An Overlapping Waveform Relaxation Preconditioner for Economic Optimal Control Problems With State Constraints

Gabriele Ciaramella and Luca Mechelli

1 Introduction

This work is concerned with the numerical solution of so-called economic optimal control problems of the parabolic type. Let $\Omega = (-1, 1)$, $T > 0$ and $\mathcal{U} := L^2(0, T; L^2(\Omega))$ endowed with its norm $\| \cdot \|_{\mathcal{U}}$. We want to solve

$$\min_{\mathcal{U}\times\mathcal{U}} \mathcal{J}(u, w) := \frac{1}{2}\|u\|^2_{\mathcal{U}} + \frac{1}{2}\|w\|^2_{\mathcal{U}}, \tag{1a}$$

subject to the PDE-constraint

$$\begin{aligned} y_t(t, x) - \Delta y(t, x) &= f(t, x) + u(t, x), && \text{in } (0, T) \times \Omega,\\ y(t, -1) = y(t, 1) &= 0, && \text{in } (0, T),\\ y(0, x) &= y_\circ(x), && \text{in } \Omega, \end{aligned} \tag{1b}$$

with $y_\circ \in L^2(\Omega)$ and $f \in \mathcal{U}$, and to mixed control-state constraints

$$|u(t, x)| \leq c_u, \quad |y(t, x) + \varepsilon w(t, x)| \leq c_y(t), \quad \text{in } (0, T) \times \Omega, \tag{1c}$$

where $c_u, \varepsilon > 0$ and $c_y \in L^2(0, T)$ with $c_y(t) > 0$ for $t \in (0, T)$. Problem (1) is related to the virtual control approach [6, 8, 9], which is a regularization technique for pointwise state-constrained problems. Under further assumptions on w, in fact, one can show that, as $\varepsilon \to 0$, the solution to (1) converges to the one of the same optimal control problem with (1c) replaced by $|u(t, x)| \leq c_u$ and $|y(t, x)| \leq c_y(t)$ in $(0, T) \times \Omega$; see, e.g., [8]. Note that there are no weights in front of the control norms in (1a). This is because of the regularization parameter ε, which is also used

Gabriele Ciaramella
Politecnico di Milano e-mail: gabriele.ciaramella@polimi.it

Luca Mechelli
Universität Konstanz e-mail: luca.mechelli@uni-konstanz.de

to tune the magnitude of the controls u and w. For example, the smaller is ε, the larger is $\|w\|_{\mathcal{U}}$. In contrast to classical optimal control problems, where the goal is to reach a precise target configuration, the focus of (1) is to find minimum-energy feasible controls such that the state solution to (1b) satisfies the bounds (1c). This difference is particularly evident in the cost functional $\mathcal{J}$ in (1a), where only the norm squared of the controls are considered, instead of typical tracking-type terms. For these reasons, problems of the type (1) are called economic optimal control problems. A typical example is the optimal heating and cooling of residual buildings [8]. Note that, for any given $u \in \mathcal{U}$, the state equation (1b) admits a unique (weak) solution $y = y(u) \in W(0,T) := \{\varphi \in L^2(0,T;H^1(\Omega)) \big| \varphi_t \in L^2(0,T;H^{-1}(\Omega))\}$; see, e.g., [10, 9]. We assume that the admissible set $\mathcal{U}_{\mathsf{ad}}^\varepsilon$ has non-empty interior, where $\mathcal{U}_{\mathsf{ad}}^\varepsilon := \{(u,w) \in \mathcal{U} \times \mathcal{U} \big| u \text{ and } y(u) + \varepsilon w \text{ satisfies (1c)}\} \subset \mathcal{U} \times \mathcal{U}$. This guarantees that (1) admits a unique solution $(\bar{u}, \bar{w}) \in \mathcal{U}_{\mathsf{ad}}^\varepsilon$ [10]. The first-order necessary and sufficient optimality system [9, 10] of problem (1) is

