

# Intégrer les connaissances physiques dans les réseaux de neurones : Application à l'apprentissage de lois de comportement de matériaux à partir de mesures de déformation par fibres optiques

Culture Sciences  
de l'Ingénieur

école  
normale  
supérieure  
paris-saclay

Antoine BENADY - Ludovic CHAMOIN - Emmanuel BARANGER

Édité le  
17/10/2022

*Cette ressource est issue d'une co-publication avec le numéro 109 de La Revue 3EI de juillet 2022 et fait partie du « Dossier Intelligence Artificielle » [7] sur Culture Sciences de l'Ingénieur. Antoine Benady est doctorant au Laboratoire de Mécanique Paris-Saclay (LMPS), Ludovic Chamoïn est Professeur des universités en génie mécanique à l'ENS Paris-Saclay et chercheur au LMPS, et Emmanuel Baranger est chargé de recherche CNRS au LMPS.*

Le travail présenté vise à utiliser des réseaux de neurones pour l'apprentissage de lois de comportement en mécanique des structures. Les réseaux de neurones classiques (au sens où il ne sont pas informés par la physique) présentent les inconvénients de nécessiter des volumes de données importants pour être entraînés, d'avoir des problèmes de généralisation à de nouvelles données et de manquer de garanties en termes d'interprétabilité.

Cette ressource présente différentes façons de tenir compte de la physique dans l'apprentissage afin de pallier ces problèmes, ainsi qu'une stratégie suivie dans le cadre d'un projet de recherche en cours.

## 1 – Introduction

L'endommagement des structures mécaniques est une préoccupation permanente en ingénierie, liée à des problématiques de durabilité et de sécurité. La thématique est actuellement l'objet de diverses activités de recherche ; un exemple caractéristique est le projet ERC DREAM-ON (site web : <https://erc-dreamon.ens-paris-saclay.fr>) dans lequel se place ce travail, qui s'intéresse aux structures mécaniques complexes et vise à aborder les challenges numériques liés au contrôle santé intégré de structures à grande échelle, afin d'aller des matériaux intelligents vers les structures intelligentes, capables de surveiller leur état de façon autonome, de prédire leur évolution à l'aide d'un jumeau numérique et de fonctionner de façon sûre même en mode dégradé (figure 1).

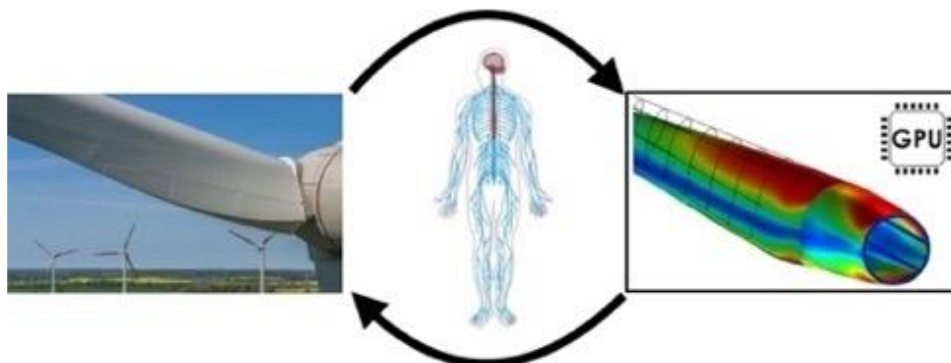


Figure 1 : Contrôle santé d'une structure complexe avec un jumeau numérique.

Des fibres optiques immergées dans la structure à surveiller sont utilisées pour mesurer les déformations dans le matériau (figure 2). L'idée est de créer un jumeau numérique hybride, combinant les modèles physiques (qui représentent une riche histoire des sciences de l'ingénieur, et qui fournissent une forte connaissance a priori) et les techniques d'apprentissage issues de l'IA (ici, les réseaux de neurones), pour le diagnostic, la prédiction et la prise de décision.

Dans le cas d'application visé, on souhaite apprendre des lois de comportement de matériaux grâce aux réseaux de neurones entraînés avec les mesures issues de fibres optiques, sans oublier des siècles de connaissances physiques : cette ressource présente les différentes façons d'intégrer l'a priori physique dans les réseaux de neurones. Dans une première partie, il est question des architectures de réseaux guidées par la physique. Une deuxième partie présente l'entraînement de réseaux avec une fonction coût physique. Dans un troisième temps, l'initialisation du réseau avec un a priori physique est traitée. Enfin on montre qu'il est possible d'utiliser les réseaux de neurones pour apprendre l'erreur d'un modèle physique. Certaines parties seront illustrées par le cas d'application visé. Bien entendu, il est possible de coupler ces techniques entre elles.

