How Does the Partition of Unity Influence SORAS Preconditioner?

Marcella Bonazzoli, Xavier Claeys, Frédéric Nataf, and Pierre-Henri Tournier

1 Introduction

The Symmetrized Optimized Restricted Additive Schwarz (SORAS) preconditioner, first introduced in [8] for the Helmholtz equation and called OBDD-H, was later studied in [6] for generic symmetric positive definite problems and viewed as a symmetric variant of ORAS preconditioner. Its convergence was rigorously analyzed in [5] for the Helmholtz equation, and in [1] we generalized this theory to generic non self-adjoint or indefinite problems. Moreover, as an illustration of our theory, we proved new estimates for the specific case of the heterogeneous reaction-convection-diffusion equation. In the numerical experiments in [1], we noticed that the number of iterations for convergence of preconditioned GMRES appears not to vary significantly when increasing the overlap width. In the present paper, we show that actually this is due to the particular choice of the partition of unity for the preconditioner. The influence of five different kinds of partition of unity on SORAS solver and preconditioner for the Laplace equation has been briefly studied in the conclusion of [4], where the method is named ORASH. Here, for the reaction-convection-diffusion equation, we focus on two kinds of partitions of unity, and study the dependence on the overlap and on the number of subdomains.

Marcella Bonazzoli
Inria, UMA, ENSTA Paris, Institut Polytechnique de Paris, Palaiseau, France,
e-mail: marcella.bonazzoli@inria.fr

Xavier Claeys, Frédéric Nataf, Pierre-Henri Tournier
Sorbonne Université, CNRS, Université Paris Cité, LJLL, Paris, France,
e-mail: xavier.claeys@sorbonne-universite.fr, frederic.nataf@sorbonne-universite.fr,
pierre-henri.tournier@sorbonne-universite.fr

2 SORAS preconditioner and two kinds of partition of unity

Let A denote the $n \times n$ matrix, not necessarily positive definite nor self-adjoint, arising from the discretization of the problem to be solved, posed in an open domain $\Omega \subset \mathbb{R}^d$. Given a set of overlapping open subdomains Ω_j, $j = 1, \dots, N$, such that $\Omega = \bigcup_{j=1}^N \Omega_j$ and each $\overline{\Omega_j}$ is a union of elements of the mesh $\mathcal{T}^h$ of Ω, we consider the set $\mathcal{N}$ of the unknowns on the whole domain, so $\#\mathcal{N} = n$, and its decomposition $\mathcal{N} = \bigcup_{j=1}^N \mathcal{N}_j$ into the non-disjoint subsets corresponding to the different overlapping subdomains $\overline{\Omega}_j \cap \Omega$, with $\#\mathcal{N}_j = n_j$. Denote by δ the width of the overlap between subdomains. The following matrices are then the classical ingredients to define overlapping Schwarz domain decomposition preconditioners (see e.g. [2, §1.3]):

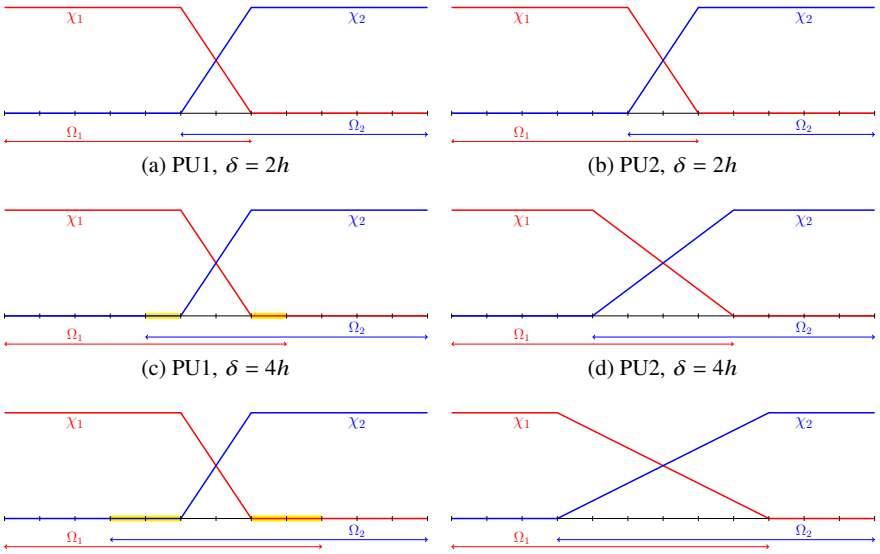
- restriction matrices R_j from Ω to $\overline{\Omega}_j \cap \Omega$, which are $n_j \times n$ Boolean matrices whose (i, i') entry equals 1 if the i-th unknown in $\mathcal{N}_j$ is the i'-th one in $\mathcal{N}$ and vanishes otherwise;
- extension by zero matrices R_j^T from $\overline{\Omega}_j \cap \Omega$ to Ω;
- partition of unity matrices D_j, which are $n_j \times n_j$ diagonal matrices with real non-negative entries such that $\sum_{j=1}^N R_j^T D_j R_j = I$ and which can be seen as matrices that properly weight the unknowns belonging to the overlap between subdomains;
- local matrices B_j, of size $n_j \times n_j$, which arise from the discretization of subproblems posed in $\overline{\Omega}_j \cap \Omega$, with for instance Robin-type or more general absorbing transmission conditions on the interfaces $\partial\Omega_j \setminus \partial\Omega$.

