



Inria

Antoine Deleforge



Sylvain Faisan



Yannick Privat



Diagnostiquer l'acoustique d'une salle grâce au traitement du signal et à l'apprentissage automatique

Les Journées Techniques Acoustique et Vibrations
Autun - 2024

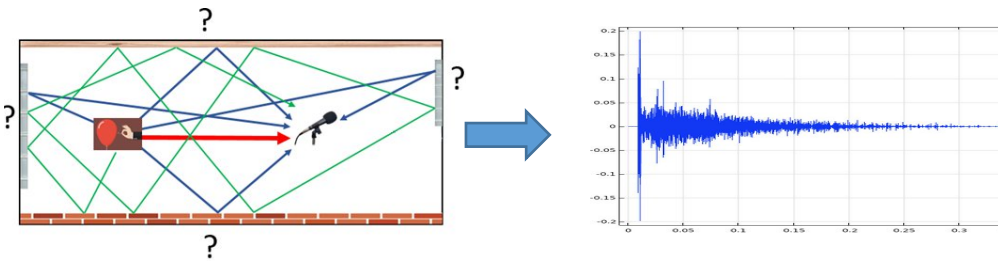


Contexte

Le diagnostic de l'existant et le choix d'une solution finale : **une approche fastidieuse !**

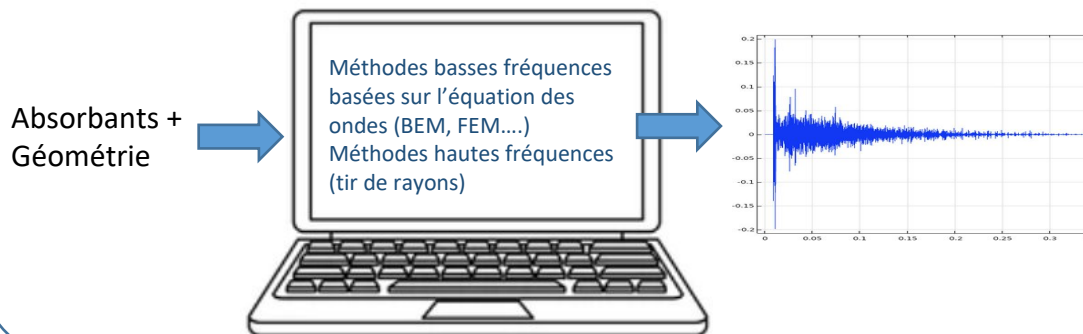
➤ Etape 1 : la mesure

la réponse impulsionnelle de la salle (RIR)

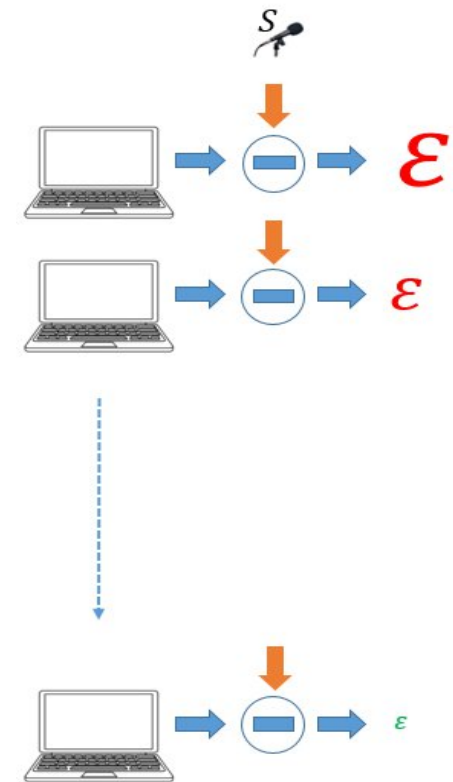


➤ Etape 2 : la simulation

➤ **les approches existantes sont uniquement directes!**

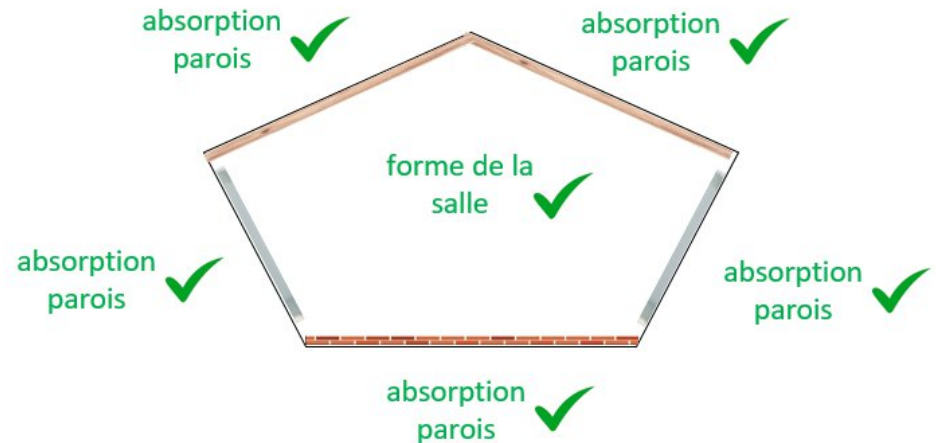
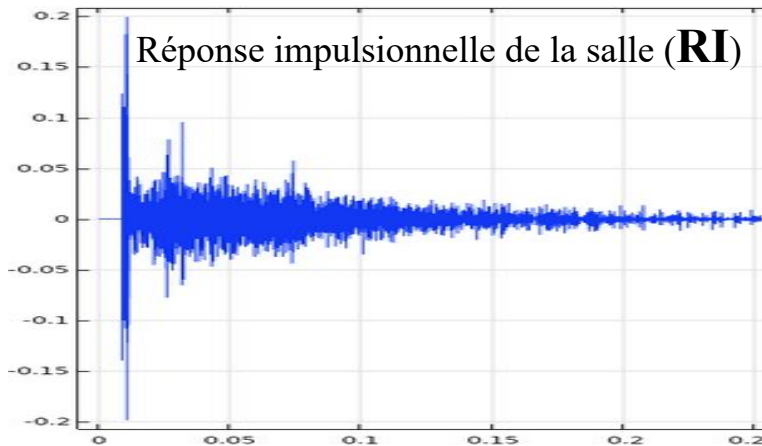


➤ Etape 3 : calage de l'absorption des matériaux par approche itérative



Optimisation du diagnostic acoustique : nouvelle approche

- ❑ Une réhabilitation acoustique optimale passe par un diagnostic fiable de l'existant.
- ❑ Développement d'une approche inverse : **un problème de physique non-linéaire**

