

On the nonlinear Dirichlet-Neumann method and preconditioner for Newton's method

F. Chaouqui, M. J. Gander, P. M. Kumbhar and T. Vanzan

1 Introduction

We consider a nonlinear Partial Differential Equation (PDE)

$$\mathcal{L}(u) = f \quad \text{in } \Omega, \quad u = g \quad \text{on } \partial\Omega, \quad (1)$$

where $\Omega \subset \mathbb{R}^d$ for $d \in \{1, 2, 3\}$ is an open bounded domain with a polygonal boundary $\partial\Omega$, and $f, g \in L^2(\Omega)$. We suppose that (1) admits a unique weak solution in some Hilbert space $u \in \mathcal{X}$ (e.g. $H^1(\Omega)$). For instance, for a quasilinear operator \mathcal{L} in divergence form, explicit assumptions can be found in [1] and references therein, see also [7, Chapter 8-9] and [5, Chapter 9]. Let us divide Ω into two nonoverlapping subdomains Ω_1 and Ω_2 and define $\Gamma_j = \partial\Omega_j \setminus \partial\Omega$, $j = 1, 2$. Let u_j be the restriction of u to Ω_j . The nonlinear Dirichlet-Neumann (DN) method starts from an initial guess λ^0 and computes for $n \geq 1$ until convergence

$$\begin{aligned} \mathcal{L}(u_1^n) &= f_1, & \text{in } \Omega_1, & & \mathcal{L}(u_2^n) &= f_2, & \text{in } \Omega_2 \\ u_1^n &= g_1, & \text{on } \partial\Omega_1 \setminus \Gamma, & & u_2^n &= g_2, & \text{on } \partial\Omega_2 \setminus \Gamma \\ u_1^n &= \lambda^n & \text{on } \Gamma, & & \mathcal{N}_2 u_2^n &= -\mathcal{N}_1 u_1^n & \text{on } \Gamma, \end{aligned} \quad (2)$$

where $\lambda^n = (1-\theta)\lambda^{n-1} + \theta u_{2|\Gamma}^{n-1}$, with $\theta \in (0, 1)$, $f_j := f|_{\Omega_j}$ and $g_j := g|_{\partial\Omega_j \setminus \Gamma}$ for $j = 1, 2$. The operators \mathcal{N}_j represent the outward nonlinear Neumann conditions that must be imposed on the interface Γ and are usually found through integration by parts of the

F. Chaouqui
Temple University, Philadelphia, USA, e-mail: faycal.chaouqui@temple.edu,

M. J. Gander
Université de Genève, Switzerland e-mail: martin.gander@unige.ch,

P. M. Kumbhar
Karlsruher Institut für Technologie, Germany e-mail: pratik.kumbhar@kit.edu,

T. Vanzan
École Polytechnique Fédérale de Lausanne, Switzerland e-mail: tommaso.vanzan@epfl.ch

variational formulation of the PDE. For instance, if $\mathcal{L}(u) = -\partial_x((1 + \alpha u^2)\partial_x u)$, then $\mathcal{N}_j u = (-1)^{j+1}(1 + \alpha u^2|_\Gamma)\partial_x u|_\Gamma$. For the well-posedness of the Dirichlet-Neumann method, we further assume that $\mathcal{N}_j u$ defines a bounded linear functional over \mathcal{X} .

System (2) can be formulated as an iteration over the substructured variable λ as

$$\lambda^n = G(\lambda^{n-1}) := (1 - \theta)\lambda^{n-1} + \theta \text{NtD}_2 \left(-\text{DtN}_1(\lambda^{n-1}, \psi_1), \psi_2 \right), \quad (3)$$

where $\psi_j = (f_j, g_j)$, $j = 1, 2$, represent the force term and boundary conditions, while the nonlinear Dirichlet-to-Neumann (DtN_j) and Neumann-to-Dirichlet operators (NtD_j) are defined as $\text{DtN}_j(\lambda, \psi_j) := \mathcal{N}_j u_j$, and $\text{NtD}_j(\phi, \psi_j) := v_j|_\Gamma$, with

$$\begin{aligned} \mathcal{L}(u_j^n) &= f_j & \text{in } \Omega_j, & & \mathcal{L}(v_j^n) &= f_j & \text{in } \Omega_j, \\ u_j^n &= g_j & \text{on } \partial\Omega_j \setminus \Gamma, & & v_j^n &= g_j & \text{on } \partial\Omega_j \setminus \Gamma, \\ u_j^n &= \lambda & \text{on } \Gamma & & \mathcal{N}_j v_j^n &= \phi & \text{on } \Gamma. \end{aligned} \quad (4)$$

