

# Adaptive Robust Control Under Model Uncertainty

Igor Cialenco

Department of Applied Mathematics

Illinois Institute of Technology

`cialenco@iit.edu`

Joint work with

T.R. Bielecki, T. Chen, A. Cousin, M. Jeanblanc

Mathematical Finance, Probability, and PDEs Conference  
Rutgers University, New Brunswick, May 17-19, 2017

# Motivations

- To minimize the risk due to model uncertainty (error in model estimation or model misspecification)
- To solve control problems when the true law of the underlying model is unknown and belongs to a family of probability laws
- Existing robust methodologies may be overly conservative when applied to unknown system
- Available robust frameworks do not consider reduction of uncertainty

## Main Goals

- We propose an **adaptive robust control methodology** for solving a discrete time Markovian control problem subject to Knightian uncertainty
- The methodology can be applied to any Markov decision process under model uncertainty, and in particular, to some financial problems
- Build the theory of recursive construction of confidence regions

## Preliminaries

- $(\Omega, \mathcal{F})$  - measurable space
- $T$  - finite time horizon
- $\mathcal{T} = \{0, 1, \dots, T\}$  and  $\mathcal{T}' = \{0, 1, \dots, T-1\}$
- $X = \{X_t, t \in \mathcal{T}\}$  - observed process
- $\mathbb{F} = \{\mathcal{F}_t, t \in \mathcal{T}\}$  - the natural filtration of  $X$
- $\{\mathbb{P}_\theta, \theta \in \Theta \subset \mathbb{R}^d\}$  - set of laws of  $X$
- $\Theta \subset \mathbb{R}^d$  - parameter space. **Model uncertainty if  $\Theta \neq \{\theta^*\}$ .**
- $\mathbb{P}_{\theta^*}$  - (unknown) true law of  $X$

# General Stochastic Control Problem

$$\inf_{\varphi \in \mathcal{A}} \mathbb{E}_{\theta^*} L(X, \varphi),$$

where  $\mathcal{A}$  is the set of admissible control processes (some  $\mathbb{F}$ -adapted processes  $\varphi = \{\varphi_t, t \in \mathcal{T}'\}$ );  $L$  is a measurable functional (loss function, utility function, etc).

**Example.** Find a self-financing portfolio that maximizes the expected utility of the terminal wealth

$$\sup_{\varphi \in \mathcal{A}} \mathbb{E}_{\theta^*} u(V_T^\varphi),$$

where  $V_T^\varphi$  is the wealth at time  $T$  using strategy  $\varphi$ ,  $u$  is an utility function,  $\mathcal{A}$  is the set of s.f. strategies.

Since the true parameter  $\theta^* \in \Theta$  is unknown, the question is how to handle the stochastic control problem subject to this type of *model uncertainty*.

# Classical Approaches

## Robust control problem:

$$\inf_{\varphi \in \mathcal{A}} \sup_{\theta \in \Theta} \mathbb{E}_{\theta}(L(X, \varphi)).$$

- Select the best strategy  $\varphi$  over the worst possible model.
- Başar and Bernhard (1995), Hansen et al. (2006), Hansen and Sargent (2008).
- If the ‘sup model’ is far from the true model, the solution could be bad.

## Strong robust control problem:

$$\inf_{\varphi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}^{\varphi, \Psi_K}} \mathbb{E}_{\mathbb{Q}}(L(S, \varphi)).$$

- $\Psi_K$  is the set of strategies chosen by Knightian adversary or the nature.
- $\mathcal{Q}^{\varphi, \Psi_K}$  is a set of probability measures on canonical space that represents all possible models resulting from  $\varphi$  and  $\Psi_K$ .
- Bayraktar, Cosso and Pham (2014), Sîrbu (2014).

