

Robustesse de l'algorithme d'identification Data-Driven vis-à-vis de données parcellaires

M. Dalémat, M. Coret, A. Leygue, E. Verron

Institut de Recherche en Génie Civil et Mécanique (GeM), UMR CNRS 6183, École Centrale de Nantes, France, {prénom.nom}@ec-nantes.fr

Résumé — Ce travail s'intéresse à l'influence des paramètres d'entrée de l'algorithme d'Identification Data-Driven (DDI, Leygue *et al.* (Comput. Mech. Appl. Mech. Engng., **331**, 184-196 [1]) sur les résultats obtenus. Cet algorithme permet de mesurer des contraintes, sans lois de comportement, à partir de champs de déplacement et des forces appliquées à une structure. Dans le cas d'expériences réelles, les données sont toujours incomplètes ; mais nous montrons ici qu'avec un traitement approprié des données expérimentales, il est possible d'identifier de façon robuste les champs de contrainte.

Mots clés — Identification data-driven, corrélation d'images numériques, mesures de contraintes.

1 Introduction

Les lois de comportement font l'objet de nombreux travaux, essentiels pour des simulations robustes. L'utilisation pratique de ces lois nécessite tout d'abord de choisir ou de concevoir un modèle de comportement qui soit représentatif de la réponse mécanique du matériau. Ensuite, l'identification de ses paramètres doit se faire idéalement en considérant plusieurs états de déformation, ce qui rend la mise en œuvre expérimentale difficile. Avec l'évolution des techniques de mesures de champs comme la corrélation d'images numériques (DIC[2]) et des *data sciences*, les méthodes d'identification ne cessent d'être améliorées (par exemple [3] pour un aperçu des techniques d'identification ou bien [4] pour un exemple d'application de réseaux de neurones pour l'identification).

C'est un chemin différent qui est choisi ici : se passer complètement de lois de comportement. En effet, il est possible d'utiliser ces outils (méthodes de champs et *data sciences*) pour créer des bases de données « riches » permettant d'identifier et de simuler sans le filtre d'une loi de comportement, évitant ainsi les difficultés mentionnées précédemment. Ici, un algorithme particulier est étudié, nommé Data Driven Identification (DDI) et développé par Leygue et al. dans [1]. Il permet d'identifier des champs de contrainte à partir de la mesure de champs de déplacement hétérogènes et de forces extérieures. L'algorithme a été validé sur des données synthétiques dans [1]. L'application à des données réelles a également été mise en œuvre dans [5], permettant une mesure originale de champs de contrainte hétérogènes sans lois de comportement à partir de champs de déplacement mesurés par DIC et de la résultante des forces de traction mesurée par une cellule de force. La difficulté dans l'application de l'algorithme DDI à des données réelles réside principalement dans le manque de données dans certaines zones, inhérent à un « vrai » problème. En effet, contrairement à un problème synthétique où tout est parfaitement connu, on ne peut mesurer ni toutes les forces, ni tous les déplacements ; ces difficultés ont été levées en faisant certains choix (à la fois sur les paramètres de l'algorithme et sur les données expérimentales).

Le présent travail propose une étude paramétrique de ces choix afin de démontrer la robustesse de la DDI appliquée à de vraies données. On fera, en premier lieu, un bref rappel de l'algorithme pour en rappeler les paramètres et les données d'entrée optimales, et des conditions expérimentales pour montrer qu'elles conduisent à des données d'entrée incomplètes. Ensuite, sera analysée l'influence (i) des paramètres intrinsèques à l'algorithme et (ii) des entrées incomplètes sur les résultats de la DDI. Pour cette étude, on utilisera des données synthétiques pour lesquelles la réponse est parfaitement connue. Une altération des données synthétiques est donc réalisée pour les rapprocher de données expérimentales réelles, c'est-à-dire incomplètes.

2 Algorithme et cas d'application

On rappelle ici la mise en œuvre de la DDI avec l'algorithme d'origine et son application à des données réelles.

2.1 Identification Data Driven

La méthode Data Driven Identification (DDI) [1] correspond à la méthode inverse de Data Driven Computational Mechanics (DDCM) initiée par Kirchdoerfer et Ortiz dans [6].

