

Numerical methods for random PDEs

1 MC under exact sampling

1.2 Problem

We are interested in calculating the following quantity

$$g = \int_{[0,1]^N} f(x) dx$$

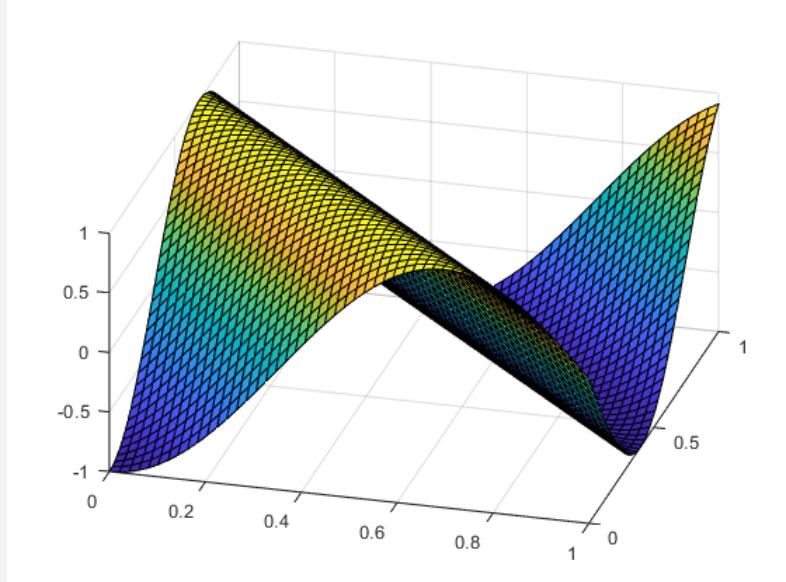
for some given function f . The functions are listed below. We added a plot of the 2D version of the corresponding function ($N = 2$).

1) Oscillator function

$$f(x) = \cos\left(2\pi w_1 + \sum_{n=1}^N c_n x_n\right)$$

with $c_n = \frac{9}{N}$, $w_n = \frac{1}{2}$ and exact solution

$$I_{ex} = \mathcal{R}\left(e^{i2\pi w_1} \prod_{n=1}^N \frac{1}{ic_n} (e^{ic_n} - 1)\right).$$

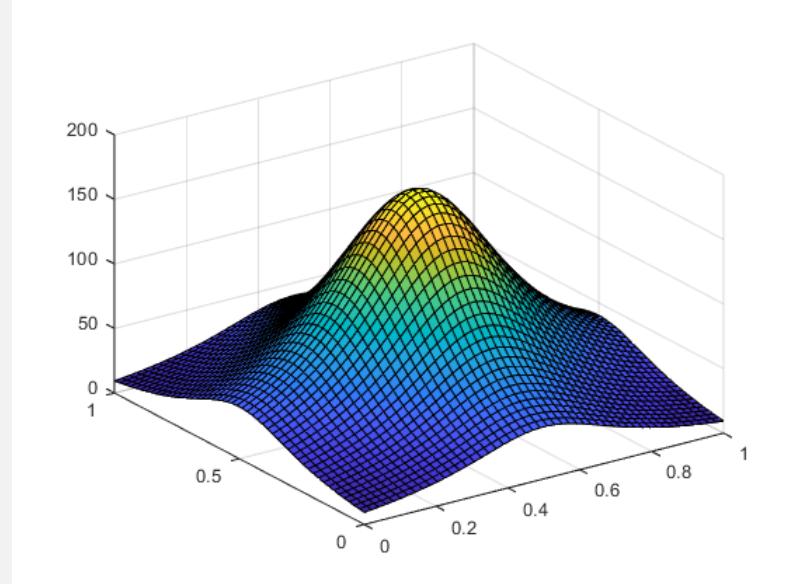


2) Product peak function

$$f(x) = \prod_{n=1}^N (c_n^{-2} + (x_n - w_n)^2)^{-1}$$

with $c_n = \frac{7.25}{N}$, $w_n = \frac{1}{2}$ and exact solution

$$I_{ex} = \prod_{n=1}^N c_n (\arctan(c_n(1 - w_n)) + \arctan(c_n w_n))$$

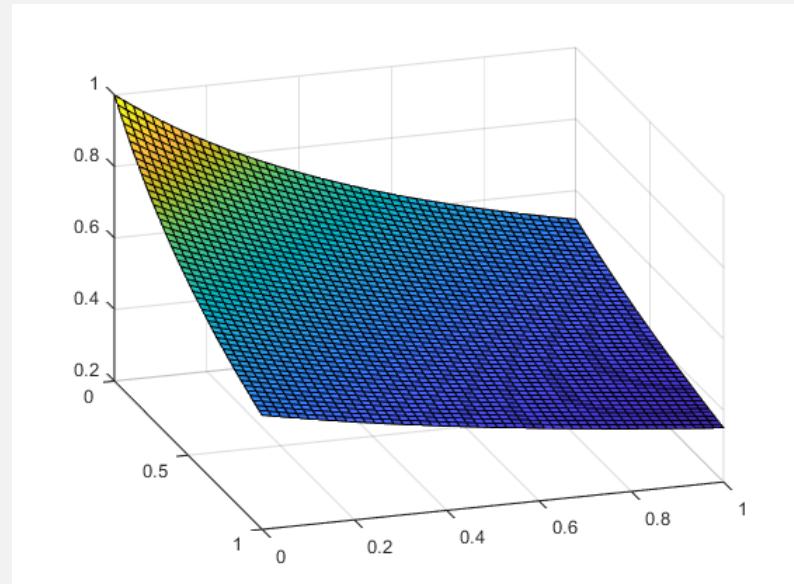


3) Corner peak function

$$f(x) = \left(1 + \sum_{n=1}^N c_n x_n\right)^{-N+1}$$

with $c_n = \frac{1.85}{N}$. We will use the reference solutions for $N = 2$ and $N = 20$

$$\begin{aligned} I_{\text{ref}}^{(2)} &= 0.541678226218786, \\ I_{\text{ref}}^{(20)} &= 0.004004658919014. \end{aligned}$$

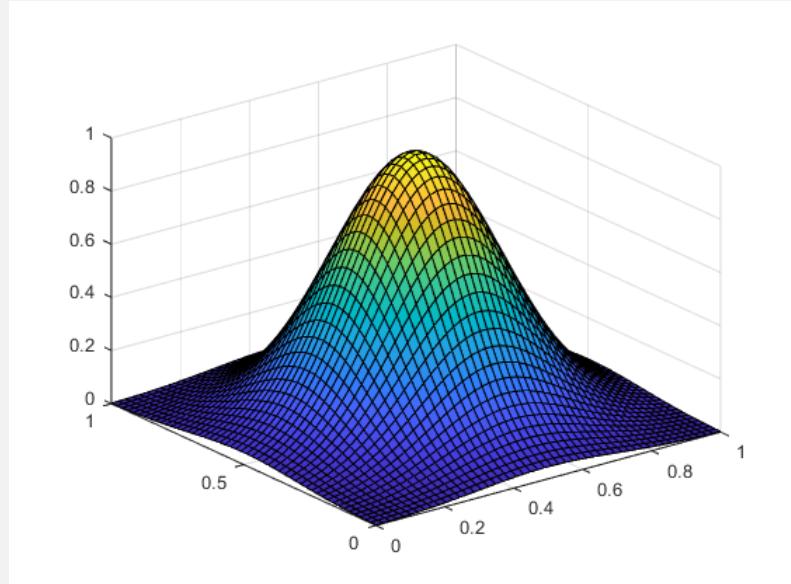


4) Gaussian function

$$f(x) = \exp\left(-\sum_{n=1}^N c_n^2 (x_n - w_n)^2\right)$$

with $c_n = \frac{7.03}{N}$, $w_n = \frac{1}{2}$ and exact solution

$$I_{ex} = \prod_{n=1}^N \frac{\sqrt{\pi}}{2c_n} (\operatorname{erf}(c_n(1-w_n)) + \operatorname{erf}(c_n w_n))$$