Then the one-level Symmetrized Optimized Restricted Additive Schwarz (SORAS) preconditioner is defined as

$$M^{-1} \coloneqq \sum_{j=1}^{N} R_j^T D_j B_j^{-1} D_j R_j. \tag{1}$$

Note that M^{-1} is not self-adjoint when B_j is not self-adjoint, even if we maintain the SORAS name, where S stands for 'Symmetrized'. In fact, this denomination was introduced in [6] for symmetric positive definite problems, since in that case SORAS preconditioner is a symmetric variant of ORAS preconditioner $\sum_{j=1}^N R_j^T D_j B_j^{-1} R_j$. Thus, the adjective 'Symmetrized' stands for the presence of the rightmost partition of unity D_j. We recall that 'Restricted' indicates the presence of the leftmost partition of unity D_j and that 'Optimized' refers to the choice of transmission conditions other than standard Dirichlet conditions in the local matrices B_j.

Here we focus on the influence exerted by the choice of partition of unity matrices D_j on the convergence of GMRES preconditioned by (1). Indeed, several definitions of the diagonal matrices D_j are possible to ensure property $\sum_{j=1}^N R_j^T D_j R_j = I$. In general, the diagonals of the D_j can be constructed by the interpolation of continuous partition of unity functions $\chi_j : \Omega \to [0, 1]$, $j = 1, \dots, N$: $\sum_{j=1}^N \chi_j = 1$ in $\overline{\Omega}$, and $\mathrm{supp}(\chi_j) \subset \Omega_j$, so in particular χ_j is zero on the subdomain interfaces $\partial\Omega_j \setminus \partial\Omega$.



(a) PU1, $\delta = 2h$ (b) PU2, $\delta = 2h$

(c) PU1, $\delta = 4h$ (d) PU2, $\delta = 4h$

(e) PU1, $\delta = 6h$ (f) PU2, $\delta = 6h$

Fig. 1 Illustration in a one-dimensional two-subdomain case of the two kinds of partition of unity functions $\chi_j : \Omega \to [0, 1]$ (PU1 on the left and PU2 on the right), with increasing width of the overlap δ from top to bottom.

In addition, in the case of ORAS fixed-point iterative solver, also the first derivatives of χ_j are required to be equal to zero on $\partial\Omega_j \setminus \partial\Omega$, because this property ensures that the continuous version of ORAS solver is equivalent to Lions' algorithm, see e.g. [2, §2.3.2] for a particular model problem. An instructive calculation for a simple one- (and two-) dimensional problem, which shows an analogous equivalence property for RAS solver, is given in [3]; a more general equivalence result for ORAS solver is proved in [10, Theorem 3.4]. This first choice of Partition of Unity (PU1), where the gradient of χ_j is zero on the subdomain interfaces $\partial\Omega_j \setminus \partial\Omega$, is illustrated in a one-dimensional two-subdomain case in Figure 1, left, and starting from an overlap $\delta = 4h$. Note that PU1 in Figure 1 is actually different from the original RAS/ORAS partition of unity, which is defined for any overlap size δ multiple of h, but essentially just at the discrete level, and takes only the values 0 or 1; in the original RAS/ORAS articles, the D_j are indeed hidden inside the definition of special extension matrices $\widetilde{R}_j^T$ related to an auxiliary non-overlapping partition of the domain (see e.g. [3, 10] and references therein). However, since the PU1 functions χ_j in Figure 1 are symmetrical to each other, defining the D_j by interpolation of the χ_j is more practical for a parallel implementation.