## Bayesian adaptive control problem:

$$\inf_{\varphi \in \mathcal{A}} \int_{\Theta} \mathbb{E}_{\theta}(L(X, \varphi)) \nu_0(d\theta).$$

- The parameter  $\theta$  is modeled as a random variable taking values in  $\Theta$  with a prior distribution  $\nu_0$ .
- Kumar, Varaiya (1986), Runggaldier et al. (2002), Corsi et al. (2007).
- Note that

$$\inf_{\varphi \in \mathcal{A}} \sup_{\theta \in \Theta} \mathbb{E}_{\theta}(L(X, \varphi)) = \inf_{\varphi \in \mathcal{A}} \sup_{\nu_0 \in \mathcal{P}(\Theta)} \int_{\Theta} \mathbb{E}_{\theta}(L(X, \varphi)) \nu_0(d\theta).$$

No reduction of uncertainty about  $\theta$ .



## Adaptive control problem:

For each  $\theta \in \Theta$  solve

$$\inf_{\varphi \in \mathcal{A}} \mathbb{E}_{\theta} (L(X, \varphi)).$$

- Let  $\varphi^{\theta}$  be a corresponding optimal control.
- At each time  $t$ , compute an  $\mathcal{F}_t$ -measurable point estimate  $\widehat{\theta}_t$  of  $\theta^*$ .
- Apply the control value  $\varphi_t^{\widehat{\theta}_t}$ .
- Known to have poor performance for finite time horizon problems
- Kumar and Varaiya (1986), Chen and Guo (1991)

- The classical robust control problem does not involve any reduction of uncertainty about  $\theta^*$ ; the parameter space is not “updated” with incoming information about the signal process  $X$ .
- Same applies to other classical approaches.
- Incorporating “learning” into the robust control paradigm appears like a good idea.
- Anderson, Hansen, Sargent (2003) state:  
*“We see three important extensions to our current investigation. Like builders of rational expectations models, we have side-stepped the issue of how decision-makers select an approximating model. ... Just as we have not formally modelled how agents learned the approximating model, neither have we formally justified why they do not bother to learn about potentially complicated misspecifications of that model. **Incorporating forms of learning would be an important extension of our work.**”*

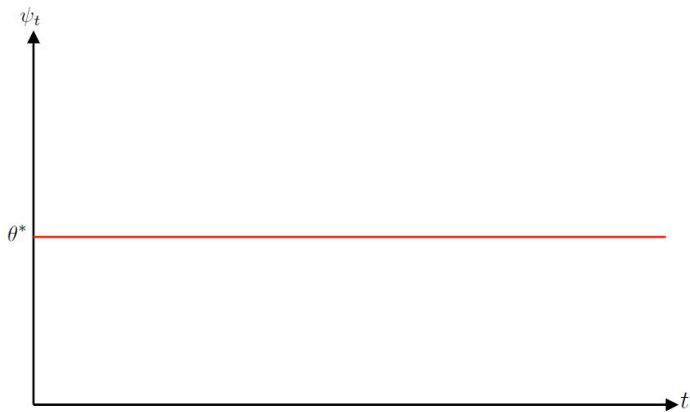
We propose **Adaptive Robust Control Methodology**

$$\inf_{\varphi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}_{h_0}^{\varphi, \Psi}} \mathbb{E}_{\mathbb{Q}} (L(X, \varphi)),$$

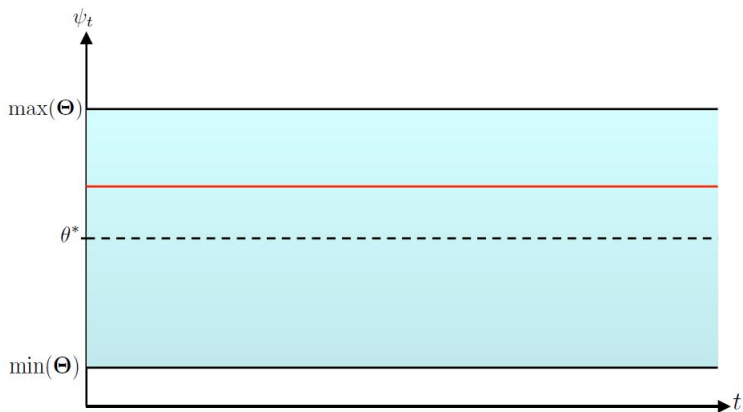
where  $\mathcal{Q}_{h_0}^{\varphi, \Psi}$  is a family of probability measures constructed in a way that allows for *dynamic reduction of uncertainty* about  $\theta^*$ .

