Model order reduction with Galerkin projection applied to nonlinear optimization with infeasible primal-dual interior point method

P.S.B. Nigro\textsuperscript{1,2} | E.T. Simões\textsuperscript{2} | P.M. Pimenta\textsuperscript{3} | J. Schröder\textsuperscript{1}

\textsuperscript{1}Institute of Mechanics, University of Duisburg-Essen, Essen, Germany
\textsuperscript{2}Virtual.PYXIS Optimization, São Caetano do Sul, Brazil
\textsuperscript{3}Department of Structural and Geotechnical Engineering, Polytechnic School at University of São Paulo, São Paulo, Brazil

Correspondence
P.S.B. Nigro, Institute of Mechanics, University of Duisburg-Essen, Universitätsstraße 15, 45141 Essen, Germany; or Virtual.PYXIS Optimization, Rua Tiradentes 160, 09541-220 São Caetano do Sul, Brazil.
Email: nigropsb@gmail.com

Funding information
Alexander von Humboldt-Stiftung, Grant/Award Number: Georg Forster Research Award; Conselho Nacional de Desenvolvimento Científico e Tecnológico, Grant/Award Number: 308142/2018-7; FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), Grant/Award Number: 2016/04460-4 and 2016/03528-4

Summary
It is not new that model order reduction (MOR) methods are employed in almost all fields of engineering to reduce the processing time of complex computational simulations. At the same time, interior point methods (IPMs), a method to deal with inequality constraint problems (which is little explored in engineering), can be applied in many fields such as plasticity theory, contact mechanics, micromechanics, and topology optimization. In this work, a MOR based in Galerkin projection is coupled with the infeasible primal-dual IPM. Such research concentrates on how to develop a Galerkin projection in one field with the interior point method; the combination of both methods, coupled with Schur complement, permits to solve this MOR similar to problems without constraints, leading to new approaches to adaptive strategies. Moreover, this research develops an analysis of error from the Galerkin projection related to the primal and dual variables. Finally, this work also suggests an adaptive strategy to alternate the Galerkin projection operator, between primal and dual variable, according to the error during the processing of a problem.

KEYWORDS
Galerkin projection, interior point methods, model order reduction, nonlinear optimization with constraint, Schur complement

1 | INTRODUCTION

Model order reduction (MOR) is a field considerably developed in computational simulations in engineering; such field has relevant applications in many other fields, like dynamics as explained in related works,\textsuperscript{1-4} multiscale analysis,\textsuperscript{5-8} and fluids analysis.\textsuperscript{9-11} Related to the development of the MOR itself, several MOR with offline analysis are described in other works,\textsuperscript{9,12-14} or in online analysis as developed in previous works\textsuperscript{1,2,10,15,16}; several of them achieving meaningful results related to the processing time. In this work, the objective is the development of a MOR from a multifield problem or constraint problem, but with an unusual way (at least to computational simulations in engineering) to deal with the inequalities constraints.

Exactly concentrated on the inequalities constraints, this research takes the decision to be part of the MOR, the interior point method (IPM). The IPM was originally developed by John von Neumann in the middle of the 20th century for linear programming. However, such method only came to have some applicability for nonlinear optimization at the end...
of the 20th century with the development of the infeasible primal-dual IPM, which is explained in other works\textsuperscript{17-24}; such variation ensures much more stability for the simulation. In some fields of computational simulations in engineering, such as topology optimization\textsuperscript{25,26} and mechanics,\textsuperscript{27,28} the infeasible primal-dual IPM already has applications; however, in the opinion of the authors, the method is still far from its true potential.

Because of that, the proposal of this work is a MOR developed from the Galerkin projection and the infeasible primal-dual IPM. The originality of this work is the application of the infeasible primal-dual IPM with special variables, with the Galerkin projection in a particular way; such configuration permits the application of a MOR with the processing similar to problems without constraints, leading to new approaches to adaptive strategies. Besides the originality of this work, the authors are concerned to demonstrate the potential and limitations of such combination of methods; and this way, permitting the applicability in fields like plasticity theory, contact mechanics, multiscale, and topology optimization to future research works, which can use this work as theoretical support and orientation. Based on all explained so far, and now using a more detailed approach, this work has the following objectives.

1. Develop a MOR to a multiconstraint problem with inequality. This MOR uses the Galerkin projection applied in only one field per time of the problem; then, because of the formulation of the infeasible primal-dual IPM, a Schur complement of this field is applied. The combination of these methods results in a MOR that permits to deal with an adaptive strategy that works equivalent in the processing as a problem without constraint.

2. Analysis of error from the primal and dual field. With the projection isolated in one field per time, the error of the respective Galerkin projection is analyzed, discriminating each parcel related to the primal or dual field. As expected, this work gives a special attention related to the parcel associated with the Karush-Kuhn-Tucker (KKT) condition.

3. The development of an adaptive strategy based on previous objectives. Now that the projected field is isolated and its error is analyzed, it is possible to propose an adaptive strategy based on the selection of the reduced field and the respective Schur complement, which combination minimizes the processing time. Two examples are tested with variations of the selection of the reduced field and the Schur complement; these examples not only have the objective to demonstrate the time reducing but also show the dependence between the error and the processing time.

After defining the objectives, the development of sections follow them. The second section introduces the idea behind the IPM; it explained the development of the infeasible primal-dual IPM, from the concept of the barrier function, passing through the classical infeasible primal-dual IPM, to the general approach with special variables, as described in the work of Svanberg.\textsuperscript{26} After the method is formulated, it is applied to the Schur complement related to the primal and dual variable. This section ends with an explanation about the IPM algorithm and the recommended adaptations of the method to its application.

The third section explains the first objective of this research; until now, the second section was only a preparation to understand this section. The third section described the general idea of the Galerkin projection. Then, it develops the formulation related to the Galerkin projection in one field per time and its respective Schur complement to the primal and dual field. Like the previous one, this section ends with an explanation about the IPM algorithm with MOR and the recommended adaptations to its application.

The fourth section describes the second objective of this research; the Galerkin error is analyzed with the formulations obtained from the second and third sections. In these Galerkin errors, from the primal and dual variable, the parcels of each error are analyzed individually. Sequentially, a Galerkin projection with both fields simultaneously is demonstrated, and a combination of both errors and the dependency between them are analyzed. Moreover, the last part of a sensibility analysis is developed.

The fifth and sixth sections, together, conclude the third objective. The fifth section describes the adaptive strategy in a complete online analysis, with many applications extracted from the works of Nigro et al.\textsuperscript{1,2} but, now, the MOR enters like the accelerator of the preconditioned conjugate gradient (PCG). This section starts from the definition of the selection of the snapshot and ends to the adaptive strategy to the selection of the Galerkin projection. Finally, these strategies are applied to the problems described in the sixth section. In the sixth section, the numerical examples are tested with different selections of Galerkin projections as well as different Schur complements. In this way, the processing time and Galerkin errors measurements provide relevant conclusions related to the advantages and disadvantages of the MOR develop in this work as well as possible applications for numerical simulations in engineering.

Before starting the next sections, it is important to explain that, to resume the formulation, this work uses some definitions of vector and matrices, ie,
\begin{align*}
  e_{\text{dim}} &\in \mathbb{R}^{\text{dim}} \quad \text{where} \quad e_{\text{dim},i} = 1 \quad \forall \quad i \in \mathbb{N} \\
o_{\text{dim}} &\in \mathbb{R}^{\text{dim}} \quad \text{where} \quad o_{\text{dim},i} = 0 \quad \forall \quad i \in \mathbb{N} \\
O_{\text{dim}} &\in \mathbb{R}^{\text{dim} \times \text{dim}} \quad \text{where} \quad O_{\text{dim},i,j} = 0 \quad \forall \quad i,j \in \mathbb{N} \\
O_{\text{dim}1 \times \text{dim}2} &\in \mathbb{R}^{\text{dim}1 \times \text{dim}2} \quad \text{where} \quad O_{\text{dim}1 \times \text{dim}2,i,j} = 0 \quad \forall \quad i,j \in \mathbb{N} \\
I_{\text{dim}} &\in \mathbb{R}^{\text{dim} \times \text{dim}} \quad \text{where} \quad I_{\text{dim}} = \text{diag}(e_{\text{dim}}).
\end{align*}

\section{Infeasible Primal-Dual IPM with Artificial Optimization Variables}

\subsection{Classical Infeasible Primal-Dual IPM}

The MOR described in this work starts with the definition of the optimization problem. It has chosen a nonlinear optimization with multiple inequality-constrained and primal variable, or design variable \(x\),

\begin{equation}
\begin{aligned}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad g \leq o_m \\
& \quad \forall \quad f(x) \in \mathbb{R}, \quad x \in \mathbb{R}^n \quad \text{and} \quad g^T = [g_1(x) \cdots g_m(x)]^T \in \mathbb{R}^m.
\end{aligned}
\tag{1}
\end{equation}

One approach to deal with inequality constraints is to apply a Lagrangian function and to use a Newton-Raphson method. In inequality constraint, the problem with this approach has an additional algorithm to detect active constraints; such algorithm changes constantly the dimension of the system and generates a considerable instability and loss of convergence. This additional algorithm has a very common application in engineering simulation, especially in a problem that involves contact mechanics and plasticity theory. Another approach, which is applied in this work, is the addition of slack variables to become from inequalities to equalities, ie,

\begin{equation}
\begin{aligned}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad g + s = o_m \\
& \quad \forall \quad s \in \mathbb{R}^m \\
& \quad s^T = [s_1 \cdots s_m]^T \in \mathbb{R}^m.
\end{aligned}
\tag{2}
\end{equation}

The slack variables named \(s\) must have positive values to ensure that they fulfill the remaining value to the equality. To deal with this property of the slack variable, it is applied a continuous, differentiable, and smooth approach to ensure a limitation in the domain. Such approach is called barrier function. The barrier function is a continuous function that has the behavior to reach a high value when close to the limit of the feasible region. Based on these modifications, the optimization problem becomes

\begin{equation}
\begin{aligned}
\text{minimize} & \quad f(x) - \mu \sum_{i=1}^{m} \ln(s_i) \\
\text{subject to} & \quad g + s = o_m \\
& \quad \forall \quad \mu \in \mathbb{R}_+.
\end{aligned}
\tag{3}
\end{equation}

The natural logarithm function described on (3) is the barrier function. It is possible to realize that, when the value inside the logarithm function reaches a value close to the defined limit (in this case, zero), it results in a high negative value, ie,

\begin{equation}
\lim_{s_i \to 0_+} \ln(s_i) = -\infty.
\tag{4}
\end{equation}

However, there is the barrier parameter \(\mu\) to ensure, during the simulation, that system (3) implies to system (2), ie,

\begin{equation}
\mu \ln(s_i) = 0 \quad \text{if} \quad \mu \to 0_+.
\tag{5}
\end{equation}

Another aspect to consider related to the slack variables and barrier function is that they belong to the infeasible version of the IPM, as applied in related works\cite{18-21} which permits the violation, eventually, of the inequality constraint. To solve
the optimization problem described in (3), it develops a Lagrangian function

$$l(x, s, \lambda) = f(x) - \mu \sum_{i=1}^{m} \ln(s_i) + \lambda^T [g + s] \quad \forall \ l \in \mathbb{R}, \ \lambda \in \mathbb{R}_+^m. \quad (6)$$

In the Lagrangian function, it is possible to realize that the slack variable can replace the value of the constraint function to satisfy the KKT condition, in which the value is imposed to be equal to zero. Such imposition is obtained, in steps, by the decrement of the barrier parameter

$$SA = \mu I_m \quad \forall \ S = \text{diag}(s) \in \mathbb{R}_+^{m \times m}, \ A = \text{diag}(\lambda) \in \mathbb{R}_+^{m \times m}. \quad (7)$$

The barrier parameter decreases during the simulation of the problem, ensuring the respect of the KKT condition at the end of the resolution.

### 2.2 Infeasible primal-dual IPM with artificial variables

In many cases, to support the application of slack variables and barrier function, the optimization problem as described in (1) has some instabilities and lack of generality. To handle with these limitations, the modification described in the work of Svanberg\textsuperscript{26} is implemented, ie,

$$\begin{align*}
\text{minimize} & \quad f(x) + a_0 z + y^T \left[ c + \frac{Dy}{2} \right] \\
\text{subject to} & \quad g - az - y \leq o_m \\
& \quad \alpha < x < \beta \\
& \quad \forall \ a_0 \in \mathbb{R}, \ z \in \mathbb{R}_+, \ a, c, d, y \in \mathbb{R}_+^m, \ D = \text{diag}(d) \in \mathbb{R}_+^{m \times m}, \ \alpha, \beta \in \mathbb{R}^n.
\end{align*} \quad (8)$$

The problem described in (8) is similar to (1), but it has advantages caused by artificial variables named “y” and “z”. The artificial variables intend to ensure stability and generality of the problem. The artificial variable “y” has the property to ensure at least one point that satisfies the KKT condition; the artificial variable “z” permits a generalization of the optimization problem, which includes minimizing the maximum (min-max) and least square problems. This generalization of the optimization problem is very useful in some engineering simulation fields, such as topology optimization. Another aspect to consider is that the problem described in (8) is applied with specific convex approximation functions described in the works of Svanberg\textsuperscript{25,26}; however, the formulation is so general that it can be applied with different locally convex functions.

Still analyzing the definition of the optimization problem, it is possible to notice that, to satisfy the KKT condition, the artificial variables must have values higher than zero. In addition, in this work and in most cases, the optimization must be made in a limited domain to respect parameter of design; to ensure that, the design variable “x” has values between “α” and “β”. To solve these domain limitation problems, there are added functions in the optimization problem (8) to penalize these variables when they are close to the limit of the domain, ie,

$$\begin{align*}
\text{minimize} & \quad f(x) + a_0 z + y^T \left[ c + \frac{Dy}{2} \right] - \mu \sum_{i=1}^{m} \ln(s_i) - \mu \ln(z) + \cdots \\
& \quad \cdots - \mu \sum_{i=1}^{m} \ln(y_i) - \mu \sum_{j=1}^{n} \ln(x_j - \alpha_j) - \mu \sum_{j=1}^{n} \ln(\beta_j - x_j) \\
\text{subject to} & \quad g - az - y + s = o_m.
\end{align*} \quad (9)$$

Problem (9) is an extension of problem (3); the idea is to ensure more stability and increase the range of application to the classical problem. Now, to start the interior method development, it is necessary to formulate a Lagrangian
function from (9), i.e.,

\[ l(x,y,z,s,\lambda) = f(x) + a_0 z + y^T \left[ c + \frac{1}{2} D y \right] - \mu \sum_{i=1}^{m} \ln (s_i) + \lambda^T [g - az - y + s] + \cdots \]

\[ \cdots - \mu \ln (z) - \mu \sum_{i=1}^{m} \ln (y_i) - \mu \sum_{j=1}^{n} \ln (x_j - a_j) - \mu \sum_{j=1}^{n} \ln (\beta_j - x_j) . \]

Now that the Lagrangian function is defined, it is necessary to find the critical point or the solution to the problem. The solution is developed by the Newton-Raphson method using primal and dual variables. Before starting the Newton-Raphson method, it is necessary to linearize the space. The linearization of the space results that each iteration of the optimization problem is developed at one point of the manifold \( \mathcal{M} \), where the manifold \( \mathcal{M} \) is defined by

\[ \mathcal{M} \subset \mathbb{R}^{n_u} \quad \forall \quad n_u \in \mathbb{N}_+. \]

In this manifold, a point is a tangent vector space \( T_u \mathcal{M} \), and, consequently,

\[ T_u \mathcal{M} \subset \mathcal{M}. \]

The geometrical representation is shown in Figure 1.

