Bayesian and Quasi Monte Carlo Spherical Integration for Illumination Integrals

Ricardo Marques[1,2], Christian Bouville[2], Kadi Bouatouch[2]

[1]Institut National de Recherche en Informatique et Automatique (INRIA)

[2]Institut de Recherche en Informatique et Systhèmes Aléatoires (IRISA)

April 7, 2014

Outline

Introduction

Monte Carlo and Quasi-Monte Carlo

- *Monte Carlo* (MC) is the base method on image synthesis but converges slowly: $(N^{-0.5})$
- Quasi-Monte Carlo (QMC)
  - Deterministic sampling for faster convergence rates:
    - $N^{-1}(\log N)^d$, d being the dimensionality, for unit hypercube integration domain
    - $N^{-0.75}$ for $d = 3$: unit sphere integration domain
  - But this convergence rate decreases when the dimensionality increases
  - Implicit assumption: smoothness of the integrand at least C^0 continuous
  - Such assumption is often not verified for illumination integrals

This leads to the following questions:

- Can we characterize the smoothness of integrands so as to better exploit this knowledge for computing more accurate integral estimates?
- Can we smooth out integrand discontinuities without loosing too much in accuracy?

We wil show:

- Problems arising when prefiltering (for smoothing the integrand) in the context of QMC,
- *Bayesian Monte Carlo* (BMC) method provides a mathematical framework to address this problem

Problem statement

- Focus on the case of hemispherical integration for illumination integrals
- Detailed analysis of the factors which determine the quality of the integral estimate:
  - Sample distribution
  - Samples' weight
  - Smoothness of the integrand
- Play with those factors to improve the quality of the estimate

Objective and Applications

Objective: synthesize physically-based photo-realistic images.

The illumination integral

$$L_o(\mathbf{x}, \omega_o) = L_e(\mathbf{x}, \omega_o) + \int_{\Omega_{2\pi}} L_i(\mathbf{x}, \omega_i)\ \rho(\mathbf{x}, \omega_i, \omega_o)\ (\omega_i \cdot \mathbf{n})\ d\Omega(\omega_i)$$

where ω is a spherical direction given by (θ, ϕ), [Kaj86].

- No analytical solution!
- Common to resort to stochastic methods (e.g., Monte Carlo).
- Massive use of sampling operations.

Direct and Indirect Light Components

$$L_o(\mathbf{x}, \omega_o) = L_e(\mathbf{x}, \omega_o) + \boxed{\int_{\Omega_{2\pi}} L_i^{ind}(\mathbf{x}, \omega_i)\ \rho(\mathbf{x}, \omega_i, \omega_o)\ (\omega_i \cdot \mathbf{n})\ d\Omega(\omega_i)}$$

$$+ \int_{\Omega_{2\pi}} L_i^{dir}(\mathbf{x}, \omega_i)\ \rho(\mathbf{x}, \omega_i, \omega_o)\ (\omega_i \cdot \mathbf{n})\ d\Omega(\omega_i)$$

Direct

Indirect

Direct + Indirect

Frequency Domain View

Frequency domain view of sampling and integration

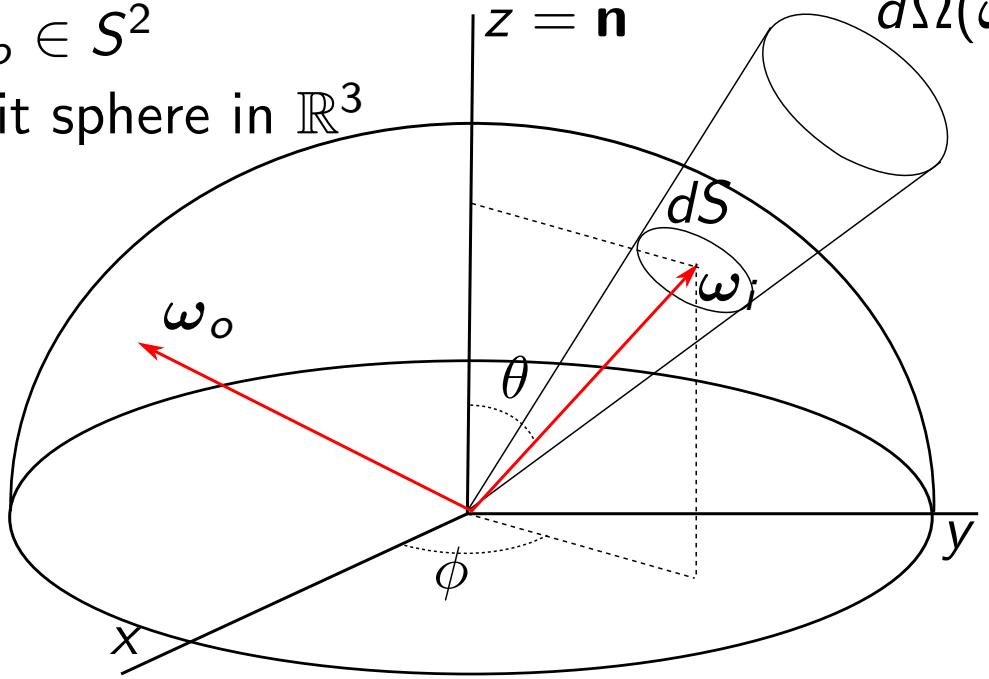
- How can we relate the integral estimate error and the Fourier spectrum of the integrand?
- Will consider mainly QMC integration (see Subr and Kautz, SIGGRAPH 2013 for the stochastic sampling case)
- For clarity, we will base our analysis on the case of circular functions in $\mathbb{R}^2$ instead of spherical functions in $\mathbb{R}^3$

The illumination integral

$\omega_i, \omega_o \in S^2$

S^2 unit sphere in $\mathbb{R}^3$

$d\Omega(\omega_i) = dS = \sin\theta \, d\theta d\phi$



$$L_o(\omega_o) = \int_{\Omega_{2\pi}} L_i(\omega_i) \; \rho(\omega_i, \omega_o) \; (\omega_i \cdot \mathbf{n}) \; \; d\Omega(\omega_i)$$

Estimate: $\tilde{L}_o(\omega_o) = \frac{1}{N} \sum_{j=1}^{N} L_i(\omega_j)$

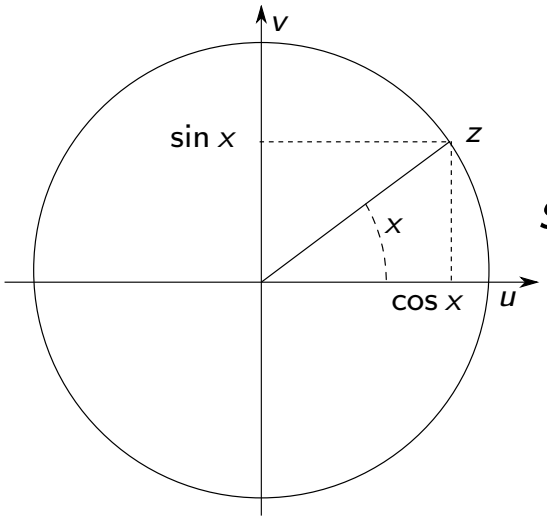
Spherical harmonics and Fourier series

- Frequency view is discrete on the unit sphere $\mathbb{S}^2$ (spherical functions are implicitly periodic)
- Basis functions are the spherical harmonics (SH): $Y_{l,m}(\theta, \phi)$ with $|m| \leq l \in \mathbb{N}$
- Projections of a function $f(\theta, \phi)$ on the $Y_{l,m}$ gives the Fourier coefficients:

$$f_{l,m} = (f, Y_{l,m}) = \int_0^{\pi} \int_0^{2\pi} f(\theta, \phi) Y_{l,m}^*(\theta, \phi) \sin\theta d\theta d\phi$$

- Fourier series equivalent to SH for circular functions $s(z)$ in $\mathbb{R}^2$, i.e. $z \in \mathbb{S}^1$.

Circular functions and Fourier series



$$s(z) = s(u, v) = s(\cos x, \sin x)$$

$f(x) := s(\cos x, \sin x)$ is 2π periodic $\Rightarrow$ Fourier series:

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{jnx}$$

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-jnx} dx$$

Fourier series view of integration

Original

$$\text{Let: } g(x) = \frac{1}{K}\sum_{k=0}^{K-1} f(x+x_k), \qquad \{x_k\} \in [-\frac{a}{2}, \frac{a}{2}] \text{ (sampling pattern)}$$

$$\text{Goal: } I = \frac{1}{a}\int_{-\frac{a}{2}}^{\frac{a}{2}} f(x)dx, \quad \text{Estimate: } \tilde{I} = \frac{1}{K}\sum_{k=0}^{K-1} f(x_k) = g(0)$$

Fourier

$$f(x) \xrightarrow{\mathcal{F}} \{a_n\} \qquad g(x) \xrightarrow{\mathcal{F}} \{a_n c_n\} \quad \text{with } c_n = \frac{1}{K}\sum_{k=0}^{K-1} e^{jnx_k}$$

$$I = \sum_{n=-\infty}^{\infty} a_n \,\mathrm{sinc}(\frac{na}{2}) \qquad \tilde{I} = g(0) = \sum_{n=-\infty}^{\infty} a_n c_n$$

Fourier series view of integration

$$I = \sum_{n=-\infty}^{\infty} a_n \,\mathrm{sinc}(\frac{na}{2}) \qquad \tilde{I} = g(0) = \sum_{n=-\infty}^{\infty} a_n c_n$$

- $\sum_{n=-\infty}^{\infty} a_n c_n$ represents the frequency distribution of the integral estimate
- Results from the *product* of the integrand spectrum ($\{a_n\}$) and the sampling pattern spectrum ($\{c_n\}$)
- In case of uniformly distributed samples on $[-a/2, a/2]$:

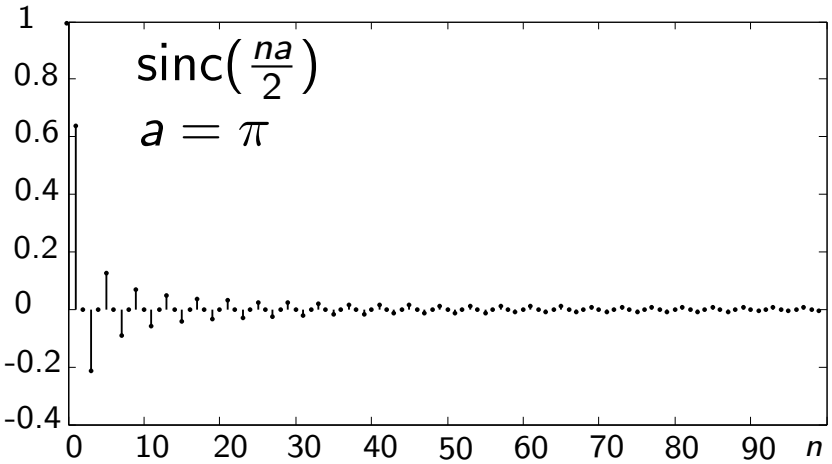
$$x_k = \frac{ka}{K} + \frac{a}{2}\frac{1-K}{K}$$

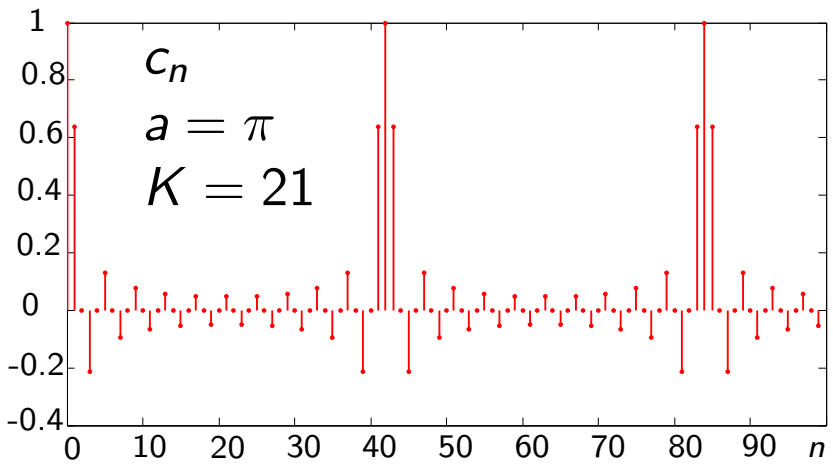
which gives: $c_n = sinc(\frac{na}{2})/\mathrm{sinc}(\frac{na}{2K})$

Frequency domain view for uniform sampling

For uniform sampling: $c_n = sinc(\frac{na}{2})/\text{sinc}(\frac{na}{2K})$

$c_n \approx 1$ when $n = mS$, $S = \frac{2K\pi}{a}$ is the sampling frequency





High frequency components have much more effect on $\tilde{I}$ than on I

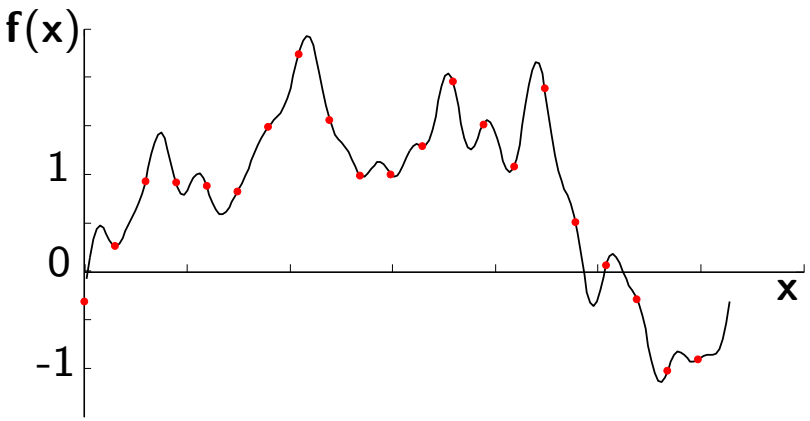
$\Rightarrow$ The error $I - \tilde{I}$ mainly depends on the high frequency components of $f(x)$

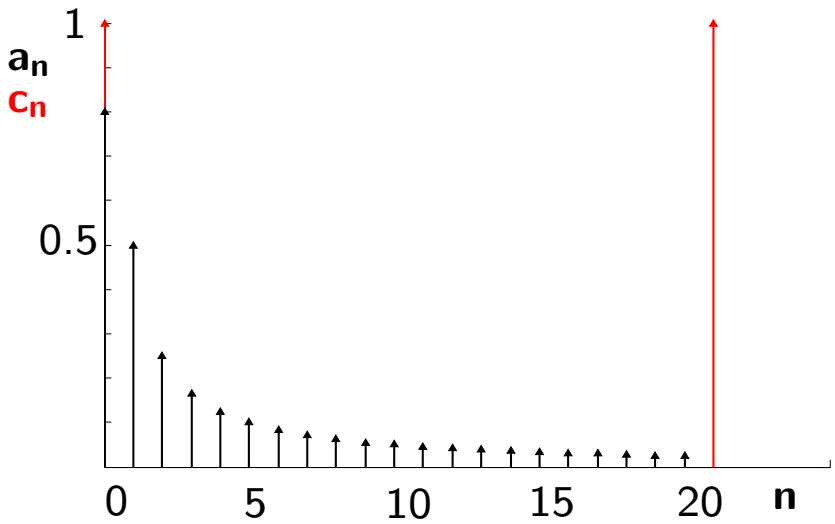
Special case: $a = 2\pi$ and uniform sampling (1)

