Frontal and multi-frontal solvers: Generalization to isogeometric finite element method

Maciej Paszynski

Department of Computer Science

AGH University of Science and Technology, Krakow, Poland

maciej.paszynski@agh.edu.pl

http://home.agh.edu.pl/paszynsk

http://www.ki.agh.edu.pl/en/staff/paszynski-maciej

http://www.ki.agh.edu.pl/en/research-groups/a2s

Main collaborators

Victor Calo (KAUST) Leszek Demkowicz (ICES, UT) David Pardo (IKERBASQUE)



INTRODUCTION B-SPLINES BASED FINITE ELEMENT METHOD

Strong formulation

$$-\frac{d}{dx}\left(A(x)\frac{du}{dx}\right) + B(x)\frac{du}{dx} + C(x)u = f(x)$$
$$u(0) = 0$$
$$A(1)\frac{du(1)}{dx} + \beta u(1) = \gamma$$

Weak formulation

Find
$$u \in V = \{u \in H^{1}(0,1) : u(0) = 0\}$$
 s.t.
 $b(v,u) = l(v), \quad \forall v \in V = \{v \in H^{1}(0,1) : v(0) = 0\}$
 $b(v,u) = \int_{0}^{1} \left[A(x)\frac{dv}{dx}\frac{du}{dx} + B(x)v(x)\frac{du}{dx} + C(x)v(x)u(x)\right]dx + \beta v(1)u(1)$
 $l(v) = \gamma v(1)$

INTRODUCTION B-SPLINES BASED FINITE ELEMENT METHOD

Using B-splines as basis functions

Find
$$u \in V = \left\{ u \in H^{1}(0,1) : u(0) = 0 \right\}$$
s.t.
 $b(v,u) = l(v), \quad \forall v \in V = \left\{ v \in H^{1}(0,1) : v(0) = 0 \right\}$
 $u(x) \approx \sum_{i} N_{i,p}(x) d_{i} \quad v(x) \leftarrow N_{j,p}(x)$
 $\sum_{i} b(N_{j,p}(x), N_{i,p}(x)) a_{i} = l(N_{j,p}(x)), \quad \forall j$
 $\int_{0}^{1} \frac{1}{1-2-3-4-5-5}$
 $\int_{0}^{1} \frac{1}{1-2-3-4-5-5-6}$

INTRODUCTION B-SPLINES BASED FINITE ELEMENT METHOD

Using B-splines as basis functions

Find
$$u \in V = \{ u \in H^1(0,1) : u(0) = 0 \}$$
s.t.
 $b(v,u) = I(v), \quad \forall v \in V = \{ v \in H^1(0,1) : v(0) = 0 \}$

$$u(x) \approx \sum_{i} N_{i,p}(x) d_{i} \qquad v(x) \leftarrow N_{j,p}(x)$$
$$\sum_{i} b(N_{j,p}(x), N_{i,p}(x)) a_{i} = I(N_{j,p}(x)), \forall j$$



GRAPH GRAMMAR PRODUCTIONS AS ATOMIC TASKS



We assign indices to grammar productions in order to localize the places where the graph grammar productions were fired

The elimination tree obtained by executing the following sequence of productions

 $(P1)-(P2)_{1}-(P2)_{2}-(P2)_{3}-(P2)_{4}-(P3)_{1}-(P3)_{2}-(P3)_{3}-(P3)_{4}-(P3)_{5}-(P3)_{6}$

Alphabet:

