

Réseau de neurones et déconvolution - Application à la simulation d'un amplificateur guitare

Pt.I - Enoncé du problème

Jean-Loup Pecquais

2024-10-16



Figure 1: *image générée par IA*

À l'heure où j'écris cet article, la simulation d'équipements "audiomusicaux" repose de plus en plus sur l'entraînement de réseaux de neurones ¹, souvent grâce à l'apprentissage profond

¹En apprentissage automatique (*machine learning*), un réseau de neurones (également appelé réseau de neu-

(*deep learning*)². Cette méthode est particulièrement séduisante, car elle ne nécessite aucune connaissance préalable sur l'équipement à modéliser. On parle alors de simulation "boîte noire" (black box) : seul le résultat final compte, peu importe comment il est produit.

Me concernant, je suis intrigué par la capacité de cette technologie à fournir des simulations convaincantes d'amplificateurs pour guitare électrique. Plus particulièrement, la modélisation de l'étage d'amplification de puissance est cruciale pour remplacer un amplificateur sur scène sans perdre son effet sur le signal. Rappelons que les amplificateurs de guitare sont, par conception, particulièrement non linéaires, que ce soit dans leur réponse en fréquence ou dans l'ajout de distorsion harmonique.

Le second aspect intéressant dans l'apprentissage profond est la possibilité de **modéliser ses équipements**. Voici alors la situation idéale : laissez votre ampli à lampe préféré à la maison, faites en une simple "emprunte" et emmener seulement une petite "boîte noire" numérique sur scène. Sur le papier, l'idée semble convaincante.

Les difficultés commencent lorsque l'on souhaite réaliser ses propres captures. En règle générale, la procédure est similaire pour la plupart des fabricants de telles simulations : on émet un signal dans notre matériel, on enregistre la sortie du périphérique, puis on entraîne le réseau de neurones en lui injectant les deux signaux, soit sur un appareil spécifique, soit en ligne, sur une plateforme telle que *Google Colab*, par exemple.

Cependant, il faut bien distinguer plusieurs types d'équipements :

1. Les périphériques utilisant des niveaux d'E/S ligne, micro ou instrument
2. Les périphériques utilisant des niveaux d'E/S bien plus important, comme un amplificateur de puissance

Dans le premier cas, l'acquisition pose moins de difficulté. On sort le signal de son interface audio, on récupère la sortie du périphérique dans une entrée en ligne ou un préamplificateur, et le tour est joué.

rones artificiels ou réseau neuronal) est un modèle inspiré par la structure et le fonctionnement des réseaux neuronaux biologiques dans les cerveaux des animaux. Les réseaux de neurones artificiels sont utilisés pour diverses tâches, notamment la modélisation prédictive, le contrôle adaptatif et la résolution de problèmes en intelligence artificielle. Ils peuvent apprendre de l'expérience et tirer des conclusions à partir d'un ensemble complexe et apparemment non lié d'informations.

²L'apprentissage profond est une sous-catégorie de l'apprentissage automatique qui utilise des réseaux de neurones avec de nombreuses couches pour modéliser des structures de données complexes.