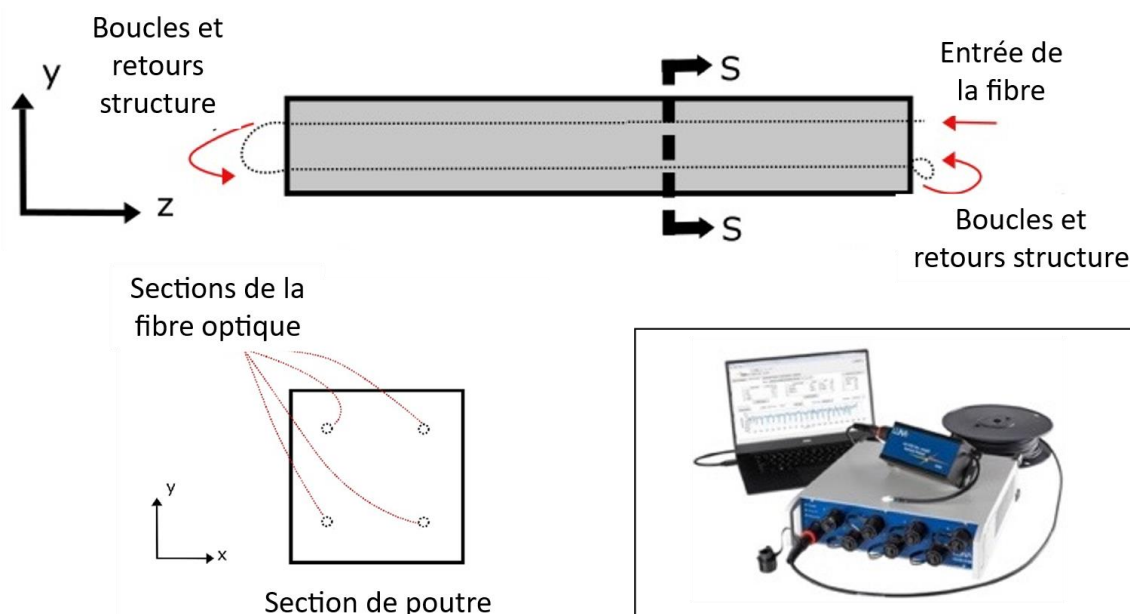


Figure 2 : Cas d'étude sur une poutre en béton et dispositif de mesure par fibre optique (système ODiSI 6108 de LUNA en bas à droite).

## 2 – Architecture du réseau guidée par la physique

Les réseaux de neurones classiques sont réputés pour ne pas offrir de garanties lors de la phase de prédiction, cela étant lié à la faible interprétabilité des résultats. Une façon d'imposer des contraintes physiques dans le réseau est de choisir une architecture qui garantit certaines propriétés. Par exemple, les couches de convolution des CNN (réseaux de neurones convolutifs, souvent utilisés en traitement d'images) présentent l'intérêt de garantir une invariance par translation des résultats.

Une autre possibilité, plus orientée vers l'interprétabilité que vers les garanties de prédiction, consiste à interpréter des grandeurs intermédiaires du réseau comme des quantités physiques [5], les grandeurs intermédiaires pouvant intervenir dans la fonction coût.

Dans le cadre du projet ERC, une architecture garantissant le respect de la thermodynamique des milieux continus est proposée. Le réseau prend en entrée une déformation (grandeur cinématique) et donne en sortie une contrainte (grandeur d'effort), traduisant le comportement du matériau. Un potentiel thermodynamique, permettant de définir la relation de comportement, est utilisé

comme grandeur intermédiaire. L'architecture utilisée est inspirée de [4] mais présente des différences, notamment au niveau de la convexité du potentiel par rapport à la déformation [1] (propriété traduisant le respect du second principe de la thermodynamique). La figure 3 présente un schéma de l'architecture utilisée. L'opération de passage du potentiel thermodynamique à la contrainte est effectuée par différentiation automatique, une technique permettant le calcul de dérivées partielles (utile notamment pour la rétropropagation), et qui est déjà implémentée dans les bibliothèques classiques de Deep Learning (PyTorch, Tensorflow).

