

Génération de champs d'ondes 3D large bande par diffusion guidée et opérateurs neuronaux

Niccolo Perrone^{1,2}, Fanny Lehmann², Stefania Fresca³, Filippo Gatti¹

¹ Université Paris-Saclay, CentraleSupélec, ENS Paris-Saclay, CNRS, Laboratoire de Mécanique Paris-Saclay UMR 9026, Gif-sur-Yvette, France, prenom.nom@centralesupelec.fr

² MOX - Department of Mathematics, Politecnico di Milano, Milano, Italy ³ ETH AI Center & Seminar for Applied Mathematics, Zurich, Switzerland, fanny.lehmann@inf.ethz.ch

⁴ Department of Mechanical Engineering - University of Washington, Seattle, WA, USA, sfresca@uw.edu

Résumé — Les simulations 3D haute fidélité sont coûteuses à grande échelle. Les opérateurs neuronaux tels que MIFNO [1] reproduisent bien les basses fréquences, mais échouent sur les hautes à cause de leur biais spectral. Un modèle de diffusion entraîné sur HEMEW^S-3D [2], guidé par MIFNO via Diffusion Posterior Sampling [7], génère des champs d'ondes large bande cohérents. Cette approche préserve phase, amplitude et continuité spatiale sans simulation complète, fournissant un outil scalable pour des scénarios 3D réalistes, utiles en ingénierie parasismique.

Mots clés — Opérateurs neuronaux, modèles de diffusion, simulation sismique, propagation des ondes, apprentissage profond

1 Introduction

Dans cette étude, nous proposons une approche basée sur l'apprentissage profond génératif, pour inférer des solutions d'élastodynamique en trois dimensions $(\mathbf{x}; t) \mapsto \mathbf{v}(\mathbf{x}; t) \in \mathbb{R}^3$, pour un couple source-propriétés mécaniques $(\mathbf{s}(\mathbf{x}; t), \boldsymbol{\mu}(\mathbf{x}))$ quelconque. L'objectif du travail est d'obtenir un outil d'inférence en temps réel, à coût de calcul négligeable, et qui ne sacrifie pas la précision de la solution. Le défi majeur est représenté par le *biais spectral* [3] affectant notre opérateur de Green neuronal \mathbb{G}_ϕ , appelé MIFNO [1], ici employé pour apprendre le mapping $(\mathbf{s}, \boldsymbol{\mu}) \mapsto \mathbf{v}(\mathbf{s}, \boldsymbol{\mu})(\mathbf{x}; t)$, *i.e.*, la solution de l'équation hyperbolique $\mathcal{A}_{(\mathbf{s}, \boldsymbol{\mu})}(\mathbf{v}, \mathbf{w}) = 0, \forall \mathbf{w} \in \mathbb{V}^1$. La Figure 1 illustre la procédure de génération de la base de données HEMEW^S-3D [2], utilisée pour entraîner MIFNO. Le biais spectral se produit dans tout modèle d'apprentissage profond, et traduit la tendance de ces modèles à apprendre prioritairement (et plutôt correctement) des *features* régulières, communes à la base de données, associées aux premiers coefficients de Fourier de la fonction à apprendre. La Figure 2a illustre ce phénomène : l'erreur spectrale du MIFNO devient plus importante à partir de 2 Hz, pour une simulation numérique de référence valable entre 0 et 5 Hz (Figure 2b). Dans le cas en objet, la régularité de ces features se traduit en une solution $\mathbf{v}(\mathbf{x}; t)$ plutôt lisse et à basses fréquences, comme on peut le remarquer à la Figure 2c. Plus le champ de propriétés mécaniques $\boldsymbol{\mu}(\mathbf{x})$ est hétérogène (*e.g.*, le coefficient de variation des propriétés aléatoires est plus grand), plus le biais spectral est prononcé. La Figure 2c montre bien ce biais, mesuré sur la valeur de pic de la solution (PGV), en fonction du coefficient de variation des propriétés mécaniques : le MIFNO \mathbb{G}_ϕ sous-estime systématiquement la PGV, et l'erreur augmente avec l'hétérogénéité du milieu.

Afin de pouvoir corriger ce biais spectral, dans cette étude nous présentons une solution basée sur l'utilisation des modèles de diffusion, inspirés par les travaux de Oommen et al., 2024 [6], portant sur la combinaison d'opérateurs neuronaux et de modèles de diffusion pour la modélisation de phénomènes de turbulence. Plus précisément, nous utilisons un modèle de diffusion conditionnel, entraîné pour corriger les solutions synthétiques générées par le MIFNO, en se basant sur la méthodologie de Diffusion Posterior Sampling (DPS) proposée par Chung

1. \mathbb{V} étant un espace de Hilbert convenable.

