A calculus for the optimal quantification of uncertainties

Houman Owhadi

- Bayesian Brittleness.
  H. Owhadi, C. Scovel, T. Sullivan. 2013. arXiv:1304.6772
- Brittleness of Bayesian inference and new Selberg formulas.
  H. Owhadi, C. Scovel. 2013 arXiv:1304.7046
- Optimal Uncertainty Quantification. H. Owhadi, C. Scovel, T. Sullivan, M. McKerns and M. Ortiz. SIAM Review, 2013

Kavli Royal Society 2014

Reduction calculus

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

$$\mathcal{G}(f,\mu) \le 0 \Leftrightarrow \begin{cases} n' \text{ generalized moment constraints on } \mu, & \mathbb{E}_\mu[\varphi_j^f] \le 0 \\ n_k \text{ generalized moment constraints on } \mu_k, & \mathbb{E}_{\mu_k}[\psi_{k,j}^f] \le 0 \end{cases}$$

Theorem

$$\sup_{(f,\mu)\in\mathcal{A}} \mathbb{E}_\mu[q_f] = \sup_{(f,\mu)\in\mathcal{A}_\Delta} \mathbb{E}_\mu[q_f]$$

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

$$\sup_{(f,\mu)\in\mathcal{A}} \mathbb{E}_\mu[q_f]$$

Non-convex and infinite dimensional optimization problems

Can be considered as a generalization of classical Chebyshev inequalities

Connection between Chebyshev inequalities and optimization theory

- Classical Markov-Krein theorem (Karlin, Studden, 1958)
- Karlin, Studden (1966, Tchebycheff systems with applications in analysis and statistics)
- Marshall & Olkin (1979, Inequalities: Theory of majorization and its applications)
- Dynkin (1978, Sufficient statistics & extreme points)
- Karr (1983, Extreme points of probability measures with applications)
- Artzner et al (1997, risk measures, value at risk, etc…)
- Betsimas & Popescu (2008, convex optimization approach to inequalities in prob. theo.)

Our proof rely on a form of Linear programming in infinite dimensional spaces

- 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)

Reduction of optimization variables

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

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

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

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

Example: Optimal concentration inequality

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

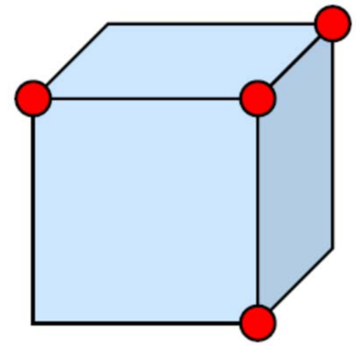
$$\mathrm{Osc}_i(f) := \sup_{(x_1,\ldots,x_m)\in\mathcal{X}} \sup_{x_i'\in\mathcal{X}_i} \big(f(\ldots,x_i,\ldots) - f(\ldots,x_i',\ldots)\big).$$

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

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

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



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

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

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

Theorem

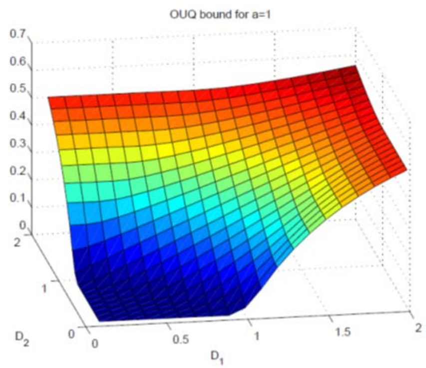
$$\mathcal{U}(\mathcal{A}_{MD}) = \mathcal{U}(\mathcal{A}_{\mathcal{C}})$$

Explicit Solution m=2

Theorem $m = 2$

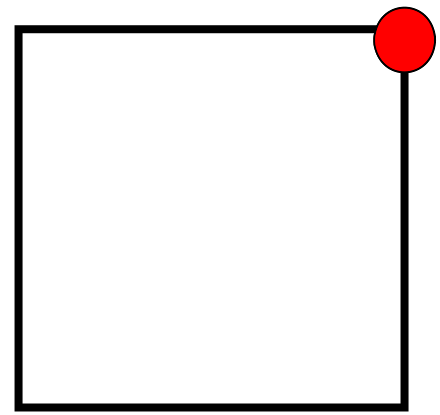
$$\mathcal{U}(\mathcal{A}_{MD}) = \begin{cases} 0 & \text{if} \quad D_1 + D_2 \leq a \\ \frac{(D_1+D_2-a)^2}{4D_1D_2} & \text{if} \quad |D_1 - D_2| \leq a \leq D_1 + D_2 \\ 1 - \frac{a}{\max(D_1,D_2)} & \text{if} \quad 0 \leq a \leq |D_1 - D_2| \end{cases}$$

