PACT: LEARNING COUPLINGS UNDER EMPIRICAL MARGINAL CONSTRAINTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Observations from different sources often provide rich information about individual distributions but limited evidence of how outcomes correspond. Learning reliable joint models from such partially paired data is a common challenge in conditional prediction and risk assessment. We introduce Pair-Aligned Conditional Transport (PACT), which combines relation learning with marginal calibration. We show that empirical marginal constraints during training not only shape model outputs but also systematically alter relation learning: adding independent endpoint observations can even reverse the direction of the leading learning bias. Building on this finding, we establish a theoretical link between objective bias and parameter displacement and derive a correction for first-order expected objective bias. Controlled experiments validate the predictions of the analysis, while applications demonstrate PACT's value for conditional prediction and cross-policy risk modeling. This work provides a statistical basis for understanding how dispersed observations jointly influence the learning of joint distributions.

1 INTRODUCTION

Many prediction and decision-making tasks require understanding both how individual outcomes are distributed and how they relate to one another, yet available observations often lack the pairings needed to learn these relationships. For example, cross-domain prediction requires identifying which target outcomes correspond to a given source input. Comparing queueing policies requires relating their delays under shared disturbances, while comparing hedging strategies requires relating their losses along the same market path. Even when the individual outcome distributions are known, different relationships between outcomes can lead to different conditional predictions and risks. These tasks therefore require a joint model that captures the correspondence between outcomes, as illustrated in Figure 1.

Conditional flows learn from pairs (Winkler et al., 2019), while unpaired matching can leave correspondence ambiguous despite cycle consistency (Zhu et al., 2017; Moriakov et al., 2020). Partial-pairing methods add aligned supervision to translation or bridge learning (Tripathy et al., 2018; Theodoropoulos et al., 2025); inverse transport and dependence models estimate association under marginal constraints (Chiu et al., 2022; Persiianov et al., 2026; Sei & Yano, 2026). These foundations raise a further question: how do finite-sample marginal constraints change the objective optimized in relation learning?

To address this problem, we introduce *Pair-Aligned Conditional Transport* (PACT), which translates task-specific relational information into compatibility scores describing how outcomes correspond, then uses marginal calibration to construct a joint model consistent with the observed marginal distributions. Building on this formulation, we develop a bias analysis of empirical calibration to explain how the combination of paired observations and independent endpoint samples affects relation learning, and derive a correction that removes first-order bias in the expected learning objective. PACT combines joint modeling from incompletely paired data with a statistical analysis of calibration, providing a methodological foundation for conditional prediction and cross-policy risk assessment. Our contributions are:

1. **The PACT joint-modeling architecture.** We integrate limited paired information with separately observed marginals to support conditional prediction and cross-policy risk assessment.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

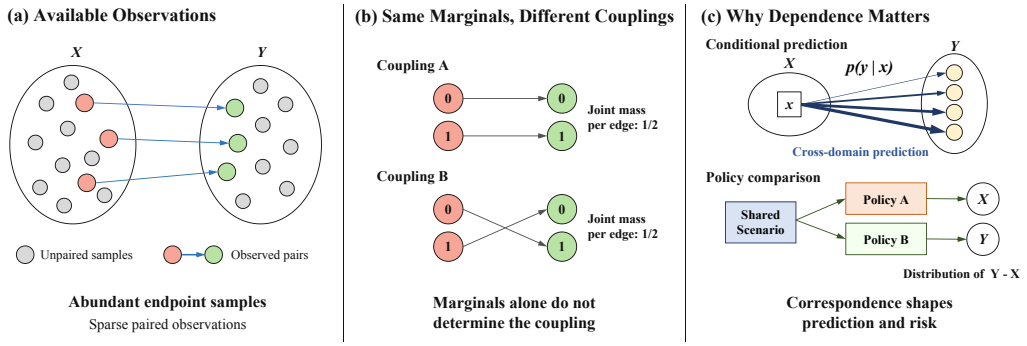

Figure 1: **Why coupling information matters.** Sparse pairs complement endpoint samples (a); identical marginals admit different couplings (b), with different conditional predictions and distributions of paired outcome differences (c). Arrow widths indicate conditional probabilities.

2. **The learning effects of empirical calibration.** We reveal how sample composition changes the direction of the leading bias and establish a theoretical link between empirical marginal constraints and local parameter displacement.
3. **First-order objective-bias cancellation.** We construct a ratio-preserving split correction that eliminates first-order expected objective bias from empirical calibration under the stated regularity conditions.

2 Related Work

Paired and unpaired learning. Paired conditional density models include mixture density networks (Bishop, 1994) and normalizing flows (Winkler et al., 2019; Papamakarios et al., 2021). CycleGAN combines distribution matching with cycle consistency (Zhu et al., 2017), but can admit multiple correspondences (Moriakov et al., 2020). Partial-pairing methods add paired supervision to adversarial translation (Tripathy et al., 2018) or guide Schrödinger bridges with pre-aligned pairs (Theodoropoulos et al., 2025). PACT studies how relational and endpoint sampling affect calibrated learning.

Inverse transport and dependence models. Fixed-marginal KL fitting builds on contingency-table estimation and information projections (Ireland & Kullback, 1968; Csiszár, 1975); Peyré & Cuturi (2019) survey computational entropic transport. Chiu et al. (2022) infer costs up to coupling-preserving equivalence; Persiianov et al. (2026) connect semi-supervised likelihood to inverse entropic transport. Association models use odds ratios with unspecified marginals (Osius, 2009). Sei & Yano (2026) develop semiparametric profile likelihood under empirical marginal constraints, including bias from same-index pairs in the empirical product measure. PACT derives calibrated pair-likelihood bias and local parameter displacement under mixed paired and independent endpoint sampling.

Minibatch transport and likelihood correction. Empirical transport theory studies plan sensitivity and limit laws (Klatt et al., 2020), and stability under marginal and cost perturbations (Ghosal et al., 2022), alongside sampling-error bounds for Sinkhorn divergences (Genevay et al., 2019). Fatras et al. (2020) establish gradient unbiasedness relative to the expected minibatch Wasserstein objective. Profile-likelihood adjustments account for nuisance fitting (Barndorff-Nielsen, 1983; Cox & Reid, 1987; McCullagh & Tibshirani, 1990); Saleh (2024) reviews a canonical log-determinant correction. Our paired finite canonical objective connects to this adjustment through its first-order score bias (Appendix B.11). Jackknife corrections cancel leading sample-size bias (Quenouille, 1956), including split-panel corrections for estimators and likelihood objectives (Dhaene & Jochmans, 2015). PACT derives the sampling-specific bias coefficient and its spectral direction changes; preserving the endpoint-to-pair ratio when splitting and recalibrating batches yields first-order cancellation through the classical jackknife combination.

3 PACT: MARGINAL CALIBRATION AND RELATION LEARNING

PACT combines correspondence and marginals, then links empirical calibration to objective bias and parameter displacement to derive a first-order correction.

3.1 PRELIMINARY

Couplings and partially paired observations. For $X \in \mathcal{X}$, $Y \in \mathcal{Y}$ with marginals μ, ν, write $\Gamma(\mu,\nu) = \{\pi : \pi_X = \mu, \pi_Y = \nu\}$ and $\pi^\star \in \Gamma(\mu,\nu)$. Pairs $(x_i, y_i) \sim \pi^\star$ reveal correspondence and marginals; independent $\widetilde{x}_j \sim \mu$, $\widetilde{y}_k \sim \nu$ add marginal information. The alignment rule defines $\pi^\star$ and paired quantities such as $Y - X$.

On finite discrete supports, endpoint weights $p \in \Delta_m$, $q \in \Delta_n$ and a coupling matrix $\Pi \geq 0$ satisfy $\Pi \mathbf{1}_n = p$ and $\Pi^\top \mathbf{1}_m = q$. Hats denote empirical estimates.

Marginal calibration. For positive marginals p, q, a finite score matrix $S \in \mathbb{R}^{m\times n}$, and temperature $\tau > 0$, marginal calibration converts the scores into a joint distribution whose row and column sums match p and q, respectively:

$$\mathcal{C}_{p,q}(S) = \arg\max_{\Pi \in \Gamma(p,q)} \{\langle \Pi, S\rangle - \tau\, \mathrm{KL}(\Pi \| pq^\top)\}. \tag{1}$$

This entropy-regularized transport problem with cost $-S$ has a unique solution (Cuturi, 2013); its exponential form and computation appear in Appendix F.1.

3.2 PACT JOINT-MODELING ARCHITECTURE

PACT turns correspondence into joint predictions. Let $\widehat{S}_{ij} = s_{\widehat{\theta}}(x_i, y_j)$ and estimate endpoint weights $\widehat{p}, \widehat{q}$ from paired or independent observations. Then

$$\widehat{\Pi} = \mathcal{C}_{\widehat{p},\widehat{q}}(\widehat{S}), \qquad \widehat{\Pi}\mathbf{1}_n = \widehat{p}, \qquad \widehat{\Pi}^\top \mathbf{1}_m = \widehat{q}. \tag{2}$$

For $\widehat{p}_i > 0$, conditional probabilities are $\widehat{\Pi}_{ij}/\widehat{p}_i$; drawing $(I, J) \sim \widehat{\Pi}$ yields paired quantities $h(x_I, y_J)$, including loss differences. Calibration preserves the cross-ratios of $\exp(\widehat{S}/\tau)$ while matching empirical marginals (Appendix B.6); prediction at new inputs appears in Appendix F.5.

Compatibility comes from paired learning or structural knowledge. Figure 2 shows calibrated-likelihood training. Queueing instead uses structural replay and endpoint matching to form a boundary coupling (Appendix F.2).

3.3 CALIBRATED-LIKELIHOOD LEARNING

Calibrated likelihood assigns high probability to observed pairs after enforcing marginals. On finite discrete supports, let $\Pi_\theta = \mathcal{C}_{p,q}(S_\theta)$. The population loss is

$$\mathcal{L}_\infty(\theta) = -\sum_{ij} \pi^\star_{ij} \log[\Pi_\theta]_{ij} = H(\pi^\star) + \mathrm{KL}(\pi^\star \| \Pi_\theta). \tag{3}$$

An attained global minimum recovers representable $\pi^\star$ or its KL projection (Appendix B); the paired finite canonical case admits a profile-likelihood interpretation (Appendix B.11).

A batch contains b aligned pairs and u additional independent observations from each endpoint. Let η be the empirical pair law, $z^X = \eta \mathbf{1}_n$, $z^Y = \eta^\top \mathbf{1}_m$, and let v^X, v^Y denote the additional endpoint frequencies. The pooled marginals are

$$\widehat{p} = \frac{bz^X + uv^X}{b+u}, \qquad \widehat{q} = \frac{bz^Y + uv^Y}{b+u}, \tag{4}$$

with the additional terms omitted when $u = 0$. The empirical loss is

$$\widehat{\mathcal{L}}^{\mathrm{NLL}}_{b,u}(\theta) = -\sum_{ij} \eta_{ij} \log[\mathcal{C}_{\widehat{p},\widehat{q}}(S_\theta)]_{ij}. \tag{5}$$

Pairs supervise the loss; independent endpoints enter through calibration, which is differentiated. Zero-mass states are excluded; observed pairs have positive endpoint mass.

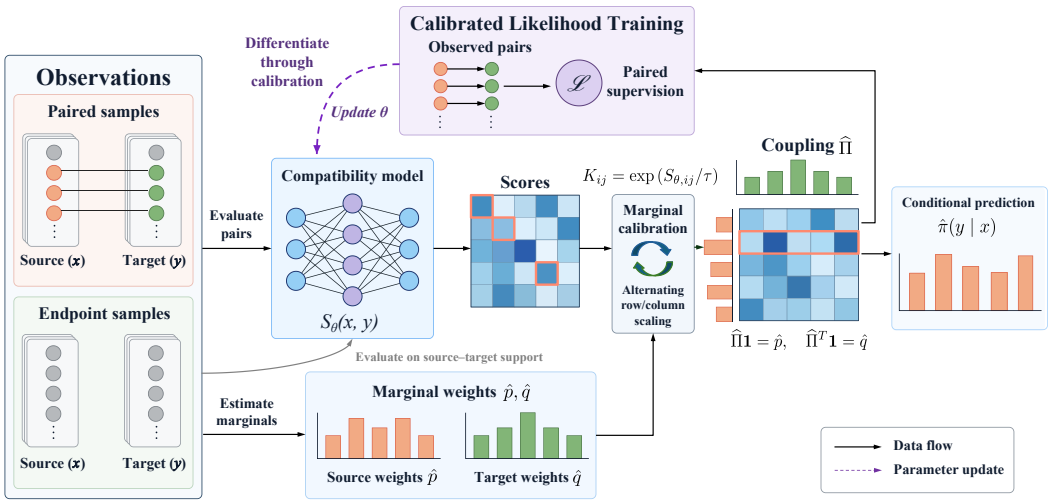

Figure 2: **Calibrated-likelihood training in PACT.** Pairs supervise calibrated joint probabilities; gradients pass through calibration. Row normalization gives conditional predictions. Solid/dashed arrows show data flow/updates; heatmaps and bars are schematic.

3.4 Sampling structure, bias, and parameter displacement

Shared observations affect both marginal estimates and paired supervision.

Let $\rho = u/b$ remain fixed and average over both sample sources. Absorb temperature into $f_\theta = S_\theta/\tau$ and define

$$\Phi_\theta(p, q) = \max_{\Pi \in \Gamma(p,q)} \{\langle \Pi, f_\theta \rangle - \mathrm{KL}(\Pi \| pq^\top)\}. \tag{6}$$

Up to parameter-independent terms, the population objective is $\mathcal{R}_\infty(\theta) = -\langle \pi^\star, f_\theta \rangle + \Phi_\theta(p, q)$. In free coordinates $r = (p_{1:m-1}, q_{1:n-1})$, let $z, \widehat{r}$ denote paired and pooled marginals. Let $\widehat{\mathcal{L}}_{b,u}$ denote Equation (5) after removing parameter-independent terms. Appendix B.1 gives its value-function form for the subsequent bias analysis.

How shared observations enter the bias. Write $p_- = p_{1:m-1}$, $q_- = q_{1:n-1}$ and define $\Sigma_p = \mathrm{diag}(p_-) - p_- p_-^\top$, $\Sigma_q = \mathrm{diag}(q_-) - q_- q_-^\top$, and $\Sigma_{pq} = \pi^\star_{1:m-1,1:n-1} - p_- q_-^\top$. Let

$$D = \begin{pmatrix} \Sigma_p & 0 \\ 0 & \Sigma_q \end{pmatrix}, \qquad O = \begin{pmatrix} 0 & \Sigma_{pq} \\ \Sigma_{pq}^\top & 0 \end{pmatrix}, \qquad H_\theta = \nabla_r^2 \Phi_\theta(p, q). \tag{7}$$

Here D records marginal sampling variance, and O records cross-covariance from shared complete pairs. The sampling structure gives

$$\mathrm{Cov}(\widehat{r}) = \frac{D}{b+u} + \frac{bO}{(b+u)^2}, \qquad \mathrm{Cov}(z, \widehat{r}) = \frac{D+O}{b+u}. \tag{8}$$

Theorem 1 (First-order minibatch bias)**.** *Assume fixed finite discrete supports with positive true marginals and $u = \rho b$ for fixed $\rho \geq 0$. On a compact parameter neighborhood, assume that f_θ and $\Phi_\theta(r)$ have continuous, uniformly bounded joint derivatives in θ, r through third order near the true marginals. Then, uniformly on that parameter neighborhood,*

$$\mathbb{E}\widehat{\mathcal{L}}_{b,\rho b}(\theta) = \mathcal{R}_\infty(\theta) + \frac{A_\rho(\theta)}{b} + o(b^{-1}), \tag{9}$$

$$A_\rho(\theta) = \frac{1}{2}\left[\frac{\mathrm{tr}(H_\theta D)}{1+\rho} + \frac{1+2\rho}{(1+\rho)^2}\,\mathrm{tr}(H_\theta O)\right]. \tag{10}$$

Different variance and covariance weights mean that adding endpoints need not reduce absolute bias monotonically.

From objective bias to parameter displacement. Under the preceding assumptions, let $\theta^\star$ be an interior local minimum of $\mathcal{R}_\infty$ with nonsingular Hessian $I_\star = \nabla^2_\theta \mathcal{R}_\infty(\theta^\star)$. If $\theta^\circ_{b,\rho b}$ is a branch of interior local minima of $\mathbb{E}\widehat{\mathcal{L}}_{b,\rho b}$ converging to $\theta^\star$, then

$$\theta^\circ_{b,\rho b} - \theta^\star = -\frac{1}{b} I_\star^{-1} \nabla_\theta A_\rho(\theta^\star) + o(b^{-1}). \tag{11}$$

This predicts expected-minimizer displacement, not finite-data risk (proofs: Appendices B.1 and B.2).

3.5 Conditions for changes in bias direction

Under the regularity conditions of Theorem 1, additionally assume exact calibration, correct specification at $\theta_\star$, and strictly positive joint probabilities. With population marginals fixed, let $C_\theta = (\Pi_\theta - pq^\top)_{1:m-1,1:n-1}$. Appendix B.11 derives the calibration Hessian used below.

Proposition 1 (Spectral direction of the leading bias)**.** *Under the preceding assumptions, let $t_k < 1$ be the squared singular values of $\Sigma_p^{-1/2} C_\theta \Sigma_q^{-1/2}$. Then $\nabla_\theta A_\rho(\theta_\star) = \nabla_\theta \sum_k b_\rho(t_k(\theta))|_{\theta=\theta_\star}$, where*

$$b_\rho(t) = \frac{(1+2\rho)\log(1-t) + 2\rho t/(1-t)}{2(1+\rho)^2}, \tag{12}$$

$$b'_\rho(t) = \frac{(1+2\rho)t - 1}{2(1+\rho)^2(1-t)^2}. \tag{13}$$

For $\rho > 0$, each spectral contribution changes sign at $t = 1/(1+2\rho)$; for $\rho = 0$, the derivative remains negative.

The threshold makes both dependence structure and sample composition explicit. For uniform binary marginals, $x, y \in \{-1, +1\}$, and $f_\theta(x, y) = \theta xy$, put $\kappa = \tanh\theta_\star > 0$. The local shift is

$$\theta^\circ_{b,\rho b} - \theta_\star = \frac{\kappa[1 - (1+2\rho)\kappa^2]}{b(1+\rho)^2(1-\kappa^2)^2} + o(b^{-1}). \tag{14}$$

When κ^2 lies below the spectral threshold, the leading shift strengthens dependence; above it, the shift weakens dependence. Additional endpoint data can therefore reverse the direction of the relation-parameter displacement.

For general models, Equation (11) maps these contributions to parameter changes. The spectral gradient identity holds at the correctly specified truth, not globally. Proofs appear in Appendix B.11.

3.6 Ratio-preserving first-order objective correction

For even b, u, split complete pairs and each independent endpoint pool into equal halves without using observed values. Each half retains u/b and is calibrated separately at the same parameter; when $u = 0$, only pairs are split. Following the classical split-jackknife principle (Dhaene & Jochmans, 2015), define

$$\widehat{\mathcal{L}}^{\mathrm{JK}}_{b,u} = 2\widehat{\mathcal{L}}_{b,u} - \tfrac{1}{2}\left\{\widehat{\mathcal{L}}^{(1)}_{b/2,u/2} + \widehat{\mathcal{L}}^{(2)}_{b/2,u/2}\right\}. \tag{15}$$

Corollary 1 (Ratio-preserving split correction)**.** *Under the assumptions of Theorem 1, uniformly on the same parameter neighborhood,*

$$\mathbb{E}\widehat{\mathcal{L}}^{\mathrm{JK}}_{b,\rho b}(\theta) = \mathcal{R}_\infty(\theta) + o(b^{-1}). \tag{16}$$

Full and half batches have leading biases A_ρ/b and $2A_\rho/b$, so their weighted difference cancels this term; see Appendix B.2. Implementation details appear in Appendix C.

4 Experimental setup

Controlled experiments test how calibration changes the learning objective and its parameter minimizers, then evaluate first-order correction on finite-data estimation. Application studies assess PACT's value for conditional prediction and cross-policy risk assessment. Tasks and metrics appear in Table 6 in Appendix C.

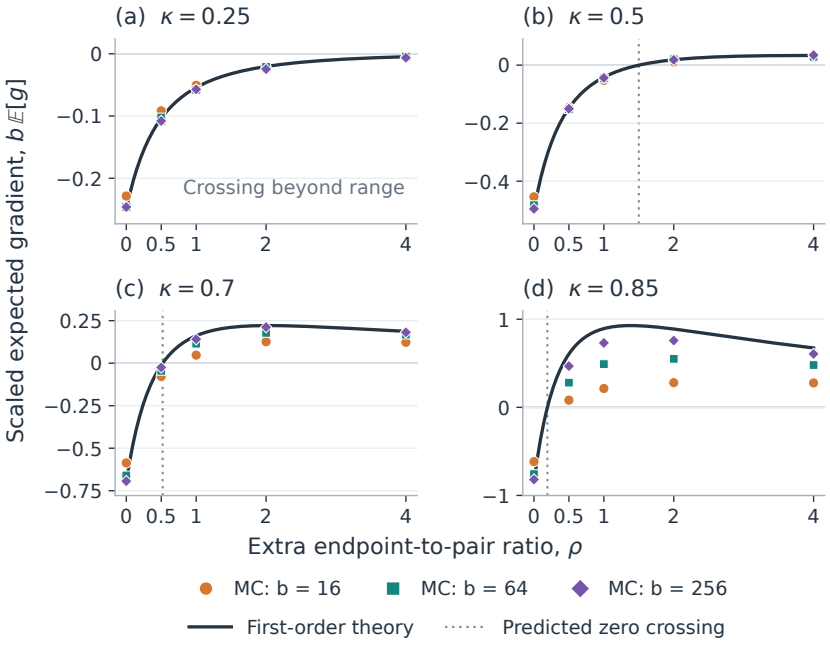

Figure 3: **Endpoint samples can reverse leading gradient bias.** Curves: first-order coefficients; points: scaled variance-reduced expected gradients with pointwise 95% Monte Carlo intervals. Dotted lines mark predicted zero crossings; the crossing in (a) lies outside the displayed range.

