



**HAL**  
open science

## Conception Interactive en Design Sonore.

Ava Souaille

► **To cite this version:**

Ava Souaille. Conception Interactive en Design Sonore.. Son [cs.SD]. École centrale de Nantes, 2023. Français. NNT : 2023ECDN0009 . tel-04189263

**HAL Id: tel-04189263**

**<https://theses.hal.science/tel-04189263>**

Submitted on 28 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# MEMOIRE DE DOCTORAT DE

## L'ECOLE CENTRALE DE NANTES

ECOLE DOCTORALE N° 602

*Sciences de l'Ingénierie et des Systèmes*

Spécialité : Acoustique

Par

**Ava SOUAILLE**

## Conception Interactive en Design Sonore

Projet de recherche doctoral présenté et soutenu à Nantes le 22 Mars 2023

Unité de recherche : UMR 6004, Laboratoire des Sciences du Numérique de Nantes (LS2N)

### Rapporteurs avant soutenance :

Étienne PARIZET      Professeur des universités, INSA Lyon  
Bernard YANNOU      Professeur des universités, CentraleSupélec Campus Paris-Saclay, Gif-sur-Yvette

### Composition du Jury :

Présidente : Mitsuko ARAMAKI      Directrice de recherche CNRS, Aix-Marseille Université  
Examineurs : Étienne PARIZET      Professeur des universités, INSA Lyon  
Bernard YANNOU      Professeur des universités, CentraleSupélec Campus Paris-Saclay, Gif-sur-Yvette

Directeur de recherches doctorales: Jean-François PETIOT      Professeur des universités, École Centrale de Nantes  
Co-dir. de recherches doctorales : Nicolas MISDARIIS      Directeur de recherche, IRCAM, Paris  
Co-enc.de recherches doctorales : Mathieu LAGRANGE      Chargé de recherche HDR, CNRS, École Centrale de Nantes

### Invité

Sébastien DENJEAN Ingénieur de recherche, Stellantis - PSA, Vélizy-Villacoublay



# REMERCIEMENTS

---

En premier lieu, je tiens à remercier Jean-François Petiot, Nicolas Misdariis et Mathieu Lagrange pour m'avoir accordé leur confiance et permis de m'investir dans ce projet de thèse. Je les remercie également pour leur implication, leur disponibilité et leur soutien pendant ces trois années. La diversité de leurs points de vue a fortement contribué à la richesse de cette expérience.

Je souhaite également remercier Bernard Yannou et Étienne Parizet pour avoir accepté d'être les rapporteurs de cette thèse, ainsi que Mitsuko Aramaki et Sébastien Denjean en tant qu'examinatrice et membre invité, respectivement. Je remercie l'ensemble des membres du jury pour les échanges que nous avons eus pendant la soutenance.

Je remercie les membres de l'équipe Perception et Design Sonores de l'Ircam avec qui j'ai pu échanger lors de mes quelques passages à Paris ou lors de conférences. Leurs perspectives m'ont permis de développer une meilleure compréhension du design sonore, son histoire et ses enjeux.

Je remercie également les membres de l'équipe Signal, IMAge et Son du LS2N, pour leur accueil et les moments conviviaux que nous avons pu passer ensemble. Je remercie en particulier Vincent Lostanlen pour nos échanges et son aide précieuse pendant cette thèse.

Je remercie ma famille et mes proches pour leur soutien pendant ces trois ans et particulièrement ma mère. Je remercie Diatomée, pour avoir toujours cru en moi malgré mes doutes et pour continuer à me pousser à me dépasser.

Enfin, je remercie tous·tes les adelphe·s et ami·e·s que j'ai rencontré·es ces trois dernières années pour la force que leur présence m'apporte.



# TABLE DES MATIÈRES

---

<b>Introduction</b>	<b>9</b>
<b>1 Contexte général</b>	<b>11</b>
1.1 Design sonore . . . . .	11
1.1.1 Définition . . . . .	11
1.1.2 Design sonore informé . . . . .	13
1.1.3 Psychoacoustique . . . . .	13
1.1.4 Étude multidimensionnelle . . . . .	14
1.2 Optimisation interactive . . . . .	16
1.2.1 Taxonomie . . . . .	16
1.2.2 Algorithmes Génétiques Interactifs . . . . .	17
1.2.3 protocole expérimental . . . . .	24
1.3 Son et mobilité . . . . .	26
1.4 Questions de recherche . . . . .	27
1.5 Conclusion du premier chapitre . . . . .	30
<b>2 Optimisation de sons de véhicule silencieux à l'aide d'un Algorithme Génétique Interactif</b>	<b>31</b>
2.1 État de l'art . . . . .	32
2.1.1 Désagrément et détectabilité des sons de VS . . . . .	32
2.1.2 Utilisation des IGAs pour l'étude de la perception sonore . . . . .	35
2.2 Expérience 1 : Optimisation bi-objectifs de sons de véhicule silencieux . . . . .	37
2.2.1 Objectif . . . . .	37
2.2.2 Matériels et Méthodes . . . . .	37
2.2.3 Résultats . . . . .	51
2.2.4 Discussion . . . . .	54
2.3 Expérience 2 : Validation des meilleurs sons individuels . . . . .	55
2.3.1 Objectifs . . . . .	55
2.3.2 Matériels et Méthodes . . . . .	55

TABLE DES MATIÈRES

---

2.3.3	Analyse . . . . .	59
2.3.4	Résultats . . . . .	59
2.3.5	Discussion . . . . .	66
2.4	Conclusion du second chapitre . . . . .	67
<b>3</b>	<b>Recommandations de design à partir des résultats d'une épreuve IGA</b>	<b>69</b>
3.1	État de l'art . . . . .	70
3.1.1	Généralisation de résultats IGA individuels . . . . .	70
3.1.2	Intégration des objectifs du problème d'optimisation . . . . .	72
3.2	Analyse des données de l'expérience 1 . . . . .	73
3.2.1	Modèles linéaires mixtes de détectabilité et de désagrément . . . . .	73
3.2.2	Analyse statistique de <i>Optimalset</i> . . . . .	77
3.3	Méthode de recommandation . . . . .	78
3.3.1	Description de la méthode . . . . .	79
3.3.2	Mise en œuvre . . . . .	81
3.4	Expérience 3 : Évaluation de la méthode de recommandation . . . . .	83
3.4.1	Objectifs . . . . .	83
3.4.2	Matériels et Méthodes . . . . .	83
3.4.3	Résultats . . . . .	85
3.4.4	Discussion . . . . .	90
3.5	Conclusion du troisième chapitre . . . . .	92
<b>4</b>	<b>Réduction de l'espace de conception (Partie 1 : Simulations)</b>	<b>95</b>
4.1	État de l'art . . . . .	96
4.2	Réduction d'espace de conception . . . . .	98
4.2.1	Simulations . . . . .	98
4.2.2	Choix de la méthode . . . . .	99
4.2.3	Formalisation générale du problème . . . . .	101
4.2.4	Élimination de niveaux . . . . .	102
4.2.5	Approximation de la fonction-objectif . . . . .	104
4.3	Expérience 4 : simulation d'IGA . . . . .	108
4.3.1	Objectifs . . . . .	108
4.3.2	Matériels et Méthodes . . . . .	108
4.3.3	Résultats . . . . .	111
4.3.4	Discussion . . . . .	116

4.4	Conclusion du quatrième chapitre . . . . .	118
<b>5</b>	<b>Réduction de l'espace de conception (Partie 2 : Expériences perceptives)</b>	<b>119</b>
5.1	Design sonore et véhicule autonome . . . . .	120
5.2	Expérience 5 : pré-expérience pour la réduction de l'espace de conception .	122
5.2.1	Objectif . . . . .	123
5.2.2	Matériels et Méthodes . . . . .	123
5.2.3	Résultats . . . . .	133
5.2.4	Discussion . . . . .	134
5.3	Expérience 6 : validation de la méthode de réduction d'espace . . . . .	136
5.3.1	Objectif . . . . .	137
5.3.2	Matériels et Méthodes . . . . .	137
5.3.3	Résultats . . . . .	139
5.3.4	Discussion . . . . .	145
5.4	Conclusion du cinquième chapitre . . . . .	148
	<b>Conclusion générale et perspectives</b>	<b>151</b>
5.5	Conclusion . . . . .	151
5.6	Perspectives . . . . .	154
<b>A</b>	<b>Précisions sur la méthode d'élimination de niveaux</b>	<b>157</b>
A.1	Démonstration du calcul de $G_i(m)$ . . . . .	157
A.2	Démonstration de $\mathbb{E}[g(D')] < \mathbb{E}[g(D)]$ . . . . .	158
<b>B</b>	<b>Détail des sections du synthétiseur de sons de VA</b>	<b>160</b>
B.1	Détail des sections . . . . .	160
B.2	Détails des modules . . . . .	163
	<b>Références bibliographiques</b>	<b>175</b>





# INTRODUCTION

---

Le design sonore est une discipline encore mal connue du grand public, malgré la place grandissante qu'elle occupe dans notre environnement quotidien. Tout autour de nous, des sons ont été pensés et conçus dans un certain but. Par exemple, les technologies numériques telles que les applications de smartphone utilisent des sons pour communiquer des informations aux utilisateur·rice·s. Ces sons accompagnent les notifications, la réception de messages ou encore la confirmation d'une action de la part de ces dernier·ère·s. Mais bien d'autres sons, moins évidents, ont également été pensés, notamment au sein de l'industrie automobile, comme par exemple le son produit par la fermeture d'une porte de voiture. Les sons font partie intégrante de l'expérience utilisateur et de l'identité de ces produits, ce qui est de plus en plus pris en compte par les industriels.

Un enjeu crucial consiste donc à pouvoir créer ou choisir un son qui produit l'effet désiré sur l'utilisateur·rice. Il peut s'agir de vouloir provoquer une sensation donnée, comme un état de relaxation, ou encore de contribuer à une image particulière, comme une impression de qualité. Cette démarche peut être guidée par l'expertise d'un·e designer·euse sonore, mais bénéficie fortement de la construction de connaissances concernant les liens entre les propriétés objectives (acoustiques) des sons et leurs propriétés subjectives (psychologiques), c'est-à-dire celles qui sont perçues par les utilisateur·rice·s. Plusieurs outils existent pour étudier ces liens et intégrer des connaissances sur la perception des utilisateur·rice·s au processus de conception, en design sonore mais aussi dans le design de produit en général.

Ces vingt dernière années, des outils interactifs ont vu le jour, notamment basés sur des Algorithmes Évolutionnaires Interactifs (IECs, *Interactive Evolutionary Computation*) et en particulier sur les Algorithmes Génétiques Interactifs (IGAs). Ces outils ont fait leurs preuves pour de nombreuses applications de design de produit, comme la forme et la couleur de robes, la forme verres à pied, la silhouette de véhicules ou encore la forme et l'affichage de tableaux de bord de voiture. Ils sont également utilisés comme aide à la créativité, pour guider un·e musicien·ne ou un·e designer·euse sonore dans l'exploration des sons produits par un synthétiseur.

La problématique de cette thèse porte sur la définition d'outils et de méthodes interac-

tives pour le design sonore, et en particulier celle d’outils intégrant l’expérience utilisateur dans le processus de conception. Un premier chapitre introduit plus en détails le contexte général dans lequel s’inscrivent les travaux présentés dans ce manuscrit. Un paradigme expérimental basé sur des expériences d’optimisation interactive y est proposé. Ce paradigme est composé de trois parties : la définition de l’espace de conception à explorer, les expériences d’optimisation interactive et l’analyse des données pour alimenter un processus de design sonore. Les quatre chapitres suivants traitent chacun d’une question de recherche portant sur ce paradigme, via deux exemples applicatifs : le design de sons d’alerte de Véhicule Silencieux (VS, chapitres 2 à 4) et le design de sons d’intérieur de Véhicule Autonome (VA, chapitre 5). Ces travaux de thèse ont donné lieu à six expériences, présentées dans ce manuscrit.

Les contributions principales de cette thèse sont :

- L’utilisation d’un IGA pour résoudre un problème bi-objectifs d’optimisation interactive de sons,
- La proposition et la validation d’une méthode d’analyse des résultats obtenus individuellement à des expériences IGA, pour produire des recommandations générales en termes de design sonore,
- La proposition et la validation d’une méthode de réduction informée d’espace de conception, préalable à une expérience d’optimisation interactive.

Les travaux effectués ont fait l’objet d’une communication dans une conférence internationale (SOUAILLE, PETIOT, LAGRANGE et al. 2021), une communication dans une conférence nationale (SOUAILLE, MISDARIIS et al. 2022), ainsi que la publication d’un article dans une revue internationale (SOUAILLE, PETIOT, MISDARIIS et al. 2022). Les chapitres 2 et 3 sont en partie adaptés de SOUAILLE, PETIOT, MISDARIIS et al. (2022) et SOUAILLE, PETIOT, LAGRANGE et al. (2021).

# CONTEXTE GÉNÉRAL

---

Ce premier chapitre présente le contexte général dans lequel s’inscrivent ces travaux de thèse. Dans un premier temps, nous présentons une définition du design sonore et des méthodes expérimentales généralement utilisées pour récolter des informations sur la perception sonore. Dans un second temps, nous détaillons un nouveau protocole expérimental basé sur la notion d’optimisation interactive, qui présente un intérêt potentiel pour le design sonore. Ensuite nous introduisons les deux applications de design sonore traitées au cours de cette thèse, à savoir le design de son d’alerte de véhicule silencieux et le design de son d’intérieur de véhicule autonome. Enfin, nous détaillons les quatre questions de recherche traitées dans ce manuscrit et leurs liens avec le protocole expérimental d’optimisation interactive et les deux applications abordées.

## 1.1 Design sonore

### 1.1.1 Définition

D’après MISDARIIS (2018), le design sonore peut être décrit comme :

[...] champ du design relativement récent qui consiste principalement à prendre en compte la dimension sonore d’un objet au cours de son processus de conception – le terme objet pouvant être pris au sens large d’objet tangible (produit manufacturé), d’objet numérique (interface homme-machine) ou d’objet spatial (environnement). Le design sonore relève donc de la création sonore appliquée à des domaines où penser le son, l’imaginer, le fabriquer et l’intégrer s’avère nécessaire d’un point de vue fonctionnel et/ou esthétique, comme notamment : l’industrie, l’architecture, l’urbanisme, le cinéma, les arts numériques, voire la communication et le marketing lorsqu’il s’agit d’associer un son à une identité de marque ou de produit.

Ce terme recouvre donc une grande diversité de pratiques et d’applications et fait appel à des savoir-faire dans de multiples disciplines, telles que l’ingénierie audio, la composition

ou la pratique musicale (ZATTRA et al. 2021), comme l’avait d’ailleurs pressenti SCHAFER (2010) :

Le véritable designer sonore doit tout comprendre de l’environnement auquel il s’attaque ; il lui faut des connaissances en acoustique, en psychologie, en sociologie, en musique, et en bien d’autres matières encore suivant les cas.

Le design sonore peut s’appliquer à deux grandes catégories de sons : les sons dits non-intentionnels (ou consécutifs) et ceux dits intentionnels (LANGEVELD et al. 2013). Les premiers sont les sons inévitablement produits par les objets manufacturés et ne sont pas nécessairement souhaités. Le design sonore a donc d’abord consisté, au cours du XX<sup>e</sup> siècle, à tenter d’améliorer a posteriori la qualité sonore de ces objets, en agissant le plus souvent sur la forme ou la structure de ces derniers. Cette approche a par exemple été appliquée au son émis par la fermeture d’une porte de voiture (PARIZET, GUYADER et NOSULENKO 2008 ; BEZAT 2007) ou encore celui d’un briquet de luxe (LAGEAT, CZELLAR et LAURENT 2003). Les sons intentionnels sont quant à eux ajoutés dans le but de communiquer une information de forme (c’est-à-dire liée à la qualité générale de l’objet) ou de fonction (liée à l’utilisation de l’objet). SUSINI, HOUIX et MISDARIIS (2014) distinguent trois types de fonctions :

- La **notification sonore**, qui sert à informer l’utilisateur·rice d’un événement extérieur. Il s’agit, par exemple, de la fonction des sons d’alarme,
- Le **retour sonore**, qui est joué suite à une action de l’utilisateur·rice. Il permet d’informer sur le résultat de l’action effectuée. Ces sons sont très utilisés dans les interfaces homme-machine (IHM). Par exemple, les lecteurs automatiques de billets de transport émettent un son au moment de la lecture du billet. Le type de son émis informe l’usager·ère de la validité (ou non) du billet,
- L’**interaction sonore**, pour laquelle le son réagit en temps réel aux actions de l’utilisateur·rice. Comme exemple d’application, nous retrouvons la sonification de la dynamique du véhicule dans un véhicule électrique (DENJEAN 2015).

En fonction du type de son (intentionnel ou non) et de sa fonction (notification, retour ou interaction), le·a designer·euse sonore dispose de différents outils :

- Modification de la structure mécanique de l’objet, choix d’un autre matériau ou modification du type d’interaction excitatrice,
- Lecture de fichiers sonores, créés en amont via des techniques de prise de son, de traitement du signal et/ou de synthèse sonore,
- Synthèse sonore en temps réel, en fonction de paramètres liés à l’état de l’objet,

de son environnement et des actions de l'utilisateur·rice.

### 1.1.2 Design sonore informé

Bien que, dans sa pratique, le·a designer·euse sonore ait souvent une approche intuitive basée sur son expérience personnelle (MISDARIIS 2018), il est intéressant de souligner l'importance d'informer le processus de design par des connaissances scientifiques sur la perception sonore (SUSINI, HOUIX et MISDARIIS 2014). En effet, lors de la conception d'un son, le·a designer·euse sonore cherche à lui donner une certaine forme et une certaine fonction. Autrement dit, le son est créé avec une certaine intention. Cette dernière est caractérisée par un ensemble d'attributs perceptifs, c'est-à-dire de qualités sonores subjectives perçues par les usager·ère·s. Ces attributs peuvent par exemple être de bas niveau, comme l'intensité sonore perçue, plutôt de haut niveau et contextuelle, comme la gêne perçue, ou encore liée au timbre, comme la brillance. La perception d'un attribut donné peut ainsi faire appel à des processus cognitifs plus ou moins complexes et être dépendante du contexte. L'étude de la perception sonore vise, entre autres, à mettre à jour les liens entre des attributs perceptifs et les propriétés acoustiques du son, afin de connaître les caractéristiques objectives qu'un son doit avoir pour répondre au cahier des charges d'un problème. Ces informations délimitent un cadre, en termes de timbre et de morphologie temporelle (variation de hauteur, durée et intensité), que le·a designer·euse peut explorer pour faire des propositions de design.

### 1.1.3 Psychoacoustique

L'étude d'attributs perceptifs de sons est historiquement ancrée dans le domaine de la psychoacoustique (SUSINI, LEMAITRE et MCADAMS 2012). Celle-ci vise à expliquer la perception d'un phénomène sonore (un percept) par des modèles objectifs, généralement issus du fonctionnement du système auditif. Parmi les quantités étudiées, nous retrouvons par exemple :

- La **sonie**, qui correspond à l'intensité sonore perçue.
- La **force de fluctuation**, qui caractérise les modulations d'un son en amplitude ou en fréquence, pour des fréquences de modulation inférieures à 20Hz.
- La **rugosité**, similaire à la force de fluctuation, mais pour des fréquences de modulation supérieures à 20Hz,

- L'**acuité**, sensation liée à la répartition de l'énergie sur la bande passante fréquentielle d'un son.

Ces descripteurs psychoacoustiques peuvent être directement calculés à partir du signal sonore étudié, par des méthodes notamment décrites par ZWICKER et FASTL (2013). Cependant, pour beaucoup d'attributs perceptifs qui peuvent avoir un intérêt dans le cadre du design sonore, un tel modèle prédictif n'est pas nécessairement disponible. Par exemple, la perception de certains attributs, comme la gêne occasionnée par un son, dépend du contexte et/ou de la source à l'origine du son. C'est également le cas pour la préférence des futur·e·s utilisateur·rice·s, qui présente un intérêt dans le cadre du design sonore et dans celui du design de produits en général, mais qui dépend de l'application considérée. L'approche psychoacoustique nécessite de définir en amont les attributs à étudier et de s'assurer qu'ils soient compréhensibles par les sujets. Cette approche est donc peu adaptée à une étude exploratoire, pourtant nécessaire pour des projets de design sonore portant sur des applications peu ou pas connues.

#### 1.1.4 Étude multidimensionnelle

À l'inverse, plusieurs approches multidimensionnelles ont été proposées pour étudier la perception sonore, avec moins de contraintes sur le choix et le nombre d'attributs perceptifs à étudier. Une première étape de ces méthodes consiste généralement à former un corpus de sons représentatifs du domaine d'intérêt. Ces sons sont alors analysés pour tenter de faire émerger les propriétés qui permettent de les caractériser et de les différencier entre eux. Parmi les méthodes utilisées, nous retrouvons (SUSINI, LEMAITRE et MCADAMS 2012; SUSINI, HOUIX et MISDARIIS 2014) :

- L'évaluation des sons selon plusieurs échelles sémantiques, comme des échelles sémantiques différentielles. Cette approche permet d'identifier les caractéristiques propres à chaque son du corpus, mais nécessite de choisir en amont des attributs perceptifs à étudier, ainsi que de s'assurer de la bonne compréhension des échelles par les sujets,
- La verbalisation libre, qui permet de faire émerger le vocabulaire utilisé pour décrire un groupe de sons. Les mots qui ressortent de cette analyse peuvent par exemple être utilisés pour construire des échelles sémantiques sur lesquelles les sons seront ensuite évalués,
- L'évaluation de la dissimilarité entre toutes les paires de sons du corpus. Cette méthode permet ensuite, via une technique de positionnement multidimensionnel

(*Multidimensional Scaling, MDS*), de placer l'ensemble des sons étudiés dans un espace de faible dimension (généralement 2 ou 3), dans lequel la distance entre les sons représente leur similarité (deux sons similaires sont proches et inversement). Ainsi, il n'est pas nécessaire de définir à l'avance combien d'attributs perceptifs serviront à caractériser le corpus, ni lesquels. Les dimensions latentes produites par la méthode de positionnement multidimensionnel sont ensuite généralement corrélées à des descripteurs psychoacoustiques. Cependant, cette approche requiert un grand nombre d'évaluation ( $N(N - 1)/2$  si  $N$  est le nombre de sons dans le corpus) et ne s'applique qu'à l'étude d'un corpus homogène,

- La classification, qui vise à grouper les sons par catégorie ou similarité selon un critère donné, par exemple en positionnant ces derniers dans un plan via une interface graphique. Cette approche permet d'étudier un corpus non homogène et est souvent utilisée pour réduire la taille d'un corpus en sélectionnant un sous-ensemble de sons représentatifs des catégories identifiées.

Plusieurs de ces outils sont souvent utilisés conjointement, permettant ainsi de caractériser le corpus étudié selon des critères à la fois objectifs, tels que des descripteurs psychoacoustiques, et subjectifs, c'est-à-dire propres à la perception des sujets et au contexte d'application. Les informations obtenues peuvent ensuite servir à prédire la qualité de nouveaux sons par rapport à celle des sons du corpus, au cours d'un processus de conception. Comme application, nous trouvons par exemple l'étude de sons de fermeture de portes de voitures (PARIZET, GUYADER et NOSULENKO 2008 ; BEZAT 2007) mentionnée précédemment, ou celle de sons de climatiseurs (SUSINI, MCADAMS et al. 2004). L'approche générale et les méthodes décrites ci-dessus sont plus largement utilisées dans le design de produits, afin d'intégrer la perception, les émotions et/ou les préférences des utilisateur·rice·s dans le processus de conception. Par exemple, PETIOT et YANNOU (2004) étudient les liens entre la forme de verres et des dimensions sémantiques. Ces méthodes permettent de caractériser l'ensemble des sons du corpus selon les quantités d'intérêt. Cependant, les liens mis à jour entre ces dernières sont généralement limités à des relations linéaires. De plus, le nombre de sons qu'il est possible d'étudier simultanément est limité, car chaque son du corpus doit être évalué.



## 1.2 Optimisation interactive

Une autre approche possible pour intégrer la perception humaine dans le processus de conception consiste à se baser sur une méthode d’optimisation interactive. Le problème de design est alors posé comme un problème d’optimisation, pour lequel un objectif doit être minimisé (ou maximisé) sous contrainte. Autrement dit, le problème peut être posé comme :

$$\inf_{x \in A} f(x), \quad (1.1)$$

où  $f$  est la fonction-objectif à minimiser et  $A$  est l’ensemble des solutions admissibles du problème. À la différence des méthodes précédentes, les designs étudiés (correspondant aux solutions  $x$ ) ne proviennent pas d’un corpus mais sont générés par un modèle paramétrique, caractérisé par plusieurs variables de conception. Nous appelons espace de conception l’espace généré par l’ensemble des variables de conception. Ces méthodes sont dites interactives, car elle font intervenir l’expertise humaine dans le processus d’optimisation. D’une manière générale, la personne qui intervient dans le processus d’optimisation peut être n’importe quelle partie prenante du projet. Ainsi, dans cette section, le terme d’utilisateur·rice se réfère non pas à l’utilisateur·rice final·e du produit, mais à l’utilisateur·rice de l’outil d’optimisation interactive.

### 1.2.1 Taxonomie

Dans leur article de revue, MEIGNAN et al. (2015) proposent une taxonomie de ces méthodes, dont le premier critère concerne le **but** de l’interaction :

- *Interaction orientée problème* : l’interaction a pour but de modifier le problème d’optimisation,
- *Interaction orientée recherche* : l’interaction a pour but d’améliorer les performances de la procédure d’optimisation.

Le deuxième critère concerne le **rôle** de l’utilisateur·rice. Lorsque l’interaction est *orientée problème*, il·elle :

- *Ajuster* : L’utilisateur·rice ajuste les contraintes ou les objectifs,
- *Enrichir* : L’utilisateur·rice définit de nouvelles contraintes ou de nouveaux objectifs.

Lorsque l’interaction est *orientée recherche*, l’utilisateur·rice peut :

- *Assister* : La procédure de recherche est effectuée par l’utilisateur·rice,

- *Guider* : L'utilisateur-riche donne des informations sur les variables de décision,
- *Régler* : L'utilisateur-riche agit sur les paramètres de la procédure d'optimisation.

Cette taxonomie montre que différents modes d'interaction sont possibles pour ce type d'approche, qui est surtout utilisé comme aide à la décision. Les catégories présentées ne sont pas exclusives, une méthode donnée pouvant se baser sur plusieurs modes d'interaction avec l'utilisateur-riche. MEIGNAN et al. (2015) proposent ensuite trois autres critères de classification, qui décrivent la méthode d'optimisation via une approche système :

- **Type d'intégration** des retours de l'utilisateur-riche :
  - *Sans modèle* : l'information est utilisée telle quelle dans la méthode d'optimisation,
  - *Avec modèle* : l'information sert à ajuster un modèle, sur lequel s'appuie la méthode d'optimisation.
- **Durée de vie de l'information** fournie par l'utilisateur-riche :
  - *Par étape* : l'information n'est utilisée qu'à l'étape à laquelle elle est fournie,
  - *Court terme* : l'information est utilisée pendant toute la durée de résolution du problème,
  - *Long terme* : l'information est utilisée sur plusieurs instances du problème d'optimisation.
- **Type de procédure** d'optimisation :
  - *Exacte* : la découverte d'un optimum est garantie,
  - *Heuristique* : procédure d'optimisation approchée, spécifique au problème traité,
  - *Métaheuristique* : méthode heuristique indépendante du problème basée sur l'exploration de l'espace de conception.

La taxonomie proposée par MEIGNAN et al. (2015) est illustrée dans la figure 1.1.

### 1.2.2 Algorithmes Génétiques Interactifs

Les algorithmes évolutionnaires sont des procédures d'optimisation métaheuristiques inspirées des mécanismes de la sélection naturelle. Ils opèrent généralement sur un groupe de solutions potentielles, qui est progressivement modifié par essai-erreur via des mécanismes d'exploration (diversification des solutions) et d'exploitation (amélioration de la qualité des solutions) (ČREPINŠEK, LIU et MERNIK 2013). Ces mécanismes sont le plus souvent basés sur des processus stochastiques. La fonction-objectif est parfois appelée fonction d'évaluation (*fitness* en anglais) et doit le cas échéant être maximisée. Les algorithmes évolutionnaires interactifs (*Interactive Evolutionary Computation*, IEC) sont des

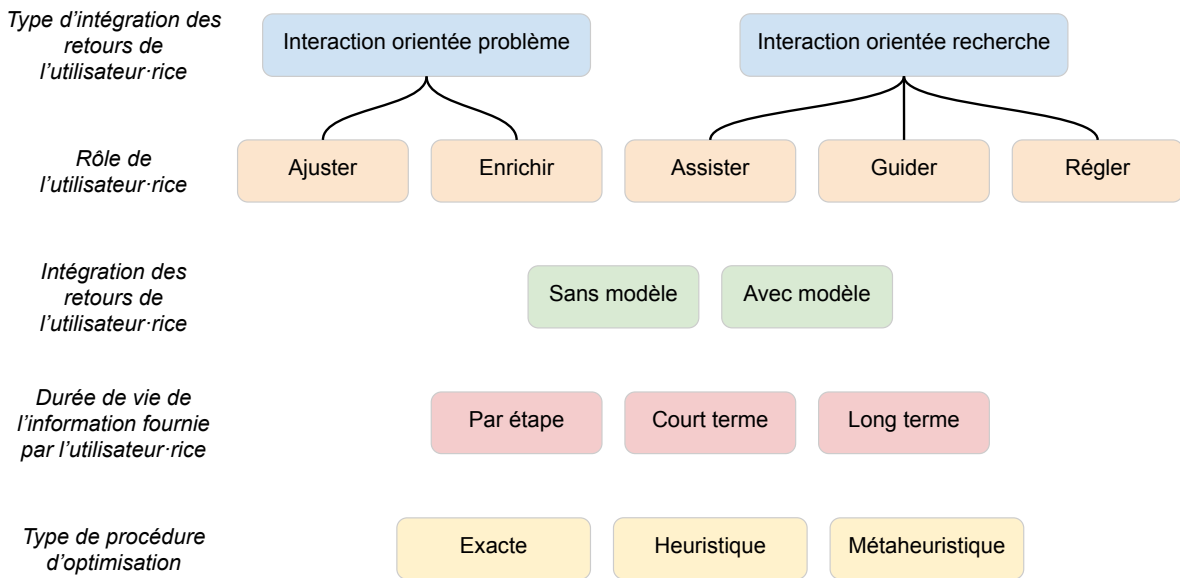


FIGURE 1.1 – Taxonomie des méthodes d’optimisation interactive proposée par MEIGNAN et al. (2015) (traduction libre).

algorithmes évolutionnaires qui font intervenir l’utilisateur·rice dans le processus d’optimisation.

Au cours de cette thèse, nous nous intéressons en particulier aux Algorithmes Génétiques Interactifs (*Interactive Genetic Algorithms*, IGA), pour lesquels les évaluations de la fonction-objectif sont fournies par l’utilisateur·rice. Ces algorithmes sont utilisés lorsque la fonction-objectif est particulièrement difficile à quantifier, ou lorsqu’il n’existe pas de modèle mathématique pour celle-ci (MEIGNAN et al. 2015 ; TAKAGI 2001). Nous retrouvons des applications de ce type d’algorithme à de nombreux problèmes, tels que :

- La conception de planches de bord de voiture « dégagées » (POIRSON, PETIOT, BOIVIN et al. 2013),
- La conception de verres à pied « élégants » (POIRSON, PETIOT, BOIVIN et al. 2013),
- Le réglage personnalisé d’appareils auditifs (TAKAGI et OHSAKI 2007) et d’implants cochléaires (LEGRAND et al. 2007).

D’autres exemples d’applications sont notamment présentés dans TAKAGI (2001) et PEI et TAKAGI (2018).

Dans un IGA, les solutions potentielles au problème sont appelées individus et définies par un ensemble de variables de conception et représentées par des chaînes de nombres

appelées chromosomes. Un chromosome est composé de plusieurs sections, appelées gènes. Dans les algorithmes abordés dans ce manuscrit, chaque gène code la valeur d'une variable de conception. Il existe cependant d'autres manières de coder les valeurs des variables de conception dans les gènes d'une solution. La figure 1.2 illustre les notions de chromosome et de gène. Les solutions sont évaluées en groupe, appelée population. L'état de la population

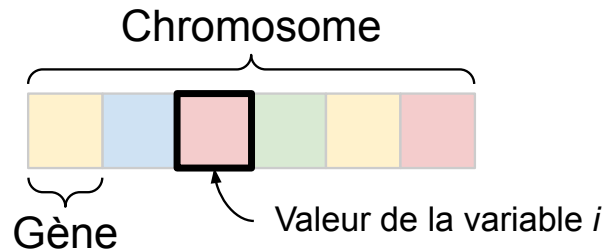


FIGURE 1.2 – Schéma d'un chromosome à six gènes. Chaque gène code la valeur d'une variable de conception. Sur le schéma, les valeurs de gènes sont représentées par des couleurs.

à une itération donnée est appelé génération. Par exemple, l'ensemble des individus à la première itération de l'algorithme est appelé première génération. La figure 1.3 illustre les notions de population et d'individu. Des opérateurs, appelés opérateurs génétiques, sont appliqués aux individus de chaque génération pour créer la suivante. Dans un IGA sous sa forme classique, ces opérateurs sont :

- La **sélection**, qui consiste à choisir les individus qui serviront à générer les individus de la génération suivante, en fonction des évaluations de l'utilisateur-riche. La sélection a pour but de favoriser la propagation des gènes des individus qui maximisent la fonction d'évaluation,
- Le **croisement**, qui consiste à mélanger les chromosomes de plusieurs individus (appelés parents), afin d'en produire de nouveaux (appelés enfants),
- La **mutation**, qui consiste à modifier un individu en changeant aléatoirement la valeur d'un de ses gènes.

Ces opérateurs peuvent être appliqués en série (chaque opérateur est appliqué à chaque individu), en parallèle (chaque individu n'est modifié que par un opérateur), ou bien dans une combinaison des deux. La première génération est le plus souvent choisie aléatoirement dans l'espace de conception. Les étapes d'évaluation et de création de la génération suivante sont répétées jusqu'à ce qu'un critère d'arrêt prédéfini soit satisfait. Le critère d'arrêt peut par exemple être d'atteindre un nombre prédéterminé d'itérations. La figure 1.4 illustre le processus d'optimisation avec un IGA.

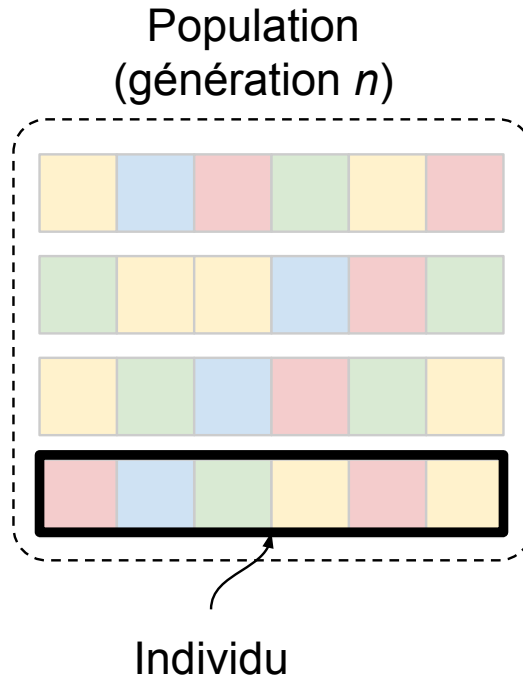


FIGURE 1.3 – Illustration des notions de population et d’individu d’un IGA. Dans cette exemple fictif, la population représentée comporte quatre individus et correspond à la génération  $n$ .

L’opérateur de sélection va tendre à uniformiser la population au cours des générations, puisqu’il favorise la transmission des gènes des meilleurs individus. Pour les IGAs, la notion de convergence se réfère à l’homogénéisation de la population au cours des générations. Celle-ci est illustrée sur la figure 1.5. La probabilité d’amélioration des solutions diminue à mesure que l’homogénéité de la population augmente. Ainsi, plus la population est homogène et moins le changement de valeur moyenne de la fonction-objectif sur la population d’une génération à l’autre est important. À l’inverse, l’opérateur de mutation va tendre à diversifier la population. Le nombre d’individus par génération, les types d’opérateurs, la manière dont ils sont appliqués et les paramètres correspondants sont choisis en fonction du problème à résoudre. Généralement, ces réglages sont effectués au moyen de simulations, faisant appel à une fonction approximant la réponse d’un-e utilisateur-riche de manière simplifiée (PEI et TAKAGI 2018). Le but de ce réglage est d’adapter la vitesse de convergence au nombre de générations disponibles et en particulier d’éviter une convergence prématurée vers un optimum local. A contrario, sans mécanisme de sélection et de transmission de gènes des meilleurs individus, l’IGA se comporte comme une recherche aléatoire.

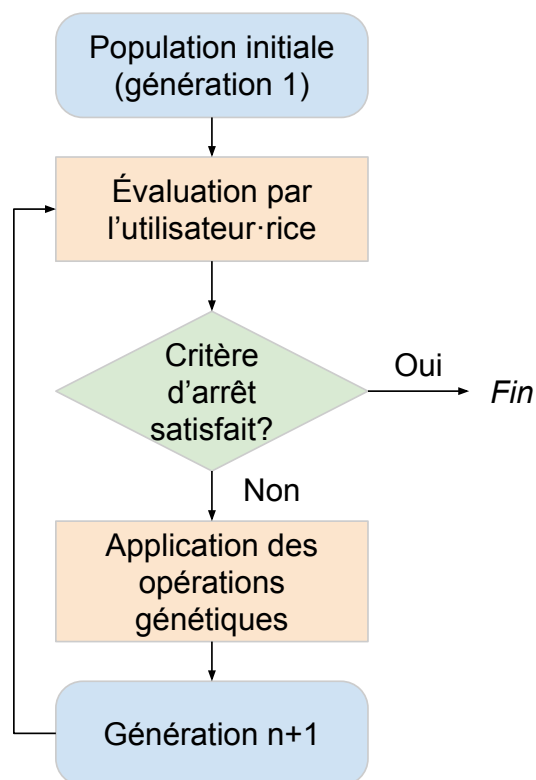


FIGURE 1.4 – Processus d'un IGA. La première génération est le plus souvent choisie aléatoirement. À chaque itération, l'utilisateur-riche évalue les individus de la population selon un attribut perceptif donné. Des opérateurs génétiques sont appliqués aux individus de la génération qui vient d'être évaluée pour créer la génération suivante. Le processus est répété jusqu'à ce que le critère d'arrêt soit satisfait. Si le critère d'arrêt est satisfait, la procédure s'arrête avant de créer la génération suivante.

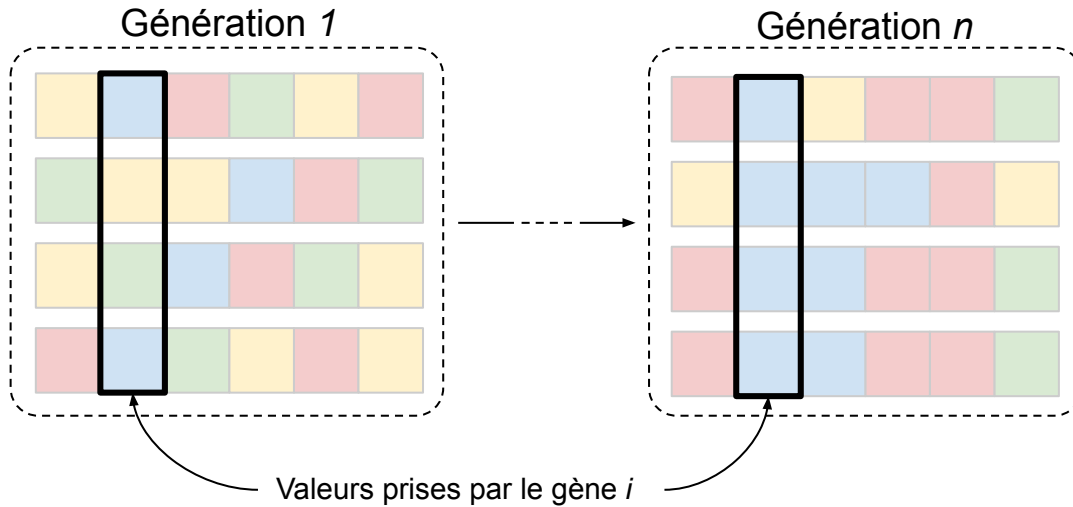


FIGURE 1.5 – Illustration de la convergence d’un Algorithme Génétique Interactif. Ici la population est constituée de quatre individus. Chaque individu est représenté par un chromosome à six gènes. La valeur de chaque gène est représentée par une couleur. À la génération  $n$ , nous observons une homogénéisation des valeurs de chaque gène au sein de la population.