OUQ bound a=1



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

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



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

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

Reduction calculus with measures over measures

$$\mathcal{M}(\mathcal{X}) \supset \mathcal{A} \xrightarrow{\quad \Psi \quad} \mathcal{Q}$$

Polish space

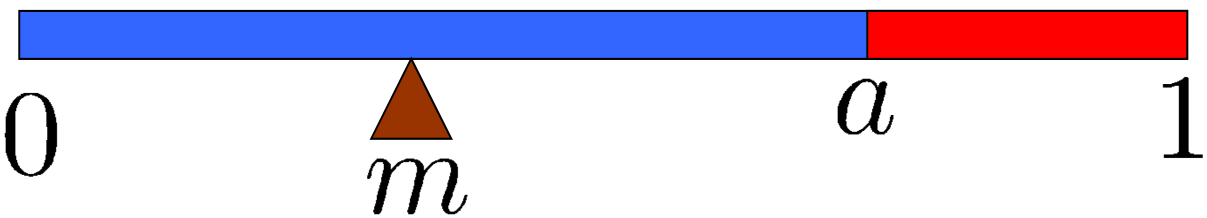
$$\mathcal{M}(\mathcal{A}) \supset \Pi \xleftarrow{\quad \Psi^{-1} \quad} \mathfrak{Q} \subset \mathcal{M}(\mathcal{Q})$$

Theorem

$$\sup_{\pi \in \Psi^{-1}\mathfrak{Q}} \mathbb{E}_{\mu\sim\pi}\Big[\Phi(\mu)\Big] = \sup_{\mathbb{Q}\in\mathfrak{Q}} \Big[\mathbb{E}_{q\sim\mathbb{Q}}\big[\sup_{\mu\in\Psi^{-1}(q)} \Phi(\mu)\big]\Big]$$

A simple example

10,000 children are given one pound of play-doh. On average, how much mass can they put above a While, on average, keeping the seesaw balanced around m?

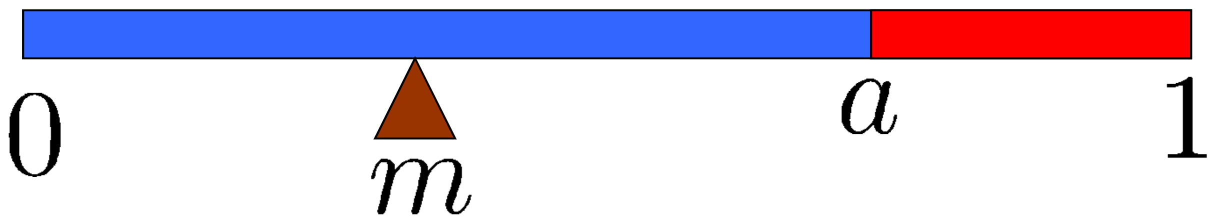


Paul is given one pound of play-doh. What can you say about how much mass he is putting above a if all you have is the belief that he is keeping the seesaw balanced around m?

What is the least upper bound on

$$\mathbb{E}_{\mu\sim\pi}\left[\mu[X \geq a]\right]$$

If all you know is $\mathbb{E}_{\mu\sim\pi}\left[\mathbb{E}_{\mu}[X]\right] = m$ **?**



