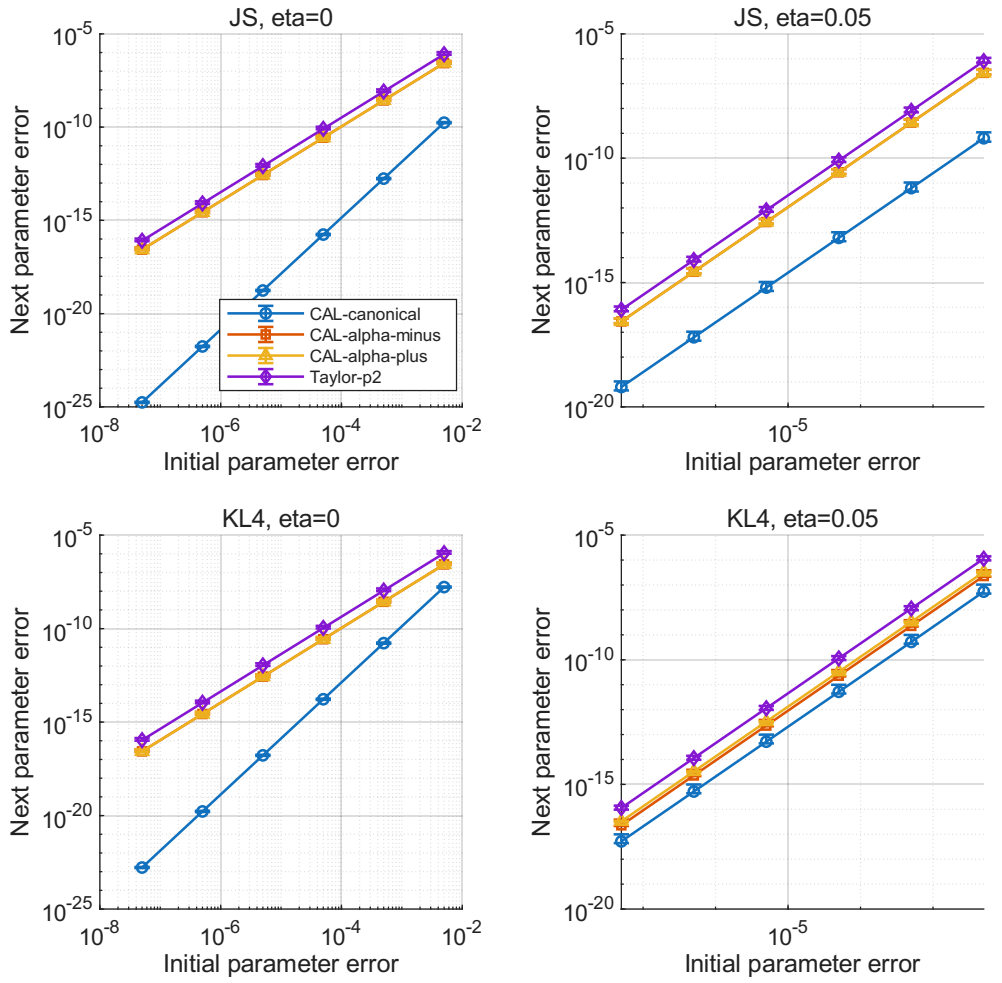

Figure 37: E10 one-step errors, seed medians/IQRs, with identical $p = 2$ calibration and cubic inner forcing. Canonical CAL-p2 is cubic at $\eta = 0$ and quadratic for positive heterogeneity. Panels show $\eta = 0, 0.05$ and all six radii; fits and audits use all four heterogeneities.

