# New time domain decomposition methods for parabolic optimal control problems II: Neumann-Neumann algorithms

Martin Jakob Gander[1], Liu-Di LU[1]

[1]*Section of Mathematics, University of Geneva, Rue du Conseil Général 7-9, 1205 Geneva, Switzerland*

**Abstract**

We present new Neumann-Neumann algorithms based on a time domain decomposition applied to unconstrained parabolic optimal control problems. After a spatial semi-discretization, the Lagrange multiplier approach provides a coupled forward-backward optimality system, which can be solved using a time domain decomposition. Due to the forward-backward structure of the optimality system, nine variants can be found for the Neumann-Neumann algorithms. We analyze their convergence behavior and determine the optimal relaxation parameter for each algorithm. Our analysis reveals that the most natural algorithms are actually only good smoothers, and there are better choices which lead to efficient solvers. We illustrate our analysis with numerical experiments.

**Keywords:** Time domain decomposition, Neumann-Neumann algorithm, Parallel in Time, Parabolic optimal control problems, Convergence analysis.

**MSCcodes:** 65M12, 65M55, 65Y05,

## 1 Introduction

As our model problem, we consider a parabolic optimal control problem: for a given target function $\hat{y} \in L^2(Q)$, $\gamma > 0$ and $\nu \geq 0$, we want to minimize the cost functional

$$J(y,u) := \frac{1}{2}\|y-\hat{y}\|^2_{L^2(Q)} + \frac{\gamma}{2}\|y(T)-\hat{y}(T)\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{U_{\mathrm{ad}}}, \tag{1}$$

subject to the linear parabolic state equation:

$$\begin{aligned} \partial_t y - \Delta y &= u && \text{in } Q := \Omega \times (0,T), \\ y &= 0 && \text{on } \Sigma := \partial\Omega \times (0,T), \\ y(0) &= y_0 && \text{on } \Sigma_0 := \Omega \times \{0\}, \end{aligned} \tag{2}$$

where $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$ is a bounded domain with boundary $\partial\Omega$, and $T$ is the fixed final time. The control $u$ on the right-hand side of the PDE is in an admissible set $U_{\text{ad}}$, and we want to control the solution of the parabolic PDE (2) toward a target state $\hat{y}$. For simplicity, we consider homogeneous boundary conditions. The parabolic optimal control problem (1)-(2) leads to necessary first-order optimality conditions (see e.g., [28, 30]), which include a forward in time primal state equation (2), a backward in time dual state equation,

$$\begin{aligned} \partial_t \lambda + \Delta\lambda &= y - \hat{y} && \text{in } Q, \\ \lambda &= 0 && \text{on } \Sigma, \\ \lambda(T) &= -\gamma(y(T) - \hat{y}(T)) && \text{on } \Sigma_T := \Omega \times \{T\}, \end{aligned} \tag{3}$$

and an algebraic equation $\lambda = \nu u$ with $\lambda$ the dual state. This forward-backward system cannot be solved by standard time-stepping methods, and has to be solved either iteratively or at once. Solving at once the space-time discretized system can be challenging, especially for spatial dimension larger than one. To overcome this challenge, one can use gradient type methods by solving sequentially forward-backward systems [20, 30]. Multigrid methods [1, 4, 17, 27], tensor product techniques [5, 16, 23, 31], model order reduction [2, 21, 22, 24], can also be applied to solve such problems. Since the role of the time variable in forward-backward optimality systems is key, it is natural to seek efficient solvers through Parallel-in-time techniques. This includes, waveform relaxation [26, 18], Parareal [29], PITA [9], PFASST [6], MGRIT [7], see also the survey paper [11]. Application of such techniques to treat parabolic optimal control problems can be found in [8, 13, 15, 19].

In [14], we considered a new time domain decomposition approach motivated by [12, 25], and analyzed the convergence behavior of Dirichlet-Neumann and Neumann-Dirichlet algorithms within this framework. We have surprisingly discovered different variants of Dirichlet-Neumann and Neumann-Dirichlet algorithms for the parabolic optimal control problem (1)-(2), when decomposing in time. This is mainly due to the forward-backward structure of the optimality system. The present paper is the sequel of [14]: our goal is to investigate Neumann-Neumann techniques [3] in the context of time domain decomposition and analyze their convergence behavior. We consider a semi-discretization in space and focus on the time variable. This consists in replacing the spatial operator $-\Delta$ by a matrix $A \in \mathbb{R}^{n\times n}$, for instance using a Finite Difference discretization in space. If $A$ is symmetric, which is natural for discretizations of $-\Delta$, then it can be diagonalized with $A = PDP^T$, and the diagonalized system reads,

$$\begin{cases} \begin{pmatrix} \dot{z}_i \\ \dot{\mu}_i \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} z_i \\ \mu_i \end{pmatrix} = \begin{pmatrix} 0 \\ -\hat{z}_i \end{pmatrix} \text{ in } (0,T), \\ \qquad\qquad z_i(0) = z_{i,0}, \\ \qquad\qquad \mu_i(T) + \gamma z_i(T) = \gamma \hat{z}_i(T), \end{cases} \tag{4}$$

where $d_i$ is the $i$-th eigenvalue of the matrix $A$, and $z_i$, $\mu_i$ as well as $\hat{z}_i$ are the $i$-th components of the vectors $\boldsymbol{z}$, $\boldsymbol{\mu}$ and $\hat{\boldsymbol{z}}$. Eliminating $\mu_i$ in (4), we obtain the

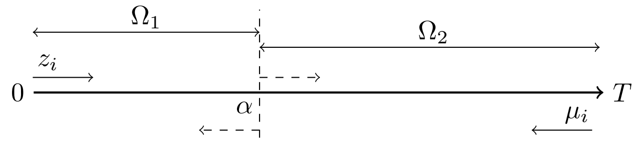


Figure 1: Illustration of the forward-backward system.

second-order ODE

$$\begin{cases} \ddot{z}_i - (d_i^2 + \nu^{-1})z_i = -\nu^{-1}\hat{z}_i \text{ in } (0,T), \\ z_i(0) = z_{i,0}, \\ \dot{z}_i(T) + (\nu^{-1}\gamma + d_i)z_i(T) = \nu^{-1}\gamma\hat{z}_i(T). \end{cases} \tag{5}$$

We refer to [14, Section 2] for more details about the transition from the PDE-constrained problem (1)-(2) to the diagonalized reduced problem (4).

The rest of the paper is structured as follows. We introduce in Section 2 our new time decomposed Neumann-Neumann algorithms and study their convergence behavior in Section 3. Numerical experiments are shown in Section 4 to support our analysis, and we draw conclusions in Section 5.

## 2 Neumann-Neumann algorithms

In this section, we apply the Neumann-Neumann technique (NN) in time to obtain our new time domain decomposition methods to solve the system (4), and investigate their convergence behavior. To focus on the error equation, we set both the initial condition $\boldsymbol{y}_0 = 0$ (i.e., $\boldsymbol{z}_0 = 0$) and the target function $\hat{\boldsymbol{y}} = 0$ (i.e., $\hat{\boldsymbol{z}} = 0$). We decompose the time domain $\Omega := (0,T)$ into two non-overlapping subdomains $\Omega_1 := (0,\alpha)$ and $\Omega_2 := (\alpha,T)$, where $\alpha$ is the interface. And we denote by $z_{j,i}$ and $\mu_{j,i}$ the restriction to $\Omega_j$, $j = 1,2$ of the states $z_i$ and $\mu_i$. Although we will focus on the two-subdomain case in our current study, the results can be extended to $N$ non-overlapping subdomains $\Omega_j := (\alpha_j, \alpha_{j+1})$, $j = 1,\ldots,N$ with $\alpha_1 = 0$ and $\alpha_{N+1} = T$.

Unlike the name of the NN algorithm suggests, it starts first with a Dirichlet step, which will be corrected by a Neumann step and then updates the transmission condition. As the system (4) is a forward-backward system, it appears natural at first glance to keep this property for the decomposed case as illustrated in Figure 2: we expect to have a final condition for the dual state $\mu_{1,i}$ in $\Omega_1$, since we already have an initial condition for $z_{1,i}$; similarly, we expect to have an initial condition for the primal state $z_{2,i}$ in $\Omega_2$, where we already have a final condition for $\mu_{2,i}$. Therefore, for iteration index $k = 1,2,\ldots$, a natural

NN algorithm first solves the Dirichlet step

$$
\begin{aligned}
&\begin{cases}
\begin{pmatrix} \dot{z}_{1,i}^k \\ \dot{\mu}_{1,i}^k \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} z_{1,i}^k \\ \mu_{1,i}^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_1, \\
z_{1,i}^k(0) = 0, \\
\mu_{1,i}^k(\alpha) = f_{\alpha,i}^{k-1},
\end{cases} \\
&\begin{cases}
\begin{pmatrix} \dot{z}_{2,i}^k \\ \dot{\mu}_{2,i}^k \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} z_{2,i}^k \\ \mu_{2,i}^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_2, \\
z_{2,i}^k(\alpha) = g_{\alpha,i}^{k-1}, \\
\mu_{2,i}^k(T) + \gamma z_{2,i}^k(T) = 0,
\end{cases}
\end{aligned}
\tag{6}
$$

then corrects the result by solving the Neumann step

$$
\begin{aligned}
&\begin{cases}
\begin{pmatrix} \dot{\psi}_{1,i}^k \\ \dot{\phi}_{1,i}^k \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} \psi_{1,i}^k \\ \phi_{1,i}^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_1, \\
\psi_{1,i}^k(0) = 0, \\
\dot{\phi}_{1,i}^k(\alpha) = \dot{\mu}_{1,i}^k(\alpha) - \dot{\mu}_{2,i}^k(\alpha),
\end{cases} \\
&\begin{cases}
\begin{pmatrix} \dot{\psi}_{2,i}^k \\ \dot{\phi}_{2,i}^k \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} \psi_{2,i}^k \\ \phi_{2,i}^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_2, \\
\dot{\psi}_{2,i}^k(\alpha) = \dot{z}_{2,i}^k(\alpha) - \dot{z}_{1,i}^k(\alpha), \\
\phi_{2,i}^k(T) + \gamma \psi_{2,i}^k(T) = 0,
\end{cases}
\end{aligned}
\tag{7}
$$

where $\psi_i$ is the primal correction state for $z_i$ and $\phi_i$ the dual correction state for $\mu_i$. Finally, we update the transmission condition by

$$
f_{\alpha,i}^k := f_{\alpha,i}^{k-1} - \theta_1\big(\phi_{1,i}^k(\alpha) + \phi_{2,i}^k(\alpha)\big), \quad g_{\alpha,i}^k := g_{\alpha,i}^{k-1} - \theta_2\big(\psi_{1,i}^k(\alpha) + \psi_{2,i}^k(\alpha)\big), \tag{8}
$$

with two relaxation parameters $\theta_1, \theta_2 > 0$.

As shown in the algorithm (6)-(7), both Dirichlet and Neumann steps have the forward-backward structure. However, this structure only appears as being the natural one at first glance. Indeed, isolating the variable in each equation in the systems (6) and (7), we find the identities

$$
\mu_i = \nu(\dot{z}_i + d_i z_i), \quad z_i = \dot{\mu}_i - d_i \mu_i, \quad \phi_i = \nu(\dot{\psi}_i + d_i \psi_i), \quad \psi_i = \dot{\phi}_i - d_i \phi_i. \tag{9}
$$

To shorten the notation, we define

$$
\sigma_i := \sqrt{d_i^2 + \nu^{-1}}, \quad \omega_i := d_i + \gamma\nu^{-1}, \quad \beta_i := 1 - \gamma d_i. \tag{10}
$$

Using (9) and (10), we can rewrite the Dirichlet step (6) in terms of the primal state $z_i$,

$$
\begin{cases}
\ddot{z}_{1,i}^k - \sigma_i^2 z_{1,i}^k = 0 \text{ in } \Omega_1, \\
z_{1,i}^k(0) = 0, \\
\dot{z}_{1,i}^k(\alpha) + d_i z_{1,i}^k(\alpha) = f_{\alpha,i}^{k-1},
\end{cases}
\qquad
\begin{cases}
\ddot{z}_{2,i}^k - \sigma_i^2 z_{2,i}^k = 0 \text{ in } \Omega_2, \\
z_{2,i}^k(\alpha) = g_{\alpha,i}^{k-1}, \\
\dot{z}_{2,i}^k(T) + \omega_i z_{2,i}^k(T) = 0.
\end{cases}
\tag{11}
$$

Similarly, the Neumann step (7) can be rewritten in terms of the primal correction state $\psi_i$,

$$\begin{cases} \ddot{\psi}_{1,i}^k - \sigma_i^2 \psi_{1,i}^k = 0 \text{ in } \Omega_1, \\ \psi_{1,i}^k(0) = 0, \\ \dot{\psi}_{1,i}^k(\alpha) + \dfrac{\sigma_i^2}{d_i}\psi_{1,i}^k(\alpha) = \big(\dot{z}_{1,i}^k(\alpha) + \dfrac{\sigma_i^2}{d_i} z_{1,i}^k(\alpha)\big) - \big(\dot{z}_{2,i}^k(\alpha) + \dfrac{\sigma_i^2}{d_i} z_{2,i}^k(\alpha)\big), \end{cases}$$
$$\begin{cases} \ddot{\psi}_{2,i}^k - \sigma_i^2 \psi_{2,i}^k = 0 \text{ in } \Omega_2, \\ \dot{\psi}_{2,i}^k(\alpha) = \dot{z}_{2,i}^k(\alpha) - \dot{z}_{1,i}^k(\alpha), \\ \dot{\psi}_{2,i}^k(T) + \omega_i \psi_{2,i}^k(T) = 0, \end{cases} \tag{12}$$

and the transmission condition (8) becomes

$$\begin{aligned} f_{\alpha,i}^k &= f_{\alpha,i}^{k-1} - \theta_1\big(\dot{\psi}_{1,i}^k(\alpha) + d_i\psi_{1,i}^k(\alpha) + \dot{\psi}_{2,i}^k(\alpha) + d_i\psi_{2,i}^k(\alpha)\big), \\ g_{\alpha,i}^k &= g_{\alpha,i}^{k-1} - \theta_2\big(\psi_{1,i}^k(\alpha) + \psi_{2,i}^k(\alpha)\big). \end{aligned} \tag{13}$$

Instead of using (6)-(8) for our analysis, we will use the equivalent formulation in system (11)-(13), in which the forward-backward structure has disappeared. Furthermore, the Dirichlet step in (6) transforms in the primal state $z_i$ to a Robin-Dirichlet (RD) step (11), and the Neumann step in (7) transforms in the primal correction state $\psi_i$ to a Robin-Neumann (RN) step (12). In other words, we analyze actually a RD step with a RN correction, although it is originally a NN algorithm. We could also have interpreted the NN algorithm (6)-(8) using the dual state $\mu_i$ and the dual correction state $\phi_i$, the algorithm would then read differently but the convergence analysis is still the same (see [14]). For the sake of consistency, we keep the interpretation with $z_i$ and $\psi_i$ for all convergence analyses.

