Calibrated α-Barycenters for Multiobjective f-Divergence Optimization

Dier Tang[1], Peijie Li[1], and Xiaoyuan Zhang[2]

[1]The University of Hong Kong , Email: *tangde_math@connect.hku.hk*, *lipeijie@hku.hk*
[2]Zhongguancun Academy , Email: *zhangxiaoyuan@bza.edu.cn*

Abstract

Matching one distribution to many targets can require repeated evaluations of every target. CAL-α-MODM aggregates a nonlinear predictor into one barycenter, calibrates its residual derivatives, and checks steps against the original f-divergence objective. Under smoothness assumptions, calibration controls parameter-step model error without requiring targets to be close to the model. Exact-oracle stationarity and local convergence guarantees are established. At common matching, canonical geometry cancels the leading quadratic iteration error, giving cubic convergence with second-order calibration and sufficiently accurate inner solves. High-precision experiments support this prediction. Sensitivity-based initialization also accelerates preference continuation along regular stationary branches. On 300-target Jensen–Shannon problems, CAL-p2 achieves median paired speedups of 3.15× over L-BFGS and 1.10× over second-order Direct-TR, across 23/25 and 24/25 objective-comparable stationary pairs. The savings arise when cheaper predictor queries offset calibration overhead.

Keywords: Multiobjective optimization, f-divergence, Information geometry, Barycenters, Trust-region methods, Distribution matching

2020 MSC: 90C29, 90C30, 65K05, 94A17

2012 ACM CCS: G.1.6, I.2.6

Contents

1 Introduction . . . 2
2 Problem formulation . . . 2
3 Related work . . . 3
4 Calibrated barycenter predictors . . . 3
5 CAL-α-MODM . . . 6
6 Convergence and Pareto guarantees . . . 8
7 Preference continuation . . . 10
8 Extension to Bregman objectives . . . 10
9 Implementation and computational scope . . . 11
10 Numerical experiments . . . 12
11 Discussion and conclusion . . . 22
Reproducibility statement . . . 23
AI use statement . . . 23
References . . . 23
A Proof of the residual derivative bounds . . . 26
B Proofs of the convergence guarantees . . . 26
C Experimental protocols and additional results . . . 28

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

$$\min_{\theta \in \Theta} (F_1^\sigma(\theta), \dots, F_m^\sigma(\theta)), \qquad F_i^\sigma(\theta) = \begin{cases} D_f(P_i \| Q_\theta), & \sigma = -, \\ D_f(Q_\theta \| P_i), & \sigma = +. \end{cases} \tag{1.1}$$

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× over L-BFGS and 1.10× over second-order Direct-TR, with 23/25 and 24/25 objective-comparable stationary pairs. Latent targets demonstrate preference-controlled generation.

Organization. Sections 2–3 introduce the setting and related work. Sections 4–6 develop the models, algorithm, and guarantees. Extensions and implementation are covered in Sections 7–9, followed by experiments in Section 10. The appendices give the remaining proofs and experimental protocols.

2 Problem formulation

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)

$$D_f(P\|Q) = \int q(x) f\left(\frac{p(x)}{q(x)}\right) \mathrm{d}\mu(x). \tag{2.1}$$

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

$$F_\lambda^\sigma(\theta) = \sum_i \lambda_i F_i^\sigma(\theta). \tag{2.2}$$

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

Throughout, the model family may be restricted: an aggregated target need not lie in $\mathcal{M}$. Unless stated otherwise, optimization takes place in Euclidean parameter coordinates with exact objective and derivative oracles. The convergence statements distinguish first-order stationarity from global minimization and from coverage of the Pareto set.

3 Related work

Divergence objectives in learning. f-divergences underlie variational classification objectives (Novello and Tonello, 2024), convex losses (Roulet et al., 2025), alignment–diversity trade-offs (Go et al., 2023), and preference optimization (Han et al., 2025). They are also used in sequence-level distillation (Wen et al., 2023), reverse-KL distillation (Gu et al., 2024), and distillation on student-generated sequences (Agarwal et al., 2024). Domain adaptation (Acuna et al., 2021) and robust offline reinforcement learning (Tang et al., 2025) use related discrepancies, although their adversarial formulations differ from our fixed-target scalarizations.

Local statistical geometry. Universal feature analysis relates local inference to low-dimensional dependence representations (Huang et al., 2024). Xu and Zheng (2024) study neural feature learning in function space, where squared dependence norms correspond to χ^2 divergence. Tang et al. (2026) analyze spectral representation learning for local tasks in χ^2 neighborhoods. Task covariance determines transfer regret and motivates a correction to spectral learning. Here, local divergence geometry is used to construct optimization predictors for fixed distributional targets.

Aggregation and centroids. Zhang et al. (2025) characterize KL and inverse-KL Pareto sets in full exponential families and develop multiobjective variational autoencoders. Bregman centroid identities reduce weighted objectives to one target (Banerjee et al., 2005; Nielsen and Nock, 2009). Power-mean annealing paths admit α-divergence characterizations (Masrani et al., 2021), while transformed Bregman divergences characterize quasi-arithmetic means (Brekelmans and Nielsen, 2024). Related constructions yield broader divergence families (Nielsen, 2022) and approximate Jeffreys centers (Nielsen, 2024). We consider general smooth f and restricted model families, using the aggregate as a predictor even when its density cannot be represented by Q_θ.

Multiobjective optimization and preferences. CAGrad (Liu et al., 2021), Nash-MTL (Navon et al., 2022), and FAMO (Liu et al., 2023) adapt task influence. Smooth Tchebycheff scalarization (Lin et al., 2024) and FERERO (Chen et al., 2024) support nonlinear or constrained preferences. Our aggregation identity instead assumes a fixed linear scalarization. For repeated preferences, hypernetworks learn preference-to-parameter maps (Navon et al., 2021), and Pareto set learning approximates solution sets in Bayesian optimization (Lin et al., 2022). E6 uses implicit differentiation (Blondel et al., 2022) to initialize a common corrector along regular stationary branches.

Calibration and model management. Trust-region model management establishes convergence for optimization with approximations (Alexandrov et al., 1998; Booker et al., 1999). Evaluation-complexity theory (Cartis et al., 2022) and local higher-order analysis (Doikov and Nesterov, 2022) provide the corresponding global and local perspectives. We specialize these tools to aggregated nonlinear predictors with calibrated residual derivatives. Stochastic trust-region methods with inexact restoration address sampled finite sums (Bellavia et al., 2023). Extending our results to sampled divergences would require derivative and acceptance-test error controls beyond the deterministic bounds proved here.

4 Calibrated barycenter predictors

4.1 Canonical local α geometry

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

$$D_\alpha(P\|Q) = \frac{4}{1-\alpha^2}\left(1 - \int p^{(1-\alpha)/2} q^{(1+\alpha)/2}\, \mathrm{d}\mu\right), \tag{4.1}$$

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

$$f_\alpha(u) = \frac{4}{1-\alpha^2}(1 - u^{(1-\alpha)/2}) + \frac{2}{1+\alpha}(u-1), \tag{4.2}$$

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

$$\alpha_f := -3 - 2\frac{f'''(1)}{f''(1)}, \qquad h_f := f - a_f f_{\alpha_f}. \tag{4.3}$$

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

$$\left| D_f(P\|Q) - a_f D_{\alpha_f}(P\|Q) \right| \le \frac{M_{f,\delta}}{24} \int q\,|p/q - 1|^4 \,\mathrm{d}\mu, \tag{4.4}$$

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

Proof. The normalization gives $f(1) = f'(1) = 0$. From the derivatives of f_α and (4.3),

$$a_f f''_{\alpha_f}(1) = f''(1), \qquad a_f f'''_{\alpha_f}(1) = -a_f\frac{3+\alpha_f}{2} = f'''(1). \tag{4.5}$$

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 (4.4). □

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

4.2 Exact aggregation

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

$$c^-_{\lambda,\alpha} = \frac{(\sum_i \lambda_i p_i^{r_\alpha})^{1/r_\alpha}}{Z^-_{\lambda,\alpha}}, \qquad Z^-_{\lambda,\alpha} = \int (\sum_i \lambda_i p_i^{r_\alpha})^{1/r_\alpha}\,\mathrm{d}\mu, \tag{4.6}$$

$$c^+_{\lambda,\alpha} = \frac{(\sum_i \lambda_i p_i^{s_\alpha})^{1/s_\alpha}}{Z^+_{\lambda,\alpha}}, \qquad Z^+_{\lambda,\alpha} = \int (\sum_i \lambda_i p_i^{s_\alpha})^{1/s_\alpha}\,\mathrm{d}\mu. \tag{4.7}$$

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

$$\sum_i \lambda_i D_\alpha(P_i\|Q) = C^-_{\lambda,\alpha} + \omega^-_{\lambda,\alpha} D_\alpha(c^-_{\lambda,\alpha}\|Q), \qquad \omega^-_{\lambda,\alpha} = (Z^-_{\lambda,\alpha})^{r_\alpha}, \tag{4.8}$$

$$\sum_i \lambda_i D_\alpha(Q\|P_i) = C^+_{\lambda,\alpha} + \omega^+_{\lambda,\alpha} D_\alpha(Q\|c^+_{\lambda,\alpha}), \qquad \omega^+_{\lambda,\alpha} = (Z^+_{\lambda,\alpha})^{s_\alpha}. \tag{4.9}$$

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.

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

$$\sum_i \lambda_i D_\alpha(P_i\|Q) = K_\alpha \left(1 - \int H^-(x) q(x)^{s_\alpha}\,\mathrm{d}\mu(x)\right). \tag{4.10}$$

By (4.6), $H^- = (Z^-_{\lambda,\alpha})^{r_\alpha}(c^-_{\lambda,\alpha})^{r_\alpha}$, and hence

$$\sum_i \lambda_i D_\alpha(P_i\|Q) = K_\alpha(1 - \omega^-_{\lambda,\alpha}) + \omega^-_{\lambda,\alpha} D_\alpha(c^-_{\lambda,\alpha}\|Q). \tag{4.11}$$

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

$$\sum_i \lambda_i \, \mathrm{KL}(P_i \| Q) = \sum_i \lambda_i \, \mathrm{KL}(P_i \| \bar{P}) + \mathrm{KL}(\bar{P} \| Q), \tag{4.12}$$

$$\sum_i \lambda_i \, \mathrm{KL}(Q \| P_i) = \mathrm{KL}(Q \| G) - \log Z_g. \tag{4.13}$$

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. The case $\alpha = 1$ exchanges them. Their scale factors are one, as also obtained from the stated 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.

4.3 Derivative calibration

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

$$\mathcal{A}(\theta) := a_f \sum_i \lambda_i D_{\alpha_f}^\sigma(P_i, Q_\theta), \tag{4.14}$$

which equals a constant plus one weighted divergence to $c_{\lambda,\alpha_f}^\sigma$ by Theorem 4.2. 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 4.3 (Order-p calibrated model)**.** For $p \in \{1, 2, 3\}$,

$$m_t^{(p)}(v) = \widehat{\mathcal{A}}_t(v) + \sum_{j=0}^{p} \frac{1}{j!} \nabla^j \widehat{\mathcal{E}}_t(0)[v]^j. \tag{4.15}$$

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

$$m_t^{(p)}(v) = T_p \widehat{F}_t(v) + (\widehat{\mathcal{A}}_t(v) - T_p \widehat{\mathcal{A}}_t(v)). \tag{4.16}$$

Calibration matches the true order-p jet (value and derivatives). The choice of α affects only the higher-order nonlinear remainder. E7 tests its finite-step accuracy, E10 its local iteration error, and E9 its computational cost.

Theorem 4.4 (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 $\|v\| \le \Delta_t$,*

$$\nabla^j m_t^{(p)}(0) = \nabla^j \widehat{F}_t(0), \quad 0 \le j \le p, \tag{4.17}$$

$$\left| \widehat{F}_t(v) - m_t^{(p)}(v) \right| \le \frac{L_{p+1}}{(p+1)!} \|v\|^{p+1}, \tag{4.18}$$

$$\left\| \nabla \widehat{F}_t(v) - \nabla m_t^{(p)}(v) \right\| \le \frac{L_{p+1}}{p!} \|v\|^p. \tag{4.19}$$

Proof. 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

$$\left| \widehat{\mathcal{E}}_t(v) - T_p \widehat{\mathcal{E}}_t(v) \right| \le \frac{L_{p+1}}{(p+1)!} \|v\|^{p+1}. \tag{4.20}$$

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 (4.19). □

Theorem 4.4 does not require target proximity. Near matching, likelihood-ratio moments further bound the residual derivatives.

4.4 Residual derivatives near matching

Proposition 4.5 (Residual derivative bounds)**.** *Under the normalization and smoothness assumptions of Theorem 4.1, 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*

$$\|\nabla R^-(\theta)\| \le C_1 \int q_\theta \, |p/q_\theta - 1|^3 \, \mathrm{d}\mu, \tag{4.21}$$

$$\left\|\nabla^2 R^-(\theta)\right\|_{\mathrm{op}} \le C_2 \int q_\theta \, |p/q_\theta - 1|^2 \, \mathrm{d}\mu. \tag{4.22}$$

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

$$\left\|\nabla^3 R^\sigma(\theta)\right\|_{\mathrm{op}} \le C_3 \int q_\theta \, |u_\sigma - 1| \, \mathrm{d}\mu, \qquad u_- = p/q_\theta, \quad u_+ = q_\theta/p. \tag{4.23}$$

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

The proof is given in Appendix A. These likelihood-ratio bounds complement the step-based guarantees of Theorem 4.4. They do not replace residual calibration when the targets are heterogeneous.

5 CAL-α-MODM

5.1 Certified trust-region steps

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

$$\mathrm{pred}_t = m_t^{(p)}(0) - m_t^{(p)}(s_t), \quad \mathrm{ared}_t = F(\theta_t) - F(\theta_t + s_t), \quad \rho_t = \frac{\mathrm{ared}_t}{\mathrm{pred}_t}. \tag{5.1}$$

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:

$$\mathrm{pred}_t \ge \kappa_c \|g_t\| \min\left\{\Delta_t, \frac{\|g_t\|}{\kappa_H}\right\}, \qquad \kappa_c \in (0, 1], \tag{5.2}$$

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 $\kappa_c = 1/2$ (Section 6.1). Larger values require verified decrease.

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

$$\Delta_{t+1} = \begin{cases} \gamma_1 \Delta_t, & \rho_t < \eta_1, \\ \min\{\gamma_2 \Delta_t, \Delta_{\max}\}, & \rho_t \ge \eta_2, \\ \Delta_t, & \text{otherwise.} \end{cases} \tag{5.3}$$

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 (4.15) and compute $\|s_t\| \le \Delta_t$ satisfying (5.2).
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 (5.3) and repeat.

5.2 Calibration order and reuse

Calibration order. Order $p = 1$ requires residual values and gradients. Order $p = 2$ adds Hessians and permits local quadratic convergence under Theorem 6.4. Order $p = 3$ permits local cubic convergence but is more expensive. A p1 inner solver may still use predictor Hessians. With $p = 2$, the retained nonlinear remainder can improve local accuracy without a true third-derivative tensor (Proposition 6.5).

Inner solves and reuse. E8/E9 use regularized Newton iterations, trust-region projection, Armijo backtracking, and a steepest-descent comparator. Section 9 gives the settings and relates the comparator to Cauchy decrease. The residual jet is reused across inner queries and rejected trials at a fixed anchor. A smaller radius tightens the error bound. Accepted steps require fresh calibration.

Computational savings. Cached aggregate values on a reusable support or grid remove the target sum from predictor queries. New nodes may require all targets to be evaluated again. Calibration and acceptance still use the true objective, so predictor savings must offset this cost. The aggregate is reused for a fixed preference and rebuilt when the preference changes.

5.3 Safeguarded lazy calibration

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

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

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,

$$\begin{aligned} F(z_{t,j}) - F(z_{t,j} + u_j) &\ge c_{\text{lazy}} \|u_j\|^2 , \\ \sum_j \|u_j\| &\le \min\{\kappa_{\text{path}} \|g_t\| , \kappa_{\text{move}} \|s_t\|\}. \end{aligned} \tag{5.4}$$

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 (5.3). The step-length budget links uncalibrated motion to certified progress. An anchor-gradient budget alone does not provide this link.

Corollary 5.1 (Lazy-search complexity)**.** *Under Assumption 6.1, including Lipschitz continuity on the lazy segments, the preceding block protocol and (5.4) 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 $O((K_{\text{lazy}} + 1)\varepsilon^{-2})$.*

E4 uses a local curvature estimate and its original path budget, without the additional step-length budget above. Its stationarity results are empirical and are not inferred from this corollary.

The proof of Corollary 5.1 is in Appendix B.3.

6 Convergence and Pareto guarantees

6.1 Exact-oracle assumptions and sufficient decrease

Assumption 6.1 (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 4.4 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 (5.2) for fixed $\kappa_c \in (0, 1]$.

(A5) The acceptance decision uses (5.1), with $\mathrm{pred}_t > 0$, and the radius follows (5.3) for fixed $0 < \eta_1 < \eta_2 < 1$ and $0 < \gamma_1 < 1 < \gamma_2$.

Condition (5.2) is attainable for the nonlinear model itself. For $g_t \neq 0$, 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 (5.2) with $\kappa_c = 1/2$. Section 9 describes a backtracking safeguard when κ_H is not known.

For every $p \ge 1$, Theorem 4.4 implies fully linear bounds

$$\left|F(\theta_t + v) - m_t^{(p)}(v)\right| \le \kappa_f \Delta_t^2, \tag{6.1}$$

$$\left\|\nabla F(\theta_t + v) - \nabla m_t^{(p)}(v)\right\| \le \kappa_g \Delta_t \tag{6.2}$$

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. Proofs of the global and local convergence results are in Appendix B. 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.

