Correct your balance heuristic: Optimizing balance-style multiple importance sampling weights

QINGQIN HUA, Saarland University, Germany
PASCAL GRITTMANN, Saarland University, Germany
PHILIPP SLUSALLEK, Saarland University, Germany and DFKI, Germany

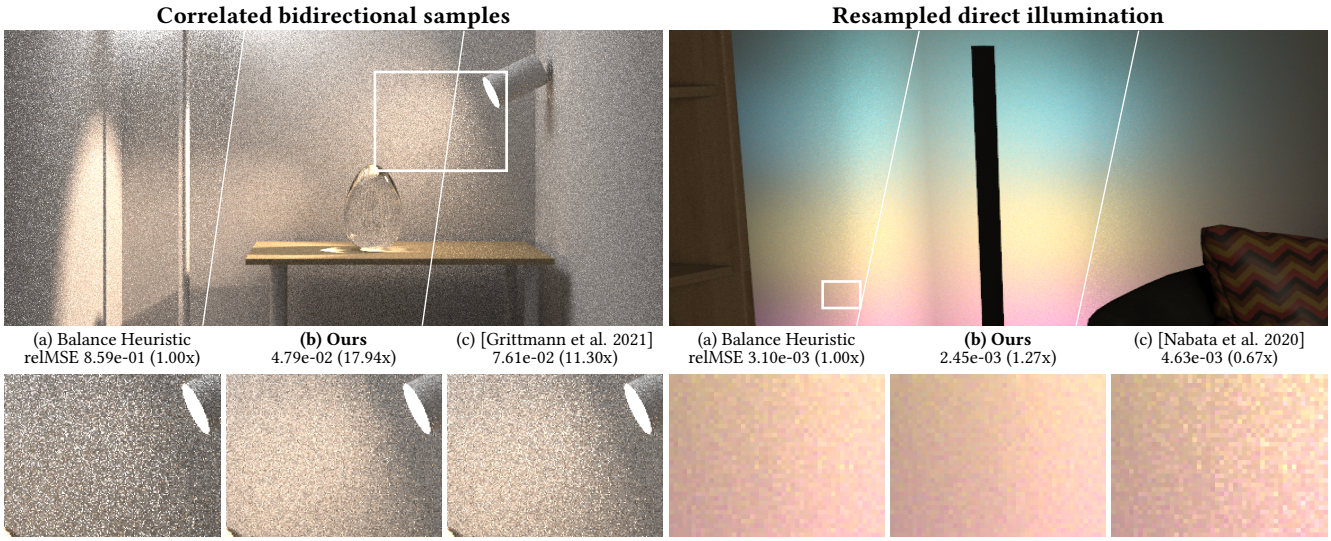

Fig. 1. Multiple importance sampling (MIS) with the common balance or power heuristic can provide inferior results. Here, we show examples from bidirectional rendering in global illumination (left) and resampled importance sampling in direct illumination (right). In both cases, prior heuristics produce unnecessary noise, as quantified by the relative mean-squared error (relMSE). We propose a simple and practical correction-factor-optimization, applicable on top of any MIS weighting heuristic, that yields consistently lower estimator noise.

Multiple importance sampling (MIS) is a vital component of most rendering algorithms. MIS computes a weighted sum of samples from many different techniques to achieve generalization, that is, to handle a wide range of scene types and lighting effects. A key factor to the performance of MIS is the choice of weighting function. The go-to default – the balance heuristic – performs well in many cases, but prior work has shown that it can yield unsatisfactory results. A number of challenges cause this suboptimal performance, including low-variance techniques, sample correlation, and unknown sampling densities. Prior work has suggested improvements for some of these problems, but a general optimal solution has yet to be found. We propose a general and practical weight correction scheme: We optimize, on-the-fly, a set of correction factors that are multiplied into any baseline MIS heuristic (e.g., balance or power). We demonstrate that this approach yields consistently better equal-time performance on two rendering applications: bidirectional algorithms and resampled importance sampling for direct illumination.

Authors' Contact Information: Qingqin Hua, Saarland University, Saarbrücken, Germany, hua@cg.uni-saarland.de; Pascal Grittmann, Saarland University, Saarbrücken, Germany, grittmann@cg.uni-saarland.de; Philipp Slusallek, Saarland University, Saarbrücken, Germany and DFKI, Saarbrücken, Germany, philipp.slusallek@dfki.de.

ACM 1557-7368/2025/8-ART98
https://doi.org/10.1145/3730819

CCS Concepts: • **Computing methodologies** → **Rendering**; **Ray tracing**.

Additional Key Words and Phrases: ray tracing, global illumination, Monte Carlo, multiple importance sampling

ACM Reference Format:
Qingqin Hua, Pascal Grittmann, and Philipp Slusallek. 2025. Correct your balance heuristic: Optimizing balance-style multiple importance sampling weights. *ACM Trans. Graph.* 44, 4, Article 98 (August 2025), 14 pages. https://doi.org/10.1145/3730819

1 Introduction

Rendering algorithms are faced with a wide range of scene types and light transport effects that must be simulated. To handle this diverse input, virtually all renderers utilize the same fundamental tool: Multiple importance sampling (MIS). MIS allows us to devise general solutions by combining many Monte Carlo sampling techniques into one joint algorithm. Thus, the combined algorithm will ideally have at least one well-suited technique for every scene type or light effect that it might encounter.

A key ingredient to the performance of MIS is the choice of weighting function. Each Monte Carlo sample in an MIS combination is multiplied by this weighting function, such that the total sum of all sample weights yields the desired, unbiased pixel estimate. Ideally, the MIS weight should be highest for the samples of the

"best" technique, i.e., the technique with lowest variance. Unfortunately, maintaining this ideal is easier said than done with common weighting heuristics such as the balance or power heuristic.

Previous work has identified a number of challenging cases, such as low-variance weighting issues [Veach and Guibas 1995; Grittmann et al. 2019], issues with sample correlation [Popov et al. 2015; Grittmann et al. 2021], or issues from unknown sample distributions [Kelemen et al. 2002; Nabata et al. 2020]. Solutions to these challenges have been proposed, but these are usually problem-specific, sometimes application-specific, and always suboptimal.

Optimal MIS weights have been derived [Kondapaneni et al. 2019] and adapted such that they can be applied to more rendering use-cases [Hua et al. 2023]. However, these require simplifying assumptions, namely uncorrelated samples and known, closed-form PDFs of all samples, limiting their applicability.

As a *general* and *practical* alternative, we propose a simple but effective weight correction scheme for MIS weighting. Specifically, we optimize correction factors that are multiplied into the MIS weight. To achieve practicality, we utilize a direct search over a tiny set of candidate correction factors, paired with aggressive filtering. We test our method on two rendering applications: Correcting for correlation in bidirectional algorithms, and correcting for unknown sampling densities in resampled direct illumination. In both cases, we achieve *consistent* equal-time performance gains over the baseline: Our method was never noticeably worse than the baseline and up to two times better in some cases (see Fig. 1 for an example).

In the following, we first recap the essential background (Section 2) and discuss the main MIS weighting challenges in typical rendering applications (Section 3). The subsequent sections introduce our method and evaluate it on the two rendering applications: bidirectional rendering with vertex connection and merging (VCM) [Georgiev et al. 2012; Hachisuka et al. 2012] and resampled importance sampling [Talbot et al. 2005] applied to direct illumination computation (inspired by ReSTIR [Bitterli et al. 2020] but simplified).

2 Preliminaries

Monte Carlo integration is the prevalent technique to numerically compute integrals such as the rendering equation [Kajiya 1986]. A Monte Carlo estimator

$$\langle F \rangle = \sum_{i=1}^{n} \frac{f(x_i)}{np(x_i)} \approx \int_X f(x)\,\mathrm{d}x = F \tag{1}$$

estimates the integral F by averaging the weights of n samples x_i following a probability density (PDF) $p(x)$. A well-chosen PDF is key to efficient Monte Carlo integration: The better $p(x)$ matches the shape of the integrand $f(x)$, the fewer samples are required.

The quality of a Monte Carlo estimate can be measured by its variance: the difference between the expectation of the square and the square of the expectation:

$$\mathbb{V}\left[\langle F \rangle\right] = \mathbb{E}\left[\left(\sum_{i=1}^{n} \frac{f(x_i)}{np(x_i)}\right)^2\right] - F^2. \tag{2}$$

In rendering, this variance manifests as noise in the image.

2.1 Multiple importance sampling (MIS)

Unfortunately, it is difficult to find a single PDF that achieves satisfactory performance. *Multiple importance sampling* (MIS) [Veach and Guibas 1995] therefore combines samples from multiple techniques $t \in \mathcal{T}$, each using a different PDF p_t. The MIS estimator

$$\langle F \rangle_{\mathrm{MIS}} = \sum_{t \in \mathcal{T}} \sum_{i=1}^{n_t} w_t(x_{t,i}) \frac{f(x_{t,i})}{n_t p_t(x_{t,i})} \tag{3}$$

sums over all samples from all techniques and multiplies their usual Monte Carlo weights by an MIS weighting function $w_t(x)$. The performance of MIS hinges on the choice of techniques, sample counts, and weighting function. We focus on the latter.

Choosing techniques. MIS is frequently applied to mix any techniques that are readily available. For instance, typical unidirectional path tracers mix BSDF importance sampling and light source sampling via MIS [Veach and Guibas 1995; Pharr et al. 2023, Chapter 13.4]. Bidirectional methods, like the vertex connection and merging (VCM) algorithm [Georgiev et al. 2012; Hachisuka et al. 2012], combine an enormous set of sampling techniques – namely, every possible combination of camera prefix and light suffix paths through connection with shadow rays or merging via photon density estimation. However, combining all techniques at ones disposal can be suboptimal: Previous work has shown that the efficiency of MIS benefits greatly from avoiding redundancy in the techniques [Karlík et al. 2019]. Finding a good set of techniques that complement each other well thus remains an important problem. For example, Grittmann et al. [2018] rely on the MIS weights w_t to prune sampling of the expensive photon mapping [Jensen 1996] technique where it is not absolutely necessary. Approaches like that stand to benefit from improved MIS weighting functions.

