Subgradient Techniques for Passivity Enforcement of Linear Device and Interconnect Macromodels

Giuseppe C. Calafiore, Alessandro Chinea and Stefano Grivet-Talocia, *Senior Member, IEEE*

***Abstract*—This paper presents a class of nonsmooth convex optimization methods for the passivity enforcement of reduced-order macromodels of electrical interconnects, packages and linear passive devices. Model passivity can be lost during model extraction or identification from numerical field solutions or direct measurements. Non-passive models may cause instabilities in transient system-level simulation, therefore a suitable postprocessing is necessary in order to eliminate any passivity violations. Different from leading numerical schemes on the subject, passivity enforcement is here formulated as a direct, frequency-domain, $\mathcal{H}_\infty$ norm minimization through perturbation of the model state-space parameters. Since the dependence of this norm on the parameters is nonsmooth but continuous and convex, we resort to the use of subdifferentials and subgradients, which are used to devise two different algorithms. We provide a theoretical proof of the global optimality for the solution computed via both schemes. Numerical results confirm that these algorithms achieve the global optimum in a finite number of iterations within a prescribed accuracy level.**

***Index Terms*—Linear macromodeling, Passivity, Nonsmooth optimization, Subgradient techniques, Convex optimization**

I. Introduction

Computer-Aided Design flows heavily rely on models for all those parts of a system that influence its performance. In common situations, such models are available through some identification process from input-output responses, which are available by direct measurements or by numerical simulations. Depending on the structure of the model, different identification strategies can be pursued. For linear structures, usually characterized by a state-space form [1], several well-consolidated time-domain and frequency-domain identification methods exist [2], [3]. Most prominent methods are based on rational approximation via iterative weighted least-squares (Vector Fitting) [4], [5], [6].

The main subject of the present work is passivity enforcement on the identified models [7], [8]. A given physical structure is passive if unable to generate energy. Examples of passive structures are the electrical interconnect networks that provide signal and power distribution in any electrical and electronic system [9], [10]. Preserving passivity also in the extracted models is very important, since numerical (transient) simulation of non-passive models may lead to instability and fail [11], [12], even if the terminations or loads are passive.

Model passivity may be lost due to numerical approximations during the identification stage, unless suitable passivity constraints are explicitly accounted for. This latter approach, however, requires very high computational costs both in terms of memory and CPU time, even for moderately complex models. Therefore, the most common approach in the literature is a two-step flow that first identifies an initial model, and then enforces passivity through a suitable perturbation stage. This work provides a new approach for this second perturbation step.

Significant efforts have been devoted to the development of robust and efficient passivity check and enforcement methods. The most notable techniques can be classified in three groups. Direct methods enforce passivity through Positive Real or Bounded Real Lemma [13] constraints [14], [15], [16]. The main advantage is the formulation of passivity enforcement as a convex optimization problem based on Linear Matrix Inequalities (LMI). This problem admits a unique global solution for which reliable optimization methods exist [17]. The main drawback of these methods, however, is the excessive computational cost, due to the introduction in the optimization problem of a large slack Lyapunov matrix variable, which prevents a good scalability to complex models characterized by a large dynamical order and/or number of inputs and outputs. A second class of methods is based on Hamiltonian eigenvalue extraction and perturbation [12], [18], [19], [20], [21], [22], [23], [24]. It can be shown that a model is not passive if and only if some associated Hamiltonian matrix has purely imaginary eigenvalues. Finding and perturbing such eigenvalues to move off the imaginary axis has been quite successful [12], [19], [24]. The main drawback of this technique is a non-convex formulation, which does not guarantee convergence. The last class of methods is based on iterative perturbation of the frequency-dependent energy gain of the model [20], [21], [25], [26]. The corresponding schemes are based on the solution of suitably constrained linear or quadratic programs at each iteration. Such "local" problems, however, are only approximated and do not guarantee that the global optimum is found. Variants of the above schemes have been presented in [27], [28], [29], [30], [31], [32]. A comprehensive comparison of main techniques is available in [33].

This paper presents a new approach to passivity enforcement. The passivity constraint is formulated as a unit bound on the $\mathcal{H}_\infty$ norm of the model, and this constraint is tackled directly in the frequency domain, thus avoiding to resort

Manuscript received ; revised.

This work was supported in part by the Italian Ministry of University (MIUR) under a Program for the Development of Research of National Interest (PRIN grant #2008W5P2K).

Giuseppe C. Calafiore is with the Department of Control and Computer Engineering, Politecnico di Torino 10129, Italy (e-mail: giuseppe.calafiore@polito.it); Alessandro Chinea and Stefano Grivet-Talocia are with the Department of Electronics, Politecnico di Torino, Torino 10129, Italy (e-mail: alessandro.chinea@polito.it, stefano.grivet@polito.it).

to the Bounded Real Lemma, which is the main source of difficulties in the usual LMI formulation, due to introduction of the large Lyapunov matrix. A closer look at the dependence of the $\mathcal{H}_\infty$ norm on the model parameters in the frequency domain reveals a convex continuous but non-smooth behavior. Therefore, standard descent methods based on gradients and derivatives are ruled out, since such quantities may not be defined everywhere in the parameter space. Therefore, we adopt a generalization of such methods based on subdifferentials and subgradients, which exist also in case of non-differentiabile but convex forms. A complete characterization of the $\mathcal{H}_\infty$ norm subgradients with respect to model parameters is derived and used to construct two schemes for passivity enforcement. These schemes provide a convex optimization framework and are thus guaranteed to attain the global optimum in a finite number of steps and within a prescribed tolerance.

This paper is organized as follows. Section II states the main problem. Section III presents preliminaries and background material. Section IV introduces a characterization for the subdifferential of the $\mathcal{H}_\infty$ norm. Sections V and VI present a projected subgradient algorithm and an alternate subgradient algorithm, respectively, for passivity enforcement. Numerical results are presented and discussed in Section VII.

II. Problem statement

We consider a *nominal* state-space macromodel characterized through its $n_y \times n_u$ transfer matrix

$$\boldsymbol{H}(\boldsymbol{0}, s) = \boldsymbol{C}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B} + \boldsymbol{D}\,, \tag{1}$$

where s is the Laplace variable, with state-space matrices $\boldsymbol{A} \in \mathbb{R}^{n,n}, \boldsymbol{B} \in \mathbb{R}^{n,n_u}, \boldsymbol{C} \in \mathbb{R}^{n_y,n}, \boldsymbol{D} \in \mathbb{R}^{n_y,n_u}$. The first argument of $\boldsymbol{H}$, which is set to $\boldsymbol{0}$ in (1), will be used in the following to parameterize a perturbation of the transfer matrix. We suppose that the macromodel (1) is available through some identification or approximation process. A very common scenario in the microwave area is the availability of frequency samples $\{(\omega_k, \boldsymbol{S}_k),\, k = 1, \ldots, K\}$ of the scattering matrix for some linear device such as a filter or an electrical interconnect, coming from direct measurement or from a full-wave electromagnetic field simulation. Common rational approximation schemes such as *Vector Fitting* [4], [5], [6] can be applied to these samples in order to find the state-space macromodel (1) with minimal deviation from the raw data. In a least-squares formulation, this amounts to solving

$$\min \sum_{k=1}^{K} \|\boldsymbol{H}(\boldsymbol{0}, \mathrm{j}\omega_k) - \boldsymbol{S}_k\|^2$$

for unknown matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}$. This problem is well addressed in the literature, so we consider the nominal macromodel (1) as our starting point.

System (1) is assumed to be asymptotically stable, and the state-space realization is assumed to be minimal. A stable system (1) is passive if and only if

$$\|\boldsymbol{H}(\boldsymbol{0})\|_{\mathcal{H}_\infty} = \sup_{\omega \in \mathbb{R}} \sigma_1(\boldsymbol{H}(\boldsymbol{0}, \mathrm{j}\omega)) \leq 1\,, \tag{2}$$

where σ_1 denotes the maximum singular value, and where the supremum $\sup \mathcal{X}$ of a set $\mathcal{X}$ represents the smallest real number that is greater than or equal to every number in $\mathcal{X}$. In case (2) does not hold, we want to perturb the state-space matrices such that the resulting perturbed system is passive, under some minimal perturbation condition. As typical in the extensive literature on the subject (see [33] for an overview), we choose to perturb only the state-space matrix $\boldsymbol{C}$, which usually stores the residues of a partial fraction expansion of $\boldsymbol{H}(\boldsymbol{0}, s)$. Matrix $\boldsymbol{A}$ is preserved in order to maintain the system poles, and matrix $\boldsymbol{B}$ does not need perturbation since it provides a static input-to-state map. Matrix $\boldsymbol{D}$, which corresponds to the high-frequency ($s \to \infty$) response, is assumed to fulfill the condition $\|\boldsymbol{D}\|_2 = \sigma_1(\boldsymbol{D}) \leq 1$, which is necessary for passivity. This condition is easy to enforce during the model identification stage.

We define the perturbed system as

$$\boldsymbol{H}(\boldsymbol{C}_p, s) = (\boldsymbol{C} + \boldsymbol{C}_p)(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B} + \boldsymbol{D}\,, \tag{3}$$

where the perturbation term $\boldsymbol{C}_p$ is unknown. Supposing that the nominal system $\boldsymbol{H}(\boldsymbol{0}, s)$ is *not* passive, our goal is to find the minimal perturbation that renders passive the perturbed system. Therefore, we need to solve the optimization problem

$$\min \|\boldsymbol{C}_p\|_F, \quad \text{s.t. } \|\boldsymbol{H}(\boldsymbol{C}_p)\|_{\mathcal{H}_\infty} < 1, \tag{4}$$

where the minimal perturbation condition is expressed without loss of generality in terms of the Frobenius norm. Other norms, including frequency-dependent and weighted norms, can be used as well [29], [30].

A more abstract formulation of (4) can be obtained by collecting all decision variables in a vector $\boldsymbol{x} = \mathrm{vec}(\boldsymbol{C}_p)$. Operator $\mathrm{vec}(\cdot)$ stacks in a single column all elements of its matrix argument. The reconstruction of $\boldsymbol{C}_p$, whose size is defined by the context, is obtained using the "inverse" operator $\boldsymbol{C}_p = \mathrm{mat}(\boldsymbol{x})$. The optimization problem (4) can thus be restated in terms of the euclidean norm

$$f(\boldsymbol{x}) = \|\boldsymbol{x}\|_2 = \|\boldsymbol{C}_p\|_F\,. \tag{5}$$

Setting now $\bar{\boldsymbol{H}}(\boldsymbol{x}, \mathrm{j}\omega) = \boldsymbol{H}(\boldsymbol{C}_p, \mathrm{j}\omega)$ and defining

$$h_\omega(\boldsymbol{x}) = \sigma_1(\bar{\boldsymbol{H}}(\boldsymbol{x}, \mathrm{j}\omega)),$$

as the maximum singular value of the perturbed transfer matrix at frequency ω, we have the following characterization of its $\mathcal{H}_\infty$ norm

$$h(\boldsymbol{x}) = \|\boldsymbol{H}(\boldsymbol{C}_p)\|_{\mathcal{H}_\infty} = \sup_{\omega \in \mathbb{R}} h_\omega(\boldsymbol{x}). \tag{6}$$

The above definitions lead the following two complementary formulations of (4).

- The minimal perturbation on $\boldsymbol{x}$ that is able to achieve system passivity according to (4) is obtained by solving the following optimization problem

  $$\min_{\boldsymbol{x}} f(\boldsymbol{x}), \quad \text{s.t. } h(\boldsymbol{x}) \leq 1. \tag{7}$$

  This problem admits a global optimum $\boldsymbol{x}^*$ if both f and h are convex and the problem is feasible.
- A second alternative formulation is based on a predefined bound on the perturbation amount. We define a set

  $$\mathcal{X}_\nu = \{\boldsymbol{x} :\ f(\boldsymbol{x}) \leq \nu\}$$

including all parameter configurations defining perturbed models that differ from the nominal system less than ν. Among all such models, we seek the one with minimal $\mathcal{H}_\infty$ norm by solving problem

$$\min_{\boldsymbol{x}} h(\boldsymbol{x}), \quad \text{s.t. } \boldsymbol{x} \in \mathcal{X}_\nu, \tag{8}$$

which admits a global optimum $\boldsymbol{x}^*$ if h is convex and $\mathcal{X}_\nu$ is a compact nonempty convex set. We shall show in Section V-B how a simple bisection loop on ν achieves macromodel passivity with the least possible perturbation, thus providing a solution to our original problem (4).

