VaR estimation and sample sizes

PillarOne Conference, Berne, September 2, 2011

Jessika Walter, PhD, Intuitive Collaboration
Markus Stricker, PhD, Intuitive Collaboration

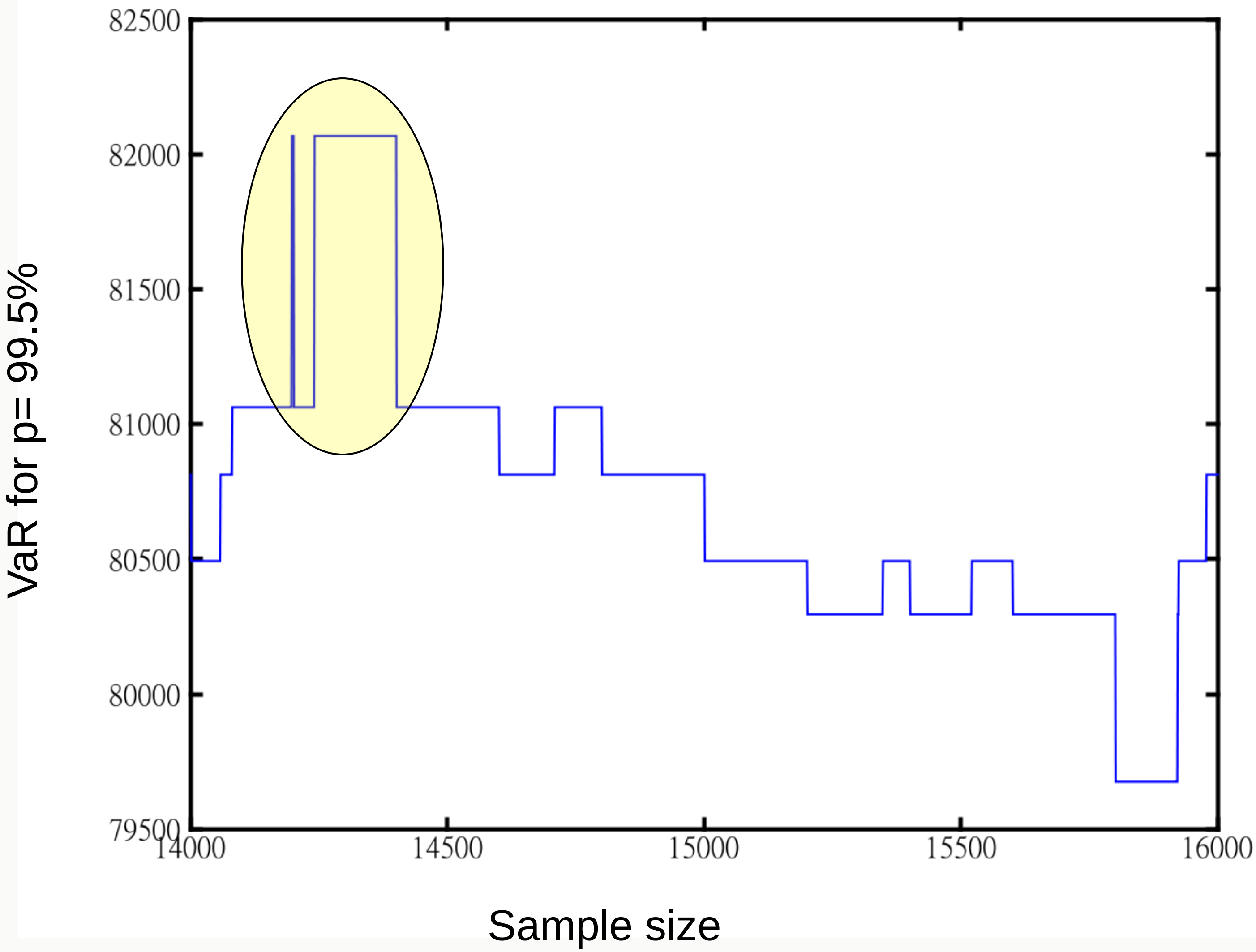
VaR fluctuations – by accident or systematic?

