CALTECH PSAAP Predictive Science Academic Alliance Program

Bridging scales with incomplete information
Optimal Uncertainty Quantification

Houman Owhadi

Clint Scovel, Tim Sullivan
Mike McKerns, Michael Ortiz

National Nuclear Security Administration

The UQ challenge in the certification context

$$G : \chi \longrightarrow \mathbb{R}$$

$$X \longrightarrow G(X)$$

$$\mathbb{P} \in \mathcal{M}(\chi)$$

You want to certify that

$$\mathbb{P}[G(X) \geq a] \leq \epsilon$$

Problem

- You don't know G.

and

- You don't know $\mathbb{P}$

The UQ challenge in the certification context

$$G : \chi \longrightarrow \mathbb{R}$$

$$X \longrightarrow G(X)$$

$$\mathbb{P} \in \mathcal{M}(\chi)$$

You want to certify that

$$\mathbb{P}[G(X) \geq a] \leq \epsilon$$

You only know

$$(G, \mathbb{P}) \in \mathcal{A}$$

$$\mathcal{A} \subset \left\{ (f, \mu) \,\middle|\, \begin{array}{c} f : \mathcal{X} \to \mathbb{R}, \\ \mu \in \mathcal{P}(\mathcal{X}) \end{array} \right\}$$

Optimal bounds on $\mathbb{P}[G(X) \geq a]$

$$\mathcal{U}(\mathcal{A}) := \sup_{(f,\mu)\in\mathcal{A}} \mu[f(X) \geq a]$$

$$\mathcal{L}(\mathcal{A}) := \inf_{(f,\mu)\in\mathcal{A}} \mu[f(X) \geq a]$$

$$\mathcal{L}(\mathcal{A}) \leq \mathbb{P}[G(X) \geq a] \leq \mathcal{U}(\mathcal{A})$$

$\mathcal{U}(\mathcal{A}) \leq \epsilon$: Safe even in worst case.

$\epsilon < \mathcal{L}(\mathcal{A})$: Unsafe even in best case.

$\mathcal{L}(\mathcal{A}) \leq \epsilon < \mathcal{U}(\mathcal{A})$: Cannot decide.
Unsafe due to lack of information.

Reduction of optimization variables

$$\{f : \mathcal{X} \to \mathbb{R},\ \mu \in \mathcal{P}(\mathcal{X})\}$$

$$\left\{ f : \mathcal{X} \to \mathbb{R},\ \mu \in \mathcal{P}(\mathcal{X}) \,\middle|\, \mu = \sum_{i=1}^{k} \alpha_k \delta_{x_k} \right\}$$

$$\{f : \{1, 2, \ldots, n\} \to \mathbb{R},\ \mu \in \mathcal{P}(\{1, 2, \ldots, n\})\}$$

$$\{\{1, 2, \ldots, q\},\ \mu \in \mathcal{P}(\{1, 2, \ldots, n\})\}$$

Application: Optimal concentration inequality

$$\mathcal{A}_{MD} := \left\{ (f,\mu) \;\middle|\; \begin{array}{c} f\colon \mathcal{X}_1 \times \cdots \times \mathcal{X}_m \to \mathbb{R}, \\ \mu \in \mathcal{M}(\mathcal{X}_1) \otimes \cdots \otimes \mathcal{M}(\mathcal{X}_m), \\ \mathbb{E}_\mu[f] \leq 0, \\ \mathrm{Osc}_i(f) \leq D_i \end{array} \right\}$$

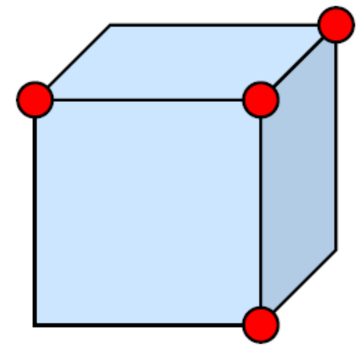
$$\mathrm{Osc}_i(f) := \sup_{(x_1,\dots,x_m)\in\mathcal{X}} \sup_{x_i'\in\mathcal{X}_i} \big(f(\dots,x_i,\dots) - f(\dots,x_i',\dots) \big).$$

$$\mathcal{U}(\mathcal{A}_{MD}) := \sup_{(f,\mu)\in\mathcal{A}_{MD}} \mu[f(X) \geq a]$$

McDiarmid inequality $\mathcal{U}(\mathcal{A}_{MD}) \leq \exp\left(-2 \frac{a^2}{\sum_{i=1}^m D_i^2} \right)$

Reduction of optimization variables

$$\mathcal{A}_{\mathcal{C}} := \left\{ (C,\alpha) \;\middle|\; \begin{array}{c} C \subset \{0,1\}^m, \\ \alpha \in \otimes_{i=1}^m \mathcal{M}(\{0,1\}), \\ \mathbb{E}_\alpha[h^C] \le 0 \end{array} \right\}$$



$$h^C : \{0,1\}^m \longrightarrow \mathbb{R}$$

$$t \longrightarrow a - \min_{s \in C} \sum_{i:\, s_i \neq t_i} D_i$$

$$\mathcal{U}(\mathcal{A}_{\mathcal{C}}) := \sup_{(C,\alpha)\in\mathcal{A}_{\mathcal{C}}} \alpha[h^C \ge a]$$

Theorem

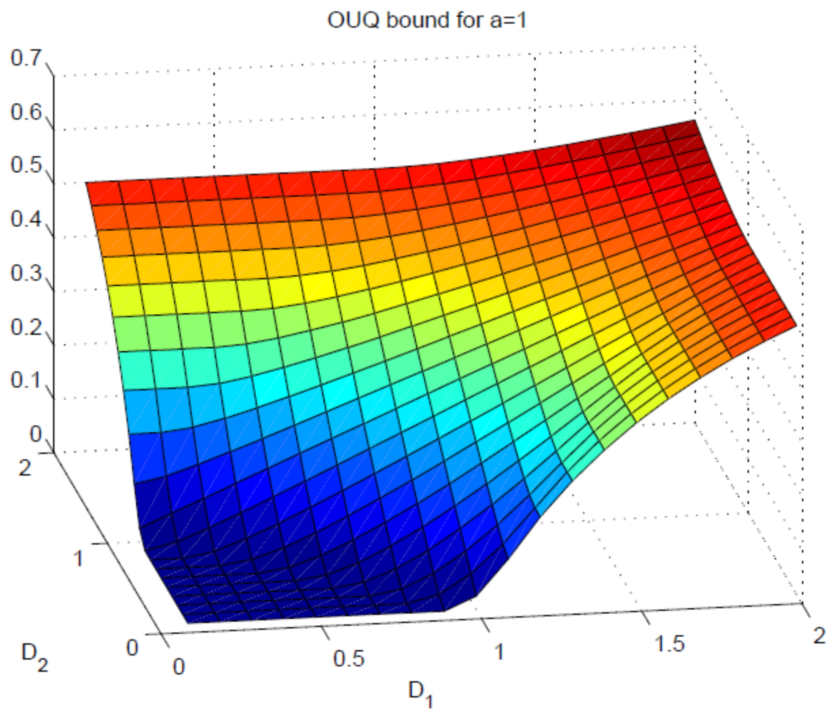
$$\mathcal{U}(\mathcal{A}_{MD}) = \mathcal{U}(\mathcal{A}_{\mathcal{C}})$$

Explicit Solution m=2

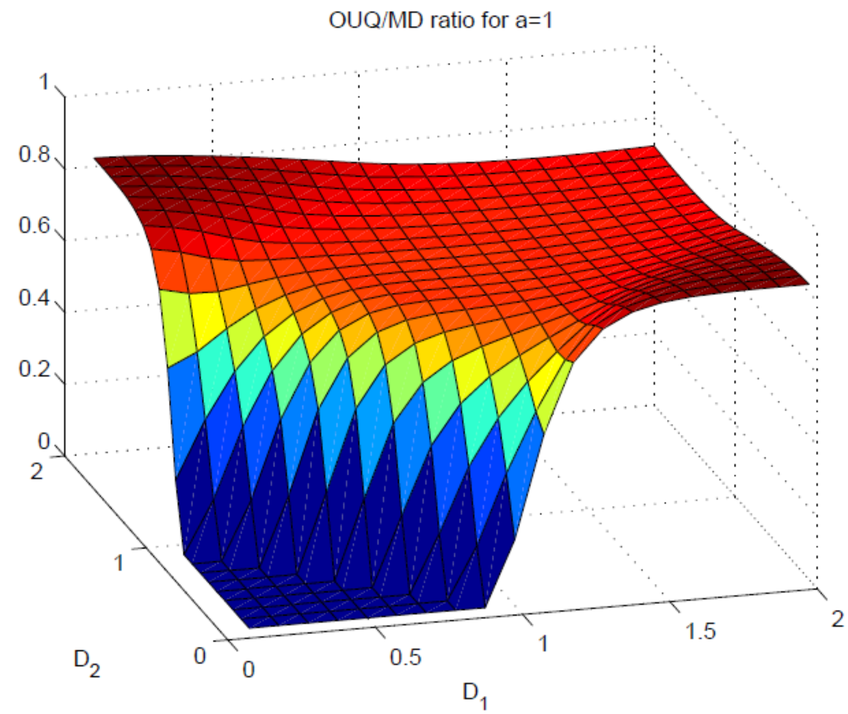
Theorem $m = 2$

$$\mathcal{U}(\mathcal{A}_{MD}) = \begin{cases} 0 & \text{if} \quad D_1 + D_2 \leq a \\ \frac{(D_1+D_2-a)^2}{4D_1D_2} & \text{if} \quad |D_1 - D_2| \leq a \leq D_1 + D_2 \\ 1 - \frac{a}{\max(D_1,D_2)} & \text{if} \quad 0 \leq a \leq |D_1 - D_2| \end{cases}$$

OUQ bound a=1



OUQ/MD a=1



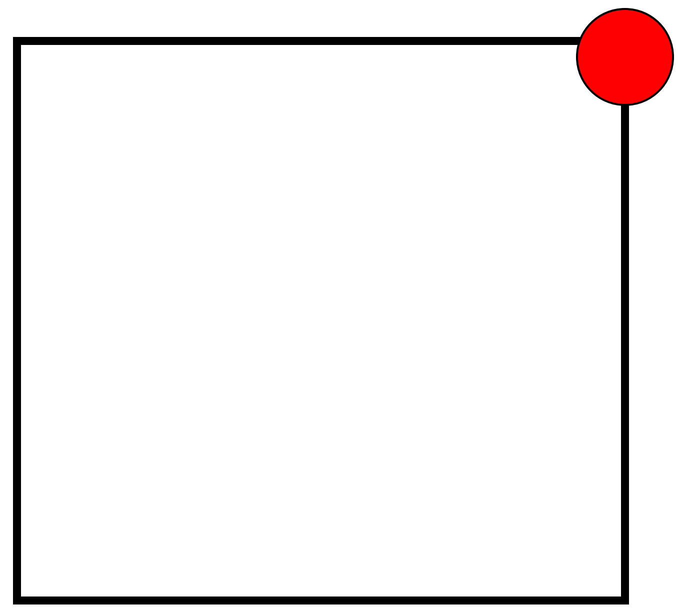
Explicit Solution m=2

Theorem $m = 2$

$$\mathcal{U}(\mathcal{A}_{MD}) = \begin{cases} 0 & \text{if} \quad D_1 + D_2 \leq a \\ \frac{(D_1+D_2-a)^2}{4D_1D_2} & \text{if} \quad |D_1 - D_2| \leq a \leq D_1 + D_2 \\ 1 - \frac{a}{\max(D_1,D_2)} & \text{if} \quad 0 \leq a \leq |D_1 - D_2| \end{cases}$$

$$C = \{(1,1)\}$$

$$h^C(s) = a - (1 - s_1)D_1 - (1 - s_2)D_2$$



Optimal Hoeffding= Optimal McDiarmid for m=2

$$\mathcal{A}_{MD} := \left\{ (f,\mu) \;\middle|\; \begin{array}{c} f\colon \mathcal{X}_1 \times \cdots \times \mathcal{X}_m \to \mathbb{R}, \\ \mu \in \mathcal{M}(\mathcal{X}_1) \otimes \cdots \otimes \mathcal{M}(\mathcal{X}_m), \\ \mathbb{E}_\mu[f] \le 0, \\ \mathrm{Osc}_i(f) \le D_i \end{array} \right\}$$

$$\mathcal{U}(\mathcal{A}_{\mathrm{MD}}) = \mathcal{U}(\mathcal{A}_{\mathrm{Hfd}})$$

$$\mathcal{A}_{\mathrm{Hfd}} := \left\{ (f,\mu) \;\middle|\; \begin{array}{c} f = X_1 + \cdots + X_m, \\ \mu \in \bigotimes_{i=1}^{m} \mathcal{M}([b_i - D_i, b_i]), \\ \mathbb{E}_\mu[f] \le 0 \end{array} \right\}$$

Explicit Solution m=2

Theorem $m = 2$

$$\mathcal{U}(\mathcal{A}_{MD}) = \begin{cases} 0 & \text{if} \quad D_1 + D_2 \leq a \\ \frac{(D_1+D_2-a)^2}{4D_1D_2} & \text{if} \quad |D_1 - D_2| \leq a \leq D_1 + D_2 \\ 1 - \frac{a}{\max(D_1,D_2)} & \text{if} \quad 0 \leq a \leq |D_1 - D_2| \end{cases}$$

Corollary If $D_1 \geq a + D_2$, then

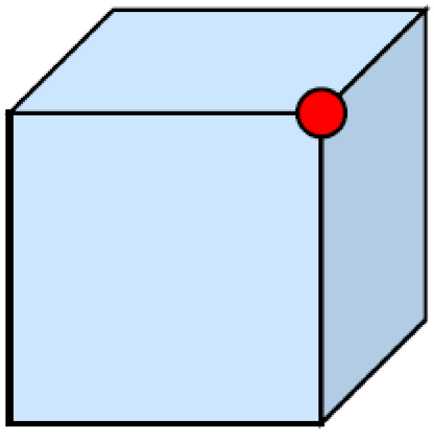
$$\mathcal{U}(\mathcal{A}_{MD})(a, D_1, D_2) = \mathcal{U}(\mathcal{A}_{MD})(a, D_1, 0)$$

Explicit Solution m=3

Theorem $m = 3 \qquad D_1 \geq D_2 \geq D_3$

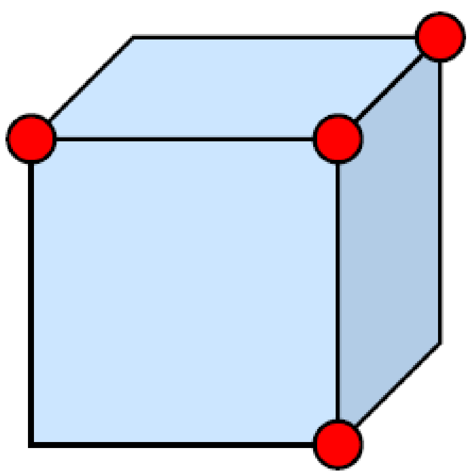
$$\mathcal{U}(\mathcal{A}_{MD}) = \max(\mathcal{F}_1, \mathcal{F}_2)$$

$\mathcal{F}_1$



$$\mathcal{U}(\mathcal{A}_{\mathrm{MD}}) = \mathcal{U}(\mathcal{A}_{\mathrm{Hfd}})$$

$\mathcal{F}_2$



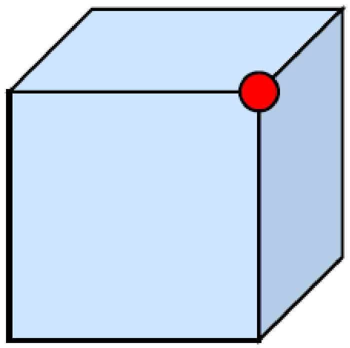
$$\mathcal{U}(\mathcal{A}_{\mathrm{MD}}) > \mathcal{U}(\mathcal{A}_{\mathrm{Hfd}})$$

Explicit Solution m=3

Theorem $m = 3 \qquad D_1 \geq D_2 \geq D_3$

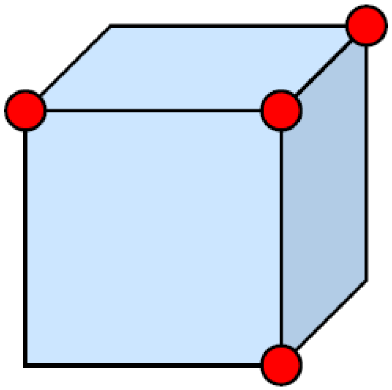
$$\mathcal{U}(\mathcal{A}_{MD}) = \max(\mathcal{F}_1, \mathcal{F}_2)$$

$\mathcal{F}_1$



$$\mathcal{F}_1 := \begin{cases} 0 & \text{if} \quad D_1 + D_2 + D_3 \leq a \\ \frac{(D_1+D_2+D_3-a)^3}{27D_1D_2D_3} & \text{if} \quad D_1 + D_2 - 2D_3 \leq a \leq D_1 + D_2 + D_3 \\ \frac{(D_1+D_2-a)^2}{4D_1D_2} & \text{if} \quad D_1 - D_2 \leq a \leq D_1 + D_2 - 2D_3 \\ 1 - \frac{a}{\max(D_1,D_2)} & \text{if} \quad 0 \leq a \leq D_1 - D_2 \end{cases}$$

$\mathcal{F}_2$



$$\mathcal{F}_2 := \max_{i \in \{1,2,3\}} \phi(\gamma_i)\psi(\gamma_i)$$

$$(1+\gamma)^3 - \frac{5D_2 - 2D_3}{2D_2 - D_3}(1+\gamma)^2 + \frac{4D_2 - a}{2D_2 - D_3} = 0,$$

$$\mathcal{F}_2 := \max_{i\in\{1,2,3\}} \phi(\gamma_i)\psi(\gamma_i)$$

$$\psi(\gamma) := \gamma^2\left(2\frac{D_2}{D_3}-1\right) - 2\gamma\left(3\frac{D_2}{D_3}-1\right) + \frac{\gamma}{1+\gamma}\left(8\frac{D_2}{D_3}-2\frac{a}{D_3}\right)$$

