On ParaDiag for BDFs

Davide Palitta   Martin Gander

Dipartimento di Matematica, Centro AM^2
Alma Mater Studiorum - Università di Bologna

PinT 2022: 11th Conference on Parallel-in-Time Integration
CIRM, July 11-15, 2022

Differential problem and all-at-once discretization

We consider

$$\begin{cases} u_t = \mathfrak{L}(u) + f, & \text{in } \Omega \times (0, T], \\ u = g, & \text{on } \partial\Omega, \\ u(0) = u_0, \end{cases}$$

$\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, $\mathfrak{L}$ linear differential operator w/ only space derivatives

All-at-once

$$(B \otimes M + I_\ell \otimes K)\mathbf{u} = \mathbf{f}$$

- $B \in \mathbb{R}^{\ell \times \ell}$, ℓ n. of time steps
- $K, M \in \mathbb{R}^{n \times n}$ stiffness and mass matrices
- $\mathbf{u} = \text{vec}([u_1, \ldots, u_\ell]), \mathbf{f} = \text{vec}([f_1, \ldots, f_\ell]) \in \mathbb{R}^{n\ell}$

Preliminaries

- $\otimes$ Kronecker product: $A \in \mathbb{R}^{n_A \times m_A}$, $B \in \mathbb{R}^{n_B \times m_B}$,

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1m_A}B \\ \vdots & & \vdots \\ a_{n_a1}B & \cdots & a_{n_Am_A}B \end{bmatrix} \in \mathbb{R}^{n_An_B \times m_Am_B}$$

- vec operator: $X = [x_1, \ldots, x_m] \in \mathbb{R}^{n \times m}$,

$$\text{vec}(X) = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \in \mathbb{R}^{nm}$$

- $(A \otimes B)\text{vec}(X) = \text{vec}(BXA^T)$

ParaDiag: a quite straightforward idea

$$(B \otimes M + I_\ell \otimes K)\mathbf{u} = \mathbf{f}$$

1. Diagonalize $B = V\Sigma V^{-1}$, $\Sigma = \text{diag}(\sigma_1, \dots \sigma_\ell)$

$$(V \otimes I_n)(\Sigma \otimes M + I_\ell \otimes K)(V^{-1} \otimes I_n)\mathbf{u} = \mathbf{f}$$

2. If $\widetilde{\mathbf{u}} = (V^{-1} \otimes I_n)\mathbf{u}$ and $\widetilde{\mathbf{f}} = (V^{-1} \otimes I_n)\mathbf{f}$, solve

$$\underbrace{\begin{bmatrix} \sigma_1 M + K & & \\ & \ddots & \\ & & \sigma_\ell M + K \end{bmatrix}}_{\Sigma \otimes M + I_\ell \otimes K} \widetilde{\mathbf{u}} = \widetilde{\mathbf{f}}$$

3. Retrieve $\mathbf{u} = (V \otimes I_n)\widetilde{\mathbf{u}}$

Main issue

What if we can **not** compute $B = V\Sigma V^{-1}$?

- B is not diagonalizable due to the adopted time integrator
- ℓ is too large
- ...

Main issue

What if we can **not** compute $B = V\Sigma V^{-1}$?

- B is not diagonalizable due to the adopted time integrator $\leftarrow$ BDFs
- ℓ is too large
- ...

Backward Differentiation Formulas (BDFs)

For the initial value problem

$$\begin{cases} \dot{y} & = & f(t, y) \\ y(t_0) & = & y_0 \end{cases}$$

a BDF of order s is given by

$$y_n - \sum_{i=1}^{s} \alpha_i y_{n-i} = \tau \beta f(t_n, y_n)$$

τ time step size, $t_i = t_0 + i\tau$, $\alpha_i = \alpha_i(s)$, $\beta = \beta(s) \in \mathbb{R}$ (known)

- Implicit methods
- Stable with order $s \leqslant 6$

Backward Differentiation Formulas (BDFs)

A BDF of order s leads to

$$B = \frac{1}{\tau\beta}\begin{bmatrix} 1 & & & & & & \\ -\alpha_1 & \ddots & & & & & \\ \vdots & \ddots & & & & & \\ -\alpha_s & & \ddots & \ddots & & & \\ 0 & \ddots & & \ddots & & & \\ \vdots & & \ddots & & \ddots & \ddots & \\ 0 & & \cdots & -\alpha_s & \cdots & -\alpha_1 & 1 \end{bmatrix}$$

which is **not** diagonalizable!

