Newton-Schwarz Optimised Waveform Relaxation Krylov Accelerators for Nonlinear Reactive Transport

Florian Haeberlein[1,2], Laurence Halpern[2], and Anthony Michel[1]

[1] IFP Energies nouvelles, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison, France, firstname.lastname@ifpen.fr
[2] Université Paris 13, CNRS, UMR 7539 LAGA, 99 av. Jean-Baptiste Clément, 93430 Villetaneuse, France, lastname@math.univ-paris13.fr

1 Introduction

Krylov-type methods are widely used in order to accelerate the convergence of Schwarz-type methods in the linear case. Authors in [2] have shown that they accelerate without overhead cost the convergence speed of Schwarz methods for different types of transmission conditions. In the nonlinear context, the well-known class of Newton-Krylov-Schwarz methods (cf. [5]) for steady-state problems or time-dependent problems uses the following strategy: time-dependent problems are discretised uniformly in time first and then one proceeds as for steady-state problems, i.e. the nonlinear problem is solved by a Newton method where the linear system at each iteration is solved by a Krylov-type method preconditioned by an algebraic Schwarz method. The major limitation is that NKS methods do not allow different time discretisations in the subdomains since the problem is discretised in time uniformly up from the beginning.

In this work, we are interested in applying the well-established technique from the linear case in the context of Schwarz Waveform Relaxation methods (SWR, cf. [8]) to nonlinear time-dependent problems in order to benefit from its accelerating properties. We emphasise the use of SWR methods since within this approach, it is possible to use different discretisations in time and space in the subdomains, even the coupling of different models is possible. In many applications, time step restrictions in implicit approaches are highly localised in space due to heterogeneity and SWR methods are perfectly suited to localise and isolate them in subdomains which are treated with different time discretisations.

Our motivation of balancing time step restrictions in the time-dependent nonlinear case on subdomains is close to the approach in [6, 11] where the balancing of nonlinearities on subdomains in the steady-state case is achieved using the permutation of domain decomposition methods and Newton's method in combination with Krylov accelerators.

The paper is organised as follows: In Sect. 2 we set up the problem to solve. In Sect. 3 we describe the Schwarz waveform relaxation (SWR) algorithm and the

R. Bank et al. (eds.), *Domain Decomposition Methods in Science and Engineering XX*, Lecture Notes in Computational Science and Engineering 91,
DOI 10.1007/978-3-642-35275-1_45,

reduction to the interface variables. The new approach is described in Sect. 4. Numerical issues and results are given in Sect. 5.

2 Problem Description

In this paper we consider the following model in $\Omega \times (0,T)$, $\Omega \subset \mathbb{R}^d$:

$$\partial_t(\phi w) + \mathscr{L} w + \mathscr{F}(w) = q \text{ in } \Omega \times (0,T), \tag{1}$$

$$w(\cdot,0) = w_0 \text{ in } \Omega, \quad \mathscr{G} w = g \text{ on } \partial\Omega \times (0,T). \tag{2}$$

where $\phi(x) > 0$ is the porosity, $w \in \mathbb{R}^s$ the vector containing the concentrations of the s chemical species. $\mathscr{L}[\cdot] = \nabla \cdot (-a\nabla + \mathbf{b})$ is a linear operator which models diffusion described by a positive scalar diffusion coefficient $a > 0$ and advection described by a Darcy field $\mathbf{b} \in \mathbb{R}^d$. The transport operator can be zero for non-mobile species. $\mathscr{F}$ is a nonlinear chemical coupling operator. We impose initial conditions on Ω given by w_0 and linear boundary conditions represented by $\mathscr{G}$, for instance Neumann or Dirichlet conditions. The data g and q are source terms depending on space and time.

