

# Couplage de réseaux récurrents avec des réseaux de neurones par graphes pour la reconstruction de champs mécaniques non-linéaires

M. R. Guevara Garban<sup>1</sup>, Y. Chemisky<sup>4</sup>, E. Prulière<sup>2</sup>, M. Clément<sup>3</sup>, M. Abendroth<sup>5</sup>, B. Kiefer<sup>5</sup>

<sup>1</sup> Univ. Bordeaux, I2M, UMR 5295 351 Cours de la Libération, 33400 Talence, France, manuel.guevara-garban@u-bordeaux.fr

<sup>2</sup> École Nationale Supérieure d'Arts et Métiers, Campus de Bordeaux, I2M, Esplanade des Arts et Métiers, etienne.pruliere@ensam.eu

<sup>3</sup> Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR 5800, 33400 Talence, France, michael.clement@labri.fr

<sup>4</sup> Univ. Grenoble Alpes, CNRS, Grenoble INP, TIMC, UMR 5525, Grenoble, France, yves.chemisky@univ-grenoble-alpes.fr

<sup>5</sup> Institute of Mechanics and Fluid Dynamics, TU Bergakademie Freiberg, Lampadiusstr. 4, 09599 Freiberg, Germany, {martin.abendroth, bjoern.kiefer}@imfd.tu-freiberg.de

**Résumé** — La simulation multi-échelle et non linéaire des structures, dont le comportement dépend de l'histoire de chargement (par exemple pour un matériau élasto-plastique), est très coûteuse. Dans ce travail, nous proposons une approche basée sur le couplage d'un réseau de neurones récurrent (RNN) et d'un réseau de neurones par graphes (GNN) afin d'apprendre le comportement mécanique d'une micro-structure. Le RNN traite les séquences pour caractériser l'effet d'histoire et prédire les champs moyens, tandis que le GNN reconstruit les champs locaux sur un maillage, avec l'objectif de retrouver des résultats similaires à ceux obtenus avec des modélisations éléments finis complètes.

**Mots clés** — réseaux de neurones par graphe, méthodes multi-échelles, localisation de champs.

## 1 Introduction

Les procédés de fabrication additive métallique permettent de concevoir des structures multi-échelles, constituées de la répétition d'une cellule unitaire périodique, éventuellement avec une densité variable pour générer des gradients de rigidité. Ces matériaux architecturés offrent la possibilité d'adapter les propriétés mécaniques en fonction des applications, d'optimiser certaines caractéristiques dans des directions spécifiques ou encore d'alléger les structures. Cependant, la simulation du comportement mécanique de ces structures est très coûteuse en ressources de calcul, en raison de la finesse de maillage nécessaire pour représenter correctement la géométrie de la cellule unitaire, dont l'échelle est généralement d'un ordre de grandeur inférieur à celle de la structure macroscopique. Les approches multi-échelles permettent de réduire en partie cette complexité, mais elles demeurent exigeantes en termes de mémoire et de temps de calcul, en particulier pour les matériaux à comportement non-linéaire. La méthode des éléments finis au carré (FE<sup>2</sup>) consiste à réaliser un calcul macroscopique à l'échelle de la structure, ou la loi de comportement en chaque point d'intégration est obtenu par résolution d'un autre calcul éléments finis complet sur la cellule unitaire, permettant d'obtenir les contraintes moyennes et les matrices tangentes. Ce calcul "microscopique" sur cellule unitaire doit être fait en chaque point d'intégration, à chaque itération en temps, et pour chaque itération du solveur non-linéaire (en général Newton-Raphson). Dans une approche force brute, ce problème que nous appelons ici de façon générique "microscopique" concentre la quasi totalité des coûts de calcul.

Il a été montré que des réseaux de neurones récurrents (RNN) permettent d'apprendre le comportement moyen d'une cellule unitaire en fonction de l'histoire du chargement au court du temps (séquence de déformations) [1]. Néanmoins, cette approche ne permet pas de retrouver les champs locaux qui sont importants pour l'utilisation de certains critères de rupture, par exemple en fatigue, qui sont fortement dépendants des contraintes locales.

Dans ce travail, nous proposons un couplage inédit entre les RNN et des réseaux de neurones en graphes (GNN) afin de reconstruire des champs de contraintes locaux pour des cellules unitaires constituées d'un matériaux à comportement non-linéaires et dépendant de l'histoire du chargement (loi de comportement élasto-plastique). Le réseau est entraîné pour être efficace avec des chargements moyens non proportionnels, telles que ceux que l'on peut obtenir dans une simulation multi-échelles. La méthode se base sur 2 réseaux de neurones :

- un RNN prédit l'évolution des contraintes moyennes d'une cellule unitaire périodique à partir d'une séquence de déformations moyennes, tandis que son état caché encode les dépendances au trajet de chargement subit par les matériau.
- Ces états cachés, associés aux contraintes moyennes prédites, sont ensuite fournis comme caractéristiques d'entrée à un GNN, qui reconstruit la distribution complète des contraintes locales au sein de la microstructure.

