Chance-constrained optimization with tight confidence bounds

Mark Cannon
University of Oxford

25 April 2017

Outline

1. Problem definition and motivation

2. Confidence bounds for randomised sample discarding

3. Posterior confidence bounds using empirical violation estimates

4. Randomized algorithm with prior and posterior confidence bounds

5. Conclusions and future directions

Chance-constrained optimization

Chance constraints limit the probability of constraint violation
in optimization problems with stochastic parameters

Applications are numerous and diverse:

> *...building control, chemical process design, finance, portfolio management, production planning, supply chain management, sustainable development, telecommunications networks ...*

Long history of stochastic programming methods

e.g. Charnes & Cooper, Management Sci., 1959
Prekopa, SIAM J. Control, 1966

but exact handling of probabilistic constraints is generally intractable

This talk: approximate methods based on random samples

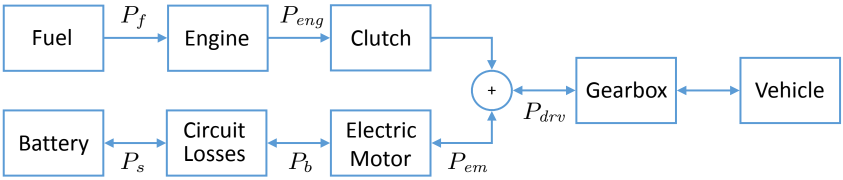
Motivation: Energy management in PHEVs

Battery: charged before trip, empty at end of trip

Electric motor: shifts engine load point to improve fuel efficiency or emissions

Control problem: optimize power flows to maximize performance (efficiency or emissions) over the driving cycle

Motivation: Energy management in PHEVs



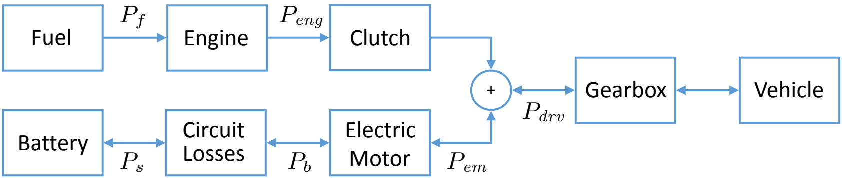
Parallel configuration with common engine and electric motor speed

$$P_{drv} = P_{eng} + P_{em}$$
$$\omega_{eng} = \omega_{em} = \omega \qquad (\text{if } \omega \geq \omega_{stall})$$

- Engine fuel consumption:

$$P_f = P_f(P_{eng}, \omega)$$

Motivation: Energy management in PHEVs



- Electric motor electrical power $P_b \leftrightarrow$ mechanical power P_{em}:

$$P_b = P_b(P_{em}, \omega)$$

- Battery stored energy flow rate $P_s \leftrightarrow P_b$:

$$P_s = P_s(P_b)$$

stored energy $E \leftrightarrow P_s$:

$$\dot{E} = -P_s$$

Motivation: Energy management in PHEVs

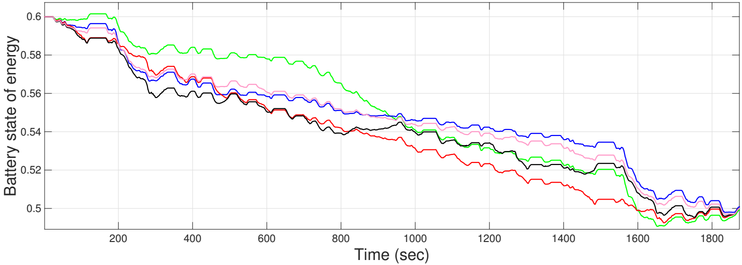
Nominal optimization, assuming known $P_{drv}(t)$, $\omega(t)$ for $t \in [0, T]$:

$$\begin{aligned}
\underset{\substack{P_{eng}(\tau),\ P_{em}(\tau)\\ \tau\in[t,T]}}{\text{minimize}} \quad & \int_0^T P_f(P_{eng}(\tau), \omega(\tau))\, \mathrm{d}\tau \\
\text{subject to} \quad & \dot{E}(\tau) = -P_s(\tau),\ \ \tau \in [t, T] \\
& E(t) = E_{now} \\
& E(T) \geq E_{end}
\end{aligned}$$

with
$$\begin{aligned}
P_{drv} &= P_{eng} + P_{em} \\
P_b &= P_b(P_{em}, \omega) \\
P_s &= P_s(P_b)
\end{aligned}$$

Motivation: Energy management in PHEVs

$P_{drv}(t)$, $\omega(t)$ depend on uncertain traffic flow and driver behaviour

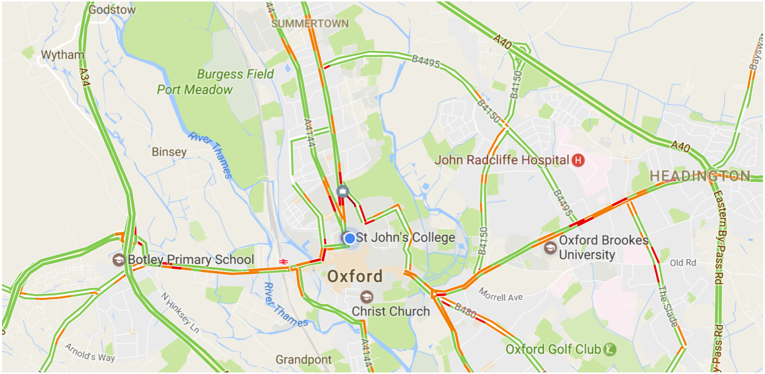


Introduce feedback by using a receding horizon implementation

but how likely is it that we can meet constraints?

Motivation: Energy management in PHEVs

$P_{drv}(t)$, $\omega(t)$ depend on uncertain traffic flow and driver behaviour



Real time traffic flow data

Motivation: Energy management in PHEVs

Reformulate as a chance-constrained optimization,
assuming known distributions for $P_{drv}(t)$, $\omega(t)$:

$$\begin{aligned}
\underset{\substack{P_{eng}(\tau),\ P_{em}(\tau)\\ \tau\in[t,T]}}{\text{minimize}} \quad & \int_0^T P_f(P_{eng}(\tau), \omega(\tau))\, \mathrm{d}\tau \\
\text{subject to} \quad & \dot{E}(\tau) = -P_s(\tau), \ \ \tau \in [t, T] \\
& E(t) = E_{now} \\
& \mathbb{P}\big\{E(T) < E_{end}\big\} \leq \epsilon
\end{aligned}$$

for a specified probability ϵ

Chance-constrained optimization

Define the chance-constrained problem

$$\text{CCP}: \quad \underset{x \in \mathcal{X}}{\text{minimize}} \;\; c^\top x \qquad \text{subject to} \;\; \mathbb{P}\big\{ f(x,\delta) > 0 \big\} \leq \epsilon.$$

$\delta \in \Delta \subseteq \mathbb{R}^d$	stochastic parameter
$\epsilon \in [0,1]$	maximum allowable violation probability

Assume

$\mathcal{X} \subset \mathbb{R}^n$ compact and convex

$f(\cdot,\delta) : \mathbb{R}^n \times \Delta \to \mathbb{R}$ convex, lower-semicontinuous for any $\delta \in \Delta$

CCP is feasible, i.e. $\mathbb{P}\{f(x,\delta) > 0\} \leq \epsilon$ for some $x \in \mathcal{X}$.

Chance-constrained optimization

Computational difficulties with chance constraints:

1. Prohibitive computation to determine feasibility of a given x

$$\mathbb{P}\{f(x,\delta) > 0\} = \int_{\delta\in\Delta} \mathbb{1}_{\{f(x,\delta)>0\}}(\delta) \ F(\mathrm{d}\delta)$$

   – Closed form expression only available in special cases

   e.g. if $\begin{cases} \delta \sim \mathcal{N}(0,I) & \text{Gaussian} \\ f(x,\delta) = b^\top\delta - 1 + (d + D^\top\delta)^\top x & \text{affine} \end{cases}$

   then

$$\mathbb{P}\{f(x,\delta) > 0\} \le \epsilon \quad \Longleftrightarrow \quad \Phi_{\mathcal{N}}^{-1}(1-\epsilon)\|b + Dx\| \le 1 - d^\top x$$

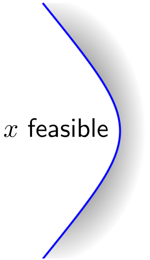
   – In general multidimensional integration is needed (e.g. Monte Carlo methods)

Chance-constrained optimization

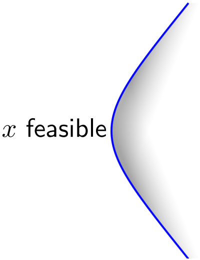
Computational difficulties with chance constraints:

2. Feasible set may be nonconvex even if $f(\cdot,\delta)$ is convex for all $\delta \in \Delta$

   e.g. $\Phi_{\mathcal{N}}^{-1}(1-\epsilon)\,\|b + Dx\| \leq 1 - d^\top x$ is convex if $\epsilon \leq 0.5$, nonconvex if $\epsilon > 0.5$



$\epsilon < 0.5$



$\epsilon > 0.5$

Sample approximation

Define the multisample $\omega_m = \{\delta^{(1)}, \ldots, \delta^{(m)}\}$ for i.i.d. $\delta^{(i)} \in \Delta$

Define the sample approximation of CCP as

$$\begin{aligned} \mathsf{SP}: \quad \underset{\substack{x \in \mathcal{X} \\ \omega \subseteq \omega_m}}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & f(x, \delta) \leq 0 \quad \text{for all } \delta \in \omega \\ & |\omega| \geq q \end{aligned}$$

$m - q$ discarded constraints, $m \geq q \geq n$

$x^*(\omega^*)$: optimal solution for x

ω^*: optimal solution for ω

Sample approximation

- Solutions of SP converge to the solution of CCP as $m \to \infty$ if $q = m(1 - \epsilon)$