Figure (1) shows the vector space \( T_u \mathcal{M} \) amplified to a prism. It is important to notice that the representation by a prism is unusual (vector spaces are represented by planes in the manifold) but the additional dimension plays a useful role in the following sections of this work. Now, in the vector space \( T_u \mathcal{M} \), the fields are linearized, i.e.,

\[ x^{\text{new}} = x + \Delta x \quad \forall \quad \Delta x \in \mathbb{R}^n \]
\[ y^{\text{new}} = y + \Delta y \quad \forall \quad \Delta y \in \mathbb{R}^m \]
\[ z^{\text{new}} = z + \Delta z \quad \forall \quad \Delta z \in \mathbb{R} \]
\[ s^{\text{new}} = s + \Delta s \quad \forall \quad \Delta s \in \mathbb{R}^m \]
\[ \lambda^{\text{new}} = \lambda + \Delta \lambda \quad \forall \quad \Delta \lambda \in \mathbb{R}^m. \]

In the same vector space, developing Taylor’s series expansions from the Lagrangian function (10), the terms of second and higher order, as the first order null terms, are ignored. As a result, Taylor’s series expansions in the critical point, arranged in the matrix form, are

\[
\begin{bmatrix}
L_{xx} & O_{n \times m} & o_n & O_{n \times m} & J^T \\
O_{n \times m}^T & L_{yy} & o_m & O_{m} & -I_m \\
o_n^T & o_m & l_{zz} & o_m & -a^T \\
O_{n \times m}^T & O_m & A & S \\
J & -I_m & -a & I_m & O_m
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta y \\
\Delta z \\
\Delta s \\
\Delta \lambda
\end{bmatrix}
= 
\begin{bmatrix}
-l_x \\
-l_y \\
-l_z \\
-l_s \\
-l_{\lambda}
\end{bmatrix}.
\]
To simplify the formulation, it is defined as

\[
l_x = \frac{\partial l}{\partial x} \in \mathbb{R}^n, \quad l_y = \frac{\partial l}{\partial y} \in \mathbb{R}^m, \quad l_c = \frac{\partial l}{\partial c} \in \mathbb{R}, \quad l_a = \frac{\partial l}{\partial a} \in \mathbb{R}^m, \quad l_s = S \frac{\partial l}{\partial s} \in \mathbb{R}^m
\]

\[
L_{xx} = \frac{\partial^2 l}{\partial x \partial x} \in \mathbb{R}^{n \times n}, \quad L_{yy} = \frac{\partial^2 l}{\partial y \partial y} \in \mathbb{R}^{m \times m}, \quad l_{cz} = \frac{\partial^2 l}{\partial c \partial z} \in \mathbb{R}.
\]

(15)

It is possible to realize in (15) that the fourth row (top to bottom) of (14) is multiplied by the slack operator defined in (7); the objective is the simplification of the system further. The gradients from (15) are

\[
l_x = f_x + J^T \lambda - \mu [\mathbb{Z}]^{-1} e_n + \mu [H]^{-1} e_n \quad \forall \quad f_x = \frac{\partial f}{\partial x} \in \mathbb{R}^n, \quad J = \frac{\partial g}{\partial x} \in \mathbb{R}^{m \times n}
\]

(16)

\[
l_y = c + Dy - \lambda - \mu Y^{-1} e_m \quad \forall \quad Y = \text{diag}(y) \in \mathbb{R}^{m \times m}
\]

(17)

\[
l_c = a_0 - \lambda^T a - \frac{\mu}{z}
\]

(18)

\[
l_s = -\mu e_m + S \lambda \quad \text{or} \quad \bar{l}_s = -\mu e_m + A S
\]

(19)

\[
l_a = g - az - y + s.
\]

(20)

In (16), the matrices shown previously are

\[
\mathbb{Z} = \text{diag}(x - \alpha) \in \mathbb{R}^{n \times n}, \quad H = \text{diag}(\beta - x) \in \mathbb{R}^{n \times n}.
\]

(21)

Moreover, the Hessians (14) are

\[
\mathbf{L}_{xx} = F_{xx} + \theta^T \lambda + \mu [\mathbb{Z}]^{-2} + \mu [H]^{-2} \forall \quad F_{xx} = \frac{\partial^2 f}{\partial x \partial x} \in \mathbb{R}^{n \times n}, \quad \theta = \frac{\partial^2 g}{\partial x \partial x} \in \mathbb{R}^{m \times n \times n}
\]

(22)

\[
\mathbf{L}_{yy} = D + \mu [Y]^{-2}
\]

(23)

\[
l_{zz} = \frac{\mu}{\epsilon^2}.
\]

(24)

As noticed, the system of (14) is not symmetric, but with the manipulation of the last two rows, it is possible to reduce the dimension and to ensure symmetry. Isolating the increment of the slack variable in the fourth row, and replacing in the fifth row, the new system is

\[
\begin{bmatrix}
-L_{xx} & O_{n \times m} & -J^T & o_n \\
O_{n \times m}^T & -L_{yy} & I & o_m \\
-J & I & A^{-1} S & a \\
o_n^T & o_m^T & a^T & -l_{zz}
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta y \\
\Delta \lambda \\
\Delta z
\end{bmatrix}
= \begin{bmatrix}
l_x \\
l_y \\
l_a(i) \\
l_s
\end{bmatrix}
\Rightarrow \Gamma \Delta \tilde{u} = \varphi
\]

(25)

\[
\forall \quad \Gamma \in \mathbb{R}^{(n+2m+1)(n+2m+1)} \quad \text{and} \quad \Delta \tilde{u}, \varphi \in \mathbb{R}^{n+2m+1},
\]

where

\[
l_{a(i)} = \mu A^{-1} e_m + g - az - y \quad \forall \quad l_{a(i)} \in \mathbb{R}^m.
\]

(26)

In the second row of system (25), it is possible to obtain directly the increment of the special variable “y”, ie,

\[
\Delta y = [L_{yy}]^{-1} [\Delta \lambda - l_y].
\]

(27)

Moreover, the increment of the slack variable is obtained directly of the fourth row from (14), ie,

\[
\Delta s = A^{-1} \mu e_m - Se_m - A^{-1} S \Delta \lambda.
\]

(28)
Matrix (25) can be considered as the curvature of a saddle function in this case because all functions related to each domain are forced to be convex. An approach that allowed to treat saddle function as convex function, at least locally, is the Schur complement. The Schur complement reduces the dimension of the system to the dimension of one domain; the price to pay is the calculation of one inverse matrix. However, such price is mitigated if the inverted matrix is a diagonal matrix. Supported by this idea, this work develops the Schur complement in two domains: one related to the primal variable and another related to the dual variable.

2.3 | Schur complement related to the primal variable

This approach of the Schur complement is indicated when the number of constraints (dual variables) is greater than the number of design variables (primal variables). It is also indicated when the inverted Hessian of the design variable is expensive. To develop this Schur complement, the increment of the artificial variable “y” is necessary as described in (27), and replacing in the third row of the system (25), ie,

$$-J\Delta x + [L_{yy}^{-1} + A^{-1}S] \Delta \lambda = l_{(s)} + [L_{yy}]^{-1}l_y.$$  \hfill (29)

The new arrangement of the system (25) becomes

$$\begin{bmatrix} -L_{xx} & -J^T & o_n \\ -J & [L_{yy}^{-1} + A^{-1}S] & a \\ o_n^T & a^T & -l_{zz} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta z \end{bmatrix} = \begin{bmatrix} l_x \\ l_{(s)} + [L_{yy}]^{-1}l_y \\ l_z \end{bmatrix}. \hfill (30)$$

Based on this arrangement and using the Schur complement as detailed in Appendix A, results in the system

$$\begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \end{bmatrix} = \begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix}, \hfill (31)$$

where

$$\Pi_{11} = -L_{xx} - J^T[L_{yy}^{-1} + A^{-1}S]^{-1}J \quad \forall \quad \Pi_{11} \in \mathbb{R}^{n \times n}$$  \hfill (32)

$$\Pi_{22} = -l_{zz} - a^T[L_{yy}^{-1} + A^{-1}S]^{-1}a \quad \forall \quad \Pi_{22} \in \mathbb{R}$$  \hfill (33)

$$\pi_{12} = J^T[L_{yy}^{-1} + A^{-1}S]^{-1}a \quad \forall \quad \pi_{12} \in \mathbb{R}^{n \times 1}$$  \hfill (34)

$$\pi_{21} = a^T[L_{yy}^{-1} + A^{-1}S]^{-1}J \quad \forall \quad \pi_{21} \in \mathbb{R}^{1 \times n}$$  \hfill (35)

$$\pi_1 = l_x + J^T[L_{yy}^{-1} + A^{-1}S]^{-1} \{ l_{(s)} + [L_{yy}]^{-1}l_y \} \quad \forall \quad \pi_1 \in \mathbb{R}^{n \times 1}$$  \hfill (36)

$$\pi_2 = l_z - a^T[L_{yy}^{-1} + A^{-1}S]^{-1} \{ l_{(s)} + [L_{yy}]^{-1}l_y \} \quad \forall \quad \pi_2 \in \mathbb{R}. \hfill (37)$$

The new system results

$$\Delta x = \left[ \Pi_{11} - \Pi_{22}^{-1} \pi_{12} \pi_{21} \right]^{-1} \left\{ \pi_1 - \pi_{12} \Pi_{22}^{-1} \pi_2 \right\}$$  \hfill (38)

$$\Delta z = \Pi_{22}^{-1} \{ \pi_2 - \pi_{21} \Delta x \}. \hfill (39)$$

Furthermore, the increment of the dual variable is found from the second row of system (30), ie,

$$\Delta \lambda = [L_{yy}^{-1} + A^{-1}S]^{-1} \{ l_{(s)} + [L_{yy}]^{-1}l_y + J\Delta x - a\Delta z \}. \hfill (40)$$

2.4 | Schur complement related to the dual variable

This approach of the Schur complement is indicated when the number of design variables (primal variables) is greater than the number of constraints (dual variables). It is not indicated when the inverted Hessian of the design variable is
To develop this Schur complement, it is necessary to separate system (25) in a specific subsystem of the three first rows, ie,

\[
\begin{bmatrix}
    -L_{xx} & O_{nxm} & -J^T \\
    O_{nxm}^T & -L_{yy} & I \\
    -J & I & \Lambda^{-1}S
\end{bmatrix}
\begin{bmatrix}
    \Delta x \\
    \Delta y \\
    \Delta \lambda
\end{bmatrix}
+ \begin{bmatrix}
    o_n \\
    o_m \\
    a\Delta z
\end{bmatrix} = \begin{bmatrix}
    l_x \\
    l_y \\
    l_{\lambda(s)}
\end{bmatrix}. \tag{41}
\]

Based in this subsystem, using the Schur complement as detailed in Appendix A and regrouping with the last row of the artificial variable “z” on system (25) result in the new arrangement

\[
\begin{bmatrix}
    \prod & a \\
    a^T & -l_{zz}
\end{bmatrix}
\begin{bmatrix}
    \Delta \lambda \\
    \Delta z
\end{bmatrix} = \begin{bmatrix}
    \pi \\
    l_z
\end{bmatrix}. \tag{42}
\]

It is possible to find

\[
\prod = \Lambda^{-1}S + J[L_{xx}]^{-1}J^T + I_m[L_{yy}]^{-1}I_m \quad \forall \prod \in \mathbb{R}^{m \times m} \tag{43}
\]

\[
\pi = l_{\lambda(s)} - J[L_{xx}]^{-1}l_x + I_m[L_{yy}]^{-1}l_y \quad \forall \pi \in \mathbb{R}^m. \tag{44}
\]

The new system results

\[
\Delta \lambda = \left[\prod + \frac{1}{l_{zz}}aa^T\right]^{-1} \left\{\pi + a\frac{1}{l_{zz}}l_z\right\} \tag{45}
\]

\[
\Delta z = \frac{1}{l_{zz}} \left[a^T \Delta \lambda - l_z\right]. \tag{46}
\]

In addition, the increment of the primal variable is found from the first row of system (25), ie,

\[
\Delta x = -[L_{xx}]^{-1} \left(l_x + J^T\Delta \lambda\right). \tag{47}
\]

2.5 | The IPM algorithm and KKT conditions

Heretofore, the IPM algorithm related to the described problem is explained. As described in the last two subsections, the procedures to update all the variables using the Schur complement can be resumed in Algorithm 1A and Algorithm 1B from system (25).

**Algorithm 1**

(A) solve $\Delta u$ - Primal variable; (B) solve $\Delta u$ - Dual variable

(A)

input: $\Gamma, \varphi$
solve system: $\Delta x^{k+1},\Delta z^{k+1}$

Schur complement - Primal

calculate: $\Delta \lambda^{k+1},\Delta y^{k+1},\Delta s^{k+1}$

output: $\Delta u^{k+1} = \begin{bmatrix}
\Delta \lambda^{k+1} \\
\Delta y^{k+1} \\
\Delta s^{k+1}
\end{bmatrix}$

(B)

input: $\Gamma, \varphi$
solve system: $\Delta \lambda^{k+1},\Delta z^{k+1}$

Schur complement - Dual

calculate: $\Delta x^{k+1},\Delta y^{k+1},\Delta s^{k+1}$

output: $\Delta u^{k+1} = \begin{bmatrix}
\Delta \lambda^{k+1} \\
\Delta y^{k+1} \\
\Delta s^{k+1}
\end{bmatrix}$
When all variables are updated, the KKT condition must be checked. In this case, it is possible to check this condition by the residuals. In the analysis, the residual is calculated from the gradients (16), (17), (18), (19), and (20), ie,

\[
\begin{align*}
\mathbf{r}_x &= f_x + J^T \lambda - \mu [\mathbf{Z}^{-1} - \mathbf{H}^{-1}] \mathbf{e}_n \quad \forall \mathbf{r}_x \in \mathbb{R}^n \\
\mathbf{r}_y &= c + D_y - \lambda - \mu \mathbf{y}^{-1} \mathbf{e}_m \quad \forall \mathbf{r}_y \in \mathbb{R}^m \\
\mathbf{r}_c &= a_0 - \lambda^T \mathbf{a} - \frac{\mu}{\mathbf{z}} \quad \forall \mathbf{r}_c \in \mathbb{R} \\
\mathbf{r}_\lambda &= \mathbf{g} - \mu \mathbf{e}_m + \lambda \mathbf{S} \quad \forall \mathbf{r}_\lambda \in \mathbb{R}^m \\
\mathbf{r}_s &= -\mu \mathbf{e}_m + \lambda \mathbf{S} \quad \forall \mathbf{r}_s \in \mathbb{R}^m.
\end{align*}
\] (48)

When the residual is defined, the IPM algorithm can be developed. In the algorithm applied in this work, the barrier parameter decreases every time that the norm of residuals described in (48) reaches values below the value of the current barrier parameter. Lastly, when the barrier parameter reaches the minimum specified value, it is considered that the KKT condition is satisfied, resulting in the solution. For a better understanding, a scheme of the IPM resolution is shown in Flowchart 1 in the next section.

3 | MOR IN THE INTERIOR POINT METHOD WITH GALERKIN PROJECTION IN ONE FIELD

3.1 | Global Galerkin projection

In an optimization with constraint, it is necessary to handle with at least two fields, represented by the design variable (primal variable) and the Lagrange multiplier (dual variable). In a common MOR, all the fields are reduced, as described in the works of Xiao et al\(^9\) and Zahr and Farhat\(^{13}\). Another approach, as described in the works of Kaulmann et al\(^{5,6}\), is a localized MOR, but an offline analysis previously is necessary. In some cases, the complete online MOR with several fields can reach a similar processing time when compared to the full model.

To start an online MOR from infeasible primal-dual IPM, the proposal is to work with the reduction, through the Galerkin projection, in only one field per time. To develop that, the application of the Schur complement to achieve this goal is shown. In the type of problem described in the previous section, this is possible with low numerical effort. During the processing of the problem, the update and the error control of the reduced model can be treated as an optimization without constraint, as described in the work of Nigro et al\(^{1,2}\).

The application of the Schur complement, in the point of view of the size of the domain, can be considered as a reduction of dimension. The Schur complement reduces the dimension of the system to the dimension of one field (usually the smaller) during the solution. In a model with two fields, when dimensions are in the same order of magnitude, the selection of the Schur complement is not so effective. Now, presenting the infeasible primal-dual IPM and the Schur complement, aims in the MOR are necessary. The objective of a MOR, based in the Galerkin projection, is to find a vector space \( \mathbf{V}_u \) such that

\[
\mathbf{V}_u \subset \mathbf{T}_u \mathbf{M}.
\] (49)

The vector space \( \mathbf{V}_u \) is the reduced space, and it can be represented by the plane inside the prism, ie, the vector space \( \mathbf{T}_u \mathbf{M} \). To develop this MOR based in the Galerkin projection, it is necessary to start with the matrix and vectors from (25). In the Galerkin projection, the increment of the domain is decomposed in two additive terms, ie, the range and null space,

\[
\begin{align*}
\Delta \hat{\mathbf{u}} &= \Delta \mathbf{u}_\psi + \Delta \mathbf{u}_\Omega \quad \forall \Delta \hat{\mathbf{u}}, \Delta \mathbf{u}_\psi, \Delta \mathbf{u}_\Omega \in \mathbf{T}_u \mathbf{M} \\
\Delta \mathbf{u}_\psi &= \mathbf{F} \Delta \mathbf{u}_r \quad \forall \Delta \mathbf{u}_r \in \mathbf{V}_u \quad \text{and} \quad \mathbf{F} : \mathbf{V}_u \rightarrow \mathbf{T}_u \mathbf{M} \\
\Delta \mathbf{u}_\Omega &= \mathbf{K} \Delta \mathbf{u}_e \quad \forall \Delta \mathbf{u}_e \in \mathbf{V}_u^\perp \quad \text{and} \quad \mathbf{K} : \mathbf{V}_u^\perp \rightarrow \mathbf{T}_u \mathbf{M} \\
\mathbf{T}_u \mathbf{M} &= \mathbf{V}_u \oplus \mathbf{V}_u^\perp.
\end{align*}
\] (50)

Vector \( \Delta \mathbf{u}_\psi \) represents the range space operator; it always belongs to the space vector \( \mathbf{V}_u \). Vector \( \Delta \mathbf{u}_\Omega \) represents the null space; it represents the error of \( \Delta \mathbf{u}_\psi \) related to \( \Delta \hat{\mathbf{u}} \), and it is always orthogonal to the space vector \( \mathbf{V}_u \). Figure 2 shows the reduction of dimension, and the reason to the point of the manifold \( \mathbf{M} \) be represented by a prism, as declared in the previous section.
Back to the explanation of (50), matrix $\Phi$ represents the global Galerkin operator, and matrix $K$ represents the global null space operator. Each global operator belongs to complementary spaces

$$\Phi^T K = O \quad \forall \quad \Phi^T \Phi = I \quad , \quad K^T K = I.$$  

(51)

To calculate the error associated to the projection, the system described in (25) is operated by $\Phi^T$ in both sides, and considering the decomposition of (50),

$$\Phi^T \Gamma \Phi \Delta u_r + \Phi^T K \Delta u_e = \Phi^T \phi.$$  

(52)

The first term on the left side in Equation (52) represents the part of the gradient $\phi$ that belongs to the range space and provides the result of the MOR; the second term on the left side in Equation (52) represents the part of the gradient $\phi$ that belongs to the null space, and, consequently, does not belong to the MOR; this last term contains the error of the reduced space related to the full space.

