

Réseaux de neurones artificiels basés sur la thermodynamique (TANN) pour la mécanique computationnelle et la modélisation multi-échelle

F. MASI^a, I. STEFANOU^a

a. Nantes Université, Ecole Centrale Nantes, CNRS, GeM, UMR 6183, F-44000 Nantes, France.
email : filippo.masi@ec-nantes.fr

Résumé :

Le comportement mécanique des matériaux complexes, affichant un couplage entre les comportements microscopique et macroscopique, tels que les milieux granulaires et les composites, est très difficile à appréhender avec des lois de comportement empiriques. À cette fin, des approches d'homogénéisation multi-échelles sont souvent utilisées, mais leur coût de calcul est prohibitif pour des applications à l'échelle réelle. Récemment, les approches pilotées par les données se sont imposées comme des alternatives prometteuses. Cependant, ces dernières manquent d'un cadre rigoureux basé sur la physique et leur application à des matériaux complexes n'est pas encore établie. Ici, nous proposons des réseaux de neurones artificiels basés sur la thermodynamique (TANN : Thermodynamics-based Artificial Neural Networks) pour la modélisation de matériaux complexes. Notre approche intègre des techniques basées sur la thermodynamique hors-équilibre pour identifier les lois de comportement et les variables d'état internes. La capacité des réseaux TANN à fournir des prédictions à haute fidélité et physiquement cohérentes est démontrée à travers plusieurs exemples pour des matériaux architecturés et des geomatériaux, avec un comportement inélastique. Enfin, nous développons un schéma d'homogénéisation à double échelle basé sur les réseaux TANN pour étudier le comportement des structures en grande échelle avec un coût de calcul de plusieurs ordres de grandeur inférieur aux méthodes classiques.

Abstract :

The mechanical behavior of complex materials, displaying coupled microscopic and macroscopic behaviors, such as granular media and meta-materials, is very difficult to predict with heuristic constitutive laws. To this end, multi-scale homogenization approaches are often used, but their computational cost is prohibitive for real-scale applications. Recently, data-driven approaches have emerged as promising candidates. However, the latter lack a rigorous framework based on physics and their application to complex materials is not yet established. Here, we propose the Thermodynamics-based Artificial Neural Networks (TANN) for the modeling of complex materials. Our approach integrates techniques based on non-equilibrium thermodynamics to identify constitutive laws and internal state variables. The ability of TANN to provide high-fidelity and physically consistent predictions is demonstrated through several examples involving lattice materials and granular media, with inelastic behavior. Finally, we develop a double-scale homogenization scheme based on TANN to study the

behavior of large-scale structures with a computational cost that is more than three-orders of magnitude smaller compared to state-of-the-art approaches.

Mots clefs : Réseaux de neurones, Thermodynamique, Variables internes, Comportement mécanique, Multi-échelle, Matériaux architecturés, Milieux granulaires.

1 Introduction

Les matériaux complexes, tels que, par exemple, les matériaux granulaires et les matériaux architecturés, affichent un fort couplage entre leur microstructure et les mécanismes de déformation. Soumis au chargement, ils sont le siège de différents phénomènes locaux qui, à leur tour, déterminent le comportement à la macro-échelle. La modélisation de tels matériaux ainsi que la prédiction de leur réponse sont particulièrement difficiles à appréhender avec des lois de comportement empiriques.

Afin de surmonter ces difficultés, des approches multi-échelle ont été développées. Celles-ci intègrent un certain nombre de méthodes à différentes échelles, telles que la méthode des éléments finis (FEM, de l'anglais *Finite Element Method*), les unes reliées aux autres par des techniques d'homogénéisation, et des approches hybrides (par exemple la méthode FE² [3] et la méthode FEM×DEM [26]). Formellement, on commence avec la résolution d'un problème aux limites, dit problème auxiliaire et défini sur un Volume Élémentaire Représentatif (VER), pour ensuite reconstruire, à travers un processus de *upscaling*, le comportement macroscopique. Néanmoins, la laborieuse et coûteuse façon avec laquelle le problème auxiliaire est résolu – à chaque pas de temps, et pour tous les points de la macrostructure – constitue le goulot d'étranglement de telles approches et limite leur utilisation [15].

À cette fin, des approches pilotées par les données, aussi appelées *data-driven*, ont commencé à apparaître ces dernières années avec comme but de fournir des descriptions enrichies du comportement matériau, tout en gardant un coût de calcul réduit par rapport aux approches multi-échelle classiques. Un changement de paradigme dans la modélisation traditionnelle en mécanique s'est ainsi produit : les données, plutôt que d'être utilisées uniquement comme des outils de vérification de modèles empiriques, ont commencé à devenir le stimulus dans la quête de nouvelles représentations constitutives. En particulier, les approches basées sur l'intelligence artificielle, principalement l'apprentissage profond, se sont révélées être des candidats efficaces et prometteurs (voir par exemple [8, 12, 24, 14, 10, 22]). Pourtant, les approches data-driven présentent des faiblesses majeures qui entravent leur application en mécanique des solides. L'absence d'un cadre basé sur la physique (effet modèle « boîte noire ») et le besoin de grandes quantités de données de haute qualité (sans bruit) sont quelques-uns des principaux problèmes menaçant la généralisation de ces méthodes à la modélisation des matériaux et à la résolution de problèmes en ingénierie (par ex. [31, 16, 17]).

Récemment, quelques approches, telles que les *physics-informed neural networks* (réseaux de neurones informés par la physique, [29]), les *structure-preserving neural networks* [9] et les *Thermodynamics-based Artificial Neural Networks* (TANN, réseaux de neurones artificiels basés sur la thermodynamique) [15, 19, 20] se sont imposés comme des excellents candidats capables de se confronter aux enjeux susmentionnés. Telles méthodes intègrent, dans leur architecture, des connaissances dérivées des principes et des lois de la physique. Pourtant, même si leur application dans plusieurs domaines de la mécanique de solides et de fluides s'est relevée être extrêmement efficace, leur généralisation à des matériaux complexes (comme définis ci-dessus) reste encore un sujet ouvert.

Dans ce travail, nous présentons une extension des réseaux appelés TANN, précédemment développée par les auteurs et collaborateurs. Les TANN représentent l'union des approches data-driven et les principes universaux de la thermodynamique. Ici, ce formalisme est étendu aux matériaux complexes. Comme démontré plus bas, les réseaux TANN peuvent en effet prédire la réponse constitutive de matériaux complexes et identifier, pour la première fois, de manière autonome, l'ensemble des variables internes, sur la base des informations relatives à la structure interne du matériau et le second principe de la thermodynamique. De plus, ces mêmes réseaux permettent d'effectuer des simulations multi-échelle, ici notées FEM×TANN, en s'appuyant sur le cadre mathématique de l'homogénéisation asymptotique. Le manuscrit est structuré comme suit. La définition et fonctionnement des réseaux de neurones ainsi que le cadre théorique de la thermodynamique sont brièvement présentés dans la Section 2. La Section 3 détaille l'architecture et le formalisme des réseaux TANN, ainsi que leur extension dans le cadre des matériaux complexes. Enfin, en Sections 4 et 5, nous démontrons la performance des TANN avec des applications relatives aux matériaux architecturés ainsi que les milieux granulaires.

