EXTENDING PATHWISE GRADIENTS TO DISCRETE RANDOM VARIABLES VIA FINITE-ORDER RELAXATION

Anonymous authors
Paper under double-blind review

ABSTRACT

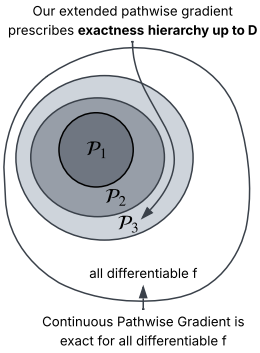
Pathwise gradients are preferred for continuous random variables because they are unbiased, low variance, and work with a single sample. For discrete variables, pathwise gradients cannot generally be exact for every differentiable function. We propose a general framework for constructing **Fini**te-order exact pathwise gradient estimators (Fini) for discrete variables, which are unbiased for all polynomials up to degree D. Among all admissible solutions, our estimator is the unique least-norm solution and has minimum weight variance. For common discrete distributions such as the Poisson, our construction yields native, closed-form estimators tailored to each distribution, implementable in a few lines of code. They preserve exact discrete forward samples without continuous relaxation, temperature tuning, or augmented representations. At matched polynomial exactness, they avoid the excess weight variance of non-native representations. We also bound bias for smooth downstream functions, showing how accuracy beyond polynomials depends on the first unmatched derivative and the sampling distribution. In experiments our low-order estimators match or outperform tuned baselines across linear, nonlinear and hierarchical latent-variable models, while out-speeding competitors in every runtime benchmark.

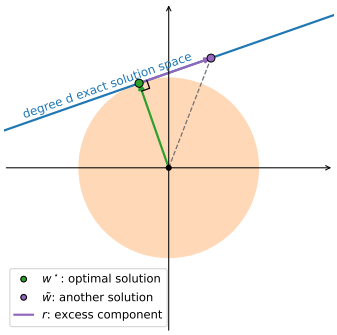
1 INTRODUCTION

For Monte Carlo (MC) gradient estimation in continuous random variables, the pathwise gradient estimator is the recommended default for its unbiasedness, typically low variance, and implementation simplicity (Mohamed et al., 2020). Pathwise gradients measure sensitivity among sample paths, combining how a realized sample changes with the distribution's parameters with how the downstream function changes with the sample. Thus it requires differentiability of both the sample path and the downstream function. However, discrete variables, commonly used in neural sequence models and deep generative models (Ranganath et al., 2015; Vafaii et al., 2024; Li et al., 2024), do not generally admit a pathwise estimator that is unbiased for all differentiable functions. This leaves a gap for MC gradient estimation in discrete distributions.

Existing gradient estimators for discrete variables navigate a bias–variance–computation tradeoff. Score-function estimators are unbiased but often have high variance; reducing it may require more samples or carefully designed control variates, adding computational or implementation cost (Ranganath et al., 2014; Mohamed et al., 2020). Measure-valued derivatives (MVD) and discrete General One-sample gradients can reduce variance through additional downstream evaluations, which become expensive in high-dimensional models (Pflug, 1989; Rosca et al., 2019; Cong et al., 2019). Biased alternatives instead approximate the backward pass, trading exactness for inexpensive gradients. Straight-through (ST) estimators preserve the hard sample but substitute a surrogate backward derivative, introducing bias (Bengio et al., 2013). For categorical variables, the Gumbel–Softmax trick (Jang et al., 2017; Maddison et al., 2017) instead replaces the hard sample with a differentiable continuous approximation controlled by temperature, allowing gradients through the relaxed sample. Lowering the temperature improves bias but can increase gradient variance, making it an additional hyperparameter to tune. Beyond categorical variables, researchers use the Gumbel–Softmax trick to approximate gradients for discrete variable such as Poisson (Li et al., 2024); however, these methods need to approximate one-dimensional variables with K-dimensional categorical variables, and inherit the temperature parameter, both adding computational overhead.

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.





Listing 1: Poisson Fini-2 and Fini-3 in torch.

```
...
rate = eta.exp()
z_hard = Poisson(rate).sample()

# order 2:
w = 0.5 * (z_hard + rate)
# order 3:
# w = (5*rate/12
#   + (8*rate+1)*z_hard/(12*rate)
#   - z_hard**2/(12*rate))

z = z_hard + w.detach() * (eta - eta.detach())
...
```

Figure 1: **Fini**te-order pathwise gradients. **Left**: pathwise gradients are unbiased for all differentiable functions - this is too rigid for discrete variables, so we relax the unbiasedness requirement to low-order polynomials; **Center**: generally, there are many solutions to the finite-order unbiasedness requirement; we use the least-norm principle to select the order-D unbiased solution, avoiding excess weight variance; **Right**: as an example, Fini-2 and 3 for Poisson reduce to a few lines of code.

Recently, Liu et al. (2023) interpret categorical ST estimators as a first-order numerical method and propose a second-order ST method, ReinMax. ReinMax is unbiased for degree-2 polynomials, whereas first-order ST is unbiased for linear functions. This improvement enables ReinMax to outmatch both biased and unbiased methods consistently on various benchmarks. Nonetheless, applying ReinMax to other discrete variables again requires approximation, which can lead to statistical and computational overheads. Can we adapt pathwise gradients to discrete variables beyond categoricals by relaxing the unbiasedness requirement in a principled way?

We introduce a general recipe, Fini, that applies to a range of discrete variables: As shown in Fig. 1, we require the pathwise estimator to be unbiased for low-degree polynomials rather than for every differentiable function; this weaker requirement admits many solutions, so we select the least-norm solution. This leads to several mathematical, statistical, and computational benefits, shown below.

Contributions:

- **Fini: Finite-order exact discrete pathwise gradients:** We introduce Fini, a general recipe that extends pathwise gradients to discrete random variables by requiring the gradient to be unbiased for polynomials up to degree D and selecting the least-norm solution. (§3).
- **Generality across discrete variables:** For common discrete variables, the estimator is solved once at design time. The estimator is simple to implement (see the Poisson example in Listing 1) and evaluation is fast and independent of support size (§3, Appendix E, §F).
- **Bias and variance analysis:** Our solution uniquely minimizes weight variance within the exactness class; we derive a non-asymptotic bias bound beyond it. We also characterize excess variance introduced by categorical and compound representations (§4).
- **Better empirical performance:** Across neural sequence models, VAEs, and hierarchical topic models, Fini matches or outperforms tuned specialized and generalist baselines (§6).

2 Background

2.1 Monte Carlo Gradient Estimation and Score Function Gradients

For a probability distribution $p_\theta(Z)$, we write the associated expectation $\mathbb{E}_{p_\theta}$ as $\mathbb{E}_\theta$ for simplicity. For an objective $\mathbb{E}_\theta[f(Z)]$, gradient-based learning requires calculating $\partial_\theta \mathbb{E}_\theta[f(Z)]$, often from MC samples. This problem emerges in supervised, unsupervised, and reinforcement learning.

One way to estimate $\partial_\theta \mathbb{E}_\theta[f(Z)]$ is through the score-function gradient, also known as REINFORCE (Williams, 1992; Ranganath et al., 2014). They use the identity $\partial_\theta \mathbb{E}_\theta[f(Z)] = \mathbb{E}_{p_\theta}[f(Z)s_\theta(Z)]$, with $s_\theta(Z) = \partial_\theta \log p_\theta(Z)$. Computing the integral analytically or numerically is generally intractable in high dimensions; MC gradient estimation instead draws N samples $z_n \sim p_\theta$ and forms unbiased

estimators of the quantity using $\sum_{n=1}^{N} f(z_n)s_\theta(z_n)/N$. The score-function estimator applies to both continuous and discrete variables and is unbiased under suitable regularity conditions. The assumptions are quite general, but the estimator can have high variance (Mohamed et al., 2020).

2.2 Pathwise Gradients

When f is differentiable, pathwise gradients instead use derivative information and integrate directly with ordinary backpropagation. For a continuous family, a pathwise derivative weight w_θ satisfies

$$\partial_\theta \mathbb{E}_\theta[f(Z)] = \mathbb{E}_\theta[w_\theta(Z)f'(Z)] \qquad \text{for every differentiable } f. \tag{1}$$

Explicit reparameterization writes out the sampling path $Z = g(\epsilon, \theta)$ for parameter-free noise ϵ and differentiates the path so $w_\theta = \partial_\theta g(\epsilon, \theta)$ (Kingma & Welling, 2014; Rezende et al., 2014). Consider a Gaussian $\mathcal{N}(\mu, \sigma)$, one valid path is $g(\epsilon, \theta = (\mu, \sigma)) = \mu + \sigma\epsilon$ where $\epsilon \sim \mathcal{N}(0, 1)$; then one can use $f'(z)\partial_\theta g(\epsilon, \theta)$ as a valid one-sample estimator for $\partial_\theta L$. For notation simplicity, we note w_θ as w going forward; it generally depends on the distribution parameters.

The pathwise estimator is the recommended default for continuous MC gradient estimation because it is unbiased, typically low variance, and works well with one sample. For common discrete variables, one can show no such w can satisfy Eq. (1) for all differentiable f. See proof in § A for details. This limitation motivates researchers to use estimators that introduce bias, high variance, or additional computations.

3 Fini: **Fini**te-Order Pathwise Gradients

To extend pathwise gradients to discrete distributions, we relax the unbiasedness requirement in Eq. (1). For a candidate pathwise weight w, instead of enforcing the unbiased pathwise identity for any differentiable f, we relax this requirement to unbiasedness with respect to polynomials up to degree D, denoted as $\mathcal{P}_D$. Since expectation is linear and the monomial basis spans the vector space of polynomials, unbiasedness over $\mathcal{P}_D$ breaks down to unbiasedness against $\{x, x^2, \cdots, x^D\}$. Thus, with the pathwise form $w_D(Z)f'(Z)$ and a prescribed order D, the relaxed requirement becomes

$$\mathbb{E}_\theta[w_D(Z)f'(Z)] = \partial_\theta \mathbb{E}_\theta[f(Z)] \quad \text{for every } f \in \{Z^d\}_{d=1}^{D} \tag{2}$$

Given we want to find a solution $w^\star$ in the functional space, we need to be able to describe the space first, so we equip the space with the inner-product $\langle u, v\rangle_\theta = \mathbb{E}_\theta[u(Z)v(Z)]$. We have now arrived at a Hilbert space $L^2(p_\theta)$ given the base law p_θ, and we see that for low orders, constraints in Eq. (2) generally leave w underdetermined. We choose the least-norm rule to resolve the ambiguity, which leads to clean derivations and computationally cheap solutions as shown later in this section, and other desirable statistical properties shown in § 4. Assuming suitable integrability and differentiability, the Hilbert space optimization problem can be written as follows:

$$\begin{aligned} &\operatorname*{arg\,min}_{w_D \in L^2(p_\theta)} \|w_D\|_2 \\ \text{s.t.} \qquad &\langle w_D, Z^{d-1}\rangle_\theta = \frac{1}{d}\partial_\theta \mathbb{E}_\theta[Z^d], \qquad d = 1, \ldots, D. \end{aligned} \tag{3}$$

The key to solving this problem is to invoke the Riesz representation theorem, and note that the monomial basis up to $D-1$ are precisely the Riesz representers to our problem. Therefore, the solution is a linear combination of monomials $v_D(Z) = [1, ..., Z^{D-1}]$, which means the resulting weight has degree at most $D-1$. Given the solution $w_D^\star$ must be a linear combination of monomials, each of the LHS in our D constraints becomes a gram matrix $G_D = \mathbb{E}_\theta[v_D(Z)v_D(Z)^\top]$ multiplied by the linear combination's coefficients. As long as the gram matrix is non-singular, we can safely invert it. Collecting the RHS of constraints in problem (3) as b_D, we see that the solution is:

$$\boxed{w_D^\star(z) = v_D(z)^\top G_D^{-1} b_D.} \tag{4}$$

We refer to solution $w_D^\star$ as the Fini-D estimator. This is the minimum-norm/least-norm solution to underspecified problems (Boyd & Vandenberghe, 2004; Trefethen & Bau, 1997). The recipe is

general: we compute moments up to $2D-2$ for the gram matrix, as well as the first D moments' derivatives for b_D, then solve the $D \times D$ linear system symbolically once to obtain the closed-form solution in Eq. (4). This requires finite moments through order $2D-2$, differentiability of the first D moments with respect to the parameter, and nonsingular G_D. The solutions are polynomials of degree at most $(D-1)$, evaluable in $O(D)$ time, independent of support size - another direct consequence of the least norm principle.