Since convexity plays a crucial role in the following derivations and in the global solvability of (7) and (8), we briefly review main definitions and concepts in next section.

III. Preliminaries

We next review some fundamental facts on convexity, subgradients, and standard subgradient-based iterative minimization algorithms for nonsmooth (i.e., non differentiable) convex optimization. This material is quite standard, but needs be synthesized for the purposes of our specific applications; relevant references include [34], [35], [36], [37].

A. Convexity

A set $\mathcal{S}$ is convex if for any $x_1, x_2 \in \mathcal{S}$ and for any ϑ with $0 \le \vartheta \le 1$, we have

$$\vartheta x_1 + (1-\vartheta)x_2 \in \mathcal{S}\,.$$

The convex hull of a set $\mathcal{S}$, denoted as $\operatorname{conv}\mathcal{S}$ is the set

$$\operatorname{conv}\mathcal{S} = \Big\{ \sum_{i=1}^{k} \vartheta_i x_i : \\ x_i \in \mathcal{S}, \vartheta_i \ge 0, i = 1,\dots,k, \sum_{i=1}^{k} \vartheta_i = 1 \Big\},$$

and it represents the smallest convex set containing $\mathcal{S}$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if its domain is a convex set and if for any $\boldsymbol{x}_1, \boldsymbol{x}_2$ in the domain and for any ϑ with $0 \le \vartheta \le 1$, we have

$$f(\vartheta \boldsymbol{x}_1 + (1-\vartheta)\boldsymbol{x}_2) \le \vartheta f(\boldsymbol{x}_1) + (1-\vartheta) f(\boldsymbol{x}_2)\,.$$

In particular, it follows from the triangle inequality that any norm is convex, therefore both $f(\boldsymbol{x})$ in (5) and $h(\boldsymbol{x})$ in (6) are convex functions.

B. Gradients and Subgradients

A vector $\boldsymbol{g} \in \mathbb{R}^n$ is a *subgradient* of a convex function f at $\boldsymbol{x}$, if for all $\boldsymbol{z}$ in the domain of f, it holds that

$$f(\boldsymbol{z}) \ge f(\boldsymbol{x}) + \boldsymbol{g}^\top (\boldsymbol{z} - \boldsymbol{x}). \tag{9}$$

If f is convex and differentiable at $\boldsymbol{x}$, then $\boldsymbol{g} = \nabla f(\boldsymbol{x})$ is the unique subgradient. However, subgradients exist also at points where f is non differentiable. The set

$$\partial f(\boldsymbol{x}) \doteq \{\boldsymbol{g} : \ \boldsymbol{g} \text{ is a subgradient of } f \text{ at } \boldsymbol{x}\}$$

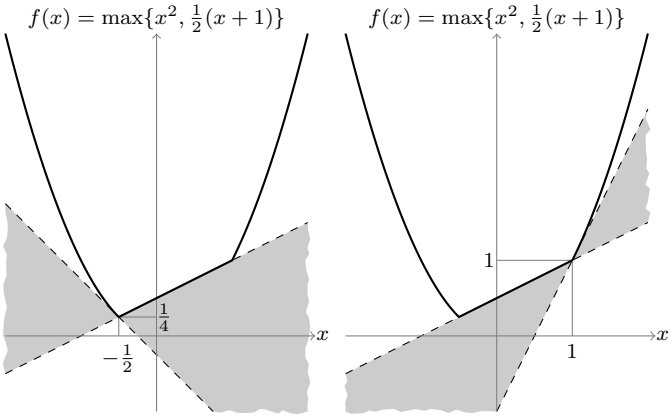

Fig. 1. A non-differentiable convex function. The shaded region represents the set of all lines defined in (9). The subdifferential of f at the two points $x = -1/2$ and $x = 1$ is the interval defined by the slopes of these lines.

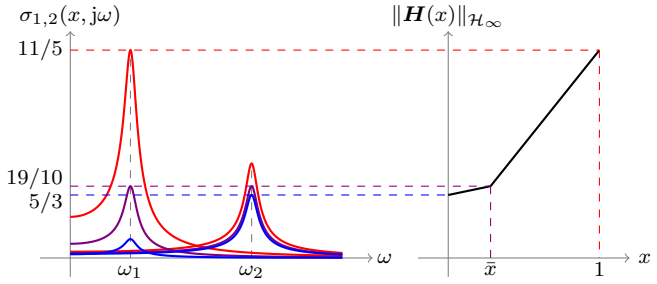

Fig. 2. Non-differentiability of $\mathcal{H}_\infty$-norm in case of multiple peaks.

is called the *subdifferential* of f at $\boldsymbol{x}$. The set $\partial f(\boldsymbol{x})$ is always closed and convex; if f is convex then, for any $\boldsymbol{x} \in \operatorname{intdom} f$, $\partial f(\boldsymbol{x})$ is also nonempty and bounded.

Figure 1 provides a graphical illustration for function

$$f(x) = \max\{x^2, \tfrac{1}{2}(x+1)\} = \begin{cases} x^2 & \text{if } x \in (-\infty, -\frac{1}{2}) \cup (1, +\infty) \\ \frac{1}{2}(x+1) & \text{if } x \in [-\frac{1}{2}, 1] \end{cases}$$

The two points where $f(x)$ is non-differentiable are $x_1 = -\frac{1}{2}$ and $x_2 = 1$, whose subdifferentials can be readily computed as $\partial f(x_1) = [-1, \frac{1}{2}]$ and $\partial f(x_2) = [\frac{1}{2}, 2]$, respectively.

C. An example

It turns out that the $\mathcal{H}_\infty$ norm $h(\boldsymbol{x})$ defined in (6) is convex but nonsmooth as a function of the perturbation parameters $\boldsymbol{x}$. We illustrate this fact through a simple example. Consider the transfer function

$$\boldsymbol{H}(x, s) = \begin{bmatrix} \frac{a_1 x + b_1}{(s-p_1)(s-p_1^*)} & 0 \\ 0 & \frac{a_2 x + b_2}{(s-p_2)(s-p_2^*)} \end{bmatrix}$$

with

$$a_1 = 1 \quad b_1 = 0.1 \quad p_1 = -0.1 + \mathrm{j}\,, \\ a_2 = 0.5 \quad b_2 = 1 \quad p_2 = -0.1 + 3\mathrm{j}\,.$$

Since $\boldsymbol{H}$ is a diagonal matrix, the singular values coincide with the magnitudes of H_{11} and H_{22},

$$\sigma_{1,2}(x, \mathrm{j}\omega) = \frac{|a_1 x + b_1|}{\sqrt{\omega^4 + \alpha_{1,2}\omega^2 + |p_{1,2}|^2}}$$

where

$$\alpha_{1,2} = 4\Re\{p_{1,2}\}^2 - 2|p_{1,2}|^2 \,.$$

A straightforward computation gives the locations of the maxima of the singular values

$$\omega_{1,2} = \sqrt{-\alpha_{1,2}/2}\,,$$

therefore, the $\mathcal{H}_\infty$-norm can be evaluated as

$$\|\boldsymbol{H}(x)\|_{\mathcal{H}_\infty} = \max\{\sigma_1(x, \mathrm{j}\omega_1), \sigma_2(x, \mathrm{j}\omega_2)\}$$

The $\mathcal{H}_\infty$-norm, as function of x, is depicted in the right panel of Fig. 2, while the left panel reports the plot of the two singular values for three different values of x. In particular, when $\sigma_1(x, \mathrm{j}\omega_1) = \sigma_2(x, \mathrm{j}\omega_2)$, that is when the parameter x is

$$\bar{x} = \frac{\beta_1 b_2 + \beta_2 b_1}{\beta_2 a_1 - \beta_1 a_2}, \quad \beta_{1,2} = 2\Re\{p_{1,2}\}^2 \sqrt{|p_{1,2}|^2 - \Re\{p_{1,2}\}}$$

the two maxima are equal, and the supremum over ω is attained at two different frequency points. As shown in Fig. 2, for $x = \bar{x}$, the $\|\boldsymbol{H}(x)\|_{\mathcal{H}_\infty}$ is non-differentiable. This example confirms that this case needs particular attention and motivates the introduction of subgradients and subdifferentials, to be derived in Section IV.

D. Subgradients and descent directions

Let f be a convex function. The *directional derivative* of f at $\boldsymbol{x}$ in direction $\boldsymbol{v} \in \mathbb{R}^n$ is defined as

$$f'(\boldsymbol{x};\ \boldsymbol{v}) = \lim_{h \to 0_+} \frac{f(\boldsymbol{x} + h\boldsymbol{v}) - f(\boldsymbol{x})}{h}. \tag{10}$$

If f is differentiable at $\boldsymbol{x}$, then $f'(\boldsymbol{x};\ \boldsymbol{v}) = \boldsymbol{v}^\top \nabla f(\boldsymbol{x})$. If f is non differentiable at $\boldsymbol{x}$, then

$$f'(\boldsymbol{x};\ \boldsymbol{v}) = \max_{\boldsymbol{g} \in \partial f(\boldsymbol{x})} \boldsymbol{v}^\top \boldsymbol{g}. \tag{11}$$

Direction $\boldsymbol{v}$ is called a *descent direction* for f at $\boldsymbol{x}$, if $f'(\boldsymbol{x};\ \boldsymbol{v}) < 0$. The meaning is indeed intuitive: if one moves away from $\boldsymbol{x}$ along a descent direction $\boldsymbol{v}$, then f locally *decreases*. If f is differentiable at $\boldsymbol{x}$, then it is immediate to verify that $\boldsymbol{v} = -\nabla f(\boldsymbol{x})$ is a descent direction (actually, it is the *steepest* descent direction), and this fact is exploited in the well-known standard *gradient descent* algorithm for minimizing f, where the solution is found by iteratively applying an update rule of the following type

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha_k \nabla f(\boldsymbol{x}^{(k)}),$$

where α_k is a suitable stepsize. Clearly, such a minimization scheme cannot be adopted as-is if f is non smooth. The idea in subgradient algorithms is to use a subgradient instead of the gradient in the update step, thus obtaining an update rule of the type

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha_k \boldsymbol{g}^{(k)}, \tag{12}$$

where $g^{(k)} \in \partial f(\boldsymbol{x}^{(k)})$. It is important to remark that this subgradient step does not in general decrease the objective value. This is due to the fact that a negative subgradient $-\boldsymbol{g}$ need not be a descent direction for f (contrary to what happens in the smooth case, where the negative gradient $-\nabla f$ is always the steepest descent direction). If desired, however, the method can be modified so that the subgradient step in (12) is indeed a descent step. In order to do this, we need to select an appropriate subgradient in the subdifferential. Indeed, if we have available the whole subdifferential of f, we can search this set for a subgradient which is also a descent direction. To this end, it suffices to minimize $f'(\boldsymbol{x};\ \boldsymbol{v})$ over all directions, and check if the minimum is negative

$$\begin{aligned}
\min_{\|\boldsymbol{v}\|=1} f'(\boldsymbol{x};\ \boldsymbol{v}) &= \min_{\|\boldsymbol{v}\|=1} \max_{\boldsymbol{g} \in \partial f(\boldsymbol{x})} \boldsymbol{v}^\top \boldsymbol{g} \\
\text{[using saddle point theorem]} &= \max_{\boldsymbol{g} \in \partial f(\boldsymbol{x})} \min_{\|\boldsymbol{v}\|=1} \boldsymbol{v}^\top \boldsymbol{g} \\
\text{[min is achieved for } \boldsymbol{v}^* = -\boldsymbol{g}/\|\boldsymbol{g}\|] &= \max_{\boldsymbol{g} \in \partial f(\boldsymbol{x})} -\|\boldsymbol{g}\| \\
&= -\min_{\boldsymbol{g} \in \partial f(\boldsymbol{x})} \|\boldsymbol{g}\| \,.
\end{aligned}$$

For additional details and proofs see [38]. Therefore, we may solve the convex optimization problem

$$\boldsymbol{g}^* = \arg \min_{\boldsymbol{g} \in \partial f(\boldsymbol{x})} \|\boldsymbol{g}\| \tag{13}$$

in order to find a subgradient which is also a descent direction.

