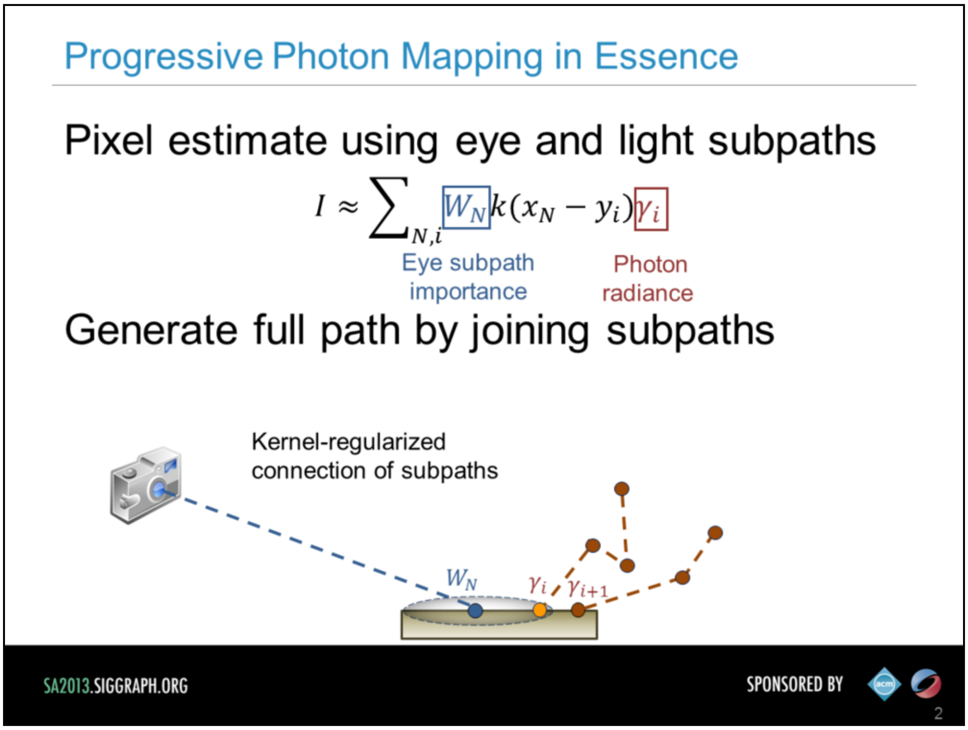
SIGGRAPH ASIA 2013

Adaptive Progressive Photon Mapping

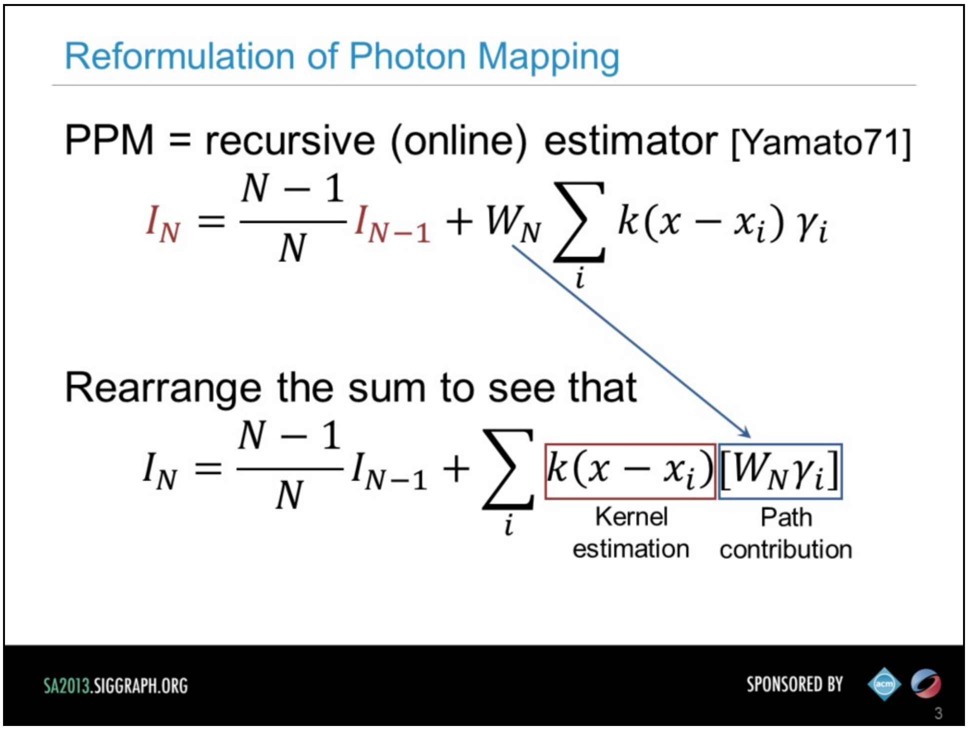
Anton S. Kaplanyan
Karlsruhe Institute of Technology, Germany



Progressive photon mapping can be interpreted as an estimator that constructs the full paths from camera to light by constantly connecting two fresh subpaths.

The estimation is a sum, where W_n is the importance of the eye subpath and gamma_i is the outgoing radiance from a light subpath (photon), as described in Probabilistic PPM [KnausZwicker2011].

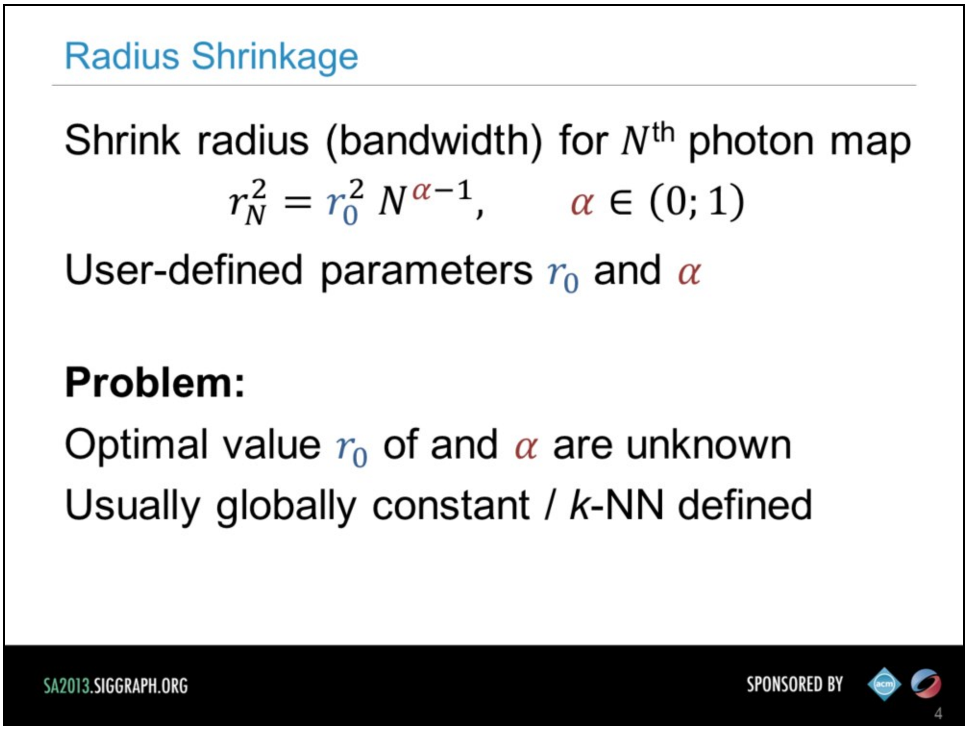
So, the actual construction is a regularized (imprecise) junction of two end vertices.



We write this estimation for every photon map in a recursive manner. This is the first step that allows us to apply the knowledge about recursive (aka online) estimation accumulated in the statistical research.

To do that, we push the W_n term into the sum to reduce the expression to the canonical form of kernel estimation for the new function, which now represents the complete path contribution.