$$
\begin{aligned}
y_t(t,x) - \Delta y(t,x) &= \mathcal{P}(q(t,x)) + f(t,x), && \text{in } (0,T) \times \Omega,\\
y(t,-1) = y(t,1) &= 0, && \text{in } (0,T),\\
y(0,x) &= y_\circ(x), && \text{in } \Omega,\\
q_t(t,x) + \Delta q(t,x) &= \mathcal{Q}^\varepsilon(y(t,x)), && \text{in } (0,T) \times \Omega,\\
q(t,-1) = q(t,1) &= 0, && \text{in } (0,T),\\
q(T,x) &= 0, && \text{in } \Omega,
\end{aligned}
\tag{2}
$$

where $\mathcal{Q}^\varepsilon(y(t,x)) := \frac{1}{\varepsilon^2}(\max\{y(t,x) - c_y(t), 0\} + \min\{y(t,x) + c_y(t), 0\})$ and $\mathcal{P}(q(t,x)) := \max\{-c_u, \min\{c_u, q(t,x)\}\}$, for all $(t,x) \in (0,T) \times \Omega$, with q the so-called adjoint variable. The pair $(\bar{y}, \bar{q})$ is the solution to (2) if and only if $(\bar{u}(t,x), \bar{w}(t,x)) = (\mathcal{P}(\bar{q}(t,x)), -\varepsilon \mathcal{Q}^\varepsilon(\bar{y}(t,x)))$, for $(t,x) \in (0,T) \times \Omega$, is the optimal solution to (1). System (2) can be rewritten in the form

$$
\mathcal{F}(y,q) = 0 \tag{3}
$$

and thus solved by using a semismooth Newton method; see, e.g., [9, 5].

As shown in [8], the semismooth Newton method lacks of convergence if the parameter ε is not sufficiently large. This is, however, in contrast with typical applications, where a sufficiently small ε is required [8, 6]. The goal of this paper is to tackle this problem by using a nonlinear preconditioning technique based on an overlapping optimized waveform-relaxation method (WRM) characterized by Robin transmission conditions [2, 3]. To the best of our knowledge, nonlinear preconditioning techniques have never been used for economic control problems. Therefore, this work aims to provide a first concrete study in order to show the applicability of WRM-based nonlinear preconditioners for this class of optimization problems. In particular, our goal is to assess the convergence behavior of the WRM nonlinear preconditioned Newton and its robustness against the regularization parameter ε. Our studies show that appropriate choices of the overlap L and of the Robin parameter p lead to a preconditioned Newton method with a robust convergence with respect to ε. Let us also mention that for elliptic optimal control problems, it is possible to consider different transmission conditions; see, e.g., [1, 4].

The paper is organized as follows. In Section 2, we introduce the WRM and present the algorithm for the proposed preconditioned generalized Newton. In Section 3, we report two numerical experiments that show the convergence behavior of the proposed computational framework in relation of the parameters characterizing problem (1) and the optimized WRM.

2 The waveform-relation and the preconditioned generalized Newton methods

Let Ω be decomposed into two overlapping subdomains $\Omega_1 = (-1, L)$ and $\Omega_2 = (-L, 1)$, where $2L \in (0, 1)$ is the size of the overlap. Moreover, let $p > 0$ and consider the operator $\mathcal{R}_j$ defined as $\mathcal{R}_j(y) := y_x + (-1)^{3-j} p y$ for $j = 1, 2$. The WRM consists in iteratively solving, for $n \in \mathbb{N}$, $n \geq 1$, the system