The previous transformation reveals that the natural NN algorithm applied to the optimality system (4) is certainly not the only option. Since there are three components in a NN algorithm: a Dirichlet step, a Neumann step and an update step, this expands our options when dealing with parabolic optimal control problems, and provides us with more choices within the NN algorithm. More precisely, instead of applying the Dirichlet step to the pair $(z_i, \mu_i)$, one can also apply it only to the primal state $z_i$ or the dual state $\mu_i$. Likewise, the Neumann step can also be applied only to the primal correction state $\psi_i$ or the dual correction state $\phi_i$. We list in Table 2 all possible new time domain decomposition NN algorithms we can obtain, together with their equivalent interpretations in terms of the states $z_i$ and $\psi_i$. According to the Dirichlet step, they can be classified into three main categories. Each category is composed of two blocks, the first block represents the Dirichlet step and the second block the three possible Neumann steps. And each step contains two rows, the first row is the algorithm applied to (4), and the second row represents the algorithm applied to (5). Note that the update step should also be adapted when modifying the Dirichlet step or the Neumann step. We will further discuss this in the next section, where we investigate the convergence of each algorithm.

Table 1: Variants of the Neumann-Neumann algorithm.

| category | step | $\Omega_1$ | $\Omega_2$ | algorithm type |
|---|---|---|---|---|
| category I: $(z_i, \mu_i)$ | Dirichlet step | $\mu_i$ | $z_i$ | (DD) |
| | | $\dot{z}_i + d_i z_i$ | $z_i$ | (RD) |
| | Neumann step | $\dot{\phi}_i$ | $\dot{\psi}_i$ | (NN) |
| | | $\ddot{\psi}_i + d_i\dot{\psi}_i$ | $\dot{\psi}_i$ | (RN) |
| | | $\dot{\psi}_i$ | $\dot{\psi}_i$ | (NN) |
| | | $\dot{\psi}_i$ | $\dot{\psi}_i$ | (NN) |
| | | $\dot{\phi}_i$ | $\dot{\phi}_i$ | (NN) |
| | | $\ddot{\psi}_i + d_i\dot{\psi}_i$ | $\ddot{\psi}_i + d_i\dot{\psi}_i$ | (RR) |
| category II: $z_i$ | Dirichlet step | $z_i$ | $z_i$ | (DD) |
| | | $z_i$ | $z_i$ | (DD) |
| | Neumann step | $\dot{\psi}_i$ | $\dot{\psi}_i$ | (NN) |
| | | $\dot{\psi}_i$ | $\dot{\psi}_i$ | (NN) |
| | | $\dot{\phi}_i$ | $\dot{\psi}_i$ | (NN) |
| | | $\ddot{\psi}_i + d_i\dot{\psi}_i$ | $\dot{\psi}_i$ | (RN) |
| | | $\dot{\phi}_i$ | $\dot{\phi}_i$ | (NN) |
| | | $\ddot{\psi}_i + d_i\dot{\psi}_i$ | $\ddot{\psi}_i + d_i\dot{\psi}_i$ | (RR) |
| category III: $\mu_i$ | Dirichlet step | $\mu_i$ | $\mu_i$ | (DD) |
| | | $\dot{z}_i + d_i z_i$ | $\dot{z}_i + d_i z_i$ | (RR) |
| | Neumann step | $\dot{\phi}_i$ | $\dot{\phi}_i$ | (NN) |
| | | $\ddot{\psi}_i + d_i\dot{\psi}_i$ | $\ddot{\psi}_i + d_i\dot{\psi}_i$ | (RR) |
| | | $\dot{\phi}_i$ | $\dot{\psi}_i$ | (NN) |
| | | $\ddot{\psi}_i + d_i\dot{\psi}_i$ | $\dot{\psi}_i$ | (RN) |
| | | $\dot{\psi}_i$ | $\dot{\psi}_i$ | (NN) |
| | | $\dot{\psi}_i$ | $\dot{\psi}_i$ | (NN) |

**Remark 1.** *Although most of the algorithms in Table 2 do not look like having the forward-backward structure, it can always be recovered by using the identities in (9). Furthermore, the transmission condition $\ddot{\psi}_i + d_i\dot{\psi}_i$ is actually a Robin type condition, considering the first equation in (12).*

**Remark 2.** *If the order in (6)-(7) is reversed, and one starts with the Neumann step, followed by the Dirichlet correction, the algorithm is then known under the name FETI (Finite Element Tearing and Interconnecting), invented by Farhat and Roux [10]. Since the two algorithms are very much related, we can also find similar variants as in Table 2 in the context of FETI algorithm.*

# 3 Convergence analysis

In this section, we will study the convergence of each algorithm listed in Table 2. Note that the two systems (11) and (12) are very similar, the only difference is in the transmission condition at $\alpha$. We can hence solve these two systems once and for all using the initial and the final condition, and find

$$\begin{aligned} z_{1,i}^k(t) &= A_i^k \sinh(\sigma_i t), \quad z_{2,i}^k(t) = B_i^k \Big(\sigma_i \cosh\big(\sigma_i(T-t)\big) + \omega_i \sinh\big(\sigma_i(T-t)\big)\Big), \\ \psi_{1,i}^k(t) &= C_i^k \sinh(\sigma_i t), \quad \psi_{2,i}^k(t) = D_i^k \Big(\sigma_i \cosh\big(\sigma_i(T-t)\big) + \omega_i \sinh\big(\sigma_i(T-t)\big)\Big). \end{aligned} \tag{14}$$

In general, the solutions (14) remain for all algorithms listed in Table 2, and the coefficients $A_i^k, B_i^k, C_i^k$ and $D_i^k$ will be determined by the transmission conditions. To stay in a compact form, we will only present the modified step for each NN variant instead of giving a complete three-step algorithm.

## 3.1 Category I

This category consists in applying the Dirichlet step to the pair $(z_i, \mu_i)$. As illustrated in Table 2, there are three variants according to the Neumann correction step.

### 3.1.1 Algorithm $\mathrm{NN}_{1a}$

This is (6)-(8), at first glance the most natural NN algorithm, which keeps the forward-backward structure both for the Dirichlet and Neumann steps. To analyze its convergence behavior, we interpret it as (11)-(13) and solve for the exact iterates. Using (14), we determine the coefficients $A_i^k$, $B_i^k$ through the transmission conditions in (11), and find

$$A_i^k = \frac{f_{\alpha,i}^{k-1}}{\sigma_i \cosh(a_i) + d_i \sinh(a_i)}, \quad B_i^k = \frac{g_{\alpha,i}^{k-1}}{\sigma_i \cosh(b_i) + \omega_i \sinh(b_i)}, \tag{15}$$

where we let $a_i := \sigma_i \alpha$ and $b_i := \sigma_i(T-\alpha)$ to simplify the notations, and $a_i + b_i = \sigma_i T$. Using once again (14), we determine the coefficients $C_i^k$, $D_i^k$

through the transmission conditions in (12)

$$C_i^k = A_i^k - B_i^k \nu^{-1} \frac{\sigma_i \gamma \sinh(b_i) + \beta_i \cosh(b_i)}{\sigma_i \sinh(a_i) + d_i \cosh(a_i)},\ D_i^k = A_i^k \frac{\cosh(a_i)}{\sigma_i \sinh(b_i) + \omega_i \cosh(b_i)} + B_i^k. \tag{16}$$

We then update the transmission condition (13), and find

$$\begin{pmatrix} f_{\alpha,i}^k \\ g_{\alpha,i}^k \end{pmatrix} = \begin{pmatrix} 1 - \theta_1 d_i E_i & \theta_1 \nu^{-1} F_i \\ -\theta_2 E_i & 1 - \theta_2 d_i F_i \end{pmatrix} \begin{pmatrix} f_{\alpha,i}^{k-1} \\ g_{\alpha,i}^{k-1} \end{pmatrix}, \tag{17}$$

with

$$E_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{\sigma_i \sinh(b_i) + \omega_i \cosh(b_i)} \frac{1}{\sigma_i \cosh(a_i) + d_i \sinh(a_i)},$$
$$F_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{\sigma_i \cosh(b_i) + \omega_i \sinh(b_i)} \frac{1}{\sigma_i \sinh(a_i) + d_i \cosh(a_i)}.$$

The characteristic polynomial associated with the iteration matrix in (17) is

$$X^2 + (\theta_1 d_i E_i + \theta_2 d_i F_i - 2)X + 1 - \theta_1 d_i E_i - \theta_2 d_i F_i + \theta_1 \theta_2 \sigma_i^2 E_i F_i.$$

We then have the following result.

**Theorem 1.** *Algorithm $NN_{1a}$ (6)-(8) converges if and only if*

$$\rho_{NN_{1a}} := \max_{d_i \in \lambda(A)} \left\{ \left| 1 - \frac{d_i(\theta_1 E_i + \theta_2 F_i) \pm \sqrt{d_i^2(\theta_1 E_i + \theta_2 F_i)^2 - 4\theta_1\theta_2\sigma_i^2 E_i F_i}}{2} \right| \right\} < 1, \tag{18}$$

*where $\lambda(A)$ is the spectrum of the matrix $A$.*

To get more insight in the convergence factor (18), we consider a few special cases. Supposing no final target (i.e., $\gamma = 0$) and a symmetric decomposition $\alpha = \frac{T}{2}$ (i.e., $a_i = b_i$), we have

$$E_i = F_i = \frac{2d_i \tanh(a_i) + \sigma_i(1 + \tanh^2(a_i))}{(\sigma_i^2 + d_i^2)\tanh(a_i) + d_i\sigma_i(1 + \tanh^2(a_i))} < \frac{1}{d_i}.$$

Letting $\theta_1 = \theta_2 = \theta$, the convergence factor (18) then becomes

$$|1 - \theta d_i E_i \pm \theta E_i \sqrt{d_i^2 - \sigma_i^2}|,$$

where the discriminant is negative due to $d_i^2 - \sigma_i^2 = -\nu^{-1}$. Thus, the convergence factor $\rho_{\mathrm{NN}_{1a}}$ in this case is

$$\sqrt{1 - 2\theta d_i E_i + \theta^2 \sigma_i^2 E_i^2} > \sqrt{1 - 2\theta + \theta^2 \sigma_i^2 E_i^2} \geq \sqrt{1 - 2\theta}.$$

**Remark 3.** *For the Laplace operator with homogeneous Dirichlet boundary conditions in our model problem (2), there is no zero eigenvalue for its discretization matrix $A$. For a zero eigenvalue, $d_i = 0$, we have from (10) that*

$$\sigma_i|_{d_i=0} = \sqrt{\nu^{-1}}, \quad \omega_i|_{d_i=0} = \gamma\nu^{-1}, \quad \beta_i|_{d_i=0} = 1. \tag{19}$$

*Substituting (19) into the convergence factor (18), we find* $\rho_{NN_{1a}}|_{d_i=0} = \{|1 \pm \sqrt{-\theta_1\theta_2(E_iF_i)|_{d_i=0}}|\}$ *with*

$$
\begin{aligned}
(E_iF_i)|_{d_i=0} = 2 + \coth(\sqrt{\nu^{-1}}\alpha)\frac{\coth(\sqrt{\nu^{-1}}(T-\alpha)) + \gamma\sqrt{\nu^{-1}}}{1+\gamma\sqrt{\nu^{-1}}\coth(\sqrt{\nu^{-1}}(T-\alpha))} \\
+ \tanh(\sqrt{\nu^{-1}}\alpha)\frac{\tanh(\sqrt{\nu^{-1}}(T-\alpha)) + \gamma\sqrt{\nu^{-1}}}{1+\gamma\sqrt{\nu^{-1}}\tanh(\sqrt{\nu^{-1}}(T-\alpha))}.
\end{aligned}
$$

*Since* $(E_iF_i)|_{d_i=0}$, $\theta_1$, $\theta_2$ *are all positive, the discriminant is once again negative, and we have* $\rho_{NN_{1a}}|_{d_i=0} = \sqrt{1+\theta_1\theta_2(E_iF_i)|_{d_i=0}}$*, which is always greater than one. In other words, the convergence behavior of algorithm* $NN_{1a}$ *for small eigenvalues is not good, and cannot be fixed with relaxation.*

**Remark 4.** *For large eigenvalues* $d_i$*, we have from (10) that*

$$
\sigma_i \sim_\infty d_i, \quad \omega_i \sim_\infty d_i, \quad \beta_i \sim_\infty -d_i, \tag{20}
$$

*and thus obtain* $E_i \sim_\infty \frac{1}{d_i}$ *and* $F_i \sim_\infty \frac{1}{d_i}$*. Substituting these into (18), we find* $\lim_{d_i\to\infty} \rho_{NN_{1a}} = \{|1-\theta_1|, |1-\theta_2|\}$*. In other words, high frequency convergence is robust with relaxation, and one can get a good smoother using* $\theta_1 = \theta_2 = 1$*.*

The above analysis reveals the fact that this most natural NN algorithm is a good smoother but not a good solver.

