CT5202: Photorealistic Rendering

Variance Reduction Techniques

Lecturer: Bochang Moon

Multiple Importance Sampling

- Proposed by Eric Veach
- Assume we have multiple (more than one) sampling techniques
- Q. How do we combine the techniques?

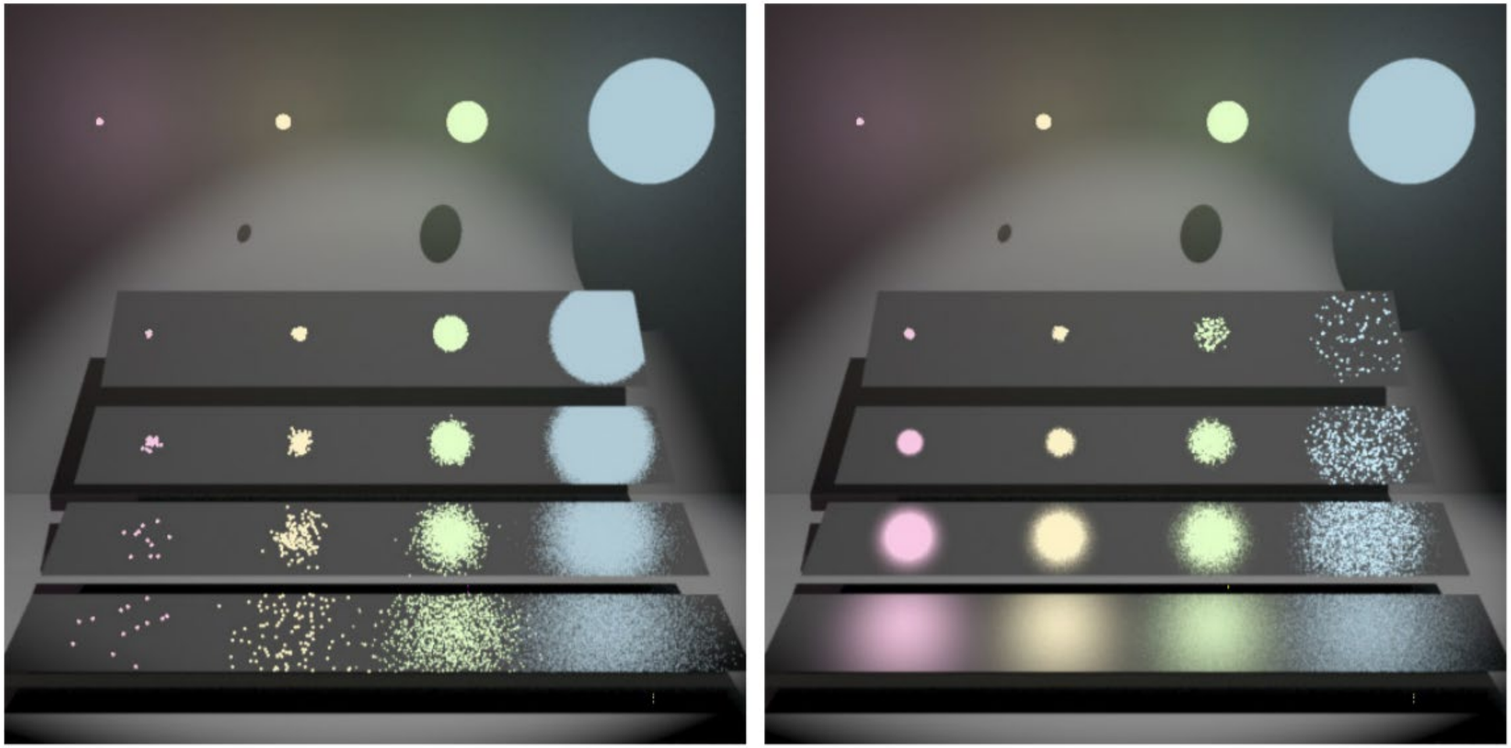
- Motivation:
  - Light transport integral is complex (most terms are unknown and should be estimated)
  - Designing a sampling technique, which works well for a variety of situations, is difficult