De manière générale, les IECs sont robustes au bruit propre aux évaluations humaines et aux fluctuations de ces dernières au cours du processus d’optimisation (TAKAGI 2001). L’optimum global trouvé avec un IEC n’est souvent pas un point, mais plutôt une zone de l’espace de conception. Les solutions trouvées à la fin du processus d’optimisation sont généralement de légères variations les unes des autres, correspondant à des valeurs similaires de la fonction-objectif.

Dans leur article de revue, PEI et TAKAGI (2018) mettent en avant trois grands sujets de recherche dans les publications concernant les algorithmes évolutionnaires interactifs : la réduction de la fatigue des utilisateur·rice·s, l’étude de la psychologie (émotions, perception) et l’application des IECs à des problèmes d’optimisation concrets. La fatigue des utilisateur·rice est une des principales limites des IECs. De nombreuses stratégies sont proposées pour y remédier, en particulier concernant les algorithmes et le type de retour demandé à l’utilisateur·rice. Pour les IGAs en particulier, beaucoup de variantes proposent par exemple aux utilisateur·rice·s d’avoir un plus grand contrôle sur le processus d’optimisation, en plus d’évaluer les solutions proposées. Cette stratégie requiert cependant que l’utilisateur·rice de l’IGA possède une bonne connaissance du problème à résoudre et/ou du processus d’optimisation. Une autre approche très répandue consiste à construire un modèle des évaluations de l’utilisateur·rice pendant le processus d’opti-

misation, afin de guider ce dernier (PEI et TAKAGI 2018). Le cas qui nous intéresse se situe à l’intersection entre les deux points restants : nous souhaitons étudier la perception humaine, afin de d’aider un processus de design sonore pour des cas d’application concrets. PEI et TAKAGI (2018) soulignent que l’intérêt des IEC réside dans leur capacité à mettre à jour des informations autrement inaccessibles et des solutions innovantes à certains problèmes. Par exemple, l’utilisation d’un IEC a permis de trouver un réglage d’implant cochléaire satisfaisant du point de vue de l’utilisateur·rice, mais ne respectant pas la stratégie habituellement suivie pour un réglage manuel (LEGRAND et al. 2007).

Dans la plupart des publications sur les IGAs, ces derniers sont utilisés dans une optique de personnalisation ou d’aide à la création. Ainsi, l’application des IGAs en audio consiste, pour une bonne partie, à aider un·e musicien·ne ou un·e designer·euse sonore à explorer l’espace des paramètres d’un synthétiseur (YEE-KING 2016 ; JOHNSON 1999 ; DAHLSTEDT 2001 ; MCDERMOTT, O’NEILL et GRIFFITH 2010). Seules quelques études proposent d’utiliser les données récoltées avec un IGA pour étudier la perception humaine, ou pour extraire des informations valables à l’échelle d’un groupe, et non plus seulement individuellement (TAKAGI 2012). En particulier, plusieurs travaux effectués au Laboratoire des Sciences du Numérique de Nantes (LS2N) proposent d’utiliser des IGAs dans ce but, en application à l’étude de l’aspect « innovant » (POIRSON, PETIOT, ALIOUAT et al. 2010), « dégagé » (POIRSON, PETIOT, BOIVIN et al. 2013) ou « sportif » (PETIOT, CERVANTES CHAVEZ, BOIVIN et al. 2015) de planches de bord, de la notion d’« élégance » de verres à pied (POIRSON, PETIOT, BOIVIN et al. 2013), ou encore de la « sévérité » de sons d’alerte (PETIOT, DENJEAN et al. 2020).

La figure 1.6 illustre la place des IGAs dans la taxonomie de MEIGNAN et al. (2015). Dans la version la plus simple d’un IGA (indiqué par les cases entourées d’un trait plein), l’utilisateur·rice ne fait qu’évaluer les solutions proposées. Les évaluations sont utilisées telles quelles pour créer la génération suivante, sans construire de modèle. L’information fournie par l’utilisateur·rice est donc uniquement valable au moment où elle est fournie. Cependant, comme nous l’avons évoqué dans cette section, de nombreuses variations existent, indiquées par les cases entourées d’un trait pointillé. Au cours de cette thèse, nous nous limitons à l’utilisation d’IGAs avec les caractéristiques les plus simples possibles (donc en trait plein dans la figure 1.6), car nous souhaitons étudier des attributs perceptifs chez des utilisateur·rice·s n’ayant a priori pas de connaissances approfondies en design sonore ou en optimisation.



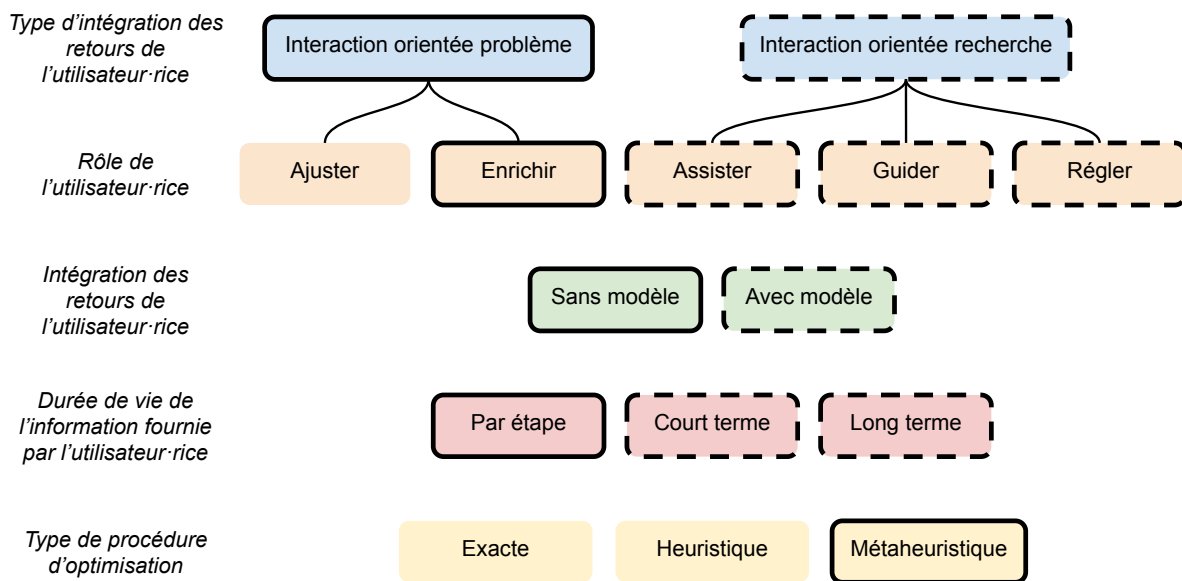


FIGURE 1.6 – Taxonomie des méthodes d'optimisation interactive proposée par MEIGNAN et al. (2015) (traduction libre). Les caractéristiques correspondant aux algorithmes génétiques interactifs, dans leur version la plus simple, sont entourées d'un trait plein. Celles qui sont souvent retrouvées dans des variantes des IGAs sont entourées d'un trait pointillé.

### 1.2.3 protocole expérimental

Dans les travaux effectués au cours de cette thèse, nous considérons le protocole expérimental basé sur une expérience d'optimisation interactive illustré par la figure 1.7. En comparaison avec un protocole basé sur l'étude d'un corpus, un protocole basé sur une expérience d'optimisation interactive permet d'explorer de plus grands espaces de conception, sans nécessité de construire de modèle prédictif des attributs perceptifs d'intérêt. Cette approche permet d'obtenir directement des exemples de solutions qui répondent aux critères choisis. Pour caractériser l'ensemble des sons de l'espace sonore d'intérêt, des approches telles que celles décrites dans les sections 1.1.2 et 1.1.4 sont en revanche plus appropriées (PETIOT, CERVANTES CHAVEZ, BOIVIN et al. 2015).

Le protocole proposé ne permet cependant pas d'explorer un corpus et nécessite la définition de variables de conception. Dans les cas d'applications traités au cours de cette thèse nous considérons uniquement des variables de conception quantitatives et discrètes ou des variables de conception catégorielles. Ainsi, l'espace de conception généré par ces variables comprend un nombre fini de designs. L'approche proposée nécessite alors de définir également les niveaux que peuvent prendre ces variables (également appelés

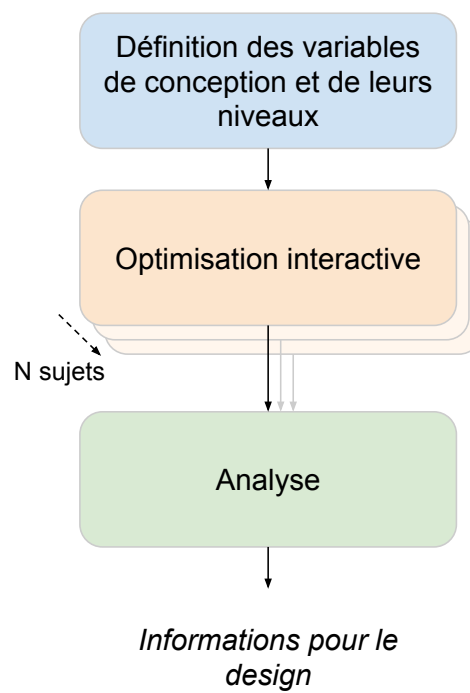


FIGURE 1.7 – Protocole expérimental de design sonore interactif.

modalités).

### 1.3 Son et mobilité

Le secteur des transports constitue un cadre d'application privilégié du design sonore. En effet, quelle que soit sa nature, un moyen de transport est nécessairement une source sonore, qui peut être pensée via une approche de design sonore (MISDARIIS 2021). Le son produit par un véhicule peut être conséquent, dû notamment à son mécanisme de propulsion (par exemple un moteur à combustion interne) ou aux interactions mécaniques entre le moyen de transport et le milieu dans lequel il se déplace (par exemple, le contact des roues d'une voiture sur une route). Cependant, des sons intentionnels peuvent être ajoutés pour des raisons de sécurité ou de fonctionnalité. Par exemple, au début du XIX<sup>e</sup> siècle les locomotives à vapeur ont commencé à être équipées d'un sifflet à vapeur pour signaler leur approche aux personnes présentes aux abords d'un chemin de fer. Ainsi, l'expérience d'usage d'un moyen de transport est intrinsèquement lié à la dimension sonore. L'industrie automobile s'est en particulier emparée de cette question depuis la fin des années 1990, intégrant le son dans l'identité du véhicule (SUSINI, HOUIX et MISDARIIS 2014), comme l'illustre l'étude du son de fermeture de porte citée précédemment (PARIZET, GUYADER et NOSULENKO 2008 ; BEZAT 2007).

Ces dernières années, des formes de mobilité voient le jour, qui soulèvent de nouvelles questions en termes de design sonore. Tout d'abord, les véhicules électriques (VE) et hybrides (VH) qui, à basse vitesse (en dessous de 30 km/h ; PARIZET, ROBART et al. (2014)) sont bien moins bruyants que des véhicules avec moteur à combustion interne (*Internal Combustion Engine*, ICE). Ces deux types de véhicule, également appelés véhicules silencieux (VS), présentent un danger potentiel pour les autres usager·ère·s de la route. Cette application constitue un cas d'étude particulièrement intéressant pour le design sonore, dans la mesure où l'ajout d'un son artificiel pour des raisons de sécurité vient poser des questions sur l'expérience utilisateur, notamment en termes de pollution sonore (MISDARIIS 2017). Depuis la fin des années 2000, plusieurs législations ont été mises en place pour imposer l'ajout d'un son d'alerte (*Acoustic Vehicle Alerting System*, AVAS) permettant de signaler la présence du VS, notamment en Europe (EU REGULATION 2014), mais également au niveau mondial via les Nations Unies (UN REGULATION 2017). Ces législations spécifient notamment un niveau sonore minimum par bande de fréquence, afin de s'assurer de l'émergence du son d'alerte, par rapport au fond sonore ambiant. Cepen-

dant, il reste une certaine liberté aux constructeurs quand au choix final du son à utiliser. Ainsi, plusieurs études tentent de mettre à jour les propriétés du timbre qui pourraient rendre un son plus détectable, à niveau sonore fixé. Le premier cas d'application considéré au cours de cette thèse de doctorat porte sur le design de sons d'alerte de véhicule silencieux. En particulier, nous nous intéressons à la problématique de la conception d'un son qui soit facilement détectable, tout en étant le moins désagréable possible (minimisant ainsi l'impact sur l'environnement sonore existant).

Les véhicules autonomes (VA) constituent un second cadre applicatif qui tend à émerger (MISDARIIS, CERA et RODRIGUEZ 2019). Bien qu'un véhicule totalement autonome reste à ce jour au stade conceptuel, laisser (même partiellement) le contrôle du véhicule à l'intelligence embarquée impose une réflexion quant aux modes d'interaction appropriés. Dans un contexte où il n'est potentiellement plus nécessaire de focaliser en permanence son attention sur la route et le tableau de bord, le son constitue une modalité pertinente pour communiquer des informations aux passager·ère·s. Cependant, comme pour les VS, se pose la question sur la manière d'ajouter des sons artificiels sans impacter négativement l'expérience utilisateur. Le deuxième cas d'application considéré dans cette thèse est la conception d'un son d'intérieur de véhicule autonome qui soit peu intrusif.

## 1.4 Questions de recherche

Au travers de ces deux exemples applicatifs, nous souhaitons tester le protocole expérimental proposé dans la section 1.2.3. Dans un premier temps, nous souhaitons savoir si les Algorithmes Génétiques Interactifs présentent un intérêt dans le cas d'un problème de design sonore. Nous intéressons ainsi à la partie « Optimisation interactive » de la figure 1.7, via une première question de recherche :

### Première question de recherche

Un Algorithme Génétique Interactif (IGA) peut-il permettre de minimiser simultanément plusieurs attributs perceptifs d'un son ?

Pour y répondre, nous nous penchons, dans le chapitre 2, sur le design d'un son d'alerte de VS au travers de deux contributions expérimentales. Dans un second temps, nous nous intéressons à la partie « Analyse » de la figure 1.7 au travers d'une seconde question de recherche :

**Seconde question de recherche**

Comment passer des résultats individuels obtenus avec une méthode d'optimisation interactive à des recommandations générales de design ?

Nous abordons également cette question au travers du problème de design d'un son d'alerte de VS. Dans le chapitre 3, nous proposons une méthode d'analyse et de recommandation, que nous testons via une troisième expérience. Puis nous nous penchons sur la première partie du protocole présenté sur la figure 1.7 (« Définition des variables de conception et de leurs niveaux »), au travers d'une troisième question de recherche :

**Troisième question de recherche**

Comment réduire l'espace de conception de manière informée, dans le cadre d'une expérience d'optimisation perceptive interactive ?

Via une quatrième expérience basée des simulations numériques du problème de conception de sons d'alerte de VS précédent, nous proposons, dans le chapitre 5, une méthode de réduction d'espace de conception. Cela nous amène à poser une quatrième et dernière question de recherche :

**Quatrième question de recherche**

La méthode de réduction d'espace de conception proposée dans le cadre d'une expérience simulée est-elle efficace dans le cadre d'une expérience réelle ?

Nous y répondons en nous intéressant, dans le chapitre 4, au problème de design d'un son d'intérieur peu intrusif pour le véhicule autonome, sur lequel portent les cinquième et sixième contributions expérimentales. La figure 1.8 résume le contenu traité dans chaque question de recherche.

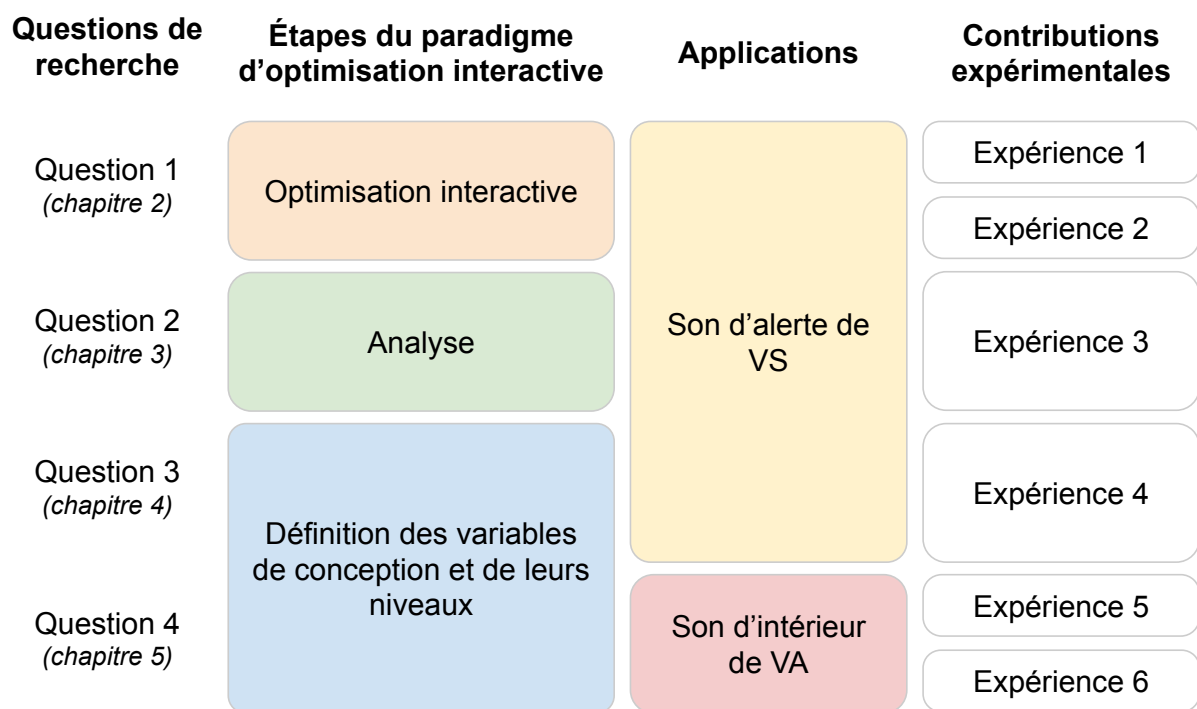


FIGURE 1.8 – Résumé des questions de recherche, avec les étapes du protocole expérimental d'optimisation interactive, les problèmes applicatifs traités et les contributions expérimentales correspondants.

## 1.5 Conclusion du premier chapitre

Dans ce chapitre, nous avons introduit le contexte général dans lequel s’inscrivent les travaux effectués pendant cette thèse de doctorat. Après une définition du design sonore, nous avons mis en évidence l’importance des études sur la perception sonore pour informer ce dernier et avons décrit quelques méthodes classiquement utilisées, qui se basent généralement sur l’étude d’un corpus de sons. Dans un second temps, nous avons décrit un autre protocole expérimental, basé sur le concept d’optimisation interactive. Ce dernier permet potentiellement d’explorer un espace sonore plus grand que le protocole précédent, car il ne vise pas à caractériser chaque son de l’espace de conception, mais à trouver ceux qui répondent au problème posé. Nous avons en particulier abordé les Algorithmes Génétiques Interactifs (IGAs), qui ont déjà été utilisés pour intégrer la perception humaine dans le cas du design d’objets. Nous avons ensuite décrit les deux cadres applicatifs traités dans ce manuscrit, à savoir le design de sons d’alerte de véhicule silencieux (VS) et celui de sons d’intérieur de véhicule autonome (VA). Enfin, nous avons présenté les quatre questions de recherche traitées au cours de cette thèse, en lien avec le protocole expérimental d’optimisation interactive présenté précédemment, ainsi que les cadres applicatifs correspondants.

# OPTIMISATION DE SONS DE VÉHICULE SILENCIEUX À L'AIDE D'UN ALGORITHME GÉNÉTIQUE INTERACTIF

---

Les travaux présentés dans ce manuscrit proposent d'intégrer l'expérience utilisateur dans le processus de conception en design sonore via une méthode d'optimisation interactive (voir chapitre 1). Ce dernier vise à récolter des connaissances sur la perception sonore afin d'informer le processus de design. Il se présente en trois parties : la définition de l'espace de conception, l'expérience d'optimisation interactive à proprement parler et enfin l'analyse des résultats obtenus par un groupe d'utilisateur·rice·s. Parmi les algorithmes d'optimisation interactive les plus utilisés figure la famille des Algorithmes Évolutionnistes, et en particulier celle des Algorithmes Génétiques Interactifs (IGAs). Or, peu d'études portent sur l'utilisation de ces derniers pour l'étude de la perception sonore. Une première étape consiste donc à déterminer si les IGAs peuvent être utilisés dans le protocole que nous proposons.

Nous étudions cette question au travers du problème du design de sons externes d'un véhicule silencieux (VS), terme qui regroupe notamment les Véhicules Électriques (VE) et les Véhicules Hybrides (VH). Ces véhicules émettent bien moins de son que les véhicules avec moteur à combustion interne (ICE, *Internal Combustion Engine*) pour des vitesses inférieures à 30km/h (PARIZET, ROBERT et al. 2014), ce qui peut présenter un danger pour les autres usager·ère·s de la route. Des législations existent donc pour imposer l'ajout d'un son d'alerte artificiel à ces véhicules (EU REGULATION 2014; UN REGULATION 2017). Cependant, un enjeu important est d'éviter que l'ajout de ce son ne contribue négativement à la pollution sonore, notamment en milieu urbain. Il s'agit d'un problème de design sonore complexe, que nous proposons de poser sous la forme d'un problème d'optimisation bi-objectifs. Ainsi, nous souhaitons optimiser un son de VS selon deux attributs perceptifs simultanément :



1. son caractère désagréable (appelé **désagrément**), que nous voulons minimiser,
2. sa détectabilité, que nous voulons maximiser. Nous mesurons celle-ci au travers du temps qu'il faut pour entendre le son du véhicule, lorsque celui-ci s'approche dans un environnement urbain (appelé **temps de réaction**). L'objectif est alors de minimiser le temps de réaction mesuré.

Dans ce chapitre, nous cherchons donc à savoir si les IGAs présentent un intérêt pour ce type d'application, au travers de la première question de recherche traitée dans ce manuscrit :

#### Première question de recherche

Un Algorithme Génétique Interactif (IGA) peut-il permettre de minimiser simultanément plusieurs attributs perceptifs d'un son ?

Les travaux présentés dans ce chapitre sont la continuation de travaux précédents effectués au sein du Laboratoire des Sciences du Numérique de Nantes sur ce sujet (PETIOT, LEGEAY et LAGRANGE 2019). Dans une première partie, nous faisons un état de l'art des études portant sur la conception de sons de VS peu désagréables et/ou rapidement détectables, ainsi que sur l'utilisation des Algorithmes Génétiques Interactifs (IGAs) pour l'étude de la perception sonore. Dans une seconde partie, nous abordons la première contribution expérimentale de cette thèse, qui porte sur l'utilisation d'un IGA pour résoudre le problème de design sonore décrit ci-dessus. Enfin, une troisième partie porte sur une deuxième expérience visant à évaluer le désagrément et la détectabilité perçus des sons obtenus avec l'IGA. Les travaux décrits dans ce chapitre ont fait l'objet d'une publication dans la revue Design Science (SOUAILLE, PETIOT, MISDARIIS et al. 2022).

## 2.1 État de l'art

### 2.1.1 Désagrément et détectabilité des sons de VS

#### Propriétés acoustiques liées au désagrément et à la détectabilité

Le désagrément (en anglais, *unpleasantness*) est souvent étudié dans les travaux portant sur la qualité de l'environnement sonore, et en particulier celle de l'environnement urbain (LAVANDIER et DEFREVILLE 2006 ; DEFREVILLE 2005). Ces travaux s'intéressent

notamment aux contributions de différentes sources sonores au désagrément perçu, dont les véhicules. Dans ce même contexte, la gêne sonore (en anglais, *annoyance*) est parfois étudiée plutôt que le désagrément (LEIBA 2017). Ces deux attributs sont proches, mais tandis que le désagrément est plutôt de l'ordre d'une préférence, la gêne est un processus qui empêche d'effectuer une tâche, généralement de nature cognitive (GUSKI, FELSCHER-SUHR et SCHUEMER 1999). Peu d'études portent sur le design de sons de VS en intégrant en même temps les percepts de désagrément et de détectabilité. C'est pourquoi nous choisissons ici d'inclure également les études traitant à la fois de la gêne et de la détectabilité de sons de VS.

Dans ROBART et al. (2013) et PARIZET, ROBART et al. (2014), les auteurs s'intéressent au temps de réaction et au désagrément causé par le son externe des VS. Ces travaux, effectués dans le cadre du projet européen eVADER (Electric Vehicle Alert for Detection and Emergency Response), cherchent à intégrer des connaissances en psychoacoustique dans le design de son d'alerte de VS. Trois facteurs sont étudiés : le nombre d'harmoniques, les modulations de fréquence au cours du temps et les modulations d'amplitude au cours du temps. Dans les expériences menées, chacun de ces trois facteurs peut prendre trois modalités. Le plan d'expérience complet est donc constitué de  $3^3 = 27$  sons différents. Les sons sont produits à partir d'un modèle de synthèse additive (ROADS 1996). Dans une première expérience visant à mesurer le temps de réaction d'une personne à l'arrivée d'un VS, le recours à un plan factoriel fractionnaire (BERENGUT 2006) permet d'étudier les effets principaux des trois facteurs en évaluant seulement neuf sons parmi les 27. Les auteurs concluent qu'un faible nombre d'harmoniques est mieux détectable qu'un grand nombre, lorsque les signaux sont égalisés en sonie (ZWICKER et FASTL 2013). Cela vient du fait que lorsque le nombre de tonalités est faible, leurs amplitudes individuelles sont plus élevées. De plus, il apparaît qu'une modulation d'amplitude permet de détecter les sons de VS plus rapidement en présence d'un bruit de fond. Dans une deuxième expérience, les auteurs demandent à des sujets de noter ces mêmes sons sur une échelle continue allant de « Pas du tout désagréable » à « Extrêmement désagréable ». Les sons perçus comme les moins désagréables sont ceux ne présentant pas de modulation d'amplitude. Les auteurs mettent ainsi en évidence l'antagonisme entre le temps de réaction et le désagrément. Il semble donc complexe de trouver un son qui soit à la fois facilement détectable et peu désagréable. Chaque son évalué constitue un compromis différent entre les deux quantités, représenté dans le plan (*désagrément, temps de réaction*). Le son d'alerte représentant le meilleur compromis est celui situé le plus proche de l'origine de ce plan. Les meilleures

performances restent cependant obtenues par un son de moteur Diesel, dit thermique ou ICE (Internal Combustion Engine), également évalué pendant l'étude.

Dans (LEE et al. 2017), les auteur·rice·s proposent deux descripteurs acoustiques pour la gêne sonore et la détectabilité de sons de VS : le « Whine Index » et le « Extended Whine Index », respectivement. Ces deux descripteurs sont basés sur la mesure du contenu tonal des sons d'alerte. Des évaluations par des sujets montrent une forte corrélation entre ces descripteurs et les sons étudiés ( $r = 0,86$  pour le « Whine Index » et  $r = 0,85$  pour le « Extended Whine Index »), mais ces derniers sont peu nombreux (quatre sons et trois conditions de fond sonore). Afin d'identifier le son qui constitue le meilleur compromis, la même méthode que (PARIZET, ROBART et al. 2014) est employée, à savoir une représentation des sons évalués dans le plan (*gêne, temps de réaction*). Le meilleur son est alors celui qui semble minimiser le mieux ces deux quantités.

Dans (STEINBACH et ALTINSOY 2019), les auteur·rice·s cherchent à expliquer la gêne perçue de sons de VS à partir de plusieurs descripteurs acoustiques classiques (Niveau sonore en dB(A), Sonie, Rugosité, Acuité, Force de fluctuation et Tonalité). Deux types de modèles de gêne sont entraînés sur des données perceptives : un modèle linéaire sans interaction ( $r = 0,79$ ) et un modèle basé sur un réseau de neurones artificiels ( $r = 0,99$ ). L'influence individuelle des descripteurs acoustiques d'entrée est également étudiée pour les deux modèles. Cependant, aucun modèle n'est proposé pour la détectabilité. Une première expérience sert à éliminer, parmi les 90 sons de départ, ceux pour lesquels le temps de réaction dépasse un seuil de sécurité.

D'autres auteur·rice·s (MISDARIIS, GRUSON et SUSINI 2013 ; POVEDA-MARTINEZ et al. 2017) se sont intéressé·e·s à la détectabilité de sons d'alerte de VS, à partir de corpus de sons. Dans (MISDARIIS, GRUSON et SUSINI 2013), deux descripteurs acoustiques sont utilisés pour caractériser le corpus initial : la planéité spectrale, qui mesure le caractère bruité d'un son, et le taux de modulation d'amplitude, qui mesure les modulations de l'enveloppe énergétique du son. Dans (POVEDA-MARTINEZ et al. 2017), les sons sont classés selon ces deux mêmes facteurs, mais de manière qualitative. Les deux études concluent que l'influence de ces deux facteurs sur le temps de réaction des sujets dépend du fond sonore.

## Méthodes expérimentales utilisées

Une première approche (PARIZET, ROBART et al. 2014) consiste donc à définir des propriétés acoustiques d'intérêt et une méthode de synthèse sonore dont les paramètres de

contrôle correspondent à ces propriétés acoustiques. La paramétrisation des sons permet de réduire le nombre d'évaluations nécessaires à l'étude grâce à un plan factoriel fractionnaire. Une étude plus poussée incluant les interactions entre ces paramètres demanderait en revanche bien plus d'évaluations. Cependant, les nombres de facteurs et de modalités étudiés restent limités. De plus, comme le désagrément et le temps de réaction sont antagonistes, la construction de modèles linéaires n'est pas suffisante pour permettre de trouver un compromis. Celui-ci est trouvé uniquement parmi les sons faisant partie du plan d'expérience et il n'est pas possible d'extrapoler aux sons qui n'ont pas été évalués.

Dans une deuxième approche (LEE et al. 2017; MISDARIIS, GRUSON et SUSINI 2013; POVEDA-MARTINEZ et al. 2017; STEINBACH et ALTINSOY 2019), la relation entre le désagrément (ou la détectabilité) et des descripteurs acoustiques d'intérêt est étudiée grâce à l'évaluation d'un corpus de sons par un groupe de sujets. L'avantage est ici que les descripteurs sont calculés à partir des sons et qu'il est donc possible d'en étudier de nouveaux sans avoir à conduire de nouvelles expériences perceptives. La qualité des informations recueillies dépend cependant de la diversité des sons du corpus vis-à-vis des descripteurs étudiés. Enfin, de même que dans l'approche précédente, la connaissance du lien individuel entre ces facteurs et les quantités à minimiser ne suffit pas pour pouvoir trouver un compromis.

Ces méthodes permettent de mettre à jour les liens entre des propriétés acoustiques et les attributs perceptifs étudiés. Cependant, l'objectif étant dans notre cas de trouver des sons qui minimisent ces attributs perceptifs, il nous est possible d'aborder le problème de design comme un problème d'optimisation. L'utilisation d'un IGA permet alors d'explorer l'espace de conception défini par des variables contrôlant un synthétiseur. Les possibles interactions entre les variables de conception sont implicitement prises en compte dans le fonctionnement de l'IGA. Il n'est donc pas nécessaire de définir en avance un modèle, comme dans la première approche décrite ci-dessus. De plus, nous pouvons minimiser conjointement le désagrément et le temps de réaction en utilisant un IGA multi-objectifs.

### 2.1.2 Utilisation des IGAs pour l'étude de la perception sonore

Dans cette section, nous soulignons le manque de données quantitatives disponibles dans la littérature concernant l'efficacité des IGAs pour minimiser (ou maximiser) un attribut perceptif d'un son, qui soit différent d'une préférence individuelle. Une des particularités du protocole expérimental que nous proposons est que les sujets choisis ne sont pas experts en design sonore, au sens où il-elle-s n'ont pas nécessairement de connais-

sances avancées dans des champs tels que l’ingénierie audio, la composition musicale ou la pratique d’un instrument de musique (ZATTRA et al. 2021). La plupart des recherches sur l’utilisation d’IGAs, ou plus généralement d’algorithmes évolutionnistes, pour l’optimisation de sons, ont pour cadre applicatif la création d’outils pour des personnes ayant une activité professionnelle dans le champ de la création sonore (musicien·ne, designer·euse, artiste, compositeur·rice, ...) ou bien une pratique musicale avancée. Dans ce cadre, l’évaluation que fournit l’utilisateur·rice est alors une note ou un choix basés sur des préférences personnelles sans spécification d’un attribut perceptif particulier à minimiser (ou maximiser). Peu de données quantitatives sont alors disponibles sur l’efficacité des algorithmes proposés. Certaines études se focalisent plutôt sur une évaluation qualitative de l’utilisation de l’outil correspondant, comparée à une exploration manuelle des paramètres de contrôle d’un synthétiseur (YEE-KING 2016) ou à d’autres outils proposés (MIKI et al. 2006 ; SEAGO 2013). De plus, lorsque les IGAs sont utilisés comme aide à la création, l’interaction avec l’utilisateur·rice ne se limite bien souvent pas à l’évaluation des sons. Les interfaces permettent généralement un contrôle avancé du processus d’évolution, par exemple en choisissant le taux de mutation ou les sons parents pour un croisement. Ces méthodes sortent donc du cadre d’étude traité dans ce manuscrit (voir figure 1.6, chapitre 1).

Au contraire, les études fournissant des données quantitatives sur les résultats du processus d’optimisation s’intéressent plutôt à des non experts. WILSON et FAZENDA (2019) utilisent un IGA comme aide au mixage pour des ingénieur·e-s du son, mais visent aussi un public novice, et observent une amélioration de la valeur de la fonction *fitness* moyenne au cours des générations. Fonction *fitness* est le nom parfois utilisé pour désigner la fonction-objectif d’un IGA, qu’il faut alors maximiser (voir section 1.2.2). Cependant, les sujets évaluent là les sons selon un critère de préférence, plutôt qu’un attribut perceptif particulier. Les auteurs soulignent dans une autre publication (WILSON, FAZENDA et al. 2016) portant sur une étude utilisant la même méthode, mais pour une autre application (le design de timbre de guitare électrique), que la fonction *fitness* pourrait tout aussi bien être un attribut tel que « Chaud » ou « Clair ». Par ailleurs, dans FUKUMOTO et HIJIKURO (2020) et FUKUMOTO, YAMAMOTO et OGAWA (2013), les auteur·rice·s observent également une augmentation de la valeur moyenne de la fonction *fitness* au cours du processus d’optimisation. Dans (FUKUMOTO, YAMAMOTO et OGAWA 2013), il s’agit du design de sons d’alerte, à partir de mélodies, où les sujets doivent évaluer la ressemblance du son entendu avec un son d’alerte. Une grande valeur de la fonction *fitness* dans ce cas cor-

respond à une grande ressemblance avec un son d’alerte. Dans FUKUMOTO et HIJIKURO 2020, les sujets doivent évaluer leur préférence pour un son ASMR (*Autonomous Sensory Meridian Response*, BARRATT et DAVIS (2015)). Il n’existe, à notre connaissance, pas d’étude portant de l’utilisation d’un IGA pour répondre à un problème de design sonore multi-objectifs, hormis des travaux précédents effectués au Laboratoire des Sciences du Numérique de Nantes (PETIOT, LEGEAY et LAGRANGE 2019).

## 2.2 Expérience 1 : Optimisation bi-objectifs de sons de véhicule silencieux

### 2.2.1 Objectif

L’objectif de cette première expérience est d’évaluer la capacité d’un IGA à optimiser des sons vis-à-vis de deux attributs perceptifs simultanément : le désagrément et la détectabilité.

### 2.2.2 Matériels et Méthodes

#### Stimuli

**Scénario** Les sons de VS à évaluer sont présentés via une simulation sonore similaire à celles utilisées dans (PARIZET, ROBART et al. 2014; MISDARIIS, GRUSON et SUSINI 2013; LEE et al. 2017; STEINBACH et ALTINSOY 2019). La simulation considère une personne qui attend devant un passage piéton de pouvoir traverser la route, comme illustré sur la figure 2.1. Un VS arrive aléatoirement de la gauche ou de la droite et passe devant le sujet à une vitesse constante de 20 km/h. Le sujet reste au même endroit et doit indiquer le moment où il-elle entend le véhicule approcher. La distance initiale entre le véhicule et le sujet est de 30 mètres.

**Sons de VS** Les sons de VS sont générés grâce à une méthode de synthèse sonore additive (ROADS 1996). La structure du modèle de synthèse est issue d’une analyse de sons implémentés par des fabricants, ainsi que de propositions faites lors de précédentes études sur le sujets (PETIOT, KRISTENSEN et MAIER 2013; PETIOT, LEGEAY et LAGRANGE 2019), dont les travaux présentés dans ce chapitre sont une continuation. Ce

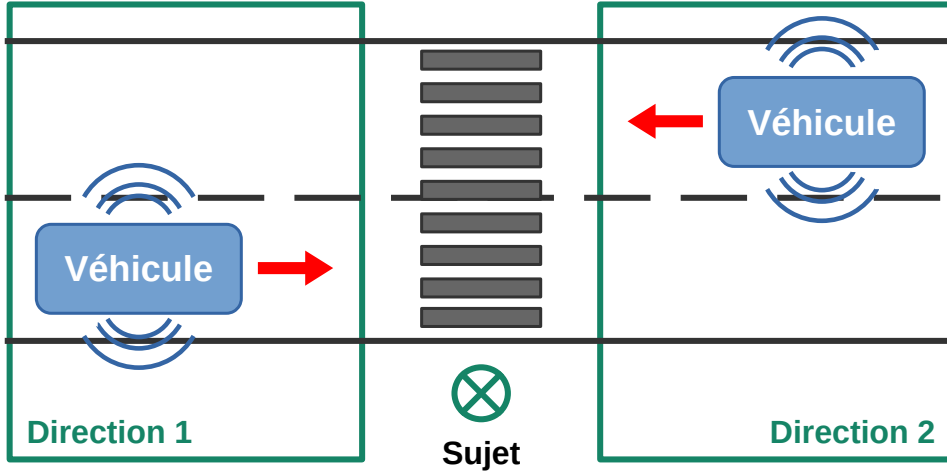


FIGURE 2.1 – Scénario de passage du VS : placement du sujet. Le sujet attend sur le trottoir, face à un passage piéton, de pouvoir traverser la route.

modèle est composé d'une partie dite « harmonique » et d'une partie bruitée. La partie « harmonique » comporte deux modules :

- $M_1$ , un son approximant celui d'un moteur à combustion interne à quatre cylindres, modélisé par une somme de composantes :

$$M_1(t) = \sum_{i=1}^6 a_{1_i} \sin(2\pi f_{1_i} t) \quad (2.1)$$

La fréquence fondamentale  $f_1$  est liée au régime moteur ( $RPM$ , en tour par minute) par la relation  $f_1 = \frac{RPM}{60}$ . Six composantes  $f_{1_i} \in \{0.5f_1, f_1, 1.5f_1, 2f_1, 4f_1, 6f_1\}$  sont considérées, avec les amplitudes  $a_{1_i} \in \{0.2, 0.4, 0.5, 0.2, 0.4, 0.6\}$  (DESOUVRE et al. 2008).

