

Time Parallelization of a Distributed Optimal Control Problem for the Wave Equation

Thanh Vuong Dang^[0009-0003-3039-1428], Bérangère Delourme^[0000-0002-8316-7572],
and Laurence Halpern^[0000-0002-7877-7130]

1 Introduction

The objective of this work is to present domain decomposition methods in time to solve a distributed optimal control problem governed by the wave equation. This problem is challenging to solve numerically. Indeed, the optimization procedure requires either to solve at each iteration in turn forward and backward equations or to solve a huge non triangular linear system. It therefore requires to be parallelized. This question has been addressed before in the same spirit for a boundary control in [5]; see also [4] for a multiple shooting approach.

The time interval is divided into overlapping subintervals, and we design an iterative method for which the subinterval problems minimize extended cost functions. For this reason the algorithm is called inherited. We introduce a matrix sine computation, very convenient to compute a convergence factor of the algorithm. This allows us to show that our new algorithm is a smoother, and study its convergence in a simpler case. To improve the convergence properties, we introduce a relaxation parameter. Finally, we introduce the natural Dirichlet Neumann algorithm (see [1, 3]) and study its properties by the same techniques.

The problem and one of the algorithms are similar to those in [5], where a boundary control problem is analyzed with the method of energy. The novelty here lies in the computation of a convergence factor, which permits to compare the convergence properties of various algorithms. This paper is a summary of some results in the unpublished thesis [7], pertaining to the sine matrix method, which extends the Fourier method.

Thanh Vuong Dang
LAGA, Université Sorbonne Paris-Nord, e-mail: dangthanh@math.univ-paris13.fr

Bérangère Delourme
LAGA, Université Sorbonne Paris-Nord, e-mail: delourme@math.univ-paris13.fr

Laurence Halpern
LAGA, Université Sorbonne Paris-Nord, e-mail: halpern@math.univ-paris13.fr

Let us describe the problem under consideration. Let $\Omega \subset \mathbb{R}^n$ for $n \in \{1, 2, 3\}$, $T > 0$ and $I = (0, T)$. We consider the following optimal control problem within the Hilbert spaces $H = L^2(\Omega)$, $H_I = L^2(I)$, $\mathcal{H} = H \times H_T$:

$$\min_{v \in \mathcal{H}} \hat{\mathcal{J}}(v) = \mathcal{J}(v, Y(v, Y^{(0)}, I)) \quad \text{with } \mathcal{J}(v, y) = \frac{1}{2} \|y - \hat{y}\|_{\mathcal{H}}^2 + \frac{\alpha^2}{2} \|v\|_{\mathcal{H}}^2,$$

where $Y(v, Y^{(0)}, I)$ is the solution $y(x, t)$ of the wave equation on $\Omega \times I$, with source term v and initial data $Y^{(0)} := (y^{(0)}, y^{(1)})$:

$$\begin{cases} \partial_{tt} y - \Delta y = v & \text{in } \Omega \times I, \\ y|_{\partial\Omega} = 0, & (y, y_t)(\cdot, 0) = Y^{(0)} \text{ in } \Omega. \end{cases} \quad (1)$$

Remark 1 This is a tracking problem : minimizing the functional $\hat{\mathcal{J}}$ requires to find a control, whose L_2 norm is not too large and such that the corresponding y is close to \hat{y} ., the coefficient α being a ponderation term. We note that in the thesis [7], a term $\gamma \|y(\cdot, T) - \hat{z}\|_H^2$ is additionally introduced to control the difference $y - \hat{y}$ at the final time T .

Classical results and derivations in optimal control , see [2, 6], show that $\hat{\mathcal{J}}'(u) = \alpha^2 u - \lambda$, where λ is the adjoint state defined by

$$\begin{cases} \partial_{tt} \lambda - \Delta \lambda = \hat{y} - y & \text{in } \Omega \times I, \\ \lambda|_{\partial\Omega} = 0, & (\lambda, \partial_t \lambda)(\cdot, T) = 0 \text{ in } \Omega. \end{cases} \quad (2)$$

The functional $\hat{\mathcal{J}}$ is α^2 -convex, and therefore has a unique minimum point u , which is the unique zero of $\hat{\mathcal{J}}'$, characterized by the optimality system

$$\hat{\mathcal{J}}'(u) = \alpha^2 u - \lambda = 0, \quad (3)$$

with the state equation (1) ($y = Y(u, Y^{(0)}, I)$), and the adjoint equation (2) for the adjoint state λ (backward wave equation, $\lambda(x, t) = \Lambda(\hat{y} - y, (0, 0), I)$). The aim of this presentation is to introduce a parallel in time decomposition algorithms to compute an approximation of u . We consider two overlapping time subintervals only, since the computation of the convergence factor we want to perform is quite difficult.