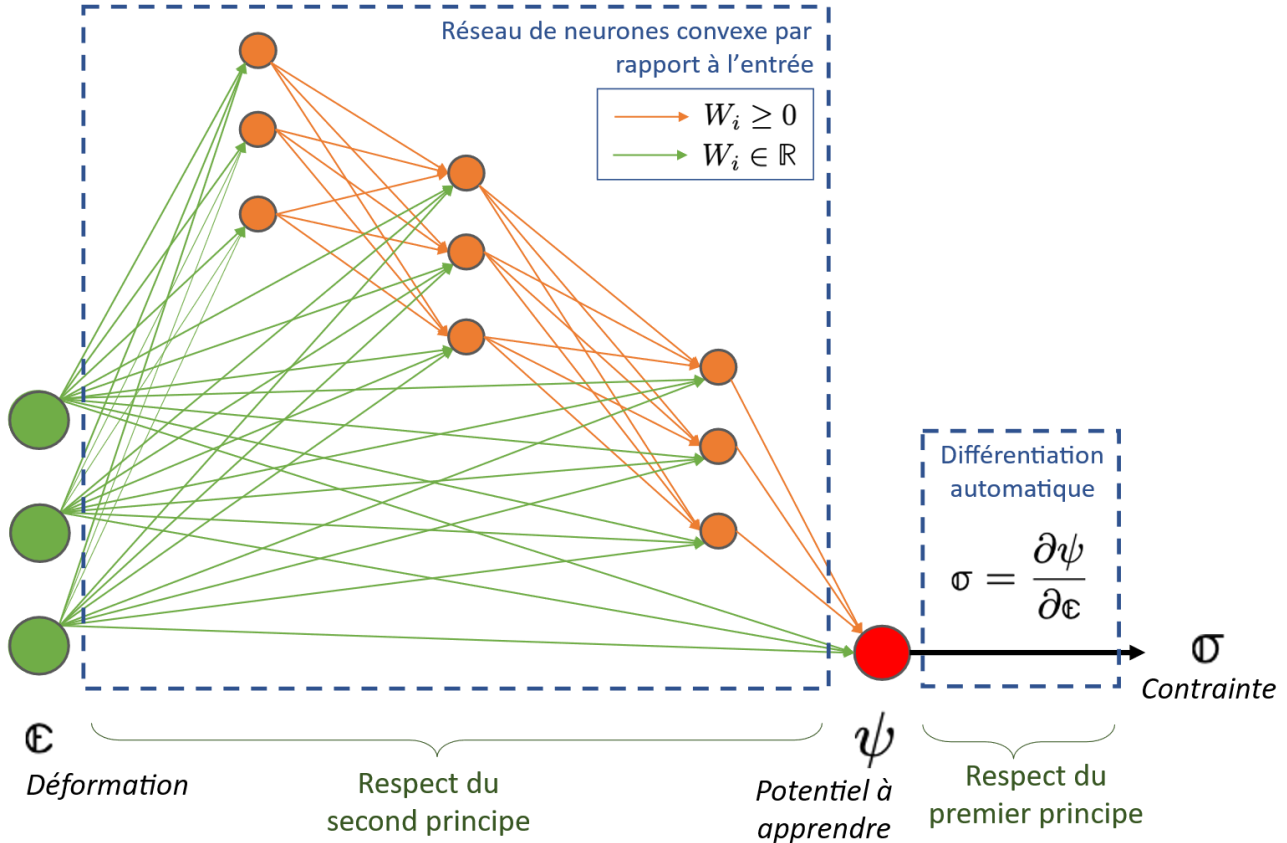


Figure 3 : Architecture utilisée pour apprendre des lois de comportement vérifiant les principes thermodynamiques, avec  $W_i$  les poids du réseau.

### 3 – Utilisation d'une fonction coût physique

Il est également possible de pénaliser le non-respect d'une équation physique en intégrant son résidu dans la fonction coût utilisée pour l'entraînement du réseau. Cette idée a initialement été proposée dans [6] dans le cas particulier de l'apprentissage de l'équation de Navier-Stokes par un réseau de neurones. Pour cela, une fonction coût avec un terme d'écart aux données et un terme de non-respect de l'équation physique est introduite :

$$\mathcal{L} = \mathcal{L}_{\text{données}} + \mathcal{L}_{\text{physique}}$$

$$\mathcal{L}_{\text{données}} = \frac{1}{N_{\text{obs}}} \sum \| \mathbf{Y}_{\text{réseau de neurones}} - \mathbf{Y}_{\text{observations}} \|^2$$

$$\mathcal{L}_{\text{physique}} = \frac{1}{N_{\text{pts}}} \sum |f(u(x,t))|^2$$

avec  $f(u(x, t))$  le résidu de l'équation physique portant sur la grandeur d'intérêt  $u$  aux  $N_{\text{pts}}$  points d'évaluation  $(x, t)$ . On note que les  $N_{\text{obs}}$  points d'évaluation du terme d'écart aux données ne sont pas nécessairement les mêmes que ceux du terme de non-respect de l'équation physique.