A second kind of Partition of Unity (PU2) is illustrated in Figure 1, right, where the χ_j functions are different from zero in the interior of the whole overlapping region. This choice is motivated by the fact that using PU1 for SORAS preconditioner can hinder the communication of information between subdomains since in (1) the matrix D_j is also applied before B_j^{-1}, that is before the local problem solve. Indeed,

the numerical experiments performed in [1], where PU1 was used, show that the number of iterations for convergence of preconditioned GMRES does not vary significantly when increasing the overlap size (see also Tables 1,2,3 in Section 4).

3 Definition of the model problem

As in the second part of [1], we consider the heterogeneous reaction-convection-diffusion problem in conservative form:

$$\begin{cases} c_0 u + \operatorname{div}(\mathbf{a}u) - \operatorname{div}(\nu \nabla u) = f & \text{in } \Omega, \\ u = 0 & \text{on } \Gamma, \end{cases} \tag{2}$$

where $\Omega \subset \mathbb{R}^d$ is an open bounded polyhedral domain, $\Gamma = \partial\Omega$, $\mathbf{n}$ is the outward-pointing unit normal vector to Γ, $c_0 \in \mathrm{L}^\infty(\Omega)$, $\mathbf{a} \in \mathrm{L}^\infty(\Omega)^d$, $\operatorname{div}\mathbf{a} \in \mathrm{L}^\infty(\Omega)$, $\nu \in \mathrm{L}^\infty(\Omega)$, $f \in \mathrm{L}^2(\Omega)$ and all quantities are real-valued. We denote $\tilde{c} \coloneqq c_0 + \operatorname{div}\mathbf{a}/2$, and suppose that there exist $\tilde{c}_- > 0$, $\tilde{c}_+ > 0$ such that

$$\tilde{c}_- \le \tilde{c}(\mathbf{x}) \le \tilde{c}_+ \text{ a.e. in } \Omega, \tag{3}$$

and that there exist $\nu_- > 0$, $\nu_+ > 0$ such that $\nu_- \le \nu(\mathbf{x}) \le \nu_+$ a.e. in Ω. The variational formulation of problem (2) is (see e.g. [1, §4]): find $u \in \mathrm{H}^1_0(\Omega)$ such that

$$a(u, v) = F(v), \quad \text{for all } v \in \mathrm{H}^1_0(\Omega), \tag{4}$$

$$a(u, v) \coloneqq \int_\Omega \Big(\tilde{c} u v + \frac{1}{2}\mathbf{a} \cdot \nabla u \, v - \frac{1}{2} u \, \mathbf{a} \cdot \nabla v + \nu \nabla u \cdot \nabla v \Big), \quad F(v) \coloneqq \int_\Omega f v.$$

On each subdomain we consider the local problem with bilinear form

$$a_j(u, v) \coloneqq \int_{\Omega_j} \Big(\tilde{c} u v + \frac{1}{2}\mathbf{a} \cdot \nabla u \, v - \frac{1}{2} u \, \mathbf{a} \cdot \nabla v + \nu \nabla u \cdot \nabla v \Big) + \int_{\partial\Omega_j \setminus \Gamma} \alpha u v,$$

where we impose an absorbing transmission condition on the subdomain interface $\partial\Omega_j \setminus \partial\Omega$ given by $\alpha(\mathbf{x}) = \sqrt{(\mathbf{a} \cdot \mathbf{n})^2 + 4 c_0 \nu}/2$ (see e.g. [7]).

4 Numerical experiments

We simulate problem (4) with Ω a rectangle $[0, N \cdot 0.2] \times [0, 0.2]$, where N is the number of subdomains. In Tables 1,2,3 we take $N = 5$ and

$$f = 100 \exp\{-10((x - 0.5)^2 + (y - 0.1)^2)\}.$$

In Table 4, we test weak scaling by varying N, with

$$f = 100 \exp\{-10((x-0.1)^2 + (y-0.1)^2)\}.$$

The problem is discretized by piece-wise linear Lagrange finite elements on a uniform triangular mesh with 60 nodes on the vertical side of the rectangle and $N \cdot 60$ nodes on the horizontal one, resulting in 18361 degrees of freedom for $N = 5$, and 7381,

Table 1 Iteration numbers for SORAS preconditioner ($N = 5$).

