Data and Models

Wolfgang Dahmen

Institut für Geometrie und Praktische Mathematik
RWTH Aachen

October 26, 2015

Outline

Outline

1. Introduction
2. 3 Scenarios
   - Electron Microscopy
   - EELS Electron Energy Loss Spectroscopy
   - Climate Modeling

Outline

1. Introduction
2. 3 Scenarios
   - Electron Microscopy
   - EELS Electron Energy Loss Spectroscopy
   - Climate Modeling
3. Theoretical Performance Bounds
   - Nonparametric Estimation - A Sample Result

Outline

What is it all about...

Objective:

...extract quantifiable information from corrupted, noisy data sets when

- off-the-shelf methods are not available

What is it all about...

Objective:

...extract quantifiable information from corrupted, noisy data sets when

- off-the-shelf methods are not available
- mathematical short-cuts are essential...

What is it all about...

Objective:

...extract quantifiable information from corrupted, noisy data sets when

- off-the-shelf methods are not available
- mathematical short-cuts are essential...

Conceptual ingredients:

- adaptivity
- nonlinear recovery
- embedding into a proper continuous framework to exploit natural "problem metrics"

General “Philosophy”...

- devise mathematical model...possibly simplified...

General "Philosophy"...

- devise mathematical model...possibly simplified...
- what can be achieved at best?...theoretical "benchmark"

General "Philosophy"...

- devise mathematical model...possibly simplified...
- what can be achieved at best?...theoretical "benchmark"
- Try to develop algorithms that realize the benchmark

General "Philosophy"...

- devise mathematical model...possibly simplified...
- what can be achieved at best?...theoretical "benchmark"
- Try to develop algorithms that realize the benchmark

...has led to "unexpected" algorithmic structures...

Outline

Precision in STEM Imaging scanning transmission electron microscopy

Collaborators: B. Berkels, N. Mevenkamp (AICES), Ernst-Ruska-Center Jülich, P. Voyles (University of Wisconsin), P. Binev, T. Vogt (Unisversity of South Carolina)

Precision in STEM Imaging scanning transmission electron microscopy

Collaborators: B. Berkels, N. Mevenkamp (AICES),
Ernst-Ruska-Center Jülich, P. Voyles (University of Wisconsin), P. Binev, T. Vogt (Unisversity of South Carolina)

Images/Models: distributions of intensities - stochastic processes

Precision in STEM Imaging scanning transmission electron microscopy

Collaborators: B. Berkels, N. Mevenkamp (AICES), Ernst-Ruska-Center Jülich, P. Voyles (University of Wisconsin), P. Binev, T. Vogt (Unisversity of South Carolina)

Images/Models: distributions of intensities - stochastic processes

Objectives: atom column positions with high precision (pm scale)
...in particular at material interfaces...

Precision in STEM Imaging scanning transmission electron microscopy

Collaborators: B. Berkels, N. Mevenkamp (AICES), Ernst-Ruska-Center Jülich, P. Voyles (University of Wisconsin), P. Binev, T. Vogt (Unisversity of South Carolina)

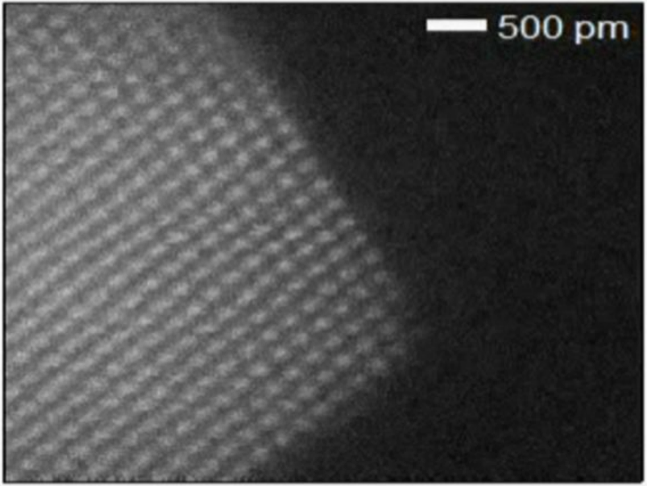
Images/Models: distributions of intensities - stochastic processes

Objectives: atom column positions with high precision (pm scale) ...in particular at material interfaces...

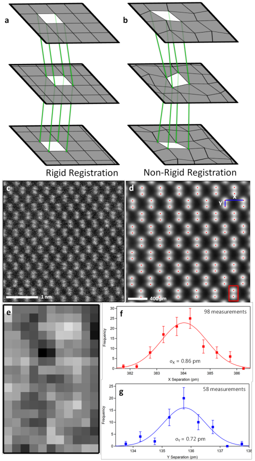
Obstructions:

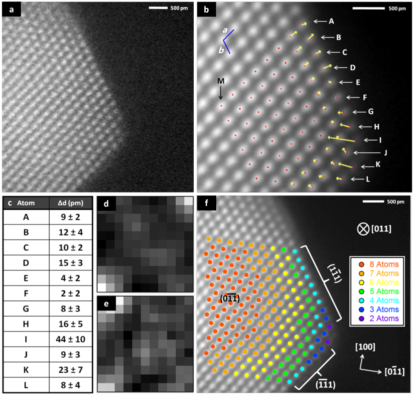
- high electron dose destroys specimen
- low electron dose causes low signal-to-noise ratio
- Poisson noise
- multiscale locally distorted movement of specimen

Pt nanocatalyst



Results - sub-picometer precision

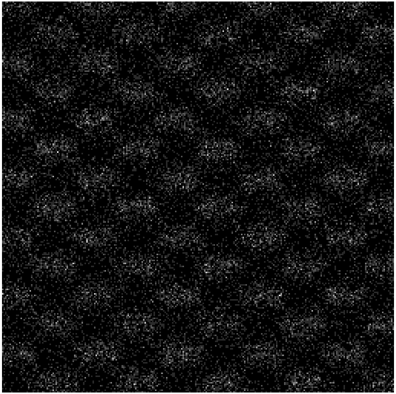


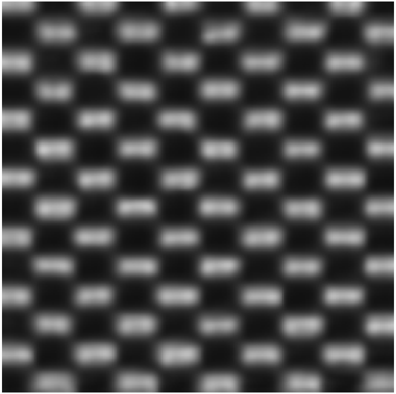


c Atom	Δd (pm)
A	9 ± 2
B	12 ± 4
C	10 ± 2
D	15 ± 3
E	4 ± 2
F	2 ± 2
G	8 ± 3
H	16 ± 5
I	44 ± 10
J	9 ± 3
K	23 ± 7
L	8 ± 4