IV. Subdifferential of $\mathcal{H}_\infty$-norm

In this section we illustrate how to compute the subdifferential and a subgradient for the $\mathcal{H}_\infty$ norm. To this end, we first state preliminary results on subgradients of functions defined as the supremum of parameterized functions. Let us thus consider a function h defined, as in (6), as the supremum over a possibly infinite family of functions:

$$h(\boldsymbol{x}) = \sup_{\omega \in \Omega} h_\omega(\boldsymbol{x}).$$

If Ω is compact and the map $\omega \mapsto h_\omega(\boldsymbol{x})$ is continuous for all $\boldsymbol{x}$, then [39]

$$\partial h(\boldsymbol{x}) = \mathrm{conv}\,\big\{ \cup \{\partial h_\omega(\boldsymbol{x}) : h_\omega(\boldsymbol{x}) = h(\boldsymbol{x})\}\big\}.$$

In words, the subdifferential of h at $\boldsymbol{x}$ is the convex hull of the union of the subdifferentials of the h_ω functions that attain the supremum. If the supremum is attained at a unique $\omega = \bar{\omega}$, then the above statement reduces to

$$\partial h(\boldsymbol{x}) = \partial h_{\bar{\omega}}(\boldsymbol{x}).$$

Let us now recall the definition of function $h(\boldsymbol{x})$ in (6):

$$h(\boldsymbol{x}) = \|\boldsymbol{H}(\boldsymbol{C}_p)\|_{\mathcal{H}_\infty} = \|\bar{\boldsymbol{H}}(\boldsymbol{x})\|_{\mathcal{H}_\infty} = \sup_{\omega \in \mathbb{R}} h_\omega(\boldsymbol{x}) \tag{14}$$

where

$$h_\omega(\boldsymbol{x}) = \sigma_1(\bar{\boldsymbol{H}}(\boldsymbol{x}, \mathrm{j}\omega))\,.$$

We shall next derive explicit expressions for the subdifferential of $h(\boldsymbol{x})$, under different cases. Some of the concepts presented next are adapted from [40] (see also the references therein), where the subgradient of $\|\mathcal{G}(x)\|_{\mathcal{H}_\infty}$ is characterized for any general smooth operator $\mathcal{G}$.

A. Simple singular value at a single frequency

We start by assuming that the supremum over ω in (14) is attained at a single frequency point $\bar{\omega}$, we will release this assumption later. Let

$$\bar{\boldsymbol{H}}(\boldsymbol{x}, \mathrm{j}\bar{\omega}) = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H \tag{15}$$

be the singular value decomposition of the transfer function evaluated at frequency $\omega = \bar{\omega}$. Further, let ℓ be the dimension of the eigenspace of $\bar{\boldsymbol{H}}(\boldsymbol{x}, \mathrm{j}\bar{\omega})\bar{\boldsymbol{H}}^H(\boldsymbol{x}, \mathrm{j}\bar{\omega})$ associated with the largest eigenvalue, and let $\boldsymbol{U}_1$, $\boldsymbol{V}_1$ be the first ℓ columns of $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively (we use the vector notation $\boldsymbol{u}_1$, $\boldsymbol{v}_1$ when $\ell = 1$). Finally, in order to make the notation more compact, we define

$$\boldsymbol{\Phi}(\mathrm{j}\omega) = (\mathrm{j}\omega\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B}\,. \tag{16}$$

The top and bottom left panels in Fig. 3 provide a graphical illustration for the cases $\ell = 1$ and $\ell = 2$, respectively.

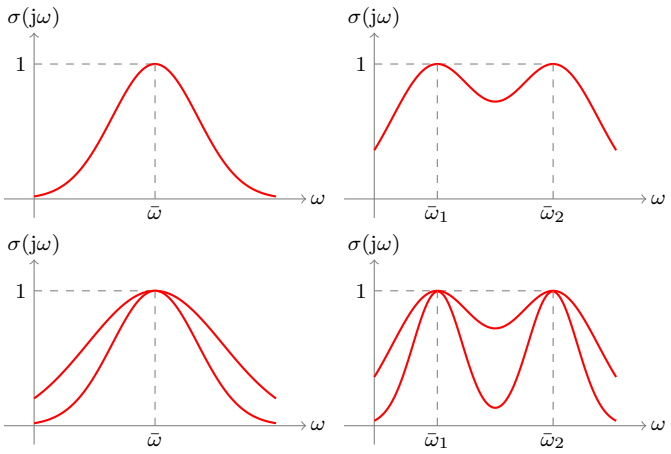

Fig. 3. Frequency-dependent trajectory of the singular values of a transfer function with unit $\mathcal{H}_\infty$ norm. Four possible cases are depicted, corresponding to a largest singular value with unit and higher multiplicity (top and bottom rows) at single and multiple frequencies (left and right panels).

In the simplest case with $\ell = 1$ the function $h(\boldsymbol{x})$ is differentiable and the gradient $\nabla h(\boldsymbol{x})$ can be easily derived. Let us denote the maximum singular value as a function of design parameters $\boldsymbol{C}_p$ as $\sigma_1(\boldsymbol{C}_p)$. If we apply a small perturbation $\mathrm{d}\boldsymbol{C}_p$, we obtain the following first order expansion of transfer function

$$\boldsymbol{H}(\boldsymbol{C}_p + \mathrm{d}\boldsymbol{C}_p, \mathrm{j}\bar{\omega}) = \boldsymbol{H}(\boldsymbol{C}_p, \mathrm{j}\bar{\omega}) + \mathrm{d}\boldsymbol{H}(\boldsymbol{C}_p, \mathrm{j}\bar{\omega})$$

where

$$\mathrm{d}\boldsymbol{H}(\boldsymbol{C}_p, \mathrm{j}\bar{\omega}) \simeq \mathrm{d}\boldsymbol{C}_p\boldsymbol{\Phi}(\mathrm{j}\bar{\omega})$$

A corresponding perturbation will be induced in the maximum singular value as

$$\sigma_1(\boldsymbol{C}_p + \mathrm{d}\boldsymbol{C}_p) = \sigma_1(\boldsymbol{C}_p) + \mathrm{d}\sigma_1(\boldsymbol{C}_p)\,.$$

Standard results on eigenvalue perturbation can be applied in order to derive a first-order approximation of $\mathrm{d}\sigma_1(\boldsymbol{C}_p)$ (see [41])

$$\begin{aligned}\mathrm{d}\sigma_1(\boldsymbol{C}_p) &\simeq \Re\left\{\boldsymbol{u}_1^H \mathrm{d}\boldsymbol{H}(\boldsymbol{C}_p, \mathrm{j}\bar{\omega})\boldsymbol{v}_1\right\} \\ &\simeq \Re\left\{\boldsymbol{u}_1^H \mathrm{d}\boldsymbol{C}_p\boldsymbol{\Phi}(\mathrm{j}\bar{\omega})\boldsymbol{v}_1\right\} \\ &= \Re\left\{\boldsymbol{v}_1^\top\boldsymbol{\Phi}(\mathrm{j}\bar{\omega})^\top \otimes \boldsymbol{u}_1^H\right\}\mathrm{vec}\,(\mathrm{d}\boldsymbol{C}_p)\end{aligned} \tag{17}$$

where $\otimes$ denotes the Kronecker product [42] and $\Re$ extracts the real part. Expression (17) can be restated in terms of the vectorized decision variables $\boldsymbol{x}$ as

$$\mathrm{d}h(\boldsymbol{x}) = \Re\left\{\boldsymbol{v}_1^\top\boldsymbol{\Phi}(\mathrm{j}\bar{\omega})^\top \otimes \boldsymbol{u}_1^H\right\}\mathrm{d}\boldsymbol{x}\,,$$

where $\mathrm{d}\boldsymbol{x} = \mathrm{vec}(\mathrm{d}\boldsymbol{C}_p)$. This result implies that the subdifferential of $h(\boldsymbol{x})$ has only one element, which corresponds to the gradient

$$\begin{aligned}\nabla h(\boldsymbol{x}) &= \Re\left\{\boldsymbol{v}_1^\top\boldsymbol{\Phi}(\mathrm{j}\bar{\omega})^\top \otimes \boldsymbol{u}_1^H\right\}^\top \\ &= \mathrm{vec}\left(\Re\{\boldsymbol{\Phi}(\mathrm{j}\bar{\omega})\boldsymbol{v}_1\boldsymbol{u}_1^H\}^\top\right).\end{aligned} \tag{18}$$

B. Simple singular value at multiple frequencies

We now consider the more general case of the supremum in (14) being attained at more than one frequency, as depicted in Fig. 3, top right panel. Let us define as $\Omega = \{\bar{\omega}_1, \ldots, \bar{\omega}_p\}$ the set of frequencies where

$$h(\boldsymbol{x}) = \sup_{\omega\in\mathbb{R}} h_\omega(\boldsymbol{x}) = h_{\bar{\omega}_i}(\boldsymbol{x}), \quad \forall\bar{\omega}_i \in \Omega\,.$$

As a generalization to (15), we denote the singular value decomposition of the transfer function at each frequency point $\bar{\omega}_i \in \Omega$ as

$$\bar{\boldsymbol{H}}(\boldsymbol{x}, \mathrm{j}\bar{\omega}_i) = \boldsymbol{U}^{(i)}\boldsymbol{\Sigma}^{(i)}\boldsymbol{V}^{(i)H}, \quad \bar{\omega}_i \in \Omega\,. \tag{19}$$

Let ℓ_i be the dimension of the eigenspace of $\bar{\boldsymbol{H}}(\boldsymbol{x}, \mathrm{j}\bar{\omega}_i)\bar{\boldsymbol{H}}^H(\boldsymbol{x}, \mathrm{j}\bar{\omega}_i)$ associated with its largest eigenvalue, and let $\boldsymbol{U}_1^{(i)}$, $\boldsymbol{V}_1^{(i)}$ be the first ℓ_i columns of $\boldsymbol{U}^{(i)}$ and $\boldsymbol{V}^{(i)}$, respectively. When the largest singular values have unit multiplicity with $\ell_1 = \cdots = \ell_p = 1$, the subdifferential can be computed following (IV) as the convex hull of the individual vectors computed as in (18) at each individual frequency $\bar{\omega}_i$. The result is

$$\partial h(\boldsymbol{x}) = \left\{\mathrm{vec}\left(\sum_{i=1}^{p} y_i\Re\{\boldsymbol{\Phi}(\mathrm{j}\bar{\omega}_i)\boldsymbol{v}^{(i)}\boldsymbol{u}^{(i)H}\}^\top\right)\right\}, \tag{20}$$

where the p-tuple $(y_1, \ldots, y_p)$ belongs to the set

$$\Sigma_p = \left\{(y_1, \ldots, y_p) : y_i \geq 0, \sum_{i=1}^{p} y_i = 1\right\}.$$

C. Multiple singular values at a single frequency

The bottom left panel of Fig. 3 depicts the case where the supremum in (14) is attained at a single frequency, but the largest singular value has multiplicity larger than one. So, the singular value decomposition in (15) holds with $\ell > 1$. In this case, it can be shown (see eq. (9) in [40] and Theorem 3 in [43]) that the subdifferential of the $\mathcal{H}_\infty$ norm can be expressed as

$$\partial h(\boldsymbol{x}) = \left\{\mathrm{vec}\left(\Re\{\boldsymbol{\Phi}(\mathrm{j}\bar{\omega})\boldsymbol{V}_1\boldsymbol{Y}\boldsymbol{U}_1^H\}^\top\right) : \boldsymbol{Y} \in \mathcal{Y}_\ell\right\}, \tag{21}$$

where $\mathcal{Y}_\ell$ is the set of positive semidefinite symmetric $\ell \times \ell$ matrices having unit trace,

$$\mathcal{Y}_\ell = \{\boldsymbol{Y} \in \mathbb{R}^{\ell,\ell} :\ \boldsymbol{Y} = \boldsymbol{Y}^\top,\ \boldsymbol{Y} \geq 0,\ \mathrm{Tr}\,\boldsymbol{Y} = 1\}\,. \tag{22}$$

