

Understanding Multivariate Computation using the Kolmogorov Superposition Theorem

Matthew Knepley and Jonas Actor

Computer Science and Engineering
University at Buffalo

MCS Seminar, Argonne National Laboratory
Chicago, IL September 19, 2018

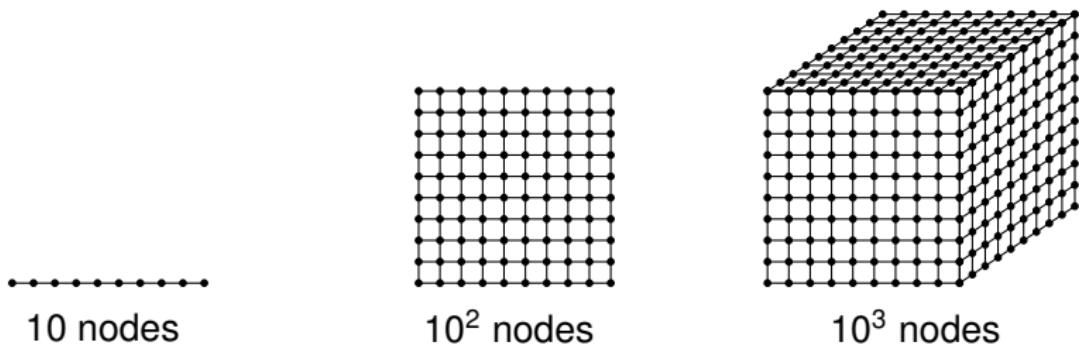


KST Representation Collaboration



Jonas Actor

Curse of Dimensionality



Cost of simulation / computation grows **exponentially**

Do functions of three variables exist at all?¹

¹Pólya and Szegő, Problems and Theorems of Analysis, 1925 (German), transl. 1945, reprinted 1978

Q: Can any function of three variables be expressed using functions of only two variables?²

²Hilbert, Göttinger Nachrichten, 1900

³Arnol'd, Dokl. Akad. Nauk SSSR 114:5, 1957

Q: Can any function of three variables be expressed using functions of only two variables?²

Ex: Cardano's Formula for roots of a cubic equation

²Hilbert, Göttinger Nachrichten, 1900

³Arnol'd, Dokl. Akad. Nauk SSSR 114:5, 1957

Q: Can any function of three variables be expressed using functions of only two variables?²

Ex: Cardano's Formula for roots of a cubic equation

A: Any *continuous* function of three variables can be expressed using *continuous* functions of only two variables.³

²Hilbert, Göttinger Nachrichten, 1900

³Arnol'd, Dokl. Akad. Nauk SSSR 114:5, 1957

Q: Can any function of three variables be expressed using only univariate functions and addition?

⁴Kolmogorov, Dokl. Akad. Nauk SSSR 114:5, 1957

Q: Can any function of three variables be expressed using only univariate functions and addition?

A: Yes⁴, for any continuous $f : [0, 1]^n \rightarrow \mathbb{R}$

$$f(x_1, \dots, x_n) = \sum_{q=0}^{2n} \chi \left(\sum_{p=1}^n \psi_{p,q}(x_p) \right).$$

⁴Kolmogorov, Dokl. Akad. Nauk SSSR 114:5, 1957

What's going on here?

$$f(x_1, \dots, x_n) = \sum_{q=0}^{2n} \chi_q \left(\sum_{p=1}^n \psi_{p,q}(x_p) \right)$$

What's going on here?

$$f(x_1, \dots, x_n) = \sum_{q=0}^{2n} \chi_q \left(\sum_{p=1}^n \psi_{p,q}(x_p) \right)$$

Original function $f : [0, 1]^n \rightarrow \mathbb{R}$

What's going on here?

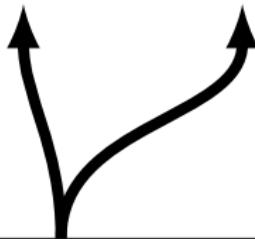
$$f(x_1, \dots, x_n) = \sum_{q=0}^{2n} \chi_q \left(\sum_{p=1}^n \psi_{p,q}(x_p) \right)$$



Inner function $\psi_{p,q} : [0, 1] \rightarrow \mathbb{R}$

What's going on here?

$$f(x_1, \dots, x_n) = \sum_{q=0}^{2n} \chi_q \left(\sum_{p=1}^n \psi_{p,q}(x_p) \right)$$



Addition

What's going on here?

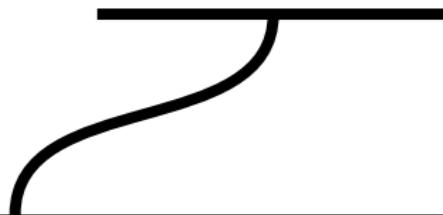
$$f(x_1, \dots, x_n) = \sum_{q=0}^{2n} \chi_q \left(\sum_{p=1}^n \psi_{p,q}(x_p) \right)$$



Outer function $\chi : \mathbb{R} \rightarrow \mathbb{R}$

What's going on here?

$$f(x_1, \dots, x_n) = \sum_{q=0}^{2n} \chi_q \left(\sum_{p=1}^n \psi_{p,q}(x_p) \right)$$



