Low Variance Sketched Finite Elements for Elliptic Equations

Nick Polydorides, Robert Lung

School of Engineering
University of Edinburgh

Conference on Fast Direct Solvers
Table of Contents

1 Introduction
   - Projection
   - Sketching based on leverage scores

2 Control variates
   - Correcting the sketch
   - Correcting the logarithm of the sketch

3 Numerical results
Motivation

- **Paradigm**: We consider the elliptic boundary value problem

\[
\nabla \cdot p \nabla u = f \quad \text{in } \Omega,
\]
\[
\alpha u + \beta p \nabla u \cdot \hat{n} = g \quad \text{on } \partial \Omega,
\]

on a simply connected domain \( \Omega \subset \mathbb{R}^d, \ d = \{2, 3\} \) with smooth boundary \( \partial \Omega \) where the unit normal is \( \hat{n} \) and \( \alpha, \beta, f \) and \( g \) are chosen such that \( u \) is unique.

- **Applications**: Engineering simulation, uncertainty propagation and statistical inverse problems.

- **Focus**: Computing a numerical approximation of \( u(p) \) for many parameter fields \( p \) (diagonal tensors).
An example in electrostatics: Neumann problem

Left: a discrete profile of $p$ on a disk with 9k nodes and 28k elements. Right: a numerical solution $u(p)$ with $f = \alpha = 0$, $\beta = 1$ and $\int_{\partial \Omega} g ds = 0$ conditions. 3D grids can have $> 10^6$ nodes.
Galerkin finite elements

- In Galerkin FEM with linear basis the BVP yields a linear system
  \[ Au = b, \]
  with
  \[ A := (P^{\frac{1}{2}} D)^T (P^{\frac{1}{2}} D) \]
  where \( P \in \mathbb{R}^{N \times N} \) is a positive diagonal, and \( D \in \mathbb{R}^{N \times n} \) a tall sparse matrix with \( i \)-th row \( D(i) \) and \( N > n \).

- The elements of \( P \) are the discretised model parameters of the PDE.

- \( A \) is \( n \times n \) real, sparse, symmetric, positive definite.

- We consider \( n \) to be very large.
Projected (again) FEM equations: POD

- Given $P$ we seek to approximate the high-dimensional solution $u_{opt}$ of

$$Au = b,$$

with $u_{reg} \in S$ that solves the projected equation

$$\Pi Au = \Pi b$$

where $\Pi : \mathbb{R}^n \to S$ is the projection onto

$$S := \{\Psi r | r \in \mathbb{R}^s\}$$

and $\Psi^T\Psi = I$ and $s \ll n$.

Assumptions:

- Choice of basis: $u_{opt} \approx \Pi u_{opt} = \Psi \Psi^T u_{opt}$,

- Existence of $u_{reg}$: $I - \Pi(I - A)$ is invertible $\iff$ $A$ is invertible for $\Psi$ ON.
Projected FEM equations

- Substituting $u_{\text{reg}} = \Psi r_{\text{reg}}$ into the projected equation yields an $s \times s$ system

$$G r = \Psi^T b,$$

where

$$G := \Psi^T A \Psi = \Psi^T (P^{1/2} D)^T (P^{1/2} D) \Psi = (P^{1/2} X)^T (P^{1/2} X)$$

and $X \in \mathbb{R}^{N \times s}$ tall having $i$-th row $X_{(i)} := D_{(i)} \Psi$ and rank$(X) = s$.

- The special case $P = I$ corresponds to the homogeneous PDE and a projected system

$$Q r = \Psi^T b,$$

and note that $G$ and $Q$ are similar

$$G = \sum_{i=1}^{N} p_i Q_{i}, \quad \text{while} \quad Q := \sum_{i=1}^{N} Q_{i}, \quad \text{with} \quad Q_{i} := X_{(i)}^T X_{(i)}.$$
Sketching the projected equations

- The plan is to estimate \( \hat{G} = (SP^{\frac{1}{2}}X)^T(SP^{\frac{1}{2}}X) \) from \( c \ll N \) iid samples \( \{i_1, \ldots, i_c\} \in \{1, \ldots, N\} \) using a suitable sketching matrix \( S \), then
  \[
  \hat{G} \hat{r} = \psi^T b \quad \longrightarrow \quad \hat{u}_{\text{reg}} = \psi \hat{G}^{-1} \psi^T b
  \]

- The sketch \( \hat{G} \) must be invertible with very high probability:
  \[
  \| \hat{G}^{-1} G - I \| \to \min
  \]

- The sketch \( \hat{G} \) should have low-variance, better than MC.