Tout au long de ce travail, nous adoptons la notation d'Einstein. On écrit le produit scalaire entre vecteurs et tenseurs du deuxième ordre respectivement : $\mathbf{a} \cdot \mathbf{b} = a_i b_i$ et $\boldsymbol{\pi} : \dot{\mathbf{f}} = \pi_{ij} \dot{f}_{ij}$.

2 Réseaux de neurones et thermodynamique hors équilibre

Cette section résume le cadre théorique général, employé dans ce travail, des réseaux de neurones ainsi que de la thermodynamique hors équilibre.

2.1 Réseaux de neurones : définition et fonctionnement

Les réseaux de neurones sont de puissants opérateurs d'approximation, inspirés du fonctionnement des neurones biologiques et constitués d'un ensemble de nœuds (neurones), assemblés en couches. Dans le cas le plus simple, un réseau de neurones, de profondeur L , a une couche d'entrée, $L - 1$ couches cachées et une couche de sortie (L -ième couche). La connexion entre différentes couches s'exprime par la relation suivante :

$$\mathbf{y}^{(k)} = \mathcal{A}^{(k)} \left(\mathbf{W}^{(k)} \cdot \mathbf{y}^{(k-1)} + \mathbf{b}^{(k)} \right), \quad (1)$$

où $\mathbf{y}^{(k)} \in \mathbb{R}^n$ est la sortie de la couche k , $\mathbf{W}^{(k)} \in \mathbb{R}^{n \times m}$ sont les *poids* entre les m nœuds de la couche k et les n nœuds de la couche $k - 1$, $\mathbf{b}^{(k)} \in \mathbb{R}^n$ sont les *biais* de la couche k et $\mathcal{A}^{(k)}$ est une fonction d'activation généralement non linéaire. Les réseaux de neurones permettent de calculer automatiquement les gradients des sorties, $\mathbf{y}^{(k)}$, par rapport aux entrées, $\mathbf{y}^{(k-1)}$, par la méthode de *differentiation automatique*, $\nabla_{\mathbf{y}^{(l-1)}} \mathbf{y}^{(l)}$.

Les poids et les biais sont ajustés, pendant la phase d'apprentissage, dans une procédure itérative (algorithme de descente de gradient [7]), afin de minimiser l'erreur entre les prédictions (sorties de la dernière couche), \mathbf{o} , et les données, $\bar{\mathbf{o}}$, à l'aide d'une fonction d'erreur \mathcal{L} , par ex. l'erreur absolue moyenne

$$\mathcal{L}(\mathbf{o}) = \sum_{i=1}^N \ell(\bar{\mathbf{o}}^i, \mathbf{o}^i), \quad \ell(\bar{\mathbf{o}}^i, \mathbf{o}^i) = \|\bar{\mathbf{o}}^i - \mathbf{o}^i\|_1 \quad (2)$$

où $k = 1, 2, \dots, N$ est la dimension des sorties \mathbf{o} et $\|\cdot\|_1$ désigne la norme L¹. La fonction d'erreur est ainsi rétro-propagée à travers chaque nœud afin d'ajuster les poids et les biais (*différentiation rétro-propagée*, backpropagation).

Une fois la phase d'apprentissage terminée, le réseau entraîné est utilisé pour effectuer des prédictions

à partir de nouvelles entrées (phase d'inférence).

2.2 Thermodynamique hors-équilibre

Considérons un corps avec une configuration de référence, \mathcal{B} . La forme générale de la conservation de l'énergie, sous forme locale et description référentielle s'exprime par

$$\dot{e} = \boldsymbol{\pi} : \dot{\mathbf{f}} - \nabla \cdot \mathbf{q} + r \quad (3)$$

où e et \dot{e} sont respectivement la densité volumique de l'énergie interne et sa dérivée temporelle locale, $\boldsymbol{\pi}$ est le premier tenseur des contraintes de Piola-Kirchhoff, \mathbf{f} et $\dot{\mathbf{f}}$ sont le gradient de déformation et sa dérivée temporelle, \mathbf{q} est le vecteur flux de chaleur, r une source volumique de chaleur; et $\nabla \cdot$ représente l'opérateur de divergence.

La forme locale du second principe de la thermodynamique s'exprime par

$$d^* = \dot{s} - \nabla \cdot \left(\frac{\mathbf{q}}{\theta} \right) - \frac{r}{\theta} \geq 0, \quad (4)$$

où d^* est la densité volumique du taux de dissipation total, s la densité volumique de l'entropie et θ la température absolue. La combinaison du premier et second principe conduit à l'inégalité fondamentale de Clausius-Duhem [6]

$$d^* = \boldsymbol{\pi} : \dot{\mathbf{f}} - \dot{e} + \theta \dot{s} - \mathbf{q} \cdot \frac{\nabla \theta}{\theta} = \boldsymbol{\pi} : \dot{\mathbf{f}} - (\dot{\psi} + \dot{\theta}s) - \mathbf{q} \cdot \frac{\nabla \theta}{\theta} \geq 0 \quad (5)$$

en ayant introduit la densité d'énergie libre $\psi = e - \theta s$.

Dans l'hypothèse que le taux de dissipation thermique, c'est-à-dire $-\mathbf{q} \cdot \frac{\nabla \theta}{\theta}$, soit défini positif, l'inégalité de Clausius-Duhem peut être reformulée comme

$$d = \boldsymbol{\pi} : \dot{\mathbf{f}} - \dot{\psi} - \dot{\theta}s, \quad d \geq 0 \quad (6)$$

définissant ainsi la densité du taux de dissipation mécanique d .

On continue en sélectionnant un volume arbitraire $\mathcal{V} \subset \mathcal{B}$ et identifions la valeur moyenne d'une quantité dans \mathcal{V} par $\langle \cdot \rangle = \frac{1}{|\mathcal{V}|} \int_{\mathcal{V}} \cdot dx$, avec $|\mathcal{V}| = \int_{\mathcal{V}} dx \neq 0$. En présence des champs de déplacement périodiques et des vecteurs de traction anti-périodiques, l'inégalité de Clausius-Duhem devient alors

$$D = \boldsymbol{\Pi} : \dot{\mathbf{F}} - \dot{\Psi} - \dot{\Theta}S, \quad D \geq 0, \quad (7)$$

avec $\boldsymbol{\Pi} \equiv \langle \boldsymbol{\pi} \rangle$, $\mathbf{F} \equiv \langle \mathbf{f} \rangle$, $\Psi \equiv \langle \psi \rangle$, $\Theta \equiv \langle \theta \rangle$, $S \equiv \langle s \rangle$, and $D \equiv \langle d \rangle$.