Cette méthode identifie la réponse complète d'une structure sans passer par une loi de comportement, à partir d'une large base de données, dans le cadre des transformations élastiques (non linéaires) en petites déformations. L'extension en grandes déformations se fait naturellement dans un cadre Eulérien comme expliqué dans [5].

Données d'entrée. On considère une géométrie 2D, que l'on choisit de mailler, sollicitée en plusieurs incréments dénotés \cdot^X . Pour cette géométrie, les données suivantes sont disponibles :

- (I-1) les déplacements nodaux \mathbf{u}_j^X , \cdot_j étant le numéro du nœud. L'information utilisée à partir des déplacements est une mesure de déformation. Ici, on adopte la déformation vraie d'Henky $\ln \mathbf{v}$, extension naturelle pour une formulation Eulérienne, définie à partir de \mathbf{b} , le tenseur de Cauchy-Green gauche, par :

$$\ln \mathbf{v} = \frac{1}{2} \ln \mathbf{b}; \quad (1)$$

- (I-2) la matrice \mathbf{B}_{ej}^X qui encode la géométrie et la connectivité, \cdot_e étant le numéro du point d'intégration de Gauss. En particulier, on peut calculer sur chacun des nœuds l'équilibre mécanique :

$$\sum_e w_e^X \mathbf{B}_{ej}^X \cdot \boldsymbol{\sigma}_e^X = \mathbf{f}_j^X \quad \forall X, j, \quad (2)$$

où w_e est le poids d'intégration ;

- (I-3) les forces nodales \mathbf{f}_j^X ;

- (I-4) les déplacements prescrits (en particulier les indices des nœuds concernés).

Données recherchées. Pour résoudre le problème mécanique, il faut :

- (O-1) les champs mécaniques $\boldsymbol{\sigma}_e^X$ qui satisfont l'équilibre mécanique en chaque élément e selon l'équation (2). Les contraintes $\boldsymbol{\sigma}_e^X$ (calculées) et les déformations $\ln \mathbf{v}_e^X$ (mesurées) constituent un état mécanique car elles sont mécaniquement admissibles (équilibrés et compatibles) ;
- (O-2) les états matériaux $(\ln \mathbf{v}_i^*, \boldsymbol{\sigma}_i^*)$, qui sont au nombre de N^* , un nombre choisi. Ces états matériaux peuvent être interprétés comme un échantillonnage de la réponse du matériau. Leur distance aux états mécaniques est défini par une norme énergétique $\|\cdot\|_{\mathbb{C}}^2$ donnée par (3) où \mathbb{C} est un tenseur du quatrième ordre également choisi.

En pratique. L'algorithme repose sur la résolution de ce problème : trouver les états matériaux les plus proches possibles des états mécaniques (ces derniers étant à moitié connus grâce aux déformations, et à moitié contraints par l'équation d'équilibre appliquée aux contraintes), selon la norme énergétique $\|\cdot\|_{\mathbb{C}}^2$ définie par :

$$\|(\ln \mathbf{v}, \boldsymbol{\sigma})\|_{\mathbb{C}}^2 = \frac{1}{2} (\ln \mathbf{v} : \mathbb{C} : \ln \mathbf{v} + \boldsymbol{\sigma} : \mathbb{C}^{-1} : \boldsymbol{\sigma}). \quad (3)$$

La solution du problème est définie comme suit :

$$\text{solution} = \arg \min_{\boldsymbol{\sigma}_e^X, \ln \mathbf{v}_e^*, \boldsymbol{\sigma}_e^*} \mathcal{E}(\boldsymbol{\sigma}_e^X, \ln \mathbf{v}_e^*, \boldsymbol{\sigma}_e^*), \quad (4)$$

avec,

$$\mathcal{E}(\boldsymbol{\sigma}_e^X, \ln \mathbf{v}_{e^X}^*, \boldsymbol{\sigma}_{e^X}^*) = \sum_X \sum_e w_e^X \|(\ln \mathbf{v}_e^X - \ln \mathbf{v}_{e^X}^*, \boldsymbol{\sigma}_e^X - \boldsymbol{\sigma}_{e^X}^*)\|_{\mathbb{C}}^2, \quad (5)$$

et sous contraintes :

$$\left\{ \begin{array}{l} \text{de respecter l'équilibre, donné en (2),} \\ \text{que l'état matériaux } (\ln \mathbf{v}_{e^X}^*, \boldsymbol{\sigma}_{e^X}^*) \text{ associé à l'élément } e^X \text{ appartienne à la base de données } (\ln \mathbf{v}_i^*, \boldsymbol{\sigma}_i^*)_{i=1}^{N^*}. \end{array} \right. \quad (6)$$