Optimal sample counts. Another key aspect is the number of samples n_t invested in each technique: Ideally, computation should be spent on the most beneficial techniques. Prior work has shown that this can be achieved through heuristics [Pajot et al. 2010; Grittmann et al. 2018] or direct optimization [Lu et al. 2013; Sbert et al. 2019; Murray et al. 2020; Müller 2019] and even for complex many-technique combinations like VCM [Grittmann et al. 2022]. The MIS weighting and the chosen sample counts influence each other; ideally, they should be optimized jointly [He and Owen 2014]. Our direct search optimization is similar to the one used by Grittmann et al. [2022], so joint optimization could be possible.

MIS weighting functions. Often, the *balance heuristic* [Veach and Guibas 1995] is a great choice for the weighting function. Its weights are proportional to the effective density of each technique,

$$w_t^{\mathrm{bal}}(x) = \frac{n_t p_t(x)}{\sum_k n_k p_k(x)} \propto n_t p_t(x). \tag{4}$$

Intuitively, this means that "good" high-density samples will receive a higher weight, while "bad" low-density samples – which are apt to cause outliers – will be weighted down. The balance heuristic has many desirable properties. Most importantly, by relying only on sample counts and PDFs, it is cheap and easy to compute, and the variance of the resulting MIS estimator has theoretically-proven

upper bounds [Veach and Guibas 1995]. Nevertheless, the balance heuristic can perform badly – sometimes surprisingly so.

3 Challenges for MIS weighting

MIS weighting with the prevalent balance heuristic faces three main challenges: low-variance techniques, correlated samples, and unknown PDFs. Solutions from prior work address different subsets of these: Optimal MIS [Kondapaneni et al. 2019] can solve the low-variance issue, ad-hoc corrections for correlated samples have been proposed [Grittmann et al. 2021; Jendersie and Grosch 2018], and workarounds for unknown PDFs exist [Nabata et al. 2020; Kelemen et al. 2002]. We propose a *practical*, *simple*, and *general* MIS weight correction that addresses all three challenges.

3.1 Variance bounds of the balance heuristic

The balance heuristic minimizes an upper bound of the variance of the MIS estimator [Veach and Guibas 1995]. Namely, *assuming independent samples*, the variance of Eq. (3),

$$\mathbb{V}\left[\langle F\rangle_{\text{MIS}}\right] = \mathbb{E}\left[\left(\sum_{t\in\mathcal{T}}\sum_{i=1}^{n_t} w_t(x_{t,i})\frac{f(x_{t,i})}{n_t p_t(x_{t,i})}\right)^2\right] - F^2, \tag{5}$$

can be written as the difference between the sum of per-technique single-sample second moments, and the per-technique squared first moments [Veach 1997, Appendix 9.A, p. 288],

$$\mathbb{V}\left[\langle F\rangle_{\text{MIS}}\right] \overset{\text{iid}}{=} \sum_{t\in\mathcal{T}}\int_{\mathcal{X}}\frac{w_t^2 f^2}{n_t p_t}\,\mathrm{d}x - \sum_{t\in\mathcal{T}}\frac{1}{n_t}\left(\int_{\mathcal{X}} w_t f\,\mathrm{d}x\right)^2. \tag{6}$$

The balance heuristic minimizes the first term,

$$w_t^{\text{bal}} = \arg\min_{w_t}\sum_{t\in\mathcal{T}}\int_{\mathcal{X}}\frac{w_t^2 f^2}{n_t p_t}\,\mathrm{d}x. \tag{7}$$

Hence, the balance heuristic performs well whenever this sum of single-sample per-technique second moments dominates the variance. Next, we will outline three cases where this is not true.

3.2 Low-variance techniques

The first challenge occurs if one of the techniques has close to zero variance. Then, by definition, its single-sample second moment will be almost identical to its first moment,

$$\mathbb{V}\left[\langle F\rangle_t\right] \approx 0 \Rightarrow \int_{\mathcal{X}}\frac{w_t^2 f^2}{n_t p_t}\,\mathrm{d}x \approx \frac{1}{n_t}\left(\int_{\mathcal{X}} w_t f\,\mathrm{d}x\right)^2. \tag{8}$$

In the extreme case of zero variance, the second-moment that the balance heuristic optimizes will thus be infinitely bigger than the actual variance of the technique. So the balance heuristic "assumes" this technique to be much worse than it really is, and assigns a too low weight to it.

This challenge is the motivation behind the power, maximum, and cut-off heuristics [Veach and Guibas 1995]. The power heuristic, for example, distorts the balance heuristic by an exponent,

$$w_t^{\text{pow}}(x) \propto \left[n_t p_t(x)\right]^{\beta}, \tag{9}$$

typically, $\beta = 2$. This further increases the weight of (presumably) "good" techniques and hence it *can* ameliorate low-variance weighting issues. However, there is no guarantee for that, and, indeed, the worst-case variance bounds of the power heuristic are worse than those of the balance heuristic [Veach and Guibas 1995].

One option is to incorporate variance estimates of each technique into the balance heuristic, as done by *variance-aware MIS (VA-MIS)* [Grittmann et al. 2019],

$$w_t^{\text{var-aware}}(x) \propto v_t n_t p_t(x), \tag{10}$$

$$\text{where}\quad v_t = \frac{\int_{\mathcal{X}}\frac{f^2}{n_t p_t}\,\mathrm{d}x}{\mathbb{E}\left[\left(\sum_{i=1}^{n_t}\frac{f(x_{t,i})}{n_t p_t(x_{t,i})}\right)^2\right] - F^2} \tag{11}$$

is the ratio between the single-sample second moment of the technique (that the balance heuristic optimizes for (7)) and its full variance. This correction factor is designed to fulfill two properties: (1) If a technique has (close-to) zero variance, it will receive unit weight, $w_t(x) = 1$. (2) If multiple techniques have identical PDFs but different variances, their weight will be proportional to their reciprocal variances – the optimal constant weighting. Beyond these properties, VA-MIS remains heuristic in nature, can perform worse than the balance heuristic, and does not support biased techniques, i.e., cases where $p_t(x) = 0$ for some x where $f(x) \neq 0$.

Better results can be achieved via optimal MIS weights [Kondapaneni et al. 2019; Hua et al. 2023]: If samples are independent, and all PDFs known in closed-form, then (approximately) optimal MIS weights can be computed on-the-fly while rendering by estimating the coefficients for a linear system of equations and solving it. In addition to these assumptions, optimal MIS weighting is also non-trivial to implement, mostly because it requires to track zero-valued samples (i.e., $f(x) = 0$) and correctly compute their PDF values.

Our weight correction achieves comparable results to VA-MIS for the low-variance challenge, but offers more robust and reliable improvements because it is not a heuristic but a direct optimization.

3.3 Correlated samples

The balance heuristic's error bound was derived under the assumption of *independent samples*. Many rendering algorithms, however, introduce some amount of correlation. For example, photon mapping methods reuse the same camera prefix path across millions of full-path samples constructed via merges with all light sub-paths (most of which have zero contribution) [Georgiev et al. 2012; Hachisuka et al. 2012]. Then, the term the balance heuristic minimizes is off by another quantity: the sample covariance Cov (cf., Eq. (6)),

$$\mathbb{V}\left[\langle F\rangle_{\text{MIS}}\right] = \sum_{t\in\mathcal{T}}\int_{\mathcal{X}}\frac{w_t^2 f^2}{n_t p_t}\,\mathrm{d}x - \sum_{t\in\mathcal{T}}\frac{1}{n_t}\left(\int_{\mathcal{X}} w_t f\,\mathrm{d}x\right)^2 + \text{Cov}. \tag{12}$$

The exact equation of this covariance depends on the nature of the correlation. Note, that positive covariance *increases* the variance. So the more the samples are correlated, the worse the balance heuristic will perform.

A computable formulation of optimal MIS weights in the presence of sample correlation has not yet been found. While variance-aware weights are general enough to tackle this challenge, their achievable improvements are limited by estimation noise and their heuristic nature. Prior work has proposed ad-hoc solutions for correlation in the context of the VCM algorithm [Grittmann et al. 2021; Jendersie and Grosch 2018], again by including heuristic correction factors

in the balance heuristic weight. For example, the *correlation-aware MIS (CA-MIS)* weights [Grittmann et al. 2021] compute

$$w_t^{\text{cov-aware}}(x) \propto c_t(x) n_t p_t(x), \tag{13}$$

where

$$c_t(x) = \frac{P_t(y)}{P_t(y) + P_t(z) - P_t(y)P_t(z)} \tag{14}$$

is a heuristic correction factor that penalizes paths $x = yz$ if their (shared) camera prefix y has a lower probability than the light suffix z. To enable that, a unitless path probability

$$P_t(y) = \prod p_t(y_i|y_{i-1})\pi r^2 \tag{15}$$

is computed, where $p_t(y_i|y_{i-1})$ is the surface-area PDF of sampling the ith vertex of the path, and the radius r is a parameter computed from the distance between the camera position y_0 and the first hit point y_1 as $r = 0.0175\|y_1 - y_0\|$; it makes the weighting scale invariant. Intuitively, this heuristic penalizes paths where the prefix y is less likely than the suffix z, because this indicates that significant covariance *might* occur.

Our approach offers a *theoretically sound, generic* method to correct the balance heuristic when samples are correlated. It can be used in isolation or on top of the heuristics proposed by prior work. Note that we only consider correlation among samples *used for the same integral*. That is, we only consider correlation *within*, not across pixels. Correlation across pixels arises due to small-step mutations in MCMC, or other algorithms that share a single high-contribution sample across a large number of nearby pixels. Whether it is possible to account for pixel correlation in MIS, and how to do so, is a question we leave for future work.

3.4 Unknown PDFs

A key advantage of the balance heuristic and its derivatives is that computation requires only the sample counts and PDF values, quantities that are *typically* easy enough to obtain efficiently.

However, there are at least two families of methods that cannot compute exact, closed-form PDFs, namely Markov-chain Monte Carlo (MCMC) [Veach and Guibas 1997; Šik et al. 2016; Kelemen et al. 2002] and resampling [Talbot et al. 2005; Popov et al. 2015; Nabata et al. 2020; Bitterli et al. 2020; Lin* et al. 2022] methods.

These generate samples that *eventually* follow a target distribution $\hat{p}(x)$. For example, resampling methods first generate candidate samples with PDF $q(x)$ and then resample a subset of these candidates. In the limit, with infinite candidate samples, the resampled sample(s) will be distributed according to the target distribution. But for a finite number of candidates M, the actual distribution after resampling is not easily computable. On top of that, the target $\hat{p}(x)$ itself is usually only known up to a normalizing constant.