Related to (25), the global Galerkin operator can obtain the projection until three domains. However, it is possible to notice that the simultaneous projections generate cross terms; the cross terms difficult a more accurate analysis of the error related to each domain. Because of that, one of the proposals of this work is the application of the projection in one domain per time (in addition to the error being measured more precisely) and there is an advantage related to the operation with the Schur complement that is demonstrated in the next two sections.

### 3.2 | MOR of the primal field

In this section, the main hypothesis (a reduction of one field) considers the decomposition of the design variable (primal variable), i.e,

$$\Delta x = \Delta x_{\Psi r} + \Delta x_{\Omega e} \quad \forall \quad \Delta x_{\Psi r}, \Delta x_{\Omega e} \in \mathbb{R}^n \quad \Delta x = \Psi x \Delta x_r + \Omega x \Delta x_e \quad \forall \quad \Psi x \in \mathbb{R}^{n \times w_n}, \quad \Omega x \in \mathbb{R}^{n \times (n-w_n)}, \Delta x_r \in \mathbb{R}^w, \Delta x_e \in \mathbb{R}^{n-w_n},$$  

(53)

(54)

where $\Psi x$ is the local range space operator related to the primal space, and $\Omega x$ is the local null space operator. The main variables are arranged in the following decomposition:

$$\begin{bmatrix} \Psi x \Delta x_r + \Omega x \Delta x_e \\ \Delta y \\ \Delta \lambda \\ \Delta z \end{bmatrix} = \begin{bmatrix} \Psi x & O_{nxm} & O_{nxm} & o_n \\ O_{mxx} & I_m & o_{nxm} & o_m \\ O_{mxn} & I_m & o_{nxm} & o_m \\ o_{w_n} & o_m & o_m & 1 \end{bmatrix} \begin{bmatrix} \Delta x_r \\ \Delta y \\ \Delta \lambda \\ \Delta z \end{bmatrix} + \begin{bmatrix} \Omega x \\ O_{mxx-nw_n} \\ O_{mxn-nw_n} \\ o_{w_n} o_m \end{bmatrix} \begin{bmatrix} \Delta x_e \end{bmatrix} \iff \Delta \tilde{u} = \Phi x \Delta u_{x_r} + K x \Delta u_{x_e}.$$  

(55)

When this global Galerkin operator and the global null space operator are replaced in (52), the parcel of the error is represented by the term

$$\Phi x^T \Gamma K x \Delta u_{x_e} = \begin{bmatrix} -\Psi x^T L_{xx} \Delta x_{\Omega e} \\ o_m \\ -\Delta x_{\Omega e} \\ 0 \end{bmatrix} \in \mathbb{R}^{(w_n+2m+1)}, \quad \Delta x_{\Omega e} = \Omega x \Delta x_e \in \mathbb{R}^n.$$

(56)
Term (56) is useful to measure the error related to the full model, including the application in an adaptive strategy to reduce the error of the MOR. Finally, the terms related to the projection are

\[
\Phi_x^T \varphi = \begin{bmatrix} \Psi_x^T I \kappa \\ I_y \\ I_{d(s)} \\ l_z \end{bmatrix} \in \mathbb{R}^{(w_n+2m+1)}
\]

(57)

\[
\Phi_x^T \Gamma \Phi_x = \begin{bmatrix} -\Psi_x^T L_{x \kappa} \Psi_x O_{w_n \times m} -\Psi_x^T J^T o_{w_n} \\
O_{w_n \times m} -L_{yy} & I_m & o_m \\
-J \Psi_x & I_m & L^{-1}S & a \\
o_{w_n} & o_m & a^T & -I_{zz} \end{bmatrix} \in \mathbb{R}^{(w_n+2m+1) \times (w_n+2m+1)}.
\]

(58)

Applying the Schur complement as developing in (31) and (37),

\[
\begin{bmatrix} (\Pi_{11})_r & (\pi_{12})_r \\ (\pi_{21})_r & \Pi_{22} \end{bmatrix} \begin{bmatrix} \Delta x_r \\ \Delta z \end{bmatrix} = \begin{bmatrix} (\pi_1)_r \\ \pi_2 \end{bmatrix},
\]

(59)

where

\[
(\Pi_{11})_r = -\Psi_x^T L_{x \kappa} \Psi_x -\Psi_x^T J^T [L_{yy}^{-1} + L^{-1}S]^{-1} J \Psi_x \Leftrightarrow
\]

\[
(\Pi_{11})_r = \Psi_x^T \Pi_{11} \Psi_x \quad \forall \quad (\Pi_{11})_r \in \mathbb{R}^{w_n \times w_n}
\]

(60)

\[
(\pi_{12})_r = \Psi_x^T \pi_{12} \quad \forall \quad (\pi_{12})_r \in \mathbb{R}^{w_n \times 1}
\]

(61)

\[
(\pi_{21})_r = a^T [L_{yy}^{-1} + L^{-1}S]^{-1} J \Psi_x \Leftrightarrow
\]

\[
(\pi_{21})_r = \pi_{21} \Psi_x \quad \forall \quad (\pi_{21})_r \in \mathbb{R}^{1 \times w_n}
\]

(62)

\[
(\pi_1)_r = \Psi_x^T I \kappa + \Psi_x^T J^T [L_{yy}^{-1} + L^{-1}S]^{-1} \left\{ J_{d(s)} + [L_{yy}]^{-1} l_y \right\} \Leftrightarrow
\]

\[
(\pi_1)_r = \Psi_x^T \pi_1 \quad \forall \quad (\pi_1)_r \in \mathbb{R}^{w_n \times 1}
\]

(63)

As demonstrated from (60) to (63), with a simple distributive property, it is possible to obtain a direct relationship between the reduced and the full Schur complement. These direct relations cannot be made in a model with two or more fields to be reduced. The choice of one reduced field, associated with the Schur complement, permits a MOR based on Galerkin projection in a similar way of processing of an optimization problem without constraints. The benefits of such an approach are the reduction of matrix operations and the application of several existing adaptive strategy already developed. Back to the formulation, the reduced system can be developed as (38) and (39), ie,

\[
\Delta x_r = \left[ (\Pi_{11})_r - \Pi_{22}^{-1} (\pi_{12})_r (\pi_{21})_r \right]^{-1} \left\{ (\pi_1)_r - (\pi_{12})_r \Pi_{22}^{-1} \pi_2 \right\}
\]

(64)

\[
\Delta x_{\psi_s} = \Psi_s \Delta x_r, \quad \Delta x_{\psi_s} \in \mathbb{R}^n
\]

(65)

\[
\Delta z_{\psi_s} = \Pi_{22}^{-1} \left\{ \pi_2 - (\pi_{21})_r \Delta x_r \right\}, \quad \Delta z_{\psi_s} \in \mathbb{R}
\]

(66)

where \( \Delta z_{\psi_s} \) represents the approximate value of increment of the special variable “\( z \)” from the range space of the primal variable. In the same way, an approximate value of increment of the dual variable from the range space of the primal variable is obtained as (40), ie,

\[
\Delta \lambda_{\psi_s} = [L_{yy}^{-1} + L^{-1}S]^{-1} \left\{ J_{d(s)} + [L_{yy}]^{-1} l_y + J \Delta x_{\psi_s} - a \Delta z_{\psi_s} \right\}, \quad \Delta \lambda_{\psi_s} \in \mathbb{R}^m.
\]

(67)

At this point, it is clear that

\[
\Delta z_{\psi_s} \neq \Delta z, \quad \Delta x_{\psi_s} \neq \Delta x \quad \text{and} \quad \Delta \lambda_{\psi_s} \neq \Delta \lambda.
\]

(68)
These variables are approximations that depend on the quality of the representation from the Galerkin projection of the primal variable space $\Psi_x$. A more detailed explanation related to the error associated with both types of variables is explained in the next section. The next subsection is related to the dual approach of the Galerkin projection.

### 3.3 MOR of the dual field

In this section, the main hypothesis (a reduction of one field) considers the decomposition of the Lagrangian multiplier (dual variable), ie,

$$\Delta \lambda = \Delta \lambda_{\Psi} + \Delta \lambda_{\Omega} \quad , \quad \Delta \lambda_{\Psi}, \Delta \lambda_{\Omega} \in \mathbb{R}^m \tag{69}$$

$$\Delta \lambda = \Psi_{\lambda} \Delta \lambda_{\Psi} + \Omega_{\lambda} \Delta \lambda_{\Omega} \quad , \quad \Psi_{\lambda} \in \mathbb{R}^{m \times w}, \quad \Omega_{\lambda} \in \mathbb{R}^{m \times m - w}, \quad \Delta \lambda_{\Psi} \in \mathbb{R}^w, \quad \Delta \lambda_{\Omega} \in \mathbb{R}^{m - w}, \tag{70}$$

where $\Psi_{\lambda}$ is the local range space operator related to the dual space, and $\Omega_{\lambda}$ is the local null space operator. The main variables are arranged in the following decomposition:

$$\Delta \mathbf{u} = \Phi_{\lambda} \Delta \mathbf{u}_{\Psi} + K_{\lambda} \Delta \mathbf{u}_{\Omega}. \tag{71}$$

When these global Galerkin operator and the global null space operator are replaced in (52), the parcel of the error is represented by the term

$$\Phi_{\lambda}^T \Gamma K_{\lambda} \Delta \mathbf{u}_{\Omega} = \begin{bmatrix} -J^T \Delta \lambda_{\Omega} \\ \Psi_{\lambda}^T \Lambda^{-1} \delta \Delta \lambda_{\Omega} \\ a^T \Delta \lambda_{\Omega} \end{bmatrix} \in \mathbb{R}^{(n+m+w,+1)} \quad , \quad \Delta \lambda_{\Omega} = \Omega_{\lambda} \Delta \lambda_{\Omega}. \tag{72}$$

As in the previous subsection, the term of (72) is extremely useful to measure the error related to the full model, including the application in an adaptive strategy to reduce the error of the MOR. Finally, the terms related to the projection are

$$\Phi_{\lambda}^T \varphi = \begin{bmatrix} \mathbf{l}_{\lambda} \\ \mathbf{l}_{\Psi}^T J_{\lambda(3)} \\ \mathbf{l}_{\Omega} \end{bmatrix} \in \mathbb{R}^{(n+m+w,+1)} \tag{73}$$

$$\Phi_{\lambda}^T \Gamma \Phi_{\lambda} = \begin{bmatrix} -L_{\lambda} & O_{\lambda} & -J^T \Psi_{\lambda} & \mathbf{o}_{\lambda} \\ O_{\lambda} & -L_{\lambda} & \Psi_{\lambda} & \mathbf{o}_{\lambda} \\ -J^T \Psi_{\lambda} & \Psi_{\lambda} & \Lambda^{-1} \delta \Psi_{\lambda} & \mathbf{a} \\ \mathbf{o}_{\lambda} & \mathbf{o}_{\lambda} & \mathbf{a} & -l_{\Omega} \end{bmatrix} \in \mathbb{R}^{(n+m+w,+1) \times (n+m+w,+1)}. \tag{74}$$

Applying the Schur complement as developing in (42) to (44),

$$\begin{bmatrix} \Pi_r \Psi_{\lambda}^T \mathbf{a} \\ a^T \Psi_{\lambda} \end{bmatrix} \Delta \lambda_r + \mathbf{l}_{\Omega} = \begin{bmatrix} (\pi)_r \\ l_{\Omega} \end{bmatrix}, \tag{75}$$

where

$$\begin{bmatrix} \Pi_r \Psi_{\lambda}^T \mathbf{a} \\ a^T \Psi_{\lambda} \end{bmatrix} = \begin{bmatrix} \Psi_{\lambda}^T \Lambda^{-1} \delta \Psi_{\lambda} + \Psi_{\lambda}^T J[L_{\lambda}]^{-1} J^T \Psi_{\lambda} + \Psi_{\lambda}^T [L_{\lambda}]^{-1} \Psi_{\lambda} \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} \Pi_r \Psi_{\lambda} \Pi_{\lambda} \end{bmatrix} \forall \begin{bmatrix} \Pi_r \end{bmatrix} \in \mathbb{R}^{w \times w} \tag{76}$$

$$\begin{bmatrix} (\pi)_r \end{bmatrix} = \begin{bmatrix} \Psi_{\lambda}^T L_{\lambda(3)} - \Psi_{\lambda}^T J[L_{\lambda}]^{-1} l_{\lambda} + \Psi_{\lambda}^T [L_{\lambda}]^{-1} l_{\lambda} \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} (\pi)_r \end{bmatrix} = \begin{bmatrix} \Psi_{\lambda}^T \mathbf{a} \end{bmatrix} \forall \begin{bmatrix} (\pi)_r \end{bmatrix} \in \mathbb{R}^{w}. \tag{77}$$
Again, with a simple distributive property from (76) and (77), it is possible to obtain a direct relationship between the reduced and full Schur complement. As in the previous subsection, these direct relations cannot be made in a model with two or more fields to be reduced. Once again, the choice of one reduced field leads to a similar way of processing of an optimization problem without constraints. In the specific case of this work, the explicit convex approximation functions generate diagonal Hessians; this matrix format reduces, even more, the number of operations to reach the inverse matrices. Returning to the formulation development, the reduced system can be developed as (45) and (46), ie,

\[
\Delta \lambda_r = \left( \prod_{i=2}^n \right)_{r} \psi_d^{T} a a^{T} \psi_d \right)^{\frac{1}{2}} \left( \pi_r + \psi_d^{T} a a^{T} \psi_d \right)^{-1} \left( \pi_r + \psi_d^{T} a a^{T} \psi_d \right) \left( \prod_{i=2}^n \right)_{r}
\]

\[
\Delta \lambda_{\psi_d} = \psi_d \Delta \lambda_r , \Delta \lambda_{\psi_d} \in \mathbb{R}^m
\]

\[
\Delta z_{\psi_d} = \frac{1}{l_{zz}} \left( a^{T} \psi_d \Delta \lambda_r - l_z \right) , \Delta z_{\psi_d} \in \mathbb{R}.
\]

where \( \Delta z_d \) represents the approximate value of increment of the special variable “z” from the range space of the dual variable. In the same way, an approximate value of increment of the primal variable from the range space of the dual variable is obtained as (47), ie,

\[
\Delta x_{\psi_d} = - \left( L_{xx} \right)^{-1} \left( l_x + J^{T} \Delta \lambda_{\psi_d} \right) , \quad x_{\psi_d} \in \mathbb{R}^n.
\]

As in the previous subsection, at this point, it is clear that

\[
\Delta z_{\psi_d} \neq \Delta z , \Delta \lambda_{\psi_d} \neq \Delta \lambda \quad \text{and} \quad \Delta x_{\psi_d} \neq \Delta x.
\]

These entire variables are approximations that depend on the quality of the representation from the Galerkin projection of the dual variable space \( \Psi_d \). A more detailed explanation related to the error associated with both types of variables is described in the next section.

### 3.4 The IPM algorithm with MOR and the KKT condition

Until now, as described in the first subsection of this work, the MOR occurs in the linearization of the manifold, the vector space \( T_u \mathcal{M} \). Such approach is common to computational mechanics; optimization without constraint (using the classical Newton-Raphson) or optimization with equality constraints can apply snapshots from previous linearized spaces without major concerns. However, in optimization methods with inequality constraints, the KKT condition is a factor that difficult the use of previous snapshots to develop a Galerkin projection, with low error, to the next step.

The source of such problem is the barrier parameter; the level of variation (that is an empirical choice) can change abruptly the values of the gradient and Hessian. In the next section, more precisely in the subsection related to the sensibility analysis, it is demonstrated that even small variations of the barrier parameter can result in meaningful variations to the primal or dual variables. This work adopts the increment of the primal and dual variables as snapshots; because of that, to avoid the increase of the error in the projection, the option adopted is to restart the Galerkin projection for each new value of barrier parameter. In Flowchart 1, the IPM algorithm with MOR is described.