If $u_{\text{ex}} \in H^1(\Omega)$ is the solution of (1), then it must have continuous Dirichlet trace and Neumann flux along the interface Γ . Defining $u_\Gamma := u_{\text{ex}}|_\Gamma$, $\phi := \mathcal{N}_1 u_{\text{ex}}|_\Gamma$ and using the operators DtN_j and NtD_j , these necessary properties are equivalent to

$$\text{DtN}_1(u_\Gamma, \psi_1) = -\text{DtN}_2(u_\Gamma, \psi_2), \quad \text{and} \quad \text{NtD}_1(\phi, \psi_1) = \text{NtD}_2(-\phi, \psi_2). \quad (5)$$

2 Nilpotent property and quadratic convergence

It is well known, see e.g. [9, 4], that if \mathcal{L} is linear and the subdomain decomposition is symmetric, then the DN method converges in one iteration for $\theta = 1/2$. Indeed, if \mathcal{L} is linear, one can work on the error equation, i.e. $\psi_j = 0$, and the symmetry of the decomposition is sufficient to guarantee $\text{DtN}_1(\cdot, 0) \equiv \text{DtN}_2(\cdot, 0)$, so that

$$\begin{aligned} \lambda^1 &= \frac{1}{2} \left(\lambda^0 + \text{NtD}_2 \left(-\text{DtN}_1(\lambda^0, 0), 0 \right), 0 \right) = \frac{1}{2} \left(\lambda^0 + \text{NtD}_2 \left(-\text{DtN}_2(\lambda^0, 0), 0 \right), 0 \right) \\ &= \frac{1}{2} \left(\lambda^0 - \text{NtD}_2(\text{DtN}_2(\lambda^0, 0), 0), 0 \right) = 0, \end{aligned} \quad (6)$$

where in the third equality we used linearity, and in the last $\text{NtD}_2(\text{DtN}_2(\lambda, \psi), \psi) = \lambda$. Can the nonlinear DN method also converge in one iteration?

On the one hand, the relation $\text{NtD}_j(\text{DtN}_j(\lambda, \psi), \psi) = \lambda$ holds even in the nonlinear case, simply because the nonlinear DtN_j operator is the inverse of the nonlinear NtD_j operator. On the other hand, due to the nonlinearity of \mathcal{L} , one cannot rely on the error equation, cannot state that $\text{NtD}_2(-\phi) = -\text{NtD}_2(\phi)$, and the symmetry of the decomposition is not sufficient to guarantee $\text{DtN}_1(\lambda, \psi_1) \equiv \text{DtN}_2(\lambda, \psi_2)$, because of the boundary conditions and the force term.

A straight forward observation is that if the nonlinear DN method converges in one iteration, then $G(\lambda) = \lambda_{\text{ex}}, \forall \lambda$, that is $G(\cdot)$ is a constant. A necessary and sufficient condition for the nonlinear DN method to converge in one iteration is then

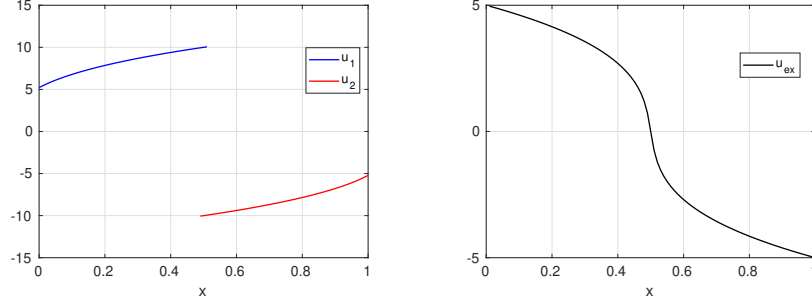


Fig. 1 Subdomain solutions of the nonlinear DN method after one iteration (left), and exact solution (right). The parameters are $g = 5$ and $k = 2$.