- Approximation accuracy characterized via bounds on:
  - probability that solution of SP is feasible for CCP
  - probability that optimal objective of SP exceeds that of CCP

Luedtke & Ahmed, *SIAM J. Optim.*, 2008
Calafiore, *SIAM J. Optim.*, 2010
Campi & Garatti, *J. Optim. Theory Appl.*, 2011

- SP can be formulated as a mixed integer program with a convex continuous relaxation

Example: minimal bounding circle

Smallest circle containing a given probability mass

$$\text{CCP}: \quad \underset{c,R}{\text{minimize}} \quad R \quad \text{subject to} \quad \mathbb{P}\big\{\|c-\delta\| > R\big\} \le \epsilon$$

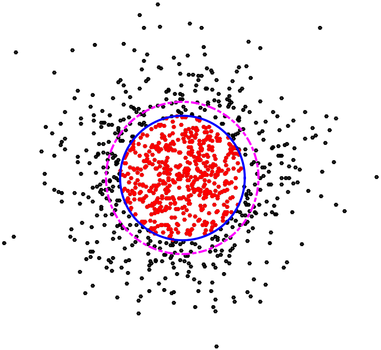
$$\text{SP}: \quad \underset{\substack{c,R\\ \omega\subseteq\omega_m}}{\text{minimize}} \quad R \quad \text{subject to} \quad \begin{array}{l}\|c-\delta\| \le R \text{ for all } \delta\in\omega \\ |\omega| \ge q\end{array}$$

e.g. $\epsilon = 0.5$, $m = 1000$, $q = 500$

	R^*	c^*
CCP	1.386	$(0,0)$
SP	1.134	$(-0.003,-0.002)$

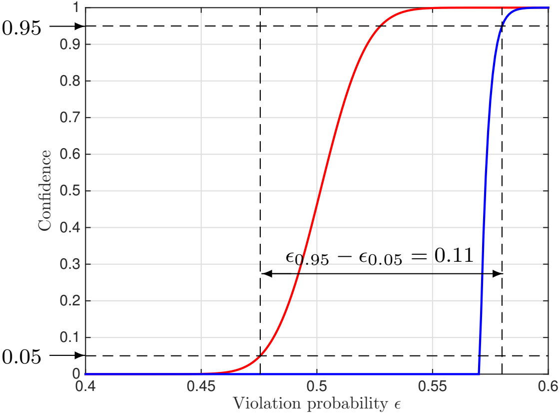
SP solved for one realisation ω_{1000}

$\delta \sim \mathcal{N}(0, I)$



Example: minimal bounding circle

Bounds on confidence that SP solution, $x^*(\omega^*)$ with $m = 1000$, $q = 500$ is feasible for CCP (ϵ) from Calafiore, 2010



$\Longrightarrow$ with 90% probability: the violation probability of $x^*(\omega^*)$ lies between 0.47 and 0.58

Example: minimal bounding circle

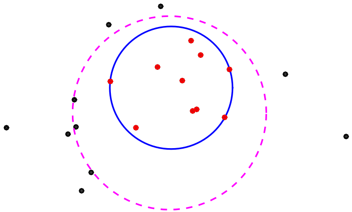
Smallest circle containing a given probability mass – observations:

- SP computation grows rapidly with m
- Bounds on confidence of feasibility of $x^*(\omega^*)$ for CCP are not tight

  (unless $q < m$ or probability distribution is a specific worst case)

  because SP solution has poor generalization properties

  e.g. $m = 20$, $q = 10$:



- Randomized sample discarding improves generalization

Outline

Level-q subsets of a multisample

Define the sampled problem for given $\omega \subseteq \omega_m$ as

$$\text{SP}(\omega): \quad \underset{x \in \mathcal{X}}{\text{minimize}} \quad c^\top x \quad \text{subject to} \quad f(x, \delta) \le 0 \ \ \text{for all } \delta \in \omega$$

$x^*(\omega)$: optimal solution of SP(ω)

Assumptions:

(i). SP(ω) is feasible for any $\omega \subseteq \omega_m$ with probability (w.p.) 1

(ii). $f\big(x^*(\omega), \delta\big) \neq 0$ for all $\delta \in \omega_m \setminus \omega$ w.p. 1

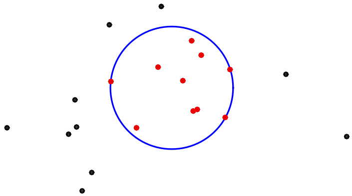
Level-q subsets of a multisample

Define the set of level-q subsets of ω_m, for $q \leq m$ as

$$\Omega_q(\omega_m) = \left\{\omega \subseteq \omega_m : |\omega| = q \text{ and } f\left(x^*(\omega), \delta\right) > 0 \text{ for all } \delta \in \omega_m \setminus \omega\right\}$$

$x^*(\omega)$ for $\omega \in \Omega_q(\omega_m)$ is called a level-q solution

e.g. minimal bounding circle problem, $m = 20$, $q = 10$:



$\omega \in \Omega_{10}(\omega_{20})$

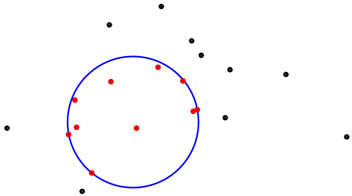
Level-q subsets of a multisample

Define the set of level-q subsets of ω_m, for $q \leq m$ as

$$\Omega_q(\omega_m) = \left\{\omega \subseteq \omega_m : |\omega| = q \text{ and } f\left(x^*(\omega), \delta\right) > 0 \text{ for all } \delta \in \omega_m \setminus \omega\right\}$$

$x^*(\omega)$ for $\omega \in \Omega_q(\omega_m)$ is called a level-q solution

e.g. minimal bounding circle problem, $m = 20$, $q = 10$:



$\omega \in \Omega_{10}(\omega_{20})$

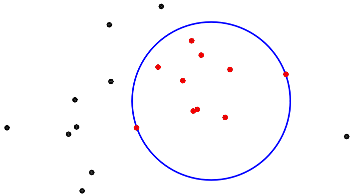
Level-q subsets of a multisample

Define the set of level-q subsets of ω_m, for $q \leq m$ as

$$\Omega_q(\omega_m) = \left\{ \omega \subseteq \omega_m : |\omega| = q \text{ and } f\big(x^*(\omega), \delta\big) > 0 \text{ for all } \delta \in \omega_m \setminus \omega \right\}$$

$x^*(\omega)$ for $\omega \in \Omega_q(\omega_m)$ is called a level-q solution

e.g. minimal bounding circle problem, $m = 20$, $q = 10$:



$\omega \in \Omega_{10}(\omega_{20})$

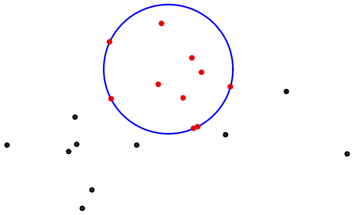
Level-q subsets of a multisample

Define the set of level-q subsets of ω_m, for $q \leq m$ as

$$\Omega_q(\omega_m) = \left\{\omega \subseteq \omega_m : |\omega| = q \text{ and } f\big(x^*(\omega), \delta\big) > 0 \text{ for all } \delta \in \omega_m \setminus \omega\right\}$$

$x^*(\omega)$ for $\omega \in \Omega_q(\omega_m)$ is called a level-q solution

e.g. minimal bounding circle problem, $m = 20$, $q = 10$:



$\omega \in \Omega_{10}(\omega_{20})$

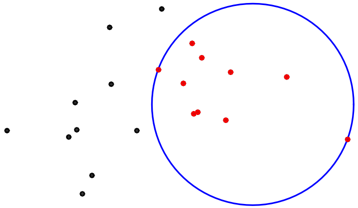
Level-q subsets of a multisample

Define the set of level-q subsets of ω_m, for $q \leq m$ as

$$\Omega_q(\omega_m) = \left\{\omega \subseteq \omega_m : |\omega| = q \text{ and } f\big(x^*(\omega), \delta\big) > 0 \text{ for all } \delta \in \omega_m \setminus \omega\right\}$$

$x^*(\omega)$ for $\omega \in \Omega_q(\omega_m)$ is called a level-q solution

e.g. minimal bounding circle problem, $m = 20$, $q = 10$:



$\omega \in \Omega_{10}(\omega_{20})$

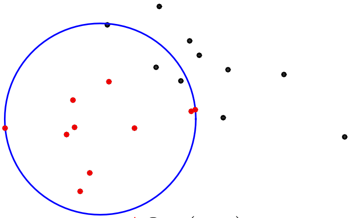
Level-q subsets of a multisample

Define the set of level-q subsets of ω_m, for $q \leq m$ as

$$\Omega_q(\omega_m) = \Big\{\omega \subseteq \omega_m : |\omega| = q \text{ and } f\big(x^*(\omega), \delta\big) > 0 \text{ for all } \delta \in \omega_m \setminus \omega\Big\}$$

$x^*(\omega)$ for $\omega \in \Omega_q(\omega_m)$ is called a level-q solution

e.g. minimal bounding circle problem, $m = 20$, $q = 10$:



$\omega \notin \Omega_{10}(\omega_{20})$

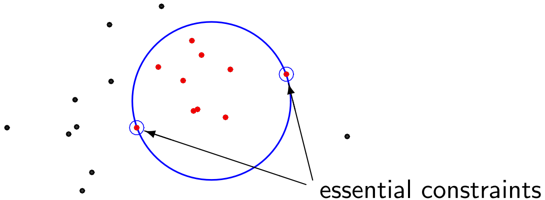
Essential constraint set

$\mathrm{Es}(\omega) = \{\delta^{(i_1)}, \ldots, \delta^{(i_k)}\} \subseteq \omega$ is an essential set of SP(ω) if