Applications

- Glossy highlights from area light sources
- Common sampling techniques
  - Sampling the light sources
  - Sampling the BRDF

- Light transport for direct lighting
  - $L_s(x, k_o) = \int_{all\ x'} \frac{\rho(k_i, k_o) L_e(x', -k_i) v(x, x') \cos\theta_i \cos\theta'}{\left|\left|x - x'\right|\right|^2} dA'$

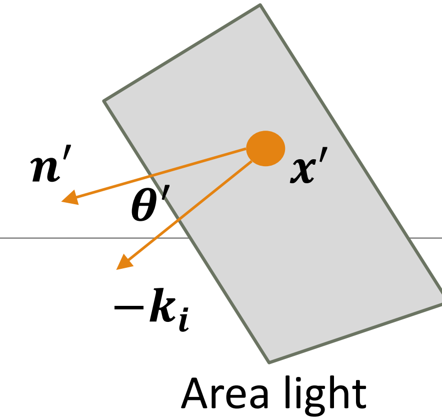
Applications



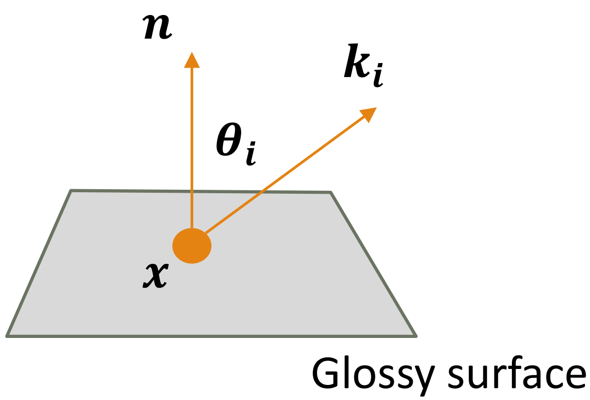
(a) Sampling the BSDF **(b)** Sampling the light sources

Images from the Veach's thesis

Applications



- Sampling the light sources
  - $p(x') \propto \frac{L_e(x', -k_i)\cos\theta_i \cos\theta'}{||x-x'||^2}$
- Sampling the BRDF
  - $p(k_i) \propto \rho(k_i, k_o)$



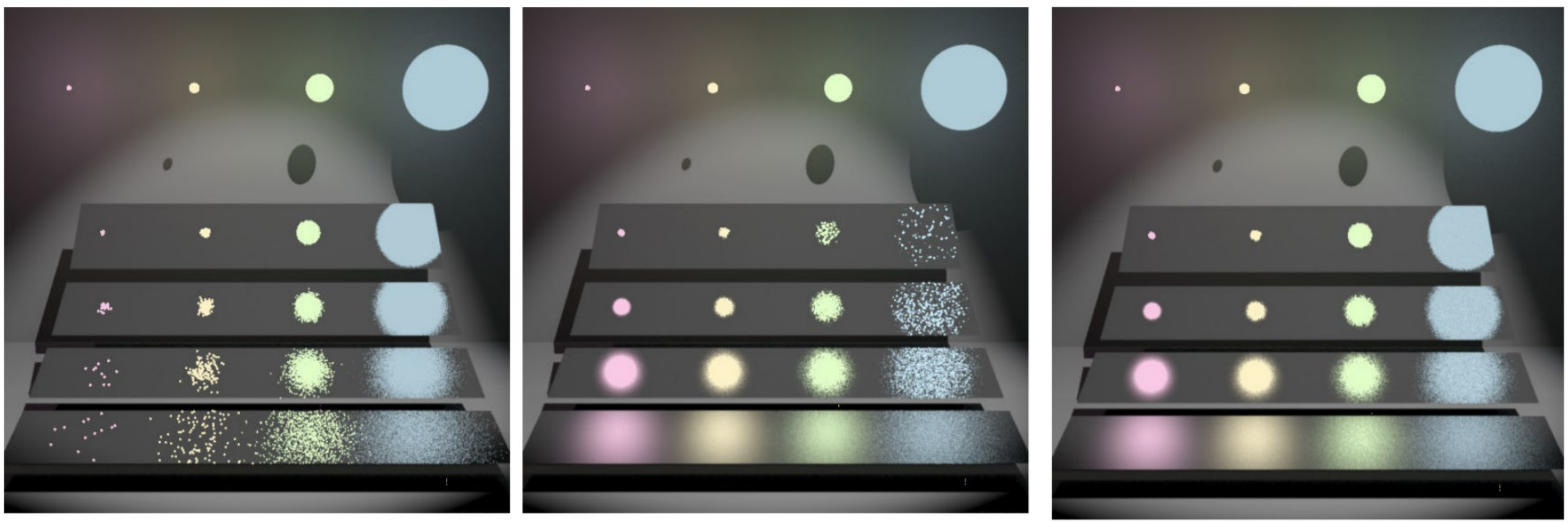
- Light transport for direct lighting
  - $L_s(x, k_o) = \int_{all\ x'} \frac{\rho(k_i, k_o) L_e(x', -k_i) v(x, x') \cos\theta_i \cos\theta'}{||x-x'||^2} dA'$

