A new ParaDiag time-parallel time integration method

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Davide Palitta   Martin Gander

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

Exploiting Algebraic and Geometric Structure
in Time-Integration Methods
Pisa - April 03-05, 2024

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
- $\mathbf{u} = \text{vec}([u_1, \dots, u_\ell]), \mathbf{f} = \text{vec}([f_1, \dots, f_\ell]) \in \mathbb{R}^{n\ell}$

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

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

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

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

Backward Euler & ParaDiag: our novel strategy

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

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

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

- $\mathbf{f} = \text{vec}([f_1, \ldots, f_\ell])$, $\widetilde{\mathbf{f}} = (F \otimes I)\mathbf{f} = \text{vec}([f_1, \ldots, f_\ell] F^T)$
- 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 $\Pi = \text{diag}(\pi_1, \ldots, \pi_\ell)$, 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 $\mathcal{K}_m(K, b) = \text{span}\{b, K, b, \dots, K^{m-1}b\}$

- $V_m = [v_1, \dots, v_m] \in \mathbf{R}^{n \times m}$ w/ orthonormal columns s.t. $\text{range}(V_m) = \mathcal{K}_m(K, b)$
- $x_m = V_m y_m$
- 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
$$\mathrm{vec}([z_1, \ldots, z_\ell]) = (\Pi \otimes I_n + I_\ell \otimes K)^{-1}\mathrm{vec}([f_1, \ldots, f_\ell]F^T)$$
in parallel
2. Set $b = [z_1, \ldots, z_\ell]F^{-T}e_\ell$
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
$$\mathrm{vec}([w_1, \ldots, w_\ell]) = (\Pi \otimes I_n + I_\ell \otimes K)^{-1}\mathrm{vec}(x_m\mathbf{1}^T)$$
in parallel
5. Set $\mathbf{u} = \mathrm{vec}(([z_1, \ldots, z_\ell] - [w_1, \ldots, w_\ell])F^{-T})$

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

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

This can get expensive!

α-acceleration

Given $\alpha > 0$, we can write

$$B = C_\alpha + \alpha e_1 e_\ell^T, \quad C_\alpha = \begin{bmatrix} 1 & & & -\alpha \\ -1 & \ddots & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{\ell \times \ell}$$

C_α is a α-circulant matrix and can be diagonalized by the *scaled* FFT

$$C_\alpha = D_\alpha^{-1} F^{-1} \Pi_\alpha F D_\alpha, \qquad D_\alpha = \begin{bmatrix} 1 & & & \\ & \alpha^{1/\ell} & & \\ & & \ddots & \\ & & & \alpha^{(\ell-1)/\ell} \end{bmatrix}, \ \Pi_\alpha = \alpha^{1/\ell} \Pi$$

Backward Euler & ParaDiag: α-accelerated strategy

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

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

Backward Euler & ParaDiag: α-accelerated strategy

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

Let's use a **tiny** α to make the coefficient matrix a small perturbation of the identity!

Backward Euler & ParaDiag: α-accelerated strategy

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

Let's use a **tiny** α to make the coefficient matrix a small perturbation of the identity!

Big issue: $\kappa(FD_\alpha) = \alpha^{-(\ell-1)/\ell}$

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

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

Competitors:

- GMRES preconditioned by $\mathcal{P} = C \otimes M + I_\ell \otimes K$ (right preconditioning), $\epsilon = 10^{-8}$ (same threshold for Galerkin)
- `Ev-Int`, an interpolation scheme that depends on two parameters: r (#Pint) and $\rho > 0$ **[Kressner et al. (2023)]**[ii]

[ii] It is suggested to use $r = 2$ and $\rho = 5 \cdot 10^{-4}$

Numerical examples

n	ℓ	ν	Us ($\alpha = 10^{-4}$) #PinT	Us ($\alpha = 10^{-4}$) Rel. Res.	GMRES #PinT	GMRES Rel. Res.	`Ev-Int` #PinT	`Ev-Int` Rel. Res.
		10^{-1}	3	8.41e-11	5	8.02e-13	2	7.26e-11
	32	10^{-2}	3	6.98e-12	5	7.28e-12	2	1.24e-11
		10^{-3}	3	2.77e-12	5	2.13e-10	2	7.22e-11
		10^{-1}	3	1.74e-11	5	1.12e-13	2	3.72e-11
16 384	64	10^{-2}	3	1.42e-12	5	1.31e-12	2	4.19e-12
		10^{-3}	3	7.41e-13	5	9.42e-12	2	1.99e-11
		10^{-1}	3	1.20e-11	5	4.51e-14	2	2.30e-11
	128	10^{-2}	3	1.01e-12	5	4.24e-13	2	2.23e-12
		10^{-3}	3	3.99e-13	5	1.88e-12	2	8.11e-12

Table: Advection-Diffusion equation: results for different values of ℓ, and ν.