$$
\begin{aligned}
y_t^{j,n}(t,x) - \Delta y^{j,n}(t,x) &= \mathcal{P}(q^{j,n}(t,x)) + f(t,x), && \text{in } (0,T)\times\Omega_j, && \text{(4a)}\\
y^{j,n}(t,(-1)^j) &= 0, && \text{in } (0,T), && \text{(4b)}\\
\mathcal{R}_j(y^{j,n})(t,(-1)^{3-j}L) &= \mathcal{R}_j(y^{3-j,n-1})(t,(-1)^{3-j}L), && \text{in } (0,T), && \text{(4c)}\\
y^{j,n}(0,x) &= y_\circ(x), && \text{in } \Omega_j, && \text{(4d)}\\
q_t^{j,n}(t,x) + \Delta q^{j,n}(t,x) &= \mathcal{Q}^{\varepsilon}(y^{j,n}(t,x)), && \text{in } (0,T)\times\Omega_j, && \text{(4e)}\\
q^{j,n}(t,(-1)^j) &= 0, && \text{in } (0,T), && \text{(4f)}\\
\mathcal{R}_j(q^{j,n})(t,(-1)^{3-j}L) &= \mathcal{R}_j(q^{3-j,n-1})(t,(-1)^{3-j}L), && \text{in } (0,T), && \text{(4g)}\\
q^{j,n}(T,x) &= 0, && \text{in } \Omega_j, && \text{(4h)}
\end{aligned}
$$

for $j = 1, 2$. We show first the well-posedness of the method.

Theorem 1 *Let $g_y^1, g_y^2, g_q^1, g_q^2 \in H^{1/4}(0,T)$ be initialization functions for the WRM, i.e., $\mathcal{R}_j(y^{j,1})(t,(-1)^{3-j}L) = g_y^j(t)$ and $\mathcal{R}_j(q^{j,1})(t,(-1)^{3-j}L) = g_q^j(t)$ for $t \in (0,T)$, with compatibility conditions $g_y^j(0) = \mathcal{R}_j(y_\circ)(t,(-1)^{3-j}L)$ and $g_q^j(0) = 0$ for $j = 1, 2$. Then the WRM* (4) *is well-posed.*

Proof For $j = 1, 2$, we define $H_j^{2,1} := L^2(0,T; H^2(\Omega_j)) \times H^1(0,T; L^2(\Omega_j))$ and $\mathcal{U}_j = L^2(0,T; L^2(\Omega_j))$. For given $g_y^j, g_q^j \in H^{1/4}(0,T)$, system (4) is the optimality system of an optimal control problem, which seeks to minimize $\mathcal{J}_{\text{aux}}(u^j, w^j) = \frac{1}{2}\|u^j\|_{\mathcal{U}_j}^2 + \frac{1}{2}\|w^j\|_{\mathcal{U}_j}^2 + \int_0^T g_q^j(t) y^j(t,(-1)^{3-j}L)\mathrm{d}t$, subject to the state equation (4a)-(4d). These auxiliary optimal control problems admit a unique optimal solution $(\bar{u}^j, \bar{w}^j) \in \mathcal{U}_j \times \mathcal{U}_j$ for $j = 1, 2$ and their optimality systems are uniquely solvable by $(\bar{y}^j, \bar{q}^j) \in H_j^{2,1} \times H_j^{2,1}$ such that

$$(\bar{u}^j(t,x),\bar{w}^j(t,x)) = (\mathcal{P}(\bar{q}^j(t,x)), -\varepsilon \mathcal{Q}^\varepsilon(\bar{y}^j(t,x))), \quad \text{in } (0,T)\times\Omega_j.$$

For more details see [10, 7, 3]. This proves well-posedness of the WRM for $n = 1$ and $j = 1, 2$. By iteratively applying the previous arguments is then easy to show that the WRM is well-posed for $n > 1$, because $y^{j,1}((-1)^j L)$, $y_x^{j,1}((-1)^j L)$, $q^{j,1}((-1)^j L)$, $q_x^{j,1}((-1)^j L) \in L^2(0,T)$. □

