CORRECTING SELECTION BIAS IN POSITIVE AND UNLABELED GENERATION WITH SELECTION RECORDS

Anonymous authors
Paper under double-blind review

ABSTRACT

Positive and unlabeled (PU) data combine labeled positives with a population sample whose classes are unknown. A generator fitted to the labeled positives inherits their selection bias: positives that are verified more often stay overrepresented, even when every label is correct. We correct this bias with selection records, which show whether objects in a population sample were selected for verification but not their classes. Under shared selection and constant reporting, the normalized density ratio of labeled positives to selected objects identifies the positive distribution without selection probabilities or the population frequency of positive labels. We bound the resulting distribution error by the ratio estimation error and guide a diffusion model with the fitted weights. Under simulated biased selection on MNIST, CelebA and CIFAR-10, the weights improve the group composition of the weighted positives, and joint class and group error falls on MNIST and CIFAR-10. On MNIST, generated composition also improves at matched fractions of predicted positives, although even weights equal to the true class leave generation error. On CelebA, longer ratio training lowers the joint error of the weighted target, but the held out classification loss does not select it. Further experiments separate the information in the records from their number and show that restricted probability models, mismatched records and class dependent selection prevent recovery.

1 INTRODUCTION

Positive and unlabeled (PU) learning uses a sample L of labeled positive objects and a sample U from the whole population whose classes are unknown, so an unlabeled object is not a known negative. Most PU methods decide which existing objects are positive (Elkan & Noto, 2008; Bekker & Davis, 2020; Chen et al., 2026); we instead model the positive class to generate new positives. This helps when positives are scarce: in mineral prospectivity mapping, known deposits are few, unexplored locations remain unlabeled, and generated positives can augment the scarce class (Farahbakhsh et al., 2023). Generative PU models and PU diffusion models extend PU learning from classification to generation (Hou et al., 2018; Na et al., 2020; Takahashi et al., 2025; Raymond et al., 2025).

Generation is harder when the labeled positives do not represent all positives. When some positives are more likely to be labeled (Bekker & Davis, 2018; Kato et al., 2019; Gong et al., 2022), the labeled positive distribution P_l differs from the distribution P_+ of all positives. In Figure 1, brighter and darker smiling faces are equally common, but brighter faces are four times as likely to be labeled, so four of five labeled faces are bright. A generator fitted to these labels produces valid smiling faces in the wrong proportions, and more labels from the same process keep this bias.

In general, L and U alone do not determine P_+: an underrepresented kind of positive may be rare or merely rarely labeled, and the two samples need not distinguish these explanations (Bekker et al., 2019; Coudray et al., 2023). Methods for such labeling therefore model or jointly estimate the labeling mechanism (Bekker & Davis, 2018; Gong et al., 2022; Rejchel et al., 2024). For generation, this mechanism is only a means to P_+. We ask whether a record of the verification process can recover P_+ without estimating each object's probability of selection.

When verification is logged, the log provides such a record: for objects in a population sample, it shows whether each was selected for verification but not its class. Kato et al. (2023) used such exposure records to correct selection bias in PU classification; the records also supply a reference sample for generation. Let $v(x)$ be the probability that an object with attributes x is selected, and f,

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.

f_+, f_l and f_A the densities of all objects, positives, labeled positives and selected objects. If labels and records arise from the same selection process, selection does not depend on the class given x, and selected positives are reported at a constant rate, then labeled positives and selected objects carry the same factor v, which cancels in their ratio:

$$f_l(x) \propto v(x) f_+(x), \qquad f_A(x) \propto v(x) f(x), \qquad \frac{f_l(x)}{f_A(x)} \propto \frac{f_+(x)}{f(x)}.$$

Weighting U by this ratio and normalizing therefore recovers P_+.

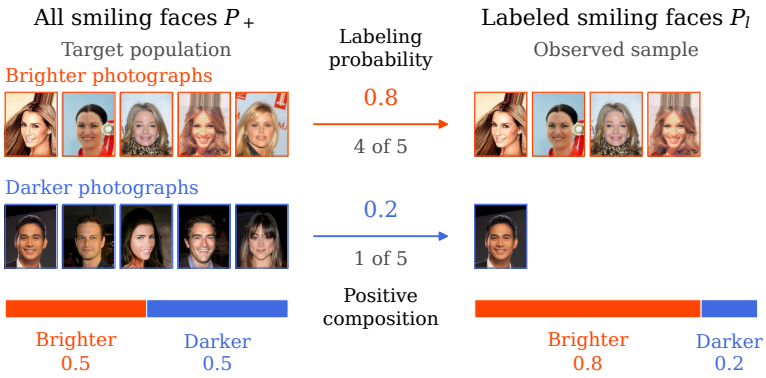

Figure 1: **Selective verification distorts the generative target.** Brighter smiling faces are verified more often, so labeled positives overrepresent them relative to the balanced population. Images are unchanged CelebA examples (Liu et al., 2015); proportions and labeling probabilities are constructed.

Contributions. We turn this cancellation into a correction for PU generation. A classifier estimates the ratio of labeled positives to selected objects without selection probabilities or the population frequency of positive labels, and Theorem 2 bounds how its estimation error propagates to the normalized distribution. The resulting weights guide a pretrained diffusion model through conditional energy guidance (Lu et al., 2023). Experiments on MNIST, CelebA and CIFAR-10 separate weight on negatives, misallocation among kinds of positives, and generation error. On MNIST, generated composition improves at matched fractions of predicted positives; on CelebA it does not, and longer ratio training improves the weighted target but not consistently the generated images.

2 Setting and Identification

2.1 Population, Labels and Records

Each object in a population has attributes X and a class $Y \in \{0, 1\}$; in our experiments, X is an image and $Y = 1$ marks the class to generate. Write P, P_+ and P_- for the distributions of X among all objects, positives and negatives, with densities f, f_+ and f_-. With the class prior $\alpha = \Pr(Y = 1) \in (0, 1)$, $f = \alpha f_+ + (1-\alpha) f_-$. The target is P_+; Appendix A collects the notation.

Labels arise through verification. An object is either selected for verification ($A = 1$) or not ($A = 0$). Verification reveals the class of a selected object; a selected positive is then either reported as a labeled positive ($S = 1$) or not, and negatives are never reported, so $S = 1$ implies $A = 1$ and $Y = 1$. A positive with attributes x is labeled with probability $e(x) = \Pr(S = 1 \mid Y = 1, X = x)$, its labeling propensity, whose mean $c = \Pr(S = 1 \mid Y = 1) > 0$ is the label frequency. Labeled positives therefore have density $f_l = e f_+ / c$ and distribution P_l.

We observe three independent samples: labeled positives L from P_l; an unlabeled sample U from the whole population P, with classes hidden; and n_V selection records $V = \{(X_i, A_i)\}_{i=1}^{n_V}$, whose attributes follow P and which show whether each object was selected but not its class. The sample sizes are set by design and carry no information about α; Appendix D.1 gives them, including records held out for validation. Write $\mathbb{E}$, $\mathbb{E}_+$ and $\mathbb{E}_l$ for expectations under P, P_+ and P_l.

Weighting P by the density ratios

$$r = \frac{f_+}{f}, \qquad r_l = \frac{f_l}{f} = \frac{re}{c}$$

gives P_+ and P_l, respectively. Unless e is constant, r_l is not proportional to r, so even an exact estimate of r_l reproduces the selection bias: estimating P_l and recovering P_+ are different tasks.

2.2 IDENTIFICATION

A target is identified when the observed distributions and the stated assumptions determine it uniquely. L and U alone do not identify P_+. Let X record only brightness, with values bright and dark, and consider two models with the same $P = (0.5, 0.5)$ and $P_l = (0.8, 0.2)$. In the first, $P_+ = (0.5, 0.5)$ and $e = (0.8, 0.2)$, as in Figure 1; in the second, $P_+ = (0.8, 0.2)$ and $e = 0.5$ for both values. Both are consistent with $\alpha = 0.5$ and $c = 0.5$, so L and U cannot distinguish unequal labeling propensities from unequal prevalence among positives (Proposition B.1).

The ambiguity arises when the classes overlap in X: if the class is a function of X and $e > 0$ on the support of P_+, then $P_+ = P(\cdot \mid r_l > 0)$ is identified. The selection records identify the selection probability $v(x) = \Pr(A = 1 \mid X = x)$ but not how selection and reporting treat positives; the next assumption links records to labels.

Assumption 1 (Shared selection and constant reporting)**.** L and V arise from the same selection process, with selection probability v. Given the attributes, selection is independent of the class, $A \perp\!\!\!\perp Y \mid X$, and a selected positive is reported with a constant probability $c_{\text{rep}} > 0$. Moreover, $v > 0$ almost surely under P_+.

Under Assumption 1, $e = c_{\text{rep}} v$, so $f_l \propto v f_+$; dividing r_l by v and normalizing gives

$$r = \frac{r_l / v}{\mathbb{E}(r_l / v)}, \qquad \mathbb{E}(r_l / v) = \mathbb{E}_l(1/v),$$

with $r_l / v = 0$ where $v = 0$, a set to which P_+ assigns no mass. If selection depends on the class or reporting varies with x, dividing by v can leave a bias that the records cannot reveal (Section 4.4); Proposition B.10 bounds the error when the reporting probability varies within a specified range.

3 ESTIMATION AND GENERATION

We first estimate weights on U that recover P_+ and bound their error, then use them to guide a pretrained diffusion model (algorithms in Appendix C.1). Each estimator trains a classifier to separate L from a reference sample. The uncorrected estimator uses all record images as reference and ignores A, so it estimates r_l. Inverse probability weighting (IPW) keeps this reference and reweights labeled positives by inverse estimated selection probabilities; the selected reference (SR) estimator uses only the selected records.

3.1 SELECTION CORRECTED WEIGHTS

Inverse probability weighting. IPW fits a model $\widehat{v}$ of the selection probability to all records and weights the ith labeled positive in the classifier loss by $\omega_i = (1/\widehat{v}_i) / \sum_j 1/\widehat{v}_j$, so rarely selected positives count more. With the true probabilities, this removes the selection bias in the population limit (Appendix B.2); estimated probabilities add error, and a model that merges groups with different selection probabilities leaves a bias that more records do not remove.

Selected reference. SR instead compares the labeled positives with the selected records. Let $P_A = P(\cdot \mid A = 1)$ be the distribution of selected objects, with density $f_A = v f / \mathbb{E}v$ and expectation $\mathbb{E}_A$. If also $v > 0$ almost surely under P, the factor v cancels, $r_A \coloneqq f_l / f_A = (\mathbb{E}v / \mathbb{E}_+ v)\, r$, and normalization over U removes the constant. Besides L and U, SR needs only a representative sample of selected objects, neither the unselected records nor the overall selection rate (Corollary B.8). In Figure 2, the uncorrected ratio overweights the right positive peak; the SR ratio restores both.

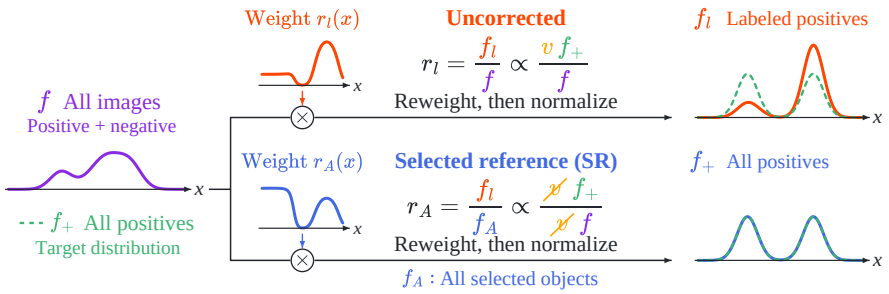

Figure 2: Graphical interpretation of selection correction for an image attribute x. Reweighting all images by r_l yields the labeled positives, whose right peak exceeds that of all positives (dashed). The SR ratio r_A uses the selected objects as reference, so $v(x)$ cancels and reweighting recovers f_+ under Assumption 1 with $v > 0$ on the population support. Density and weight plots each share a scale.

Estimating the ratio. A classifier that predicts whether an image comes from L or from the selected records estimates r_A: with $\ell(z) = \log(1 + e^{-z})$, the balanced logistic loss of a log ratio g,

$$\mathcal{L}_A(g) = \tfrac{1}{2}\mathbb{E}_l \ell(g(X)) + \tfrac{1}{2}\mathbb{E}_A \ell(-g(X)), \tag{1}$$

has population minimizer $g = \log r_A$, with $g = -\infty$ where $r_A = 0$. The fitted e^g is divided by its mean over U and truncated at $a_{\max}$, which gives fixed image weights a (Algorithm 1).

3.2 Error in the recovered distribution

The next result bounds the total variation distance (TV) between the normalized weighted distribution and P_+ by the ratio error and by the excess population loss.[1]

Theorem 2 (Positive distribution recovery)**.** *Suppose Assumption 1 holds and $v > 0$ almost surely under P. For a nonnegative estimate $\widehat{r}_A$ with $0 < \mathbb{E}\widehat{r}_A < \infty$, put $\widehat{Q} = (\widehat{r}_A/\mathbb{E}\widehat{r}_A)P$. Then*

$$\mathrm{TV}(\widehat{Q}, P_+) \le \frac{\mathbb{E}\,|\widehat{r}_A - r_A|}{\mathbb{E} r_A}. \tag{2}$$

If also $v \ge v_{\min} > 0$ almost surely under P and $0 \le r_A, \widehat{r}_A \le B < \infty$, then

$$\mathrm{TV}(\widehat{Q}, P_+) \le 2\sqrt{\frac{\mathbb{E}_+ v\, J_B \Delta}{v_{\min}}}, \qquad J_B = 2(1 + B), \tag{3}$$

where $\Delta = \mathcal{L}_A(\log \widehat{r}_A) - \inf_g \mathcal{L}_A(g)$ is the excess population loss in (1)*, with the infimum over all measurable g, including $g = -\infty$ where $r_A = 0$.*

Normalization removes the unknown scale $\mathbb{E}v/\mathbb{E}_+ v$, so only the relative L^1 error of $\widehat{r}_A$ matters. The second bound ties this error to the classifier's loss, with a constant that grows as $v_{\min}$ shrinks.

3.3 Guided generation

Each estimator yields fixed weights $a \ge 0$ on clean images, that is, images without added noise; these define the target $Q_a = aP/\mathbb{E}a$ with $0 < \mathbb{E}a < \infty$. A denoising diffusion probabilistic model (DDPM) generates images by gradually removing noise (Ho et al., 2020); our base model is an unconditional DDPM for P, trained on U or on a larger sample from the same population. Let X_t be a Gaussian perturbation of a clean image $X_0 \sim P$ at noise level t, with density p_t; the base model estimates the score $\nabla \log p_t$, the denoising direction (Song et al., 2021b). Following Lu et al. (2023), we fit $\phi_a(t, x) = \log \mathbb{E}[a(X_0) \mid X_t = x]$ with a weighted logistic loss at every noise level; adding its gradient to the base score gives the score of the noised target (Proposition C.3). The base model and ϕ_a stay fixed during sampling. Because the weights are fitted on clean images, where v cancels, guidance only carries the corrected target to each noise level.

Generation adds its own error, since the base score and ϕ_a are estimated and sampling uses finitely many steps. Corollary B.9 separates this error from the effect of truncating the weights, Appendix C bounds it for a specified kernel estimator, and RQ1 measures it (Appendix D.10).

[1] Proposition B.7 in Appendix B.4 proves both bounds.

4 EXPERIMENTS

We ask what selection correction recovers in the weights on U and how much of it reaches the generated images. The positive classes are even digits in MNIST, smiling faces in CelebA and vehicles in CIFAR-10. Selection is simulated from image intensity (MNIST, CIFAR-10) or the Male annotation (CelebA); its probabilities are known to us but given to no estimator.

Design and evaluation. MNIST isolates estimation error, and a finite population of four types isolates violated assumptions (RQ4). CelebA and CIFAR-10 use unconditional models pretrained on their populations and shared by all methods.* The positive reference is a held out set of positive images. Precision is the fraction of generated images near some reference image, and coverage the fraction of reference images near some generated image, where an image is near a reference image if it lies within the distance from that reference image to its fifth nearest reference neighbor in a feature space ($k = 5$). Squared maximum mean discrepancy (MMD^2) on MNIST and Fréchet and kernel Inception distances (FID, KID) on CelebA and CIFAR-10 compare whole feature distributions; smaller is better (Appendix D.3). Comparisons are repeated five times with new seeds for fitting and sampling; the main CelebA and CIFAR-10 comparisons also redraw L and V. We report means $\pm$ standard error (SE), paired for differences; proportions are in %, their differences in pp, and empirical TV values are multiplied by 100. Appendices D.1, D.2 and D.6 give sample sizes, training and the additional repetitions.

All four methods in Table 1 share one base model and sampler; the three guided methods differ only in the weights on U that define their target (Section 3). **Unguided** draws samples from the base model without guidance. **Uncorrected** fits the ratio classifier with all record images as reference and ignores A, so it estimates r_l. **IPW** keeps this reference and weights each labeled positive by the inverse of its estimated selection probability, which on MNIST comes from one probability per intensity quartile, a correctly specified model, and on CelebA and CIFAR-10 from a network on pixels. **SR** uses only the selected records as reference.

Table 1: Precision and coverage (%) of generated images, $k = 5$: mean $\pm$ SE over five repetitions; MNIST biased selection is strong selection. Bold and underlined: highest and second highest mean per condition and column. Unguided samples do not depend on selection and are shared.

Selection	Method	MNIST		CelebA		CIFAR-10	
		Precision	Coverage	Precision	Coverage	Precision	Coverage
Constant	Unguided	64.84±0.68	86.11±0.66	78.68±0.32	79.24±0.22	38.09±0.49	79.22±1.17
	Uncorrected	**82.05**±0.59	<u>92.11</u>±0.67	<u>80.79</u>±0.39	<u>82.20</u>±0.19	**61.76**±1.22	<u>92.66</u>±0.46
	IPW	<u>80.21</u>±0.67	**92.93**±0.47	**80.97**±0.36	**82.38**±0.27	61.15±1.12	92.21±0.83
	SR	78.42±1.43	91.95±0.51	80.66±0.24	82.16±0.53	<u>61.41</u>±1.20	**92.99**±0.77
Biased	Unguided	64.84±0.68	86.11±0.66	78.68±0.32	79.24±0.22	38.09±0.49	79.22±1.17
	Uncorrected	**82.81**±0.69	88.34±0.91	<u>80.79</u>±0.29	**82.34**±0.26	**62.77**±1.48	91.82±0.87
	IPW	<u>80.20</u>±0.59	<u>91.23</u>±1.09	**80.86**±0.37	82.22±0.32	<u>61.43</u>±1.16	<u>92.15</u>±0.50
	SR	78.75±0.88	**92.91**±0.86	80.50±0.31	<u>82.28</u>±0.40	60.68±0.96	**92.23**±0.76

4.1 RQ1: DOES SELECTION CORRECTION IMPROVE THE WEIGHTS AND GENERATED IMAGES?

We evaluate the weights on U, then images generated with a common base model and sampler.

Positive mass and composition. The true annotations of U, which no estimator sees, measure the share of weight on positives, called positive mass, and how this weight is spread over the selection groups (Table 2). Under biased selection, SR lowers the group TV on all three datasets, and the joint TV, which also counts weight on negatives, on MNIST and CIFAR-10. On CelebA, the Male fraction among weighted smiling faces rises from 20.16% to 33.87%, near the 34.92% in U, but positive mass falls from 62.71% to 57.90%, raising the joint TV before generation.

The CelebA gain in composition is not explained by the lower positive mass: mixing the uncorrected weights with uniform weights on U until the positive mass equals that of SR gives 24.15% Male

among smiling faces, against 33.87% for SR. This comparison uses the annotations to set the mixture and serves only for evaluation (Appendix D.4).

Table 2: **Weights on U under biased selection.** Mean ± SE over five repetitions, evaluated with true annotations against the positives in U. Group and subclass TV compare the selection groups and the positive classes (digits or vehicle types) among weighted positives with those among the positives in U; CelebA has the single positive class Smiling, so its subclass TV is zero (−) and its Male composition appears under Group. Joint TV compares class and group over all images. Effective size is $100\,\mathrm{ESS}/|U|$; larger values mean less concentrated weights. Bold: smallest mean error per dataset.

Dataset	Method	Positive mass (%) ↑	Distribution error (100 TV) ↓			Effective size (%)
			Group	Subclass	Joint	
MNIST	Uncorrected	98.88±0.16	27.68±3.25	14.19±2.53	29.98±3.20	12.16±1.27
	IPW	98.03±1.02	26.06±3.89	9.92±1.43	27.49±3.68	10.08±0.98
	SR	97.96±0.27	**4.10**±1.19	**6.00**±0.66	**9.84**±0.73	7.59±0.47
CelebA	Uncorrected	62.71±1.48	14.76±1.60	–	**37.29**±1.48	53.45±5.19
	IPW	60.07±0.71	8.30±0.95	–	39.93±0.71	62.68±2.69
	SR	57.90±0.75	**1.10**±0.41	–	42.10±0.75	74.85±1.38
CIFAR-10	Uncorrected	83.83±1.65	23.78±0.94	7.19±1.11	29.46±1.26	33.13±1.62
	IPW	83.93±1.23	4.86±1.21	4.62±0.45	17.20±1.04	39.02±1.82
	SR	83.77±1.48	**4.02**±1.47	**3.40**±0.87	**16.96**±1.09	40.98±2.02

Generated images. Weaker guidance alone also changes composition and coverage. We thus vary guidance strength (Figure 3) and compare SR with a mixture of uncorrected samples at the two strengths whose predicted positive fractions bracket that of SR, mixed to match it in each repetition (Appendix D.6).

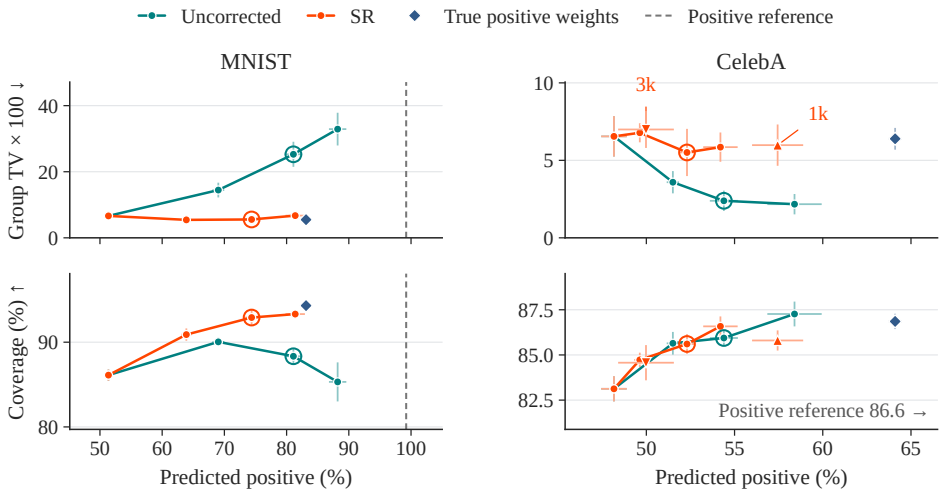

Figure 3: **Composition and coverage across guidance strengths.** Curves connect strengths $0, 0.5, 1, 1.5$, starting from the shared unguided samples; rings mark strength 1. The horizontal axis is the percentage of images predicted positive, which the dashed line marks for the reference. Group TV ($\times 100$) compares the intensity groups (MNIST) or the Male fraction (CelebA) of predicted positives with the reference. Diamonds weight each image by its true class; triangles show SR at strength 1 after 1,000 or 3,000 ratio training iterations. Mean ± SE over five repetitions.

At the matched fraction, SR lowers the MNIST group TV by 14.09 ± 2.95, in every repetition; against held out labeled positives, which share the bias, the sign reverses. On CelebA, the Male TV instead

increases by 2.76 ± 1.48, in four of five repetitions. Five further MNIST repetitions, reusing the L and V of the weight evaluation, also improve composition and coverage at lower precision.

Generation adds its own error. With each image weighted by its true class, 83.1% of generated images are predicted positive on MNIST and 64.1% on CelebA, against 99.2% and 86.6% of the reference, and the MNIST sampler misses known weighted targets (Appendices D.6 and D.10). On CelebA, 53.36% of the SR target on U and 52.30% of its generated images are predicted Smiling. Training the ratio classifier for 1,000 or 3,000 instead of 200 iterations raises the target to 72.92% and 73.67%, but generation reaches only 57.44% and 49.98% (Table 12). Annotation results motivated this training; the held out classification loss, free of evaluation labels, always selects 200 iterations.

