Simulation of soft tissue deformation in real-time using domain decomposition
Ryadh Haferssas, Pierre-Henri Tournier, Frédéric Nataf, Stéphane Cotin

To cite this version:
Ryadh Haferssas, Pierre-Henri Tournier, Frédéric Nataf, Stéphane Cotin. Simulation of soft tissue deformation in real-time using domain decomposition. 2019. hal-02400982v2

HAL Id: hal-02400982
https://hal.archives-ouvertes.fr/hal-02400982v2
Preprint submitted on 11 Dec 2019
Simulation of soft tissue deformation in real-time using domain decomposition

Ryadh Haferssas\textsuperscript{1}, Pierre-Henri Tournier\textsuperscript{2}, and Frédéric Nataf\textsuperscript{2} Stéphane Cotin\textsuperscript{1}

\textsuperscript{1} INRIA France
\{ryadh.haferssas,stephane.cotin\}@inria.fr
\textsuperscript{2} Sorbonne Université
\{frederic.nataf,tournier\}@sorbonne-universite.fr

Abstract. Non-linear models used in dynamic simulations usually require the solution of multiple large and sparse linear systems in a successive manner. In this paper, we conduct a study of numerical solvers in the framework of real-time soft tissue deformation. Domain Decomposition paradigm has the potential of providing parallelism at both levels of equation assembling and linear system solving. In our case domain decomposition is employed to solve a non-linear model in a dynamic simulation in order to meet real-time computation by using parallel architecture. Numerical test on liver deformations using a non-linear deformation model is presented to evaluate the acceleration impact of the domain decomposition paradigm. Performances tests clearly show the efficiency of using a domain decomposition approach for real-time feedback.

Keywords: soft tissue deformation · non-linear model · linear systems · domain decomposition

1 Introduction

Image-guided therapy has revolutionized medicine, in its ability to provide care that is both efficient and effective. However, images acquired during an intervention are either incomplete, under-exploited, or can induce adverse outcomes. This can be due, for instance, to the lack of dimensionality of X-ray images and the associated radiation exposure for the patient. In the same time, the scientific computing community developed a particular interest in medical models which attempt to provide numerical simulation to reproduce living anatomy or physiology of the specific patient. The main challenge is to combine numerical models with data extracted from intra-operative images and deliver efficient per-operative guidance to clinicians.

In this paper, we are mainly interested in the simulation of soft tissue, in the context described above. Whether we consider augmented reality or simply real-time elastic registration, the constraints are similar. Models that aim to mimic the mechanical behavior of complex anatomical structures must be accurate enough
to predict the location of internal structures invisible in the intra-operative image while providing visual feedback in real-time. This makes the rise of simulation constraints at several levels: mechanical modeling, equation discretization, and linear solvers. As a result, any simulation attempt would be necessarily looking for the best compromise between accuracy and computation time. At the mechanical level, there are several instances in the literature of real-time simulation of deformable object which rely on linear elasticity like in [3]. The first limitation is inherent to the small strain assumption: to keep a linear model, the magnitude of deformation must be very small. Non-linear models do not suffer from this constraint but they introduce an extra computational cost. At the time discretization level, one of the most known strategies to deal with real-time computation of the liver tissue deformations is based on an explicit integration scheme, as proposed in [14, 13]. Explicit integration methods are particularly well suited for some applications like the real-time non-rigid registration of the brain shift during surgery [9]. However, liver tissues are often very soft, but in most of the pathologic cases, these tissues are much stiffer. That makes the critical time step very small (about $10^{-6}$ sec) which would not correspond to the real dynamics. Implicit integration allows the use of larger time steps (about $2 \cdot 10^{-2}$ sec) without any stability issues. The counterpart is a heavier computation at each time step.