D. Multiple singular values at multiple frequencies

The most general case, depicted in the bottom right panel of Fig. 3, occurs when the supremum in (14) is attained at multiple frequencies $\bar{\omega}_i \in \Omega$, and at those frequencies the multiplicity of the largest singular value is larger than one. So, the singular value decomposition (19) applies with $\ell_i > 1$ for some i. For this case, it can be shown by combining (20) and (21) that the subdifferential of the $\mathcal{H}_\infty$ norm can be expressed as

$$\partial h(\boldsymbol{x}) = \Big\{ \operatorname{vec}\Big(\sum_{i=1}^{p} \Re\{ \boldsymbol{\Phi}(\mathrm{j}\bar{\omega}_i) \boldsymbol{V}^{(i)} \boldsymbol{Y}_i \boldsymbol{U}^{(i)H} \}^\top \Big) \Big\}, \quad (23)$$

where the p-tuple $(\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_p)$ ranges over the set

$$\mathcal{Y}_{(\ell_1,\ldots,\ell_p)} = \{ \boldsymbol{Y} : \boldsymbol{Y}_i \in \mathbb{R}^{\ell_i,\ell_i},\ \boldsymbol{Y}_i = \boldsymbol{Y}_i^\top,\ \boldsymbol{Y}_i \geq 0,\ \sum_{i=1}^{p} \operatorname{Tr} \boldsymbol{Y}_i = 1 \}. \quad (24)$$

E. Descent direction in the case of simple eigenvalues

The most prominent case (i.e., the one that occurs in our practical application) is that either the $\mathcal{H}_\infty$ norm is attained at a single frequency, or it is attained at multiple frequencies, but in both situations the maximum eigenvalue happens to be simple. It is seen from (20) that, in the case when $\ell_i = 1$, $i = 1, \ldots, p$, all subgradients are found as convex combinations (weighted average) of the individual 'gradients' at $\bar{\omega}_i$. Let $\boldsymbol{y} = \big[y_1, \ldots, y_p\big]^\top$, and

$$\boldsymbol{g}_i = \sum_{i=1}^{p} y_i \Re\{ \boldsymbol{\Phi}(\mathrm{j}\bar{\omega}_i) \boldsymbol{v}^{(i)} \boldsymbol{u}^{(i)H} \}^\top,$$

$$\Sigma_p = \{ \boldsymbol{y} : y_i \geq 0, i = 1, \ldots, p, \sum_{i=1}^{p} y_i = 1 \}.$$

In order to find the steepest descent direction we need to solve a simple quadratic optimization problem (13)

$$\min \| \sum_{i=1}^{p} y_i \boldsymbol{g}_i \|_F^2, \quad \text{s.t. } \boldsymbol{y} \in \Sigma_p,$$

and then choose $\boldsymbol{g}^* = \sum_{i=1}^{p} y_i^* \boldsymbol{g}_i$ as a descent direction, where y_i^* are the coefficients of the optimal solution. By using the definition of Frobenius norm and the properties of vec operator, the optimization problem can be written as

$$\min \boldsymbol{y}^\top \boldsymbol{G} \boldsymbol{y}, \quad \text{s.t. } \boldsymbol{y} \in \Sigma_p,$$

where $\boldsymbol{G} \in \mathbb{R}^{p,p}$ is the matrix with the entries

$$(\boldsymbol{G})_{ij} = \operatorname{vec}(\boldsymbol{g}_i)^\top \operatorname{vec}(\boldsymbol{g}_j).$$

V. A Projected Subgradient Algorithm

A. General description

Consider the generic optimization problem

$$\min_{\boldsymbol{x}} h(\boldsymbol{x}), \quad \text{s.t. } \boldsymbol{x} \in \mathcal{X}, \quad (25)$$

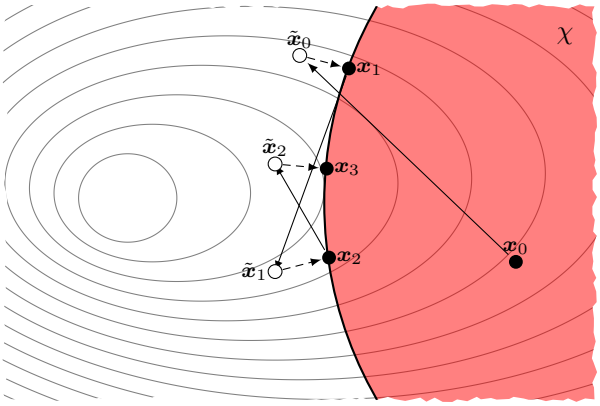

Fig. 4. Projected subgradient algorithm. The contour lines of the objective function $h(\boldsymbol{x})$ are depicted in gray and the feasible set $\mathcal{X}$ is the red shaded region. The first three iterations of the subgradient algorithm are depicted, distinguishing between the step towards the direction $-\boldsymbol{g}^{(k)}$ (black solid line) and the projection step onto the feasible set $\mathcal{X}$ (black dashed line).

where $h(\boldsymbol{x})$ and $\mathcal{X}$ are a convex function and a nonempty convex compact set, respectively. This problem admits a global optimal solution that can be found with a simple iterative scheme, as illustrated in Fig. 4: we pick a generic initial point $\boldsymbol{x}^{(0)}$, and we generate the next point by performing a step in the direction $-\boldsymbol{g}^{(0)}$

$$\tilde{\boldsymbol{x}}^{(0)} = \boldsymbol{x}^{(0)} - \alpha_0 \boldsymbol{g}^{(0)}$$

where $\boldsymbol{g}^{(0)}$ is a subgradient of the function $h(\boldsymbol{x})$ in $\boldsymbol{x}_0$ and α_0 is a suitable stepsize (the specific rule for computing the stepsizes is reported in eq. (37) in the Appendix). Generally, $\tilde{\boldsymbol{x}}^{(0)}$ does not belong to the feasible set $\mathcal{X}$ (highlighted in red in Fig. 4), therefore we project the $\tilde{\boldsymbol{x}}^{(0)}$ on the set $\mathcal{X}$ obtaining the new candidate solution

$$\boldsymbol{x}^{(1)} = [\tilde{\boldsymbol{x}}^{(0)}]_{\mathcal{X}} = [\boldsymbol{x}^{(0)} - \alpha_0 \boldsymbol{g}^{(0)}]_{\mathcal{X}}$$

where $[\cdot]_{\mathcal{X}}$ is the operator that performs an orthogonal Euclidean projection of its argument onto $\mathcal{X}$. The above process is repeated following the iterative scheme

$$\boldsymbol{x}^{(k+1)} = [\boldsymbol{x}^{(k)} - \alpha_k \boldsymbol{g}^{(k)}]_{\mathcal{X}} \quad (26)$$

until convergence. Technical details on fundamental assumptions of this method and a proof of its convergence are reported in Appendix A.

B. Passivity enforcement via a projected subgradient algorithm

We illustrate in this section how the basic projected subgradient algorithm described in Section V-A can be applied in order to solve our main passivity enforcement problem (4). Recalling the problem statement (8)

$$\min_{\boldsymbol{x}} h(\boldsymbol{x}), \quad \text{s.t. } \boldsymbol{x} \in \mathcal{X}_\nu\,. \quad (27)$$

we consider the set of all macromodels $\mathcal{X}_\nu$ that differ less than ν from the nominal system, where ν is some prescribed and controlled accuracy. We search this set for the macromodel with minimum $\mathcal{H}_\infty$ norm by applying the above projected

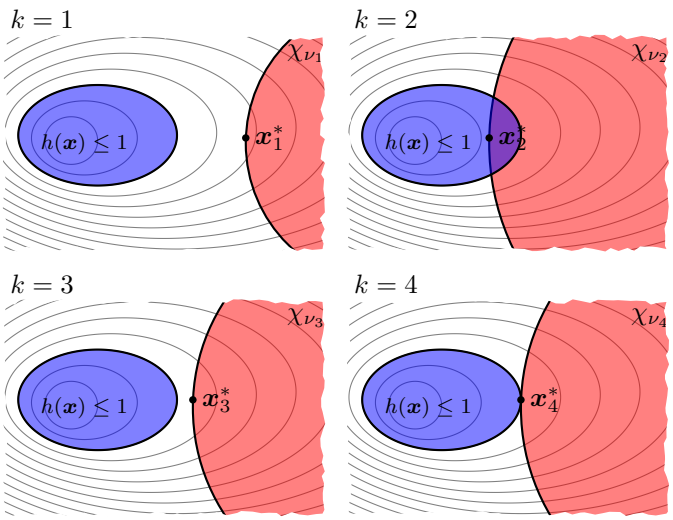

Fig. 5. First three iterations of the outer bisection loop on the macromodel perturbation amount ν. For each iteration, $\boldsymbol{x}^*$ denotes the optimal solution of problem (27). Bottom right panel depicts the situation at convergence, where ν^* denotes the optimal accuracy of the optimal passive macromodel $\bar{\boldsymbol{H}}(\boldsymbol{x}^*, s)$.

subgradient scheme, and we denote the optimal solution as $\boldsymbol{x}^*$. The following two cases may apply

1) if $h(\boldsymbol{x}^*) \le 1$, we have found a passive macromodel with controlled accuracy with respect to nominal macromodel; in other words, problem (27) with the additional passivity constraint $h(\boldsymbol{x}) \le 1$ is feasible;
2) if instead $h(\boldsymbol{x}^*) > 1$, we can conclude that there exist no passive macromodel which deviates less than ν from the original model.

We then argue that there exists an optimal accuracy ν^* such that problem

$$\min_{\boldsymbol{x}} h(\boldsymbol{x}), \quad \text{s.t. } \boldsymbol{x} \in \mathcal{X}_{\nu^*}, \quad h(\boldsymbol{x}) \le 1 \tag{28}$$

is feasible. We will look for the optimal accuracy ν^* by an outer bisection loop, as illustrated in Fig. 5 and described below.

Let us assume that at the first iteration $k = 1$ (top left panel) problem (28) is not feasible. Therefore, the accuracy ν_1 is too stringent and the set $\mathcal{X}_{\nu_1}$ is too small. We then need to relax the accuracy to a larger value $\nu_2 > \nu_1$, which makes problem (28) feasible. The top right panel in Fig. 5 illustrates this situation, highlighting that the intersection of sets $\mathcal{X}_{\nu_2}$ and $\{\boldsymbol{x} : h(\boldsymbol{x}) \le 1\}$ is nonempty. The optimal accuracy is such that $\nu^* \in [\nu_1, \nu_2]$. We then define $\nu_3 = (\nu_1 + \nu_2)/2$ and solve problem (28) again (bottom left panel). This bisection process on ν is repeated until convergence (bottom right panel). We remark that we do not need to obtain the optimal solution $\boldsymbol{x}^*$ of the projected subgradient problem (28). Rather, we need to determine only the feasibility of this problem. If the problem is feasible, we decrease ν. If the problem is infeasible, we increase ν.

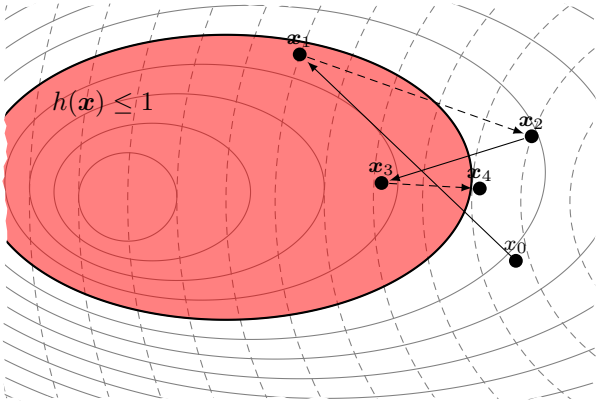

Fig. 6. Alternate subgradient algorithm. The level sets of $h(\boldsymbol{x})$ and $f(\boldsymbol{x})$ are depicted with solid and dashed gray lines, respectively. The region $h(\boldsymbol{x}) \le 1$ is colored in red. The iteration steps in a direction in $\partial h(\boldsymbol{x})$ and $\partial f(\boldsymbol{x})$ are depicted with solid and dashed arrows, respectively.

