31

A Domain Decomposition Method for Control Problems

Jean-David Benamou

1 Introduction

A DDM (Domain Decomposition Method) for the optimal control of systems governed by PDEs (Partial Differential Equations) is presented. The general framework is based on a nonoverlapping spatial decomposition of the domain (time is not decomposed for evolution equations) and the introduction of *skew-symmetric, Robin, iterative transmission conditions* between subdomains which couple the direct and adjoint states appearing in the optimality system (derived from the control problem). A reference paper on this family of DDM is [Lio90]; see also [RG90b, Des91, Des93, BD96] on the extension of this method to the Helmholtz equation.

Because of the natural coupling between direct and adjoint states in optimality systems, a general strategy for proving convergence of the DDM is available which is independent of the nature of the governing equation. In our opinion, this general convergence property makes the SRC (Skew-symmetric, Robin, Coupled) transmissions conditions the natural choice to domain decompose control problems.

Unlike for PDEs, the resolution of optimal control problem using DDMs has received little attention (at least in the previous edition of this conference). On this precise subject we can only cite [AB73, Bou, Leu96]. A multigrid method can also be found in [W.H79]. Any DDM, relevant to solve a PDE, can of course be used as the kernel of an optimization algorithm (of gradient type for instance). The originality of our approach, developed in [Ben93, Ben96a, Ben96b, BD96, Ben95b, Ben95a], consists in decomposing the full optimality system. This means that our DDM solves concurrently the equations and the optimization problem.

In the following two sections we briefly present the method on simple model problems. We then describe a general proof of convergence. We discuss, in a fourth section, the possible extension of this method to more complicated and also different control problems. Based on our own experiments, we finally give a few remarks on the implementation of this method.

Ninth International Conference on Domain Decomposition Methods
Editor Petter E. Bjørstad, Magne S. Espedal and David E. Keyes

2 The Model Problems

The goal of the control problem is to minimize a cost function:

$$\min_{u \in U} J(u, y(u)) \tag{1}$$

where u is a an admissible control which acts on $y(u)$ the solution of a PDE which can be either elliptic, parabolic or hyperbolic (to avoid unnecessary complications in the notations, we drop the dependence of y in u in the remainder of the paper):

$$\begin{aligned}
&-\Delta y(x) = f(x) + u(x) \quad \text{on } \Omega \\
&(\frac{\partial}{\partial t} - \Delta)y(t,x) = f(t,x) + u(t,x) \quad \text{on }]0,T[\times\Omega, \text{ with initial condition} \\
&y(0,x) = y_0(x). \\
&(\frac{\partial^2}{\partial t^2} - \Delta)y(t,x) = f(t,x) + u(t,x) \quad \text{on }]0,T[\times\Omega, \text{ with initial conditions} \\
&y(0,x) = y_0(x) \;\; \frac{\partial}{\partial t}y(0,x) = y_1(x).
\end{aligned} \tag{2}$$

The spatial domain is Ω and the time domain $]0,T[$. We choose, for simplicity, a Dirichlet boundary condition on $\Gamma = \partial\Omega$:

$$\begin{aligned}
&y(x) = g(x) \;\; \textit{on } \Gamma && \text{or} \\
&y(t,x) = g(t,x) \;\; \textit{on }]0,T[\times\Gamma && \text{for the last two equations}
\end{aligned} \tag{3}$$

In the first case y is independent of time, U is taken as a convex subset of $L^2(\Omega)$ and

$$J(u,y) = \frac{1}{2}\int_\Omega |y(x)|^2 + \alpha|u(x)|^2 dx. \tag{4}$$

For the time-dependent problems we take U as a convex subset of $L^2(]0,T[\times\Omega)$ and

$$J(u,y) = \frac{1}{2}\int_{]0,T[\times\Omega} |y(x)|^2 + \alpha|u(t,x)|^2 dxdt. \tag{5}$$