In the main comparison (Table 1), SR lowers FID and KID on CelebA, but neither CelebA nor CIFAR-10 gains coverage (Appendix D.5), and both lose coverage relative to resampling U with the same weights (Appendix D.9). RQ2 separates the size of the SR reference from its composition.

4.2 RQ2: Do selection indicators help beyond reducing the reference size?

The selected reference differs from all records in both composition and size. We therefore also fit against a uniformly random subset of all records, as large as the selected subset: random minus all measures the effect of size, and selected minus random that of the indicators at equal size (Figure 4). The weights are evaluated by resampling U, which excludes generation error.

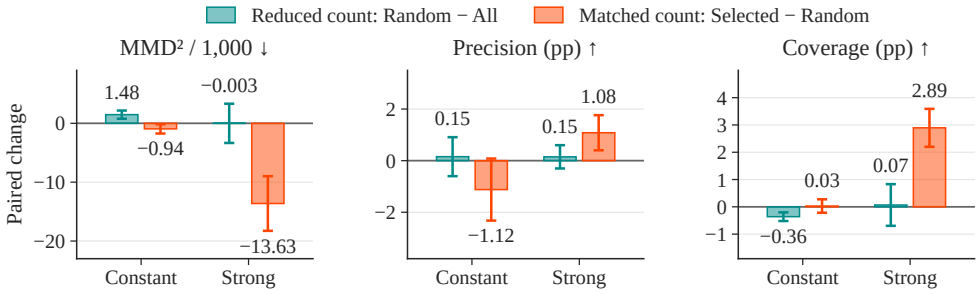

Figure 4: **Reference size and composition on MNIST.** Bars are paired changes in weighted resamples of U; Random − All changes only the reference size, and Selected − Random only its composition. Mean ± SE over five repetitions with new L and V, each averaging twenty resamples.

At equal size, the selected reference increases MNIST coverage by 2.89 ± 0.70 pp under strong selection and by 0.03 ± 0.25 pp under constant selection, where mean precision is lower (Table 13). Under biased selection, it increases coverage by 1.86 ± 0.30 pp on CelebA and 1.68 ± 0.69 pp on CIFAR-10 and lowers KID on both, while CelebA precision falls by 1.60 ± 0.49 pp. On CIFAR-10, the smaller size alone lowers coverage by 1.46 ± 0.60 pp, offsetting most of the gain. Under constant selection, neither CelebA nor CIFAR-10 gains mean coverage from the selected reference (Appendix D.7). The indicators thus help through the composition of the reference under biased selection and add little under constant selection; RQ3 turns to their use in fitting selection probabilities.

4.3 RQ3: Can more indicators compensate for a misspecified probability model?

IPW instead uses the indicators to fit selection probabilities. On MNIST, we keep all 8,000 record images as reference and reveal 500, 2,000 or 8,000 indicators. The model has either the four intensity groups that define the simulated selection or two groups that each merge two of them and so cannot represent strong selection. Figure 5 compares the mean squared error (MSE) of the estimated probabilities with the resulting weighted distributions.

Probability estimation. Under strong selection, the four group MSE ($\times 10^4$) falls from 12.86 with 500 indicators to 0.83 with 8,000, whereas the two group MSE approaches its approximation error, 181.25: more records reduce estimation error, not the error of merging distinct probabilities. Under constant selection, both models are correct and two groups give lower mean error (Table 16).

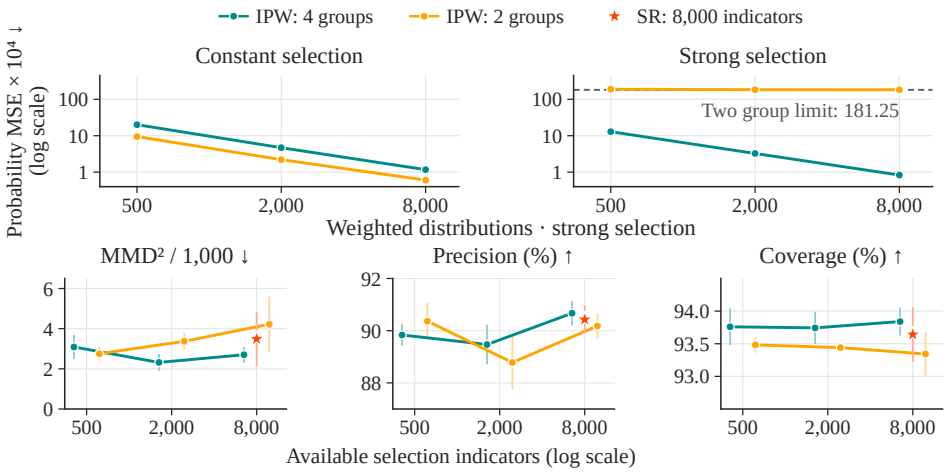

Figure 5: **Selection probability models on MNIST.** The upper panels show the MSE ($\times 10^4$) of probabilities estimated with four or two groups, mean $\pm$ Monte Carlo SE over 200 record samples; the dashed line is the smallest MSE two groups can attain. The lower panels evaluate weighted resamples of U under strong selection, mean $\pm$ SE over five repetitions of twenty resamples.

Weighted distributions. Better probability estimates do not uniformly give better weighted resamples: with four groups, MMD^2 does not decrease monotonically with the number of indicators, and coverage barely changes. SR, which fits no probability model, differs from four group IPW at 8,000 indicators by less than one SE in all three metrics, including coverage (-0.20 ± 0.36 pp; Table 15).

4.4 RQ4: Does SR still help when its assumptions are violated?

The image experiments satisfy Assumption 1 by construction; when the class is a function of X, as for digits, $A \perp\!\!\!\perp Y \mid X$ holds for every selection rule. We therefore use four equally likely types i that each contain both classes, with $\eta_i = \Pr(Y = 1 \mid X = i) = (0.2, 0.4, 0.6, 0.8)_i$ and $v = (0.2, 0.5, 0.5, 0.8)$, and report the TV to P_+ as mean $\pm$ Monte Carlo SE over 1,000 repetitions with 1,000 labeled positives and 8,000 records (Appendix D.8).

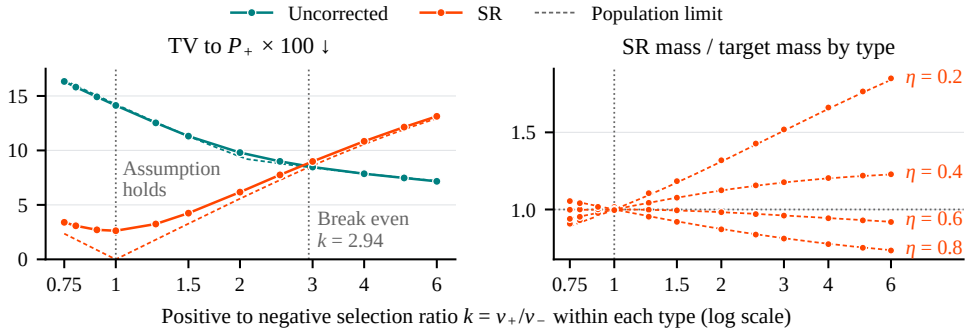

Figure 6: **Class dependent selection in the four type population.** The left panel shows the TV ($\times 100$) of each estimate to P_+; the right panel shows, for each type labeled by its positive rate η, the mass SR assigns to it divided by its mass under P_+. Markers are means $\pm$ Monte Carlo SE over 1,000 repetitions with 8,000 records; dashed lines are population limits from the exact P_l and v.

Records from another selection rule. Let the records follow a rule that shifts the selection probabilities of the two middle types in opposite directions, keeping the overall rate. Relative to the uncorrected estimate, SR reduces TV by 11.60 ± 0.06 when the rules match, but the two most extreme mismatches increase it by 7.93 ± 0.10 and 10.37 ± 0.10.

Class dependent selection. Next, within each type, positives are selected k times as often as negatives while v is held fixed, so the records are unchanged. Labeled positives then follow $f_l \propto v_+ f_+$, with v_+ the selection probability of positives, and SR and IPW with the true v both target $Q_k(i) \propto P_+(i)\, k/\{1 + \eta_i(k-1)\}$, which for $k > 1$ overweights positives in types where they are rare (Figure 6). TV rises from 2.63 ± 0.04 at $k = 1$ to 6.17 ± 0.05 at $k = 2$ and 13.13 ± 0.06 at $k = 6$; beyond $k = 2.94$, the uncorrected weights are closer to P_+ in this population.

In both cases, more records converge to the wrong target: at $k = 2$, TV only falls from 7.32 ± 0.09 with 500 records to 6.08 ± 0.05 with exact v. Each violating model gives (L, U, V) the same distribution as a model satisfying Assumption 1, so no test on these samples can detect either violation (Proposition D.1).

5 Related work

PU learning and validation. Representative labeled positives support estimation of the class posterior and the classification risk (Elkan & Noto, 2008; du Plessis et al., 2014; 2015; Kiryo et al., 2017). Work on PU validation covers sampling corrections (Claesen et al., 2015; Jain et al., 2017), model selection (Wang et al., 2026), biased references (Molotkov & Artomov, 2023) and signal testing (Xu & Ackerman, 2024). Labeling that depends on the attributes motivates propensity weighting and joint estimation (Bekker & Davis, 2018; Bekker et al., 2019; Gong et al., 2022; Rejchel et al., 2024), with risk bounds for known propensities (Coudray et al., 2023). These works concern classification; our target is a distribution, whose frequencies a ranking does not determine.

Selection records and density ratios. Kato et al. (2023) use exposure records without estimating propensities, but their classification objective requires the population frequency of positive labels; normalizing the positive distribution removes this scale. Direct density ratio estimation is established (Nguyen et al., 2010; Menon & Ong, 2016); we bound the distribution error by the ratio estimation loss under overlap and compare corrections when the propensity model is restricted.

Generation and guidance. GenPU, CGenPU and *VAE-PU* connect generation with PU classification (Hou et al., 2018; Papič et al., 2023; Na et al., 2020). PU diffusion models avoid generating sensitive examples (Takahashi et al., 2025) or guide sampling with positive to unlabeled ratios (Raymond et al., 2025). Corrections of selection bias in generative models use importance weights (Diesendruck et al., 2019; Grover et al., 2019), a representative sample of the target (Choi et al., 2020; Kim et al., 2024), or propensity and outcome models (Luedtke & Fukumizu, 2025; Melnychuk & Feuerriegel, 2026); we instead use selection records and compare generation with resampling. Conditional energies and related guidance methods sample weighted targets (Lu et al., 2023; Kim et al., 2023; Garg et al., 2024; Chang et al., 2026), and diffusion theory bounds sampling and estimation error (Chen et al., 2023; Benton et al., 2024; Oko et al., 2023); our Wasserstein bound adds the error of the weights for a specified kernel estimator (Appendix C).

6 Conclusion

We corrected the selection bias of labeled positives with selection records, which show which objects were selected for verification but not their classes. Under shared selection and constant reporting, the normalized ratio of labeled positives to selected objects identifies the positive distribution without selection probabilities, its estimation error bounds the distribution error, and the fitted weights guide a pretrained diffusion model. On three image datasets, the records improve the composition of the recovered positives without a probability model. We hope that selection records, which logged verification already produces, become a routine input to PU generation.

Limitations. The correction rests on shared selection and constant reporting, which are properties of how records and labels are collected rather than of the samples, so they are justified by design; our experiments impose selection on public image datasets, and records from an actual verification process remain to be studied. The correction fixes the target distribution, while generated images also depend on the base model and sampler, so faithful generation from a weighted target is a separate problem, as is choosing the ratio training length without evaluation labels.

REPRODUCIBILITY STATEMENT

Appendix D specifies the observed samples, training, evaluation and all comparisons, including the sampling checks. Models, samples and paired measurements are retained. The PU estimators use only L, U and V; comparisons that use true labels, such as weights equal to the true class, and diagnostics added after the main comparisons are identified separately.

ETHICS STATEMENT

This work involves no human subjects and collects no new data; MNIST, CelebA and CIFAR-10 are publicly available benchmarks. The CelebA Male annotation only simulates selection; it is a dataset attribute, not a verified gender label.

AI USE STATEMENT

In this work, we used generative AI tools for editing the manuscript to improve readability, including grammar checking and wording. We have not used generative AI tools for developing theoretical models, formulating mathematical claims, proving claims, designing methodology or experiments, implementing methods, generating data, or interpreting results. We have reviewed all AI edited text. We take responsibility for the final content of this work, including text, claims and artifacts produced with the aid of generative AI.

REFERENCES

Jessa Bekker and Jesse Davis. Learning from positive and unlabeled data under the selected at random assumption. In *Proceedings of the Second International Workshop on Learning with Imbalanced Domains: Theory and Applications*, volume 94 of *Proceedings of Machine Learning Research*, pp. 8 to 22, 2018. URL https://proceedings.mlr.press/v94/bekker18a.html.

Jessa Bekker and Jesse Davis. Learning from positive and unlabeled data: A survey. *Machine Learning*, 109(4):719 to 760, 2020. doi: 10.1007/s10994-020-05877-5.

Jessa Bekker, Pieter Robberechts, and Jesse Davis. Beyond the selected completely at random assumption for learning from positive and unlabeled data. In *Machine Learning and Knowledge Discovery in Databases*, volume 11907, pp. 71 to 85, 2019. doi: 10.1007/978-3-030-46147-8_5.

Joe Benton, Valentin De Bortoli, Arnaud Doucet, and George Deligiannidis. Nearly d-linear convergence bounds for diffusion models via stochastic localization. In *International Conference on Learning Representations*, 2024. URL https://arxiv.org/abs/2308.03686.

Mikołaj Bińkowski, Danica J. Sutherland, Michael Arbel, and Arthur Gretton. Demystifying MMD GANs. In *International Conference on Learning Representations*, 2018. URL https://arxiv.org/abs/1801.01401.

Jinyuan Chang, Chenguang Duan, Yuling Jiao, Yi Xu, and Jerry Zhijian Yang. Inference-time alignment for diffusion models via variationally stable doob's matching. *arXiv preprint arXiv:2601.06514*, 2026. URL https://arxiv.org/abs/2601.06514.

Qiuyi Chen, Haiyang Zhang, Leqi Zhang, Changchun Li, Jia Wang, and Wei Wang. PU-Bench: A unified benchmark for rigorous and reproducible PU learning. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=tb8DabMbMq.

Sitan Chen, Sinho Chewi, Jerry Li, Yuanzhi Li, Adil Salim, and Anru R. Zhang. Sampling is as easy as learning the score: theory for diffusion models with minimal data assumptions. In *International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2209.11215.

Kristy Choi, Aditya Grover, Trisha Singh, Rui Shu, and Stefano Ermon. Fair generative modeling via weak supervision. In *ICML*, volume 119, pp. 1887 to 1898, 2020. URL https://proceedings.mlr.press/v119/choi20a.html.

Marc Claesen, Jesse Davis, Frank De Smet, and Bart De Moor. Assessing binary classifiers using only positive and unlabeled data. *arXiv preprint arXiv:1504.06837*, 2015. URL https://arxiv.org/abs/1504.06837.

Olivier Coudray, Christine Keribin, Pascal Massart, and Patrick Pamphile. Risk bounds for positive-unlabeled learning under the selected at random assumption. *Journal of Machine Learning Research*, 24(107):1 to 31, 2023. URL https://www.jmlr.org/papers/v24/22-067.html.

Maurice Diesendruck, Ethan R. Elenberg, Rajat Sen, Guy W. Cole, Sanjay Shakkottai, and Sinead A. Williamson. Importance weighted generative networks. In *Machine Learning and Knowledge Discovery in Databases*, pp. 249 to 265, 2019. URL https://ecmlpkdd-storage.s3.eu-central-1.amazonaws.com/former-websites/2019/downloads/paper/21.pdf.

Marthinus du Plessis, Gang Niu, and Masashi Sugiyama. Convex formulation for learning from positive and unlabeled data. In *ICML*, volume 37, pp. 1386 to 1394, 2015. URL https://proceedings.mlr.press/v37/plessis15.html.

Marthinus C. du Plessis, Gang Niu, and Masashi Sugiyama. Analysis of learning from positive and unlabeled data. In *Advances in Neural Information Processing Systems 27*, pp. 703 to 711, 2014. URL https://proceedings.neurips.cc/paper_files/paper/2014/hash/f032bc3f1eb547f716df87edb523b8f0-Abstract.html.

Charles Elkan and Keith Noto. Learning classifiers from only positive and unlabeled data. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 213 to 220, New York, NY, USA, 2008. doi: 10.1145/1401890.1401920.

Ehsan Farahbakhsh, Jack Maughan, and R. Dietmar Müller. Prospectivity modelling of critical mineral deposits using a generative adversarial network with oversampling and positive-unlabelled bagging. *Ore Geology Reviews*, 162:105665, 2023. doi: 10.1016/j.oregeorev.2023.105665.

Jhanvi Garg, Xianyang Zhang, and Quan Zhou. Soft-constrained Schrödinger bridge: a stochastic control approach. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 4429 to 4437, 2024. URL https://proceedings.mlr.press/v238/garg24a.html.

Chen Gong, Qizhou Wang, Tongliang Liu, Bo Han, Jane You, Jian Yang, and Dacheng Tao. Instance-dependent positive and unlabeled learning with labeling bias estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(8):4163 to 4177, 2022. doi: 10.1109/TPAMI.2021.3061456.

Aditya Grover, Jiaming Song, Alekh Agarwal, Kenneth Tran, Ashish Kapoor, Eric Horvitz, and Stefano Ermon. Bias correction of learned generative models using likelihood-free importance weighting. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/hash/d76d8deea9c19cc9aaf2237d2bf2f785-Abstract.html.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770 to 778, 2016. doi: 10.1109/CVPR.2016.90.

Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. GANs trained by a two time-scale update rule converge to a local nash equilibrium. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a1d694707eb0fefe65871369074926d-Abstract.html.

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

Ming Hou, Brahim Chaib-draa, Chao Li, and Qibin Zhao. Generative adversarial positive-unlabelled learning. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pp. 2255 to 2261, 2018. doi: 10.24963/ijcai.2018/312.

Shantanu Jain, Martha White, and Predrag Radivojac. Recovering true classifier performance in positive-unlabeled learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, pp. 2066 to 2072, 2017. doi: 10.1609/aaai.v31i1.10937.

Masahiro Kato, Takeshi Teshima, and Junya Honda. Learning from positive and unlabeled data with a selection bias. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=rJzLciCqKm.

Masahiro Kato, Shuting Wu, Kodai Kureishi, and Shota Yasui. Automatic debiased learning from positive, unlabeled, and exposure data. *arXiv preprint arXiv:2303.04797*, 2023. URL https://arxiv.org/abs/2303.04797.

Dongjun Kim, Yeongmin Kim, Se Jung Kwon, Wanmo Kang, and Il-Chul Moon. Refining generative process with discriminator guidance in score-based diffusion models. In *ICML*, volume 202, pp. 16567 to 16598, 2023. URL https://proceedings.mlr.press/v202/kim23i.html.

Yeongmin Kim, Byeonghu Na, Minsang Park, JoonHo Jang, Dongjun Kim, Wanmo Kang, and Il chul Moon. Training unbiased diffusion models from biased dataset. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=39cPKijBed.

Ryuichi Kiryo, Gang Niu, Marthinus C. du Plessis, and Masashi Sugiyama. Positive-unlabeled learning with non-negative risk estimator. In *Advances in Neural Information Processing Systems 30*, pp. 1675 to 1685, 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/hash/7cce53cf90577442771720a370c3c723-Abstract.html.

Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In *Proceedings of the IEEE International Conference on Computer Vision*, 2015. URL https://mmlab.ie.cuhk.edu.hk/projects/CelebA.html.

Cheng Lu, Huayu Chen, Jianfei Chen, Hang Su, Chongxuan Li, and Jun Zhu. Contrastive energy prediction for exact energy-guided diffusion sampling in offline reinforcement learning. In *ICML*, volume 202, 2023. URL https://proceedings.mlr.press/v202/lu23d.html.

Alex Luedtke and Kenji Fukumizu. Doublegen: Debiased generative modeling of counterfactuals. *arXiv preprint arXiv:2509.16842*, 2025. URL https://arxiv.org/abs/2509.16842v1.

Valentyn Melnychuk and Stefan Feuerriegel. GDR-learners: Orthogonal learning of generative models for potential outcomes. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=bbmcIaEmJG.

Aditya Krishna Menon and Cheng Soon Ong. Linking losses for density ratio and class-probability estimation. In *ICML*, volume 48, pp. 304 to 313, 2016. URL https://proceedings.mlr.press/v48/menon16.html.

Ivan Molotkov and Mykyta Artomov. Detecting biased validation of predictive models in the positive-unlabeled setting: disease gene prioritization case study. *Bioinformatics Advances*, 3(1):vbad128, 2023. doi: 10.1093/bioadv/vbad128.

Byeonghu Na, Hyemi Kim, Kyungwoo Song, Weonyoung Joo, Yoon-Yeong Kim, and Il-Chul Moon. Deep generative positive-unlabeled learning under selection bias. In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management*, pp. 1155 to 1164, 2020. doi: 10.1145/3340531.3411971.

Muhammad Ferjad Naeem, Seong Joon Oh, Youngjung Uh, Yunjey Choi, and Jaejun Yoo. Reliable fidelity and diversity metrics for generative models. In *ICML*, volume 119, pp. 7176 to 7185, 2020. URL https://proceedings.mlr.press/v119/naeem20a.html.

XuanLong Nguyen, Martin J. Wainwright, and Michael I. Jordan. Estimating divergence functionals and the likelihood ratio by convex risk minimization. *IEEE Transactions on Information Theory*, 56(11):5847 to 5861, 2010. URL https://arxiv.org/abs/0809.0853.

Kazusato Oko, Shunta Akiyama, and Taiji Suzuki. Diffusion models are minimax optimal distribution estimators. In *ICML*, volume 202, pp. 26517 to 26582, 2023. URL https://proceedings.mlr.press/v202/oko23a.html.

Aleš Papič, Igor Kononenko, and Zoran Bosnić. Conditional generative positive and unlabeled learning. *Expert Systems with Applications*, 224:120046, 2023. doi: 10.1016/j.eswa.2023.120046.

Harish Ramaswamy, Clayton Scott, and Ambuj Tewari. Mixture proportion estimation via kernel embeddings of distributions. In *ICML*, volume 48, pp. 2052 to 2060, 2016. URL https://proceedings.mlr.press/v48/ramaswamy16.html.

Matt Raymond, Yilun Zhu, Jianxin Zhang, Angela Violi, and Clayton Scott. Joint diffusion sampling via positive-unlabeled guidance for multi-modal data. In *Proceedings of the ICML 2025 Workshop on Multi-modal Foundation Models and Large Language Models for Life Sciences*, July 2025. URL https://openreview.net/forum?id=1o89PN84NH.

Wojciech Rejchel, Paweł Teisseyre, and Jan Mielniczuk. Joint empirical risk minimization for instance-dependent positive-unlabeled data. *Knowledge-Based Systems*, 304:112444, 2024. doi: 10.1016/j.knosys.2024.112444.

Paul R. Rosenbaum and Donald B. Rubin. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society: Series B (Methodological)*, 45(2):212 to 218, 1983. doi: 10.1111/j.2517-6161.1983.tb01242.x.

Donald B. Rubin. Inference and missing data. *Biometrika*, 63(3):581 to 592, 1976. doi: 10.1093/biomet/63.3.581.

Yuta Saito, Shunsuke Aihara, Megumi Matsutani, and Yusuke Narita. Open bandit dataset and pipeline: Towards realistic and reproducible off-policy evaluation. In *NeurIPS Datasets and Benchmarks*, volume 1, 2021. URL https://arxiv.org/abs/2008.07146.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021a. URL https://arxiv.org/abs/2010.02502.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021b. URL https://openreview.net/forum?id=PxTIG12RRHS.

Hiroshi Takahashi, Tomoharu Iwata, Atsutoshi Kumagai, Yuuki Yamanaka, and Tomoya Yamashita. Positive-unlabeled diffusion models for preventing sensitive data generation. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=jKcZ4hF4s5.

Wei Wang, Dong-Dong Wu, Ming Li, Jingxiong Zhang, Gang Niu, and Masashi Sugiyama. Accessible, realistic, and fair evaluation of positive-unlabeled learning algorithms. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/pdf?id=5R11h5o44C.

