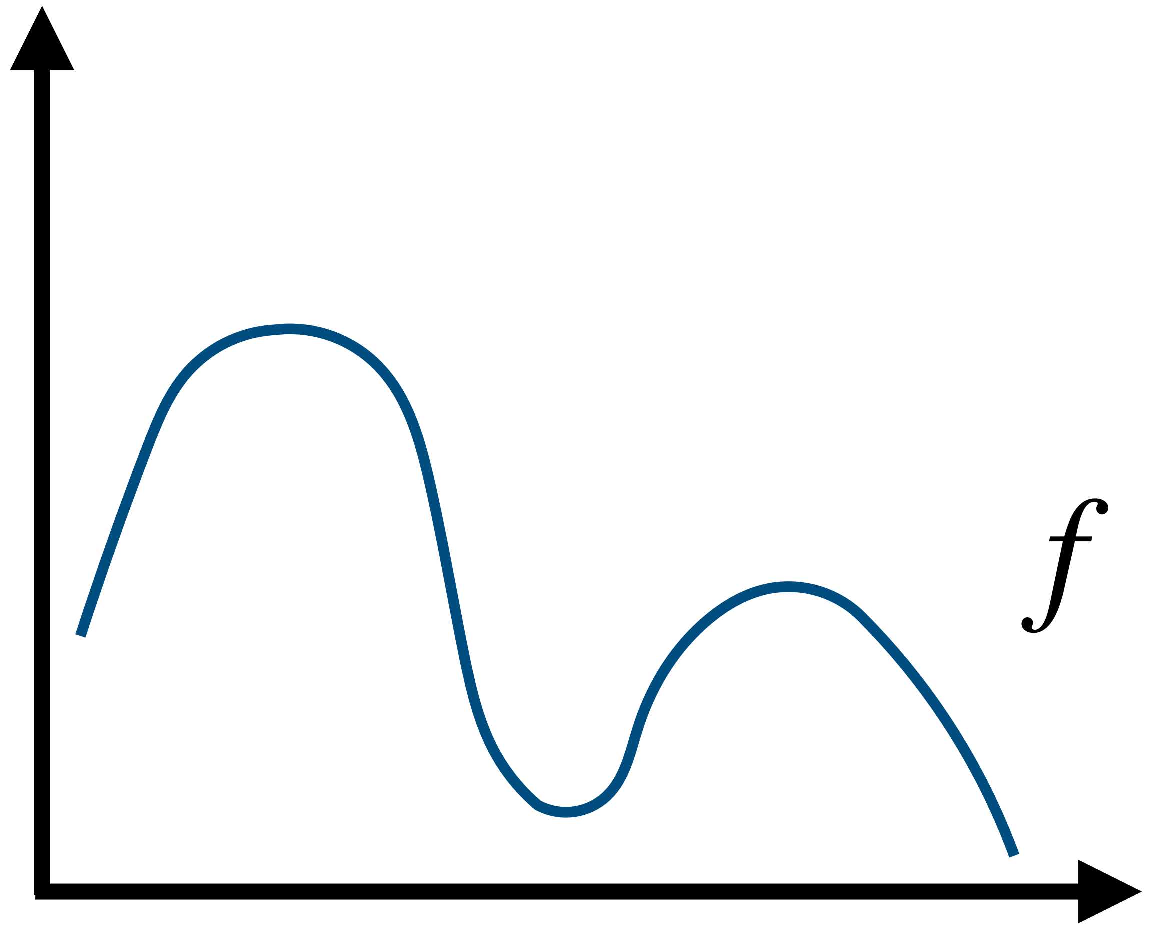
Multiple importance sampling++

UCSD CSE 272
Advanced Image Synthesis

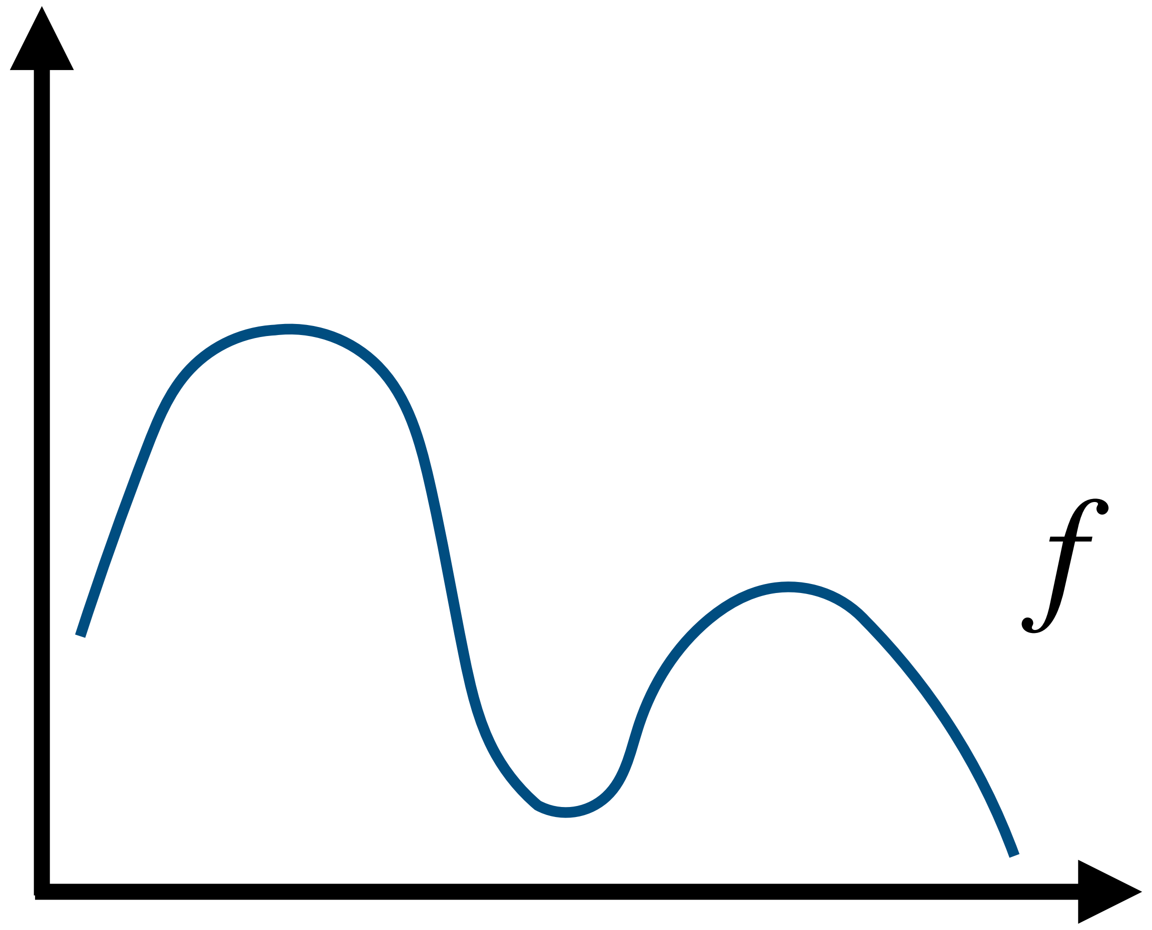
Tzu-Mao Li

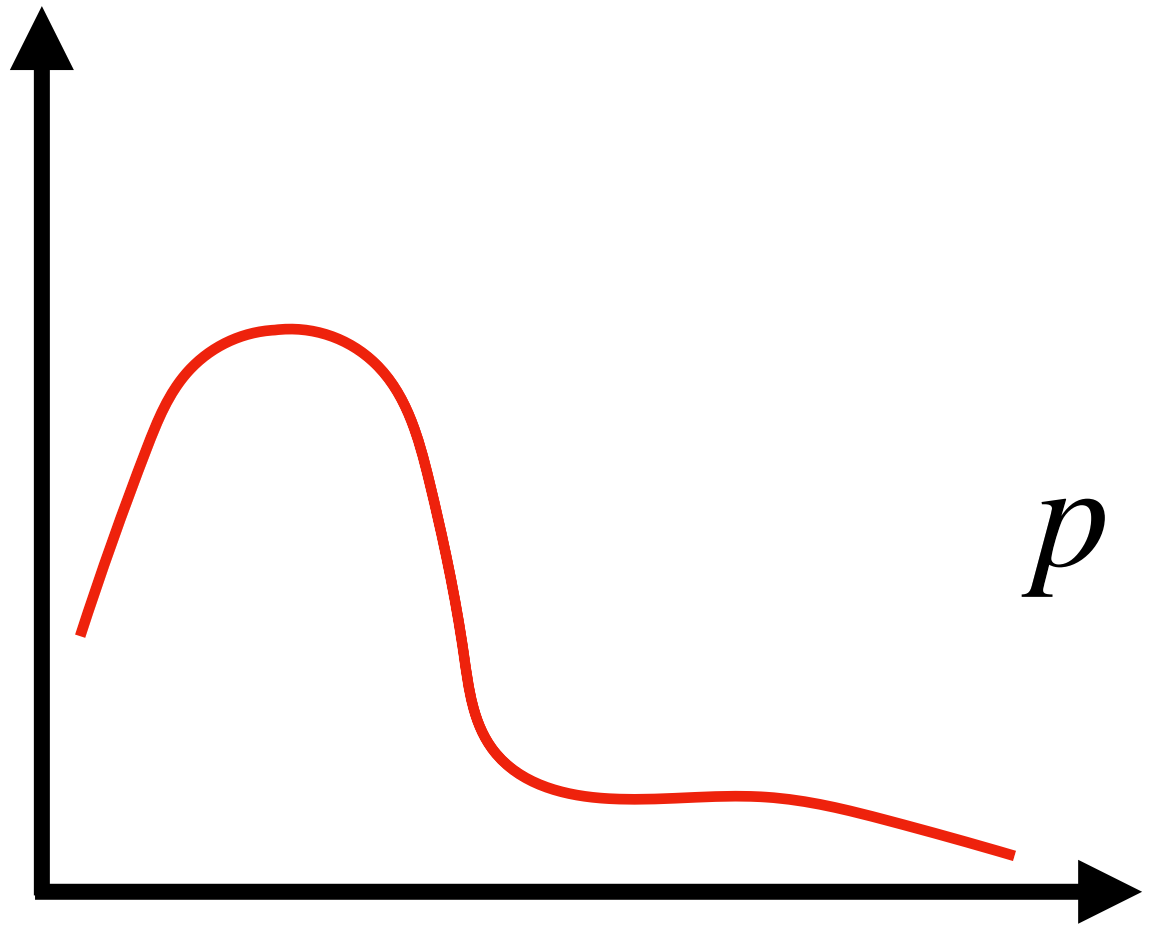
Monte Carlo integration



$$F = \int f(x)\mathrm{d}x$$

Monte Carlo integration



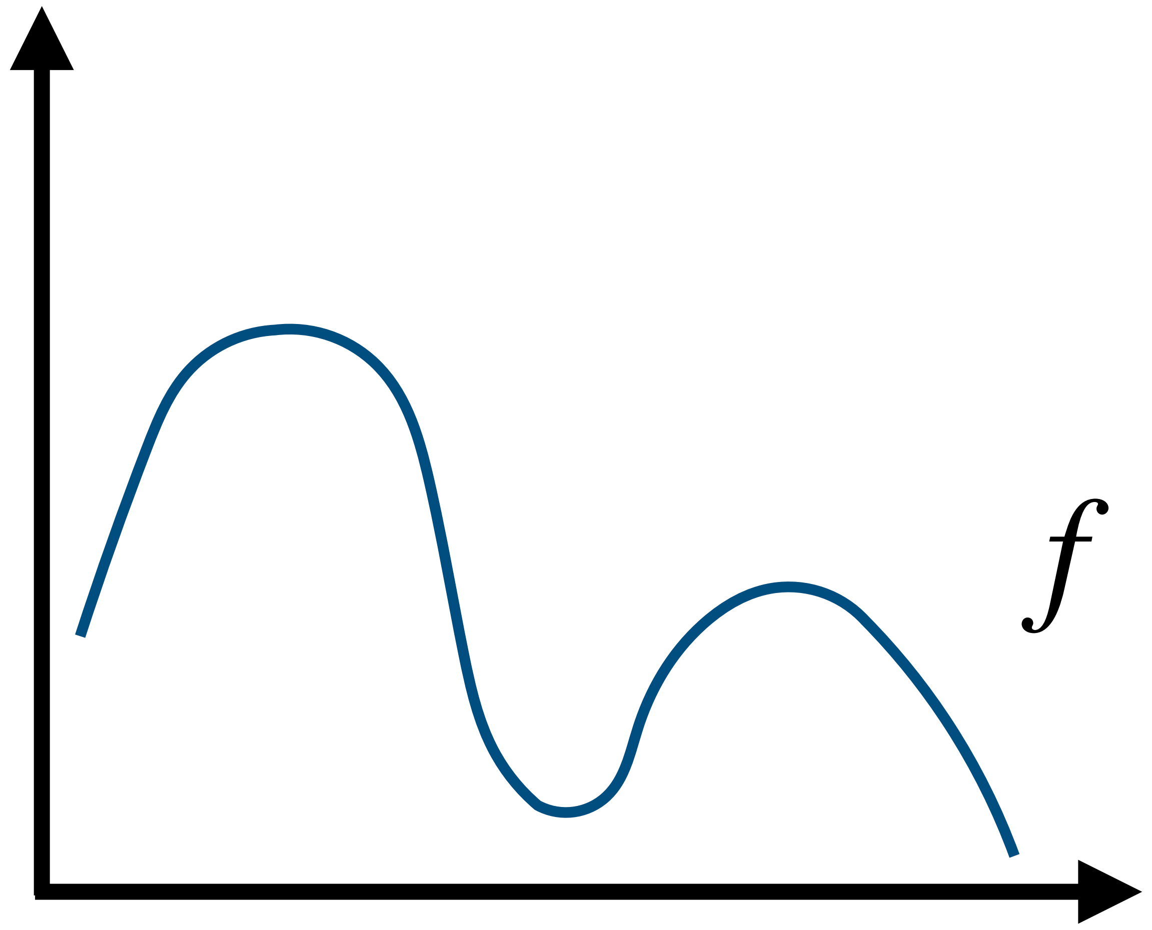


$$\langle F \rangle = \frac{1}{N} \sum_i \frac{f(X_i)}{p(X_i)}$$

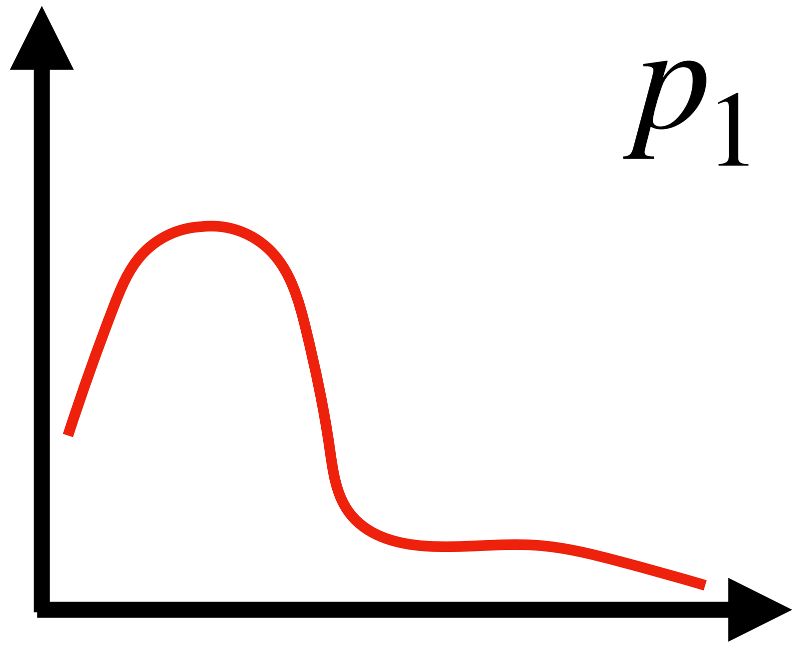
$$F = \int f(x)\mathrm{d}x$$

$$X_i \sim p$$

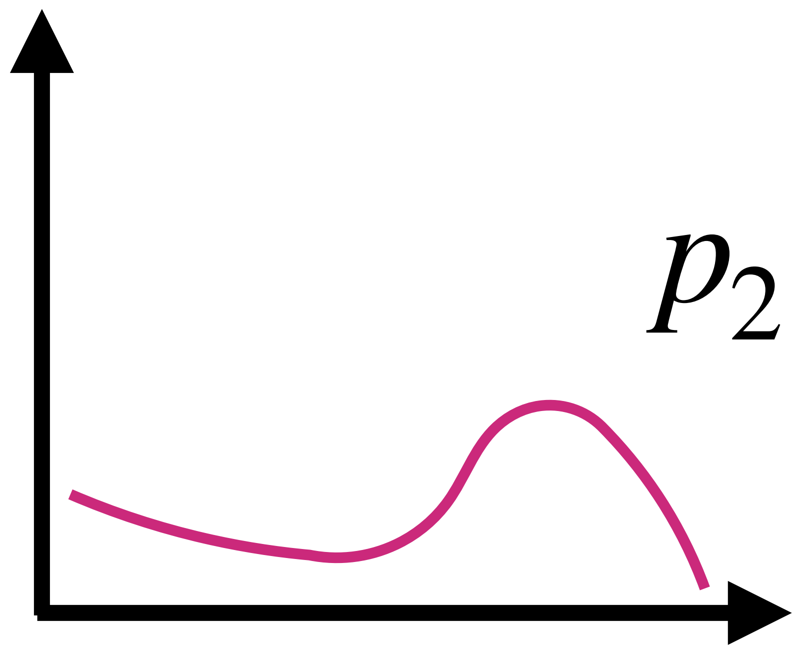
Monte Carlo integration



$$F = \int f(x)\mathrm{d}x$$



$$X_{1,i} \sim p_1$$

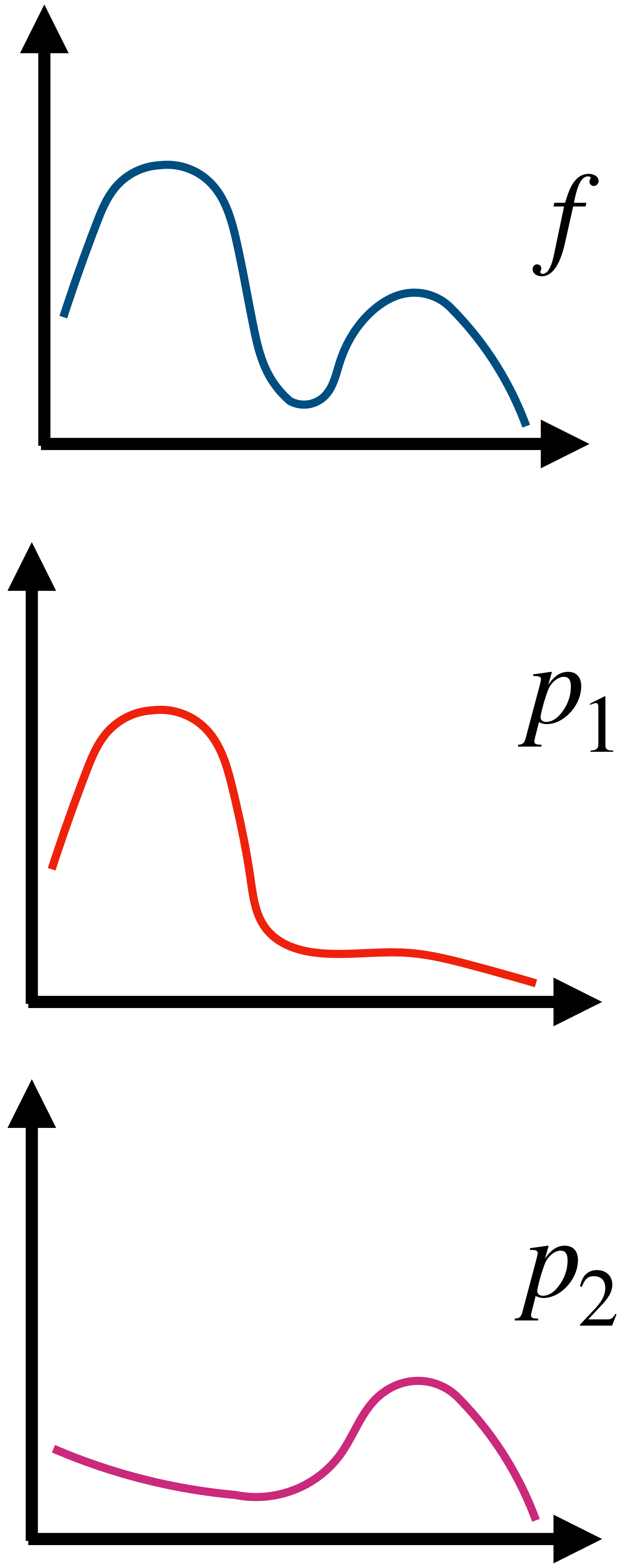


$$X_{2,j} \sim p_2$$

$$\langle F \rangle_1 = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)}$$

