Representation of Distributionally Robust Chance-constraints

Jean B. Lasserre, Tillmann Weisser

Presenter: Chen Li

Nov, 2019

Outline

Chance-constraints

- Decision $\boldsymbol{x} \in \boldsymbol{X} \subset \mathbb{R}^n$
- Uncertainty $\boldsymbol{\omega} \in \boldsymbol{\Omega} \subset \mathbb{R}^p$ with probability measure $\mathbb{P}$
- Constraint function $f : \boldsymbol{X} \times \boldsymbol{\Omega} \to \mathbb{R}$
- Chance-constraints

$$\mathbb{P}\left[f(\boldsymbol{x}, \tilde{\boldsymbol{\omega}}) > 0\right] > 1 - \epsilon \tag{1}$$

- Distributionally robust chance-constraints

$$\mathbb{P}\left[f(\boldsymbol{x}, \tilde{\boldsymbol{\omega}}) > 0\right] > 1 - \epsilon, \quad \forall \mathbb{P} \in \mathcal{F} \tag{2}$$

An Example: Newsvendor Problem

- Decision: inventory level x
- Uncertain demand: $\tilde{d} \sim \mathbb{P}$
- Parameters: price p, unit inventory cost c, profit target t
- Objective: trying to achieve the profit target with high probability

$$\mathbb{P}\left[p\min\{x,\tilde{d}\} - cx > t\right] \geq 1 - \epsilon$$

- If $\mathbb{P}$ is not known exactly but $\mathbb{P} \in \mathcal{F}$, we have a distributionally robust chance-constraint:

$$\inf_{\mathbb{P}\in\mathcal{F}} \mathbb{P}\left[p\min\{x,\tilde{d}\} - cx > t\right] \geq 1 - \epsilon$$

Research Objective

- Get a deterministic approximation of

$$\boldsymbol{X}_{\epsilon}^{*} \triangleq \{\boldsymbol{x} \in \boldsymbol{X} : \mathbb{P}\left[f(\boldsymbol{x}, \tilde{\boldsymbol{\omega}}) > 0\right] > 1 - \epsilon, \quad \forall \mathbb{P} \in \mathcal{F}\} \tag{3}$$

- How general the problem can we expect to approximate? What assumptions can we make on $\boldsymbol{X}, \boldsymbol{\Omega}, f$ and the ambiguity set $\mathcal{F}$?
- How do we measure the quality of the approximation?
- How efficient can we solve the deterministic approximation?

Uncertainty structure: Mixed ambiguity set

- Parametric probability distributions: For any $\boldsymbol{a} \in \boldsymbol{A} \subset \mathbb{R}^t$,

$$\mu_{\boldsymbol{a}} \in \mathcal{P}(\boldsymbol{\Omega})$$

- Mixture of parametric probability distributions: Given $\{\mu_{\boldsymbol{a}}\}_{\boldsymbol{a}\in\boldsymbol{A}} \subset \mathcal{P}(\Omega)$ and $\varphi \in \mathcal{P}(\boldsymbol{A})$, define

$$\mu(B) \triangleq \int_{\boldsymbol{A}} \mu_{\boldsymbol{a}}(B)\mathrm{d}\varphi(\boldsymbol{a}), \quad \forall B \in \mathcal{B}(\boldsymbol{\Omega})$$

- Ambiguity set of mixture of parametric probability distributions:

$$\mathcal{M}_{\boldsymbol{a}} \triangleq \left\{\mu : \mu(B) \triangleq \int_{\boldsymbol{A}} \mu_{\boldsymbol{a}}(B)\mathrm{d}\varphi(\boldsymbol{a}), \quad \varphi \in \mathcal{P}(\boldsymbol{A})\right\} \tag{4}$$

Examples: Mixed ambiguity set

Example 1.2 *(Mixtures of Gaussians)* $\mathbf{\Omega} = \mathbb{R}$, $\mathbf{a} = (a, \sigma) \in \mathbf{A} := [\underline{a}, \overline{a}] \times [\underline{\sigma}, \overline{\sigma}] \subset \mathbb{R}^2$, with $\underline{\sigma} > 0$, and

$$d\mu_{\mathbf{a}}(\omega) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\omega - a)^2}{2\sigma^2}\right) d\omega,$$

that is μ is a mixture of Gaussian probability measures with mean-deviation couple $(a, \sigma) \in \mathbf{A}$.

