Can discovery be computed?

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
- H. Owhadi, Bayesian Numerical Homogenization (2014). arXiv:1406.6668
- H Owhadi, C. Scovel, Scientific Computation of Optimal Statistical Estimators, to appear

Arlington 2014

Quantity of Interest $\Phi(\mu^\dagger) = \mu^\dagger[X \geq a]$

$\mu^\dagger$: Unknown or partially known measure of probability on $\mathbb{R}$

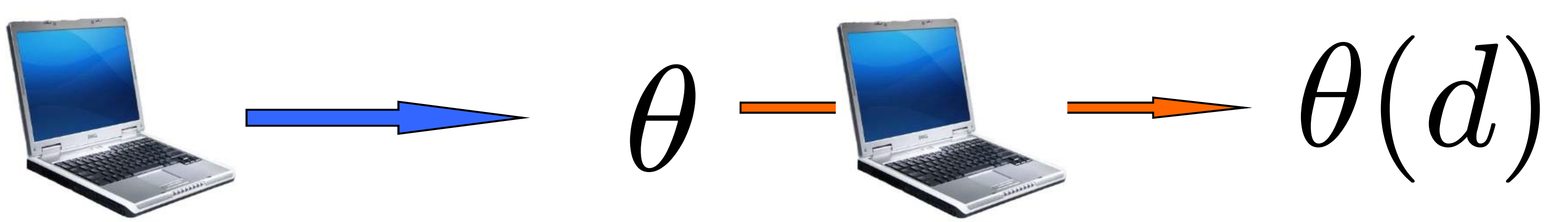
You know $\mu^\dagger \in \mathcal{A}$

You observe $d = (d_1, \ldots, d_n) \in \mathbb{R}^n$

n i.i.d samples from $\mu^\dagger$

Problem:

Compute the best estimate of $\Phi(\mu^\dagger)$



Player A

Chooses $\mu^\dagger \in \mathcal{A}$

Player B

Chooses θ

$$\mathcal{E}(\mu^\dagger, \theta)$$

Mean squared error

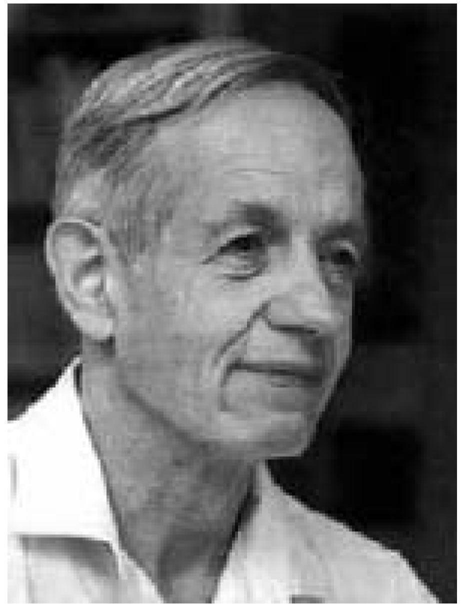
$$\mathcal{E}(\mu^\dagger, \theta) = \mathbb{E}_{d \sim (\mu^\dagger)^n}\left[\left[\theta(d) - \Phi(\mu^\dagger)\right]^2\right]$$

Confidence error

$$\mathcal{E}(\mu^\dagger, \theta) = \mathbb{P}_{d \sim (\mu^\dagger)^n}\left[\left|\theta(d) - \Phi(\mu^\dagger)\right| \geq r\right]$$

Game theory and statistical decision theory

John Von Neumann



John Nash

Abraham Wald

The best strategy is to play at random

Obtained by finding the worst prior in the Bayesian class of estimators

Player A

Player B

Chooses $\mu^\dagger \in \mathcal{A}$

Chooses θ

$$\mathcal{E}(\mu^\dagger, \theta)$$

Best strategy for A $\quad \mu^\dagger \sim \pi_A \in \mathcal{M}(\mathcal{A})$

The best strategy for B

$$\theta_{\pi_B}(d) = \mathbb{E}_{\mu \sim \pi_B,\, d' \sim \mu^n}\left[\Phi(\mu) \middle| d' = d\right]$$

The best strategy for A and B = worst prior for B

$$\max_{\pi \in \mathcal{M}(\mathcal{A})} \mathbb{E}_{\mu \sim \pi}\left[\mathcal{E}(\mu, \theta_\pi)\right]$$