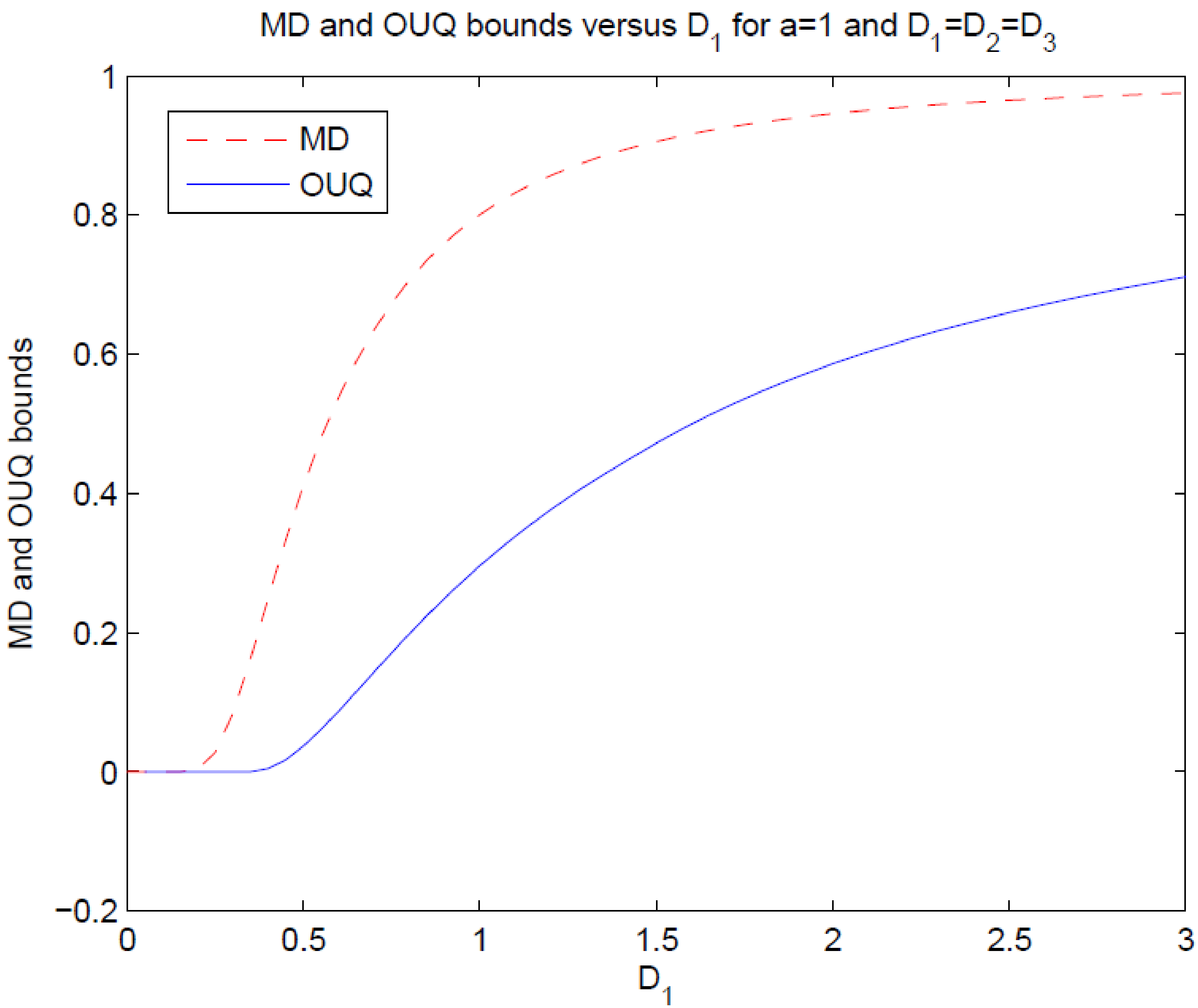
$$\phi(\gamma) := \begin{cases} 1, & \text{if } \gamma\in(0,1) \text{ and } \theta(\gamma)\in(0,1), \\ 0, & \text{otherwise,} \end{cases}$$

$$\theta(\gamma) := 1 - \frac{a}{D_3(1-\gamma^2)} + \frac{D_2}{D_3}\frac{1-\gamma}{1+\gamma}.$$

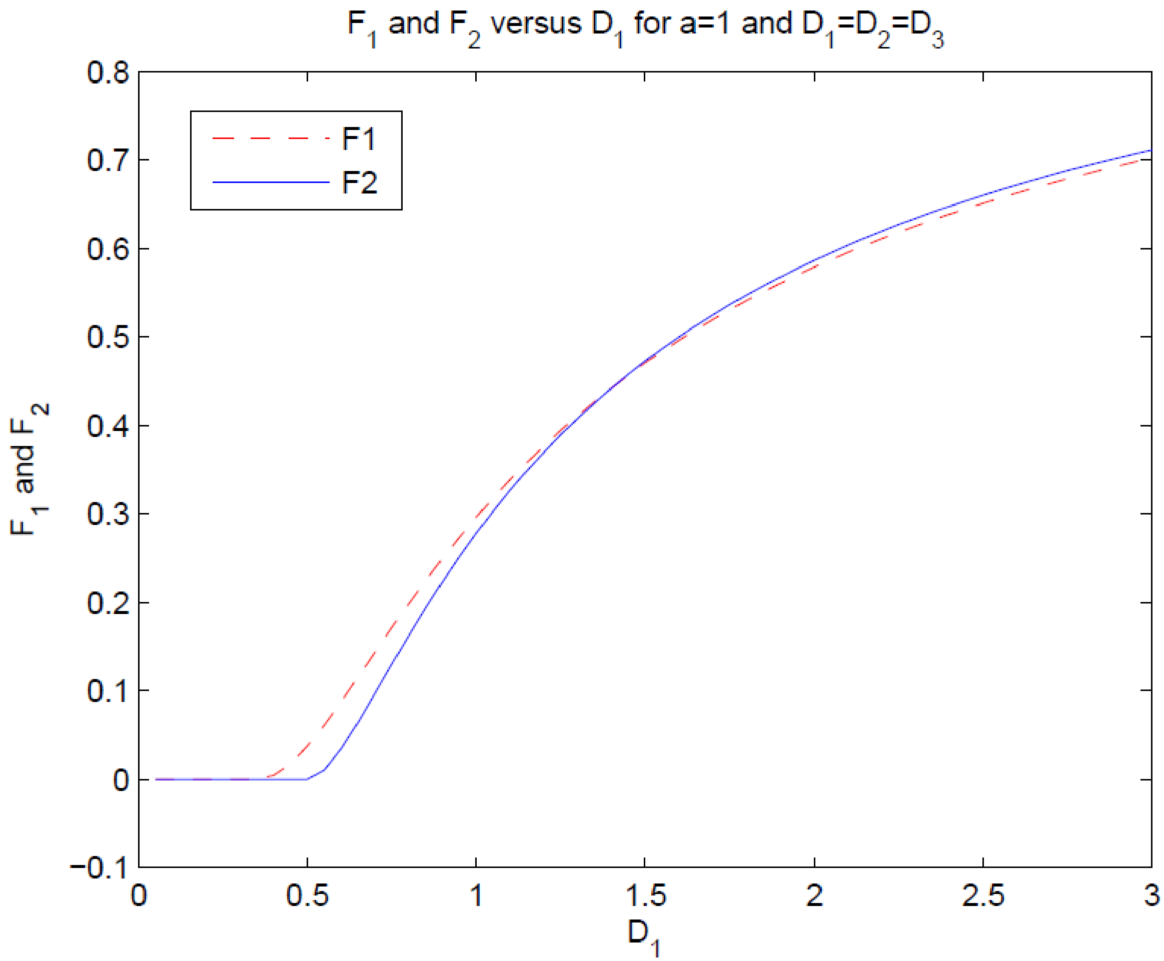
$\gamma_1, \gamma_2, \gamma_3$ are the roots of the cubic polynomial

$$(1+\gamma)^3 - \frac{5D_2-2D_3}{2D_2-D_3}(1+\gamma)^2 + \frac{4D_2-a}{2D_2-D_3} = 0,$$

OUQ vs McD m=3 $D_1 = D_2 = D_3$



a=1



$$m = 3$$

$\mathcal{F}_1$ and $\mathcal{F}_2$ vs D_1

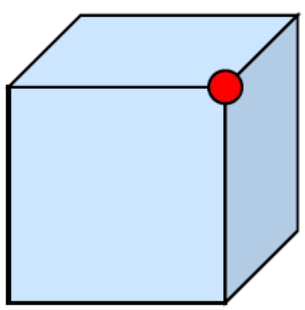
$$D_1 = D_2 = D_3$$

$\mathcal{F}_2 > \mathcal{F}_1$ for D_1 large enough

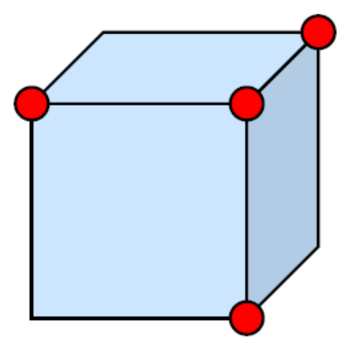
$$\mathcal{U}(\mathcal{A}_{\mathrm{MD}}) = \mathcal{U}(\mathcal{A}_{\mathrm{Hoeffding}})$$

$$\mathcal{U}(\mathcal{A}_{\mathrm{MD}}) > \mathcal{U}(\mathcal{A}_{\mathrm{Hoeffding}})$$

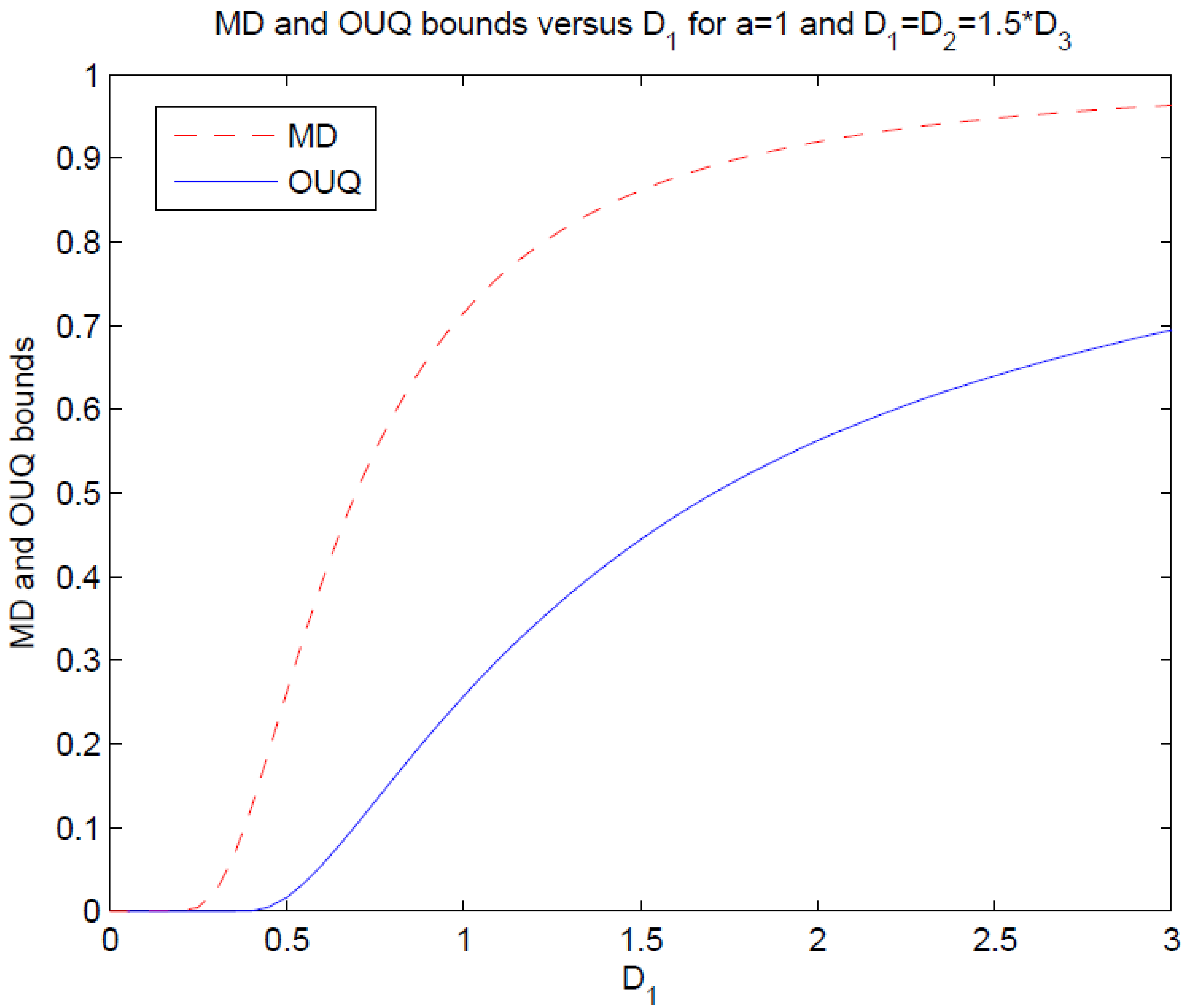
$\mathcal{F}_1$



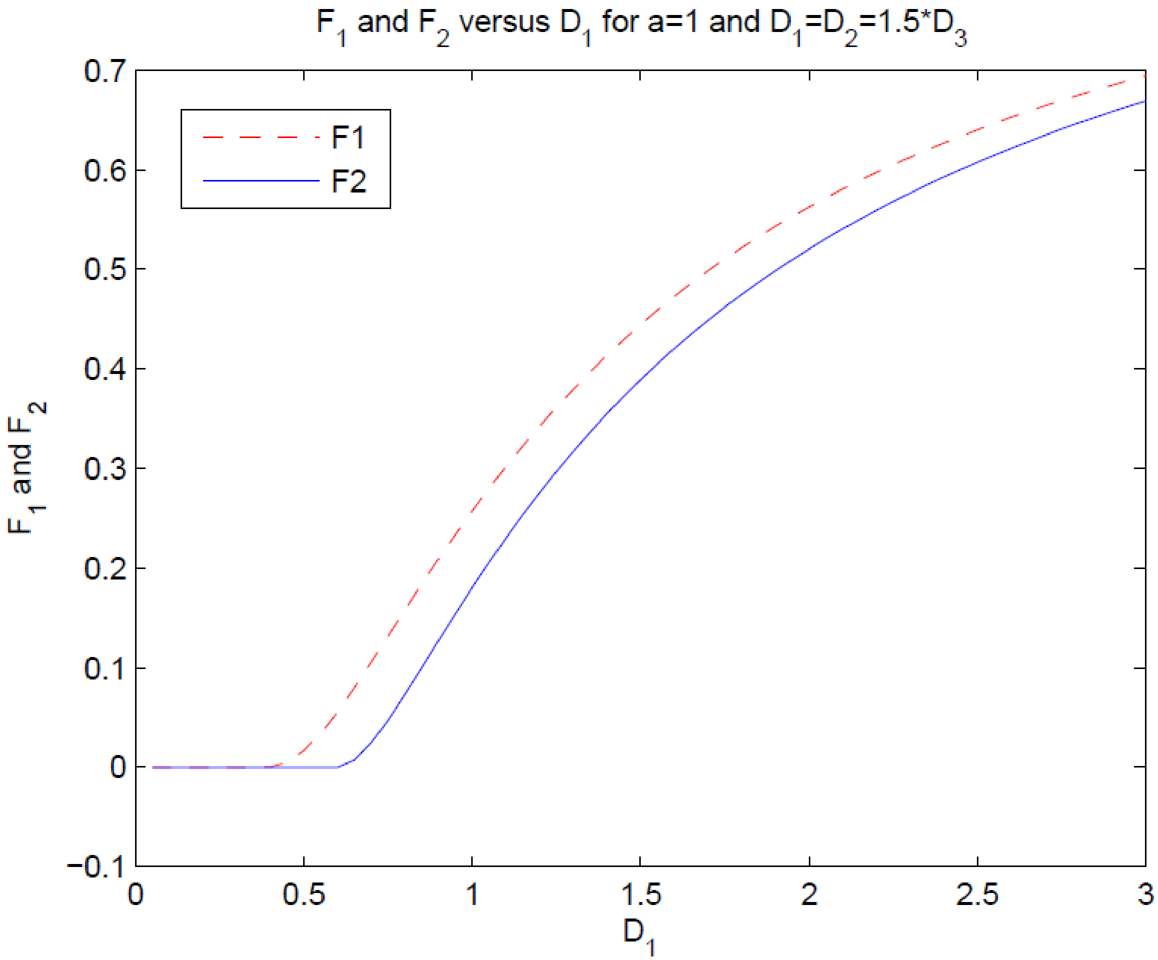
$\mathcal{F}_2$



OUQ vs McD m=3 $D_1 = D_2 = \frac{3}{2} D_3$



a=1



$$m = 3$$

$\mathcal{F}_1$ and $\mathcal{F}_2$ vs D_1

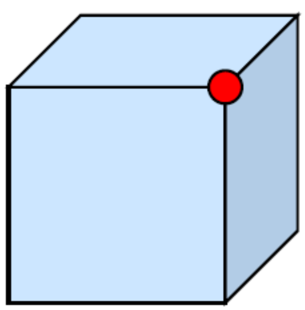
$$D_1 = D_2 = \frac{3}{2} D_3$$

$\mathcal{F}_2 < \mathcal{F}_1$ away from the diagonal

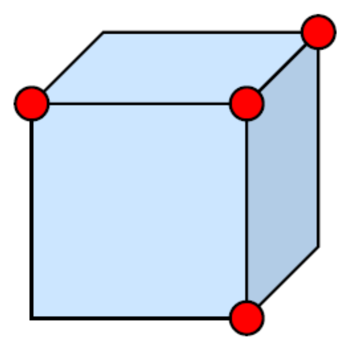
$$\mathcal{U}(\mathcal{A}_{\text{MD}}) = \mathcal{U}(\mathcal{A}_{\text{Hoeffding}})$$

$$\mathcal{U}(\mathcal{A}_{\text{MD}}) > \mathcal{U}(\mathcal{A}_{\text{Hoeffding}})$$

$\mathcal{F}_1$



$\mathcal{F}_2$



Dimension m

Theorem $D_1 \geq D_2 \geq \cdots \geq D_m$

$$a \geq \sum_{j=1}^{m-2} D_j + D_m$$

$$\mathcal{U}(\mathcal{A}_{MD}) = \begin{cases} 0, & \text{if } \sum_{j=1}^{m} D_j \leq a, \\ \dfrac{(\sum_{j=1}^{m} D_j - a)^m}{m^m \prod_{j=1}^{m} D_j}, & \text{if } \sum_{j=1}^{m} D_j - mD_m \leq a \leq \sum_{j=1}^{m} D_j, \\ \dfrac{(\sum_{j=1}^{k} D_j - a)^k}{k^k \prod_{j=1}^{k} D_j}, & \text{if for } k \in \{1, \ldots, m-1\} \\ & \sum_{j=1}^{k} D_j - kD_k \leq a \leq \sum_{j=1}^{k+1} D_j - (k+1)D_{k+1}. \end{cases}$$

Other cases

Direct computation with optimization variables in

$$\{1, \ldots, [(m+1)/2]\} \times [0,1]^m$$

Reduction theorems

χ_i: Suslin spaces.