These cost functions are a compromise between the desired damping of a physical quantity y and the cost of controlling the system represented by u term. A positive penalization parameter α controls this trade-off. The set of admissible controls U takes into account the possible constraints on the control which we restrict to be local in space and time and linear. For more on optimal control problems and also on mathematical issues such as well posedness, we refer to [Lio68] where can be found, in particular, the following reformulation of these problems as *optimality systems*. The solution is given by the resolution of (2)–(3), called the direct equations, together with (respectively)

$$\begin{aligned}
&-\Delta p(x) = y(x) \quad \text{on } \Omega \\
&(-\frac{\partial}{\partial t} - \Delta)p(t,x) = y(t,x) \quad \text{on }]0,T[\times\Omega, \text{ with initial condition} \\
&p(T,x) = 0. \\
&(\frac{\partial^2}{\partial t^2} - \Delta)p(t,x) = y(t,x) \quad \text{on }]0,T[\times\Omega, \text{ with initial conditions} \\
&p(T,x) = 0 \;\; \frac{\partial}{\partial t}p(T,x) = 0.
\end{aligned} \tag{6}$$

$$
\begin{aligned}
&p(x) = 0 \quad \text{on} \quad \Gamma && \text{or} \\
&p(t,x) = 0 \quad \text{on} \quad]0,T[\times\Gamma && \text{for the last two equations}
\end{aligned} \tag{7}
$$

called the adjoint equations (backward in time) and the optimality conditions

$$
\begin{aligned}
&\int_{\Omega} (p + \alpha\, u)(v - u)\ dx \geq 0,\ \ \forall v \in U, && \text{or} \\
&\int_{]0,T[\times\Omega} (p + \alpha\, u)(v - u)\ dxdt \geq 0,\ \ \forall v \in U && \text{for the evolution problems.}
\end{aligned} \tag{8}
$$

From now on, we drop the dependence in space and time in the notation, these being implicitly defined by the type of problem we consider.

3 The Domain Decomposition Method

We decompose the domain as follows: Let $\Omega_i,\ i = 1, m$ be a partition of Ω (i.e. $\Omega = \cup_i \bar{\Omega}_i$ and $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$). We denote $\Gamma_i = \partial\Omega_i \cap \partial\Omega$ as the 'exterior' boundary of the subdomains and $\Sigma_{ij} = \partial\Omega_i \cap \partial\Omega_j$ as the interfaces. The external normal of $\partial\Omega_i$ is noted ν_i. We assume that the geometry of this decomposition is, as Ω, regular enough to ensure the well posedness of the global problems and the local subproblems defined by the DDM.

We schematically describe the *iterative* methods. At each step $n+1$ we perform local resolutions of (2), (3), (6), (7) and (8) where Ω has systematically been replaced by Ω_i in (2), (6) and (8) and Γ by Γ_i in (3) and (7). We call the local solutions at step $n+1$ on subdomain Ω_i: $(y_i^{n+1}, p_i^{n+1}, u_i^{n+1})$.

For instance, the local optimality conditions on Ω_i at step $n+1$ are obtained by replacing (8) with

$$
\begin{aligned}
&\int_{\Omega_i} (p_i^{n+1} + \alpha\, u_i^{n+1})(v_i - u_i^{n+1})\ dx \geq 0,\ \ \forall v_i \in U_i, && \text{or} \\
&\int_{]0,T[\times\Omega_i} (p_i^{n+1} + \alpha\, u_i^{n+1})(v - u_i^{n+1})\ dxdt \geq 0,\ \ \forall v_i \in U_i, &&
\end{aligned} \tag{9}
$$

where U_i is a set of local admissible controls satisfying the same local constraints as the elements of U.

