A Parallel in Time Algorithm Based on ParaExp for Optimal Control Problems.

Djahou N. Tognon, INRIA Paris & Sorbonne Université (Paris, France)
Joint work with **Félix Kwok**, Université Laval (Québec, Canada)

63rd IEEE Conference on Decision and Control (CDC), 16-19 Dec 2024, Milan, Italy

Outline

We consider the linear quadratic optimal control problem given by

$$\min_{\nu} \mathcal{J}(\nu) := \frac{1}{2} \left\| y(T) - y_{tg} \right\|^2 + \frac{\alpha}{2} \int_0^T \left\| \nu(t) \right\|^2 dt,$$

$$\textit{subject to} \quad \dot{y}(t) = \mathcal{L} y(t) + \nu(t), \quad y(0) = y_{in}, \quad t \in (0, T].$$

- $y : [0, T] \longrightarrow \mathbb{R}^r$ the state function,
- $\nu : [0, T] \longrightarrow \mathbb{R}^r$ the control,
- y_{in} and y_{tg} the initial and final states,
- $\mathcal{L} \in \mathbb{R}^{r \times r}$ comes from a semi-discretization in space.

Introduction: Linear quadratic control problem

- Using an adjoint variable λ, the Lagrange operator becomes

$$\mathfrak{L}(\nu, y, \lambda) = \mathcal{J}(\nu) - \int_0^T (\dot{y}(t) - \mathcal{L}y(t) - \nu(t))^T \cdot \lambda(t) dt.$$

- Taking $\nabla \mathfrak{L} = 0$, we get the optimality system

$$\begin{cases} \dot{y}(t) - \mathcal{L}y(t) = \nu(t) \\ y(0) = y_{in}, \end{cases} \qquad \begin{cases} \dot{\lambda}(t) + \mathcal{L}^T \lambda(t) = 0 \\ \lambda(T) = y(T) - y_{tg}, \end{cases}$$

$$\alpha \nu(t) = -\lambda(t).$$

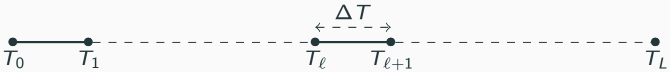
Reduced optimality system

$$\dot{y}(t) = \mathcal{L}y(t) - \frac{1}{\alpha}\lambda(t), \qquad \dot{\lambda}(t) = -\mathcal{L}^T \lambda(t), \tag{Opt-Syst}$$

with $y(0) = y_{in}$ and $\lambda(T) = y(T) - y_{tg}$.

ParaExp Algorithm [M. Gander and S. Güttel, 2013]

- For a given initial value problem: $\dot{y}(t) = \mathcal{L}y(t) + f(t), \;\; y(0) = y_0, \quad t \in \; (0, T],$
- We consider L sub-intervals of $[0, T]$ given by $(T_{\ell-1}, T_\ell) \;\; \ell = 1, \ldots, L \;\; T_\ell = \ell \Delta T, \;\; \Delta T = T/L.$



Sub-problems

Inhomogeneous sub-problems on y: For $\ell = 1, \ldots, L$,

$$\dot{w}_\ell(t) = \mathcal{L}w_\ell(t) + f(t), \;\; w_\ell(T_{\ell-1}) = 0, \;\; t \in \; (T_{\ell-1}, T_\ell],$$

Homogeneous sub-problems: $\dot{u}_1(t) = \mathcal{L}u_1(t), \;\; u_1(T_0) = y_0, \; t \in \; (T_0, T_L]$ and for $\ell = 2, \ldots, L$.

$$\dot{u}_\ell(t) = \mathcal{L}u_\ell(t), \;\; u_\ell(T_{\ell-1}) = w_{\ell-1}(T_{\ell-1}), \;\; t \in \; (T_{\ell-1}, T_L]$$

Superposition principle: For $t \in \; [T_{\ell-1}, T_\ell], \; \ell = 1, \ldots, L,$

$$y(t) = w_\ell(t) + \sum_{j=1}^{\ell} u_j(t).$$

Parallel in Time Algorithm: ParaExp idea.