Backward Euler & ParaDiag

Let's focus on $s = 1$: Backward Euler

$$B = \frac{1}{\tau}\begin{bmatrix} 1 & & & \\ -1 & \ddots & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}$$

Backward Euler & ParaDiag

ParaDiag state-of-the-art strategies

- Different τ_j at each t_j,

$$B = \begin{bmatrix} 1/\tau_1 & & & \\ -1/\tau_2 & 1/\tau_2 & & \\ & \ddots & \ddots & \\ & & -1/\tau_\ell & 1/\tau_\ell \end{bmatrix}$$

diagonalizable but very ill-conditioned eigenvector matrix **[Maday, Rønquist, 2008]**, **[Gander et al, 2016]**

Backward Euler & ParaDiag

ParaDiag state-of-the-art strategies

- Hybryd time discretization

$$B = \frac{1}{\tau}\begin{bmatrix} 0 & 1/2 & & & \\ -1/2 & 0 & 1/2 & & \\ & \ddots & \ddots & \ddots & \\ & & -1/2 & 0 & 1/2 \\ & & & -1 & 1 \end{bmatrix}$$

works well **[Liu et al, 2021]** but not very *flexible*: we'd like to have an effective approach for a class of time integrators!

Backward Euler & ParaDiag

ParaDiag state-of-the-art strategies

- Iterative solution of

$$(B \otimes M + I_\ell \otimes K)\mathbf{u} = \mathbf{f}$$

  and confine ParaDiag to the preconditioning step only

$$\mathcal{P} = C \otimes M + I_\ell \otimes K$$

  C diagonalizable approximation to B **[Bertaccini, 2000]**, **[Mcdonald, Pestana, Wathen, 2018]**, ...

$$B = \frac{1}{\tau}\begin{bmatrix} 1 & & & \\ -1 & \ddots & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix} = \underbrace{\frac{1}{\tau}\begin{bmatrix} 1 & & & -1 \\ -1 & \ddots & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}}_{C} + \frac{1}{\tau} e_1 e_\ell^T$$

Backward Euler & ParaDiag: our novel strategy

$M = I$ (FD) for sake of simplicitly, $C = F^{-1}\Pi F$, F FFT

$$(B \otimes I_n + I_\ell \otimes K)\mathbf{u} = \mathbf{f}$$
$$\Downarrow$$
$$(C \otimes I_n + \frac{1}{\tau} e_1 e_\ell^T \otimes I_n + I_\ell \otimes K)\mathbf{u} = \mathbf{f}$$
$$\Downarrow$$
$$(\Pi \otimes I_n + \frac{1}{\tau} F e_1 e_\ell^T F^{-1} \otimes I_n + I_\ell \otimes K)\widetilde{\mathbf{u}} = \widetilde{\mathbf{f}}$$
$$\widetilde{\mathbf{u}} := (F \otimes I)\mathbf{u}, \quad \widetilde{\mathbf{f}} = (F \otimes I)\mathbf{f}$$

Backward Euler & ParaDiag: our novel strategy

$$\left(\underbrace{\Pi \otimes I_n + I_\ell \otimes K}_{=P} + \underbrace{\left(\frac{1}{\tau} F e_1 \otimes I_n \right) \left(F^{-T} e_\ell \otimes I_n \right)^T}_{=MN^T} \right) \widetilde{\mathbf{u}} = \widetilde{\mathbf{f}}$$

