Compression, inversion and sparse approximate PCA of dense kernel matrices in near linear computational complexity

Florian Schäfer

ICERM 2017

Compression, inversion and approximate PCA of dense kernel matrices in near linear computational complexity

Florian Schäfer, T.J. Sullivan, Houman Owhadi

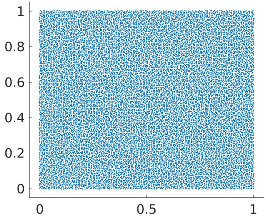
http://arxiv.org/abs/1706.02205

Outline

1. A numerical experiment
2. Disintegration of measure and Gaussian elimination
3. Near-linear complexity algorithms using the theory of Gamblets
4. Further numerical results

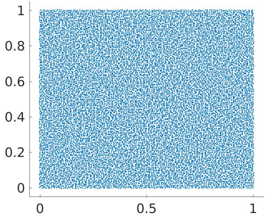
A numerical experiment

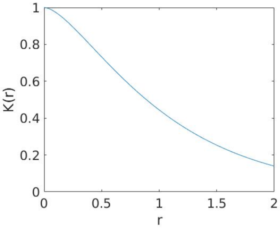
- $\{x_i\}_{i \in I} \subset [0,1]^2$, with $\# I = N = 16641$



A numerical experiment

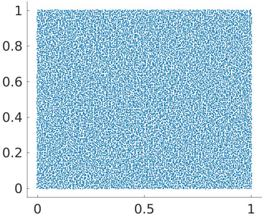
- $\{x_i\}_{i \in I} \subset [0,1]^2$, with $\#I = N = 16641$
- Define $K(r)$ as Matérn kernel with smoothness parameter $\nu = 1$ and lengthscale $l = 0.4$.

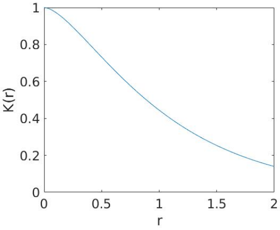


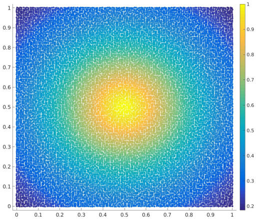


A numerical experiment

- $\{x_i\}_{i \in I} \subset [0,1]^2$, with $\#I = N = 16641$
- Define $K(r)$ as Matérn kernel with smoothness parameter $\nu = 1$ and lengthscale $l = 0.4$.
- $\Gamma_{i,j} := K\left(\|x_i - x_j\|\right)$.







A numerical experiment

- Γ, interpreted as covariance matrix, describes a Gaussian field with second order smoothness.

A numerical experiment

- Γ, interpreted as covariance matrix, describes a Gaussian field with second order smoothness.
- Alternatively, K can be seen as the Green's function of a fourth order elliptic PDE, on the whole space.

A numerical experiment

- Γ, interpreted as covariance matrix, describes a Gaussian field with fourth order smoothness.
- Alternatively, K can be seen as the Green's function of a fourth order elliptic PDE, on the whole space.
- Matrices of this kind appear in both statistics and scientific computing.

A numerical experiment

- Γ, interpreted as covariance matrix, describes a Gaussian field with fourth order smoothness.
- Alternatively, K can be seen as the Green's function of a fourth order elliptic PDE, on the whole space.
- Matrices of this kind appear in both statistics and scientific computing.
- We need to apply the Matrix and its inverse, and compute its determinant.

A numerical experiment

- Γ, interpreted as covariance matrix, describes a Gaussian field with fourth order smoothness.
- Alternatively, K can be seen as the Green's function of a fourth order elliptic PDE, on the whole space.
- Matrices of this kind appear in both statistics and scientific computing.
- We need to apply the Matrix and its inverse, and compute its determinant.
- Γ is dense, and hence has N^2 storage cost. Direct inversion via Gaussian elimination has $\mathcal{O}\left(N^3\right)$ complexity in time.

A numerical experiment

- Can we be more efficient?

A numerical experiment

- Can we be more efficient?
- Many existing methods: Quadrature formulas, subsampling, randomised approximations, low rank approximations, fast multipole methods, hierarchical matrices, wavelet methods, inducing points, covariance tapering . . . .

A numerical experiment

- Can we be more efficient?
- Many existing methods: Quadrature formulas, subsampling, randomised approximations, low rank approximations, fast multipole methods, hierarchical matrices, wavelet methods, inducing points, covariance tapering ....
- We provide a simple algorithm, with rigorous error bounds and near-linear complexity.

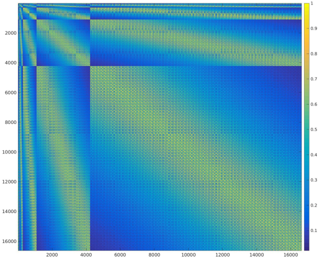
A numerical experiment

- Even writing down the matrix has N^2 complexity.

A numerical experiment

- Even writing down the matrix has N^2 complexity.
- Therefore, we subsample Γ:

$$\tilde{\Gamma}_{i,j} := \begin{cases} \Gamma_{i,j}, & \text{for } (i,j) \in S_2 \\ 0, & \text{else} \end{cases}$$

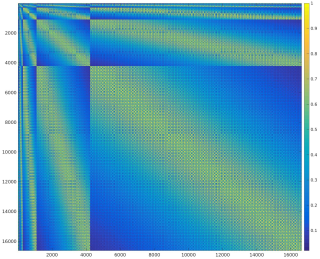


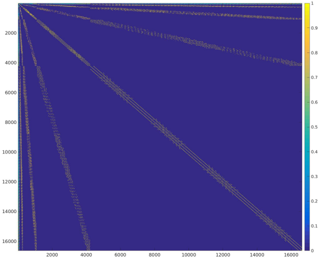
A numerical experiment

- Even writing down the matrix has N^2 complexity.
- Therefore, we subsample Γ:

$$\tilde{\Gamma}_{i,j} := \begin{cases} \Gamma_{i,j}, & \text{for } (i,j) \in S_2 \\ 0, & \text{else} \end{cases}$$

- $\# S_2 = 5528749 = 0.0189 N^2$. We have thrown away all but 2 percent of the entries, without even touching them!



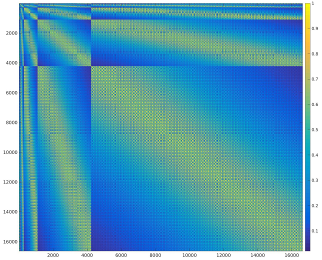


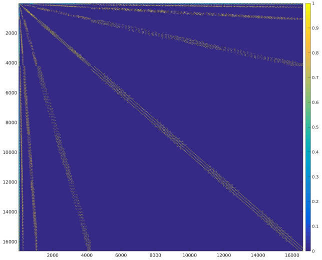
A numerical experiment

- Even writing down the matrix has N^2 complexity.
- Therefore, we subsample Γ:

$$\tilde{\Gamma}_{i,j} := \begin{cases} \Gamma_{i,j}, & \text{for } (i,j) \in S_2 \\ 0, & \text{else} \end{cases}$$

- $\# S_2 = 5528749 = 0.0189 N^2$. We have thrown away all but 2 percent of the entries, without even touching them!
- We will see later: S_2 does not depend on the entries of Γ.





A numerical experiment

- We have compressed Γ to 2 percent of its original size.

A numerical experiment

- We have compressed Γ to 2 percent of its original size.
- How much information have we retained?

A numerical experiment

- We have compressed Γ to 2 percent of its original size.
- How much information have we retained?
- Consider relative error in operator norm:

A numerical experiment

- We have compressed Γ to 2 percent of its original size.
- How much information have we retained?
- Consider relative error in operator norm:
- $$\frac{\left\|\Gamma - \tilde{\Gamma}\right\|}{\|\Gamma\|} = 0.9662$$

A numerical experiment

- We have compressed Γ to 2 percent of its original size.
- How much information have we retained?
- Consider relative error in operator norm:
- $$\frac{\left\| \Gamma - \tilde{\Gamma} \right\|}{\left\| \Gamma \right\|} = 0.9662$$
- $\tilde{\Gamma}$ is a very bad approximation of Γ.

