Multidisciplinary optimization formulation for the optimization of multirate systems

Antoine Pierquin, Stephane Brisset, Thomas Henneron

Multidisciplinary optimization formulation for the optimization of multirate systems

A. Pierquin[1], S. Brisset[2] and T. Henneron[1]

[1] L2EP, Université de Lille 1, Cité Scientifique, 59655 Villeneuve d'Ascq, France
[2] L2EP, EC Lille, Cité Scientifique, CS 20048, 59651 Villeneuve d'Ascq, France

Multidisciplinary optimization strategies are widely used in static case and can be extended to a problem with a time-domain model in order to reduce optimization time. The waveform relaxation method is a fixed-point approach applied to waveforms which allows the coupling of dynamic models. By using the individual discipline feasibility strategy, the coupling is transferred to the optimization problem and leads to a high decrease of the number of model evaluations compared to the multidisciplinary feasibility strategy. The drawback of this approach might be the increased number of optimization variables but it is coped through an efficient way to compute the derivatives of time-dependent variables.

***Index Terms*—Electromagnetic coupling, multi-physics and coupled problems, waveform relaxation method, multidisciplinary optimization.**

I. Introduction

MULTIDISCIPLINARY optimization (MDO) considers the different ways to insert and use a multiphysic model into an optimization problem. Indeed, the model can be more or less integrated to the optimization process by changing the formulation of the optimization problem. Different levels of integration are possible when the multiphysic modeling is based on a fixed-point approach. In this case, each subsystem (physical phenomenon, device or discipline) of the multiphysic model is solved with an adapted numerical model. Then, the coupling is done by using a solution from a subsystem as a source for another subsystem. Lastly, the consistency of the coupling between the subsystems is guaranteed in an iterative way. In optimization, the fixed-point approach can be exploited by adding the fixed-point as an integral part of the optimization problem.

These considerations are known and used since a long time in optimization in the static case [1]–[3]. Two main approaches are available: the multidisciplinary feasibility (MDF) and the individual disciplinary feasibility (IDF) [4]. In the first one, the fixed-point loop is performed systematically when the multiphysic model is evaluated. In the second one, the fixed-point criterion is added as a constraint of the problem. Consequently, with the MDF approach, the consistency of the coupling is guaranteed at each evaluation of the model whereas with the IDF approach, the consistency of the model is only guaranteed for the optimal solution while the number of model evaluations is reduced.

In the dynamic case, fixed-point approach also exists, known as waveform relaxation method (WRM) [5], [6] that is well-adapted for multirate system, a system whose subsystems have different time-constants. However, to the best of our knowledge, MDO strategies have not been studied in dynamic case. Authors already studied the MDF [7] and IDF approaches but the use of the last one was problematic for gradient-based optimization methods. Indeed, the computation of the gradient was too time-consuming because it was computed by finite differences and the number of optimisation variables is high with IDF. This paper demonstrates that the partial derivatives of the objective and constraints versus all time-dependant variables can be obtained by using only one additional evaluation of the model. Thus, the computation of the gradient requires a small number of model evaluations, allowing a reduction of the optimization time compared to the MDF approach.

In a first time, the WRM will be presented, then the MDF and IDF strategies in dynamic case. Lastly, they will be applied on an academic example and analysed in terms of complexity and computation time.

II. Waveform relaxation method

The global system $E\left(\dot{w}(t), w(t)\right) = 0$ is considered to be decomposed in r subsystems such that each subsytem verifies

$$E_j\left(\dot{w}_j(t), w_j(t), d_j(t)\right) = 0, \tag{1}$$

with $d_j(t) = [w_1(t), \ldots, w_{j-1}(t), w_{j+1}(t), \ldots, w_r(t)]$.

The WRM, also called *dynamic iteration*, solves iteratively the r subsystems of differential algebraic equations to produce an approximation of the exact solution. At the k^{th} iteration and for the subsystem j, the following problem is solved:

$$E_j^k\left(\dot{w}_j^k(t), w_j^k(t), d_j^k(t)\right) = 0, \tag{2}$$

with $d_j^k(t) = \left[w_1^k(t), \ldots, w_{j-1}^k(t), w_{j+1}^{k-1}(t), \ldots, w_r^{k-1}(t)\right]$ with a Gauss-Seidel scheme (see [5] for other schemes). The solution w^k is obtained from the previous solution w^{k-1}, so the solution at the k^{th} iteration can be written by introducing an operator $\mathbf{\Psi}$ as

$$w^k = \mathbf{\Psi}(w^{k-1}), \tag{3}$$

Manuscript received July 5, 2015; revised May 15, 2015 and June 1, 2015; accepted July 1, 2015. Date of publication July 10, 2015; date of current version July 31, 2015.