Theorem 1 implies that (4) admits a unique solution $(y^{j,n}, p^{j,n}) \in H_j^{2,1} \times H_j^{2,1}$ for $j = 1, 2$ and $n \geq 1$. Note that, at each iteration of the WRM, the solution at iteration n depends on the one at iteration $n-1$. Therefore, we can define the solution mappings $\mathcal{S}_j : H_{3-j}^{2,1} \times H_{3-j}^{2,1} \to H_j^{2,1} \times H_j^{2,1}$ for $j = 1, 2$ as

$$\begin{aligned} &(y^1, q^1) = \mathcal{S}_1(y^2, q^2) \text{ solves (4) for } j = 1,\ y^{2,n-1} = y^2 \text{ and } q^{2,n-1} = q^2, \\ &(y^2, q^2) = \mathcal{S}_2(y^1, q^1) \text{ solves (4) for } j = 2,\ y^{1,n-1} = y^1 \text{ and } q^{1,n-1} = q^1, \end{aligned} \tag{5}$$

and the preconditioned form of (3) as

$$\mathcal{F}_P(y^1, q^1, y^2, q^2) = (\mathcal{F}_1(y^1, q^1, y^2, q^2), \mathcal{F}_2(y^1, q^1, y^2, q^2)) = 0, \tag{6}$$

where $\mathcal{F}_j(y^1, q^1, y^2, q^2) = (y^j, q^j) - \mathcal{S}_j(y^{3-j}, q^{3-j})$, for $j = 1, 2$. To solve (6), we apply a generalized Newton method. To do so, we assume that the maps $\mathcal{S}_j$, $j = 1, 2$, admit derivative[1] $D\mathcal{S}_j$. This allows us to characterize the derivative $D\mathcal{F}_P$ and its application to a direction $\mathbf{d}^{3-j} = (d_y^{3-j}, d_q^{3-j}) \in H_{3-j}^{2,1} \times H_{3-j}^{2,1}$, which is needed for the generalized Newton method. Let $z^j := (y^j, q^j) \in H_j^{2,1} \times H_j^{2,1}$ for $j = 1, 2$. Thus, we have that $z^j = \mathcal{S}_j(z^{3-j})$, according to the definition of the mapping $\mathcal{S}_j$ in (5). Moreover, we have that $\mathcal{F}_j(\mathcal{S}_j(z^{3-j}), z^{3-j}) = 0$. From this we formally obtain

$$D_1\mathcal{F}_j(\mathcal{S}_j(z^{3-j}), z^{3-j}) D\mathcal{S}_j(z^{3-j})(\mathbf{d}^{3-j}) + D_2\mathcal{F}_j(\mathcal{S}_j(z^{3-j}), z^{3-j})(\mathbf{d}^{3-j}) = 0,$$

which leads to $D\mathcal{S}_j(y^{3-j}, q^{3-j})(\mathbf{d}^{3-j}) = (\widetilde{y}^j, \widetilde{q}^j)$ where $(\widetilde{y}^j, \widetilde{q}^j)$ solves

$$\begin{aligned} \widetilde{y}_t^j(t,x) - \Delta \widetilde{y}^j(t,x) &= \widetilde{q}^j(t,x)\chi_{\mathcal{I}(q^j)}(t,x), && \text{in } (0,T)\times\Omega_j, \\ \widetilde{y}^j(t,(-1)^j) &= 0, && \text{in } (0,T), \\ \mathcal{R}_j(\widetilde{y}^j)(t,(-1)^{3-j}L) &= \mathcal{R}_j(d_y^{3-j})(t,(-1)^{3-j}L), && \text{in } (0,T), \\ \widetilde{y}^j(0,x) &= 0, && \text{in } \Omega_j, \end{aligned} \tag{7a}$$

[1] Since the functions $\mathcal{S}_j$ are implicit functions of semismooth functions, one cannot directly invoke the implicit function theorem to obtain the desired regularity. Hence, investigating the existence and regularity of $D\mathcal{S}_j$ requires a detailed theoretical analysis, which is beyond the scope of this short manuscript.