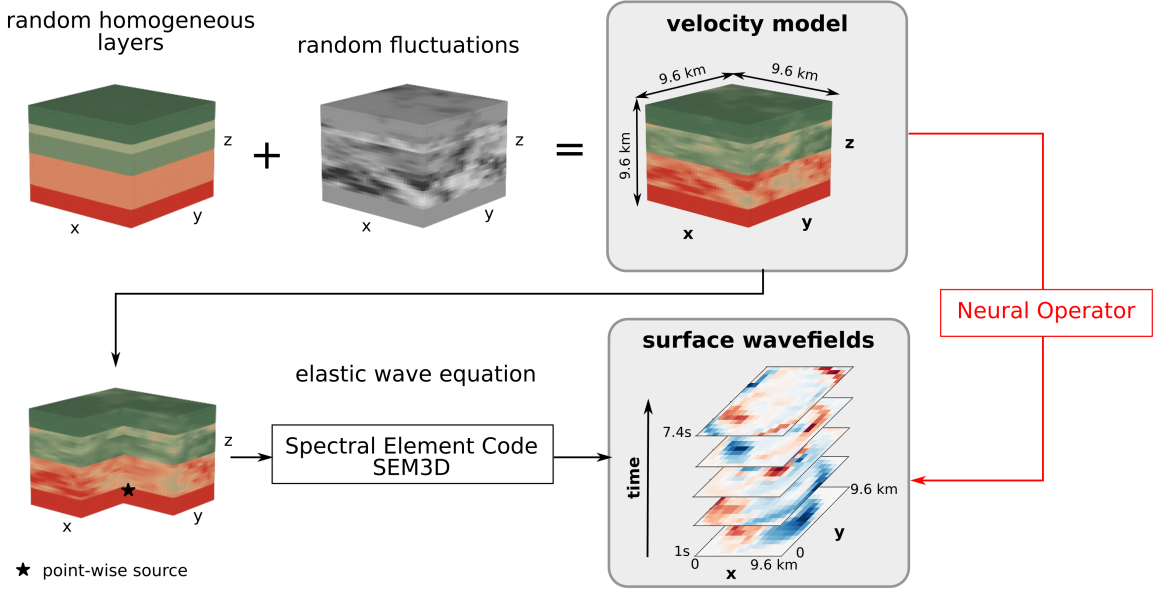


Figure 1 – Schéma de génération de la base de données HEMEW^S-3D [2], utilisée pour entraîner l’opérateur neuronal de Fourier multi-entrées, (acronyme anglais : MIFNO), pour l’élastodynamique en trois dimensions (3D) [1]. Les propriétés mécaniques sont générées sur un domaine ouvert $\Omega \in \mathbb{R}^3$, représentant un morceau de demi-espace à surface libre. Leur distribution spatiale est obtenue en partant d’un modèle multicouches horizontales (à épaisseur et propriétés moyennes aléatoires) et y ajoutant une fluctuation aléatoire selon un modèle de corrélation de von Karman anisotrope, avec longueurs de corrélations aléatoires. La source ponctuelle est localisée aléatoirement dans le domaine, et représentée par un double-couple $\mathbf{m}(\mathbf{x};t)$ (un tenseur d’ordre 2). Les propriétés mécaniques $\boldsymbol{\mu}(\mathbf{x})$ sont représentées par les deux coefficients de Lamé $\lambda(\mathbf{x})$, $\mu(\mathbf{x})$, en élasticité linéaire et isotrope : $\hat{\sigma}(\mathbf{x};t) = \lambda(\mathbf{x}) \operatorname{div}_x \mathbf{v} + \mu(\mathbf{x}) \mathbf{v} \otimes_s \nabla_x$. La solution $\mathbf{u}(\mathbf{x};t)$ est obtenue numériquement, en résolvant l’équation $\operatorname{div}_x(\boldsymbol{\sigma}(\mathbf{v}) + \mathbf{m}) = \rho \partial_t \mathbf{v}$ par éléments finis spectraux (code SEM3D [4]). La solution est enregistrée à la surface libre Γ_n , sur une grille régulière de récepteurs. Les bords latéraux et inférieur du domaine sont équipés de conditions aux limites absorbantes de type Perfectly Matched Layers. Dans [1], seulement la vitesse des ondes de cisaillements $V_S(\mathbf{x}) = \sqrt{\frac{\mu(\mathbf{x})}{\rho}}$ est faite varier. Réimprimé à partir de la Figure 2 de [5].

et al., 2024 [7] et repris par Du et al., 2024 [8]. La stratégie adoptée consiste en l’utilisation du MIFNO comme *guide* pour un modèle de diffusion préentraîné de manière non conditionnée (*i.e.*, l’architecture proposée dans [9]). En phase d’inférence, le modèle de diffusion est guidé par le MIFNO dans son processus markovien de génération du champ \mathbf{v} . Le MIFNO accompagne ce processus, de manière plus ou moins contraignante, en fonction de sa fidélité par rapport à la simulation numérique. Le MIFNO guide le modèle de diffusion en corrigeant son score (modélisé par un réseau de neurones profond, selon l’architecture de [9]), tout en se rapprochant de la condition imposée (la solution basée sur la physique, issue du MIFNO) mais sans perdre la capacité de généralisation du modèle non conditionné. La méthode est décrite dans la Section 2, et les résultats préliminaires sont présentés dans la Section 3.

2 Méthodes

2.1 Multiple-Input Fourier Neural Operator (MIFNO)

Le problème d’élastodynamique 3D dans un domaine $\Omega \subset \mathbb{R}^3$ s’écrit :

$$\mathcal{L}[\boldsymbol{\mu}, \mathbf{v}](\mathbf{x};t) = \mathbf{f}_s(\mathbf{x};t), \quad \forall (\mathbf{x};t) \in \Omega \times]0, T] \quad (1)$$

où $\mathcal{L}[\boldsymbol{\mu}, \mathbf{v}]$ est l’équation des ondes élastiques 3D paramétrée par les paramètres de Lamé $\boldsymbol{\mu}(\mathbf{x})$. $\mathbf{f} \in L^2(\Omega \times]0, T[, \mathbb{R}^3)$ est la distribution de forces volumiques, exprimée comme $\operatorname{div}_x \mathbf{m}(\mathbf{x}) \cdot s(t)$, la divergence d’un tenseur densité de moment en double couple $\mathbf{m} \in H^1(\Omega)^6$, localisé en un point \mathbf{x}_s et dont les angles de direction (*strike*), pendage (*dip*) et glissement (*rake*) sont regroupés dans