Numerical examples

$\bar{n}$	ℓ	ν	#PinT= 2 Us ($\alpha = 10^{-4}$)	#PinT= 2 Ev-Int	#PinT= 1 Us ($\alpha = 10^{-6}$)	#PinT= 1 Ev-Int
		10^{-1}	8.44e-11	7.26e-11	1.87e-8	8.95e-6
	32	10^{-2}	6.97e-12	1.24e-11	3.97e-8	1.98e-5
		10^{-3}	3.14e-12	7.22e-11	5.88e-8	2.94e-5
		10^{-1}	1.76e-11	3.72e-11	1.09e-8	5.45e-6
16 384	64	10^{-2}	1.43e-12	4.19e-12	2.45e-8	1.22e-5
		10^{-3}	8.43e-13	1.99e-11	3.99e-8	1.99e-5
		10^{-1}	1.20e-11	2.30e-11	8.23e-9	3.55e-6
	128	10^{-2}	1.01e-12	2.23e-12	1.59e-8	7.96e-6
		10^{-3}	4.30e-13	8.11e-12	2.79e-8	1.39e-5

Table: Advection-Diffusion equation: results for different values of ℓ, and ν by fixing #PinT.

Conclusions

- New solution framework for ParaDiag w/ BDFs but it can be generalized to different time integrators (e.g., Runge-Kutta)
- Competitive w.r.t. state-of-the-art approaches

Not shown here

- Numerical study on the impact of α
- 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

Reference: *A new ParaDiag time-parallel time integration method*
M. J. Gander and D. Palitta
SIAM J. Sci. Comput., 46 (2), A697 - A718 (2024)

Wave equation - 1D

$$\begin{cases} u_{tt} &= c^2 u_{xx}, \quad \text{in } \Omega \times (0, T], \\ u &= 0, \quad \text{on } \partial\Omega, \\ u(0) &= 0, \\ u_t(0) &= \exp(-100(x-1/2)^2) \end{cases}$$

$\|U_{march} - U_{par}\|_F / \|U_{march}\|_F = \mathcal{O}(10^{-8})$ w/ 2 PinT

Wave equation - 2D

$$\begin{cases} u_{tt} = c^2 \Delta u, \quad \text{in } \Omega \times (0, T], \\ u = 0, \quad \text{on } \partial\Omega, \\ u(0) = 0, \\ u_t(0) = \exp(-100(x-1/2)^2 - 100(y-1/2)^2) \end{cases}$$

$\|U_{march} - U_{par}\|_F / \|U_{march}\|_F = \mathcal{O}(10^{-8})$ w/ 2 PinT

Wave equation w/ viscoelastic dumping - 1D

$$\begin{cases} u_{tt} &= c^2 u_{xx} + g u_{txx}, \quad \text{in } \Omega \times (0, T], \\ u &= 0, \quad \text{on } \partial\Omega, \\ u(0) &= \sin(\pi x), \\ u_t(0) &= 0 \end{cases}$$

$\|U_{march} - U_{par}\|_F / \|U_{march}\|_F = \mathcal{O}(10^{-7})$ w/ 2 PinT

Wave equation w/ viscoelastic dumping - 2D

$$\begin{cases} u_{tt} = c^2 u_{xx} + g\Delta u_t, & \text{in } \Omega \times (0, T], \\ u = 0, & \text{on } \partial\Omega, \\ u(0) = \sin(\pi(x+y)), \\ u_t(0) = 0 \end{cases}$$

$\|U_{march} - U_{par}\|_F / \|U_{march}\|_F = \mathcal{O}(10^{-7})$ w/ 2 PinT