We show an example Poisson derivation below. Furthermore, we show for continuous variables at order 2, Fini recovers classic pathwise gradients. Going forward, we use exactness and unbiasedness interchangeably, and call an estimator order-D when it is unbiased for every polynomial of degree at most D.

Poisson example: Let $Z \sim \mathrm{Pois}(\exp(\eta))$. The order-3 Gram matrix and b_3 are

$$G_3 = \begin{pmatrix} 1 & \lambda & \lambda^2+\lambda \\ \lambda & \lambda^2+\lambda & \lambda^3+3\lambda^2+\lambda \\ \lambda^2+\lambda & \lambda^3+3\lambda^2+\lambda & \lambda^4+6\lambda^3+7\lambda^2+\lambda \end{pmatrix}, \qquad b_3 = \begin{pmatrix} \lambda \\ \lambda^2+\lambda/2 \\ \lambda^3+2\lambda^2+\lambda/3 \end{pmatrix}. \quad (5)$$

The order-2 system uses the leading 2×2 block of G_3 and the first two entries of b_3. Solving the two systems and substituting into $w_D^\star(z) = v_D(z)^\top G_D^{-1} b_D$ yields

$$\boxed{w_2^\star(Z) = \frac{Z+\lambda}{2}, \qquad w_3^\star(Z) = \frac{5\lambda}{12} + \frac{8\lambda+1}{12\lambda} Z - \frac{1}{12\lambda} Z^2.} \quad (6)$$

These are the Fini-2 and 3 weights used in Listing 1. The construction extends to any prescribed order; Appendix E gives the Fini-4 weight, whose variance under the log-rate parameterization diverges as $\lambda \downarrow 0$. This illustrates the variance cost of higher exactness.

Poisson MVD and quadrature. The same weights also follow from the difference identity

$$\partial_\eta \mathbb{E}[f(Z)] = \lambda \mathbb{E}[f(Z+1) - f(Z)] = \lambda \mathbb{E}\left[\int_0^1 f'(Z+t)\, dt\right].$$

Direct use of the difference identity requires an extra evaluation of f at $Z+1$. For order 2, trapezoidal quadrature approximates the increment by $(f'(Z) + f'(Z+1))/2$. The Poisson shift identity $\lambda \mathbb{E}[g(Z+1)] = \mathbb{E}[Z g(Z)]$ then replaces $f'(Z+1)$ by $(Z/\lambda) f'(Z)$ in expectation, recovering $w_2^\star$. The order-3 rule uses interpolation nodes at $t = 0, 1, 2$; Appendix C.2 gives the algebra.

Exact Pathwise Recoveries and ST recoveries: Fini-2 recovers fully exact pathwise gradients in the following cases: for scalar location–scale families, $w_2^\star(Z) = \partial_\theta Z$; for Gaussian random vectors, we recover the minimum-energy affine pathwise field. Both estimators are exact for every sufficiently regular f, although we impose exactness only for quadratics. Furthermore for categoricals, if we replace the min-norm rule with a symmetry principle, we recover ReinMax at temperature $= 1$, and Fini-2 corresponds to hard one-hot MVE in Hooper & Shekhovtsov (2026). See § C for details.

4 Theory

We establish uniqueness and weight-variance optimality, characterize two sources of excess variance, and bound bias outside the prescribed class. Proofs and details are in § D and § B. We start by showing our solutions are unique, and the solved weights are variance optimal.

Proposition 4.1. *Let $V_D = \mathrm{span}\{1, Z, \ldots, Z^{D-1}\}$ and assume G_D is nonsingular. Every degree-D exact weight $w \in L^2(p_\theta)$ satisfies*

$$w = w_D^\star + r, \qquad r \perp V_D, \qquad \mathrm{Var}_\theta[w] = \mathrm{Var}_\theta[w_D^\star] + \|r\|_2^2. \quad (7)$$

Thus $w_D^\star$ uniquely minimizes both norm and weight variance in the exactness class.

Prop. 4.1 states that any other admissible solution must have an excess null component, which increases weight variance; given ReinMax applied to Poisson with categorical representation is order 2, we can characterize its excess weight variance.

Corollary 4.2. *For $Z \sim \mathrm{Pois}(\lambda)$ and $\eta = \log \lambda$, the large-truncation limit of categorical ReinMax at $\tau = 1$ is degree-2 exact, with*

$$w_{\mathrm{RM}}(Z) = \underbrace{\frac{Z+\lambda}{2}}_{w_2^\star(Z)} + \underbrace{\frac{(Z-\lambda)^2 - (Z-\lambda) - \lambda}{2}}_{r_{\mathrm{RM}}(Z) \perp V_2}, \qquad \mathrm{Var}[w_{\mathrm{RM}}] - \mathrm{Var}[w_2^\star] = \frac{\lambda^2}{2}. \tag{8}$$

The gap in weight variance is due to the null component r_{RM}. Its expression and proof are in Prop. D.2, and a general theorem about Reinmax applied to scalar exponential families is in Prop. D.1, pointing out excess variance in combining ReinMax with categorical approximation.

Rao-Blackwellization (RB) is a variance reduction technique for compound representations. Proposition below shows compounding can lead to excess variance, and when RB can neutralize it.

Proposition 4.3. *Let $p_\theta(z, u)$ have marginal $p_\theta(z)$, and let $\widetilde{w}_D(U, Z) \in L^2$ be degree-D exact under the joint law. Then $\bar{w}_D(Z) = \mathbb{E}_\theta[\widetilde{w}_D \mid Z]$ is also degree-D exact, and*

$$\mathrm{Var}_\theta[\widetilde{w}_D] = \mathrm{Var}_\theta[w_D^\star] + \mathbb{E}_\theta[\mathrm{Var}_\theta(\widetilde{w}_D \mid Z)] + \|\bar{w}_D - w_D^\star\|_2^2. \tag{9}$$

RB removes auxiliary noise; least-norm optimization removes the remaining null component. Rao–Blackwellization suffices only when $\bar{w}_D = w_D^\star$, which we demonstrate below with the Gamma–Poisson compound representation of Negative Binomial variables.

Corollary 4.4. *Let $\Lambda \sim \mathrm{Gamma}(r, \mathrm{rate} = (1-p)/p)$, $Z \mid \Lambda \sim \mathrm{Pois}(\Lambda)$, and $\eta = \mathrm{logit}\, p$. For the compound order-D Poisson construction, conditioning its weight on Z gives*

$$\bar{w}_2 = w_2^\star, \qquad \bar{w}_3 = w_3^\star + r_3, \qquad r_3 \perp V_3, \quad r_3 \not\equiv 0, \tag{10}$$

where $w_D^\star$ is our Negative-Binomial weight.

Rao–Blackwellization closes the weight-variance gap at order 2 but not at order 3. The full statement and variance identities are in § B.1 and § B.2.

With variance characterized, we bound the bias for smooth functions beyond our prescribed class.

Proposition 4.5. *For a degree-D exact weight w, let $B_w(f) = \mathbb{E}_\theta[w(Z)f'(Z)] - \partial_\theta \mathbb{E}_\theta[f(Z)]$. Fix $a \in \mathbb{R}$ independently of θ and write $m_k(a) = \mathbb{E}_\theta[|Z-a|^k]$. Under the smoothness, moment, and score-identity conditions in § D.2,*

$$|B_w(f)| \le \|f^{(D+1)}\|_\infty \left[\frac{m_{2D}(a)^{1/2}}{D!} \|w\|_2 + \frac{m_{2D+2}(a)^{1/2}}{(D+1)!} \|s_\theta\|_2 \right]. \tag{11}$$

Exactness cancels the degree-D Taylor polynomial, leaving bias controlled by the first unmatched derivative. Least norm minimizes the bound's only weight-dependent term. The bound therefore predicts small bias when the downstream computation is locally well approximated by low-degree polynomials, a regime we expect to be common for sufficiently smooth neural objectives.

5 Related Work

Unbiased discrete gradient estimators. Score-function estimators use control variates, multi-sample baselines, and antithetic, coupled, or Stein identities to reduce variance (Mohamed et al., 2020; Mnih & Gregor, 2014; Tucker et al., 2017; Grathwohl et al., 2018; Mnih & Rezende, 2016; Yin & Zhou, 2019; Dong et al., 2020; 2021; Shi et al., 2022). MVD and GO gradients instead exploit distributional identities for lower variance, at the cost of additional f evaluations (Pflug, 1989; Rosca et al., 2019; Cong et al., 2019). For Poisson, §3 shows that quadrature of the MVD identity yields our weights. These methods retain arbitrary-function unbiasedness but require many downstream function evaluations, whereas Fini has low bias and works with a single downstream evaluation.

Gradient estimators for the categorical distribution Categorical gradient methods provide a common interface for discrete variables, or *categorical representation*, but their computation grows with the number of discrete states (Li et al., 2024). Unlike categorical representation, Fini applies

directly to the target law, avoiding excess compute overheads. Gumbel–Softmax uses temperature-controlled continuous relaxations (Jang et al., 2017; Maddison et al., 2017), whereas ST estimators preserve the hard sample and substitute its backward derivative (Bengio et al., 2013).

ReinMax extends the ST estimator (Liu et al., 2023). It shows ST is a linear approximation, analogous to Euler's method in numerical quadrature. Given Euler's method's first-order bias can be improved to second-order, ST can also be improved similarly. ReinMax thus replaces the linear ST by a quadratic one, and demonstrates consistently excellent empirical performance across benchmarks (Liu et al., 2023), making it our principal generalist comparator. Hooper & Shekhovtsov (2026) extend this line of work by axiomatizing a broader ST family and deriving a second-order minimum-variance estimator (MVE). Their estimator has a temperature knob that interpolates between ST and their second-order estimator. Although theoretically MVE has lower weight variance than ReinMax, MVE does not consistently outperform ReinMax empirically. Our order-2 solutions recover their one-hot and 1D estimators for categorical variables.

Other distribution-specific estimators and compound estimators. Although most discrete MC gradient estimation research focuses on the categorical distribution, many tailored estimators have been developed for distributions such as the Poisson. Tailored estimators exploit distribution-specific structure and can be highly effective. For Poisson, EAT-cubic differentiates a relaxed inter-arrival construction with cubic interpolation and is our principal specialist comparator (Knuth, 1997; Vafaii et al., 2024; Ibrahim et al., 2026). Compound representations transfer an estimator through auxiliary variables; Gamma–Poisson, for example, reuses a Poisson estimator for a Negative-Binomial variable. Proposition 4.3 and Corollary 4.4 quantify the resulting variance overhead. Our construction instead applies the general recipe directly to the target variable.

6 Numerical Experiments

We test four empirical claims: (i) Fini is unbiased for functions inside its prescribed exactness class; (ii) For Poisson, low-order Fini matches or outperforms competitors on nonlinear and hierarchical scenarios beyond their prescribed exactness class; (iii) Fini performs well for other common discrete variables such as Negative Binomial and Binomial; and (iv) across GPUs and benchmarks, low-order Fini consistently outspeeds all competitors.