Reduction calculus with measures over measures

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

Polish space

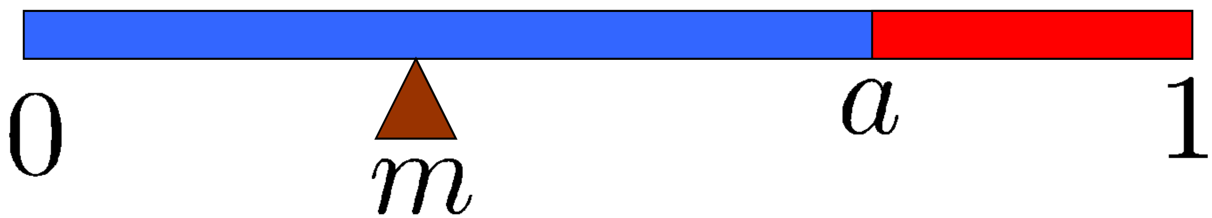
$$\mathcal{M}(\mathcal{A}) \supset \Pi \xleftarrow{\quad\Psi^{-1}\quad} \mathfrak{Q} \subset \mathcal{M}(\mathcal{Q})$$

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

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

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

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

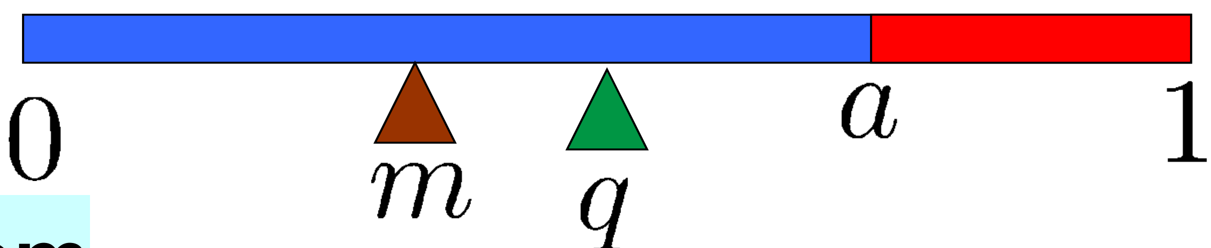
Answer

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

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

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

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

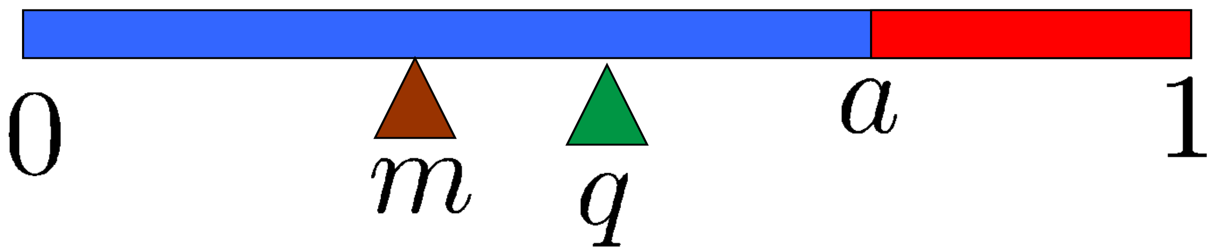


Theorem

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

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

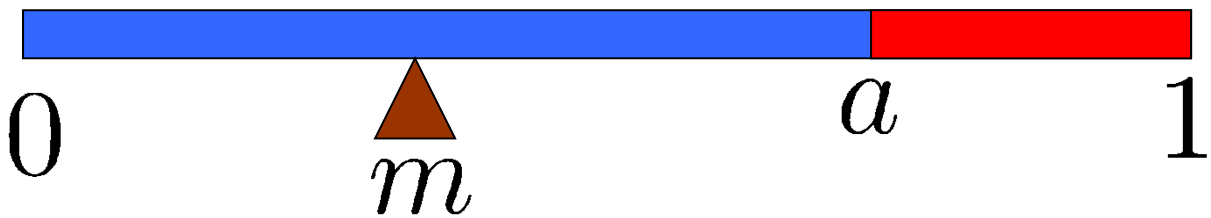
$$\Pi := \left\{\pi \in \mathcal{M}\big(\mathcal{M}([0,1])\big) \;:\; \mathbb{E}_{\mu\sim\pi}\left[\mathbb{E}_\mu[X]\right] = m\right\}$$