2 Time domain decomposition: alternate inherited algorithm

The time interval is split into two overlapping subintervals $I_1 = (0, T_1 + \delta)$ and $I_2 = (T_1, T)$. Introduce the cost functions in the subdomains $\Omega \times I_j$, with $H_j = L^2(I_j)$ and $\mathcal{H}_j = H \times H_j$:

$$\begin{aligned}
\mathcal{J}_1(v_1, y_1) &= \frac{1}{2} \|y_1 - \widehat{y}_1\|_{\mathcal{H}_1}^2 + \frac{\alpha^2}{2} \|v_1\|_{\mathcal{H}_1}^2 - (\mu_0, \partial_t y_1(T_1 + \delta))_H + (\mu_1, y_1(T_1 + \delta))_H, \\
\mathcal{J}_2(v_2, y_2) &= \frac{1}{2} \|y_2 - \widehat{y}_2\|_{\mathcal{H}_2}^2 + \frac{\alpha^2}{2} \|v_2\|_{\mathcal{H}_2}^2, \\
\widehat{\mathcal{J}}_1(v_1) &= \mathcal{J}(v_1, Y(v_1, Y^{(0)}, I_1)), \quad \widehat{\mathcal{J}}_2(v_2) = \mathcal{J}(v_2, Y(v_2, Y^{(2)}, I_2)).
\end{aligned} \tag{4}$$

The same analysis as before shows that $\widehat{\mathcal{J}}_1$ and $\widehat{\mathcal{J}}_2$ have a unique minimum, characterized by their corresponding Euler or optimality system. If furthermore at step n ,

$$(\mu_0, \mu_1) = (\lambda_2^n, \partial_t \lambda_2^n)(\cdot, T_1 + \delta), \quad Y^{(2)} = (y_1^{n+1}, \partial_t y_1^{n+1})(\cdot, T_1),$$

then the optimality systems in the subintervals are $\alpha^2 u_j^n = \lambda_j^n$ in $\Omega \times I_j$, with

$$\begin{cases} \partial_{tt} y_1^n - \Delta y_1^n = u_1^n, \\ t = 0, \quad (y_1^n, \partial_t y_1^n) = Y^{(0)}, \\ \partial_{tt} \lambda_1^n - \Delta \lambda_1^n = \widehat{y} - y_1^n, \\ t = T_1 + \delta, \quad (\lambda_1^n, \partial_t \lambda_1^n) = (\lambda_2^{n-1}, \partial_t \lambda_2^{n-1}). \end{cases} \quad \begin{cases} \partial_{tt} y_2^n - \Delta y_2^n = u_2^n \\ t = T_1, \quad (y_2^n, \partial_t y_2^n) = (y_1^n, \partial_t y_1^n), \\ \partial_{tt} \lambda_2^n - \Delta \lambda_2^n = \widehat{y} - y_2^n \\ t = T, \quad (\lambda_2^n, \partial_t \lambda_2^n) = 0. \end{cases} \tag{5}$$

The algorithm (5) is well-defined with an initial guess $(\lambda_2^0, \partial_t \lambda_2^0)(t = T_1 + \delta)$. We call the algorithm inherited, because it uses the transmission conditions inherited from the control problem. It is different to the algorithm for a boundary control in [5] where the transmission conditions are linear combinations of y and λ , and convergence is proved by energy estimates.