A numerical experiment

- Can we obtain a better approximation of Γ, using only $\tilde{\Gamma}$?

A numerical experiment

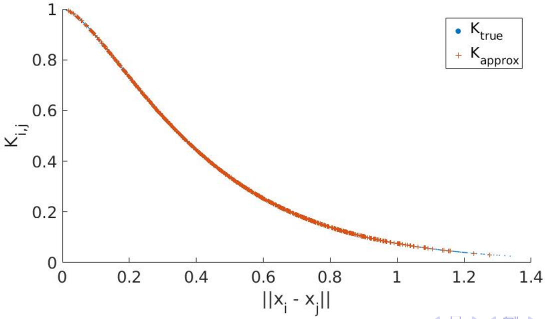
- Can we obtain a better approximation of Γ, using only $\tilde{\Gamma}$?
- Consider $L = \text{ICHOL}\left(\tilde{\Gamma}\right)$, the incomplete Cholesky factorisation of $\tilde{\Gamma}$, ignoring all fill-in.

A numerical experiment

- Can we obtain a better approximation of Γ, using only $\tilde{\Gamma}$?
- Consider $L = \mathrm{ICHOL}\left(\tilde{\Gamma}\right)$, the incomplete Cholesky factorisation of $\tilde{\Gamma}$, ignoring all fill-in.
- $$\frac{\left\|\Gamma - LL^T\right\|}{\|\Gamma\|} = 3.0676\mathrm{e}{-04}$$

A numerical experiment

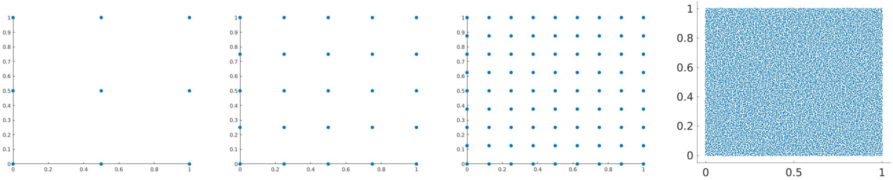
- Can we obtain a better approximation of Γ, using only $\tilde{\Gamma}$?
- Consider $L = \text{ICHOL}\left(\tilde{\Gamma}\right)$, the incomplete Cholesky factorisation of $\tilde{\Gamma}$, ignoring all fill-in.
- $$\frac{\left\|\Gamma - LL^T\right\|}{\left\|\Gamma\right\|} = 3.0676\mathrm{e}{-04}$$



A numerical experiment

- Decompose $\{x_i\}_{i \in I}$ into a nested hierarchy as:

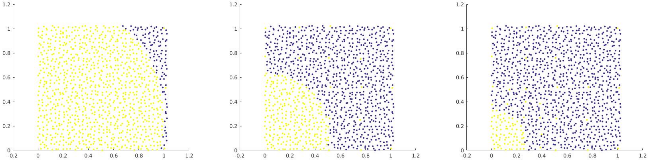
$$\{x_i\}_{i \in I^{(1)}} \subset \{x_i\}_{i \in I^{(2)}} \subset \{x_i\}_{i \in I^{(3)}} \subset \cdots \subset \{x_i\}_{i \in I^{(q)}} = \{x_i\}_{i \in I} \qquad (1.1)$$



A numerical experiment

- We define $J^{(k)} := I^{(k)} \setminus I^{(k-1)}$ and define the sparsity pattern:

$$S_2 := \left\{ (i,j) \in I \times I \middle| i \in J^{(k)}, j \in J^{(l)}, \operatorname{dist}\left(x_i, x_j\right) \leq 2 * 2^{-\min(k,l)} \right\}.$$

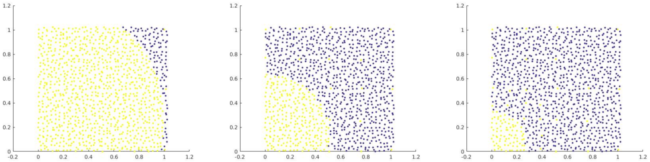


A numerical experiment

- We define $J^{(k)} := I^{(k)} \setminus I^{(k-1)}$ and define the sparsity pattern:

$$S_2 := \left\{ (i,j) \in I \times I \middle| i \in J^{(k)}, j \in J^{(l)}, \operatorname{dist}\left(x_i, x_j\right) \leq 2 * 2^{\min(k,l)} \right\}.$$

- We order the elements of I from coarse to fine, that is from $J^{(1)}$ to $J^{(q)}$.



A numerical experiment

- L provides a good approximation of Γ at only 2 percent of the storage cost.

A numerical experiment

- L provides a good approximation of Γ at only 2 percent of the storage cost.
- Can be computed in near linear complexity in time and space.

A numerical experiment

- L provides a good approximation of Γ at only 2 percent of the storage cost.
- Can be computed in near linear complexity in time and space.
- Allows for approximate evaluation of Γ, Γ^{-1}, and $\det(\Gamma)$ in near-linear time.

A numerical experiment

- L provides a good approximation of Γ at only 2 percent of the storage cost.
- Can be computed in near linear complexity in time and space.
- Allows for approximate evaluation of Γ, Γ^{-1}, and $\det(\Gamma)$ in near-linear time.
- Allows for sampling of $X \sim N(0, \Gamma)$ in near-linear time.

A numerical experiment

- In this work, we

A numerical experiment

- In this work, we
- prove that this phaenomenon holds whenever the covariance function K is the Green's function of an elliptic boundary value problem.

A numerical experiment

- In this work, we
- prove that this phaenomenon holds whenever the covariance function K is the Green's function of an elliptic boundary value problem.
- prove that it leads to an algorithm with computational complexity of $\mathcal{O}\left(N\log^2(N)\left(\log(1/\epsilon)+\log^2(N)\right)^{4d+1}\right)$ in time and $\mathcal{O}(N\log(N)\log^d(N\frac{1}{\epsilon}))$ in space for an approximation error of ϵ.

A numerical experiment

- In this work, we
- prove that this phaenomenon holds whenever the covariance function K is the Green's function of an elliptic boundary value problem.
- prove that it leads to an algorithm with computational complexity of $\mathcal{O}\left(N\log^2(N)\left(\log(1/\epsilon)+\log^2(N)\right)^{4d+1}\right)$ in time and $\mathcal{O}(N\log(N)\log^d(N\frac{1}{\epsilon}))$ in space.
- show that even though the Matérn family is not covered rigorously by our theoretical results, we get good approximation results, in particular in the interior of the domain.

A numerical experiment

- In this work, we
- prove that this phaenomenon holds whenever the covariance function K is the Green's function of an elliptic boundary value problem.
- prove that it leads to an algorithm with computational complexity of $\mathcal{O}\left(N\log^2(N)\left(\log(1/\epsilon)+\log^2(N)\right)^{4d+1}\right)$ in time and $\mathcal{O}(N\log(N)\log^d(N\frac{1}{\epsilon}))$ in space.
- show that even though the Matérn family is not covered rigorously by our theoretical results, we get good approximation results, in particular in the interior of the domain.
- show that as a byproduct of our algorithm we obtain a sparse approximate PCA with near optimal approximation property.

Disintegration of Gaussian Measures and the *Screening Effect*

- Let X be a centered Gaussian vector with covariance Θ.

Disintegration of Gaussian Measures and the *Screening Effect*

- Let X be a centered Gaussian vector with covariance Θ.
- Assume we want to compute $\mathbb{E}\left[f\left(X\right)\right]$ for some function f.

