Nonparametric Distributional Robustness in Multistage Stochastic Optimization

Georg Ch. Pflug/A. Pichler/D. Wozabal/B. Analui/M. Glanzer/M. Pohl

March 7, 2018

Optimal management and scenario models

Traditionally, optimal decision making under uncertainty is done two steps:

- Step 1: Estimation of a probability model for the random scenarios
- Step 2: Finding the best decision given the estimated model

According to Ellsberg (1961) we face here two types of non-determinism:
Uncertainty: the probabilistic model is known, but the realizations of the random variables are unknown ("aleatoric uncertainty")
Ambiguity: the probability model itself is not fully known ("epistemic uncertainty").
Ambiguity sets $\mathcal{P}$: A family of probability models $\mathcal{P}$ which are all plausible models for the reality and we are uncertain about which concrete $P \in \mathcal{P}$ is the true one.

Problem formulation: Ambiguity

Let the basic problem be

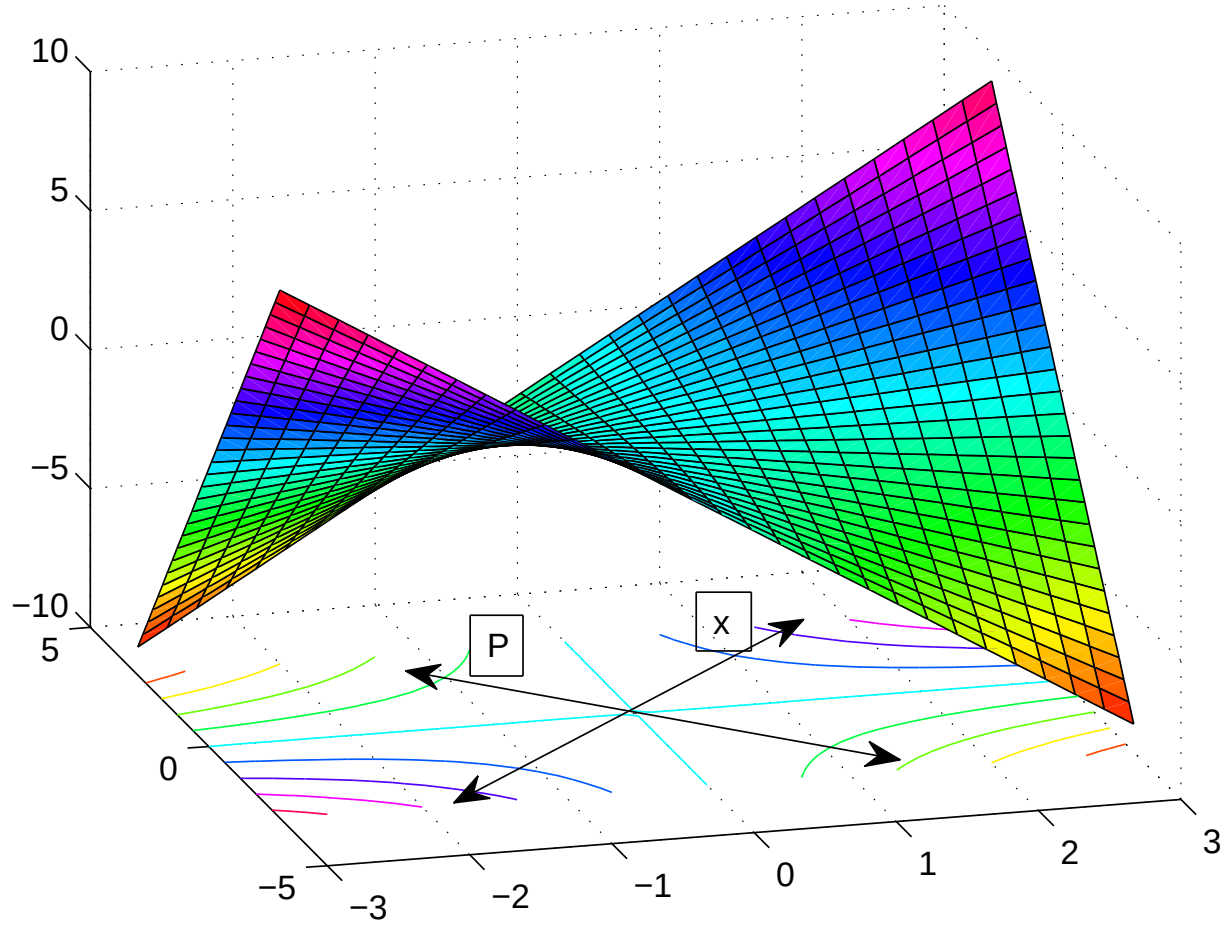
$$\min \{\mathcal{R}_{P_0}[Q(x,\xi)] : x \in \mathbb{X}\}$$

and let $\mathcal{P}$ be the ambiguity set. Then the ambiguity problem is

$$\min \{\max \{\mathcal{R}_P[Q(x,\xi)] : P \in \mathcal{P}\} : x \in \mathbb{X}\}.$$

Find the pair of optimal decision $x^* \in X$ which is good for all models $P \in \mathcal{P}$, among which there is a worst case model $P^* \in \mathcal{P}$.

The pair (x^*, P^*) forms a saddle point



Wasserstein distance

In order to measure the distance of two scenario distributions we use the transportation distance (Kantorovich (1958) distance, Wasserstein (Vasserstein 1969) distance, earth mover distance) between random distributions on $\mathbb{R}^m = (\Omega, d)$ where d is a distance on $\mathbb{R}^m$.
Wasserstein distance of order r:

$$\mathrm{d}_r\left(P_1, P_2; d\right) := \left(\inf_{\pi}\left\{\int_{\Omega\times\Omega} d\left(\omega_1, \omega_2\right)^r \pi\left[\mathrm{d}\omega_1, \mathrm{d}\omega_2\right]\right\}\right)^{\frac{1}{r}},$$

where the infimum is taken over all (bivariate) probability measures π on $\Omega \times \Omega$ which have respective marginals, that is

$$\pi\left[A \times \Omega\right] = P_1\left[A\right] \text{ and } \pi\left[\Omega \times B\right] = P_2\left[B\right]$$

for all measurable sets $A \subseteq \Omega$ and $B \subseteq \Omega$.
We shall call such a measure π a *transportation plan*.

Why Wasserstein balls?

- It is very flexible, since the distance d on $\mathbb{R}^m$ can be chosen, e.g. to take more care about tails in extreme value problems.
- If the supports of P_1 and P_2 are finite and fixed, we have polyhedrality.
- If the scenario values may vary, the Wasserstein ball is more involved, but can be attacked by DC-algorithms (D. Wozabal)
- The Wasserstein distance metricizes the weak topology on uniformly r-integrable sets of probability measures. In particular, $\mathrm{d}_r(P, \hat{P}_n) \to 0$ for the empirical measure $\hat{P}_n$.
- For any Lipschitz L-function f

$$\left| \int f(u)\, dP_1(u) - \int f(u)\, dP_2(u) \right| \leq L \cdot \mathrm{d}_1(P_1, P_2)$$

Implications of closedness in Wasserstein distance

Assume that $X \sim P$ and $\tilde{X} \sim \tilde{P}$. Then

1. $\left|\mathbb{E}|X|^p - \mathbb{E}|\tilde{X}|^p\right| \leq$
   $p \cdot \mathrm{d}_r\left(P, \tilde{P}\right) \cdot \max\left\{\mathbb{E}^{\frac{r-1}{r}}\left[|X|^{r\cdot\frac{p-1}{r-1}}\right], \mathbb{E}^{\frac{r-1}{r}}\left[|\tilde{X}|^{r\cdot\frac{p-1}{r-1}}\right]\right\},$
2. $|\mathbb{E}(X^p) - \mathbb{E}(X^p)| \leq$
   $p \cdot \mathrm{d}_r\left(P, \tilde{P}\right) \cdot \max\left\{\mathbb{E}^{\frac{r-1}{r}}\left[|X|^{r\cdot\frac{p-1}{r-1}}\right], \mathbb{E}^{\frac{r-1}{r}}\left[|\tilde{X}|^{r\cdot\frac{p-1}{r-1}}\right]\right\}$ for p an integer,
3. $\left|\mathbb{E}X^2 - \mathbb{E}\tilde{X}^2\right| \leq 2 \cdot \mathrm{d}_2\left(P, \tilde{P}\right) \cdot \max\left\{\mathbb{E}^{\frac{1}{2}}\left[X^2\right], \mathbb{E}^{\frac{1}{2}}\left[\tilde{X}^2\right]\right\},$
4. $\left|\mathbb{E}|X|^r - \mathbb{E}|\tilde{X}|^r\right| \leq$
   $r \cdot \mathrm{d}_r\left(P, \tilde{P}\right) \cdot \max\left\{\mathbb{E}^{\frac{r-1}{r}}\left[|X|^r\right], \mathbb{E}^{\frac{r-1}{r}}\left[|\tilde{X}|^r\right]\right\}$ and
5. $\left|\mathbb{E}|X|^p - \mathbb{E}|\tilde{X}|^p\right| \leq$
   $p \cdot \mathrm{d}_2\left(P, \tilde{P}\right) \cdot \max\left\{\mathbb{E}^{\frac{1}{2}}\left[|X|^{2(p-1)}\right], \mathbb{E}^{\frac{1}{2}}\left[|\tilde{X}|^{2(p-1)}\right]\right\},$

where $p \geq 1$ and $r > 1$.