Distribution: Pareto, alpha = 1.2, beta = 1000

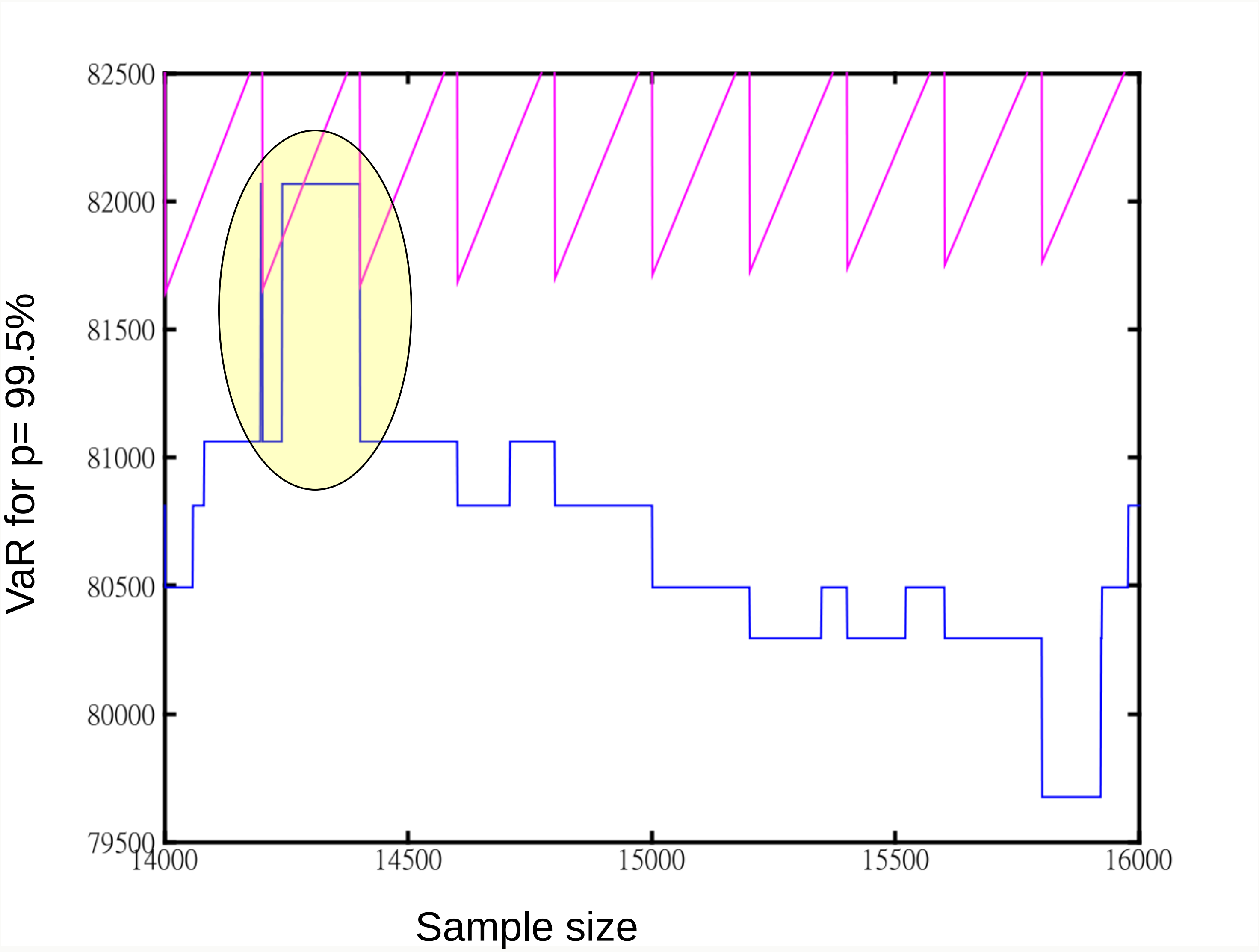
p = 99.5%

VaR derived from sample size n = 14‘250: 82‘100
VaR derived from sample size n = 14‘400: 81‘100

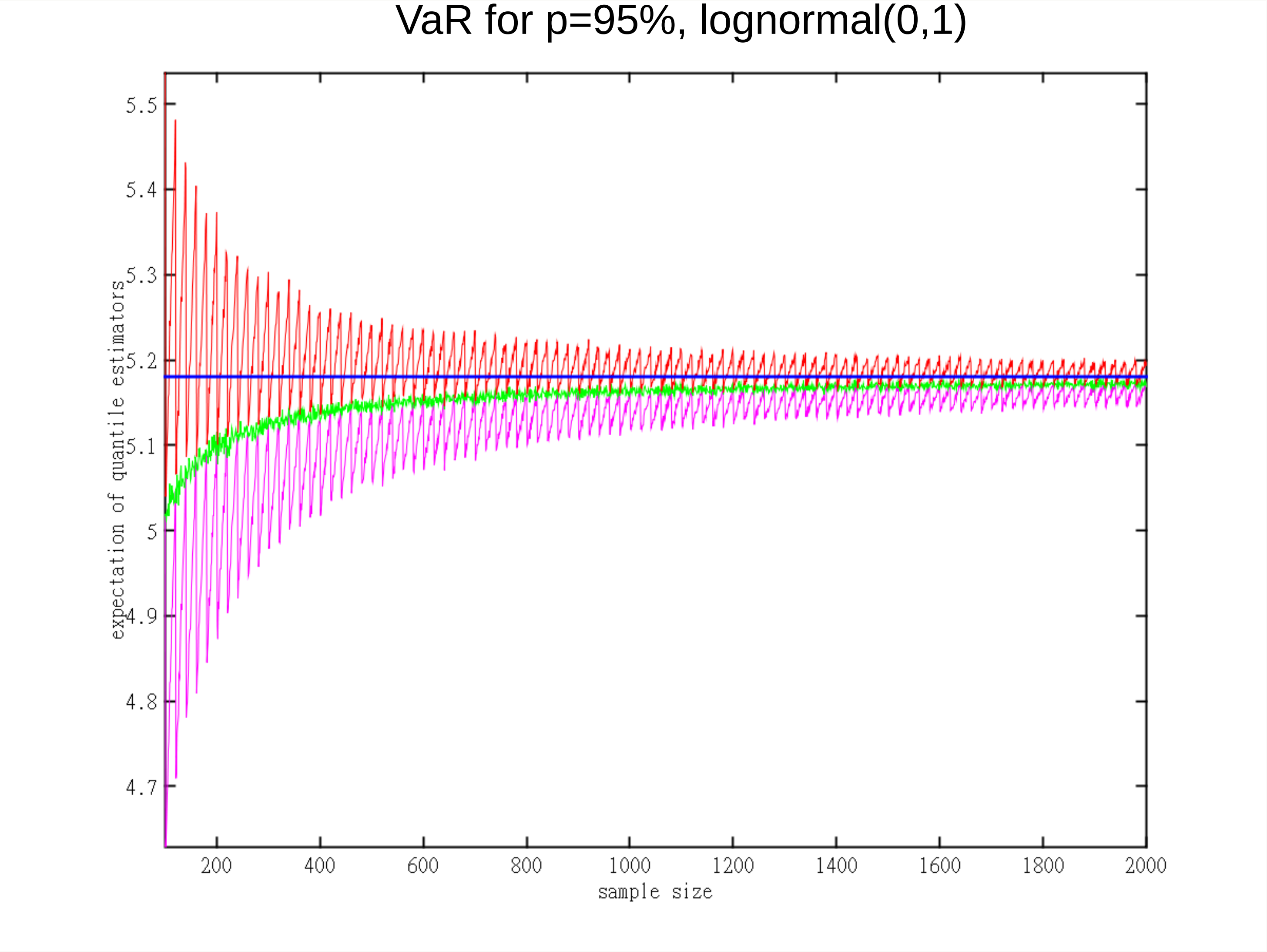
VaR fluctuations – by accident or systematic?