Function composition

2D: $n = 2$

$$\begin{aligned}f(x, y) = & \chi_0(\psi_{x,0}(x) + \psi_{y,0}(y)) \\& + \chi_1(\psi_{x,1}(x) + \psi_{y,1}(y)) \\& + \chi_2(\psi_{x,2}(x) + \psi_{y,2}(y)) \\& + \chi_3(\psi_{x,3}(x) + \psi_{y,3}(y)) \\& + \chi_4(\psi_{x,4}(x) + \psi_{y,4}(y))\end{aligned}$$

$$\begin{aligned}f(x, y) = & \chi_0(\lambda_x \psi(x) + \lambda_y \psi(y)) \\& + \chi_1(\lambda_x \psi(x + \epsilon) + \lambda_y \psi(y + \epsilon)) \\& + \chi_2(\lambda_x \psi(x + 2\epsilon) + \lambda_y \psi(y + 2\epsilon)) \\& + \chi_3(\lambda_x \psi(x + 3\epsilon) + \lambda_y \psi(y + 3\epsilon)) \\& + \chi_4(\lambda_x \psi(x + 4\epsilon) + \lambda_y \psi(y + 4\epsilon))\end{aligned}$$

Single ψ function (Sprecher)

$$\begin{aligned}f(x, y) = & \chi(\lambda_x \psi(x) + \lambda_y \psi(y)) \\& + \chi(\lambda_x \psi(x + \epsilon) + \lambda_y \psi(y + \epsilon) + \delta) \\& + \chi(\lambda_x \psi(x + 2\epsilon) + \lambda_y \psi(y + 2\epsilon) + 2\delta) \\& + \chi(\lambda_x \psi(x + 3\epsilon) + \lambda_y \psi(y + 3\epsilon) + 3\delta) \\& + \chi(\lambda_x \psi(x + 4\epsilon) + \lambda_y \psi(y + 4\epsilon) + 4\delta)\end{aligned}$$

Single χ function (Lorentz) and ψ function (Sprecher)