2.2.1 Variables et lois d'état

Tout processus satisfaisant l'inégalité de Clausius-Duhem est défini thermodynamiquement admissible. La question principale se réduit alors au choix des variables destinées à la description du comportement thermo-mécanique des matériaux. Nous suivons ici la méthode de l'état local [1]. Ainsi, l'état d'un milieu matériel, à un instant donné, est complètement défini par la connaissance, à cet instant, des valeurs d'un certain nombre de variables d'état. Ces dernières regroupent les variables *observables* et les variables *internes*. Les variables observables ce sont la déformation totale et la température. Les variables internes, $Z = \{Z_0, Z_1, \dots, Z_s\}$, sont des représentations moyennes de l'état microscopique

de la matière (densité de dislocations, micro-structure, etc.) [25]. Elles n'interviennent pas dans les lois de conservation et leur cinématique est décrite à travers des lois d'évolution, régies par le second principe [21]. Ici, par souci de clarté, nous ne postulons ni la nature physique ni la nature tensorielle de ces variables (voir paragraphe 3.2).

Afin de dériver les lois d'état, nous postulons l'existence d'un potentiel thermodynamique, l'énergie libre, tel que [13]

$$\Psi = \tilde{\Psi}(\Theta, \mathbf{F}, \mathbf{Z}). \quad (8)$$

En remplaçant la dérivée temporelle de l'énergie libre dans l'équation (6), le caractère arbitraire de $\dot{\Theta}$, $\dot{\mathbf{F}}$ et $\dot{\mathbf{Z}}$ conduit aux restrictions constitutives suivantes basées sur le premier et second principe de la thermodynamique

$$-\underbrace{\left(\frac{\partial \Psi}{\partial \Theta} + S \right)}_{S = -\frac{\partial \Psi}{\partial \Theta}} \dot{\Theta} + \underbrace{\left(\frac{\partial \Psi}{\partial \mathbf{F}} - \boldsymbol{\Pi} \right)}_{\boldsymbol{\Pi} = \frac{\partial \Psi}{\partial \mathbf{F}}} : \dot{\mathbf{F}} + \underbrace{\left(D + \frac{\partial \Psi}{\partial \mathbf{Z}} \bullet \dot{\mathbf{Z}} \right)}_{D = \mathbf{A} \bullet \dot{\mathbf{Z}} \geq 0} = 0, \quad \forall \dot{\Theta}, \dot{\mathbf{F}}, \dot{\mathbf{Z}} \quad (9)$$

où "•" désigne un simple produit, un produit scalaire ou un produit tensoriel, selon la nature des variables internes, $\mathbf{A} \equiv -\frac{\partial \Psi}{\partial \mathbf{Z}}$ les forces thermodynamiques et $\dot{\mathbf{Z}}$ le taux des variables internes (ou flux thermodynamiques). Il convient de noter que le cadre susmentionné peut être facilement étendu aux milieux continus généralisés [5, 4].

3 TANN : réseaux de neurones basés sur la thermodynamique

Sur la base du cadre théorique présenté en Section 2, les auteurs et collaborateurs ont précédemment développé les réseaux de neurones artificiels basés sur la thermodynamique (TANN : *Thermodynamics-based Artificial Neural Networks*, voir [19, 20]). Les TANN relient le pouvoir d'approximation des réseaux de neurones et les lois universelles de la thermodynamique. Elles sont composées par plusieurs éléments de base, présentés en détail dans la suite.

3.1 Réseaux élémentaires

Le bloc principale des TANN, nommé réseau thermodynamique, NN_ψ , prédit l'énergie libre à partir de la connaissance des variables d'état, équation (8), voir Figure 1. Le calcul des contraintes, de l'entropie et du taux de dissipation est effectué en s'appuyant sur la différenciation automatique de l'opérateur NN_ψ par rapport à ses entrées. Le réseau est entraîné en minimisant la fonction d'erreur présentée en Figure 1 qui permet, à apprentissage effectué, de prédire le comportement thermo-mécanique des matériaux en respectant, par construction, le premier et second principe de la thermodynamique.

Toutefois, le réseau thermodynamique n'est pas suffisant pour prédire le comportement matériel. Une loi d'évolution des variables internes, \mathbf{Z} , doit être ainsi identifiée. Cette dernière peut s'exprimer, de façon générale [21], sous la forme

$$\dot{\mathbf{Z}} = g(\mathbf{Z}, \mathbf{F}, \Theta, \dot{\mathbf{F}}, \dot{\Theta}). \quad (10)$$

Dans le cadre de réseaux TANN, ceci se traduit dans la construction d'un réseau, NN_Z , qui permet découvrir, à partir des données, la fonction $g(\cdot)$ et donc la loi d'évolution, voir Figure 2.

En assemblant les deux réseaux, NN_ψ et NN_Z , nous obtenons l'architecture des réseaux TANN. On renvoie le lecteur intéressé au travail originel [19], où des analyses détaillées sur l'efficacité, la

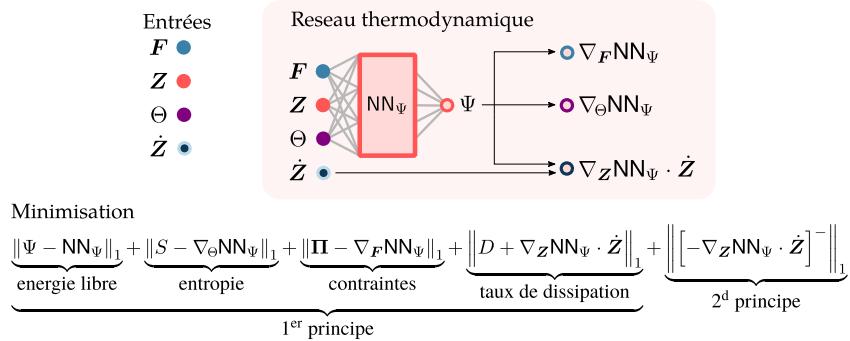


FIGURE 1 – Architecture du réseau thermodynamique et forme de la fonction d’erreur, avec $[x]^-=\min(x, 0)$.