Corresponding author: A. Pierquin (e-mail: ant.pierquin@gmail.com)

Digital Object Identifier (inserted by IEEE).

and for each subsystem $w_j^k = \mathbf{\Psi}_j(w^{k-1})$, $j = 1$ to r. Note that if $w^{k-1} = w^k$, then w^k verifies (1) and is solution of the global system. So the opterator $\mathbf{\Psi}$ is a fixed-point operator [6]. Finally, the algorithm stops when $w^k(t)$ is close enough to $w^{k-1}(t)$. At the end of the iterative process, the solution is $w^K(t) = \mathbf{\Psi}^K\left(w^0(t)\right)$.

To apply a gradient-based algorithm to solve the fixed-point problem, or later in optimization in the IDF formulation, the $\mathbf{\Psi}$-operator will be approximated in the neighbourhood of $\bar{w}(t)$ by the first order Taylor expansion

$$\mathbf{\Psi}(w(t)) \simeq \mathbf{\Psi}(\bar{w}(t)) + \nabla\mathbf{\Psi}.(w(t) - \bar{w}(t)). \tag{4}$$

III. Multidisciplinary optimization

A. Multidisciplinary feasibility

Let us consider the optimization problem

$$x^\star = \arg\min_x f(x) \text{ such that } g(x) \leq 0, \tag{5}$$

$$\text{and } h(x) = 0, \tag{6}$$

with x the degrees of freedom, f the objective function and g and h the constraints. We suppose that to obtain the output $g(x)$ and $h(x)$, a system modeled by WRM is solved. With an MDF approach, at each evaluation of g, the iterative procedure of WRM is performed until the fixed-point criterion is verified. Thus, the consistency of the coupling is insured for all evaluations of the model. The Fig. 1 shows the principle with two subsystems Σ_1 and Σ_2: the waveform solution w_1 of Σ_1 is the source of Σ_2, and reciprocally the waveform solution w_2 of Σ_2 is the source of Σ_1.

If the operator $\mathbf{\Psi}$ defined in (3) is applied K times on average, the number of evaluations of $\mathbf{\Psi}$ is approximately $n_{\text{eval}} \times K$ at the end of the optimization process, with n_{eval} the number of model evaluations during the optimization process.

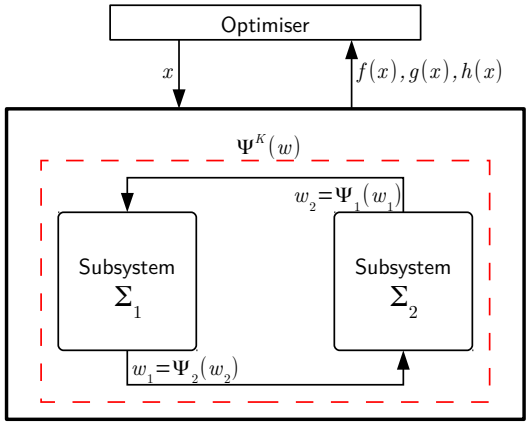


Fig. 1. Scheme of the MDF principle.

B. Individual discipline feasibility

With an IDF approach, the initial problem (6) is modified. The discretized waveform $w(t)$ is added to the optimization variables and the fixed-point condition is also added as a constraint of the problem. Only one evaluation of $\mathbf{\Psi}$ is done per model resolution, the consistency of the coupling being guaranteed at the end of the optimization process. The optimization problem to solve is

$$[\mathbf{x}^\star, w^\star] = \arg\min_{x,w} f(x, w) \text{ such that} \tag{7}$$