To use resampling or MCMC in an MIS combination, two options present themselves. We could use the candidate PDF $q(x)$ as a surrogate for MIS weighting purposes. That, however, neglects the improvements gained from the MCMC or resampling method. Hence, the weight will be too low whenever resampling is effective. The other extreme is to pretend that the samples already follow the target distribution $\hat{p}(x)$ [Kelemen et al. 2002]. This requires us to numerically estimate or approximate the normalizing constant, as we generally only know $p^*(x) \propto \hat{p}$. Additionally, the samples will only actually follow the target given infinite candidates / mutations. To account for that, previous work has suggested an interpolation between the candidate and target PDFs [Nabata et al. 2020]:

$$w_t^{\text{RIS}}(x) \propto \left(\frac{1}{M}\frac{1}{q(x)} + \left(1 - \frac{1}{M}\right)\frac{\langle P^*\rangle}{p^*(x)}\right)^{-1}, \tag{16}$$

where M is the number of candidate samples, q the candidate PDF, $p^*(x)$ the unnormalized target function, and $\langle P^*\rangle$ an estimate of its normalization factor (i.e., $\langle P^*\rangle \approx \int_{\mathcal{X}} p^*(x)\,\mathrm{d}x$) computed from prior, *independent* samples. Our method can be used to either replace or improve upon this heuristic weighting.

4 Our Approach

Prior work has successfully improved upon the balance heuristic by including heuristic correction factors $\beta_t(x)$, i.e.,

$$w_t(x) \propto \beta_t(x) n_t p_t(x). \tag{17}$$

Examples include the above-mentioned $c_t(x)$ (14) and v_t (11), or the power heuristic (where $\beta_t(x) = n_t p_t(x)$). We improve or replace these with an optimized correction factor γ, i.e., our weights read:

$$w_t(x) \propto \gamma_t \beta_t(x) n_t p_t(x). \tag{18}$$

To achieve practicality, γ_t are constant per technique (or per group of techniques). We find *an approximation* of the optimal set

$$\{\gamma_t^{\text{opt}}\} \approx \underset{\{\gamma_t\}}{\arg\min}\, \mathbb{V}\left[\langle F\rangle_{\text{MIS}}\right] \tag{19}$$

via a direct search over variance estimates. This approach avoids issues with non-convexity and correlated samples, and facilitates a fast convergence rate. A key insight from our work is that such a simple and seemingly crude approach achieves robust and significant improvements.

4.1 Simplifications

We perform a direct search over the variance estimates of a small number of correction factors. Those factors are shared over multiple integrals (e.g., all path lengths in a pixel) and applied to a group of techniques (e.g., all merging techniques).

Direct search. We opt for direct-search rather than gradient-based optimization for three main benefits. First, the variance (5) is easy to compute for any MIS estimator, while its gradients can become challenging due to, e.g., covariance. Second, direct-search enables a single-step optimization and does not require prolonged iterative refinement (though such refinement is, of course, still possible if desired). Finally, provided $\gamma_t = 1$ is included in the candidates, we can guarantee robust and consistent improvements over the plain balance heuristic.

Small number of candidates. A drawback of a discrete direct search solution is the less precise optimization outcome. However, we found that this is actually another benefit in our setting. We observed that fine-grained tuning of the correction factors has little benefit. Fig. 2 illustrates this on a challenging example from our VCM application. The figure compares optimization outcomes with the small set of just four candidates we use in practice ("Practical"), to a more fine-grained and wider-range alternative ("Extended"). Results for

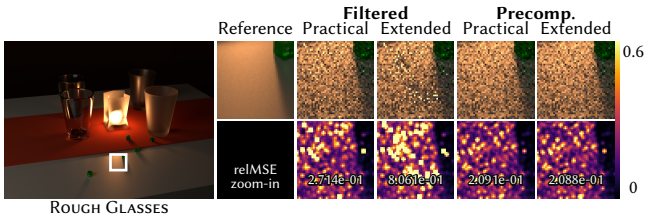

Fig. 2. Equal-sample comparison between a small set of candidate correction factors ("Practical") and a larger one ("Extended"). The "Filtered" version selects the best factor based on single-sample estimates, the "Precomp." one utilizes more converged variance estimates. The top row shows rendered images, the bottom row visualizes their error (relMSE) in false color. The fine-granular control of the "Extended" pool has next to no benefits with precomputation. In the practical "Filtered" version, it makes the optimization more susceptible to noise.

each set are shown with our practical on-the-fly estimation ("Filtered") and a ground-truth obtained from precomputed variances ("Precomp."). Even with accurate precomputation, the extended candidate pool does not provide measurable improvements. When used in a practical fashion (optimized from a single sample per pixel) the extended pool exacerbates issues arising from estimation error in the optimizer and produces inferior results.

Dimensionality reduction. To minimize overhead, we want to minimize the number of correction factors. For that, we found two effective solutions. First, we can group similar techniques together. For example, our VCM application applies the same correction factor to all merging techniques in the same pixel. Another easy trick is to optimize ratios of correction factors. For example, consider a combination of two (groups of) techniques $p(x)$ and $q(x)$. Any pair of correction factors a and b applied to both is equivalent to their ratio c applied only to p (or its reciprocal applied to q):

$$\frac{ap(x)}{ap(x)+bq(x)} = \frac{cp(x)}{cp(x)+q(x)}, \text{ where } c = \frac{a}{b}. \tag{20}$$

We use this to reduce the number of correction factors and thus the dimensionality of the optimization problem.

Pruning with prior knowledge. In most applications, we can benefit from the knowledge gained in prior work to further reduce overhead. For example, our VCM application only corrects the MIS weights of the merging and light tracing techniques, because prior work [Grittmann et al. 2019, 2021] has shown that these are the most problematic ones. Additionally, we know that the weight of those techniques is typically too large, so we further reduce the candidate pool by only considering $\gamma_t \leq 1$. Asides from performance gains, we found that leaving out unlikely candidates also improves optimization robustness (also shown in Fig. 2).

4.2 Variance estimation

To pick the best correction factor, we estimate the variance of all candidate choices. Since we want to support correlated samples, we *do not* utilize the single-sample formulation used by the balance heuristic (6). Rather, we directly estimate Eq. (5). Since the squared first moment F^2 of the *full* estimator is a constant w.r.t γ, we can neglect it and compute only an estimate of the second moment,

$$M_\gamma = \mathbb{E}\left[\left(\sum_{t\in\mathcal{T}}\sum_{i=1}^{n_t} w_{t,\gamma}(x_{t,i})\frac{f(x_{t,i})}{n_t p_t(x_{t,i})}\right)^2\right]. \tag{21}$$

Ground-truth computation. Accurate second moment estimates can be obtained by rendering many (thousands) of iterations using, e.g., the plain balance heuristic. For every candidate correction factor γ (a vector of γ_t if correction is applied to multiple techniques), we accumulate one first moment estimate

$$\langle F_\gamma\rangle = \sum_{t\in\mathcal{T}}\sum_{i=1}^{n_t} \frac{w_{t,\gamma}(x_{t,i})}{w_t(x_{t,i})} w_t(x_{t,i})\frac{f(x_{t,i})}{n_t p_t(x_{t,i})}. \tag{22}$$

This is identical to the usual pixel value computation, except that each sample is multiplied by the highlighted term: the ratio between the corrected MIS weight $w_{t,\gamma}$ of this candidate and the MIS weight that is used during variance estimation. At the end of each iteration, the squared values of these first moments are then accumulated in a running sum. So with N iterations, we compute

$$\langle M_\gamma\rangle = \frac{1}{N}\sum_{i=1}^{N}\langle F_\gamma\rangle_i^2. \tag{23}$$

Then, we simply compare these values to pick the best γ.

Practical computation. In practice, we avoid costly precomputation and instead estimate the second moments on-the-fly from the first rendering iteration. For that, we compute $\langle F_\gamma\rangle$ the same way as for a ground-truth optimization (22). Since we have only a single sample per pixel, directly comparing the squared values of these estimates is not useful. Instead, we apply filtering, essentially averaging $\langle F_\gamma\rangle^2$ over a pixel neighborhood to make squaring meaningful. This filtering also reduces noise in the estimates.

Filtering. Fig. 3 provides an overview of our filter pipeline, showing the optimized factor and resulting rendering after each stage. We start by blurring $\langle F_\gamma\rangle^2$ with a simple Gaussian filter. This makes the squared values meaningful to compare at 1spp. The result at this stage somewhat resembles the ground-truth (rightmost column), but still shows drastic noise and artifacts in this challenging scene.

We found that this can be greatly improved by blurring *relative* moments instead. For that, we compute a filtered version of the rendered image and square the pixel values to obtain an approximation of F^2. The filtered second moment estimates of each pixel are then divided by this value. We noticed two benefits from this approach: First, it improved results around discontinuities. Second, the division of two correlated quantities, namely the estimates of the first and second moments that are computed from the same samples, seems to reduce the overall noise. This introduces some additional bias to the variance estimation, but we did not notice any adverse effects.

Finally, after the best correction factor is picked in each pixel by comparing the relative moments, we blur this factor image. This avoids abrupt changes in the MIS weighting and thus prevents visible discontinuities in the noise pattern of the rendered image.