$$0 = G'(\lambda) = \frac{1}{2} + \frac{1}{2} (\text{NtD}_2(-\text{DtN}_1(\lambda, \psi_1), \psi_2))' \implies (\text{NtD}_2(-\text{DtN}_1(\lambda, \psi_1), \psi_2))' = -1. \quad (7)$$

Clearly, (7) is satisfied if $\text{NtD}_2(-\text{DtN}_1(\lambda, \psi_1), \psi_2) = -\lambda$. We consider a toy example in which this condition is satisfied. Let $\mathcal{L} = -\partial_x \left((1 + u^2) \partial_x u \right)$, $u(0) = g \in \mathbb{R}^+$, $u(1) = -g$ and $f(x) = \sin((2k)\pi x)$. On the left plot of Fig. 1, we show the subdomain solutions u_1 and u_2 obtained from (2) after the first iteration. The two contributions sum to zero, which is the value of λ_{ex} . Thus, after one iteration we obtain the exact solution shown in the right panel.

Even though the nilpotent property does not hold in general, we show in the following Theorem that the nonlinear DN method can exhibit quadratic convergence.

Theorem 1 (Quadratic Convergence of nonlinear DN) *For any one-dimensional nonlinear problem $\mathcal{L}(u) = f$ such that $\text{DtN}'_1(\lambda_{\text{ex}}, \psi_1) \cdot \text{DtN}'_2(\lambda_{\text{ex}}, \psi_2) > 0$ with $\lambda_{\text{ex}} := u_{\text{ex}}|_{\Gamma}$, there exists a $\theta \in (0, 1)$ such that the nonlinear Dirichlet-Neumann method converges quadratically.*

Proof A sufficient condition for quadratic convergence is that the Jacobian of $G(\cdot)$, defined in (3), is zero at $\lambda_{\text{ex}} := u_{\text{ex}}|_{\Gamma}$, that is $G'(\lambda_{\text{ex}}) = 0$. A direct calculation shows

$$G'(\lambda) = (1 - \theta) + \theta \text{NtD}'_2(-\text{DtN}_1(\lambda, \psi_1), \psi_2) \cdot (-\text{DtN}'_1(\lambda, \psi_1)). \quad (8)$$

Setting $\lambda = \lambda_{\text{ex}}$ and using the optimality condition $\text{DtN}_1(\lambda_{\text{ex}}, \psi_1) = -\text{DtN}_2(\lambda_{\text{ex}}, \psi_2)$ of (5), the above equation changes to

$$G'(\lambda_{\text{ex}}) = (1 - \theta) + \theta \text{NtD}'_2(\text{DtN}_2(\lambda_{\text{ex}}, \psi_2), \psi_2) \cdot (-\text{DtN}'_1(\lambda_{\text{ex}}, \psi_1)). \quad (9)$$

If $\text{DtN}'_1(\lambda_{\text{ex}}, \psi_1) = \text{DtN}'_2(\lambda_{\text{ex}}, \psi_2)$ held true, then using the identity $\text{NtD}'_2(\text{DtN}_2(\lambda, \psi_2), \psi_2) \cdot (\text{DtN}'_2(\lambda, \psi_2)) = 1$, obtained by differentiating $\text{NtD}_j(\text{DtN}_j(\lambda, \psi_j), \psi_j) = \lambda$, we would easily get that $\theta = 1/2$ leads to $G'(\lambda_{\text{ex}}) = 0$. Nevertheless, variational calculus shows that to calculate $\text{DtN}'_j(\lambda, \psi_j)$, one has to solve a linear PDE which does not depend on ψ_j anymore, but whose coefficients still depend on the subdomain so-