3 Modal analysis in dimension 1+1

In order to analyze the convergence of the algorithm, we compute through Fourier series the solutions of the optimality system for which $\widehat{y} = 0$ and $Y^{(0)} = 0$. We restrict ourselves to the one dimensional case $\Omega = (a, b)$. We write the system $\partial_{tt} y - \Delta y - \frac{1}{\alpha^2} \lambda = 0$, $\partial_{tt} \lambda - \Delta \lambda + y = 0$ as a Klein-Gordon vector equation

$$Y = \begin{pmatrix} y \\ \lambda \end{pmatrix}, \quad J = \begin{pmatrix} 0 & -\frac{1}{\alpha} \\ \alpha & 0 \end{pmatrix}, \quad \partial_{tt} Y - \Delta Y + \frac{1}{\alpha} J Y = 0.$$

Then, we expand y and λ on the eigenmodes of the Dirichlet-Laplace operator

$$Y(x, t) = \sum_{k \geq 1} \widehat{Y}(k, t) \Phi_k(x), \quad \Phi_k(x) = \sin(\xi(k)x), \quad k \geq 1, \quad \xi(k) = \frac{k\pi}{b-a}. \tag{6}$$

Introducing the matrix $K(k) := \xi^2(k)I + \frac{1}{\alpha} J$, we obtain the second order vector differential equation in time $\partial_{tt} \widehat{Y}(k, t) + K(k) \widehat{Y}(k, t) = 0$. The general solution is

$$\hat{Y}(k, t) = \cos(N(k)t)Y(k, 0) + N(k)^{-1} \sin(N(k)t)\partial_t Y(k, 0), \quad (7)$$

where $N(k)$ is the square root of $K(k)$. To be more specific, let

$$\mu(k) = \xi^2(k) + \frac{i}{\alpha} \in \mathbb{C} \setminus \mathbb{R}_-, \quad \nu(k) = \sqrt{\mu(k)}.$$

Since $J^2 = -I$, we can expand all matrices on the basis (I, J) , starting with

$$N(k) = \operatorname{Re} \nu(k)I + \operatorname{Im} \nu(k)J.$$

In what follows, if no confusion is to be feared, we will omit the dependency in the k variable. In (7), the trigonometric functions $\cos Nt$ and $\sin Nt$ are defined by their Taylor series, and the derivatives are calculated like in the real scalar analysis, for instance $\frac{d}{dt} \cos Nt = -N \sin Nt$. Introducing the complex functions $c(t) = \cos(\nu t)$, $s(t) = \sin(\nu t)$, we can write

$$\cos Nt = \operatorname{Re} c(t)I + \operatorname{Im} c(t)J, \quad \sin Nt = \operatorname{Re} s(t)I + \operatorname{Im} s(t)J.$$

A long computation (see [7, Theorem 3.2.2]) shows that the convergence of the algorithm is defined by the convergence matrix $M = M_1 M_2$, with $M_1 = \tilde{S}(\delta + T_1)^{-1} S(\delta - T_2)$ and $M_2 = -\tilde{S}^{-1}(-T_2) S(T_1)$, $T_2 = T - T_1$, and

$$\tilde{S}(t) = \begin{pmatrix} \operatorname{Re} c(t) & \operatorname{Re} \nu^{-1} s(t) \\ -\operatorname{Re} \nu s(t) & \operatorname{Re} c(t) \end{pmatrix}, \quad S(t) = \begin{pmatrix} \operatorname{Im} c(t) & \operatorname{Im} \nu^{-1} s(t) \\ -\operatorname{Im} \nu s(t) & \operatorname{Im} c(t) \end{pmatrix}.$$

Trigonometric identities imply that the determinant of the matrix $\tilde{S}(t)$ is $d(t) = \frac{\nu_1^2 \cosh^2(\nu_2 t) + \nu_2^2 \cos^2(\nu_1 t)}{|\nu|^2}$ with $\nu_1 = \operatorname{Re} \nu$ and $\nu_2 = \operatorname{Im} \nu$, which shows that the matrix \tilde{S} is invertible, and gives easily the inverse of \tilde{S} . We have not been able yet to compute the eigenvalues of M in the general case. However we have the important general result:

Theorem 1 *For any times T and $T_1 < T$, for any overlap $\delta \geq 0$, for large k , the iteration matrix is equivalent to $\frac{T_1(T-T_1-\delta)}{(2\alpha\xi)^2} I$. Therefore the algorithm is a smoother: it damps large frequencies.*