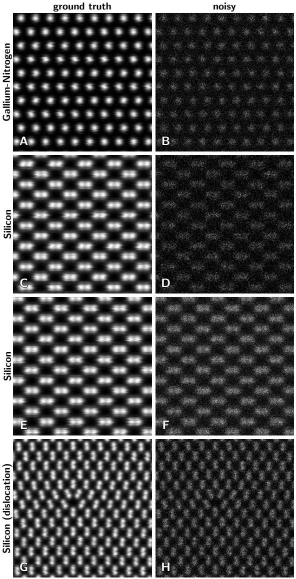
Result: improve precision by factors 5 to 10

...before and after...Perfect Silicon





...before and after...Perfect Silicon Gallium Nitride



Methods

Key ingredients:

- multilevel variational non-rigid registration scheme
  (Benjamin Berkels, AICES)

Methods

Key ingredients:

- multilevel variational non-rigid registration scheme (Benjamin Berkels, AICES)
- Nonlocal means - BM3D algorithm

Methods

Key ingredients:

- multilevel variational non-rigid registration scheme (Benjamin Berkels, AICES)
- Nonlocal means - BM3D algorithm
- Poisson noise removal

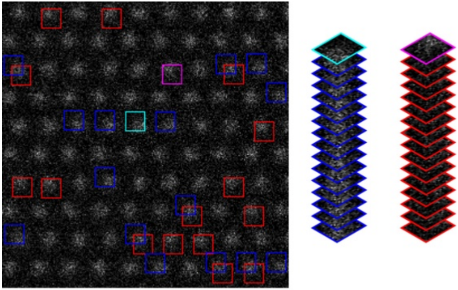
Methods

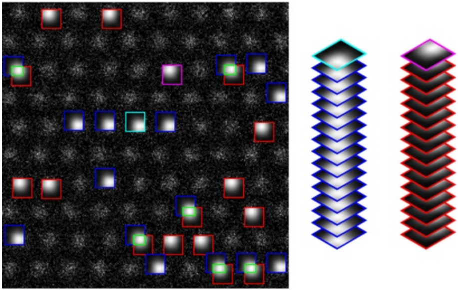
Key ingredients:

- multilevel variational non-rigid registration scheme (Benjamin Berkels, AICES)
- Nonlocal means - BM3D algorithm
- Poisson noise removal

A. B. Yankovich, B. Berkels, W. Dahmen, P. Binev, S.I. Sanchez, S.A. Bradley, and P.M. Voyles, Picometer precision STEM imaging of Pt Nanocatalysts, Nature Communications 5, Article number: 4155 (2014) doi:10.1038/ncomms5155

N. Mevenkamp, P. Binev, W. Dahmen, P.M. Voyles, A.B. Yankovich, and B. Berkels, Poisson noise removal from high-resolution STEM images based on periodic block matching, Advanced Structural and Chemical Imaging, DOI 10.1186/s40679-015-0004-8



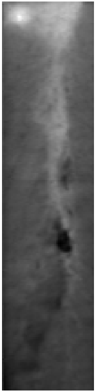


Outline

EELS-Images

electron energy loss spectroscopy ⇝ "function valued pixels"

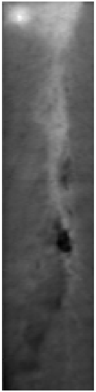
Collaborations: B. Berkels (AICES), M. Duchamps (FZ Jülich), C. Bajaj, ICES, University of Texas at Austin

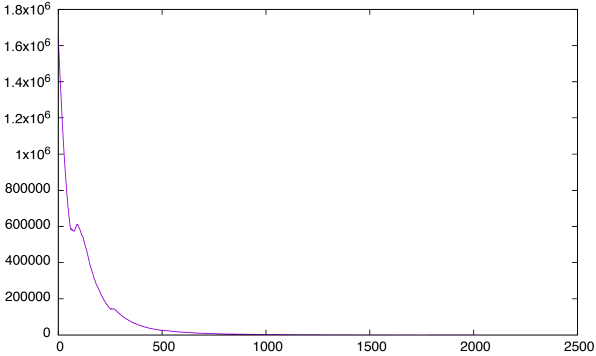


Example task: unmixing of materials

EELS-Images electron energy loss spectroscopy ⇝ “function valued pixels”

Collaborations: B. Berkels (AICES), M. Duchamps (FZ Jülich), C. Bajaj, ICES, University of Texas at Austin



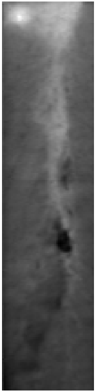


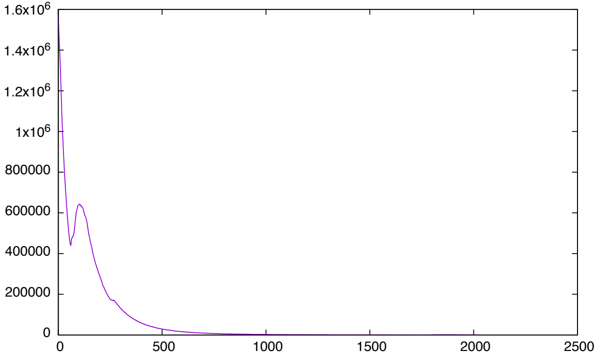
Example task: unmixing of materials

EELS-Images

electron energy loss spectroscopy ⇝ "function valued pixels"

Collaborations: B. Berkels (AICES), M. Duchamps (FZ Jülich), C. Bajaj, ICES, University of Texas at Austin