$$\Psi^q(x_1, \dots, x_n) = \sum_{p=1}^n \psi_{p,q}(x_p)$$

is independent of f , so that

$$KST : f \rightarrow \chi$$

Outline

1 Constructive KST

2 Abstract KST

3 Concrete KST

4 Conclusions

Topographic KST

$f(x, y) = \text{elevation in Chugash Mountains at } (x, y)$



<https://viewer.nationalmap.gov/basic/>

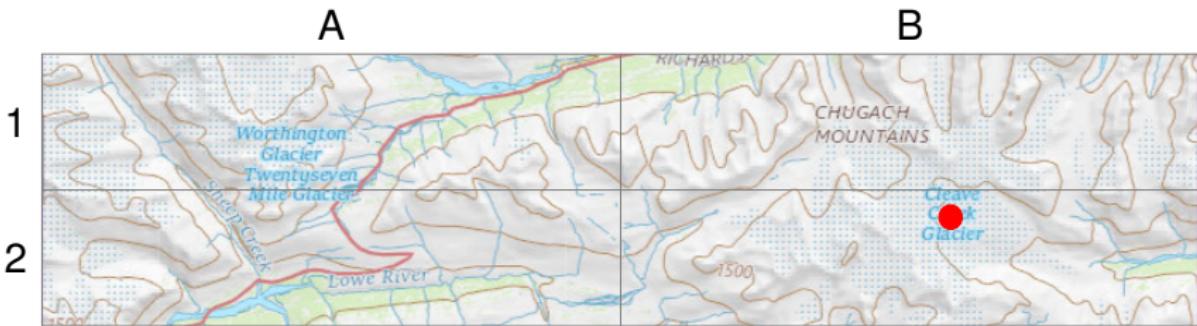
Topographic KST

$f(\text{Cleave Creek Glacier}) = ?$



<https://viewer.nationalmap.gov/basic/>

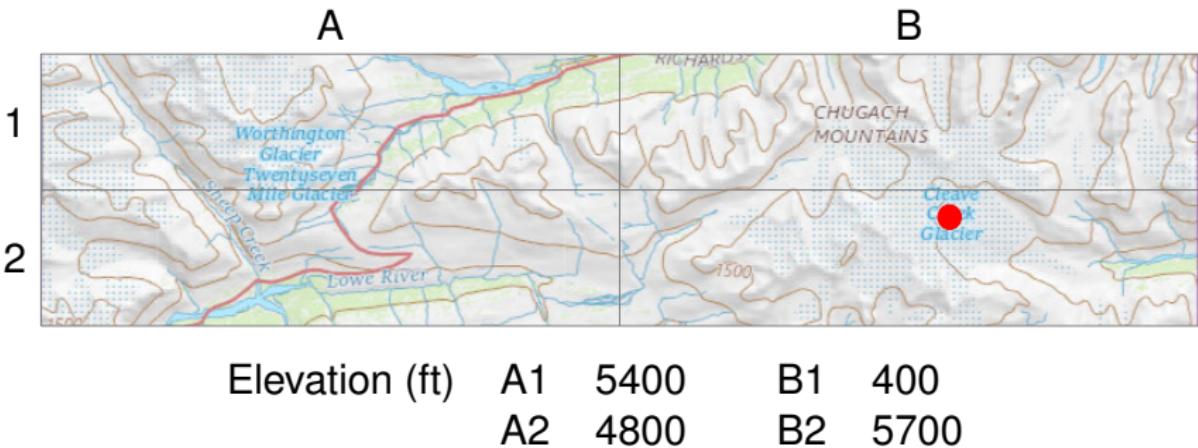
Topographic KST



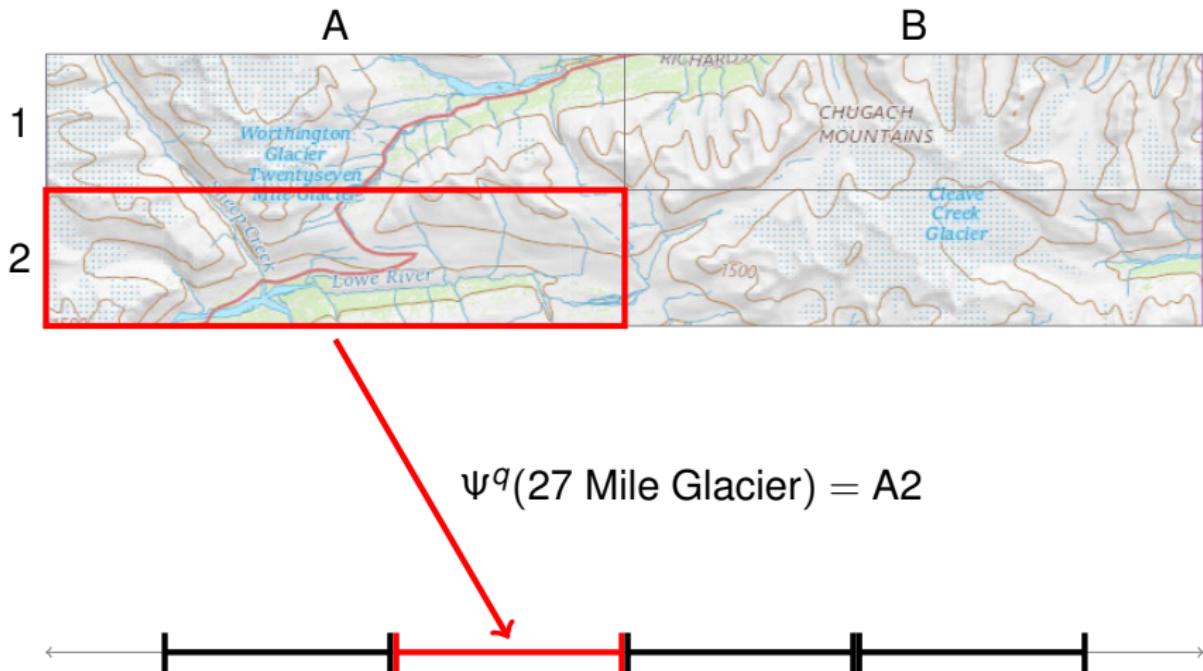
Tsina Glacier A1
27 Mile Glacier A2

Tsina River B1
Chugash Mtns. B2

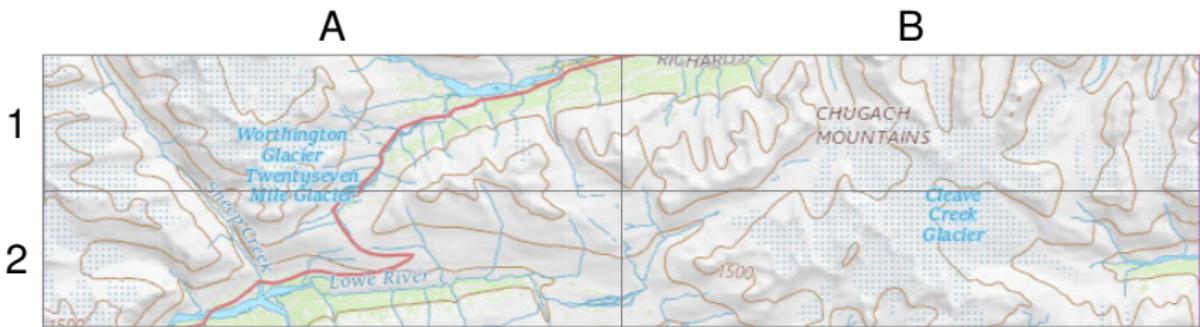
Topographic KST



Topographic KST



Topographic KST



$$f(27 \text{ Mile Glacier}) = \chi(A2)$$



Topographic KST

A

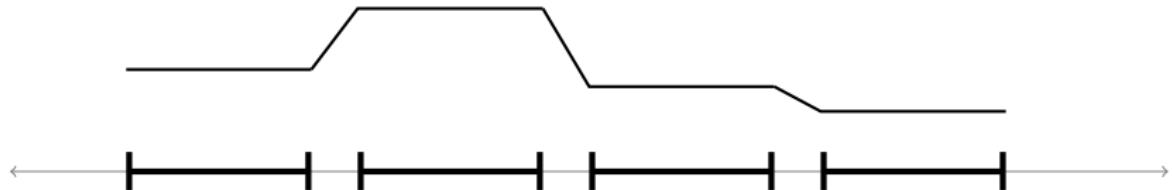
1



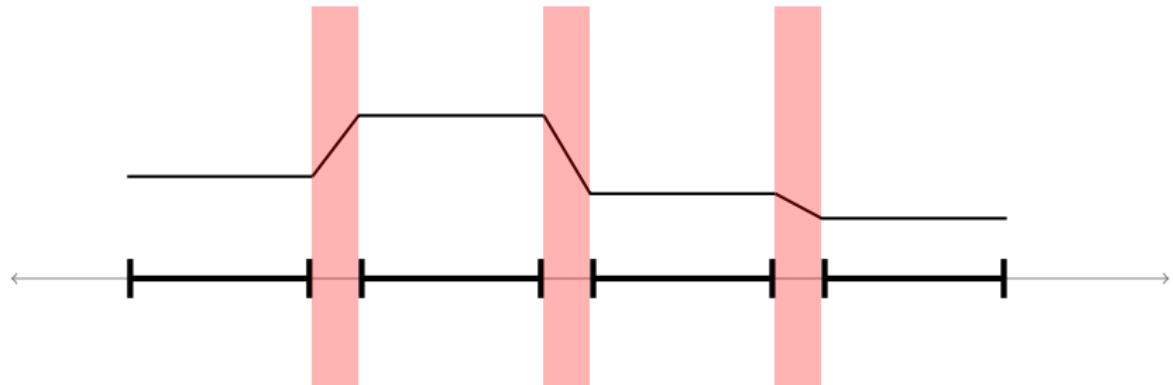
2

**B**

Topographic KST



Topographic KST



Outline

1 Constructive KST

2 Abstract KST

- Spatial Decomposition
- Inner Functions

3 Concrete KST

4 Conclusions

Kolmogorov Strategy

- Approximate with *constants*
- Insert *gaps* to preserve continuity
- *Duplicate and shift* to cover domain