Cette combinaison d'un RNN et d'un GNN permet une reconstruction efficace de champs non linéaires dépendants de l'histoire, avec un gain de temps significatif par rapport aux simulations par éléments finis (FEM) classiques.

## 2 Modèle mécanique

La réponse mécanique des métaux et autres matériaux ductiles est généralement décrite dans le cadre des modèles constitutifs élasto-plastiques [2], qui tiennent compte de la coexistence de déformations réversibles (élastiques) et irréversibles (plastiques). Dans le régime des petites déformations, le tenseur de déformation totale est décomposé de manière additive en contributions élastiques et plastiques :

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^e + \boldsymbol{\varepsilon}^p. \quad (1)$$

Dans le cas d'un matériaux isotrope, la déformation élastique  $\boldsymbol{\varepsilon}^e$  est donnée par la loi de Hooke :

$$\boldsymbol{\sigma} = \mathbb{C} : \boldsymbol{\varepsilon}^e, \quad (2)$$

où  $\boldsymbol{\sigma}$  est le tenseur des contraintes de Cauchy et  $\mathbb{C}$  le tenseur d'élasticité d'ordre quatre, qui dépend du module d'Young  $E$  et du coefficient de Poisson  $\nu$ .

La déformation plastique débute lorsque l'état de contrainte satisfait le critère de plasticité. Dans cette étude, le critère de von Mises est retenu, selon lequel l'écoulement plastique se produit lorsque la contrainte équivalente de Von Mises  $\sigma_{vm}$  atteint la limite d'élasticité  $\sigma_y$  avec :

$$\sigma_{vm} = \sqrt{\frac{3}{2} \boldsymbol{s} : \boldsymbol{s}}, \quad (3)$$

où  $\boldsymbol{s}$  est le déviateur du tenseur des contraintes, défini par :

$$\boldsymbol{s} = \boldsymbol{\sigma} - \frac{1}{3} \text{tr}(\boldsymbol{\sigma}) \boldsymbol{I}, \quad (4)$$

avec  $\text{tr}(\boldsymbol{\sigma})$  la contrainte hydrostatique et  $\boldsymbol{I}$  le tenseur identité. Le critère de plasticité s'écrit alors :

$$f(\boldsymbol{\sigma}, \bar{\boldsymbol{\varepsilon}}^p) = \sigma_{vm}(\boldsymbol{\sigma}) - \sigma_y(\bar{\boldsymbol{\varepsilon}}^p) \geq 0, \quad (5)$$

avec  $\bar{\boldsymbol{\varepsilon}}^p$  la déformation plastique équivalente cumulée.

On suppose que le matériau présente un écrouissage isotrope, caractérisé par l'expansion uniforme de la surface de charge lors de l'augmentation de la déformation plastique. L'évolution de la limite élastique est modélisée par une loi d'écrouissage de type loi puissance.

$$\sigma_y(\bar{\boldsymbol{\varepsilon}}^p) = \sigma_{y0} + K (\bar{\boldsymbol{\varepsilon}}^p)^n, \quad (6)$$

avec les paramètres  $K$ ,  $n$  et  $\sigma_{y0}$  représentant la limite élastique initiale du matériau. Cette relation permet de bien décrire les écrouissages non-linéaires généralement observées dans les matériaux métallique ductiles.

La loi de comportement ainsi définie dépend de l'histoire du chargement. Ainsi, l'état de contraintes à un temps donné ne dépend pas uniquement des déformations actuelles mais aussi de toute l'histoire des déformations plastiques, décrit ici par la variable d'état  $\bar{\boldsymbol{\varepsilon}}^p$ . Cette dépendance intrinsèque au chemin et le coût de calcul associé à l'évaluation d'une loi de comportement en chaque point matériel d'une microstructure motivent les développements méthodologiques proposés dans ces travaux : en couplant des réseaux de neurones récurrents (RNN) avec des réseaux de neurones en graphes (GNN), il devient possible de capturer l'évolution temporelle des variables dépendantes de l'histoires et de reconstruire les champs mécaniques locaux, tout allégeant significativement les coûts de calcul par rapport à une simulation par éléments finis.