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

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

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



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

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

Identification of accurate bases for numerical homogenization with optimal recovery properties

Identification of New Reproducing Kernel Hilbert Spaces and Selberg Integral formulas

Bayesian Numerical Homogenization

(1) $$\begin{cases} -\operatorname{div}(a\nabla u) = g, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases}$$

$$\Omega \subset \mathbb{R}^d \quad \partial\Omega \text{ is piec. Lip.}$$

$$a \text{ unif. ell.} \quad a_{i,j} \in L^\infty(\Omega)$$

$d \leq 3$

We want to homogenize (1)

We need $g \in L^2(\Omega)$

$$\begin{cases} -\operatorname{div}(a\nabla u) = g, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases}$$

$$g \longrightarrow u$$

$$\mathcal{H}^{-1}(\Omega) \longrightarrow \mathcal{H}_0^1(\Omega)$$

$$L^2(\Omega) \longrightarrow V$$

$$V \subset\subset \mathcal{H}_0^1(\Omega) \qquad V \sim \mathcal{H}^2(\Omega)$$

Q: How to approximate V with a finite dimensional space?

Numerical Homogenization Approach

Work hard to find good basis functions

Harmonic Coordinates Babuska, Caloz, Osborn, 1994

Kozlov, 1979 Allaire Brizzi 2005; Owhadi, Zhang 2005

MsFEM [Hou, Wu: 1997]; [Efendiev, Hou, Wu: 1999]

[Fish - Wagiman, 1993] [Gloria 2010] Arbogast, 2011: Mixed MsFEM

Projection based method Nolen, Papanicolaou, Pironneau, 2008

HMM

Engquist, E, Abdulle, Runborg, Schwab, et Al. 2003-...

Flux norm Berlyand, Owhadi 2010; Symes 2012

Localization [Chu-Graham-Hou-2010] (limited inclusions)

[Efendiev-Galvis-Wu-2010] (limited inclusions or mask)

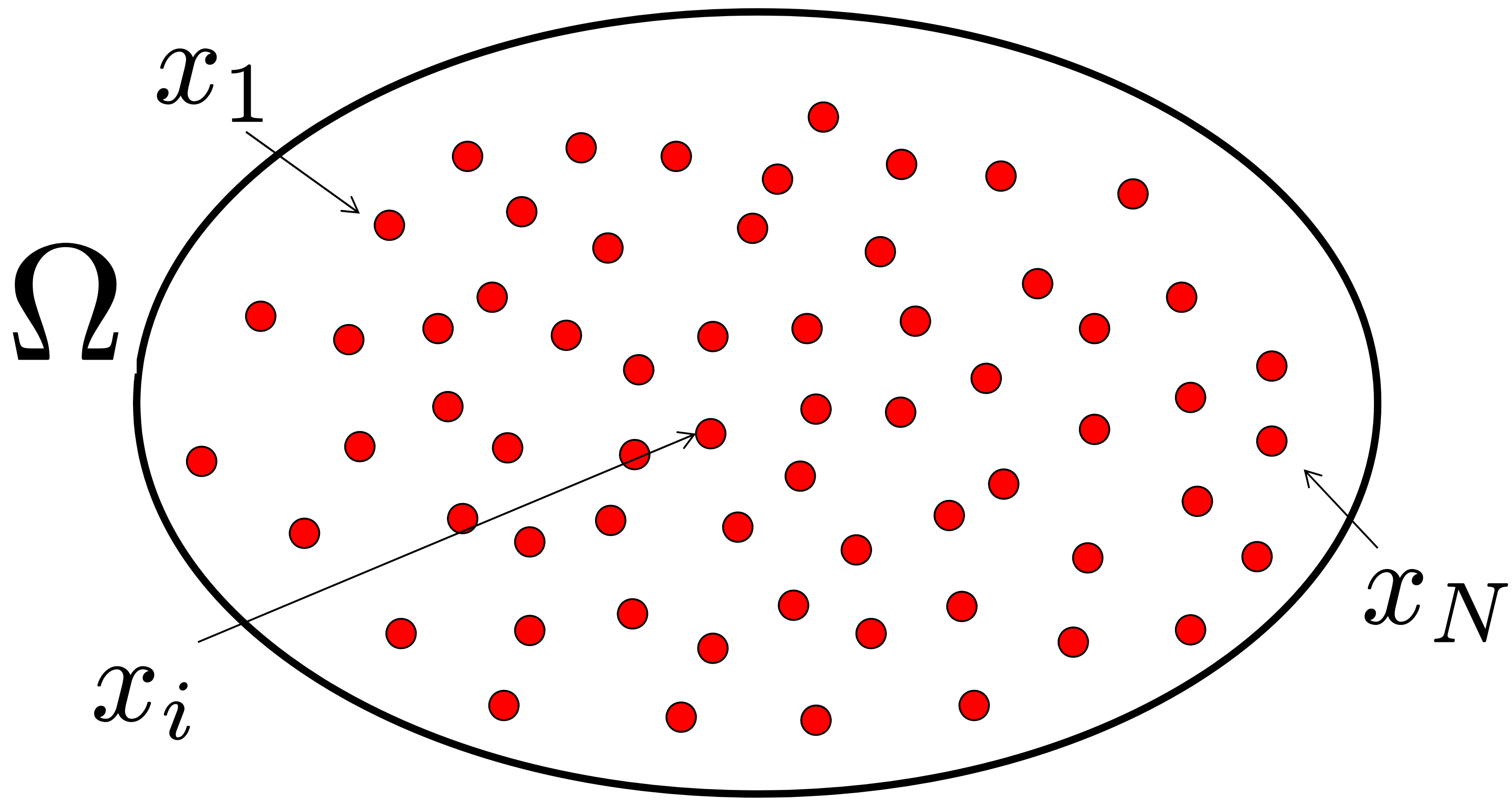
[Babuska-Lipton 2010] (local boundary eigenvectors)