In this case: $I = a_0$ and $c_n = 1$ if $n = mK$, $c_n = 0$ elsewhere

$$\Longrightarrow \tilde{I} = \sum_{m=-\infty}^{\infty} a_{mK}$$

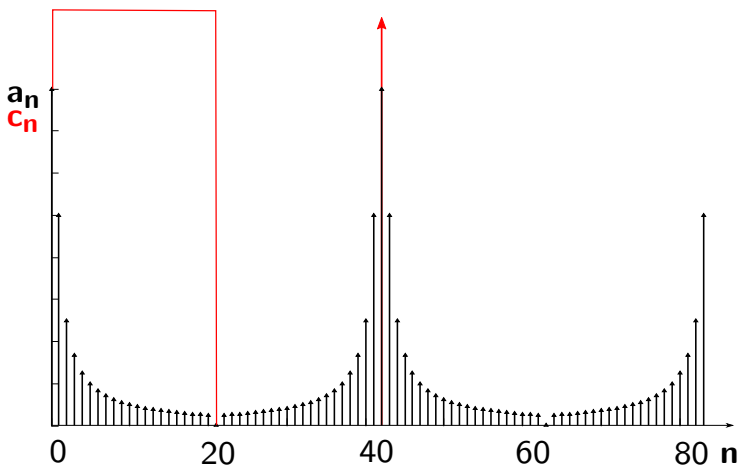
$\Longrightarrow I = \tilde{I}$ if $f(x)$ band-limited (BL) to N harmonics and $N < K$.
Example: $K = 21$ samples and $f(x)$ has $N = 20$ harmonics

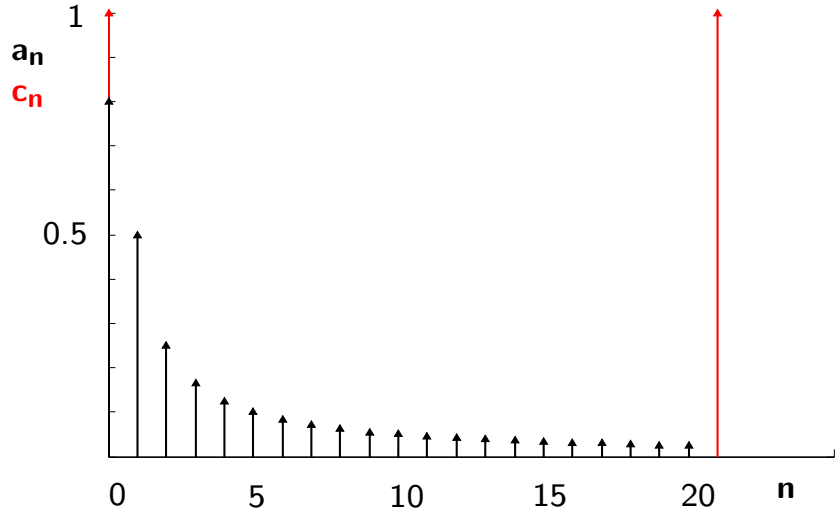




Integration and sampling theorem

Example: $f(x)$ has $N = 20$ harmonics





- 42 samples for exact reconstruction but only 21 samples for exact integration
- a samples set that enables exact integration on $\mathbb{S}^d$ in $\mathbb{R}^{d+1}$ is a *spherical design*

A glimpse at the case of non-periodic integrands

Discrete sums become integrals for non-periodic integrands:

$$I = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)\mathrm{sinc}(\omega a/2)d\omega \quad \tilde{I} = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)\frac{\mathrm{sinc}(\omega a/2)}{\mathrm{sinc}(\omega a/2K)}d\omega$$

$\Longrightarrow$ The Fourier transform of the sampling pattern is $\frac{\mathrm{sinc}(\omega a/2)}{\mathrm{sinc}(\omega a/2K)}$

It becomes a Dirac comb only if $a \to \infty$ at fixed sampling period $T = a/K$ and then (Poisson summation formula):

$$\int_{-\infty}^{\infty} f(x)dx = T\sum_{n=-\infty}^{\infty} f(nT)$$

with $f(x)$ band-limited to $\omega_M < 2\pi/T$
$\Longrightarrow$ exact integration impossible in practice

Special case: $a = 2\pi$ and uniform sampling (2)

- If $f(x)$ has N harmonics and samples number is K:
- Uniform sampling pattern yields $c_n = 0$ for $0 < |n| < K$, which enables exact integration of BL functions
- Exact integration of BL function requires $K > N$
- Exact reconstruction would require $K > 2N$ (Sampling theorem)
- Exact integration of BL spherical functions also exists on the $\mathbb{S}^2$ sphere with spherical designs [DGS77]
  A point set $\{x_0, \ldots, x_{K-1}\}$ on $\mathbb{S}^2$ is a spherical design if:

$$\frac{1}{4\pi} \int_{S^2} f(x) dS(x) = \frac{1}{K} \sum_{k=0}^{K-1} f(x_k)$$

Spherical designs in $\mathbb{S}^2$

- In the circular functions case ($\mathbb{S}^1$), $c_n = 0$ for $0 < |n| < K$ entails:

$$c_n = \frac{1}{K} \sum_{k=0}^{K-1} e^{jnx_k} = 0$$

- Equivalently, in $\mathbb{R}^3$, using the SH basis functions $Y_{l,m}(\theta, \phi)$ (c_n becomes $c_{l,m}$), for spherical designs [ACSW10]:

$$c_{l,m} = \frac{1}{K} \sum_{k=0}^{K-1} Y_{l,m}(\theta_k, \phi_k) = 0 \quad 0 < l < L, \quad |m| \leq l$$

- For a function band-limited to L harmonics, exact integration is possible if [DGS77]:

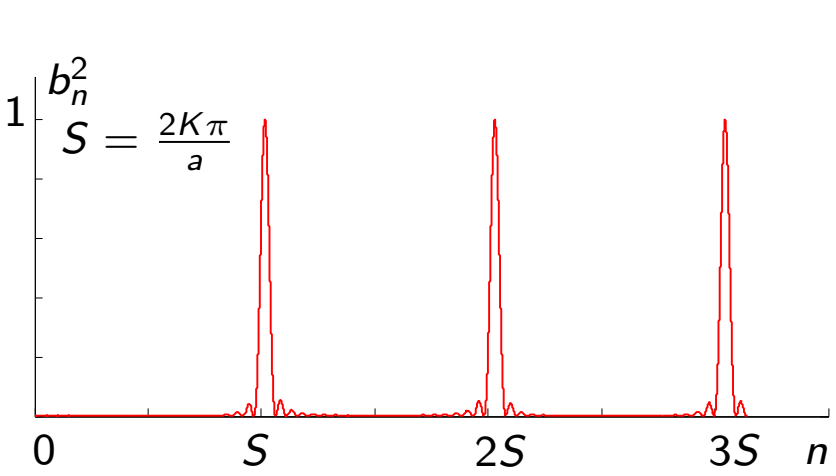
$$K \geq \frac{(L+1)(L+3)}{4} \text{ if } L \text{ odd, and } K \geq \frac{(L+2)^2}{4} \text{ if } L \text{ even}$$

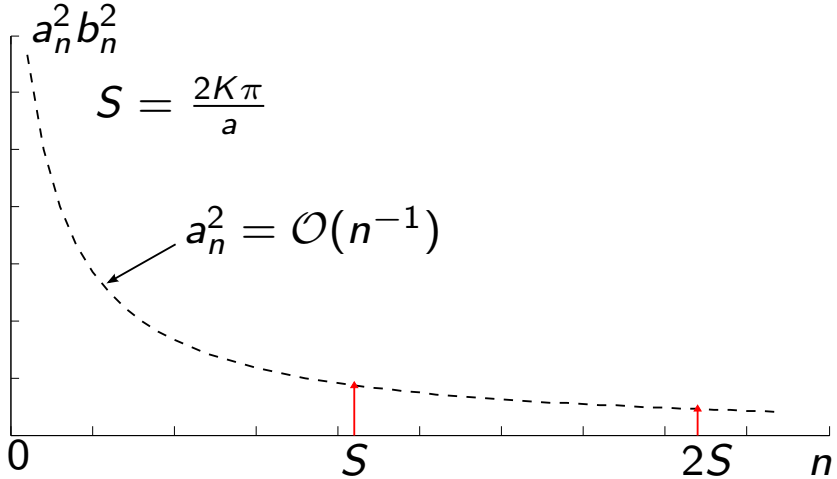
Error analysis for uniform sampling

When $a < 2\pi$:

$$|I - \tilde{I}| = \left| \sum_{n=-\infty}^{\infty} a_n b_n \right| = 2 \left| \sum_{n=1}^{\infty} \Re(a_n b_n) \right|$$

with: $b_n = \text{sinc}(\frac{na}{2}) - c_n \qquad c_n = sinc(\frac{na}{2})/\text{sinc}(\frac{na}{2K})$





The error depends on the **sampling frequency S** and **the rate of decay of the a_n**

Frequency distribution of error noise in the case of uniform scrambling

- Scrambling is often necessary in rendering to avoid regular patterns
- Uniform scrambling when $a = 2\pi$:
  $X_k = x_k + \delta \quad \forall k$, and $\delta \sim U(\delta| -\pi, \pi)$
- Preserves optimality and has no bias.
- Power spectrum of error function: $(I - \tilde{I}_\delta)$: $\{|a_n b_n|^2\}$
- Error variance:

$$E[(I - \tilde{I}_\delta)^2)] = 2\sum_{n=1}^{\infty} |a_n b_n|^2 = 2\sum_{m=1}^{\infty} |a_{mk}|^2$$

Optimality of samples set as viewed in the frequency domain

- Optimality is obtained when, given set size K, c_n small in $0 < n < N$ for the largest possible N
- For uniform sampling, the largest N corresponds to the sampling frequency $S = 2K\pi/a$
- When $a = 2\pi$, $c_n = 0$ for all n up to $n = S = K$
- On the $\mathbb{S}^2$ sphere, spherical designs are optimal: $c_{l,m} = 0$ up to $l = L$ with $L = 2\sqrt{K} - 2$ at best [DGS77]
- Best sphere packing (min. $\text{dist}(x_i, x_j) = \delta_K$) is not generally a spherical design but has small $c_{l,m}$ up to the "sampling frequency" [SK97] (and using the Jeans' rule):

$$L \approx \frac{2\pi}{\arccos(1 - \delta_K^2/2)} - \frac{1}{2} \quad \text{with} \quad \delta_K = \sqrt{\frac{8\pi K}{\sqrt{3}}}$$

Modeling the integrand in the frequency domain (1)

- Classic QMC theory uses *variation of f in the sense of Hardy and Krause* [Nie92] $\Longrightarrow$ no direct frequency interpretation
- Spherical QMC uses Sobolev spaces $\mathbb{H}^s(\mathbb{S}^2)$ [BSSW12, Sob05]
  The parameter s characterizes the smoothness of the function

Modeling the integrand in the frequency domain (2)

- Sobolev spaces $\mathbb{H}^s(\mathbb{S}^2)$ on the $\mathbb{S}^2$ sphere are defined as the space of functions $f \in \mathbb{L}_2(\mathbb{S}^2)$ whose Fourier coefficients satisfy:

$$\sum_{l=0}^{\infty} \sum_{m=-l}^{m=l} (1 + l + l^2)^s \, |f_{l,m}|^2 < \infty$$

- Spectrum rate of decay:
  If $f \in \mathbb{H}^s(\mathbb{S}^2)$, $a_{l,m}$ decreases in $\mathcal{O}(l^{-s})$ at least
- Continuity (Embedding theorem): $f \in C^k(\mathbb{S}^2)$ if $s > k + 1$
- If f defined over $\mathbb{S}^2$ and $(s = 3/2) \Longrightarrow f$ is continuous
- $\mathbb{H}^s(\mathbb{S}^2) \subset \mathbb{H}^{s'}(\mathbb{S}^2)$ if $s > s'$

Modeling the integrand in the frequency domain (3)

- Alternative approach: model $f(x)$ by a random process $f_r(x)$ characterized by a stationary covariance function $k(\tau)$, $\tau = x - x'$
- $\Longrightarrow$ the integrand is considered as random but the sampling pattern is considered as deterministic
- A power spectrum $F_r(\omega)$ of $f_r(x)$ can be derived by the Wiener-Khintchine theorem:

$$F_r(\omega) = \mathcal{F}(k) = \int k(\tau) e^{-j\omega\tau} d\tau$$

- Common approach in machine learning, image processing and coding
- Will be used in Bayesian Monte-Carlo

Worst case error over $\mathbb{H}^s(\mathbb{S}^2)$

If $f \in \mathbb{H}^s(\mathbb{S}^2)$ and $s > 1$ [ACSW10]:

$$WCE = \left(\sum_{l=1}^{\infty} a_l^{(s)} \sum_{m=-l}^{m=l} |c_{l,m}|^2 \right)^{1/2}$$

$\{c_{l,m}\}$ are the Fourier coefficients of the sampling pattern:

$$c_{l,m} = \frac{1}{K} \sum_{k=0}^{K-1} Y_{l,m}(\theta_k, \phi_k)$$

$\{a_l^{(s)}\}$ represents the $\mathbb{H}^s(\mathbb{S}^2)$ frequency behavior:

$$a_l^{(s)} = k_s(1+l)^{-2s}$$

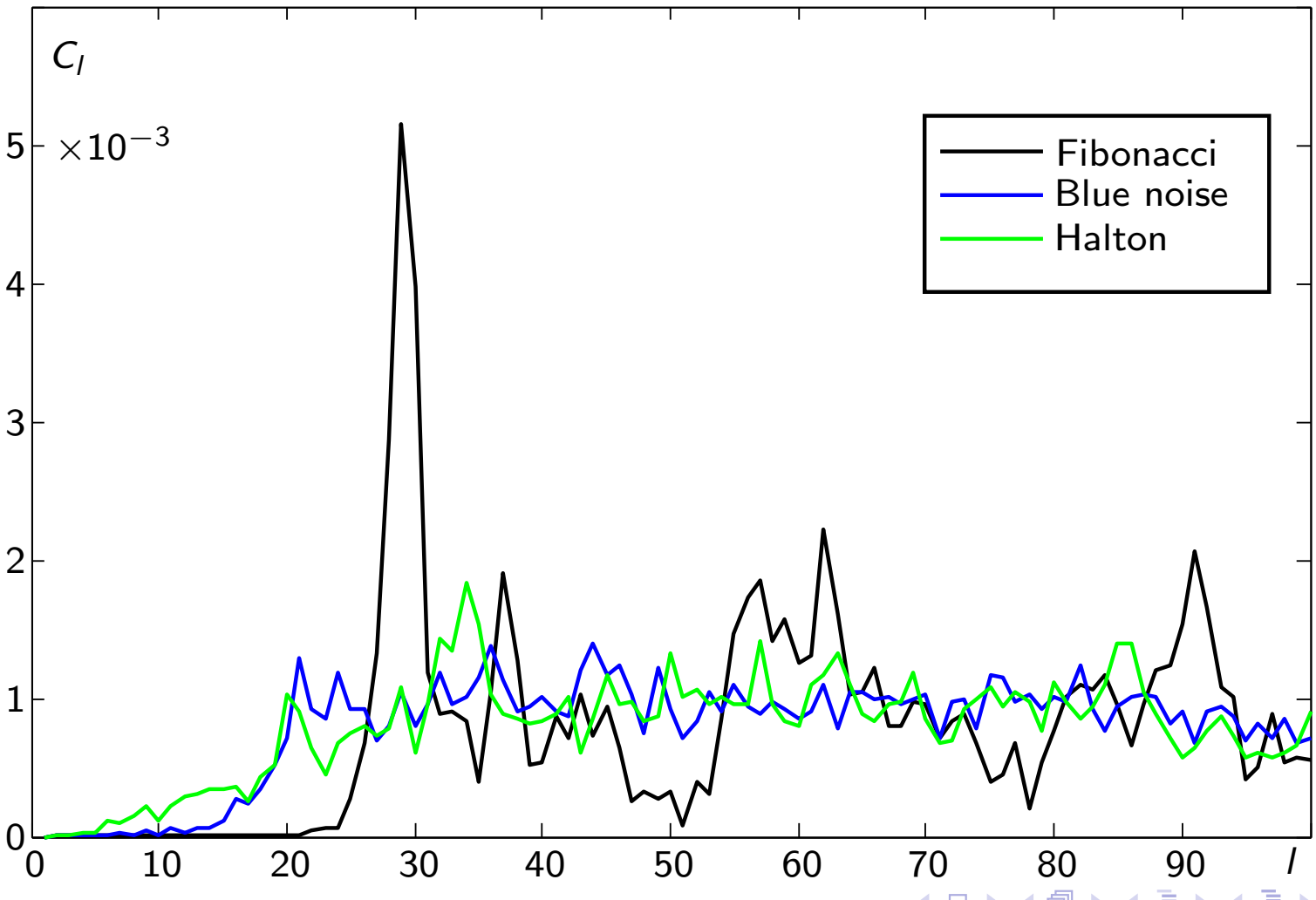
Visualization of spectrum of spherical point sets