Disintegration of Gaussian Measures and the *Screening Effect*

- Let X be a centered Gaussian vector with covariance Θ.
- Assume we want to compute $\mathbb{E}\left[f\left(X\right)\right]$ for some function f.
- Use Monte Carlo, but for Θ large, each sample is expensive.

Disintegration of Gaussian Measures and the *Screening Effect*

- Let X be a centered Gaussian vector with covariance Θ.
- Assume we want to compute $\mathbb{E}\left[f\left(X\right)\right]$ for some function f.
- Use Monte Carlo, but for Θ large, each sample is expensive.
- Idea: use disintegration of measure:

$$\mathbb{E}\left[f(X)\right] = \mathbb{E}\left[\mathbb{E}[f(X)|Y]\left(Y\right)\right].$$

Disintegration of Gaussian Measures and the *Screening Effect*

- Let X be a centered Gaussian vector with covariance Θ.
- Assume we want to compute $\mathbb{E}\left[f\left(X\right)\right]$ for some function f.
- Use Monte Carlo, but for Θ large, each sample is expensive.
- Idea: use disintegration of measure:

$$\mathbb{E}\left[f(X)\right] = \mathbb{E}\left[\mathbb{E}[f(X)|Y]\left(Y\right)\right].$$

- Choose Y, such that Y and $\mathbb{E}[f(X)|Y]$ can be sampled cheaply.

Disintegration of Gaussian Measures and the *Screening Effect*

- Consider $X \in \mathbb{R}^N$, $\{x_i\}_{1 \leq i \leq N} \subset [0, 1]$ and $\Theta_{i,j} := \exp\left(-|x_i - x_j|\right)$.

Disintegration of Gaussian Measures and the *Screening Effect*

- Consider $X \in \mathbb{R}^N$, $\{x_i\}_{1 \leq i \leq N} \subset [0, 1]$ and $\Theta_{i,j} := \exp\left(-|x_i - x_j|\right)$.
- Corresponds to a prior on the space $H^1(0, 1)$ of mean square differentiable functions.

Disintegration of Gaussian Measures and the *Screening Effect*

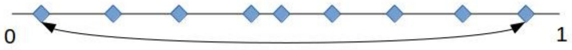
- Consider $X \in \mathbb{R}^N$, $\{x_i\}_{1 \leq i \leq N} \subset [0, 1]$ and $\Theta_{i,j} := \exp\left(-|x_i - x_j|\right)$.
- Corresponds to a prior on the space $H^1(0, 1)$ of mean square differentiable functions.
- Assume x_i are ordered in increasing order and $x_{\lfloor N/2 \rfloor} \approx 1/2$

Disintegration of Gaussian Measures and the *Screening Effect*

- Consider $X \in \mathbb{R}^N$, $\{x_i\}_{1 \leq i \leq N} \subset [0,1]$ and $\Theta_{i,j} := \exp\left(-|x_i - x_j|\right)$.
- Corresponds to a prior on the space $H^1\left(0,1\right)$ of mean square differentiable functions.
- Assume x_i are ordered in increasing order and $x_{\lfloor N/2 \rfloor} \approx 1/2$
- We then have, for $i < \lfloor N/2 \rfloor < j$: $\operatorname{Cov}\left[X_i, X_j | X_{\lfloor N/2 \rfloor}\right] \approx 0$.

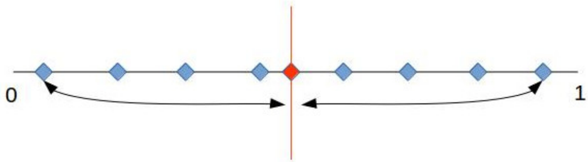
Disintegration of Gaussian Measures and the *Screening Effect*

- Consider $X \in \mathbb{R}^N$, $\{x_i\}_{1 \leq i \leq N} \subset [0,1]$ and $\Theta_{i,j} := \exp\left(-|x_i - x_j|\right)$.
- Corresponds to a prior on the space $H^1\left(0,1\right)$ of mean square differentiable functions.
- Assume x_i are ordered in increasing order and $x_{\lfloor N/2 \rfloor} \approx 1/2$
- We then have, for $i < \lfloor N/2 \rfloor < j$: $\operatorname{Cov}\left[X_i, X_j | X_{\lfloor N/2 \rfloor}\right] \approx 0$.



Disintegration of Gaussian Measures and the *Screening Effect*

- Consider $X \in \mathbb{R}^N$, $\{x_i\}_{1 \leq i \leq N} \subset [0,1]$ and $\Theta_{i,j} := \exp\left(-|x_i - x_j|\right)$.
- Corresponds to a prior on the space $H^1(0,1)$ of mean square differentiable functions.
- Assume x_i are ordered in increasing order and $x_{\lfloor N/2 \rfloor} \approx 1/2$
- We then have, for $i < \lfloor N/2 \rfloor < j$: $\operatorname{Cov}\left[X_i, X_j | X_{\lfloor N/2 \rfloor}\right] \approx 0$.



Disintegration of Gaussian Measures and the *Screening Effect*

- For two observation sites x_i, x_j, the covariance conditional on the obervation sites inbetween is small.

Disintegration of Gaussian Measures and the *Screening Effect*

- For two observation sites x_i, x_j, the covariance conditional on the obervation sites inbetween is small.
- Known as *screening effect* in the spatial statistics community. Analysed by Stein (2002). Used, among others, by Banerjee et al. (2008) and Katzfuss (2015) for efficient approximation of Gaussian processes.

Disintegration of Gaussian Measures and the *Screening Effect*

- For two observation sites x_i, x_j, the covariance conditional on the obervation sites inbetween is small.
- Known as *screening effect* in the spatial statistics community. Analysed by Stein (2002). Used, among others, by Banerjee et al. (2008) and Katzfuss (2015) for efficient approximation of Gaussian processes.
- Let us take $Y = X_{\lfloor N/2 \rfloor}$. Then Y is cheap to sample, and the covariance matrix of $X|Y$ has only $2\,(N/2)^2$ noneglegible entries.

Disintegration of Gaussian Measures and the *Screening Effect*

- For two observation sites x_i, x_j, the covariance conditional on the obervation sites inbetween is small.
- Known as *screening effect* in the spatial statistics community. Analysed by Stein (2002). Used, among others, by Banerjee et al. (2008) and Katzfuss (2015) for efficient approximation of Gaussian processes.
- Let us take $Y = X_{\lfloor N/2 \rfloor}$. Then Y is cheap to sample, and the covariance matrix of $X|Y$ has only $2\,(N/2)^2$ noneglegible entries.
- When using Cholesky decomposition, this yields a factor 4 improvement of computational speed.

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- Look at a single step of Block Cholesky decomposition:

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- Look at a single step of Block Cholesky decomposition:
- This corresponds to:

$$
\begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{pmatrix}
= \begin{pmatrix} \mathrm{Id} & 0 \\ \Theta_{21}\Theta_{11}^{-1} & \mathrm{Id} \end{pmatrix}
\begin{pmatrix} \Theta_{11} & 0 \\ 0 & \Theta_{22} - \Theta_{21}\Theta_{11}^{-1}\Theta_{12} \end{pmatrix}
\begin{pmatrix} \mathrm{Id} & \Theta_{11}^{-1}\Theta_{12} \\ 0 & \mathrm{Id} \end{pmatrix}
$$

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- Look at a single step of Block Cholesky decomposition:
- This corresponds to:

$$
\begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{pmatrix}
= \begin{pmatrix} \mathrm{Id} & 0 \\ \Theta_{21}\Theta_{11}^{-1} & \mathrm{Id} \end{pmatrix}
\begin{pmatrix} \Theta_{11} & 0 \\ 0 & \Theta_{22} - \Theta_{21}\Theta_{11}^{-1}\Theta_{12} \end{pmatrix}
\begin{pmatrix} \mathrm{Id} & \Theta_{11}^{-1}\Theta_{12} \\ 0 & \mathrm{Id} \end{pmatrix}
$$

- Note, that for $\left(\Theta_{21}\Theta_{11}^{-1}\right) b = \mathbb{E}\left[X_2 | X_1 = b\right]$, and $\Theta_{22} - \Theta_{21}\Theta_{11}^{-1}\Theta_{12} = \mathrm{Cov}\left[X_2 | X_1\right]$.