(i). $x\big(\{\delta^{(i_1)}, \ldots, \delta^{(i_k)}\}\big) = x^*(\omega)$,

(ii). $x(\omega \setminus \delta) \neq x^*(\omega)$ for all $\delta \in \{\delta^{(i_1)}, \ldots, \delta^{(i_k)}\}$.

e.g. minimal bounding circle problem



Support dimension

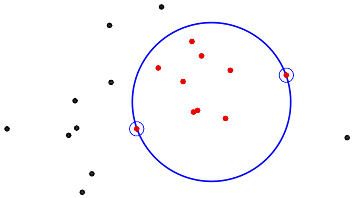
Define:

maximum support dimension: $\bar{\zeta} = \operatorname{ess\,sup}_{\substack{\omega \in \Delta^m \\ m \geq 1}} |\mathrm{Es}(\omega)|$

minimum support dimension: $\underline{\zeta} = \operatorname{ess\,inf}_{\substack{\omega \in \Delta^m \\ m \geq \bar{\zeta}}} |\mathrm{Es}(\omega)|$

then $\quad \bar{\zeta} \leq \dim(x) = n \quad$ and $\quad \underline{\zeta} \geq 1$

e.g. minimal bounding circle problem



$$|\mathrm{Es}(\omega)| = 2 = \underline{\zeta}$$

Support dimension

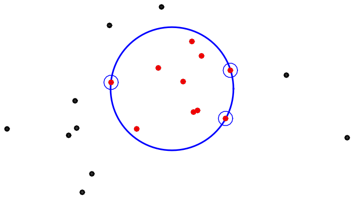
Define:

maximum support dimension: $\bar{\zeta} = \operatorname{ess\,sup}_{\substack{\omega \in \Delta^m \\ m \geq 1}} |\mathrm{Es}(\omega)|$

minimum support dimension: $\underline{\zeta} = \operatorname{ess\,inf}_{\substack{\omega \in \Delta^m \\ m \geq \bar{\zeta}}} |\mathrm{Es}(\omega)|$

then $\bar{\zeta} \leq \dim(x) = n$ and $\underline{\zeta} \geq 1$

e.g. minimal bounding circle problem



$$|\mathrm{Es}(\omega)| = 3 = \bar{\zeta}$$

Confidence bounds for randomly selected level-q subsets

Theorem 1:

For any $\epsilon \in [0,1]$ and $q \in [\bar{\zeta}, m]$,

$$\mathbb{P}^m\big\{V\big(x^*(\omega)\big) \le \epsilon \mid \omega \in \Omega_q(\omega_m)\big\} \ge \Phi(q - \bar{\zeta}; m, 1-\epsilon).$$

For any $\epsilon \in [0,1]$ and $q \in [\underline{\zeta}, m]$,

$$\mathbb{P}^m\big\{V\big(x^*(\omega)\big\} \le \epsilon \mid \omega \in \Omega_q(\omega_m)\big\} \le \Phi(q - \underline{\zeta}; m, 1-\epsilon).$$

where

violation probability: $V(x) = \mathbb{P}\big\{f(x,\delta) > 0\big\}$

binomial c.d.f.: $\Phi(n; N, p) = \sum_{i=0}^{n} \binom{N}{i} p^i (1-p)^{N-i}$

Confidence bounds for randomly selected level-q subsets

Proof of Thm. 1 relies on ...

Lemma 1: For any $v \in [0,1]$ and $k \in [\underline{\zeta}, \bar{\zeta}]$,

$$\mathbb{P}^m\big\{V\big(x^*(\omega_k)\big) \le v \;\big|\; |\mathrm{Es}(\omega_k)| = k\big\} = F(v) = v^k$$

(e.g. Calafiore, 2010)

Proof (sketch): for all $q \in [k, m]$,

$$\mathbb{P}^m\big\{\omega_k = \mathrm{Es}(\omega_q) \cap V\big(x^*(\omega_k)\big) = v \;\big|\; |\mathrm{Es}(\omega_k)| = k\big\} = (1-v)^{q-k}\,\mathrm{d}F(v)$$

$$\implies \quad \mathbb{P}^m\big\{\omega_k = \mathrm{Es}(\omega_q) \;\big|\; |\mathrm{Es}(\omega_k)| = k\big\} = \int_0^1 (1-v)^{q-k}\,\mathrm{d}F(v)$$

hence $\displaystyle\int_0^1 (1-v)^{q-k}\,\mathrm{d}F(v) = \binom{q}{k}^{-1}$ — Hausdorff moment problem

Confidence bounds for randomly selected level-q subsets

Proof of Thm. 1 relies on ...

Lemma 1: For any $v \in [0,1]$ and $k \in [\underline{\zeta}, \bar{\zeta}]$,

$$\mathbb{P}^m\left\{V\left(x^*(\omega_k)\right) \leq v \mid |\mathrm{Es}(\omega_k)| = k\right\} = F(v) = v^k$$

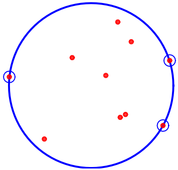
(e.g. Calafiore, 2010)

Proof (sketch): for all $q \in [k, m]$,

$$\mathbb{P}^m\left\{\omega_k = \mathrm{Es}(\omega_q) \cap V\left(x^*(\omega_k)\right) = v \mid |\mathrm{Es}(\omega_k)| = k\right\} = (1-v)^{q-k}\,\mathrm{d}F(v)$$

$$\Longrightarrow \quad \mathbb{P}^m\left\{\omega_k = \mathrm{Es}(\omega_q) \mid |\mathrm{Es}(\omega_k)| = k\right\} = \int_0^1 (1-v)^{q-k}\,\mathrm{d}F(v)$$

e.g. $k = 3$, $q = 10$:



hence $\displaystyle\int_0^1 (1-v)^{q-k}\,\mathrm{d}F(v) = \binom{q}{k}^{-1}$ — Hausdorff moment problem

Confidence bounds for randomly selected level-q subsets

…and

Lemma 2: For any $k \in [\underline{\zeta}, \bar{\zeta}]$ and $q \in [k, m]$,

$$\mathbb{P}^m \big\{ \omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m) \;\big|\; |\mathrm{Es}(\omega_k)| = k \big\} \\ = k \binom{m-k}{q-k} B(m-q+k, q-k+1)$$

($B(\cdot,\cdot)$ = beta function)

Note:

- $\omega_k = \{\delta^{(1)}, \ldots, \delta^{(k)}\}$ is statistically equivalent to a randomly selected subset of ω_m

- $\big\{ \omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m) \;\big|\; |\mathrm{Es}(\omega_k)| = k \big\}$ occurs iff:

  $q-k$ of the samples in $\omega_m \setminus \omega_k$ satisfy $f(x^*(\omega_k), \delta) \le 0$
  and the remaining $m-q$ samples satisfy $f(x^*(\omega_k), \delta) > 0$

  hence Lem. 2 follows from Lem. 1 and the law of total probability

Confidence bounds for randomly selected level-q subsets

...and

Lemma 2: For any $k \in [\underline{\zeta}, \bar{\zeta}]$ and $q \in [k, m]$,

$$\mathbb{P}^m\{\omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m) \mid |\mathrm{Es}(\omega_k)| = k\} = k\binom{m-k}{q-k} B(m-q+k, q-k+1)$$

($B(\cdot,\cdot)$ = beta function)

Note:

- $\omega_k = \{\delta^{(1)}, \ldots, \delta^{(k)}\}$ is statistically equivalent to a randomly selected subset of ω_m

- $\{\omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m) \mid |\mathrm{Es}(\omega_k)| = k\}$ occurs iff:

  $q - k$ of the samples in $\omega_m \setminus \omega_k$ satisfy $f(x^*(\omega_k), \delta) \le 0$
  and the remaining $m - q$ samples satisfy $f(x^*(\omega_k), \delta) > 0$

  hence Lem. 2 follows from Lem. 1 and the law of total probability

Confidence bounds for randomly selected level-q subsets

Proof of Thm. 1 (sketch):

- $$\mathbb{P}^m\{\omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m) \cap V\big(x^*(\omega_k)\big) \le \epsilon \;\big|\; |\mathrm{Es}(\omega_k)| = k\}$$
$$= k\binom{m-k}{q-k} B(\epsilon; m-q+k, q-k+1)$$

  ($B(\cdot,\cdot,\cdot)$ = incomplete beta function)

- $$\mathbb{P}^m\{V\big(x^*(\omega_k)\big) \le \epsilon \;\big|\; \omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m)\}$$
$$= \frac{\mathbb{P}^m_k\{\omega_k = \mathrm{Es}(\omega) \text{ for } \omega \in \Omega_q(\omega_m) \cap V\big(x^*(\omega_k)\big) \le \epsilon \;\big|\; |\mathrm{Es}(\omega_k)| = k\}}{\mathbb{P}^m_k\{\omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m) \;\big|\; |\mathrm{Es}(\omega_k)| = k\}}$$
$$= \Phi(q-k; m, 1-\epsilon)$$
  (Baye's law)

- Hence

  $$\mathbb{P}^m\{V\big(x^*(\omega)\big) \le \epsilon \;\big|\; \omega \in \Omega_q(\omega_m) \cap |\mathrm{Es}(\omega)| = k\} = \Phi(q-k; m, 1-\epsilon)$$

  Thm. 1 follows from total law of probability, since, for all $k \in [\underline{\zeta}, \bar{\zeta}]$,

  $$\Phi(q-\bar{\zeta}; m, 1-\epsilon) \le \Phi(q-k; m, 1-\epsilon) \le \Phi(q-\underline{\zeta}; m, 1-\epsilon)$$

Confidence bounds for randomly selected level-q subsets

Proof of Thm. 1 (sketch):

- $$\mathbb{P}^m\{\omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m) \cap V\big(x^*(\omega_k)\big) \le \epsilon \;\big|\; |\mathrm{Es}(\omega_k)| = k\}$$
$$= k\binom{m-k}{q-k} B(\epsilon; m-q+k, q-k+1)$$
$(B(\cdot,\cdot,\cdot) =$ incomplete beta function$)$