Controlled studies. Finite-support experiments vary dependence strength, sampling ratio, and batch size to test gradient bias, expected-objective minimizer displacement, and finite-data KL/TV error. Comparisons share score bases, optimizers, and parameter domains; classical correction comparisons select regularization separately on development data. Independent Swiss-roll experiments assess correction at fixed and separately tuned evaluation temperatures.

Applications and ablations. Swiss roll follows established protocols for baseline comparability (Tenenbaum et al., 2000; Persiianov et al., 2026); Colored MNIST tests multimodal prediction with unknown hue offsets on MNIST (LeCun et al., 1998). Finance and queueing use aligned trajectories and known dynamics, respectively, to evaluate cross-policy risk. These applications use task-adapted objectives to evaluate the PACT framework; correction is evaluated separately in the controlled and neural comparisons. Matched-architecture ablations compare paired contrastive learning, paired calibrated likelihood, and mixed-sample calibrated likelihood with common paired data, training and selection budgets, and final marginal constraints. Each procedure separately selects its learning rate, evaluation temperature, and applicable checkpoint under the same development rules.

Finance crosses five independent training datasets with five independent market datasets, estimating uncertainty along both axes; its objective ablation uses five separate independent confirmation blocks. Data sources, baselines, selection rules, and inference procedures appear in Appendices C, E, and F.

5 Results and Analysis

PACT combines limited correspondence information with separately observed distributions to construct joint models for conditional prediction and cross-policy risk assessment. The results first explain how empirical calibration changes relation learning, then examine bias correction and objective selection, and finally demonstrate PACT's application value.

Table 1: Paired estimation ($b = 20$, bounds $[-4, 4]^2$): JK minus raw error, averaged over 500 datasets per truth, with paired 95% t intervals (499 degrees of freedom). Negative favors correction.

Truth	Δ KL (95% interval)	Δ TV (95% interval)
Moderate	-0.0245 $[-0.0403, -0.0087]$	-0.0075 $[-0.0106, -0.0044]$
Strong	-0.0883 $[-0.1009, -0.0757]$	-0.0198 $[-0.0230, -0.0167]$

Table 2: Regularized finite-support KL (256 confirmation datasets). U/N: uniform/nonuniform marginals; MP: classical modified profile likelihood. Each method selects regularization on development data; intervals appear in Table 20.

m	Margins	b	Raw	JK	MP
3	U	64	0.022394	0.023516	0.023295
3	U	256	0.007425	0.007435	0.007399
3	N	64	0.015272	0.015275	0.014936
3	N	256	0.005238	0.005264	0.005251
6	U	64	0.025288	0.026406	0.026352
6	U	256	0.007738	0.007702	0.007688
6	N	64	0.024869	0.025722	0.025722
6	N	256	0.007067	0.007028	0.007038

5.1 Effects of Sample Composition on Calibration Bias and Parameter Learning

PACT reveals how additional marginal observations change relation learning. Even with the true joint distribution fixed, adding unpaired samples for marginal estimation can reverse first-order gradient bias. Estimated expected gradients in the binary model exhibit the predicted direction changes; see Figure 3. PACT explains this effect through the covariance between marginal estimation and paired supervision: additional samples change both the marginal estimates and their statistical association with paired supervision, thereby affecting the objective optimized by the relation model.

PACT further connects this objective change to learned relation parameters. Under paired-only sampling, estimated expected-objective minimizer displacements approach the first-order prediction as batches grow, and split correction brings the mean minimizer closer to truth in every tested condition; see Table 23 in Appendix D.10. Interior solutions and positive-definite numerical Hessians verify local minima underlying these shifts; see Table 24.

Mixed sampling further shows why correction effects depend on sampling conditions: under moderate dependence, correction increases parameter displacement at smaller batches and reduces it at larger batches. Complete displacements, approximation accuracy, and numerical checks appear in Figure 9 and Tables 23 and 24 in Appendix D.10.

5.2 Implications of Calibration Analysis for Objective Correction and Training Strategies

PACT translates its bias analysis into a direct correction of the training objective. In controlled paired settings, split correction reduces both mean KL divergence and total variation error, with paired intervals excluding zero; see Table 1. Correction also reduces both errors across both dependence strengths with a wider parameter range; see Table 10 in Appendix D.

Stronger comparisons establish the scope of these gains. With development-selected regularization, correction significantly increases KL in some configurations; the remaining comparisons are inconclusive, and none establishes superiority over classical modified profile likelihood; see Tables 2 and 20. In independent Swiss-roll experiments, correction lowers coupling KL at a fixed evaluation temperature; separately tuning both methods' evaluation temperatures yields no clear KL or conditional-MMD difference; see Tables 25 and 26 in Appendix D.11.

PACT separates relation learning from final marginal matching so that each task can use a suitable learning objective. Matched-architecture ablations compare independently tuned procedures under common networks, data, marginal constraints, and development budgets. Paired calibrated likelihood

Table 3: Objective ablations with matched networks, pairs, budgets, selection rules, and final calibration. Confirmation mean ± sample SD; bold: lowest mean.

Relation objective	Swiss MMD	CMNIST NLL	Finance dependence
Paired contrastive	.2306 ± .1672	**.0823 ± .0355**	0.0877 ± 0.0023
Paired likelihood	**.1309 ± .1036**	.5868 ± .4604	**0.0804 ± 0.0033**
Mixed likelihood	.2663 ± .1953	.5068 ± .1987	0.0895 ± 0.0035
Mixed (128)	.3269 ± .2609	–	–
Mixed (512)	.3081 ± .2514	–	–

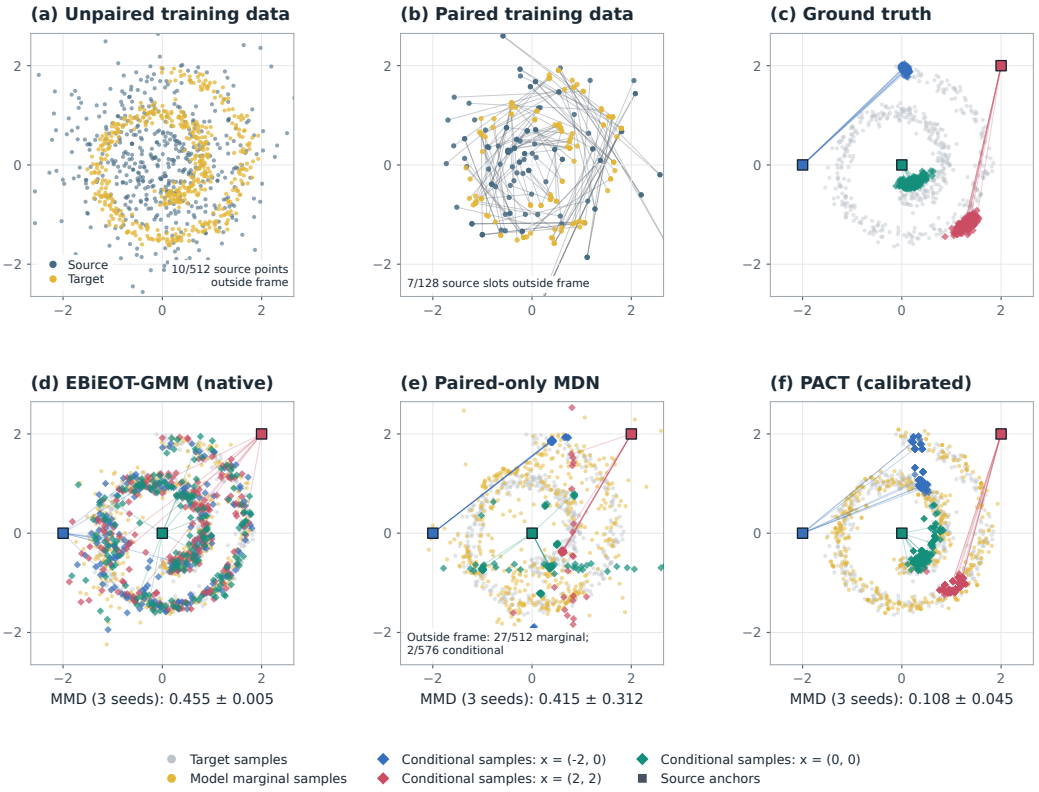

Figure 4: **Swiss-roll conditional prediction.** In (c–f), gray/gold dots show target/model marginals; colored diamonds match square source anchors. Panels show replicate 1; MMD reports mean ± sample SD over three replicates. Axes and sample counts are shared; out-of-frame counts are labeled. All replicates and full-range views: Figures 6 and 7.

attains the lowest mean error on Swiss roll and finance, whereas contrastive learning outperforms both likelihood variants on Colored MNIST; see Tables 3 and 29. In finance, paired calibrated likelihood recovers loss dependence more accurately than contrastive learning, with statistical significance retained after adjustment for multiple comparisons; see Table 31. PACT thus improves prediction and joint estimation through task-adapted relation learning under common final marginal constraints.

On Swiss roll, paired likelihood outperforms every tested mixed-sample setting, while unpaired samples determine the final marginals; see Table 29. PACT thus uses unpaired data to estimate distributions without requiring them in relation learning. Complete objective comparisons and statistical intervals appear in Appendix E.

Table 4: Colored-MNIST conditional recovery (mean ± sample SD over three seeds). Lower NLL and higher Recall@1 are better. "Calibrated" denotes marginal calibration to numerical tolerance.

Method	Pair NLL ↓	R@1 ↑
PACT calibrated	$.0978 \pm .0407$	$.9818 \pm .0137$
Stabilized EBiEOT calibrated	$6.2338 \pm .0059$	$.0020 \pm .0020$
Circular conditional MLE (specified)	$\mathbf{.0276 \pm .0036}$	$\mathbf{1.0000 \pm .0000}$

Table 5: Finance under common empirical marginals: five training datasets crossed with five independent market datasets. Lower is better except Top-1; inverse EOT and projected diffusion are the primary comparators.

Method	Dependence	Joint	CVaR	Regret	Top-1
PACT	0.1016	0.2704	1.0192	0.1403	0.6608
Inverse EOT	0.1095	0.2723	1.0694	0.1373	0.6600
Pair diffusion, projected	0.1192	0.2779	1.0549	0.4061	0.3692
Independent product	0.9033	0.3231	6.2851	0.2956	0.5067

5.3 Conditional Prediction and Joint Risk Assessment with Limited Correspondence

PACT models the overall target distribution and uses limited pairs to learn which outcomes correspond to a given input. On Swiss roll, PACT achieves lower conditional MMD than the baselines in every confirmation replicate; see Figure 4 and Table 14. Even when a baseline fits the target marginal slightly better, PACT retains more accurate conditional predictions; see Table 15. This comparison demonstrates the value of learning correspondence: beyond reproducing the target distribution, PACT predicts outcomes for specific inputs more accurately.

On Colored MNIST, PACT extends this capability to multiple possible outputs for each input. Without knowing the true hue offsets, it learns multimodal relationships and improves likelihood and retrieval over the tested baseline; see Table 4. The circular reference is more accurate with the correct functional form specified; PACT learns these relationships from pairs without known offsets.

PACT connects data describing separate policy outcomes into a joint model of how losses vary across policies. In finance, crossed evaluation over independent training datasets and market scenarios shows more accurate recovery of loss dependence than inverse transport and projected diffusion, with both differences remaining statistically significant after multiplicity adjustment; see Table 32. Under common empirical marginals, PACT also achieves lower mean joint-distribution error, providing a more accurate joint loss model for cross-policy risk assessment; see Table 5.

In queueing, PACT uses known dynamics without paired labels and achieves lower mean CVaR error than the compared baselines under common empirical marginals in all three mechanisms. Native replay without marginal matching is more accurate; see Table 16 in Appendix D.5.

6 Conclusion

PACT combines limited correspondence information with separately observed marginals to construct joint models for conditional prediction and cross-policy risk assessment. Its theoretical analysis reveals how the composition of paired and independent marginal samples changes the learning objective, predicts local parameter displacement and changes in the direction of first-order gradient bias, and yields a ratio-preserving split correction that eliminates first-order bias in the expected learning objective. Controlled experiments support these theoretical predictions and characterize the correction's estimation gains and their scope; applications demonstrate the value of recovering correspondence beyond marginal fit. Together, these results clarify how data composition and empirical marginal constraints shape joint-distribution learning.

Reproducibility Statement

Appendix A provides pseudocode for calibrated-likelihood training, optional ratio-preserving split correction, and calibrated conditional prediction. Appendix C specifies data generation, architectures, training and selection budgets, baselines, and evaluation metrics. Proofs appear in Appendix B; additional results and matched-architecture comparisons are reported in Appendices D and E. Appendix F details numerical calibration, Queue's endpoint map, and Finance uncertainty estimation.

AI Use Statement

In this work, we used generative AI tools to assist with language polishing, literature review, coding, and the review of mathematical proofs and scientific claims. The authors reviewed all AI-assisted work and take full responsibility for the final manuscript.

References

Ole Barndorff-Nielsen. On a formula for the distribution of the maximum likelihood estimator. *Biometrika*, 70(2):343–365, 1983.

Christopher M. Bishop. Mixture density networks. Neural Computing Research Group Report NCRG/94/004, Aston University, 1994. URL https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/bishop-ncrg-94-004.pdf.

Charlotte Bunne, Andreas Krause, and Marco Cuturi. Supervised training of conditional monge maps. *Advances in Neural Information Processing Systems*, 35:6859–6872, 2022.

Wei-Ting Chiu, Pei Wang, and Patrick Shafto. Discrete probabilistic inverse optimal transport. In *International Conference on Machine Learning*, pp. 3925–3946. PMLR, 2022.

David Roxbee Cox and Nancy Reid. Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society: Series B (Methodological)*, 49(1):1–18, 1987.

Imre Csiszár. I-divergence geometry of probability distributions and minimization problems. *The annals of probability*, pp. 146–158, 1975.

Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. *Advances in neural information processing systems*, 26, 2013.

Geert Dhaene and Koen Jochmans. Split-panel jackknife estimation of fixed-effect models. *The Review of Economic Studies*, 82(3):991–1030, 2015.

Kilian Fatras, Younes Zine, Rémi Flamary, Remi Gribonval, and Nicolas Courty. Learning with minibatch Wasserstein: asymptotic and gradient properties. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 2131–2141. PMLR, 2020. URL https://proceedings.mlr.press/v108/fatras20a.html.

David Firth. Bias reduction of maximum likelihood estimates. *Biometrika*, pp. 27–38, 1993.

Aude Genevay, Lénaic Chizat, Francis Bach, Marco Cuturi, and Gabriel Peyré. Sample complexity of sinkhorn divergences. In *The 22nd international conference on artificial intelligence and statistics*, pp. 1574–1583. PMLR, 2019.

Promit Ghosal, Marcel Nutz, and Espen Bernton. Stability of entropic optimal transport and schrödinger bridges. *Journal of Functional Analysis*, 283(9):109622, 2022.

Arthur Gretton, Karsten M Borgwardt, Malte J Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *The journal of machine learning research*, 13:723–773, 2012.

Michael Gutmann and Aapo Hyvärinen. Noise-contrastive estimation: A new estimation principle for unnormalized statistical models. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pp. 297–304. JMLR Workshop and Conference Proceedings, 2010.

Jonathan Ho, Ajay N. Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html.

C Terrance Ireland and Solomon Kullback. Contingency tables with given marginals. *Biometrika*, 55 (1):179–188, 1968.

Marcel Klatt, Carla Tameling, and Axel Munk. Empirical regularized optimal transport: Statistical theory and applications. *SIAM Journal on Mathematics of Data Science*, 2(2):419–443, 2020.

Dominik Klein, Théo Uscidda, Fabian Theis, and Marco Cuturi. Genot: Entropic (gromov) wasserstein flow matching with applications to single-cell genomics. *Advances in Neural Information Processing Systems*, 37:103897–103944, 2024.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=PqvMRDCJT9t.

Piyushi Manupriya, Rachit K Das, Sayantan Biswas, and SakethaNath N Jagarlapudi. Consistent optimal transport with empirical conditional measures. In *International Conference on Artificial Intelligence and Statistics*, pp. 3646–3654. PMLR, 2024.

Peter McCullagh and Robert Tibshirani. A simple method for the adjustment of profile likelihoods. *Journal of the Royal Statistical Society: Series B (Methodological)*, 52(2):325–344, 1990.

Nikita Moriakov, Jonas Adler, and Jonas Teuwen. Kernel of CycleGAN as a principal homogeneous space. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=B1eWOJHKvB.

Gerhard Osius. Asymptotic inference for semiparametric association models. *The Annals of Statistics*, pp. 459–489, 2009.

Art B. Owen. The pigeonhole bootstrap. *The Annals of Applied Statistics*, 1(2):386–411, 2007. doi: 10.1214/07-AOAS122. URL https://doi.org/10.1214/07-AOAS122.

Art B. Owen and Dean Eckles. Bootstrapping data arrays of arbitrary order. *The Annals of Applied Statistics*, 6(3):895–927, 2012. doi: 10.1214/12-AOAS547. URL https://doi.org/10.1214/12-AOAS547.

George Papamakarios, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. *Journal of Machine Learning Research*, 22(57):1–64, 2021.

Edouard Pauwels and Samuel Vaiter. The derivatives of sinkhorn–knopp converge. *SIAM Journal on Optimization*, 33(3):1494–1517, 2023.

Mikhail Persiianov, Arip Asadulaev, Nikita Andreev, Nikita Starodubcev, Dmitry Baranchuk, Anastasis Kratsios, Evgeny Burnaev, and Alexander Korotin. Inverse entropic optimal transport solves semi-supervised learning via data likelihood maximization. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=0p617sK4Z4.

Gabriel Peyré and Marco Cuturi. Computational optimal transport: With applications to data science. *Foundations and Trends in Machine Learning*, 11(5-6):355–607, 2019. doi: 10.1561/2200000073. URL https://doi.org/10.1561/2200000073.

Maurice H Quenouille. Notes on bias in estimation. *Biometrika*, 43(3/4):353–360, 1956.

R Tyrrell Rockafellar, Stanislav Uryasev, et al. Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42, 2000.

Asma Saleh. Reduced bias estimation of the log odds ratio. *Statistical Papers*, 65(8):5293–5331, 2024. doi: 10.1007/s00362-024-01593-7. URL https://doi.org/10.1007/s00362-024-01593-7.

John K Salmon, Mark A Moraes, Ron O Dror, and David E Shaw. Parallel random numbers: as easy as 1, 2, 3. In *Proceedings of 2011 international conference for high performance computing, networking, storage and analysis*, pp. 1–12, 2011.

Lee W. Schruben. Common random numbers. In *Wiley Encyclopedia of Operations Research and Management Science*. John Wiley & Sons, Ltd, 2011. doi: 10.1002/9780470400531.eorms0166.

Tomonari Sei and Keisuke Yano. Estimation of bivariate minimum information dependence models using the density-ratio framework. *Annals of the Institute of Statistical Mathematics*, pp. 1–25, 2026.

Liangliang Shi, Gu Zhang, Haoyu Zhen, Jintao Fan, and Junchi Yan. Understanding and generalizing contrastive learning from the inverse optimal transport perspective. In *International conference on machine learning*, pp. 31408–31421. PMLR, 2023.

Richard Sinkhorn and Paul Knopp. Concerning nonnegative matrices and doubly stochastic matrices. *Pacific Journal of Mathematics*, 21(2):343–348, 1967.

Joshua B Tenenbaum, Vin de Silva, and John C Langford. A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323, 2000.

Panagiotis Theodoropoulos, Nikolaos Komianos, Vincent Pacelli, Guan-Horng Liu, and Evangelos Theodorou. Feedback schrödinger bridge matching. In *International Conference on Learning Representations*, volume 2025, pp. 58939–58967, 2025.

Soumya Tripathy, Juho Kannala, and Esa Rahtu. Learning image-to-image translation using paired and unpaired training samples. In *Asian conference on computer vision*, pp. 51–66. Springer, 2018.

Christina Winkler, Daniel Worrall, Emiel Hoogeboom, and Max Welling. Learning likelihoods with conditional normalizing flows. *arXiv preprint arXiv:1912.00042*, 2019.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *2017 IEEE international conference on computer vision (ICCV)*, pp. 2242–2251. Ieee, 2017.