$$g(x, w) \leq 0, \tag{8}$$

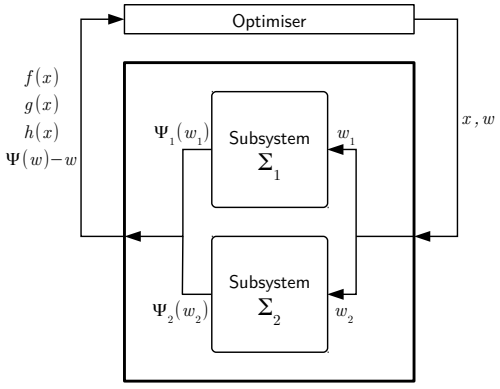
$$h(x, w) = 0, \tag{9}$$

$$w^{\mathrm{T}} - \mathbf{\Psi}(w) = 0. \tag{10}$$

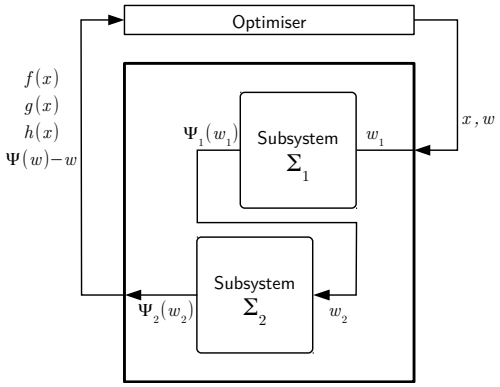
One negative point of the IDF approach is to considerably increase the number of optimization variables. The size of the problem can become too high to be solved by non-linear optimization methods.

The Fig. 2 shows the principle with the same two subsystems Σ_1 and Σ_2 of Fig. 1, in two possible configurations. The subsystems are solved in parallel from the input given by the optimiser in Fig. 2(a); the subsystems are solved sequentially in Fig. 2(b). With this second configuration, the waveform w added as optimization variables can be reduced to w_1.

At the end of the optimization process, the number of evaluations of $\mathbf{\Psi}$ is $\tilde{n}_{\text{eval}}$, with $\tilde{n}_{\text{eval}}$ the number of model evaluations. The optimization duration will be reduced if $\tilde{n}_{\text{eval}} < n_{\text{eval}} \times K$.



(a) IDF principle with parallel solving of the subsystems.



(b) IDF principle with sequential solving of the subsystems.

Fig. 2. Scheme of the IDF principle.

IV. Application example

A. Problem formulations

A LC filter supplying a transformer is considered. The voltage source v_s is a 50 Hz fundamental signal and a 20 kHz

sine perturbation. Different time-steps can be used to model this device by WRM (Fig. 3), with a 2D finite element model (FEM) for the simulation of the transformer. The time-step into the FEM is fitted to the 50 Hz dynamic, whereas in the circuit part, it is fitted to the 20 kHz dynamic. With a Gauss-Seidel scheme, the circuit model and the FEM are solved sequentially. At each WRM iteration, a current waveform $i^{k-1}(t)$ is imposed as a source into the circuit model. Its resolution gives the voltage waveform $v^k(t)$ imposed to the transformer, then the resolution of the FEM gives the current $i^k(t)$ that will be the current source of the circuit model at the next iteration. The time-discretisation being different, linear interpolation is used to obtain the waveform for all time t.

At each iteration $\mathbf{i}^k = \mathbf{\Psi}(\mathbf{i}^{k-1})$, with $\mathbf{i}^k = [i^k(t_1), i^k(t_2), \ldots, i^k(t_N)]^T$ and $t_n,\ n = 1$ to N the time discretisation, and so the algorithm produces $\mathbf{i}^K = \mathbf{\Psi}^K(\mathbf{i}^0)$.

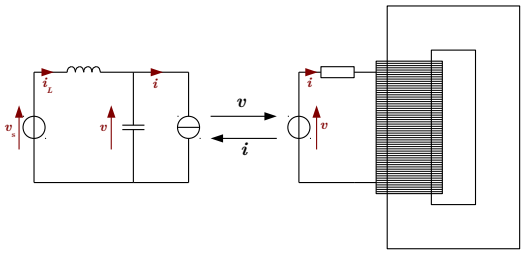


Fig. 3. Device split to apply the WRM.

