Error estimation and adaptive computation for linear partial differential equations with randomly perturbed data

Axel Målqvist,

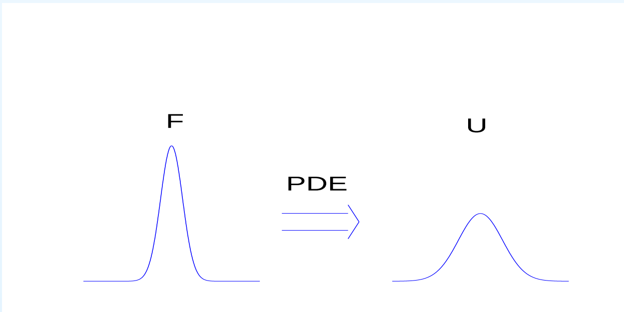
Donald Estep, and Simon Tavener

axel@cam.ucsd.edu

Department of Mathematics, University of California at San Diego

Outline

- Model problem with randomly perturbed data
- Our method for computing stochastic quantities (such as $E[\cdot]$ and Var$(\cdot)$) of a linear functional of the solution, (U, ψ)
- A posteriori error analysis (Discretization and Sample size)
- Adaptive algorithm that tunes method parameters
- Numerical examples
- Conclusions and related projects



Given samples of F the goal is to cheaply compute samples of (U, ψ).

The model problem

Strong form:

The Dirichlet problem with multiple right hand sides,

$F_j = \sum_{i=1}^{m} A_i^j v_i(x)$, $v_i \in L^2(\Omega)$, A_i^j random numbers, $j = 1, \ldots, N$, F_j are independent identically distributed,

$$-\triangle U_j = F_j \quad \text{in } \Omega,$$
$$U_j = 0 \quad \text{on } \Gamma.$$

Weak form:

Let $(\cdot, \cdot)$ denote the $L^2(\Omega)$ scalar product. Find $U_j \in V = H_0^1(\Omega)$ such

that,

$$(\nabla U_j, \nabla v) = (F_j, v) \quad \text{for all } v \in V.$$

The corresponding adjoint problem

Assume we are interested in a specific functional of the solution (U, ψ), where $\psi \in L^2(\Omega)$ is deterministic.

Strong form:

We let ϕ solve the adjoint problem with ψ in the right hand side,

$$-\triangle \phi = \psi \quad \text{in } \Omega,$$
$$\phi = 0 \quad \text{on } \Gamma.$$

Weak form:

Find $\phi \in V$ such that,

$$(\nabla w, \nabla \phi) = (w, \psi) \quad \text{for all } w \in V.$$

Note that the adjoint problem is deterministic.

Simple observation

Using Green's identity we get,

$$(U_j, \psi) = (\nabla U_j, \nabla \phi) = (F_j, \phi) \quad \text{for } j \in 1 \ldots, N.$$

- We can derive the distribution for (U_j, ψ) by just solving one partial differential equation instead of N.
- This works as long as the differential operator is linear and the adjoint problem is deterministic, i.e. for a wide range of problems.
- In particular it works for randomly perturbed initial and boundary conditions.
- This leads to a cheap way to study how sensitive a particular linear functional of the solution is to perturbations in data.

However, we need to solve the adjoint problem.

Discretization using the finite element method

Let $V_h \subset V$ be some finite element space.

Find $\phi_h \in V_h$ such that,

$$(\nabla w, \nabla \phi_h) = (w, \psi) \quad \text{for all } w \in V_h.$$

- We now have a computable approximation of $(U_j, \psi) = (F_j, \phi)$ namely (F_j, ϕ_h).
- Given the samples of (F_j, ϕ_h) we are interested in computing approximations of stochastic quantities of $X_j = (U_j, \psi)$, $m(X)$ cheaply.
- We need control over errors committed by discretization (h) (since we approximate ϕ by ϕ_h) and also committed by using a smaller sample size ($n \ll N$) in order to compute approximations to the stochastic quantities.

Error estimation

Let $X_j = (U_j, \psi) = (F_j, \phi)$ and $X_{h,j} = (F_j, \phi_h)$.

Let $m(X)$ denote the exact stochastic quantity and let $M(X)$ be an unbiased estimator computed using n realizations of X i.e.,

$$E[M(X)] = m(X).$$

We divide the error into two parts,

$$m(X) - M(X_h) = (m(X_h) - M(X_h)) + (m(X) - m(X_h))$$

- We call the first part *Stochastic error* (n).
- We call the second part *Discretization error* (h).

Stochastic error $m(X_h) - M(X_h)$

Chebyshev inequality:

$$P\left(|Y - E[Y]| \geq \delta\right) < \text{Var}(Y)/\delta^2$$

or by choosing $\delta = \sqrt{\text{Var}(Y)/\epsilon}$ and turning the inequalities around,

$$P\left(|Y - E[Y]| < \sqrt{\text{Var}(Y)/\epsilon}\right) \geq 1 - \epsilon.$$

We let $Y = M(X_h)$ and use that M is an unbiased estimotor for m i.e. $E[Y] = E[M(X_h)] = m(X_h)$ to get,

$$P\left(|m(X_h) - M(X_h)| < \sqrt{\text{Var}(M(X_h))/\epsilon}\right) \geq 1 - \epsilon.$$

Example 1: $m(X) = E[X]$

We study the special case when m is the expected value closer.

Let $m(X_h) = E[X_h]$ then $M(X_h) = \bar{X}_h$. We have,

$$P\left(|E[X_h] - \bar{X}_h| < \sqrt{\text{Var}(\bar{X}_h)/\epsilon}\right) \geq 1 - \epsilon.$$

$$\text{Var}(\bar{X}_h) = \text{Var}\left(\sum_{i=1}^{n} X_{h,j}/n\right) = \text{Var}(X_{h,j})/n.$$

Since F_j are independent identically distributed X_j and $X_{h,j}$ will also be iid. Let $\sigma = \text{Var}(X_{h,j})$ then,

$$P\left(|E[X] - \bar{X}| < \sigma/\sqrt{n\epsilon}\right) \geq 1 - \epsilon.$$

Example 2: $m(X) = \text{Var}(X)$

We let $M(X) = S_n^2(X) = \sum_{j=1}^{n} (X_j - \bar{X}_j)^2/(n-1)$.

We need to estimate the variance of $S_n^2(X_h)$.

$$\text{Var}(S_n^2(X_h)) = \text{Var}\left(\sum_{j=1}^{n} (X_{h,j} - \bar{X}_{h,j})^2\right)/(n-1)^2.$$

We assume $(X_{h,j} - \bar{X}_{h,j})^2$ to be almost independent,

$$\text{Var}(S_n^2(X_h)) \approx \sum_{j=1}^{n} \text{Var}\left((X_{h,j} - \bar{X}_{h,j})^2\right)/(n-1)^2.$$

Example 2: $m(X) = \text{Var}(X)$

Given $\{X_{h,j}\}_{j=1}^{n}$ we can get a good approximation of $\text{Var}((X_{h,j} - \bar{X}_{h,j})^2)$ using the unbiased estimator on this particular sample, lets call it

$$s_n^2((X_h - \bar{X}_h)^2) \approx \text{Var}((X_{h,j} - \bar{X}_{h,j})^2),$$

then,

$$\text{Var}(S_n^2(X_h)) \approx n s_n^2((X_h - \bar{X}_h)^2)/(n-1)^2,$$

so,

$$P\left(|\text{Var}(X_h) - S_n^2(X_h)| < \frac{C_V}{\sqrt{n\epsilon}}\right) \geq 1 - \epsilon,$$

where a good approximation to C_V can be computed.

Error estimation

Remember the error has two contributions,

$$m(X) - M(X_h) = (m(X_h) - M(X_h)) + (m(X) - m(X_h))$$

We turn our attention to the discretization error.

Discretization error $m(X) - m(X_h)$

Remember that,

$$m(X) - m(X_h) = m((F, \phi)) - m((F, \phi_h)).$$

Since ϕ is not known and we only have access to M we need to estimate this quantity.

Let $\phi_{\gamma h}$, $0 < \gamma < 1$, be an improved version of ϕ_h. Then,

$$\begin{aligned} m(X) - m(X_h) = {} & M((F, \phi_{\gamma,h})) - M((F, \phi_h)) \\ & + M((F, \phi)) - M((F, \phi_{\gamma h})) \\ & + m((F, \phi)) - m((F, \phi_h)) - M((F, \phi)) + M((F, \phi_h)) \end{aligned}$$

Example 1: $m(X) = E[X]$

In this case m is linear which leads to,

$$\begin{aligned} m(X) - m(X_h) &= E[X - X_h] = E[(F, \phi - \phi_h)] \\ &= (\bar{F}, \phi - \phi_h) + (E[F] - \bar{F}, \phi - \phi_h) \\ &= (\bar{F}, \phi_{\gamma h} - \phi_h) + (\bar{F}, \phi - \phi_{\gamma h}) + (E[F] - \bar{F}, \phi - \phi_h). \end{aligned}$$

However is this case we have another option,

$$m(X) - m(X_h) = (\bar{F}, \phi - \phi_h) + \text{higher order term}$$

We can construct an adjoint problem to take care of the $(\bar{F}, \phi - \phi_h)$-term.