Example 1.7 With $\mathbf{\Omega} = \mathbb{R}$ one is given a finite family of probability measures $(\mu_i)_{i=1,\ldots,p} \subset \mathscr{P}(\mathbf{\Omega})$. Then $\mathbf{a} = a \in \mathbf{A} := \{1, \ldots, p\} \subset \mathbb{R}$, and

$$d\mu(\omega) = \sum_{a=1}^{p} \lambda_a \, d\mu_a(\omega); \quad \sum_{a \in \mathbf{A}} \lambda_a = 1; \quad \lambda_a \geq 0,$$

that is, μ is a finite convex combination of the probability measures $(\mu_{\mathbf{a}})$.

Assumptions on measurability, integrability, regularity

(i) For every $B \in \mathcal{B}(\mathbf{\Omega})$, the function $\mathbf{a} \mapsto \mu_{\mathbf{a}}(B)$ is measurable.

(ii) For every $\beta \in \mathbb{N}^p$:

$$\int_{\mathbf{\Omega}} \omega^\beta \, d\mu_{\mathbf{a}}(\omega) = p_\beta(\mathbf{a}), \quad \forall \mathbf{a} \in \mathbf{A}, \tag{2.2}$$

for some polynomial $p_\beta \in \mathbb{R}[\mathbf{a}]$.

(iii) For every $\mathbf{a} \in \mathbf{A}$ and every polynomial $g \in \mathbb{R}[\omega]$, $\mu_{\mathbf{a}}(\{\omega : g(\omega) = 0\}) = 0$.

(iv) For every bounded measurable (resp. bounded continuous) function q on $\mathbf{X} \times \mathbf{\Omega}$, the function

$$(\mathbf{x}, \mathbf{a}) \mapsto Q(\mathbf{x}, \mathbf{a}) := \int_{\mathbf{\Omega}} q(\mathbf{x}, \omega) \, d\mu_{\mathbf{a}}(\omega),$$

is bounded measurable (resp. bounded continuous) on $\mathbf{X} \times \mathbf{A}$.

(If $\mathbf{\Omega} \subset \mathbb{R}^p$ is unbounded):

There exists $c, \gamma > 0$ such that for every $i = 1, \ldots, p$:

$$\sup_{\mathbf{a} \in \mathbf{A}} \int_{\mathbf{\Omega}} \exp(c\,|\omega_i|) \, d\mu_{\mathbf{a}}(\omega) < \gamma. \tag{2.3}$$

Contributions

- If $\boldsymbol{X}, \boldsymbol{\Omega}, \boldsymbol{A}$ are basic semi-algebraic sets, f is a polynomial, $\mu_{\boldsymbol{a}}$ has polynomial moments, and some bounded, measurable, and integrable assumptions, a sequence of monotone inner approximations

$$\boldsymbol{X}_\epsilon^d \triangleq \{\boldsymbol{x} \in \boldsymbol{X} : w_d(\boldsymbol{x}) < \epsilon\} \tag{5}$$

  can be obtained where w_d is a polynomial of degree at most d and $\boldsymbol{X}_\epsilon^d \subset \boldsymbol{X}_\epsilon^{d+1} \subset \boldsymbol{X}_\epsilon^*$ for all d.
- Asymptotic guarantee: convergence in Lebesgue measure λ,

$$\lim_{d \to +\infty} \lambda(\boldsymbol{X}_\epsilon^* \backslash \boldsymbol{X}_\epsilon^d) = 0$$

- w_d can be solved from a hierarchy of semidefinite relaxations
- Extension to joint chance-constraints approximations
- Acceleration of convergence via Stokes' theorem

Outline

Preliminary

- Compact basic semi-algebraic sets:

$$
\begin{aligned}
\boldsymbol{X} &\triangleq \{\boldsymbol{x} \in \mathbb{R}^n : g_j(\boldsymbol{x}) \geq 0, j = 1, ..., m\} \\
\boldsymbol{\Omega} &\triangleq \{\boldsymbol{\omega} \in \mathbb{R}^q : s_l(\boldsymbol{\omega}) \geq 0, l = 1, ..., \bar{s}\} \\
\boldsymbol{A} &\triangleq \{\boldsymbol{a} \in \mathbb{R}^t : q_l(\boldsymbol{a}) \geq 0, l = 1, ..., L\}
\end{aligned}
$$