$$\langle F \rangle_2 = \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)}$$

Multiple importance sampling



idea: weighted average of the two estimators

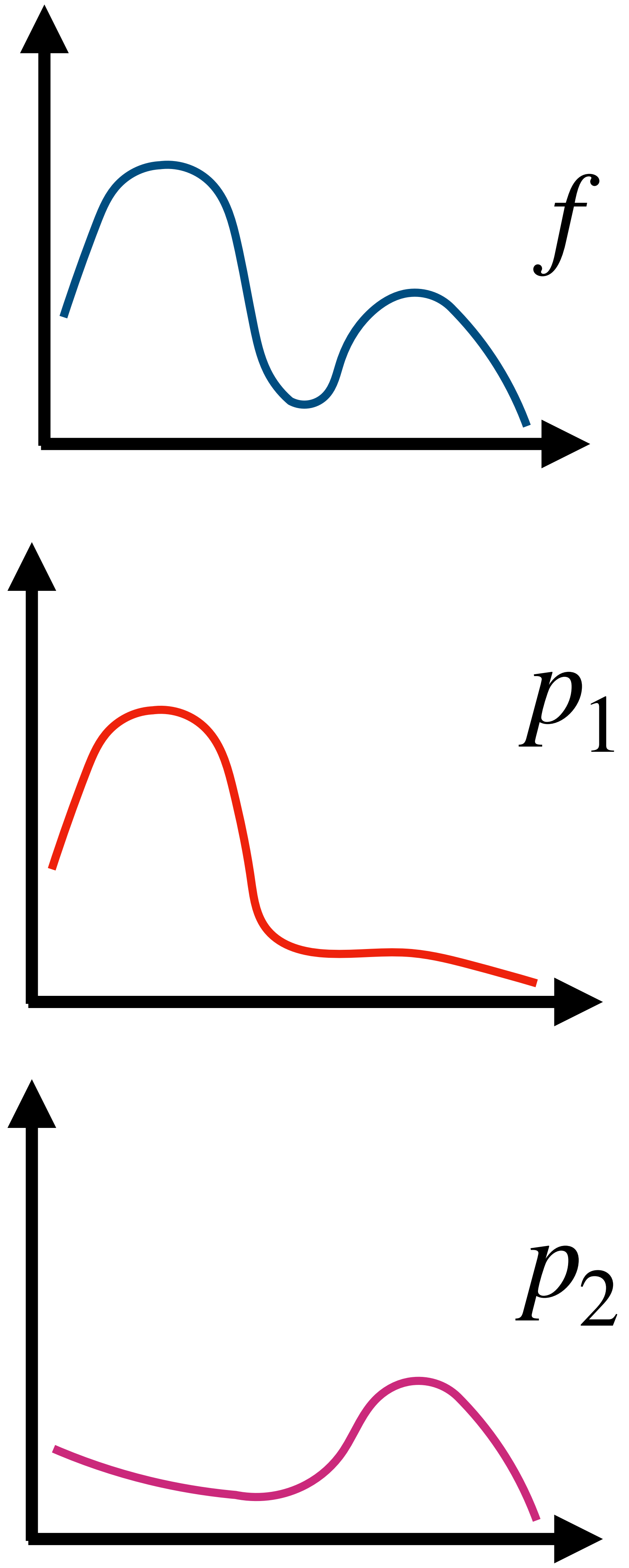
$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right) + \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

Optimally Combining Sampling Techniques for Monte Carlo Rendering

Eric Veach *Leonidas J. Guibas*

Computer Science Department
Stanford University

Multiple importance sampling

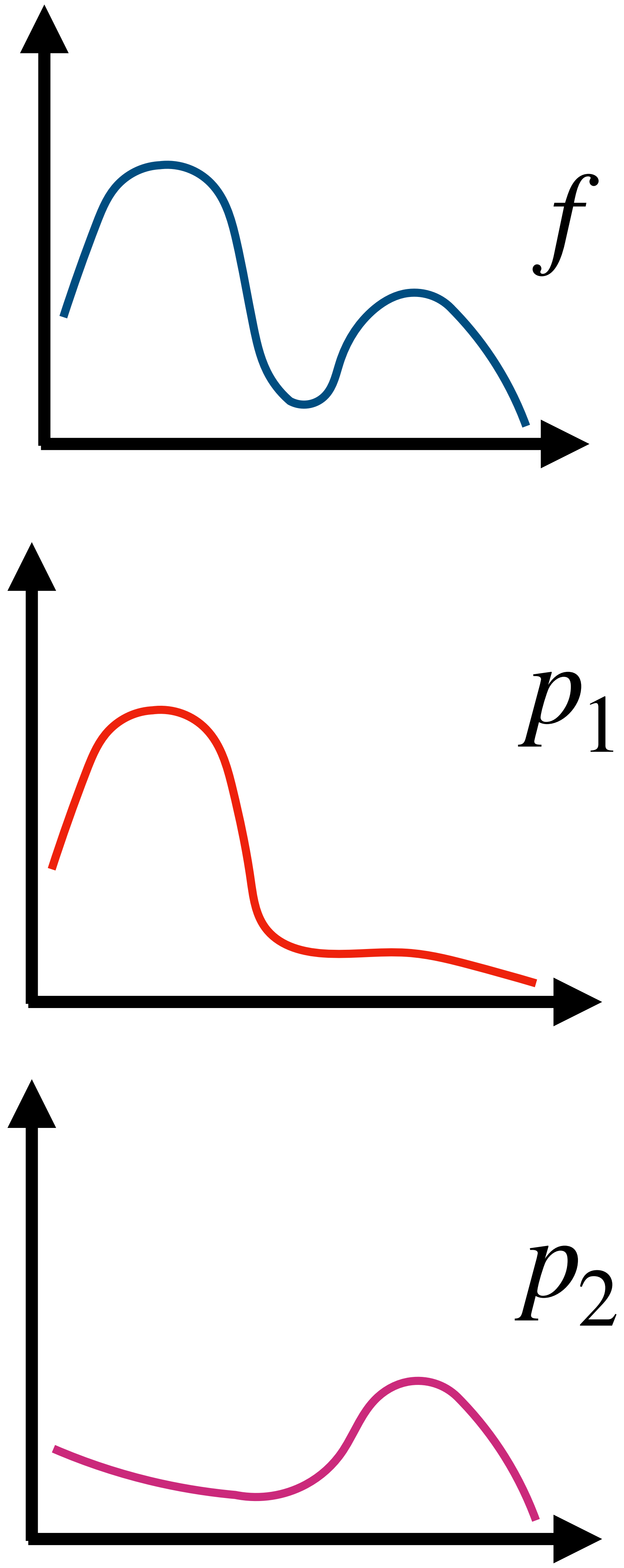


idea: weighted average of the two estimators

$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right) + \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

quiz: when will $\langle F \rangle$ be an unbiased estimator?