This works since it is a linear functional of the error $\phi - \phi_h$.

Example 1: Another adjoint problem

We let $\chi \in V$ solve,

$$
\begin{aligned}
-\triangle \chi = \bar{F} \quad & \text{in } \Omega, \\
\chi = 0 \quad & \text{on } \Gamma,
\end{aligned}
$$

in order to get hold the error in the linear functional $(\bar{F}, \phi - \phi_h)$. This leads to the following error representation formula,

$$
(\phi - \phi_h, \bar{F}) = (\nabla(\phi - \phi_h), \nabla \chi) = (\psi, \chi) - (\nabla \phi_h, \nabla \chi) := (R(\phi_h), \chi - \pi_h \chi),
$$

where R denotes the residual of the ϕ_h-equation.

This only works because $E[\cdot]$ i linear.

Example 1: $m(X) = E[X]$

The discretization part of the error when $m(X) = E[X]$ consists of two terms,

$$E[X] - E[X_h] = (R(\phi_h), \chi - \pi_h \chi) + (E[F] - \bar{F}, \phi - \phi_h).$$

We can proceed with an interpolation estimate if we assume enough regularity in χ,

$$|E[X] - E[X_h]| \leq C\|h^2 \mathcal{R}(\phi_h)\| + |(E[F] - \bar{F}, \phi - \phi_h)|,$$

where $\mathcal{R}$ is a bound of the residual that includes the jump terms.

We can now combine this estimate with the estimate of the stochastic error contribution.

Example 1: Total error estimate

We want to estimate $E[X] - \bar{X}_h$. If we combine the two results we get: the probability that,

$$|E[(U,\psi)] - (\bar{U}_h, \psi)| \leq \sqrt{\mathrm{Var}((F,\phi_h))/(n\epsilon)} \\ + C\|h^2 R(\phi_h)\| + |(E[F] - \bar{F}, \phi - \phi_h)|.$$

holds is greater then $1 - \epsilon$.

If we skip the higher order term and assume enough regularity we have that,

$$|E[(U,\psi)] - (\bar{U}_h, \psi)| \leq C_1/\sqrt{n\epsilon} + C_2 h^2, \quad n \sim h^{-4},$$

holds with probability $1 - \epsilon$.

In general

For an arbitrary moment m we will have to compute $\phi_{\gamma h}$ and an approximation to Var$(M(X_h))$ in order to get an the following approximate bound:

$$|m(X) - M(X_h)| \leq \sqrt{\text{Var}(M(X_h))/\epsilon} + |M((F, \phi_{\gamma h})) - M((F, \phi_h))|,$$

holds approximately with probability $1 - \epsilon$.

The higher order terms are neglected here.

Same method for related problems: Different operator

Different deterministic linear operator,

$$LU_j = F_j \quad \text{in } \Omega,$$
$$U_j = 0 \quad \text{on } \Gamma.$$

The adjoint problem reads,

$$L^*\phi = \psi \quad \text{in } \Omega,$$
$$\phi = 0 \quad \text{on } \Gamma.$$

The framework extends easily to deterministic linear operators. We get the following formula for computing samples,

$$(U_j, \phi) = (U_j, L^*\phi) = (LU_j, \phi) = (F_j, \phi).$$

Same method for related problems: Time dependent

Stochastic initial condition, primal and dual:

$$\begin{aligned} \dot{U}_j - \triangle U_j &= 0 \quad \text{in } \Omega, \quad t > 0, \\ U_j &= 0 \quad \text{on } \Gamma, \quad t > 0, \\ U_j &= F_j \quad \text{for } t = 0. \end{aligned}$$

$$\begin{aligned} -\dot{\phi} - \triangle \phi &= 0 \quad \text{in } \Omega, \quad t < T, \\ \phi &= 0 \quad \text{on } \Gamma, \quad t < T, \\ \phi &= \psi \quad \text{for } t = T. \end{aligned}$$

Again we can use that the adjoint problem is deterministic to get a simple formula to compute the distribution of a linear functional of the solution.

$$(U_j(T), \psi) = (F_j, \phi(0)).$$

Same method for related problems: Boundary cond.

Randomly perturbed boundary condition:

$$-\triangle U_j = 0 \quad \text{in } \Omega,$$
$$-\partial_n U_j = k U_j + F_j \quad \text{on } \Gamma,$$

the adjoint problem reads,

$$-\triangle \phi = \psi \quad \text{in } \Omega,$$
$$-\partial_n \phi = k\phi \quad \text{on } \Gamma.$$

In this case we get the following formula for computing samples of a linear functional of the solution,

$$(U_j, \psi) = -\langle F_j, \phi \rangle.$$

One example when the idea does not work.

We consider the Poisson equation with randomly perturbed diffusion coefficient,

$$-\nabla \cdot \mathcal{A}_j \nabla U_j = f \quad \text{in } \Omega,$$
$$U_j = 0 \quad \text{on } \Gamma,$$

The corresponding adjoint problem also has randomly perturbed coefficient,

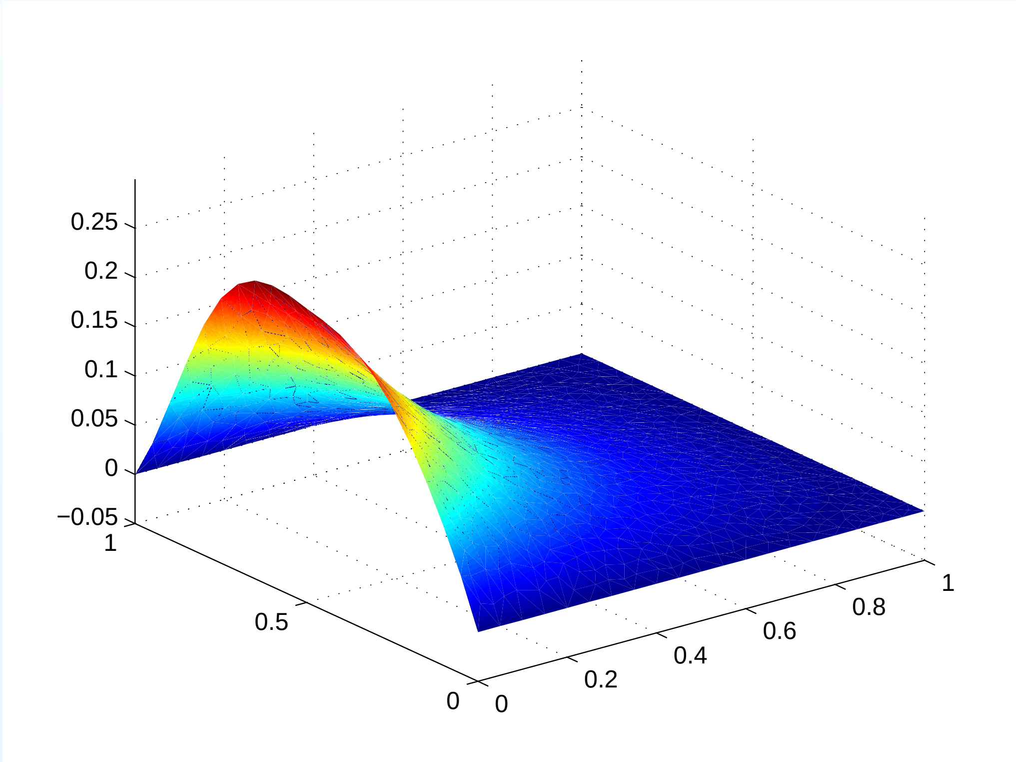
$$-\nabla \cdot \mathcal{A}_j \nabla \Phi_j = \psi \quad \text{in } \Omega,$$
$$\Phi_j = 0 \quad \text{on } \Gamma,$$

We can do the same trick but this time we do not gain anything since the adjoint problem is equally hard to solve as the original,

$$(U_j, \psi) = (f, \Phi_j).$$

Numerical examples: m(X)=E[X]