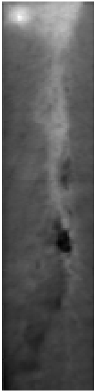


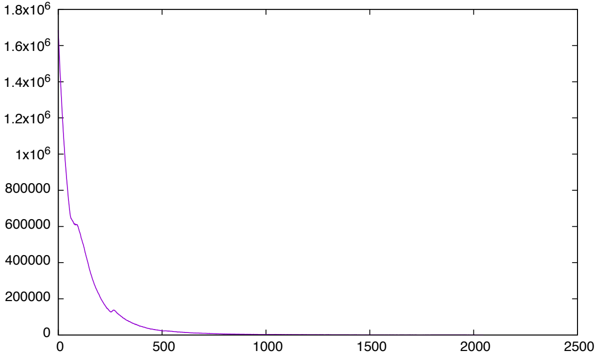
Example task: unmixing of materials

EELS-Images

electron energy loss spectroscopy ⇝ "function valued pixels"

Collaborations: B. Berkels (AICES), M. Duchamps (FZ Jülich), C. Bajaj, ICES, University of Texas at Austin



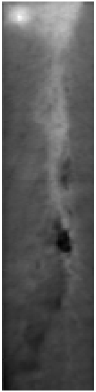


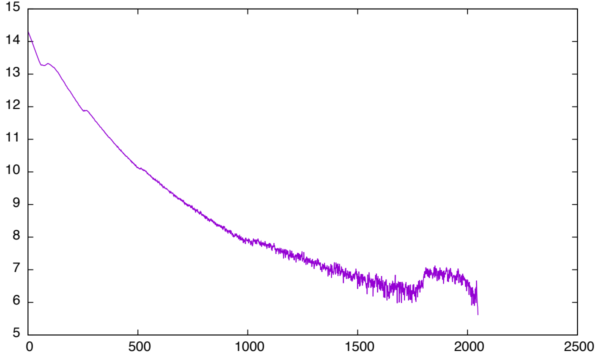
Example task: unmixing of materials

EELS-Images

electron energy loss spectroscopy ⇝ "function valued pixels"

Collaborations: B. Berkels (AICES), M. Duchamps (FZ Jülich), C. Bajaj, ICES, University of Texas at Austin



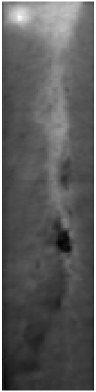


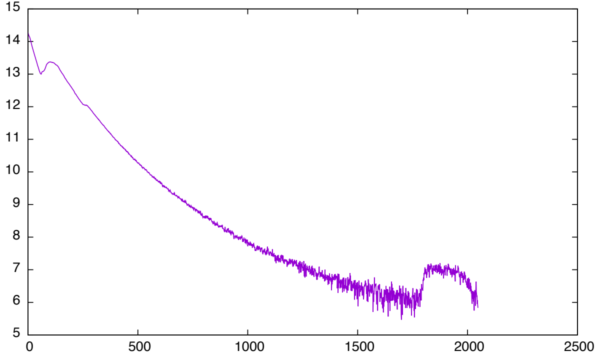
Example task: unmixing of materials

EELS-Images

electron energy loss spectroscopy ⇝ "function valued pixels"

Collaborations: B. Berkels (AICES), M. Duchamps (FZ Jülich), C. Bajaj, ICES, University of Texas at Austin



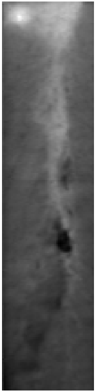


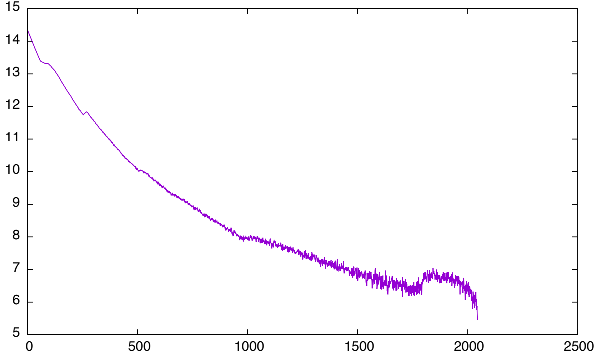
Example task: unmixing of materials

EELS-Images

electron energy loss spectroscopy ⇝ "function valued pixels"

Collaborations: B. Berkels (AICES), M. Duchamps (FZ Jülich), C. Bajaj, ICES, University of Texas at Austin





Example task: unmixing of materials

Obstructions, Methods

Obstructions:

- very large, noisy data sets

Obstructions, Methods

Obstructions:

- very large, noisy data sets
- highly heterogeneous spectrum structure

Obstructions, Methods

Obstructions:

- very large, noisy data sets
- highly heterogeneous spectrum structure
- noise structure ?

Obstructions, Methods

Obstructions:

- very large, noisy data sets
- highly heterogeneous spectrum structure
- noise structure ?

Methods:

- nonlocal means, BMD3-algorithm

Obstructions, Methods

Obstructions:

- very large, noisy data sets
- highly heterogeneous spectrum structure
- noise structure ?

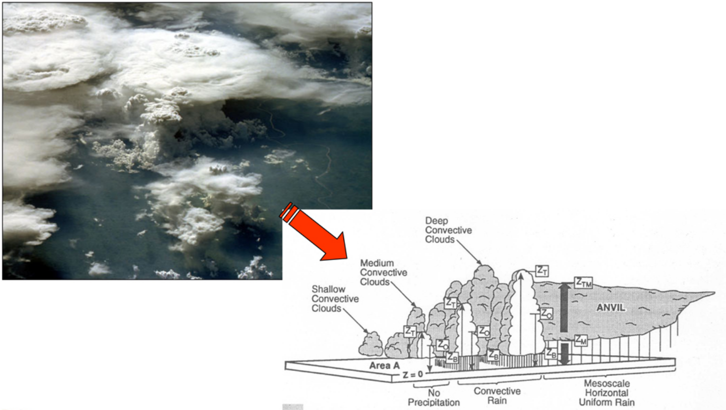
Methods:

- nonlocal means, BMD3-algorithm
- future: compressed sensing, dictionary learning, tensor methods,...

