Performance analysis of

Bridge Monte-Carlo Estimator

Oleg V. Lukashenko†, **Evsey V. Morozov**†, **Michele Pagano**♮

† **Karelian Research Centre of RAS, Petrozavodsk State University, Russia**

This work is partly supported by Russian Foundation for Basic research, project No 10-07-00017

♮ **Dipartimento di Ingegneria dell'Informazione, Università di Pisa, Italy**

Trondheim

June 2012

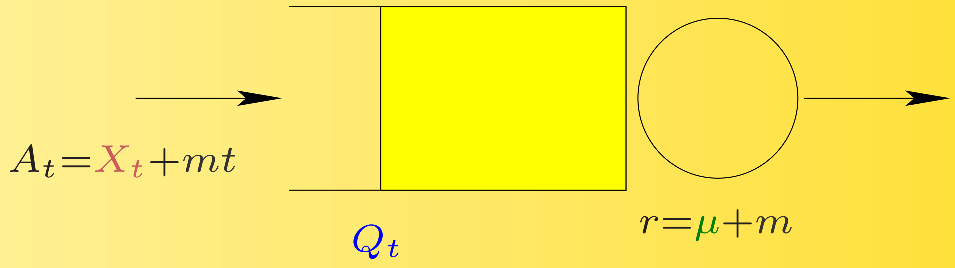
Outlines

- Network motivations and theoretical background
  - Performance issues in broadband networks
  - Problem statement and system parameters
  - Overview of key LDT concepts
- **Bridge Monte Carlo (BMC)**
  - Definition and heuristic interpretation
  - Asymptotic approximation of the estimator expression
- Simulation Results
  - Analysis with different input processes: FBM, combination of FBMs, IOU
  - Comparison with LDT asymptotics
  - Justification of the asymptotic approximation
- Conclusions

- Broadband traffic exhibits Long Range Dependence (LRD), which has a deep impact on performance
- Wide-area networks handle heterogeneous traffic flows with a variety of Quality of Service (QoS) requirements and a primary QoS parameter is the Packet Loss Rate
- Typical values of the loss rate can be very small and therefore hard to estimate through standard Monte Carlo simulation
- We focus on the efficient simulation of a single server queue equipped with an infinite buffer and fed by Gaussian inputs
  - Flexibility and parsimony: a broad range of correlation structures can be described by few parameters
  - Possibility of accurately modelling network data traffic
  - Central-limit-type arguments: in a wide-area network a large number of independent sources are multiplexed and it is reasonable to argue that the aggregate traffic converges to a Gaussian process
  - Fractional Brownian Motion (FBM) has become a canonical model in the context of LRD traffic
  - Integrated Ornstein-Uhlenbeck process (IOU)

Problem statement and system parameters

☞ We refer to a single server queue



➾ Input traffic

$$A_t = X_t + mt$$

➠ m is the mean input rate

➠ $\{X_t\}_t$ is a random centred Gaussian Component with variance $v_t \triangleq \mathbb{D}X_t$

Covariance function $\Gamma_{ts} \triangleq \mathbb{E}[X_t X_s] = \frac{1}{2}\left[v_t + v_s - v_{|t-s|}\right]$

➾ Deterministic service rate

$$r = m + \mu \quad \text{with } \mu > 0$$

Overflow probability

- We consider an upper bound for the loss rate, namely the overflow probability, defined as *the probability that the steady-state queue-length Q exceeds a given threshold b*
- From Lindley's recursion, the overflow probability can be rewritten as

$$\mathbb{P}(Q \geq b) = \mathbb{P}\left(\sup_{t\geq 0}(X_t - \mu t) \geq b\right) = \mathbb{P}\left(\sup_{t\geq 0}(X_t - \varphi_t) \geq 0\right) \quad \text{where } \varphi_t = b + \mu t$$

  - In general, this probability has not a closed form
  - In applications (finance, telecommunications) usually it is very small
- To study the behaviour of the estimators when the probability of interest is small, we introduce a *smallness* parameter ε in the problem and consider the probabilities p_ε defined as

$$p_\varepsilon = \mathbb{P}\left(\sup_{t\in\mathcal{I}}(\varepsilon X_t - \varphi_t) \geq 0\right) = \mathbb{P}(A_\varepsilon)$$

  where $\mathcal{I}$ is a finite (simulation horizon is finite) index set: the process X is just a random vector in $\mathcal{X} = \mathbb{R}^n$, where $n = |\mathcal{I}|$ (cardinality of $\mathcal{I}$)
- For the trivial MC estimator $\widehat{p}_{\varepsilon,\text{MC}}$ it is easy to show that when $p_\varepsilon \to 0$, the number N of samples to obtain a reliable estimate grows as p_ε^{-1}

☞ *Roughly*, the Large Deviation Principle for Gaussian Processes states that given an event B

$$-\varepsilon^2 \log \mathbb{P}(\varepsilon X \in B) \simeq \frac{1}{2} \inf_{\rho \in B} |\rho|^2_{\mathcal{H}} \qquad \text{as} \quad \varepsilon \to 0$$

For a finite-dimensional Gaussian process X we have the explicit expression

$$|\rho|^2_{\mathcal{H}} = \langle \rho, \rho \rangle_{\mathcal{H}} = \langle \rho, \Gamma^{-1} \rho \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_i \rho_j (\Gamma^{-1})_{i\,j}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean scalar product of $\mathcal{X}$ and Γ^{-1} is the inverse of the $n \times n$ covariance matrix $\{\Gamma_{ij}\}_{i,j=1,\ldots,n}$

☞ Heuristics behind Large Deviations for Gaussian processes

In the finite dimensional case $X \in \mathbb{R}^n$, $B \subset \mathbb{R}^n$ and