### 3.1.2 Algorithm $\mathrm{NN}_{1b}$

We apply now the Neumann step only to the primal correction state $\psi_i$. For $k = 1, 2, \ldots$, we consider the algorithm that first solves the Dirichlet step (6), and then corrects it by solving the Neumann step

$$
\begin{cases}
\begin{pmatrix}\dot\psi^k_{1,i}\\ \dot\phi^k_{1,i}\end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1}\\ -1 & -d_i\end{pmatrix}\begin{pmatrix}\psi^k_{1,i}\\ \phi^k_{1,i}\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \text{ in } \Omega_1,\\
\psi^k_{1,i}(0) = 0,\\
\dot\psi^k_{1,i}(\alpha) = \dot z^k_{1,i}(\alpha) - \dot z^k_{2,i}(\alpha),
\end{cases}
\\
\begin{cases}
\begin{pmatrix}\dot\psi^k_{2,i}\\ \dot\phi^k_{2,i}\end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1}\\ -1 & -d_i\end{pmatrix}\begin{pmatrix}\psi^k_{2,i}\\ \phi^k_{2,i}\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \text{ in } \Omega_2,\\
\dot\psi^k_{2,i}(\alpha) = \dot z^k_{2,i}(\alpha) - \dot z^k_{1,i}(\alpha),\\
\phi^k_{2,i}(T) + \gamma\psi^k_{2,i}(T) = 0.
\end{cases}
\tag{21}
$$

As for the update step, let us first consider keeping the same update as (8).

Unlike the Dirichlet step (6), the Neumann step (21) does not have the forward-backward structure in the current form, but this can be recovered using the identities in (9). More precisely, we can rewrite the transmission condition $\dot\psi^k_{1,i}(\alpha) = \dot z^k_{1,i}(\alpha) - \dot z^k_{2,i}(\alpha)$ as

$$
\dot\phi^k_{1,i}(\alpha) - \frac{\sigma_i^2}{d_i}\phi^k_{1,i}(\alpha) = (\dot\mu^k_{1,i}(\alpha) - \frac{\sigma_i^2}{d_i}\mu^k_{1,i}(\alpha)) - (\dot\mu^k_{2,i}(\alpha) - \frac{\sigma_i^2}{d_i}\mu^k_{2,i}(\alpha)),
$$

which is a Robin type condition. In other words, when the forward-backward structure is recovered with this interpretation, the Neumann step (21) becomes a RN step.

Compared with algorithm $\mathrm{NN}_{1a}$, only the Neumann step is modified, which can be transformed into

$$\begin{cases} \ddot{\psi}_{1,i}^k - \sigma_i^2 \psi_{1,i}^k = 0 \text{ in } \Omega_1, \\ \psi_{1,i}^k(0) = 0, \\ \dot{\psi}_{1,i}^k(\alpha) = \dot{z}_{1,i}^k(\alpha) - \dot{z}_{2,i}^k(\alpha), \end{cases} \quad \begin{cases} \ddot{\psi}_{2,i}^k - \sigma_i^2 \psi_{2,i}^k = 0 \text{ in } \Omega_2, \\ \dot{\psi}_{2,i}^k(\alpha) = \dot{z}_{2,i}^k(\alpha) - \dot{z}_{1,i}^k(\alpha), \\ \dot{\psi}_{2,i}^k(T) + \omega_i \psi_{2,i}^k(T) = 0. \end{cases} \tag{22}$$

The convergence analysis is then given by solving explicitly (11), (22) and (13) for one step. In this form, we are actually analyzing here a RD step with a NN correction step. Using (14), we can solve (22) and determine the coefficients

$$C_i^k = A_i^k + B_i^k \frac{\sigma_i \sinh(b_i) + \omega_i \cosh(b_i)}{\cosh(a_i)}, \; D_i^k = A_i^k \frac{\cosh(a_i)}{\sigma_i \sinh(b_i) + \omega_i \cosh(b_i)} + B_i^k. \tag{23}$$

Combining with (15), we update the transmission condition (13) and find

$$\begin{pmatrix} f_{\alpha,i}^k \\ g_{\alpha,i}^k \end{pmatrix} = \begin{pmatrix} 1 - \theta_1 d_i E_i & -\theta_1 d_i F_i \\ -\theta_2 E_i & 1 - \theta_2 F_i \end{pmatrix} \begin{pmatrix} f_{\alpha,i}^{k-1} \\ g_{\alpha,i}^{k-1} \end{pmatrix}, \tag{24}$$

with

$$\begin{aligned} E_i &= \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{\sigma_i \sinh(b_i) + \omega_i \cosh(b_i)} \frac{1}{\sigma_i \cosh(a_i) + d_i \sinh(a_i)}, \\ F_i &= \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{\sigma_i \cosh(b_i) + \omega_i \sinh(b_i)} \frac{1}{\cosh(a_i)}. \end{aligned}$$

In particular, the eigenvalues of the iteration matrix in (24) are 1 and $1 - (\theta_1 d_i E_i + \theta_2 F_i)$, meaning that the algorithm (6), (21), (8) stagnates in its current form, and cannot be fixed even with relaxation.

Note that we choose to keep the same Dirichlet and update steps in the algorithm (6), (21), (8), although the Neumann step has been changed comparing to algorithm $\mathrm{NN}_{1a}$. We also observe from the Neumann correction step (21) that $\dot{\psi}_{1,i}^k(\alpha) + \dot{\psi}_{2,i}^k(\alpha) = 0$, which implies that in this case, the update step (8) in terms of the primal correction state (13) is actually

$$f_{\alpha,i}^k = f_{\alpha,i}^{k-1} - \theta_1 d_i \big(\psi_{1,i}^k(\alpha) + \psi_{2,i}^k(\alpha)\big), \quad g_{\alpha,i}^k = g_{\alpha,i}^{k-1} - \theta_2 \big(\psi_{1,i}^k(\alpha) + \psi_{2,i}^k(\alpha)\big). \tag{25}$$

In other words, we update both $f_{\alpha,i}^k$ and $g_{\alpha,i}^k$ only by $\psi_i^k(\alpha)$. This observation leads to the idea to consider a modified NN algorithm. More precisely, we first remove $d_i$ in (25) as

$$f_{\alpha,i}^k = f_{\alpha,i}^{k-1} - \theta_1 (\psi_{1,i}^k(\alpha) + \psi_{2,i}^k(\alpha)), \quad g_{\alpha,i}^k = g_{\alpha,i}^{k-1} - \theta_2 (\psi_{1,i}^k(\alpha) + \psi_{2,i}^k(\alpha)). \tag{26}$$

In the case when $f_{\alpha,i}^0 = g_{\alpha,i}^0$ and $\theta_1 = \theta_2 = \theta$, we have $f_{\alpha,i}^k = g_{\alpha,i}^k$, $\forall k \in \mathbb{N}$. In this way, we consider the modified NN algorithm which solves first the Dirichlet step

$$
\begin{cases}
\begin{pmatrix} \dot{z}_{1,i}^k \\ \dot{\mu}_{1,i}^k \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} z_{1,i}^k \\ \mu_{1,i}^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_1, \\
z_{1,i}^k(0) = 0, \\
\mu_{1,i}^k(\alpha) = f_{\alpha,i}^{k-1},
\end{cases}
$$
$$
\begin{cases}
\begin{pmatrix} \dot{z}_{2,i}^k \\ \dot{\mu}_{2,i}^k \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} z_{2,i}^k \\ \mu_{2,i}^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_2, \\
z_{2,i}^k(\alpha) = f_{\alpha,i}^{k-1}, \\
\mu_{2,i}^k(T) + \gamma z_{2,i}^k(T) = 0,
\end{cases} \tag{27}
$$

then corrects the result by solving the Neumann step (21) and updates the transmission condition by

$$
f_{\alpha,i}^k = f_{\alpha,i}^{k-1} - \theta(\psi_{1,i}^k(\alpha) + \psi_{2,i}^k(\alpha)), \quad \theta > 0. \tag{28}
$$

For this modified NN algorithm, we find the following result.

**Theorem 2.** *Algorithm $NN_{1b}$ (27), (21), (28) converges if and only if*

$$
\rho_{NN_{1b}} := \max_{d_i \in \lambda(A)} \left|1 - \theta(E_i + F_i)\right| < 1. \tag{29}
$$

Compared to the algorithm (6), (21), (8), algorithm $\text{NN}_{\text{1b}}$ converges with a proper choice of $\theta$. More precisely, for a zero eigenvalue, substituting (19) into (29), we find $\dot{\psi}_{1,i}^k(\alpha) = \dot{z}_{1,i}^k(\alpha) - \dot{z}_{2,i}^k(\alpha)$ as

$$
\dot{\phi}_{1,i}^k(\alpha) - \frac{\sigma_i^2}{d_i}\phi_{1,i}^k(\alpha) = (\dot{\mu}_{1,i}^k(\alpha) - \frac{\sigma_i^2}{d_i}\mu_{1,i}^k(\alpha)) - (\dot{\mu}_{2,i}^k(\alpha) - \frac{\sigma_i^2}{d_i}\mu_{2,i}^k(\alpha)),
$$

meaning that small eigenvalue convergence is good with relaxation. For large eigenvalues $d_i$, using (20), we have $E_i \sim_\infty \frac{1}{d_i}$ and $F_i \sim_\infty 2$. Thus, we obtain $\lim_{d_i \to \infty} \rho_{\text{NN}_{\text{1b}}} = |1 - 2\theta|$, which is independent of the interface $\alpha$. So high frequency convergence is robust with relaxation, and one can get a good smoother using $\theta = 1/2$. By equioscillating the convergence factor for small (i.e., $\rho_{\text{NN}_{\text{1b}}}|_{d_i=0}$) and large (i.e., $\rho_{\text{NN}_{\text{1b}}}|_{d_i\to\infty}$) eigenvalues, we obtain

$$
\theta^*_{\text{NN}_{\text{1b}}} := \frac{2}{3 + \sqrt{\nu}\left(\tanh(\sqrt{\nu^{-1}}\alpha) + \frac{1+\gamma\sqrt{\nu^{-1}}\tanh(\sqrt{\nu^{-1}}(T-\alpha))}{\gamma\sqrt{\nu^{-1}}+\tanh(\sqrt{\nu^{-1}}(T-\alpha))}\right) + \tanh(\sqrt{\nu^{-1}}\alpha)\frac{\gamma\sqrt{\nu^{-1}}+\tanh(\sqrt{\nu^{-1}}(T-\alpha))}{1+\gamma\sqrt{\nu^{-1}}\tanh(\sqrt{\nu^{-1}}(T-\alpha))}}, \tag{30}
$$

which is smaller than $2/3$. However, it is not clear under what condition $\theta^*_{\text{NN}_{\text{1b}}}$ is the optimal relaxation parameter. Indeed, the monotonicity of $E_i$ and $F_i$ with respect to $d_i$ may change according to the parameter values $\alpha$, $\gamma$ and $\nu$. Thus, the variation of $E_i + F_i$ to $d_i$ is less clear even in the case with $\gamma = 0$. Generally, algorithm $\text{NN}_{\text{1b}}$ is a good smoother and can also be a good solver with a proper relaxation parameter $\theta$.

**Remark 5.** *Instead of considering the update step as in (26), we could have also modified (25) to*

$$f_{\alpha,i}^k = f_{\alpha,i}^{k-1} - \theta_1 d_i(\psi_{1,i}^k(\alpha) + \psi_{2,i}^k(\alpha)), g_{\alpha,i}^k = g_{\alpha,i}^{k-1} - \theta_2 d_i(\psi_{1,i}^k(\alpha) + \psi_{2,i}^k(\alpha)).$$

*Using then the same arguments as above, we end up with* $g_{\alpha,i}^k \equiv f_{\alpha,i}^k = f_{\alpha,i}^{k-1}(1-\theta d_i(E_i + F_i))$*. However, the convergence of the algorithm can no longer be guaranteed with this update. More precisely, for a zero eigenvalue* $d_i = 0$*, the convergence factor is one, and cannot be improved with relaxation. As for large eigenvalues, using once again the equivalence relation of* $E_i$ *and* $F_i$*, we find the convergence factor goes to infinity when* $d_i$ *is large.*

In general, the above analysis shows that the update step should also be adapted when modifying the Neumann step.