### 3 Modèle IA

#### 3.1 Réseaux de neurones récurrents

Les réseaux de neurones récurrents (RNN) sont conçus pour modéliser des données séquentielles grâce à un état caché qui conserve l'historique des entrées. Cependant, les RNN classiques sont limités par le problème du "vanishing gradient" [3], qui réduit leur capacité à apprendre des dépendances à long terme. Pour surmonter cette difficulté, les réseaux LSTM (Long Short-Term Memory) [4] introduisent une mémoire interne contrôlée avec un mécanisme de "portes" (entrée, oubli, sortie) permettant de préserver la mémoire sur des longues séquences. En pratique, cela permet aux réseaux LSTM de capturer la dépendance à l'histoire, essentielle pour modéliser les réponses non linéaires des matériaux dépendants du trajet de chargement. L'architecture d'un réseau LSTM est représentée Figure 1.

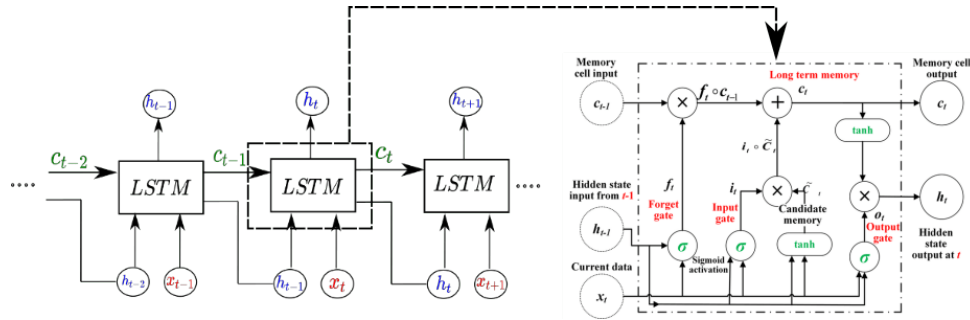


FIGURE 1 – À gauche : architecture d'un réseau LSTM déroulé dans le temps ; à droite : représentation du mécanisme interne d'une cellule LSTM, d'après [5].

De façon analogue à ce qui a été proposé dans les travaux d'A. Danoun [5], le comportement non linéaire homogénéisé d'une microstructure dans le cadre de la méthode  $FE^2$  est approché par 2 couches de LSTM suivi d'un réseau dense (Figure 2). Ainsi, à partir d'une séquence de déformations moyennes, le réseau prédit en sortie les contraintes moyennes sur la microstructure. Il est possible d'utiliser un autre réseau LSTM indépendant pour prédire la matrice tangente nécessaire à la résolution du problème local dans un calcul multi-échelle, comme cela a été fait dans [5].

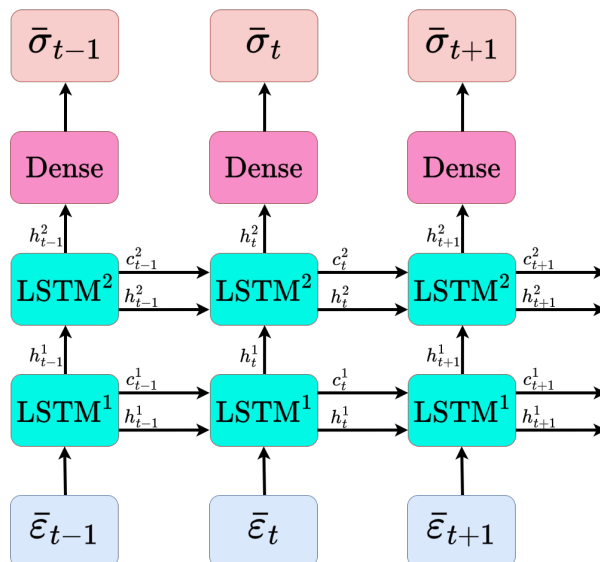


FIGURE 2 – Architecture LSTM implémentée, en prenant une séquence de déformations moyennes  $\bar{\epsilon}$  en entrée et prédit une séquence de contraintes moyennes  $\bar{\sigma}$  en sortie.

## 4 Couplage avec un réseau de neurones par graphe

Le modèle LSTM présenté précédemment encode l'évolution temporelle des déformations moyennes dans un ensemble de variables internes jouant un rôle analogue à celui des variables de plasticité cumulée dans un modèle élasto-plastique à écrouissage isotrope. Ces états cachés  $(h_t^i, c_t^i)$ , définis pour l'ensemble de la microstructure, ne possèdent toutefois pas d'interprétation physique directe. Bien que ce modèle séquentiel reproduise fidèlement les contraintes moyennes, il ne renseigne pas sur leur distribution locale. Pour restituer cette information spatiale, on lui associe un réseau de neurones par graphe (GNN).