[Owhadi-Zhang 2011] (localized transfer property)

[Malqvist-Peterseim 2012] Volume averaged interpolation

Alternative Approach

$$\text{Select } \{x_1, \ldots, x_N\} \subset \Omega$$



Player A

Chooses
$g \in L^2(\Omega)$

Player B

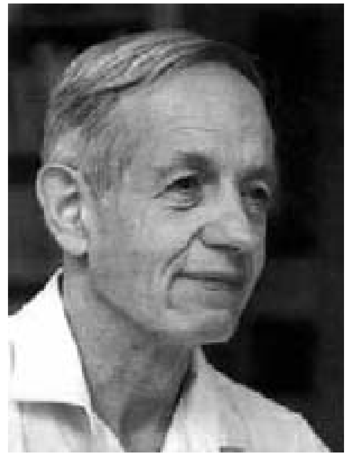
Sees
$u(x_1), \ldots, u(x_N)$

Chooses θ

$$\mathcal{E}(g,\theta) = \Big| u(x) - \theta\big(u(x_1), \ldots, u(x_N)\big) \Big|^2$$

Game theory and statistical decision theory

John Von Neumann



John Nash

Abraham Wald

The best strategy is to play at random

Obtained by finding the worst prior in the Bayesian class of estimators

Replace g by a stochastic field ξ

(2)
$$\begin{cases} -\operatorname{div}(a\nabla u) = \xi, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases}$$

$g \in L^2(\Omega)$ ⟺ ξ: white noise

$g \in H^{\pm s}(\Omega)$ ⟺ $\xi = \Delta^{\mp s/2}$white noise

Best strategy

$$\theta = \mathbb{E}\left[u(x) \middle| u(x_1), \ldots, u(x_N)\right]$$

Replace g by a stochastic field ξ

(2)
$$\begin{cases} -\operatorname{div}(a\nabla u) = \xi, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases}$$

$g \in L^2(\Omega)$ ⟷ ξ: white noise

$g \in H^{\pm s}(\Omega)$ ⟷ $\xi = \Delta^{\mp s/2}$white noise

Best strategy

$$\theta = \mathbb{E}\left[u(x) \middle| u(x_1), \ldots, u(x_N)\right]$$

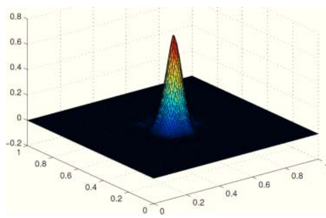
Theorem u: sol of (2)

$$\mathbb{E}\big[u(x)\big|u(x_1),\ldots,u(x_N)\big] = \sum_{i=1}^N u(x_i)\phi_i(x)$$

$a = I_d$ ⟺ ϕ_i: Polyharmonic splines

[Harder-Desmarais, 1972]