Shiwei Xu and Margaret E. Ackerman. Leveraging permutation testing to assess confidence in positive-unlabeled learning applied to high-dimensional biological datasets. *BMC Bioinformatics*, 25:218, 2024. doi: 10.1186/s12859-024-05834-2.

A NOTATION

Uppercase letters denote random variables, and densities are taken with respect to a common dominating measure. The unlabeled sample is drawn from P, not conditioned on the absence of a label.

Symbol	Meaning
X, Y, S	Attributes, binary class and labeled positive indicator; $S = 1$ implies $A = 1$ and $Y = 1$.
L, U, V	Labeled positives, unlabeled population sample and selection records (X, A); classes are unavailable in U and V.
$A, v(x)$	Selection indicator and probability $\Pr(A = 1 \mid X = x)$.
f, f_+, f_-, f_l, f_A	Densities of the population, positives, negatives, labeled positives and selected objects.
$\alpha, e(x), c$	Class prior $\Pr(Y = 1)$, labeling propensity $\Pr(S = 1 \mid Y = 1, X = x)$ and label frequency $\Pr(S = 1 \mid Y = 1)$.
c_{rep}	Reporting probability for a selected positive; $e = c_{\mathrm{rep}} v$ under Assumption 1.
$\mathbb{E}, \mathbb{E}_+, \mathbb{E}_l, \mathbb{E}_A$	Expectations under P, P_+, P_l, P_A.
$\mathbb{E}_{\mathrm{data}}$	Expectation over the samples used to fit an estimator, conditional on any explicitly fixed quantities.
r, r_l, r_A	Density ratios f_+/f, f_l/f, and f_l/f_A.
ω_i	Normalized inverse probability weight of the ith labeled positive in the IPW loss.
g, ℓ	Log density ratio and logistic margin loss $\ell(u) = \log(1 + e^{-u})$.
$a, a_{\max}, Q_a$	Fitted image weight, fixed before guidance; its truncation level; its normalized target $aP/\mathbb{E}a$.
$\widehat{r}_A, \widehat{Q}$	Estimated selected reference ratio and its normalized target $\widehat{r}_A P/\mathbb{E}\widehat{r}_A$.
w, Q_w, Z_w	Guidance strength, target $r^w P/Z_w$ and mass $Z_w = \mathbb{E}r^w$.
n_L, n_U, n_V	Numbers of labeled positives, unlabeled objects and selection records used for estimation.
$X_0, X_t, \boldsymbol{z}$	Clean image, noised image and independent standard Gaussian noise.
$\bar{\alpha}_t, \sigma_t, K_t$	Cumulative signal variance, noise standard deviation and transition kernel.
$p_t, p_{a,t}$	Densities after noising P and Q_a.
$\boldsymbol{s}_t, \boldsymbol{s}_{a,t}$	Their scores $\nabla \log p_t$ and $\nabla \log p_{a,t}$.
ϕ_a	Log conditional weight $\log \mathbb{E}[a(X_0) \mid X_t = x]$.

Constants and auxiliary functions used only within a proof are defined locally.

B SELECTION CORRECTION AND DISTRIBUTION RECOVERY

This appendix proves the identification and recovery results and examines inverse probability weighting, truncation and variable reporting. Throughout, $P = \alpha P_+ + (1 - \alpha)P_-$, $0 < \alpha < 1$, and $r = \mathrm{d}P_+/\mathrm{d}P$. A positive is labeled with probability $e(x) = \Pr(S = 1 \mid X = x, Y = 1)$, so

$$\mathrm{d}P_l = \frac{e}{c}\mathrm{d}P_+, \qquad r_l = \frac{\mathrm{d}P_l}{\mathrm{d}P} = \frac{re}{c}, \qquad c = \mathbb{E}_+ e. \tag{B.1}$$

The expectations $\mathbb{E}$, $\mathbb{E}_+$ and $\mathbb{E}_l$ refer to P, P_+ and P_l. For a fixed weighting strength, define

$$Q_w = \frac{r^w}{Z_w}P, \qquad Z_w = \mathbb{E}r^w, \qquad w > 0. \tag{B.2}$$

Here $Q_1 = P_+$ and $0 < Z_w < \infty$ because $r \le \alpha^{-1}$. All logarithms are natural, and $\ell(z) = \log(1 + e^{-z})$.

B.1 IDENTIFICATION FROM POSITIVE AND UNLABELED SAMPLES

Here the observations are independent samples from P_l and P of fixed sizes; neither the prior α nor the label frequency c is observed. Under constant e, the *selected completely at random* (SCAR) assumption, $P_l = P_+$. General positive selection need not identify P_+.

Proposition B.1 (Observational equivalence). *Assume $M := \operatorname{ess\,sup}_P r_l < \infty$. Every such observed pair (P, P_l) has a SCAR representation: take $P'_+ = P_l$, choose $0 < \alpha' < \min\{1, 1/M\}$, and put*

$$P'_- = \frac{P - \alpha' P_l}{1 - \alpha'}, \qquad e'(x) \equiv c', \quad 0 < c' \le 1.$$

For $0 < \delta_0 < 1$, any test of size at most δ_0 uniformly over these SCAR models has power at most δ_0 at every selected model with this representation. Selection records change the observed data; Proposition B.6 shows when they restore identification.

The ratio r is not identified: two models can have the same observed distributions and different r, so no estimator is consistent throughout the unrestricted selection model.

Proof. The measure $P - \alpha' P_l$ is nonnegative because $\alpha' r_l \le 1$ and has mass $1 - \alpha'$. Constant labeling preserves P'_+, leaving both observed distributions unchanged. Every statistic thus has the same distribution under the original model and its SCAR representation, where the size condition bounds rejection probability.

For nonidentification, take $P = (1/2, 1/2)$ on two states and $\alpha = 1/2$ in both models. Set $P_+ = P$, $e = (1/2, 1/2)$ in the first, and $P_+ = (1/4, 3/4)$, $e = (1, 1/3)$ in the second. Both have $c = 1/2$, $P_l = P$, and valid negative class distributions, but their ratios are $(1, 1)$ and $(1/2, 3/2)$. Their labeling frequencies $\alpha c = 1/4$ also agree. By the triangle inequality, any estimator's larger expected $L^1(P)$ error is at least half the $L^1(P)$ distance between these ratios. Under the identical data distribution, disjoint neighborhoods of the two ratios cannot both have probability tending to one, ruling out consistency throughout the class. □

Proposition B.2 (Ranking and generation targets). *Suppose $e = br^\gamma$ on $\{r > 0\}$, where $\gamma \ge 0$ and $b > 0$ makes e a valid propensity. Then*

$$r_l = \frac{r^{1+\gamma}}{Z_{1+\gamma}}, \qquad \frac{r_l^w}{\mathbb{E} r_l^w} P = Q_{w(1+\gamma)}. \tag{B.3}$$

The ratios r_l and r have the same ranking, including ties. Thus their population AUCs agree on the same evaluation distribution. For $w > 0$ and $\gamma > 0$, their generation targets agree if and only if r is constant P almost surely on $\{r > 0\}$.

Proof. Since $c = b\mathbb{E} r^{1+\gamma}$, substitution into (B.1) gives both identities. The map $u \mapsto u^{1+\gamma}$ is strictly increasing, which preserves both comparisons and ties in AUC. Equality of the target densities is equivalent to constancy of $r^{w\gamma}$ on $\{r > 0\}$. □

This ordering invariance (Kato et al., 2019) does not recover probability mass within the positive class. With deterministic membership and positive region $\mathcal{X}_+ = \{x : Y(x) = 1\}$, $r = \alpha^{-1} \mathbf{1}_{\mathcal{X}_+}$ and $Q_w = P_+$ for every $w > 0$. For general $e > 0$ on $\mathcal{X}_+$, r and r_l both have population AUC one, yet the uncorrected target is $e^w P_+ / \mathbb{E}_+ e^w$; since exact distributions identify $P_+ = P(\cdot \mid r_l > 0)$ here, the distortion comes from uncorrected weighting, not from nonidentification.

B.2 Inverse probability weighting with known propensities

Let $\mathrm{d}P_l = (e/c)\,\mathrm{d}P_+$, with $e > 0$ almost surely under P_+ and $c = \mathbb{E}_+ e > 0$. For $W = 1/e(X)$, $X \sim P_l$, we have $\mathbb{E}_l W = 1/c$; assume $\mathbb{E}_l W^2 < \infty$. Given independent observations $X_i^L \sim P_l$ and $X_j^U \sim P$, define

$$\widehat{\mathcal{L}}_e(g) = \frac{\sum_{i=1}^{n_L} W_i \ell(g(X_i^L))}{2 \sum_{i=1}^{n_L} W_i} + \frac{1}{2n_U} \sum_{j=1}^{n_U} \ell(-g(X_j^U)), \qquad \ell(z) = \log(1 + e^{-z}).$$

Write $\omega_i = W_i / \sum_j W_j$ for the normalized positive weights. Under constant reporting, $e = c_{\mathrm{rep}} v$, so normalizing $1/e$ or $1/v$ gives the same ω_i. The estimator $\widehat{g}_e$ minimizes $\widehat{\mathcal{L}}_e(g) + \lambda \|g\|_{\mathcal{H}}^2 / 2$ in a reproducing kernel Hilbert space $\mathcal{H}$ with $k(x, x) \le \kappa^2$. Its population risk and approximation error are

$$\mathcal{L}_+(g) = \tfrac{1}{2}\mathbb{E}_+ \ell(g) + \tfrac{1}{2}\mathbb{E} \ell(-g), \qquad \Delta_+(g) = \mathcal{L}_+(g) - \mathcal{L}_+^*,$$
$$\mathcal{E}(\lambda) = \inf_{g \in \mathcal{H}} \{\Delta_+(g) + \lambda \|g\|_{\mathcal{H}}^2 / 2\}.$$

Here $\mathcal{L}_+^* = \inf_g \mathcal{L}_+(g)$ is the unrestricted optimum.

Proposition B.3 (Correction and effective sample size)**.** *For every integrable ψ, $\frac{\mathbb{E}_l[W\psi]}{\mathbb{E}_l W} = \mathbb{E}_+\psi$. The corrected population risk has Bayes logit $\log r = \log(\mathrm{d}P_+/\mathrm{d}P)$, with the extended value when $r = 0$. Multiplying all weights by a positive constant leaves the empirical risk unchanged. Define the population effective sample size and its empirical counterpart by*

$$n_{\mathrm{eff}} = n_L \frac{(\mathbb{E}_l W)^2}{\mathbb{E}_l W^2}, \qquad \widehat{n}_{\mathrm{eff}} = \frac{(\sum_i W_i)^2}{\sum_i W_i^2}.$$

Then $n_{\mathrm{eff}} \le n_L$, with equality exactly when e is constant P_l almost surely. Also $1 \le \widehat{n}_{\mathrm{eff}} \le n_L$, with upper equality exactly when all sample weights are equal. For constant e, the corrected empirical objective and its unique regularized minimizer equal their uncorrected counterparts for every sample. If $\mathbb{E}_l W^2 < \infty$, then $\widehat{n}_{\mathrm{eff}}/n_{\mathrm{eff}} \to 1$ almost surely.

Proof. Change of measure gives $\mathbb{E}_l[W\psi] = \mathbb{E}_+\psi/c$ and $\mathbb{E}_l W = 1/c$; pointwise minimization gives $e^{g^*} = r$. The population inequality $(\mathbb{E}_l W)^2 \le \mathbb{E}_l W^2$ is an equality exactly when W has zero variance. Cauchy Schwarz gives the sample bounds and upper equality condition. Constant weights normalize to $1/n_L$, and strong convexity identifies the minimizers. The final claim follows from almost sure convergence of the sample means of W and W^2. □

Theorem B.4 (Risk under self normalized importance weighting)**.** *For $n_L \ge 2$, $\lambda > 0$, and $\mathbb{E}_l W^2 < \infty$, the corrected estimator obeys*

$$\mathbb{E}_{\mathrm{data}} \Delta_+(\widehat{g}_e) \le \mathcal{E}(\lambda) + \frac{20\kappa^2/\lambda + 56\log 2}{n_{\mathrm{eff}}} + \frac{\kappa^2}{2\lambda n_U}. \tag{B.4}$$

The expectation covers both independent samples, and n_{eff} is a population quantity.

Proof. Normalize $\widetilde{W} = W/(\mathbb{E}_l W)$, so $\mathbb{E}_l \widetilde{W} = 1$ and $\mathbb{E}_l \widetilde{W}^2 = n_L/n_{\mathrm{eff}}$. Conditional on the unlabeled sample D, define

$$\varphi_D(g, x) = \tfrac{1}{2}\ell(g(x)) + \tfrac{1}{2}\widehat{\mathbb{E}}_{U,D}\ell(-g) + \tfrac{\lambda}{2}\|g\|_{\mathcal{H}}^2.$$

Here $\widehat{\mathbb{E}}_{U,D}$ is the empirical mean over D. Multiplying the objective by $\overline{\widetilde{W}} = n_L^{-1}\sum_i \widetilde{W}_i$ leaves its minimizer unchanged and gives $n_L^{-1}\sum_i \widetilde{W}_i \varphi_D(g, X_i^L)$. In the original objective, each loss term equals $\log 2/2$ at zero. Nonnegativity gives $\|\widehat{g}_e\|_{\mathcal{H}} \le R_\lambda := \sqrt{2\log 2/\lambda}$. On this ball, φ_D is L_λ Lipschitz with range $[0, B_\lambda]$, where $L_\lambda = \kappa + \lambda R_\lambda$ and $B_\lambda = 2\log 2 + \kappa R_\lambda$.

Replace labeled observation i by an independent copy with weight $\widetilde{W}_i'$, and denote the minimizers by $\widehat{g}$ and $\widehat{g}^{(i)}$. The weighted objective is $\lambda\overline{\widetilde{W}}$ strongly convex, with modulus at least $\lambda/4$ on $\mathcal{E}_i = \{\sum_{j\ne i}\widetilde{W}_j \ge (n_L - 1)/2\}$. Comparing gradients gives

$$\left\|\widehat{g} - \widehat{g}^{(i)}\right\|_{\mathcal{H}} \le \frac{4L_\lambda(\widetilde{W}_i + \widetilde{W}_i')}{n_L\lambda}.$$

The replacement identity gives conditional expected generalization error

$$\frac{1}{n_L}\sum_i \mathbb{E}_{\mathrm{data}}\left[\widetilde{W}_i\{\varphi_D(\widehat{g}^{(i)}, X_i^L) - \varphi_D(\widehat{g}, X_i^L)\} \mid D\right].$$

On $\mathcal{E}_i$, the expected summand is at most $8L_\lambda^2\mathbb{E}_l\widetilde{W}^2/(n_L\lambda)$, since $\mathbb{E}_{\mathrm{data}}[\widetilde{W}_i(\widetilde{W}_i + \widetilde{W}_i')] = \mathbb{E}_l\widetilde{W}^2 + 1 \le 2\mathbb{E}_l\widetilde{W}^2$. On $\mathcal{E}_i^c$, use the range B_λ, independence of $\widetilde{W}_i$ and $\mathcal{E}_i$, and Chebyshev's inequality:

$$\Pr(\mathcal{E}_i^c) \le \frac{4\,\mathrm{Var}_l(\widetilde{W})}{n_L - 1} \le \frac{8\mathbb{E}_l\widetilde{W}^2}{n_L}.$$

The conditional generalization error is at most $(8L_\lambda^2/\lambda + 8B_\lambda)/n_{\mathrm{eff}}$. For a fixed comparator, the multiplied objective is unbiased conditional on D. Empirical optimality transfers this bound to the regularized excess risk with population positive and empirical unlabeled expectations.

Conditional on the labeled sample, replacing one unlabeled observation gives the regularized logistic stability bound $\kappa^2/(2\lambda n_U)$, including the loss coefficient $1/2$. For any fixed comparator g, the bound is $\Delta_+(g) + \lambda \|g\|^2/2$ plus these two generalization terms. Take the infimum over g and use

$$\frac{L_\lambda^2}{\lambda} \le \frac{2\kappa^2}{\lambda} + 4\log 2, \qquad B_\lambda \le 3\log 2 + \frac{\kappa^2}{2\lambda}.$$

This gives (B.4). □

Positivity does not imply a finite second inverse moment. For P_+ uniform on $(0,1)$ and $e(x) = x^2$, we have $f_l(x) = 3x^2$, $\mathbb{E}_l W = 3$, but $\mathbb{E}_l W^2 = \infty$, so the risk bound does not apply however large the realized effective count is. For a fixed positive estimate $\widehat{e}$ with $0 < \mathbb{E}_+(e/\widehat{e}) < \infty$, the population weighted distribution is proportional to $(e/\widehat{e})P_+$. A misspecified probability model can therefore retain bias as sampling error vanishes.

B.3 From logistic loss to weight error

Following the use of Hellinger losses in density ratio estimation (Nguyen et al., 2010), we bound weight error by balanced logistic regret for a fixed positive power.

Proposition B.5 (Logistic regret and powered weights)**.** *Let $r = \mathrm{d}P_+/\mathrm{d}P$, let g be any real measurable logit, and write $\Delta_+(g) = \mathcal{L}_+(g) - \mathcal{L}_+^*$ for the balanced logistic excess risk defined above. For $B > 0$, put*

$$D^* = \frac{r}{1+r}, \qquad D_g = (1+e^{-g})^{-1}, \qquad H_B = 2(1+B)^3,$$
$$b = \min\{e^g, B\}, \qquad r_B = \min\{r, B\}.$$

For every $w > 0$, with $\nu_w = \min\{w, 1/2\}$,

$$\mathbb{E}(\sqrt{b} - \sqrt{r_B})^2 \le H_B \Delta_+(g), \tag{B.5}$$
$$\mathbb{E}|b^w - r^w| \le C_{w,B}^{\mathrm{H}} \Delta_+(g)^{\nu_w} + \mathbb{E}(r^w - B^w)_+, \tag{B.6}$$

where

$$C_{w,B}^{\mathrm{H}} = \begin{cases} H_B^w, & 0 < w \le 1/2, \\ 2wB^{w-1/2}\sqrt{H_B}, & w > 1/2. \end{cases}$$

The remainder vanishes when $B \ge \|r\|_\infty$; no lower bound on r is needed.

Proof. Let $\overline{P} = (P_+ + P)/2$. The regret identity is

$$\Delta_+(g) = \mathbb{E}_{\overline{P}} \mathrm{KL}\{\mathrm{Bern}(D^*) \| \mathrm{Bern}(D_g)\}.$$

For Bernoulli distributions, use $H^2(p,q) = (\sqrt{p} - \sqrt{q})^2 + (\sqrt{1-p} - \sqrt{1-q})^2$. Then $H^2(p,q) \le \mathrm{KL}\{\mathrm{Bern}(p) \| \mathrm{Bern}(q)\}$. Capping both probabilities at $p_B = B/(1+B)$ cannot increase either square root difference. On $[0, \sqrt{p_B}]$, the map $F(u) = u/\sqrt{1-u^2}$ has derivative $F'(u) = (1-u^2)^{-3/2} \le (1+B)^{3/2}$, so

$$|\sqrt{b} - \sqrt{r_B}|^2 \le (1+B)^3 H^2(D_g, D^*).$$

Since $\mathrm{d}P/\mathrm{d}\overline{P} \le 2$, integration proves (B.5).

For $0 < w \le 1/2$, $u \mapsto u^{2w}$ is Hölder with constant one on $[0,\infty)$. Jensen's inequality gives

$$\mathbb{E}|b^w - r_B^w| \le \mathbb{E}|\sqrt{b} - \sqrt{r_B}|^{2w} \le \{\mathbb{E}(\sqrt{b} - \sqrt{r_B})^2\}^w.$$

For $w > 1/2$, its Lipschitz constant on $[0, \sqrt{B}]$ is $2wB^{w-1/2}$; Cauchy Schwarz gives the bound. Add $\mathbb{E}|r_B^w - r^w| = \mathbb{E}(r^w - B^w)_+$ to obtain (B.6). □

B.4 Recovery from selection records

Let $A = 1$ denote selection for verification and $S = 1$ a reported positive label. Write

$$v(x) = \Pr(A = 1 \mid X = x), \qquad v_+(x) = \Pr(A = 1 \mid X = x, Y = 1),$$
$$c_{\mathrm{rep}}(x) = \Pr(S = 1 \mid A = 1, Y = 1, X = x),$$

Only selected positives can have $S = 1$, so $e(x) = v_+(x)c_{\mathrm{rep}}(x)$.

Proposition B.6 (What selection records identify)**.** *Independent records (X, A) with $X \sim P$ identify v almost surely under P, but additional assumptions are needed to identify v_+. A sufficient condition is $A \perp Y \mid X$, which gives $v_+ = v$ on the positive support. Under this condition, $v > 0$ on that support, and unknown constant reporting $c_{\mathrm{rep}}(x) \equiv c_{\mathrm{rep}} > 0$, every integrable ψ satisfies*

$$\frac{\mathbb{E}_l[\psi(X)/v(X)]}{\mathbb{E}_l[1/v(X)]} = \mathbb{E}_+\psi(X). \tag{B.7}$$

The constant c_{rep} cancels. Estimating v is binary regression with response A; the records contain no class labels Y.

Proof. The records identify $\mathbb{E}[A \mid X] = v$; conditional independence gives $v_+ = v$. Under constant reporting, $\mathrm{d}P_l = c_{\mathrm{rep}} v \mathrm{d}P_+ / c$; dividing by v and normalizing proves (B.7).

Without conditional independence, let X be uniform on two states with $\Pr(Y = 1 \mid X) = 1/2$, and use selection probabilities $(1/5, 4/5)$ for positives and $(4/5, 1/5)$ for negatives. Then $v \equiv 1/2$, but v_+ is nonconstant, as is the labeling propensity under complete reporting. A constant observed selection rate therefore does not establish SCAR labeling. □

B.4.1 Density ratio estimation with selected observations

This proposition gives the identification formula in Section 3.1 and proves Theorem 2.

Proposition B.7 (Selected reference identification and target error)**.** *Let selection records have marginal P and $v(x) = \Pr(A = 1 \mid X = x)$. Assume $A \perp\!\!\!\perp Y \mid X$, constant positive reporting with probability $c_{\mathrm{rep}} > 0$, the same selection probability for L and the records, and $v > 0$ almost surely under P. The selected distribution $P_A = P(\cdot \mid A = 1)$, with density f_A, satisfies*

$$r_A := \frac{\mathrm{d}P_l}{\mathrm{d}P_A} = \frac{\mathbb{E}v}{\mathbb{E}_+ v} r, \qquad \mathbb{E}r_A = \frac{\mathbb{E}v}{\mathbb{E}_+ v}, \qquad P_+ = \frac{r_A}{\mathbb{E}r_A} P. \tag{B.8}$$

For any nonnegative measurable estimate $\widehat{r}_A$ with $0 < \mathbb{E}\widehat{r}_A < \infty$, let $\widehat{Q} = (\widehat{r}_A / \mathbb{E}\widehat{r}_A)P$. Then

$$\mathrm{TV}(\widehat{Q}, P_+) \le \frac{\mathbb{E}|\widehat{r}_A - r_A|}{\mathbb{E}r_A}. \tag{B.9}$$

If additionally $v \ge v_{\min} > 0$ almost surely under P and $0 \le r_A, \widehat{r}_A \le B$ for a finite B, define

$$\mathcal{L}_A(g) = \tfrac{1}{2}\mathbb{E}_l \ell(g) + \tfrac{1}{2}\mathbb{E}_A \ell(-g), \qquad \Delta = \mathcal{L}_A(\log \widehat{r}_A) - \inf_g \mathcal{L}_A(g),$$

where $\ell(z) = \log(1 + e^{-z})$, with the limiting convention at zero. For $J_B = 2(1 + B)$,

$$\mathrm{TV}(\widehat{Q}, P_+) \le 2\sqrt{\frac{(\mathbb{E}_+ v) J_B \Delta}{v_{\min}}}. \tag{B.10}$$

The infimum is the unrestricted population optimum.