SUPPLEMENTARY CONTENTS

A Core algorithms **15**

B Proofs for the PACT theory **17**

B.1 Objective reformulation and local parameter displacement 17

B.2 Proofs of the bias expansion, parameter displacement, and split correction 17

B.3 Latent-gauge non-identification 18

B.4 Population identification and preservation 18

B.5 Proof of Theorem 2 19

B.6 Proof of Theorem 3 19

B.7 Marginal estimation and downstream risk 19

B.8 Proof of Theorem 4: estimated marginals 20

B.9 Continuation of Theorem 4: downstream bounds 20

B.10 Scenario-preserving empirical marginal replacement 21

B.11 Proof of the spectral characterization 21

C Experimental protocols **22**

C.1 Data splits and model selection 22

C.2 Finite-minibatch study 23

C.3 Swiss roll 24

C.4 Colored MNIST 24

C.5 Cross-domain architecture-matched strategy comparison 25

C.6 Queue 26

C.7 Finance 27

C.8 Evaluation metrics and aggregation 28

C.9 Binary verification of the bias direction 29

C.10 Regularized comparison with modified profile likelihood 29

C.11 Expected-objective minimizers 30

C.12 Independent evaluation-temperature confirmation 30

D Additional results and diagnostics **31**

D.1 Finite-minibatch expansion and estimation risk 31

D.2 Learned-score transfer of the split correction 32

D.3 Conditional application results and visualizations 32

D.4 Native and calibrated conditional outputs 33

D.5 Queue endpoint matching and full baseline comparison 34

D.6 Finance baseline comparisons 35

D.7 Boundary targets 35

D.8 Spectral verification 36

D.9 Regularized classical-correction comparison 36
D.10 Expected-minimizer displacements and numerical checks 38
D.11 Fixed and tuned evaluation temperatures 39

E Matched-architecture comparisons **39**

F Calibration and statistical details **41**
F.1 Positive finite-temperature calibration 41
F.2 Queue's replicated-endpoint boundary map 44
F.3 Finance uncertainty under crossed seeds 44
F.4 Compatibility implementations and structural relations 45
F.5 Conditional prediction from calibrated scores 45

G Additional connections to prior work **46**

H Discussion and limitations **46**
H.1 Turning marginal information into relation learning 46
H.2 Limitation & future work 46

SUPPLEMENTARY MATERIALS

A CORE ALGORITHMS

Algorithm 1 specifies minibatch calibrated-likelihood training with optional split correction. Algorithm 2 constructs finite-temperature predictions from a fitted score, including scores learned by the contrastive objectives in Section C.5. Queue uses the structural replay and endpoint map in Section F.2.

Algorithm 1 Calibrated-likelihood training with optional split correction

Input: Paired data and independent endpoint pools; groups G; per-group batch sizes b, u; training temperature $\tau_{\mathrm{tr}} > 0$; correction flag; training budget T, optimizer, prescribed checkpoint schedule and development selection rule; optional loss penalty $R(\theta)$ (zero if unused).

Output: Parameters $\widehat{\theta}$ at the development-selected checkpoint.

function BATCHLOSS(θ, P, U_X, U_Y)
  Form pair frequencies η_P and pooled marginals $\widehat{p}, \widehat{q}$ by Equation (4).
  Keep positive-mass endpoint states; form $S_{\theta,ij} = s_\theta(x_i, y_j)$ on their Cartesian product.
  Compute $\log \Pi$ for $\Pi = \mathcal{C}_{\widehat{p},\widehat{q}}(S_\theta)$ at τ_{tr} using the prescribed training solver.
  return $-\sum_{ij:\eta_{P,ij}>0} \eta_{P,ij} \log \Pi_{ij}$.
end function
Initialize θ and the prescribed optimizer.
for $t = 1, \ldots, T$ **do**
  for $g = 1, \ldots, G$ **do**
    Draw b complete pairs P_g and u independent observations per endpoint, $U_{X,g}, U_{Y,g}$.
    $\ell_g \leftarrow$ BATCHLOSS($\theta, P_g, U_{X,g}, U_{Y,g}$).
    if split correction is enabled **then**
      Split pair indices independently of observed values to form equal halves $P_g^{(1)}, P_g^{(2)}$.
      Split endpoint indices independently, without using values, into equal halves $U_{X,g}^{(h)}, U_{Y,g}^{(h)}$, $h = 1, 2$.
      $\ell_g \leftarrow 2\ell_g - \frac{1}{2}\sum_{h=1}^{2}$ BATCHLOSS($\theta, P_g^{(h)}, U_{X,g}^{(h)}, U_{Y,g}^{(h)}$).
    end if
  end for
  $J \leftarrow G^{-1}\sum_{g=1}^{G} \ell_g + R(\theta)$.
  Differentiate J through every full- and half-batch calibration.
  Update θ once with the prescribed optimizer; retain scheduled checkpoints.
end for
return the checkpoint fixed by the development protocol.

For correction, b, u are even; $u = 0$ gives empty endpoint pools, whose terms are omitted. Each half rebuilds its empirical pair law and marginals at the same θ and τ_{tr}, preserving u/b. Splits retain repeated observations and never exchange samples across contexts. Groups are calibrated separately before averaging. The algorithm uses joint NLL; conditional normalization on uniform endpoint banks changes only parameter-independent terms. An explicit loss penalty is added once; optimizer-side weight decay follows the declared protocol. Calibration differentiation is implicit or through iterations as specified in Section C. Finite-support fits optimize the same losses on fixed datasets with the stated numerical solvers.

Here $\mathrm{LSE}(z) = \log \sum_j e^{z_j}$. Algorithm 2 displays log-domain alternating scaling; the solver variants, residual definitions, and acceptance criteria follow Section C. Fixed-iteration settings use all K iterations. Prediction uses the target potential fitted at the same evaluation temperature. On a positive-mass source-bank point, the softmax normalizes the computed row; it equals $\widehat{\Pi}_{ij}/\widehat{p}_i$ when the marginal constraints hold.

Algorithm 2 Marginal calibration and conditional prediction

Input: Fitted score $s_{\widehat{\theta}}$; weighted endpoint banks $\{(x_i, \widehat{p}_i)\}$, $\{(y_j, \widehat{q}_j)\}$; query x; evaluation temperature $\tau_{\mathrm{ev}} > 0$; positive integer iteration cap K and declared stopping rule.

Output: Joint mass matrix $\widehat{\Pi}$, conditional probabilities $\widehat{P}_j(x)$, and marginal residual δ.

Restrict to positive-mass states; retain their normalized weights and original indices.
$S_{ij} \leftarrow s_{\widehat{\theta}}(x_i, y_j)$; initialize potentials $\alpha_i \leftarrow 0$, $\beta_j \leftarrow 0$.
for $k = 1, \ldots, K$ **do**
    $\alpha_i \leftarrow -\tau_{\mathrm{ev}} \, \mathrm{LSE}_j[\log \widehat{q}_j + (S_{ij} + \beta_j)/\tau_{\mathrm{ev}}]$ for all i.
    $\beta_j \leftarrow -\tau_{\mathrm{ev}} \, \mathrm{LSE}_i[\log \widehat{p}_i + (S_{ij} + \alpha_i)/\tau_{\mathrm{ev}}]$ for all j.
    $\log \widehat{\Pi}_{ij} \leftarrow \log \widehat{p}_i + \log \widehat{q}_j + (S_{ij} + \alpha_i + \beta_j)/\tau_{\mathrm{ev}}$.
    $\widehat{\Pi} \leftarrow \exp(\log \widehat{\Pi})$ elementwise; compute the protocol's marginal residual δ.
    if residual-based stopping is prescribed and its tolerance is met **then**
        break
    end if
end for
$z_j \leftarrow \log \widehat{q}_j + (s_{\widehat{\theta}}(x, y_j) + \beta_j)/\tau_{\mathrm{ev}}$.
$\widehat{P}_j(x) \leftarrow \exp[z_j - \mathrm{LSE}_k(z_k)]$.
Restore zero mass at excluded rows, columns, and target entries.
return $\widehat{\Pi}$, $\widehat{P}(x)$, δ.

B PROOFS FOR THE PACT THEORY

B.1 OBJECTIVE REFORMULATION AND LOCAL PARAMETER DISPLACEMENT

With $f_\theta = S_\theta/\tau$ and free marginal coordinates $r = (p_{1:m-1}, q_{1:n-1})$, let $z, \widehat{r}$ denote the paired and pooled marginals. On the interior, the empirical objective equals, up to parameter-independent terms,

$$\widehat{\mathcal{L}}_{b,u}(\theta) = -\langle \eta, f_\theta \rangle + \Phi_\theta(\widehat{r}) + (z - \widehat{r})^\top \nabla_r \Phi_\theta(\widehat{r}). \tag{17}$$

The final term vanishes when supervised and calibration marginals coincide. The local displacement in Equation (11) concerns a convergent branch of interior minima of the expected objective under the main-text assumptions. A parameter-constant leading bias causes no first-order displacement. The derivations follow below.

B.2 PROOFS OF THE BIAS EXPANSION, PARAMETER DISPLACEMENT, AND SPLIT CORRECTION

We use free coordinates obtained by dropping the last category of each marginal. All derivatives below are taken in these coordinates. Put $r = (p_{1:m-1}, q_{1:n-1})$, $M = b+u$, $\delta = \widehat{r} - r$, and $\xi = z - \widehat{r}$.

Empirical objective. Aggregate the sample-index coupling by endpoint category, and write its log-density ratio as

$$\log \frac{\Pi_{ij}}{p_i q_j} = f_{\theta,ij} + a_i + b_j,$$

where the potentials here absorb temperature. The optimum value is $\Phi_\theta(p, q) = -p^\top a - q^\top b$. By the envelope theorem, in the free coordinates,

$$\partial_{p_i} \Phi_\theta = -(a_i - a_m), \qquad \partial_{q_j} \Phi_\theta = -(b_j - b_n).$$

The reference-product derivatives are constant on each simplex and cancel in these differences. At the empirical marginals, the exact pair NLL is

$$-\langle \eta, f_\theta \rangle - (z^X)^\top a - (z^Y)^\top b - (z^X)^\top \log \widehat{p} - (z^Y)^\top \log \widehat{q}.$$

The last two terms are independent of θ. Subtracting and adding $-\widehat{p}^\top a - \widehat{q}^\top b = \Phi_\theta(\widehat{r})$ gives Equation (17). This makes explicit that $\widehat{\mathcal{L}}_{b,u}$ and the NLL have the same parameter derivatives, although their values differ by sample-dependent constants. When $u = 0$, $z = \widehat{r}$ and the potential correction vanishes.

Covariances. Let Σ_p and Σ_q be the multinomial covariance matrices in free coordinates and $\Sigma_{pq} = \pi^\star - pq^\top$ the paired cross-covariance, with the last row and column removed. Directly expanding the paired and independent sample averages yields

$$\mathrm{Cov}(\widehat{r}) = \frac{D}{M} + \frac{bO}{M^2}, \qquad \mathrm{Cov}(z, \widehat{r}) = \frac{D + O}{M}. \tag{18}$$

The diagonal blocks of the first identity use all M observations at each endpoint; only the b aligned observations contribute to its off-diagonal blocks. For the second identity, z uses the paired sample alone.

Expansion. The score average is unbiased: $\mathbb{E}\langle \eta, f_\theta \rangle = \langle \pi^\star, f_\theta \rangle$. On the event that $\widehat{r}$ remains in the positive neighborhood from the theorem assumptions, Taylor expansion at r gives

$$\begin{aligned} \Phi_\theta(r + \delta) &= \Phi_\theta(r) + \nabla \Phi_\theta(r)^\top \delta + \tfrac{1}{2} \delta^\top H_\theta \delta + R_1, \\ \nabla \Phi_\theta(r + \delta) &= \nabla \Phi_\theta(r) + H_\theta \delta + R_2, \end{aligned}$$

where the uniform derivative bound gives $|R_1| \le C\|\delta\|^3$ and $\|R_2\| \le C\|\delta\|^2$. Since $\mathbb{E}\delta = \mathbb{E}\xi = 0$,

$$\mathbb{E}\widehat{\mathcal{L}}_{b,u}(\theta) = \mathcal{R}_\infty(\theta) + \mathrm{tr}\{H_\theta \,\mathrm{Cov}(z, \widehat{r})\} - \tfrac{1}{2} \,\mathrm{tr}\{H_\theta \,\mathrm{Cov}(\widehat{r})\} + o(b^{-1}).$$

Indeed, multinomial third moments make $\mathbb{E}|R_1| + \mathbb{E}|\xi^\top R_2| = O(b^{-3/2})$ when u/b is fixed. Positivity of the true marginals makes the complement of the Taylor neighborhood exponentially unlikely. On

the original sample-index support, bounded scores bound the theta-dependent likelihood difference uniformly, so this complement is $o(b^{-1})$ as well.

Substituting Equation (18) gives

$$\frac{\operatorname{tr}(H_\theta D)}{2M} + \frac{b+2u}{2M^2}\operatorname{tr}(H_\theta O) + o(b^{-1}).$$

With $u = \rho b$, this is Equations (9) and (10). Uniformity follows from the uniform derivative bounds. Applying Taylor expansion also to $\nabla_\theta \Phi_\theta$ gives the differentiated expansion: continuity of the joint third derivatives makes its quadratic remainder uniformly $o(\|\delta\|^2)$, and the corresponding potential remainder contributes $o(b^{-1})$ in expectation. Thus the objective expansion can be used in the local score equation. Expanding the first-order condition $\nabla_\theta \mathbb{E}\widehat{\mathcal{L}}_{b,\rho b}(\theta^\circ_{b,\rho b}) = 0$ around $\theta^\star$ and using nonsingularity of $I_\star$ proves Equation (11).

For the split correction, a half sample has the same ratio ρ and hence expectation $\mathcal{R}_\infty + 2A_\rho/b + o(b^{-1})$. Therefore

$$2\{\mathcal{R}_\infty + A_\rho/b + o(b^{-1})\} - \{\mathcal{R}_\infty + 2A_\rho/b + o(b^{-1})\} = \mathcal{R}_\infty + o(b^{-1}),$$

which proves Corollary 1. This argument does not assert an $O(b^{-2})$ remainder without stronger fourth-order conditions.

B.3 Latent-gauge non-identification

Let $Z \sim \nu$ and let T_a satisfy $(T_a)_\# \nu = P_a$. If $(G_a)_\# \nu = \nu$, then $(T_a \circ G_a)_\# \nu = P_a$. Choosing different G_a across conditions may change $(T_a \circ G_a, T_b \circ G_b)_\# \nu$. For example, take $\nu = \mathcal{N}(0,1)$, $T_a = T_b$ equal to the identity, $G_a(z) = z$, and $G_b(z) = -z$. The endpoint marginals are unchanged while the correlation changes from 1 to -1.

B.4 Population identification and preservation

Theorem 2 (Identification by calibrated likelihood)**.** *Assume $\pi^\star \in \Gamma(p,q)$ and $p_i, q_j, \pi^\star_{ij} > 0$, and let $S^\star_{ij} = \tau \log[\pi^\star_{ij}/(p_i q_j)]$. Then $\mathcal{C}_{p,q}(S^\star) = \pi^\star$. Moreover, for $\mathcal{R}(S) = -\sum_{ij} \pi^\star_{ij} \log[\mathcal{C}_{p,q}(S)]_{ij}$,*

$$\mathcal{R}(S) - \mathcal{R}(S^\star) = \mathrm{KL}(\pi^\star \| \mathcal{C}_{p,q}(S))\,. \tag{19}$$

Thus the pair law is recovered whenever the score class contains the target log density ratio up to endpoint gauges.

The following preservation statement uses positive p, q, finite scores, and $\tau > 0$, as in Equation (1). The score supplies interaction; calibration selects the coupling with the requested marginals. Endpoint samples alone cannot identify the interaction.

Theorem 3 (Gauge invariance, interaction preservation, and stability)**.** *For arbitrary row and column vectors r, c,*

$$\mathcal{C}_{p,q}(S + r\mathbf{1}^\top + \mathbf{1}c^\top) = \mathcal{C}_{p,q}(S). \tag{20}$$

If $\Pi = \mathcal{C}_{p,q}(S)$, every 2×2 interaction obeys

$$\log \frac{\Pi_{ij}\Pi_{k\ell}}{\Pi_{i\ell}\Pi_{kj}} = \frac{S_{ij} + S_{k\ell} - S_{i\ell} - S_{kj}}{\tau}. \tag{21}$$

Finally, with $\delta(S,T) = \inf_{r,c} \|S - T - r\mathbf{1}^\top - \mathbf{1}c^\top\|_\infty$,

$$\mathrm{TV}(\mathcal{C}_{p,q}(S), \mathcal{C}_{p,q}(T)) \le \min\{1, \delta(S,T)/(2\tau)\}. \tag{22}$$

The first two statements follow from Equation (34). For stability, row and column potentials cancel when the log-plan ratio is integrated against the difference of plans with common marginals; symmetrized KL and Pinsker's inequality yield Equation (22). The detailed argument is in Section B.6.

B.5 Proof of Theorem 2

Because KL is strictly convex on the feasible affine set and $\tau > 0$, the maximization objective is strictly concave; hence the optimizer is unique. The Lagrangian first-order condition for Eq. (1) is

$$\log \frac{\Pi_{ij}}{p_i q_j} = \frac{S_{ij} + a_i + b_j}{\tau},$$

after absorbing an additive constant into either potential. This proves Eq. (34). Substituting

$$S^\star_{ij} = \tau \log \frac{\pi^\star_{ij}}{p_i q_j}$$

and taking zero potentials gives $\Pi = \pi^\star$, which is feasible and therefore the unique optimizer.

For any score S, let $\Pi_S = \mathcal{C}_{p,q}(S)$. Then

$$\mathcal{R}(S) - \mathcal{R}(S^\star) = \sum_{ij} \pi^\star_{ij} \log \frac{\pi^\star_{ij}}{(\Pi_S)_{ij}} = \mathrm{KL}(\pi^\star \| \Pi_S).$$

The difference is nonnegative and vanishes exactly when the two plans agree.

For the negative-sampling objective associated with Equation (35), the row-wise population objective is

$$-\sum_{ij} \pi^\star_{ij} f_{ij} + \sum_i p_i \log \sum_j q_j e^{f_{ij}}.$$

For each row it is cross-entropy between $\pi^\star(j \mid i)$ and $q_j e^{f_{ij}} / \sum_\ell q_\ell e^{f_{i\ell}}$. Its minimizers therefore satisfy $f_{ij} = \log[\pi^\star_{ij}/(p_i q_j)] + c_i$. The column-wise statement is identical after swapping the endpoints.

B.6 Proof of Theorem 3

For every $\Pi \in \Gamma(p, q)$,

$$\langle \Pi, r\mathbf{1}^\top + \mathbf{1}c^\top \rangle = \langle p, r \rangle + \langle q, c \rangle,$$

which is constant over the feasible set. Adding row or column functions therefore leaves the unique optimizer unchanged.

Using Eq. (34), all endpoint masses and scaling potentials cancel in a 2×2 cross ratio, leaving

$$\frac{\Pi_{ij}\Pi_{k\ell}}{\Pi_{i\ell}\Pi_{kj}} = \exp\left\{ \frac{S_{ij} + S_{k\ell} - S_{i\ell} - S_{kj}}{\tau} \right\}.$$

For stability, write $\Pi = \mathcal{C}_{p,q}(S)$ and $P = \mathcal{C}_{p,q}(T)$. Their log ratio has the form

$$\log \frac{\Pi_{ij}}{P_{ij}} = \frac{S_{ij} - T_{ij}}{\tau} + u_i + v_j.$$

Because Π and P have the same marginals, the row and column terms cancel when integrated against $\Pi - P$. For any r, c, with $E = S - T - r\mathbf{1}^\top - \mathbf{1}c^\top$,

$$\begin{aligned} J(\Pi, P) &:= \mathrm{KL}(\Pi \| P) + \mathrm{KL}(P \| \Pi) \\ &= \frac{1}{\tau} \langle \Pi - P, E \rangle \le \frac{2\|E\|_\infty}{\tau} \mathrm{TV}(\Pi, P). \end{aligned}$$

Pinsker's inequality in both directions gives $J(\Pi, P) \ge 4\mathrm{TV}(\Pi, P)^2$. If the total variation is nonzero, division and minimization over r, c give $\mathrm{TV}(\Pi, P) \le \delta(S, T)/(2\tau)$; the zero case is immediate.

B.7 Marginal estimation and downstream risk

Theorem 4 (Estimated marginals and downstream risk). *Fix $\beta > 0$ and $B < \infty$. On the finite-support class in which all endpoint masses are at least β and the gauge-adjusted score norm is at*

most B, where $\|S\|_{\mathrm{g}} = \inf_{r,c} \|S - r\mathbf{1}^\top - \mathbf{1}c^\top\|_\infty$, there is a finite constant $L = L(\beta, B, \tau, m, n)$ such that, for estimates sufficiently close and remaining in this class,

$$\mathrm{TV}(\mathcal{C}_{\widehat{p},\widehat{q}}(\widehat{S}), \pi^\star) \le L(\|\widehat{p} - p\|_1 + \|\widehat{q} - q\|_1) + \delta(\widehat{S}, S^\star)/(2\tau). \tag{23}$$

Let $\widehat{\pi} = \mathcal{C}_{\widehat{p},\widehat{q}}(\widehat{S})$. If a downstream scalar h is L_h-Lipschitz on a pair metric space of diameter D, then for $0 \le \alpha < 1$,

$$|\,\mathrm{CVaR}_\alpha(h_\#\widehat{\pi}) - \mathrm{CVaR}_\alpha(h_\#\pi^\star)| \le \frac{L_h}{1-\alpha} W_1(\widehat{\pi}, \pi^\star) \le \frac{L_h D}{1-\alpha}\,\mathrm{TV}(\widehat{\pi}, \pi^\star). \tag{24}$$

Uniform action-risk error at most ϵ yields decision regret at most 2ϵ.