The only missing ingredients are the transmission conditions on the interfaces between subdomains (i.e. the new boundaries generated by the decomposition of the domain). They provide the mechanism which links the resolutions between neighboring subdomains at successive iteration steps. The choice of these two boundary conditions (one for the direct equation and one for the adjoint) is discussed in detail in the above mentioned references and are the same SRC transmissions conditions for all considered PDEs:

$$
\begin{aligned}
\frac{\partial}{\partial\nu_i} y_i^{n+1} + \beta\, p_i^{n+1} &= -\frac{\partial}{\partial\nu_j} y_j^n + \beta\, p_j^n \quad \textit{on } \Sigma_{ij}, \\
\frac{\partial}{\partial\nu_i} p_i^{n+1} - \beta\, y_i^{n+1} &= -\frac{\partial}{\partial\nu_j} p_j^n - \beta\, y_j^n \quad \textit{on } \Sigma_{ij}.
\end{aligned} \tag{10}
$$

The choice of β, a positive parameter, is discussed in section 6. We take $\beta = 1$ in the next section to simplify the presentation of a unified proof of convergence.

Let us finally mention that these local problems can be reformulated as optimal control problems and local cost functions can be derived depending on the original quantities to be minimized but also on a new "transmission" cost. This new cost arises because of the coupling in the transmissions conditions (10). This shows, as stated in the introduction, that we actually decompose the full optimization problem.

4 Convergence

The convergence is established using the equation on the local errors:

$$(\tilde{y}_i^{n+1}, \tilde{p}_i^{n+1}, \tilde{u}_i^{n+1}) = (y, p, u) - (y_i^{n+1}, p_i^{n+1}, u_i^{n+1}) \tag{11}$$

on each subdomain. These errors satisfy the same equations as $(y_i^{n+1}, p_i^{n+1}, u_i^{n+1})$ with $f = g = y_0 = y_1 = 0$. We *assume* (see above mentioned references for more details on mathematical issues) that the regularity of the global and local solutions allow the computations made in this section.

We introduce the following notations

$$\begin{aligned} \|z\|_{ij}^2 &= \int_{\Sigma_{ij}} z^2 \, d\sigma \ , \quad (z, z')_{ij} = \int_{\Sigma_{ij}} z z' \, d\sigma \\ \|z\|_i^2 &= \int_{\Omega_i} z^2 \, dx \ , \quad (z, z')_i = \int_{\Omega_i} z z' \, dx \end{aligned} \tag{12}$$

for the elliptic problem and

$$\begin{aligned} \|z\|_{ij}^2 &= \int_{]0,T[\times\Sigma_{ij}} z^2 \, d\sigma dt \ , \quad (z, z')_i = \int_{]0,T[\times\Sigma_{ij}} z z' \, d\sigma dt \\ \|z\|_i^2 &= \int_{]0,T[\times\Omega_i} z^2 \, dx dt \ , \quad (z, z')_i = \int_{]0,T[\times\Omega_i} z z' \, dx dt \end{aligned} \tag{13}$$

for the last two cases.

We are now able to give a general proof of convergence. To this end we introduce the following "energy" with support on the interfaces (note that each interface is counted twice, we have indeed $\Sigma_{ij} = \Sigma_{ji}$):

$$E^{n+1} = \sum_{i \neq j, \, s.t. \Sigma_{ij} \neq \emptyset} \{ \| \frac{\partial}{\partial \nu_i} \tilde{y}_i^{n+1} \|_{ij}^2 + \| \tilde{p}_i^{n+1} \|_{ij}^2 + \| \frac{\partial}{\partial \nu_i} \tilde{p}_i^{n+1} \|_{ij}^2 + \| \tilde{y}_i^{n+1} \|_{ij}^2 \}. \tag{14}$$

Using (10), the energies can be shown to satisfy (recall that we take $\beta = 1$)

$$\begin{aligned} E^{n+1} = E^n - 2 \\ \sum_{i \neq j, \, s.t. \Sigma_{ij} \neq \emptyset} \{ (\frac{\partial}{\partial \nu_i} \tilde{y}_i^{n+1}, \tilde{p}_i^{n+1})_{ij} - (\frac{\partial}{\partial \nu_i} \tilde{p}_i^{n+1}, \tilde{y}_i^{n+1})_{ij} \\ + (\frac{\partial}{\partial \nu_i} \tilde{y}_i^n, \tilde{p}_i^n)_{ij} - (\frac{\partial}{\partial \nu_i} \tilde{p}_i^n, \tilde{y}_i^n)_{ij} \}. \end{aligned} \tag{15}$$