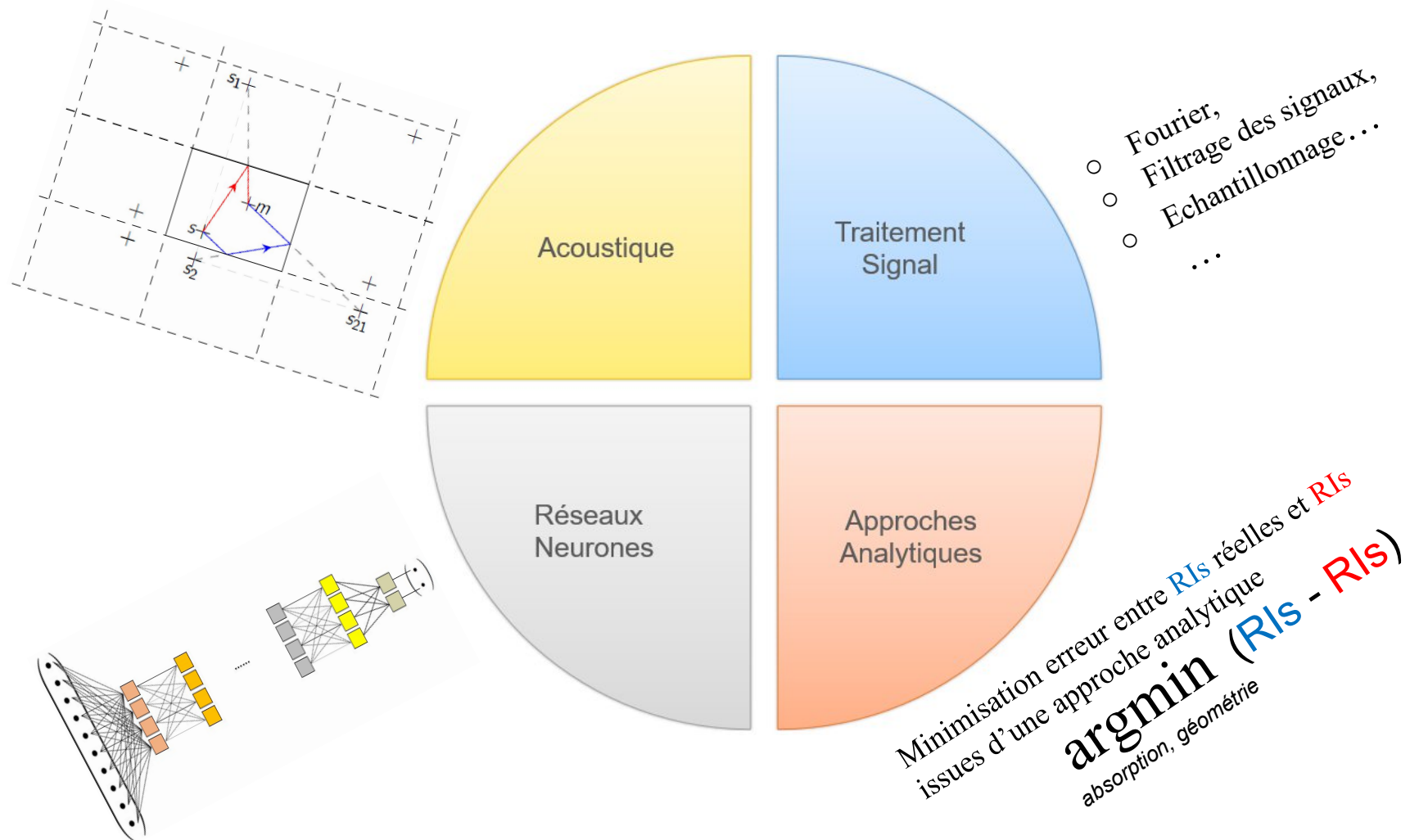


Est - il possible, en utilisant des « **mesures** » **ponctuelles** de champ sonore d'estimer précisément et automatiquement les **paramètres acoustiques et géométriques d'une salle** ?

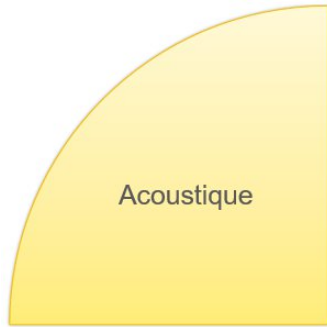
- ❑ La réponse impulsionnelle d'une salle contient toutes les informations (acoustique et géométrique) ayant impacté le champ sonore, **encore faut-il pouvoir les extraire!**

L'approche

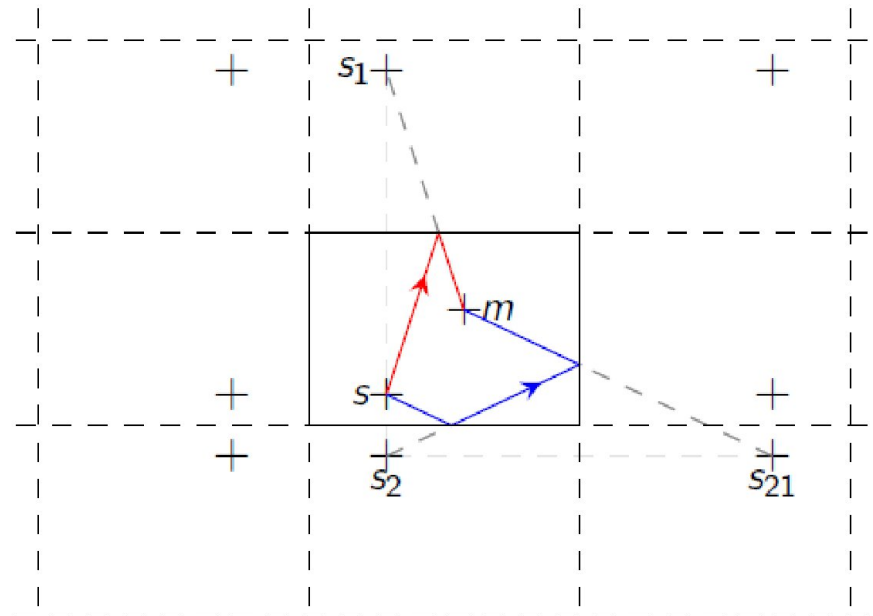
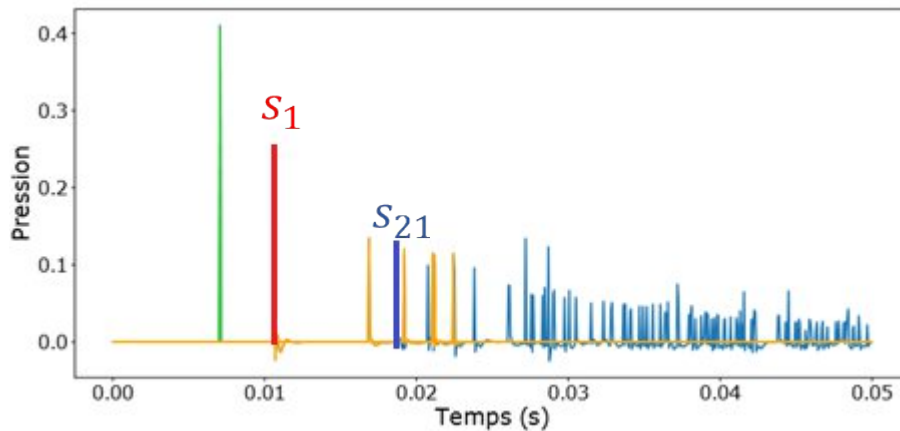
Approche nouvelle novatrice : combinaison de traitement du signal et de méthodes d'apprentissage automatique et ce, spécifiquement au cas de l'acoustique des salles