Using this form, we can apply many statistical results from recursive estimation field.



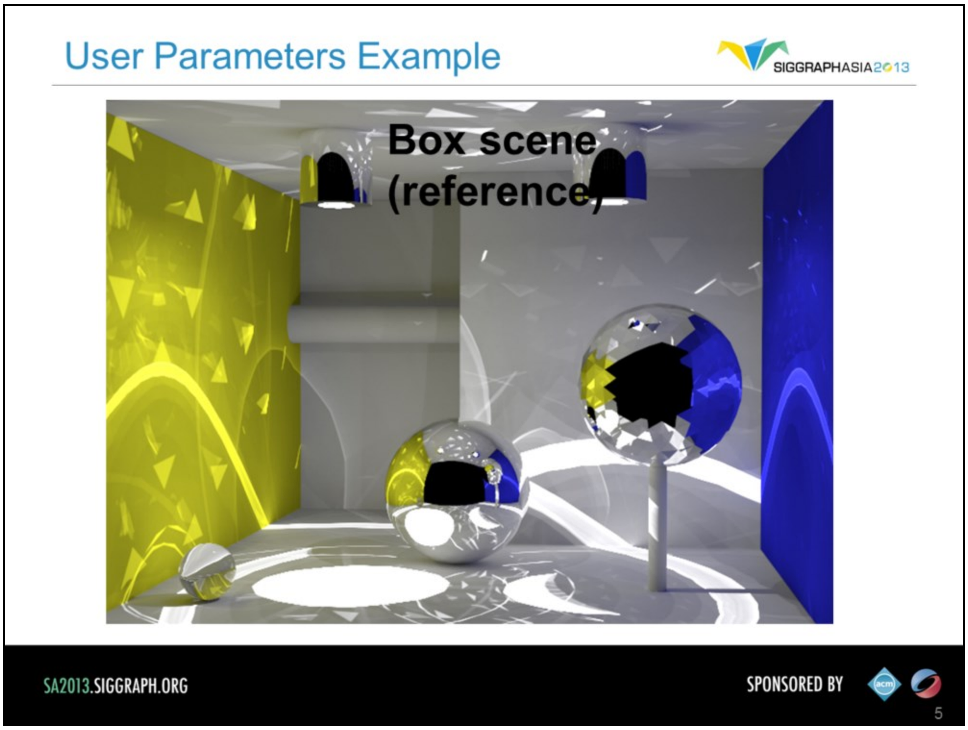
Radius, or kernel bandwidth, is the central parameter for efficiency of the kernel estimation in general.

Knaus and Zwicker shown that the radius must be decreased exponentially.

This equation has two extremely important parameters.

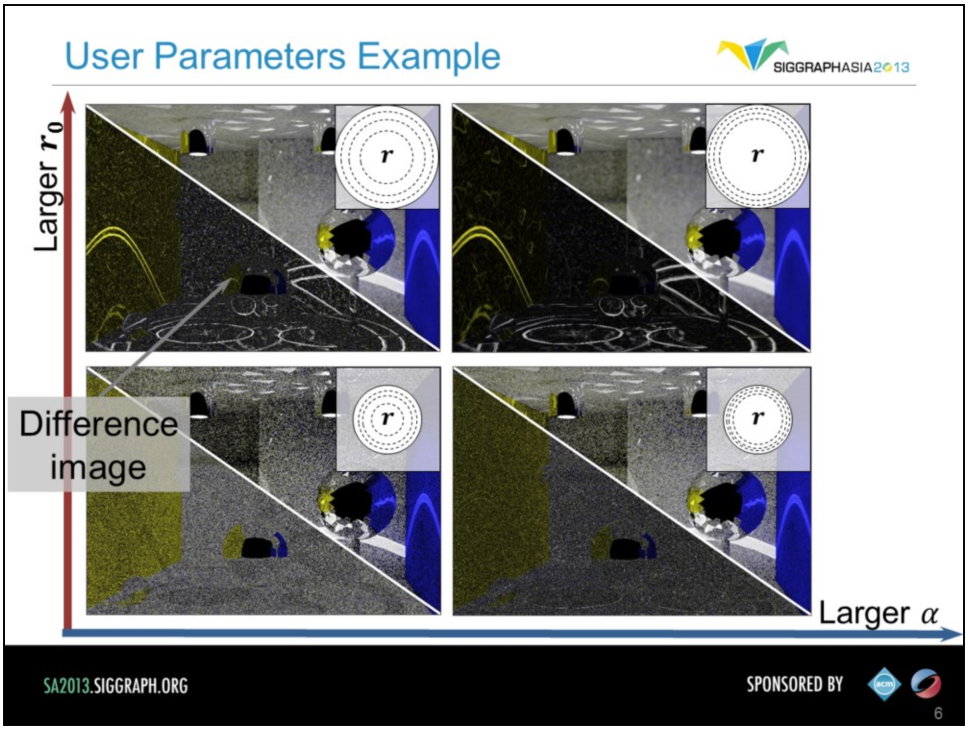
It is an initial radius; and a shrinkage rate alpha.

The alpha parameter is usually defined by user, while the initial radius can be selected using *k*-NN.



This is a Box scene with complicated illumination.

We show how these two parameters affect the rendering.



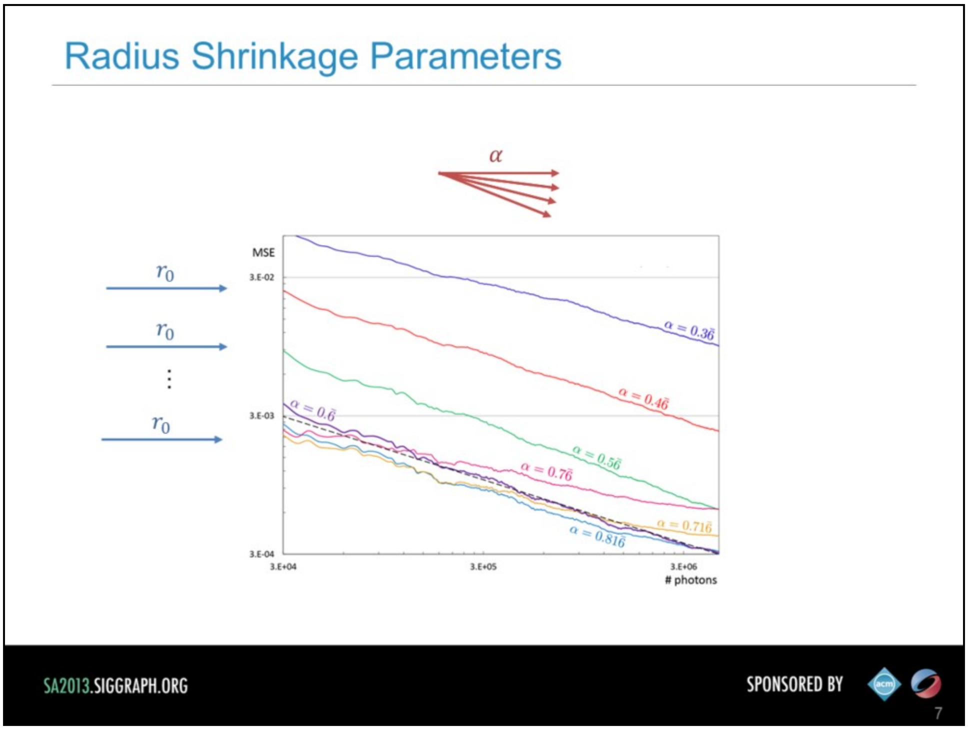
First, we initialize the algorithm with a large initial *k*NN-selected radius and a small alpha of 0.5.

This leads to a lot of bias, like smoothened caustics due to the large initial radius, yet the noise quickly becomes apparent due to a rapid radius shrinkage.

