Estimation of probability density functions by the Maximum Entropy Method

Claudio Bierig, Alexey Chernov

Institute for Mathematics
Carl von Ossietzky University Oldenburg

alexey.chernov@uni-oldenburg.de

3. Workshop of the GAMM AGUQ

Dortmund, 14 March 2018

Forward uncertainty propagation in complex systems:

$$X = S(\xi)$$

- $\xi =$ input (random variable, known pdf)
- $X =$ output (random variable, unknown pdf ρ)
- $S =$ System's action (deterministic, known)

Aim: determine the pdf ρ of X (nonintrusively!)

Note: only a finite amount of information can be pushed through S.

Outline:

- Methodology
- Error Analysis
- Experiments (i.a. rough random obstacle problem)

Idea 1: ($\rightarrow$ Fabio's talk on Monday)

- Determine point values $\rho(x_i)$, for $x_0 < x_1 < \cdots < x_N$
- Interpolate for $x \in (x_i, x_{i+1})$

M. B. GILES, T. NAGAPETYAN, AND K. RITTER, Multilevel Monte Carlo Approximation of Distribution Functions and Densities, SIAM/ASA *J. Uncert. Quantif.*, 3 (2015), pp. 267–295

Adv.: More point values $\rightarrow$ better approximation (stability, conv.).

Drawback: Many unknown parameters in the algorithm (complex).

See also the "antiderivative approach":

S. KRUMSCHEID AND F. NOBILE, Multilevel Monte Carlo approximation of functions, MATHICSE technical report Nr. 12.2017

Idea 2 (explore in this talk):

- Determine the (generalized) moments $\mu_1, \ldots, \mu_R$ of X

$$\mu_k = \mathbb{E}[\phi_k(X)]$$

- Reconstruct η that satisfies the moment constraints:
  a) $\mu_k = \int \phi_k(x)\eta(x)\,dx$
  b) $\eta(x) \geq 0$ and $\int \eta(x)\,dx = 1.$

C. Bierig and A. Chernov, Approximation of probability density functions by the Multilevel Monte Carlo Maximum Entropy method, *J. Comput. Physics*, 314 (2016), 661–681 [a]

[a] Based on earlier works [Csiszár'75], [Barron, Sheu'91], [Borwein, Lewis'91]

Advantage: Only a few parameters (simple)

Drawback: More moments $\rightarrow$ better approximation? (stability?).

- **Observe:** If the moments $\mu_1, \ldots, \mu_R$ are consistent, the reconstructed density η is usually not uniquely determined!
- **How to select** the "most appropriate density"?
- **The Maximum Entropy (ME) method:**

Find $\rho_R = \underset{\eta}{\operatorname{argmax}} \left(- \int \eta(x) \ln \eta(x) \, dx \right)$ under constraints:

a) $\mu_k = \int \phi_k(x) \eta(x) \, dx, \qquad k = 1, \ldots, R$

b) $\eta(x) \geq 0$ and $\int \eta(x) \, dx = 1.$

"It is least biased estimate possible on the given information; i.e., it is maximally noncommittal with regard to missing information."
[E.T. Jaynes, 1957]

- **The Maximum Entropy (ME) method:**

  The solution to this problem can be equiv. characterized as

$$\rho_R \propto \exp\left(\sum_{k=0}^{R} \lambda_k \phi_k(x)\right), \qquad \lambda_k \in \mathbb{R}.$$

where $\lambda_0, \ldots, \lambda_R$ satisty the constraints **(moment matching)**:

$$\mu_k = \int \phi_k(x) \rho_R(x)\, dx, \qquad k = 0, \ldots, R,$$

with $\mu_0 = 1$, $\phi_0(x) = 1$.

In this sense:

Entropy maximization

$$\Updownarrow$$

moment matching

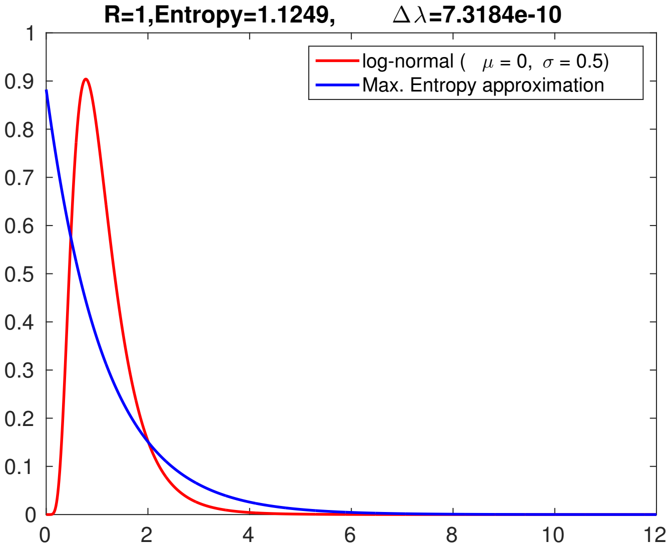
Test example:

- ρ is the log-normal distribution with $\mu = 0$ and $\sigma = 0.5$ and 0.2
- Estimation of moments $\mu_1, \ldots, \mu_R$ by MC with 10^8 samples
- $\lambda = (\lambda_0, \ldots, \lambda_R)$ determined by the Newton-Raphson method

$$\int \phi_k(x)\rho_R(\lambda, x)\, dx = \mu_k, \qquad k = 0, \ldots R.$$

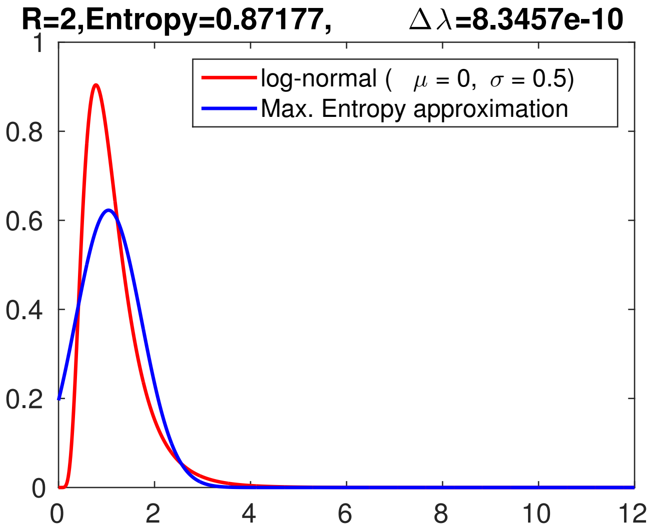
Stopping parameters for the Newton-Raphson Method:

- $\Delta\lambda \leq 10^{-9}$ (convergence)
- $\Delta\lambda \geq 10^{3}$ (no convergence)
- $\#\texttt{iter} \geq 1000$ (no convergence)



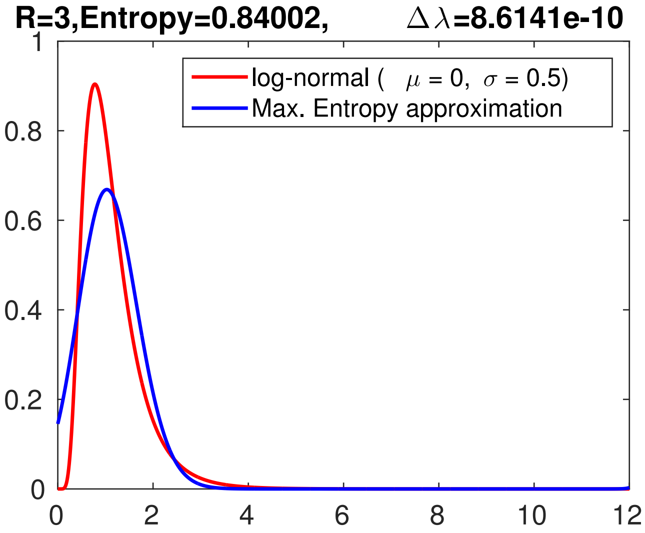
Monomial Moments:

- Unstable for $R \geq 5$



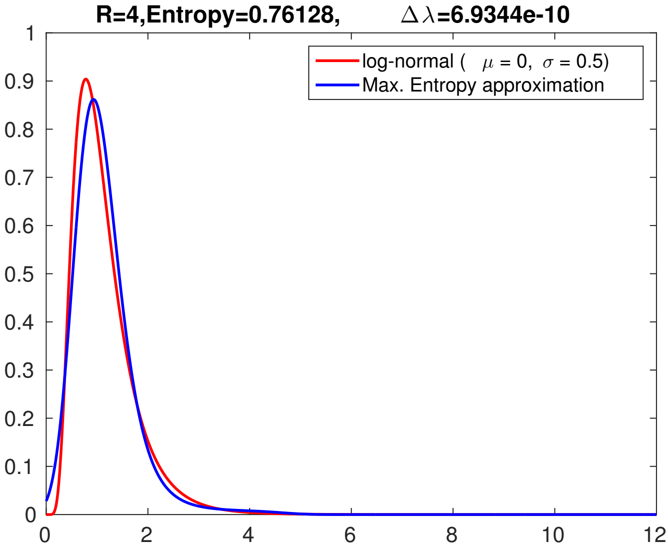
Monomial Moments:

- Unstable for $R \geq 5$



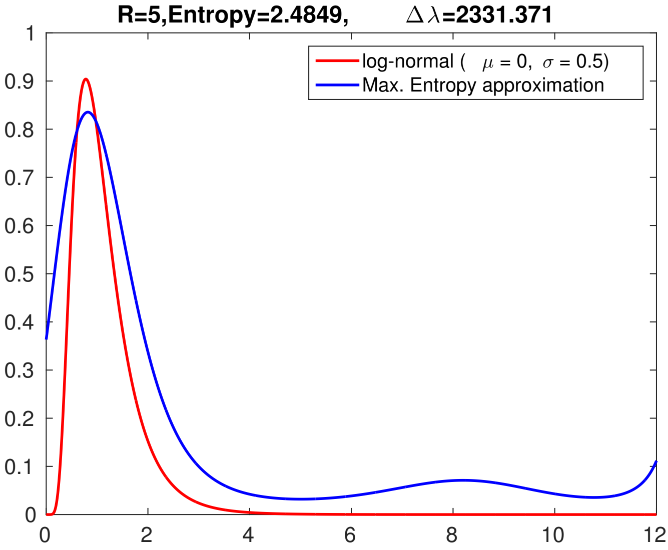
Monomial Moments:

- Unstable for $R \geq 5$



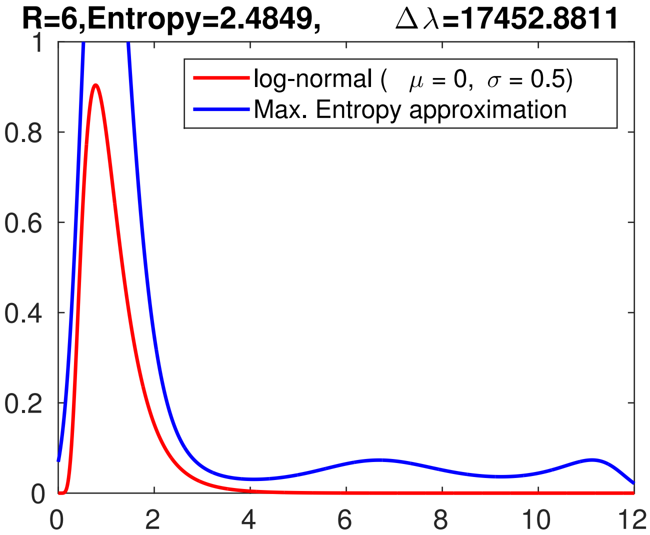
Monomial Moments:

- Unstable for $R \geq 5$



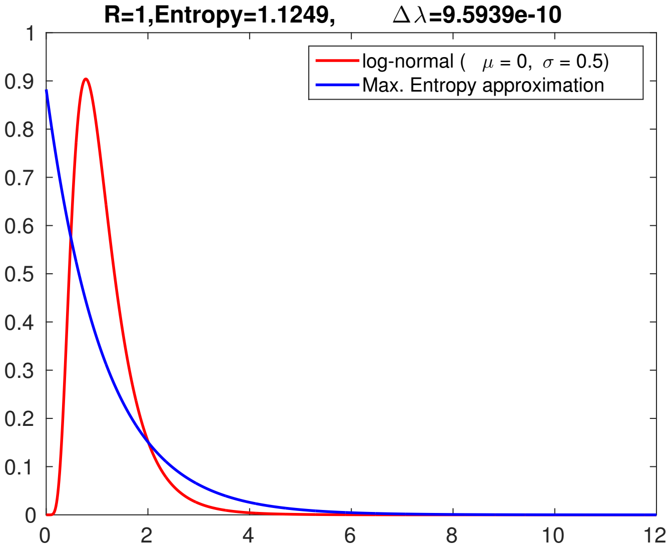
Monomial Moments:

- Unstable for $R \geq 5$



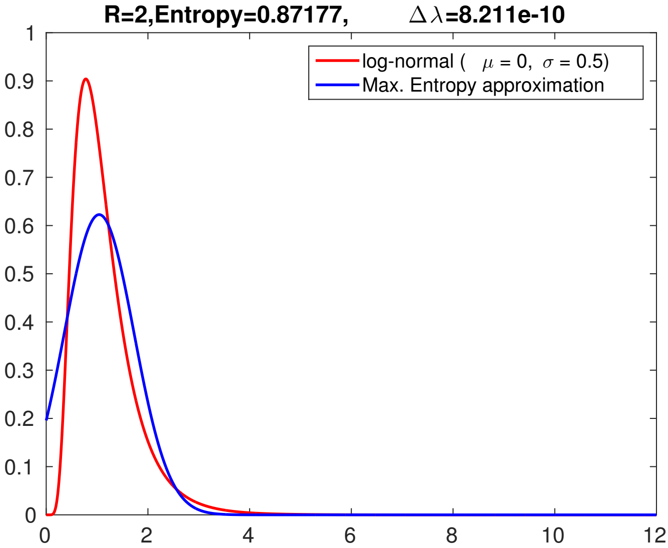
Monomial Moments:

- Unstable for $R \geq 5$



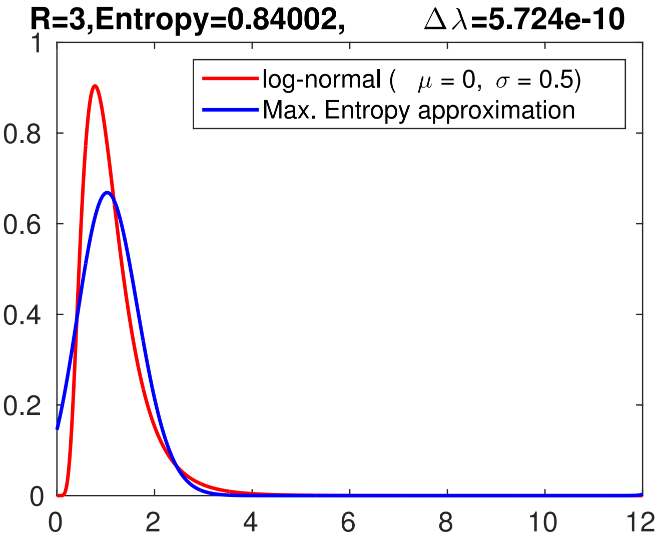
Legendre Moments:

- Stable for $R \leq 8$
- Entropy is may decrease even when the Newton Method does not converge



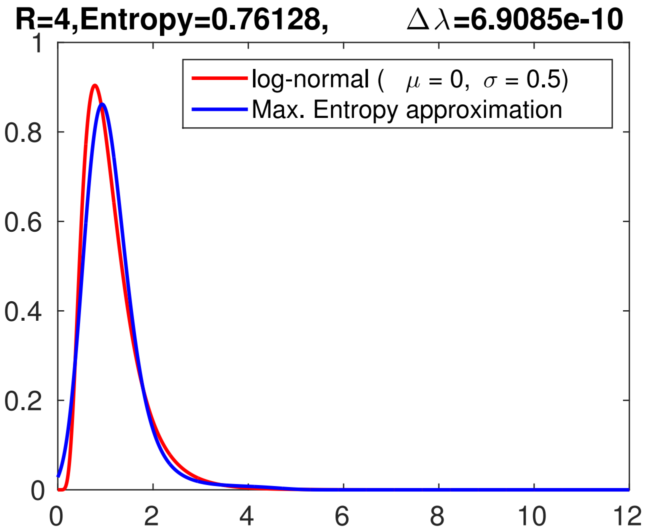
Legendre Moments:

- Stable for $R \leq 8$
- Entropy is may decrease even when the Newton Method does not converge



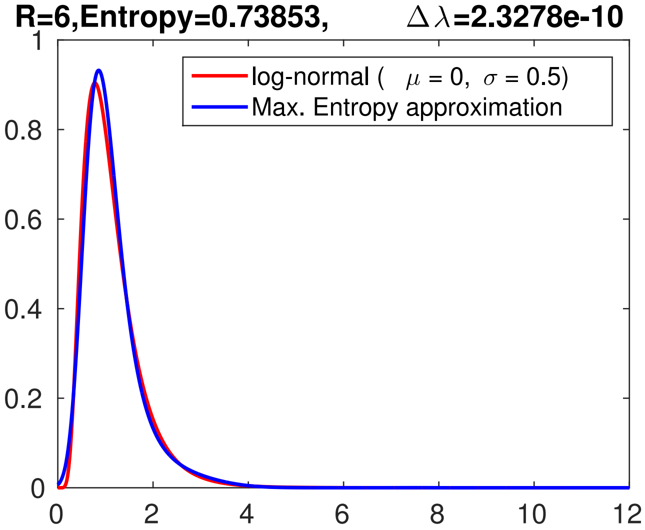
Legendre Moments:

- Stable for $R \leq 8$
- Entropy is may decrease even when the Newton Method does not converge