Les sorties du problème sont donc :

- les états mécaniques ;
- la base de données des états matériaux ;
- le nuage de points (ou le *mapping*) d'états mécaniques associés à chaque état matériaux.

Dans le travail initial [1], la validité de la méthode a été démontrée avec des données synthétiques complètes (I-1) à (I-4). Les deux paramètres intrinsèques à la méthode sont donc le nombre d'états matériaux N^* et \mathbb{C} définissant la norme énergétique.

2.2 Définition du cas d'étude

Problème théorique. On applique la méthode précédente à une membrane hyperélastique perforée soumise à une traction uniaxiale. Le problème mécanique et les différentes notations sont fournies sur la Figure 1. L'altération des données synthétiques se fait au regard des conditions expérimentales usuelles : on ne peut pas mesurer les forces nodales mais une résultante ; les déplacements sont parfois manquants dans des zones appelées *clusters*.

Modification des données théoriques pour simuler des données expérimentales incomplètes. Avec de vraies données, il convient de réécrire le problème mécanique en considérant les zones où des données sont manquantes [5]. Suivant les données manquantes, plusieurs possibilités sont proposer pour traiter les zones concernées.

- Dans la zone proche du mors, où la force est mesurée :
 - (m-1) Dans le cas synthétique, on connaît chaque force nodale \mathbf{f}_j^X à la frontière avec le mors Γ_m ;
 - (m-2) Dans le cas réel, on ne connaît que la somme des forces \mathbf{f}_j^X sur la frontière Γ_m dans la direction de sollicitations \mathbf{n}_{sol} :

$$\sum_{j \in \Gamma_m} \mathbf{f}_j^X \cdot \mathbf{n}_{\text{sol}} = F_{\text{cell}}^X. \quad (7)$$

On peut donc définir une condition d'équilibre sur la frontière :

$$\sum_{j \in \Gamma_m} \sum_e w_e^X \mathbf{B}_{ej}^X \cdot \boldsymbol{\sigma}_e^X \cdot \mathbf{n}_{\text{sol}} = F_{\text{cell}}^X \quad \forall X; \quad (8)$$

- Pour les *clusters* de déplacements manquants, on ne peut pas avoir accès à des informations à l'intérieur de ces zones ; on ne considère que la frontière des *clusters* Γ_c :
 - (c-1) Une solution simple et naïve consiste à considérer cette frontière comme une frontière où s'appliquent des conditions aux limites de Dirichlet, c'est-à-dire à imposer les déplacements sans connaître les forces et à ne pas considérer l'équation d'équilibre pour ces nœuds ;
 - (c-2) Une solution mécaniquement optimale consiste à considérer que la frontière des *clusters* est le frontière d'un sous-ensemble équilibré mécaniquement. En effet, la matière du *cluster* est en équilibre, ce qui revient à imposer une condition d'auto-équilibre sur Γ_c :

$$\sum_{j \in \Gamma_c} \sum_e w_e^X \mathbf{B}_{ej}^X \cdot \boldsymbol{\sigma}_e^X = \mathbf{0}. \quad (9)$$

2.3 Données d'entrée à étudier

Les données d'entrée de l'algorithme sont :

- Les paramètres de l'algorithme (intrinsèques à la méthode de résolution) : N^* et \mathbb{C} ;
- Les données mesurées, en particulier les déplacements et les forces, qui sont incomplètes.

On commencera par analyser l'effet des paramètres intrinsèques sur un cas où les données d'entrée sont complètes. Ensuite, on comparera l'influence des données mesurées dans les cas (m-1) et (m-2) relatifs au mors, puis dans les cas (c-1) et (c-2) relatifs aux *clusters* de données manquantes.