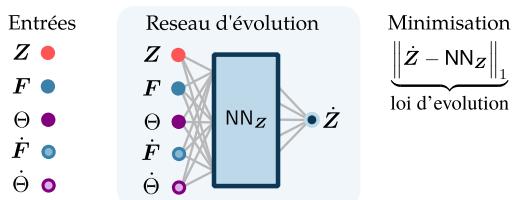


FIGURE 2 – Architecture du réseau d’évolution et forme de la fonction d’erreur.

performance et la robustesse des réseaux TANN sont présentées dans le cas des matériaux homogènes inélastiques.

3.2 Variables internes et leur identification

L’ingrédient principal du formalisme TANN que nous venons de présenter est la connaissance des variables internes. Néanmoins, en présence de matériaux complexes, comme définis dans l’introduction (Section 1), l’ensemble des variables internes n’est pas connu a priori. Il n’existe pas non plus de cadre rigoureux pour le choix de ces dernières. Cette difficulté est généralement contournée en utilisant des variables dites *d’histoire* (voir, par ex., [11, 27]) et/ou de variables internes prédefinies. Toutefois, de telles solutions peuvent conduire à des approches de modélisation qui demandent un choix de la part de l’utilisateur et semblent difficilement adaptables au large spectre des matériaux complexes.

À cette fin, nous proposons une modification du schéma original des réseaux TANN en introduisant ce que nous définissons ici comme des *coordonnées internes*, ξ . Les coordonnées internes, ce sont des variables décrivant l’état du matériau à l’échelle microscopique/internes – par ex. champs de déplacement, de vitesse, de quantité de mouvement, forces. À travers des techniques de réduction de dimensionnalité, ces coordonnées internes peuvent être transformées en une représentation thermodynamiquement admissible des variables internes.

Ici, nous utilisons des auto-encodeurs, ce sont de réseaux de neurones composés d’un encodeur, qui transforme les données d’entrée (les coordonnées internes) en une représentation intrinsèque (de faible dimension), et un décodeur, qui effectue la transformation inverse. Ce couplage, uni au schéma de TANN et les principes de la thermodynamique, permet de découvrir, pour la première fois de manière autonome, une représentation intrinsèque qui coïncide avec un ensemble thermodynamiquement admissible des variables internes (voir aussi [20]).

En assemblant l’auto-encodeur avec le reste de l’architecture présentée dans le paragraphe ci-dessus, on obtient une nouvelle formulation des réseaux de neurones basés sur la thermodynamique, voir Figure

3. Cette extension, basée sur une formulation incrémentale, prévoit ainsi l'identification des variables internes en effectuant l'apprentissage de trois blocs élémentaires (auto-encodeur, réseaux d'évolution et réseaux thermodynamique) au même temps.

Une fois entraîné, le réseau TANN est utilisé pour effectuer des prédictions (phase d'inférence). À

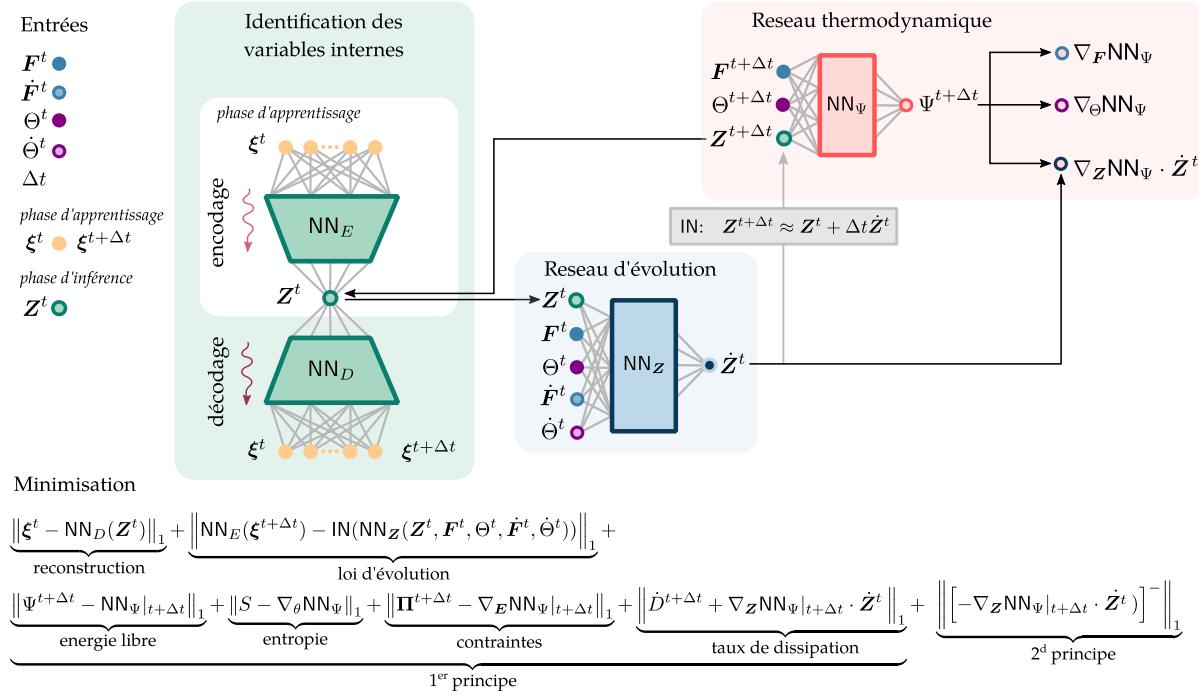


FIGURE 3 – Architecture du réseau d'évolution et forme de la fonction d'erreur.

noter que seuls les réseaux thermodynamique et d'évolution sont nécessaires pour la prédiction de l'état matériel au nouveau temps, $t + \Delta t$, à partir de l'état au temps précédent, t . Cependant, le décodeur peut être utilisé si on souhaite retransformer les variables internes en coordonnées internes, en permettant ainsi de suivre l'évolution de la structure interne du matériau, sans besoin de résoudre des équations de bilan et/ou de mouvement.

À ce stade, le réseau TANN peut être utilisé pour effectuer des analyses aux éléments finis, en remplaçant directement les modèles constitutifs standards, aux points de Gauss, par le réseau entraîné. Les exemples et applications présentés ci-dessous démontrent la capacité des TANN à fournir des descriptions du comportement mécanique des matériaux à haute fidélité et (toujours) thermodynamiquement cohérentes, dépassant ainsi les limites des approches data-driven classiques. De plus, les modèles de réduction de dimensionnalité basés sur la thermodynamique permettent de découvrir, automatiquement et pour la première fois, des ensembles thermodynamiquement admissibles de variables d'état internes pour des matériaux complexes. Inversement, les variables internes ainsi identifiées peuvent être utilisées pour prédire l'évolution de la microstructure. Une telle caractéristique peut accompagner la découverte et la compréhension de la mécanique des matériaux complexes, via des techniques d'extraction de caractéristiques (*features extraction*).