6.2 Global convergence and Pareto stationarity

Theorem 6.2 (Global convergence and complexity)**.** *Under Assumption 6.1, for fixed λ and orientation, the zero-tolerance version of Algorithm 1 either stops at an exact stationary point or satisfies*

$$\lim_{t\to\infty} \|\nabla F_\lambda^\sigma(\theta_t)\| = 0. \tag{6.3}$$

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

$$N(\varepsilon) \le C_0 + C_1 \varepsilon^{-2} \tag{6.4}$$

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 (6.4) 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 6.3 (Pareto stationarity)**.** *For C^1 components and $\lambda \in \mathrm{ri}(\Delta_m)$, every accumulation point in Theorem 6.2 is Pareto stationary with a strictly positive multiplier. If each component is convex, that point is globally Pareto optimal.*

6.3 Local convergence order

Theorem 6.4 (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 B.1. Assume the subproblem solver returns steps in the local neighborhood of Lemma B.1 satisfying*

$$\left\|\nabla m_t^{(p)}(s_t)\right\| \le \kappa_s \left\|\nabla F(\theta_t)\right\|^p . \tag{6.5}$$

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

$$\left\|\theta_{t+1} - \theta^\star\right\| \le C_p \left\|\theta_t - \theta^\star\right\|^p . \tag{6.6}$$

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

The forcing condition (6.5) provides a sufficient inner-accuracy requirement for our local-rate guarantee, in the spirit of inexact Newton methods (Dennis and Moré, 1974; Dembo et al., 1982). Lazy safeguards appear in Section 5.3. Common-generator Bregman objectives are zero-residual special cases (Section 8).

6.4 Leading error and canonical cancellation

Proposition 6.5 (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 6.4, impose the stronger forcing condition $\left\|\nabla m_\theta^{(2)}(s)\right\| = O(\|\nabla F(\theta)\|^3)$. Once untruncated and accepted, with $e = \theta - \theta^\star$,*

$$e_+ := \theta + s - \theta^\star = \tfrac{1}{2} H_\star^{-1} \nabla^3 \mathcal{E}(\theta^\star)[e,e] + O(\|e\|^3). \tag{6.7}$$

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 sufficient for cancellation. Heterogeneous targets need not give a cubic rate. The stronger forcing condition uses only model gradients and Hessians.

Proof. 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

$$\nabla F(y) = r + \tfrac{1}{2} \nabla^3 \mathcal{E}(x)[s,s] + O(\|s\|^3). \tag{6.8}$$

The local stationary model branch from Lemma B.1 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 (6.8), 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 (6.8) proves (6.7). These same local steps satisfy the $p = 2$ forcing condition and are eventually accepted as in the proof of Theorem 6.4.

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 $j \le 3$, which vanishes by Theorem 4.1.

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. □

Table 1: 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 $p = 2, 3$

7 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

$$\frac{\partial \theta^\star}{\partial \lambda_j} = -H_\lambda^{-1}(\nabla F_j^\sigma(\theta^\star) - \nabla F_m^\sigma(\theta^\star)), \quad j < m. \tag{7.1}$$

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.

Predictor–corrector use. For a feasible increment with $\mathbf{1}^\top \delta\lambda = 0$, the coordinate-free form is

$$\theta_{\mathrm{pred}} = \theta^\star(\lambda) - H_\lambda^{-1} \sum_i \delta\lambda_i \nabla F_i^\sigma(\theta^\star(\lambda)).$$

At the new preference, rebuild the aggregate and apply the calibrated corrector using the new true scalarization. The predictor accelerates initialization along a regular branch. It does not certify branch coverage or remove the need for a true-gradient stopping test.

8 Extension to Bregman objectives

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

$$\sum_i \lambda_i B_\phi(x_i, y) = C_\lambda^- + B_\phi(\bar{x}_\lambda, y), \qquad \bar{x}_\lambda = \sum_i \lambda_i x_i, \tag{8.1}$$

$$\sum_i \lambda_i B_\phi(y, x_i) = C_\lambda^+ + B_\phi(y, \widetilde{x}_\lambda), \qquad \nabla\phi(\widetilde{x}_\lambda) = \sum_i \lambda_i \nabla\phi(x_i). \tag{8.2}$$

Expanding the definition proves (8.1). Collecting the linear terms in y proves (8.2). 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 6.2 and 6.4 then transfer whenever their Euclidean assumptions hold.

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

$$\Phi_\lambda = \sum_i \lambda_i \phi_i, \qquad b_\lambda = \sum_i \lambda_i \nabla\phi_i(a_i). \tag{8.3}$$

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

$$\sum_i \lambda_i B_{\phi_i}(v, a_i) = C_\lambda + B_{\Phi_\lambda}(v, a_\lambda). \tag{8.4}$$

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.

9 Implementation and computational scope

9.1 The computational regime

Group queries by returned derivative order r (values, gradients, or higher jets):

$$T_{\mathrm{CAL}} = T_{\mathrm{setup}} + \sum_r n_r^F \bar{c}_r^F + \sum_r n_r^{\mathcal{A}} \bar{c}_r^{\mathcal{A}} + T_{\mathrm{other}}. \tag{9.1}$$

Here $n_r^F, n_r^{\mathcal{A}}$ count true-objective and predictor calls, $\bar{c}_r^F, \bar{c}_r^{\mathcal{A}}$ are their mean costs, and T_{other} includes subproblem algebra and control overhead. Setup is charged once per preference. A net gain requires true-oracle savings to exceed setup, predictor-query, and model-management costs. E9 measures this crossover as the target count grows.

9.2 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 (6.5).

9.3 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

$$m(-\alpha g) - m(0) \leq -\tfrac{1}{4}\alpha \|g\|^2. \tag{9.2}$$

The center Hessian supplies an initial scale, not a certified bound over the ball. To connect this test to the theory, suppose $\|\nabla^2 m(v)\|_{\mathrm{op}} \leq \kappa_H$ on the ball, taking $\kappa_H \geq 1$ without loss of generality. Taylor's inequality ensures (9.2) whenever $\alpha \leq 3/(2\kappa_H)$. In exact arithmetic, uncapped halving therefore terminates with

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

Any returned step with at least this model decrease satisfies (5.2) 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.

9.4 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. Order $p = 3$ also requires third derivatives. Tensor storage is not inherent: for $p = 2$,

$$\nabla_v^2 m_\theta^{(2)}(v)u = \nabla^2\mathcal{A}(\theta + v)u + [\nabla^2 F(\theta) - \nabla^2\mathcal{A}(\theta)]u. \tag{9.3}$$

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.

9.5 Integration dimension and sampled oracles

Aggregation holds for integrable densities independently of discretization. For $\alpha \neq \pm 1$, 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.

10 Numerical experiments

The experiments test model fidelity (E1–E5, E7), local iteration error (E10), total cost (E9), learned targets (E8), and repeated preferences (E6). Table 3 summarizes the controlled evidence. Appendix C gives the protocols. Forward/reverse denote $\sigma = -/+$.

10.1 Experimental overview and generator conventions

The labels JS, KLChi, and KL4 denote

$$\begin{aligned} f_{\mathrm{JS}}(u) &= u\log u - (u+1)\log\left(\frac{u+1}{2}\right), \\ f_{\mathrm{KLChi}}(u) &= u\log u - u + 1 + \tfrac{1}{2}(u-1)^2, \\ f_{\mathrm{KL4}}(u) &= u\log u - u + 1 + (u-1)^4. \end{aligned}$$

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 $M = (P+Q)/2$.

10.2 Why compression needs calibration

Geometry, exact aggregation, and step accuracy (E1–E3). E1 recovers residual orders $(4, 3, 2)$ near matched distributions, supporting the canonical local geometry (Figure 1). E2 verifies 15,680 aggregation identities, with maximum scaled error 3.77×10^{-15}. Its repeated-query benchmark measures the cost of constructing and reusing an exact aggregate: at $m = 256$, end-to-end speedups are 2.46–3.14×, with break-even counts of 5–100 queries (Figure 2). These are aggregation microbenchmarks, rather than full optimization timings.

E3 recovers value/directional-derivative orders $(p+1, p)$ even with median maximum absolute log-likelihood ratio 3.94 (Figure 3). Thus the predictor is exactly compressible, and calibration controls parameter-step errors beyond nearby targets.

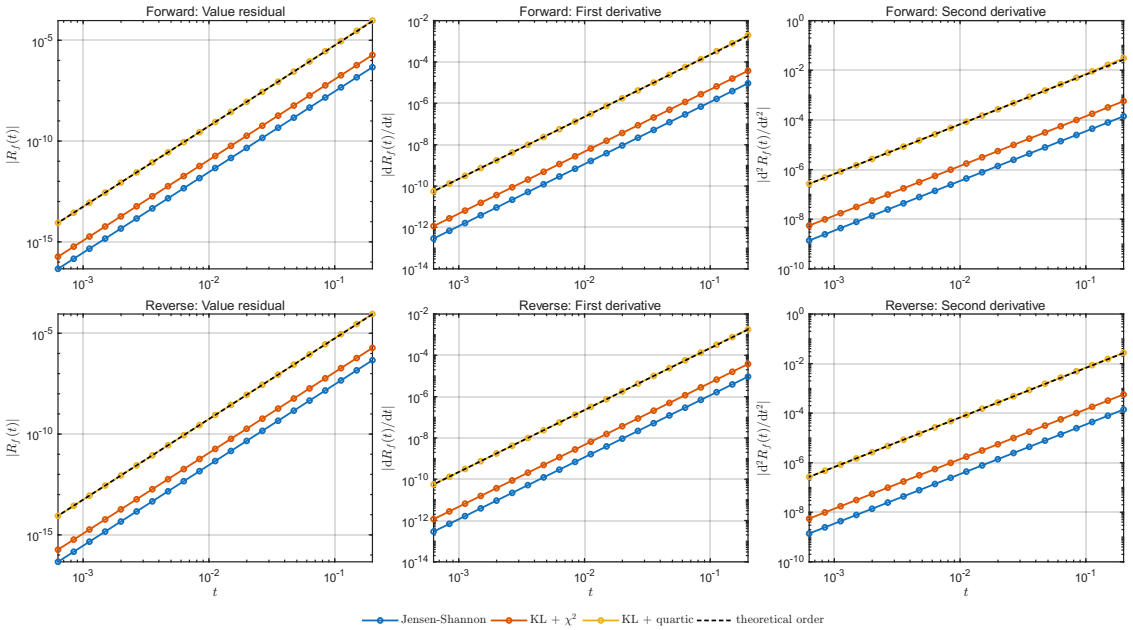


Figure 1: Local normal-form scaling in the forward and reverse orientations. Curves show medians over 20 seeds. Dashed references have slopes 4, 3, and 2. Both axes are logarithmic.

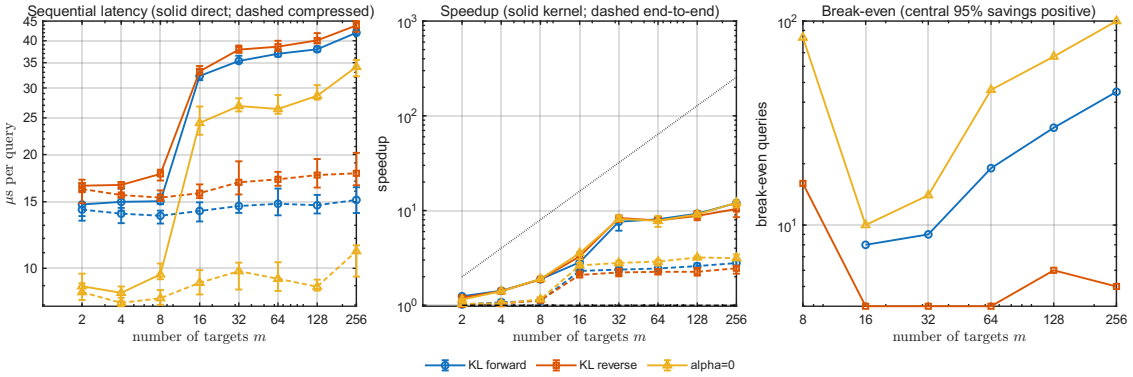


Figure 2: E2 amortized cost, $K = 2048$, 30 paired repeats. Left: direct (solid) and aggregated (dashed) time per query, with IQRs. Middle: target-kernel (solid) and end-to-end (dashed) speedups, with m-fold and unit references. Right: break-even query counts. Gaps indicate a nonpositive 2.5th percentile of paired savings.

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. Figure 4 reports the E4 success and cost comparisons. These comparisons show why a compressible approximation needs calibration to serve the specified divergence.

Local order under controlled inner accuracy (E5). E5 uses a positive affine probability chart, ten seeds, two generators, both orientations, and 120-digit arithmetic. Independent one-step maps recover orders two and three for $p = 2$ and $p = 3$, respectively (Figure 5). All 160 primary trajectories reach the true-gradient tolerance. Condition ablations show why inner accuracy, eventually inactive truncation, and nondegeneracy matter. The full protocol and actual-iterate orders are in Appendix C.5.

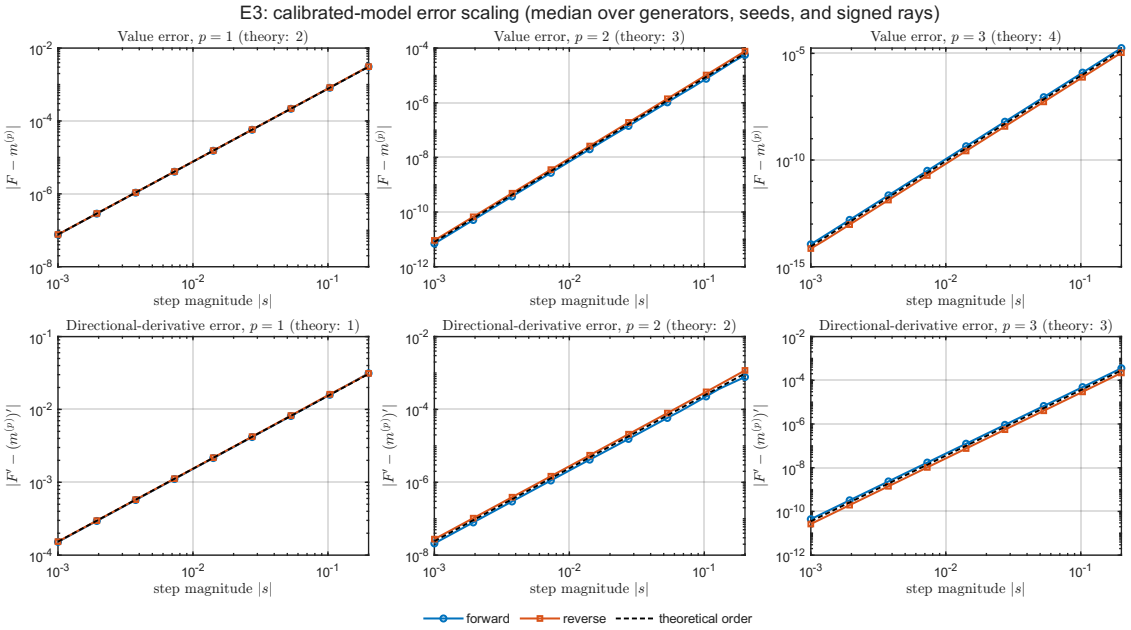


Figure 3: Step-local calibrated-model error scaling. Columns correspond to $p = 1, 2, 3$. Rows show value and directional-derivative errors. Curves are medians over three generators, 20 seeds, and two signed rays. Dashed references have slopes $p + 1$ and p.

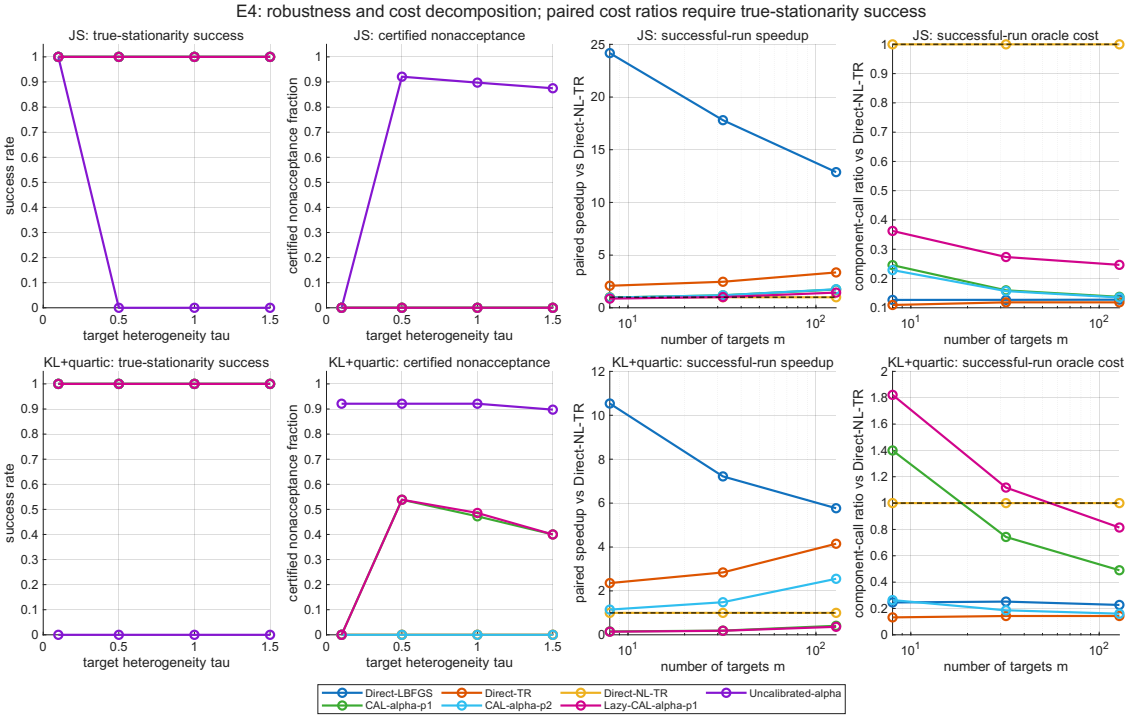


Figure 4: E4 robustness and cost, 30 paired seeds. Rows: JS and KL+quartic. Left: success rates and median certified nonacceptance fractions, including failed subproblems and rejected trials (not applicable to L-BFGS). Right: median paired runtime speedups and component-call ratios against Direct-NL-TR in the target-count sweep. Both runs must succeed.