#### 3.1.3 Algorithm $\mathbf{NN_{1c}}$

Instead of applying the Neumann step to the primal correction state $\psi_i$, we can also apply it only to the dual correction state $\phi_i$. For $k = 1, 2, \ldots$, we consider the algorithm that first solves the Dirichlet step (6), then corrects it by solving the Neumann step

$$\begin{cases} \begin{pmatrix} \dot{\psi}_{1,i}^k \\ \dot{\phi}_{1,i}^k \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} \psi_{1,i}^k \\ \phi_{1,i}^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_1, \\ \psi_{1,i}^k(0) = 0, \\ \dot{\phi}_{1,i}^k(\alpha) = \dot{\mu}_{1,i}^k(\alpha) - \dot{\mu}_{2,i}^k(\alpha), \end{cases} \\ \begin{cases} \begin{pmatrix} \dot{\psi}_{2,i}^k \\ \dot{\phi}_{2,i}^k \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} \psi_{2,i}^k \\ \phi_{2,i}^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_2, \\ \dot{\phi}_{2,i}^k(\alpha) = \dot{\mu}_{2,i}^k(\alpha) - \dot{\mu}_{1,i}^k(\alpha), \\ \phi_{2,i}^k(T) + \gamma\psi_{2,i}^k(T) = 0. \end{cases} \tag{31}$$

Once again, let us first consider keeping the same update step (8).

The Neumann step (31) does not seem to have the forward-backward structure due to the transmission condition on the second domain $\Omega_2$. Using (9), we can rewrite it as

$$\dot{\psi}_{2,i}^k(\alpha) + \frac{\sigma_i^2}{d_i}\psi_{1,i}^k(\alpha) = (\dot{z}_{2,i}^k(\alpha) + \frac{\sigma_i^2}{d_i} z_{2,i}^k(\alpha)) - (\dot{z}_{1,i}^k(\alpha) + \frac{\sigma_i^2}{d_i} z_{1,i}^k(\alpha)),$$

which then becomes a NR step with the usual forward-backward structure.

Once again, only the Neumann step is modified and can be transformed into

$$\begin{cases} \ddot{\psi}_{1,i}^k - \sigma_i^2 \psi_{1,i}^k = 0 \text{ in } \Omega_1, \\ \psi_{1,i}^k(0) = 0, \\ \dot{\psi}_{1,i}^k(\alpha) + \frac{\sigma_i^2}{d_i}\psi_{1,i}^k(\alpha) = \big(\dot{z}_{1,i}^k(\alpha) + \frac{\sigma_i^2}{d_i} z_{1,i}^k(\alpha)\big) - \big(\dot{z}_{2,i}^k(\alpha) + \frac{\sigma_i^2}{d_i} z_{2,i}^k(\alpha)\big), \end{cases}$$
$$\begin{cases} \ddot{\psi}_{2,i}^k - \sigma_i^2 \psi_{2,i}^k = 0 \text{ in } \Omega_2, \\ \dot{\psi}_{2,i}^k(\alpha) + \frac{\sigma_i^2}{d_i}\psi_{1,i}^k(\alpha) = \big(\dot{z}_{2,i}^k(\alpha) + \frac{\sigma_i^2}{d_i} z_{2,i}^k(\alpha)\big) - \big(\dot{z}_{1,i}^k(\alpha) + \frac{\sigma_i^2}{d_i} z_{1,i}^k(\alpha)\big), \\ \dot{\psi}_{2,i}^k(T) + \omega_i \psi_{2,i}^k(T) = 0. \end{cases} \tag{32}$$

The convergence analysis is thus given for a RD step (11) with a RR correction step (32). We can solve (32) using (14) and determine the coefficients

$$C_i^k = A_i^k - B_i^k \nu^{-1} \frac{\sigma_i \gamma \sinh(b_i) + \beta_i \cosh(b_i)}{\sigma_i \sinh(a_i) + d_i \cosh(a_i)}, D_i^k = B_i^k - \nu A_i^k \frac{\sigma_i \sinh(a_i) + d_i \cosh(a_i)}{\sigma_i \gamma \sinh(b_i) + \beta_i \cosh(b_i)}. \tag{33}$$

Combining with (15), we update the transmission condition (13) and find

$$\begin{pmatrix} f_{\alpha,i}^k \\ g_{\alpha,i}^k \end{pmatrix} = \begin{pmatrix} 1-\theta_1 E_i & \theta_1 \nu^{-1} F_i \\ \theta_2 \nu d_i E_i & 1-\theta_2 d_i F_i \end{pmatrix} \begin{pmatrix} f_{\alpha,i}^{k-1} \\ g_{\alpha,i}^{k-1} \end{pmatrix}, \tag{34}$$

with

$$E_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{\sigma_i \gamma \sinh(b_i) + \beta_i \cosh(b_i)} \frac{1}{\sigma_i \cosh(a_i) + d_i \sinh(a_i)},$$
$$F_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{\sigma_i \cosh(b_i) + \omega_i \sinh(b_i)} \frac{1}{\sigma_i \sinh(a_i) + d_i \cosh(a_i)}.$$

In particular, the eigenvalues of the iteration matrix in (34) are 1 and $1-(\theta_1 E_i + \theta_2 d_i F_i)$. Once again, the algorithm (6), (31), (8) stagnates, and cannot be fixed with relaxation. Similar as in Section 3.1.2, we can adapt the transmission condition (8) and make this algorithm converge. More precisely, we first consider the update

$$f_{\alpha,i}^k = f_{\alpha,i}^{k-1} - \theta(\phi_{1,i}^k(\alpha) + \phi_{2,i}^k(\alpha)), g_{\alpha,i}^k = g_{\alpha,i}^{k-1} - \theta(\phi_{1,i}^k(\alpha) + \phi_{2,i}^k(\alpha)).$$

In the case when $f_{\alpha,i}^0 = g_{\alpha,i}^0$ and $\theta_1 = \theta_2 = \theta$, we have $g_{\alpha,i}^k = f_{\alpha,i}^k$, $\forall k \in \mathbb{N}$ and

$$f_{\alpha,i}^k = f_{\alpha,i}^{k-1} - \theta(\phi_{1,i}^k(\alpha) + \phi_{2,i}^k(\alpha)). \tag{35}$$

This leads to the following result.

**Theorem 3.** *Algorithm $NN_{1c}$ (27), (31), (35) converges if and only if*

$$\rho_{NN_{1c}} := \max_{d_i \in \lambda(A)} |1 - \theta(E_i - \nu^{-1} F_i)| < 1. \tag{36}$$

Compared to the algorithm (6), (31), (8), algorithm $\mathrm{NN}_{1c}$ may converge with a proper choice of $\theta$. More precisely, for a zero eigenvalue, $d_i = 0$, we find

$$\rho_{\mathrm{NN}_{1c}}|_{d_i=0} = |1 - \theta(1 + \tanh(\sqrt{\nu^{-1}}\alpha)\frac{\gamma\sqrt{\nu^{-1}} + \tanh(\sqrt{\nu^{-1}}(T-\alpha))}{\gamma\sqrt{\nu^{-1}}\tanh(\sqrt{\nu^{-1}}(T-\alpha)) + 1} - \sqrt{\nu^{-1}}(\coth(\sqrt{\nu^{-1}}\alpha) + \frac{\gamma\sqrt{\nu^{-1}} + \tanh(\sqrt{\nu^{-1}}(T-\alpha))}{1 + \gamma\sqrt{\nu^{-1}}\tanh(\sqrt{\nu^{-1}}(T-\alpha))}))|.$$

Depending on the values of $\nu$, $\gamma$ and $\alpha$, $(E_i - \nu^{-1}F_i)|_{d_i=0}$ could be negative, then $\rho_{\mathrm{NN}_{1c}}|_{d_i=0}$ would be greater than one since $\theta > 0$. In other words, the convergence for small eigenvalues could be not good, and cannot be fixed even with relaxation. For large eigenvalues $d_i$, using (20), we find $E_i \sim_\infty 2$ and $F_i \sim_\infty \frac{1}{d_i}$. Thus, we obtain $\lim_{d_i\to\infty} \rho_{\mathrm{NN}_{1c}} = |1 - 2\theta|$, which is independent of the interface $\alpha$. So large eigenvalue convergence is robust with relaxation, and one can get a good smoother using $\theta = 1/2$. Moreover, we observe that algorithms $\mathrm{NN}_{1b}$ and $\mathrm{NN}_{1c}$ share similar behavior for large eigenvalues. By equioscillating the convergence factor for small (i.e., $\rho_{\mathrm{NN}_{1c}}|_{d_i=0}$) and large (i.e., $\rho_{\mathrm{NN}_{1c}}|_{d_i\to\infty}$) eigenvalues, we obtain

$$\theta^*_{\mathrm{NN}_{1c}} := \frac{2}{3 + \tanh(\sqrt{\nu^{-1}}\alpha)\frac{\gamma\sqrt{\nu^{-1}} + \tanh(\sqrt{\nu^{-1}}(T-\alpha))}{\gamma\sqrt{\nu^{-1}}\tanh(\sqrt{\nu^{-1}}(T-\alpha)) + 1} - \sqrt{\nu^{-1}}(\coth(\sqrt{\nu^{-1}}\alpha) + \frac{\gamma\sqrt{\nu^{-1}} + \tanh(\sqrt{\nu^{-1}}(T-\alpha))}{1 + \gamma\sqrt{\nu^{-1}}\tanh(\sqrt{\nu^{-1}}(T-\alpha))})}. \tag{37}$$

Note that when $(E_i - \nu^{-1}F_i)|_{d_i=0} < 0$, the relaxation cannot improve the convergence for small eigenvalues, thus, (37) could also be negative and cannot provide the optimal value of $\theta$ in this case. One may use however a negative relaxation parameter $\theta$ to make the algorithm converge for small eigenvalues, but this will induce divergence for large eigenvalues. Based on the analysis, algorithm $\mathrm{NN}_{1c}$ is a good smoother but not necessarily a good solver.

**Remark 6.** *One could also consider the update step* (28) *instead of* (35)*, and the convergence factor* (36) *will be* $\max_{d_i\in\lambda(A)} |1 - \theta d_i(F_i - \nu E_i)|$*. For a similar reason as in Remark 5, the algorithm diverges with this choice of update step.*

Together with the analysis in Section 3.1.2, we observe that keeping the same update step (8) leads to divergent algorithms, when modifying the Neumann step. Thus, we should also adapt the update step according to the Neumann step.

## 3.2 Category II

We now study the algorithms in Category II which run the Dirichlet step only on the primal state $z_i$.

#### 3.2.1 Algorithm NN$_{2a}$

The most natural way is to correct $z_i$ by the primal correction state $\psi_i$. For $k = 1, 2, ...$, algorithm NN$_{2a}$ first solves the Dirichlet step

$$
\begin{cases}
\begin{pmatrix} \dot{z}^k_{1,i} \\ \dot{\mu}^k_{1,i} \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} z^k_{1,i} \\ \mu^k_{1,i} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_1, \\
z^k_{1,i}(0) = 0, \\
z^k_{1,i}(\alpha) = f^{k-1}_{\alpha,i},
\end{cases}
\\
\begin{cases}
\begin{pmatrix} \dot{z}^k_{2,i} \\ \dot{\mu}^k_{2,i} \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} z^k_{2,i} \\ \mu^k_{2,i} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_2, \\
z^k_{2,i}(\alpha) = f^{k-1}_{\alpha,i}, \\
\mu^k_{2,i}(T) + \gamma z^k_{2,i}(T) = 0,
\end{cases}
\tag{38}
$$

then corrects the result by solving the Neumann step (21), and updates the transmission condition by (28)

**Remark 7.** *Here, it is more natural to consider the transmission condition only for $f^k_{\alpha,i}$. This is due to the continuity of the primal state $z^k_i$ at the interface $\alpha$. In general, we can show that an update step as (35) will lead to divergence for a similar reason as in Remark 5. We can also show that a pair of transmission conditions $(f^k_{\alpha,i}, g^k_{\alpha,i})$ will lead to non-convergent behavior (see Appendix A).*

For algorithm NN$_{2a}$, neither the Dirichlet (38) nor the Neumann step (21) has the forward-backward structure in its current form. We have seen in Section 3.1.2 that we can recover this structure for the Neumann step (21) which becomes a RN step. Using the same idea, we can interpret $z^k_{1,i}(\alpha) = f^{k-1}_{\alpha,i}$ as $\dot{\mu}^k_{1,i}(\alpha) - d_i \mu^k_{1,i}(\alpha) = f^{k-1}_{\alpha,i}$ to recover the forward-backward structure, and the Dirichlet step (38) then becomes a ND step.