4 Applications aux matériaux architecturés

Dans les applications suivantes, nous considérons des processus isothermes, dans l'hypothèse des petites perturbations, $|\nabla u| \ll 1$. Ainsi, $\mathbf{P} \approx \Sigma$, $\text{Sym} \mathbf{F} = \mathbf{E} + \mathbf{I}$, $\mathbf{E} = \langle \text{Sym} \nabla \mathbf{u} \rangle$, avec Σ étant la valeur

moyenne du tenseur des contraintes de Cauchy, E la valeur moyenne du tenseur des déformations infinitésimales, et $\text{Sym}F$ désignant la partie symétrique du gradient de déformation.

4.1 Modélisation au point matériel

Considérons un volume élémentaire représentatif d'un matériau architecturé, dont la microstructure est montrée en Figure 4(a). Chaque barre a une section circulaire et présente un comportement élasto-plastique, avec un module de Young égal à 100 GPa, une limite d'élasticité $\Sigma_y = 300 \text{ MPa}$ et un module d'écrouissage $H = 10 \text{ GPa}$. Pour cet exemple, nous prenons comme coordonnées internes les champs des déplacements microscopiques et des forces internes. Cependant, les coordonnées internes et, par conséquent, l'approche ici proposée ne se limitent pas à ce choix particulier.

L'apprentissage du réseau TANN est effectué à partir de données générées en résolvant un problème incrémental aux limites sur le VER, en terme des champs de déplacement microscopiques, à l'aide d'un code développé par les auteurs [20]. En nous appuyant sur des conditions aux limites périodiques, nous appliquons des chemins de déformation aléatoires, avec des taux de déformation constants, dans chaque pas de temps. Les états matériaux générés aléatoirement sont ensuite divisés en deux ensembles d'incrément, les premiers à l'instant t et les deuxièmes à l'instant $t + \Delta t$. Ceux-ci sont en suite mélangés de manière aléatoire et utilisés pour l'apprentissage.

En Figure 4(b), nous montrons les prédictions du réseau TANN pour un chemin de chargement bi-axial, jamais vu par le réseau lors de l'apprentissage. Malgré l'apparente complexité de la microstructure, les TANN parviennent à (1) identifier un espace réduit de variables d'état internes qui caractérisent les champs microstructuraux, (2) apprendre l'énergie libre et le taux de dissipation, et (3) prédire avec précision la contrainte moyenne par rapport au modèle micromécanique onéreux computationnel. Des exemples supplémentaires, y compris le cas de structures irrégulières, peuvent être trouvés dans un travail précédent [20].

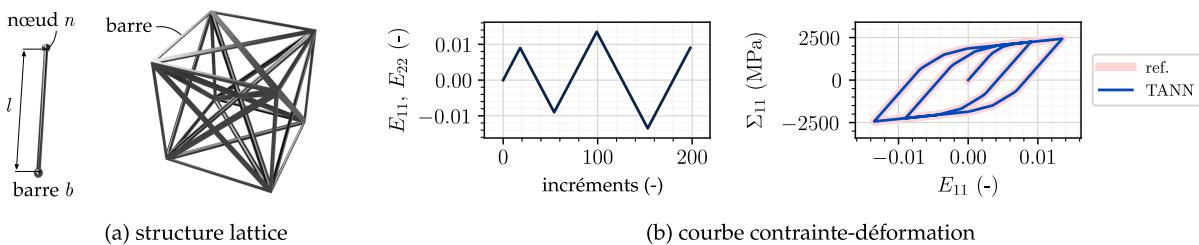


FIGURE 4 – Volume élémentaire représentatif d'un matériau architecturé (a) et sa réponse à une déformation cyclique bi-axiale en termes de simulations micromécaniques et de prédictions du réseau TANN.

4.2 Modélisation multi-échelle

Après avoir démontré les capacités des TANN à prédire le comportement au point matériel, nous considérons des structures à grande échelle, caractérisées par de distributions spatialement (quasi-) périodiques du VER en Figure 4(a). En vertu de l'homogénéisation asymptotique [28], nous faisons l'hypothèse de séparation d'échelles, c'est-à-dire on suppose l'existence de deux échelles indépendantes x et $y = x/\epsilon$, avec ϵ la dimension du VER, associées respectivement au comportement macroscopique et microscopique. L'homogénéisation asymptotique linéaire classique est étendue aux problèmes non linéaires au moyen d'une formulation incrémentale de la non-linéarité (voir [23]).

Étant donné que le réseau TANN a été entraîné sur un VER avec des conditions aux limites périodiques, celui-ci peut être utilisé pour prédire la solution du problème auxiliaire d'homogénéisation, tout en respectant les lois de la thermodynamique. Ainsi, les TANN permettent d'effectuer des analyses multi-échelles – notées ici FEM×TANN, pour les distinguer de schémas classiques type FE² et FEM×DEM – par un simple remplacement des modèles de comportement classiques. De plus, dans des analyses implicites, comme celles présentées ci-dessous, le jacobien est également nécessaire. Ce dernier est calculé par le réseau grâce à la différenciation automatique [19].

Nous comparons les résultats obtenus par l'approche FEM × TANN avec les simulations micromécaniques de l'ensemble de la structure, pour de valeurs différentes du paramètre d'échelle ϵ . Dans toutes les simulations, le maillage des modèles aux éléments finis est composé d'éléments tétraédriques linéaires, de taille uniforme et égale à 0.5 cm, sélectionnés lors d'analyses de sensibilité au maillage.

4.2.1 Application I

Comme première application, nous considérons une structure de $10 \times 10 \times 20$ cm³ (largeur × profondeur × hauteur). Un déplacement correspondant à un angle de torsion $\omega = 60^\circ$ est appliqué à la frontière supérieure, laissant libres tous les autres degrés de liberté afin de permettre un éventuel gauchissement. En figure 5, on compare les résultats de l'approche FEM×TANN aux simulations micromécaniques (exactes). L'approche proposée prédit avec une très bonne précision les réponses microscopique et moyenne, avec une erreur par rapport aux simulations micromécaniques aussi faible que 2%, pour $\epsilon \rightarrow 0$. De plus, la précision du réseau est telle que le taux de convergence quadratique, typique des schémas aux EF implicites, est préservé.