- $M_2$ , un son musical, composé de trois sons harmoniques. Les 3 notes choisies sont une fondamentale ( $f_2$ ), la tierce majeure ( $\frac{5}{4}f_2$ ), et la quinte juste ( $\frac{3}{2}f_2$ ), formant ainsi un accord parfait majeur. Chaque note est composée de six harmoniques. Par exemple, pour la note fondamentale, nous avons les harmoniques  $f_{2_i} \in \{f_2, 2f_2, 3f_2, 4f_2, 5f_2, 6f_2\}$ , avec les amplitudes  $a_{2_i} \in \{1, 0.4, 0.4, 0.1, 0.1, 0.1\}$ .

$$M_2(t) = \sum_{i=1}^6 a_{2_i} \left( \sin(2\pi f_{2_i} t) + \sin(2\pi \frac{5}{4} f_{2_i} t) + \sin(2\pi \frac{3}{2} f_{2_i} t) \right) \quad (2.2)$$

La partie bruitée est composée de deux modules :

- $M_3$ , créé à partir d'une somme de 150 sinusoides de fréquences et de phases aléatoires, tirées de lois uniformes :

$$M_3(t) = \sum_{i=1}^{150} \sin(2\pi f_{3_i} t + \phi_{3_i}) \quad (2.3)$$

avec  $f_{3_i} \in [0, 2f_3]$  et  $\phi_{3_i} \in [0, 2\pi]$ .

- $M_4$ , identique à  $M_3$ , mais couvrant une plus grande bande de fréquences ( $f_4 > f_3$ ) :

$$M_4(t) = \sum_{i=1}^{150} \sin(2\pi f_{4_i} t + \phi_{4_i}) \quad (2.4)$$

avec  $f_{4_i} \in [0, 2f_4]$  et  $\phi_{4_i} \in [0, 2\pi]$ .

La son résultant  $s(t)$  est une somme pondérée de ces quatre modules, modulée en amplitude avec une indice de modulation  $h$  et une fréquence  $f_m$  :

$$s(t) = (1 + h \cdot \sin(2\pi f_m t)) \cdot (a_1 \cdot M_1(t) + a_2 \cdot M_2(t) + a_3 \cdot M_3(t) + a_4 \cdot M_4(t)) \quad (2.5)$$

Les amplitudes  $a_i$  des quatre modules sont normalisés de telle sorte que :

$$\sqrt{\sum_{i=1}^4 a_i^2} = 1 \quad (2.6)$$

Le rapport entre  $a_3$  et  $a_4$  est fixé à 0.5. La figure 2.2 montre des spectrogrammes correspondant à chacun des quatre modules<sup>1</sup>.

Le synthétiseur possède en tout plus de 70 paramètres de contrôle indépendants. Par exemple, les fréquences et amplitudes de tous les modules sont modifiables au cours du temps, ce qui permet de générer une grande variété de sons. Cependant, il n'est pas réaliste d'explorer un espace de design si grand avec un IGA. Six variables de conception ont donc été définies après plusieurs essais, via l'utilisation d'un outil de prototypage. Cet outil, créé pour l'occasion, consiste en une interface graphique composée de boutons rotatifs assignables aux paramètres de contrôle du synthétiseur, permettant ainsi de tester l'effet des variables choisies sur le son produit de manière interactive. Les six variables de conception choisies contrôlent le contenu fréquentiel et les amplitudes des quatre modules  $M_i$ , ainsi que la modulation d'amplitude appliquée au signal total. Elle peuvent chacune prendre quatre modalités (ou niveaux), détaillées dans la table 2.1, choisies de sorte à

---

1. Les sons correspondants peuvent être écoutés à l'adresse <https://souaille.github.io/conception/#ch2>



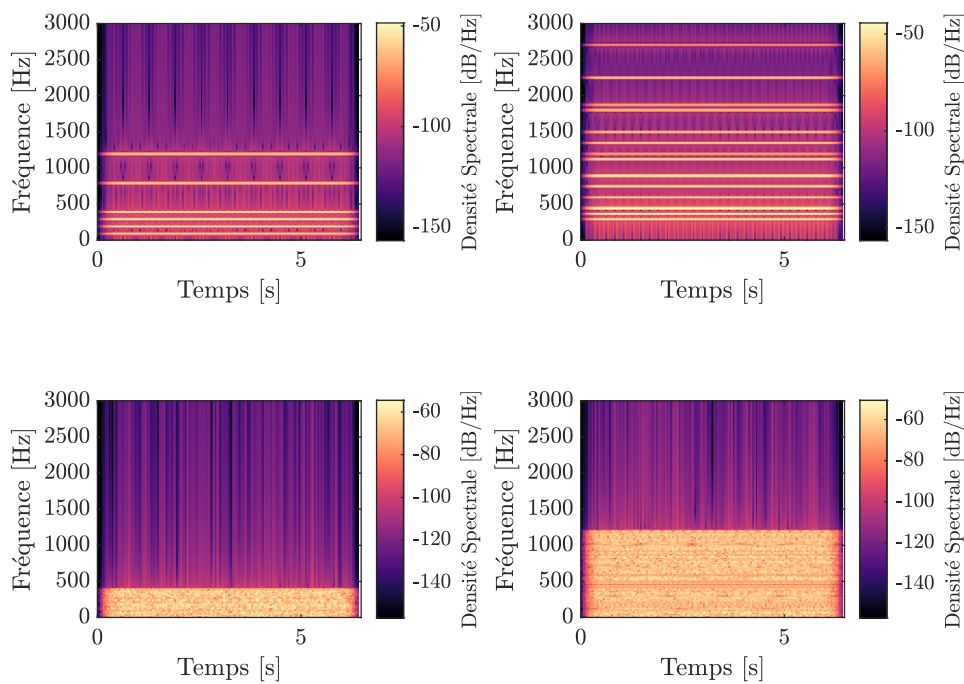


FIGURE 2.2 – Spectrogrammes des quatre modules du synthétiseur de son d’alerte de VS, sans modulation d’amplitude. En haut à gauche : son de moteur  $M_1$ , avec  $f_1 = 200\text{Hz}$ . En haut à droite : son musical  $M_2$ , avec  $f_2 = 300\text{Hz}$ . En bas à gauche : première partie bruitée  $M_3$ , avec  $f_3 = 200\text{Hz}$ . En bas à droite : deuxième partie bruitée  $M_4$ , avec  $f_4 = 600\text{Hz}$ .

TABLE 2.1 – Description des variables de conception manipulées par l’IGA. Les valeurs indiquées correspondent à une vitesse de 20km/h.

Variable de conception	Niveau 1	Niveau 2	Niveau 3	Niveau 4
$V_1$ : Proportion Moteur/accord	$a_2 = 0$	$\frac{a_1}{a_2} = 2$	$\frac{a_1}{a_2} = \frac{1}{2}$	$a_1 = 0$
$V_2$ : Fréquence	$f_1 = 80\text{Hz}$	$f_1 = 120\text{Hz}$	$f_1 = 160\text{Hz}$	$f_1 = 200\text{Hz}$
	$f_2 = 120\text{Hz}$	$f_2 = 180\text{Hz}$	$f_2 = 240\text{Hz}$	$f_2 = 300\text{Hz}$
	$f_3 = 80\text{Hz}$	$f_3 = 120\text{Hz}$	$f_3 = 160\text{Hz}$	$f_3 = 200\text{Hz}$
	$f_4 = 240\text{Hz}$	$f_4 = 360\text{Hz}$	$f_4 = 480\text{Hz}$	$f_4 = 600\text{Hz}$
$V_3$ : Proportion composantes harmoniques/bruit	$a_1 = a_2 = 0$	$\frac{a_3+a_4}{a_1+a_2} = 2$	$\frac{a_3+a_4}{a_1+a_2} = \frac{1}{2}$	$a_3 = a_4 = 0$
$V_4$ : Nombre d’harmoniques	Moteur : 1	Moteur : 2	Moteur : 3	Moteur : 6
	accord : 1	accord : 2	accord : 3	accord : 6
$V_5$ : Fréquence de la modulation d’amplitude	$f_m = 0.5\text{Hz}$	$f_m = 2\text{Hz}$	$f_m = 5\text{Hz}$	$f_m = 10\text{Hz}$
$V_6$ : Taux de modulation d’amplitude	$m = 0\%$	$m = 17\%$	$m = 33\%$	$m = 50\%$

pouvoir produire une grande diversité de sons. Le nombre total de configurations possibles (plan factoriel complet) est donc de  $4^6 = 4096$ . Les paramètres restants du synthétiseur sont fixés par essai-erreur, de manière à produire des sons satisfaisants.

Ce choix de variables a été motivé par des études montrant que la complexité harmonique et la modulation d’amplitude ont un effet sur la détectabilité (PARIZET, ROBERT et al. 2014 ; LEE et al. 2017 ; MISDARIIS, GRUSON et SUSINI 2013 ; POVEDA-MARTINEZ et al. 2017). Ces propriétés sont affectées par les variables  $V_3$  (Proportion de sons harmoniques/bruit) et  $V_4$  (Nombre d’harmoniques), ainsi que les variables  $V_5$  (Fréquence de la modulation d’amplitude) et  $V_6$  (Taux de modulation d’amplitude). De plus, PETIOT, KRISTENSEN et MAIER (2013) ont montré que la ressemblance d’un son de véhicule silencieux avec celui d’un véhicule à moteur à combustion interne a une influence sur le désagrément perçu. Celle-ci peut être positive ou négative en fonction des personnes. Les variables  $V_1$  (Proportion son de moteur/accord parfait majeur),  $V_3$  (Proportion de sons harmoniques/bruit) et  $V_4$  (Nombre d’harmoniques) permettent notamment de contrôler la ressemblance du son produit avec un tel son. Dans la suite de ce manuscrit, les sons produits par le synthétiseur sont désignés par le 6-uplet des valeurs prises par les six variables de conception. Par exemple, le 6-uplet (1, 4, 3, 2, 2, 4) désigne le son pour lequel  $V_1 = 1$ ,  $V_2 = 4$ ,  $V_3 = 3$ ,  $V_4 = 2$ ,  $V_5 = 2$  et  $V_6 = 4$

Enfin, certains paramètres du synthétiseur, comme les fréquences de contrôle des modules, sont normalement fonction de la vitesse du véhicule, afin d’en informer les autres

usager·ère·s de la route (PETIOT, KRISTENSEN et MAIER 2013). Cependant, pour l'expérience considérée ici, le véhicule se déplace à vitesse constante, ce pourquoi les valeurs des paramètres présentées dans la table 2.1 sont fixes.

**Spatialisation** La position spatiale du véhicule est simulée via les traitements suivants :

- Le niveau sonore du VS dépend de la distance entre ce dernier et le sujet. Nous assimilons le véhicule à une source ponctuelle, omnidirectionnelle et appliquons donc un gain inversement proportionnel à la distance.
- Un effet Doppler est simulé par une variation de la fréquence des sinusôides composant le modèle de synthèse, calculé à partir de la vitesse relative du véhicule par rapport au sujet.
- Le panoramique est contrôlé de sorte que la source soit complètement à gauche (ou à droite) au début du scénario, puis passe progressivement vers le côté opposé, en fonction de la position du véhicule. L'amplitude de chaque canal stéréo est calculée par une « loi du sinus » (« *sine-law panning* », PULKKI 2001).

Pour des raisons de simplicité, le modèle n'inclut pas d'autres sources sonores, telles que les bruits de roulement des pneus sur la route. Afin de réduire la durée de chaque évaluation, l'atténuation du son de VS avec la distance est asymétrique au cours du temps, comme dans (MISDARIIS, GRUSON et SUSINI 2013). À l'approche du véhicule, une loi en  $\frac{1}{r}$  est bien utilisée, mais la diminution de l'amplitude est en  $\frac{1}{r^2}$  une fois le véhicule passé devant le sujet. Ceci est la raison de la forme asymétrique de l'amplitude du son de VS au cours du temps, sur la figure 2.3. Ce choix n'affecte pas la détectabilité, qui est évaluée à l'approche du véhicule. Nous faisons l'hypothèse que ce choix n'affecte pas non plus le jugement de désagrément.

**Fond sonore** Pour plus de réalisme, un fond sonore urbain est présent pendant toute la durée de la scène. Celui-ci est construit à partir d'enregistrements effectués à un carrefour dans Paris intra-muros. Les sons trop émergents (KERBER et FASTL 2008), tels que les bruits de klaxons, en ont été retirés, ainsi que les passages de véhicule proche, qui pourraient être confondus avec le son à évaluer. Ces modifications ont pour but de rendre la piste de fond sonore la plus amorphe possible (MAFFIOLO 1999). La piste obtenue après modifications a une durée de 42 secondes. Afin de limiter l'impression de répétitivité, qui joue sur la fatigue du sujet, un extrait d'environ 10 secondes est sélectionné au hasard

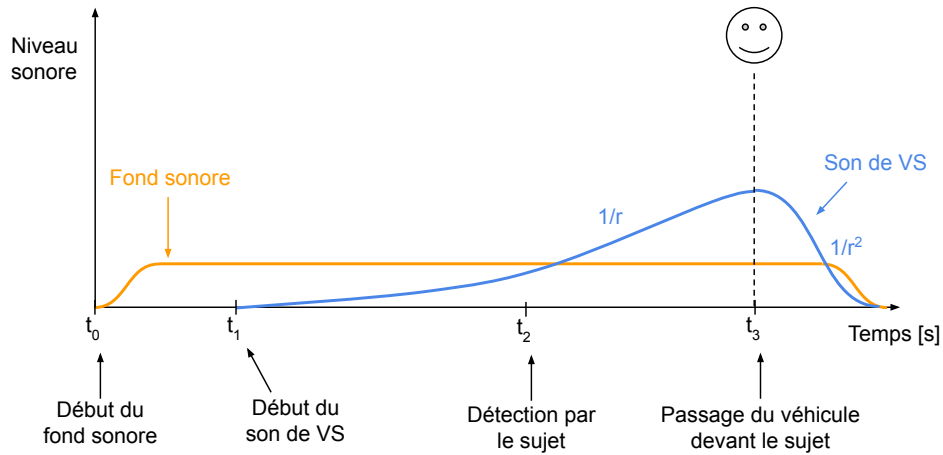


FIGURE 2.3 – Scénario de passage du VS : vue temporelle. Les amplitudes du fond sonore (pointillés) et du son de VS (trait plein) sont affichées au cours du temps.

parmi cette piste à chaque évaluation. L'évolution du niveau du fond sonore au cours du temps figure également sur la figure 2.3. La durée totale de chaque stimulus est de 10 secondes. La figure 2.4 montre le spectrogramme d'un son produit par le synthétiseur, spatialisé et mélangé au fond sonore<sup>2</sup>. Le changement de fréquence dû à l'effet Doppler est particulièrement visible autour du temps  $t_3$ , correspondant au passage devant le sujet.

## Plan d'expérience

Tous les sujets passent la même condition expérimentale, qui consiste à utiliser un IGA pour optimiser des sons de VS grâce à des évaluations perceptives. Chaque sujet procède à l'évaluation de 99 sons, générés pendant le processus d'optimisation.

## Participant·e·s

32 élèves ingénieur·e·s de l'École Centrale de Nantes (16 hommes, 16 femmes) participent à cette première expérience.

2. Le son correspondant peut être écouté à l'adresse <https://souaille.github.io/conception/#ch2>

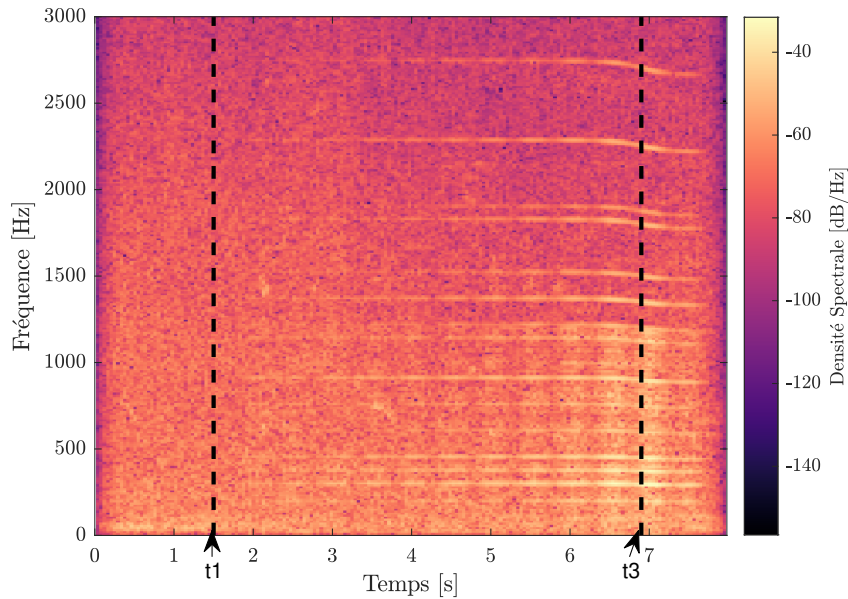


FIGURE 2.4 – Spectrogramme d'un son de VS avec le fond sonore. Le son de VS est spatialisé. Le temps  $t_1$  est le temps auquel de son de VS commence à être joué. Le temps  $t_3$  est le temps auquel le VS passe devant le sujet. Les lignes horizontales correspondent aux différentes composantes du son de VS, qui émerge puis disparaît dans le fond sonore.

## Apparatus

L'expérience a lieu sur les ordinateurs des salles informatiques de l'École Centrale de Nantes. Les sujets écoutent le scénario avec un casque audio Beyerdynamics DT-990 (casques fermés, d'une impédance de 250 Ohm). La carte son interne de l'ordinateur est utilisée. Le niveau sonore en sortie du casque est calibré avec un sonomètre, de sorte à ce que le fond sonore soit à 69 dB(A). Ce niveau a été choisi pour être cohérent avec le niveau sonore en dB(A) mesuré lors de l'enregistrement du fond sonore. Un même gain est appliqué à tous les sons de VS, choisi afin de ne pas avoir trop de sons détectés trop tôt ou trop tard. Ce gain a été ajusté empiriquement, à l'écoute de sons choisis aléatoirement parmi tous les sons de l'espace de conception. Le niveau moyen des sons de VS est compris entre 59 dB(A) et 89 dB(A). 95% des sons ont un niveau au dessus de 68 dB(A).

## Protocole

**Implémentation de l'IGA** L'IGA implémenté s'inspire de l'algorithme NSGA-II (DEB et al. 2000), conçu pour les problèmes d'optimisation multi-objectifs. Dans cet algorithme, les solutions sont comparées via la notion de dominance de Pareto. Une solution

en domine une autre, au sens de Pareto, si elle est meilleure ou équivalente pour tous les objectifs du problème et s'il existe au moins un objectif pour lequel elle est strictement meilleure. Un exemple est présenté sur la figure 2.5, où les solutions sont représentées dans un plan formé de deux objectifs à minimiser. Sur la figure de gauche (figure 2.5(a)), la solution en noir domine toutes les solutions en blanc. On appelle front de Pareto l'ensemble des solutions non dominées, au sens de Pareto, représentées sur la figure de droite (figure 2.5(b)). Les solutions du front de Pareto sont également appelées solution de rang 1. En ignorant ces solutions, les solutions non dominées parmi les solutions restantes sont les solutions dites de rang 2. Les solutions des rangs suivants sont définies de la même manière, comme illustré sur la figure 2.5(b). L'algorithme NSGA-II a pour but de trouver

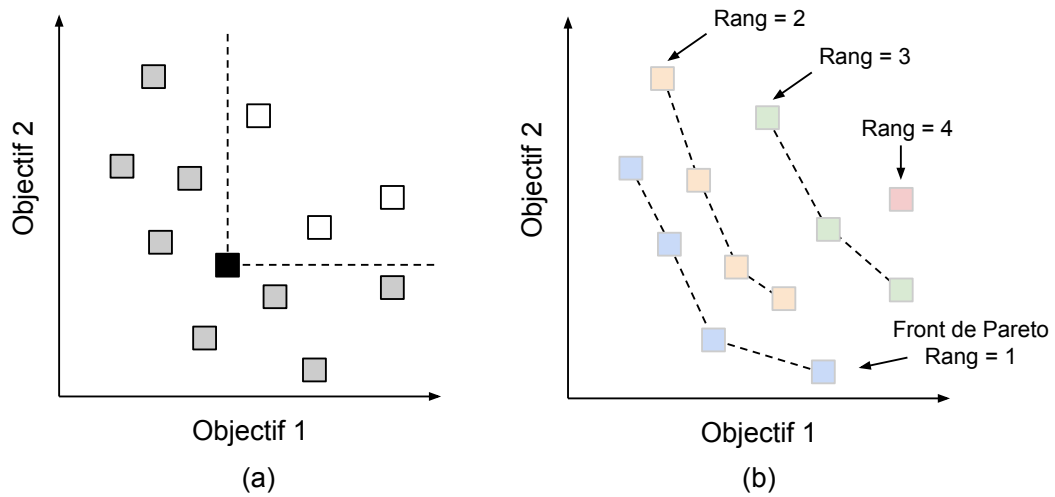


FIGURE 2.5 – Exemple de domination au sens de Pareto, pour un problème bi-objectifs, où les deux objectifs doivent être minimisés. Dans la figure (a), le carré noir domine tous les carrés blancs, mais aucun carré gris. La figure (b) montre les rangs de non domination, numérotés à partir du front de Pareto.

une approximation du front des solutions Pareto-optimales au problème d'optimisation. Une première génération de  $N$  sons est créée via un échantillonnage par hypercube latin (LOH 1996). Chaque son de cette première génération est évalué par le sujet selon les deux propriétés perceptives considérées. Les sons évalués sont ensuite ordonnés selon un opérateur appelé *crowded-comparison operator*, qui fonctionne de la manière suivante :

1. Les sons sont d'abord classés selon leur rang de Pareto.
2. Au sein d'un même rang, les sons sont ordonnés en fonction de leur distance les uns des autres dans le plan (*désagrément, temps de réaction*), via une métrique appelée

*crowding-distance*. Les sons les plus éloignés des autres sont mieux classés. Ceci a pour but de favoriser un espacement régulier des sons au sein de chaque rang. La figure 2.6 illustre le principe de *crowding-distance*

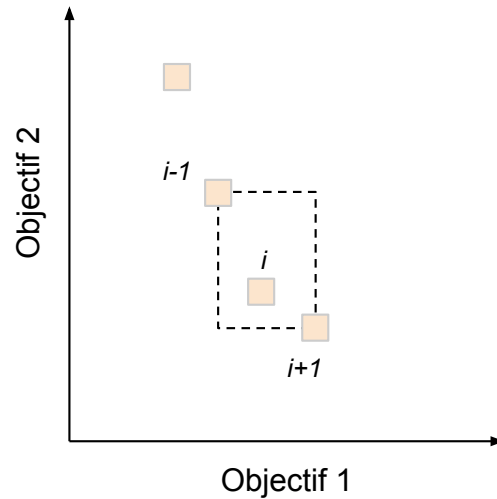


FIGURE 2.6 – Illustration du principe de *crowding-distance*. La valeur de cette métrique pour la solution  $i$  correspond à la moyenne des distances entre les solutions adjacentes  $i - 1$  et  $i + 1$  du même rang sur tous les objectifs. Il s'agit donc la valeur moyenne d'une arête de la boîte indiquée en pointillés. Plus cette valeur est grande et plus la solution est considérée comme éloignée des solutions adjacentes appartenant au même rang. La figure est adaptée de DEB et al. (2000).

Dans cet algorithme, un registre des meilleures solutions est mis à jour après chaque génération. Celui-ci contient le même nombre  $N$  de solutions que dans une génération et est initialisé avec les  $N$  sons de la première génération. À chaque génération, une fois tous les sons évalués, le registre est mis à jour en gardant les  $N$  meilleurs sons, selon la *crowded-comparison operator*, parmi les sons déjà présents dans celui-ci et les sons qui viennent d'être évalués.

Un fois le registre mis à jour, un opérateur est appliqué à chaque son qu'il contient pour créer les  $N$  sons de la génération suivante. Cette partie diffère de l'algorithme NSGA-II original. Elle est adaptée de l'algorithme utilisé lors d'une étude précédente conduite au LS2N, portant sur le même problème, et que poursuivent les travaux présentés dans ce chapitre (PETIOT, LEGEAY et LAGRANGE 2019).

Les opérateurs sont :

- **Copie** : la solution est copiée telle quelle dans la génération suivante.
- **Mutation** : identique à la copie, mais la valeur d'un des gènes de la solution est

changée aléatoirement.

- **Croisement simple** : une deuxième solution est sélectionnée par tournoi binaire, à partir du classement obtenu via le *crowded-comparison operator*. C'est à dire que deux solutions sont tirées au hasard parmi le registre et que la mieux classée est sélectionnée pour effectuer le croisement. Les chromosomes représentant les deux solutions, appelés parents, sont alors recombinaés, pour créer deux nouveaux chromosomes, appelés enfants. Le processus est illustré sur la figure 2.7, pour la création d'un des deux chromosomes enfant. Le choix du chromosome enfant à garder est fait aléatoirement.

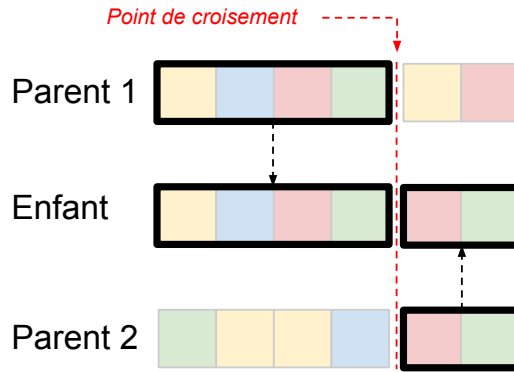


FIGURE 2.7 – Exemple de croisement simple avec deux parents. Chaque case représente un gène dont la valeur est représentée par une couleur. Chaque chromosome parent est séparé en deux au niveau d'un point de croisement. La première section du parent 1 est recombinaée avec la deuxième section du parent 2 pour créer le chromosome enfant.

Chaque opérateur peut s'appliquer avec une probabilité  $p_{co}$ ,  $p_{mu}$  et  $p_{cr}$ , correspondant à la copie, à la mutation, et au croisement, respectivement. La somme de ces probabilités vaut  $p_{co} + p_{mu} + p_{cr} = 1$ . Le choix de l'opérateur à appliquer se fait par génération aléatoire d'un nombre  $r$  compris entre zéro et un, avec une répartition uniforme :

- Si  $0 \leq r \leq p_{co}$ , une copie est appliquée.
- Si  $p_{co} < r \leq p_{co} + p_{mu}$ , une mutation est appliquée.
- Si  $p_{co} + p_{mu} < r \leq 1$ , un croisement avec un autre son est appliqué.

Il est possible qu'après l'application des opérateurs, une solution qui a déjà été évaluée par le passé soit de nouveau présente dans la prochaine génération. Il y a donc un risque que le registre se peuple rapidement de sons identiques, réduisant la diversité des sons proposés au sujet par l'algorithme. Afin d'éviter cela, le registre ne contient qu'un seul exemplaire de chaque son, associé à son évaluation la plus récente.



Pour cette expérience, les probabilités d'application des opérateurs de l'algorithme génétique sont  $p_{co} = 0.05$ ,  $p_{mu} = 0.7$  et  $p_{cr} = 0.25$ . Ces valeurs ont été choisies empiriquement via des simulations du processus d'optimisation interactive dans un premier temps, puis par essai-erreur en testant l'IGA. Les valeurs choisies semblaient permettre une amélioration des solutions trouvées tout au long du processus, avec une stabilisation dans les dernières générations. Le taux de mutation est très élevé afin de préserver la diversité des sons, malgré le faible nombre de sons au sein d'une génération, et d'éviter ainsi une convergence prématurée.

**Interface de test et procédure expérimentale** L'interface utilisée pour l'évaluation est représentée en figure 2.8. Le son à évaluer est généré en cliquant sur le bouton « Initialiser ». Le sujet lance la lecture d'un extrait du fond sonore en appuyant sur la barre « Espace ». Ce moment correspond au temps  $t_0$  de la figure 2.3. Après un délai de durée aléatoire, le VS commence à approcher (temps  $t_1$ ). Le sujet doit alors appuyer la touche « A » de son clavier, dès qu'il-elle entend le véhicule, si celui-ci arrive de la gauche. Si le véhicule arrive par la droite, le sujet doit appuyer sur la touche « E ». Le moment où la touche « A » ou « E » est pressée correspond au temps  $t_2$ . Ce choix des touches « A » et « E » a été effectué pour permettre à l'utilisateur-riche d'évaluer les sons sans avoir à décoller la main du clavier (clavier AZERTY). La correspondance « "A" situé à gauche » et « "E" situé à droite » par rapport à la direction de provenance du véhicule (« A » pour gauche, « E » pour droite) permet une évaluation instinctive de cette direction de provenance. Aucun sujet n'a rapporté de difficultés dans l'utilisation de l'interface. Sur la figure 2.8, un seul son est proposé pour évaluation sur une même page de l'interface. Cependant, dans cette première expérience, chaque page de l'interface présente l'ensemble des sons de la génération en cours.

Afin d'éviter un effet d'habitude, le temps  $t_1$  auquel commence le son de VS est choisi aléatoirement dans l'intervalle  $[1, 3]$  secondes. Étant donnée l'évolution du son au cours du temps (figure 2.3), le temps de réaction vaut :

$$R_t = t_2 - t_1 \tag{2.7}$$

L'intervalle de valeurs que peut prendre  $t_1$  étant restreint, l'arrivée du véhicule reste prévisible, mais un intervalle plus grand rallongerait trop la durée de chaque écoute. Nous considérons que le protocole permet d'évaluer une borne inférieure du temps de réaction, car dans une situation réelle, un véhicule peut arriver à tout moment, sans que le sujet ne

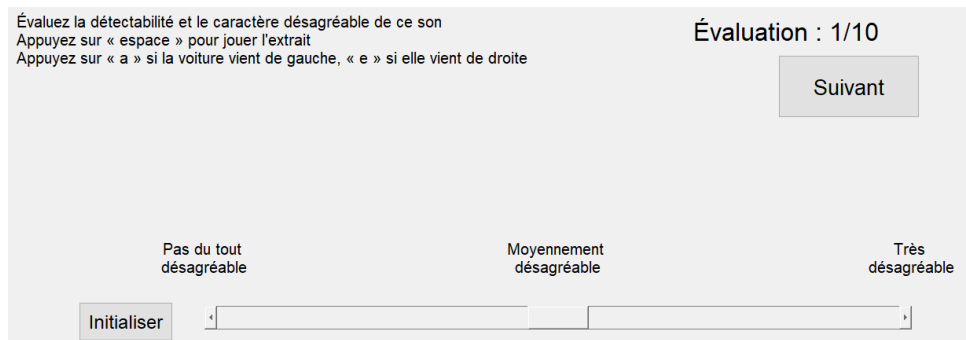


FIGURE 2.8 – Interface pour l’évaluation des sons de VS. L’évaluation du temps de réaction se fait via les touches du clavier. L’évaluation du désagrément se fait grâce à un curseur horizontal sur une échelle continue structurée (*slider*).

s’y attende. Le temps de réaction en situation réelle serait donc vraisemblablement plus grand.

Après avoir écouté le passage du véhicule, le sujet doit évaluer le désagrément causé par le son correspondant, sur une échelle sémantique continue allant de « Pas du tout désagréable » à « Très désagréable ». Cette évaluation se fait grâce à un curseur horizontal, comme celui présent sur la figure 2.8, qui convertit le jugement en une note allant de 0 (« Pas du tout désagréable ») à 10 (« Très désagréable »). Afin d’expliquer la notion de désagrément, le contexte suivant est donné au sujet : « Pour évaluer cette grandeur, imaginez que vous êtes en train de vous reposer, ou de lire, et que vous entendez le bruit de ce VS. Le bruit est-il perturbant, désagréable ? ». Le sujet peut rejouer le stimulus autant de fois que nécessaire, pour effectuer l’évaluation de désagrément, mais la détectabilité est seulement mesurée à la première écoute. Au début du test, le sujet suit un tutoriel pour se familiariser avec l’interface et avec le type de son de VS produit par le synthétiseur.

Chaque sujet évalue 11 générations de 9 sons (soit 99 sons en tout). L’expérience dure environ 30 minutes. Chaque page de l’interface contient donc 9 sons, afin de favoriser la cohérence des jugements fournis au sein d’une même génération. À la fin du test, pour chaque sujet  $i$ , nous obtenons les informations suivantes :

- La liste des 99 sons évalués au cours de l’expérience, ainsi que les valeurs de temps de réaction et de désagrément correspondantes.
- Le nombre d’erreurs commises au cours de l’expérience.
- La liste des sons qui sont sur le front de Pareto du registre après la dernière générations, appelée *Optimset<sup>i</sup>*.

## Analyse

**Gestion des erreurs** Si le sujet appuie sur la touche « A » ou « E » avant que le son de VS n'ait effectivement commencé (c'est-à-dire avant  $t_1$ ), ceci est compté comme une erreur, qui est enregistrée. Nous appelons ce type d'erreur, Détection Prématurée (DP). Le temps de réaction prend alors une valeur par défaut de  $R_t = \frac{t_3 - t_1}{2}$ . Si le VS n'est pas détecté ou est détecté après son passage devant le sujet (c'est-à-dire après  $t_3$ ), le temps de réaction enregistré vaut  $R_t = t_3 - t_1$  et l'erreur, que nous appelons Détection Tardive (DT), est enregistrée. Enfin, si le sujet n'identifie pas correctement la direction d'arrivée du VS, l'erreur est également enregistrée, mais aucune action n'est prise. Nous appelons ce type d'erreur une Erreur de Direction (ED). En rapportant le nombre d'erreurs de chaque type au nombre total de sons évalués pendant l'expérience, nous obtenons des taux qui nous permettent d'évaluer la fiabilité des sujets :

- Le taux d'erreur de direction ( $TED$ ) : proportion des stimuli détectés comme provenant de la mauvaise direction.
- Le taux de détection prématurée ( $TDP$ ) : proportion des stimuli détectés avant le début du son du véhicule (avant  $t_1$ , voir figure 2.3).
- Le taux de détection tardive ( $TDT$ ) : proportion des stimuli détectés après le temps de passage du véhicule devant le sujet (après  $t_3$ , voir figure 2.3).

Nous définissons des valeurs limites pour chaque taux d'erreur, en prenant en compte qu'il est inévitable de faire quelques erreurs lors du test complet. Ces limites sont fixées afin d'identifier les sujets qui ne possèdent pas d'aptitudes auditives suffisantes pour la détection, voire qui n'ont pas réalisé le test sérieusement, pour pouvoir écarter leurs données de l'analyse ultérieure. Pour que les données d'un sujet soient considérées comme valides, il faut que les taux d'erreur soient tous en dessous des limites suivantes :  $TED_{lim} = 40\%$ ,  $TDP_{lim} = 20\%$  et  $TDT_{lim} = 85\%$ . Pour le  $TED$  la limite est proche des 50%. Nous cherchons ainsi à identifier les sujets qui choisiraient la direction d'approche au hasard. La limite du  $TDP$  est assez basse (20%), car nous considérons que des pressions anticipées sur le clavier sont un signe que le-a participant-e cherche à finir le test rapidement, et n'est pas suffisamment impliqué-e. La limite pour le  $TDT$  est relativement élevée, car il n'était pas possible de prendre en compte la totalité des sons du plan d'expérience pour la calibration (4096 sons en tout). En conséquence, nous pensons qu'il est possible que certains sujets doivent évaluer des sons particulièrement difficiles à détecter, sans que cela ne soit dû à leurs aptitudes auditives ou à leur concentration. Si au moins une des conditions n'est pas remplie, les données du sujet correspondant sont écartées.

**Distance de sécurité** La distance entre le véhicule et le sujet au moment de la détection peut être calculée grâce à la vitesse  $V$  du véhicule, car celle-ci est constante :

$$D_d = (t_3 - t_2) \times V \quad (2.8)$$

Nous pouvons alors comparer cette distance à une distance de sécurité, définie comme la distance d'arrêt du véhicule. Cette distance, qui prend en compte le temps de réaction du ou de la conducteur·rice, dépend de la vitesse du véhicule. Dans POVEDA-MARTINEZ et al. (2017), les auteur·rice·s donnent des valeurs de 3,30m, 7,50m et 12,50m à 10km/h, 20km/h et 30km/h, respectivement. Dans notre cas, le véhicule roulant à 20km/h, la distance de sécurité vaut  $D_{d,s} = 7,50\text{m}$ , ce qui correspond à un temps de réaction maximum de  $R_{t,s} = 4,05\text{s}$ .

**Convergence de l'IGA** Afin de visualiser l'évolution de la qualité des solutions au cours du processus d'optimisation, l'indicateur hypervolume  $H_v$  (ZITZLER et THIELE 1999) est calculé à partir des solutions du registre après chaque génération. Cet indicateur est calculé comme l'union des hypervolumes dominés par chaque solution du front de Pareto. Pour calculer ces hypervolumes, un point de référence est défini en amont. Ce point est généralement la pire solution possible, du point de vue de tous les objectifs. La figure 2.9 montre un exemple d'indicateur hypervolume pour l'exemple de problème bi-objectifs utilisé dans la figure 2.5. Une augmentation de  $H_v$  au cours du processus d'optimisation indique une amélioration de la qualité des solutions trouvées (au sens de Pareto), ainsi que de leur diversité (ZITZLER 1999).

## 2.2.3 Résultats

### Fiabilité des sujets

La table 2.2 montre les valeurs moyennes des taux d'erreur pour les sujets de l'expérience 1, ainsi que les écart-types correspondants. Les distributions correspondantes sont également affichées sur la figure 2.10. Le taux moyen d'erreur de direction ( $TED$ ) est faible (7,1%). Seul un sujet a fait un grand nombre d'erreurs de ce type (plus de 40%), que nous avons interprété comme une mauvaise compréhension des instructions. Les données de ce sujet n'ont pas été prises en compte dans l'analyse ultérieure des résultats. Le taux de détection prématurée moyen ( $TDP$ ) est également très faible (0,4%), ce qui montre que les sujets n'ont pas précipité le test. Le taux moyen de détection tardive

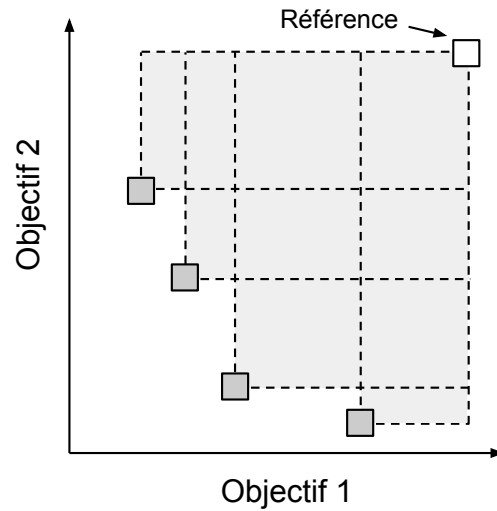


FIGURE 2.9 – Illustration de l'indicateur hypervolume  $H_v$ , pour un problème à deux dimension, dans lequel chaque objectif doit être minimisé. La valeur de  $H_v$  correspond à la surface en gris, qui est l'union des surfaces comprises entre chaque point du front de Pareto (carrés gris) et un point de référence (carré blanc).

TABLE 2.2 – Taux d'erreurs moyens des sujets pour l'expérience 1, avec les écart-types et le nombre de sujets écartés à cause de performances de détection jugées trop faibles (voir section 2.2.2). *TED* : Taux d'erreur de direction. *TDP* : Taux de détection prématurée. *TDT* : Taux de détection tardive.

	Erreur Moyenne (%)	Écart-Type (%)	Sujets Écartés
<i>TED</i>	7,1	10,1	1
<i>TDP</i>	0,4	2,0	0
<i>TDT</i>	17,2	25,3	3

(*TDT*) est en revanche assez élevé (17,2%), ce qui montre que certains sons peuvent être particulièrement difficiles à détecter. Les données de trois sujets, pour lesquels le taux de détection tardive est supérieur à 85%, ont été écartées. Au total, les données de quatre sujets ont été écartées, soit 12,5%, sur les 32 sujets ayant participé à l'expérience.