- $$\mathbb{P}^m\{V\big(x^*(\omega_k)\big) \le \epsilon \;\big|\; \omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m)\}$$
$$= \frac{\mathbb{P}_k^m\{\omega_k = \mathrm{Es}(\omega) \text{ for } \omega \in \Omega_q(\omega_m) \cap V\big(x^*(\omega_k)\big) \le \epsilon \;\big|\; |\mathrm{Es}(\omega_k)| = k\}}{\mathbb{P}_k^m\{\omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m) \;\big|\; |\mathrm{Es}(\omega_k)| = k\}}$$
$$= \Phi(q-k; m, 1-\epsilon)$$ (Baye's law)

- Hence
$$\mathbb{P}^m\{V\big(x^*(\omega)\big) \le \epsilon \;\big|\; \omega \in \Omega_q(\omega_m) \cap |\mathrm{Es}(\omega)| = k\} = \Phi(q-k; m, 1-\epsilon)$$
Thm. 1 follows from total law of probability, since, for all $k \in [\underline{\zeta}, \bar{\zeta}]$,
$$\Phi(q-\bar{\zeta}; m, 1-\epsilon) \le \Phi(q-k; m, 1-\epsilon) \le \Phi(q-\underline{\zeta}; m, 1-\epsilon)$$

Confidence bounds for randomly selected level-q subsets

Proof of Thm. 1 (sketch):

- $$\mathbb{P}^m\left\{\omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m) \cap V\big(x^*(\omega_k)\big) \le \epsilon \;\Big|\; |\mathrm{Es}(\omega_k)| = k\right\}$$
$$= k\binom{m-k}{q-k} B(\epsilon; m-q+k, q-k+1)$$
$(B(\cdot,\cdot,\cdot) = \text{incomplete beta function})$

- $$\mathbb{P}^m\left\{V\big(x^*(\omega_k)\big) \le \epsilon \;\Big|\; \omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m)\right\}$$
$$= \frac{\mathbb{P}^m_k\left\{\omega_k = \mathrm{Es}(\omega) \text{ for } \omega \in \Omega_q(\omega_m) \cap V\big(x^*(\omega_k)\big) \le \epsilon \;\Big|\; |\mathrm{Es}(\omega_k)| = k\right\}}{\mathbb{P}^m_k\left\{\omega_k = \mathrm{Es}(\omega) \text{ for some } \omega \in \Omega_q(\omega_m) \;\Big|\; |\mathrm{Es}(\omega_k)| = k\right\}}$$
$$= \Phi(q-k; m, 1-\epsilon) \qquad \text{(Baye's law)}$$

- Hence
$$\mathbb{P}^m\left\{V\big(x^*(\omega)\big) \le \epsilon \;\Big|\; \omega \in \Omega_q(\omega_m) \cap |\mathrm{Es}(\omega)| = k\right\} = \Phi(q-k; m, 1-\epsilon)$$
Thm. 1 follows from total law of probability, since, for all $k \in [\underline{\zeta}, \bar{\zeta}]$,
$$\Phi(q-\bar{\zeta}; m, 1-\epsilon) \le \Phi(q-k; m, 1-\epsilon) \le \Phi(q-\underline{\zeta}; m, 1-\epsilon)$$

Confidence bounds for randomly selected level-q subsets

Consider the relationship between optimal costs: $J^o(\epsilon)$ of CCP
$J^*(\omega)$ of SP(ω)

Prop. 1: $\mathbb{P}^m\big\{J^*(\omega) \geq J^o(\epsilon) \;\big|\; \omega \in \Omega_q(\omega_m)\big\} \geq \Phi(q - \bar{\zeta}; m, 1-\epsilon)$

Proof: $J^*(\omega) \geq J^o(\epsilon)$ if the solution x of SP(ω) is feasible for CCP, so

$$\mathbb{P}^m\big\{V\big(x^*(\omega)\big) \leq \epsilon \;\big|\; \omega \in \Omega_q(\omega_m)\big\} \leq \mathbb{P}^m\big\{J^*(\omega_r) \geq J^o(\epsilon) \;\big|\; \omega \in \Omega_q(\omega_m)\big\}$$

Prop. 2: $\mathbb{P}^m\big\{J^*(\omega) > J^o(\epsilon) \;\big|\; |\omega| = q\big\} \leq \Phi(q-1; q, 1-\epsilon)$

Proof: $J^*(\omega) \leq J^o(\epsilon)$ if the solution x^o of CCP is feasible for SP(ω)
i.e. if $f\big(x^o(\epsilon), \delta\big) \leq 0$ for all $\delta \in \omega$, so

$$\mathbb{P}^m\big\{J^*(\omega) \leq J^o(\epsilon)\big\} \geq \Big(\mathbb{P}\big\{f\big(x^o(\epsilon), \delta\big) \leq 0\big\}\Big)^q$$
$$\geq (1-\epsilon)^q = 1 - \Phi(q-1; q, 1-\epsilon)$$

Confidence bounds for randomly selected level-q subsets

Consider the relationship between optimal costs: $J^o(\epsilon)$ of CCP
$J^*(\omega)$ of SP(ω)

Prop. 1: $\mathbb{P}^m\{J^*(\omega) \geq J^o(\epsilon) \mid \omega \in \Omega_q(\omega_m)\} \geq \Phi(q - \bar{\zeta}; m, 1-\epsilon)$

Proof: $J^*(\omega) \geq J^o(\epsilon)$ if the solution x of SP(ω) is feasible for CCP, so

$$\mathbb{P}^m\{V(x^*(\omega)) \leq \epsilon \mid \omega \in \Omega_q(\omega_m)\} \leq \mathbb{P}^m\{J^*(\omega_r) \geq J^o(\epsilon) \mid \omega \in \Omega_q(\omega_m)\}$$

Prop. 2: $\mathbb{P}^m\{J^*(\omega) > J^o(\epsilon) \mid |\omega| = q\} \leq \Phi(q-1; q, 1-\epsilon)$

Proof: $J^*(\omega) \leq J^o(\epsilon)$ if the solution x^o of CCP is feasible for SP(ω)
i.e. if $f(x^o(\epsilon), \delta) \leq 0$ for all $\delta \in \omega$, so

$$\mathbb{P}^m\{J^*(\omega) \leq J^o(\epsilon)\} \geq \big(\mathbb{P}\{f(x^o(\epsilon), \delta) \leq 0\}\big)^q$$
$$\geq (1-\epsilon)^q = 1 - \Phi(q-1; q, 1-\epsilon)$$

Confidence bounds for randomly selected level-q subsets

Consider the relationship between optimal costs: $J^o(\epsilon)$ of CCP
$J^*(\omega)$ of SP(ω)

Prop. 1: $\mathbb{P}^m\big\{J^*(\omega) \geq J^o(\epsilon) \;\big|\; \omega \in \Omega_q(\omega_m)\big\} \geq \Phi(q - \bar{\zeta}; m, 1 - \epsilon)$

Proof: $J^*(\omega) \geq J^o(\epsilon)$ if the solution x of SP(ω) is feasible for CCP, so

$$\mathbb{P}^m\big\{V\big(x^*(\omega)\big) \leq \epsilon \;\big|\; \omega \in \Omega_q(\omega_m)\big\} \leq \mathbb{P}^m\big\{J^*(\omega_r) \geq J^o(\epsilon) \;\big|\; \omega \in \Omega_q(\omega_m)\big\}$$

Prop. 2: $\mathbb{P}^m\big\{J^*(\omega) > J^o(\epsilon) \;\big|\; |\omega| = q\big\} \leq \Phi(q - 1; q, 1 - \epsilon)$

Proof: $J^*(\omega) \leq J^o(\epsilon)$ if the solution x^o of CCP is feasible for SP(ω)
i.e. if $f\big(x^o(\epsilon), \delta\big) \leq 0$ for all $\delta \in \omega$, so

$$\begin{aligned}\mathbb{P}^m\big\{J^*(\omega) \leq J^o(\epsilon)\big\} &\geq \Big(\mathbb{P}\big\{f\big(x^o(\epsilon), \delta\big) \leq 0\big\}\Big)^q \\ &\geq (1-\epsilon)^q = 1 - \Phi(q-1; q, 1-\epsilon)\end{aligned}$$

Comparison with deterministic discarding strategies

For optimal sample discarding (i.e. SP) or a greedy heuristic we have

$$\mathbb{P}^m\big\{V\big(x^*(\omega^*)\big) \le \epsilon\big\} \ge \Psi(q, \bar{\zeta}; m, \epsilon),$$

$$\Psi(q, \bar{\zeta}; m, \epsilon) = 1 - \binom{m - q + \bar{\zeta} - 1}{m - q}\Phi(m - q + \bar{\zeta} - 1; m, \epsilon)$$

Calafiore (2010), Campi & Garatti (2011)

this bound is looser than the lower bound for random sample discarding (i.e. SP(ω), $\omega \in \Omega_q(\omega_m)$)

since

$$1 - \Psi(q, \bar{\zeta}; m, \epsilon) = \binom{m - q + \bar{\zeta} - 1}{m - q}\big(1 - \Phi(q - \bar{\zeta}; m, 1 - \epsilon)\big)$$

implies

- $\Psi(q, \bar{\zeta}; m, \epsilon) \le \Phi(q - \bar{\zeta}; m, 1 - \epsilon)$
- $\Psi(q, \bar{\zeta}; m, \epsilon) = \Phi(q - \bar{\zeta}; m, 1 - \epsilon)$ only if $\bar{\zeta} = 1$ or $q = m$ (in which case $|\Omega_q(\omega_m)| = 1$)

Comparison with deterministic discarding strategies

For optimal sample discarding (i.e. SP) or a greedy heuristic we have

$$\mathbb{P}^m\{V\big(x^*(\omega^*)\big) \le \epsilon\} \ge \Psi(q,\bar\zeta;m,\epsilon),$$

$$\Psi(q,\bar\zeta;m,\epsilon) = 1 - \binom{m-q+\bar\zeta-1}{m-q}\Phi(m-q+\bar\zeta-1;m,\epsilon)$$