[Duchon 1976, 1977,1978]



$a_{i,j} \in L^\infty(\Omega)$ ⟺ ϕ_i: Rough Polyharmonic splines

[Owhadi-Zhang-Berlyand 2013]

Theorem u: sol of (1), $\sigma(x)$: SD of $u(x)|u(x_i)$

$$\Big|u(x) - \sum_{i=1}^N u(x_i)\phi_i(x)\Big| \leq \sigma(x)\|g\|_{L^2(\Omega)}$$

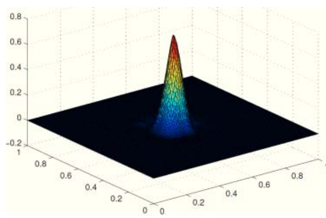
Theorem u: sol of (2)

$$\mathbb{E}\left[u(x)\middle|u(x_1),\ldots,u(x_N)\right] = \sum_{i=1}^{N} u(x_i)\phi_i(x)$$

$a = I_d$ $\Longleftrightarrow$ ϕ_i: Polyharmonic splines

[Harder-Desmarais, 1972]

[Duchon 1976, 1977,1978]



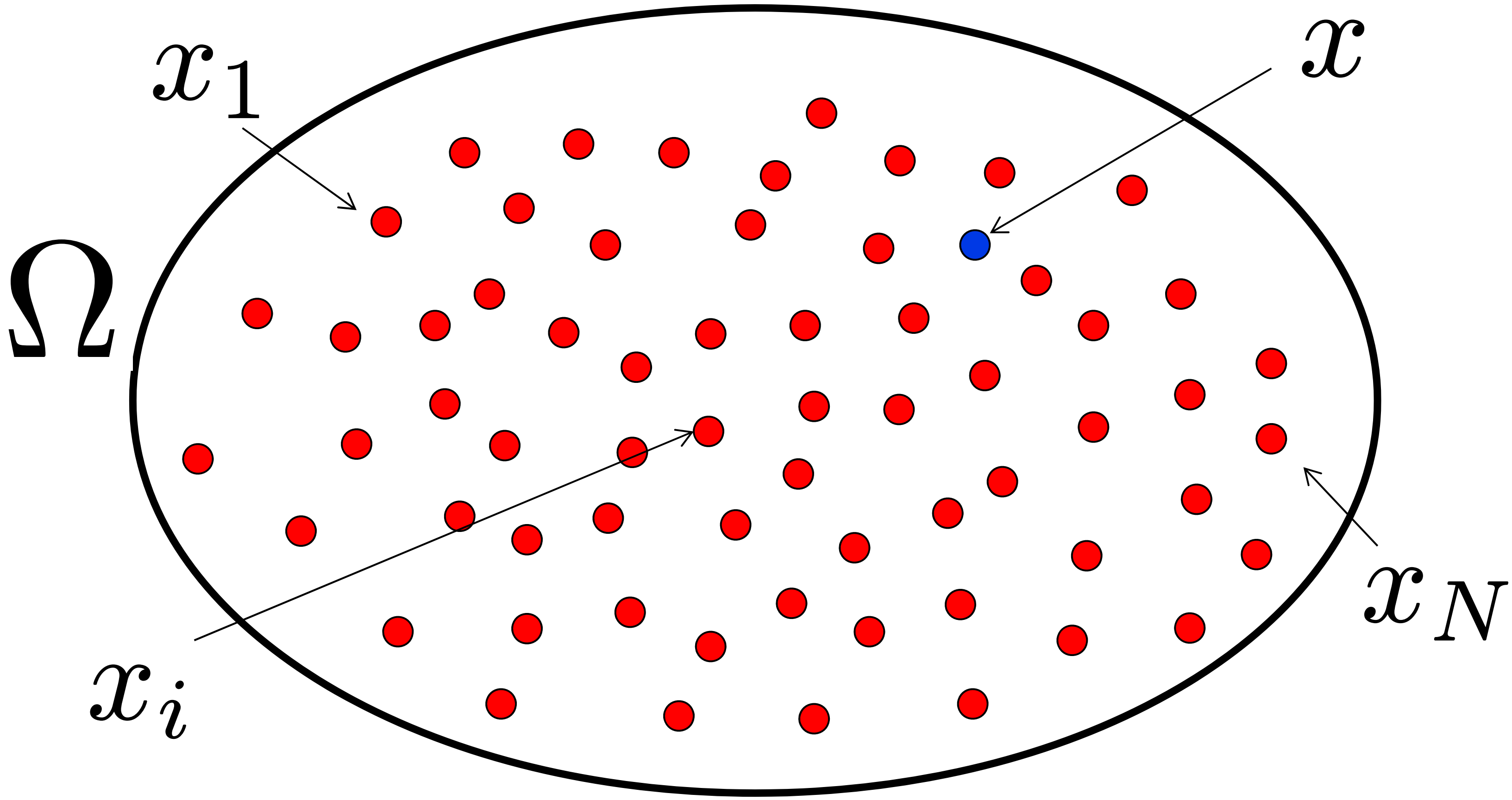
$a_{i,j} \in L^\infty(\Omega)$ $\Longleftrightarrow$ ϕ_i: Rough Polyharmonic splines