Multi-Sample Estimator

- A combination strategy to average samples from multiple sampling techniques
- $\int_\Omega f(x) d\mu(x)$
  - $f : \Omega \rightarrow R$
- Samples ($j = 1, \dots, n_i$) from i-th sampling
  - $X_{i,j}$
- Multi-sample estimator
  - $F = \sum_{i=1}^{n} 1/n_i \sum_{j=1}^{n_i} w_i(X_{i,j}) \frac{f(X_{i,j})}{p_i(X_{i,j})}$
  - Conditions for unbiasedness
    - $\sum_{i=1}^{n} w_i(x) = 1$ whenever $f(x) \neq 0$
    - $w_i(x) = 0$ whenever $p_i(x) = 0$
    - See the Veach's thesis for the proof.

Multi-Sample Estimator

- The balance heuristic
  - $\hat{w}_i(x) = \frac{n_i p_i(x)}{\Sigma_\mathrm{k} n_k p_k(x)}$



(a) Sampling the BSDF

(b) Sampling the light sources

Images from the Veach's thesis

Antithetic Variates

- $\theta = \int_{\Omega} f(x) d\mu(x)$
- Monte Carlo estimator
  - $\hat{\theta} = \frac{1}{\mathrm{N}} \Sigma_{\mathrm{i}} \mathrm{f}(\mathrm{X}_{\mathrm{i}})$
- Suppose:
  - $pdf(X) = pdf(-X)$
  - i.e., a symmetric pdf of X
- Define a variable
  - $Y_i \equiv \frac{f(X_i) + f(-X_i)}{2}$
  - Y_i is an unbiased estimate of θ, $E[Y] = E[f(X)] = \theta$
  - $\hat{\theta}^{av} = \frac{1}{\mathrm{N}} \Sigma_{\mathrm{i}} Y_i$