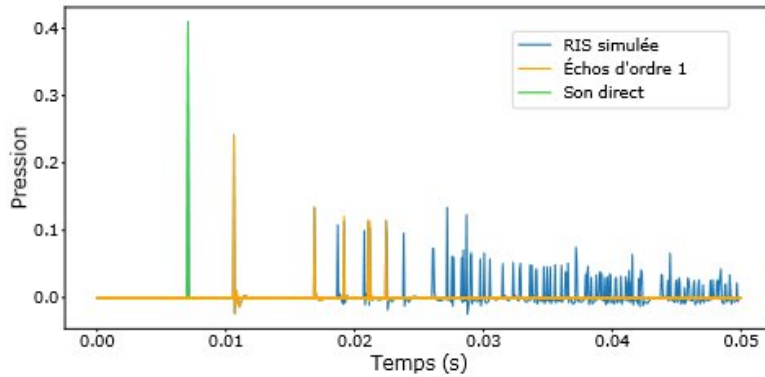
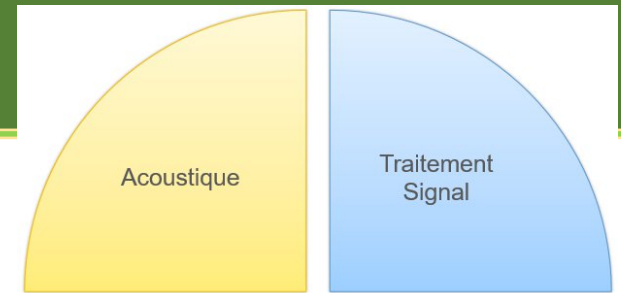
Génération de bases de données annotées



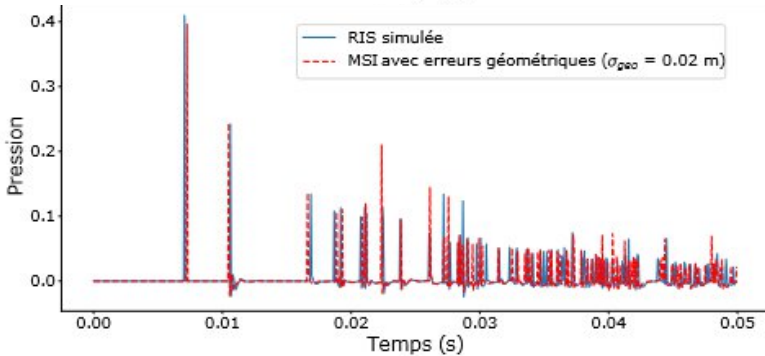
- Première difficulté : avoir des bases de données de Ris mesurées spécifiquement annotées pour nos travaux
- Création de bases de données simulées « réalistes »
- **Le modèle des sources-images** : une réponse impulsionnelle peut être vue comme la superposition des réponses d'une infinité de sources virtuelles; un écho d'une RI correspond à une SI



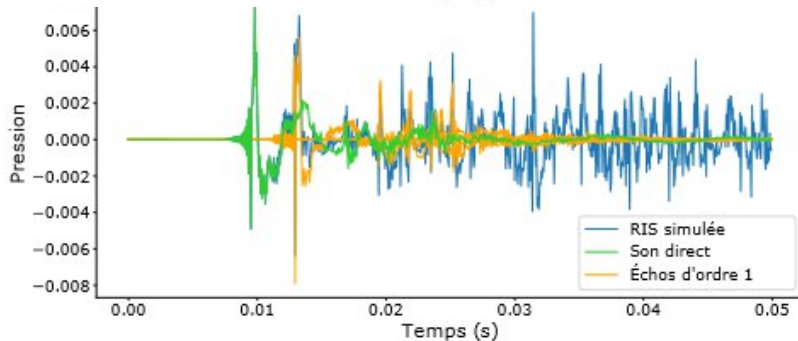
Génération de bases de données annotées



Modèle acoustique



Modèle acoustique + erreur géométrique de mesurage (loi normale)



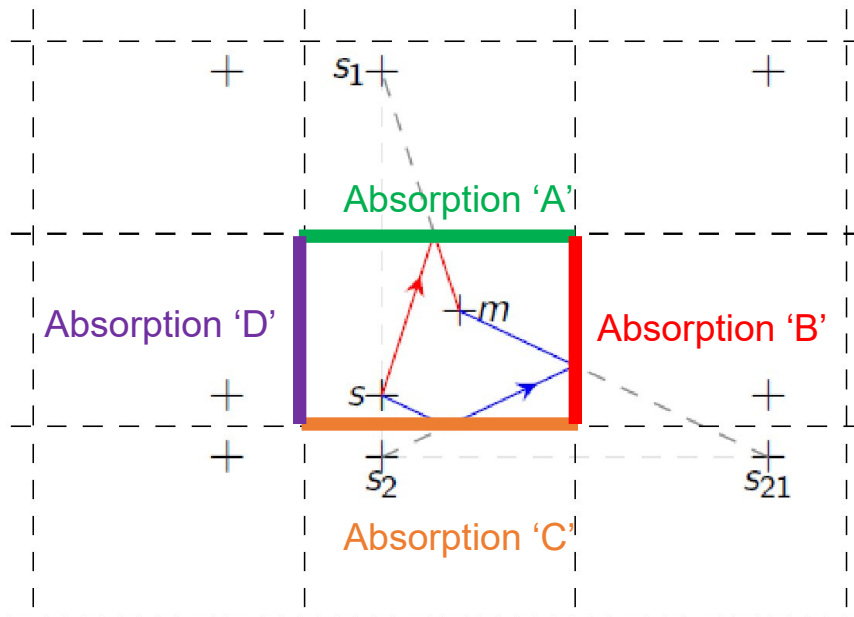
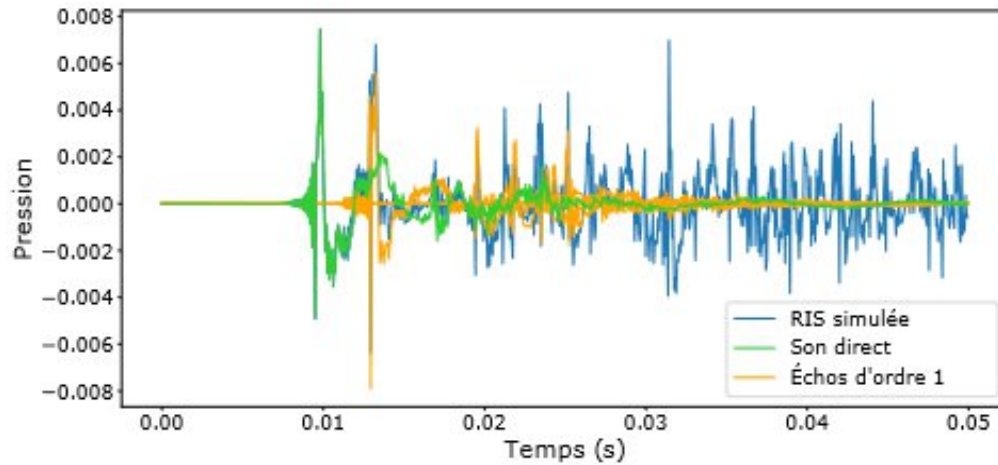
Modèle acoustique + Réponse de la source et du microphone