La méthode mise en place dans le cadre du projet ERC pour apprendre des lois de comportement s'inspire de cette structure de fonction coût, avec l'utilisation de la fonctionnelle « erreur en

relation de comportement modifiée » (mCRE) pour entraîner le réseau de la figure 3. Cette fonctionnelle, utilisée depuis les années 90 en mécanique dans le contexte de recalage de paramètres [2, 3], se décompose en 2 termes :

- un terme d'écart aux observations ;
- un terme d'erreur de modèle fondé sur l'erreur en relation de comportement. Ce terme quantifie le niveau de non-vérification des relations de comportement pour un couple déplacement-contrainte admissible (vérifiant l'équilibre et les conditions aux limites) et fait intervenir le potentiel  $\psi$  recherché.

#### 4 – Initialisation du réseau par la physique

Un problème couramment rencontré lors de l'entraînement d'un réseau de neurones concerne la quantité de données nécessaire d'autant plus lorsqu'on souhaite apprendre à partir de mesures expérimentales. Une façon d'augmenter la quantité de données à disposition consiste à procéder à un entraînement du réseau en deux temps. Dans un premier temps, un entraînement est réalisé sur des données synthétiques qui sont générées à partir d'un modèle physique a priori, supposé proche de la réalité. À l'issue de ce premier entraînement, le réseau est initialisé proche de la réponse observée dans les données expérimentales et un deuxième entraînement, cette fois-ci avec les données mesurées (en quantité moindre) permet de converger vers un modèle recalé. La démarche est détaillée dans la figure 4.

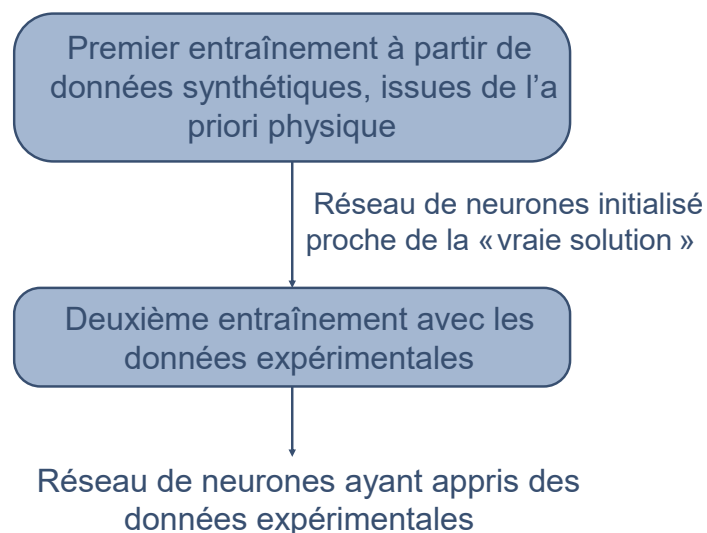


Figure 4 : Prise en compte de l'a priori physique dans l'initialisation du réseau.

#### 5 – Réseau de neurones utilisé pour apprendre l'erreur du modèle physique

Enfin, on trouve aussi dans la littérature l'utilisation de réseaux de neurones pour apprendre l'erreur d'un modèle physique. Si on s'intéresse à une grandeur d'intérêt  $\mathbf{u}$ , on peut la décomposer comme suit :

$$\mathbf{u} = \mathbf{u}_{phy} + \mathbf{u}_{NN}$$

avec  $\mathbf{u}_{phy}$  la grandeur prédite par un modèle physique (qui par définition ne fait que modéliser une partie de la réalité) et  $\mathbf{u}_{NN}$  le complément prédit par un réseau entraîné de manière à minimiser l'écart à des données mesurées.

Par exemple, si on souhaite apprendre une loi de comportement avec un endommagement générant une rigidité du matériau différente en traction et en compression (à gauche de la figure 5), une possibilité est de supposer la loi de comportement élastique linéaire (au milieu sur la figure 5) et de faire apprendre le complément c'est-à-dire la non-linéarité au réseau de neurones (à droite de la figure 5).