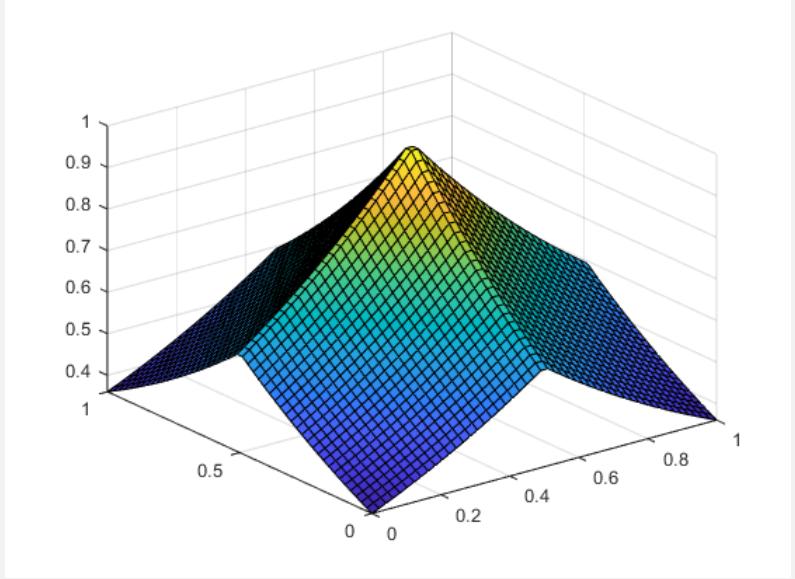


5) Continuous function

$$f(x) = \exp\left(-\sum_{n=1}^N c_n |x_n - w_n|\right)$$

with $c_n = \frac{2.04}{N}$, $w_n = \frac{1}{2}$ and exact solution

$$I_{ex} = \prod_{n=1}^N \frac{1}{c_n} \left(2 - e^{-c_n w_n} - e^{-c_n(1-w_n)}\right)$$

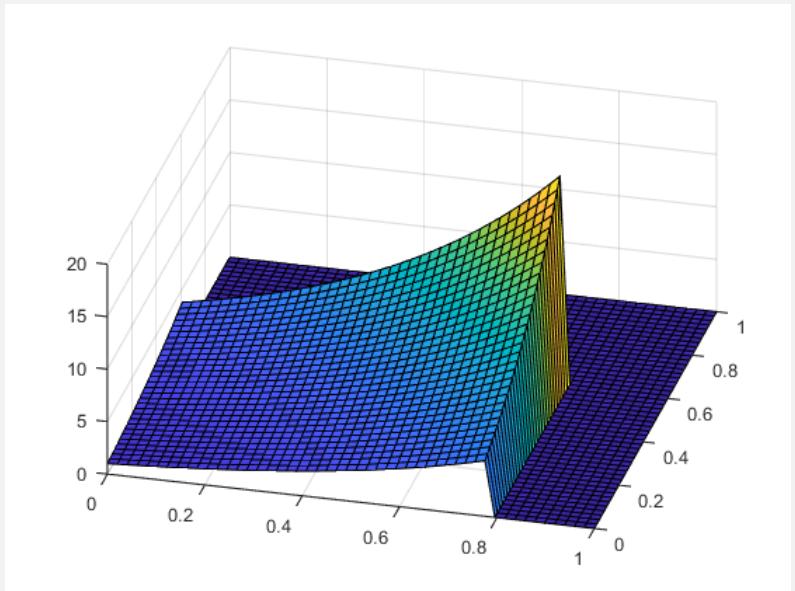


6) Discontinuous function

$$f(x) = 0, \quad \text{if } x_1 > w_1 \text{ or } x_2 > w_2, \quad f(x) = \exp\left(\sum_{n=1}^N c_n w_n\right), \quad \text{otherwise}$$

with $c_n = \frac{4.3}{N}$, $w_1 = \frac{\pi}{4}$, $w_2 = \frac{\pi}{5}$ and exact solution

$$I_{ex} = \frac{1}{\prod_{n=1}^N c_n} (e^{c_1 w_1} - 1)(e^{c_2 w_2} - 1) \prod_{n=1}^N (e^{c_n} - 1)$$



1.4 Numerical tasks

First we will estimate the error using the **Central Limit Theorem**, i.e. we will use the fact that for $\varepsilon_M := \sum_{i=1}^M \frac{f(x_i)}{M} - \int_{[0,1]^N} f(x) dx$ the CLT yields

$$\sqrt{M} \varepsilon_M \xrightarrow{} \sigma \mathcal{N}(0, 1).$$

And thus we get an asymptotic error estimate for a given confidence level constraint C_α , $0 < \alpha \ll 1$ and large M

$$\mathbb{P}\left(|\varepsilon_M| \leq \frac{C_\alpha \sigma}{\sqrt{M}}\right) \approx 1 - \alpha.$$

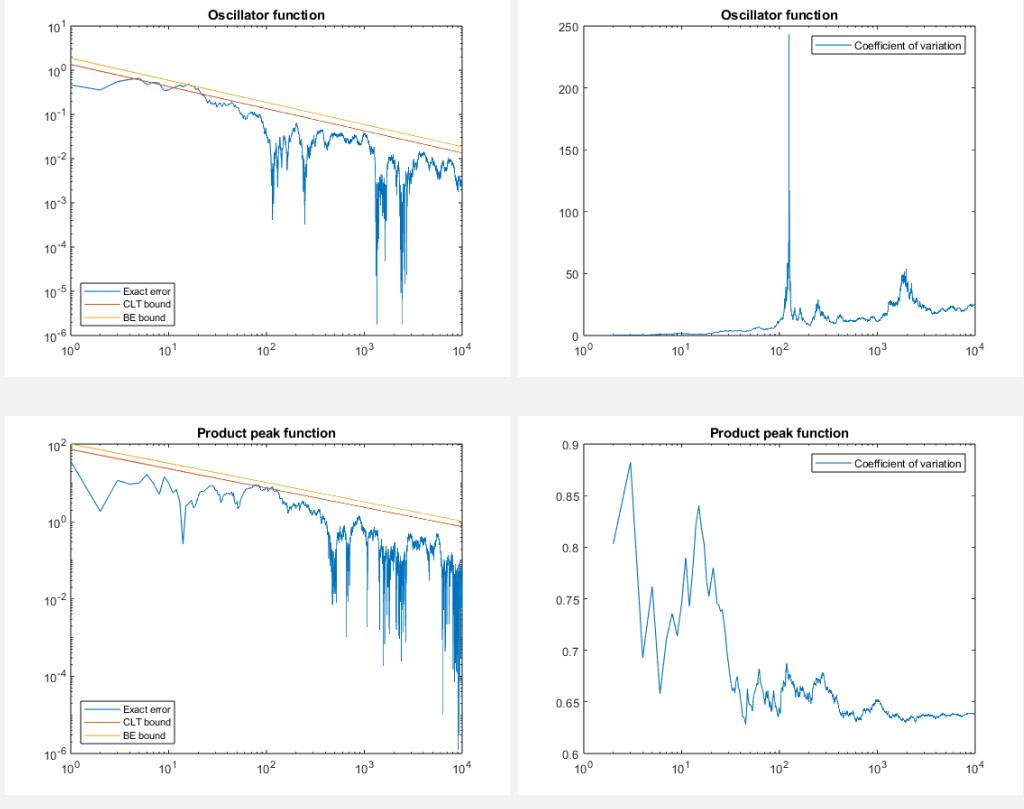
A similar result from the Berry Esseen theorem yields an error estimate (which is not asymptotic)