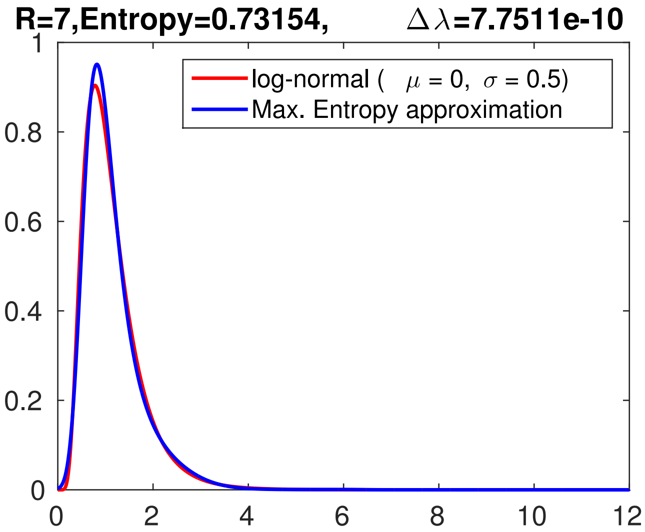
Legendre Moments:

- Stable for $R \leq 8$
- Entropy is may decrease even when the Newton Method does not converge



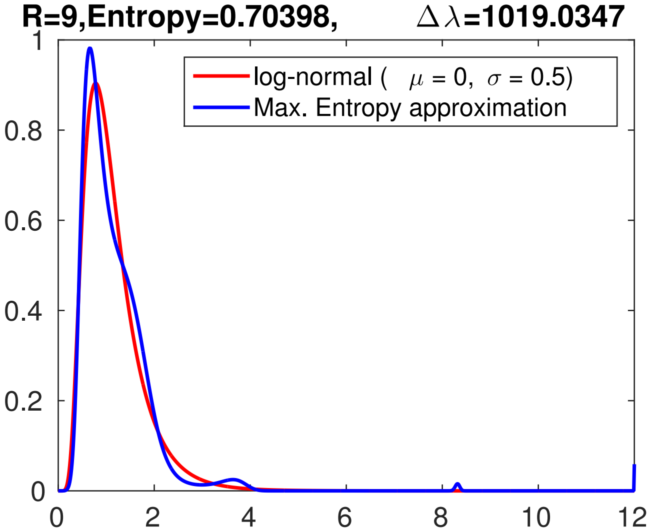
Legendre Moments:

- Stable for $R \leq 8$
- Entropy is may decrease even when the Newton Method does not converge



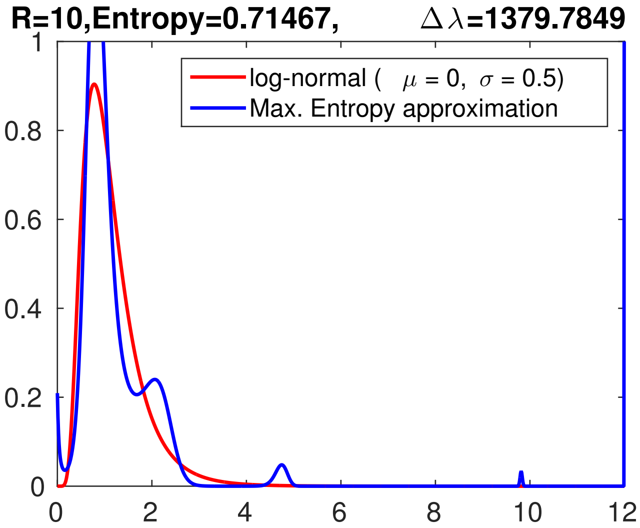
Legendre Moments:

- Stable for $R \leq 8$
- Entropy is may decrease even when the Newton Method does not converge



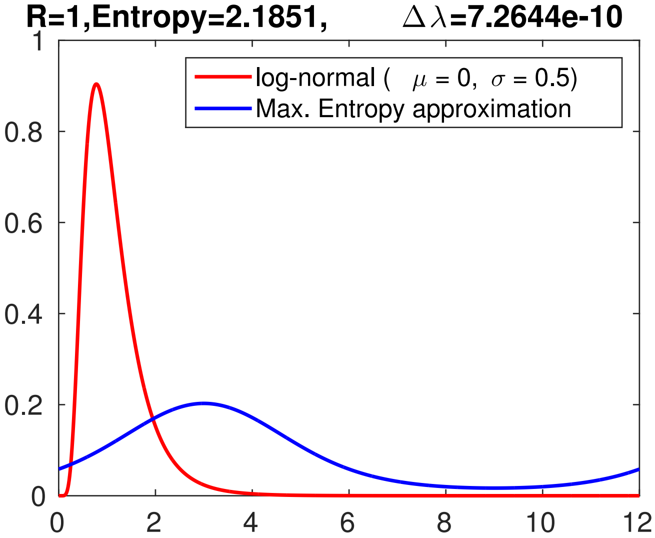
Legendre Moments:

- Stable for $R \leq 8$
- Entropy is may decrease even when the Newton Method does not converge



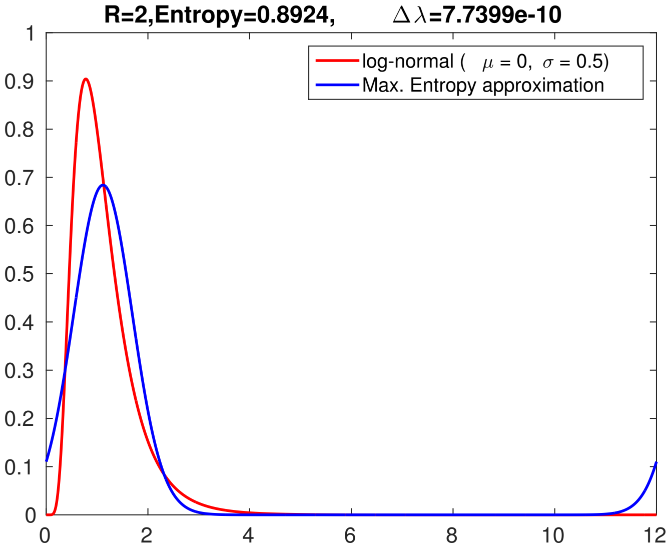
Legendre Moments:

- Stable for $R \leq 8$
- Entropy is may decrease even when the Newton Method does not converge



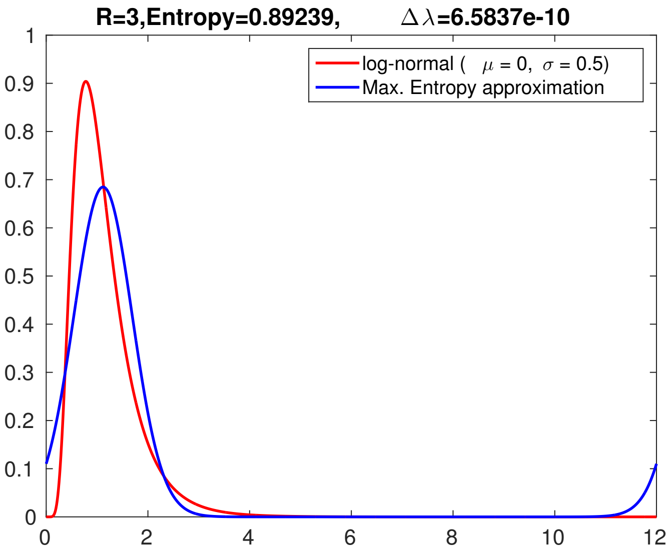
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



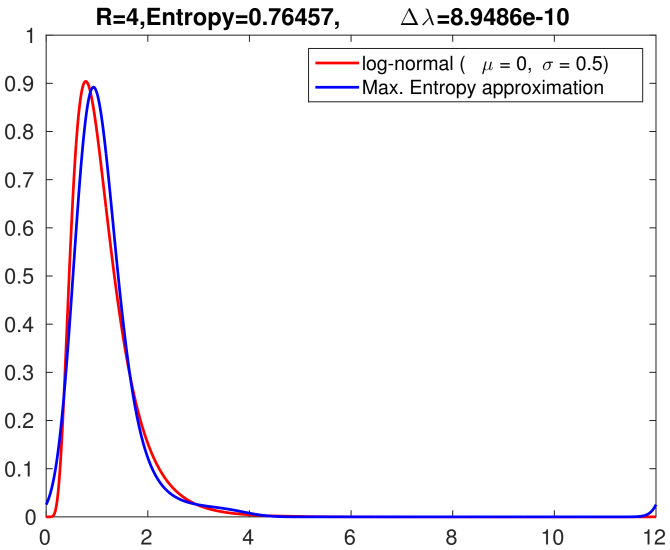
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



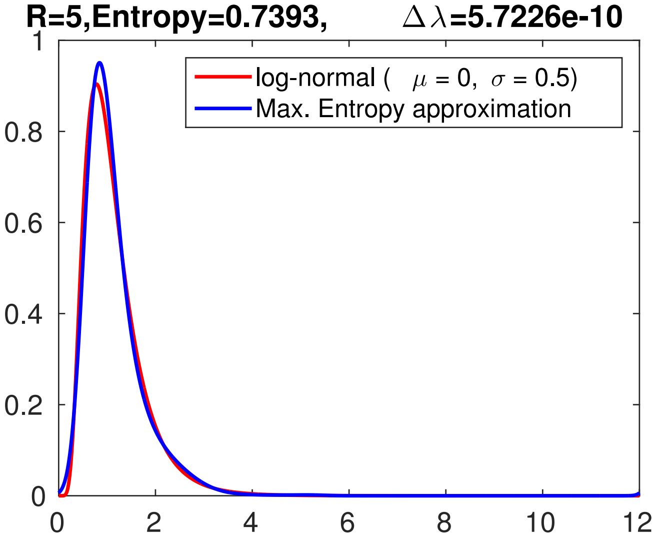
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



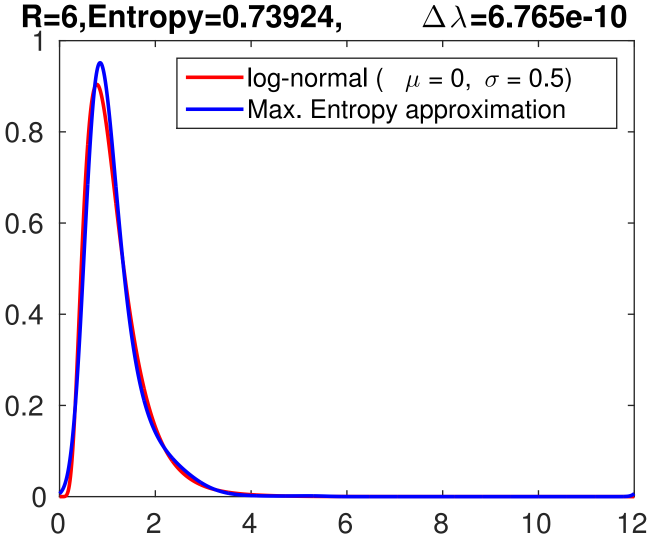
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



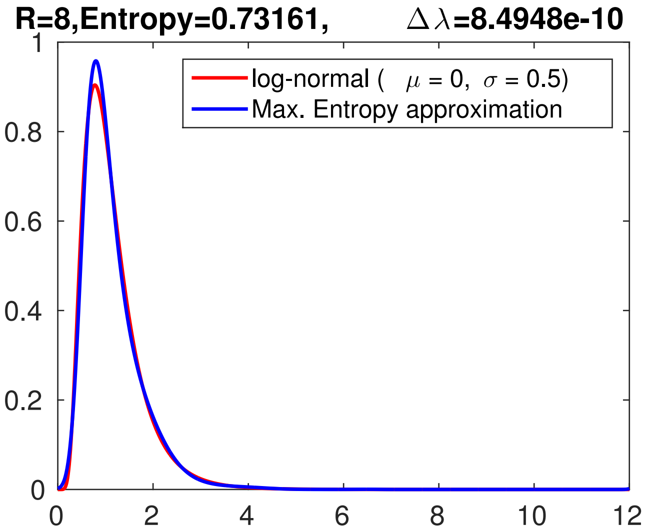
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



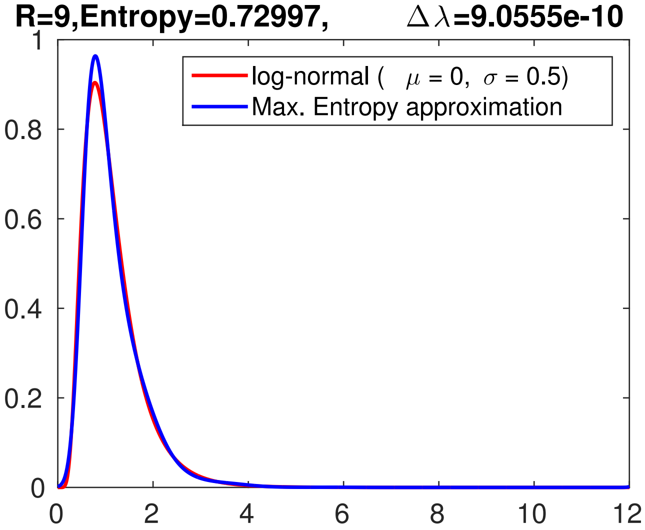
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



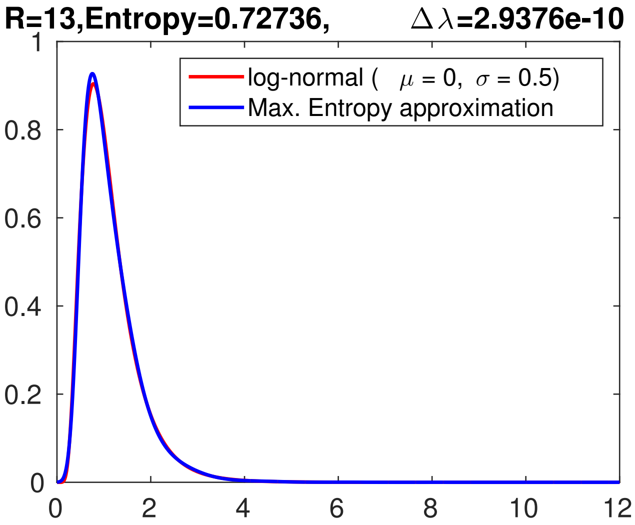
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



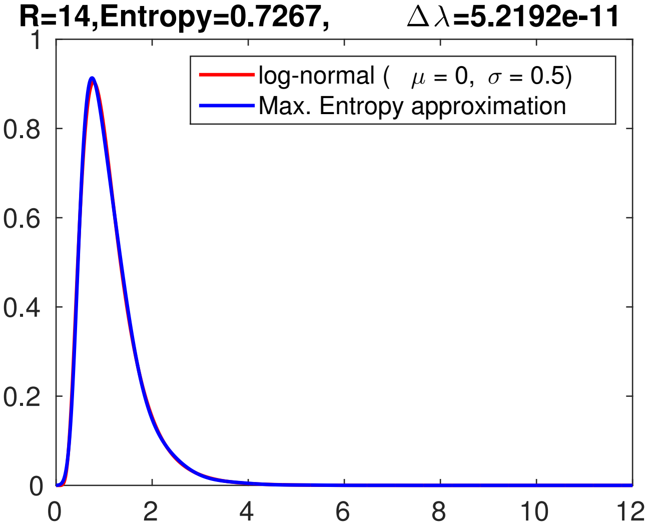
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



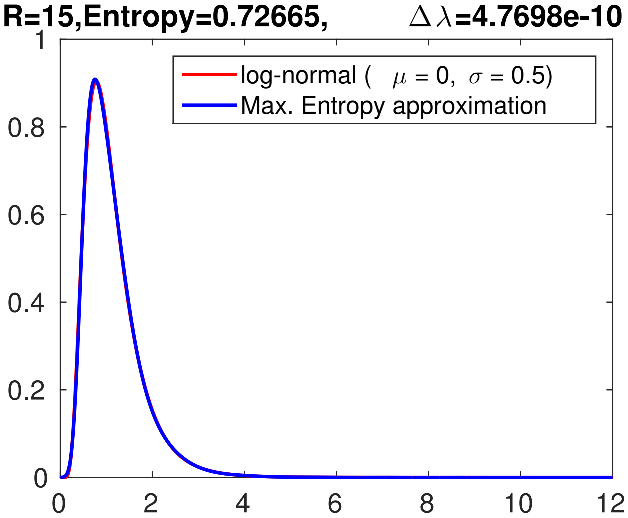
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



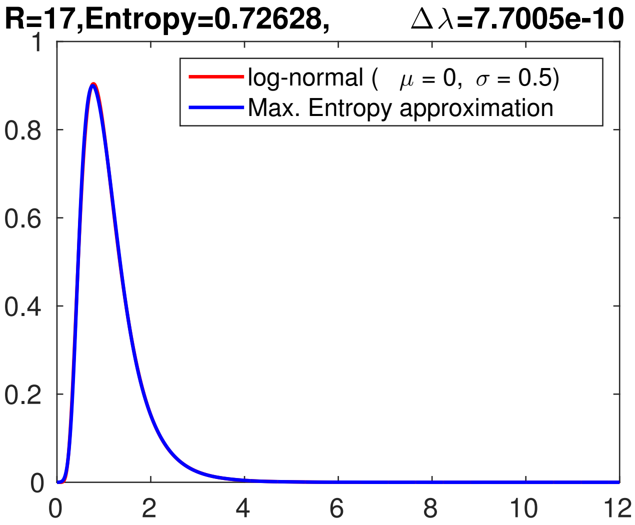
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



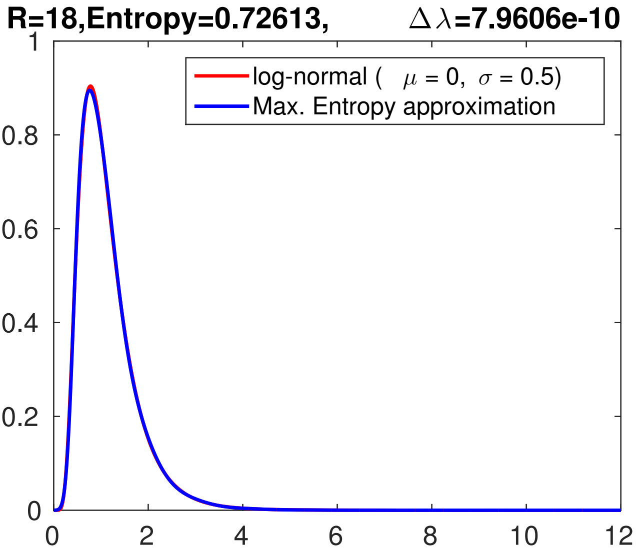
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