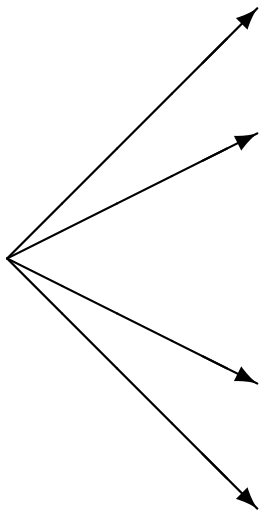
Statistical estimation and confidence sets

- For compatibility with the convergence of the empirical distribution, the distance should metricize the weak (weak*) convergence.
- It should allow to construct statistical confidence sets, i.e. random sets, which contains the true distribution with prescribed level of confidence.

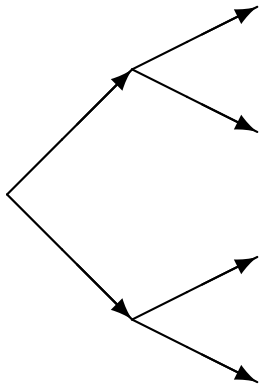
One typical Theorem:
Theorem (Boley, Guilin, Villani). Let $d(x,y) = \|x-y\|$. Suppose that $\int \exp(\alpha\|x\|)\, P(dx) < \infty$. Then for $m' < m$, there exist constants k and N_0 such that for $\epsilon > 0$ and $n \geq N_0 \max(\epsilon^{-(2r+m')}, 1)$

$$P\{\mathrm{d}_r(\hat{P}_n, P) \geq \epsilon\} \leq \exp(-Kn^{1/r} \min(\epsilon, \epsilon^2)).$$

Single-,two- and multistage



Single- or twostage



Multistage

Single- or twostage Stochastic optimization problems involve a probability distribution for the scenarios. Multistage problems involve a stochastic process for which the conditional distributions are relevant, not just the joint distribution. It is crucial to encode the *information structure into the scenario process in order to model non-anticipativity.*

Nested distance for multistage situations

A generalization of the Wasserstein distance for multistage situations: Let $\mathbb{P} = (\Omega, (\mathcal{F}_t)_{0\le t\le T}, P)$ and $\tilde{\mathbb{P}} = (\tilde{\Omega}, (\tilde{\mathcal{F}}_t)_{0\le t\le T}, \tilde{P})$ two filtered probability spaces. The nested distance of order $r \ge 1$ between $\mathbb{P}$ and $\tilde{\mathbb{P}}$, denoted by $\mathrm{dl}(\mathbb{P}, \tilde{\mathbb{P}})$, is defined as the optimal value of the optimization problem

$$\inf_{\pi} \left(\iint_{\mathcal{F}_T \otimes \tilde{\mathcal{F}}_T} d(\omega, \tilde{\omega})^r \pi(d\omega, d\tilde{\omega}) \right)^{1/r}$$

$$\text{s.t. } \pi\left(A \times \tilde{\Omega} | \mathcal{F}_t \otimes \tilde{\mathcal{F}}_t\right) = \mathbb{P}\left[A \middle| \mathcal{F}_t\right] \quad A \in \mathcal{F}_T;\ t = 1, \dots, T$$

$$\pi\left(\Omega \times B | \mathcal{F}_t \otimes \tilde{\mathcal{F}}_t\right) = \mathbb{P}\left[B | \tilde{\mathcal{F}}_t\right] \quad B \in \tilde{\mathcal{F}}_T;\ t = 1, \dots, T.$$

We define model neigborhoods as balls in nested distance around a baseline model P_0:

$$\mathcal{P} = \{\mathbb{P} :\ \mathrm{dl}(\mathbb{P}, \mathbb{P}_0) \le \epsilon\}$$

The main Lipschitz property

Theorem. (A. Pichler, G.P. 2009)
Let $\mathbb{P} := \left(\Omega, (\mathcal{F}_t)_{t=0,\dots T}, P\right)$, ($\tilde{\mathbb{P}} := \left(\tilde{\Omega}, (\tilde{\mathcal{F}}_t)_{t=0,\dots T}, \tilde{P}\right)$, resp.) be filtered probability spaces. Consider the multistage stochastic optimization problem

$$v(\mathbb{P}) := \inf \left\{ \mathbb{E}_P Q(\xi, x) : \; x \lhd \mathcal{F} \right\},$$

where Q is convex in x for any ξ fixed, and Lipschitz with constant L in ξ for any x fixed. Then

$$\left| v(\mathbb{P}) - v(\tilde{\mathbb{P}}) \right| \leq L \cdot \mathrm{d}_r(\mathbb{P}, \tilde{\mathbb{P}})$$

for every $r \geq 1$.
The constraint $x \lhd \mathcal{F}$ expresses the fact that the stochastic decision process x must be adapted to the filtration $\mathcal{F}$, i.e.must be *nonanticipative*.

Confidence sets for the nested distance

If $\hat{\mathbb{P}}_n$ is the empirical distribution of a set of n paths $(\xi_1^{(i)}, \dots, \xi_T^{(i)})_{i=1,\dots,n}$. Then $\hat{\mathbb{P}}_n$ does NOT converge in nested distance to the true distribution.

Theorem. Let $\hat{\mathbb{P}}_n * k_{h_n}$ be a smoothed version of the empirical process, where the conditional distributions are smoothed with the kernel function k and a bandwidth h_n. Suppose that

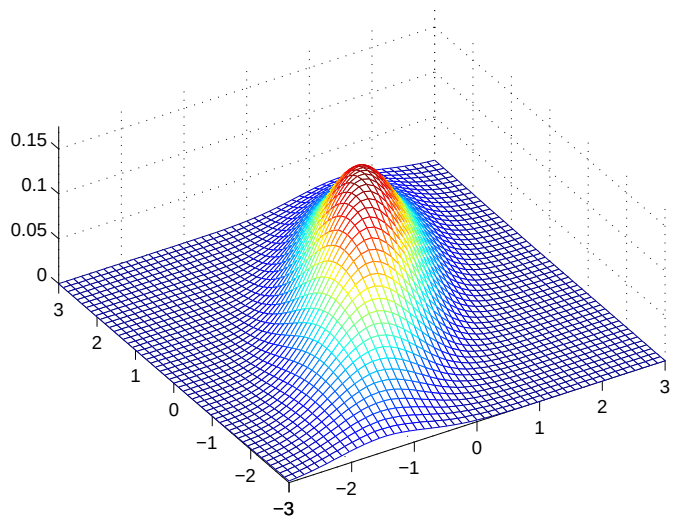
- the stochastic process takes its values in $\mathbb{R}^m$ and the bandwidth constants satisfy

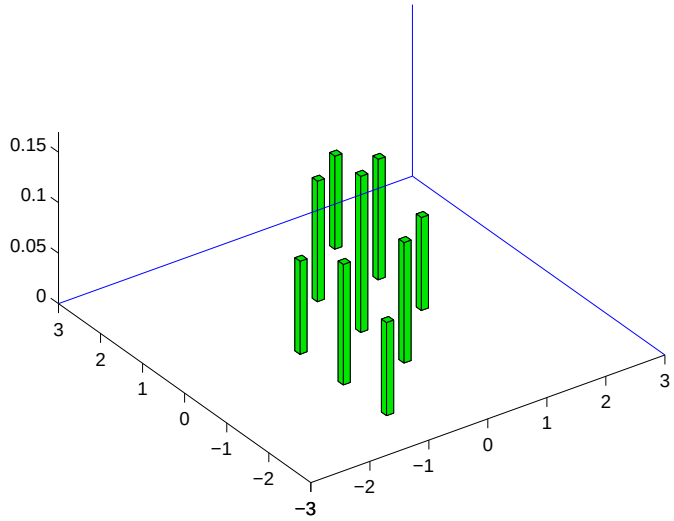
$$h_n \to 0, \ \frac{n h_n^m}{|\log h_n|} \to \infty, \ \frac{|\log h_n|}{\log\log n} \to \infty, \ \text{and } n h_n^m \to \infty,$$

- the conditional densities have a compact and convex support and are bounded from above and from below.