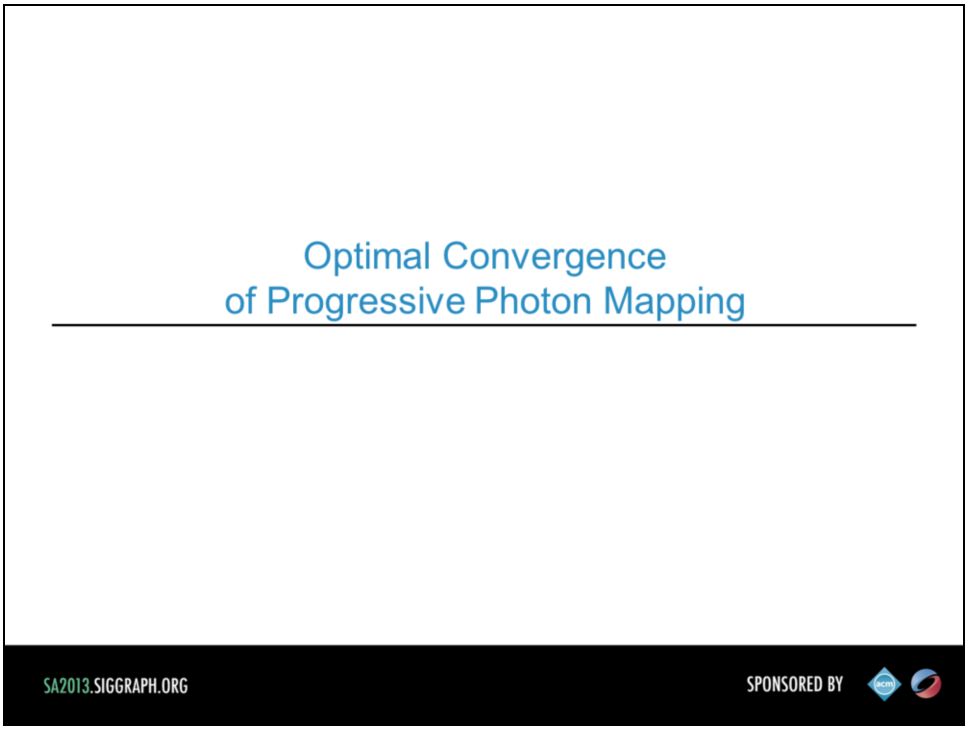
The second case suffers from high image noise due to a small and quickly shrinking initial radius.

The third case demonstrates high bias that appears due to initially large and slowly shrinking radius.

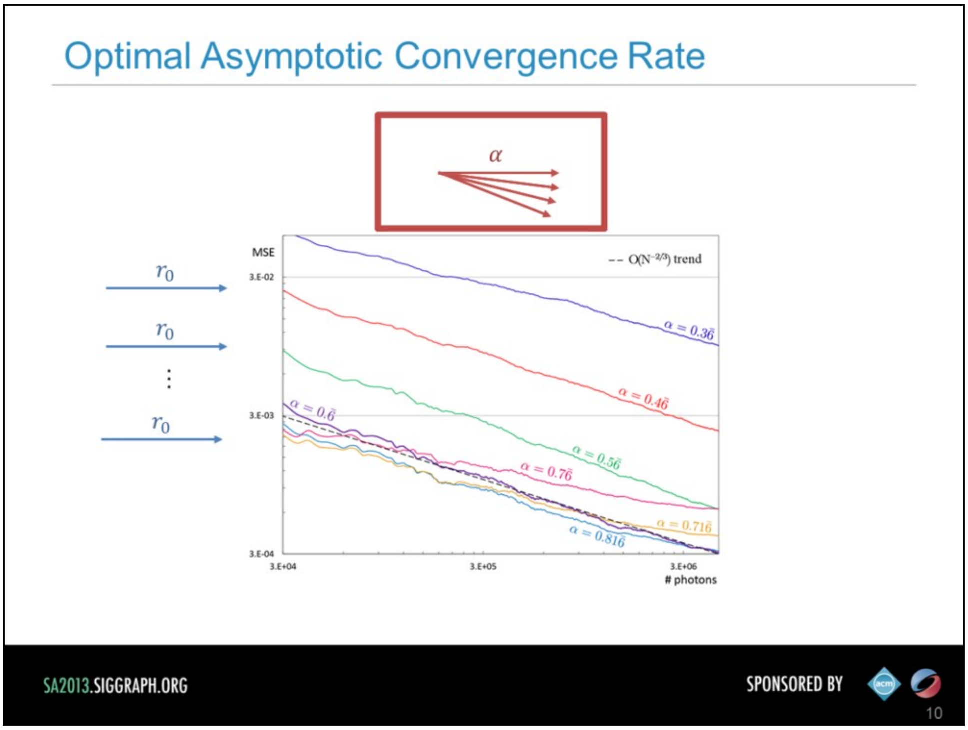
In the fourth, a too-small initial radius can lead to high noise, while bias becomes apparent due to slow shrinkage rate.



If we measure the convergence rate with different parameters alpha and r_0 (vertically – MSE, horizontally – number of photons), it appears that the parameter alpha influences the slope of the plot (in log-log scale) in a non-trivial manner. And parameter r_0 influences the starting error.



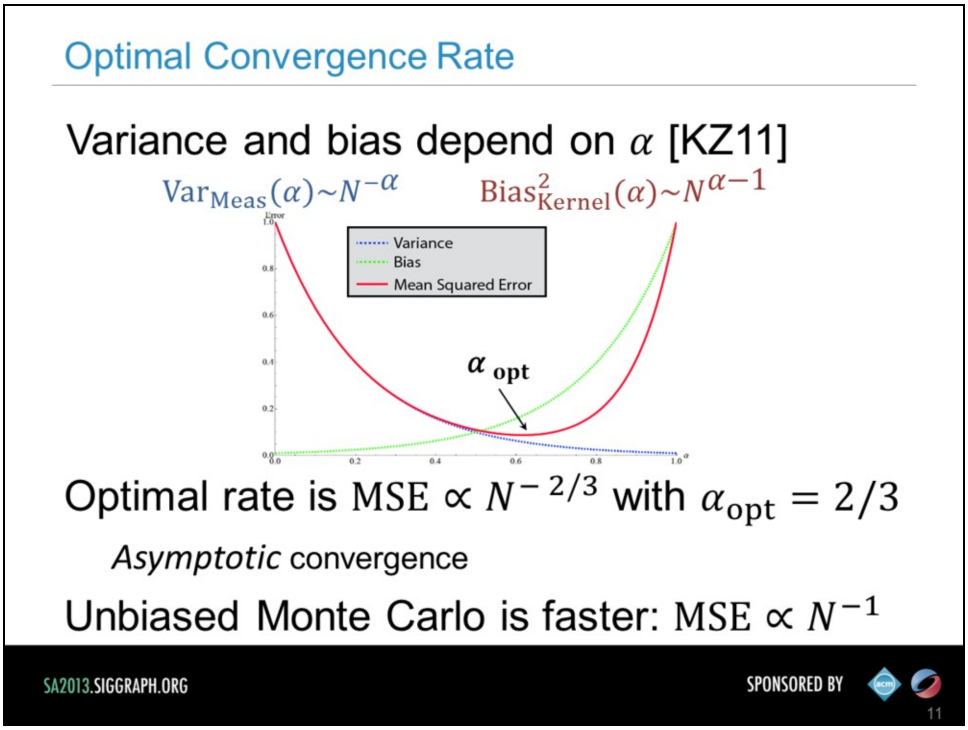
First, we will show how to choose alpha parameter for asymptotically optimal convergence.



Here, we will do an asymptotic analysis, focusing on the optimal alpha parameter.

Note that under the asymptotic assumptions, there is no dependency on the other parameter r_0,

since we analyze the convergence slope, but not the absolute value of the error.