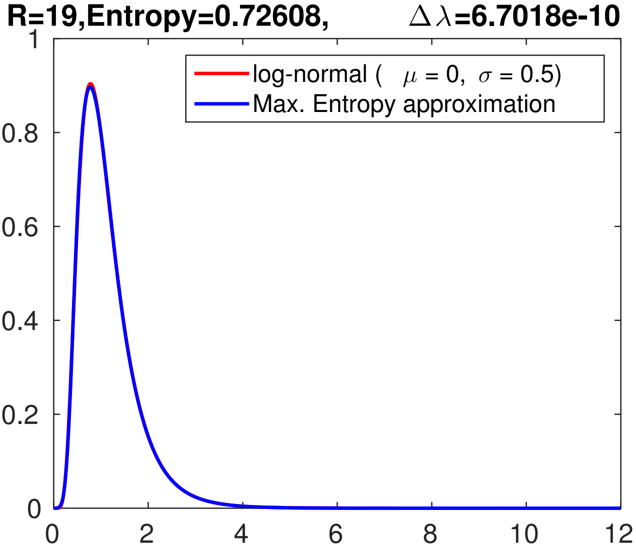
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



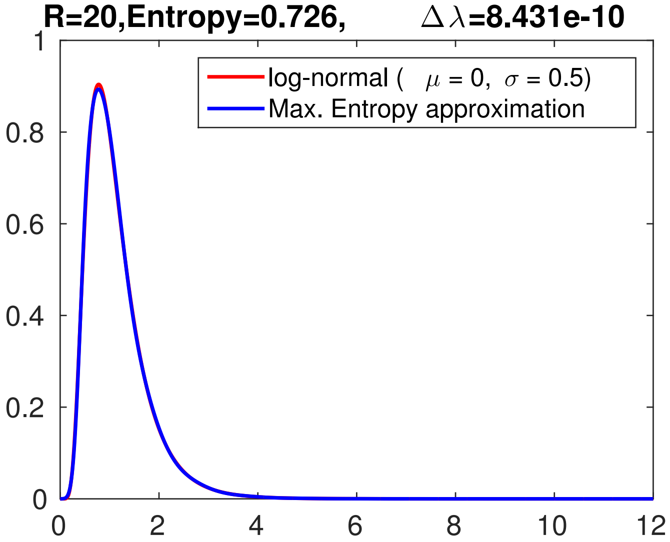
Fourier Moments:

- Stable
- Entropy is monotonously decreasing



Fourier Moments:

- Stable
- Entropy is monotonously decreasing



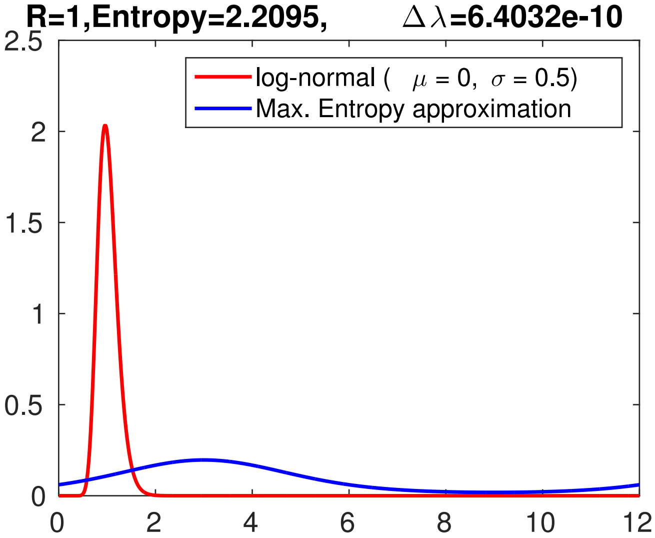
Fourier Moments:

- Stable
- Entropy is monotonously decreasing

Breaking convergence for the Fourier basis by choosing a more concentrated density!

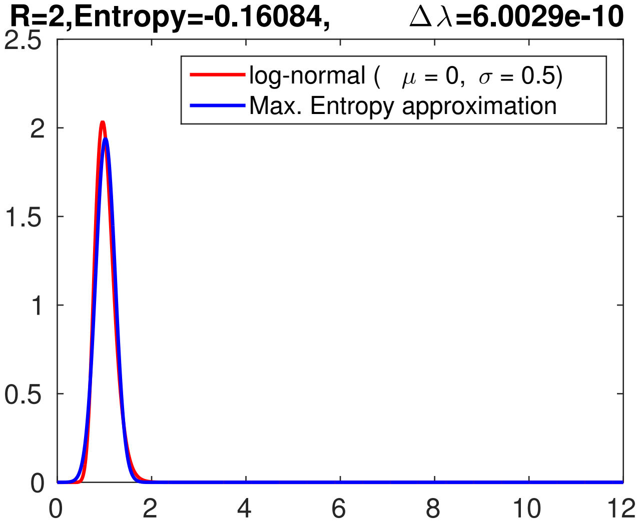
e.g. log-normal with $\mu = 0$, $\sigma = 0.2$

skip



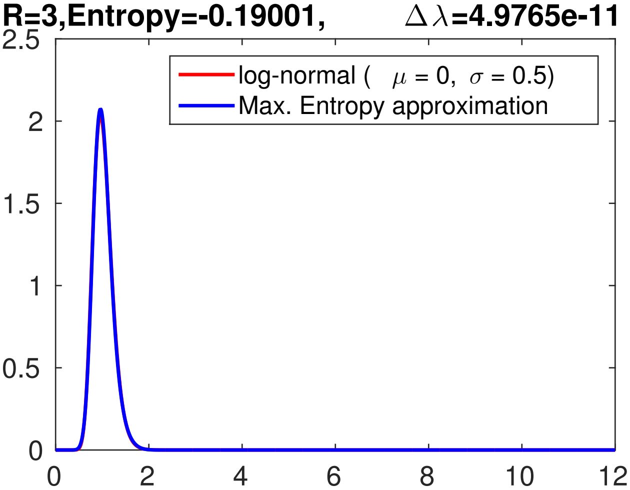
Fourier Moments ($\sigma = 0.2$, $[a, b] = [0, 12]$):

- Remain stable even without convergence!
- Entropy is still monotonously decreasing!



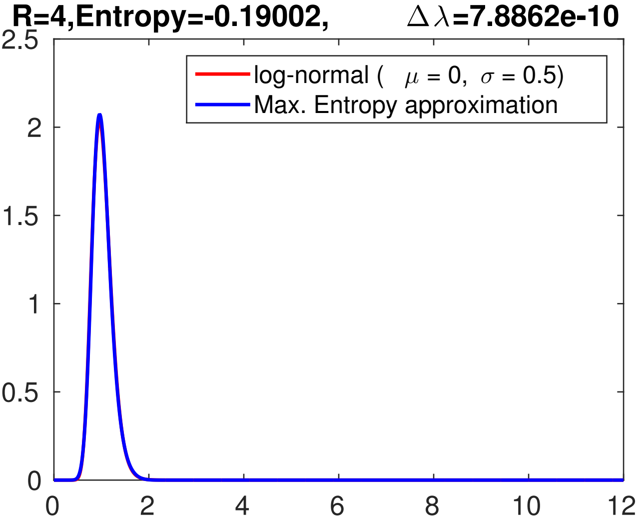
Fourier Moments ($\sigma = 0.2$, $[a, b] = [0, 12]$):

- Remain stable even without convergence!
- Entropy is still monotonously decreasing!



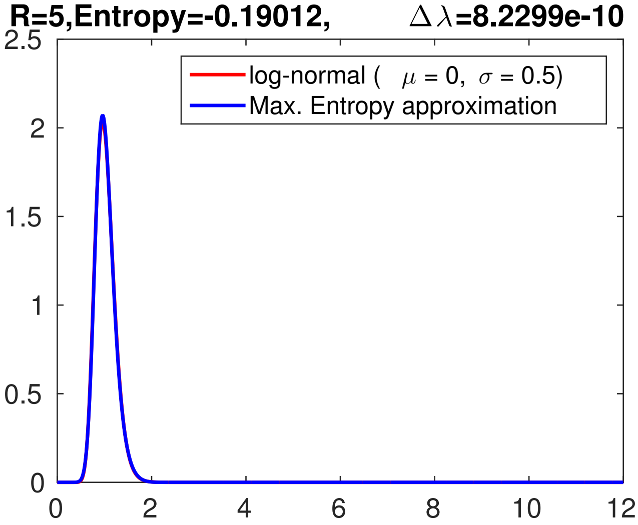
Fourier Moments ($\sigma = 0.2$, $[a, b] = [0, 12]$):

- Remain stable even without convergence!
- Entropy is still monotonously decreasing!



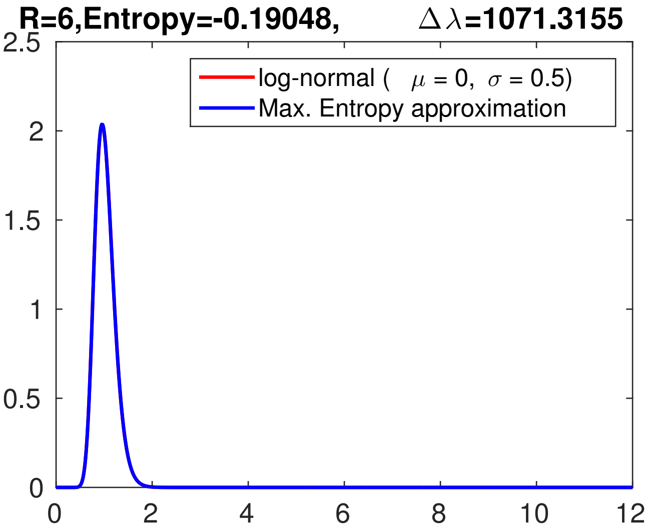
Fourier Moments ($\sigma = 0.2$, $[a, b] = [0, 12]$):

- Remain stable even without convergence!
- Entropy is still monotonously decreasing!



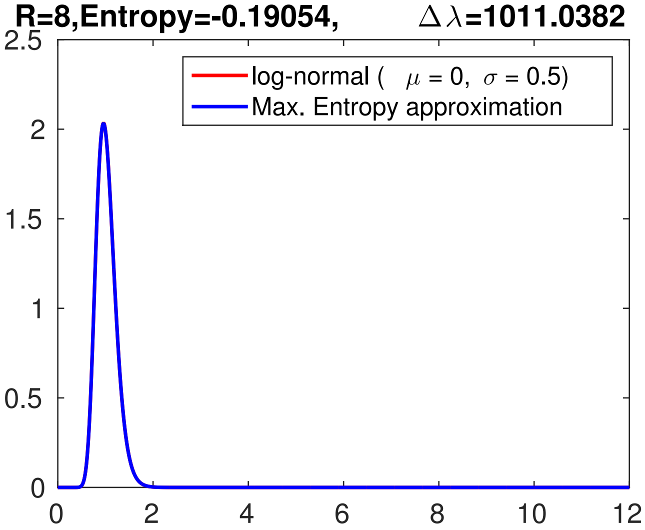
Fourier Moments ($\sigma = 0.2$, $[a, b] = [0, 12]$):

- Remain stable even without convergence!
- Entropy is still monotonously decreasing!



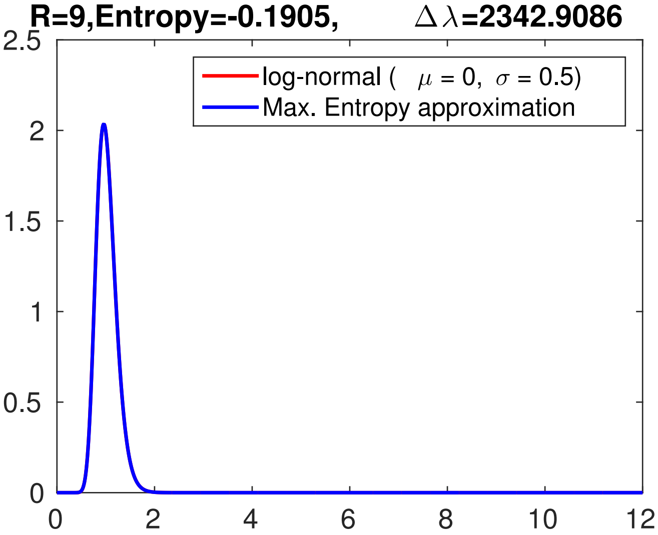
Fourier Moments ($\sigma = 0.2$, $[a, b] = [0, 12]$):

- Remain stable even without convergence!
- Entropy is still monotonously decreasing!



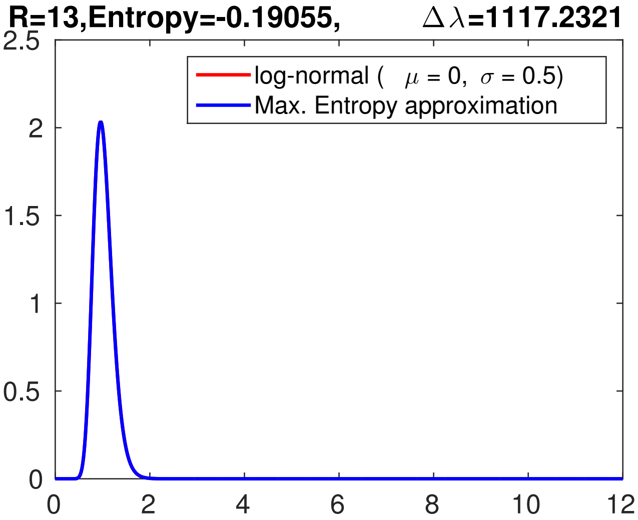
Fourier Moments ($\sigma = 0.2$, $[a, b] = [0, 12]$):

- Remain stable even without convergence!
- Entropy is still monotonously decreasing!



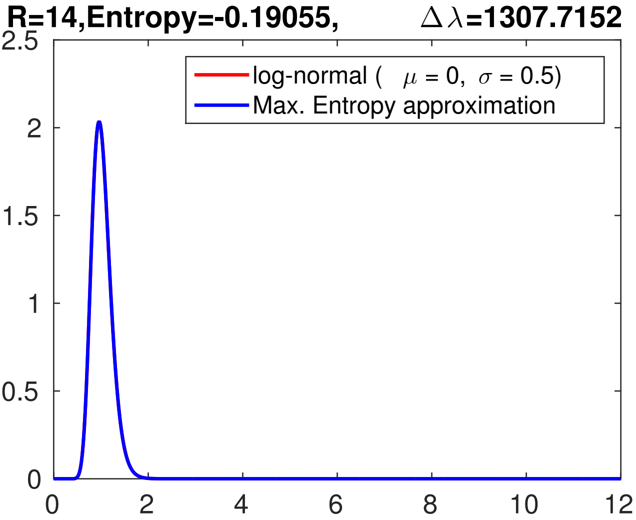
Fourier Moments ($\sigma = 0.2$, $[a, b] = [0, 12]$):

- Remain stable even without convergence!
- Entropy is still monotonously decreasing!



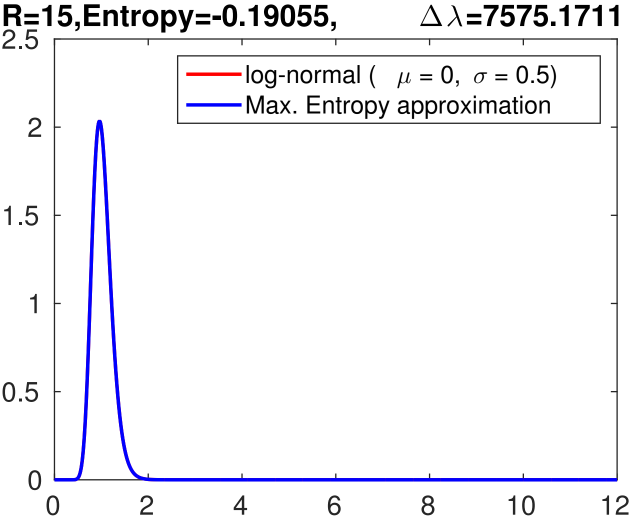
Fourier Moments ($\sigma = 0.2$, $[a, b] = [0, 12]$):

- Remain stable even without convergence!
- Entropy is still monotonously decreasing!



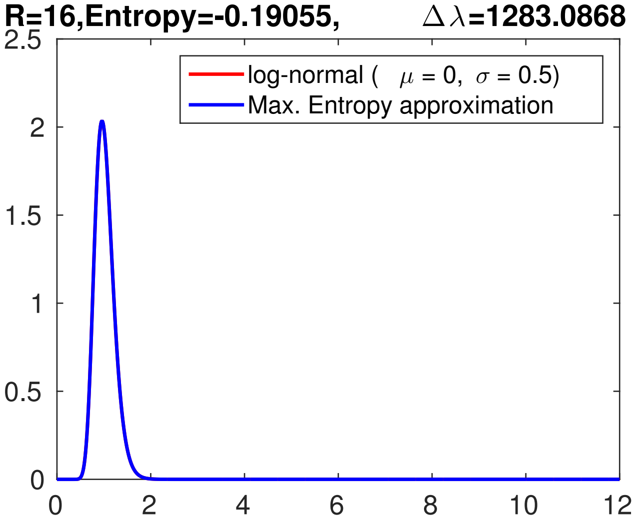
Fourier Moments ($\sigma = 0.2$, $[a, b] = [0, 12]$):

- Remain stable even without convergence!
- Entropy is still monotonously decreasing!



Fourier Moments ($\sigma = 0.2$, $[a, b] = [0, 12]$):

- Remain stable even without convergence!
- Entropy is still monotonously decreasing!