VaR fluctuations – by accident or systematic?



Saw-tooth oscillation of the bias of a simple one-point estimators



Content

- Decomposition of the bias for
  - one-point estimators for uniform distributions
  - smoothed two-point estimator
- Hints to construct optimized sample estimators
- Derive heuristics for required sample sizes

Results from the paper „On VaR estimation and sample sizes“, submitted to the European Actuarial Journal.

Decomposition of the bias
Uniform distribution, one point estimator

pillar1 Risk Management meets Open Source

Let $x_1, x_2, \ldots, x_n$ be a sample of size n drawn from independent, uniformly distributed random variables.

Denote with $x_{(1)}, \ldots, x_{(n)}$ the sample ordered by increasing size (order statistics).

Let p be the percentile, $g := pn + o(n)$ and the one-point estimator $\hat{x}_p := X_{(\lfloor g \rfloor)}$.

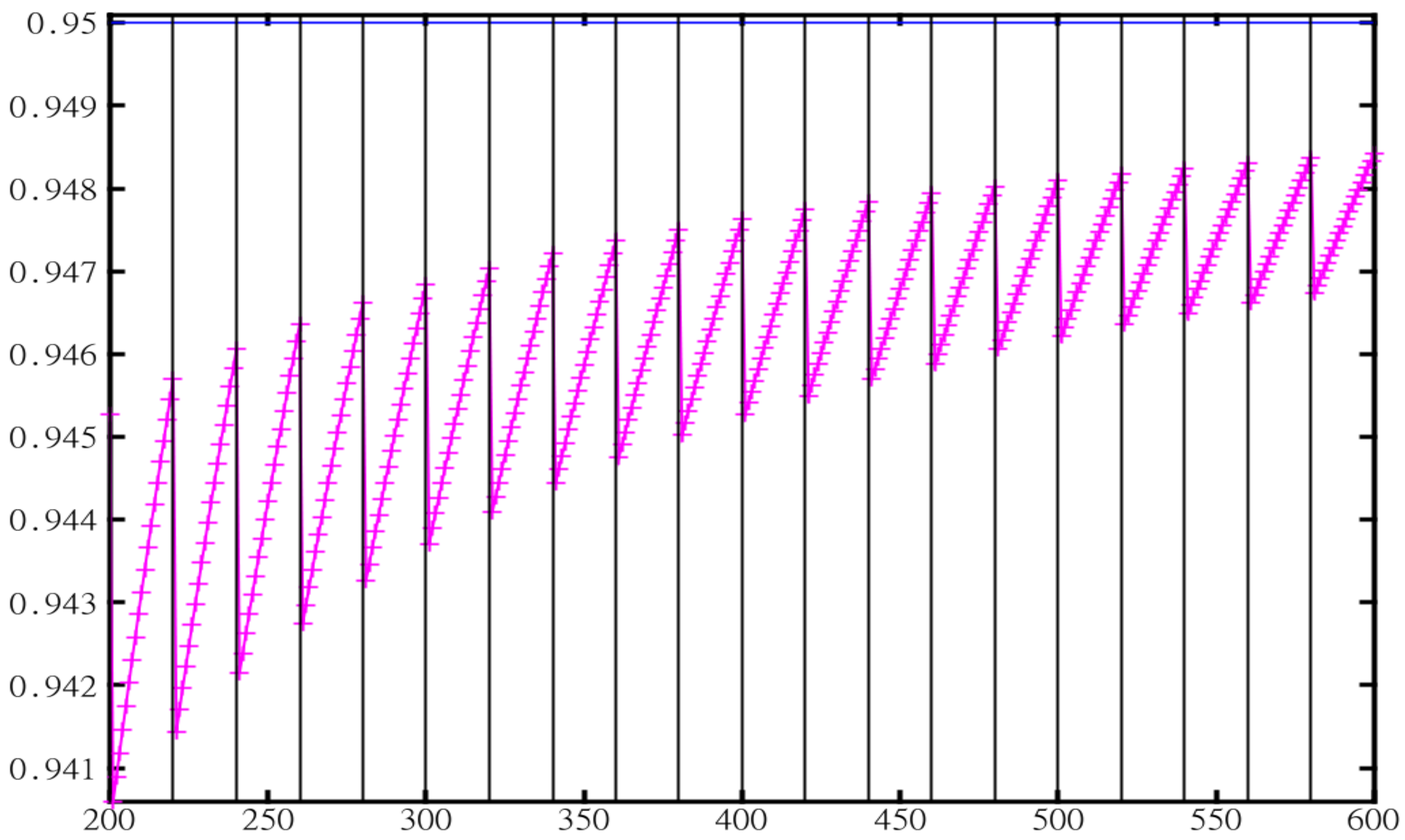
Discretization error of $\hat{x}_p$: $DE(\hat{x}_p) := -\dfrac{g - \lfloor g \rfloor}{n+1} \leq 0$

