Asynchronous 'Parareal' in time discretization for partial differential equations

Frédéric Magoulès, Guillaume Gbikpi-Benissan

April 7, 2016

CentraleSupélec – IRT SystemX

Outline of the presentation

01 INTRODUCTION

1 - Introduction

Massively parallel platforms

- Different models of hardware (extension of existing parallel machines)
- Different levels of interconnection (processors, nodes, clusters, ...)
- $\Rightarrow$ Heterogeneous computer and network capabilities
- $\Rightarrow$ Exacerbation of efficiency problems (load balancing, interprocess dependency)
- Particularly for iterative methods: frequent global synchronization

Asynchronous iterative methods

- Self-adapting to bad load balance and communication delays
- $\Rightarrow$ Interesting for overcoming current limitations
- Main focus on space domain decomposition for fast solution of various mathematical problems

What about asynchronous iterations for time domain decomposition?
Case of the 'Parareal' algorithm

"Possibly the kind of methods which will allow the next generation of parallel machines to attain the expected potential." – *Frommer & Szyld, 2000.*

OUTLINE

02 ASYNCHRONOUS METHODS

Chaotic iteration

Relaxation methods: $Ax = b \iff x = Tx + c$

- First numerical results, with the SOR method, on an electrical network problem
  (Rosenfeld, 1967, 1969)
- Sufficient and necessary: $|T|v \leq \alpha v, \quad v > 0, \quad 0 < \alpha < 1 \quad (\iff \rho(|T|) < 1)$
  - A symmetric and strictly diagonally dominant
  - A irreducibly diagonally dominant
  - A SPD with non-positive off-diagonal entries

  (Chazan & Miranker, 1969)

General operator: $x = f(x), \quad f : D(f) \subset E \to E$

- Sufficient: $\eta(f(x) - f(y)) \leq T\eta(x - y), \quad \rho(T) < 1, \quad T$ nonnegative
  η: canonical vectorial norm (η-contracting or T-contracting operator)
  (Robert et al., 1975; Miellou, 1975)
- Extension to first “asynchronous iteration” models (unbounded delays)
  Introduction of “m-contracting” operators ($f : D(f) \subset E^m \to E$)
  (Baudet, 1978)

Asynchronous iteration

$$x_i(k+1) = \begin{cases} f_i(x_1(\tau_1(k)),\dots,x_n(\tau_n(k))), & \forall i \in P(k) \\ x_i(k), & \forall i \notin P(k) \end{cases}$$

(a) $P(k) \subseteq \{1,\dots,n\}, P(k) \neq \emptyset$
(b) $\tau_j(k) \leq k$
(c) $\mathrm{card}\{k \in \mathbf{N} | i \in P(k)\} = +\infty$
(d) $\lim_{k\to+\infty} \tau_j(k) = +\infty$

- Sufficient: $\|f(x) - f(x^*)\|_\infty^v \leq \alpha \|x - x^*\|_\infty^v, \quad v > 0, \quad 0 < \alpha < 1$
  $\|\|_\infty^v$: weighted maximum norm (contracting operator)
  Generalization of T-contraction and m-contraction frameworks
  (El Tarazi, 1982)
- "Totally asynchronous iteration": $f_i(x_1(\tau_1^i(k)),\dots,x_n(\tau_n^i(k)))$
  Sufficient conditions based on nested sets (generalization of contraction framework)
  (Bertsekas, 1983; Bertsekas & Tsitsiklis, 1989)
- Extension to two-stage iteration: $x(k+1) = f^k(x(k))$
  Contraction conditions for all f^k
  (Frommer & Szyld, 1994)

2 - Asynchronous methods

Asynchronous two-stage iteration

Implicit operator: $x(k) = f(x(k)) \quad \sim \quad x(k+1) = \tilde{f}^k(x(k))$

- "Totally asynchronous two-stage iteration": $f(x) = \tilde{f}(x, x, \ldots)$ (multiple copies of E)
  $\Rightarrow \tilde{f}_i((x_1(\tau_{11}(k)), \ldots, x_n(\tau_{n1}(k))), \ldots, (x_1(\tau_{1m_k}(k)), \ldots, x_n(\tau_{nm_k}(k))))$
  Sufficient conditions based on nested infinite product spaces
  (Frommer & Szyld, 1994)

- "Asynchronous iteration with flexible communication": $f(x) = \tilde{f}(x, x)$
  $\Rightarrow$ Two set of delays: $\tilde{f}_i((x_1(\tau_1^i(k)), \ldots, x_n(\tau_n^i(k))), (x_1(\rho_1^i(k)), \ldots, x_n(\rho_n^i(k))))$
  Sufficient (extension of El Tarazi):
  $\|\tilde{f}(x, y) - x^*\|_\infty^v \leq \alpha \max(\|x - x^*\|_\infty^v, \|y - x^*\|_\infty^v), \quad v > 0, \quad 0 \leq \alpha < 1$
  (Miellou et al., 1994, 1998; El Baz et al., 1996; Frommer & Szyld, 1998)

- Application to waveform relaxation methods
  (Martin, 1999; Frommer & Szyld, 2000)

Application to the Parareal in time discretization?