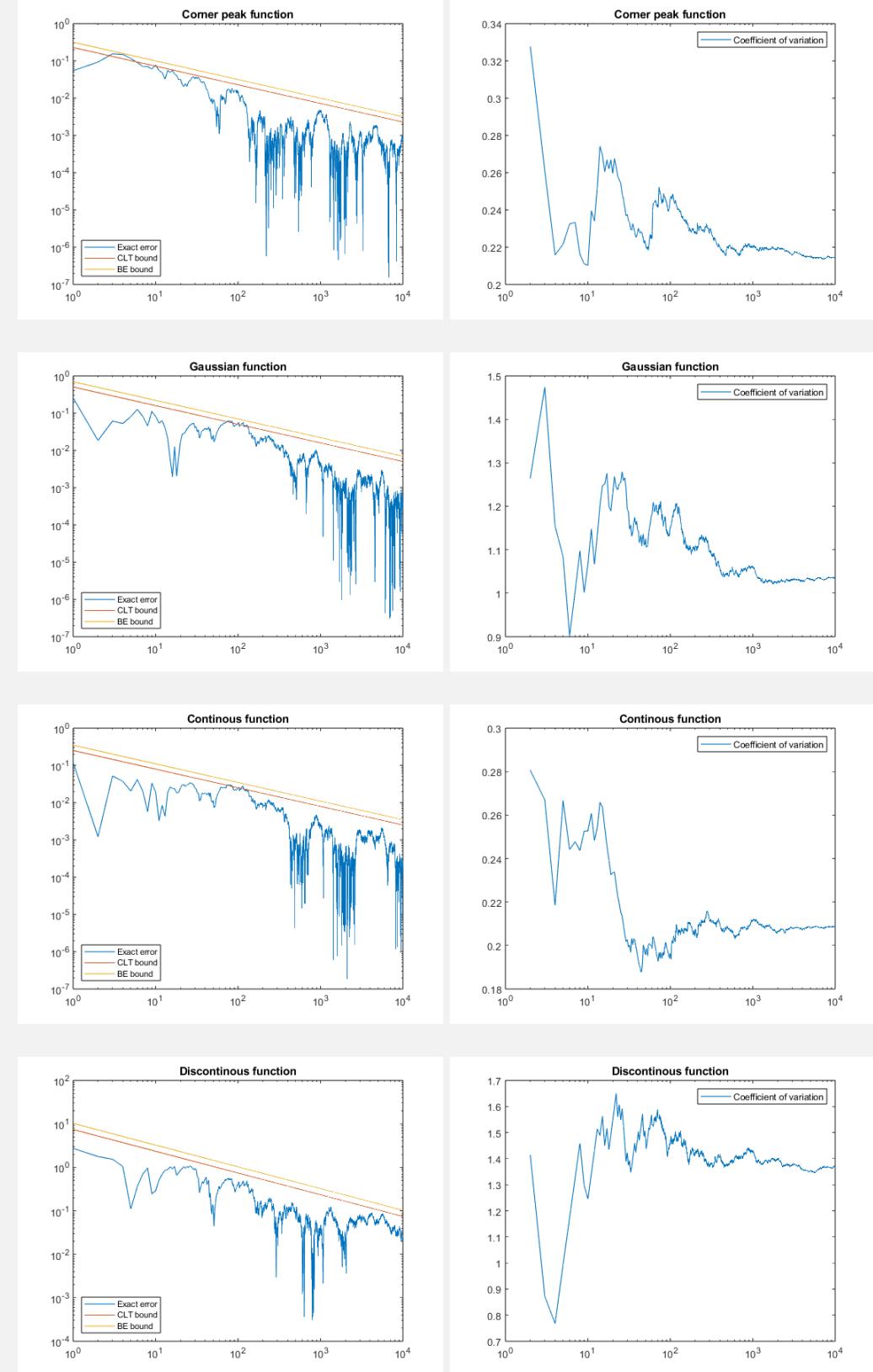
$$\mathbb{P}\left(|\varepsilon_S(Y, M)| \leq c_0 \frac{\sigma_Y}{\sqrt{M}}\right) \geq 2\Phi(c_0) - 1 - 2 \frac{C_{BE} \lambda^3}{(1 + c_0)^3 \sqrt{M}},$$

where $\varepsilon_S(Y, M)$ is the statistical error of Y resulting from the Berry Esseen theorem. For the CLT bound we will choose $\alpha = 0.05$ such that we get a 95%-confidence interval. For the Berry Esseen bound we want a similar result, so we solve

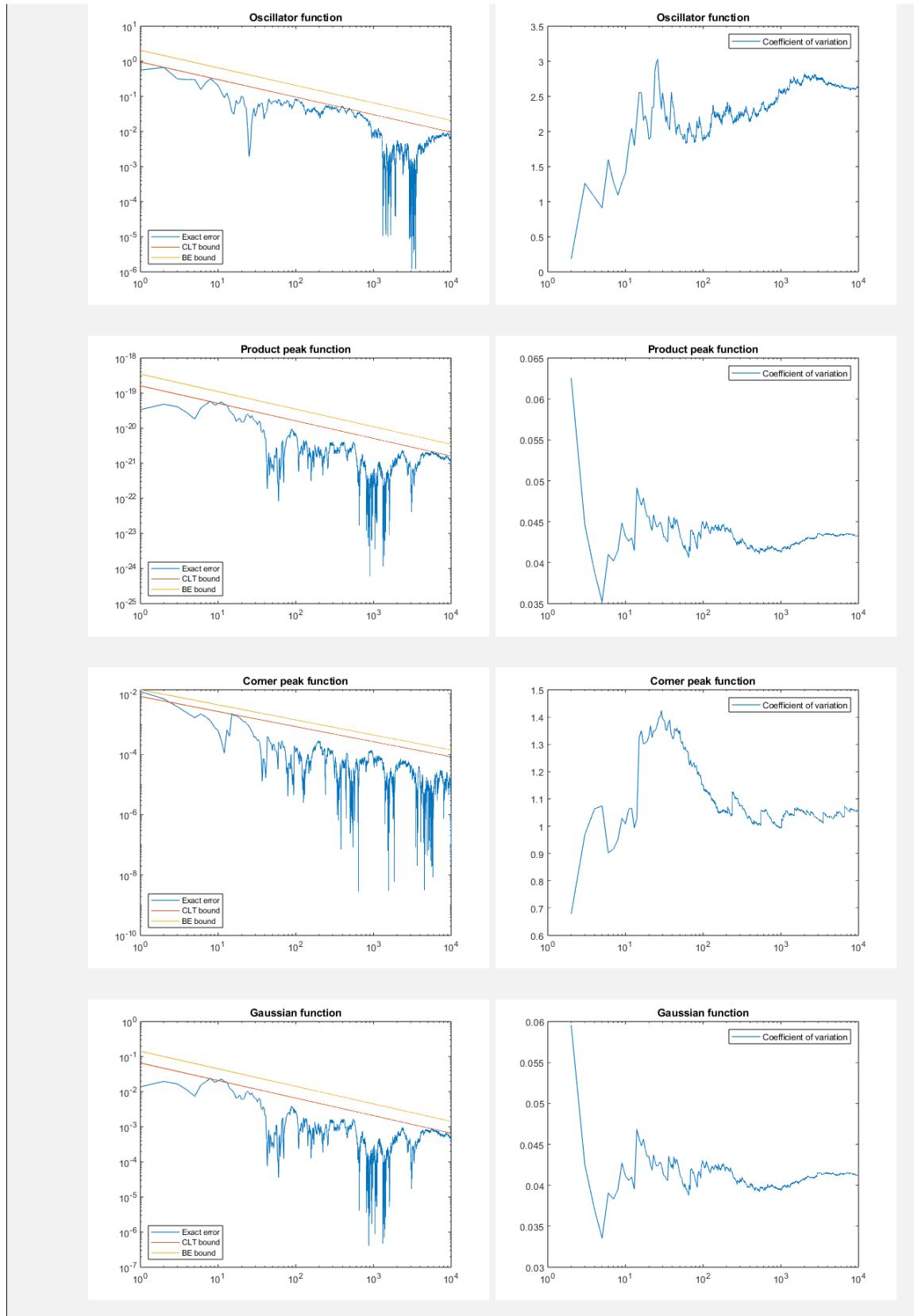
$$2\Phi(c_0) - 1 - 2 \frac{C_{BE} \lambda^3}{(1 + c_0)^3 \sqrt{M}} = 0.95$$