Equation (23) separates marginal and compatibility error; Equation (24) connects coupling recovery to tail risk. The constants depend on the finite support and minimum mass. Boundary smoothing is described in Section D.

B.8 Proof of Theorem 4: estimated marginals

For fixed (p, q, S), use the smooth convex dual

$$D(a, b) = \tau \sum_{ij} p_i q_j \exp\left(\frac{S_{ij} + a_i + b_j}{\tau}\right) - \langle p, a\rangle - \langle q, b\rangle.$$

At its optimum, the exponential terms form the positive plan Π and the Hessian is

$$H = \frac{1}{\tau}\begin{pmatrix} \mathrm{diag}(p) & \Pi \\ \Pi^\top & \mathrm{diag}(q) \end{pmatrix}.$$

For perturbations (u, v) its quadratic form is

$$(u, v)^\top H(u, v) = \frac{1}{\tau}\sum_{ij} \Pi_{ij}(u_i + v_j)^2.$$

Because every $\Pi_{ij} > 0$, the null space consists exactly of $u = c\mathbf{1}_m$, $v = -c\mathbf{1}_n$, the row–column gauge direction. Fixing one gauge, for example $\langle p, a\rangle = 0$, makes the restricted Hessian positive definite. The implicit-function theorem therefore makes the potentials, and hence the calibrated plan, continuously differentiable in (p, q, S) throughout the positive interior.

The specified lower mass bound and two-sided gauge-adjusted score bound define a compact subset of that interior. The derivative norm is consequently bounded there by a finite constant L. Holding the score fixed and applying the mean value theorem bounds the change caused by $(p, q) \mapsto (\widehat{p}, \widehat{q})$. Holding the true marginals fixed and applying Theorem 3 bounds the score change. The triangle inequality gives Eq. (23).

For a boundary target, convexity gives

$$\mathrm{TV}((1-\eta)\pi^\star + \eta pq^\top, \pi^\star) = \eta\,\mathrm{TV}(pq^\top, \pi^\star) \le \eta.$$

B.9 Continuation of Theorem 4: downstream bounds

The variational representation of upper-tail CVaR is

$$\mathrm{CVaR}_\alpha(Z) = \inf_t \left\{ t + \frac{1}{1-\alpha}\mathbb{E}[(Z - t)_+] \right\}.$$

For fixed t, the hinge is one-Lipschitz. Kantorovich–Rubinstein duality and the inequality $|\inf f - \inf g| \le \sup|f - g|$ imply

$$|\mathrm{CVaR}_\alpha(P) - \mathrm{CVaR}_\alpha(Q)| \le \frac{1}{1-\alpha} W_1(P, Q).$$

An L_h-Lipschitz pushforward contracts Wasserstein distance up to L_h, and on a space of diameter D, $W_1 \le D\,\mathrm{TV}$. This proves Eq. (24).

For action selection, let $a^\star = \arg\min_a R_a$ and $\widehat{a} = \arg\min_a \widehat{R}_a$. If $\sup_a |R_a - \widehat{R}_a| \le \epsilon$, then

$$R_{\widehat{a}} \le \widehat{R}_{\widehat{a}} + \epsilon \le \widehat{R}_{a^\star} + \epsilon \le R_{a^\star} + 2\epsilon.$$

B.10 SCENARIO-PRESERVING EMPIRICAL MARGINAL REPLACEMENT

The queue implementation uses a finite-sample counterpart of structural posterior replay. Draw coherent tuples $(Y_1^{(s)}, \dots, Y_K^{(s)})$, $s \in [N]$, using one posterior disturbance per tuple. For each condition a, expand its M-point frozen endpoint pool into an $N = MR$ element multiset by repeating each pool position R times. Assign that multiset to tuple indices in the deterministic native ordered-slot order. Every output marginal is then exactly the replicated empirical pool, while all coordinates retain the same tuple slot. The assignment can retain its latent tie-breaking order as an auxiliary label, but repeated numeric values create ties: ranks and the empirical copula recomputed from the numeric output are generally coarsened and need not equal those of native replay. Exact posterior-replay values, value-based distances, and CVaR need not be preserved; this is precisely the Queue-specific structural-replay-to-fixed performance gap reported in the experiments. This rank-map statement is separate from the row-softmax native diagnostic used for positive finite-temperature scores.

B.11 PROOF OF THE SPECTRAL CHARACTERIZATION

Assume the fixed finite-support setting of Theorem 1, positive joint probabilities, exact calibration, fixed positive population marginals, and correct specification at $\theta_\star$. In reduced endpoint-indicator coordinates, write

$$K_\theta = \begin{pmatrix} \Sigma_p & C_\theta \\ C_\theta^\top & \Sigma_q \end{pmatrix}, \qquad C_\theta = (\Pi_\theta - pq^\top)_{1:m-1,1:n-1}, \qquad D = \operatorname{diag}(\Sigma_p, \Sigma_q).$$

These covariance matrices are positive definite in the reduced coordinates. The Hessian of the calibration value with respect to the marginals is

$$H_\theta = D^{-1} - K_\theta^{-1}. \tag{25}$$

Define

$$B_\rho(\theta) = \frac{(1+2\rho)\log\det K_\theta + \rho\,\operatorname{tr}(K_\theta^{-1}D)}{2(1+\rho)^2}.$$

Then the first-order score bias at a correctly specified parameter satisfies

$$\nabla_\theta A_\rho(\theta_\star) = \nabla_\theta B_\rho(\theta_\star).$$

This equality is local at the truth; the two functions need not agree away from the truth, even up to a constant.

Let t_k be the squared singular values of $R_\theta = \Sigma_p^{-1/2} C_\theta \Sigma_q^{-1/2}$. Each $t_k < 1$. Up to a parameter-independent constant,

$$B_\rho(\theta) = \sum_k b_\rho(t_k), \qquad b_\rho(t) = \frac{(1+2\rho)\log(1-t) + 2\rho t/(1-t)}{2(1+\rho)^2},$$

and

$$b_\rho'(t) = \frac{(1+2\rho)t - 1}{2(1+\rho)^2(1-t)^2}.$$

For $\rho > 0$, a dependence mode's contribution changes sign at $t = 1/(1+2\rho)$. The derivative stays negative for $\rho = 0$. This is a statement about spectral contributions to the expected score bias. General parameter displacements additionally depend on the score parameterization and the inverse population Hessian.

Proof. The maximum joint-entropy-plus-score value is the negative convex conjugate of the endpoint-potential log partition. Its Hessian in marginal mean coordinates is $-K_\theta^{-1}$. Subtracting the marginal entropies contributes D^{-1}, giving the Hessian identity. Hold the true data distribution, hence $O_\star$, fixed while differentiating. For $M_\star = D/(1+\rho) + (1+2\rho)O_\star/(1+\rho)^2$,

$$dA_\rho = \tfrac{1}{2}\operatorname{tr}[K_\theta^{-1}(dK_\theta)K_\theta^{-1}M_\star].$$

Only after differentiation, use correct specification to substitute $K_\star = D + O_\star$. Thus $M_\star = ((1+2\rho)K_\star - \rho D)/(1+\rho)^2$. The standard differential identities for the log determinant and inverse

Table 6: Evidence roles of the experimental studies.

Question	Comparison	Evaluation target
Bias direction	Binary spectral grid	Expected gradients
Parameter displacement	Raw / corrected expected loss	Local minimizers
Estimation accuracy	Raw / split / classical correction	KL, TV, bias and variance
Additional neural gains	Fixed / tuned temperatures	Coupling KL, conditional MMD
Application value	Four application tasks	Prediction and cross-policy risk
Objective choice	Matched relation learners	Common-marginal task error

yield $dA_\rho = dB_\rho$ at $\theta_\star$. The whitened covariance $D^{-1/2}K_\theta D^{-1/2}$ has blocks $I, R_\theta, R_\theta^\top, I$, so its nontrivial eigenvalues are $1 \pm \sqrt{t_k}$. Consequently

$$\log\det K_\theta = \log\det D + \sum_k \log(1-t_k), \qquad \operatorname{tr}(K_\theta^{-1}D) = m+n-2+\sum_k \frac{2t_k}{1-t_k}.$$

Substitution and scalar differentiation complete the proof. Repeated singular values are handled through the smooth matrix function above; the argument does not require choosing differentiable singular vectors.

An explicit binary consequence. For $x, y \in \{-1, +1\}$, uniform marginals and $f_\theta(x, y) = \theta xy$, put $\kappa = \tanh\theta_\star$. Then

$$\partial_\theta A_\rho(\theta_\star) = \frac{\kappa[(1+2\rho)\kappa^2 - 1]}{(1+\rho)^2(1-\kappa^2)}.$$

Since the population Hessian is $1-\kappa^2$, the local expected-objective minimizer has leading displacement

$$\theta^\circ_{b,\rho b} - \theta_\star = \frac{\kappa[1-(1+2\rho)\kappa^2]}{b(1+\rho)^2(1-\kappa^2)^2} + o(b^{-1}).$$

For a positive truth this strengthens dependence below the threshold and weakens it above the threshold. It describes a local minimum of the expected objective, not the mean of a finite-dataset optimizer or its prediction risk.

Relation to likelihood adjustment. For paired observations and a linear canonical score, calibration at empirical marginals has an exact profile-likelihood interpretation. Write

$$P_{\theta,\gamma}(i,j) = \exp\{\theta^\top F_{ij} + \gamma^\top z_{ij} - \psi(\theta,\gamma)\},$$

where z_{ij} contains reduced row and column indicators. At fixed θ, the nuisance score equation matches $E[z]$ to the observed marginal frequencies. Profiling γ therefore gives the calibrated paired likelihood. The nuisance observed information is bK_θ. The canonical modified-profile adjustment (Saleh, 2024), expressed as average negative log likelihood, is consequently

$$L_{\mathrm{MP}}(\theta) = L_{\mathrm{raw}}(\theta) - \frac{1}{2b}\log\det K_\theta,$$

up to parameter-independent terms. Here K_θ uses empirical marginals; it is the nuisance covariance, not the full Fisher information. At $\rho = 0$, the local potential $B_0 = \frac{1}{2}\log\det K_\theta$ recovers the corresponding classical adjustment at the level of the first-order score bias at the truth. This does not identify the split estimator with modified profile likelihood at finite sample size. For mixed sampling, the additional inverse-trace term and shared-pair covariance require the analysis above: imposing pooled marginals on paired supervision is not automatically profiling the full observed-data likelihood.

C Experimental protocols

C.1 Data splits and model selection

Architecture, optimizer, temperature, and hyperparameters are selected on development data and frozen before confirmation. Endpoint pools and aligned truth are disjoint where the simulator permits. Queue comparator selection and its unadjusted interval are specified below.

C.2 FINITE-MINIBATCH STUDY

Split-corrected training differentiates each calibration and adds any regularizer once. Prediction uses Equation (2).

The finite-support model uses two doubly centered orthonormal interactions on 3×3 support, uniform marginals, unit temperature, and truths $(.25, 1)$ and $(.25, 2)$. The coefficient study crosses paired batch sizes $\{8, 16, 20, 32, 64\}$ with additional-endpoint ratios $\rho \in \{0, 1\}$. Three Monte Carlo blocks (seeds 92001–92003) use 20,000 independent archives per condition; the mixed diagnostic uses 50,000. Expected half-objective gradients use independent half-size archives. See Tables 8 and 9 for gradients and Table 10 for fixed-data optimization diagnostics.

Risk compares raw calibrated likelihood and split-jackknife on 500 datasets per truth, with paired batch size 20, seeds 93001 and 93002, and one deterministic complementary 10/10 split. Both parameter boxes $[-4, 4]^2$ and $[-6, 6]^2$ use all datasets. These define constrained estimators: boundary fits are outside the interior local displacement formula.

Score basis and fitting. For states $0, 1, 2$, identity I, $F_{ij} = 1\{j = (i + 1) \bmod 3\}$, and $B = F^\top$, use

$$U_1 = \frac{I - \mathbf{1}\mathbf{1}^\top/3}{\sqrt{2/9}}, \qquad U_2 = \frac{F - B}{\sqrt{2/3}}, \qquad S_\theta = \theta_1 U_1 + \theta_2 U_2.$$

The basis is doubly centered and orthonormal under $\langle A, B\rangle/9$. At unit temperature, $\pi_{ij}(\theta) = e^{S_{\theta,ij}}/(3\sum_k e^{S_{\theta,ik}})$. For $[-4, 4]^2$, the six starts are $(0, 0), (0, 1), (0, 2), (.5, 1), (.5, 2), (-.5, 0)$; $[-6, 6]^2$ additionally uses $(0, 4), (.5, 4)$. Both losses use analytic-gradient L-BFGS-B: 600 iterations, function tolerance 10^{-13}, gradient tolerance 10^{-8}, and 40 line-search steps. Active-support reduced-dual Newton calibration uses absolute marginal tolerance 10^{-12} and an iteration cap of 300. The lowest finite exact objective selects the fit, requiring this tolerance and projected-gradient residual $\|\theta - \mathrm{clip}(\theta - \nabla L, -a, a)\|_\infty \le 10^{-6}$ for box radius a. Coordinates within .001 of a box limit count as boundary hits.

Three corrected fits required refinement: a restart with function tolerance zero, gradient tolerance 10^{-10}, 2,000 iterations and 80 line-search steps; all four corners; and scalar minimization on 16 equal subintervals per edge, with tolerance 10^{-12} and 1,000 iterations. Selection uses the same objective, never population risk; failed candidates are retained. All 4,000 final fits meet both numerical criteria, establishing constrained first-order accuracy, not global optimality.

Risk uses the analytic coupling at the fitted score and true uniform marginals. For corrected-minus-raw risk differences d_i, pointwise paired t intervals are $\bar{d} \pm t_{.975,499} s_d/\sqrt{500}$ for each truth and box; they are not simultaneous. Parameter bias, variance and MSE use all fitted vectors, with variance normalized by 500. The swap comparator requires no calibration, is not refitted, and is excluded from the 4,000 numerical checks.

Learned-score transfer check. Raw paired likelihood and its ratio-preserving split correction share a three-hidden-layer width-128 MLP, initial weights and paired stream within replicate, two groups of 64 pairs per update, AdamW learning rate and weight decay 10^{-4}, gradient clip 5, training/evaluation temperatures .05 and .08, and 6,000 updates. Correction combines one full and two complementary half objectives; a separate deterministic RNG partitions existing observations.

Three development blocks select each arm's global checkpoint from $\{500, 1000, 2000, 4000, 6000\}$ by mean fixed-output conditional MMD, requiring projection residual below 10^{-5}. Both select update 6,000 before five fresh confirmation blocks. The primary success criterion requires negative mean corrected-minus-raw MMD, wins in at least four of five blocks, and paired-bootstrap 95% upper bound below zero. Interaction error is secondary. All ten fits are retained, sharing endpoint pools and truth within blocks; model and output streams are separate from data streams.

For the interaction diagnostic, $\mathrm{H}(Z)$ subtracts row and column means and adds the grand mean on the first 128 training endpoints per side. Relative RMSE is $\|\mathrm{H}(S_\theta/.08) - \mathrm{H}(-C_{\mathrm{rot}}/.05)\|_F / \|\mathrm{H}(-C_{\mathrm{rot}}/.05)\|_F$, where C_{rot} is the benchmark rotation cost. Double centering removes endpoint gauges. This oracle training-bank diagnostic measures score interaction, not neural weights or held-out prediction.

C.3 SWISS ROLL

The native EBiEOT protocol supplies 128 aligned samples and 1,024 unpaired samples per endpoint. Confirmation uses fresh data seeds 969100–969102, method-specific optimization seeds, and truth seeds 979100–979102. Conditional truth uses the official repeated 1,024-by-1,024 minibatch OT construction, with 1,024 endpoints per official starting point. Comparators are the official EBiEOT-GMM configuration (Persiianov et al., 2026) and a comparable-capacity paired-only MDN (Bishop, 1994). Conditional MMD is primary; conditional Sinkhorn proxy and unconditional MMD are secondary.

Application implementation. PACT maps the 12-dimensional features $[x, y, x \odot y, x - y, |x - y|, x + y]$ through three width-128 SiLU layers to a scalar score. Symmetric paired cross-entropy uses 128 pairs per update, AdamW learning rate 3×10^{-4}, weight decay 10^{-4}, gradient clip 10, 15,000 updates, and temperature .05. Dual-only calibration uses 300 log-Sinkhorn iterations on 1,024-point endpoint banks and records marginal residuals. The MDN has 16 diagonal-Gaussian components, two width-128 SiLU layers, and log standard deviations clipped to $[-4.5, 1.5]$; paired conditional NLL training uses 25,001 updates, batch size 128, Adam learning rate 3×10^{-4}, and weight decay 10^{-6}. The official GMM uses 50 model and 25 cost potentials with empty hidden-channel lists, alternating unpaired/paired Adam updates at .001/.0003, batch size 128, for 25,001 steps. Figure 4 and Table 14 compare calibrated PACT with native baseline samplers. The separate baseline-projection diagnostic (Section D.4) does not give all methods a common exact marginal-matching output.

Learning-objective ablation. Five arms use paired contrastive learning, paired calibrated likelihood, or mixed-support calibrated likelihood with nested endpoint pools of 128, 512, or 1,024. They share a three-hidden-layer width-128 MLP, initialization and paired stream within replicate, AdamW settings, gradient clip 5, 6,000 updates, training temperature .05, 128 paired draws per update, and final full-pool projection. Mixed arms add 128 independent draws per endpoint per update: pool size changes diversity, not batch size. Two development blocks select each arm's learning rate from $\{10^{-4}, 3\times10^{-4}, 10^{-3}\}$ and evaluation temperature from $\{.03, .05, .08\}$.

For Table 3, the selected learning rate and temperature are retained. Three further development blocks with disjoint training data, validation endpoints, and conditional truth select each arm's global checkpoint from $\{500, 1000, 2000, 4000, 6000\}$ by mean calibrated validation MMD, subject to projection-residual tolerance. Every trajectory runs 6,000 updates. Five fresh confirmation blocks then use disjoint model/output seeds; all mixed pools are reported. Their data differ from the application comparison (Table 14), precluding a cross-protocol checkpoint-effect comparison.

C.4 COLORED MNIST

Unlike the single-offset colored-digit task of Persiianov et al. (2026), our MNIST task (LeCun et al., 1998) has bimodal circular hue dependence.

Source hues are uniform on $[0, 360^\circ)$, with target hue

$$h_y = (h_x + 120^\circ z) \bmod 360^\circ, \qquad z \in \{-1, +1\}.$$

Offsets are balanced and randomly permuted within paired sets; unpaired hues are independent uniform draws. Grayscale values are divided by 255, bilinearly resized to 32×32 with `align_corners=False`, and rescaled to $[-1, 1]$. PACT receives no generating offsets.

Training uses MNIST-train digits 2 and 3: 500 freshly colored pairs and complete independently colored endpoint pools. Confirmation uses the first 512 digit-2/digit-3 MNIST-test identities, fresh hues and branches, and three fresh seeds. PACT trains a circular Fourier head for 1,500 steps; prediction temperature is $\tau = .08$.

The head extracts $\operatorname{atan2}(\sqrt{3}(g - b), 2r - g - b)$ from image-summed RGB. Target-minus-source angle d gives features $[\cos(kd), \sin(kd)]_{k=1}^{12}$, followed by two width-64 SiLU layers and a scalar output. Symmetric paired cross-entropy uses 128 pairs per update, training temperature 1, AdamW learning rate .001, weight decay 10^{-4}, and gradient clip 5. Double-precision final calibration requires absolute residual 2×10^{-7} within 30,000 iterations.

The development-selected EBiEOT CNN uses a rank-generic sampler broadcast repair, gradient clip 5, 500 steps, batch size 64, and learning rate 3×10^{-5}. Energy is normalized by score standard deviation on the unlabeled confirmation bank; development selects $\tau = .03$. The HSV reference assumes the correct symmetric two-offset form, estimates offset magnitude by median absolute circular hue difference in rendered training pairs, scores both signs, and calibrates to the same bank. PACT does not receive this form or the offsets. Methods report native row-softmax and converged fixed-marginal outputs; only the latter are invariant to endpoint score gauges.

C.5 Cross-domain architecture-matched strategy comparison

Three arms share score architecture, initial weights, paired data and sampling stream, 1,500 updates, AdamW weight decay 10^{-4}, gradient clip 5, and final projection. They use symmetric paired InfoNCE, paired calibrated likelihood, or mixed-support calibrated likelihood. Unpaired draws use a separate RNG. Each arm selects learning rate and evaluation temperature with equal development budgets.