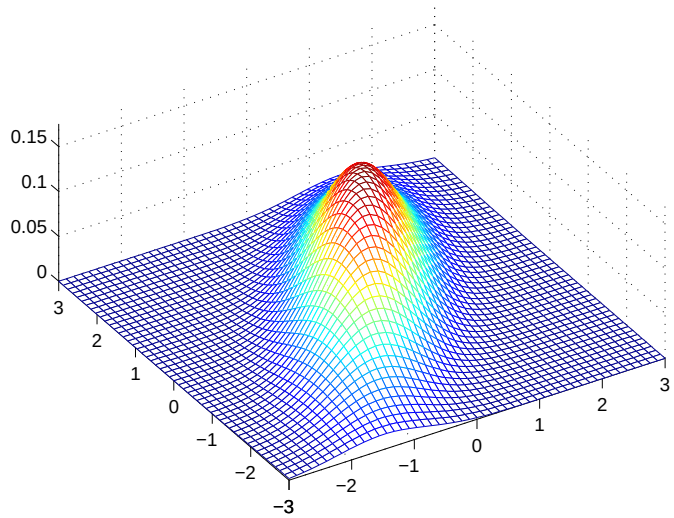
Then the nested distance between the smoothed empirical distribution $\mathbb{P}_n^k := \hat{\mathbb{P}}_n * k_{h_n}$ and the true model $\mathbb{P}$ satisfies

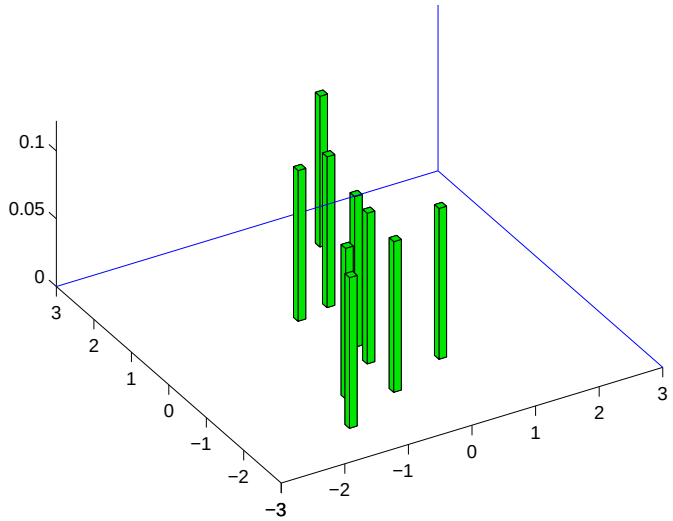
$$\mathbb{P}\left(\mathbf{dl}\left(\mathbb{P}, \mathbb{P}_n^k\right) > C \cdot \varepsilon\right) \leq T \cdot \varepsilon$$





The nested distance is $d(\mathbb{P}, \tilde{\mathbb{P}}) = 0.82$.
The distance of the multiperiod distributions is $d(P, \tilde{P}) = 0.68$.
This approximation is calculated as a stagewise optimal transportation problem (a stagewise optimal facility location problem), not by Monte Carlo. Consequently, all scenarios get different probability values, not just $1/n$ as by crude MC.





The nested distance is $d(\mathbb{P}, \tilde{\mathbb{P}}) = 1.12$.
The distance of the multiperiod distributions is $d(P, \tilde{P}) = 0.67$.

Examples for robust decisions

- Pricing of contingent claims
- Portfolio optimization (single-stage)
- Management of hydro reservoirs

Pricing of contingent claims in discrete time

Consider a $d+1$-dimensional price process S for the underlying defined on a filtered probability space $(\Omega, \mathcal{F})$

$$S_t = (S_t^{(0)}, S_t^{(1)}, \ldots, S_t^{(d)}),\ t = 1, \ldots, T.$$

A contingent claim C consists of a sequence of cash flows $C_1, \ldots, C_T$ which is adapted to the filtration generated by the underlying.
The consistent no-arbitrage ask-price of the contingent claim by a pointwise replication strategy is given by solution of the multistage optimization problem

$$\begin{aligned} \pi_{a:} := \min_{x,w}\ & w \\ \text{s.t. } & x_0^\top S_0 \le w \\ & x_{t-1}^\top S_t - x_t^\top S_t - C_t \ge 0 \quad \forall t = 1, \cdots, T-1; \\ & x_{T-1}^\top S_T - C_T \ge 0. \end{aligned}$$

Acceptability functionals

An acceptability functional is a mapping
$\mathcal{A} : L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R} \cup \pm\infty$ which is: concave, monotonic, translation equivariant and positively homogeneous.
By duality, such a functional can be represented as

$$\mathcal{A}(Y) = \inf_{Z \in \mathcal{Z}_{\mathcal{A}}} \{\mathbb{E}[YZ]\},$$

where $\mathcal{Z}_{\mathcal{A}}$ is the *supergradient set of densities.*
The (upper) Average Value-at-Risk has supergradient set is
$\mathcal{Z} = \{Z : 0 \leq Z \leq 1/\alpha; \mathbb{E}(Z) = 1\}$.
Special cases of the AV@R are

- $AV@R_0(Y) = essinf(Y)$, the essental infimum. Acceptability w.r.t. essinf means almost sure (super-) hedging, the "classical" hedging condition.
- $AV@R_1(Y) = \mathbb{E}(Y)$, the expectation. Acceptability w.r.t. the expectation is the weakest form of acceptability: expectation hedging.

The acceptable ask price

The acceptable ask-price is given as the optimal solution of the optimization problem S

$$\begin{aligned}\pi_a(\mathcal{A}_1,\dots,\mathcal{A}_T) := \min_{x,w}\ & w \\ \text{s.t. } & x_0^\top S_0 \le w \\ & \mathcal{A}_t(x_t^\top S_t - x_{t-1}^\top S_t - C_t) \ge 0 \quad \forall t = 1,\cdots,T-1; \\ & \mathcal{A}_T(x_{T-1}^\top S_T - C_T) \ge 0.\end{aligned}$$

A similar problem determines the acceptable bid-price.

Acceptability under model ambiguity

The robust acceptable ask-price is defined as the optimal solution of the optimization problem

$$\min_{x_t, w} \quad w$$

s.t.

$$x_0^\top S_0 \leq w$$

$$\mathcal{A}_t^{\mathbb{P}}(x_{t-1}^\top S_t - x_t^\top S_t - C_t) \geq 0 \quad \forall \mathbb{P} \in \mathcal{P};\ \forall t = 1, \ldots, T-1$$

$$\mathcal{A}_T^{\mathbb{P}}(x_{T-1}^\top S_T - C_T) \geq 0 \quad \forall \mathbb{P} \in \mathcal{P}$$

The optimal bid price is defined in an analogous way.