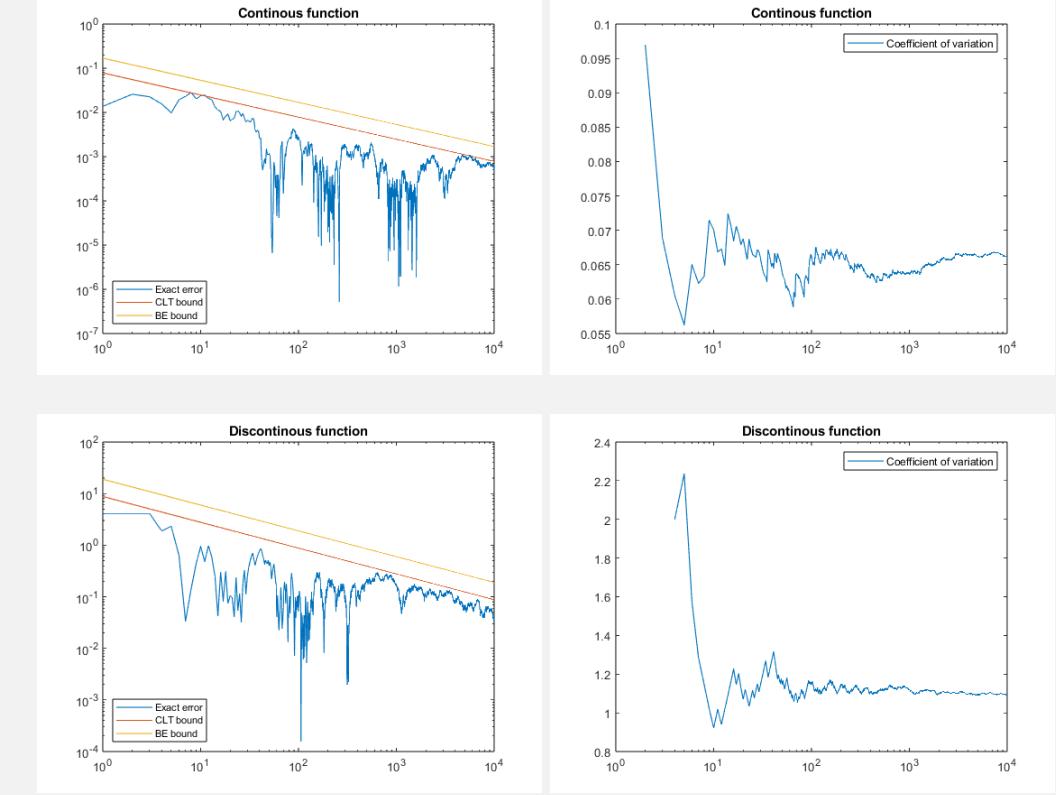
for $c_0 > 0$. For that we estimate the third moment using Monte Carlo with $M = 10^4$ samples. Below are the plots of the exact error, the CLT bound and the BE bound plotted against the number of samples considered. Additionally to all the plots we added a plot of the coefficient of variation also plotted against the number of samples used. We used $M = 10^4$ samples. **First the plots for $N = 2$ dimensions.**





Now the plots for $N = 20$ dimensions.





Finally we will consider approximations for the integral, for values $K = 3$ and $K = 6$,

$$\frac{1}{2\pi} \int_{\mathbb{R}^2} \max(e^{x_1} + e^{x_2} - K, 0) e^{\frac{-(x_1^2+x_2^2)}{2}} dx_1 dx_2.$$

We first note that we can rewrite this integral as an expected value

$$\frac{1}{2\pi} \int_{\mathbb{R}^2} \max(e^{x_1} + e^{x_2} - K, 0) e^{\frac{-(x_1^2+x_2^2)}{2}} dx_1 dx_2 = \mathbb{E}[\max(e^{X_1} + e^{X_2} - K, 0)],$$

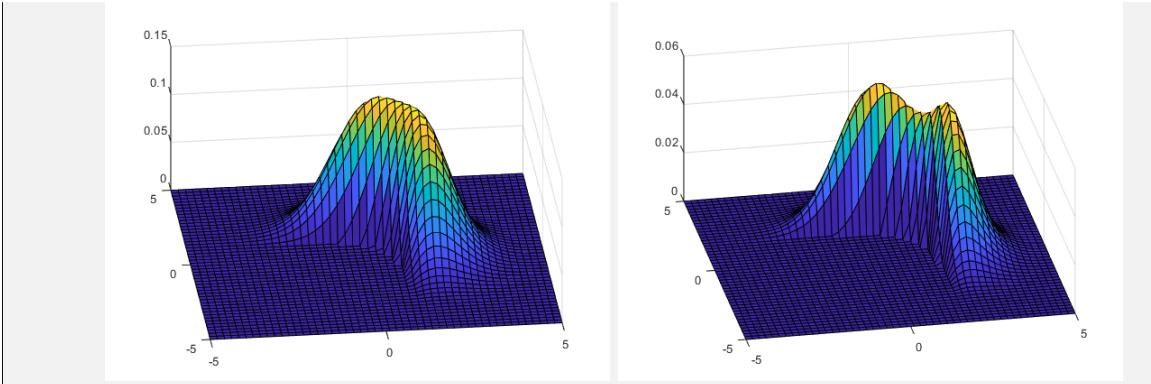
where $X_1, X_2 \sim \mathcal{N}(0, 1)$ independent. First we will approximate the expected value using plain Monte Carlo. After that we use Importance Sampling with the Shift Dilatation technique. For that we want to solve the problem

$$(\mu_1^*, \mu_2^*) = \operatorname{argmax}_{x_1, x_2} \max(e^{X_1} + e^{X_2} - K, 0) \rho_{x_1}(x_1) \rho_{x_2}(x_2)$$

where

$$\rho_{x_1}(x) = \rho_{x_2}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

If we look at the plots below for $K = 3$ (on the left) and $K = 6$ (on the right),



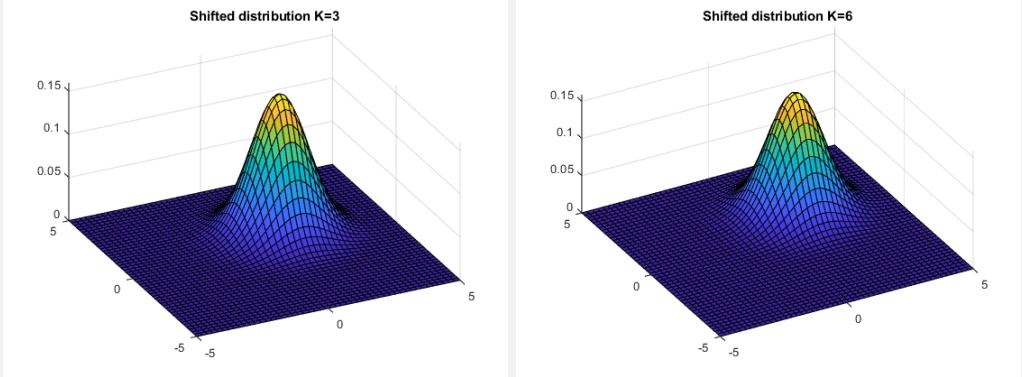
we can see that an optimal shift for $K = 3$ is along the diagonal of x and y axis into the positive direction, i.e. we choose for $K = 3$

$$(\mu_1^*, \mu_2^*) = (1.12245, 1.12245)$$

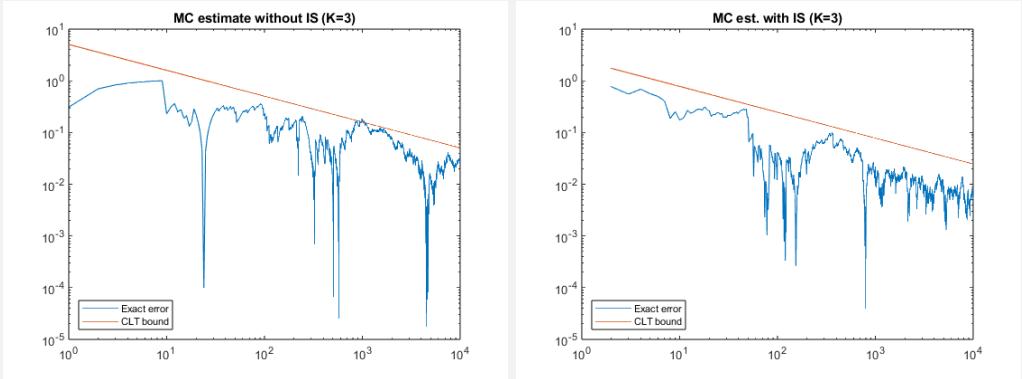
and for the other case ($K = 6$), we choose