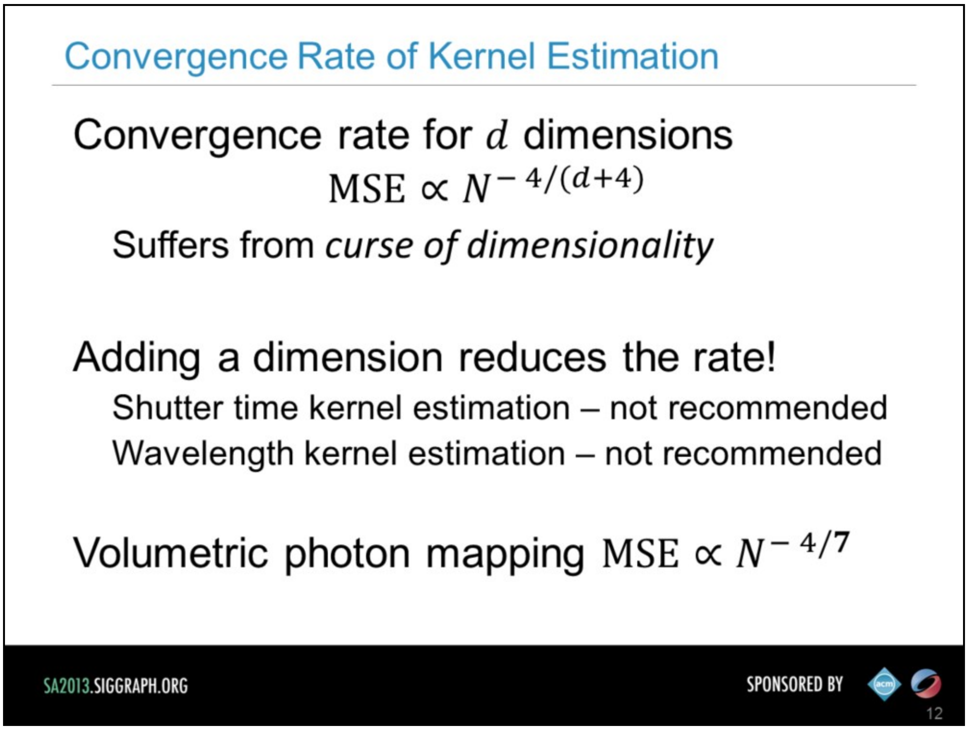


As it was explained by Toshiya Hachisuka in his part of the course, both bias and variance depend on this parameter.

If we put them together into MSE (with some additional analysis described in the Adaptive Progressive Photon Mapping paper), we can deduce the asymptotic behavior of MSE with respect to the parameter alpha. The MSE plot has a minimum.

The optimal asymptotic convergence rate is achieved with alpha = 2/3.

Interestingly, kernel-estimation-based methods are asymptotically slower than unbiased methods.



Moreover, the asymptotic convergence rate of kernel estimation methods depends on the number of dimensions involved in the estimation.

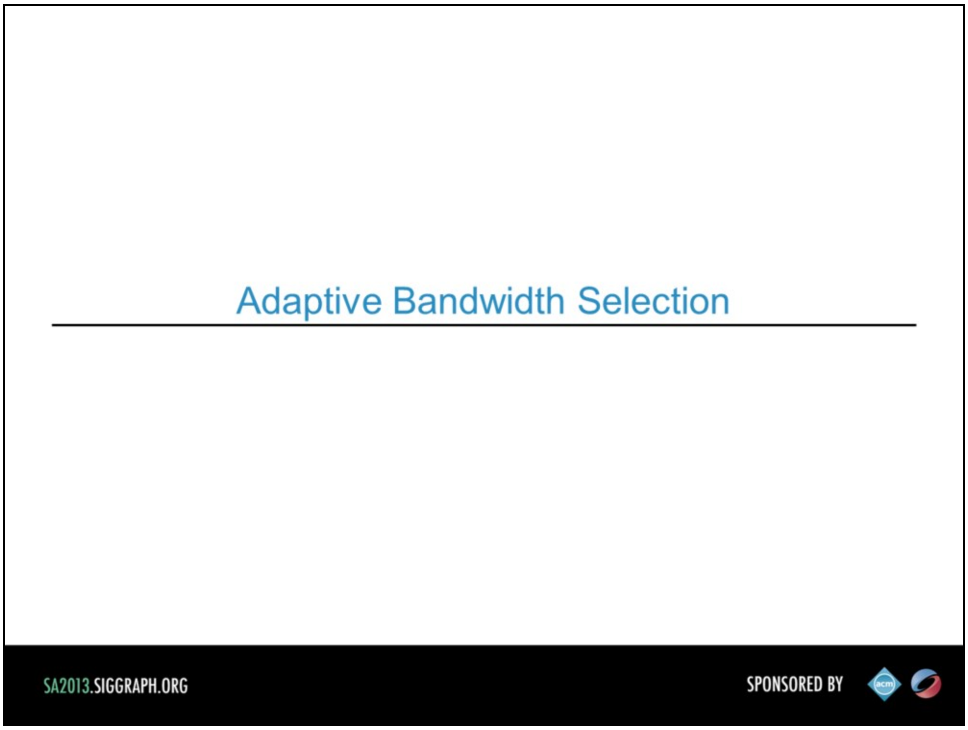
This is a well-known formula in statistics for computing the asymptotic convergence rate of d-dimensional kernel estimation.

Note that unlike, for example, Monte Carlo methods, kernel estimation suffers from curse of dimensionality.

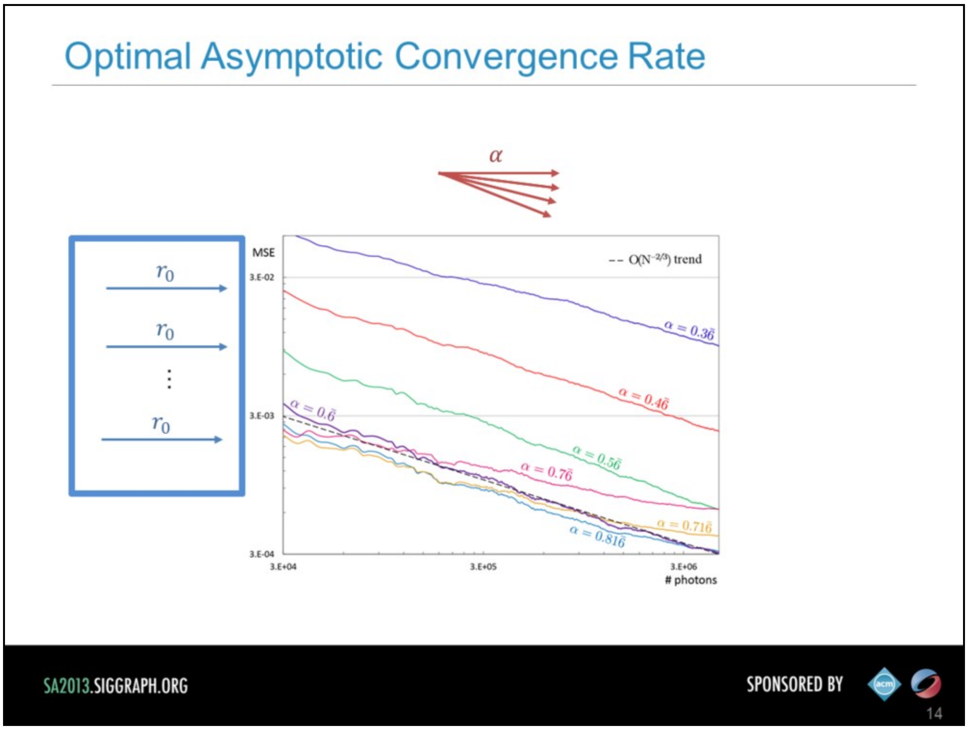
What does that mean for rendering? Every time we add a new dimension, we decrease the convergence rate and quickly face the curse of dimensionality.