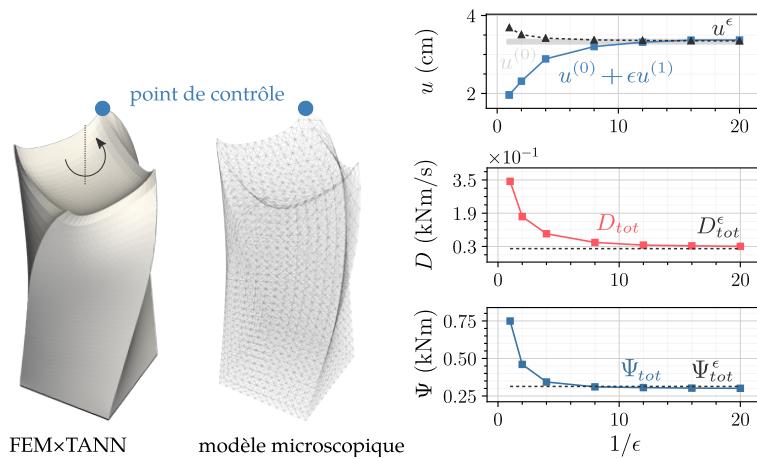


FIGURE 5 – Comparaison entre l'approche FEM×TANN et les simulations micromécaniques (exposant ϵ) pour un chargement en torsion. À gauche : déformée. À droite : déplacement vertical au point de contrôle (à gauche) et énergie totale (à droite). Les déplacements sont calculés avec une approximation du premier ordre, $\mathbf{u} = \mathbf{u}^{(0)} + \epsilon \mathbf{u}^{(1)}$, avec $\mathbf{u}^{(0)}$ le terme d'ordre zéro et $\mathbf{u}^{(1)}$ le terme du premier ordre, coïncidant avec les coordonnées internes, obtenues en décodant les variables internes.

4.2.2 Application II

Nous considérons, dans cette deuxième application, des déplacements cycliques d'amplitude croissante, voir Figure 6. Un incrément final prévoit le déchargement de la structure. Ce déchargement permet d'observer des déplacements permanents dus à la plasticité. Les conditions aux limites et la géométrie coïncident avec celles de l'application I.

En Figure 6, on peut observer que les résultats de l'approche FEM×TANN sont en accord avec la solution du modèle micromécanique, malgré les charges exigeantes appliquées. Particulièrement intéressante est la comparaison à l'incrément final, où un déchargement soudain est imposé. L'approche proposée prédit avec précision le déplacement vertical résiduel non nul au point de contrôle et le taux de dissipation total.

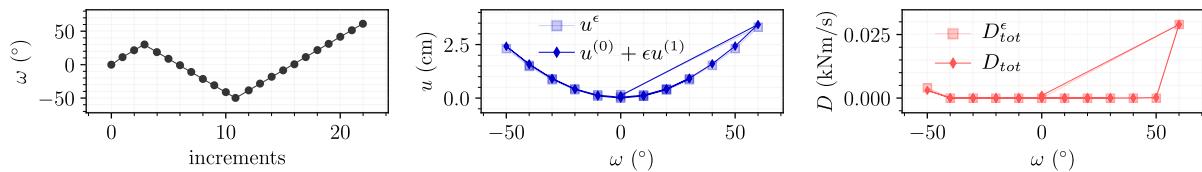


FIGURE 6 – Comparaison entre l'approche FEM×TANN et le modèle micromécanique (exposant ϵ) pour un chemin de chargement de torsion cyclique (gauche) : déplacements au point de contrôle (centre) et taux de dissipation (droite).

En Figure 7, on peut observer l'évolution du champ de déplacements et d'une des variables d'état internes, calculées par l'approche proposée. Après le déchargement, la variable interne a une valeur non nulle. Ceci est une conséquence directe du caractère dissipatif des variables internes, automatiquement identifié avec les TANN.

De plus, en comparant le coût computationnel, l'approche FEM×TANN s'avère fournir des accélérations computationnelles aussi élevées que $10^2\text{-}10^3$ par rapport au modèle micromécanique.

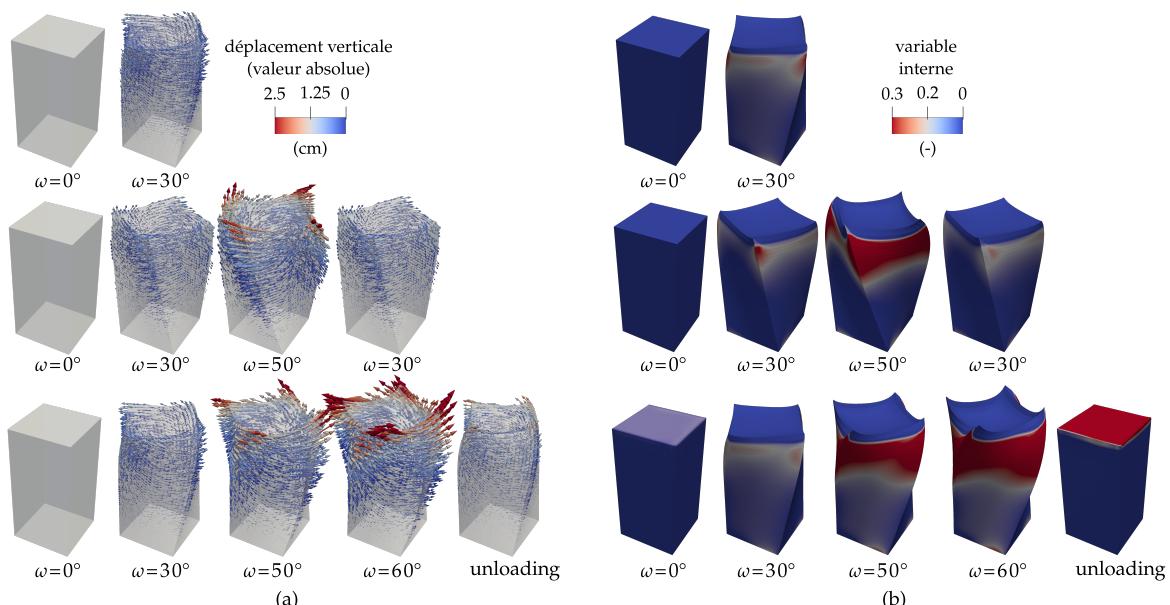


FIGURE 7 – Évolution du champ de déplacements (a) et d'une des variables d'état internes (b) obtenues par l'approche FEM×TANN, pour un chargement cyclique de la Figure 4.

5 Applications aux milieux granulaires

Comme dernière application de la méthode développée dans ce travail, nous considérons le cas emblématique des milieux granulaires. Plus précisément, en s'appuyant sur la méthode des éléments discrets (DEM), nous modélisons des assemblages granulaires polydisperses denses, avec un comportement qui est représentatif de celui des sols. Le modèle est composé de grains (sphériques) infiniment rigides interagissant les uns avec les autres à travers de contacts modélisées avec une loi linéaire élasto-plastique [2].