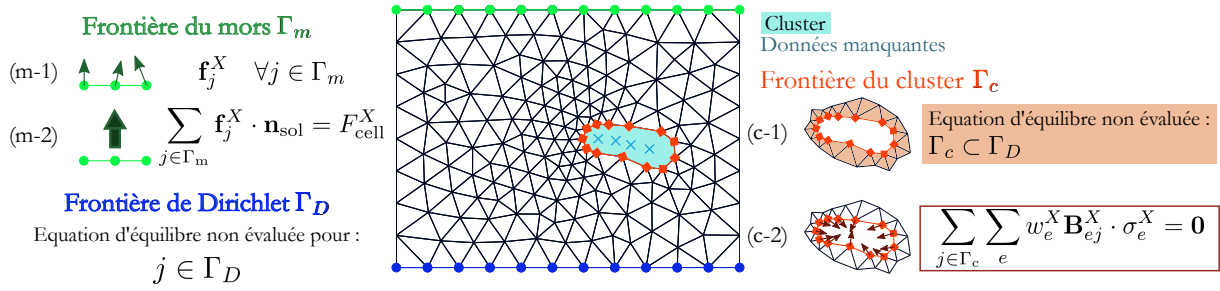


FIGURE 1 – Problème théorique traité avec 3 frontières particulières (mors, *cluster* et Dirichlet). Les bords libres ne sont pas considérés car l'équation d'équilibre s'y applique normalement. On y représente les différentes données d'entrée qui seront analysées.

3 Étude paramétrique

Les maillages non-déformés et déformés du problème sont présentés sur la Figure 2. Premièrement,

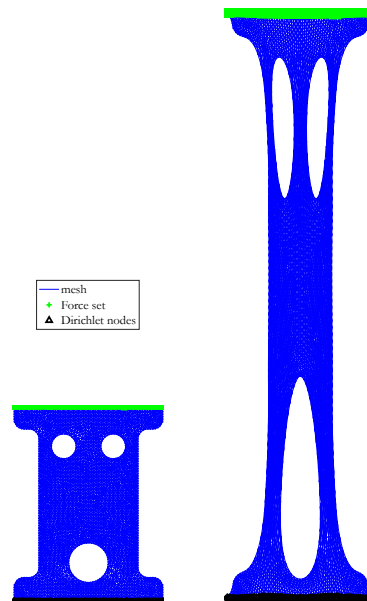


FIGURE 2 – Membrane hyperélastique trouée sollicitée en traction (200% de déformation) : états initial et final. Tous les nœuds colorés (verts et noirs) sont bloqués en déplacement, mais sur les nœuds verts une information sur la force est disponible.

les contraintes dans chaque élément σ_{FE} et les forces résultantes sont calculées par méthode éléments finis classique. On applique ensuite la méthode DDI avec les paramètres et les données d'entrée tels que décrits précédemment ce qui permet d'obtenir des contraintes identifiées σ_{DDI} . On calcule ensuite l'erreur entre les contraintes identifiées par la DDI et les contraintes de référence calculées par éléments finis, pour chaque incrément de chargement \cdot^X (ici, $N_X = 21$) et chaque élément \cdot_e (ici au nombre de $N_e = 6108$) :

$$\text{err} = \frac{\sum_{X,e} \|\sigma_{\text{FE},e^X} - \sigma_{\text{DDI},e^X}\|_2}{\sum_{X,e} \|\sigma_{\text{FE},e^X}\|_2}. \quad (10)$$