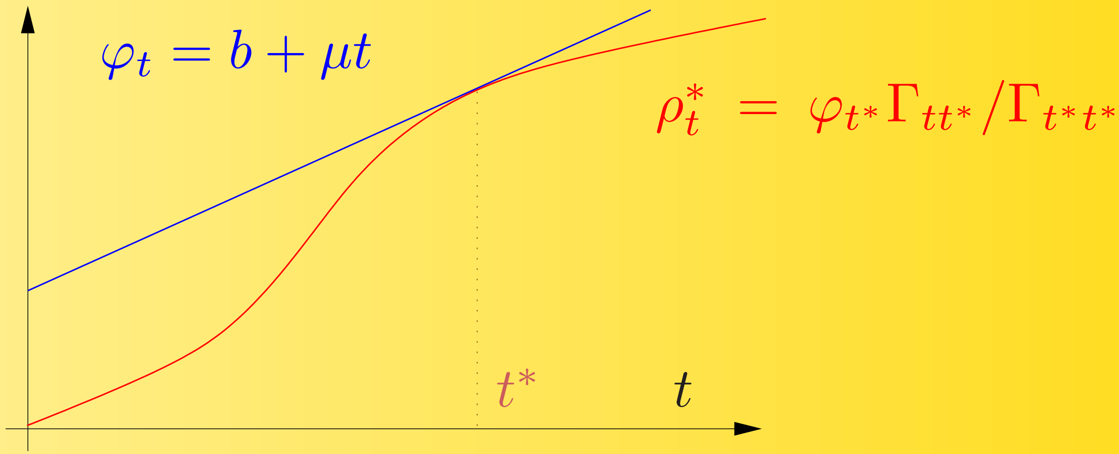
$$\mathbb{E}\left[1_{\{\varepsilon X \in B\}}\right] = C\varepsilon^{-n/2} \int_B e^{-\varepsilon^{-2} \frac{1}{2} |x|^2_{\mathcal{H}}} d^n x \simeq C\varepsilon^{-n/2} e^{-\varepsilon^{-2} \inf_{x \in B} \frac{1}{2} |x|^2_{\mathcal{H}}}$$

The Path to Buffer Overflow

☞ For the particular structure of the event A_ε we have:

$$-\lim_{\varepsilon\to 0}\varepsilon^2\log p_\varepsilon = -\lim_{\varepsilon\to 0}\varepsilon^2\log\mathbb{P}(A_\varepsilon) = \frac{1}{2}\inf_{\rho\in A}|\rho|^2_{\mathcal{H}} = \inf_{t\in\mathcal{I}}\frac{\varphi_t^2}{2\Gamma_{tt}} \triangleq \inf_{t\in\mathcal{I}} I_t = \frac{1}{2}|\rho^*|^2_{\mathcal{H}}$$

- The value t^* of t which minimizes I_t is called most-likely time
- The value of ρ which reaches the minimum is the most-likely path ρ^*: in the large deviation regime, the majority of the samples of the process which attains the level φ are concentrated around ρ^*



The most-likely time t^* can be evaluated when X is a FBM $\Rightarrow$ $t^* = \dfrac{bH}{\mu(1-H)}$

The Bridge Monte-Carlo (BMC) approach

- An alternative approach can be derived by expressing the overflow probability as the expectation of a function of the Bridge Y of the Gaussian process X

- The Bridge Y is the process obtained by conditioning X to reach a certain level (in our case the level 0) at some prefixed time τ; in the following we will assume that $\tau = t^*$

The Gaussian Bridge

☞ Fix τ and consider the following centred Gaussian process

$$Y_t = X_t - \psi_t X_\tau \qquad \text{where} \quad \psi_t \triangleq \frac{\Gamma_{t\tau}}{\Gamma_{\tau\,\tau}}$$

and suppose that $\psi_0 = 0$ and $\psi_t > 0$ for all other values of $t \in \mathcal{I}$

➾ These conditions are automatically satisfied if v_t is an increasing function

☞ The joint process (X, Y) is still Gaussian and the process Y is independent of X_τ since

$$\text{Cov}(X_\tau, Y_t) = \mathbb{E}[X_\tau Y_t] = \Gamma_{\tau t} - \frac{\Gamma_{t\tau}}{\Gamma_{\tau\tau}}\Gamma_{\tau\tau} = 0 \quad \text{for any} \quad t$$

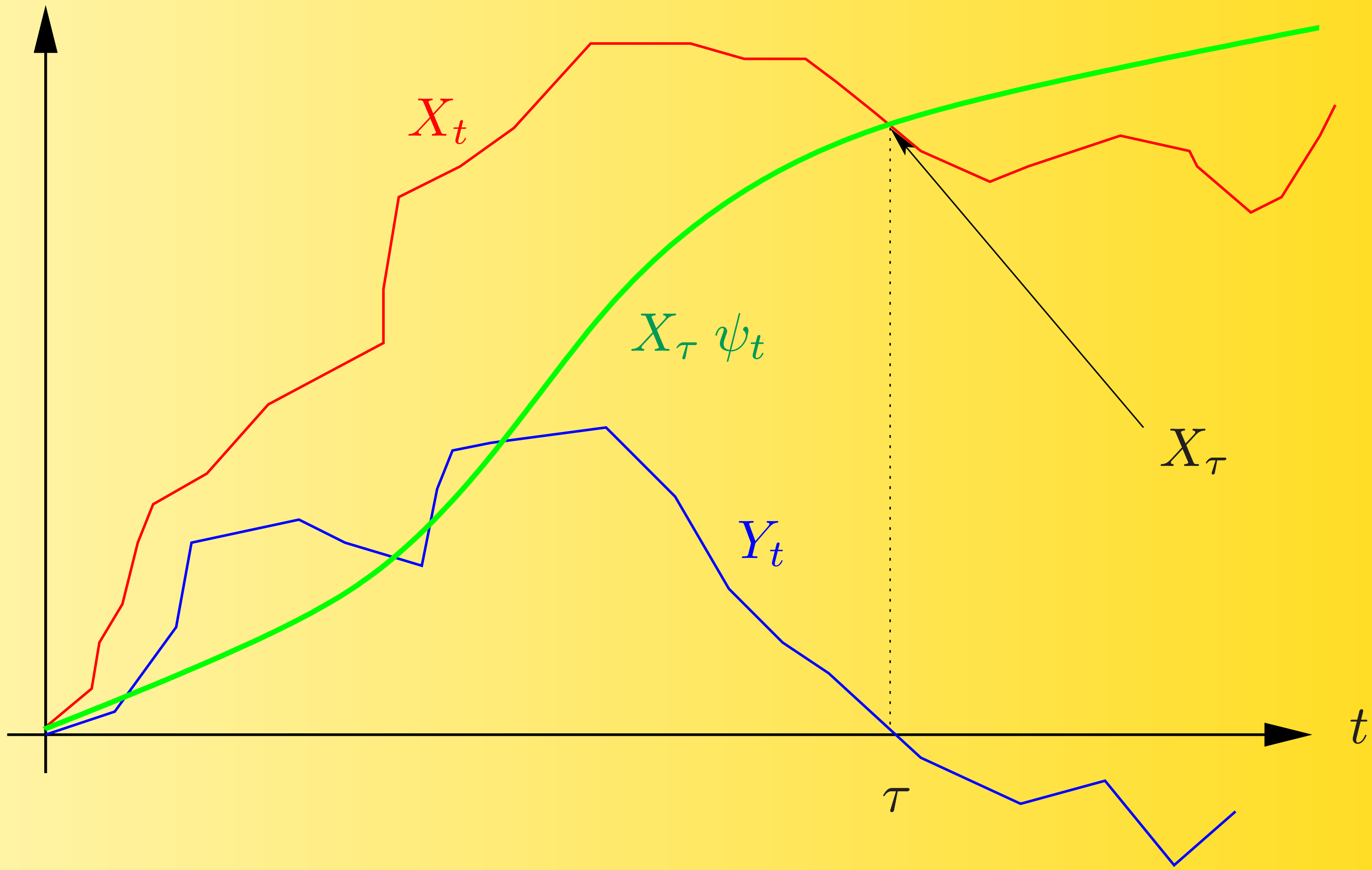
☞ Y has covariance function $\widetilde{\Gamma}$ given by

$$\widetilde{\Gamma}_{ts} = \Gamma_{ts} - \frac{\Gamma_{t\,\tau}\Gamma_{s\,\tau}}{\Gamma_{\tau\,\tau}}$$