Kolmogorov Strategy

- Approximate with *constants*
 - Insert *gaps* to preserve continuity
 - *Duplicate and shift* to cover domain
- Ψ : balances continuity against
discriminating between different points

Kolmogorov Strategy

- Approximate with *constants*
- Insert *gaps* to preserve continuity
- *Duplicate and shift* to cover domain

Ψ : regularity vs. point separation

χ assigns values to shifted sums of our R^{2n+1} embedding to match f

Kolmogorov Strategy

- Approximate with *constants*
- Insert *gaps* to preserve continuity
- *Duplicate and shift* to cover domain

Ψ : regularity vs. point separation

χ : approximates f

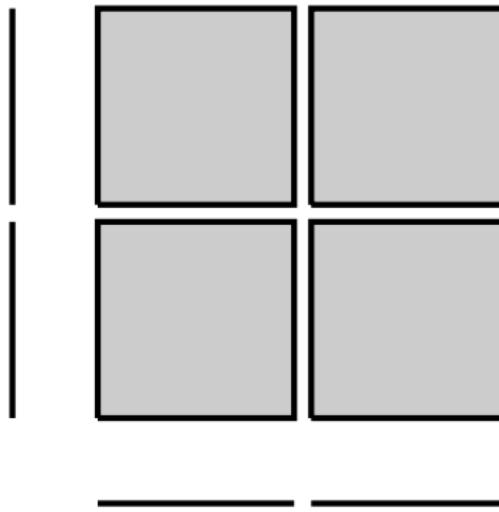
Outline

- 2 Abstract KST
 - Spatial Decomposition
 - Inner Functions

Spatial Decomposition

Single town

Cartesian product \mathcal{S} of a set of intervals \mathcal{I}
that nearly cover $[0, 1]$

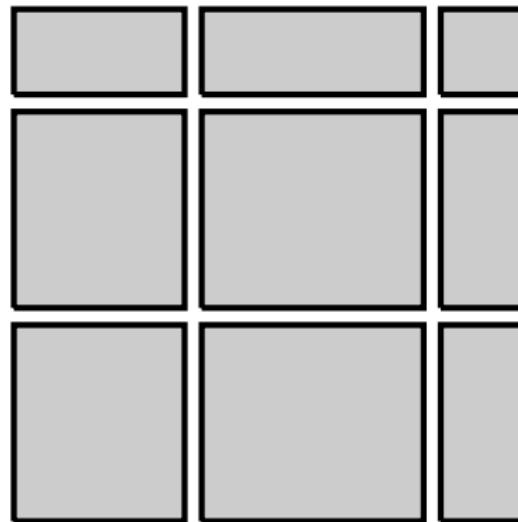


Spatial Decomposition

Refined towns

\mathcal{S}^k : Near partition of $[0, 1]^d$

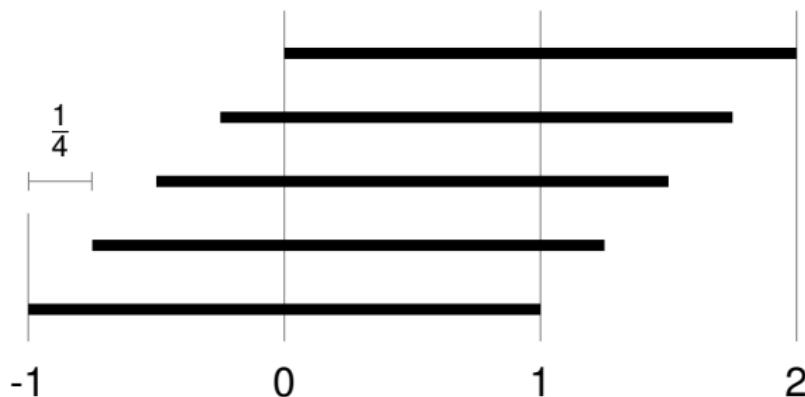
$\text{diam}(\mathcal{S}^k) \rightarrow 0$ as $k \rightarrow \infty$



Spatial Decomposition

Cover gaps

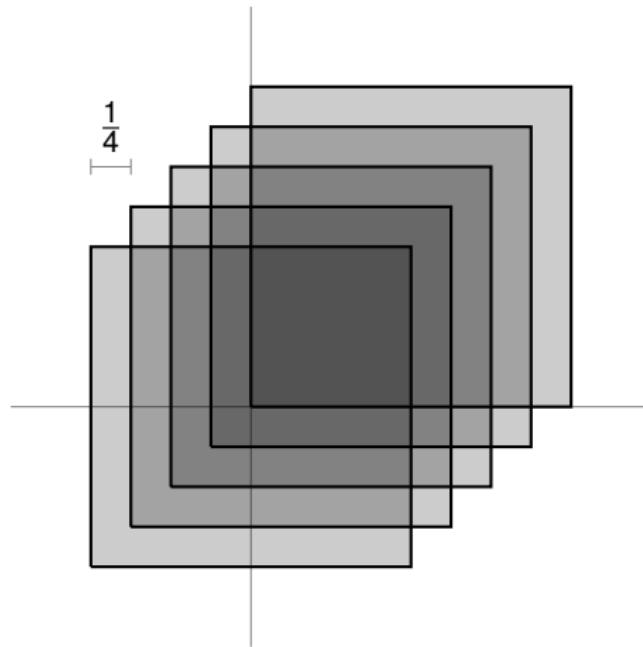
$2n + 1$ copies \mathcal{I}_q of the same set of intervals \mathcal{I}
shifted by ϵ