The optimization aims to minimize the mass of the transformer by acting on the width L and the height H of the transformer (Fig. 4). Moreover, the root mean square current i_{rms} into the transformer has to be equal to 3 A. The initial problem, solved with an MDF approach, is

$$\begin{cases} \min\limits_{H,L} m(H,L), \\ 20\text{cm} \leq H \leq 40\text{cm}, \\ 12\text{cm} \leq L \leq 24\text{cm}, \\ H - \frac{2L}{3} > 0, \\ i_{\text{rms}} = 3\ A. \end{cases} \tag{11}$$

The sequential solving of the models (circuit then FE) allows an IDF formulation with only one waveform added as optimization variables, like in Fig. 2(b). The discretized current waveform $\mathbf{i}$ is added to the optimization variables. The sequential solving of the circuit and of the FEM produces the output current waveform $\mathbf{i}^{\text{out}} = \mathbf{\Psi}(\mathbf{i})$: the fixed-point criterion $\mathbf{i}^{\text{out}} = \mathbf{i}$ is added as a constraint into the optimization problem. The waveform of the voltage v is juste an intermediary variable that does not take part of the optimization problem. By adding only i to the problem, it avoids increasing too much the dimension of the optimization problem. We note $i_n = i(t_n)$ and $i_n^{\text{out}} = i^{\text{out}}(t_n)$. With these notations, the problem to solve with the IDF formulation is

$$\begin{cases} \min\limits_{H,L,i} m(H,L), \\ 20\text{cm} \leq H \leq 40\text{cm}, \\ 12\text{cm} \leq L \leq 24\text{cm}, \\ H - \frac{2L}{3} > 0, \\ i_{\text{rms}} = 3\ A, \\ i_n - i_n^{\text{out}} = 0, \forall n = 1, \ldots, N. \end{cases} \tag{12}$$

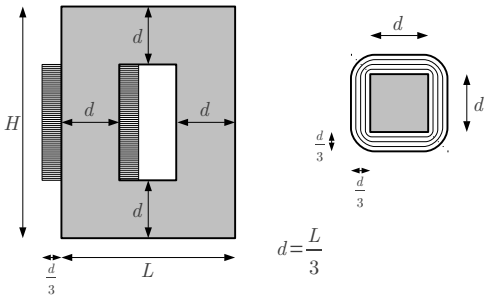


Fig. 4. Geometry of the transformer.

B. Computation of the jacobian

The discretized waveform of the current is added to the optimization variables and the constraints $i_n - i_n^{\text{out}} = 0$, for $n = 1$ to N ensure $\mathbf{i} - \mathbf{\Psi}(\mathbf{i}) = 0$, namely the consistency of the coupling. One difficulty is to obtain the jacobian of the operator $\mathbf{\Psi}$ to use a gradient-based optimization algorithm. Most of the time, the gradient is computed by finite differences for each optimization variable. Nevertheless, this is not required for the gradient with respect to i_n

Due to the implicit Euler scheme used to solve the differential equations, the jacobian $\nabla\mathbf{\Psi}$ is a triangular matrix if the time-discretisation is the same for all the subsystems. Indeed, $\frac{\partial i_q^{\text{out}}}{\partial i_p} = 0$ if $q < p$, because a perturbation at the moment t_p in input does not have any influence on the output for the previous moments.

Moreover, we have $\frac{\partial i_p^{\text{out}}}{\partial i_p} = \frac{\partial i_q^{\text{out}}}{\partial i_q},\ p \neq q$ and consequently $\frac{\partial i_{p+\ell}^{\text{out}}}{\partial i_p} = \frac{\partial i_{q+\ell}^{\text{out}}}{\partial i_q},\ \forall \ell \geq 0$. This means that the matrix $\nabla\Psi$ has identical values onto its diagonals, and only $\frac{\partial \mathbf{i}^{\text{out}}}{\partial i_1}$ is needed to complete the matrix.