The last terms in the above expression are evaluated as follows. On each subdomain, the direct equation (in $\tilde{y}_i^{n+1}$) is multiplied by $\tilde{p}_i^{n+1}$ and the adjoint equation (in $\tilde{p}_i^{n+1}$)

is multiplied by $-\tilde{y}_i^{n+1}$. We then integrate by part in space and integrate in space and time in the case of evolution equations. Adding the results, we note that the terms involving time derivatives and gradients vanish and finally obtain, for all i (regardless of the considered PDE):

$$-\sum_{j,s.t.\,\Sigma_{ij}\neq\emptyset}\{(\frac{\partial}{\partial\nu_i}\tilde{y}_i^{n+1},\tilde{p}_i^{n+1})_{ij}-(\frac{\partial}{\partial\nu_i}\tilde{p}_i^{n+1},\tilde{y}_i^{n+1})_{ij}\} = -\|\tilde{y}_i^{n+1}\|_i^2+(\tilde{p}_i^{n+1},\tilde{u}_i^{n+1})_i \tag{16}$$

We now use the optimality conditions. We choose $v = u_i^{n+1}$ on Ω_i and 0 elsewhere in (8) and $v_i = u$ in (9) and subtract the two inequalities. This yields the estimate

$$(\tilde{p}_i^{n+1},\tilde{u}_i^{n+1})_i \leq -\alpha\|\widetilde{u_i^{n+1}}\|_i^2. \tag{17}$$

Combining (16) and (17) (and similar results for step n) and (15), we establish a law of decrease for the energy:

$$E^{n+1} \leq E^n - 2\sum_i\{\|\tilde{y}_i^{n+1}\|_i^2+\alpha\|\tilde{u}_i^{n+1}\|_i^2+\|\tilde{y}_i^n\|_i^2+\alpha\|\tilde{u}_i^n\|_i^2\}. \tag{18}$$

Summing over n gives straightforwardly a first result of convergence on the errors on each subdomain Ω_i:

$$\|\tilde{y}_i^{n+1}\|_i \xrightarrow{n} 0 \quad \|\tilde{u}_i^{n+1}\|_i \xrightarrow{n} 0. \tag{19}$$

This result can be improved using the equations and the uniform boundedness of E^n.

5 Other Problems

The convergence of this method relies on the SRC transmission conditions and the structure of optimal control problems reformulated as coupled system formed of a direct and adjoint equations and a optimality condition. The *convexity* of the cost function provides the necessary coercivity for this system. This explains why this method can be applied to a wide range of linear optimal control problems involving more complicated (possibly nonsymmetric and inhomogeneous) operators for the equation and the boundary conditions. Boundary observation and control problems can also be treated. Dealing with nonlocal observation is also possible. It however couples the resolution of subproblems set on the domain of observation.

This DDM can be applied to noncoercive PDEs such as the Helmholtz equation. The proof of convergence remains formally the same but the usual bilinear form is replaced by a sesquilinear form.

We are currently working on the adaptation of this method to the decomposition of the HUM method for exact controllability problems [Lio88].

Let us mention that it is also possible, at least formally, to use the same SRC transmision conditions to domain decompose in time.

6 Remarks

Numerical Implementation and Speed of Convergence

A good way of discretizing these problem is to use mixed hybrid finite elements (see [CJ86] or [RG90a, RG88] on the use of mixed finite elements in domain decomposition methods). This approach is well suited to our problem for it uses in particular, as degrees of freedom, the fluxes of the normal derivatives and the average values of the trace of the direct and adjoint states on the interfaces which are the natural unknowns of our transmission conditions. The proof of convergence in the continuous case is easily extended to this mixed hybrid discrete formulation as it allows an exact discrete integration by part.

The parameter β has a decisive influence on the speed of convergence. We always choose it proportional to $\frac{1}{h}$ where h is the size of the finite elements. The discrete transmission conditions are in these case adimensional.