Sample error of $\hat{x}_p$: $SE(\hat{x}_p) := \dfrac{g - p(n+1)}{n+1}$

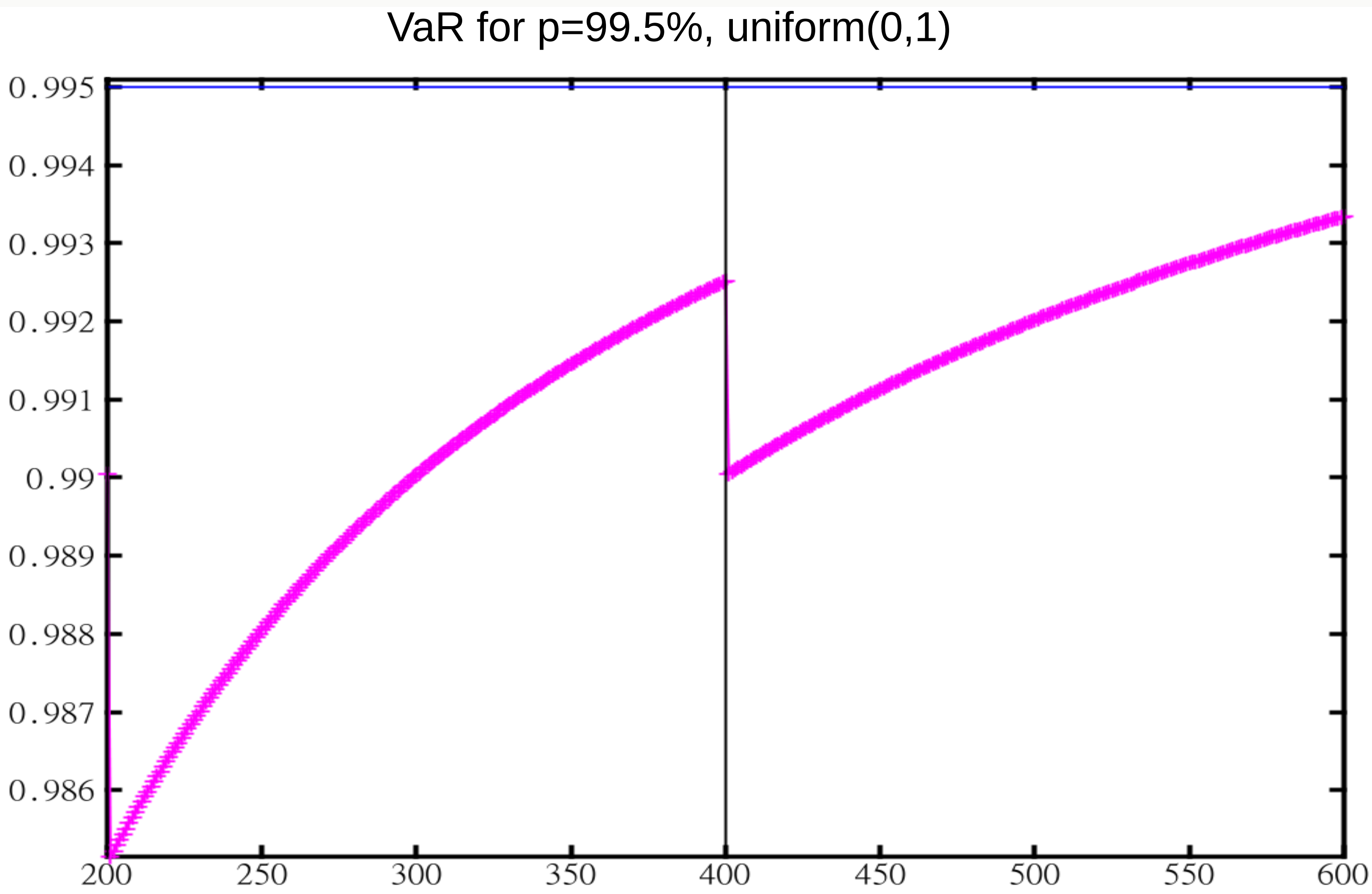
Lemma: The bias decompositon is $E[\hat{x}_p] = p + DE(\hat{x}_p) + SE(\hat{x}_p)$

