Multigrid Preconditioners for Linear Systems Arising in PDE Constrained Optimization

ANDREI DRĂGĂNESCU, COSMIN PETRA, ANA MARIA SOANE, and JYOTI SARASWAT

UMBC
Department of MATHEMATICS & STATISTICS

Abstract model problem

Original abstract problem:

$$\begin{cases} \text{minimize } J(y,u) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + R(u,y), \\ \text{subj. to } \quad u \in U_{ad} \subset U, \;\; y \in Y_{ad} \subset Y = L^2(\Omega), \\ \qquad e(y,u) = 0 \,. \end{cases} \tag{1}$$

- U_{ad} and Y_{ad} – sets of admissible controls resp. states (convex, closed, non-empty).
- Ex.: $U_{ad} = \{u \in U \;:\; \underline{u} \le u \le \overline{u}\}$, $Y_{ad} = \{y \in Y \;:\; \underline{y} \le y \le \overline{y}\}$.
- Equality constraint is a well-posed PDE: for all $u \in U$ there is a unique $y \in Y$ (depending continuously on u), so that

$$e(y,u) = 0, \quad K(u) \stackrel{\text{def}}{=} y \,.$$

Reduced problem: If $Y_{ad} = Y$

$$\begin{cases} \text{minimize } \hat{J}(u) = \frac{1}{2}\|K(u) - y_d\|^2_{L^2(\Omega)} + \frac{\beta}{2}\|Lu\|^2, \\ \text{subj. to } \quad u \in U_{ad} \subset U, \quad L = I \text{ or } \nabla \,. \end{cases} \tag{2}$$

Background: linear PDE, no control constraints

- Assume K is a linear smoothing operator (e.g., solution operator of elliptic PDE).
- Discretization of problem (2) is equivalent to the regularized normal equations

$$G_h u \stackrel{\text{def}}{=} (\beta I + K_h{}^* \cdot K_h)u = K_h{}^* \pi_h y_d \,.$$

- Two-grid preconditioner:

$$T_h = G_{2h}\pi_{2h} + \beta(I - \pi_{2h}) \,. \tag{3}$$

Theorem 1 (Drăgănescu, Dupont 2004) *For h sufficiently small and $u \in V_h$*

$$1 - C\frac{h^p}{\beta} \le \frac{\langle (T_h)^{-1}u, u\rangle}{\langle (G_h)^{-1}u, u\rangle} \le 1 + C\frac{h^p}{\beta} \,,$$

where p is the order of the discrete method.

A. Semi-linear elliptic PDE, no control constraints

Optimal control problem:

$$\begin{cases} \text{minimize} \quad \frac{1}{2}\|y - y_d\|^2 + \frac{\beta}{2}\|u\|^2 \,, \\ \text{subj. to} \quad -\Delta y + \alpha y^3 = u \,, \;\; u \in L^2(\Omega) \,. \end{cases} \tag{4}$$

- K is twice differentiable ⇒ use Newton's method – mesh independent number of iterations:

$$u_{n+1} = u_n - \text{Hessian}^{-1}\text{gradient} \,.$$

- *Grid-sequencing* used to obtain good initial guess.
- *Adjoint methods* used to obtain gradients and the Hessian-vector multiplication:

$$\begin{aligned} \text{Linearization}: \quad & L = L(u) &&= -\Delta + 3y^2(u) \,, \\ \text{Gradient}: \quad & \nabla_u \hat{J}(u) &&= (L^*)^{-1}(y(u) - y_d) + \beta u \,, \\ \text{Hessian}: \quad & G(u) &&= (L^*)^{-1}(1 - 6K(u)Q(u))L^{-1} + \beta I \,, \\ \text{where}: \quad & Q(u) &&= (L^*)^{-1}(K(u) - y_d) \,. \end{aligned}$$

- Proposed two grid preconditioner:

$$T_h = G_{2h}(\pi_{2h}u)\pi_{2h} + \beta(I - \pi_{2h}) \,.$$

Theorem 2 (Drăgănescu, Saraswat 2011) *On a quasi-uniform mesh and under usual elliptic regularity assumptions*

$$\|(G_h(u) - T_h(u))v\| \le Ch^2\|v\|, \quad \forall v \in L^2(\Omega) \,,$$

with C independent of h.