	#PU1(PU2)			
$\mathbf{a} = 2\pi[-(y-0.1), (x-0.5)]^T$	$\delta = 2h$	$\delta = 4h$	$\delta = 6h$	$\delta = 8h$
$c_0 = 1,\ \nu = 1$	21(21)	20(17)	20(15)	19(14)
$c_0 = 1,\ \nu = 0.001$	14(14)	13(11)	12(11)	12(10)
$c_0 = 0.001,\ \nu = 1$	21(21)	20(18)	20(15)	19(14)
$c_0 = 0.001,\ \nu = 0.001$	15(15)	14(12)	13(11)	13(11)

Table 2 Repeat of Table 1 but with $\mathbf{a} = [-x, -y]^T$. In this case $\operatorname{div}\mathbf{a} = -2$ is negative and $\tilde{c} = c_0 - 1$ does not verify condition (3).

	#PU1(PU2)			
$\mathbf{a} = [-x, -y]^T$	$\delta = 2h$	$\delta = 4h$	$\delta = 6h$	$\delta = 8h$
$c_0 = 1,\ \nu = 1$	21(21)	21(19)	20(17)	20(15)
$c_0 = 1,\ \nu = 0.001$	16(16)	16(14)	16(13)	16(13)
$c_0 = 0.001,\ \nu = 1$	22(22)	22(19)	22(17)	21(16)
$c_0 = 0.001,\ \nu = 0.001$	17(17)	16(15)	16(14)	16(13)

Table 3 Repeat of Table 1 but with $\mathbf{a} = [1, 0]^T$ and with Streamline Upwind Petrov-Galerkin stabilization for the Galerkin approximation.

	#PU1(PU2)			
$\mathbf{a} = [1, 0]^T$	$\delta = 2h$	$\delta = 4h$	$\delta = 6h$	$\delta = 8h$
$c_0 = 1,\ \nu = 1$	20(20)	20(18)	20(16)	20(15)
$c_0 = 1,\ \nu = 0.001$	11(11)	11(12)	11(12)	11(12)
$c_0 = 0.001,\ \nu = 1$	20(20)	20(18)	20(16)	20(15)
$c_0 = 0.001,\ \nu = 0.001$	12(12)	12(12)	12(13)	12(12)

Table 4 Iteration numbers in a weak scaling test ($\delta = 4h$).

	#PU1(PU2)					
$\mathbf{a} = [1, 0]^T$	$N = 2$	$N = 4$	$N = 8$	$N = 16$	$N = 32$	$N = 64$
$c_0 = 1,\ \nu = 1$	18(15)	23(20)	28(24)	35(28)	36(29)	36(29)
$c_0 = 1,\ \nu = 0.001$	8(8)	10(12)	16(16)	23(24)	37(37)	63(61)
$c_0 = 0.001,\ \nu = 1$	18(15)	23(20)	29(25)	35(29)	36(29)	36(29)
$c_0 = 0.001,\ \nu = 0.001$	8(8)	10(12)	16(17)	24(25)	40(40)	71(71)

14701, 29341, 58621, 117181, 234301 degrees of freedom for $N = 2, 4, 8, 16, 32, 64$ respectively. The domain is partitioned into N vertical strips, then each subdomain is augmented with mesh elements layers of size $\delta/2$ to obtain the overlapping decomposition: the total width of the overlap between two subdomains is then δ. In particular, for $\delta = 2h, 4h, 6h, 8h$ the ratio between the subdomain width ($60h$) and δ is equal to 30, 15, 10, 7.5. We use GMRES with right preconditioning, with a zero initial guess in Tables 1,2,3 and a random initial guess in Table 4. The stopping criterion is based on the relative residual, with a tolerance of 10^{-6}. To apply the preconditioner, the local problems in each subdomain are solved with the direct solver MUMPS[1]. All the computations are done in the `ffddm` framework [11] of FreeFEM[2].

We compare the number of iterations for convergence (denoted by # in the tables) using the two kinds of partition of unity: the results for PU1 were also included in [1] and the results for PU2 are reported inside brackets in Tables 1–4. In `ffddm` framework, the first partition of unity is selected by the flag `-raspart`, while the second partition of unity is the one used by default.