VI. An Alternate Subgradient Algorithm

A. General description

Consider the following optimization problem

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}), \quad \text{s.t. } h(\boldsymbol{x}) \le 1. \tag{29}$$

where, $h(\boldsymbol{x})$ and $f(\boldsymbol{x})$ are convex functions. Figure 6 depicts an alternate subgradient algorithm for solving this problem. We start from some initial point $\boldsymbol{x}_0$, which is assumed, for the purpose of the example, to be outside the feasible region $h(\boldsymbol{x}) \le 1$ highlighted in red. The first iteration step has thus the purpose of decreasing $h(\boldsymbol{x})$. We perform a step

$$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} - \alpha_0 \boldsymbol{g}^{(0)}, \tag{30}$$

where $\boldsymbol{g}^{(0)}$ is a subgradient of $h(\boldsymbol{x})$ at $\boldsymbol{x}^{(0)}$ and α_0 is some suitable stepsize (see eq. (33)). The updated point $\boldsymbol{x}^{(1)}$ may satisfy the constraint $h(\boldsymbol{x}) - 1 \le 0$ (as in the example depicted in Figure 6), but it is not necesarily the point corresponding to the minimum of the objective function $f(\boldsymbol{x})$. The second step is thus performed, with a suitable stepsize α_1, so to decrease the objective:

$$\boldsymbol{x}^{(2)} = \boldsymbol{x}^{(1)} - \alpha_1 \boldsymbol{g}^{(1)}, \tag{31}$$

where $\boldsymbol{g}^{(1)}$ is a subgradient of the objective function $f(\boldsymbol{x})$. The process is repeated by performing steps

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha_k \boldsymbol{g}^{(k)},$$

where the direction of each step is defined according to the alternate subgradient rule

$$\boldsymbol{g}^{(k)} \in \begin{cases} \partial f(\boldsymbol{x}^{(k)}) & \text{if } h(\boldsymbol{x}^{(k)}) \le 1 \\ \partial h(\boldsymbol{x}^{(k)}) & \text{if } h(\boldsymbol{x}^{(k)}) > 1. \end{cases} \tag{32}$$

Technical details on the implementation of this scheme, and a proof of its convergence are reported in Appendix B.

B. Passivity enforcement via alternate subgradients

Macromodel passivity can be easily achieved by applying the alternate subgradient scheme described in Section VI-A. We can choose as the objective function the square of the

Frobenius norm of the perturbation, (5): $f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{x}\|_2^2$. Our main problem thus becomes

$$\min \frac{1}{2}\|\boldsymbol{x}\|_2^2, \quad \text{s.t. } h(\boldsymbol{x}) < 1\,,$$

which is fully equivalent to (4). Since $f(\boldsymbol{x})$ is smooth and differentiable, its subdifferential contains only one element, which coincides with its gradient $\boldsymbol{x}$. Each step of the alternate subgradient algorithm thus applies the following update

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha_k \boldsymbol{g}^{(k)}$$

where

$$\boldsymbol{g}^{(k)} = \begin{cases} \boldsymbol{g}_f^{(k)} = \boldsymbol{x}^{(k)} & \text{if } h(\boldsymbol{x}^{(k)}) \le 1\,, \\ \boldsymbol{g}_h^{(k)} \in \partial h(\boldsymbol{x}^{(k)}) & \text{if } h(\boldsymbol{x}^{(k)}) > 1\,. \end{cases}$$

The step size α_k is computed by using the adaptive method described in Appendix C, i.e.,

$$\alpha_k = \frac{-G_k\zeta_{k-1} + \sqrt{G_k^2\zeta_{k-1}^2 + R^2 + \xi_{k-1}}}{G_k} \tag{33}$$

where

$$G_k^2 = \|\boldsymbol{g}_f^{(k)}\|^2 + \|\boldsymbol{g}_h^{(k)}\|^2\,,$$

constants ζ_{k-1} and ξ_{k-1} are defined in (36) and R is any constant such that $\|x^{(1)} - x^*\| \le R$. In other words, for each step, feasible or infeasible, it is necessary to calculate both subgradients in order to derive the correct value for the step size.

VII. Numerical Experiments

A. PCB interconnect

We applied the proposed passivity enforcement algorithms to a practical case. The 4-port scattering matrix of a coupled Printed Circuit Board interconnect has been measured from DC up to 20 GHz with resolution 10 MHz, obtaining the raw data $S_{ij}(\mathrm{j}\omega_k)$, $i,j \in \{1,\ldots,4\}$ and $k \in \{1,\ldots,2000\}$. These samples have been processed by the well known Vector Fitting algorithm [4], [5], [6] to obtain an initial macromodel (1), with $n = 272$ states and $n_y = n_u = 4$ inputs/outputs. Figure 7 demonstrates the accuracy of this initial model by comparing its responses to the raw data. However, even with this aggressive accuracy the model exhibits some non-passive bands between 0 and 4 GHz highlighted in Figure 10, even if all singular values of the raw data are unitary bounded at all frequencies.

The projected subgradient method in Section V-B has been applied to the model in order to enforce its passivity. The initial Frobenius norm of the state matrix under perturbation is $\|\boldsymbol{C}\|_F = 3.33 \times 10^6$. We then selected an accuracy $\varepsilon = 1$ on the unknown perturbation term in order to stop the outer bisection loop on ν. The algorithm required 22 outer iterations before reaching this accuracy. The number of inner iterations resulted highly dependent on the current value of ν. During early outer iterations, when the current solution estimate is still far from the optimal solution, only few inner iterations (about 20) are sufficient to establish if the problem (27) is feasible or not. When ν approaches its optimal value ν^*, the number of inner iterations increases. This effect is intrinsic in the algorithm structure.

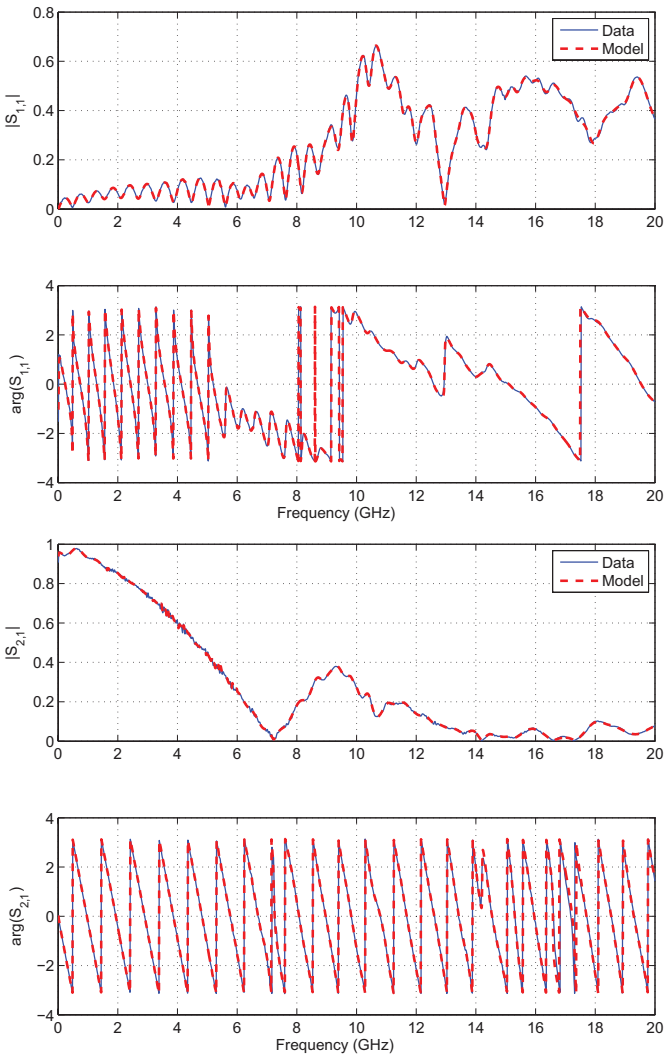

Fig. 7. Comparisons between original data and model generated with Vector Fitting algorithm [4], [5], [6].

The alternate subgradient algorithm is more efficient, since significantly less total iterations are required. If a stopping condition is applied to enforce the theoretical error estimate δ_k (see (35) in Appendix C) to be less than ε, about 900 total iterations are needed. However, very accurate estimates are available after only few tens of iterations. As a comparison, one of the best passivity enforcement algorithms available in the literature [12], [19], [24] obtains, for this case, a solution with $\|\boldsymbol{C}_p^*\|_F \simeq 500$. The presented scheme reaches a better accuracy after about 100 iterations, as depicted in Figure 8. This figure reports with different colors (blue and red, respectively) the values of the objective function $\|\boldsymbol{C}_p^*\|_F$ for both feasible and infeasible iterations. Figure 9 reports the evolution of the $\mathcal{H}_\infty$ norm through iterations.

Finally, Figure 10 and Figure 11 compare the frequency-dependent singular values of original and optimal perturbed models obtained by the two proposed schemes. The two figures are almost undistinguishable since the two algorithms provably

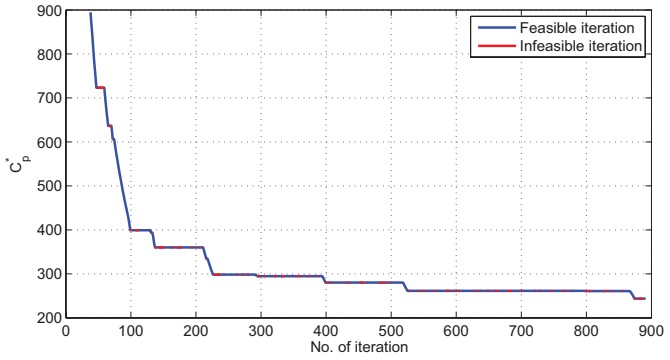

Fig. 8. Trend of objective function C_p^* for alternate subgradient algorithm of Section VI. Feasible and infeasible iterations are plotted in blue and red, respectively.

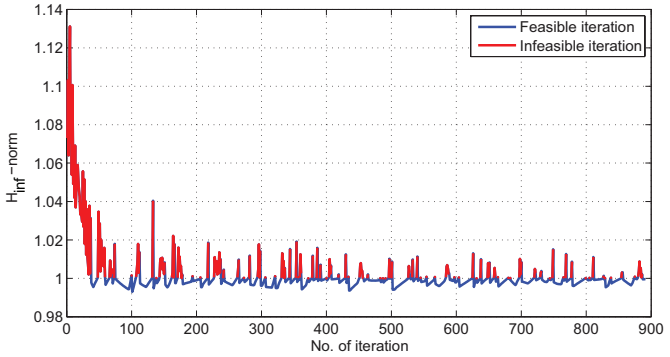

Fig. 9. Trend of $\mathcal{H}_\infty$-norm for alternate subgradient algorithm of Section VI. Feasible and infeasible iterations are plotted in blue and red, respectively.

converge to the same optimal solution.

For this example, the computational cost is about 0.5 seconds per iteration on a standard laptop with a 2 GHz clock. Passivity enforcement is thus achieved in few minutes with both projected and alternate subgradient techniques. It is important to note that the total runtime can be traded with accuracy with both schemes. Since upper bounds on the macromodel perturbation are available at each step of the algorithm, iterations can be stopped at any time as soon as this upper bound is satisfactory, even if successive iterations would further improve the solution. This possibility is ruled out for common non-convex passivity enforcement schemes.