$$J = \begin{bmatrix} \frac{\partial i_1^{\text{out}}}{\partial i_1} & 0 & 0 & \ldots & 0 \\ \frac{\partial i_2^{\text{out}}}{\partial i_1} & \frac{\partial i_1^{\text{out}}}{\partial i_1} & 0 & \ldots & 0 \\ \frac{\partial i_3^{\text{out}}}{\partial i_1} & \frac{\partial i_2^{\text{out}}}{\partial i_1} & \frac{\partial i_1^{\text{out}}}{\partial i_1} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \frac{\partial i_N^{\text{out}}}{\partial i_1} & \frac{\partial i_{N-1}^{\text{out}}}{\partial i_1} & \frac{\partial i_{N-2}^{\text{out}}}{\partial i_1} & \ldots & \frac{\partial i_1^{\text{out}}}{\partial i_1} \end{bmatrix} \tag{13}$$

To explain the last point, let us consider more generally the WRM applied with a Gauss-Seidel scheme to the following linear differential systems

$$\frac{\partial}{\partial t} X^1 = A_1 X^1 + B_1 X^{\text{in}} \tag{14}$$

$$\frac{\partial}{\partial t} X^{\text{out}} = A_2 X^1 + B_2 X^{\text{out}}. \tag{15}$$

where $A_1,\ A_2,\ B_1,\ B_2$ are constant matrix. With an Euler scheme in time, the systems become

$$X_n^1 = \left(\frac{1}{\delta t} I_1 - A_1\right)^{-1} \left(\frac{1}{\delta t} X_{n-1}^1 + B_1 X_n^{\text{in}}\right) \tag{16}$$

$$X_n^{\text{out}} = \left(\frac{1}{\delta t} I_2 - B_2\right)^{-1} \left(\frac{1}{\delta t} X_{n-1}^{\text{out}} + A_2 X_n^1\right), \tag{17}$$

with $I_1,\ I_2$ the identities matrix with the adadted size.

The derivative of the previous expressions with respect to X_n^{in} gives

$$\frac{\partial X_n^1}{\partial X_n^{\text{in}}} = \left(\frac{1}{\delta t} I_1 - A_1\right)^{-1} B_1 \tag{18}$$

$$\frac{\partial X_n^{\text{out}}}{\partial X_n^{\text{in}}} = \left(\frac{1}{\delta t} I_2 - B_2\right)^{-1} \left(A_2 \frac{\partial X_n^1}{\partial X_n^{\text{in}}}\right), \tag{19}$$

and so

$$\frac{\partial X_n^{\text{out}}}{\partial X_n^{\text{in}}} = \left(\frac{1}{\delta t} I - B_2\right)^{-1} A_2 \left(\frac{1}{\delta t} I_1 - A_1\right)^{-1} B_1, \ \forall n. \tag{20}$$

This proves that $\frac{\partial i_p^{\text{out}}}{\partial i_p} = \frac{\partial i_q^{\text{out}}}{\partial i_q}$, $p \neq q$ and consequently $\frac{\partial i_{p+\ell}^{\text{out}}}{\partial i_p} = \frac{\partial i_{q+\ell}^{\text{out}}}{\partial i_q}$, $\forall \ell \geq 0$ by induction. This can be proved with more than 2 subsystems if the solving of the subsystems is totally sequential, when the solution of a subsystem is used as a source for the next one.

Finally, the computation of $\frac{\partial i_\ell^{\text{out}}}{\partial i_1}$ is obtained for all ℓ by only one calculation by finite differences, and allows to have the complete matrix $\nabla \mathbf{\Psi}$. The application $\mathbf{\Psi}$ is evaluated with the waveform input $\mathbf{i}$:

$$\mathbf{i}^{\text{out}} = \mathbf{\Psi}(\mathbf{i}). \tag{21}$$

Then we subject the input i_1 to a small perturbation δi, before to evaluate Ψ with the perturbed waveform:

$$\mathbf{i}_\delta^{\text{out}} = \mathbf{\Psi}(\mathbf{i} + \delta i \mathbf{e}_1), \tag{22}$$

with $\mathbf{e}_1$ the canonical basis. Lastly, the gradient can be approximated

$$\frac{\partial \mathbf{i}}{\partial i_1} \simeq \frac{\mathbf{i}_\delta^{\text{out}} - \mathbf{i}^{\text{out}}}{\delta i}. \tag{23}$$

C. Results

Optimizations using the sequential quadratic programming (SQP) method of Matlab® software are done for the MDF and IDF formulations. SQP finds local optimum and needs the computation of the gradient of the objective and constraint functions. The gradient of the constraints is computed by finite differences with the technique presented in section IV-B for $\nabla \mathbf{\Psi}$. So at each call of the constraint function, the application Ψ is evaluated four times: one times to have the value of the function at a given point, then three times to obtain the derivative with respect to H, L and $\nabla \mathbf{\Psi}$.