The coefficients are averaged in the $\{m\}$ dimension [LWSF10]:

$$C_l = \frac{1}{2l+1} \sum_{m=-l}^{m=l} |c_{l,m}|^2$$

Spectrum of the sampling pattern: $\{C_l\}$, $l = 0, \infty$

Comparison of spectra of different spherical points distribution



Integration on the hemisphere

- Recall that the illumination integral is over an hemisphere of $\mathbb{S}^2$
- Similar to the $\mathbb{S}^1$ case with $a = \pi$ i.e., all frequency components of the integrand contribute to the integral value
- The weights on the Fourier coefficients of the integrand are not given by sinc functions but decrease in $\mathcal{O}(l^{-1})$ as well
- Details will be given in a future paper

Integration error and smoothness

- Exact integration not possible when $a < 2\pi$ and so on the hemisphere in $\mathbb{R}^3$
- Estimation error depends on the rate of decrease of a_n
- Smoothness, continuity and rate of decrease of a_n are linked (Sobolev space theory): roughly $\mathcal{O}(l^{-3/2})$ for continuous function on $\mathbb{S}^2$
- Can we cut off high frequencies by a lowpass filter before integration? [CK07]
- On $\mathbb{S}^1$, bandwidth must be lower than S/2 to avoid aliasing
- Additional error if the integration domain is not a full period ($a = 2\pi$ in the $\mathbb{S}^1$ case)

Lowpass prefiltering and samples weighting (1)

Lowpass filter: $h(x)$, Prefiltered function: $f_p(x)$.

$$f_p(x) = \int_{-\pi}^{\pi} f(x')h(x - x')dx'$$

If $I = \int_{-a/2}^{a/2} f(x)dx \approx \int_{-a/2}^{a/2} f_p(x)$, then:

$$I \approx \int_{-\pi}^{\pi} f(x')p(x')dx'$$

with:

$$p(x') = \int_{-\frac{a}{2}}^{\frac{a}{2}} h(x - x')dx$$

$\Longrightarrow$ amounts to samples weighting with $p(x)$
but integration is now a full period instead of $[-a/2, a/2]$

Lowpass prefiltering and samples weighting (2)

Problem: how to deal with samples outside the integration domain $[-a/2, a/2]$? Tricky compromise:

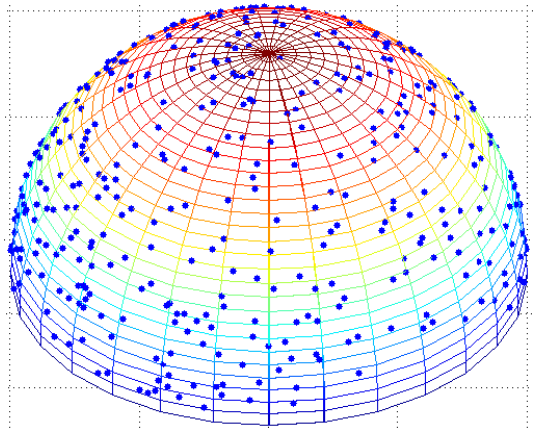
- Sampling outside $[-a/2, a/2]$ (if possible) means extra computation
- Setting $f(x) = 0$ outside the integration domain increases high frequencies and ignoring them increases the error
- Reducing filter support (i.e. its length) impacts filter efficiency
- Extrapolation could be a solution but prior model is required $\Longrightarrow$ Bayesian Monte Carlo

Quasi Monte Carlo for Illumination Integrals

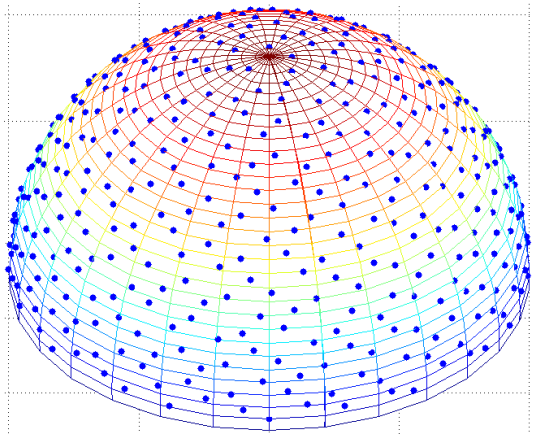
[MBR+13a, Mar13]

Spherical Quasi-Monte Carlo: A Brief Presentation

- Classical Monte Carlo (CMC): samples position is randomly generated.
  - Convergence rate: $N^{-1/2}$



- Quasi-Monte Carlo (QMC): the samples position is deterministic.
  - Resort to low discrepancy sample sets.
  - Best theoretical convergence rate: $N^{-3/4}$ (spherical integration).



Overview of QMC

- The spherical QMC rules are not well known to the CG community.
- We give a comprehensive description of such rules.
  - Relate upper bound of the error to the sum of distances between the samples.
  - Show that the sum of distances can be used to quickly assess the quality of a samples set.
  - More general metrics than spherical discrepancy.
- Present and apply the *spherical Fibonacci point sets* [HN04] to illumination integral evaluation.
  - Point sets which minimize the w.c.e. (worst case integration error)

The Spherical QMC Estimator

- Given a spherical function $f(\omega)$, the QMC estimator for

$$\frac{1}{2\pi}\int_{\Omega_{2\pi}} f(\omega)d\omega \quad \text{is given by} \quad \frac{1}{N}\sum_{j=1}^{N} f(\omega_{j,N}),$$

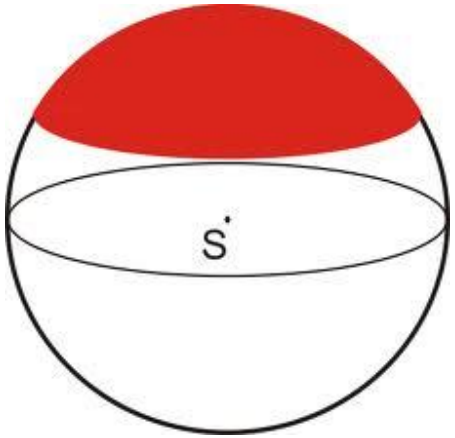
  where $P_N = \{\omega_{j,N} \in \mathcal{S}^2, j = 1 \dots N\}$ is a uniformly distributed samples set.

- The worst case integration error (w.c.e.) of P_N is defined as:

$$\text{w.c.e.} := e(P_N) = \sup_f \left| \frac{1}{N}\sum_{j=1}^{N} f(\omega_{j,N}) - \frac{1}{4\pi}\int_{\mathbb{S}^2} f(\omega)d\Omega(\omega) \right|$$

Spherical Discrepancy

- Let $\mathcal{C}(\boldsymbol{\omega}, t)$ be a spherical cap centered in $\boldsymbol{\omega}$ and containing all the points $\{\boldsymbol{\omega}_c \in \mathbb{S}^2 : \boldsymbol{\omega}_c \cdot \boldsymbol{\omega} \leq t\}$.



- Given a set $P_N = \{\boldsymbol{\omega}_{j,N}\}$, the spherical cap L_2 discrepancy is defined as [BSSW12]:

$$L_2(P_N) = \left[\int_{-1}^{1} \int_{\mathbb{S}^2} \left| \frac{\mathrm{Card}\{j : \boldsymbol{\omega}_{j,N} \in \mathcal{C}(\boldsymbol{\omega}, t)\}}{N} - \frac{\Omega(\mathcal{C}(\boldsymbol{\omega}, t))}{4\pi} \right|^2 d\Omega(\boldsymbol{\omega}) dt \right]^{\frac{1}{2}}$$

Worst Case (Integration) Error

- In the general case the w.c.e. relates to the sum of all the distances between samples by [Brauchart 2012]:

$$e(P_N) = \left[V_s(\mathbb{S}^2) - \frac{1}{N^2} \sum_{r=1}^{N} \sum_{k=1}^{N} |\omega_r - \omega_k|^{2s-2} \right]^{\frac{1}{2}}$$

where s is the smoothness parameter of the Sobolev space $\mathbb{H}^s(\mathbb{S}^2)$ so that $s > n+1$ for $f(\omega)$ C^n continuous.

$$V_s(\mathbb{S}^2) = \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} \left|\omega - \omega'\right|^{2s-2} d\sigma(\omega) d\sigma(\omega')$$

- So, minimizing $e(P_N)$ amounts to maximizing the sum of the distances (to the power $(2s-2)$) between the samples position
- If an information on n is known, then we can adapt the sampling

Stolarsky's Invariance Principle

- In the particular case where f is C^0 continuous, we have:

$$e(P_N) = \left[\frac{4}{3} - \frac{1}{N^2}\sum_{r=1}^{N}\sum_{k=1}^{N}|\omega_r - \omega_k|\right]^{\frac{1}{2}}$$

because $2s - 2 = 1$ then $s = 3/2$ and $n = 0$.

- The Stolarsky's invariance principle for f C^0 continuous, states that [Brauchart 2012]:

$$\frac{1}{N^2}\sum_{r=1}^{N}\sum_{k=1}^{N}|\omega_r - \omega_k| + 4L_2^2(P_N) = \frac{4}{3}$$

- Consequently, if f is C^0 continuous ($f \in \mathbb{H}^s(\mathbb{S}^2)$), we have:

$$e(P_N) = 2L_2(P_N)$$

- So, in this particular case, minimizing the discrepancy $L_2(P_N)$ amounts to minimizing the w.c.e.

Assess the Quality of the Samples Set

- The sum of distances is an interesting criterion to assess the quality of the distribution of a spherical point set P_N.
  - Tightly related to the w.c.e.
  - Fast to compute.
- We can thus define a distance-based criterion to measure the quality of the samples set:

$$E_N(P_N) = \left(\frac{4}{3} - \frac{1}{N^2} \sum_{r=1}^{N} \sum_{k=1}^{N} |\omega_r - \omega_k| \right)^{\frac{1}{2}}$$

Related Work: QMC Spherical Integration in CG

- Current approach in CG [Dut03, PH10]:
  - Produce a unit square samples set (LD or BN).
  - Apply a spherical projection.
  - Perform QMC using the resulting spherical samples set.
- Problem: the quality of the distribution is impaired by the spherical projection!

Spherical Fibonacci: Motivation

- In the following, we present a strategy to generate point sets directly on the sphere.
  - No spherical projection is needed.
- Several spherical point sets have been compared in [BSSW12].
- The spherical Fibonacci are a good compromise between complexity and efficiency.
  - Good behavior regarding w.c.e. [BSSW12].
  - Applied in other research fields with promising results [HN04].

Spherical Fibonacci Point Sets

- The spherical Fibonacci point sets are directly defined on the sphere [HN04].

$$\left.\begin{array}{rcl} \theta_j & = & \arccos(1-2j/F_m) \\ \phi_j & = & 2\pi\left\{j\frac{F_{m-1}}{F_m}\right\} \end{array}\right\} 0 \le j < F_m,$$

  where F_m is the m^{th} Fibonacci number.

- The points are evenly distributed over the vertical axis $\mathbf{z} = \cos(\theta) = 1 - 2j/F_m$.

- The Fibonacci ratio $\left(\frac{F_m}{F_{m-1}}\right)^{-1}$ is used to compute the angle of rotation ϕ.

Spherical Fibonacci Point Sets

- The Fibonacci ratio quickly approaches the golden ratio Φ as m increases.

$$\lim_{m\to\infty} \frac{F_m}{F_{m-1}} = (1+\sqrt{5})/2 = \Phi$$

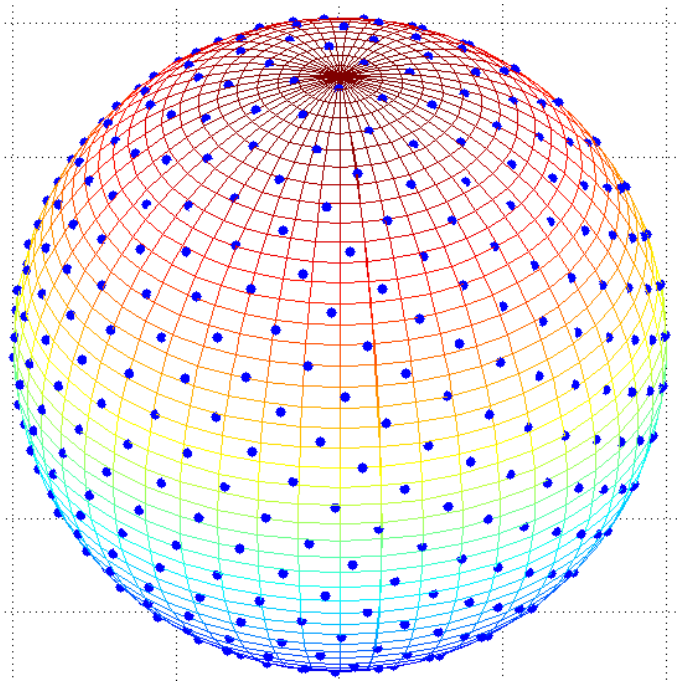
- We can thus replace the Fibonacci ratio by Φ (no more need for N to be a Fibonacci number).

$$\left.\begin{array}{rcl} \theta_j & = & \arccos\left(1-2j/N\right) \\ \phi_j & = & 2\pi\left\{j\Phi^{-1}\right\} \end{array}\right\} 0 \le j < N$$

Spherical Fibonacci point sets

- By introducing an offset of $1/N$ to the z coordinates, [SJP06] achieved a more uniform distribution near the poles.
- The SF point set is then given by:

$$\left.\begin{aligned} \theta_j &= \arccos\left(1 - \frac{2j+1}{N}\right) \\ \phi_j &= 2\pi\left\{j\Phi^{-1}\right\} \end{aligned}\right\} 0 \leq j < N$$



Applying SF to Illumination Integrals

- To evaluate the $L_o^{ind}(\omega_o)$:
  - Generate the samples set on the hemisphere.
  - Morph the point set to fit the BRDF shape.
- Image synthesis requires computing millions of integrals.
  - Using the same point set might result in visible patterns!
  - This problem appears if two consecutive pixels have a correlated noise.
- To avoid a visible structure in the images, we apply a random rotation to the samples set about the vertical axis.