$$\mu \in \mathcal{A} := \mathcal{M}\big([0,1]\big)$$

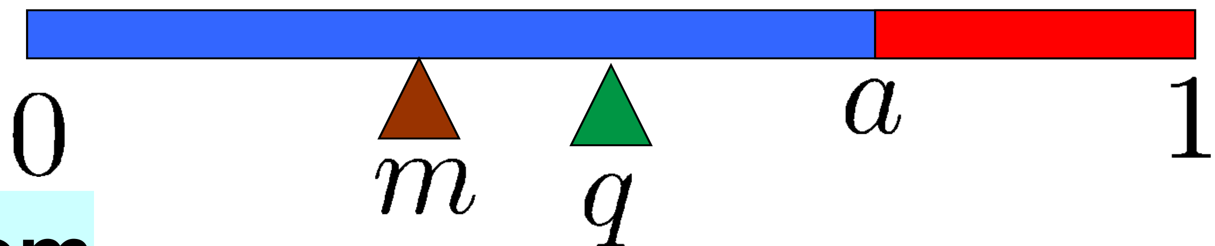
Answer

$$\sup_{\pi\in\Pi} \mathbb{E}_{\mu\sim\pi}\left[\mu[X \geq a]\right]$$

$$\Pi := \left\{\pi \in \mathcal{M}(\mathcal{A}) \,:\, \mathbb{E}_{\mu\sim\pi}\left[\mathbb{E}_{\mu}[X]\right] = m\right\}$$

$$\sup_{\pi \in \Pi} \mathbb{E}_{\mu \sim \pi} \left[\mu[X \geq a] \right]$$

$$\Pi := \left\{ \pi \in \mathcal{M}(\mathcal{M}([0,1])) \; : \; \mathbb{E}_{\mu \sim \pi}\left[\mathbb{E}_\mu[X]\right] = m \right\}$$

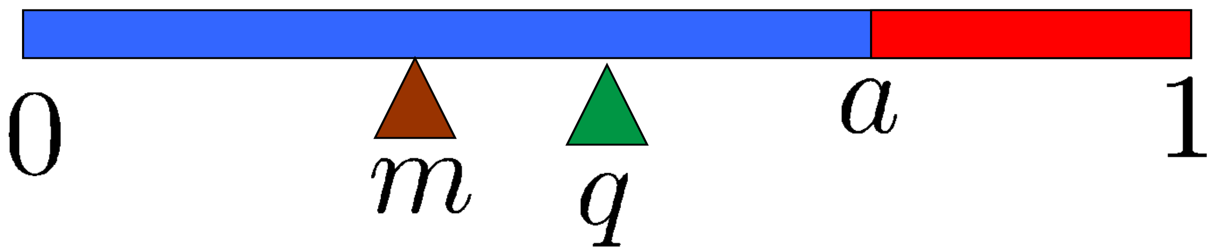


Theorem

$$\sup_{\pi \in \Pi} \mathbb{E}_{\mu \sim \pi} \left[\mu[X \geq a] \right] = \sup_{\mathbb{Q} \in \mathcal{M}([0,1]) \,:\, \mathbb{E}_{\mathbb{Q}}[q] = m} \mathbb{E}_{q \sim \mathbb{Q}} \left[\sup_{\mu \in \mathcal{M}([0,1]) \,:\, \mathbb{E}_\mu[X] = q} \mu[X \geq a] \right]$$

$$\sup_{\pi \in \Pi} \mathbb{E}_{\mu \sim \pi}\big[\mu[X \geq a]\big]$$

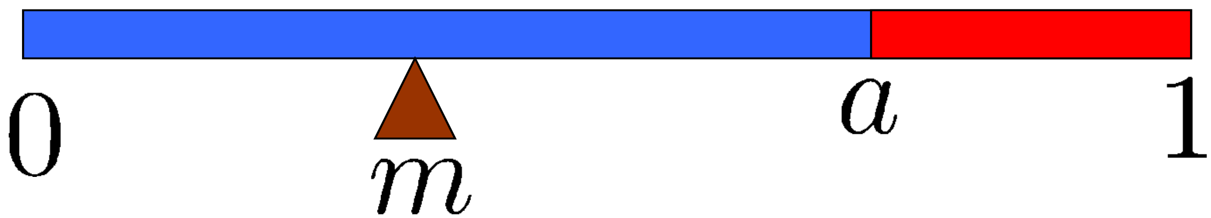
$$\Pi := \Big\{\pi \in \mathcal{M}\big(\mathcal{M}([0,1])\big) \;:\; \mathbb{E}_{\mu \sim \pi}\big[\mathbb{E}_{\mu}[X]\big] = m\Big\}$$



$$\sup_{\pi \in \Pi} \mathbb{E}_{\mu \sim \pi}\big[\mu[X \geq a]\big] = \sup_{\mathbb{Q} \in \mathcal{M}([0,1]) \,:\, \mathbb{E}_{\mathbb{Q}}[q] = m} \mathbb{E}_{q \sim \mathbb{Q}}\left[\min\left(\frac{q}{a}, 1\right)\right]$$