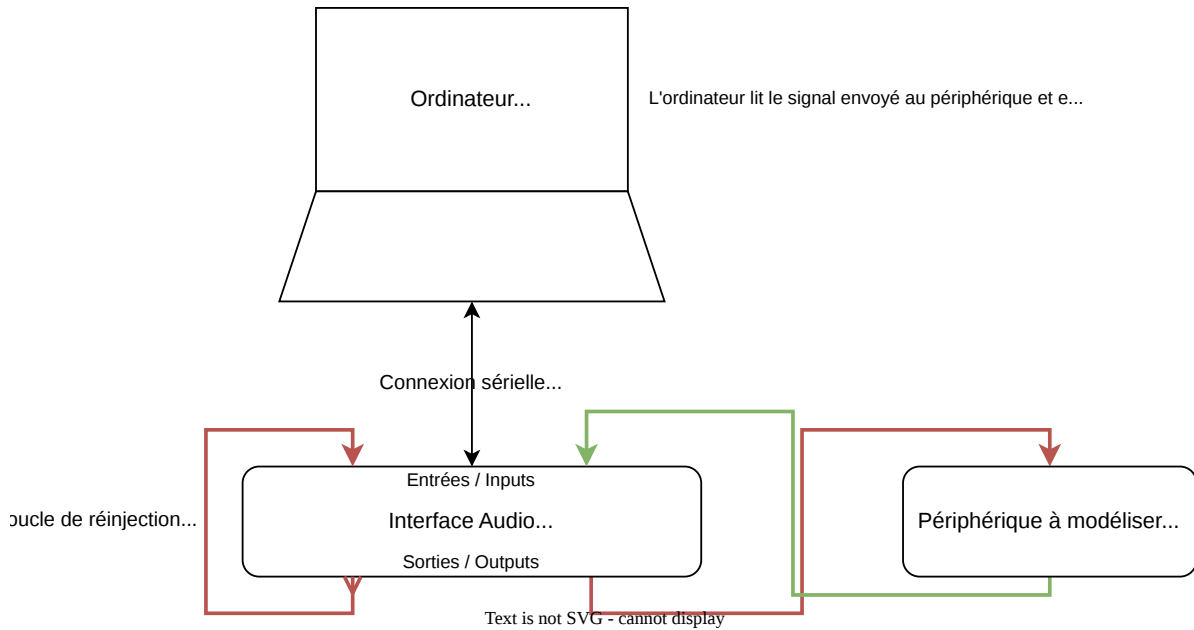


Figure 2: Exemple de câble pour illustrer ce premier cas. L'interface audio joue le rôle de préamplificateur et de convertisseur.

! Important

Le schéma ci-dessus propose l'utilisation d'une boucle de rétroaction. L'avantage de cette méthode est de ne pas tenir compte des éventuelles altérations sur la réponse en fréquence des étages de préamplification et de conversion A/N et N/A. Aussi, cela permet de s'assurer de la bonne synchronisation temporelle des enregistrements et de ne pas être victime d'un décalage dû à la latence inhérente aux systèmes numériques. Cette boucle n'est cependant pas indispensable dès lors que l'on juge ses équipements de mesure suffisamment "transparent" et que l'on est certain de pouvoir garantir la bonne synchronisation des enregistrements audio.

Dans le second cas, les choses se corsent. Un ampli de puissance à lampe possède un transformateur de sortie qui doit impérativement voir une charge particulière à ses bornes, sinon, on risque la défaillance de l'équipement. De plus, le niveau de sortie est tel qu'il est inimaginable de câbler la sortie d'un amplificateur de puissance dans une entrée en ligne. On a alors besoin d'un équipement qui va à la fois jouer le rôle de charge passive du haut-parleur (loadbox) et d'atténuateur vers un niveau en ligne.