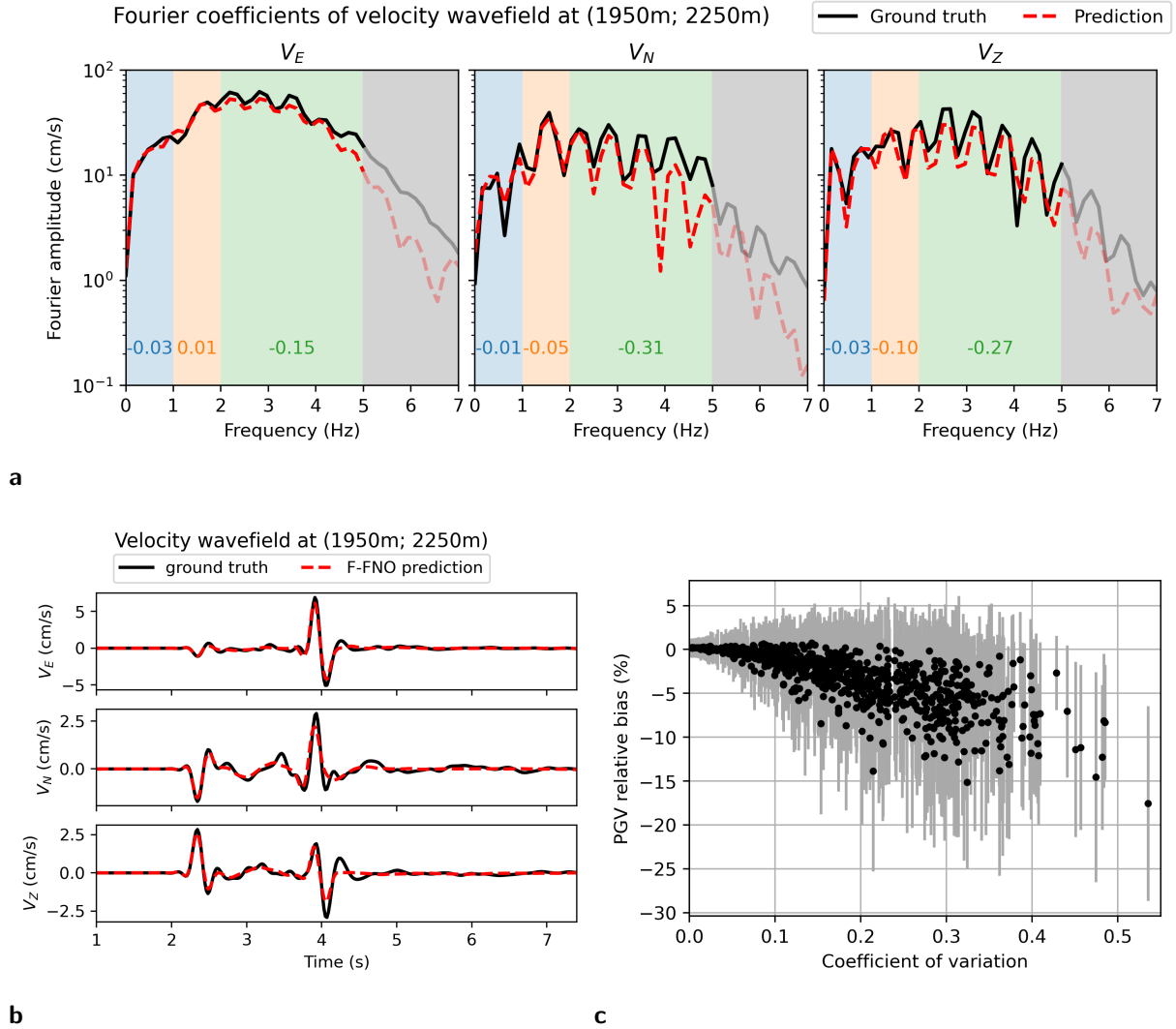


Figure 2 – (a) Illustration du biais spectral affectant l’opérateur neuronal MIFNO [1] par comparaison de l’amplitude du spectre de Fourier généré (rouge) et de référence (noir), selon les trois directions de l’espace (Est-ouest, Nord-sud, vertical Z). Figure reprise de la Figure 4 dans [5]. (b) Exemple de solution $\mathbf{v}(\mathbf{x}_m, t)$ générée par le MIFNO \mathbb{G}_ϕ (rouge), comparée à la solution de référence obtenue par simulation numérique (*ground truth* en noir). Figure reprise de la Figure 3 de [5]. (c) Biais de prédiction de la Peak Ground Velocity (PGV, correspondante à $\sup_{t \in [0, T]} \|\mathbf{v}(\cdot, t)\|$) estimé par \mathbb{G}_ϕ en fonction du coefficient de variation choisi pour la génération du champ aléatoire des propriétés mécaniques $(\lambda(\mathbf{x}), \mu(\mathbf{x}))$. Figure réimprimé de la Figure 13 dans [5].