Empirical objectives. Group $g \in [G]$ contains b aligned pairs $\{(x_i^{(g)}, y_i^{(g)})\}_{i=1}^{b}$, and write $S_{ij}^{(g,M)} = s_\theta(x_i^{(g)}, y_j^{(g)})$ on M endpoints per side. Paired contrastive training ($M = b$) uses symmetric InfoNCE:

$$\widehat{L}_{\mathrm{pc}}(\theta) = -\frac{1}{2Gb} \sum_{g=1}^{G} \sum_{i=1}^{b} \left[\log \frac{e^{S_{ii}^{(g,b)}/\tau_{\mathrm{tr}}}}{\sum_{j=1}^{b} e^{S_{ij}^{(g,b)}/\tau_{\mathrm{tr}}}} + \log \frac{e^{S_{ii}^{(g,b)}/\tau_{\mathrm{tr}}}}{\sum_{j=1}^{b} e^{S_{ji}^{(g,b)}/\tau_{\mathrm{tr}}}} \right]. \tag{26}$$

Let $\mathbf{u}_M$ denote uniform weights on M points and $\Pi_\theta^{(g,M)} = \mathcal{C}_{\mathbf{u}_M, \mathbf{u}_M}(S^{(g,M)})$. Paired likelihood uses the same $b \times b$ support:

$$\widehat{L}_{\mathrm{pb}}(\theta) = -\frac{1}{Gb} \sum_{g=1}^{G} \sum_{i=1}^{b} \log\Big(b[\Pi_\theta^{(g,b)}]_{ii} \Big). \tag{27}$$

Mixed likelihood appends u independent endpoints per side: $M = b + u$, with positives restricted to the first b diagonal entries:

$$\widehat{L}_{\mathrm{mb}}(\theta) = -\frac{1}{Gb} \sum_{g=1}^{G} \sum_{i=1}^{b} \log\Big(M[\Pi_\theta^{(g,M)}]_{ii} \Big). \tag{28}$$

Factors b and M convert joint masses to conditional probabilities. Unpaired endpoints affect normalization and gradients in Equation (28), without positive labels. Both likelihoods differentiate through the converged dual solution.

Two development replicates select learning rates $\{3 \times 10^{-4}, 10^{-3}\}$ and calibration temperatures $\{.05, .08, .2, .5, 1\}$ for Colored MNIST or $\{.25, .5, 1\}$ for finance, minimizing validation pair NLL or scaled dependence energy, respectively. Choices are frozen before five confirmation blocks. Finance uses separate development, training, validation, and confirmation random streams, sharing only within comparison blocks.

Colored MNIST arms share the 12-harmonic width-64 cyclic score, 500 training pairs, disjoint MNIST-train validation identities 1000–1255, and the first 512 MNIST-test confirmation identities with fresh colors. Updates draw 64 pairs at training temperature 1; mixed likelihood adds 64 endpoints per side. In arm order (contrastive, paired, mixed), selected learning rates are $.001, .001, .0003$ and evaluation temperatures $.05, .5, .05$.

Finance arms share the width-128 `SummaryPairCritic`. Training has 96 contexts, 200 rollouts per action, and 20 aligned rollouts per edge. Updates use 16 groups of all 20 pairs at temperature .5; mixed training adds 20 independent endpoints per side from the remaining 180 rollouts. All arms select learning rate .0003; evaluation temperatures are $.25, .5, .25$. Validation/confirmation use 16/48 independent contexts. Each confirmation context has disjoint 32-sample endpoint pools and 128 further aligned truth samples. Dependence energy averages one-dimensional paired-difference errors divided by training feature standard deviations; CVaR uses a training cost scale. These pools and normalization differ from the finance application, precluding direct numerical comparison.

Calibrated arms use double-precision implicit dual differentiation, checked by finite differences, zero row/column gradient sums, and the full-pair envelope identity. Retained training residuals are below 10^{-4} relative; final residuals are below 2×10^{-7} absolute.

Primary comparisons use mean confirmation error and paired replicate differences: 20,000 bootstrap draws for Colored MNIST; 50,000 draws and 97.5% marginal intervals for finance's two prespecified comparisons. Finance CVaR and image retrieval are secondary; all arms are reported.

C.6 QUEUE

Dynamics, policies, and loss. For $c = (\lambda_0, A, \phi, Q_0, \mu_0, B)$ and normalized time t/H, $H = 24$, arrivals are

$$A_t \sim \mathrm{Pois}(\lambda_0[1 + A\sin(2\pi(t + .5)/H + \phi)]) + I_t\,\mathrm{Pois}(\lambda_{\mathrm{burst}}), \qquad I_t \sim \mathrm{Bernoulli}(p_{\mathrm{burst}}). \tag{29}$$

Capacity noise is $Z_t = \rho Z_{t-1} + \sqrt{1-\rho^2}\epsilon_t$ with deterministic multipliers u_t^a for policy a and updates

$$\begin{aligned} K_t^a &= \mu_0 u_t^a \exp(.14Z_t), & S_t^a &= \min\{Q_t^a + A_t, K_t^a\}, \\ O_t^a &= (Q_t^a + A_t - S_t^a - B)_+, & Q_{t+1}^a &= Q_t^a + A_t - S_t^a - O_t^a. \end{aligned} \tag{30}$$

The three multipliers are steady $(1, 1)$, front-loaded $(1.3, .7)$, and back-loaded $(.7, 1.3)$ over horizon halves. Summaries are mean/maximum queue, total overflow, and mean queue in the upper time quintile. Cost is mean queue plus twice overflow; upper-tail error is defined below.

Three policies are tested under three mechanisms. Nominal capacity shocks are Gaussian AR(1), $\rho = .72$. Stress increases arrival/service ratio and bursts, reduces buffers, and uses the correctly supplied $\rho = .85$. Misspecification uses Student-$t(4)$ innovations, $\rho = .35$, and larger bursts; PACT retains Gaussian $\rho = .72$ inference.

Nominal and misspecified contexts draw $\lambda_0 \in [3.5, 6.5]$, $A \in [.05, .35]$, $Q_0 \in [0, 8]$, $\mu_0 \in [4.5, 6.5]$, and $B \in [12, 24]$, with uniform phase; burst parameters are $(p_{\mathrm{burst}}, \lambda_{\mathrm{burst}}) = (.06, 4)$ and $(.09, 5)$, respectively. Stress uses the corresponding ranges $[6, 8.5]$, $[.15, .45]$, $[2, 10]$, $[4.2, 6]$, and $[10, 20]$, with $(.12, 6)$ bursts. Innovations have unit variance: Gaussian except for standardized Student-$t(4)$ under misspecification.

Each mechanism has 128 training contexts with 64 rollouts per policy and 32 validation contexts with 16. Confirmation uses two fresh seeds, 24 contexts per seed, three disjoint 128-sample marginal pools, and separate 256-sample aligned truth per context. PACT receives no alignment labels; neural baselines receive eight aligned pairs per observed star edge and all unpaired training rollouts. Comparators include raw-path inverse EOT, matching critics, edge-specific/shared encoders, three tree completions, official-architecture EBiEOT ($N = 50, M = 5$), paired/joint CFM, pair diffusion, independence, comonotonicity, tuned cost rank, and Gaussian/Student-t factor MLE. EBiEOT is adapted to this task.

Selected non-neural baselines. Cost rank selects temperature from $\{.1, .5, 1, 2\}$ by validation CRPS of policy-cost differences. Gaussian/Student-t factors fit correlations to Gaussian rank scores of training costs; validation likelihood selects shrinkage from $\{0, .25, .5, .75, 1\}$ and Student-t degrees of freedom from $\{4, 8, 16\}$. Sampled costs are mapped to the same endpoint pools.

Neural baselines and evaluation. Training observes star edges $(0, 1), (0, 2)$, holding out $(1, 2)$. Paths have 24 steps and three channels (queue, service, overflow), standardized on training data. Pair critics use three width-64 SiLU temporal convolutions (kernels 5, 3, 3), mean/max pooling, a width-32 context layer, and two width-64 scalar-head layers with LayerNorm after the first activation. Shared encoders use the same temporal widths, a width-32 condition layer, and a width-64 projection with LayerNorm to 32 dimensions; edge-specific encoders fit each observed edge separately. Matching/inverse-EOT variants share temperature .5, AdamW learning rate 3×10^{-4}, weight decay 10^{-5}, 16-group batches, 80-epoch cap, and validation patience 10. Inverse EOT uses 40 training Sinkhorn iterations. Edge plans form a maximum-entropy tree law; independent, shared-index, and path-metric tree constructions are also evaluated.

Paired/joint flow matching and conditional diffusion use three width-128 SiLU layers, AdamW learning rate 3×10^{-4}, weight decay .01, batches of 512, 80-epoch cap, and validation patience 10.

Conditional generators train in both directions. Flow integration/diffusion use 32 steps; diffusion scores use four Monte Carlo samples and joint-flow scores four samples per endpoint. EBiEOT uses 50 potentials, five cost terms, potential widths $(128, 128)$, cost widths $(256, 256)$, softplus diagonals, paired/unpaired learning rates 3×10^{-4}/.001, and the same epoch, patience, and batch budgets. Generator edge scores use kernel temperature .5 and common balanced endpoints. Graph evaluation allows 10,000 Sinkhorn iterations, checks tolerance 10^{-6} every 20 iterations, and draws 256 coupling samples; factor/empirical-copula samplers use 8,192 slots. Optimizer seeds are 2026091901 and 2026091902; true-pair/shuffled controls share initialization and sampling. Full means retain shuffled controls outside baseline competition (Tables 16 and 17).

Posterior disturbance reconstruction. Queue, service, and overflow recover arrivals as $A_t = Q_{t+1} + S_t + O_t - Q_t$. When $Q_t + A_t - S_t > 10^{-4}$, observed service fixes capacity noise $Z_t = \log[\max(S_t, 10^{-8})/(\mu_0 u_t^a)]/.14$; otherwise this is a lower bound. Gaussian AR(1) Gibbs updates draw unobserved coordinates from normal conditionals truncated below these bounds: interior mean $\rho(Z_{t-1} + Z_{t+1})/(1+\rho^2)$ and variance $(1-\rho^2)/(1+\rho^2)$; endpoint mean ρ times its neighbor and variance $1-\rho^2$. Observed coordinates stay fixed. Initialization clips the Gaussian conditional mean to the bounds. Eight draws using 24 sweeps replay shared arrivals and disturbances under every policy, providing reconstructed correspondence without observed pair labels. Cost tuples are mapped to endpoint pools via Section F.2.

Each confirmation context replays 128 trajectories per source policy, yielding $3 \times 128 \times 8 = 3{,}072$ tuples. Each of 128 endpoint values is replicated 24 times; Appendix F.2 specifies assignment and ties.

The primary fixed-margin CVaR$_{.9}$ error for $\max(C_1 - C_0, C_2 - C_0)$ uses a training-only cost scale. After validation tuning, the pooled nominal/stress comparator is the baseline family with lowest pooled confirmation mean: tuned cost rank. Its 20,000 paired bootstrap draws (seed 20260919) resample 24 contexts per mechanism/evaluation-seed stratum, holding the comparator fixed without adjusting for its selection. No optimizer-replicate resampling is needed for this non-neural comparator. KPG-RL optimizes policies rather than estimating the joint law of frozen policies and is not compared.

C.7 Finance

Market, hedge rules, and loss. Option contexts draw spot, strike, variance, spread, liquidity, and jump intensity/scale. All hedge rules share the market path, with $\Delta t = 1/H$:

$$\log(S_{t+1}/S_t) = -\tfrac{1}{2} V_t \Delta t + \sqrt{V_t \Delta t}\, Z_t^S + J_t,$$
$$V_{t+1} = \left[V_t + 3(.04 - V_t)\Delta t + .55\sqrt{V_t \Delta t}\, Z_t^V \right]_+, \qquad \mathrm{Corr}(Z_t^S, Z_t^V) = -.62, \quad (31)$$

Jumps occur with probability $\lambda_J \Delta t$ and have distribution $\mathcal{N}(-.35\sigma_J, \sigma_J^2)$. The horizon is $H = 48$. Contexts use $S_0 \sim \mathrm{LogNormal}(\log 100, .04^2)$, $\log(K/S_0) \sim \mathcal{N}(0, .08^2)$, $V_0 \sim \mathrm{LogNormal}(\log .035, .35^2)$, $s_0 \sim \mathrm{LogNormal}(\log .0025, .40^2)$, $\ell_0 \sim \mathrm{LogNormal}(0, .35^2)$, $\lambda_J \sim \mathrm{Unif}(.4, 2.2)$, and $\sigma_J \sim \mathrm{Unif}(.015, .055)$. For $R_t = |\log(S_{t+1}/S_t)| / \max\{\sqrt{V_t \Delta t}, 10^{-4}\}$, spread and liquidity evolve as

$$\log s_{t+1} = .94 \log s_t + .06 \log s_0 + .10(R_t - .8) + .06\xi_t^s,$$
$$\log \ell_{t+1} = .93 \log \ell_t + .07 \log \ell_0 - .11(R_t - .8) + .07\xi_t^\ell, \quad (32)$$

with independent standard-normal innovations and bounds $s_t \in [2\times10^{-4}, .035]$ and $\ell_t \in [.08, 8]$.

Frozen rules track Black–Scholes delta with adaptive no-trade bands, fixed rebalancing intervals, and trade caps. Trade h_t costs

$$\mathrm{TC}_t = \mathrm{spread}_t |h_t| S_t + .0015 h_t^2 S_t / \mathrm{liquidity}_t. \quad (33)$$

The four rules use $(.080, 4, .80, .20, .35)$, $(.012, 1, 0, .35, .65)$, $(.035, 1, 2.40, .25, .55)$, and $(.025, 1, .35, 1.40, .75)$ for base band, rebalancing interval, spread sensitivity, volatility sensitivity, and trade cap. Path statistics are terminal loss, maximum drawdown, mean/maximum absolute hedge error, total transaction cost, turnover, and $L^{\mathrm{risk}} = L_T + .35 D_{\max} + .10\,\mathrm{TC}$. Coupling metrics use this vector; CVaR uses changes in L^{risk} between rules on shared paths.

Table 7: Independent Finance inverse-EOT optimization traces and maximum final calibration residuals across market evaluations. Final convergence does not certify fixed-iteration training convergence.

Rep.	Epochs	Selected	Initial NLL	Selected NLL	Row residual	Col. residual
1	100	99	2.8268	0.2338	1.00e-05	3.47e-17
2	100	100	2.8758	0.2428	1.00e-05	3.47e-17
3	100	100	2.8349	0.2299	1.00e-05	3.47e-17
4	100	100	2.8060	0.2300	1.00e-05	3.47e-17
5	100	100	2.7814	0.2103	1.00e-05	3.47e-17

Five models trained on independent datasets are crossed with five independent evaluation market datasets. Training/validation/evaluation use 96/32/48 contexts; training and validation have 200 rollouts per rule, including 20 aligned per edge, and evaluation has 32. Separate random-stream identifiers distinguish data roles, replicates, contexts, and paths; path identifiers are invariant to requested sample count. Methods share data within blocks, and policies intentionally share paths within scenarios.

PACT uses temperature .5 and adaptive log-Sinkhorn projection. Primary comparators are inverse EOT and forward pair diffusion with exact marginal projection. Other baselines are backward/mixture diffusion, bidirectional/mixture paired CFM, joint CFM, shared-latent flow, task-matched inverse-EOT likelihood, shuffled controls, and independence. Training-standardized path summaries x enter a width-128 critic

$$s_\theta(x, y, c) = \mathrm{MLP}_\theta([x, y, y - x, c]),$$

with two SiLU layers and LayerNorm after the first. PACT averages row/column cross-entropies on aligned diagonals, using within-context off-diagonal endpoints as negatives. AdamW uses learning rate 3×10^{-4}, 100-epoch cap, batch size 64, and validation patience 12. Inverse EOT shares critic/data but uses calibrated plan likelihood through 40 training Sinkhorn iterations. Temperature-scaled scores are clipped to $[-12, 12]$ during training and final calibration; final residuals do not certify fixed-iteration training gradients. Scaled dependence energy is primary, with two prespecified comparisons; CVaR, joint error, regret, and Top-1 are secondary. Every baseline is independently trained under its declared architecture, budget, and stopping rule. This application uses contrastive PACT, not split correction. Inference resamples training/market axes separately (Appendix F.3).

The finite training budget does not establish convergence of the critic.

C.8 Evaluation metrics and aggregation

Aggregation follows each task's confirmation design. Application comparisons and matched ablations use different data protocols, so their numerical levels do not measure a before-and-after effect.

Conditional prediction. Swiss conditional MMD averages results over source points $(-2, 0), (2, 2), (0, 0)$. Each compares 1,024 predicted and 1,024 reference samples using $k(x, y) = \frac{1}{5}\sum_{j=0}^{4} \exp[-\|x - y\|^2/(2\sigma_j^2)]$, where $\sigma_j = 2^j d_{\mathrm{med}}/4$ and d_{med} is the median positive pairwise distance in the concatenated evaluation endpoint pools. For $X_{1:N}$ and $Y_{1:M}$, unbiased squared MMD (Gretton et al., 2012) is

$$\frac{\sum_{i \neq j} k(X_i, X_j)}{N(N-1)} + \frac{\sum_{i \neq j} k(Y_i, Y_j)}{M(M-1)} - \frac{2\sum_{i,j} k(X_i, Y_j)}{NM}.$$

This estimate can be negative. Unconditional MMD uses this kernel on full evaluation pools. Colored MNIST probabilities P_{ij} use the candidate bank, with aligned targets at column i. Pair NLL is $-N^{-1}\sum_i \log \max(P_{ii}, 10^{-12})$; Recall@1 is the fraction of rows maximized at i. For uniform calibrated plans, $P = N\Pi$. Marginal residual is maximum absolute row/column deviation from $1/N$.

Coupling estimation. Finite-support risks are $\mathrm{KL}(\pi^\star \| \Pi_{\widehat{\theta}})$ and $\frac{1}{2}\|\pi^\star - \Pi_{\widehat{\theta}}\|_1$, where $\Pi_{\widehat{\theta}}$ uses true marginals (uniform for 3×3) to isolate score estimation. Finance energy discrepancy for predictions

x_i, weights w_i, and M references y_j is

$$\mathcal{E}_w(x, y) = \frac{2}{M} \sum_{i,j} w_i \|x_i - y_j\| - \sum_{i,k} w_i w_k \|x_i - x_k\| - \frac{1}{M^2} \sum_{j,k} \|y_j - y_k\|.$$

Dependence energy averages this discrepancy across target-minus-source summary coordinates, divided in the application by training mean absolute paired differences, floored at 10^{-6}. Matched ablations use training feature standard deviations (Section C.5). Joint energy uses concatenated summaries standardized on training data, with Monte Carlo samples from the predicted weighted distribution.

Tail risk and decisions. CVaR measures upper-tail risk (Rockafellar et al., 2000). Weighted CVaR averages the highest $1 - \alpha$ mass, including fractional cutoff mass, with $\alpha = .9$. Unweighted Queue CVaR averages the largest $\lceil (1-\alpha)N \rceil$ values. Queue evaluates $\max(C_1 - C_0, C_2 - C_0)$ and divides the absolute CVaR difference by the training mean absolute cost difference across policy pairs, floored at 10^{-5}. Finance evaluates the difference in risk-adjusted shortfall and uses its corresponding training mean absolute difference scale. Finance actions minimize estimated CVaR of changes from rule 0. Regret subtracts minimum true risk from the chosen action's true risk, divided by the same training scale; Top-1 is the fraction of contexts with matching predicted/true minimizers. Crossed-model uncertainty is defined in Section F.3.

C.9 Binary verification of the bias direction

The binary law has uniform marginals, $\pi^\star(x, y) = (1 + \kappa xy)/4$, $f_\theta(x, y) = \theta xy$, $\theta_\star = \operatorname{arctanh}(\kappa)$, and unit temperature. The grid crosses $\kappa \in \{.25, .50, .70, .85\}$, $\rho \in \{0, .5, 1, 2, 4\}$, and $b \in \{16, 64, 256\}$, using 50,000 independent batches per configuration at seeds 20260924, 20260925, 20260926 (9,000,000 total). Paired counts are multinomial; endpoint counts are independently binomial with $u = \rho b$ draws per side. Marginals follow Equation (4); paired frequencies weight the likelihood. No model is fitted.

For first-category marginals p, q, the plan $(z, p - z, q - z, 1 - p - q + z)$ has odds ratio $e^{4\theta}$. The feasible quadratic root gives z, with interior derivative $dz/d\theta = 4/(1/z + 1/(p - z) + 1/(q - z) + 1/(1 - p - q + z))$. Differentiation holds paired frequencies fixed. Degenerate margins have a unique active-support plan and zero gradient; all 41 such batches are retained. Independent bisection and centered finite differences verify calibration/gradients; maximum marginal residual is 1.11×10^{-16}.

The control variate $g_{\text{cv}} = g_{\text{raw}} - (\kappa - b^{-1} \sum_i x_i y_i)$ subtracts a zero-expectation term with fixed coefficient one, using synthetic truth only for Monte Carlo variance reduction. Each configuration pools 150,000 batches; pointwise 95% normal intervals use empirical standard deviation divided by the square root of this count. Section D.8 summarizes sign agreement for both estimators and plots the variance-reduced discrepancies across the complete grid. These intervals measure Monte Carlo uncertainty, not trained-model performance or simultaneous coverage; the control variate is not a training correction.