❑ Erreur de mesurage, Réponse des sources et microphones, Bruit résiduel, Diffusion des parois

❑ Logiciels Roosim, Pyroomacoustic

❑ RIs échantillonnées à 16 kHz et tronquées

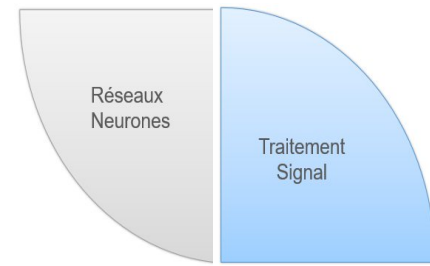
Problématique



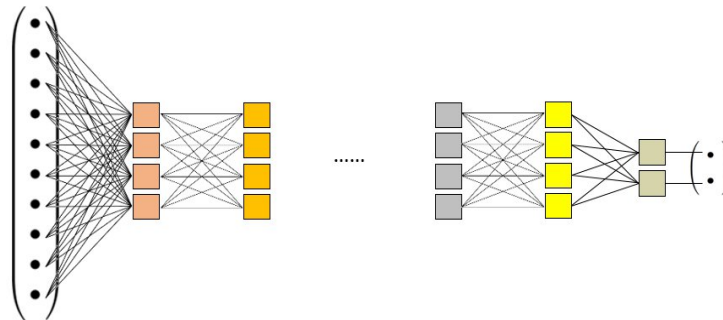
- Pic ordre 1 S_1 'A'
- Pic ordre 2 S_{21} 'B'. 'C'
- ...
- ...
- Pic ordre 1 'B'
- Pic ordre 2 'C'. 'D'
- ...
- ...
- Pic ordre 3 'A'. 'B'. 'A'
- Pic ordre 3 'D'. 'C'. 'A'
- ...
- ...



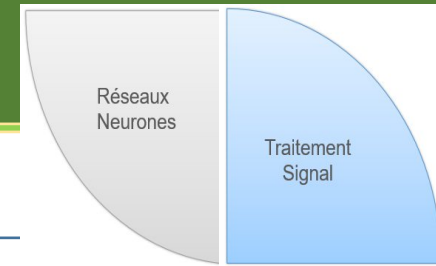
'A' 'B' 'C' 'D'



Réseaux de neurones : les avancées



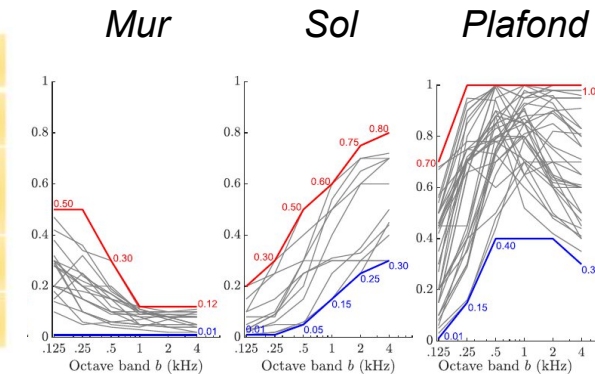
Approche temporelle *via* des réseaux de neurones



□ Absorption moyenne de la salle : études de faisabilité et empirique



Bruit de fond	PSNR (10-50 dB) Gaussien
Erreur géométrique	✗
Réponse matériel	courte, indépendante direction et fréquence
Absorption paroi	✓
Diffusion paroi	✓



- 15000 géométries par base de données
- 1 RI par salle

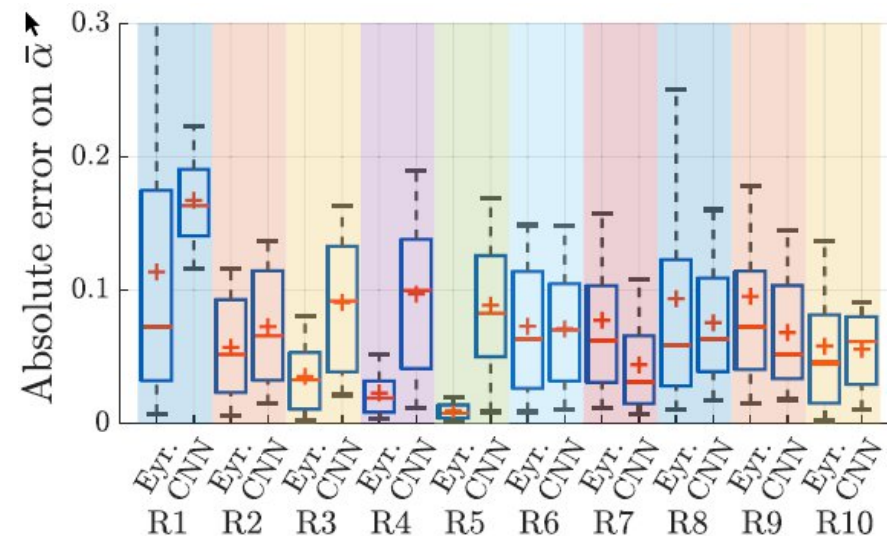
□ Deux réseaux de neurones (MLP, CNN)

□ CNN meilleur que le MLP

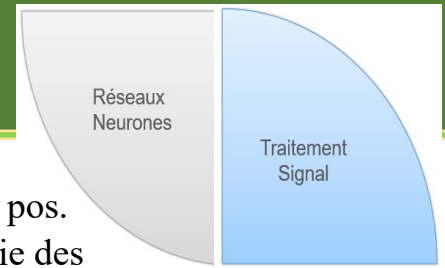
□ Résultats similaires à Eyring = $f(TR, V, S)$