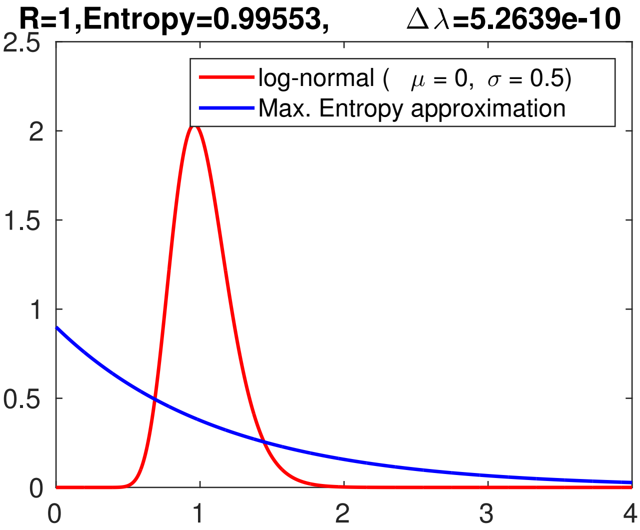


Fourier Moments ($\sigma = 0.2$, $[a, b] = [0, 12]$):

- Remain stable even without convergence!
- Entropy is still monotonously decreasing!

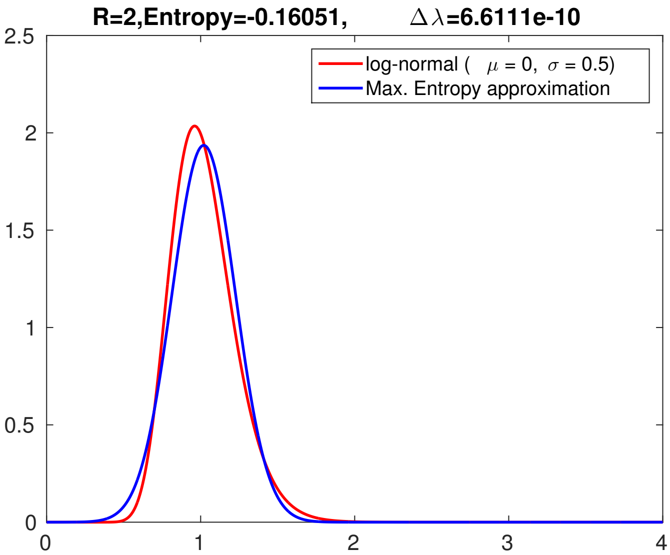
Regain stability of the Legendre basis
by choosing a smaller approximation interval!

e.g. $[a, b] = [0, 4]$



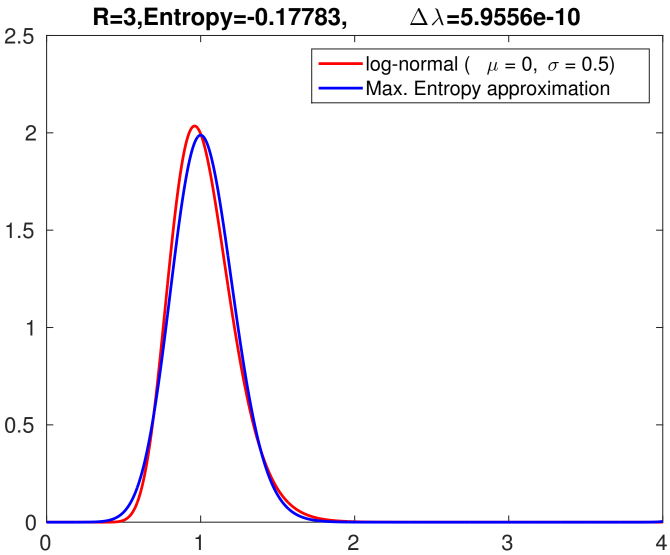
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



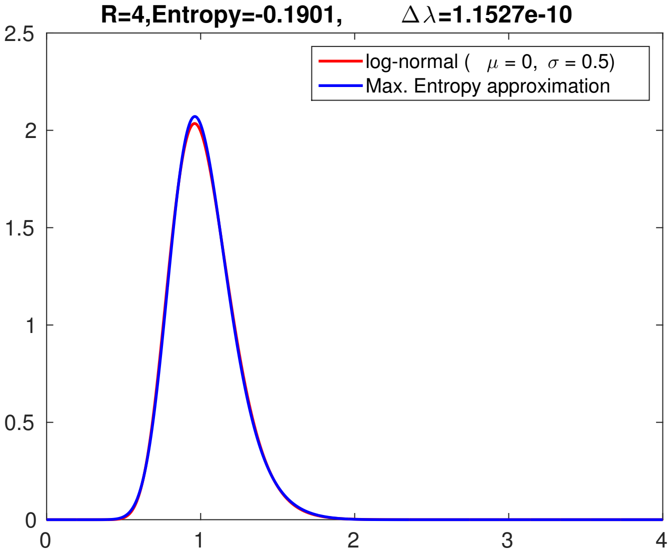
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



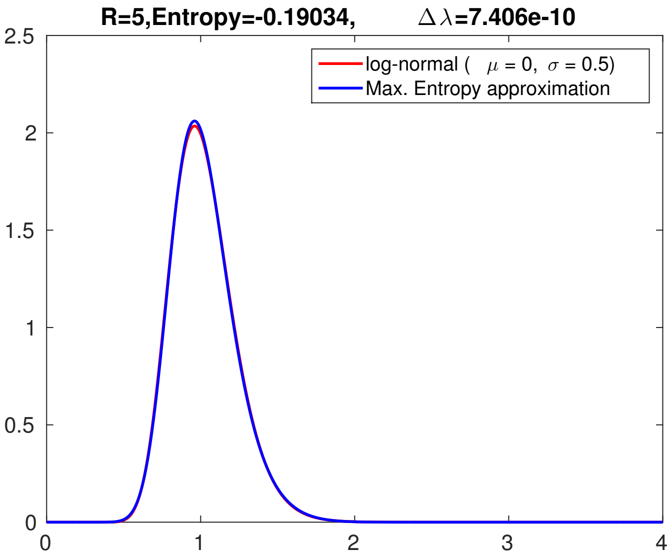
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



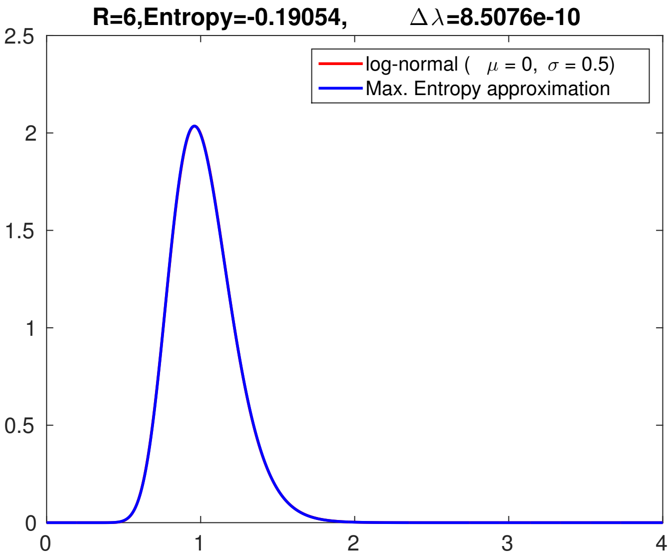
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



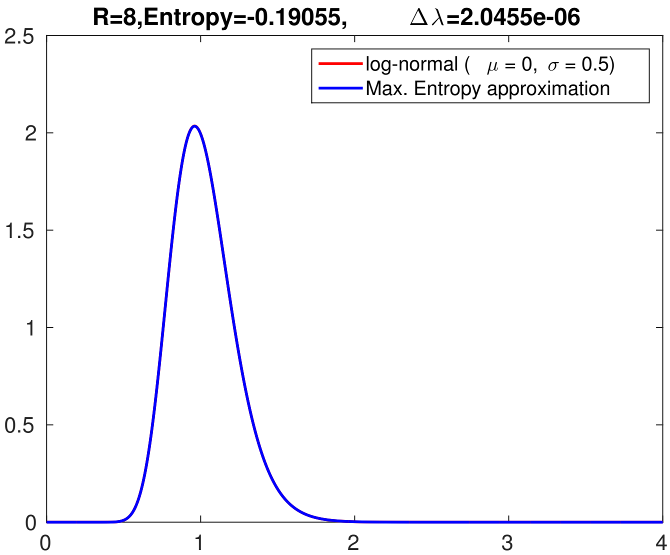
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



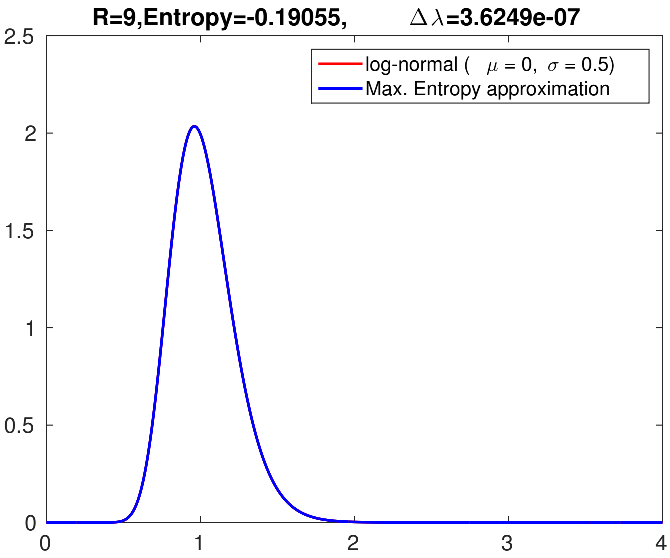
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



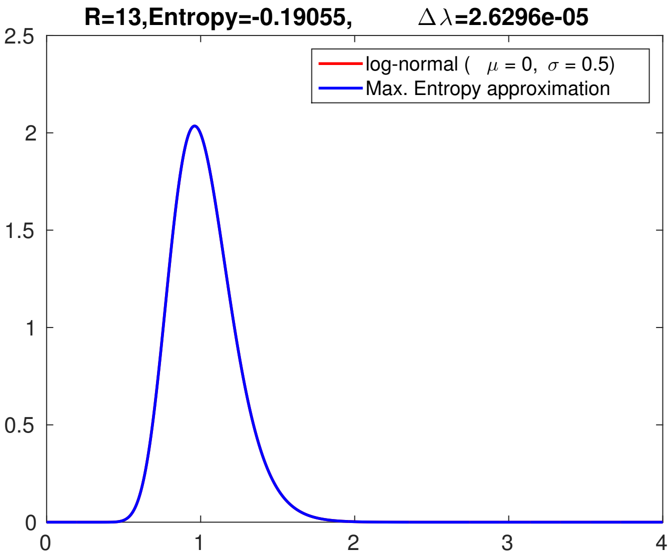
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



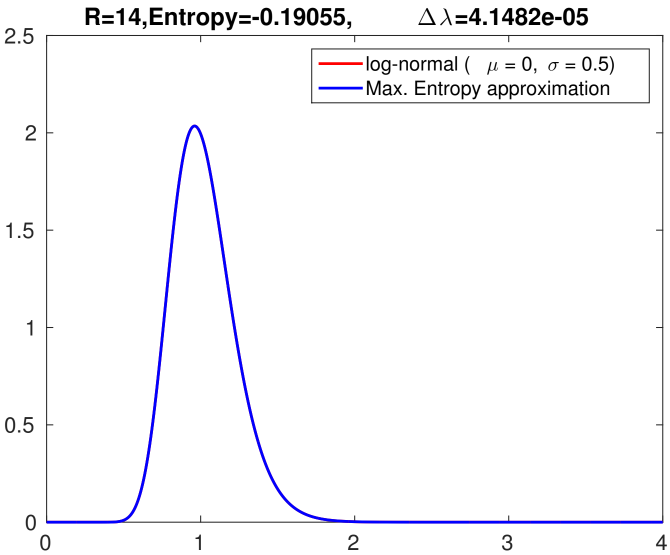
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



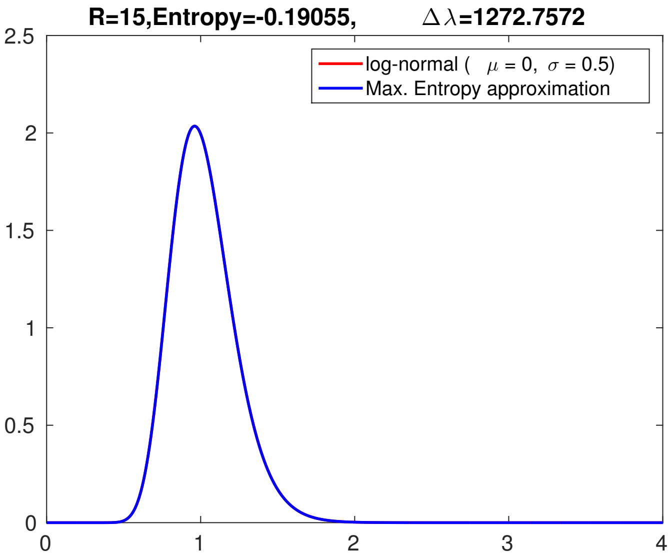
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



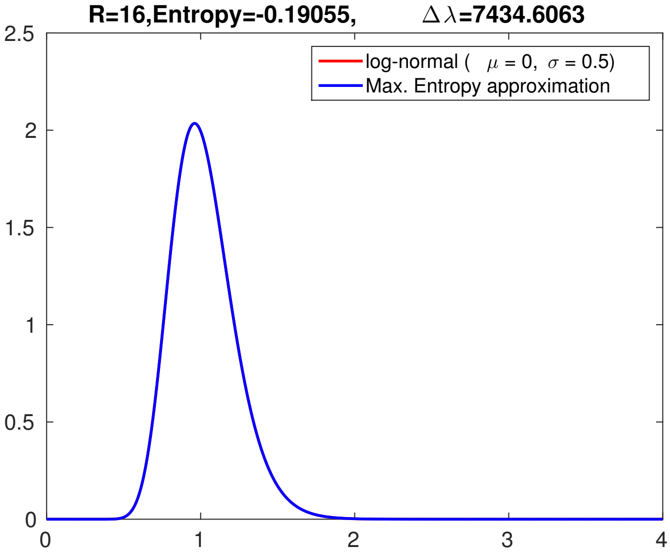
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



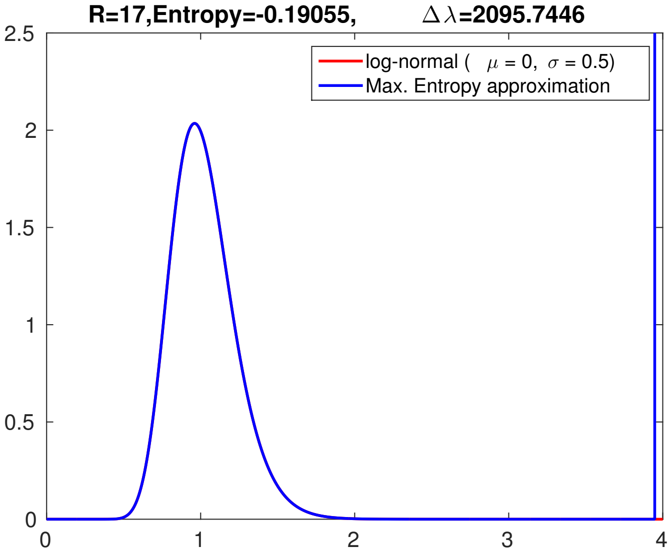
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



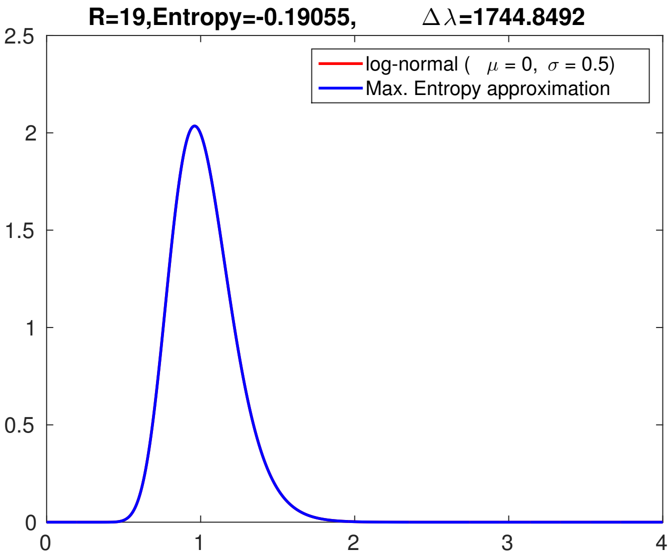
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



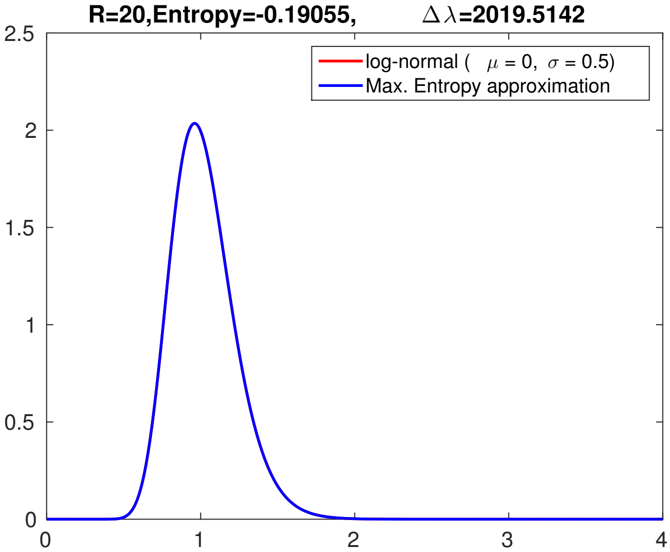
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



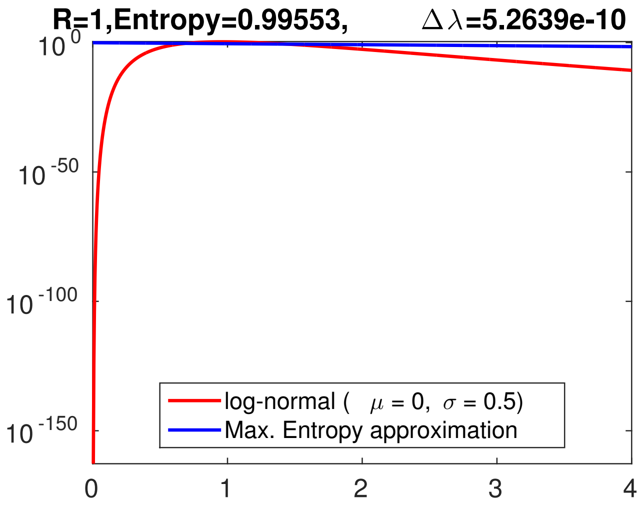
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



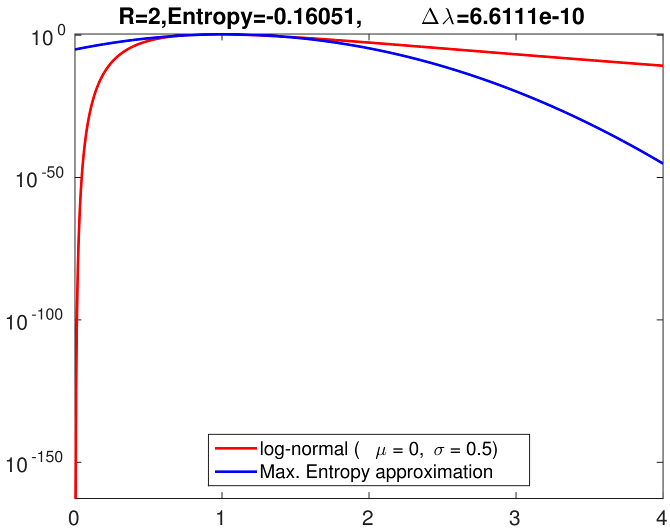
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$):

- Still quite stable even without convergence!
- Entropy is still monotonously decreasing!