Proof. The observation assumptions give $\mathrm{d}P_l = (v/\mathbb{E}_+ v)\mathrm{d}P_+$ and $\mathrm{d}P_A = (v/\mathbb{E}v)\mathrm{d}P$. Positivity of v almost surely under P makes P and P_A equivalent. Their ratio gives (B.8). Adding and subtracting $\widehat{r}_A / \mathbb{E}r_A$ gives

$$2\,\mathrm{TV}(\widehat{Q}, P_+) \le \frac{|\mathbb{E}\widehat{r}_A - \mathbb{E}r_A| + \mathbb{E}|\widehat{r}_A - r_A|}{\mathbb{E}r_A} \le \frac{2\mathbb{E}|\widehat{r}_A - r_A|}{\mathbb{E}r_A}.$$

For $0 < t, s \le B$, the pointwise logistic regret is

$$d(t,s) = \tfrac{1}{2}\left[t\log(t/s) + (1+t)\log\frac{1+s}{1+t}\right] = \tfrac{1}{2}\int_t^s \frac{z-t}{z(1+z)}\,\mathrm{d}z \ge \frac{s-t+t\log(t/s)}{2(1+B)}.$$

The inequality holds in either integration direction. The numerator is at least $(\sqrt{s}-\sqrt{t})^2$, since their difference equals $2t\{\sqrt{s/t}-1-\log\sqrt{s/t}\} \ge 0$. At $t=0$, $d(0,s) = \frac{1}{2}\log(1+s)$ and $d(0,0)=0$; $s=0<t$ gives infinite regret. The inequalities extend to these endpoints. Integrating $d(r_A, \widehat{r}_A)$ under P_A gives Δ, hence $\mathbb{E}_A(\sqrt{\widehat{r}_A}-\sqrt{r_A})^2 \le J_B\Delta$. Because $\mathrm{d}P/\mathrm{d}P_A = (\mathbb{E}v)/v \le (\mathbb{E}v)/v_{\min}$,

$$\|\sqrt{\widehat{r}_A}-\sqrt{r_A}\|_{L^2(P)} \le \sqrt{\frac{(\mathbb{E}v)J_B\Delta}{v_{\min}}}.$$

The triangle and reverse triangle inequalities give $\|x/\|x\| - y/\|y\|\| \le 2\|x-y\|/\|y\|$ for nonzero Hilbert vectors. Apply this to $x=\sqrt{\widehat{r}_A}$ and $y=\sqrt{r_A}$ in $L^2(P)$. Cauchy Schwarz bounds TV by the distance between normalized square roots, so

$$\mathrm{TV}(\widehat{Q}, P_+) \le \frac{2}{\sqrt{\mathbb{E}r_A}}\sqrt{\frac{(\mathbb{E}v)J_B\Delta}{v_{\min}}} = 2\sqrt{\frac{(\mathbb{E}_+v)J_B\Delta}{v_{\min}}}. \qquad \square$$

Corollary B.8 (Recovery using a representative selected sample)**.** *Under the observation assumptions of Proposition B.7, the observed distributions P_l, P and P_A identify P_+. A sample $V_1 \sim P_A$ represents all selected objects regardless of whether they receive a positive report; neither U nor V_1 supplies class labels. Records with $A=0$, the selection rate $\mathbb{E}v$, and the population frequency of positive labels are not required.*

Proof. L and V_1 identify r_A under P_A, hence under the equivalent measure P. Equation (B.8) identifies P_+ by normalization, canceling $\mathbb{E}v/\mathbb{E}_+v$. The independent U estimates the fitted ratio's normalizer. $\square$

Selection logs and observation assumptions. Conditional independence at selection parallels missing at random (Rubin, 1976), and selective reporting motivates sensitivity analysis (Rosenbaum & Rubin, 1983). Logs of selection beyond positive outcomes exist in practice: the Open Bandit Dataset records displayed items, contexts, clicks and display probabilities (Saito et al., 2021), although its nonclicks are not verified negatives. Our model withholds verdicts from V rather than assuming that negatives are never verified, so V does not identify c_{rep}. Repeated verification with a latent history would require dependence assumptions beyond independent sampling.

Distribution recovery and posterior identification. Identifying r need not identify the posterior $\Pr(Y=1 \mid X) = \alpha r$ (Ramaswamy et al., 2016): any $0<\alpha<1$ with $\alpha \operatorname{ess\,sup}_P r \le 1$ gives a valid $P_- = (P-\alpha P_+)/(1-\alpha)$ and the same observations, but different posteriors and Bayes classifiers. Normalization removes this unknown scale for generation. If V_1 comes from n_V independent records, its size is $\mathrm{Binomial}(n_V, \mathbb{E}v)$ and, given the size, its objects follow P_A; a rate for Δ further depends on the function class and the optimization. With a sample source prior $\pi \in (0,1)$, the loss $\pi\mathbb{E}_l\ell(g) + (1-\pi)\mathbb{E}_A\ell(-g)$ has optimum $g = \log r_A + \log\{\pi/(1-\pi)\}$; normalization removes the offset, and our regret bound uses $\pi = 1/2$.

Corollary B.9 (Clipping and generator error)**.** *Under the premises of* (B.10)*, let $\widehat{Z}_A > 0$ be the mean of $\widehat{r}_A$ over U, and let $a_{\max} > 0$ be the truncation level applied to $\widehat{r}_A/\widehat{Z}_A$. Here B bounds unnormalized ratios and $a_{\max}$ truncates normalized weights. For the fitted weight $a = \min\{\widehat{r}_A/\widehat{Z}_A, a_{\max}\}$, put*

$$\widehat{r} = \frac{\widehat{r}_A}{\mathbb{E}\widehat{r}_A}, \qquad K = \frac{a_{\max}\widehat{Z}_A}{\mathbb{E}\widehat{r}_A}, \qquad \tau = \mathbb{E}(\widehat{r}-K)_+, \qquad Q_a = \frac{\min(\widehat{r},K)}{\mathbb{E}\min(\widehat{r},K)}P.$$

For any generated distribution G,

$$\mathrm{TV}(G, P_+) \le \mathrm{TV}(G, Q_a) + \tau + 2\sqrt{\frac{(\mathbb{E}_+v)J_B\Delta}{v_{\min}}}. \tag{B.11}$$

Proof. Normalizing $\min(\widehat{r}_A/\widehat{Z}_A, a_{\max})$ or $\min(\widehat{r}, K)$ gives the same distribution. Since $\mathbb{E}\widehat{r} = 1$, the remaining mass is $1-\tau > 0$. For $\tau > 0$, the density $(\widehat{r}-K)_+/\tau$ defines a probability distribution R, and $\widehat{Q} = (1-\tau)Q_a + \tau R$. Hence $\mathrm{TV}(\widehat{Q}, Q_a) \le \tau$, also valid when $\tau = 0$. The triangle inequality and (B.10) prove (B.11). □

The first term collects the errors of the finite U, the base model, the potential and the sampler; it can differ between estimators that share a sampler. For example, $B = 16$, $v_{\min} = 0.1$ and $\mathbb{E}_+ v \le 1$ give a coefficient of at most $\sqrt{1360} \approx 36.88$ in (B.10), so the bound is below one only if $\Delta < 1/1360 \approx 7.35 \times 10^{-4}$. In the neural experiments, the empirical truncation does not establish B, and the held out loss is not Δ.

B.4.2 Sensitivity to positive reporting

Proposition B.10 (Sensitivity to reporting probabilities). *Assume $v_+ = v > 0$ on the positive support. Correcting selection probabilities identifies*

$$\mathrm{d}\widetilde{P} = \frac{v^{-1}\mathrm{d}P_l}{\mathbb{E}_l v^{-1}}, \qquad \widetilde{r} = \mathrm{d}\widetilde{P}/\mathrm{d}P.$$

For a candidate reporting function $0 < c_{\mathrm{rep}} \le 1$, assume $0 < \mathbb{E}_{\widetilde{P}} c_{\mathrm{rep}}^{-1} < \infty$ and $0 < \mathbb{E}_{Q_w^0} c_{\mathrm{rep}}^{-w} < \infty$, with Q_w^0 defined below. Set

$$\mathrm{d}P_+^{c_{\mathrm{rep}}} = \frac{c_{\mathrm{rep}}^{-1}\mathrm{d}\widetilde{P}}{\mathbb{E}_{\widetilde{P}} c_{\mathrm{rep}}^{-1}}, \tag{B.12}$$

$$Q_w^{c_{\mathrm{rep}}}(\mathrm{d}x) = \frac{c_{\mathrm{rep}}(x)^{-w} Q_w^0(\mathrm{d}x)}{\mathbb{E}_{Q_w^0} c_{\mathrm{rep}}^{-w}}, \qquad Q_w^0(\mathrm{d}x) = \frac{\widetilde{r}(x)^w P(\mathrm{d}x)}{\mathbb{E}\widetilde{r}^w}. \tag{B.13}$$

The sensitivity set consists of these distributions for all c_{rep} satisfying reporting constraints, mixture feasibility, and any known constraints on the prior or labeling frequency. If $0 < c_{\min} \le c_{\mathrm{rep}}(x) \le c_{\max} \le 1$ and $\Gamma = c_{\max}/c_{\min}$, then for any measurable event B with $p = Q_w^0(B)$,

$$\frac{p}{p + \Gamma^w(1-p)} \le Q_w^{c_{\mathrm{rep}}}(B) \le \frac{\Gamma^w p}{1 - p + \Gamma^w p}. \tag{B.14}$$

The bounds are sharp over the unrestricted reporting band; prior or mixture constraints can narrow the compatible set. The lower bound $c_{\min}$ ensures finite normalizers. Moreover,

$$\mathrm{TV}(Q_w^{c_{\mathrm{rep}}}, Q_w^0) \le \frac{\Gamma^{w/2} - 1}{\Gamma^{w/2} + 1}. \tag{B.15}$$

On a support of diameter D, the W_1 distance is at most D times the same bound.

Proof. Dividing P_l by v normalizes $c_{\mathrm{rep,true}} P_+$ for the actual reporting function. Its ratio to P is bounded by $1/(\alpha \mathbb{E}_+ c_{\mathrm{rep,true}})$, ensuring Q_w^0 exists. Dividing by a candidate c_{rep}, taking powers and normalizing gives (B.12) and (B.13). Extremal weights $c_{\max}^{-w}, c_{\min}^{-w}$ on B, B^c give (B.14). Maximizing either deviation from p gives $(\sqrt{\Gamma^w} - 1)/(\sqrt{\Gamma^w} + 1)$; bounded support then gives the Wasserstein bound by coupling. □

C Weighted targets and diffusion sampling

Noise and score conventions. For $X_0 \sim P$, write $X_t = \sqrt{\bar{\alpha}_t} X_0 + \sigma_t \boldsymbol{z}$, $\boldsymbol{z} \sim \mathcal{N}(0, I)$, using the DDPM schedule in Appendix D.2. Negative predicted noise divided by σ_t estimates the score $\nabla \log p_t$; guidance adds $\nabla_x \phi_\psi(t, x)$.

C.1 Stagewise estimation and generation

Algorithms 1, 2 and 3 specify the three stages. The base estimates P from U or unconditional pretraining. Use $\ell(u) = \log(1 + e^{-u})$ and nonempty L, U and $V_1 = \{X : (X, A) \in V,\ A = 1\}$. Appendix D specifies training and sampling settings.

Algorithm 3 uses every noise level without clipping the predicted clean image. Time spacing and clipping for each dataset are specified in Appendix D.2.

Algorithm 1 Selection corrected weight estimation

Input: Labeled positives L, unlabeled images U, selection records $V = (X, A)$; batch size b, truncation level $a_{\max} > 0$, iteration count K_g.
Output: Fixed image weights $\{a(x) : x \in U\}$.

Form $V_1 \leftarrow \{X : (X, A) \in V,\ A = 1\}$; initialize log ratio g_θ.
for $k = 1, \ldots, K_g$ **do**
  Draw batches $\{x_i^L\}_{i=1}^b$ from L and $\{x_i^A\}_{i=1}^b$ from V_1, uniformly with replacement.
  Compute the balanced logistic loss:

$$\widehat{\mathcal{L}}_A(\theta) = \frac{1}{2b} \sum_{i=1}^{b} \big[\ell(g_\theta(x_i^L)) + \ell(-g_\theta(x_i^A))\big].$$

  Update θ by AdamW using $\nabla_\theta \widehat{\mathcal{L}}_A(\theta)$.
end for
Set $\widehat{r}_A(x) \leftarrow e^{g_\theta(x)}$ on U; normalize, then truncate:

$$a(x) = \min\left\{ \frac{\widehat{r}_A(x)}{|U|^{-1} \sum_{x' \in U} \widehat{r}_A(x')},\ a_{\max} \right\}, \qquad x \in U.$$

Algorithm 2 Guidance potential estimation

Input: Unlabeled images U, fixed weights a from Algorithm 1; noise schedule $(\bar{\alpha}_t, \sigma_t)_{t=1}^T$, batch size b, iteration count K_ϕ.
Output: Fitted guidance potential $\phi_\psi(t, x)$.

Initialize ϕ_ψ.
for $k = 1, \ldots, K_\phi$ **do**
  Draw b images x_0^i uniformly from U with replacement.
  Draw t_i uniformly from $\{1, \ldots, T\}$ and independent $\boldsymbol{z}_i \sim \mathcal{N}(0, I)$.
  Form $x_{t_i}^i \leftarrow \sqrt{\bar{\alpha}_{t_i}} x_0^i + \sigma_{t_i} \boldsymbol{z}_i$.
  Compute the weighted logistic loss using the same noisy images in both terms:

$$\widehat{\mathcal{L}}_a(\psi) = \frac{1}{2b} \sum_{i=1}^{b} \big[a(x_0^i)\ell(\phi_\psi(t_i, x_{t_i}^i)) + \ell(-\phi_\psi(t_i, x_{t_i}^i))\big].$$

  Update ψ by AdamW using $\nabla_\psi \widehat{\mathcal{L}}_a(\psi)$.
end for

Algorithm 3 Guided image generation

Input: Fixed base score $\widehat{\boldsymbol{s}}_t$, fitted potential ϕ_ψ from Algorithm 2; noise schedule $(\bar{\alpha}_t, \sigma_t)_{t=1}^T$, image count N_{gen}.
Output: N_{gen} generated images.

Draw N_{gen} initial states independently from $\mathcal{N}(0, I)$.
for $t = T, \ldots, 1$ **do**
  Evaluate $\boldsymbol{s}_{\text{guided}}(x, t) = \widehat{\boldsymbol{s}}_t(x) + \nabla_x \phi_\psi(t, x)$ at each current image x.
  Set $\beta_t = 1 - \bar{\alpha}_t / \bar{\alpha}_{t-1}$ and $\widetilde{\beta}_t = \beta_t(1 - \bar{\alpha}_{t-1})/(1 - \bar{\alpha}_t)$, with $\bar{\alpha}_0 = 1$.
  Draw independent $\boldsymbol{z} \sim \mathcal{N}(0, I)$ for each image and update

$$x \leftarrow \frac{x + \beta_t \boldsymbol{s}_{\text{guided}}(x, t)}{\sqrt{1 - \beta_t}} + \sqrt{\widetilde{\beta}_t} \boldsymbol{z}.$$

end for
return the final images, projected to $[-1, 1]$.

Conditional sampling analysis. The analysis below concerns a kernel estimator with projected sampling. For fixed $w > 0$, compare Q_a with $Q_w = r^w P/Z_w$ from (B.2); $Q_1 = P_+$. Assume P is supported in the known ball $B_{R_X} \subset \mathbb{R}^d$, $0 < R_X < \infty$; P need not have a density.

C.2 Conditional target and the two estimators

Let $a : \mathbb{R}^d \to [0, a_{\max}]$ be measurable, with $0 < a_{\max} < \infty$, fixed conditional on an independent weight estimation sample, and set

$$Z_a = \mathbb{E}a > 0, \qquad m_2 = \mathbb{E}a^2 \le a_{\max} Z_a, \qquad Q_a = \frac{aP}{Z_a}. \tag{C.1}$$

Let K_t be the Ornstein Uhlenbeck (OU) noising kernel defined by

$$\mathrm{d}X_t = -X_t\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}\boldsymbol{w}_t, \qquad \bar{\alpha}_t = e^{-2t}, \quad \sigma_t^2 = 1 - \bar{\alpha}_t, \quad \gamma_t = \sqrt{\bar{\alpha}_t}/\sigma_t^2. \tag{C.2}$$

Here $\boldsymbol{w}_t$ is standard d dimensional Brownian motion. Conditional expectations $\mathbb{E}_{Q,K_t}$ use $X_0 \sim Q$ and $X_t \mid X_0 \sim K_t(\cdot \mid X_0)$. For $t > 0$, write p_t for the density of $K_t P$ and $p_{a,t}$ for the density of $K_t Q_a$. Define

$$h_a(t,x) = \frac{K_t[aP](x)}{p_t(x)}, \qquad \phi_a(t,x) = \log h_a(t,x), \qquad \boldsymbol{s}_{a,t} = \nabla \log p_{a,t} = \boldsymbol{s}_t + \nabla\phi_a, \tag{C.3}$$

where $\boldsymbol{s}_t = \nabla \log p_t$. The positive Gaussian kernel and $Z_a > 0$ imply $h_a > 0$ everywhere, even where the clean weight vanishes. Also,

$$h_a \le a_{\max}, \qquad p_{a,t} = p_t h_a / Z_a \le (a_{\max}/Z_a) p_t. \tag{C.4}$$

Fix a fitting ball $\Omega = B_{R+1}$ with $R \ge 1$ and noise times $t_k > 0$. Use three mutually independent unlabeled samples, independent of the weight estimation sample: X_i^a, $1 \le i \le n_a$; X_i^0, $1 \le i \le n_0$; and X_i^s, $1 \le i \le n_s$. For each time and each sample, draw independent standard Gaussian noises and put $X_{i,k}^b = \sqrt{\bar{\alpha}_{t_k}} X_i^b + \sigma_{t_k} \boldsymbol{z}_{i,k}^b$, $b \in \{a, 0, s\}$. The superscript a denotes the weighted term, not positive class membership. Clean samples may be reused across noise times; the analysis requires independence only within each time level. We write $\mathbb{E}_{\mathrm{data}}$ for expectation over these three samples and their Gaussian noises, conditional on the weight estimation sample; unsubscripted $\mathbb{E}$ integrates under P.

Let $\mathcal{H}$ and $\mathcal{H}_s$ be restrictions to Ω of fixed full space Matérn RKHSs equivalent to $H^m(\mathbb{R}^d)$, where

$$m > d/2 + 2, \qquad M = \lceil m \rceil. \tag{C.5}$$

The kernels and length scales are fixed across R, sample sizes and time grids, so their evaluation bounds are independent of R:

$$|f(x)| \le \kappa_0 \|f\|_{\mathcal{H}}, \quad \|\nabla f(x)\| \le \kappa_1 \|f\|_{\mathcal{H}}, \quad \|\nabla^2 f(x)\|_{\mathrm{op}} \le \kappa_2 \|f\|_{\mathcal{H}}, \quad |v(x)| \le \kappa_s \|v\|_{\mathcal{H}_s}. \tag{C.6}$$

These bounds pass to restriction spaces by taking the infimum over global extensions. The vector space $\mathcal{H}_s^d$ has the sum of squared component norms.

Definition C.1 (Restricted weighted logistic estimator)**.** Let $\ell(u) = \log(1 + e^{-u})$. At each t_k, minimize

$$\widehat{\phi}_{a,k} = \operatorname*{arg\,min}_{f \in \mathcal{H}} \left\{ \frac{1}{2n_a} \sum_{i=1}^{n_a} \mathbf{1}_{\{X_{i,k}^a \in \Omega\}} a(X_i^a) \ell(f(X_{i,k}^a)) \right.$$
$$\left. + \frac{1}{2n_0} \sum_{i=1}^{n_0} \mathbf{1}_{\{X_{i,k}^0 \in \Omega\}} \ell(-f(X_{i,k}^0)) + \frac{\lambda_k}{2} \|f\|_{\mathcal{H}}^2 \right\}, \qquad \lambda_k > 0. \tag{C.7}$$

Denote the unregularized population risk by $\mathcal{R}_k$ and its regularized minimizer by $\phi_{a,\lambda,k}$. Its Bayes logit on Ω is $\phi_a(t_k, \cdot)$, since

$$\mathcal{R}_k(f) = \frac{1}{2} \int_\Omega p_{t_k}(x) \{ h_a(t_k, x) \ell(f(x)) + \ell(-f(x)) \}\,\mathrm{d}x.$$

The denominators are the original sample sizes, including observations outside Ω; the weighted term is not normalized by its weight sum.

Definition C.2 (Conditional mean kernel ridge estimator)**.** On the third unlabeled sample define

$$\widehat{m}_k = \operatorname*{arg\,min}_{v \in \mathcal{H}_s^d} \left\{ \frac{1}{2n_s} \sum_{i=1}^{n_s} \mathbf{1}_{\{X_{i,k}^s \in \Omega\}} \|v(X_{i,k}^s) - X_i^s\|^2 + \frac{\lambda_{s,k}}{2} \|v\|_{\mathcal{H}_s^d}^2 \right\}, \qquad \lambda_{s,k} > 0. \tag{C.8}$$

Set $\widehat{\boldsymbol{s}}_{t_k}(x) = \gamma_{t_k} \widehat{m}_k(x) - x/\sigma_{t_k}^2$. The population regression target is $m_t(x) = \mathbb{E}_{P,K_t}[X_0 \mid X_t = x]$.

Both objectives are strictly convex, with unique minimizers even when no observations fall in the fitting ball. By the representer theorem, each empirical minimizer is a finite sum of retained kernel sections; the same sum with the full space kernel defines its global extension. The following cutoff argument bounds restricted Sobolev norms without assuming global Sobolev membership of the population targets.

C.3 Gaussian posterior identities and source norms

Proposition C.3 (Regularity for a bounded terminal weight)**.** *Under* (C.1)*, $\phi_a(t,\cdot)$ is smooth on $\mathbb{R}^d$ for every $t > 0$. For $R' > 0$, set*

$$V_t(R') = |\log Z_a| + |\log a_{\max}| + \frac{2\sqrt{\bar{\alpha}_t} R_X R' + \bar{\alpha}_t R_X^2/2}{\sigma_t^2}.$$

Then, for $\|x\| \le R'$,

$$\log Z_a - \frac{2\sqrt{\bar{\alpha}_t} R_X \|x\| + \bar{\alpha}_t R_X^2/2}{\sigma_t^2} \le \phi_a(t,x) \le \log a_{\max}, \qquad |\phi_a(t,x)| \le V_t(R'). \tag{C.9}$$

Writing $m_{a,t}, \Sigma_{a,t}$ and m_t, Σ_t for the posterior means and covariance matrices under Q_a and P, respectively,

$$\begin{aligned} \nabla \phi_a &= \gamma_t (m_{a,t} - m_t), & \nabla^2 \phi_a &= \gamma_t^2 (\Sigma_{a,t} - \Sigma_t), \\ \|\nabla \phi_a\| &\le 2\gamma_t R_X, & \|\nabla^2 \phi_a\|_{\mathrm{op}} &\le 2\gamma_t^2 R_X^2. \end{aligned} \tag{C.10}$$

Also $\boldsymbol{s}_t = \gamma_t m_t - x/\sigma_t^2$ and $\boldsymbol{s}_{a,t} = \gamma_t m_{a,t} - x/\sigma_t^2$. For $j \ge 1$ define

$$C_j = \sum_{\mathcal{P} \text{ a partition of } \{1,\dots,j\}} (|\mathcal{P}| - 1)!.$$

Then $\sup_x |\partial^\beta \phi_a(t,x)| \le 2C_j(\gamma_t R_X)^j$ for every multi-index β with $|\beta| = j \ge 1$. Derivatives of order $j \ge 0$ of each coordinate of m_t are bounded by $C_{j+1}\gamma_t^j R_X^{j+1}$.

Proof. Expanding the Gaussian exponent expresses ϕ_a as the difference of the log Laplace transforms at $\theta = \gamma_t x$ of two finite nonzero measures:

$$e^{-\bar{\alpha}_t \|y\|^2/(2\sigma_t^2)} a(y) P(\mathrm{d}y), \qquad e^{-\bar{\alpha}_t \|y\|^2/(2\sigma_t^2)} P(\mathrm{d}y).$$

Compact support permits differentiation of every order, and the transforms are positive for real θ. For $y, y' \in B_{R_X}$,

$$\left| \|x - \sqrt{\bar{\alpha}_t} y\|^2 - \|x - \sqrt{\bar{\alpha}_t} y'\|^2 \right| \le 4\sqrt{\bar{\alpha}_t} R_X \|x\| + \bar{\alpha}_t R_X^2.$$

The smallest numerator and largest denominator kernels give the lower bound in (C.9); $a \le a_{\max}$ gives the upper bound.