☞ BMC does not rely on any change of measure

A graphical interpretation of the Gaussian Bridge

$$Y_t = X_t - \psi_t X_\tau$$



The Gaussian Bridge at Work

☞ We can express the probability $\mathbb{P}_L$ of the event $L = \left\{ \sup_{t \in \mathcal{I}} [X_t - \varphi_t] \geq 0 \right\}$ as follows

$$\mathbb{P}_L = \mathbb{P}\left(\sup_{t \in \mathcal{I}} [X_t - \varphi_t] \geq 0 \right) = \mathbb{P}\left(\sup_{t \in \mathcal{I}} [Y_t + \psi_t X_\tau - \varphi_t] \geq 0 \right)$$
$$= \mathbb{P}\left(\inf_{t \in \mathcal{I}} (\varphi_t - Y_t - \psi_t X_\tau) \leq 0 \right)$$

☞ The events

$$A = \left\{ \inf_{s \in \mathcal{I}} \left(\varphi_s - Y_s - \psi_s X_\tau \right) \leq 0 \right\} \quad \text{and} \quad B = \left\{ \inf_{t \in \mathcal{I}} \psi_t^{-1} \left(\varphi_t - Y_t \right) \leq X_\tau \right\}$$

are equivalent (see next slide)

☞ Denote

$$\overline{Y} \triangleq \inf_{t \in \mathcal{I}} \frac{\varphi_t - Y_t}{\psi_t} \quad \text{and} \quad \Phi(x) \triangleq \int_x^\infty \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy$$

By the independence between $\overline{Y}$ and X_τ, where $X_\tau \in \mathcal{N}(0, \Gamma_{\tau\tau})$

$$\mathbb{P}_L = \mathbb{P}\left(\overline{Y} \leq X_\tau \right) = \mathbb{E}\left[\Phi\left(\frac{\overline{Y}}{\sqrt{\Gamma_{\tau\tau}}} \right) \right]$$

Equivalence between sets A and B

$$A = \left\{ \inf_{s\in\mathcal{I}} (\varphi_s - Y_s - \psi_s X_\tau) \le 0 \right\} \qquad B = \left\{ \inf_{t\in\mathcal{I}} \psi_t^{-1} (\varphi_t - Y_t) \le X_\tau \right\}$$

- Let us show that $A \subseteq B$
  - Fix $\omega \in A$ and let $s^* = \mathrm{argmin}(\varphi_\mathrm{s} - \mathrm{Y_s} - \psi_\mathrm{s}\mathrm{X}_\tau)$
  - Then

$$\varphi_{s^*} - Y_{s^*}(\omega) - \psi_{s^*} X_\tau(\omega) = \inf_{s\in\mathcal{I}} (\varphi_s - Y_s(\omega) - \psi_s X_\tau(\omega)) \le 0$$

  - Consequently

$$\inf_{t\in\mathcal{I}} \psi_t^{-1}[\varphi_t - Y_t(\omega)] \le \psi_{s^*}^{-1}[\varphi_{s^*} - Y_{s^*}(\omega)] \le X_\tau(\omega) \quad \Rightarrow \omega \in B$$

  - Then, $A \subseteq B$
- Similarly it is easy to check that $B \subseteq A$
- Thus these two events are equivalent

The bridge estimator

- MC can be seen as a numerical scheme to perform integration in a large number of variables
- BMC performs one of these integrations exactly exploiting properties of Gaussian processes.

  The rest of the integrations are still performed using a MC scheme

$$\mathbb{P}_L = \mathbb{E}\left[\Phi\left(\frac{\overline{Y}}{\sqrt{\Gamma_{\tau\,\tau}}}\right)\right]$$

- In the full space the characteristic function of the rare event has support on a region with small probability and this renders MC ineffective. However BMC smooth out the function to be integrated allowing a more efficient estimation by the MC part
- Given an iid sequence $\{Y^{(i)}, i = 1, \ldots, N\}$ distributed as Y, the bridge estimator $\widehat{p}^N$ for $\mathbb{P}_L$ is

$$\widehat{p}^N \triangleq \frac{1}{N}\sum_{i=1}^{N}\Phi\left(\frac{\overline{Y}^{(i)}}{\sqrt{\Gamma_{\tau\,\tau}}}\right) \qquad \text{where} \qquad \overline{Y}^{(i)} \triangleq \inf_{t\in\mathcal{I}} \frac{\varphi_t - Y_t^{(i)}}{\psi_t}$$

Asymptotic approximation of the BMC estimator

☞ For *some values* of the system parameters, the infimum in the expression of

$$\overline{Y} \triangleq \inf_{t \in \mathcal{I}} \frac{\varphi_t - Y_t}{\psi_t} \triangleq \inf_{t \in \mathcal{I}} G_t$$

is attained *near* the most-likely time τ, with $G_\tau = \varphi_\tau$

☞ Assume that $\overline{Y}^{(i)} \in [\varphi_\tau - h, \varphi_\tau]$; then

$$\Phi\left(\frac{\varphi_\tau}{\sqrt{\Gamma_{\tau\,\tau}}}\right) \leq \widehat{p}^N \leq \Phi\left(\frac{\varphi_\tau - h}{\sqrt{\Gamma_{\tau\,\tau}}}\right)$$

☞ The lower bound, in the approximation $\Phi(x) \approx e^{-x^2/2}$, corresponds to the well-known LDT asymptotic bound