Un volume élémentaire représentatif constitué de 8000 grains, voir Fig. 8(a), dans un état initial de compression triaxiale isotropique, est utilisé pour la génération des données pour l'apprentissage du réseau. Les paramètres de modélisation sont choisis en accord avec l'état de l'art : module de Young $E = 10$ GPa, coefficient de poisson $\nu = 0.3$, angle de frottement $\varphi = 30^\circ$, et diamètre des grains $d = 0.9 \div 2$ mm. Le VER est soumis à des chemins aléatoires en cisaillement simple, en régime quasi-statique (conditions aux limites périodiques), voir Figure 8(b). Les données ainsi générées sont ensuite divisées en deux ensembles d'incrément, comme présenté en Section 4.

À la différence des exemples ci-dessus, nous considérons les champs des forces internes (produit des interactions intergranulaires) comme coordonnées internes, afin d'identifier les variables internes nécessaires pour décrire le comportement mécanique.

Une fois entraîné, le réseau est employé pour prédire le comportement matériel sous des histoires de chargement différentes de celles présentes dans les données d'apprentissage. En Figure 8(c-e), on compare les prédictions du réseau avec les résultats du modèle DEM pour des chargements cycliques. Les TANN, une fois de plus, se révèlent extrêmement fidèles au modèle microscopique, tout en respectant les lois de la thermodynamique. De plus, quand comparé aux simulations DEM, le réseau TANN permet d'atteindre des accélérations computationnelles de plus de trois ordres de grandeur.

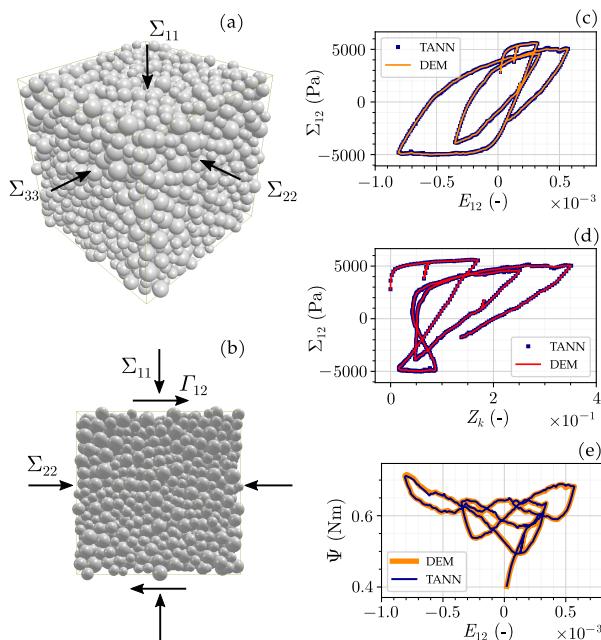


FIGURE 8 – Modélisation d'un milieu granulaire par le réseau TANN : (a,b) VER composée de 8000 grains, (c-e) comparaison de prédictions du réseau avec les résultats DEM en terme des contraintes et de l'énergie libre.

6 Conclusions

La limitation principale des approches multi-échelle est la résolution du problème auxiliaire qui est laborieuse et coûteuse, limitant ainsi leur utilisation dans des applications à grande échelle.

Dans ce travail, nous développons une méthode alternative qui, à l'aide des réseaux de neurones artificiels basés sur la thermodynamique (TANN), permet de prédire le comportement mécanique des matériaux complexes. Ceci est rendu possible grâce à l'intégration du premier et second principe de la thermodynamique, qui assurent des résultats toujours thermodynamiquement admissibles. De plus, l'approche, intégrée dans un schéma d'homogénéisation à double échelle, permet d'effectuer des simulations multi-échelle, avec un coût de calcul extrêmement réduit par rapport aux méthodes classiques.

Un autre point clé de ce travail est l'identification des variables internes à partir des coordonnées internes, propres de la structure interne d'un matériau. Ces dernières peuvent inclure les déplacements, les vitesses, les contraintes et tout autre champ caractéristique de la microstructure. L'identification des variables internes a été réalisée en développant une nouvelle architecture des réseaux TANN qui implique des techniques de réduction de dimensionnalité.

La méthode proposée est générale et peut être appliquée à une large gamme de matériaux complexes, puisque le cadre thermodynamique ici utilisé reste valable, indépendamment de la complexité et de la nature du matériau.

Références

- [1] Coleman, B.D., Gurtin, M.E., 1967. Thermodynamics with internal state variables. *Journal of Chemical Physics* 47, 597–613. doi : [10.1063/1.1711937](https://doi.org/10.1063/1.1711937)
- [2] Cundall, P.A. and Strack, O.D., 1979. A discrete numerical model for granular assemblies. *Géotechnique*, 29(1), 47-65. doi : [10.1680/geot.1979.29.1.47](https://doi.org/10.1680/geot.1979.29.1.47)
- [3] Feyel, F., 2003. A multilevel finite element method (FE^2) to describe the response of highly non-linear structures using generalized continua. *Computer Methods in Applied Mechanics and Engineering* 192, 3233–3244. doi : [10.1016/S0045-7825\(03\)00348-7](https://doi.org/10.1016/S0045-7825(03)00348-7)
- [4] Forest, S., 2020. Continuum thermomechanics of nonlinear micromorphic, strain and stress gradient media. *Philosophical Transactions of the Royal Society A* 378, 20190169. doi : [10.1098/rsta.2019.0169](https://doi.org/10.1098/rsta.2019.0169)
- [5] Germain, P., 1973. The method of virtual power in continuum mechanics. part 2 : Microstructure. *SIAM Journal on Applied Mathematics* 25, 556–575. doi : [10.1137/0125053](https://doi.org/10.1137/0125053)
- [6] Germain, P. 1973. Cours de mécanique des milieux continus. Masson.
- [7] Géron, A., 2019. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow : Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media
- [8] Ghaboussi, J., Garrett, J.H., Wu, X., 1991. Knowledge-based modeling of material behavior with neural networks. *Journal of Engineering Mechanics* 117, 132–153. doi : [10.1061/\(ASCE\)0733-9399\(1991\)117:1\(132\)](https://doi.org/10.1061/(ASCE)0733-9399(1991)117:1(132))
- [9] Hernandez, Q., Badia, A., Gonzalez, D., Chinesta, F., Cueto, E., 2021. Deep learning of thermodynamics-aware reduced-order models from data. *Computer Methods in Applied Mechanics and Engineering* 379, 113763. doi : [10.1016/j.cma.2021.113763](https://doi.org/10.1016/j.cma.2021.113763)