The additional variables in the flowchart are

\[
\mathbf{r} = \begin{bmatrix} r_x^T & r_y^T & r_z^T & r_i^T \end{bmatrix} \in \mathbb{R}^{n+3m+1}
\]

\[
\mathbf{u} = \begin{bmatrix} x^T & y^T & z^T & \lambda^T \end{bmatrix} \in \mathbb{R}^{n+3m+1}.
\]

The difference between the full and MOR algorithms resides in obtaining the new increment through the development of the Newton Method. The application of the MOR in Flowchart 1 justifies the snapshots acquired from the increments. It is important to highlight that the option of discard the snapshots with a new barrier parameter is based in some possibilities of high variations demonstrated in the sensibility analysis from Section 4; it is not possible to ensure that this option always result in a lower number of corrections to the Galerkin projection.
4 | ERROR RELATED TO THE GALERKIN PROJECTION

4.1 Error related to the primal variable

In the previous section, the parcel of the error related to the projection from the primal variable is obtained from the vector described in (56); based on this vector, it is possible to associate the error of the Galerkin projection related to the primal variable. To achieve that, the idea is to isolate the vector from the null space of the primal variable. This approach is possible, exposing all the equations from (52) with null space of the primal variable different of a null vector, i.e.,

\[
-\Psi_x^T L_{xx} \Delta x_{\Psi_x} - \Psi_x^T J^T \Delta \lambda - \Psi_x^T L_{xx} \Delta x_{\Omega_x} = \Psi_x^T I_x
\]

\[
-J \Delta x_{\Psi_x} + I_m \Delta y + \Lambda^{-1} S \Delta \lambda - J \Delta x_{\Omega_x} = \nabla \lambda(x) I .
\]

Isolating the projection value in (84),

\[
\Delta x_{\Psi_x} = -\Psi_x \left[ \Psi_x^T L_{xx} \Psi_x \right]^{-1} \Psi_x^T \left( I_x + J^T \Delta \lambda + L_{xx} \Delta x_{\Omega_x} \right).
\]

Now, replacing the projection value (86) in (85) and isolating by field,

\[
A_{11(x)} \Delta x_{\Omega_x} + A_{12(x)} \Delta y + A_{13(x)} \Delta \lambda + A_{14(x)} \Delta z = a_{1(x)}
\]

\forall A_{11(x)} \in \mathbb{R}^{m \times n} , A_{12(x)} , A_{13(x)} \in \mathbb{R}^{m \times m} , A_{14(x)} \in \mathbb{R}^{m \times 1} \text{ and } a_{1(x)} \in \mathbb{R}^m .
\]

Related to the error analysis, the operators \( A_{11(x)} , A_{12(x)} , A_{13(x)} , A_{14(x)} \), and \( a_{1(x)} \) are less relevant; the description of such operators can be seen in Appendix B. The relevant operator is related to the vector of the null space, and in this case, it is the operator \( A_{11(x)} \), where

\[
A_{11(x)} = -J \left( I_n - \Psi_x \left[ \Psi_x^T L_{xx} \Psi_x \right]^{-1} \Psi_x^T L_{xx} \right).
\]

In the null space perspective, it is possible to see that all the other terms must belong to

\[
a_1 - A_{12} \Delta y - A_{13} \Delta \lambda - A_{14} \Delta z \subset A_{11} .
\]
In addition, if the Galerkin projection converges to the full space,

$$\lim_{n\to\infty} A_{31}(x) (\Psi_n) = O_{n \times n}.$$  \hfill (90)

This occurs because of the parcel of (88), ie,

$$\lim_{n\to\infty} \left\{ I_n - \Psi_n [\Psi_n^T L_{\text{ex}} \Psi_n]^{-1} \Psi_n^T L_{\text{ex}} \right\} = O_{n \times n}.$$  \hfill (91)

Based on this parcel, it is possible to realize that the error is dependent on the Hessian of the Lagrangian function related to the primal variable, and the curvature of this function in the primal field must be convex to ensure the convergence of the error (the matrix must be invertible). This Hessian, described in (22), is partially dependent on the dual variables, and consequently has some levels of dependency to fulfill the KKT condition.

### 4.2 | Error related to the dual variable

Furthermore, in the previous section, the parcel of the error related to the projection from the dual variable is obtained from the vector described in (72); based on this vector, it is possible to associate the error of the Galerkin projection related to the dual variable. As in the previous subsection, the idea is to isolate the vector from the null space of the dual variable. This approach is possible, exposing all the equations from (52) with null space of the dual variable different of a null vector, ie,

\[
-L_{xx} \Delta x - J^T \Delta \lambda_{\mathcal{D}_1} - J^T \Delta \lambda_{\Omega_2} = l_x \tag{92}
\]

\[
-L_{yy} \Delta y + \Delta \lambda_{\mathcal{D}_1} + \Delta \lambda_{\Omega_2} = l_y \tag{93}
\]

\[
-\Psi_{\mathcal{D}_1}^T J \Delta x + \Psi_{\mathcal{D}_1}^T \Delta y + \Psi_{\mathcal{D}_1}^T A^{-1} S \Delta \lambda_{\mathcal{D}_1} + \Psi_{\mathcal{D}_1}^T a \Delta z = \Psi_{\mathcal{D}_1}^T l_{\mathcal{D}_1} \tag{94}
\]

\[
a^T \Delta \lambda_{\mathcal{D}_1} + a^T \Delta \lambda_{\Omega_2} - l_{\mathcal{D}_2} \Delta z = l_z. \tag{95}
\]

Isolating the projection value in (94),

\[
\Delta \lambda_{\mathcal{D}_1} = \Psi_{\mathcal{D}_1} \left[ \Psi_{\mathcal{D}_1}^T A^{-1} S \Psi_{\mathcal{D}_1} \right]^{-1} \Psi_{\mathcal{D}_1}^T \left\{ l_{\mathcal{D}_1} + J \Delta x - \Delta y - A^{-1} S \Delta \lambda_{\Omega_2} - a \Delta z \right\}. \tag{96}
\]

Now, replacing the projection value (96) in (92) and isolating by field,

$$A_{11}(\lambda) \Delta x + A_{12}(\lambda) \Delta y + A_{13}(\lambda) \Delta \lambda_{\Omega_2} + A_{14}(\lambda) \Delta z = a_{1}(\lambda) \tag{97}$$

\% \ A_{11}(\lambda) \in \mathbb{R}^{n \times n}, \ A_{12}(\lambda), A_{13}(\lambda) \in \mathbb{R}^{n \times m}, \ A_{14}(\lambda) \in \mathbb{R}^{n \times 1} \text{ and } a_{1}(\lambda) \in \mathbb{R}^{n}.

As in the previous subsection, all the operators that are not related of the null space vector from the dual variable are described in Appendix B. The relevant operator is related to the vector of the null space, and in the case of (97), it is the operator $A_{13}(\lambda)$, where

$$A_{13}(\lambda) = -J^T \left\{ I_m - \Psi_{\mathcal{D}_1} \left[ \Psi_{\mathcal{D}_1}^T A^{-1} S \Psi_{\mathcal{D}_1} \right]^{-1} \Psi_{\mathcal{D}_1}^T A^{-1} S \right\}. \tag{98}
\]

The same way, related to the others operators, it is possible to achieve replacing the projection value (96) in (93) and isolating again the fields, ie,

$$A_{21}(\lambda) \Delta x + A_{22}(\lambda) \Delta y + A_{23}(\lambda) \Delta \lambda_{\Omega_2} + A_{24}(\lambda) \Delta z = a_{2}(\lambda) \tag{99}$$

\% \ A_{21}(\lambda) \in \mathbb{R}^{m \times n}, \ A_{22}(\lambda), A_{23}(\lambda) \in \mathbb{R}^{m \times m}, \ A_{24}(\lambda) \in \mathbb{R}^{m \times 1} \text{ and } a_{2}(\lambda) \in \mathbb{R}^{m},

and the operator $A_{23}(\lambda)$, associated with the null space,

$$A_{23}(\lambda) = \left\{ I_m - \Psi_{\mathcal{D}_1} \left[ \Psi_{\mathcal{D}_1}^T A^{-1} S \Psi_{\mathcal{D}_1} \right]^{-1} \Psi_{\mathcal{D}_1}^T A^{-1} S \right\}. \tag{100}
\]
Finally, replacing the projection value (96) in (95) and isolating the fields,

\[
A_{31(\lambda)} \Delta x + A_{32(\lambda)} \Delta y + A_{33(\lambda)} \Delta \lambda + A_{34(\lambda)} \Delta z = a_{3(\lambda)} \\
\forall \quad A_{31(\lambda)} \in \mathbb{R}^{m \times n}, \quad A_{32(\lambda)}, A_{33(\lambda)}, A_{34(\lambda)} \in \mathbb{R}^{m \times m}, \quad A_{34(\lambda)} \in \mathbb{R}^{m \times 1} \text{ and } a_{3(\lambda)} \in \mathbb{R}^{m}.
\]

(101)

Results in the operator \(A_{33(\lambda)}\), ie,

\[
A_{33(\lambda)} = a^T \left\{ I_m - \Psi_\lambda \Psi_\lambda^T A^{-1} S \Psi_\lambda \right\}^{-1} \Psi_\lambda^T A^{-1} S.
\]

(102)

As in the previous subsection, it is possible to see that all the other terms must belong to the null space operators

\[
a_{1(\lambda)} - A_{11(\lambda)} \Delta x - A_{12(\lambda)} \Delta y - A_{14(\lambda)} \Delta z \subset A_{13(\lambda)} \\
a_{2(\lambda)} - A_{21(\lambda)} \Delta x - A_{22(\lambda)} \Delta y - A_{24(\lambda)} \Delta z \subset A_{23(\lambda)} \\
a_{3(\lambda)} - A_{31(\lambda)} \Delta x - A_{32(\lambda)} \Delta y - A_{34(\lambda)} \Delta z \subset A_{33(\lambda)}.
\]

(103)

In all operators, when the Galerkin projection converges to the full space,

\[
\lim_{\Psi_\lambda, \Psi_\lambda^T \rightarrow I_m} \hat{A}_{33(\lambda)}(\Psi_\lambda) = O_{n \times m} \\
\lim_{\Psi_\lambda, \Psi_\lambda^T \rightarrow I_m} \hat{A}_{23(\lambda)}(\Psi_\lambda) = O_{m \times m} \\
\lim_{\Psi_\lambda, \Psi_\lambda^T \rightarrow I_m} \hat{A}_{33(\lambda)}(\Psi_\lambda) = O_{1 \times m}.
\]

(104)

This occurs because of the parcel that is common to all operators of the null space, ie,

\[
\lim_{\Psi_\lambda, \Psi_\lambda^T \rightarrow I_m} \left\{ I_m - \Psi_\lambda \Psi_\lambda^T A^{-1} S \Psi_\lambda \right\}^{-1} \Psi_\lambda^T A^{-1} S = O_{m \times m}.
\]

(105)

Moreover, in this case, it is possible to take two conclusions from the error related of Galerkin projection in the dual variable. The first conclusion is that the null space is directly dependent of the variables that compound the KKT condition. The direct dependency of the dual and slack variables results, in the case of IPM algorithm, a higher dependency of the barrier parameter value. The consequence of such dependency is demonstrated in the last subsection. In the next subsection, it is possible to notice the advantage of the Galerkin projection applied in one field per time when compared with the application of both Galerkin projections at the same time.

4.3 Error related to the primal and dual variable at the same time

In the last two subsections, the analysis of the null space operators demonstrates that a Galerkin projection in a unique field generates no direct error between the primal and dual variables. To understand better the advantage of the selection between the primal and dual fields, once per time, it is important to develop an error analysis with the Galerkin projection in two fields at the same time. To start such analysis, the fields that are reduced are defined, ie,

\[
\begin{bmatrix}
\Psi_x \Delta x + \Omega_x \Delta z \\
\Delta y \\
\Psi_\lambda \Delta \lambda + \Omega_\lambda \Delta \lambda
\end{bmatrix}
= 
\begin{bmatrix}
\Psi_x & O_{n \times m} & O_{n \times w_n} & o_n \\
o_m & I_m & O_{m \times w_n} & o_m \\
o_m & O_{m \times m} & \Psi_\lambda & o_m \\
o_w & o_m^T & o_m^T & I
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta y \\
\Delta \lambda
\end{bmatrix}
+ 
\begin{bmatrix}
\Omega_x & O_{n \times m-w_n} \\
o_m & I_m & O_{m \times m-w_n} \\
o_m & O_{m \times m-w_n} & \Omega_\lambda \\
o_{n-w_n} & o_{n-w_n}^T & o_{n-w_n}^T
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta y \\
\Delta \lambda
\end{bmatrix}
\]

(106)

\[\Rightarrow \Delta u = \Theta_{x,\lambda} \Delta u_{r,\lambda} + K_{x,\lambda} \Delta u_{r,\lambda}^T.\]
Now, the global Galerkin projection and the global null space operator are defined in (106), they are replaced in (52); the parcel of the error is represented by the term

$$\Theta_{14} = \Delta u_x \Delta \lambda_{\Omega}$$

Like in the previous two subsections, the idea is to isolate the two vectors from the null space of the primal and dual variables. Again, this approach is possible, exposing all the equations from (52) with null spaces of the primal and dual variables different of null vectors, ie,

$$-\Psi_{x}^T L_{xx} \Delta x_{\psi_x} - \Psi_{x}^T J^T \Delta \lambda_{\psi_x} - \Psi_{x}^T J^T \Delta \lambda_{\Omega} = \Psi_{x}^T I_x$$

$$-L_{yy} \Delta y + \Delta \lambda_{\psi_x} + \Delta \lambda_{\Omega} = I_y$$

$$-\Psi_{x}^T J \Delta x_{\psi_x} - \Psi_{x}^T J \Delta x_{\Omega} + \Psi_{x}^T J^T \Delta y + \Psi_{x}^T A^{-1} S \Delta \lambda_{\psi_x} + \cdots$$

$$\cdots + \Psi_{x}^T A^{-1} S \Delta \lambda_{\Omega} + \Psi_{x}^T a \Delta z = \Psi_{x}^T I_{(\delta)}$$

$$a^T \Delta \lambda_{\psi_x} + a^T \Delta \lambda_{\Omega} - I_{zz} \Delta z = I_z.$$ (111)

Developing a similar manipulation described in the previous two subsections, but now isolating the two null space and the other two fields in one equation,

$$A_{11(x,\lambda)} \Delta x_{\Omega} + A_{12(x,\lambda)} \Delta y + A_{13(x,\lambda)} \Delta \lambda_{\Omega} + A_{14(x,\lambda)} \Delta z = a_{1(x,\lambda)}.$$ (112)

The questions related in how to achieve Equation (112), as such as the values of all operators, can be answered in Appendix B. Once again, the important operators to the error analysis are related to both null spaces

$$A_{11(x,\lambda)} = a^T \Psi_{x} \left[ \Psi_{x}^T A^{-1} S \Psi_{x} \right]^{-1} \Psi_{x}^T J \left\{ I_n - \Psi_{x} \left[ \Psi_{x}^T L_{xx} \Psi_{x} \right]^{-1} \Psi_{x}^T L_{xx} \right\}$$

$$A_{13(x,\lambda)} = a^T \left\{ I_m - \Psi_{x} \left[ \Psi_{x}^T A^{-1} S \Psi_{x} \right]^{-1} \Psi_{x}^T A^{-1} S \right\}.$$ (113)

It is possible to see that all the other terms must to belong to

$$a_{1(x,\lambda)} - A_{12(x,\lambda)} \Delta y - A_{14(x,\lambda)} \Delta z \subset \left[ A_{11(x,\lambda)} \cup A_{13(x,\lambda)} \right].$$ (115)

When the range space converges to the full space,

$$\lim_{\Psi, \Psi' \to I_n} \hat{A}_{11(x,\lambda)} (\Psi, \Psi') = O_{1 \times n}$$

$$\lim_{\Psi, \Psi' \to I_m} \hat{A}_{13(x,\lambda)} (\Psi) = O_{1 \times m}. $$ (116)

It is possible to realize that (113) and (114) contain the parcels described in (91) and (105), respectively. In the point of view of the dual variable, the null space operator, even with both projections applied at the same time, remains with the same behaviors described in the previous subsection. However, in the point of view of the primal variable, the behavior of the operator is different when compared to the first subsection. Now, the operator in (113) accumulates both errors related to the Galerkin projection in the primal and dual variables. In the case that an increase of error in the dual field exists, both fields are affected. Furthermore, in the case that an increase of error in the primal field exists, such field must deal with the error from the dual field too, resulting more time processing to reduce the error of the range space in both cases.

As can be seen in the first subsections, when the Galerkin projection is applied in the primal field, the error in the primal field is only partially dependent on the KKT condition. These property permits, when the barrier parameter is
not properly chosen, lower errors in the primal field when compared to the dual field, which this last one is completely dependent on the barrier parameter. The level of dependency with the KKT condition of each parcel, described in (91) and (105), can be seen in more details in the next subsection, through the sensibility analysis.

4.4 | IPM algorithm sensibility that affects the MOR error

In this subsection, the idea is a sensibility analysis related to the operators that compound the parcels (91) and (105). Based on such parcels, this subsection analyzes the sensibility from the Hessian of the Lagrangian function related to the primal field, the Lagrange multipliers (dual field) and the slack variables. Although the range space converging to the full space eliminates all null space operators, before this happens, the operation between the inverse of the reduced operator with the original operator can amplify, in this meantime, the distance between the range and full spaces.

Others operators from the parcels (91) and (105) are the Galerkin projections from the primal and dual variables. Related to these operators, there are two points to consider. The first point is that the sensibility analysis depends on the adaptive strategy from the MOR; the choice by the generation or update is made by eigenvectors, singular vectors, or Ritz vectors, as such as the criteria of selection change the sensibility analysis. The second point is that the sensibility of primal and dual variables changes the snapshots that generate the Galerkin projections, becoming the process extremely complicate to analyze. Because of these both particularities of the adaptive strategy, this subsection focuses on the operators from the infeasible primal-dual IPM with artificial variables.