- Sketching linear equations involving the Laplacian matrix of a graph. (Drineas & Mahoney, 2010)
Sketching invertible matrices

- Consider first $Q = X^T X$ with $u_{\text{reg}} = \Psi Q^{-1} \Psi^T b$, $\hat{u}_{\text{reg}} = \Psi \hat{Q}^{-1} \Psi^T b$ and $X = U_X \Sigma_X V_X^T$. The sketching error is bounded by

$$\|u_{\text{reg}} - \hat{u}_{\text{reg}}\| \leq \|\hat{Q}^{-1} Q - I\| = \|\Sigma_X^{-1} (U_X S^T S U_X)^{-1} \Sigma_X - I\|,$$

conditioned on $\hat{Q} = (S X)^T S X$ being invertible.

- How do we choose $S$?

- We argue $S$ must be such that $U_X^T S^T S U_X \approx I$ in spectral norm, which for $\|U_X^T S^T S U_X - I\| < \epsilon < 1$ guarantees

$$1 - \epsilon \leq \frac{\|U_X^T S^T S U_X - I\|}{\|(U_X^T S^T S U_X)^{-1} - I\|} \leq 1 + \epsilon.$$
Leverage score sampling without replacement

- $\hat{Q}^{-1} \to \|\hat{u}_{\text{reg}} - u_{\text{reg}}\|$ bounded $\to U_X^T S^T S U_X \approx I$ in spectral norm $\to$ design sketch $S$.

- Let $\ell_i(X) = \|U_{X(i)}\|^2$ be the leverage score of $X_{(i)}$ and $\xi$ a distribution with element

$$\xi_i = \ell_i(X)/s > 0, \quad i = 1, \ldots, N,$$

then sampling each row of $X$ independently with probability

$$\eta_i = \min\{1, c'\xi_i\}$$

where $c'$ is an upper bound on the sample size, then by (Tropp, 2015)

$$\mathbb{P}(\|U_X^T S^T S U_X - I\| \geq \epsilon) \leq 2s \exp\left(-\frac{3c'\epsilon^2}{6s + 2s\epsilon}\right), \quad \forall \epsilon > 0.$$
Approximate leverage scores

- Sampling based on $\ell(X)$ yields virtually always an invertible $\hat{Q}$. We are however interested in $\hat{G} = (SP^{1/2}X)^T(SP^{1/2}X)$ not $\hat{Q} = (SX)^T SX$.

- The desirable invertibility is preserved even when the rows of $X$ are re-weighted by positive scalars through $P^{1/2}$.

- **Proposition:** Let $S$ be a sketching sparse diagonal matrix with rows

$$S(i) = \frac{\gamma_i}{\sqrt{\eta_i}} e_i^T, \quad i = 1, \ldots, N,$$

where $e_i$ the $i$-th column of $I$, and $\gamma_i$ is a Bernoulli variable with $P(\gamma_i = 1) = \eta_i$ then

$$P(\hat{G}^{-1} \text{ exists}) = P(\hat{Q}^{-1} \text{ exists}) \geq 1 - 2s \exp\left(-\frac{3c}{8s}\right).$$
Approximate leverage scores - invertibility guarantees

- **Key idea:** To sketch $G$ based on the leverage scores of $X$ which can be pre-computed offline.

- We can show that $\hat{G} \succ 0$ when $\hat{Q} \succ 0$ by exploiting the commutative property of diagonal matrices

  \[ \hat{Q} \succ 0 \iff U_X^T S^T S U_X \succ 0 \]

- With $P \succ 0$ and $\text{rank}(X) = s \implies U_X^T S^T P S U_X \succ 0$ since

  \[ \hat{G} = X^T P_\frac{1}{2} S^T S P_\frac{1}{2} X = X^T S^T P S X = V_X \Sigma_X (U_X^T S^T P S U_X) \Sigma_X V_X^T \]

- Rescaling the rows of $X$ by some positive values $P_\frac{1}{2}$ preserves the invertibility iff $U_X^T S^T S U_X \succ 0$. 

Nick Polydorides (UoEdinburgh)
Controlling complexity

- To get \( c \approx s \log s + m \) samples we sample without replacement using \( \eta_i = \min\{1, c' \xi_i\} \) where \( c' \) is an upper bound on samples.

- For a given \( c' \) the invertibility probability bound depends on the ratio \( c/s \), where \( c \) is the actual number of samples.

- For a target error \( \epsilon \) in \( \mathbb{P}(\|U_X^T S^T S U_X - I\| \geq \epsilon) \) the choice of \( c' \) should be made independently of the high dimension \( N \) and around \( \mathcal{O}(\epsilon^{-2} s \log s) \).