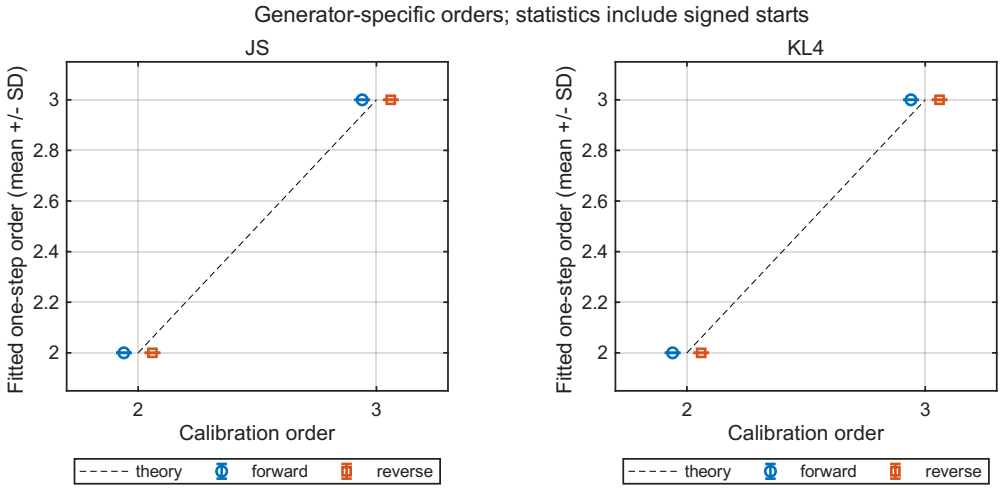


Figure 5: E5 one-step orders: means and one standard deviation over ten seeds and two paired signs (bars smaller than markers, not confidence intervals). Dashed lines mark order p. Actual-iterate orders appear in Table 7.

Template choice at fixed calibration order (E7). E7 holds the matched quadratic jet fixed and varies the nonlinear template. Canonical geometry improves the displayed finite-step model errors in several regimes (Figure 6), but does not guarantee the smallest error for every target configuration. All calibrated models and Taylor-p2 attain true-gradient stationarity in 480/480 cases. Canonical CAL's group-median paired speedups are 0.470–0.816 relative to Taylor-p2, so these accuracy improvements do not imply lower runtime in this experiment.

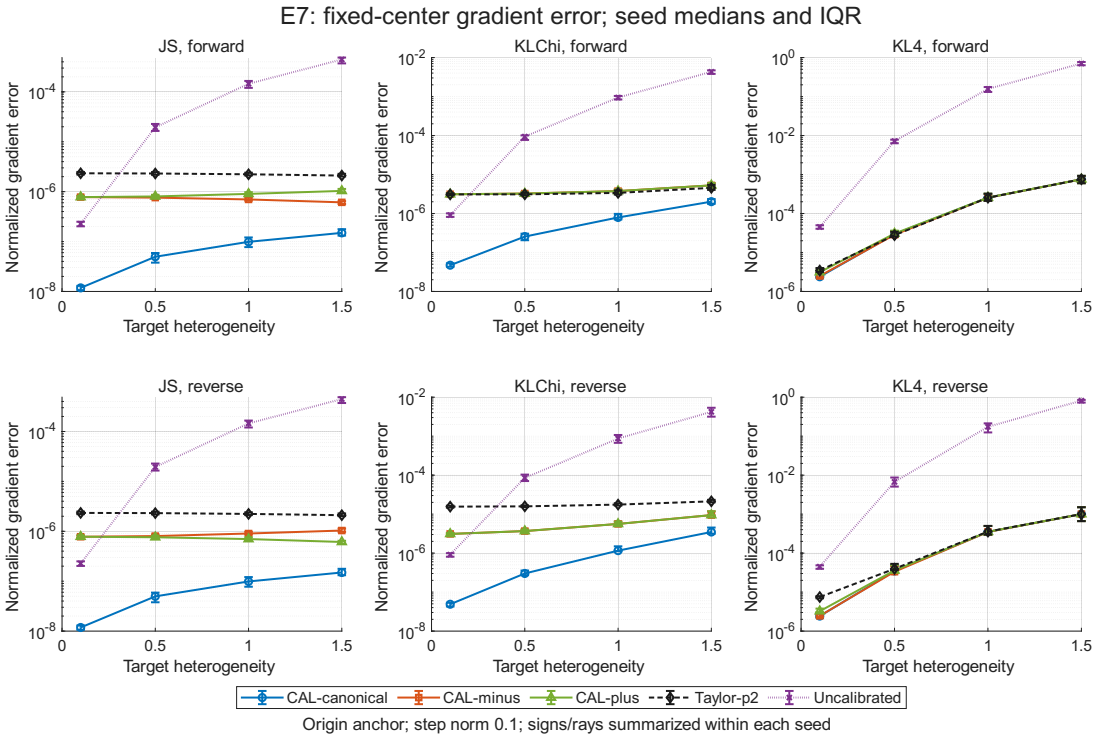


Figure 6: E7 normalized gradient errors (C.13) at $\theta = 0$, $\|s\| = 0.1$. Directions and signs are summarized within seeds, then medians/IQRs across 20 seeds. CAL-minus/plus use $\alpha_f \mp 1$. All calibrated models and Taylor-p2 match through order two. JS orientations coincide.

10.3 Canonical geometry and local iteration error (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 (6.7) leaves a cubic remainder for every method, directly testing the coefficient formula.

For $\eta > 0$, fitted orders return to approximately two. Canonical CAL-p2 has smaller one-step errors than Taylor-p2 in all 360 paired probes. These tests measure local accuracy at fixed calibration order, not runtime or a universal ranking of templates (Appendix C.10).

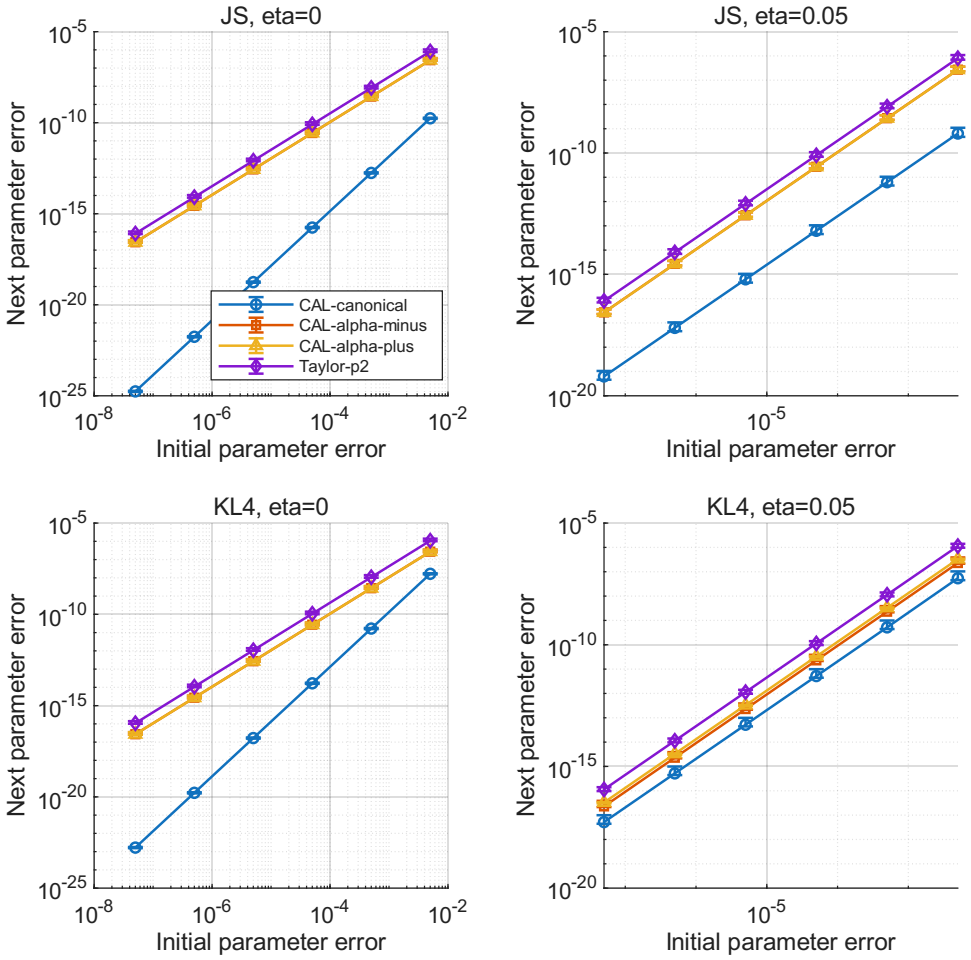


Figure 7: E10 one-step errors, seed medians/IQRs, with identical $p = 2$ calibration and cubic inner forcing. Canonical CAL-p2 is cubic at $\eta = 0$ and quadratic for positive heterogeneity. Panels show $\eta = 0, 0.05$ and all six radii. Fits and audits use all four heterogeneities.

Error coefficients. Table 2 compares the leading directional errors in (6.7). At $\eta = 0.05$, canonical CAL-p2 reduces the seed-median coefficient by a factor of 2.37×10^3 for JS and 31.8 for KL4. The models have identical center values, gradients, and Hessians. Their different nonlinear remainders account for the change in local error. E7 tests finite-step model accuracy, and E9 examines whether aggregation lowers runtime.

Table 2: E10 directional error ratio $\|b_{\text{Taylor}}(d_s)\| \,/\, \|b_{\text{CAL}}(d_s)\|$. Medians over five seeds. Values above one favor canonical CAL-p2. The denominator vanishes 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

Table 3: 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, $p = 3$

10.4 When compression reduces total cost (E9)

Design and cost accounting. E9 forms nested banks of $m \in \{3, 10, 30, 100, 300\}$ Gaussian-mixture targets by perturbing three fitted latent targets. Five seeds, five positive preferences, and forward JS or KL are used. The model has two Gaussian components and five free parameters. CAL-p1/p2 are compared with direct L-BFGS and a quadratic true-objective trust-region solver (Direct-TR). All methods share target caches, vectorized evaluations, initial points, and stationarity thresholds. The direct KL baselines use exact aggregation, as does the predictor.

Total time includes scalarization setup, calibration, inner solves, and true-objective acceptance. Common data preparation and independent audits are excluded. Three unprofiled repeats are summarized within each case. Timing pairs must be stationary and have objective values within 2×10^{-6}. Ratios are summarized over preferences within seeds, then by medians and IQRs across five seeds. All 1,000 primary runs pass the stationarity audit. Different local solutions account for incomplete timing-pair coverage.

Cost crossover. CAL-p2 becomes faster than L-BFGS at $m = 30$. At $m = 300$, median paired speedups are 3.15× over L-BFGS and 1.10× over Direct-TR (Figure 8). CAL-p2 uses median counts of five true value/gradient/Hessian calls and four additional value calls, compared with six and five for Direct-TR. The aggregate handles the repeated inner queries. Pairs reaching different local solutions are excluded from timing ratios but retained in the endpoint audits and reference-gap plots.

Table 4: E9 CAL-p2 speedups on JS. Parentheses give 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 9). This separation offsets calibration and inner-solver overhead. An implementation ablation keeps the optimization rules fixed and verifies agreement of acceptance decisions

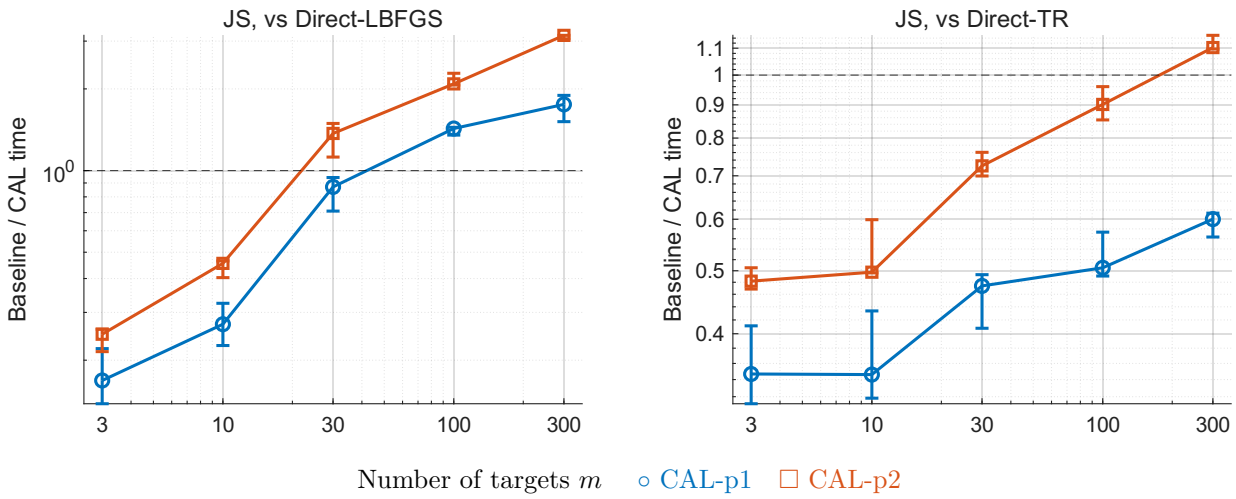


Figure 8: E9: measured JS acceleration relative to direct L-BFGS (left) and Direct-TR (right). Values above one favor CAL. Points and bars are seed medians and IQRs of objective-comparable successful pairs. These are the JS panels of the complete comparison in Figure 35. Table 4 reports CAL-p2 pair coverage.

and endpoints while requesting only the needed derivative order, confirming the cost of unnecessary template Hessians (Figure 10).

Derivative-demand control. 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 10 shows the timing and template-Hessian savings.

Direct-TR matches CAL-p2's true derivative order. The larger speedup over L-BFGS also reflects different use of curvature. CAL-p2 is slower than Direct-TR through $m = 100$. At $m = 300$, the JS speedup IQRs are $[3.06, 3.15]$ against L-BFGS and $[1.10, 1.15]$ against Direct-TR. The solve time includes median counts of 56 predictor-value calls and 14 predictor-Hessian calls. Target count alone does not ensure savings: E4's direct solvers remain faster at $m = 128$ (Table 6). The evidence supports a cost crossover at comparable attained objectives, without implying better minima or faster exact-KL optimization.

10.5 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 (Figure 12).

Preference response. Changing λ changes which learned target is prioritized (Figure 13). These weights specify divergence objectives, not generated class proportions.

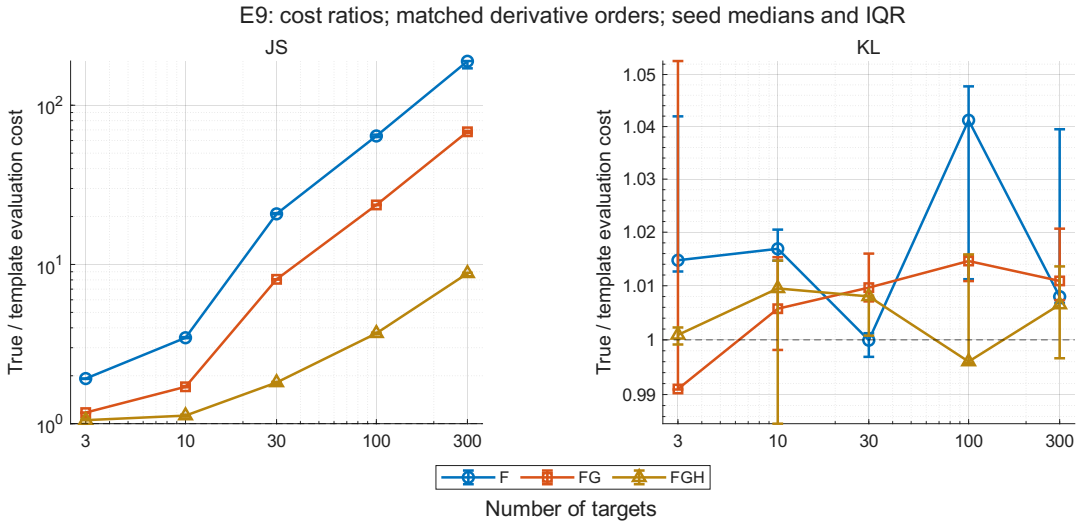


Figure 9: E9: true-objective cost divided by aggregated-predictor cost, at matched derivative orders. F, FG, and FGH include values, values/gradients, and values/gradients/Hessians. The JS cost separation grows with the target count. KL stays near one because both implementations aggregate exactly. Points and bars are seed medians and IQRs across five seeds.

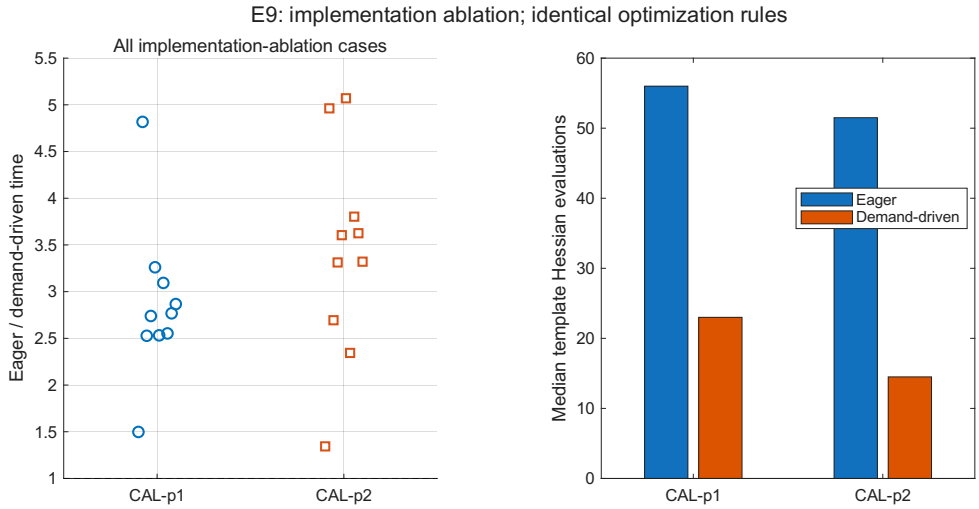


Figure 10: E9 implementation ablation with identical optimization rules. Left: eager/demand-driven time ratios for CAL-p1/p2. Right: median template-Hessian counts. This compares two implementations of the same method, not CAL against a direct optimizer.