Numerical results:

- Approximation order $O(h^2)$ confirmed by "in-vitro" experiments.
- Two-dimensional, "in-vivo" experiments: $\alpha = 1, \beta = 10^{-4}$; showing: no. T_h-PCG iterations (no. unpreconditioned CG iterations):

iterate N	16	32	64	128
1	7 (12)	6 (12)	4 (12)	4 (12)
2	7 (11)	5 (11)	4 (11)	4 (11)
3	4 (5)	3 (5)	2 (6)	1 (6)

B. Linear elliptic PDE, box-constraints on controls

Discrete optimal control problem: If $U_{ad} = \{u \in U \;:\; \underline{u} \le u \le \overline{u}\}$ in (2), solve

$$\begin{cases} \text{minimize} \quad \frac{1}{2}\|K_h u - y_{d,h}\|^2_h + \frac{\beta}{2}\|u\|^2_h \\ \text{subj. to} \quad u \in V_h, \;\; \underline{u}_h(P) \le u(P) \le \overline{u}_h(P), \;\; \forall \text{ node } P \,, \end{cases} \tag{5}$$

where discrete norms have diagonal mass matrices (mass-lumping).

Optimization algorithms (outer iteration):

- Interior point methods (IPM), semi-smooth Newton methods (active-set type strategies).
- Each requires solving one/two linear systems at each outer iteration.

B 1. Interior point methods

- At each outer iteration we have $\mathbf{U}, \mathbf{V}$ diagonal, positive; assume diag($\mathbf{U}^{-1}\mathbf{V}$) represents a relatively smooth function λ.
- Need to invert matrices of the form:

$$(\underbrace{\beta\mathbf{I} + \mathbf{U}^{-1}\mathbf{V}}_{D_{\beta+\lambda}} + \mathbf{K}^T\mathbf{K}) = (D_{\beta+\lambda} + \mathbf{K}^T\mathbf{K}) = A^{-1}(\underbrace{I + A\mathbf{K}^T\mathbf{K}A}_{G = G_h})A^{-1} \,, \qquad A = D_{\sqrt{1/(\beta+\lambda)}} \,,$$

where D_μ is the multiplication operator with the function μ (or a diagonal matrix with diagonal μ).

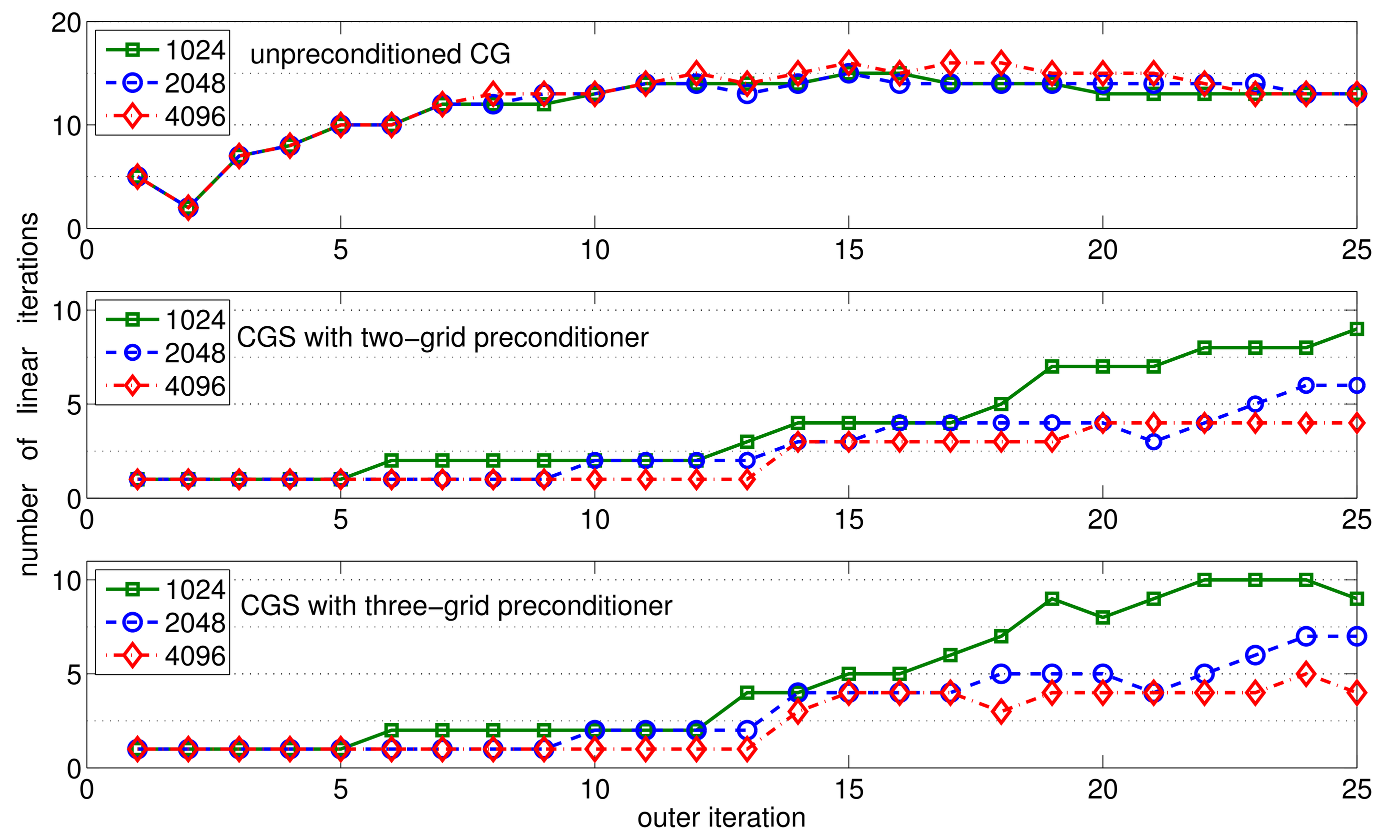
- Define preconditioner for G_h as in (3) with $\beta = 1$.