$$\mathcal{A} = \left\{ (f,\mu) \;\middle|\; \begin{array}{c} f : \mathcal{X}_1 \times \cdots \times \mathcal{X}_m \to \mathbb{R}, \\ \mu = \mu_1 \otimes \cdots \otimes \mu_m, \\ \mathcal{G}(f,\mu) \leq 0 \end{array} \right\}$$

$$\mathcal{G}(f,\mu) \leq 0 \Leftrightarrow \begin{cases} \mathbb{E}_\mu[g_j^f(X_1,\ldots,X_m)] \leq 0 & 1 \leq j \leq n' \\ \mathbb{E}_{\mu_1}[g_j^{f,1}(X_1)] \leq 0 & 1 \leq j \leq n_1 \\ \cdots & \cdots \\ \mathbb{E}_{\mu_m}[g_j^{f,m}(X_m)] \leq 0 & 1 \leq j \leq n_m \end{cases}$$

$$\mathcal{U}(\mathcal{A}) := \sup_{(f,\mu)\in\mathcal{A}} \mathbb{E}_\mu[r_f]$$

Reduction to products of convex linear combinations of Dirac masses

$$\mathcal{A} = \left\{ (f, \mu) \;\middle|\; \begin{array}{c} f : \mathcal{X}_1 \times \cdots \times \mathcal{X}_m \to \mathbb{R}, \\ \mu = \mu_1 \otimes \cdots \otimes \mu_m, \\ \mathcal{G}(f, \mu) \leq 0 \end{array} \right\}$$

$$\mathcal{G}(f,\mu) \leq 0 \Leftrightarrow \begin{cases} n' \text{ generalized moment constraints on } \mu \\ n_k \text{ generalized moment constraints on } \mu_k \end{cases}$$

Theorem

$$\mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_\Delta)$$

$$\mathcal{A}_\Delta = \left\{ (f, \mu) \in \mathcal{A} \;\middle|\; \begin{array}{c} \mu_k \text{ is a sum of at most} \\ n' + n_k + 1 \text{ weighted} \\ \text{Dirac measures on } \chi_k \end{array} \right\}$$

Reduction to products of convex linear combinations of Dirac masses

For each f, let $r_f : \chi \to \mathbb{R}$ be integrable for each μ such that $\mathcal{G}(f,\mu)$ is well defined.

$$\mathcal{U}(\mathcal{A}) := \sup_{(f,\mu)\in\mathcal{A}} \mathbb{E}_\mu[r_f]$$

Theorem

$$\mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_\Delta)$$

$$\mathcal{A}_\Delta = \left\{ (f,\mu) \;\middle|\; \begin{matrix} f: \mathcal{X}_1 \times \cdots \times \mathcal{X}_m \to \mathbb{R}, \\ \mu \in \Delta_{n_1+n'}(\mathcal{X}_1) \otimes \cdots \otimes \Delta_{n_m+n'}(\mathcal{X}_m), \\ G(f,\mu) \le 0 \end{matrix} \right\}$$

$$\Delta_k(\mathcal{X}) := \left\{ \sum_{j=0}^{k} \alpha_j \delta_{x^j} \;\middle|\; x^j \in \mathcal{X},\ \alpha_j \ge 0,\ \sum_{j=0}^{k} \alpha_j = 1 \right\}$$

Application to McDiarmid's inequality assumptions

$$\mathcal{A}_{MD} := \left\{ (f,\mu) \;\middle|\; \begin{array}{c} f : \mathcal{X}_1 \times \cdots \times \mathcal{X}_m \to \mathbb{R}, \\ \mu \in \mathcal{M}(\mathcal{X}_1) \otimes \cdots \otimes \mathcal{M}(\mathcal{X}_m), \\ \mathbb{E}_\mu[f] \leq 0, \\ \mathrm{Osc}_i(f) \leq D_i \end{array} \right\}$$

$$\boxed{r_f(x) := 1_{f(x) \geq a}}$$

$$\mathcal{A}_{\Delta} := \left\{ (f,\mu) \;\middle|\; \begin{array}{c} f : \mathcal{X}_1 \times \cdots \times \mathcal{X}_m \to \mathbb{R}, \\ \mu \in \Delta_1(\mathcal{X}_1) \otimes \cdots \otimes \Delta_1(\mathcal{X}_m), \\ \mathbb{E}_\mu[f] \leq 0, \\ \mathrm{Osc}_i(f) \leq D_i \end{array} \right\}$$

Second reduction (positions of the Diracs)

$$G(f,\mu) \leq 0 \Leftrightarrow \mathbb{E}_\mu[g_j \circ f] \leq 0 \quad 1 \leq j \leq n$$

$$\mathcal{U}(\mathcal{A}) := \sup_{(f,\mu)\in\mathcal{A}} \mathbb{E}_\mu[r \circ f]$$

$$\mathcal{A} = \{(f,\mu) \in \mathcal{G} \times \otimes_{i=1}^m \mathcal{M}(\chi_i) \,|\, \mathcal{G}(f,\mu) \leq 0\}$$

$$\mathcal{G} \subset \mathcal{F}$$

$\mathcal{F}$: Set of real-valued measurable functions on $\chi := \chi_1 \times \cdots \times \chi_m$

$\mathcal{F}_\mathcal{D}$: Real functions on $\mathcal{D} := \{0,\ldots,n\}^m$

$$\mathcal{G}_\mathcal{D} \subset \mathcal{F}_\mathcal{D}$$

$$\mathcal{A}_\mathcal{D} = \{(h,\alpha) \in \mathcal{G}_\mathcal{D} \times \otimes_{i=1}^m \mathcal{M}(\mathcal{D}) \,|\, \mathcal{G}(f,\mu) \leq 0\}$$

$$\mathcal{U}(\mathcal{A}_\mathcal{D}) := \sup_{(h,\alpha)\in\mathcal{A}_\mathcal{D}} \mathbb{E}_\alpha[r \circ h]$$

Second reduction (positions of the Diracs)

Theorem If

$$\mathbb{F}\left[\mathcal{G} \times \otimes_{i=1}^{m} \Delta_n(\chi_i)\right] = \mathcal{G}_{\mathcal{D}}$$

$$\mathbb{F} \colon \mathcal{F} \times \otimes_{i=1}^{m} \Delta_n(\mathcal{X}_i) \longrightarrow \mathcal{F}_{\mathcal{D}}$$

$$(f, \otimes_{i=1}^{m}(\sum_{k=0}^{n} \alpha_k^i \delta_{x_i^k})) \longrightarrow (s_1, \ldots, s_m) \to f(x_1^{s_1}, \ldots, x_m^{s_m})$$

Then

$$\mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_{\mathcal{D}})$$

Application to McDiarmid's inequality assumptions

$$\mathcal{A}_{MD} := \left\{ (f,\mu) \,\middle|\, \begin{array}{c} f \colon \mathcal{X}_1 \times \cdots \times \mathcal{X}_m \to \mathbb{R}, \\ \mu \in \mathcal{M}(\mathcal{X}_1) \otimes \cdots \otimes \mathcal{M}(\mathcal{X}_m), \\ \mathbb{E}_\mu[f] \leq 0, \\ \mathrm{Osc}_i(f) \leq D_i \end{array} \right\}$$

$$r \circ f(x) := 1_{f(x) \geq a}$$

$$\mathcal{A}_{\mathcal{D}} := \left\{ (h,\alpha) \,\middle|\, \begin{array}{c} h \colon \{0,1\}^m \to \mathbb{R}, \\ \alpha \in \mathcal{M}(\{0,1\}) \otimes \cdots \otimes \mathcal{M}(\{0,1\}) \\ \mathbb{E}_\alpha[h] \leq 0, \\ \mathrm{Osc}_i(h) \leq D_i \end{array} \right\}$$

Third reduction: lattice structure of the function space

$$\mathcal{A}_{\mathcal{D}} := \left\{ (h,\alpha) \;\middle|\; \begin{array}{c} h\colon \{0,1\}^m \to \mathbb{R}, \\ \alpha \in \mathcal{M}(\{0,1\}) \otimes \cdots \otimes \mathcal{M}(\{0,1\}) \\ \mathbb{E}_\alpha[h] \leq 0, \\ \mathrm{Osc}_i(h) \leq D_i \end{array} \right\}$$

$\mathcal{F}_{\mathcal{D}}$ is a lattice.

$\mathcal{G}_{\mathcal{D}}$ is a sub-lattice.

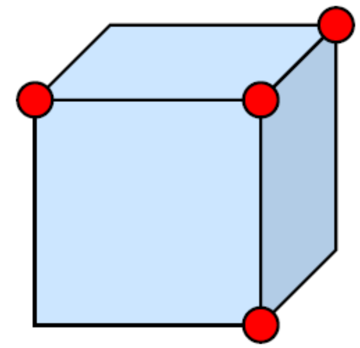
$$(h,\alpha) \in \mathcal{A}_{\mathcal{D}} \Rightarrow \big(\min(h,a),\alpha\big) \in \mathcal{A}_{\mathcal{D}}$$

For each $C \in \mathcal{C} := \{0,1\}^m$

$C_{\mathcal{D}} := \{h \in \mathcal{G}_{\mathcal{D}} \,:\, \{s \,:\, h(s) = a\} = C\}$ is a sub-lattice

Reduction of optimization variables

$$\mathcal{A}_{\mathcal{C}} := \left\{ (C,\alpha) \;\middle|\; \begin{array}{c} C \subset \{0,1\}^m, \\ \alpha \in \otimes_{i=1}^m \mathcal{M}(\{0,1\}), \\ \mathbb{E}_\alpha[h^C] \leq 0 \end{array} \right\}$$



$$h^C \colon \{0,1\}^m \longrightarrow \mathbb{R}$$

$$t \longrightarrow a - \min_{s \in C} \sum_{i\colon s_i \neq t_i} D_i$$

$$\mathcal{U}(\mathcal{A}_{\mathcal{C}}) := \sup_{(C,\alpha)\in\mathcal{A}_{\mathcal{C}}} \alpha[h^C \geq a]$$

Theorem

$$\mathcal{U}(\mathcal{A}_{MD}) = \mathcal{U}(\mathcal{A}_{\mathcal{C}})$$

Literature

$$\mathcal{U}(\mathcal{A}) := \sup_{(f,\mu)\in\mathcal{A}} \mu[f(X) \geq a]$$

Non-convex and infinite dimensional optimization problems

Can be considered as a **generalization of classical Chebyshev inequalities**

History of classical inequalities: Karlin, Studden (1966, Tchebycheff systems with applications in analysis and statistics)

Connection between Chebyshev inequalities and optimization theory

- Mulholland & Rogers (1958, Representation theorems for distribution functions)
- Godwin (1973, Manipulation of voting schemes: a general result)
- Isii (1959, On a method for generalization of Tchebycheff's inequality
  1960, The extrema of probability determined by generalized moments
  1962, On sharpness of Techebycheff-type inequalities)
- Olhin & Pratt (1958, A multivariate Tchebycheff inequality)
- Classical Markov-Krein theorem (Karlin, Studden, 1958)
- Dynkin (1978, Sufficient statistics & extreme points)
- Karr (1983, Extreme points of probability measures with applications)

Literature

$$\mathcal{U}(\mathcal{A}) := \sup_{(f,\mu)\in\mathcal{A}} \mu[f(X) \geq a]$$

Our work: Further generalization to

- Independence constraints
- More general domains (Suslin spaces) (non metric, non compact)
- More general classes of functions (measurable) (non continuous, non-bounded)
- More general classes of probability measures
- More general constraints (inequalities, on measures and functions)

Theory of majorization

- Marshall & Olkin (1979, Inequalities: Theory of majorization and its applications)

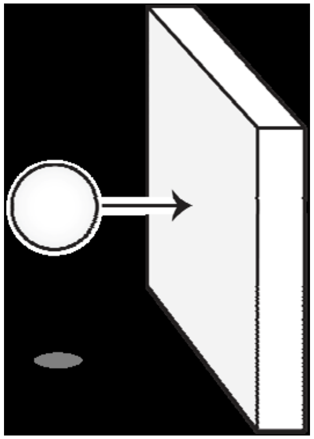
Inequalities of

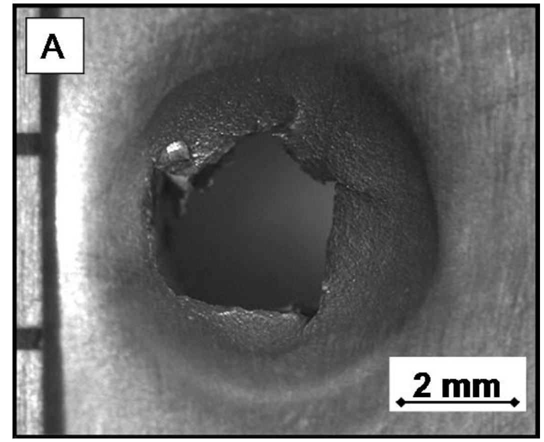
- Anderson (1955, the integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities)
- Hoeffding (1956, on the distribution of the number of successes in independent trials)
- Joe (1987, Majorization, randomness and dependence for multivariate distributions)
- Bentkus, Geuze, Van Zuijlen (2006, Optimal Hoeffding like inequalities under a symmetry assumption)
- Pinelis (2007, Exact inequalities for sums of asymmetric random variables with applications.
  2008, On inequalities for sums of bounded random variables)

Our proof rely on

- Winkler (1988, Extreme points of moment sets)
- Follows from an extension of Choquet theory (Phelps 2001, lectures on Choquet's theorem) by Von Weizsacker & Winkler (1979, Integral representation in the set of solutions of a generalized moment problem)
- Kendall (1962, Simplexes & Vector lattices)

Caltech Small Particle Hypervelocity Impact Range





$$(h, \theta, v) \xrightarrow{G} G(h, \theta, v)$$

Plate thickness

Plate Obliquity

Projectile velocity

Perforation area

We want to certify that

$$\mathbb{P}[G = 0] \leq \epsilon$$

Caltech Hypervelocity Impact Surrogate Model

Plate thickness $h \in \mathcal{X}_1 := [1.524, 2.667]\,\mathrm{mm},$

Plate Obliquity $\theta \in \mathcal{X}_2 := [0, \frac{\pi}{6}],$

Projectile velocity $v \in \mathcal{X}_3 := [2.1, 2.8]\,\mathrm{km}\cdot\mathrm{s}^{-1}.$

Deterministic surrogate model for the perforation area (in mm^2)

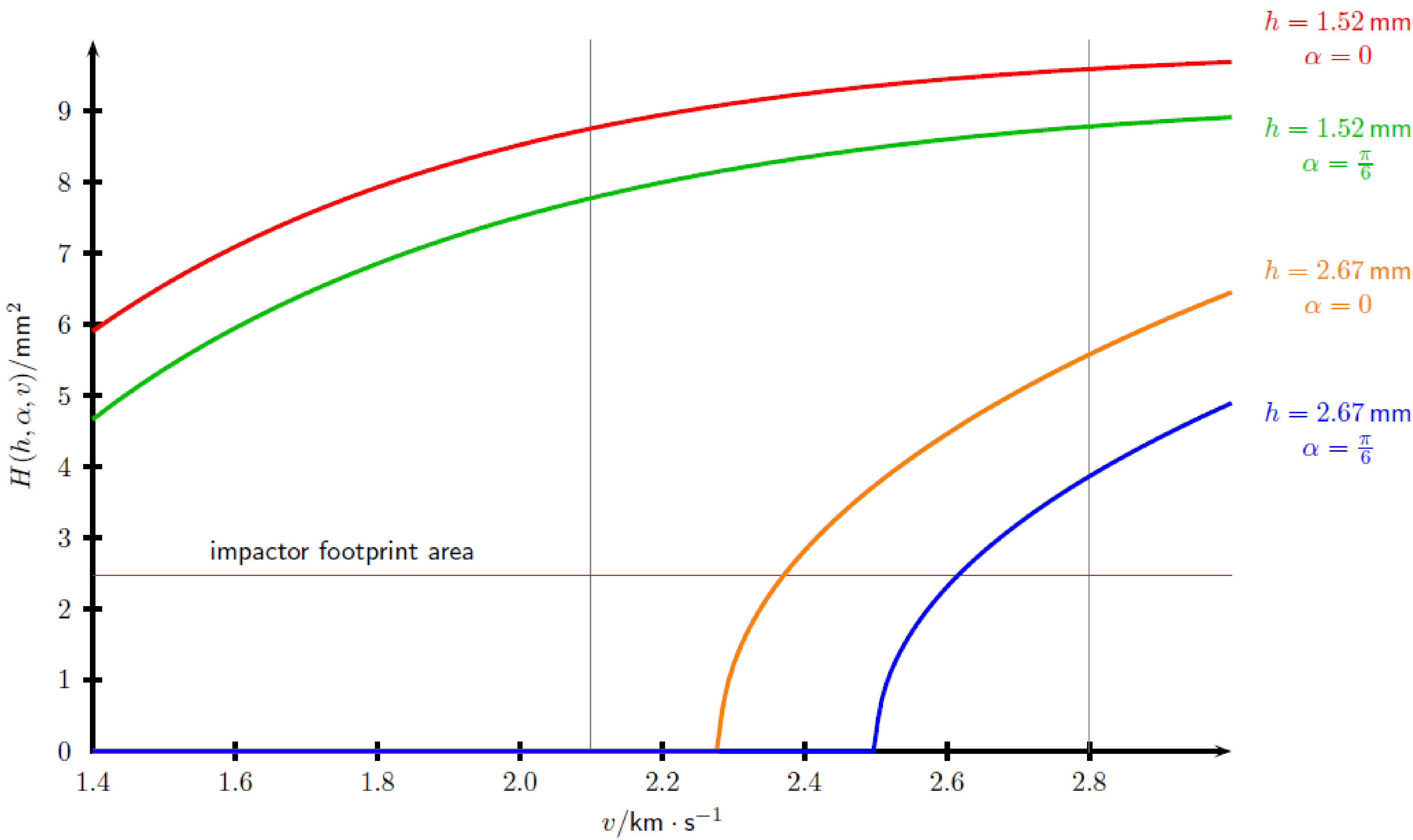
$$H(h, \theta, v) = K \left(\frac{h}{D_\mathrm{p}}\right)^p (\cos\theta)^u \left(\tanh\left(\frac{v}{v_\mathrm{bl}} - 1\right)\right)_+^m,$$

$H_0 = 0.5794\,\mathrm{km}\cdot\mathrm{s}^{-1}$, $s = 1.4004$, $n = 0.4482$, $K = 10.3936\,\mathrm{mm}^2$,
$p = 0.4757$, $u = 1.0275$, $m = 0.4682$.