Un graphe est défini comme un couple  $G = (\mathcal{V}, \mathcal{E})$ , où  $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$  est un ensemble fini de nœuds, et  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  un ensemble d'arêtes reliant ces nœuds. Chaque arête  $e_{i,j} = (v_i, v_j)$  connecte deux nœuds du graphe. Dans ce cadre, un maillage éléments finis  $\mathcal{M}$  peut naturellement être interprété comme un graphe  $\mathcal{M} = (\mathcal{V}, \mathcal{E})$ .

L'approche consiste à utiliser le LSTM pour extraire les caractéristiques temporelles du chargement. À chaque pas de temps  $t$ , le LSTM traite la déformation moyenne  $\bar{\epsilon}_t$  et produit (i) une prédiction de contrainte moyenne  $\bar{\sigma}_t$  et (ii) un état caché  $h_t$  encodant l'historique du trajet de chargement. Ces informations globales servent ensuite d'entrée au GNN, chargé de propager et de localiser l'information au sein de la microstructure.

Les caractéristiques nodales injectées dans le graphe sont définies par :

$$\mathbf{v}_{i,t} = (h_t, \bar{\sigma}_t, \mathbf{x}_i, \alpha_i),$$

où  $\bar{\sigma}_t \in \mathbb{R}^3$  contient les composantes du tenseur des contraintes moyennes,  $\mathbf{x}_i \in \mathbb{R}^2$  regroupe les coordonnées spatiales du nœud  $i$ , et  $\alpha_i \in \{-1, 0, 1\}$  indique si un nœud appartient à une surface interne, à un bord périodique ou à une zone intérieure (Figure 3). Afin de préserver la périodicité du domaine, des arêtes supplémentaires sont ajoutées pour relier les nœuds situés sur les faces opposées du maillage.

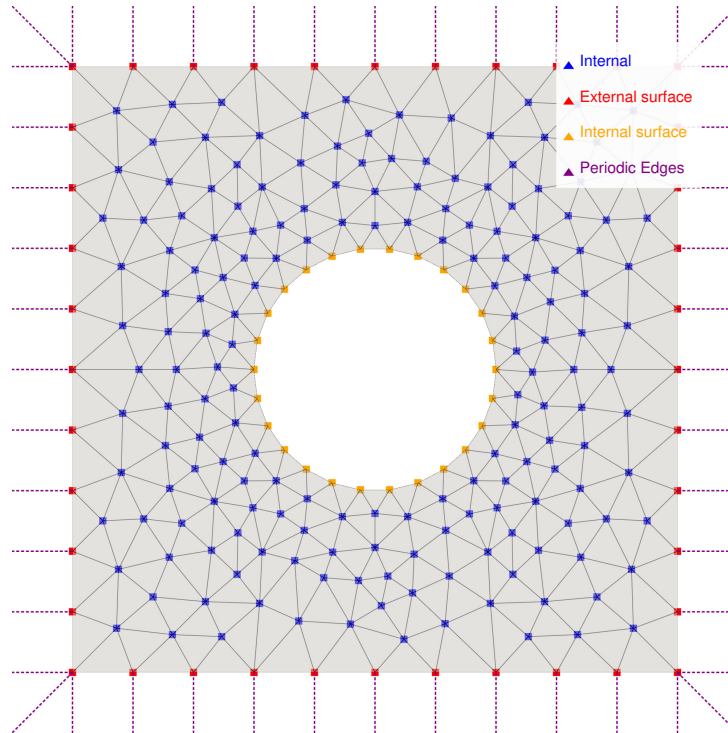


FIGURE 3 – Représentation d'un maillage périodique sous forme de graphe avec les différents labels de nœuds ajoutés pour renforcer l'apprentissage du GNN par *Message Passing*.

Après un encodage préliminaire par un MLP, le GNN réalise une étape de *Message Passing* qui combine les caractéristiques temporelles extraites par le LSTM et les relations spatiales imposées par le graphe. Ce processus permet de reconstruire le champ local de contraintes  $\sigma(\mathbf{x}, t)$  correspondant à l'état de chargement au temps  $t$ .

La Figure 4 illustre l'architecture du modèle proposé : le LSTM fournit une représentation temporelle globale de l'historique de chargement, tandis que le GNN en assure la localisation au sein de la

microstructure.