- Alternatively we may fix the expected number of sample \( c_e = \sum_{i=1}^{N} \eta_i \) and compute the corresponding \( c' \) by finding the root of the monotonic

\[
c' = \arg \left\{ c_e - \sum_{j=1}^{N} \min\{1, c' \xi_j\} \right\} = 0.
\]
Remarks on leverages

- Sampling $\mathcal{O}(s \log s) \ll N$ rows of $(P^{1/2}X)$ the probability of invertibility failure is infinitesimally small.

- These remarks are consistent to the results in (Cohen et al., 2015) describing the change in leverage scores & matrix coherence after re-weighting a single row.

- Invertibility breaks down if the elements of $P^{1/2}$ vary wildly. This causes $A = (P^{1/2}D)^T(P^{1/2}D)$ to be ill-conditioned, $u_{opt}$ unstable.

- Using the leverage scores suited for $Q$ to sketch $G$, invertibility is preserved at the cost of higher variance.

- Estimating the leverage scores on-the-fly when solving over-determined LS problems, e.g. (Drineas et al., 2012).
Table of Contents

1 Introduction
   - Projection
   - Sketching based on leverage scores

2 Control variates
   - Correcting the sketch
   - Correcting the logarithm of the sketch

3 Numerical results
The elements of $\hat{G} = (SP^{\frac{1}{2}}X)^T(SP^{\frac{1}{2}}X)$ and $\hat{Q} = (SX)^T(SX)$ are positively correlated.

Variance is similarly distributed between $\hat{G}_{ij}$ and $\hat{Q}_{ij}$.

Since $Q$ does not depend on $P$ we can compute it a priori, and subsequently sketch it along with $G$.

Compute a new estimator with lower variance after applying an element-wise correction to the sketched $\hat{G}$ as

$$\tilde{G} = \hat{G} - W \circ (Q - \hat{Q}),$$

where $\circ$ denotes Shur product, and $W$ is $s \times s$ symmetric

$$W_{ij} := \text{arg min} \text{ Var}(\tilde{G}_{ij}) = \frac{\text{Cov}(\hat{G}_{ij}, \hat{Q}_{ij})}{\text{Var}(\hat{Q}_{ij})}.$$
Control variates

- Considering the control variates estimator

\[ \tilde{G} = \hat{G} - W \circ (Q - \hat{Q}), \]

notice that although \( \hat{G} \succ 0 \) with very high-probability, \( \tilde{G} \) is indefinite and thus \( \tilde{G}^{-1} \) may not exist.

- To preserve invertibility and reduce variance we may correct the matrix logarithm of \( \hat{G} \) instead

\[ \tilde{\log G} = \log \hat{G} - W \circ (\log Q - \log \hat{Q}). \]

- **Rational:** Compute an estimator whose expectation is \( \log G \) and then take its matrix exponential to get a positive definite estimator of \( G \).
Logarithmic control variates

- The log control variates estimator
  \[
  \log \hat{G} = \log \hat{G} - W \circ (\log Q - \log \hat{Q}).
  \]
  has two important shortcomings:
  
  - Bias(\(\log \hat{G}\)) \(\neq 0\), and it is not computationally tractable.
  
  - The variances and covariances needed for \(W_{ij}\) are only available for sample batches, i.e. \(\log Q_i = \log (X^T(i)X(i))\) is not well defined.

- To rectify this we propose to work with a finite expansion of the Neumann series for the matrix log,

  \[
  \log(M) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} (M - I)^k \approx (M - I) - \frac{1}{2} (M - I)^2 := \mathcal{F}(M)
  \]
Preconditioning

- To ensure that the transform $\mathcal{F}$ converges to the log fast we sketch instead
  
  $$
  \mathcal{F}(C^T_0 QC_0), \quad \text{and} \quad \mathcal{F}(C^T GC),
  $$

  for some choices of invertible preconditioners $C_0, C \in \mathbb{R}^{s \times s}$ such that

  $$
  C^T_0 QC_0 \approx I \quad \text{and} \quad C^T GC \approx I.
  $$

- This yields an estimator

  $$
  \hat{\log}(C^T GC) = \left( \mathcal{F}(C^T \hat{G} C) - B_1 \right) - W \circ \left( \mathcal{F}(C^T_0 \hat{Q} C_0) - B_2 \right)
  $$

  for some bias correction matrices $B_1$ and $B_2$ and thus arriving at the sought

  $$
  \hat{G}^{-1} = C \exp(\hat{\log}(C^T GC)) C^T
  $$
A two-sample estimator

- The optimal choice of preconditioners $C_0$ and $C$ requires knowledge of $Q$ and $G$.

- $Q$ is known a priori but $G$ is not as it depends on $P$.

- A way around this is to utilise two independent samples based on the same Bernoulli probabilities.

- Use the first sample to obtain a sketched approximation of $G$ in order to get $C$ and $C_0$ (involves one SVD of an $s \times s$ matrix).