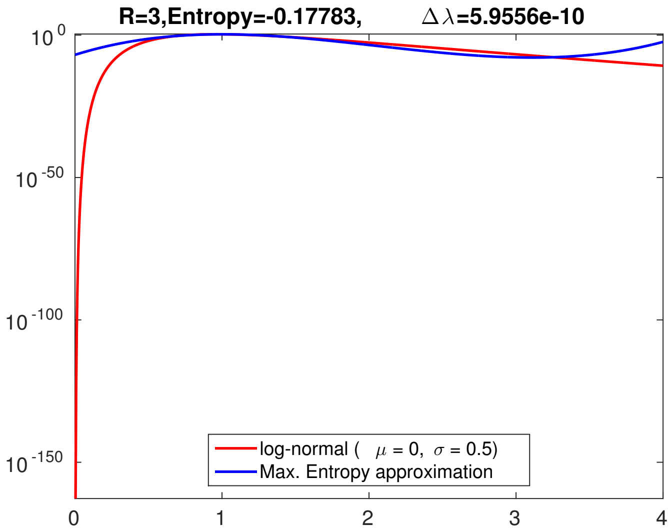
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



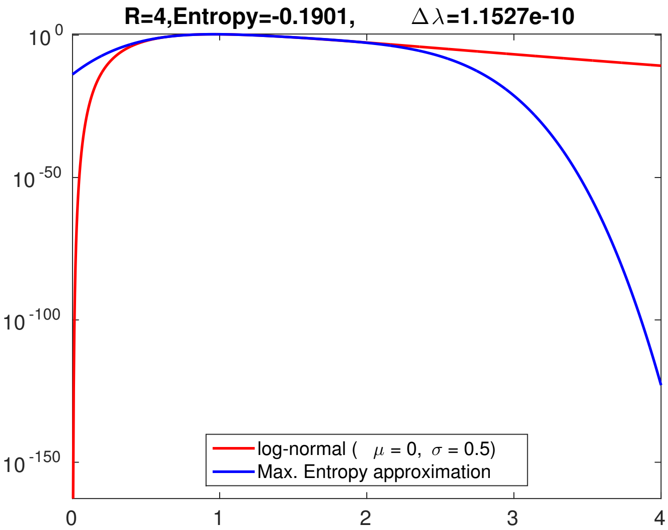
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



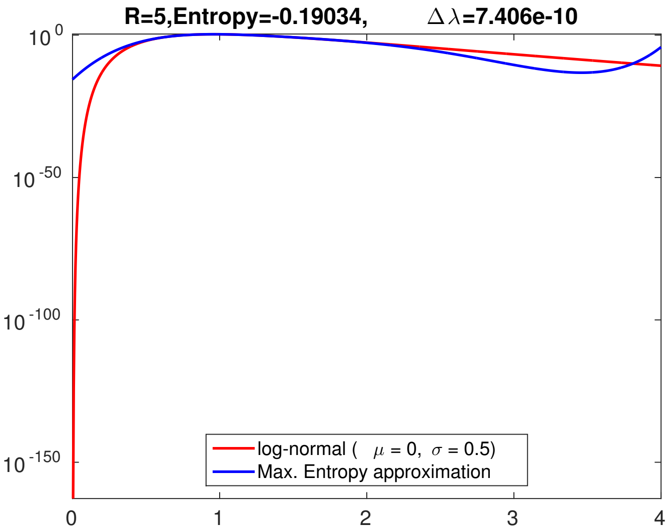
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



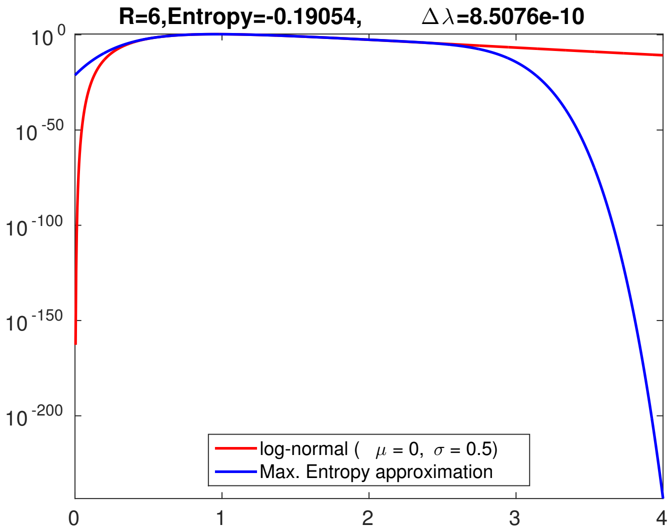
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



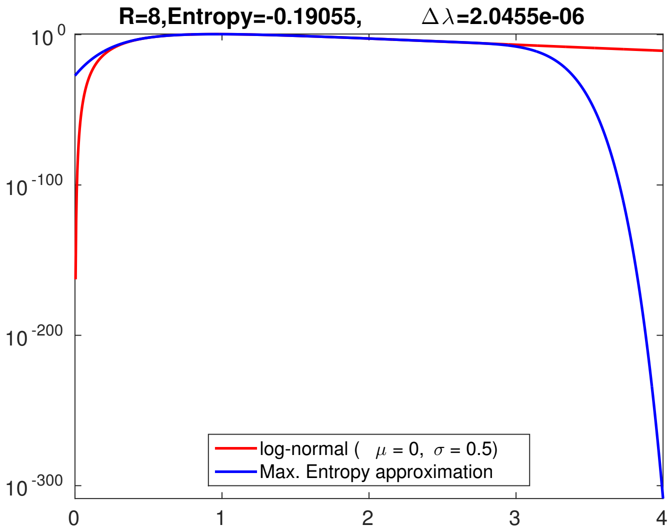
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



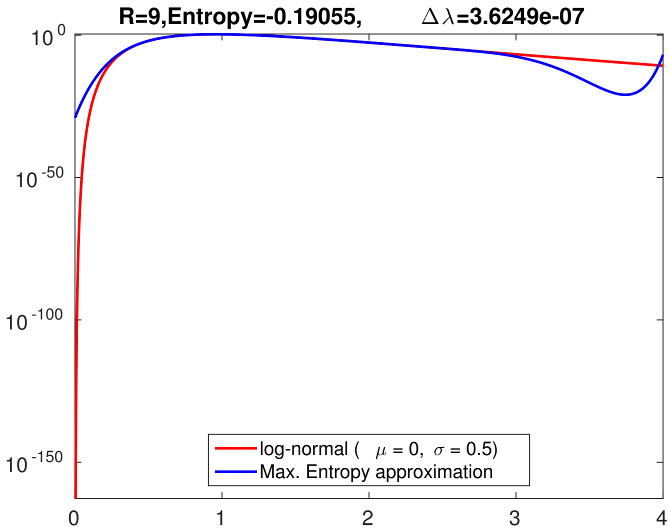
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



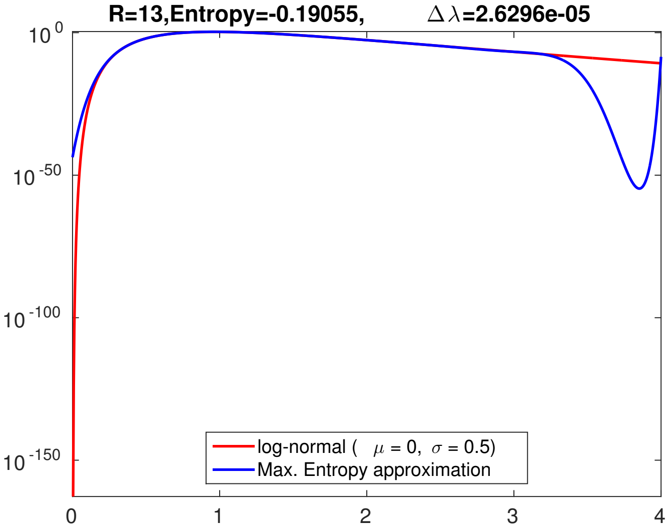
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



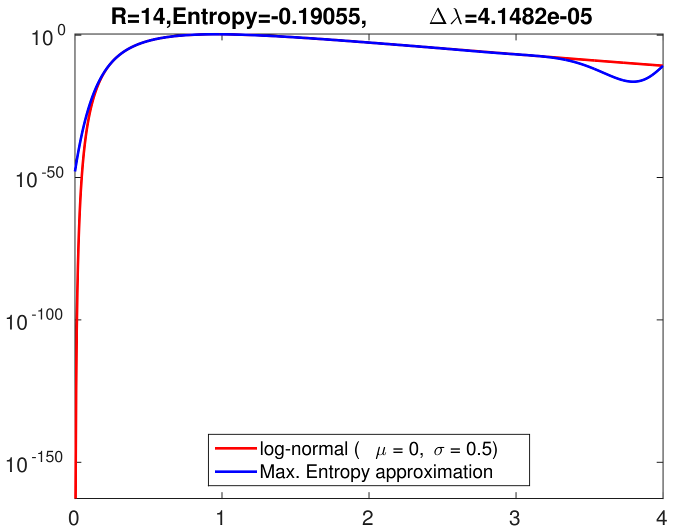
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



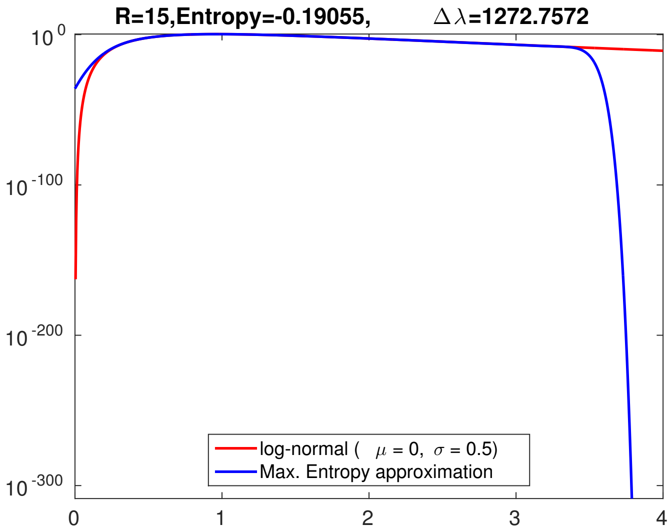
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



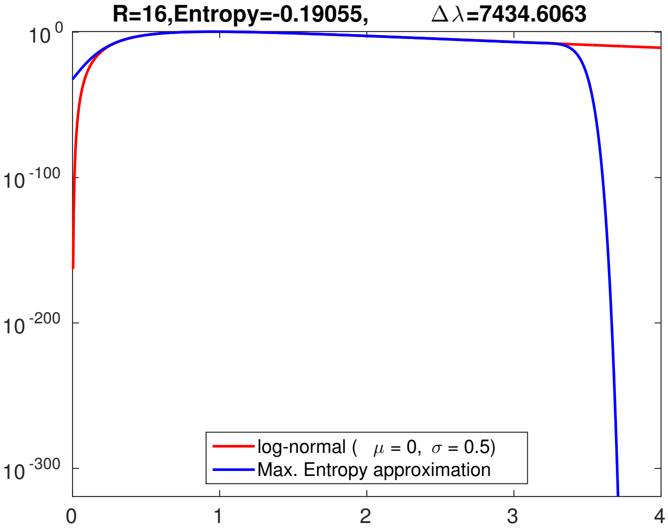
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



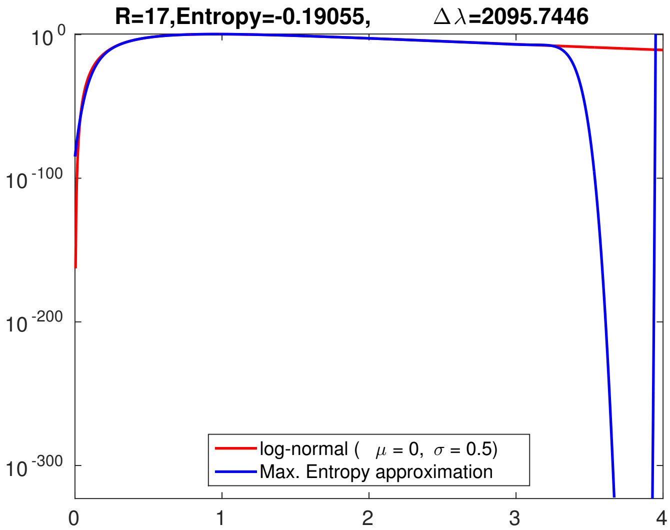
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



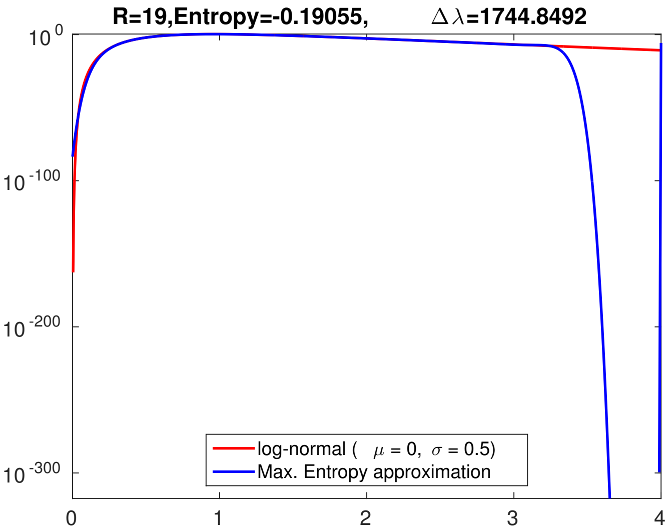
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



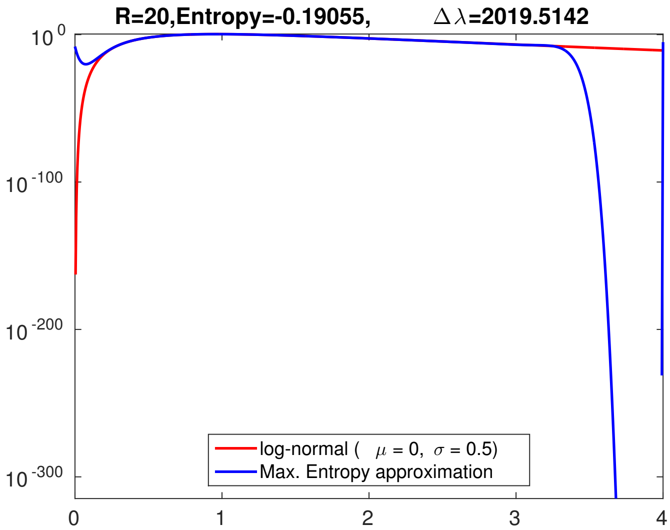
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



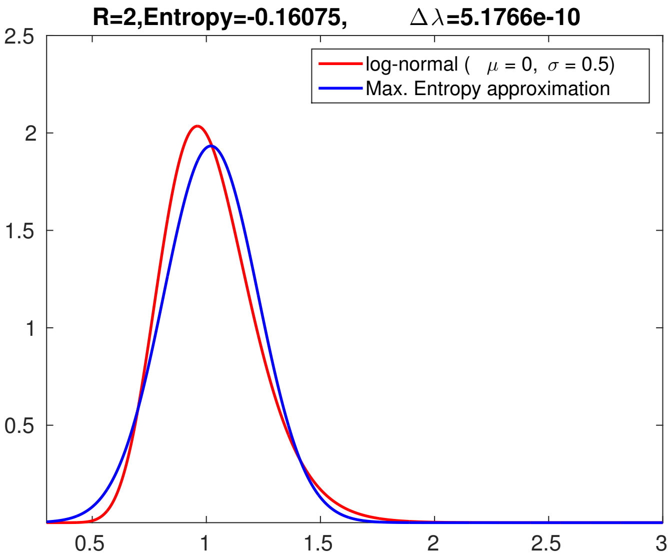
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