The sensibility analysis starts in the IPM algorithm. In such algorithm, as described in Flowchart 1, the barrier parameter must achieve the value lower than a default tolerance; in the meantime, the barrier parameter assumes different values, decreasing until the default tolerance value. Considering now two sequential barrier parameters, from step “j” to step “j + 1”, as the imposition of two incomplete KKT conditions,

\[ \mu^j I_m = A^j S^j \]
\[ \mu^{j+1} I_m = [A^j + \Delta A^{j+1}] [S^j + \Delta S^{j+1}] \]
\[ \forall \Delta A = \text{diag}(\Delta \lambda) \in \mathbb{R}^{m \times m}, \Delta S = \text{diag}(\Delta s) \in \mathbb{R}^{m \times m}. \]

Assuming the difference between (118) and (117) results

\[ [\mu^{j+1} - \mu^j] I_m = A^j \Delta S^{j+1} + \Delta A^{j+1} S^j + \Delta A^{j+1} \Delta S^{j+1}. \]

It is possible to realize that the sensibility (variation of the dual variable and slack variable) of the IPM algorithm is dependent on the variation (or decrease) from the barrier parameter. Usually, there are adaptive strategies to determine the variation of the barrier parameter to achieve a better convergence; however, there is no exact function to describe the barrier parameter variation. Because of that, the convergence of the IPM algorithm, as such as the error described by the parcel represented by (105), is dependent on an empirical estimation of the variation from the barrier parameter. In addition, it is important to highlight that a low difference between barrier parameters is not able to ensure a low variation of both variables; when low differences are imposed, three situations can occur, ie,

\[ \lim_{\Delta A \to a_m} [\mu^{j+1} - \mu^j] I_m = A^j \Delta S^{j+1} \quad \exists \quad \lambda^{j+1}_i \approx 0 \quad \text{or} \quad \Delta s^{j+1}_i \approx 0 \]
\[ \lim_{\Delta S \to a_m} [\mu^{j+1} - \mu^j] I_m = \Delta A^{j+1} S^j \quad \exists \quad \Delta \lambda^{j+1}_i \approx 0 \quad \text{or} \quad s^{j}_i \approx 0 \]
\[ \lim_{\Delta A \to a_m, \Delta S \to a_m} [\mu^{j+1} - \mu^j] I_m = O_m. \]

Based in (120) and (121), it is possible to realize that the low variation can occur in only one variable. Although the term from (105) is compound with the inverse of the dual variable, the sensibility analysis represented by (120), (121), and (122) are enough to demonstrate the dependency of (105) related to the barrier parameter. In the case of the term from (91), the dependency of the barrier parameter is not as clear as (105), but it is possible to realize that some terms that compound the Hessian of the Lagrangian function related to the primal variable have direct influence with the barrier parameter.
To make the sensibility analysis clearer, the terms related to the superior and inferior limit of the primal variable from (22) are represented by

$$B_{xx} = |H|^{-2} + |Z|^{-2} \quad \forall \quad B_{xx} \in \mathbb{R}^{n \times n}. \quad (123)$$

Replacing (123) in (22),

$$L_{xx} = F_{xx} + \theta^T \lambda + \mu B_{xx}. \quad (124)$$

As the previous analysis of sensibility, considering now two sequential barrier parameters, from step “j” to step “j + 1”, and replacing (117) and (118) for each step, respectively,

$$L_{xx}' = F_{xx}' + [\theta']^T \lambda' + A'S'B_{(x)}' \quad (125)$$

$$L_{xx}^{j+1} = F_{xx}^{j+1} + [\theta^{j+1}]^T [\lambda^{j} + \Delta \lambda^{j+1}] + [A^{j} + \Delta A^{j+1}] [S^{j} + \Delta S^{j+1}] B_{(x)}^{j+1}. \quad (126)$$

Again, assuming the difference between (126) and (125) results

$$\Delta F_{xx}^{j+1} = F_{xx}^{j+1} - F_{xx}^{j} \in \mathbb{R}^{n \times n}, \quad \Delta \theta^{j+1} = \theta^{j+1} - \theta^{j} \in \mathbb{R}^{n \times n \times n} \quad (127)$$

and $\Delta B_{(x)}^{j+1} = B_{(x)}^{j+1} - B_{(x)}^{j} \in \mathbb{R}^{n \times n}$. \quad (128)

Splitting Equation (127) in one parcel referent to the variation from the primal variable, and another one related to the variation from the dual and slack variables,

$$L_{xx}^{j+1} - L_{xx}^{j} = \Delta L_{xx^{(j+1)+1}} + \Delta L_{xx^{(j+1)+1}} \quad \forall \quad \Delta L_{xx^{(j+1)+1}}, \Delta L_{xx^{(j+1)+1}} \in \mathbb{R}^{n \times n}. \quad (129)$$

where

$$\Delta L_{xx^{(j+1)+1}} = \Delta F_{xx^{(j+1)+1}} + [\Delta \theta^{j+1}]^T \lambda^{j} + A'S'\Delta B_{(x)}^{j+1} \quad (130)$$

$$\Delta L_{xx^{(j+1)+1}} = [\theta^{j+1}]^T \Delta \lambda^{j+1} + \Delta A^{j+1} S^{j+1} + \Delta A^{j+1} \Delta S^{j+1}] B_{(x)}^{j+1}. \quad (131)$$

Operator (130) is directly dependent on the variation from the primal variable, and it is clear that, when the variation of the primal variable tends to zero, such operator vanishes. However, in operator (131), the analysis of the variation follows the conditions from (120), (121), and (122), ie,

$$\lim_{\Delta \lambda \to 0_n} \Delta L_{xx^{(j+1)+1}} = A' \Delta S^{j+1} B_{(x)}^{j+1} \quad (132)$$

$$\lim_{\Delta S \to 0_n} \Delta L_{xx^{(j+1)+1}} = [\theta^{j+1}]^T \Delta \lambda^{j+1} + \Delta A^{j+1} S^{j+1} B_{(x)}^{j+1} \quad (133)$$

$$\lim_{\Delta \lambda \to 0_n, \Delta S \to 0_n} \Delta L_{xx^{(j+1)+1}} = 0_n. \quad (134)$$

In the same way, a low difference between barrier parameters cannot ensure a higher convergence related to the primal variable. Another condition, with no relation to the barrier parameter, is the high variation generated by (123). Such condition comes from the definition of the problem, which is in two extremes situations that can be achieved from a specific value of the primal variable, ie,

$$\lim_{x_i \to \hat{x}_i} B_{(x_i),i} = \frac{1}{[\hat{x}_i - x_i]^2} + \infty \quad \Rightarrow \quad B_{(x_i),i} = \infty \quad (135)$$

$$\lim_{x_i \to \hat{x}_i} B_{(x_i),i} = \infty + \frac{1}{[x_i - \alpha_i]^2} \quad \Rightarrow \quad B_{(x_i),i} = \infty. \quad (136)$$
The variations described in (135) and (136) are related to the constraints of the problem. In this case, it is possible to obtain a low variation from (131) only if condition (134) is fulfilled. Based on what has been shown so far, it is possible to realize that a selection of the barrier parameter variation that ensures a fast convergence is essential to achieve quickly the low variation of parcels (91) and (105). Once a low variation is achieved, the snapshots from the previous points (vector spaces) from the manifold \( \mathcal{M} \) are closer to each other, ensuring Galerkin projections with a better representation of the current vector space \( \mathcal{T}_u \).

To reduce the MOR error and achieve a faster convergence, this work recommends the application of two empirical approaches. The first one is, obviously, an adaptive strategy to the barrier parameter variation that leads to a convergence closer in the quadratic rate. The second approaches is the normalization of the objective and constraints functions; in case of both types of functions have similar proportion, the variation of the Lagrange multiplier (dual variable) is lower, resulting also in a lower variation of the term (131). In consequence of different behaviors from both kind of null space operator and by the lack of objective approach to control them, the next section of this work suggests an adaptive strategy that changes the field of Galerkin projection, according to the comparison between errors.

5 | MOR AND ADAPTIVE STRATEGY

5.1 | Introduction to the MOR and adaptive strategy

Although the MOR presented until now can be applied as a surrogate model for several steps, the proposal mentioned previously is to generate an initial guess to the PCG, in a similar way as described in the works of Tromeur-Dervout and Vassilevska\(^{10}\) and Markovinovic and Jansen.\(^{15}\) However, different of the works cited previously, there is a selection based on error related to each projection. Although such an approach can be applied in several types of analysis, an efficient processing time only is achieved to a low-cost inverse Hessian. An option to obtain a low-cost inverse Hessian is to apply an explicit convex approximation function as described in the work of Bruyneel et al.\(^{29}\) Such approximation function generates a diagonal (explicit property) and invertible (convex property) Hessian.

The purpose of this section is to use the models developed in Sections 2 and 3 and, through an accurate initial guess, to provide an improvement of convergence to the PCG method. To facilitate the understanding of how to select the models, this section is divided into three subsections. The first subsection is how to generate and restart (if necessary) the operator of the Galerkin projection. The second subsection is related on how to select the initial guess from a Galerkin projection with lower error. The last subsection is how to update and correct the operator of the Galerkin projection. Flowchart 3 and Flowchart 4 show, respectively, the MOR with adaptive strategy and only the MOR inside the previous flowchart.

5.2 | Initialization and restart of the Galerkin projection

To develop the initial projection operators, it is necessary to collect snapshots of both fields of the problem. In this work, where all the analyses are developed online, the gathering starts after the acquisition of the new value of the barrier parameter. In a specific barrier parameter value, a predetermined number of iterations \( k_v \) is selected; each iteration obtains two snapshots, ie, the first one from the increment of the primal and the second one from dual variable.

Flowchart 2  Adaptative strategy – from Flowchart 1. MOR, model order reduction [Colour figure can be viewed at wileyonlinelibrary.com]
Flowchart 3  MOR - from Flowchart 2. MOR, model order reduction; PCG, preconditioned conjugate gradient [Colour figure can be viewed at wileyonlinelibrary.com]

Such choice for snapshots is based on the property that increments are not accumulative values, increasing the chance of a higher sensibility for iteration. When the predetermined quantity of iterations $k_I$ is reached, it is applied to the classical POD. The classical POD selected a quantity of singular vector from dimensions $n$ and $m$ based on the largest singular values in each decomposition. In Algorithm 2, this procedure is demonstrated more clearly.

**Algorithm 2 Initial $\Psi$**

| input: $\Delta x^{k+1}$, $V_x$, $\Delta \lambda^{k+1}$, $V_\lambda$ |
| $V_x^{\text{new}} = [V_x \, \Delta x^{k+1}]$, $V_\lambda^{\text{new}} = [V_\lambda \, \Delta \lambda^{k+1}]$ |
| $n_v = n_v + 1$ |
| if $k_v > n_v$ |
| $V_x = V_x^{\text{new}}$, $V_\lambda = V_\lambda^{\text{new}}$ |
| else |
| $\Psi_x^{\text{new}} = POD(V_x^{\text{new}})$ |
| $\Psi_\lambda^{\text{new}} = POD(V_\lambda^{\text{new}})$ |
| $w_n = w_n^{\text{new}}$ |
| end if |

output: $\Psi_x$, $V_x$, $\Psi_\lambda$, $V_\lambda$

The result of the procedure described in Algorithm 2 is the initial Galerkin projection for each field (commands inside box means that the process can be parallelized). After such procedure, the MOR is initialized. During the MOR, the matrices described in Algorithm 2 suffer updates; these updates store the last $k_I$ snapshots. Such matrices are applied when a Galerkin projection operator reaches a maximum dimension $w_n^{\text{max}}$ or $w_m^{\text{max}}$, then, a new classical POD is applied to obtain a new Galerkin projection. In Algorithm 3, this procedure is demonstrated more clearly.

**Algorithm 3 Restart $\Psi$**

| input: $\Delta x^{k+1}$, $V_x$, $\Psi_x$, $\Delta \lambda^{k+1}$, $V_\lambda$, $\Psi_\lambda$ |
| $V_x^{\text{new}} = [\Delta x^{k+1}, \ldots, \Delta x^{k+1}]$, $V_\lambda^{\text{new}} = [\Delta \lambda^{k+1}, \ldots, \Delta \lambda^{k+1}]$ |
| if $w_n^{\text{max}} > w_n$ |
| $V_x = V_x^{\text{new}}$ |
| else |
| $\Psi_x^{\text{new}} = POD(V_x^{\text{new}})$ |
| $\Psi_\lambda = \Psi_\lambda^{\text{new}}$, $w_n = w_n^{\text{new}}$ |
| $V_\lambda = V_\lambda^{\text{new}}$ |
| end if |

if $w_m^{\text{max}} > w_m$ |
| $V_\lambda = V_\lambda^{\text{new}}$ |
| else |
| $\Psi_\lambda^{\text{new}} = POD(V_\lambda^{\text{new}})$ |
| $\Psi_\lambda = \Psi_\lambda^{\text{new}}$, $w_m = w_m^{\text{new}}$ |
| $V_\lambda = V_\lambda^{\text{new}}$ |
| end if |

output: $\Psi_x$, $V_x$, $\Psi_\lambda$, $V_\lambda$
Between these two processes described in Algorithm 2 and Algorithm 3, the Galerkin projection operators suffer variations of the dimensions to reach the next points (vector spaces $T_{\Psi}M$) in the manifold $M$. To obtain the variation of dimension, the solution by the projection of the fields (range space) and the full solution of the fields are necessary. The next subsection describes how to obtain both solutions.

## 5.3 Selection of the initial guess and PCG

In this work, the solution of the reduced space provides the initial guess of the PCG. A contribution of this research is the initial guess selected by the Galerkin projection with lower error. Before the description of the process, it is important to highlight that the efficiency of this approach depends on inverse Hessian with a simple solution, as described previously. Consider that, in such approach, both reduced Schur complements are processed per iteration, as described in Algorithm 4.

**Algorithm 4** Solve $\Delta u_{\Psi}$ to initial guess

input: $\Psi_x, \Psi_\lambda, \Gamma, \varphi$

- solve reduced system: $\Delta x_{\Psi_x}, \Delta z_{\Psi_x}$, Schur complement - Primal calculate: $\Delta \lambda_{\Psi}$
- solve reduced system: $\Delta \lambda_{\Psi_x}, \Delta z_{\Psi_x}$, Schur complement - Dual calculate: $\Delta x_{\Psi}$

output: $\Delta x_{\Psi_x}, \Delta z_{\Psi_x}, \Delta \lambda_{\Psi_x}, \Delta \lambda_{\Psi}, \Delta z_{\Psi}, \Delta x_{\Psi}$

Once the range of space vectors of the primal and dual variables is obtained, the original dimension of each field is analyzed. The original dimension of the field selects the Schur complement applied in the process. As a recommended criterion, the lower original dimension selects if the Schur complement is applied to the primal or dual variable. Based on the previous criterion, if the Schur complement with a primal variable is chosen ($n < m$), the initial guess is between Equations (65) and (81). In (65), the equation comes from the range space of the primal variable, and Equation (81) comes from the range space of the dual variable.

However, if the Schur complement with a dual variable is chosen ($n > m$), the initial guess is between Equations (67) and (79). In (67), the equation comes from the range space of the dual variable, and Equation (79) comes from the range space of the primal variable. This work tests two kinds of criteria; the criteria to select between Equations (65) and (81), or between Equations (67) and (79). Based on that, to the problem solved with the Schur complement from the primal variable,

$$\min (\varepsilon_{xx}, \varepsilon_{x\lambda}) \quad \forall \quad \varepsilon_{xx}, \varepsilon_{x\lambda} \in \mathbb{R}^+.$$  \hfill (137)

To the Schur complement from the dual variable,

$$\min (\varepsilon_{\lambda x}, \varepsilon_{\lambda\lambda}) \quad \forall \quad \varepsilon_{\lambda x}, \varepsilon_{\lambda\lambda} \in \mathbb{R}^+.$$  \hfill (138)

In the first criterion, the values from (137) and (138) are related to the lower error from the initial guess or the lower error from the Galerkin projection, ie,

$$\varepsilon_{xx} = \| \Delta x - \Delta x_{\Psi_x} \|_2, \quad \varepsilon_{xx} = \| \Delta x - \Delta x_{\Psi_x} \|_2, \quad \varepsilon_{\lambda x} = \| \Delta \lambda - \Delta \lambda_{\Psi_x} \|_2, \quad \varepsilon_{\lambda x} = \| \Delta \lambda - \Delta \lambda_{\Psi} \|_2.$$  \hfill (139)

Moreover, the second criterion, which comes from (56) and (72), is defined by the minimum Galerkin error

$$\varepsilon_{xx} = \varepsilon_{\lambda x} = \left\| \Phi_x^T \Gamma K_x \Delta u_e \right\|_2, \quad \varepsilon_{xx} = \varepsilon_{\lambda x} = \left\| \Phi_x^T \Gamma K_x \Delta u_e \right\|_2.$$  \hfill (140)

Unlike (139), the selection from (140) does not depend on the field related to the Schur complement to solve the full system. In Algorithm 5, the procedure of the selection described previously is demonstrated more clearly.
In this stage, with a solution in the range space and the full solution, it is possible to find the solution of the null space; such space is applied to update and correct the Galerkin projection operator to both fields. The next subsection describes such a process.

5.4 | Update and correction of the Galerkin projection

As described in the previous subsection, with the solution from the range and full space, it is possible to obtain the solution of the null space by the difference of both. To update the Galerkin projection, it is applied the concept of Ritz vector. In this case, the Ritz vector is developed by the orthonormalization (by Gram-Schmidt process) of the null space vector, i.e.,

\[
\psi_x = \text{orthonormalization}(\Delta x_{\Omega}), \quad \psi_x, \in \mathbb{R}^n, \quad \text{and} \quad \psi_\lambda = \text{orthonormalization}(\Delta \lambda_{\Omega}), \quad \psi_\lambda, \in \mathbb{R}^m. \tag{141}
\]

In relations (141), the orthonormalization occurs related to the Galerkin projection operator vectors. Consequently, after the development of the Ritz vectors, such vectors are added to the Galerkin projection. In the Galerkin projection operators, the Ritz vector increases the number of columns; the number of columns represents the dimension of the reduced space. During the update of the Galerkin projection, the norm of errors is calculated, and the error can be developed at the same time of the update from the Galerkin projection operators. This procedure is demonstrated more clearly in Algorithm 6.