Proof. We expand all functions for large ξ which is asymptotically equivalent to k . Define the two matrices

$$R_1(t) = \begin{pmatrix} \cos \xi t & -\frac{1}{\xi} \sin \xi t \\ \xi \sin \xi t & \cos \xi t \end{pmatrix}, \quad R_2(t) = \begin{pmatrix} \sin \xi t & -\frac{1}{\xi} \cos \xi t \\ \xi \cos \xi t & \sin \xi t \end{pmatrix}.$$

A Taylor expansion shows that for fixed t , for large ξ ,

$$\tilde{S}^{-1}(t) \sim R_1(t), \quad S(t) \sim -\frac{t}{2\alpha\xi} R_2(t).$$

Furthermore $R_1(t)R_2(t') = R_2(t' - t)$. Therefore

$$M_1 \sim \frac{T_2 - \delta}{2\alpha\xi} R_2(-T), \quad M_2 \sim -\frac{T_1}{2\alpha\xi} R_2(T).$$

It is now easy to see that $R_2(t)R_2(-t) = -I$, which gives the result. \square

In a simpler case, we were able to calculate explicitly the eigenvalues of the iteration matrix M :

Theorem 2 *Suppose $\delta = 0$ and $T_1 = T/2$. Then the eigenvalues of M are*

$$\begin{aligned} \zeta^\pm(k, T_1) &= -\frac{1}{d^2(T_1)} \left[\sqrt{\varphi} \pm \frac{1}{4|v|^2} (\cosh(2v_2 T_1) - \cos(2v_1 T_1)) \right]^2, \\ \varphi(k, T_1) &= \frac{1}{4|v|^2} \left[v_1^2 \sinh^2(2v_2 T_1) - v_2^2 \sin^2(2v_1 T_1) \right]. \end{aligned} \quad (8)$$

Consequently, the spectral radius of the matrix M is

$$\rho(k, T_1) = \left(\frac{\sqrt{\varphi} + \frac{1}{4|v|^2} (\cosh(2v_2 T_1) - \cos(2v_1 T_1))}{d(T_1)} \right)^2. \quad (9)$$

The proof is long and technical and can be found in [7, Theorem 3.3.1]. The domain of time in which the algorithm converges has not been found yet. However we have two existence results.

Theorem 3 *There exists $T_0 > 0$ such that, for any $T < T_0$, the algorithm converges.*

Theorem 4 *For fixed k , as T tends to infinity,*

$$\max_k \rho(k, T_1) \sim \left(\frac{|v|}{v_1} + \frac{1}{2v_1^2} \right)^2 > 1. \quad (10)$$

Therefore there exists a time \bar{T} such that the algorithm diverges for $T > \bar{T}$.

Theorems 3 and 4 are illustrated by Figure 1. On the left panel, we plot ρ as a function of ξ (see (6), ξ is proportional to k) and T . We can see that the convergence factor is very small for large ξ , but can be larger than one for low frequencies when T is large. On the right panel, we plot the function $T \mapsto \max_{k \in \mathbb{R}^+} \rho(T, \xi)$. The red curve corresponds to a numerical convergence factor computed by running a finite differences code. These two convergence factors are in good agreement. In the two cases, the convergence factor becomes larger than one for large T .

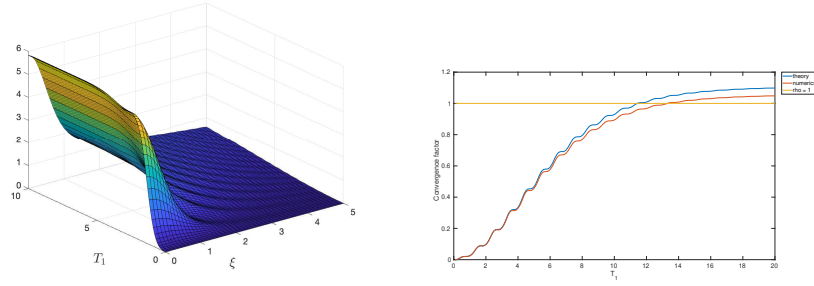


Fig. 1 ρ as a function of ξ and T_1 (left), Convergence factor with respect to T_1 (right)