To compare the two approaches, 10 random initial points are generated. For 8 points, the IDF formulation converges to the same solution than the MDF one. For these 8 points, the average relative error is 0.0012%. But with the IDF approach, the number of evaluations of the Ψ operator is reduced (Fig. 5). Each evaluation of Ψ corresponds to the solving of a FE problem in the time-domain, that is time-consuming. So the reduction of the number of FE problem solvings leads most of the time to a decrease of the optimization time. Nevertheless, the optimization process can evaluate a lot of times the objective function without computing the constraints. This explains that the optimization duration stays superior with the IDF formulation for the 7th point. Finally, the optimization duration is globally lower with the IDF approach. The speedup factor is 3 on average, and even greater than 7 for one point (Fig. 6).

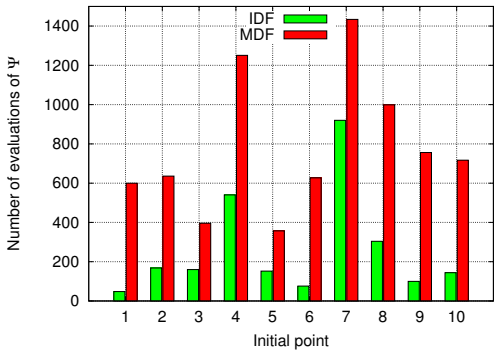


Fig. 5. Number of evaluations of $\mathbf{\Psi}$ operator.

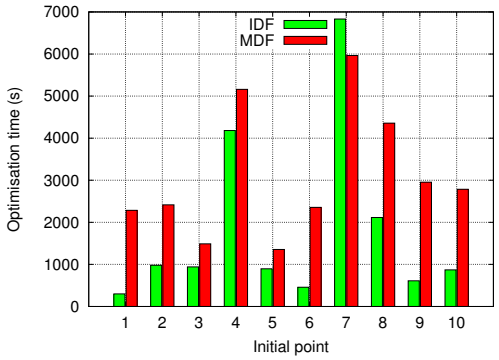


Fig. 6. Optimization time.

V. Conclusion

The WRM used to co-simulate multirate systems has been added into the optimization problem thanks to the IDF formulation. The dimension of the problem is increased but the number of evaluations of the model decreases through an efficient way to compute the derivatives of time-dependent variables. This leads to a significant reduction of the optimization time.

References

[1] R. D. Braun and I. M. Kroo, "Development and application of the collaborative optimization architecture in a multidisciplinary design environment," in *Multidisciplinary Design Optimization: State of the Art*. SIAM, 1997, pp. 98–116.

[2] J. Sobieszczanski-Sobieski and R. Haftka, "Multidisciplinary aerospace design optimization: survey of recent developments," *Structural optimization*, vol. 14, pp. 1–23, 1997.

[3] R. Ayed and S. Brisset, "Multidisciplinary optimization formulations benefits on space mapping techniques," *COMPEL - The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 31, no. 3, pp. 945–957, 2012.

[4] E. Cramer, J. Dennis, Jr., P. Frank, R. Lewis, and G. Shubin, "Problem formulation for multidisciplinary optimization," *SIAM Journal on Optimization*, vol. 4, no. 4, pp. 754–776, 1994.

[5] E. Lelarasmee, A. Ruehli, and A. Sangiovanni-Vincentelli, "The waveform relaxation method for time-domain analysis of large scale integrated circuits," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 1, no. 3, pp. 131–145, july 1982.

[6] A. Bartel, M. Brunk, M. Günther, and S. Schöps, "Dynamic iteration for coupled problems of electric circuits and distributed devices," *SIAM Journal on Scientific Computing*, vol. 35, no. 2, pp. B315–B335, 2013.

[7] A. Pierquin, S. Brisset, T. Henneron, and S. Clénet, "Benefits of waveform relaxation method and output space mapping for the optimization of multirate systems," *Magnetics, IEEE Transactions on*, vol. 50, no. 2, pp. 653–656, Feb 2014.