 $A = \{(P1), (P2)_1, (P2)_2, (P2)_3, (P2)_4, (P3)_1, (P3)_2, (P3)_3, (P3)_4, (P3)_5, (P3)_6\}$







Dependency graph

TRACE THEORY BASED SCHEDULER

 $(P1)-(P2)_{1}-(P2)_{2}-(P2)_{3}-(P2)_{4}-(P3)_{1}-(P3)_{2}-(P3)_{3}-(P3)_{4}-(P3)_{5}-(P3)_{6}$

Foata Normal Form

 $\begin{bmatrix} a_1^1 a_2^1 \dots a_{l_1}^1 \end{bmatrix} \begin{bmatrix} a_1^2 a_2^2 \dots a_{l_1}^2 \end{bmatrix} \dots \begin{bmatrix} a_1^n a_2^n \dots a_{l_n}^n \end{bmatrix}$ $a_i^k \in A \text{ (alphabet)}$ $\forall k \ \forall i, j \in \{1, \dots, J_k\} \quad a_i^k I a_j^k \text{ i<>j where } I = A x A \setminus D$ $\forall k \ \forall i \in \{1, \dots, J_k\} \exists j \in \{1, \dots, J_{k-1}\} \quad a_i^{k-1} D a_j^k$

Scheduling according to Foata Normal Form:

 $[(P1)][(P2)_1(P2)_2][(P2)_3(P2)_4(P3)_5(P3)_6]][(P3)_1(P3)_2(P3)_3(P3)_4]$

Thus, the execution of the solver consists of several steps, where independent tasks are executed in concurrent, interchanged with the synchronization barriers.

GRAMMAR BASED NUMERICAL INTEGRATION

$$b(N_{j,p}(x), N_{i,p}(x)) = \int_{0}^{1} \left[A(x) \frac{dN_{j,p}(x)}{dx} \frac{dN_{i,p}(x)}{dx} + B(x)N_{j,p}(x) \frac{dN_{i,p}(x)}{dx} + C(x)N_{j,p}(x)N_{i,p}(x) \right] dx + \beta N_{j,p}(1)N_{i,p}(1)$$

$$I(N_{i,p}(x)) = \gamma N_{i,p}(1)$$

using Gaussian quadrature the integration over the domain can be substituted by a weighted summation over Gauss points

$$\int_{0}^{1} \left[A\left(x\right) \frac{dN_{i,p}\left(x\right)}{dx} \frac{dN_{j,p}\left(x\right)}{dx} + B\left(x\right) \frac{dN_{i,p}\left(x\right)}{dx} N_{j,p}\left(x\right) + C\left(x\right) N_{i,p}\left(x\right) N_{j,p}\left(x\right) \right] dx = \sum_{l} W_{l} \left[A\left(x_{l}\right) \frac{dN_{i,p}\left(x_{l}\right)}{dx} \frac{dN_{j,p}\left(x_{l}\right)}{dx} + B\left(x_{l}\right) \frac{dN_{i,p}\left(x_{l}\right)}{dx} N_{j,p}\left(x_{l}\right) + C\left(x_{l}\right) N_{i,p}\left(x_{l}\right) N_{j,p}\left(x_{l}\right) \right] dx = \sum_{l} W_{l} \left[A\left(x_{l}\right) \frac{dN_{i,p}\left(x_{l}\right)}{dx} \frac{dN_{j,p}\left(x_{l}\right)}{dx} + B\left(x_{l}\right) \frac{dN_{i,p}\left(x_{l}\right)}{dx} N_{j,p}\left(x_{l}\right) + C\left(x_{l}\right) N_{i,p}\left(x_{l}\right) N_{j,p}\left(x_{l}\right) \right] dx = \sum_{l} W_{l} \left[A\left(x_{l}\right) \frac{dN_{i,p}\left(x_{l}\right)}{dx} \frac{dN_{i,p}\left(x_{l}\right)}{dx} + B\left(x_{l}\right) \frac{dN_{i,p}\left(x_{l}\right)}{dx} N_{j,p}\left(x_{l}\right) + C\left(x_{l}\right) N_{i,p}\left(x_{l}\right) N_{j,p}\left(x_{l}\right) \right] dx$$

GRAMMAR BASED NUMERICAL INTEGRATION

Cox-de Boor formula

$$N_{i,0}(\xi) = I_{[\xi_i,\xi_{i+1}]}$$
$$N_{i,p}(\xi) = \frac{\xi - \xi_i}{x_{i+p} - \xi_i} N_{i,p-1}(\xi) + \frac{\xi_{i+p+1} - \xi}{x_{i+p+1} - \xi_{i+1}} N_{i+1,p-1}(\xi)$$