4 Inherited relaxation algorithm

Suppose $\delta = 0$ and $T_1 = T/2$. We introduce new unknowns $(\underline{\lambda}^n, \underline{\lambda}'^n)$ and modify the iterates as follows:

- Step 1, domain 1: $\lambda_1^{n+1}(T_1) = \underline{\lambda}^n$, $\partial_t \lambda_1^{n+1}(T_1) = \underline{\lambda}'^n$,
- Step 2, domain 2: $y_2^{n+1}(T_1) = y_1^{n+1}(T_1)$, $\partial_t y_2^{n+1}(T_1) = \partial_t y_1^{n+1}(T_1)$,
- Step 3, relaxation:

$$(\underline{\lambda}^{n+1}, \underline{\lambda}'^{n+1}) = \theta(\lambda_2^{n+1}, \partial_t \lambda_2^{n+1}) + (1 - \theta)(\underline{\lambda}^n, \underline{\lambda}'^n).$$

Theorem 5 Let $\theta^* = \frac{2}{1 + \max_{k \in [k_{\min}, k_{\max}]} |\zeta_-(k, T_1)|}$, where $\zeta_-(k, T_1)$ is the eigenvalue of the iteration matrix M defined in (8). For any $\theta \in (0, \theta^*)$, the relaxed inherited algorithm converges. In particular, if the algorithm converges without relaxation, the relaxed algorithm converges better for any $\theta \in (0, 1)$.

For the proof see [7, Theorem 4.2.1]. In Figure 2, the limit relaxation parameter θ^* is denoted with a star. The set of θ such that the algorithm converges is larger for $T = 1$ than for $T = 10$. Furthermore there is an optimal parameter, represented by a circle.

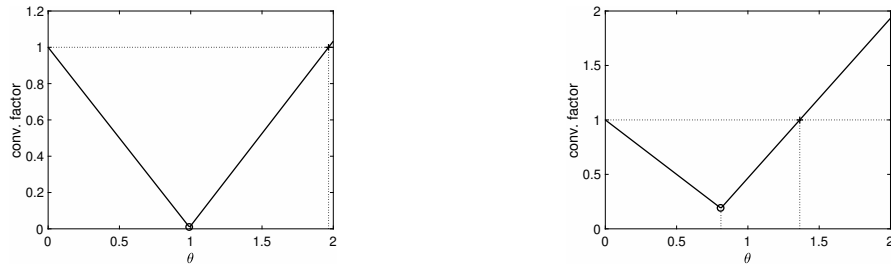


Fig. 2 Convergence factor with respect to θ (Left : $T = 1$, Right : $T = 10$)

5 Dirichlet-Neumann algorithm

The Dirichlet-Neumann algorithm is a nonoverlapping algorithm, invented by Bjorstad and Widlund in 1986 [1] for elliptic problems. The symmetry of the subdomains plays an important role, this is why we translate the intervals in time. The time interval for the control problem is $(-T_1, T_2)$, and will be divided into $(-T_1, 0)$ and $(0, T_2)$. Motivated by the paper by Gander and Yongxiang Liu [3], we create the Dirichlet Neumann algorithm as follows. We solve the optimality system by changing the transmission conditions in (5). The sequence $G^n(x)$ is define recursively by

- Step 1, domain 1: final transmission conditions in $(-T_1, 0)$:

$$(y_1^{n+1}, \lambda_1^{n+1})(\cdot, 0) = G^n.$$

- Step 2, domain 2: initial transmission conditions in $(0, T_2)$:

$$(\partial_t y_2^{n+1}, \partial_t \lambda_2^{n+1})(\cdot, \cdot) = (\partial_t y_1^{n+1}, \partial_t \lambda_1^{n+1})(\cdot, \cdot)$$

- Step 3, relaxation step:

$$G^{n+1} = \theta(y_2^{n+1}, \lambda_2^{n+1})(\cdot, 0) + (1 - \theta)G^n.$$

In the inherited algorithm, the well-posedness of the subproblems was given by the fact that they were the optimality systems of convex functions. It does not seem to be the case here, so we must manage differently. We use the sine decomposition defined in Section 3, for the well-posedness and for the convergence too.