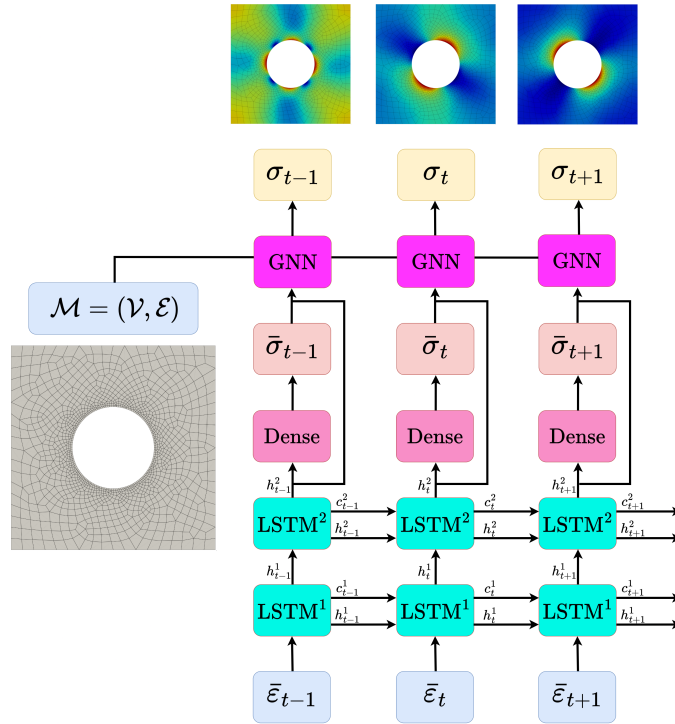


FIGURE 4 – Schéma du modèle couplé LSTM–GNN. Le LSTM traite les séquences de déformations moyennes pour produire des états cachés  $h_t^2$  et des prédictions de contraintes moyennes  $\bar{\sigma}_t$ . Ces sorties sont utilisées par le GNN comme caractéristiques d’entrée des nœuds afin de reconstruire le champ local de contraintes sur la microstructure.

#### 4.1 Génération de la base de données pour matériaux architecturés

Pour entraîner le modèle couplé LSTM–GNN, une base de données par éléments finis contenant 10000 cas de chargement distincts et non proportionnels a été générée. Chaque cas de chargement comporte quatre incréments de déformation, avec des composantes cibles de déformation  $(\bar{\epsilon}_{xx}, \bar{\epsilon}_{yy}, \bar{\epsilon}_{xy})$  échantillonnées aléatoirement dans l’intervalle  $[-0.05, 0.05]$ , et 25 étapes intermédiaires définies entre deux états de déformation successifs. Des exemples représentatifs de ces trajectoires de déformation sont illustrés dans la Figure 6. La base de données est ensuite divisée en ensembles d’entraînement et de validation (70% et 30% respectivement). Les simulations ont été réalisées dans un cadre de déformation plane avec des conditions aux limites périodiques, en utilisant le modèle constitutif élasto-plastique isotrope décrit dans la Section 2, avec des paramètres matériaux, correspondant à un alliage de titane, spécifiés dans le Tableau 1. La géométrie correspond à une plaque percée discrétisée avec des éléments quadrangles, comme illustré dans la Figure 5, et le maillage a été construit à l’aide de la bibliothèque *Gmsh* [6]. Les éléments quadrangles ont été privilégiés par rapport aux éléments triangles en raison de leur meilleure précision pour capturer les gradients locaux des champs.

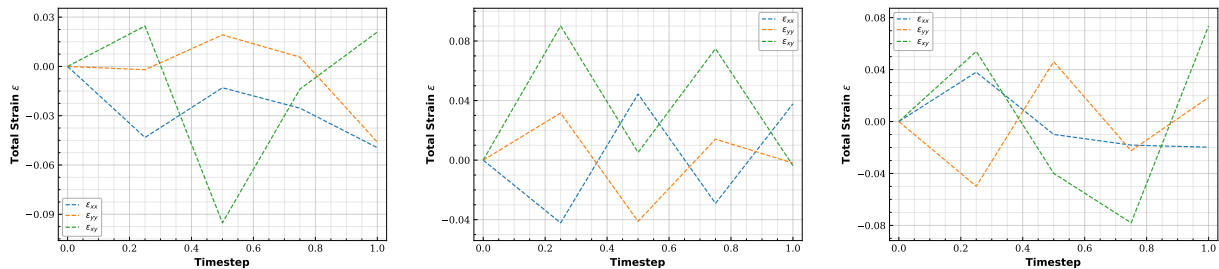


FIGURE 6 – Exemples de trajets de déformations choisis aléatoirement dans la base de données.

Parameter	Value
Young's modulus $E$	$10^5$ MPa
Poisson's ratio $\nu$	0.3
Yield stress $\sigma_{y0}$	300 MPa
Hardening parameter $K$	1000 MPa
Plastic hardening exponent $n$	0.3
Plate width/height $L$	1
Hole radius $r$	0.2

TABLE 1 – Propriétés mécaniques du matériau constitutif et propriétés géométriques de la plaque trouée.