le vecteur $\phi_s \in \mathbb{R}^3$. $s(t) \in C^0([0, T])$ représente l’amplitude temporelle de la source au cours de la durée T . Enfin, $\mathbf{v} \in \mathcal{C}^1(\Omega \times]0, T[, \mathbb{R}^3)$ est le champ de vitesse particulaire, solution du problème. Le MIFNO est un modèle neuronal substitut $\mathbb{G}_\phi : (\boldsymbol{\mu}, \mathbf{x}_s, \phi_s) \mapsto \mathbf{v}|_{\partial\Omega_{top}}$ de l’opérateur de Green, paramétré par des poids ϕ , prédisant le champ de vitesse \mathbf{v} à la surface de la Terre $\partial\Omega_{top}$, à partir de la seule connaissance de la géologie $\boldsymbol{\mu}$ et des caractéristiques de la source \mathbf{x}_s, ϕ_s ². Comme montré dans la Figure 3, le MIFNO est alimenté par des champs de propriétés mécaniques 3D complexes et hétérogènes $\boldsymbol{\mu}(\mathbf{x})$, ainsi que par des données décrivant les caractéristiques de la source sismique ponctuelle (\mathbf{x}_s, ϕ_s) . La dépendance temporelle $s(t)$ est fixée et identique pour toutes les simulations de la base de données HEMEW^S-3D [2]. L’architecture dorsale du MIFNO repose sur les couches de Fourier factorisées (F-Fourier, Factorized Fourier Neural Operator, F-FNO) proposés par Tran et al., [10]. Des détails supplémentaires sont fournis dans la Figure 3. Le MIFNO s’appuie sur le F-FNO [10] pour apprendre le champ de vitesse obtenu via des simulations numériques basées sur la physique. Les simulations de référence ont été calculées avec le logiciel open-source de calcul haute performance SEM3D [4], développé par le CEA,

2. <https://github.com/lehmannfa/MIFNO>

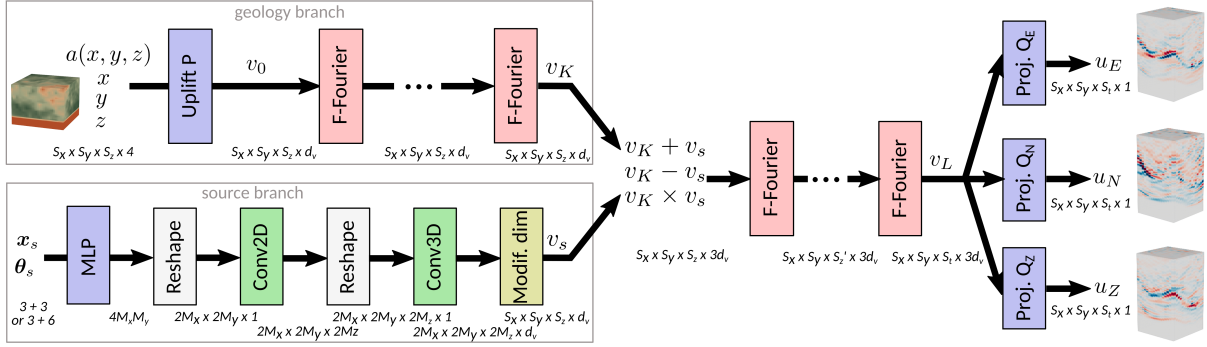


Figure 3 – Le MIFNO est constitué d’une *branche géologie* qui encode la géologie au moyen de couches de Fourier factorisées (F-Fourier), et d’une *branche source* qui transforme le vecteur des paramètres de source (\mathbf{x}_s, ϕ_s) en une variable 4D v_s correspondant aux dimensions de la sortie v_K de la *branche géologie*. Les sorties de chaque branche sont concaténées après des opérations mathématiques élémentaires, puis les couches de Fourier factorisées restantes sont appliquées. Les blocs d’élévation P et de projection Q_E, Q_N, Q_Z sont identiques à ceux du F-FNO : ils intègrent le champ géologique et projettent l’espace latent dans l’espace solution, respectivement. Reproduit de [1].

CentraleSupélec, le CNRS et l’Institut de Physique du Globe de Paris³. Le MIFNO est entraîné avec la base de données ouverte HEMEW^S-3D [2] afin de minimiser l’erreur entre les champs de vitesse prédits à la surface $\hat{\mathbf{v}}$ et ceux simulés \mathbf{v} . HEMEW^S-3D comprend 30 000 simulations sismiques SEM3D dans diverses géologies hétérogènes $\boldsymbol{\mu}(\mathbf{x})$, incluant des positions de source \mathbf{x}_s et des orientations ϕ_s tirées aléatoirement (voir Figure 1). Lehmann et al., 2025 [1] ont montré que le MIFNO produit des accélérogrammes jugés bons à excellents selon les métriques de qualité (*Goodness-Of-Fit*, GOF) de [11], comparés aux signaux SEM3D de référence. La précision des temps d’arrivée et de la propagation des fronts d’ondes est particulièrement notable, avec 80% des prédictions présentant un score de phase GOF exceptionnel supérieur à 8. De plus, les fluctuations d’amplitude sont jugées satisfaisantes pour 87% des prédictions (avec une enveloppe GOF dépassant 6) et excellentes pour 28% des cas. Le MIFNO démontre une remarquable capacité de généralisation pour des sources situées en dehors du domaine d’entraînement, et montre également de solides performances de généralisation lorsqu’il est appliqué à un scénario géologique réel complexe caractérisé par des structures en chevauchement. De plus, lorsqu’on se concentre sur des régions d’intérêt spécifiques, l’apprentissage par transfert s’avère efficace pour améliorer la précision des prédictions sans coûts significatifs. Cependant, la précision des prédictions se dégrade lorsque la fréquence augmente. Cela est dû aux phénomènes physiques complexes associés aux caractéristiques haute fréquence, et reflète également le biais bien connu des réseaux neuronaux, selon lequel les structures à petite échelle sont plus difficiles à capturer. Cette limitation peut freiner l’utilisation du MIFNO dans les applications sismiques de génie nucléaire.