Theorem 6 *For $T \leq +\infty$, the subproblems in the intervals are well-posed.*

Proof. We use the sine decomposition defined in Section 3 to decompose the initial or final time problems for each frequency in space. We show that the series are uniformly convergent. For details see [7, Section 4.3.3]. \square

Theorem 7 *For $T_1 = T_2 = +\infty$, the algorithm is convergent for all $\theta \in (0, 1)$. For $T_1 = T_2 < +\infty$, the convergence factor has for large k the asymptotic behavior:*

$$\rho(k, T_1) \sim 4 \frac{\xi^2}{T_1^2}.$$

Therefore the algorithm is divergent for all $\theta \in (0, 1)$.

Proof. For $T_1 = T_2 = +\infty$, it is easy to see that, as in the elliptic case, $G^{n+1} = (1 - 2\theta)G^n$. Therefore for all $\theta \in (0, 1)$, the sequence converges. Furthermore it converges in two iterations for $\theta = \frac{1}{2}$. The behavior is radically different when $T_1 = T_2 < +\infty$. We can obtain explicit formulas for the convergence factor, which imply the asymptotics for fixed T_1 and large ξ :

$$\rho(k, T_1) \sim 4 \frac{\xi^2}{T_1^2}.$$

This shows that for any T_1 , there is a k_0 such that $\rho(k, T_1) > 1$ for $k > k_0$. \square

Remark 2 This is an unfortunate non-existence result, compared to the elliptic case. However, when using the algorithm in a discrete setting, not all the frequencies are present. There is a largest frequency $k_{\max} = \pi/\Delta x$. Then if T_1 is large the convergence factor is smaller than 1, and there is a θ_0 such that for $0 < \theta < \theta_0$, the algorithm converges. We present on Figure 3 the convergence factor of the relaxed algorithm for $T_1 = 25$ and $k_{\max} = 20$. The parameter θ has to be chosen very small to make the algorithm converge: indeed, θ_0 goes to 0 as k_{\max} increases.

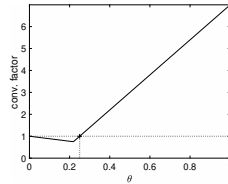


Fig. 3 Convergence factor with respect to θ for the Dirichlet-Neumann algorithm

Remark 3 The Dirichlet-Neumann algorithm is better converging for low frequency, while the inherited algorithm converges better for low frequencies.

In conclusion, we just described promising algorithms and analysis. It still needs to be extended to more general situations. e.g multi-subdomains, 2-D configurations, more general transmission conditions, and used in a two-grid algorithm.

Acknowledgements The authors acknowledge the ANR – FRANCE (French National Research Agency) for its financial support of the project ALLOWAPP ANR-19-CE46-0013-01.

References

1. Bjorstad, P., Widlund, O.: Iterative methods for the solution of elliptic problems on regions partitioned into substructures. *SIAM Journal on Numerical Analysis* **23**(6), 1097–1120 (1986)
2. Chavent, G.: *Nonlinear least squares for inverse problems: theoretical foundations and step-by-step guide for applications*. Springer Science & Business Media (2010)
3. Gander, M.J., Liu, Y.: Is there more than one Dirichlet–Neumann algorithm for the biharmonic problem? *SIAM Journal on Scientific Computing* **43**(3), A1881–A1906 (2021)
4. Kwok, F.: On the time-domain decomposition of parabolic optimal control problems. In: *Domain decomposition methods in science and engineering XXIII. Proceedings of the 23rd international conference, Jeju Island, South Korea, July 6–10, 2015*, pp. 55–67. Cham: Springer (2017)
5. Lagnese, J.E., Leugering, G.: Time-domain decomposition of optimal control problems for the wave equation. *Systems & Control Letters* **48**(3-4), 229–242 (2003)
6. Lions, J.L.: *Optimal control of systems governed by partial differential equations*. Springer (1971)
7. Vuong, D.T.: *Algorithmes de parallélisation en temps pour des problèmes de contrôle optimal de l’équation des ondes*. Phd thesis, University Paris 13 Sorbonne Paris-Nord (2023)