Uniform distribution, one point estimator

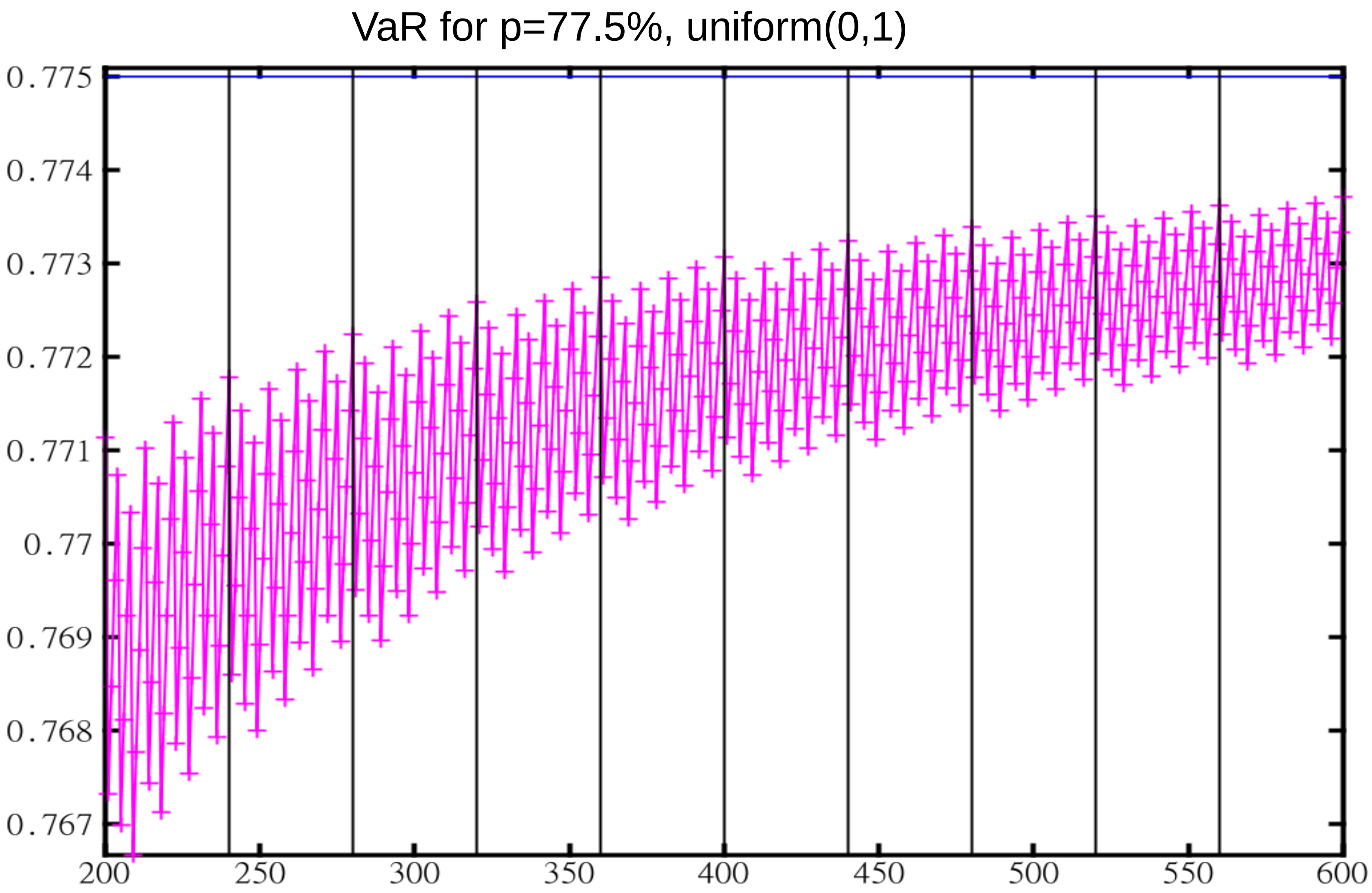
VaR for p=95%, uniform(0,1)



Uniform distribution, one point estimator



Uniform distribution, one point estimator



Decomposition of the bias
Uniform distribution, smoothed estimator

Let $g := pn + c_p$ and the weight factor $\gamma \in [0,1]$, we define the smoothedtwo-point estimatoras

$$\hat{x}_p := \gamma \cdot X_{(\lfloor g \rfloor)} + (1-\gamma) \cdot X_{(\lceil g \rceil)}$$

Lemma: For suitablygeneralized definitions of discretization error andsampleerror the biasdecompositonis

$$E[\hat{x}_p] = p + DE(\hat{x}_p) + SE(\hat{x}_p)$$

with

$$SE(\hat{x}_p) = \frac{g - p(n+1)}{n+1} = \frac{c_p - p}{n+1}$$

$$DE(\hat{x}_p) = -\frac{g - \lfloor g \rfloor}{n+1} + \frac{\gamma}{n+1}(1 - \delta(g - \lfloor g \rfloor))$$

$DE(\hat{x}_p) = 0$ if andonlyif $n \in N$ suchthat $\gamma = g - \lfloor g \rfloor$

Decomposition of the bias
General distribution, smoothed estimator

Proposition: The bias decompositon is

$$E\left[\hat{x}_p\right] = x_p + \frac{DE(\hat{x}_p) + SE(\hat{x}_p)}{f(x_p)} - \frac{p(1-p)}{2(n+2)} R + O(n^{-2})$$

with R of order one and a function of f' and f^{-3}

For the uniform distribution this decompositon the earlier lemma.