The first two derivatives of a log Laplace transform are the tilted mean and covariance. Support in B_{R_X} bounds their norms by R_X and R_X^2, respectively. Differentiating the Gaussian convolution gives the score identities. For either tilted posterior ν, the joint cumulant of j coordinates is

$$\sum_{\mathcal{P}} (-1)^{|\mathcal{P}|-1} (|\mathcal{P}| - 1)! \prod_{J \in \mathcal{P}} \mathbb{E}_\nu \prod_{i \in J} X_{c_i}.$$

Each product of moments has absolute value at most R_X^j; the chain rule contributes γ_t^j, and the two transforms contribute a factor of two. A conditional mean coordinate is a first derivative of one transform, so its derivative bound uses cumulants of order $j + 1$. □

Lemma C.4 (Norm bounds for the restricted targets)**.** *For constants C_ϕ, C_m depending only on the dimension, Sobolev order and fixed kernels,*

$$\|\phi_a(t,\cdot)\|_{\mathcal{H}} \le C_\phi (R+2)^{d/2} \left\{ V_t(R+2) + \sum_{j=1}^{M} (\gamma_t R_X)^j \right\} =: \Phi_t, \tag{C.11}$$

$$\|m_t\|_{\mathcal{H}_s^d} \le C_m (R+2)^{d/2} R_X \sum_{j=0}^{M} (\gamma_t R_X)^j =: \Phi_{s,t}. \tag{C.12}$$

These bounds are polynomial in R, $1 + |\log Z_a| + |\log a_{\max}|$ and σ_t^{-2M}, with fixed R_X. The constants are independent of R, R_X, the noise time, the fitted weight and its truncation level, sample sizes, and the number of noise levels.

Proof. Choose a smooth cutoff $\eta_R = 1$ on B_{R+1} and zero outside B_{R+2}. A fixed transition in $\|x\| - R - 1$, outside B_1, bounds its derivatives through order M independently of $R \ge 1$. The product rule and Proposition C.3 bound the corresponding L^2 derivative norms of $\eta_R \phi_a$ and each coordinate of $\eta_R m_t$, with volume factor $O((R+2)^{d/2})$. Apply $\|f\|_{H^m} \le \|f\|_{H^M}$, equivalence of the fixed Matérn and Sobolev norms, and restriction to Ω. The polynomial bounds follow from $\gamma_t \le \sigma_t^{-2}$. □

C.4 Bounds on score estimation error

Proposition C.5 (Hilbert space concentration and recovery)**.** *Conditional on a, the regularized empirical logistic gradient at $\phi_{a,\lambda,k}$ has mean zero and satisfies*

$$\mathbb{E}_{\text{data}} \|\nabla \widehat{\mathcal{R}}_{\lambda,k}(\phi_{a,\lambda,k})\|_{\mathcal{H}}^2 \le \frac{\kappa_0^2}{4} \left(\frac{m_2}{n_a} + \frac{1}{n_0} \right),$$
$$\mathbb{E}_{\text{data}} \|\widehat{\phi}_{a,k} - \phi_{a,\lambda,k}\|_{\mathcal{H}}^2 \le \frac{\kappa_0^2}{4\lambda_k^2} \left(\frac{m_2}{n_a} + \frac{1}{n_0} \right). \tag{C.13}$$

If the Bayes target norm is bounded by Φ_{t_k}, then

$$\|\phi_{a,\lambda,k}\|_{\mathcal{H}} \le \Phi_{t_k}, \qquad \mathcal{R}_k(\phi_{a,\lambda,k}) - \mathcal{R}_k(\phi_a) \le \lambda_k \Phi_{t_k}^2/2. \tag{C.14}$$

Proof. At a fixed function, the weighted empirical gradient averages independent Hilbert valued vectors $\frac{1}{2} a(X_i^a) \ell'(f(X_{i,k}^a)) k_{X_{i,k}^a}$, taken to be zero outside Ω. Since $|\ell'| \le 1$, their second moments are bounded by $\kappa_0^2 m_2/4$, and those of the reference vectors by $\kappa_0^2/4$. Population optimality cancels the penalty in expectation. The two centered averages are independent; each has expected squared norm equal to the centered individual second moment divided by its sample size.

Strong convexity, stationarity at $\widehat{\phi}_{a,k}$ and Cauchy Schwarz give

$$\lambda_k \|\widehat{\phi}_{a,k} - \phi_{a,\lambda,k}\|_{\mathcal{H}} \le \|\nabla \widehat{\mathcal{R}}_{\lambda,k}(\phi_{a,\lambda,k})\|_{\mathcal{H}}.$$

This proves (C.13). Comparing the regularized population objective at $\phi_{a,\lambda,k}$ and the Bayes target gives (C.14), since the latter minimizes the unregularized risk. □

Proposition C.6 (Conditional mean prediction error)**.** *Set $C_{s,k} = \kappa_s^2 (\kappa_s \Phi_{s,t_k} + R_X)^2$ and*

$$U_k = \frac{\kappa_s^2 C_{s,k}}{n_s \lambda_{s,k}^2} + \lambda_{s,k} \Phi_{s,t_k}^2.$$

The estimator in (C.8) *obeys*

$$\mathbb{E}_{\text{data}} \int_\Omega p_{a,t_k} \|\widehat{\boldsymbol{s}}_{t_k} - \boldsymbol{s}_{t_k}\|^2 \le \frac{2 a_{\max}}{Z_a} \gamma_{t_k}^2 U_k. \tag{C.15}$$

Proof. Let $m_{\lambda,k}$ minimize the regularized population squared risk $\mathcal{Q}_k + \lambda_{s,k}\|\cdot\|^2/2$. Conditional expectation yields

$$\mathcal{Q}_k(v) - \mathcal{Q}_k(m_{t_k}) = \frac{1}{2}\int_\Omega p_{t_k}\|v - m_{t_k}\|^2.$$

Population comparison implies $\|m_{\lambda,k}\| \le \Phi_{s,t_k}$ and $\int_\Omega p_{t_k}\|m_{\lambda,k} - m_{t_k}\|^2 \le \lambda_{s,k}\Phi_{s,t_k}^2$. The empirical gradient at $m_{\lambda,k}$ is a centered average with second moment at most $C_{s,k}/n_s$, since each observation contributes a vector of norm at most $\kappa_s(\kappa_s\Phi_{s,t_k} + R_X)$. Strong convexity gives $\mathbb{E}_{\mathrm{data}}\|\widehat{m}_k - m_{\lambda,k}\|^2_{\mathcal{H}_s^d} \le C_{s,k}/(n_s\lambda_{s,k}^2)$. The evaluation bound contributes κ_s^2, and the restricted p_{t_k} measure has mass at most one. The squared triangle inequality then bounds prediction error by $2U_k$; multiplication by $\gamma_{t_k}^2$ and (C.4) give the result. □

Lemma C.7 (Weighted calibration and a Dirichlet identity)**.** *Fix $t > 0$, write $q = p_{a,t}$, $p = p_t$, $h = h_a(t,\cdot)$ and $\phi = \log h$, and let $e = f - \phi$. On a measurable region Ω, the excess of the restricted unnormalized logistic risk satisfies*

$$\mathcal{R}_t(f) - \mathcal{R}_t(\phi) \ge \frac{Z_a}{4(1+a_{\max})}\int_\Omega q\,\frac{e^2}{1+|e|}. \tag{C.16}$$

For $b \in C_c^2(\mathbb{R}^d)$, put $\mathcal{L}_q b = \Delta b + \nabla\log q\cdot\nabla b$. Then

$$\int q\|\nabla b\|^2 \le \|b\|_{L^2(q)}\|\mathcal{L}_q b\|_{L^2(q)}, \qquad \int q\|\nabla\log q\|^2 \le d/\sigma_t^2. \tag{C.17}$$

In particular, $\|\nabla b\|_\infty \le G$ and $\|\nabla^2 b\|_{\mathrm{op},\infty} \le H$ imply $\|\mathcal{L}_q b\|_{L^2(q)} \le dH + \sqrt{d}\,G/\sigma_t$.

Proof. The pointwise risk integrand is $F_x(u) = p(x)\{h(x)\ell(u) + \ell(-u)\}/2$. It has $F_x'(\phi) = 0$, $F_x''(\phi) = ph/[2(1+h)]$, and $|(\log F_x'')'| \le 1$. For $v = |e|$, Taylor's integral formula thus gives

$$F_x(f) - F_x(\phi) \ge F_x''(\phi)(v - 1 + e^{-v}) \ge \frac{ph}{4(1+a_{\max})}\frac{v^2}{1+v}.$$

The last inequality integrates $1 - e^{-u} \ge u/(1+u) \ge u/(1+v)$ over $[0,v]$. Since $ph = Z_a q$, integration proves (C.16). Integration by parts without boundary terms gives $\int q\|\nabla b\|^2 = -\int qb\mathcal{L}_q b$. For $X_t = \sqrt{\bar\alpha_t}X_0 + \sigma_t \boldsymbol{z}$ with $X_0 \sim Q_a$ and independent standard Gaussian $\boldsymbol{z}$, differentiation of the convolution gives $\nabla\log q(X_t) = -\mathbb{E}_{Q_a,K_t}[\boldsymbol{z} \mid X_t]/\sigma_t$. Conditional Jensen yields the Fisher bound. Finally $|\Delta b| \le dH$ and Cauchy Schwarz give the stated graph norm bound. □

Proposition C.8 (Localized potential gradient error)**.** *Choose a smooth $0 \le \chi_R \le 1$, equal to one on B_R and compactly supported in B_{R+1}, with gradient and Hessian bounds c_1, c_2 independent of $R \ge 1$. At time $t = t_k$ define*

$$E_{0,t} = \kappa_0\Phi_t + V_t(R+1), \qquad E_{1,t} = \kappa_1\Phi_t + 2\gamma_t R_X,$$
$$E_{2,t} = \kappa_2\Phi_t + 2\gamma_t^2 R_X^2, \qquad G_t = E_{1,t} + c_1 E_{0,t},$$
$$H_t = E_{2,t} + 2c_1E_{1,t} + c_2E_{0,t}, \qquad B_t = \Phi_t\sqrt{\frac{2(1+a_{\max})(1+E_{0,t})}{Z_a}},$$
$$D_t = B_t\left(dH_t + \frac{\sqrt{d}\,G_t}{\sigma_t}\right).$$

Then

$$\mathbb{E}_{\mathrm{data}}\int_{B_R} p_{a,t_k}\|\nabla\widehat{\phi}_{a,k} - \nabla\phi_a(t_k,\cdot)\|^2 \le 2V_k, \qquad V_k = \frac{\kappa_1^2\kappa_0^2}{4\lambda_k^2}\left(\frac{m_2}{n_a} + \frac{1}{n_0}\right) + D_{t_k}\sqrt{\lambda_k}. \tag{C.18}$$

Proof. Let $e = \phi_{a,\lambda,k} - \phi_a(t_k,\cdot)$ on $\Omega = B_{R+1}$. Equations (C.6), (C.9), (C.10) and (C.14) give $|e| \le E_{0,t}$, $\|\nabla e\| \le E_{1,t}$ and $\|\nabla^2 e\|_{\mathrm{op}} \le E_{2,t}$. The local calibration inequality therefore yields

$$\int_{B_{R+1}} p_{a,t}e^2 \le \frac{4(1+a_{\max})(1+E_{0,t})}{Z_a}\frac{\lambda_k\Phi_t^2}{2} = B_t^2\lambda_k.$$

Extend $b = \chi_R e$ by zero. Its first and second derivative bounds are G_t, H_t and it equals e on B_R. The Dirichlet and Fisher bounds imply

$$\int_{B_R} p_{a,t} \|\nabla e\|^2 \le \|b\|_{L^2(p_{a,t})} \|\mathcal{L}_{p_{a,t}} b\|_{L^2(p_{a,t})} \le D_t \sqrt{\lambda_k}.$$

For the empirical term, (C.13), derivative evaluation and the mass bound $\int_{B_R} p_{a,t} \le 1$ give the first term in V_k. The squared triangle inequality gives (C.18). □

The proof cutoff leaves estimator (C.7) unchanged; gradient control requires both value regret and derivative bounds.

C.5 Global score evaluation and the two projections

Evaluate the two finite kernel expansions globally and put $\widehat{\boldsymbol{s}}_{a,t_k} = \widehat{\boldsymbol{s}}_{t_k} + \nabla \widehat{\phi}_{a,k}$. With Euclidean projection Π_{R_X} onto B_{R_X}, project the conditional mean associated with this total score:

$$\begin{aligned} \widehat{m}_k^{\mathrm{total}}(x) &= \Pi_{R_X}\left(\frac{\sigma_{t_k}^2 \widehat{\boldsymbol{s}}_{a,t_k}(x) + x}{\sqrt{\bar{\alpha}_{t_k}}} \right), \\ \widehat{\boldsymbol{s}}_{a,t_k}^{\mathrm{proj}}(x) &= -\frac{x}{\sigma_{t_k}^2} + \gamma_{t_k} \widehat{m}_k^{\mathrm{total}}(x). \end{aligned} \tag{C.19}$$

Corollary C.9 (Discrete score risk on the full space)**.** *For any nonnegative deterministic time weights v_k, define*

$$\mathcal{T}_k(R) = \exp\left[-\frac{1}{2} \left\{ \frac{(R - \sqrt{\bar{\alpha}_{t_k}} R_X)_+}{\sigma_{t_k}} - \sqrt{d} \right\}_+^2 \right].$$

With U_k, V_k defined above,

$$\begin{aligned} \mathbb{E}_{\mathrm{data}} \sum_k v_k \int p_{a,t_k} \|\widehat{\boldsymbol{s}}_{a,t_k}^{\mathrm{proj}} - \boldsymbol{s}_{a,t_k}\|^2 \le 4 \sum_k v_k \left\{ \frac{a_{\max}}{Z_a} \gamma_{t_k}^2 U_k + V_k \right\} \\ + 4 R_X^2 \sum_k v_k \gamma_{t_k}^2 \mathcal{T}_k(R). \end{aligned} \tag{C.20}$$

Expectations are conditional on the fitted weight.

Proof. Projection is nonexpansive and fixes the true mean $m_{a,t_k}(x) \in B_{R_X}$, giving

$$\|\widehat{\boldsymbol{s}}_{a,t_k}^{\mathrm{proj}}(x) - \boldsymbol{s}_{a,t_k}(x)\| \le \|\widehat{\boldsymbol{s}}_{a,t_k}(x) - \boldsymbol{s}_{a,t_k}(x)\|, \qquad \|\widehat{\boldsymbol{s}}_{a,t_k}^{\mathrm{proj}}(x) - \boldsymbol{s}_{a,t_k}(x)\| \le 2\gamma_{t_k} R_X.$$

On B_R, combine (C.15) and (C.18) with the squared triangle inequality. Outside B_R, use the second bound above. The coupling $X_{t_k} = \sqrt{\bar{\alpha}_{t_k}} X_0 + \sigma_{t_k} \boldsymbol{z}$, $\|X_0\| \le R_X$, and Gaussian concentration $\Pr(\|\boldsymbol{z}\| > \sqrt{d} + u) \le e^{-u^2/2}$ imply $\int_{B_R^c} p_{a,t_k}(x)\,\mathrm{d}x \le \mathcal{T}_k(R)$. Multiply by v_k and sum expectations; independence across times is unnecessary. □

The smooth finite kernel expansions and Lipschitz projection make (C.19) globally defined, locally Lipschitz and of at most linear growth on each fixed interval $[\varepsilon, T]$. Neither the fitting region nor the proof cutoffs stop paths at a spatial boundary; a separate projection is applied to the final sample.

C.6 Initialization, stopping and discrete sampling

Lemma C.10 (Initialization and early stopping)**.** *Let $X \sim Q$ have $M_{2,Q} = \mathbb{E}_Q \|X\|^2 < \infty$. For the OU noising in* (C.2)*,*

$$\mathrm{KL}(K_T Q \| \mathcal{N}(0, \sigma_T^2 I_d)) \le \frac{\bar{\alpha}_T M_{2,Q}}{2\sigma_T^2}, \tag{C.21}$$

$$\mathrm{KL}(K_T Q \| \mathcal{N}(0, I_d)) \le \tfrac{1}{2}\{\bar{\alpha}_T M_{2,Q} + d(\sigma_T^2 - 1 - \log \sigma_T^2)\}, \tag{C.22}$$

$$\mathsf{W}_1(K_\varepsilon Q, Q) \le \sqrt{(1 - e^{-\varepsilon})^2 M_{2,Q} + d(1 - e^{-2\varepsilon})} =: \rho_Q(\varepsilon). \tag{C.23}$$

In particular $\rho_Q(\varepsilon) \le \varepsilon \sqrt{M_{2,Q}} + \sqrt{2d\varepsilon}$.

Proof. The first two bounds follow from Gaussian relative entropy conditional on X and convexity in its first argument. For early stopping, couple X and $e^{-\varepsilon}X + \sqrt{1-e^{-2\varepsilon}}\boldsymbol{z}$ using independent standard Gaussian $\boldsymbol{z}$. The mean squared displacement is the quantity under the square root in (C.23); apply $\mathsf{W}_1 \le \mathsf{W}_2$ and $1 - e^{-u} \le u$. □

Proposition C.11 (Final projection converts KL to Wasserstein distance)**.** *If Q is supported in B_{R_X} and $\mathrm{KL}(K_\varepsilon Q \| \widehat{Q}) \le H$, then*

$$\mathsf{W}_1((\Pi_{R_X})_{\#}\widehat{Q}, Q) \le \rho_Q(\varepsilon) + R_X\sqrt{2H}. \tag{C.24}$$

No moment assumption on $\widehat{Q}$ is needed.

Proof. Pinsker's inequality and contraction under projection bound the projected total variation distance by $\sqrt{H/2}$. Support of diameter $2R_X$ then bounds the Wasserstein distance by $2R_X\sqrt{H/2}$. Projection also contracts Wasserstein distance and fixes Q; add (C.23). □

For reverse nodes $0 = u_0 < u_1 < \cdots < u_N = T - \varepsilon$, set $h_k = u_{k+1} - u_k$ and $\tau_k = T - u_k$. The exponential integrator treats the linear drift exactly and holds the score fixed within each step:

$$Y_{k+1} = e^{h_k}Y_k + 2(e^{h_k} - 1)\widehat{\boldsymbol{s}}_{Q,\tau_k}(Y_k) + \sqrt{e^{2h_k} - 1}\,\boldsymbol{z}_k, \qquad \boldsymbol{z}_k \sim \mathcal{N}(0, I_d). \tag{C.25}$$

Its continuous interpolation satisfies $\mathrm{d}Y_u = \{Y_u + 2\widehat{\boldsymbol{s}}_{Q,\tau_k}(Y_{u_k})\}\mathrm{d}u + \sqrt{2}\mathrm{d}\overline{\boldsymbol{w}}_u$ on $[u_k, u_{k+1}]$. The Gaussian noises are independent of the initialization and each other.

Proposition C.12 (Discrete reverse diffusion)**.** *Let Q be supported in B_{R_X}, with noised density $p_{Q,t}$ and score $\boldsymbol{s}_{Q,t} = \nabla \log p_{Q,t}$. The exact reverse OU process, initialized from $p_{Q,T}$, satisfies*

$$\mathrm{d}Y_u = \{Y_u + 2\boldsymbol{s}_{Q,T-u}(Y_u)\}\mathrm{d}u + \sqrt{2}\mathrm{d}\overline{\boldsymbol{w}}_u, \qquad 0 \le u \le T - \varepsilon,$$

where $\overline{\boldsymbol{w}}$ is standard d dimensional Brownian motion. Write $\mathbb{P}$ for its path distribution and $H_T = \mathrm{KL}(p_{Q,T}\|\mu_T) < \infty$ for the initialization error of (C.25)*. Any measurable score estimate finite everywhere gives a finite recursion and nonexplosive interpolation. If the following lagged energy is finite, then*

$$\begin{aligned}\mathrm{KL}(p_{Q,\varepsilon}\|\widehat{Q}_N) &\le H_T + \sum_{k=0}^{N-1}\int_{u_k}^{u_{k+1}} \mathbb{E}_{\mathbb{P}}\|\boldsymbol{s}_{Q,T-u}(Y_u) - \widehat{\boldsymbol{s}}_{Q,\tau_k}(Y_{u_k})\|^2\,\mathrm{d}u \\ &\le H_T + 2\mathcal{E}_{\mathrm{grid}} + 2\mathcal{D}_{\mathrm{grid}},\end{aligned} \tag{C.26}$$

where

$$\begin{aligned}\mathcal{E}_{\mathrm{grid}} &= \sum_k h_k \mathbb{E}_{p_{Q,\tau_k}}\|\boldsymbol{s}_{Q,\tau_k} - \widehat{\boldsymbol{s}}_{Q,\tau_k}\|^2, \\ \mathcal{D}_{\mathrm{grid}} &= \sum_k \int_{u_k}^{u_{k+1}} \mathbb{E}_{\mathbb{P}}\|\boldsymbol{s}_{Q,T-u}(Y_u) - \boldsymbol{s}_{Q,\tau_k}(Y_{u_k})\|^2\,\mathrm{d}u.\end{aligned}$$

Proof. Conditional on the past, the interpolation drift $y + 2\widehat{\boldsymbol{s}}_{Q,\tau_k}(Y_{u_k})$ is affine with a finite constant, so the process is nonexplosive. The drift difference is twice the score difference; Girsanov's relative entropy formula with diffusion coefficient $\sqrt{2}$ therefore gives the integrated squared score difference with coefficient one. For finite lagged energy, the localization argument of Benton et al. (2024, Proposition 3 and Appendix F) justifies the comparison without a global Novikov assumption. Initialization adds H_T, and data processing gives the final relative entropy. Insert $\boldsymbol{s}_{Q,\tau_k}(Y_{u_k})$, apply the squared triangle inequality, and use the exact marginal p_{Q,τ_k} at u_k. □

Corollary C.13 (Nonuniform discretization)**.** *Suppose $M_{2,Q} < \infty$, $T \ge 1$, $0 < \varepsilon < 1$, and*

$$h_k \le \kappa \min\{1, \tau_{k+1}\} \quad (0 \le k < N) \tag{C.27}$$

for $0 < \kappa \le 1$. The general second moment bound of Benton et al. (2024, Lemma 2 and Appendix C) gives, for a numerical constant C,

$$\mathcal{D}_{\mathrm{grid}} \le C\{\kappa^2 dN + \kappa dT + \kappa M_{2,Q}\}. \tag{C.28}$$

Hence

$$\mathrm{KL}(p_{Q,\varepsilon}\|\widehat{Q}_N) \le H_T + 2\mathcal{E}_{\mathrm{grid}} + 2C\{\kappa^2 dN + \kappa dT + \kappa M_{2,Q}\}. \tag{C.29}$$

The initialization error is bounded by Lemma C.10.

For $T > 1$ and even $N \ge \max\{2(T-1), 2\log(1/\varepsilon)/\log 2, 2\}$, an admissible grid uses $N/2$ equal reverse steps on $[0, T-1]$, then $\tau_{N/2+j} = \varepsilon^{2j/N}$ for $0 \le j \le N/2$. It satisfies (C.27) *for*

$$\kappa = \max\left\{\frac{2(T-1)}{N}, \varepsilon^{-2/N} - 1\right\} \le C_0 \frac{T + \log(1/\varepsilon)}{N}.$$

For this grid, the discretization part of the KL bound is at most a constant times

$$\frac{d\{T + \log(1/\varepsilon)\}^2 + M_{2,Q}\{T + \log(1/\varepsilon)\}}{N}. \tag{C.30}$$

Proof. The discretization inequality is the cited bound. The first half of the grid has $h_k = 2(T-1)/N$ and $\tau_{k+1} \ge 1$; the second has $h_k/\tau_{k+1} = \varepsilon^{-2/N} - 1$, verifying the relative step condition. The lower bound on N gives $2\log(1/\varepsilon)/N \le \log 2$, where $e^y - 1 \le 2y$. Substitute this bound on κ into (C.28) to obtain (C.30); (C.26) then gives (C.29). □

Corollary C.14 (A conditional PU sampling bound)**.** *Let $a \ge 0$ be fitted on an independent weight estimation sample, with $Z_a > 0$, and put $\delta_g = \mathbb{E}\,|a - r^w|$. Assume clean support in B_{R_X}. Conditional on this weight, apply Corollary C.13 to $Q = Q_a$, using independent data to estimate $\boldsymbol{s}_{a,t} = \nabla \log K_t Q_a$. Let $\mathcal{E}_{\mathrm{grid},a}$ be the energy in* (C.26) *for Q_a, let $M_{2,a} = \mathbb{E}_{Q_a}\|X\|^2 \le R_X^2$, and set*

$$H_a = \mathrm{KL}(K_T Q_a\|\mu_T) + 2\mathcal{E}_{\mathrm{grid},a} + 2C\{\kappa^2 dN + \kappa dT + \kappa M_{2,a}\}.$$

For the projected exponential integrator output,

$$\mathsf{W}_1((\Pi_{R_X})_{\#}\widehat{Q}_N, Q_w) \le \frac{2R_X\delta_g}{Z_w} + \rho_{Q_a}(\varepsilon) + R_X\sqrt{2H_a}. \tag{C.31}$$

The variance matched Gaussian initialization error is bounded by $e^{-2T}R_X^2/\{2(1-e^{-2T})\}$ uniformly over the realized weight estimation sample. The sampling comparison conditions on the fitted weight and score estimates; statistical bounds on δ_g and $\mathcal{E}_{\mathrm{grid},a}$ then apply with their stated probabilities.