Antithetic Variates

- (original) Monte Carlo estimator
  - $\hat{\theta} = \frac{1}{\mathrm{N}} \Sigma_{\mathrm{i}} \mathrm{f}(\mathrm{X}_{\mathrm{i}})$
- New estimator
  - $\hat{\theta}^{av} = \frac{1}{\mathrm{N}} \Sigma_{\mathrm{i}} Y_i$
- Suppose I.I.D:
  - $\mathrm{var}\big(\mathrm{f}(\mathrm{X}_{\mathrm{i}})\big) = \mathrm{var}\left(\mathrm{f}\big(\mathrm{X}_{\mathrm{j}}\big)\right) = \sigma^2$
- Variance of estimators
  - $var\big(\hat{\theta}\big) = var\left(\frac{1}{\mathrm{N}} \Sigma_{\mathrm{i}} \mathrm{f}(\mathrm{X}_{\mathrm{i}})\right) = \frac{1}{\mathrm{N}^2} \Sigma_i var\big(f(X_i)\big) = \frac{1}{N^2} \Sigma_i \sigma^2 = \frac{\sigma^2}{N}$
  - $var\big(\hat{\theta}^{av}\big) = var\left(\frac{1}{\mathrm{N}} \Sigma_{\mathrm{i}} Y_i\right) = \frac{1}{\mathrm{N}^2} \Sigma_i var(Y_i)$

Antithetic Variates

- Variance of estimators
  - $var(\hat{\theta}) = var\left(\frac{1}{\mathrm{N}}\Sigma_{\mathrm{i}}\mathrm{f}(\mathrm{X_i})\right) = \frac{1}{\mathrm{N}^2}\Sigma_i var(f(X_i)) = \frac{1}{N^2}\Sigma_i \sigma^2 = \frac{\sigma^2}{N}$
  - $var(\hat{\theta}^{av}) = var\left(\frac{1}{\mathrm{N}}\Sigma_{\mathrm{i}} Y_i\right) = \frac{1}{\mathrm{N}^2}\Sigma_i var(Y_i)$
  - $var(Y_i) = var\left(\frac{f(X_i)+f(-X_i)}{2}\right) = \frac{1}{4}\left[var(f(X_i)) + var(f(-X_i)) + 2cov(f(X_i), f(-X_i))\right]$

  - Putting $var(Y_i)$ into $var(\hat{\theta}^{av})$:
  - $var(\hat{\theta}^{av}) = \frac{1}{4N^2}\Sigma_i\{2\sigma^2 + 2cov(f(X_i), f(-X_i))\} = \frac{1}{2N^2}\Sigma_i\{\sigma^2 + cov(f(X_i), f(-X_i))\}$
  - If there is no correlation
    - $var(\hat{\theta}^{av}) = \frac{\sigma^2}{2N} = \frac{var(\hat{\theta})}{2}$
    - No actual gain here since we use 2N samples instead of N samples
  - What if there is a negative correlation?
    - $var(\hat{\theta}^{av}) < \frac{var(\hat{\theta})}{2}$

Antithetic Variates

- Antithetic variates introduces a negative correlation for monotonically increasing functions
  - $cov\left(f(X_i), f(-X_i)\right) < 0$
  - e.g., linear functions – ideal case

- Properties
  - Very simple to implement it even for high-dimensional cases
  - Some applications in rendering:
    - Direct lighting
    - Pixel estimator in PSS?

Common Random Numbers (CRN)

- Suppose we want to estimate a difference between two functions
  - $\theta_1 = \int_\Omega f_1(x) d\mu(x)$
  - $\theta_2 = \int_\Omega f_2(x) d\mu(x)$
  - $\theta = \theta_1 - \theta_2$

- MC estimator
  - $\hat{\theta} = \frac{1}{N} \Sigma_i f_1(X_i) - \frac{1}{N} \Sigma_j f_2(X_j)$

- CRN estimator

- $\hat{\theta} = \frac{1}{N} \Sigma_i f_1(X_i) - \frac{1}{N} \Sigma_i f_2(X_i) = \frac{1}{N} \Sigma_i \left(f_1(X_i) - f_2(X_i) \right)$

Common Random Numbers (CRN)