Backward Euler & ParaDiag: our novel strategy

$$\left(\underbrace{\Pi \otimes I_n + I_\ell \otimes K}_{=P} + \underbrace{\left(\frac{1}{\tau} F e_1 \otimes I_n \right) \left(F^{-T} e_\ell \otimes I_n \right)^T}_{=MN^T} \right) \widetilde{\mathbf{u}} = \widetilde{\mathbf{f}}$$

Backward Euler & ParaDiag: our novel strategy

$$\left(\underbrace{\Pi \otimes I_n + I_\ell \otimes K}_{=P} + \underbrace{\left(\frac{1}{\tau}\mathbf{1} \otimes I_n\right)\left(F^{-T} e_\ell \otimes I_n\right)^T}_{=MN^T} \right) \widetilde{\mathbf{u}} = \widetilde{\mathbf{f}}$$

Backward Euler & ParaDiag: our novel strategy

$$\left(\underbrace{\Pi \otimes I_n + I_\ell \otimes K}_{=P} + \underbrace{\left(\frac{1}{\tau}\mathbf{1} \otimes I_n\right)\left(F^{-T} e_\ell \otimes I_n\right)^T}_{=MN^T} \right) \widetilde{\mathbf{u}} = \widetilde{\mathbf{f}}$$

SMW

$$\widetilde{\mathbf{u}} = P^{-1}\widetilde{\mathbf{f}} - P^{-1}M(I + N^T P^{-1} M)^{-1} N^T P^{-1}\widetilde{\mathbf{f}}$$

then

$$\mathbf{u} = (F^{-1} \otimes I)\widetilde{\mathbf{u}}$$

Backward Euler & ParaDiag: our novel strategy

$$\widetilde{\mathbf{u}} = P^{-1}\widetilde{\mathbf{f}} - P^{-1}M(I + N^T P^{-1} M)^{-1} N^T P^{-1}\widetilde{\mathbf{f}}$$

- $\mathbf{f} = \text{vec}([f_1, \dots, f_\ell])$, $\widetilde{\mathbf{f}} = (F \otimes I)\mathbf{f} = \text{vec}([f_1, \dots, f_\ell]F^T)$
- P is block diagonal, $P^{-1}\widetilde{\mathbf{f}}$ in parallel but $P^{-1}M$ is too expensive
- Exploit the Kronecker structure of $M = 1/\tau \mathbf{1} \otimes I_n$ and $N = F^{-T} e_\ell \otimes I_n$

$$N^T \text{vec}(Y) = Y F^{-T} e_\ell \quad M\text{vec}(X) = \text{vec}(1/\tau X \mathbf{1}^T)$$

- Main issue: $(I + N^T P^{-1} M)^{-1}$

Backward Euler & ParaDiag: our novel strategy

If $\Pi = \mathrm{diag}(\pi_1, \ldots, \pi_\ell)$, and $F^{-T} e_\ell = (\gamma_1, \ldots, \gamma_\ell)^T$

$$I + N^T P^{-1} M = I + \sum_{i=1}^{\ell} \gamma_i ((1 - \pi_i) I + \tau K)^{-1}$$

we need to solve

$$\underbrace{\left(I + \sum_{i=1}^{\ell} \gamma_i ((1 - \pi_i) I + \tau K)^{-1} \right)}_{=: J_\ell} x = b$$

Backward Euler & ParaDiag: our novel strategy

$$\left(I + \sum_{i=1}^{\ell} \gamma_i ((1-\pi_i)I + \tau K)^{-1}\right) x = b$$

Galerkin w/ Krylov subspace $\mathcal{K}_m(K,b) = \text{span}\{b, K, b, \ldots, K^{m-1}b\}$

- $V_m = [v_1, \ldots, v_m] \in \mathrm{R}^{n\times m}$ w/ orthonormal columns s.t. range$(V_m) = \mathcal{K}_m(K,b)$
- $x_m = V_m y_m$
- Compute $y_m \in \mathbb{R}^m$ by imposing

$$r_m \perp \mathcal{K}_m(K,b), \qquad r_m = J_\ell x_m - b$$
$$\Downarrow$$
$$\left(I_m + \textstyle\sum_{i=1}^{\ell} \gamma_i \left(I_m - t_{m+1,m} V_m^T h_i e_m^T\right) M_i\right) y_m = \theta e_1$$