$$
\begin{aligned}
\widetilde{q}^j_t(t,x) + \Delta\widetilde{q}^j(t,x) &= \frac{\widetilde{y}^j(t,x)}{\varepsilon^2}\chi_{\mathcal{A}(y^j)}(t,x), && \text{in } (0,T)\times\Omega_j,\\
\widetilde{q}^j(t,(-1)^j) &= 0, && \text{in } (0,T),\\
\mathcal{R}_j(\widetilde{q}^j)(t,(-1)^{3-j}L) &= \mathcal{R}_j(d_q^{3-j})(t,(-1)^{3-j}L), && \text{in } (0,T),\\
\widetilde{q}^{j,n}(T,x) &= 0, && \text{in } \Omega_j,
\end{aligned}
\tag{7b}
$$

for $j = 1, 2$, with $\chi_{\mathcal{I}(q^j)}$ and $\chi_{\mathcal{A}(y^j)}$ the characteristic functions of the sets

$$
\begin{aligned}
\mathcal{I}(q^j) &:= \{(t,x)\in(0,T)\times\Omega_j \,\big|\, |q^j(t,x)| \le c_u\},\\
\mathcal{A}(y^j) &:= \{(t,x)\in(0,T)\times\Omega_j \,\big|\, |y^j(t,x)| > c_y(t)\}.
\end{aligned}
$$

Note that (7) is a linearization of the WRM subproblems (4). Now, we can resume our preconditioned generalized Newton method in Algorithm 1.

Algorithm 1 WRM-preconditioned generalized Newton method

```
Data: Initial guess y^{j,0} and q^{j,0} for j = 1,2, tolerance τ.
Perform one WRM step to compute S_j(y^{3-j,0}, q^{3-j,0});
Assemble F_P(y^{1,0}, q^{1,0}, y^{2,0}, q^{2,0}) and set k = 0;
while ||F_P(y^{1,k}, q^{1,k}, y^{2,k}, q^{2,k})|| ≥ τ do
    Compute d^1, d^2 solving DF_P(y^1, q^1, y^2, q^2)(d^1, d^2) = −F_P(y^1, q^1, y^2, q^2)
    by using a matrix-free Krylov method, e.g., GMRES, and considering that
    DF_P(y^1, q^1, y^2, q^2)(d^1, d^2) = (d^1 − (ỹ^1, q̃^1), d^2 − (ỹ^2, q̃^2)), with (ỹ^j, q̃^j) solution
    to the linearized subproblems (7) for j = 1,2;
    Update (y^{j,k+1}, q^{j,k+1}) = (y^{j,k}, q^{j,k}) + d^j and set k = k + 1;
    Perform one WRM step to compute S_j(y^{3-j,k}, q^{3-j,k});
    Assemble F_P(y^{1,k}, q^{1,k}, y^{2,k}, q^{2,k});
end while
```

3 Numerical experiments

In this section, we study the behavior of the preconditioned generalized Newton method (Algorithm 1) and its robustness against the Robin parameter p, the regularization ε and the overlap L. It is well known that the convergence of the semismooth Newton method applied to (3) deteriorates fast for decreasing values of ε, since the solution approaches the one of a pure pointwise state-constrained problem, whose adjoint variable q lacks of L^2-regularity; cf. [10, 8]. The focus is on understanding if the WRM can be a valid (nonlinear) preconditioner and in which cases. We will perform two numerical experiments. In both tests we discretize the domain Ω with $n_x = 161$ points and we apply a centered finite-difference scheme. Furthermore, we consider $n_t = 21$ time discretization points and apply the implicit Euler method. The initial guesses $y^{j,0}$ and $q^{j,0}$ are chosen randomly but feasible, i.e. such that $\big(\mathcal{P}(q^{j,0}(t,x)), -\varepsilon Q^{\varepsilon}(y^{j,0}(t,x))\big) \in \mathcal{U}^{\varepsilon}_{\mathrm{ad}}$, since we noticed that choosing feasible initial guesses improves the convergence of the method. We set the stopping tol-

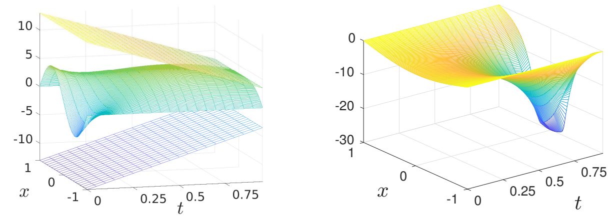

Fig. 1: Test1: Optimal state with bound c_y (left) and control (right) for $\varepsilon = 5 \times 10^{-4}$.