- CRN estimator
  - $\hat{\theta} = \frac{1}{\mathrm{N}}\Sigma_{\mathrm{i}}\mathrm{f}_1(\mathrm{X}_{\mathrm{i}}) - \frac{1}{\mathrm{N}}\Sigma_{\mathrm{i}}\mathrm{f}_2(\mathrm{X}_{\mathrm{i}}) = \frac{1}{N}\Sigma_{\mathrm{i}}\big(f_1(X_i) - f_2(X_i)\big)$
  - $var(\hat{\theta}) = var\left(\frac{1}{N}\Sigma_{\mathrm{i}}\big(f_1(X_i) - f_2(X_i)\big)\right) = 1/N^2\Sigma_i var\big(f_1(X_i) - f_2(X_i)\big)$
  - $var\big(f_1(X_i) - f_2(X_i)\big) = var\big(f_1(X_i)\big) + var\big(f_2(X_i)\big) - 2cov\big(f_1(X_i), f_2(X_i)\big)$
- What if $cov\big(f_1(X_i), f_2(X_i)\big) = 0$?
  - No actual gain over the ordinary MC estimator.

Common Random Numbers (CRN)

- When the two functions tend to increase (or decrease) together,
  - $cov\left(f_1(X_i), f_2(X_i)\right) > 0$
  - e.g., both functions are linear whose derivatives have the same sign.

- Applications in rendering
  - Estimating image gradients
  - (screened) Poisson reconstruction takes the image gradients to output a reconstructed image

- Q. can we decide whether or not we apply the CRN?
  - In practice, it is hard to know if there is such correlation in advance.
  - However, implementing and testing CRN are very easy.

CRN examples

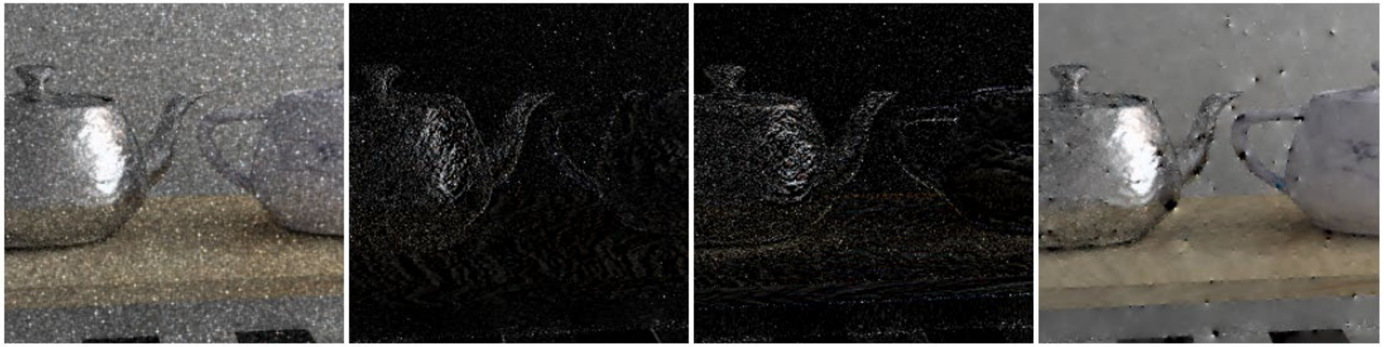
Path tracing with CRN numbers, 76 samples per pixel

Applications of Correlated Sampling: Gradient-Domain Rendering

- Image gradients can be estimated via correlated sampling
  - $\mathrm{I}(\mathrm{x}+1,\mathrm{y})-\mathrm{I}(\mathrm{x},\mathrm{y})$
  - $\mathrm{I}(\mathrm{x},\mathrm{y}+1)\ -\ \mathrm{I}(\mathrm{x},\mathrm{y})$
- The variance of the estimated gradients can be smaller than the pixel color when then the covariance term is positive
  - $var(\mathrm{I}(\mathrm{x}+1,\mathrm{y})-\mathrm{I}(\mathrm{x},\mathrm{y}))=var\big(\mathrm{I}(\mathrm{x}+1,\mathrm{y})\big)+var(\mathrm{I}(\mathrm{x},\mathrm{y}))-2cov(\mathrm{I}(\mathrm{x}+1,\mathrm{y}),\mathrm{I}(\mathrm{x},\mathrm{y}))$