Proof. The sampling results apply to Q_a, which has finite second moment. Normalization gives

$$\left\|\frac{a}{Z_a} - \frac{r^w}{Z_w}\right\|_{L^1(P)} \le \frac{|Z_w - Z_a| + \delta_g}{Z_w} \le \frac{2\delta_g}{Z_w}.$$

Their supports have diameter $2R_X$, so $\mathsf{W}_1(Q_a, Q_w) \le 2R_X\,\mathrm{TV}(Q_a, Q_w) \le 2R_X\delta_g/Z_w$. Add this to (C.24) for Q_a, and use (C.29). Score error remains measured under K_tQ_a, without a density comparison to K_tQ_w. □

C.7 Conditional convergence rate

Theorem C.15 (Convergence rate for projected sampling)**.** *Let P be supported in the known B_{R_X} and let $r = \mathrm{d}P_+/\mathrm{d}P$ with $0 < Z_w = \mathbb{E}r^w < \infty$ for a fixed $w > 0$. Let an independent sample yield the estimated weight $a_n \in [0, a_{\max,n}]$, where, for fixed constants $C_{\max}, L_{\max} < \infty$,*

$$1 \le a_{\max,n} \le C_{\max}\{\log(e+n)\}^{L_{\max}}, \qquad \delta_{g,n} = \mathbb{E}|a_n - r^w|, \qquad \Pr(\delta_{g,n} \le Z_w/2) \longrightarrow 1.$$

Apply Definitions C.1 and C.2 to three fresh independent unlabeled samples, with $n = n_a + n_0 + n_s$ and $\min(n_a, n_0, n_s) \ge c_0 n$ for a fixed $c_0 > 0$. Use the fixed global kernels of (C.5)*, and put*

$$M = \lceil m \rceil, \quad D = 10M + 7, \quad \varepsilon_n = n^{-1/D}, \quad T_n = \log n, \quad R_n = R_X + \sqrt{d} + \sqrt{8\log n}.$$

For sufficiently large n, fit on B_{R_n+1} with

$$\lambda_k = n^{-(5M+3)/D}, \qquad \lambda_{s,k} = n^{-1/3}, \qquad \kappa_n = n^{-1/D}. \tag{C.32}$$

Choose the linear then geometric grid of Corollary C.13 with

$$N_n = 2\left\lceil \frac{2\{T_n + \log(1/\varepsilon_n)\}}{\kappa_n} \right\rceil, \qquad t_k = \tau_k = T_n - u_k. \tag{C.33}$$

Use the global representer expansions and the total conditional mean projection (C.19)*. Initialize* (C.25) *from* $\mathcal{N}(0, \sigma_{T_n}^2 I_d)$*, use* $\widehat{s}_{a_n,t_k}^{\mathrm{proj}}$ *at its nodes, and project its final sample onto* B_{R_X}*. Denote its conditional output distribution by* $\widehat{Q}_n^{\mathrm{proj}}$*. Then*

$$\mathsf{W}_1(\widehat{Q}_n^{\mathrm{proj}}, Q_w) \le \frac{2R_X \delta_{g,n}}{Z_w} + \widetilde{O}_p\big(n^{-1/(20M+14)}\big). \tag{C.34}$$

Here $\widetilde{O}_p$ *hides only fixed logarithmic powers of* n*. Constants depend only on dimension, support radius, kernel order and kernels,* Z_w*, and the stated sample size and truncation bounds. They are uniform over grid sizes and weight estimates satisfying the displayed event. If* $\delta_{g,n} = O_p(d_n)$ *with deterministic* $d_n \to 0$*, the convergence rate is* $O_p(d_n) + \widetilde{O}_p(n^{-1/(20M+14)})$*.*

Proof. On $\delta_{g,n} \le Z_w/2$, the weight moments satisfy

$$Z_{a_n} \in [Z_w/2, 3Z_w/2], \qquad \mathbb{E}a_n^2 \le a_{\max,n} Z_{a_n}.$$

Condition on a weight estimate satisfying this event; the following bounds are uniform over such estimates. The source bounds give

$$\Phi_t, \Phi_{s,t} = \widetilde{O}(\sigma_t^{-2M}), \qquad D_t = \widetilde{O}(\sigma_t^{-(5M+1)}), \qquad \varepsilon_n \le t \le T_n.$$

Since $M \ge 3$, the quantities $E_{0,t}, E_{1,t}, E_{2,t}$ are bounded by σ_t^{-2M} up to fixed logarithmic factors. Thus B_t contributes σ_t^{-3M} and $dH_t + \sqrt{d}\, G_t/\sigma_t$ contributes σ_t^{-2M-1}; dependence on $a_{\max,n}$ and R_n adds only fixed logarithmic powers.

In (C.20), set $v_k = h_k$, whose sum is $T_n - \varepsilon_n \le \log n$. Since $\sigma_{\varepsilon_n}^2 \asymp \varepsilon_n$,

$$\mathbb{E}_{\mathrm{data}}\mathcal{E}_{\mathrm{grid},a_n} \le \widetilde{O}\Bigg[\frac{1}{n\lambda_k^2} + \sqrt{\lambda_k}\,\varepsilon_n^{-(5M+1)/2} + \left(\lambda_{s,k} + \frac{1}{n\lambda_{s,k}^2}\right)\varepsilon_n^{-(2M+2)} + \varepsilon_n^{-2} n^{-4}\Bigg] = \widetilde{O}(n^{-1/D}).$$

The first two exponents are exactly $1/D$. For the base term, $C_{s,k} = \widetilde{O}(\sigma_{t_k}^{-4M})$ and $\gamma_{t_k}^2 \le \sigma_{t_k}^{-4}$; its decay exponent is $1/3 - (2M+2)/D = (4M+1)/(3D) > 1/D$. The bounds $\sigma_t \le 1$, $\sqrt{\bar{\alpha}_t} \le 1$ and the chosen R_n give $\mathcal{T}_k(R_n) \le n^{-4}$ uniformly. Conditional Markov's inequality gives the same rate in probability for $\mathcal{E}_{\mathrm{grid},a_n}$.

For sufficiently large n, grid (C.33) satisfies $h_k \le \kappa_n \min\{1, \tau_{k+1}\}$: its geometric exponent is at most $\kappa_n/2$, where $e^y - 1 \le 2y$. Its size $N_n = \widetilde{O}(n^{1/D})$ and $M_{2,a_n} \le R_X^2$ give a discretization term of $\widetilde{O}(n^{-1/D})$ in (C.29). The initialization KL is $O(n^{-2})$, and $\rho_{Q_{a_n}}(\varepsilon_n) = O(n^{-1/(2D)})$. Equation (C.31) now yields (C.34), since $2D = 20M + 14$. The conditioning event has probability tending to one, so the unconditional assertion follows. □

Weight estimation enters through the truncation level and the $L^1(P)$ error. Truncating the ratio at a fixed $B \ge \|r\|_\infty$ gives $a = \min(e^{\widehat{g}}, B)^w$ and $a_{\max} = \max\{1, B^w\}$; with a growing truncation level, its cost remains in $\delta_{g,n}$.

D Experimental settings and supplementary results

The main comparisons generate 65 image sets on MNIST (66,560 images), with 80 density ratio estimates, and 35 on each of CelebA and CIFAR-10; the diagnostics of Appendix D.6 add 76 sets, for 211 in total, and the MNIST studies of Appendix D.7 use 80 further ratio estimates. Unless stated otherwise, comparisons use five repetitions and $k = 5$. For R repetitions, the standard error (SE) is the sample standard deviation divided by $\sqrt{R}$, and paired SEs use differences within repetitions; TV in the theory keeps its $[0, 1]$ scale.

D.1 Data and observation designs

Table 3 gives the official splits and the samples of one observation draw; L is drawn from the labeling candidates. The records $V = (X, A)$ carry no class labels, the number of selected records varies with condition and draw, and the reserved records are not used for model selection.

Table 3: Dataset splits and the samples of one observation draw. Classifier images train the evaluation classifier. Set size applies to generated sets, resamples and references.

Dataset and official splits

Dataset	Positive class	Negative class	Total	Train	Valid.	Test	Image $H \times W \times C$
MNIST	Even digits	Odd digits	70,000	60,000	0	10,000	$28 \times 28 \times 1$
CelebA	Smiling	Not smiling	202,599	162,770	19,867	19,962	$64 \times 64 \times 3$
CIFAR-10	Vehicles	Animals	60,000	50,000	0	10,000	$32 \times 32 \times 3$

Samples assigned to the study

Dataset	Labeled L	Unlabeled U	Records $V = (X, A)$	Reserved records	Labeling candidates	Classifier images	Set size
MNIST	1,000	30,000	8,000	2,000	10,000	10,000	1,024
CelebA	1,000	100,000	8,000	2,000	52,770	19,867	5,000
CIFAR-10	1,000	20,000	8,000	2,000	10,000	10,000	1,024

Classifier images come from the official training split for MNIST and CIFAR-10 and from the official validation split for CelebA; official test images supply the evaluation references. Image sizes refer to evaluation; the MNIST diffusion model pads images to 32×32 and crops them after sampling.

Selection rules. Constant selection uses probability 0.6 on every dataset, and reporting is complete, $c_{\text{rep}} = 1$, so $e = v$. Quartiles of mean pixel intensity on U define the MNIST intensity groups and the CIFAR-10 brightness groups. From the lowest to the highest quartile, strong selection on MNIST and biased selection on CIFAR-10 use probabilities $(0.1, 0.25, 0.55, 0.9)$, and weak selection on MNIST uses $(0.55, 0.565, 0.595, 0.63)$. CelebA assigns 0.05 to images annotated Male and 0.8 otherwise; Male is a dataset annotation, not verified gender, and the IPW model receives pixels and selection indicators, not the annotation. Vehicles are airplanes, automobiles, ships and trucks; animals are birds, cats, deer, dogs, frogs and horses. MNIST and CIFAR-10 take the first labeled positives in a random order of the candidates; CIFAR-10 uses a single selection draw.

Repetitions and comparisons. The MNIST generation comparison keeps the observed samples fixed and varies the seeds for fitting and sampling; its weight evaluation redraws the observed samples five times. The RGB comparisons redraw L and V with each seed and keep U, the evaluation classifiers, both references and the base model fixed, so their SEs exclude variation in the base model and references. Each RGB dataset has 30 guided sets, five unguided sets and ten resampling comparisons with random references, from 40 ratio estimates and 30 potentials.

D.2 Estimation and generation

Estimators. All methods share a base model and sampler. The uncorrected estimator and SR contrast L with all record images and with the selected subset V_1, respectively; IPW keeps all record images as reference and weights the positive loss term by normalized inverse estimates of v. On MNIST, these are binomial maximum likelihood estimates within four intensity groups or two merged groups, with floor 0.025; both models represent constant selection, and only four groups represent strong selection. Ratios are divided by their mean over U and truncated at $a_{\max} = 16$. Unguided sampling uses neither a ratio nor a potential.

MNIST models and training. The DDPM base follows Song et al. (2021b): 128 initial channels, multipliers $(1, 2, 2, 2)$, two residual blocks per resolution, attention at resolution 16, dropout 0.1, and time embedding dimension 512. MNIST padding uses background pixels, without horizontal flips. The ratio classifier and the noisy potential are residual convolutional networks with widths $(32, 64, 64)$, global averaging, a scalar output and time embeddings of dimension 128; the ratio classifier sets time to zero, and the RGB versions change only the number of input channels. Table 4 gives the training settings. Both networks clip gradient norms at 1 and use the final iterate, without moving averages or selection by evaluation metrics.

Table 4: Training settings. Durations are epochs for RGB classifiers and iterations otherwise. The MNIST base uses 5,000 warmup iterations and the final moving average with maximum coefficient $1-10^{-4}$. The MNIST classifier uses its final iterate; RGB classifiers select epochs as in Appendix D.3.

Dataset	Model	Duration	Optimizer	Batch	Learning rate	Weight decay
MNIST	Base diffusion	200,000	Adam	128	2×10^{-4}	0
	Ratio classifier (generation)	1,500	AdamW	256	2×10^{-4}	10^{-4}
	Noisy potential	20,000	AdamW	256	2×10^{-4}	10^{-4}
	Ratio classifier (RQ2, RQ3)	200	AdamW	256	2×10^{-4}	10^{-4}
	Evaluation classifier	1,500	AdamW	256	10^{-3}	10^{-4}
RGB	Ratio and propensity	200	AdamW	128	2×10^{-4}	10^{-4}
	Noisy potential	20,000	AdamW	128	2×10^{-4}	10^{-4}
CIFAR-10	Base adaptation	5,000	Adam	128	2×10^{-5}	0
	Evaluation classifier	100	SGD	128	0.1	5×10^{-4}
CelebA	Evaluation classifier	20	AdamW	128	3×10^{-4}	10^{-4}

RGB base models. The RGB comparisons hold one pretrained base fixed per dataset*. CIFAR-10 uses the unconditional DDPM of Song et al. (2021b), with moving average parameters after 1,300,000 iterations on all 50,000 official training images. Pretraining sees no class labels but includes pixels later assigned to L, V and the evaluation classifier; its empirical distribution differs from U. Adaptation on U restarts the optimizer and the moving average, whose final maximum coefficient is $1 - 10^{-4}$, with 500 warmup iterations, gradient norm clipping at 1 and horizontal flips with probability 0.5.

CelebA uses the unconditional model of Song et al. (2021a), with moving average parameters after 500,000 iterations on official training images and no adaptation. It also sees pixels beyond U, without attribute labels. Images use the original 128×128 crop from row 57 and column 25, resized to 64×64 by bilinear interpolation; CIFAR-10 keeps 32×32 images. Neither base sees official test images.

RGB correction. IPW fits the selection indicators from pixels with a logistic loss, two folds and floor 0.025; predictions on L use the model fitted on the other fold, and correct specification is not assumed. Ratio classifiers hold out 20% of the labeled and reference images for the held out loss. The CIFAR-10 propensity model uses all records, whereas CelebA excludes the held out reference images. These models clip gradient norms at 1 and use the final iterate; the guidance strength is 1.

Sampling. All datasets use 1,000 forward steps with β_t linearly spaced from 10^{-4} to 0.02, $\bar{\alpha}_t = \prod_{s=1}^{t}(1 - \beta_s)$, and $\sigma_t^2 = 1 - \bar{\alpha}_t$. MNIST and CIFAR-10 use all 1,000 reverse steps; CelebA uses 200 respaced ancestral steps. All methods use posterior reverse variance and shared noise. MNIST projects the final images to $[-1, 1]$ without clipping intermediate predictions of the clean image; CelebA and CIFAR-10 clip these predictions at every step, and CelebA pixels are stored in float16. Weighted resampling draws from U with replacement, using random numbers shared across methods.

*Official model releases. CIFAR-10 (`vp/cifar10_ddpm`, final EMA): `https://github.com/yang-song/score_sde_pytorch`. CelebA (64×64, EMA): `https://github.com/ermongroup/ddim`.

Checks before the RGB comparison. Two separate observation draws cover both conditions and all four estimators, including random references; a ratio classifier passes if its held out balanced loss does not exceed the loss at zero logits. The sampler is checked on constant weights and on normalized inverse probabilities $(0.8, 0.4, 0.2, 0.05)$ over the brightness quartiles of U: in each of 16 regions defined by quartiles of brightness and pixel standard deviation with target mass at least 0.005, the generated count must fall in a binomial interval, with error level 0.05 split across regions and targets. The comparison started only after every classifier and region passed.

Approximation limits. On MNIST, the sampler did not reproduce known weighted targets on U, so generation comparisons do not separate ratio error from the errors of the base model, potential and sampler; weighted resampling and the exact conditional quantities of Appendix D.10 separate them. The RGB datasets have their own checks and resampling comparisons (Table 19), whose discrepancies the MNIST diagnostics do not explain.

D.3 Evaluation quantities

The MNIST evaluation classifier, trained on the 10,000 classifier images, has convolutional widths 32 and 64, max pooling, 128 features and ten output classes. Class annotations are used only for evaluation, and every generated image is evaluated without filtering.

Let $\boldsymbol{R}_i$, $i = 1, \ldots, N_R$, and $\boldsymbol{G}_j$, $j = 1, \ldots, N_G$, be reference and evaluated feature vectors. Let $d_i^{(k)}$ be the Euclidean distance from $\boldsymbol{R}_i$ to its kth other reference vector. Following Naeem et al. (2020),

$$\text{precision} = \frac{1}{N_G} \sum_j \mathbf{1}\{\exists i : \|\boldsymbol{G}_j - \boldsymbol{R}_i\| < d_i^{(k)}\},$$

$$\text{coverage} = \frac{1}{N_R} \sum_i \mathbf{1}\{\min_j \|\boldsymbol{G}_j - \boldsymbol{R}_i\| < d_i^{(k)}\}.$$

We report $k = 3, 5, 10$. Each reference excludes itself as a neighbor; other duplicates and zero radii are retained. These quantities depend on the feature encoder and the sample sizes as well as on the distributions. MMD squared is the complete unbiased U statistic for $K(\boldsymbol{x}, \boldsymbol{y}) = (1 + \boldsymbol{x}^\top \boldsymbol{y}/128)^3$ (Bińkowski et al., 2018), retaining negative estimates; tables report $\mathrm{MMD}^2/1{,}000$.

Each MNIST reference contains 1,024 of the 4,926 even digits in the official test split: the positive reference is drawn without replacement, and the labeled reference follows the selection rule, independently of L. The two references may overlap and stay fixed across methods. The feature metrics were chosen after the histogram diagnostics and before any metric was computed.

CelebA and CIFAR-10 evaluation. Generated images, resamples and references are compared through Inception features of dimension 2,048, using Fréchet Inception distance (FID), kernel Inception distance (KID), and precision, recall, density and coverage at $k = 3, 5, 10$ (Heusel et al., 2017; Bińkowski et al., 2018; Naeem et al., 2020). KID uses the same unbiased cubic kernel estimate and retains negative values; lower FID and KID mean closer feature distributions. Equal sample counts support comparisons within, not across, datasets.

The CIFAR-10 evaluation classifier is a ResNet18 (He et al., 2016) trained on 8,000 classifier images, with the epoch selected by cross entropy on the other 2,000. Its first convolution is 3×3 with stride 1 and no initial max pooling. SGD uses cosine decay and momentum 0.9; augmentation uses random 32×32 crops after reflection padding by four pixels. Argmax predictions give class and vehicle fractions. Reference comparisons use both annotations and predictions: all reference images are annotated vehicles, but only 97.36% are predicted vehicles.

The CelebA evaluation classifier is a ResNet18 trained on 80% of the official validation images, with the epoch selected by binary cross entropy on the rest. Threshold 0.5 yields Smiling, Male, and Wearing_Lipstick predictions. Joint Smiling/Male frequencies and marginal lipstick frequency complement the feature metrics. Test accuracy over 19,962 images is 90.23%, 95.00%, and 90.52%, respectively. All population reference images are annotated Smiling, but only 85.86% are predicted Smiling; comparisons use both reference definitions. Both RGB classifiers are trained from scratch with horizontal flips of probability 0.5 (Table 4) and stay fixed during evaluation.

D.4 Positive mass and composition of the weighted unlabeled sample

We evaluate all 225 fitted weight vectors with the annotations of U, without refitting or resampling. For final weights a_i on U,

$$Q_a(B) = \frac{\sum_{i \in U} a_i \mathbf{1}_B(i)}{\sum_{i \in U} a_i},$$

and the target is the uniform distribution on the annotated positives in U. We report $100\,\mathrm{ESS}/|U|$ with $\mathrm{ESS} = (\sum_i a_i)^2 / \sum_i a_i^2$; larger values mean less concentrated weights. Let J be the dataset class and G the selection group: intensity quartile on MNIST, brightness quartile on CIFAR-10 and Male on CelebA. Group, subclass and conditional joint TV compare G, J and (J, G) within positives, and full joint TV includes the negative classes; all compare distributions over categories. CelebA has a single positive class, so its subclass TV is zero and its Male composition appears under group TV. Writing $p = Q_a(Y = 1)$, with conditional joint frequencies c and target frequencies t,

$$\mathrm{TV}_{\mathrm{joint}} = \frac{1-p}{2} + \frac{1}{2}\|pc - t\|_1 \geq 1 - p.$$

SEs condition on U and the references. In the independent MNIST study at 8,000 indicators under strong selection, SR minus IPW with four groups has joint TV difference 0.45 ± 0.80 (first optimization seed).

Changing the evaluation target. Table 5 compares all nine conditions with the positives in U and with an independent positive reference; on CelebA, their Male frequencies are 34.92% and 30.70%. Every CelebA weight vector assigns less mass to each positive joint category than either target, so its joint TV equals $1 - Q_a(Y = 1)$.

Table 5: SR minus uncorrected weights, all nine conditions. Each condition has two reference rows: positives in U and an independent positive reference (Indep.). Positive mass is shared across the two rows. CelebA has one positive class, so its subclass TV is zero ($-$). The independent MNIST study (rows MNIST clean) uses all records as the uncorrected reference, 8,000 indicators and the first optimization seed.

Selection	Ref.	Δ Positive mass (pp)	Δ Group TV (×100)	Δ Subclass TV (×100)	Δ Conditional joint TV (×100)	Δ Joint TV (×100)
MNIST						
Constant	U	$-0.33_{\pm 0.23}$	$-4.56_{\pm 1.81}$	$-8.82_{\pm 2.79}$	$-8.80_{\pm 3.62}$	$-8.65_{\pm 3.54}$
	Indep.		$-3.57_{\pm 1.84}$	$-7.80_{\pm 2.63}$	$-6.38_{\pm 3.10}$	$-6.24_{\pm 2.97}$
Weak	U	$-0.12_{\pm 0.35}$	$-6.57_{\pm 1.34}$	$-1.55_{\pm 3.00}$	$-3.31_{\pm 3.10}$	$-3.33_{\pm 3.01}$
	Indep.		$-2.65_{\pm 1.19}$	$-0.55_{\pm 2.18}$	$-2.88_{\pm 2.08}$	$-2.92_{\pm 1.97}$
Strong	U	$-0.93_{\pm 0.32}$	$-23.58_{\pm 2.64}$	$-8.20_{\pm 2.30}$	$-20.82_{\pm 2.67}$	$-20.14_{\pm 2.60}$
	Indep.		$-22.65_{\pm 2.98}$	$-6.92_{\pm 2.79}$	$-18.94_{\pm 2.88}$	$-18.36_{\pm 2.79}$
CelebA						
Constant	U	$-1.83_{\pm 0.74}$	$-0.27_{\pm 0.54}$	$-$	$-0.27_{\pm 0.54}$	$+1.83_{\pm 0.74}$
	Indep.		$+0.56_{\pm 0.67}$	$-$	$+0.56_{\pm 0.67}$	$+1.83_{\pm 0.74}$
Biased	U	$-4.81_{\pm 1.04}$	$-13.66_{\pm 1.42}$	$-$	$-13.66_{\pm 1.42}$	$+4.81_{\pm 1.04}$
	Indep.		$-7.37_{\pm 1.87}$	$-$	$-7.37_{\pm 1.87}$	$+4.81_{\pm 1.04}$
CIFAR-10						
Constant	U	$-0.56_{\pm 0.91}$	$+0.71_{\pm 0.68}$	$-0.26_{\pm 0.22}$	$+0.45_{\pm 0.26}$	$+0.67_{\pm 0.96}$
	Indep.		$+0.72_{\pm 0.75}$	$-0.35_{\pm 0.27}$	$+0.37_{\pm 0.35}$	$+0.69_{\pm 0.87}$
Biased	U	$-0.05_{\pm 1.53}$	$-19.75_{\pm 1.06}$	$-3.79_{\pm 0.83}$	$-18.37_{\pm 0.98}$	$-12.50_{\pm 1.45}$
	Indep.		$-17.57_{\pm 0.81}$	$-3.27_{\pm 0.77}$	$-14.11_{\pm 1.08}$	$-10.41_{\pm 1.26}$
MNIST clean						
Constant	U	$-1.08_{\pm 1.29}$	$+0.38_{\pm 0.84}$	$+0.42_{\pm 1.06}$	$+0.68_{\pm 1.09}$	$+1.28_{\pm 0.75}$
	Indep.		$-0.01_{\pm 0.48}$	$+0.05_{\pm 1.29}$	$+0.30_{\pm 0.64}$	$+1.35_{\pm 0.54}$
Strong	U	$+0.18_{\pm 0.50}$	$-18.94_{\pm 2.64}$	$-4.44_{\pm 1.32}$	$-15.67_{\pm 2.07}$	$-14.35_{\pm 2.17}$
	Indep.		$-17.97_{\pm 3.04}$	$-4.28_{\pm 1.28}$	$-12.90_{\pm 2.44}$	$-12.12_{\pm 2.23}$

Matching positive mass. To separate positive mass from composition within positives, we mix the uncorrected weights with the uniform distribution P_U on U:

$$Q_{\text{mix}} = \lambda Q_{\text{raw}} + (1-\lambda)P_U, \qquad \lambda = \frac{p_{\text{SR}} - p_U}{p_{\text{raw}} - p_U},$$

where p_ν is the positive mass under ν. This comparison uses the true annotations and serves only for evaluation. A pair is comparable when $\lambda \in [0,1]$; this holds for 31 of 45 pairs, and the other 14 are reported as not applicable (Table 6). SEs are taken over the comparable pairs.

