UQOP 2019
Uncertainty Quantification & Optimization Conference
18-20 March, Paris, France

Recycling Krylov subspace strategies to solve stochastic elliptic equations

Nicolas Venkovic, Paul Mycek

CERFACS, Toulouse, France
venkovic@cerfacs.fr, mycek@cerfacs.fr

Luc Giraud

INRIA-Bordeaux, France
luc.giraud@inria.fr

Olivier Le Maître

LIMSI-CNRS, Université Paris-Saclay, France
olm@limsi.fr

Keywords: Stochastic PDEs, Iterative solvers, Deflation, Monte Carlo, MCMC.

ABSTRACT

This work addresses the iterative resolution of linear systems which arise upon sampling discretized stochastic PDEs of the form

$$\nabla \cdot [\kappa(\mathbf{x}, \theta) \nabla u(\mathbf{x}, \theta)] = -f(\mathbf{x}), \ \mathbf{x} \in \Omega, \ \theta \in \Theta \tag{1}$$

with random coefficient field $\kappa(\mathbf{x}, \theta)$. Following a Monte Carlo (MC) approach, each realization of $\kappa(\mathbf{x}, \theta)$ leads to a linear system whose solution is used to compute statistics of the random solution $u(\mathbf{x}, \theta)$. To reduce the error of those statistics, a large number of realizations needs to be drawn resulting in a long sequence of linear systems whose resolution dominates the overall computational cost and time of uncertainty analyses.

When solving those systems with Krylov-based iterative techniques, it is possible to extract spectral information of previous systems to accelerate the convergence of subsequent resolutions. For instance, the deflation method constrains the search of solution to iterates with residuals in the orthogonal complement of an invariant subspace. Realistically, this deflation subspace (DS) is spanned by eigenvector approximations of the sampled operator associated with the eigenvalues which most hinder convergence. As a system is resolved, a basis of the augmented Krylov subspace is stored and used to approximate relevant eigenvectors of the next sampled operator, see [1]. These two steps (deflated resolution followed by approximate spectral reduction) are repeated throughout the sequence of sampled systems. In practice, the deflation is applied to preconditioned systems where the preconditioner can be (i) the same for all systems, or (ii)

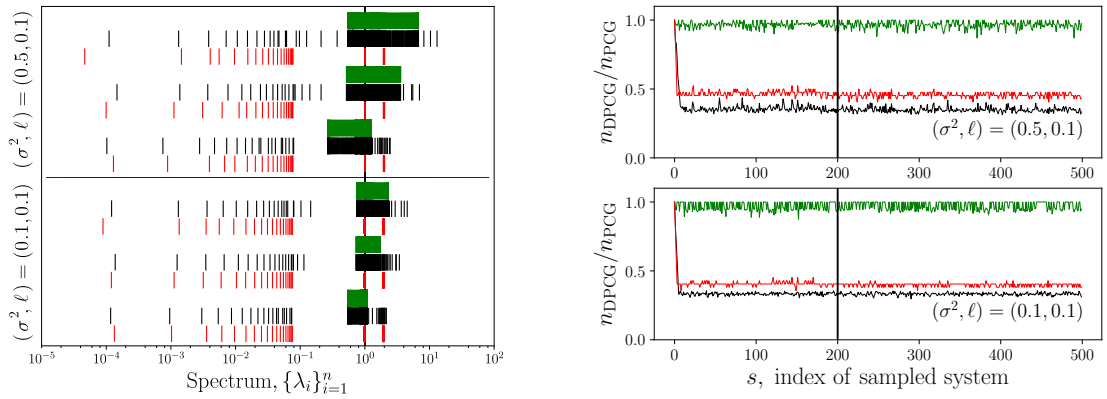

Figure 1: *Left: Spectra of preconditioned operators. Right: Relative decrease of iterations of DPCG over PCG.*

system-dependent. While the former may reduce the cost of setting up and computing potential factorizations, it will statistically not act as efficiently as a system dependent approach.

This work investigates the effect of the preconditioner and sampling strategy on the performance of the deflated method. We consider the 1D case of Eq. (1) over a domain of unit length with stationary log-normal coefficient field, such that $\log \kappa(\mathbf{x}, \theta)$ has a squared exponential covariance with variance σ^2 and correlation length ℓ. All sampled problems are discretized on the same finite element mesh, resulting in symmetric positive definite systems. Truncated Karhunen-Loève (KL) expansions of $\log \kappa$ are used to sample the coefficient field. The random variables in the KL expansion are sampled (i) by MC, or (ii) by Markov chains MC (MCMC), leading to sequences of independent or correlated realizations of κ, respectively. The rationale behind the MCMC sampling is to improve the relevance of the computed DS by using a correlated sequence of operators.

We consider two constant and one system-dependent preconditioners, namely, the median system, the block-Jacobi (BJ) of the median system, and the BJ of each system. Fig. 1 compares the spectra of preconditioned systems for 3 independent samples of κ at two variance values (left plot). Preconditioning with each system-BJ (in red) shows efficient clustering of the eigenvalues around unity with a number (related to the number of blocks) of eigenvalues isolated away from one. Median-BJ (in black) exhibits similar spectra, but with less effective clustering around one. System-BJ leads to faster PCG resolutions than median-BJ, and increasingly so as the variance is increased. Preconditioning with the median (in green) clusters efficiently all the eigenvalues around one, leads to fastest PCG resolutions, but is not a viable option for high-dimensional problems. Fig. 1 shows the acceleration measured by the ratio $n_{\mathrm{DPCG}}/n_{\mathrm{PCG}}$ of iterations needed to converge with or without deflation (right plot). For the median preconditioner (in green), DPCG yields no significant improvement; in contrast, deflating with median-BJ (in black) exhibits a sharp drop in the iteration numbers, asymptotically more than halved compared to PCG. This suggests a convergence to an almost system-independent DS which can be exploited to save computational resources, ceasing to update the DS after it has converged. For instance, reusing the DS of the 200-th system for all subsequent computations induced here no deterioration of the performance. Although not plotted here, MCMC sampling of κ results in similar accelerations as MC, so that MC should be preferred for its higher sampling efficiency.

References

[1] Y. Saad *et al* (2000). A deflated version of the conjugate gradient algorithm, *SIAM J. Sc. Comp.*.