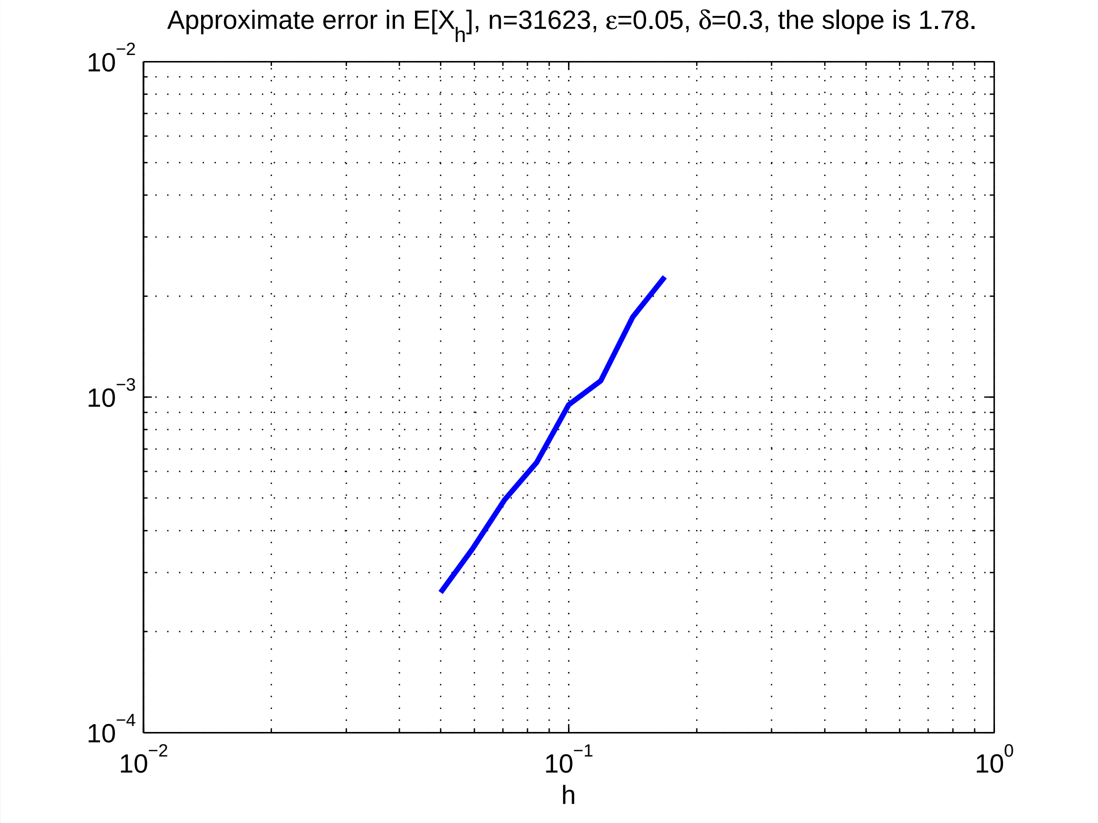
Let $-\triangle U_j = 0$ and $-\partial_n U_j = kU_j + F_j$, where $F_j = -\sin(\pi y) + \delta_j^{[5]}$ is randomly perturbed



Quasi uniform mesh with meshsize $h \approx 0.05$, $\delta = 0.6$, $k = 0$ *on Neumann part.*

Test of convergence ($P(|E[X] - \bar{X}_h| < y) \geq 0.95$)

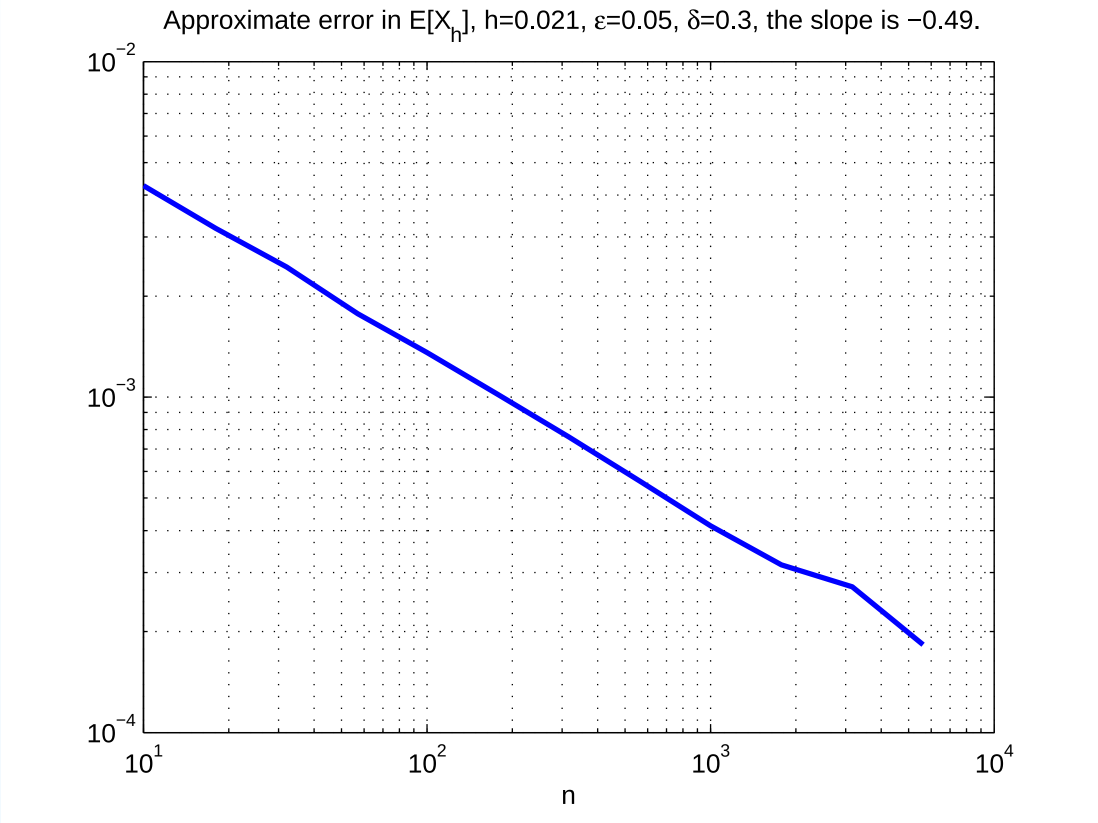
Error computed using reference solution, $n_{\text{ref}} = 5 \cdot 10^5$ and $h_{\text{ref}} = 0.01$. We let $\epsilon = 0.05$.



For each h we compute ≈ 160 realizations of the error using $n \approx 3 \cdot 10^4$ (big) and pick the 95% worst value.

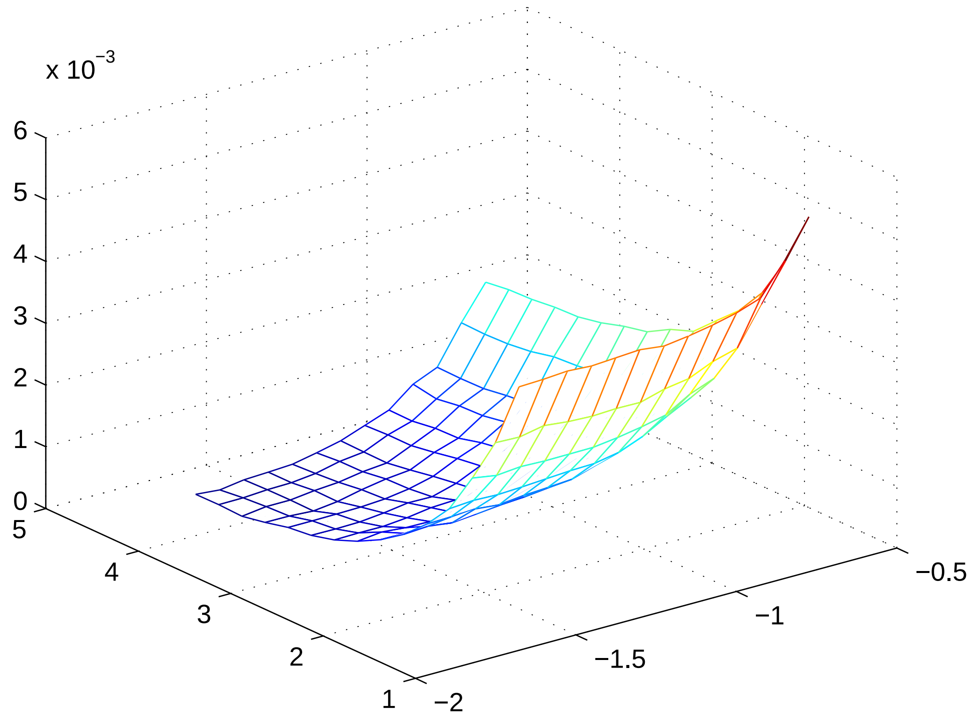
Test of convergence ($P(|E[X] - \bar{X}_h| < y) \geq 0.95$)

Error computed using reference solution, $n_{\text{ref}} = 5 \cdot 10^5$ and $h_{\text{ref}} = 0.01$.
We let $\epsilon = 0.05$.