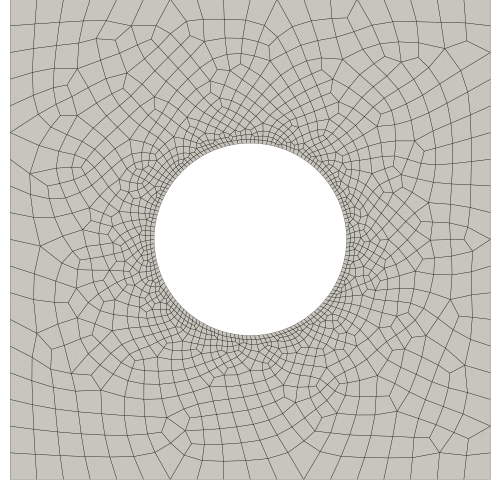


FIGURE 5 – Maillage éléments finis de la plaque trouée (1512 nodes et 1627 éléments quadrangulaires).

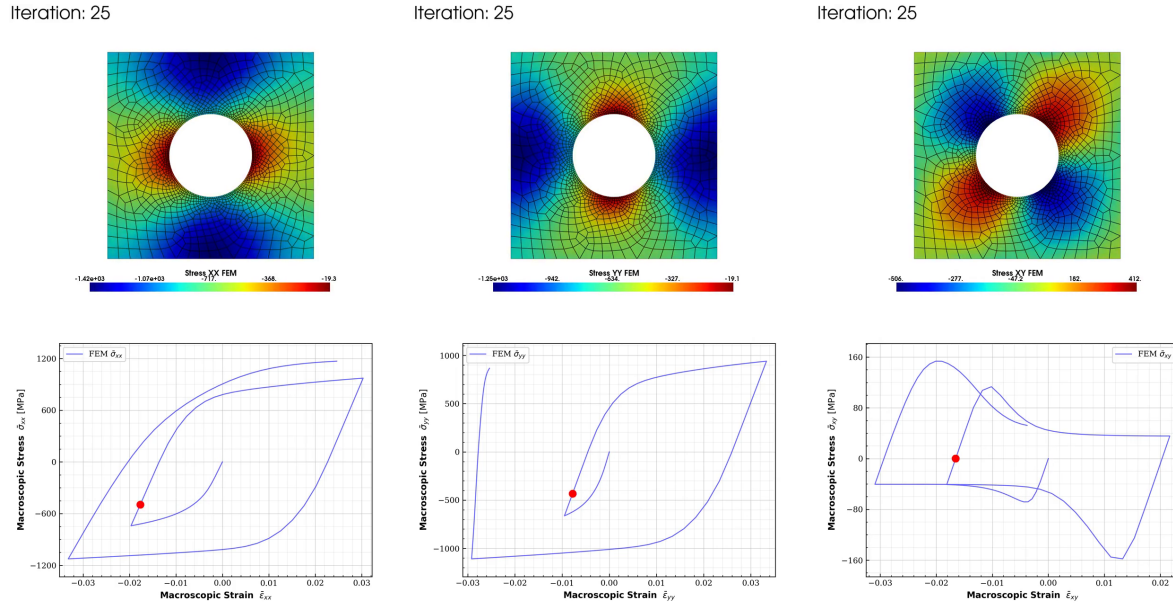


FIGURE 7 – Exemple de champ de contraintes local avec les courbes contraintes moyennes /déformations moyennes correspondantes. Le champ local est affiché à l'itération marquée en rouge sur la courbe.

À partir de chaque simulation, les grandeurs de champ macroscopiques (moyennes) et microscopiques (locales) ont été enregistrées à chaque pas de temps. Ces données constituent la base d'apprentissage pour le modèle LSTM à partir des séquences des valeurs homogénéisées, puis pour le modèle GNN, qui utilise les distributions spatiales des champs. Un exemple de champ local de contraintes à un pas de chargement donné, accompagné de sa réponse contrainte-déformation moyenne correspondante, est présenté dans la Figure 7.

## 4.2 Entraînement et résultats

L'apprentissage du modèle complet LSTM-GNN s'effectue en deux étapes. D'abord, le LSTM est entraîné pour prédire les contraintes moyennes à partir des séquences de déformations. L'optimiseur Adam [7] est utilisé avec une fonction coût de type MSE (Mean Square Error). A titre indicatif, l'apprentissage du LSTM converge avec environ 2000 itérations (ou *epochs*), une itération correspondant à un passage complet sur l'ensemble de la base de données. Sur un GPU RTX A4000, cela représente environ six minutes d'entraînement.