L	ε \ p	10^{-1}	5×10^{-2}	10^{-2}	5×10^{-3}	10^{-3}	5×10^{-4}
Δx	10^{-6}	4(5–2)	4(6–2)	5(12–2)	6(13–2)	7(35–2)	8(45–2)
Δx	10^{-4}	4(5–2)	4(6–2)	5(13–2)	6(13–2)	7(34–2)	8(45–2)
Δx	10^{-2}	4(6–2)	4(6–2)	5(11–2)	6(13–2)	7(30–2)	8(43–2)
Δx	10^{0}	5(4–2)	5(5–2)	5(9–2)	6(12–2)	max(112–2)	max(123–3)
Δx	10^{2}	6(4–2)	6(5–2)	8(8–2)	9(9–2)	6(22–2)	9(37–2)
Δx	10^{4}	6(5–2)	6(5–2)	9(7–2)	9(10–2)	8(23–2)	max(65–4)
Δx	10^{6}	6(5–2)	6(5–2)	9(7–2)	9(10–2)	max(33–2)	max(92–3)
$2\Delta x$	10^{-6}	4(5–2)	4(7–2)	5(11–2)	6(13–2)	7(39–2)	6(51–2)
$2\Delta x$	10^{-4}	4(5–2)	4(7–2)	5(11–2)	6(13–2)	7(41–2)	6(48–2)
$2\Delta x$	10^{-2}	4(6–2)	4(7–2)	5(12–2)	5(13–2)	7(23–2)	6(54–2)
$2\Delta x$	10^{0}	5(4–2)	5(6–2)	5(9–2)	6(11–2)	7(27–2)	max(107–3)
$2\Delta x$	10^{2}	6(4–2)	6(5–2)	8(8–2)	8(10–2)	8(26–2)	9(37–2)
$2\Delta x$	10^{4}	6(5–2)	6(5–2)	8(8–2)	9(10–2)	8(19–2)	9(41–2)
$2\Delta x$	10^{6}	6(5–2)	6(5–2)	8(8–2)	9(9–2)	8(19–2)	9(41–2)
$4\Delta x$	10^{-6}	4(5–2)	4(7–2)	5(11–2)	6(13–2)	6(30–2)	max(126–6)
$4\Delta x$	10^{-4}	4(5–2)	4(7–2)	5(11–2)	6(13–2)	6(30–2)	max(98–4)
$4\Delta x$	10^{-2}	4(5–2)	4(7–2)	5(12–2)	6(13–2)	6(30–2)	11(124–2)
$4\Delta x$	10^{0}	4(5–2)	4(6–2)	5(9–2)	6(11–2)	6(27–2)	max(152 – 5)
$4\Delta x$	10^{2}	6(4–2)	6(5–2)	8(8–2)	8(10–2)	10(23–2)	15(40 – 2)
$4\Delta x$	10^{4}	6(4–2)	6(5–2)	8(8–2)	8(10–2)	9(26–2)	max(183–3)
$4\Delta x$	10^{6}	6(4–2)	6(5–2)	8(8–2)	8(10–2)	9(26–2)	max(45–2)
Sem. New.		4	5	10	13	30	44

Table 1: Test1: Number of outer iterations (maximum number - minimum number of inner iterations) for preconditioned generalized Newton varying L, p and ε and number of iterations for the semismooth Newton applied to (3) (last row).

erance $\tau = 10^{-10}$ for the norm of the Newton residual (see Algorithm 1) and the maximum number of outer (inner) iterations to 200 (500). For the first test we choose $T = 1$, $y_\circ(x) = 5\sin(\pi x)$, $f(t,x) = 20$, $c_u = 30$ and $c_y(t) = 10(1-t) + 3$ for all $(t,x) \in (0,1) \times \Omega$. As one can see from Table 1, for a decreasing ε the number of iterations of the semismooth Newton method applied to (3) increases and its convergence deteriorates fast. On the contrary, the number of iterations of Algorithm 1is