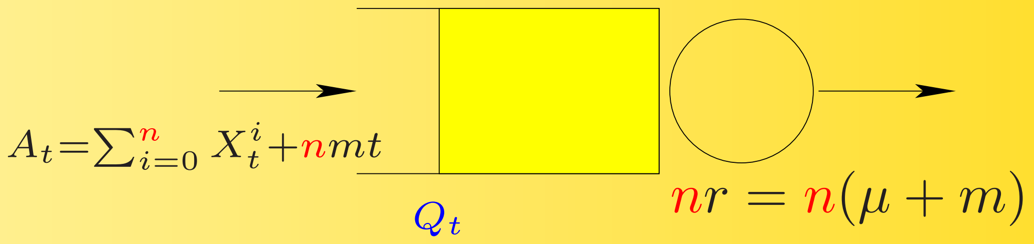
$$\mathbb{P}_L \approx e^{-\varphi_\tau^2/2\Gamma_{\tau\,\tau}}$$

☞ The difference between upper and lower bounds

$$\Delta \triangleq \Phi\left(\frac{\varphi_\tau - h}{\sqrt{\Gamma_{\tau\,\tau}}}\right) - \Phi\left(\frac{\varphi_\tau}{\sqrt{\Gamma_{\tau\,\tau}}}\right) \approx -\Phi'\left(\frac{\varphi_\tau}{\sqrt{\Gamma_{\tau\,\tau}}}\right)\frac{h}{\sqrt{\Gamma_{\tau\,\tau}}} = \frac{h}{\sqrt{2\pi\,\Gamma_{\tau\,\tau}}}\, e^{-\varphi_\tau^2/2\Gamma_{\tau\,\tau}}$$

Bridge estimator in the many-sources regime

- Many-sources regime
  - n i.i.d. Gaussian sources
  - The queueing resources (buffer size and service rate) are linearly scaled with n
  - Buffer overflow (over level nb) becomes a rare event when $n \to \infty$



- The overflow probability, in this case, is given by

$$\mathbb{P}\left(Q \geq nb\right) = \mathbb{P}\left(\sup_{t \in \mathcal{I}} \left(\sqrt{1/n}\, X_t - \varphi_t\right) \geq 0\right) \triangleq \mathbb{P}\left(\sup_{t \in \mathcal{I}} \left(\varepsilon X_t - \varphi_t\right) \geq 0\right)$$

- Given an iid sequence $\{Y^{(i)},\ i = 1, \ldots, N\}$ distributed as Y, the bridge estimator for p_ε is

$$\widehat{p}_\varepsilon^N \triangleq \frac{1}{N} \sum_{i=1}^{N} \Phi\left(\frac{\overline{Y}_\varepsilon^{(i)}}{\varepsilon \sqrt{\Gamma_{\tau\tau}}}\right) \qquad \text{where} \qquad \overline{Y}_\varepsilon^{(i)} \triangleq \inf_{t \in \mathcal{I}} \frac{\varphi_t - \varepsilon Y_t^{(i)}}{\psi_t}$$

Simulation Scenario

- Input traffic
  - Fractional Brownian Motion (FBM) $\Rightarrow \quad v_t = t^{2H}$
  - Superposition of two independent FBMs $\Rightarrow \quad v_t = t^{2H_1} + t^{2H_2}$
  - Integrated Ornstein-Uhlenbeck process (IOU) $\Rightarrow \quad v_t = t - 1 + e^{-t}$
- Number of generated sample paths: $N = 10^4$
- Comparison with LDT asymptotics
  - Large-buffer regime

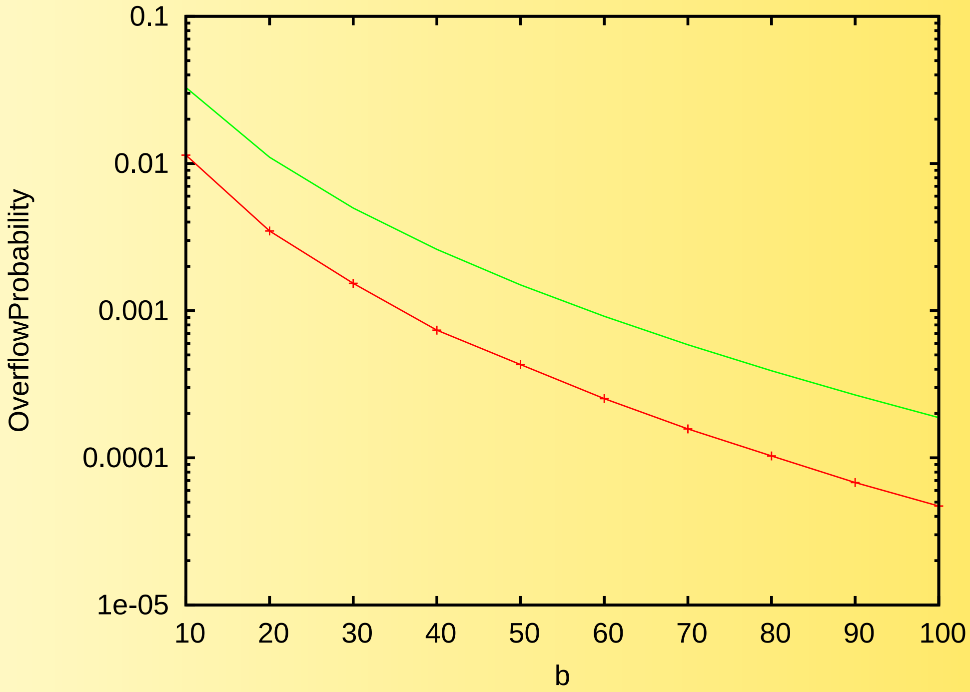
$$\log \mathbb{P}(Q > b) \approx -\inf_{t \geq 0} \frac{\varphi_t^2}{2\,\Gamma_{tt}}$$