Our main Poisson-specific comparator is EAT-cubic, which differentiates a relaxed Poisson inter-arrival construction using cubic interpolation (Ibrahim et al., 2026). Our general comparator is ReinMax, an ST estimator for categorical variables (Liu et al., 2023); we apply it to count distributions (Poisson, Binomial, and Negative-Binomial) through categorical approximation where feasible and through the Gamma–Poisson representation for Negative-Binomial variables. For unbounded categorical approximations, we choose the truncation so that the omitted tail has probability ≤ 0.001; the appendix gives more detail. Both comparators have a tuned temperature τ. We also include EAT-sigmoid, Gumbel–Softmax, and score-function estimators where they are part of the original benchmark comparison. EAT-sigmoid and Gumbel–Softmax require a temperature.

We follow the metrics in prior works for the corresponding benchmarks. Higher ELBO and lower perplexity are better. Parenthesized numbers beside temperature-based methods give the selected τ; Fini and the score-function estimator do not use a temperature.

6.1 Linear Poisson VAE: Testing Order-2 Exactness

In a Poisson VAE, an encoder specifies a Poisson distribution over latent counts Z given an observation, and a decoder maps those counts back to an observation distribution. We use the linear-decoder van Hateren benchmark of Ibrahim et al. (2026), which has an analytic gradient for comparison. With a Gaussian observation likelihood and a linear decoder, the benchmark's downstream objective is quadratic in the latent counts, so Fini-2 is exact here. We tune EAT-cubic over its reported temperature grid. In Figure 2, the observed mean gradient errors at both checkpoints are near zero, while Fini-2 has about one-fifth the gradient-error standard deviation of tuned EAT-cubic. Its validation-ELBO trajectory also lies closer to the analytic-gradient reference during training.

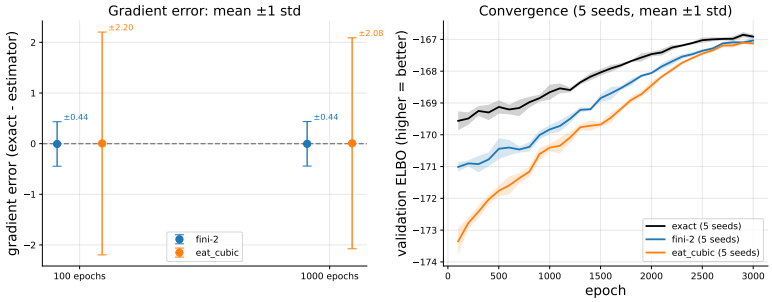


Figure 2: Linear Poisson VAE with 512 latent dimensions. **Left**: mean single-sample gradient error relative to the analytic gradient early (epoch 100) and partway through training (epoch 1,000), shown as dots with standard-deviation bars. Both means are near zero; Fini-2 has about one-fifth the standard deviation of tuned EAT-cubic. **Right**: validation ELBO over training, mean ± standard deviation across five matched seeds. Fini-2 follows the analytic-gradient trajectory more closely.

6.2 Performance on Realistic Benchmarks

Proposition 4.5 bounds bias outside the prescribed class, but can low-order Fini perform well in realistic applications, where functions can be nonlinear, and have multiple stochastic layers?

Neural sequence model. The partially observed generalized linear model (POGLM) is an autoregressive latent variable model for inferring connectivities from partially observed spike trains. As a dynamic graphical model, it has both observed and latent Poisson variables whose rates depend on recent count histories (Li et al., 2024). We reproduce the synthetic benchmark in Ibrahim et al. (2026), following the same temperature grids and tuning protocol across 30 matched independent seeds. Table 1 compares validation ELBOs for Fini 2 and 3, EAT-cubic, Gumbel–Softmax, EAT-sigmoid, and the score-function estimator on the synthetic POGLM benchmark. The temperature-free Fini 2 and 3 have mean ELBOs close to tuned EAT-cubic, while the other baselines are worse.

Table 1: Synthetic POGLM validation ELBO (higher is better), mean ± std. over 30 seeds. Parentheses give the selected temperature τ. Boldface marks the three highest means, which differ by at most 0.02.

Dataset	Fini-3	Fini-2	EAT-cubic (0.1)	GSM (0.5)	EAT-sigmoid (0.5)	Score
Synthetic	$\mathbf{-330.61 \pm 0.79}$	$\mathbf{-330.60 \pm 0.75}$	$\mathbf{-330.62 \pm 0.79}$	-331.94 ± 0.95	-332.18 ± 0.97	-333.23 ± 1.51

Nonlinear Poisson VAEs. Following the discrete-VAE setup of Liu et al. (2023), we replace Bernoulli latents with Poisson latents on MNIST, Fashion-MNIST, and Omniglot. We use dynamically binarized images with Bernoulli likelihoods and whitened continuous images with Gaussian likelihoods (Appendix G.3). Tab. 2 compares Fini-2 and Fini-3 with EAT-cubic and ReinMax. Fini-3 leads in four of six settings, including all Bernoulli-likelihood cases, with Fini-2 close behind. Both orders occupy the leading tier across all six settings, but their ranking varies under Gaussian likelihoods: higher exactness does not uniformly improve optimization.

Hierarchical topic models. We next test our estimators on Deep Exponential Families (DEF), a deep generative model used for topic modeling and recommender systems (Ranganath et al., 2015). A DEF generates observed data through multiple stochastic layers, where each layer draws from an exponential-family distribution whose natural parameters come from the previous layer. We train amortized two-layer sparse Poisson DEFs on 20 Newsgroups (20 NG) and RCV1 using the inference architecture of Mnih & Gregor (2014); Appendix G.5 gives the full setup.

Tab. 3 shows the results. Fini 2 and 3 have the best train and test ELBOs on both datasets. ReinMax ranks first on heldout perplexity for the smaller 20NG, and Fini-3 leads on RCV1. To diagnose, we split the ELBO into log-likelihood and the KL terms, and use the low variance MVD as the ground truth. ReinMax at $\tau = 1.3$ tilts toward worse KL and better log likelihood, which strongly correlates with perplexity; Fini closely tracks MVD on both datasets, suggesting ReinMax's 20NG win may reflect bias rather than better gradients. See § G.6 for details.

Table 2: Nonlinear Poisson-VAE training ELBO ↑, mean ± standard deviation over five matched seeds. Columns pair each dataset with a Bernoulli or Gaussian observation likelihood. First is bold, second is underlined.

	Bernoulli Likelihoods			Gaussian Likelihoods		
	MNIST	Fashion-MNIST	Omniglot	MNIST	Fashion-MNIST	Omniglot
Fini-3	$\mathbf{-118.60 \pm 0.26}$	$\mathbf{-251.98 \pm 0.40}$	$\mathbf{-129.38 \pm 0.07}$	$\underline{595.48 \pm 3.72}$	$\mathbf{133.74 \pm 2.80}$	$\underline{403.14 \pm 2.46}$
Fini-2	$\underline{-118.88 \pm 0.19}$	$\underline{-252.26 \pm 0.18}$	$\underline{-129.93 \pm 0.12}$	594.47 ± 3.92	132.09 ± 3.29	$\mathbf{403.80 \pm 2.00}$
EAT-cubic	-120.06 ± 0.17 (0.5)	-253.04 ± 0.22 (0.5)	-130.76 ± 0.07 (0.2)	$\mathbf{595.66 \pm 3.60}$ (0.5)	$\underline{132.42 \pm 1.70}$ (0.5)	400.84 ± 2.05 (0.5)
ReinMax	-125.87 ± 0.38 (1.1)	-260.56 ± 0.57 (1.1)	-140.07 ± 0.40 (1.0)	589.21 ± 2.78 (1.0)	104.06 ± 2.51 (1.1)	396.63 ± 1.53 (1.0)

Table 3: Poisson DEF results on train and test ELBO (↑) and heldout perplexity (PPL, ↓), mean ± standard deviation over five matched seeds. Best is bold; second is underlined.

		20 Newsgroups			RCV1		
Model	Training method	Train ELBO ↑	Test ELBO ↑	PPL ↓	Train ELBO ↑	Test ELBO ↑	PPL ↓
Poisson DEF	Fini-3	$\mathbf{-179.30 \pm 2.81}$	$\underline{-199.31 \pm 2.04}$	660.7 ± 7.6	$\underline{-269.27 \pm 3.16}$	$\underline{-267.64 \pm 3.73}$	$\mathbf{379.3 \pm 5.4}$
	Fini-2	$\mathbf{-179.30 \pm 2.62}$	$\mathbf{-198.98 \pm 1.92}$	$\underline{659.7 \pm 6.7}$	$\mathbf{-269.20 \pm 3.05}$	$\mathbf{-267.60 \pm 3.73}$	$\underline{381.9 \pm 4.6}$
	EAT-cubic (0.2)	$\underline{-180.94 \pm 2.59}$	-199.77 ± 2.17	664.6 ± 8.5	-271.92 ± 3.18	-270.16 ± 3.82	388.4 ± 4.1
	ReinMax (1.3)	-191.25 ± 2.88	-210.01 ± 2.37	$\mathbf{638.1 \pm 10.5}$	-301.86 ± 2.83	-300.05 ± 3.27	384.4 ± 4.8

Summary. Fini-2 and Fini-3 match tuned EAT-cubic on POGLM, improve most of its ELBOs, and outperform ReinMax on every displayed ELBO. The preferred order of Fini remains task dependent.

6.3 Generality Across Common Discrete Variables

We evaluate Poisson, Binomial, and Negative-Binomial VAEs on MNIST with latent dimensions 2, 16, and 128. ReinMax uses categorical representation for Binomial and Poisson (with truncation). Overdispersion in Negative-Binomial makes direct truncation impractical, so we use the Gamma–Poisson representation. Results are in Tab. 4. Fini-3 leads at higher latent dimensions, while Fini-2 leads in two of the three 2-dimensional settings - exactness order remains problem-dependent.

Table 4: Count-VAE training ELBO, mean ± standard deviation over five independent matched seeds. Best is bold; second is underlined.

Family	Latent dim.	Fini-2	Fini-3	ReinMax
binomial $n = 8, p = 0.5$	2	$\mathbf{-174.51 \pm 0.45}$	$\underline{-176.91 \pm 1.07}$	-179.67 ± 0.67 (1.0)
	16	$\underline{-123.38 \pm 0.14}$	$\mathbf{-123.11 \pm 0.10}$	-125.55 ± 0.10 (1.0)
	128	$\underline{-118.73 \pm 0.22}$	$\mathbf{-118.43 \pm 0.17}$	-121.45 ± 0.15 (1.1)
negbin $r = 8, p = 0.5$	2	$\underline{-177.34 \pm 0.15}$	$\mathbf{-177.31 \pm 0.24}$	-182.58 ± 0.48 (1.0)
	16	$\underline{-129.29 \pm 0.16}$	$\mathbf{-129.25 \pm 0.20}$	-136.58 ± 0.71 (1.0)
	128	$\underline{-122.58 \pm 0.06}$	$\mathbf{-122.54 \pm 0.36}$	-124.41 ± 0.24 (1.1)
poisson $\lambda_0 = 0.1$	2	$\mathbf{-190.37 \pm 0.66}$	$\underline{-191.77 \pm 1.57}$	-199.33 ± 2.57 (1.0)
	16	$\underline{-146.74 \pm 0.62}$	$\mathbf{-145.68 \pm 0.31}$	-170.60 ± 0.39 (1.0)
	128	$\underline{-123.20 \pm 0.30}$	$\mathbf{-122.80 \pm 0.40}$	-129.99 ± 1.02 (1.0)

Compound representation adds statistical overhead. Negative-Binomial latents can reuse a Poisson estimator through an auxiliary rate Λ. Corollary 4.4 predicts that conditioning recovers Fini-2 exactly, while at order 3 a nonzero Z-only component can remain.