For the convergence analysis, we transform the Dirichlet step (38) using (9) and (10), and find

$$
\begin{cases}
\ddot{z}^k_{1,i} - \sigma_i^2 z^k_{1,i} = 0 \text{ in } \Omega_1, \\
z^k_{1,i}(0) = 0, \\
z^k_{1,i}(\alpha) = f^{k-1}_{\alpha,i},
\end{cases}
\qquad
\begin{cases}
\ddot{z}^k_{2,i} - \sigma_i^2 z^k_{2,i} = 0 \text{ in } \Omega_2, \\
z^k_{2,i}(\alpha) = f^{k-1}_{\alpha,i}, \\
\dot{z}^k_{2,i}(T) + \omega_i z^k_{2,i}(T) = 0.
\end{cases}
\tag{39}
$$

The Neumann step becomes (22), and we keep the same update step (28). In particular, the convergence analysis also proceeds on a NN algorithm (39), (22), (28). Using (14), we can solve (39) and determine the coefficients,

$$
A^k_i = \frac{f^{k-1}_{\alpha,i}}{\sinh(a_i)}, \quad B^k_i = \frac{f^{k-1}_{\alpha,i}}{\sigma_i \cosh(b_i) + \omega_i \sinh(b_i)}. \tag{40}
$$

Combining them with (23), we update the transmission condition (28) and find $f^k_{\alpha,i} = f^{k-1}_{\alpha,i} - \theta f^{k-1}_{\alpha,i}(E_i + F_i)$, with

$$
E_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{(\sigma_i \sinh(b_i) + \omega_i \cosh(b_i)) \sinh(a_i)}, F_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{(\sigma_i \cosh(b_i) + \omega_i \sinh(b_i)) \cosh(a_i)}.
$$

This leads to the following result.

**Theorem 4.** *Algorithm $NN_{2a}$ (38), (21), (28) converges if and only if*

$$\rho_{NN_{2a}} := \max_{d_i \in \lambda(A)} |1 - \theta(E_i + F_i)| < 1. \tag{41}$$

In particular, for a zero eigenvalue, substituting (19) into (41), we have

$$\rho_{\mathrm{NN_{2a}}}|_{d_i=0} = \Big|1 - \theta\Big(2 + \coth(\sqrt{\nu^{-1}}\alpha)\frac{\coth\left(\sqrt{\nu^{-1}}(T-\alpha)\right) + \gamma\sqrt{\nu^{-1}}}{1 + \gamma\sqrt{\nu^{-1}}\coth\left(\sqrt{\nu^{-1}}(T-\alpha)\right)} + \tanh(\sqrt{\nu^{-1}}\alpha)\frac{\tanh\left(\sqrt{\nu^{-1}}(T-\alpha)\right) + \gamma\sqrt{\nu^{-1}}}{1 + \gamma\sqrt{\nu^{-1}}\tanh\left(\sqrt{\nu^{-1}}(T-\alpha)\right)}\Big)\Big|. \tag{42}$$

For large eigenvalues $d_i$, using (20), we find $E_i \sim_\infty 2$ and $F_i \sim_\infty 2$. Thus, we obtain $\lim_{d_i \to \infty} \rho_{\mathrm{NN_{2a}}} = |1 - 4\theta|$, which is independent of the interface $\alpha$. So the convergence for high frequencies is robust with relaxation, and one can get a good smoother using $\theta = 1/4$. By equioscillating the convergence factor for small (i.e., $\rho_{\mathrm{NN_{2a}}}|_{d_i=0}$) and large (i.e., $\rho_{\mathrm{NN_{2a}}}|_{d_i\to\infty}$) eigenvalues, we obtain the relaxation parameter

$$\theta^*_{\mathrm{NN_{2a}}} := \frac{2}{6 + \coth(\sqrt{\nu^{-1}}\alpha)\frac{\coth(\sqrt{\nu^{-1}}(T-\alpha)) + \gamma\sqrt{\nu^{-1}}}{1+\gamma\sqrt{\nu^{-1}}\coth(\sqrt{\nu^{-1}}(T-\alpha))} + \tanh(\sqrt{\nu^{-1}}\alpha)\frac{\tanh(\sqrt{\nu^{-1}}(T-\alpha)) + \gamma\sqrt{\nu^{-1}}}{1+\gamma\sqrt{\nu^{-1}}\tanh(\sqrt{\nu^{-1}}(T-\alpha))}}, \tag{43}$$

which is smaller than 1/3. In the case with no final state, i.e., $\gamma = 0$, we have

$$\theta^*_{\mathrm{NN_{2a}}}|_{\gamma=0} = \frac{2}{6 + \coth(\sqrt{\nu^{-1}}\alpha)\coth(\sqrt{\nu^{-1}}(T-\alpha)) + \tanh(\sqrt{\nu^{-1}}\alpha)\tanh(\sqrt{\nu^{-1}}(T-\alpha))}.$$

Using properties of the hyperbolic tangent and cotangent, we find

$$\coth(\sqrt{\nu^{-1}}\alpha)\coth(\sqrt{\nu^{-1}}(T-\alpha)) + \tanh(\sqrt{\nu^{-1}}\alpha)\tanh(\sqrt{\nu^{-1}}(T-\alpha)) \geq \coth^2(\sqrt{\nu^{-1}}\frac{T}{2}) + \tanh^2(\sqrt{\nu^{-1}}\frac{T}{2}) > 2,$$

thus $\theta^*_{\mathrm{NN_{2a}}} < \frac{1}{4}$. Based on the analysis, algorithm $\mathrm{NN_{2a}}$ is a good smoother and can also be a good solver. However, it is less clear under what condition $\theta^*_{\mathrm{NN_{2a}}}$ is the optimal relaxation parameter, since the monotonicity of the convergence factor with respect to the eigenvalues $d_i$ is not clear even in the case $\gamma = 0$. This has been observed in our numerical experiments.

### 3.2.2 Algorithm $\mathrm{NN_{2b}}$

We can also keep the Dirichlet step (38), but apply the Neumann step only to the dual correction state $\phi_i$ as in (31). As for the update step, we first consider to take the same update as for algorithm $\mathrm{NN_{2a}}$, i.e., (28).

For the convergence analysis, we actually solve a DD step (39) and correct by a RR step (32). Using (40) and (33), we update the transmission condition (28) and find $f_{\alpha,i}^k = f_{\alpha,i}^{k-1}(1 - \theta d_i(F_i - \nu E_i))$ with

$$E_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{\sigma_i \gamma \sinh(b_i) + \beta_i \cosh(b_i)} \frac{1}{\sinh(a_i)},$$
$$F_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega \sinh(\sigma_i T)}{(\sigma_i \cosh(b_i) + \omega_i \sinh(b_i))(\sigma_i \sinh(a_i) + d_i \cosh(a_i))}.$$

We then obtain the convergence factor

$$\rho_{\mathrm{NN}_{2\mathrm{b}}} := \max_{d_i \in \lambda(A)} |1 - \theta d_i(F_i - \nu E_i)| < 1. \tag{44}$$

To get more insight, we first study the extremal cases. For a zero eigenvalue, $d_i = 0$, substituting (19) into (44), we have $(F_i - \nu E_i)|_{d_i=0} = 0$. Hence, we find $\rho_{\mathrm{NN}_{2\mathrm{b}}}|_{d_i=0} = 1$, which is independent of the relaxation parameter. In other words, the convergence behavior of algorithm $\mathrm{NN}_{2\mathrm{b}}$ is not good for small eigenvalues, and the relaxation cannot fix this problem. For large eigenvalues $d_i$, using (20), we find $E_i \sim_\infty 4d_i$ and $F_i \sim_\infty \frac{1}{d_i}$. Thus, we obtain $1 - \theta d_i(F_i - \nu E_i) \sim_\infty 4\nu\theta d_i^2$ and $\lim_{d_i\to\infty} \rho_{\mathrm{NN}_{2\mathrm{b}}} = \infty$, which is divergent, and cannot be fixed with relaxation. Generally, we have the following result.

**Theorem 5.** *Algorithm $NN_{2b}$ (38) (31) (28) always diverges.*

*Proof.* Using the formula of $E_i$ and $F_i$, we find $F_i - \nu E_i = \frac{-\nu d_i}{\sigma_i \sinh(a_i) + d_i \cosh(a_i)} \frac{(\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T))^2}{\sinh(a_i)(\sigma_i \gamma \sinh(b_i) + \beta_i \cosh(b_i))(\sigma_i \cosh(b_i) + \omega_i \sinh(b_i))}$ which is negative or zero (if $d_i = 0$). Since $\theta$ and $\nu$ are both positive, $1 - \theta d_i(F_i - \nu E_i) \geq 1$ which concludes the proof. ☐

The above result shows that algorithm $\mathrm{NN}_{2\mathrm{b}}$ diverges with a positive relaxation parameter $\theta$. Moreover, this divergence cannot be fixed even with a negative $\theta$, since the convergence factor is one for a zero eigenvalue, and is equivalent to $4\nu|\theta|d_i^2$ for large eigenvalues. In general, algorithm $\mathrm{NN}_{2\mathrm{b}}$ is neither a good smoother nor a good solver.

**Remark 8.** *Compared with algorithm $NN_{2a}$, we change the Neumann step but keep the same update step. One can also consider the update step (35), since the Neumann correction (31) is only applied to the dual correction state $\phi_i$. Following the same computation, the convergence factor (44) then becomes*

$$\max_{d_i \in \lambda(A)} |1 - \theta(E_i - \nu^{-1} F_i)|,$$

*with $E_i - \nu^{-1} F_i \geq 0$. However, this does not change the poor convergence behavior for both small and large eigenvalues. Indeed, we still have $(E_i - \nu^{-1} F_i)|_{d_i=0} = 0$, hence $\rho_{NN_{2b}}|_{d_i=0} = 1$, and $\lim_{d_i\to\infty} \rho_{NN_{2b}} = \infty$. Thus, the modified algorithm stays divergent. Furthermore, for a similar reason as mentioned in Appendix A, the algorithm is also divergent when considering the update step (8) with a pair of transmission conditions $(f_{\alpha,i}^k, g_{\alpha,i}^k)$.*

Based on the analysis, we cannot find a good NN algorithm when combining the Dirichlet step (38) with the Neumann step (31).

### 3.2.3 Algorithm $\mathbf{NN_{2c}}$

If we apply the correction to the pair $(\psi_i, \phi_i)$, then the Neumann step immediately has the forward-backward structure. In this way, algorithm $\mathrm{NN_{2c}}$ solves first the Dirichlet step (38), next the Neumann step (7) and updates the transmission condition by (28).

For the convergence analysis, we solve a DD step (39) followed by a RN correction step (12). Using (40) and (16), we update the transmission condition (28) and find $f^k_{\alpha,i} = f^{k-1}_{\alpha,i}(1-\theta(E_i + d_i F_i))$ with

$$E_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{(\sigma_i \sinh(b_i) + \omega_i \cosh(b_i))\sinh(a_i)},$$
$$F_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{(\sigma_i \cosh(b_i) + \omega_i \sinh(b_i))(\sigma_i \sinh(a_i) + d_i \cosh(a_i))}.$$

We then obtain the following result.

**Theorem 6.** *Algorithm $NN_{2c}$ (38), (7), (28) converges if and only if*

$$\rho_{NN_{2c}} := \max_{d_i \in \lambda(A)} |1 - \theta(E_i + d_i F_i)| < 1. \tag{45}$$

For a zero eigenvalue $d_i = 0$, substituting the identities (19) into (45), we find

$$\rho_{\mathrm{NN_{2c}}}|_{d_i=0} = \Big|1 - \theta\big(1 + \coth(\sqrt{\nu^{-1}}\alpha)\frac{\coth(\sqrt{\nu^{-1}}(T-\alpha)) + \gamma\sqrt{\nu^{-1}}}{1 + \gamma\sqrt{\nu^{-1}}\coth(\sqrt{\nu^{-1}}(T-\alpha))}\big)\Big|. \tag{46}$$

For large eigenvalues $d_i$, using (20), we find $E_i \sim_\infty 2$ and $F_i \sim_\infty \frac{1}{d_i}$. Thus, we obtain $\lim_{d_i\to\infty} \rho_{\mathrm{NN_{2c}}} = |1 - 3\theta|$, which is independent of the interface $\alpha$. So the convergence for high frequencies is robust with relaxation, and one can get a good smoother using $\theta = 1/3$. By equioscillating the convergence factor for small (i.e., $\rho_{\mathrm{NN_{2c}}}|_{d_i=0}$) and large (i.e., $\rho_{\mathrm{NN_{2c}}}|_{d_i\to\infty}$) eigenvalues, we obtain

$$\theta^*_{\mathrm{NN_{2c}}} := \frac{2}{4 + \coth(\sqrt{\nu^{-1}}\alpha)\frac{\coth(\sqrt{\nu^{-1}}(T-\alpha)) + \gamma\sqrt{\nu^{-1}}}{1 + \gamma\sqrt{\nu^{-1}}\coth(\sqrt{\nu^{-1}}(T-\alpha))}}, \tag{47}$$

which is smaller than $1/2$. In the case $\gamma = 0$, the relaxation parameter $\theta^*_{\mathrm{NN_{2c}}}$ is bounded by $2/5$. However, it is also not clear under what condition $\theta^*_{\mathrm{NN_{2c}}}$ is the optimal relaxation parameter, since the monotonicity of $E_i + d_i F_i$ with respect to $d_i$ is less clear, and depends on the parameter values $\alpha$, $\gamma$ and $\nu$. Generally, algorithm $\mathrm{NN_{2c}}$ is both a good smoother and a good solver with a well-chosen $\theta$.

**Remark 9.** *Instead of choosing* (28) *as the update step, one could have considered the update step* (35)*. Following the same computation, the convergence factor becomes* $\max_{d_i\in\lambda(A)}|1-\theta(d_iE_i-\nu^{-1}F_i)|$*, which diverges for large eigenvalues. Furthermore, the algorithm will also be divergent when considering the update step* (8) *with a pair transmission conditions* $(f^k_{\alpha,i}, g^k_{\alpha,i})$ *as mentioned in Appendix A.*

## 3.3 Category III

The algorithms in Category III run the Dirichlet step only on the dual state $\mu_i$, and according to the Neumann step, there are three variants.