GRAMMAR BASED NUMERICAL INTEGRATION





Generation of frontal matrices at leaves of the eliminaton tree expressed as the execution of graph grammar productions $(A1)-(A)^4-(AN)$



Graph grammar productions generating local frontal matrices for left boundary, interior and right boundary nodes for linear B-splines



Graph grammar productions merging element frontal matrices at parent level



Graph grammar production eliminating fully assembled row at parent level



Graph grammar production for merging element frontal matrices at root level



Graph grammar production for solution at root level



Graph grammar production for recursive backward substitution



Expression of the solver execution by graph grammar productions

(A1)-(A)⁴-(AN) (generation of frontal matrices at leaves of the elimination trees)

(A2)³ (merging contributions at father nodes)

(E2)³ (elimination of fully assembled nodes)

(A2) – (E2) (merging at parent node followed by elimination)

(Aroot) – (Eroot) (merging at root node followed by full forward elimination)

(BS)⁴ (backward substitutions)

Alphabet:

 $\begin{array}{l} \mathsf{A} = \{(\mathsf{A1}), \, (\mathsf{A})_1 \, , \, (\mathsf{A})_2 \, , \, (\mathsf{A})_3 \, , \, (\mathsf{A})_4 \, , \, (\mathsf{AN}), \, (\mathsf{A2})_1 \, , \, (\mathsf{A2})_2 \, , \, (\mathsf{A2})_3 \, , \, (\mathsf{E2})_1 \, , \, (\mathsf{E2})_2 \, , \, (\mathsf{E2})_3 \, , \, (\mathsf{A2})_4 \, , \\ (\mathsf{E2})_4 \, , \, (\mathsf{Aroot}) \, , \, (\mathsf{Eroot}) \, , \, (\mathsf{BS})_1 \, , \, (\mathsf{BS})_2 \, , \, (\mathsf{BS})_3 \, , \, (\mathsf{BS})_4 \, \} \end{array}$

Dependency relation for the solver algorithm

 $\{(A1), (A)_1\}D(A2)_1$ $\{(A)_2, (A)_3\}D(A2)_2$ $\{(A)_4, (AN)\}D(A2)_3$ $(A2)_1D(E2)_1$ $(A2)_{2}D(E2)_{2}$ $(A2)_{3}D(E2)_{3}$ Aroot $\{(E2)_1, (E2)_2\}D(A2)_4$ root Eroot $(A2)_{4}D(E2)_{4}$ $\{(E2)_{3}(E2)_{4}\}D(Aroot)$ (A2)4 (Aroot)D(Eroot) int (E2)4 $(\text{Eroot})D\{(BS)_1, (BS)_2\}$ (BS)₁D{(BS)₃,(BS)₄} (A2)3 (E2)3 (A2)2 int int int (E2)2 (E2) (A1) (A)1 (A)3 (AN) (A) $(\mathbf{A})4$ int int int int int int node node (node) node (node) node node node node node (node) node

Dependency graph



Dependency graph



TRACE THEORY BASED SCHEDULER

 $(A1)-(A)_1-(A)_2-(A)_3-(A)_4-(AN)-(A2)_1-(A2)_2-(A2)_3-(E2)_1-(E2)_2-(E2)_3-(A2)_4-(E2)_4-(Aroot)-(Eroot)-(BS)_1-(BS)_2-(BS)_3-(BS)_4$