B. Guaranteed convergence

We show in this second example the reliability of the new proposed strategy by processing a model for which the state-of-the-art passivity enforcement methods [12], [19], [24] fail. The nominal macromodel is obtained by applying the Vector Fitting algorithm to the scattering responses of a sharp filter. The model order is $n = 60$, with a number of ports $n_y = n_u = 2$.

We analyze first the strategy presented in [12], based on an iterative perturbation of the model, where the constraints are derived from a linearized expression of the singular values as a function of residues. Figure 12 depicts the singular values

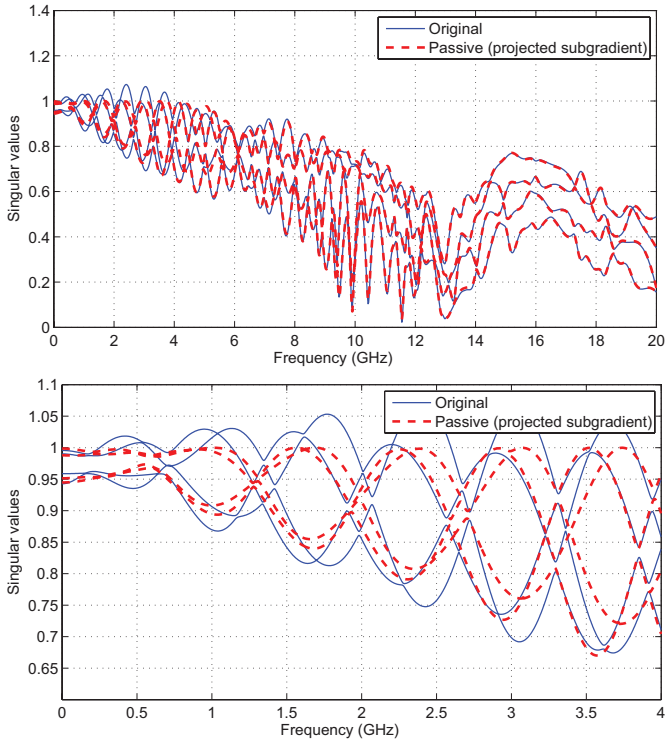

Fig. 10. Passivity enforcement of an electrical interconnect model via the projected subgradient algorithm of Section V. Singular values of original non-passive (solid blue) and perturbed passive (dashed red) models are plotted versus frequency.

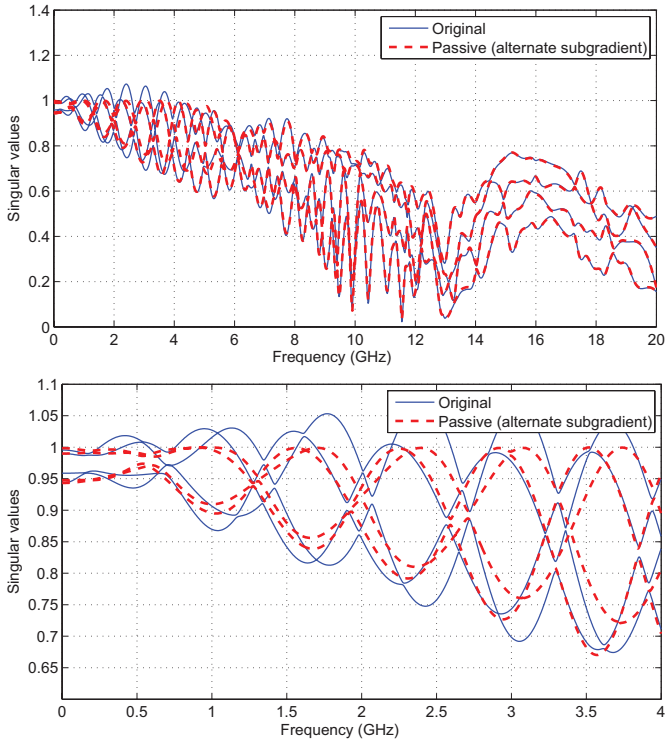

Fig. 11. As in Fig. 10, but using the alternate subgradient algorithm of Section VI.

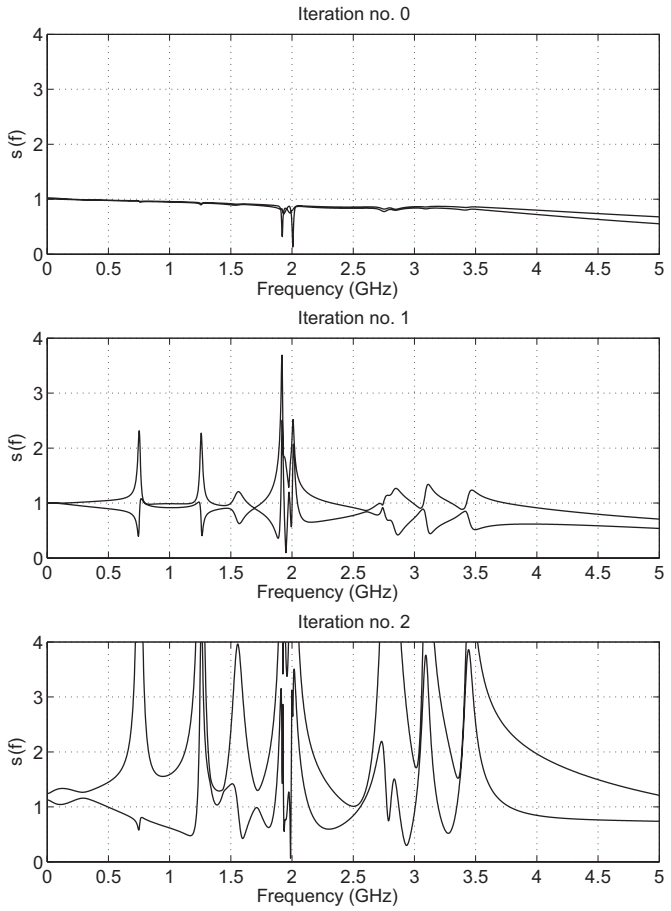

Fig. 12. Singular values trajectories of the first iterations of the passivity enforcement scheme [12]. Similar results are obtained with the scheme [19].

trajectories during the first iterations of the tentative passivity enforcement. The top panel shows that the original singular values of the unperturbed model are only slightly larger than one, with a corresponding small non-passive band at low frequencies. Starting from the first iteration, we see that very large perturbations of the singular values and model responses are induced throughout the frequency axis. This is a clear evidence of the ill-conditioned nature of this scheme, which diverges in very few iterations. We obtained the same negative results (not shown here) with the Hamiltonian perturbation scheme of [19]. Both these schemes solve a nonconvex formulation of the passivity enforcement. Therefore, convergence is not guaranteed.

We then applied to this model the alternate subgradient algorithm presented in Section VI. After 30 iterations we obtained a passive model with a very good accuracy, as demonstrated in figure 13, where the scattering parameters of original and passive model are compared.

For this example, we investigated also an acceleration strategy on the algorithm, which is useful to improve the convergence speed. We implemented a method where the update direction is a conic combination of the current subgradient and the last search direction,

$$\boldsymbol{s}^{(k)} = \boldsymbol{g}^{(k)} + \beta_k \boldsymbol{s}^{(k-1)} .$$

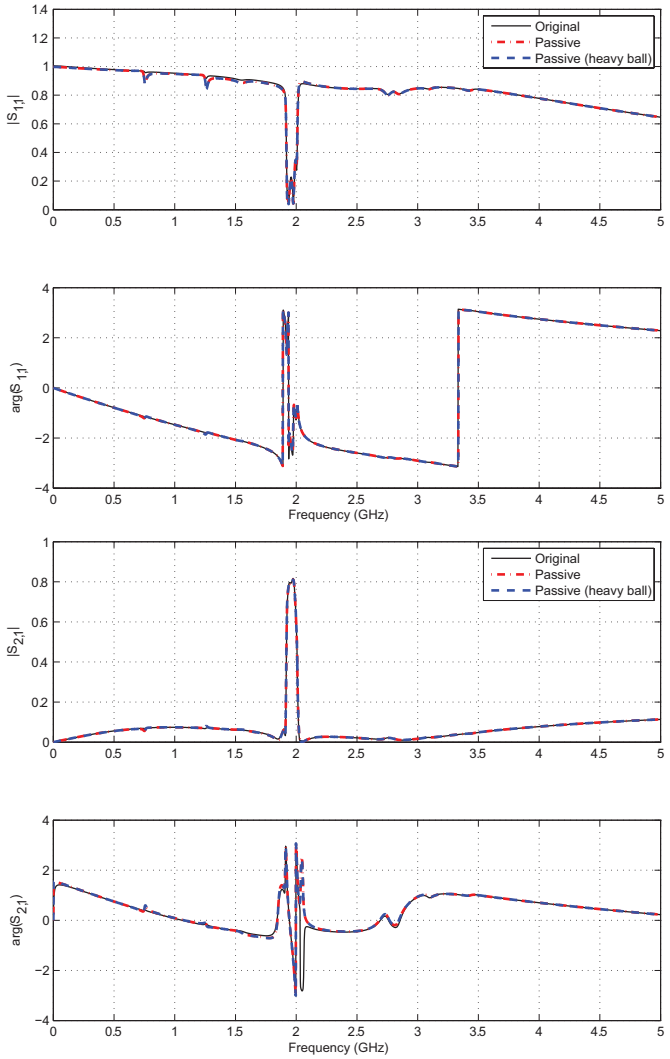

Fig. 13. Comparison between original and perturbed scattering responses.

In the literature, these acceleration techniques are referred as the "heavy ball" methods. In particular we use the solution proposed in [44], where

$$\beta_k = \max\{0, -\gamma_k (\boldsymbol{s}^{(k-1)})^\top \boldsymbol{g}^{(k)} / \|\boldsymbol{s}^{(k-1)}\|_F^2)\}$$

and with $\gamma_k = 1.5$ (as the authors suggest). This method guarantees that this update has a smaller angle towards the optimal set than the standard negative subgradient direction. Figure 14 shows a comparison between the standard update method and the heavy ball method. A relative perturbation norm of 5.34×10^{-5} is reached at iteration 270 with standard negative subgradient and at iteration 75 with the heavy ball strategy, with a speed-up factor of about 3.6.

This last example shows that the proposed algorithms based on subgradient techniques are able to manage cases where other state-of-art methodologies fail to converge.

VIII. Conclusion and Discussion

This paper introduced a novel formulation of passivity enforcement schemes for linear macromodels in state-space

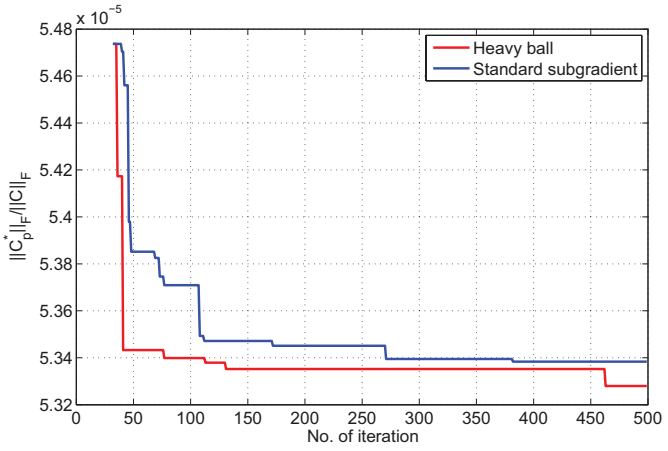

Fig. 14. Relative error perturbation norm of the standard update technique (blue solid line) and heavy ball method (red solid line).

form. The theoretical framework that we have discussed in the paper shows that the problem of finding the least perturbed macromodel under passivity constraints is *convex*. Therefore, when solved through a convex optimization scheme, such as the projected subgradient method, or the alternate subgradient method, there is a theoretical guarantee that the global optimal solution is found up to any prescribed accuracy within a finite number of steps or iterations. This fact is a distinctive advantage over most existing schemes in the literature. Some of these schemes are not convex at all, and are not guaranteed a priori to converge. Some other schemes perform some approximation such as linearization, projection, or similar, which lead to locally convex problems at each iteration. Yet, global convexity is lost during the approximation stage, thus loosing global optimality.