For example, if we add kernel estimation in time domain for support of motion blur, then in this case we add one more dimension, and the convergence rate decreases. The same holds if we attempt to do spectral rendering by kernel estimation of wavelength during gathering. Both methods are not practically recommended, since they decrease the asymptotic convergence rate. In order to support motion blur and spectral rendering, one can pre-sample these domains before each pass, as was described by Toshiya Hachisuka in his part of the course.

Another interesting detail is that original volumetric photon mapping suffers from slower convergence due to 3-dimensional kernel estimation (as opposed to 2 dimensions for regular photon mapping). Thus, using photon beams as a primitive is more desirable.

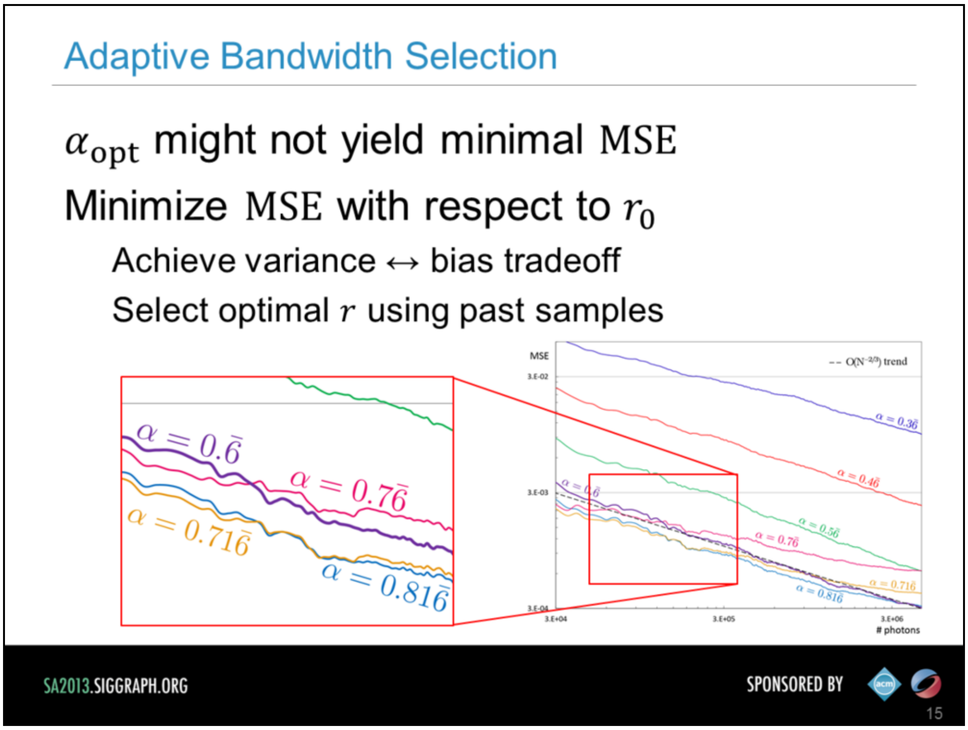


This was an asymptotic analysis. In practice, we have to worry about behavior of MSE on a finite number of photons.
We will show how to select the bandwidth (radius) parameter *r* to achieve faster convergence in finite time.



To minimize the MSE on a finite set of samples, we optimize it w.r.t. the second free parameter, kernel radius (bandwidth) *r*.

In recursive estimation in statistics, this is also a well-known approach called bandwidth selection.

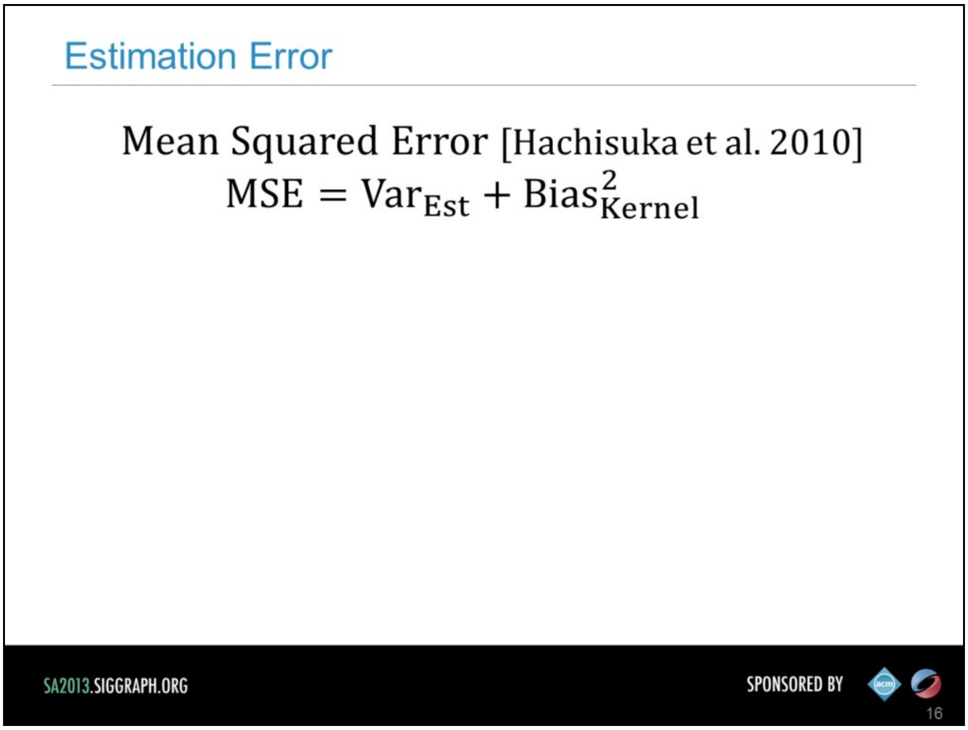


As we can see from this plot, minimizing MSE with respect to the parameter alpha is not enough for practical rendering.

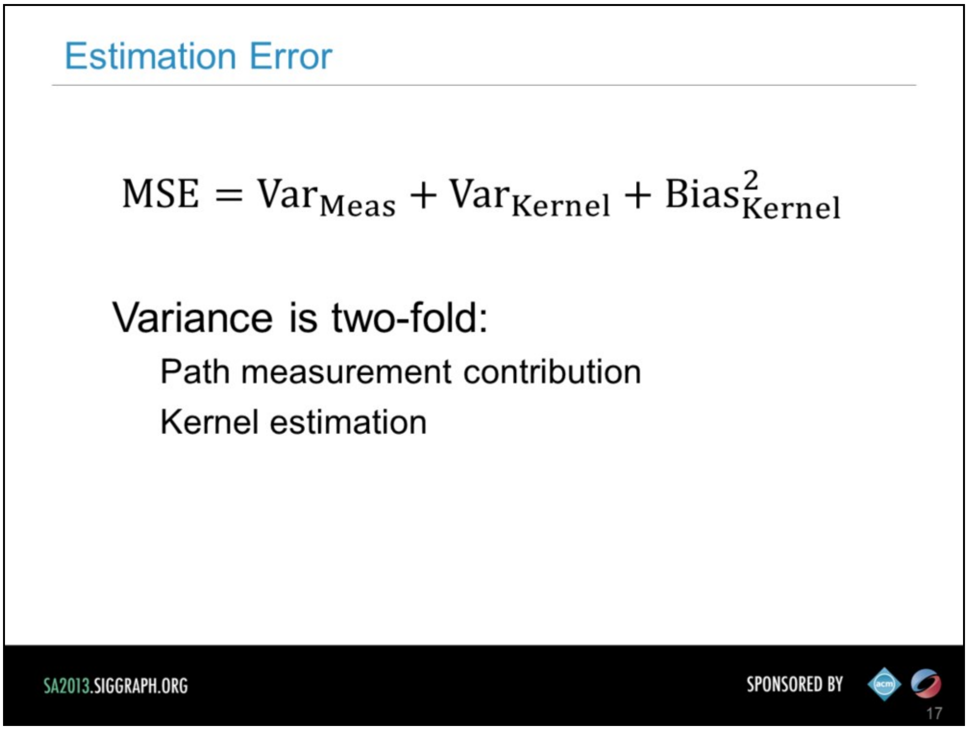
Even the optimal value of alpha, 2/3 (magenta), might lead to high MSE at early stages of rendering, while providing the steepest slope.