The *ballistic limit velocity* (the speed below which no perforation area occurs) is given by

$$v_\mathrm{bl} := H_0 \left(\frac{h}{(\cos\theta)^n}\right)^s$$

Caltech Hypervelocity Impact Surrogate Model



Bound on the probability of non perforation

$$\mathcal{A}_{\text{McD}} := \left\{ (f,\mu) \,\middle|\, \begin{array}{c} \mu = \mu_1 \otimes \mu_2 \otimes \mu_3, \\ 5.5\, mm^2 \leq \mathbb{E}_\mu[f] \leq 7.5\, mm^2, \\ \mathrm{Osc}_i(f) \leq \mathrm{Osc}_i(H) \text{ for } i = 1,2,3 \\ f \geq 0 \end{array} \right\}$$

$$\mathrm{Osc}_i(f) := \sup_{(x_1,\ldots,x_m)\in\mathcal{X}} \sup_{x_i'\in\mathcal{X}_i} \big(f(\ldots,x_i,\ldots) - f(\ldots,x_i',\ldots) \big).$$

$$\mathcal{U}(\mathcal{A}_{\text{McD}}) := \sup_{(f,\mu)\in\mathcal{A}} \mu[f(X)=0]$$

$$\mathbb{P}[H=0] \leq \mathcal{U}(\mathcal{A}_{\text{McD}}) \leq \exp\left(-\frac{2m_1^2}{\sum_{i=1}^3 \mathrm{Osc}_i(H)^2} \right) = 66.4\%.$$

Optimal bound on the probability of non perforation

$$\mathcal{A}_{\mathrm{McD}} := \left\{ (f,\mu) \;\middle|\; \begin{array}{c} \mu = \mu_1 \otimes \mu_2 \otimes \mu_3, \\ 5.5\, mm^2 \leq \mathbb{E}_\mu[f] \leq 7.5\, mm^2, \\ \mathrm{Osc}_i(f) \leq \mathrm{Osc}_i(H) \text{ for } i = 1,2,3 \\ f \geq 0 \end{array} \right\}$$

$$\mathcal{U}(\mathcal{A}_{\mathrm{McD}}) := \sup_{(f,\mu)\in\mathcal{A}} \mu[f(X) = 0]$$

$$\mathbb{P}[H = 0] \leq \mathcal{U}(\mathcal{A}_{\mathrm{McD}}) = 43.7\%.$$

Optimal bound on the probability of non perforation

$$\mathcal{A} := \left\{ (f,\mu) \;\middle|\; \begin{array}{c} \mu = \mu_1 \otimes \mu_2 \otimes \mu_3, \\ m_1 \leq \mathbb{E}_\mu[H] \leq m_2 \\ f = H \end{array} \right\}$$

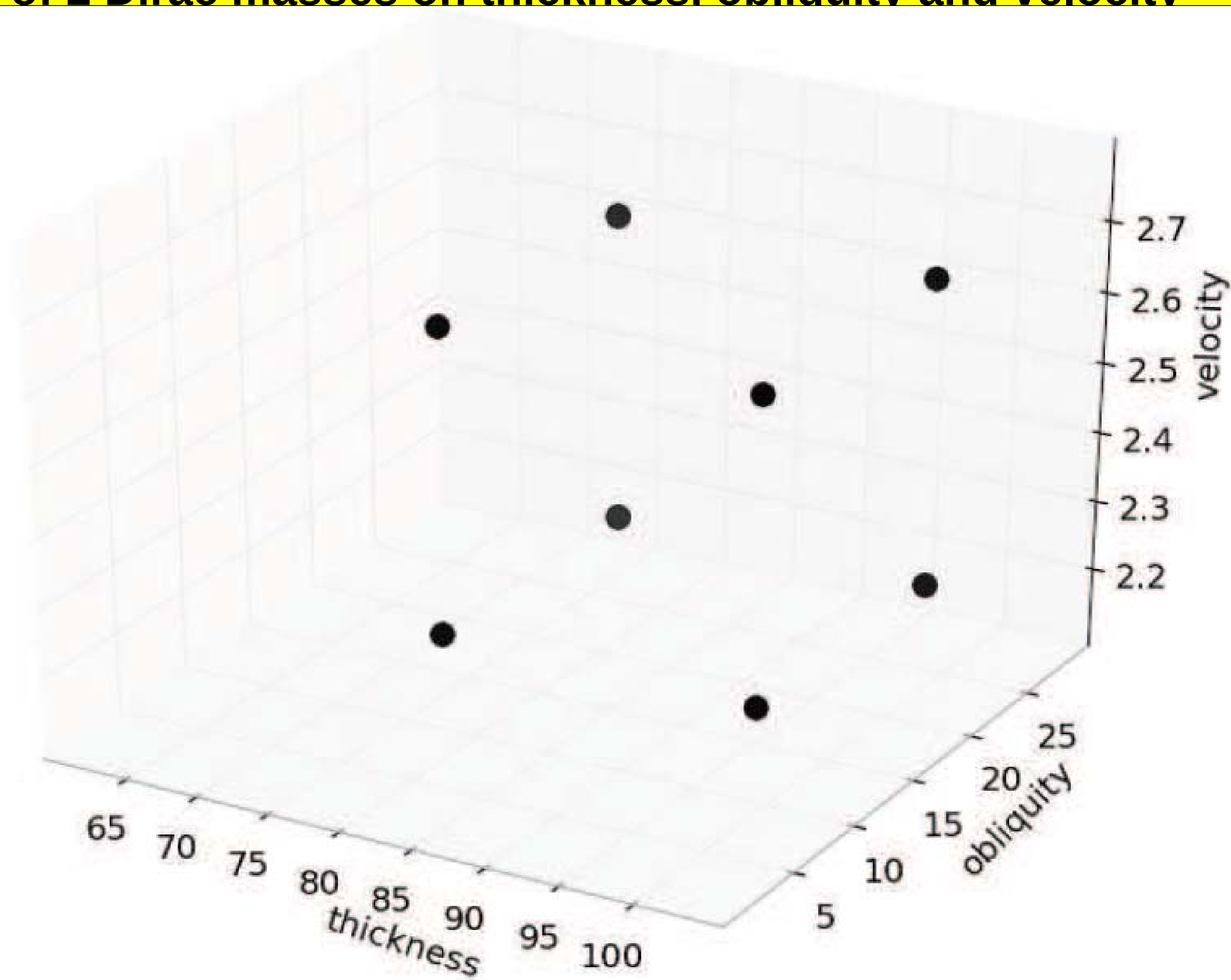
$$\mathcal{U}(\mathcal{A}) := \sup_{(f,\mu)\in\mathcal{A}} \mu[f(X) = 0]$$

Application of the reduction theorem

The measure of probability can be reduced to the tensorization of 2 Dirac masses on thickness, obliquity and velocity

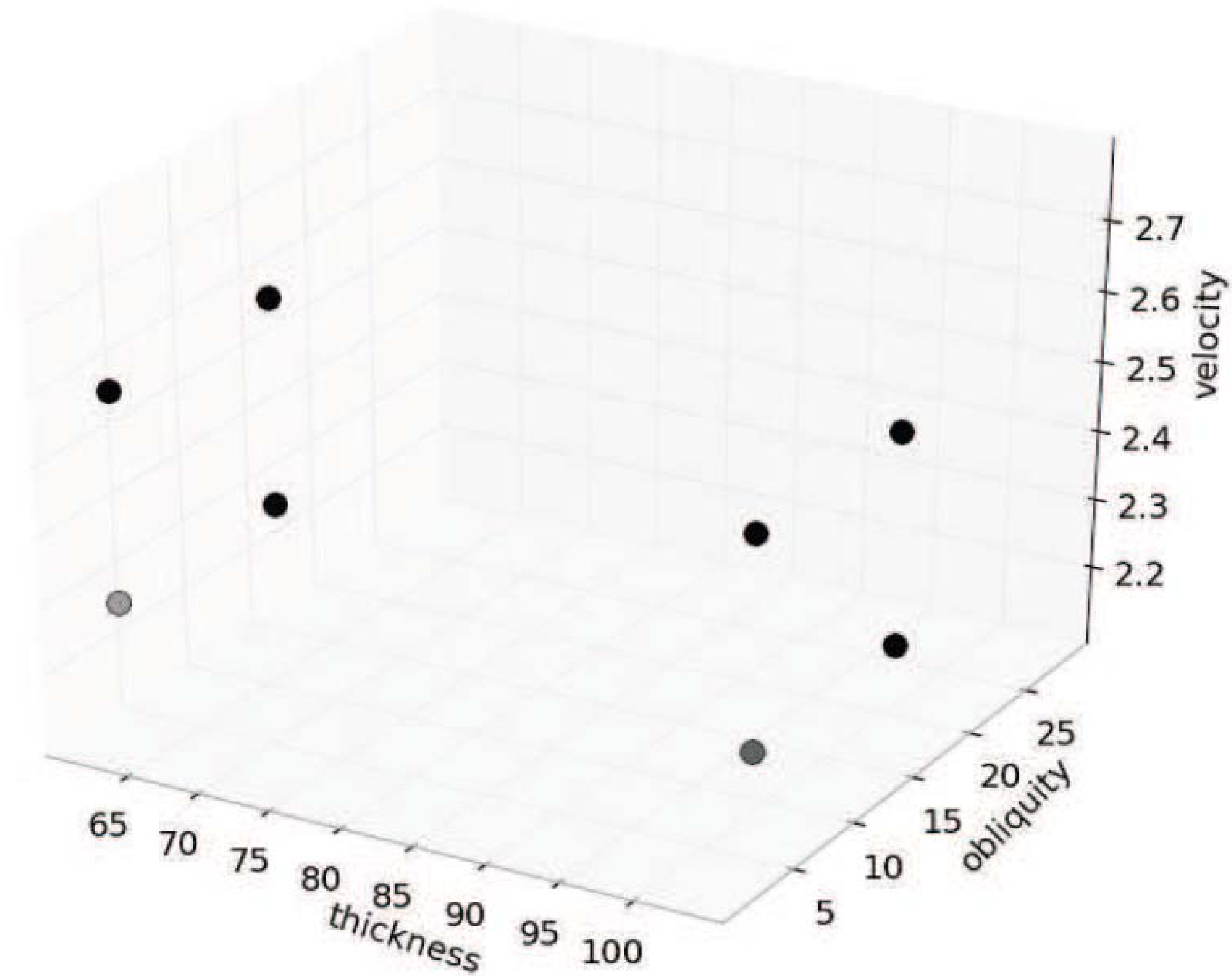
$$\mathcal{U}(\mathcal{A}) \overset{\text{num}}{=} 37.9\%$$

The optimization variables can be reduced to the tensorization of 2 Dirac masses on thickness, obliquity and velocity



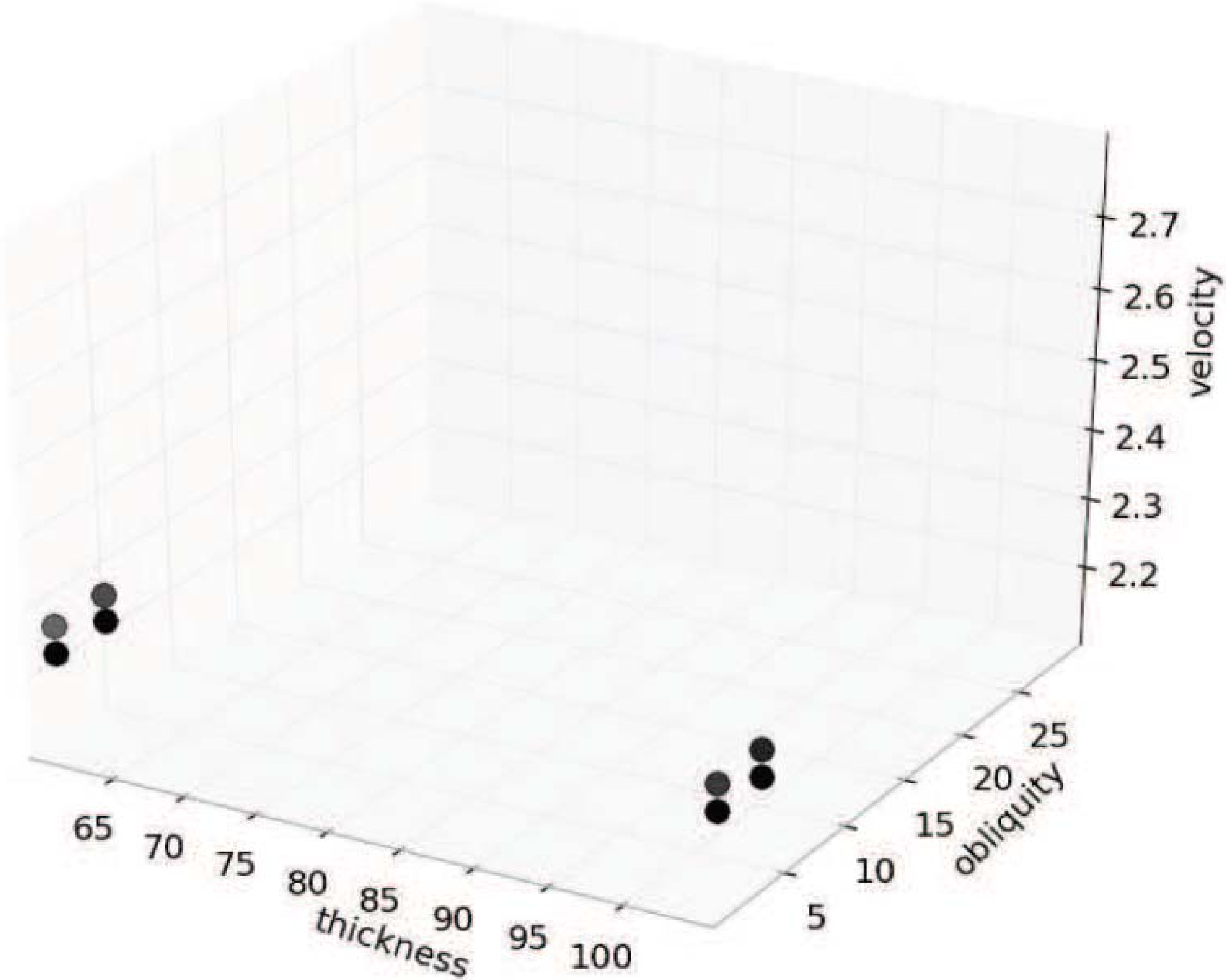
Support Points at iteration 0

Numerical optimization



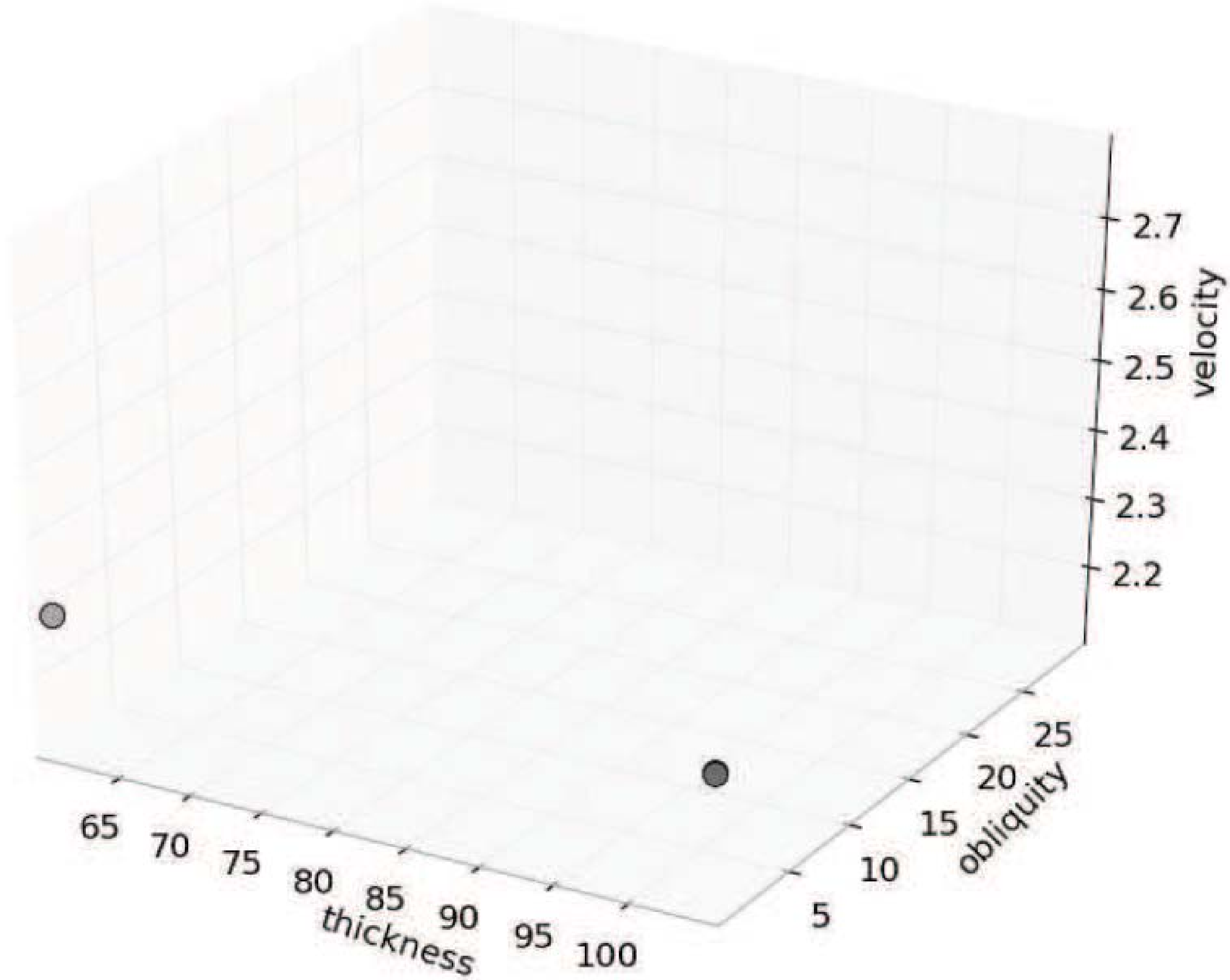
Support Points at iteration 150

Numerical optimization



Support Points at iteration 200

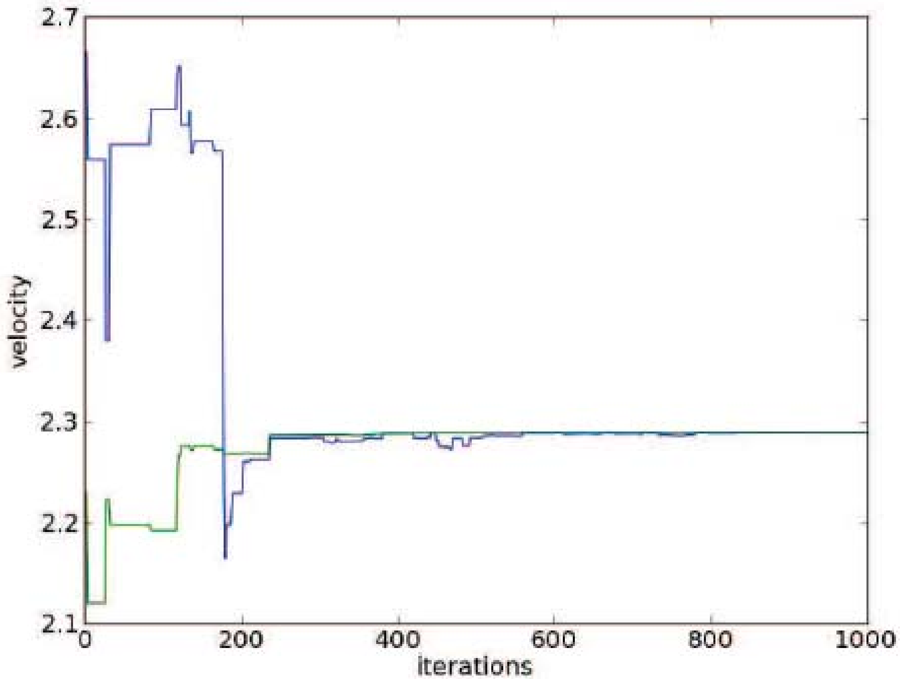
Velocity and obliquity marginals each collapse to a single Dirac mass. The plate thickness marginal collapses to have support on the extremes of its range.