- [10] Huang, D.Z., Xu, K., Farhat, C., Darve, E., 2020. Learning constitutive relations from indirect observations using deep neural networks. *Journal of Computational Physics* 416, 109491. doi : [10.1016/j.jcp.2020.109491](https://doi.org/10.1016/j.jcp.2020.109491)
- [11] Kirchdoerfer, T., Ortiz, M., 2016. Data-driven computational mechanics. *Computer Methods in Applied Mechanics and Engineering* 304, 81–101. doi : [10.1016/j.cma.2016.02.001](https://doi.org/10.1016/j.cma.2016.02.001)
- [12] Lefik, M., Schrefler, B.A., 2003. Artificial neural network as an incremental non-linear constitutive model for a finite element code. *Computer methods in applied mechanics and engineering* 192, 3265–3283. doi : [10.1016/S0045-7825\(03\)00350-5](https://doi.org/10.1016/S0045-7825(03)00350-5)
- [13] Lemaitre, J., Chaboche, J.-L., Benallal, A., Desmorat R. 2019. Mécanique des matériaux solides, 3e Édition, Dunod.
- [14] Liu, Z., Wu, C., 2019. Exploring the 3D architectures of deep material network in data-driven multiscale mechanics. *Journal of the Mechanics and Physics of Solids* 127, 20–46. doi : [10.1016/j.jmps.2019.03.004](https://doi.org/10.1016/j.jmps.2019.03.004)
- [15] Masi, F., 2020. Fast-dynamic response and failure of masonry structures of non-standard geometry subjected to blast loads (Doctoral dissertation, École centrale de Nantes).
- [16] Masi, F., Stefanou, I. and Vannucci, P., 2018. A study on the effects of an explosion in the Pantheon of Rome. *Engineering structures*, 164, 259-273. doi : [10.1016/j.engstruct.2018.02.082](https://doi.org/10.1016/j.engstruct.2018.02.082)
- [17] Masi, F., Stefanou, I., Maffi-Berthier, V. and Vannucci, P., 2020. A Discrete Element Method based-approach for arched masonry structures under blast loads. *Engineering Structures*, 216, 110721. doi : [10.1016/j.engstruct.2020.110721](https://doi.org/10.1016/j.engstruct.2020.110721)
- [18] Masi, F., Stefanou, I., Vannucci, P., Maffi-Berthier, V., 2021a. Material modeling via thermodynamics-based artificial neural networks, in : Barbaresco, F., Nielsen, F. (Eds.), *Geometric Structures of Statistical Physics, Information Geometry, and Learning*, Springer International Publishing, Cham. pp. 308–329. doi : [10.1007/978-3-030-77957-3_16](https://doi.org/10.1007/978-3-030-77957-3_16)
- [19] Masi, F., Stefanou, I., Vannucci, P., Maffi-Berthier, V., 2021b. Thermodynamics-based artificial neural networks for constitutive modeling. *Journal of the Mechanics and Physics of Solids* 147, 104277. doi : [10.1016/j.jmps.2020.104277](https://doi.org/10.1016/j.jmps.2020.104277)
- [20] Masi, F., Stefanou, I., 2021. Thermodynamics-based artificial neural networks (TANN) for multiscale modeling of materials with inelastic microstructure. *arXiv preprint arXiv* : [2108.13137](https://arxiv.org/abs/2108.13137)
- [21] Maugin, G.A., Muschik, W., 1994. Thermodynamics with internal variables. Part I. General concepts. *Journal of Non-Equilibrium Thermodynamics* 19, 217–249. doi : [10.1515/jnet.1994.19.3.217](https://doi.org/10.1515/jnet.1994.19.3.217)
- [22] Mianroodi, J.R., Siboni, N.H., Raabe, D., 2021. Teaching solid mechanics to artificial intelligence—a fast solver for heterogeneous materials. *npj Computational Materials* 7, 1–10. doi : [10.1038/s41524-021-00571-z](https://doi.org/10.1038/s41524-021-00571-z)
- [23] Miehe, C., 2002. Strain-driven homogenization of inelastic microstructures and composites based on an incremental variational formulation. *International Journal for Numerical Methods in Engineering* 55, 1285–1322. doi : [10.1002/nme.515](https://doi.org/10.1002/nme.515)
- [24] Mozaffar, M., Bostanabad, R., Chen, W., Ehmann, K., Cao, J., Bessa, M., 2019. Deep learning predicts path-dependent plasticity. *Proceedings of the National Academy of Sciences* 116, 26414–26420. doi : [10.1073/pnas.1911815116](https://doi.org/10.1073/pnas.1911815116)

- [25] Muschik, W., 1993. Fundamentals of nonequilibrium thermodynamics, in : Non-Equilibrium Thermodynamics with Application to Solids. Springer, pp. 1–63. doi : [10.1007/978-3-7091-4321-6_1](https://doi.org/10.1007/978-3-7091-4321-6_1).
- [26] Nitka, M., Combe, G., Dascalu, C., Desrues, J., 2011. Two-scale modeling of granular materials : a DEM-FEM approach. *Granular Matter* 13, 277–281. doi : [10.1007/s10035-011-0255-6](https://doi.org/10.1007/s10035-011-0255-6)
- [27] Wu, L., Noels, L., 2022. Recurrent neural networks (RNNs) with dimensionality reduction and break down in computational mechanics ; application to multi-scale localization step. *Computer Methods in Applied Mechanics and Engineering* 390, 114476. doi : [10.1016/j.cma.2021.114476](https://doi.org/10.1016/j.cma.2021.114476)
- [28] Pinho-da Cruz, J., Oliveira, J., Teixeira-Dias, F., 2009. Asymptotic homogenisation in linear elasticity. Part I : Mathematical formulation and finite element modelling. *Computational Materials Science* 45, 1073–1080. doi : [10.1016/j.commatsci.2009.02.025](https://doi.org/10.1016/j.commatsci.2009.02.025)
- [29] Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks : A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* 378, 686–707. doi : [10.1016/j.jcp.2018.10.045](https://doi.org/10.1016/j.jcp.2018.10.045)
- [30] Rattez, H., Stefanou, I., Sulem, J., Veveakis, M. and Poulet, T., 2018. The importance of Thermo-Hydro-Mechanical couplings and microstructure to strain localization in 3D continua with application to seismic faults. Part II : Numerical implementation and post-bifurcation analysis. *Journal of the Mechanics and Physics of Solids*, 115, 1-29. doi : [10.1016/j.jmps.2018.03.004](https://doi.org/10.1016/j.jmps.2018.03.004)
- [31] Van den Eijnden, A.P., Bésuelle, P., Collin, F., Chambon, R. and Desrues, J., 2017. Modeling the strain localization around an underground gallery with a hydro-mechanical double scale model; effect of anisotropy. *Computers and Geotechnics*, 85, 384-400. doi : [10.1016/j.compgeo.2016.08.006](https://doi.org/10.1016/j.compgeo.2016.08.006)