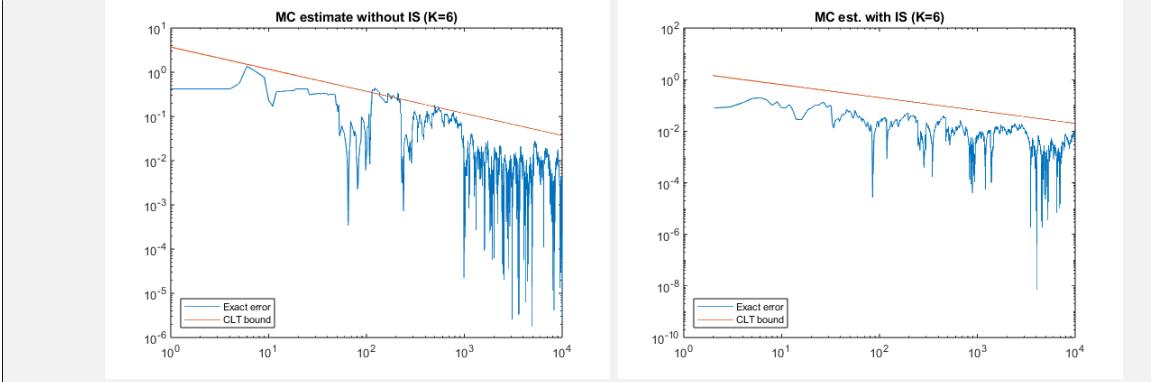
$$(\mu_1^*, \mu_2^*) = (1.57, 1.57).$$

Since we are supposed to plot the shifted densities, here they are:

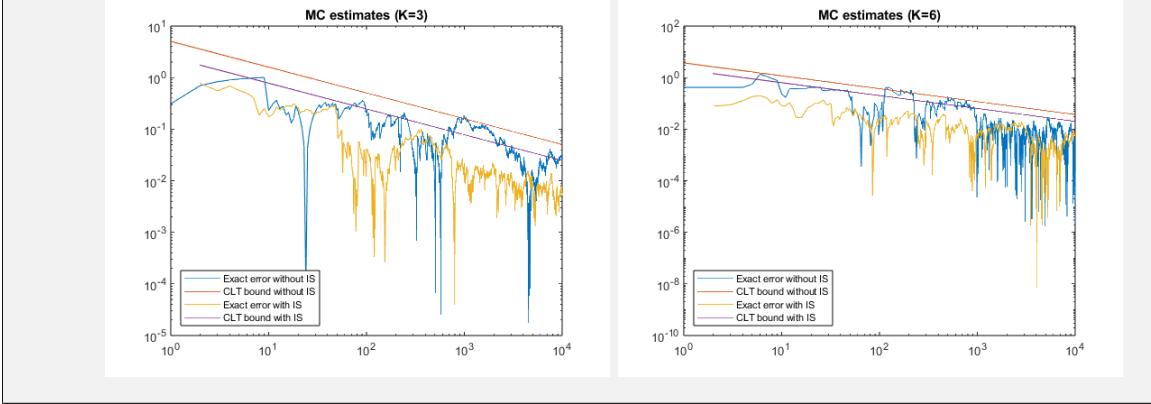


Now we using a reference solution (MC estimate with $M = 10^6$ samples and IS shifting) we can compute the exact error and plot it with the CLT bound against the number of samples. On the left are the plain MC estimates and on the right are the results of Monte Carlo using Importance Sampling.





And in the end also the plots of Monte Carlo estimations with and without Importance sampling in one plot for comparison.



2 MC for PDEs using approximate sampling

In this part we look at a solution to the one dimensional boundary value problem

$$-(a(x, w)u'(x, w))' = \underbrace{4\pi^2 \cos(2\pi x)}_{=: f(x)}, \quad \text{for } x \in (0, 1)$$

with $u(0, \cdot) = 0$ and $u(1, \cdot) = 0$. In this case we are interested in the quantity of interest

$$Q(u(w)) = \int_0^1 u(x, w) dx.$$

We use an $I + 1$ uniform grid $0 = x_0 < x_1 < \dots < x_I = 1$ on $[0, 1]$ with uniform spacing $h = x_i - x_{i-1} = \frac{1}{I}$, $i = 1, \dots, I$. Using this grid we can build a piecewise linear FEM approximation

$$u_h(x, w) = \sum_{i=1}^{I-1} \mathbf{u}_{h,i}(w) \varphi_i(x),$$

yielding a tridiagonal linear system for nodal values, $A(w)\mathbf{u}_h(w) = F$, with

$$\begin{aligned} A_{i,i-1}(w) &= -\frac{a(x_{i-1/2}, w)}{h^2} \\ A_{i,i}(w) &= \frac{a(x_{i-1/2}, w) + a(x_{i+1/2}, w)}{h^2} \\ A_{i,i+1}(w) &= -\frac{a(x_{i+1/2}, w)}{h^2} \end{aligned}$$

and

$$F_i = f(x_i).$$

Here we used the notation $x_{i+1/2} = \frac{x_i+x_{i+1}}{2}$ and $x_{i-1/2} = \frac{x_i+x_{i-1}}{2}$. The integral $Q(u_h)$ can now be computed exactly by a trapezoidal rule, yielding

$$Q(u_h) = h \sum_{i=1}^{I-1} \mathbf{u}_{h,i}.$$

For the random field $a(x, w)$ we use two different models

1. Piecewise constant coefficients

$$a(x, w) = 1 + \sigma \sum_{i=1}^N Y_n(w) \mathbf{1}_{[\hat{x}_{n-1}, \hat{x}_n]}(x)$$

with equidistant nodes $\hat{x}_n = \frac{n}{N}$ for $0 \leq n \leq N$ and i.i.d. uniform random variables $Y_n \sim \mathcal{U}([- \sqrt{3}, \sqrt{3}])$.

2. Log-normal

$$a(x, w) = \exp(\kappa(x, w))$$

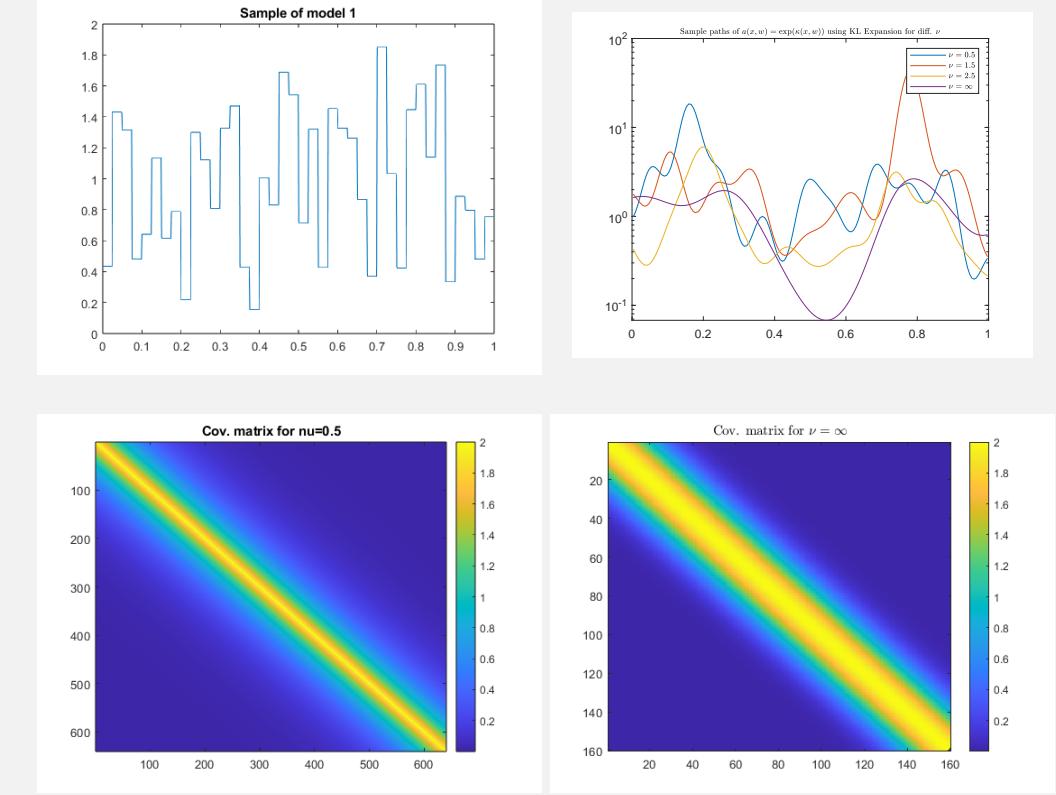
where $\kappa(x, w)$ is a stationary Gaussian random field with mean zero and the Matérn covariance function

$$C(x, y) = \sigma^2 \frac{1}{\Gamma(\nu) 2^{\nu-1}} \left(\sqrt{2\nu} \frac{|x-y|}{\rho} \right)^\nu K_\nu \left(\sqrt{2\nu} \frac{|x-y|}{\rho} \right)$$

where Γ is the Gamma function and K_ν is the modified Bessel function of the second kind. We will look at special cases

$$\begin{aligned} \nu = 0.5 &\implies C(x, y) = \sigma^2 \exp\left(-\frac{|x-y|}{\rho}\right) \\ \nu = 1.5 &\implies C(x, y) = \sigma^2 \left(1 + \frac{\sqrt{3}|x-y|}{\rho}\right) \exp\left(-\frac{\sqrt{3}|x-y|}{\rho}\right) \\ \nu = 2.5 &\implies C(x, y) = \sigma^2 \left(1 + \frac{\sqrt{5}|x-y|}{\rho} + \frac{\sqrt{3}|x-y|}{\rho^2}\right) \exp\left(-\frac{\sqrt{5}|x-y|}{\rho}\right) \\ \nu = \infty &\implies C(x, y) = \sigma^2 \exp\left(-\frac{|x-y|^2}{2\rho^2}\right) \end{aligned}$$

Moreover we choose $\sigma^2 = 2$ and $\rho = 0.1$. Below are plots of a sample for both models, we used $N = 40$, $\sigma = 0.5$ for Model 1 and $\nu \in \{0.5, 2.5\}$. Also a plot of a sample covariance function for those values of ν are shown.