Multiple importance sampling



idea: weighted average of the two estimators

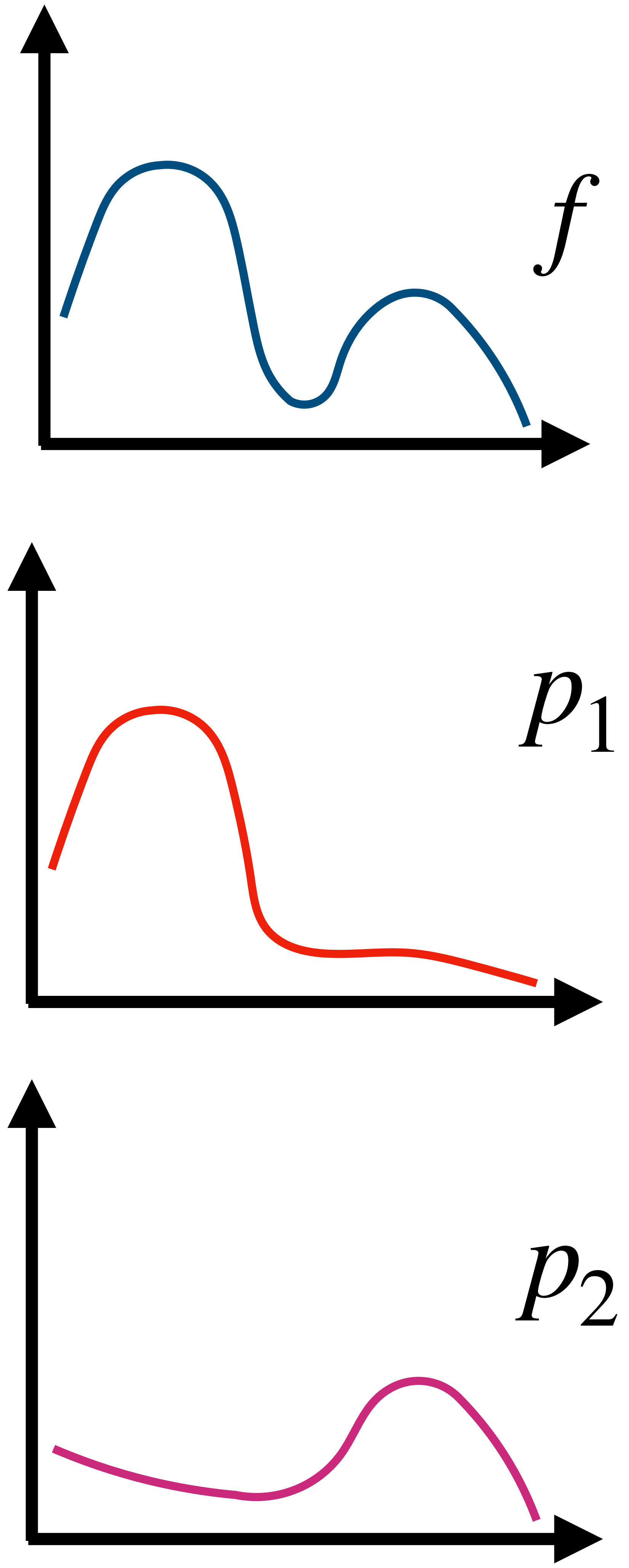
$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right)$$

$$+ \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

if $w_1(x) + w_2(x) = 1$,

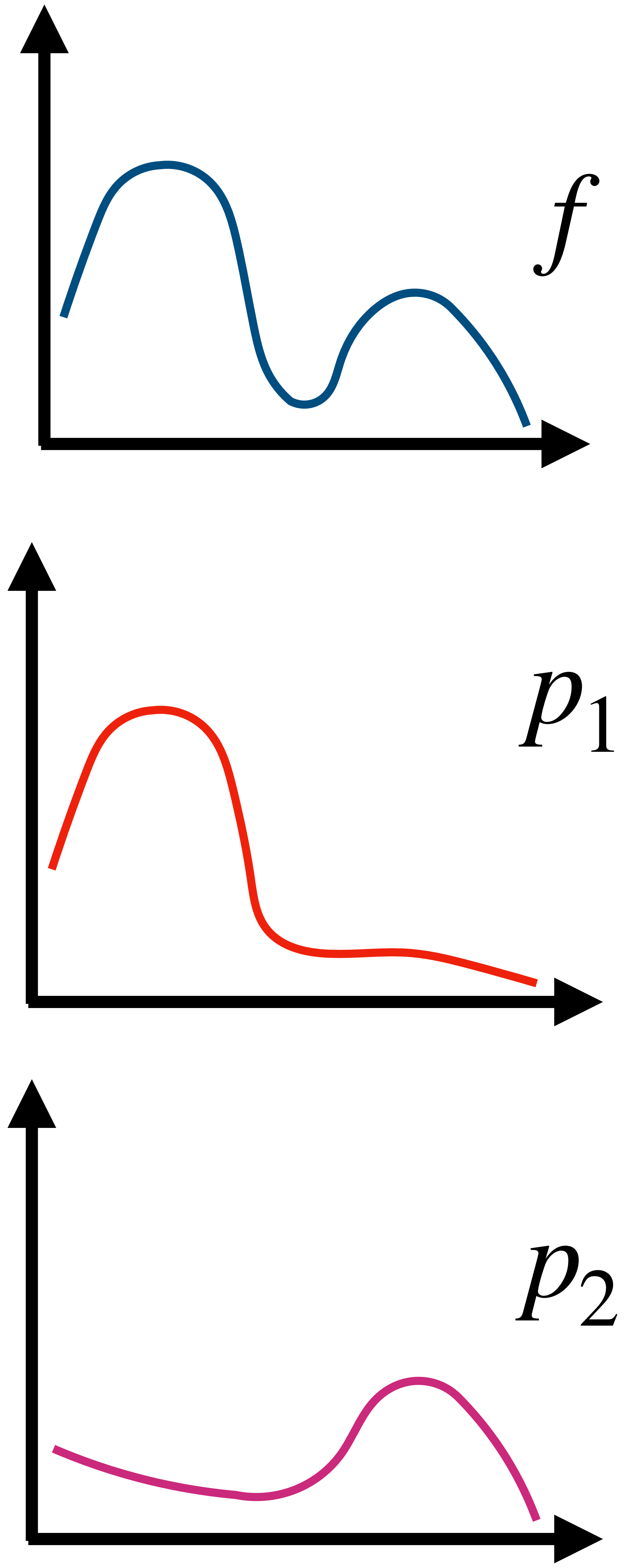
$$E\left[w_1 \frac{f(x(u_1))}{p_1(x(u_1))} + w_2 \frac{f(x(u_2))}{p_2(x(u_2))}\right] = w_1 E\left[\frac{f(x(u_1))}{p_1(x(u_1))}\right] + w_2 E\left[\frac{f(x(u_2))}{p_2(x(u_2))}\right] = \int f(x)\mathrm{d}x$$

How do we determine w_1 and w_2?



$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right) + \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

How do we determine w_1 and w_2?



goal: choose w_1 and w_2 such that $\mathrm{Var}\left[\langle F \rangle\right]$ is minimized

$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right)$$

$$+ \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

Veach's strategy

$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right) + \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

trick 1: assuming $X_{1,i}$ and $X_{2,j}$ are uncorrelated

$$\mathrm{Var}\left[\langle F \rangle\right] = \mathrm{Var}\left[\frac{1}{N_1} \sum \frac{f w_1}{p_1}\right] + \mathrm{Var}\left[\frac{1}{N_2} \sum \frac{f w_2}{p_2}\right]$$

Veach's strategy

$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right) + \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

trick 1: assuming $X_{1,i}$ and $X_{2,j}$ are uncorrelated

$$\mathrm{Var}\left[\langle F \rangle\right] = \mathrm{Var}\left[\frac{1}{N_1} \sum \frac{f w_1}{p_1}\right] + \mathrm{Var}\left[\frac{1}{N_2} \sum \frac{f w_2}{p_2}\right]$$

trick 2: minimize upper bound of the variance

$$\mathrm{Var}\left[X\right] = E\left[X^2\right] - E\left[X\right]^2 \leq E\left[X^2\right]$$

$w_i \propto N_i p_i$ is a good choice [Veach 1995]

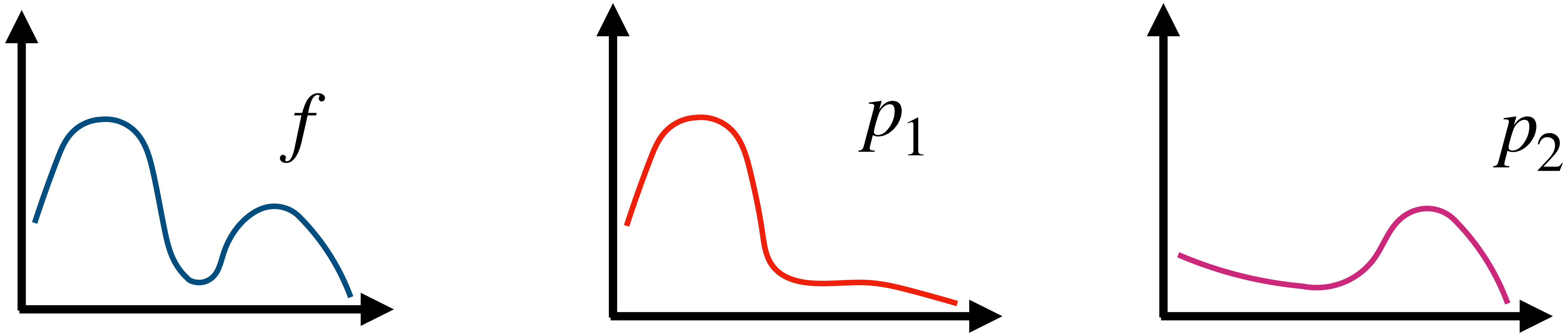
aka balance heuristic

$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right) + \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

$$w_1(x) = \frac{N_1 p_1(x)}{N_1 p_1(x) + N_2 p_2(x)} \text{ and } w_2(x) = \frac{N_2 p_2(x)}{N_1 p_1(x) + N_2 p_2(x)}$$

minimizes $E\left[\langle F \rangle^2\right]$

intuition: higher weight for higher sampling density



see CSE 168 for the proof

Can we do better than Veach?

$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right)$$

$$+ \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

trick 1: assuming $X_{1,i}$ and $X_{2,j}$ are uncorrelated

$$\mathrm{Var}[\langle F \rangle] = \mathrm{Var}\left[\frac{1}{N_1} \sum \frac{f w_1}{p_1}\right] + \mathrm{Var}\left[\frac{1}{N_2} \sum \frac{f w_2}{p_2}\right]$$

trick 2: minimize upper bound of the variance

$$\mathrm{Var}[X] = E\left[X^2\right] - E[X]^2 \leq E\left[X^2\right]$$

Can we do better than Veach?

$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f(X_{1,i})}{p_1(X_{1,i})} w_1(X_{1,i}) + \frac{1}{N_2} \sum_j \frac{f(X_{2,j})}{p_2(X_{2,j})} w_2(X_{2,j})$$

Optimal Multiple Importance Sampling

IVO KONDAPANENI*, Charles University, Prague
PETR VÉVODA*, Charles University, Prague and Render Legion, a. s.
PASCAL GRITTMANN, Saarland University
TOMÁŠ SKŘIVAN, IST Austria
PHILIPP SLUSALLEK, Saarland University and DFKI
JAROSLAV KŘIVÁNEK, Charles University, Prague and Render Legion, a. s.

trick 1: assuming $X_{1,i}$ and $X_{2,j}$ are uncorrelated

$$\mathrm{Var}[\langle F \rangle] = \mathrm{Var}\left[\frac{1}{N_1} \sum \frac{fw_1}{p_1}\right] + \mathrm{Var}\left[\frac{1}{N_2} \sum \frac{fw_2}{p_2}\right]$$

trick 2: minimize upper bound of the variance

$$\mathrm{Var}[X] = E\left[X^2\right] - E[X]^2 \leq E\left[X^2\right]$$

can we minimize variance directly?