C.10 Regularized comparison with modified profile likelihood

Raw paired likelihood, ratio-preserving split jackknife (JK), and modified profile likelihood (MP) are compared on eight configurations: $m \in \{3, 6\}$, uniform or nonuniform marginals, and $b \in \{64, 256\}$, with $\rho = 0$ throughout. At category angles $a_i = 2\pi i/m$, define $v_i = (\sqrt{2} \cos a_i, \sqrt{2} \sin a_i)$. Four features are ordered outer products $v_{ir} v_{js}$, $r, s \in \{1, 2\}$, and $f_\theta(i, j) = \theta^\top F_{ij}$ with $\theta_\star = (.8, -.4, .3, .7)$. Truth is the unit-temperature calibrated plan. Uniform marginals are $1/m$; nonuniform marginals are $\operatorname{softmax}(\operatorname{linspace}(-.5, .5, m))$ and $\operatorname{softmax}(\operatorname{linspace}(.4, -.4, m))$. Datasets sum two independent size-$b/2$ multinomial tables; methods share full tables and complementary splits.

MP minimizes $L_{\text{raw}} - \log\det K_\theta/(2b)$, with reduced endpoint-indicator covariance K_θ at empirical calibration (Appendix B.11). Each complete objective adds $\lambda\|\theta\|^2/2$ once, without a parameter box. For each method/configuration, 48 development datasets select $\lambda \in \{.001, .01, .1\}$ by simulation-oracle population-margin KL, ties favoring smaller λ. Evaluation uses 256 fresh paired datasets per configuration; selected values appear in Appendix D.9.

Unbounded L-BFGS-B starts from zero and $\pm.25$ times two unit vectors obtained by normalizing standard Gaussian four-vectors (NumPy default generator seed 730120260924). Settings are 1,000 iterations, function tolerance zero, gradient tolerance 10^{-8}, and 60 line-search steps; the lowest finite objective selects the fit. Double-precision calibration requires absolute marginal residual at most 10^{-12} within 300 iterations and selected outer gradient infinity norm at most 10^{-6}. All 6,144 confirmation fits pass; stationarity does not certify global optimality. Zero-margin rows/columns are removed separately in full/half tables. No pseudocount or covariance ridge is added; one indicator per active endpoint is dropped, with zero-dimensional determinant one. This boundary convention is separate from the positive-interior theory. NumPy SeedSequence keys are $(20260924, k, c, i)$: $k = 7301$ for development, 7302 for confirmation; zero-based c orders m, marginal type (uniform first), then b; i indexes datasets.

Primary KL and descriptive TV use true population marginals to isolate relation estimation. Empirical-margin TV and parameter squared bias, variance, and MSE are also reported. Across 16 paired primary contrasts (JK minus Raw/MP in eight configurations), pointwise and Bonferroni intervals use $t_{.975,255}$ and $t_{1-.05/(2\cdot16),255}$, respectively, times the sample standard deviation of 256 differences divided by $\sqrt{256}$. Zero-crossing intervals do not establish equivalence. Data and selection budgets are shared; computational costs differ.

C.11 Expected-objective minimizers

The 3×3 centered basis, uniform marginals, and truths $(.25, 1)$ and $(.25, 2)$ are crossed with $\rho \in \{0, 1\}$ and $b \in \{16, 32, 64, 128\}$. Three independent Monte Carlo repetitions each average 32,768 complete batches, formed from independent ratio-preserving halves. Optimization minimizes the average objective, not average dataset-specific minimizers. A zero-expectation linear score control variate uses synthetic truth to reduce integration variance without changing the expectation; it is absent from application training.

Starts are the truth and predicted displacement; local boxes have radii 1 and 1.5. Checks cover stationarity, numerical Hessians, boundary hits, and start/region sensitivity. Double-precision calibration uses tolerance 2×10^{-11} and cap 10,000. Finite differences and independent likelihood evaluation verify zero-count cases. Uncertainty is measured across the three Monte Carlo approximations, not optimizer tolerance.

C.12 Independent evaluation-temperature confirmation

Paired likelihood and split correction share a three-hidden-layer width-128 MLP, 128 training pairs, initialization and pair stream within block. Updates use two 64-pair groups, training temperature .05, AdamW learning rate/weight decay 10^{-4}, gradient clip 5, and 6,000 updates. Evaluation regimes reuse models, changing prediction score scale only.

Three development blocks select temperature per objective/task from $\{.02, .04, .08, .16, .32, .64, 1.28\}$, allowing two factor-two boundary extensions per side and at most 16 bounded log-temperature refinement evaluations (tolerance .005). Fixed temperature is .08. Checkpoints/selections precede ten confirmation blocks; data, initialization, and reference draws are fresh between blocks and shared within blocks.

KL averages oracle-to-model divergence across eight disjoint 128×128 banks with uniform marginals; the oracle uses known rotation cost and temperature .05. Conditional MMD uses 1,024 endpoints and the original sampled prediction/reference target. Development integrates model-output sampling for smooth selection; confirmation retains the sampled metric. Double-precision calibration uses relative tolerance 10^{-8}, reporting acceptance 10^{-6}, and a residual-checked convex-dual fallback.

Two primary contrasts compare tuned corrected versus tuned raw, one per task, averaging banks/starting points within blocks. Paired bootstrap intervals and exact sign-flip tests use Bonferroni intervals and Holm adjustment across these endpoints. Sign-flip inference assumes null symmetry of paired differences. Fixed-temperature comparisons are secondary; stopping and selection do not use confirmation outcomes.

Table 8: Paired-only gradient validation across batch sizes. Relative error compares $b\nabla\mathbb{E}\widehat{\mathcal{L}}_{b,0}(\theta^\star)$ with $\nabla A_0(\theta^\star)$; cosine compares their directions. JK/raw is the norm ratio of the cross-seed mean gradients.

Truth	b	Coefficient rel. error	Direction cosine	JK/raw gradient	Min. JK Hessian eig.
Moderate	8	.306	.978	.636	.2589
	16	.149	.997	.277	.2417
	20	.132	.996	.257	.2379
	32	.058	.999	.236	.2328
	64	.031	1.000	.020	.2304
Strong	8	.483	.998	1.169	.0422
	16	.261	.999	.599	.0346
	20	.223	1.000	.456	.0322
	32	.134	1.000	.298	.0296
	64	.062	1.000	.171	.0287

Table 9: Mixed-support diagnostics across all tested batch sizes with an equal number of paired and additional endpoint observations. The expansion predicts the gradient direction asymptotically, but the split correction is not uniformly beneficial at small batch sizes.

Truth	b	Coefficient rel. error	Direction cosine	JK/raw gradient
Moderate	8	.506	.865	1.359
Moderate	16	.647	.777	1.091
Moderate	20	.479	.933	1.308
Moderate	32	.126	.996	.721
Moderate	64	.569	.927	.342
Strong	8	.964	.310	2.079
Strong	16	.727	.921	1.602
Strong	20	.628	.962	1.317
Strong	32	.455	.987	.923
Strong	64	.246	.997	.516

Table 10: Finite-support risk on all 500 datasets per truth. Each row reports raw likelihood / split-jackknife. Both parameter ranges use the same datasets. Every refit satisfies the calibration and projected-KKT criteria; boundary counts do not imply interior or global optimality.

Box	Truth	Mean KL (raw / JK)	Mean TV (raw / JK)	Boundary (raw / JK)
±4	moderate	0.258903/0.234383	0.135280/0.127766	188/174
±4	strong	0.295664/0.207370	0.108020/0.088209	461/407
±6	moderate	0.432939/0.382416	0.138491/0.129260	186/167
±6	strong	0.526243/0.321792	0.119223/0.089714	461/399

D Additional results and diagnostics

D.1 Finite-minibatch expansion and estimation risk

The coefficient study uses independent Monte Carlo archives for each batch size and three frozen seeds. The decreasing coefficient error and stable direction support the first-order expansion. The smallest strong-dependence batch lies outside the useful correction regime and is retained in the table.

Both objectives use the same 500 datasets and splits in both parameter boxes, with no exclusions. The numerical checks certify calibration residuals and constrained first-order optimality, not global minima of the nonconvex corrected objective. Strong-dependence estimates remain frequently on the boundary, so these risk results are not a direct verification of the interior minimizer-displacement formula. Parameter variance increases under strong dependence even though squared bias, MSE, KL,

Table 11: Paired mean differences (JK minus raw) with pointwise 95% t intervals, using 500 datasets and 499 degrees of freedom.

Box	Truth	Δ KL	Δ TV
± 4	moderate	-0.02452 $[-0.04031, -0.00873]$	-0.00751 $[-0.01059, -0.00444]$
± 4	strong	-0.08829 $[-0.10088, -0.07571]$	-0.01981 $[-0.02295, -0.01667]$
± 6	moderate	-0.05052 $[-0.07829, -0.02276]$	-0.00923 $[-0.01279, -0.00567]$
± 6	strong	-0.20445 $[-0.22900, -0.17991]$	-0.02951 $[-0.03347, -0.02555]$

Table 12: Descriptive parameter-error decomposition (raw / JK). Variance uses normalization by the number of datasets, so squared bias plus variance equals MSE. These are parameter statistics, not a decomposition of KL or TV.

Box	Truth	Squared bias	Variance	MSE
± 4	moderate	$1.3127/1.1308$	$3.1535/3.0978$	$4.4662/4.2286$
± 4	strong	$3.2035/2.3865$	$2.8968/3.3168$	$6.1003/5.7033$
± 6	moderate	$3.8519/3.3315$	$8.2535/8.0138$	$12.1053/11.3452$
± 6	strong	$13.0778/10.1579$	$7.0860/8.7736$	$20.1637/18.9315$

Table 13: Frozen Swiss-roll transfer check over five paired confirmation blocks. Lower is better. The fixed conditional MMD is the prespecified primary endpoint. The paired interval is a bootstrap 95% interval for corrected minus uncorrected conditional MMD. Interaction error is an oracle secondary diagnostic and was analyzed separately from the primary metric.

Objective	Fixed conditional MMD	Native conditional MMD	Interaction rel. RMSE
Uncorrected calibrated likelihood	$.3054 \pm .1883$	$.3588 \pm .1355$	4.7754 ± 1.0093
Ratio-preserving split correction	$.3107 \pm .1916$	$.3573 \pm .1356$	$4.3960 \pm .9554$
Paired corrected−uncorrected	$.0053$ $[-.0001, .0114]$	$-.0015$	$-.3793$
Corrected wins	$1/5$	$2/5$	$5/5$

Table 14: Swiss application confirmation replicates. The compared outputs are PACT's calibrated conditional sampler and the baselines' native samplers. The residual is PACT's maximum absolute marginal error on its calibration bank.

Replicate	PACT MMD	EBiEOT MMD	Paired MDN MMD	PACT residual
1	0.140909	0.456428	0.307743	$1.63e-08$
2	0.056108	0.459346	0.765756	$6.52e-09$
3	0.126479	0.448737	0.170816	$5.12e-08$

and TV decrease on average. The gradient experiments in Tables 8 and 9 are separate from these risk refits.

D.2 Learned-score transfer of the split correction

Correction improves interaction recovery in all five blocks but does not improve the prespecified conditional-prediction endpoint: its mean error is higher, only one block improves, and the paired interval crosses zero. This distinction does not identify the approximation, optimization, or sampling errors responsible for the difference.

D.3 Conditional application results and visualizations

Table 14 reports all Swiss confirmation replicates and calibration residuals; the complete conditional-task comparison, including the specified-form circular reference, appears in Table 4.

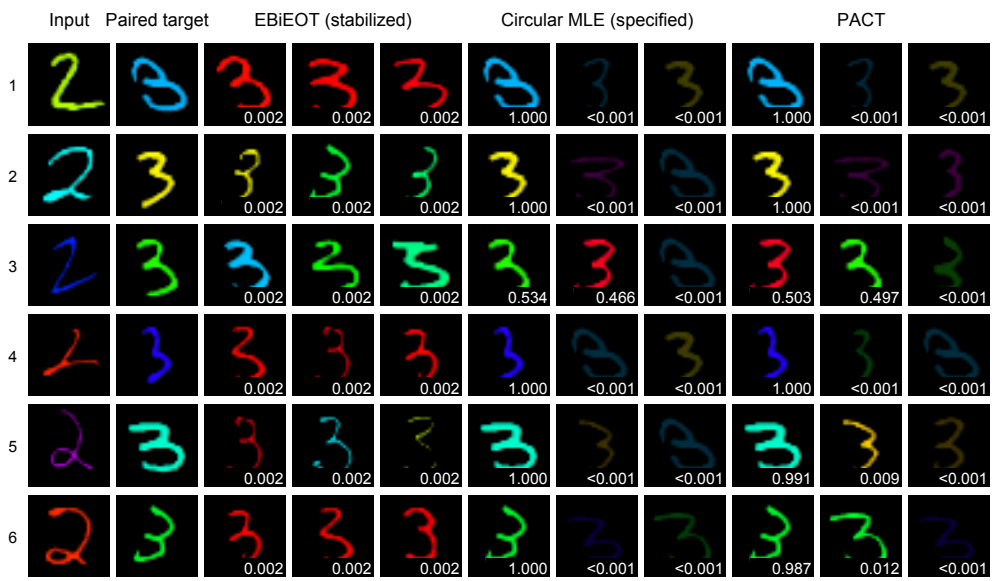

Figure 5: Colored-MNIST retrieval for six fixed test inputs. Each method's three columns show its top-ranked candidates, from left to right, with their conditional probabilities; these are retrieved images, not generated samples. Candidates with probability below 0.001 are dimmed identically across methods.

Colored-MNIST retrieval. Figure 5 shows PACT's learned color correspondences alongside an inverse-transport baseline and the specified-form circular reference. The first six inputs of the first confirmation replicate are fixed before inspection; all methods use the same 512-target bank and frozen application models, without split correction. A recorded paired target is one valid correspondence; multimodal predictions can assign probability to other targets.

Swiss-roll visualization protocol. Figures 4, 6 and 7 use frozen application models from Section C.3, separately from the correction and objective-ablation studies. The main figure uses the first replicate, fixed before sampling; the supplementary views retain all three. PACT retains its MLP, temperature .05, target bank, and calibration potentials; baselines use their native samplers. Ground truth comes from the original evaluation, while model outputs are redrawn from the frozen models.

Each method draws 1,024 marginal outputs at the original test inputs and 1,024 conditional outputs at each anchor $(-2, 0), (2, 2), (0, 0)$. Panels display the first 512 marginal and 192 conditional draws per anchor, without value-based selection or jitter; segments show the first eight conditional draws. Out-of-frame counts and full-range views retain outliers. MMD annotations are the original three-replicate means and sample standard deviations, not metrics recomputed from the displayed samples.

D.4 Native and calibrated conditional outputs

The Swiss baseline native row uses its generative sampler. Its projection diagnostic fits a balanced OT plan between native and target reference samples, takes the row-wise maximizing target index, and maps new outputs through their nearest native reference point. This many-to-one operation does not guarantee an exact target multiset. Colored-MNIST raw-score rows use row-softmax; these depend on column score gauges and are not intrinsic measures of compatibility quality.

The Swiss projection changes unconditional MMD with little change in conditional error. Colored-MNIST matrix calibration satisfies the endpoint constraints to numerical tolerance, while conditional accuracy depends on the learned scores. The Swiss projection diagnostic does not guarantee exact empirical marginal matching.

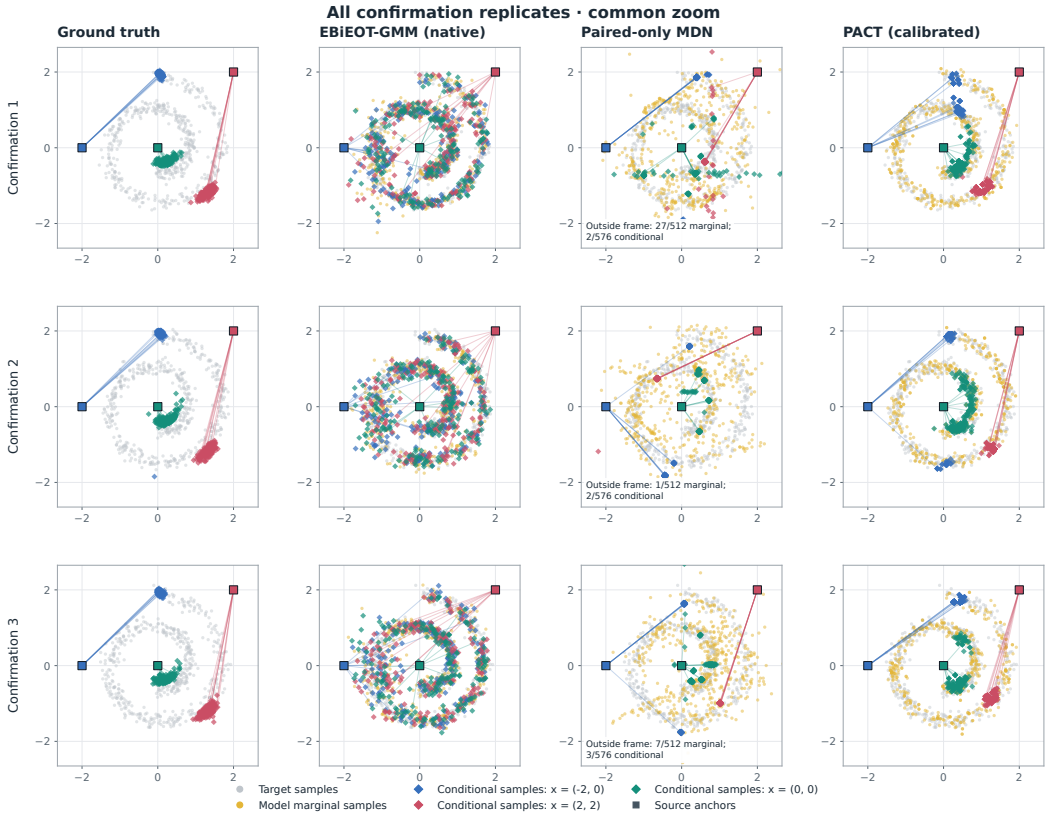

Figure 6: Swiss-roll conditional predictions across all three confirmation replicates, using a common viewing window. Colors and symbols follow Figure 4; rows follow the original replicate order. Out-of-frame sample counts are shown for each affected model panel.

Table 15: Conditional-task results before and after endpoint calibration. Lower is better except retrieval. The unbiased unconditional MMD estimate can be slightly negative.

Task	Output	Primary error	Auxiliary error	Retrieval
Swiss	PACT calibrated	Cond. MMD .1078	Uncond. MMD .0016	–
Swiss	EBiEOT native	Cond. MMD .4548	Uncond. MMD .0012	–
Swiss	EBiEOT projection	Cond. MMD .4547	Uncond. MMD $-.0008$	–
CMNIST	PACT raw score	NLL .4723	Marg. err. .00310	R@1 .9017
CMNIST	PACT calibrated	NLL .0978	Marg. err. 2×10^{-7}	R@1 .9818
CMNIST	EBiEOT raw score	NLL 24.6677	Marg. err. .44078	R@1 .0020
CMNIST	EBiEOT calibrated	NLL 6.2338	Marg. err. $< 10^{-7}$	R@1 .0020

D.5 Queue endpoint matching and full baseline comparison

PACT's endpoint-matched output has lower mean normalized CVaR error than all evaluated baselines in each mechanism (see Table 16). Native replay remains more accurate; endpoint matching increases its error by .1361, .3913, and .1582 for nominal, stress, and misspecified mechanisms, respectively. The endpoint-matched model wins 28/48, 31/48, and 26/48 contexts against the respective lowest-mean baseline. It also has the lowest mean worst-cost energy discrepancy among matched methods with this metric reported (see Table 17).

Across nominal and stress contexts, PACT's mean CVaR error is .3668, compared with .5537 for tuned cost rank, the lowest pooled-mean baseline. The paired difference is $-.1869$ with a stratified paired 95% bootstrap interval $[-.2476, -.1279]$. This interval concerns that pooled contrast, not each

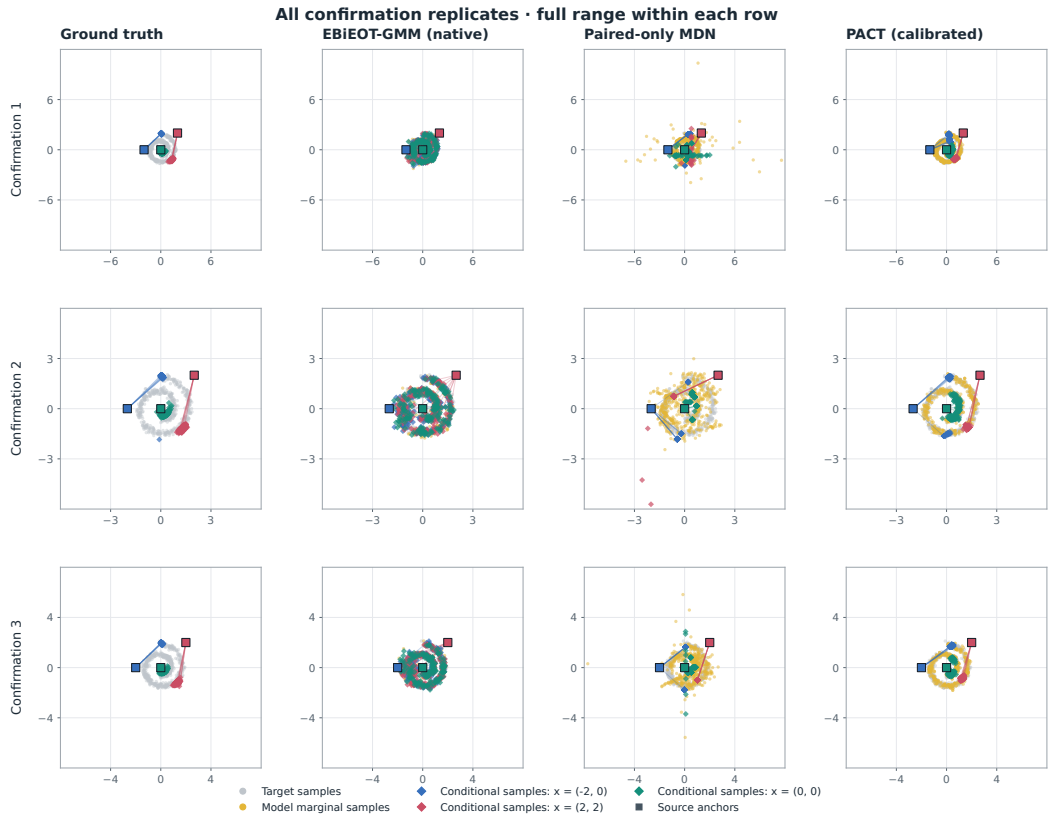

Figure 7: Full-range views of the samples in Figure 6. Axis limits are shared across methods within each row and include every displayed sample, retaining the MDN's outlying outputs. Colors and symbols are unchanged.

mechanism separately, and holds the selected baseline family fixed (Appendix C.6). The misspecified mechanism is reported separately. All endpoint-matched graph outputs have zero recorded empirical marginal error.