I choose the possibility of using the Karhunen-Loeve expansion. We can write the random field $a(x, w)$ as an infinite series. Given a compact domain $D \subseteq \mathbb{R}^d$, continuous covariance function Cov_a , a sequence of values $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ with $\lim_{k \rightarrow \infty} \lambda_k = 0$ and corresponding sequence of functions $b_i : D \rightarrow \mathbb{R}$, $i = 1, 2, \dots$ such that they solve the eigenvalue problem

$$\int_D \text{Cov}_a(x, y) b_i(y) dy = \lambda_i b_i(x)$$

and

$$\int_D b_i(x) b_j(x) dx = \delta_{ij}$$

for all $x \in D$. Define, now, the sequence of random variables $y_i(w)$, $i = 1, 2, \dots$

$$y_i(w) = \frac{1}{\sqrt{\lambda_i}} \int_D (a(x, w) - \mathbb{E}[a](x)) b_i(x) dx$$

which are uncorrelated with zero mean and unit variance. Then, the random field $a(x, w)$ can be represented as the infinite series

$$a(x, w) = \mathbb{E}[a](x) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} b_i(x) y_i(w).$$

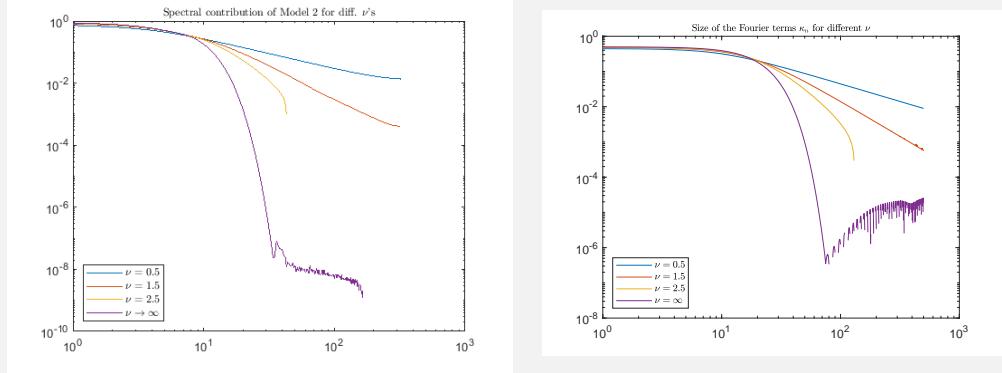
Moreover, the Karhunen-Loeve expansion is the best N -terms approximation in terms of variance

$$\{y_n, b_n\}_{n=1}^N = \operatorname{argmin}_{(\chi_n, \psi_n), \int_D \chi_n \psi_m = \delta_{mn}} \mathbb{E} \left[\int_D \left(a(x, \cdot) - \mathbb{E}[a](x) - \sum_{n=1}^N \chi_n(\cdot) \psi_n(x) \right)^2 dx \right].$$

The convergence rate of the N -term truncation

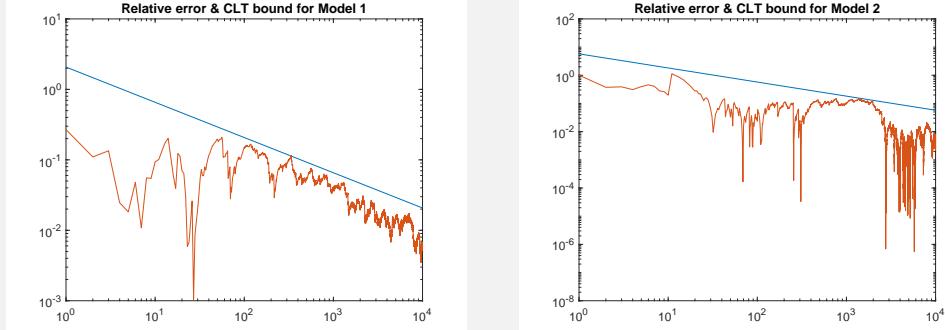
$$a_N(x, w) = \mathbb{E}[a](x) + \sum_{n=1}^N \sqrt{\lambda_n} b_n(x) y_n(w)$$

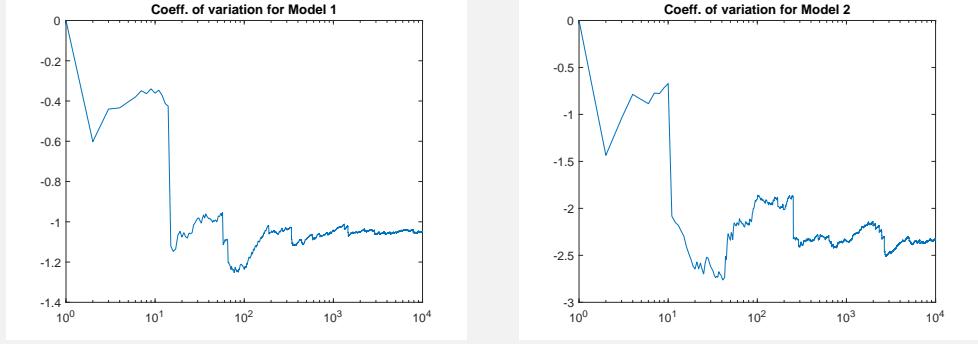
depends on the decay of the eigenvalues λ_n , which, in turn, depends on the smoothness of the integrant. Below you can see a plot of the spectral contribution of C of the Log-Normal model to the Karhunen-Loeve expansion for the different parameter values for ν . Next to this you can see a plot of the size of the Fourier terms κ_n for different values of ν using the presented Fourier expansion method for model 2 (I only include the plot because from now on I will only work with the KL expansion).



Numerical approximations of $\mathbb{E}[Q(u)]$

1. We will repeat the analysis of task 1 for $\mathbb{E}[Q(u_h)]$ for model 1 with $\sigma = 0.5$ and model 2 with $\nu = 0.5$. For that we will use plain Monte Carlo with $M = 5 \cdot 10^4$ samples and $I = 64$. As a reference solution we used a Monte Carlo estimate with $M = 10^5$ samples. The plots of the relative CLT bound and the relative error using the reference solutions are shown below. In addition we plotted the coefficient of variation just as in Task 1.