### Convergence de l'IGA

La figure 2.11 montre l'évolution de la valeur moyenne de l'indicateur hypervolume  $H_v$  sur tous les sujets, après chaque génération. Nous observons que cet indicateur augmente au fur et à mesure du processus d'optimisation, ce qui traduit une amélioration, en moyenne, de la qualité des solutions proposées par l'algorithme. L'amélioration des solutions est plus importante au début du processus et semble stagner à partir de la sixième

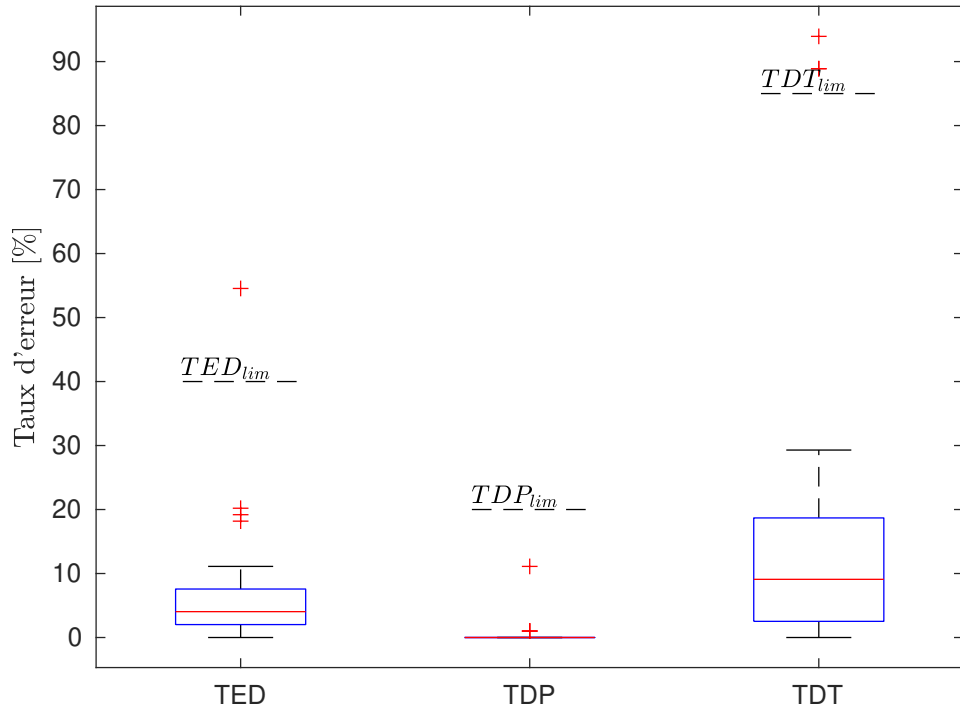


FIGURE 2.10 – Boîtes à moustaches des taux d'erreurs pour l'expérience 1. Les lignes rouges représentent les valeurs médianes des taux d'erreur. Les premiers et troisièmes quartiles sont représentés par les extrémités inférieures et supérieures des boîtes bleues. Les moustaches vont des plus petites aux plus grandes valeurs qui ne sont pas considérées comme des outliers. Les outliers sont représentés par des croix rouges et correspondent aux valeurs situées à plus de 1,5 écarts inter-quartile en-dessous du premier quartile ou au-dessus du troisième quartile. Les lignes noires pointillées horizontales correspondent aux taux d'erreur limites fixés pour que les données du sujet soient considérées comme recevables.  $TED$  : Taux d'erreur de direction.  $TDP$  : Taux de détection prématurée.  $TDT$  : Taux de détection tardive.

génération.

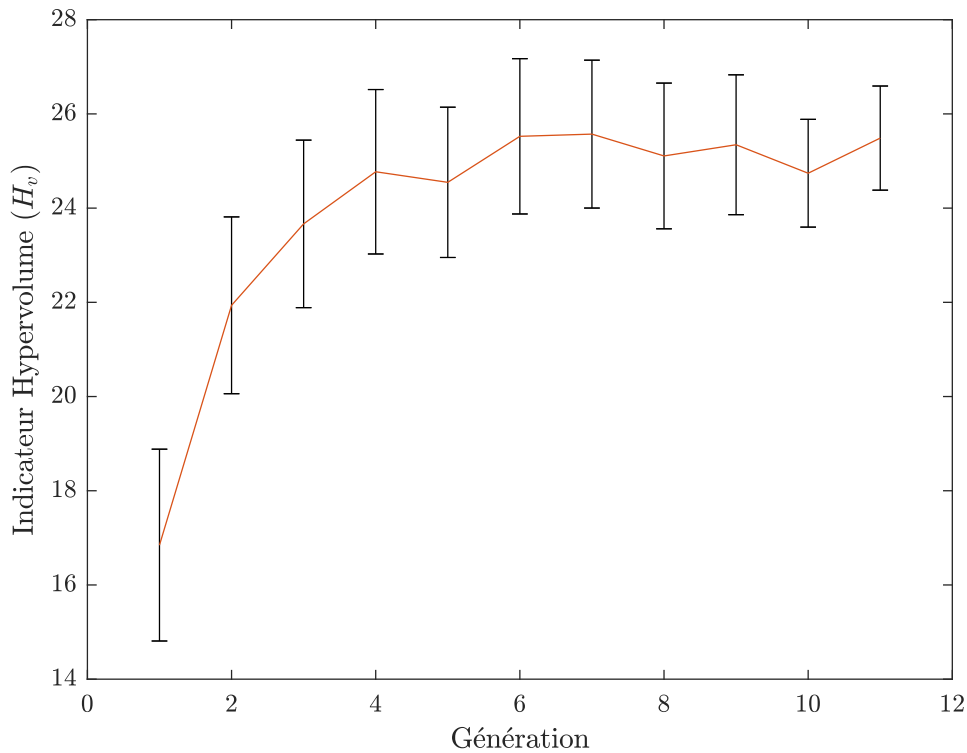


FIGURE 2.11 – Indicateur hypervolume  $H_v$  moyen à chaque génération, avec les erreurs standards correspondantes.

## 2.2.4 Discussion

Cette première expérience montre l'intérêt de l'IGA utilisé pour la minimisation conjointe du désagrément et du temps de réaction de sons de VS. Elle conforte les résultats déjà observés lors d'une précédente étude (PETIOT, LEGEAY et LAGRANGE 2019) sur le même sujet, ainsi que d'autres études sur l'optimisation de sons avec des IGAs (WILSON et FAZENDA 2019; FUKUMOTO, YAMAMOTO et OGAWA 2013; FUKUMOTO et HIJIKURO 2020). Cependant, bien que nous observions une amélioration des solutions trouvées au cours du processus d'optimisation, cette première expérience ne nous permet pas de comparer l'approche IGA avec d'autres méthodes de design sonore. C'est ce que nous souhaitons faire au travers d'une deuxième expérience. Une analyse du reste des résultats de cette première expérience est présentée dans le chapitre 3, notamment concernant les liens entre les variables de conception et les attributs perceptifs que nous cherchons à

minimiser.

## 2.3 Expérience 2 : Validation des meilleurs sons individuels

Nous souhaitons étudier l'intérêt des Algorithmes Génétiques Interactifs (IGA) pour la minimisation simultanée de plusieurs attributs perceptifs de sons. Au cours de la première expérience (expérience 1), nous avons observé une diminution du désagrément perçu et du temps de réaction des sons explorés avec l'IGA. Afin de valider ces résultats, nous souhaitons comparer les sons obtenus grâce à l'IGA avec des sons de référence.

### 2.3.1 Objectifs

Le but de cette expérience est d'évaluer les performances des sons trouvés individuellement avec l'IGA lors de l'expérience 1, comparativement à celles de sons choisis empiriquement et de sons choisis aléatoirement.

### 2.3.2 Matériels et Méthodes

#### Stimuli

**Définition du meilleur son de chaque sujet  $IGA_{opt}^i$**  À l'issue de l'expérience 1, chaque sujet obtient un ensemble de sons appelé  $Optimset^i$ . Au cours du processus d'optimisation, l'IGA met à jour une archive des meilleurs sons. Après que la dernière génération de sons soit évaluée, l'ensemble  $Optimset^i$  est constitué en sélectionnant, parmi cette archive, les sons appartenant au front de Pareto. Un unique meilleur son, que nous appelons  $IGA_{opt}^i$ , est sélectionné dans  $Optimset^i$  pour chaque sujet  $i$ , via la méthode suivante.

Tout d'abord, les solutions trop extrêmes sont éliminées de l'ensemble  $Optimset^i$ . D'une part, les solutions avec un temps de réaction  $R_t$  supérieur à  $R_{t,s} = 4,05s$  sont éliminées, car ces derniers sont détectés trop tard pour respecter la distance de sécurité  $D_{d,s} = 7,5m$  à 20km/h. D'autre part, les sons trop désagréables, relativement aux évaluations de désagrément du sujet sur l'ensemble des sons écoutés, sont également éliminés. Pour cela, une normalisation « Min-Max » est effectuée sur toutes les évaluations du sujet.



Les sons de  $Optimset^i$  dont le désagrément normalisé est situé dans le tiers supérieur de la plage de valeurs ainsi obtenue sont éliminés.

Ensuite, la méthode TOPSIS (HWANG et YOON 1981) est utilisée pour sélectionner un son unique  $IGA_{opt}^i$  parmi les sons restants. Le principe de cette méthode est le suivant. Nous construisons une matrice  $x_{kj}$  à partir des évaluations des sons restants. Chaque indice de ligne de  $x_{kj}$  correspond à un son  $k$  et chaque colonne correspond à un attribut perceptif (ici,  $j = 2$ ). Chaque colonne de  $x_{kj}$  est ensuite normalisée par  $\sum_k x_{kj}^2$  et multipliée par un poids. Nous donnons le même poids de 0.5 aux deux colonnes. Deux « solutions idéales » sont alors définies :

- La meilleure solution idéale (MSI) qui a, pour chaque attribut perceptif, la plus petite valeur au sein des solutions considérées :  $MSI = \{\min(x_{k1}), \min(x_{k2})\}$
- La pire solution idéale (PSI) qui a, pour chaque attribut perceptif, la plus grande valeur au sein des solutions considérées :  $PSI = \{\max(x_{k1}), \max(x_{k2})\}$

Nous calculons, pour chaque son  $k$ , la distance euclidienne de chaque solution restante avec les deux solutions idéales (appelées  $d_{MSI,k}$  et  $d_{PSI,k}$ ). Nous calculons alors le coefficient de proximité  $C$  suivant :

$$C_k = \frac{d_{PSI,k}}{d_{PSI,k} + d_{MSI,k}} \quad (2.9)$$

La solution choisie  $IGA_{opt}^i$ , pour le sujet  $i$ , est celle qui maximise ce coefficient  $C$ , qui vaut 1 si le son correspond à la MSI et 0 s'il correspond à la PSI. Nous obtenons ainsi un son unique pour chaque sujet.

**Sons *Design*** Nous avons créé les sons *Design* après avoir développé une bonne compréhension du synthétiseur et des possibilités offertes par celui-ci. Ces deux sons ont été conçus avec l'intention de proposer deux compromis intéressants, mais différents du point de vue du timbre. Le processus de conception a consisté à expérimenter différentes valeurs des variables de conception tout en écoutant les sons produits, jusqu'à trouver un timbre satisfaisant. Cette démarche, finalement identique à la prise en main de n'importe quel synthétiseur dans le cadre d'une création sonore, a été effectuée avant d'avoir connaissance des résultats de l'expérience 1. Cette approche nous a permis de développer une intuition experte mais personnelle concernant la relation entre les variables de conception et notre interprétation subjective des notions de désagrément et de détectabilité.

Les deux sons retenus sont appelés *design1* et *design2*. Leurs représentations par le

codage sur les variables (voir tableau 2.1) sont respectivement  $design1 = (2, 4, 4, 2, 1, 1)$  et  $design2 = (1, 3, 1, 1, 4, 2)$ . La figure 2.12 présente une description des caractéristiques de ces deux sons.

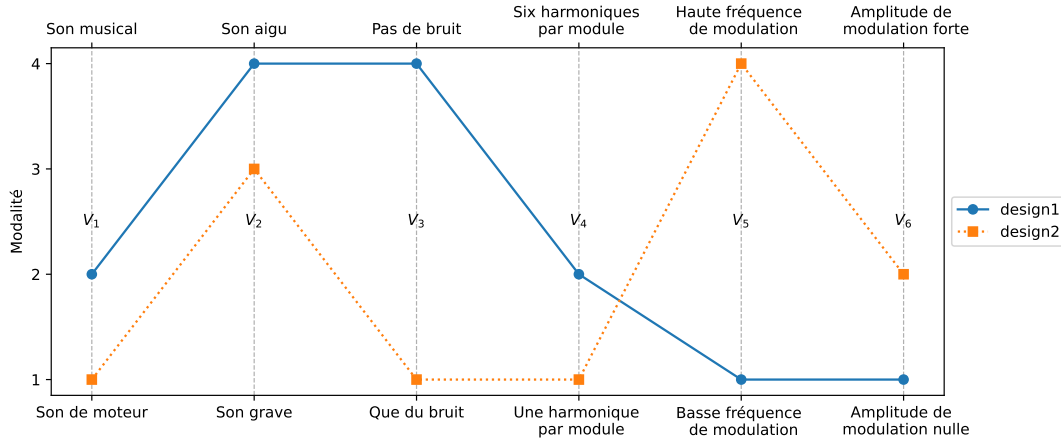


FIGURE 2.12 – Description des sons  $design1$  et  $design2$ .

**Sons *Random*** En plus de ces deux sons, quatre sons *Random* ( $rand1$ ,  $rand2$ ,  $rand3$  et  $rand4$ ) sont générés à partir d'un échantillonnage par hypercube latin. Ces sons aléatoires servent de référence pour évaluer les performances des sons  $design1$  et  $design2$ , ainsi que celles des sons  $IGA_{opt}^i$ . Leurs codages sont respectivement  $rand1 = (3, 4, 1, 2, 2, 3)$ ,  $rand2 = (1, 3, 2, 4, 1, 1)$ ,  $rand3 = (4, 1, 3, 1, 4, 4)$  et  $rand4 = (2, 2, 4, 3, 3, 2)$ . La figure 2.13 présente une description de leurs caractéristiques<sup>3</sup>.

### Plan d'expérience

Chaque sujet  $i$  évalue tous les sons des catégories *Design* et *Random*, ainsi que le son  $IGA_{opt}^i$  extrait à la suite de sa passation de l'expérience 1, ce qui fait sept sons au total. Chaque son est évalué quatre fois, ce qui correspond donc à  $7 \times 4 = 28$  évaluations en tout par sujet. Les sons sont présentés dans un ordre aléatoire.

3. Les sons *Design* et *Random* peuvent être écoutés à l'adresse <https://souaille.github.io/conception/#sec23>

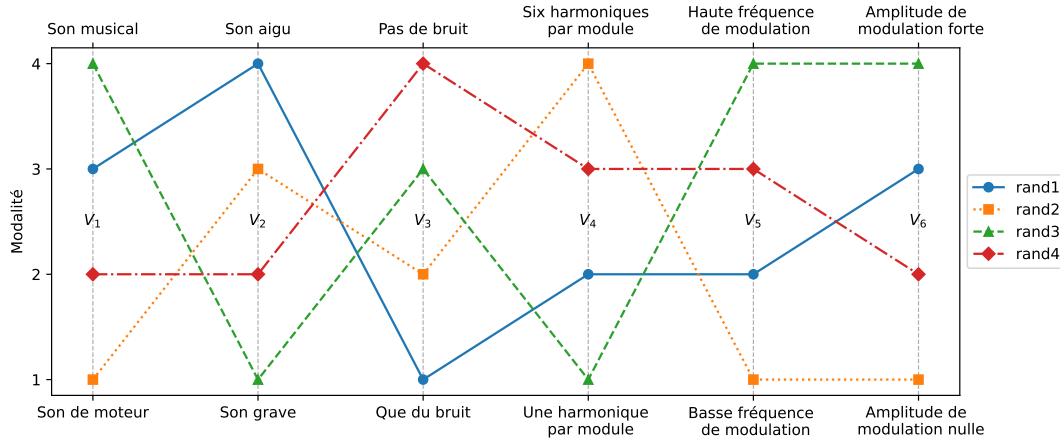


FIGURE 2.13 – Description des sons *rand1* à *rand4*.

## Participant.e.s

24 des 32 sujets ayant participé à l'expérience 1 (voir section 2.2) participent à cette seconde expérience, qui a lieu une semaine après la première. Quatre des huit sujets absents sont les sujets qui avaient été exclus de l'analyse à cause de leurs performances. Les quatre autres ne se sont pas présentés pour la seconde expérience.

## Apparatus

L'expérience a lieu sur les ordinateurs des salles informatiques de l'École Centrale de Nantes. Les sujets écoutent le scénario avec un casque audio Beyerdynamics DT-990 (casques fermés, d'une impédance de 250 Ohm). La carte son utilisée est la carte son interne de l'ordinateur. La calibration est la même que pour l'expérience 1 (voir section 2.2).

## Protocole

**Interface et procédure expérimentale** L'interface et la procédure de présentation et d'évaluation des sons sont les mêmes que pour l'expérience 1. La seule différence est qu'un seul son est présenté par page, et non plus 9. Les sujets n'ont aucune indication visuelle leur permettant d'identifier quel son leur est présenté. À la fin du test, pour chaque sujet  $i$ , nous obtenons 4 évaluations de désagrément (sur une échelle de 0 à 10) et 4 de temps de réaction (en secondes) par son.

TABLE 2.3 – Taux d’erreurs moyens des sujets pour l’expérience 2, avec les écart-types et le nombre de sujets écartés à cause de performances de détection jugées trop faibles (voir section 2.2.2). *TED* : Taux d’erreur de direction. *TDP* : Taux de détection prématurée. *TDT* : Taux de détection tardive.

	Erreur	Moyenne (%)	Écart-Type (%)	Sujets Écartés
<i>TED</i>		2,5	3,2	0
<i>TDP</i>		0	0	0
<i>TDT</i>		6,1	9,1	0

### 2.3.3 Analyse

Nous comparons les performances des sons selon deux critères :

1. La significativité statistique des différences de désagrément moyen, entre chaque paire de sons parmi ( $IGA_{opt}^i$ , *design1*, *design2*, *rand1*, *rand2*, *rand3* et *rand4*). De même pour le temps de réaction. Pour cela nous utilisons un test de comparaison multiple de Tukey,
2. Leur appartenance au front de Pareto, à partir des valeurs moyennes des évaluations fournies par les sujets. Cette analyse ne fait appel à aucun test statistique.

Bien que les sons nommés  $IGA_{opt}^i$  soient différents pour chaque sujet, une seule valeur moyenne de désagrément et de temps de réaction moyens est calculée pour l’ensemble des sujets. À l’inverse, tous les sons *Design* et *Random* sont évalués par l’ensemble des sujets. Le but est de comparer les performances globales de la catégorie des sons  $IGA_{opt}^i$  à celles des autres sons évalués. Nous procédons également à une analyse de la fiabilité des sujets, identique à celle effectuée pour l’expérience 1.

### 2.3.4 Résultats

#### Fiabilité des sujets

La table 2.3 montre les taux d’erreur moyens des sujets pour l’expérience 2. La figure 2.14 montre les distributions correspondantes. Tous les taux d’erreurs moyens sont faibles au regard des critères fixés, et aucun sujet n’a été écarté.

#### Désagrément

Nous calculons les valeurs moyennes de désagrément et de temps de réaction de chaque son, sur tous les sujets et toutes les évaluations. Les moyennes et les erreurs standard des

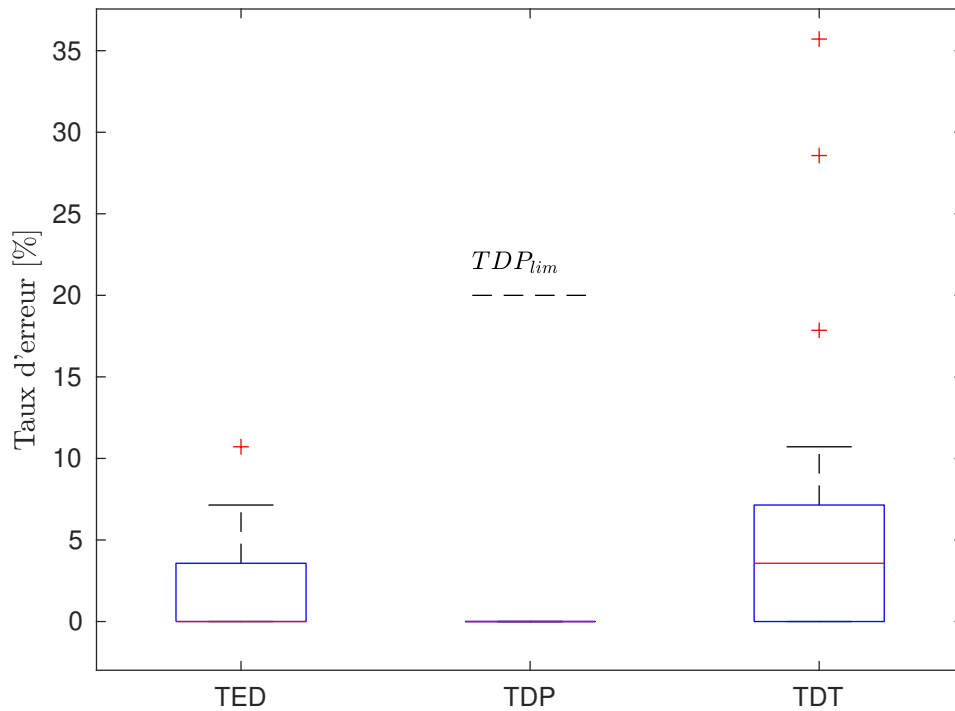


FIGURE 2.14 – Boîtes à moustaches des taux d'erreurs pour l'expérience 2. Les lignes rouges représentent les valeurs médianes des taux d'erreur. Les premiers et troisièmes quartiles sont représentés par les extrémités inférieures et supérieures des boîtes bleues. Les moustaches vont des plus petites aux plus grandes valeurs qui ne sont pas considérées comme des outliers. Les outliers sont représentés par des croix rouges et correspondent aux valeurs situées à plus de 1,5 écarts inter-quartile en-dessous du premier quartile ou au-dessus du troisième quartile. Les lignes noires pointillées horizontales correspondent aux taux d'erreur limites fixés pour que les données du sujet soient considérées comme recevables. Seule la limite pour de TDP apparaît car les autres sont supérieures au taux d'erreur maximal représenté sur la figure. *TED* : Taux d'erreur de direction. *TDP* : Taux de détection prématurée. *TDT* : Taux de détection tardive.

évaluations du désagrément pour chacun des sept sons du corpus ( $IGA_{opt}^i$ , les deux sons *Design* et les quatre sons *Random*) sont affichées dans la figure 2.15.

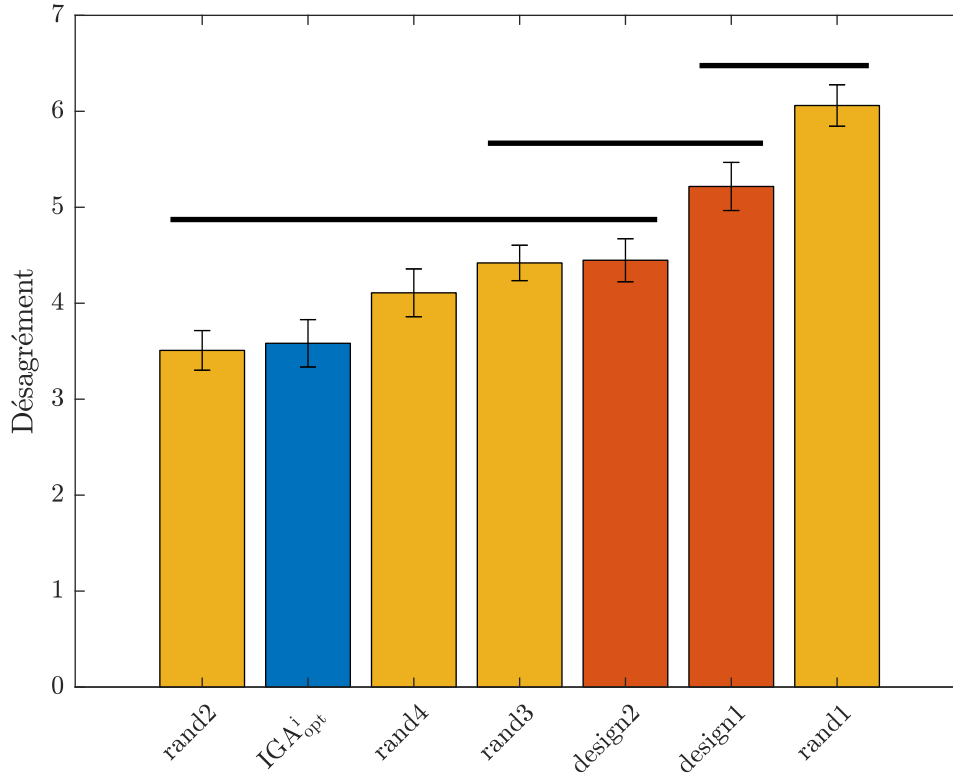


FIGURE 2.15 – Désagrément moyen des sons évalués pendant l’expérience 2, avec les erreurs standards correspondantes et le résultats de l’analyse post-hoc. Les lignes noires horizontales indiquent les groupes de sons au sein desquels les moyennes de désagrément ne sont pas significativement différentes (Test HSD de Tukey,  $p > 0,05$ ). Le son en bleu est  $IGA_{opt}^i$ , les sons en orange sont les sons *Design* et ceux en jaune sont les sons *Random*.

En ce qui concerne le désagrément (figure 2.15), il n’y a pas de différence significative au sein des trois groupes suivants :  $(rand2, IGA_{opt}^i, rand4, rand3, design2)$ ,  $(rand3, design2, design1)$  et  $(design1, rand1)$ . Bien que les sons  $IGA_{opt}^i$  ne soient pas significativement moins désagréables que tous les autres sons, ils font partie des sons les moins désagréables. Ils sont en particulier significativement moins désagréables que le son *design1*.

### Temps de réaction

Les moyennes et les erreurs standard des mesures du temps de réaction pour chacun des sept sons du corpus ( $IGA_{opt}^i$ , les deux sons *Design* et les quatre sons *Random*) sont affichées dans la figure 2.16.

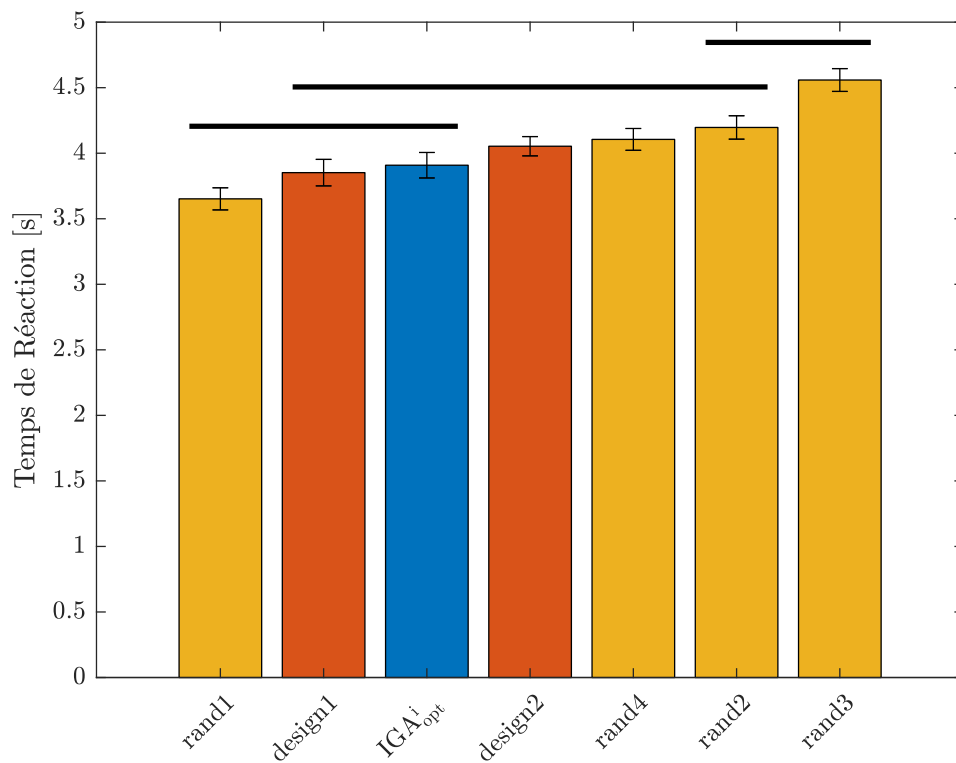


FIGURE 2.16 – Temps de réaction moyen des sons évalués pendant l'expérience 2, avec les erreurs standards correspondantes et le résultats de l'analyse post-hoc. Les lignes noires horizontales indiquent les groupes de sons au sein desquels les moyennes de temps de réaction ne sont pas significativement différentes (Test HSD de Tukey,  $p > 0,05$ ). Le son en bleu est  $IGA^i_{opt}$ , les sons en orange sont les sons *Design* et ceux en jaune sont les sons *Random*.

Nous observons qu'il n'y a pas de différence significative au sein des trois groupes de sons suivants :  $(rand1, design1, IGA_{opt}^i)$ ,  $(design1, IGA_{opt}^i, design2, rand4, rand2)$  et  $(rand2, rand3)$ . De même que pour le désagrément, les sons  $IGA_{opt}^i$  ne sont pas significativement plus rapidement détectables que tous les autres, mais ils sont tout de même parmi les trois sons pour lesquels le temps de réaction est le plus court. En particulier, le temps de réaction pour  $IGA_{opt}^i$  n'est pas significativement différent de celui des deux sons  $design1$  et  $design2$ .

Le temps de réaction  $R_t$  peut être converti en distance entre le véhicule et le sujet au moment de la détection  $D_d$ , via l'équation 2.8. Nous obtenons ainsi les distance moyennes pour chaque son, représentées sur la figure 2.17. Nous pouvons y observer que les sons  $rand2$ ,  $rand3$  et  $rand4$  ne permettent pas à un-e piéton-ne d'entendre le VS de suffisamment loin. Une personne atteinte d'une déficience visuelle risque par exemple de s'engager sur la chaussée trop tard pour éviter que le-a conducteur-ric-e aie le temps de freiner et d'éviter une collision. Les sons  $IGA_{opt}^i$  et  $design1$  sont détectés avant la limite, tandis que le son  $design2$  en est très proche.

### Analyse bi-objectifs

Les représentations unidimensionnelles des figures 2.15, 2.16 et 2.17 ne permettent pas de rendre compte des relations de dominance, au sens Pareto (voir section 2.2.2), entre les sons. Pour cela, une représentation où figurent tous les objectifs du problème est nécessaire. Ainsi, la figure 2.18 montre simultanément les valeurs moyennes de désagrément (axe des abscisses, correspondant à la figure 2.15) et de temps de temps de réaction (axes des ordonnées, correspondant à la figure 2.16) pour chaque son, avec les erreurs standards associées.

Les résultats montrent que les sons  $IGA_{opt}^i$  constituent en moyenne un très bon compromis entre désagrément et temps de réaction. En effet, bien qu'ils ne dominent pas tous les autres sons au sens de Pareto, ils sont tout de même situés sur le front de Pareto, avec les sons  $rand1$ ,  $rand2$  et  $design1$ . Parmi ces derniers, le son  $rand2$  est peu désagréable, mais au prix d'un temps de réaction élevé, au dessus de la limite de sécurité  $R_{t,s}$  décrite précédemment. Au contraire, le son  $rand1$  est très désagréable, mais avec un temps de réaction très court. Les sons  $IGA_{opt}^i$  dominant en moyenne le son  $design2$  au sens de Pareto, mais sont de même rang que le son  $design1$ . Les sons  $IGA_{opt}^i$  et  $design1$  constituent donc des compromis intéressants entre désagrément et temps de réaction. À l'inverse, les sons *Random* sont soit dominés par les sons  $IGA_{opt}^i$ , soit situés aux extrêmes du front de



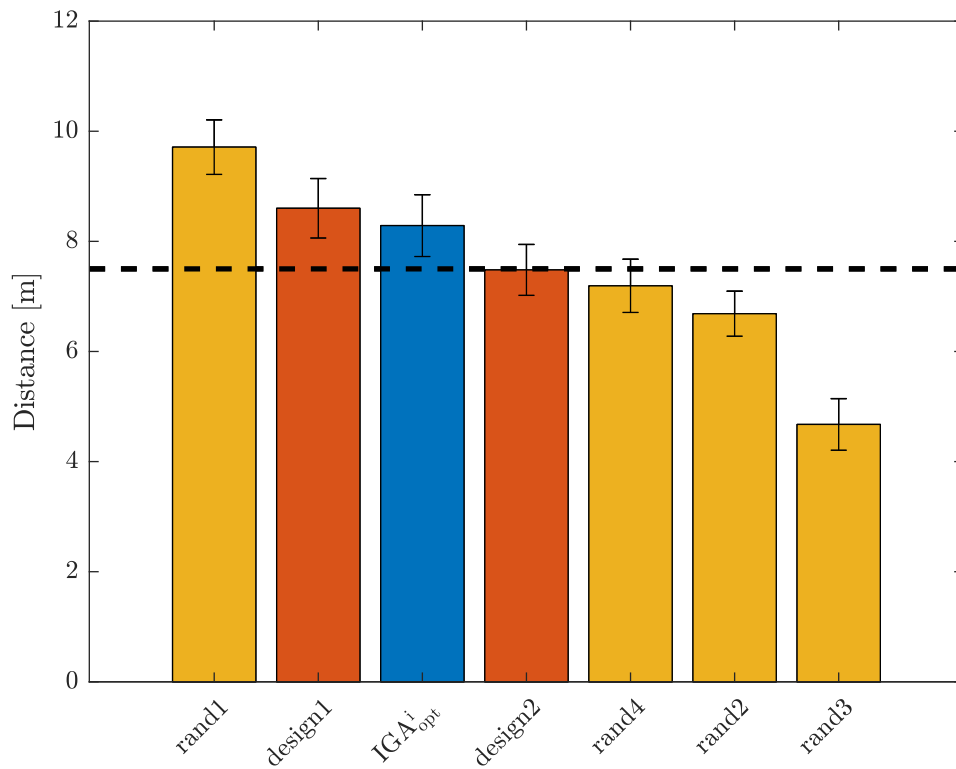


FIGURE 2.17 – Distance moyenne au moment de la détection, pour les sons évalués lors de l'expérience 2. Les barres d'erreur correspondent à l'erreur standard. La ligne pointillée horizontale représente la distance de sécurité  $D_{d,s}$ . Le son en bleu est  $IGA_{opt}^i$ , les sons en orange sont les sons *Design* et ceux en jaune sont les sons *Random*.

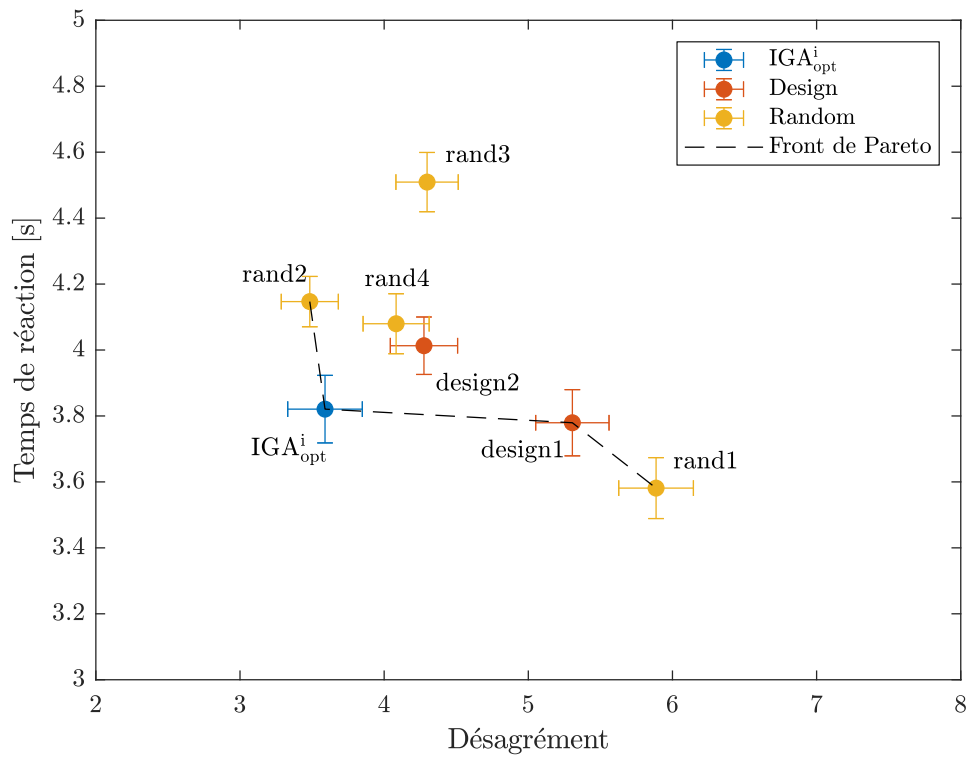


FIGURE 2.18 – Performances moyennes des sons pour les deux objectifs, pour l'expérience 2. Les barres d'erreur correspondent à l'erreur standard. La ligne pointillée représente le front de Pareto.

TABLE 2.4 – Nombre de sujets pour lesquels chaque son fait partie du front de Pareto.

Son	Nombre de sujets
<i>rand1</i>	16
<i>rand2</i>	12
<i>rand3</i>	1
<i>rand4</i>	13
<i>design1</i>	13
<i>design2</i>	12
$IGA_{opt}^i$	17

Pareto et constituent donc des compromis peu intéressants.

Contrairement aux sons des catégories *Design* et *Random*, le son  $IGA_{opt}^i$  est différent pour chaque sujet  $i$ . Le point  $IGA_{opt}^i$  sur la figure 2.18 ne correspond donc pas à un unique son, mais représente les évaluations moyennes fournies pour l'ensemble des sons  $IGA_{opt}^i$ . C'est pourquoi nous procédons également à une analyse sujet par sujet, consistant à identifier pour chacun d'entre eux les sons qui appartiennent au front de Pareto. Pour chaque sujet, nous calculons ainsi les valeurs moyennes des évaluations de désagrément et de temps de réaction de chaque son, puis nous listons les sons appartenant au front de Pareto. Le tableau 2.4 présente le nombre de sujets pour lesquels chaque son fait partie du front de Pareto.

Nous observons que les sons  $IGA_{opt}^i$  sont les plus représentés, suivis par le son *rand1*. Cela conforte les résultats de la figure 2.18 concernant la qualité des sons  $IGA_{opt}^i$ . Cependant, cette analyse ne permet pas de distinguer les sons situés aux extrêmes du front de Pareto de ceux situés dans sa zone centrale. En effet, le son *rand1* est presque autant présent sur les front de Pareto des sujets que les sons  $IGA_{opt}^i$ , alors qu'il constitue en moyenne un moins bon compromis, étant rapidement détectable, mais très désagréable. Parmi les sons restants, tous apparaissent un nombre de fois similaire, hormis le son *rand3*. Ce dernier apparaît dans le front de Pareto d'un seul sujet, ce qui est cohérent avec ses très mauvaises performances moyennes (figure 2.18).

### 2.3.5 Discussion

Ces résultats vont dans le sens des résultats de la première expérience et de travaux précédents (PETIOT, LEGEAY et LAGRANGE 2019) et montrent l'intérêt de l'utilisation d'un IGA pour la minimisation simultanée de plusieurs attributs perceptifs d'un son. En faisant évaluer par chaque sujet  $i$  le meilleur son  $IGA_{opt}^i$  trouvé avec l'IGA et sélectionné

avec la méthode TOPSIS, nous étudions la capacité de cet outil à produire des solutions satisfaisantes d'un point de vue individuel. Les résultats de l'expérience 2 montrent que chaque son  $IGA_{opt}^i$  constitue un bon compromis pour le problème de design, selon le sujet  $i$  correspondant.

Aucune étude, à notre connaissance, ne procède à une comparaison similaire. Dans la littérature, les solutions trouvées à l'aide d'un IGA sont généralement comparées avec les solutions trouvées via d'autres méthodes d'optimisation, ou des variantes d'un même algorithme. Par exemple, FUKUMOTO, YAMAMOTO et OGAWA (2013) comparent les sons trouvés avec un IGA avec des sons trouvés via un autre algorithme, dit à évolution différentielle (DE). Bien que ce second algorithme permette de trouver de meilleures solutions que l'IGA vis-à-vis de la fonction fitness utilisée, le nombre d'évaluations par génération et, de fait, la durée de l'expérience, sont deux fois plus élevés. Les comparaisons avec des méthodes empiriques comme dans (YEE-KING 2016) sont, comme mentionné dans la section 2.2.4, d'ordre qualitatif et se concentrent plutôt sur l'expérience d'utilisation de l'outil.