We chose the family  $\mathcal{Q}_{h_0}^{\varphi, \Psi}$  in terms of *confidence regions* for the parameter  $\theta^*$ .

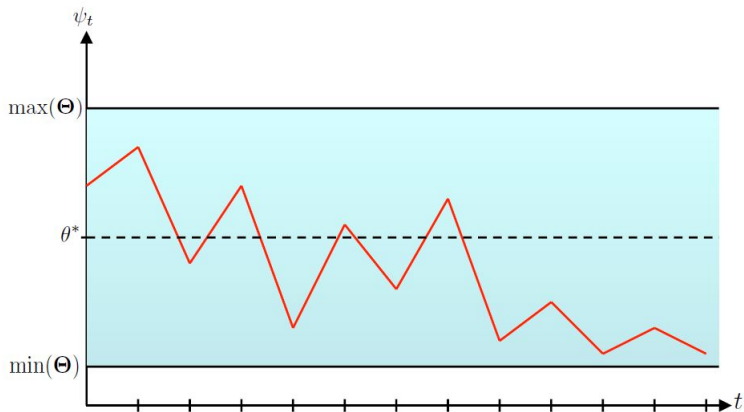
## Without uncertainty



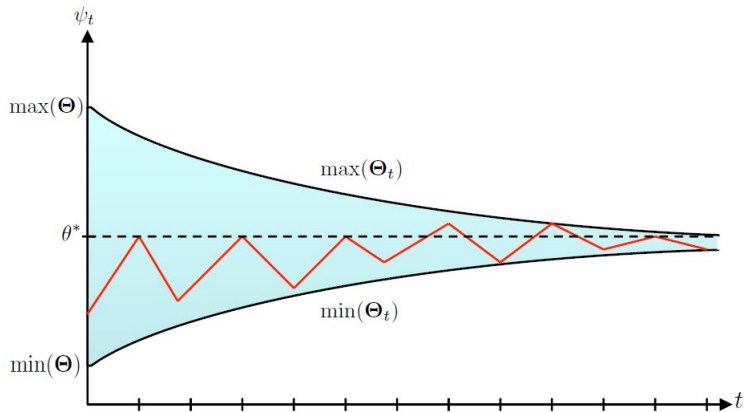
# Robust



## Strong robust



## Adaptive Robust



We assume that the observed state process  $X$  follows the dynamics

$$\begin{aligned} X_0 &= x_0, \\ X_{t+1} &= f(X_t, \varphi_t, Z_{t+1}), \quad t \in \mathcal{T}, \end{aligned}$$

where  $Z = \{Z_t\}_{t \in \mathcal{T}'}$  is an  $\mathbb{R}^m$ -valued random sequence that is

- $\mathbb{F}$ -adapted,
- observed,
- i.i.d. under  $\mathbb{P}_\theta$ ,  $\theta \in \Theta \subset \mathbb{R}^d$ ,
- the law of  $Z_1$  is unknown but it belongs to a family of parameterized distributions  $\{\mathbb{P}_\theta, \theta \in \Theta\}$ .

Example:  $Z_t$  is the excess return on risk assets in the optimal portfolio selection problem.



The adaptive robust control methodology relies *essentially* on **recursive construction of confidence regions**. In Bielecki, C., Chen (2016), for time-homogenous Markov chains  $Z$ , we showed that:

- A point estimator  $\hat{\theta}_t$  of  $\theta^*$  can be computed recursively

$$\begin{aligned}\hat{\theta}_0 &= \theta_0, \\ \hat{\theta}_{t+1} &= R(t, \hat{\theta}_t, Z_{t+1}),\end{aligned}$$

where  $R(t, c, z)$  is a deterministic measurable function.