  - Compactness of $\boldsymbol{\Omega}$ can be relaxed.
  - Lebesgue measure λ can be normalized to probability measure
- Violation sets:

$$
\begin{aligned}
\boldsymbol{K} &\triangleq \{(\boldsymbol{x}, \boldsymbol{\omega}) \in \boldsymbol{X} \times \boldsymbol{\Omega} : f(\boldsymbol{x}, \boldsymbol{\omega}) \leq 0\} \\
\boldsymbol{K}_{\boldsymbol{x}} &\triangleq \{\boldsymbol{\omega} \in \boldsymbol{\Omega} : (\boldsymbol{x}, \boldsymbol{\omega}) \in \boldsymbol{K}\}
\end{aligned}
$$

Methodology: GMP model and Moment-SOS approach

Notations:

- $\mathbb{R}[\boldsymbol{x}]$: ring of polynomials; $\mathbb{R}[\boldsymbol{x}]_d$: polynomials of degree at most d with dimension $s(d) \triangleq \binom{n+d}{n}$
- $\mathcal{B}(\mathcal{X}), \mathcal{P}(\mathcal{X}), \mathscr{B}(\mathcal{X}), \mathcal{M}(\mathcal{X}), \mathcal{M}_+(\mathcal{X})$
- Given $\boldsymbol{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n}$, define
  - the functional $L_{\boldsymbol{y}} : f \in \mathbb{R}[\boldsymbol{x}] \to L_{\boldsymbol{y}}(f) \triangleq \sum_\alpha f_\alpha y_\alpha$
  - Moment matrix $\boldsymbol{M}_d(\boldsymbol{y})$ where $\boldsymbol{M}_d(\boldsymbol{y})(\alpha, \beta) \triangleq L_{\boldsymbol{y}}(\boldsymbol{x}^{\alpha+\beta}) = y_{\alpha+\beta}$ for all $\alpha, \beta \in \mathbb{N}_d^n$ given d
  - Localizing moment matrix $\boldsymbol{M}_d(g\boldsymbol{y})$ where $\boldsymbol{M}_d(g\boldsymbol{y})(\alpha, \beta) \triangleq L_{\boldsymbol{y}}(g(\boldsymbol{x})\boldsymbol{x}^{\alpha+\beta}) = \sum_\gamma g_\gamma y_{\alpha+\beta+\gamma}$ for all $\alpha, \beta \in \mathbb{N}_d^n$ given a polynomial g
- Given a measure ψ on $\boldsymbol{X} \times \boldsymbol{A}$, define ψ' on $\boldsymbol{X} \times \boldsymbol{A} \times \boldsymbol{\Omega}$ as $\mathrm{d}\psi' = \mathrm{d}\mu_{\boldsymbol{a}}(\boldsymbol{\omega})\mathrm{d}\psi(\boldsymbol{x}, \boldsymbol{a})$ with marginal $\psi'_{\boldsymbol{x},\boldsymbol{a}} = \psi$ and disintegration $\hat{\psi}'(\cdot \mid \boldsymbol{x}, \boldsymbol{a}) = \mu_{\boldsymbol{a}}$

Examples: Moment matrix

Example 2.1 For illustration, consider the case $n = 2$, $d = 2$. Then:

$$M_2(\mathbf{y}) = \begin{pmatrix} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{pmatrix}.$$

Moment matrix associated with $\boldsymbol{y}$

Example 2.3 For illustration, consider the case $n = 2$, $d = 1$. Then the localization matrix associated with $\mathbf{y}$ and $g = x_1 - x_2$, is:

$$M_1(g\mathbf{y}) = \begin{pmatrix} y_{10} - y_{01} & y_{20} - y_{11} & y_{11} - y_{02} \\ y_{20} - y_{11} & y_{30} - y_{21} & y_{21} - y_{12} \\ y_{11} - y_{02} & y_{21} - y_{12} & y_{12} - y_{03} \end{pmatrix}.$$