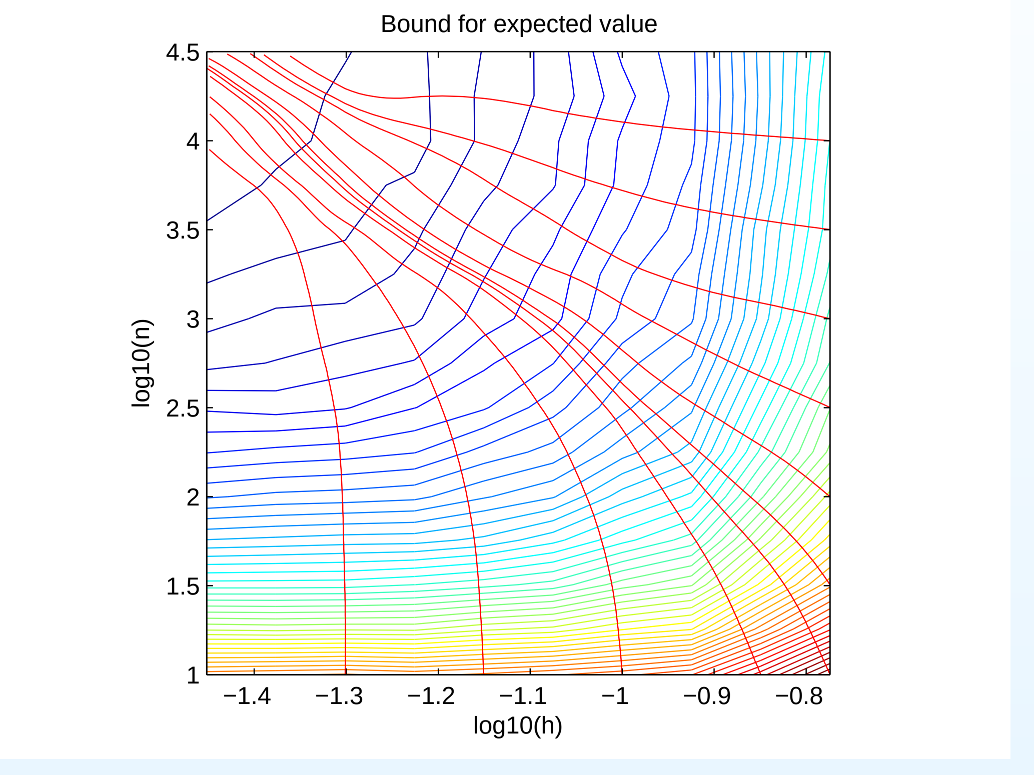
For each n compute $\approx 5 \cdot 10^5/n$ real. of the error using $h \approx 0.02$ (small) and pick the 95% worst.

Test of convergence in both h and n



The 95% probability bound of $|E[X_{h_{ref}}] - \bar{X}_h|$ for different choices of h and n.

Test of convergence in both h and n



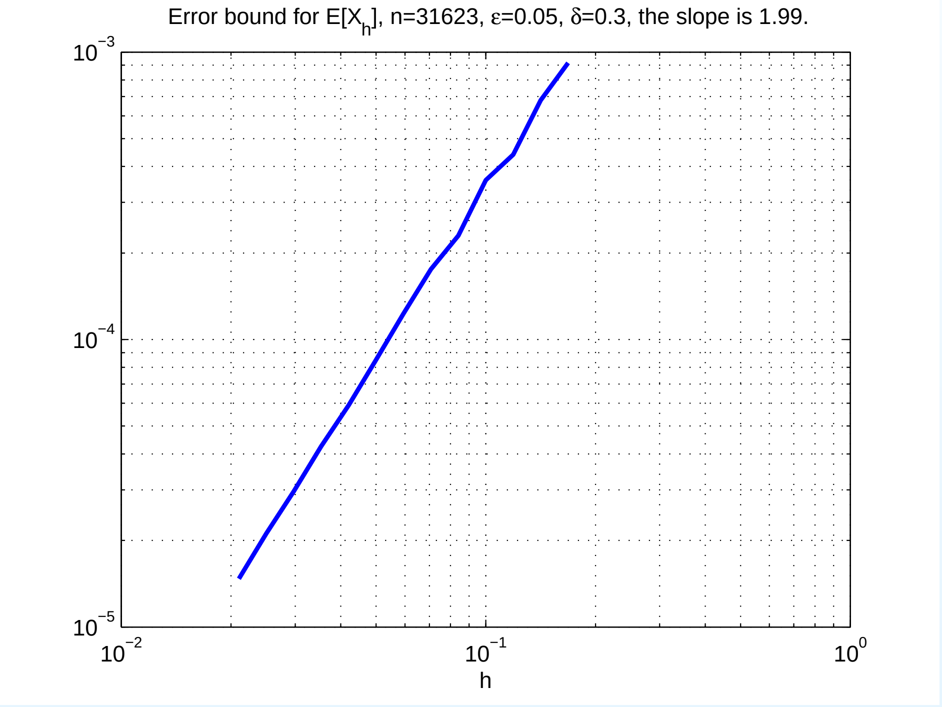
Contour plot with steepest decent paths indicating the dependence between h and n when trying to minimize the error.

The slope seem to be close to -4 which we expected.

The error bound

Remember $P(|E[X]-\bar{X}_h| < \sigma/\sqrt{n\epsilon} + Ch^2\|\mathcal{R}(\phi_h)\|) \leq 0.95$.

We let $n = 31623$ (big) and vary h between $0.05 \leq h \leq 0.168$.

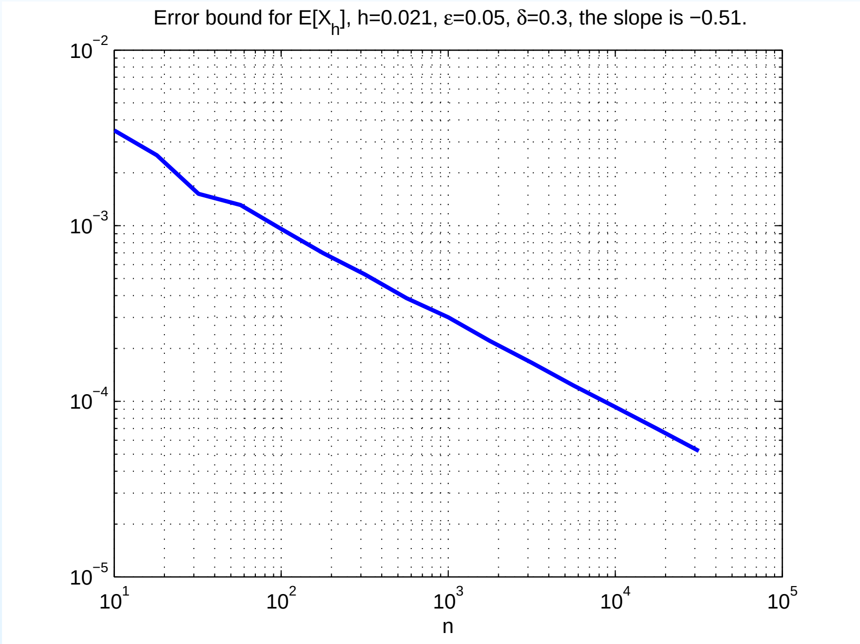


We note that the 95% bound of $|\bar{X}-\bar{X}_h|$ approximately depends on h^2.

The error bound

Remember $P(|E[X] - \bar{X}_h| < \sigma/\sqrt{n\epsilon} + Ch^2\|\mathcal{R}(\phi_h)\|) \leq 0.95$.

Let $h = 0.021$ (small) and compute the bound of the stochastic contribution of the error.



We note that the 95% *bound of* $|\bar{X} - \bar{X}_h|$ *approximately depends on* $1/\sqrt{n}$.