□ Test sur jeux de données réelles : bon accord excepté en basses fréquences

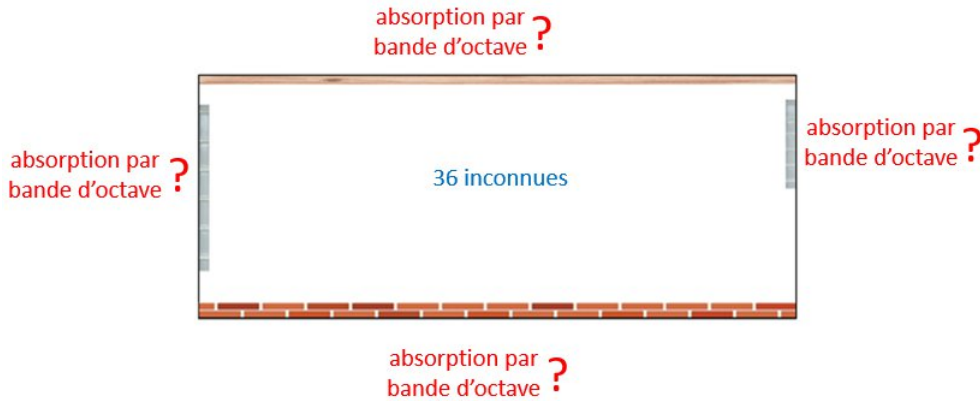
□ Erreur plus élevée si PSNR < 20 dB, Diffusion < 0.2, Absorption > 0.3



Approche temporelle *via* des réseaux de neurones



- ❑ **Absorption par bande d'octave** : étude paramétrique (géométrie (dim. salle, pos. source et microphone), bruit résiduel, diffusion parois, erreur géométrique, taille et partie des RIs)



Bruit de fond	PSNR (40-50 dB) Gaussien
Erreur géométrique	écart-type 2 cm
Réponse matériel	courte, indépendante direction et fréquence
Absorption paroi	✓
Diffusion paroi	✓

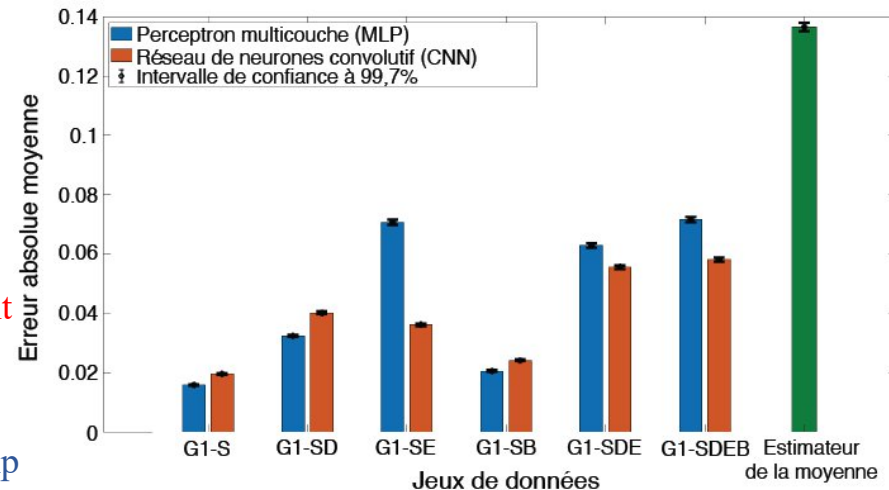
- RIs tronquées [5ms : 500ms], filtre passe-haut (50Hz)
- 15000RIs par base de données, 1 RI par salle

- ❑ Deux réseaux de neurones (MLP, CNN)

- ❑ Pas de généralisation à une géométrie quelconque (l'entraînement doit être refait pour chaque géométrie)
- ❑ L'erreur géométrique et la diffusion ont un impact important

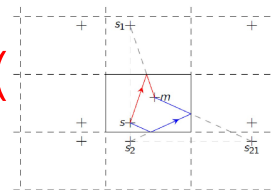
- ❑ Robustesse au bruit résiduel

- ❑ L'information utile est dans la partie précoce des RIs (champ direct, SI d'ordres petits)

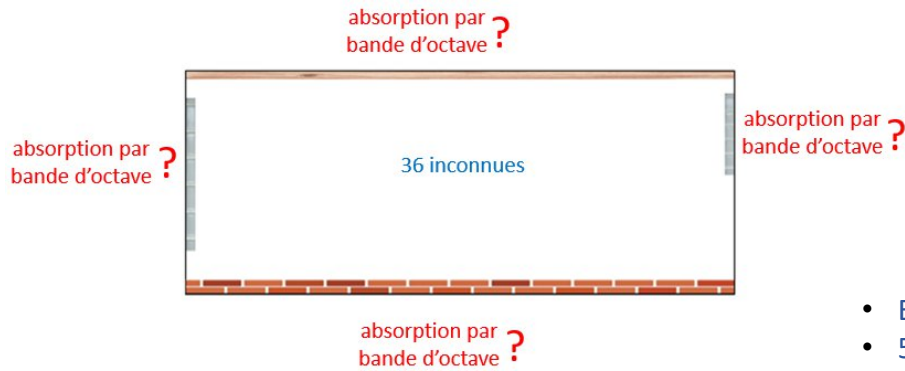


Les approches analytiques: les avancées

$\operatorname{argmin} (RIs - RIs($
absorption, géométrie



- ❑ **Absorption par bande d'octave** : identifier des fenêtres temporelles incluant un unique pic isolé



Bruit de fond	PSNR (50-60 dB) Gaussien
Erreur géométrique	écart-type 2-5 cm
Réponse matériel	courte, indépendante direction et fréquence
Absorption paroi	✓
Diffusion paroi	✗

- Echogramme : $|TFDC|^2$ de la RI
- 500 salles, 25 RIs max par salle
- Sélection des fenêtres pertinentes : probabilité de présence d'un écho + **méthode itérative robuste RANSAC (outliers)**