Theorem 3 (Drăgănescu, Petra 2010) *On a uniform grid, if λ_h = interpolate(λ),*

$$\rho(I - T_h{}^{-1}G_h) \le Ch^2\|(\beta + \lambda)^{-\frac{1}{2}}\|_{W^2_\infty} \,.$$

Numerical results:

- Approximation order $O(h^2)$ confirmed by "in-vitro" experiments.
- Tested with linear 2D-elliptic, 1D parabolic PDEs.
- Below: results with initial value control of parabolic PDE, y_d is the end-time state.



B 2. Semi-smooth Newton methods (SSNM)

- The SSNM produces a sequence of sets $(\mathcal{A}_k, \mathcal{I}_k)_{k=1,2,\ldots}$ that approximate the exact active/inactive sets $(\mathcal{A}, \mathcal{I})$.
- The reduced system at each SSNM iteration has the form

$$\mathbf{G}^{\mathcal{I}}\mathbf{u}_{\mathcal{I}} \stackrel{\text{def}}{=} (\beta\mathbf{I} + \mathbf{K}^T\mathbf{K})^{\mathcal{I}\mathcal{I}}\mathbf{u}_{\mathcal{I}} = \mathbf{b}_{\mathcal{I}} \,.$$

where $\mathcal{I}$ is the current guess at the inactive set.

- Similar preconditioning as in (3); challenge is to find a coarse space $V_{2h}{}^{\mathcal{I}} \subset V_h{}^{\mathcal{I}}$.

$$\mathbf{T}_h = \beta(\mathbf{I} - \pi_{2h}{}^{\mathcal{I}}) + \mathbf{G}_h{}^{\mathcal{I}}\pi_{2h}{}^{\mathcal{I}} \,.$$

Theorem 4 (Drăgănescu 2011) *On a uniform mesh*

$$\rho(I - \mathbf{T}_h{}^{-1}\mathbf{G}_h) \le C\beta^{-1}\left(h^2 + \sqrt{\mu_h{}^{\text{in}}}\right) \,, \tag{6}$$

where $\mu_h{}^{\text{in}}$ is the Lebesgue measure of the set $\partial_{\text{n}}\Omega_h{}^{\text{in}}$, denoting the numerical boundary the inactive domain relative to the coarse grid .

- Preconditioner is expected to be of suboptimal quality:

$$\rho(I - T_h{}^{-1}G_h) \le Ch^{\frac{1}{2}} \,.$$

C. Stokes control

Optimal control problem constrained by the Stokes system:

$$\begin{cases} \text{minimize } \frac{\gamma_u}{2}\|\vec{u} - \vec{u}_d\|^2 + \frac{\gamma_p}{2}\|p - p_d\|^2 + \frac{\beta}{2}\|\vec{f} - \vec{f}_0\|^2 \\ \text{subj. to} \qquad -\nu\Delta\vec{u} + \nabla p = \vec{f} \,, \\ \qquad \text{div } \vec{u} = 0 \,, \;\; \vec{u}|_\Omega = \vec{0} \end{cases} \tag{7}$$

- Hessian of reduced functional (matrix of reduced KKT system):

$$G_h = \beta I + \gamma_u U_h{}^* U_h + \gamma_p P_h{}^* P_h \,,$$

where U_h, P_h are the solution operators (velocity resp. pressure as function of force).

- The proposed two-grid preconditioner is defined as in (3).

Theorem 5 (Drăgănescu, Soane 2011) *If standard finite element approximations*

$$\|(U - U_h)(f)\| \le Ch^p\|f\|, \;\; \|(P - P_h)(f)\| \le Ch^q\|f\|$$

hold, and under standard regularity assumptions,

$$\rho(I - T_h{}^{-1}G_h) \le \frac{C}{\beta}\left(\gamma_u h^p + \gamma_p h^q\right) \,,$$

with C independent of h, β, provided the coarsest grid is sufficiently fine.