3.1 Influence des paramètres intrinsèques à l'algorithme

Le nombre d'états matériaux N^* est le paramètre qui permet d'échantillonner plus ou moins finement le comportement du matériau à partir des états mécaniques. Il est à comparer au nombre total de degrés de liberté du problème : $N_e \times N_X = 128\,268$. On peut définir le ratio $r^* = (N_e \times N_X) / N^*$ et on le fait varier entre 2 et 10^4 . La distance aux états mécaniques $(\varepsilon_e^X, \sigma_e^X)$ est définie par la norme énergétique $\|\cdot\|_{\mathbb{C}}^2$ de

l'Eq. (3), où \mathbb{C} est un tenseur du quatrième ordre. La forme la plus simple de ce tenseur consiste à choisir un tenseur sphérique défini par son amplitude C , qui peut s'écrire, en adoptant la notation de Voigt :

$$[\mathbb{C}] = C \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (11)$$

Cette forme vise à privilégier la convergence de toutes les composantes à la fois. L'amplitude du tenseur est définie relativement au module d'élasticité tangent choisi pour le modèle de comportement utilisé dans les simulations éléments finis : $C_0 = 2.5 \cdot 10^6$ Pa. Ainsi, on choisit des valeurs de C variant de C_0 à $10^6 C_0$.

La figure 3 présente les résultats de l'erreur après convergence. La sous-figure de gauche 3 montre

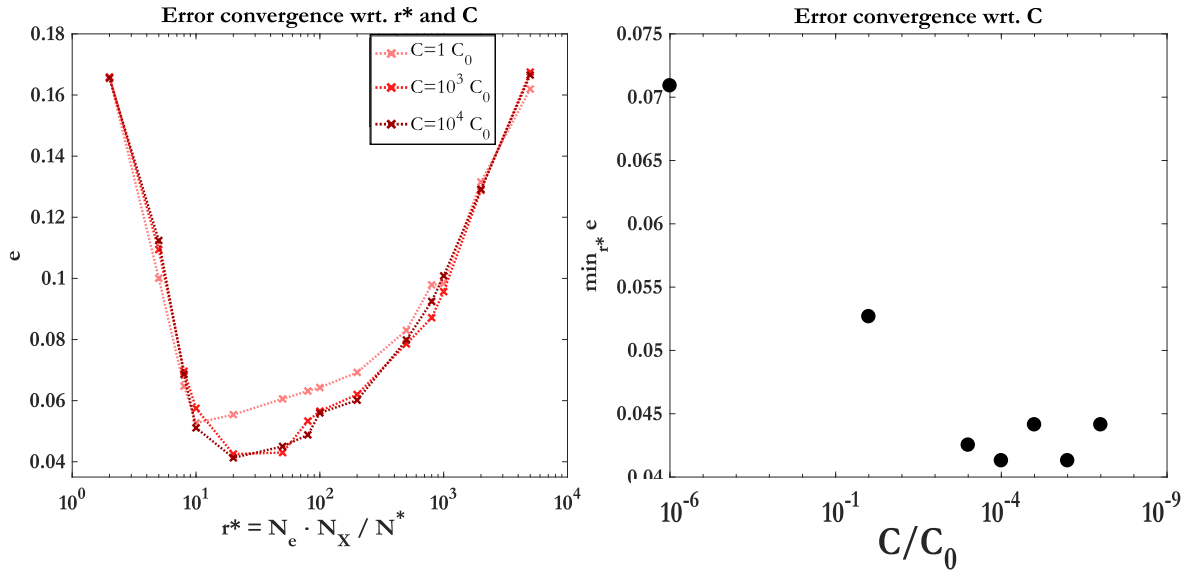


FIGURE 3 – Influence des paramètres intrinsèques N^* (lié à r^*) et C dans un cas sans données manquantes. On compare l'erreur par rapport à r^* , et l'erreur minimale par rapport à C .

l'erreur de convergence en fonction du ratio r^* pour différentes valeurs de C . Pour chaque valeur de C , on reporte la valeur de l'erreur minimale par rapport à r^* et on présente, sur la sous-figure de droite, l'erreur minimale (pour le r^* optimal) en fonction de C/C_0 . L'erreur s'avère minimale pour un $r^* \approx 20$ (10 à 50 en fonction de la valeur de C). Un ratio trop important (trop peu d'états matériaux) implique un sous-échantillonnage de la réponse et donc une erreur importante. Réciproquement, un trop faible ratio (trop d'états matériaux) produit un effet comparable à l'*overfitting* en régression : on ne « moyenne » plus assez le comportement, ce qui conduit aussi à une erreur importante. Il convient donc de se placer entre ces deux extrêmes ; ce sont des résultats similaires à ceux obtenus dans [1]. De plus, C contribue de manière importante à la convergence : plus on l'augmente (jusqu'à un certain point), plus l'erreur est faible. Ceci est simplement expliqué par le fait que la distance définie par la norme influe sur l'appariement entre les états matériaux et les états mécaniques. En choisissant une grande valeur de C , l'appariement sur les déformations est privilégié, ce qui est intéressant puisqu'elles sont mesurées (et donc fiables), contrairement aux contraintes qui évoluent pendant la convergence de l'algorithme. On remarque finalement que le choix de N^* est plus influent que celui de C : sans données manquantes, un « mauvais choix » de N^* ne sera jamais compensé par un « bon » C .