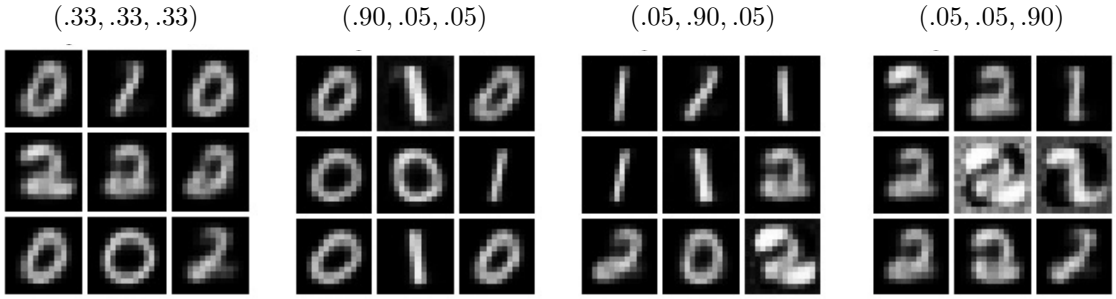


Figure 11: E8: saved decoded mean images for CAL-p2, JS, seed 1, with a fixed decoder. Each preference λ has nine latent draws, without reranking. All-method samples and complete held-out image diagnostics are in Appendix C.8.

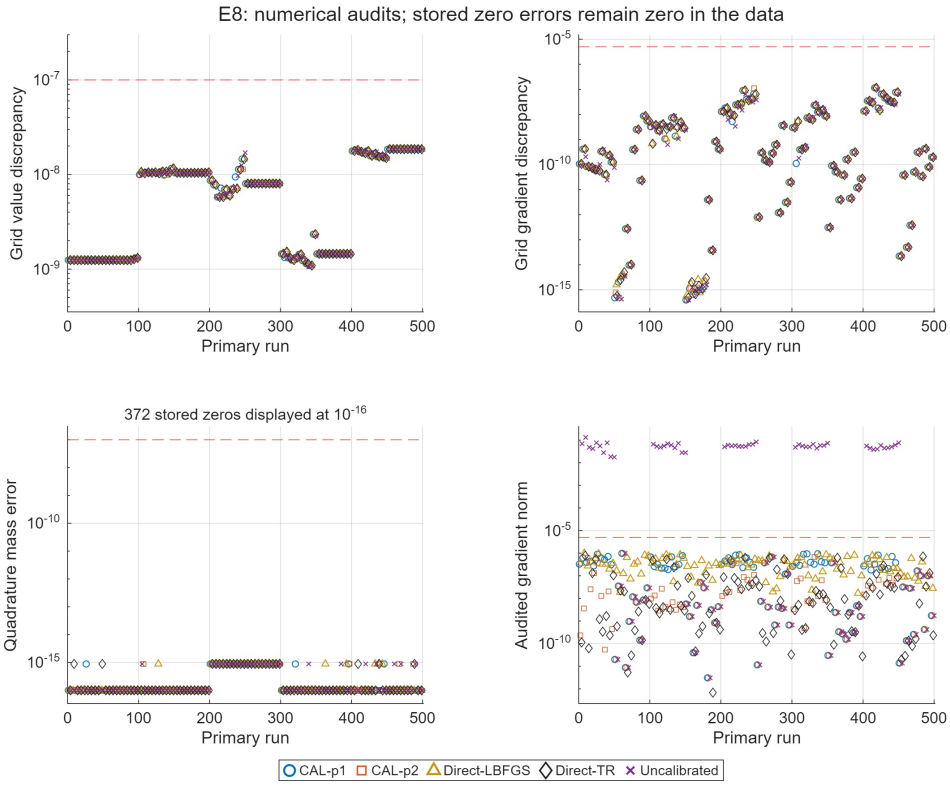


Figure 12: E8 quadrature and endpoint audits. Dashed lines mark the configured thresholds. The mass panel reports normalization error of the discrete masses. Stored zeros are displayed at 10^{-16} without changing the data. Large true gradients of uncalibrated JS endpoints explain their failed audits.

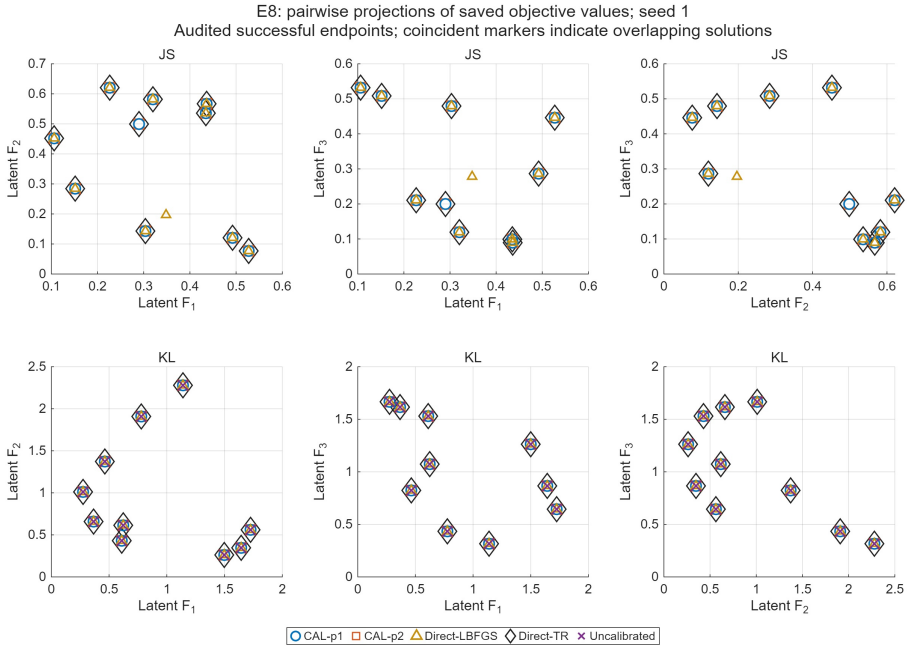


Figure 13: E8 pairwise projections of latent objectives at audited successful endpoints, seed 1. Coincident markers indicate overlapping solutions. The points are sampled stationary trade-offs, not a certified full Pareto front.

Decoded consequences. Figure 11 shows the changes in decoded digits. Independent class proportions and image-space MMD^2 (Figure 14) describe the resulting distributions. They do not enter the optimized objective or certify stationarity. The targets have six Gaussian components each, while the model has two with a fixed shared covariance. Thus the aggregate need not be representable by the model. E8 validates a restricted low-dimensional fit, without establishing better image quality or faster three-target optimization.

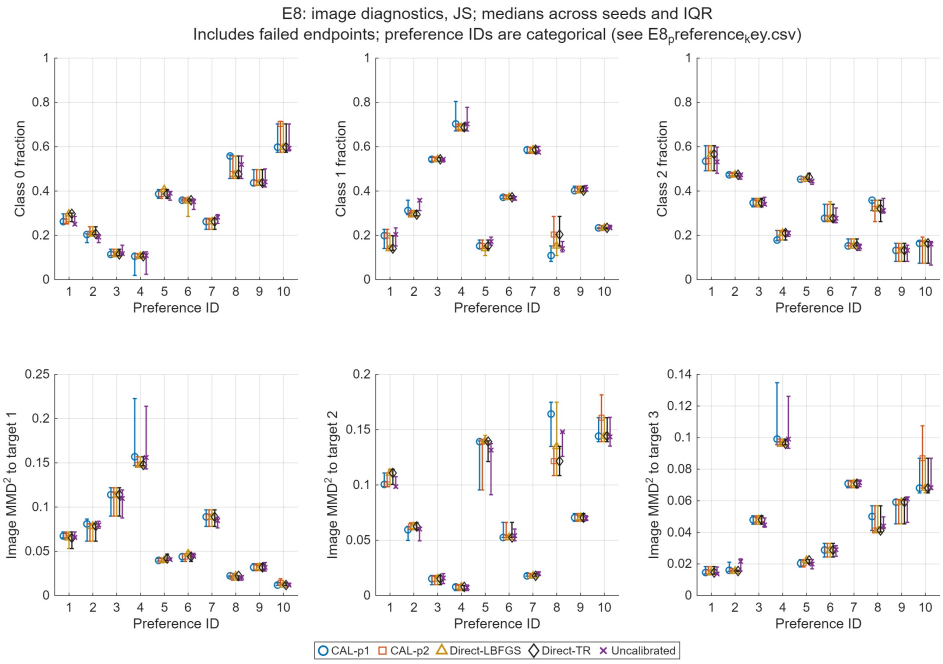


Figure 14: E8 held-out image diagnostics: class fractions and image MMD^2 under JS, summarized by seed medians and IQRs. Preference IDs are categorical and follow the saved preference key. Failed endpoints are included. These quantities describe decoded distributions rather than certify optimality.

10.6 Complementary acceleration across preferences (E6)

Standard implicit differentiation provides a complementary initializer for repeated preferences, independently of α-aggregation (Section 7). 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× over Warm and 2.03–2.10× over Cold, including aggregation, tangent solves, and correction. Reusing regular stationary branches thus reduces the work of the calibrated corrector (Table 5).

Table 5: 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 (Figure 15). The same corrector and stopping test are used throughout (Appendix C.6).

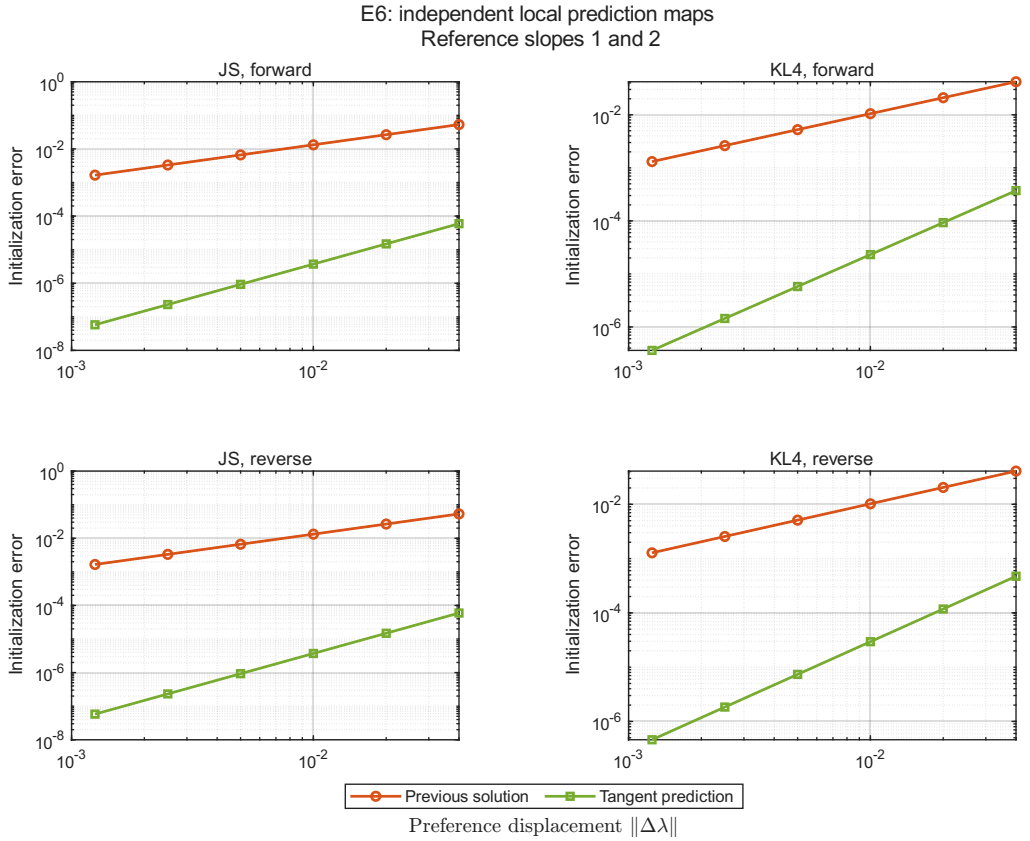


Figure 15: E6 independent prediction errors: medians over signs within each seed, then over ten seeds. Warm uses the base reference. Tangent adds its sensitivity prediction. Dashed slopes one and two nearly overlap the curves. The horizontal axis measures preference displacement, not corrector iteration.

11 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.

Computational scope. E9 identifies a regime where cheaper predictor queries offset calibration overhead. E6 adds initialization savings for repeated preferences. When the true objective 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 (Section 9.4). The guarantees concern stationarity, not full Pareto coverage.

Reproducibility statement

Section 4 proves the model identities and error bounds. Section 6 states the exact-oracle assumptions, and Appendices A–B give the remaining proofs. Appendix C records all ten MATLAB experimental protocols, including seeds, tolerances, timing boundaries, and audits. E5 and E10 use 120-digit arithmetic. Histories, paired timings, and validation records were retained as described in the protocols.

AI use statement

Generative AI tools were used to assist with polishing the language, improving readability and organization, and formatting the manuscript in LaTeX. They were also used to assist in generating and debugging MATLAB code for the numerical experiments. All AI-generated content was carefully reviewed, revised, and verified by the authors, who assume full responsibility for the final manuscript.

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*, pages 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, pages 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, pages 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*, pages 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*, pages 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, pages 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*, pages 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, pages 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, pages 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*, pages 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*, pages 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*, pages 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*, pages 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*, pages 58842–58882, 2025. URL https://proceedings.mlr.press/v267/tang25f.html.

Dier Tang, Jing Yee Tan, and Guangyue Han. Sharp regret bounds and a task-covariance correction for spectral representation learning, 2026. URL https://arxiv.org/abs/2609.15825. arXiv:2609.15825.

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)*, pages 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*, pages 75413–75426, 2025. URL https://proceedings.mlr.press/v267/zhang25as.html.

A Proof of the residual derivative bounds

Proof of Proposition 4.5. Write $h = h_f$, $u = p/q_\theta$, $z = \nabla \log q_\theta$, and $J = \nabla^2 \log q_\theta$. For the forward orientation,

$$\nabla R^-(\theta) = \int q_\theta (h(u) - uh'(u)) z \, \mathrm{d}\mu, \tag{A.1}$$

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

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 $u = q_\theta/p$,

$$\nabla R^+(\theta) = \int q_\theta h'(u) z \, \mathrm{d}\mu, \tag{A.3}$$

$$\nabla^2 R^+(\theta) = \int q_\theta [(h'(u) + uh''(u)) zz^\top + h'(u) J] \, \mathrm{d}\mu, \tag{A.4}$$

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,

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

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 (4.23). Weighted sums give the corresponding scalarized residual bound. □

B Proofs of the convergence guarantees

B.1 Proof of Theorem 6.2

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

$$|\mathrm{ared}_t - \mathrm{pred}_t| \le \kappa_f \Delta_t^2. \tag{B.1}$$

Define

$$c_\rho := \min \left\{ \frac{1}{\kappa_H}, \frac{(1-\eta_2)\kappa_c}{\kappa_f} \right\}. \tag{B.2}$$

If $g_t \neq 0$ and

$$\Delta_t \le c_\rho \, \|g_t\| \,, \tag{B.3}$$

then (5.2)–(B.1) 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,

$$\Delta_t \ge \underline{\Delta}_\varepsilon := \gamma_1 c_\rho \varepsilon. \tag{B.4}$$

Indeed, an unsuccessful iteration must have $\Delta_t > c_\rho \, \|g_t\| > c_\rho \varepsilon$, so its contracted successor exceeds (B.4). 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

$$\eta_1 \kappa_c \varepsilon \min\{\gamma_1 c_\rho \varepsilon, \varepsilon/\kappa_H\} = c_s \varepsilon^2. \tag{B.5}$$

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 (B.4) imply

$$N_u|\log\gamma_1| \le N_v \log\gamma_2 + \log\frac{\Delta_{\max}}{\gamma_1 c_\rho \varepsilon}. \tag{B.6}$$

Since N_v is no larger than the successful count and $|\log\varepsilon| = O(\varepsilon^{-2})$ on $(0,1)$, this proves (6.4). 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

$$b_\varepsilon = \eta_1\kappa_c\varepsilon\min\{1, \varepsilon/(\kappa_H\Delta_{\max})\} > 0. \tag{B.7}$$

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

$$F(\theta_t) - F(\theta_{t+1}) \ge b_\varepsilon\Delta_t \ge b_\varepsilon\|s_t\|. \tag{B.8}$$

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 (B.8) 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 6.3. 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. □

B.2 Local stationary branch and proof of Theorem 6.4

For a generic anchor θ, write $m_\theta^{(p)}$ for (4.15) and define $G_p(\theta, s) = \nabla_s m_\theta^{(p)}(s)$.

Lemma B.1 (Local model stationary branch)**.** *Under the hypotheses of Theorem 6.4, 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*

$$\|s_p(\theta)\| \le C_s\|\theta - \theta^\star\|, \qquad \nabla_s^2 m_\theta^{(p)}(s) \succeq \frac{\mu}{2}I \tag{B.9}$$

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 (B.9). Continuity gives the Hessian bound. Strong monotonicity of $G_p(\theta, \cdot)$ then gives the final inequality. □

Proof of Theorem 6.4. 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

$$c_1 e_t \le \|\nabla F(\theta_t)\| \le c_2 e_t. \tag{B.10}$$

Lemma B.1 and (6.5) give $\left\|s_t - s_p^t\right\| = O(e_t^p)$ and hence $\|s_t\| \le c_3 e_t$. At the trial point, (4.19) and (6.5) give

$$\|\nabla F(\theta_t + s_t)\| \le \frac{L_{p+1}}{p!}\|s_t\|^p + \kappa_s\|\nabla F(\theta_t)\|^p \le c_4 e_t^p. \tag{B.11}$$

Equation (B.10) 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 $\left\|s_p^t\right\| \geq \|g_t\|/M$. Strong convexity gives $m_t^{(p)}(0) - m_t^{(p)}(s_p^t) \geq \mu\left\|s_p^t\right\|^2/4$, 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 (6.6) applies. □

B.3 Proof of Corollary 5.1

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 6.2 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 (B.8), 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. □

C Experimental protocols and additional results

The generators and their normalization are defined in Section 10.1.

C.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

$$Q_t(k) = \frac{P_k \exp(tv_k)}{\sum_\ell P_\ell \exp(tv_\ell)}, \qquad Q_0 = P, \tag{C.1}$$

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}.