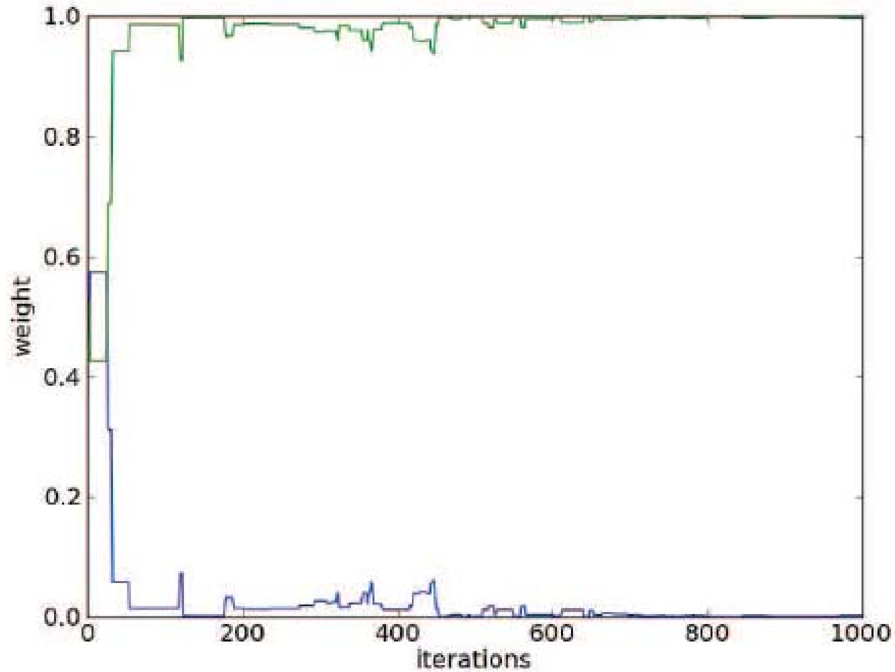
Iteration 1000

The probability of non-perforation is maximized by a distribution supported on the minimal, not maximal, impact obliquity.

Velocity



Position of Dirac Masses



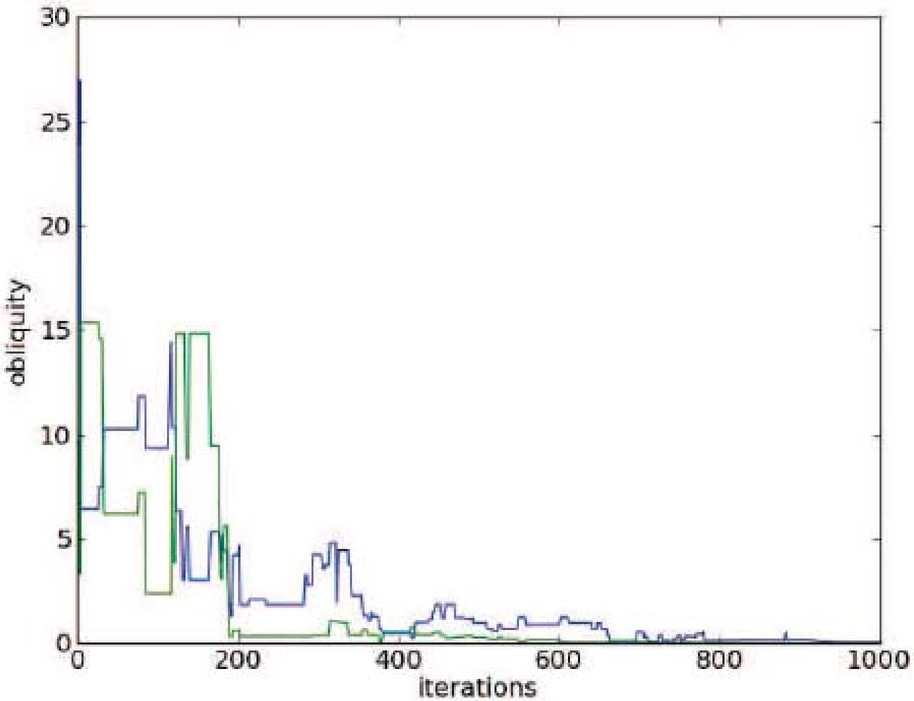
Weight of on Dirac Masses

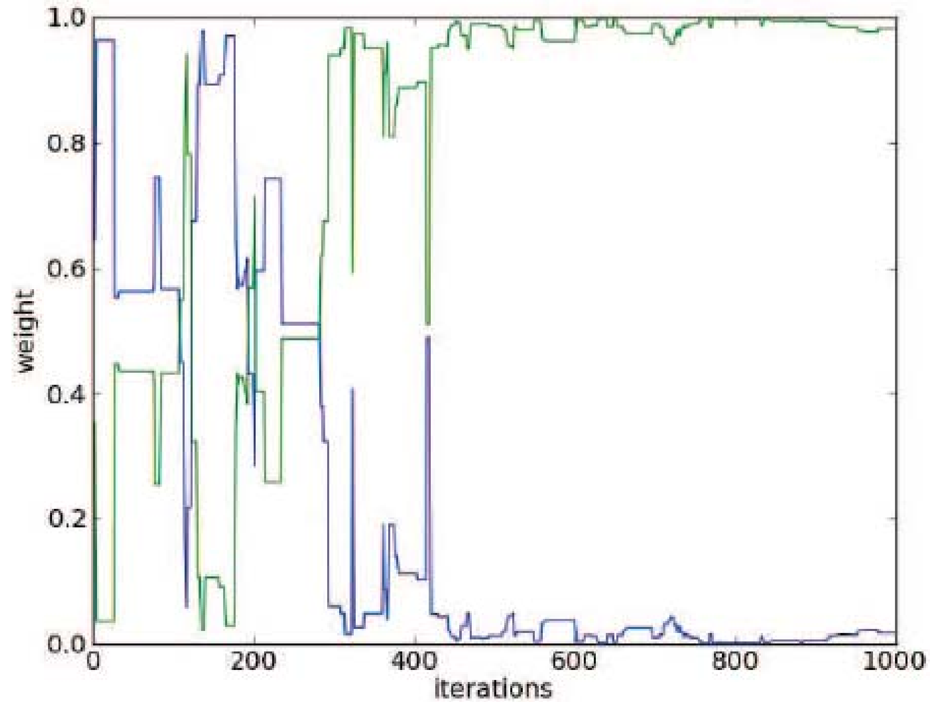
Position and weight vs Iteration

Converges towards non extreme value at $2.289 \ \mathrm{km} \cdot \mathrm{s}^{-1}$

Reducing the velocity range does not decrease the optimal bound on the probability of non perforation

Obliquity





Position of Dirac Masses

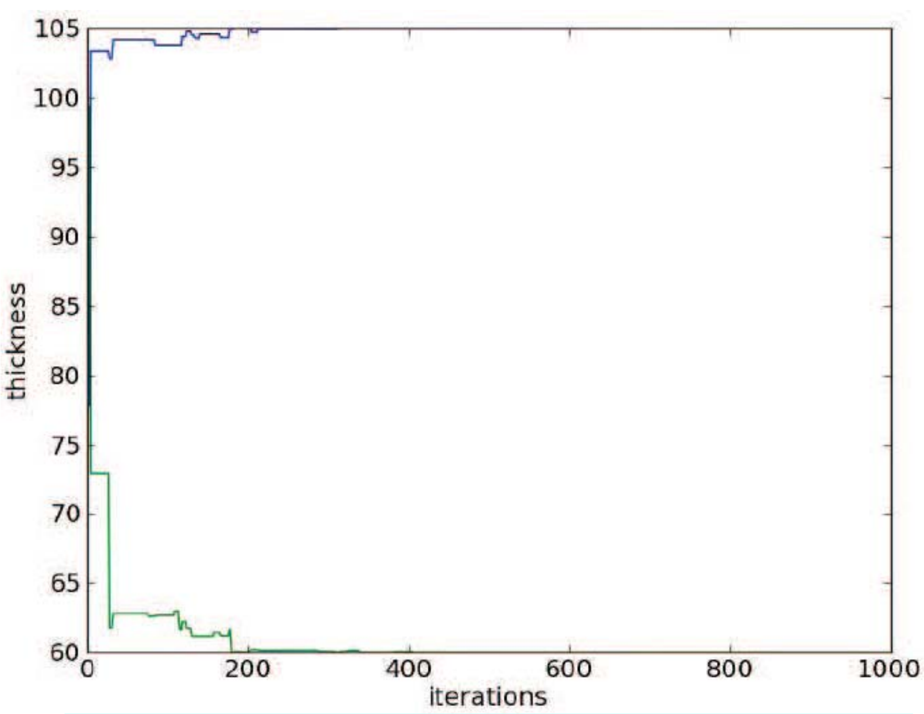
Weight of on Dirac Masses

Position and weight vs Iteration

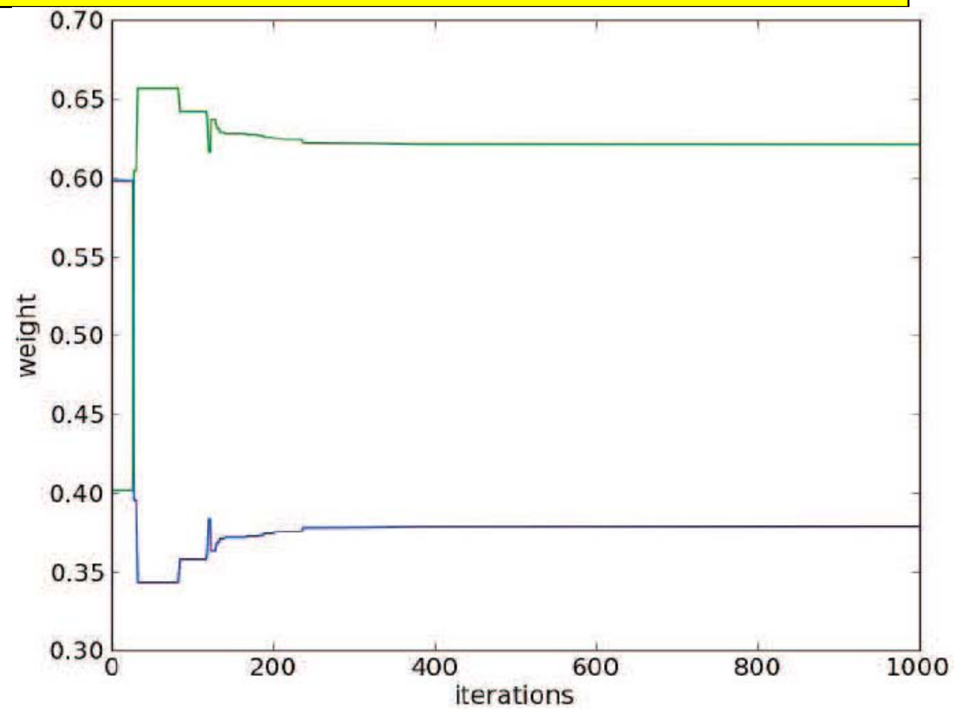
Converges towards 0 obliquity

Reducing maximum obliquity does not decrease the optimal bound on the probability of non perforation

Thickness



Position of Dirac Masses



Weight of on Dirac Masses

Position and weight vs Iteration

Converges towards the extremes of its range

Reducing uncertainty in thickness will decrease the optimal bound on the probability of non perforation