Foata Normal Form

 $\begin{bmatrix} a_1^1 a_2^1 \dots a_{l_1}^1 \end{bmatrix} \begin{bmatrix} a_1^2 a_2^2 \dots a_{l_1}^2 \end{bmatrix} \dots \begin{bmatrix} a_1^n a_2^n \dots a_{l_n}^n \end{bmatrix}$ $a_i^k \in A \quad \text{(alphabet)}$ $\forall k \; \forall i, j \in \{1, \dots, J_k\} \quad a_i^k \; Ia_j^k$ $\forall k \; \forall i \in \{1, \dots, J_k\} \; \exists j \in \{1, \dots, J_{k-1}\} \quad a_i^{k-1} Da_j^k$

Scheduling according to Foata Normal Form:

 $[(A1)(A)_{1}(A)_{2}(A)_{3}(A)_{4}(AN)][(A2)_{1}(A2)_{2}(A2)_{3}][(E2)_{1}(E2)_{2}(E2)_{3}][(A2)_{4}][(E2)_{4}]][(Eroot)][(Aroot)][(BS)_{1}(BS)_{2}][(BS)_{3}(BS)_{4}]]$

Thus, the execution of the solver consists of several steps, where independent tasks are executed in concurrent, interchanged with the synchronization barriers.

GRAPH GRAMMAR PRODUCTIONS EXPRESSING THE SOLVER ALGORITHM



GRAPH GRAMMAR PRODUCTIONS EXPRESSING THE SOLVER ALGORITHM



NUMERICAL EXPERIMENTS

1D NUMERICAL RESULTS LINEAR B-SPLINES



1D NUMERICAL RESULTS QUADRATIC B-SPLINES



1D NUMERICAL RESULTS CUBIC B-SPLINES

0.0025



1D NUMERICAL RESULTS QINTIC B-SPLINES



COMPARISON WITH CPU MUMPS SOLVER



NVidia Tesla c2070, 6GB memory, 448 CUDA cores, each one with 1.15GHz clock Intel(R) Core(TM)2 Quad CPU Q9400 with 2.66GHz clock, 8GB of memory

S





























2D NUMERICAL INTEGRATION



2D NUMERICAL INTEGRATION

For p=2 there are 3x3=9 two dimensional B-splines

$b(B_{k-2,l-2;1},B_{k-2,l-2;1})$		$b(B_{k,l;1}, B_{k-2,l-2;1})$			
$b(B_{k-2,l-2;1},B_{k,l;1})$		$b\left(B_{k,l;1},B_{k,l;1}\right)$			
so we need to compute 3x3=9 two dimensional B-splines					

	I						
$B_{k,l;1}(x_1, x_2) = N_{k;1}(x_1)N_{l;1}(x_2)$	$B_{k,l-1;1}(x_1,x_2) = N_{k;1-1}(x_1)N_{l;1}(x_2)$	$B_{k,l-2,1}(x_1,x_2) = N_{k,1}(x_1)N_{l-2,1}(x_2)$					
$B_{k-1,l;1}(x_1, x_2) = N_{k-1;1}(x_1)N_{l;1}(x_2)$	$B_{k-1,l-1;1}(x_1,x_2) = N_{k-1;1-1}(x_1)N_{l;1}(x_2)$	$B_{k-1,l-2;1}(x_1,x_2) = N_{k-1;1-1}(x_1)N_{l-2;1}(x_2)$					
$B_{k-2,l;1}(x_1,x_2) = N_{k-2,1}(x_1)N_{l;1}(x_2)$	$B_{k-2,l-1;1}(x_1,x_2) = N_{k-2;1-1}(x_1)N_{l;1}(x_2)$	$B_{k-2,l-2,1}(x_1,x_2) = N_{k-2,1-1}(x_1)N_{l-2,1}(x_2)$					
so we need to compute 3+3=6 one dimensional B-splines							
$N_{k,1}(x_1)$	$N_{k-1;1}(x_1)$	$N_{k-2;1}(x_1)$					
$N_{l;1}(x_2)$	$N_{l-1;1}(x_2)$	$N_{l-2;1}(x_2)$					