3 The Schwarz Waveform Relaxation Algorithm and the Classical Approach

We decompose the domain Ω into two non-overlapping domains Ω_1 and Ω_2 and call the common boundary $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$ the interface between the subdomains. We introduce the following SWR algorithm with Robin transmission conditions to approximate the solution of (1): given the iterate w_i^{k-1} which is equal to an initial guess for the first iteration, then one step of the algorithm consists in computing in parallel w_i^k for subdomains $\Omega_i = 1,2$, with data coming from the neighbouring subdomain $\Omega_\times$, with $\tilde{1} = 2$ and $\tilde{2} = 1$.

$$\partial_t(\phi w_i^k) + \mathscr{L} w_i^k + \mathscr{F}(w_i^k) = q \quad \text{in } \Omega_i \times (0,T), \tag{3}$$

$$(\partial_{n_i} + p) w_i^k = (\partial_{n_i} + p) w_\times^{k-1} \quad \text{on } \Gamma \times (0,T), \tag{4}$$

$$w_i^k(\cdot,0) = w_0 \text{ in } \Omega_i, \quad \mathscr{G} w_i^k = g \text{ on } \partial\Omega_i \setminus \Gamma \times (0,T), \tag{5}$$

with n_i the unit outward normal of Ω_i on Γ and $p \in \mathbb{R}$, $p > 0$ a constant.

It is possible to reduce algorithm (3)–(5) to the so-called interface variables. Define the operators $\mathscr{M}_i : (\lambda_i, f) \mapsto w_i$ solution of

$$\partial_t(\phi w_i) + \mathscr{L} w_i + \mathscr{F}(w_i) = q \quad \text{in } \Omega_i \times (0,T), \tag{6}$$

$$(\partial_{n_i} + p) w_i = \lambda_i \quad \text{on } \Gamma \times (0,T), \tag{7}$$

$$w_i^k(\cdot,0) = w_0 \text{ in } \Omega_i, \quad \mathscr{G} w_i^k = g \text{ on } \partial\Omega_i \setminus \Gamma \times (0,T). \tag{8}$$

Here $f = (q, w_0, g)$ represents all source terms except the ones on the interface Γ that are represented separately by λ_i. With these definitions, the transmission conditions (4) can be written as $\lambda_i^{k+1} = -\lambda_\times^k + 2p\mathscr{M}_\times(\lambda_\times^k, f)$, and as a system

$$\begin{pmatrix}\lambda_1^k\\ \lambda_2^k\end{pmatrix} = \begin{pmatrix}-\lambda_2^{k-1}+2p\mathscr{M}_2(\lambda_2^{k-1},f)\\ -\lambda_1^{k-1}+2p\mathscr{M}_1(\lambda_1^{k-1},f)\end{pmatrix}. \tag{9}$$

The SWR algorithm (3) is therefore a fixed point algorithm for the nonlinear interface problem

$$\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix} = \begin{pmatrix}-\lambda_2+2p\mathscr{M}_2(\lambda_2,f)\\ -\lambda_1+2p\mathscr{M}_1(\lambda_1,f)\end{pmatrix}. \tag{10}$$

AQ1 Each iterate requires solving the nonlinear problem (6)–(8). This can be achieved by a Newton method, or a semi-implicit discretisation in time. The latter method has been implemented in [4] for the advection diffusion reaction equation, where the convergence of the fixed point algorithm has been proved. The extension of the proof to the system (1) should be easy.

4 Newton-Schwarz Optimised Waveform Relaxation

The new approach consists first in solving the system (10) by a Newton algorithm. If the interface problem is well-posed, and if the initial data for Newton is sufficiently closed to the solution, the algorithm converges to that solution. According to the interface problem (10), we seek the zeros of the nonlinear function

$$\Theta(\lambda) := -(\lambda_1+\lambda_2)\begin{pmatrix}1\\1\end{pmatrix} + 2p\Upsilon(\lambda), \quad \Upsilon(\lambda) := \begin{pmatrix}\mathscr{M}_2(\lambda_2,f)\\ \mathscr{M}_1(\lambda_1,f)\end{pmatrix}.$$

One step $k-1 \to k$ of Newton's algorithm consists in solving the linear system $\Theta'(\lambda^{k-1})\cdot(\lambda^k-\lambda^{k-1}) = -\Theta(\lambda^{k-1})$. To evaluate the derivative of Θ, we must calculate the derivative of the functions $\lambda_i \mapsto \mathscr{M}_i(\lambda_i,f)$. If $w_i = \mathscr{M}_i(\lambda_i,f)$ and $W_i = \mathscr{M}_i(\lambda_i+\tilde{\lambda}_i,f)$, we see by subtracting equations (6) for w_i and W_i, that $W_i - w_i$ is solution of

$$\partial_t(\phi(W_i-w_i)) + \mathscr{L}(W_i-w_i) + \mathscr{F}(W_i) - \mathscr{F}(w_i) = 0.$$