Par ailleurs, nous n'avons pas pris en compte de contrainte technique ou légale dans notre approche du design de son de VS, telles que celles imposées dans les réglementations internationales (EU REGULATION 2014 ; UN REGULATION 2017). Celles-ci pourraient facilement être intégrées à la méthode proposée, soit comme contraintes au problème d'optimisation, ou bien directement dans le choix des variables de conception et des modalités correspondantes.

## 2.4 Conclusion du second chapitre

Dans ce chapitre, nous avons étudié l'intérêt de l'utilisation d'un IGA pour l'optimisation interactive de sons, afin de savoir si cette famille d'algorithme pouvait être utilisée dans le protocole expérimental proposé au chapitre 1 (voir section 1.2.3). Pour cela, nous avons étudié le problème du design de sons externes de véhicules silencieux (VS). Il s'agit d'un problème bi-objectif complexe, du fait de l'antagonisme de ses deux objectifs (désagrément et temps de réaction) et des possibles interactions entre les variables de conception. Une approche d'optimisation interactive semblait donc intéressante, car celle-ci permet de chercher des solutions au problème sans spécifier en avance de modèle des attributs perceptifs à minimiser en fonction des variables de conception. Nous avons ainsi poursuivi une première étude effectuée au LS2N (PETIOT, LEGEAY et LAGRANGE

2019) qui semblait montrer l'intérêt des IGAs pour ce problème.

Au cours d'une première expérience, un groupe de sujets a participé à l'optimisation interactive de sons de VS, au travers d'évaluations perceptives guidées par un IGA. Cette première expérience a montré une amélioration de la qualité des solutions au cours du processus d'optimisation. Nous avons ensuite conduit une seconde expérience avec le même groupe de sujets, afin de comparer les sons trouvés individuellement par les sujets grâce à l'IGA et la méthode TOPSIS (sons  $IGA_{opt}^i$ ) avec des sons conçus empiriquement (catégorie *Design*) et des sons choisis aléatoirement (catégorie *Random*). Cette deuxième expérience a montré que, du point de vue chaque sujet  $i$ , le son  $IGA_{opt}^i$  a de bonnes performances comparé aux sons des catégories *Design* et *Random*.

Ces deux premières contributions expérimentales nous permettent donc de donner une réponse positive à la première question de recherche :

#### Première question de recherche

Un Algorithme Génétique Interactif (IGA) peut-il permettre de minimiser simultanément plusieurs attributs perceptifs d'un son ?

#### Points clés à retenir

- Une première expérience a montré une amélioration des sons de VS au cours du processus d'optimisation interactive, au regard de deux objectifs antagonistes que sont le désagrément et le temps de réaction.
- Une seconde expérience a permis de montrer que les sons trouvés avec l'IGA obtiennent de très bonnes performances, d'un point de vue individuel, en comparaison de sons choisis empiriquement ou aléatoirement.

# RECOMMANDATIONS DE DESIGN À PARTIR DES RÉSULTATS D'UNE ÉPREUVE IGA

---

Les résultats des expériences 1 et 2, présentés dans le chapitre 2, nous ont montré qu'un Algorithme Génétique Interactif (IGA) peut permettre à un·e utilisateur·rice de trouver des solutions pertinentes à un problème d'optimisation perceptive interactive, d'un point de vue individuel. À l'issue de l'expérience 1, nous obtenons pour chaque sujet  $i$  un ensemble de sons que nous avons appelé *Optimset<sup>i</sup>*, constitué des meilleurs sons évalués au cours du processus d'optimisation. A ce stade, nous n'avons donc pas d'information sur les niveaux des variables de conception susceptibles de correspondre à des sons qui seraient considérés, en moyenne, comme de bonnes solutions au problème. Pour pouvoir produire des recommandations plus générales en termes de design, une analyse des sons obtenus par les sujets lors de l'expérience 1 est donc nécessaire (FUKUMOTO et HIJIKURO 2020). C'est le sujet de la seconde question de recherche de cette thèse, qui traite de la troisième étape du protocole expérimental proposé dans ce manuscrit (voir section 1.2.3) :

## Seconde question de recherche

Comment passer des résultats individuels obtenus avec une méthode d'optimisation interactive à des recommandations générales de design ?

Une première partie porte sur l'état de l'art concernant l'utilisation de résultats issus d'expériences utilisant des IGAs, pour informer un processus de design. Ensuite, une méthode d'analyse et de recommandation de sons est proposée et appliquée au problème de design de sons d'alerte de VS abordé dans le chapitre 2. Enfin, une troisième expérience est présentée, au cours de laquelle les sons recommandés sont évalués par un nouveau panel

de sujets, nous permettant ainsi d'évaluer perceptivement les performances de la méthode d'analyse. Les travaux présentés dans ce chapitre ont fait l'objet de deux publications (SOUAILLE, PETIOT, LAGRANGE et al. 2021 ; SOUAILLE, PETIOT, MISDARIIS et al. 2022).

## 3.1 État de l'art

### 3.1.1 Généralisation de résultats IGA individuels

Dans (WAKEFIELD et al. 2005), les auteur·rice·s soulignent que même si les solutions obtenues avec un IGA répondent très bien au problème d'un point de vue individuel, il reste difficile d'exploiter les données récoltées lors d'une telle expérience pour produire des recommandations plus générales. Peu d'études abordent ce sujet, la plupart des applications des algorithmes évolutionnistes interactifs visant soit à assister des personnes dans un processus créatif, soit à permettre la personnalisation d'un design par rapport à des préférences individuelles. Par ailleurs, certains algorithmes d'optimisation interactive intègrent directement les évaluations de plusieurs sujets pendant le processus d'optimisation. C'est par exemple ce qui est fait dans MATA et al. (2018), qui porte sur le design de volants de voiture grâce à un IGA multi-objectifs. Les fonction-objectifs sont cinq affordances (NORMAN DONALD 2013), c'est-à-dire les caractéristiques d'un objet qui suggèrent son mode d'usage à l'utilisateur·rice. Une particularité de l'IGA utilisé ici est que tous les sujets participent aux évaluations de volants de voiture générés par un IGA unique. Celui-ci fonctionne depuis un serveur et attend, à chaque génération, que tous les volants aient été évalués pour créer la génération suivante. Ainsi, les solutions obtenues à la fin du processus d'optimisation ont mécaniquement intégré les contributions individuelles d'un groupe de sujet. Des approches similaires ont été proposées pour l'optimisation interactive et collaborative de mélodies (NOMURA et FUKUMOTO 2017) ou encore de plans d'architecte et d'interfaces graphiques (BANERJEE, QUIROZ et LOUIS 2008).

Cependant, il peut être intéressant d'extraire des informations qui vont au-delà de la simple connaissance des solutions répondant le mieux au problème d'optimisation. Pour cela, une première approche consiste à créer un modèle de la quantité à minimiser (ou maximiser) en fonction des variables de conception, c'est d'ailleurs qui est fait dans MATA et al. (2018). L'ensemble des évaluations effectuées au cours du processus d'optimisation est utilisé pour construire des modèles entre chaque variable de conception, qui définissent la forme du volant, et les affordances évaluées. Ces modèles permettent de prédire, pour

chaque affordance, la probabilité que chaque variable de conception ait un impact positif sur celle-ci, en fonction de sa valeur. Comme les variables de conception sont continues, il est possible de déterminer des plages de valeurs pour lesquelles elles ont un impact positif sur les affordances. Cependant, les solutions explorées étant définies par l'algorithme au fur et à mesure du processus, la qualité et la complexité du modèle construit est moindre qu'en définissant un plan d'expériences adapté en amont. En effet, les observations sur lesquelles le modèle est ajusté ne possèdent pas nécessairement les propriétés requises d'un plan d'expériences (orthogonalité, conditionnement correct de la matrice d'expérience). Un tel modèle peut également être construit durant le processus d'optimisation, afin de faciliter ce dernier (TAKAGI 2001 ; PEI et TAKAGI 2018). Cette approche est très largement utilisée pour tenter de répondre au problème de la fatigue des utilisateur·rice·s, mais nous ne connaissons pas d'étude qui propose d'utiliser le modèle a posteriori pour étudier des aspects liés à la perception humaine.

Un second type d'approche se concentre sur l'analyse de l'ensemble des meilleures solutions trouvées individuellement par un groupe de personnes ayant utilisé un algorithme d'optimisation interactive. Dans ce cas, il s'agit de faire émerger des propriétés communes à l'ensemble des solutions étudiées, généralement au regard des variables de conception. Dans (POIRSON, PETIOT, BOIVIN et al. 2013) par exemple, les auteur·rice·s s'intéressent au design de verres à pied « élégants » et de planches de bord « dégagées ». Concernant le design des verres, une Classification Ascendante Hiérarchique (CAH) est effectuée à partir des meilleurs designs obtenus par les sujets. Cette classification permet d'établir trois groupes de sujets pour lesquels des verres « élégants » ont des propriétés similaires. Pour chaque groupe de sujets, une analyse du nombre d'occurrences des modalités de chaque variable de conception permet alors de faire émerger un type de verre caractéristique du groupe en question. Pour les planches de bord, une analyse statistique univariée vise à identifier, pour chaque variable, si certaines modalités sont significativement plus représentées que d'autres, parmi les planches de bord les plus « dégagées ». Une analyse similaire est ensuite effectuée sur les paires de variables de conception jugées comme potentiellement en interaction dans le jugement de la forme de la planche de bord évaluée. Ces deux analyses permettent de faire émerger des propriétés de la planche de bord qui contribuent à ce que celle-ci soit perçue comme « dégagée », selon un groupe de sujets. Une CAH est ensuite effectuée pour identifier des groupes de sujets. Une analyse univariée est également effectuée pour chaque groupe, permettant d'identifier les variables de conception pour lesquelles un consensus existe quant à la modalité correspondant à une



planche de bord « dégagée ». Une analyse similaire, basée sur le nombre d’occurrences des modalités de chaque variable est effectuée dans (PETIOT, CERVANTES CHAVEZ, BOIVIN et al. 2015), qui porte sur le design d’un affichage de planche de bord « sportif ». De même, dans (PETIOT, DENJEAN et al. 2020), qui porte sur le design de sons d’alerte, la modalité la plus représentée parmi les meilleurs sons individuels est choisie pour chaque variable de conception, afin de définir le meilleur son commun à un groupe de sujets. Nous ne connaissons pas d’étude appliquant ces approches à un problème d’optimisation multi-objectifs.

### 3.1.2 Intégration des objectifs du problème d’optimisation

Pour le problème d’optimisation bi-objectifs auquel nous nous intéressons, la généralisation des résultats obtenus individuellement avec un IGA est d’autant plus complexe qu’elle doit intégrer les deux attributs perceptifs qui caractérisent les sons. Nous souhaitons en effet pouvoir produire des recommandations de sons qui représentent des compromis intéressants entre désagrément et temps de réaction à l’échelle d’un groupe de sujets. Dans la littérature, peu d’études s’intéressent à l’optimisation de produits selon plusieurs attributs perceptifs. Dans MATA et al. (2018), la construction de modèles reliant les variables de conception et les affordances du volant de voiture à optimiser permet de mettre à jour que certaines affordances sont affectées de manière opposée par certaines variables de conception. La recherche d’un compromis entre les différentes affordances par le·a designer·euse à partir de ces modèles est facilitée par le fait que les variables de conception soient continues.

Par ailleurs, dans leur article de revue de la littérature, XIN et al. (2018) montrent que les méthodes d’optimisation multi-objectifs interactives ne visent bien souvent pas à minimiser (ou maximiser) conjointement plusieurs attributs perceptifs, mais plutôt des fonctions dont la valeur peut être calculée numériquement. Le rôle de l’utilisateur·rice est alors de guider le processus d’optimisation, en informant l’algorithme sur l’importance relative des différentes fonctions-objectif, ou bien encore en fournissant un classement partiel des solutions du front de Pareto trouvées par l’algorithme (voir la taxonomie proposée par MEIGNAN et al. (2015), section 1.2, chapitre 1). Lorsque des attributs perceptifs sont en jeu, ces derniers ne constituent qu’un objectif à minimiser (ou maximiser), conjointement avec un objectif qui peut être calculé numériquement. Dans (BRINTRUP, RAMSDEN et TIWARI 2007 ; BRINTRUP, RAMSDEN, TAKAGI et al. 2008), les auteur·rice·s appellent ces deux types d’objectifs, objectif qualitatif et objectif quantitatif, respectivement. Leurs

études, qui portent sur l'agencement d'espace et le design de chaises ergonomiques, font appel à un algorithme génétique multi-objectifs pour minimiser conjointement un objectif quantitatif avec un objectif qualitatif. Plusieurs notions subjectives doivent être évaluées par les sujets, mais leur jugement global est synthétisé par un score unique. Dans (TSENG, CAGAN et KOTOVSKY 2012), les auteurs font évaluer la forme de représentations schématiques de profils de voiture selon quatre attributs subjectifs. Ils entraînent ensuite un réseau de neurones artificiels pour chacun des sujets, visant à prédire les quatre attributs évalués précédemment en fonction de la forme de la voiture. Plusieurs problèmes bi-objectifs sont ensuite abordés, à chaque fois visant à minimiser conjointement un jugement subjectif et une quantité physique calculable numériquement. Ici, les auteurs choisissent cependant d'utiliser un algorithme génétique non interactif, en calculant les valeurs des attributs perceptifs grâce aux modèles construits auparavant. Ces approches ne visent pas à extraire une information sur la perception d'un groupe de sujets, mais s'inscrivent plutôt dans l'utilisation répandue des algorithmes génétiques interactifs comme outils d'aide au design pour un·e expert·e, abordée au chapitre 2.

Il n'existe donc pas, à notre connaissance, de méthode dans la littérature qui permettrait de généraliser les résultats obtenus individuellement par un groupe de sujets ayant participé à un expérience IGA multi-objectifs. Dans ce chapitre nous mettons en oeuvre les méthodes d'analyse trouvées dans la littérature, puis nous proposons une méthode, mieux adaptée aux spécificités de notre problème de design sonore.

## 3.2 Analyse des données de l'expérience 1

Nous cherchons dans un premier temps à analyser les données de l'expérience 1 au travers des deux approches décrites dans la section 3.1.1, c'est-à-dire la construction de modèles des attributs perceptifs à partir des évaluations des sujets et l'analyse statistique des meilleurs sons trouvés individuellement par ces derniers. Ceci nous permet de mettre à jour les limites de ces méthodes dans le cas de notre problème.

### 3.2.1 Modèles linéaires mixtes de détectabilité et de désagrément

Afin d'étudier l'influence des variables de design sur la détectabilité et le désagrément, un modèle linéaire mixte sans interaction est ajusté sur les données de l'expérience. Le

modèle est un modèle linéaire mixte (comme dans une ANOVA), avec les six variables  $V_i$ , une variable « sujet » et un terme d'erreur aléatoire (MATHEW, SINHA et KHURI 2011). Dans ce modèle mixte, les variables sont considérées avec un effet fixe tandis que le facteur sujet est considéré comme aléatoire.

L'expression du modèle est :

$$y_{ijklmnop} = \mu + V_{1i} + V_{2j} + V_{3k} + V_{4l} + V_{5m} + V_{6n} + S_o + \epsilon_{ijklmnop} \quad (3.1)$$

avec :

- $y_{ijklmnop}$  : l'évaluation du désagrément (ou du temps de réaction) de l'observation  $p$  du son ( $V_{1i}, V_{2j}, V_{3k}, V_{4l}, V_{5m}, V_{6n}$ ) par le sujet  $o$ ,
- $\mu$  : terme constant,
- $V_{1i}$  : coefficient du niveau  $i$  de la variable  $V_1$ , avec  $\sum_i V_{1i} = 0$ . Les autres coefficients correspondent aux variables  $V_2, V_3, V_4, V_5, V_6$ ,
- $S_o$  : coefficient du sujet  $o$ ,
- $\epsilon_{ijklmnop}$  : terme d'erreur,  $\epsilon \sim \mathcal{N}(0, \sigma^2)$ .

Nous nous intéressons au coefficient de détermination  $R^2$  des modèles, ainsi qu'à l'importance de chaque variable et leurs effets, via un test de Fisher. Le pourcentage d'importance  $I_{V_1}$  de la variable  $V_1$  dans le modèle est donné par :

$$I_{V_1}(\%) = \frac{\max_i(V_{1i}) - \min_i(V_{1i})}{\sum_j(\max_k(V_{jk}) - \min_k(V_{jk}))} \quad (3.2)$$

Les expressions sont similaires pour les autres variables.

Un modèle est ajusté pour chaque attribut perceptif : nous obtenons donc un modèle pour le désagrément et un modèle pour le temps de réaction. Une analyse des paramètres du modèle nous permet de comprendre les effets principaux des variables de conception sur chaque attribut perceptif. Le tableau 3.1 donne les résultats de la construction des deux modèles en considérant toutes les évaluations effectuées dans l'expérience 1 (en supprimant les sujets non valides (voir section 2.2.3)). Ces modèles sont construits à partir de l'ensemble des évaluations fournies par les sujets considérés comme valides pendant l'expérience 1, soit  $28 \text{ sujets} \times 9 \text{ sons} \times 11 \text{ générations} = 2772 \text{ observations}$ . Pour chacun des deux modèles, le coefficient de détermination  $R^2$  est peu élevé (43% pour le désagrément et 46% pour la détectabilité). Ces modèles ne permettent donc qu'une analyse des tendances générales dans les données, et ne permettent pas de faire une prédiction fiable des attributs

TABLE 3.1 – Coefficients, p-valeurs des tests de Fisher et importance des variables pour les deux modèles (Désagrément et Temps de réaction).

Variable	Niveau	Désagrément ( $R^2 = 43\%$ )			Temps de réaction ( $R^2 = 46\%$ )		
		Coeff.	Signif.	I(%)	Coeff.	Signif.	I(%)
$V_1$	1	-0.99	p<0.001	21.3%	0.23	p<0.001	16.7%
	2	-0.17			0.13		
	3	0.41			-0.12		
	4	0.74			-0.24		
$V_2$	1	-0.96	p<0.001	27.2%	0.48	p<0.001	34.3%
	2	-0.58			0.19		
	3	0.28			-0.20		
	4	1.25			-0.47		
$V_3$	1	-0.49	p<0.001	15.3%	0.22	p<0.001	18.3%
	2	-0.50			0.15		
	3	0.26			-0.09		
	4	0.74			-0.28		
$V_4$	1	-0.42	p<0.001	13.4%	0.13	p<0.001	10.9%
	2	-0.47			0.10		
	3	0.26			-0.06		
	4	0.62			-0.17		
$V_5$	1	-0.34	p<0.001	6.4%	-0.04	p<0.001	7.1%
	2	0.17			-0.05		
	3	0.17			-0.06		
	4	-0.01			0.14		
$V_6$	1	-0.50	p<0.001	16.3%	0.17	p<0.001	12.6%
	2	-0.43			0.08		
	3	0.10			-0.06		
	4	0.83			-0.18		

perceptifs étudiés. Néanmoins, toutes les variables ont un effet significatif ( $p < 0,001$ ), et ce pour les deux modèles. Nous pouvons donc avoir une information sur leurs contributions respectives dans la variabilité des attributs perceptifs étudiés. Nous observons ainsi que toutes variables ont une importance similaire d'environ  $\frac{1}{6}$ , soit 16.6%, excepté les variables  $V_5$  et  $V_2$ . La variable  $V_5$ , correspondant à la fréquence de modulation d'amplitude, a une influence plus faible avec  $I_{V_5} = 7.1\%$  pour le désagrément et  $I_{V_5} = 6.4\%$  pour le temps de réaction. La variable  $V_2$ , correspondant à la fréquence fondamentale/centrale, a quant à elle une plus grande influence ( $I_{V_2} = 27.2\%$  pour le désagrément et  $I_{V_2} = 34.3\%$  pour la détectabilité).

La figure 3.1 montre le graphique des effets principaux des six variables pour chaque

modèle. Nous pouvons très clairement y observer l'opposition entre le désagrément et la détectabilité. Selon les modèles, les niveaux les plus bas des variables correspondent en moyenne à des sons moins désagréables, mais aussi moins rapidement détectables, et l'inverse pour les niveaux les plus hauts. La variable  $V_5$  a cependant une influence un peu plus complexe. Pour cette variable, le niveau 4 n'est pas le plus désagréable ni le plus rapidement détectable.

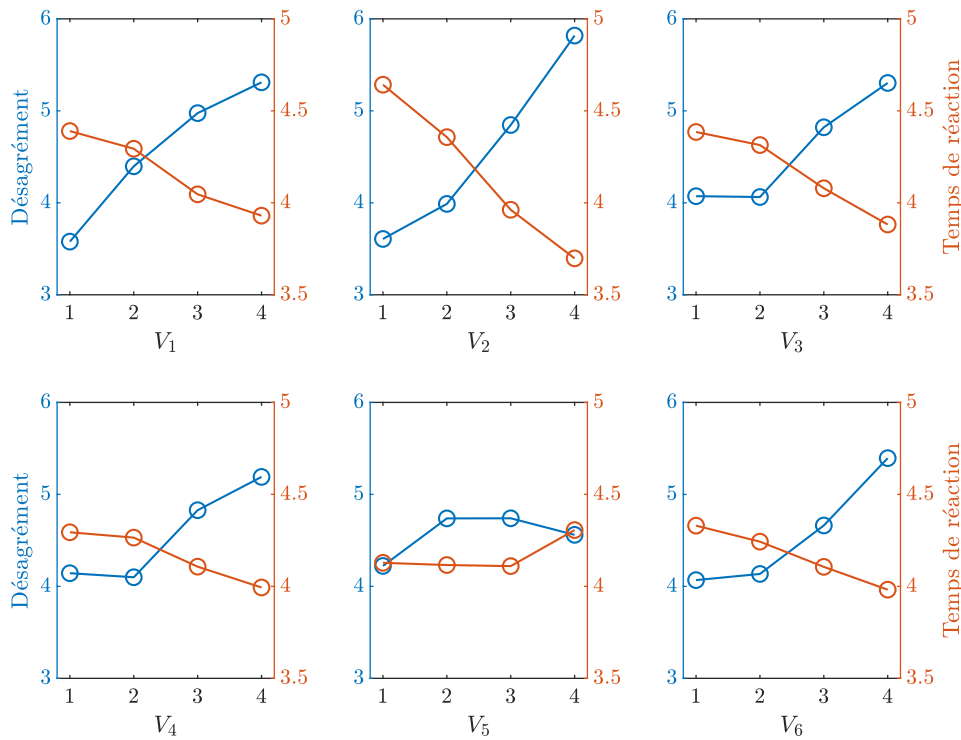


FIGURE 3.1 – Graphes des effets principaux des modèles linéaires mixtes de désagrément (en bleu) et de temps de réaction (en orange).

Un examen individuel des variables montre une cohérence avec de précédentes études (PARIZET, ROBERT et al. 2014 ; LEE et al. 2017 ; PETIOT, KRISTENSEN et MAIER 2013). En moyenne, les propriétés suivantes correspondent à des sons peu désagréables :

- La ressemblance avec un son de moteur à combustion interne ( $V_1 = 1$ )
- Un contenu surtout dans les basse fréquences ( $V_2 = 1$ )
- Un contenu peu tonal ( $V_3 = 1$ )
- Peu de modulation d'amplitude ( $V_6 = 1$ )

En revanche, les propriétés suivantes correspondent plutôt à des sons détectables rapidement :

- Un contenu fréquentiel plutôt dans les hautes fréquences ( $V_2 = 4$ )
- Une forte richesse harmonique ( $V_1 = 4$ )
- De la modulation d'amplitude ( $V_6 = 4$ )

Cependant, l'opposition entre le désagrément et le temps de réaction n'est qu'une tendance générale, révélée par un modèle linéaire qui explique moins de 50% de la variance. Par conséquent, ces modèles ne nous permettent pas de savoir s'il existe des combinaisons de niveaux de variables qui minimisent à la fois le désagrément et le temps de réaction.

### 3.2.2 Analyse statistique de *Optimalset*

L'ensemble *Optimalset* est l'union des ensembles *Optimset<sub>i</sub>*, qui sont, pour chaque sujet *i*, le front de Pareto du registre obtenu à la fin du processus d'optimisation (voir section 2.3.2). Pour chaque participant·e, l'ensemble *Optimset<sub>i</sub>* contient entre un et huit sons. L'ensemble *Optimalset* contient quant à lui  $N = 113$  sons. La distribution du nombre de sons dans les ensembles *Optimset<sub>i</sub>* est indiquée sur la figure 3.2. Chacun de ces sons représente un compromis différent entre désagrément et temps de réaction. Chaque son n'est présent qu'une seule fois dans *Optimalset*, sauf le son  $(1, 4, 4, 4, 2, 1)$ , présent deux fois, et le son  $(2, 4, 4, 4, 1, 1)$ , présent trois fois. Ces sons font partie de plusieurs ensembles *Optimset<sub>i</sub>*, ce qui veut dire qu'ils présentent probablement un intérêt pour répondre au problème d'optimisation. C'est pourquoi lorsqu'un son est présent dans plusieurs ensembles *Optimset<sub>i</sub>*, il est représenté le même nombre de fois dans *Optimalset*.

La table 3.2 montre le nombre d'occurrences des niveaux de chaque variable dans *Optimalset*. Nous observons par exemple que concernant la variable  $V_3$ , le niveau 4 apparaît 50 fois. Afin de savoir si, pour chaque variable, la répartition des niveaux diffère significativement d'une répartition aléatoire, nous conduisons des tests multinomiaux, dont les résultats sont indiqués dans la table 3.2. Seules les répartitions des niveaux des variables  $V_2$  et  $V_3$  sont significativement différentes de répartitions aléatoires ( $p < 0,05$ ). Concernant la fréquence centrale/fondamentale des sons ( $V_2$ ), le niveau 1 (basse fréquence) est sous-représenté. Pour la proportion son harmonique/bruit ( $V_3$ ), le niveau 4 (pas de bruit) est sur-représenté. Si nous considérons l'ensemble des participant·e-s, il semble donc que les sons trop graves et les bruits large bande sont à éviter pour pouvoir minimiser à la fois le désagrément et le temps de réaction. Pour les autres variables, il n'est pas possible de tirer de conclusion à partir de cette approche. En effet, il se peut qu'il y ait des groupes de sujets en désaccord au sujet des niveaux adéquats pour ces variables. Cependant, il ne nous est pas possible d'identifier ces potentiels groupes, car le protocole n'inclut pas

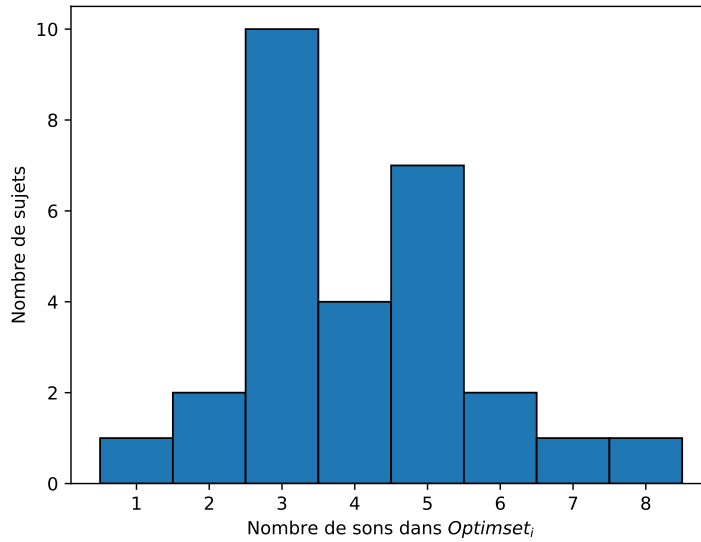


FIGURE 3.2 – Histogramme du nombre de sons dans les ensembles  $Optimset_i$ .

TABLE 3.2 – Nombre d’occurrences des niveaux de chaque variable parmi les sons de  $Optimalset$  et résultats de chaque test multinomial (\* :  $p < 0.05$ ; \*\* :  $p < 0.01$ ; N.S. : non significatif). Pour les variables dont la distribution est significativement différente d’une distribution uniforme, les modalités sur ou sous représentées sont indiquées en gras.

Variable	$V_1$	$V_2$	$V_3$	$V_4$	$V_5$	$V_6$
Niveau 1	35	<b>14</b>	19	27	39	36
Niveau 2	20	36	21	28	23	33
Niveau 3	31	33	23	25	28	24
Niveau 4	27	30	<b>50</b>	33	23	20
Test multinomial	N.S.	*	**	N.S.	N.S.	N.S.

l’évaluation de sons communs à tous les sujets. Avec l’IGA, chaque sujet évalue un groupe de sons qui lui est propre. Il se peut également que les variables de conception aient peu d’influence individuellement sur la probabilité d’appartenance d’un son à  $Optimalset$ , et qu’il faille étudier leurs interactions.

### 3.3 Méthode de recommandation

Les méthodes décrites ci-avant ne permettent pas, dans notre cas, d’extraire suffisamment d’informations pour pouvoir proposer des recommandations à un.e designer.euse sonore. Dans cette section, nous proposons donc une méthode d’analyse de la constitu-

tion de l'ensemble *Optimalset*, afin d'en extraire des recommandations.

### 3.3.1 Description de la méthode

La méthode proposée vise à identifier les combinaisons de niveaux de variables ayant la plus forte probabilité de présence dans *Optimalset*. Pour cela, nous considérons le processus de construction de l'ensemble *Optimalset* au cours de l'expérience IGA comme un processus stochastique dépendant d'une distribution de probabilités discrète. L'ensemble *Optimalset*, de taille  $N$ , est un sous-ensemble de l'univers  $\Omega$ , constitué de l'ensemble des solutions de l'espace de design (ce qui correspond à un plan factoriel complet). À partir des sons appartenant à *Optimalset*, nous pouvons estimer les paramètres de la distribution de probabilité sous-jacente. Grâce à ces paramètres, il est ensuite possible d'inférer la probabilité d'appartenance à *Optimalset* pour n'importe quel son de  $\Omega$ .

#### Loi conjointe

A partir de l'univers  $\Omega$  et des variables  $V_i$  ( $i = 1$  à  $6$ ) qui décrivent un son, une première approche consiste à considérer que la probabilité qu'un son  $d$  appartienne à *Optimalset* dépend de toutes les variables et de leurs interactions :

$$P(D = d) = P(V_1 = v_1, \dots, V_6 = v_6) = \frac{\text{Card}(\{D \in \text{Optimalset} | D = d\})}{\text{Card}(\text{Optimalset})} \quad (3.3)$$

$\text{Card}(\text{Optimalset})$  désigne le cardinal de l'ensemble *Optimalset*, c'est-à-dire le nombre de sons différents qu'il contient. Par exemple, si un son est présent une fois dans *Optimalset*, la probabilité correspondante sera  $P = \frac{1}{\text{Card}(\text{Optimalset})}$ . Si un son n'appartient pas à *Optimalset*, la probabilité correspondante vaut  $P = 0$ . Cependant, cette loi conjointe à six dimensions ne permet pas de recommander d'autres sons que ceux appartenant à *Optimalset*. Pour pouvoir faire des recommandations au niveau des variables  $V_i$ , il est nécessaire de tester leur indépendance mutuelle vis-à-vis du processus de constitution de *Optimalset*, afin de simplifier la loi conjointe.

#### Probabilité marginale

Si nous faisons l'hypothèse que les variables  $V_i$  sont mutuellement indépendantes vis-à-vis du processus de construction de *Optimalset* (autrement dit, qu'il n'y a pas d'interaction entre elles), alors la probabilité de la loi conjointe se simplifie en le produit des proba-



bilités marginales (LEJEUNE 2004). La probabilité que le son  $d = (V_1 = v_1, \dots, V_6 = v_6)$  appartienne à *Optimalset* devient :

$$P(V_1 = v_1, \dots, V_6 = v_6) = \prod_{i=1}^6 P(V_i = v_i) \quad (3.4)$$

Dans ce cas, le son avec la plus grande probabilité d'appartenir à *Optimalset* est constitué des valeurs les plus représentées de chaque variable  $V_i$ . Ainsi, un son ne doit pas nécessairement appartenir à *Optimalset* pour obtenir une probabilité non nulle. Cependant l'indépendance mutuelle de toutes les variables est une hypothèse très forte. Dans la réalité, en particulier en design industriel, il est peu probable qu'il n'y ait pas d'interaction entre les variables  $V_i$ , étant donné que l'évaluation globale d'un design n'est généralement pas égale à la somme des évaluations de chacune des variables qui le constitue (SYLCOTT, MICHALEK et CAGAN 2015).

### Test d'indépendance des variables

Pour pouvoir tout de même simplifier l'expression de la probabilité jointe, et pouvoir faire des inférences pour des designs qui ne font pas partie de l'ensemble *Optimalset*, nous proposons de vérifier, à l'aide d'un test statistique portant sur l'ensemble *Optimalset*, l'indépendance de toutes les paires de variables  $V_i$ . Par exemple, supposons que les tests statistiques montrent une indépendance entre les deux groupes de variables  $\{V_1, V_2, V_3\}$  et  $\{V_4, V_5, V_6\}$ . Grâce à cette information, nous pouvons simplifier l'expression de la loi conjointe de l'équation 3.3 :

$$P(V_1 = v_1, \dots, V_6 = v_6) = P(V_1 = v_1, V_2 = v_2, V_3 = v_3) \cdot P(V_4 = v_4, V_5 = v_5, V_6 = v_6) \quad (3.5)$$

Où :

$$P(V_1 = v_1, V_2 = v_2, V_3 = v_3) = \frac{\text{Card}(\{D \in \text{Optimalset} | V_1 = v_1, V_2 = v_2, V_3 = v_3\})}{\text{Card}(\text{Optimalset})} \quad (3.6)$$

$$P(V_4 = v_4, V_5 = v_5, V_6 = v_6) = \frac{\text{Card}(\{D \in \text{Optimalset} | V_4 = v_4, V_5 = v_5, V_6 = v_6\})}{\text{Card}(\text{Optimalset})} \quad (3.7)$$

TABLE 3.3 – Matrice des valeurs-p des tests du  $\chi^2$  pour chaque paire de variables. Les valeurs-p pour lesquelles l’hypothèse d’indépendance est rejetée sont indiquées en gras. Une correction de Bonferroni est appliqué au seuil de significativité pour prendre en compte les tests des 15 paires de variables simultanément ( $\alpha = 0,05/15 = 0,3\%$ ).

	$V_1$	$V_2$	$V_3$	$V_4$	$V_5$	$V_6$
$V_1$	/	<b>0.002</b>	0.003	0.953	0.264	0.393
$V_2$		/	0.015	<b>0.001</b>	0.052	0.109
$V_3$			/	<b>0.001</b>	0.53	0.466
$V_4$				/	0.219	0.026
$V_5$					/	<b>0.001</b>
$V_6$						/

Il est alors possible de calculer une probabilité pour chaque son  $d = (V_1 = v_1, \dots, V_6 = v_6)$  du plan complet. Si toutes les variables sont dépendantes deux à deux, le calcul est identique à celui de l’équation 3.3.

## Recommandations

Afin de produire des recommandations de sons, nous estimons la probabilité d’appartenir à *Optimalset* pour tous les sons de  $\Omega$ . Les meilleurs sons (ceux à recommander) sont ceux pour lesquelles cette probabilité est la plus élevée. Certains d’entre eux seront présents dans *Optimalset*, mais il est possible que des sons n’en faisant pas partie obtiennent également une probabilité élevée.

### 3.3.2 Mise en œuvre

#### Test d’indépendance des variables

À partir de l’ensemble des sons présents dans *Optimalset*, nous formons des tableaux de contingence pour chaque paire de variables. Pour chaque tableau, nous utilisons un test du  $\chi^2$  d’indépendance pour déterminer s’il existe un lien statistique significatif entre les deux variables correspondantes. Les valeurs-p de chaque test sont indiquées dans la table 3.3. Une correction de Bonferroni est appliquée au seuil de significativité pour prendre en compte les tests des 15 paires de variables simultanément ( $\alpha = 0,05/15 = 0,3\%$ ). Une valeur-p au dessus du seuil pour une paire de variables donnée, signifie que l’hypothèse d’indépendance entre ces deux variables ne peut pas être rejetée.

A partir de cette table, nous construisons le graphe des dépendances entres les va-

riables, présenté en figure 3.3. Ce graphe met en évidence deux groupes de variables mutuellement indépendants :  $\{V_1, V_2, V_3, V_4\}$  et  $\{V_5, V_6\}$ . Cela suggère qu’il existe des interactions entre les variables au sein de ces deux groupes, vis-à-vis du processus de constitution de l’ensemble *Optimalset*. Nous ne pouvons donc pas faire de recommandations en considérant une variable à la fois. La loi de probabilité empirique devient alors :

$$P(V_1 = v_1, V_2 = v_2, V_3 = v_3, V_4 = v_4, V_5 = v_5, V_6 = v_6) \approx P(V_1 = v_1, V_2 = v_2, V_3 = v_3, V_4 = v_4).P(V_5 = v_5, V_6 = v_6) \quad (3.8)$$

La probabilité  $P(V_1 = v_1, V_2 = v_2, V_3 = v_3, V_4 = v_4)$  est calculée comme le nombre de sons de *Optimalset* pour lesquels  $V_1 = v_1, V_2 = v_2, V_3 = v_3$  et  $V_4 = v_4$ , divisé par le nombre total de sons dans *Optimalset*. De même pour  $P(V_5 = v_5, V_6 = v_6)$ .

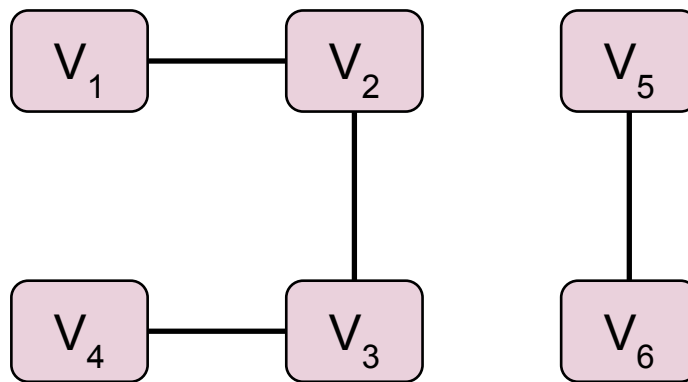


FIGURE 3.3 – Graphe d’indépendance entre les variables de conception. Une ligne noire entre deux variables signifie qu’il existe un lien statistique significatif entre celles-ci.

### Recommandations

La table 3.4 montre les huit sons avec la plus grande probabilité d’appartenance à *Optimalset*, nommés *reco1* à *reco8* (catégorie *Reco*)<sup>1</sup>. Nous considérons que les sons avec les plus grandes probabilités sont des solutions potentiellement pertinentes au problème de design et peuvent être recommandés à un-e designer-euse. Les valeurs des probabilités présentées dans la table 3.4 sont basses ( $< 1\%$ ), car la combinaison de  $\{V_1, V_2, V_3, V_4\}$  la plus représentée dans *Optimalset* est présente 15 fois, tandis que pour  $\{V_5, V_6\}$ , plusieurs

---

1. Les sons *Reco* peuvent être écoutés à l’adresse <https://souaille.github.io/conception/#ch3>

TABLE 3.4 – Liste des huit sons avec la plus grande probabilité estimée d’appartenir à *Optimalset*. La ligne pointillée illustre l’indépendance entre les groupes de variables  $\{V_1, V_2, V_3, V_4\}$  et  $\{V_5, V_6\}$ .

Son	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>	V <sub>6</sub>	Probabilité	Présent dans <i>Optimalset</i> ?
<i>reco1</i>	3	2	2	2	1	1	0.58%	Non
<i>reco2</i>	1	4	4	4	1	1	0.58%	Oui
<i>reco3</i>	2	4	4	4	1	1	0.58%	Oui
<i>reco4</i>	3	2	2	2	3	2	0.5%	Oui
<i>reco5</i>	1	4	4	4	3	2	0.5%	Oui
<i>reco6</i>	2	4	4	4	3	2	0.5%	Non
<i>reco7</i>	1	2	4	2	1	1	0.46%	Non
<i>reco8</i>	4	3	4	2	1	1	0.46%	Non

combinaisons sont présentes 5 fois. La probabilité maximale d’appartenance à *Optimalset* vaut donc  $\frac{15}{113} \cdot \frac{5}{113} = 0,58\%$ . Dans la suite de ce manuscrit, l’indépendance entre les groupes  $\{V_1, V_2, V_3, V_4\}$  et  $\{V_5, V_6\}$  sera visuellement marquée par un trait vertical dans la notation des sons (par exemple, (1, 4, 4, 4|2, 1) pour le son *design1*).