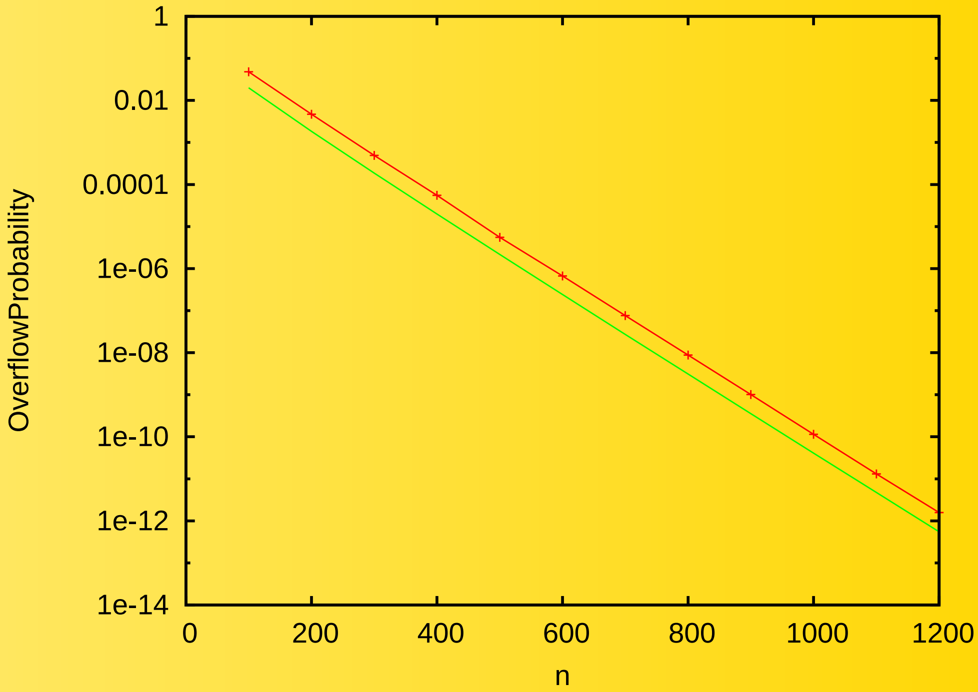
  - Many-sources regime

$$\log \mathbb{P}(Q > nb) \approx -n \inf_{t \geq 0} \frac{\varphi_t^2}{2\,\Gamma_{tt}}$$

- Analysis of the upper and lower bounds for $\widehat{p}^N$

Simulation Results: FBM (H=0.8)





Large-buffer regime

$$\mu = 1$$

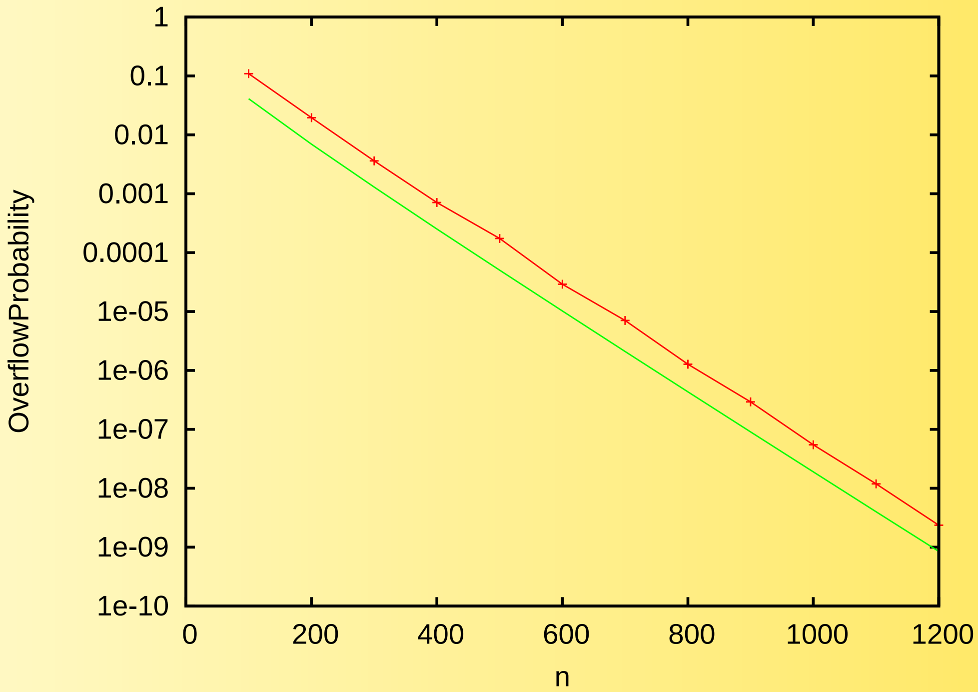
Many-sources regime

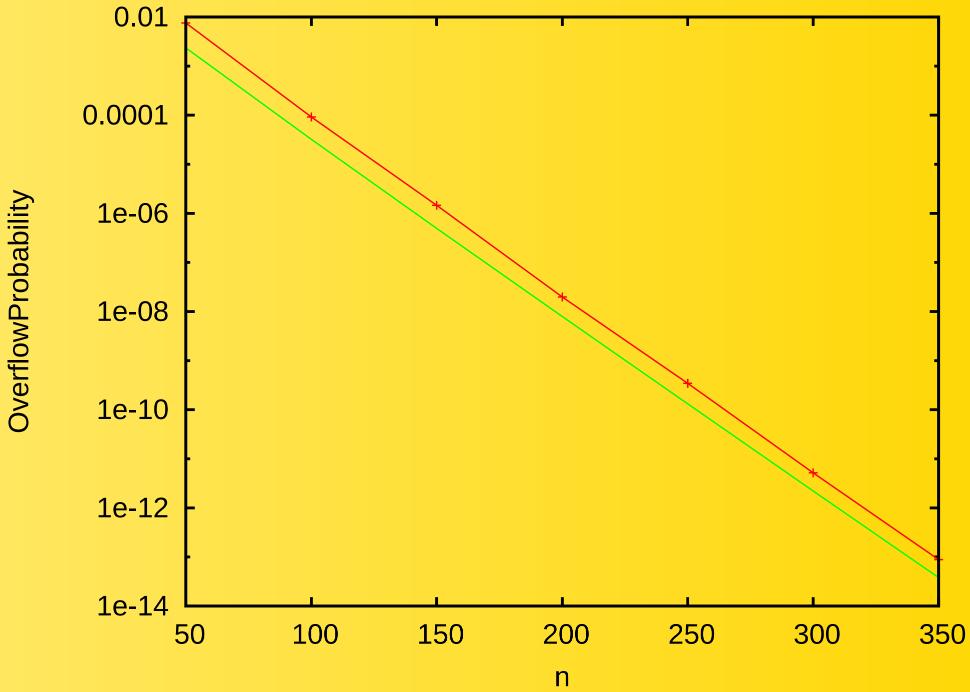
$$\mu = 0.1 \quad b = 0.3$$

For FBM, the most-likely time $t^* = \tau$ is given by $t^* = \dfrac{bH}{\mu(1-H)}$ and

$$\inf_{t \geq 0} \frac{\varphi_t^2}{2\,\Gamma_{tt}} = \frac{1}{2} \left(\frac{b}{1-H} \right)^{2-2H} \left(\frac{\mu}{H} \right)^{2H}$$