Introducing the derivative of $\mathscr{F}$, $\mathscr{F}(W_i)-\mathscr{F}(w_i) = \mathscr{F}'(w_i)(W_i-w_i) + \mathscr{O}((W_i-w_i)^2)$, and therefore $W_i - w_i = \tilde{w}_i + o(\tilde{w}_i^2)$, where $\tilde{w}_i$ is solution of the linear equation

$$\partial_t(\phi\tilde{w}_i) + \mathscr{L}\tilde{w}_i + \mathscr{F}'(w_i)\tilde{w}_i = 0. \tag{11}$$

$$(\partial_{n_i}+p)\tilde{w}_i = \tilde{\lambda}_i \tag{12}$$

$$\tilde{w}_i(x,0) = 0 \text{ in } \Omega_i, \quad \mathscr{G}\tilde{w}_i = 0 \text{ on } \partial\Omega_i \setminus \Gamma \times (0,T). \tag{13}$$

Therefore $\partial_{\lambda_i}\mathscr{M}_i(\lambda_i,f)\cdot\tilde{\lambda}_i = \tilde{w}_i := \mathscr{M}^{lin}(\mathscr{F}'(w_i);\tilde{\lambda}_i)$, and

$$\Theta'(\lambda)\cdot\tilde{\lambda} = -(\tilde{\lambda}_1+\tilde{\lambda}_2)\begin{pmatrix}1\\1\end{pmatrix} + 2p\begin{pmatrix}\mathscr{M}^{lin}(\mathscr{F}'(w_2);\tilde{\lambda}_2)\\ \mathscr{M}^{lin}(\mathscr{F}'(w_1);\tilde{\lambda}_1)\end{pmatrix}.$$

After these computations, the algorithm takes the form

$$w_i^{k-1} = \mathscr{M}_i(\lambda_i^{k-1}, f),$$

$$-\sum_{i=1}^{2}(\lambda_i^k - \lambda_i^{k-1})\begin{pmatrix}1\\1\end{pmatrix} + 2p\begin{pmatrix}\mathscr{M}_2^{lin}(\mathscr{F}'(w_2^{k-1});\lambda_2^k - \lambda_2^{k-1})\\ \mathscr{M}_1^{lin}(\mathscr{F}'(w_1^{k-1});\lambda_1^k - \lambda_1^{k-1})\end{pmatrix} = \tag{14}$$
$$-\sum_{i=1}^{2}\lambda_i^k\begin{pmatrix}1\\1\end{pmatrix} + 2p\begin{pmatrix}\mathscr{M}_2(\lambda_2^{k-1}, f)\\ \mathscr{M}_1(\lambda_1^{k-1}, f)\end{pmatrix}$$

The approach requires in every iteration to solve two nonlinear problems in the subdomains. Therefore, a nested iterative procedure is necessary (Newton, or semi-implicit time stepping). Once this is done, $\lambda^{n+1} - \lambda^n$ is a solution of a linear problem solved in parallel in the subdomains.

5 Implementation Using Newton-Krylov Methods and Numerical Results

We have implemented both the classical and the new approach for a special case of problem (1). We assume that s = 2 and $w = (u, v)$ where u denotes a mobile species and v denotes a fixed species. The nonlinear function $\mathscr{F}$ is given by $\mathscr{F}(w) = (R(u,v), -R(u,v))$ where $R(u,v)$ is the overall reaction rate of the reversible reaction $u \leftrightarrows v$.

For the computation of $\mathscr{M}_i(\lambda_i^{k-1}, f)$, we use an implicit Euler scheme in time and a hybrid finite volume scheme (based on [7]) in space. The nonlinear systems are then treated with a global implicit approach by means of Newton's method with exact LU-decomposition. The linear interface problems (14) for λ_i^k are solved using GMRES as Krylov-type method with a precision strategy in the spirit of inexact Newton methods: we adapt the precision of the linear solver with respect to the residuals of the Newton iterates and save therefore costly subdomain evaluations.

Concerning the stopping criterion for the Newton-Schwarz optimised algorithm, it is classically controlled by both the residual and the correction ($\Delta\lambda$) norm. The Schwarz optimised algorithm is only controlled by the correction norm.