In practice, implicit integration schemes require to solve large systems of linear or non-linear equations. In both cases, we need to use a solver of linear systems in a successive manner to compute a solution that represents the state of the simulated model. The computation time of this part is the most important fold of the overall simulation computation time. The main interest which has driven our work in this paper is the efficiency in solving a large algebraic linear system in parallel computation. During these last few years, the computational power available within the hardware of computers is evolving in a different way. Due to frequency and heat-dissipation limits, the current trend is focused on increasing the number of computation units rather than their individual speed. Nowadays, one or two quadcore processors can be found in standard desktop computers. Two families of methods are traditionally used to solve a large linear system arising from discretization of mechanical models on a parallel machine: direct and iterative solvers. Direct solvers are known to be very robust. However, the memory requirement becomes significant with larger systems. On the other hand, iterative solvers [15] e.g., GMRES, CG, are less memory consuming and naturally parallel but they suffer from a lack of robustness. Domain Decomposition method [19, 4] as well as multigrid [8] method are hybrid methods that can take advantage of direct solvers and iterative solvers in the same algorithm. These two groups of approaches are described as hybrid methods because they are ultimately used as a preconditioner for the linear system during an iterative method, but direct methods are also used within the definition of the global preconditioner on some smaller subsystems or auxiliary problems. Such a hybridization provides highly concurrent methods that are robust enough to solve complex real-life problems [12, 6, 18].
In this paper, we aim to assess and achieve performance gain in a typical case of liver deformation simulation using a non-linear model with implicit integration scheme. To this end, we propose to adopt a domain decomposition paradigm that introduces parallelism at two different levels, first at the model assembling then at equation solves. Such a strategy opens new opportunities to deal with accurate real-time simulations of soft organs, in particular in the context of complex interactions.

2 Method

In this section, we summarize the continuum framework, introduce a constitutive model along with the boundary value problem, and its numerical discretization.

2.1 Biomechanical model

To describe the mechanical behavior of the liver, we use a total Lagrangian formalism. In general, we consider a body whose reference configuration is $\Omega_0$ at time $t_0$, subjected to a force per unit mass $f$, its boundary surface $\partial \Omega$ is divided into a Dirichlet part $\Gamma_D^0$ constrained by a displacement $y$ and a Neumann part $\Gamma_N^0$ subjected to a traction force $T$, the continuum equations stated in the strong form are

$$\begin{align*}
\rho_0 \ddot{y} - \nabla \cdot \Sigma = \rho_0 f & \quad \text{in } \Omega_0, \\
\Sigma \cdot n = T & \quad \text{on } \Gamma_N^0, \\
y = y & \quad \text{on } \Gamma_D^0.
\end{align*}$$

(1)

In these relations $\rho_0$ is the density, $\Sigma$ is the second Piola-Kirchhoff stress tensor, and $n$ is the unit surface normal in the reference configuration.

Space integration In order to discretize problem (1) by a finite element method, consider a tetrahedral mesh $\{T_h\}_{h>0}$ of the computational domain $\Omega_0$. The discretized finite element formulation results in a nonlinear system of algebraic equations

$$M \ddot{Y} + F^{int}(Y) = F^{ext}, \quad \forall t \in [0,T],$$

(2)

where initial, internal and external forces are respectively given by

$$
\begin{align*}
M \ddot{Y} &= \int_{T_h} \rho_0 N_i N_j dV \ddot{Y}, \\
F^{int}_i &= \int_{T_h} \Sigma : \nabla N_i dV, \\
F^{ext}_i &= \int_{T_h} \rho_0 f \cdot N_i dV + \int_{\Gamma_N^0} T \cdot N^S_i dS,
\end{align*}
$$

(3)

where $M$ is the mass matrix, $N_i$ is the conventional shape function corresponding to node $i \in [1,N]$ with $N$ the number of nodes. $Y \in \mathbb{R}^N$ is the vector of the current nodal positions.
Time integration: We chose a conventional implicit integration scheme provided by Newmark [1]. To this end, we consider a positive integer $N$ and define $\Delta t = T/N$, $t_n = n \Delta t$, for $n = 0, 1, ..., N$. We compute the approximation $Y^n$ by using the following second-order Newmark scheme

\begin{align*}
Y^{n+1} &= Y^n + \Delta t \dot{Y}^n + (1/4) \Delta t^2 \ddot{Y}^n, \\
\dot{Y}^{n+1} &= Y^n - (1/2) \Delta t \ddot{Y}^n, \\
M \ddot{Y}^{n+1} + F^{\text{int}}(Y^{n+1}) &= F^{\text{ext}}_{n+1}.
\end{align*}

(4)

In hyperelastic models, the internal forces are provided by a non-linear function $F^{\text{int}}$. Here, we use the Newton-Raphson method to address the non-linearity at each time step.

The fully discretized problem (4) gives rise to a linear system of the form $Au = b$ which needs to be solved for each simulation time step and more than ones in case of non-linear models. Solving such a linear system could become extremely expensive from the computational point of view. Medical simulations are constrained by the need for real-time computation to enable interactivity of the simulation, this requirement translates into solving concurrently multiple linear systems under a very challenging time constraint.