### 3.3.1 Algorithm $\mathrm{NN_{3a}}$

As in Section 3.2.1, the most natural way is to correct the dual state $\mu_i$ only by the dual correction state $\phi_i$. In this way, for $k = 1, 2, ...$, algorithm $\mathrm{NN_{3a}}$ first solves the Dirichlet step

$$\begin{cases} \begin{pmatrix} \dot z^k_{1,i} \\ \dot\mu^k_{1,i}\end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i\end{pmatrix}\begin{pmatrix} z^k_{1,i} \\ \mu^k_{1,i}\end{pmatrix} = \begin{pmatrix} 0 \\ 0\end{pmatrix} \text{ in } \Omega_1, \\ z^k_{1,i}(0) = 0, \\ \mu^k_{1,i}(\alpha) = f^{k-1}_{\alpha,i}, \end{cases}$$
$$\begin{cases} \begin{pmatrix} \dot z^k_{2,i} \\ \dot\mu^k_{2,i}\end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i\end{pmatrix}\begin{pmatrix} z^k_{2,i} \\ \mu^k_{2,i}\end{pmatrix} = \begin{pmatrix} 0 \\ 0\end{pmatrix} \text{ in } \Omega_2, \\ \mu^k_{2,i}(\alpha) = f^{k-1}_{\alpha,i}, \\ \mu^k_{2,i}(T) + \gamma z^k_{2,i}(T) = 0, \end{cases} \tag{48}$$

then corrects the above result by solving the Neumann step (31), and updates the transmission condition by (35).

Similar to Remark 7, we choose here the update step (35) because of the continuity of the dual state $\mu^k_i$ at the interface $\alpha$, since other choices of the update step will induce divergence behavior. Regarding the forward-backward structure for the Dirichlet step (48), we can recover it by interpreting $\mu^k_{2,i}(\alpha) = f^{k-1}_{\alpha,i}$ as $\dot z^k_{2,i}(\alpha) + d_i z^k_{2,i}(\alpha) = f^{k-1}_{\alpha,i}$. The Dirichlet step (48) then becomes a NR step.

To analyze algorithm $\mathrm{NN_{3a}}$, we can rewrite the Dirichlet step (48) using (9) and (10), and find

$$\begin{cases} \ddot z^k_{1,i} - \sigma_i^2 z^k_{1,i} = 0 \text{ in } \Omega_1, \\ z^k_{1,i}(0) = 0, \\ \dot z^k_{1,i}(\alpha) + d_i z^k_{1,i}(\alpha) = f^{k-1}_{\alpha,i}, \end{cases} \quad \begin{cases} \ddot z^k_{2,i} - \sigma_i^2 z^k_{2,i} = 0 \text{ in } \Omega_2, \\ \dot z^k_{2,i}(\alpha) + d_i z^k_{2,i}(\alpha) = f^{k-1}_{\alpha,i}, \\ \dot z^k_{2,i}(T) + \omega_i z^k_{2,i}(T) = 0. \end{cases} \tag{49}$$

We then correct the above RR step by a RR correction (32), which is also the equivalent of the Neumann step (31). And the update step (35) becomes

$$f^k_{\alpha,i} = f^{k-1}_{\alpha,i} - \theta\big(\dot\psi^k_{1,i}(\alpha) + d_i\psi^k_{1,i}(\alpha) + \dot\psi^k_{2,i}(\alpha) + d_i\psi^k_{2,i}(\alpha)\big). \tag{50}$$

Using (14), we can solve explicitly (49) and determine the coefficients

$$A_i^k = \frac{f_{\alpha,i}^{k-1}}{\sigma_i \cosh(a_i) + d_i \sinh(a_i)}, \quad B_i^k = -\nu \frac{f_{\alpha,i}^{k-1}}{\sigma_i \gamma \cosh(b_i) + \beta_i \sinh(b_i)}. \tag{51}$$

Combining with (33), we update the transmission condition (50) and obtain $f_{\alpha,i}^k = f_{\alpha,i}^{k-1} - \theta f_{\alpha,i}^{k-1}(E_i + F_i)$ with

$$E_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{\sigma_i \gamma \sinh(b_i) + \beta_i \cosh(b_i)} \frac{1}{\sigma_i \cosh(a_i) + d_i \sinh(a_i)},$$
$$F_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega \sinh(\sigma_i T)}{\sigma_i \gamma \cosh(b_i) + \beta_i \sinh(b_i)} \frac{1}{\sigma_i \sinh(a_i) + d_i \cosh(a_i)}.$$

Thus, we have the following result.

**Theorem 7.** *Algorithm $NN_{3a}$ (48), (31), (35) converges if and only if*

$$\rho_{NN_{3a}} := \max_{d_i \in \lambda(A)} |1 - \theta(E_i + F_i)| < 1. \tag{52}$$

We consider some special cases to get more insight in the convergence factor (52). Assuming no final target (i.e., $\gamma = 0$) and a symmetric decomposition $\alpha = \frac{T}{2}$ (i.e., $a_i = b_i$), we find that $E_i$ and $F_i$ are actually the same as for algorithm $\mathrm{NN_{2a}}$ in Section 3.2.1. Hence, the convergence factor (52) is as (41) under this assumption, and $\mathrm{NN_{2a}}$ and $\mathrm{NN_{3a}}$ are actually the same algorithm. Moreover, for a zero eigenvalue, substituting (19) into (52), we find exactly the same formula as (42). Thus, the two algorithms $\mathrm{NN_{2a}}$ and $\mathrm{NN_{3a}}$ share the same behavior for small eigenvalues. On the other hand, using (20) for large eigenvalues $d_i$, we find $E_i \sim_\infty 2$ and $F_i \sim_\infty 2$. This implies that $\lim_{d_i \to \infty} \rho_{\mathrm{NN_{3a}}} = |1 - 4\theta|$, which is the same as for algorithm $\mathrm{NN_{2a}}$. Once again, the two algorithms $\mathrm{NN_{2a}}$ and $\mathrm{NN_{3a}}$ share the same behavior for large eigenvalues. Hence, we obtain the same relaxation parameter $\theta^*_{\mathrm{NN_{3a}}} = \theta^*_{\mathrm{NN_{2a}}}$ as defined in (43). In general, algorithm $\mathrm{NN_{3a}}$ seems to be very similar to $\mathrm{NN_{2a}}$, and we could also expect it to be a good smoother and solver.

### 3.3.2 Algorithm $\mathrm{NN_{3b}}$

The second variant in Category III consists in applying the Neumann step to the primal correction state $\psi_i$. In this way, we consider the algorithm that first solves the Dirichlet step (48), followed by the Neumann step (21), and updates the transmission condition by (35).

For the convergence analysis, we solve a RR step (49) and correct by a NN step (22). Using (51) and (23), we can update the transmission condition (50) and find $f_{\alpha,i}^k = f_{\alpha,i}^{k-1} - f_{\alpha,i}^{k-1} \theta d_i (E_i - \nu F_i)$ with

$$E_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{(\sigma_i \sinh(b_i) + \omega_i \cosh(b_i))(\sigma_i \cosh(a_i) + d_i \sinh(a_i))},$$
$$F_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega \sinh(\sigma_i T)}{(\sigma_i \gamma \cosh(b_i) + \beta_i \sinh(b_i)) \cosh(a_i)}.$$

This leads to the convergence factor

$$\rho_{\mathrm{NN}_{3\mathrm{b}}} := \max_{d_i \in \lambda(A)} |1 - \theta d_i (E_i - \nu F_i)| < 1. \tag{53}$$

We first study the extreme cases. For a zero eigenvalue, substituting the identities (19) into (53), we find $(E_i - \nu F_i)|_{d_i=0} = 0$, and hence $\rho_{\mathrm{NN}_{3\mathrm{b}}}|_{d_i=0} = 1$. This is once again independent of the relaxation parameter. In other words, the convergence of this algorithm is not good for small eigenvalues, and the relaxation cannot fix this problem. For large eigenvalues $d_i$, using (20), we find $E_i \sim_\infty \frac{1}{d_i}$ and $F_i \sim_\infty 4d_i$. Thus, we obtain $\rho_{\mathrm{NN}_{3\mathrm{b}}} \sim_\infty 4\nu\theta d_i^2$ and $\lim_{d_i \to \infty} \rho_{\mathrm{NN}_{3\mathrm{b}}} = \infty$, which is divergent and cannot be fixed with relaxation. In general, we have the following result.

**Theorem 8.** *Algorithm $NN_{3b}$ (48), (21), (35) always diverges.*

*Proof.* Following the same idea as in the proof of Theorem 5, we can show that $E_i - \nu F_i$ is always negative or zero, and this concludes the proof. □

**Remark 10.** *One could have also applied a similar strategy as in Remark 8, that is, considering the update step (28) instead of (35). The convergence factor (53) then becomes $\max_{d_i \in \lambda(A)} |1 - \theta(E_i - \nu F_i)|$. Once again, this does not change the poor convergence behavior for both small and large eigenvalues.*

Similar to algorithm $\mathrm{NN}_{2\mathrm{b}}$, algorithm $\mathrm{NN}_{3\mathrm{b}}$ is neither a good smoother nor a good solver, and other choices of the update step will not change this. Together with Section 3.2.2, we observe that, applying the Dirichlet step to the primal state $z_i$ (resp. dual state $\mu_i$) and correcting the result by a Neumann step to the dual correction state $\phi_i$ (resp. primal correction state $\psi_i$), will lead to divergent algorithms, and cannot be fixed even by adapting the update step.

### 3.3.3 Algorithm $\mathrm{NN}_{3\mathrm{c}}$

The last variant consists in applying the Neumann step to the pair $(\psi_i, \phi_i)$. In this way, the $\mathrm{NN}_{3\mathrm{b}}$ algorithm solves first the Dirichlet step (48), next the Neumann step (7) which also has the forward-backward structure. Then it updates the transmission condition by (35).

For the convergence analysis, we solve a RR step (49) followed by a NR correction (12). Using (51) and (16), we update the transmission condition (50) and find $f_{\alpha,i}^k = f_{\alpha,i}^{k-1}(1 - \theta(d_i E_i + F_i))$ with

$$E_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega_i \sinh(\sigma_i T)}{(\sigma_i \sinh(b_i) + \omega_i \cosh(b_i))(\sigma_i \cosh(a_i) + d_i \sinh(a_i))},$$
$$F_i = \frac{\sigma_i \cosh(\sigma_i T) + \omega \sinh(\sigma_i T)}{(\sigma_i \gamma \cosh(b_i) + \beta_i \sinh(b_i))(\sigma_i \sinh(a_i) + d_i \cosh(a_i))}.$$

We thus find the following result.

**Theorem 9.** *Algorithm $NN_{3c}$ (48), (7), (35) converges if and only if*

$$\rho_{NN_{3c}} := \max_{d_i \in \lambda(A)} |1 - \theta(d_i E_i + F_i)| < 1. \tag{54}$$

We consider some special cases to get more insight. Assuming no final target (i.e., $\gamma = 0$) and a symmetric decomposition $\alpha = \frac{T}{2}$ (i.e., $a_i = b_i$), we find that $E_i$ is actually the same as the $F_i$ for algorithm $\mathrm{NN_{2c}}$, and $F_i$ is the same as the $E_i$ for algorithm $\mathrm{NN_{2c}}$ in Section 3.2.3. Hence, $\mathrm{NN_{2c}}$ and $\mathrm{NN_{3c}}$ are the same algorithm under this assumption. For a zero eigenvalue, $d_i = 0$, substituting the identities (19) into (54), we find $\rho_{\mathrm{NN_{3c}}}|_{d_i=0} = \rho_{\mathrm{NN_{2c}}}|_{d_i=0}$ as in (45). In other words, algorithms $\mathrm{NN_{2c}}$ and $\mathrm{NN_{3c}}$ have a similar behavior for small eigenvalues. For large eigenvalues $d_i$, using (20), we find $E_i \sim_\infty \frac{1}{d_i}$ and $F_i \sim_\infty 2$. Thus, we obtain $\lim_{d_i \to \infty} \rho_{\mathrm{NN_{3c}}} = |1 - 3\theta|$, which is independent of the interface $\alpha$. So the convergence for large eigenvalues is robust with relaxation, and one can get a good smoother using $\theta = 1/3$. Furthermore, we find again similar behavior between algorithms $\mathrm{NN_{2c}}$ and $\mathrm{NN_{3c}}$ for large eigenvalues. Using hence equioscillation, we obtain $\theta^*_{\mathrm{NN_{3c}}} = \theta^*_{\mathrm{NN_{2c}}}$ as defined in (47). Based on all these similarities with algorithm $\mathrm{NN_{2c}}$, algorithm $\mathrm{NN_{3c}}$ is also a good smoother and solver. Also for a similar reason as explained in Remark 9, other choices of the update step will lead to divergent behavior.