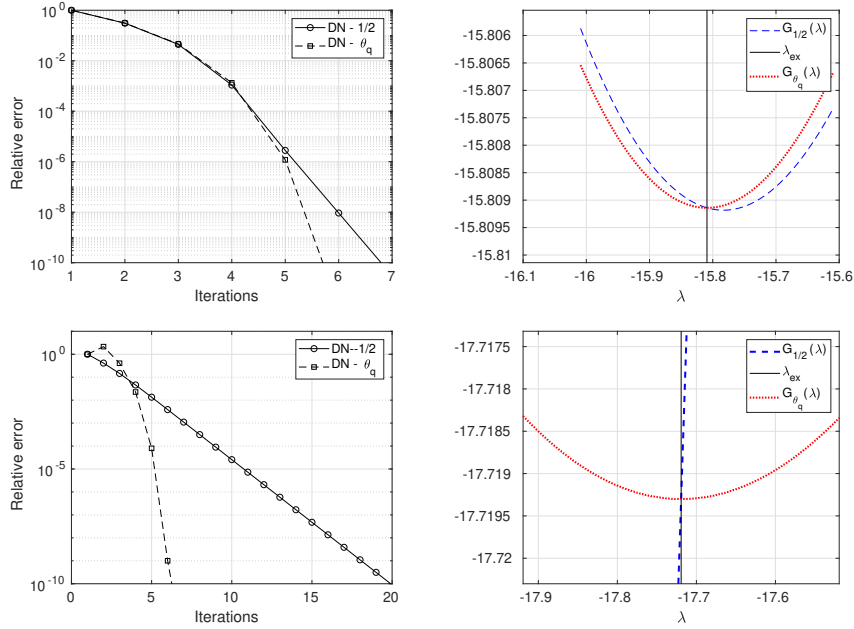


Fig. 2 In the left panels, we show the convergence curves, and in the right panels we plot $G_{\theta}(\lambda)$. The top-row refers to a symmetric decomposition, and the bottom-row to an asymmetric one.

lutions $u_{\text{ex}}|_{\Omega_1}$ and $u_{\text{ex}}|_{\Omega_2}$. In general then, $\text{DtN}'_1(\lambda_{\text{ex}}, \psi_1) \neq \text{DtN}'_2(\lambda_{\text{ex}}, \psi_2)$. However, DtN_j being one dimensional functions, we have $\text{DtN}'_1(\lambda_{\text{ex}}, \psi_1) = \delta \text{DtN}'_2(\lambda_{\text{ex}}, \psi_2)$, for some $\delta \in \mathbb{R}^+$ if $\text{DtN}'_1(\lambda_{\text{ex}}, \psi_1) \cdot \text{DtN}'_2(\lambda_{\text{ex}}, \psi_2) > 0$. Inserting this into (9), we obtain $G'(\lambda_{\text{ex}}) = 0$ if $\theta = \frac{1}{1+\delta} \in (0, 1)$. \square

To illustrate Theorem 1 numerically, we consider $\mathcal{L}(u) = -\partial_x((1 + \alpha u^2)\partial_x u)$, $\Omega = (0, 1)$, $f(x) = 100x$, $u(0) = 0$ and $u(1) = -20$. In the top-row of Fig. 2, we set the interface Γ to $x = 1/2$. In the left panel, we plot the convergence curves for $\theta = 1/2$ and for $\theta_q := \frac{1}{1+\delta}$. In this setting, $\delta = 1.006$ and $\theta_q = 0.498$, so due to the symmetry of the decomposition, θ_q is still very close to $1/2$. In the right panel, we plot $G_{\theta}(\lambda)$ and see that as θ changes, the minimum of $G_{\theta}(\lambda)$ moves, such that it is attained at $\lambda = \lambda_{\text{ex}}$ for $\theta = \theta_q$.

Next, in the bottom row of Fig 2, we consider the same equation and boundary conditions, but Γ is now at $x = 0.3$. The decomposition is asymmetric, with $\delta = 0.43$ and $\theta_q = 0.699$. The left panel shows clearly that for $\theta = 1/2$ the convergence is linear, while for $\theta = \theta_q$, the DN method converges quadratically. In the right panel, we observe that $G_{1/2}(\lambda)$ does not have a local extremum at $\lambda = \lambda_{\text{ex}}$, while $G_{\theta_q}(\lambda)$ does. Theorem 1 does not easily generalize to higher dimensions, since DtN'_j are then matrices, and the relaxation parameter would have to be an operator. Numerically we observed for symmetric decompositions fast convergence for $\theta = 0.5$, while for asymmetric decompositions, θ needs to be tuned for good performance.