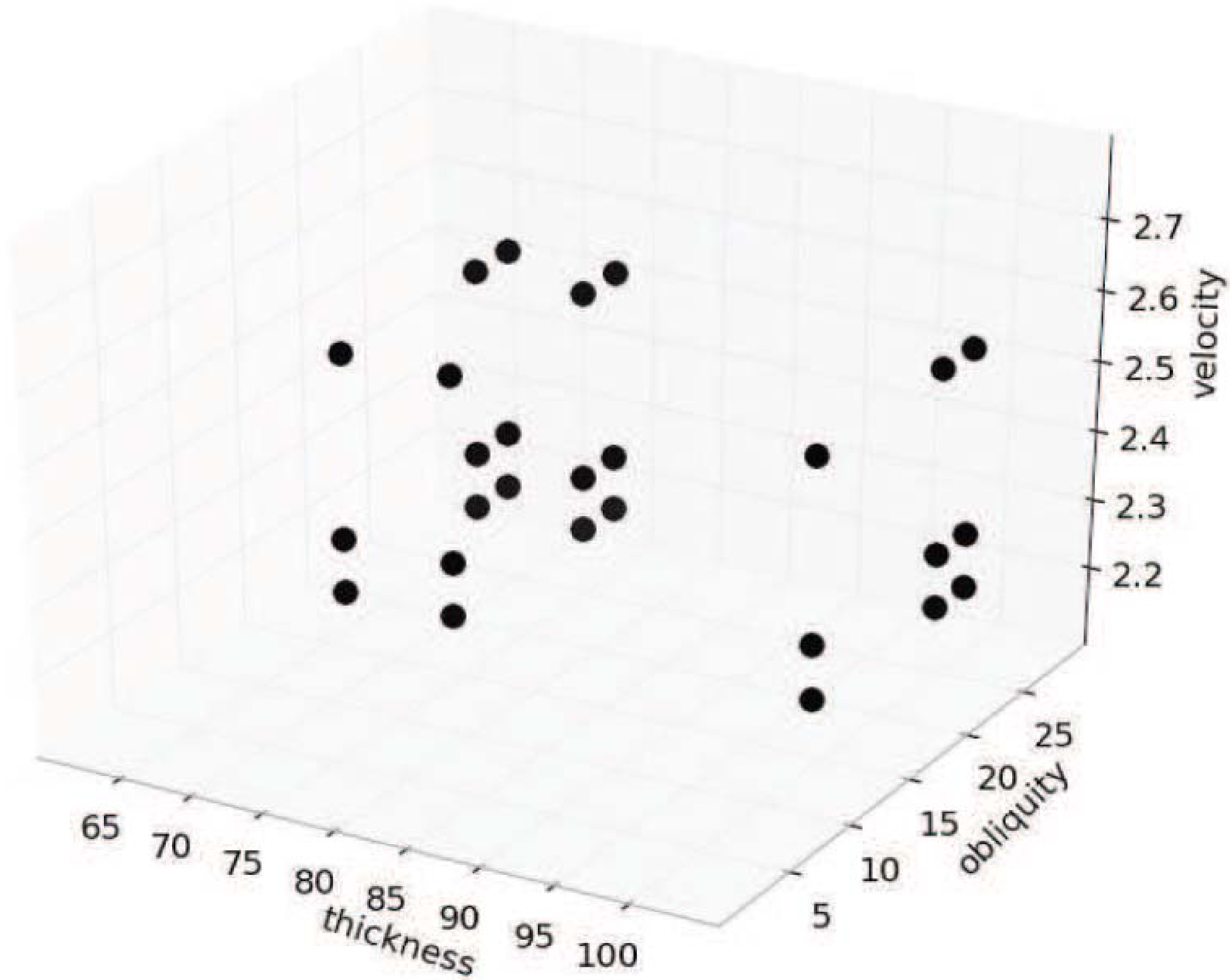
Important observations

Extremizers are singular

They identify key players
i.e. vulnerabilities of the physical system

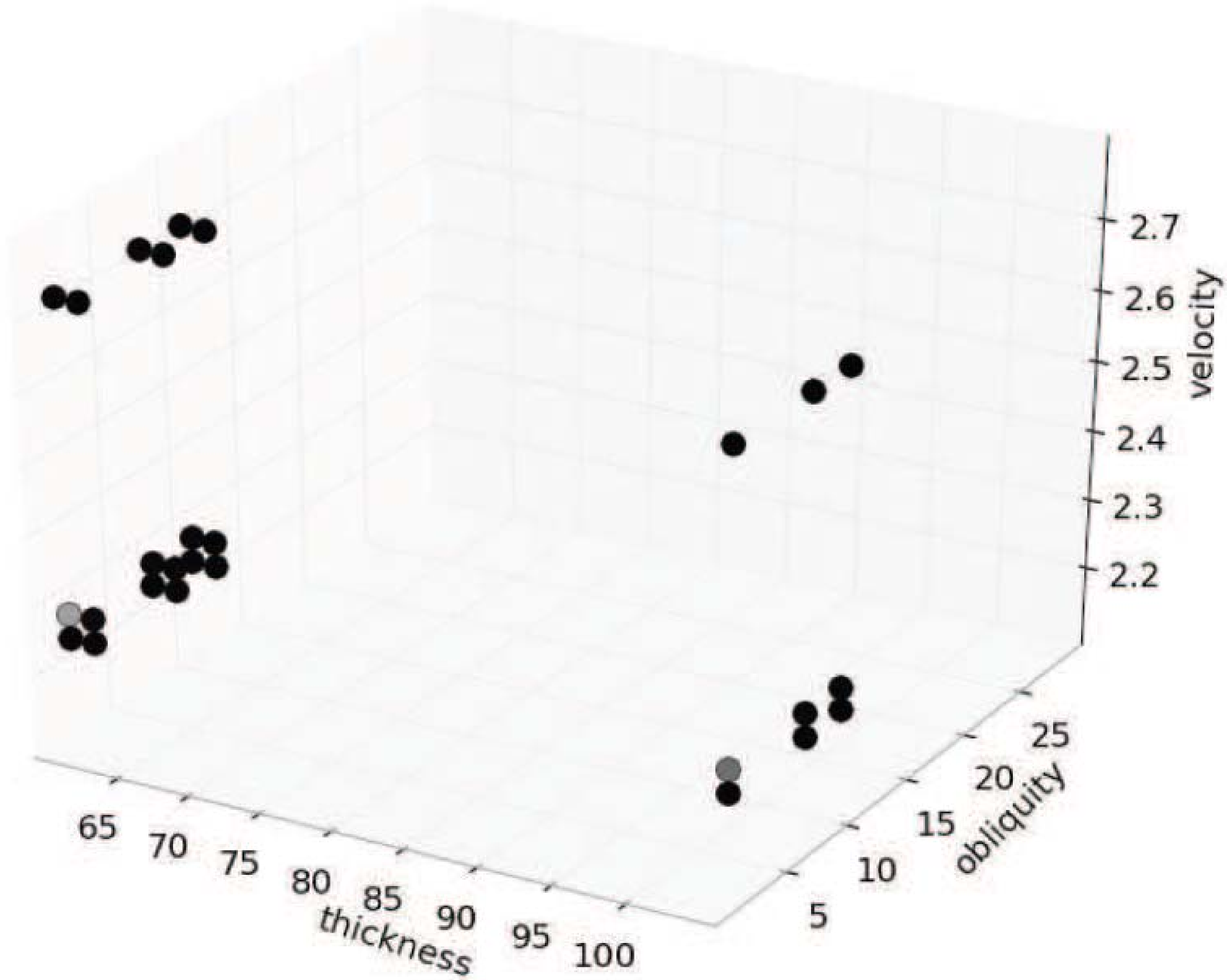
Extremizers are attractors

Initialization with 3 support points per marginal



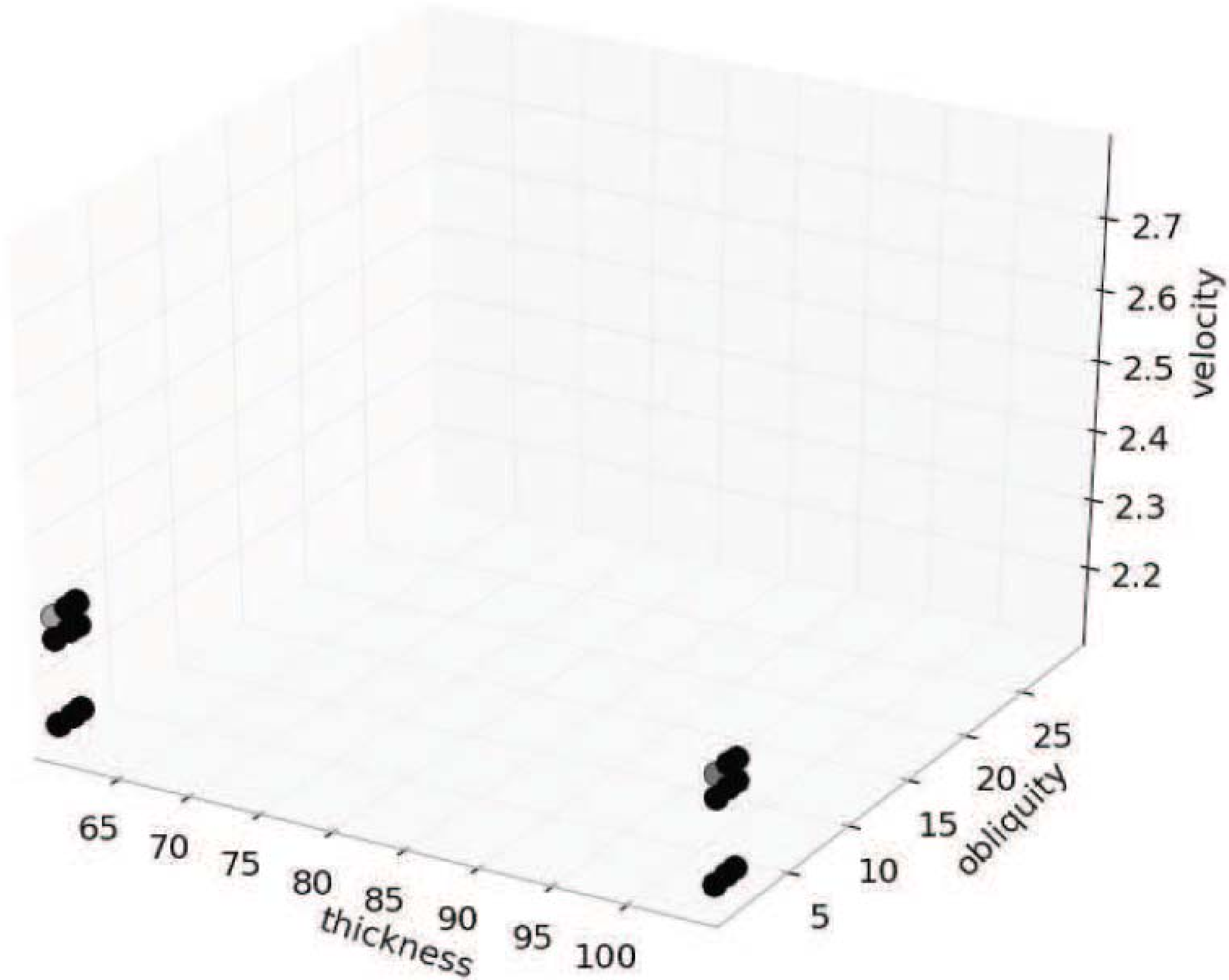
Support Points at iteration 0

Initialization with 3 support points per marginal



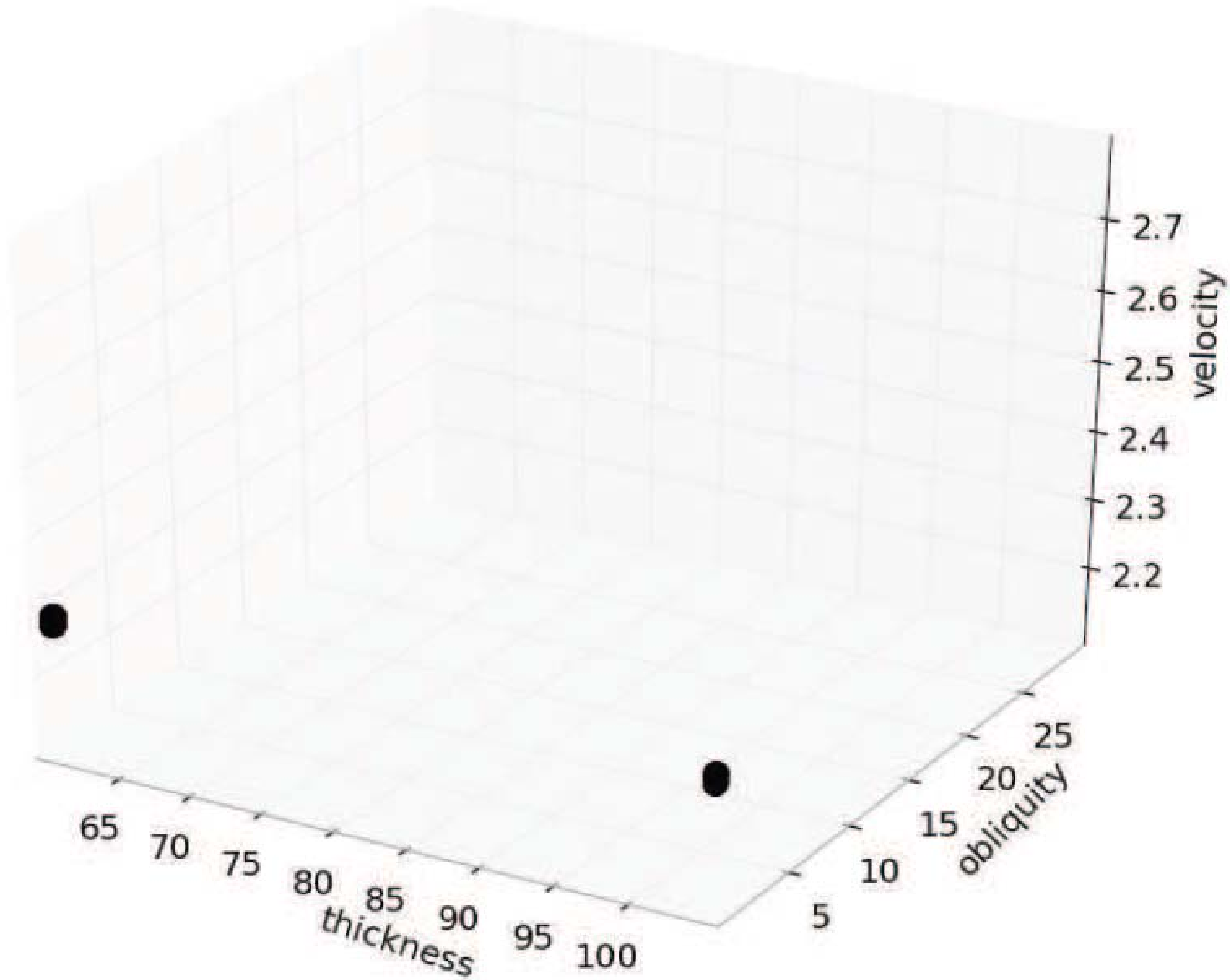
Support Points at iteration 500

Initialization with 3 support points per marginal



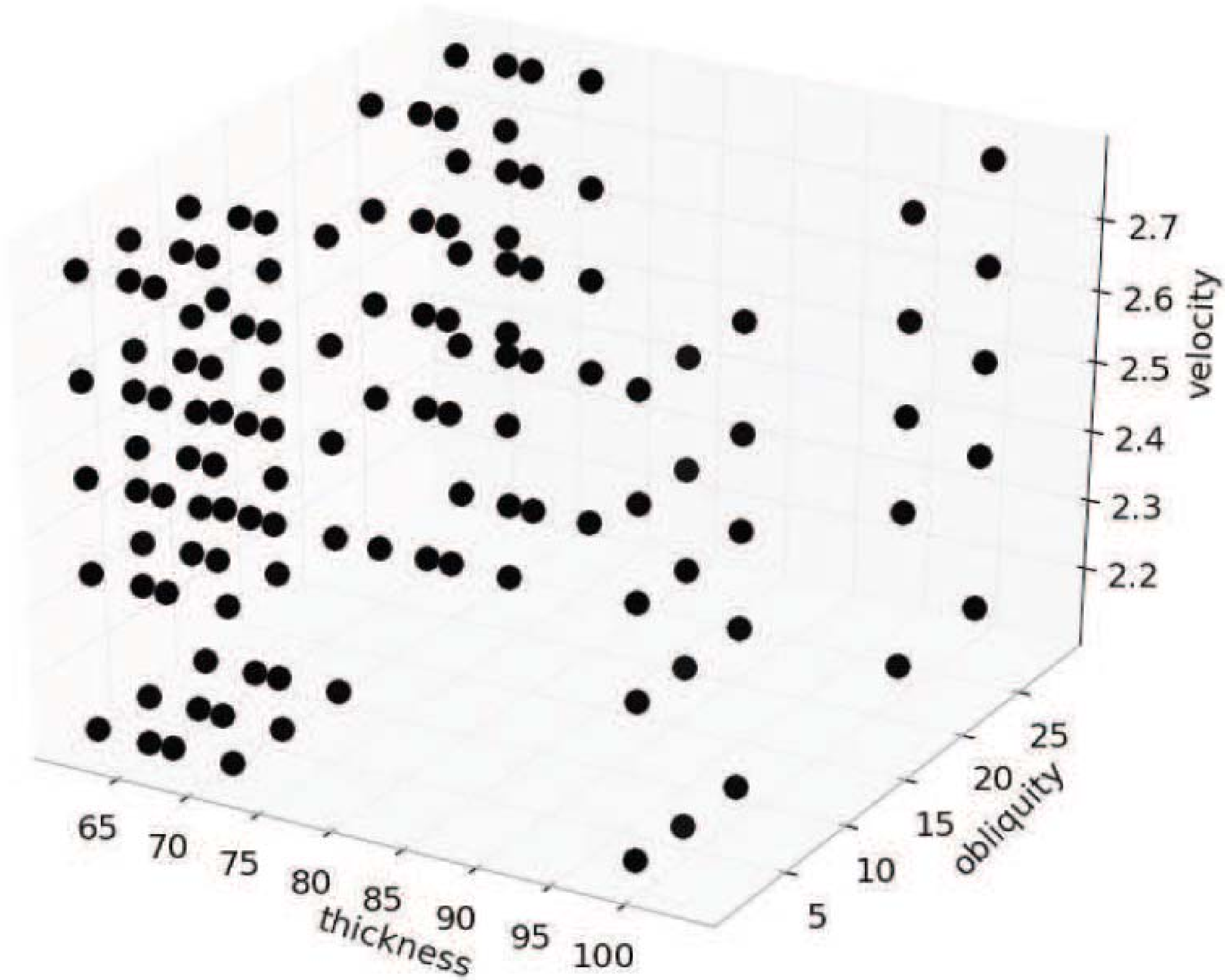
Support Points at iteration 1000

Initialization with 3 support points per marginal



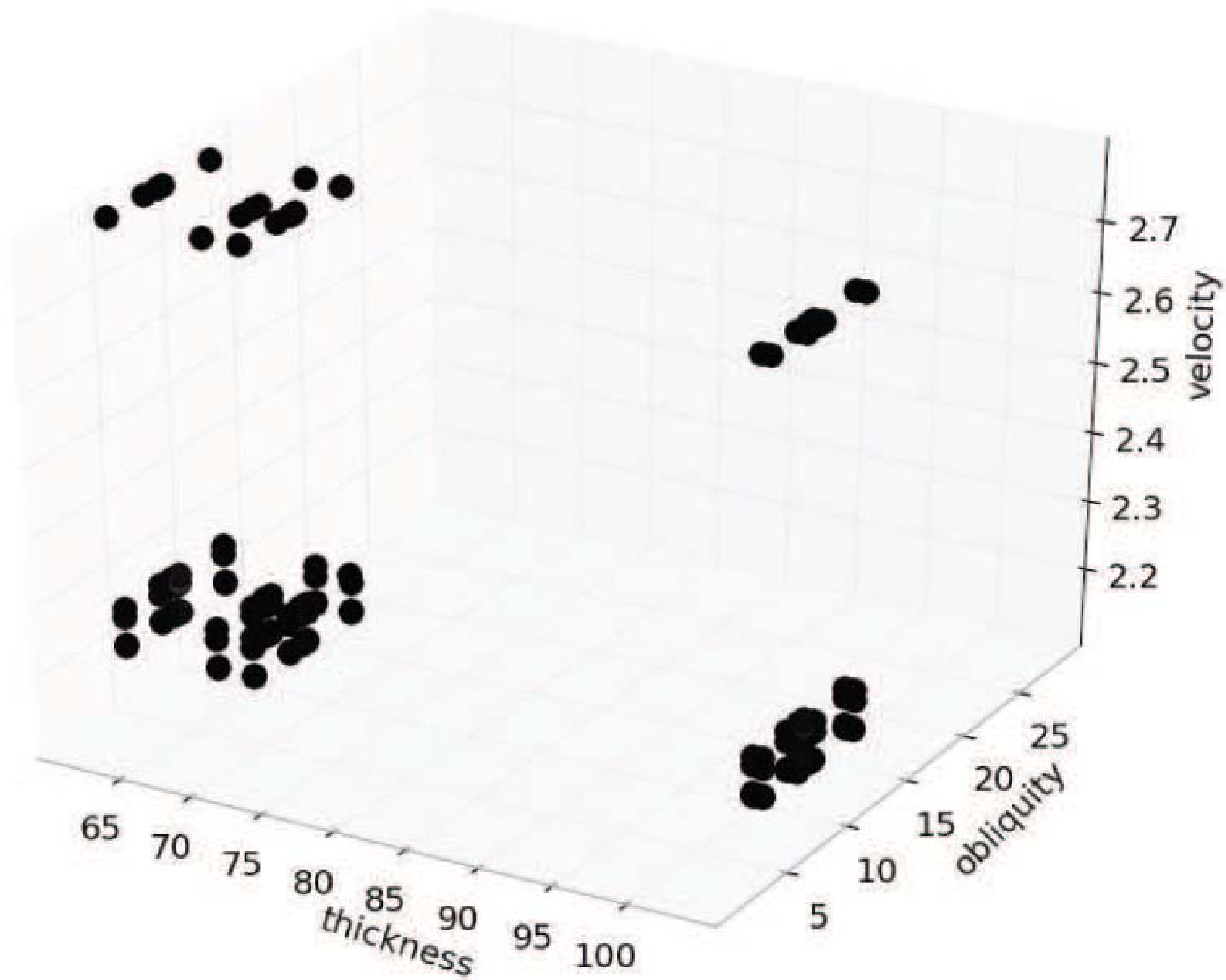
Support Points at iteration 2155

Initialization with 5 support points per marginal



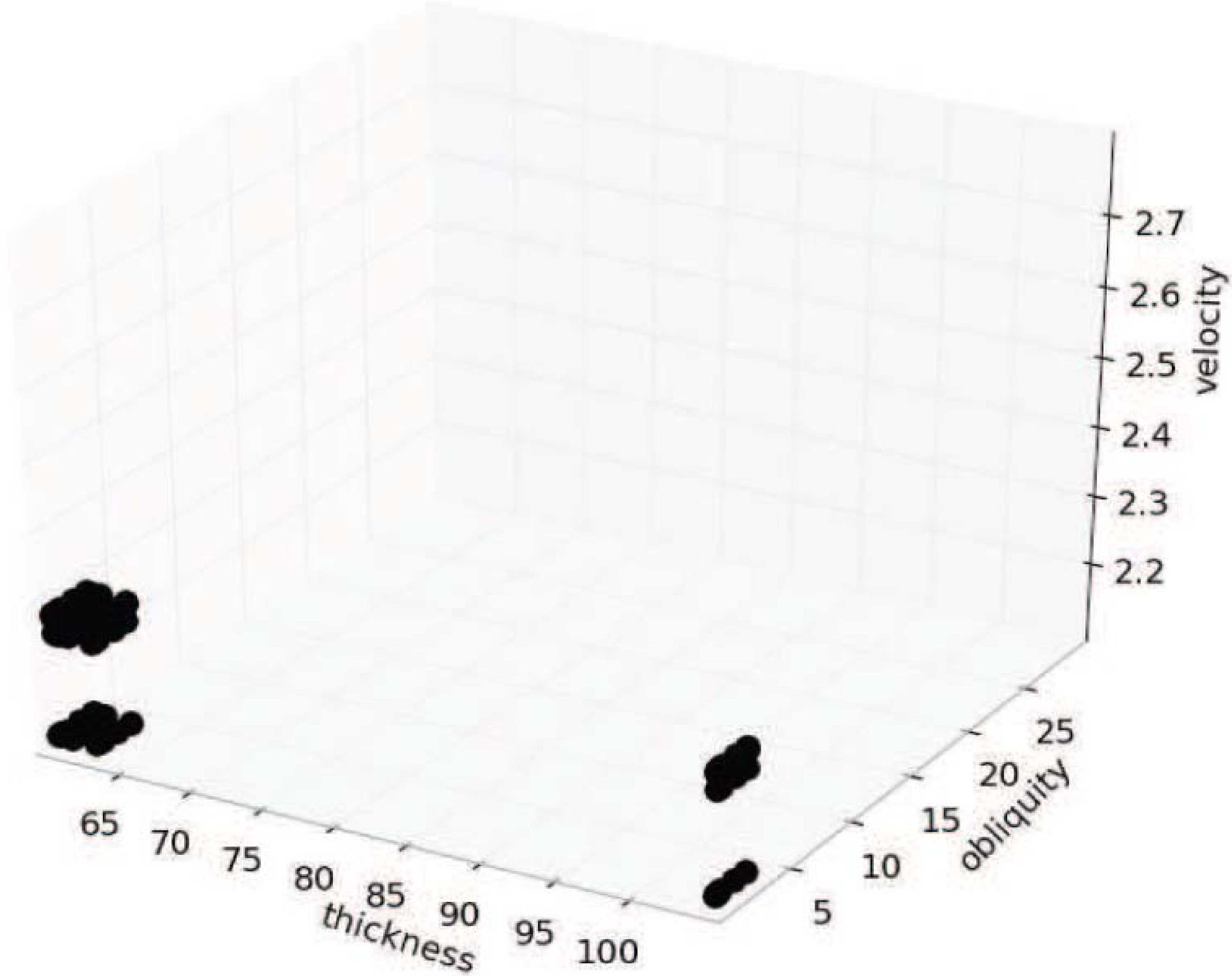
Support Points at iteration 0

Initialization with 5 support points per marginal



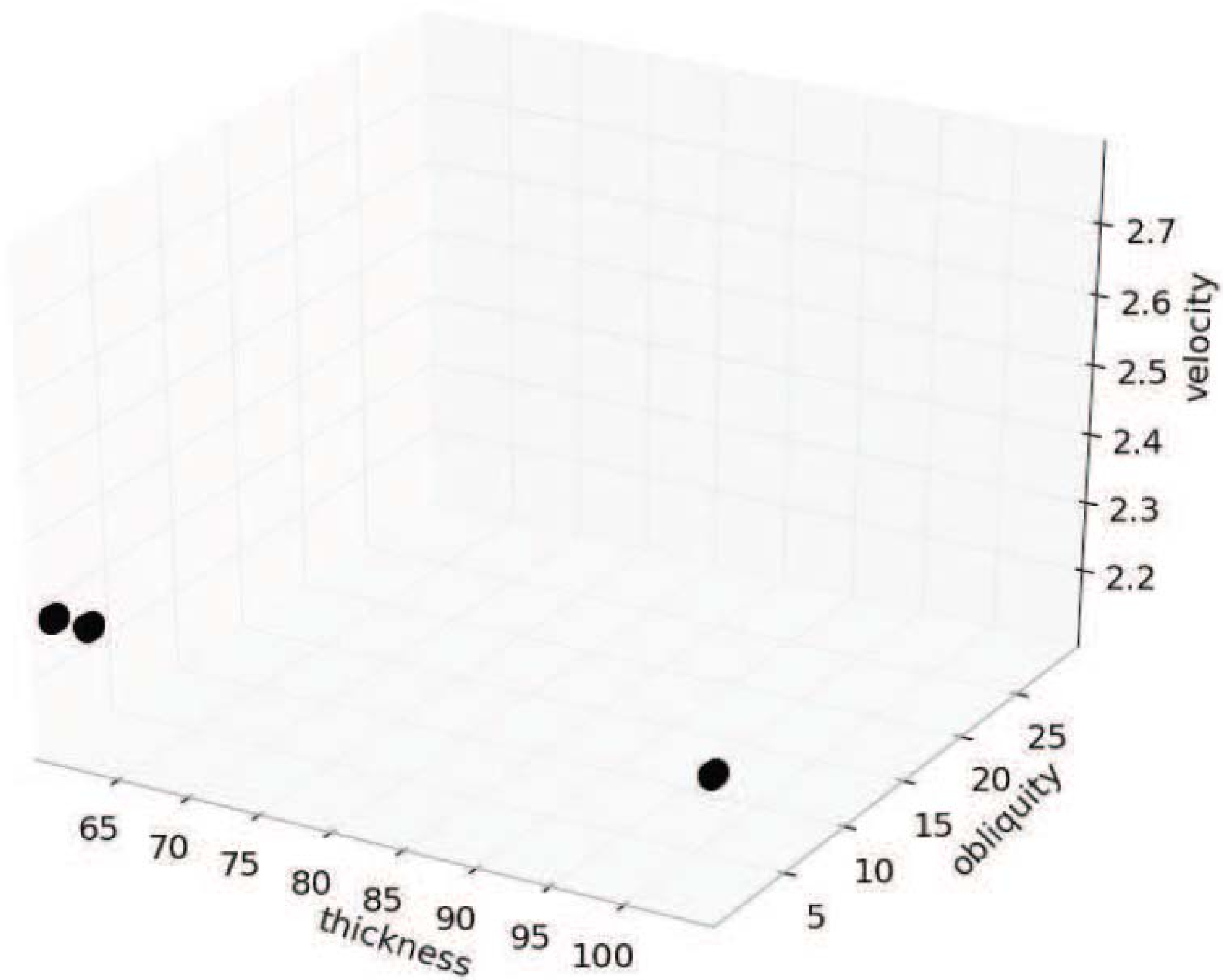
Support Points at iteration 1000

Initialization with 5 support points per marginal



Support Points at iteration 3000

Initialization with 5 support points per marginal



Support Points at iteration 7100

Optimal bounds for other admissible sets

Admissible scenarios, $\mathcal{A}$	$\mathcal{U}(\mathcal{A})$	Method
$\mathcal{A}_{\text{McD}}$: independence, oscillation and mean constraints (exact response H not given)	$\leq 66.4\%$ $= 43.7\%$	McD. ineq. Opt. McD.
$\mathcal{A} := \{(f,\mu) \mid f = H \text{ and } \mathbb{E}_\mu[H] \in [5.5, 7.5]\}$	$\overset{\text{num}}{=} 37.9\%$	OUQ
$\mathcal{A} \cap \left\{(f,\mu) \,\middle\vert\, \begin{array}{c}\mu\text{-median velocity} \\ = 2.45\,\text{km}\cdot\text{s}^{-1}\end{array}\right\}$	$\overset{\text{num}}{=} 30.0\%$	OUQ
$\mathcal{A} \cap \{(f,\mu) \mid \mu\text{-median obliquity} = \frac{\pi}{12}\}$	$\overset{\text{num}}{=} 36.5\%$	OUQ
$\mathcal{A} \cap \{(f,\mu) \mid \text{obliquity} = \frac{\pi}{6}\ \mu\text{-a.s.}\}$	$\overset{\text{num}}{=} 28.0\%$	OUQ

Should we compare those bounds to the true P.O.F.?