Localizing moment matrix associated with $\boldsymbol{y}$ and g

Characterize worst-case distribution

Lemma

For each $\boldsymbol{x} \in \boldsymbol{X}$ there exists measurable mappings $\boldsymbol{x} \to \boldsymbol{a}(\boldsymbol{x}) \in \boldsymbol{A}$ and $\boldsymbol{x} \to \kappa(\boldsymbol{x})$ such that:

$$\kappa(\boldsymbol{x}) = \max\{\mu(\boldsymbol{K_x}) : \mu \in \mathcal{M}_{\boldsymbol{a}}\} = \max\{\mu_{\boldsymbol{a}}(\boldsymbol{K_x}) : \boldsymbol{a} \in \boldsymbol{A}\} = \mu_{\boldsymbol{a}(\boldsymbol{x})}(\boldsymbol{K_x})$$

Hence

$$\boldsymbol{X}^*_\epsilon = \{\boldsymbol{x} \in \boldsymbol{X} : \kappa(\boldsymbol{x}) < \epsilon\} \tag{6}$$

- How to approximate $\kappa(\boldsymbol{x})$?
- Using duality! Consider a certain infinite dimension LP with the important property that any feasible solution of its dual provides the coefficients of some polynomial which is an upper bound of $\kappa(\boldsymbol{x})$.

GMP model

- Consider the infinite-dimensional LP:

$$\begin{array}{rl} \sup_{\phi,\psi} & \langle 1_{\boldsymbol{K}}, \phi \rangle \\ \text{s.t.} & \phi \leq T^* \psi \\ & \psi_{\boldsymbol{x}} = \lambda \\ & \phi \in \mathcal{M}_+(\boldsymbol{K}), \psi \in \mathcal{P}(\boldsymbol{X} \times \boldsymbol{A}) \end{array} \quad \text{(Infdim-LP)}$$

where

$$\begin{array}{l} T : g(\boldsymbol{x}, \boldsymbol{\omega}) \to (Tg)(\boldsymbol{x}, \boldsymbol{a}) \triangleq \int_\Omega g(\boldsymbol{x}, \boldsymbol{\omega}) \mathrm{d}\mu_{\boldsymbol{a}}(\omega) \\ \langle g, T^* \psi \rangle = \langle Tg, \psi \rangle \end{array}$$

Theorem

The model Infdim-LP has optimal value $\rho^* = \int_{\boldsymbol{K}} \kappa(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$. And the feasible pair (ϕ^*, ψ^*) with

$$\mathrm{d}\phi^*(\boldsymbol{x}, \boldsymbol{\omega}) \triangleq 1_{\boldsymbol{K}}(\boldsymbol{x}, \boldsymbol{\omega}) \mu_{\boldsymbol{a}(\boldsymbol{x})}(\mathrm{d}\omega) \mathrm{d}\lambda(\boldsymbol{x}), \quad \mathrm{d}\psi^*(\boldsymbol{x}, \boldsymbol{a}) \triangleq \delta_{\boldsymbol{a}(\boldsymbol{x})}(\mathrm{d}\boldsymbol{a}) \mathrm{d}\lambda(\boldsymbol{x})$$

is an optimal solution of model Infdim-LP.

GMP model

Dual:

$$\begin{array}{ll} \inf_{w,h} & \langle \lambda, w\rangle \\ \text{s.t.} & w \geq Th \quad \text{on } \boldsymbol{X} \times \boldsymbol{A} \\ & h \geq 1 \text{ on } \boldsymbol{K} \\ & h \geq 0 \text{ on } \boldsymbol{X} \times \boldsymbol{\Omega} \\ & h \in \mathbb{R}[\boldsymbol{x}, \boldsymbol{\omega}], w \in \mathbb{R}[\boldsymbol{x}] \end{array} \qquad \text{(Infdim-LPdual)}$$

Theorem

By weak duality, the optimal value of model Infdim-LPdual $\rho_D^* \geq \rho^*$. Also, for any feasible (w, h) in Infdim-LPdual, we have

$$w(\boldsymbol{x}) \geq \kappa(\boldsymbol{x}), \forall \boldsymbol{x} \in \boldsymbol{X}$$

Hence $\boldsymbol{X}_w \triangleq \{\boldsymbol{x} : w(\boldsymbol{x}) < \epsilon\} \subset \boldsymbol{X}_\epsilon^*$. Moreover, suppose no duality gap, i.e., $\rho^* = \rho_D^*$, and let (w_n, d_n) be a minimizing sequence of model Infdim-LPdual, then with $\|\cdot\|_1$ the norm of $L_1(\boldsymbol{X}, \lambda)$:

$$\lim_{n\to+\infty} \|w_n - \kappa\|_1 = 0, \qquad \lim_{n\to+\infty} \lambda(\boldsymbol{X}_\epsilon^* \backslash X_{w_n}) = 0$$