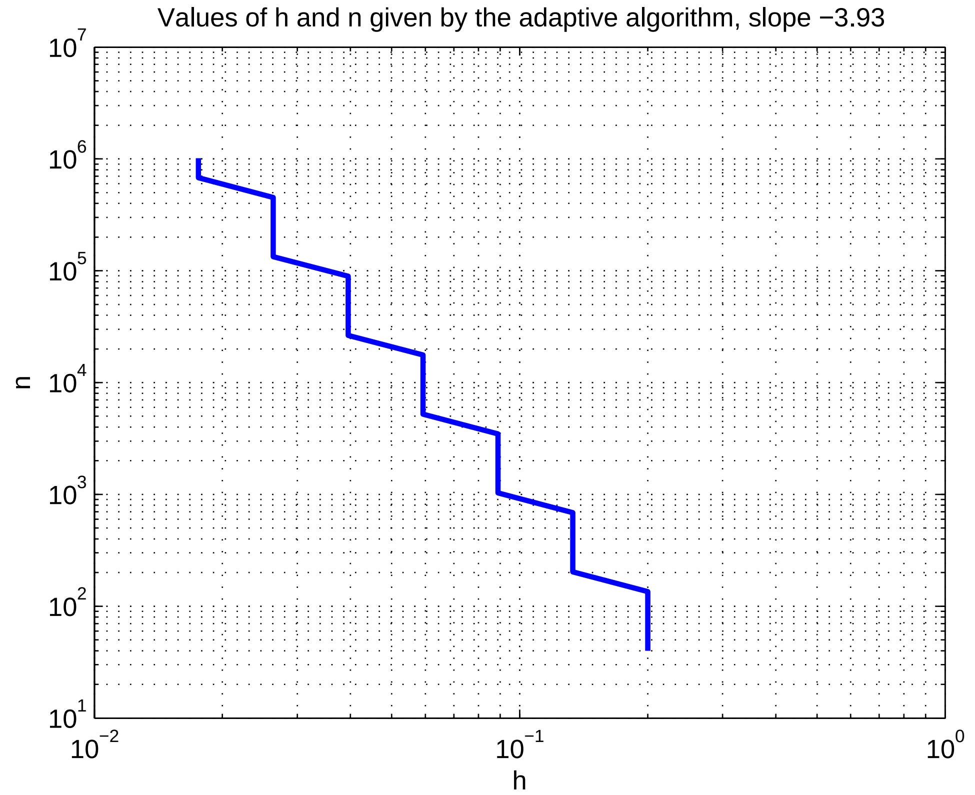
Adaptivity

Remember $P(|E[X] - \bar{X}_h| < \sigma/\sqrt{n\epsilon} + Ch^2\|\mathcal{R}(\phi_h)\|) \leq 0.95$. We now present an adaptive algorithm based on the a posteriori error estimate that tunes the method parameters, h and n automatically,

1. Choose ϵ, TOL, $r > 1$, $h = h_{\text{start}}$, and $n = n_{\text{start}}$.
2. Compute the solutions $X_{h,j} = (F_j, \phi_h)$, $1 \leq j \leq n$.
3. Compute $S = \sigma/\sqrt{n\epsilon}$.
4. Solve adjoint problem (χ) on finer mesh or using higher order method.
5. Compute $D = |(\psi, \chi - \pi_h\chi) - (\nabla\phi_h, \nabla(\chi - \pi_h\chi)) - \langle k\phi_h, \chi - \pi_h\chi\rangle|$ or $D = Ch^2\|\mathcal{R}\phi_h\|$.
6. If $D + S < TOL$ stop.
7. If $D > rS$ then let $h := h/r$ and $n := n$. If $S > rD$ then let $h := h$ and $n := r \cdot n$. Otherwise $h := h/r$ and $n := n \cdot r$.
8. Goto 2.

Adaptivity in h and n

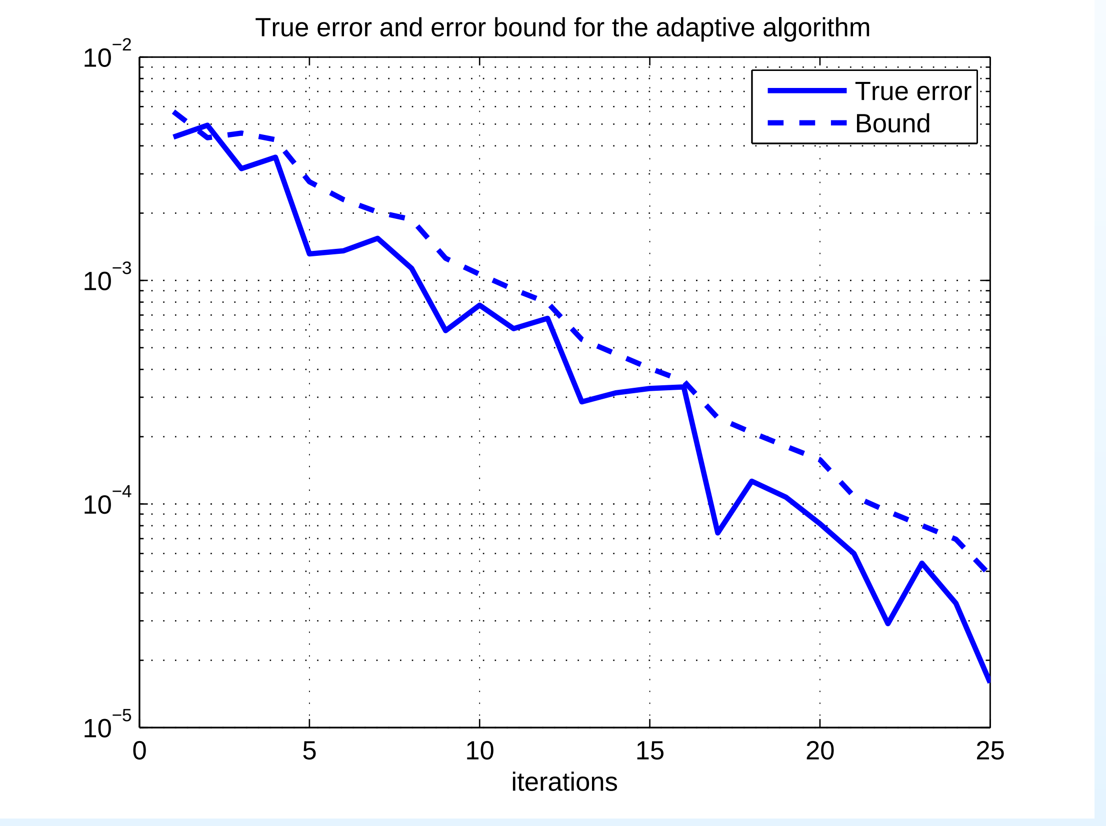
Let $\epsilon = 0.05$, $r = 1.5$, $TOL = 5 \cdot 10^{-5}$, (relative error less then 0.1%), $h_{\text{start}} = 0.2$, and $n_{\text{start}} = 40$.



We see clearly how the algorithm enforces $n \sim h^{-4}$.

Bound verses true error (using reference solution)

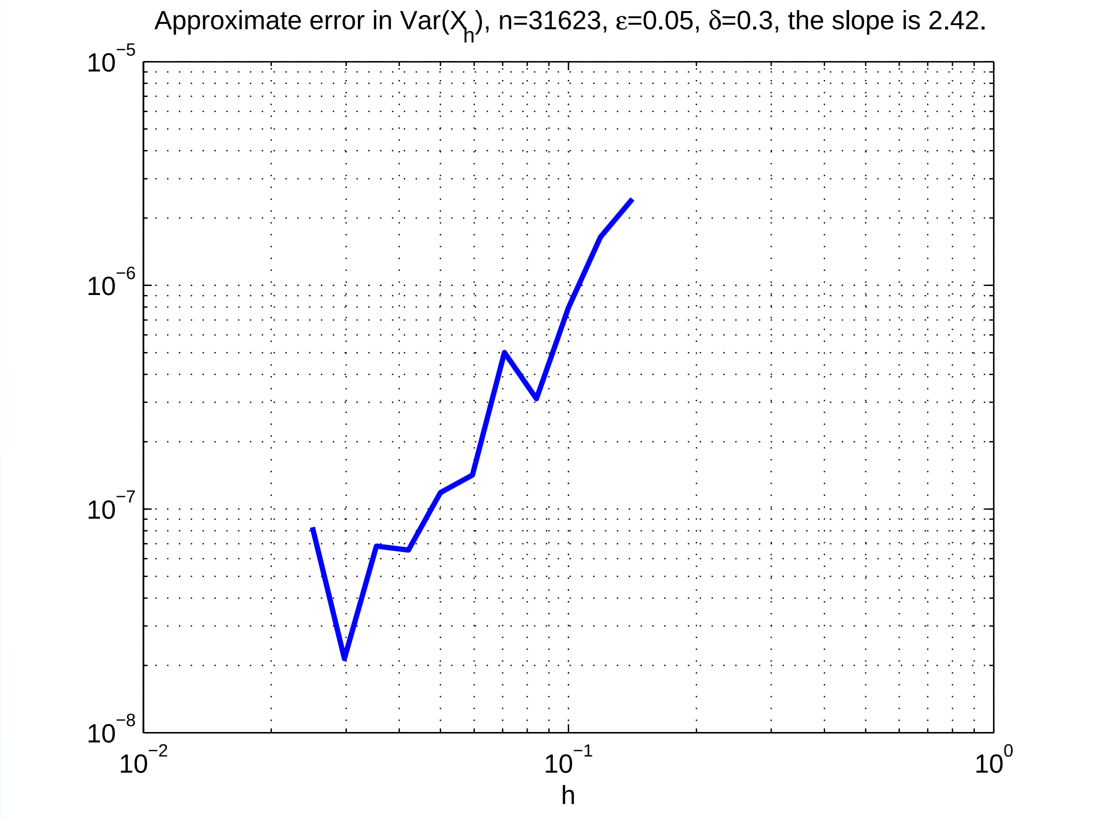
Error compared to reference solution and error bound after each iterations in the adaptive algorithm.