The nice theoretical features of the proposed schemes come with a cost. Although convergence to the optimum is guaranteed, this convergence may require many iterations. This is mainly due to the global structure of the problem, ultimately to the non-smooth behavior of the $\mathcal{H}_\infty$ norm with respect to the decision variables in the optimization process. This lack of regularity called for generalizations of standard descent methods involving subgradients and subdifferentials. The result is a possibly slower convergence rate with respect to regular Newton-like methods for smooth problems. However, a substantial speedup of subgradient-based methods seems to be achievable via simple modifications, such as the heavy-ball technique, as demonstrated in the examples.

APPENDIX A
CONVERGENCE RESULT FOR ALGORITHM (26)

Proposition 1: Assume that

(a) problem (25) admits an optimal solution $\boldsymbol{x}^*$;
(b) there exist a finite constant G such that $\|\boldsymbol{g}^{(k)}\| \leq G$ for all k;
(c) a constant R is known such that $\|\boldsymbol{x}^{(1)} - \boldsymbol{x}^*\| \leq R$.

Let h_k^* denote the best value achieved by algorithm (26) up to iteration k (note that this need not be the value of h at iteration k),

$$h_k^* = \min\{h(x^{(i)}) : i = 1, \ldots, k\}\,.$$

Then,

$$h_k^* - h^* \leq \frac{R^2 + G^2 \sum_{i=1}^k \alpha_i^2}{2\sum_{i=1}^k \alpha_i}. \tag{34}$$

In particular, if the α_k sequence is non-summable and diminishing (i.e., $\alpha_k \to 0$ as $k \to \infty$, and $\sum_{k=1}^\infty \alpha_k = \infty$; for example, $\alpha_k = \gamma/k$ for some $\gamma > 0$), then

$$\lim_{k\to\infty} h_k^* - h^* = 0.$$

Proof: A proof of the previous proposition can be found in [35], see also [34]. ■

It can be checked that, given the total number of iterations K, the upper bound in (34) is minimized by the choice of a constant stepsize $\alpha_i = \alpha = (R/G)/\sqrt{K}$, for which we obtain

$$h_k^* - h^* \leq \frac{RG}{\sqrt{K}}.$$

This means that if we stop the algorithm when exit accuracy $f_K^* - f^* \leq \frac{RG}{\sqrt{k}} \leq \epsilon$ is achieved, then we need at least $O(1/\epsilon^2)$ iterations.

APPENDIX B
CONVERGENCE RESULT FOR ALGORITHM (VI-A), (32)

Proposition 2: Assume that

(a) problem (29) admits an optimal solution $\boldsymbol{x}^*$;
(b) there exist a strictly feasible point $\boldsymbol{x}_{\text{sf}}$;
(c) there exist a finite constant G such that $\|\boldsymbol{g}^{(k)}\| \leq G$ for all k;
(d) a constant R is known such that $\|\boldsymbol{x}^{(1)} - \boldsymbol{x}^*\| \leq R$ and $\|\boldsymbol{x}^{(1)} - \boldsymbol{x}_{\text{sf}}\| \leq R$.

We denote with f_k^* the value of the best feasible point up to iteration k in algorithm (VI-A), (32) :

$$f_k^* = \min\{f(\boldsymbol{x}^{(i)}) : h(\boldsymbol{x}^{(i)}) \leq 1, i = 1, \ldots, k\}.$$

Then,

$$f_k^* - f^* \leq \epsilon_k, \quad h_k^* \leq 1 - \mu_k,$$

where

$$\epsilon_k + \mu_k \leq \frac{R^2 + G^2 \sum_{i=1}^k \alpha_i^2}{\sum_{i=1}^k \alpha_i}\,.$$

Proof: We proceed by contradiction. Assume that there exist $\epsilon > 0$ such that $f_k^* \geq f^* + 2\epsilon$ for all k, so that

$$f(\boldsymbol{x}^{(k)}) \geq f^* + 2\epsilon$$

for all k for which $\boldsymbol{x}^{(k)}$ is feasible. We will show that this assumption leads to a contradiction. Let first

$$\tilde{\boldsymbol{x}} = (1-\theta)\boldsymbol{x}^* + \theta \boldsymbol{x}_{\text{sf}}, \quad \theta \in (0,1).$$

By convexity of f we have

$$f(\tilde{\boldsymbol{x}}) \leq (1-\theta) f^* + \theta f(\boldsymbol{x}_{\text{sf}}).$$

By choosing $\theta = \min\{1, \epsilon/(f(\boldsymbol{x}_{\rm sf}) - f^*)\}$, this implies that

$$f(\tilde{\boldsymbol{x}}) \leq f^* + \epsilon,$$

that is, $\tilde{\boldsymbol{x}}$ is ϵ-suboptimal. Moreover

$$h(\tilde{\boldsymbol{x}}) \leq (1-\theta)h(\boldsymbol{x}^*) + \theta h(\boldsymbol{x}_{\rm sf}) \leq \theta h(\boldsymbol{x}_{\rm sf}),$$

therefore

$$h(\tilde{\boldsymbol{x}}) \leq 1 - \mu, \quad 1 - \mu = \theta h(\boldsymbol{x}_{\rm sf})\,.$$

Consider now an $i \in \{1, \ldots, k\}$ for which $\boldsymbol{x}^{(i)}$ is feasible. Then, $\boldsymbol{g}_f^{(i)} \in \partial f(\boldsymbol{x}^{(i)})$ and $f(\boldsymbol{x}^{(i)}) \geq f^* + 2\epsilon$ which, from (B), gives

$$f(\boldsymbol{x}^{(i)}) - f(\tilde{\boldsymbol{x}}) \geq \epsilon,$$

and therefore

$$\begin{aligned}
\|\boldsymbol{x}^{(i+1)} - \tilde{\boldsymbol{x}}\|^2 &= \\
&= \|\boldsymbol{x}^{(i)} - \tilde{\boldsymbol{x}}\|^2 - 2\alpha_i \boldsymbol{g}_f^{(i)\top}(\boldsymbol{x}^{(i)} - \tilde{\boldsymbol{x}}) + \alpha_i^2 \|g_f^{(i)}\|^2 \\
&\leq \|\boldsymbol{x}^{(i)} - \tilde{\boldsymbol{x}}\|^2 - 2\alpha_i (f(\boldsymbol{x}^{(i)}) - f(\tilde{\boldsymbol{x}})) + \alpha_i^2 \|\boldsymbol{g}_f^{(i)}\|^2 \\
&\leq \|\boldsymbol{x}^{(i)} - \tilde{\boldsymbol{x}}\|^2 - 2\alpha_i \epsilon + \alpha_i^2 \|\boldsymbol{g}_f^{(i)}\|^2.
\end{aligned}$$

Suppose instead that $i \in \{1, \ldots, k\}$ is such that $\boldsymbol{x}^{(i)}$ is infeasible. Then, $\boldsymbol{g}_h^{(i)} \in \partial h(\boldsymbol{x}^{(i)})$ with $h(\boldsymbol{x}^{(i)}) > 1$, and from (B) we have that

$$h(\boldsymbol{x}^{(i)}) - h(\tilde{\boldsymbol{x}}) \geq \mu.$$

Repeating the above derivation yields

$$\|\boldsymbol{x}^{(i+1)} - \tilde{\boldsymbol{x}}\|^2 \leq \|\boldsymbol{x}^{(i)} - \tilde{\boldsymbol{x}}\|^2 - 2\alpha_i \mu + \alpha_i^2 \|\boldsymbol{g}_h^{(i)}\|^2\,.$$

Therefore, for any iteration (feasible or infeasible), we have

$$\|\boldsymbol{x}^{(i+1)} - \tilde{\boldsymbol{x}}\|^2 \leq \|\boldsymbol{x}^{(i)} - \tilde{\boldsymbol{x}}\|^2 - \alpha_i \beta + \frac{1}{2}\alpha_i^2 G_i^2\,,$$

where

$$\beta = \epsilon + \mu, \quad G_i^2 = \|\boldsymbol{g}_f^i\|^2 + \|\boldsymbol{g}_h^i\|^2\,.$$

Applying this inequality recursively for $i = 1, \ldots, k$, we obtain

$$\|\boldsymbol{x}^{(k+1)} - \tilde{\boldsymbol{x}}\|^2 \leq \|\boldsymbol{x}^{(1)} - \tilde{\boldsymbol{x}}\|^2 - \beta \sum_{i=1}^{k} \alpha_i + \frac{1}{2}\sum_{i=1}^{k} \alpha_i^2 G_i^2.$$

From this it follows that

$$\beta \sum_{i=1}^{k} \alpha_i \leq R^2 + G^2 \sum_{i=1}^{k} \alpha_i^2,$$

hence

$$\beta \leq \frac{R^2 + G^2 \sum_{i=1}^{k} \alpha_i^2}{\sum_{i=1}^{k} \alpha_i}.$$

Now, if the stepsize sequence is diminishing and non-summable, then the right-hand side of the previous expression goes to zero as $k \to \infty$, thus leading to a contradiction. ■

APPENDIX C
ADAPTIVE STEPSIZES

We observe that if we use a constant step-size $\alpha = (R/G)/\sqrt{K}$ in algorithm (26), then we can estimate a lower bound of the optimal solution, in particular

$$h^* \geq h_k^* - \frac{RG}{\sqrt{k}}$$

where h^* is the optimal solution and h_k^* is the best value of $h(\boldsymbol{x})$ over the first k iterations. To perform this estimation we have to determine explicitly the constants R and G. In particular, we have that $R = \nu$, and an upper bound on the norm of the subgradient can be found from (23). Let φ be an element of $\partial h(\boldsymbol{x})$, then

$$\begin{aligned}
\|\varphi\|_2 &\leq \sum_{i=1}^{p} \|\boldsymbol{\Phi}(\mathrm{j}\bar{\omega}_i)\boldsymbol{V}_1^{(i)}\boldsymbol{Y}_i\boldsymbol{U}_1^{(i)H}\|_F \\
&\leq \sum_{i=1}^{p} \|\boldsymbol{\Phi}(\mathrm{j}\bar{\omega}_i)\|_F \|\boldsymbol{V}_1\boldsymbol{Y}_i\boldsymbol{U}_1^H\|_F \\
&= \sum_{i=1}^{p} \|\boldsymbol{\Phi}(\mathrm{j}\bar{\omega}_i)\|_F \sqrt{\mathrm{Tr}(\boldsymbol{V}_1\boldsymbol{Y}_i\boldsymbol{U}_1^H\boldsymbol{U}_1\boldsymbol{Y}_i\boldsymbol{V}_1^H)} \\
&\leq \|\boldsymbol{\Phi}\|_{\mathcal{H}_\infty} \sum_{i=1}^{p} \sqrt{\mathrm{Tr}\,\boldsymbol{Y}_i^2} \\
&\leq \|\boldsymbol{\Phi}\|_{\mathcal{H}_\infty},
\end{aligned}$$

where the last inequality follows from the fact that $(\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_p)$ belongs to the set $\mathcal{Y}_{(\ell_1,\ldots,\ell_p)}$ defined in (24).

Even if the constant step-size method is formally correct, it may require too many iterations to reach the requested accuracy. Therefore, we use in the implementation another strategy based on an adaptive selection of the step-size. Since for the subgradient method it holds that

$$h_{k-1}^* - h^* \leq \frac{R^2 + \sum_{i=1}^{k-1} \|\boldsymbol{g}^{(i)}\|^2 \alpha_i^2}{2\sum_{i=1}^{k-1} \alpha_i}$$

then at step k we can bound the optimality gap as

$$h_k^* - h^* \leq \delta_k, \quad \delta_k = \frac{R^2 + \|\boldsymbol{g}^{(k)}\|^2 \alpha_k^2 + \xi_{k-1}}{2(\alpha_k + \zeta_{k-1})}\,, \tag{35}$$

where

$$\xi_{k-1} = \sum_{i=1}^{k-1} \|\boldsymbol{g}^{(i)}\|^2 \alpha_i^2\,, \quad \zeta_{k-1} = \sum_{i=1}^{k-1} \alpha_i\,. \tag{36}$$

and with $\xi_0 = \zeta_0 = 0$. The uncertain δ_k in (35) is minimized when $\delta_k'(\alpha_k) = 0$, that is for

$$\alpha_k = \frac{-\|\boldsymbol{g}^{(k)}\|\zeta_{k-1} + \sqrt{\|\boldsymbol{g}^{(k)}\|^2 \zeta_{k-1}^2 + R^2 + \xi_{k-1}}}{\|\boldsymbol{g}^{(k)}\|} \tag{37}$$

REFERENCES

[1] T. Kailath, *Linear systems*, Prentice Hall, Englewood Cliffs, NJ, 1980.