3.2 Influence des données d'entrée incomplètes

Dans un premier temps, on considère les forces : soit avec l'ensemble des forces nodales ($m-1$), soit moyennées par leur résultante sur la frontière ($m-2$) : l'influence de N^* sur l'erreur est fournie sur la figure 4 (on choisit de représenter les résultats pour une valeur de C grande). Les courbes d'erreur sont sensiblement identiques : un ratio d'états matériaux par rapport au nombre de degrés de liberté de 20 à 50 est préférable. Ces résultats montrent que l'algorithme DDI n'est que faiblement influencé par la façon

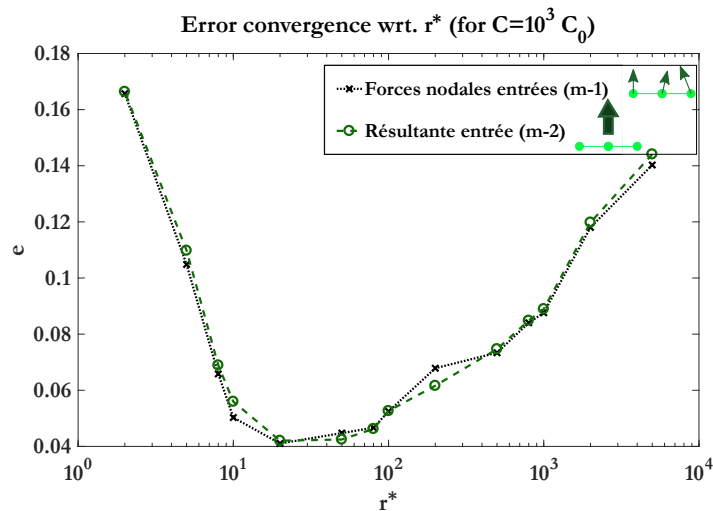


FIGURE 4 – Influence des forces entrées sur l’erreur en fonction de r^* des forces nodales données et de la résultante donnée.

d’imposer les conditions aux limites de type force imposée : les contraintes sont fournies par l’équation d’équilibre et sont définies à une constante multiplicative près, levée par l’information de force. Dans un cas simple (symétrique et uniaxiale), l’information moyennée F_{cell}^X est suffisante pour retrouver les contraintes.

Ensuite, on observe l’influence des données d’entrée liées à un *cluster* : en le traitant de manière naïve et de manière mécaniquement optimale. L’influence de N^* et C est présentée sur la figure 5. Dans