We will show that both variance and bias depend on *kernel bandwidth* (radius) r.

It is possible to select the optimal bandwidth in a way to achieve the trade-off between variance and bias, using the past photon maps.

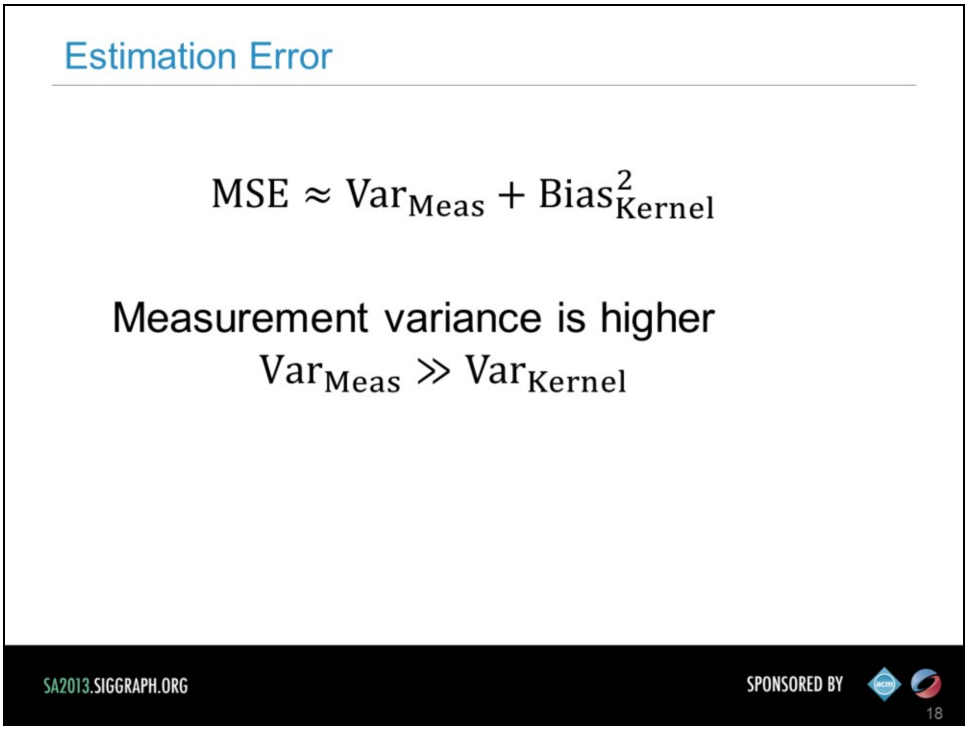


As was mentioned by Toshiya Hachisuka earlier in this course, we know that the mean squared error of such recursive kernel estimator can be decomposed into variance and a bias squared.

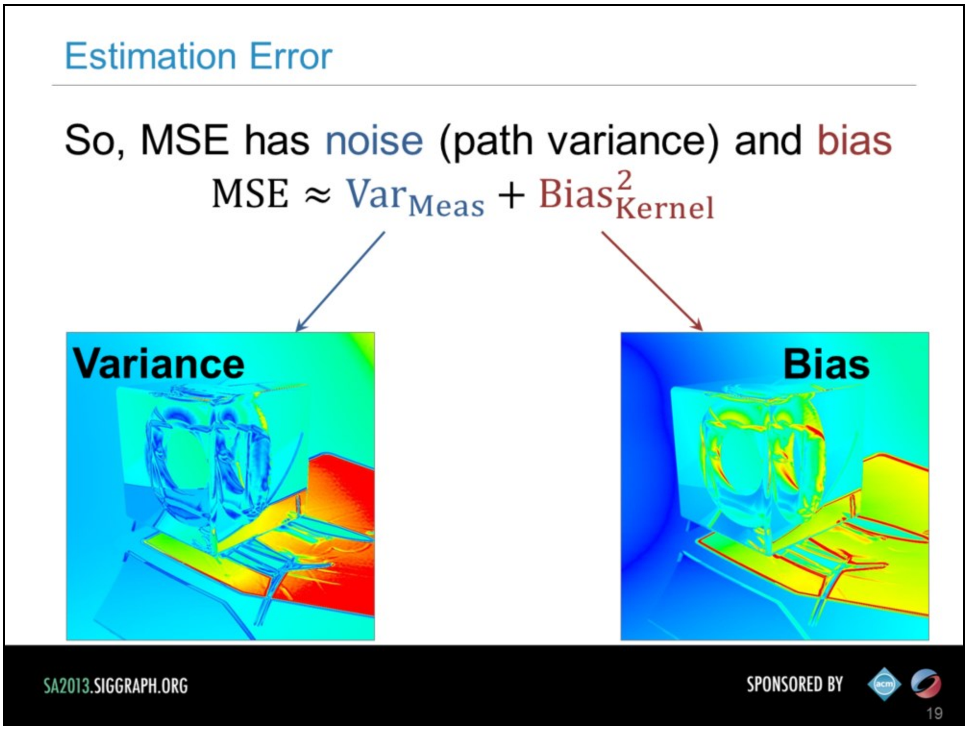


The variance in our case consists of two subterms:

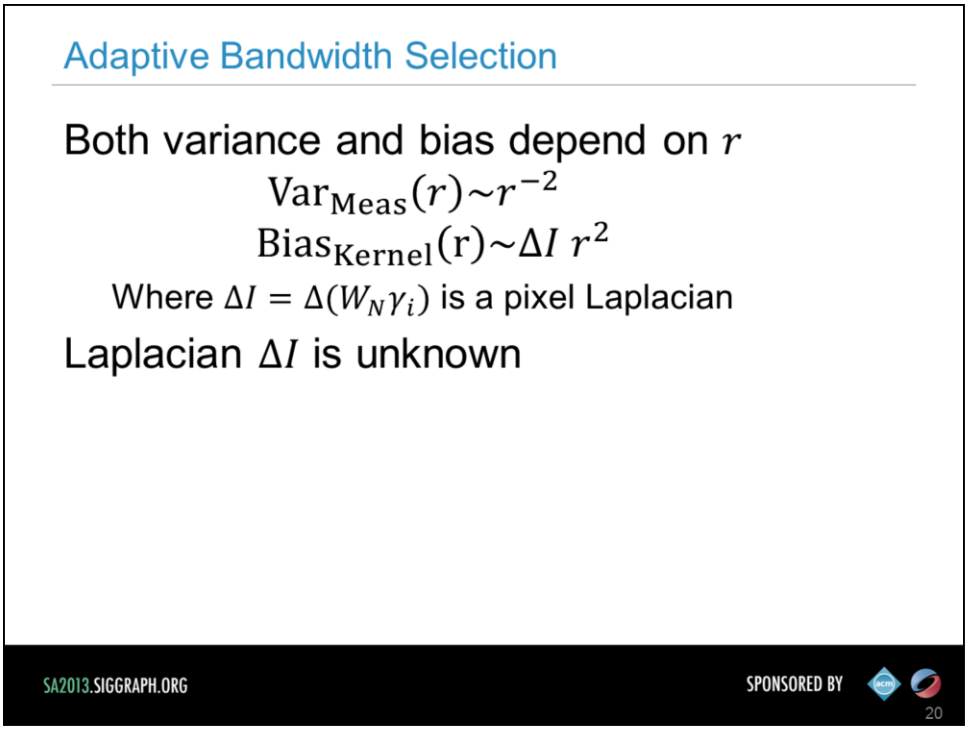
- the variance directly related to the kernel estimation (that is, the noisy distribution of photons);
- and the variance coming from the high-dimensional path integral.