For all tests, we set the simulation domain to $\Omega = [0,1] \times [0,1] \subset \mathbb{R}^2$ with the subdomains $\Omega_1 = [0, 0.5] \times [0, 1]$ and $\Omega_2 = [0.5, 1] \times [0, 1]$. The time window considered is $t \in [0, 1]$. Physical parameters are $\phi = 1$, $a = 1.5$, $(b_x, b_y) = (5 \cdot 10^{-2}, 1 \cdot 10^{-3})$. The nonlinear coupling term is defined by $R(u,v) = k(v - \Psi(u))$ where the function Ψ is a BET isotherm law defined by

$$\Psi(u) = \frac{Q_s K_L u}{(1 + K_L u - K_S u)(1 - K_S u)}.$$

BET theory is a rule for the physical adsorption of gas molecules on a solid surface and serves as the basis for an important analysis technique for the measurement of the specific surface area of a material (cf. [3]). This law is insofar mathematically interesting as it is neither convex nor concave (cf. Fig. 1) and is therefore a challenging problem for standard nonlinear solvers like Newton's method. We set $k = 100$,

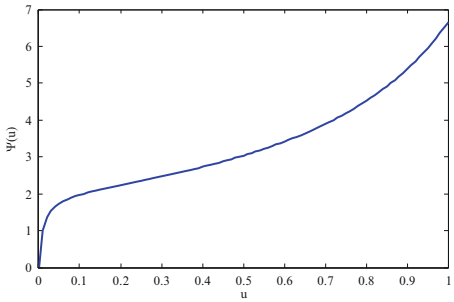

Fig. 1. BET Isotherm law function Ψ with $Q_S = 2$, $K_S = 0.7$, $K_L = 100$

$Q_S = 2$, $K_S = 0.7$ and $K_L = 100$. Initial values are set to $(u_0, v_0) = (\frac{1}{2}, \frac{1}{3})$. By defining the function $g(x,y,t) = (\sin(\pi x)\cos(\pi y)\cos(2\pi t) + \cos(\pi x)\sin(\pi y)\cos(2\pi t) + \cos(\pi x)\cos(\pi y)\sin(2\pi t) + 1)/2$ we impose Dirichlet boundary conditions with values set to $u(x,y,t) = g(x,y,t)$ for $(x,y) \in \partial\Omega$.

As a first experiment, we are interested in the sensitivity of the new approach with respect to the parameter p of the Robin transmission condition. Indeed the theory of optimised Schwarz waveform relaxation for linear problem relies on the fact that the convergence properties of the algorithm heavily depend on this parameter. A best parameter for the advection diffusion reaction equation can be found analytically by solving a best approximation problem, see [1, 8]. No such analysis is available for the nonlinear problem, it is therefore interesting to study the issue numerically.

We discretise the numerical domain with $\Delta x = \Delta y = 1/40$ and $\Delta t = 1/10$ and impose a random initial guess on the interface for the first iteration. As both subdomains are the same size, the number of overall matrix inversions is a meaningful criterion for measuring the numerical performance. We run the two approaches for different parameters p of the Robin transmission condition and plot in Fig. 2 (left) the number of matrix inversions as a function of the parameter p in the Robin transmission condition. One observes first that the performance of the classical approach depends highly on the parameter p of the Robin transmission condition, as in the linear case. The best parameter is $p^* \approx 40$. We observe that the new approach also shows the best performance at p^* but is much less sensitive to the choice of the parameter. The loss of sensitivity with respect to the parameter is still an open question.

It turns out that the new method has a cost overhead, that becomes non negligible if space discretisations are chosen too coarse. For this reason, we study the asymptotic behaviour of the two approaches using always the optimal parameter of the classical approach. We refine the problem in space using always $\Delta x = \Delta y$. Note that we keep the time step constant at $\Delta t = 0.1$. Refining the discretisation also in time would lead to a problem that is quasi stationary at every time step since we use a global implicit approach. We measure again the overall number of matrix inversions in the two approaches and plot them in Fig. 2 (right) versus the discretisation size. One observes that the overhead cost of the new approach compared to the classical approach becomes negligible starting at a discretisation with about 150 grid points

per dimension for the new method. For problems finer than the respective thresholds, the new approach is always faster than the classical approach with the best parameter for the transmission condition. Moreover, the finer the discretisation, the larger the problem, the more important the accelerating property of the new approach. Note that the new approaches has a slope of $O(N^{1/7})$ in the asymptotic behaviour which is considerably less than the slope of the classical approach which behaves like $O(N^{1/2.75})$. The slopes have been determined graphically, no theoretical justification is available. However, this plot shows that the method is much less dependent of the size of the problems than the classical one.