Dualization

Let $\mathcal{P}$ be a convex set of probability models, which is spanned by a sequence of models $\mathbb{P}_1, \mathbb{P}_2, \ldots$, that are all absolutely continuous w.r.t. some baseline model $\hat{\mathbb{P}}$. Moreover, for all $t = 1, \ldots, T$, let $\mathcal{A}_t$ be acceptability functionals with corresponding supergradient sets $\mathcal{Z}_{\mathcal{A}_t}$. Then, strong duality holds between the given ambiguity problem and its dual given by

$$\pi_a^{\mathcal{P}}(\mathcal{A}_1, \ldots, \mathcal{A}_T) = \sup_{\mathbb{Q}} \ \mathbb{E}^{\mathbb{Q}} \left[\sum_{t=1}^{T} \tilde{C}_t \right]$$

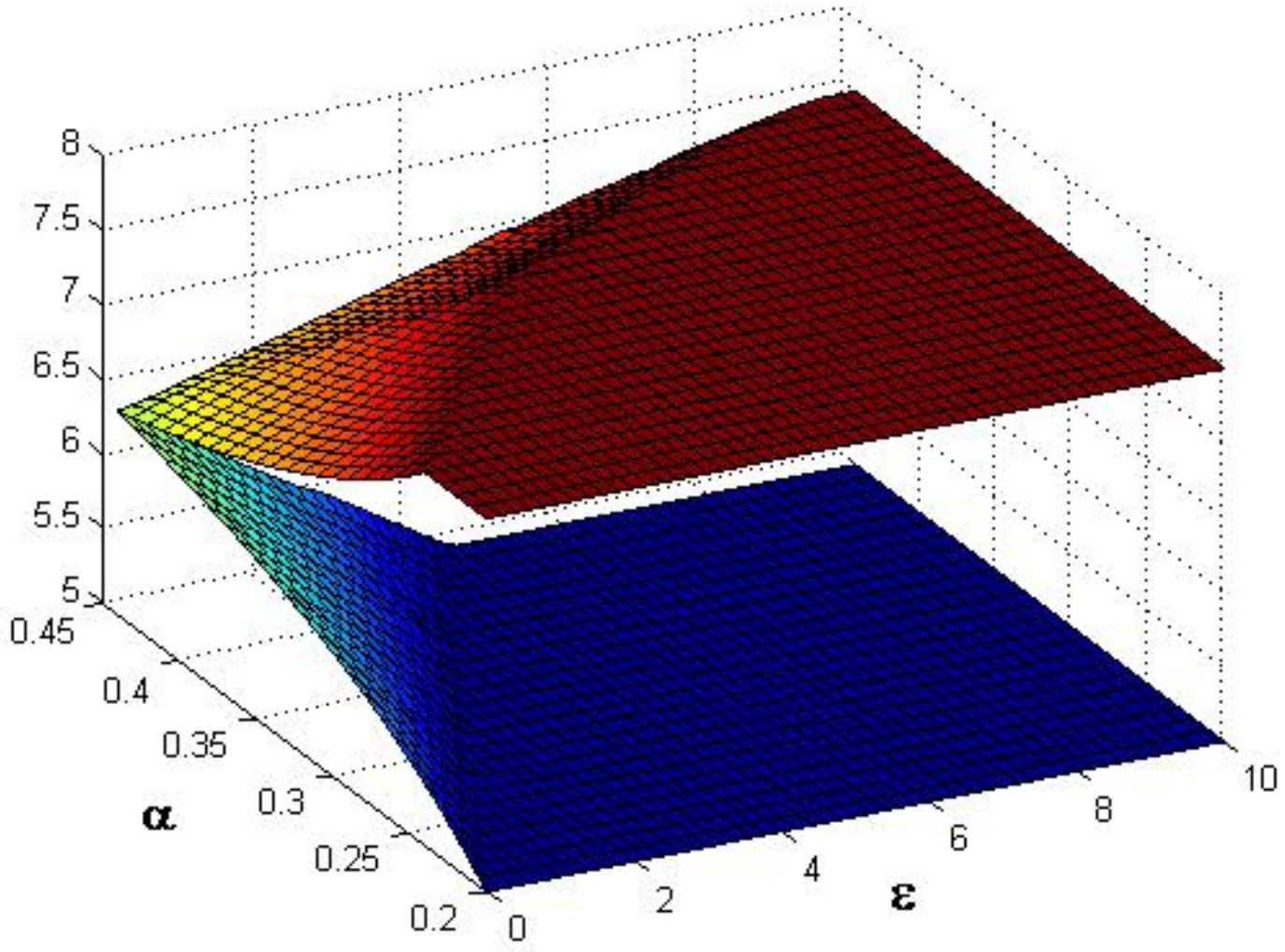
$$\text{s.t. } \mathbb{E}^{\mathbb{Q}} \left[\tilde{S}_{t+1} \middle| \mathcal{F}_t \right] = \tilde{S}_t \quad \forall t = 0, \ldots, T-1$$

$$\forall t = 1, \ldots, T \ \exists \ \mathbb{P} \in \mathcal{P} : \ \left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} \in \mathcal{Z}_{\mathcal{A}_t^{\mathbb{P}}}.$$

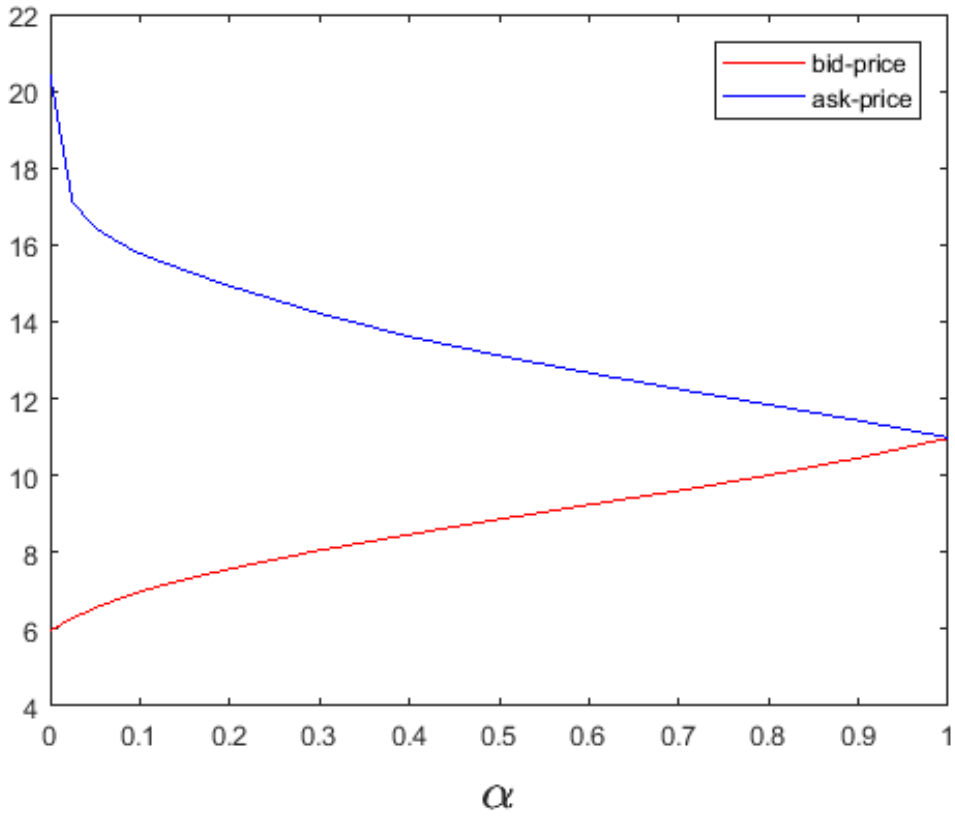
Observations

- Weakening the acceptance level decreases the ask-price and increases the bid-price
- Increasing the ambiguity radius increases the ask-price and decreases the bid-price
- The bid-ask spread decreases with weakening the acceptance level and increases with the ambiguity level