One should be careful with such comparisons in presence of asymmetric information

The real question is how to construct a selective information set A.

Selection of the most decisive experiment

$$\mathcal{A} = \mathcal{A}_{\text{safe}} \cup \mathcal{A}_{\text{unsafe}}$$

$$\mathcal{A}_{\text{safe}} = \{(\mu, f) \in \mathcal{A} \colon \mu[f(X) \geq a] \leq \epsilon\}$$

$$\mathcal{A}_{\text{unsafe}} = \{(\mu, f) \in \mathcal{A} \colon \mu[f(X) \geq a] > \epsilon\}$$

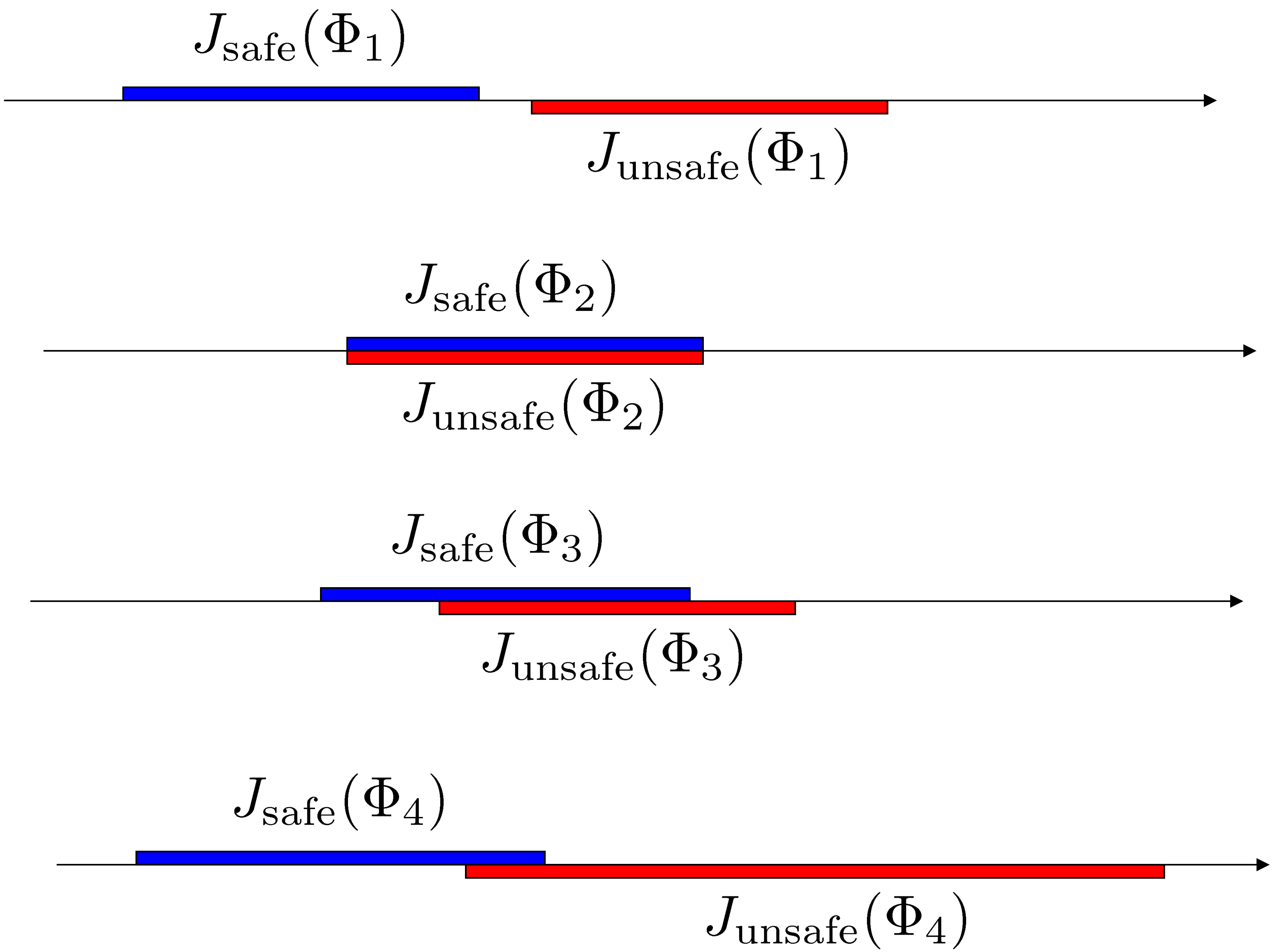
Experiments $\Phi(G, \mathbb{P})$

Ex: $\Phi_1(G, \mathbb{P}) = \mathbb{P}[X \in A]$ $\qquad \Phi_2(G, \mathbb{P}) = \mathbb{E}_{\mathbb{P}}[G]$

$$J_{\text{safe}}(\Phi) := \left[\inf_{f,\mu \in \mathcal{A}_{\text{safe}}} \Phi(f, \mu), \sup_{f,\mu \in \mathcal{A}_{\text{safe}}} \Phi(f, \mu) \right]$$

$$J_{\text{unsafe}}(\Phi) := \left[\inf_{f,\mu \in \mathcal{A}_{\text{unsafe}}} \Phi(f, \mu), \sup_{f,\mu \in \mathcal{A}_{\text{unsafe}}} \Phi(f, \mu) \right]$$

Selection of the most decisive experiment

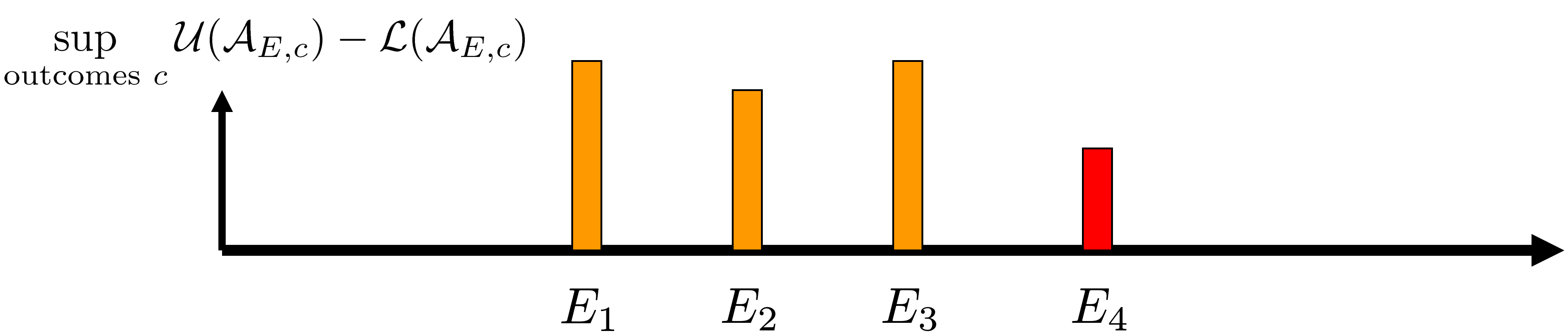


Selection of the most predictive experiment

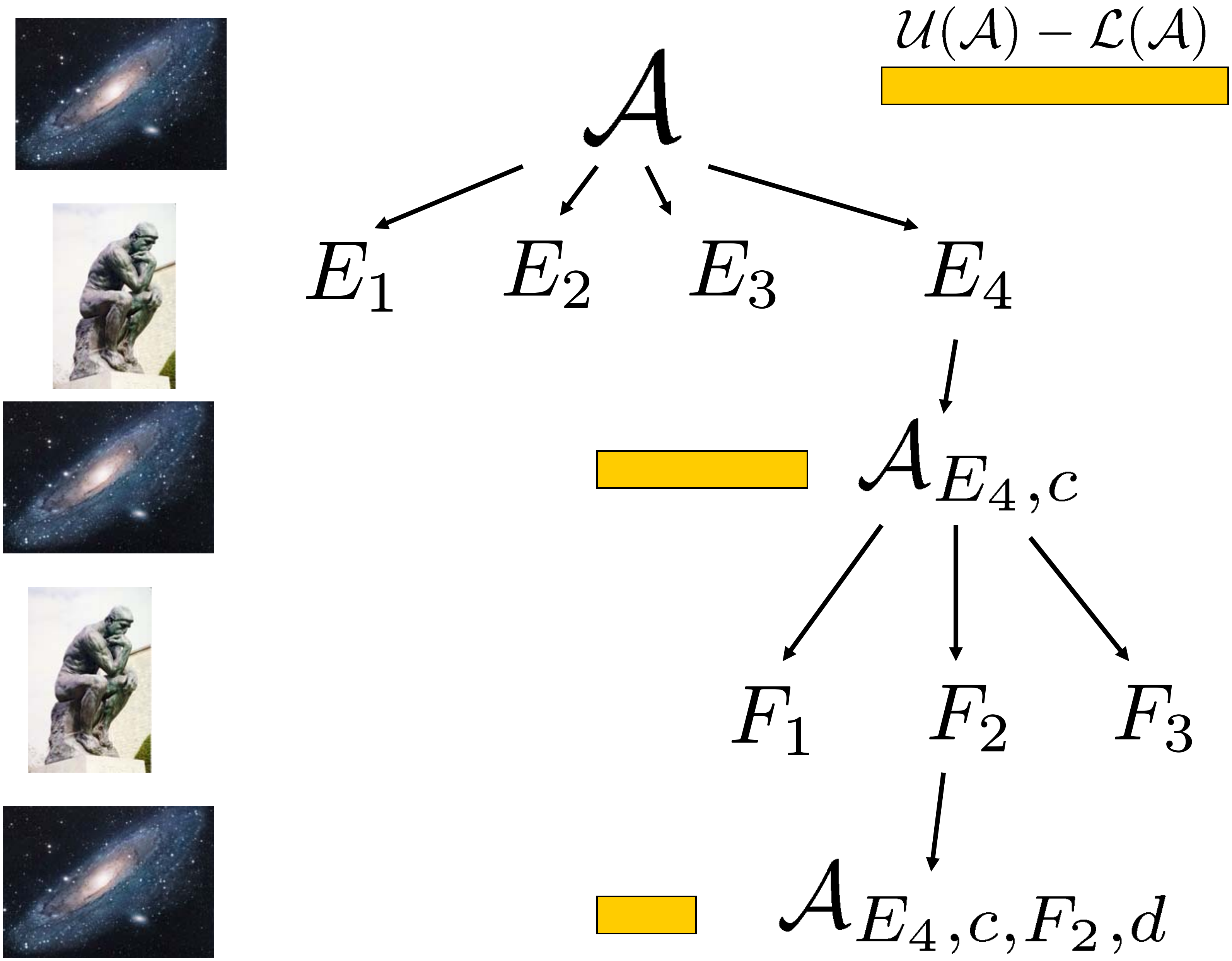
$$\mathcal{L}(\mathcal{A}) \leq \mathbb{P}[G(X) \geq a] \leq \mathcal{U}(\mathcal{A})$$

- If your objective is to have an "accurate" prediction of $\mathbb{P}[G(X) \leq \theta]$ in the sense that $\mathcal{U}(\mathcal{A}) - \mathcal{L}(\mathcal{A})$ is small, then proceed as follows:
- Let $\mathcal{A}_{E,c}$ denote those scenarios in $\mathcal{A}$ that are compatible with obtaining outcome c from experiment E.
- The experiment that is most predictive even in the worst case is defined by a minimax criterion: we seek

$$E^* \in \operatorname*{arg\,min}_{\text{experiments } E} \left(\sup_{\text{outcomes } c} \big(\mathcal{U}(\mathcal{A}_{E,c}) - \mathcal{L}(\mathcal{A}_{E,c}) \big) \right).$$



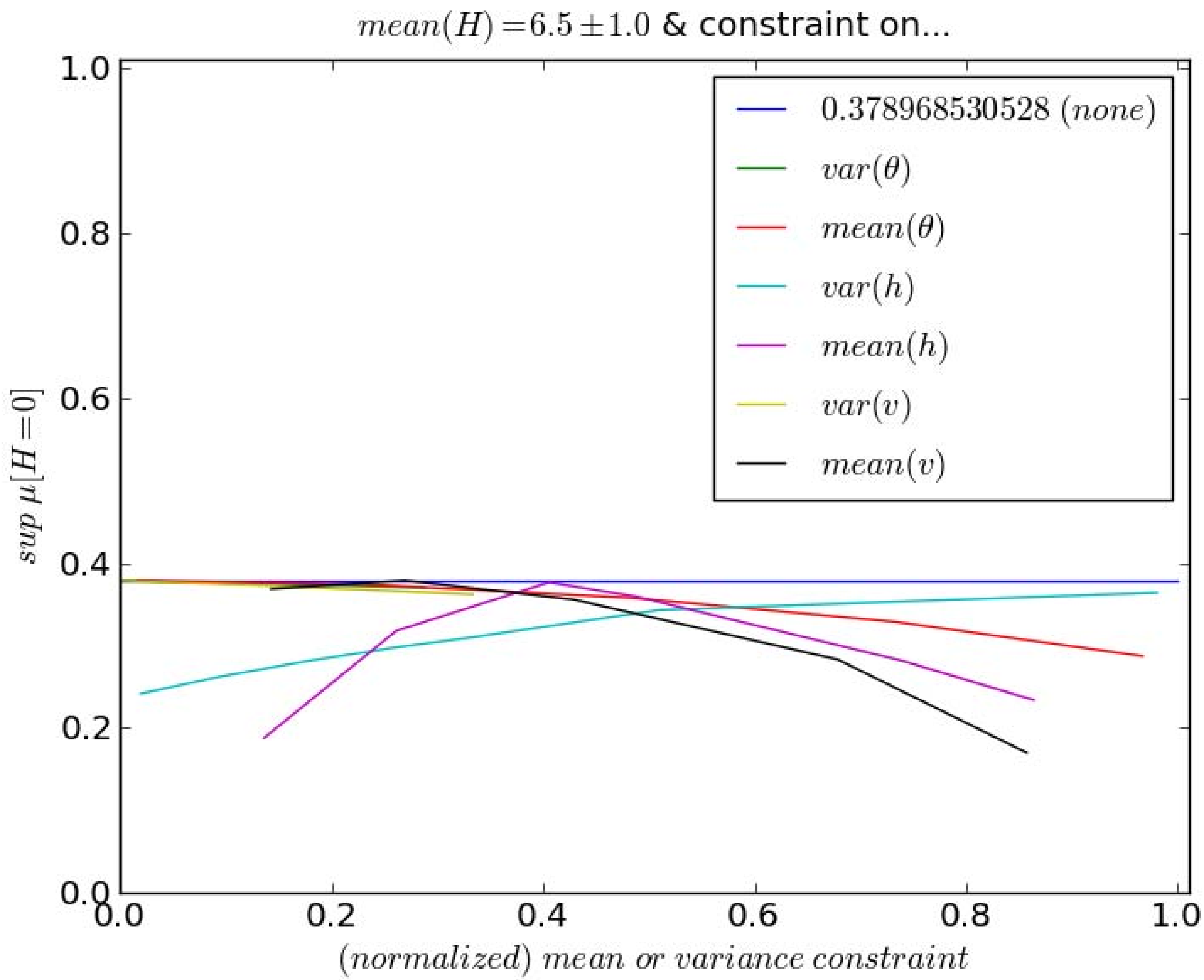
- This idea of experimental selection can be extended to plan several experiments in advance, *i.e.* to plan campaigns of experiments.