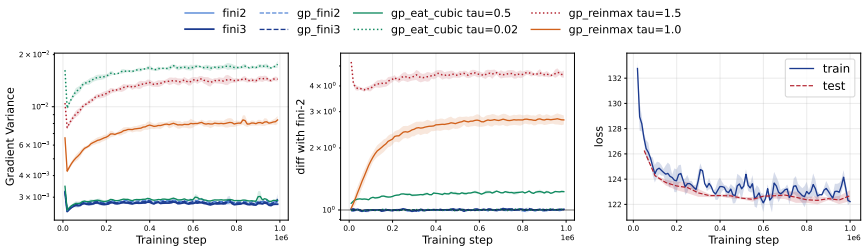


Figure 3: Negative-Binomial representation diagnostic. Native Fini-2 and Fini-3 have lower gradient variance than matched Gamma–Poisson counterparts; ReinMax has higher variance.

We reproduce the variance study of Dong et al. (2020) with Negative-Binomial latents. The 128-dimensional variance results are in Fig. 3. In this benchmark, each native estimator has lower gradient variance than its matched Gamma–Poisson counterpart, while ReinMax has higher variance.

Categorical representation adds computational overhead. For Binomial latents, the released hard 1D MVE implementation of Hooper & Shekhovtsov (2026) operates over $K = n + 1$ categorical states, whereas our fixed-order weights' costs are independent of the support size. Figure 4 shows the resulting support-size scaling. The compound and categorical results therefore test different costs of non-native representations: statistical overhead from auxiliary variables and computational overhead from categorical embedding.

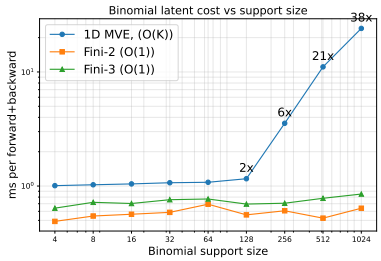


Figure 4: Categorical support scaling. At 1024, the released 1D MVE implementation takes 38× as long as Fini-2.

6.4 Speed Comparisons

Fini 2 and 3 require one forward–backward pass and evaluate affine or quadratic polynomials. Across three benchmarks on L4 and A100 GPUs, Fini-2 and 3 are first and second in every workload–hardware pair (Table 5); Appendix G.7 gives the five-repeat post-warm-up protocol. Fini-2 is consistently faster because its weight is affine rather than quadratic.

Table 5: Wall-clock(s) per epoch (↓), mean ± standard deviation over five repeats.

	Linear VAE		Nonlinear VAE		DEF	
Estimator	L4	A100	L4	A100	L4	A100
Fini-3	0.326 ± 0.010	0.361 ± 0.008	1.868 ± 0.074	2.016 ± 0.083	0.928 ± 0.021	0.907 ± 0.016
Fini-2	$\mathbf{0.309 \pm 0.005}$	$\mathbf{0.338 \pm 0.004}$	$\mathbf{1.768 \pm 0.017}$	$\mathbf{1.848 \pm 0.046}$	$\mathbf{0.863 \pm 0.014}$	$\mathbf{0.835 \pm 0.016}$
EAT-cubic	1.427 ± 0.005	0.581 ± 0.010	2.680 ± 0.040	2.942 ± 0.073	1.546 ± 0.039	1.565 ± 0.015
ReinMax	1.144 ± 0.006	0.568 ± 0.009	2.692 ± 0.035	2.859 ± 0.104	1.451 ± 0.024	1.610 ± 0.134

7 Discussion

Limitations: Fini requires relevant moments and their parameter derivatives, and order-D exactness requires at least D support points. Outside the prescribed class, Proposition 4.5 requires smoothness and finite moments; polynomial unbiasedness alone does not guarantee small bias. Finally, higher D adds an orthogonal correction and cannot decrease the raw second moment; the fourth-order Poisson example in Appendix E.1 shows that its variance can worsen at low rates. Thus, estimator order controls a bias–variance tradeoff rather than providing a uniform improvement.

Conclusion: We propose Fini, a recipe for extending pathwise gradients to discrete distributions by prescribing degree-D polynomial exactness and then choosing the least-norm solution. For common discrete variables, the resulting polynomial weights preserve the exact forward sample, require one forward–backward pass, and use neither temperature tuning nor auxiliary randomness. Moreover, this recipe recovers classical pathwise gradients for continuous random variables. Our theory extends the guarantees beyond the prescribed exactness class through a non-asymptotic bias bound for smooth functions and decomposes excess weight variance into auxiliary-variable noise and orthogonal null components. Empirically, Fini-2 and 3 remain in the leading performance tier across increasingly realistic benchmarks, while ranking first and second in every runtime benchmark. Thus, we provide a general derivation principle for extending pathwise gradient estimators beyond continuous random variables, yielding low-bias, low-variance estimators with simple, efficient implementations.

AI Statement:

We used LLMs to periodically and systematically search for related work. We used LLM to improve paper organization and grammar, but not to generate full sections from scratch. We also used LLM to

reorganize human-written proofs and math prose into more concise statements. All LLM assisted text and codes have been verified by us, and we take full responsibility for the end results.

ETHICS STATEMENT:

This work develops a general-purpose gradient estimation method and does not involve human subjects or sensitive data. Its reduced computational overhead may lower barriers to experimenting with discrete latent-variable models, though we do not directly measure monetary or environmental costs.

REPRODUCIBILITY STATEMENT:

We show worked-out estimator formulas for all variables in Appendix E, a torch implementation in Listing 1 and multi-parameter implementation in Appendix F. We provide all experiment details in Appendix G. We plan to release the code after the review process.

REFERENCES

Yoshua Bengio, Nicholas Léonard, and Aaron Courville. Estimating or propagating gradients through stochastic neurons for conditional computation, 2013.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

Yulai Cong, Miaoyun Zhao, Ke Bai, and Lawrence Carin. GO gradient for expectation-based objectives. In *International Conference on Learning Representations*, 2019.

Zhe Dong, Andriy Mnih, and George Tucker. DisARM: An antithetic gradient estimator for binary latent variables. In *Advances in Neural Information Processing Systems*, volume 33, pp. 18637–18647, 2020.

Zhe Dong, Andriy Mnih, and George Tucker. Coupled gradient estimators for discrete latent variables. In *Advances in Neural Information Processing Systems*, volume 34, pp. 24498–24508, 2021.

Will Grathwohl, Dami Choi, Yuhuai Wu, Geoffrey Roeder, and David Duvenaud. Backpropagation through the void: Optimizing control variates for black-box gradient estimation. In *International Conference on Learning Representations*, 2018.

James Hooper and Alexander Shekhovtsov. Generalized and optimal straight-through estimators. In *Proceedings of the 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, pp. 3187–3195. PMLR, 2026. AISTATS 2026 Spotlight.

Michael Ibrahim, Hanqi Zhao, Eli Sennesh, Zhi Li, Anqi Wu, Jacob L. Yates, Chengrui Li, and Hadi Vafaii. A hitchhiker's guide to poisson gradient estimation. In *Proceedings of the 43rd International Conference on Machine Learning*, volume 306 of *Proceedings of Machine Learning Research*. PMLR, 2026.

Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with gumbel-softmax. In *International Conference on Learning Representations*, 2017.

Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. In *International Conference on Learning Representations*, 2014.

Donald E. Knuth. *The Art of Computer Programming, Volume 2: Seminumerical Algorithms*. Addison-Wesley, 3 edition, 1997.

Chengrui Li, Weihan Li, Yule Wang, and Anqi Wu. A differentiable partially observable generalized linear model with forward-backward message passing. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 28095–28111. PMLR, 2024.

Liyuan Liu, Chengyu Dong, Xiaodong Liu, Bin Yu, and Jianfeng Gao. Bridging discrete and backpropagation: Straight-through and beyond. In *Advances in Neural Information Processing Systems*, volume 36, pp. 12291–12311, 2023.

Chris J. Maddison, Andriy Mnih, and Yee Whye Teh. The concrete distribution: A continuous relaxation of discrete random variables. In *International Conference on Learning Representations*, 2017.

Andriy Mnih and Karol Gregor. Neural variational inference and learning in belief networks. In *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *Proceedings of Machine Learning Research*, pp. 1791–1799. PMLR, 2014.

Andriy Mnih and Danilo Rezende. Variational inference for monte carlo objectives. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 2188–2196. PMLR, 2016.

Shakir Mohamed, Mihaela Rosca, Michael Figurnov, and Andriy Mnih. Monte carlo gradient estimation in machine learning. *Journal of Machine Learning Research*, 21(132):1–62, 2020.

Georg Ch. Pflug. Sampling derivatives of probabilities. *Computing*, 42(4):315–328, 1989. doi: 10.1007/BF02243227.

Rajesh Ranganath, Sean Gerrish, and David M. Blei. Black box variational inference. In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics*, volume 33 of *Proceedings of Machine Learning Research*, pp. 814–822. PMLR, 2014.

Rajesh Ranganath, Linpeng Tang, Laurent Charlin, and David M. Blei. Deep exponential families. In *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, volume 38 of *Proceedings of Machine Learning Research*, pp. 762–771. PMLR, 2015.

Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and approximate inference in deep generative models. In *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *Proceedings of Machine Learning Research*, pp. 1278–1286. PMLR, 2014.

Mihaela Rosca, Michael Figurnov, Shakir Mohamed, and Andriy Mnih. Measure-valued derivatives for approximate bayesian inference. In *Fourth Workshop on Bayesian Deep Learning*, 2019.

Jiaxin Shi, Yuhao Zhou, Jessica Hwang, Michalis K. Titsias, and Lester W. Mackey. Gradient estimation with discrete stein operators. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Michalis K. Titsias and Jiaxin Shi. Double control variates for gradient estimation in discrete latent variable models. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pp. 6134–6151. PMLR, 2022.

Lloyd N. Trefethen and David Bau. *Numerical Linear Algebra*. SIAM, Philadelphia, PA, 1997. ISBN 978-0-89871-361-9.

George Tucker, Andriy Mnih, Chris J. Maddison, Dieterich Lawson, and Jascha Sohl-Dickstein. REBAR: Low-variance, unbiased gradient estimates for discrete latent variable models. In *Advances in Neural Information Processing Systems*, 2017.

Hadi Vafaii, Dekel Galor, and Jacob L. Yates. Poisson variational autoencoder. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Ronald J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8(3–4):229–256, 1992. doi: 10.1007/BF00992696.

Mingzhang Yin and Mingyuan Zhou. ARM: Augment-REINFORCE-merge gradient for stochastic binary networks. In *International Conference on Learning Representations*, 2019.

Mingzhang Yin, Yuguang Yue, and Mingyuan Zhou. ARSM: Augment-REINFORCE-swap-merge estimator for gradient backpropagation through categorical variables. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 7095–7104. PMLR, 2019.

Extending Pathwise Gradients to Discrete Random Variables via Finite-Order Relaxation

Appendix

A Discrete all-function pathwise obstruction **14**

B Solving the Optimization Problem **14**
B.1 Augmented representations and Rao–Blackwellization 14
B.2 Gamma–Poisson specialization 15

C Connection with other methods **15**
C.1 Recovering pathwise derivatives 15
C.2 Connection to measure-valued derivatives and quadrature 16
C.3 Connection to ReinMax and generalized straight-through methods 16

D Theory section Proofs **19**
D.1 Proof of Proposition 4.1 19
D.2 Proof of Proposition 4.5 19
D.3 Order-2 ReinMax in a scalar natural exponential family 19
D.4 Poisson ReinMax under increasing categorical truncation 20

E Additional closed-form estimators **21**
E.1 Poisson: order 4 21
E.2 Binomial: $\eta = \operatorname{logit} p$ 21
E.3 Negative binomial: $\eta = \operatorname{logit} p$ 21
E.4 Negative binomial: $\rho = \log r$ 22
E.5 Beta-binomial: $\eta = \operatorname{logit} \pi$ 22
E.6 Beta-binomial: $\rho = \log \kappa$ 23

F Two-Parameter Straight-Through Implementation **23**

G Additional Experimental Details **24**
G.1 Linear Poisson VAE 24
G.2 Neural sequence model 24
G.3 Nonlinear Poisson and count VAEs 24
G.4 Binomial support-size timing 24
G.5 DEF experiment details 24
G.6 ReinMax and the rate–distortion tradeoff 25
G.7 Timing protocol 25

A Discrete all-function pathwise obstruction

Proposition A.1 (No universal exact sample pathwise weight on isolated support)**.** *Let p_θ be a differentiable family on a fixed discrete support $\mathcal{S} \subset \mathbb{R}$. Suppose that, at the parameter value under consideration, some $z_0 \in \mathcal{S}$ is isolated and $\partial_\theta p_\theta(z_0) \neq 0$. Then there is no Z-measurable weight $w_\theta(Z)$ satisfying*

$$\mathbb{E}_\theta[w_\theta(Z) f'(Z)] = \partial_\theta \mathbb{E}_\theta[f(Z)]$$

for every smooth f for which the expectation exists.