Traditionally, to solve these linear systems, two types of approaches are used: direct and iterative solvers. Direct solvers provide the solution in a fixed number of steps. It mainly involves two phases: first, the factorization phase, e.g., $LU, LDL^T$, then, the solving phase. The factorization phase is independent of the right hand side and is computationally more expensive than the solving phase. Iterative solvers, e.g., GMRES, on the other hand, do not modify the matrix and rely solely on matrix-vector products and other basic algebra operations. However, for an iterative solver to be efficient, choosing a good preconditioner [15] is imperative, but in some cases finding a good preconditioner is a difficult task.

To overcome the disadvantages of iterative solvers and to take advantage of the desirable features of direct solvers in the framework of parallel computing, there has been an increasing focus on the so-called hybrid methods such as domain decomposition and multigrid methods. For this paper, we adopt in the numerical implementation a parallel strategy based on domain decomposition method.

2.2 Domain decomposition solver

Domain decomposition methods are known to be a divide & conquer paradigm to accelerate numerical simulations. In our simulation context, we choose to use an overlapping Schwarz method. To describe it, we first divide the mesh $\{T_h\}_{h>0}$ in $N$ non-overlapping meshes (the sub-domains) $\{T_i\}_{1 \leq i \leq N}$ using standard graph partitioners, e.g., METIS [11]. If $\delta$ is a positive integer, the overlapping decomposition $\{T_i^\delta\}_{1 \leq i \leq N}$ is defined recursively as follows: $T_i^\delta$ is obtained by including
all elements of \( T_{i}^{\delta-1} \) plus all adjacent elements of \( T_{i}^{\delta} \). For \( \delta = 0, T_{i}^{\delta} = T_{i} \). Let \( \{ V_{i}^{\delta} \}_{1 \leq i \leq N} \) be the local deformation FE spaces defined on \( \{ T_{i}^{\delta} \}_{1 \leq i \leq N} \).

Now, consider the restrictions \( \{ R_{i} \}_{1 \leq i \leq N} \) from \( V_{h} \) to \( \{ V_{i}^{\delta} \}_{1 \leq i \leq N} \) and a local partitions of unity \( \{ D_{i} \}_{1 \leq i \leq N} \) such that:

\[
\sum_{j=1}^{N} R_{i}^{T} D_{j} R_{j} = I_{d_{i} \times d_{i}},
\]

where \( I_{d} \) denotes the identity matrix and \( n \) is the global number of unknowns in the deformation space. Algebraically speaking, if \( n \) is the global number of deformation unknowns and \( \{ n_{i} \}_{1 \leq i \leq N} \) are the numbers of degrees of freedom in each local deformation FE space, then \( R_{i} \) is a Boolean matrix of size \( n_{i} \times n_{i} \), and \( D_{i} \) is a diagonal matrix of size \( n_{i} \times n_{i} \), for all \( 1 \leq i \leq N \).

Using the partition of unity, one can use the one-level preconditioner, Restricted Additive Schwarz (RAS) method, proposed by Cai and Sarkis [2]:

\[
M_{\text{RAS}}^{-1} = \sum_{i=1}^{N} R_{i}^{T} D_{i} A_{i}^{-1} R_{i},
\]

where the \( \{ A_{i} \}_{1 \leq i \leq N} \) are local operators defined by the submatrices \( \{ R_{i} A R_{i}^{T} \}_{1 \leq i \leq N} \).

In this case, we thus chose to use a more sophisticated multilevel domain decomposition method using the GenEO approach [17, 6]. This preconditioner, \( M_{\text{GenEO}}^{-1} \), uses a spectral coarse grid to better couple all sub-domains.

### 3 Results

This section aims to assess the efficiency of linear solvers described in the previous section in the presence of non-linear deformation model. To this end, we consider Saint-Venant Kirchhoff as a constitutive law to model the liver mechanical response. The Saint-Venant Kirchhoff law is given by the following potential:

\[
W(e) = \frac{\lambda}{2} (\text{Tr}(e))^{2} + \mu (\text{Tr}(e^{2})),
\]

where \( e \) is the Green Lagrange tensor \( e = \frac{1}{2} \left( \nabla y + (\nabla y)^{T} + (\nabla y)^{T} \cdot \nabla y \right) \). Then the second Piola stress tensor is given by \( \Sigma = \frac{\partial W}{\partial e} \). \( \lambda \) and \( \mu \) are the Lamé coefficients that can be determined from the Youngs modulus \( E \) and Poissons ratio \( \nu \).