Outline

NCAR National Center of Atmospheric Research CAM Community Atmospheric Model

Collaborators: P. Binev, R. DeVore, P. Lamby, (M. Fox-Rabinowitz, V. Krasnopolsky)



NCAR

NCAR National Center of Atmospheric Research CAM Community Atmospheric Model

Collaborators: P. Binev, R. DeVore, P. Lamby, (M. Fox-Rabinowitz, V. Krasnopolsky)

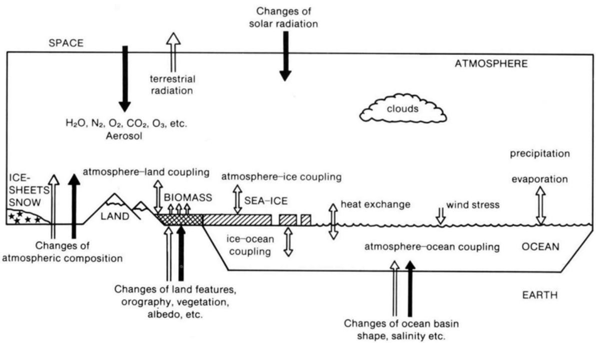


Figure 3.1: Schematic illustration of the components of the coupled atmosphere-ocean-ice-land climatic system. The full arrows are examples of external processes, and the open arrows are examples of internal processes in climatic change (from Houghton, 1984).

NCAR

igpm

NCAR Model general Circulation Model

$$d\overline{\mathbf{V}}/dt + fk \times \overline{\mathbf{V}} + \nabla\overline{\phi} = \mathbf{F}, \qquad \textit{(horizontal momentum)}$$

$$d\overline{T}/dt - \kappa\overline{T}\omega/p = Q/c_p, \qquad \textit{(thermodynamic energy)}$$

$$\nabla \cdot \overline{\mathbf{V}} + \partial\overline{\omega}/\partial p = 0, \qquad \textit{(mass continuity)}$$

$$\partial\overline{\phi}/\partial p + R\overline{T}/p = 0, \qquad \textit{(hydrostatic equilibrium)}$$

$$d\overline{q}/dt = S_q. \qquad \textit{(water vapor mass continuity)}$$

Harmless looking terms F, Q, and S_q $\Longrightarrow$ "physics"

NCAR CAM

General Circulation Model: fluid dynamics equation on the sphere

$$\partial_t \psi + D(\psi) = P(\psi)$$

NCAR CAM

General Circulation Model: fluid dynamics equation on the sphere

$$\partial_t \psi + D(\psi) = P(\psi)$$

- ψ set of variables representing temperature, wind, pressure, moisture, etc.;

NCAR CAM

General Circulation Model: fluid dynamics equation on the sphere

$$\partial_t \psi + D(\psi) = P(\psi)$$

- ψ set of variables representing temperature, wind, pressure, moisture, etc.;
- D represents the model dynamics constituted by the processes of motion and thermodynamics;

NCAR CAM

General Circulation Model: fluid dynamics equation on the sphere

$$\partial_t \psi + D(\psi) = P(\psi)$$

- ψ set of variables representing temperature, wind, pressure, moisture, etc.;
- D represents the model dynamics constituted by the processes of motion and thermodynamics;
- P stands for model physics (e.g., atmospheric radiation, turbulence, convection and large-scale precipitations, clouds, Interactions between land and ocean, etc.)

NCAR CAM

General Circulation Model: fluid dynamics equation on the sphere

$$\partial_t \psi + D(\psi) = P(\psi)$$

- ψ set of variables representing temperature, wind, pressure, moisture, etc.;
- D represents the model dynamics constituted by the processes of motion and thermodynamics;
- P stands for model physics (e.g., atmospheric radiation, turbulence, convection and large-scale precipitations, clouds, Interactions between land and ocean, etc.)

See: National Center of Atmospheric Research Community Atmospheric Model (NCAR CAM)

NCAR CAM

General Circulation Model: fluid dynamics equation on the sphere

$$\partial_t \psi + D(\psi) = P(\psi)$$

- ψ set of variables representing temperature, wind, pressure, moisture, etc.;
- D represents the model dynamics constituted by the processes of motion and thermodynamics;
- P stands for model physics (e.g., atmospheric radiation, turbulence, convection and large-scale precipitations, clouds, Interactions between land and ocean, etc.)

See: National Center of Atmospheric Research Community Atmospheric Model (NCAR CAM)

Bottleneck for numerical simulations: modeling physics/chemistry
Long Wave Radiation: 70% – 90% of total cost

Problem and Goal

$$\partial_t \psi + D(\psi) = P(\psi)$$

Problem and Goal

$\partial_t \psi + D(\psi) = P(\psi)$ "Learn" $P(\psi)$ from data

Problem and Goal

$$\partial_t \psi + D(\psi) = P(\psi) \qquad \text{“Learn” } P(\psi) \text{ from data}$$

For LWR recovery $\rightsquigarrow$

$$f : X \subset \mathbb{R}^{220} \to \mathbb{R}^{33}$$

from data $\mathbf{z} = \{z^i = (x^i, y^i) : i = 1, \ldots, N\} \subset X \times Y$, where

$$N \sim 10^5$$

Inputs $x \in X$ consist of 10 vertical profiles of local physical properties and gas concentrations plus a surface characteristic.

Outputs y consist of a vertical profile of heating rates and some radiation fluxes

outputs $y = f(x)$ are computed via solving the physical model with input x

Key Issues

- Severe undersampling - Classical approximation schemes are not feasible
- Recover functions of many variables - curse of dimensionality

Key Issues

- Severe undersampling - Classical approximation schemes are not feasible
- Recover functions of many variables - curse of dimensionality
- How to measure the quality of an approximation?

$$\text{accuracy } \varepsilon \quad \leftrightarrow \quad \text{computational work } O(\varepsilon^{-d/\alpha})$$

Key Issues

- Severe undersampling - Classical approximation schemes are not feasible
- Recover functions of many variables - curse of dimensionality
- How to measure the quality of an approximation?

$$\text{accuracy } \varepsilon \quad \leftrightarrow \quad \text{computational work } O(\varepsilon^{-d/\alpha})$$

- Meaningful results are only possible if the data are highly correlated - implicit lower dimensionality

Key Issues

- Severe undersampling - Classical approximation schemes are not feasible
- Recover functions of many variables - curse of dimensionality
- How to measure the quality of an approximation?

  accuracy $\varepsilon \quad \leftrightarrow \quad$ computational work $O(\varepsilon^{-d/\alpha})$

- Meaningful results are only possible if the data are highly correlated - implicit lower dimensionality
- Key question: How much of the structure does an algorithm have to know in order to take advantage of it?