Minimizing variance for multiple importance sampling

goal: choose w_1 and w_2 to minimize $\mathrm{Var}\left[\langle F\rangle\right]$ s.t. $w_1 + w_2 = 1$

$$\langle F\rangle = \frac{1}{N_1}\sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right) + \frac{1}{N_2}\sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

Minimizing variance for multiple importance sampling

goal: choose w_1 and w_2 to minimize $\text{Var}\left[\langle F \rangle\right]$ s.t. $w_1 + w_2 = 1$

$$\langle F \rangle = \frac{f\left(X_1\right)}{p_1\left(X_1\right)} w_1\left(X_1\right) + \frac{f\left(X_2\right)}{p_2\left(X_2\right)} w_2\left(X_2\right)$$

simplified without loss of generality

Minimizing variance for multiple importance sampling

goal: choose w_1 and w_2 to minimize $\mathrm{Var}\left[\langle F \rangle\right]$ s.t. $w_1 + w_2 = 1$

$$\langle F \rangle = \frac{f\left(X_1\right)}{p_1\left(X_1\right)} w_1\left(X_1\right) + \frac{f\left(X_2\right)}{p_2\left(X_2\right)} w_2\left(X_2\right)$$

simplified without loss of generality

$$\mathrm{Var}\left[\langle F \rangle\right] = \sum_i \int w_i^2 \frac{f^2}{p_i} - \left[\int w_i f\right]^2$$

Minimizing variance for multiple importance sampling

choose w_i to minimize

$$\mathrm{Var}\left[\langle F \rangle\right] = \sum_i \int w_i^2 \frac{f^2}{p_i} - \left[\int w_i f\right]^2$$

s.t.

$$\sum w_i = 1$$

Minimizing variance for multiple importance sampling

choose w_i and λ to minimize

$$V = \sum_i \int w_i^2 \frac{f^2}{p_i} - \left[\int w_i f \right]^2 + \lambda \left(\sum_i w_i - 1 \right)$$

https://en.wikipedia.org/wiki/Lagrange_multiplier

Minimizing variance for multiple importance sampling

choose w_i and λ to minimize

$$V = \sum_i \int w_i^2 \frac{f^2}{p_i} - \left[\int w_i f \right]^2 + \lambda \left(\sum_i w_i - 1 \right)$$

need to solve this using “calculus of variations”
https://en.wikipedia.org/wiki/Calculus_of_variations

generalizes differentiation with **vectors** to differentiation with **functions**

Minimizing variance for multiple importance sampling

choose w_i and λ to minimize

$$V = \sum_i \int w_i^2 \frac{f^2}{p_i} - \left[\int w_i f \right]^2 + \lambda \left(\sum_i w_i - 1 \right)$$

set derivatives to zero

$$\frac{\partial V}{\partial w_i} = 2w_i \frac{f^2}{p} - 2f \int w_i f + \lambda = 0 \qquad \frac{\partial V}{\partial \lambda} = \sum_i w_i - 1 = 0$$

Minimizing variance for multiple importance sampling

choose w_i and λ to solve

$$2w_i\frac{f^2}{p} - 2f\int w_i f + \lambda = 0$$

$$\sum_i w_i - 1 = 0$$

Minimizing variance for multiple importance sampling

choose w_i and λ to solve

$$2w_i\frac{f^2}{p} - 2fa_i + \lambda = 0 \qquad a_i = \int w_i f$$

$$\sum_i w_i - 1 = 0$$

Minimizing variance for multiple importance sampling

choose w_i and λ to solve

$$w_i = \frac{p_i}{f}a_i - \frac{1}{2}\frac{p_i}{f^2}\lambda \qquad a_i = \int w_i f$$

$$\sum_i w_i - 1 = 0$$

Minimizing variance for multiple importance sampling

choose w_i and λ to solve

$$w_i = \frac{p_i}{f} a_i - \frac{1}{2} \frac{p_i}{f^2} \lambda \qquad a_i = \int w_i f$$

$$\sum_i \frac{p_i}{f} a_i - \frac{1}{2} \frac{p_i}{f^2} \lambda = 1$$

Minimizing variance for multiple importance sampling

choose w_i and λ to solve

$$w_i = \frac{p_i}{f} a_i - \frac{1}{2} \frac{p_i}{f^2} \lambda \qquad a_i = \int w_i f$$

$$\lambda = \frac{2 \sum_i p_i a_i f - f^2}{\sum_i p_i}$$

Minimizing variance for multiple importance sampling

choose w_i to solve

$$w_i = \frac{p_i}{f} a_i - p_i \frac{\sum_j \frac{p_j}{f} a_j - 1}{\sum_j p_j} \qquad a_i = \int w_i f$$

Minimizing variance for multiple importance sampling

choose w_i to solve

$$w_i = \frac{p_i}{f} a_i - p_i \frac{\sum_j \frac{p_j}{f} a_j - 1}{\sum_j p_j} \qquad a_i = \int w_i f$$

integrate both side with f

$$\int w_i f = \alpha_i = a_i \int p_i - \int p_i \frac{\sum_j p_j a_j - f}{\sum_j p_j}$$

Minimizing variance for multiple importance sampling

choose w_i to solve

$$w_i = \frac{p_i}{f} a_i - p_i \frac{\sum_j \frac{p_j}{f} a_j - 1}{\sum_j p_j} \qquad a_i = \int w_i f$$

$$\int p_i \frac{\sum_j p_j a_j}{\sum_j p_j} = \int \frac{f p_i}{\sum_j p_j}$$

Minimizing variance for multiple importance sampling

choose w_i to solve

$$w_i = \frac{p_i}{f} a_i - p_i \frac{\sum_j \frac{p_j}{f} a_j - 1}{\sum_j p_j}$$

$$A_{ij} = \int \frac{p_i p_j}{\sum_k p_k}$$

$$\begin{bmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$b_i = \int \frac{f p_i}{\sum_k p_k}$$

Minimizing variance for multiple importance sampling

the MIS weight that minimizes the variance!

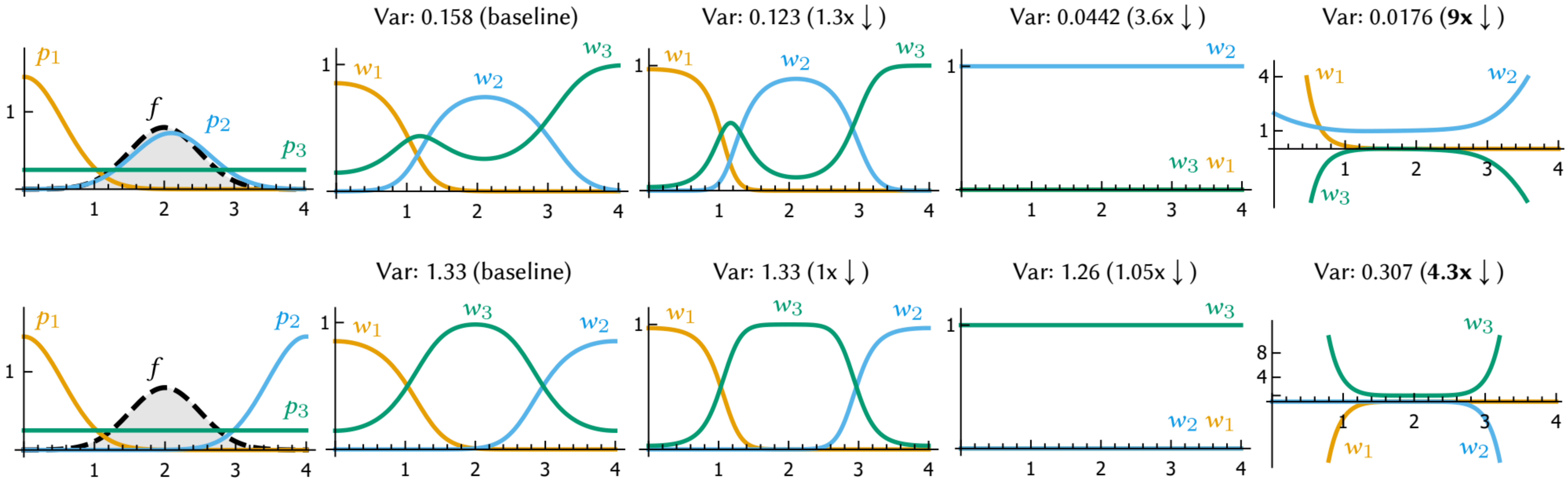
$$w_i = \frac{p_i}{f} a_i - p_i \frac{\sum_j \frac{p_j}{f} a_j - 1}{\sum_j p_j}$$

$$A_{ij} = \int \frac{p_i p_j}{\sum_k p_k}$$

$$\begin{bmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$b_i = \int \frac{f p_i}{\sum_k p_k}$$

Optimal MIS weight requires negative weights



a) integrand and sampling techniques

b) balance heuristic weights

c) power heuristic weights

d) best-technique heuristic weights

e) optimal weights (unconstrained sign)

Inuitition: control variates of mixture PDFs lead to optimal MIS

control variates: variance reduction using known integrals

$$\int f(x)\mathrm{d}x = \int f(x) - g(x)\mathrm{d}x + \int g(x)\mathrm{d}x = \int f(x) - g(x)\mathrm{d}x + G$$

Inuitition: control variates of mixture PDFs lead to optimal MIS

control variates: variance reduction using known integrals

$$\int f(x)\mathrm{d}x = \int f(x) - \sum_i \alpha_i p_i(x)\mathrm{d}x + \sum_i \alpha_i$$

Inuitition: control variates of mixture PDFs lead to optimal MIS

control variates: variance reduction using known integrals

$$\int f(x)\mathrm{d}x = \int f(x) - \sum_i \alpha_i p_i(x)\mathrm{d}x + \sum_i \alpha_i$$

optimal α_i satisfies

$$\begin{bmatrix} \int \frac{p_1 p_1}{p_1 + p_2} & \int \frac{p_1 p_2}{p_1 + p_2} \\ \int \frac{p_2 p_1}{p_1 + p_2} & \int \frac{p_2 p_2}{p_1 + p_2} \end{bmatrix} \alpha = \begin{bmatrix} \int \frac{f p_1}{p_1 + p_2} \\ \int \frac{f p_2}{p_1 + p_2} \end{bmatrix}$$

Inuitition: control variates of mixture PDFs lead to optimal MIS

control variates: variance reduction using known integrals

$$\int f(x)\mathrm{d}x = \int f(x) - \sum_i \alpha_i p_i(x)\mathrm{d}x + \sum_i \alpha_i$$

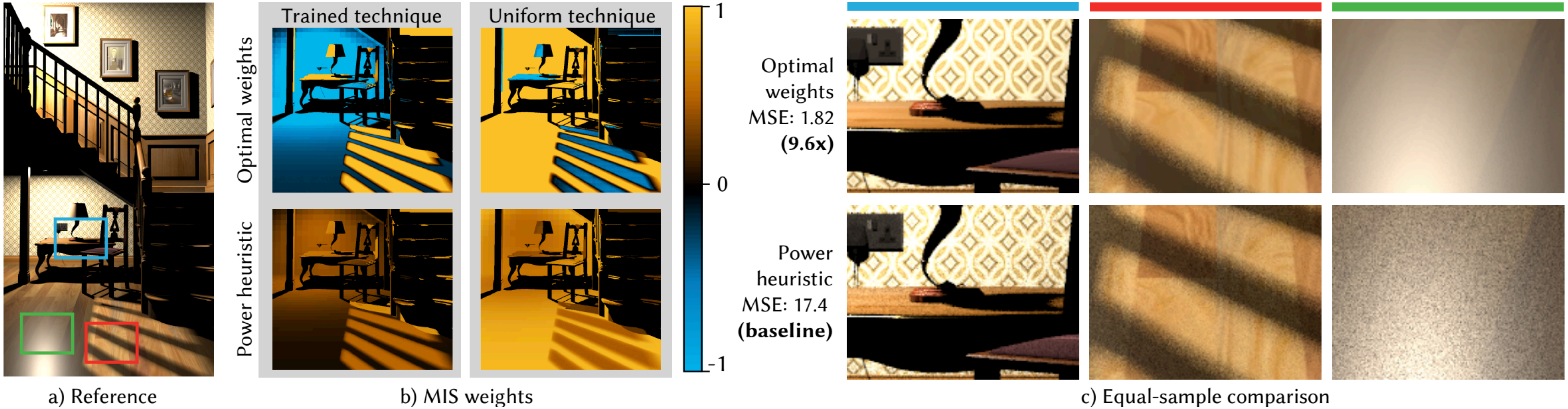
optimal α_i satisfies

$$\begin{bmatrix} \int \frac{p_1 p_1}{p_1+p_2} & \int \frac{p_1 p_2}{p_1+p_2} \\ \int \frac{p_2 p_1}{p_1+p_2} & \int \frac{p_2 p_2}{p_1+p_2} \end{bmatrix} \alpha = \begin{bmatrix} \int \frac{f p_1}{p_1+p_2} \\ \int \frac{f p_2}{p_1+p_2} \end{bmatrix}$$