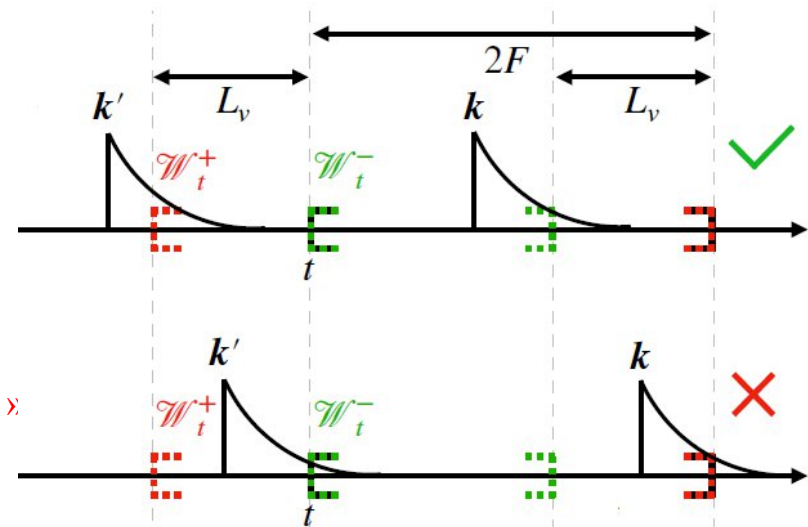
- ❑ $\tau_{exact(inconnu)} = N(\tau_{mesure(erreur)}, \text{écart-type})$

- ❑ RANSAC améliore les estimations en exploitant les SI d'ordre 2

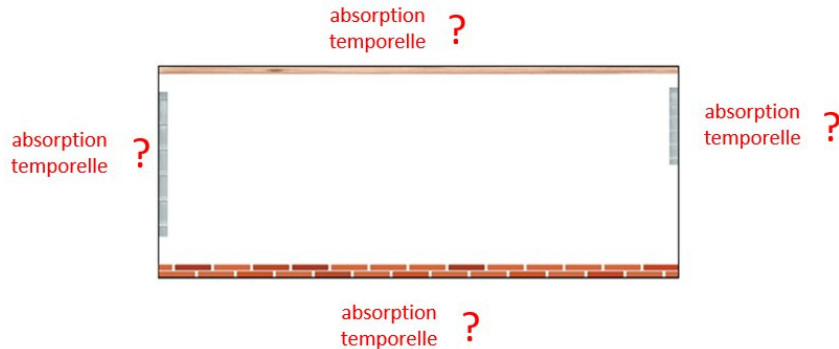
- ❑ Plus il y a de RIs dans la salle plus on réduit l'erreur et meilleure est la robustesse

- ❑ La géométrie doit être connue (temps d'arrivée « erronés » des pics)

- ❑ Pas valable en basses fréquences (< 400Hz)



□ Absorption temporelle : définir les réponses temporelles des murs



Bruit de fond	PSNR (50 dB) Gaussien
Erreur géométrique	écart-type 2 cm
Réponse matériel	✓
Absorption paroi	✓
Diffusion paroi	✗

- 100 salles, 25 RIs max par salle
- **Réponse dépendante de la direction et de la fréquence (source Genelec 8020, micro AKG C414)**

□ Extension du modèle des sources-images : introduction d'un terme corrigeant l'impact de l'erreur géométrique sur l'atténuation géométrique et sur le filtre fractionnaire

$$\left(\text{Diagram of source-image model} \right) + (\tau_{exact} - \tau_{mesure})$$

□ Problème de minimisation du délai puis, des réponses temporelles des murs (descente de gradient)

□ Soutenance de Thèse de Stéphane Dilungana (4 septembre 2024)

Approche temporelle 2 (1/4)

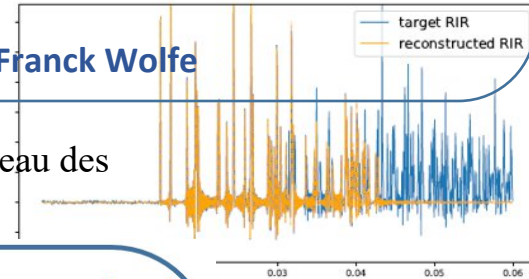
Traitement Signal

Approches Analytiques

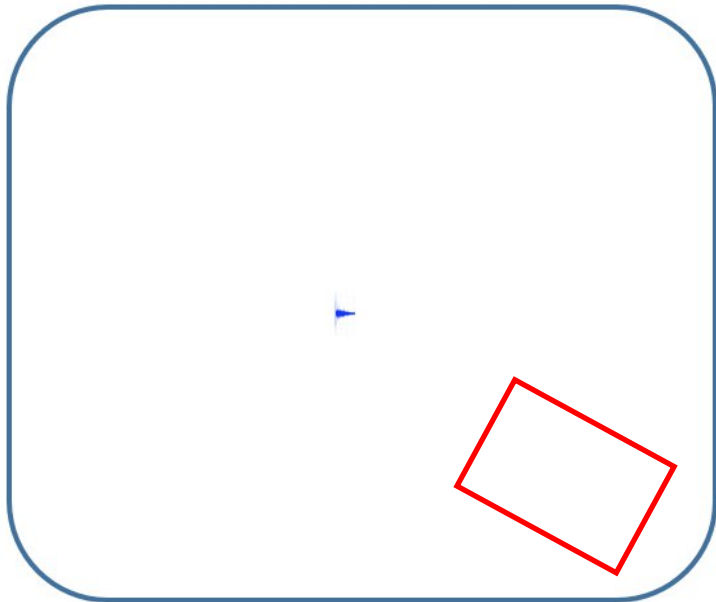
Construction de la géométrie (salle, source) ?

Bruit de fond	PSNR (50 dB) Gaussien
Erreur géométrique	✗
Réponse matériel	courte, indépendante direction et fréquence
Absorption paroi	constante ✗
Diffusion paroi	✗

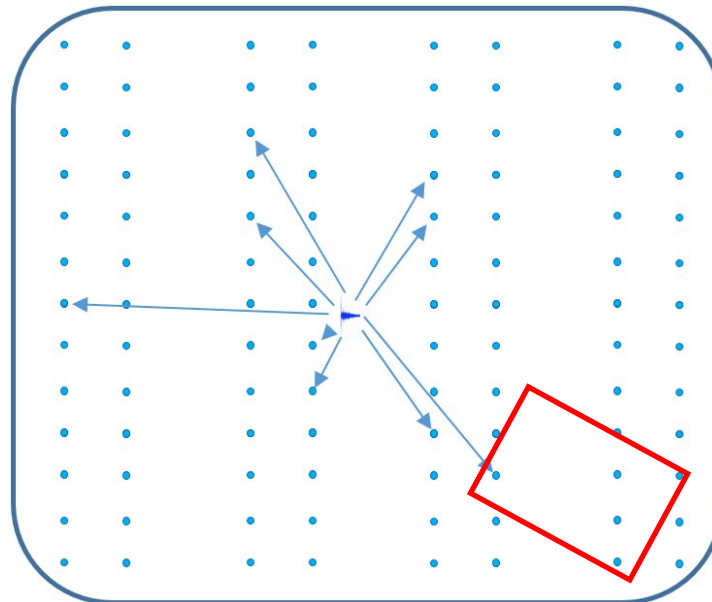
- 200 salles, 32 RI par salle
- **Algorithme itératif de Sliding Franck Wolfe**