A Hierarchy of Semidefinite Relaxations

- Handle measure constraints through moments (truncated at degree $2d$): $(\phi, \varphi := T^*\psi - \phi, \psi \to \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v})$

$$\begin{array}{rl} \sup_{\boldsymbol{y},\boldsymbol{u},\boldsymbol{v}} & y_{00} \\ \text{s.t.} & L_{\boldsymbol{y}+\boldsymbol{u}}(\boldsymbol{x}^\alpha, \boldsymbol{\omega}^\beta) - L_{\boldsymbol{v}}(\boldsymbol{x}^\alpha p_\beta(\boldsymbol{a})) = 0 \\ & L_{\boldsymbol{v}}(\boldsymbol{x}^\alpha) = \lambda_\alpha \\ & \boldsymbol{M}_d(\boldsymbol{y}), \boldsymbol{M}_d(\boldsymbol{u}), \boldsymbol{M}_d(\boldsymbol{v}) \succeq \mathbf{0} \\ & \boldsymbol{M}_{d-d_{m+1}}(g_{m+1}\boldsymbol{y}) \succeq \mathbf{0} \\ & \boldsymbol{M}_{d-d_j^g}(g_j\boldsymbol{y}), \boldsymbol{M}_{d-d_j^g}(g_j\boldsymbol{u}), \boldsymbol{M}_{d-d_j^g}(g_j\boldsymbol{v}) \succeq \mathbf{0} \forall j = 1, ...m \\ & \boldsymbol{M}_{d-d_l^s}(s_l\boldsymbol{y}), \boldsymbol{M}_{d-d_l^s}(s_l\boldsymbol{u}) \succeq \mathbf{0} \forall l = 1, ..., \bar{s} \\ & \boldsymbol{M}_{d-d_l^q}(q_l\boldsymbol{v}) \succeq \mathbf{0} \forall l = 1, ..., L \end{array}$$

(SDP-relaxation)

where $g_{m+1} = -f, d_j^g = \lceil \deg(g_j)/2 \rceil, d_l^s = \lceil \deg(s_l)/2 \rceil, d_l^q = \lceil \deg(q_l)/2 \rceil$, and $\boldsymbol{y} = (y_{\alpha,\beta}), \boldsymbol{u} = (u_{\alpha,\beta}), (\alpha, \beta) \in \mathbb{N}^n \times \mathbb{N}^p$ and $\boldsymbol{v} = (v_{\alpha,\eta}), (\alpha, \eta) \in \mathbb{N}^n \times \mathbb{N}^t$ and $p_\beta(\boldsymbol{a}) = \int_{\boldsymbol{\Omega}} \boldsymbol{\omega}^\beta \mathrm{d}\mu_{\boldsymbol{a}}(\boldsymbol{\omega})$

Dual of SDP relaxation

- Interpretation: replace the positivity constraints in (Infdim-LPdual) by SOS constraints