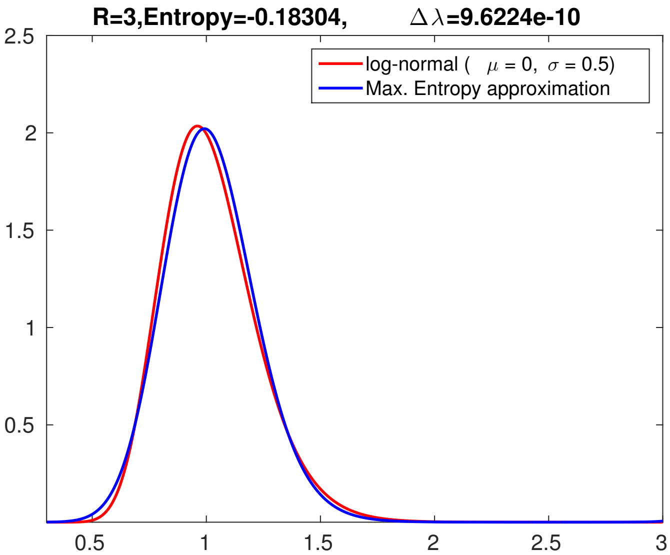
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0, 4]$, semilog):

- Oscillations in the negative domain
- $\Rightarrow$ stability of the density



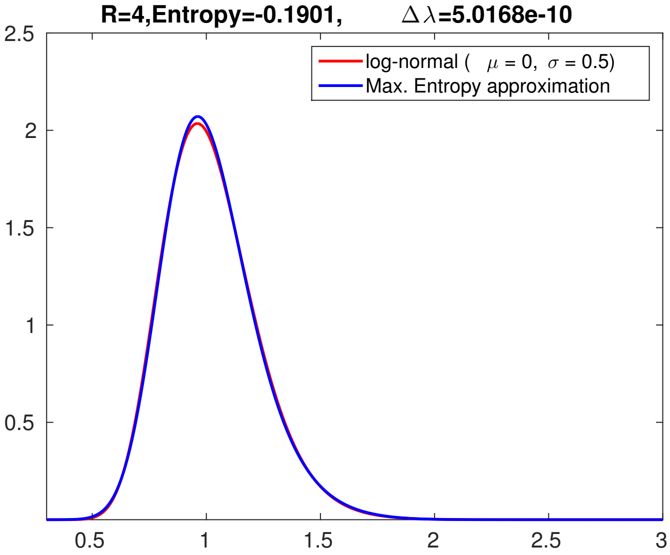
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



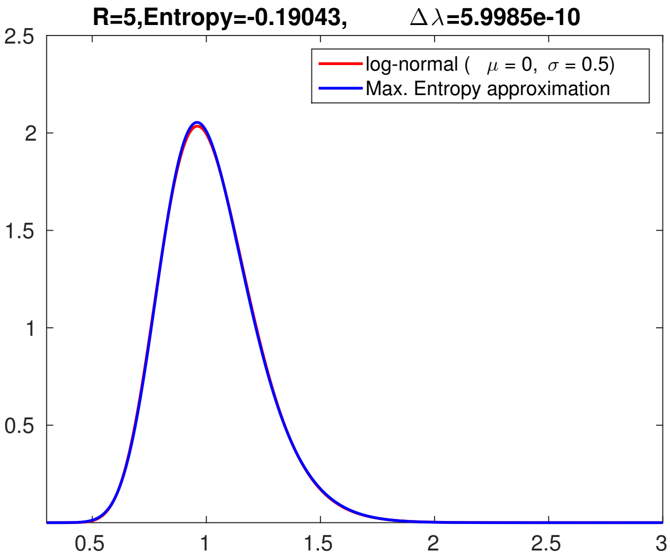
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



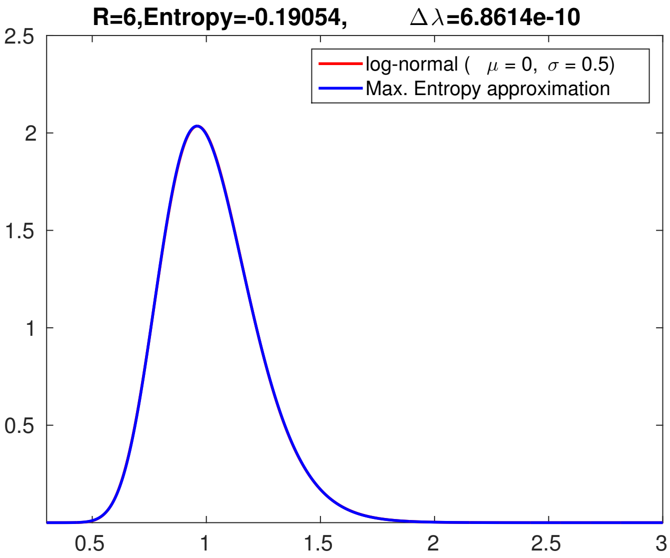
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



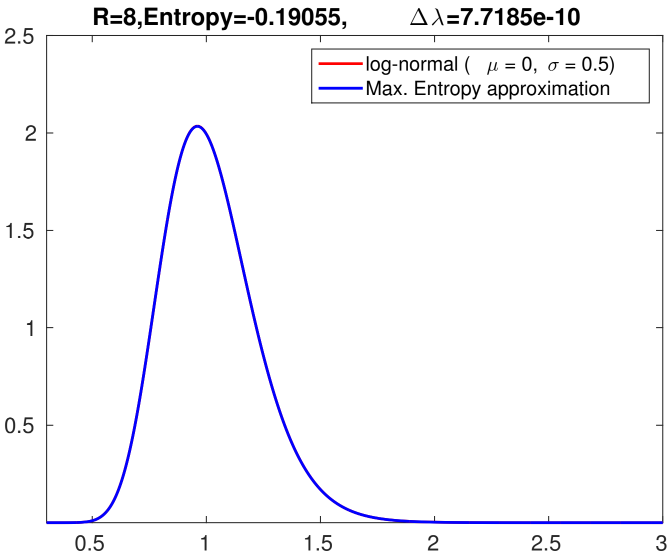
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



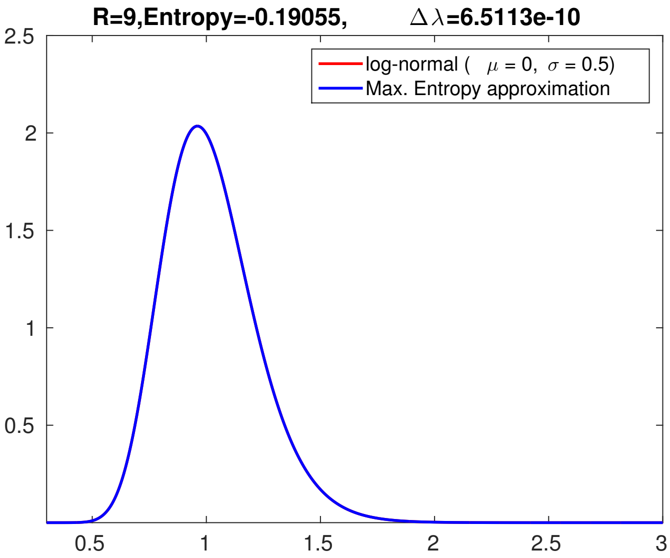
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



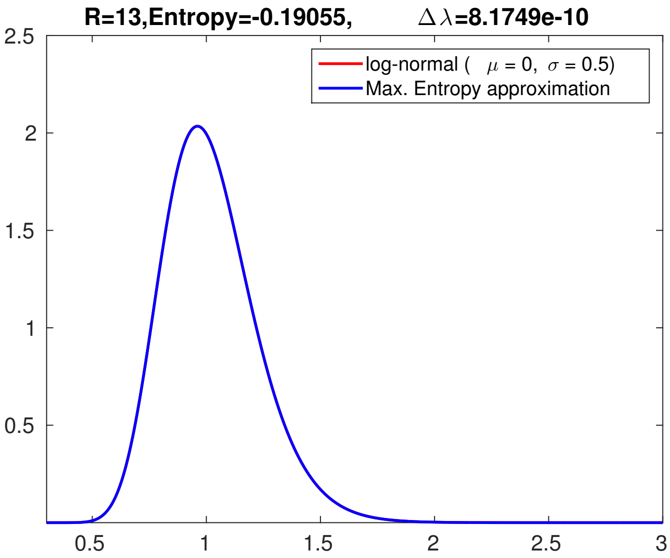
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



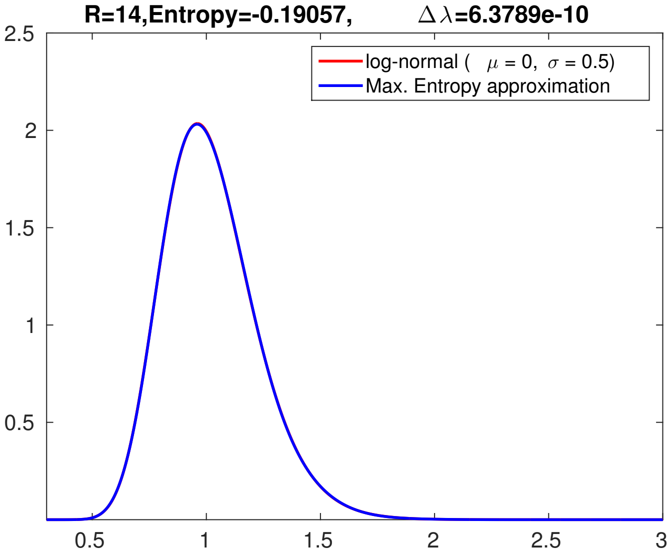
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



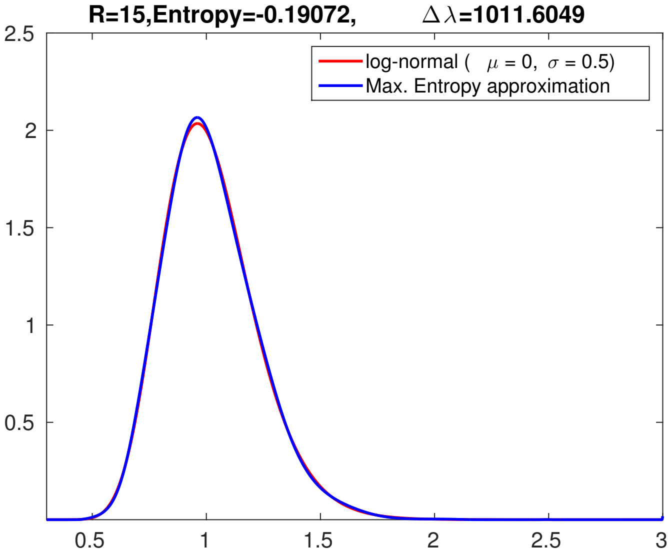
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



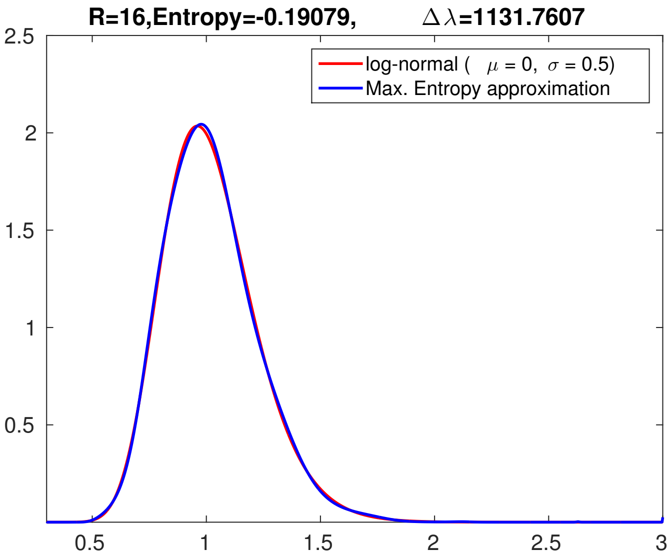
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



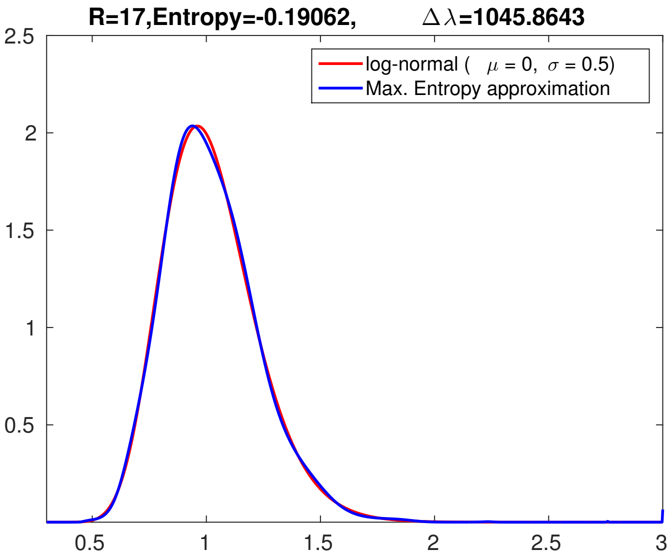
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



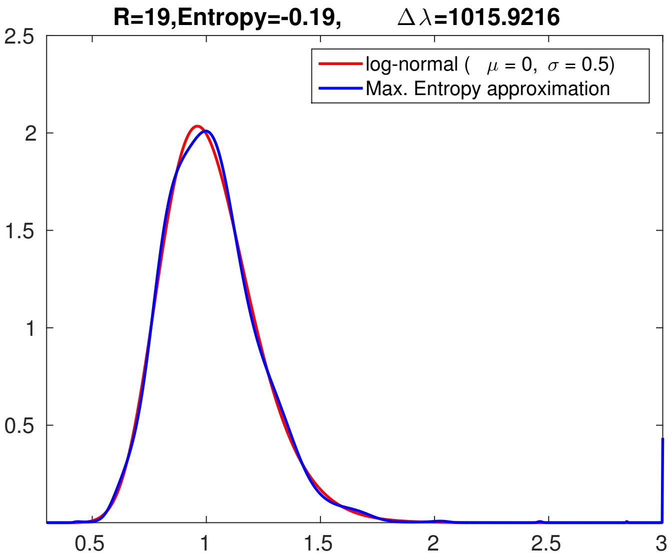
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



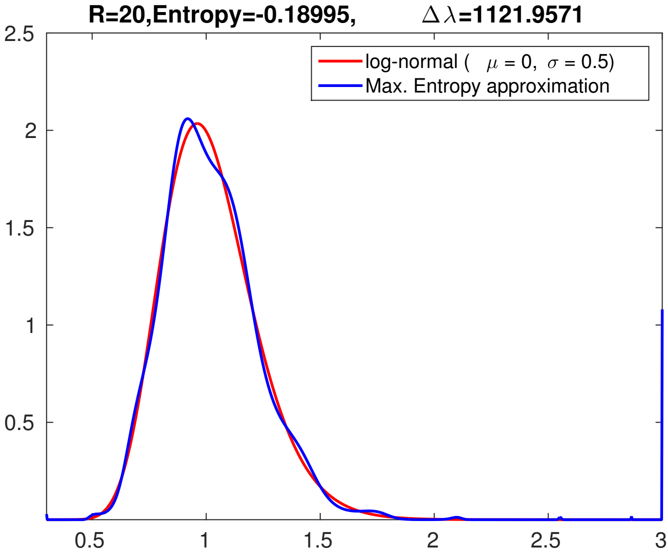
Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)



Legendre Moments ($\sigma = 0.2$, $[a, b] = [0.3, 3]$):

- Is stable and convergent for a bigger range $R \leq 14$!
- Entropy is monoton. decr. (stat. accuracy of MC is reached?)

- **Observation:** μ_k are not known exactly, we only have

$$\tilde{\mu}_k \approx \mu_k, \qquad k = 1, \ldots, R$$

(approx. moments $\tilde{\mu}_k$ may be computed e.g. by MC or MLMC).

$$\boxed{\begin{array}{c}\text{Moment matching with } \tilde{\mu}_k \quad \Rightarrow \quad \text{perturbed density} \\[1em] \tilde{\rho}_R \in \mathcal{E} := \left\{ p_\theta \propto \exp\left(\sum_{k=0}^{R} \theta_k \phi_k(x)\right), \theta_k \in \mathbb{R} \right\}.\end{array}}$$

- **Question:** What is the distance between ρ and $\tilde{\rho}_R$?

Consider the natural "metric"

$$D_{KL}(p\|q) := \int p(x) \ln \frac{p(x)}{q(x)}\, dx \qquad \left(\begin{matrix}\text{Kullback-Leibler divergence}\\ \text{or relative entropy}\end{matrix}\right)$$

"Pythagoras theorem"

$$D_{KL}(\rho\|\tilde{\rho}_R) = \underbrace{D_{KL}(\rho\|\rho_R)}_{\text{trunc. error}} + \underbrace{D_{KL}(\rho_R\|\tilde{\rho}_R)}_{\text{estim. error}}.$$

- ρ is the exact density
- $\rho_R \in \mathcal{E}$ is the ME solution for **exact moments** $\mu_1, \ldots, \mu_R$
- $\tilde{\rho}_R \in \mathcal{E}$ is the ME solution for **approx. moments** $\tilde{\mu}_1, \ldots, \tilde{\mu}_R$

- Truncation error: $D_{KL}(\rho\|\rho_R) \to 0$ when $R \to \infty$
- Estimation error: $D_{KL}(\rho_R\|\tilde{\rho}_R) \to 0$ when $\tilde{\mu}_k \to \mu_k$.

Our aim is the rigorous quantification of these statements.

From D_{KL} to L^p-norms:

Relation to L^p-norms

$$i)\ \frac{1}{2}\|\rho-\eta\|_{L^1}^2 \le D_{KL}(\rho\|\eta); \quad ii)\ C_p\|\rho-\eta\|_{L^p}^p \le D_{KL}(\rho\|\eta), \quad \rho,\eta \in L^\infty.$$

Relation to L^2-norms of the log-density

For two pdf's ρ and η with $\ln(\rho/\eta) \in L^\infty$

$$D_{KL}(\rho\|\eta) \ge \frac{1}{2} e^{-\|\ln(\rho/\eta)\|_{L^\infty}} \int \left|\ln(\rho/\eta)\right|^2 \rho\, dx \qquad \text{and}$$

$$D_{KL}(\rho\|\eta) \le \frac{1}{2} e^{\|\ln(\rho/\eta)-c\|_{L^\infty}} \int \left|\ln(\rho/\eta)-c\right|^2 \rho\, dx \qquad (\#)$$

for any $c \in \mathbb{R}$.