2.2 Modèle de diffusion

Un modèle de diffusion \mathbb{D}_θ exploite l’échantillonnage Monte Carlo par chaînes de Markov pour générer des données réalistes (dans notre cas, le champ de solution $\mathbf{v}(t) \sim p(\mathbf{V}_0(t); 0)$, correspondant à la distribution empirique induite sur HEMEW^S-3D), en partant d’un échantillon tiré d’une distribution normale $\mathbf{v}(t) \sim p(\mathbf{V}_\tau; \mathcal{T}) = \mathcal{N}(\mathbf{0}, \sigma_{\max} \mathbb{I})$ [12], qu’on ne peut pas distinguer d’un échantillon tiré de $p(\mathbf{V}_0(t); 0)$, car $\sigma_{\max} \gg \sigma_{\text{HEMEW}^S\text{-3D}}$. La chaîne de Markov d’un modèle de diffusion se compose de deux processus discrets : le *processus direct*, ajoutant un bruit aux données originales $\mathbf{v}_0(t)$ jusqu’à obtenir un bruit blanc, et le *processus inverse* débruitant les données bruitées $\mathbf{v}_\tau(t) \sim \mathcal{N}(\mathbf{0}, \sigma_{\max} \mathbb{I})$ vers un nouvel échantillon $\hat{\mathbf{v}}(t) \sim p(\mathbf{v}_0(t); 0)$ ressemblant aux données originales (voir Figure 4). En corrompant des données propres $\mathbf{v}_0(t)$ via un processus de bruitage direct, dont les conditionnelles $p_{\tau|0}(\mathbf{V}_\tau | \mathbf{v}_0; \tau)$ sont connues (les noyaux de transition gaussiens sont paramétrés par la densité proxy $q_{(a(\tau), \sigma(\tau))}(\mathbf{V}_\tau | \mathbf{v}_{\tau-1})$), \mathbb{D}_θ représente,

3. <https://github.com/sem3d/sem>

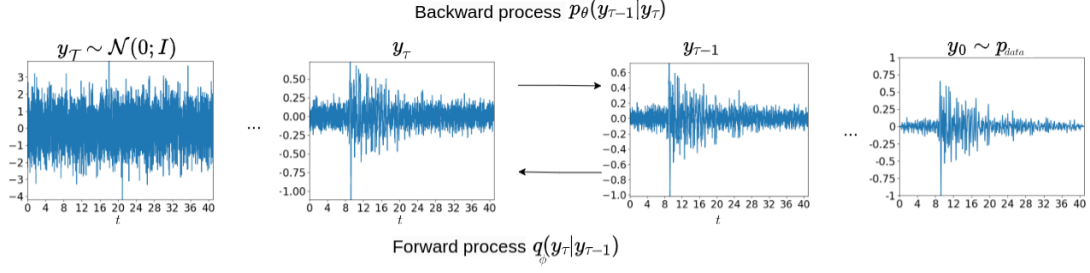


Figure 4 – Diagramme du processus de diffusion. Le processus direct est associé à un noyau de transition $q_{(a(\tau), \sigma(\tau))}(\mathbf{v}_\tau | \mathbf{v}_{\tau-1})$, tandis que le processus inverse est associé au noyau de transition $p_\theta(\mathbf{v}_\tau | \mathbf{v}_{\tau+1})$.

dans sa version *denoising score matching* [13], un modèle paramétrique du score $\nabla_{\mathbf{V}_\tau} \ln p(\mathbf{V}_\tau; \tau)$ (paramètres θ) qui guide le processus de débruitage, et le flot de la distribution de probabilité $p(\mathbf{V}_\tau; \tau)$ vers $p(\mathbf{V}_0; 0)$.

En particulier, suivant [14], le processus de bruitage de données de la base HEMEW^S-3D suit le processus markovien défini par l'équation suivante :

$$d\mathbf{V}_\tau = \frac{\dot{a}(\tau)}{a(\tau)} \mathbf{V}_\tau d\tau + a(\tau) \sqrt{2\sigma(\tau)\dot{\sigma}(\tau)} d\mathbf{W}_\tau \quad \mathbf{v}_0 \sim p(\mathbf{V}_0; 0) \quad (2)$$

avec \mathbf{W}_τ un processus de Wiener standard, $\tau \mapsto \sigma(\tau)$ et $\tau \mapsto a(\tau) \in \mathbb{R}_0^+$ les fonctions de planification du bruitage et de mise à l'échelle des données, respectivement. $\frac{\dot{a}(\tau)}{a(\tau)} \mathbf{V}_\tau$ est appelé *drift*, $a(0) = 1$ et $\sigma(0) = 0$. Le processus (2) a pour moments conditionnels $\mathbb{E}[\mathbf{V}_\tau | \mathbf{v}_0] = a(\tau) \mathbf{v}_0$ et $\mathbb{C}[\mathbf{V}_\tau | \mathbf{v}_0] = a^2(\tau) \sigma^2(\tau) \mathbb{I}$. En intégrant (2), le processus de bruitage se planifie selon le processus gaussien $\mathbf{V}_\tau = a(\tau) \left(\mathbf{V}_0 + \int_0^\tau \sqrt{2\sigma(s)\dot{\sigma}(s)} d\mathbf{W}_s \right)$. En inversant l'axe temporel de la diffusion, $\tau \mapsto \mathcal{T} - \tau$, on obtient le processus génératif de $\mathbf{V}_\tau = \mathbf{V}_{\mathcal{T}-\tau}$, qui s'écrit alors de la même manière que (2), mais avec un *drift* corrigé par le *score*, selon l'expression suivante [15, 16] :

$$d\mathbf{V}_\tau = \left(-\frac{\dot{a}(\tau)}{a(\tau)} \mathbf{V}_\tau + 2a^2(\tau)\sigma(\tau)\dot{\sigma}(\tau) \nabla_{\mathbf{v}} \ln p(\mathbf{V}_\tau, \tau) \right) d\tau + a(\tau) \sqrt{2\sigma(\tau)\dot{\sigma}(\tau)} d\mathbf{W}_\tau \quad (3)$$