C.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 (4.8)–(4.9), 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

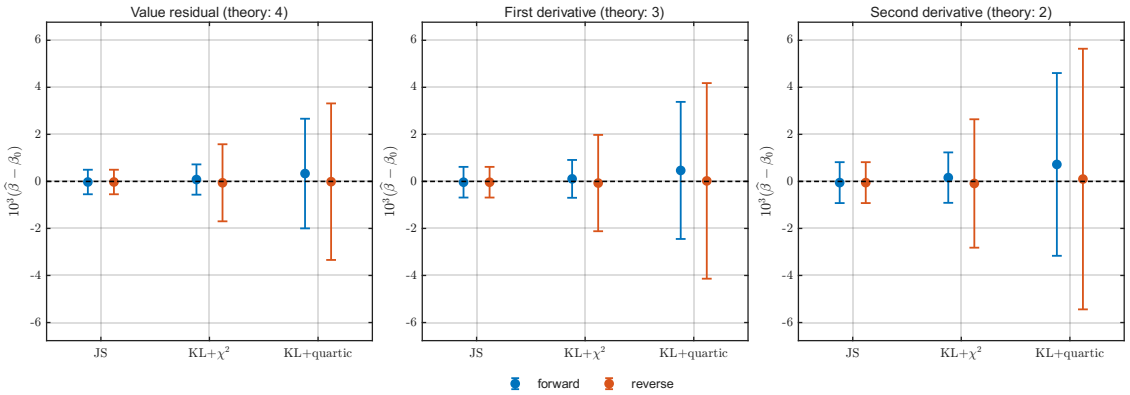


Figure 16: Fitted-exponent deviations for E1, scaled by 10^3. Markers show means over 20 seeds, error bars one standard deviation, and the dashed line exact agreement.

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.

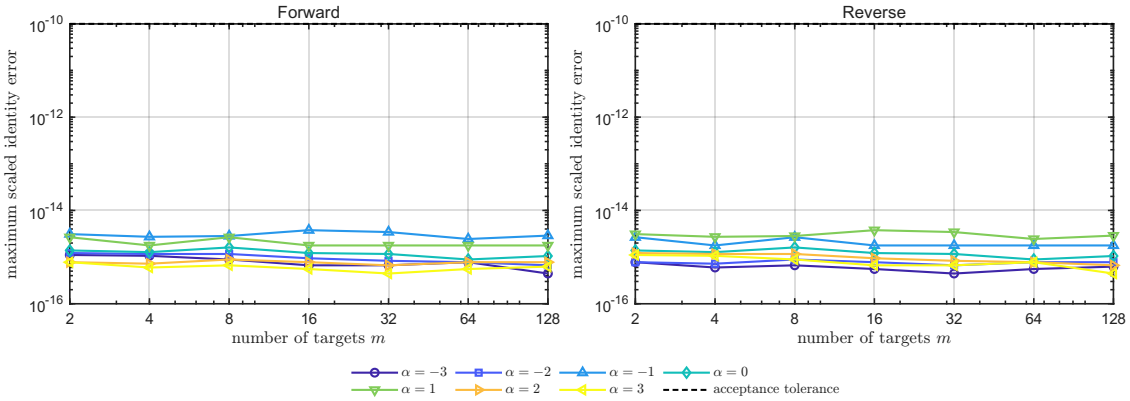


Figure 17: E2 aggregation identity errors in both orientations: maximum scaled discrepancy over 20 seeds and eight queries per seed, including the aggregate's constant and scale. The dashed line marks tolerance 10^{-10}. Errors remain at double-precision roundoff.

C.3 Calibrated-model order diagnostics

E3 uses $K = 64$ categories, 16 targets, 20 seeds, and the three generators in Section 10.1, each in both orientations. For each seed, draw $z_c \sim \mathcal{N}(0, 0.45^2 I_K)$ and $z_i \sim \mathcal{N}(0, I_K)$ independently. Set $Q_0 = \mathrm{softmax}(z_c)$ and $P_i = \mathrm{softmax}(z_i)$. Normalize weights $\lambda_i \propto 0.25 + U_i$, with independent $U_i \sim \mathrm{Unif}(0, 1)$. Center a Gaussian direction d under Q_0 and rescale it to $\|d\|_\infty = 1$. Both signed rays are used.

The 25 step magnitudes are logarithmically spaced in $[10^{-3}, 0.2]$. Log–log fits use $[2 \times 10^{-3}, 3 \times 10^{-2}]$ and exclude errors below 10^{-13}. If fewer than six points remain, the upper endpoint is extended to 6×10^{-2}. The predictor includes the aggregation constant and scale. Residual coefficients through degree three are computed with fourth-order truncated-series arithmetic.

For $p = 1, 2, 3$, value and directional-derivative fits recover orders $(p + 1, p)$ (Figure 3). The median maximum absolute log-likelihood ratio is 3.94, and the median target-to-center KL is 0.540. Each of the 18 generator–orientation–order cells has 40 trials (20 seeds, two signs). All 720 trials yield valid value and derivative fits. The largest mean-order deviation is 1.67×10^{-2} and the largest standard deviation is 0.148 (Figure 18). At $p = 2$, one value fit and two directional-derivative fits have $R^2 < 0.95$. The two derivative residuals change sign within the window. All fits are retained.

All 80 exact-α controls with $\alpha = 0.4$ pass. Their maximum value, directional-derivative, and

residual-jet discrepancies are 1.94×10^{-15}, 1.25×10^{-16}, and 1.78×10^{-15}.

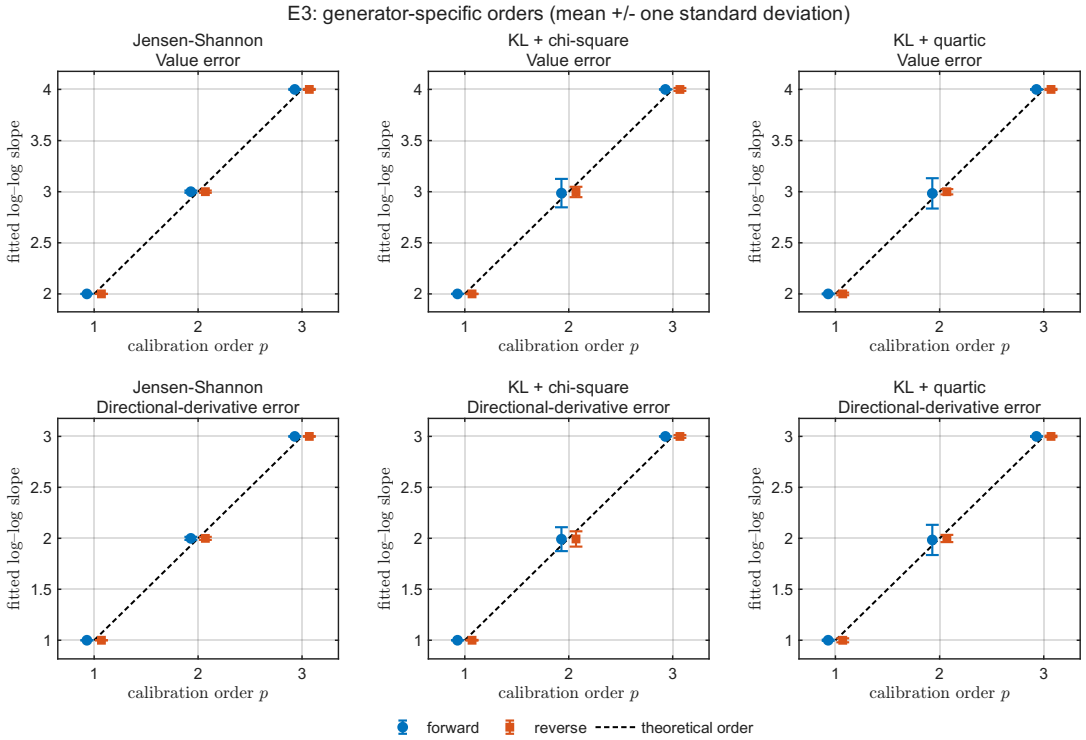


Figure 18: E3 fitted model-error orders by generator (columns): value errors (top, order $p+1$) and directional-derivative errors (bottom, order p). Markers and bars show means and one standard deviation over 20 seeds and two signed rays, not confidence intervals. Blue circles/orange squares denote forward/reverse orientations, offset for visibility. Dashed lines mark theoretical orders.

C.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 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

$$P_i = \mathrm{softmax}(b + \tau B\xi_i + 0.1\tau\Pi z_i)\,, \tag{C.2}$$

$$\theta_0 = -\frac{0.75\tau}{m}\sum_{i=1}^{m}\xi_i + 0.8\max\{\tau, 0.25\}\zeta. \tag{C.3}$$

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 in Section 10.1, 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 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 $c_{\text{lazy}} = 10^{-4}$. Its path coefficient is selected as $0.25/L_{\text{est}}$, where

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

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 Section 5.3 was not enforced here, so these measurements evaluate the implemented heuristic, rather than verify every hypothesis of Corollary 5.1.

Metrics and paired comparisons. Success is defined exclusively by $\|\nabla F^{\sigma}_{\lambda}(\theta_{\text{final}})\| \le 10^{-6}$. 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.

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

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 19 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.

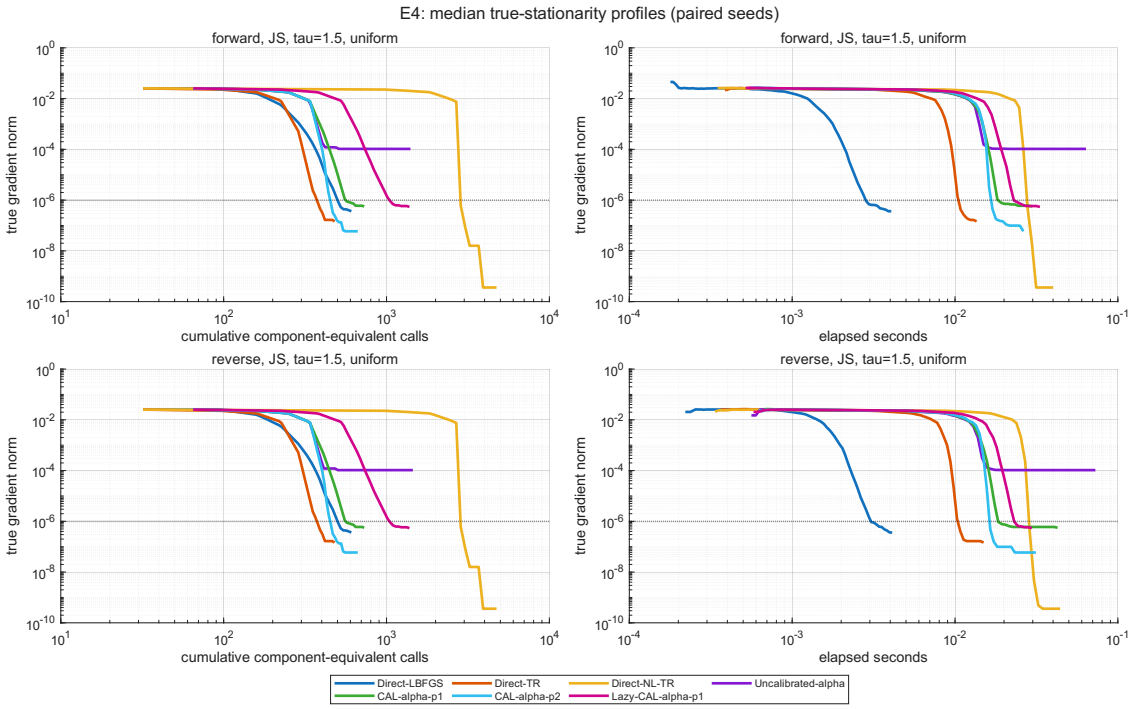


Figure 19: E4 true-gradient profiles: JS, $(K, d, m) = (128, 16, 32)$, $\tau = 1.5$, uniform preferences. Rows separate orientations. Budgets are component-equivalent calls (left) and time (right). Curves are pointwise medians over 30 seeds after log-coordinate interpolation. Final gradients are carried forward, and budgets before the first record are unavailable. The dotted line marks 10^{-6}.

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.

C.5 Local high-order convergence and condition ablations

Problem construction. E5 tests Theorem 6.4 on the affine chart $q_\theta = q_0 + B\theta > 0$, with $(K, d, m) = (6, 2, 4)$. For seed $s \in \{1, \ldots, 10\}$, the MATLAB random stream is `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

$$q_0 = \mathrm{softmax}(z), \qquad \widetilde{G} = G - \mathbf{1}(q_0^\top G), \qquad B = \frac{0.4\ \mathrm{diag}(q_0)\widetilde{G}}{\max_k \left\|\widetilde{G}_{k,:}\right\|_2}. \tag{C.4}$$

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 Section 10.1, in both orientations. The same primitive random draws are used across generators, orientations, and calibration orders within each seed.

Precision and local solves. E5 uses 120-digit arithmetic. Independent damped Newton solves obtain 40 references with $\|\nabla F(\theta^\star)\| \le 10^{-90}$ and minimum Hessian eigenvalues in $[2.206\times10^{-3}, 0.665]$. References are used only for initialization and error measurement. Acceptance and stopping use the true objective and gradient. Recomputing the first seed's four references at 160 digits gives a maximum parameter discrepancy of 4.88×10^{-108}. The diagnostic error floor is 10^{-80}, without an interval certificate.

Primary runs use analytic residual derivatives through $p = 2$ or 3. At most 35 damped Newton iterations, initialized at zero and maintaining positive model Hessians, seek the local stationary branch. The forcing condition (6.5) uses $\kappa_s = 0.01$. An arithmetic floor of 10^{-105} is recorded separately. Untruncated primary steps are accepted only if the original forcing condition holds. The initial and maximum radii are 1 and 4, with $(\eta_1, \eta_2) = (0.1, 0.75)$. Rejection halves the radius, and a very successful step doubles it. Each applied ball is limited to 0.45 times the distance to the positive-chart boundary. The Cauchy test uses an analytic ballwise curvature bound and $\kappa_c = 0.1$. Runs stop at true-gradient norm 10^{-30} or after 60 trials.

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

$$r_\star = \min_{k:\|B_{k,:}\|_2>0} \frac{q_{\theta^\star}(k)}{\|B_{k,:}\|_2}, \qquad c_\star = \min\{1, 0.2r_\star\}, \qquad \theta_{\epsilon,h} = \theta^\star + \epsilon h c_\star v, \quad \epsilon \in \{-1, 1\}. \tag{C.5}$$

For each problem, order, and sign, independent starts at $h \in \{10^{-2}, \dots, 10^{-8}\}$ give 1,120 probes. Log next-error versus log initial-error fits use the four fixed radii $h \in \{10^{-3}, \dots, 10^{-6}\}$. All four points must be accepted, untruncated, forcing-compliant, and above the error floor. All 1,120 probes and 160 fits qualify, without slope-based filtering or adaptive windows. The largest 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 20 and 5 report the results. Signed starts share a seed. JS orientations coincide and are not independent replications.

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

$$\widehat{q}_t = \frac{\log(e_{t+1}/e_t)}{\log(e_t/e_{t-1})}. \tag{C.6}$$

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.

Table 7: E5 actual-iterate orders: mean ± 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

Condition ablations and singular control. Four $p = 3$ controls modify one condition on the first forward KL+quartic problem with positive initial direction (Figure 22). A fixed inner tolerance

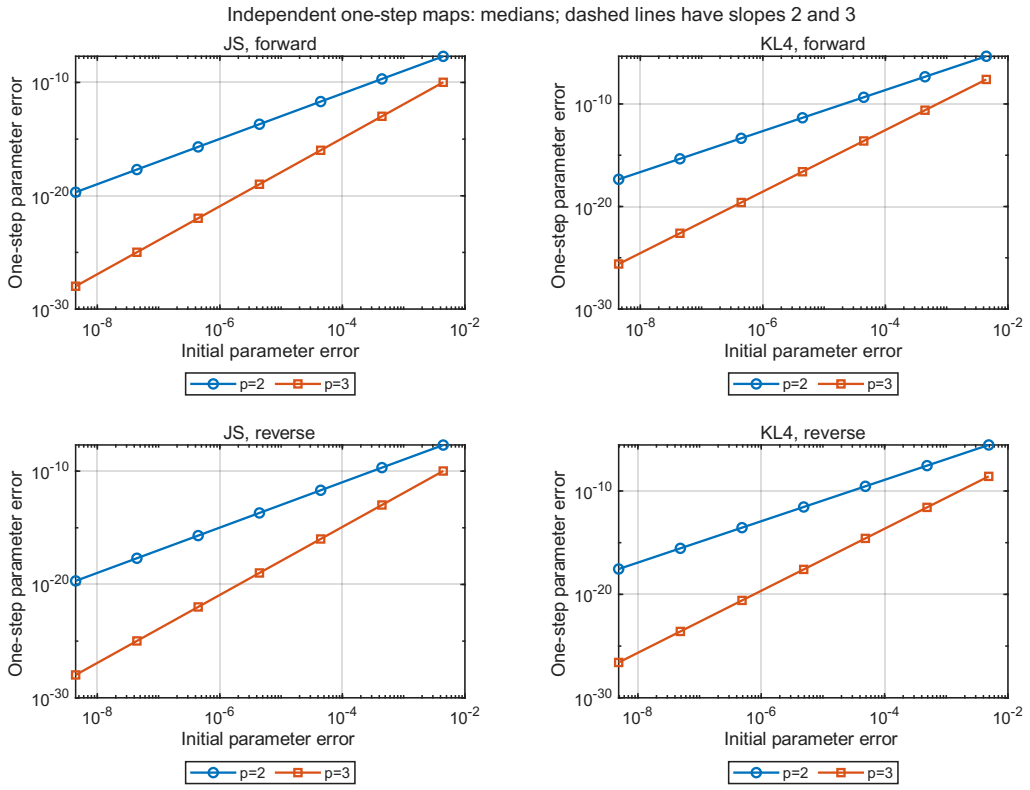


Figure 20: E5 independent one-step errors by generator (columns) and orientation (rows). Coordinates are medians over ten seeds and two signs. Dashed slopes two and three largely overlap the curves. Fits use each seed/sign's prespecified window, not the median curves.