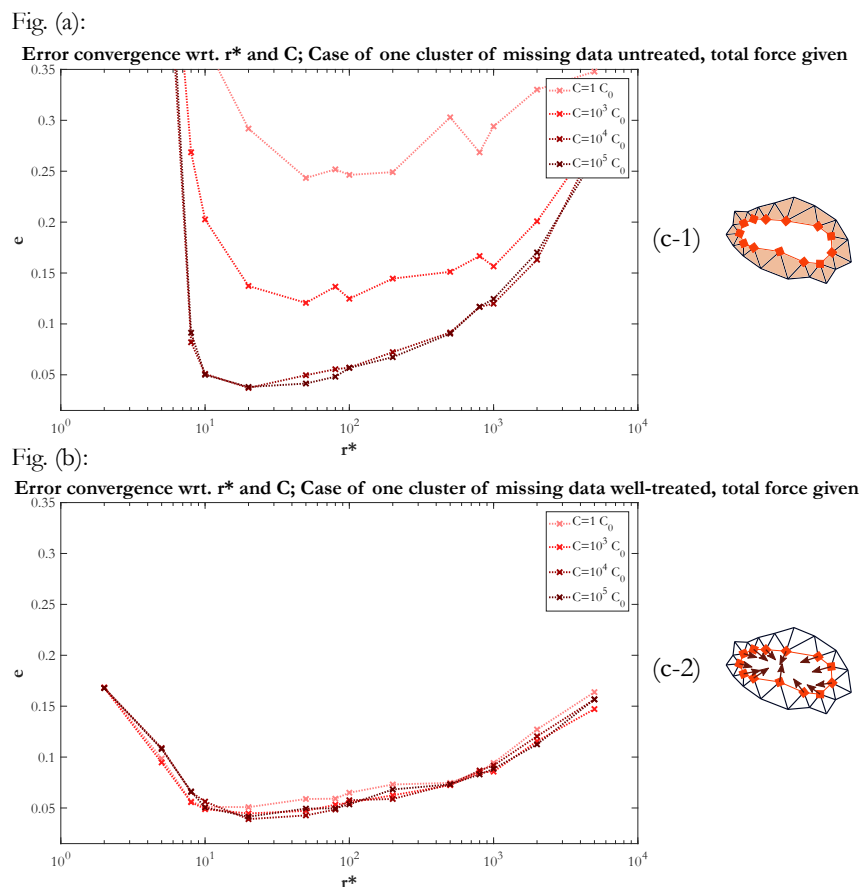


FIGURE 5 – Influence des données d’entrée dans le cas d’un *cluster* sur l’erreur en fonction de r^* et C pour les cas (a) naïf et (b) mécaniquement optimal

le cas d’un *cluster* traité naïvement, il est plus difficile d’obtenir une erreur faible. Un trop grand ou

trop petit nombre d'états matériaux entraîne des erreurs plus importantes. Ici, le choix de C est primordial : plus il est grand (dans une certaine limite) et plus on se rapproche d'un appariement basé sur les déformations, qui sont connues. En ajoutant simplement la condition d'auto-équilibre sur la frontière Eq. (9), on retrouve une robustesse similaire à précédemment ; dans ce cas le choix de C est beaucoup moins important que celui de N^* . En d'autres termes, dans le cas naïf, les éléments du problème ayant un nœud dans la frontière Γ_c (en orange sur la figure 9 (a)) n'ont aucune information pour les équilibrer mécaniquement. Ceci est corrigé dans le cas mécaniquement optimal.

En conclusion, on présente un résultat des contraintes identifiées (pour les 3 composantes) avec ces corrections sur la figure 6.