Le GNN est entraîné dans un deuxième temps à reconstruire les distributions locales de contraintes  $\sigma_t$  à partir des informations du maillage  $\mathcal{M}$ , les contraintes moyennes  $\bar{\sigma}_t$  et les états cachés  $h_t^2$  du LSTM comme données d'entrée, en utilisant une version normalisé de la MSE, appelé ici NMSE

$$\text{NMSE}(\sigma_t, \hat{\sigma}_t) = \frac{1}{n_c} \sum_{c=1}^{n_c} \frac{\sum_{i=1}^n (\sigma_{c,i} - \hat{\sigma}_{c,i})^2}{\sum_{i=1}^n (\sigma_{c,i} - \frac{1}{n} \sum_{j=1}^n \sigma_{c,j})^2}, \quad (7)$$

où  $c \in \{xx, yy, xy\}$  représente la composante du tenseur de contraintes,  $\hat{\sigma}$  la contrainte locale prédite par le modèle et  $\sigma$  la contrainte locale de référence calculée par FEM au temps de chargement  $t$ .

Cette phase est plus coûteuse en calcul car le réseau doit s'entraîner à reproduire de champs de contraintes locaux pour chaque pas de chargement. La convergence est observée pour 100 *epochs* qui nécessitent environ 20 heures d'entraînement sur un GPU RTX A4000.

Pour évaluer le modèle proposé, les champs macroscopiques moyens et les champs locaux microscopiques prédits par le cadre LSTM-GNN sont comparés aux solutions obtenues par éléments finis (FEM) sur un trajet de chargement arbitraire qui n'appartient pas à la base d'entraînement. Comme on peut voir dans la Figure 8, le modèle LSTM-GNN prédit des champs macroscopiques et locaux très proches à aux champs obtenus par simulation aux éléments finis. En Figure 9 le champ d'erreurs local est affiché au niveau des noeuds du maillage en montrant une faible différence entre les prédictions du GNN et la solution FEM de référence.