D.6 Finance baseline comparisons

The independent rerun retains every output variant, including native generators and shuffled controls; see Table 18 and Table 19. The prespecified comparisons in Table 5 concern common empirical marginals. Some native generator outputs have lower distributional errors, so this comparison does not establish a ranking across all output types.

PACT has lower dependence and joint error than both prespecified comparators. Decision outcomes depend on the comparator: regret and Top-1 favor PACT over projected diffusion but are unresolved against inverse EOT. Full intervals appear in Table 32.

D.7 Boundary targets

Deterministic correspondences lie on the boundary of the transport polytope. For $\pi^\star_\eta = (1-\eta)\pi^\star + \eta pq^\top$, the positive theory applies and $\mathrm{TV}(\pi^\star_\eta, \pi^\star) \le \eta$. This bound holds for fixed positive η. Allowing η to decrease with sample size additionally requires uniform control of the regularity bounds and expansion remainder.

Table 16: Queue normalized CVaR error for the full confirmation suite. Each mechanism contains 48 contexts; learned-method means average two optimizer replicates. Lower is better. Native replay is a reference, and shuffled rows are controls excluded from the primary baseline competition.

Method	Nominal	Stress	Misspecified
PACT (endpoint-matched)	0.2211	0.5125	0.2313
PACT (native replay)	0.0850	0.1212	0.0731
Tuned cost rank	0.2696	0.8377	0.2954
Comonotone copula	0.4968	0.7467	0.5275
Gaussian factor MLE	0.6442	0.6700	0.7275
Student-t factor MLE	0.6530	0.6974	0.7376
Edge encoder, inverse EOT	1.2433	4.7961	1.4226
Edge encoder, matching	1.3368	4.7070	1.3130
Shared encoder, matching	1.3893	5.0920	1.3918
Shared encoder, inverse EOT	1.4534	5.2695	1.4161
Paired CFM	1.6579	7.4179	1.7642
Path-metric tree	1.7387	7.4632	1.8589
Pair critic, inverse EOT	1.7803	7.1868	1.7868
Pair critic, matching	1.8937	7.4694	1.9411
Joint CFM	2.0508	8.0548	2.2643
Pair diffusion	2.0950	8.1418	2.3380
EBiEOT adapter	2.1248	8.1330	2.3407
Shared-index tree	2.1410	8.1679	2.3782
Independent copula	2.1582	8.2877	2.3744
Independent tree	2.1625	8.1327	2.3697
Paired CFM (shuffled)	1.6630	7.4443	1.7804
Joint CFM (shuffled)	2.0674	8.0805	2.2814
Pair diffusion (shuffled)	2.1023	8.1424	2.3393
EBiEOT adapter (shuffled)	2.1248	8.1330	2.3407
Shared encoder, inverse EOT (shuffled)	2.1274	8.1301	2.3389
Shared encoder, matching (shuffled)	2.1276	8.1269	2.3440
Edge encoder, matching (shuffled)	2.1283	8.1289	2.3405
Edge encoder, inverse EOT (shuffled)	2.1361	8.1228	2.3374
Pair critic, matching (shuffled)	2.2065	8.2930	2.3810
Pair critic, inverse EOT (shuffled)	2.2128	8.3736	2.4213

D.8 Spectral verification

All 60 variance-reduced point estimates have the predicted sign, with their pointwise intervals lying on that side of zero. The raw estimates match in 56 configurations. No configuration lies exactly at the asymptotic threshold, so the grid checks the directional pattern rather than locating an exact finite-batch zero crossing. Mean absolute discrepancy from the leading coefficient decreases across batch sizes, but only 14 of 20 fixed (κ, ρ) groups decrease monotonically; 18 are closer at $b = 256$ than at $b = 16$. At $b = 256$, 10 of 20 intervals contain the leading coefficient. Strong-dependence discrepancies therefore remain material.

D.9 Regularized classical-correction comparison

The complete comparison includes all eight prespecified configurations (see Table 2). The simultaneous intervals identify higher JK KL than Raw in configurations 0, 4, and 6; all other JK–Raw intervals and all JK–MP intervals cross zero (see Table 20). The latter do not establish equivalence. These results assess a different four-feature, regularized model from the earlier two-feature box-constrained study and are not a controlled ablation of regularization alone.

Table 21 retains the secondary metrics. Primary risk uses true marginals; empirical-margin TV also includes marginal estimation error. Selection minimizes simulation-oracle development KL over $\lambda \in \{.001, .01, .1\}$ using 48 datasets per configuration. All methods select .1 in configurations 0, 4, and 6, and .01 otherwise. These results therefore concern the stated grid and truth, not exhaustive

Table 17: Queue worst-cost energy discrepancy for the full confirmation suite. Lower is better. Rows and aggregation match the primary comparison; native replay and shuffled controls are reported separately from claims about endpoint-matched baselines. A dash indicates an unavailable secondary metric.

Method	Nominal	Stress	Misspecified
PACT (endpoint-matched)	0.0228	0.1803	0.0401
PACT (native replay)	0.0048	0.0045	0.0059
Tuned cost rank	–	–	–
Comonotone copula	0.0814	0.2634	0.1330
Gaussian factor MLE	0.0421	0.2459	0.0635
Student-t factor MLE	0.0411	0.2417	0.0625
Edge encoder, inverse EOT	0.1177	1.0419	0.1434
Edge encoder, matching	0.1203	1.0419	0.1304
Shared encoder, matching	0.1357	1.1179	0.1389
Shared encoder, inverse EOT	0.1468	1.1701	0.1468
Paired CFM	0.1946	1.6883	0.2183
Path-metric tree	0.2127	1.7098	0.2345
Pair critic, inverse EOT	0.1583	1.5357	0.1856
Pair critic, matching	0.1932	1.6077	0.2219
Joint CFM	0.2911	1.8377	0.3341
Pair diffusion	0.2962	1.8635	0.3491
EBiEOT adapter	0.3120	1.8618	0.3611
Shared-index tree	0.3302	1.9084	0.3695
Independent copula	0.3101	1.8886	0.3555
Independent tree	0.3303	1.9156	0.3563
Paired CFM (shuffled)	0.1988	1.6980	0.2214
Joint CFM (shuffled)	0.2971	1.8436	0.3392
Pair diffusion (shuffled)	0.2977	1.8628	0.3490
EBiEOT adapter (shuffled)	0.3120	1.8618	0.3611
Shared encoder, inverse EOT (shuffled)	0.3125	1.8612	0.3590
Shared encoder, matching (shuffled)	0.3129	1.8625	0.3601
Edge encoder, matching (shuffled)	0.3126	1.8619	0.3586
Edge encoder, inverse EOT (shuffled)	0.3154	1.8602	0.3557
Pair critic, matching (shuffled)	0.3346	1.9246	0.3726
Pair critic, inverse EOT (shuffled)	0.3438	1.9349	0.3921

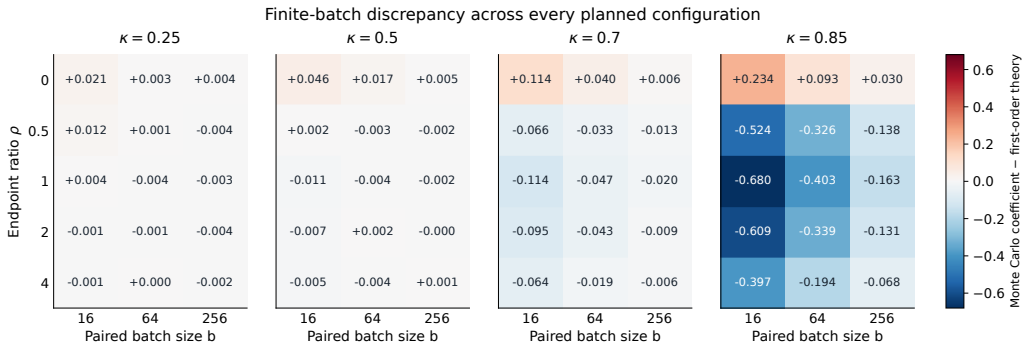

Figure 8: Finite-batch deviations from the first-order gradient coefficient for every tested configuration. The same scaled Monte Carlo estimates as Figure 3 are compared with theory; stronger dependence retains larger magnitude deviations. No configuration is omitted.

tuning or a universal performance ordering. Equal data and selection budgets do not imply equal

Table 18: All 34 Finance output variants in the independent rerun. Projected denotes explicit endpoint projection; PACT and inverse EOT use calibrated plans. The table retains both native and projected generator outputs. Lower is better except Top-1.

Output variant	Dep.	Joint	CVaR	Regret	Top-1
independent product	0.9033	0.3231	6.2851	0.2956	0.5067
Inverse EOT shuffled	0.9070	0.3234	6.2917	0.3080	0.4983
Inverse EOT	0.1095	0.2723	1.0694	0.1373	0.6600
joint CFM shuffled	1.1435	0.3794	6.6788	0.4522	0.3567
joint CFM shuffled (projected)	0.9122	0.3515	5.9415	0.4188	0.3875
joint CFM	0.1828	0.3324	1.3575	0.4170	0.3900
joint CFM (projected)	0.1943	0.2803	1.8494	0.3777	0.4175
PACT shuffled	0.9056	0.3234	6.2921	0.3128	0.4958
PACT	0.1016	0.2704	1.0192	0.1403	0.6608
pair diffusion shuffled backward	1.1506	0.4862	5.8861	0.4218	0.3717
pair diffusion shuffled backward (projected)	0.9298	0.3551	6.0381	0.4230	0.3625
pair diffusion shuffled forward	1.1035	0.3973	5.1046	0.4708	0.3375
pair diffusion shuffled forward (projected)	0.9312	0.3560	6.0748	0.4291	0.3725
pair diffusion shuffled mixture	0.9750	0.3311	5.7956	0.4434	0.3492
pair diffusion shuffled mixture (projected)	0.9660	0.3615	6.1161	0.4133	0.3917
pair diffusion backward	0.1354	0.3088	1.6436	0.4167	0.3850
pair diffusion backward (projected)	0.1273	0.2779	1.1536	0.3934	0.3942
pair diffusion forward	0.1210	0.2806	1.0278	0.4839	0.2942
pair diffusion forward (projected)	0.1192	0.2779	1.0549	0.4061	0.3692
pair diffusion mixture	0.0920	0.2562	1.2060	0.4414	0.3408
pair diffusion mixture (projected)	0.2364	0.2848	2.2687	0.4068	0.3750
paired CFM shuffled backward	0.2373	0.3370	1.6102	0.2411	0.5608
paired CFM shuffled backward (projected)	0.1569	0.2842	1.7859	0.1309	0.6650
paired CFM shuffled forward	0.2179	0.3146	1.6971	0.1310	0.6667
paired CFM shuffled forward (projected)	0.1567	0.2848	1.7957	0.1364	0.6667
paired CFM shuffled mixture	0.1721	0.2895	1.5826	0.1624	0.6400
paired CFM shuffled mixture (projected)	0.1719	0.2890	1.7498	0.2895	0.5242
paired CFM backward	0.1336	0.2849	1.5537	0.1129	0.6842
paired CFM backward (projected)	0.1335	0.2779	1.4504	0.1541	0.6442
paired CFM forward	0.1322	0.2816	1.5241	0.0943	0.7050
paired CFM forward (projected)	0.1323	0.2780	1.4284	0.1716	0.6308
paired CFM mixture	0.1213	0.2640	1.4956	0.0766	0.7192
paired CFM mixture (projected)	0.1377	0.2793	1.3715	0.2542	0.5442
shared latent flow	0.1598	0.2844	1.8251	0.1428	0.6525

computation: JK recalibrates three components per objective evaluation and MP differentiates a nuisance-information term.

Full tables retain every endpoint in both development and confirmation; some half tables lose endpoint categories, especially at $m = 6, b = 64$ (see Table 22). These datasets remain in the comparison under the active-support convention.

D.10 Expected-minimizer displacements and numerical checks

Under paired-only sampling, the mean of the three estimated expected-objective minimizers under correction is closer to the truth throughout the tested grid. Moderate-dependence mixed sampling at smaller batches instead increases displacement. The complete grid and checks below retain these unfavorable conditions. These results concern minimizers of expected objectives, not risks of estimators fitted to individual datasets.

Table 19: Remaining Finance distributional metrics for every output variant. Cost dependence and marginal energy are both errors (lower is better).

Output variant	Cost dependence	Marginal energy
independent product	1.1818	0.0188
Inverse EOT shuffled	1.1877	0.0188
Inverse EOT	0.1517	0.0188
joint CFM shuffled	1.5114	0.0310
joint CFM shuffled (projected)	1.1905	0.0188
joint CFM	0.1993	0.0484
joint CFM (projected)	0.1706	0.0188
PACT shuffled	1.1849	0.0188
PACT	0.1340	0.0188
pair diffusion shuffled backward	1.4783	0.0934
pair diffusion shuffled backward (projected)	1.2253	0.0188
pair diffusion shuffled forward	1.4467	0.0665
pair diffusion shuffled forward (projected)	1.2285	0.0188
pair diffusion shuffled mixture	1.2551	0.0443
pair diffusion shuffled mixture (projected)	1.2805	0.0188
pair diffusion backward	0.1672	0.0356
pair diffusion backward (projected)	0.1388	0.0188
pair diffusion forward	0.1642	0.0306
pair diffusion forward (projected)	0.1368	0.0188
pair diffusion mixture	0.1081	0.0540
pair diffusion mixture (projected)	0.2326	0.0188
paired CFM shuffled backward	0.3044	0.0453
paired CFM shuffled backward (projected)	0.2723	0.0188
paired CFM shuffled forward	0.2873	0.0385
paired CFM shuffled forward (projected)	0.2717	0.0188
paired CFM shuffled mixture	0.2340	0.0649
paired CFM shuffled mixture (projected)	0.1711	0.0188
paired CFM backward	0.2129	0.0226
paired CFM backward (projected)	0.2060	0.0188
paired CFM forward	0.2053	0.0254
paired CFM forward (projected)	0.2065	0.0188
paired CFM mixture	0.1966	0.0574
paired CFM mixture (projected)	0.1391	0.0188
shared latent flow	0.2786	0.0188

D.11 Fixed and tuned evaluation temperatures

The original fixed-procedure MMD study and this temperature confirmation evaluate different training and evaluation procedures; both are retained. At fixed temperature, coupling KL improves in all ten blocks. With independently tuned evaluation temperatures, neither primary comparison establishes superiority. Intervals containing zero do not establish equivalence or identify score scaling as the sole source of fixed-temperature gains. At fixed evaluation temperature, the corrected-minus-raw 95% intervals are $[-0.5163, -0.3193]$ for coupling KL and $[-0.0140, 0.0011]$ for conditional MMD.

E Matched-architecture comparisons

The learning-objective comparison is reported in Table 3.

The training and selection protocols appear in Appendices C.3 and C.5. Table 28 complements the main comparison with native-score diagnostics for Swiss roll and Colored MNIST; Finance outputs are consolidated in Table 30.

Colored MNIST calibrated Recall@1 is .9875 for paired contrastive training, .7367 for likelihood on pairs, and .8902 for likelihood on mixed support.

Table 20: All paired primary KL differences, multiplied by 10^4 for readability. Negative favors JK. Pointwise and Bonferroni simultaneous 95% t intervals use 255 degrees of freedom. Configurations 0–7 follow the row order of Table 2; simultaneous coverage concerns all 16 comparisons.

Config.	Comparison	Mean	Pointwise interval	Simultaneous interval
0	JK–Raw	11.213	[7.634, 14.791]	[5.790, 16.635]
0	JK–MP	2.206	[−1.305, 5.718]	[−3.113, 7.526]
1	JK–Raw	0.100	[−0.830, 1.030]	[−1.309, 1.508]
1	JK–MP	0.358	[−0.457, 1.174]	[−0.877, 1.594]
2	JK–Raw	0.023	[−2.447, 2.493]	[−3.719, 3.765]
2	JK–MP	3.387	[0.958, 5.816]	[−0.294, 7.067]
3	JK–Raw	0.252	[−0.162, 0.667]	[−0.376, 0.881]
3	JK–MP	0.123	[−0.292, 0.537]	[−0.506, 0.751]
4	JK–Raw	11.182	[7.307, 15.058]	[5.311, 17.054]
4	JK–MP	0.547	[−3.216, 4.310]	[−5.154, 6.248]
5	JK–Raw	−0.363	[−1.113, 0.387]	[−1.499, 0.773]
5	JK–MP	0.145	[−0.520, 0.809]	[−0.862, 1.152]
6	JK–Raw	8.528	[4.856, 12.201]	[2.965, 14.092]
6	JK–MP	−0.009	[−3.773, 3.755]	[−5.711, 5.693]
7	JK–Raw	−0.390	[−1.057, 0.277]	[−1.400, 0.620]
7	JK–MP	−0.104	[−0.703, 0.495]	[−1.012, 0.803]

Table 21: Descriptive secondary confirmation metrics. TV uses either true or empirical marginals. Parameter variance has denominator 256, so squared bias plus variance equals MSE. Configuration order matches the primary table.

Config.	Method	λ	TV true	TV empirical	Bias^2	Variance	MSE
0	Raw	0.1	0.06115	0.11771	0.05749	0.06706	0.12455
0	JK	0.1	0.06373	0.11905	0.06123	0.06771	0.12894
0	MP	0.1	0.06381	0.11910	0.06354	0.06439	0.12793
1	Raw	0.01	0.03075	0.05826	0.00050	0.05194	0.05244
1	JK	0.01	0.03107	0.05852	0.00054	0.05152	0.05206
1	MP	0.01	0.03090	0.05839	0.00053	0.05148	0.05200
2	Raw	0.01	0.04994	0.11364	0.04035	0.14271	0.18306
2	JK	0.01	0.05010	0.11369	0.03761	0.14340	0.18101
2	MP	0.01	0.04927	0.11336	0.03813	0.14094	0.17906
3	Raw	0.01	0.02604	0.05771	0.01834	0.08928	0.10762
3	JK	0.01	0.02615	0.05780	0.01855	0.08921	0.10776
3	MP	0.01	0.02605	0.05770	0.01850	0.08925	0.10775
4	Raw	0.1	0.07804	0.17578	0.02745	0.07249	0.09994
4	JK	0.1	0.07985	0.17655	0.03184	0.07224	0.10408
4	MP	0.1	0.08022	0.17659	0.03373	0.06988	0.10361
5	Raw	0.01	0.03970	0.08949	0.00009	0.03641	0.03650
5	JK	0.01	0.03967	0.08947	0.00008	0.03602	0.03610
5	MP	0.01	0.03968	0.08944	0.00010	0.03591	0.03601
6	Raw	0.1	0.07922	0.17861	0.02954	0.07751	0.10705
6	JK	0.1	0.08076	0.17937	0.03327	0.07735	0.11062
6	MP	0.1	0.08094	0.17922	0.03564	0.07488	0.11052
7	Raw	0.01	0.03831	0.08970	0.00016	0.03578	0.03594
7	JK	0.01	0.03827	0.08969	0.00027	0.03532	0.03559
7	MP	0.01	0.03833	0.08970	0.00027	0.03532	0.03558

Finance hyperparameters were selected separately for each objective under the same development budget. The paired-versus-mixed comparison was not a prespecified primary contrast. Native-output interpretation is given in Appendix F.1.

Full Finance means appear in Table 30; both interval levels appear in Table 31.

Table 22: Active-support diagnostics for the classical comparison. Counts are datasets with at least one absent endpoint category in the full or half table. There are 48 development and 256 confirmation datasets per configuration. Configurations 0, 1, 3, 5, and 7 have zero counts in every column; all nonzero configurations are shown. No dataset is excluded.

	Development			Confirmation		
Config.	Full	Half A	Half B	Full	Half A	Half B
2	0	0	0	0	1	1
4	0	1	4	0	8	8
6	0	5	5	0	26	24

Table 23: Expected-minimizer verification: three Monte Carlo objectives per cell, each averaging 32,768 batches. Relative prediction error and direction cosine refer to the raw displacement; JK/raw is the ratio of displacement norms. It is not a finite-data risk ratio.