Afin de débruiter des nouveaux échantillons, selon le processus inverse (3), la fonction de score est calculé grâce à la formule de Tweedie [17] :

$$\nabla_{\mathbf{v}} \ln p(\mathbf{v}, \tau) = \frac{a(\tau) \mathbb{E}[\mathbf{V}_0 | \mathbf{v}_\tau = \mathbf{v}] - \mathbf{v}}{a^2(\tau) \sigma^2(\tau)}, \quad \tau > 0 \quad (4)$$

Dans la formulation (4), proposée par [14], le score représente, à un facteur multiplicatif près, la quantité de bruit à retirer du signal bruité \mathbf{v}_τ pour converger, en moyenne, vers un échantillon $\mathbf{v}_0 \sim p(\mathbf{V}_0; 0)$. En pratique, le terme $\mathbb{E}[\mathbf{V}_0 | \mathbf{v}_\tau = \mathbf{v}]$ est remplacé par un réseau de neurones $\mathbb{D}_\theta(\mathbf{v}_\tau; \tau)$ [13]. En remplaçant (4) dans (3), et avec les changements de variable $a(\tau) = 1$ et $\tau = \sigma(\tau)$, suggéré dans [14], le processus d'échantillonnage de nouveaux champs \mathbf{v} peut être écrit de manière *déterministe*, avec une équation différentielle ordinaire (EDO), obtenue en retirant le processus de Wiener dans (3). La distribution marginale de la solution de l'EDO déterministe (appelée *probability flow ODE*) est la même que celle obtenue en résolvant l'équation de Fokker-Planck associée à (3). L'EDO utilisée pour le débruitage déterministe est :

$$\frac{d\mathbf{v}}{d\sigma}(\sigma) = \frac{\mathbb{D}_\theta(\mathbf{v}(\sigma); \sigma) - \mathbf{v}(\sigma)}{\sigma}, \quad \sigma > 0 \quad (5)$$

Afin de bien moduler le niveau de bruit σ en phase de débruitage, le modèle de diffusion \mathbb{D}_θ suggérée par [14] s'écrit :

$$\mathbb{D}_\theta(\mathbf{v}; \sigma) = c_{\text{skip}}(\sigma) \mathbf{v} + c_{\text{out}}(\sigma) \mathbf{F}_\theta(c_{\text{in}}(\sigma) \mathbf{v}; c_{\text{noise}}(\sigma)), \quad (6)$$

Dans notre cas, le réseau de neurones \mathbf{F}_θ correspond au ResUNet conçu dans [9] et adapté à la taille des données provenant de HEMEW^S-3D. La formulation (6) permet de mieux contrôler

l'importance relative des termes de bruit et de signal dans l'entrée du réseau de neurones \mathbf{F}_θ , en fonction du niveau de bruit σ . Les fonctions $c_{\text{in}}(\sigma)$, $c_{\text{out}}(\sigma)$, $c_{\text{noise}}(\sigma)$ et $c_{\text{skip}}(\sigma)$ sont définies dans [14]⁴. Les paramètres θ du modèle de diffusion \mathbb{D}_θ sont obtenus en minimisant la fonction coût *score-matching* suivante [14] :

$$\begin{aligned} \mathcal{L}(\theta) &= \mathbb{E}_{\sigma \sim p_\sigma, \mathbf{N} \sim \mathcal{N}(\mathbf{0}, \mathbb{I}), \mathbf{V}_0} \left[\lambda(\sigma) \left\| \mathbb{D}_\theta(\mathbf{V}_0 + \sigma \mathbf{N}; \sigma) - \mathbf{V}_0 \right\|_2^2 \right] = \\ &= \mathbb{E}_{\sigma, \mathbf{N}, \mathbf{V}_0} \left[\lambda(\sigma) c_{\text{out}}^2(\sigma) \left\| \mathbf{F}_\theta(c_{\text{in}}(\sigma)(\mathbf{V}_0 + \mathbf{N}); c_{\text{noise}}(\sigma)) - \frac{\mathbf{V}_0 - c_{\text{skip}}(\sigma)(\mathbf{V}_0 + \mathbf{N})}{c_{\text{out}}(\sigma)} \right\|_2^2 \right] \end{aligned} \quad (7)$$

L'équation (7) implique le fait que minimiser la fonction coût $\mathcal{L}(\theta)$ revient à trouver un point stable de l'EDO (5). Karras et al., 2022 [14] proposent de prendre un niveau de bruit tiré aléatoirement et suivant une log-normale, *i.e.*, $\ln \sigma \sim \mathcal{N}(-1.2, 1.2^2)$. $\lambda(\sigma)$ est choisi en fonction du niveau de pondération de \mathbf{F}_θ , *i.e.*, $\lambda(\sigma) = 1/c_{\text{out}}(\sigma)$, avec une distribution log-uniforme des niveaux de bruit $p_\sigma(\ln \sigma) = \mathcal{U}(\ln \sigma_{\text{min}}, \ln \sigma_{\text{max}})$ $\sigma_{\text{min}} = 0.001$, $\sigma_{\text{max}} = 80$.