- An approximate  $1 - \alpha$ -confidence region  $\Theta_t$  of  $\theta^*$  can be constructed by a deterministic rule:

$$\Theta_t = \tau_\alpha(t, \hat{\theta}_t)$$

where  $\tau_\alpha(t, \cdot) : \mathbb{R}^d \rightarrow 2^\Theta$  is a *deterministic* set valued function,  $\mathbb{P}_{\theta^*}(\theta^* \in \Theta_t) \approx 1 - \alpha$ , and  $\lim_{t \rightarrow \infty} \Theta_t = \{\theta^*\}$ .

We consider the augmented state process

$$Y_t := (X_t, \widehat{\theta}_t), \quad t \in \mathcal{T},$$

with state space  $E_Y := \mathbb{R}^n \times \mathbb{R}^d$ , and dynamics

$$Y_{t+1} = \mathbf{T}(t, Y_t, \varphi_t, Z_{t+1})$$

with  $\mathbf{T}(t, y, a, z) := (f(x, a, z), R(t, c, z))$  and  $y = (x, c)$ .

We define the histories

$$H_t := (Y_0, \dots, Y_t) \in \mathbf{H}_t := \underbrace{E_Y \times E_Y \times \dots \times E_Y}_{t+1 \text{ times}}.$$

The sequence of confidence regions  $\Theta_t$ ,  $t \in \mathcal{T}$ , will represent the learning about  $\theta^*$  based on the observed history  $H_t$ ,  $t \in \mathcal{T}$ .

A robust control problem can be viewed as a game between a controller and the nature.

The controller plays history-dependent strategies  $\varphi$  that belong to

$$\mathcal{A} = \{(\varphi_t)_{t \in \mathcal{T}'} \mid \varphi_t : \mathbf{H}_t \rightarrow A, t \in \mathcal{T}'\}.$$

**Adaptive robust control:** the nature plays history-dependent strategies  $\psi$  that belong to

$$\Psi = \{(\psi_t)_{t \in \mathcal{T}'} \mid \psi_t : \mathbf{H}_t \rightarrow \Theta_t, t \in \mathcal{T}'\}.$$

**Remark.** The *strong robust case* corresponds to

$$\Psi_K = \{(\psi_t)_{t \in \mathcal{T}'} \mid \psi_t : \mathbf{H}_t \rightarrow \Theta, t \in \mathcal{T}'\}.$$

Define the Markov transition probability kernel on  $\mathcal{E}_Y$

$$Q(B \mid t, y, a, \theta) := \mathbb{P}_\theta(Y_{t+1} \in B \mid Y_t = y, \varphi_t = a),$$

for each  $(t, y, a, \theta) \in \mathcal{T}' \times E_Y \times A \times \Theta$ ,

Using Ionescu-Tulcea theorem, define the canonical law of the state process  $Y$  on  $E_Y^{T+1}$  as

$$\begin{aligned} \mathbb{Q}_{h_0}^{\varphi, \psi}(B_0, B_1, \dots, B_T) \\ &= \int_{B_0} \cdots \int_{B_T} Q(dx_T \mid T-1, x_{T-1}, \varphi_{T-1}(h_{T-1}), \psi_{T-1}(h_{T-1})) \\ &\quad \cdots Q(dx_1 \mid 0, x_0, \varphi_0(h_0), \psi_0(h_0)) \delta_{h_0}(dx_0) \end{aligned}$$

## The adaptive robust hedging problem

$$\inf_{\varphi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}_{\mu}^{\varphi, \Psi}} \mathbb{E}_{\mathbb{Q}} L(X_T^{\varphi}),$$

where

$$\mathcal{Q}_{\mu}^{\varphi, \Psi} := \{\mathbb{Q}_{\mu}^{\varphi, \psi}, \psi \in \Psi\},$$

and

$$\Psi = \{(\psi_t)_{t \in \mathcal{T}'} \mid \psi_t : \mathbf{H}_t \rightarrow \Theta_t, t \in \mathcal{T}'\}.$$