Under biased CelebA selection, Male frequencies within positives are 24.15% for the matched mixture and 33.87% for SR. SR reduces conditional group TV by 9.67 ± 0.76 against the same U and by 3.38 ± 1.34 against the independent reference, while the joint TV difference is zero; the reduction against the independent reference occurs in four of five pairs. Constant selection gives no uniform composition improvement; CIFAR-10 has larger mean group and joint TV under both targets. Biased CIFAR-10 and the strong independent MNIST study have only two comparable pairs of five.

Table 6: SR minus the mixture with matched positive mass: mean $\pm$ paired SE over n pairs. Reference labels follow Table 5; both rows use the same n pairs and λ. On CelebA, subclass TV is zero and joint TV equals $1 - Q_a(Y=1)$ (Appendix D.4), so matching the positive mass makes both differences zero ($-$). Negative differences favor SR.

Selection	$n/5$	λ (%)	Ref.	Δ Group TV ($\times 100$)	Δ Subclass TV ($\times 100$)	Δ Conditional joint TV ($\times 100$)	Δ Joint TV ($\times 100$)
MNIST							
Constant	4/5	$99.02_{\pm 0.41}$	U	$-5.43_{\pm 1.98}$	$-9.79_{\pm 3.31}$	$-10.43_{\pm 4.06}$	$-10.43_{\pm 4.04}$
			Indep.	$-4.46_{\pm 2.02}$	$-8.51_{\pm 3.21}$	$-7.81_{\pm 3.44}$	$-7.83_{\pm 3.37}$
Weak	4/5	$99.18_{\pm 0.50}$	U	$-7.48_{\pm 1.19}$	$-3.79_{\pm 2.51}$	$-5.88_{\pm 2.12}$	$-5.95_{\pm 2.12}$
			Indep.	$-3.27_{\pm 1.26}$	$-2.09_{\pm 1.94}$	$-4.27_{\pm 1.93}$	$-4.40_{\pm 1.83}$
Strong	4/5	$97.66_{\pm 0.56}$	U	$-25.13_{\pm 2.36}$	$-8.84_{\pm 2.70}$	$-22.53_{\pm 2.13}$	$-21.98_{\pm 2.12}$
			Indep.	$-24.41_{\pm 2.69}$	$-7.55_{\pm 3.37}$	$-20.66_{\pm 2.52}$	$-20.30_{\pm 2.41}$
CelebA							
Constant	4/5	$79.13_{\pm 6.46}$	U	$+0.13_{\pm 0.72}$	$-$	$+0.13_{\pm 0.72}$	$-$
			Indep.	$+0.23_{\pm 0.86}$	$-$	$+0.23_{\pm 0.86}$	$-$
Biased	5/5	$68.46_{\pm 3.96}$	U	$-9.67_{\pm 0.76}$	$-$	$-9.67_{\pm 0.76}$	$-$
			Indep.	$-3.38_{\pm 1.34}$	$-$	$-3.38_{\pm 1.34}$	$-$
CIFAR-10							
Constant	3/5	$95.55_{\pm 0.64}$	U	$+1.68_{\pm 0.56}$	$-0.33_{\pm 0.35}$	$+0.88_{\pm 0.19}$	$+0.28_{\pm 0.09}$
			Indep.	$+1.64_{\pm 0.88}$	$-0.42_{\pm 0.44}$	$+0.85_{\pm 0.42}$	$+0.61_{\pm 0.11}$
Biased	2/5	$92.31_{\pm 3.03}$	U	$-20.94_{\pm 1.38}$	$-2.40_{\pm 1.00}$	$-18.54_{\pm 2.13}$	$-10.17_{\pm 2.63}$
			Indep.	$-17.05_{\pm 2.05}$	$-2.04_{\pm 0.72}$	$-13.10_{\pm 2.37}$	$-8.68_{\pm 2.59}$
MNIST clean							
Constant	3/5	$93.13_{\pm 1.17}$	U	$+1.09_{\pm 1.34}$	$-0.48_{\pm 1.19}$	$-0.05_{\pm 1.47}$	$-0.74_{\pm 0.79}$
			Indep.	$+0.20_{\pm 0.66}$	$-1.22_{\pm 1.48}$	$-0.50_{\pm 0.34}$	$+0.08_{\pm 0.52}$
Strong	2/5	$98.36_{\pm 1.43}$	U	$-16.52_{\pm 3.89}$	$-6.65_{\pm 2.51}$	$-15.63_{\pm 4.92}$	$-13.50_{\pm 4.87}$
			Indep.	$-14.93_{\pm 5.92}$	$-5.98_{\pm 2.99}$	$-10.97_{\pm 6.28}$	$-10.20_{\pm 5.48}$

D.5 Selection correction in generated images

Table 7 reports changes in quality and predicted composition relative to uncorrected generation for every selection condition; Table 1 gives the shared precision and coverage comparisons, and Table 8 adds weak MNIST selection and IPW with two groups. On MNIST, SR increases mean coverage and decreases precision under strong selection, but does not increase mean coverage under constant or weak selection. These directions persist across neighborhood sizes (Table 9).

CelebA quality and attributes. Under biased selection, SR lowers mean FID/KID without increasing mean coverage; under constant selection, mean FID/KID increase. On both RGB datasets, precision differences remain negative at $k = 3, 5, 10$, while small coverage differences change

sign (Table 9). Closer features do not mean closer attributes: in the main comparison, with 5,000 images per repetition, predicted Smiling declines from $(54.76 \pm 0.81)\%$ for uncorrected guidance to $(52.33 \pm 0.33)\%$ for SR, and the joint Smiling and Male TV against reference predictions rises from 31.10 ± 0.81 to 33.53 ± 0.33, a paired increase of 2.43 ± 0.62. The same increase holds against reference annotations; SEs exclude classifier error.

Table 7: SR minus uncorrected generation in all seven selection conditions. The Fréchet distance uses MNIST evaluator features on MNIST and Inception features (FID) on RGB datasets. Discrepancy is MMD squared divided by 1,000 on MNIST and KID multiplied by 1,000 on RGB datasets. Composition TV compares predictions on generated images with the reference: joint Smiling/Male on CelebA and ten classes on CIFAR-10 against reference predictions, and ten classes on MNIST against reference labels.

Selection	Fréchet	Discrepancy	Predicted positive (pp)	Composition TV ($\times 100$)
MNIST				
Constant	$+62.55_{\pm 39.14}$	$+1.91_{\pm 2.12}$	$-5.31_{\pm 2.14}$	$+4.94_{\pm 2.19}$
Weak	$+61.38_{\pm 26.22}$	$+2.68_{\pm 1.34}$	$-4.12_{\pm 1.58}$	$+3.98_{\pm 1.39}$
Strong	$-22.30_{\pm 45.98}$	$-11.11_{\pm 5.51}$	$-6.72_{\pm 0.52}$	$+2.77_{\pm 1.71}$
CelebA				
Constant	$+0.22_{\pm 0.17}$	$+0.29_{\pm 0.22}$	$-1.07_{\pm 0.41}$	$+1.07_{\pm 0.41}$
Biased	$-1.51_{\pm 0.23}$	$-2.11_{\pm 0.31}$	$-2.43_{\pm 0.62}$	$+2.43_{\pm 0.62}$
CIFAR-10				
Constant	$+0.38_{\pm 0.75}$	$+0.25_{\pm 0.27}$	$-0.45_{\pm 0.73}$	$+0.25_{\pm 0.57}$
Biased	$+0.72_{\pm 0.92}$	$-0.80_{\pm 0.43}$	$-2.30_{\pm 1.19}$	$+0.80_{\pm 0.64}$

Table 8: MNIST comparisons beyond Table 1. The shared unguided result is unchanged across conditions. IPW parentheses give the number of groups.

Selection and method	Precision (%)	Coverage (%)
Weak: Uncorrected	$82.09_{\pm 0.66}$	$93.12_{\pm 0.75}$
Weak: IPW (4 groups)	$80.80_{\pm 0.57}$	$93.52_{\pm 0.36}$
Weak: IPW (2 groups)	$81.70_{\pm 0.80}$	$92.54_{\pm 0.22}$
Weak: SR	$79.55_{\pm 0.79}$	$92.52_{\pm 0.72}$
Constant: IPW (2 groups)	$79.98_{\pm 1.03}$	$92.66_{\pm 0.55}$
Strong: IPW (2 groups)	$79.92_{\pm 1.01}$	$89.55_{\pm 1.12}$

Table 9: SR minus uncorrected generation against population positives: neighborhood sensitivity in pp. All 21 combinations of dataset, selection, and neighborhood size k are retained.

Selection	Precision $k=3$	Coverage $k=3$	Precision $k=5$	Coverage $k=5$	Precision $k=10$	Coverage $k=10$
MNIST						
Constant	$-3.91_{\pm 1.73}$	$-0.96_{\pm 0.60}$	$-3.63_{\pm 1.70}$	$-0.16_{\pm 0.63}$	$-3.24_{\pm 1.57}$	$-0.16_{\pm 0.22}$
Weak	$-2.89_{\pm 1.43}$	$-2.68_{\pm 0.92}$	$-2.54_{\pm 1.33}$	$-0.61_{\pm 0.60}$	$-1.95_{\pm 1.05}$	$+0.00_{\pm 0.22}$
Strong	$-3.67_{\pm 0.52}$	$+4.18_{\pm 1.33}$	$-4.06_{\pm 0.62}$	$+4.57_{\pm 1.63}$	$-2.91_{\pm 0.60}$	$+1.45_{\pm 0.82}$
CelebA						
Constant	$-0.16_{\pm 0.39}$	$+0.20_{\pm 0.59}$	$-0.13_{\pm 0.35}$	$-0.03_{\pm 0.41}$	$-0.53_{\pm 0.14}$	$+0.04_{\pm 0.11}$
Biased	$-0.32_{\pm 0.25}$	$+0.40_{\pm 0.53}$	$-0.30_{\pm 0.08}$	$-0.06_{\pm 0.35}$	$-0.41_{\pm 0.19}$	$-0.49_{\pm 0.28}$
CIFAR-10						
Constant	$-0.37_{\pm 0.94}$	$-0.35_{\pm 0.29}$	$-0.35_{\pm 0.94}$	$+0.33_{\pm 0.38}$	$-0.08_{\pm 0.89}$	$+0.00_{\pm 0.19}$
Biased	$-1.70_{\pm 1.08}$	$-0.47_{\pm 0.74}$	$-2.09_{\pm 0.96}$	$+0.41_{\pm 0.72}$	$-2.25_{\pm 1.21}$	$-0.06_{\pm 0.16}$

CIFAR-10 quality and attributes. All guided methods have lower mean FID/KID and higher precision, coverage, and predicted vehicle fraction than the unguided base. SR has no uniform advan-

tage over uncorrected guidance or IPW; under biased selection, its FID difference from uncorrected guidance is 0.72 ± 0.92.

D.6 Separating weight estimation and generation

These diagnostics were added after the main comparisons, with settings fixed in advance. They use five repetitions, the same base models and classifiers, and 1,024 images per set; reused CelebA sets are the first 1,024 of the original 5,000, so feature metrics are not comparable across set sizes.

Guidance strength and additional information. Strong selection on MNIST and biased selection on CelebA use strengths $s \in \{0, 0.5, 1, 1.5\}$ with shared sampling noise, reusing the unguided samples at $s = 0$ and the original samples at $s = 1$. The strength scales the guidance gradient; the potential is not refitted for a^s. Two further targets use information unavailable to PU estimators: weights equal to the true class on U and, on MNIST, the true selection probabilities. The latter still pass through ratio fitting, normalization and truncation, which leave target error.

Observed samples and training length. The five further MNIST repetitions reuse the draws of L and V from the weight evaluation with new ratio classifiers, potentials and sampling noise, so their SE combines observation, fitting and sampling variation. On CelebA, ratio classifiers trained for 200, 1,000 and 3,000 iterations share one held out split; before any annotation was examined, the smallest held out balanced logistic loss chose the training length, and for SR it chose 200 iterations in every repetition. The annotation results then motivated generation from both longer SR classifiers with the original base model, potential training and references; this comparison stays within SR and gives no label free rule for the training length.

Matching predicted positive fractions. Within each repetition, adjacent uncorrected strengths have predicted positive fractions p_0, p_1 and conditional compositions c_0, c_1. When they bracket the fraction p_S of SR at $s = 1$, their empirical mixture has

$$\lambda = \frac{p_S - p_0}{p_1 - p_0}, \qquad c_{\text{mix}} = \frac{(1-\lambda)p_0 c_0 + \lambda p_1 c_1}{p_S}.$$

Compositions use the classifier's predictions on generated and reference images. Every bracketing pair is used, without extrapolation, including the whole interval when $p_0 = p_1 = p_S$. The mixture is a comparison within a repetition, not a sample generated at an intermediate strength.

Table 10: MNIST diagnostics at strength 1: mean $\pm$ SE over five repetitions, each with 1,024 generated and 1,024 reference images. Group TV compares intensity groups among predicted positives. The true class indicator and known selection probabilities use information unavailable to PU estimators; rows marked other draws use the further draws of L and V.

Method	Predicted positive (%)	Group TV ($\times 100$)	Precision (%)	Coverage (%)
Uncorrected	$81.09_{\pm 1.43}$	$25.28_{\pm 3.79}$	$82.81_{\pm 0.69}$	$88.34_{\pm 0.91}$
SR	$74.38_{\pm 1.31}$	$5.56_{\pm 0.71}$	$78.75_{\pm 0.88}$	$92.91_{\pm 0.86}$
True class indicator	$83.12_{\pm 0.66}$	$5.49_{\pm 0.52}$	$85.39_{\pm 0.64}$	$94.32_{\pm 0.48}$
Known selection probabilities	$79.94_{\pm 0.72}$	$19.40_{\pm 2.56}$	$81.80_{\pm 0.29}$	$89.43_{\pm 1.28}$
Uncorrected (other draws)	$80.51_{\pm 1.12}$	$26.72_{\pm 3.33}$	$81.74_{\pm 0.43}$	$88.11_{\pm 1.40}$
SR (other draws)	$74.82_{\pm 1.57}$	$5.65_{\pm 1.35}$	$79.67_{\pm 0.68}$	$91.29_{\pm 0.86}$

What the true class weights resolve. On CelebA, weights equal to the true class raise the predicted Smiling fraction of generated images by 11.80 ± 0.22 pp over SR after 200 iterations, but on neither dataset do the generated images reach the predicted positive fraction of the reference (Figure 3). With the true MNIST selection probabilities, truncation removes 13.7% to 36.0% of the weight mass.

Sensitivity to the observed samples. Across the five further MNIST draws, SR changes coverage by 3.18 ± 0.90 pp, precision by -2.07 ± 1.04 pp, conditional group TV by -21.07 ± 2.75 (lower in every pair) and conditional digit and group TV by -16.34 ± 2.68, while the predicted positive fraction falls by 5.68 ± 2.21 pp and the encoder Fréchet distance rises by 12.53 ± 47.07.

Table 11: CelebA ratio classifiers by training length (updates are iterations): mean ± SE over five repetitions. Selected counts the repetitions in which the held out classification loss chose that length before annotations were examined. Positive mass and Male TV use the true annotations of U; joint TV compares class and Male with the true positives in U.

Method	Updates	Selected	Held out loss (unscaled)	True positive mass (%)	Group TV (×100)	Joint TV (×100)
Uncorrected	200	4/5	$0.663_{\pm 0.013}$	$62.71_{\pm 1.48}$	$14.76_{\pm 1.60}$	$37.29_{\pm 1.48}$
Uncorrected	1,000	1/5	$0.813_{\pm 0.056}$	$87.49_{\pm 0.97}$	$28.93_{\pm 0.64}$	$29.70_{\pm 0.50}$
Uncorrected	3,000	0/5	$2.275_{\pm 0.256}$	$87.01_{\pm 1.85}$	$28.81_{\pm 0.77}$	$29.66_{\pm 0.58}$
SR	200	5/5	$0.679_{\pm 0.007}$	$57.90_{\pm 0.75}$	$1.10_{\pm 0.41}$	$42.10_{\pm 0.75}$
SR	1,000	0/5	$0.842_{\pm 0.048}$	$76.17_{\pm 1.94}$	$6.50_{\pm 1.63}$	$23.83_{\pm 1.94}$
SR	3,000	0/5	$2.049_{\pm 0.133}$	$76.73_{\pm 1.37}$	$5.97_{\pm 1.78}$	$23.27_{\pm 1.37}$

Table 12: CelebA generation against its own weighted target, with one classifier run under identical settings: mean ± SE over five repetitions. Weighted predictions are exact sums over U, and each generated set has 1,024 images. Δ is generated minus target predicted Smiling, and joint TV compares their predicted Smiling and Male distributions.

Method (updates)	Weighted target positive (%)	Generated positive (%)	Δ (pp)	Joint TV (×100)
Uncorrected (200)	$58.64_{\pm 1.64}$	$54.39_{\pm 0.80}$	$-4.25_{\pm 0.99}$	$8.69_{\pm 1.45}$
SR (200)	$53.36_{\pm 0.79}$	$52.30_{\pm 0.41}$	$-1.06_{\pm 0.49}$	$2.56_{\pm 0.34}$
SR (1,000)	$72.92_{\pm 2.10}$	$57.44_{\pm 1.45}$	$-15.48_{\pm 0.99}$	$15.52_{\pm 0.95}$
SR (3,000)	$73.67_{\pm 1.54}$	$49.98_{\pm 1.58}$	$-23.69_{\pm 1.53}$	$23.74_{\pm 1.51}$
True Smiling weights	$85.04_{\pm 0.00}$	$64.10_{\pm 0.36}$	$-20.94_{\pm 0.36}$	$20.94_{\pm 0.36}$

Composition at a matched predicted positive fraction. Every repetition has one uncorrected mixture that matches the predicted positive fraction of SR. MNIST conditional group TV falls in all five pairs (-14.09 ± 2.95), and conditional digit and group TV changes by -11.33 ± 3.17; against labeled positives, group TV instead rises by 11.59 ± 1.61. CelebA Male TV rises by 2.76 ± 1.48, worse in four pairs, so its gain on U does not carry over to generated composition.

Longer ratio training. Longer SR training improves positive mass and joint TV but worsens Male TV (Table 11); against equally trained uncorrected classifiers, joint TV falls in every pair (-5.87 ± 2.19 at 1,000 iterations and -6.38 ± 1.74 at 3,000), despite lower positive mass. With 1,000 training iterations, predicted Smiling among generated images rises from 52.30% to 57.44% (paired $+5.14 \pm 1.07$ pp); with 3,000 it falls to 49.98% (-2.32 ± 1.35 pp). Male TV against held out positive predictions worsens from 5.51 to 5.98 and 6.99, and coverage changes by 0.20 ± 0.73 and -1.04 ± 1.41 pp. SR after 200 iterations generates close to its own target, whose predicted Smiling fraction is low (Table 12); generation from the longer classifiers falls below its targets by 15.48 ± 0.99 and 23.69 ± 1.53 pp, and generation from the true Smiling weights by 20.94 ± 0.36 pp.

D.7 Selection information and model restrictions

Five draws of the observed samples each contain 1,000 labeled positives and 8,000 records, with a fixed U of 30,000 images. Under constant and strong selection, ratio classifiers against all records, the selected records and a random subset of the same size give 30 weightings. Under strong selection, IPW with four or two groups keeps all 8,000 record images as reference and observes 500, 2,000 or 8,000 nested indicators, giving 30 more; SR, with 8,000 indicators, matches IPW in information only at that count. Twenty further weightings repeat SR, random subsets and both IPW models at 8,000 records with a second optimization seed.

Each weighting is evaluated by averaging twenty resamples of 1,024 images; Tables 13 and 15 report all nineteen contrasts, and truncation removes no weight mass.

RGB reference composition. Each weighting is evaluated by one resample, of 5,000 images on CelebA and 1,024 on CIFAR-10, so SEs include observation, fitting and resampling variation. Under

biased selection, SR improves mean FID/KID and coverage over random references matched in count (Table 14). Reducing count alone largely offsets the coverage difference on CIFAR-10; CelebA instead has lower precision with SR. Under constant selection, SR raises mean FID/KID and lowers precision and coverage on both datasets. These comparisons contain no generated images.

Table 13: MNIST reference comparisons: first minus second, mean ± paired SE over five observation repetitions with twenty resamples per weighting. MMD squared is divided by 1,000; precision and coverage use pp. Random references match selected counts.

Selection	Comparison	MMD^2	Precision	Coverage
Constant	Selected − Random subset	$-0.94_{\pm 0.79}$	$-1.12_{\pm 1.20}$	$0.03_{\pm 0.25}$
Constant	Random subset − All records	$1.48_{\pm 0.70}$	$0.15_{\pm 0.76}$	$-0.36_{\pm 0.16}$
Constant	Selected − All records	$0.53_{\pm 1.10}$	$-0.97_{\pm 1.00}$	$-0.33_{\pm 0.17}$
Strong	Selected − Random subset	$-13.63_{\pm 4.64}$	$1.08_{\pm 0.68}$	$2.89_{\pm 0.70}$
Strong	Random subset − All records	$0.00_{\pm 3.33}$	$0.15_{\pm 0.45}$	$0.07_{\pm 0.76}$
Strong	Selected − All records	$-13.63_{\pm 3.58}$	$1.23_{\pm 0.34}$	$2.96_{\pm 0.53}$

Table 14: RGB weighted resampling: first reference minus second, mean ± paired SE. KID is multiplied by 1,000; precision and coverage use pp. The contrasts share the random reference.

Selection	Contrast	FID	KID $\times 10^3$	Precision	Coverage
CelebA					
Constant	Random minus All	$-0.07_{\pm 0.06}$	$-0.01_{\pm 0.06}$	$+0.05_{\pm 0.59}$	$+0.54_{\pm 0.18}$
	Selected minus Random	$+0.25_{\pm 0.13}$	$+0.17_{\pm 0.12}$	$-0.48_{\pm 0.63}$	$-0.06_{\pm 0.24}$
Biased	Random minus All	$-0.72_{\pm 0.60}$	$-0.95_{\pm 0.81}$	$+0.22_{\pm 0.21}$	$+0.27_{\pm 0.59}$
	Selected minus Random	$-1.88_{\pm 0.32}$	$-2.92_{\pm 0.42}$	$-1.60_{\pm 0.49}$	$+1.86_{\pm 0.30}$
CIFAR-10					
Constant	Random minus All	$-0.71_{\pm 1.07}$	$-0.02_{\pm 0.25}$	$+1.72_{\pm 0.88}$	$+0.76_{\pm 0.52}$
	Selected minus Random	$+2.08_{\pm 1.47}$	$+0.29_{\pm 0.37}$	$-3.71_{\pm 1.17}$	$-0.43_{\pm 0.48}$
Biased	Random minus All	$+1.52_{\pm 1.12}$	$+0.74_{\pm 0.29}$	$-1.43_{\pm 2.23}$	$-1.46_{\pm 0.60}$
	Selected minus Random	$-2.62_{\pm 0.99}$	$-1.35_{\pm 0.35}$	$+0.61_{\pm 1.60}$	$+1.68_{\pm 0.69}$

Table 15: MNIST model and optimization comparisons under strong selection: first estimate minus second, mean ± paired SE. MMD squared is divided by 1,000; precision and coverage use pp. Counts denote indicators; all 8,000 images remain the ratio reference. SR uses 8,000 indicators.