Safe and Effective Importance Sampling

Art OWEN and Yi ZHOU

Optimal MIS can outperform MIS



Downside of Optimal MIS

Downside of Optimal MIS

computing the weights requires solving integrals!

$$w_i = \frac{p_i}{f} a_i - p_i \frac{\sum_j \frac{p_j}{f} a_j - 1}{\sum_j p_j}$$

$$A_{ij} = \int \frac{p_i p_j}{\sum_k p_k}$$

$$\begin{bmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$b_i = \int \frac{f p_i}{\sum_k p_k}$$

Variance-aware MIS

- combine balance heuristic with empirical variance

$$w_1 = \frac{\nu_1 p_1}{\nu_1 p_1 + \nu_2 p_2}$$

$$\nu_i = \frac{E\left[\left(\frac{f_i}{p_i}\right)^2\right]}{\mathrm{Var}\left[\frac{f_i}{p_i}\right]}$$

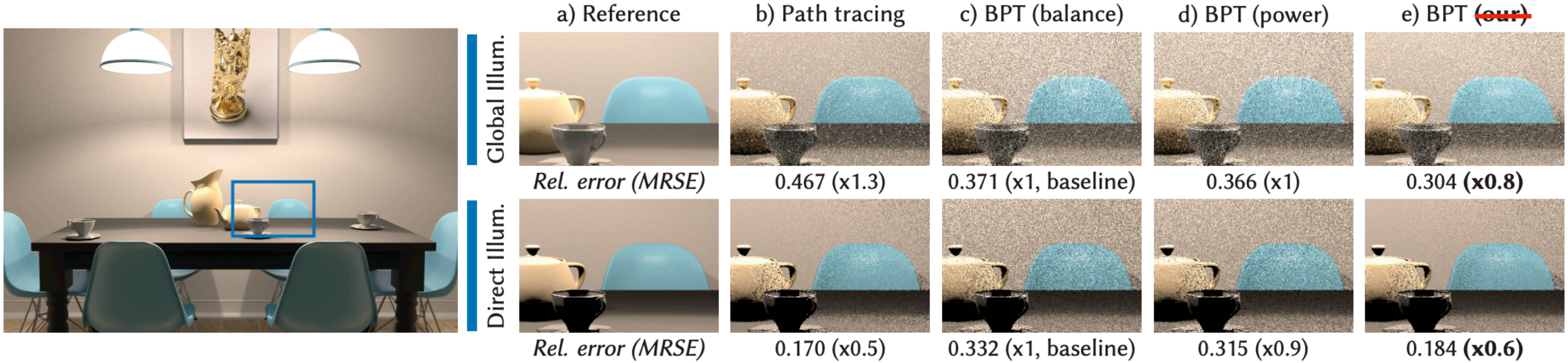
if $\nu_i = 1$, balance heuristic is optimal, if $\nu_i \to 0$, variance is high and we should distrust technique i

Variance-Aware Multiple Importance Sampling

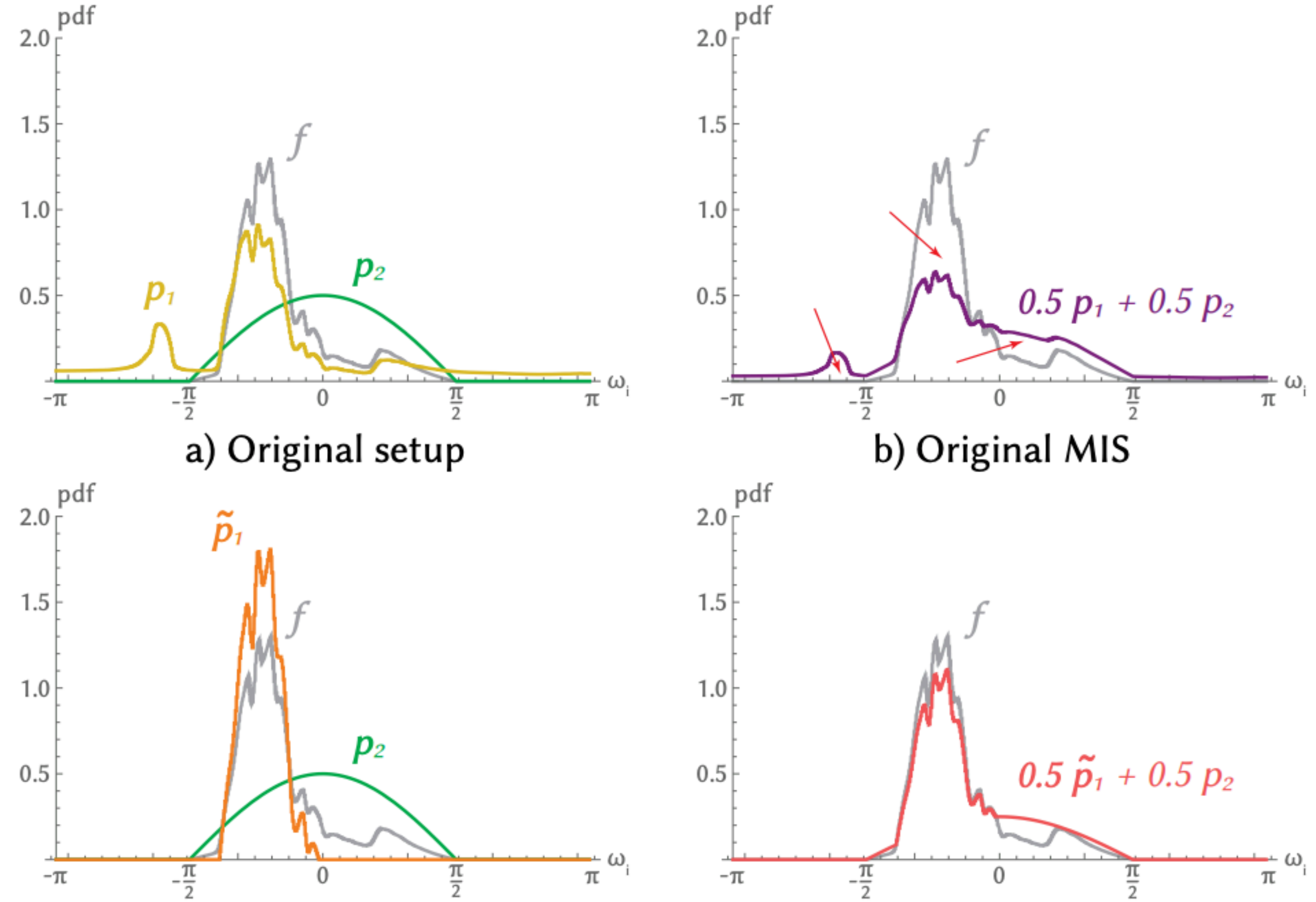
PASCAL GRITTMANN, Saarland University, Germany
ILIYAN GEORGIEV, Autodesk, United Kingdom
PHILIPP SLUSALLEK, DFKI and Saarland University, Germany
JAROSLAV KŘIVÁNEK, Charles University and Chaos Czech a. s., Czech Republic

Variance-aware MIS takes stratification into consideration

(Gritmann et al.)



MIS compensation: modify a PDF based on other techniques



a) Original setup

b) Original MIS

c) Setup with the optimized technique

d) MIS with the optimized technique

MIS Compensation: Optimizing Sampling Techniques in Multiple Importance Sampling

ONDŘEJ KARLÍK, Chaos Czech a. s.
MARTIN ŠIK, Chaos Czech a. s.
PETR VÉVODA, Charles University, Prague and Chaos Czech a. s.
TOMÁŠ SKŘIVAN, IST Austria
JAROSLAV KŘIVÁNEK, Charles University, Prague and Chaos Czech a. s.

MIS compensation: modify a PDF based on other techniques

super simple to implement for discrete PMFs!

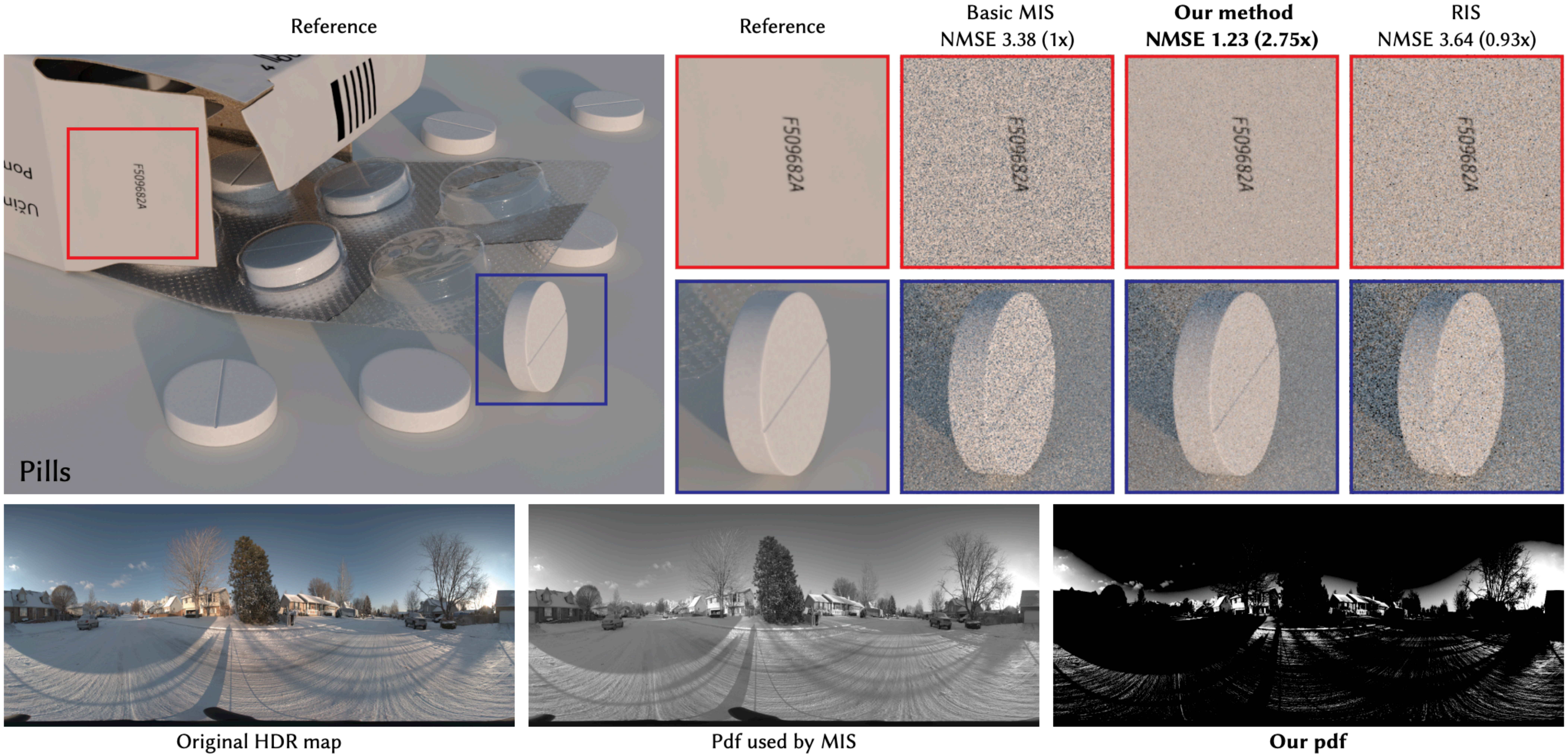
```
void MIS_compensation()
{
    for (int i = 0; i < N; ++i) {
        probability[i] = max(probability[i] - averageValue, 0.f);
    }
}
```

https://www.iliyan.com/publications/RenderingCourse2020/RenderingCourse2020_Notes_rev1.pdf

MIS Compensation: Optimizing Sampling Techniques in Multiple Importance Sampling

ONDŘEJ KARLÍK, Chaos Czech a. s.
MARTIN ŠIK, Chaos Czech a. s.
PETR VÉVODA, Charles University, Prague and Chaos Czech a. s.
TOMÁŠ SKŘIVAN, IST Austria
JAROSLAV KŘIVÁNEK, Charles University, Prague and Chaos Czech a. s.

MIS compensation: modify a PDF based on other techniques



Can we do better than Veach?