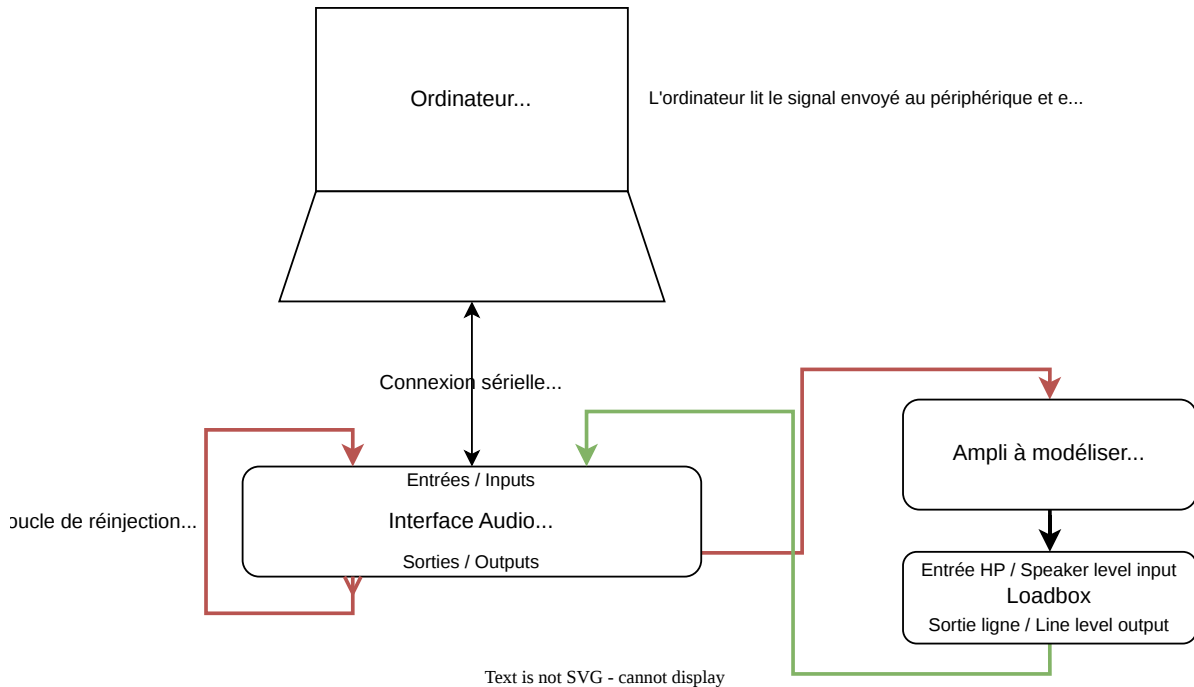


Figure 3: Même schéma, mais avec l'utilisation d'une loadbox pour la mesure d'un amplificateur.

Dans cette famille d'équipement, que nous appellerons simplement *loadbox*, on distingue deux catégories principales:

- Les *loadbox* à charges purement résistives
- Les *loadbox* à impédance ³ variable en fonction de la fréquence

Ces dernières sont les plus "réalistes" par rapport au comportement d'un haut-parleur réel, qui, lui-même, ne possède pas une impédance d'entrée constante en fonction de la fréquence.

Si les *loadbox* à charges purement résistives sont relativement abordables, les autres le sont beaucoup moins. C'est donc un investissement conséquent à rajouter en plus de son équipement à modéliser et de son équipement permettant de réaliser le traitement du signal à partir d'un modèle issu d'un apprentissage profond.

Qui plus est, cette approche tend à considérer le haut-parleur guitare comme un transducteur purement linéaire. Typiquement, une chaîne de traitement numérique "moderne" pour guitariste inclut une simulation par réseau de neurones pour la partie amplification et une

³L'impédance est une mesure de l'opposition d'un circuit électrique au passage d'un courant alternatif. Elle est une combinaison de la résistance et de la réactance.

opération de convolution pour la simulation du haut-parleur (convolution ⁴ réalisée avec la réponse impulsionnelle faite d'un haut-parleur). Rappelons ici que l'opération de convolution est une opération strictement linéaire. Cette méthode minimise ainsi le rôle du haut-parleur dans les non-linéarités introduites sur le signal.

La question est donc: comment pourrait-on faire mieux, ou, du moins, autrement ?

La méthode proposée ici est de ne pas court-circuiter le haut-parleur dans l'étape de la mesure. Ainsi, au lieu de relier l'amplificateur de puissance à une *loadbox*, ce dernier est normalement connecté à son/ses haut-parleur(s). Dès lors, on enregistre simplement ce qu'émet le haut-parleur à l'aide d'un microphone.

De prime à bord, on inclut donc les non-linéarités du haut-parleur et ses interactions avec l'amplificateur de puissance, mais aussi la réponse en fréquence du transducteur ! On perd alors la flexibilité de pouvoir changer "virtuellement" de haut-parleur en changeant de réponse impulsionnelle dans un moteur de convolution.