[2] D. Luenberger, *Optimization by Vector Space Methods*, John Wiley and Sons, Inc., New York, 1969

[3] R. Pintelon, J. Schoukens, *Identification of Linear Systems: A Practical Guideline to Accurate Modeling*, Pergamon, 1991.

[4] B. Gustavsen, A. Semlyen, "Rational approximation of frequency responses by vector fitting", *IEEE Trans. Power Delivery*, Vol. 14, N. 3, pp. 1052–1061, July 1999.
[5] B. Gustavsen, A. Semlyen, "A robust approach for system identification in the frequency domain", *IEEE Trans. Power Delivery*, Vol. 19, N. 3, pp. 1167–1173, July 2004.
[6] D. Deschrijver, B. Haegeman, T. Dhaene, "Orthonormal Vector Fitting: A Robust Macromodeling Tool for Rational Approximation of Frequency Domain Responses", *IEEE Transactions on Advanced Packaging*, Vol. 30, No. 2, pp. 216–225, May 2007.
[7] M. R. Wohlers, *Lumped and Distributed Passive Networks*, Academic Press, 1969.
[8] P. Triverio, S. Grivet-Talocia, M. S. Nakhla, F. Canavero, R. Achar, "Stability, Causality, and Passivity in Electrical Interconnect Models" , *IEEE Trans. Adv. Packaging*, Vol. 30, No. 4, pp. 795–808, Nov. 2007.
[9] M. Celik, L. Pileggi, A. Obadasioglu, *IC Interconnect Analysis*, Kluwer, 2002.
[10] M. Nakhla and R. Achar, "Simulation of High-Speed Interconnects", *Proc. IEEE*, Vol. 89, No. 5, pp. 693–728, May 2001.
[11] S. Grivet-Talocia, "On driving non-passive macromodels to instability", *Int. J. Circuit Theory Appl.*, 2009, in press.
[12] S. Grivet-Talocia, A. Ubolli "On the Generation of Large Passive Macromodels for Complex Interconnect Structures", *IEEE Trans. Adv. Packaging*, vol. 29, No. 1, pp. 39–54, Feb. 2006
[13] S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan, *Linear matrix inequalities in system and control theory*, *SIAM studies in applied mathematics*, SIAM, Philadelphia, 1994.
[14] C. P. Coelho, J. Phillips, L. M. Silveira, "A Convex Programming Approach for Generating Guaranteed Passive Approximations to Tabulated Frequency-Data", *IEEE Trans. Computed-Aided Design of Integrated Circuits and Systems*, Vol. 23, No. 2, pp. 293–301, Feb. 2004.
[15] H. Chen, J. Fang, "Enforcing Bounded Realness of S parameter through trace parameterization", in *12th IEEE Topical Meeting on Electrical Performance of Electronic Packaging*, October 27–29, 2003, Princeton, NJ, pp. 291–294.
[16] B. Dumitrescu, "Parameterization of Positive-Real Transfer Functions With Fixed Poles", *IEEE Trans. CAS-I*, vol. 49, n. 4, pp. 523-526, Apr. 2002.
[17] Stephen Boyd, Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
[18] S. Boyd, V. Balakrishnan, P. Kabamba, "A bisection method for computing the H_∞ norm of a transfer matrix and related problems", *Math. Control Signals Systems*, Vol. 2, No. 3, pp. 207–219, Sept. 1989.
[19] S. Grivet-Talocia, "Passivity enforcement via perturbation of Hamiltonian matrices", *IEEE Trans. CAS-I*, Vol. 51, No. 9, pp. 1755-1769, Sept. 2004
[20] D. Saraswat, R. Achar and M. Nakhla, "A Fast Algorithm and Practical Considerations For Passive Macromodeling Of Measured/Simulated Data", *IEEE Transactions on Components, Packaging and Manufacturing Technology*, Vol. 27, No. 1, pp. 57–70, Feb. 2004.
[21] D. Saraswat, R. Achar and M. Nakhla, "Global Passivity Enforcement Algorithm for Macromodels of Interconnect Subnetworks Characterized by Tabulated Data", *IEEE Transactions on VLSI Systems*, Vol. 13, No. 7, pp. 819–832, July 2005.
[22] A. Semlyen, B. Gustavsen, "A half-size singularity test matrix for fast and realiable passivity assessment of rational models", *IEEE Trans. Power Delivery*, Vol. 24, No. 1, pp. 345–351, Jan. 2009.
[23] B. Gustavsen, A. Semlyen, "Fast passivity assessment for S-parameter rational models via a half-size test matrix", *IEEE Trans. Microwave Theory and Techniques*, Vol. 56, No. 12, pp. 2701–2708, Dec. 2008.
[24] S. Grivet-Talocia, "An adaptive sampling technique for passivity characterization and enforcement of large interconnect macromodels," *IEEE Trans. Adv. Packaging*, Vol. 30, No. 2, pp. 226–237, May 2007.
[25] B. Gustavsen, A. Semlyen, "Enforcing passivity for admittance matrices approximated by rational functions", *IEEE Trans. Power Systems*, Vol. 16, N. 1, pp. 97–104, Feb. 2001.
[26] B. Gustavsen, "Computer Code for Passivity Enforcement of Rational Macromodels by Residue Perturbation," *IEEE Trans. Adv. Packaging*, vol. 30, No. 2, pp. 209–215, May 2007.
[27] B. Gustavsen, "Fast passivity enforcement of Rational Macromodels by Perturbation of Residue Matrix Eigenvalues", *11th IEEE Workshop on Signal Propagation on Interconnects, May 13–16, 2007, Ruta di Camogli, Genova, Italy*, pp. 71–74.
[28] A Lamecki and M. Mrozowski, "Equivalent SPICE Circuits With Guaranteed Passivity From Nonpassive Models," IEEE Transactions on Microwave Theory And Techniques, Vol. 55, No. 3, pp. 526–532, Mar. 2007.
[29] S. Grivet-Talocia, A. Ubolli, "Passivity Enforcement With Relative Error Control" , *IEEE Trans. Microwave Theory and Techniques*, Vol. 55, No. 11, pp. 2374–2383, Nov. 2007.
[30] A. Ubolli, S. Grivet-Talocia, "Weighting Strategies for Passivity Enforcement Schemes", *16th IEEE Topical Meeting on Electrical Performance of Electronic Packaging, Atlanta, GA, 29–31 October, 2007*
[31] C. S. Saunders, Jie Hu, C. E. Christoffersen, M. B. Steer, "Inverse Singular Value Method for Enforcing Passivity in Reduced-Order Models of Distributed Structures for Transient and Steady-State Simulation, " *IEEE Trans. Microwave Theory and Techniques*, Vol. 59, No. 4, pp. 837–847, Apr. 2011.
[32] L. De Tommasi, M. de Magistris, D. Deschrijver, T. Dhaene, "An algorithm for direct identification of passive transfer matrices with positive real fractions via convex programming," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, Vol. 24, N. 4, pp. 375–386, 2011.
[33] S. Grivet-Talocia and A. Ubolli, "A comparative study of passivity enforcement schemes for linear lumped macromodels," *IEEE Trans. Advanced Packaging*, vol. 31, pp. 673–683, Nov 2008.
[34] S. Boyd, *Lecture notes and slides for EE364b, Convex Optimization II*, Stanford University.
[35] B.T. Polyak, *Introduction to Optimization*, Optimization Software, 1987.
[36] D.P. Bertsekas, A. Nedic, A.E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, 2003.
[37] Yu. Nesterov, *Introductory Lectures on Convex Optimization. A Basic Course*, Kluwer, 2004.
[38] H. Nikaidô, "On von Neumann's minimax theorem," *Pacific J. Math*, Vol. 4, No. 1, pp. 65–72, 1954.
[39] F.H. Clarke, Optimization and Nonsmooth Analysis, SIAM Classics in Applied Mathematics, 1990.
[40] P. Apkarian and D. Noll, "Nonsmooth H_∞ Synthesis," *IEEE Trans. Automatic Control*, Vol. 51, No. 1, pp. 71–86, Dec. 2006.
[41] J. H. Wilkinson, "The algebraic eigenvalue problem," Oxford University Press, 1965.
[42] C. F. Van Loan, "The ubiquitous Kronecker product," *J. Comput. Appl. Math.*, Vol. 123, 2000, pp. 85–100.
[43] M.L. Overton, "Large-scale Optimization of eigenvalues," *Siam J. Optimization*, Vol. 2, No. 1, pp. 88–120, 1992.
[44] P. Camerini, L. Fratta, and F. Maffioli, "On improving relaxation methods by modifying gradient techniques," *Math. Programming Study*, Vol. 3, pp. 26–34, 1975.

Giuseppe Calafiore G.C. Calafiore received the "Laurea" degree in Electrical Engineering from Politecnico di Torino in 1993, and the Ph.D in Information and System Theory from Politecnico di Torino, in 1997. Since 1998 he was with the faculty of Dipartimento di Automatica e Informatica, Politecnico di Torino, where he currently serves as an Associate Professor.

Dr. Calafiore held several visiting positions at international institutions: at Information Systems Laboratory, Stanford University, California, in 1995; at Ecole Nationale Supérieure de Techniques Avanceés (ENSTA), Paris, in 1998; and at the University of California at Berkeley, in 1999, 2003 and 2007, at the Institute of Pure and Applied Mathematics, University of California at Los Angeles, 2010, and at the John Von Neumann Institute, Vietnam National University in 2011 and 2012. He is an Associate Editor for the IEEE Transactions on Systems, Man, and Cybernetics (T-SMC), for the IEEE Transactions on Automation Science and Engineering (T-ASE), and for Journal Européen des Systèmes Automatisés. Dr. Calafiore is the author of more than 130 journal and conference proceedings papers, and of seven books. His research interests are in the fields of convex optimization, randomized algorithms, identification, and control of uncertain systems, with applications ranging from signals and systems, networks and finance to robust control, pattern recognition and robotics.

Alessandro Chinea received the Laurea Specialistica (M.Sc.) and Ph.D. degrees in electronic engineering from Politecnico di Torino, Italy in 2006 and 2010, respectively. In 2009 he spent a period at the Department of Information Technology (INTEC) of the Ghent University, Belgium, working under the supervision of the professors T. Dhaene and L. Knockaert. Since 2012 he works at the IdemWorks s.r.l. as a senior engineer. His research interests concern passive macromodeling of electrical interconnects for electromagnetic compatibility and signal/power integrity problems. Dr. Chinea received the Optime Award from the Unione Industriale di Torino and he was selected for the IBM EMEA Best Student Recognition Event 2006.

Stefano Grivet-Talocia (M'98–SM'07) received the Laurea and the Ph.D. degrees in electronic engineering from Politecnico di Torino, Italy. From 1994 to 1996, he was with the NASA/Goddard Space Flight Center, Greenbelt, MD, USA. Currently, he is an Associate Professor of Circuit Theory with Politecnico di Torino. His research interests are in passive macromodeling of lumped and distributed interconnect structures, modeling and simulation of fields, circuits, and their interaction, wavelets, time-frequency transforms, and their applications. He is author of more than 120 journal and conference papers. He is co-recipient of the 2007 Best Paper Award of the IEEE Trans. Advanced Packaging. He received the IBM Shared University Research (SUR) Award in 2007, 2008 and 2009. Dr. Grivet-Talocia served as Associate Editor for the IEEE TRANSACTIONS ON ELECTROMAGNETIC COMPATIBILITY from 1999 to 2001.