Neglecting the higher order term does not cause any trouble in this example.

Numerical examples: m(X)=Var(X)

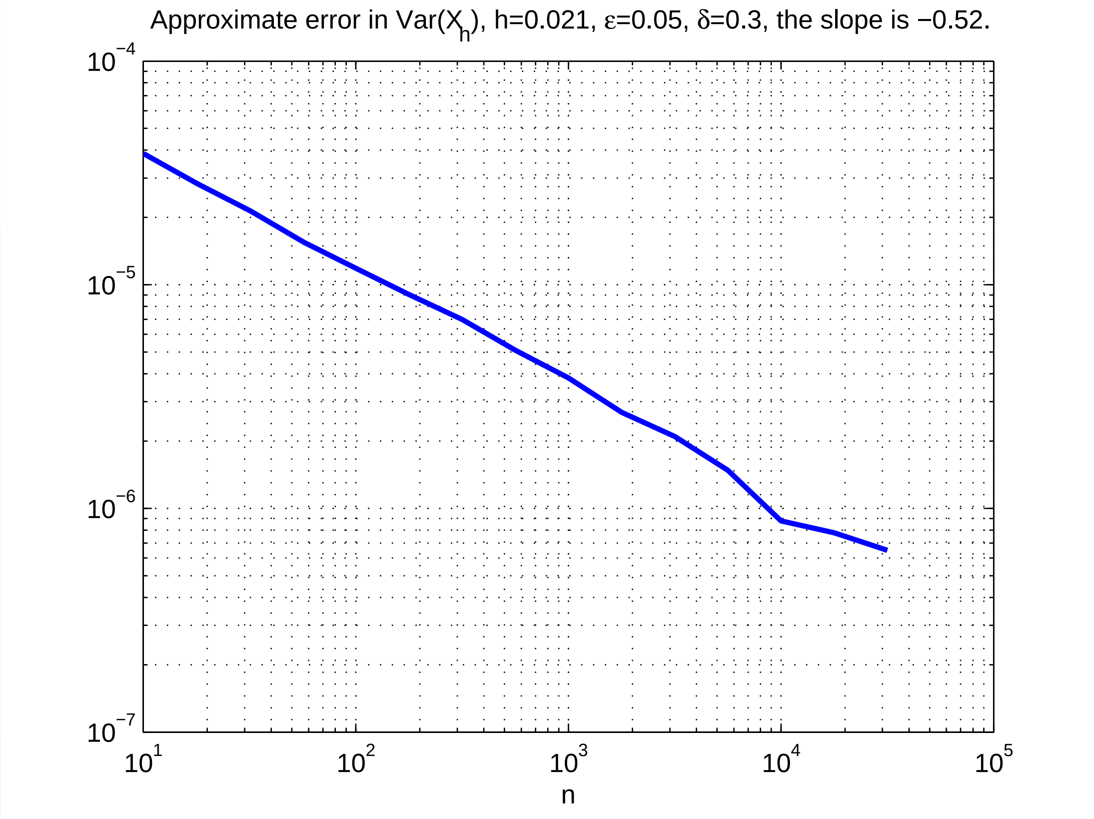
We start with the h-dependence, i.e. let n be big. Remember $P(|\text{Var}(X) - S_n^2(X_h)| < y) \geq 0.95$.



We note that the 95% probability bound of depends roughly on $h^{2.4}$.

Convergence for Var(n)

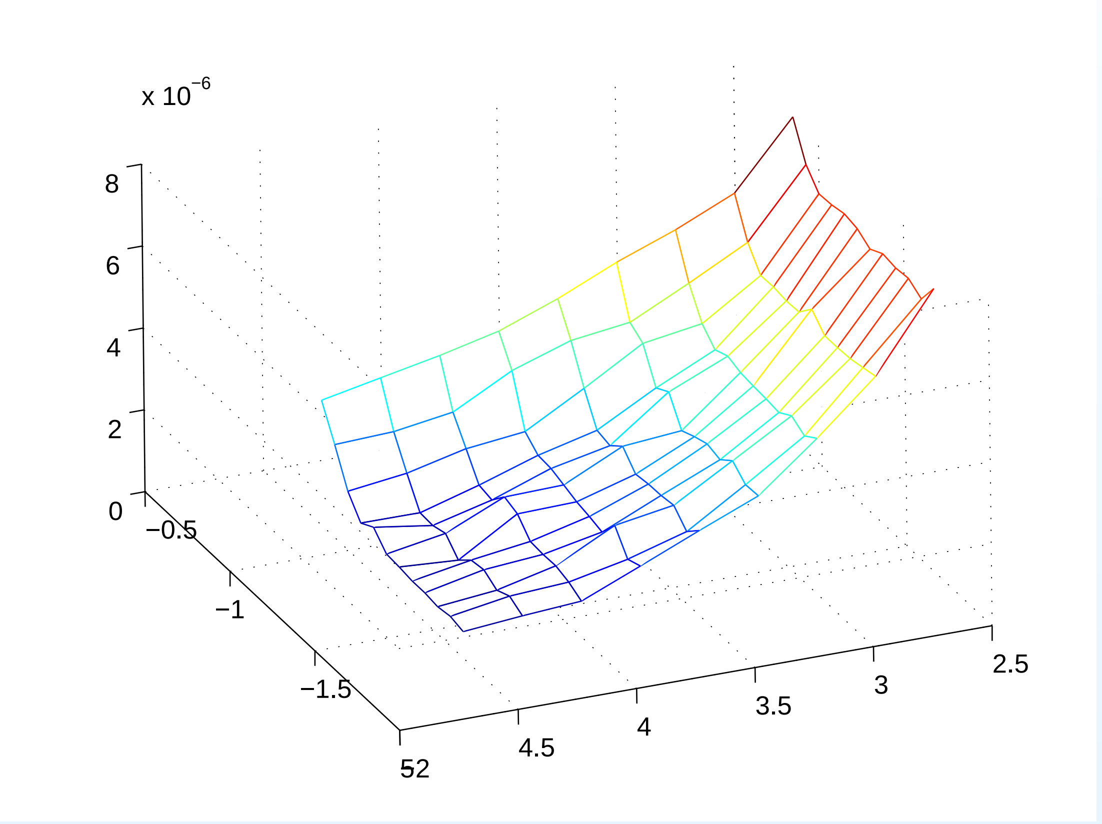
Then the n-dependence, i.e. let h be small. Remember $P(|\text{Var}(X) - S_n^2(X_h)| < y) \geq 0.95$.



We note that the 95% bound of approximately depends on $1/\sqrt{n}$.

Test of convergence in h and n

Next we study the surface we get from varying both h and n.