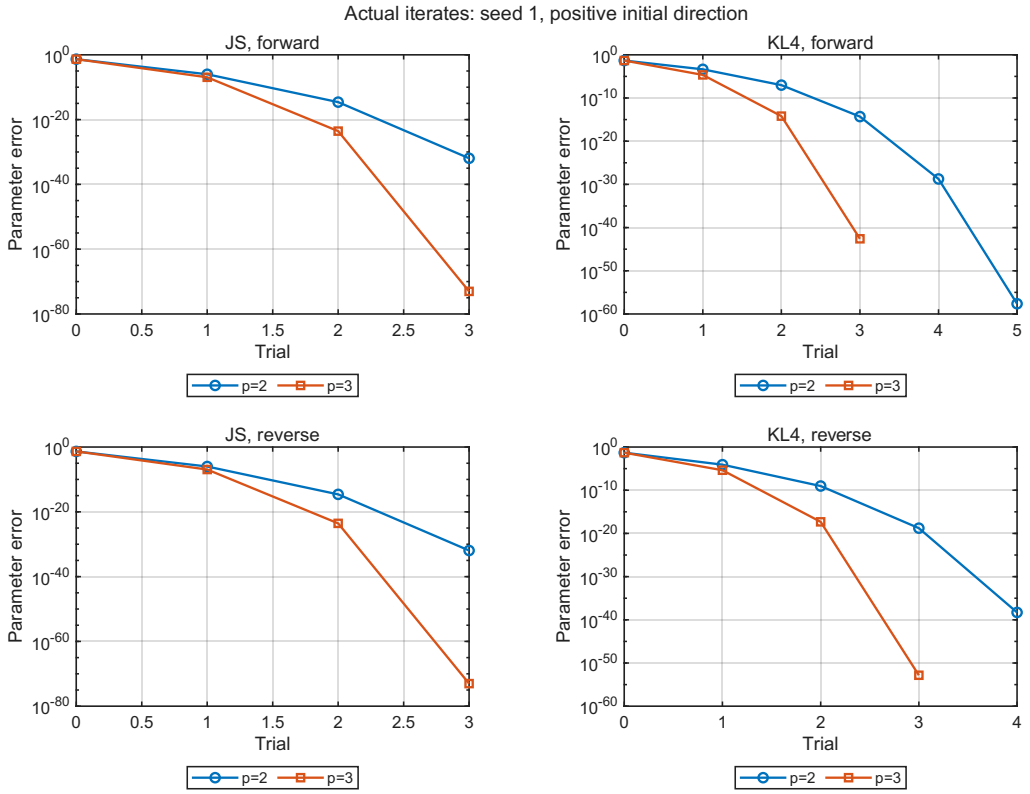


Figure 21: E5 parameter-error trajectories, seed 1, positive initial direction, 120-digit arithmetic. Panels separate generators and orientations. Table 7 summarizes all runs.

of 10^{-8} stalls at true-gradient norm 6.23×10^{-11}. Capping each step at one quarter of the branch step gives late error ratio 0.75 and order 1.00, exhausting the 60-trial budget. Residual Hessians differenced with fixed step 2×10^{-4} still yield stationarity, but the last observed order is 2.257. This short run shows third-jet sensitivity rather than a finite-difference asymptotic rate. With initial radius 10^{-3}, five truncated steps are followed by recovery to observed order 2.982.

Setting the second column of B to zero creates a flat direction. From $\theta^\star + (0.02, 0.1)^\top$, this control reaches gradient norm 3.64×10^{-58} while remaining 0.1 from the selected reference. Stationarity therefore need not identify a unique parameter without nondegeneracy. The five controls bring the total to 165 trajectories.

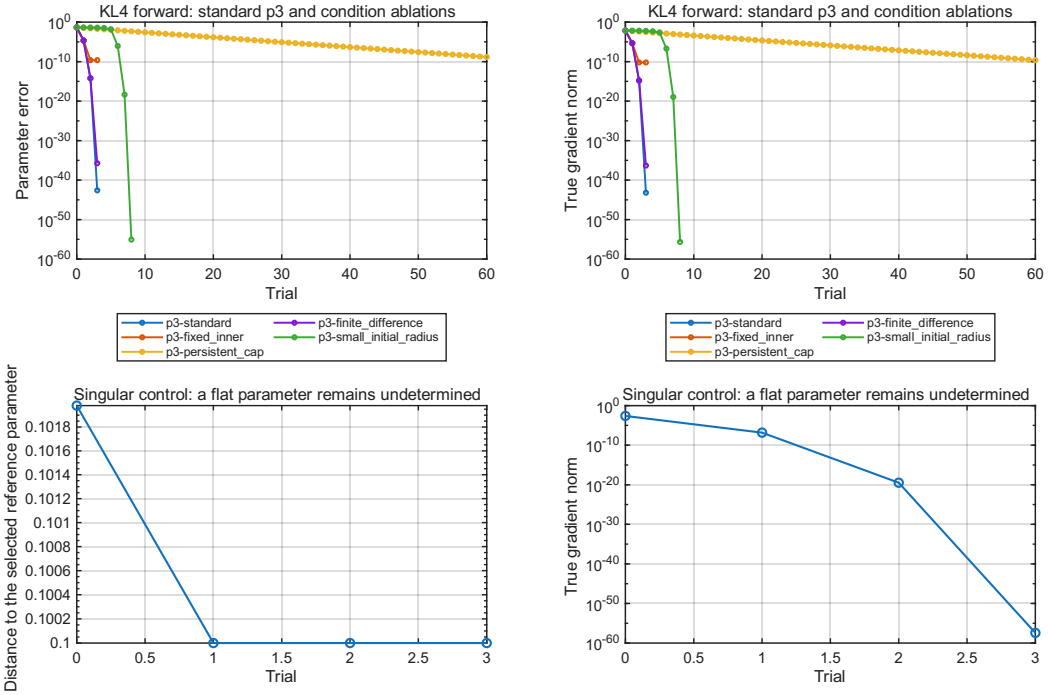


Figure 22: E5 individual condition ablations on the first forward KL+quartic problem. Top: parameter errors and true gradients for standard $p = 3$ and four controls. Bottom: a flat parameter direction permits vanishing gradient without convergence to the selected reference. The text reports each control's outcome.

C.6 Preference continuation and sensitivity prediction

Problem construction. E6 evaluates the nonsingular preference sensitivity formula in Section 7. We use the affine probability chart $q_\theta = q_0 + B\theta > 0$, with $(K, d, m) = (24, 3, 3)$. For seed $s \in \{1, \dots, 10\}$, the MATLAB random stream is `rng(61000+s,'twister')`. Draw $z \sim \mathcal{N}(0, 0.35^2 I_K)$ and a standard Gaussian matrix $A \in \mathbb{R}^{K \times d}$, and define

$$q_0 = \mathrm{softmax}(z), \qquad \widetilde{A} = A - \mathbf{1}(q_0^\top A), \qquad B = \frac{0.4\ \mathrm{diag}(q_0)\widetilde{A}}{\max_k \left\| \widetilde{A}_{k,:} \right\|_2}. \tag{C.7}$$

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

$$\vartheta_i = 1.2\, \frac{u_i}{\|u_i\|_2}, \qquad P_i = q_0 + B\vartheta_i, \qquad i = 1, 2, 3. \tag{C.8}$$

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 Section 10.1. 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

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

where the component objectives are defined in (1.1).

At subsequent preferences, the three initializers are

$$\begin{aligned} \theta_{j,0}^{\text{Cold}} &= 0, \\ \theta_{j,0}^{\text{Warm}} &= \theta_{j-1}, \\ \theta_{j,0}^{\text{Tangent}} &= \theta_{j-1} - H_{j-1}^{-1} G_{j-1} \delta\lambda_j. \end{aligned} \tag{C.9}$$

Since $\mathbf{1}^\top \delta\lambda_j = 0$, Tangent equals (7.1). 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 from zero target gradient norm 10^{-13}. The maximum reference gradient is 9.88×10^{-14}, and the minimum Hessian eigenvalue is 1.96×10^{-3}. Derivative, aggregation, center-calibration, and stable-difference checks all pass. The maximum normalized finite-difference discrepancy for the preference derivative is 8.24×10^{-9}.

The primary design has $10 \times 2 \times 2 \times 2 \times 3 \times (13 + 25 + 49) = 20{,}880$ point solves, indexed by seed, generator, divergence orientation, traversal, initializer, and grid. Three timing repetitions give 62,640 solves. All pass stationarity and reference-objective checks, with identical work counts across repetitions. The largest primary final gradient is 9.96×10^{-9}, and the largest absolute reference objective gap 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 $h = 0.04\, 2^{-r}$, $r = 0, \ldots, 5$,

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

$$e_{\mathrm{W}}(h) = \|\theta^\star(\lambda_0) - \theta^\star(\lambda_0 + \delta\lambda)\|\,, \tag{C.10}$$

$$e_{\mathrm{T}}(h) = \left\|\theta^\star(\lambda_0) - H_0^{-1}G_0\delta\lambda - \theta^\star(\lambda_0 + \delta\lambda)\right\|. \tag{C.11}$$

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.

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

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 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 5 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 24).

All corrected paths agree numerically with independent references (Figure 25). 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.

C.7 Template choice after second-order calibration

Compared models. E7 varies the nonlinear template while fixing $p = 2$. For $\mathcal{A}_\alpha = a_f \sum_i \lambda_i D_\alpha^\sigma(P_i, Q_\theta)$, the decomposition (4.16) becomes

$$m^{(2)}_{\alpha,\theta}(s) = T_{2,\theta}F(s) + \mathcal{A}_\alpha(\theta + s) - T_{2,\theta}\mathcal{A}_\alpha(s), \tag{C.12}$$

where $T_{2,\theta}$ is the quadratic Taylor polynomial at θ. The templates share the true quadratic expansion and differ in their higher-order terms.

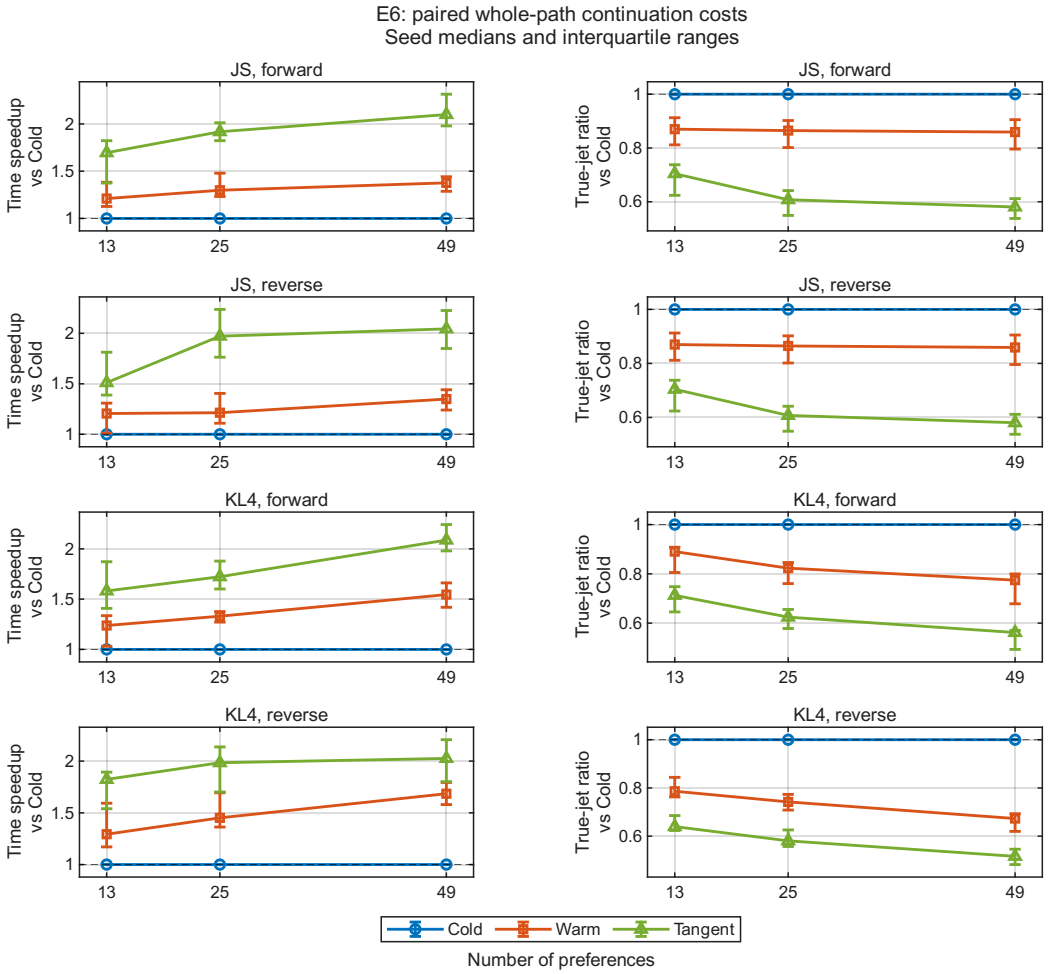


Figure 23: E6 whole-path costs by generator/orientation: runtime speedup against Cold (left) and true-jet pass ratio (right). Paired ratios use median times over three repetitions, then medians over traversals within seeds and medians/IQRs over ten seeds. Initialization and the common first solve are included.

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

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

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

Problem construction. E7 uses $Q_\theta = \mathrm{softmax}(b + B\theta)$ with $(K, d, m) = (128, 16, 32)$. As in E4, $b_k \sim \mathcal{N}(0, 0.25^2)$ and $B = 1.2U$, where U is the thin QR factor of a column-centered Gaussian matrix. Targets and initialization follow (C.2)–(C.3), with $\tau \in \{0.1, 0.5, 1, 1.5\}$ and Dirichlet$(1, \ldots, 1)$ preferences. Twenty seeds use MATLAB's Twister stream initialized by 71000 + seed. Primitive draws are shared across heterogeneities, generators, orientations, and methods. JS, KLChi, and KL4 in both orientations give 480 cases and 2,400 primary method runs. JS orientations coincide and are not independent replications.

Common optimization protocol. All runs use double precision, a true-gradient tolerance of 10^{-6}, and limits of 360 trials and 120 accepted steps. The initial, maximum, and minimum radii are 0.5, 2, and 10^{-12}. The thresholds are $(\eta_1, \eta_2) = (0.1, 0.75)$. Rejection halves the radius, while $\rho \geq \eta_2$ doubles it up to the maximum. The Cauchy-decrease check uses coefficient 0.1.

All methods use at most 15 regularized Newton iterations with ball projection, Armijo backtrack-

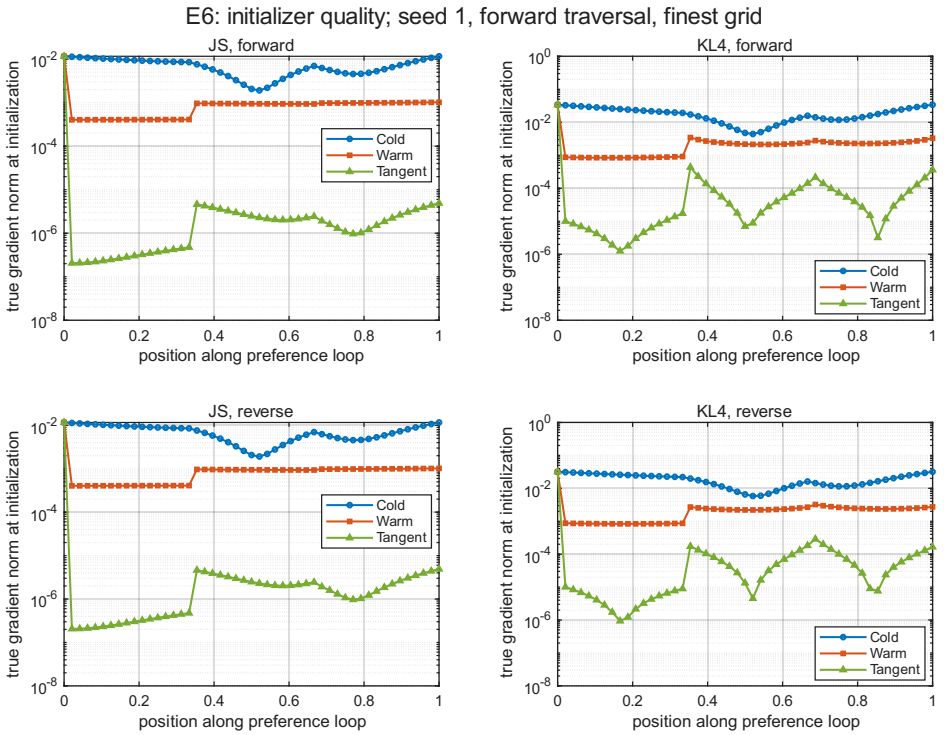


Figure 24: E6 pre-correction gradient norms, seed 1, forward traversal of the 49-point loop, both divergence orientations. All methods share the first cold initialization. The reference line is the corrector tolerance 10^{-8}.

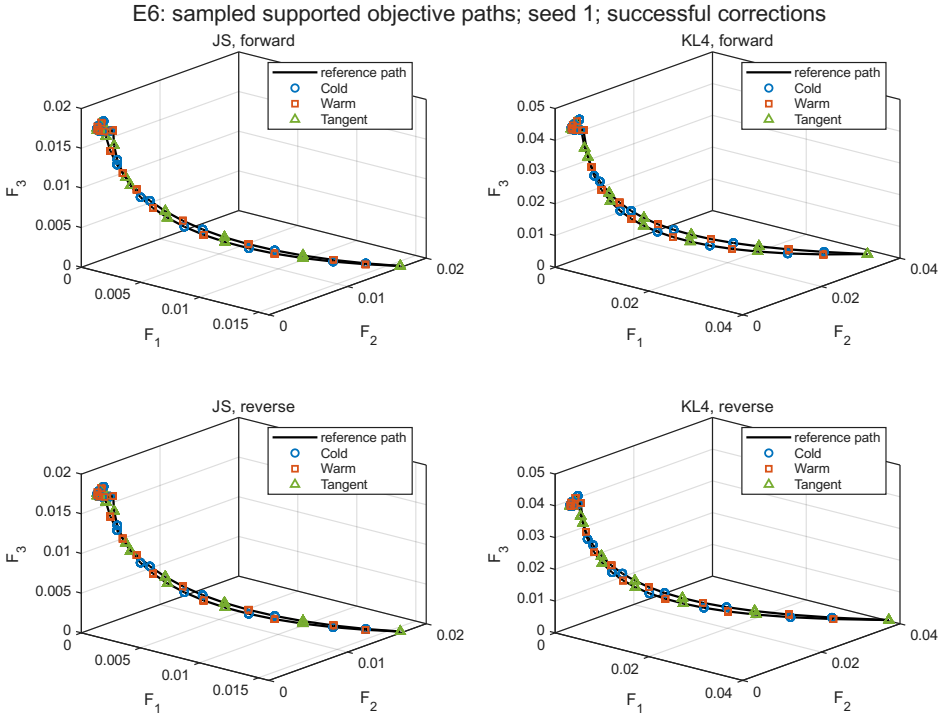


Figure 25: E6 corrected objective paths, seed 1, finest grid, forward traversal. The black line connects independent references. Staggered markers show Cold, Warm, and Tangent solutions. All sampled points enter validation. The loop does not cover the full Pareto surface.

ing, a Cauchy candidate, and model-gradient threshold $\max\{10^{-10}, 0.01\,\|\nabla m(0)\|\}$. Acceptance uses the true objective, and success uses its gradient. An uncalibrated run that stops at surrogate stationarity but fails the true-gradient test is unsuccessful. Taylor-p2 uses the same inner routine to isolate template effects, rather than a separate quadratic solver.