Applying SF to Illumination Integrals

Without random rotation.

With random rotation.

A Direct Application of Spherical QMC

- Recall that we want to approximate the value of:

$$L_o^{ind}(\omega_o) = \int_{\Omega_{2\pi}} L_i^{ind}(\omega_i)\ \rho(\omega_i, \omega_o)\ (\omega_i \cdot \mathbf{n})\ \ d\Omega(\omega_i)$$

- We could directly apply QMC by saying that:

$$f(\omega_i) = L_i^{ind}(\omega_i)\ \rho(\omega_i, \omega_o)\ (\omega_i \cdot \mathbf{n})$$

- Inefficient! Few samples bring an effective contribution.

Morphing the Samples Set

- Let $\rho(\omega_i, \omega_o) = k_s \cos^n(\omega_i, \mathbf{n})/(\omega_i \cdot \mathbf{n})$.
- Make an integration variable substitution to distribute the samples on $\rho(\omega_i, \omega_o)$ $(\omega_i \cdot \mathbf{n})$.
- Using $\theta_i' = g^{-1}(\theta_i) = \arccos\left(\cos^{1/(n+1)}(\theta_i)\right)$, we have:

$$L_o^{ind}(\omega_o) = \frac{k_s}{n+1} \int_{\Omega_{2\pi}} L_i^{ind}(g(\theta_i'), \phi_i')\ d\Omega(\omega_i'), \qquad \omega_i' = (\theta_i', \phi_i')$$

- $g(\theta_i')$ is called *morphing function*: this is what we learn from the scene.

The QMC Estimator for Illumination Integrals

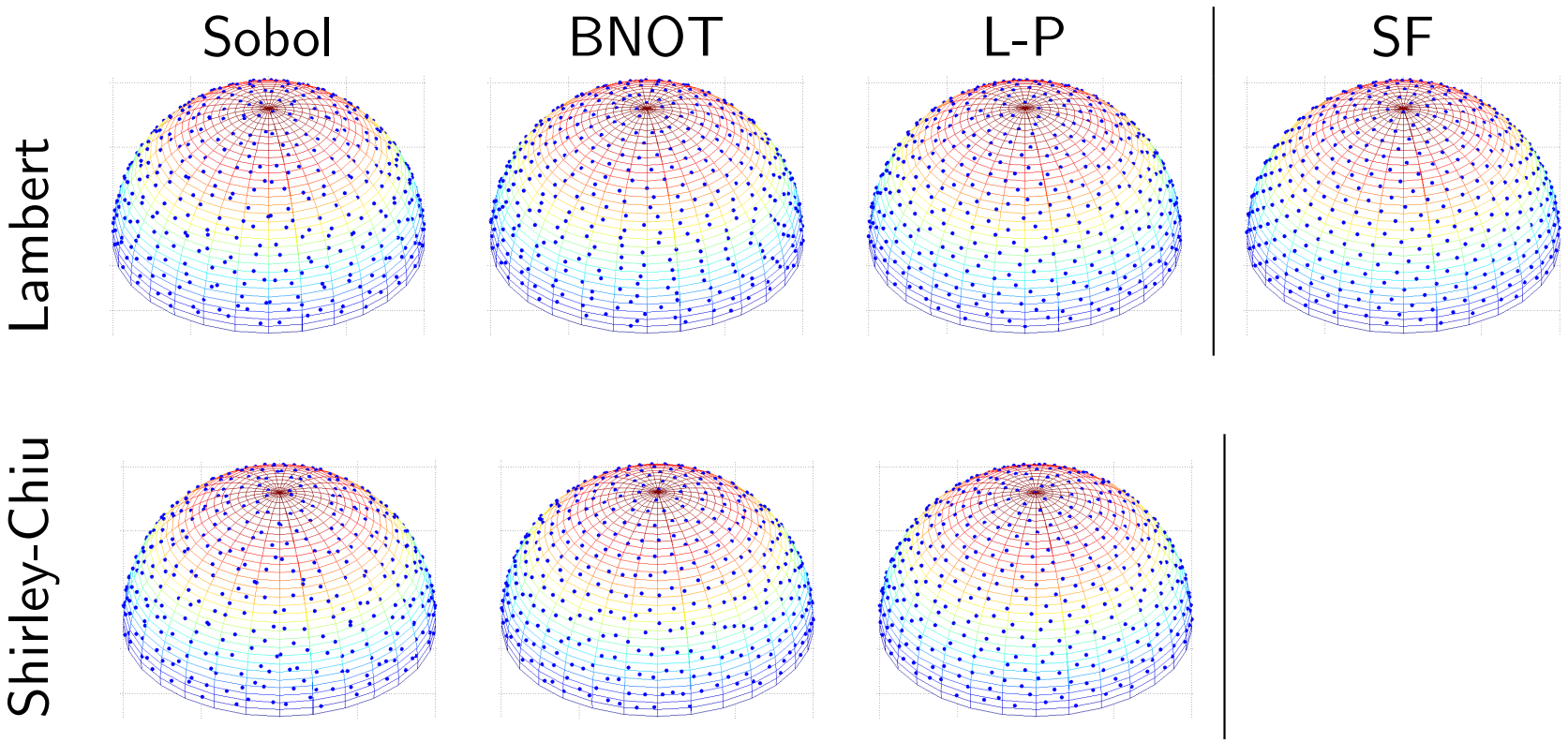
- We can now use QMC to efficiently estimate $L_o(\omega_o)$:

$$L_o^{ind}(\omega_o) \approx \frac{2\pi k_s}{N(n+1)} \sum_{j=1}^{N} L_i^{ind}(g(\theta'_j), \phi'_j)$$

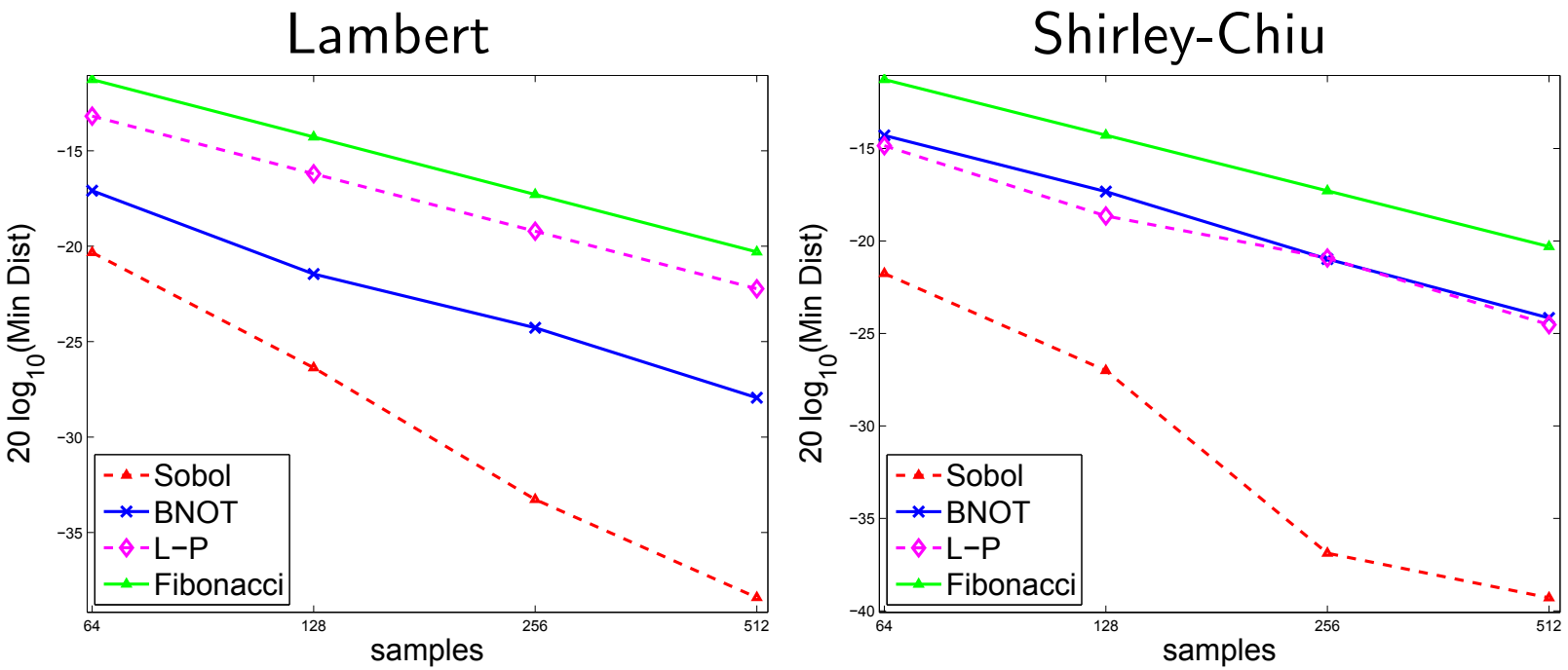
where $\{(\theta'_j, \phi'_j)\} = \{\omega'_j\}$ is asymptotically uniformly distributed.

- The error of the estimate depends on $E_N(\{\omega'_j\})$.

Experimental Set Up

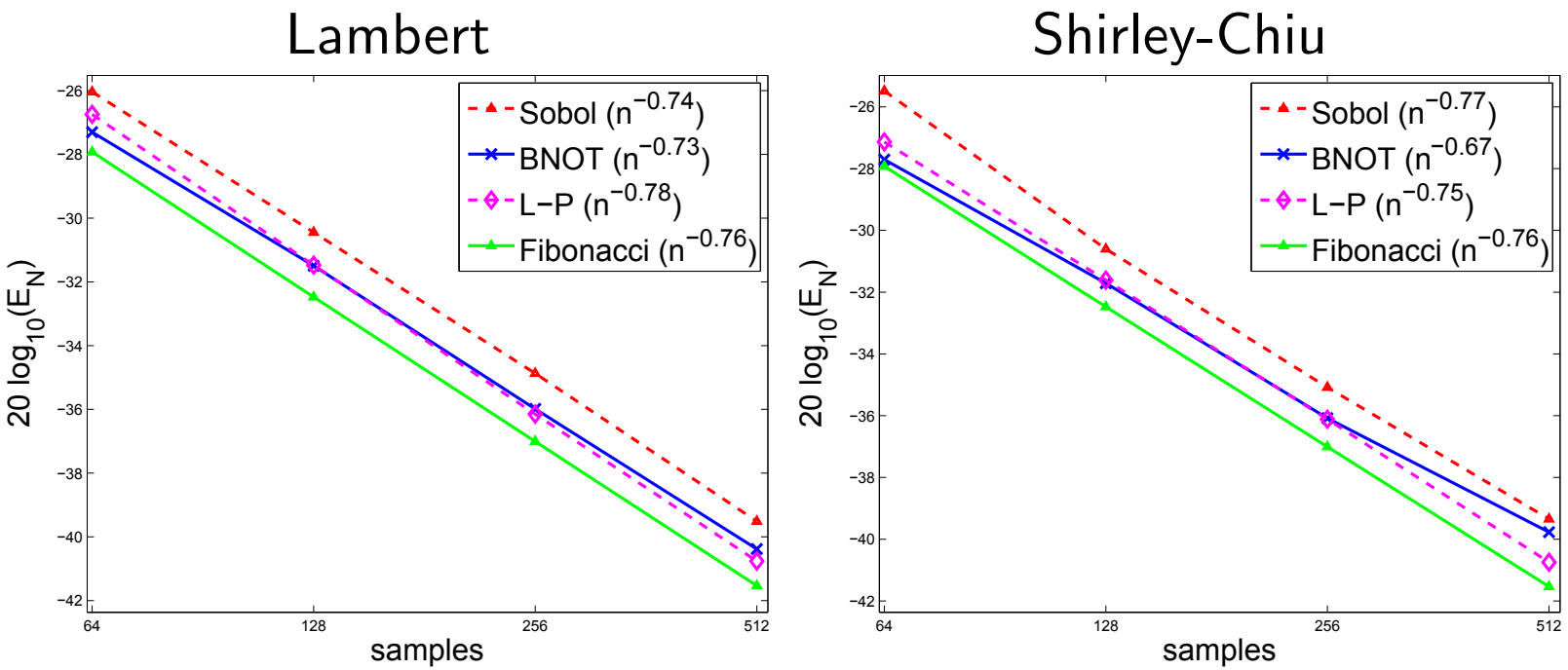


Minimum Inter-Samples Distance



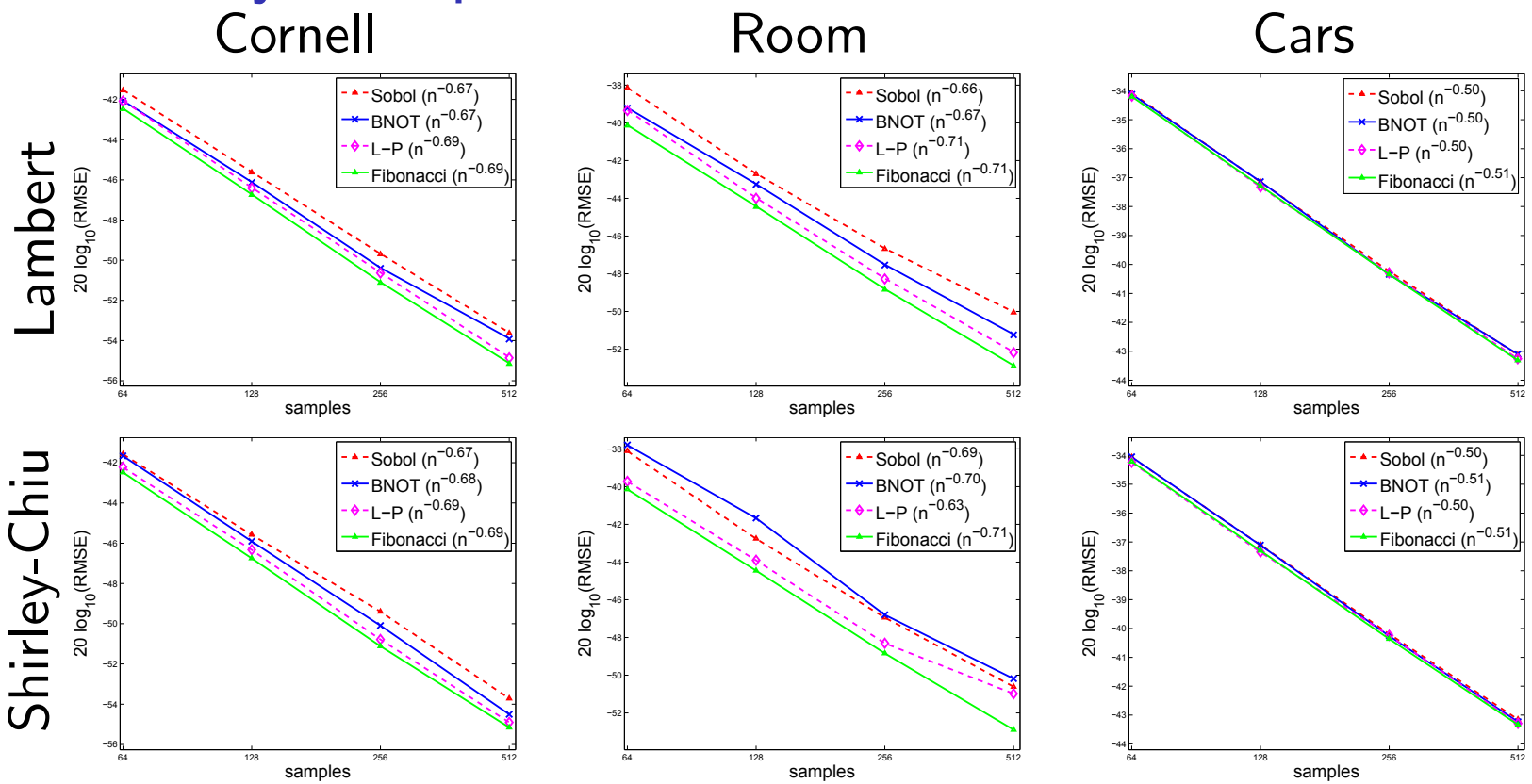
- The distance between closest samples on a SF point set is larger than in the other tested point sets.