Proof. Because z_0 is isolated, choose a smooth bump function f supported in an interval containing no other point of $\mathcal{S}$, with $f(z_0) = 1$ and $f'(z_0) = 0$. Then $f'(Z) = 0$ almost surely, and hence $\mathbb{E}_\theta[w_\theta(Z) f'(Z)] = 0$ for every weight w_θ. On the other hand, $\mathbb{E}_\theta[f(Z)] = p_\theta(z_0)$, so

$$\partial_\theta \mathbb{E}_\theta[f(Z)] = \partial_\theta p_\theta(z_0) \neq 0,$$

a contradiction. □

B Solving the Optimization Problem

Proposition B.1 (Least-norm solution)**.** *Let*

$$v_D(Z) = (1, Z, \ldots, Z^{D-1})^\top, \qquad G_D = \mathbb{E}_\theta[v_D(Z) v_D(Z)^\top],$$

and let b_D denote the degree-D exactness constraints, so that every feasible weight satisfies $\mathbb{E}_\theta[w(Z) v_D(Z)] = b_D$. If G_D is nonsingular, the unique least-norm degree-D exact weight in $L^2(p_\theta)$ is

$$w_D^\star(z) = v_D(z)^\top G_D^{-1} b_D. \tag{12}$$

Proof. Let $\mathcal{V}_D = \operatorname{span}\{1, Z, \ldots, Z^{D-1}\}$. Since every exactness constraint is an inner product with an element of $\mathcal{V}_D$, orthogonal projection onto $\mathcal{V}_D$ preserves feasibility and cannot increase the $L^2(p_\theta)$ norm. Hence the minimum-norm weight has the form $w_D^\star(z) = v_D(z)^\top c$. Its exactness constraints reduce to

$$G_D c = b_D, \qquad \text{hence} \qquad c = G_D^{-1} b_D,$$

which proves (12); nonsingularity of G_D gives uniqueness. □

B.1 Augmented representations and Rao–Blackwellization

Suppose $p_\theta(z, u)$ is an augmented representation of the marginal law $p_\theta(z)$, and let $\widetilde{w}_D(U, Z) \in L^2$ be any degree-D exact compound weight:

$$\mathbb{E}_\theta[\widetilde{w}_D(U, Z) f'(Z)] = \partial_\theta \mathbb{E}_\theta[f(Z)], \qquad f \in \mathcal{P}_D. \tag{13}$$

Let $\bar{w}_D(Z) = \mathbb{E}_\theta[\widetilde{w}_D(U, Z) \mid Z]$ denote its Rao–Blackwellization.

Proposition B.2 (Rao–Blackwellization and native projection)**.** *Let $w_D^\star$ be the native minimum-$L^2(p_\theta)$ degree-D exact weight. Then $\bar{w}_D$ is degree-D exact and*

$$\begin{aligned} \|\widetilde{w}_D\|_2^2 &= \|w_D^\star\|_2^2 + \mathbb{E}_\theta[\operatorname{Var}_\theta(\widetilde{w}_D \mid Z)] + \|\bar{w}_D - w_D^\star\|_2^2, \\ \operatorname{Var}_\theta[\widetilde{w}_D] &= \operatorname{Var}_\theta[w_D^\star] + \mathbb{E}_\theta[\operatorname{Var}_\theta(\widetilde{w}_D \mid Z)] + \|\bar{w}_D - w_D^\star\|_2^2. \end{aligned} \tag{14}$$

Hence the native construction has no greater weight variance than the compound estimator or its Rao–Blackwellization. If $\bar{w}_D \in \mathcal{V}_D$, then $\bar{w}_D = w_D^\star$, so the native estimator is exactly the Rao–Blackwellized compound estimator.

Proof. By the tower property,

$$\mathbb{E}_\theta[\bar{w}_D(Z) f'(Z)] = \mathbb{E}_\theta[\widetilde{w}_D(U, Z) f'(Z)],$$

so $\bar{w}_D$ preserves degree-D exactness. Moreover, $\widetilde{w}_D - \bar{w}_D$ is orthogonal to every square-integrable function of Z, while exactness of $\bar{w}_D$ and $w_D^\star$ gives $\bar{w}_D - w_D^\star \perp \mathcal{V}_D$. Since $w_D^\star \in \mathcal{V}_D$, these three components are mutually orthogonal:

$$\widetilde{w}_D = w_D^\star + (\bar{w}_D - w_D^\star) + (\widetilde{w}_D - \bar{w}_D). \tag{15}$$

Orthogonality gives the first identity in (14), with

$$\|\widetilde{w}_D - \bar{w}_D\|_2^2 = \mathbb{E}_\theta[\mathrm{Var}_\theta(\widetilde{w}_D \mid Z)].$$

Degree-D exactness fixes the mean of all three weights through the linear test function $f(z) = z$, giving the variance identity. □

B.2 Gamma–Poisson specialization

For

$$\Lambda \sim \mathrm{Gamma}\left(r, \mathrm{rate} = \frac{1-p}{p}\right), \qquad Z \mid \Lambda \sim \mathrm{Pois}(\Lambda), \qquad \eta = \mathrm{logit}\, p,$$

write $\Lambda = e^\eta G$ with $G \sim \mathrm{Gamma}(r, \mathrm{rate} = 1)$. Since $\partial_\eta \log \Lambda = 1$, applying the order-$D$ Poisson log-rate weight conditionally on Λ gives a degree-D exact compound weight.

At order 2, $\widetilde{w}_2(Z, \Lambda) = (Z + \Lambda)/2$. Conjugacy gives

$$\Lambda \mid Z = z \sim \mathrm{Gamma}\left(r + z, \mathrm{rate} = \frac{1}{p}\right),$$

and therefore

$$\begin{aligned} \mathbb{E}[\widetilde{w}_2(Z, \Lambda) \mid Z = z] &= \frac{z + p(r+z)}{2} \\ &= \frac{rp}{2} + \frac{1+p}{2} z = w_2^\star(z). \end{aligned} \tag{16}$$

Thus the native order-2 Negative-Binomial weight is exactly the Rao–Blackwellization of its Gamma–Poisson counterpart.

At order 3, the conditional Poisson weight can be written as

$$\widetilde{w}_3(Z, \Lambda) = \frac{5\Lambda}{12} + \frac{2Z}{3} - \frac{Z(Z-1)}{12\Lambda}. \tag{17}$$

Under the posterior above, $\mathbb{E}[\Lambda \mid Z = z] = p(r + z)$ and, for $z \geq 2$, $\mathbb{E}[\Lambda^{-1} \mid Z = z] = 1/[p(r + z - 1)]$. Because $z(z-1) = 0$ for $z \in \{0, 1\}$, conditioning (17) gives, for all $z \in \mathbb{N}_0$,

$$\bar{w}_3(z) = \frac{5p(r+z)}{12} + \frac{2z}{3} - \frac{z(z-1)}{12p(r+z-1)}, \tag{18}$$

where the final term is set to zero for $z < 2$. Both $\bar{w}_3$ and the native $w_3^\star$ are degree-3 exact, so

$$r_3 := \bar{w}_3 - w_3^\star \perp \mathrm{span}\{1, Z, Z^2\}.$$

Moreover, (18) is $O(z)$ as $z \to \infty$, whereas the native formula in Appendix E.3 satisfies

$$w_3^\star(z) = -\frac{(1-p)^3}{12p(r+1)} z^2 + O(z).$$

Hence $r_3 \not\equiv 0$ for $r > 0$ and $0 < p < 1$. Proposition 4.3 therefore separates the compound order-3 variance gap into auxiliary conditional variance and the additional native null-component norm $\|r_3\|_2^2$.

C Connection with other methods

C.1 Recovering pathwise derivatives

Proof of Scalar location–scale recovery. Because $w_2^\star \in V_2 = \mathrm{span}\{1, Z\}$, write $w_2^\star(z) = a + bz$. Linear and quadratic exactness impose

$$a + b\mu_\theta = \dot{\mu}, \qquad a\mu_\theta + b(\mu_\theta^2 + \sigma_\theta^2) = \mu_\theta \dot{\mu} + \sigma_\theta \dot{\sigma}. \tag{19}$$

Thus $b = \dot{\sigma}/\sigma_\theta$ and $a = \dot{\mu} - \mu_\theta \dot{\sigma}/\sigma_\theta$, so

$$w_2^\star(z) = \dot{\mu} + \frac{\dot{\sigma}}{\sigma_\theta}(z - \mu_\theta). \tag{20}$$

Since $Z = \mu_\theta + \sigma_\theta \varepsilon$, $\partial_\theta Z = \dot{\mu} + \dot{\sigma}\varepsilon = w_2^\star(Z)$, which gives exactness for every sufficiently regular f. □

For a pure location family, $\dot{\sigma} = 0$, so degree-1 exactness already suffices.

Proof of Gaussian-vector recovery. The gradients of quadratic polynomials span

$$\mathcal{V}_2 = \{a + A(z - \mu_\theta) : a \in \mathbb{R}^K,\ A = A^\top\}.$$

Hence the least-norm solution is $W_2^\star(z) = a + A(z - \mu_\theta)$ with $A = A^\top$. Linear and quadratic exactness yield

$$a = \dot{\mu}, \qquad A\Sigma_\theta + \Sigma_\theta A = \dot{\Sigma}. \tag{21}$$

Because $\Sigma_\theta \succ 0$, the Sylvester equation has a unique symmetric solution A, giving the weight results. Gaussian integration by parts then gives

$$\mathbb{E}_\theta[W_2^\star(Z)^\top \nabla f(Z)] = \dot{\mu}^\top \mathbb{E}_\theta[\nabla f(Z)] + \frac{1}{2}\operatorname{tr}\Big(\dot{\Sigma}\,\mathbb{E}_\theta[\nabla^2 f(Z)]\Big) = \partial_\theta \mathbb{E}_\theta[f(Z)], \tag{22}$$

for every sufficiently regular f. □

C.2 Connection to measure-valued derivatives and quadrature

For $Z \sim \mathrm{Pois}(\lambda)$ with $\eta = \log\lambda$,

$$\partial_\eta \mathbb{E}[f(Z)] = \lambda \mathbb{E}[f(Z+1) - f(Z)] = \lambda \mathbb{E}\left[\int_0^1 f'(Z+t)\,dt\right].$$

Replacing the integral with polynomially exact quadrature and returning the shifted terms to the original Poisson law yields the same weights as the moment construction.

For degree-2 exactness, f' is linear, so the two-node rule at $t = 0, 1$ is exact:

$$\begin{aligned}\partial_\eta \mathbb{E}[f(Z)] &= \frac{\lambda}{2}\mathbb{E}[f'(Z) + f'(Z+1)] \\ &= \mathbb{E}\left[\frac{\lambda + Z}{2} f'(Z)\right],\end{aligned} \tag{23}$$

where $\lambda\mathbb{E}[g(Z+1)] = \mathbb{E}[Zg(Z)]$. Thus quadrature recovers $w_2^\star(Z) = (\lambda + Z)/2$.