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- Look at a single step of Block Cholesky decomposition:
- This corresponds to:

$$\begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{pmatrix} = \begin{pmatrix} \mathrm{Id} & 0 \\ \Theta_{21}\Theta_{11}^{-1} & \mathrm{Id} \end{pmatrix} \begin{pmatrix} \Theta_{11} & 0 \\ 0 & \Theta_{22} - \Theta_{21}\Theta_{11}^{-1}\Theta_{12} \end{pmatrix} \begin{pmatrix} \mathrm{Id} & \Theta_{11}^{-1}\Theta_{12} \\ 0 & \mathrm{Id} \end{pmatrix}$$

- Note, that for $\left(\Theta_{21}\Theta_{11}^{-1}\right) b = \mathbb{E}\left[X_2 | X_1 = b\right]$, and $\Theta_{22} - \Theta_{21}\Theta_{11}^{-1}\Theta_{12} = \mathrm{Cov}\left[X_2 | X_1\right]$.
- (Block-)Cholesky decomposition is computationally equivalent to the disintegration of Gaussian measures.

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

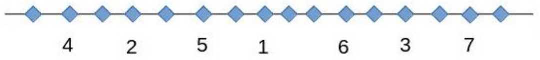
- Look at a single step of Block Cholesky decomposition:
- This corresponds to:

$$\begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{pmatrix}$$
$$= \begin{pmatrix} \mathrm{Id} & 0 \\ \Theta_{21}\Theta_{11}^{-1} & \mathrm{Id} \end{pmatrix} \begin{pmatrix} \Theta_{11} & 0 \\ 0 & \Theta_{22} - \Theta_{21}\Theta_{11}^{-1}\Theta_{12} \end{pmatrix} \begin{pmatrix} \mathrm{Id} & \Theta_{11}^{-1}\Theta_{12} \\ 0 & \mathrm{Id} \end{pmatrix}$$

- Note, that for $\left(\Theta_{21}\Theta_{11}^{-1}\right) b = \mathbb{E}\left[X_2 | X_1 = b\right]$, and $\Theta_{22} - \Theta_{21}\Theta_{11}^{-1}\Theta_{12} = \mathrm{Cov}\left[X_2 | X_1\right]$.
- (Block-)Cholesky decomposition is computationally equivalent to the disintegration of Gaussian measures.
- Follows immediately from well known formulas, but rarely used in the literature. One Example: Bickson (2008).

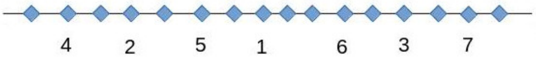
Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- This suggests to choose a *bisective* elimination ordering:



Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

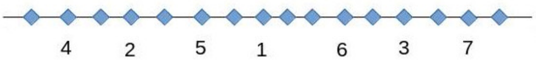
- This suggests to choose a *bisective* elimination ordering:



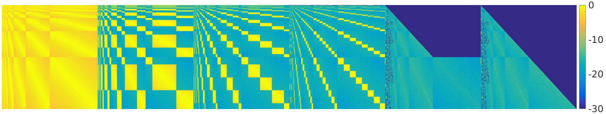
- Lets start compting the Cholesky decomposition

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- This suggests to choose a *bisective* elimination ordering:



- Lets start compting the Cholesky decomposition
- We observe a *fade-out* of entries!

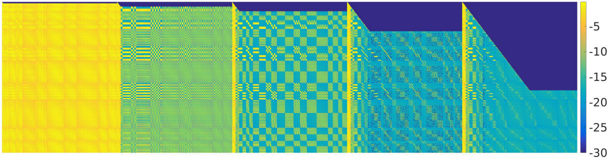


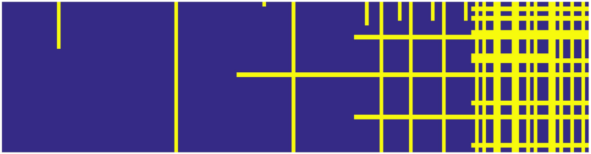
Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- What about higher dimensional examples?

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- How about higher dimensional examples?
- In 2d, use quadsection:





Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- We know that the result of the factorisation is sparse, but can we compute it efficiently?

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- We know that the result of the factorisation is sparse, but can we compute it efficiently?
- Key observation: The entry (i, j) is used for the first time with the $\min(i, j)$-th pivot.

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- We know that the result of the factorisation is sparse, but can we compute it efficiently?
- Key observation: The entry (i, j) is used for the first time with the $\min(i, j)$-th pivot.
- If we know they will be negligible untill we use them, we don't have to update them, nor know them in the first place.

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- We know that the result of the factorisation is sparse, but can we compute it efficiently?
- Key observation: The entry (i, j) is used for the first time with the $\min(i, j)$-th pivot.
- If we know they will be negligible untill we use them, we don't have to update them, nor know them in the first place.

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- Bisective/Quadsective ordering is the reverse of nested dissection.
- Indeed, for P the order-reversing permutation matrix, we have:

$$(\Theta)^{-1} = \left(LL^T\right)^{-1} = L^{-T}L^{-1}$$
$$\implies P(\Theta)^{-1}P = PL^{-T}PPL^{-1}P = \left(PL^{-T}P\right)\left(PL^{-T}P\right)^T,$$

- But we have $L^{-1} = L^T(\Theta)^{-1}$.
- For a sparse elimination ordering of Θ, the reverse ordering leads to sparse factorisation of $(\Theta)^{-1}$

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- We obtain a very simple algorithm:

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- We obtain a very simple algorithm:
- Given a positive definite matrix Θ and a Graph G, such that Θ^{-1} is sparse according to G.

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- We obtain a very simple algorithm:
- Given a positive definite matrix Θ and a Graph G, such that Θ^{-1} is sparse according to G.
- Obtain inverse nested dissection ordering for G.

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- We obtain a very simple algorithm:
- Given a positive definite matrix Θ and a Graph G, such that Θ^{-1} is sparse according to G.
- Obtain inverse nested dissection ordering for G.
- Set entries (i, j) that are separated after pivot number $min(i, j)$ to zero.

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- We obtain a very simple algorithm:
- Given a positive definite matrix Θ and a Graph G, such that Θ^{-1} is sparse according to G.
- Obtain inverse nested dissection ordering for G.
- Set entries (i, j) that are separated after pivot number $min(i, j)$ to zero.
- Compute incomplete Cholesky factorisation.

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- Remaining problems with our approach:

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- Remaining problems with our approach:
- Nested dissection does not lead to near-linear complexity algorithms

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- Remaining problems with our approach:
- Nested dissection does not lead to near-linear complexity algorithms
- Precision matrix will not be exactly sparse. How is it localised?

Sparse Factorisation of Dense Matrices: *fade-out* instead of *fill-in*

- Remaining problems with our approach:
- Nested dissection does not lead to near-linear complexity algorithms
- Precision matrix will not be exactly sparse. How is it localised?
- The answer can be found in the recent literature on numerical homogenisation:

Sparse factorisation of dense matrices using gamblets

- “Gamblet” bases have been introduced as part of the game theoretical approach to numerical PDE (Owhadi (2017), Owhadi and Scovel (2017)).

Sparse factorisation of dense matrices using gamblets

- "Gamblet" bases have been introduced as part of the game theoretical approach to numerical PDE (Owhadi (2017), Owhadi and Scovel (2017)).
- Assume our covariance matrix is

$$\Theta_{i,j} = \int_{[0,1]^2} \phi_i^{(q)}(x)\, G(x,y)\, \phi_j^{(q)}(y) \,\mathrm{d}x\,\mathrm{d}y$$

For $\phi_i^{(q)} := \mathbb{1}_{[(i-1)h^q, ih^q]}$ and G the Green's function of a second order elliptic PDE.