# 4 Numerical results

We illustrate now our nine new time domain decomposition algorithms with numerical experiments. As mentioned in the convergence analysis, some algorithms are much more sensitive to the chosen parameters than others. To well illustrate and compare these algorithms, we consider two different test cases,

**case A:** The time interval $\Omega = (0, 1)$ is subdivided into $\Omega_1 = (0, 0.5)$, $\Omega_2 = (0.5, 1)$ (i.e., symmetric), and the objective function has no explicit final target term ($\gamma = 0$). The regularization parameter is $\nu = 0.1$.

**case B:** The time interval $\Omega = (0, 5)$ is subdivided into $\Omega_1 = (0, 1)$, $\Omega_2 = (1, 5)$ (i.e., asymmetric), and the objective function has a final target term with $\gamma = 10$. The regularization parameter is $\nu = 10$.

For each test, we will investigate the performance by plotting the convergence factor as a function of the eigenvalues $d_i \in [10^{-2}, 10^2]$.

## 4.1 Convergence factor of $\mathrm{NN_{2b}}$ and $\mathrm{NN_{3b}}$

We first illustrate the behavior of $\mathrm{NN_{2b}}$ and $\mathrm{NN_{3b}}$ separately, since their convergence analyses are very similar, and both algorithms are divergent. Figure 2 shows the behavior of the convergence factor as a function of the eigenvalues for these two algorithms. More precisely, both algorithms diverge in the case $\theta = 0.25$. And for both test cases A and B, the two algorithms diverge violently

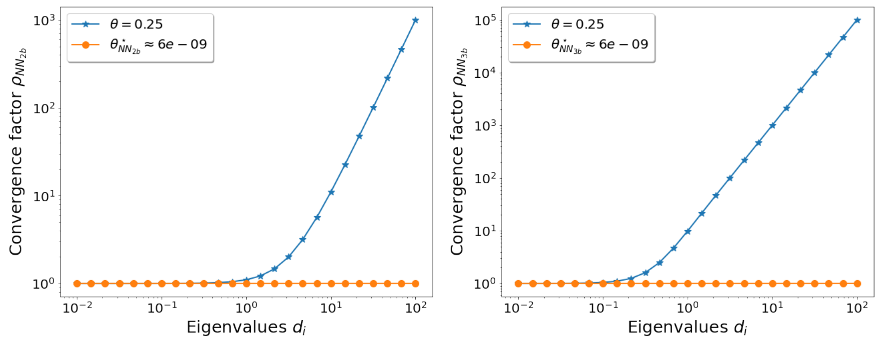


Figure 2: Convergence factor with $\theta = 0.25$ of $\mathrm{NN}_{2b}$ and $\mathrm{NN}_{3b}$ as a function of the eigenvalues $d_i \in [10^{-2}, 10^2]$. Left: case A for $\mathrm{NN}_{2b}$. Right: case B for $\mathrm{NN}_{3b}$.

for large eigenvalues with the scale of $10^3$ for $\mathrm{NN}_{2b}$ and $10^5$ for $\mathrm{NN}_{3b}$. This corresponds to our estimate $4\nu\theta d_i^2$. By applying optimization[1], we find the optimal relaxation parameter is approximately zero for both algorithms in the test cases. As shown in our analysis, the best one can do is to choose $\theta = 0$ to compensate the bad large eigenvalue behavior, yet the algorithms are still divergent. Note that $\mathrm{NN}_{2b}$ and $\mathrm{NN}_{3b}$ in the case $\theta = 0$ are actually a classical Schwarz type algorithm, which does not converge without overlap. Therefore, $\mathrm{NN}_{2b}$ and $\mathrm{NN}_{3b}$ are not good algorithms and cannot be improved with relaxation.

## 4.2 Convergence factor of $\mathrm{NN}_{1a}$ with different $\theta$

The second test is dedicated to the most natural Neumann-Neumann algorithm $\mathrm{NN}_{1a}$. Based on our analysis, $\mathrm{NN}_{1a}$ is only a good smoother but not a good solver. Therefore, we choose some different relaxation parameters $\theta$ and show the behavior of the convergence factor as a function of the eigenvalues in Figure 3. For both test cases A and B, $\mathrm{NN}_{1a}$ has similar behavior for the tested parameters $\theta$. In the case $\theta = [0.8, 0.2]$ and $\theta = [1.2, 1.8]$, the convergence behavior is the same for large eigenvalues. Indeed, our analysis shows that $\lim_{d_i \to \infty} \rho_{\mathrm{NN}_{1a}} = \{|1-\theta_1|, |1-\theta_2|\}$, and in this case equals to 0.8 for both $\theta$. Furthermore, we observe that $\mathrm{NN}_{1a}$ is a good smoother with the choice $\theta = [1, 1]$. By using optimization, we find that the optimal relaxation parameter has the form that one goes to zero and the other one goes to two, yet with a poor convergence. Therefore, $\mathrm{NN}_{1a}$ can be a good smoother but not a good solver.

[1]We use in this paper the optimization toolbox *scipy.optimize.fmin* in python.

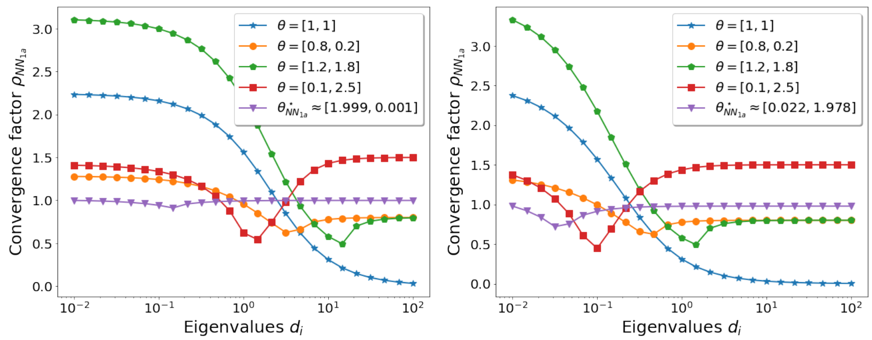


Figure 3: Convergence factor with different relaxation parameters $\theta$ of $\mathrm{NN}_{1\mathrm{a}}$ as a function of the eigenvalues $d_i \in [10^{-2}, 10^2]$. Left: case A. Right: case B.

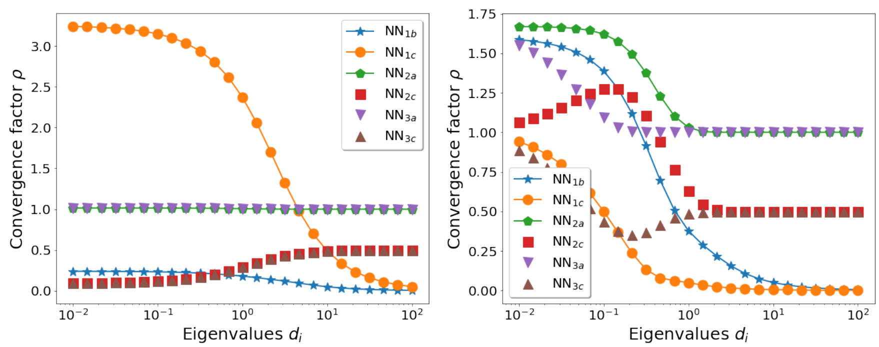


Figure 4: Convergence factor with $\theta = 1/2$ of the six algorithms as a function of the eigenvalues $d_i \in [10^{-2}, 10^2]$. Left: case A. Right: case B.

### 4.3 Convergence factor with $\theta = 1/2$

We now focus on the remaining six algorithms $\mathrm{NN}_{1\mathrm{b}}$, $\mathrm{NN}_{1\mathrm{c}}$, $\mathrm{NN}_{2\mathrm{a}}$, $\mathrm{NN}_{2\mathrm{c}}$, $\mathrm{NN}_{3\mathrm{a}}$ and $\mathrm{NN}_{3\mathrm{c}}$. Based on our analysis, all six algorithms have shown the potentiel of being a good solver, we thus compare them with a given relaxation parameter $\theta = 1/2$ in two test cases. Figure 4 shows the behavior of the convergence factor as a function of the eigenvalues for the six algorithms. In case A, we observe that $\mathrm{NN}_{2\mathrm{a}}$ and $\mathrm{NN}_{3\mathrm{a}}$ have identical behavior, and similar for $\mathrm{NN}_{2\mathrm{c}}$ and $\mathrm{NN}_{3\mathrm{c}}$. Indeed, as explained in our analysis, the convergence factors are the same in case A for $\mathrm{NN}_{2\mathrm{a}}$ and $\mathrm{NN}_{3\mathrm{a}}$, and also for $\mathrm{NN}_{2\mathrm{c}}$ and $\mathrm{NN}_{3\mathrm{c}}$. Furthermore, $\mathrm{NN}_{1\mathrm{b}}$ and $\mathrm{NN}_{1\mathrm{c}}$ have similar behavior for large eigenvalues, which has also been pointed out in our analysis. And as expected, these two algorithms are good smoothers with $\theta = 1/2$. In particular, $\mathrm{NN}_{1\mathrm{b}}$ outperforms the other five algorithms in case A, that is both a good smoother and solver. However, this changes in case B. More precisely, $\mathrm{NN}_{2\mathrm{a}}$ and $\mathrm{NN}_{3\mathrm{a}}$ have rather a symmetric behavior, as well as $\mathrm{NN}_{2\mathrm{c}}$ and $\mathrm{NN}_{3\mathrm{c}}$. And as shown in our analysis, both $\mathrm{NN}_{2\mathrm{a}}$ and $\mathrm{NN}_{3\mathrm{a}}$ have the

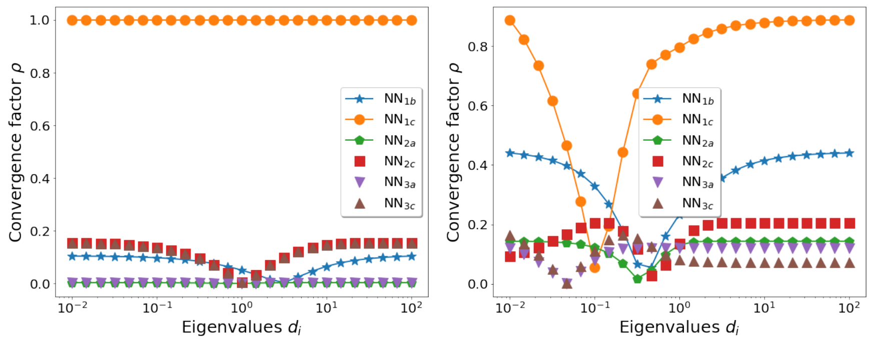


Figure 5: Convergence factor with optimal relaxation parameter $\theta^\star$ of the six algorithms as a function of the eigenvalues $d_i \in [10^{-2}, 10^2]$. Left: case A. Right: case B.

same behavior for large eigenvalues, and also $\mathrm{NN_{2c}}$ and $\mathrm{NN_{3c}}$. Moreover, $\mathrm{NN_{1b}}$ and $\mathrm{NN_{1c}}$ are both good smoothers, and $\mathrm{NN_{1c}}$ has a better performance than $\mathrm{NN_{1b}}$ this time.

### 4.4 Convergence factor with optimal $\theta$

We then show the convergence behavior of each algorithm using their optimal relaxation parameter $\theta^\star$ determined by optimization. Figure 5 shows the behavior of the convergence factor as a function of the eigenvalues for the six algorithms. In case A, $\mathrm{NN_{2a}}$ and $\mathrm{NN_{3a}}$ have once again identical behavior. Indeed, their convergence factors are the same in case A, and both $\mathrm{NN_{2a}}$ and $\mathrm{NN_{3a}}$ have the same optimal relaxation parameter $\theta^\star_{\mathrm{NN_{2a}}} = \theta^\star_{\mathrm{NN_{3a}}}$, which corresponds to the theoretical value $\theta^*_{\mathrm{NN_{2a}}} \approx 0.249$ as determined by (43). For the same reason, we observe the same behavior for $\mathrm{NN_{2c}}$ and $\mathrm{NN_{3c}}$, where the optimal relaxation parameter $\theta^\star_{\mathrm{NN_{2c}}} = \theta^\star_{\mathrm{NN_{3c}}} = \theta^*_{\mathrm{NN_{2c}}} \approx 0.385$ as determined by (47). As for $\mathrm{NN_{1b}}$, we find that the optimal relaxation parameter $\theta^\star_{\mathrm{NN_{1b}}} = \theta^*_{\mathrm{NN_{1b}}} \approx 0.446$ as determined by (30). However, the optimal relaxation parameter for $\mathrm{NN_{1c}}$ is $\theta^\star_{\mathrm{NN_{1c}}} \approx 0$, which cannot be determined by (37). As explained in our analysis, the term $E_i - \nu^{-1} F_i$ in (36) is negative in case A, thus the best option is to choose $\theta = 0$ which becomes then a Schwarz type algorithm without overlap. In general, all algorithms except $\mathrm{NN_{1c}}$ have very good performance in case A, and both $\mathrm{NN_{2a}}$ and $\mathrm{NN_{3a}}$ outperform the others with a convergence factor around $10^{-3}$. Once again, the behavior of the six algorithms becomes much different in case B. While $\mathrm{NN_{1c}}$ diverges in case A, it converges in the test case B with the optimal relaxation parameter $\theta^\star_{\mathrm{NN_{1c}}} = \theta^*_{\mathrm{NN_{1c}}} \approx 0.944$ as determined by (37). $\mathrm{NN_{1b}}$ rather keeps a similar performance with the optimal relaxation parameter $\theta^\star_{\mathrm{NN_{1b}}} = \theta^*_{\mathrm{NN_{1b}}} \approx 0.278$ as determined by (30). $\mathrm{NN_{2a}}$ and $\mathrm{NN_{3a}}$ also have the same optimal relaxation parameter $\theta^\star_{\mathrm{NN_{2a}}} = \theta^\star_{\mathrm{NN_{3a}}} = \theta^*_{\mathrm{NN_{2a}}} \approx 0.214$ as determined by (43). However, for $\mathrm{NN_{2c}}$ and $\mathrm{NN_{3c}}$, the optimal relaxation parameter of $\theta^\star_{\mathrm{NN_{2c}}} \approx 0.265$ is rather different from $\theta^\star_{\mathrm{NN_{3c}}} \approx 0.307$, and both are different from the value determined by (47) using equioscillation $\theta^*_{\mathrm{NN}_{2c}} \approx 0.285$. Indeed, $\mathrm{NN}_{2c}$ rather equioscillates the convergence value between large eigenvalues with some eigenvalue in the interval $[0.1, 1]$, whereas $\mathrm{NN}_{3c}$ equioscillates the convergence value between small eigenvalues with some eigenvalue in the interval $[0.1, 1]$. In general, all six algorithms converge in case B, $\mathrm{NN}_{2a}$ and $\mathrm{NN}_{3a}$ still outperform the others with $\mathrm{NN}_{3a}$ slightly better than $\mathrm{NN}_{2a}$.