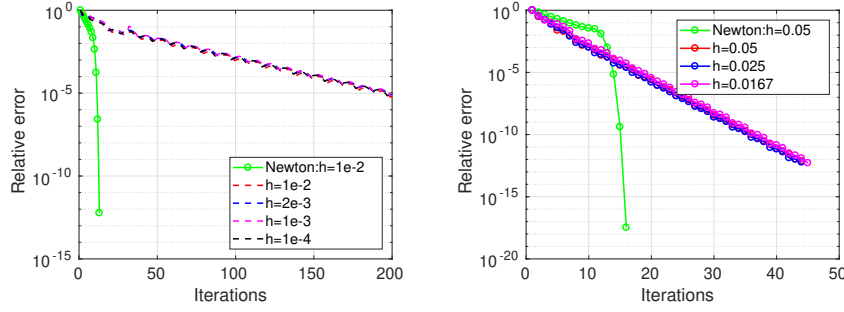


Fig. 3 Convergence behavior of nonlinear DN for different mesh sizes in 1D (left) and 2D (right).

3 Mesh independent convergence

One of the attractive features of the DN method for linear problems is that it achieves mesh independent convergence. Does this also hold for the nonlinear DN method (2)? We first define the nonlinear DN method for multiple subdomains. Motivated by the definition of the DN method for the linear case in [2], we divide the domain $\Omega := (0, L) \times (0, L)$ into N nonoverlapping subdomains $\Omega_j = (\Gamma_{j-1}, \Gamma_j) \times (0, L)$, with $\Gamma_0 = 0$ and $\Gamma_N = L$. The nonlinear DN method for multiple subdomains is then defined for the interior subdomains by

$$\begin{aligned} \mathcal{L}(u_j) &= f_j && \text{in } \Omega_j, \\ \mathcal{N}_j u_j^n(\Gamma_{j-1}, \cdot) &= -\mathcal{N}_{j-1} u_{j-1}^n(\Gamma_{j-1}, \cdot) && \text{on } \Gamma_{j-1}, \\ u_j^n(\Gamma_j) &= (1 - \theta) u_j^{n-1}(\Gamma_j, \cdot) + \theta u_{j+1}^{n-1}(\Gamma_j, \cdot) && \text{on } \Gamma_j, \end{aligned}$$

where $\theta \in (0, 1)$, and for the left and right most subdomains by

$$\begin{aligned} \mathcal{L}(u_1^n) &= f_1, && \text{in } \Omega_1, && \mathcal{L}(u_N^n) &= f_N, && \text{in } \Omega_N, \\ u_1^n(\Gamma, \cdot) &= g(0), && && \mathcal{N}_N u_N^n(\Gamma_{N-1}, \cdot) &= -\mathcal{N}_{N-1} u_{N-1}^n(\Gamma_{N-1}, \cdot), \\ u_1^n(\Gamma_1, \cdot) &= (1 - \theta) u_1^{n-1}(\Gamma_1, \cdot) + \theta u_2^{n-1}(\Gamma_1, \cdot), && && u_N^n(L, \cdot) &= g(L). \end{aligned}$$

We perform two experiments, one in 1D and one in 2D. For the 1D case, we consider the nonlinear diffusion equation $-\partial_x \left((1 + u^2) \partial_x u \right) = 0$, with $u(0) = 0$ and $u(1) = 20$. We divide the domain $\Omega = (0, 1)$ into ten equal subdomains. We then plot the relative error of the nonlinear DN for four different mesh sizes $h = 1e - 2$, $h = 2e - 3$, $h = 1e - 3$, and $h = 1e - 4$. The left plot in Fig. 3 shows that the convergence rate of the nonlinear DN is independent of mesh size, while it is quadratic for Newton's Method. We repeat a similar experiment in 2D, but now the domain $\Omega = (0, 1) \times (0, 1)$ is divided into four equal subdomains. Even in 2D, we observe the mesh independent convergence of the nonlinear DN method, see the right plot of Fig. 3.

4 Dirichlet-Neumann Preconditioned Exact Newton (DNPEN)

In Section 2, we observed that under some special conditions on the exact solution of the nonlinear problem and θ , the nonlinear DN method (2) can be nilpotent. Moreover, the nonlinear DN method can also converge quadratically. But to achieve this, we need to tune the parameter θ according to some a-priori knowledge of the exact solution of the nonlinear problem. Thus in general, the nonlinear DN method converges linearly (as shown in Fig 3).