- Use the second sample to estimate $F(C_0^T \hat{Q} C_0)$, $F(C^T \hat{G} C)$ and compute weights

$$W_{ij} = \frac{\text{Cov}(F(C^T \hat{G} C)_{ij}, F(C_0^T \hat{Q} C_0)_{ij})}{\text{Var}(F(C_0^T \hat{Q} C_0)_{ij})}$$
Further implementation details

- The choice of projection basis $\Psi$ (in $X = D\Psi$) requires solving a large-scale eigenvalue problem off-line, or using a snapshots-derived ON basis.

- The low-dimensional bias correction matrices $B_1(\eta, X, P)$ and $B_2(\eta, X)$ are needed. $B_2$ can be computed off-line but $B_1$ must be approximated.

- Sketching $C_0 Q C_0$ and $C^T G C$ is equivalent to sampling the rows of two tall matrices with ON columns. This is not the case in sampling directly $Q$ and $G$. 

Nick Polydorides (UoEdinburgh)
Table of Contents

1 Introduction
   • Projection
   • Sketching based on leverage scores

2 Control variates
   • Correcting the sketch
   • Correcting the logarithm of the sketch

3 Numerical results
Tests: 2D toy problem

Two dimensional circular grid with $n = 8830$ and $N = 52224$.

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
s & c/N & \|\hat{u}_{\text{reg}} - u_{\text{reg}}\|/\|u_{\text{reg}}\| & \|\tilde{u}_{\text{reg}} - u_{\text{reg}}\|/\|u_{\text{reg}}\| & \|\hat{u}_{\text{reg}} - u_{\text{opt}}\|/\|u_{\text{opt}}\| & \|\tilde{u}_{\text{reg}} - u_{\text{opt}}\|/\|u_{\text{opt}}\| \\
\hline
100 & 0.125 & 0.0503 & 0.0040 & 0.0546 & 0.0218 \\
500 & 0.166 & 0.0675 & 0.0037 & 0.0675 & 0.0046 \\
\hline
\end{array}
\]

where

\[
\hat{u}_{\text{reg}} = \hat{G}^{-1}\psi^Tb, \quad \tilde{u}_{\text{reg}} = \tilde{G}^{-1}\psi^Tb, \quad u_{\text{reg}} = G^{-1}\psi^Tb, \quad u_{\text{opt}} = A^{-1}b
\]

- Error figures are based on averages of 100 solves for the same $b$. The 100 $P$ profiles where sampled from a mixture of Gaussians.
- Note the errors in the last two columns are inclusive of the subspace approximation error.
2D sketched solution and error

Left: a sketched solution and right: the log profile of the relative error. Solution is with $s = 500, c/N = 0.166$. 
Tests: 3D problem

Three dimensional spherical mesh with \( n = 315743 \) and \( N = 5066607 \).

| \( s \) | \( c/N \) | \( ||\hat{u}_{\text{reg}} - u_{\text{reg}}|| / ||u_{\text{reg}}|| \) | \( ||\tilde{u}_{\text{reg}} - u_{\text{reg}}|| / ||u_{\text{reg}}|| \) | \( ||\hat{u}_{\text{reg}} - u_{\text{opt}}|| / ||u_{\text{opt}}|| \) | \( ||\tilde{u}_{\text{reg}} - u_{\text{opt}}|| / ||u_{\text{opt}}|| \) |
|---|---|---|---|---|---|
| 50 | 0.020 | 0.0193 | 0.0024 | 0.0629 | 0.0595 |
| 150 | 0.020 | 0.0249 | 0.0036 | 0.0383 | 0.0298 |
| 150 | 0.100 | 0.0102 | 0.0015 | 0.0313 | 0.0297 |

where

\[
\hat{u}_{\text{reg}} = \hat{G}^{-1}\Psi^Tb, \quad \tilde{u}_{\text{reg}} = \tilde{G}^{-1}\Psi^Tb, \quad u_{\text{reg}} = G^{-1}\Psi^Tb, \quad u_{\text{opt}} = A^{-1}b
\]

- Averages of 100 solves with same right hand side \( b \). The 100 \( P \) profiles where sampled from a lognormal random field with a smooth Whittle-Matérn covariance function.
- Note the errors in the last two columns are inclusive of the subspace approximation error.
Conclusions

- Our approach decouples invertibility and accuracy of the sketched projected matrix estimator.

- Empirical results show the CV estimator suppresses sketching error by an order of magnitude.

- Low variance pays off when the subspace approximation error is small.

- Is it more efficient than estimating quickly the leverage scores?

- Further accuracy improvements via few iterations of a ‘smoother’ Jacobi iterative method.
References