Applications of Correlated Sampling: Gradient-Domain Rendering

- Rendering estimates three images:
  - Primal colors (e.g., standard path tracing)
  - Image gradients (e.g., correlated samplings such as CRNs, shift mapping, and path reusing)



(a) Input color image, 512 spp relMSE 0.1318 (b) Input gradients in horizontal direction (c) Input gradients in vertical direction (d) L2 relMSE 0.0184

Images from [Ha et al. 2019]

Applications of Correlated Sampling: Gradient-Domain Rendering

- Screened Poisson Reconstruction
  - $\hat{y} = \underset{\bar{y}}{\operatorname{argmin}} \sum_{i=1}^{n} \|\alpha(y_i - \bar{y}_i)\|^2 + \sum_{i=1}^{n} \left\|g_i^{dx} - D^{dx}\bar{y}_i\right\|^2 + \sum_{i=1}^{n} \left\|g_i^{dy} - D^{dy}\bar{y}_i\right\|^2$
- g_i^{dx}, g_i^{dy}: Estimated gradients at pixel i in x and y directions
- y_i: Pixel color at i
- α: user-parameter (e.g., 0.2)
- D^{dx}, D^{dy}: differential operator in x and y directions (i.e., finite differences)
- Has a closed-form solution (i.e., normal equation) when the norm is $L2$

Applications of Correlated Sampling: Gradient-Domain Rendering

- When L2 reconstruction is used, the output is unbiased.
- One may use a neural network that takes the three inputs
  - e.g., Deep Convolutional Reconstruction for Gradient-Domain Rendering, Kettunen et al. 2019
- More information:
  - EG STAR paper 2019: A Survey on Gradient-Domain Rendering, Hua et al. 2019

Control Variates

- $\theta = \int_{\Omega} f(x) d\mu(x)$
- Monte Carlo estimator
  - $\hat{\theta} = \frac{1}{\mathrm{N}} \Sigma_{\mathrm{i}} \mathrm{f}(\mathrm{X}_{\mathrm{i}})$
- Define a control variate $g(x)$ whose integration G is known.
  - $\theta = \int_{\Omega} f(x) - \alpha g(x) d\mu(x) + \alpha G$
  - $\hat{\theta}^{cv} = \frac{1}{\mathrm{N}} \Sigma_{\mathrm{i}} \left(\mathrm{f}(\mathrm{X}_{\mathrm{i}}) - \alpha \mathrm{g}(\mathrm{X}_{\mathrm{i}})\right) + \alpha G$

Control Variates

- When $\alpha = 1$
  - $\hat{\theta}^{cv} = \frac{1}{\mathrm{N}}\Sigma_{\mathrm{i}}\left(\mathrm{f}(\mathrm{X_i}) - \mathrm{g}(\mathrm{X_i})\right) + G$
  - $var(\hat{\theta}^{cv}) = \frac{1}{\mathrm{N}^2}\Sigma_{\mathrm{i}} var\left(\mathrm{f}(\mathrm{X_i}) - \mathrm{g}(\mathrm{X_i})\right) = 1/N^2\{\Sigma_{\mathrm{i}} var\left(f(X_i)\right) + var\left(g(X_i)\right) - 2cov\left(f(X_i), g(X_i)\right)\}$
  - Assume that:
    - $var\left(f(X_i)\right) = var\left(g(X_i)\right) = \sigma^2$
    - $cov\left(f(X_i), g(X_i)\right) = \sigma_{f,g}^2$
    - $corr\left(f(X_i), g(X_i)\right) = corr_{f,g} = \frac{\sigma_{f,g}^2}{\sigma_f \sigma_g} = \sigma_{f,g}^2/\sigma^2$
  - $var(\hat{\theta}^{cv}) = \frac{\Sigma_{\mathrm{i}} 2\sigma^2}{N^2} - \frac{2\Sigma_i \sigma_{f,g}^2}{N^2} = \frac{2\sigma^2}{N} - \frac{2\sigma_{f,g}^2}{N}$
  - Variance of the original estimator with the same sample count N
    - $var(\hat{\theta}) = \frac{\sigma^2}{2N}$
  - Condition for $var(\hat{\theta}^{cv}) < var(\hat{\theta})$
    - $\frac{2\sigma^2}{N} - \frac{2\sigma_{f,g}^2}{N} < \frac{\sigma^2}{2N}$
    - $\frac{3\sigma^2}{4} < \sigma_{f,g}^2$
    - $\frac{3}{4} < corr_{f,g}$