$$\sup_{\pi \in \Pi} \mathbb{E}_{\mu \sim \pi}\left[\mu[X \geq a]\right]$$

$$\Pi := \left\{\pi \in \mathcal{M}(\mathcal{M}([0,1])) \;:\; \mathbb{E}_{\mu \sim \pi}\left[\mathbb{E}_{\mu}[X]\right] = m\right\}$$



$$\sup_{\pi \in \Pi} \mathbb{E}_{\mu \sim \pi}\left[\mu[X \geq a]\right] = \frac{m}{a}$$

Can this form of calculus in infinite dimensional spaces facilitate the process of scientific discovery?

New Reproducing Kernel Hilbert Spaces and Selberg Integral formulas

Forrester and Warnaar 2008

The importance of the Selberg integral

“Used to prove outstanding conjectures in Random matrix theory and cases of the Macdonald conjectures”

“Central role in random matrix theory, Calogero-Sutherland quantum many-body systems, Knizhnik-Zamolodchikov equations, and multivariable orthogonal polynomial theory”

The truncated moment problem

$$\mathcal{M}[0,1] \xrightarrow{\quad \Psi \quad} \mathbb{R}^k$$

$$\mu \longrightarrow \left(\mathbb{E}_{X\sim\mu}[X], \mathbb{E}_{X\sim\mu}[X^2], \ldots, \mathbb{E}_{X\sim\mu}[X^k]\right)$$

$$\text{Study of the geometry of } M_k := \Psi\big(\mathcal{M}([0,1])\big)$$

P. L. Chebyshev
1821-1894

A. A. Markov
1856-1922

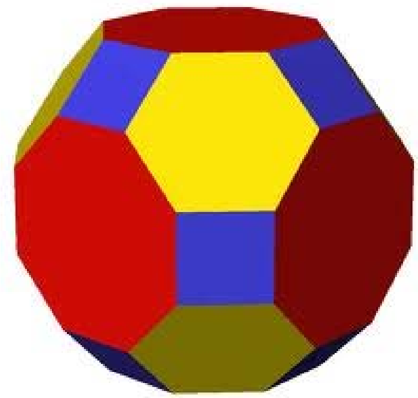
M. G. Krein
1907-1989

$$\mathcal{M}[0,1] \xrightarrow{\quad \Psi \quad} \mathbb{R}^k$$

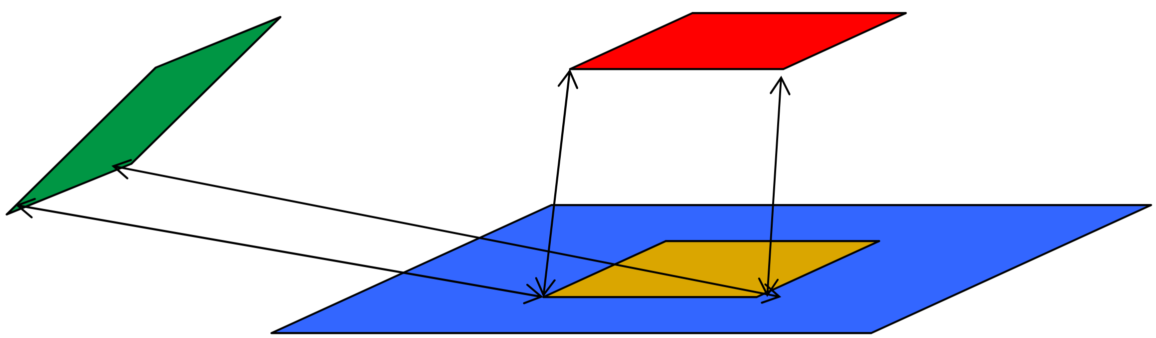
$$\mu \longrightarrow \left(\mathbb{E}_{X\sim\mu}[X], \mathbb{E}_{X\sim\mu}[X^2], \ldots, \mathbb{E}_{X\sim\mu}[X^k]\right)$$