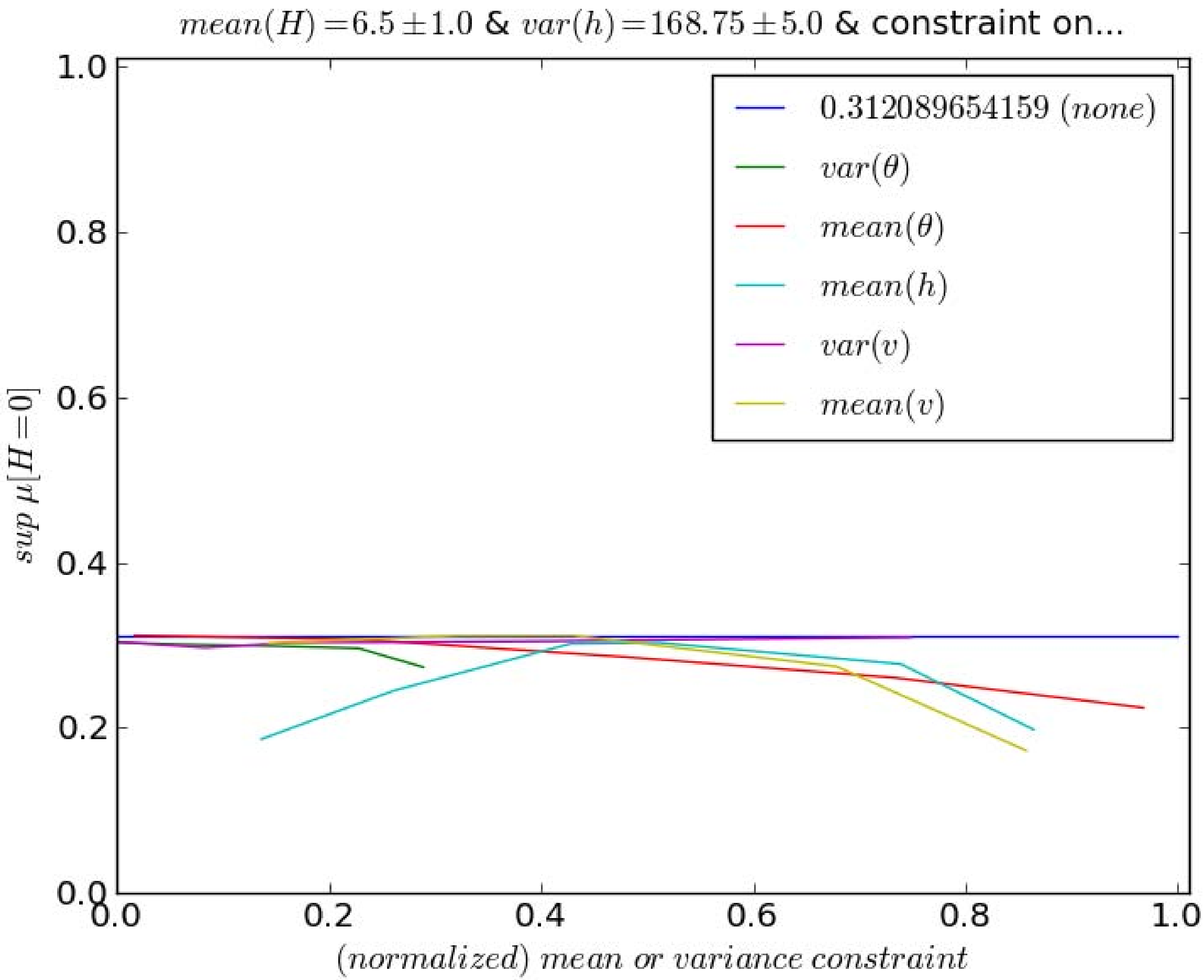
Plan several experiments in advance, i.e. campaigns of experiments

- This is a kind of infinite-dimensional *Cluedo*, played on spaces of admissible scenarios, against our lack of perfect information about reality, and made tractable by the reduction theorems.

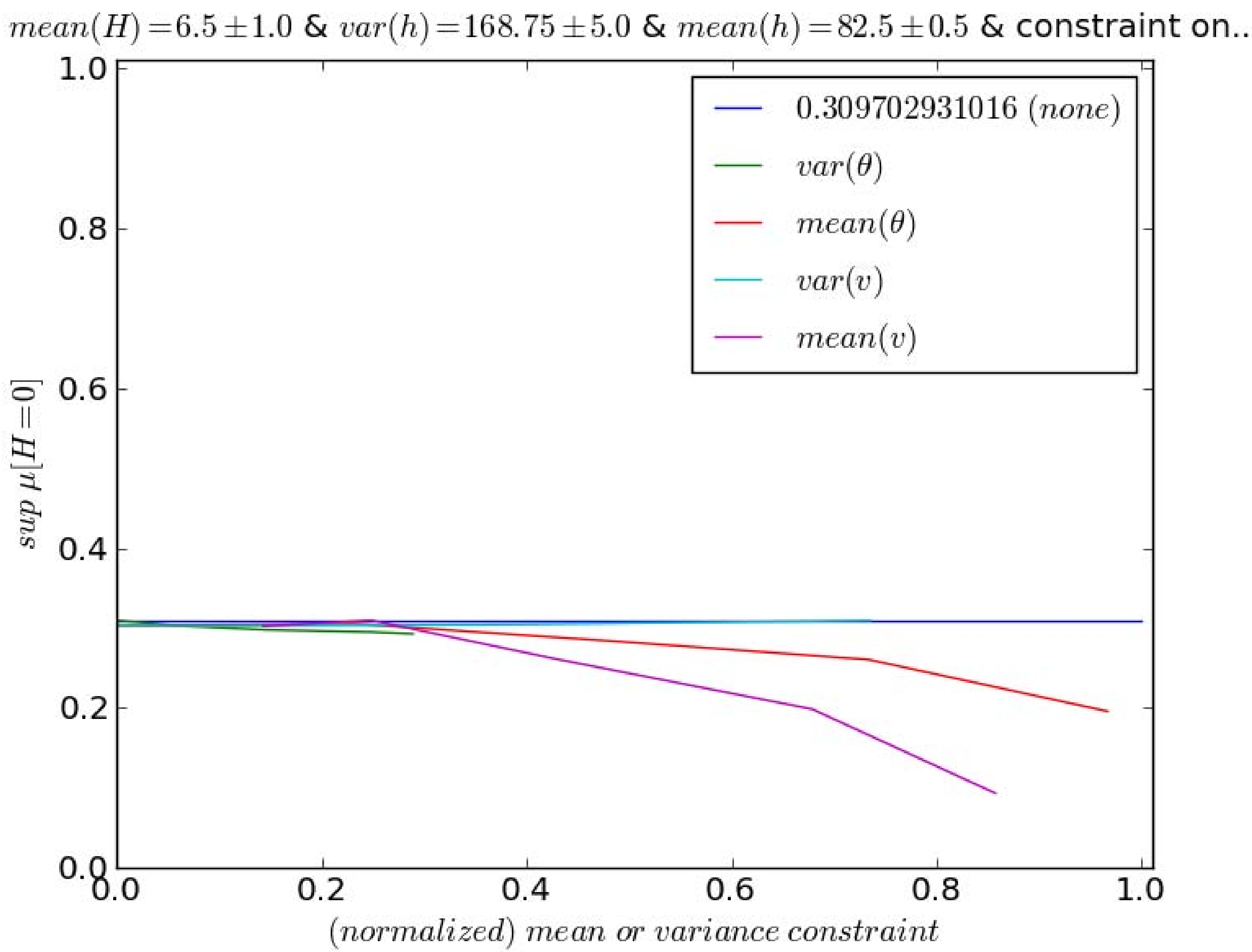
Let's play Clue



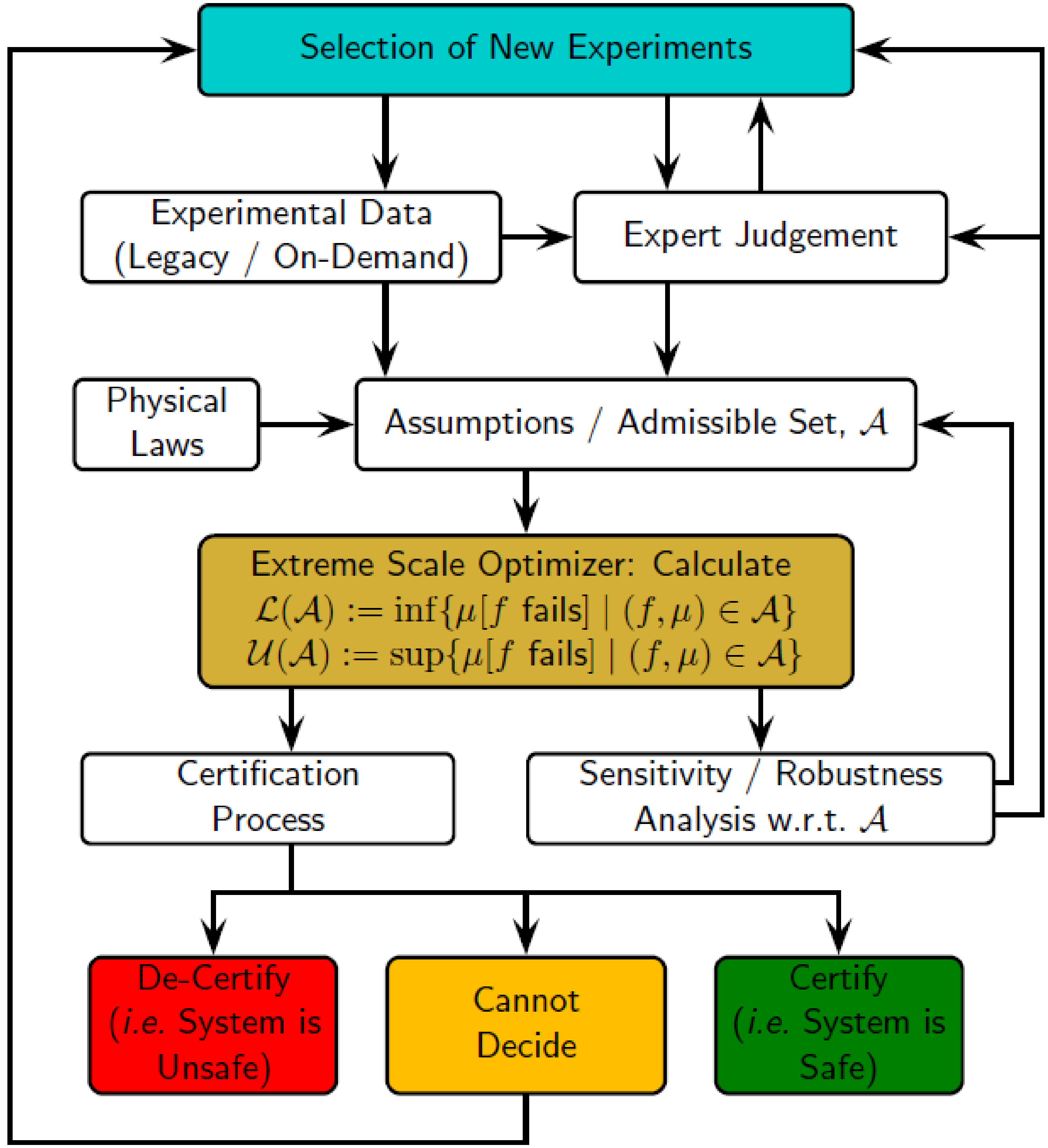
Let's play Clue



Let's play Clue



The UQ Loop



Selection of New Experiments
Experimental Data (Legacy / On-Demand)
Expert Judgement
Physical Laws
Assumptions / Admissible Set, $\mathcal{A}$
Extreme Scale Optimizer: Calculate
$\mathcal{L}(\mathcal{A}) := \inf\{\mu[f \text{ fails}] \mid (f,\mu) \in \mathcal{A}\}$
$\mathcal{U}(\mathcal{A}) := \sup\{\mu[f \text{ fails}] \mid (f,\mu) \in \mathcal{A}\}$
Certification Process
Sensitivity / Robustness Analysis w.r.t. $\mathcal{A}$
De-Certify (*i.e.* System is Unsafe)
Cannot Decide
Certify (*i.e.* System is Safe)

Bridging scales with incomplete information

$$\begin{cases} -\operatorname{div}(a(x,\omega)\nabla u(x,\omega)) = f(x,\omega), & x \in \mathcal{D} \\ u(x,\omega) = 0, & x \in \partial\mathcal{D} \end{cases}$$

$\mathbb{P}$: Measure of probability on (a, f)

You want to certify that

$$\mathbb{P}[\ln u(x_0,\omega) \geq \mathbb{E}[\ln u(x_0,\omega)] + \tau] \leq \epsilon$$

Problem

- You don't know $\mathbb{P}$

You only know

$$\boxed{\mathbb{P} \in \mathcal{A}}$$

$$\mathcal{A} \subset \{\mu \text{ meas. of prob on } (a, f)\}$$

Bridging scales with incomplete information

$$\begin{cases} -\operatorname{div}(a(x,\omega)\nabla u(x,\omega)) = f(x,\omega), & x \in \mathcal{D} \\ u(x,\omega) = 0, & x \in \partial\mathcal{D} \end{cases}$$

$\mathbb{P}$: Measure of probability on (a, f)

$\mathcal{A} \subset \{\mu \text{ meas. of prob on } (a, f)\}$

$$\mathcal{U}(\mathcal{A}) := \sup_{\mu\in\mathcal{A}} \mu[\ln u(x_0,\omega) \geq \mu[\ln u(x_0,\omega)] + \tau]$$

Bridging scales with incomplete information

$$\begin{cases} -\operatorname{div}(a(x,\omega)\nabla u(x,\omega)) = f(x,\omega), & x \in \mathcal{D} \\ u(x,\omega) = 0, & x \in \partial\mathcal{D} \end{cases}$$

$$\mathcal{A} := \left\{ \mu \,\middle|\, \begin{array}{c} a, f \ \text{independent} \ , \\ A(x) \leq a(x,\omega) \leq e^{D_1} A(x), \\ F(x) \leq f(x,\omega) \leq e^{D_2} F(x) \end{array} \right\}$$

$$\begin{gathered} D_1, D_2 \geq 0, \\ A, F \in L^\infty(\mathcal{D}), \\ \operatorname{essinf} A > 0,\ F \geq 0,\ \int_{\mathcal{D}} F > 0 \end{gathered}$$

Non propagation of information across scales

$$\mathcal{U}(\mathcal{A}) := \sup_{\mu\in\mathcal{A}} \mu[\ln u(x_0,\omega) \geq \mu[\ln u(x_0,\omega)] + \tau]$$

$$\mathcal{U}(\mathcal{A}_{MD}) = \begin{cases} 0 & \text{if} \quad D_1 + D_2 \leq \tau \\ \frac{(D_1+D_2-\tau)^2}{4D_1D_2} & \text{if} \quad |D_1 - D_2| \leq \tau \leq D_1 + D_2 \\ 1 - \frac{\tau}{\max(D_1,D_2)} & \text{if} \quad 0 \leq \tau \leq |D_1 - D_2| \end{cases}$$

Corollary If $D_1 \geq \tau + D_2$, then

$$\mathcal{U}(\mathcal{A}_{MD})(\tau, D_1, D_2) = \mathcal{U}(\mathcal{A}_{MD})(\tau, D_1, 0)$$

Corollary If $D_2 \geq \tau + D_1$, then

$$\mathcal{U}(\mathcal{A}_{MD})(\tau, D_1, D_2) = \mathcal{U}(\mathcal{A}_{MD})(\tau, 0, D_2)$$

Bridging scales with incomplete information

$$\begin{cases} -\operatorname{div}(a(x,\omega)\nabla u(x,\omega)) = f(x,\omega), & x \in \mathcal{D} \\ u(x,\omega) = 0, & x \in \partial\mathcal{D} \end{cases}$$

$$\mathcal{A} := \left\{ \mu \,\middle|\, \begin{array}{c} a = a_1 a_2, \\ a_1, a_2 \text{ independent }, \\ a_1 \text{ smooth }, \|\partial_k a_1\|_{L^\infty} \leq 1 \\ a_2 \text{ periodic of period }, \delta \ll 1 \\ A_1(x) \leq a_1(x,\omega) \leq e^{D_1} A_1(x), \\ A_2(x) \leq a_2(x,\omega) \leq e^{D_2} A_2(x), \end{array} \right\}$$

Non propagation of information across scales

$$\mathcal{U}(\mathcal{A}) := \sup_{\mu \in \mathcal{A}} \mu[\ln u(x_0, \omega) \geq \mu[\ln u(x_0, \omega)] + \tau]$$

$$\mathcal{U}(\mathcal{A}_{MD}) = \begin{cases} 0 & \text{if} \quad D_1 + D_2 \leq \tau \\ \frac{(D_1 + D_2 - \tau)^2}{4 D_1 D_2} & \text{if} \quad |D_1 - D_2| \leq \tau \leq D_1 + D_2 \\ 1 - \frac{\tau}{\max(D_1, D_2)} & \text{if} \quad 0 \leq \tau \leq |D_1 - D_2| \end{cases}$$

Corollary If $D_1 \geq \tau + D_2$, then

$$\mathcal{U}(\mathcal{A}_{MD})(\tau, D_1, D_2) = \mathcal{U}(\mathcal{A}_{MD})(\tau, D_1, 0)$$

List what you know and what you assume on $\mathbb{P}$ and G

Experimental data

Physical laws

Expert Judgment

Information
Assumptions
on G and P