Many-sources regime





Superposition of two independent FBMs

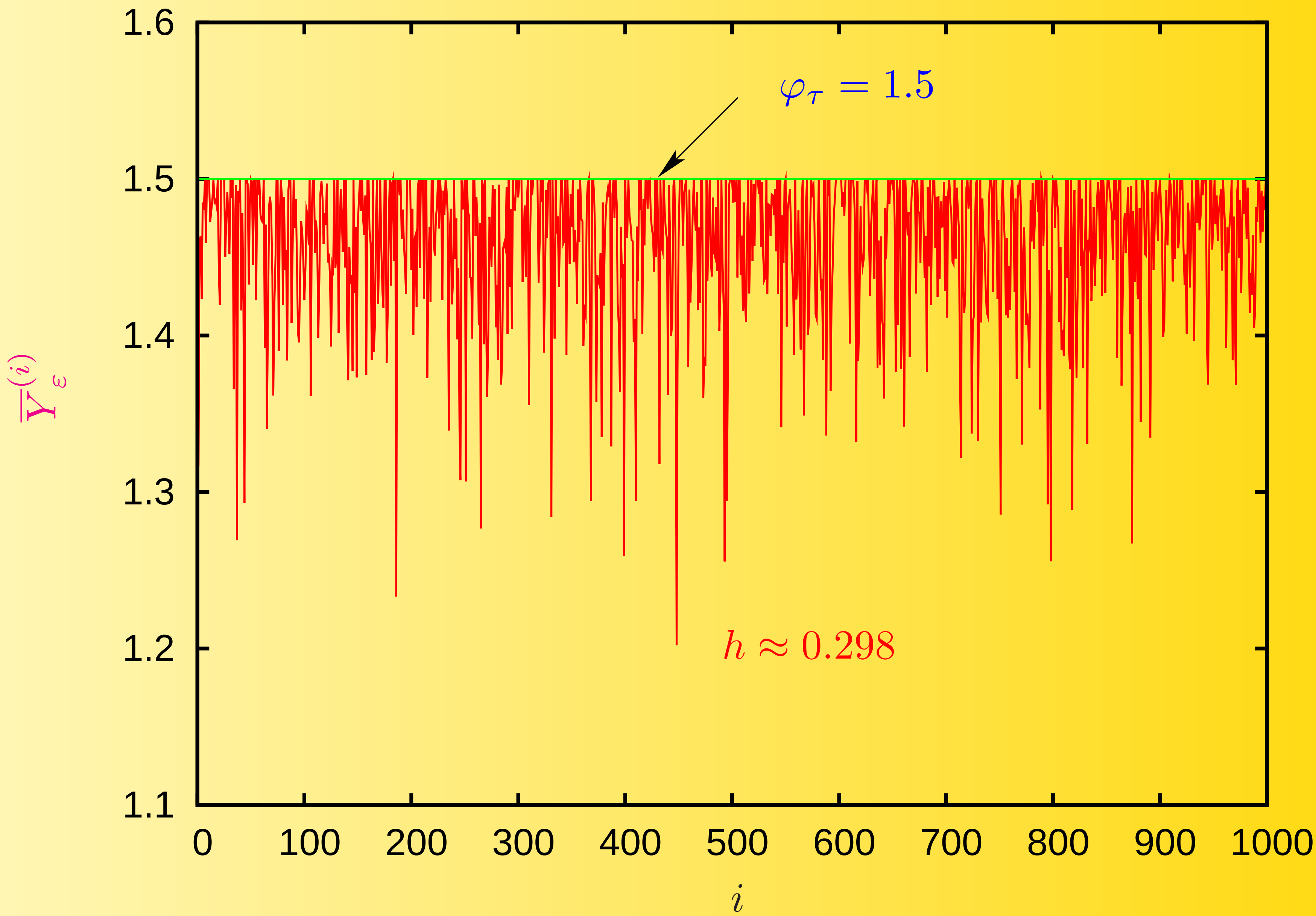
$$H_1 = 0.8 \quad H_2 = 0.6$$

$$\mu = 0.1 \quad b = 0.3$$

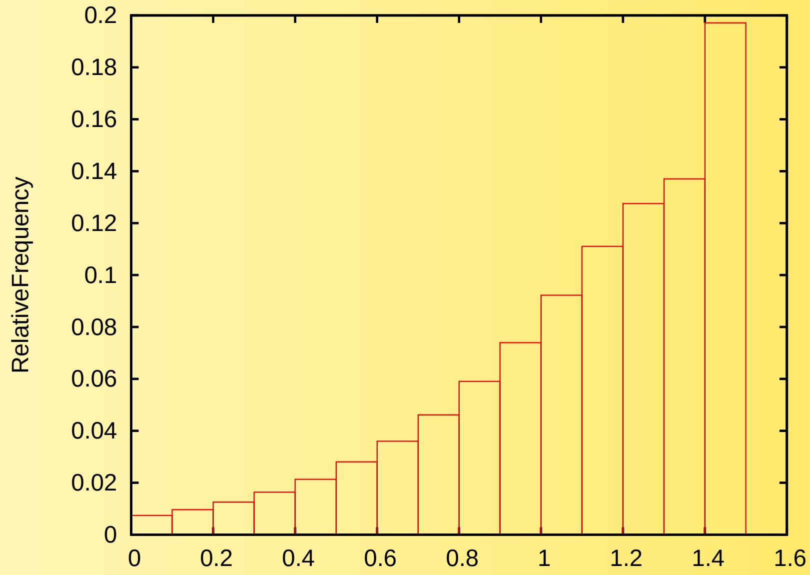
Integrated Ornstein-Uhlenbeck process

$$\mu = 0.1 \quad b = 0.3$$

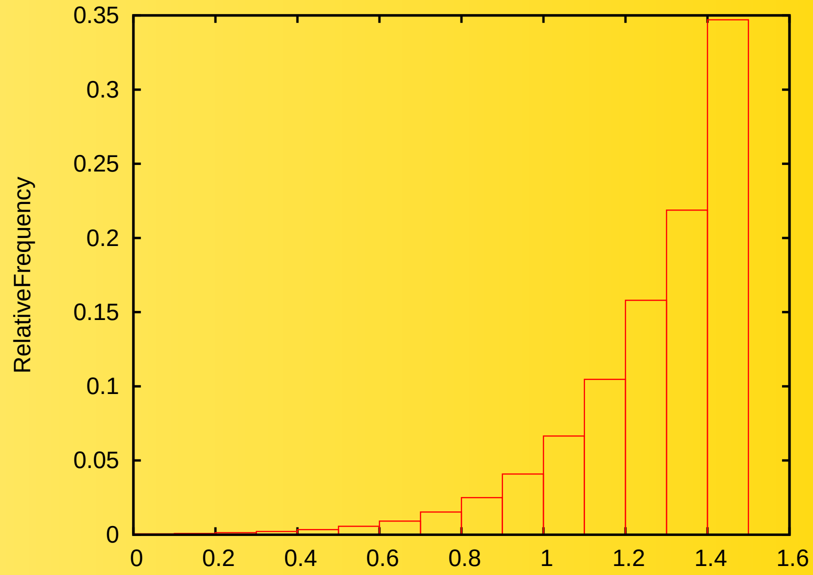
Variability of $\overline{Y}^{(i)}$ – sample path ($n = 500$)