In order to exemplify the accelerating property of the new approach, we perform a simulation with $N_x = N_y = 200$ points in each dimension keeping the number of time steps constant and compare the convergence behaviour of the stopping criteria of the two methods. In Fig. 3 we plot the convergence criterion versus the number of matrix inversions. Note that, for a better comparison, we set the residual norm of the nonlinear interface problem evaluated at the initial guess for both methods at zero matrix inversions. The classical approach exhibit a linear convergence followed by a superlinear convergence, similar to the behaviour of the linear algorithm. We observe the quadratic convergence of the new approach, the characteristic feature of the Newton algorithm.

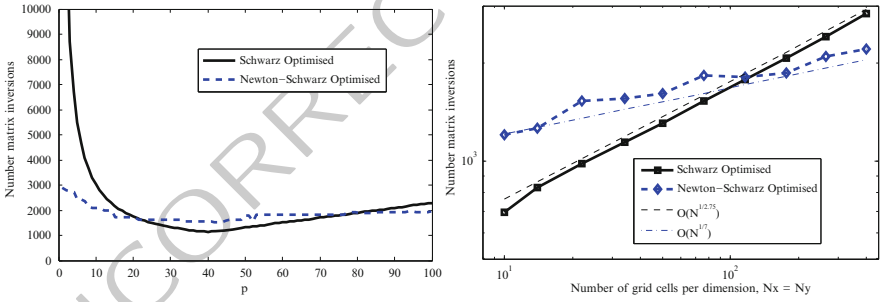

Fig. 2. Number of matrix inversions for the classical approach and new approach, synthetic test case. *Left*: Varying parameter p of the Robin transmission conditions with fixed discretisation in space and time. *Right*: Varying the number of discrete points per dimension ($N_x = N_y$) with fixed discretisation in time and optimal parameter for the Robin transmission condition

Finally, we want to apply the new approach to a benchmark test case in the context of CO_2 geological storage. The 3D test case is based on the benchmark for the SHPCO2 project (Simulation haute performance pour le stockage géologique du CO_2) which is described in [10]. The global domain is set to $\Omega = [0, 4{,}750] \times [0, 3{,}000] \times [-1{,}100, -1{,}000]$ with $(38, 24, 8)$ grid cells in (x, y, z)-direction. The domain is decomposed into the two nonoverlapping subdomains $\Omega_1 = [1{,}000, 2{,}500] \times [0, 3{,}000] \times [-1{,}050, -1{,}000]$ and $\Omega_2 = \Omega \setminus \Omega_1$. We call Ω_1 the reactive subdomain since in this subdomain an injection of the mobile species u is modelled by a source term. The initial state is zero for the mobile and immobile species. We consider

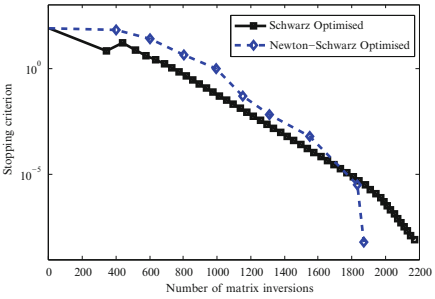

Fig. 3. Convergence history with 200 points per space dimension for the classical approach and new approach, synthetic test case