$$\boxed{M_k := \Psi\big(\mathcal{M}([0,1])\big)}$$

Origin of these new Selberg integral formulas and new RKHS



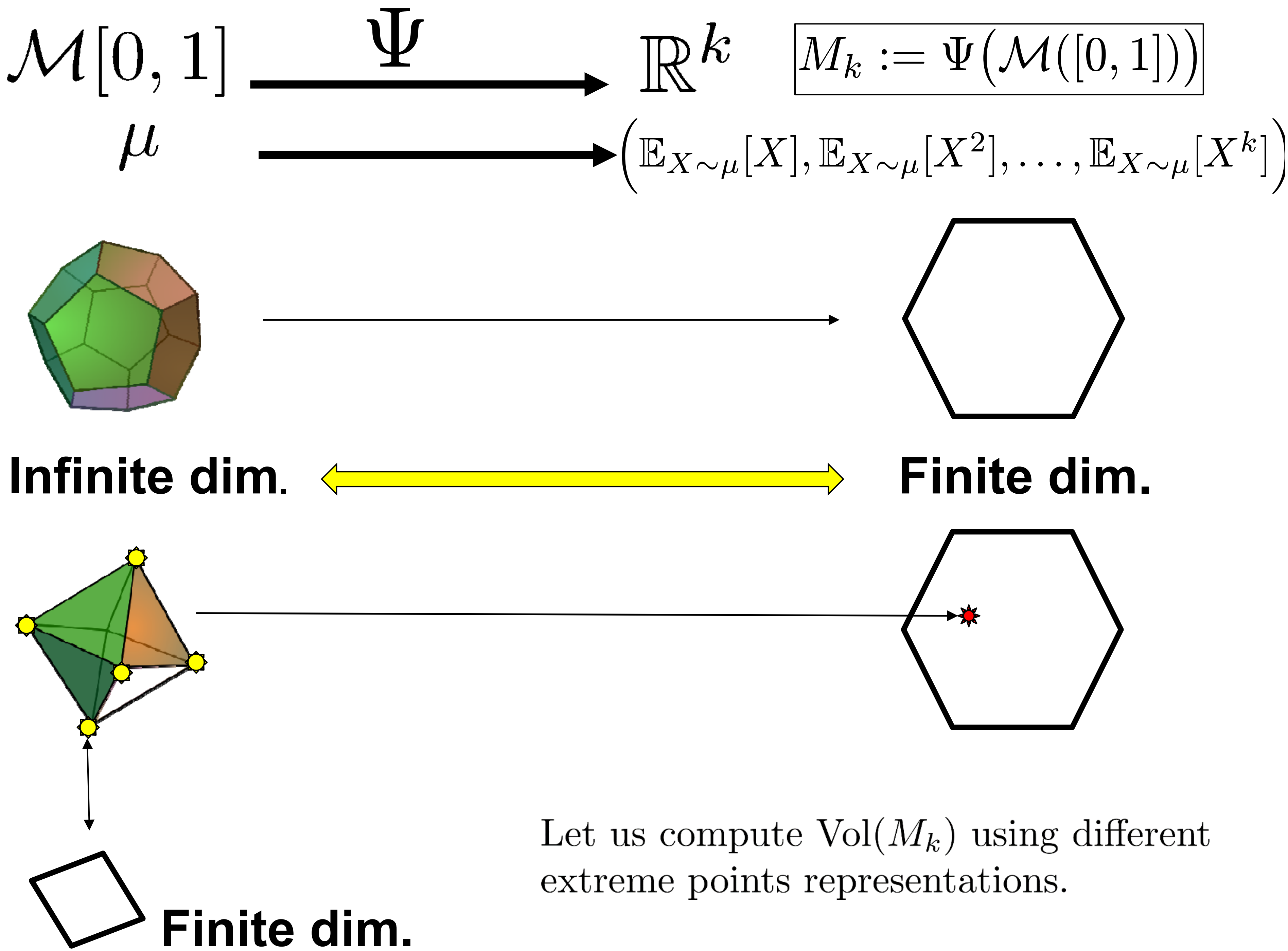
Compute $\mathrm{Vol}(M_k)$ using different (finite-dimensional) representations in $\mathcal{M}\big([0,1]\big)$