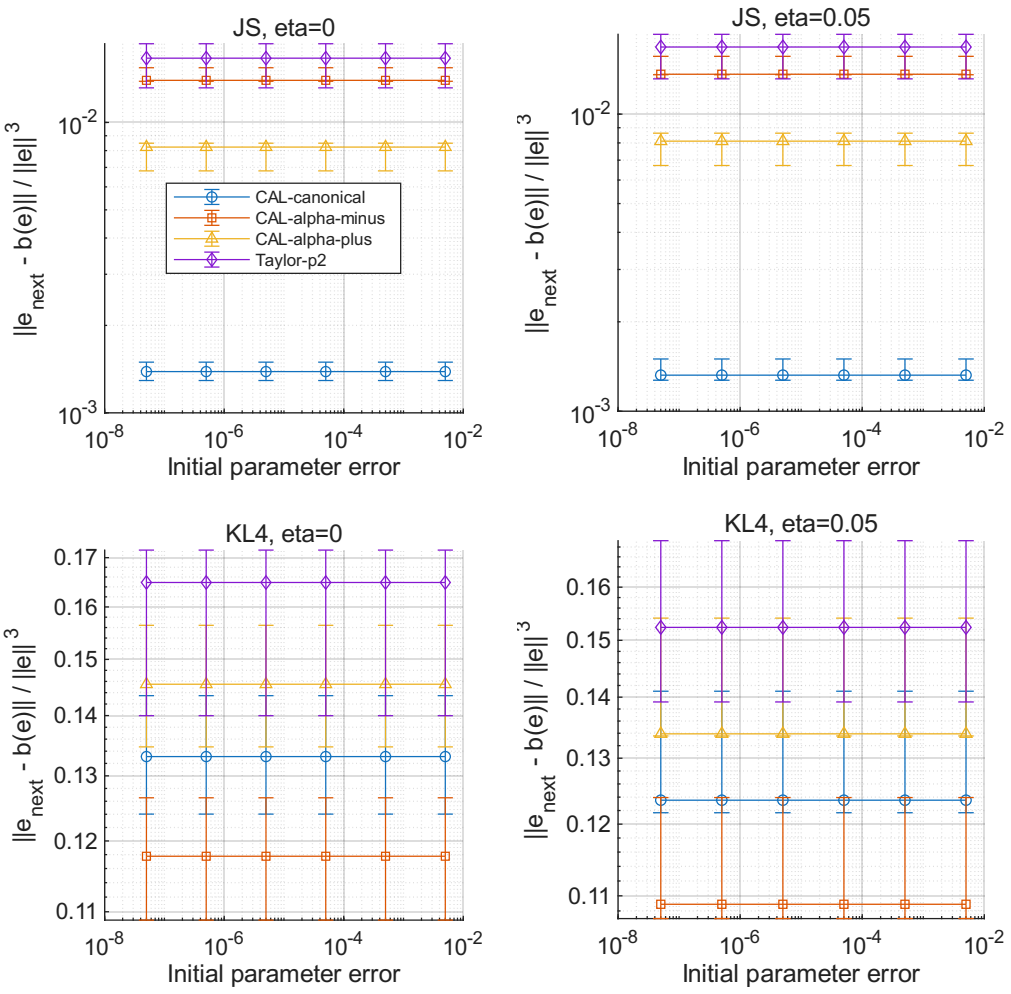

Figure 38: E10 expansion audit: subtracting the predicted quadratic vector leaves an approximately constant $\|e_+ - b(e)\| \,/\, \|e\|^3$. Panels show $\eta = 0, 0.05$; fits cover all four heterogeneities. Third derivatives enter only the offline prediction and audit.

Actual trajectories and interpretation. Separate trajectories begin at $\theta^\star \pm 0.01 r_s d_s$ and recalibrate after every accepted step, stopping at true-gradient norm 10^{-30} or ten steps. All 320 trajectories reach that tolerance; the stored 1,009 steps are accepted and meet the same forcing and feasibility checks. E10 supports geometry-dependent local accuracy without third-order calibration, not runtime gains, global optimality, or generic cubic convergence for conflicting targets; E9 supplies the runtime evidence.