Spatial Decomposition

Overlapping towns

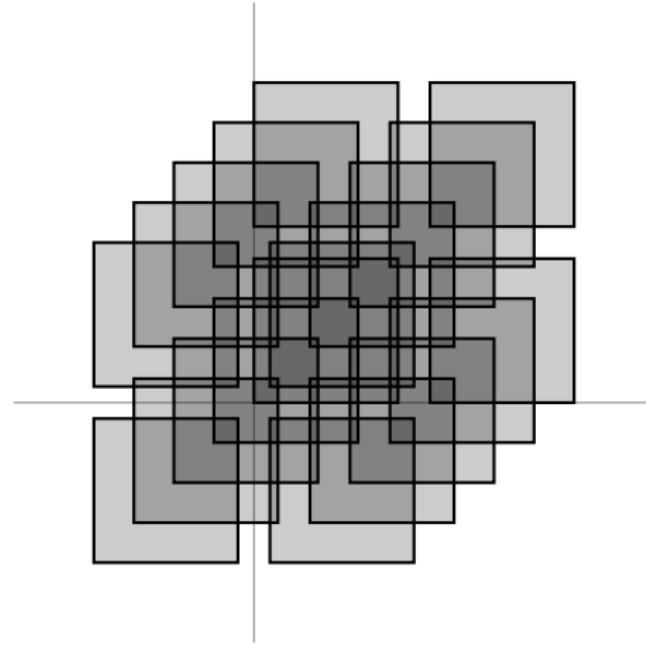
Cartesian product \mathcal{S}_q of shifted intervals \mathcal{I}_q



Spatial Decomposition

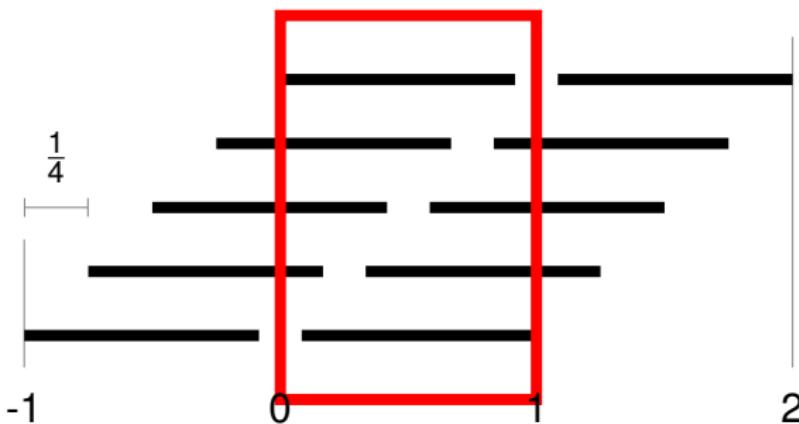
Refined overlapping towns

$$\text{diam}(S_q^k) \rightarrow 0 \text{ as } k \rightarrow \infty$$



Gap Requirement

Every $x \in [0, 1]$ in **All But One** of \mathcal{I}_q



All But One of the intervals implies
More than Half of the squares

Outline

2

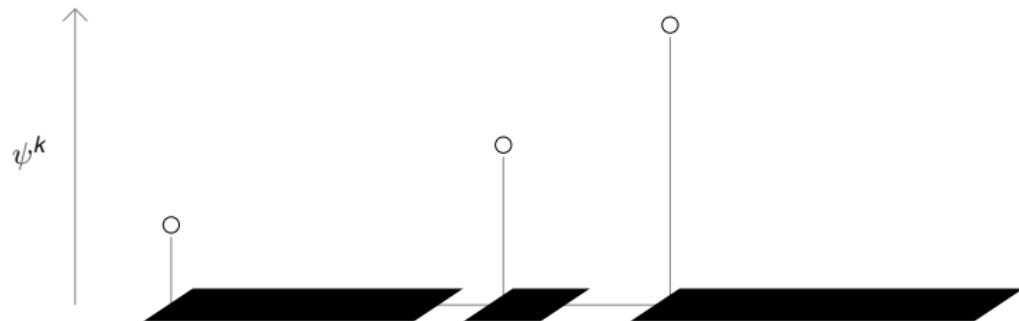
Abstract KST

- Spatial Decomposition
- Inner Functions

Inner Function ψ

At each refinement level k :

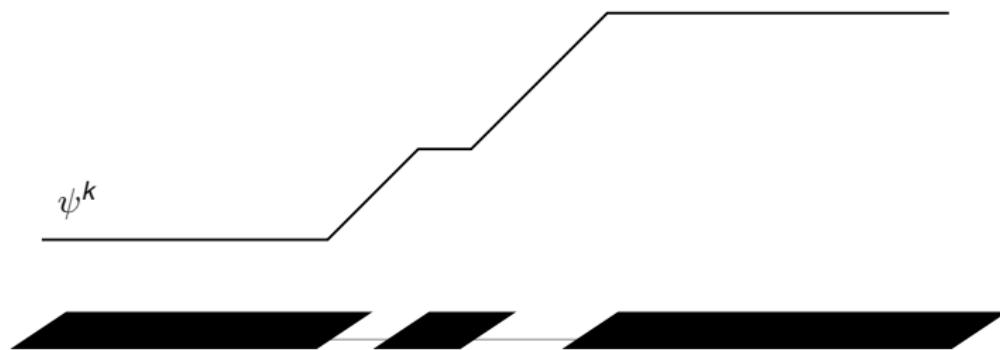
- Assign a value of ψ^k at the lower left corner of each square
- Value is fixed for all future k