For degree-3 exactness, interpolate $f'(Z+t)$ at $t = 0, 1, 2$. Exact integration of quadratics on $[0, 1]$ gives weights $5/12$, $2/3$, and $-1/12$:

$$\begin{aligned}\partial_\eta \mathbb{E}[f(Z)] &= \lambda \mathbb{E}\left[\frac{5}{12} f'(Z) + \frac{2}{3} f'(Z+1) - \frac{1}{12} f'(Z+2)\right] \\ &= \mathbb{E}\left[\left(\frac{5\lambda}{12} + \frac{2}{3}Z - \frac{Z(Z-1)}{12\lambda}\right) f'(Z)\right],\end{aligned} \tag{24}$$

using $\lambda^2\mathbb{E}[g(Z+2)] = \mathbb{E}[Z(Z-1)g(Z)]$. Hence

$$w_3^\star(Z) = \frac{5\lambda}{12} + \frac{8\lambda + 1}{12\lambda} Z - \frac{1}{12\lambda} Z^2,$$

again matching the least-norm moment solution.

The identity above is the measure-valued-derivative representation of the Poisson gradient: a finite difference between neighboring laws. The same argument applies to other distributions with finite-difference derivative identities; the Binomial quadratic weight, for example, follows from the analogous two-node construction.

C.3 Connection to ReinMax and generalized straight-through methods

Lemma C.1 (Categorical quadratic exactness)**.** *Let $Z \sim \mathrm{Cat}(p)$ be represented by $Z \in \{e_1, \dots, e_K\}$, where $p = \mathrm{softmax}(\eta)$ has $p_i > 0$ for all i, and let $J = \mathrm{diag}(p) - pp^\top$. A same-path estimator*

$$\widehat{g}(Z) = W(Z)\nabla f(Z), \qquad W_i := W(e_i) \in \mathbb{R}^{K\times K},$$

is exact for every polynomial of degree at most two if and only if

$$W_i e_i = \frac{1}{2}(e_i - p), \qquad p_i W_i e_j + p_j W_j e_i = 0 \quad (i \neq j), \qquad \sum_i p_i W_i = J. \tag{25}$$

Proof. Write $f(z) = a^\top z + \frac{1}{2} z^\top H z$ with $H = H^\top$. Quadratic exactness is equivalent to

$$\sum_i p_i W_i (a + H e_i) = J\left(a + \frac{1}{2}\operatorname{diag}(H)\right) \qquad \text{for all } a \text{ and } H = H^\top. \tag{26}$$

The coefficient of a gives the final constraint in (25). Taking $H = e_i e_i^\top$ and $H = e_i e_j^\top + e_j e_i^\top$ gives the first two constraints. Conversely, these matrices span the symmetric matrices, so (25) implies (26). □

Proposition C.2 (Symmetry characterizes ReinMax)**.** *Let W be quadratic-exact under the assumptions of Lemma C.1. Then*

$$W_i = W_i^\top \quad \textit{for every } i$$

if and only if

$$W_i = \frac{1}{2}\left[J + (e_i - p)(e_i - p)^\top\right], \qquad i = 1, \dots, K. \tag{27}$$

Thus sample-wise symmetry uniquely selects the ReinMax weight at $\tau = 1$ within the quadratic-exact class.

Proof. Write $d_i = e_i - p$ and denote the right-hand side of (27) by W_i^{RM}. Since $J e_i = p_i d_i$ and $\sum_i p_i d_i d_i^\top = J$, its quadratic-exactness follows from Lemma C.1:

$$\begin{aligned} W_i^{\mathrm{RM}} e_i &= \frac{1}{2}\left[p_i d_i + (1 - p_i) d_i\right] = \frac{1}{2} d_i, \\ W_i^{\mathrm{RM}} e_j &= \frac{p_j}{2}(e_j - e_i), \qquad i \neq j, \\ \sum_i p_i W_i^{\mathrm{RM}} &= \frac{1}{2}\left[J + \sum_i p_i d_i d_i^\top\right] = J. \end{aligned} \tag{28}$$

The matrices W_i^{RM} are symmetric, so it remains to show uniqueness under symmetry.

Let $R_i := W_i - W_i^{\mathrm{RM}}$ and define the linear matrix-valued map

$$\mathcal{R}(u) := \sum_i p_i u_i R_i, \qquad u \in \mathbb{R}^K.$$

Both W and W^{RM} are quadratic-exact. Applying their difference to $f_u(z) = \frac{1}{2}(u^\top z)^2$ gives $\mathcal{R}(u)u = 0$ for every u. Polarization therefore gives

$$\mathcal{R}(u)v = -\mathcal{R}(v)u \qquad \text{for all } u, v \in \mathbb{R}^K. \tag{29}$$

Since $R_i = R_i^\top$, each $\mathcal{R}(u)$ is symmetric. Hence, for arbitrary $r, u, v \in \mathbb{R}^K$,

$$\begin{aligned} r^\top \mathcal{R}(u) v &= v^\top \mathcal{R}(u) r = -v^\top \mathcal{R}(r) u = -u^\top \mathcal{R}(r) v \\ &= u^\top \mathcal{R}(v) r = r^\top \mathcal{R}(v) u = -r^\top \mathcal{R}(u) v. \end{aligned} \tag{30}$$

Thus $\mathcal{R}(u)v = 0$ for all u, v. Taking $u = e_i$ gives $p_i R_i v = 0$ for every v, hence $R_i = 0$ because $p_i > 0$. Therefore $W_i = W_i^{\mathrm{RM}}$ for every i. □

Proposition C.3 (Least-norm categorical quadratic estimator)**.** *Let $Z \sim \mathrm{Cat}(p)$ be represented by $Z \in \{e_1, \dots, e_K\}$, where $p = \mathrm{softmax}(\eta)$ has $p_i > 0$ for all i, and let $J = \operatorname{diag}(p) - pp^\top$. Consider quadratic-exact same-path estimators*

$$\widehat{g}(Z) = W(Z)\nabla f(Z), \qquad W_i := W(e_i) \in \mathbb{R}^{K \times K}.$$

Among all such weights, there is a unique minimizer of $\mathbb{E}\|W(Z)\|_F^2$.

For $i \neq j$, define $\omega_{ij} = p_i p_j/(p_i + p_j)$, and let $G \in \mathbb{R}^{K \times K}$ have entries

$$G_{ij} = -\omega_{ij} \quad (i \neq j), \qquad G_{ii} = \sum_{j \neq i} \omega_{ij}.$$

If $C \in \mathbb{R}^{K\times K}$ satisfies

$$GC = \frac{1}{2}J,$$

then the minimum-norm weight is

$$W_i^\star e_j = \begin{cases} \frac{1}{2}(e_i - p), & j = i, \\ \frac{p_j}{p_i + p_j} C^\top (e_j - e_i), & j \neq i. \end{cases}$$

In particular, one may take $C = \frac{1}{2}G^\dagger J$.

Proof. By Lemma C.1, quadratic exactness imposes (25). The first relation fixes the diagonal columns, while the second identifies the two off-diagonal columns associated with each pair $i < j$. We therefore use $(W_i e_j)_{i<j}$ as coordinates for the remaining degrees of freedom and equip them with the inner product

$$\langle W, V\rangle_{\mathcal{H}} := \sum_{i<j} \frac{p_i(p_i + p_j)}{p_j} (W_i e_j)^\top (V_i e_j).$$

Indeed, using $p_j W_j e_i = -p_i W_i e_j$, the original objective reduces to

$$\mathbb{E}\|W(Z)\|_F^2 = \text{const} + \sum_{i<j} \frac{p_i(p_i + p_j)}{p_j} \|W_i e_j\|^2 = \text{const} + \|W\|_{\mathcal{H}}^2. \tag{31}$$

The remaining exactness constraints define a linear map $\mathcal{A} : \mathcal{H} \to \mathbb{R}^{K\times K}$ by

$$e_j^\top \mathcal{A} W := \left(\sum_{i<j} p_i W_i e_j - \sum_{i>j} p_j W_j e_i \right)^\top. \tag{32}$$

After subtracting the fixed diagonal contribution in (25), quadratic exactness is $\mathcal{A}W = \frac{1}{2}J$.

For $C \in \mathbb{R}^{K\times K}$, the adjoint is determined by

$$\begin{aligned} \langle \mathcal{A}W, C\rangle_F &= \sum_{i<j} p_i (W_i e_j)^\top C^\top (e_j - e_i), \\ (\mathcal{A}^* C)_i e_j &= \frac{p_j}{p_i + p_j} C^\top (e_j - e_i), \qquad i < j. \end{aligned} \tag{33}$$

Consequently, the normal operator is

$$e_j^\top \mathcal{A}\mathcal{A}^* C = \sum_{i\neq j} \omega_{ij} (e_j - e_i)^\top C = e_j^\top GC, \qquad \mathcal{A}\mathcal{A}^* C = GC. \tag{34}$$

The minimum-norm solution lies in $\text{range}(\mathcal{A}^*) = \ker(\mathcal{A})^\perp$. Hence $W^\star = \mathcal{A}^* C$, where the coefficient matrix satisfies the normal system

$$GC = \frac{1}{2}J.$$

Equation (33) gives the off-diagonal columns displayed in Proposition C.3, and the first constraint in (25) gives the diagonal columns. Since $G\mathbf{1} = 0$ and $J\mathbf{1} = 0$, one may take $C = \frac{1}{2}G^\dagger J$; adding $\mathbf{1}a^\top$ to C leaves $C^\top(e_j - e_i)$ unchanged. Uniqueness follows from uniqueness of the minimum-norm element of the feasible affine subspace. □

Corollary C.4 (Relation to Hooper and Shekhovtsov)**.** *The least-norm quadratic-exact estimator in Proposition C.3 coincides with the $\tau \to \infty$ categorical minimum-variance estimator of Hooper & Shekhovtsov (2026).*

Relation to the 1D MVE of Hooper and Shekhovtsov. Let Z have a finite scalar support, with $m_1 = \mathbb{E}_\theta[Z]$, $m_2 = \mathbb{E}_\theta[Z^2]$, and $v = m_2 - m_1^2 > 0$. Then the degree-2 least-norm weight coincides with the hard quadratically-unbiased 1D MVE of Hooper & Shekhovtsov (2026).

The degree-2 least-norm construction gives $w_2^\star(z) = a + bz$, with

$$a + bm_1 = \dot{m}_1, \qquad am_1 + bm_2 = \frac{1}{2}\dot{m}_2.$$

Hence

$$w_2^\star(z) = \frac{m_2 - zm_1}{v}\,\dot{m}_1 + \frac{z - m_1}{2v}\,\dot{m}_2,$$

which corresponds to Hooper & Shekhovtsov (2026, Prop. C.3). The correspondence is specific to order 2. For a scalar finite-support law, the construction extends to degree-D exactness whenever the support contains at least D distinct points.

D Theory section Proofs

Throughout this appendix, a dot denotes differentiation with respect to the parameter coordinate under consideration.