Calafiore (2010), Campi & Garatti (2011)

this bound is looser than the lower bound for random sample discarding (i.e. SP(ω), $\omega \in \Omega_q(\omega_m)$)

since

$$1-\Psi(q,\bar\zeta;m,\epsilon) = \binom{m-q+\bar\zeta-1}{m-q}\big(1-\Phi(q-\bar\zeta;m,1-\epsilon)\big)$$

implies

- $\Psi(q,\bar\zeta;m,\epsilon) \le \Phi(q-\bar\zeta;m,1-\epsilon)$
- $\Psi(q,\bar\zeta;m,\epsilon) = \Phi(q-\bar\zeta;m,1-\epsilon)$ only if $\bar\zeta = 1$ or $q = m$ (in which case $|\Omega_q(\omega_m)| = 1$)

Comparison with deterministic discarding strategies

The lower confidence bound for optimal sample discarding corresponds to the worst case assumption:

$$\mathbb{P}^m\Big\{V\big(x^*(\omega)\big) > \epsilon\Big\} = 0 \text{ for all } \omega \in \Omega_q(\omega_m) \text{ such that } \omega \neq \omega^*$$

since

- $\star$ Thm. 1 implies

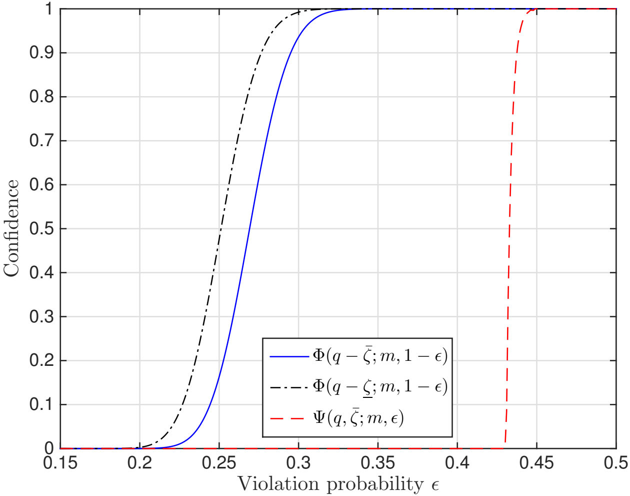
$$\frac{1}{\mathbb{E}^m\big\{|\Omega_q(\omega_m)|\big\}}\mathbb{E}^m\left\{\sum_{\omega\in\Omega_q(\omega_m)} \mathbb{P}^m\big\{V\big(x^*(\omega)\big) > \epsilon\big\}\right\} \leq 1 - \Phi(q - \bar{\zeta}; m, 1-\epsilon)$$

- $\star$ Lem. 2 implies

$$\mathbb{E}^m\big\{|\Omega_q(\omega_m)|\big\} \leq \binom{m - q + \bar{\zeta} - 1}{m - q}$$

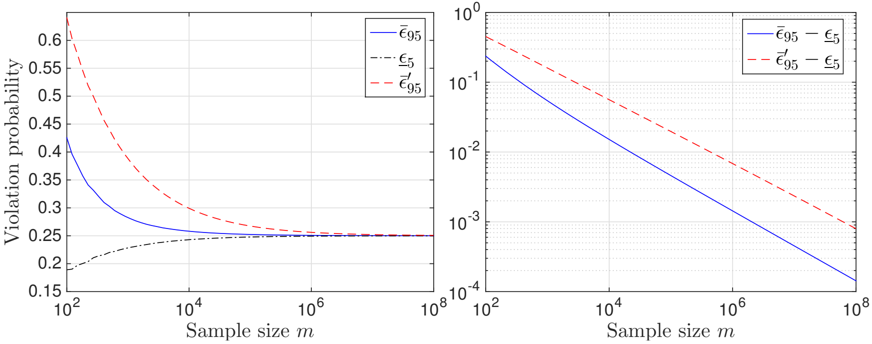
Comparison with deterministic discarding strategies

Confidence bounds for $m = 500$ with $q = \lceil 0.75m \rceil = 375$, $\underline{\zeta} = 1$, $\bar{\zeta} = 10$



Comparison with deterministic discarding strategies

Values of ϵ lying on 5% and 95% confidence bounds
for varying m with $q = \lceil 0.75m \rceil$, $\underline{\zeta} = 1$, $\bar{\zeta} = 10$



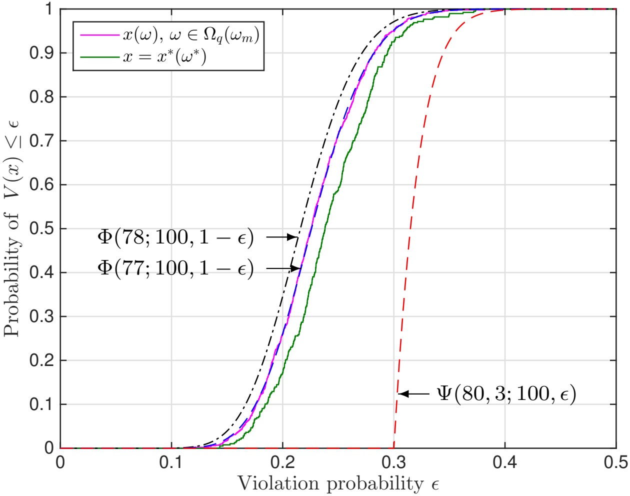
probablities $\underline{\epsilon}_5$, $\bar{\epsilon}_{95}$ and $\bar{\epsilon}'_{95}$ are defined by

$$0.05 = \Phi(q - \underline{\zeta}; m, 1 - \underline{\epsilon}_5), \quad 0.95 = \Phi(q - \bar{\zeta}; m, 1 - \bar{\epsilon}_{95})$$

$$0.95 = \Psi(q, \bar{\zeta}; m, \bar{\epsilon}'_{95})$$

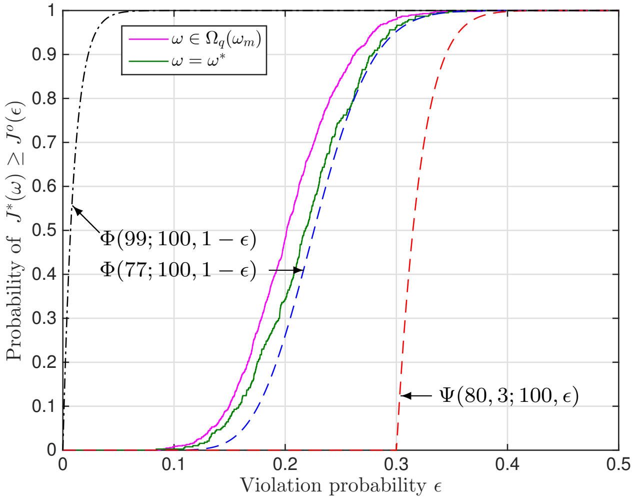
Example: minimal bounding circle

Empirical distribution of violation probability for smallest bounding circle with: $\delta \sim \mathcal{N}(0, I)$, $q = 80$, $m = 100$, and 500 realisations of ω_m



Example: minimal bounding circle

Empirical cost distributions for smallest bounding circle
with: $\delta \sim \mathcal{N}(0, I)$, $q = 80$, $m = 100$, and 500 realisations of ω_m



Outline

Random constraint selection strategy

How can we select $\omega \in \Omega_q(\omega_m)$ at random?

Summary of approach:

For given ω_m, let $\omega_r = \{\delta^{(1)}, \ldots, \delta^{(r)}\}$, and

- solve SP(ω_r) for $x^*(\omega_r)$
- count $q = \left|\left\{\delta \in \omega_m : f\big(x^*(\omega_r), \delta\big) \le 0\right\}\right|$

then $\left\{\delta \in \omega_m : f\big(x^*(\omega_r), \delta\big) \le 0\right\} \in \Omega_q(\omega_m)$

Choose r to maximize the probability
that q lies in the desired range

e.g. $r = 10$
$m = 50$
$q = 13 + 10 = 23$

Random constraint selection strategy

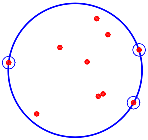
How can we select $\omega \in \Omega_q(\omega_m)$ at random?

Summary of approach:

For given ω_m, let $\omega_r = \{\delta^{(1)}, \ldots, \delta^{(r)}\}$, and

- solve $\mathrm{SP}(\omega_r)$ for $x^*(\omega_r)$
- count $q = \left|\left\{\delta \in \omega_m : f\big(x^*(\omega_r), \delta\big) \le 0\right\}\right|$

then $\left\{\delta \in \omega_m : f\big(x^*(\omega_r), \delta\big) \le 0\right\} \in \Omega_q(\omega_m)$



e.g. $r = 10$

$m = 50$

$q = 13 + 10 = 23$

Choose r to maximize the probability that q lies in the desired range

Random constraint selection strategy

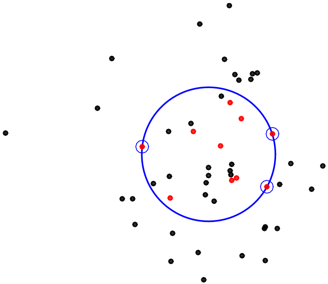
How can we select $\omega \in \Omega_q(\omega_m)$ at random?

Summary of approach:

For given ω_m, let $\omega_r = \{\delta^{(1)}, \ldots, \delta^{(r)}\}$, and

- solve SP(ω_r) for $x^*(\omega_r)$
- count $q = \left|\left\{\delta \in \omega_m : f\big(x^*(\omega_r), \delta\big) \le 0\right\}\right|$

then $\left\{\delta \in \omega_m : f\big(x^*(\omega_r), \delta\big) \le 0\right\} \in \Omega_q(\omega_m)$



e.g. $r = 10$

$m = 50$

$q = 13 + 10 = 23$

Choose r to maximize the probability that q lies in the desired range

Random constraint selection strategy

How can we select $\omega \in \Omega_q(\omega_m)$ at random?