# Adaptive Robust DPP

## Theorem

The solution  $\varphi^* = (\varphi_t^*(h_t))_{t \in \mathcal{T}'}$  of

$$\inf_{\varphi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}_{\mu}^{\varphi, \Psi}} \mathbb{E}_{\mathbb{Q}} L(X_T^{\varphi}),$$

is the same as the solution of the following Bellman equations:

$$W_T(y) = \ell(x),$$

$$W_t(y) = \inf_{a \in A} \sup_{\theta \in \mathcal{T}_{\alpha}(t, \tilde{\theta})} \int_{E_Y} W_{t+1}(y) Q(dy \mid t, y, a, \theta),$$

for any  $y = (x, \tilde{\theta}) \in E_Y$  and  $t = T - 1, \dots, 0$ .

Note that the optimal strategy at time  $t$  is such that  $\varphi_t^*(h_t) = \varphi_t^*(y_t)$ .

## Example: Dynamic Optimal Portfolio Selection

An investor is deciding on investing in a risky asset and a risk-free banking account by maximizing the expected utility  $u(V_T)$  of the terminal wealth:

$$\inf_{\varphi \in \mathcal{A}} \sup_{Q \in \mathcal{Q}_{h_0}^{\varphi, \Psi}} \mathbb{E}_Q[-u(V_T)]$$

- $r$  - the constant risk free rate
- $Z_t$  - the excess return on the risky asset
- Assume that  $Z_t = \mu + \sigma \varepsilon_t$ , where  $\varepsilon_t$  are i.i.d.  $\mathcal{N}(0, 1)$
- The dynamics of the wealth process produced by a s.f. strategy

$$V_{t+1} = V_t(1 + r + \varphi_t Z_{t+1}), \quad t \in \mathcal{T}',$$

with  $V_0 = v_0$ , and where  $\varphi_t$  is the proportion of the wealth invested in the risky asset from  $t$  to  $t + 1$ .

The MLE  $\hat{\theta}_t = (\hat{\mu}_t, \hat{\sigma}_t^2)$  of the unknown parameter  $\theta^* = (\mu^*, (\sigma^*)^2)$  can be expressed in the following recursive way:

$$\begin{aligned}\hat{\mu}_{t+1} &= \frac{t}{t+1}\hat{\mu}_t + \frac{1}{t+1}Z_{t+1}, \\ \hat{\sigma}_{t+1}^2 &= \frac{t}{t+1}\hat{\sigma}_t^2 + \frac{t}{(t+1)^2}(\hat{\mu}_t - Z_{t+1})^2\end{aligned}$$

Due to asymptotic normality of the MLEs, we have

$$\frac{t}{\hat{\sigma}_t^2}(\hat{\mu}_t - \mu^*)^2 + \frac{t}{2\hat{\sigma}_t^4}(\hat{\sigma}_t^2 - (\sigma^*)^2)^2 \xrightarrow[t \rightarrow \infty]{d} \chi_2^2.$$

Consequently, the recursive  $1 - \alpha$  confidence regions take the form

$$\Theta_t = \tau_\alpha(t, \hat{\mu}, \hat{\sigma}^2) := \left\{ (\mu, \sigma^2) \in \mathbb{R}^2 : \frac{t}{\hat{\sigma}^2}(\hat{\mu} - \mu)^2 + \frac{t}{2\hat{\sigma}^4}(\hat{\sigma}^2 - \sigma^2)^2 \leq \kappa_\alpha \right\}$$

with  $\kappa_\alpha$  being the  $(1 - \alpha)$ -quantile of the  $\chi_2^2$  distribution.