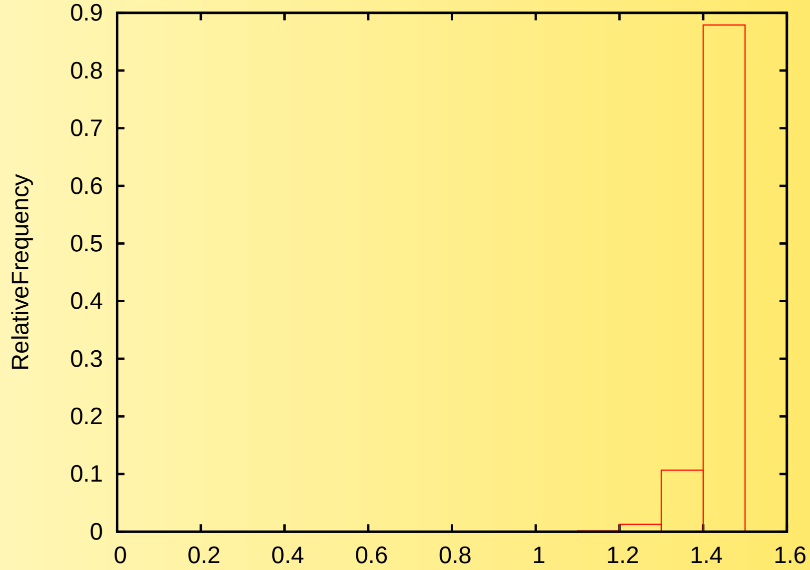
Variability of $\overline{Y}^{(i)}$ – histograms