It was shown in [Des93] (for the direct Helmholtz equation) that the eigenvalues of the discrete iteration operator may be close to 1. This explains the observed bad convergence behavior of the algorithm. A simple way to remedy to this situation (still [Des93]) is to use and under-relaxed version of the transmission conditions. For control problems they take the form

$$\begin{aligned}\frac{\partial}{\partial \nu_i} y_i^{n+1} + \beta\, p_i^{n+1} &= \gamma\Big(-\frac{\partial}{\partial \nu_j} y_j^n + \beta\, p_j^n\Big) + (1-\gamma)\Big(\frac{\partial}{\partial \nu_i} y_i^n + \beta\, p_i^n\Big) \quad \textit{on } \Sigma_{ij},\\ \frac{\partial}{\partial \nu_i} p_i^{n+1} - \beta\, y_i^{n+1} &= \gamma\Big(-\frac{\partial}{\partial \nu_j} p_j^n - \beta\, y_j^n\Big) + (1-\gamma)\Big(\frac{\partial}{\partial \nu_i} p_i^n - \beta\, y_i^n\Big) \quad \textit{on } \Sigma_{ij} \qquad (20)\end{aligned}$$

The relaxation parameter γ has to belong to $]0,1[$ (we usually choose $\gamma = \frac{1}{2}$). The theoretical convergence proof can also be established with (20) instead of (10).

More pragmatically, the convergence speed observed in actual simulations is always proportional to the number of subdomains. Multilevel or multigrid method are much faster for standard direct elliptic problems (see [SBG96] for instance). This is not so evident for hyperbolic problems or noncoercive elliptic equations such as the Helmholtz equation without even speaking of extending this methods to the resolution of control problems.

Resolution of the Subproblems

Our own approach on the decomposition of the domain and the resolution of the subproblems was to take the smallest possible subdomains, i.e. each finite element is a subdomain. This is of course the worst possible choice in terms of number of iterations but it restricts the number of degrees of freedom on each subdomain to a minimum and allows in most cases an analytical resolution of the subproblems. This strategy is well suited to an implementation on a massively parallel machine but requires a structured meshing of the domain.

In the case of complicated geometries, a domain decomposition in simple shapes may be a good motivation to use this domain decomposition method. A method of resolution of the subproblems is still needed.

As already mentioned in section 2, in the case of evolution equations the adjoint equation is backward in time. Gradient-type methods rely on iterating successively forward integrations for the direct equation and backward integrations for the adjoint equation. The adjoint variable providing a descent direction for the optimization method. An other possibility is to compute the *feed-back law* which express the linear relationship between p and y. It can *formally* be written $\{p(t), \frac{\partial}{\partial t}p(t)\} = A(t) * \{y(t), \frac{\partial}{\partial t}y(t)\} + \{r0(t), r1(t)\}$ [Lio68]. At each instant t, an operator $A(t)$ maps a space-dependent function into an other space-dependent function. If A and r can be computed, the feed-back law can be used to eliminate p in the direct equation. The operators $A(t)$ satisfy a nonlinear Riccati differential equation which can be difficult to solve at least because of the size of the discretization of $A(t)$. The resolution of a PDE depending on the second hand terms of the original system (and on A) gives r. A similar feed-back law $\{p_i^{n+1}(t), \frac{\partial}{\partial t}p_i^{n+1}(t)\} = A_i^{n+1}(t) * \{y_i^{n+1}(t), \frac{\partial}{\partial t}y_i^{n+1}(t)\} + \{r0_i^{n+1}(t), r1_i^{n+1}(t)\}$ can also be defined for each of the subproblems of our DDM. The important remark here is that A_i^{n+1} *does not depend on the iterative process* but only on the geometry of the decomposition and only need to be computed once. This is because it defines the homogeneous part of the feed back law. The DDM therefore implicitly decompose the computation of the feed-back law. A numerical implementation of the method using this technique is presented in [Ben97].

REFERENCES

[AB73] A. Bensoussan R. Glowinsky J. L. (1973) Méthode de décomposition appliquée au contrôle optimal de systèmes distribués. In *5th IFIP Conference on Optimization techniques.* Lecture Notes in Computer Science.