Independent Taylor-p2 references start from zero, use tolerance 10^{-10}, and have larger budgets. All 480 converge, with maximum gradient norm 9.71×10^{-11} and minimum observed Hessian eigenvalue 3.95×10^{-3}. They are local numerical references, not global-optimality certificates.

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

$$\begin{aligned} e_v(\theta, s) &= \frac{|F(\theta + s) - m_\theta(s)|}{1 + |F(\theta)|}, \\ e_g(\theta, s) &= \frac{\|\nabla F(\theta + s) - \nabla_s m_\theta(s)\|}{1 + \|\nabla F(\theta + s)\|}. \end{aligned} \tag{C.13}$$

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

Figures 6 and 26 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. The remaining 247 are unavailable in the records. Neither slope agreement nor R^2 determines eligibility.

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 27): better model accuracy does not yield lower runtime here. Figure 28 gives representative gradient histories. Success statistics use all seeds.

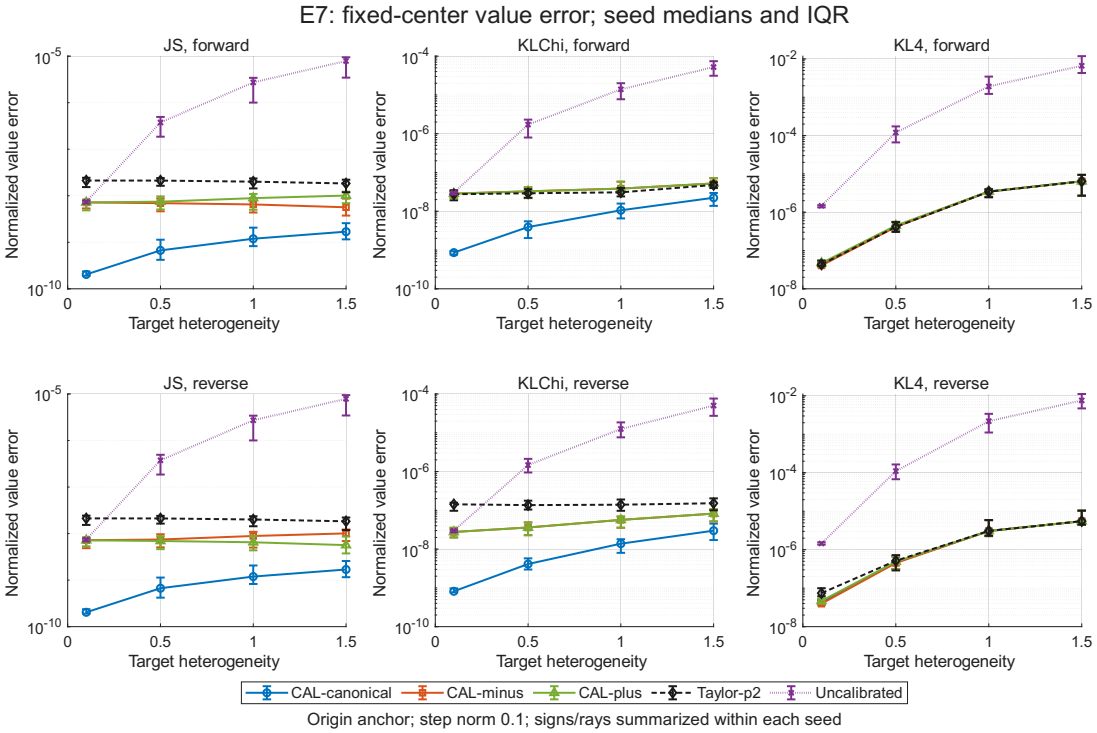


Figure 26: E7 normalized value errors with the probes and aggregation of Figure 6. Canonical median errors improve on JS and KL+chi-square. Calibrated and Taylor curves nearly coincide for KL+quartic at larger heterogeneities. The horizontal axis varies targets, not step size.

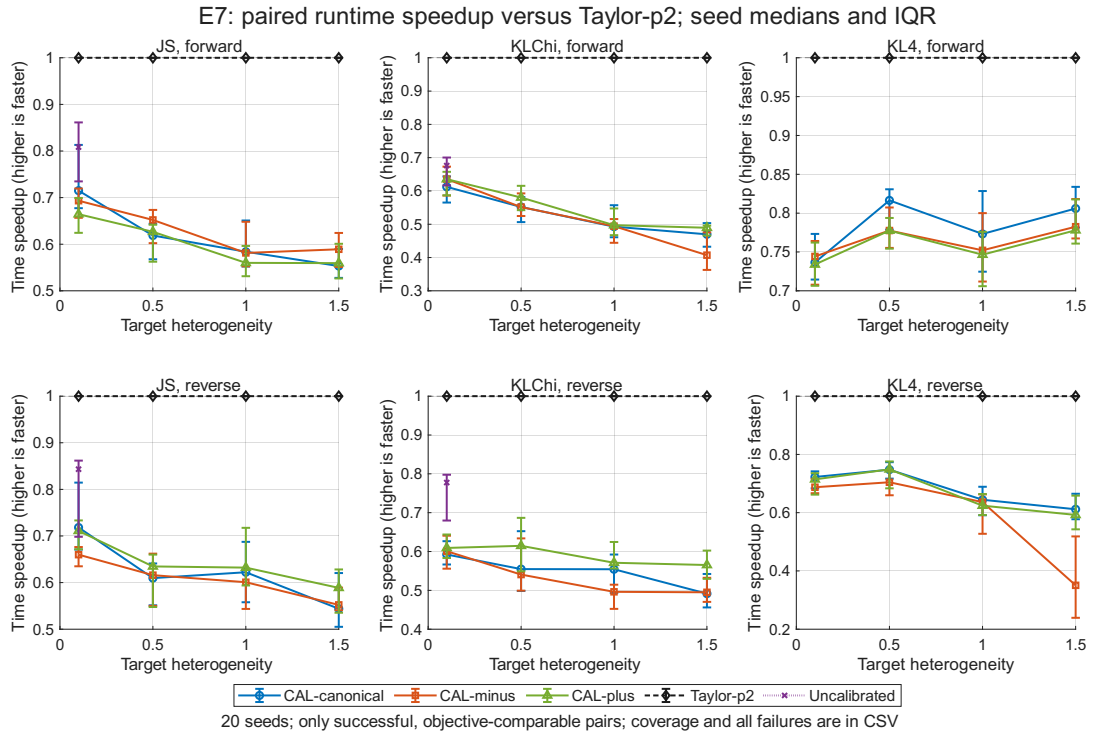


Figure 27: E7 paired speedups against Taylor-p2 (above one is faster), using median times over three repetitions and seed medians/IQRs. Each calibrated group has 20 successful, objective-comparable pairs. Uncalibrated pairs exist only for JS and KL+chi-square at $\tau = 0.1$. Missing points indicate no eligible pairs.

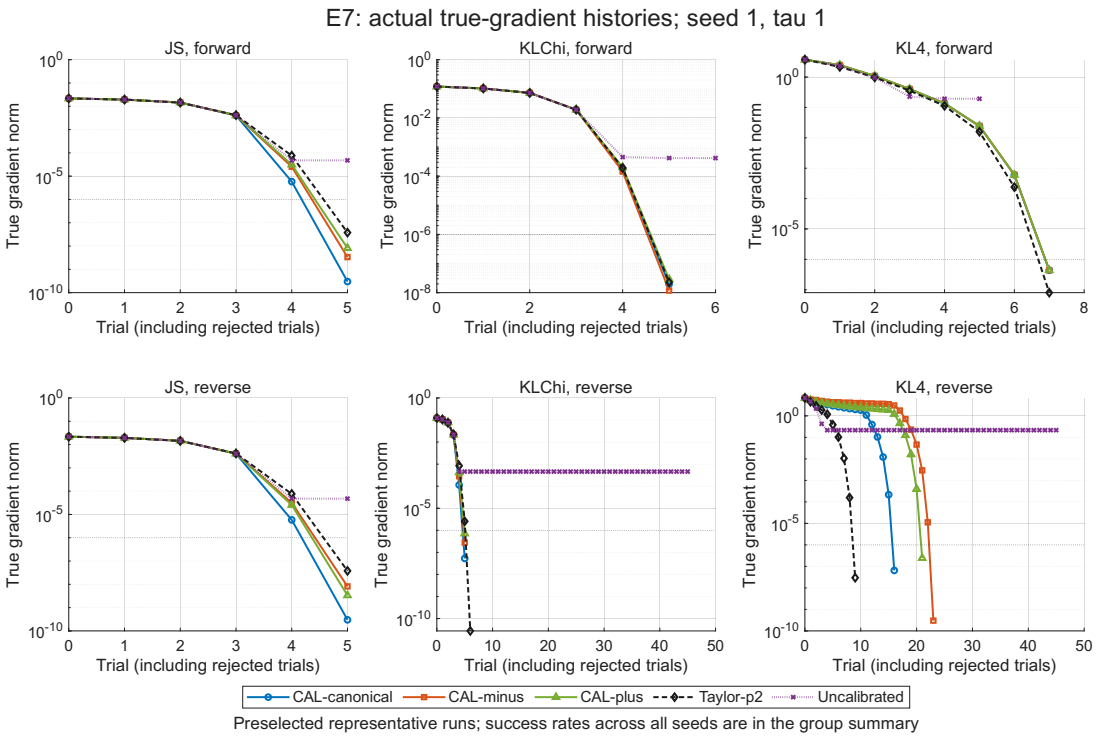


Figure 28: E7 true-gradient histories, preselected seed 1 and $\tau = 1$, including rejected trials. These representative histories do not measure asymptotic local order or replace all-seed success statistics.

C.8 E8: learned latent distribution matching and generation

Restricted model family. Each target has six Gaussian components, while Q_θ has two with a fixed shared covariance. Even the exact forward-KL aggregate may lie outside this family. For example, preference $(.90, .05, .05)$ gives $.645R_0 + .220R_1 + .135R_2$. These are weights of fitted latent laws, not prescribed image-class frequencies.

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.

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

$$\begin{aligned} P_1 &= .7R_0 + .2R_1 + .1R_2, \\ P_2 &= .1R_0 + .7R_1 + .2R_2, \\ P_3 &= .2R_0 + .1R_1 + .7R_2. \end{aligned} \tag{C.14}$$

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.

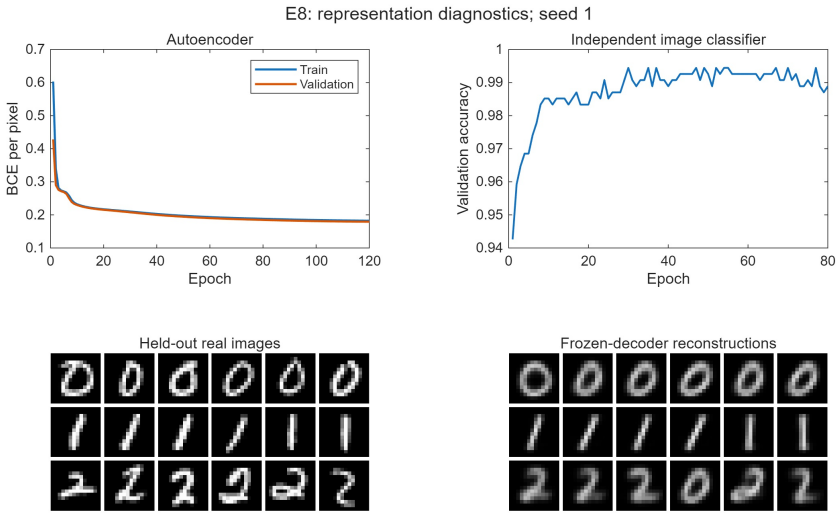


Figure 29: E8 representation diagnostics for seed 1: training and validation reconstruction loss, validation accuracy of an independent image classifier, and held-out inputs with frozen-decoder reconstructions.

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

$$\mathcal{A}(\theta) = 2\left(1 - \sum_x H_\lambda(x)\sqrt{q_\theta(x)}\right), \qquad H_\lambda(x) = \sum_i \lambda_i \sqrt{p_i(x)}, \tag{C.15}$$

where p_i and q_θ denote probability masses on the quadrature grid. This expression retains the scale and constant in Theorem 4.2.

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 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.

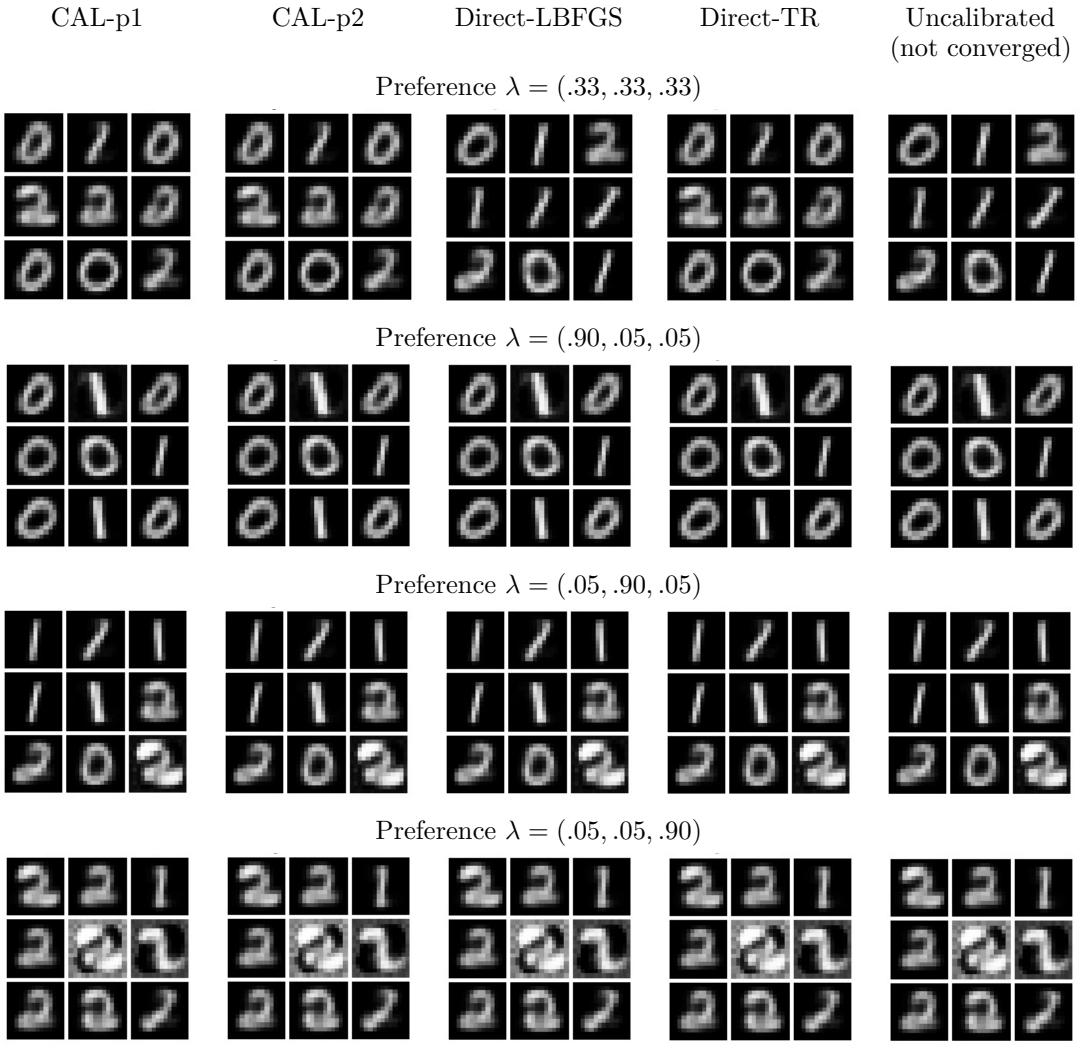


Figure 30: E8 saved decoded samples, four JS preferences, seed 1. Columns compare methods. All displayed uncalibrated endpoints failed to converge. Mosaics retain the saved samples, cropping only surrounding whitespace. CAL-p2 also appears in Figure 11.

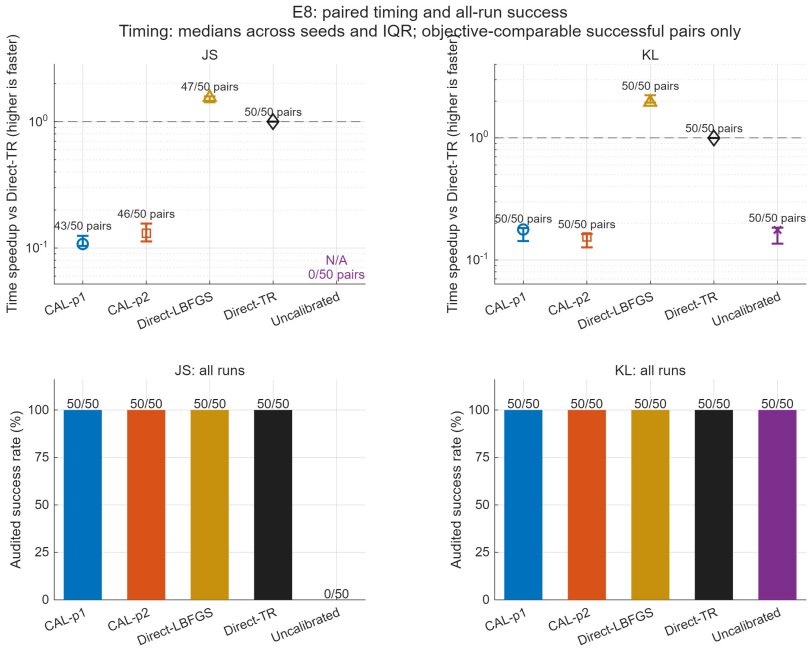


Figure 31: E8 paired runtime and all-run audited success. Timing uses only successful, objective-comparable pairs and reports pair counts. Success uses all attempted cases. N/A denotes absence of eligible timing pairs.