$\theta = \|b\|$, $M_i = ((1-\pi_i)I + \tau T_m)^{-1}$, $T_m = V_m^T K V_m$,
$t_{m+1,m} = v_{m+1}^T K V_m$, $h_i = ((1-\pi_i)I + \tau K)^{-1} v_{m+1}$

Backward Euler & ParaDiag: our novel strategy

$$(B \otimes I_n + I_\ell \otimes K)\mathbf{u} = \mathbf{f}$$

1. Compute
$$\mathrm{vec}([z_1, \dots, z_\ell]) = (\Pi \otimes I_n + I_\ell \otimes K)^{-1}\mathrm{vec}([f_1, \dots, f_\ell]F^T)$$
in parallel
2. Set $b = [z_1, \dots, z_\ell]F^{-T}e_\ell$
3. Obtain x_m by using Galerkin to solve
$$\left(I + \sum_{i=1}^{\ell} \gamma_i((1-\pi_i)I + \tau K)^{-1}\right)x = b$$
4. Compute
$$\mathrm{vec}([w_1, \dots, w_\ell]) = (\Pi \otimes I_n + I_\ell \otimes K)^{-1}\mathrm{vec}(x_m\mathbf{1}^T)$$
in parallel
5. Set $\mathbf{u} = \mathrm{vec}(([z_1, \dots, z_\ell] - [w_1, \dots, w_\ell])F^{-T})$

Backward Euler & ParaDiag: our novel strategy

Algorithmic considerations (m number of Galerkin iterations)

- $m + 2$ parallel-in-time loops in general
- $m/d + 2$ parallel-in-time loops if we check the residual norm in Galerkin every d iterations

Numerical examples

$$
\begin{cases}
u_t - \nu\Delta u + \vec{w}\cdot\nabla u &= 0, & \text{in } \Omega\times(0,1],\ \Omega :\ = (0,1)^2 \\
u &= g(x,y), & \text{on } \partial\Omega \\
u_0 = u(x,y,0) &= g(x,y) & \text{if } (x,y)\in\partial\Omega \\
u_0 = u(x,y,0) &= 0 & \text{otherwise}
\end{cases}
$$

- $\nu = 1/20$
- $\vec{w} = (2y(1-x^2), -2x(1-y^2))$
- $g(1,y) = g(x,0) = g(x,1) = 0$, $g(0,y) = 1$
- K obtained by IFISS[i]

[i] `cd_testproblem`, n. 4 with the default setting

Numerical examples

$$(B \otimes M + I_\ell \otimes K)\mathbf{u} = \mathbf{f}, \quad B = C + \frac{1}{\tau} e_1 e_\ell^T$$

Competitor: GMRES preconditioned by $\mathcal{P} = C \otimes M + I_\ell \otimes K$ (right preconditioning), $\epsilon = 10^{-8}$ (same threshold for Galerkin)

		ParaDiag	GMRES	
n	ℓ	$m+2$	$p+1$	$\Vert\mathbf{u}_{ParaDiag} - \mathbf{u}_{GMRES}\Vert / \Vert\mathbf{u}_{GMRES}\Vert$
289	32	9	50	7.12e-9
	128	9	50	7.28e-9
	512	9	50	7.35e-9
1089	32	9	78	6.88e-9
	128	8	78	2.14e-8
	512	8	78	2.00e-8
4225	32	8	117	2.82e-8
	128	8	117	2.14e-8
	512	8	117	2.01e-8

Conclusions

So far

- New solution framework for ParaDiag w/ BDFs
- Competitive w.r.t. state-of-the-art approaches

Not done yet

- Detailed derivation of the scheme for BDFs of order $s > 1$
  - Same exact procedure but J_ℓ is now a $s \times s$ block matrix
  - Galerkin needs to take into account this structure
- Broader numerical testing (e.g., comparison w/ Daniel et al.)