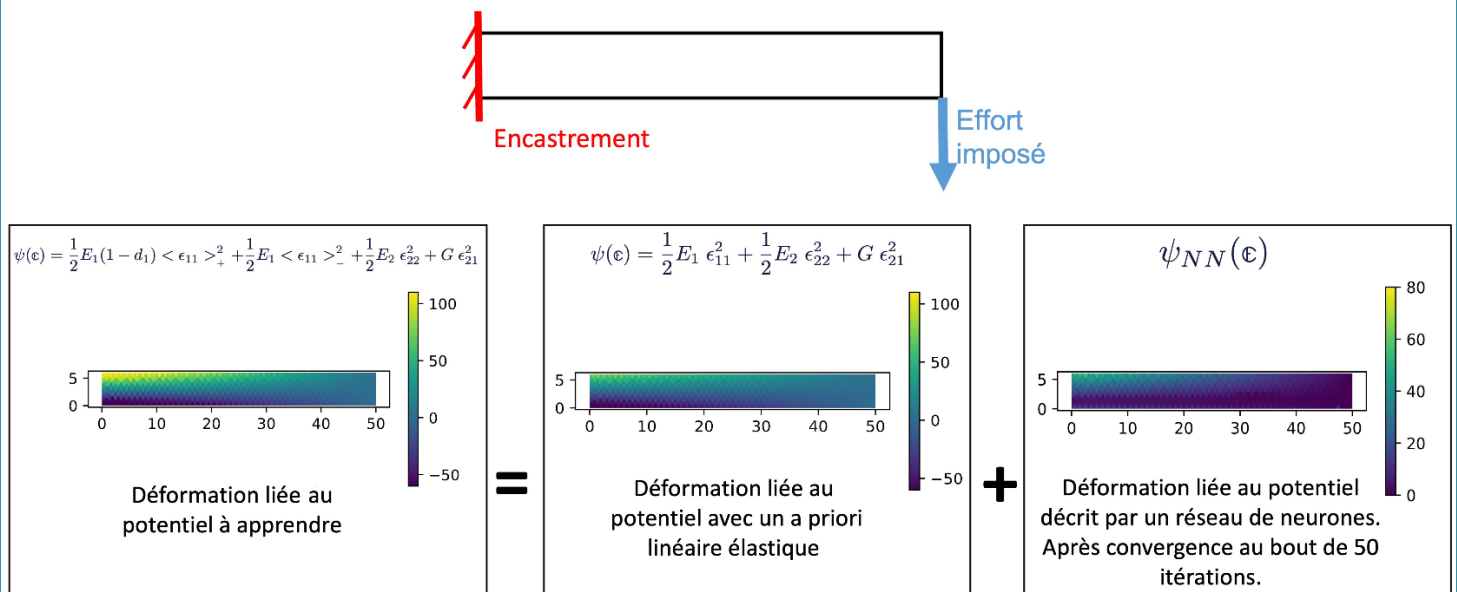


Figure 5 : Illustration de la décomposition d'une grandeur en un terme provenant d'un modèle physique et un terme provenant d'un réseau de neurones.

À gauche une déformation liée à un potentiel à apprendre qui modélise l'endommagement suivant l'axe 1, au milieu une déformation liée à un potentiel décrivant un modèle linéaire élastique, à droite une déformation liée à un potentiel décrit par un réseau de neurones.

## Références :

- [1]: Amos, Brandon and Xu, Lei and Kolter, J. Zico, Input Convex Neural Networks, <https://arxiv.org/abs/1609.07152> ; (2016)
- [2]: Ladevèze, P., Nedjar, D., and Reynier, M. Updating of finite element models using vibration tests. *AIAA Journal*, 32(7) :1485-1491. (1994)
- [3]: Ladevèze, P. A modelling error estimator for dynamical structural model updating. In Elsevier, editor, *Advances in Adaptive Computational Methods in Mechanics*. P. Ladevèze, J.T. Oden eds. (1998)
- [4]: Masi F, Stefanou I, Vannucci P, Maffi-Berthier V, Thermodynamics-Based Artificial Neural Networks for Constitutive Modeling, *Journal of the Mechanics and Physics of Solids* 147 :104277 (2021)
- [5]: Nikhil Muralidhar, Jie Bu, Ze Cao, Long He, Naren Ramakrishnan, Danesh Tafti, and Anuj Karpatne. Physics-Guided Deep Learning for Drag Force Prediction in Dense Fluid-Particulate Systems. *Big Data*, 8(5) :431-449, October 2020.
- [6]: Raissi M, Perdikaris P, Karniadakis G.E., 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 (2019)
- [7]: Dossier Intelligence Artificielle, juin 2022, [https://eduscol.education.fr/sti/si-ens-paris-saclay/ressources\\_pedagogiques/dossier-intelligence-artificielle](https://eduscol.education.fr/sti/si-ens-paris-saclay/ressources_pedagogiques/dossier-intelligence-artificielle)

Ressource publiée sur Culture Sciences de l'Ingénieur : <https://eduscol.education.fr/sti/si-ens-paris-saclay>