As in [1], we examine several configurations for the coefficients in (2). First, in Table 1 we consider a rotating convection field $\mathbf{a} = 2\pi[-(y - 0.1), (x - 0.5)]^T$ and small/large values for the reaction coefficient c_0 and the viscosity ν. We can see that a larger overlap helps the convergence of the preconditioner, especially with PU2, while with PU1 the number of iterations does not vary significantly. Moreover, with both kinds of partition of unity, the number of iterations appears not very sensitive to the reaction coefficient c_0, while it increases when the viscosity ν is larger.

Then, in Table 2 we take $\mathbf{a} = [-x, -y]^T$, which has negative divergence $\operatorname{div} \mathbf{a} = -2$, to test the robustness of the method when condition (3) on the positiveness of $\tilde{c}$ is violated: in this case, $\tilde{c} = c_0 - 1$, so $\tilde{c} = 0$, $\tilde{c} = -0.999$ for $c_0 = 1$, $c_0 = 0.001$ respectively. We can still observe a convergence behavior similar to the one of Table 1.

Finally, in Table 3 we consider a horizontal convection field $\mathbf{a} = [1, 0]^T$, which is normal to the interfaces between subdomains. Since in this case non-physical numerical instabilities appear in the solution, we stabilize the discrete variational formulation using the Streamline Upwind Petrov-Galerkin (SUPG) method (see for instance [9, §11.8.6]). In this configuration for the convection field, for low viscosity $\nu = 0.001$ the dependence of the iteration number on the overlap size δ appears to be not significant, even with PU2.

Again in this third configuration with $\mathbf{a} = [1, 0]^T$ and SUPG stabilization, we perform a weak scaling test by taking $\Omega = [0, N \cdot 0.2] \times [0, 0.2]$ for increasing number of subdomains N, and $\delta = 4h$. We can see that especially in the cases with low viscosity $\nu = 0.001$, convergence deteriorates with N, as expected since we are testing a one-level preconditioner.

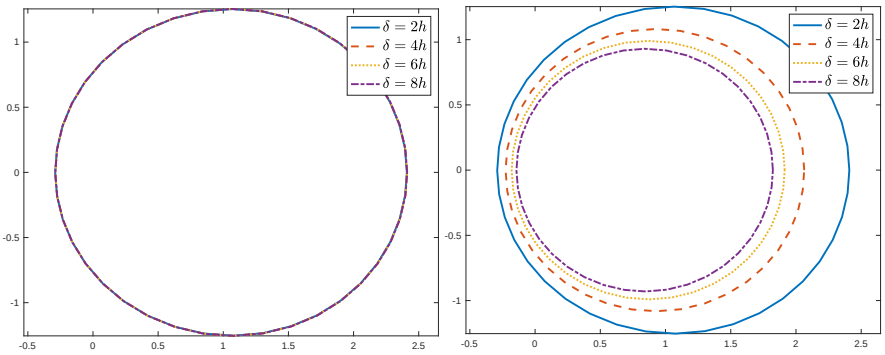
In summary, our numerical investigation shows that, for the considered SORAS preconditioner, PU2 generally improves the iteration counts obtained with PU1. Moreover, the first kind of partition of unity (PU1), which would be the natural choice for ORAS solver instead, yields for SORAS preconditioner iterations counts

[1] `http://mumps.enseeiht.fr/`

[2] `https://freefem.org/`

Table 5 Minimum and maximum eigenvalues of the preconditioned operator.

	PU1(PU2)			
$\mathbf{a} = [0,0]^T$	$\delta = 2h$	$\delta = 4h$	$\delta = 6h$	$\delta = 8h$
$\lambda_{\min}$	0.50 (0.50)	0.50 (0.50)	0.50 (0.50)	0.50 (0.50)
$\lambda_{\max}$	11.25 (11.25)	10.61 (5.98)	10.07 (4.01)	9.60 (3.02)



(a) PU1 (b) PU2

Fig. 2 Numerical range of the preconditioned operator ($\mathbf{a} = [-x, -y]^T$).

that do not vary significantly when increasing the overlap width, whereas using the second kind of partition of unity (PU2) a larger overlap gives faster convergence.