Energy of the Sample Sets



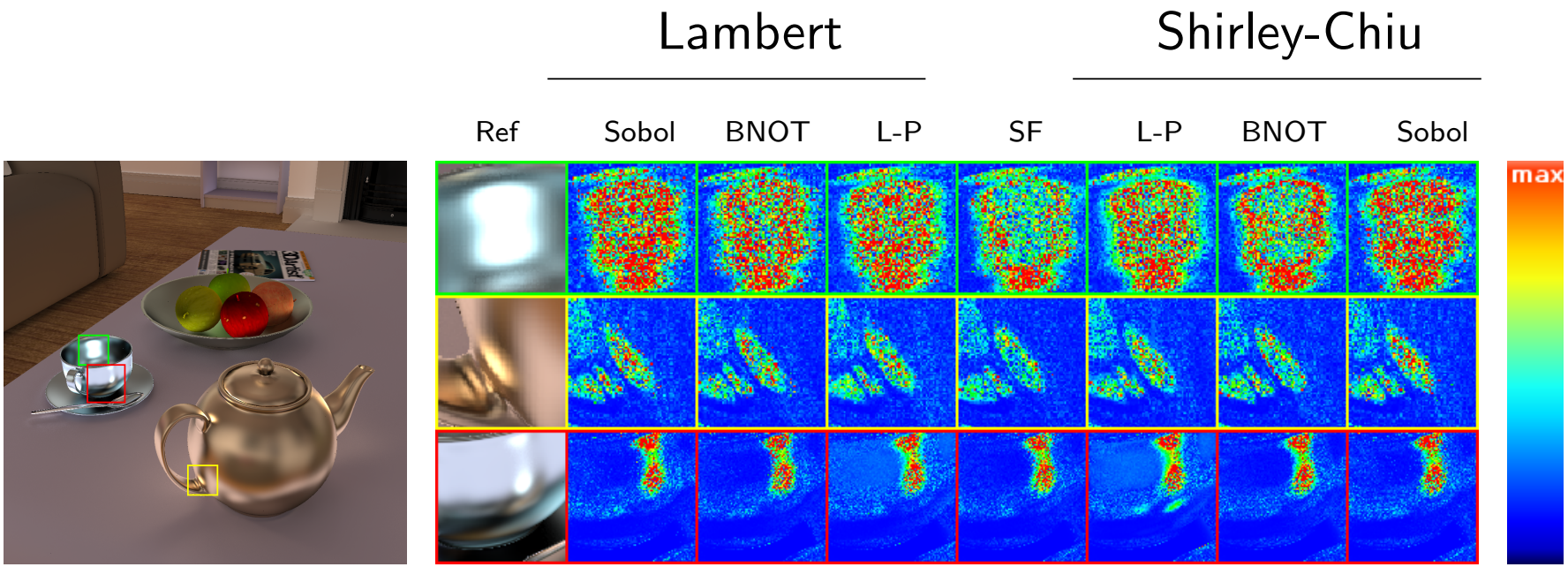
- Spherical Fibonacci yields a smaller energy E_N than the other tested point sets.

RMSE Glossy Component



- The RMSE with SF point sets is consistently smaller.
- The gap is larger when the convergence slope is steeper.
  - SF better exploits continuous functions.

Visual Results: Room Scene



Reference ($\times 4$) Close-up views

Conclusion

- The most important characteristic of an estimation method is its capacity to incorporate existing information.
- QMC only incorporates deterministic knowledge (but no probabilistic knowledge).
- Examples:
  - Morph a samples set to follow the BRDF shape (QMC).
  - Continuity assumption regarding the integrand (QMC).
- Reduction of performance due to discontinuity or lack of smoothness: use methods to smooth the integrand

Conclusion

- We have presented the QMC spherical rules to CG community.
  - Put together very recent advances in the field of QMC spherical integration.
  - The obtained results are in line with the presented theory.
- We applied SF point sets to illumination integrals.
  - Outperforms traditional QMC point sets.
  - Simple to generate and a single point set is needed.
  - Limitation: similar to L-P, adaptive sampling with SF is not possible.

Bayesian Monte Carlo

Introduction and Motivation

Representing a function using a smooth model

- Consider the evaluation of the integral

$$I = \int f(\mathbf{x})p(\mathbf{x})\,d\mathbf{x},\ \mathbf{x} \in \mathbb{R}^D,$$

  where $p(\mathbf{x})$ is analytically known and $f(\mathbf{x})$ is unknown before any sampling.

- The quality of the integral approximation depends on integrand smothness [BSSW12, MBR+13a].
- Recall some conclusions from the previous presentations:
  - Integration error depends on the rate of decay of the integrand's power spectrum (frequency view).
  - Discontinuities or lack of smoothness in the integrand impairs performance (QMC theory view).

Representing a function using a smooth model

$$I = \int f(\mathbf{x}) p(\mathbf{x})\, d\mathbf{x},\ \mathbf{x} \in \mathbb{R}^D,$$

- In our case:
  - $f(\mathbf{x}) = L_i(\boldsymbol{\omega})$ is the incident radiance, $\boldsymbol{\omega}$ being the incident direction.
  - $p(\mathbf{x}) = brdf \times cos(\theta)$, θ being the incident angle.
- Objective: substitute a smooth model $\tilde{f}(\mathbf{x})$ to the original $f(\mathbf{x})$ while keeping the integration error as small as possible.
- In the following, we show in a progressive way how this can be done:
  - Deterministic, Gaussian and Bayesian linear models.
  - *Non-parametric Bayesian approach* (Gaussian Process model).

Linear Basis Functions Model

- Let $\mathcal{D} = \{(\mathbf{x}_i, Y_i) \mid i = 1, \ldots, n\}$ be a sample set where $\mathbf{x} \in \mathbb{R}^D$ and $Y \in \mathbb{R}$.
  - $\mathbf{x}_i$ is a sample position.
  - Y_i is the observed function value at $f(\mathbf{x}_i)$.
- Let $\phi_j : \mathbb{R}^D \to \mathbb{R}$ be a set of smooth functions and $\mathbf{w} \in \mathbb{R}^F$ be a column vector of weights.
- We want to approximate the data $\mathcal{D}$ using a linear model

$$f(\mathbf{x}) \approx \mathbf{w}^t \Phi(\mathbf{x}),$$

  where $\Phi(\mathbf{x}) = [\phi_1(\mathbf{x}), \ldots, \phi_F(\mathbf{x})]^t$.
- Question: how to determine $\mathbf{w}$?

Deterministic Approach (Least Squares)

Model Parameterization

- We define a quadratic error criterion:

$$C(\mathbf{w}) = \sum_{i=1}^{n} \| Y_i - \mathbf{w}^t \Phi(\mathbf{x}_i) \|^2 = \| \mathbf{Y} - \mathbf{\Phi}^t \mathbf{w} \|^2$$

  where $\mathbf{\Phi} = [\Phi(\mathbf{x}_1), \ldots, \Phi(\mathbf{x}_n)]$ and $\mathbf{Y} = [Y_1, \ldots, Y_n]^t$.

- Objective: minimize $C(\mathbf{w})$, also called residual.

- Because we chose a quadratic error criterion $\hat{\mathbf{w}} = \arg\min_{\mathbf{w}} C(\mathbf{w})$ has a unique solution given by:

$$\hat{\mathbf{w}} = \left(\mathbf{\Phi}\mathbf{\Phi}^t\right)^{-1} \mathbf{\Phi}\mathbf{Y}$$

Gaussian Linear Model (Maximizing Likelihood)

Model Parameterization

- Assumption: Y_i contains an i.i.d. Gaussian noise such that:

$$Y_i = \mathbf{w}^t \Phi(\mathbf{x}_i) + \varepsilon_i, \ \varepsilon = \mathcal{N}(0, \sigma_n^2)$$

- ε_i quantifies the part of the observations not explained by the model (data accommodation).
- This assumption gives rise to the *likelihood* $L(\mathbf{w})$:

$$L(\mathbf{w}) = p(\mathbf{Y}|\mathbf{w}; \mathbf{X}) = \prod_{i=1}^{n} p(Y_i|\mathbf{w}; \mathbf{x_i})$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_n}} e^{-\frac{\|Y_i - \mathbf{w}^t \Phi(\mathbf{x}_i)\|^2}{2\sigma_n^2}} = \frac{1}{(2\pi\sigma_n)^{n/2}} e^{-\frac{C(\mathbf{w})}{2\sigma_n^2}}$$

- Goal: find the *maximum likelihood* (ML):

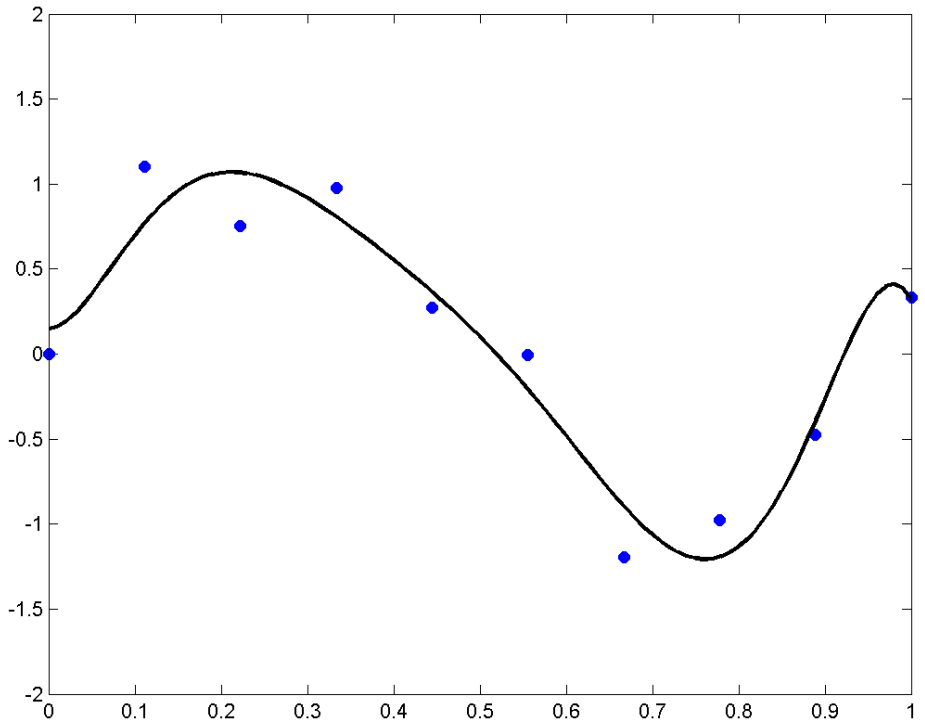
$$\hat{\mathbf{w}} = \arg\max_{\mathbf{w}} L(\mathbf{w}) = \arg\min_{\mathbf{w}} C(\mathbf{w})$$

Gaussian Linear Model (Maximizing Likelihood)

Predictions

- For a new input value $\mathbf{x}_*$ the model prediction Y_* is given by:

$$Y_* = \hat{\mathbf{w}}^t \Phi(\mathbf{x}_*)$$



Bayesian Linear Model

- Bayesian reasoning: all forms of uncertainty are modeled by probability.
  - $\mathbf{w}$ is unknown and thus uncertain (considered as random).
- We use the observations $\mathbf{Y}$ to make an inference about $\mathbf{w}$:
  - Put a prior probability on $\mathbf{w}$.
  - Apply the *Baye's rule*.
  - $\mathbf{w}$ is given by the *maximum a posteriori* (MAP) instead of the maximum likelihood.

$$p(\mathbf{w}|\mathbf{Y};\mathbf{X}) = \frac{p(\mathbf{Y}|\mathbf{w};\mathbf{X})\,p(\mathbf{w})}{p(\mathbf{Y};\mathbf{X})} \qquad posterior = \frac{likelihood \times prior}{marginal\ likelihood}$$

Bayesian Linear Model

Model Parameterization

- Let the prior $p(\mathbf{w})$ be a multivariate Gaussian distribution with mean 0 and covariance matrix Σ_p.

- The *posterior* distribution of the weights is then given by:

$$p(\mathbf{w}|\mathbf{Y};\mathbf{X}) = \frac{p(\mathbf{Y}|\mathbf{w};\mathbf{X})\,p(\mathbf{w})}{p(\mathbf{Y};\mathbf{X})} = \mathcal{N}(\bar{\mathbf{w}} = \sigma_n^{-2} A^{-1} \mathbf{\Phi Y}, A^{-1})$$

where $A = \sigma_n^{-2}\mathbf{\Phi}\mathbf{\Phi}^t + \Sigma_p^{-1}$.