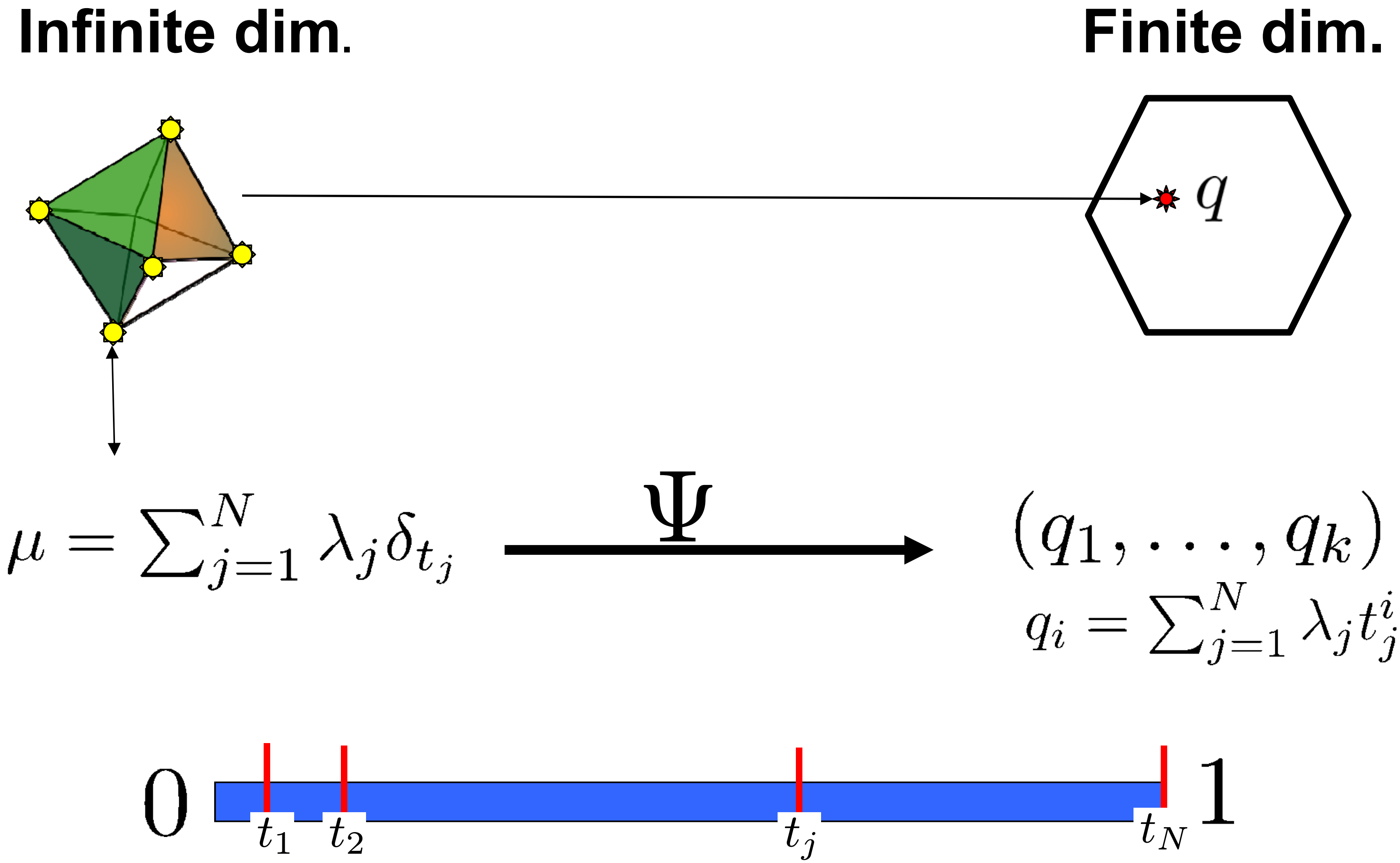
Infinite dim.

↕

Finite dim.



$\mathcal{M}[0,1]$
Ψ
$\mathbb{R}^k$
$M_k := \Psi(\mathcal{M}([0,1]))$
μ
$\left(\mathbb{E}_{X\sim\mu}[X], \mathbb{E}_{X\sim\mu}[X^2], \ldots, \mathbb{E}_{X\sim\mu}[X^k]\right)$
Infinite dim.
Finite dim.
Finite dim.
Let us compute $\mathrm{Vol}(M_k)$ using different extreme points representations.