Now we are interested in estimating the bias using Richardson extrapolation. We can write the bias

$$\mathbb{E}[Q(u_h) - Q(u)] = a_1 h^p + O(h^r)$$

for some $r > p$. Using Richardson extrapolation, we can obtain an approximation of the bias that is r^{th} -order accurate. For example

$$\begin{aligned} \text{(I)} \quad Q(u_h) - Q(u) &= a_1 h^p + a_2 h^r + \text{h.o.t.} \\ \text{(II)} \quad Q(u_{h/2}) - Q(u) &= \frac{a_1}{2^p} h^p + \frac{a_2}{2^p} h^r + \text{h.o.t.} \end{aligned}$$

Subtracting (II) from (I) yields

$$Q(u_h) - Q(u_{h/2}) = \left(1 - \frac{1}{2^p}\right) a_1 h^p + O(h^r)$$

such that the bias is approximately

$$\text{Bias} \approx \frac{1}{\left(1 - \frac{1}{2^p}\right)} \mathbb{E}[(Q(u_h) - Q(u_{h/2}))]$$

The relative bias is given by

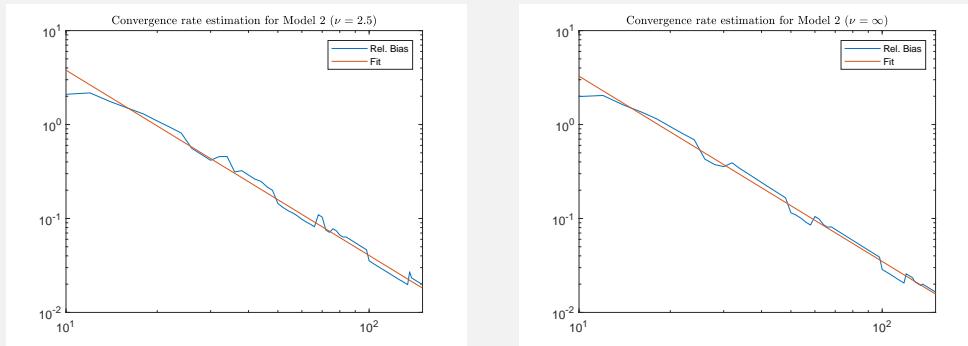
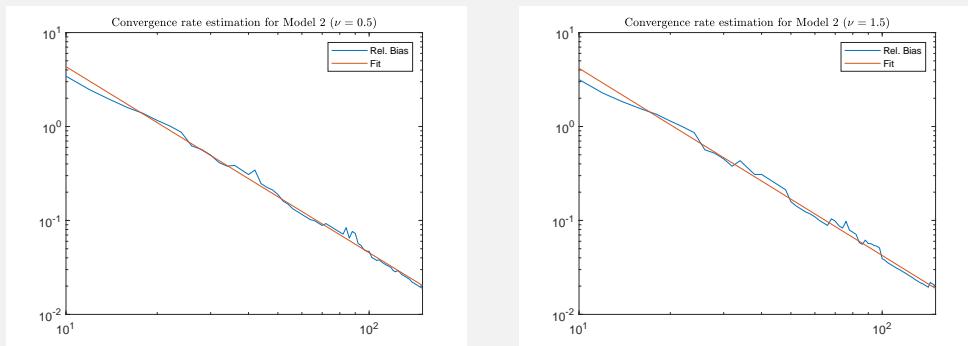
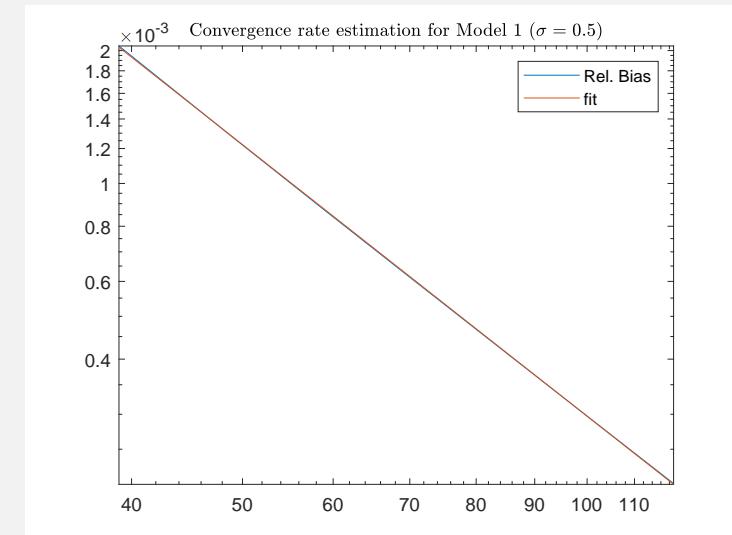
$$\varepsilon_{bias} = \frac{|\mathbb{E}[Q(u_h)] - \mathbb{E}[Q(u)]|}{\mathbb{E}[Q(u)]}$$

and the relative statistical error

$$\varepsilon_{stat} = \frac{c_\alpha \sqrt{\text{Var}[Q(u_h)]}}{\mathbb{E}[Q(u)] \sqrt{M}}.$$

For an estimation of p in Model 1, I calculated $u_{\frac{h}{2}}$ from a sample of the random field using a fine grid of some size I and then u_h with every second entry of the random field sample. After that, using $M = 10^3$ samples I computed the approximate bias $\mathbb{E}[Q_{h/2} - Q_h]$ and fitted a line on the log-log data to receive an approximation for p for model 1. Now for model 2 I used another technique, namely, similarly to model 1, calculating a sample of the random field using a fine grid of size I for the $Q_{h/2}$ and using the same sample but averaging out the redundant values by their neighbours to get a coarser mesh. Fitting a line to the log-data will result in a line with slope p . Below you can see the relative bias and the fit. The approximate values of p for different parameters and models can be seen in the table below.

	Parameter	<i>p</i> -value
Model 1	$\sigma = 0.5$	2.0435
Model 2	$\nu = 0.5$	1.9811
	$\nu = 1.5$	1.9924
	$\nu = 2.5$	1.9725
	$\nu = \infty$	1.9724