We tuned the filter parameters per application: The VCM experiments use a radius of 8 pixels, the resampling ones use a radius

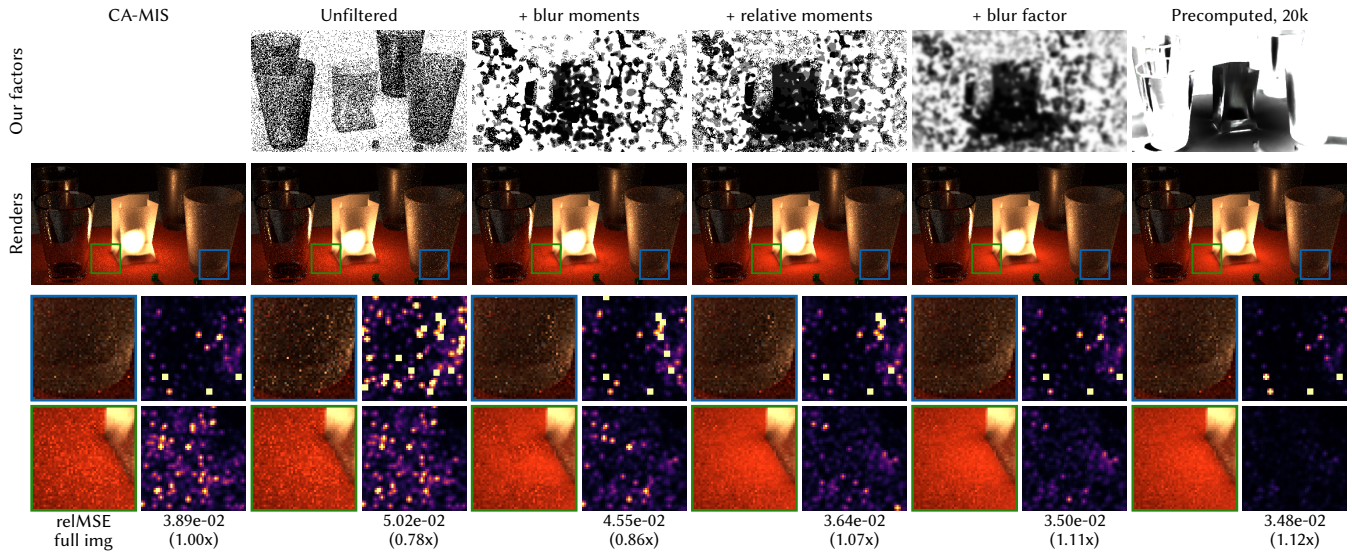

Fig. 3. Our filtering pipeline on an example from the VCM application. The top row shows the (hypothetical) correction factors γ_t at each stage. The second row shows the rendered image when using these factors. We highlight two interesting regions in the crops below: The top one on the glass is a place where the baseline (correlation-aware MIS, leftmost column) works well. The bottom one on the table cloth is an instance where further weight reduction is desired. Unfiltered moments (second column) are too noisy to be useful. Applying just a simple blur (third column) yields visible artifacts. By computing relative moments and blurring the resulting correction factors, we can achieve practical, single-sample optimization results that approach the precomputed ground-truth (rightmost column) in quality.

of 32 pixels. That is because VCM combines many techniques and thus has a lower noise level after just one iteration. In future work, a more advanced filtering pipeline – like a state-of-the-art denoiser – can be constructed. Section C discusses results from an initial experiment with bilateral filtering.

4.3 Summary

To summarize, our practical applications employ the following steps:

(1) Render one sample per pixel; estimate $\langle F_\gamma \rangle$ per pixel and candidate
(2) Filter pixel values (approximates F^2)
(3) Square the candidate means $\langle F_\gamma \rangle$ and blur over image (approximates M_γ)
(4) In each pixel, pick γ with lowest M_γ / F^2; blur over the image
(5) Render remaining samples with optimized w_γ

The individual applications differ only in the set of techniques and filter parameters.

All experiments were run on an AMD Ryzen 9 5900X; images are rendered at a 640×480 resolution with a maximum path length of ten rays[1]. We also include renderings with 1920×1440 resolution in the supplemental. The error metric we use is the relative mean-squared error (relMSE), i.e., the squared error of each pixel is divided by the squared ground-truth prior to averaging.

In addition to our main applications, we also show a quick test for correlations in BDPT in Section B and abstract 1D examples in Section A.

[1]The maximum depth is necessary for the comparison to VA-MIS.

5 Application 1: VCM

Our first application is the vertex connection and merging (VCM) [Georgiev et al. 2012; Hachisuka et al. 2012] algorithm. Previous work has demonstrated two main MIS weighting issues in the context of VCM [Grittmann 2023, Chapter 4]: First, a particularly strong case of the "low-variance technique" challenge (see Section 3) arises in the presence of simple direct illumination effects. Second, the merging (aka photon mapping [Jensen 1996]) technique introduces severe sample correlation and thus poor MIS weights.

Approach. We tackle these challenges by optimizing two *independent* correction factors. For the direct illumination in each pixel, i.e., paths comprising two rays, we find γ_l, a scaling factor to reduce the weight of the "light tracing" technique, i.e., direct connections of light paths to the camera. For indirect illumination, we optimize a joint γ_m, a scaling factor to reduce the weight of all merging techniques in this pixel. The former γ_l addresses the low-variance challenge. The latter γ_m also addresses this challenge (e.g., if simple direct illumination is visible in a mirror) and also the sample correlation challenge. For both cases, we only need to reduce the weight. We thus found the following set of candidates sufficient:

$$\gamma_l, \gamma_m \in \{0.01, 0.1, 0.5, 1.0\}. \tag{24}$$

For our ground-truth comparisons, we use a larger number of candidate values in $[0.01, 100]$, i.e., we also allow increasing the weight.

A benefit of our approach is that it can be combined with prior work. We explore this by using our correction on top of the correlation-aware MIS weights (CA-MIS) [Grittmann et al. 2021]. Unless indicated otherwise, all our results use this combined approach.

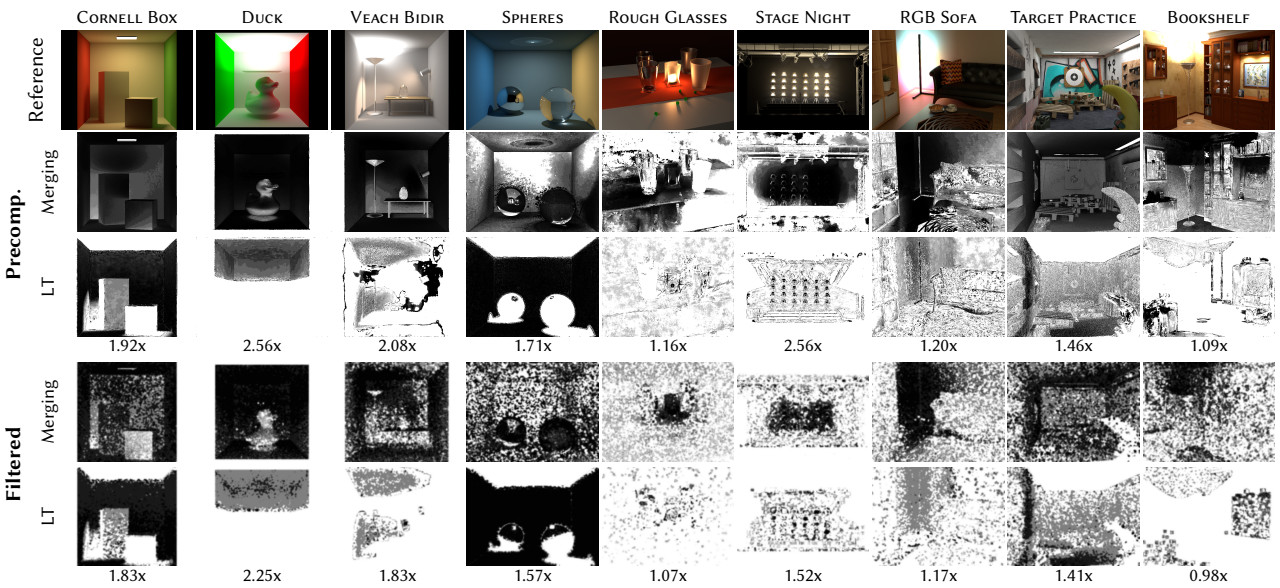

Fig. 4. Comparison of our MIS weight correction from a 20k spp precomputation ("Precomp.") to the filtered 1spp solution we use in practice ("Filtered"). The merge factor γ_m and light tracer factor γ_l are shown in rows "Merging" and "LT" respectively. White pixels retain the baseline MIS weights [Grittmann et al. 2021], darker ones have the weight of the corresponding technique reduced. The numbers below each pair of rows are the equal-sample speed-up when using the corresponding weight correction, compared to the baseline. While our crude 1spp factors are quite far from converged, they do successfully identify the regions in most dire need of weight correction (close to black pixels).

5.1 Impact of filtering

Filtering prevents artifacts from noisy variances but could potentially cause artifacts of its own. In fact, the filtered factors output by our method look very noisy, as can be seen in the bottom two rows of Fig. 4. The figure compares the ground-truth and 1spp correction factors for a diverse set of test scenes. The numbers in Fig. 4 indicate the equal-sample improvement over the baseline (CA-MIS) using the precomputed or filtered solution. While the more accurate solution can achieve significant further improvements in some challenging scenes, like Stage Night, most scenes do not stand to benefit enough from it to justify investing further compute resources.

The rightmost column in Fig. 4 shows the worst-case scene we found: Variance estimates in the Bookshelf are extremely noisy. Our practical solution here results in a (just barely measurable) performance degradation. The precomputed values, in comparison, yield a minor speed-up of 10%. Future work could improve upon this via better filtering schemes or iterative refinement. However, care must be taken to avoid offsetting the achievable gains with the added overhead.

Fig. 5 shows the adverse impact of noisy correction factors in the worst-case example, in equal time. While our method achieves significant improvements in some spots (top row of crops), the inaccurate correction factors cause visible increases in noise elsewhere. This is slightly worsened by the computational overhead of our approach, amounting to one fewer iteration being rendered here. While the overall performance is still on-par with the baseline, it could be around ten percent better, as shown in Fig. 4.

5.2 Comparison to variance-aware MIS

On the surface, our approach is similar to variance-aware MIS (VA-MIS) [Grittmann et al. 2019]: Both utilize variance estimates from the first sample per pixel to enhance MIS weighting of subsequent iterations. But compared to VA-MIS, our method (1) requires fewer correction factors, (2) can be combined with other heuristics, and (3) provides further improvements thanks to being an optimization rather than a heuristic. In the following, we discuss the first two cases, the general improvements are discussed in the main equal-time comparison.

Low-variance weighting. VA-MIS was introduced to rectify the "low variance technique" weighting issue. An example is shown in Fig. 6, where direct illumination is rendered using an MIS combination of next-event estimation and light tracing. The former has next to no variance thanks to the tiny light, but the balance heuristic does not reflect that. VA-MIS can rectify this issue, but its correction factor must be applied to all techniques. This can be seen by comparing columns (b) and (c), where (b) only corrects the light tracer and (c) is the full solution. Our approach can optimize a ratio (γ_l here) and thus only needs a single factor to achieve similar results. In the simple direct illumination case here, the overhead reduction is insignificant. However, if we move to full global illumination in a bidirectional algorithm, where the number of techniques scales quadratically with the path length, cost quickly becomes an issue for VA-MIS. Our method can freely control the number of correction factors and obtains meaningful values for any setup.