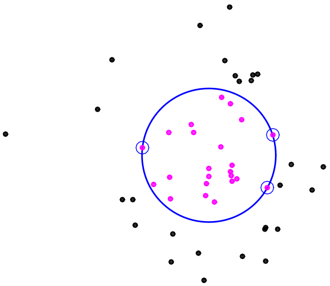
Summary of approach:

For given ω_m, let $\omega_r = \{\delta^{(1)}, \ldots, \delta^{(r)}\}$, and

- solve SP(ω_r) for $x^*(\omega_r)$
- count $q = \left|\left\{\delta \in \omega_m : f\big(x^*(\omega_r), \delta\big) \le 0\right\}\right|$

then $\left\{\delta \in \omega_m : f\big(x^*(\omega_r), \delta\big) \le 0\right\} \in \Omega_q(\omega_m)$

Choose r to maximize the probability that q lies in the desired range



e.g. $r = 10$
$m = 50$
$q = 13 + 10 = 23$

Posterior confidence bounds

Define $\theta_r(\omega_m) = \left|\left\{\delta \in \omega_m : f\big(x^*(\omega_r), \delta\big) \le 0\right\}\right|$

Theorem 2:

For any $\epsilon \in [0,1]$ and $\bar{\zeta} \le r \le q \le m$,

$$\mathbb{P}^m\left\{V\big(x^*(\omega_r)\big) \le \epsilon \mid \theta_r(\omega_m) = q\right\} \ge \Phi(q - \bar{\zeta}; m, 1-\epsilon).$$

For any $\epsilon \in [0,1]$ and $\underline{\zeta} \le r \le q \le m]$,

$$\mathbb{P}^m\left\{V\big(x^*(\omega_r)\big) \le \epsilon \mid \theta_r(\omega_m) = q\right\} \le \Phi(q - \underline{\zeta}; m, 1-\epsilon).$$

- $\theta_r(\omega_m)$ provides an empirical estimate of violation probability
- Thm. 2 gives *a posteriori* confidence bounds (i.e. for given $\theta_r(\omega_m)$)
- These bounds are tight for general probability distributions, i.e.
  - they cannot be improved if ζ can take any value in $[\underline{\zeta}, \bar{\zeta}]$
  - they are exact for problems with $\underline{\zeta} = \bar{\zeta}$

Posterior confidence bounds

Define $\theta_r(\omega_m) = \left|\left\{\delta \in \omega_m : f\big(x^*(\omega_r), \delta\big) \leq 0\right\}\right|$

Theorem 2:

For any $\epsilon \in [0,1]$ and $\bar{\zeta} \leq r \leq q \leq m$,

$$\mathbb{P}^m\left\{V\big(x^*(\omega_r)\big) \leq \epsilon \mid \theta_r(\omega_m) = q\right\} \geq \Phi(q - \bar{\zeta}; m, 1-\epsilon).$$

For any $\epsilon \in [0,1]$ and $\underline{\zeta} \leq r \leq q \leq m]$,

$$\mathbb{P}^m\left\{V\big(x^*(\omega_r)\big) \leq \epsilon \mid \theta_r(\omega_m) = q\right\} \leq \Phi(q - \underline{\zeta}; m, 1-\epsilon).$$

- $\theta_r(\omega_m)$ provides an empirical estimate of violation probability
- Thm. 2 gives *a posteriori* confidence bounds (i.e. for given $\theta_r(\omega_m)$)
- These bounds are tight for general probability distributions, i.e.
  - they cannot be improved if ζ can take any value in $[\underline{\zeta}, \bar{\zeta}]$
  - they are exact for problems with $\underline{\zeta} = \bar{\zeta}$

Probability of selecting a level-q subset of ω_m

Selection of r is based on

Lemma 3: For any $\bar{\zeta} \leq r \leq q \leq m$,

$$\mathbb{P}^m\{\theta_r(\omega_m) = q\} \geq \binom{m-r}{q-r} \min_{\zeta \in [\underline{\zeta}, \bar{\zeta}]} \frac{B(m-q+\zeta, q-\zeta+1)}{B(\zeta, r-\zeta+1)}$$

$$\mathbb{P}^m\{\theta_r(\omega_m) = q\} \leq \binom{m-r}{q-r} \max_{\zeta \in [\underline{\zeta}, \bar{\zeta}]} \frac{B(m-q+\zeta, q-\zeta+1)}{B(\zeta, r-\zeta+1)}$$

Proof (sketch):

1. From Lem. 1 & Lem. 2, for $k \in [\underline{\zeta}, \bar{\zeta}]$ we have

$$\mathbb{P}^m\{V(x^*(\omega_r)) = v \mid |\mathrm{Es}(\omega_r)| = k\} = \frac{(1-v)^{r-k} v^{k-1}}{B(k, r-k+1)}\, \mathrm{d}v$$

2. $\theta_r(\omega_m) = q$ iff

   $q-r$ of the $m-r$ samples in $\omega_m \setminus \omega_r$ satisfy $f(x^*(\omega_r), \delta) \leq 0$
   and the remaining $m-q$ samples satisfy $f(x^*(\omega_r), \delta) > 0$

   Hence Lem. 3 follows from 1 and the law of total probability

Probability of selecting a level-q subset of ω_m

Selection of r is based on

Lemma 3: For any $\bar{\zeta} \le r \le q \le m$,

$$\mathbb{P}^m\{\theta_r(\omega_m) = q\} \ge \binom{m-r}{q-r} \min_{\zeta \in [\underline{\zeta}, \bar{\zeta}]} \frac{B(m-q+\zeta, q-\zeta+1)}{B(\zeta, r-\zeta+1)}$$

$$\mathbb{P}^m\{\theta_r(\omega_m) = q\} \le \binom{m-r}{q-r} \max_{\zeta \in [\underline{\zeta}, \bar{\zeta}]} \frac{B(m-q+\zeta, q-\zeta+1)}{B(\zeta, r-\zeta+1)}$$

Proof (sketch):

1. From Lem. 1 & Lem. 2, for $k \in [\underline{\zeta}, \bar{\zeta}]$ we have

$$\mathbb{P}^m\{V(x^*(\omega_r)) = v \mid |\mathrm{Es}(\omega_r)| = k\} = \frac{(1-v)^{r-k} v^{k-1}}{B(k, r-k+1)} \, \mathrm{d}v$$

2. $\theta_r(\omega_m) = q$ iff

   $q - r$ of the $m - r$ samples in $\omega_m \setminus \omega_r$ satisfy $f(x^*(\omega_r), \delta) \le 0$
   and the remaining $m - q$ samples satisfy $f(x^*(\omega_r), \delta) > 0$

   Hence Lem. 3 follows from 1 and the law of total probability

Posterior confidence bounds (contd.)

Proof of Thm. 2 (sketch):

- $$\mathbb{P}^m\left\{\theta_r(\omega_m) = q \cap V\big(x^*(\omega_r)\big) \leq \epsilon \;\big|\; |\mathrm{Es}(\omega_r)| = k\right\} = \binom{m-r}{q-r} \frac{B(\epsilon; m-q+k, q-k+1)}{B(k, r-k+1)}$$

- Hence from Lem. 3 and Bayes' law

$$\mathbb{P}^m\left\{V\big(x^*(\omega_r)\big) \leq \epsilon \;\big|\; \theta_r(\omega_m) = q \cap |\mathrm{Es}(\omega_r)| = k\right\} = \Phi(q-k; m, 1-\epsilon)$$

so the total law of probability implies

$$\mathbb{P}^m\left\{V\big(x^*(\omega_r)\big) \leq \epsilon \;\big|\; \theta_r(\omega_m) = q\right\} \geq \Phi(q-\bar{\zeta}; m, 1-\epsilon) \sum_{k=\underline{\zeta}}^{\bar{\zeta}} \mathbb{P}^m\left\{|\mathrm{Es}(\omega_r)| = k\right\}$$

$$\mathbb{P}^m\left\{V\big(x^*(\omega_r)\big) \leq \epsilon \;\big|\; \theta_r(\omega_m) = q\right\} \leq \Phi(q-\underline{\zeta}; m, 1-\epsilon) \sum_{k=\underline{\zeta}}^{\min\{r,\bar{\zeta}\}} \mathbb{P}^m\left\{|\mathrm{Es}(\omega_r)| = k\right\}$$

Posterior confidence bounds (contd.)

Proof of Thm. 2 (sketch):

- $$\mathbb{P}^m\{\theta_r(\omega_m) = q \cap V\big(x^*(\omega_r)\big) \le \epsilon \;\big|\; |\mathrm{Es}(\omega_r)| = k\} = \binom{m-r}{q-r}\frac{B(\epsilon; m-q+k, q-k+1)}{B(k, r-k+1)}$$

- Hence from Lem. 3 and Bayes' law

$$\mathbb{P}^m\{V\big(x^*(\omega_r)\big) \le \epsilon \;\big|\; \theta_r(\omega_m) = q \cap |\mathrm{Es}(\omega_r)| = k\} = \Phi(q-k; m, 1-\epsilon)$$

so the total law of probability implies

$$\mathbb{P}^m\{V\big(x^*(\omega_r)\big) \le \epsilon \;\big|\; \theta_r(\omega_m) = q\} \ge \Phi(q-\bar{\zeta}; m, 1-\epsilon)\sum_{k=\underline{\zeta}}^{\bar{\zeta}} \mathbb{P}^m\{|\mathrm{Es}(\omega_r)| = k\}$$

$$\mathbb{P}^m\{V\big(x^*(\omega_r)\big) \le \epsilon \;\big|\; \theta_r(\omega_m) = q\} \le \Phi(q-\underline{\zeta}; m, 1-\epsilon)\sum_{k=\underline{\zeta}}^{\min\{r,\bar{\zeta}\}} \mathbb{P}^m\{|\mathrm{Es}(\omega_r)| = k\}$$

Posterior confidence bounds (contd.)