Iterative methods can be used as preconditioners to achieve faster convergence, see [4, Section 4.1] for the linear case, and [8] for a historical introduction including also the nonlinear case. It was proposed in [6, 3] to use the nonlinear Restricted Additive Schwarz (RAS) and nonlinear Substructured RAS (SRAS) methods as preconditioner for Newton's method. We use the same idea here and apply Newton's method to the fixed point equation of the nonlinear DN method (3), which represents a systematic way of constructing non-linear preconditioners [8]. The fixed point version of (3) can be written as

$$\mathcal{F}(\lambda) := \lambda - G(\lambda) = \theta\lambda - \theta\text{NtD}_2(-\text{DtN}_1(\lambda, \psi_1), \psi_2). \quad (10)$$

Applying Newton to (10) we obtain a new method called Dirichlet Neumann Preconditioned Exact Newton (DNPEN) method.

We saw in Section 2 that the DN method can be nilpotent in certain cases. Can DNPEN still be nilpotent? Let λ_{ex} denote the fixed point of the iteration (3). Let us assume that the Dirichlet Neumann method converges in one iteration. This means that G defined in (3) satisfies $\lambda_{\text{ex}} = G(\lambda^0)$ for any initial guess λ^0 . This shows that the map G is constant, and hence $\mathcal{F}'(\lambda)$ reduces to the identity matrix. Moreover, one step of Newton's method applied to (3) can then be written as

$$\lambda^1 = \lambda^0 - (\mathcal{F}'(\lambda^0))^{-1}\mathcal{F}(\lambda^0) = \lambda^0 - \mathcal{F}(\lambda^0) = G(\lambda^0) = \lambda_{\text{ex}},$$

and hence DNPEN will also be nilpotent in that case. We further have also the following result.

Theorem 2 *The convergence of DNPEN does not depend on the relaxation parameter θ in the DN preconditioner.*

Proof The function \mathcal{F} from (10) corresponding to DNPEN can be rewritten as $\mathcal{F}(\lambda) = \theta\mathcal{K}(\lambda, \psi_1, \psi_2)$, where $\mathcal{K}(\lambda, \psi_1, \psi_2) := \lambda - \text{NtD}_2(-\text{DtN}_1(\lambda, \psi_1), \psi_2)$. Thus, Newton's iteration reads

$$\lambda^{k+1} = \lambda^k - (J\mathcal{F}(\lambda^k))^{-1}\mathcal{F}(\lambda^k) = \lambda^k - (\theta J\mathcal{K}(\lambda^k))^{-1}\theta\mathcal{K}(\lambda^k) = \lambda^k - (J\mathcal{K}(\lambda^k))^{-1}\mathcal{K}(\lambda^k),$$

which shows that the Newton correction does not depend on the relaxation parameter θ . The iterates of Newton's method will thus only depend on \mathcal{K} , and DNPEN has θ independent convergence. \square

The above theorem shows that when using DNPEN, one does not need to search for an optimal choice of θ , in contrast to the nonlinear DN method (2).

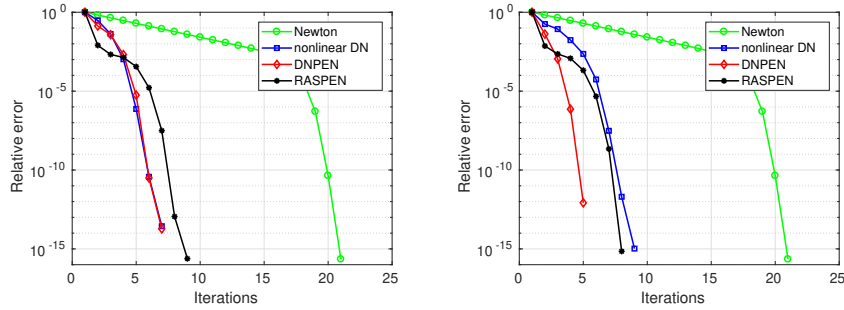


Fig. 4 Comparison of DNPEN (with optimal θ) with unpreconditioned Newton, nonlinear DN (with optimal θ) and RASPEN for a symmetric partition (left) and an asymmetric partition (right).

We now compare the convergence of DNPEN, the unpreconditioned Newton method, the nonlinear DN method (2) and RASPEN [6]. We consider the nonlinear diffusion problem $-\partial_x((1+u^2)\partial_x u) = f$ on $\Omega = (0, 1)$ decomposed into two equally sized subdomains, with $u(0) = 0, u(1) = 10$ and $f(x) = \sin(10\pi x)$. For both DN and DNPEN, we choose the optimal relaxation parameter provided in Theorem 1. The left plot in Fig. 4 shows that the iterative DN converges quadratically using the optimal parameter and is very similar to DNPEN with no significant gain in the number of iterations. The convergence curves also show that the unpreconditioned Newton method is slower than all preconditioned ones, and DNPEN has a slight advantage over RASPEN.