Procedural Definition of Function Recovery

learn f from $\mathbf{z} = \{(x^1, y^1), \ldots, (x^N, y^N)\}$

Procedural Definition of Function Recovery

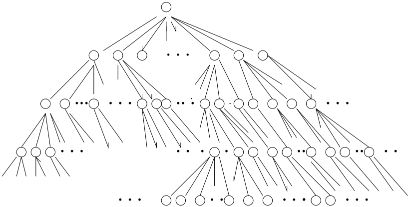
learn f from $\mathbf{z} = \{(x^1, y^1), \dots, (x^N, y^N)\}$
Candidates:

- kernel methods, artificial neural networks
- Recovery based on k nearest neighbor search
- Tree based schemes: random forests, sparse occupancy trees

Sparse Occupancy Trees

Hierarchy of nested partitions of X

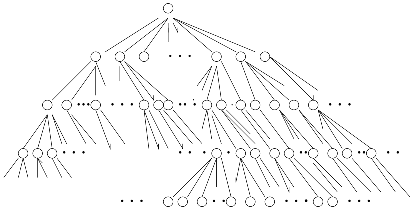
$$X = \mathcal{P}_\emptyset \prec \mathcal{P}_0 \prec \mathcal{P}_1 \prec \cdots, \qquad X_{j,k} \in \mathcal{P}_j \rightsquigarrow X_{j,k} = \bigcup_{i \in \mathcal{I}_{j,k}} X_{j+1,i}$$



Sparse Occupancy Trees

Hierarchy of nested partitions of X

$$X = \mathcal{P}_\emptyset \prec \mathcal{P}_0 \prec \mathcal{P}_1 \prec \cdots, \qquad X_{j,k} \in \mathcal{P}_j \rightsquigarrow X_{j,k} = \bigcup_{i \in \mathcal{I}_{j,k}} X_{j+1,i}$$



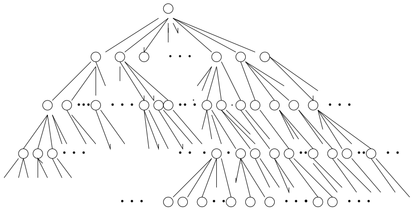
Cost Summary:

- $\mathcal{T}(X)$ has at most $1 + LN$ nodes

Sparse Occupancy Trees

Hierarchy of nested partitions of X

$$X = \mathcal{P}_\emptyset \prec \mathcal{P}_0 \prec \mathcal{P}_1 \prec \cdots, \qquad X_{j,k} \in \mathcal{P}_j \rightsquigarrow X_{j,k} = \bigcup_{i \in \mathcal{I}_{j,k}} X_{j+1,i}$$



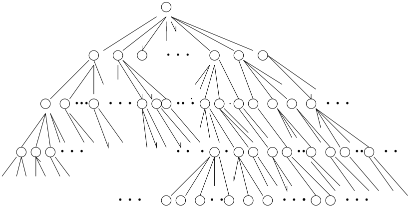
Cost Summary:

- $\mathcal{T}(X)$ has at most $1 + LN$ nodes
- $O(LdN)$ operations to collect information about X

Sparse Occupancy Trees

Hierarchy of nested partitions of X

$$X = \mathcal{P}_\emptyset \prec \mathcal{P}_0 \prec \mathcal{P}_1 \prec \cdots, \qquad X_{j,k} \in \mathcal{P}_j \rightsquigarrow X_{j,k} = \bigcup_{i \in \mathcal{I}_{j,k}} X_{j+1,i}$$



Cost Summary:

- $\mathcal{T}(X)$ has at most $1 + LN$ nodes
- $O(LdN)$ operations to collect information about X
- $O(N \log N)$ sorting operations of bitstreams of length $\leq Ld$

Simulation Results

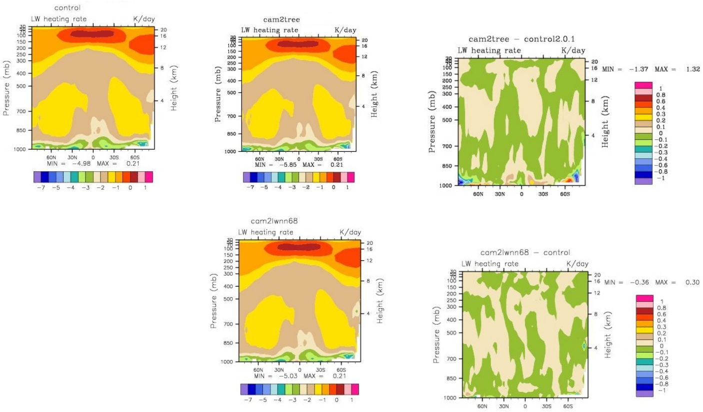


Figure: Comparison of the predicted annual zonal means of the LWR heating rates computed with the original parameterization (1 in 1. row), a tree based emulation (2, 3 in 1. row) and a neural network emulation (2. row). The right column plots the difference between the simulation and the control.

Simulation Results

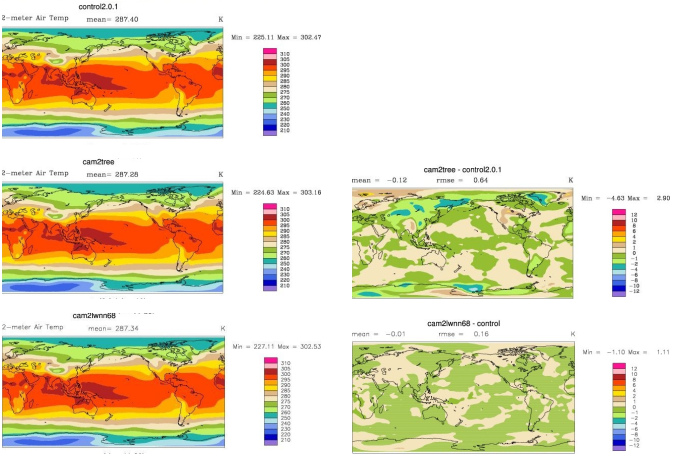
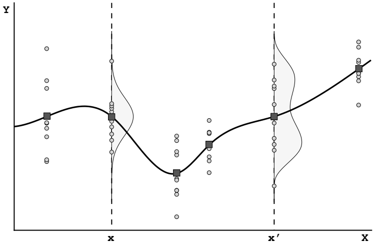


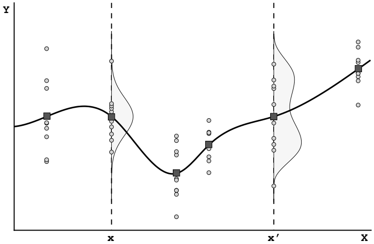
Figure: Comparison of the predicted annual means of the two meter air temperatures computed with the original parameterization (top row), a tree-based emulation (center row) and a neural network emulation (bottom row). The right column plots the difference between the simulation and the control.