Infinite dim.
Finite dim.
q
$\mu = \sum_{j=1}^{N} \lambda_j \delta_{t_j}$
Ψ
$(q_1, \ldots, q_k)$
$q_i = \sum_{j=1}^{N} \lambda_j t_j^i$
0
1
t_1
t_2
t_j
t_N

$$\mu = \sum_{j=1}^{N} \lambda_j \delta_{t_j}$$

Index $i(\mu)$: Number of support points of μ

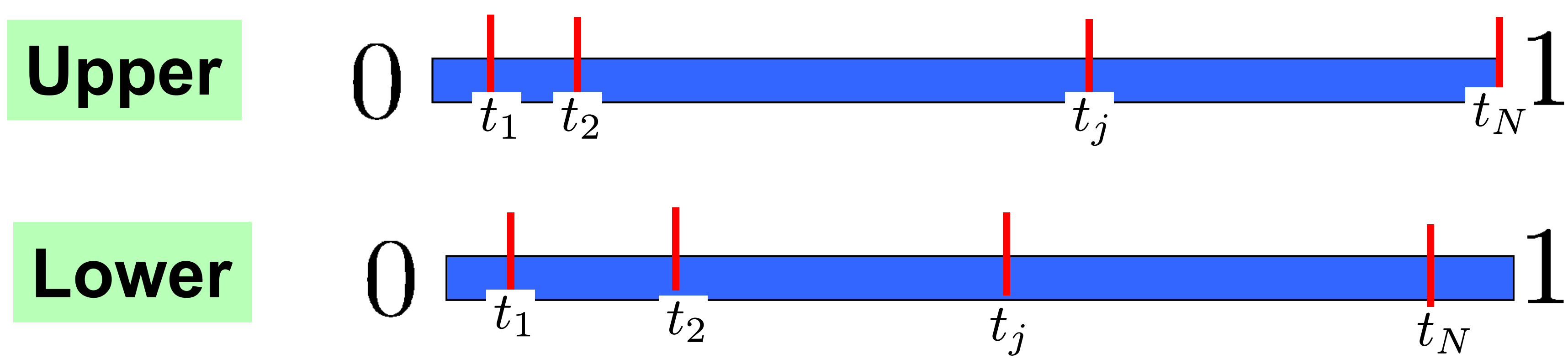
Counting interior points with weight 1 and boundary points with weight $\frac{1}{2}$

μ is called

- principal if $i(\mu) = \frac{k+1}{2}$
- canonical if $i(\mu) = \frac{k+2}{2}$
- upper if support points include 1
- lower if support points do not include 1

Theorem

Every point $q \in \mathrm{Int}(M_k)$ has a unique upper and lower principal representation.



$\mathrm{Vol}(M_{2m-1})$ using Upper Rep. $=$ $\mathrm{Vol}(M_{2m-1})$ using Lower Rep.

$$\frac{1}{(m-1)!}S_{m-1}(3,3,2) = \frac{1}{m!}S_m(1,1,2)$$

$\mathrm{Vol}(M_{2m})$ using Upper Rep. $=$ $\mathrm{Vol}(M_{2m})$ using Lower Rep.

$$S_m(1,3,2) = S_m(3,1,2)$$

Selberg Identities

$$S_n(\alpha,\beta,\gamma) = \prod_{j=0}^{n-1} \frac{\Gamma(\alpha+j\gamma)\Gamma(\beta+j\gamma)\Gamma(1+(j+1)\gamma)}{\Gamma(\alpha+\beta+(n+j-1)\gamma)\Gamma(1+\gamma)}$$

$$S_n(\alpha,\beta,\gamma) := \int_{[0,1]^n} \prod_{j=1}^{n} t_j^{\alpha-1}(1-t_j)^{\beta-1}|\Delta(t)|^{2\gamma}dt\,.$$

$$\Delta(t) := \prod_{j<k}(t_k - t_j)$$

Brittleness of Bayesian inference and new Selberg formulas. H. Owhadi and C. Scovel. Communications in Mathematical Sciences (2015). arXiv:1304.7046