---

Algorithm 5 PCG to solve \( \Delta u \)

input: \( \Delta x_{\psi}, \Delta z_{\psi}, \Delta \lambda_{\psi}, \Delta \lambda_{\psi}, \Delta x_{\psi}, \Gamma, \varphi, \varepsilon_{xx}, \varepsilon_{\lambda x} \)

if (Schur complement Primal = true) or \((n > m)\)

\[
\begin{align*}
&\text{if } \varepsilon_{xx} < \varepsilon_{\lambda x} \\
&\quad \text{solve full system: } \Delta x^{k+1}, \Delta z^{k+1} - \text{PCG initial guess } \Delta x_{\psi} \\
&\text{else} \\
&\quad \text{solve full system: } \Delta x^{k+1}, \Delta z^{k+1} - \text{PCG initial guess } \Delta x_{\psi} \\
&\end{align*}
\]

Schur complement – Primal

end if

calculate: \( \Delta \lambda \)

else

\[
\begin{align*}
&\text{if } \varepsilon_{\lambda x} < \varepsilon_{xx} \\
&\quad \text{solve full system: } \Delta \lambda^{k+1}, \Delta z^{k+1} - \text{PCG initial guess } \Delta \lambda_{\psi} \\
&\text{else} \\
&\quad \text{solve full system: } \Delta \lambda^{k+1}, \Delta z^{k+1} - \text{PCG initial guess } \Delta \lambda_{\psi} \\
&\end{align*}
\]

Schur complement – Dual

end if

calculate: \( \Delta x \)

end if

calculate: \( \Delta y^{k+1}, \Delta s^{k+1} \)

output: \( \Delta x^{k+1}, \Delta \lambda^{k+1}, \Delta z^{k+1}, \Delta y^{k+1}, \Delta s^{k+1} \)

---

Algorithm 6 Update \( \Psi \) and \( \varepsilon \)

input: \( \Psi_x, \Psi_\lambda, \Gamma, \varphi, \Delta x_{\psi}, \Delta x^{k+1} \)

\[
\begin{align*}
\Delta x_{\Omega} &= \Delta x^{k+1} - \Delta x_{\psi} \\
\text{calculate: } \Psi_x, \varepsilon_{xx}, \varepsilon_{\lambda x} \\
\Psi_{\text{new}} &= [\Psi_x, \Psi_\lambda] \\
\Psi_x &= \Psi_{\text{new}}, w_n = w_n + 1 \\
\Delta \lambda_{\Omega} &= \Delta \lambda^{k+1} - \Delta \lambda_{\psi} \\
\text{calculate: } \Psi_\lambda, \varepsilon_{\lambda x}, \varepsilon_{\lambda x} \\
\Psi_{\text{new}} &= [\Psi_\lambda] \\
\Psi_\lambda &= \Psi_{\text{new}}, w_m = w_m + 1
\end{align*}
\]

output: \( \Psi_x, \varepsilon_{xx}, \varepsilon_{\lambda x}, \Psi_\lambda, \varepsilon_{\lambda x} \)
The process represented by Algorithm 6 is repeated per iteration. During the IPM algorithm, the repetition of this process, described in Flowchart 3, is possible to generate a nonnull kernel in the reduced matrices from Equations (64) and (78), resulting in singular matrices. To avoid this behavior, the determinants of the matrices from Equations (64) and (78) are calculated; when a determinant reaches a value extremely close to zero, the last Ritz vector is discarded. In this case, the Galerkin projection operator has the number of columns reduced in one dimension with the attempt to restore the bijectivity of the matrices cited previously. The reduction of the number of columns, and consequently the dimension of the reduced space, is repeated until the values of the determinant from each matrix increases again. This procedure is demonstrated more clearly in Algorithm 7.

**Algorithm 7 Correction $\Psi$**

input: $\Psi_x, \Psi_A, \Gamma$

```
while det $\to$ 0
    $\Psi_x = [\Psi_x^{\text{new}}, \Psi_x]$
    $\Psi_x = [\Psi_x^{\text{new}}, \Psi_x = w_n - 1$
    $\det \left( \left( \Pi_{11} \right)_r - \left( \Pi_{12} \right)_r \frac{(\Pi_{13})_r}{\Pi_{13}} \right)$
end while
output: $\Psi_x, \Psi_A$
```

All the processes described in this section, from Flowchart 2 to Algorithm 7, show that this work is concerned to simplify the adaptive strategy. The classical POD and Ritz vectors are basic strategies to the Galerkin projection generation, and the reduction of this one is developed only to ensure the solvability of the reduced systems. These works assume this approach to focus the analysis in the reduction of one field per time and the improvement of the convergence from the PCG by the selection of the initial guess generated by the MOR. The next section shows the results of simulations based on this adaptive strategy.

6 | EXAMPLES

6.1 Definition of the examples

As mentioned in the last section, this work developed two examples to demonstrate the advantages and disadvantages of the adaptive strategy described previously. These examples are an extension of the example described in the work of Svanberg\(^{26}\); the main difference is related to the number of constraint functions, but properties as convexity and limited domain in the primal variable are the same, ie,

\[
\begin{align*}
\text{minimize} & \quad x^T W x \\
\text{subject to} & \quad b_1 - x^T [P + R_k] x \leq 0 \quad \text{for} \quad 1 \leq k \leq m_1 \\
& \quad b_2 - x^T [Q + T_k] x \leq 0 \quad \text{for} \quad m_1 < k \leq m_2 \\
& \quad b_3 - x^T [P + T_k] x \leq 0 \quad \text{for} \quad m_2 < k \leq m_3 \\
& \quad b_4 - x^T [Q + R_k] x \leq 0 \quad \text{for} \quad m_3 < k \leq m \\
& \quad -e_n < x < e_n \\
\forall & \quad k, m_1, m_2, m_3 \in \mathbb{N} \quad , \quad b_1, b_2, b_3, b_4 \in \mathbb{R} \quad \text{and} \quad W, P, Q, R_k, T_k \in \mathbb{R}^{n \times n}.
\end{align*}
\]

(142)

A modification is made related to the aforementioned work\(^{26}\), the explicit convex function adopted is the classical method of moving asymptotes (MMA), as described in the work of Svanberg\(^{25}\), and in the work of Bruyneel et al.\(^{29}\).
It is important to highlight also, different of the works of Wächter and Biegler\textsuperscript{24} and Herskovits,\textsuperscript{30} the IPM algorithm is processing without line-search. These both modifications are made to facilitate the reproducibility of the simulations. More details can be seen in Appendix C and Appendix D.

Moreover, it is important to mention that all reduced linear systems are solved by Cholesky decomposition, and the preconditioned of the CG method, different of more advanced approaches as in the works of Ghosh et al.,\textsuperscript{31} Risler and Rey,\textsuperscript{32} and Anzt et al\textsuperscript{33} is always the diagonal of the Schur complement matrix. Related to the simulated models, this work compares four kinds of simulations: the simulation with the full model, MOR with a fixed field, MOR with field selection based of initial guess, and MOR with field selection based of the Galerkin error. Each simulation measured the time and the number of iterations of the PCG per step.

Related to the procedure, both examples are simulated following the procedures.

1. Both examples are simulated by applying the Schur complement in the primal form (59) and dual form (75). The application of the Schur complement occurs in separate simulations and they are indicated by the term “primal” or “dual” in the title of the graph.

2. In the full model, when simulated with the Schur complement in the primal form, it applies the previous value of $\Delta x$ as initial guess; when simulated with the Schur complement in the dual form, it applies the previous value of $\Delta \lambda$ as initial guess.

3. The simple MOR simulation is made with the Galerkin projection always in the same field. When the MOR is simulated with the Schur complement in the primal form, it applies the value of (65) as initial guess; when simulated with the Schur complement in the dual form, it applies the value of (79) as initial guess. The result from this simple MOR simulation is described in the legend of the graphs by “x” to the primal form and “lambda” to the dual form;

4. The MOR simulation proposed in this work is made with Galerkin projection alternating the fields. When the MOR is simulated with the Schur complement in the primal form, it can apply between the equations from (65) and (81) as initial guesses that are dependents of a criterion; when simulated with the Schur complement in the dual form, it can apply between the equations from (67) and (79) as initial guesses that are dependents of the same criterion. The result from this MOR simulation is described in the legend of the graphs by “x” and “lambda”.

5. In the MOR simulation proposed in this work, the criteria to select the initial guess can be made between (139) and (140). The first one is based on the minimum error between the initial guess and the final result of the PCG; the second is based on the minimum error between the Galerkin error fields. In the graphs, result from (139) is described in the legend by “initial guess error”, and the result from (140) is described in the legend by “Galerkin error”. More details are presented in Algorithm 5.

6. Related to the demonstration of the results in the graphs, one part of graphs is demonstrated related to the steps; a step represents the output of Flowchart 1, and in the graph, it is represented by

$$[\text{Value of the graph}]^i = \sum_{k=1}^{n_{\text{total}}} [\text{Value of the algorithm}]^{k,i}$$

$$n_{\text{total}} = \text{number of iterations in the step } i.$$  \hfill (143)

7. Another part of them is demonstrated related to the internal iterations of the IPM algorithm that occur inside the algorithm from Flowchart 1, and in the graph, it is represented by

$$[\text{Value of the graph}]^k = \sum_{i=1}^{n_{\text{total}}} [\text{Value of the algorithm}]^{k,i}$$

$$n_{\text{total}} = \text{number of steps that exist the iteration } k.$$  \hfill (144)

The internal iterations result in a step; each step has a specific number of internal iterations. Because of that, the number of iterations showed in the graph is determined by the step with more iteration. In the graphs, the differentiation of both approaches is showed through the name of the horizontal axis as “step” or “iteration”. The different representations help to understand global aspects of the results (steps), as well as local aspects, which occur inside the IPM algorithm (iterations).
6.2 | Example 1

The first example is a simulation of a problem with more design variables (primal) than constraints (dual). Such example represents the majority of problems in computational mechanic problems with multifields. The data are

\[
\begin{align*}
    n & = 2000, & m & = 1000, & w_{\text{max}}^n & = 60, & w_{\text{max}}^m & = 60, & k_v & = 5 \\
    m_1 & = 250, & m_2 & = 500, & m_3 & = 750, & x^0 & = 0.25e_n.
\end{align*}
\] (145)

In (145), they are respectively the number of design variables, the number of constraints, the initial value of each design variable, and the maximum dimension of each reduced space.

6.3 | Example 2

The second example is a simulation of a problem with more constraints (dual) than design variables (primal). Such example does not represent the majority of problems in computational mechanic problems with multifields, but it has a behavior that is important to show for this work. The data are

\[
\begin{align*}
    n & = 1000, & m & = 2000, & w_{\text{max}}^n & = 60, & w_{\text{max}}^m & = 60, & k_v & = 5 \\
    m_1 & = 500, & m_2 & = 1000, & m_3 & = 1500, & x^0 & = 0.25e_m.
\end{align*}
\] (146)

In (145), they are respectively the number of design variables, the number of constraints, the initial value of each design variable, and the maximum dimension of each reduced space.

GRAPH 1  A, Ex 1 – Primal – PCG iterations per step; B, Ex 1 – Dual – PCG iterations per step; C, Ex 2 – Primal – PCG iterations; D, Ex 2 – Dual – PCG iterations. MOR, model order reduction; PCG, preconditioned conjugate gradient [Colour figure can be viewed at wileyonlinelibrary.com]
6.4 Results of the examples

The first results in Graph 1 show the total number of iterations in the PCG per step.

It is possible to realize, in these examples, that Graphs 1B and 1D show reductions related to the MOR with the selection of the fields if compared to the full model and MOR only with the dual field. However, in Graphs 1A and 1C, the reductions are only meaningful related to the full model; the selection of the field does not show relevant reduction during the simulation. The explanation of the different behavior by the different application of the Schur complement is described at the end of these examples through an analysis of error. Table 1 shows the accumulative number of iterations in the PCG.

For now, a more direct analysis is made related to the processing time; the second type of results compare the processing time of the PCG from the full model with the processing time of the PCG from the three others reduced models. However, for a fair comparison, each reduced model is added the processing time of the reduced linear system and the norm of errors.

Graphs 2A, 2B, 2C, and 2D do not show relevant changing of results when compared to the PCG iteration in Graphs 1A, 1B, 1C, and 1D. This means, at least for the dimensions selected to the both reduced space, that the additional processing time caused to the adaptive strategy is not relevant to the total processing time of the problem. Table 2 shows the accumulative processing time during the entire simulation.

---

**TABLE 1**  Total iterations of the preconditioned conjugate gradient

| Total iterations | Ex. 1 - Primal | Ex. 1 - Dual | Ex. 2 - Primal | Ex. 2 - Dual |
|------------------|---------------|--------------|----------------|--------------|
| Full             | 213 618       | 385 285      | 156 049        | 392 979      |
| MOR - x or λ     | 130 946       | 301 895      | 94 591         | 332 952      |
| MOR - x and λ - Galerkin | 127 630 | 141 124      | 94 134         | 105 921      |
| MOR - x and λ - Init. Guess | 120 436 | 131 962      | 93 259         | 96 693       |

Abbreviations: MOR, model order reduction.

---

**GRAPH 2**  A, Ex 1 – Primal – Processing time per step; B, Ex 1 – Dual – Processing time per step; C, Ex 2 – Primal – Processing time per step; D, Ex 2 – Dual – Processing time per step. MOR, model order reduction [Colour figure can be viewed at wileyonlinelibrary.com]
Table 3 shows the accumulative processing time, only of the adaptive strategy, during the entire simulation. Although the selection of the fields increases the processing time, it is realized that the adaptive strategy described in this work shows some better results compared to the full model and the MOR with only one reduced space; because of that, the next results are focused on the MOR with the selection of the field. The next type of results shows the number that each field is selected per step based on the adaptive strategy described in the last section.

In Graphs 3A and 3C, it is possible to notice that the selected field of the norm from initial guess error changes considerably (the curves intersect) if compared to the selected field of the norm from the Galerkin error. This behavior occurs in the primal field because the adaptive strategy imposes a reduction described in Algorithm 7. The reduction in the local

| Processing time (s) | Ex. 1 - Primal | Ex. 1 - Dual | Ex. 2 - Primal | Ex. 2 - Dual |
|---------------------|----------------|--------------|----------------|--------------|
| Full                | 4091.152       | 1992.631     | 928.147        | 7584.814     |
| MOR - x or λ        | 2595.037       | 1578.512     | 692.46         | 7614.148     |
| MOR - x and λ - Galerkin | 2587.803   | 839.608      | 647.911        | 2231.762     |
| MOR - x and λ - Init. Guess | 2611.418 | 855.063      | 632.362        | 2689.959     |

Abbreviations: MOR, model order reduction.

| Processing time (s) | Ex. 1 - Primal | Ex. 1 - Dual | Ex. 2 - Primal | Ex. 2 - Dual |
|---------------------|----------------|--------------|----------------|--------------|
| Full                | 0              | 0            | 0              | 0            |
| MOR - x or λ        | 0.983          | 0.615        | 1.374          | 0.943        |
| MOR - x and λ - Galerkin | 36.612     | 35.563       | 52.489         | 54.439       |
| MOR - x and λ - Init. Guess | 40.019    | 45.69        | 43.405         | 59.164       |

Abbreviations: MOR, model order reduction.

GRAPH 3  A, Ex 1 – Primal – Selections of the field per step; B, Ex 1 – Dual – Selections of the field per step; C, Ex 2 – Primal – Selections of the field per step; D, Ex 2 – Dual – Selections of the field per step [Colour figure can be viewed at wileyonlinelibrary.com]
TABLE 4 Selections of the field according to the adaptive strategy

| Selection of the field | Ex. 1 - Primal | Ex. 1 - Dual | Ex. 2 - Primal | Ex. 2 - Dual |
|------------------------|---------------|------------|---------------|-------------|
| MOR \( \times \) and \( \lambda \) - Galerkin | 876 | 323 | 870 | 315 |
| MOR \( \times \) and \( \lambda \) - Init. Guess | 473 | 730 | 1008 | 191 |

Abbreviations: MOR, model order reduction.

Galerkin projection acts directly of the norm from initial guess error. Such behavior occurs only in some parcels of the norm from the Galerkin error; the result is a norm much less sensitive than the norm from initial guess error. Table 4 shows the total value of the selected fields from the examples.

To explain better what happens to the primal field, the next type of results demonstrates the reduction suffered by the adaptive strategy described in Algorithm 7; they show the average of the reduced dimension of the MOR in each step.

The decrement of reduced dimensions, related on to initial guess, in Graphs 4A and 4C, changes completely the selection of the fields described in Graphs 3A and 3C. However, the adaptation to the new local Galerkin projections does not change considerably the number of the PCG iterations, as demonstrated in Graphs 1A and 1C. Table 5 shows the total average of the reduced dimension from the examples.

In the next type of results, the norm from initial guess error is not analyzed; the choice is made by the major instability, gives by the correction in Algorithm 6, and low changing from the previous results. The next type of results shows the evolution of the selected fields during the internal iterations of the IPM algorithm; the criterion of the selection is the Galerkin error, as described in (140).