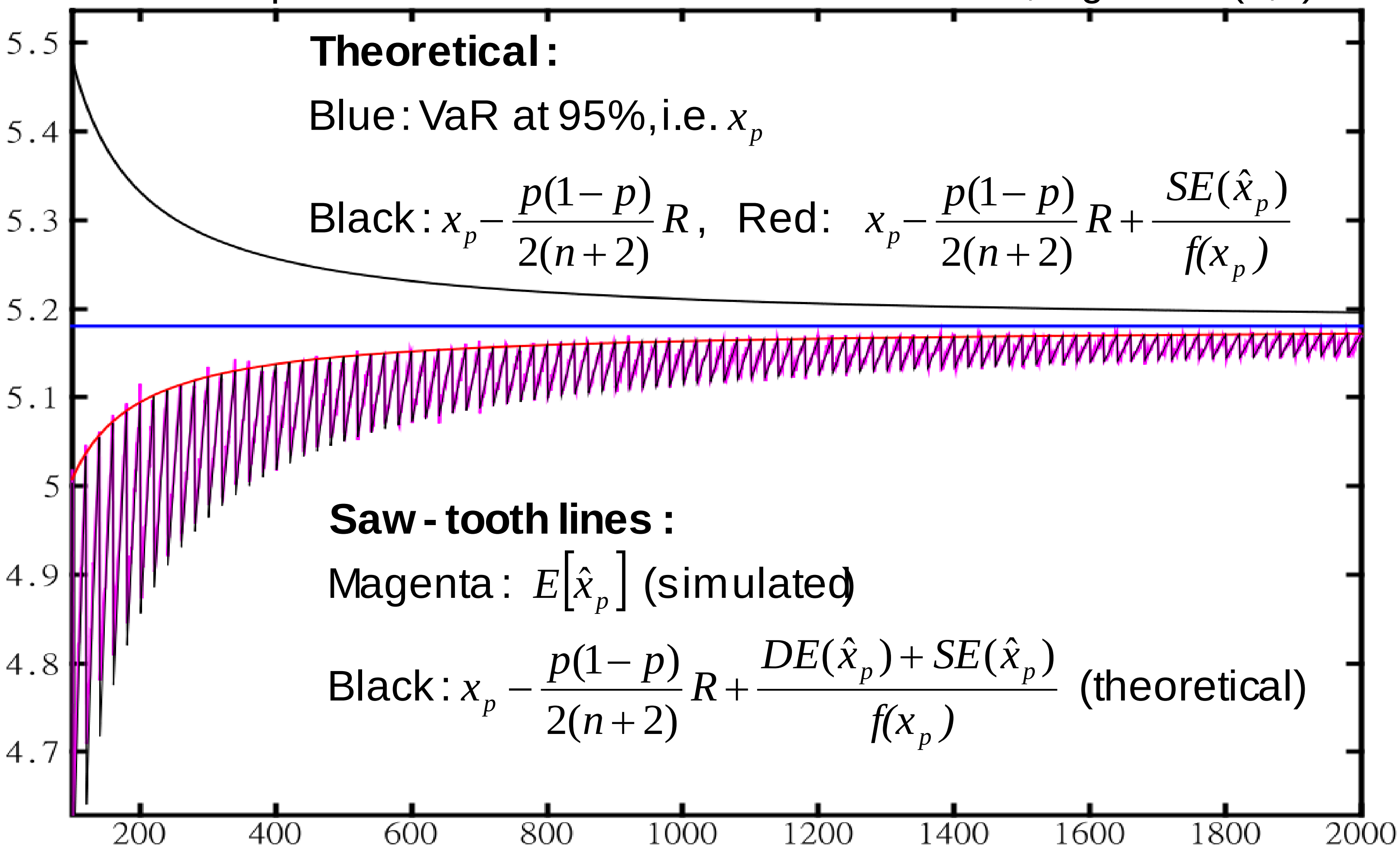
The position g which yields the unbiased estimator for the uniform distribution is not necessarily a good choice for general distributions. Good choices are dependent of f and exploit the fact that the error terms have different signs.

One can also optimize the median of the error instead of its expectation. Hyndman and Fan's chose the g in such a way:

$\hat{x}_p := (1-\gamma) x_{(\lfloor g \rfloor)} + \gamma\, x_{(\lceil g \rceil)}$ with $g = p(n+\frac{1}{3}) + \frac{1}{3}$ and $\gamma = g - \lfloor g \rfloor$