Bayesian Linear Regression

Predictions

- Since the *maximum a posteriori* of a Gaussian distribution is also its expected value, we have:

$$\arg\max_{\mathbf{w}} p(\mathbf{w}|\mathbf{Y};\mathbf{X}) = \bar{\mathbf{w}}$$

- For a new input value $\mathbf{x}_*$ the model prediction Y_* has a Gaussian distribution [Bis06]:

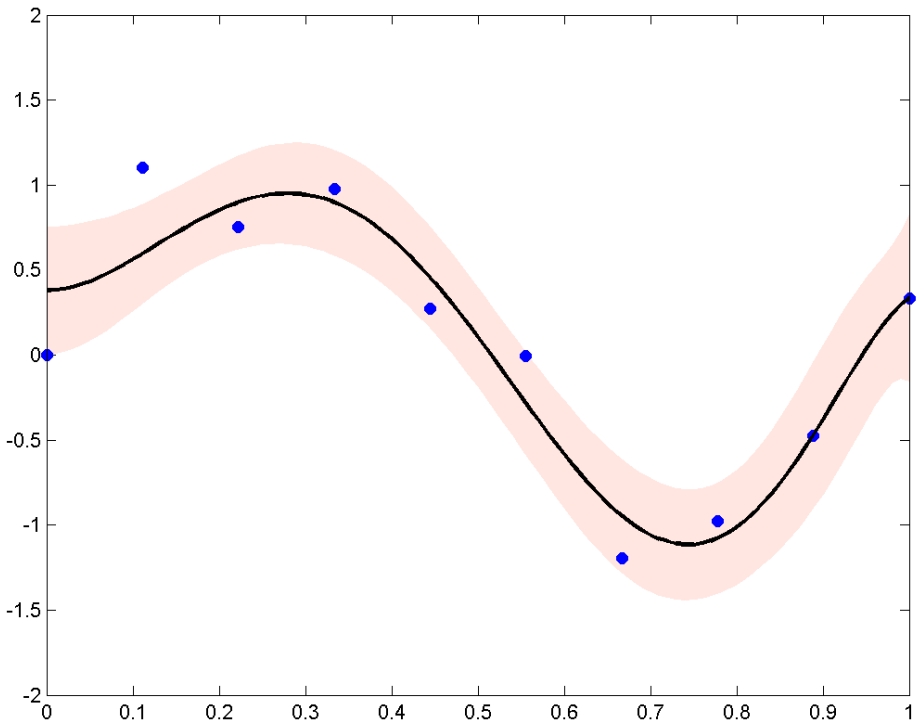
$$\mathcal{N}(\bar{\mathbf{w}}^t\Phi(\mathbf{x}_*), \Phi(\mathbf{x}_*)^t A^{-1}\Phi(\mathbf{x}_*))$$

- The most probable value of Y_* is given by $\bar{\mathbf{w}}^t\Phi(\mathbf{x}_*)$.

Bayesian Linear Regression

Predictions

$$Y_* = \bar{\mathbf{w}}^t \Phi(\mathbf{x}_*)$$



Brief Summary

- The Bayesian approach uses more information than the classic probabilistic approach by resorting to a prior $p(\mathbf{w})$.
- Nevertheless, both approaches share the same limitation:
  - Need to choose a family of basis functions.
  - Need to specify the number of basis functions (length of $\mathbf{w}$).
- In the following, we will present a Bayesian non-parametric approach to the problem of regression (*Bayesian Regression*).
- We will also show how this approach can be used for integral estimation (*Bayesian Monte Carlo*).

Theoretical Background

Bayesian Regression

- Goal: given a set of samples, approximate the value of the unknown function $f(\mathbf{x})$.
- Since the value of $f(\mathbf{x})$ is uncertain, it is modeled through probabilities using a Gaussian process (GP).
  - No need for basis functions
  - Only a mean function and a covariance function are specified.

Gaussian Process

- GP is a collection of random variables, any finite number of which has a joint Gaussian distribution [RW06].
- A GP is completely defined by its mean function $\bar{f}(\mathbf{x})$ and its covariance function $k(\mathbf{x}, \mathbf{x}')$, which must be positive definite:

$$\begin{aligned} \bar{f}(\mathbf{x}) &= \mathrm{E}[f(\mathbf{x})] \\ k(\mathbf{x}, \mathbf{x}') &= \mathrm{E}[(f(\mathbf{x}) - \bar{f}(\mathbf{x}))(f(\mathbf{x}') - \bar{f}(\mathbf{x}'))] \end{aligned}$$

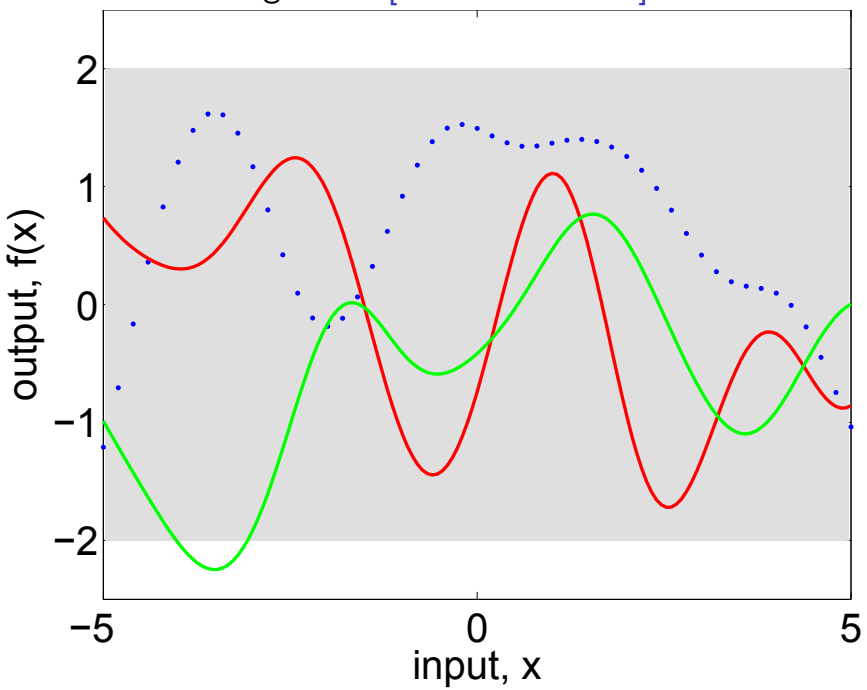
- We denote a GP by:

$$f(\mathbf{x}) \sim \mathcal{GP}[\bar{f}(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')]$$

The Prior Gaussian Process

- Covariance function: defines the smoothness and the variance of the GP.
- Characterized by a lengthscale l and a variance σ_f.
- l and σ_f are hyperparameters of the GP model.

Image from [Rasmussen 2006]

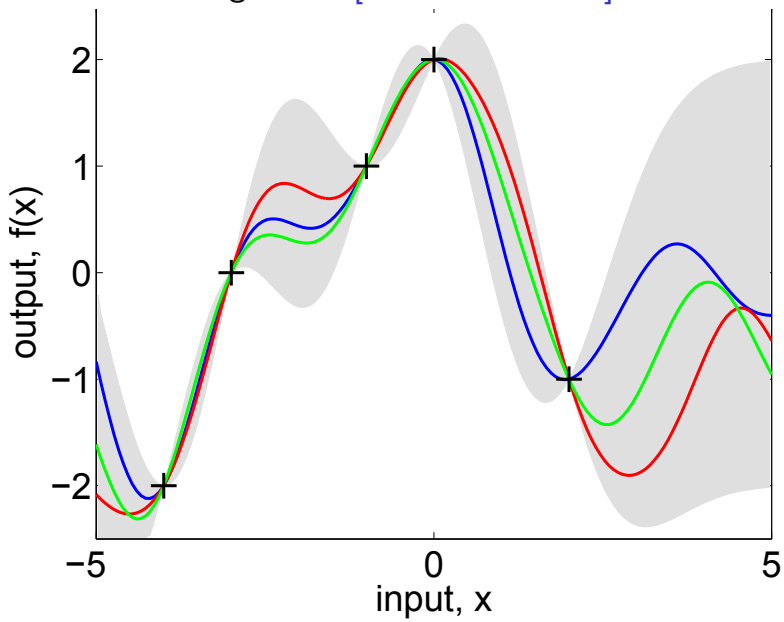


- Large lengthscale implies smooth realizations of the GP.
- Large variance implies large uncertainty zone (gray).

The Posterior Gaussian Process

- The posterior GP results from refining the prior using a set of observations $\mathcal{D}$.
- Uncertainty region (gray) smaller than in the prior.

Image from [Rasmussen 2006]



$$\mathcal{D} = \{(\mathbf{x}_i, Y_i) \,|\, i = 1, \ldots, n\} \quad \text{with} \quad Y_i = f(\mathbf{x}_i) + \varepsilon_i,$$

- $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$ accommodates observations to smooth model.
- Besides l and σ_f, σ_n is another hyperparameter of the model.

$$f(\mathbf{x}|\mathcal{D}) \sim \mathcal{GP}\left[\mathrm{E}\left[f(\mathbf{x})\,|\mathcal{D}\right], \mathrm{cov}\left(f(\mathbf{x}), f(\mathbf{x}')\,|\mathcal{D}\right)\right]$$

Bayesian Regression Equations

- The Bayesian Regression equations result from conditioning the prior GP to the observations $\mathcal{D}$ [RW06].

$$\begin{aligned} \tilde{f}(\mathbf{x}) = \mathrm{E}[f(\mathbf{x})|\mathcal{D}] &= \bar{f}(\mathbf{x}) + \mathbf{k}(\mathbf{x})^t Q^{-1}(\mathbf{Y} - \bar{\mathbf{F}}) \\ \mathrm{cov}[f(\mathbf{x}), f(\mathbf{x}')|\mathcal{D}] &= k(\mathbf{x}, \mathbf{x}') - \mathbf{k}(\mathbf{x})^t Q^{-1} \mathbf{k}(\mathbf{x}') \end{aligned}$$

with:

$$\begin{aligned} \mathbf{k}(\mathbf{x}) &= (k(\mathbf{x}_1, \mathbf{x}), \ldots, k(\mathbf{x}_n, \mathbf{x}))^t \\ K_{i,j} &= k(\mathbf{x}_i, \mathbf{x}_j) \qquad \text{with } (i,j) \in [1,n]^2 \\ Q &= (K + \sigma_n^2 I_n) \\ \mathbf{Y} &= (Y_1, \ldots, Y_n)^t \\ \bar{\mathbf{F}} &= \left(\bar{f}(\mathbf{x}_1), \ldots, \bar{f}(\mathbf{x}_n)\right)^t \end{aligned}$$

Bayesian Monte Carlo

- The Bayesian quadrature equations are given by [RG02]:

$$\begin{aligned}\hat{I}_{BMC} = \mathrm{E}(I|\mathcal{D}) &= \int \tilde{f}(\mathbf{x})p(\mathbf{x})\,d\mathbf{x} \\ &= \bar{I} + \mathbf{c}^t(\mathbf{Y} - \bar{\mathbf{F}}) \\ &= \bar{I} + \mathbf{z}^t Q^{-1}(\mathbf{Y} - \bar{\mathbf{F}})\end{aligned}$$

where $\tilde{f}(\mathbf{x})$ is a smooth function resulting from a *Bayesian Regression*, and:

$$\bar{I} = \int \bar{f}(\mathbf{x})p(\mathbf{x})d\mathbf{x} \qquad \mathbf{z} = \int \mathbf{k}(\mathbf{x})p(\mathbf{x})d\mathbf{x}$$

$$\bar{\mathbf{F}} = \left(\bar{f}(\mathbf{x}_1), \ldots, \bar{f}(\mathbf{x}_n)\right) \qquad Q \text{ is the covariance matrix}$$

- In fact, σ_n and σ_f are replaced by the ratio $\sigma'_n = \frac{\sigma_n}{\sigma_f}$.

Variance of the Interal Estimate

- The variance of the BMC estimate is given by [RG02]:

$$\mathrm{Var}(I|\mathcal{D}) = \bar{V} - \mathbf{z}^t Q^{-1} \mathbf{z},$$

with

$$\bar{V} = \iint k(\mathbf{x}, \mathbf{x}') p(\mathbf{x}) p(\mathbf{x}') \; d\mathbf{x} d\mathbf{x}'$$

- $\mathrm{Var}(I|\mathcal{D})$ depends on the location $\mathbf{x}$ of the samples and on the covariance function k.
- We can thus previously select the $\{\mathbf{x}_i | i = 1, \ldots, n\}$ that minimizes $\mathrm{Var}(I|\mathcal{D})$. Such a set is called *optimal samples set*.

Theoretical Background: Summary

- BMC allows acting on the three factors which determine the quality of the estimate.
- Pdf-free: any sample distribution can be used.
  - The prior knowledge can be used to determine an optimal samples set.
- Performs a careful samples weighting.
  - Exploits all the information available: prior probabilistic knowledge and samples location.
- Uses a smooth model of the integrand.
  - Leads to better estimates.

BMC for Rendering (Our Approach)

[MBR+13b, Mar13]

Choosing the model

Mean Function

- Recall:

$$f(\mathbf{x}) \sim \mathcal{GP}[\bar{f}(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')]$$

- The first step to apply BMC is to specify $\bar{f}$ and k for the prior model.
- The mean function contains our belief about the value of f
  - It can be seen as a rough a priori approximation of f which will be refined using the observations.
- In our approach we consider $\bar{f}$ constant [BBL+09, MBR+13b].
  - Simplifies the quadrature computation (details later).

Choosing the Model

Covariance Function

- The covariance function determines the smoothness of the GP model.
- It characterizes the power spectrum of the GP:
  - Smooth covariance $\Leftrightarrow$ higher rate of decay of the power spectrum
  - Exponential decay in case of a Gaussian covariance function.
- We chose a Gaussian covariance function (squared exponential):
  - Smooth (infinitely derivable).
  - Leads to a smooth model of f.
- This choice simplifies the quadrature computation (details later).

BMC Algorithm Overview

1. Learn the hyperparameters associated with the covariance function
2. Select the optimal samples set $\{\mathbf{x}_i\}$
3. Compute and invert the covariance matrix Q
4. Compute the vector $\mathbf{z} = \int \mathbf{k}(\mathbf{x})p(\mathbf{x})\ d\mathcal{S}(\mathbf{x})$
5. Collect the observation $\mathbf{Y}_i$ for each sampling position $\mathbf{x}_i$
6. Compute the prior mean value vector $\bar{\mathbf{F}}$
7. Compute the posterior estimate $\mathrm{E}(I|\mathcal{D}) = \bar{I} + \mathbf{z}^t Q^{-1}(\mathbf{Y} - \bar{\mathbf{F}})$

1. Learn the Hyperparameters

Problem Statement

- To compute a BMC estimate the hyperparameters l, σ_f and σ_n of the covariance function k must be specified:

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp(-\frac{|\mathbf{x} - \mathbf{x}'|^2}{l^2}) + \sigma_n^2 \delta(\mathbf{x}, \mathbf{x}'),$$

  where $\delta(\mathbf{x}, \mathbf{x}')$ is the Kronecker delta.

- The hyperparameters can be learned from the observations.

- Problem: learning the hyperparameters for each integration would be too costly.

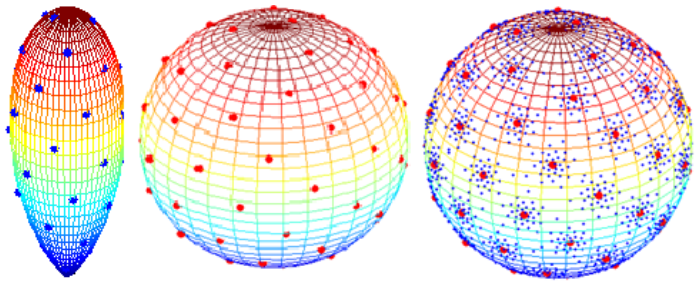
1. Learn the Hyperparameters

Tackling the Problem

- As we shall see later, the hyperparameters value mainly depends on the BRDF shape
  - Shininess in the case of a Phong model.
- We can thus learn the hyperparameters off-line for different BRDF shininesses. Example:
  - Define a set $\mathbf{M}$ of materials with different shininesses.
  - Learn the hyperparameters for each $m \in \mathbf{M}$, when illuminated by different environment maps.
- To render an object with a given shininess we use interpolated values of the previously learned hyperparameters.

1. Learn the Hyperparameters

Off-line Learning Algorithm



- The covariance is statistically evaluated across the scene using a point set distributed on the current BRDF lobe.
  - Statistically generate a covariance matrix $\tilde{Q}$.
- The hyperparameters (l, σ') are then extracted by fitting the covariance function $k(\mathbf{x}, \mathbf{x}'; l, \sigma')$ to $\tilde{Q}$ [BBL+09, MBR+13b, Mar13].