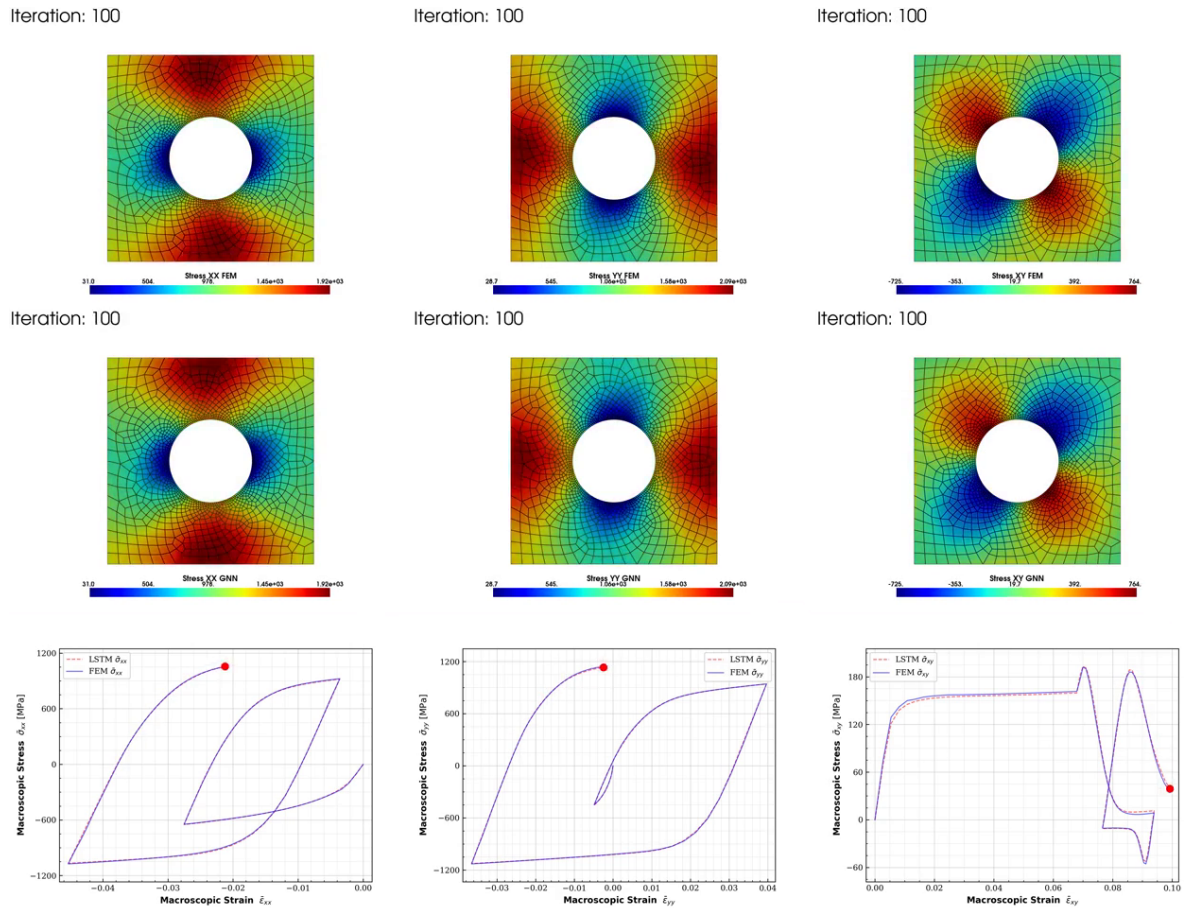


FIGURE 8 – Résultats de champs locaux obtenus par éléments finis pour chaque composante du tenseur de contraintes ( $\sigma_{xx}, \sigma_{yy}, \sigma_{xy}$ ) en haut, résultats de champs locaux obtenus par GNN au milieu, courbe de contraintes déformation moyennes en bas, comparant la prédiction du LSTM vs FEM en fin de chargement (itération 100).

Iteration: 100

Iteration: 100

Iteration: 100

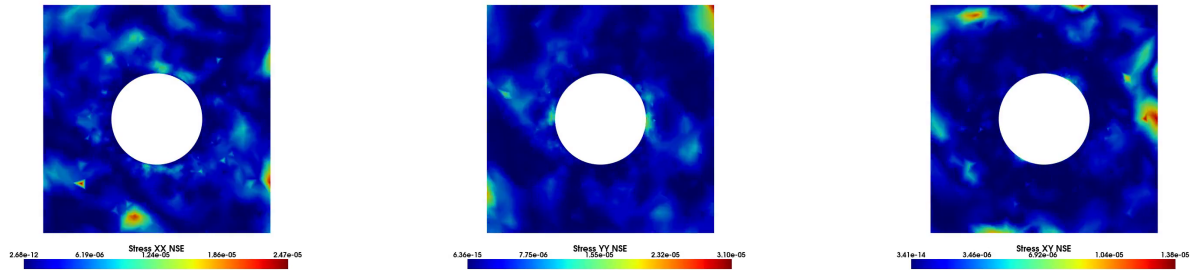


FIGURE 9 – Comparaison de l’erreur locale entre la solution éléments finis et la prédiction du modèle LSTM-GNN en fin de chargement (itération 100).

## 5 Conclusion

Le couplage LSTM–GNN développé dans ce travail permet de prédire avec précision et efficacité l’évolution des champs locaux sous un trajet de chargement non proportionnel piloté en déformation moyenne. Le LSTM capture la dépendance temporelle des contraintes moyennes, tandis que le GNN reconstruit les réponses mécaniques localisées au sein du maillage. Les résultats obtenus confirment le potentiel de cette approche pour accélérer les simulations non linéaires tout en conservant un niveau de précision proche de celui des éléments finis, ouvrant ainsi la voie à son intégration dans des cadres de calcul multi-échelles dédiés à l’identification de critères de fracture et de fatigue sur des microstructures complexes. Ces travaux ouvrent également la perspective d’étendre l’usage des GNN à des maillages tridimensionnels et d’intégrer de manière conjointe LSTM et GNN dans des environnements visant l’optimisation du choix de microstructures pour la simulation multi-échelle.

### 5.1 Références bibliographiques

#### Références

- [1] Aymen Danoun, Etienne Prulière, and Yves Chemisky. Fe- lstm : A hybrid approach to accelerate multiscale simulations of architected materials using recurrent neural networks and finite element analysis. *Computer Methods in Applied Mechanics and Engineering (CMAME)*, 429 :117192, 2024.
- [2] Jean Lemaitre and Jean-Louis Chaboche. *Mechanics of Solid Materials*. Cambridge University Press, 1990.
- [3] Sepp Hochreiter. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6 :107–116, 04 1998.
- [4] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8) :1735–1780, 1997.
- [5] Aymen Danoun. *Numerical simulation of heterogeneous materials combining Artificial Intelligence and physics-based modeling*. PhD thesis, 2022.
- [6] Christophe Geuzaine and Jean-François Remacle. Gmsh : A 3-D finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11) :1309–1331, 2009.
- [7] Diederik P. Kingma and Jimmy Ba. Adam : A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.