## 3.4 Expérience 3 : Évaluation de la méthode de recommandation

### 3.4.1 Objectifs

Le but de cette expérience est d’évaluer les performances des sons recommandés avec la méthode décrite dans la section précédente auprès d’un nouveau panel de sujets, et de les comparer aux performances d’autres sons. Nous souhaitons ainsi tester la capacité de la méthode d’analyse proposée à généraliser les résultats obtenus par les sujets qui ont utilisé l’IGA pendant l’expérience 1.

### 3.4.2 Matériels et Méthodes

#### Stimuli

Les stimuli sont présentés au travers du même scénario que pour les expériences 1 et 2 (voir section 2.2). 18 sons sont évalués :

- huit sons choisis aléatoirement (*rand1* à *rand8*, catégorie *Random*)<sup>2</sup>. Les sons *rand1* à *rand4* sont les mêmes que pour l’expérience 2 (voir section 2.3).
- les deux sons *design1* et *design2* (catégorie *Design*), également évalués au cours de l’expérience 2.
- huit sons recommandés (*reco1* à *reco8*, catégorie *Reco*), présentés dans la table 3.4. Ce sont les sons pour lesquels la probabilité estimée d’appartenance à l’ensemble *Optimalset* est la plus grande.

## Plan d’expérience

Tous les sujets passent la même condition expérimentale et évaluent l’ensemble des 18 sons. Chaque son est présenté quatre fois, donc chaque sujet procède à  $4 \times 18 = 72$  évaluations en tout. Les sons sont présentés dans un ordre aléatoire, avec un seul son par page.

## Participant·e·s

17 étudiant·e·s de l’École Centrale de Nantes (14 Hommes, 3 Femmes), ne rapportant aucune déficience auditive, sont sollicité·e·s pour participer à cette expérience.

## Apparatus

L’expérience a lieu sur les ordinateurs des salles informatiques de l’École Centrale de Nantes. Les sujets écoutent le scénario avec un casque audio Beyerdynamics DT-990 (casques fermés, d’une impédance de 250 Ohm). La carte son utilisée est la carte son interne de l’ordinateur. La calibration est la même que pour les expériences 1 et 2 (voir section 2.2).

## Protocole

**Interface et procédure expérimentale** L’interface et la procédure de présentation et d’évaluation des sons sont les mêmes que pour les expériences 1 et 2 (voir section 2.2). Comme pour l’expérience 2, un seul son est présenté par page. Les sujets n’ont aucune indication visuelle leur permettant d’identifier quel son leur est présenté. À la fin du

---

<sup>2</sup>. Les sons *rand5* à *rand8* peuvent être écoutés à l’adresse <https://souaille.github.io/conception/#sec34>

TABLE 3.5 – Taux d’erreurs moyens des sujets pour l’expérience 3, avec les écart-types et le nombre de sujets écartés à cause de performances de détection jugées trop faibles (voir section 2.2.2). *TED* : Taux d’erreur de direction. *TDP* : Taux de détection prématurée. *TDT* : Taux de détection tardive.

	Erreur	Moyenne (%)	Écart-Type (%)	Sujets Écartés
<i>TED</i>		3,8	4,1	0
<i>TDP</i>		0,2	0,7	0
<i>TDT</i>		10,0	11,2	0

test, pour chaque sujet  $i$ , nous obtenons 4 évaluations de désagrément (sur une échelle de 0 à 10) et 4 de temps de réaction (en seconde) par son.

### Analyse

Les taux d’erreur décrits dans la section 2.2.2 sont utilisés pour évaluer la fiabilité des sujets vis-à-vis du protocole. De même que pour l’expérience 2, les sons sont comparés selon deux critères :

1. La significativité statistique des différences de désagrément moyen entre chaque paire de sons. De même pour le temps de réaction. Nous utilisons pour cela un test de comparaison multiple de Tukey.
2. Leur appartenance au front de Pareto, au regard des évaluations moyennes de désagrément et de temps de réaction. Cette analyse ne fait appel à aucun test statistique.

### 3.4.3 Résultats

La table 3.5 montre les taux d’erreur moyens pour l’expérience 3. La figure 3.4 montre les distributions correspondantes. Les taux d’erreur moyens sont relativement bas. Le *TED* et le *TDT* moyens sont légèrement supérieurs à ceux observés pendant l’expérience 2 (2,5% et 6,1% respectivement, voir section 2.3.4). Aucun sujet n’a dépassé les seuils fixés, les données de tous les sujets ont donc été utilisées pour la suite de l’analyse.

### Désagrément

Les valeurs moyennes de désagrément de chaque son sont présentées dans la figure 3.5, avec les erreurs standard correspondantes. Nous observons que trois sons recommandés (*reco1*, *reco4* et *reco7*) font partie du groupe de sons les moins désagréables, qui contient

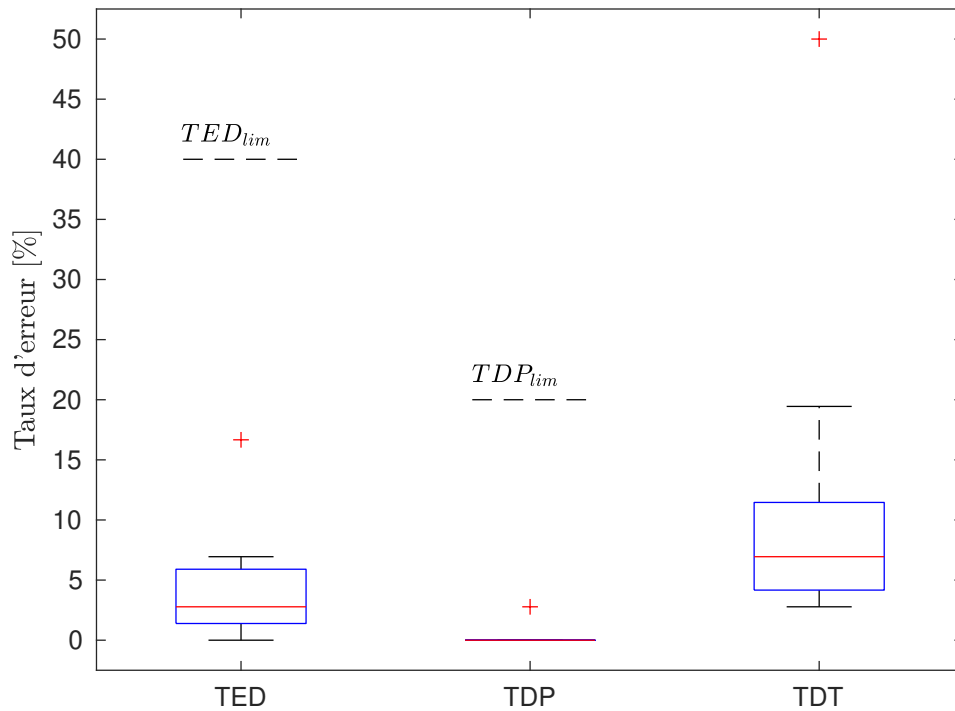


FIGURE 3.4 – Boîtes à moustaches des taux d'erreurs pour l'expérience 3. Les lignes rouges représentent les valeurs médianes des taux d'erreur. Les premiers et troisièmes quartiles sont représentés par les extrémités inférieures et supérieures des boîtes bleues. Les moustaches vont des plus petites aux plus grandes valeurs qui ne sont pas considérées comme des outliers. Les outliers sont représentés par des croix rouges et correspondent aux valeurs situées à plus de 1,5 écarts inter-quartile en-dessous du premier quartile ou au-dessus du troisième quartile. Les lignes noires pointillées horizontales correspondent aux taux d'erreur limites fixés pour que les données du sujet soient considérées comme recevables. La limite pour le TDT n'apparaît pas car elle est supérieure au taux d'erreur maximal représenté sur la figure. *TED* : Taux d'erreur de direction. *TDP* : Taux de détection prématurée. *TDT* : Taux de détection tardive.

également les deux sons *design1* et *design2*. Les sons *Design* sont proches de la valeur médiane en termes de désagrément. Les sons *Random* et *Reco* sont répartis le long de toutes les valeurs possibles de désagrément, mais les sons *Reco* tendent à être plus désagréables que les sons *Random*.

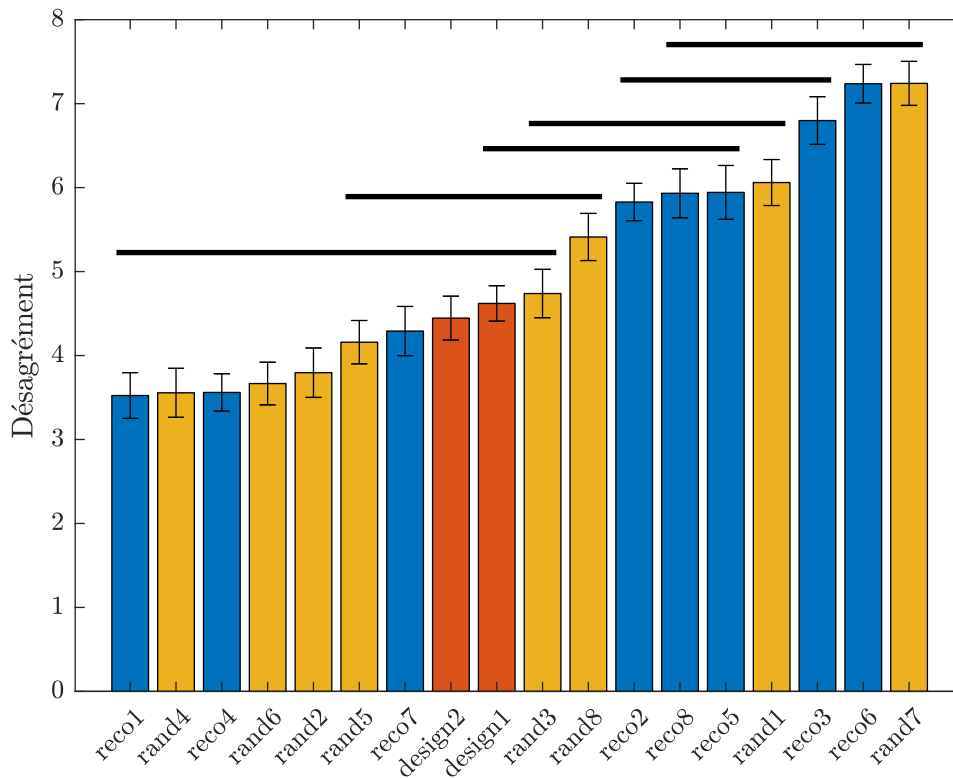


FIGURE 3.5 – Désagrément moyen des sons évalués pendant l’expérience 3, avec les erreurs standards correspondantes et le résultat de l’analyse post-hoc. Les lignes noires horizontales indiquent les groupes de sons au sein desquels les moyennes de désagrément ne sont pas significativement différentes (Test HSD de Tukey,  $p > 0,05$ ). Les sons en bleu sont les sons *Reco*, les sons en orange sont les sons *Design* et ceux en jaune sont les sons *Random*.

### Temps de réaction

La figure 3.6 montre les valeurs moyennes de temps de réaction. Quatre sons *Reco* constituent le groupe des sons les plus rapidement détectables. De même que pour le désagrément, le son *design2* est proche de la valeur médiane, mais le son *design1* se distingue par un temps de réaction plus court. Les sons *Random* correspondent à des temps de réaction élevés, tandis que la majorité des sons *Reco* ont des temps de réaction



plutôt courts. Pour les sons *reco1*, *reco4* et *reco7*, le temps de réaction est en revanche assez élevé.

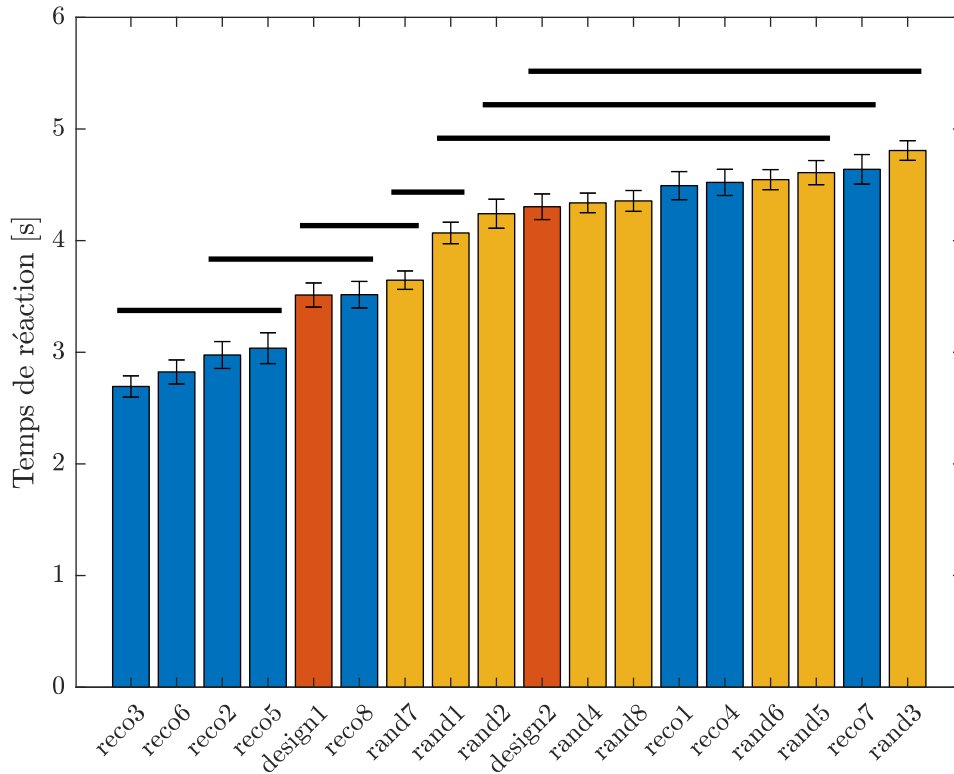


FIGURE 3.6 – Temps de réaction moyen des sons évalués pendant l’expérience 3, avec les erreurs standards correspondantes et le résultat de l’analyse post-hoc. Les lignes noires horizontales indiquent les groupes de sons au sein desquels les moyennes de temps de réaction ne sont pas significativement différentes (Test HSD de Tukey,  $p > 0,05$ ). Les sons en bleu sont les sons *Reco*, les sons en orange sont les sons *Design* et ceux en jaune sont les sons *Random*.

### Analyse bi-objectifs

La figure 3.7 montre les valeurs moyennes de désagrément et de temps de réaction de chaque son. Nous pouvons tout d’abord observer que la position relative des sons *design1* et *design2* est la même que lors de l’expérience 2 (voir figure 2.18, page 65). En effet, le son *design1* est plus désagréable, mais détectable plus rapidement que le son *design2*. De plus, le sons *design1* appartient au front de Pareto dans les deux expériences. La position relative des sons *rand1* à *rand4* est également similaire pour les deux expériences. Ces résultats montrent une cohérence entre les jugements effectués par les deux panels de

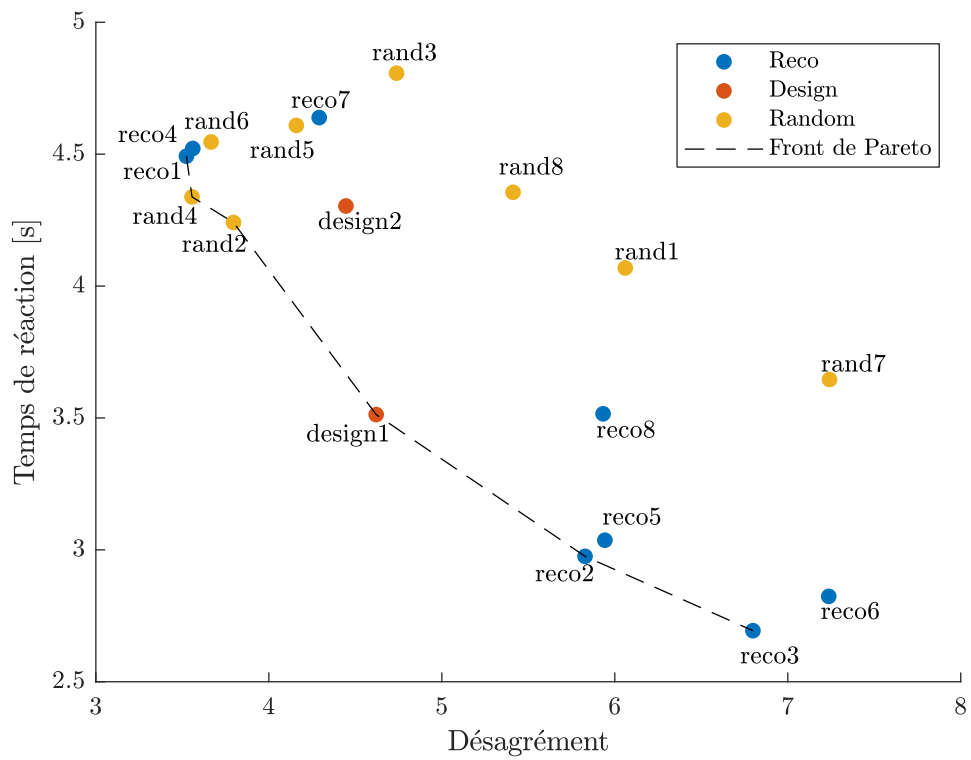


FIGURE 3.7 – Performances moyennes des sons pour les deux objectifs, pour l'expérience 3. Les barres d'erreur correspondent à l'erreur standard. La ligne pointillée représente le front de Pareto.

participant·e·s et sont le signe de la fiabilité du protocole expérimental et des panels utilisés.

Ensuite, nous pouvons voir que le front de Pareto est constitué de trois sons *Reco* (*reco1*, *reco2* et *reco3*), un son *Design* (*design1*) et deux sons *Random* (*rand2* et *rand4*). Les sons (*reco1*, *reco2* et *reco3*) sont ceux pour lesquels la probabilité estimée d’appartenir à *Optimalset* est la plus grande. De plus, les trois sons *Reco* suivants au regard de cette même probabilité (*reco4*, *reco5* et *reco6*) sont proches du front de Pareto.

Par ailleurs, les sons *Random* couvrent une grande plage de valeur pour le désagrément, mais ont dans l’ensemble un temps de réaction élevé. Cela semble indiquer que des sons rapidement détectables ont peu de chance d’être trouvés en choisissant des sons aléatoirement dans le plan complet. Au contraire, les sons *Reco* couvrent de grandes plages de valeurs pour les deux attributs perceptifs. Les recommandations sont donc variées, ce qui est un atout en termes de design. De plus, les sons les plus détectables sont des sons recommandés (*reco2*, *reco5*, *reco3*, *reco6*). Ni les sons *Design*, ni les sons *Random* ne donnent des temps de réaction aussi courts, ce qui suggère que la méthode de recommandation proposée a permis d’extraire certaines propriétés favorisant la détection.

Nous notons enfin que les sons *Reco* tendent à être situés aux extrêmes du front de Pareto. La méthode de recommandation n’a pas réussi à trouver des sons situés dans la région centrale de ce dernier, caractérisée par la présence du son *design1*, qui constitue quant à lui un compromis intéressant.

### 3.4.4 Discussion

Grâce à la méthode d’analyse proposée, nous avons pu faire des recommandations générales de design, au niveau des variables de conception. Ces recommandations prennent la forme d’un score attribué à chaque son du plan complet, correspondant à la probabilité qu’a ce son d’appartenir à l’ensemble *Optimalset*. Cette expérience montre que les sons avec un score élevé sont des solutions pertinentes au problème d’optimisation. Les sons recommandés peuvent servir de points de départ à un·e designer·euse, qui peut ensuite affiner le design pour y intégrer d’autres notions, comme par exemple l’image que le constructeur de VS souhaite véhiculer. La diversité des sons recommandés constitue un avantage pour le·a designer·euse, qui peut alors explorer plusieurs pistes de design. De plus, l’évaluation par un nouveau panel montre qu’il est possible de généraliser les résultats obtenus avec un IGA par un groupe de sujets donné à une population plus large. N’importe quelles données obtenues par un groupe de sujets avec une méthode d’optimisation multi-

objectifs interactive peut être analysée avec la méthode proposée, qui n’est pas limitée au design sonore. Il serait intéressant de faire une analyse plus fine, qui permettrait de mettre à jour des caractéristiques de sons représentatifs de différents compromis entre désagrément et temps de réaction, mais ce n’est pas possible à cause de la complexité du problème bi-objectifs traité ici.

La présence du son *design1* nous indique qu’il devrait être possible de recommander plus de sons situés dans la zone centrale du front de Pareto (figure 3.7). La moins grande densité de sons recommandés dans cette zone peut être due à la méthode de recommandation, mais également à l’algorithme utilisé pendant le processus d’optimisation (sa structure et/ou sa paramétrisation), ou encore à un désaccord entre les sujets pendant l’expérience 1. En effet, l’efficacité de la méthode de recommandation dépend du nombre d’occurrences de sons similaires (du point de vue des variables de conception) dans *Optimalset*. Pendant l’expérience 1, une forte variabilité inter-sujets quant à la caractérisation des sons situés dans la zone centrale du front de Pareto rendrait donc difficile l’extraction de propriétés communes. Cependant, comme mentionné dans la section 3.2.2, il ne nous est pas possible d’identifier de tels groupes, car les sujets n’ont pas évalué les mêmes sons. Il est également possible que le processus d’optimisation tende à trouver plus de sons situés aux extrêmes du front de Pareto. D’autres études sont nécessaires pour confirmer ou infirmer ces hypothèses. Par exemple, l’accord inter-sujets dans les évaluations des attributs perceptifs pourrait être testé en incluant des sons qui seraient évalués par tous les sujets au cours de la procédure IGA. Ces sons n’auraient pas d’influence sur le processus d’optimisation, mais permettraient de comparer les sujets entre eux et éventuellement d’identifier s’il existe des groupes de sujets homogènes dans leurs évaluations des sons. Des simulations numériques avec des modèles de sujets virtuels pourraient permettre d’étudier l’effet de certains facteurs, comme le bruit dans les évaluations faites par un même sujet ou l’accord inter-sujet dans les évaluations des sons, sur le processus d’optimisation et les recommandations faites avec la méthode d’analyse proposée. Ainsi, il serait possible de savoir si des améliorations pourraient être apportées à l’IGA ou à la méthode d’analyse.

La méthode d’analyse que nous proposons permet d’intégrer les résultats IGAs individuels d’un groupe de sujet, en prenant en compte les interactions entre les variables de conception. Elle permet ainsi de faire émerger des informations qui n’étaient pas accessibles avec les méthodes statistiques habituellement utilisées (POIRSON, PETIOT, BOIVIN et al. 2013; PETIOT, CERVANTES CHAVEZ, BOIVIN et al. 2015; PETIOT, DENJEAN et

al. 2020). Il est intéressant noter que l'approche, notamment proposée dans MATA et al. (2018), consistant à utiliser un seul IGA qui prend en compte l'ensemble des évaluations des sujets permet de s'affranchir de la nécessité d'une méthode d'analyse pour trouver des solutions pertinentes à l'échelle d'un groupe. Cette approche pourrait être implémentée dans notre cas. Cependant, dans le problème bi-objectifs auquel nous nous intéressons, la méthode d'analyse sert aussi à identifier les propriétés des sons qui en font de bons compromis entre désagrément et temps de réaction. Dans leur article, MATA et al. (2018) ne proposent pas une telle analyse, mais plutôt une étude des liens qui existent entre chaque variable de conception et chaque affordance évaluée. Cette analyse est similaire à la construction du modèle linéaire effectué dans la section 3.2.1. Cependant, nous avons vu que dans notre cas, toutes les variables de conception (à l'exception de  $V_5$ ) ont un effet opposé sur les deux objectifs du problème. Une analyse supplémentaire serait tout de même nécessaire après avoir utilisé un IGA similaire à celui utilisé par MATA et al. (2018).

### 3.5 Conclusion du troisième chapitre

Au chapitre 2, nous avons montré l'intérêt des IGAs pour la minimisation de plusieurs attributs perceptifs d'un son extérieur de VS. Cependant, afin de produire une information utile pour un·e designer·euse sonore, une analyse des résultats obtenus individuellement pendant l'expérience 1 (voir section 2.2) est nécessaire. Dans ce chapitre, nous avons ainsi cherché à répondre à la deuxième question de recherche :

#### Deuxième question de recherche

Comment passer des résultats individuels obtenus avec une méthode d'optimisation interactive à des recommandations générales de design ?

Après avoir appliqué les méthodes habituellement trouvées dans la littérature et montré leurs limites, nous avons proposé une nouvelle approche, plus adaptée à notre problème. Celle-ci nous a permis de produire des recommandations satisfaisantes et variées, que nous avons validées perceptivement au moyen d'une troisième expérience. Cependant, beaucoup de sons recommandés restent situés proche des extrêmes du front de Pareto des sons évalués pendant l'expérience 3. L'inclusion de sons communs à tous les sujets dans

le processus IGA pourrait permettre de savoir si ces résultats sont dûs à la présence de groupes de sujets discordants dans leurs évaluations de désagrément et de temps de réaction. Par ailleurs, la conduite de simulations d'expériences IGA pourrait permettre de savoir si ces résultats sont dûs à l'algorithme utilisé (ou sa configuration), ou à la méthode de recommandation elle-même. Enfin, bien que des méthodes IGA intégrant directement les évaluations de tout un groupe de sujets dans le processus d'optimisation permettent d'obtenir des solutions satisfaisantes à l'échelle de ce groupe, une analyse telle que celle proposée dans ce chapitre serait tout de même nécessaire pour produire des recommandations qui prennent en compte l'antagonisme des deux attributs perceptifs à minimiser dans notre cas.

#### Points clés à retenir

- La minimisation conjointe de plusieurs attributs perceptifs est une tâche complexe, peu traitée dans la littérature. L'utilisation d'IGAs pour étudier les attributs perceptifs à l'échelle d'un groupe de sujets se limite la plupart du temps à des problèmes mono-objectif.
- Nous avons proposé une méthode d'analyse des données d'individus ayant participé à l'expérience 1, grâce à laquelle nous avons fait des recommandations de sons pour répondre au problème d'optimisation.
- Une expérience de validation (Expérience 3) nous a montré que les sons recommandés sont variés et obtiennent de bonnes performances vis-à-vis des attributs étudiés, comparés à des sons générés aléatoirement et des sons choisis empiriquement.



# RÉDUCTION DE L'ESPACE DE CONCEPTION (PARTIE 1 : SIMULATIONS)

---

Les chapitres précédents ont permis de montrer l'intérêt d'utiliser un IGA dans un protocole expérimental d'étude de la perception sonore (voir chapitre 1, section 1.2.3) et de proposer une méthode d'analyse et de généralisation des résultats obtenus individuellement lors d'expériences d'optimisation interactive (chapitre 3). Nous nous intéressons à présent à la définition des variables de conception à manipuler et des niveaux pris par celles-ci, après qu'une méthode de synthèse sonore ait été choisie. Celle-ci, est importante car elle conditionne les informations qui pourront être récoltées au cours de l'expérience, ainsi que les recommandations de design sonore qui pourront être faites. Cependant, ce choix est souvent fait sur la base de l'expertise individuelle des personnes qui mettent en place l'expérience. Le protocole expérimental proposé (voir section 1.2.3, chapitre 1) bénéficierait ainsi de l'intégration d'une méthode de réduction de l'espace de conception informée et générique, permettant ensuite d'explorer des solutions qui soient les plus pertinentes possible pour le problème traité. Pour cela, nous nous intéressons à la troisième question de recherche :

### Troisième question de recherche

Comment réduire l'espace de conception de manière informée, dans le cadre d'une expérience d'optimisation perceptive interactive ?

Ne sachant pas quelle méthode est la mieux appropriée dans le cas de notre protocole expérimental, il nous est nécessaire d'en tester plusieurs. Afin de limiter le coût expérimental d'une telle étude, nous choisissons de passer par des simulations d'expériences d'optimisation perceptive interactive. Cette approche nous permet d'identifier les méthodes de réduction qui pourraient être utilisées en amont d'une expérience d'optimisation perceptive réelle.



Ce chapitre débute par un état de l'art concernant les stratégies proposées pour limiter la fatigue des utilisateur·rice·s lors d'une expérience IGA, puis les outils utilisés dans la littérature pour réduire l'espace de conception préalablement à une expérience d'optimisation. Une seconde partie décrit l'approche qui a été explorée dans le contexte de cette thèse. Enfin, une troisième partie porte sur une étude expérimentale de l'approche proposée au travers de simulations numériques.

## 4.1 État de l'art

De manière très générale, l'approche expérimentale en science consiste à réaliser des expérimentations en contrôlant certains paramètres appelés facteurs expérimentaux. En particulier, la mise en place d'une expérience perceptive nécessite de choisir ces facteurs expérimentaux, ainsi que les niveaux (ou modalités) que ceux-ci peuvent prendre. La théorie des plans d'expériences (*Design of Experiments*, DOE) permet ensuite de planifier les mesures qui seront nécessaires pour étudier l'effet des facteurs expérimentaux de la manière la plus efficace possible, en termes de nombre de sujets, nombre de mesures par sujet et nombre total de mesures. Dans le cas des algorithmes génétiques interactifs (IGA), il est notamment important de restreindre la taille de l'espace de conception, car une même personne peut seulement réaliser un nombre limité d'évaluations fiables au cours d'une même expérience (TAKAGI 2001). À notre connaissance, les études sur les IGAs tentent de répondre à ce problème, non pas en proposant des méthodes de choix informé de l'espace de conception, mais en tentant d'améliorer les performances de l'algorithme afin de converger plus rapidement. En effet, selon (PEI et TAKAGI 2018), les publications scientifiques sur les IGAs portent généralement sur trois types de contributions :

1. Des proposition concernant la structure de l'algorithme et les modes d'interaction avec l'utilisateur·rice·, dans le but de limiter sa fatigue,
2. L'application des IGAs à des problèmes d'optimisation de design,
3. L'application des IGAs à l'étude de la psychologie et/ou la perception humaine.

Le premier type de contribution est motivé par une volonté d'améliorer l'ergonomie des outils basés sur des IGAs, ainsi que l'efficacité du processus d'optimisation, dans un contexte où le nombre de configurations de design qu'il est possible d'évaluer est très limité. Ce type de contribution est le plus représenté parmi les publications que nous avons pu trouver. Dans le deuxième cas, il s'agit de montrer l'intérêt des IGAs pour s'attaquer

à un problème industriel donné, généralement avec une proposition d'algorithme adapté. Dans ce type de contribution, peu d'informations sont données sur les motivations qui sous-tendent le choix de l'espace de conception, si ce n'est concernant le choix du type d'objet à optimiser. Très peu d'articles portent sur le troisième type d'application, dans lequel l'IGA n'est pas pensé comme un outil d'aide à la conception, mais comme une alternative à un plan d'expérience pour étudier des dimensions liées à la psychologie et/ou la perception. Le choix des variables indépendantes se fait alors en fonction du but de l'étude.

Notre cadre d'application se situe à mi-chemin entre les deux derniers points. En effet, nous cherchons d'une part à optimiser des sons selon une dimension perceptive, et d'autre part à analyser les solutions obtenues pour produire des recommandations pour un·e designer·euse sonore. Comme nous l'avons dit, nous n'avons pas pu trouver d'étude abordant ce sujet dans le cas de problèmes d'optimisation interactive. Cependant, des problèmes similaires sont présents dans certains domaines d'ingénierie, tels que l'aéronautique ou l'industrie automobile (SHAN et WANG 2010). En effet, de la même manière que la fatigue des sujets limite le nombre d'évaluations pendant une expérience d'optimisation interactive, le temps de calcul nécessaire à la simulation précise de certains processus physiques peut parfois prendre plusieurs jours. De plus les modèles étudiés dans ces domaines dépendent généralement d'un grand nombre de variables. Il est donc inenvisageable d'optimiser un design modélisé ainsi par des méthodes numériques. Enfin, les calculs nécessaires à ces simulations sont complexes et la fonction à minimiser ne possède pas nécessairement d'expression analytique simple.

Dans (SHAN et WANG 2010), les auteur·rice·s font une revue des approches existantes pour répondre à ce type de problème, caractérisé par un espace de conception de grande dimension, des évaluations longues ou coûteuses et une absence d'expression analytique de la fonction-objectif. Il·elle·s mentionnent notamment quatre stratégies :

1. Utiliser un algorithme adapté,
2. Donner un plus grand contrôle au·à la concepteur·rice sur le processus d'optimisation,
3. Construire un modèle de substitution de la quantité à minimiser,
4. Réduire la complexité du problème à résoudre, grâce à une des méthodes décrites ci-dessous (*mapping, decomposition, screening, space reduction*).

Ainsi, ces approches sont similaires à celles proposées dans le cas des problèmes d'optimisation interactive. En effet, les deux premiers points sont semblables au premier type

de contribution mentionné au début de cette section, c'est-à-dire l'amélioration des performances du processus d'optimisation grâce à la modification de l'algorithme utilisé et du mode d'interaction avec l'utilisateur-riche. Le troisième point est très largement exploré pour tenter de palier le problème de la fatigue des utilisateur-riche-s (PEI et TAKAGI 2018), en remplaçant une partie des évaluations du sujet par celles d'un modèle de substitution. Concernant le quatrième point, qui nous intéresse ici, les auteur-riche-s mentionnent les stratégies suivantes :

1. *Mapping* (également appelé extraction de caractéristiques) : Créer de nouvelles variables à partir des variables initiales,
2. *Decomposition* : Décomposer le problème en sous-problèmes de plus petites dimensions,
3. *Screening* (aussi appelé sélection de caractéristiques) : Identifier les variables les plus influentes et fixer celles qui le sont peu à des valeurs arbitraires. Il existe notamment des plans d'expériences spécifiquement conçus dans ce but (ANTONY 2014),
4. *Space reduction* : Réduire les plages de valeurs (ou les nombres de niveaux, dans le cas discret) que peuvent prendre les variables de conception.

Toutes ces méthodes ont pour but de choisir un espace de conception, comme sous-espace d'un espace plus grand qu'il est inenvisageable d'explorer avec les outils d'optimisation à disposition. Afin de déterminer quelles méthodes sont appropriées à notre problème nous souhaitons faire des simulations d'expériences d'optimisation interactive. Ainsi, nous pouvons étudier l'efficacité des différentes méthodes avec un faible coût expérimental.

## 4.2 Réduction d'espace de conception

### 4.2.1 Simulations

Nous voulons étudier différentes méthodes de sélection informée d'espace de conception, au regard de leur impact sur un processus d'optimisation interactive. Comme mentionné dans (PEI et TAKAGI 2018), afin de comparer les performances du processus d'optimisation sous de multiples conditions, il est commun de simuler ce dernier afin de réduire le coût expérimental. Nous construisons donc des modèles de sujets qui nous serviront de sujets virtuels pendant des exécutions d'IGA.

Pour cela, nous utilisons les données de l'expérience 1 (voir chapitre 2), pour laquelle les 28 sujets fiables ont chacun fourni 99 évaluations de désagrément et de temps de réaction. Les expériences simulées traiteront donc du même problème de design de son d'alerte de véhicule silencieux (VS) que dans les chapitres 2 et 3. Nous construisons un modèle linéaire sans interaction pour chaque sujet et chacun des deux attributs perceptifs. L'expression de ces modèles est similaire à celle du modèle construit dans le chapitre 3 pour l'analyse des données de l'expérience 1 (équation 3.1, page 74), à l'exception de l'absence du facteur sujet :

$$y_{ijklmno} = \mu_o + V_{1io} + V_{2jo} + V_{3ko} + V_{4lo} + V_{5mo} + V_{6no} + \epsilon_{ijklmno} \quad (4.1)$$

avec :

- $y_{ijklmno}$  : l'évaluation de désagrément (ou de temps de réaction) du son ( $V_1 = i, V_2 = j, V_3 = k, V_4 = l, V_5 = m, V_6 = n$ ) par le sujet  $o$ ,
- $\mu_o$  : intercept,
- $V_{1io}$  : coefficient du niveau  $i$  de la variable  $V_1$  pour le modèle du sujet  $o$ , avec  $\sum_i V_{1io} = 0$  (paramétrage centré). Les autres coefficients correspondent aux variables  $V_2, V_3, V_4, V_5, V_6$ ,
- $\epsilon_{ijklmno}$  : terme d'erreur,  $\epsilon \sim \mathcal{N}(0, \sigma_o^2)$ .

Chaque modèle est caractérisé par 19 coefficients à estimer (une constante et trois paramètres par variable de conception ; comme nous utilisons un paramétrage centré, la valeur du coefficient correspondant au quatrième niveau d'une variable se déduit des valeurs des trois autres). Ces coefficients sont ajustés par la méthode des moindres carrés. Pour le désagrément, le  $R^2$  ajusté médian sur les 28 modèles construits vaut 0.52. Il vaut 0.34 pour le temps de réaction. Les modèles construits nous permettent de simuler deux expériences d'optimisation indépendantes, à savoir d'une part la minimisation du désagrément et d'autre part la minimisation du temps de réaction, chacune avec une cohorte de 28 sujets virtuels. Le modèle de synthèse sonore du problème simulé est le même que pendant l'expérience 1, contrôlé par six variables de conception de quatre modalités chacune.

## 4.2.2 Choix de la méthode

Parmi les quatre approches pour réduire la taille de l'espace de conception décrites dans la section 4.1, nous nous concentrons sur la dernière, à savoir celle visant à réduire les plages de valeurs (ou les nombres de niveaux, dans le cas discret) que peuvent prendre

les variables de conception.

Les deux premières approches (*mapping* et *decomposition*) ne sont pas applicables à notre problème, en particulier parce que nos variables de conception sont quantitatives discrètes et/ou qualitatives, afin de s’assurer d’une différence significative de timbre entre deux configurations possibles du synthétiseur. Cela implique qu’il n’est pas possible de réduire le nombre de variables de conception par *mapping*. En effet, bien que des contributions récentes (ESLING et al. 2019), basées sur des méthodes d’apprentissage machine, aient abordé ce problème dans le cas de la synthèse sonore, les approches proposées nécessitent de pouvoir contrôler le synthétiseur avec des variables continues.

Par ailleurs, la nature des variables de conception, le faible nombre de données par modèle de sujet et le bruit important dans ces données (erreur expérimentale) nous limitent à la construction de modèles linéaires sans interaction pour la simulation du processus d’optimisation interactive. Il ne nous est donc pas possible de tester l’effet d’une décomposition du problème en sous-problèmes indépendants. En effet, à cause de l’absence d’interaction entre les variables de conception dans les modèles de sujets, le problème simulé pourrait, de fait, être résolu en considérant chaque variable de conception séparément. Par exemple, considérons le problème d’optimisation suivant :

$$\inf_{x \in A} f(x), \tag{4.2}$$

où  $f$  est la fonction-objectif à minimiser et  $A$  est l’ensemble des solutions admissibles du problème. Posons  $x = [x_1 \cdots x_I]$ , où  $I$  est le nombre de variables de conception. Si la fonction-objectif est un modèle linéaire sans interactions, alors elle peut être décomposée sous la forme  $f(x) = \sum_{i=1}^I f_i(x_i)$ , où chaque fonction  $f_i$  ne dépend que de la variable  $x_i$ . Pour minimiser  $f$ , il est donc possible de minimiser les  $f_i$  indépendamment les uns des autres.

De premières expériences ne nous ont pas donné de résultats probants avec la méthode de *screening*, qui consiste à identifier et fixer à une valeur arbitraire les variables qui ont peu d’influence sur la fonction-objectif. Nous n’avons en effet pas observé d’amélioration du processus d’optimisation après avoir réduit l’espace de conception avec cette approche. Nous supposons que ces résultats sont dûs à l’algorithme d’optimisation utilisé. Il est fort probable que cette approche soit pertinente pour d’autres algorithmes, notamment ceux dont le nombre d’évaluations de la fonction-objectif à chaque itération est dépendant du nombre de dimensions du problème, comme par exemple un algorithme de descente de

gradient. Ainsi, nous présentons ici les résultats concernant la dernière méthode (*space reduction*), qui consiste à réduire les plages de valeurs (ou les nombres de modalités, dans le cas discret) que peuvent prendre les variables de conception.