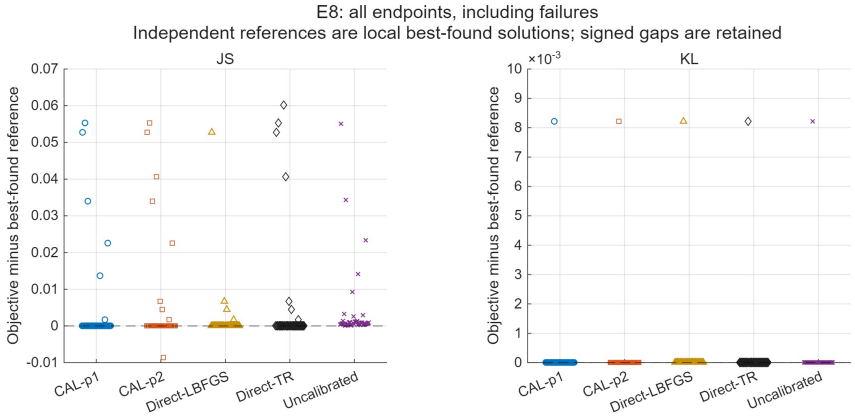


Figure 32: E8 signed objective gaps to independent local best-found references, including failed endpoints. A negative value means that the tested run improved on its reference.

C.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 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.

Baseline scope. CAL-p2 and Direct-TR use true values, gradients, and Hessians. L-BFGS uses values and gradients. Adaptive weighting or constrained preferences change the problem (Section 3). Full-exponential-family Pareto constructions (Zhang et al., 2025) do not directly solve restricted-mixture JS matching. The KL baselines already aggregate exactly. Target-subsampled solvers are untested here.

The derivative-demand control (Stage A) is described in Section 10.4 and Figure 10.

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 33 and 9.

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

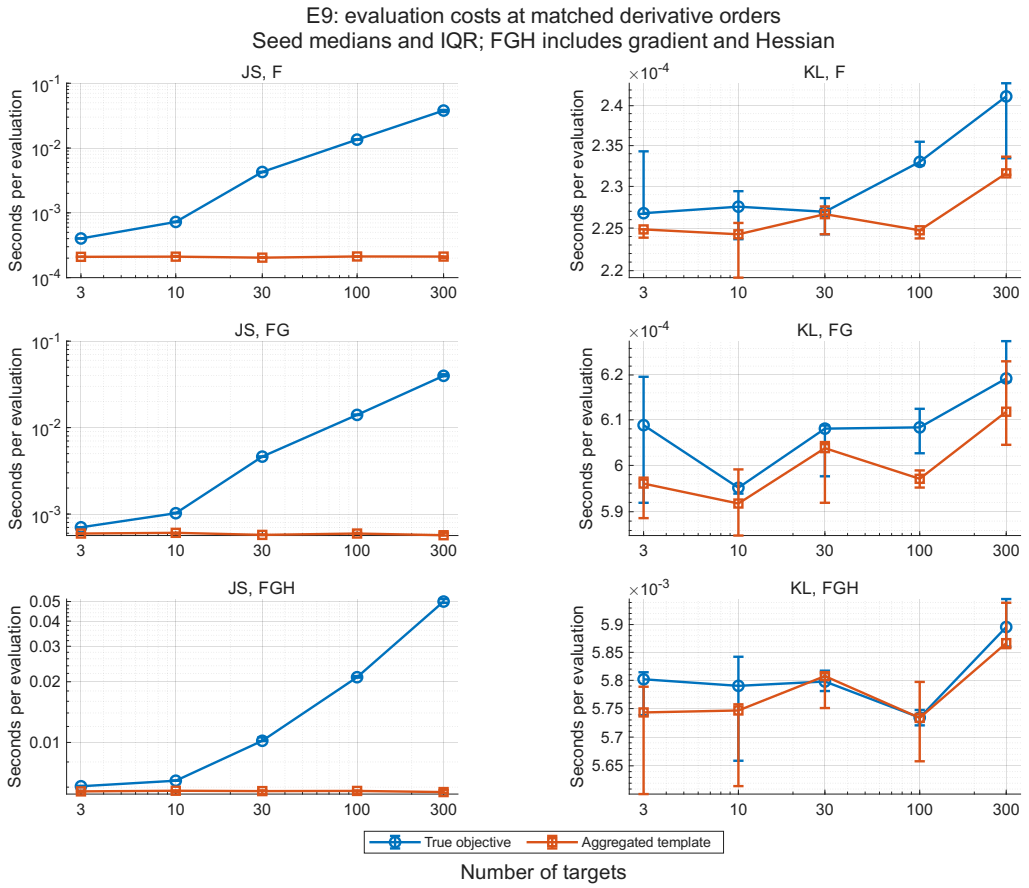


Figure 33: E9 absolute evaluation costs at matched derivative orders, summarized by seed medians and IQRs. FGH includes both gradients and Hessians. Predictor preprocessing is accounted for separately and included in total solves.

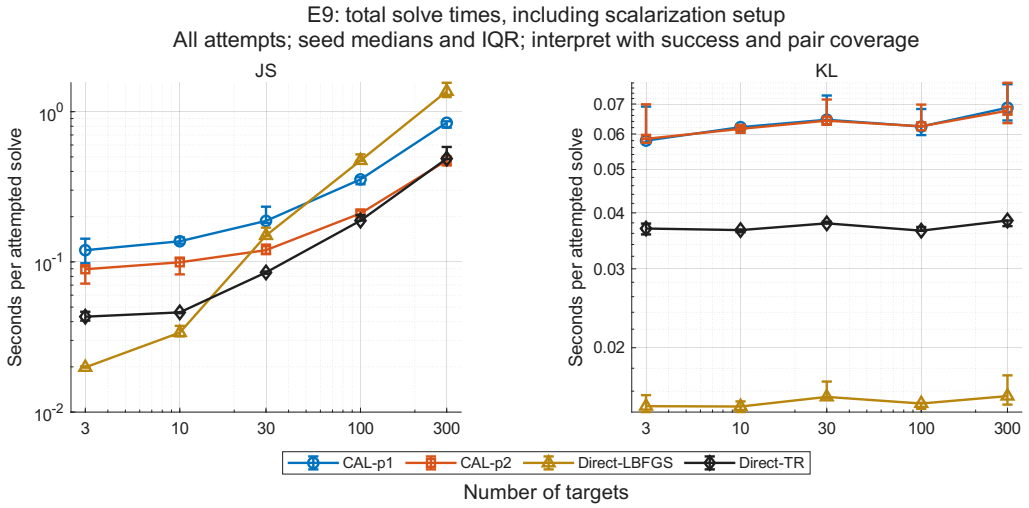


Figure 34: E9 total solve times including preference-dependent scalarization setup, over all attempts. Because different methods can reach different local solutions, these marginal medians supplement rather than replace paired speedups.

B, a pair is eligible if both runs pass the endpoint audit and

$$\left|F_\lambda(\theta^M_{s,k}) - F_\lambda(\theta^B_{s,k})\right| \le 2 \times 10^{-6}. \tag{C.16}$$

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.

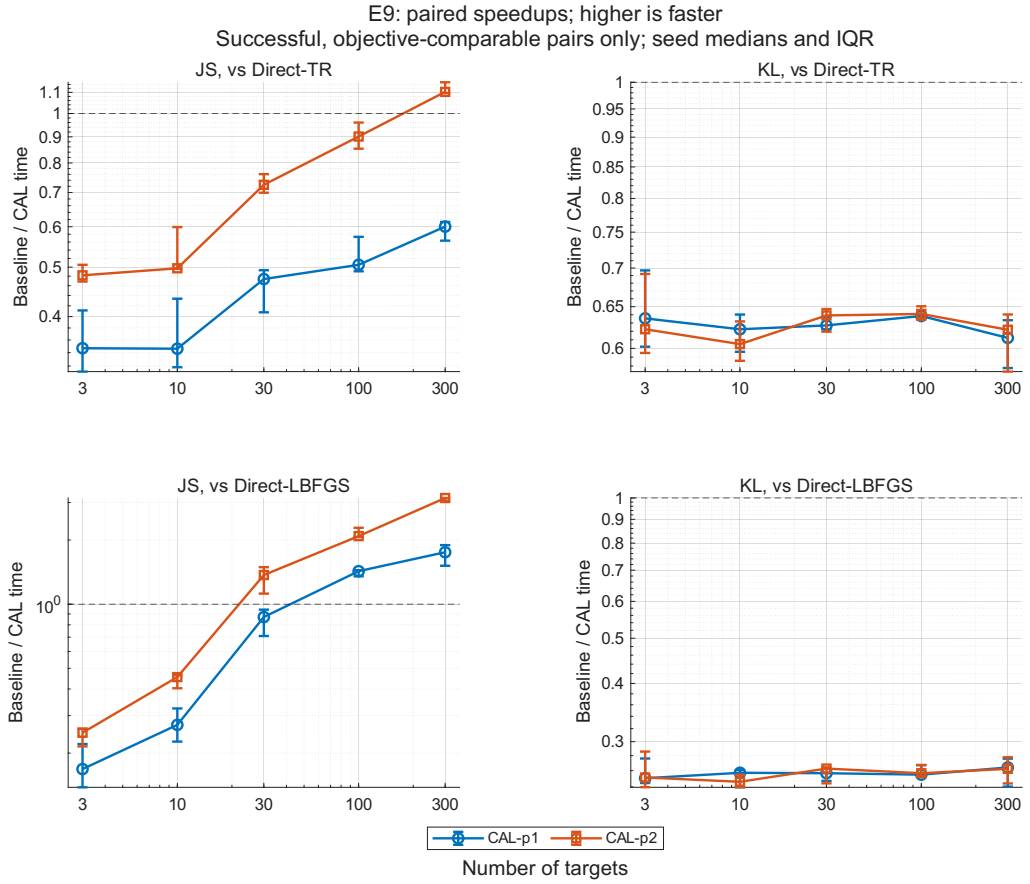


Figure 35: E9 full paired speedups for JS and the exact-aggregation KL control, against Direct-TR and Direct-LBFGS. Curves summarize successful objective-comparable pairs by seed medians and IQRs. CAL-p2 JS pair coverage is given in Table 4. Raw records retain every attempted case.

Diagnostic cost breakdown. Separate instrumented replays measure preparation, true/template evaluations, linear algebra, and other work (Figure 36). 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.

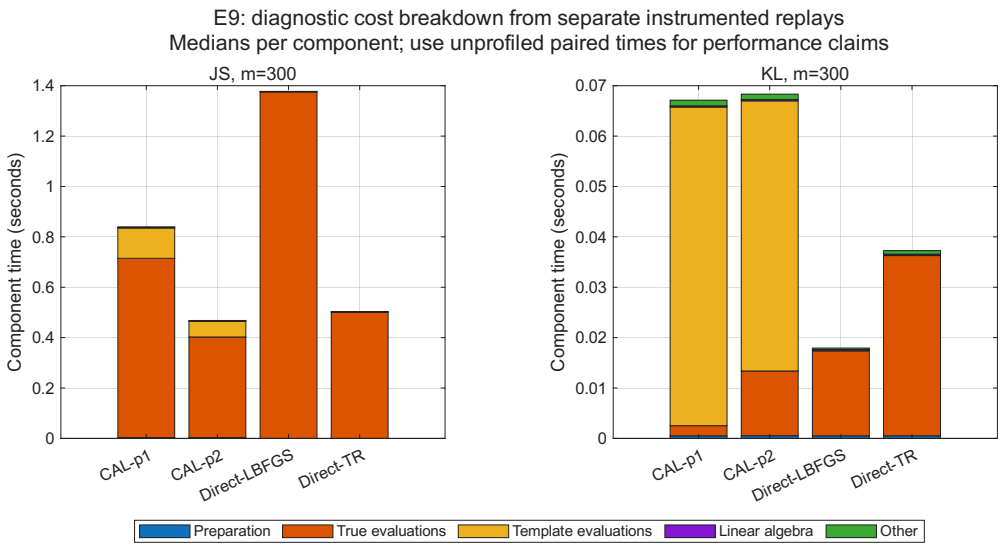


Figure 36: E9 diagnostic cost breakdown at $m = 300$ from separate instrumented replays. Each component is summarized by its median. Performance claims use the unprofiled paired timing experiment.

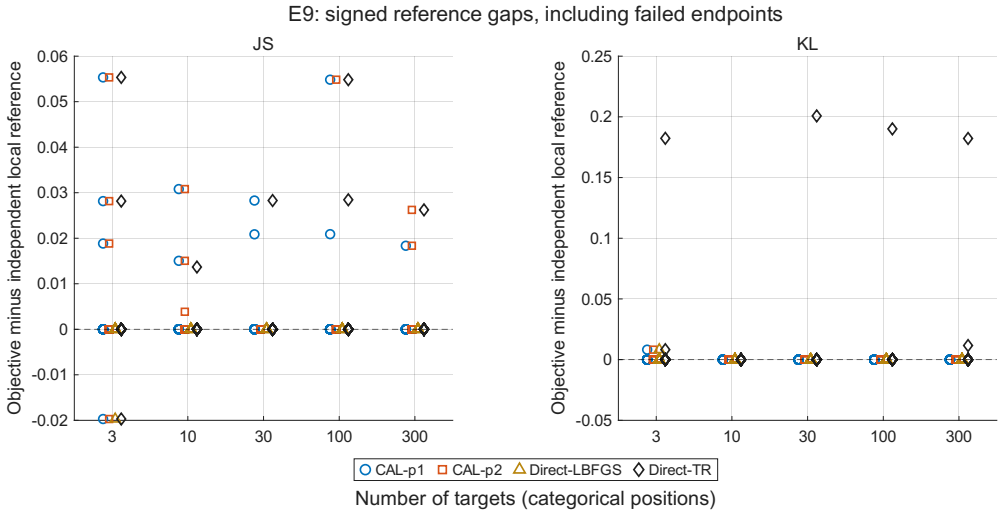


Figure 37: E9 signed objective gaps relative to independent local references, retaining all endpoints. Negative gaps are possible because references are best-found local solutions. These plots report solution variation separately from stationarity success and timing-pair coverage.

C.10 E10: canonical geometry and the leading local error

E10 tests the leading-error cancellation in Proposition 6.5 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 in Section 10.1. 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

$$\|\nabla m_\theta(s)\| \le 0.01 \|\nabla F(\theta)\|^3 . \tag{C.17}$$

This strengthens Theorem 6.4's $p = 2$ forcing to reveal the coefficient in (6.7). 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 (C.17), 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 7 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 2 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).

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.

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.

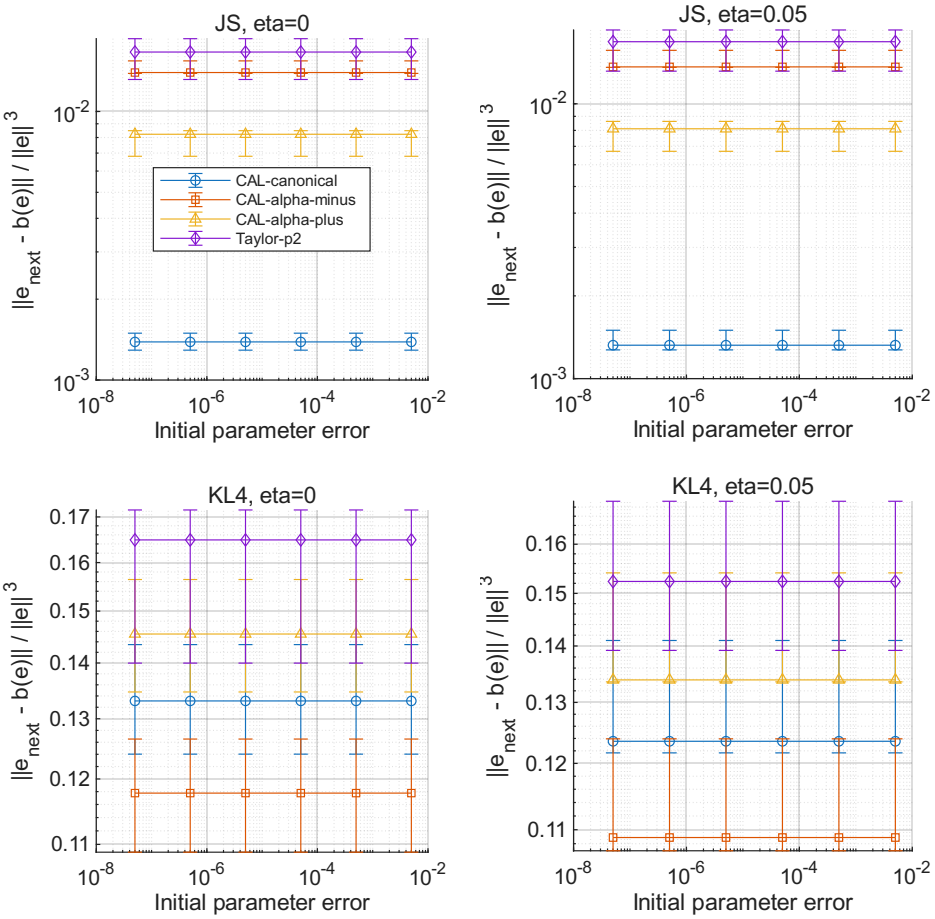


Figure 38: E10 expansion audit: subtracting the predicted quadratic vector leaves an approximately constant $\|e_+ - b(e)\| / \|e\|^3$. Panels show $\eta = 0, 0.05$. Fits cover all four heterogeneities. Third derivatives enter only the offline prediction and audit.