Nous notons d'abord que l'opérateur neuronal $\mathbf{v}_{\text{MIFNO}}$ est ajusté (*fine-tuned*) sur des cibles passe-bas, obtenues en appliquant un filtre de Butterworth d'ordre quatre, de fréquence de coupure 2.5 Hz, aux champs de vitesse de vérité-terrain. Durant l'entraînement par diffusion, nous utilisons donc comme entrée de conditionnement la prédiction MIFNO correspondante filtrée en passe-bas, $\mathbf{v}_{\text{LP}}(t) = \mathcal{B}_{2.5\text{Hz}} * \mathbf{v}_{\text{MIFNO}}(t)$ ⁵. Nous définissons le champ résiduel large bande par $\mathbf{r} = \mathbf{v}_0 - \mathbf{v}_{\text{LP}}$, et nous notons p_{res} la distribution empirique de tels résidus sur l'ensemble d'entraînement. La fonction coût utilisée pour entraîner le débruiteur conditionnel \mathbb{D}_θ est alors donnée par

$$\mathcal{L}(\mathbb{D}_\theta) = \mathbb{E}_{\sigma, \mathbf{r} \sim p_{\text{res}}} \mathbb{E}_{\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbb{I})} \left[\lambda(\sigma) \left\| \mathbb{D}_\theta(\mathbf{r} + \boldsymbol{\eta}, \sigma \mid \mathbf{v}_{\text{LP}}) - \mathbf{r} \right\|_2^2 \right] \quad (8)$$

3 Résultats

Figure 5 présente des résultats préliminaires obtenus avec la méthodologie décrite dans la Section 2. On observe que le modèle de diffusion conditionnel améliore significativement la prédiction du MIFNO sur les trois composantes. La Figure 5(b) compare le champ d'onde à l'instant qui réalise la vitesse maximale (PGV), avec la solution de référence. On constate que le modèle de diffusion conditionnel corrige efficacement le biais spectral du MIFNO, en particulier dans les zones où la PGV est sous-estimée par le MIFNO. Par construction, cette formulation résiduelle empêche le modèle de diffusion de réapprendre la géométrie basse fréquence déjà capturée par l'opérateur neuronal et concentre plutôt sa capacité sur un raffinement conditionnel haute fréquence.

4 Conclusions

Dans cette étude, nous avons présenté une méthodologie combinant un opérateur neuronal MIFNO et un modèle de diffusion conditionnel pour améliorer la modélisation de la propagation des ondes sismiques en 3D dans des géologies hétérogènes. Le MIFNO, bien qu'efficace pour capturer les caractéristiques globales de la propagation des ondes, souffre d'un biais spectral qui limite sa capacité à modéliser les hautes fréquences. En intégrant un modèle de diffusion conditionnel, nous avons pu corriger ce biais et améliorer la précision des prédictions, en particulier pour les composantes horizontales du champ de vitesse. Les résultats préliminaires montrent que cette approche hybride est prometteuse pour surmonter les limitations des opérateurs neuronaux dans la modélisation sismique, ouvrant la voie à des applications plus précises dans le domaine du génie sismique et de la sécurité nucléaire. Des travaux futurs incluront une évaluation plus approfondie de la performance de cette méthodologie sur un ensemble plus large de scénarios

4. $c_{\text{textskip}}(\sigma) = \frac{\sigma_{\text{data}}^2}{\sigma^2 + \sigma_{\text{data}}^2}$, $c_{\text{in}}(\sigma) = \frac{1}{\sqrt{\sigma^2 + \sigma_{\text{data}}^2}}$, $c_{\text{out}}(\sigma) = \frac{\sigma \sigma_{\text{data}}}{\sqrt{\sigma^2 + \sigma_{\text{data}}^2}}$, $c_{\text{noise}}(\sigma) = \frac{1}{4} \log \sigma$, $\lambda(\sigma) = \frac{1}{\sigma^2} + \frac{1}{\sigma_{\text{data}}^2}$

5. * indique le produit de convolution.

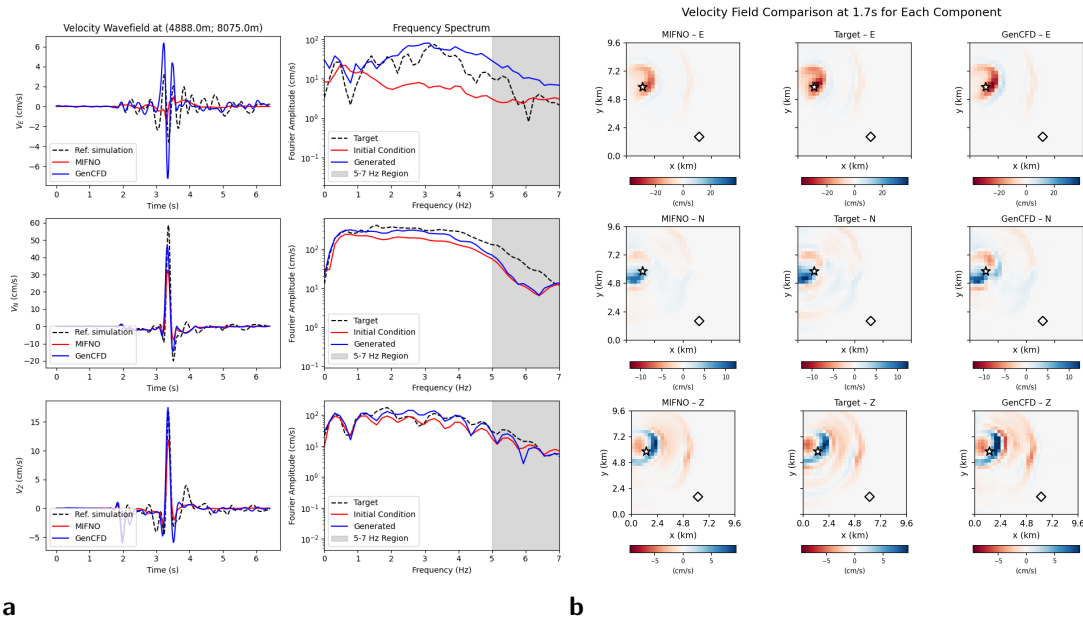


Figure 5 – (a) Comparaison des sismogrammes synthétiques à la station générés par le MIFNO (rouge), le modèle de diffusion conditionnel (bleu) et la solution de référence SEM3D (noir). (b) Champ de vitesse au temps $t=1.7s$ à la surface, comparant la solution de référence SEM3D (gauche), la prédiction MIFNO (milieu) et la prédiction par le modèle de diffusion conditionnel (droite).

sismiques et l’exploration de différentes architectures de modèles de diffusion pour optimiser davantage les résultats.