An illustration for the bid-ask spread



Challenging the Black-Scholes formula



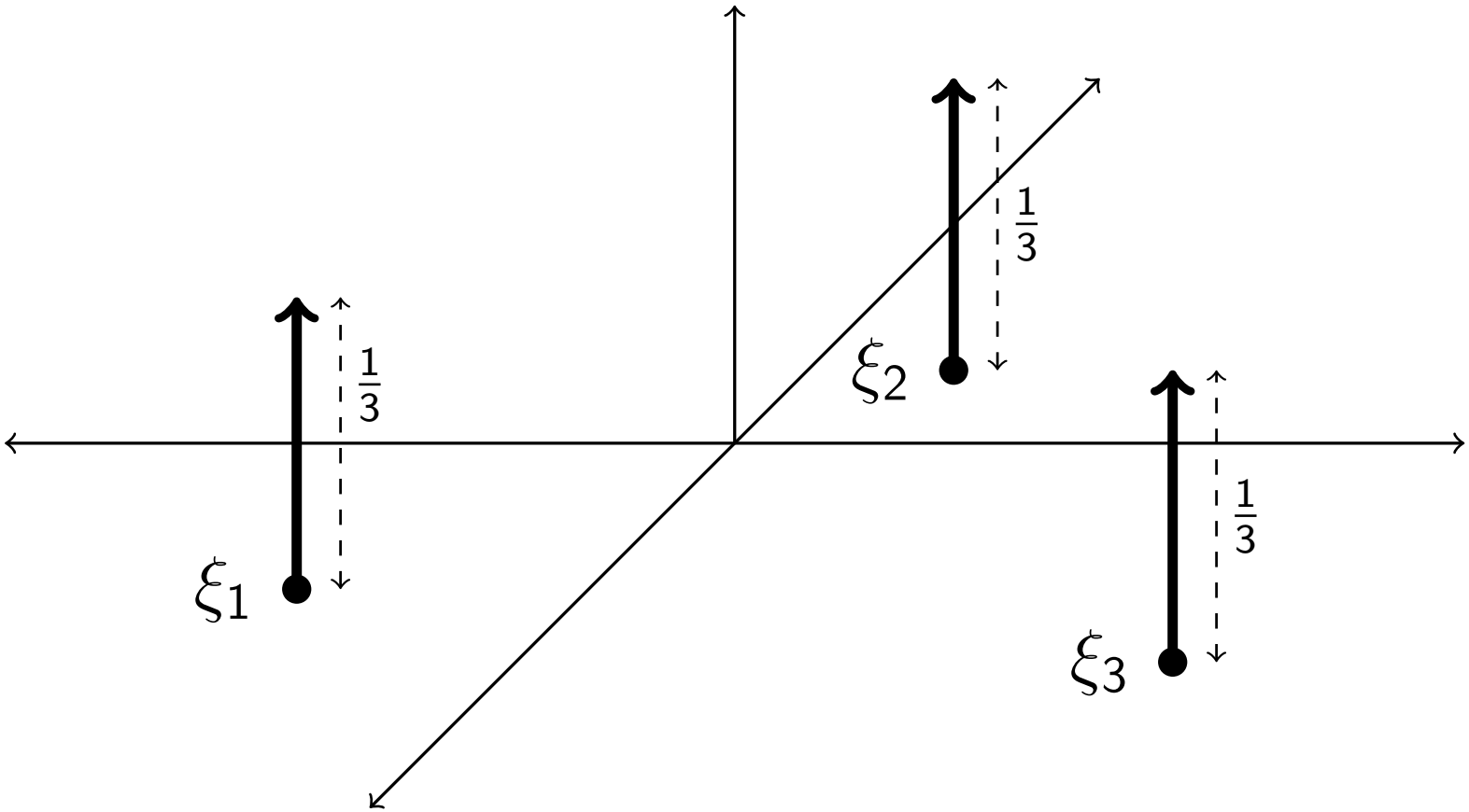
Call option: strike $K = 95$, $S_0 = 100$, T=1yr, vola 20%, $r = 1\%$
Continuous time BS price 11.0602
Discrete price on a bushy tree , which carries lots of martingale measures: 11.0593

Portfolio selection with AV@R risk under model ambiguity

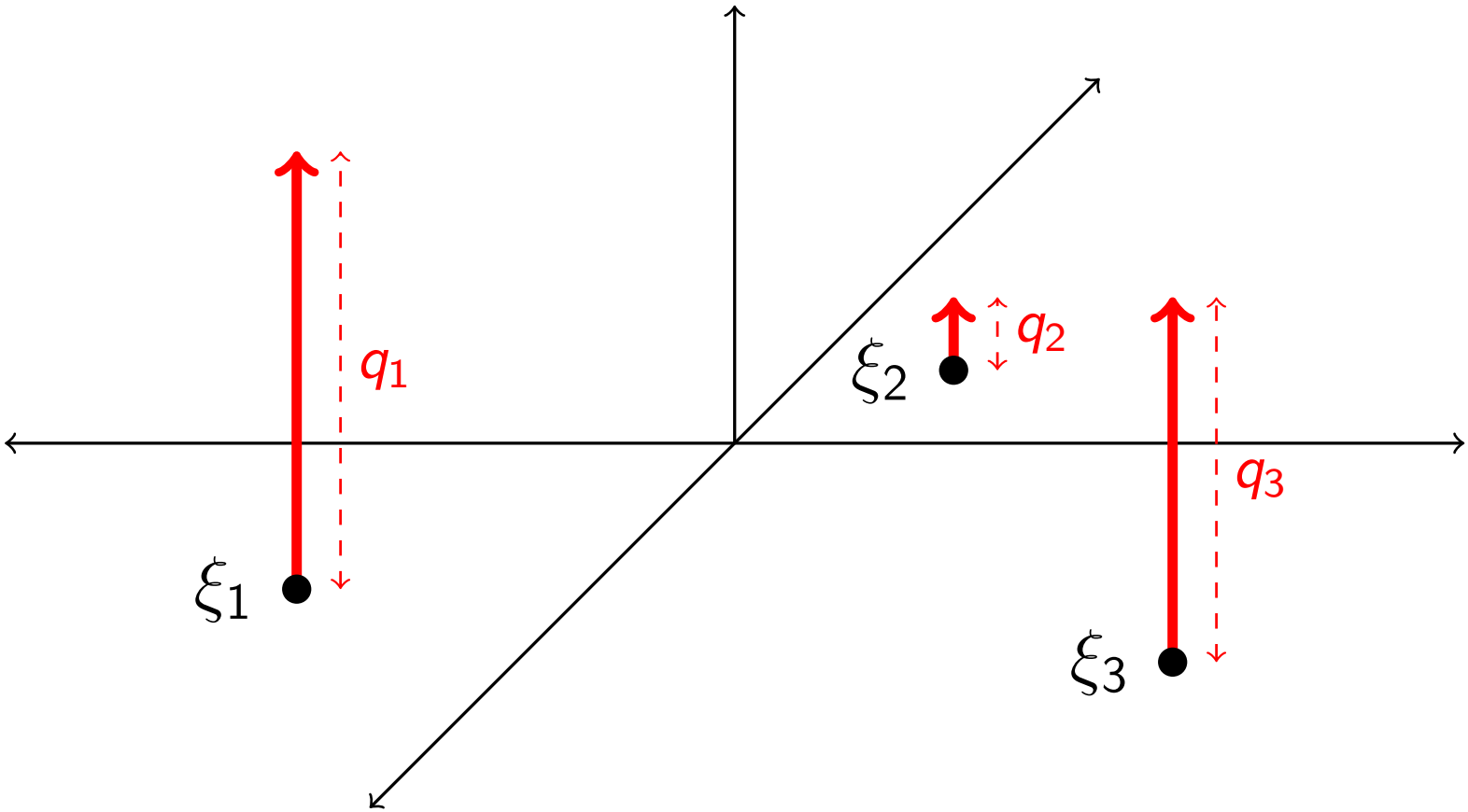
$$\max_{x\in\mathbb{X}} \min_{P\in\mathcal{P}} \mathbb{E}\left(x^\top \xi^P\right) - \lambda \mathsf{AV@R}_\alpha\left(-x^\top \xi^P\right),$$

$\mathcal{P} := \{P : \mathrm{d}_1(P, P_0) \le \epsilon\}$,
P_0... reference/baseline distribution,
ϵ... level of model ambiguity,
$\mathrm{d}_1(\cdot,\cdot)$... Wasserstein distance.