<mark>16</mark>	2	2	2	2	2	2	2	2	$\begin{bmatrix} u_{i,j} \end{bmatrix} \begin{bmatrix} 9 \end{bmatrix}$	
2	8	2	0	2	1	1	0	0	$u_{i-1,j} = 0$	
2	2	8	2	0	0	1	1	0	$u_{i,j+1}$ 3	
2	0	2	8	2	0	0	1	1	$u_{i+1,j}$ 6	
2	2	0	2	8	1	0	0	1	$u_{i,j-1} = 6$	
2	1	0	0	1	4	0	0	0	$u_{i-1,j-1}$ 3	
2	1	1	0	0	0	4	0	0	$u_{i-1,j+1}$ -	3
2	0	1	1	0	0	0	4	0	$u_{i+1,i+1}$ 3	
2	0	0	1	1	0	0	0	4	$u_{i+1,j-1}$ 3	







2D ELIMINATION



Tasks related to single row subtractions

2D ELIMINATION



Tasks related to single row subtractions

2D ELIMINATION



Tasks related to single row subtractions

2D NUMERICAL RESULTS LINEAR B-SPLINES



2D NUMERICAL RESULTS QUADRATIC B-SPLINES



2D NUMERICAL RESULTS CUBIC B-SPLINES



COMPARISON WITH CPU MUMPS ASSEMBLY



NVidia Tesla c2070, 6GB memory, 448 CUDA cores, each one with 1.15GHz clock Intel(R) Core(TM)2 Quad CPU Q9400 with 2.66GHz clock, 8GB of memory

COMPARISON WITH CPU MUMPS FACTORIZATION



NVidia Tesla c2070, 6GB memory, 448 CUDA cores, each one with 1.15GHz clock Intel(R) Core(TM)2 Quad CPU Q9400 with 2.66GHz clock, 8GB of memory

COMPARISON WITH CPU MUMPS TOTAL TIME



NVidia Tesla c2070, 6GB memory, 448 CUDA cores, each one with 1.15GHz clock Intel(R) Core(TM)2 Quad CPU Q9400 with 2.66GHz clock, 8GB of memory

COMPUTATIONAL COST OF SERIAL C⁰ and C^{p-1} SOLVERS

COMPUTATIONAL COST OF SERIAL C⁰ and C^{p-1} SOLVERS

List of crimes

- Consider a regular grid with same number of elements in each direction
- Number of elements assumed to be sufficiently large
- Ignore orthogonality between basis functions sharing same support
- To simplify analysis, only consider limiting cases of continuity C^0 and C^{p-1}
- Numerical results consider Laplace equation over a unit cube geometry

COMPUTATIONAL COST OF SERIAL C⁰ and C^{p-1} SOLVERS FLOPS (Time) Estimates

$$\begin{aligned} &\mathsf{FLOPS}\left(2\mathsf{D}, C^{0}\right) &= 2^{2s}p^{6} + \sum_{i=1}^{s-1} 2^{2(s-i)}2^{3i}p^{3} = \mathcal{O}(Np^{4}) + \mathcal{O}(N^{1.5}) \\ &\mathsf{FLOPS}\left(2\mathsf{D}, C^{p-1}\right) = 2^{2s}p^{4} + \sum_{i=1}^{s-1} 2^{2(s-i)}2^{3i}p^{6} = \mathcal{O}(N^{1.5}p^{3}) \\ &\mathsf{FLOPS}\left(3\mathsf{D}, C^{0}\right) &= 2^{3s}p^{9} + \sum_{i=1}^{s-1} 2^{3(s-i)}2^{6i}p^{6} = \mathcal{O}(Np^{6}) + \mathcal{O}(N^{2}) \\ &\mathsf{FLOPS}\left(3\mathsf{D}, C^{p-1}\right) = 2^{3s}p^{6} + \sum_{i=1}^{s-1} 2^{3(s-i)}2^{6i}p^{9} = \mathcal{O}(N^{2}p^{3}) \end{aligned}$$

For large problems, C^0 continuity is p^3 faster

N.Collier, D. Pardo, L. Dalcin, M. Paszynski, V.Calo; (2012) The cost of continuity: a study of performance of isogeometric finite elements using direct solvers, **Computer Methods in Applied Mechanics and Engineering**, 213-216, p. 353-361

COMPUTATIONAL COST OF SHARED MEMORY C⁰ and C^{p-1} SOLVERS

COMPUTATIONAL COST OF SHARED MEMORY C⁰ and C^{p-1} SOLVERS

List of parallel crimes

- Assume infinitely large number of processors.
- Assume zero communication cost.
- Assume infinite memory.