03 Asynchronous Parareal algorithm

3 - Asynchronous Parareal algorithm

Parareal in time discretization

Lions et al., 2001; Bal & Maday, 2002

$$\begin{aligned} \frac{\partial u}{\partial t} + Au &= f, \quad t \in [0,T] \\ u(t=0) & \; \texttt{given} \end{aligned} \quad \rightarrow \quad \begin{aligned} \frac{\partial u_n}{\partial t} + Au_n &= f_n, \quad t \in [T_n, T_{n+1}] \\ u_n(t=T_n) &= U_n \end{aligned}$$

$\Delta T = T/N, \quad T_n = n\Delta T \quad (0..T_1..T_2 \dots T_{N-1}..T)$
$U_0 = u(t=0)$

- Coarse solver $G_{\Delta T}(u_0, 0, T)$: 'predict' $U_1, \dots, U_N$
- Fine solver $F_{\Delta t}(U_n, T_n, T_{n+1})$: parallel-compute 'exact' $u_0(T_1), \dots, u_{N-1}(T_N = T)$, according to predictions
- Better predict by accounting gap between previous predictions and exact values

$\Rightarrow$ Parallel iterative formulation:

$$U_{n+1}^{k+1} = G_{\Delta T}(U_n^{k+1}, T_n, T_{n+1}) + F_{\Delta t}(U_n^k, T_n, T_{n+1}) - G_{\Delta T}(U_n^k, T_n, T_{n+1})$$

Based on multiple shooting and multi-grid approaches in parallel-in-time discretization
(Chartier & Philippe, 1993; Horton & Vandewalle 1995)

3 - Asynchronous Parareal algorithm

Parareal iteration

$$U_{n+1}^{k+1} = G_{\Delta T}(U_n^{k+1}, T_n, T_{n+1}) + F_{\Delta t}(U_n^k, T_n, T_{n+1}) - G_{\Delta T}(U_n^k, T_n, T_{n+1})$$

Algebraic form:

$$x_{i+1}^{k+1} = A^{(i)} x_i^{k+1} + B^{(i)} x_i^k - A^{(i)} x_i^k$$

$$(I - A)x^{k+1} = (B - A)x^k$$

(Gauss-Seidel)-type relaxation for:

$$(I - B)x = 0$$

$\Rightarrow$ Asynchronous two-stage iteration:

$$U_{n+1}^{k+1} = G_{\Delta T}(U_n^{\tau_n(k)}, T_n, T_{n+1}) + F_{\Delta t}(U_n^{\rho_n(k)}, T_n, T_{n+1}) - G_{\Delta T}(U_n^{\rho_n(k)}, T_n, T_{n+1}),$$

$$\rho_n(k) \leq \tau_n(k), \quad \rho_n(k) \leq k, \quad \tau_n(k) \leq k+1$$

Contracting integration scheme for linear problems:

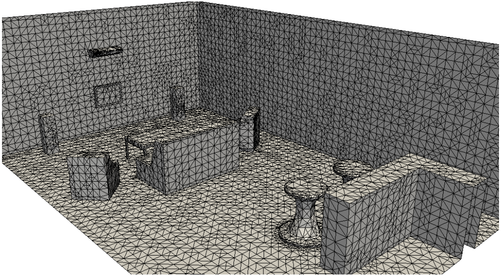
- Implicit Euler for $G_{\Delta T}$ and $F_{\Delta t}$
  (Mathew et al., 2010)
- Implicit Euler for $G_{\Delta T}$ and second-order scheme (e.g. trapezoidal rule) for $F_{\Delta t}$
  (Wu, 2015)

04 Numerical experiments

4 - Numerical experiments

Problem and settings

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta u &= f, \quad t \in [0, T_{max}] \\ u(0) &= c \end{aligned}$$



- Implicit Euler for $G_{\Delta T}$
  Trapezoidal rule for $F_{\Delta t}$
- $\Delta t = 0.0001, \quad \Delta T = 0.01$
- 33 796 nodes, 171 478 elements
- C++ implementation using an MPI-based communication library
  (Magoulès & G.-Benissan, JACK: an asynchronous communication kernel library for iterative algorithms. The Journal of Supercomputing, 2016)
- Cluster Altix ICE 8400 LX, 68 nodes (two 6-cores Intel Xeon, 2.4 GHz), InfiniBand
  MPI library: SGI-MPT 2.01

4 - Numerical experiments

Results

		Sync. Parareal			Async. Parareal		
#Proc.	T_{max}	Time (s)	#Iter.	Residual	Time (s)	#Iter.	Residual
18	0.18	1.16	10	1.74e-07	1.57	27	9.78e-08
24	0.24	1.65	14	1.85e-07	1.97	34	1.06e-07
36	0.36	3.62	23	6.45e-11	2.74	45	2.94e-11
48	0.48	4.59	32	4.94e-11	3.62	59	3.33e-11
72	0.72	5.88	50	9.01e-08	5.65	73	5.25e-10
90	0.90	10.51	64	4.75e-11	7.61	149	4.77e-11

$$U_{n+1}^{k+1} = G_{\Delta T}(U_n^{\tau_n(k)}, T_n, T_{n+1}) + F_{\Delta t}(U_n^{\rho_n(k)}, T_n, T_{n+1}) - G_{\Delta T}(U_n^{\rho_n(k)}, T_n, T_{n+1}),$$
$$\rho_n(k) \leq \tau_n(k), \quad \rho_n(k) \leq k, \quad \tau_n(k) \leq k+1$$

Asynchronous 'Parareal' in time discretization for partial differential equations

Frédéric Magoulès, Guillaume Gbikpi-Benissan

April 7, 2016

CentraleSupélec – IRT SystemX