Références

- [1] Lehmann, F. ; Gatti, F. ; Clouteau, D. *Multiple-Input Fourier Neural Operator (MIFNO) for Source-Dependent 3D Elastodynamics*. Journal of Computational Physics 2025, 527, 113813. <https://doi.org/10.1016/j.jcp.2025.113813>.
- [2] Lehmann, F. ; Gatti, F. ; Bertin, M. ; Clouteau, D. Synthetic Ground Motions in Heterogeneous Geologies from Various Sources : The HEMEW S -3D Database. Earth Syst. Sci. Data 2024, 16 (9), 3949-3972. <https://doi.org/10.5194/essd-16-3949-2024>.
- [3] Rahaman, N. ; Baratin, A. ; Arpit, D. ; Draxler, F. ; Lin, M. ; Hamprecht, F. A. ; Bengio, Y. ; Courville, A. *On the Spectral Bias of Neural Networks*. In Proceedings of the 36 th International Conference on Machine Learning ; Long Beach, California, 2019 ; Vol. PMLR 97. <https://proceedings.mlr.press/v97/rahaman19a.html>
- [4] Touhami, S. ; Gatti, F. ; Lopez-Caballero, F. ; Cottreau, R. ; de Abreu Corrêa, L. ; Aubry, L. ; Clouteau, D. SEM3D : A 3D High-Fidelity Numerical Earthquake Simulator for Broadband (0-10 Hz) Seismic Response Prediction at a Regional Scale. Geosciences 2022, 12 (3), 112. <https://doi.org/10.3390/geosciences12030112>.
- [5] Lehmann, F. ; Gatti, F. ; Bertin, M. ; Clouteau, D. 3D Elastic Wave Propagation with a Factorized Fourier Neural Operator (F-FNO). Computer Methods in Applied Mechanics and Engineering 2024, 420 (116718), 1-27. <https://doi.org/10.1016/j.cma.2023.116718>.
- [6] Oommen, V. ; Bora, A. ; Zhang, Z. ; Karniadakis, G. E. *Integrating Neural Operators with Diffusion Models Improves Spectral Representation in Turbulence Modeling*. arXiv September 13, 2024. <http://arxiv.org/abs/2409.08477>.
- [7] Chung, H. ; Kim, J. ; Mccann, M. T. ; Klasky, M. L. ; Ye, J. C. *Diffusion Posterior Sampling for General Noisy Inverse Problems*. arXiv May 20, 2024. <https://doi.org/10.48550/arXiv.2209.14687>.
- [8] Du, P. ; Parikh, M. H. ; Fan, X. ; Liu, X.-Y. ; Wang, J.-X. *CoNF \hat{i} LD : Conditional Neural Field Latent Diffusion Model Generating Spatiotemporal Turbulence*. arXiv March 15, 2024. <http://arxiv.org/abs/2403.05940>.

- [9] Molinaro, R.; Lanthaler, S.; Raonic, B.; Rohner, T.; Armegioiu, V.; Wan, Z. Y.; Sha, F.; Mishra, S.; Zepeda-Nunez, L. *Generative AI for Fast and Accurate Statistical Computation of Fluids*. arXiv September 27, 2024. <https://doi.org/10.48550/arXiv.2409.18359>.
- [10] Tran, A.; Mathews, A.; Xie, L.; Ong, C. S. *Factorized Fourier Neural Operators*. arXiv March 2, 2023. <http://arxiv.org/abs/2111.13802> (accessed 2024-02-27).
- [11] Kristekova, M., Kristek, J., and Moczo, P. *Time-frequency misfit and goodness-of-fit criteria for quantitative comparison of time signals*. *Geophysical Journal International*, 178(2) :813–825, August 2009. 1365-246X.
- [12] J. Ho, A. Jain, and P. Abbeel. *Denoising Diffusion Probabilistic Models*, 33 :6840-6851, 2020. <https://arxiv.org/abs/2006.11239>.
- [13] Song, Y.; Sohl-Dickstein, J.; Kingma, D. P.; Kumar, A.; Ermon, S.; Poole, B. Score-Based Generative Modeling through Stochastic Differential Equations. arXiv February 10, 2021. <http://arxiv.org/abs/2011.13456>.
- [14] Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. *Advances in Neural Information Processing Systems*, 35 :26565-26577, 2022. <https://doi.org/10.48550/arXiv.2206.00364>.
- [15] Brian D.O. Anderson. *Reverse-time diffusion equation models*. *Stochastic Processes and their Applications*, 12(3) :313-326, 1982. [https://doi.org/10.1016/0304-4149\(82\)90051-5](https://doi.org/10.1016/0304-4149(82)90051-5).
- [16] Ulrich G. Haussmann and Etienne Pardoux. *Time reversal of diffusions*. *The Annals of Probability*, pages 1188-1205, 1986. <https://doi.org/10.1214/aop/1176992362>.
- [17] Bradley Efron. *Tweedie’s formula and selection bias*. *Journal of the American Statistical Association*, 106(496) :1602-1614, 2011. <https://doi.org/10.1198/jasa.2011.tm11181>