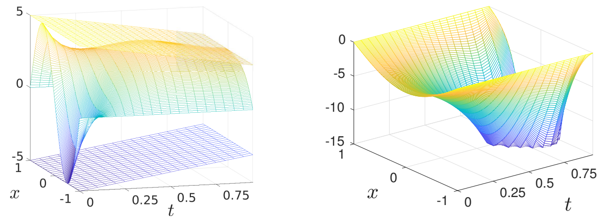

Fig. 2: Test2: Optimal state with bound c_y (left) and control (right) for $\varepsilon = 5 \times 10^{-4}$.

L	p \ ε	10^{-1}	5×10^{-2}	10^{-2}	5×10^{-3}	10^{-3}	5×10^{-4}
Δx	10^{-6}	5(5–2)	6(7–2)	10(10–2)	max(61–2)	max(102–2)	max(297–4)
Δx	10^{-4}	5(5–2)	6(7–2)	10(10–2)	max(32–2)	max(246–2)	max(145–2)
Δx	10^{-2}	5(5–2)	6(7–2)	8(10–2)	max(25–2)	max(max–2)	max(max–4)
Δx	10^{0}	5(5–2)	6(6–2)	6(10–2)	9(11–2)	max(122–4)	max(193–2)
Δx	10^{2}	6(4–2)	7(5–2)	9(8–2)	9(10–2)	9(20–2)	10(25–2)
Δx	10^{4}	6(4–2)	7(5–2)	9(8–2)	9(11–2)	11(20–2)	max(32–2)
Δx	10^{6}	6(4–2)	7(5–2)	9(8–2)	9(11–2)	11(20–2)	max(67–4)
$2\Delta x$	10^{-6}	5(6–2)	6(7–2)	12(11–2)	max(29–2)	max(123–2)	max(206–3)
$2\Delta x$	10^{-4}	5(6–2)	6(7–2)	12(11–2)	max(28–2)	max(91–2)	max(196–3)
$2\Delta x$	10^{-2}	5(6–2)	6(7–2)	11(11–2)	max(25–2)	max(max–4)	max(max–4)
$2\Delta x$	10^{0}	5(5–2)	6(6–2)	6(9–2)	7(10–2)	max(166–5)	max(183–2)
$2\Delta x$	10^{2}	6(4–2)	7(5–2)	8(8–2)	9(11–2)	9(20–2)	10(29–2)
$2\Delta x$	10^{4}	6(4–2)	7(5–2)	9(7–2)	9(11–2)	10(20–2)	9(26–2)
$2\Delta x$	10^{6}	6(4–2)	7(5–2)	9(7–2)	9(11–2)	10(19–2)	10(26–2)
$4\Delta x$	10^{-6}	5(5–2)	6(7–2)	10(11–2)	max(32–2)	max(313–4)	max(187–4)
$4\Delta x$	10^{-4}	5(5–2)	6(7–2)	10(11–2)	max(27–2)	max(145–4)	max(148–4)
$4\Delta x$	10^{-2}	6(5–2)	6(7–2)	9(11–2)	max(35–3)	max(296 – 4)	max(max – 4)
$4\Delta x$	10^{0}	5(5–2)	5(6–2)	6(8–2)	8(11–2)	max(136 – 3)	max(max – 3)
$4\Delta x$	10^{2}	6(4–2)	7(5–2)	6(8–2)	8(11–2)	11(20–2)	14(44–2)
$4\Delta x$	10^{4}	6(4–2)	7(5–2)	8(8–2)	8(11–2)	10(20–2)	12(26–2)
$4\Delta x$	10^{6}	6(4–2)	7(5–2)	8(8–2)	8(11–2)	10(20–2)	13(25–2)
Sem. New.		4	6	10	12	23	30