It can be shown that the measurement variance is significantly higher (see the original paper for details).



Thus the major contributors to the error of photon mapping class methods are the variance stemming from the path sampling (image noise); and the bias introduced by kernel estimation (blurred illumination). This error directly depends on the two parameters (alpha and r) we described earlier.



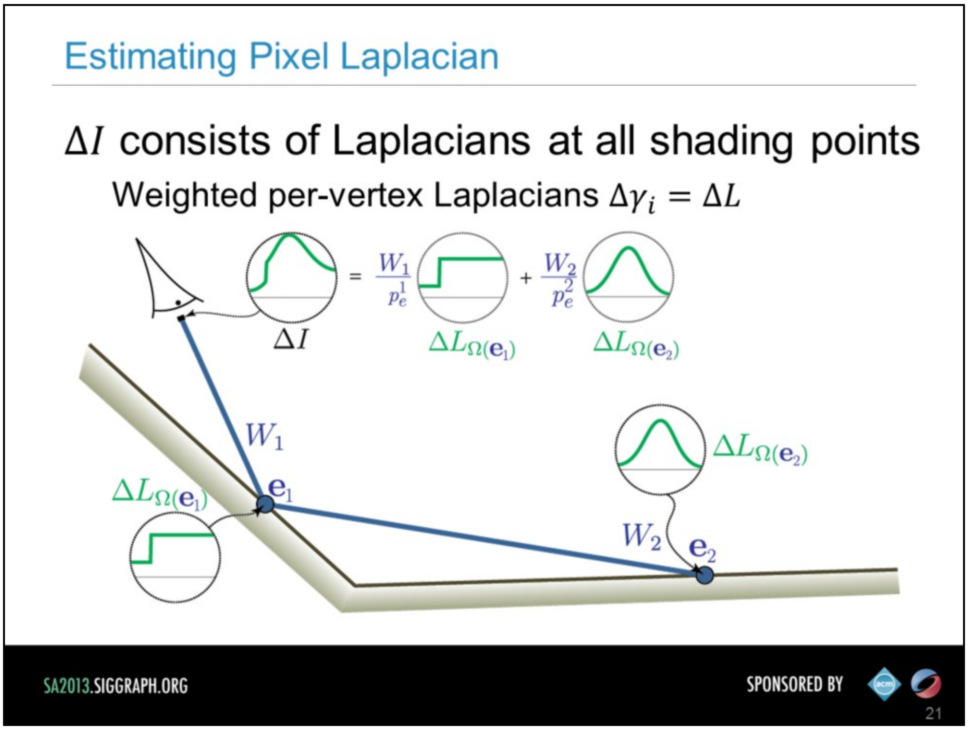
As was explained by Toshiya Hachisuka in his error analysis part of the course, both variance and bias depend on kernel bandwidth.

Interestingly, both terms have a counter-dependence on the bandwidth, which suggests that we can balance between variance and bias just by changing this bandwidth *r*.

We generalize Toshiya's error analysis for stochastic progressive photon mapping case, where the kernel estimation point can also vary.

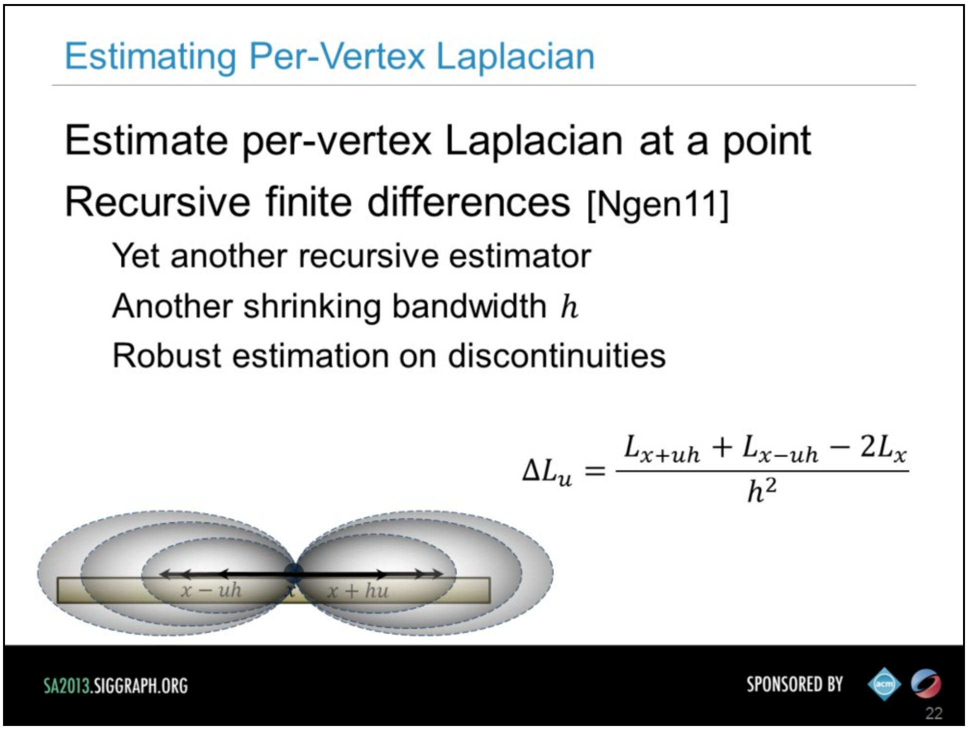
In this case, bias depends on the so-called pixel Laplacian ΔI, which can be intuitively explained as an average curvature of the estimated radiance around the all estimation points within the pixel.

This pixel Laplacian is unknown and needs to be robustly estimated.



In order to estimate pixel Laplacian, we expand it into Laplacians at each potential shading point.

This way we can estimate Laplacian with kernel method at a shading point and then compute a Monte-Carlo estimate of the resulting pixel Laplacian by taking into account the importance weight of the camera subpath and the probability of sampling this shading point.



In order to estimate Laplacian at each shading point, Toshiya mentioned that the derivatives of the kernel can be used.

However, we found that such an estimation is instable on the radiance discontinuities, such as, for example, a hard edge of a shadow.

Thus, we use a recent approach of recursive finite differencing along tangent axes of the surface.

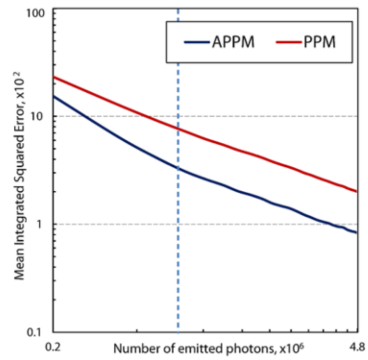
This is a standalone recursive estimator, which has multiple kernel estimators and computes the Laplacian as a recursive finite difference of them.

This estimator has its own independent shrinking bandwidth parameter *h*, which can be selected in a more simplified way and does not influence the quality of bandwidth selection too much (please see the original paper for more details on it).

Unlike kernel derivatives estimator, this recursive finite differencing estimator is robust on discontinuities, which is an important property for our bandwidth selection method.

Adaptive Bandwidth Selection

- Estimate all unknowns
  - Path variance
  - Pixel Laplacian
- Minimize MSE as MSE(r)
- Lower initial error
  - Keeps noise-bias balance
  - Data-driven bandwidth selector



Now, that we have all the unknowns to select the optimal bandwidth on the fly, we can achieve faster convergence in practical scenarios of finite time.