Sparse factorisation of dense matrices using gamblets

- "Gamblet" bases have been introduced as part of the game theoretical approach to numerical PDE (Owhadi (2017), Owhadi and Scovel (2017)).
- Assume our covariance matrix is

$$\Theta_{i,j} = \int_{[0,1]^2} \phi_i^{(q)}(x)\, G(x,y)\, \phi_j^{(q)}(y)\, \mathrm{d}x\, \mathrm{d}y$$

  For $\phi_i^{(q)} := \mathbb{1}_{[(i-1)h^q, ih^q]}$ and G the Green's function of a second order elliptic PDE.

- Corresponds to $X_i(\omega) = \int_0^1 \phi_i^{(q)}(x)\, u(x,\omega)\, \mathrm{d}x$, with $u(\omega)$ solution to elliptic SPDE with Gaussian forcing.

- Similiar to our case, only with $\mathbb{1}_{[(i-1)h^q, ih^q]}$ instead of dirac mesure.

Sparse factorisation of dense matrices using gamblets

- Similiar to our case, only with $\mathbb{1}_{[(i-1)h^q, ih^q]}$ instead of dirac mesure.
- For $\phi_i^{(k)} := \mathbb{1}_{[(i-1)h^k, ih^k]}$, Owhadi and Scovel (2017) shows:

- Similiar to our case, only with $\mathbb{1}_{[(i-1)h^q, ih^q]}$ instead of dirac mesure.
- For $\phi_i^{(k)} := \mathbb{1}_{\left[(i-1)h^k, ih^k\right]}$, Owhadi and Scovel (2017) shows:
- $\psi_i^{(k)} := \mathbb{E}\left[u \middle| \int_0^1 u(x) \phi_j^{(k)}(x) \,\mathrm{d}x = \delta_{i,j}\right]$ is exponentially localised, on a scale of h^k:

Sparse factorisation of dense matrices using gamblets

- Similiar to our case, only with $\mathbb{1}_{[(i-1)h^q, ih^q]}$ instead of dirac mesure.
- For $\phi_i^{(k)} := \mathbb{1}_{\left[(i-1)h^k, ih^k\right]}$, Owhadi and Scovel (2017) shows:
- $\psi_i^{(k)} := \mathbb{E}\left[u \middle| \int_0^1 u(x)\,\phi_j^{(k)}(x)\,\mathrm{d}x = \delta_{i,j}\right]$ is exponentially localised, on a scale of h^k:
- Main idea: Estimate on exponential decay of a conditional expectation implies exponential decay of a Cholesky factors.

Sparse factorisation of dense matrices using gamblets

- Transform to multiresolution basis to obtain block matrix:

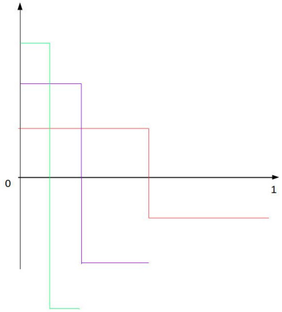
$$\left(\Gamma_{k,l}\right)_{i,j} = \int\limits_{[0,1]^2} \phi_i^{(k),\chi}(x)\, G(x,y)\, \phi_j^{(l),\chi}(y)\, \mathrm{d}x\, \mathrm{d}y$$

Sparse factorisation of dense matrices using gamblets

- Transform to multiresolution basis to obtain block matrix:

$$\left(\Gamma_{k,l}\right)_{i,j} = \int_{[0,1]^2} \phi_i^{(k),\chi}(x)\, G(x,y)\, \phi_j^{(l),\chi}(y)\, \mathrm{d}x\, \mathrm{d}y$$

- Where the $\left\{\phi_j^{(k),\chi}\right\}_{j \in J^{(k)}}$ are chosen as Haar basis functions.



- Then the results of Owhadi (2017) and Owhadi and Scovel (2017) imply that:

Sparse factorisation of dense matrices using gamblets

- Then the results of Owhadi (2017) and Owhadi and Scovel (2017) imply that:
- $\chi_i^{(k)} := \mathbb{E}\left[u \middle| \int_0^1 u(x)\, \phi_j^{(l),\chi}(x)\, \mathrm{d}x = \delta_{i,j}\delta_{k,l}, \forall l \leq k\right]$ is exponentially localised, on a scale of h^k:

$$\left|\chi_i^{(k)}\left(x - x_i^{(k)}\right)\right| \leq C \exp\left(-\frac{\gamma}{h^k}\left\|x - x_i^{(k)}\right\|\right).$$

Sparse factorisation of dense matrices using gamblets

- Then the results of Owhadi (2017) and Owhadi and Scovel (2017) imply that:
- $\chi_i^{(k)} := \mathbb{E}\left[u \middle| \int_0^1 u(x)\,\phi_j^{(l),\chi}(x)\,\mathrm{d}x = \delta_{i,j}\delta_{k,l}, \forall l \leq k\right]$ is exponentially localised, on a scale of h^k:

$$\left|\chi_i^{(k)}\left(x - x_i^{(k)}\right)\right| \leq C \exp\left(-\frac{\gamma}{h^k}\left\|x - x_i^{(k)}\right\|\right).$$

- Furthermore, the stiffness matrices decay exponentially on each level:

$$B_{i,j}^{(k)} := \int_0^1 \chi_i^{(k)}(x)\, G^{-1} \chi_j^{(k)}(x)\,\mathrm{d}x \leq \exp\left(-\gamma \left\|x_i - x_j\right\|\right)$$

Sparse factorisation of dense matrices using gamblets

- Then the results of Owhadi (2017) and Owhadi and Scovel (2017) imply that:
- $\chi_i^{(k)} := \mathbb{E}\left[u \middle| \int_0^1 u(x)\,\phi_j^{(l),\chi}(x)\,\mathrm{d}x = \delta_{i,j}\delta_{k,l}, \forall l \leq k\right]$ is exponentially localised, on a scale of h^k:

$$\left|\chi_i^{(k)}\left(x - x_i^{(k)}\right)\right| \leq C \exp\left(-\frac{\gamma}{h^k}\left\|x - x_i^{(k)}\right\|\right).$$

- Furthermore, the stiffness matrices decay exponentially on each level:

$$B_{i,j}^{(k)} := \int_0^1 \chi_i^{(k)}(x)\, G^{-1} \chi_j^{(k)}(x)\,\mathrm{d}x \leq \exp\left(-\gamma \left\|x_i - x_j\right\|\right)$$

- Finally, we have for a constant κ:

$$\operatorname{cond}\left(B^{(k)}\right) \leq \kappa, \ \forall k$$

Sparse factorisation of dense matrices using gamblets

- The above properties will allow us to show localisation of the (block) Cholesky factors:

- The above properties will allow us to show localisation of the (block) Cholesky factors:
- Consider the two-scale case:

$$\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} = \begin{pmatrix} \mathrm{Id} & 0 \\ \Gamma_{21}\Gamma_{11}^{-1} & \mathrm{Id} \end{pmatrix} \begin{pmatrix} \Theta_{11} & 0 \\ 0 & \Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12} \end{pmatrix} \begin{pmatrix} \mathrm{Id} & \Gamma_{11}^{-1}\Gamma_{12} \\ 0 & \mathrm{Id} \end{pmatrix}$$