$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right) + \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

trick 1: assuming $X_{1,i}$ and $X_{2,j}$ are uncorrelated

$$\mathrm{Var}\left[\langle F \rangle\right] = \mathrm{Var}\left[\frac{1}{N_1} \sum \frac{f w_1}{p_1}\right] + \mathrm{Var}\left[\frac{1}{N_2} \sum \frac{f w_2}{p_2}\right]$$

trick 2: minimize upper bound of the variance

$$\mathrm{Var}\left[X\right] = E\left[X^2\right] - E\left[X\right]^2 \leq E\left[X^2\right]$$

Can we do better than Veach?

$$\langle F \rangle = \frac{1}{\boxed{N_1}} \sum_i \frac{f(X_{1,i})}{p_1(X_{1,i})} w_1(X_{1,i})$$

$$+ \frac{1}{\boxed{N_2}} \sum_j \frac{f(X_{2,j})}{p_2(X_{2,j})} w_2(X_{2,j})$$

how do we choose N_1 and N_2?

trick 1: assuming $X_{1,i}$ and $X_{2,j}$ are uncorrelated

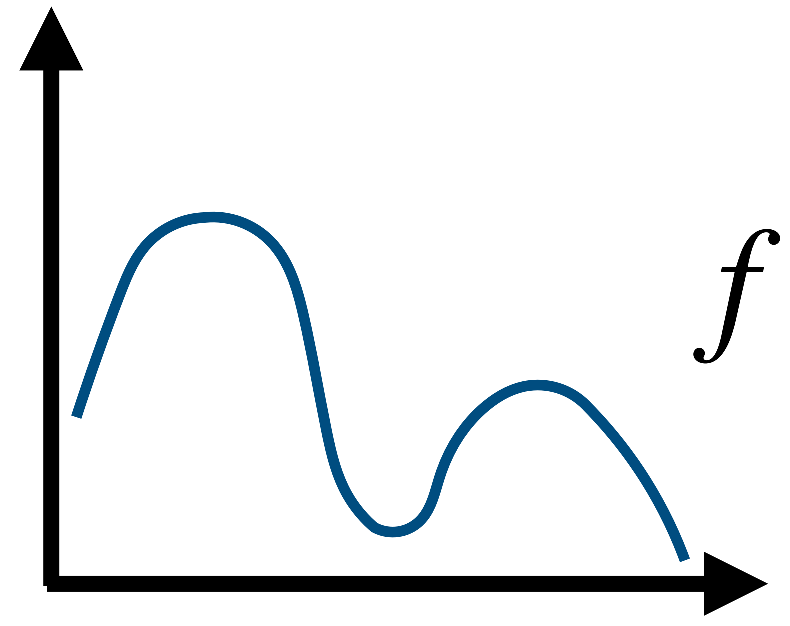
$$\mathrm{Var}[\langle F \rangle] = \mathrm{Var}\left[\frac{1}{N_1} \sum \frac{fw_1}{p_1}\right] + \mathrm{Var}\left[\frac{1}{N_2} \sum \frac{fw_2}{p_2}\right]$$

trick 2: minimize upper bound of the variance

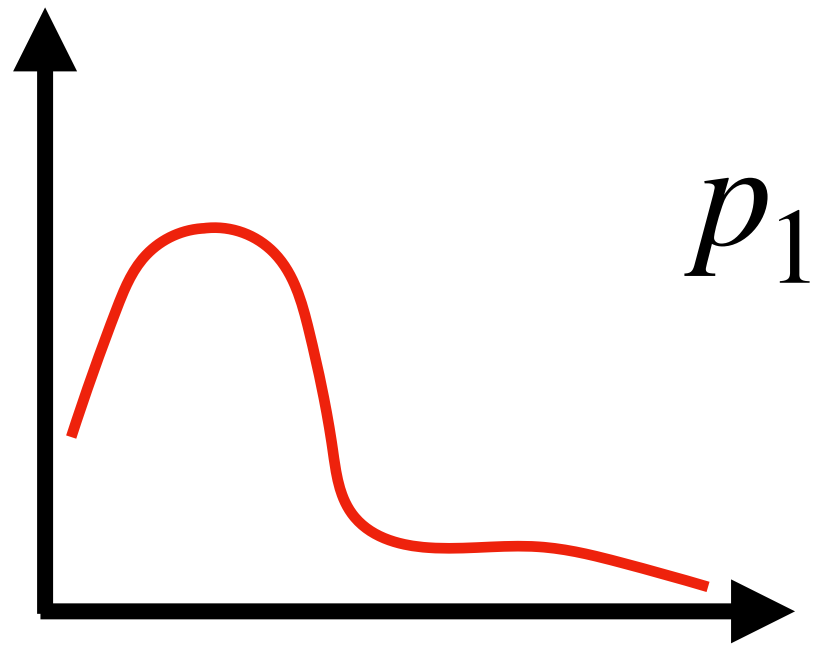
$$\mathrm{Var}[X] = E[X^2] - E[X]^2 \leq E[X^2]$$

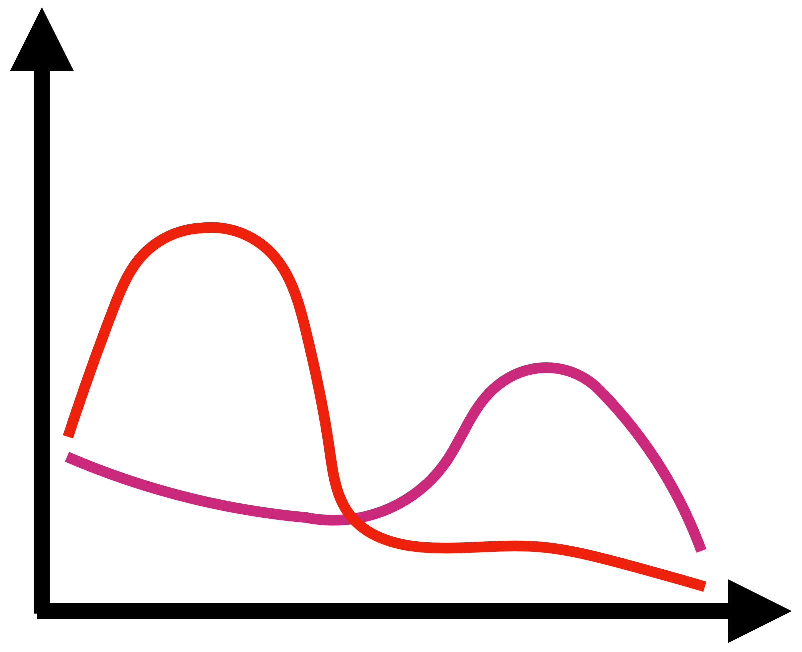
Can we randomly choose one technique?

effectively blending the two PDFs into one

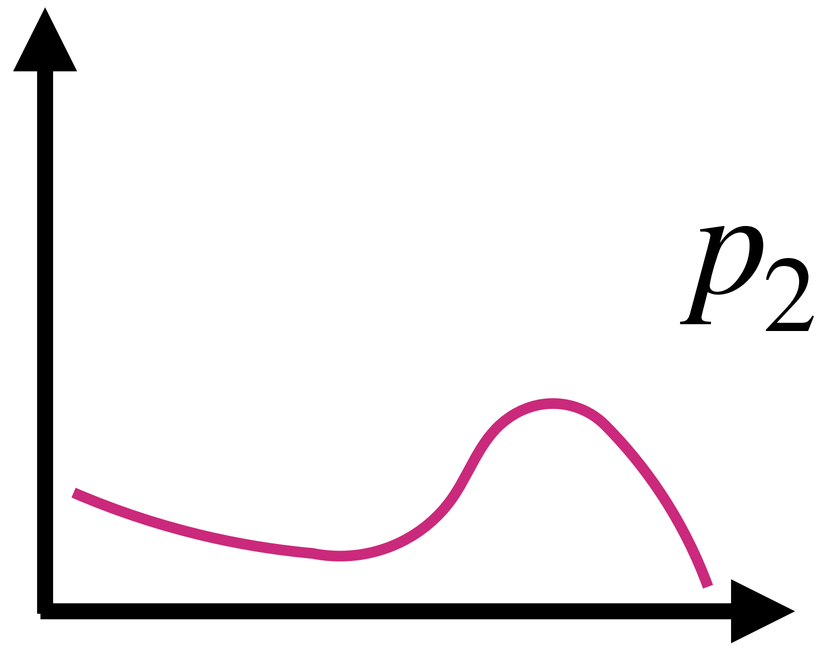


$$\langle F \rangle = \frac{1}{N} \sum_i \frac{f(X_i)}{p(X_i)}$$





$$p = \frac{w_1 p_1 + w_2 p_2}{2}$$



One-sample MIS

- instead of sampling from both p_1 and p_2, we randomly choose one of them

$$\langle F \rangle_{\text{ms}} = \frac{f}{p_1} w_1 + \frac{f}{p_2} w_2$$

"multi-sample" MIS

$$\langle F \rangle_{\text{os}} = w_i \frac{f}{\frac{1}{2} p_i}$$

"one-sample" MIS

Balance heuristic is optimal in one-sample MIS

$$\langle F \rangle_{\text{OS}} = w_i \frac{f}{\frac{1}{2} p_i}$$

$$w_i = \frac{p_i}{p_1 + p_2}$$

Balance heuristic is optimal in one-sample MIS

$$\langle F \rangle_{\text{OS}} = w_i \frac{f}{\frac{1}{2} p_i} = \frac{f}{\frac{1}{2}\left(p_1 + p_2\right)}$$

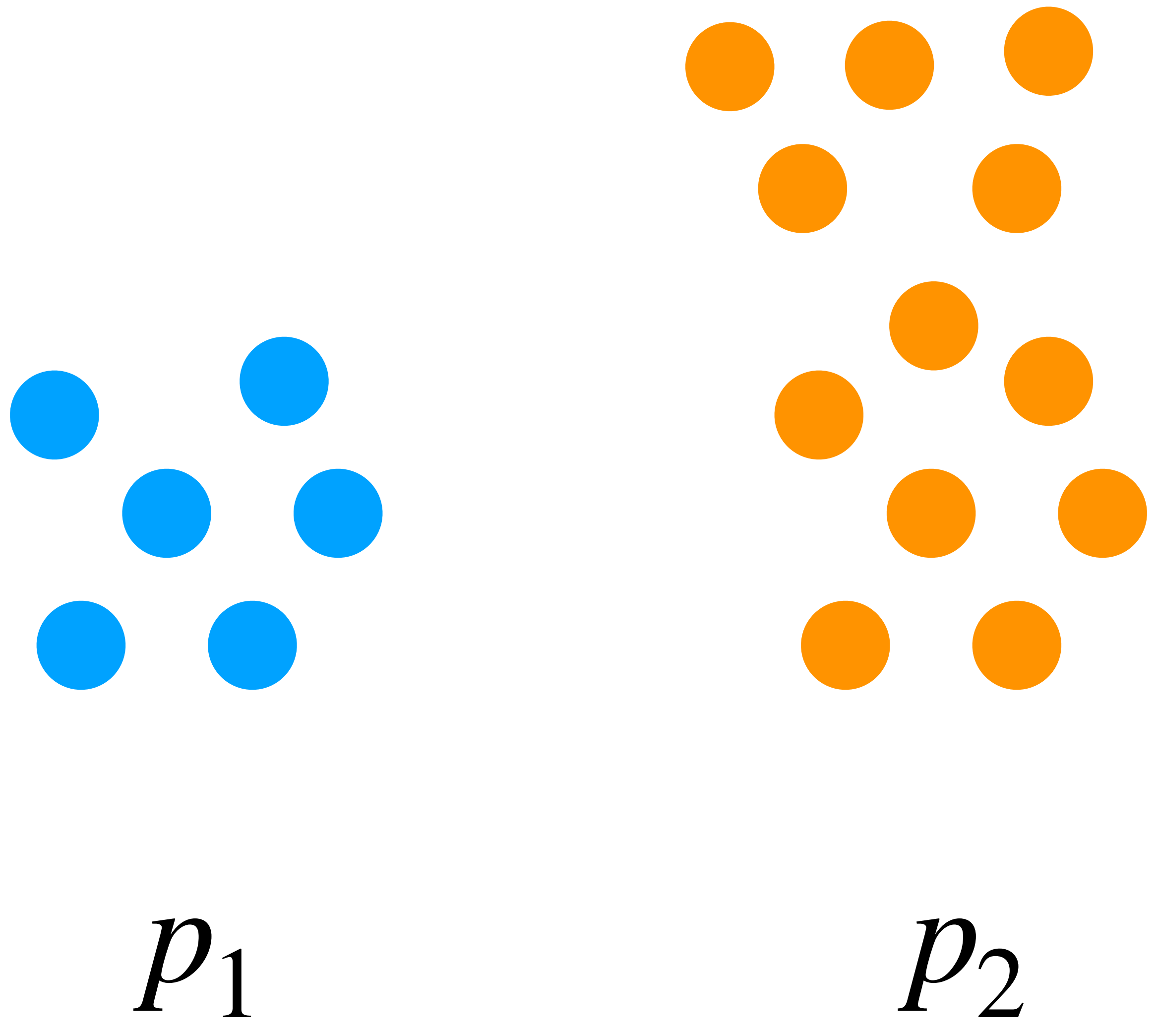
$$w_i = \frac{p_i}{p_1 + p_2}$$

- one-sample MIS = just average the distribution
  - not really doing anything!

MIS is helpful because of stratification!