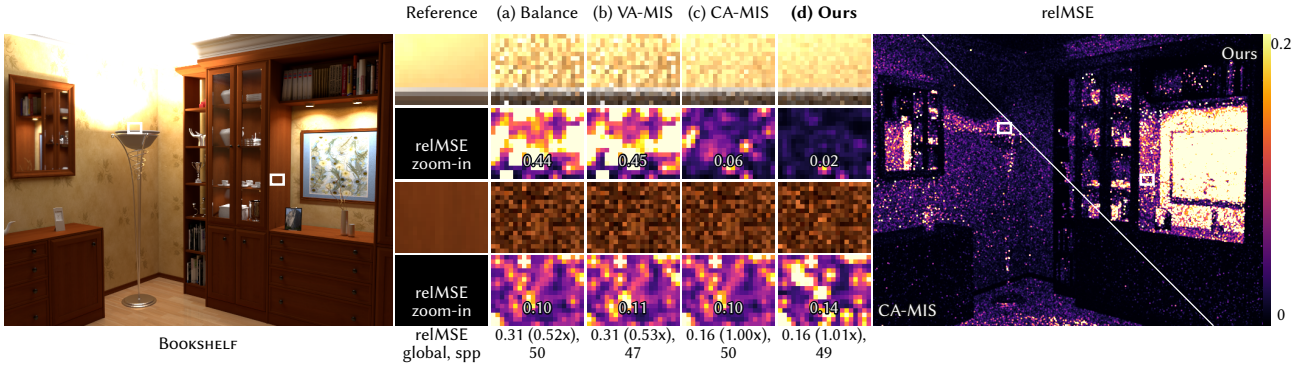

Fig. 5. Equal-time comparison of our method to prior work in the *worst case* scene that we found. The false-color images visualize the error (relative squared error), the numbers report the mean error in each crop and over the entire image, as well as the number of iterations rendered. Due to noisy estimates in the optimization, we only improve some spots (top row of crops) while slightly degrading others (bottom row of crops).

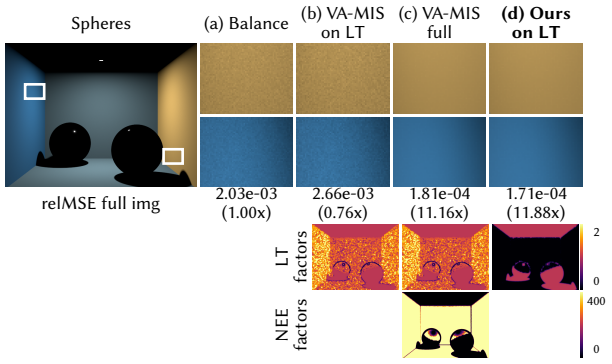

Fig. 6. Equal time (10s) comparison between balance heuristic, variance-aware MIS, and ours for low-variance direct illumination. We compare two variants of variance-aware: (b) only applied to light tracing and (c) applied also to next event. Our method (d) only applies a factor to light tracing. The false color images at the bottom visualize the correction factors. We achieve similar results compared to (c) with fewer factors being computed.

Flexibility and composite solutions. A major advantage of our approach is the design freedom in how it should be used. Fig. 7 compares different combined MIS weighting schemes on a difficult setup. The Veach Bidir scene features severe MIS weighting issues due to covariance in the merging techniques, as evident in the high noise of the balance heuristic combination (a). VA-MIS can be applied to this case (b) but it was not designed for it: Its weight correction is too hesitant and the improvements achieved are rather meagre. In contrast, using our optimization (c) directly on the balance heuristic produces significant improvements, held back only by the noise in the 1spp correction factor estimation. The CA-MIS heuristic [Grittmann et al. 2021] was specifically designed for this problem and handles it well (d), although some residual noise is evident. Our method also enables a theoretically sound way to further enhance the CA-MIS weights (f). While VA-MIS can also be arbitrarily applied on top of CA-MIS (e), this does not yield noticeable improvements – unsurprisingly, since there is no theoretical reason why this combination should perform well.

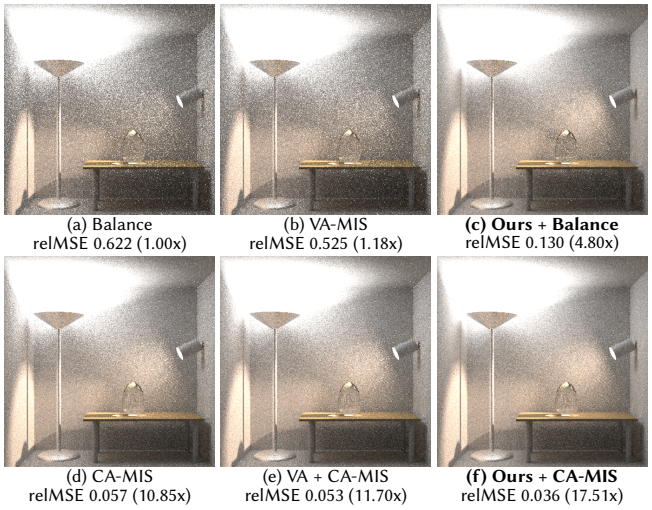

Fig. 7. Equal-time (10s) study of different MIS weights for global illumination with VCM. Our approach can be applied directly to the balance heuristic, or on top of other heuristics like correlation-aware MIS (CA-MIS). We achieve consistent improvements in all cases, though the best results are achieved in conjunction with the best baseline; here, that is CA-MIS.

5.3 Equal-time performance

Thanks to the cheap 1spp single factor optimization, our method compares favorably to prior work in equal-time comparisons. A selection of results is shown in Fig. 8. Full renderings of all scenes can be found in the supplemental HTML files.

The first example is the Veach Bidir scene. It features challenging indirect illumination, caustics, and severe correlation issues. CA-MIS rectifies the worst of those. Our optimized version of it achieves further improvements across most of the scene by further reducing the weight of the problematic merging technique in many places.

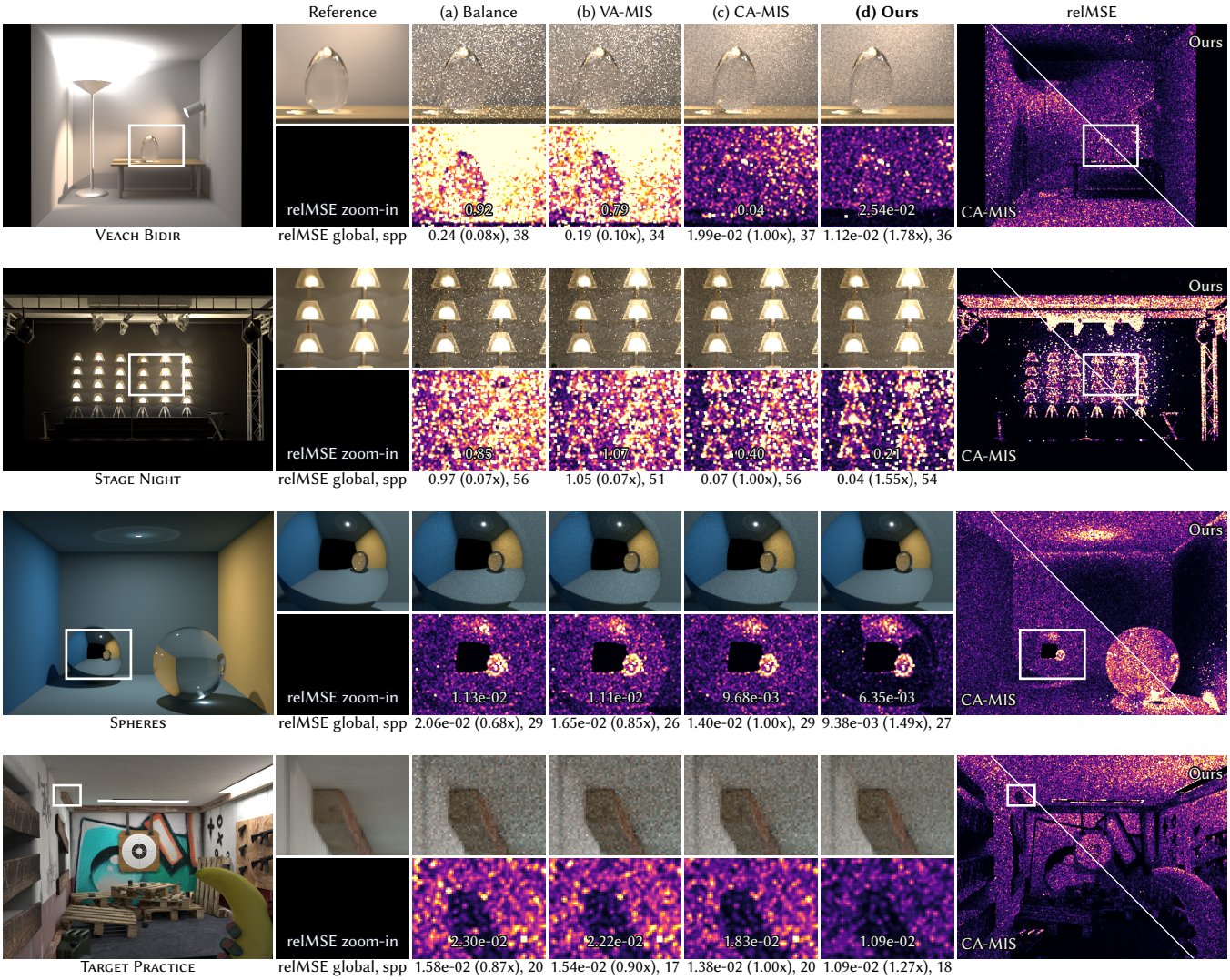

Fig. 8. Equal-time results (30 seconds) from our VCM application. The false color images visualize the error (relative squared error) in each crop and over the entire image. The inset texts provide the mean error values and speed-up compared to CA-MIS in each crop, the text below each row are the errors and speed-ups over the entire image. Our method achieves visible improvements over a wide range of scene types and illumination effects, while never being noticeably worse than the baseline.

The Stage Night shows a similar correlation-induced challenge, this time paired with a large number of light sourcesinside lampshades. Our correction removes noticeable outliers on the wall behind the lamps.