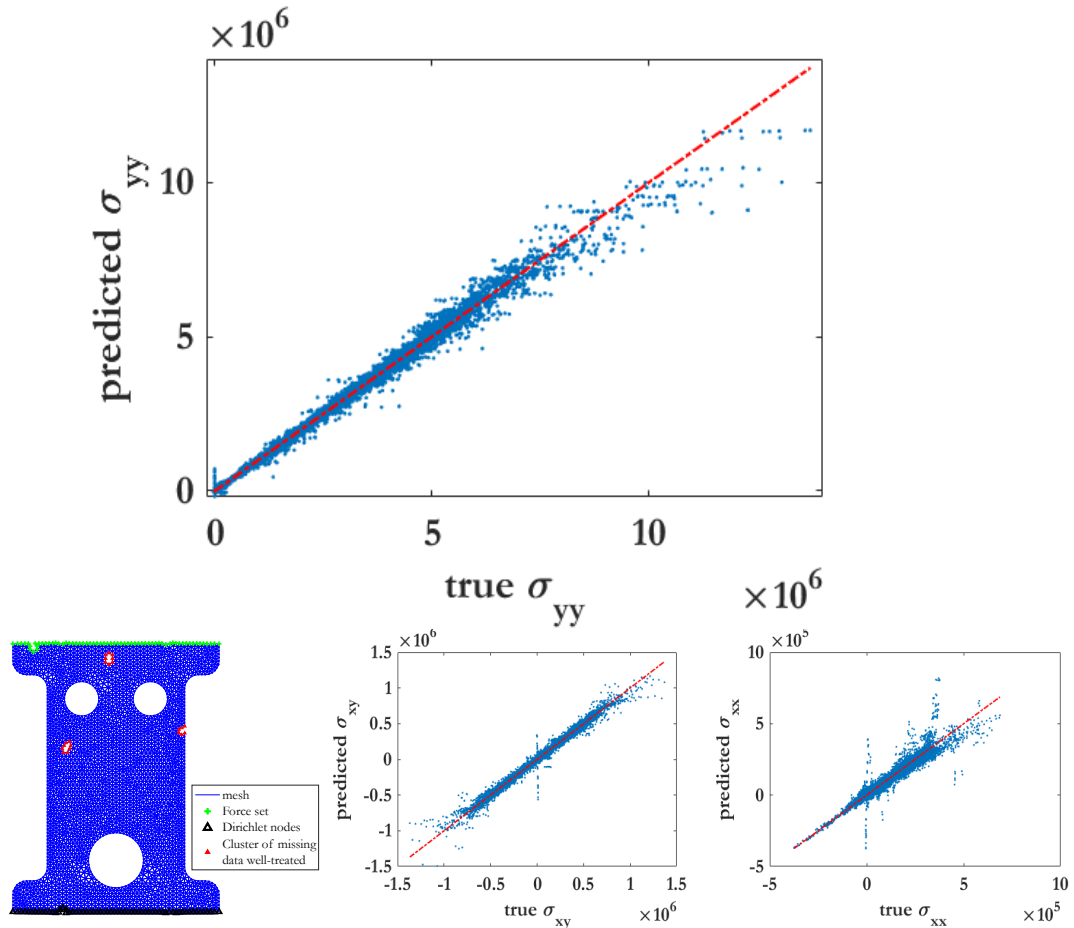


FIGURE 6 – Comparaisons entre les contraintes réelles et les contraintes identifiées ; pour une géométrie avec des *clusters* bien traités et une information de force moyennée

4 Conclusion

Dans ces travaux, on s'intéresse aux paramètres d'entrée de l'algorithme DDI. Une étude de ses paramètres intrinsèques confirment les travaux précédents. En particulier, la prise en compte des données incomplètes (fréquentes pour des données expérimentales) est analysée à travers deux aspects du problème : les forces résultantes et les *clusters* (zones de quelques pixels/éléments) de données manquantes. On montre ainsi que la robustesse de la méthode est assurée quand les données incomplètes sont gérées avec un œil de mécanicien. En effet, le remplacement des forces nodales par leur résultante (mesurée) ne pose aucun problème supplémentaire : il suffit de bien choisir les paramètres de l'algorithme. Pour les *clusters* de données manquantes, des paramètres bien choisis permettent de réduire l'erreur mais la solution la plus robuste consiste à ajouter une condition d'équilibre sur les données perdues. Tous les cas de données incomplètes n'ont pas été traités ici : imperfection des zones sur lesquelles appliquer les conditions aux limites, bruit sur la mesure des déplacements.... Malgré cela, la présente étude valide

l'application de la DDI à des données réelles, incomplètes mais bien traitées.

Références

- [1] Leygue A., Coret M., Réthoré J., Stainier L. & Verron E., *Data-based derivation of material response*, Computer Methods in Applied Mechanics and Engineering, 331, 184–196, 2018.
- [2] Sutton M. A., Orteu J. J. & Schreier H. *Image Correlation for Shape, Motion and Deformation Measurements : Basic Concepts, Theory and Applications*. Springer Science & Business Media, 2009.
- [3] Avril S. et al. *Overview of Identification Methods of Mechanical Parameters Based on Full-field Measurements*. Exp Mech 48, 381 (2008).
- [4] Furukawa T. & Yagawa G. *Implicit constitutive modelling for viscoplasticity using neural networks*. Int. J. Numer. Meth. Engng 43, 195–219, 1998.
- [5] Dalémat M., Coret M., Leygue A. & Verron E., *Measuring heterogeneous stress fields without constitutive equations*, To be submitted.
- [6] Kirchdoerfer T. & Ortiz M., *Data-driven computational mechanics*. Computer Methods in Applied Mechanics and Engineering, 304, 81–101, 2016.