## 5 Conclusion

We introduced and analyzed nine new time domain decomposition methods based on Neumann-Neumann techniques for parabolic optimal control problems. Our analysis shows that the Neumann correction step and the update step should be well-adjusted to the Dirichlet step to avoid potential divergence. Moreover, while it seems natural at first glance to preserve the forward-backward structure in the time subdomains as well, there are better choices that lead to substantially faster algorithms which can still be identified to be of forward-backward structure using changes of variables. We also found many interesting mathematical connections between these algorithms, for instance the algorithms in Categories II and III have rather similar convergence behavior. In terms of the performance, $\mathrm{NN}_{2b}$ and $\mathrm{NN}_{3b}$ are bad algorithms, the most natural algorithm $\mathrm{NN}_{1a}$ is rather a good smoother, and $\mathrm{NN}_{2a}$ and $\mathrm{NN}_{3a}$ with optimized relaxation parameter are much faster than the other algorithms and can be considered as highly efficient solvers.

Our study was restricted to the two subdomain case, but the algorithms can all naturally be written for many subdomains, and then one can also run them in parallel. They can also be used for more general parabolic constraints than the heat equation. Extensive numerical results will appear elsewhere.

## References

[1] Dirk Abbeloos, Moritz Diehl, Michael Hinze, and Stefan Vandewalle. Nested multigrid methods for time-periodic, parabolic optimal control problems. *Computing and Visualization in Science*, 14:27–38, 2011.

[2] Alessandro Alla and Stefan Volkwein. Asymptotic stability of POD based model predictive control for a semilinear parabolic PDE. *Advances in Computational Mathematics*, 41:1073–1102, 2015.

[3] Petter E. Bjørstad and Olof B. Widlund. Iterative methods for the solution of elliptic problems on regions partitioned into substructures. *SIAM Journal on Numerical Analysis*, 23(6):1097–1120, 1986.

[4] A. Borzì and V. Schulz. *Computational Optimization of Systems Governed by Partial Differential Equations*. Society for Industrial and Applied Mathematics, 2011.

[5] A. Bünger, S. Dolgov, and M. Stoll. A low-rank tensor method for PDE-constrained optimization with isogeometric analysis. *SIAM Journal on Scientific Computing*, 42(1):A140–A161, 2020.

[6] M. Emmett and M. Minion. Toward an efficient parallel in time method for partial differential equations. *Communications in Applied Mathematics and Computational Science*, 7(1):105 – 132, 2012.

[7] R. D. Falgout, S. Friedhoff, Tz. V. Kolev, S. P. MacLachlan, and J. B. Schroder. Parallel time integration with multigrid. *SIAM Journal on Scientific Computing*, 36(6):C635–C661, 2014.

[8] Liang Fang, Stefan Vandewalle, and Johan Meyers. A parallel-in-time multiple shooting algorithm for large-scale PDE-constrained optimal control problems. *Journal of Computational Physics*, 452:110926, 2022.

[9] Charbel Farhat and Marion Chandesris. Time-decomposed parallel time-integrators: theory and feasibility studies for fluid, structure, and fluid–structure applications. *International Journal for Numerical Methods in Engineering*, 58(9):1397–1434, 2003.

[10] Charbel Farhat and Francois-Xavier Roux. A method of finite element tearing and interconnecting and its parallel solution algorithm. *International Journal for Numerical Methods in Engineering*, 32(6):1205–1227, 1991.

[11] M. J. Gander. 50 years of time parallel time integration. In T. Carraro, M. Geiger, S. Körkel, and R. Rannacher, editors, *Multiple Shooting and Time Domain Decomposition Methods*, pages 69–114. Springer, Heidelberg, 2015.

[12] Martin J. Gander and Felix Kwok. Schwarz methods for the time-parallel solution of parabolic control problems. In Thomas Dickopf, Martin J. Gander, Laurence Halpern, Rolf Krause, and Luca F. Pavarino, editors, *Domain Decomposition Methods in Science and Engineering XXII*, pages 207–216, Cham, 2016. Springer International Publishing.

[13] Martin J. Gander, Felix Kwok, and Julien Salomon. Paraopt: A parareal algorithm for optimality systems. *SIAM Journal on Scientific Computing*, 42(5):A2773–A2802, 2020.

[14] Martin J. Gander and Liu-Di Lu. New time domain decomposition methods for parabolic optimal control problems I: Dirichlet-Neumann and Neumann-Dirichlet algorithms. Accepted with minor revision in SIAM Journal on Numerical Analysis, 2023.

[15] Sebastian Götschel and Michael L. Minion. An efficient parallel-in-time method for optimization with parabolic PDEs. *SIAM Journal on Scientific Computing*, 41(6):C603–C626, 2019.

[16] Max D. Gunzburger and Angela Kunoth. Space-time adaptive wavelet methods for optimal control problems constrained by parabolic evolution equations. *SIAM Journal on Control and Optimization*, 49(3):1150–1170, 2011.

[17] W. Hackbusch. Numerical solution of linear and nonlinear parabolic control problems. In Alfred Auslender, Werner Oettli, and Josef Stoer, editors, *Optimization and Optimal Control*, pages 179–185, Berlin, Heidelberg, 1981. Springer Berlin Heidelberg.

[18] Laurence Halpern and Jérémy Szeftel. Optimized and quasi-optimal Schwarz waveform relaxation for the one-dimensional Schrödinger equation. *Mathematical Models and Methods in Applied Sciences*, 20(12):2167–2199, 2010.

[19] Matthias Heinkenschloss. A time-domain decomposition iterative method for the solution of distributed linear quadratic optimal control problems. *Journal of Computational and Applied Mathematics*, 173(1):169–198, 2005.

[20] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*. Springer Dordrecht, 2009.

[21] Laura Iapichino, Stefan Trenz, and Stefan Volkwein. Reduced-order multiobjective optimal control of semilinear parabolic problems. In Bülent Karasözen, Murat Manguoğlu, Münevver Tezer-Sezgin, Serdar Göktepe, and Ömür Uğur, editors, *Numerical Mathematics and Advanced Applications ENUMATH 2015*, pages 389–397, Cham, 2016. Springer International Publishing.

[22] Eileen Kammann, Fredi Tröltzsch, and Stefan Volkwein. A posteriori error estimation for semilinear parabolic optimal control problems with application to model reduction by POD. *ESAIM: Mathematical Modelling and Numerical Analysis*, 47(2):555–581, 2013.

[23] M. Kollmann, M. Kolmbauer, U. Langer, M. Wolfmayr, and W. Zulehner. A robust finite element solver for a multiharmonic parabolic optimal control problem. *Computers & Mathematics with Applications*, 65(3):469–486, 2013. Efficient Numerical Methods for Scientific Applications.

[24] K. Kunisch, S. Volkwein, and L. Xie. HJB-POD-based feedback design for the optimal control of evolution problems. *SIAM Journal on Applied Dynamical Systems*, 3(4):701–722, 2004.

[25] Felix Kwok. On the time-domain decomposition of parabolic optimal control problems. In Chang-Ock Lee, Xiao-Chuan Cai, David E. Keyes, Hyea Hyun Kim, Axel Klawonn, Eun-Jae Park, and Olof B. Widlund, editors, *Domain Decomposition Methods in Science and Engineering XXIII*, pages 55–67, Cham, 2017. Springer International Publishing.

[26] E. Lelarasmee, A. E. Ruehli, and A. L. Sangiovanni-Vincentelli. The waveform relaxation method for time-domain analysis of large scale integrated circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 1(3):131–145, 1982.

[27] Buyang Li, Jun Liu, and Mingqing Xiao. A new multigrid method for unconstrained parabolic optimal control problems. *Journal of Computational and Applied Mathematics*, 326:358–373, 2017.

[28] J.-L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations.* 170. Springer-Verlag Berlin Heidelberg, 1 edition, 1971.

[29] Jacques-Louis Lions, Yvon Maday, and Gabriel Turinici. A parareal in time procedure for the control of partial differential equations. *Comptes Rendus Mathematique*, 335(4):387–392, 2002.

[30] Fredi Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, volume 112. Graduate Studies in Mathematics, 2010.

[31] Sergey Repin Ulrich Langer and Monika Wolfmayr. Functional a posteriori error estimates for time-periodic parabolic optimal control problems. *Numerical Functional Analysis and Optimization*, 37(10):1267–1294, 2016.

# A Pair transmission conditions

Let us consider a modified algorithm $\mathrm{NN}_{2\mathrm{a}}$, that is, we first solve the Dirichlet step

$$\begin{cases} \begin{pmatrix} \dot{z}_{1,i}^k \\ \dot{\mu}_{1,i}^k \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} z_{1,i}^k \\ \mu_{1,i}^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_1, \\ z_{1,i}^k(0) = 0, \\ z_{1,i}^k(\alpha) = f_{\alpha,i}^{k-1}, \end{cases}$$

$$\begin{cases} \begin{pmatrix} \dot{z}_{2,i}^k \\ \dot{\mu}_{2,i}^k \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} z_{2,i}^k \\ \mu_{2,i}^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_2, \\ z_{2,i}^k(\alpha) = g_{\alpha,i}^{k-1}, \\ \mu_{2,i}^k(T) + \gamma z_{2,i}^k(T) = 0, \end{cases}$$

and then correct the result by the Neumann step

$$\begin{cases} \begin{pmatrix} \dot\psi_{1,i}^k \\ \dot\phi_{1,i}^k \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} \psi_{1,i}^k \\ \phi_{1,i}^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_1, \\ \psi_{1,i}^k(0) = 0, \\ \dot\psi_{1,i}^k(\alpha) = \dot z_{1,i}^k(\alpha) - \dot z_{2,i}^k(\alpha), \end{cases}$$

$$\begin{cases} \begin{pmatrix} \dot\psi_{2,i}^k \\ \dot\phi_{2,i}^k \end{pmatrix} + \begin{pmatrix} d_i & -\nu^{-1} \\ -1 & -d_i \end{pmatrix} \begin{pmatrix} \psi_{2,i}^k \\ \phi_{2,i}^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ in } \Omega_2, \\ \dot\psi_{2,i}^k(\alpha) = \dot z_{2,i}^k(\alpha) - \dot z_{1,i}^k(\alpha), \\ \phi_{2,i}^k(T) + \gamma\psi_{2,i}^k(T) = 0. \end{cases}$$

and update the transmission condition by

$$f_{\alpha,i}^k := f_{\alpha,i}^{k-1} - \theta_1\big(\psi_{1,i}^k(\alpha) + \psi_{2,i}^k(\alpha)\big), \quad g_{\alpha,i}^k := g_{\alpha,i}^{k-1} - \theta_2\big(\psi_{1,i}^k(\alpha) + \psi_{2,i}^k(\alpha)\big),$$

with $\theta_1, \theta_2 > 0$. Following the same analysis as in Section 3.2.1, we find,

$$\begin{pmatrix} f_{\alpha,i}^k \\ g_{\alpha,i}^k \end{pmatrix} = \begin{pmatrix} 1-\theta_1 E_i & -\theta_1 F_i \\ -\theta_2 E_i & 1-\theta_2 F_i \end{pmatrix} \begin{pmatrix} f_{\alpha,i}^{k-1} \\ g_{\alpha,i}^{k-1} \end{pmatrix}.$$

In particular, the eigenvalues of the iteration matrix are 1 and $1-(\theta_1 E_i + \theta_2 F_i)$. Thus, the modified algorithm $\mathrm{NN}_{2\mathrm{a}}$ does not converge in this form. This divergence still stays even by considering the update step (8) for the pair transmission conditions. More generally, we have the same behavior for $\mathrm{NN}_{2\mathrm{b}}$, $\mathrm{NN}_{2\mathrm{c}}$, $\mathrm{NN}_{3\mathrm{a}}$, $\mathrm{NN}_{3\mathrm{b}}$ and $\mathrm{NN}_{3\mathrm{c}}$, if we keep a pair of transmission conditions $(f_{\alpha,i}^k, g_{\alpha,i}^k)$.