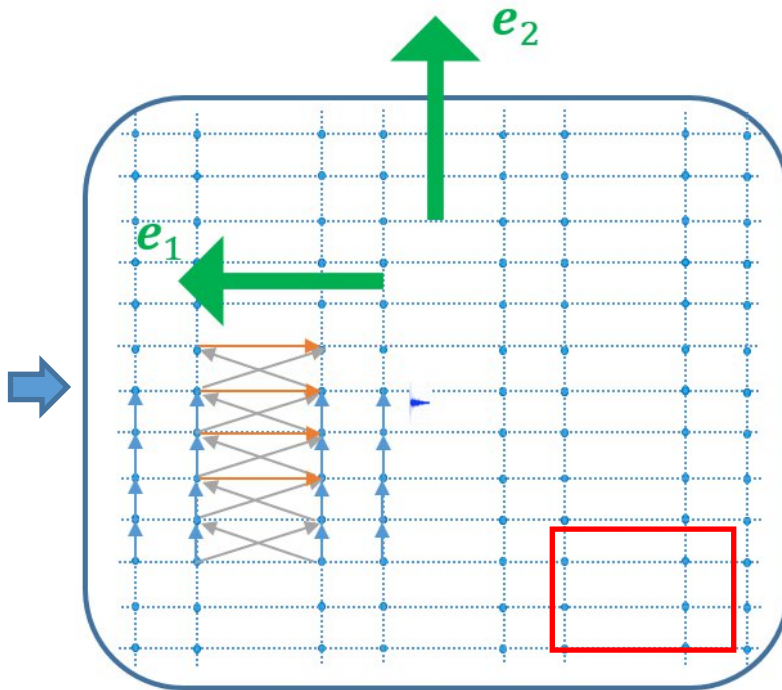
Etape 0 : Seule la réponse impulsionnelle est connue



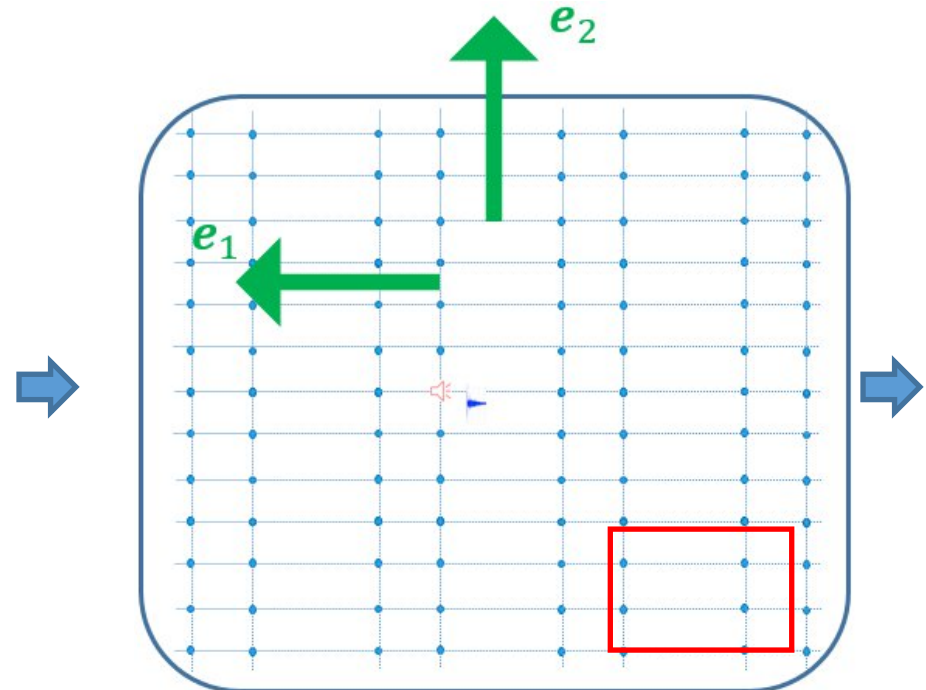
Etape 1 / 6 : construction du réseau des sources-images par SFW



Etape 2/6 : Obtention des normales aux parois de la salle (maximisation des orthogonalités) : la salle est orientée



Etape 3/6 : obtention de la source réelle



T. Sprunck, K. Chahdi, C. Foy, E. Franck, A. Deleforge, Reconstruction de la forme d'une pièce par super-résolution à l'aide de réponses impulsionnelles, CFA, France, 2022.

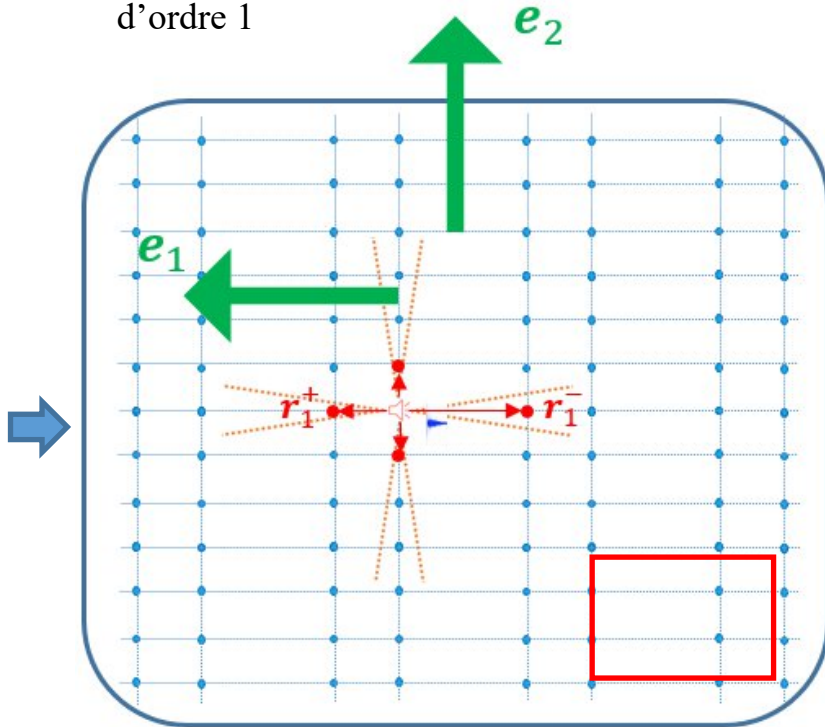
T. Sprunck, A. Deleforge, C. Foy, Y. Privat, Room Shape Reconstruction using Acoustic Super-Resolution (Poster), CANUM, France, 2022.