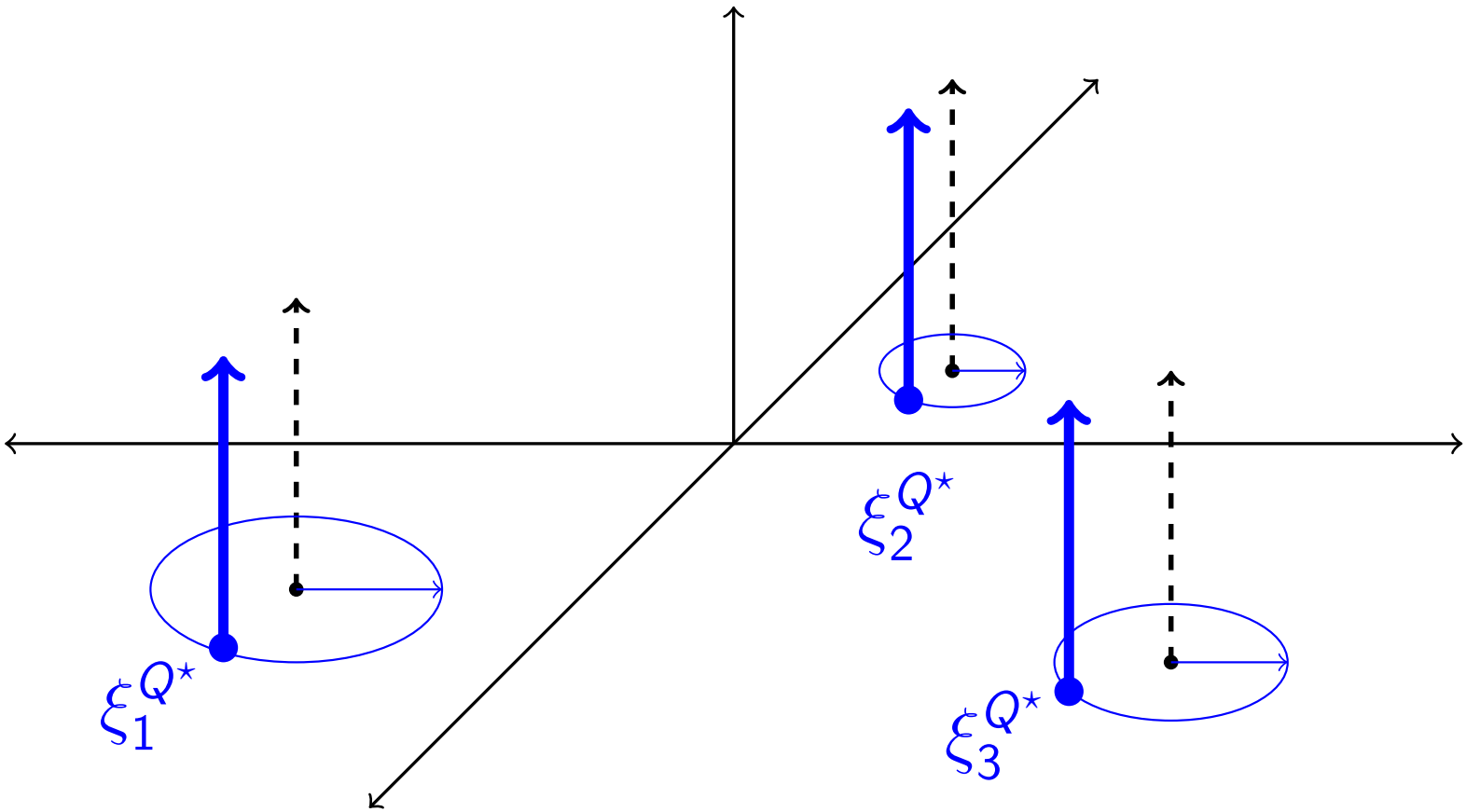
Consider the empirical distribution $P_0 = \hat{P}_n$



Pflug and Wozabal (2007)



Esfahani and Kuhn (2015)



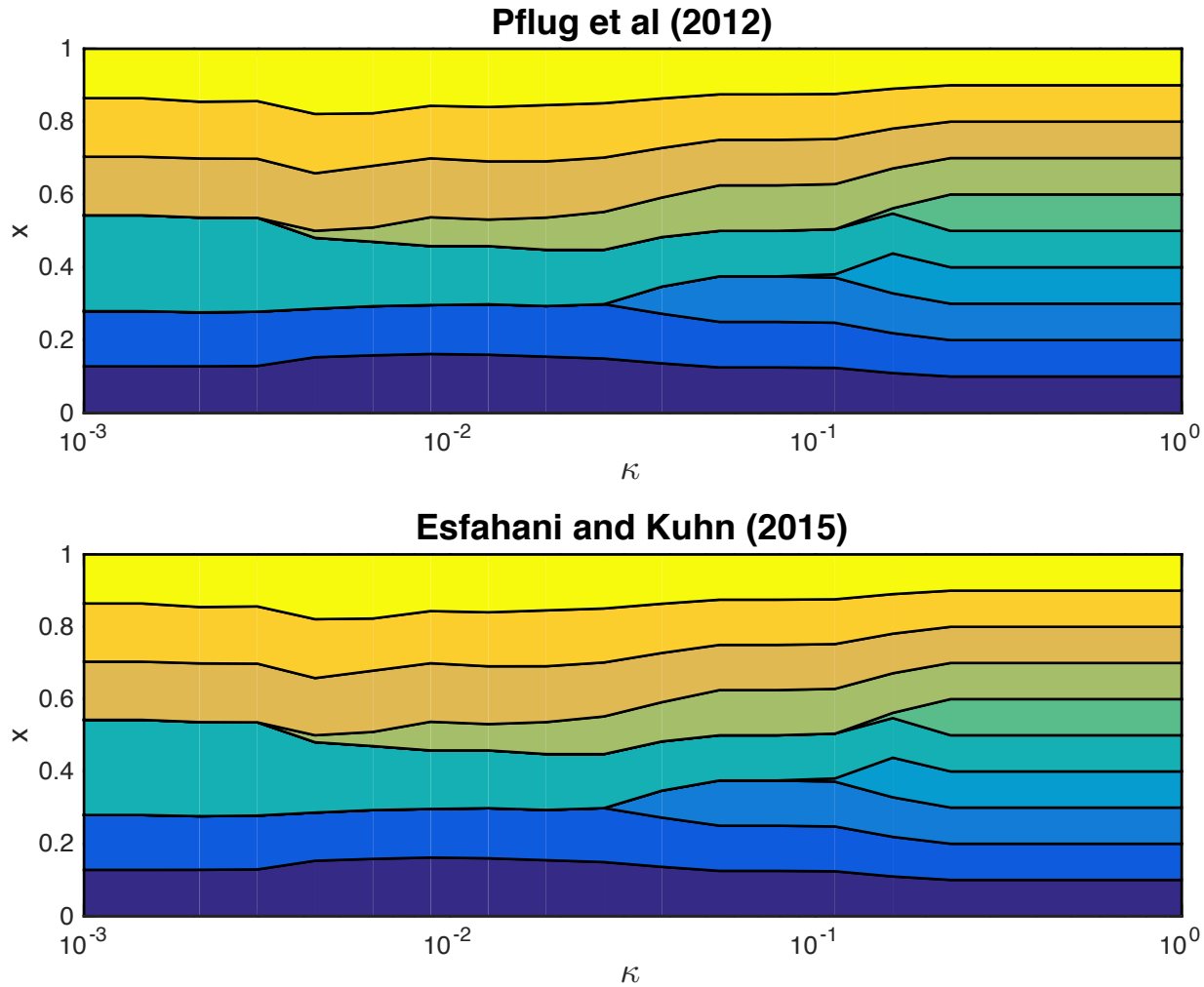


Figure: Optimal portfolio composition as a function of the level of model ambiguity κ.

Equal weights is maximin for large ambiguity

With this insight, we may prove a remarkable result for distortion functionals:

$$\lim_{K\to\infty} \operatorname{argmax}_{\{\sum x_i=1, x_i\geq 0\}} \min_{\mathrm{d}_r(P,P_0)\leq K} \mathcal{A}_P(x^\top \xi) = \frac{1}{M}\mathbb{1}.$$

Under large ambiguity, the optimal decision is the "equal weights" allocation.
The same result holds for the Markovitz model, if the distance is d_2.
Distortion utility functional: $\mathcal{A}(Y) = \int_0^1 F_Y^{-1}(p)h(p)\,dp$
Average value-at-risk: $\mathbb{A}\mathrm{V@R}(Y) = \frac{1}{\alpha}\int_0^\alpha F_Y^{-1}(p)\,dp$

Ambiguity only in dependency

ξ^C is the distribution of the return vector, where C is the copula while the marginals of ξ_i are fixed. Let C_0 be the baseline copula.

$$\max_{x\in\mathbb{X}} \min_{\mathrm{d}_1(C,C_0)\leq\epsilon} \mathbb{E}\left(-x^\top \xi^C\right) - \lambda\mathcal{R}\left(x^\top \xi^C\right)$$