D.1 Proof of Proposition 4.1

Proof. Let $v_D(Z) = (1, Z, \dots, Z^{D-1})^\top$ and $r = w - w_D^\star$. Since w and $w_D^\star$ satisfy the same exactness constraints, $\mathbb{E}_\theta[r(Z)v_D(Z)] = 0$, hence $r \perp V_D$. Since $w_D^\star \in V_D$,

$$\begin{aligned} \mathbb{E}_\theta[w(Z)^2] &= \mathbb{E}_\theta[w_D^\star(Z)^2] + \mathbb{E}_\theta[r(Z)^2] = b_D^\top G_D^{-1} b_D + \mathbb{E}_\theta[r(Z)^2], \\ \mathrm{Var}_\theta[w(Z)] &= \mathrm{Var}_\theta[w_D^\star(Z)] + \mathbb{E}_\theta[r(Z)^2]. \end{aligned} \tag{35}$$

The second identity follows because degree-D exactness fixes $\mathbb{E}_\theta[w(Z)]$. Equality in either objective requires $r = 0$ almost surely, proving uniqueness. □

D.2 Proof of Proposition 4.5

Assume $f \in C^{D+1}$ with bounded $f^{(D+1)}$, $m_{2D+2}(a) < \infty$, $w, s_\theta \in L^2(p_\theta)$, and that the score identity holds for f and degree-D polynomials. Let $m_k(a) = \mathbb{E}_\theta[|Z - a|^k]$. The constants in Proposition 4.5 are

$$C_{D,w}(a) = \frac{m_{2D}(a)^{1/2}}{D!}, \qquad C_{D,s}(a) = \frac{m_{2D+2}(a)^{1/2}}{(D+1)!}. \tag{36}$$

Proof. Fix an expansion point a independent of θ, and let T_D be the degree-D Taylor polynomial of f about a. Degree-D exactness annihilates T_D, so the score identity reduces the bias to

$$B_w(f) = \mathbb{E}_\theta\big[w(Z)\big(f'(Z) - T_D'(Z)\big) - s_\theta(Z)\big(f(Z) - T_D(Z)\big)\big]. \tag{37}$$

Taylor's theorem gives $|f'(z) - T_D'(z)| \le \|f^{(D+1)}\|_\infty |z-a|^D/D!$ and $|f(z) - T_D(z)| \le \|f^{(D+1)}\|_\infty |z-a|^{D+1}/(D+1)!$. Cauchy–Schwarz therefore yields

$$|B_w(f)| \le \|f^{(D+1)}\|_\infty \left[\frac{m_{2D}(a)^{1/2}}{D!}\|w\|_2 + \frac{m_{2D+2}(a)^{1/2}}{(D+1)!}\|s_\theta\|_2\right]. \tag{38}$$

□

D.3 Order-2 ReinMax in a scalar natural exponential family

The finite-support result isolates the component introduced by ReinMax before specializing to Poisson.

Proposition D.1 (ReinMax excess component). *Let $p_\eta(z) = h(z)\exp\{\eta z - A(\eta)\}$ be a scalar natural exponential family represented categorically over a finite support. Write $Y = Z - \mu$, $V = \mathbb{E}_\eta[Y^2]$, $M_3 = \mathbb{E}_\eta[Y^3]$, and $M_4 = \mathbb{E}_\eta[Y^4]$. At $\tau = 1$,*

$$w_{\mathrm{RM}}(Z) = \frac{1}{2}(V + Y^2) = \underbrace{V + \frac{M_3}{2V}Y}_{w_2^\star(Z)} + \underbrace{\frac{1}{2}\left(Y^2 - \frac{M_3}{V}Y - V\right)}_{r_{\mathrm{RM}}(Z)}, \tag{39}$$
$$r_{\mathrm{RM}} \perp \mathrm{span}\{1, Z\}, \qquad \mathrm{Var}_\eta[w_{\mathrm{RM}}] - \mathrm{Var}_\eta[w_2^\star] = \frac{1}{4}\left(M_4 - V^2 - \frac{M_3^2}{V}\right).$$

The variance gap is strictly positive whenever the law has at least three support points of positive probability.

Proof. Let $C \sim \mathrm{Cat}(p)$, let x contain the support values, and write $Z = x^\top C$. With $J(p) = \mathrm{diag}(p) - pp^\top$, the $\tau = 1$ ReinMax matrix weight reduces to

$$A_C = \frac{1}{2}\left[J(p) + (C - p)(C - p)^\top\right], \qquad x^\top A_C x = \frac{1}{2}(V + Y^2). \tag{40}$$

For a natural exponential family, $\dot\mu = V$ and $\dot V = M_3$, so the native order-2 weight is $w_2^\star = V + (M_3/(2V))Y$. The residual in (39) satisfies $\mathbb{E}_\eta[r_{\mathrm{RM}}] = \mathbb{E}_\eta[Y r_{\mathrm{RM}}] = 0$, and Proposition 4.1 gives

$$\mathrm{Var}_\eta[w_{\mathrm{RM}}] - \mathrm{Var}_\eta[w_2^\star] = \|r_{\mathrm{RM}}\|_2^2 = \frac{1}{4}\left(M_4 - V^2 - \frac{M_3^2}{V}\right). \tag{41}$$

The gap can vanish only if $Y^2 - (M_3/V)Y - V = 0$ almost surely; a nonzero quadratic cannot vanish on three support points. □

D.4 Poisson ReinMax under increasing categorical truncation

The Poisson result follows by letting the categorical truncation grow to infinite support.

Proposition D.2 (Poisson categorical limit). *Let Z_K have the Poisson law truncated to $\{0, \dots, K\}$ and renormalized, with mean μ_K and variance V_K. Under categorical embedding with logits $\ell_k = k\eta - \log(k!)$, $\eta = \log\lambda$, the $\tau = 1$ ReinMax weight is*

$$w_{\mathrm{RM},K}(Z_K) = \frac{1}{2}\left[V_K + (Z_K - \mu_K)^2\right].$$

Extend $w_{\mathrm{RM},K}(z) = \frac{1}{2}[V_K + (z - \mu_K)^2]$ to $z \in \mathbb{N}_0$. Then $w_{\mathrm{RM},K} \to w_{\mathrm{RM}}$ in $L^2(p_\lambda)$, where $w_{\mathrm{RM}}(z) = \frac{1}{2}[\lambda + (z - \lambda)^2]$. The limiting weight satisfies

$$w_{\mathrm{RM}}(Z) = \underbrace{\frac{Z + \lambda}{2}}_{w_2^\star(Z)} + \underbrace{\frac{1}{2}\left[(Z - \lambda)^2 - (Z - \lambda) - \lambda\right]}_{r_{\mathrm{RM}}(Z)}, \tag{42}$$
$$r_{\mathrm{RM}} \perp \mathrm{span}\{1, Z\}, \qquad \|r_{\mathrm{RM}}\|_2^2 = \frac{\lambda^2}{2}.$$

Proof. Proposition D.1 applied to the truncated Poisson law gives $w_{\mathrm{RM},K} = \frac{1}{2}[V_K + (Z_K - \mu_K)^2]$. The truncated moments give $\mu_K \to \lambda$ and $V_K \to \lambda$. Hence the coefficients of the quadratic weight functions converge; since the Poisson law has a finite fourth moment, the convergence holds in $L^2(p_\lambda)$.

For $Y = Z - \lambda$, the limiting decomposition in (42) has $\mathbb{E}[Y] = 0$, $\mathbb{E}[Y^2] = \mathbb{E}[Y^3] = \lambda$, and $\mathbb{E}[Y^4] = \lambda + 3\lambda^2$. Therefore

$$\mathbb{E}[r_{\mathrm{RM}}] = \mathbb{E}[Z r_{\mathrm{RM}}] = 0, \qquad \|r_{\mathrm{RM}}\|_2^2 = \frac{1}{4}\mathbb{E}[(Y^2 - Y - \lambda)^2] = \frac{\lambda^2}{2}. \tag{43}$$

□

E ADDITIONAL CLOSED-FORM ESTIMATORS

For a scalar discrete variable X and parameter coordinate θ, write the order-m same-path estimator as

$$\widehat{\partial_\theta \mathbb{E}[f(X)]} = v_\theta^{(m)}(X)\, f'(X). \tag{44}$$

Here order m means exactness for every polynomial f of degree at most m. The formulas below are the minimum-$L^2(p_\theta)$ solutions in the stated parameter coordinates: order 2 is affine, $v^{(2)}(x) = a_0 + a_1 x$, and order 3 is quadratic, $v^{(3)}(x) = a_0 + a_1 x + a_2 x^2$.

E.1 POISSON: ORDER 4

For $Z \sim \mathrm{Pois}(\lambda)$ with $\eta = \log\lambda$, order-4 exactness gives

$$w_4^\star(z) = v_4(z)^\top G_4^{-1} b_4, \qquad v_4(z) = (1, z, z^2, z^3)^\top.$$

Using moments through order 6 and $(b_4)_d = d^{-1}\partial_\eta \mathbb{E}[Z^d]$, the 4×4 system gives

$$G_4^{-1} b_4 = \begin{pmatrix} 3\lambda/8 \\ (19\lambda^2 + 5\lambda + 2)/(24\lambda^2) \\ -(5\lambda+3)/(24\lambda^2) \\ 1/(24\lambda^2) \end{pmatrix}.$$

The resulting weight and variance are

$$\boxed{w_4^\star(Z) = \frac{3\lambda}{8} + \frac{19\lambda^2 + 5\lambda + 2}{24\lambda^2} Z - \frac{5\lambda+3}{24\lambda^2} Z^2 + \frac{1}{24\lambda^2} Z^3 \qquad \mathrm{Var}(w_4^\star) = \frac{\lambda}{4} + \frac{1}{72} + \frac{1}{96\lambda}}$$

E.2 BINOMIAL: $\eta = \mathrm{logit}\, p$

Let

$$X \sim \mathrm{Binomial}(n, p), \qquad q = 1 - p, \qquad \eta = \mathrm{logit}\, p. \tag{45}$$

The order-2 estimator is

$$\boxed{v_\eta^{(2)}(x) = \frac{np}{2} + \frac{1-2p}{2}\, x.} \tag{46}$$

For $n \geq 2$, the order-3 estimator is

$$\boxed{\begin{aligned} v_\eta^{(3)}(x) = {} & \frac{np(5-6p)}{12q} \\ & + \left[\frac{1-2p}{2} + \frac{2np+1-2p}{12(n-1)pq}\right] x \\ & - \frac{1}{12(n-1)pq}\, x^2. \end{aligned}} \tag{47}$$

The corresponding raw-p weight follows from parameterization equivariance:

$$v_p^{(m)}(x) = \frac{v_\eta^{(m)}(x)}{pq}. \tag{48}$$

E.3 NEGATIVE BINOMIAL: $\eta = \mathrm{logit}\, p$

We use the convention

$$\Pr(X = x) = \frac{\Gamma(r+x)}{\Gamma(r)x!}(1-p)^r p^x, \qquad q = 1 - p, \tag{49}$$

so that $\mathbb{E}[X] = rp/q$. For $\eta = \mathrm{logit}\, p$, the order-2 estimator is

$$\boxed{v_\eta^{(2)}(x) = \frac{rp}{2} + \frac{1+p}{2}\, x.} \tag{50}$$

The order-3 estimator is

$$\boxed{\begin{aligned} v_\eta^{(3)}(x) = \frac{rp(p+5)}{12} \\ + \left[\frac{1+p}{2} + \frac{q^2(2rp+1+p)}{12p(r+1)}\right] x \\ - \frac{q^3}{12p(r+1)}\, x^2. \end{aligned}} \tag{51}$$

The corresponding raw-p weight is

$$v_p^{(m)}(x) = \frac{v_\eta^{(m)}(x)}{pq}. \tag{52}$$

E.4 Negative binomial: $\rho = \log r$

Under the same negative-binomial convention, let

$$\rho = \log r. \tag{53}$$

The order-2 estimator is

$$\boxed{v_\rho^{(2)}(x) = \frac{rp}{2q} + \frac{1}{2}\, x.} \tag{54}$$

The order-3 estimator is

$$\boxed{\begin{aligned} v_\rho^{(3)}(x) = \frac{rp(5-p)}{12q} \\ + \left[\frac{1}{2} + \frac{(1+p)(2rp+1+p)}{12p(r+1)}\right] x \\ - \frac{q(1+p)}{12p(r+1)}\, x^2. \end{aligned}} \tag{55}$$