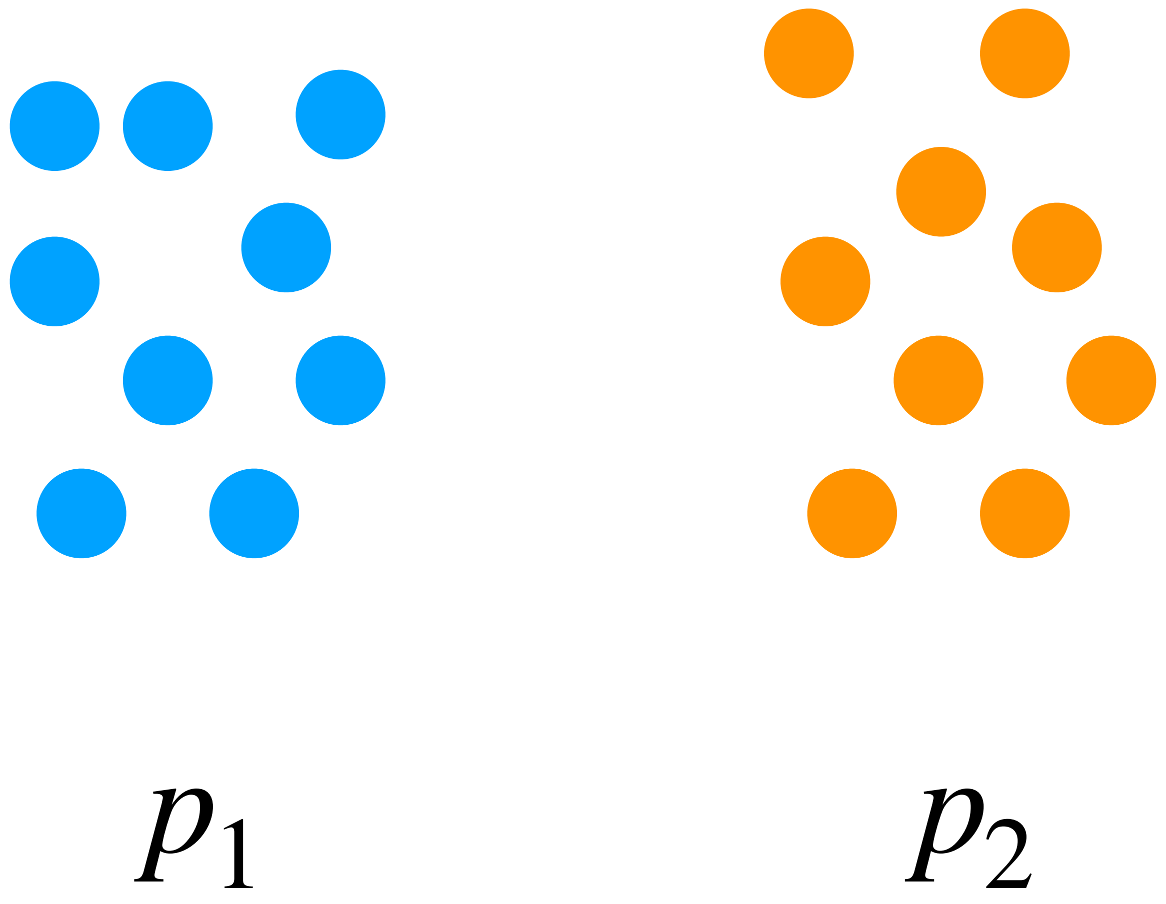
- stratification ensures we have the same amount of samples for each sampling distribution

one-sample MIS



extra variance comes from uneven sample counts

multi-sample MIS



more variance reduction

Many alternatives between one-sample and multi-sample

$\mathcal{W}_1$: $\varphi_{\mathcal{P}_n}(\mathbf{x}_n) = \varphi_{j_{1:n-1}}(\mathbf{x}_n) = p(\mathbf{x}_n|j_{1:n-1})$
Since the sampling process is sequential, this option is of particular interest. It interprets the proposal pdf as the conditional density of $\mathbf{x}_n$ given all the previous proposal indexes of the sampling process.

$\mathcal{W}_2$: $\varphi_{\mathcal{P}_n}(\mathbf{x}_n) = \varphi_{j_n}(\mathbf{x}_n) = p(\mathbf{x}_n|j_n) = q_{j_n}(\mathbf{x}_n)$
It interprets that if the index j_n is known, $\varphi_{\mathcal{P}_n}$ is the proposal q_{j_n}.

$\mathcal{W}_3$: $\varphi_{\mathcal{P}_n}(\mathbf{x}_n) = p(\mathbf{x}_n)$
It interprets that $\mathbf{x}_n$ is a realization of the marginal $p(\mathbf{x}_n)$. This is probably the most "natural" option (as it does not assume any further knowledge in the generation of $\mathbf{x}_n$) and is a usual choice for the calculation of the weights in some of the existing MIS schemes (see Section 5).

$\mathcal{W}_4$: $\varphi_{\mathcal{P}_n}(\mathbf{x}_n) = \varphi_{j_{1:N}}(\mathbf{x}_n) = f(\mathbf{x}_n|j_{1:N}) = \frac{1}{N}\sum_{k=1}^{N} q_{j_k}(\mathbf{x}_n)$
This interpretation makes use of the distribution of the r.v. $\mathbf{X}$ conditioned on the whole set of indexes (defined in Section 3.5).

$\mathcal{W}_5$: $\varphi_{\mathcal{P}_n}(\mathbf{x}_n) = \varphi(\mathbf{x}_n) = f(\mathbf{x}_n) = \frac{1}{N}\sum_{k=1}^{N} q_k(\mathbf{x}_n)$
This option considers that all the $\mathbf{x}_n$ are realizations of the r.v. $\mathbf{X}$ defined in Section 3.5 (see Appendix A for a thorough discussion of this interpretation).

`[R1]`: *Sampling with replacement, $\mathcal{S}_1$, and weight denominator $\mathcal{W}_2$:*
For the weight calculation of the n-th sample, only the proposal selected for generating the sample is evaluated in the denominator.

`[R2]`: *Sampling with replacement, $\mathcal{S}_1$, and weight denominator $\mathcal{W}_4$:*
With the N selected indexes j_n, for $n = 1, \ldots, N$, one forms a mixture comprising all the corresponding proposal pdfs. The weight calculation of the n-th sample considers this *a posteriori* mixture evaluated at the n-th sample in the denominator, i.e., some proposals might be used more than once while other proposals might not be used.

`[R3]`: *Sampling with replacement, $\mathcal{S}_1$, and weight denominator $\mathcal{W}_1$, $\mathcal{W}_3$, or $\mathcal{W}_5$:*
For the weight calculation of the n-th sample, the denominator applies the value of the n-th sample to the whole mixture ψ composed of the set of initial proposal pdfs (i.e., the function in the denominator of the weight does not depend on the sampling process). This is the approach followed by the so called mixture PMC method [Cappé et al., 2008].

`[N1]`: *Sampling without replacement (random or deterministic), $\mathcal{S}_2$ or*
weight denominator $\mathcal{W}_2$ (for $\mathcal{S}_2$) or $\mathcal{W}_1$, $\mathcal{W}_2$, or $\mathcal{W}_3$ (for $\mathcal{S}_3$):
For calculating the denominator of the n-th weight, the specific p
used for the generation of the sample is used. This is the approach fre
used in particle filtering [Gordon et al., 1993] and in the standar
method [Cappé et al., 2004].

`[N2]`: *Sampling without replacement (random), $\mathcal{S}_2$, and weight denomina*
This MIS implementation draws one sample from each proposal,
order matters (it must be random) since the calculation of the n-th
uses for the evaluation of the denominator the mixture pdf formed
proposal pdfs that were still available at the generation of the n-th

`[N3]`: *Sampling without replacement (random or deterministic), $\mathcal{S}_2$ or*
weight denominator $\mathcal{W}_3$, $\mathcal{W}_4$, or $\mathcal{W}_5$ (for $\mathcal{S}_2$), or $\mathcal{W}_4$ or $\mathcal{W}_5$ (for $\mathcal{S}_3$
In the calculation of the n-th weight, one uses for the denomina
whole mixture. This is the approach, for instance, of [Martino et al
Cornuet et al., 2012]. As shown in Section 6, this scheme has several
over the others.

Generalized Multiple Importance Sampling

Víctor Elvira[1*], Luca Martino[2], David Luengo[3], and Mónica F. Bugallo[4]

[1]University of Edinburgh (United Kingdom), [2]Universidad Rey Juan Carlos (Spain), [3]Universidad Politécnica de Madrid (Spain), [4]Stony Brook University (USA)

Choosing N_i to minimize variance

$$\langle F \rangle = \frac{1}{\boxed{N_1}} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right) + \frac{1}{\boxed{N_2}} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

can we pick optimal N_i?

Choosing N_i to minimize variance

$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f(X_{1,i})}{p_1(X_{1,i})} w_1(X_{1,i}) + \frac{1}{N_2} \sum_j \frac{f(X_{2,j})}{p_2(X_{2,j})} w_2(X_{2,j})$$

need to jointly optimize w_i and N_i
(given a total budget N)

Choosing N_i to minimize variance

$$\langle F \rangle = \frac{1}{\boxed{N_1}} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} \boxed{w_1\left(X_{1,i}\right)} + \frac{1}{\boxed{N_2}} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} \boxed{w_2\left(X_{2,j}\right)}$$

need to jointly optimize w_i and N_i
(given a total budget N)

Many people have shown that this can be formulated as a convex optimization problem!

On Learning the Best Local Balancing Strategy

D. Murray[1] and S. Benzait[1] and R. Pacanowski[2] and X. Granier[2,3]

Optimal Deterministic Mixture Sampling

Mateu Sbert[1] & Vlastimil Havran[2] & László Szirmay-Kalos[3]

Optimal mixture weights in multiple importance sampling

Hera Y. He, Stanford University
Art B. Owen, Stanford University

Choosing N_i to minimize **efficiency**

$$\langle F \rangle_{N_1,N_2} = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right) + \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

minimize

$$C(N_1, N_2)\mathrm{Var}\left[\langle F \rangle_{N_1,N_2}\right]$$

C = how long it takes to render

Efficiency-aware multiple importance sampling for bidirectional rendering algorithms

PASCAL GRITTMANN, Saarland University, Germany
ÖMERCAN YAZICI, Saarland University, Germany
ILIYAN GEORGIEV, Autodesk, United Kingdom
PHILIPP SLUSALLEK, Saarland University, Germany and DFKI, Germany

Grittmann et al.'s strategy: just try out different combinations of Ns!

$$\langle F \rangle_{N_1,N_2} = \frac{1}{N_1}\sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right) + \frac{1}{N_2}\sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

minimize

$$C(N_1, N_2)\mathrm{Var}\left[\langle F \rangle_{N_1,N_2}\right]$$

C = how long it takes to render

estimate

$$C(10\,\%\,,90\%)\mathrm{Var}\left[\langle F \rangle_{10\%,90\%}\right]$$

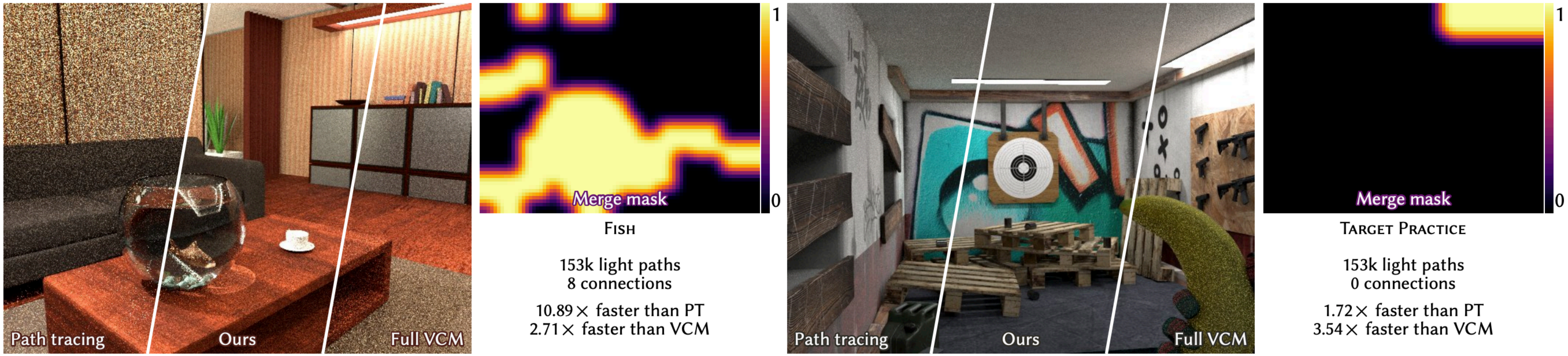
$$C(30\,\%\,,70\%)\mathrm{Var}\left[\langle F \rangle_{30\%,70\%}\right]$$

…

Efficiency-aware multiple importance sampling for bidirectional rendering algorithms

PASCAL GRITTMANN, Saarland University, Germany
ÖMERCAN YAZICI, Saarland University, Germany
ILIYAN GEORGIEV, Autodesk, United Kingdom
PHILIPP SLUSALLEK, Saarland University, Germany and DFKI, Germany

Choosing number of samples for MIS is crucial for bidirectional methods



Can we do better than Veach?

$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right)$$

$$+ \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

trick 1: assuming $X_{1,i}$ and $X_{2,j}$ are uncorrelated

$$\mathrm{Var}\left[\langle F \rangle\right] = \mathrm{Var}\left[\frac{1}{N_1} \sum \frac{f w_1}{p_1}\right] + \mathrm{Var}\left[\frac{1}{N_2} \sum \frac{f w_2}{p_2}\right]$$

trick 2: minimize upper bound of the variance

$$\mathrm{Var}\left[X\right] = E\left[X^2\right] - E\left[X\right]^2 \leq E\left[X^2\right]$$

Can we do better than Veach?

$$\langle F \rangle = \frac{1}{N_1} \sum_i \frac{f\left(X_{1,i}\right)}{p_1\left(X_{1,i}\right)} w_1\left(X_{1,i}\right) + \frac{1}{N_2} \sum_j \frac{f\left(X_{2,j}\right)}{p_2\left(X_{2,j}\right)} w_2\left(X_{2,j}\right)$$

what if the samples are correlated?