COMPUTATIONAL COST OF SHARED MEMORY C⁰ and C^{p-1} SOLVERS

 $\begin{aligned} & \mathsf{FLOPS}\,(\mathsf{1D}, C^0) &= p^2 + \sum_{\substack{i=1 \\ s-1}}^{s-1} 1 = & \mathcal{O}(p^2) + & \mathcal{O}(\log(N/p)), \\ & \mathsf{FLOPS}\,(\mathsf{1D}, C^{p-1}) &= p + \sum_{\substack{i=1 \\ s-1}}^{s-1} p^2 = & \mathcal{O}(p^2 \log(N/p)), \end{aligned}$ $\begin{aligned} & \mathsf{FLOPS}\,(\mathsf{2D},\,C^0) &= p^4 + \sum_{\substack{i=1\\s-1}}^{s-1} 2^{2i} p^2 = \ \mathcal{O}(p^4) + \ \mathcal{O}(N) \\ & \mathsf{FLOPS}\,(\mathsf{2D},\,C^{p-1}) \ = p^2 + \sum_{\substack{i=1\\s-1}}^{s-1} 2^{2i} p^4 = \ \mathcal{O}(Np^2) \end{aligned}$ FLOPS (3D, C^0) = $p^6 + \sum_{\substack{i=1 \ s-1}}^{r} 2^{4i} p^4 = \mathcal{O}(p^6) + \mathcal{O}(N^{1.33})$ FLOPS (3D, C^{p-1}) = $p^4 + \sum_{i=1}^{r} 2^{4i} p^6 = \mathcal{O}(N^{1.33} p^2)$

For large problems, C^0 continuity is p^2 faster

1D LINEAR B-SPLINES



execution time [s]

1D QUADRATIC B-SPLINES



1D CUBIC B-SPLINES



execution time [s]

1D QINTIC B-SPLINES



2D LINEAR B-SPLINES



2D QUADRATIC B-SPLINES



execution time [s]

2D CUBIC B-SPLINES



PAPERS

Multi-frontal solver for IGA discretizations in GPUs

K. Kuznik, M. Paszynski, V. Calo, D.Pardo (2013) Multi-frontal solver for IGA discretizations in GPUs, submitted to **Computers and Mathematics with Applications**

Graph grammar based 2D isogeometric FEM solver:

K. Kuznik, M. Paszynski, V. Calo (2012) Graph Grammar-Based Multi-Frontal Parallel Direct Solver for Two-Dimensional Isogeometric Analysis, *Procedia Computer Science*, 9, p. 1454-1463

Computational costs for 1D/2D/3D sequential isogeometric FEM:

 N. Collier, D. Pardo, L. Dalcin, M. Paszynski, V.Calo; (2012) The cost of continuity: a study of performance of isogeometric finite elements using direct solvers, *Computer Methods in Applied Mechanics and Engineering*, 213-216, p. 353-361

Graph grammar based 2D FEM solver for distributed memory linux cluster:

M. Paszyński, R. Schaefer (2010) Graph grammar-driven parallel partial differential equation solver *Concurrency & Computations, Practise & Experience* 22 (9) p.1063-1097

Graph grammar based 3D FEM solver for distributed memory linux cluster:

M. Paszyński, D. Pardo, A. Paszynska (2010) Parallel multi-frontal solver for p adaptive finite element modeling of multi-physics computational problems, *Journal of Computational Science* 1 (1) p.48-54