The corresponding raw-r weight is

$$v_r^{(m)}(x) = \frac{v_\rho^{(m)}(x)}{r}. \tag{56}$$

E.5 Beta-binomial: $\eta = \operatorname{logit} \pi$

Let

$$X \sim \mathrm{BetaBinomial}(n, \alpha, \beta), \qquad \alpha = \kappa\pi, \qquad \beta = \kappa(1-\pi), \tag{57}$$

and write

$$q = 1 - \pi, \qquad \eta = \operatorname{logit} \pi. \tag{58}$$

The order-2 estimator is

$$\boxed{v_\eta^{(2)}(x) = \frac{n\pi}{2} + \frac{1-2\pi}{2}\, x.} \tag{59}$$

For order 3, define

$$S := \frac{2\kappa n\pi - 2\kappa\pi + \kappa + 2n}{\kappa + 2}, \tag{60}$$

$$T := \frac{\kappa n\pi(n-1)(\kappa\pi + 1)}{(\kappa+1)(\kappa+2)}, \tag{61}$$

and

$$C_\eta := -\frac{(\kappa+1)(\kappa+2)(\kappa+3)(\kappa+2n)}{12\kappa(n-1)(\kappa\pi+1)(\kappa q+1)(\kappa+n+1)}. \tag{62}$$

Then

$$\boxed{\begin{aligned} v_\eta^{(3)}(x) = &\left(\frac{n\pi}{2} + C_\eta T\right) \\ &+ \left(\frac{1-2\pi}{2} - C_\eta S\right) x \\ &+ C_\eta x^2. \end{aligned}} \tag{63}$$

The corresponding raw-π weight is

$$v_\pi^{(m)}(x) = \frac{v_\eta^{(m)}(x)}{\pi(1-\pi)}. \tag{64}$$

E.6 Beta-binomial: $\rho = \log \kappa$

Under the same mean–concentration parameterization, let

$$\rho = \log \kappa. \tag{65}$$

Retain the definitions of q, S, and T from Section E.5, and define

$$D := \frac{\kappa(n-1)}{2(\kappa+1)(\kappa+n)}. \tag{66}$$

The order-2 estimator is

$$\boxed{v_\rho^{(2)}(x) = Dn\pi - Dx.} \tag{67}$$

For order 3, define

$$C_\rho := \frac{\kappa(\kappa+3)(2\pi-1)(3\kappa+2n+2)}{12(\kappa+n)(\kappa\pi+1)(\kappa q+1)(\kappa+n+1)}. \tag{68}$$

Then

$$\boxed{\begin{aligned} v_\rho^{(3)}(x) = &\left(Dn\pi + C_\rho T\right) \\ &+ \left(-D - C_\rho S\right) x \\ &+ C_\rho x^2. \end{aligned}} \tag{69}$$

The corresponding raw-κ weight is

$$v_\kappa^{(m)}(x) = \frac{v_\rho^{(m)}(x)}{\kappa}. \tag{70}$$

F Two-Parameter Straight-Through Implementation

Consider a hard scalar sample $Z_{\text{hard}} \sim p_{a,b}$ with optimized coordinates a and b. Suppose the desired same-path estimators are

$$\partial_a \mathbb{E}[f(Z)] \approx \mathbb{E}[u(Z) f'(Z)]\,, \tag{71}$$

$$\partial_b \mathbb{E}[f(Z)] \approx \mathbb{E}[v(Z) f'(Z)]\,. \tag{72}$$

Here u and v may be any minimum-norm weights above, expressed in the model's optimized coordinates a and b.

Let $\mathrm{sg}(\cdot)$ denote stop-gradient. The corresponding straight-through surrogate is

$$\boxed{Z_{\text{st}} = \mathrm{sg}(Z_{\text{hard}}) + \mathrm{sg}(u(Z_{\text{hard}}))\big(a - \mathrm{sg}(a)\big) + \mathrm{sg}(v(Z_{\text{hard}}))\big(b - \mathrm{sg}(b)\big).} \tag{73}$$

For a Negative-Binomial variable parameterized by

$$a = \eta = \operatorname{logit} p, \qquad b = \rho = \log r, \tag{74}$$

one evaluates

$$u(Z_{\text{hard}}) = v_\eta^{(m)}(Z_{\text{hard}}), \qquad v(Z_{\text{hard}}) = v_\rho^{(m)}(Z_{\text{hard}}), \tag{75}$$

and inserts both weights into the same straight-through expression.

In PyTorch pseudocode, the scalar two-parameter construction is

```
z_st = (
    z_hard.detach()
    + u.detach() * (a - a.detach())
    + v.detach() * (b - b.detach())
)
```

G Additional Experimental Details

Unless stated otherwise, means and standard deviations are computed over independently seeded runs, with seeds matched across estimator configurations. One additional detail on the tested method, is that for ReinMax, we take the largest rate of the batch, and calculate truncation using that rate and probability $p = 0.999$.

G.1 Linear Poisson VAE

We follow the benchmark in Appendix G.1 of Ibrahim et al. (2026). EAT-cubic is tuned over the reported temperature grid $[0.02, 0.05, 0.2]$; 0.05 performs best in our replication. Across five independent matched seeds, order 2 obtains -167.02 ± 0.05, EAT-cubic obtains -167.11 ± 0.06 at temperature 0.05 ($p = 0.034$, two-sided Welch t-test), and the exact reference gives -166.99 ± 0.04. Because final ELBO saturates in this controlled benchmark, we use it primarily to validate exactness and gradient variance rather than as a headline performance comparison.

G.2 Neural sequence model

We follow the synthetic spike-train benchmark, temperature grids, and tuning protocol in Appendix G.2 of Ibrahim et al. (2026), except for initialization. In our implementation, the stated $N(0, 0.5^2)$ initialization did not reproduce the reported ELBO range: with either one or five Monte Carlo samples, runs converged around -427 to -429. The repository-default `Conv1d` initialization reproduced results near -330.5, so we used it for every estimator and held it fixed across methods. Reported statistics use 30 independent matched seeds.

G.3 Nonlinear Poisson and count VAEs

We reuse the experimental design of Titsias & Shi (2022), which follows earlier discrete-gradient control-variate studies (Yin & Zhou, 2019; Dong et al., 2020; Yin et al., 2019), and replace only the latent distribution with Poisson. The prior is $\mathrm{Pois}(1.0)$: no proper uninformative Poisson prior exists, and rate 1.0 corresponds to natural parameter 0. For the count-family comparison, we retain the same design but restrict the experiments to dynamically binarized MNIST. Binomial and Negative-Binomial latents use $p = 0.5$, while the sparse Poisson setup uses $\lambda = 0.1$. ReinMax temperatures are tuned over $[1.0, 1.1, 1.3, 1.5]$ and EAT-cubic temperatures over $[0.02, 0.05, 0.2, 0.5]$. Both experiment families report five independent matched seeds per estimator configuration.

G.4 Binomial support-size timing

The benchmark evaluates a batch of 200×128 Binomial latents on an NVIDIA L4. We sweep $n \in \{4, 8, 16, 32, 64, 128, 256, 512, 1024\}$, corresponding to categorical support size $K = n+1$. The comparison includes native order-2 and order-3 weights and the released 1D MVE implementation of Hooper & Shekhovtsov (2026) at $\tau = 10^6$. Figure 4 reports milliseconds per forward–backward evaluation.

G.5 DEF experiment details

The DEF experiments use a reduced grid relative to the nonlinear-VAE experiments. ReinMax temperatures are $\{1, 1.3, 1.5\}$, EAT-cubic temperatures are $\{0.02, 0.05, 0.2\}$, and learning rates are $\{3 \times 10^{-3}, 10^{-3}, 3 \times 10^{-4}\}$. We use AdamW with batch size 64, train for 500 epochs on 20 Newsgroups and 100 epochs on RCV1, and report five independent matched seeds for each selected configuration.

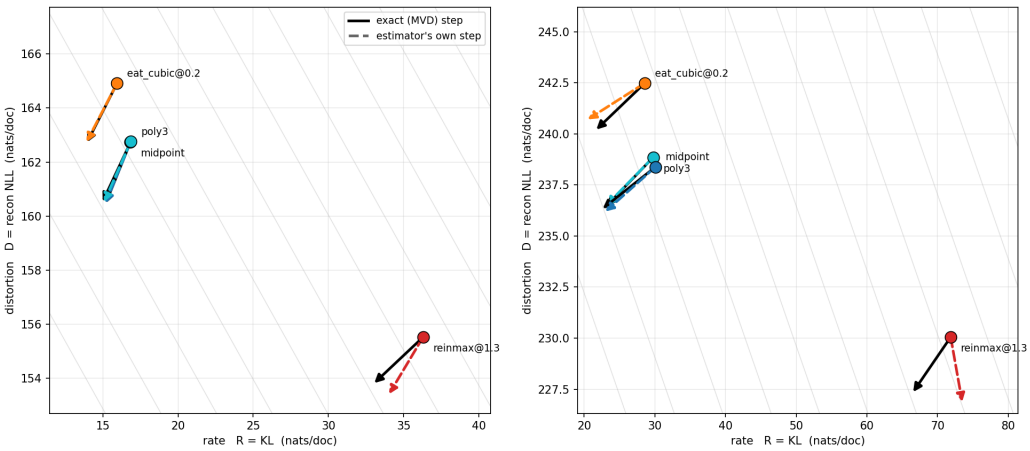


Figure 5: Local rate–distortion update directions in selected high-performing runs. Our estimators closely track the near-exact MVD reference; EAT-cubic is close on 20 Newsgroups but deviates more on RCV1, while ReinMax tilts toward lower distortion and higher rate.

We use amortized inference because storing document-specific latent variables for RCV1 would be memory-intensive. The decoder follows the Gaussian-weight, softplus-link Poisson DEF of Ranganath et al. (2015); its weights are point-estimated, with 30 supertopics and 100 topics. The encoder is a two-layer neural network mirroring the decoder. As in the original model, a $\mathrm{Pois}(0.1)$ prior encourages sparsity.

For 20 Newsgroups, we use the master-branch data from `https://github.com/akashgit/autoencoding_vi_for_topic_models`, also used by the AVITM paper. The vocabulary has 1,995 terms, with 11,258 training and 7,487 test documents. For RCV1, we use the stemmed token files from `http://www.ai.mit.edu/projects/jmlr/papers/volume5/lewis04a/a12-token-files/`, retain the 10,000 most frequent terms, pool the documents, and randomly split 794,414 for training and 10,000 for testing.

G.6 ReinMax and the rate–distortion tradeoff

On 20 Newsgroups, ReinMax attains better held-out perplexity despite substantially worse train and test ELBO. We therefore inspect local update directions in selected high-performing runs.

In Figure 5, our estimators and EAT-cubic closely track the near-exact reference in the displayed runs. ReinMax is broadly aligned but places more weight on decreasing distortion, especially in the displayed 20 Newsgroups run, where it moves toward a higher-rate, lower-distortion solution.

Lower distortion improves reconstruction, whereas the ELBO also penalizes KL rate. ReinMax's local bias therefore acts like a reweighting toward reconstruction. Its perplexity advantage should not be interpreted as more accurate ELBO-gradient estimation; the displayed direction diagnostic points in the opposite direction.

G.7 Timing protocol

Each value in Table 5 is the mean $\pm$ standard deviation over five timing repeats after warm-up. Within a workload, the model, batch size, hardware, and training configuration are fixed across estimators.

- **Linear Poisson VAE:** warm up for two epochs, record cumulative elapsed time at epochs 8 and 24, and divide the difference by 16.
- **Nonlinear Poisson VAE:** warm up for two epochs, record cumulative elapsed time at epochs 3 and 9, and divide the difference by 6.
- **Sparse Poisson DEF:** warm up for two epochs, record cumulative elapsed time at epochs 4 and 12, and divide the difference by 8.

We repeat this procedure five times for each estimator–workload–hardware combination.