From D_{KL} to L^p-norms:

Relation to L^p-norms

$$i)\ \frac{1}{2}\|\rho-\eta\|_{L^1}^2 \le D_{KL}(\rho\|\eta); \quad ii)\ C_p\|\rho-\eta\|_{L^p}^p \le D_{KL}(\rho\|\eta), \quad \rho,\eta\in L^\infty.$$

Relation to L^2-norms of the log-density

For two pdf's ρ and η with $\ln(\rho/\eta)\in L^\infty$

$$D_{KL}(\rho\|\eta) \ge \frac{1}{2}e^{-\|\ln(\rho/\eta)\|_{L^\infty}}\int \left|\ln(\rho/\eta)\right|^2\rho\,dx \qquad \text{and}$$

$$D_{KL}(\rho\|\eta) \le \frac{1}{2}e^{\|\ln(\rho/\eta)-c\|_{L^\infty}}\int \left|\ln(\rho/\eta)-c\right|^2\rho\,dx \qquad (\#)$$

for any $c\in\mathbb{R}$.

Truncation error is driven by the smoothness of ρ:

polynomial moments $\mathrm{span}\{1, \phi_1, \ldots, \phi_R\} = \mathcal{P}_R$	$\rightsquigarrow$	**polynomial best approximation**

Conv. of the truncation error

$$D_{KL}(\rho \| \rho_R) \leq \frac{1}{2} \inf_{v \in \mathcal{P}_R} \left[\exp\left\{ \| \ln \rho - v \|_{L^\infty} \right\} \| \ln \rho - v \|^2_{L^2(\rho)} \right]$$

$$\lesssim \begin{cases} R^{-2s}, & \text{when } \ln \rho \in H^s, s \geq 1 \\ \exp(-bR), & \text{when } \ln \rho \text{ is analytic.} \end{cases}$$

Proof: Choose $\eta = e^v / \int e^v$ for $\forall v \in \mathcal{P}_R$ in $(\#)$.

Truncation error is driven by the smoothness of ρ:

polynomial moments $\mathrm{span}\{1, \phi_1, \ldots, \phi_R\} = \mathcal{P}_R$	$\rightsquigarrow$	**polynomial best approximation**

Conv. of the truncation error

$$D_{KL}(\rho \| \rho_R) \le \frac{1}{2} \inf_{v \in \mathcal{P}_R} \left[\exp\left\{ \| \ln \rho - v \|_{L^\infty} \right\} \| \ln \rho - v \|^2_{L^2(\rho)} \right]$$

$$\lesssim \begin{cases} R^{-2s}, & \text{when } \ln \rho \in H^s, s \ge 1 \\ \exp(-bR), & \text{when } \ln \rho \text{ is analytic.} \end{cases}$$

Proof: Choose $\eta = e^v / \int e^v$ for $\forall v \in \mathcal{P}_R$ in $(\#)$.

Estimation error is driven by the perturbation of moments:

Conv. of the estimation error (a.s.-version)

Suppose $\|\mu - \tilde{\mu}\| \leq (2A_R C_R)^{-1}$ where A_R, C_R are explicit constants†, then $\tilde{\rho}_R$ exists and there holds

$$D_{KL}(\rho_R \| \tilde{\rho}_R) \leq C_R \|\mu - \tilde{\mu}\|^2$$

Estimation error is driven by the perturbation of moments:

Conv. of the estimation error (a.s.-version)

Suppose $\|\mu - \tilde{\mu}\| \leq (2A_R C_R)^{-1}$ where A_R, C_R are explicit constants[†],

then $\tilde{\rho}_R$ exists and there holds

$$D_{KL}(\rho_R \| \tilde{\rho}_R) \leq C_R \|\mu - \tilde{\mu}\|^2$$

[†] $A_R = \max\left\{\|v\|_{L^\infty}/\|v\|_{L^2} \ : \ v \in \mathcal{P}_R\right\}$ and $C_R = 2\exp\left\{1 + \|\ln \rho_R\|_\infty\right\}$.

Estimation error is driven by the perturbation of moments:

Conv. of the estimation error (a.s.-version)

Suppose $\|\mu - \tilde{\mu}\| \le (2A_R C_R)^{-1}$ where A_R, C_R are explicit constants†,

then $\tilde{\rho}_R$ exists and there holds

$$D_{KL}(\rho_R \| \tilde{\rho}_R) \le C_R \|\mu - \tilde{\mu}\|^2$$

Caution:

If $\tilde{\mu}_k$ is computed by MC $\quad \Rightarrow \quad \tilde{\rho}_R$ may fail to exist (with some prob.)

Estimation error is driven by the perturbation of moments:

Conv. of the estimation error (probabilistic version, simplified)

Suppose $\mathbb{E}\|\mu - \tilde{\mu}\|^2 \leq (2A_R C_R)^{-2}$ with the same A_R, C_R,

then $\tilde{\rho}_R$ exists with prob. $1 - p$ (for $p \in [p_*, 1]$) and it holds

$$D_{KL}(\rho_R \| \tilde{\rho}_R) \leq C_R p^{-1} \mathbb{E}\|\mu - \tilde{\mu}\|^2 \qquad (*)$$

Estimation error is driven by the perturbation of moments:

Conv. of the estimation error (probabilistic version, simplified)

Suppose $\mathbb{E}\|\mu - \tilde{\mu}\|^2 \leq (2A_R C_R)^{-2}$ with the same A_R, C_R,

then $\tilde{\rho}_R$ exists with prob. $1 - p$ (for $p \in [p_*, 1]$) and it holds

$$D_{KL}(\rho_R \| \tilde{\rho}_R) \leq C_R p^{-1} \mathbb{E}\|\mu - \tilde{\mu}\|^2 \qquad (*)$$

Here $\mathbb{E}\|\mu - \tilde{\mu}\|^2 = \sum_{k=1}^{R} \underbrace{\mathbb{E}|\mu_k - \tilde{\mu}_k|^2}_{\text{standard } \mathrm{MSE}(\tilde{\mu}_k)}$ $\leftarrow$ $\boxed{\text{can be handled by the standard (ML)MC theory}}$

Structure of the estimate:

$$D_{KL}(\rho\|\tilde{\rho}_R) = D_{KL}(\rho\|\rho_R) + D_{KL}(\rho_R\|\tilde{\rho}_R)$$

$$\Downarrow \quad \text{assuming } \ln\rho \in H^1$$

$$D_{KL}(\rho\|\tilde{\rho}_R) \lesssim \boxed{\begin{array}{c}\text{polynomial}\\ \text{best approx.}\end{array}} + \frac{C_R}{p}\sum_{k=1}^{R} \boxed{\mathrm{MSE}(\tilde{\mu}_k)}$$

$$\Downarrow \quad \text{assuming } \left\{ \begin{array}{c} \ln\rho \in H^s, s \geq 1 \\ \ln\rho \text{ is analytic} \end{array} \right\}$$

$$D_{KL}(\rho\|\tilde{\rho}_R) \lesssim \left\{ \begin{array}{c} R^{-2s} \\ \exp(-bR) \end{array} \right\} + \frac{C_R}{p}\sum_{k=1}^{R} \mathrm{Bias}(\tilde{\mu}_k)^2 + \mathrm{Var}(\tilde{\mu}_k)$$

- C_R is unif. bounded when $\ln\rho \in H^s, s > 1$
- Error vs. cost theorems $\leftarrow$ balancing the error contributions

Structure of the estimate:

$$D_{KL}(\rho \| \tilde{\rho}_R) = D_{KL}(\rho \| \rho_R) + D_{KL}(\rho_R \| \tilde{\rho}_R)$$

$$\Downarrow \quad \text{assuming } \ln \rho \in H^1$$

$$D_{KL}(\rho \| \tilde{\rho}_R) \lesssim \boxed{\begin{array}{c}\text{polynomial}\\ \text{best approx.}\end{array}} + \frac{C_R}{p} \sum_{k=1}^{R} \boxed{\mathrm{MSE}(\tilde{\mu}_k)}$$

$$\Downarrow \quad \text{assuming} \left\{ \begin{array}{c} \ln \rho \in H^s, s \geq 1 \\ \ln \rho \text{ is analytic} \end{array} \right\}$$

$$D_{KL}(\rho \| \tilde{\rho}_R) \lesssim \left\{ \begin{array}{c} R^{-2s} \\ \exp(-bR) \end{array} \right\} + \frac{C_R}{p} \sum_{k=1}^{R} \mathrm{Bias}(\tilde{\mu}_k)^2 + \mathrm{Var}(\tilde{\mu}_k)$$

- C_R is unif. bounded when $\ln \rho \in H^s, s > 1$
- Error vs. cost theorems $\leftarrow$ balancing the error contributions

Structure of the estimate:

$$D_{KL}(\rho\|\tilde{\rho}_R) = D_{KL}(\rho\|\rho_R) + D_{KL}(\rho_R\|\tilde{\rho}_R)$$

$$\Downarrow \quad \text{assuming } \ln\rho \in H^1$$

$$D_{KL}(\rho\|\tilde{\rho}_R) \lesssim \boxed{\begin{array}{c}\text{polynomial}\\ \text{best approx.}\end{array}} + \frac{C_R}{p}\sum_{k=1}^{R} \boxed{\mathrm{MSE}(\tilde{\mu}_k)}$$

$$\Downarrow \quad \text{assuming } \left\{ \begin{array}{c} \ln\rho \in H^s, s \geq 1 \\ \ln\rho \text{ is analytic} \end{array} \right\}$$

$$D_{KL}(\rho\|\tilde{\rho}_R) \lesssim \left\{ \begin{array}{c} R^{-2s} \\ \exp(-bR) \end{array} \right\} + \frac{C_R}{p}\sum_{k=1}^{R} \mathrm{Bias}(\tilde{\mu}_k)^2 + \mathrm{Var}(\tilde{\mu}_k)$$

- C_R is unif. bounded when $\ln\rho \in H^s, s > 1$
- Error vs. cost theorems $\leftarrow$ balancing the error contributions

Structure of the estimate:

$$D_{KL}(\rho\|\tilde{\rho}_R) = D_{KL}(\rho\|\rho_R) + D_{KL}(\rho_R\|\tilde{\rho}_R)$$

$$\Downarrow \quad \text{assuming } \ln\rho \in H^1$$

$$D_{KL}(\rho\|\tilde{\rho}_R) \lesssim \boxed{\begin{array}{c}\text{polynomial}\\ \text{best approx.}\end{array}} + \frac{C_R}{p}\sum_{k=1}^{R} \boxed{\mathrm{MSE}(\tilde{\mu}_k)}$$

$$\Downarrow \quad \text{assuming } \left\{ \begin{array}{c} \ln\rho \in H^s, s \geq 1 \\ \ln\rho \text{ is analytic} \end{array} \right\}$$

$$D_{KL}(\rho\|\tilde{\rho}_R) \lesssim \left\{ \begin{array}{c} R^{-2s} \\ \exp(-bR) \end{array} \right\} + \frac{C_R}{p}\sum_{k=1}^{R} \mathrm{Bias}(\tilde{\mu}_k)^2 + \mathrm{Var}(\tilde{\mu}_k)$$

- C_R is unif. bounded when $\ln\rho \in H^s, s > 1$
- Error vs. cost theorems $\quad\leftarrow\quad$ balancing the error contributions

Theorem (Monte Carlo-ME: Accuracy / Cost, simplified)

Assume in addition that

$$\text{a) } \mathbb{E}[(X - X_\ell)^2] \lesssim N_\ell^{-\beta}, \quad \text{b) } Cost(X_\ell) \lesssim N_\ell^{\gamma}$$

Then $D_{KL}(\rho \| \tilde{\rho}_R) < \varepsilon$ with probability $\geq 1 - p$ and for the

- **single level** approximation of Legendre moments:

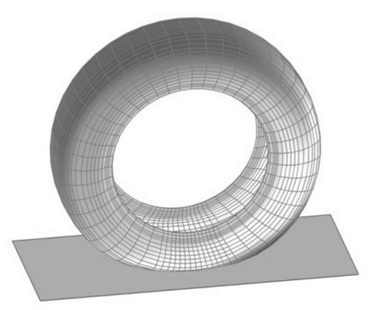
$$Cost(\tilde{\rho}_R) \lesssim (p\varepsilon)^{-\frac{\beta+\gamma}{\beta}} \begin{cases} \varepsilon^{-\frac{\beta+\gamma(\delta+1)}{2s\beta}} & \text{if } \ln(\rho) \in H^s, \\ |\ln(\varepsilon)|^{\frac{\beta+\gamma(\delta+1)}{\beta}} & \text{if } \ln(\rho) \text{ is analytic;} \end{cases}$$

- **multilevel** approximation of Legendre moments:

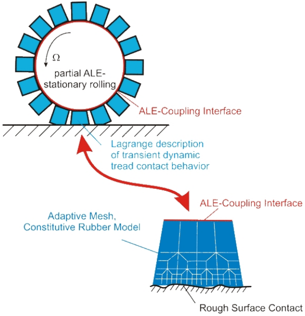
$$Cost(\tilde{\rho}_R) \lesssim (p\varepsilon)^{-\frac{\max(\beta,\gamma)}{\beta}} \begin{cases} \varepsilon^{-\frac{(\delta+1)\max(\beta,\gamma)}{2s\beta}} & \text{if } \ln(\rho) \in H^s, \\ |\ln(\varepsilon)|^{\frac{(\delta+1)\max(\beta,\gamma)}{\beta}} & \text{if } \ln(\rho) \text{ is analytic.} \end{cases}$$

Here δ is such that $\mathbb{E}[(\phi_k(X) - \phi_k(X_\ell))^2] \lesssim k^\delta N_\ell^{-\beta}$ and $\beta \neq \gamma$.

Contact with rough surfaces



$\psi(x)$



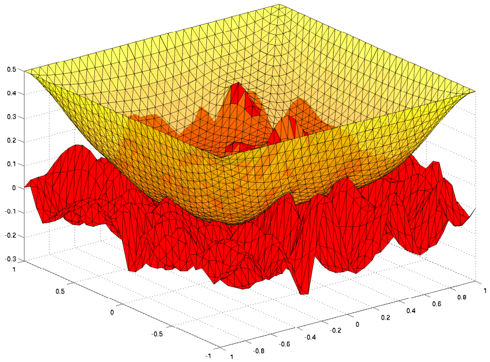
Courtesy: Prof. Udo Nackenhorst, IBNM, Univ. Hannover

Unknown parameter: $\psi(x)$ is the road surface profile.
(irregular microstructure)

Model: Contact of an elastic membrane with a rough surface (2d)

$$\left.\begin{array}{c} -\Delta u \geq f, \quad u \geq \psi, \\ (\Delta u + f)(u - \psi) = 0, \end{array}\right\} \quad \text{in } D,$$

$$u = 0 \qquad \text{on } \partial D.$$



$\left(\begin{array}{c}\text{here} \\ D = [-1, 1]^2\end{array}\right)$

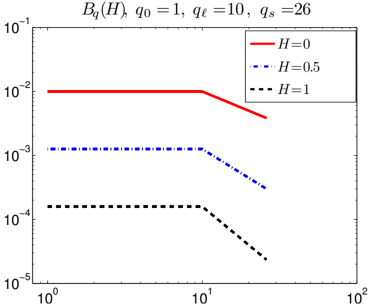
QoI: Membrane deformation $u(x)$; Contact Area $\Lambda(\omega) = \{x : u(x) = \psi(x)\}$.

Example: Rough obstacle models

Power spectrum [Persson et al.'05]:

$$\psi(x) = \sum_{q_0 \le |q| \le q_s} B_q(H) \cos(q \cdot x + \varphi_q)$$

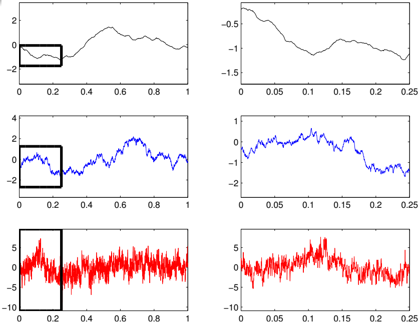
where $B_q(H) = \frac{\pi}{5}(2\pi \max(|q|, q_l))^{-H-1} \rightarrow$



Many materials in Nature and technics obey this law for amplitudes.

$H \sim \mathcal{U}(0, 1)$ random roughness

$\varphi_q \sim \mathcal{U}(0, 2\pi)$ random phase

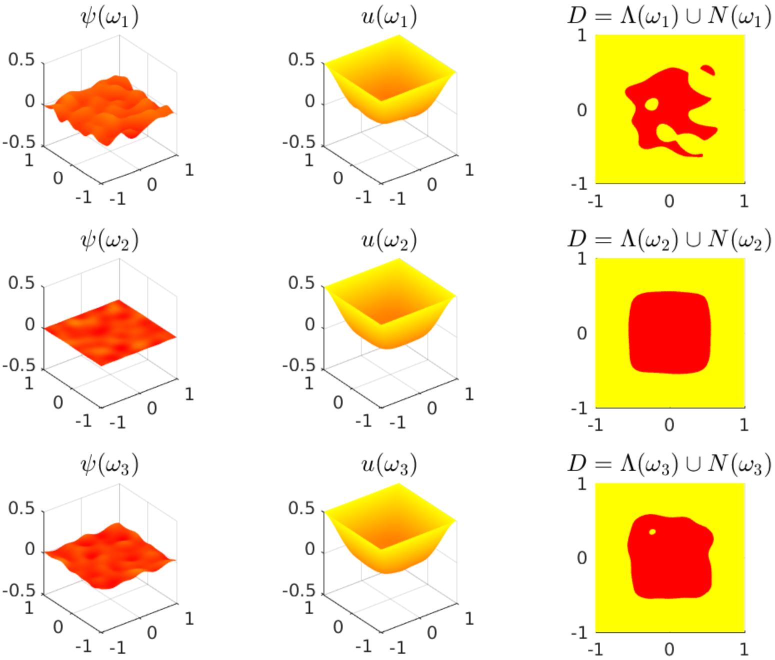


Forward solver:

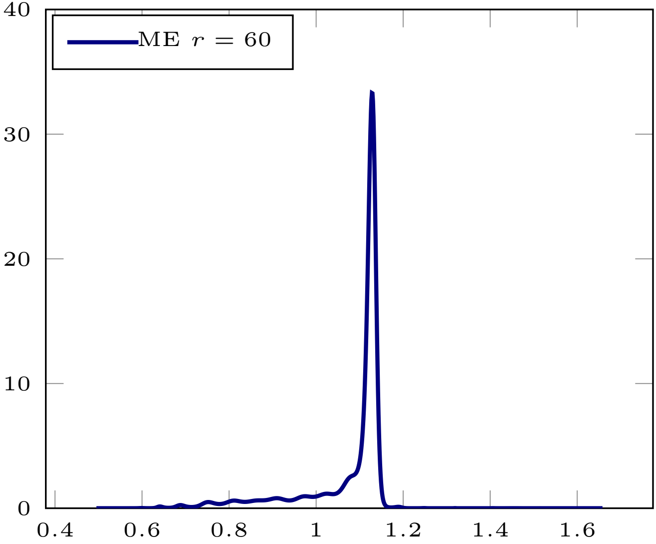
Own implementation of MMG (TNNM)

[Kornhuber'94,...]