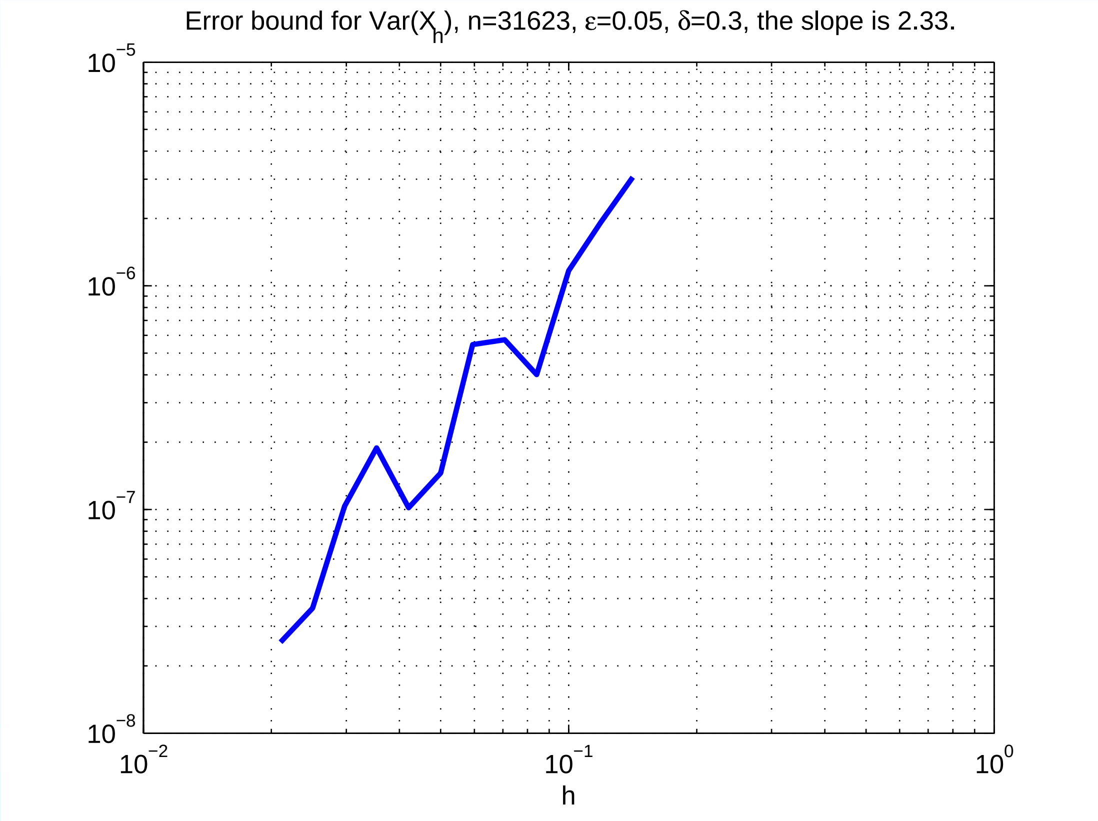


The 95% *probability bound of* $|Var(X) - S_n^2(X_h)|$ *versus* h *and* n*.*

Error bound, h-part no reference

The h-dependent part. We let $\gamma = 0.5$.

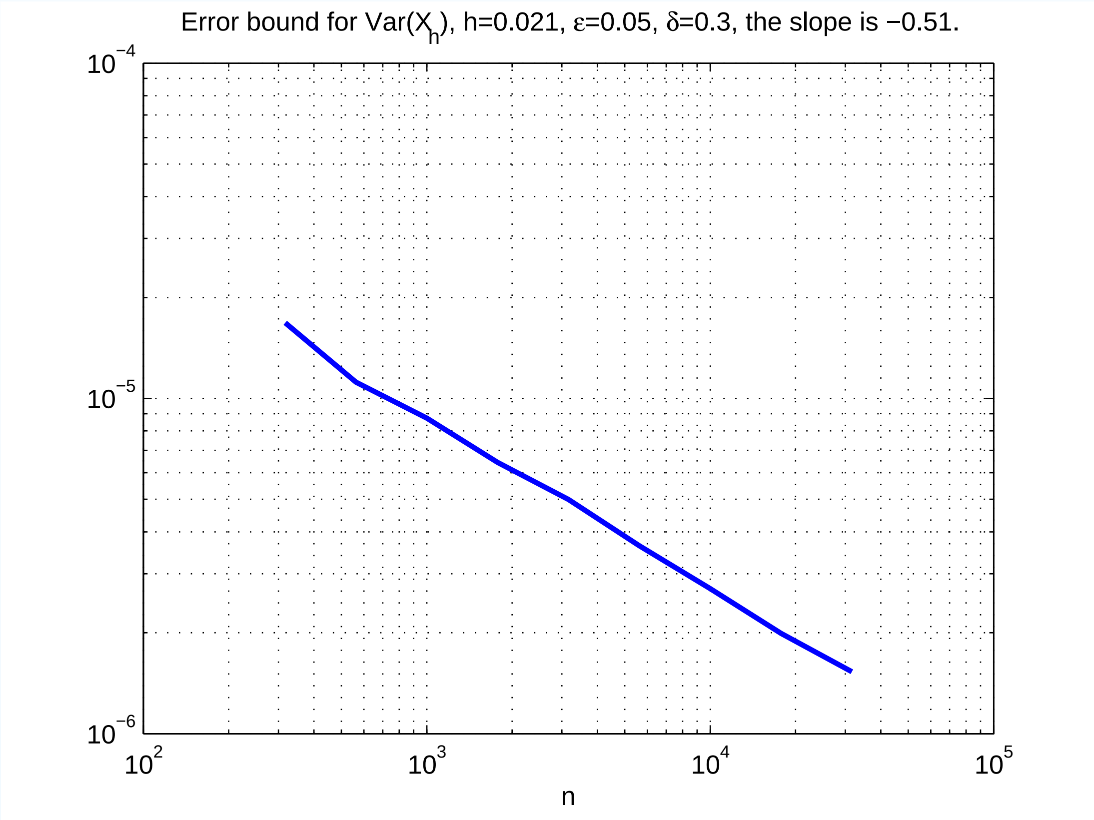
$$P(|\text{Var}(X) - S_n^2(X_h)| < C_V/\sqrt{\epsilon n} + |M(X_{\gamma h}) - M(X_h|) \geq 0.95.$$



Error bound, n-part no reference

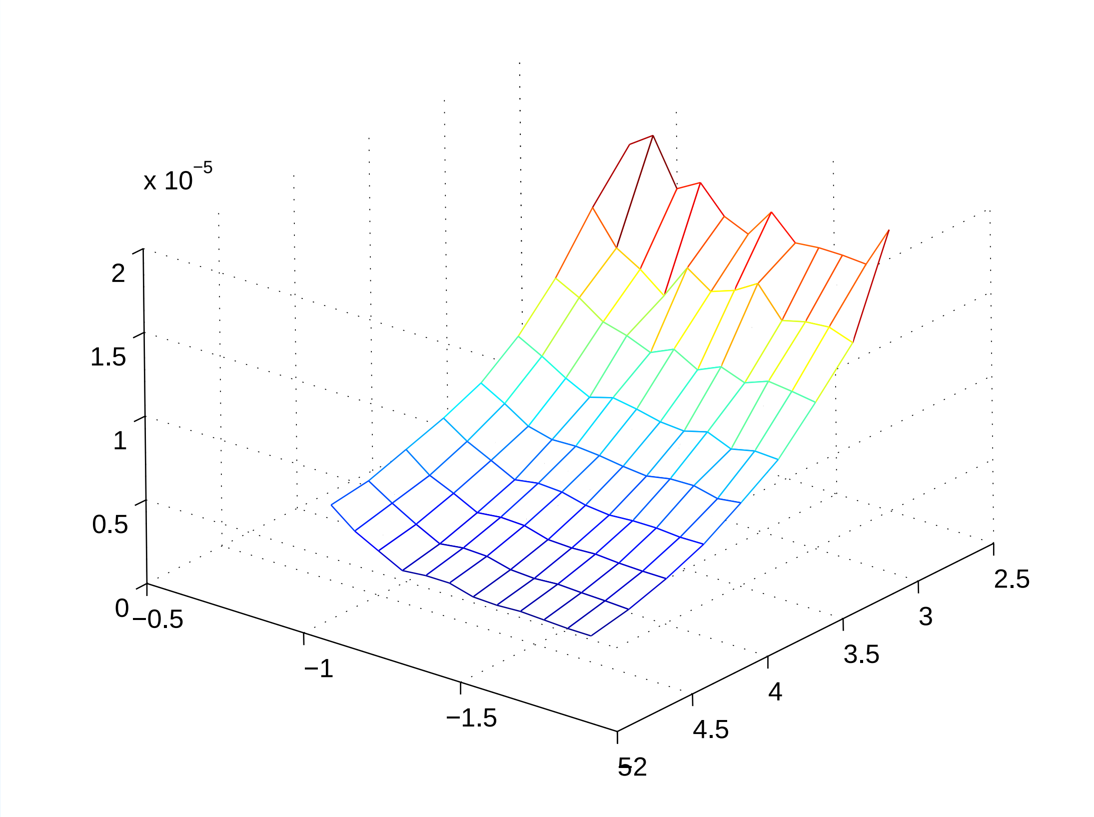
The n-dependency is again easier to capture.

$$P(|\text{Var}(X) - S_n^2(X_h)| < C_V/\sqrt{\epsilon n} + |M(X_{\gamma h}) - M(X_h|) \geq 0.95.$$