The Markov decision process  $Y_t = (V_t^\varphi, \hat{\mu}_t, \hat{\sigma}_t^2)$  has dynamics

$$Y_{t+1} = \mathbf{T}(t, Y_t, \varphi_t, Z_{t+1})$$

where

$$\mathbf{T}(t, v, \hat{\mu}, \hat{\sigma}^2, a, z) = \left( v(1 + r^f + az), \frac{t}{t+1}\hat{\mu} + \frac{1}{t+1}z, \frac{t}{t+1}\hat{\sigma}^2 + \frac{t}{(t+1)^2}(\hat{\mu} - z)^2 \right)$$

The corresponding robust Bellman equation becomes

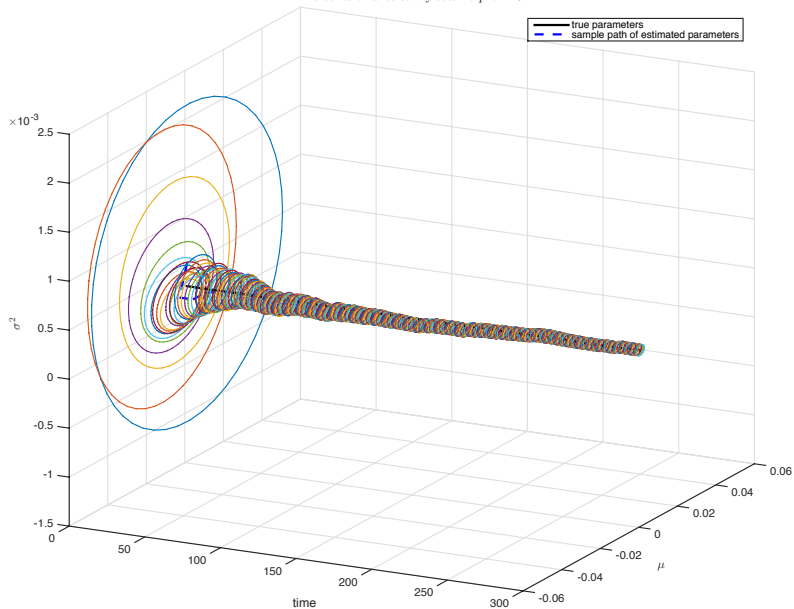
$$W_T(v, \hat{\mu}, \hat{\sigma}^2) = u(v),$$

$$W_t(v, \hat{\mu}, \hat{\sigma}^2) = \sup_{a \in A} \inf_{(\mu, \sigma^2) \in \tau_\alpha(t, \hat{\mu}, \hat{\sigma}^2)} \mathbb{E}_{\mu, \sigma} [W_{t+1}(\mathbf{T}(t, v, \hat{\mu}, \hat{\sigma}^2, a, Z_{t+1}))]$$

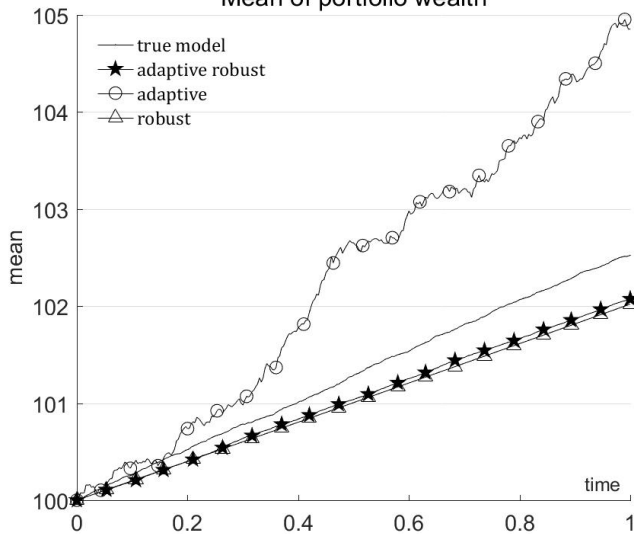
We consider the following setting:

- CRRA utility function  $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$ , with  $\gamma = 40$ ,
- Return of the risk-free asset :  $r = 0$ ,
- Excess return true mean and standard deviation :  $\mu^* = 1\%$ ,  $\sigma^* = 4\%$
- Initial guess on model parameters :  $\widehat{\mu}_0 = 2\%$  and  $\widehat{\sigma}_0 = 1\%$
- Initial endowment :  $V_0 = 100$
- $A = [0, 1]$ , i.e. optimal strategies at time  $t$  to be between 0 and 1.