2. Select the Optimal Samples Set

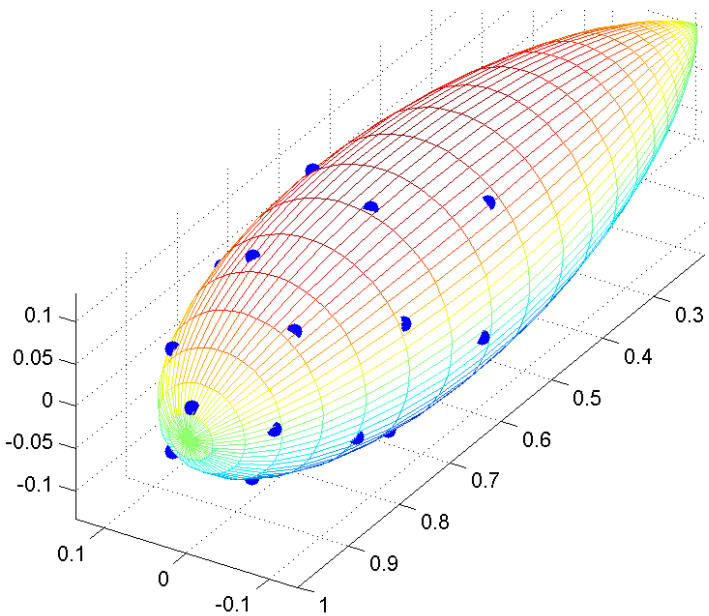
- Recall that $\mathrm{Var}(I|\mathcal{D}) = \bar{V} - \mathbf{z}^t Q^{-1} \mathbf{z}$
  - Depends on the samples position $\{\mathbf{x}_i\}$;
  - And on the hyperparameters (l, σ') of the covariance function $k(\mathbf{x}, \mathbf{x}'; l, \sigma')$.
- Since the hyperparameters are already known (see step 1.) $\mathrm{Var}(I|\mathcal{D})$ becomes a function of $\{\mathbf{x}_i\}$ only.
- We can thus determine the optimal samples set by computing:

$$\arg\min_{\{\mathbf{x}_i\}} \mathrm{Var}(I|\mathcal{D})$$

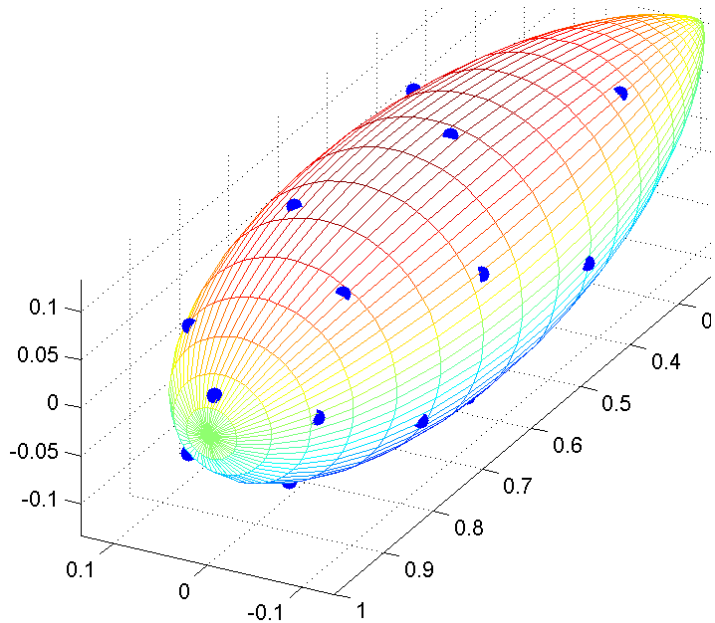
- Done for each BRDF in a preprocessing step.

2. Select the Optimal Samples Set

- The result is strongly dependent on the lengthscale l of the covariance function k.
  - Stronger covariance (large l) $\rightarrow$ Larger distance between samples (each sample brings more information about $f(\mathbf{x})$)



(a) $l = 0.05$



(b) $l = 0.17$

3. Compute and Invert Q

- The covariance matrix Q accounts for the relative position between the samples:
  - Nearby samples $\rightarrow$ High correlation $\rightarrow$ Redundant information
  - Highly correlated samples have a lower weight in the quadrature due to the matrix Q inversion.
- Q only depends on the hyperparameters of the covariance function (l, σ') and on the samples position $\{\mathbf{x}_i\}$.
  - The hyperparameters are already known (see step 1.).
  - So are the samples position (see step 2.).
- We can thus also precompute and invert Q.

4. Compute the $\mathbf{z}$ Vector

- Recall that $\mathbf{z}$ is a vector given by

$$\mathbf{z} = \begin{bmatrix} z(\mathbf{x}_1) \\ z(\mathbf{x}_2) \\ \vdots \\ z(\mathbf{x}_n) \end{bmatrix} = \begin{bmatrix} \int k(\mathbf{x}_1, \mathbf{x})\, p(\mathbf{x})\, d\mathcal{S}(\mathbf{x}) \\ \int k(\mathbf{x}_2, \mathbf{x})\, p(\mathbf{x})\, d\mathcal{S}(\mathbf{x}) \\ \vdots \\ \int k(\mathbf{x}_n, \mathbf{x})\, p(\mathbf{x})\, d\mathcal{S}(\mathbf{x}) \end{bmatrix}$$

- In the diffuse BRDF case $\mathbf{z}$ is precomputed [BBL+09].
- For non-diffuse BRDFs $\mathbf{z}$ must be evaluated during rendering.
  - $\rho(\mathbf{x})$ becomes dependent on the viewing direction.
  - Resort to spherical Gaussian functions (SGF) to model both the BRDF ρ and the covariance function k [MBR+13b, Mar13].

4. Compute the **z** Vector

- Each element z_i of $\mathbf{z}$ becomes:

$$z_i = \int G_k(\mathbf{x}_i - \mathbf{x}; \sigma'^2, l)\ G_\rho(\mathbf{r} - \mathbf{x}; k_s, w)\ d\mathcal{S}(\mathbf{x})$$

- Since the product of two SGFs is a SGF, each element z_j of $\mathbf{z}$ is reduced to a *spherical Gaussian integral* (SGI).

$$z_i = \int G(\mathbf{x}_m - \mathbf{x}; c_m, l_m)\ d\mathcal{S}(\mathbf{x})$$

- A SGI varies smoothly, its value can be easily tabulated.
- Computing z_i now amounts to querying a 2D look-up table.
  - Details in [MBR+13b, Mar13].

5. Collect the observations

- The sampling directions are determined by the optimal samples set $\{\mathbf{x}_i\}$.
- Compute the vector of observations

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

6. Compute the Prior Mean Value

- $\bar{f}(\mathbf{x})$ is used in the following terms of the Bayesian quadrature:

$$\bar{\mathbf{F}} = \left[\bar{f}(\mathbf{x}_1), \ldots, \bar{f}(\mathbf{x}_n)\right] \qquad \bar{I} = \int \bar{f}(\mathbf{x}) p(\mathbf{x}) \, d\mathcal{S}(\mathbf{x})$$

- Recall: we consider a model with a constant mean function $\bar{f}(\mathbf{x}) = \bar{f}$ (see slides Choosing the model).
- Such a choice simplifies the computation of $\bar{\mathbf{F}}$ and $\bar{I}$.

$$\bar{I} = \bar{f} \int G_\rho(\mathbf{r} - \mathbf{x}; k_s, w) \, d\mathcal{S}(\mathbf{x}) = \bar{f} \times SGI$$

- Problem: how to choose the value of the constant mean function? $\bar{f} = ?$

6. Compute the Prior Mean Value

- In [RW06] the authors suggest to express $\bar{f}(\mathbf{x})$ as a weighted sum of F basis functions $h_j(\mathbf{x})$:

$$\bar{f}(\mathbf{x}) = \boldsymbol{\beta}^t \mathbf{h}(\mathbf{x}), \quad \text{where} \quad \mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}), \ldots, h_F(\mathbf{x})]^t$$

- The weights $\boldsymbol{\beta}$ are determined based on the observations:

$$\boldsymbol{\beta} = \frac{HQ'^{-1}}{HQ'^{-1}H^t}\mathbf{Y}, \quad \text{where} \quad H = [\mathbf{h}(\mathbf{x}_n), \ldots, \mathbf{h}(\mathbf{x}_n)]$$

- Such an approach allows *local adaptation* of the prior model.
- We use a single basis function $h_1(\mathbf{x}) = 1$ in which case:
  - $H = [1, \ldots, 1]$ and $\bar{f}(\mathbf{x}) = \boldsymbol{\beta}$ is constant.

7. Compute the BMC Estimate

- All the terms of the BMC quadrature are known.
- We only have to apply the BMC equation.

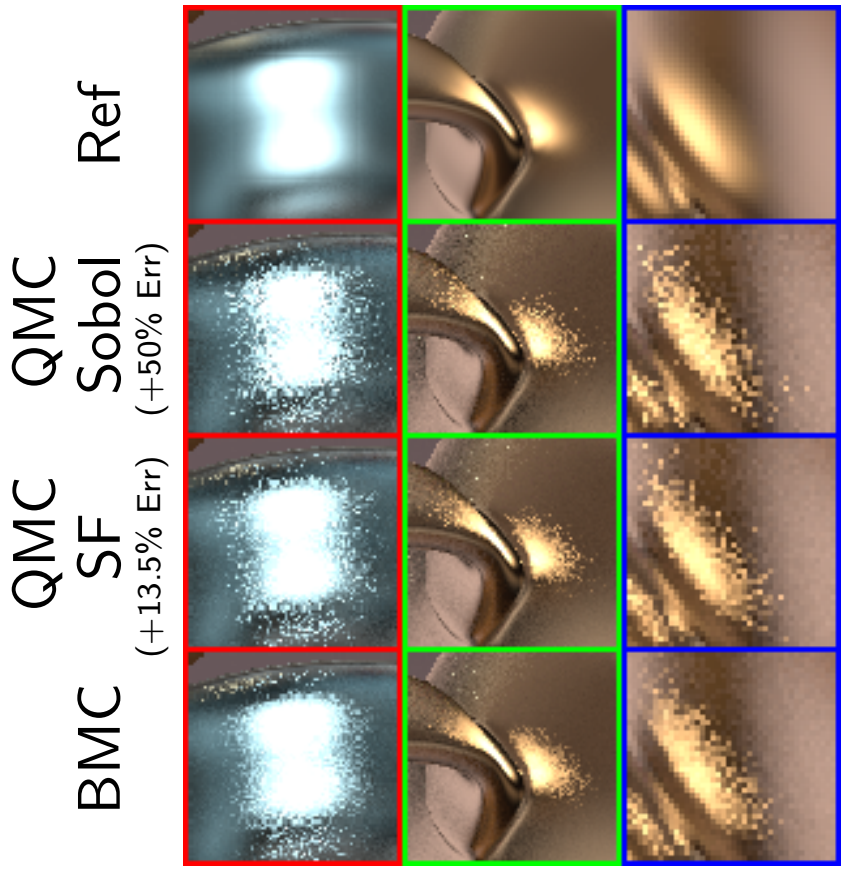
$$\hat{I}_{BMC} = \bar{I} + \mathbf{z}^t Q^{-1}(\mathbf{Y} - \bar{\mathbf{F}})$$

- In the particular case in which $\bar{f}$ is expressed based on the samples, we can write:

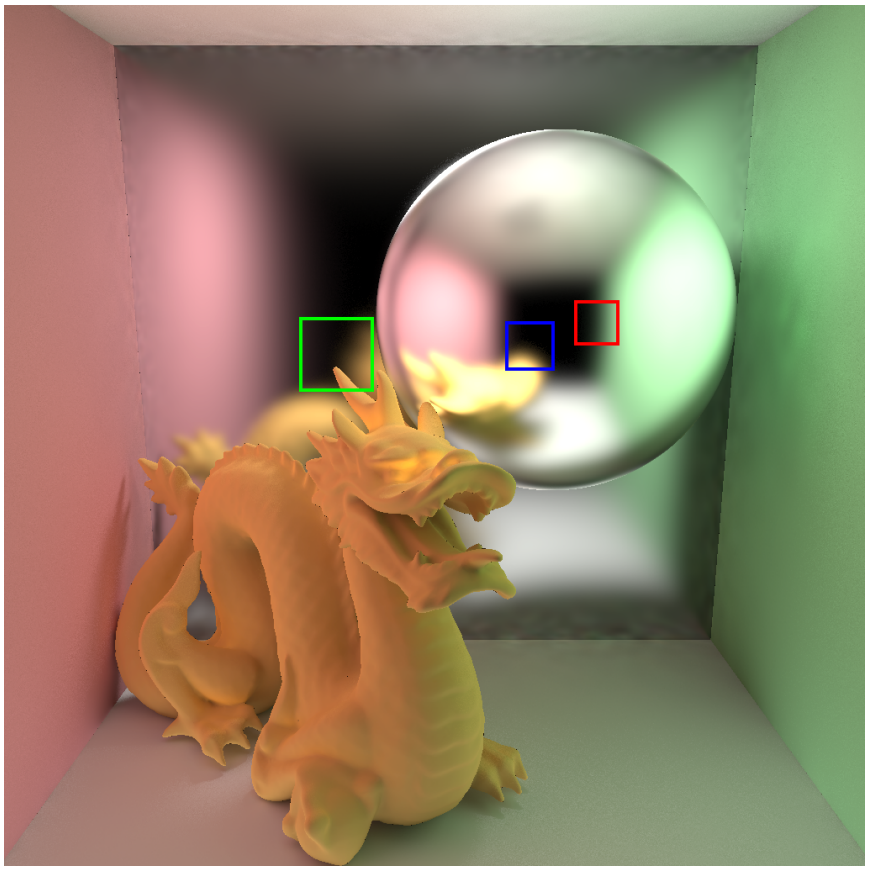
$$\hat{I}_{BMC} = \mathbf{w}^t \mathbf{Y} = \sum_{i=1}^{n} w_i Y_i$$