Sparse factorisation of dense matrices using gamblets

- The above properties will allow us to show localisation of the (block) Cholesky factors:
- Consider the two-scale case:

$$\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}$$

$$= \begin{pmatrix} \mathrm{Id} & 0 \\ \Gamma_{21}\Gamma_{11}^{-1} & \mathrm{Id} \end{pmatrix} \begin{pmatrix} \Gamma_{11} & 0 \\ 0 & \Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12} \end{pmatrix} \begin{pmatrix} \mathrm{Id} & \Gamma_{11}^{-1}\Gamma_{12} \\ 0 & \mathrm{Id} \end{pmatrix}$$

$$\left(\Gamma_{21}\Gamma_{11}^{-1}\right)_{i,j} = \mathbb{E}\left[\int u\phi_i^{(2),\chi}\,\mathrm{d}x \middle| \int u\phi_m^{(1),\chi}\,\mathrm{d}x = \delta_{j,m}\right] = \int \phi_i^{(2),\chi}\chi_j^{(1)}\,\mathrm{d}x$$

$$\Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12} = \mathrm{Cov}\left[\int u\phi^{(2),\chi}\,\mathrm{d}x \middle| \int u\phi^{(1),\chi}\,\mathrm{d}x\right] = \left(B^{(2)}\right)^{-1}$$

Sparse factorisation of dense matrices using gamblets

- $\left(\Gamma_{21}\Gamma_{11}^{-1}\right)_{i,j} = \int \phi_i^{(2),\chi}\chi_j^{(1)}\,\mathrm{d}x \leq C\exp\left(-\frac{\gamma}{h}\left\|x_i^{(2)} - x_j^{(1)}\right\|\right)$

Sparse factorisation of dense matrices using gamblets

- $\left(\Gamma_{21}\Gamma_{11}^{-1}\right)_{i,j} = \int \phi_i^{(2),\chi} \chi_j^{(1)} \, \mathrm{d}x \leq C \exp\left(-\frac{\gamma}{h}\left\|x_i^{(2)} - x_j^{(1)}\right\|\right)$
- Fact: Inverses (Demko (1984), Jaffard (1990)) and Cholesky factors (Benzi and Tůma (2000), Krishtal et al. (2015)) of well-conditioned and banded/exponentially localised matrices are exponentially localised.

Sparse factorisation of dense matrices using gamblets

- $\left(\Gamma_{21}\Gamma_{11}^{-1}\right)_{i,j} = \int \phi_i^{(2),\chi} \chi_j^{(1)} \,\mathrm{d}x \leq C \exp\left(-\frac{\gamma}{h}\left\|x_i^{(2)} - x_j^{(1)}\right\|\right)$
- Fact: Inverses (Demko (1984), Jaffard (1990)) and Cholesky factors (Benzi and Tůma (2000), Krishtal et al. (2015)) of well-conditioned and banded/exponentially localised matrices are exponentially localised.
- Therefore: $\left(\left(B^{(2)}\right)^{-1}\right)_{i,j} \leq C \exp\left(-\frac{\gamma}{h^2}\left\|x_i^2 - x_j^{(2)}\right\|\right)$.

Sparse factorisation of dense matrices using gamblets

- $\left(\Gamma_{21}\Gamma_{11}^{-1}\right)_{i,j} = \int \phi_i^{(2),\chi} \chi_j^{(1)} \,\mathrm{d}x \leq C \exp\left(-\frac{\gamma}{h}\left\|x_i^{(2)} - x_j^{(1)}\right\|\right)$
- Fact: Inverses (Demko (1984), Jaffard (1990)) and Cholesky factors (Benzi and Tůma (2000), Krishtal et al. (2015)) of well-conditioned and banded/exponentially localised matrices are exponentially localised.
- Therefore: $\left(\left(B^{(2)}\right)^{-1}\right)_{i,j} \leq C \exp\left(-\frac{\gamma}{h^2}\left\|x_i^2 - x_j^{(2)}\right\|\right)$.
- Argument can can be extended to multiple scales. Results in exponentially decaying (block-)Cholesky factors.

Sparse factorisation of dense matrices using gamblets

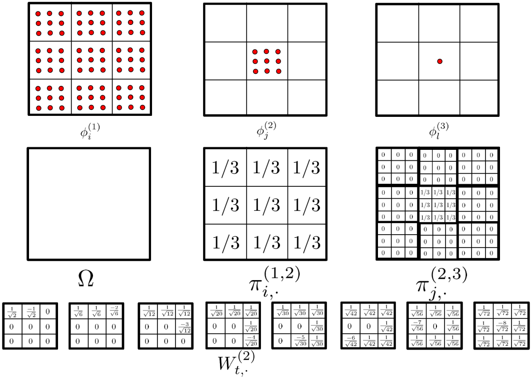
- $\left(\Gamma_{21}\Gamma_{11}^{-1}\right)_{i,j} = \int \phi_i^{(2),\chi}\chi_j^{(1)}\,\mathrm{d}x \leq C\exp\left(-\frac{\gamma}{h}\left\|x_i^{(2)} - x_j^{(1)}\right\|\right)$
- Fact: Inverses (Demko (1984), Jaffard (1990)) and Cholesky factors (Benzi and Tůma (2000), Krishtal et al. (2015)) of well-conditioned and banded/exponentially localised matrices are exponentially localised.
- Therefore: $\left(\left(B^{(2)}\right)^{-1}\right)_{i,j} \leq C\exp\left(-\frac{\gamma}{h^2}\left\|x_i^2 - x_j^{(2)}\right\|\right)$.
- Argument can can be extended to multiple scales. Results in exponentially decaying (block-)Cholesky factors.
- These factors can be approximated in time complexity by (block-)Cholesky decomposition in computational complexity of $\mathcal{O}\left(N\log^2(N)\left(\log(1/\epsilon) + \log^2(N)\right)^{4d+1}\right)$ in time and $\mathcal{O}(N\log(N)\log^d(N\frac{1}{\epsilon}))$ in space for an approximation error of ϵ.

Sparse factorisation of dense matrices using gamblets

- How about $\phi_i^{(q)} = \delta_{x_i^{(q)}}$, i.e. pointwise sampling?

Sparse factorisation of dense matrices using gamblets

- How about $\phi_i^{(q)} = \delta_{x_i^{(q)}}$, i.e. pointwise sampling?
- In Owhadi and Scovel (2017), analogue results for pointwise samples are obtained using averaging:



- We are left with a simple algorithm:

Sparse factorisation of dense matrices using gamblets

- We are left with a simple algorithm:
- Let Γ be Θ expressed in multiresolution basis.

Sparse factorisation of dense matrices using gamblets

- We are left with a simple algorithm:
- Let Γ be Θ expressed in multiresolution basis.
- Throw away all entries outside of S_ρ, defined as

$$S_\rho := \left\{ (i,j) \in I \times I \middle| i \in J^{(k)}, j \in J^{(l)}, \operatorname{dist}\left(x_i^{(k)}, x_j^{(l)} \right) \leq \rho * h^{\min(k,l)} \right\}.$$

Sparse factorisation of dense matrices using gamblets

- We are left with a simple algorithm:
- Let Γ be Θ expressed in multiresolution basis.
- Throw away all entries outside of S_ρ, defined as

$$S_\rho := \left\{ (i,j) \in I \times I \middle| i \in J^{(k)}, j \in J^{(l)}, \operatorname{dist}\left(x_i^{(k)}, x_j^{(l)} \right) \leq \rho * h^{\min(k,l)} \right\}.$$

- Compute incomplete (block-)Cholesky decomposition of Γ restricted to S_ρ.

Sparse factorisation of dense matrices using gamblets

- We are left with a simple algorithm:
- Let Γ be Θ expressed in multiresolution basis.
- Throw away all entries outside of S_ρ, defined as

$$S_\rho := \left\{ (i,j) \in I \times I \middle| i \in J^{(k)}, j \in J^{(l)}, \operatorname{dist}\left(x_i^{(k)}, x_j^{(l)}\right) \leq \rho * h^{\min(k,l)} \right\}.$$

- Compute incomplete (block-)Cholesky decomposition of Γ restricted to S_ρ.
- Factorisation can be done in $\mathcal{O}\left(N \operatorname{poly}\left(\rho \log\left(N\right)\right)\right)$, error decays exponentially with ρ.