Proof of Thm. 2 (sketch):

- $$\mathbb{P}^m\{\theta_r(\omega_m) = q \cap V\big(x^*(\omega_r)\big) \le \epsilon \;\big|\; |\mathrm{Es}(\omega_r)| = k\}$$
  $$= \binom{m-r}{q-r} \frac{B(\epsilon; m-q+k, q-k+1)}{B(k, r-k+1)}$$

- Hence from Lem. 3 and Bayes' law
  $$\mathbb{P}^m\{V\big(x^*(\omega_r)\big) \le \epsilon \;\big|\; \theta_r(\omega_m) = q \cap |\mathrm{Es}(\omega_r)| = k\} = \Phi(q-k; m, 1-\epsilon)$$
  so the total law of probability implies
  $$\mathbb{P}^m\{V\big(x^*(\omega_r)\big) \le \epsilon \;\big|\; \theta_r(\omega_m) = q\}$$
  $$\ge \Phi(q-\bar{\zeta}; m, 1-\epsilon) \sum_{k=\underline{\zeta}}^{\bar{\zeta}} \mathbb{P}^m\{|\mathrm{Es}(\omega_r)| = k\}$$
  $$\mathbb{P}^m\{V\big(x^*(\omega_r)\big) \le \epsilon \;\big|\; \theta_r(\omega_m) = q\}$$
  $$\le \Phi(q-\underline{\zeta}; m, 1-\epsilon) \sum_{k=\underline{\zeta}}^{\min\{r,\bar{\zeta}\}} \mathbb{P}^m\{|\mathrm{Es}(\omega_r)| = k\}$$

Outline

Parameters and target confidence bounds

Consider the design of an algorithm to generate

$\hat{x}$: an approximate CCP solution

$\hat{q}$: the number of constraints satisfied out of m sampled constraints

Impose prior confidence bounds:

$$V(\hat{x}) \in (\underline{\epsilon}, \bar{\epsilon}] \text{ with probability } p_{\textit{prior}}$$

Impose posterior confidence bounds, given the value of $\hat{q}$:

$$|V(\hat{x}) - \hat{q}/m| \leq \Delta\epsilon \text{ with probability } p_{\textit{post}}$$

Parameters and target confidence bounds

$\hat{q}$ cannot be specified directly, but, for given m and $\underline{q}, \bar{q}$:

- ▷ Lem. 3 gives the probability, p_{trial}, of $\theta_r(\omega_m) \in [\underline{q}, \bar{q}]$, for given r
- ▷ hence SP(ω_r) must be solved N_{trial} times
  in order to ensure that $\hat{q} \in [\underline{q}, \bar{q}]$ with a given prior probability
- ▷ choose $r = r^*$, the maximizer of p_{trial} (possibly subject to $r \leq r_{\max}$)
  in order to minimize N_{trial}

Selection of m and $\underline{q}, \bar{q}$:

- ▷ choose m so that posterior confidence bounds are met:

$$\Phi\big(m(1-\bar{\epsilon}) - \bar{\zeta}; m, 1 - \bar{\epsilon} - \Delta\epsilon/2\big) \geq \tfrac{1}{2}(1 + p_{post})$$
$$\Phi\big(m(1-\bar{\epsilon}) - \underline{\zeta}; m, 1 - \bar{\epsilon} + \Delta\epsilon/2\big) \leq \tfrac{1}{2}(1 - p_{post}).$$

- ▷ Thm. 2 allows $\underline{q}, \bar{q}$ to be chosen so that prior confidence bounds hold

Randomized algorithm with tight confidence bounds

Algorithm:

(i). Given bounds $\underline{\epsilon}$, $\bar{\epsilon}$, probabilities $p_{prior} < p_{post}$, and sample size m, determine

$$\underline{q} = \min\ q \quad \text{subject to} \quad \Phi(q - \bar{\zeta}; m, 1 - \bar{\epsilon}) \geq \tfrac{1}{2}(1 + p_{post})$$

$$\bar{q} = \max\ q \quad \text{subject to} \quad \Phi(q - \underline{\zeta}; m, 1 - \underline{\epsilon}) \leq \tfrac{1}{2}(1 - p_{post})$$

$$r^* = \arg\max_r \sum_{q=\underline{q}}^{\bar{q}} \binom{m-r}{q-r} \min_{\zeta \in [\underline{\zeta}, \bar{\zeta}]} \frac{B(m-q+\zeta, q-\zeta+1)}{B(\zeta, r-\zeta+1)}$$

$$p_{trial} = \sum_{q=\underline{q}}^{\bar{q}} \binom{m-r^*}{q-r^*} \min_{\zeta \in [\underline{\zeta}, \bar{\zeta}]} \frac{B(m-q+\zeta, q-\zeta+1)}{B(\zeta, r^*-\zeta+1)}$$

$$N_{trial} = \left\lceil \frac{\ln(1 - p_{prior}/p_{post})}{\ln(1 - p_{trial})} \right\rceil$$

Randomized algorithm with tight confidence bounds

Algorithm (cont.):

(ii). Draw N_{trial} multisamples, $\omega_m^{(i)}$, and for each $i = 1, \ldots, N_{trial}$:

- compute $x^*(\omega_{r^*}^{(i)})$
- count $\theta_{r^*}(\omega_m^{(i)}) = \left|\left\{\delta \in \omega_m^{(i)} : f\big(x^*(\omega_{r^*}^{(i)}), \delta\big) \leq 0\right\}\right|$

(iii). Determine $i^* \in \{1, \ldots, N_{trial}\}$:

$$i^* = \arg\min_{i \in \{1,\ldots,N_{trial}\}} \left|\tfrac{1}{2}(\bar{q} + \underline{q}) - \theta_{r^*}(\omega_m^{(i)})\right|,$$

and return $\begin{cases} \text{the solution estimate, } \hat{x} = x^*(\omega_{r^*}^{(i^*)}) \\ \text{the number of satisfied constraints, } \hat{q} = \theta_{r^*}(\omega_m^{(i^*)}) \end{cases}$

Randomized algorithm with tight confidence bounds

Proposition 3: $\hat{x}$, $\hat{q}$ satisfy the prior confidence bounds

$$\mathbb{P}^{N_{trial}\, m}\big\{V(\hat{x}) \in (\underline{\epsilon}, \bar{\epsilon}\,]\big\} \geq p_{prior}$$

and the posterior confidence bounds

$$\Phi(q - \bar{\zeta}; m, 1-\epsilon) \leq \mathbb{P}^{N_{trial}\, m}\big\{V(\hat{x}) \leq \epsilon \mid \hat{q} = q\big\} \leq \Phi(q - \underline{\zeta}; m, 1-\epsilon).$$

- The posterior bounds follow directly from Thm. 2
- The choice of $\underline{q}, \bar{q}$ gives $\mathbb{P}^{m}\big\{V(\hat{x}) \in (\underline{\epsilon}, \bar{\epsilon}\,] \mid \hat{q} = q\big\} \geq p_{post}\ \forall q \in [\underline{q}, \bar{q}]$,

  $$\implies \mathbb{P}^{N_{trial}\, m}\big\{V(\hat{x}) \in (\underline{\epsilon}, \bar{\epsilon}\,]\big\} \geq p_{post}\, \mathbb{P}^{N_{trial}\, m}\big\{\hat{q} \in [\underline{q}, \bar{q}]\big\}$$

  but the definition of i^* implies $\mathbb{P}^{N_{trial}\, m}\big\{\hat{q} \in [\underline{q}, \bar{q}]\big\} \geq 1 - (1 - p_{trial})^{N_{trial}}$

  $$\implies \mathbb{P}^{N_{trial}\, m}\big\{V(\hat{x}) \in (\underline{\epsilon}, \bar{\epsilon}\,]\big\} \geq p_{post}\big(1 - (1 - p_{trial})^{N_{trial}}\big)$$

  so the choice of N_{trial} ensures the prior bound

Randomized algorithm with tight confidence bounds

- Large m is computationally cheap (hence small $\Delta\epsilon$ and $p_{post} \approx 1$)
- p_{trial} is a tight bound on the probability of $\hat{q} \in [\underline{q}, \bar{q}]$
  and is exact if $\underline{\zeta} = \bar{\zeta}$
- The posterior confidence bounds are tight for all distributions
  and are exact if $\underline{\zeta} = \bar{\zeta}$
- Repeated trials and empirical violation tests have been used before
  e.g. Calafiore (2017), Chamanbaz et al. (2016)
- The lower posterior confidence bound derived here coincides with Calafiore (2017) for fully supported problems ($\underline{\zeta} = \bar{\zeta} = n$)

  but our approach:

  - shows that this bound is exact for fully supported problems
  - provides tight bounds in all other cases

Randomized algorithm with tight confidence bounds

Variation of r^* and N_{trial} with $\underline{\zeta}, \bar{\zeta}$

for $m = 10^5$, $\underline{\epsilon} = 0.19$, $\bar{\epsilon} = 0.21$, $\Delta\epsilon = 0.005$, and $p_{post} = 0.5(1 + p_{prior})$:

$(\underline{\zeta}, \bar{\zeta})$	$(2,5)$				$(7,10)$				$(17,20)$				$(47,50)$				$(97,100)$			
p_{prior}	p_1	p_2	p_3	p_4	p_1	p_2	p_3	p_4	p_1	p_2	p_3	p_4	p_1	p_2	p_3	p_4	p_1	p_2	p_3	p_4
r^*	15				40				91				241				492			
N_{trial}	84	109	176	291	37	48	77	128	22	29	46	76	13	16	26	43	8	11	17	29