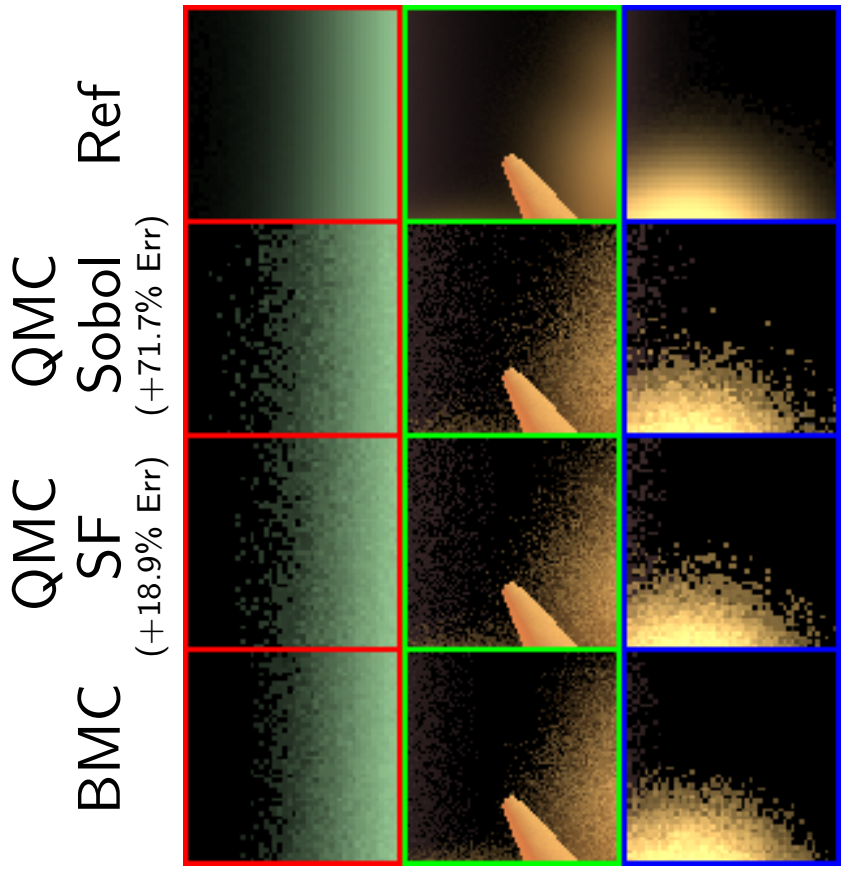
Results

Visual Comparison

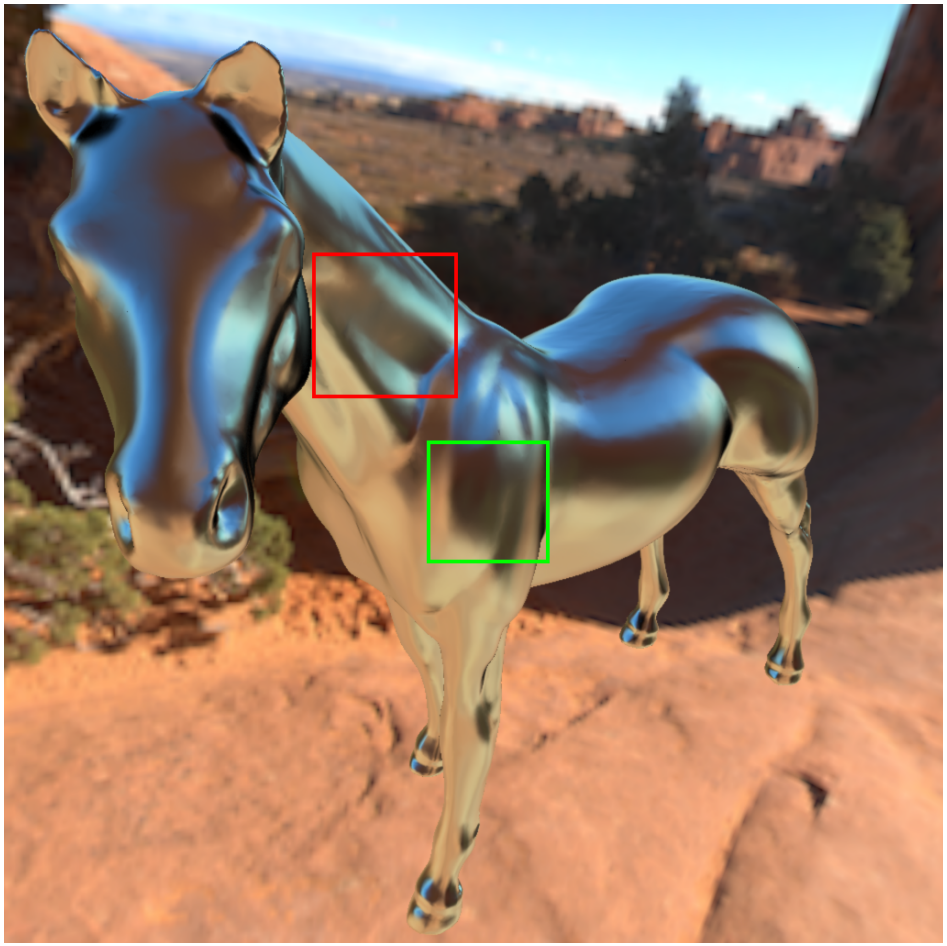


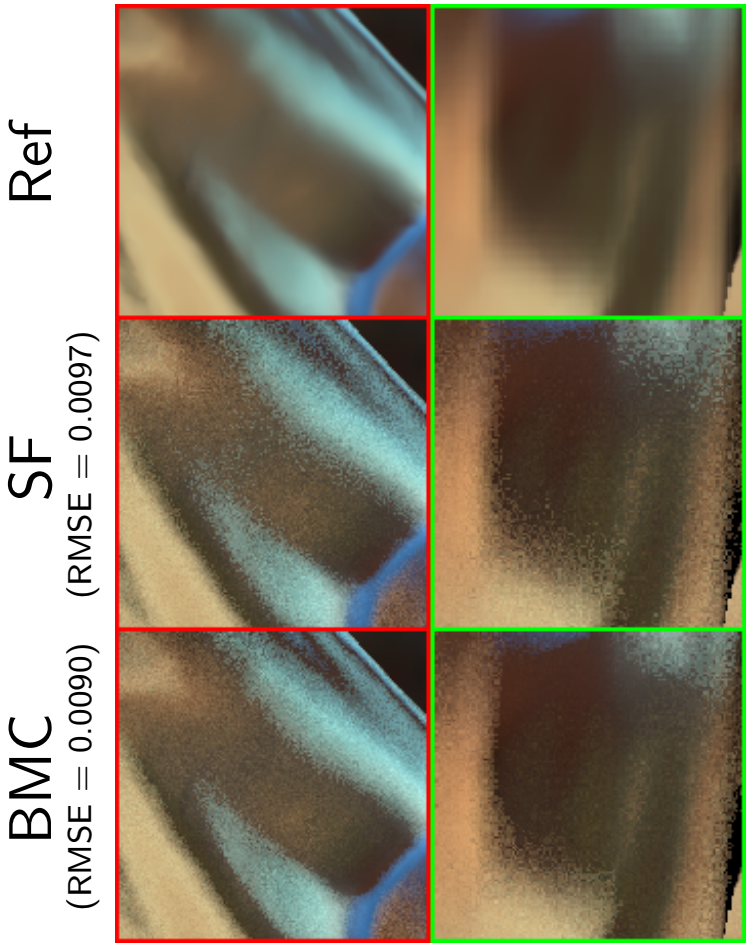
Visual Comparison



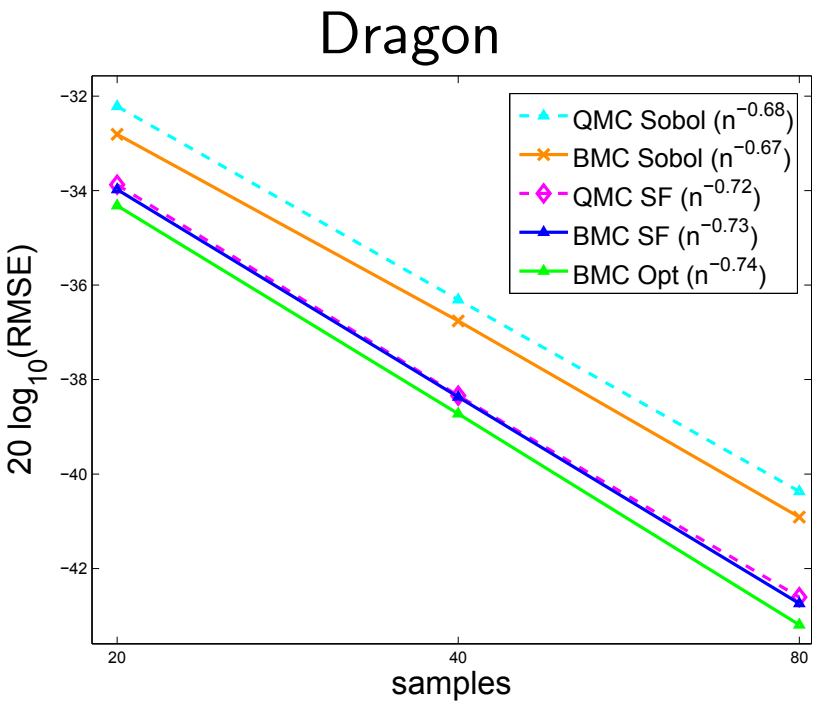


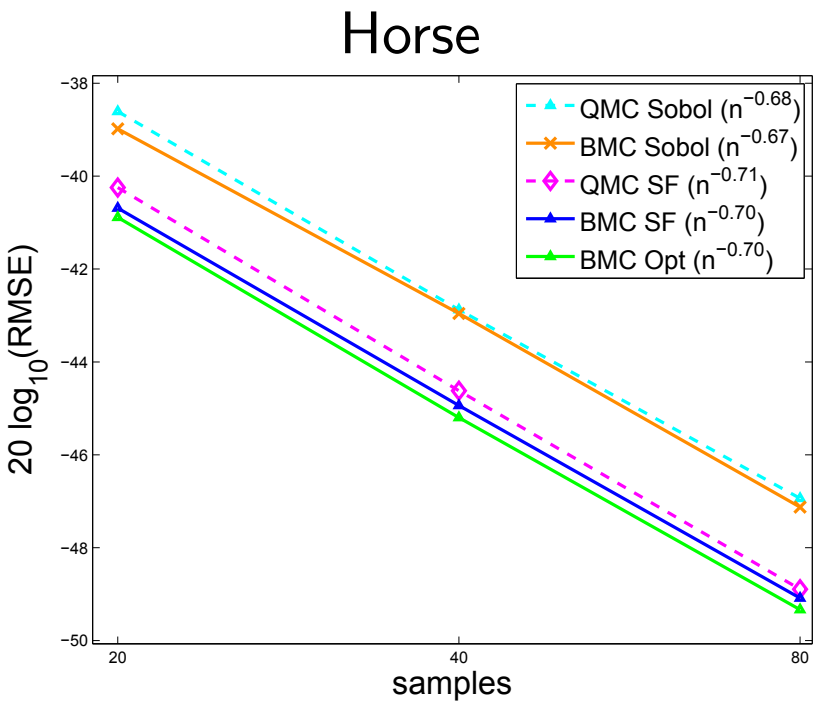
Visual Comparison with QMC Spherical Fibonacci





RMSE Comparison with QMC





Efficiency Comparison with QMC

	Dragon		Horse	
	RMSE	Same quality rays needed	RMSE	Same quality rays needed
QMC Sobol	+38.4%	+61.6%(≈ 129)	+31.8%	+58.5%(≈ 127)
BMC Sobol	+30.0%	+48.6%(≈ 119)	+28.9%	+49.9%(≈ 120)
QMC SF	+7.0%	+9.8%(≈ 88)	+5.2%	+7.1%(≈ 86)
BMC SF	+5.3%	+7.2%(≈ 86)	+2.9%	+4.0%(≈ 83)

- Efficiency of other methods relative to that of BMC using 80 samples per spherical integration with an optimal distribution.

Overall Conclusion

Conclusion I

- The most important characteristic of an estimation method is its capacity to incorporate existing information.
- CMC and QMC only incorporate deterministic knowledge (but no probabilistic knowledge).
- Examples:
  - Information regarding incident radiance for product sampling (CMC).
  - Morph a samples set to follow the BRDF shape (QMC).
  - Continuity assumption regarding the integrand (QMC).

Conclusion II

- BMC has proved to be the most flexible regarding knowledge introduction.
  - Deterministic knowledge:
    - Through the known part of the integrand $p(\mathbf{x})$.
  - Probabilistic knowledge:
    - Through a probabilistic model of unknown part of the integrand.
    - Covariance between the samples.
    - Mean function $\bar{f}(\mathbf{x})$, an approximation of the unknown function $f(\mathbf{x})$.

Future Research directions

- Yet many research directions to be explored such as:
  - Local adaptation of the hyperparameters.
  - Application to problems of higher dimensionality.

Questions

References I

Congpei AN, Xiaojun Chen, Ian H. Sloan, and Robert S. Womersley, *Well conditioned spherical designs for integration and interpolation on the two-sphere*, SIAM J. Numer. Anal. **48** (2010), no. 6, 2135–2157.

Jonathan Brouillat, Christian Bouville, Brad Loos, Charles Hansen, and Kadi Bouatouch, *A bayesian monte carlo approach to global illumination*, Computer Graphics Forum **28** (2009), no. 8, 2315–2329.

Christopher M. Bishop, *Pattern recognition and machine learning (information science and statistics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

References II

J. S. Brauchart, E. B. Saff, I. H. Sloan, and R. S. Womersley, *QMC designs: optimal order Quasi Monte Carlo Integration schemes on the sphere*, ArXiv e-prints (2012), http://arxiv.org/abs/1208.3267.

Mark Colbert and Jaroslav Křivánek, *Real-time shading with filtered importance sampling*, ACM SIGGRAPH 2007 sketches, ACM, 2007, p. 71.

P. Delsarte, J.M. Goethals, and J.J. Seidel, *Spherical codes and designs*, Geometriae Dedicata **6** (1977), no. 3, 363–388 (English).

Philip Dutré, *Global illumination compendium*, 2003, http://people.cs.kuleuven.be/~philip.dutre/GI/TotalCompendium.pdf.

References III

J H Hannay and J F Nye, *Fibonacci numerical integration on a sphere*, Journal of Physics A: Mathematical and General **37** (2004), no. 48, 11591.

James T. Kajiya, *The rendering equation*, SIGGRAPH Comput. Graph. **20** (1986), no. 4, 143–150.

Hongwei Li, Li-Yi Wei, Pedro V. Sander, and Chi-Wing Fu, *Anisotropic blue noise sampling*, ACM Trans. Graph. **29** (2010), no. 6, 167:1–167:12.

Ricardo Marques, *Bayesian and Quasi Monte Carlo Spherical Integration for Global Illumination*, Ph.d. thesis, Université de Rennes 1, October 2013.

References IV

R. Marques, C. Bouville, M. Ribardière, L. P. Santos, and K. Bouatouch, *Spherical fibonacci point sets for illumination integrals*, Computer Graphics Forum **32** (2013), no. 8, 134–143.

_____, *A spherical gaussian framework for bayesian monte carlo rendering of glossy surfaces*, IEEE Transactions on Visualization and Computer Graphics **19** (2013), no. 10, 1619–1632.

H. Niederreiter, *Random number generation and quasi-monte carlo methods*, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, 1992.

References V

Matt Pharr and Greg Humphreys, *Physically based rendering, second edition: From theory to implementation*, 2nd ed., Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2010.

Carl Edward Rasmussen and Zoubin Ghahramani, *Bayesian monte carlo*, Neural Information Processing Systems, MIT Press, 2002, pp. 489–496.

C. E. Rasmussen and C. K. I. Williams, *Gaussian process for machine learning*, MIT Press, 2006.

Richard Swinbank and R. James Purser, *Fibonacci grids: A novel approach to global modelling*, Quarterly Journal of the Royal Meteorological Society **132** (2006), no. 619, 1769–1793.

References VI

E. Saff and A. Kuijlaars, *Distributing many points on a sphere*, The Mathematical Intelligencer **19** (1997), 5–11.

Sobolev spaces, Theoretical Numerical Analysis, Texts in Applied Mathematics, vol. 39, Springer New York, 2005, pp. 273–322.