Sparse factorisation of dense matrices using gamblets

- We are left with two closely related problems:

Sparse factorisation of dense matrices using gamblets

- We are left with two closely related problems:
- The multiresolution basis, in order to satisfy the conditions of the proof of bounded condition numbers given in Owhadi and Scovel (2017) needs to satisfy the vanishing moment condition:

$$\int_{\tau_i^{(k)}} p\phi_i^{(k),\chi} \,\mathrm{d}x = 0, \forall p \in \mathcal{P}_{s-1}\left(\tau_i^{(k)}\right),$$

for a $\tau_i^{(k)}$ of diameter $\approx h^k$ and $2s$ the order of the elliptic operator.

- We are left with two closely related problems:
- The multiresolution basis, in order to satisfy the conditions of the proof of bounded condition numbers given in Owhadi and Scovel (2017) needs to satisfy the vanishing moment condition:

$$\int_{\tau_i^{(k)}} p\phi_i^{(k),\chi} \,\mathrm{d}x = 0, \forall p \in \mathcal{P}_{s-1}\left(\tau_i^{(k)}\right),$$

  for a $\tau_i^{(k)}$ of diameter $\approx h^k$ and $2s$ the order of the elliptic operator.
- Therefore, the multiresolution basis depends on the operator.

Sparse factorisation of dense matrices using gamblets

- We are left with two closely related problems:
- The multiresolution basis, in order to satisfy the conditions of the proof of bounded condition numbers given in Owhadi and Scovel (2017) needs to satisfy the vanishing moment condition:

$$\int_{\tau_i^{(k)}} p \phi_i^{(k),\chi} \, \mathrm{d}x = 0, \forall p \in \mathcal{P}_{s-1}\left(\tau_i^{(k)}\right),$$

  for a $\tau_i^{(k)}$ of diameter $\approx h^k$ and $2s$ the order of the elliptic operator.
- Therefore, the multiresolution basis depends on the operator.
- Also, averaging over large regions required for coarse basis functions. Leads to $\mathcal{O}\left(N^2\right)$ complexity of basis transform.

Sparse factorisation of dense matrices using gamblets

- Can we get rid of vanishing moment condition?

Sparse factorisation of dense matrices using gamblets

- Can we get rid of vanishing moment condition?
- Conditions in Owhadi and Scovel (2017) are (roughly speaking):

$$\frac{1}{C} H^k \leq \lambda_{min} \left(\Theta|_{\Phi^{(k)}} \right)$$
$$\lambda_{\max} \left(\Theta|_{\perp \Phi^{(k-1)}} \right) \leq C H^{k-1}.$$

Sparse factorisation of dense matrices using gamblets

- Can we get rid of vanishing moment condition?
- Conditions in Owhadi and Scovel (2017) are (roughly speaking):

$$\frac{1}{C}H^k \leq \lambda_{min}\left(\Theta|_{\Phi^{(k)}}\right)$$
$$\lambda_{\max}\left(\Theta|_{\perp\Phi^{(k-1)}}\right) \leq CH^{k-1}.$$

- Moving to finer scales, the discrete space contains more and more oscillatory functions (small eigenvalues).

- Can we get rid of vanishing moment condition?
- Conditions in Owhadi and Scovel (2017) are (roughly speaking):

$$\frac{1}{C}H^k \leq \lambda_{min}\left(\Theta|_{\Phi^{(k)}}\right)$$
$$\lambda_{\max}\left(\Theta|_{\perp\Phi^{(k-1)}}\right) \leq CH^{k-1}.$$

- Moving to finer scales, the discrete space contains more and more oscillatory functions (small eigenvalues).
- But its in the orthogonal complement, of a larger space, low modes are "projected out".

Sparse factorisation of dense matrices using gamblets

- Can we get rid of vanishing moment condition?
- Conditions in Owhadi and Scovel (2017) are (roughly speaking):

$$\frac{1}{C} H^k \leq \lambda_{min} \left(\Theta|_{\Phi^{(k)}} \right)$$
$$\lambda_{\max} \left(\Theta|_{\perp \Phi^{(k-1)}} \right) \leq C H^{k-1}.$$

- Moving to finer scales, the discrete space contains more and more oscillatory functions (small eigenvalues).
- But its in the orthogonal complement, of a larger space, low modes are “projected out”.
- Balance of these effects leads to bounded condition numbers.

Sparse factorisation of dense matrices using gamblets

- Gamblets are more robust!

Sparse factorisation of dense matrices using gamblets

- Gamblets are more robust!
- Can replace the conditions with (roughly speaking):

$$\frac{1}{C} H^k \leq \lambda_{min} \left(\Theta|_{\Phi^{(k)}} \right)$$

$$\max_{\phi \in \Phi^k, \|\phi\| = 1} \min_{\varphi \in \Phi^{k-1} : \|\varphi\| \leq C} (\phi - \varphi)^T \Theta (\phi - \varphi) \leq C H^{k-1}.$$

Sparse factorisation of dense matrices using gamblets

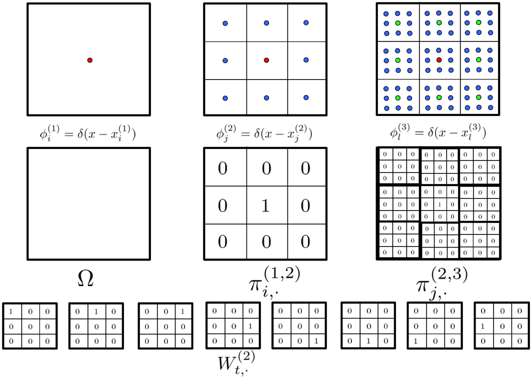
- Gamblets are more robust!
- Can replace the conditions with (roughly speaking):

$$\frac{1}{C} H^k \leq \lambda_{min} \left(\Theta|_{\Phi^{(k)}} \right)$$

$$\max_{\phi \in \Phi^k, \|\phi\| = 1} \min_{\varphi \in \Phi^{k-1} : \|\varphi\| \leq C} (\phi - \varphi)^T \Theta (\phi - \varphi) \leq C H^{k-1}.$$

- The gamblets find the optimal orthogonalisation themselves!

- We can use subsampling as an aggregation scheme!



Sparse factorisation of dense matrices using gamblets

- Our algorithm now consists of three steps:
  1. Reorder the variables hierarchically
  2. Obtain the entries in S_2 (or more generally S_ρ), set other entries to zero.
  3. Compute the incomplete Cholesky decomposition

Sparse factorisation of dense matrices using gamblets

- Our algorithm now consists of three steps:
  1. Reorder the variables hierarchically
  2. Obtain the entries in S_2 (or more generally S_ρ), set other entries to zero.
  3. Compute the incomplete Cholesky decomposition
- At this point, for theoretical guarantuees we need to replace step three with an incomplete Block factorisation. All numerical evidence indicates that this is not necessary.

Two additional results

- As observed in Owhadi 2017, Hou and Zhang 2017, gamblets provide a near-optimal sparse PCA. We obtain a PCA with the same approximation property, by keeping only the first k columns of L.

Two additional results

- As observed in Owhadi 2017, Hou and Zhang 2017, gamblets provide a near-optimal sparse PCA. We obtain a PCA with the same approximation property, by keeping only the first k columns of L.
- By reversing the elimination ordering, we obtain a near linear complexity Cholesky factorisation of the sparse/exponentially decaying inverse of Θ.