We repeat the same experiment but now using an asymmetric partition of the domain Ω . The right plot in Fig. 4 shows that for this configuration, DNPEN is the fastest while again unpreconditioned Newton is the slowest among the methods considered. Moreover, DNPEN is significantly faster than the nonlinear DN method.

Finally, we illustrate numerically that the convergence of DNPEN does not depend on θ . We know that in general, the nonlinear DN method converges linearly, and it is not always possible to find an optimal θ such that it converge quadratically. We again consider the symmetric partition of the domain and use the same boundary conditions and force term as above. However, instead of the optimal θ , we consider two non-optimal θ 's, namely $\theta = 0.1$ and $\theta = 0.9$. The left plot in Fig. 5 shows the linear convergence of nonlinear DN for both $\theta = 0.1$, and $\theta = 0.9$, and both are slower than the unpreconditioned Newton method. However, DNPEN converges much faster than Newton's method and in the same number of iterations for the two different values $\theta = 0.1$ and $\theta = 0.9$. The right plot in Fig. 5 shows that DNPEN is still faster than RASPEN for both values θ considered.

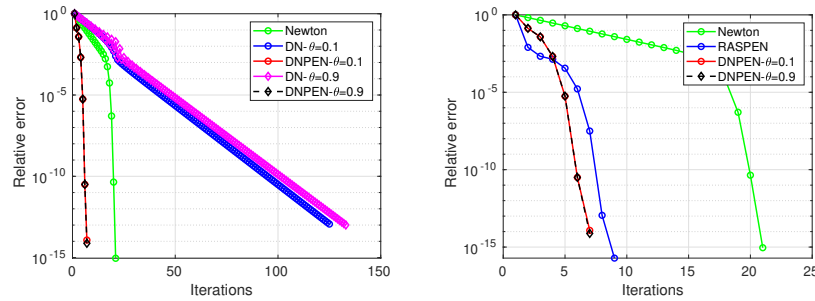


Fig. 5 Comparison of DNPEN with the unpreconditioned Newton method and nonlinear DN (left) and with RASPEN (right) for two different non optimal θ 's.

5 Conclusion

While iterative DN methods are known to converge linearly, we proved that one can obtain quadratic converge for some one-dimensional nonlinear problems and for a well chosen relaxation parameter θ . Under specific conditions, the nonlinear DN method can also become a direct solver, like in the linear case. We then extended DN to multiple subdomains and numerically showed that its convergence is mesh independent. We finally introduced the nonlinear preconditioner DNPEN, proved that the convergence of DNPEN does not depend on the relaxation parameter θ , and observed numerically that DNPEN is faster than unpreconditioned Newton, nonlinear DN and RASPEN in all our examples.

References

1. X-C. Cai and M. Dryja. Domain decomposition methods for monotone nonlinear elliptic problems. *Domain Decomposition Methods in Scientific and Engineering*, 180:21–27, 1994.
2. F. Chaouqui, G. Ciaramella, M. J. Gander, and T. Vanzan. On the scalability of classical one-level domain-decomposition methods. *Vietnam J. Math*, 46(6):1053–1088, 2018.
3. F. Chaouqui, M. J. Gander, P.M. Kumbhar, and T. Vanzan. Linear and Nonlinear Substructured Restricted Additive Schwarz Iterations and Preconditioning. *arXiv:2103.16999*, 2021.
4. G. Ciaramella and M.J Gander. *Iterative Methods and Preconditioners for Systems of Linear Equations*. Accepted for publication in SIAM, 2020.
5. P.G. Ciarlet. *Linear and Nonlinear Functional Analysis with Applications*. Applied mathematics. SIAM, Philadelphia, PA, 2013.
6. 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 Journal on Scientific Computing*, 38(6):A3357–A3380, 2016.
7. L.C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.
8. Martin J. Gander. On the origins of linear and non-linear preconditioning. In *Domain Decomposition Methods in Science and Engineering XXIII*, pages 153–161. Springer, 2017.
9. Alfio Quarteroni and Alberto Valli. *Domain decomposition methods for partial differential equations*. Oxford University Press, 1999.