La solution proposée est donc de réaliser, en plus de la mesure nécessaire à l'entraînement du réseau de neurones, la réponse impulsionnelle du haut-parleur, puis de **déconvoluer** ⁵ cette réponse impulsionnelle des enregistrements réalisés en vue de l'apprentissage profond. Afin de ne pas altérer la réponse en fréquence de l'amplificateur, on réalise la réponse impulsionnelle avec un autre amplificateur, que l'on sait suffisamment linéaire.

⁴Pour les lectrices et lecteurs moins aguerris en traitement du signal, l'opération de convolution est d'une importance majeure dans la matière. Voici un bon [point de départ](#).

⁵La déconvolution est une opération mathématique utilisée pour inverser les effets de la convolution, permettant de retrouver le signal d'origine à partir d'un signal convolué.

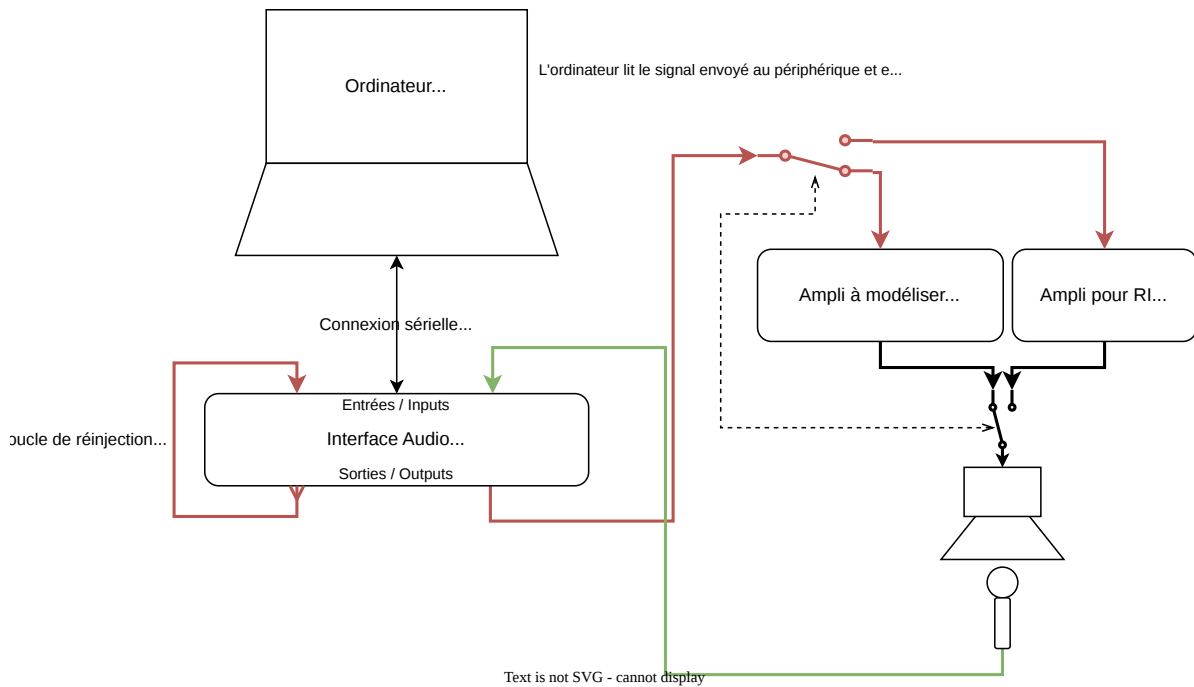


Figure 4: Mesure à partir d'un microphone. Un deuxième ampli est prévu pour la capture de la réponse impulsionnelle.

Quels sont les avantages de cette méthode ? Premièrement, il semble que l'on “extrait” ici plus d'information sur le comportement de l'amplificateur que l'on cherche à modéliser. Deuxièmement, un microphone de mesure et un amplificateur relativement “transparent” sont des équipements moins onéreux qu'une *loadbox* à charge variable.

Dans le prochain article, nous aborderons dans le détail le protocole de mesure ainsi que les enjeux entourant la déconvolution.