Inner Function ψ

ψ^k (near) constant on squares, linear on gaps

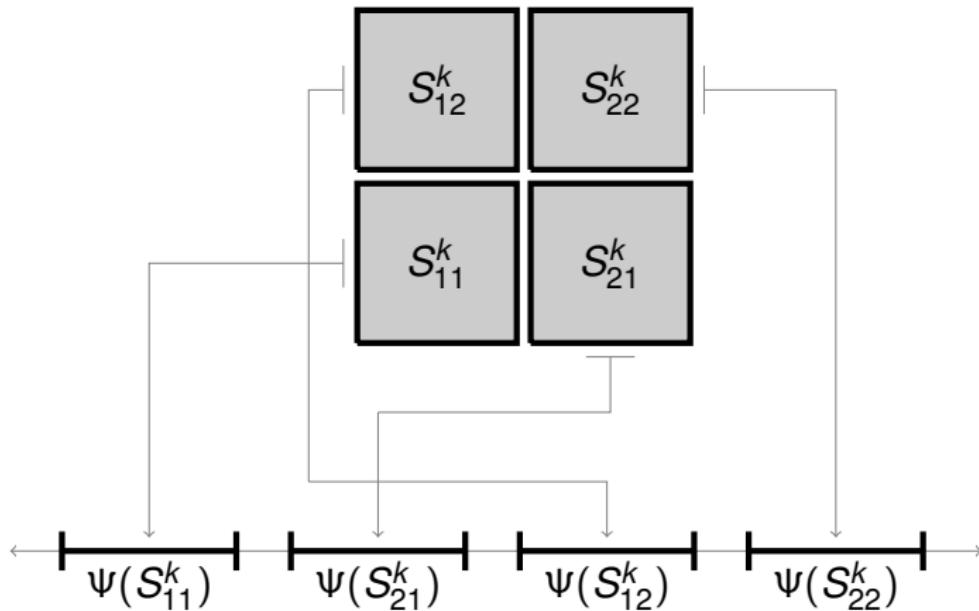
$$\psi = \lim_{k \rightarrow \infty} \psi^k \text{ uniformly}$$



Small gaps \rightarrow steep ψ^k , but
Large gaps can violate **All But One**

Disjoint Image Condition

$$\forall S, S' \in \mathcal{S}^k, \Psi(S) \cap \Psi(S') = \emptyset$$



Kolmogorov Requirements for the Inner Function

Refinement:

$$\text{diam}(\mathcal{S}_q^k) \rightarrow 0 \text{ as } k \rightarrow \infty$$

More Than Half:

$\forall x \in [0, 1]^n, x \in \mathcal{S}_q^k$ for at least $n + 1$ of the q 's

Disjoint Image:

$$\forall S, S' \in \mathcal{S}_q^k, \Psi(S) \cap \Psi(S') = \emptyset$$

Monotonicity:

Function ψ is strictly monotonic increasing

Regularity of Inner Function

KST trades smoothness for variables

- KST not feasible for $\psi_{p,q} \in C^1([0, 1])$ ⁵
- Possible to construct $\psi_{p,q} \in \text{Lip}([0, 1])$ ⁶
- Only known construction $\psi_{p,q} \in \text{Hölder}([0, 1])$

⁵Vituskin, DAN, 95:701–704, 1954.

⁶Fridman, DAN, 177:1019–1022, 1967.

Outline

1 Constructive KST

2 Abstract KST

3 Concrete KST

- Sprecher Construction
- Lipschitz Construction
- Outer Functions

4 Conclusions

Outline

3

Concrete KST

- Sprecher Construction
- Lipschitz Construction
- Outer Functions

Sprecher Construction $\widehat{\psi}$ ⁷

- Inner function $\widehat{\psi}$ is the uniform limit $\lim_{k \rightarrow \infty} \widehat{\psi}^k$
- Fix values at numbers with k digits in base- γ expansion
 - Almost flat for most points
 - Large increase for expansions ending in $\gamma - 1$
- Linearly interpolate between fixed values

⁷Sprecher, J. Constr. Approx. 1995

Sprecher Construction $\widehat{\psi}$ ⁸

Radix $\gamma \geq 2d + 1$

$$\begin{aligned}\mathcal{D}^k &= \left\{ \frac{i}{\gamma^k} : i = 0, \dots, \gamma^k \right\} \\ &= \left\{ 0.i_0i_1\dots i_k : i_\ell \in [0, \gamma - 1], \ell \in [0, k] \right\}\end{aligned}$$