[BD96] Benamou J. and Després B. (1996) A domain decomposition method for the Helmholtz equation and related optimal control problems. *Submitted to J. Comp. Physics and INRIA Tech. Report 2791* .

[Ben93] Benamou J. (1993) Décomposition de domaine pour le contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles elliptiques. *C. R .Acad. Sci. Paris* 317: 205–209.

[Ben95a] Benamou J. (1995) A domain decomposition method for the optimal control of system governed by the helmholtz equation. In Cohen G. (ed) *Third international conference on mathematical and numerical wave propagation phenomena (Juan-les-Pins.* SIAM.

[Ben95b] Benamou J. (1995) A massively parallel algorithm for the optimal control of systems governed by elliptic p.d.e.'s. In *seventh SIAM conference on parallel processing for scientific computing.* SIAM.

[Ben96a] Benamou J. (1996) Décomposition de domaine pour le contrôle de systèmes gouvernés par des équations d'évolution. *Submitted to C. R .Acad. Sci. Paris* .

[Ben96b] Benamou J. (1996) Domain decomposition methods with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations. *to appear in SINUM and INRIA Tech. Report 2246* .

[Ben97] Benamou J.-D. (1997) Optimal control of systems governed by the wave equation : resolution of a test case using a domain decomposition method. Technical report, INRIA.

[Bou] Bounaim A.In *This proceedings.*

[CR91] Chavent G. and Roberts J. (December 1991) A unified physical presentation of mixed, mixed-hybrid finite elements and standard finite difference approximations for the determination of velocities in waterflow problems. *Advances in Water Ressources* 14(6): 329–348.

[Des91] Després B. (1991) Domain decomposition method and the Helmholtz problem. In Cohen G., Halpern L., and Joly P. (eds) *Mathematical and Numerical Aspects of*

Wave Propagation Phenomena, pages 44–52. SIAM.

[Des93] Després B. (1993) Domain decomposition method and the Helmholtz problem (part ii). In Kleinman R., Angell T., Colton D., Santosa F., and Stackgold I. (eds) *Second international conference on mathematical and numerical aspects of wave propagation phenoma.* SIAM.

[Leu96] Leugering G. (1996) On dynamic domain decomposition of controlled networks of strings and joint-masses.

[Lio68] Lions J. (1968) *Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles.* Dunod, Paris.

[Lio88] Lions J. (1988) *Contrôlabilité exacte perturbations et stabilisation de systèmes distribués, tome 1.* Masson, Paris.

[Lio90] Lions P. (1990) On the Schwarz alternating method 3. In Chan T. F., Glowinski R., Périaux J., and Widlund O. B. (eds) *Third international symposium on domain decomposition methods for partial differential equations.* SIAM.

[RG88] R. Glowinski M. W. (1988) Domain decomposition and mixed finite element methods for elliptic problems. In Glowinski R., Golub G. H., Meurant G. A., and Périaux J. (eds) *First international symposium on domain decomposition methods for partial differential equations.* SIAM.

[RG90a] R. Glowinski M.Kinton M. W. (1990) Acceleration of domain decomposition algorithms for mixed finite element methods by multi level methods. In Chan T. F., Glowinski R., Périaux J., and Widlund O. B. (eds) *Third international symposium on domain decomposition methods for partial differential equations.* SIAM.

[RG90b] R. Glowinski P. L. (1990) Augmented lagrangian interpretation of the nonoverlapping Schwarz alternating method. In Chan T. F., Glowinski R., Périaux J., and Widlund O. B. (eds) *third international symposium on domain decomposition methods for partial differential equations.* SIAM.

[SBG96] Smith B., Bjorstad P., and Gropp W. (1996) *Parallel Multilevel Methods for Elliptic Partial Differential Equations.* Cambridge University Press.

[W.H79] W.Hackbush (1979) On the fast solving of parabolic boundary control problems. *SIAM J. Control and Optimization* 17: 231–244.