Analyzing Graphs 5A, 5B, 5C, and 5D, it is possible to realize that the error related to the primal field is lower during the most part of the simulation. In the middle of the simulation, it is the only part that the error of the dual field becomes
Reduced dimension - average

|                | Ex. 1 - Primal | Ex. 1 - Dual | Ex. 2 - Primal | Ex. 2 - Dual |
|----------------|---------------|--------------|---------------|--------------|
| MOR - $\mathbf{x}$ and $\lambda$ - Galerkin | 17.8          | 18.9         | 18.8          | 17.6         | 18.9         | 18.9         | 18.9         |
| MOR - $\mathbf{x}$ and $\lambda$ - Init. Guess | 17.8          | 18.9         | 18.9          | 17.7         | 18.9         | 18.9         | 18.9         |

Abbreviations: MOR, model order reduction.

TABLE 5 Average of the reduced dimensions

It is possible to realize, in Graphs 6A, 6B, 6C, and 6D, that the value of 1E-3 results in a faster convergence (the number of this value is reduced related to the value of 1 and 1E-6) and consequently reduces the error of the dual field. However, the value of 1E-6 in the final part of the simulation is critical in this case; in the IPM algorithm, a small value of the barrier parameter, at the end of the simulation, increases the condition number of the system. Moreover, an ill-conditioned system induces a poor convergence to the PCG, and a poor initial guess (higher Galerkin error-damages, even more, the situation. Table 6 shows the accumulated selection of the field according to the barrier parameter.

Finally, the last type of results shows the PCG iterations per internal iteration of the IPM algorithm:

As expected, as the final part of the simulation, as shown in Graphs 7A, 7B, 7C, and 7D, the values of the PCG iterations increase considerably. Furthermore, it is possible to notice that the number of the PCG iterations in the dual examples reduces dramatically when the primal field is selected, specially, in the last half of simulation (Graph 5B against Graph 7B...
Related to the primal examples, no relevant reduction of PCG iterations is noticed during the simulations, not even in the middle of the simulation (the lowest error of the dual field); an explanation is that the barrier parameter is not small enough to result in an ill-conditioned system, and consequently, a better initial guess is not so relevant to reduce the number of PCG iterations.

Until now, at least in these two examples, the different behaviors from both Schur complements are clear. The part of the explanation can be found in Section 4, through the error and sensitivity analysis. Subsection 4.1, specifically in (91), shows that the error of the primal field is not directly dependent on the KKT condition, suffering a less impact of the barrier parameter selection, unlike the dual field, showed in subsection 4.2, specifically in (105), which a direct dependency of the error with KKT condition is verified. The influence of the KKT condition is even more evident at the end of simulation, when low values of barrier parameters cause a slower convergence; because of that, no relevant reduction in the PCG iteration is noticed in the first half of the simulation.

Nonetheless, it is important to highlight in these two examples that, despite an error analysis and a sensibility analysis, they are not enough to ensure that the selection of the dual field always results in higher error and lower convergence; however, it is possible to conclude that there exist higher risk of high error and low convergence in the dual and Graph 5D against Graph 7D). Related to the primal examples, no relevant reduction of PCG iterations is noticed during the simulations, not even in the middle of the simulation (the lowest error of the dual field); an explanation is that the barrier parameter is not small enough to result in an ill-conditioned system, and consequently, a better initial guess is not so relevant to reduce the number of PCG iterations.
field, if compared to the primal field, because of the direct dependency of the empirical selection from the barrier parameter.

7 | CONCLUSION

In this work, it possible to realize that the combination of the Galerkin projection in one field, applied to the infeasible primal-dual interior point Method, can be simplified with the approach of the Schur complement. However, both Galerkin projections still have, even in different levels, an error dependency of the KKT conditions. Nevertheless, a MOR using Galerkin projection can be applied with a considerable improvement of the processing time if some approaches are made.

1. As demonstrated in both examples, in problems with explicit approximation convex functions, the adaptive strategy using the selection of the field can generate a considerable reduction in the processing time in a simulation with the Schur complement in the dual form. This approach has a broad application in topology optimization, especially in problems with a great number of inequality constraints. In the case of the topology optimization, the use of explicit approximation convex functions result in diagonal Hessians from the objective function; the calculus of the inverse hessian to apply in the Schur complement in the dual form becomes extremely cheap.

2. In problems with only differentiable convex objective functions, where the diagonal Hessian is not possible, an interesting approach can be the Galerkin projection only in the primal field. Although the simulation of this kind of problem is not showed in this work, all the formulations presented here, including the error and sensibility analysis, permit such approach. In the Schur complement only in the primal field, it does not require the calculation of the inverse Hessians from the objective function; the only inverse that is required comes from the KKT condition that is naturally diagonal. Moreover, as demonstrated in Section 4, the Galerkin error from the primal field is less dependent from the...
barrier parameter selection. Because of these arguments, the processing time from the MOR based in Galerkin projection only in the primal field can result in some reduction of time, if compared to the full model. Such an approach can be applied in some kind of problems in mechanics, as contact mechanics and plastic behavior of the material.

It is important to mention that the analysis of the MOR described in this work is much wider than what is simulated here. An analysis with different adaptive strategies to the barrier parameter variations, to improve the convergence of the IPM algorithm and reduces the MOR error, could generate another work. Furthermore, problems with minimum maximum (min-max), reduction of other fields (related to “y”, for example), and the combination of reduced fields, which the parcel of error is related to the primal and dual field as demonstrated, are only some of the possibilities to develop in the next research works. Finally, to further research works, problems with original convex function, as described in the last topic, are interesting applications to the IPM algorithm with Galerkin projection.

ACKNOWLEDGEMENTS

The first and second authors acknowledge the support by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) under grants 2016/03528-4 and 2016/04460-4, which made possible the development of part of this research. The first and third authors acknowledge the support of the Alexander von Humboldt Foundation for the Georg Forster Research Award, which made possible their stay at the University of Duisburg-Essen.

ORCID

P.S.B. Nigro https://orcid.org/0000-0001-7817-6362

REFERENCES

1. Nigro PSB, Anndif M, Teixeira Y, Pimenta PM, Wriggers P. An adaptive model order reduction with quasi-Newton method for nonlinear dynamical problems. *Int J Numer Methods Eng Nonlinear Dyn Probl*. 2015;106:740-759.
2. Nigro PSB, Anndif M, Teixeira Y, Pimenta PM, Wriggers P. An adaptive model order reduction by proper snapshot selection for nonlinear dynamical problems. *Computational Mechanics*. 2016;57:537-554.
3. Rathinam M, Petzold LR. Dynamic iteration using reduced order models: a method for simulation of large scale modular systems. *SIAM J Numer Anal*. 2002;40(4):1446-1474.
4. Xie X, Zheng H, Jonckheere S, van de Walle A, Playmbs B, Desmet W. Adaptive model reduction technique for large-scale dynamical systems with frequency-dependent damping. *Comput Methods Appl Mech Eng*. 2018;332:363-381.
5. Kaulmann S, Flemisch B, Haasdonk B, Lie K-A, Ohlberger M. The localized reduced basis multiscale method for two-phase flows in porous media. *Int J Numer Methods Eng*. 2015;102:1018-1040.
6. Kaulmann S, Ohlberger M, Haasdonk B. A new local reduced basis discontinuous Galerkin approach for heterogeneous multiscale problems. *Comptes Rendus Math*. 2011;349:1233-1238.
7. Yuan Z, Fish J. Multiple scale eigendeformation-based reduced order homogenization. *Comput Methods Appl Mech Eng*. 2009;198:2016-2038.
8. Oliver J, Caicedo M, Huespe AE, Hernandez JA, Roubin E. Reduced order modeling strategies for computational multiscale fracture. *Comput Methods Appl Mech Eng*. 2017;313:560-595.
9. Xiao D, Fang F, Buchan AG, et al. Non-linear model reduction for the Navier-Stokes equations using residual DEIM method. *J Comput Phys*. 2014;263:1-18.
10. Tromeur-Dervout D, Vassilevski Y. Choice of initial guess in iterative solution of series of systems arising in fluid flow simulations. *J Comput Phys*. 2006;219:210-227.
11. Pasetto D, Feronato M, Putti M. A reduced order model-based preconditioner for the efficient solution of transient diffusion equations. *Int J Numer Methods Eng*. 2017;109:1159-1179.
12. Carlberg K, Farhat C. An adaptive POD-Krylov reduced-order model for structural optimization. Paper presented at: 8th World Congress on Structural and Multidisciplinary Optimization; 2009; Lisbon, Portugal.
13. Zahr MJ, Farhat C. Progressive construction of a parametric reduced-order model for PDE-constrained optimization. *Int J Numer Methods Eng*. 2015;102:1111-1135.
14. Ghavamian F, Tiso P, Simone A. POD–DEIM model order reduction for strain-softening viscoplasticity. *Comput Methods Appl Mech Eng*. 2017;317:458-479.
15. Markovinovic R, Jansen JD. Accelerating iterative solution methods using reduced-order models as solution predictors. *Int J Numer Methods Eng*. 2006;68:525-541.
16. Kerfriden P, Gosselet P, Adhikari S, Bordas SPA. Bridging proper orthogonal decomposition methods and augmented Newton–Krylov algorithms: an adaptive model order reduction for highly nonlinear mechanical problems. *Comput Methods Appl Mech Eng*. 2011;200:850-866.
17. Wright MH. Why a pure primal Newton barrier step may be infeasible. *SIAM J Optim*. 1995;5(1):1-12.
The Schur complement of a matrix divided into four blocks can be obtained in two ways: to isolate $x_2$, the multiplication of the first row by $A_{21}^{-1}$ and subtracting the same related to the second row are necessary; and to isolate $x_1$, the multiplication of the second row by $A_{12}^{-1}$ and subtracting the same related to the first row are necessary. These manipulations result in

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad \text{where} \quad \begin{bmatrix} A_{11} - A_{12} A_{22}^{-1} A_{21} \\ A_{22} - A_{21} A_{11}^{-1} A_{12} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 - A_{12} A_{22}^{-1} b_2 \\ b_2 - A_{21} A_{11}^{-1} b_1 \end{bmatrix}$$

\hspace{1cm} \forall A_{11} \in \mathbb{R}^{n_1 \times n_1}, A_{22} \in \mathbb{R}^{n_2 \times n_2}, A_{12}, A_{21}^T \in \mathbb{R}^{n_1 \times n_2}, b_1, x_1 \in \mathbb{R}^{n_1}, b_2, x_2 \in \mathbb{R}^{n_2}. \quad (A1)

To simplify the number of variables, it is defined

$$\begin{align*}
\overline{A}_1 x_1 &= \overline{b}_1 \quad \text{where} \quad \overline{A}_1 = [A_{11} - A_{12} A_{22}^{-1} A_{21}] \quad \text{and} \quad \overline{b}_1 = [b_1 - A_{12} A_{22}^{-1} b_2], \\
\overline{A}_2 x_2 &= \overline{b}_2 \quad \text{where} \quad \overline{A}_2 = [A_{22} - A_{21} A_{11}^{-1} A_{12}] \quad \text{and} \quad \overline{b}_2 = [b_2 - A_{21} A_{11}^{-1} b_1].
\end{align*} \hspace{1cm} (A2)$$

\[ \forall \overline{A}_1 \in \mathbb{R}^{n_1 \times n_1}, \overline{A}_2 \in \mathbb{R}^{n_2 \times n_2}, \overline{b}_1 \in \mathbb{R}^{n_1}, \overline{b}_2 \in \mathbb{R}^{n_2}. \]
Now, considering the equivalent system as (30) defined in Section 2.3,

\[
\begin{bmatrix}
A_{11} & A_{12} & A_{13} \\
A_{21} & A_{22} & A_{23} \\
a_{31} & a_{32} & a_{33}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
= 
\begin{bmatrix}
b_1 \\
b_2 \\
b_3
\end{bmatrix}
\quad \forall \quad a_{13}, a_{31}^T \in \mathbb{R}^{n_1 \times 1}, \quad a_{23}, a_{32}^T \in \mathbb{R}^{n_2 \times 1}, \quad a_{33}, x_3, b_3 \in \mathbb{R}.
\]  
(A3)

The system described in (31) can be achieved in two parts. In the first part, the first row and the second row from (A3) being separate and manipulated, ie,

\[
\begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
= 
\begin{bmatrix}
b_1 - a_{13} x_3 \\
b_2 - a_{23} x_3
\end{bmatrix}
\quad \text{where} \quad \vec{b}_1 = [b_1 - a_{13} x_3] - A_{12} A_{22}^{-1} [b_2 - a_{23} x_3].
\]  
(A4)

Arranging \(\vec{b}_1\) to isolate \(x_3\) and replacing the terms by the terms from (36) and (34),

\[
\vec{b}_1 = [b_1 - A_{12} A_{22}^{-1} b_2] - [a_{13} - A_{12} A_{22}^{-1} a_{23}] x_3 \quad \text{where} \quad \pi_1 = [b_1 - A_{12} A_{22}^{-1} b_2], \quad \pi_{12} = [a_{13} - A_{12} A_{22}^{-1} a_{23}].
\]  
(A5)

Replacing also \(\vec{A}_1, x_1,\) and \(x_3\) by the equivalent term from (32) and replaces in (A4) results in the first row of (31), ie,

\[
\Pi_{11} \Delta x = \pi_1 - \pi_{12} \Delta z \quad \text{where} \quad \Pi_{11} = \vec{A}_1, \quad \Delta x = x_1, \quad \Delta z = x_3.
\]  
(A6)

Related to the second row, the second row and the third row from (A3) being separated and manipulated, ie,

\[
\begin{bmatrix}
a_{33} & a_{32} \\
a_{23} & A_{22}
\end{bmatrix}
\begin{bmatrix}
x_3 \\
x_2
\end{bmatrix}
= 
\begin{bmatrix}
b_3 - a_{31} x_3 \\
b_2 - A_{22} x_3
\end{bmatrix}
\quad \text{where} \quad \vec{b}_1 = [b_3 - a_{31} x_3] - a_{32} A_{22}^{-1} [b_2 - A_{22} x_3].
\]  
(A7)

Arranging \(\vec{b}_1\) to isolate \(x_1\) and replacing the terms by the terms from (37) and (35),

\[
\vec{b}_1 = [b_3 - a_{32} A_{22}^{-1} b_2] - [a_{31} - a_{32} A_{22}^{-1} A_{21}] x_1 \quad \text{where} \quad \pi_2 = [b_3 - a_{32} A_{22}^{-1} b_2], \quad \pi_{21} = [a_{31} - a_{32} A_{22}^{-1} A_{21}].
\]  
(A8)

Replacing also \(\vec{A}_1, x_1,\) and \(x_3\) by the equivalent term from (33) and replaces in (A7) results in the second row of (31), ie,

\[
\Pi_{22} \Delta z = \pi_2 - \pi_{21} \Delta x \quad \text{where} \quad \Pi_{22} = \vec{A}_1, \quad \Delta x = x_1, \quad \Delta z = x_3.
\]  
(A9)

Now, considering the equivalent system as (41) defined in Section 2.4,

\[
\begin{bmatrix}
A_{11} & O_{n_1 \times n_2} & A_{13} \\
O_{n_2 \times n_1} & A_{22} & A_{23} \\
A_{31} & A_{32} & A_{33}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
+ 
\begin{bmatrix}
o_{n_1} \\
o_{n_2} \\
o_{34}
\end{bmatrix}
= 
\begin{bmatrix}
b_1 \\
b_2 \\
b_3
\end{bmatrix}
\quad \forall \quad A_{23}, A_{32}^T \in \mathbb{R}^{n_2 \times n_1}, A_{33} \in \mathbb{R}^{n_3 \times n_3},
\]  
(A10)

Now, multiplying the first row by \(A_{31} A_{11}^{-1}\), the second row by \(A_{32} A_{22}^{-1}\), and subtract both related to the third row,

\[
\vec{A}_3 x_3 + a_{34} x_4 = \vec{b}_3 \quad \text{where} \quad \vec{A}_3 = [A_{33} - A_{31} A_{11}^{-1} A_{13} - A_{32} A_{22}^{-1} A_{23}] \in \mathbb{R}^{n_3 \times n_3}, \quad \vec{b}_3 = [b_3 - A_{31} A_{11}^{-1} b_1 - A_{32} A_{22}^{-1} b_2] \in \mathbb{R}^{n_3}.
\]  
(A11)

Replacing also \(\vec{A}_3, a_{34}, x_3,\) and \(x_4\) by the equivalent term from (43) and (44), and replaces in (A11), results in the first row of (42), ie,

\[
\Pi \Delta \lambda + a \Delta z = \pi \quad \text{where} \quad \Pi = \vec{A}_3, \quad \pi = \vec{b}_3, a_{34} = a, \Delta \lambda = x_3, \quad \Delta z = x_4.
\]  
(A12)
APPENDIX B

The rest of the operators from Section 4.1 are

\[ A_{12(x)} = I_m \]
\[ A_{13(x)} = \left( J\Psi_x [\Psi_x^T L_{xx} \Psi_x]^{-1} \Psi_x^T J + \Lambda^{-1} S \right) \]
\[ A_{14(x)} = o_m \]
\[ a_{1(x)} = \nabla_{\Delta(\delta)} I - J\Psi_x [\Psi_x^T L_{xx} \Psi_x]^{-1} \Psi_x^T l_x. \] (B1)

In Section 4.2, the rest of the operators, from the first row, are

\[ A_{11(\delta)} = - \left( L_{xx} + J^T \Psi_A [\Psi_A^T \Lambda^{-1} S \Psi_A]^{-1} \Psi_A^T J \right) \]
\[ A_{12(\delta)} = - J^T \Psi_A [\Psi_A^T \Lambda^{-1} S \Psi_A]^{-1} \Psi_A^T a \]
\[ A_{13(\delta)} = l_x + J^T \Psi_A [\Psi_A^T \Lambda^{-1} S \Psi_A]^{-1} \Psi_A^T l_{\Delta(\delta)}. \] (B2)