Outline

Setting - Model



Setting - Model



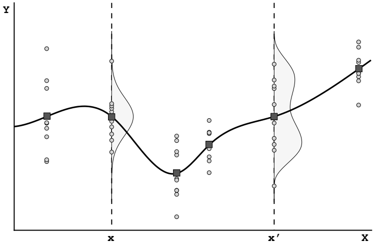
ρ unknown measure on

$Z := X \times Y$

$d\rho(x, y) = d\rho(y|x)d\rho_X(x)$

$\mathrm{supp}\,\rho(\cdot|x) \subseteq [-M.M],\ x \in X$

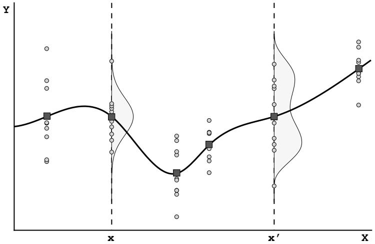
Setting - Model



ρ unknown measure on
$Z := X \times Y$
$d\rho(x, y) = d\rho(y|x)d\rho_X(x)$
$\operatorname{supp} \rho(\cdot|x) \subseteq [-M.M], \; x \in X$
Goal: estimate the regression function

$$f_\rho(x) := \int_Y y d\rho(y|x) = \mathbb{E}(y|x)$$

Setting - Model

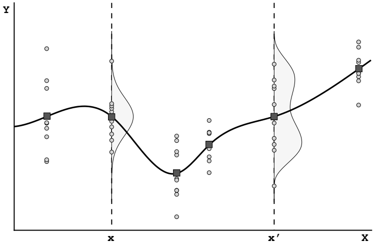


ρ unknown measure on
$Z := X \times Y$
$d\rho(x, y) = d\rho(y|x)d\rho_X(x)$
$\text{supp}\, \rho(\cdot|x) \subseteq [-M.M],\ x \in X$
Goal: estimate the regression function

$$f_\rho(x) := \int_Y y d\rho(y|x) = \mathbb{E}(y|x)$$

Risk functional: $\mathcal{E}(f) := \int_Z (y - f(x))^2 d\rho$

Setting - Model



ρ unknown measure on
$Z := X \times Y$
$d\rho(x,y) = d\rho(y|x)d\rho_X(x)$
$\operatorname{supp} \rho(\cdot|x) \subseteq [-M.M],\ x \in X$
Goal: estimate the regression function

$$f_\rho(x) := \int_Y y d\rho(y|x) = \mathbb{E}(y|x)$$

Risk functional: $\mathcal{E}(f) := \int_Z (y - f(x))^2 d\rho \rightsquigarrow$

$$\mathcal{E}(f) = \mathcal{E}(f_\rho) + \|f - f_\rho\|^2_{L_2(X,\rho_X)}, \quad \|\cdot\| := \|\cdot\|_{L_2(X,\rho_X)}$$

Estimators

Key issue: proper balance of bias and variance

- Adaptive Tree Partitioning
- Greedy algorithms

$\rightsquigarrow$ universality, i.e., realization of best rates without a priori knowledge about the searched regression function

A Performance Theorem

Bench mark: $f \in L_2(X, \rho_X) \rightsquigarrow \mathcal{T}(f, \eta)$ via $L_2(X, \rho_X)$- projections

$$|f|^p_{\mathcal{B}^s} := \sup_{\eta > 0} \eta^p \#(\mathcal{T}(f, \eta)), \quad \text{where } p := (s + 1/2)^{-1} \quad \rightsquigarrow$$

A Performance Theorem

Bench mark: $f \in L_2(X, \rho_X) \rightsquigarrow \mathcal{T}(f, \eta)$ via $L_2(X, \rho_X)$- projections

$$|f|_{\mathcal{B}^s}^p := \sup_{\eta > 0} \eta^p \#(\mathcal{T}(f, \eta)), \quad \text{where } p := (s + 1/2)^{-1} \quad \rightsquigarrow$$

$$\|f - P_\eta(f)\| \le C_s |f|_{\mathcal{B}^s}^{\frac{1}{2s+1}} \eta^{\frac{2s}{2s+1}} \le C_s |f|_{\mathcal{B}^s} N^{-s}, \quad N := \#(\mathcal{T}(f, \eta))$$

A Performance Theorem

Bench mark: $f \in L_2(X, \rho_X) \rightsquigarrow \mathcal{T}(f, \eta)$ via $L_2(X, \rho_X)$- projections

$$|f|^p_{\mathcal{B}^s} := \sup_{\eta > 0} \eta^p \#(\mathcal{T}(f, \eta)), \quad \text{where } p := (s + 1/2)^{-1} \quad \rightsquigarrow$$

$$\|f - P_\eta(f)\| \le C_s |f|^{\frac{1}{2s+1}}_{\mathcal{B}^s} \eta^{\frac{2s}{2s+1}} \le C_s |f|_{\mathcal{B}^s} N^{-s}, \quad N := \#(\mathcal{T}(f, \eta))$$

THEOREM:

If $f_\rho \in \mathcal{B}^s$ for some $s > 0$ then for $\beta > 0$ (arbitrary fixed), n samples, then

and need not know s.