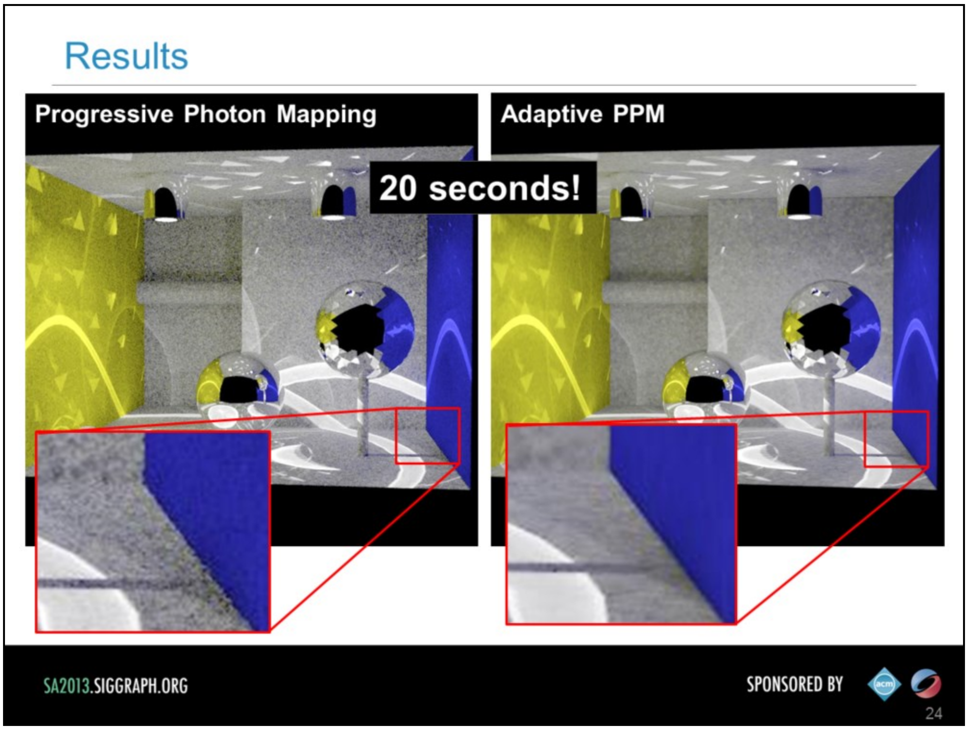
After estimating all the unknowns, we minimize the MSE with respect to the parameter *r*.

After a few photon maps, MSE becomes already smaller than with the best manually-chosen parameters.

However, at some point, the lines become parallel, which suggests that the asymptotic convergence rate is the same.

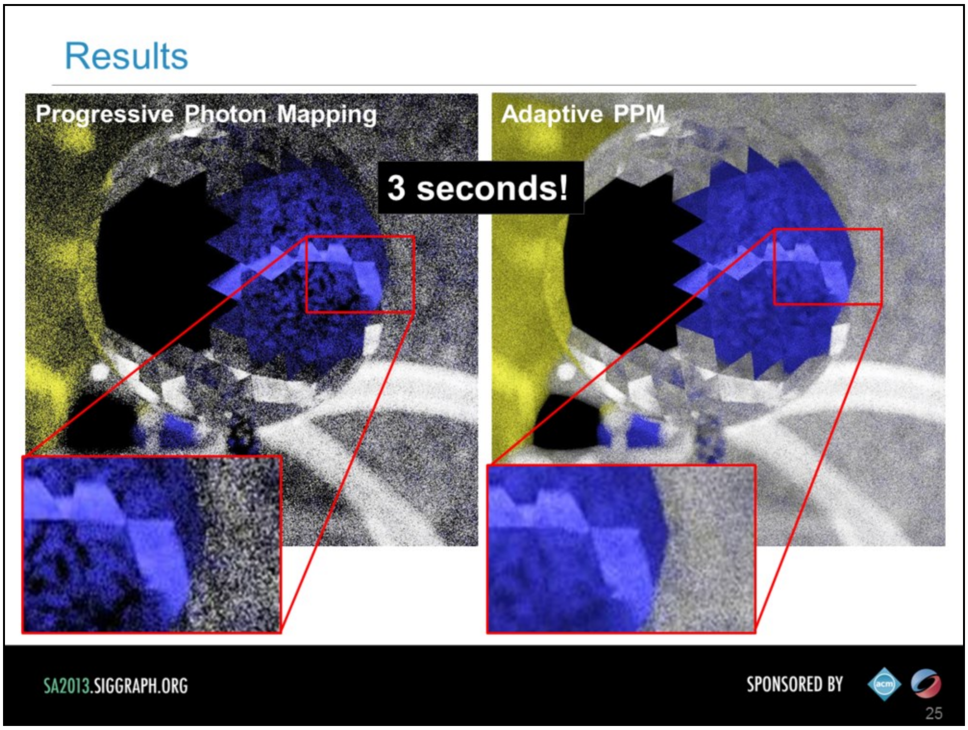
The proposed selector always tries to keep the balance between image noise and bias.

The selector uses past samples to estimate several auxiliary values, so it learns from the data.

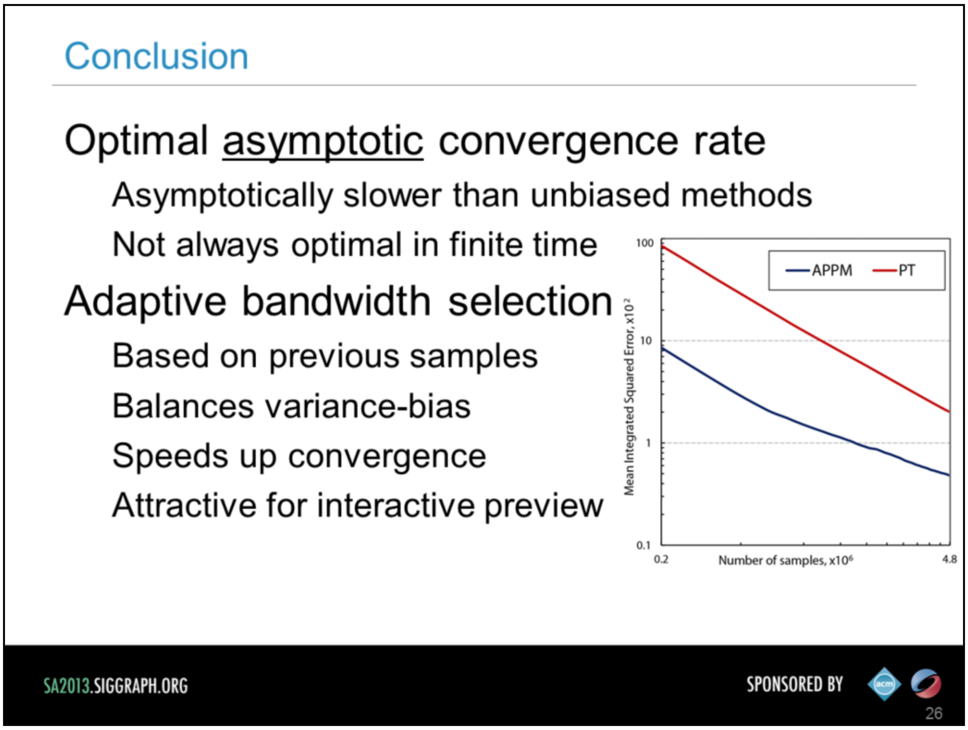


Here are some results.

Equal time comparison. Less variance in flat regions, yet caustic edges are sharp.



Equal time comparison. Heavy out of focus region (different shading points within pixel support).
Less variance, caustic edges remain sharp even out of focus, yet gathering radius is enlarged where possible.



To conclude my part, we have deduced the **asymptotic** convergence rate for photon mapping class methods.

The asymptotic convergence is slower comparing to unbiased methods whenever one applies a kernel estimation (please notice that the slopes of the plots on the figure are different).

However, that does not necessarily mean the results are inferior – in many practical cases the initial error is much lower because of efficient variance reduction; and might still stay low after long rendering time.

The second part is the adaptive bandwidth (radius) selection.

The method uses the same photon map data to estimate the pixel Laplacian, thus an additional sampling is not needed.

This selection technique allows to significantly cut the amount of bias and variance in the image starting already from the first photon map.

This property makes it attractive for fast interactive preview renderers, but also reduces the bias in final images.

Thank you for your attention.