In the second row,

\[ A_{21(\delta)} = - \Psi_A [\Psi_A^T \Lambda^{-1} S \Psi_A]^{-1} \Psi_A^T J \]
\[ A_{22(\delta)} = - \left( \nabla_{yx} L + \Psi_A [\Psi_A^T \Lambda^{-1} S \Psi_A]^{-1} \Psi_A^T \right) \]
\[ A_{23(\delta)} = - \Psi_A [\Psi_A^T \Lambda^{-1} S \Psi_A]^{-1} \Psi_A^T a \]
\[ a_{2(\delta)} = l_y - \Psi_A [\Psi_A^T \Lambda^{-1} S \Psi_A]^{-1} \Psi_A^T l_{\Delta(\delta)}. \] (B3)

Moreover, in the third row,

\[ A_{31(\delta)} = a^T \Psi_A [\Psi_A^T \Lambda^{-1} S \Psi_A]^{-1} \Psi_A^T J \]
\[ A_{32(\delta)} = - a^T \Psi_A [\Psi_A^T \Lambda^{-1} S \Psi_A]^{-1} \Psi_A^T \]
\[ A_{33(\delta)} = - \left( a^T \Psi_A [\Psi_A^T \Lambda^{-1} S \Psi_A]^{-1} \Psi_A^T a + l_{zz} \right) \]
\[ a_{3(\delta)} = l_z - a^T \Psi_A [\Psi_A^T \Lambda^{-1} S \Psi_A]^{-1} \Psi_A^T l_{\Delta(\delta)}. \] (B4)

Related to Section 4.3, now is shown the development to achieve the two null space operators. The development starts isolating the projection value of the primal variable in (108), ie,

\[ \Delta x_{\Psi_x} = - \Psi_x [\Psi_x^T L_{xx} \Psi_x] \Psi_x^T \left( I_x + L_{xx} \Delta x_{\Omega_x} + J^T \Delta \lambda_{\Psi_x} + J^T \Delta \lambda_{\Omega_x} \right). \] (B5)

In addition, isolating the projection value of the dual variable in (109),

\[ \Delta \lambda_{\Psi_x} = \nabla_{y} I + L_{yy} \Delta y - \Delta \lambda_{\Omega_y}. \] (B6)

Replacing (B6) in (B5), it leads to

\[ \Delta x_{\Psi_x} = - \Psi_x [\Psi_x^T L_{xx} \Psi_x] \Psi_x^T \left( I_x + L_{xx} \Delta x_{\Omega_x} + J^T l_y + J^T L_{yy} \Delta y \right). \] (B7)

Now, isolating the projection value of the dual variable in (110), ie,

\[ \Delta \lambda_{\Psi_x} = \Psi_A [\Psi_A^T \Lambda^{-1} S \Psi_A]^{-1} \Psi_A^T \left( I_{\Delta(\delta)} + J \Delta x_{\Psi_x} + J \Delta x_{\Omega_x} - \Delta y - \Lambda^{-1} S \Delta \lambda_{\Omega_x} - a \Delta z \right). \] (B8)
Replacing (B8) in (B7),

\[ \Delta \lambda_{\psi} = \Psi_{\lambda} [\Psi_{\lambda}^T A^{-1} \Psi_{\lambda}]^{-1} \Psi_{\lambda}^T \left\{ b_1 + B_{11} \Delta x_\omega + B_{12} \Delta y + B_{13} \Delta \lambda_{\omega} + B_{14} \Delta z \right\} , \]  

(B9)

where

\[ b_1(x, \lambda) = L_{\lambda(\psi)} - J \Psi_{\lambda} [\Psi_{\lambda}^T L_{\omega} \Psi_{\lambda}]^{-1} \Psi_{\lambda}^T \left\{ L_k + J^T I_2 \right\} \in \mathbb{R}^m \]

\[ B_{11}(x, \lambda) = -J \Psi_{\lambda} [\Psi_{\lambda}^T L_{\omega} \Psi_{\lambda}]^{-1} \Psi_{\lambda}^T L_{\omega} + J \in \mathbb{R}^{m \times n} \]

\[ B_{12}(x, \lambda) = -J \Psi_{\lambda} [\Psi_{\lambda}^T L_{\omega} \Psi_{\lambda}]^{-1} \Psi_{\lambda}^T J^T L_{\nu} - I_m \in \mathbb{R}^{m \times n} \]  

(B10)

Finally, replacing (B9) in (111) and isolating by field,

\[ A_{11}(x, \lambda) \Delta x_\omega + A_{12}(x, \lambda) \Delta y + A_{13}(x, \lambda) \Delta \lambda_{\omega} + A_{14}(x, \lambda) \Delta z = a_{1}(x, \lambda) , \]

(B11)

where

\[ a_{1}(x, \lambda) = l_k - a^T \Psi_{\lambda} [\Psi_{\lambda}^T A^{-1} \Psi_{\lambda}]^{-1} \Psi_{\lambda}^T b_{1}(x, \lambda) \]

\[ A_{11}(x, \lambda) = a^T \Psi_{\lambda} [\Psi_{\lambda}^T A^{-1} \Psi_{\lambda}]^{-1} \Psi_{\lambda}^T B_{11}(x, \lambda) \]

\[ A_{12}(x, \lambda) = a^T \Psi_{\lambda} [\Psi_{\lambda}^T A^{-1} \Psi_{\lambda}]^{-1} \Psi_{\lambda}^T B_{12}(x, \lambda) \]

\[ A_{13}(x, \lambda) = a^T \left\{ \Psi_{\lambda} [\Psi_{\lambda}^T A^{-1} \Psi_{\lambda}]^{-1} \Psi_{\lambda}^T B_{13}(x, \lambda) + I_m \right\} \]

\[ A_{14}(x, \lambda) = a^T \Psi_{\lambda} [\Psi_{\lambda}^T A^{-1} \Psi_{\lambda}]^{-1} \Psi_{\lambda}^T B_{14}(x, \lambda) - l_{\omega z} \]  

(B12)

APPENDIX C

The problem defined in section 6.1 has the initial data per step with the following values:

\[ a_0 = 1 , \quad a = o_m , \quad c = 1000 e_m , \quad d = e_m \]

\[ \lambda^0 = e_m , \quad z^0 = 1 , \quad y^0 = e_m , \quad s^0 = o_m , \quad \mu^0 = 1 \]

\[ tol_{POD} = tol_{PCG} = tol_{p} = 10^{-7} \]

where the barrier parameter variation is

\[ \mu^{i+1} = 0.001 \mu^i . \]  

(C2)

Furthermore, the normalization of the functions is defined as

\[ \tilde{f}(x) = f(x) \left\| \frac{\partial f}{\partial x} \right\|_2^{-1} \quad \text{and} \quad \tilde{g}_i(x) = g_i(x) \left\| \frac{\partial g_i}{\partial x} \right\|_2^{-1} \]  

(C3)

In the next step, the problems start with the same values (only x starts with the value of the previous step). The values of the matrices \( W, P, \) and \( Q \) are

\[ w_{i,j} = \frac{2 + \sin \left( 4 \pi a_{i,j} \right)}{\left| 1 + |i - j| \right| \ln n} , \quad p_{i,j} = \frac{1 + 2 a_{i,j}}{\left| 1 + |i - j| \right| \ln n} , \quad q_{i,j} = \frac{3 - 2 a_{i,j}}{\left| 1 + |i - j| \right| \ln n} . \]  

(C4)

These matrix values are the same as the classical example defined in the work of Svanberg. The values from the matrices \( R^k \) and \( T^k \) defined in this work are

\[ r_{i,j,k} = \frac{\cos \left( 2 \pi \left[ 2n \exp \left( a_{i,j} \right) - m \beta_j \right] \right)}{\left| 1 + |i - j| \right| \ln n} , \quad t_{i,j,k} = \frac{\cos \left( 2 \pi \left[ 3m \gamma_{ik} - n \exp \left( a_{i,j} \right) \right] \right)}{\left| 1 + |i - j| \right| \ln n} . \]  

(C5)
Also defined in this work, the limit values of the constraints functions are

\[
\begin{align*}
    b_1 &= \left[ \frac{1}{m_1} \sum_{k=1}^{m_1} \beta_k \right] \left[ \frac{n + m}{8} \right], \quad b_2 = \left[ \frac{1}{m_2 - m_1} \sum_{k=m_1+1}^{m_2} \gamma_k \right] \left[ \frac{n + m}{8} \right], \\
    b_3 &= \left[ \frac{1}{m_3 - m_2} \sum_{k=m_1+1}^{m_3} \beta_k \right] \left[ \frac{n + m}{8} \right], \quad b_4 = \left[ \frac{1}{m_3 - m_2} \sum_{k=m_1+1}^{m_3} \gamma_k \right] \left[ \frac{n + m}{8} \right].
\end{align*}
\] (C6)

Finally, the matrix values defined in (C4), (C5), and (C6) have the support variables \(a^{ij}, \beta^{ij}\), and \(\gamma^{ij}\) defined as

\[
\begin{align*}
    a_{ij} &= \frac{i + j - 2}{2n - 2}, \quad \beta_k = \exp \left( \frac{k}{m} \frac{1}{4} - \frac{k}{m} \right), \quad \gamma_k = \exp \left( \frac{k}{m} \frac{3}{4} - \frac{k}{m} \right), \\
    \forall \quad a_{ij}, \beta_k, \gamma_k \in \mathbb{R}.
\end{align*}
\] (C7)

where the variable \(a^{ij}\) is also defined in the work of Svanberg. For a better reproducibility of the examples, the results of both examples are shown in Table C1.

**APPENDIX D**

Sections 6.2 and 6.3 described the examples 1 and 2, respectively. As described previously, both examples use the explicit convex approximation functions; in this case, the MMA, as described in the works of Svanberg and Bruyneel et al. The basic idea of the MMA is a convex function limited by an inferior and superior asymptote. To describe the MMA, it is necessary to describe the following operators:

\[
\begin{align*}
    X_i &= \text{diag}(x - x_i) \in \mathbb{R}^{n \times n}, \quad X_u = \text{diag}(x_u - x) \in \mathbb{R}^{n \times n}, \\
    X_i^{\text{initial}} &= \text{diag}(x^{\text{initial}} - x_i) \in \mathbb{R}^{n \times n}, \quad X_u^{\text{initial}} = \text{diag}(x_u - x^{\text{initial}}) \in \mathbb{R}^{n \times n}, \\
    \forall \quad x_i, x_u, x^{\text{initial}} \in \mathbb{R}^n.
\end{align*}
\] (D1)

In (D1), vectors \(x_i\) and \(x_u\) represent the inferior and superior limit, respectively; vector \(x^{\text{initial}}\) represents the initial value of the primal variable in the step. Define the operators that contain the asymptotes, the objective function by the MMA can be described as

\[
\begin{align*}
    \bar{f} &= f(x^{\text{initial}}) + f_u^T \left[ \left( X_u \right)^{-1} - \left( X_u^{\text{initial}} \right)^{-1} \right] e_n + f_l^T \left[ \left( X_l \right)^{-1} - \left( X_l^{\text{initial}} \right)^{-1} \right] e_n \forall \quad \bar{f} \in \mathbb{R} \\
    f_u &= \left( X_u^{\text{initial}} \right)^2 \nabla x^+ \in \mathbb{R}^n, \quad f_l = \left( X_l^{\text{initial}} \right)^2 \nabla x^- \in \mathbb{R}^n, \quad \nabla \bar{f} = \nabla x^+ + \nabla x^- \forall \quad \nabla x^+ \in \mathbb{R}^n_+, \quad \nabla x^- \in \mathbb{R}^n_-
\end{align*}
\] (D2)

Moreover, the constraint function can be defined as

\[
\begin{align*}
    \bar{g} &= g^{\text{initial}} + J_u \left[ \left( X_u \right)^{-1} - \left( X_u^{\text{initial}} \right)^{-1} \right] e_n + J_l \left[ \left( X_l \right)^{-1} - \left( X_l^{\text{initial}} \right)^{-1} \right] e_n \forall \quad \bar{g}, g^{\text{initial}} \in \mathbb{R}^m \\
    J_u &= \left( J^+ \right) \left( X_u^{\text{initial}} \right)^2 \in \mathbb{R}^{m \times n}, \quad J_l = \left( J^- \right) \left( X_l^{\text{initial}} \right)^2 \in \mathbb{R}^{m \times n}, \quad J = J^+ + J^- \forall \quad J^+ \in \mathbb{R}^{m \times n}_+, \quad J^- \in \mathbb{R}^{m \times n}_-
\end{align*}
\] (D4)

The super index with “+” or “−” represents only values higher and lower than zero, respectively; otherwise, the value of the vector or matrix is fulfilled with zero. Defining the functions, the gradients of (D2) and (D4) related to the primal variable are

\[
\begin{align*}
    \nabla \bar{f} &= \left( X_u \right)^{-2} f_u - \left( X_l \right)^{-2} f_l \quad \forall \quad \nabla \bar{f} \in \mathbb{R}^n \\
\end{align*}
\] (D6)

| Example results – 20 steps | Ex. 1 - Primal | Ex. 1 - Dual | Ex. 2 - Primal | Ex. 2 - Dual |
|----------------------------|---------------|-------------|---------------|-------------|
| Objective function value   | 371.774       | 369.028     | 279.769       | 279.771     |
| Respected constraints      | 1000          | 1000        | 2000          | 2000        |

**TABLE C1** Objective function values and number of respected constraints
\[ \mathbf{J} = \mathbf{J}_u \mathbf{X}_u^{-2} - \mathbf{J}_l \mathbf{X}_l^{-2} \quad \forall \mathbf{J} \in \mathbb{R}^{m \times n}. \tag{D7} \]

The Hessians of (D2) and (D4) related to the primal variable are
\[
\nabla_{xx}^2 \mathbf{F} = \text{diag} \left( 2 \left\{ \mathbf{X}_u^{-3} \mathbf{f}_u + \mathbf{X}_l^{-3} \mathbf{f}_l \right\} \right) \in \mathbb{R}^{n \times n}. \tag{D8} \]

Because of the property of the explicit function, the crossed terms are nulls; the result is a diagonal matrix. In the Hessian of (D4), it becomes a third-order tensor; again, the property of the explicit function permits, when operated with the dual variable vector, a diagonal matrix, i.e.,
\[
\mathbf{g}^T \lambda = \text{diag} \left( 2 \left\{ \mathbf{X}_u^{-3} \mathbf{J}_u^T + \mathbf{X}_l^{-3} \mathbf{J}_l^T \right\} \lambda \right) \in \mathbb{R}^{n \times n}. \tag{D9} \]

As described in the work of Svanberg,\textsuperscript{25} the heuristic part of this version of the MMA is the definition of the limit values applied to the asymptotes; they are
\[
\begin{align*}
\mathbf{x}_l^k &= \max \left( \mathbf{x}^{k-1} - 0.5 \left( \mathbf{x}_{\max} - \mathbf{x}_{\min} \right) \mathbf{x}_{\min} \right) \quad \text{if} \quad k \leq 2, \\
\mathbf{x}_u^k &= \min \left( \mathbf{x}^{k-1} + 0.5 \left( \mathbf{x}_{\max} - \mathbf{x}_{\min} \right) \mathbf{x}_{\max} \right)
\end{align*} \tag{D10} \]

In addition, from the third step,
\[
\begin{align*}
\mathbf{x}_l^k &= \max \left( \mathbf{x}^k - \text{rate} \left[ \mathbf{x}^{k-1} - \mathbf{x}_{l}^{k-1} \right] \mathbf{x}_{l}^{\min} \right), \\
\mathbf{x}_u^k &= \min \left( \mathbf{x}^k + \text{rate} \left[ \mathbf{x}_{u}^{k-1} - \mathbf{x}^{k-1} \right] \mathbf{x}_{u}^{\max} \right) \\
\text{rate} &= 0.7 \quad \text{if} \quad \left[ \mathbf{x}^k - \mathbf{x}^{k-2} \right] < \left[ \mathbf{x}^{k} - \mathbf{x}^{k-1} \right] \quad \text{and} \quad k > 2, \tag{D11} \\
\text{rate} &= 1.2 \quad \text{if} \quad \left[ \mathbf{x}^k - \mathbf{x}^{k-2} \right] > \left[ \mathbf{x}^{k} - \mathbf{x}^{k-1} \right] \quad \text{and} \quad k > 2.
\end{align*} \]

Furthermore, the inferior and superior limits of the barrier functions are defined by
\[
\begin{align*}
\alpha^k &= \max \left( \mathbf{x}_l^k + 0.1 \left[ \mathbf{x}^k - \mathbf{x}_l^k \right] \mathbf{x}_l^k \right), \\
\beta^k &= \min \left( \mathbf{x}_u^k - 0.1 \left[ \mathbf{x}^k - \mathbf{x}_u^k \right] \mathbf{x}_u^k \right). \tag{D12} \]

The intention is keeping the relation
\[
\mathbf{x}_{\min} \leq \mathbf{x}_l \leq \mathbf{\alpha} < \mathbf{x} < \mathbf{\beta} \leq \mathbf{x}_u \leq \mathbf{x}_{\max}. \tag{D13} \]

It is important to highlight again that all the description of the MMA is to facilitate the reproduction of the examples 1 and 2 from Section 6. The formulation presented in this work is general; this means that, theoretically, it can be applied to any kind of convex function.