A Performance Theorem

Bench mark: $f \in L_2(X,\rho_X) \rightsquigarrow \mathcal{T}(f,\eta)$ via $L_2(X,\rho_X)$- projections

$$|f|^p_{\mathcal{B}^s} := \sup_{\eta>0} \eta^p \#(\mathcal{T}(f,\eta)), \quad \text{where } p := (s+1/2)^{-1} \quad \rightsquigarrow$$

$$\|f - P_\eta(f)\| \le C_s |f|^{\frac{1}{2s+1}}_{\mathcal{B}^s} \eta^{\frac{2s}{2s+1}} \le C_s |f|_{\mathcal{B}^s} N^{-s}, \quad N := \#(\mathcal{T}(f,\eta))$$

THEOREM:

If $f_\rho \in \mathcal{B}^s$ for some $s > 0$ then for $\beta > 0$ (arbitrary fixed), n samples, then

$$\mathbb{P}_{\rho^n}\left\{\mathbf{z} : \|f_\rho - \hat{f}_{\mathbf{z}}\| \ge c\left(\frac{d \log n}{n}\right)^{\frac{s}{2s+1}}\right\} \le Cn^{-\beta}$$

and need not know s.

A Performance Theorem

Bench mark: $f \in L_2(X, \rho_X) \rightsquigarrow \mathcal{T}(f, \eta)$ via $L_2(X, \rho_X)$- projections

$$|f|^p_{\mathcal{B}^s} := \sup_{\eta > 0} \eta^p \#(\mathcal{T}(f, \eta)), \quad \text{where } p := (s + 1/2)^{-1} \quad \rightsquigarrow$$

$$\|f - P_\eta(f)\| \le C_s |f|^{\frac{1}{2s+1}}_{\mathcal{B}^s} \eta^{\frac{2s}{2s+1}} \le C_s |f|_{\mathcal{B}^s} N^{-s}, \quad N := \#(\mathcal{T}(f, \eta))$$

THEOREM:

If $f_\rho \in \mathcal{B}^s$ for some $s > 0$ then for $\beta > 0$ (arbitrary fixed), n samples, then

$$\mathbb{P}_{\rho^n}\left\{ \mathbf{z} : \|f_\rho - \hat{f}_{\mathbf{z}}\| \ge c \left(\frac{d \log n}{n} \right)^{\frac{s}{2s+1}} \right\} \le C n^{-\beta} \quad \Rightarrow$$

$$\mathbb{E}(\|f_\rho - f_{\mathbf{z}}\|^2) \le C \left(\frac{d \log N}{N} \right)^{\frac{2s}{2s+1}}$$

Moreover, the scheme is universally consistent and need not know s.

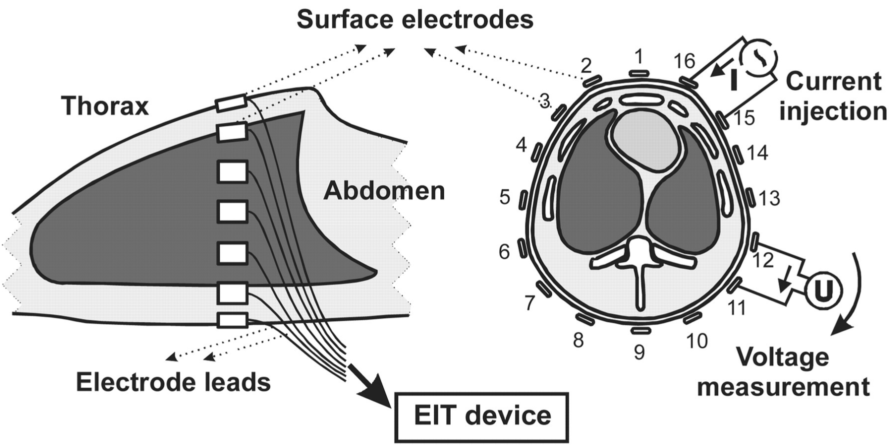
Comments

- the convergence rates are best possible for a given "regularity" order $s > 0$;
- the smaller s the weaker the hypothesis, $\mathcal{B}^0 = L_2(X, \rho_X)$;
- to achieve this rate the algorithm does not need to know the property $f_\rho \in \mathcal{B}^s$ but automatically exploits such a property (adaptivity $\rightsquigarrow$ universality);
- for arbitrary densities ρ and piecewise polynomial estimators the estimates in probability hold only for piecewise constants, higher orders require restrictions on ρ, the rate in expectation holds for higher order piecewise polynomial estimators;
- analogous results are valid for tree based adaptive classifiers

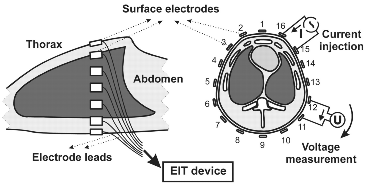
References:

- P. Binev, A. Cohen, W. Dahmen, R. DeVore, V. Temlyakov, Universal algorithms for learning theory - Part I : piecewise constant functions, Journal of Machine Learning Research (JMLR), 6(2005), 1297–1321.
- A. Barron, A. Cohen, W. Dahmen, R. DeVore, Approximation and learning by greedy algorithms, Annals of Statistics, 3(No 1)(2008), 64–94.
- P. Binev, A. Cohen, W. Dahmen, R. DeVore, Classification algorithms using adaptive partitioning, Annals of Statistics, 42 (No. 6)(2014), 2141-2163.
- P. Binev, W. Dahmen, P. Lamby, Fast High-Dimensional Approximation with Sparse Occupancy Trees, Journal of Computational and Applied Mathematics, 235 (2011), pp. 2063-2076.

Example: Electron Impedance Tomography: “many parameters...”

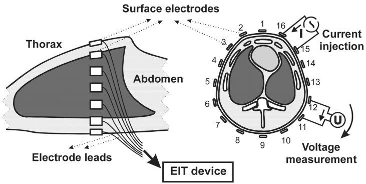


Example: EIT: Model



$a(\omega, x) : \Omega \times D \to \mathbb{R}$ (unknown) random conductivity field

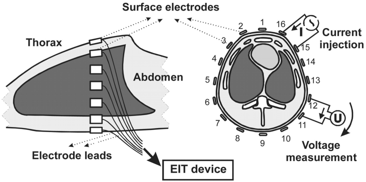
Example: EIT: Model



$a(\omega, x) : \Omega \times D \to \mathbb{R}$ (unknown) random conductivity field

Forward problem: given $\boldsymbol{I} = (I_1, \ldots, I_M) \in \mathbb{R}_0^M$, find $(u, \boldsymbol{U}) \in L^2_\rho(\Omega; X)$, $X := H^1(D) \times \mathbb{R}_0^M$, s.t.