### 4.2.3 Formalisation générale du problème

Nous considérons un problème d'optimisation dans lequel un attribut perceptif doit être minimisé à partir d'évaluations perceptives. Nous appelons  $f$  cet attribut, qui est donc la fonction-objectif du problème d'optimisation. L'espace de conception  $E$  est composé de  $I$  variables de conception, qui peuvent chacune prendre  $M_i$  niveaux différents. Nous avons :

$$E = \prod_{i=1}^I \llbracket 1; M_i \rrbracket \quad (4.3)$$

Un son est défini par le  $I$ -uplet  $D = (V_1, \dots, V_I) \in E$ , correspondant à une configuration du synthétiseur choisi pour générer les sons. Le nombre total de sons dans  $E$  vaut :

$$\prod_{i=1}^I M_i = M_1 \times M_2 \times \dots \times M_I \quad (4.4)$$

Dans notre cas, nous avons,  $I = 6$  et  $M_i = 4, \forall i \in \llbracket 1; 6 \rrbracket$ . Le nombre total de sons dans le plan complet est donc de 4096. Résoudre le problème d'optimisation revient à trouver la configuration  $D$  qui minimise  $f(D)$ . Idéalement, la méthode de réduction d'espace de conception devrait donc permettre de préserver cet optimum. Nous souhaitons donc réduire l'ensemble des sons  $E$  à un sous-ensemble  $E' \subset E$ , tel que le minimum de  $f$  dans  $E$  et dans  $E'$  soit le même. Cependant, ce problème étant complexe nous nous ramenons à une condition plus simple à respecter : nous voulons choisir un sous-ensemble  $E' \subset E$ , tel que  $\bar{f}(E') < \bar{f}(E)$ , où  $\bar{f}$  est la valeur moyenne de  $f$ . De cette manière, notre objectif est de diminuer la valeur moyenne de la fonction-coût des meilleures solutions trouvées par l'IGA, au moins au début du processus d'optimisation. Avec cette approche nous prenons cependant le risque que l'optimum dans  $E$  ne soit plus dans  $E'$ , en particulier s'il existe des interactions entre les variables de conception. En revanche, s'il n'y a pas d'interactions, l'optimum sera préservé.

Réduire la taille de l'espace de conception à partir de  $f(D)$  demanderait un grand nombre d'évaluations humaines. Or, la réduction de l'espace de conception est une étape préliminaire à une expérience d'optimisation interactive, elle même coûteuse en évaluations. Afin de limiter le coût expérimental, nous proposons donc de réduire l'espace de

conception en nous basant sur une quantité  $g(D)$  approchant  $f(D)$  (un approximateur de  $f(D)$ ).

#### 4.2.4 Élimination de niveaux

Nous cherchons à diminuer la taille de l'espace de conception en réduisant le nombre de niveaux par variables. Cela revient à sélectionner, pour chaque variable  $V_i$ , un sous-ensemble des modalités de départ. Pour cela, nous considérons une configuration du synthétiseur comme étant une variable aléatoire  $D = (V_1, \dots, V_I)$ , qui suit une distribution discrète uniforme sur tout l'espace de conception  $E = \prod_{i=1}^I \llbracket 1; M_i \rrbracket$ . Nous appelons  $G_i(m)$  l'espérance conditionnelle de  $g(D)$  sachant  $V_i$ , c'est-à-dire l'espérance de  $g(D)$  lorsque  $V_i = m$ .

$$G_i(m) = \mathbb{E} [g(V_1, \dots, V_I) \mid V_i = m] \quad (4.5)$$

Celle-ci peut être calculée comme :

$$G_i(m) = \frac{1}{\prod_{j \neq i} M_j} \sum_{v_1} \dots \sum_{v_{i-1}} \sum_{v_{i+1}} \dots \sum_{v_I} g(v_1, \dots, v_{i-1}, m, v_{i+1}, \dots, v_I) \quad (4.6)$$

Les détails de ce calcul sont donnés dans l'annexe A, section A.1. Pour chaque variable  $V_i$ , nous réduisons le nombre de modalités à  $k_i$ , avec  $1 \leq k_i < M_i$ , en écartant les  $M_i - k_i$  modalités pour lesquelles la valeur de  $G_i(m)$  est la plus grande. La figure 4.1 illustre le processus de sélection.

Soit  $D'$  la variable aléatoire correspondant à l'espace de conception réduit  $E' = \prod_{i=1}^I \llbracket 1; k_i \rrbracket$ . En éliminant les niveaux  $V_i$  qui maximisent  $G_i(m)$ , nous réduisons l'espérance de  $g$  :

$$\mathbb{E} [g(D')] < \mathbb{E} [g(D)] \quad (4.7)$$

La preuve de ce résultat est donnée dans l'annexe A, section A.2.

Dans cette étude, nous calculons les valeurs de  $G_i(m)$  avec l'équation 4.6, à partir des valeurs de  $g$  obtenues sur le plan factoriel complet. Nous avons donc défini une méthode pour réduire le nombre de modalités par variable, en estimant l'espérance conditionnelle  $G_i(m)$ , à partir d'une quantité  $g(D)$  approchant la fonction à minimiser  $f(D)$ . Nous éliminons ainsi les modalités pour lesquelles la valeur de  $g(D)$  est, en moyenne, la plus élevée.

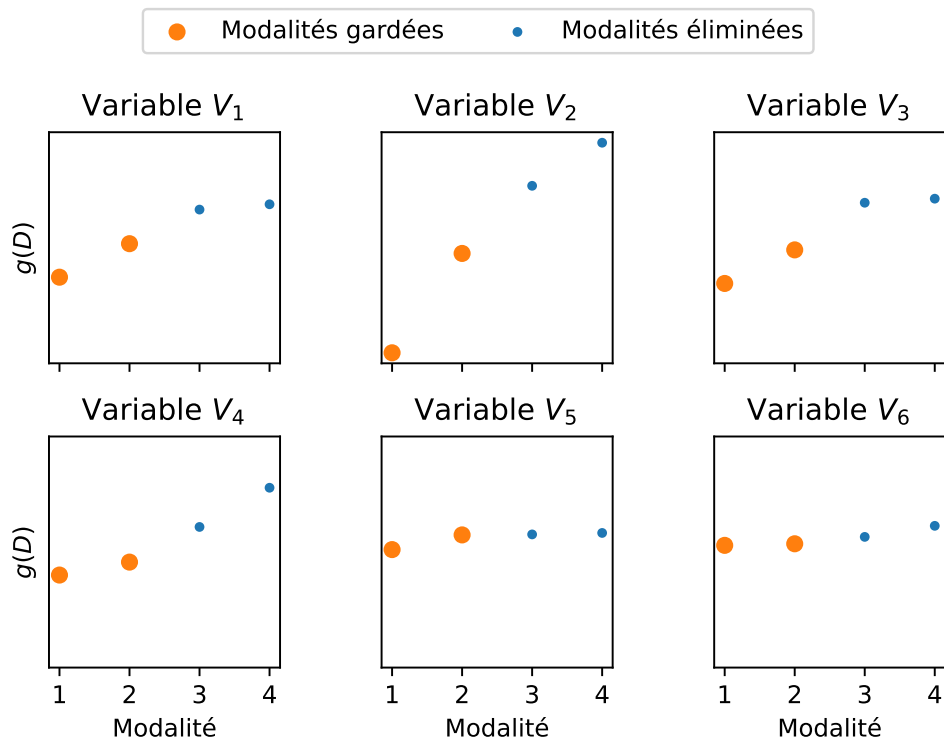


FIGURE 4.1 – Illustration du processus de sélection de variables, lorsque le nombre de variable est de  $I = 6$ , le nombre de modalités par variable est de  $M_i = 4, \forall i$ , et le nombre de modalités à garder vaut  $k_i = 2, \forall i$ . Chaque point représente une moyenne conditionnelle  $G_i(m)$ . Les modalités sont ordonnées et numérotées de sorte que les  $G_i(m)$  soient par ordre croissant.



## 4.2.5 Approximation de la fonction-objectif

La fonction  $g(D)$  est un approximateur de l'attribut perceptif à minimiser. En utilisant  $g(D)$  pour réduire l'espace de conception, nous souhaitons éliminer les modalités pour lesquelles la valeur de  $f(D)$  est également la plus élevée en moyenne. Il est donc important que  $g(D)$  constitue une bonne approximation de  $f(D)$  et que nous soyons en mesure d'évaluer la qualité de cette approximation. Nous rappelons que même dans la situation où  $g$  est une parfaite modélisation de  $f$ , la méthode que nous proposons présente le risque d'éliminer l'optimum de l'espace de conception, s'il existe des interactions entre les variables de conception.

Nous nous intéressons à deux types d'approximation en particulier :

1. **Approche descripteur (AD)** :  $g$  est un descripteur acoustique ou un modèle corrélé à  $f$  selon la littérature, comme par exemple la sonie (voir chapitre 1).
2. **Approche sujet (AS)** :  $g$  est un modèle de  $f$  construit à partir de données perceptives propres au problème traité, récoltées lors d'une pré-expérience.

Ces deux approches correspondent à deux cas de figure qui peuvent être rencontrés pendant un problème de design sonore. Si l'attribut perceptif d'intérêt est déjà connu, il se peut que des ressources suffisantes soient présentes dans la littérature (**AD**). A l'inverse, s'il s'agit d'un attribut encore peu ou pas étudié, il est nécessaire de récolter des données en amont, via par exemple des expériences perceptives (**AS**).

### Approche descripteur (AD)

**Désagrément** Comme dans le chapitre 2, nous présentons ici des études portant sur le désagrément, mais aussi sur la gêne sonore. En effet, bien que ces deux attributs soient différents, les études sur la qualité sonore des véhicules traitent parfois de l'un ou de l'autre. Nous considérons que des informations sur les descripteurs acoustiques propres à la gêne sonore peuvent nous aiguiller dans le choix d'un descripteur approprié du désagrément. Pour ce dernier, notre choix s'est porté sur la sonie de Zwicker (ZWICKER, FASTL et al. 1991) et la rugosité de Daniel et Weber (DANIEL et WEBER 1997). En effet, plusieurs auteur·rice·s (LAVANDIER et DEFREVILLE 2006 ; ELLERMEIER, MADER et DANIEL 2004 ; MOREL, MARQUIS-FAVRE et GILLE 2016 ; KACZMAREK et PREIS 2010) observent une forte corrélation entre ces deux quantités et le désagrément ou la gêne provoqués par un véhicule à moteur. En particulier, dans (LAVANDIER et DEFREVILLE 2006), 22 sujets évaluent des enregistrements de trafic urbain. Les auteur·rice·s trouvent un coefficient

de corrélation de Pearson de ( $r = 0,95$ ;  $p < 0,05$ ) entre la sonie de Zwicker moyenne et le désagrément perçu. Dans (ELLERMEIER, MADER et DANIEL 2004), un modèle de désagrément est construit à partir de comparaisons paires par 60 sujets. Parmi les 10 sons qui constituent le corpus, seul un son de camion est représentatif du cadre d'application étudié ici. La rugosité moyenne des sons évalués est corrélée de manière significative au désagrément prédit par le modèle ( $r = 0,97$ ;  $p < 0,001$ ). Une corrélation significative ( $r = 0,71$ ;  $p = 0,021$ ) est également trouvée avec la sonie médiane. Des évaluations de gêne sont fournies par 30 sujets, pour 7 catégories de véhicules, dans (MOREL, MARQUIS-FAVRE et GILLE 2016). Chaque catégorie est une combinaison entre un ou plusieurs types de véhicule (bus, deux-roues, véhicules légers et lourds) et un profil de conduite (accélération, vitesse constante, décélération). Pour cinq des sept catégories étudiées, la sonie de Zwicker est le facteur considéré qui contribue le plus à la gêne perçue. Enfin, dans (KACZMAREK et PREIS 2010), 19 sujets évaluent la gêne causée par quatre scénarios de passage de véhicule. Une forte corrélation est trouvée entre la gêne et la sonie moyenne ( $r = 0,96$ ;  $p < 0,05$ ), la rugosité ( $r = 0,97$ ;  $p < 0,05$ ) et la force de fluctuation ( $r = 0,93$ ;  $p < 0,05$ ).

Nous faisons l'hypothèse que les sons avec une sonie élevée ou une forte rugosité sont également très désagréables, et inversement. La sonie de Zwicker est calculée avec le module python *AudioCommons Timbral Models*<sup>1</sup>, qui implémente l'algorithme proposé dans (ZWICKER, FASTL et al. 1991). La rugosité est calculée grâce au module python *MoSQITo*<sup>2</sup>, qui implémente l'algorithme de (DANIEL et WEBER 1997). Dans la suite de ce manuscrit, nous appelons  $S$  la sonie et  $R$  la rugosité.

**Temps de réaction** Nous n'avons pas trouvé d'étude mentionnant l'existence d'un descripteur acoustique simple, qui serait corrélé avec le temps de réaction à l'écoute d'un stimulus sonore. Les études qui tentent d'établir un lien entre les propriétés acoustiques du son d'alerte d'un véhicule et le temps de réaction d'une personne à son approche le font au moyen de modèles ajustés sur des données qui leurs sont propres (LEE et al. 2017). Nous nous avons donc choisi de construire deux descripteurs à partir de notions intervenant dans le processus de détection : la saillance et l'émergence.

Dans (HUANG et ELHILALI 2017), les auteurs étudient la saillance auditive de 20 scènes sonores, via un protocole d'écoute dichotique. 50 sujets indiquent en temps réel

1. [https://github.com/AudioCommons/timbral\\_models](https://github.com/AudioCommons/timbral_models)

2. <https://github.com/Eomys/MoSQITo>

laquelle des deux scènes écoutées simultanément attire leur attention (celle de gauche ou celle de droite). À partir des évaluations de tous les sujets, des événements saillants sont mis à jour et leur niveau de saillance est estimé. Les auteur·rice·s observent une corrélation positive entre le niveau de saillance d'un événement et le changement de sonie dans le temps précédant l'évènement ( $r = 0,44$ ;  $p < 1,0 \cdot 10^{-23}$ ). De plus il·elle·s observent que le temps de réaction des sujets est plus court pour les événements ayant un niveau de saillance élevé. Nous choisissons donc de calculer le changement de sonie précédent le passage du VS, d'une manière similaire à (HUANG et ELHILALI 2017). Nous calculons la sonie moyenne dans deux fenêtres de temps,  $f_1 = [t_3 - 2; t_3 - 1,5]$  et  $f_2 = [t_3 - 0,5; t_3]$ , où  $t_3$  est le temps (en secondes) auquel le VS passe devant le sujet, suivant la convention choisie dans chapitre 2. La position des deux fenêtres est illustrée dans la figure 4.2. La différence de sonie  $dS$  vaut  $dS = S_2 - S_1$ , où  $S_1$  et  $S_2$  sont les niveaux de sonie moyens dans les fenêtres  $f_1$  et  $f_2$ , respectivement.

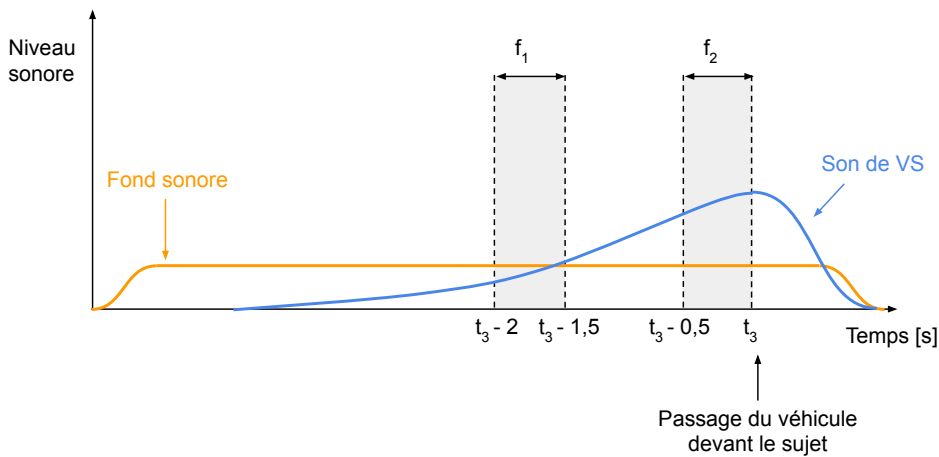


FIGURE 4.2 – Illustration du calcul de la différence de sonie. Deux fenêtres de 0,5 secondes,  $f_1$  et  $f_2$ , sont fixées en amont du passage du véhicule (en gris sur la figure). La sonie moyenne est calculée pour chaque fenêtre. Le descripteur  $dS$  est la différence entre les deux sonies moyennes.

Par ailleurs, les réglementations imposées par les Nation Unies (UN REGULATION 2017) sur le son d'alerte de VS requièrent un niveau sonore minimum, en dB(SPL), par bande de tiers d'octave. Cette contrainte a pour but d'assurer une émergence suffisante du son du VS dans un fond sonore urbain. Nous avons choisi de construire un deuxième descripteur suivant cette logique, basé sur la différence de puissance acoustique par bande entre le son de VS et le fond sonore. Pour cela, nous calculons le spectrogramme du signal d'approche du VS, avec 128 bandes mel entre 0 et 24kHz (ZWICKER et FASTL 2013). La puissance

du spectrogramme est ensuite moyennée en temps, pour obtenir le spectre de puissance moyen par bande mel. La même chose est faite pour un extrait du fond sonore, de même durée. Nous calculons ensuite la différence entre les deux spectres et sommes les valeurs correspondant aux bandes mel pour lesquelles la puissance du son de VS est supérieure à celle du fond sonore. Nous calculons ainsi la puissance sonore du signal de VS qui émerge du fond sonore, comme illustré par la zone en gris sur la figure 4.3.

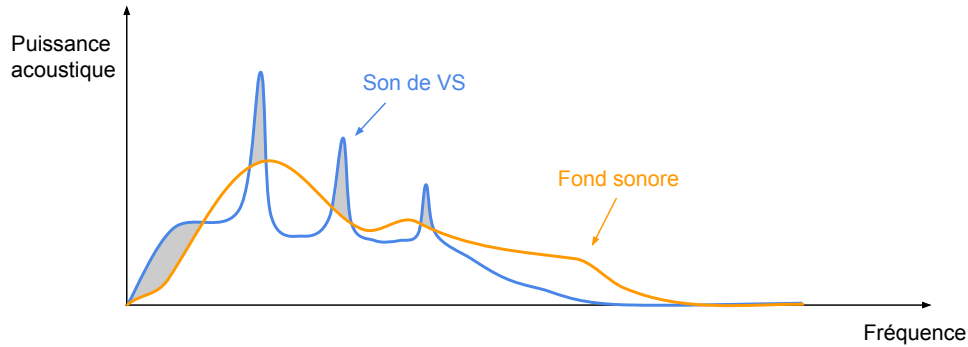


FIGURE 4.3 – Illustration du calcul de la différence de puissance. Les courbes bleue et rouge représentent les spectre de puissance moyen du son de VS et du fond sonore, respectivement. Le descripteur  $dP$  est calculé comme la puissance acoustique du son de VS qui émerge du fond sonore, représentée par les zones en gris sur la figure.

**Prise en compte du fond sonore** Pendant l'expérience 1, le fond sonore était choisi aléatoirement parmi un extrait de 42 secondes. Afin de prendre cela en compte, nous sélectionnons plusieurs extraits du fond sonore : un toutes les deux secondes en partant de  $t = 0$ . Nous obtenons 18 extraits de six secondes, correspondant à la durée du passage du véhicule. Les quatre descripteurs ( $S$ ,  $R$ ,  $dS$  et  $dP$ ) sont calculés en combinant le son de VS avec chacun des 18 extraits, puis en gardant la valeur moyenne sur l'ensemble des extraits.

### Approche sujet (AS)

Pour cette approche nous souhaitons construire un modèle de l'attribut perceptif à minimiser à partir de données récoltées lors d'une pré-expérience. Dans le cas de cette étude simulée, nous utilisons les données de l'expérience 1 et construisons, pour chaque attribut perceptif, un modèle à partir des évaluations de l'ensemble des sujets qui ont participé à l'expérience 1. Nous utilisons pour cela des modèles linéaires sans interaction,

comme ceux utilisés pour construire les sujets virtuels (voir équation 4.1). Cependant, ici nous ne construisons qu'un seul modèle par attribut perceptif, représentant l'ensemble du groupe de sujets. Pour les simulations, nous utilisons une approche similaire à la méthode de validation croisée (REFAEILZADEH, TANG et LIU 2009). Par exemple, lors des simulations avec le modèle du sujet  $o$ , le modèle de désagrément  $g$  utilisé pour réduire l'espace de conception est construit à partir des données de tous les sujets  $p \neq o$ . Cette façon de faire est opérationnelle, car elle correspond à ce que nous ferions préalablement à une expérience d'optimisation perceptive interactive réelle pour avoir un modèle perceptif approché, c'est-à-dire prendre les évaluations d'un groupe de sujets.

## 4.3 Expérience 4 : simulation d'IGA

Maintenant que nous avons défini une méthode de réduction informée d'espace de conception, nous la mettons en oeuvre lors d'une expérience d'optimisation perceptive interactive simulée avec des modèles de sujets virtuels.

### 4.3.1 Objectifs

Les objectifs de cette étude sont de :

1. Déterminer si la réduction de la taille de l'espace de conception par l'élimination de modalités permet d'améliorer la qualité des solutions de design trouvées au cours du processus d'optimisation,
2. Comparer les performances des sons trouvés en fonction de l'approche de réduction d'espace choisie (**AD** ou **AS**).

Pour cela, nous simulons un processus d'optimisation perceptive interactive après avoir réduit l'espace de conception, pour les deux approches **AD** et **AS**. Nous simulons également le processus d'optimisation sur l'espace complet (**EC**), qui nous sert de référence. La figure 4.4 illustre le protocole expérimental.

### 4.3.2 Matériels et Méthodes

#### Stimuli

Les stimuli sont les sons générés par le synthétiseur utilisé lors de l'expérience 1 (voir section 2.2). Le fond sonore est également le même. Les modèles de sujets virtuels ont été

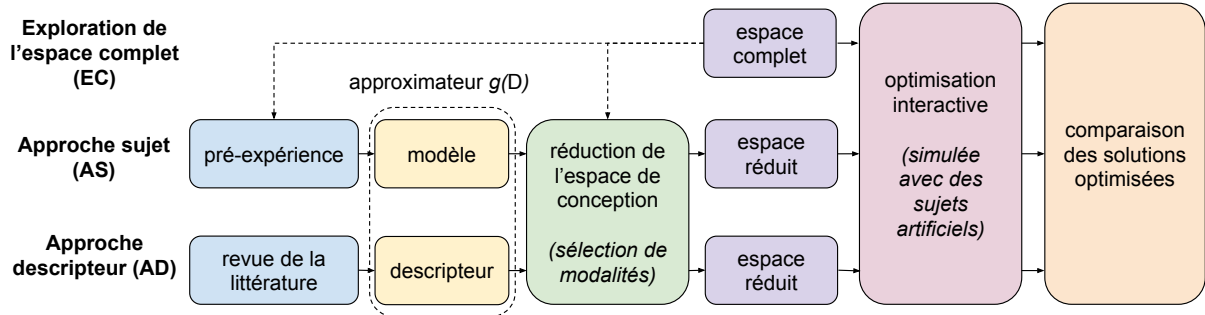


FIGURE 4.4 – Illustration du protocole de l'expérience 4.

construits à partir d'évaluations de ces stimuli pendant l'expérience 1, mais produisent des prédictions de désagrément ou de temps de réaction à partir des valeurs des variables de conception (voir équation 4.1).

### Plan d'expérience

Tous les sujets virtuels passent toutes des conditions expérimentales :

- Approches **EC**, **AS**, **AD-S** et **AD-R** pour le désagrément,
- Approches **EC**, **AS**, **AD-dS** et **AD-dP** pour le temps de réaction.

À chaque fois, nous étudions les cas où deux et trois modalités sur quatre sont gardées par variable après réduction de l'espace de conception.

### Participant·e·s

28 sujets virtuels (voir section 4.2.1) sont utilisés pour simuler le processus d'optimisation interactive. En considérant les prédictions faites par ces modèles (voir équation 4.1) comme des évaluations fournies par des sujets, nous pouvons simuler le processus d'optimisation interactive avec IGA plusieurs fois par sujet. Lorsque les simulations sont effectuées pour un modèle de sujet donné, le modèle utilisé dans l'approche (**AS**) pour réduire l'espace de conception est construit à partir des données des 27 sujets restants. Cette approche est similaire à une méthode de validation croisée.

### Protocole

**Implémentation de l'IGA** L'IGA utilisé pour ce problème est mono-objectif. La première génération de  $N$  sons est générée via un échantillonnage par hypercube latin. Une

fois tous les sons de la première génération évalués, des opérateurs de copie, de croisement et de mutation leur sont appliqués. Contrairement à l'algorithme utilisé pour l'expérience 1, les opérateurs de croisement et de mutation sont appliqués de manière séquentielle, et non pas de manière exclusive :

1. Les meilleurs sons, aussi appelés **élites**, sont directement copiés dans la génération suivante. Le nombre de meilleurs sons est calculé à partir d'un taux d'élitisme  $p_{el}$ , comme étant  $p_{el} \times N$ .
2. Deux parents sont sélectionnés par **tournoi binaire** parmi les  $N$  sons de génération qui vient d'être évaluée, à partir de leur valeur de fonction-objectif.
3. Deux nouveaux sons sont créés à partir des deux parents, via un **croisement uniforme**, illustré dans la figure 4.5. Pour ce type de croisement, les valeurs de chaque gène des deux chromosomes parent sont distribuées aléatoirement entre les deux chromosomes enfant.
4. Les deux chromosomes enfant sont ensuite **mutés** avec un taux de mutation  $p_{mu}$ . Cela signifie que chaque gène a une probabilité de  $p_{mu}$  de prendre une autre valeur, choisie aléatoirement.
5. Les deux chromosomes enfant sont ajoutés à la génération suivante, à moins qu'il ne manque qu'un individu pour atteindre  $N$  sons, auquel cas seul un des deux enfant, choisi aléatoirement, est ajouté.
6. Les étapes 2 à 5 sont répétées jusqu'à avoir  $N$  individus pour la prochaine génération.

Afin de favoriser la diversité génétique et d'éviter une convergence prématurée, l'algorithme est construit de telle sorte qu'il n'est pas possible d'avoir deux sons identiques dans une même génération. Le nombre d'individus par générations (9) et le nombre de générations (11) sont les mêmes que pendant l'expérience 1 (voir section 2.2, chapitre 2). Le taux de mutation de l'IGA est de  $p_{mu} = 0,05$  et le taux d'élites est de  $p_{el} = 0,2$ , soit deux élites à chaque génération.

**Procédure expérimentale** L'optimisation avec un IGA étant un processus stochastique, nous utilisons la méthode de Monte-Carlo pour estimer les performances des sons trouvés (ABDI et al. 2009). Pour chaque sujet virtuel, nous effectuons 50 simulations du processus.

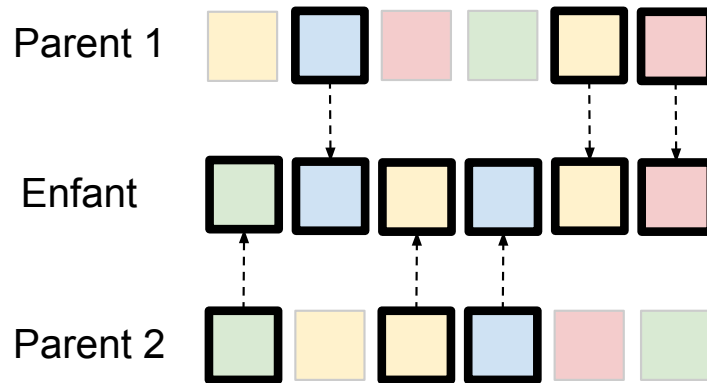


FIGURE 4.5 – Exemple de croisement uniforme avec deux parents. Chaque case est un gène dont la valeur est représentée par une couleur. Le chromosome enfant est créé en choisissant aléatoirement chaque gène parmi ceux des deux parents.

### Analyse

Au cours d'une simulation, le score minimum de désagrément/temps de réaction est enregistré à chaque génération. Nous faisons ensuite la moyenne de ce score sur l'ensemble des simulations. Enfin, nous comparons les scores obtenus avec chaque méthode à chaque génération au moyen de tests statistiques (test U de Mann-Whitney unilatéral).

### 4.3.3 Résultats

#### Désagrément

La figure 4.6 montre les résultats des simulations du processus d'optimisation pour le désagrément. Cette figure montre tout d'abord que plus l'espace de conception est réduit, plus la différence de désagrément entre la première et la dernière génération est faible. Un test de Shapiro-Wilk nous indique qu'aucun des points de la figure ne correspond à une distribution normale (avec  $\alpha = 0,05$ ). Pour la suite de l'analyse, nous utilisons donc un test non paramétrique, le test U de Mann-Whitney unilatéral. Nous faisons un test statistique à chaque génération et notons la dernière génération à laquelle les méthodes considérées donnent des résultats significativement meilleurs que la méthode de référence (**EC**). Les résultats de ces tests sont présentés dans le tableau 4.1. Nous observons que réduire l'espace de conception permet de trouver des sons moins désagréables dès la première génération, quelle que soit l'approche choisie. Le nombre de générations pour lesquelles nous observons une amélioration est plus petit lorsque deux modalités



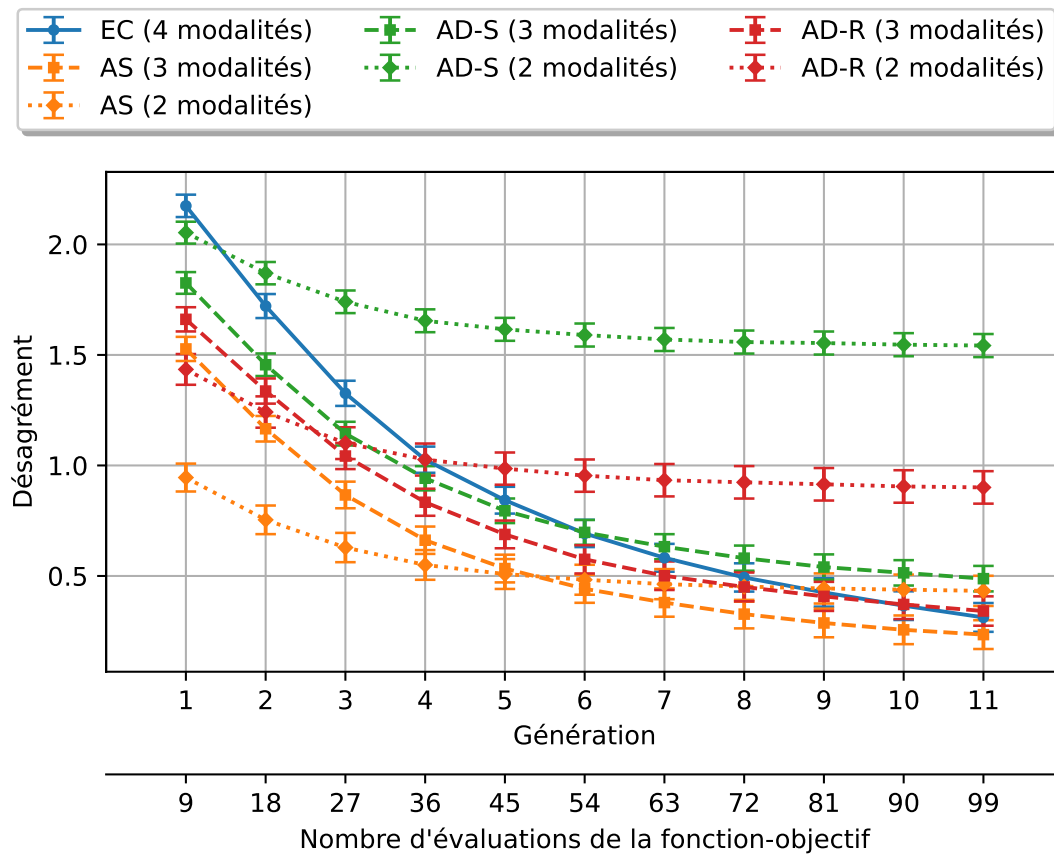


FIGURE 4.6 – Résultats des simulations pour la minimisation du désagrément. Chaque point correspond à la moyenne, sur toutes les simulations (28 sujet virtuels  $\times$  50 simulations par sujet), de la plus petite valeur de désagrément à chaque génération. L'axe horizontal représente le nombre de générations, ainsi que le nombre d'évaluations de la fonction-coût nécessaire. Le nombre d'évaluation correspond au nombre de générations, multiplié par le nombre de sons par génération, qui est ici de neuf. **EC** : exploration de l'espace complet (référence), **AS** : approche sujet, **AD-S** : approche descripteur, avec la sonie, **AD-R** : approche descripteur, avec la rugosité. Le nombre de modalités gardées pour chaque courbe est indiqué entre parenthèses dans la légende.

sont gardées, comparé au cas où trois modalités sont gardées. Cela illustre le fait que les descripteurs/modèles utilisés pour réduire l'espace conception restent des approximations de la fonction objectif, et qu'il n'est donc pas possible de faire une trop grande réduction à partir de leurs prédictions. Les meilleurs résultats sont obtenus avec l'approche **AS**, quel que soit le nombre de modalités gardées. Quand trois modalités sont gardées par variable, l'approche **AS** permet de trouver de meilleurs sons jusqu'à la neuvième génération, ce qui est presque le nombre maximum de générations pour cette expérience. Si suffisamment de générations sont considérées, l'exploration de l'espace complet finit nécessairement par permettre d'atteindre de meilleures solutions que l'exploration de l'espace réduit. Cependant, la génération à partir de laquelle l'approche **EC** donne de meilleures solutions dépend de l'approximateur utilisé pour réduire l'espace de conception.

TABLE 4.1 – **Dernière** génération à laquelle réduire l'espace de conception permet de trouver des sons significativement **moins désagréables** ( $p < 0.05$ ). **AD-S** : approche descripteur, avec la sonie, **AD-R** : approche descripteur, avec la rugosité, **AS** : approche sujet.

Méthode	Nombre de modalités gardées	
	3	2
<b>AD-S</b>	5	1
<b>AD-R</b>	4	3
<b>AS</b>	9	6

Le tableau 4.2 montre les générations à partir desquelles chaque approche donne des sons significativement plus désagréables que la référence (**EC**). L'approche descripteur fait moins bien que la référence à partir de la troisième génération pour la sonie (**AD-S**) et à partir de la sixième génération pour la rugosité (**AD-R**). Aucune des autres approches testées n'est cependant surpassée par la référence pour le nombre maximum de générations considéré ici. En conclusion, dans la plupart des cas testés, réduire l'espace de design permet de trouver des solutions qui sont au moins de qualité équivalente à celles trouvées lors d'une exploration de l'espace complet et ainsi d'atteindre plus rapidement des solutions d'une qualité donnée.

### Temps de réaction

La figure 4.7 montre les résultats des simulations pour le temps de réaction. De la même manière que pour la figure 4.6, plus la réduction de l'espace de conception est grande, plus la différence de temps de réaction moyen des meilleurs sons entre la première

TABLE 4.2 – **Première** génération à laquelle réduire l'espace de conception amène à trouver des sons significativement **plus désagréables** ( $p < 0.05$ ). **AD-S** : approche descripteur, avec la sonie, **AD-R** : approche descripteur, avec la rugosité, **AS** : approche sujet.

Méthode	Nombre de modalités gardées	
	3	2
<b>AD-S</b>	-	3
<b>AD-R</b>	-	6
<b>AS</b>	-	-

et la dernière génération est faible. En premier lieu, nous constatons que le descripteur basé sur la variation de sonie avant le passage du véhicule (**AD-dS**) constitue un mauvais approximateur du temps de réaction. En effet, les solutions trouvées lors du processus d'optimisation sont moins bonnes que celles trouvées avec la méthode de référence (**EC**) dès le début du processus d'optimisation, quel que soit le nombre de modalités gardées. À première vue, l'approche **AS** semble aboutir à des performances similaires que dans le cas du désagrément. Le second descripteur choisi (la différence de puissance par bande, **AD-dP**), semble réduire l'espace de conception de manière pertinente, puisque les sons trouvés au cours du processus de simulation ont des performances proches de ceux trouvés après avoir réduit l'espace avec l'approche **AS**. Cette fois-ci encore, un test de Shapiro-Wilk nous indique qu'aucun des points de la figure ne correspond à une distribution normale ( $\alpha = 0,05$ ). Un test U de Mann-Whitney unilatéral est donc utilisé pour les comparaisons. La table 4.3 indique les générations jusqu'auxquelles les méthodes testées permettent de trouver des meilleurs sons que la méthode de référence (**EC**). Ces résultats confirment les observations de la figure 4.7. L'approche descripteur, avec la différence de sonie (**AD-dS**), ne fait jamais mieux que la référence (**EC**). L'approche descripteur, avec la différence de puissance (**AD-dP**), permet d'obtenir des résultats similaires à ceux obtenus avec l'approche **AS**. Pour cette dernière, les performances sont encore meilleures que dans le cas du désagrément, en particulier lorsque trois modalités sont conservées.

La table 4.4 montre les générations à partir desquelles les approches testées donnent des résultats significativement moins bons que la référence. L'approche (**AD-dS**) est rapidement dépassée, dès la première génération lorsque deux modalités sont gardées et à partir de la deuxième génération lorsque trois modalités sont gardées. En gardant trois modalités, l'approche **AD-dP** produit des résultats similaires à l'approche **AS**, jusqu'à la dernière génération considérée. Cependant, la méthode **AS** produit toujours les meilleurs

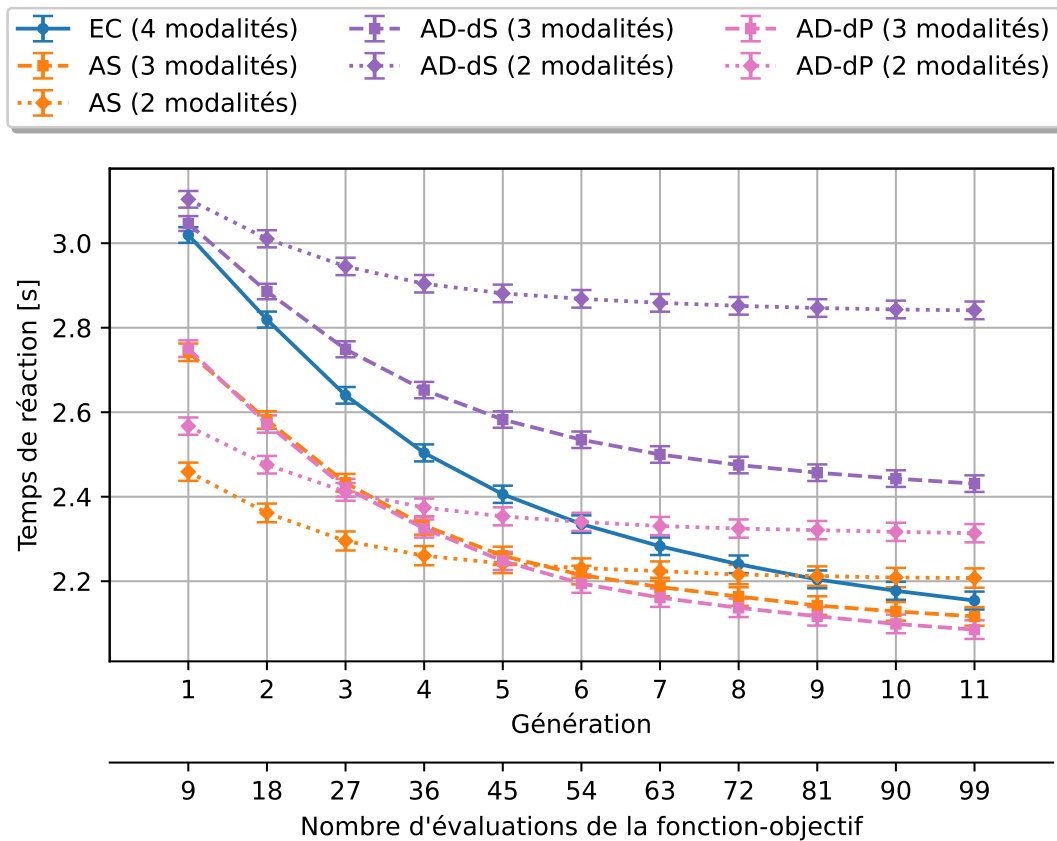


FIGURE 4.7 – Résultats des simulations pour la minimisation du temps de réaction. Chaque point correspond à la moyenne, sur toutes les simulations (28 sujet virtuels  $\times$  50 simulations par sujet), de la plus petite valeur de désagrément à chaque génération. L'axe horizontal représente le nombre de générations, ainsi que le nombre d'évaluations de la fonction-coût nécessaire. Le nombre d'évaluation correspond au nombre de générations, multiplié par le nombre de sons par génération, qui est ici de neuf. **EC** : exploration de l'espace complet (référence), **AS** : approche sujet, **AD-dS** : approche descripteur, avec la différence de sonie, **AD-dP** : approche descripteur, avec la différence de puissance. Le nombre de modalités gardées pour chaque courbe est indiqué entre parenthèses dans la légende.