- Let $Y_\ell \approx y(T_\ell)$, $\ell = 1, \ldots, L$, $\quad \Lambda_\ell \approx \lambda(T_\ell)$, $\ell = 1, \ldots, L-1$.

Sub-problems

Homogeneous sub-problems on λ: For $\ell = 1, \ldots, L-1$,

$$\dot{\lambda}_\ell(t) = -\mathcal{L}^T \lambda_\ell(t), \quad \lambda_\ell(T_L) = \Lambda_L, \quad t \in [T_{\ell-1}, T_L), \qquad (H_\lambda)$$

Inhomogeneous sub-problems on y: For $\ell = 1, \ldots, L$,

$$\dot{w}_\ell(t) = \mathcal{L} w_\ell(t) - \frac{1}{\alpha}\lambda_\ell(t), \quad w_\ell(T_{\ell-1}) = 0, \quad t \in (T_{\ell-1}, T_\ell], \qquad (IH_y)$$

Homogeneous sub-problems on y: for $\ell = 2, \ldots, L$.

$$\begin{aligned} \dot{u}_1(t) &= \mathcal{L} u_1(t), \quad u_1(T_0) = y_{in}, \quad t \in (T_0, T_L], \\ \dot{u}_\ell(t) &= \mathcal{L} u_\ell(t), \quad u_\ell(T_{\ell-1}) = w_{\ell-1}(T_{\ell-1}), \quad t \in (T_{\ell-1}, T_L], \end{aligned} \qquad (H_y)$$

Parallel in Time Algorithm: Solution propagators

Optimal trajectory

$$y(T_\ell) = w_\ell(T_\ell) + \sum_{j=1}^{\ell} u_j(T_\ell), \quad \ell = 1, \dots, L.$$

Solution operators

- $\mathcal{Q}_\ell$: exponential propagator that solves (H_λ) such that

$$\lambda_\ell(t) = \mathcal{Q}_\ell(t) \cdot \Lambda_L, \;\; t \in [T_{\ell-1}, T_L), \quad \ell = 1, \dots, L-1,$$

- $\mathcal{R}_\ell$: solution operator that solves (IH_y) such that

$$w_\ell(t) = -\frac{1}{\alpha}\mathcal{R}_\ell(t) \cdot \Lambda_L, \;\; t \in [T_{\ell-1}, T_\ell], \quad \ell = 1, \dots, L,$$

- $\mathcal{P}_\ell$: exponential propagator that solves (H_y) such that $u_1(t) = \mathcal{P}_1(t) \cdot y_{in}, \;\; t \in [T_0, T_L]$, and

$$u_\ell(t) = -\frac{1}{\alpha}\mathcal{P}_\ell(t) \cdot \mathcal{R}_{\ell-1}(T_{\ell-1}) \cdot \Lambda_L, \;\; t \in [T_{\ell-1}, T_L], \quad \ell = 2, \dots, L.$$

Parallel in Time Algorithm: Discrete optimality system

Optimality system

- Discrete optimal trajectory:

$$Y_\ell = \mathcal{P}_\ell(T_\ell) \cdot y_{in} - \frac{1}{\alpha}\mathcal{R}_\ell(T_\ell) \cdot \Lambda_L - \frac{1}{\alpha}\sum_{\ell=2}^{\ell} \mathcal{P}_j(T_\ell) \cdot \mathcal{R}_{j-1}(T_{j-1}) \cdot \Lambda_L, \quad \ell = 1, \dots, L.$$

- Final condition : $\Lambda_L - Y_L + y_{tg} = 0.$

Linear system on Λ_L

We substitute Y_L into the final condition and obtain :

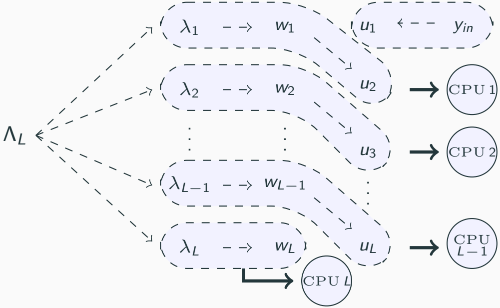
$$\mathcal{M} \cdot \Lambda_L = b,$$

$$\text{where} \quad \mathcal{M} := I + \frac{1}{\alpha}\mathcal{R}_L(T_L) + \frac{1}{\alpha}\sum_{j=2}^{L} \mathcal{P}_j(T_L) \cdot \mathcal{R}_{j-1}(T_{j-1}), \quad b = y_{tg} - \mathcal{P}_1(T_L) \cdot y_{in}.$$