$$\mu = \sum_{j=1}^{N} \lambda_j \delta_{t_j}$$

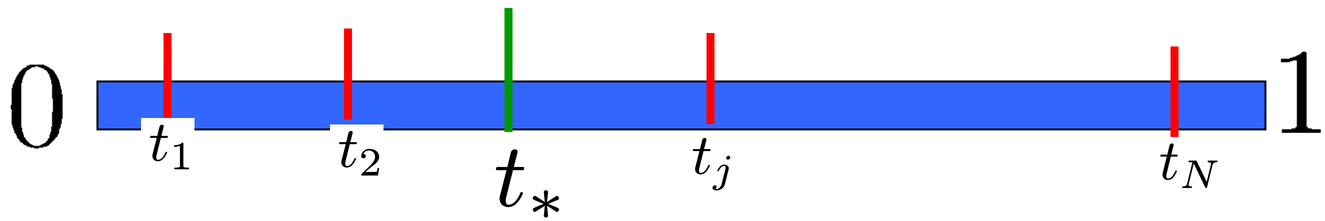
Index $i(\mu)$: Number of support points of μ

Counting interior points with weight 1 and boundary points with weight $\frac{1}{2}$

μ is called
- principal if $i(\mu) = \frac{k+1}{2}$
- canonical if $i(\mu) = \frac{k+2}{2}$
- upper if support points include 1
- lower if support points do not include 1

Theorem

For $t_* \in (0,1)$, every point $q \in \mathrm{Int}(M_k)$ has a unique canonical representation whose support contains t_*. When $t_* = 0$ or 1, there exists a unique principal representation whose support contains t_*.



New Reproducing Kernel Hilbert Spaces and Selberg Integral formulas related to the Markov-Krein representations of moment spaces.

$$\mathcal{M}[0,1] \xrightarrow{\quad\Psi\quad} [0,1]^k$$

$$\mu \longrightarrow \left(\mathbb{E}_{X\sim\mu}[X], \mathbb{E}_{X\sim\mu}[X^2], \ldots, \mathbb{E}_{X\sim\mu}[X^k]\right)$$

$$\int_{I^m} \Sigma t^{-1} \cdot \prod_{j=1}^m t_j^2(1-t_j)^2 \Delta_m^4(t)dt = \frac{S_m(5,1,2)-S_m(3,3,2)}{2}$$

$$\int_{I^m} \Sigma t^{-1} \cdot \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t)dt = \frac{m}{2} S_{m-1}(5,3,2)$$

$$\Delta_m(t) := \prod_{j<k} (t_k - t_j) \quad I := [0,1]$$

$$(\Sigma\phi)(t) := \sum_{j=1}^m \phi(t_j), \quad t \in I^m$$

$$S_n(\alpha,\beta,\gamma) = \prod_{j=0}^{n-1} \frac{\Gamma(\alpha+j\gamma)\Gamma(\beta+j\gamma)\Gamma(1+(j+1)\gamma)}{\Gamma(\alpha+\beta+(n+j-1)\gamma)\Gamma(1+\gamma)}$$

$$e_j(t) := \sum_{i_1 < \cdots < i_j} t_{i_1} \cdots t_{i_j}$$

Π_0^n: n-th degree polynomials which vanish on the boundary of $[0,1]$

$M_n \subset \mathbb{R}^n$: set of $q = (q_1, \ldots, q_n) \in \mathbb{R}^n$ such that there exists a probability measure μ on $[0,1]$ with $\mathbb{E}_\mu[X^i] = q_i$ with $i \in \{1, \ldots, n\}$.

Theorem

Bi-orthogonal systems of Selberg Integral formulas

Consider the basis of Π_0^{2m-1} consisting of the associated Legendre polynomials $Q_j, j = 2, .., 2m-1$ of order 2 translated to the unit interval I. For $k = 2, .., 2m-1$ define

$$a_{jk} := \frac{(j+k+k^2)\Gamma(j+2)\Gamma(j)}{\Gamma(j+k+2)\Gamma(j-k+1)}, \quad k \leq j \leq 2m-1$$

$$\tilde{h}_k(t) := \sum_{j=k}^{2m-1} (-1)^{j+1} a_{jk} e_{2m-1-j}(t,t) \, .$$

Then for $j = k \bmod 2$, $j, k = 2, .., 2m-1$, we have

$$\int_{I^{m-1}} \tilde{h}_k(t) \Sigma Q_j(t) \prod_{j'=1}^{m-1} t_{j'}^2 \cdot \Delta_{m-1}^4(t) dt = Vol(M_{2m-1})(2m-1)!(m-1)! \frac{(k+2)!}{(8k+4)(k-2)!} \delta_{jk} \, .$$