Control Variates

- When $\alpha = \frac{\sigma_{f,g}^2}{\sigma_g^2} = \frac{\sigma_{f,g}^2}{\sigma^2}$
  - $\hat{\theta}^{cv} = \frac{1}{\mathrm{N}}\Sigma_{\mathrm{i}}\left(\mathrm{f}(\mathrm{X_i}) - \alpha\mathrm{g}(\mathrm{X_i})\right) + \alpha G$
  - $var(\hat{\theta}^{cv}) = \frac{1}{\mathrm{N}^2}\Sigma_{\mathrm{i}} var\left(\mathrm{f}(\mathrm{X_i}) - \alpha\mathrm{g}(\mathrm{X_i})\right) = 1/N^2\{\Sigma_{\mathrm{i}} var\left(f(X_i)\right) + \alpha^2 var\left(g(X_i)\right) - 2\alpha cov\left(f(X_i), g(X_i)\right)\}$
  - $= \frac{\sigma^2}{N} + \frac{\alpha^2\sigma^2}{N} - \frac{2\alpha\sigma_{f,g}^2}{N}$
  - $= \frac{\sigma^2}{N} + \frac{corr_{f,g}^2\sigma^2}{N} - \frac{2corr_{f,g}^2\sigma^2}{N}$
  - $= \frac{\sigma^2}{N}\left(1 - corr_{f,g}^2\right)$

  - Condition for $var(\hat{\theta}^{cv}) < var\left(\hat{\theta}\right)$
    - $\frac{\sigma^2}{N}\left(1 - corr_{f,g}^2\right) < \frac{\sigma^2}{2N}$
    - $1/2 < corr_{f,g}^2$
    - $\frac{1}{4} < \left|corr_{f,g}\right|$

Resampled Importance Sampling

- $I = \int_D f(x) d\mu(x)$
- $\hat{I} = \frac{1}{N} \Sigma_{i=1}^{N} \frac{f(y_i)}{q(y_i)}$
- For importance sampling,
  - q should be normalized (i.e., a valid pdf)
  - Able to sample y_i easily (efficiently) via inverse CDF or rejection sampling
  - Otherwise, we need a workaround (e.g., resampled importance sampling)

Resampled Importance Sampling

- Procedure
  - Generate M ($M \geq 1$) proposal samples from the source pdf p
    - $X = \{x_1, \ldots, x_M\}$
    - Assumption: we can easily sample from the p, but the p may be not a good approximation of the f
  - Compute a weight $w(x_i)$ for x_i
    - $w(x_i) = \frac{q(x_i)}{p(x_i)}$
  - (resampling) Generate N samples with replacement from X
    - Probability of selecting x_i is proportional to $w(x_i)$
    - $Y = \{y_1, \ldots, y_N\}$
  - Estimation
    - $\hat{I} = \frac{1}{N} \Sigma_{i=1}^{N} \frac{f(y_i)}{q(y_i)} \frac{1}{M} \Sigma_{j=1}^{M} w(x_j)$

- Note:
  - Can use an unnormalized q (also hard to sample from it), but it approximates f well

- Ref.
  - Importance Resampling for Global Illumination [Talbot and Cline 2005]