Parallel in Time Algorithm: Parallel distribution

Parallel computation of $\mathcal{M} \cdot \Lambda_L$

$$\mathcal{M} \cdot \Lambda_L = \Lambda_L + \frac{1}{\alpha}\mathcal{R}_L(T_L) \cdot \Lambda_L + \frac{1}{\alpha}\sum_{j=2}^{L} \mathcal{P}_j(T_L) \cdot \mathcal{R}_{j-1}(T_{j-1}) \cdot \Lambda_L.$$



$$y_{in} \longrightarrow u_1 : \quad \mathcal{P}_1(T_L) \cdot y_{in}, \quad \Lambda_L \longrightarrow \lambda_L \longrightarrow w_L : \quad \mathcal{R}_L(T_L) \cdot \Lambda_L,$$

$$\Lambda_L \longrightarrow \lambda_\ell \longrightarrow w_\ell \longrightarrow u_{\ell+1} : \quad \mathcal{P}_{\ell+1}(T_L) \cdot \mathcal{R}_\ell(T_\ell) \cdot \Lambda_L, \quad \ell = 1, \ldots, L-1.$$

Preconditioner: 1D heat equation

- We consider 1D heat equation $\dot{y} = \Delta y + \nu$ on $[0,1] \times (0,T]$ with $y(x,0) = y_{in}(x)$ and Dirichlet boundary condition.A semi-discretization using second-order centered finite difference gives

$$\dot{y}(t) = \mathcal{L}y(t) + \nu(t), \quad y(0) = y_{in}, \quad t \in (0,T].$$

- From the continuous form of $\mathcal{M}$ given by $I + \frac{1}{\alpha}\mathcal{L}^{-1}\left(e^{2T\mathcal{L}} - I\right)$ we obtain:

Preconditioner

$$\widehat{\mathcal{M}}^{-1} = \mathcal{L}\left(\mathcal{L} - \frac{2}{\alpha}I\right)^{-1}.$$

Each application of $\widehat{\mathcal{M}}^{-1}$ only requires a multiplication by $\mathcal{L}$ and the solving of an elliptic problem of the form $(\mathcal{L} - \frac{1}{2\alpha})v = f$, which can be done cheaply using algebraic multigrid.

Theorem

Let N be given and $\mathcal{R}_\ell$ be approximated using implicit Euler with N fine sub-intervals over each $[T_{\ell-1}, T_\ell]$. Then any eigenvalue μ of $\mathcal{M}\widehat{\mathcal{M}}^{-1}$ satisfies

$$1 < \mu < 1 + \frac{\delta t}{\alpha}, \qquad \delta t = T/LN.$$

Numerical results: 1D heat equation

Test data

In our numerical test, we set $T = 1, r = 100, \alpha = 10^{-4}$. SDIRK is the Runge-Kutta method of stages $(1/2 + \sqrt{(}3)/6, 1/2 - \sqrt{(}3)/6)$ and weights $(1/2, 1/2)$. We set $N = 1000$. We use Euler and SDIRK to approximate $\mathcal{R}_\ell$ and the function `expm` in MATLAB to get $\mathcal{P}_\ell$ and $\mathcal{Q}_\ell$. GMRES `tol=1e` $-$ 8.

Efficiency of the preconditioner in GMRES

Unpreconditioned system $\mathcal{M}$

	σ_{max}	# iters	Res
Euler	4.9e2	500	1.3e-8
SDIRK	4.9e2	500	4.08e-7

Preconditioned system $\mathcal{M}\widehat{\mathcal{M}}^{-1}$

	σ_{max}	# iters	Res
Euler	1.78	9	7.7e-9
SDIRK	1.0	2	9.64e-9

Number of iterations of the preconditioned system remains bounded as $r \longrightarrow \infty$.