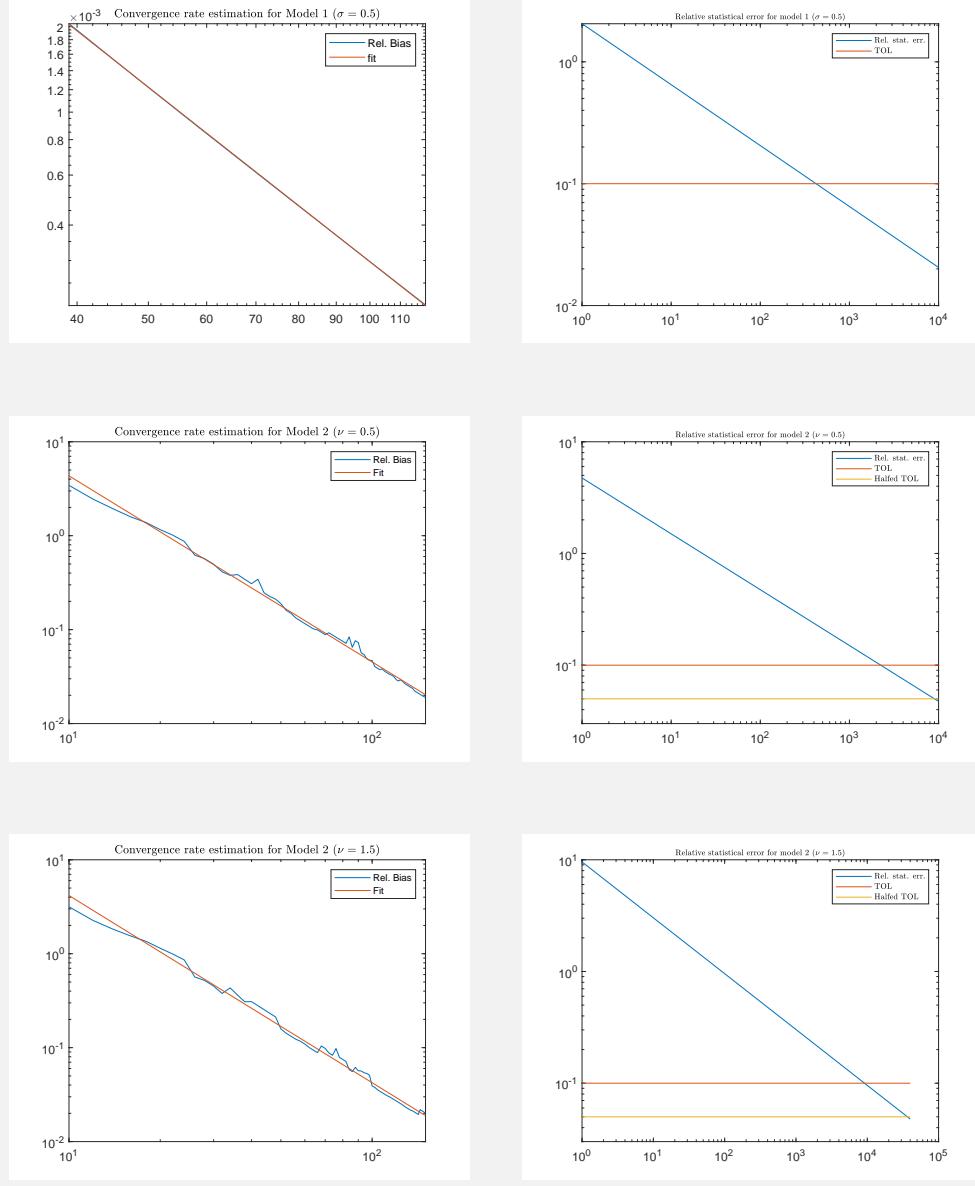


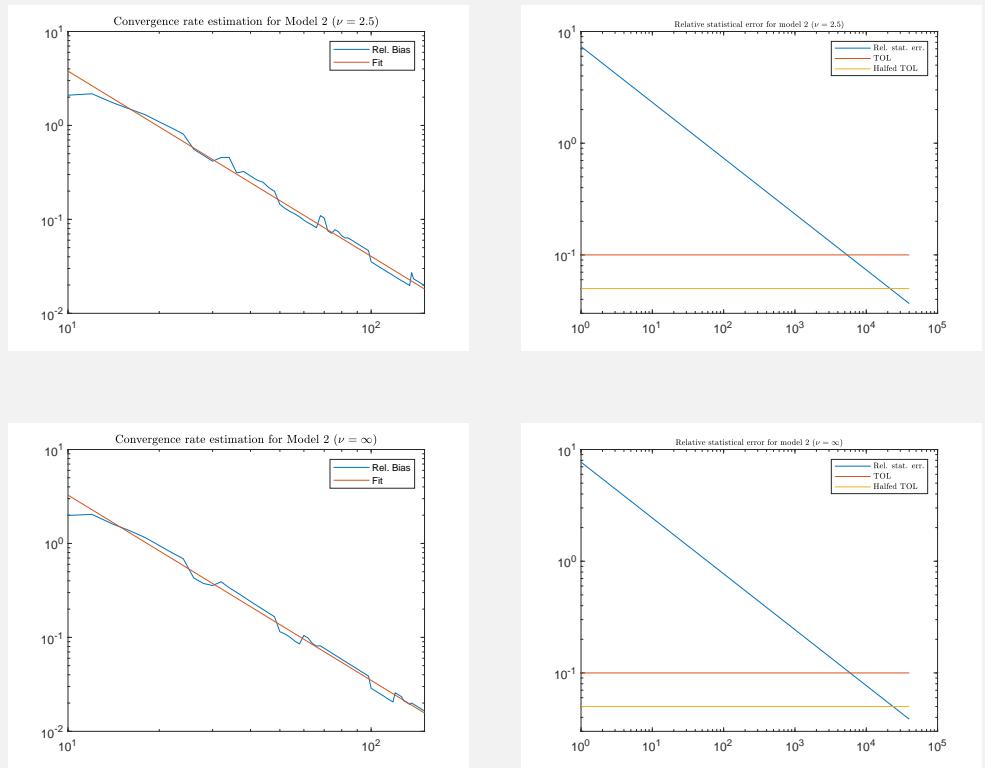
Now using the results on estimating the bias, we can choose M and h such that the total

relative error

$$\varepsilon_{total} = \varepsilon_{bias} + \varepsilon_{stat}$$

is below $TOL = 0.1$. For that we will first look at the plots of the relative bias and the relative statistical error for both models and all the parameters of model 2.





We can see that choosing $I = 100$, i.e. $h = 0.01$ results in a relative bias of min. half the tolerance of $TOL_{half} = 0.05$. Now we can look at the statistical error on the right plots and see that we have different sample sizes for each model and parameter to result in a total error of maximum $TOL = 0.1$. The table below shows an example of how to choose h and M for each model to satisfy $\varepsilon_{total} < TOL$.

	Parameter	h	M
Model 1	$\sigma = 0.5$	0.01	10^3
Model 2	$\nu = 0.5$	0.01	10^3
	$\nu = 1.5$	0.01	$3 \cdot 10^3$
	$\nu = 2.5$	0.01	$3 \cdot 10^3$
	$\nu = \infty$	0.01	$2 \cdot 10^3$

2. Now we will use control variates to see if we can achieve variance reduction in the Monte Carlo sampling and see if it pays in terms of computational work. Therefore we introduce the generalized control variate

$$\frac{1}{M_1} \sum_{i=1}^{M_1} Q^{(i)}(u_{2h}) + \frac{1}{M_2} \sum_{i=1}^{M_2} (Q(u_h) - Q(u_{2h}))^{(j)}$$

where we choose h as in task 1, i.e. $h = 0.01$ or $I = 100$. Now for this given h the variance of the estimator is

$$V = \frac{\text{Var}[Q(u_{2h}(w))]}{M_1} + \frac{\text{Var}[Q(u_h(w)) - Q(u_{2h}(w))]}{M_2}$$

and the work of the estimator

$$\text{Work} = M_1 C_1 + M_2 C_2$$

where C_1 is the cost of one realization of $Q(u_{2h}(w))$ and C_2 is the cost of one realization of $(Q(u_h)(w) - Q(u_{2h}))$. We want to solve

$$\begin{cases} \min_{\{M_1, M_2\}} & (M_1 C_1 + M_2 C_2) \\ \text{s.t.} & V \approx \frac{\text{TOL}^2}{C_a^2}. \end{cases}$$

We will first run pilots to estimate C_1 and C_2 by running $M = 10^3$ sample generations and devide the time it took to generate M samples and divide it by M to get a good average computing time of each sample. In addition to that, we will run crude Monte Carlo sampling for $Q(u_h)$ to estimate the work constant C_{CMC} for a comparison. The average times in seconds are shown in the table below.

	Parameter	C_1	C_2	C_{CMC}
Model 2	$\nu = 0.5$	$2.69 \cdot 10^{-3}$	$7.29 \cdot 10^{-3}$	$5.06 \cdot 10^{-3}$
	$\nu = 1.5$	$2.82 \cdot 10^{-3}$	$7.46 \cdot 10^{-3}$	$4.86 \cdot 10^{-3}$
	$\nu = 2.5$	$3.13 \cdot 10^{-3}$	$7.73 \cdot 10^{-3}$	$5.01 \cdot 10^{-3}$
	$\nu = \infty$	$2.86 \cdot 10^{-3}$	$7.48 \cdot 10^{-3}$	$4.83 \cdot 10^{-3}$

Now to get below a certain TOL in our Monte Carlo estimates we need different sample sizes, namely the sizes shown in the table below. Additionally the work for the control variate estimate and the crude Monte Carlo work is displayed using the samplesizes below. We can see that using this approach pays off in terms of computational work since it almost reduces the work by around 40%.

	Parameter	M_1	M_2	M_{CMC}	Work CV	Work CMC
Model 2	$\nu = 0.5$	2500	100	2600	7.454	13.156
	$\nu = 1.5$	7500	300	7800	23.388	37.908
	$\nu = 2.5$	7500	300	7800	25.794	39.078
	$\nu = \infty$	5000	200	5200	15.796	25.116

- For this part we are interested in approximating $\mathbb{P}(Q(u) < K)$ using importance sampling. We will use the mesh size as in task 1, i.e. $I = 100$ or $h = 0.01$ and different values of $K \in \{-5, -10, -20\}$. For this matter, we will rewrite the probability as an expectation

$$\mathbb{P}(Q(u) < K) = \mathbb{E} [\mathbf{1}_{\{Q(u) < K\}}] \approx \frac{1}{M} \sum_{m=1}^M \mathbf{1}_{\{Q(u_h(w_m)) < K\}}$$

where $u_h(w_m) \equiv u_h(Y^{(m)})$ and $Y^{(m)} \sim \mathcal{N}(0, 1)$. Similar to the first part of this homework we use a shift dilation technique and solve the N -dimensional optimization problem

$$\mu^* = \operatorname{argmax}_y \left\{ \mathbf{1}_{\{Q(u) < K\}} \prod_{n=1}^N \exp \left(-\frac{1}{2} y_n^2 \right) \right\}$$

We will also use the hint, that we can reduce the dimension of this problem by reducing the truncation scheme of the Karhunen-Loeve expansion. For this we choose $N = 5$. Solving the minimization problem above using MATLAB's build in function fmincon yields optimal shifts μ_K^* for each K and ν . A comparison of the resulting relative errors are shown below. First Model 2 with parameter $\nu = 0.5$, then Model 2 with parameter $\nu = \infty$. The number of samples are chosen accordingly such that the sum of stat. err. and relative bias error is less than $\text{TOL} = 0.1$.