$$
\begin{array}{rl}
\inf_{h,w,\sigma_j^i} & \int_{\boldsymbol{X}} w(\boldsymbol{x})\mathrm{d}\lambda(\boldsymbol{x}) \\
\text{s.t.} & h(\boldsymbol{x},\boldsymbol{\omega}) - 1 = \sum_{j=1}^{m+1} \sigma_j^1 g_j + \sum_{l=1}^{\bar{s}} \sigma_l^1 s_l, \quad \forall(\boldsymbol{x},\boldsymbol{\omega}) \\
& h(\boldsymbol{x},\boldsymbol{\omega}) = \sum_{j=1}^{m} \sigma_j^2 g_j + \sum_{l=1}^{\bar{s}} \sigma_l^2 s_l, \quad \forall(\boldsymbol{x},\boldsymbol{\omega}) \\
& w(\boldsymbol{x}) - \sum_{\alpha,\beta} h_{\alpha,\beta}\boldsymbol{x}^\alpha p_\beta(\boldsymbol{a}) = \sum_{j=1}^{m} \sigma_j^3 g_j + \sum_{l=1}^{L} \sigma_l^3 q_l \\
& \qquad \forall(\boldsymbol{x},\boldsymbol{a}) \\
& h(\boldsymbol{x},\boldsymbol{\omega}) = \sum_{|\alpha+\beta|\leq 2d} h_{\alpha,\beta}\boldsymbol{x}^\alpha\boldsymbol{\omega}^\beta \\
& w(\boldsymbol{x}) = \sum_{|\alpha|\leq 2d} w_\alpha \boldsymbol{x}^\alpha \\
& \sigma_j^1 \in \Sigma[\boldsymbol{x},\boldsymbol{\omega}]_{d-d_j^g}, j = 1,\ldots,m+1 \\
& \sigma_j^2, \sigma_j^3 \in \Sigma[\boldsymbol{x},\boldsymbol{\omega}]_{d-d_j^g}, j = 1,\ldots,m \\
& \sigma_l^1, \sigma_l^2 \in \Sigma[\boldsymbol{x},\boldsymbol{\omega}]_{d-d_l^s}, l = 1,\ldots,\bar{s} \\
& \sigma_l^3 \in \Sigma[\boldsymbol{x},\boldsymbol{a}]_{d-d_l^q}, l = 1,\ldots,\bar{s}
\end{array}
\qquad \text{(SDP-dual)}
$$

where $\Sigma[\boldsymbol{x},\boldsymbol{\omega}]$ is the space of SOS polynomials

Power of Moment-SOS approach

Theorem

Under mild bounded and measurable conditions, and assume $\boldsymbol{K}, \boldsymbol{A}, \boldsymbol{X} \times \boldsymbol{\Omega}, \boldsymbol{X} \times \boldsymbol{\Omega} \backslash \boldsymbol{K}$ *have nonempty interior. Then*

- *Slater's condition holds for* (SDP-relaxation) *so strong duality holds. Namely, the optimal value of* (SDP-relaxation) $\rho_d^* = \rho_{Dd}^*$, *the optimal value of* (SDP-dual).
- *Let* $\boldsymbol{X}_\epsilon^d := \{\boldsymbol{x} \in \boldsymbol{X} : w_d(\boldsymbol{x}) < \epsilon\}$, *then* $\boldsymbol{X}_\epsilon^d \subset \boldsymbol{X}^\epsilon$. *In addition,* $\lim_{d\to+\infty} \rho_d^* = \rho^*, \lim_{d\to+\infty} \|w_d(\boldsymbol{x}) - \kappa(\boldsymbol{x})\|_{L^1} = 0, \lim_{d\to+\infty} \lambda(X_\epsilon^* \backslash \boldsymbol{X}_\epsilon^d) = 0$,

Acceleration via Stoke's Theorem

- Adding constraints satisfied at optimal solution to facilitate convergence
- For mixed Gaussian distribution $\mathrm{d}\mu_{\boldsymbol{a}}(\omega) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\omega-a)^2}{2\sigma^2}} \mathrm{d}\omega$, since $f(\boldsymbol{x},\omega) = 0$ on the boundary of $\boldsymbol{K_x}$, Stoke's Theorem yields:

$$\int_{\boldsymbol{K_x}} q_\beta(\boldsymbol{x},\omega,\boldsymbol{a})\mathrm{d}\mu_a(\omega) = 0, \forall \boldsymbol{x} \in \boldsymbol{X}, \forall \beta = 0, 1, ...$$

where $q_\beta(\boldsymbol{x},\omega,\boldsymbol{a}) = \frac{\sigma^2 \partial(\omega^\beta f(\boldsymbol{x},\omega))}{\partial\omega} - \omega^\beta f(\boldsymbol{x},\omega)(\omega - a)$ Then

$$\int_{\boldsymbol{K}} \boldsymbol{x}^\alpha \boldsymbol{a}(\boldsymbol{x})^\gamma q_\beta(\boldsymbol{x},\omega,\boldsymbol{a}(\boldsymbol{x}))\mathrm{d}\phi^*(\boldsymbol{x},\omega) = 0, \forall \alpha \in \mathbb{N}^n, \gamma \in \mathbb{N}^2, \beta = 0, 1, ...$$