[Owhadi-Zhang-Berlyand 2013]

Theorem u: sol of (1)

$$\left\| u - \sum_{i=1}^{N} u(x_i)\phi_i \right\|_{H_0^1(\Omega)} \leq H \|g\|_{L^2(\Omega)}$$

Accuracy of RPS as an interpolation basis



The accuracy depends only on

$$H := \sup_{x \in \Omega} \min_i \|x - x_i\|$$

Theorem

$$\mathbb{E}\big[u(x)\big| \int_\Omega u(y)\chi_1(y)\,dy, \ldots, \int_\Omega u(y)\chi_N(y)\,dy\big] = \sum_{i=1}^N \phi_i(x) \int_\Omega u(y)\chi_i(y)\,dy$$

$-\operatorname{div}(a\nabla)$ ⟷ Abritrary linear integro-differential operator $\mathcal{L}$

Observations $u(x_1), \ldots, u(x_N)$ ⟷ Abritrary linear observations $\int_\Omega u(y)\chi_i(y)\,dy$

$g \in L^2(\Omega)$ ⟷ $g \in H^{\pm s}(\Omega)$

ξ: white noise ⟷ $\xi = \Delta^{\mp s/2}$white noise

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

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

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

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

P. L. Chebyshev
1821-1894

A. A. Markov
1856-1922

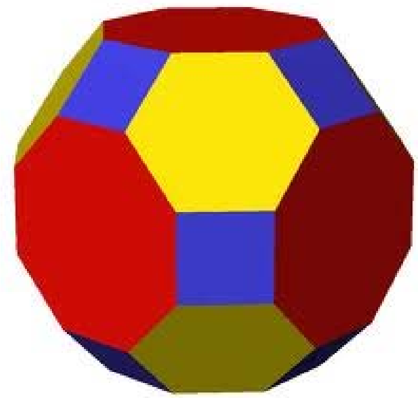
M. G. Krein
1907-1989

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

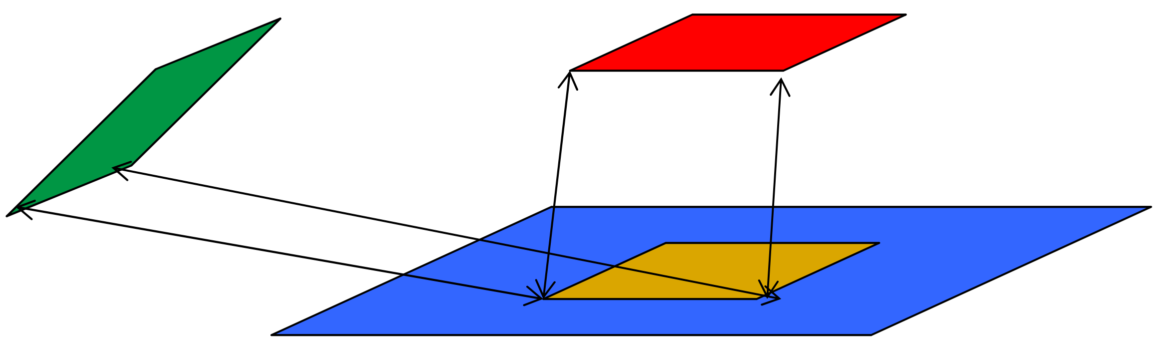
$$\mu \longrightarrow \left(\mathbb{E}_{X\sim\mu}[X], \mathbb{E}_{X\sim\mu}[X^2], \ldots, \mathbb{E}_{X\sim\mu}[X^k]\right)$$

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

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.