General distribution – Verification of results

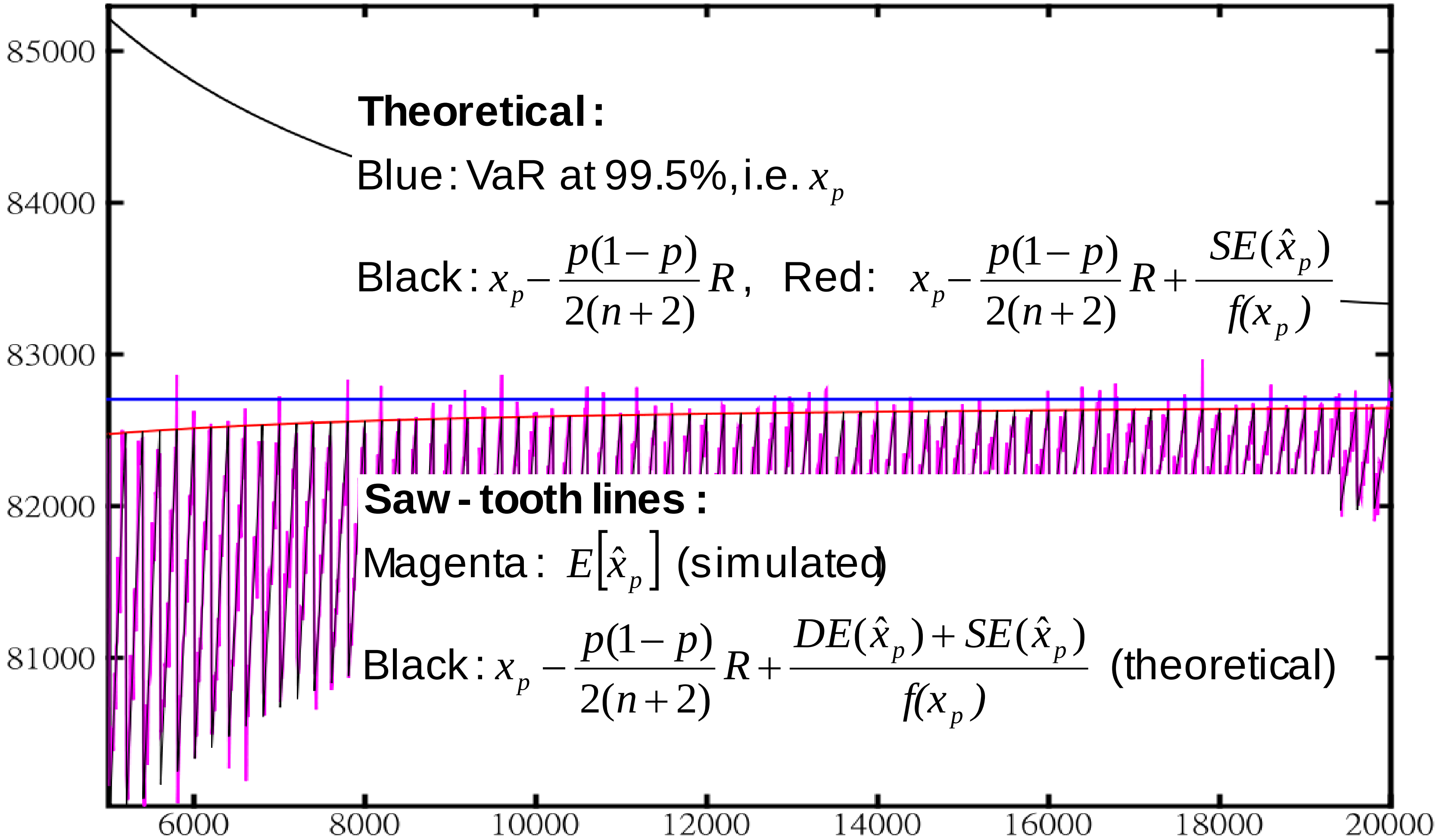
Bias decomposition for low side estimator: VaR 95%, lognormal(0,1)



Simulated results are visually almost indistinguishable from the theoretical results