Truth	ρ	b	Rel. error	Cosine	JK/raw
moderate	0	16	0.2821	0.9975	0.1919
moderate	0	32	0.1325	0.9993	0.1536
moderate	0	64	0.0659	0.9998	0.0999
moderate	0	128	0.0322	1.0000	0.0477
moderate	1	16	1.2249	0.0953	1.5354
moderate	1	32	0.7583	0.6535	1.8112
moderate	1	64	0.3316	0.9582	0.9427
moderate	1	128	0.2447	0.9999	0.6331
strong	0	16	1.1759	0.9595	0.2551
strong	0	32	0.4830	0.9844	0.1889
strong	0	64	0.2194	0.9954	0.1533
strong	0	128	0.1077	0.9985	0.1021
strong	1	16	1.5026	0.6741	0.6156
strong	1	32	1.2993	0.7588	0.6354
strong	1	64	0.9033	0.8530	0.6484
strong	1	128	0.5582	0.9285	0.5790

Table 24: Numerical acceptance checks across all 16 expected-minimizer conditions, covering both objectives and all three approximations. No fit reaches a boundary. Extrema summarize the original conditionwise checks.

Check	Worst case
Maximum gradient norm	1.76e-08
Minimum Hessian eigenvalue	1.26e-02
Maximum search-region disagreement	0.00e+00
Maximum initialization disagreement	3.13e-07
Maximum calibration residual	2.00e-11
Boundary fits	0

F Calibration and Statistical Details

F.1 Positive Finite-Temperature Calibration

Fixed-marginal calibration builds on minimum-discrimination-information estimation for contingency tables (Ireland & Kullback, 1968) and information projections (Csiszár, 1975). The calibrated coupling has the exponential form

$$[\mathcal{C}_{p,q}(S)]_{ij} = p_i q_j \exp\{(S_{ij} + a_i + b_j)/\tau\}, \tag{34}$$

where potentials a, b enforce the marginals. Adding row or column terms to S changes the objective only by a constant and leaves the coupling unchanged.

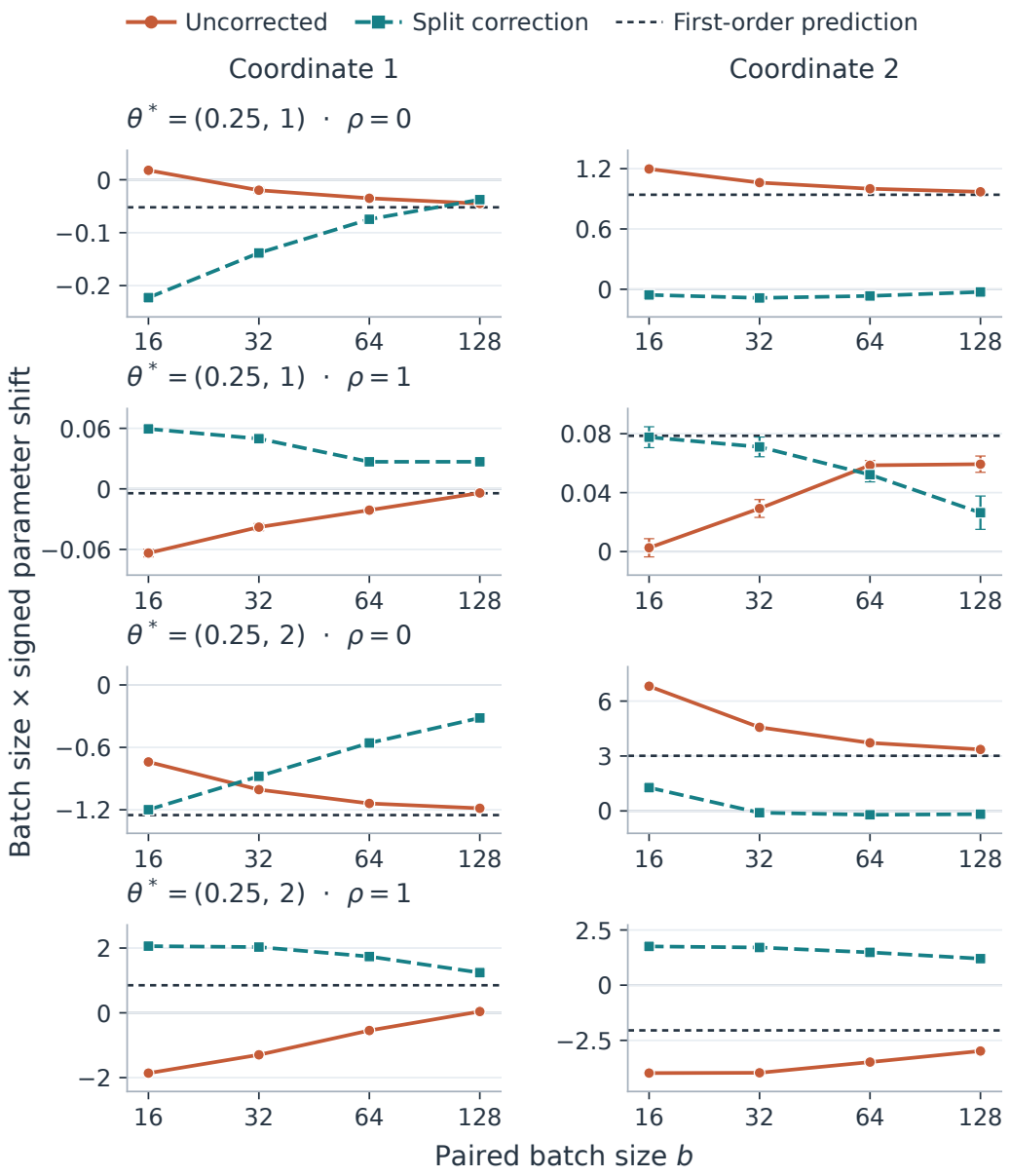

Figure 9: **From calibration bias to parameter displacement.** Panels show batch-scaled signed displacement of expected-objective minimizers for both truths, sampling ratios, and coordinates. Error bars are Monte Carlo standard errors across three independent objective approximations; dashed lines show first-order predictions for the uncorrected objective and gray lines mark zero.

For a frozen score matrix S, positive endpoint weights p, q, and temperature τ, implementation of Equation (1) alternates the log-domain dual updates, corresponding to alternating matrix scaling (Sinkhorn & Knopp, 1967; Cuturi, 2013),

$$a_i \leftarrow -\tau \log \sum_j q_j \exp\{(S_{ij} + b_j)/\tau\},$$

$$b_j \leftarrow -\tau \log \sum_i p_i \exp\{(S_{ij} + a_i)/\tau\}.$$

Under smooth positive problem data, derivatives of the Sinkhorn coupling converge with the iterations (Pauwels & Vaiter, 2023). These updates describe the calibration operator; iteration budgets and

Table 25: Full temperature confirmation (mean ± sample standard deviation). Native outputs are row-normalized scores; calibrated outputs match the empirical marginals.

Metric	Objective	Regime	τ	Native	Calibrated
KL	uncorrected	fixed	0.0800	68.4038 ± 24.9543	4.8310 ± 1.8850
KL	uncorrected	tuned	0.6343	8.3272 ± 3.0167	1.1723 ± 0.1656
KL	corrected	fixed	0.0800	62.2177 ± 22.7799	4.4174 ± 1.8753
KL	corrected	tuned	0.5856	8.2192 ± 2.9821	1.1764 ± 0.1667
MMD	uncorrected	fixed	0.0800	0.5979 ± 0.1559	0.3718 ± 0.2436
MMD	uncorrected	tuned	0.0294	0.6060 ± 0.1586	0.3832 ± 0.2547
MMD	corrected	fixed	0.0800	0.6457 ± 0.1718	0.3667 ± 0.2369
MMD	corrected	tuned	0.0066	0.6688 ± 0.1859	0.3825 ± 0.2463

Table 26: Primary tuned-corrected minus tuned-raw comparisons. Both prespecified superiority tests fail. Holm adjustment is across the two sign-flip tests; sign-flip inference assumes symmetric paired differences under the null.

Metric	Difference	95% interval	97.5% interval	Wins	Holm p
KL	0.0040	$[-0.0044, 0.0131]$	$[-0.0055, 0.0144]$	4/10	0.9492
MMD	-0.0007	$[-0.0168, 0.0117]$	$[-0.0195, 0.0130]$	4/10	0.9492

Table 27: Mean wall times in seconds. Evaluation-loop timing excludes preceding score, kernel, and oracle-plan computation.

Objective	Training	Evaluation loop
uncorrected	192.8878	3.1023
corrected	436.4467	7.5285

Table 28: Raw-score diagnostics and calibrated primary errors in the matched comparison. Raw-score values depend on the trained score gauge; calibrated values are the method outputs. Swiss mixed-support rows vary only the available endpoint-pool size.

Task	Training rule	Raw score	Calibrated
Swiss	Paired contrastive	$.2938 \pm .3035$	$.2306 \pm .1672$
Swiss	Likelihood on pairs	$.6967 \pm .1977$	$.1309 \pm .1036$
Swiss	Mixed support (128)	$.6646 \pm .2018$	$.3269 \pm .2609$
Swiss	Mixed support (512)	$.5274 \pm .2806$	$.3081 \pm .2514$
Swiss	Mixed support (1,024)	$.5651 \pm .2396$	$.2663 \pm .1953$
Colored MNIST	Paired contrastive	$.5715 \pm .6067$	$.0823 \pm .0355$
Colored MNIST	Likelihood on pairs	$1.5862 \pm .8524$	$.5868 \pm .4604$
Colored MNIST	Likelihood on mixed support	$1.2804 \pm .8443$	$.5068 \pm .1987$

Table 29: Paired differences for the Swiss-roll and Colored-MNIST objective ablations. Differences are left minus right on the primary error; negative values favor the left rule. Intervals are 95% intervals from 20,000 paired bootstrap draws over five confirmation replicates per task. Wins count lower-error replicates for the left rule.

Task	Comparison	Mean	Interval	Wins
Swiss	Paired likelihood − mixed (128)	-0.1960	$[-0.3616, -0.0338]$	$5/5$
Swiss	Paired likelihood − mixed (512)	-0.1772	$[-0.3240, -0.0304]$	$4/5$
Swiss	Paired likelihood − mixed (1,024)	-0.1353	$[-0.2470, -0.0248]$	$4/5$
CMNIST	Contrastive − paired likelihood	-0.5046	$[-0.8794, -0.1744]$	$5/5$
CMNIST	Contrastive − mixed likelihood	-0.4245	$[-0.5861, -0.2630]$	$5/5$

Table 30: Finance matched-architecture results (mean ± sample standard deviation across five independent confirmation blocks). Native denotes a parameterization-dependent diagnostic; fixed denotes common empirical marginals.

Objective	Output	Dependence	CVaR
Contrastive	native	0.06876 ± 0.00646	0.70827 ± 0.06396
Contrastive	fixed	0.08775 ± 0.00233	1.04560 ± 0.04105
Paired likelihood	native	0.07287 ± 0.00725	0.47364 ± 0.07510
Paired likelihood	fixed	0.08039 ± 0.00325	0.95549 ± 0.06276
Mixed likelihood	native	0.07045 ± 0.00958	0.58308 ± 0.16400
Mixed likelihood	fixed	0.08952 ± 0.00353	1.14191 ± 0.04943

Table 31: Both prespecified Finance matched-architecture contrasts, with 50,000 paired bootstrap draws. The 97.5% marginal intervals control the two primary contrasts; wins count lower-error blocks for contrastive learning.

Comparison	Difference	95% interval	97.5% interval	Wins
Contrastive − paired	0.007356	$[0.006297, 0.008236]$	$[0.006218, 0.008323]$	$0/5$
Contrastive − mixed	-0.001773	$[-0.004782, 0.001060]$	$[-0.004877, 0.001774]$	$4/5$

tolerances depend on the experiment (Section C). The original Swiss application uses 300 fixed iterations; Colored MNIST uses a residual-based stop; the finance application uses 40 iterations during inverse-EOT training and adaptive calibration at evaluation. The architecture-matched likelihood ablations use converged duals with implicit differentiation. On uniform endpoint banks, native diagnostics row-normalize $\exp(S/\tau)$ without the dual potentials. They depend on the trained score parameterization and are not invariant to column gauges. Native–calibrated comparisons measure the effect of endpoint matching on task performance, not intrinsic compatibility quality.

F.2 Queue's replicated-endpoint boundary map

Let $z_a^{(s)}$ be coordinate $a \in [K]$ of posterior-replay scenario $s \in [N]$, and let $w_{a,(1)} \le \cdots \le w_{a,(M)}$ be the sorted frozen endpoint pool. Queue has $N = MR$ tuple slots. After deterministically breaking ties in the native replay, let $r_a(s) \in [N]$ be the resulting ordered-slot label. Queue reports

$$\widetilde{z}_a^{(s)} = w_{a,(\lceil r_a(s)/R \rceil)}.$$

With $M = 128$, $R = 24$, and $N = 3{,}072$, each endpoint position appears in 24 slots, exactly matching the empirical marginal while retaining scenario-slot correspondence. Repeated endpoint values create ties and coarsen observable ranks: exact order preservation applies to the retained latent labels, not to the rank copula recomputed from numeric outputs. Distances, tail losses, and CVaR are recomputed after mapping.

This finite boundary coupling differs from the strictly positive solution of Equation (1); the cross-ratio identity in Theorem 3 does not justify Queue's map. Positive approximations are described in Section D.

F.3 Finance uncertainty under crossed seeds

The application crosses five independent training datasets with five independent market datasets. These 25 cells are not independent training repetitions. Methods are paired within each cell. Each of 50,000 bootstrap draws independently resamples the training and market indices with replacement, then averages the selected Cartesian product. This is the crossed-factor resampling construction of the pigeonhole and array bootstraps (Owen, 2007; Owen & Eckles, 2012).

The matched-architecture study instead resamples paired differences across five independent confirmation blocks. Each study prespecifies two primary dependence comparisons and reports 95% intervals and 97.5% marginal intervals for Bonferroni control. Secondary intervals do not control all reported endpoints jointly. Precision remains limited by the number of independent blocks.

Table 32: PACT minus each prespecified Finance comparator. Intervals use 50,000 two-way bootstrap draws. The 97.5% marginal intervals control the two primary dependence comparisons by Bonferroni; other metrics are secondary. Negative favors PACT except Top-1.

Comparator	Metric	Difference	95% interval	97.5% interval
Inverse EOT	Dependence (primary)	-0.0079	$[-0.0091, -0.0070]$	$[-0.0092, -0.0069]$
Inverse EOT	Joint	-0.0018	$[-0.0023, -0.0014]$	$[-0.0024, -0.0013]$
Inverse EOT	Cost dependence	-0.0176	$[-0.0198, -0.0160]$	$[-0.0202, -0.0158]$
Inverse EOT	Marginal energy	0.0000	$[0.0000, 0.0000]$	$[0.0000, 0.0000]$
Inverse EOT	CVaR	-0.0503	$[-0.0949, -0.0043]$	$[-0.1005, 0.0022]$
Inverse EOT	Regret	0.0030	$[-0.0224, 0.0299]$	$[-0.0273, 0.0350]$
Inverse EOT	Top-1	0.0008	$[-0.0275, 0.0267]$	$[-0.0333, 0.0300]$
Projected diffusion	Dependence (primary)	-0.0176	$[-0.0209, -0.0147]$	$[-0.0214, -0.0144]$
Projected diffusion	Joint	-0.0075	$[-0.0082, -0.0068]$	$[-0.0083, -0.0067]$
Projected diffusion	Cost dependence	-0.0028	$[-0.0058, -0.0001]$	$[-0.0065, 0.0003]$
Projected diffusion	Marginal energy	0.0000	$[-0.0000, 0.0000]$	$[-0.0000, 0.0000]$
Projected diffusion	CVaR	-0.0357	$[-0.0770, 0.0043]$	$[-0.0826, 0.0106]$
Projected diffusion	Regret	-0.2658	$[-0.3602, -0.1727]$	$[-0.3737, -0.1592]$
Projected diffusion	Top-1	0.2917	$[0.1975, 0.3825]$	$[0.1808, 0.3958]$

PACT improves both primary application dependence comparisons. Relative to inverse EOT, the CVaR 95% interval excludes zero but the 97.5% interval includes zero; regret and Top-1 are unresolved. Relative to projected diffusion, secondary regret and Top-1 comparisons favor PACT, while CVaR remains unresolved.

F.4 Compatibility implementations and structural relations

Aligned pairs. Given aligned pairs and endpoint negatives, a symmetric contrastive objective or a calibrated pair likelihood can learn whether two endpoints belong together. At population level, row-wise classification against negatives from q has optimum

$$S_{ij}/\tau = \log \frac{\pi^\star_{ij}}{p_i q_j} + c_i, \tag{35}$$

with an analogous column offset in the reverse direction. The score network can therefore follow the task: we use an MLP for Swiss roll, a circular Fourier model for Colored MNIST, and a trajectory-pair critic for finance.

The three empirical training objectives and their matched comparison protocol are specified in Appendix C.5.

Structural relation. If known transition laws share a disturbance U and an observed history H, posterior replay defines

$$\pi^{\mathrm{nat}}_{1:K}(dy_{1:K} \mid x, H) = \int \prod_{a=1}^{K} P_a(dy_a \mid x, u) P(du \mid x, H). \tag{36}$$

This supplies correspondence without observed pair labels. Queue's replicated-endpoint mapping and its distinction from positive-temperature calibration are defined in Appendix F.2.

F.5 Conditional prediction from calibrated scores

Prediction and paired evaluation. The calibrated joint distribution determines both conditional predictions and paired outcome distributions. For a source support point x_i with $\widehat{p}_i > 0$, conditional probabilities are given by $\widehat{\Pi}_{ij}/\widehat{p}_i$. The scaling form in Equation (34) also gives a conditional extension to a new source x against the target bank:

$$\widehat{P}(Y = y_j \mid X = x) = \frac{\widehat{q}_j \exp\{(s_{\widehat{\theta}}(x, y_j) + \widehat{b}_j)/\tau\}}{\sum_{k=1}^{n} \widehat{q}_k \exp\{(s_{\widehat{\theta}}(x, y_k) + \widehat{b}_k)/\tau\}}, \tag{37}$$

where $\widehat{b}$ is obtained from calibration on the endpoint supports. On the source support this agrees with the calibrated conditional law; for new sources it uses the learned score with the fitted target potential. The same coupling supports paired evaluation: drawing $(I, J) \sim \widehat{\Pi}$ and computing $h(x_I, y_J)$ yields the model's distribution of a paired quantity, such as a difference in outcomes or losses.

G Additional connections to prior work

Contrastive learning and inverse transport. Noise-contrastive estimation learns unnormalized models by distinguishing data from noise (Gutmann & Hyvärinen, 2010). Contrastive learning also admits an inverse-OT interpretation (Shi et al., 2023). This complements the inverse-transport and likelihood connections in Section 2; the population negative-sampling relation used here appears in Equation (35).

Conditional generators and structural models. Conditional OT, flow matching, and diffusion models provide flexible parameterizations of conditional or paired distributions (Bunne et al., 2022; Manupriya et al., 2024; Klein et al., 2024; Lipman et al., 2023; Ho et al., 2020). We compare PACT with paired density, flow, diffusion, inverse EOT, and factor models under task-matched information budgets. PACT does not prescribe one network family: it specifies the object learned from relational evidence and the operation that combines it with empirical marginals.

When known dynamics share a partly observed disturbance, posterior replay supplies structural information about correspondence. Common-random-number simulation uses the same principle to compare policies under a shared path (Schruben, 2011), and counter-based random number generators make such paths reusable (Salmon et al., 2011). Our queueing experiment uses this source of relational evidence and then matches the replayed scenarios to separately observed policy marginals.

Objective bias and estimator bias. Firth (1993) reduces first-order asymptotic bias of maximum-likelihood estimators in regular parametric models by adjusting the score. Our expansion tracks bias in the expected calibrated objective and the resulting local optimizer displacement. These are distinct from the bias and variance of an estimator fitted to a finite dataset; the risk experiments therefore evaluate both separately.

H Discussion and limitations

H.1 Turning marginal information into relation learning

PACT distinguishes marginal precision from relation-learning bias: more accurate marginal estimates do not alone ensure better correspondence learning. Additional observations matter through how they enter the objective, not only their number. Calibration therefore shapes both the output distribution and the statistical learning process.

H.2 Limitation & future work

Marginals alone do not identify correspondence; valid pairs or credible structural knowledge remain necessary, and a cross-policy coupling alone does not establish causal identification. The bias analysis assumes fixed finite supports and the stated regularity conditions; the spectral result additionally assumes correct specification. Controlled studies distinguish objective-bias cancellation, expected-minimizer displacement, and finite-data risk. Classical correction baselines and neural temperature comparisons do not establish a universal correction advantage. Finance inference is based on a limited number of independent blocks, and application inverse EOT retains the declared finite training budget.

Future work can extend the analysis to continuous spaces and neural scores, examine misspecification and approximate calibration, and incorporate variance, regularization, and computational cost into correction selection. A further direction is to design endpoint-data use around the sampling structure and task objective, so that additional marginal information yields verifiable improvements in relation estimation.