The geometry of the model is segmented from a patient pre-operative Computed Tomography (CT) image. The domain is then meshed in a set of linear tetrahedral elements using GMSH [5]. In all the following simulations, we use two different mesh discretizations. An initial tetrahedral mesh with 3316 elements which yields a linear system \( A \) with 2000 unknowns. Then, we refine the same mesh by splitting each element into multiple smaller elements to get a
finer mesh with 22208 elements, which yields a linear system $A$ with 12321 unknowns. For domain decomposition purpose, the mesh is decomposed using the graph partitioner METIS [11] (figure 1). The resulting finite element linear system is preconditioned with $M_{\text{GenEO}}^{-1}$, and the GMRES solver is stopped when the relative preconditioned residual is lower than $10^{-6}$.

To implement the physical model, we have employed the open source simulation software FreeFem [7]. The linear solvers and the preconditioners are implemented in HPDDM [10]. We used the PARDISO [16] library for direct solver. Results were obtained on a standard desktop machine equipped with Intel with 6 Intel cores clocked at 3.2 GHz.

![Liver domain decomposition in 4 sub domains.](image1)

![Liver deformation estimated by a non linear model.](image2)

**Fig 1**: Left: liver computational domain decomposed in 4 sub-domains using the graph partitioner METIS [11] - Right: Liver deformation estimated by the non-linear Saint-Venant Kirchhoff law (the initial configuration is represented by blue points).

### 3.1 Static non-linear deformation

In this paragraph we evaluate the performances of pure direct solver versus domain decomposition solver for static deformation using a Newton-Raphson algorithm. We simulated an entire liver deformation. A volumic force of $100 \, Pa$ in the $(x + y)$ direction is uniformly applied to the liver while several selected vertices of a plane are fixed (representing the ligament and veins). We considered Young modulus $E = 3 \cdot 10^5$ and Poisson ratio $\nu = 0.35$. We do not consider the liver as a fully incompressible material in this test. The simulated deformation is shown in Fig. 1b, where the deformed mesh is plotted as well as the initial liver configuration (represented by blue points).
Table 1: Breakdown of the time spent in solver steps for 3D non-linear solver with respect to the number of subdomains, the second column corresponds to the maximum number of unknowns per subdomain, the third column is the time spent in building the DDM preconditioner, and the fourth column corresponds to the time spent in solving the multiple Newton inner linear systems.

<table>
<thead>
<tr>
<th>$N$</th>
<th># d.o.f per sub.</th>
<th>Prec(ms)</th>
<th>Solves(ms)</th>
<th>Total(ms)</th>
<th># Newton iterations</th>
<th># Iterations per Newton it</th>
<th>Speedup</th>
<th># of d.o.f</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1998</td>
<td>0.00</td>
<td>47.96</td>
<td>47.96</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>1263</td>
<td>16.50</td>
<td>11.05</td>
<td>27.55</td>
<td>3</td>
<td>8</td>
<td>1.7</td>
<td>2.00 · 10^3</td>
</tr>
<tr>
<td>4</td>
<td>897</td>
<td>12.88</td>
<td>7.89</td>
<td>20.77</td>
<td>3</td>
<td>9</td>
<td>2.3</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>12321</td>
<td>0.00</td>
<td>421.97</td>
<td>421.97</td>
<td>4</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>7056</td>
<td>142.68</td>
<td>105.86</td>
<td>248.54</td>
<td>4</td>
<td>10</td>
<td>1.7</td>
<td>12.32 · 10^3</td>
</tr>
<tr>
<td>4</td>
<td>4494</td>
<td>99.00</td>
<td>94.92</td>
<td>193.91</td>
<td>4</td>
<td>13</td>
<td>2.1</td>
<td>-</td>
</tr>
</tbody>
</table>

In Table 1, we report the time spent in all subroutines included during the multiples Newton iterations solve with respect to the number of subdomains $N$. The case $N = 1$ corresponds direct solver case, where the system is first factorized with and $LDL^T$ than solved. For $N \geq 2$, the system is solved using GMRES with a domain decomposition preconditioner $M^{-1}_{GenEO}$. Very few Newton iterations (column number 6) are needed for the solver to converge, independently of the number of subdomains (first column). The scalability of the solving approach is reported in the table using the run of the direct solver (a.k.a $N = 1$) as a reference. For each mesh, we ensure that we are calculating the same solution regardless of the linear solver type (direct or domain decomposition solver). To do that, we make sure that $L_2$ norm of the final deformation is the same during each scalability test.

We notice that domain decomposition approach is already providing a reasonable speedup with respect to the number of subdomains. Moreover, we observe that using a simple direct solver approach requires to build the factorization again for each inner Newton iteration with the same high cost. Whereas with domain decomposition approach the same preconditioner is reused for all the Newton iterations with no significant impact on the GMRES solver, thanks to the robustness of the DD preconditioner, the number of Krylov iterations remains stable. This suggests that we can take more benefits in a scenario of dynamic deformations, where we need to solve more linear systems successively through both Newton iterations and time integration algorithm.