Obstacle surfaces of variable/random roughness $\psi = \psi(x, \omega)$:



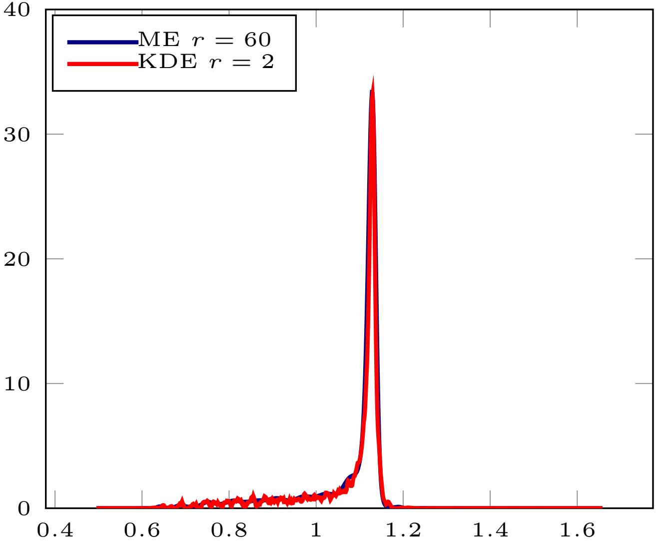
Estimation of the PDF ρ_X of the contact area $X = |\Lambda|$ by the Maximum Entropy method



The peak(s) corresponds to ca. 28.2% of the membrane in contact with the surface

More experiments and rigorous error analysis in [Bierig/Chernov, JCP, 2016].

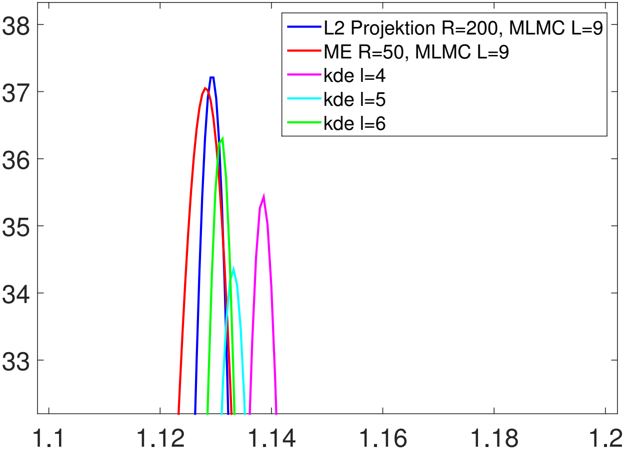
Estimation of the PDF ρ_X of the contact area $X = |\Lambda|$ by the Maximum Entropy method



The peak(s) corresponds to ca. 28.2% of the membrane in contact with the surface

More experiments and rigorous error analysis in [Bierig/Chernov, JCP, 2016].

Estimation of the PDF ρ_X of the contact area $X = |\Lambda|$ by the Maximum Entropy method



The peak(s) corresponds to ca. 28.2% of the membrane in contact with the surface

More experiments and rigorous error analysis in [Bierig/Chernov, JCP, 2016].

Summary:

- Rigorous convergence analysis for the MLMC-Maximum Entropy Method for compactly suported densities
- Num. experiments: smoothness assumptions may be relaxed
- Open issue: how to select the number of moment constraints R in practical computations (in the pre-asymptotic regime)? Adaptivity?
- With appropriate R the method is able to produce good approximations for a broad class of densities.

Thank you for your attention!

Summary:

- Rigorous convergence analysis for the MLMC-Maximum Entropy Method for compactly suported densities
- Num. experiments: smoothness assumptions may be relaxed
- Open issue: how to select the number of moment constraints R in practical computations (in the pre-asymptotic regime)? Adaptivity?
- With appropriate R the method is able to produce good approximations for a broad class of densities.

Thank you for your attention!

Summary:

- Rigorous convergence analysis for the MLMC-Maximum Entropy Method for compactly suported densities
- Num. experiments: smoothness assumptions may be relaxed
- Open issue: how to select the number of moment constraints R in practical computations (in the pre-asymptotic regime)? Adaptivity?
- With appropriate R the method is able to produce good approximations for a broad class of densities.

Thank you for your attention!

Summary:

- Rigorous convergence analysis for the MLMC-Maximum Entropy Method for compactly suported densities

- Num. experiments: smoothness assumptions may be relaxed

- Open issue: how to select the number of moment constraints R in practical computations (in the pre-asymptotic regime)? Adaptivity?

- With appropriate R the method is able to produce good approximations for a broad class of densities.

Thank you for your attention!

Summary:

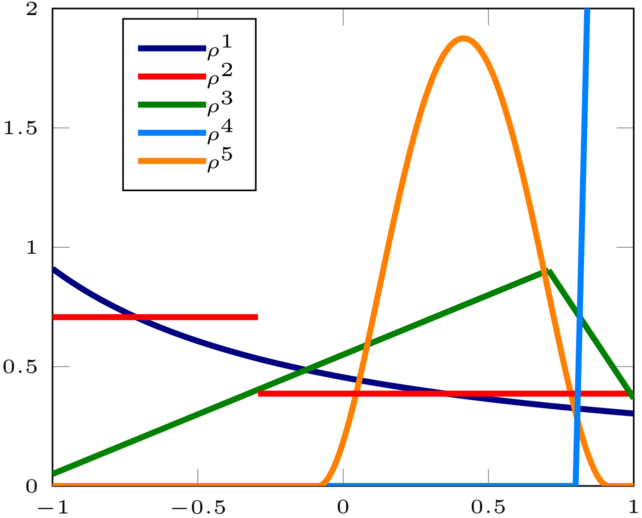
- Rigorous convergence analysis for the MLMC-Maximum Entropy Method for compactly suported densities
- Num. experiments: smoothness assumptions may be relaxed
- Open issue: how to select the number of moment constraints R in practical computations (in the pre-asymptotic regime)? Adaptivity?
- With appropriate R the method is able to produce good approximations for a broad class of densities.

Thank you for your attention!

References:

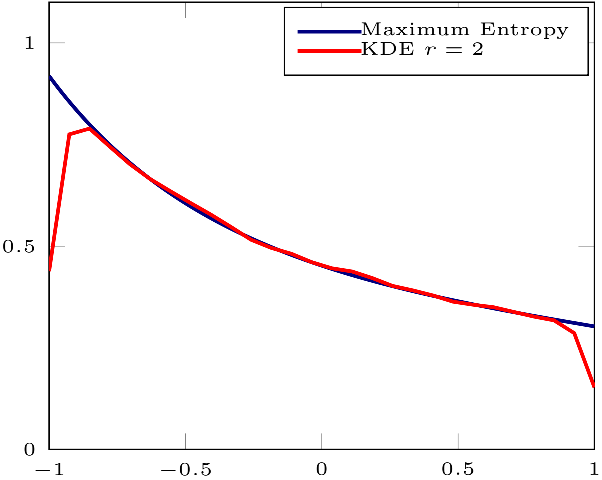
- C. Bierig and A. Chernov, Approximation of probability density functions by the Multilevel Monte Carlo Maximum Entropy method, *J. Comput. Physics*, 314 (2016), 661–681
- M. B. Giles, T. Nagapetyan, and K. Ritter, Multilevel Monte Carlo Approximation of Distribution Functions and Densities, SIAM/ASA *J. Uncert. Quantif.*, 3 (2015), pp. 267–295

Numerical examples for synthetic pdf's of different smoothness:



Exact synthetic pdf's $\rho_1, \ldots, \rho_5$

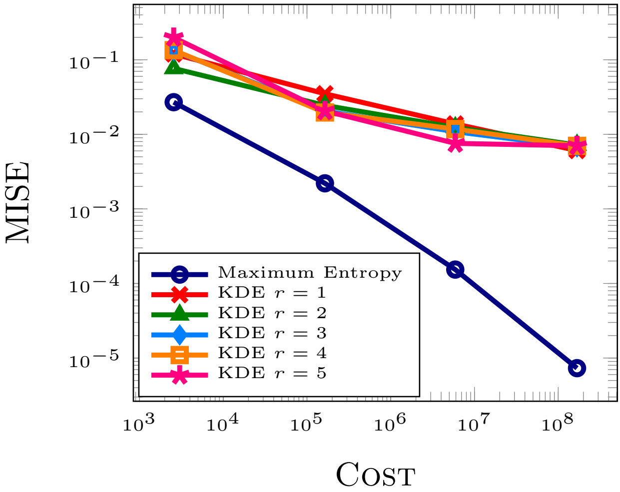
Numerical examples for synthetic pdf's of different smoothness:



$\ln \rho^1$ is analytic: smoothness assumptions are satisfied

Best maximum entropy approximation over $R = 1, \ldots, 60$ moments

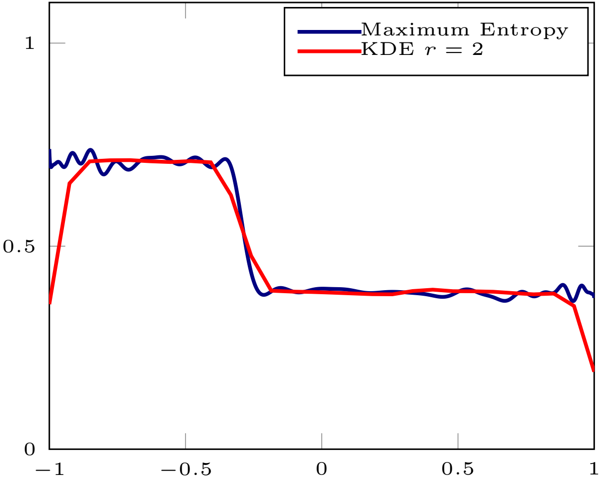
Numerical examples for synthetic pdf's of different smoothness:



$\ln \rho^1$ is analytic: smoothness assumptions are satisfied

Best maximum entropy approximation over $R = 1, \ldots, 60$ moments

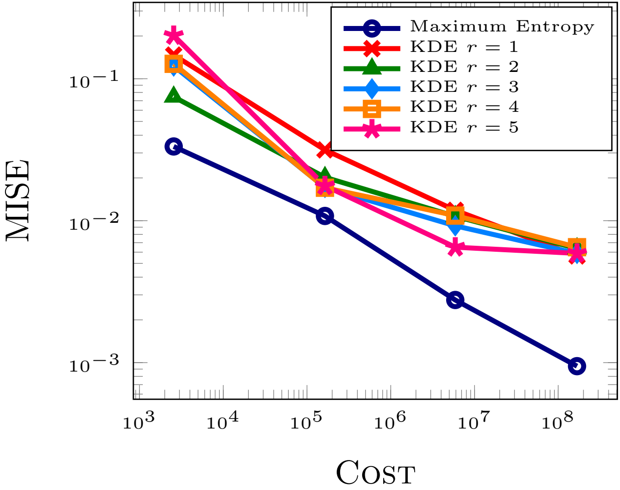
Numerical examples for synthetic pdf's of different smoothness:



In $\rho^2 \in H^{1/2-\epsilon}(-1,1)$: smoothness assumptions are not satisfied

Best maximum entropy approximation over $R = 1, \ldots, 60$ moments

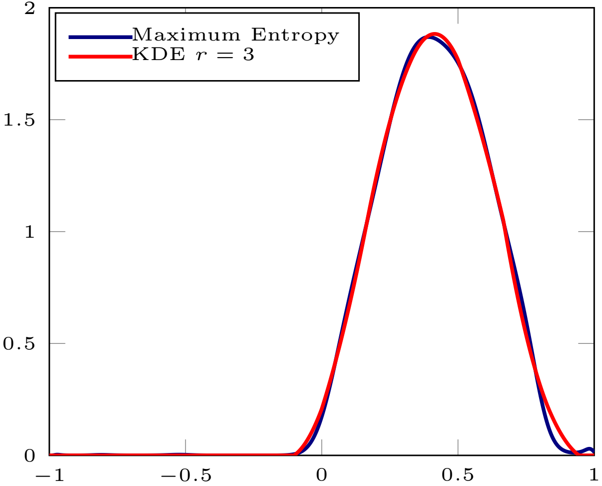
Numerical examples for synthetic pdf's of different smoothness:



In $\rho^2 \in H^{1/2-\epsilon}(-1,1)$: smoothness assumptions are not satisfied

Best maximum entropy approximation over $R = 1, \dots, 60$ moments

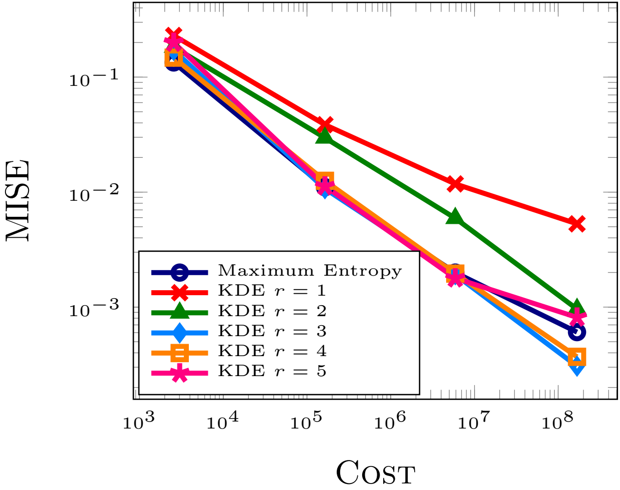
Numerical examples for synthetic pdf's of different smoothness:



In ρ^5 is unbounded: smoothness assumptions are not satisfied

Best maximum entropy approximation over $R = 1, \ldots, 60$ moments

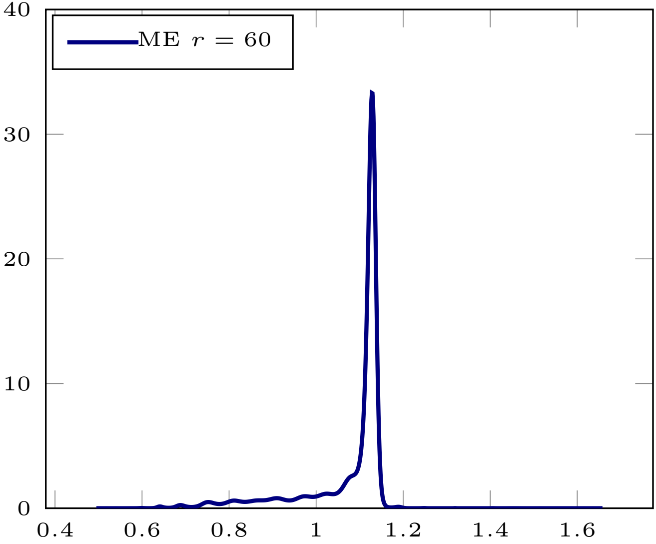
Numerical examples for synthetic pdf's of different smoothness:



In ρ^5 is unbounded: smoothness assumptions are not satisfied

Best maximum entropy approximation over $R = 1, \ldots, 60$ moments

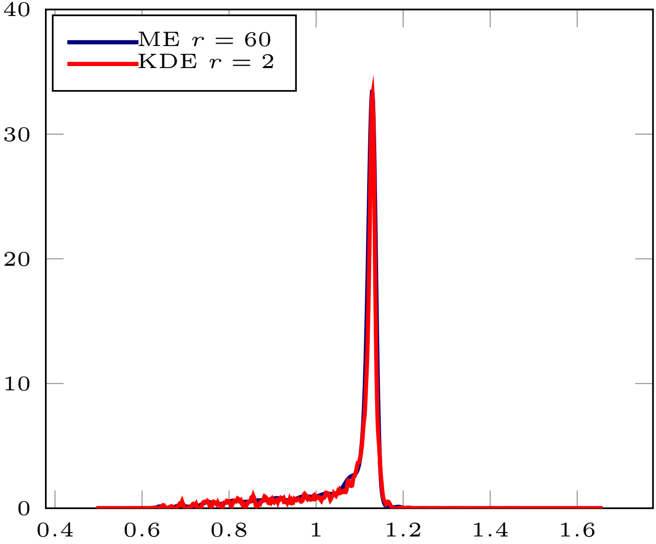
Estimation of the PDF ρ_X of the contact area $X = |\Lambda|$ by the Maximum Entropy method



The peak(s) corresponds to ca. 28.2% of the membrane in contact with the surface

More experiments and rigorous error analysis in [Bierig/Chernov, JCP, 2016].

Estimation of the PDF ρ_X of the contact area $X = |\Lambda|$ by the Maximum Entropy method



The peak(s) corresponds to ca. 28.2% of the membrane in contact with the surface

More experiments and rigorous error analysis in [Bierig/Chernov, JCP, 2016].