Example: EIT: Model



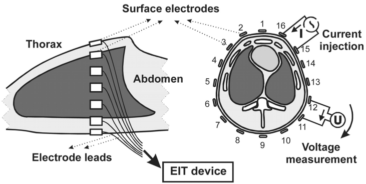
$a(\omega, x) : \Omega \times D \to \mathbb{R}$ (unknown) random conductivity field

Forward problem: given $\boldsymbol{I} = (I_1, \ldots, I_M) \in \mathbb{R}_0^M$, find $(u, \boldsymbol{U}) \in L_\rho^2(\Omega; X)$, $X := H^1(D) \times \mathbb{R}_0^M$, s.t.

$$\mathrm{div}(a\nabla u) = 0, \quad \text{in } D, \quad \partial_n u = 0 \quad \text{on } \partial D \setminus \bar{E}$$

$$u + z_m a\partial_n u = U_m \text{ on } E_m, \quad \int_{E_m} a\partial_n u ds = I_m, \quad m = 1, \ldots, M$$

Example: EIT: Model



$a(\omega, x) : \Omega \times D \to \mathbb{R}$ (unknown) random conductivity field

Forward problem: given $\boldsymbol{I} = (I_1, \ldots, I_M) \in \mathbb{R}_0^M$, find $(u, \boldsymbol{U}) \in L^2_\rho(\Omega; X)$, $X := H^1(D) \times \mathbb{R}_0^M$, s.t.

$$\mathrm{div}(a \nabla u) = 0, \quad \text{in } D, \quad \partial_n u = 0 \quad \text{on } \partial D \setminus \bar{E}$$

$$u + z_m a \partial_n u = U_m \text{ on } E_m, \quad \int_{E_m} a \partial_n u ds = I_m, \quad m = 1, \ldots, M$$

Inverse problem: find $a(\omega, x) = a_0(x) + \sum_{j \in \mathbb{N}} y_j \psi_j(x) = a(y, x)$

Solution Manifold...

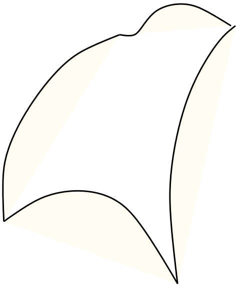
Find $u \in \mathcal{H}$, s.t.

$$F(u, y) = 0, \quad y \in \mathcal{Y} \quad \rightsquigarrow \mathcal{M} := \{u(y) \in \mathcal{H} : F(u(y), y) = 0\}$$

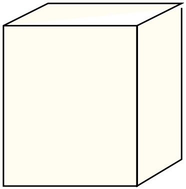
Solution Manifold...

Find $u \in \mathcal{H}$, s.t.

$$F(u, y) = 0, \quad y \in \mathcal{Y} \quad \rightsquigarrow \mathcal{M} := \{u(y) \in \mathcal{H} : F(u(y), y) = 0\}$$



solution manifold $\mathcal{M}$

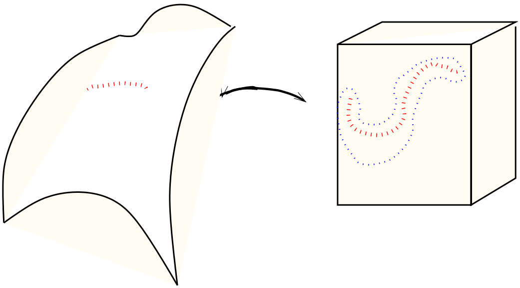


parameter domain $\mathcal{Y}$

Solution Manifold...

Find $u \in \mathcal{H}$, s.t.

$$F(u, y) = 0, \quad y \in \mathcal{Y} \quad \rightsquigarrow \mathcal{M} := \{u(y) \in \mathcal{H} : F(u(y), y) = 0\}$$



data: $d = \ell(u(y^*)) \in \mathbb{R}^d$ find the parameters $y^* \in \mathcal{Y}$

The Role of Reduced Models...

- highly underdetermined, severely ill-posed

The Role of Reduced Models...

- highly underdetermined, severely ill-posed
- use sparsity, smoothness of $\mathcal{M} \rightsquigarrow$ highly accurate certified reduced models

The Role of Reduced Models...

- highly underdetermined, severely ill-posed
- use sparsity, smoothness of $\mathcal{M}$ $\rightsquigarrow$ highly accurate certified reduced models
- better regularization?

The Role of Reduced Models...

- highly underdetermined, severely ill-posed
- use sparsity, smoothness of $\mathcal{M} \rightsquigarrow$ highly accurate certified reduced models
- better regularization?
- train reduced models through conditioned sampling, MCMC methods,

The Role of Reduced Models...

- highly underdetermined, severely ill-posed
- use sparsity, smoothness of $\mathcal{M}$ $\rightsquigarrow$ highly accurate certified reduced models
- better regularization?
- train reduced models through conditioned sampling, MCMC methods,
- combination with classification schemes...in progress...

The Role of Reduced Models...

- highly underdetermined, severely ill-posed
- use sparsity, smoothness of $\mathcal{M} \rightsquigarrow$ highly accurate certified reduced models
- better regularization?
- train reduced models through conditioned sampling, MCMC methods,
- combination with classification schemes...in progress...

P.Binev, A.Cohen, W. Dahmen, R.DeVore, G. Petrova, P. Wojtaszczyk, Convergence Rates for Greedy Algorithms in Reduced Basis Methods, SIAM J. Math. Anal., 43 (2011), 1457–1472.

W. Dahmen, C. Plesken, G. Welper, Double Greedy Algorithms: Reduced Basis Methods for Transport Dominated Problems, ESAIM: Mathematical Modelling and Numerical Analysis, 48(3) (2014), 623–663.
DOI 10.1051/m2an/2013103, http://arxiv.org/abs/1302.5072.

Peter Binev, Albert Cohen, Wolfgang Dahmen, Ronald DeVore, Guergana Petrova, Przemyslaw Wojtaszczyk, Data Assimilation in Reduced Modeling, Preprint June 2015,

http://arxiv.org/abs/1506.04770 [math.NA].