T. Sprunck, Y. Privat, C. Foy, A. Deleforge, Room Shape Reconstruction Using Acoustic Super-Resolution, ICA, Korea, 2022,

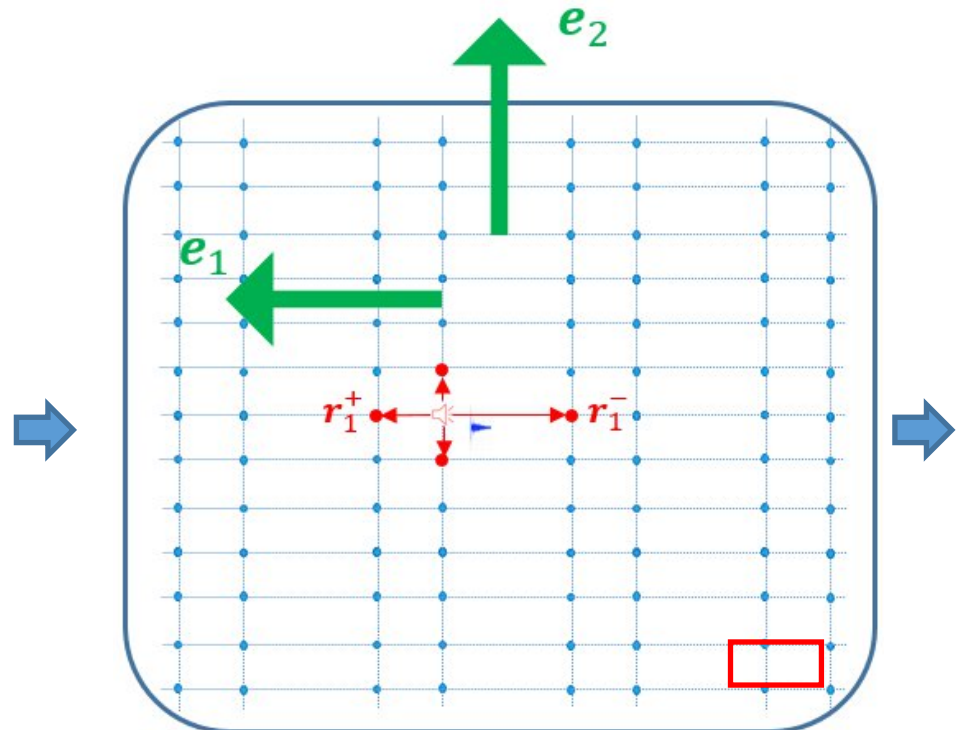
T. Sprunck, Y. Privat, C. Foy, A. Deleforge, Gridless 3D Recovery of Images Sources from Room Impulse Responses, IEEE Signal Processing Letters, 2022

T. Sprunck, A. Deleforge, Y. Privat, C. Foy, Fully Reversing the shoebox image source method: from impulse responses to room parameters, IEEE (soumis)

Etape 4/6 : obtention des sources-images d'ordre 1

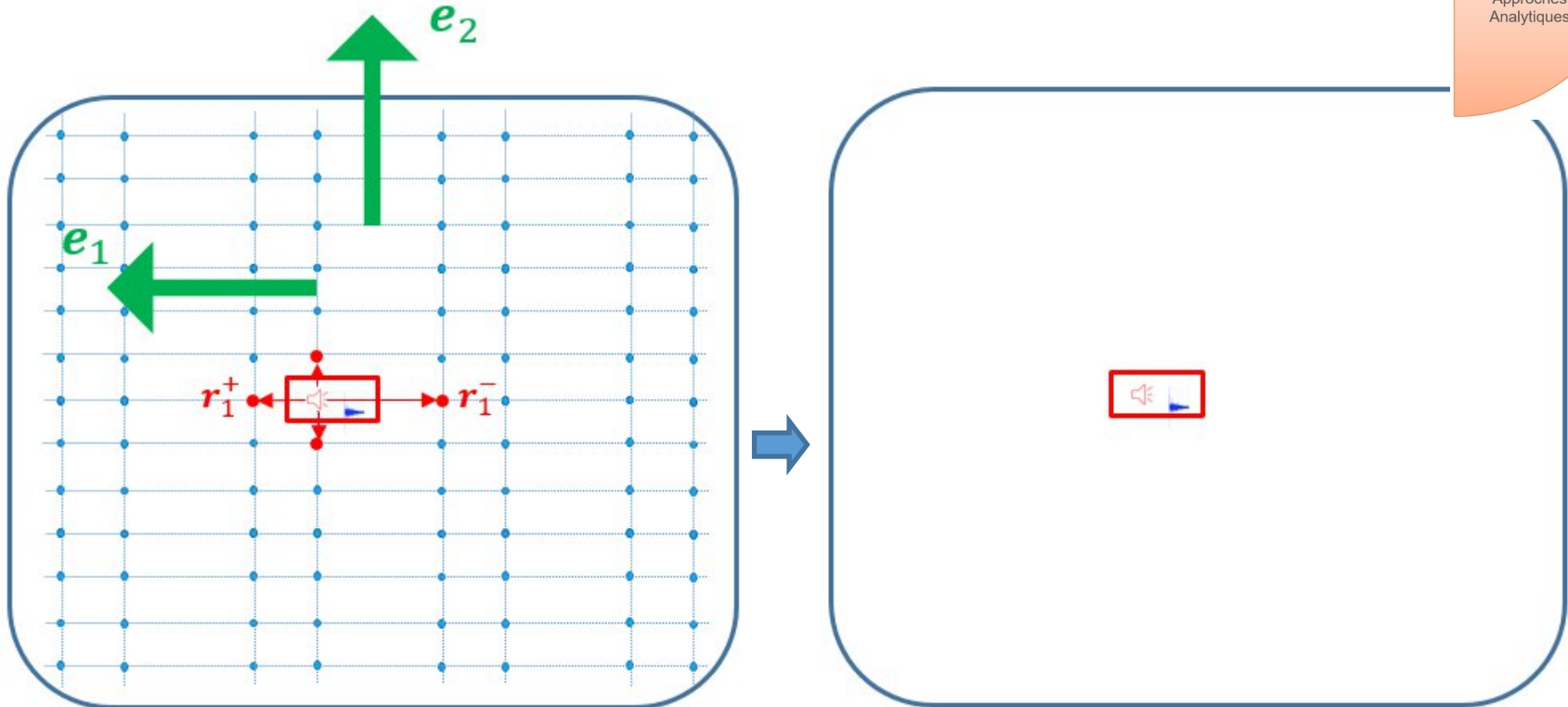


Etape 5/6 : dimensionnement de la salle



$$L_1 = \frac{1}{2} \mathbf{e}_1 \cdot (\mathbf{r}_1^+ - \mathbf{r}_1^-)$$

Etape 6/6 : positionnement de la salle



$$T = \frac{1}{2} \begin{pmatrix} e_1 \cdot (r_0 - r_1^-) \\ e_2 \cdot (r_0 - r_2^-) \end{pmatrix}$$

r_0 : distance source-microphone

Merci de votre attention