In Spheres we observe the low-variance weighting issue on the walls and in their specular reflections. Additionally, severe outliers arise again from sample correlation (caused by indirect light from the caustic underneath the glass ball). VA-MIS struggles to improve upon this, and CA-MIS cannot compensate the low-variance issue by design. Our method visibly reduces noise: on the walls and on their reflections.

Finally, Target Practice features medium-difficulty diffuse illumination from large light sources. Unnecessary noise from the merging technique pollutes the rendering with the balance heuristic. CA-MIS cannot remove it effectively: The uniform, yet high-variance illumination violates an underlying assumption behind the heuristic that CA-MIS employs, namely that sample density is an accurate indicator of variance. Since our method is based on optimization rather than heuristics, it does not suffer such issues.

The overhead of our method is mainly due to the moment computation and filtering in the first iteration. It amounts to roughly the cost of 1-2 rendering iterations with VCM. This is insignificant compared to the achievable speed-ups. Even in the worst case, when no speed-ups can be had (like Fig. 5) the overhead is low enough to not visibly degrade the image compared to the baseline.

6 Application 2: Resampled Importance Sampling

Our second application is Resampled Importance Sampling (RIS) [Talbot et al. 2005; Bitterli et al. 2020]. RIS is an example where MIS must make due with only an approximation of the true sample density: Resampling methods generate a set of candidate samples and then resample a subset based on some target function. Hence, the sampling density *asymptotically* follows the target density with infinite candidates. However, the actual density for a finite number of candidate samples is unknown. Furthermore, the target density itself is also only known up to a normalization factor.

6.1 Setup

We render direct illumination with an MIS combination of BSDF- and light sampling. For the light sampling, we generate 32 candidate samples via uniform area sampling of the light sources. From these, we resample one sample with a target function based on the unoccluded, BSDF-weighted contribution, i.e.,

$$p^*(y|x,\omega_o) = L_e(y \to x)G(x,y)B(x,y,\omega_o), \tag{25}$$

where x is the shading point, ω_o the direction from which x is observed, y the sampled point on a light, L_e the emitted radiance, $G(x,y)$ the geometry term [Veach 1997] (excluding visibility) and B the BSDF. MIS is used to combine this with one BSDF sample.

For MIS weighting, we consider three alternatives to our approach: Using the balance heuristic with the candidate PDF, the approach of Nabata et al. [2020] who estimate the normalization constant of the target density to use it for the balance heuristic (see Section 3), as well as the VA-MIS weights.

The baseline for our method is the balance heuristic using the candidate PDF. We modify it with a factor

$$\gamma_r \in \{0.01, 0.1, 0.5, 1.0\} \tag{26}$$

multiplied onto the BSDF sampling weight. Since resampling improves the light sampling, we aim increase its weight by reducing the weight of BSDF sampling. We use the same pool of candidate values as for the VCM application.

6.2 Equal-time comparison

Fig. 9 shows results for three representative scenes. For each scene, we show equal-time results with the practical filtering approach in the first row, and equal-sample results with precomputation in the second row. Fig. 10 shows the corresponding correction factors computed by our method.

The Modern Hall scenes features many lights, complex occlusion, and glossy surfaces. In the equal-time setting (first row) VA-MIS struggles with the sever noise, causing visible artifacts. With precomputation, the heuristic nature of VA-MIS becomes apparent: It was not designed for this challenge and hence performs unsatisfactorily. The domain-specific solution of Nabata et al. [2020] works well here, albeit with slightly higher noise in the practical setup compared to our solution.

The Rgb Sofa showcases a low-variance problem arising due to resampling. VA-MIS handles this case well, as it was designed for it. The approach of Nabata et al. [2020] only tackles the unknown resampling PDF and inherits the low-variance problem from the underlying balance heuristic. In this example, that actually amplifies the low-variance weighting issue. Our method performs on-par with VA-MIS.

The Veach MIS showcases a limitation of our method. First, our crude filtering pipeline is not sensitive to the sharp edges of the reflection. Thus, in equal-time, our method is worse than Nabata et al. around those edges. Second, the cropped region shows a well-known low-variance problem where higher weights for BSDF sampling are desired [Veach and Guibas 1995]. However, we are only considering candidates that reduce the BSDF weight. Including additional candidates is possible, but introduces higher error in the optimization and a higher overhead.

Overall, our results show that none of the prior solutions perform satisfactorily for every scene. Our method performs on-par with the best solution in each scene, thus it is the only *consistent* solution of the options we have considered here. Note that all options – VA-MIS, Nabata et al., and ours – require the use of filtered noisy estimates. Namely, the technique variances, normalization factor, and candidate variances, respectively. Finding an alternative with no added memory cost remains an open problem.

7 Limitations and future work

Filtering. The main limitation of our method is that we operate with variance estimates from a single sample per pixel. Hence, the filtering we apply is a crucial component. It is also the source of the majority of our overhead. A better filtering approach could offer two benefits: First, the quality can be improved by more accurate filtering, such as using a modern denoiser. Second, the overhead can be reduced by reducing the filtering cost. Unfortunately, these two goals contradict each other, so a good trade-off is needed.

Adaptive candidate selection. A key step to make our method practical was to carefully choose the set of allowed candidates. Allowing too many results in undue overhead; allowing too severe weight adjustment in the wrong direction (i.e., increases of weight for a very poor technique) risks image degradation. So, each practical application of our method must carefully review this choice of candidates. To increase accuracy without undue overhead, an adaptive, iterative scheme like the Nelder-Mead method could be used instead of an a-priori fixed set. For robustness against extreme adverse weight changes, statistical tests could replace simple comparisons. For example, a chi-squared test can be used to avoid picking a poor candidate with severely underestimated variance.

Spatial structure. In our rendering application, we computed the correction factor in image space, where filtering is easy and the behavior of the method is easy to understand and easy to compare to others. It could be interesting to explore how the approach behaves if factors are optimized in 3D scene space instead. Doing so allows for finer control over the MIS weights, but complicates filtering and increases cost.

Real-time application. We showed our simple optimization scheme is effective for resampling methods. It would be interesting to apply our approach to real-time applications like ReSTIR [Bitterli et al. 2020; Lin* et al. 2022], which inherit the same issue from unknown PDFs. However, this will require some form of carefully designed temporal reuse and probably a cheaper filtering approach.

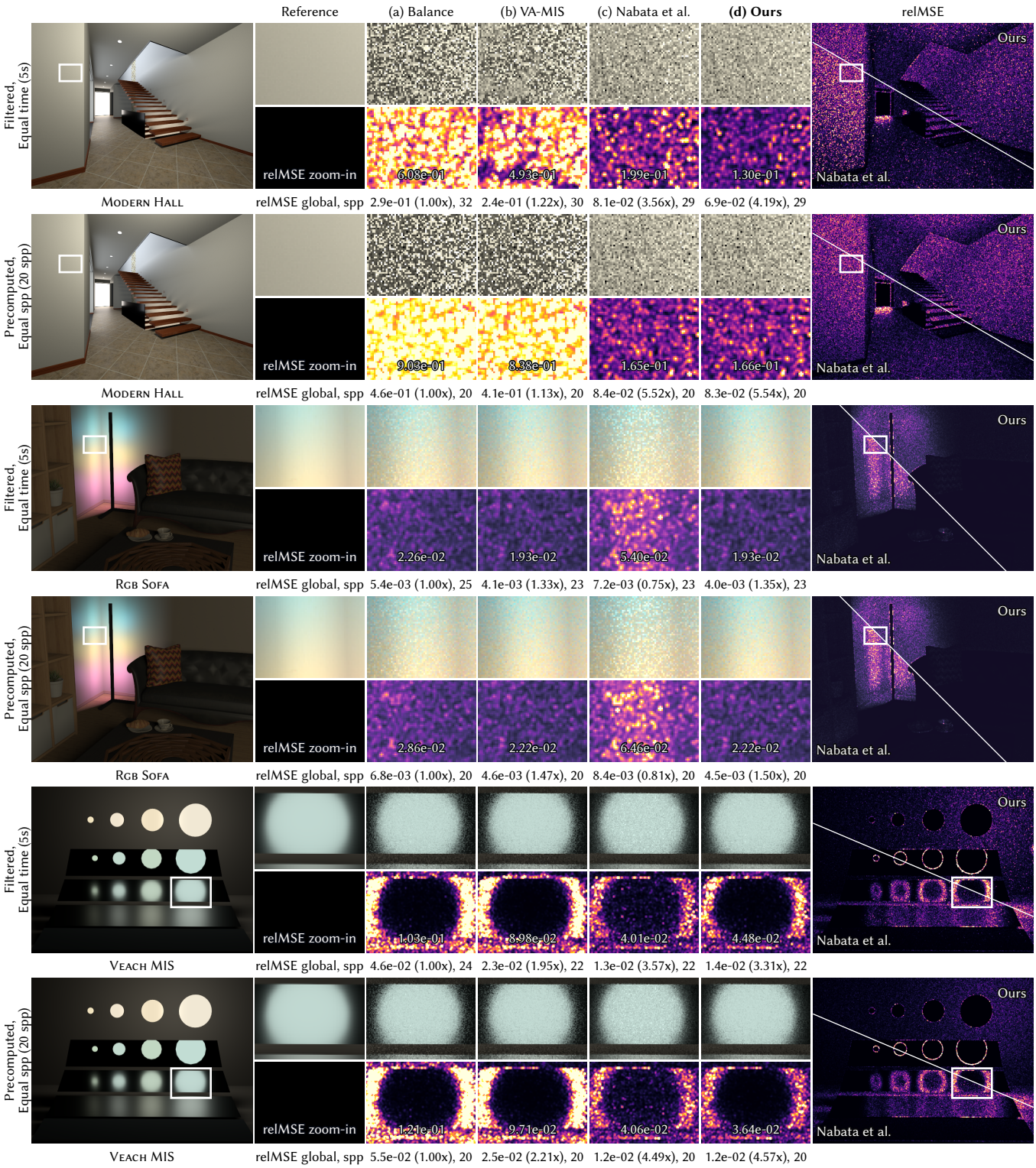

Fig. 9. Representative results for the resampled direct illumination application. For each scene, we show equal-time (5 seconds) results with filtering in the first row, and equal-sample (20) with precomputation in the second row. "Balance" uses the candidate PDF for MIS weighting and is the only method that requires neither auxiliary memory nor precomputation. Our method performs on-par with the best prior alternative in each scene and is thus the only one that *consistently* improves upon the balance heuristic.