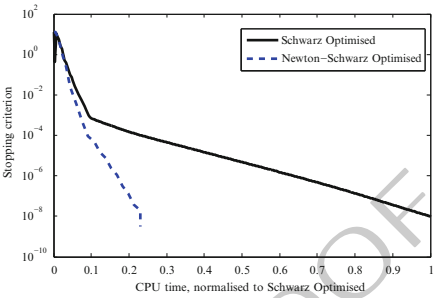

Fig. 4. Convergence history for the classical approach and new approach, SHPCO2 benchmark case

again the BET isotherm law as nonlinear coupling term. The injected mobile species is partially adsorbed by the reaction and partially transported by mainly advection. Simulation time is $[0, 100]$. The SWR approach allows us to use different discretisations in the subdomains. We choose to use ten time steps in the reactive subdomain Ω_1 and only five time steps in the subdomain Ω_2. This choice is insofar justified since the rapid injection in the reactive subdomain restricts the time step size by imposing a maximum number of Newton iterates of ten. As in the subdomain Ω_2, the mobile species appears only by transport processes on a slower time scale than the injection, one can choose a larger time step in order to respect the maximum number of Newton iterations. Concerning the parameter of the Robin transmission condition, we use a low frequency approximation of the optimal parameter. The initial guess on the interface is zero for both subdomain interfaces. In Fig. 4 we plot the convergence histogram, i.e. the stopping criterion in a logarithmic scale versus the CPU time (normalised to the CPU time of the classical approach). Note that both subdomains have a different size of unknowns and therefore the number of matrix inversions, as used in the previous examples, is no longer a valid tool to measure the effort. One observes that the new approach needs only about 20 % of the CPU time of the classical approach.

6 Conclusion

Based on a nonlinear coupled reactive transport system we have developed a new approach for solving the interface problem in the nonlinear case using Krylov-accelerators. In contrast to NKS methods the use of SWR methods allows us to use different time discretisations in the subdomains and so to localise time stepping constraints. We have implemented and tested the method, comparative results with the classical approach have been provided.

The numerical tests showed that, besides an overhead cost for coarse space discretisations, the method has an accelerating property and shows much less sensitivity with respect to the choice of the parameter of the Robin condition. The quadratic

convergence behaviour of the new approach outperforms the superlinear convergence behaviour of the classical approach. Nevertheless, the new approach does have significant overhead costs that are not negligible in the case of coarse problems. Note that a third approach is possible, namely to start with a Newton algorithm for the nonlinear problem, and to solve the so obtained linear problem by a Schwarz-Krylov algorithm (cf. [9]).

Bibliography

[1] D. Bennequin, M. Gander, and L. Halpern. A Homographic Best Approximation Problem with Application to Optimized Schwarz Waveform Relaxation. *Math. Comp.*, 78(265):185–223, 2009.

[2] E. Brakkee and P. Wilders. The Influence of Interface Conditions on Convergence of Krylov-Schwarz Domain Decomposition for the Advection-Diffusion Equation. *Journal of Scientific Computing*, 12:11–30, 1997.

[3] S. Brunauer, P. H. Emett, and E. Teller. Adsorption of Gases in Multimolecular Layers. *Journal American Chemical Society*, 60(2):309–319, 1938.

[4] F. Caetano, L. Halpern, M. Gander, and J. Szeftel. Schwarz waveform relaxation algorithms for semilinear reaction-diffusion. *Networks and heterogeneous media*, 5(3):487–505, 2010.

[5] X. C. Cai, W. D. Gropp, D. E. Keyes, and M. D. Tidriri. Parallel implicit methods for aerodynamics. In *In Keyes*, pages 465–470. American Mathematical Society, 1994.

[6] P. Cresta, O. Allix, C. Rey, and S. Guinard. Nonlinear localization strategies for domain decomposition methods: application to post-buckling analyses. *CMAME*, 196(8):1436–1446, 2007.

[7] R. Eymard, T. Gallouët, and R. Herbin. Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes sushi: a scheme using stabilization and hybrid interfaces. *IMA Journal of Numerical Analysis*, 30(4):1009–1043, 2010.

[8] 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.

[9] F. Haeberlein. *Time-Space Domain Decomposition Methods for Reactive Transport Applied to CO_2 Geological Storage*. PhD Thesis, University Paris 13, 2011.

[10] F. Haeberlein, A. Michel, and L. Halpern. A test case for multi-species reactive-transport in heterogeneous porous media applied to CO_2 geological storate. *http://www.ljll.math.upmc.fr/mcparis09/Files/haeberlein_poster.pdf*, 2009.

[11] J. Pebrel, C. Rey, and P. Gosselet. A nonlinear dual domain decomposition method: application to structural problems with damage. *international journal of multiscale computational engineering*, 6(3):251–262, 2008.