r	# Iters(Euler)		# Iters(SDIRK)	
	$L = 10^3$	$L = 3 \times 10^3$	$L = 10^3$	$L = 3 \times 10^3$
100	9	6	3	2
200	10	7	3	3
250	11	7	3	3
600	11	7	3	3

Preconditioner: 1D wave equation

- We now consider 1D wave equation given by $\partial_{tt} v = \Delta v + \nu,\ \text{on}\ [0,1] \times (0, T]$, with $v(x,0) = v_0(x), \partial_t v(x,0) = 0, x \in [0,1]$. A semi-discretization in space with second-order centered finite-difference leads

$$\dot{y} = \mathcal{L} y + \mathcal{B}\nu, \quad y = \begin{bmatrix} v \\ \partial_t v \end{bmatrix},\ \mathcal{L} = \begin{bmatrix} 0 & I \\ \Delta_h & 0 \end{bmatrix} \text{and } \mathcal{B} = \begin{bmatrix} 0 \\ I \end{bmatrix}.$$

- The continuous form of $\mathcal{M}$ is given by

$$I + \frac{1}{\alpha} \int_0^T \exp(s\mathcal{L}) \mathcal{B}\mathcal{B}^T \exp(s\mathcal{L}^T)\, ds = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix},$$

where for $A = -\Delta_h,\ M_{21} = M_{12}, \quad M_{12} = \frac{1}{2\alpha} A^{-1}(I - \cos(2TA^{1/2}))$

$$M_{11} = I + \frac{T}{2\alpha} A^{-1} - \frac{1}{4\alpha} A^{-3/2} \sin(2TA^{1/2}), \quad M_{22} = (1 + \frac{T}{2\alpha}) I + \frac{1}{4\alpha} A^{-1/2} \sin(2TA^{1/2}).$$

Preconditioner

$$\widehat{\mathcal{M}}^{-1} = I - \begin{bmatrix} (aI + bA)^{-1} & 0 \\ 0 & cI \end{bmatrix}, a = T, b = 2\alpha \text{ and } c = (T + 2\alpha)^{-1}.$$

Test data

In our numerical test, we set $T = 1, r = 100, \alpha = 10^{-6}$ and $N = 1000$. We use Euler and SDIRK to approximate $\mathcal{R}_\ell$ and the function `expm` in MATLAB to get $\mathcal{P}_\ell$ and $\mathcal{Q}_\ell$. GMRES `tol=1`$e - 8$.

Efficiency of the preconditioner in GMRES

Unpreconditioned system $\mathcal{M}$

	cond	# iters	Res
Euler	3.8e4	84	8.74e-9
SDIRK	3.8e4	84	8.74e-9

Preconditioned system $\mathcal{M}\widehat{\mathcal{M}}^{-1}$

	cond	# iters	Res
Euler	2.59	4	1.58e-9
SDIRK	2.59	4	1.58e-9

Number of iterations of the preconditioned system remains bounded as $r \longrightarrow \infty$

r	# iters $\mathcal{M}$	# iters $\mathcal{M}\widehat{\mathcal{M}}^{-1}$
10	10	5
150	76	3
350	104	3

Number of iterations of the preconditioned system for various α

α	# iters $\mathcal{M}$	# iters $\mathcal{M}\widehat{\mathcal{M}}^{-1}$
1e-3	20	3
1e-1	9	3
1e1	4	2

Conclusion and Ongoing works

Conclusion

- We introduced a new time parallel algorithm for time dependent linear quadratic optimal control problem when a cheap exponential integrator is available,
- We proposed two preconditioners for 1D heat equation and 1D wave equation.

Ongoing works

- We are currently studying the behavior of the preconditioners when $\mathcal{M}$ is obtained from an explicit method respecting CFL condition,
- We are also working on more general convergence properties of the algorithm, its error analysis, and on understanding its performance compared to existing parallel-in-time algorithms.

References

M. Gander, S. Güttel
Paraexp: A parallel integrator for linear initial-value problems.
SIAM, 2013.

M. Gander, F. Kwok, J. Salomon
ParaOpt: A Parareal algorithm for optimality systems.
SIAM, 2020.

Colin S.C. Tsang,
Preconditioners for linear parabolic optimal control problems. Master's thesis,
Hong Kong, China, 2017.