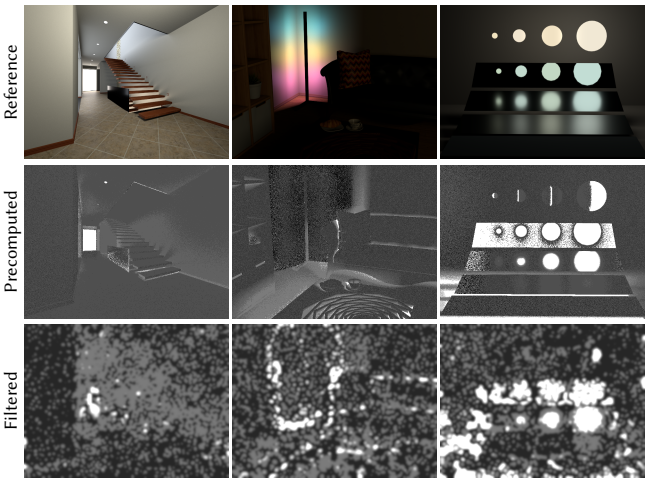

Fig. 10. Comparsion of our MIS weight correction factors from a 20k spp precomputation (center) to the filtered 1spp solution we use in practice. White pixels retain closer to balance heuristic. Darker pixels are the regions where Bsdf weights are reduced.

Single-sample optimization. Our usage of pixel variance approximations from a single-sample iteration might be insufficient in some settings. One case are adaptive algorithms, like path guiding, MCMC methods, or temporal sample reuse, where sampling quality improves over time. These will likely require at least sporadic incremental updates of our optimization factors. Another challenge are low-discrepancy samples: If these are not sufficiently decorrelated across pixels, then the pixel variances estimated in early iterations might be distorted by the sampling artifacts. This is again an issue that can be resolved via iterative refinement, although the more desirable solution would be to simply use uncorrelated sampling sequences to begin with.

Pixel correlation. Some rendering techniques can exhibit severe *pixel* correlation. Examples include most Markov chain Monte Carlo (MCMC) methods [Šik and Křivánek 2018] or spatiotemporal reuse schemes [Bitterli et al. 2020]. These often exhibit splotchy artifacts in the image from reusing the same, high-contribution sample over many neighboring pixels. This type of correlation is not accounted for by our approach: Like all prior work on MIS, we consider only the variance *within* pixels, thus ignoring correlation between different pixels. In an MIS context, it might be beneficial to reduce the weight of techniques with severe pixel correlation. Sadly, this objective cannot be achieved with our current formulation, as the pixel variance, by definition, does not contain correlation between pixels. Future work could investigate if MIS weights can (and should) be altered to encourage low, or even negative, correlation between pixels. This is closely related to blue-noise error distribution methods [Georgiev and Fajardo 2016].

Negative weights. Optimal MIS weights [Kondapaneni et al. 2019] show significant improvements by allowing negative weights. Our 1D examples in Section A show that such benefits can theoretically also be obtained via our optimization. However, in rendering practice, we did not find a case where negative weighting was beneficial. It is possible that these improvements only occur when the MIS weights encode a control variate [Kondapaneni et al. 2019; Hua et al. 2023; Owen and Zhou 2000] – that would explain why our corrected balance heuristic did not profit from negative weighting. However, further research is required to see if negative weighting can be utilized in a practical way in complex MIS combinations like the VCM algorithm.

Optimal weights for correlation. Finally, the answer to a major question continues to elude us is: What are the truly optimal MIS weights for cases where the samples are correlated within and across the different techniques? While we can find the optimal corrected balance heuristic that accounts for correlation, it remains unclear how much improvements could be had by moving away from balance-style MIS weights altogether.

8 Conclusion

We present a generic recipe to adaptively correct MIS weights while rendering. We start with an initial guess of a well-suited MIS weighting function (e.g., the balance or power heuristic), and a set of candidate correction factors. The latter are chosen based on prior knowledge or mathematical analysis of the application. Through direct search over those candidate correction factors, we can then find the best MIS weighting function on-the-fly. By minimizing the number of allowed candidates and applying careful filtering, we can run our method off of just a single sample per pixel. Thus, we achieve significant, consistent improvements over prior work in equal-time. We demonstrate these improvements on two practical applications: Bidirectional rendering with the VCM algorithm, and resampled direct illumination.

Acknowledgments

We thank the following artists for sharing the test scenes: Wig42 (Dining Room), NewSee21035 (Modernhall), and Vladislav Hnatovskiy and Mira Niemann (Target Practice). This project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 540768663.

References

Benedikt Bitterli, Chris Wyman, Matt Pharr, Peter Shirley, Aaron Lefohn, and Wojciech Jarosz. 2020. Spatiotemporal reservoir resampling for real-time ray tracing with dynamic direct lighting. *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 39, 4 (July 2020). doi:10/gg8xc7

Iliyan Georgiev and Marcos Fajardo. 2016. Blue-noise Dithered Sampling. *ACM SIGGRAPH 2016 Talks* (2016). doi:10.1145/2897839.2927430

Iliyan Georgiev, Jaroslav Křivánek, Tomáš Davidovič, and Philipp Slusallek. 2012. Light transport simulation with vertex connection and merging. *ACM Trans. Graph.* 31, 6 (2012), 192–1.

Pascal Grittmann. 2023. *Rethinking multiple importance sampling for general and efficient Monte Carlo rendering.* Ph. D. Dissertation. Saarland University.

Pascal Grittmann, Iliyan Georgiev, and Philipp Slusallek. 2021. Correlation-Aware Multiple Importance Sampling for Bidirectional Rendering Algorithms. *Comput. Graph. Forum (EG 2021)* 40, 2 (2021).

Pascal Grittmann, Iliyan Georgiev, Philipp Slusallek, and Jaroslav Křivánek. 2019. Variance-Aware Multiple Importance Sampling. *ACM Trans. Graph. (SIGGRAPH Asia 2019)* 38, 6 (Nov. 2019), 152:1–152:9.

Pascal Grittmann, Arsène Pérard-Gayot, Philipp Slusallek, and Jaroslav Křivánek. 2018. Efficient Caustic Rendering with Lightweight Photon Mapping. In *Comput. Graph. Forum (EGSR '18)*, Vol. 37. Wiley Online Library, 133–142.

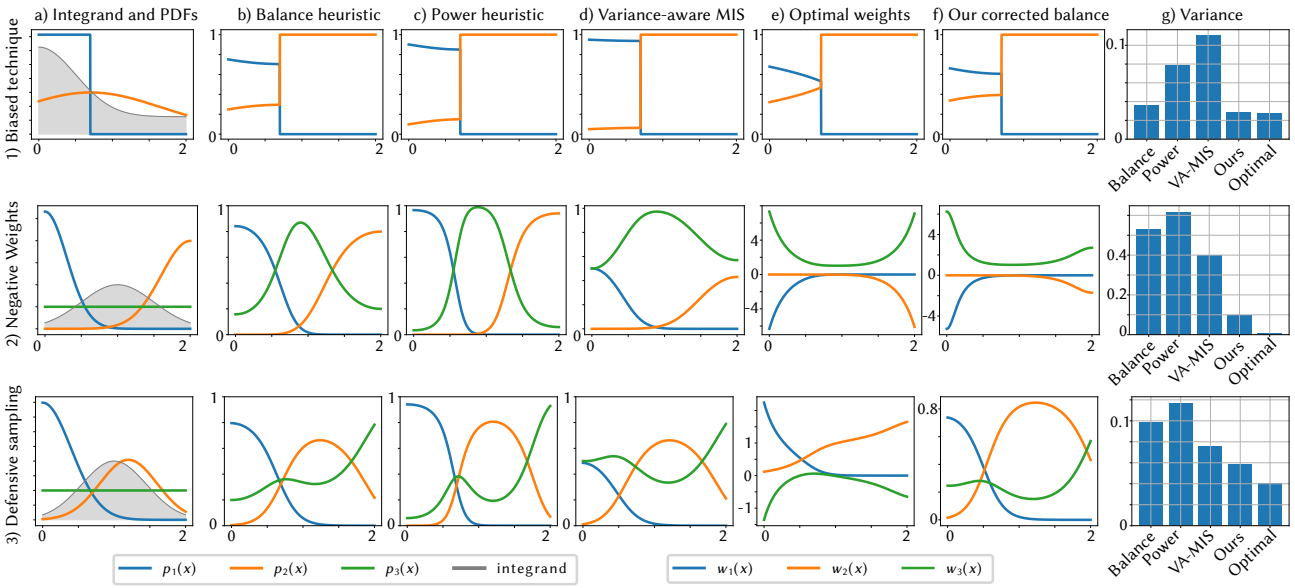

Fig. 11. Illustrative 1D integration problems. Top row: A combination of a biased technique and an unbiased one. Middle row: A setup where an unconstrained sign of the weighting function can be beneficial. Bottom row: One technique is nearly proportional to the integrand. Columns (b-f) show the different MIS weighting functions, (g) compares the resulting variances (lower is better). While not always optimal, ours is consistently better than all prior weights.

Pascal Grittmann, Ömercan Yazici, Iliyan Georgiev, and Philipp Slusallek. 2022. Efficiency-Aware Multiple Importance Sampling for Bidirectional Rendering Algorithms. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2022)* 41, 4, Article 80 (jul 2022), 12 pages. doi:10.1145/3528223.3530126

Toshiya Hachisuka, Jacopo Pantaleoni, and Henrik Wann Jensen. 2012. A path space extension for robust light transport simulation. *ACM Trans. Graph. (TOG)* 31, 6 (2012), 191.

Hera Y He and Art B Owen. 2014. Optimal mixture weights in multiple importance sampling. *arXiv preprint arXiv:1411.3954* (2014).

Qingqin Hua, Pascal Grittmann, and Philipp Slusallek. 2023. Revisiting Controlled Mixture Sampling for Rendering Applications. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2023)* 42, 4 (jul 2023). doi:10.1145/3592435

Johannes Jendersie and Thorsten Grosch. 2018. An Improved Multiple Importance Sampling Heuristic for Density Estimates in Light Transport Simulations.. In *EGSR (EI&I)*. 65–72.