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

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

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

Origin of these new Selberg integral formulas and new RKHS

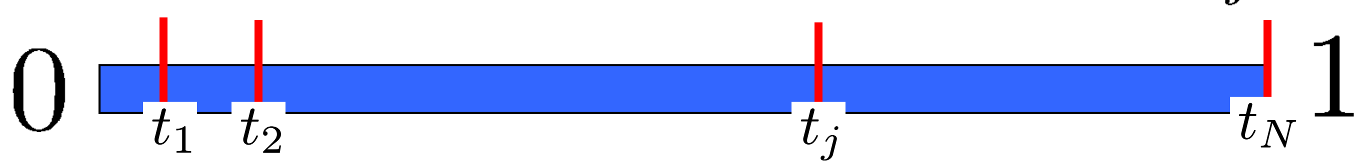
Compute $\mathrm{Vol}(M_k)$ using different (finite-dimensional) representations in $\mathcal{M}\big([0,1]\big)$

$$0 \le t_1 < t_2 < \cdots < t_N \le 1$$

$$\lambda_1, \ldots, \lambda_N > 0,\ \textstyle\sum_{j=1}^N \lambda_j = 1$$

$$\mu = \textstyle\sum_{j=1}^N \lambda_j \delta_{t_j} \xrightarrow{\quad\Psi\quad} (q_1, \ldots, q_k)$$

$$q_i = \textstyle\sum_{j=1}^N \lambda_j t_j^i$$



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

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

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

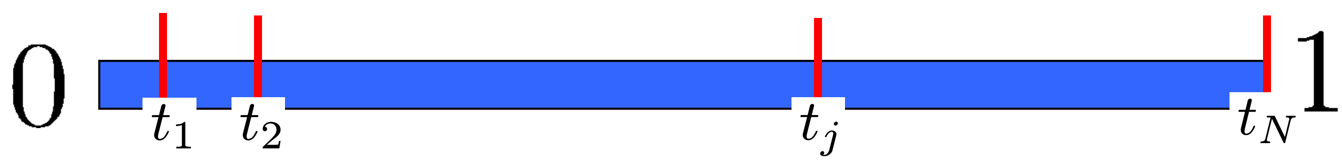
μ is called

- 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

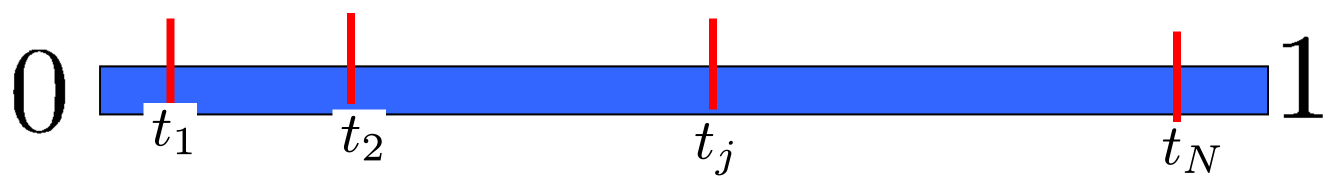
Theorem

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

Upper



Lower



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

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

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

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

Selberg Identities

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

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

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

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

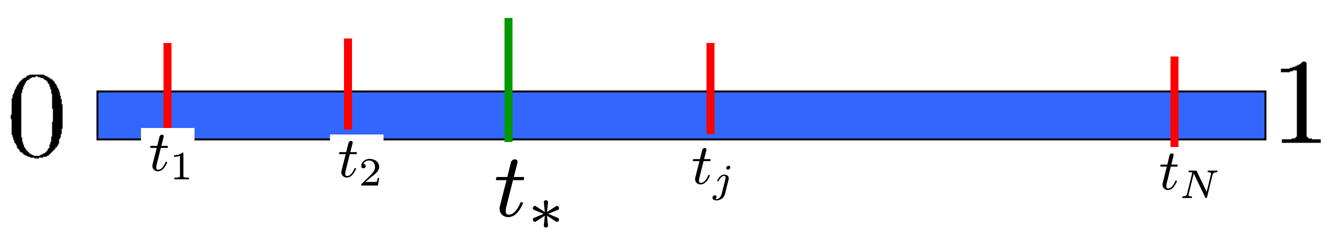
Index $i(\mu)$: Number of support points of μ

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

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

Thank you