$$\beta(k) = \frac{n^k - 1}{n - 1}$$

⁸Sprecher, J. Constr. Approx. 1995

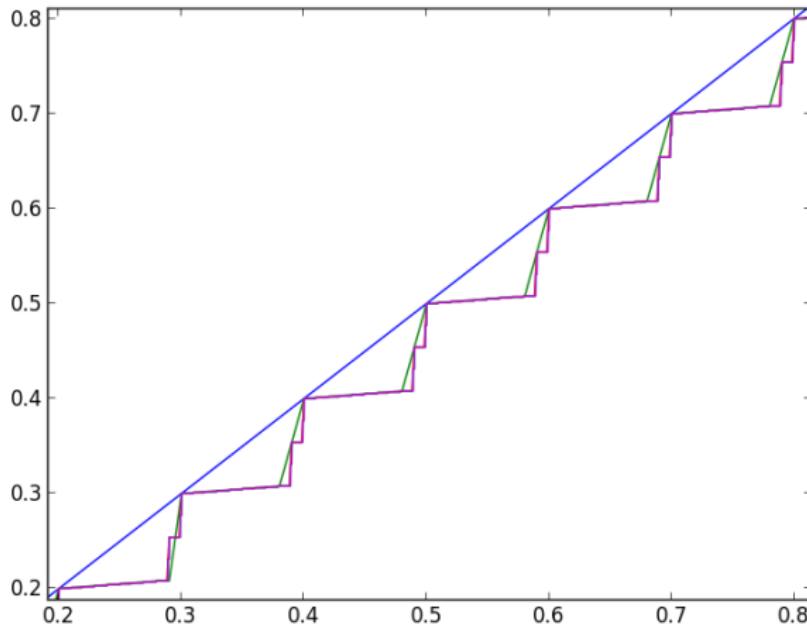
Köppen Construction $\widehat{\psi}^9$

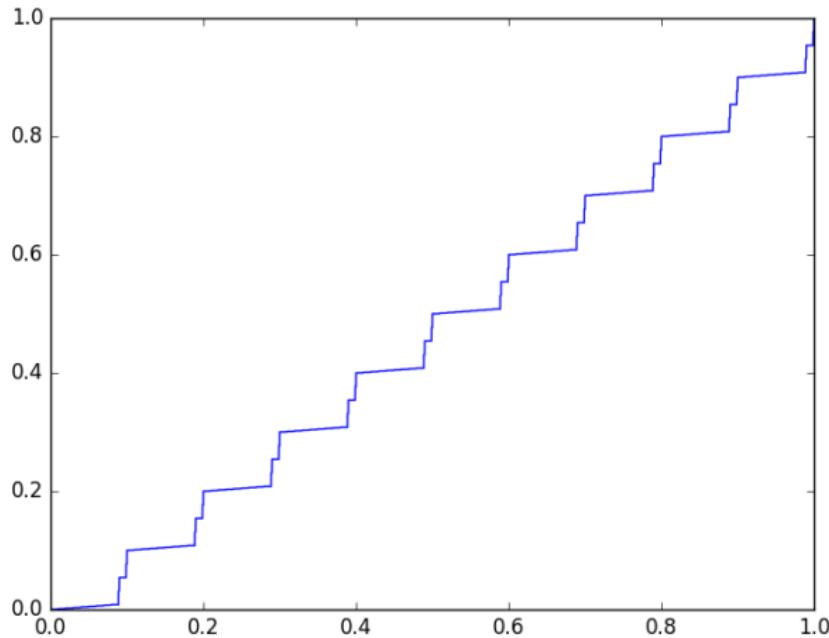
$$\widehat{\psi}^k(d_k) = \begin{cases} d_k & k = 1 \\ \widehat{\psi}^{k-1} \left(d_k - \frac{i_k}{\gamma^k} \right) + \frac{i_k}{\gamma^{\beta(k)}} & k > 1, i_k < \gamma - 1 \\ \frac{1}{2} \left(\widehat{\psi}^k \left(d_k - \frac{1}{\gamma^k} \right) + \widehat{\psi}^{k-1} \left(d_k + \frac{1}{\gamma^k} \right) \right) & k > 1, i_k = \gamma - 1 \end{cases}$$

Interpolate linearly to extend $\widehat{\psi}^k$ from \mathcal{D}^k to $[0, 1]$

$$\widehat{\psi} = \lim_{k \rightarrow \infty} \widehat{\psi}^k$$

⁹Köppen, ICANN 2002, LNCS 2415, 2002

Köppen Construction $\hat{\psi}^{10}$ ¹⁰Köppen, ICANN 2002, LNCS 2415, 2002

Köppen Construction $\widehat{\psi}^{11}, k = 7$ ¹¹Köppen, ICANN 2002, LNCS 2415, 2002

Köppen Construction $\widehat{\psi}$

Approximability

While $\widehat{\psi}$ satisfies all Kolomogorov requirements¹², it's not locally Lipschitz on **any** open interval $I \subset [0, 1]$.

In fact, $\psi \in \text{Hölder}_\alpha([0, 1])$ for $\alpha = \log_{2n+2} 2$,

$$|\psi(x) - \psi(y)| < \epsilon \quad \text{if} \quad |x - y| < \epsilon^{1/\alpha}.$$

For $n = 2$, $\alpha^{-1} \approx 2.5$, but

for $n = 10$ we have $\alpha^{-1} \approx 4.5$.

¹²Braun and Griebel, Constr. Approx., 30:653-675, 2009

Outline

3

Concrete KST

- Sprecher Construction
- Lipschitz Construction
- Outer Functions

Reasoning for Lipschitz Construction

KST inner functions are

- ... strictly monotonically increasing
- ... so they are of Bounded Variation
- ... so they define rectifiable curves
- ... which have Lipschitz reparameterizations.