Henrik Wann Jensen. 1996. Global illumination using photon maps. In *Rendering Techniques' 96*. Springer, 21–30.

James T. Kajiya. 1986. The Rendering Equation. *SIGGRAPH Comput. Graph.* 20, 4 (Aug. 1986), 143–150.

Ondřej Karlík, Martin Šik, Petr Vévoda, Tomáš Skřivan, and Jaroslav Křivánek. 2019. MIS Compensation: Optimizing Sampling Techniques in Multiple Importance Sampling. *ACM Trans. Graph. (SIGGRAPH Asia '19)* 38, 6 (2019), 12 pages.

Csaba Kelemen, László Szirmay-Kalos, György Antal, and Ferenc Csonka. 2002. A simple and robust mutation strategy for the metropolis light transport algorithm. In *Comput. Graph. Forum*, Vol. 21. Wiley Online Library, 531–540.

Ivo Kondapaneni, Petr Vévoda, Pascal Grittmann, Tomáš Skřivan, Philipp Slusallek, and Jaroslav Křivánek. 2019. Optimal Multiple Importance Sampling. *ACM Trans. Graph. (SIGGRAPH 2019)* 38, 4 (July 2019), 37:1–37:14.

Daqi Lin*, Markus Kettunen*, Benedikt Bitterli, Jacopo Pantaleoni, Cem Yuksel, and Chris Wyman. 2022. Generalized Resampled Importance Sampling: Foundations of ReSTIR. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2022)* 41, 4, Article 75 (07 2022), 23 pages. doi:10.1145/3528223.3530158 (*Joint First Authors).

Heqi Lu, Romain Pacanowski, and Xavier Granier. 2013. Second-Order Approximation for Variance Reduction in Multiple Importance Sampling. In *Computer Graphics Forum*, Vol. 32. Wiley Online Library, 131–136.

Thomas Müller. 2019. "Practical Path Guiding" in Production. In *ACM SIGGRAPH Courses: Path Guiding in Production, Chapter 10* (Los Angeles, California). ACM, New York, NY, USA, 18:35–18:48. doi:10.1145/3305366.3328091

David Murray, Sofiane Benzait, Romain Pacanowski, and Xavier Granier. 2020. On Learning the Best Balancing Strategy. In *Eurographics 2020*, Vol. 20. 1–4.

Kosuke Nabata, Kei Iwasaki, and Yoshinori Dobashi. 2020. Resampling-aware Weighting Functions for Bidirectional Path Tracing Using Multiple Light Sub-Paths. *ACM Trans. Graph. (SIGGRAPH 2020)* 39, 2 (2020), 1–11.

Art Owen and Yi Zhou. 2000. Safe and Effective Importance Sampling. *J. Amer. Statist. Assoc.* 95, 449 (2000), 135–143.

Anthony Pajot, Loic Barthe, Mathias Paulin, and Pierre Poulin. 2010. Representativity for robust and adaptive multiple importance sampling. *IEEE transactions on visualization and computer graphics* 17, 8 (2010), 1108–1121.

Matt Pharr, Wenzel Jakob, and Greg Humphreys. 2023. *Physically based rendering: From theory to implementation.* MIT Press. https://pbr-book.org/4ed/

Stefan Popov, Ravi Ramamoorthi, Fredo Durand, and George Drettakis. 2015. Probabilistic connections for bidirectional path tracing. In *Computer Graphics Forum*, Vol. 34. Wiley Online Library, 75–86.

Mateu Sbert, Vlastimil Havran, and László Szirmay-Kalos. 2019. Optimal Deterministic Mixture Sampling.. In *Eurographics (Short Papers)*. 73–76.

Martin Šik and Jaroslav Křivánek. 2018. Survey of Markov Chain Monte Carlo Methods in Light Transport Simulation. *IEEE Transactions on Visualization and Computer Graphics* (2018), 1–1.

Martin Šik, Hisanari Otsu, Toshiya Hachisuka, and Jaroslav Křivánek. 2016. Robust light transport simulation via metropolised bidirectional estimators. *ACM Trans. Graph. (TOG)* 35, 6 (2016), 245.

Justin Talbot, David Cline, and Parris Egbert. 2005. Importance Resampling for Global Illumination. In *Eurographics Symposium on Rendering (2005)*, Kavita Bala and Philip Dutre (Eds.). The Eurographics Association. doi:/10.2312/EGWR/EGSR05/139-146

Eric Veach. 1997. *Robust Monte Carlo methods for light transport simulation.* Stanford University PhD thesis.

Eric Veach and Leonidas J Guibas. 1995. Optimally Combining Sampling Techniques for Monte Carlo Rendering. In *SIGGRAPH '95*. ACM, 419–428.

Eric Veach and Leonidas J Guibas. 1997. Metropolis light transport. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 65–76.

A Discussion in 1D

Fig. 11 compares the performance of different MIS weighting schemes in illustrative 1D examples. Each row corresponds to a different integration problem. The first column (a) plots the integrand and the sampling techniques. Columns (b-f) show the MIS weights with different methods. We compare, from left to right, the original balance and power heuristics [Veach and Guibas 1995], the variance-aware

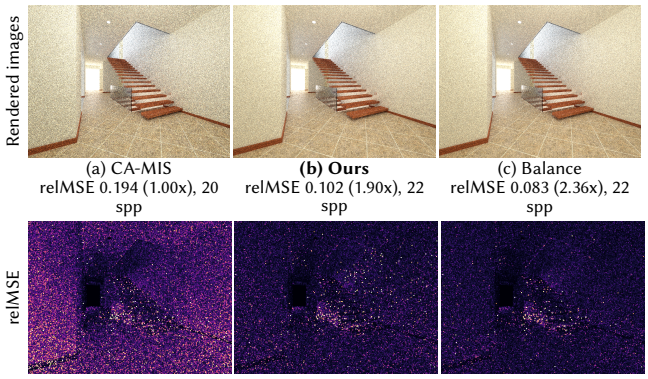

Fig. 12. An equal-time (60s) evaluation of our correction applied to a correlated bidirectional path tracer with multiple shadow rays. In this worst-case for CA-MIS, we can almost fully restore the better performance of the balance heuristic through our correction factors.

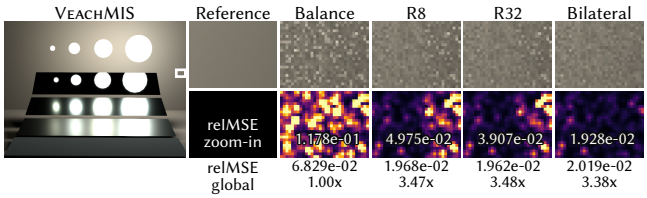

Fig. 13. An equal-sample (20 spp) comparison of our filtering pipeline with different blur radii (R8, R32) and a bilateral filter. Due to the noisy correction factors, our filtering pipeline may generate inconsistent noise in some areas. Using bilateral filtering can eliminate those, but this requires carefully selected parameters and introduces additional overhead.

balance heuristic [Grittmann et al. 2019], the optimal weights [Kondapaneni et al. 2019] and our corrected balance heuristic. For our method, we consider correction factor candidates in $[-1, 1]$. We use a regular subdivision and 20 candidates for each technique. The last column (g) compares the resulting variances.

Biased techniques. In the first example (top), the blue technique is biased: it only samples a portion of the domain. This setup is not supported by the variance-aware weights, resulting in a higher error than with the balance heuristic. Our approach supports arbitrary setups and here achieves close to optimal MIS weighting.

Negative weights. The middle row shows a setup where a negative weight can reduce the variance significantly. Here, p_1 (blue) and p_2 (orange) are extremely bad techniques, p_3 (green) is uniform, i.e., not great but not horrible either. By allowing negative weights – through our correction factors – we can achieve drastic improvements over the balance heuristic here.

Defensive sampling. A classic failure case of the balance heuristic arises from the combination of (potentially) low-variance techniques with other techniques [Veach and Guibas 1995; Grittmann et al. 2019]. The bottom row shows one such case. Here, p_2 (orange) is almost proportional to the integrand, but not ideally so, necessitating the need for defensive sampling [Owen and Zhou 2000]. Here, all MIS weights except for the optimal ones have similar shape. The main difference occurs around the tail on the right side. There, the balance, power and variance-aware weights are too high for p_3 (green). In contrast, optimal MIS relies more heavily on p_2 (orange) in that part, additionally utilizing negative weights. While our correction does not yield a negative shape as optimal MIS, we do reduce the weight assigned to p_3 correctly. VA-MIS moves in a similar direction, but too meekly.

B Correlation in Bidirectional path tracing

We revisit a failure case of CA-MIS in Fig. 12. Here, the integrator is bidirectional path tracing (i.e., no merging) with multiple shadow rays used for next event estimation (we use 50). This also causes sample correlation, albeit at a much lower scale than with merging. The Modern Hall scene is lit quite evenly by 300 lights from multiple directions. CA-MIS here incorrectly "assumes" that the low light density indicates high variance, when in reality it is a consequence of the more uniform distribution of light. As a result, performance is much worse than with the plain balance heuristic. Our method can be used to restore most of the better performance of the balance heuristic, as demonstrated with the ground-truth precomputed factors in the rightmost column. Here, we apply a global correction factor to the next-event techniques with candidates chosen from $\{1.0, 10, 30, 50, 100\}$. Our practical 1spp solution struggles to achieve the best result possible here: since the required weight correction is quite drastic, it overshoots the target and yields unnecessary outliers above the staircase. Such cases where very severe weight correction is required would benefit from improved denoising of the variance estimates, or incremental refinement of the correction factors.

C Bilateral Filter

We found that a primitive filter pipeline, based on aggressive Gaussian blurring, performs satisfactorily. Better filters, like bilateral ones or even deep-learning-based solutions, can yield slightly better results but at a higher cost and complexity. We have experimented with bilateral filtering, using surface normals and distance as auxiliary features. Fig. 13 shows one of the results with the most significant improvement from the RIS application. In this scene, the variance estimates are very noisy. This causes a small spot in the image to have visibly higher noise than its surroundings. A bilateral filter can help to remove these by blurring more aggressively without harming the quality around discontinuities. Across most of the image, the result with bilateral filtering is the same as with our simpler approach, as evident from the approximately identical equal-sample error values in the figure. Full resolution images can be found in the supplemental.