To conclude, we wish to provide a deeper explanation of the observed effects. First, we examine the symmetric positive definite case, with $\mathbf{a} = [0,0]^T$, $c_0 = 1$, $\nu = 1$ (so $\tilde{c} = c_0 > 0$), and report in Table 5 the largest and smallest eigenvalues of the preconditioned operator. We take $N = 2$ and 40 nodes on the vertical side of the rectangle, $2 \cdot 40$ nodes on the horizontal one. Note that SORAS preconditioner for generic symmetric positive definite problems was analyzed in [6], but no explicit discussion about the influence of the partition of unity was included there. On the one hand, the largest eigenvalue of the preconditioned operator is controlled by the modes of the local generalized eigenvalue problems defined in [6, Definition 3.1], where the partition of unity matrices appear in the local operator on the left-hand side: Table 5 shows that indeed $\lambda_{\max}$ is smaller for PU2, which is less steep than PU1, especially when increasing the overlap width δ (see Fig. 1). Moreover, with PU1, the dependence of $\lambda_{\max}$ on δ is much less significant than with PU2. On the other hand, the smallest eigenvalue of the preconditioned operator is controlled by the modes of the local generalized eigenvalue problems defined in [6, Definition 3.2], where the partition of unity is not involved: in Table 5 we can see that indeed $\lambda_{\min}$ is independent of the partition of unity.

For the non-symmetric case, with $\mathbf{a} = [-x, -y]^T$, $c_0 = 0.001$, $\nu = 0.001$ (so $\tilde{c} = c_0 - 1 < 0$), we plot in Fig. 2 the contour of the numerical range of the preconditioned operator for overlap widths that range from $\delta = 2h$ to $\delta = 8h$, for the two types of partition of unity. We can remark that for PU1 (Fig. 2, left) the numerical ranges practically coincide for the different overlap widths, whereas for PU2 (Fig. 2, right) the numerical range gets smaller for larger overlap width. This explains the more favorable convergence properties of preconditioned GMRES with PU2 when increasing the overlap width, and the much less significant influence of the overlap in the case of PU1.

References

1. Bonazzoli, M., Claeys, X., Nataf, F., and Tournier, P.-H. Analysis of the SORAS domain decomposition preconditioner for non-self-adjoint or indefinite problems. *Journal of Scientific Computing* **89**(1), 19 (2021).
2. Dolean, V., Jolivet, P., and Nataf, F. *An introduction to domain decomposition methods: algorithms, theory and parallel implementation*. SIAM, Philadelphia, PA (2015).
3. Efstathiou, E. and Gander, M. J. Why Restricted Additive Schwarz converges faster than Additive Schwarz. *BIT Numerical Mathematics* **43**(5), 945–959 (2003).
4. Gander, M. J. Does the partition of unity influence the convergence of Schwarz methods? In: *Domain Decomposition Methods in Science and Engineering XXV*, 3–15. Springer International Publishing, Cham (2020).
5. Graham, I. G., Spence, E. A., and Zou, J. Domain Decomposition with Local Impedance Conditions for the Helmholtz Equation with Absorption. *SIAM J. Numer. Anal.* **58**(5), 2515–2543 (2020).
6. Haferssas, R., Jolivet, P., and Nataf, F. A robust coarse space for optimized Schwarz methods: SORAS-GenEO-2. *C. R. Math. Acad. Sci. Paris* **353**(10), 959–963 (2015).
7. Japhet, C., Nataf, F., and Rogier, F. The optimized order 2 method: Application to convection–diffusion problems. *Future Generation Computer Systems* **18**(1), 17–30 (2001).
8. Kimn, J.-H. and Sarkis, M. Restricted overlapping balancing domain decomposition methods and restricted coarse problems for the Helmholtz problem. *Comput. Methods Appl. Mech. Engrg.* **196**(8), 1507–1514 (2007).
9. Quarteroni, A. M. *Numerical models for differential problems*, vol. 2. Springer (2009).
10. St-Cyr, A., Gander, M. J., and Thomas, S. J. Optimized Multiplicative, Additive, and Restricted Additive Schwarz preconditioning. *SIAM Journal on Scientific Computing* **29**(6), 2402–2425 (2007).
11. Tournier, P.-H. and Nataf, F. FFDDM: FreeFEM Domain Decomposition Methods. https://doc.freefem.org/documentation/ffddm/index.html (2019).