Total dependency ambiguity: Portfolio Concentration

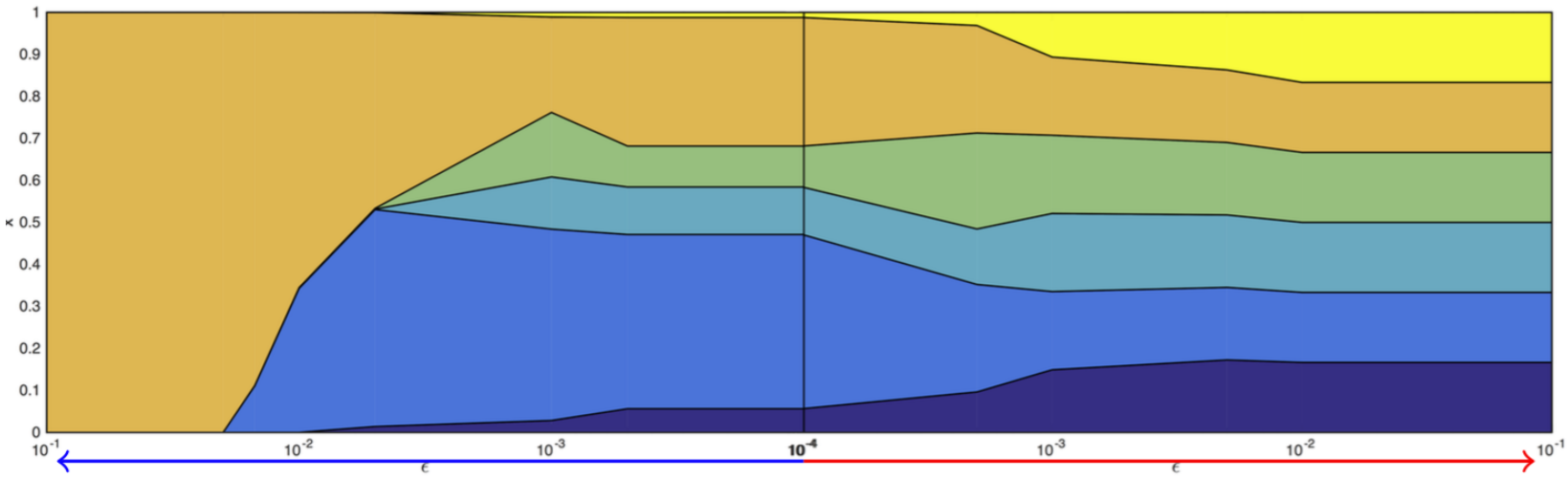
Let $\mathcal{C}$ be the family of all copulas.
If $\mathcal{R}$ is subadditive, comonotone additive and positive homogeneous, then

$$\max_{x\in\mathbb{X}} \min_{C\in\mathcal{C}} \mathbb{E}\left(x^\top \xi^C\right) - \lambda\mathcal{R}\left(x^\top \xi^C\right)$$
$$= \max_{i\in\{1,\dots,m\}} \mathbb{E}[\xi_i] - \lambda\mathcal{R}(\xi_i).$$

Thus the maximin portfolio is to invest everything in just one the asset i^*, where

$$i^* = \operatorname{argmax}_{i\in\{1,\dots,m\}} \mathbb{E}[\xi_i] - \lambda\mathcal{R}(\xi_i).$$

Concentration vs Diversification



Ambiguity in the dependence structure Ambiguity in the joint distribution

Data: 6 Indices: S&P 500, TOPIX, FTSE China B35, EURO STOXX 50, FTSE 100 and NIFTY 500; observations Jan 1 - Dec 13, 2016

Price of Ambiguity and Reward for Robustness

Let $\hat{\mathbb{P}}$ be the baseline model and let $x^*(\mathbb{P})$ be the optimal solution of the baseline problem. Likewise, let $\mathcal{P}$ be the ambiguity set and let $x^*(\mathcal{P})$ be the solution of the minimax problem. Under convex-concavity, the solution $x^*(\mathcal{P})$ of the minimax problem together with the worst case model $\mathbb{P}^*$ form a saddle point, meaning that the following inequality is valid for all feasible x and all $\mathbb{P} \in \mathcal{P}$

$$\mathbb{E}_{\mathbb{P}}[Q(x^*(\mathcal{P}), \xi)] \leq \mathbb{E}_{\mathbb{P}^*}[Q(x^*(\mathcal{P}), \xi)] \leq \mathbb{E}_{\mathbb{P}^*}[Q(x, \xi)].$$

Let us call $\mathbb{E}_{\mathbb{P}^*}[Q(x^*(\mathcal{P}), \xi)]$ the minimax value.

Define:

- The Price of Ambiguity.

$$\mathbb{E}_{\hat{\mathbb{P}}}[Q(x^*(\mathcal{P}), \xi)] - \mathbb{E}_{\hat{\mathbb{P}}}[Q(x^*(\hat{\mathbb{P}}), \xi)] \geq 0.$$

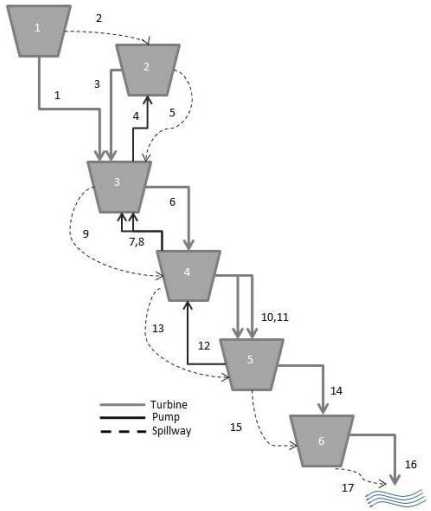
"How much do I loose by implementing the minimax strategy $x^*(\mathcal{P})$ instead of the best strategy for the baseline model, if in fact the baseline model is true?"

- Reward for robust decisions.

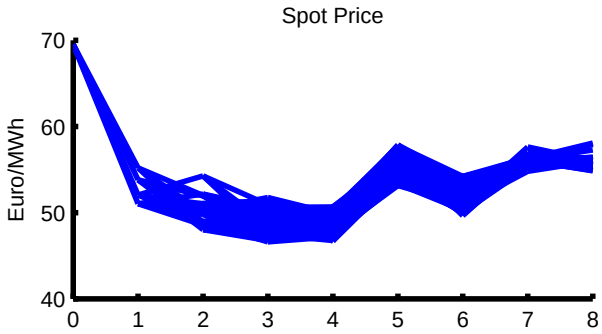
$$\mathbb{E}_{\mathbb{P}^*}[Q(x^*(\mathbb{P}), \xi)] - \mathbb{E}_{\mathbb{P}^*}[Q(x^*(\mathcal{P}), \xi)] \geq 0.$$

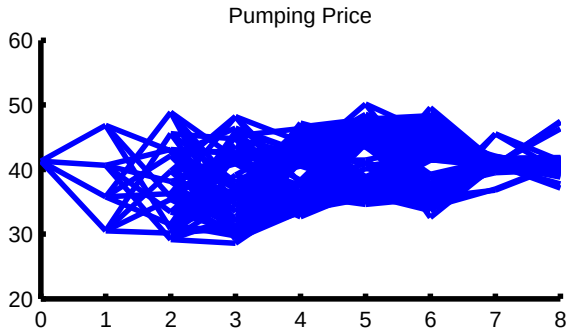
"How much do I gain, when I implement the minimax strategy $x^*(\mathcal{P})$ instead of the best strategy for the baseline model, if in fact the worst case model is true?"

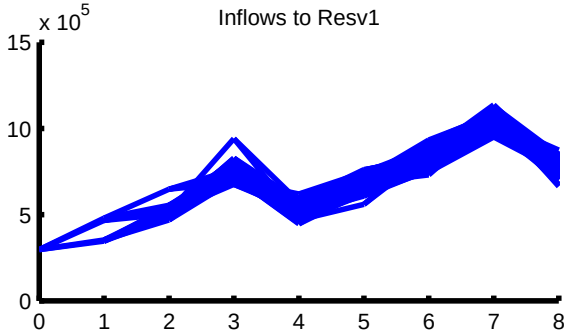
Management of a hydrosystem in the Austrian Alps



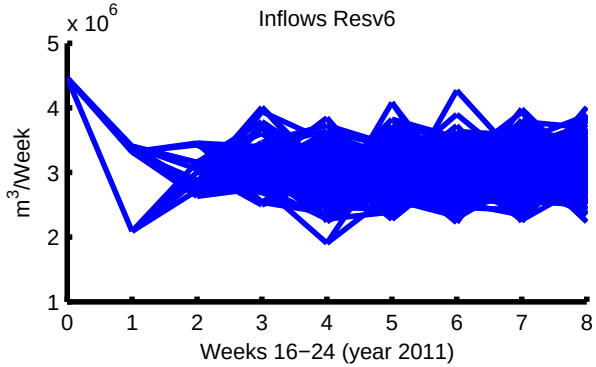
The scenario process is 5-dimensional: Spot prices, Pumping prices, Inflows for 3 reservoirs.
The decision process is 2-dimensional: Turbining and Pumping
The number of decision stages is 8 (weeks).