Error bound varying both h and n

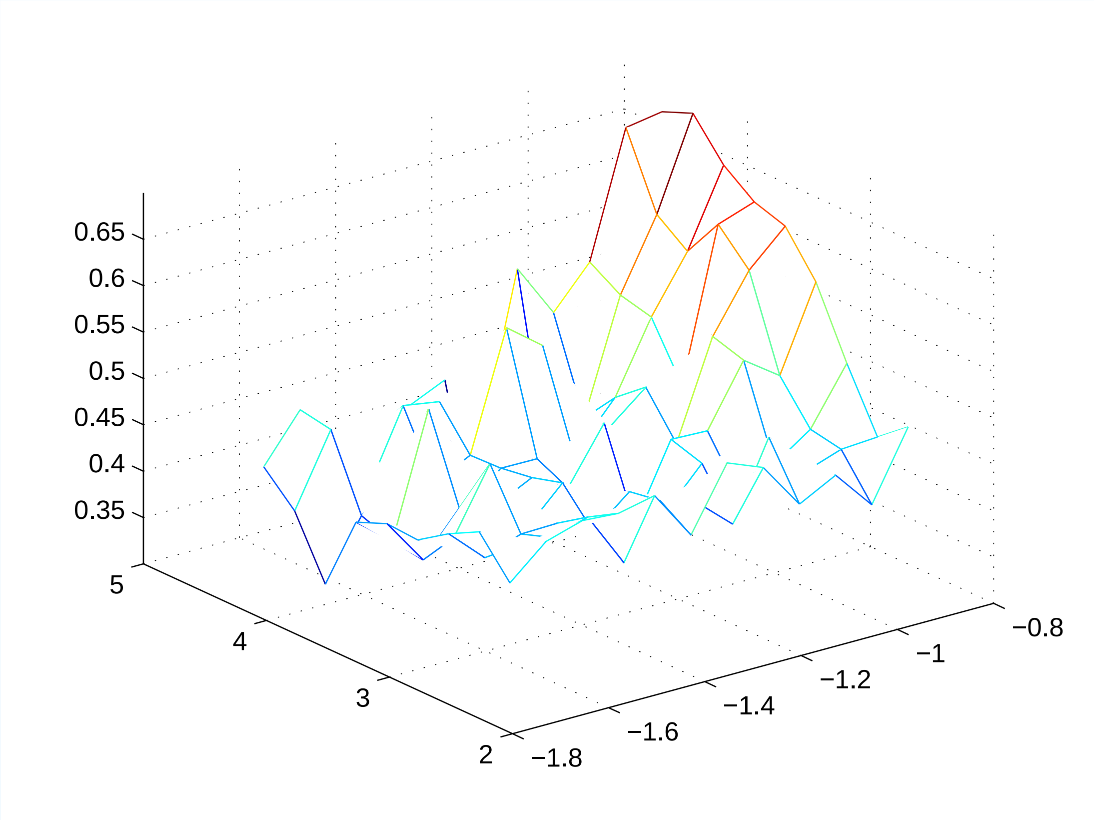
We plot the surface of the bound we get by varying h and n.



The error bound of $|$*Var*$(X) - S_n^2(X_h)|$.

Efficiency

Efficiency of the estimate. Pick the 95% worst value of $|\text{Var}(X) - S_n^2(X_h)|/(C_V/\sqrt{n\epsilon} + |M(X_{\gamma h}) - M(X_h|)$.

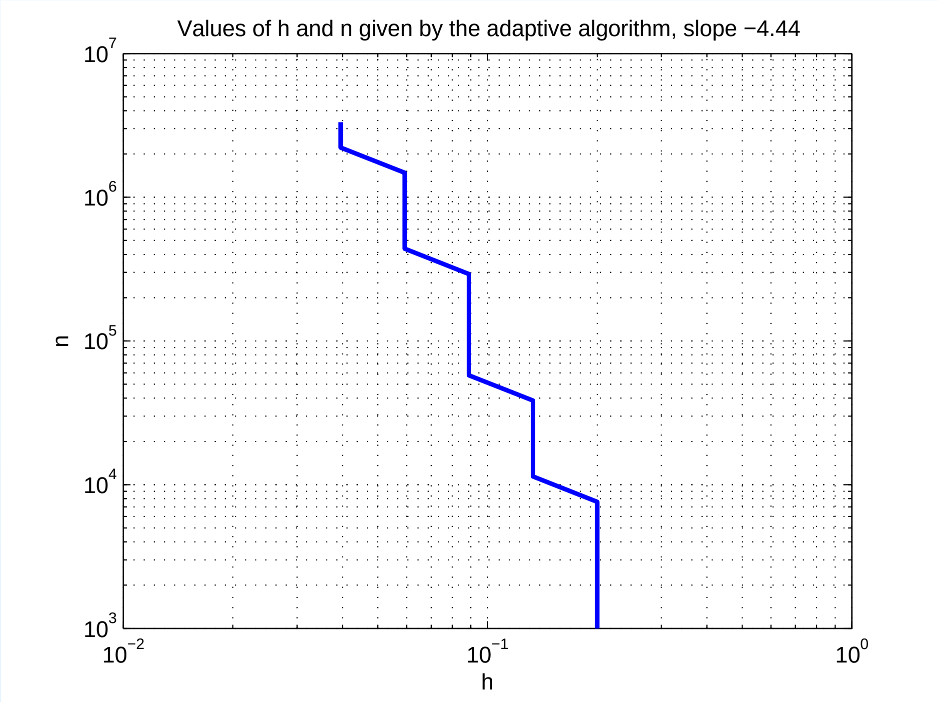


The error compared to the reference solution divided by the error bound.

Adaptive algorithm

$$P(|\text{Var}(X) - S_n^2(X_h)| < C_V/\sqrt{\epsilon n} + |M(X_{\gamma h}) - M(X_h|) \geq 0.95.$$

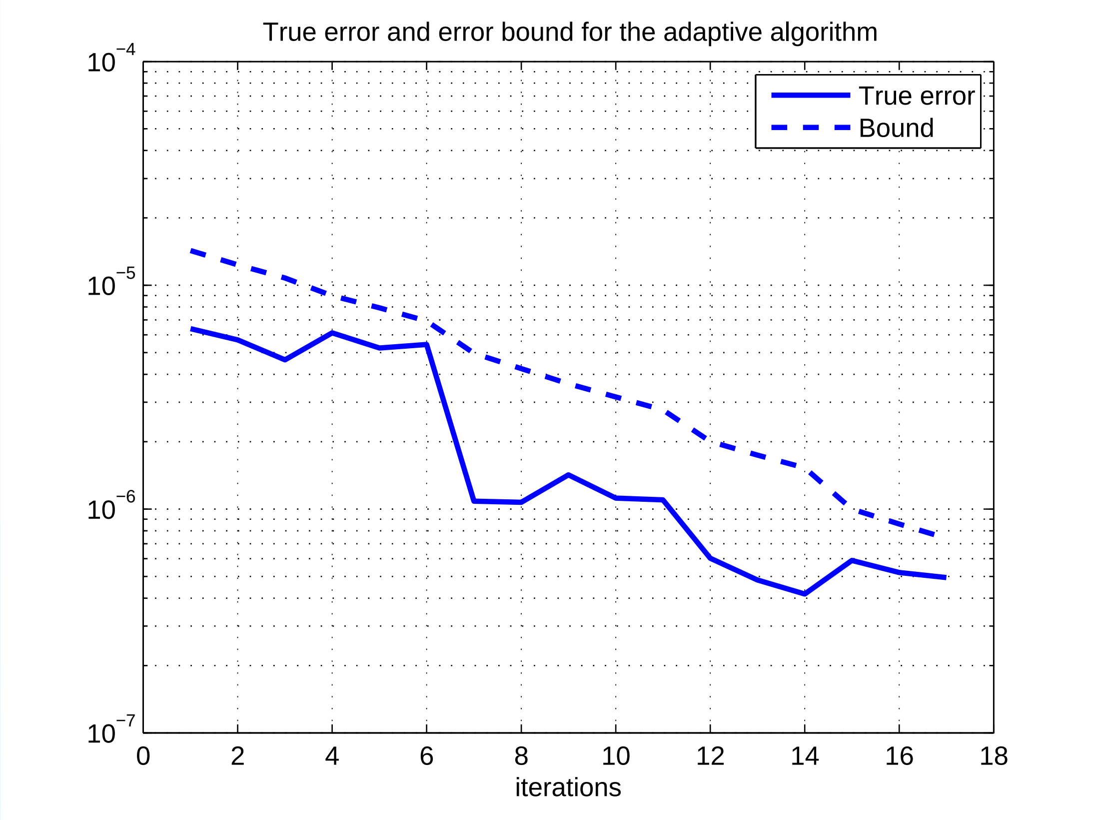
Let $\epsilon = 0.05$, $r = 1.5$, $TOL = 10^{-7}$, (relative error $< 0.2\%$), $h_{\text{start}} = 0.2$, and $n_{\text{start}} = 10^3$.



We see how the algorithm gives us roughly $n \sim h^{-4.4}$.

Error bound versus ”true” error

We compare the error bound with the error compared to a reference solution.



The solid line is error compared to a reference solution and the dashed line is the error bound.

Conclusion and related projects

- Quick method for computing arbitrary stochastic quantities of linear functionals of solution to a wide class of linear PDE's.
- Error analysis that takes both the discretization error and the stochastic error into account.
- Numerical results that agrees with theory.

Lately we have also studied random perturbation in the diffusion coefficient,

$$-\nabla \cdot \mathcal{A}^s \nabla U^s = f \quad \text{in } \Omega,$$
$$U^s = 0 \quad \text{on } \Gamma,$$

with applications in oil reservoir simulation.