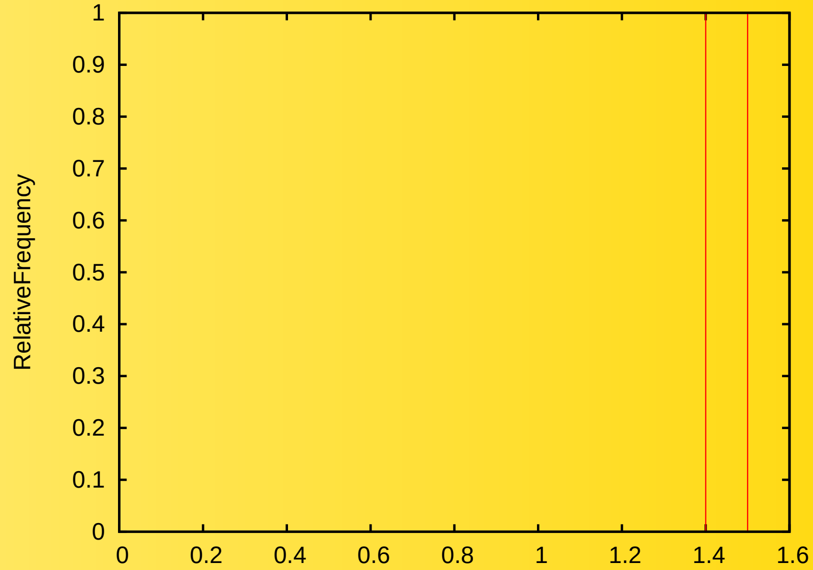
$n=50$



$n=100$

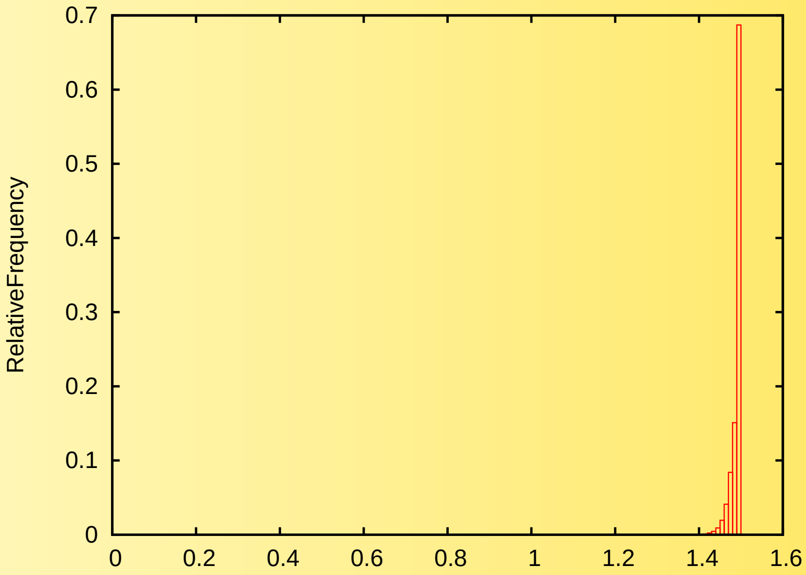


$n=500$

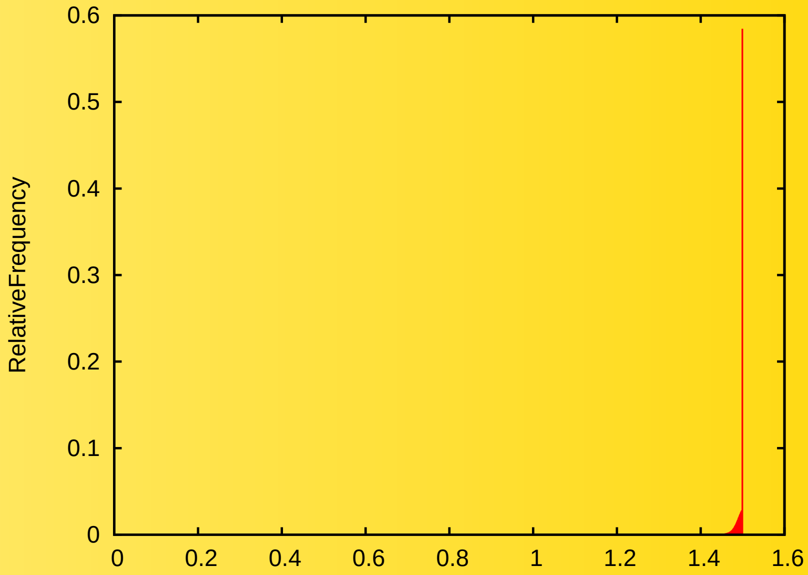


$n=2000$

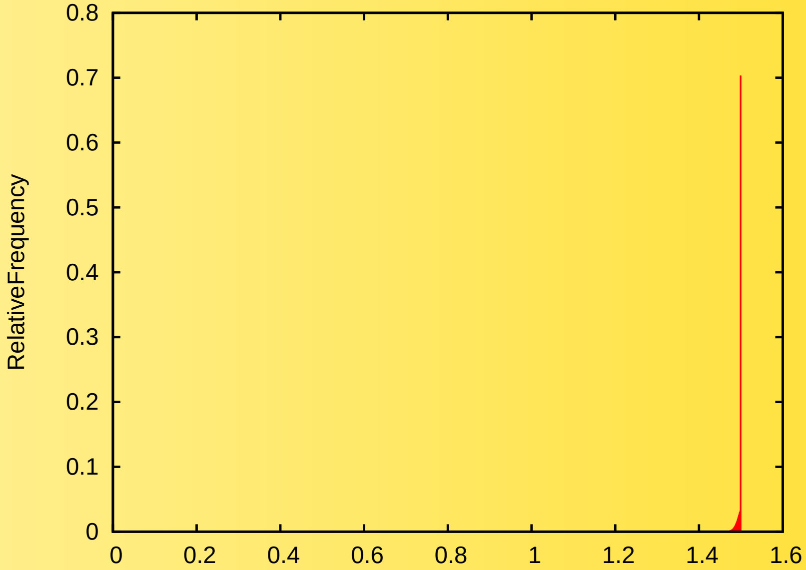
Variability of $\overline{Y}^{(i)}$ – histograms



n=2000

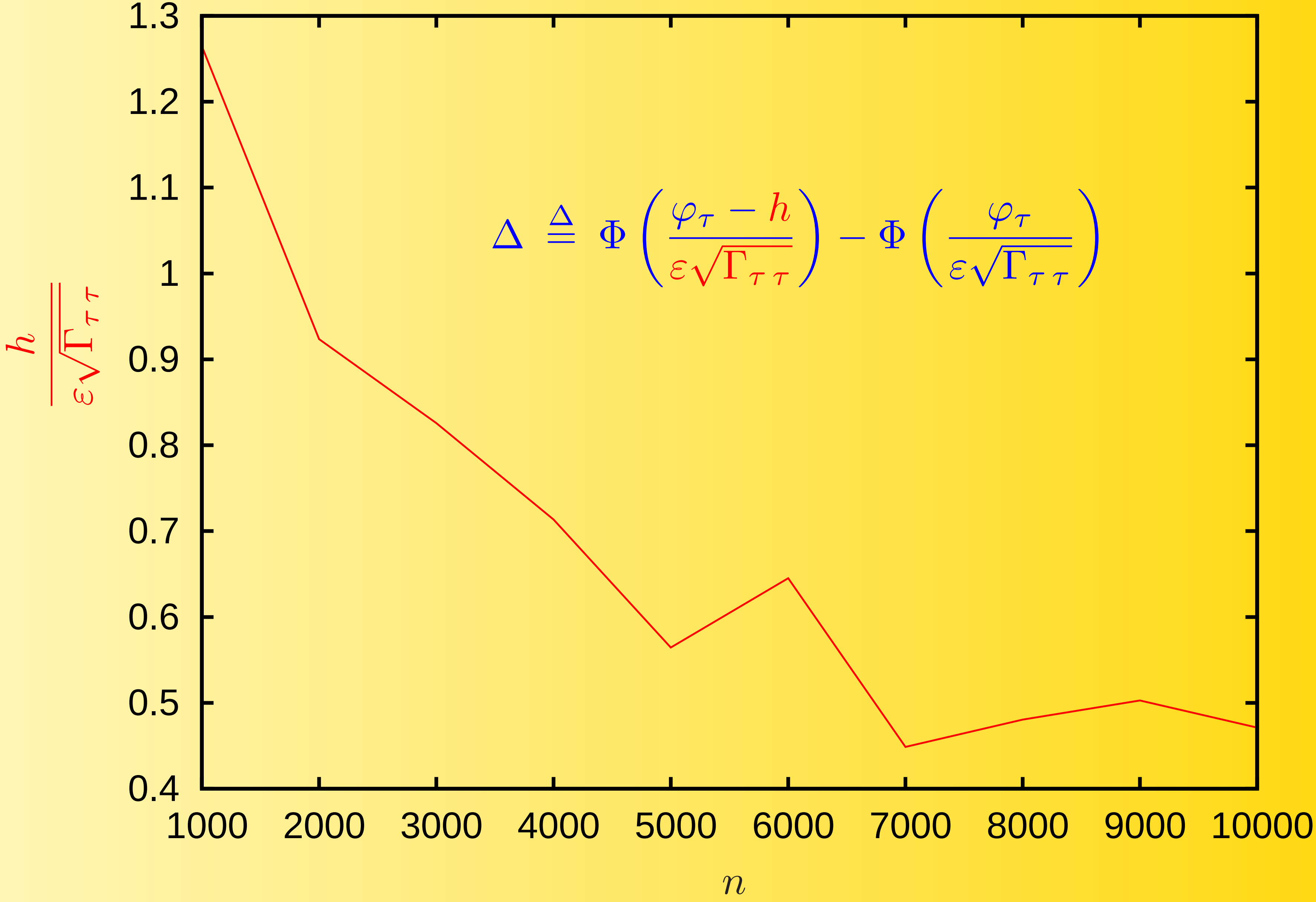


n=3000



n=5000

Variability of $\overline{Y}^{(i)}$ – normalized width of the interval



Conclusions

- We analysed the performance of the Bridge Monte-Carlo (BMC) estimator, which does not rely on Importance Sampling and does not need any refined preventive theoretical analysis
- Key features of BMC
  - Although BMC is not asymptotically efficient, for any choice of the rarity parameter ε, BMC performs *better* than single-twist IS, even when the change of measure is based on the most likely path ρ^*

    For any $\varepsilon > 0$ and any twist η of the form $\eta_t = \lambda\psi_t$ ($\alpha \in \mathbb{R}$):

    $$\sigma^2_{\text{BMC},\varepsilon} \leq \sigma^2_{\text{twist},\varepsilon}$$

  - The computational cost of BMC is comparable to that of simple IS
  - The principle underlying the BMC method can be applied to any Gaussian process
  - BMC could be generalized with more than one conditioning or with dynamic choice of the parameters
- Analysis of upper and lower bounds of the overflow probability (based on the expression of the BMC estimator)