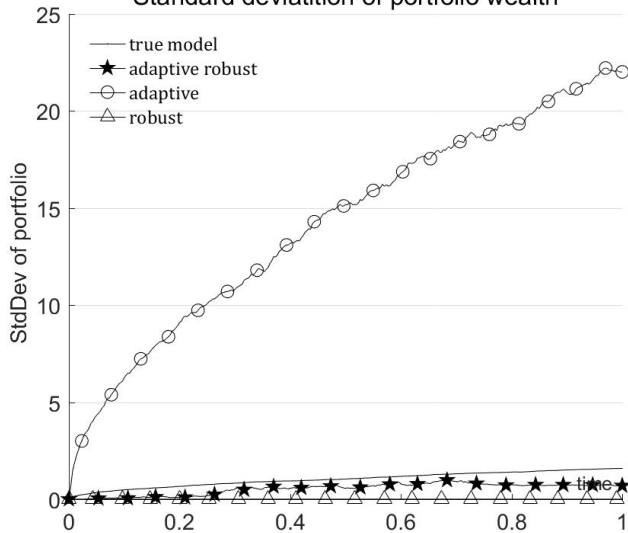
Time-series of uncertainty sets -  $\alpha = 0.1$



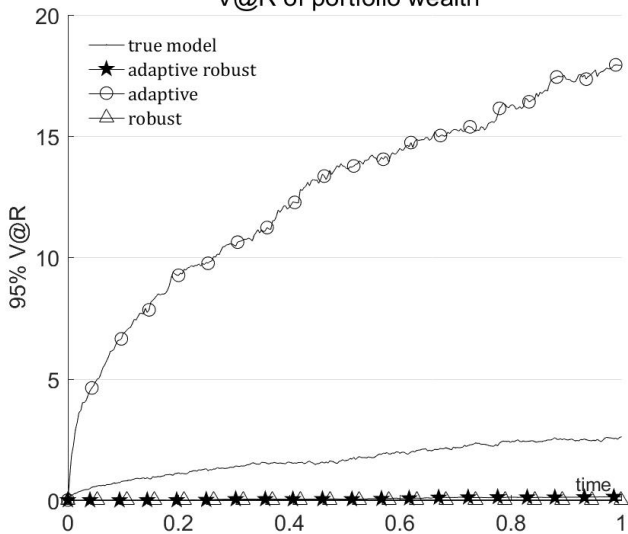
Mean of portfolio wealth

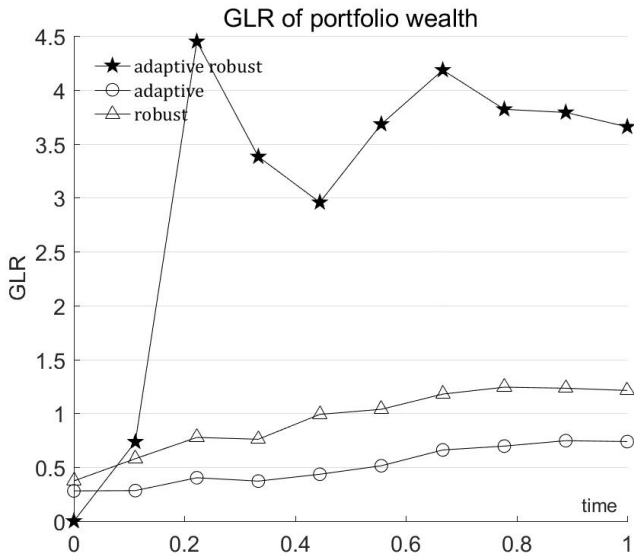


Standard deviation of portfolio wealth



# V@R of portfolio wealth





**Thank You !**

The end of the talk...  
but not of the story.