$(\underline{\zeta}, \bar{\zeta})$	$(1,2)$				$(1,5)$				$(1,10)$			
p_{prior}	p_1	p_2	p_3	p_4	p_1	p_2	p_3	p_4	p_1	p_2	p_3	p_4
r^*	5				12				22			
N_{trial}	96	125	200	331	189	246	396	655	1022	1329	2116	3465

where $p_1 = 0.9$, $p_2 = 0.95$, $p_3 = 0.99$, $p_4 = 0.999$

Example I: minimal bounding hypersphere

Find smallest hypersphere in $\mathbb{R}^4$ containing $\delta \sim \mathcal{N}(0, I)$ with probability 0.8

- Support dimension: $\underline{\zeta} = 2$, $\bar{\zeta} = 5$
- Prior bounds:
  $$V(\hat{x}) \in (0.19, 0.21] \text{ with probability } p_{prior} = 0.9$$
- Posterior bounds:
  $$|V(\hat{x}) - \hat{q}/m| \leq 0.005 \text{ with probability } p_{post} = 0.95$$
- Parameters: $m = 10^5$, $\underline{q} = 79257$, $\bar{q} = 80759$
  $r^* = 15$, $p_{trial} = 0.0365$, $N_{trial} = 84$

Compare SP solved approximately using greedy discarding (assuming one constraint discarded per iteration), with $m = 420$, $q = 336$:

$$V(x^*(\omega^*)) \in (0.173, 0.332] \text{ with probability } 0.9$$

Example I: minimal bounding hypersphere

Find smallest hypersphere in $\mathbb{R}^4$ containing $\delta \sim \mathcal{N}(0, I)$ with probability 0.8

- Support dimension: $\underline{\zeta} = 2$, $\bar{\zeta} = 5$
- Prior bounds:
$$V(\hat{x}) \in (0.19, 0.21] \text{ with probability } p_{\mathit{prior}} = 0.9$$
- Posterior bounds:
$$|V(\hat{x}) - \hat{q}/m| \leq 0.005 \text{ with probability } p_{\mathit{post}} = 0.95$$
- Parameters: $m = 10^5$, $\underline{q} = 79257$, $\bar{q} = 80759$
$r^* = 15$, $p_{\mathit{trial}} = 0.0365$, $N_{\mathit{trial}} = 84$

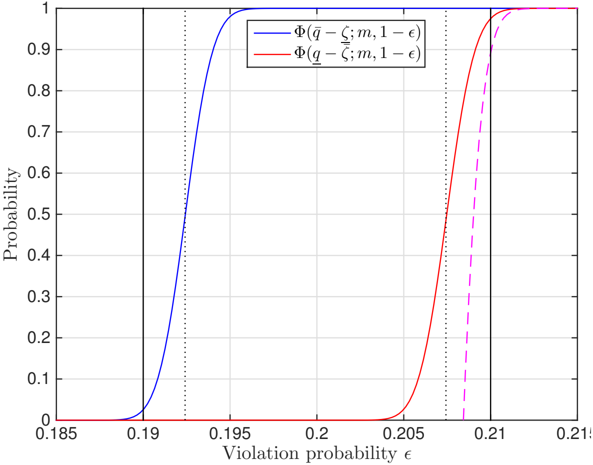
Compare SP solved approximately using greedy discarding (assuming one constraint discarded per iteration), with $m = 420$, $q = 336$:

$$V\big(x^*(\omega^*)\big) \in (0.173, 0.332] \text{ with probability } 0.9$$

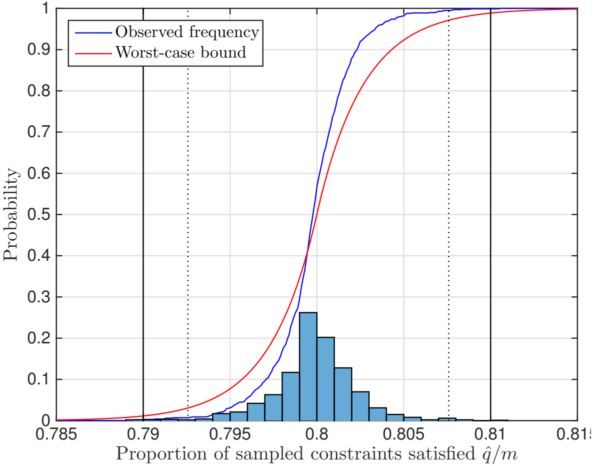
Example I: minimal bounding hypersphere

Posterior confidence bounds given $\hat{q} \in [\underline{q}, \bar{q}]$:

$\mathbb{P}\{V(\hat{x}) \leq \epsilon \mid \hat{q} = \bar{q}\}$ (blue) $\mathbb{P}\{V(\hat{x}) \leq \epsilon \mid \hat{q} = \underline{q}\}$ (red)



Example I: minimal bounding hypersphere



- Empirical probability distribution of $\hat{q}$ (blue) for 1000 repetitions
- Worst case bound F (red), $\mathbb{P}\{|\hat{q}/m - 0.8| \leq t\} \geq F(0.8+t) - F(0.8-t)$

Example II: finite horizon robust control design

Control problem (from Calafiore 2017):

Determine a control sequence to drive the system state to a target over a finite horizon with high probability despite model uncertainty

Discrete-time uncertain linear system

$$z(t+1) = A(\delta)z(t) + Bu(t), \qquad A(\delta) = A_0 + \sum_{i,j=1}^{n_z} \delta_{i,j} e_i e_j^\top$$

$\delta_{i,j} \in [-\rho, \rho]$ uniformly distributed

Example II: finite horizon robust control design

Control problem (from Calafiore 2017):

Determine a control sequence to drive the system state to a target over a finite horizon with high probability despite model uncertainty

Discrete-time uncertain linear system

$$z(t+1) = A(\delta)z(t) + Bu(t), \qquad A(\delta) = A_0 + \sum_{i,j=1}^{n_z} \delta_{i,j} e_i e_j^\top$$

$\delta_{i,j} \in [-\rho, \rho]$ uniformly distributed

Example II: finite horizon robust control design

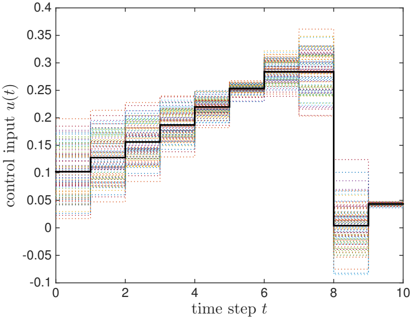
Control objective:

$$\mathsf{CCP}: \quad \underset{\gamma, u}{\text{minimize}} \ \gamma$$
$$\text{subject to} \ \mathbb{P}\{\|z_T - z(T,\delta)\|^2 + \lambda\|u\|^2 > \gamma\} \leq \epsilon$$

with $u = \{u(0), \ldots, u(T-1)\}$: predicted control sequence

$$z(T,\delta) := A(\delta)^T z(0) + \begin{bmatrix} A(\delta)^{T-1}B & \cdots & B \end{bmatrix} u$$

z_T: target state



Example II: finite horizon robust control design

- Problem dimensions: $n_z = 6,\ T = 10,\ n = 11,\ \underline{\zeta} = 1,\ \bar{\zeta} = 3$
  and parameters: $\rho = 10^{-3},\ \lambda = 0.005$

- Single-sided prior probability bounds with $\bar{\epsilon} = 0.005$:

$$\mathbb{P}\{V(x) \in (0, \bar{\epsilon}]\} \geq p_{\mathit{prior}} = 0.99$$

Posterior uncertainty: $\Delta\epsilon = 0.003$
and probability: $p_{\mathit{post}} = 1 - 10^{-12}$

$$\left.\begin{array}{l} \Delta\epsilon = 0.003 \\ p_{\mathit{post}} = 1 - 10^{-12} \end{array}\right\} \implies m = 63 \times 10^3$$

- Choose number of sampled constraints in each trial to maximize p_{trial}:

$$r^* = 2922,\ p_{\mathit{trial}} = 0.482,\ N_{\mathit{trial}} = 9$$

if r^* is limited to 2000, we get $p_{\mathit{trial}} = 0.414,\ N_{\mathit{trial}} = 9$

Example II: finite horizon robust control design

Compare number of trials needed for $\hat{q} \geq \underline{q}$

expected using bounds of Calfiore (2017):	9.6
expected using proposed approach $(1/p_{\mathit{trial}})$:	2.4
observed:	1.4

Residual conservativism is due to use of bounds: $\underline{\zeta} = 1$, $\bar{\zeta} = 3$ rather than the actual distribution for ζ:

support dimension:	1	2	3
observed frequency:	5%	59%	36%

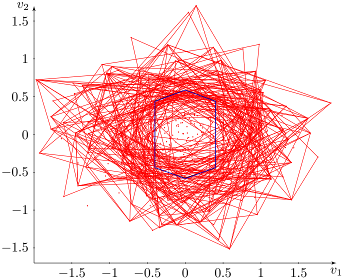
Future directions

- ▷ Alternative methods for randomly selecting level-q solutions

- ▷ Methods of checking feasibility of suboptimal points

- ▷ Estimation of violation probability over a set of parameters

  application to computing invariant sets for constrained dynamics



- ▷ Application to chance-constrained energy management in PHEVs

Conclusions

- Solutions of sampled convex optimization problems with randomised sample selection strategies have tighter confidence bounds than optimal sample discarding strategies
- Randomised sample selection can be performed by solving a robust sampled optimization and counting violations of additional sampled constraints
- We propose a randomised algorithm for chance-constrained convex programming with tight *a priori* and *a posteriori* confidence bounds

Thank you for your attention!

Acknowledgements: James Fleming, Matthias Lorenzen,
Johannes Buerger, Rainer Manuel Schaich . . .

Research directions

- Optimal control of uncertain systems
  - chance-constrained optimization
  - stochastic predictive control
- Robust control design
  - invariant sets and uncertainty parameterization
- Applications:
  - energy management in hybrid electric vehicles (BMW)
  - ecosystem-based fisheries management (BAS)