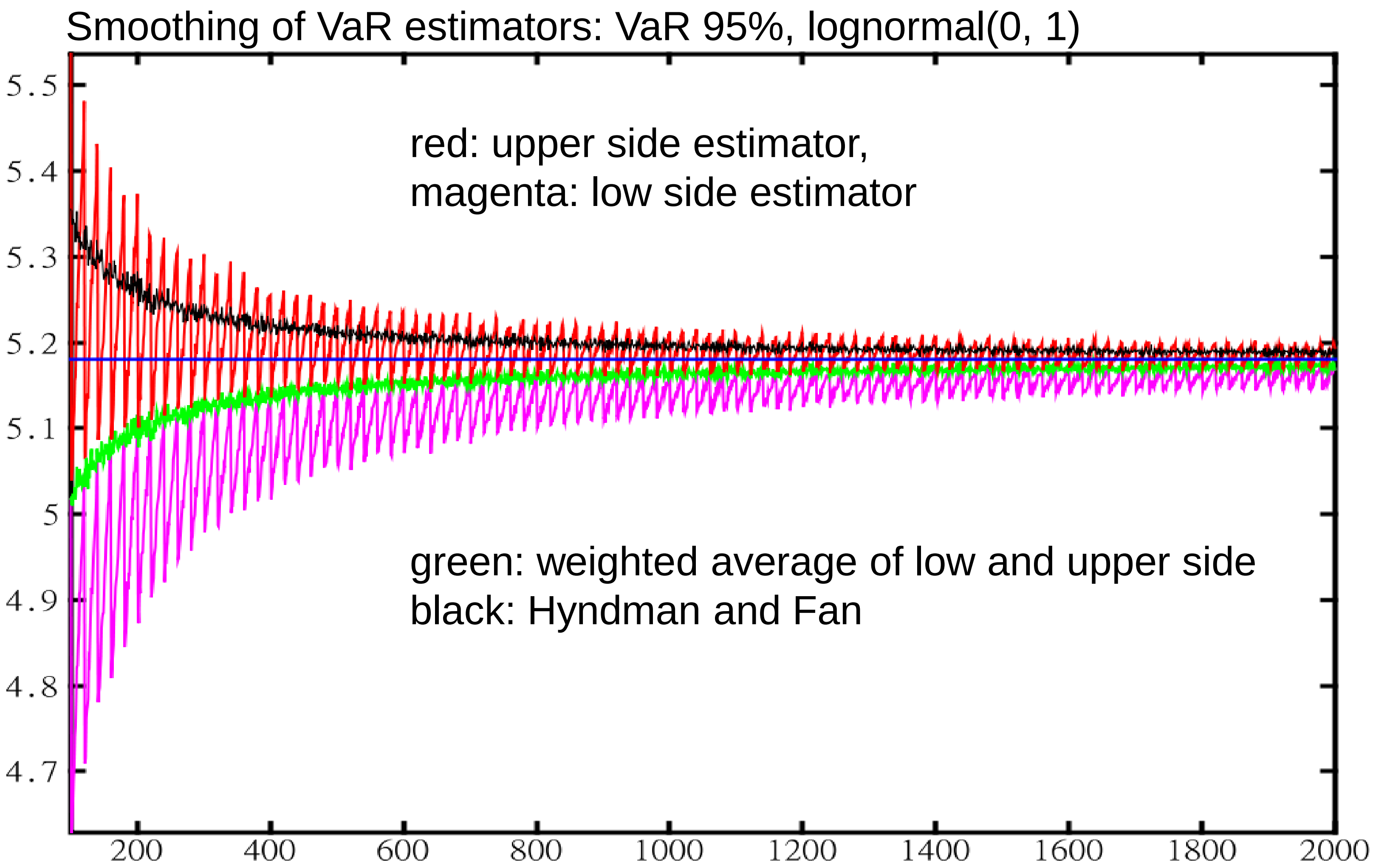
Examples

Bias decomposition for low side estimator: VaR 99.5%, Pareto(1.2, 1000)



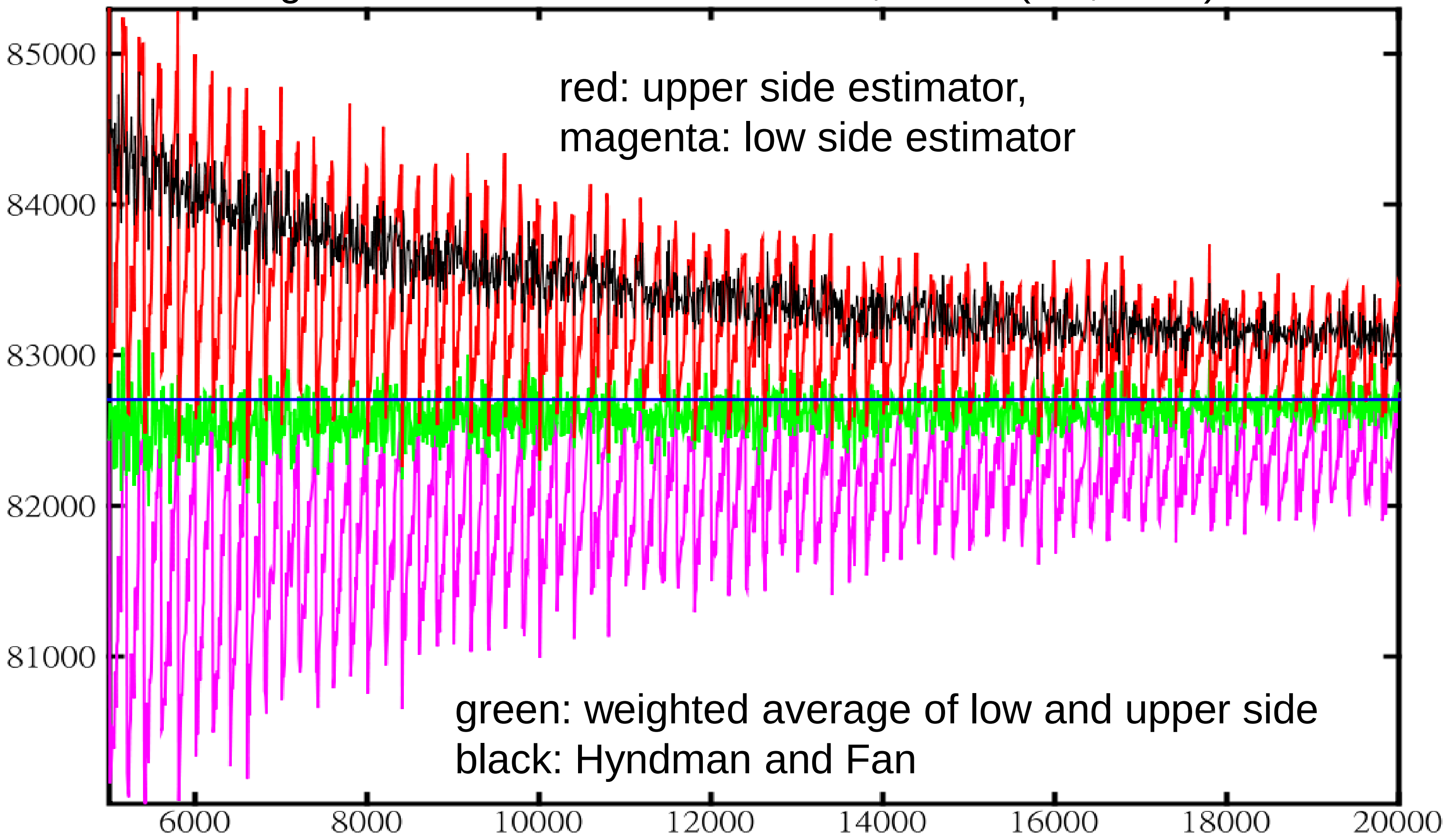
Simulated results are visually almost indistinguishable from the theoretical results

The best smoothing depends on tail behavior and p



The best smoothing depends on tail behavior and p

Smoothing of VaR estimators: VaR 99.5%, Pareto(1.2, 1000)



For heavy tail distributions, the Hyndman and Fan is not the best estimator.

Conclusions

- If at all, one-point estimators should only be used for very specific order statistics
- A good smoothing depends on the tail behavior of the distribution
- A good smoothing depends on the VaR level