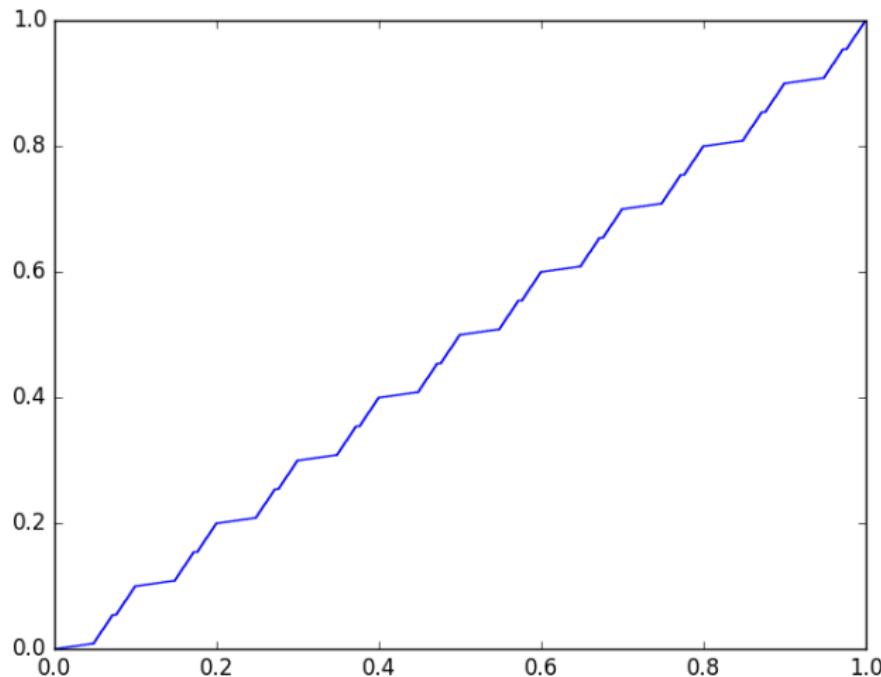
Lipschitz Reparameterization

Reparameterization $\sigma : [0, 1] \rightarrow [0, 1]$

$$\sigma(x) = \frac{\text{arclength of } \widehat{\psi} \text{ from 0 to } x}{\text{total arclength of } \widehat{\psi} \text{ from 0 to 1}}$$

$$\psi(x) = \widehat{\psi}(\sigma^{-1}(x))$$

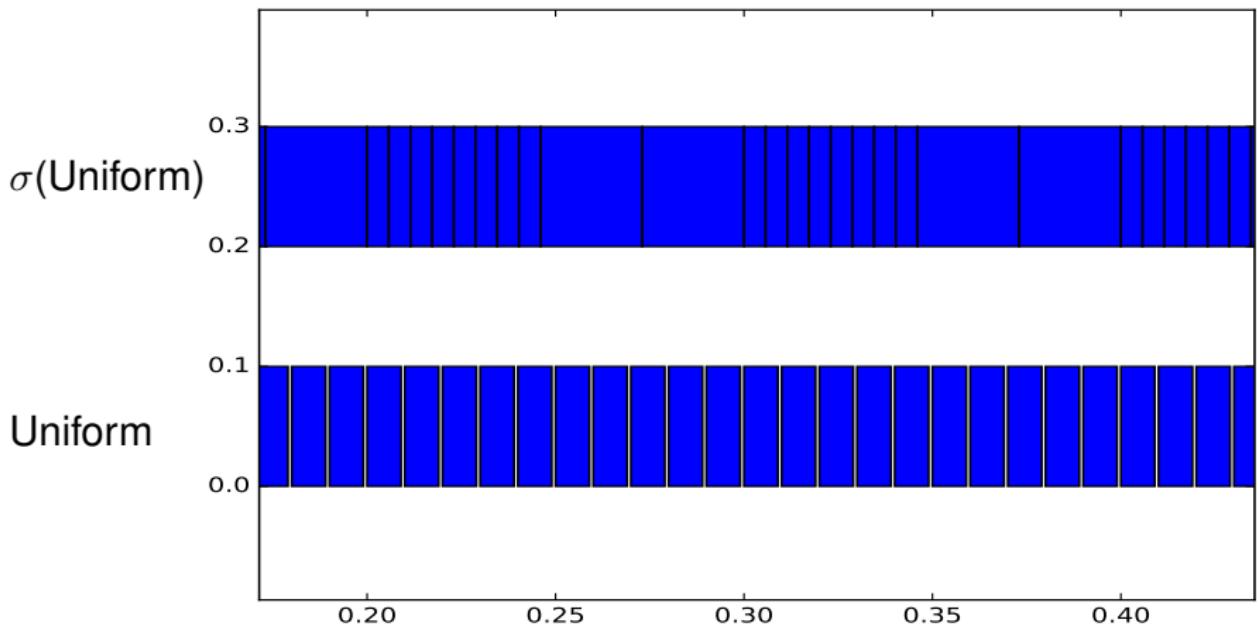
Lipschitz Reparameterization $\psi, k = 7$



KST Proof with Lipschitz ψ

- $\psi^k = \widehat{\psi^k} \circ (\sigma^k)^{-1}$ converges uniformly to ψ
- ψ satisfies the Kolmogorov requirements
- $\psi \in \text{Lip}_2([0, 1])$

Comparing $\widehat{\mathcal{S}}^k$ and \mathcal{S}^k



Outline

3

Concrete KST

- Sprecher Construction
- Lipschitz Construction
- Outer Functions

Kolmogorov Outer Function

Define a residual

$$f_r = f_{r-1} - \sum_{q=0}^{2d} \chi_{r-1} \circ \Psi^q$$

and choose k_r , with $r \in \mathcal{N}$, so that the oscillation of

$$f_{r-1} - \sum_{q=0}^{2d} \chi_{r-1} \circ \Psi^q$$

is bounded by $\frac{\|f_{r-1}\|_\infty}{d+1}$. Then by a Fixed Point Theorem,

$$\chi = \lim_{r \rightarrow \infty} \chi_r.$$

Outer Function Construction

$\chi_r : \mathbb{R} \rightarrow \mathbb{R}$ interpolates

$$\left\{ \left(\Psi^q(\mathbf{d}_{k_r}), \frac{1}{d+1} f(\mathbf{d}_{k_r}) \right) : \mathbf{d}_{k_r} \in \prod_{p=1}^d \sigma(\mathcal{D}_{k_r} + q\varepsilon) \right\}$$

Outline

- 1 Constructive KST
- 2 Abstract KST
- 3 Concrete KST
- 4 Conclusions

Conclusions

- KST leverages superpositions for dimension reduction
- Practical KST computation requires Lipschitz inner functions
- Need to develop adaptive theory for outer function construction

Thank You!

<http://cse.buffalo.edu/~knepley>