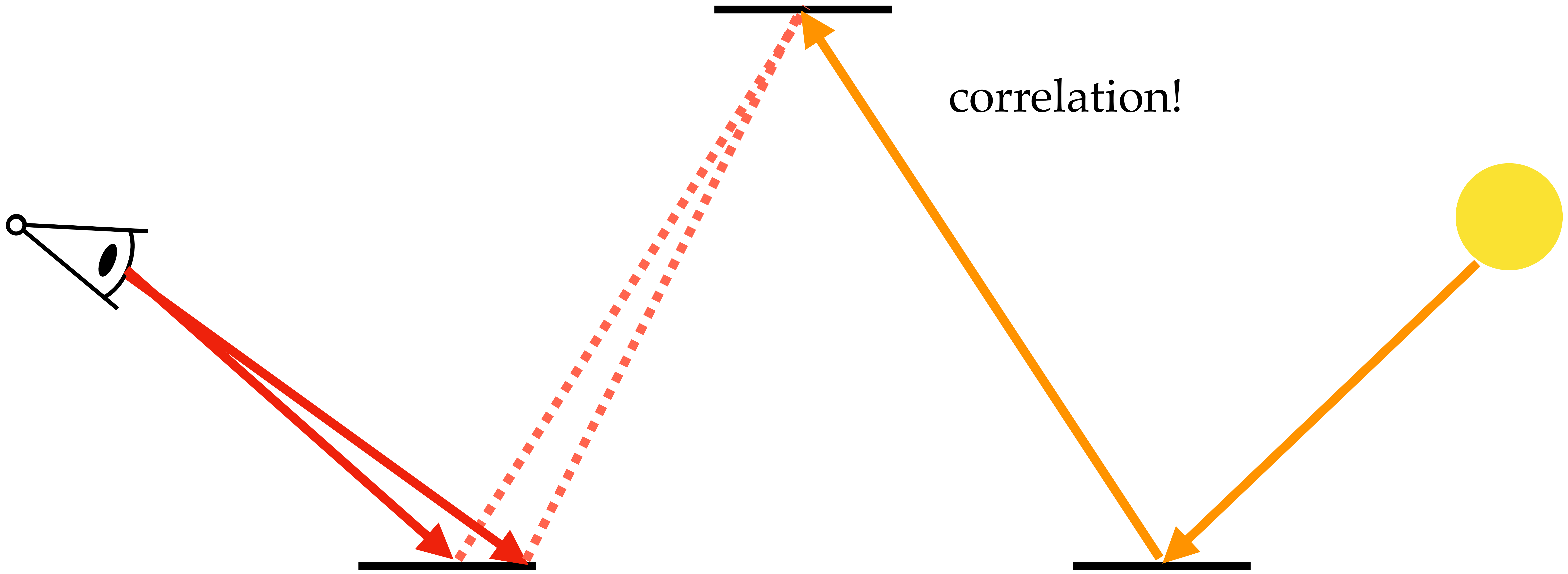
trick 1: assuming $X_{1,i}$ and $X_{2,j}$ are uncorrelated

$$\mathrm{Var}\left[\langle F \rangle\right] = \mathrm{Var}\left[\frac{1}{N_1} \sum \frac{f w_1}{p_1}\right] + \mathrm{Var}\left[\frac{1}{N_2} \sum \frac{f w_2}{p_2}\right]$$

trick 2: minimize upper bound of the variance

$$\mathrm{Var}\left[X\right] = E\left[X^2\right] - E\left[X\right]^2 \leq E\left[X^2\right]$$

Correlation occurs in bidirectional path tracing when camera subpaths are shared by a light subpath



Correlation-aware MIS

- no satisfactory solution yet, only heuristics exist

Probabilistic Connections for Bidirectional Path Tracing

Stefan Popov [1] Ravi Ramamoorthi [2] Fredo Durand [3] George Drettakis [1]

[1] Inria [2] UC San Diego [3] MIT CSAIL

minimize a very loose upper bound

$$V[\tilde{I}_l] \leq \sum_{i \in S_u} \frac{1}{n_i} V[F_{i1}] + \sum_{i \in S_c} V[F_{i1}]$$

Correlation-Aware Multiple Importance Sampling for Bidirectional Rendering Algorithms

Pascal Grittmann[1] Iliyan Georgiev[2] Philipp Slusallek[1,3]

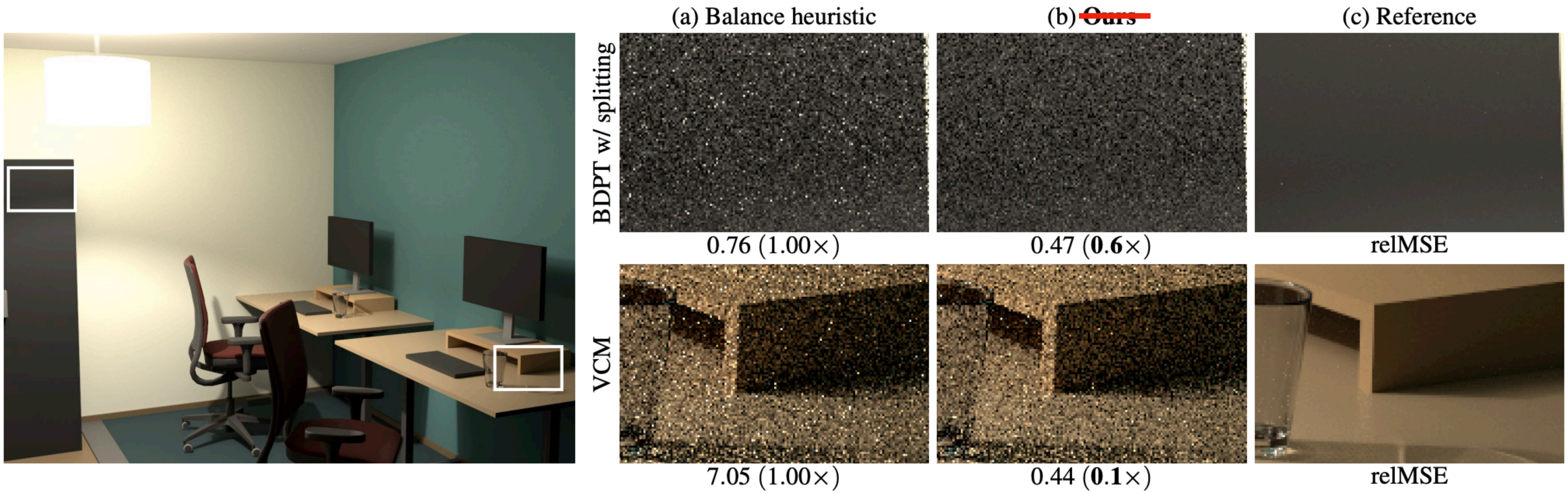
[1]Saarland University, Germany [2]Autodesk, United Kingdom [3]DFKI, Germany

add a "correlation factor" c

$$w_t(\overline{\mathbf{x}}) = \frac{c_t(\overline{\mathbf{x}}) n_t p_t(\overline{\mathbf{x}})}{\sum_k c_k(\overline{\mathbf{x}}) n_k p_k(\overline{\mathbf{x}})}$$

Correlation-aware MIS

Grittmann et al.'s



Correlation-Aware Multiple Importance Sampling for Bidirectional Rendering Algorithms

Pascal Grittmann[1] Iliyan Georgiev[2] Philipp Slusallek[1,3]

[1]Saarland University, Germany [2]Autodesk, United Kingdom [3]DFKI, Germany

Continuous MIS

- instead of a finite amount of distributions, we can consider uncountably many distributions

$$\langle F \rangle_{\text{MIS}} = \sum \frac{f}{p_i} w_i$$

$$\langle F \rangle_{\text{CMIS}} = \int \frac{f}{p(i)} w(i)$$

Continuous Multiple Importance Sampling

REX WEST, The University of Tokyo, Japan
ILIYAN GEORGIEV, Autodesk, United Kingdom
ADRIEN GRUSON, McGill University, Canada
TOSHIYA HACHISUKA, The University of Tokyo, Japan

MIS is frequently used in Bayesian inference

Adaptive Multiple Importance Sampling

JEAN-MARIE CORNUET
Centre de Biologie et Gestion des Populations
INRA, Montpellier

JEAN-MICHEL MARIN
Institut de Mathématiques et Modélisation de Montpellier,
(UMR CNRS 5149), Université Montpellier 2

ANTONIETTA MIRA
Department of Economics, University of Lugano, Switzerland

CHRISTIAN P. ROBERT
Université Paris Dauphine, CEREMADE,
IUF, and CREST, Paris

Implicitly adaptive importance sampling

Topi Paananen[1] · Juho Piironen[1] · Paul-Christian Bürkner[1] · Aki Vehtari[1]

Generalized Multiple Importance Sampling

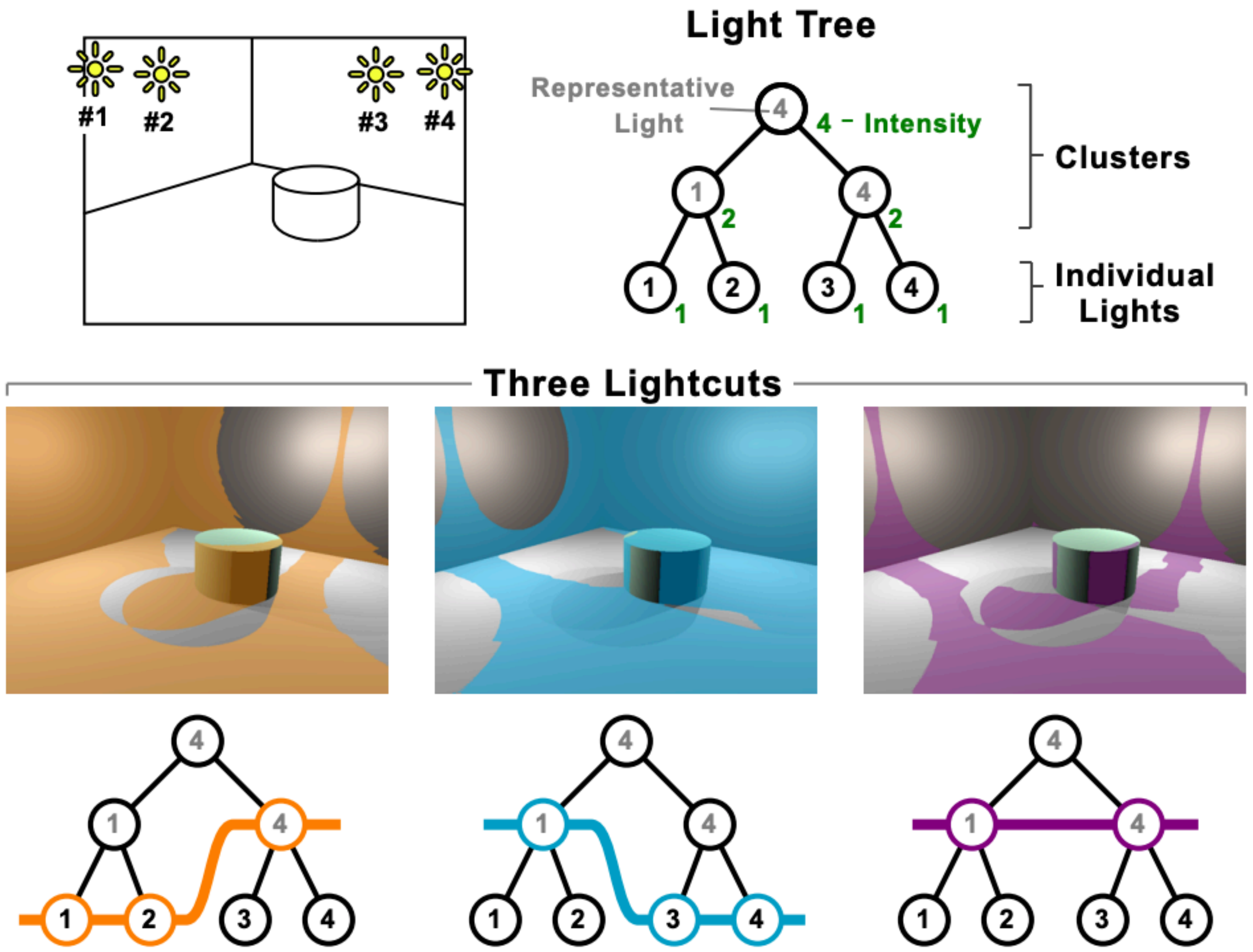
Víctor Elvira[1*], Luca Martino[2], David Luengo[3], and Mónica F. Bugallo[4]

[1]University of Edinburgh (United Kingdom), [2]Universidad Rey Juan Carlos (Spain), [3]Universidad Politécnica de Madrid (Spain), [4]Stony Brook University (USA)

A layered multiple importance sampling scheme for focused optimal Bayesian experimental design*

Chi Feng* and Youssef M. Marzouk†

Next: many-lights sampling



Lightcuts: A Scalable Approach to Illumination

Bruce Walter   Sebastian Fernandez   Adam Arbree   Kavita Bala   Michael Donikian   Donald P. Greenberg

*Program of Computer Graphics, Cornell University**