- Since $\boldsymbol{a}(\boldsymbol{x})$ is not a polynomial of $(\boldsymbol{x},\omega)$, we need to lift to $(\boldsymbol{x},\omega,\boldsymbol{a})$ by introducing φ on $\boldsymbol{K} \times \boldsymbol{A}$ and constraints

$$\varphi_{\boldsymbol{x},\omega} = \phi, \varphi_{\boldsymbol{x},\boldsymbol{a}} \leq \psi, \int_{\boldsymbol{K}\times\boldsymbol{A}} \boldsymbol{x}^\alpha \boldsymbol{a}^\gamma q_\beta(\boldsymbol{x},\omega,\boldsymbol{a})\mathrm{d}\varphi(\boldsymbol{x},\omega,\boldsymbol{a}) = 0, \forall \alpha, \gamma, \beta$$

Outline

1. Problem
2. Approach
   - Basic notations
   - Approximation via a GMP
   - Solve the GMP via SDP relaxation
3. Numerical Study

Numerical Experiment I

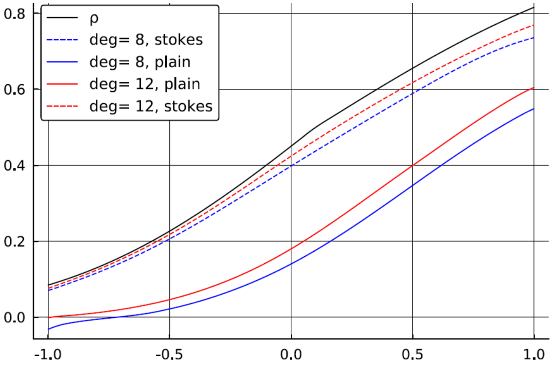

Fig. 1 Approximation of $\rho(\mathbf{x})$ (black) by polynomials $1 - w_4(\mathbf{x})$ (blue) and $1 - w_6(\mathbf{x})$ (red), dashed/solid lines correspond to with/without Stokes constraints (color figure online)

Figure 1: Comparison of approximation with and without acceleration: $\boldsymbol{X} = [-1, 1]$, $f(x, \omega) = x - \omega$, Gaussian distribution with parameters $\boldsymbol{A} = [-0.1, 0.1] \times [0.8, 1]$. $\rho(x) = 1 - \kappa(x) = \inf_{\mathbb{P} \in \mathcal{M}_a} \mathbb{P}\left[f(x, \omega) > 0\right]$

Numerical Experiment II

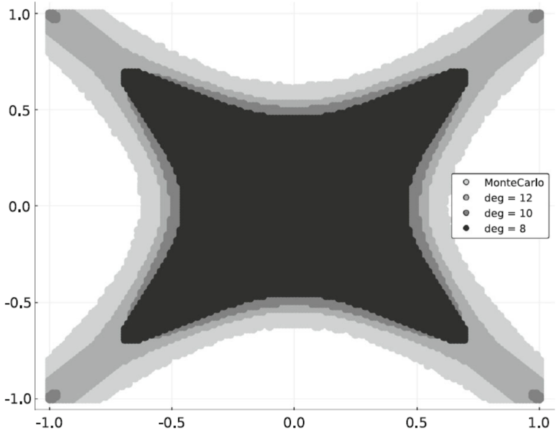

Fig. 2 Monte Carlo simulation (light grey) of $\mathbf{X}_\varepsilon^*$ and inner approximations $\mathbf{X}_\varepsilon^d$ for $d = 4, 5, 6$, in decreasing intensity

Figure 2: Inner approximations from various relaxations

Numerical Experiment III

$(d, \text{time})\backslash\varepsilon$	50%	25%	12.5%	6.25%	3.125%
4 (10 s)	97.0%	82.3%	68.5%	21.3%	0%
5 (107 s)	99.9%	86.0%	72.2%	43.19%	2.4%
6 (1042 s)	100.0%	89.4%	78.8%	60.3%	27.4%

Figure 3: Polynomial approximations versus Monte Carlo simulation with different violation levels. $\boldsymbol{X} = [-1, 1]^3$, $\boldsymbol{\Omega} = \mathbb{R}$, $f(x, \omega) = -2\omega x_1^2 + 2\omega x_2^2 - 2\omega x_3^2 - 1$. MC time: 2643s

Conclusion

- A systematic numerical scheme which provides a monotone sequence (a hierarchy) of inner approximations with strong asymptotic guarantees
- Pros: generality, no convexity assumptions
- Cons: computationally demanding, not scalable