The scalability is impacted by the number of unknowns per subdomain, which is not scaling linearly with the number subdomains. This fact is first due to the load balancing provided by the graph partitioner and also due to the overlapping regions between subdomains, which seems to be more critical in case of small meshes. We also notice that the increase in the number of unknowns from 2000 to 12321 leads to the increase of the computation time. This happens because the convergence of the iterative solvers is influenced by the condition number.
of the stiffness matrix, and the condition number will increase with a decreased element size for a given object.

### 3.2 Dynamic non-linear deformation

The main objective of this paragraph is to show that the cost of the domain decomposition preconditioner is quickly amortized in the scenario of successive solutions of linear systems. Typically in case of stiff elastic deformation where the implicit integration method is more appropriate. To this end, we solve the entire discretized problem (4), where, a Newmark, implicit time integration scheme is used with a time step of 0.01s, in this scenario, the same volume force as the one used in the static case is again uniformly applied to the liver (Fig. 1b) for 0.03s than released to let the system reaches the equilibrium state. We simulate the liver deformation for both discretizations with coarse mesh yielding a system of 2000 unknowns and refined mesh yielding a system of 12321 unknowns. In Fig 2, we show the behavior of computational time spent in solving the successive linear system. The red and blue curves represent, respectively, the coarse and fine mesh discretization. The global domain has been decomposed in 4 subdomains allowing the simulation to run on 4 processors. For each time step, the Newton-Raphson algorithm is performed to update the deformation state. We use the domain decomposition approach as a solver, where preconditioner $M^{-1}_{\text{GenEO}}$ is built at the first time step than used for preconditioning all the following Newton inner linear systems.

![Fig 2: Performances of domain decomposition solver during the simulation of a dynamic non-linear deformation liver response. Using the Saint-Venant Kirchhoff model. Red and blue curves represent the computation time per Newton iteration over time steps with resp to coarse mesh (2000 unknowns) and refined mesh (12321 unknowns).]
3.3 Contact & deformation

While keeping the context of simulating the behavior of soft tissues, we can point out the importance of providing methods and models to simulate the mechanical interactions on the organ model: interactions with the surrounding anatomical structures (i.e. contacts), interactions with different surgical instruments leading to contacts or other types of complex interactions. Simulating these interactions necessitate to detect them, to model them, and then to solve them with adapted numerical methods. In practice the numerical system to solve is an augmented version of linear systems that we have solved in previous paragraphs. If we consider the case of a two-body interaction, the augmented system takes the following form

\[ AU = b + J^T \lambda \]  

which can be formulated as

\[
\begin{bmatrix}
A_1 & 0 & J_1^T \\
0 & A_2 & J_1^T \\
J_1 & J_2 & 0
\end{bmatrix}
\begin{bmatrix}
y_1 \\
y_2 \\
\lambda
\end{bmatrix}
= 
\begin{bmatrix}
b_1 \\
b_2 \\
-\delta
\end{bmatrix}
\]  

where \( A_1 \) and \( A_2 \) correspond to the linear operators of each body, \( J_1 \) model the interaction, then the unknowns are the \( y_1, y_2 \) are respectively the displacement of the two bodies and \( \lambda \) is the vector of contact forces. We believe that in such case if interaction, domain decomposition approach has a tremendous potential for solving problem (7). And more than accelerating the linear solves, domain decomposition can be specifically designed to consider the interaction area as a single sub-domain domain. This possibility would allow a partial updating of the global operator, and is likely to lead to a substantial gain in simulation time.

4 Conclusion

In this work, we have investigated the computational expense of solving linear systems resulting from a combination of non-linear model and dynamic integration. We showed that employing hybrid solver like a domain decomposition method has a real potential to harness the capability of small parallel machines since it takes full advantage in making the solving procedure fully parallel. On the other hand, the robustness of domain decomposition preconditioners makes it possible to reuse the preconditioner for successive solves. These two strategies combined open up the possibility to significantly accelerate the computation for complex simulation and meet the real-time feedback, which is a hard constraint in surgical training or intra-operative guidance.

The next step will be to integrate the domain decomposition paradigm with fast hyperelastic FEM models and implicit contact schemes. We will also investigate further the limited scalability of the current approach when dealing with real-time applications, which is likely due to load balancing. This can be improved by a better tuning of the graph partitioner.
Bibliography