Comparison	MMD^2	Precision	Coverage
Probability models and indicator counts			
2 groups − 4 groups (500)	$-0.34_{\pm 0.72}$	$0.53_{\pm 0.38}$	$-0.28_{\pm 0.20}$
2 groups − 4 groups (2,000)	$1.05_{\pm 0.29}$	$-0.69_{\pm 1.02}$	$-0.30_{\pm 0.21}$
2 groups − 4 groups (8,000)	$1.52_{\pm 1.60}$	$-0.49_{\pm 0.41}$	$-0.50_{\pm 0.48}$
IPW (4 groups): 8,000 − 500	$-0.39_{\pm 0.42}$	$0.83_{\pm 0.18}$	$0.08_{\pm 0.16}$
IPW (4 groups): 8,000 − 2,000	$0.38_{\pm 0.58}$	$1.20_{\pm 0.41}$	$0.10_{\pm 0.24}$
SR − IPW (4 groups)	$0.78_{\pm 1.40}$	$-0.24_{\pm 0.71}$	$-0.20_{\pm 0.36}$
IPW (2 groups): 8,000 − 500	$1.47_{\pm 1.39}$	$-0.19_{\pm 0.64}$	$-0.14_{\pm 0.40}$
IPW (2 groups): 8,000 − 2,000	$0.85_{\pm 0.98}$	$1.40_{\pm 0.89}$	$-0.10_{\pm 0.37}$
SR − IPW (2 groups)	$-0.73_{\pm 2.22}$	$0.25_{\pm 0.68}$	$0.30_{\pm 0.55}$
Optimization sensitivity: second minus first optimization seed			
SR	$0.44_{\pm 1.85}$	$0.65_{\pm 0.63}$	$0.06_{\pm 0.31}$
Random subset	$-1.18_{\pm 5.31}$	$-0.67_{\pm 1.63}$	$0.00_{\pm 0.90}$
IPW (4 groups)	$1.19_{\pm 0.84}$	$0.18_{\pm 0.64}$	$-0.47_{\pm 0.24}$
IPW (2 groups)	$0.73_{\pm 0.91}$	$1.41_{\pm 0.28}$	$-0.44_{\pm 0.46}$

Probability estimation. A separate simulation draws 200 independent record samples for each of six selection/count conditions and fits both probability models, yielding 2,400 estimates. Group frequencies equal those in U; comparisons across counts use nested records. Table 16 averages squared error over these frequencies. Under strong selection, estimation with two groups approaches the projection error $29/1600$: additional records reduce variance but cannot separate merged probabilities.

Table 16: Selection probability MSE multiplied by 10,000: mean ± Monte Carlo SE over 200 record samples per condition and count.

Selection	Records	Four groups	Two groups
Constant	500	$20.07_{\pm 1.03}$	$9.48_{\pm 0.61}$
Constant	2,000	$4.69_{\pm 0.23}$	$2.20_{\pm 0.15}$
Constant	8,000	$1.17_{\pm 0.06}$	$0.60_{\pm 0.04}$
Strong	500	$12.86_{\pm 0.76}$	$188.63_{\pm 0.50}$
Strong	2,000	$3.26_{\pm 0.18}$	$183.07_{\pm 0.13}$
Strong	8,000	$0.83_{\pm 0.05}$	$181.73_{\pm 0.03}$

D.8 Violated observation assumptions

Both studies use four equally probable types, $P_i = 1/4$, with $\eta_i = \Pr(Y = 1 \mid X = i) = (0.2, 0.4, 0.6, 0.8)_i$, so $P_+ = (0.1, 0.2, 0.3, 0.4)$, and complete reporting, $c_{\text{rep}} = 1$. Unlike the image experiments, each type contains both classes, so $A \perp\!\!\!\perp Y \mid X$ can fail.

Proposition D.1 (Undetectable violations)**.** *Let L, U and V be independent samples of fixed sizes from P_l, P and records with covariate distribution P and selection probability $v = \Pr(A = 1 \mid X) > 0$, however P_l was produced. Suppose $\mathbb{E}_l v^{-1} < \infty$ and $Q = (v^{-1}\mathrm{d}P_l)/\mathbb{E}_l v^{-1}$ has bounded density with respect to P. Then a model satisfying Assumption 1, with positive distribution Q and selection probability v for both classes, generates the same distribution of (L, U, V). SR and IPW with the true v both target Q, and any test of size at most δ_0 uniformly over models satisfying the assumption has power at most δ_0 against the original model.*

Proof. Choose $\alpha' > 0$ with $\alpha' \operatorname{ess\,sup}_P \mathrm{d}Q/\mathrm{d}P \le 1$ and put $P'_- = (P - \alpha' Q)/(1 - \alpha')$. Under the new model, labeled positives follow $vQ/\mathbb{E}_Q v = P_l$, since $\mathbb{E}_Q v = 1/\mathbb{E}_l v^{-1}$; U and V are unchanged. Equation (B.7) identifies Q under this model, and equal observation distributions give the power bound as in Proposition B.1. □

Mismatched records. Selection probabilities $v_0 = (0.2, 0.5, 0.5, 0.8)$ give $P_l = (2, 10, 15, 32)/59$. Records instead follow

$$v_\delta = (0.2,\ 0.5 + \delta,\ 0.5 - \delta,\ 0.8), \qquad \delta \in \{-0.30, -0.25, \ldots, 0.30\}.$$

Thus P, P_l, mean selection probability 0.5, and bounds $0.2 \le v_\delta \le 0.8$ are unchanged; only the middle types' probabilities vary. The corrected population is $Q_\delta(i) \propto P_+(i) v_0(i)/v_\delta(i)$. Negative δ overweights type 2; positive δ overweights type 3. Their unequal target masses, 0.20 and 0.30, and normalization explain the asymmetry between extreme mismatches.

Each of 1,000 independent repetitions draws $n_L = 1{,}000$ labeled positives and nested samples with $n_V \in \{500, 2{,}000, 8{,}000\}$. Shared random numbers keep the selected count fixed across mismatch conditions. With labeled counts N_i, selected counts M_i, and pseudocount $h = 0.5$, the estimates are

$$\widehat{p}_{l,i} = \frac{N_i + h}{n_L + 4h}, \qquad \widehat{Q}(\{i\}) = \frac{P_i (N_i + h)/(M_i + h)}{\sum_{j=1}^4 P_j (N_j + h)/(M_j + h)}.$$

All 39 finite conditions use only these counts and the known type masses P_i.

The paired difference is $D = \mathrm{TV}(\widehat{Q}, P_+) - \mathrm{TV}(\widehat{P}_l, P_+)$; negative values favor correction. The infinite record comparison retains $\widehat{P}_l$ and uses $\widehat{Q}_\infty(\{i\}) \propto \widehat{p}_{l,i}/v_{\delta,i}$. Table 17 reports every finite sample condition and the analytic limit using exact P_l. Both extremes remain adverse as $n_V \to \infty$, whereas some intermediate mismatches retain improvement.

Table 17: Record mismatch, measured by $100D$. Columns for finite L report mean $\pm$ Monte Carlo SE over 1,000 paired repetitions, retaining each labeled sample across record counts. The last column uses exact P_l. Negative values favor correction.

	Finite labeled sample				Exact P_l
δ	$n_V = 500$	$n_V = 2{,}000$	$n_V = 8{,}000$	$n_V \to \infty$	$n_V \to \infty$
-0.30	$8.53_{\pm 0.19}$	$8.11_{\pm 0.12}$	$7.93_{\pm 0.10}$	$7.96_{\pm 0.10}$	7.87
-0.25	$3.28_{\pm 0.16}$	$2.59_{\pm 0.11}$	$2.38_{\pm 0.10}$	$2.39_{\pm 0.09}$	2.13
-0.20	$-0.77_{\pm 0.13}$	$-1.61_{\pm 0.10}$	$-1.84_{\pm 0.09}$	$-1.88_{\pm 0.09}$	-2.42
-0.15	$-3.97_{\pm 0.12}$	$-4.86_{\pm 0.09}$	$-5.12_{\pm 0.08}$	$-5.18_{\pm 0.08}$	-6.13
-0.10	$-6.61_{\pm 0.11}$	$-7.62_{\pm 0.08}$	$-7.89_{\pm 0.07}$	$-7.94_{\pm 0.07}$	-9.24
-0.05	$-8.53_{\pm 0.09}$	$-9.97_{\pm 0.07}$	$-10.36_{\pm 0.07}$	$-10.43_{\pm 0.06}$	-11.65
0.00	$-9.26_{\pm 0.09}$	$-11.01_{\pm 0.06}$	$-11.60_{\pm 0.06}$	$-11.78_{\pm 0.06}$	-14.24
0.05	$-8.50_{\pm 0.10}$	$-9.83_{\pm 0.08}$	$-10.22_{\pm 0.07}$	$-10.36_{\pm 0.07}$	-11.40
0.10	$-6.46_{\pm 0.12}$	$-7.28_{\pm 0.09}$	$-7.51_{\pm 0.08}$	$-7.62_{\pm 0.08}$	-8.24
0.15	$-3.56_{\pm 0.14}$	$-4.05_{\pm 0.10}$	$-4.21_{\pm 0.09}$	$-4.28_{\pm 0.09}$	-4.64
0.20	$0.19_{\pm 0.15}$	$-0.12_{\pm 0.11}$	$-0.24_{\pm 0.09}$	$-0.28_{\pm 0.09}$	-0.49
0.25	$4.86_{\pm 0.17}$	$4.67_{\pm 0.11}$	$4.55_{\pm 0.10}$	$4.51_{\pm 0.09}$	4.41
0.30	$10.50_{\pm 0.19}$	$10.45_{\pm 0.12}$	$10.37_{\pm 0.10}$	$10.36_{\pm 0.09}$	10.31

Class dependent selection. Each type keeps selection probability $v_0(i)$, while its positives and negatives are selected with probabilities

$$v_-(i) = \frac{v_0(i)}{1 + \eta_i(k-1)}, \qquad v_+(i) = k\, v_-(i),$$

for the twelve values of k in Table 18. Thus $\eta_i v_+(i) + (1-\eta_i)v_-(i) = v_0(i)$, and $k = 1$ satisfies Assumption 1. The endpoints are the feasible limits: $v_-(4) = 1$ at $k = 0.75$ and $v_+(2) = 1$ at $k = 6$. Labeled positives follow $P_l(i) \propto P_+(i)v_+(i)$, whereas the records have the same distribution for every k. By Proposition D.1, SR and IPW with the true v_0 target

$$Q_k(i) \propto \frac{P_l(i)}{v_0(i)} \propto P_+(i)\, \frac{k}{1 + \eta_i(k-1)},$$

which does not depend on v_0.

Each of 1,000 repetitions draws $n_L = 1{,}000$ labeled positives from P_l and nested records with $X \sim P$, $Y \sim \text{Bernoulli}(\eta_X)$ and $A \sim \text{Bernoulli}\{v_Y(X)\}$; estimators receive only X and A. Random numbers are shared across k. Estimates follow the mismatch study, and the infinite record column uses $\widehat{Q}_\infty(\{i\}) \propto \widehat{p}_{l,i}/v_{0,i}$, which equals IPW with the true probabilities. Uncorrected weighting becomes closer to P_+ beyond $k = 2.94$ in the population limit and $k = 2.85$ with 8,000 records, interpolating Table 18. It improves with k because $P_l(i) \propto P_+(i)v_0(i)\, k/\{1 + \eta_i(k-1)\}$, whose last factor decreases with η_i, whereas v_0 increases with η_i here.

D.9 Generation and weighted resampling

Each generated set is compared with one independent resample of U under the same normalized and truncated weights: 1,024 images on MNIST and CIFAR-10 and 5,000 on CelebA. Resampling uses random numbers shared across methods and independent of the diffusion noise, and both samples use the same encoder and references. Table 19 reports SR in all seven selection conditions. On MNIST, every guided method has higher mean coverage and lower precision than resampling, and SR also has higher MMD squared in all three conditions. On CelebA, all three guided methods have higher mean precision, lower coverage and higher FID/KID in both conditions; on CIFAR-10, all three have higher mean FID/KID and lower precision and coverage. With the weights fixed, these differences measure generation error as a whole, without separating the base model, potential and sampler.

D.10 Denoising and guidance approximations

These MNIST diagnostics use known weights, which separates generation error from weight estimation. They are exploratory: they were chosen while developing the sampler, use pixels only and rest on single draws, so they add no confirmatory evidence to the PU comparisons.

Table 18: Class dependent selection, measured by TV to P_+ ($\times 100$). Columns for finite L report mean $\pm$ Monte Carlo SE over 1,000 repetitions, retaining each labeled sample across record counts. The last two columns use exact P_l.

	Finite labeled sample				Exact P_l	
	Uncorrected	SR			Uncorrected	SR
k		$n_V = 500$	$n_V = 8{,}000$	$n_V \to \infty$		$n_V \to \infty$
0.75	$16.33_{\pm 0.05}$	$5.14_{\pm 0.07}$	$3.40_{\pm 0.04}$	$3.29_{\pm 0.04}$	16.45	2.36
0.80	$15.81_{\pm 0.05}$	$5.00_{\pm 0.07}$	$3.08_{\pm 0.04}$	$2.95_{\pm 0.04}$	15.95	1.82
0.90	$14.92_{\pm 0.05}$	$4.83_{\pm 0.07}$	$2.71_{\pm 0.04}$	$2.54_{\pm 0.03}$	15.04	0.85
1.00	$14.13_{\pm 0.05}$	$4.81_{\pm 0.07}$	$2.63_{\pm 0.04}$	$2.44_{\pm 0.03}$	14.24	0.00
1.25	$12.53_{\pm 0.05}$	$5.19_{\pm 0.08}$	$3.24_{\pm 0.04}$	$3.08_{\pm 0.04}$	12.59	1.80
1.50	$11.31_{\pm 0.05}$	$5.83_{\pm 0.08}$	$4.24_{\pm 0.05}$	$4.12_{\pm 0.05}$	11.31	3.28
2.00	$9.80_{\pm 0.05}$	$7.32_{\pm 0.09}$	$6.17_{\pm 0.05}$	$6.08_{\pm 0.05}$	9.43	5.57
2.50	$9.00_{\pm 0.04}$	$8.66_{\pm 0.10}$	$7.75_{\pm 0.05}$	$7.68_{\pm 0.05}$	8.79	7.28
3.00	$8.48_{\pm 0.05}$	$9.77_{\pm 0.10}$	$8.99_{\pm 0.06}$	$8.93_{\pm 0.05}$	8.42	8.61
4.00	$7.86_{\pm 0.05}$	$11.47_{\pm 0.11}$	$10.84_{\pm 0.06}$	$10.78_{\pm 0.05}$	7.87	10.56
5.00	$7.47_{\pm 0.05}$	$12.71_{\pm 0.11}$	$12.14_{\pm 0.06}$	$12.09_{\pm 0.06}$	7.49	11.93
6.00	$7.16_{\pm 0.05}$	$13.60_{\pm 0.11}$	$13.13_{\pm 0.06}$	$13.08_{\pm 0.06}$	7.20	12.94

Table 19: SR generation minus resampling with the same weights in all seven conditions, against population positives. Units follow Table 7; positive fraction differences use pp (even digits on MNIST, Smiling on CelebA, vehicles on CIFAR-10).

Selection	Fréchet	Discrepancy	Precision	Coverage	Positive
MNIST					
Constant	$+185.04_{\pm 29.87}$	$+6.61_{\pm 1.61}$	$-15.98_{\pm 1.60}$	$+1.25_{\pm 0.76}$	$-25.06_{\pm 2.16}$
Weak	$+141.88_{\pm 35.30}$	$+3.75_{\pm 1.99}$	$-14.77_{\pm 0.97}$	$+1.88_{\pm 1.08}$	$-23.95_{\pm 1.54}$
Strong	$+177.26_{\pm 24.87}$	$+5.99_{\pm 1.18}$	$-14.90_{\pm 0.75}$	$+2.05_{\pm 1.39}$	$-23.50_{\pm 1.49}$
CelebA					
Constant	$+7.61_{\pm 0.33}$	$+9.33_{\pm 0.40}$	$+6.49_{\pm 0.40}$	$-11.24_{\pm 0.60}$	$+0.22_{\pm 0.70}$
Biased	$+7.58_{\pm 0.21}$	$+9.27_{\pm 0.23}$	$+6.04_{\pm 0.67}$	$-10.89_{\pm 0.29}$	$-0.93_{\pm 0.62}$
CIFAR-10					
Constant	$+10.27_{\pm 1.00}$	$+5.86_{\pm 0.57}$	$-5.82_{\pm 1.25}$	$-2.58_{\pm 1.08}$	$-11.46_{\pm 1.09}$
Biased	$+11.54_{\pm 0.94}$	$+5.89_{\pm 0.58}$	$-8.73_{\pm 1.10}$	$-3.61_{\pm 0.86}$	$-12.01_{\pm 1.22}$

A known conditional distribution. Let P_U put equal mass on the 30,000 images in U, and let $Q_{a,U} = aP_U/\mathbb{E}_{P_U} a$. Gaussian corruption gives a finite mixture whose conditional mean can be computed directly. Its exact noise predictor and correction satisfy

$$\boldsymbol{\epsilon}_a^*(t,x) = \frac{x - \sqrt{\bar{\alpha}_t}\mathbb{E}_{Q_{a,U}}[X_0 \mid X_t = x]}{\sigma_t}, \qquad \nabla\phi_a(t,x) = \frac{\boldsymbol{\epsilon}_1^*(t,x) - \boldsymbol{\epsilon}_a^*(t,x)}{\sigma_t}.$$

Weights are constant or proportional to $(0.1, 0.25, 0.55, 0.9)^{-1}$ across ascending intensity quartiles, so the targets require no labels. Quartiles of mean intensity and vertical ink centroid define 16 regions, over which TV compares generated and target histograms.

Replacing the two approximations. We replace the learned base predictor, the guidance, or both by exact conditionals, generating 1,024 images with 1,000 steps and posterior reverse variance. Unlike in the main comparison, predicted clean images are clipped to $[-1, 1]$ at each step. The base model is trained for 200,000 iterations and the potential for 100,000; exact quantities are computed in double precision and sampling in single precision.

Neither replacement alone lowers histogram TV in this draw; replacing both lowers it from 13.82 to 4.56 (Table 20), and only this joint replacement meets all 16 binomial intervals of the sampling check, each at level $1/640$. With constant weights, exact and learned predictions give TV 5.37 and 10.06. At low noise, exact conditioning on the finite U concentrates on individual images of U, so a large denoising error need not mean poor visual quality.

Table 20: Exact versus learned components under known nonconstant weights: one paired draw of 1,024 images. TV compares generated and target histograms over 16 pixel regions; exact quantities use the finite U.

Base predictor	Guidance	Histogram TV $\times 100$	Regions outside intervals
Learned	Learned	13.82	3
Learned	Exact	14.51	4
Exact	Learned	17.66	8
Exact	Exact	4.56	0

Training length and numerical precision. Potentials trained for 20,000 and 100,000 iterations are compared on 1,024 shared saved states at nine steps of the learned weighted sampler, by squared correction error per pixel against $\sigma_t \nabla \phi_a$. Longer training reduces the mean error by 34.2%, 59.9% and 66.1% at steps 600, 800 and 1000, but only by 2.4% at step 400, and raises it by 0.2% at step 200. Full float32 instead of TensorFloat-32 in the base model changes the mean denoising error by at most 1.43% on 9,216 saved states.

Variation in the training target. Under forward corruption of images in U, sampled noise $\boldsymbol{z}$ has conditional mean $\boldsymbol{\epsilon}_1^*(t, X_t)$. For $d = 784$ pixels, let $v_t = \mathbb{E}\|\boldsymbol{z} - \boldsymbol{\epsilon}_1^*(t, X_t)\|^2/d$ and $e_t = \mathbb{E}\|\widehat{\boldsymbol{\epsilon}}(t, X_t) - \boldsymbol{\epsilon}_1^*(t, X_t)\|^2/d$. Then $\mathbb{E}\|\widehat{\boldsymbol{\epsilon}}(t, X_t) - \boldsymbol{z}\|^2/d = e_t + v_t$. On 1,024 queries per level, v_t/e_t is 0.29, 88.25 and 81.39 at steps 400, 600 and 800, so conditional target variance can hide small fitting improvements at high noise. Replacing sampled noise by $\boldsymbol{\epsilon}_1^*(t, X_t)$ removes v_t from expected squared loss without changing its expected parameter gradient. For fixed weights, put $m_t(X_t) = \mathbb{E}_{P_U}[a(X_0) \mid X_t]$. For $h = \phi_\psi(t, X_t)$, linearity in the weight gives $\mathbb{E}[a(X_0)\ell(h) + \ell(-h) \mid X_t] = m_t(X_t)\ell(h) + \ell(-h)$, so conditional weights preserve the logistic objective while removing weight variation at fixed X_t; retraining with these targets has not been evaluated.

D.11 Sensitivity to the evaluation reference

With images, models, encoder and references fixed, let $D_+ = M_{P_+}(\text{method}) - M_{P_+}(\text{uncorrected})$ for a metric M, and define D_l with P_l; SEs exclude reference sampling (Table 21).

Table 21: SR minus uncorrected generation against labeled positives (D_l), and its paired change from population positives ($D_l - D_+$). Discrepancy units follow Table 7; precision and coverage use pp.

Selection	Discrepancy D_l	Discrepancy $D_l - D_+$	Precision D_l	Precision $D_l - D_+$	Coverage D_l	Coverage $D_l - D_+$
MNIST						
Constant	$+1.36_{\pm 2.14}$	$-0.55_{\pm 0.22}$	$-3.46_{\pm 1.66}$	$+0.18_{\pm 0.32}$	$-0.04_{\pm 0.94}$	$+0.12_{\pm 0.45}$
Weak	$+3.09_{\pm 1.53}$	$+0.41_{\pm 0.26}$	$-2.17_{\pm 1.19}$	$+0.37_{\pm 0.35}$	$+0.10_{\pm 0.47}$	$+0.70_{\pm 0.24}$
Strong	$+23.97_{\pm 2.42}$	$+35.08_{\pm 7.32}$	$-3.59_{\pm 0.48}$	$+0.47_{\pm 0.46}$	$-2.17_{\pm 0.52}$	$-6.74_{\pm 1.64}$
CelebA						
Constant	$+0.26_{\pm 0.19}$	$-0.02_{\pm 0.03}$	$+0.01_{\pm 0.18}$	$+0.14_{\pm 0.32}$	$+0.12_{\pm 0.32}$	$+0.16_{\pm 0.22}$
Biased	$+2.75_{\pm 0.41}$	$+4.86_{\pm 0.32}$	$-0.47_{\pm 0.30}$	$-0.17_{\pm 0.27}$	$-1.39_{\pm 0.59}$	$-1.33_{\pm 0.51}$
CIFAR-10						
Constant	$+0.21_{\pm 0.22}$	$-0.04_{\pm 0.05}$	$-1.00_{\pm 0.42}$	$-0.64_{\pm 0.54}$	$-0.12_{\pm 0.36}$	$-0.45_{\pm 0.44}$
Biased	$+0.65_{\pm 0.38}$	$+1.45_{\pm 0.22}$	$-2.58_{\pm 1.21}$	$-0.49_{\pm 0.69}$	$-2.30_{\pm 0.43}$	$-2.71_{\pm 0.68}$

With labeled positives as reference, strong MNIST's coverage difference reverses from 4.57 ± 1.63 to -2.17 ± 0.52 pp, as does its MMD squared difference; constant and weak selection show no general reversal. Biased CelebA's FID and KID differences change from -1.51 ± 0.23 and $(-2.11 \pm 0.31) \times 10^{-3}$ to 2.04 ± 0.28 and $(2.75 \pm 0.41) \times 10^{-3}$, and biased CIFAR-10's coverage and KID differences reverse sign, while precision stays lower in all three cases. With the images unchanged, the reference decides which composition is rewarded.