Table 2: Test2: Number of outer iterations (maximum number - minimum number of inner iterations) for preconditioned generalized Newton varying L, p and ε and number of iterations for the semismooth Newton applied to (3) (last row).

almost constant as ε varies (when it converges). Choosing $p = 10^2$ guarantees that the method is convergent for any choice of ε and L. In particular, for small ε, such as 10^{-3} and 5×10^{-4}, the speed-up in terms of number of iterations is also significant. According to Table 1, there are some combinations for which Algorithm 1reaches a maximum number of iterations (indicated in the tables with max). This issue can be related to the fact that $y^{j,k}$ and $q^{j,k}$ might become unfeasible during Algorithm 1and when traced to the interface of the other subdomain might cause oscillations. For the second test we choose $T = 1$, $y_\circ(x) = 5\sin(\pi x)$, $f(t,x) = 18$, $c_u = 15$ and

$c_y(t) = 2(1 - t) + 3$ for $(t, x) \in (0, 1) \times \Omega$. In this case, there are more points in the space-time domain for which both bounds become active (cf. Figures 1-2). This makes the problem even more difficult to be solved by the WRM, since its nonlinearities are more strongly activated. In Table 2, in fact, the number of cases for which Algorithm 1does not converge increases with respect to the first numerical experiment, particularly for ε small. We observe that transmission conditions of Dirichlet type and large-enough overlap L guarantee that the number of unfeasible points at the interface is significantly reduced, so that Algorithm 1converges. This confirms the previous remark on the importance of having feasible iterations. As a rule of thumb, if the regularization ε is small, we suggest to choose a sufficiently large parameter p (e.g., $p \geq 10^2$) so that the Dirichlet part of the transmission conditions of the WRM dominates the Neumann part. Note that, also in the second test, there always exists a combination of p and L for which Algorithm 1is faster than the semismooth Newton method, in particular for a small ε.

In conclusion, the WRM is a valid preconditioner for solving (3), although there are combinations of p and L for which the method may not converge. As observed, a crucial point for the convergence is to keep the iteration feasible. Preserving such a feasibility, together with other important aspects (e.g., multiple subdomains decomposition and the study of an optimal parameter p) will be the focus of a future work.

References

1. J.-D. Benamou. A domain decomposition method with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations. *SIAM J. Numer. Anal.*, 33(6):2401–2416, 1996.
2. V. Dolean, M. J. Gander, W. Kheriji, F. Kwok, and R. Masson. Nonlinear preconditioning: How to use a nonlinear Schwarz method to precondition Newton's method. *SIAM J. Sci. Comput.*, 38(6):A3357–A3380, 2016.
3. M. J. Gander and L. Halpern. Optimized Schwarz waveform relaxation methods for advection reaction diffusion problems. *SIAM J. Numer. Anal.*, 45(2):666–697, 2007.
4. M. Heinkenschloss and H. Nguyen. Neumann–Neumann domain decomposition preconditioners for linear-quadratic elliptic optimal control problems. *SIAM J. Sci. Comput.*, 28(3):1001–1028, 2006.
5. M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13(3):865–888, 2002.
6. K. Krumbiegel and A. Rösch. A virtual control concept for state constrained optimal control problems. *Comput. Optim. Appl.*, 43:213–233, 2009.
7. J. L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications (Vol II)*. Die Grundlehren der mathematischen Wissenschaften. Springer-Verlag Berlin Heidelberg, 1972.
8. L. Mechelli. *POD-based state-constrained economic Model Predictive Control of convection-diffusion phenomena*. PhD thesis, University of Konstanz, 2019.
9. L. Mechelli and S. Volkwein. POD-based economic optimal control of heat-convection phenomena. In M. Falcone, R. Ferretti, L. Grüne, and W. M. McEneaney, editors, *Numerical Methods for Optimal Control Problems*, pages 63–87, Cham, 2018. Springer International Publishing.
10. F. Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*. American Mathematical Society, 2010.