Problems at the boundary

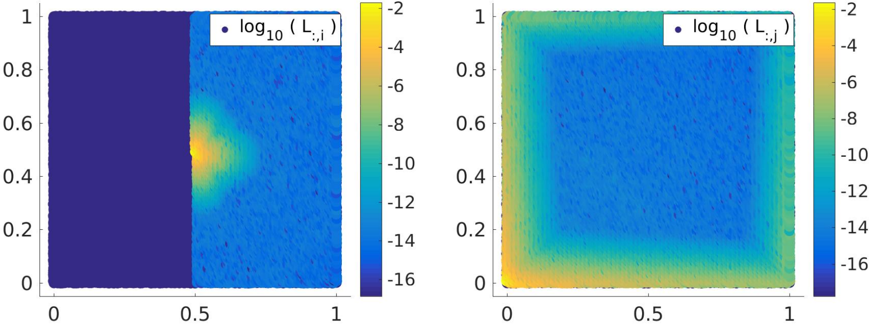
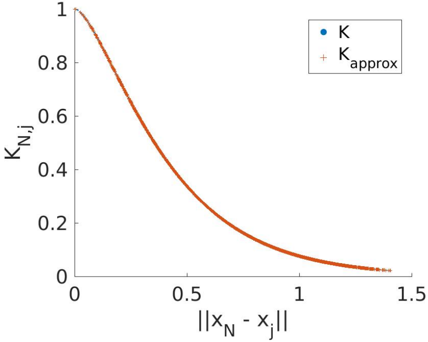
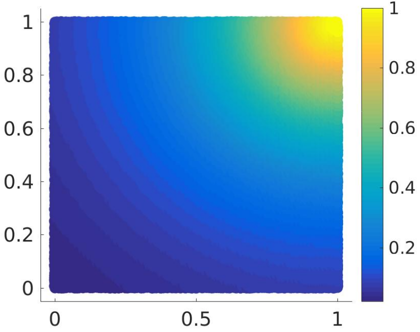


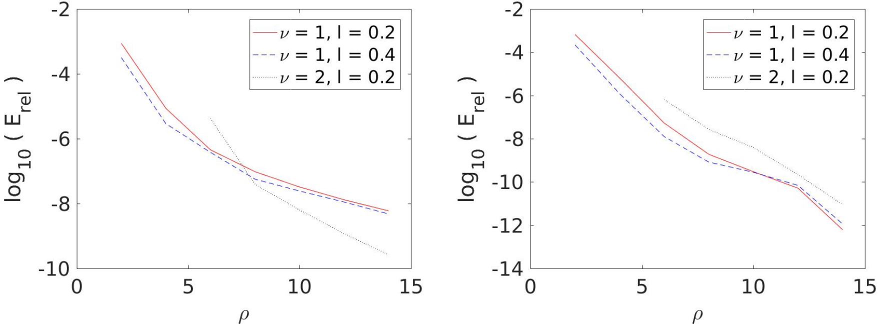
Figure: $\nu = 1$, $l = 0.4$

Problems at the boundary





Decay of approximation error



Sparse approximate PCA

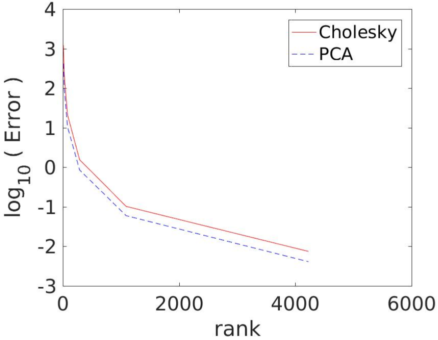
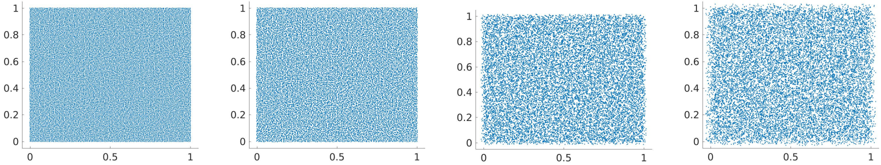


Figure: Near optimal sparse PCA: First panel: $\nu = 1$, $l = 0.2$, $\delta_x = 0.2$ and $\rho = 6$. Second panel: $\nu = 2$, $l = 0.2$ and $\delta_x = 0.2$ and $\rho = 8$.

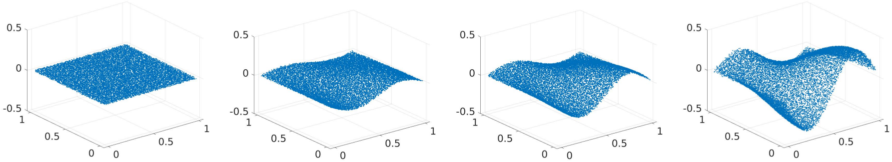
Perturbation of the Mesh



δ_x	$\\|\Gamma^\rho - \Gamma\\|$	$\\|\Gamma^\rho - \Gamma\\| / \\|\Gamma\\|$	$\\|\Gamma^\rho - \Gamma\\|_{\mathrm{Fro}}$	$\\|\Gamma^\rho - \Gamma\\|_{\mathrm{Fro}} / \\|\Gamma\\|_{\mathrm{Fro}}$	$\#S$	$\#S/N^2$
0.2	4.336e-03	1.560e-06	1.669e-02	1.026e-06	2.125e+07	7.675e-02
0.4	4.495e-03	1.617e-06	1.706e-02	1.063e-06	2.128e+07	7.683e-02
2.0	4.551e-03	1.638e-06	1.820e-02	1.077e-06	2.127e+07	7.682e-02
4.0	8.158e-03	2.940e-06	2.976e-02	1.933e-06	2.119e+07	7.652e-02

Table: Compression and accuracy for $q = 7,\ l = 0.2,\ \rho = 5,\ \nu = 1$ and different values of δ_x.

Data on low dimensional manifold



δ_z	$\|\Gamma^\rho - \Gamma\|$	$\|\Gamma^\rho - \Gamma\| / \|\Gamma\|$	$\|\Gamma^\rho - \Gamma\|_{\text{Fro}}$	$\|\Gamma^\rho - \Gamma\|_{\text{Fro}} / \|\Gamma\|_{\text{Fro}}$	$\#S$	$\#S/N^2$
0.0	5.049e-03	1.560e-06	1.885e-02	1.026e-06	2.126e+07	7.677e-02
0.1	6.341e-02	1.648e-06	1.232e-01	1.077e-06	2.083e+07	7.521e-02
0.2	1.204e-01	1.749e-06	2.203e-01	1.126e-06	1.976e+07	7.137e-02
0.4	1.954e-01	3.550e-06	5.098e-01	2.197e-06	1.722e+07	6.218e-02

Table: Compression and accuracy for $q = 7,\ l = 0.2,\ \rho = 5,\ \nu = 1,\ \delta_x = 2$ and different values of δ_z.

Fractional Operators

ν	$\|\Gamma^{\rho} - \Gamma\|$	$\|\Gamma^{\rho} - \Gamma\| / \|\Gamma\|$	$\|\Gamma^{\rho} - \Gamma\|_{\text{Fro}}$	$\|\Gamma^{\rho} - \Gamma\|_{\text{Fro}} / \|\Gamma\|_{\text{Fro}}$	$\#S$	$\#S/N^2$
1.0	1.266e-03	4.556e-07	4.987e-03	2.995e-07	2.776e+07	1.003e-01
1.1	1.813e-03	6.423e-07	6.216e-03	4.190e-07	2.776e+07	1.003e-01
1.3	3.235e-03	1.129e-06	1.039e-02	7.312e-07	2.776e+07	1.003e-01
1.5	5.245e-03	1.811e-06	1.652e-02	1.166e-06	2.776e+07	1.003e-01
1.6	6.800e-03	2.333e-06	2.148e-02	1.498e-06	2.776e+07	1.003e-01
1.8	9.891e-03	3.362e-06	3.088e-02	2.147e-06	2.776e+07	1.003e-01
2.0	1.238e-02	4.180e-06	3.892e-02	2.662e-06	2.776e+07	1.003e-01

Table: Compression and accuracy for $q = 7,\ l = 0.2,\ \rho = 6,\ \delta_x = 0.2$ and different values of ν.