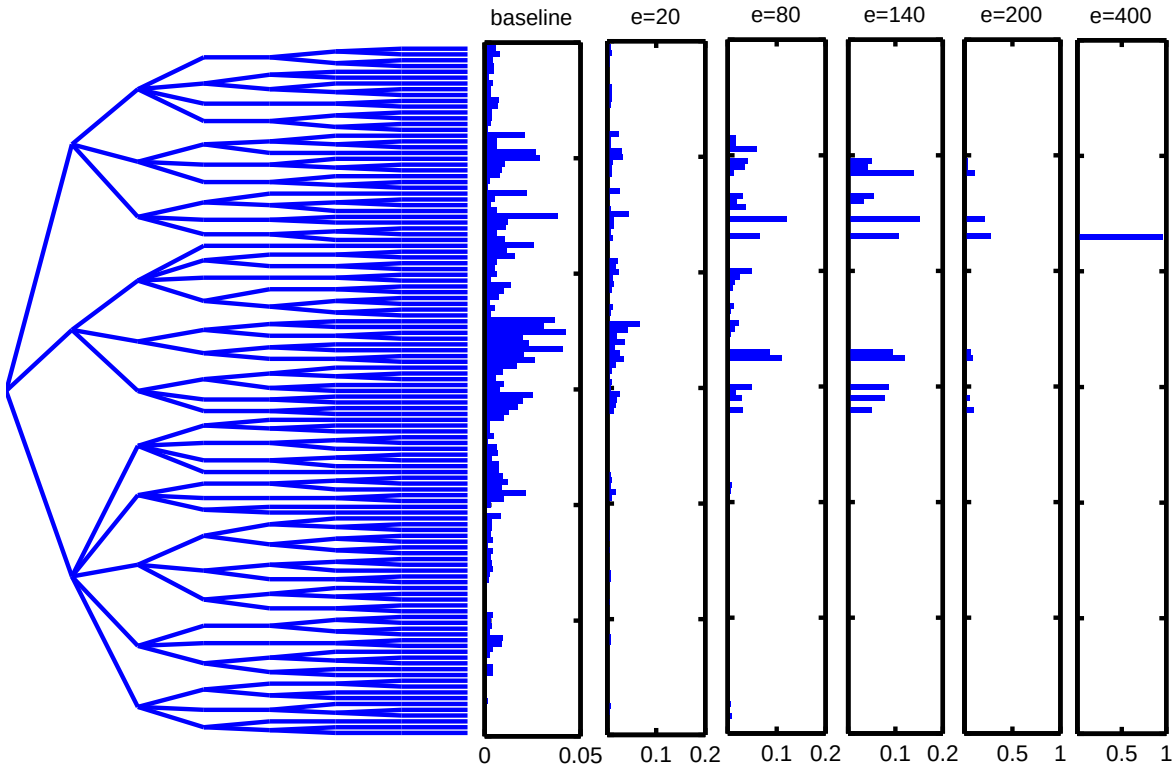

Spot Price
Euro/MWh
70
60
50
40
0 1 2 3 4 5 6 7 8


Pumping Price
60
50
40
30
20
0 1 2 3 4 5 6 7 8

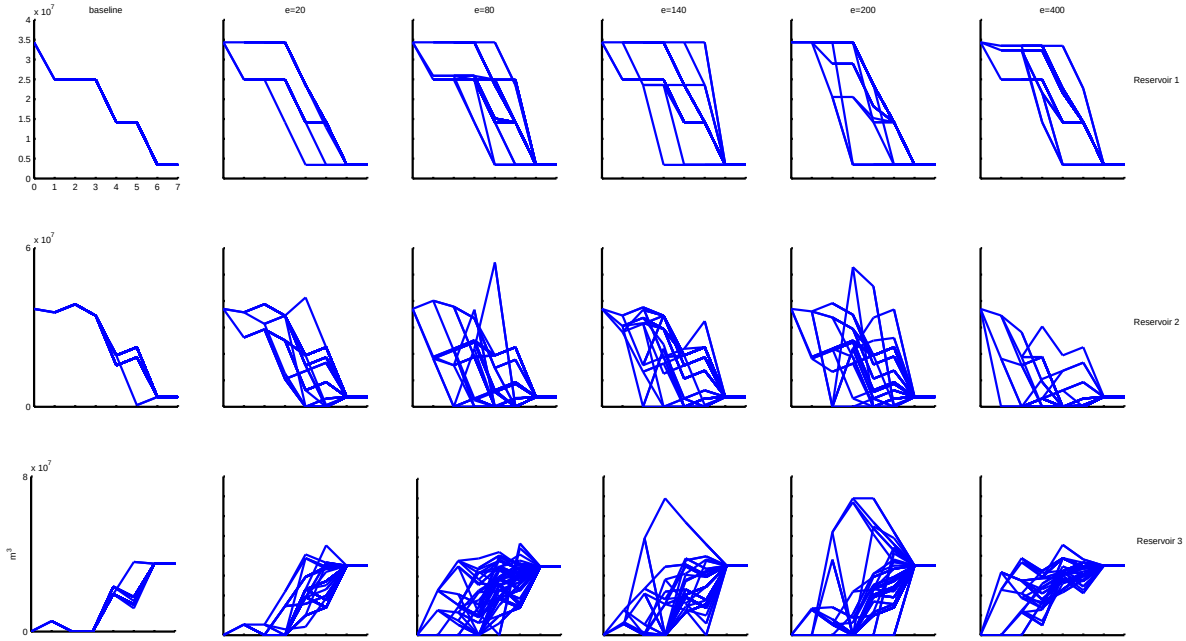

x 10^5
Inflows to Resv1
15
10
5
0
0 1 2 3 4 5 6 7 8


x 10^5
Inflows to Resv2
5
4
3
2
0 1 2 3 4 5 6 7 8


x 10^6
Inflows Resv6
m^3/Week
5
4
3
2
1
0 1 2 3 4 5 6 7 8
Weeks 16−24 (year 2011)



The typical picture: The larger is the ambiguity radius, the simpler



The minimax decisions: They get more complicated with increasing ambiguity radius: Decisions lying on bounds are avoided.
Price of ambiguity: 2.3%.
Reward for robustness: 7.5%.

Conclusions

- In order to capture scenario uncertainty (aleatoric uncertainty) and probability ambiguity (epistemic uncertainty) we use a maximin approach.
- The ambiguity neighborhood are chosen in such a way that they form statistical confidence regions for which bounds for the covering probability are available.
- For single stage problems we use the Kantorovich-Wasserstein distance, for multistage problems we use the nested distance to quantify the epistemic uncertainty.
- If the ambiguity radius is increased, then the saddle point changes typically in the following way:
  - The robust decision strategy becomes more complicated and "diversified".
  - The worst case model gets more simpler.
- Often the price for ambiguity is smaller than the reward for robustness.

References

- M. Glanzer, Pflug, G. Incorporating statistical model error into the calculation of acceptability prices of contingent claims. Manuscript, submitted.
- Pflug, G., Pohl, M. (2017). A review on ambiguity in stochastic portfolio optimization. Set-Valued and Variational Analysis.
- Pflug, G., Pichler, A. (2016). From empirical observations to models for Stochastic Optimization: Convergence properties. SIAM Journal on Optimization, 26(3), 1715-1740.
- Pflug, G., Pichler, A. (2014). Multistage Stochastic Optimization. (Springer Series in Operations Research and Financial Engineering). Springer.
- Analui, B., Pflug, G. (2014). On Distributionally Robust Multiperiod Stochastic Optimization. Computational Management Science, 11, 197-220. DOI: 10.1007/s10287-014-0213-y
- Pflug, G., Pichler, A., Wozabal, D. (2012). The 1/N investment strategy is optimal under high model ambiguity. Journal of Banking and Finance, 36(2), 410-417. DOI:
- Wozabal, D., Pflug, G. (2007). Ambiguity in portfolio selection. Quantitative Finance, 7(4), 435-442.