TABLE 4.3 – Dernière génération à laquelle réduire l'espace de conception permet de trouver des sons avec un temps de réaction significativement **plus court** ( $p < 0.05$ ). **AD-dS** : approche descripteur, avec la différence de sonie, **AD-dP** : approche descripteur, avec la différence de puissance, **AS** : approche sujet.

Méthode	Nombre de modalités gardées	
	3	2
<b>AD-dS</b>	-	-
<b>AD-dP</b>	11	5
<b>AS</b>	10	7

résultats, car elle ne fait jamais significativement moins bien que la méthode de référence, quel que soit le nombre de modalités gardées.

TABLE 4.4 – Première génération à laquelle réduire l'espace de conception amène à trouver des sons avec un temps de réaction significativement **plus long** ( $p < 0.05$ ). **AD-dS** : approche descripteur, avec la différence de sonie, **AD-dP** : approche descripteur, avec la différence de puissance, **AS** : approche sujet.

Méthode	Nombre de modalités gardées	
	3	2
<b>AD-dS</b>	2	1
<b>AD-dP</b>	-	8
<b>AS</b>	-	-

#### 4.3.4 Discussion

Nous observons qu'un choix judicieux d'approximateur de la fonction-objectif permet, en moyenne, de trouver de meilleurs sons sur les premières générations qu'explorer l'espace complet (**EC**) et d'accélérer ainsi le processus d'optimisation. Nous avons pu trouver des études démontrant un lien significatif entre le désagrément lié aux sons de véhicules et plusieurs descripteurs acoustiques. Cela nous a permis de faire un choix de descripteurs informé, dont l'impact positif sur le processus d'optimisation se reflète dans la figure 4.6. En effet, nous remarquons une amélioration des solutions trouvées en début d'optimisation dans la plupart des cas testés. À l'inverse, nous avons construit des approximateurs du temps de réaction à partir d'informations trouvées dans la littérature, sans cependant avoir d'assurance de liens de corrélation. Ainsi, la figure 4.7 montre des résultats plus variables en fonction du descripteur. Là où l'un permet une réduction judicieuse de l'espace de conception (approche descripteur avec la différence de puissance, **AD-dP**), l'autre

(approche descripteur avec la différence de sonie, **AD-dS**) ne fait au contraire jamais mieux que la référence. En utilisant des données propre au problème (**AS**), nous voyons une amélioration des solutions trouvées dans les deux problèmes traités.

Ces approches illustrent différents cas de figure qui peuvent être rencontrés dans un problème de design sonore. Là où il est peu surprenant que l'approche **AS** donne de meilleurs résultats que l'approche **AD**, il est intéressant de noter que des descripteurs qui fournissent une approximation suffisamment bonne de la fonction-objectif peuvent tout de même être trouvés dans la littérature. Le choix entre les différentes approches proposées dépend bien sûr de la disponibilité ou de l'absence d'informations pertinentes dans la littérature, mais également du nombre d'évaluations humaines disponibles pour l'étude. En effet, il semble, d'après la figure 4.6, que s'il n'est possible d'obtenir qu'un nombre très limité d'évaluations par des personnes, il peut y avoir un bénéfice à réduire l'espace de conception en amont grâce à un approximateur approprié trouvé dans la littérature. Si, au contraire, un plus grand nombre d'évaluations sont disponibles, l'approche **AS** semble mieux adaptée pour améliorer la qualité des solutions trouvées sur un plus grand nombre de générations. Il y a cependant un coût expérimental additionnel dû à l'acquisition des données qui servent à construire le modèle. Nous n'avons pas inclus ce coût dans l'analyse de cette expérience, car les données utilisées n'ont pas été obtenues à partir d'un plan d'expériences choisi en fonction du modèle construit. Une approche expérimentale en deux étapes, incluant plan d'expérience et algorithme d'optimisation interactive, semble être une piste intéressante, comme cela a été mentionné dans (PETIOT, CERVANTES CHAVEZ, BOIVIN et al. 2015) et (PETIOT, DENJEAN et al. 2020).

Nous avons ici choisi d'éliminer le même nombre de modalités par variable de conception. Nous pouvons voir que l'impact de la réduction de l'espace de conception sur le processus d'optimisation dépend fortement du nombre de modalités éliminées. Un choix judicieux est donc important, car éliminer un trop grand nombre de modalités augmente le risque d'éliminer des solutions pertinentes, si l'approximateur  $g$  n'est pas suffisamment fidèle à la fonction-objectif  $f$ . Enfin, nous avons ici utilisé comme sujets virtuels des modèles simples, qui ne prennent pas en compte les interactions éventuelles entre les variables de conception. Ainsi, les performances obtenues avec l'approche **AS** pendant les simulations dépendent essentiellement de l'accord ou du désaccord entre les sujets. En pratique, si l'expérience d'optimisation faisait intervenir des sujets réels, il se pourrait qu'il y ait un fort effet d'interaction entre les variables de conception. En éliminant des modalités sur la base des moyennes conditionnelles correspondantes de l'approximateur, il y a un risque

supplémentaire que nous éliminons en même temps les solutions optimales de l'espace de conception.

## 4.4 Conclusion du quatrième chapitre

Dans ce chapitre, nous avons proposé et étudié deux approches pour réduire la taille d'un espace de conception, afin de répondre à la troisième question de recherche :

### Troisième question de recherche

Comment réduire l'espace de conception de manière informée, dans le cadre d'une expérience d'optimisation perceptive interactive ?

Le première approche (**AD**) est basée sur l'utilisation d'informations provenant de la littérature scientifique. La seconde (**AS**) est basée sur l'exploitation de données perceptives propres au problème d'optimisation à résoudre. Dans les deux cas, un approximateur de la fonction-objectif est utilisé pour éliminer les modalités qui donnent, en moyenne, les moins bonnes solutions (selon l'approximateur choisi). Nous avons pu montrer l'intérêt des approches proposées via des simulations d'expériences perceptives, traitant du problème de design de sons d'alerte de véhicule silencieux abordé au chapitre 2. Deux problèmes d'optimisation interactive distincts ont été abordés : la minimisation du désagrément et la minimisation du temps de réaction.

### Points clés à retenir

- Une étude basée sur des simulations numériques d'expériences nous confirme que la réduction d'espace de conception par élimination de modalités permet de trouver plus rapidement des solutions intéressantes au cours d'un processus d'optimisation interactive.
- En fonction des informations disponibles en amont et du budget expérimental, un gain de performances plus ou moins important peut être attendu.

# RÉDUCTION DE L'ESPACE DE CONCEPTION (PARTIE 2 : EXPÉRIENCES PERCEPTIVES)

---

Dans le chapitre 4, nous avons pu, au travers de simulations, définir une méthode de réduction d'espace de conception adaptée à un problème d'optimisation perceptive interactive. Nous nous choisissons de procéder en éliminant, pour chaque variable de conception, les modalités les moins susceptibles de produire des sons qui minimisent le critère d'optimisation du problème. Cette approche semble donner des résultats prometteurs pour des simulations numériques. Cependant, les modèles de sujets utilisés pendant ces simulations sont des modèles simplifiés qui ne rendent pas forcément compte de la complexité des processus perceptifs en jeu dans l'évaluation d'un son. De plus, ces modèles ne rendent pas compte du bruit dans les évaluations faites par les sujets. Or, ces facteurs sont susceptibles d'influencer à la fois la qualité de la réduction d'espace de conception et le processus d'optimisation perceptive interactive. Dans ce chapitre, nous souhaitons donc répondre à la quatrième question de recherche :

## Quatrième question de recherche

La méthode de réduction d'espace de conception proposée dans le cadre d'une expérience simulée est-elle efficace dans le cadre d'une expérience réelle ?

Pour répondre à cette question, nous nous intéressons à un nouveau problème de design sonore : celui du design de sons d'intérieur de véhicule autonome peu intrusifs. Nous souhaitons ainsi mettre en oeuvre la méthode de réduction proposée, au travers d'un problème innovant, encore peu étudié dans la littérature.



Ce chapitre débute par un état de l'art concernant le design de sons de véhicule autonomes, qui nous servira donc de cadre applicatif. Les deux parties suivantes traitent des expériences effectuées pour tenter de répondre à la question de recherche. La seconde partie traite de la pré-expérience servant à obtenir les informations nécessaires à la réduction de l'espace de conception. La troisième partie aborde l'expérience d'optimisation interactive visant à valider la réduction d'espace effectuée suite à l'expérience précédente.

## 5.1 Design sonore et véhicule autonome

Le terme de véhicule autonome (VA) désigne des véhicules de type automobile pour lesquels la conduite est totalement ou partiellement déléguée à une intelligence embarquée (un automate). La SAE (*Society of Automotive Engineers*) International définit une classification des véhicules autonomes en six niveaux (INTERNATIONAL 2021) :

- Niveau 0 : Pas d'automatisation.
- Niveau 1 (« *Hands On* ») : Conduite assistée.
- Niveau 2 (« *Hands off* ») : Automatisation partielle.
- Niveau 3 (« *Eyes off* ») : Automatisation conditionnelle.
- Niveau 4 (« *Mind off* ») : Automatisation élevée.
- Niveau 5 (« *Steering wheel optional* ») : Automatisation complète.

Les niveaux 1 et 2 sont déjà atteints dans des véhicules commercialisés, par exemple avec l'utilisation d'un régulateur automatique de vitesse. Pour les niveaux 3 et 4, la surveillance de la route en continu par le·a conducteur·rice n'est plus requise dans certains contextes. Au niveau 5, le véhicule est complètement autonome et ne nécessite aucune action de la part des passager·ère·s. En 2022, seuls quelques véhicules de niveau 3 ont été commercialisés. Bien que des véhicules complètement autonomes (niveau 5) ne soient pas encore sur le marché, FAGERLÖNN, LARSSON et MACULEWICZ 2020 soulignent que la confiance que leur accordent les passager·ère·s est l'un des principaux défis à relever pour leur adoption en masse. Plusieurs auteur·rice·s (BEATTIE, LYNNE et al. 2013 ; BEATTIE, BAILLIE, HALVEY et MCCALL 2014 ; BEATTIE, BAILLIE et HALVEY 2017 ; CHEN et CHEN 2021 ; DU et al. 2019 ; GANG et al. 2018 ; FAGERLÖNN, LARSSON et MACULEWICZ 2020 ; LARSSON, OPPERUD et al. 2009 ; LARSSON et VÄSTFJÄLL 2013 ; LARSSON, MACULEWICZ et al. 2019 ; MISDARIIS, CERA et RODRIGUEZ 2019) proposent de sonifier certaines actions du VA afin, entre autres, d'améliorer le niveau de confiance des passager·ère·s. Le design sonore intérieur de véhicules complètement autonomes (niveau 5) constitue un

cadre d'application nouveau et original (MISDARIIS, CERA et RODRIGUEZ 2019), auquel nous nous intéressons dans ce chapitre. De plus, cette application est propice à la mise en oeuvre de la méthode sujet proposée au chapitre 4. En effet, ce nouveau terrain d'étude offre la possibilité de manipuler une grandeur relativement haut niveau (l'intrusivité), peu connue de la littérature et pour laquelle il n'existe pas de métrique psychoacoustique qui permettrait de la représenter.

FAGERLÖNN, LARSSON et MACULEWICZ 2020 présentent une approche constituée de plusieurs strates sonores. Par exemple, une première strate, dite de perception, indique aux passager·ère·s que le VA est en mesure de percevoir son environnement, en jouant par exemple des sons à l'approche d'autres usager·ère·s de la route. Une autre strate, dite d'intention, sonifie les actions du véhicule à venir, comme des accélérations, des virages ou des freinages. Au travers de ces sons, les passager·ère·s peuvent développer une meilleure compréhension du fonctionnement du VA et une plus grande confiance dans la capacité de ce dernier à se déplacer de manière sécurisée et à réagir de manière appropriée quand cela est nécessaire (LARSSON, MACULEWICZ et al. 2019; MISDARIIS, CERA et RODRIGUEZ 2019; BEATTIE, LYNNE et al. 2013). Ces sons sont purement informatifs et n'ont donc pas pour vocation de susciter une réaction de la part des passager·ère·s. FAGERLÖNN, LARSSON et MACULEWICZ 2020 remarquent cependant qu'il est important d'éviter de surcharger les passager·ère·s d'informations. Il·elle·s évoquent par ailleurs qu'il serait pertinent de prendre en compte comment les informations sonifiées s'inscrivent dans l'environnement sonore général du VA, de sorte à pouvoir cohabiter avec d'autres sons. Leur étude vise donc à créer des sons subtils, agréables et moins saillants que les sons d'alerte habituellement utilisés dans les véhicules.

De même, dans (MISDARIIS, CERA et RODRIGUEZ 2019) les auteurs soulèvent, au travers de leur travail sur une demo-car de Renault (Symbioz), l'importance du concept de faible intrusivité (*low-intrusiveness*), en particulier pour les sons informatifs, ne nécessitant pas d'action particulière de la part des passager·ère·s (appelés *brain sounds*). Cette notion émerge dans un contexte où de nombreux signaux sonores artificiels sont inspirés de sons musicaux (CERA et MISDARIIS 2021). Les auteurs avancent que cette approche du design sonore n'est pas optimale et génère potentiellement certains effets indésirables, comme de la surprise, de l'intrusivité, de la distraction ou de la gêne. L'intrusivité est comprise comme étant la contribution à la gêne perçue ne relevant pas du niveau sonore (PREIS 1987), qui intervient à la première écoute d'un son (MISDARIIS, CERA et RODRIGUEZ 2019). En particulier, CERA et MISDARIIS 2021 lient l'intrusivité d'un son,

au contraste entre le spectre de ce son et celui du reste de l'environnement sonore, à son timbre, à sa durée, à la facilité avec laquelle il est reconnaissable et peut être mémorisé, ainsi qu'au contexte dans lequel il apparaît. En se focalisant sur la conception de sons peu intrusifs, le but est donc de créer des sons qui communiquent efficacement une information, en perturbant le moins possible l'environnement sonore. L'utilisation de sons peu intrusifs est également recommandée dans (LARSSON, MACULEWICZ et al. 2019), afin de favoriser une expérience positive pour les utilisateur·rice·s.

Ces différentes contributions ne donnent cependant pas les mêmes recommandations pour la conception de sons peu intrusifs. Par exemple, dans (FAGERLÖNN, LARSSON et MACULEWICZ 2020) et (LARSSON, MACULEWICZ et al. 2019), les auteur·rice·s utilisent des sons continus pour la strate d'intention, là où MISDARIIS, CERA et RODRIGUEZ 2019 privilégient des sons courts, aigus et de faible amplitude. Quelques pistes sont avancées dans (CERA et MISDARIIS 2021) concernant les propriétés d'un son peu intrusif, mais il n'existe pour l'instant pas de règles de conception communément admises. Enfin, comme cela est souligné dans (CERA et MISDARIIS 2021), la conception d'un son peu intrusif s'apparente à un problème d'optimisation avec de multiples contraintes, parfois contradictoires. Il s'agit en effet de créer un son qui ne soit ni trop discret, ni trop gênant, ni trop reconnaissable, mais qui transmette malgré tout une information claire.

## 5.2 Expérience 5 : pré-expérience pour la réduction de l'espace de conception

Nous allons traiter du problème du design de sons d'intérieur de VA peu intrusifs au travers du protocole expérimental développé jusqu'à présent, afin d'évaluer sa pertinence sur un cas d'étude nouveau et relativement inédit. L'intrusivité est un percept de niveau similaire au désagrément étudié au chapitre 4, mais sur lequel il existe bien moins d'informations dans la littérature. Nous mettons en oeuvre la méthode de réduction d'espace de conception proposée au chapitre 4, en utilisant l'approche sujet (**AS**), car nous avons besoin de récolter des données perceptives pour pouvoir estimer l'intrusivité des sons à optimiser.

### 5.2.1 Objectif

L'objectif de cette expérience est de construire un modèle simple d'intrusivité perçue en fonction des variables de conception du problème. Ce modèle sert ensuite à prédire l'intrusivité moyenne des sons en fonction des niveaux des variables de conception, afin d'informer un processus de sélection de modalités visant à réduire la taille de l'espace de conception.

### 5.2.2 Matériels et Méthodes

#### Stimuli

**Synthèse des sons de VA** Pour ce nouveau problème, nous avons la chance de collaborer avec le designer sonore Andrea Cera<sup>1</sup> qui a déjà travaillé avec l'équipe Perception et Design Sonores (PDS) de l'Ircam sur la conception de sons de véhicules silencieux (VS) et véhicules autonomes (VA) (MISDARIIS, CERA, LEVALLOIS et al. 2012 ; MISDARIIS et CERA 2013 ; MISDARIIS 2017 ; MISDARIIS, CERA et RODRIGUEZ 2019). Celui-ci nous transmet plusieurs programmes, appelés *patches*, conçus sur le logiciel Max/MSP<sup>2</sup> qui lui ont servi de base pendant son travail sur de précédents projets de sonification de VA. Un des *patches* sert à générer des sons de VA à partir d'un modèle de synthèse additive, tandis que deux autres servent à générer un bruit d'intérieur d'habitacle avec une méthode de synthèse granulaire. Max/MSP est un logiciel de programmation graphique dédié à la création de programmes multimédia interactifs. Ce logiciel est très utilisé pour des applications musicales, comme par exemple le prototypage de synthétiseurs virtuels. Sa méthode de programmation graphique permet une grande flexibilité d'utilisation. Max/MSP étant cependant un logiciel propriétaire, nous choisissons d'adapter le *patch* de départ dans sa version libre, le logiciel Puredata<sup>3</sup>. Grâce à plusieurs échanges avec Andrea Cera, nous adaptons à notre problème les *patches* que celui-ci nous a fournis.

Souhaitant partir sans a priori sur les caractéristiques d'un son peu intrusif, nous choisissons une architecture similaire de celle d'un synthétiseur modulaire, offrant par là une grande expressivité. Cette structure est en partie influencée par l'utilisation du logiciel Puredata pour le prototypage, celui-ci se présentant sous forme de modules interconnectés. Deux types de données existent dans Puredata : des signaux audio, échantillonnés à une

---

1. <https://www.andreacera.info/>

2. <https://cycling74.com/products/max>

3. <http://puredata.info/>

fréquence élevée (par exemple 44100 Hz), et des données de contrôle, échantillonnées à basse fréquence. Plus d'informations sur le logiciel Puredata et les outils de traitement du signal utilisés pour créer le synthétiseur peuvent être trouvées dans (PUCKETTE 2007).

La structure générale du synthétiseur est schématisée sur la figure 5.1. Celui-ci est composé de quatre sections, elle-mêmes composées de modules détaillés en annexe. Le synthétiseur comporte deux sources sonores. La première, correspondant à la section *SYNTHÈSE ADDITIVE* (voir figure B.2, page 161, en annexe pour le détail), consiste en une somme d'oscillateurs sinusoïdaux modulés en fréquence et en amplitude. La seconde, qui correspond à la section *BRUIT FILTRÉ* (voir figure B.3, page 161), est composée d'un générateur de bruit rose passant par un filtre passe-bande, dont les paramètres sont modulés en fréquence et en amplitude. Une section *MÉLANGEUR* (voir figure B.4, page 162) sert à définir l'amplitude relative des deux signaux générés  $y_{syn}(t)$  et  $y_{bf}(t)$ . Avant d'atteindre cette section, les signaux  $y_{syn}(t)$  et  $y_{bf}(t)$  sont tous les deux normalisés pour avoir un niveau sonore identique, en dB(A). Enfin, une section *FILTRAGE* (voir figure B.5, page 162) applique un filtre passe-bas au mélange  $y_{mix}(t)$  pour obtenir le signal final  $y_{total}(t)$ . Le signal  $y_{total}(t)$  est finalement égalisé en sonie (sonie de Zwicker ; ZWICKER, FASTL et al. 1991), afin que tous les sons produits par le synthétiseur soient approximativement au même niveau sonore perçu.

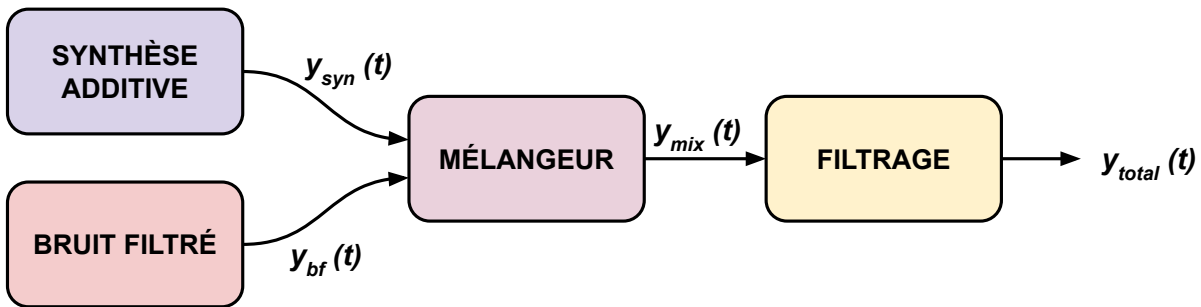


FIGURE 5.1 – Structure générale du modèle de synthèse utilisé pour construire les sons de VA.

En tout, le synthétiseur est contrôlé par 19 variables de conception, décrites qualitativement dans le tableau 5.1. Les sections, les modules et les variables de conception sont présentées plus en détails dans l'annexe B. Afin de soulager la notation de la formalisation mathématique dans la suite de ce manuscrit, un nom de variable générique  $V_i$ , pour  $i \in \{1, \dots, 19\}$ , est associé à chaque variable de conception. Celles-ci peuvent prendre entre 3 et 5 modalités, choisies empiriquement pour leurs capacités à produire

des sons perceptivement distincts. Le but est ainsi de faire en sorte, autant que possible, que chaque variable de conception ait un effet perceptible sur le son produit. Ces choix ont été faits par essai-erreur, en manipulant les contrôles du synthétiseur dans le logiciel Puredata. Pour les variables quantitatives, cette procédure a consisté à définir des valeurs minimales et maximales, puis des niveaux dans ces plages de valeurs. La structure générale du synthétiseur permet de générer des sons très différents à partir de variables de contrôle qui restent interprétables. Cela est notamment possible grâce au mélange entre plusieurs méthodes de synthèse, que permet l'approche modulaire. La figure 5.2 montre 12 exemples de sons générés par le synthétiseur<sup>4</sup>. Nous pouvons y observer la diversité des sons qui peuvent être produits. Les modulations de fréquence et d'amplitude, ainsi que la présence ou non de bruit ou de contenu harmonique sont clairement visibles.

**Synthèse du fond sonore** Le fond sonore est quant à lui généré à partir d'un enregistrement de son de roulage dans un habitacle, en utilisant une méthode de synthèse granulaire (ROADS 2004). Celle-ci génère un son en combinant de très courts échantillons d'un signal source. Elle permet de générer un son non répétitif mais dont la texture est stable dans le temps. La figure 5.3 montre le spectrogramme d'un extrait de 10 secondes du fond sonore.

## Plan d'expérience

Nous souhaitons construire un modèle qui nous servira à identifier les modalités des variables de conception à éliminer, à partir d'évaluations perceptives récoltées lors de la pré-expérience. Nous construisons donc un plan d'expériences pour cette dernière, car il n'est pas raisonnable de faire évaluer l'ensemble des sons du plan complet par les sujets. La construction du plan d'expériences consiste à définir un tableau  $X$ , dont chaque ligne  $n \in \{1, \dots, N\}$  correspond à une instanciation des 19 variables de conception. En première approche, nous souhaitons faire un plan d'expériences en blocs incomplets équilibrés (BIE), c'est-à-dire un plan dans lequel tous les sujets n'évaluent pas les mêmes sons. D'après PAGÈS et PÉRINEL (2007), lorsqu'un BIE n'existe pas pour une configuration donnée (dans notre cas, une configuration est caractérisée par le nombre de sujets, le nombre de sons au total et le nombre de sons évalués par chaque sujet), une stratégie consiste à construire un plan à partir du BIE le plus proche, en ajoutant des sujets ou

---

4. Les sons correspondants et le fond sonore peuvent être écoutés à l'adresse <https://souaille.github.io/conception/#ch5>

TABLE 5.1 – Description des variables de conception manipulées par l'IGA.

Variable	Description	Nombre de modalités
<b>Contrôle fréquentiel</b>		
$V_1 = E_f$	Enveloppe de fréquence	5
$V_2 = f_{syn}$	Fréquence cible du premier oscillateur	5
$V_3 = f_{bf}$	Fréquence cible du bruit filtré	3
$V_4 = f_{filt}$	Fréquence cible du filtre passe-bas	3
<b>Modulation de fréquence</b>		
$V_5 = r_{mf,syn}$	Fréquence de modulation de la fréquence de contrôle des oscillateurs	5
$V_6 = r_{\Delta,syn}$	Écart de fréquence autour de la fréquence de contrôle des oscillateurs	3
$V_7 = r_{mf,bf}$	Fréquence de modulation de la fréquence de contrôle du bruit filtré	3
$V_8 = r_{\Delta,bf}$	Écart de fréquence autour de la fréquence de contrôle du bruit filtré	3
<b>Oscillateurs sinusoïdaux</b>		
$V_9 = F$	Forme d'onde générée par la somme des oscillateurs	4
$V_{10} = p_{inh}$	Pourcentage d'inharmonicité des oscillateurs	3
<b>Filtres</b>		
$V_{11} = \zeta_{bf}$	Facteur d'atténuation du filtre passe-bande du bruit filtré	3
$V_{12} = \zeta_{filt}$	Facteur d'atténuation du filtre passe-bas de la section <i>FILTRAGE</i>	3
<b>Modulation d'amplitude</b>		
$V_{13} = f_{ma}$	Fréquence de modulation d'amplitude	4
$V_{14} = h_{syn}$	Taux de modulation de d'amplitude du signal de synthèse additive	3
$V_{15} = h_{bf}$	Taux de modulation d'amplitude du bruit filtré	3
<b>Enveloppe temporelle</b>		
$V_{16} = E_{a,syn}$	Enveloppe d'amplitude du signal de synthèse additive	5
$V_{17} = E_{a,bf}$	Enveloppe d'amplitude du signal de bruit filtré	5
<b>Mélangeur</b>		
$V_{18} = G$	Gain du signal de synthèse additive	5
$V_{19} = S$	Sélection de la source sonore	3

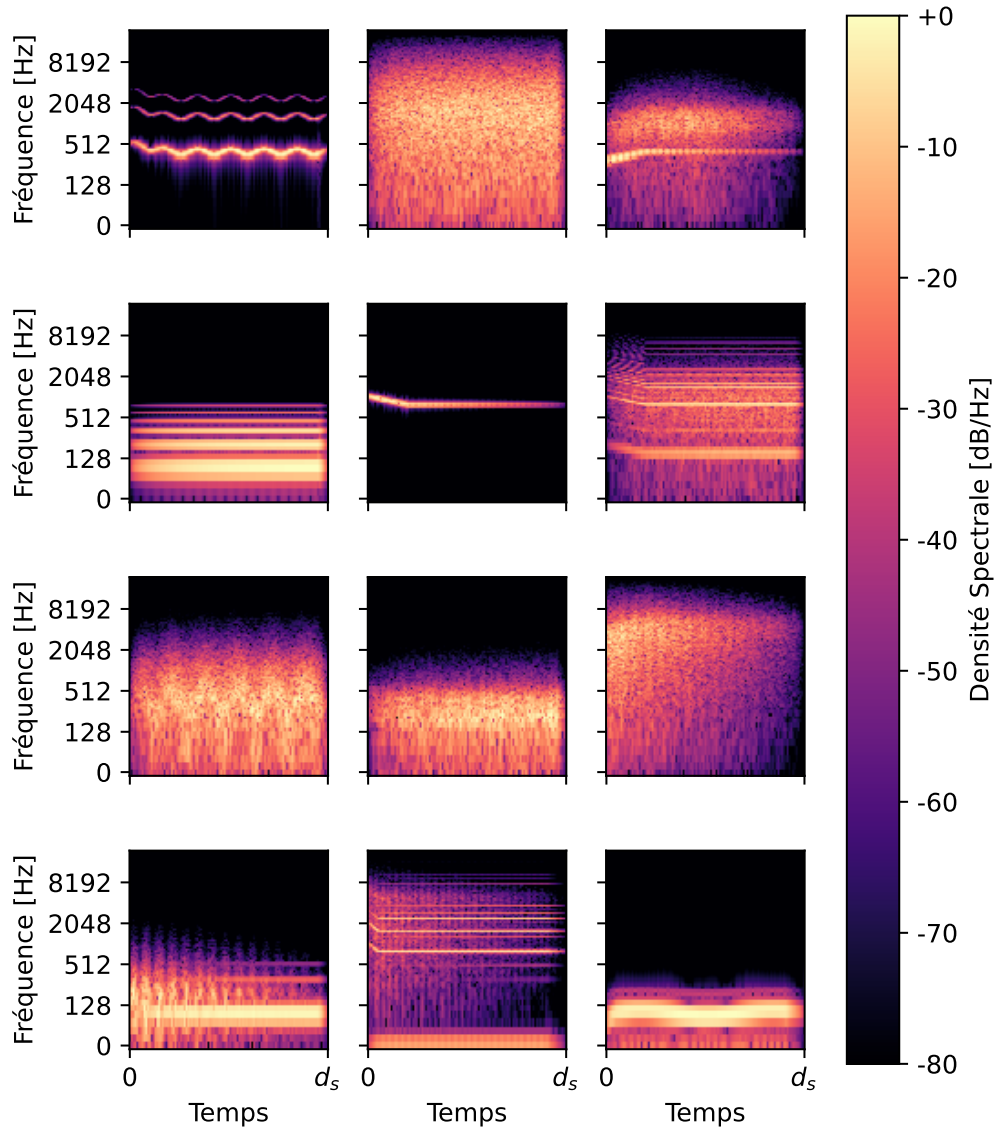


FIGURE 5.2 – Exemples de sons générés par le synthétiseur. 12 sons ont été générés via un échantillonnage par hypercube latin. L'échelle de chaque spectrogramme est normalisée par sa valeur maximale.



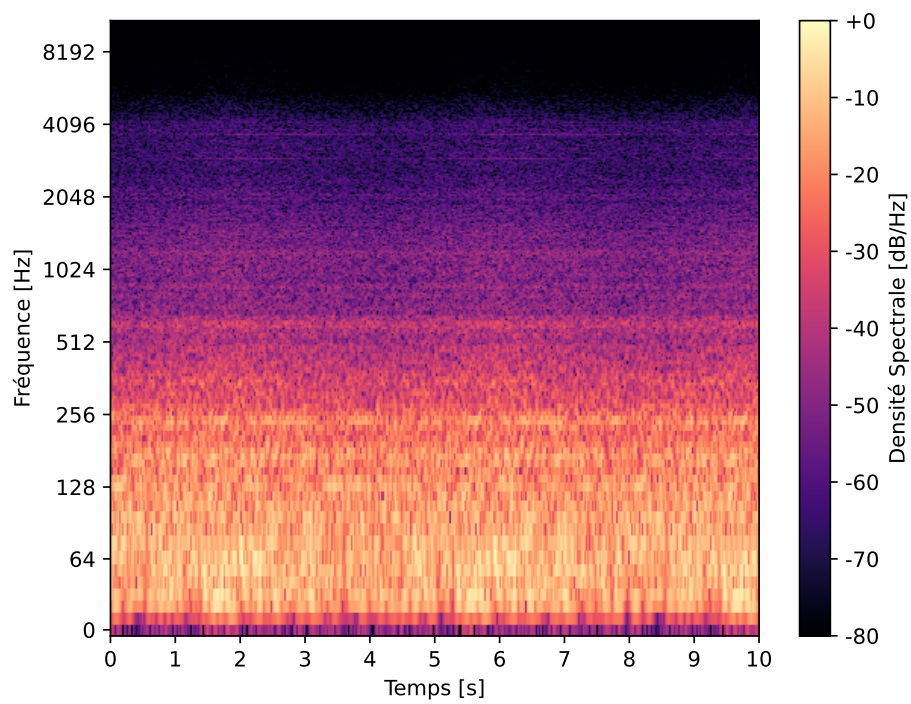


FIGURE 5.3 – Spectrogramme du fond sonore. Spectrogramme d'un extrait de 10 secondes du fond sonore. L'échelle du spectrogramme est normalisée par sa valeur maximale.

des évaluations tout en essayant de respecter au mieux les conditions d'équilibre du plan. Cependant, cette approche n'est envisageable que lorsque les valeurs des paramètres du plan sont faibles. Pour des plan volumineux, l'utilisation d'un algorithme de construction de plans optimaux permet d'obtenir un plan d'une efficacité acceptable, qui prend en compte l'ensemble des observations potentiellement disponibles. C'est pourquoi nous décidons plutôt d'utiliser un algorithme de construction de plans optimaux, en introduisant un facteur sujet.

**Tableau disjonctif complet** Lorsque les variables de conception sont qualitatives, le tableau  $X$  est un tableau disjonctif complet  $X = [X_1, \dots, X_I]$  (LEBART, MORINEAU et PIRON 1995), où les  $X_i$ , pour  $i \in \{1, \dots, 19\}$ , sont des sous-tableaux correspondant à chaque variable  $V_i$ . Nous incluons un facteur sujet dans notre modèle, dont les modalités sont codées dans le sous-tableau  $X_{20}$ . Chaque sous-tableau  $X_i$  possède  $M_i$  colonnes, soit une par modalité de la variable  $V_i$ . Dans le cas du sous-tableau  $X_{20}$ , le nombre de colonnes est égal au nombre de sujets pris en compte par le modèle. Chaque ligne du sous-tableau  $X_i$  possède un 1 dans la colonne correspondant à la modalité prise par la variable  $V_i$  pour cette observation, et des 0 ailleurs.

**Tableau de plein rang** Le tableau  $X$  n'est cependant pas de plein rang (la somme des termes des lignes est égal à 1 sur le tableau disjonctif), c'est pourquoi chaque sous-tableau est recodé en retirant la colonne  $M_i$ , après l'avoir retranchée aux autres colonnes. Chaque sous-tableau  $\hat{X}_i$  du nouveau tableau  $\hat{X} = [\hat{X}_1, \dots, \hat{X}_I]$  possède à présent  $M_i - 1$  colonnes (paramétrage centré, classique en ANOVA).

**Modèle linéaire** Pour notre analyse, nous supposons que l'évaluation d'intrusivité (la réponse du modèle, notée  $y$ ) est une fonction linéaire des variables de conception et qu'il n'y a pas d'interaction (modèle ANOVA à plusieurs facteurs). L'expression de ce modèle sous forme matricielle est :

$$y = L\delta + \epsilon \tag{5.1}$$

avec :

- $L = [1, \hat{X}]$  : tableau des variables explicatives recodé, avec une colonne de 1 à gauche,

- $\delta = [\mu, \alpha_{1,1}, \dots, \alpha_{1,M_1-1}, \dots, \alpha_{I,1}, \dots, \alpha_{I,M_I-1}]$  : vecteur contenant l'ensemble des coefficients correspondants aux  $M_i - 1$  modalités de chaque variable de conception  $V_i$  (y compris le facteur sujet), ainsi qu'un terme constant  $\mu$ ,
- $\epsilon$  : terme d'erreur,  $\epsilon \sim \mathcal{N}(0, \sigma^2)$ .

**Plan D-optimal** Afin de prédire l'intrusivité moyenne des sons produits par le synthétiseur en fonction des modalités des variables de conception, nous construisons un plan d'expériences selon le critère de D-optimalité. Ce dernier vise à minimiser le déterminant de la matrice de dispersion  $|L'L|^{-1}$  (ou, de manière équivalente, à maximiser le déterminant de la matrice d'information  $|L'L|$ ). Nous utilisons un algorithme d'échange pour construire  $L$ , qui consiste à :

1. Générer une matrice  $L$  aléatoirement (tirage aléatoire de  $N$  configurations du synthétiseurs dans le plan complet).
2. Remplacer au hasard une observation (c'est-à-dire une ligne de  $L$ ) par une nouvelle observation, générée aléatoirement.
3. Garder la nouvelle matrice  $L$  si la nouvelle observation permet d'augmenter  $|L'L|$ , garder l'ancienne matrice sinon.
4. Répéter les points 2 et 3 un nombre prédéfini de fois.

Nous souhaitons construire une plan d'expériences pour 24 sujets, où chaque sujet fait 20 observations. Pour cela, nous exécutons l'algorithme pour 35000 itérations et gardons le meilleur plan obtenu jusque là (celui pour lequel la valeur de  $|L'L|$  est la plus grande). La contrainte de l'équilibre du nombre d'observations par sujet est intégrée à l'algorithme.

**Prise en compte des variables inactives** Certaines variables ont pour effet de potentiellement rendre inactives d'autres variables en fonction de leur modalité. Par exemple, lorsque  $r_{\Delta, syn} = 0$  (modalité 1), il n'y a plus de modulation de fréquence dans la section *SYNTHÈSE ADDITIVE* et la variable  $r_{mf, syn}$  n'a alors plus d'effet. Nous prenons cela en compte au moment du calcul de  $|L'L|$  en laissant, pour chaque variable  $V_i$ , toutes les colonnes du sous-tableaux  $X_i$  à zéro, pour les observations où  $V_i$  est inactive à cause de la modalité prise par une autre variable. Ainsi, les configurations des variables de conceptions correspondant à des sons identiques sont codées de la même manière. Cela permet d'intégrer cette information dans le calcul du critère d'optimisation du plan d'expériences. La tableau 5.2 liste des conditions pour lesquelles certaines variables sont inactives.

TABLE 5.2 – Liste des conditions rendant certaines variables de conception inactives.

Condition	Variables inactives
$S = (1, 0)$	$f_{bf}, \zeta_{bf}, r_{mf,bf}, r_{\Delta,bf}, h_{bf}, G, E_{a,bf}$
$S = (0, 1)$	$p_{inh}, F, r_{mf,syn}, r_{\Delta,syn}, h_{syn}, G, E_{a,syn}$
$F = (1, 0, 0, 0, 0, 0)$	$p_{inh}$
$r_{\Delta,syn} = 0$	$r_{mf,syn}$
$r_{\Delta,bf} = 0$	$r_{mf,bf}$
$(h_{syn} = 0 \text{ ou } S = (0, 1)) \text{ et } (h_{bf} = 0 \text{ ou } S = (1, 0))$	$f_{ma}$

Pour construire le plan d'expériences, nous avons donc modifié un algorithme d'échanges pour tenir compte des variables inactives et de la contrainte sur le nombre de sons évalués par chaque sujet. Nous obtenons alors un plan d'expérience pour 24 sujets, où chaque sujet évalue 20 configurations du synthétiseur, qui maximise la D-optimalité selon le modèle linéaire postulé.

### Participant·e·s

24 étudiant·e·s de l'École Centrale de Nantes participent à cette expérience.

### Apparatus

L'expérience a lieu sur les ordinateurs des salles informatiques de l'École Centrale de Nantes. Les sons sont écoutés via les cartes sons internes des ordinateurs et des casques audio Beyerdynamics DT-990 (casques fermés, d'une impédance de 250 Ohm). Le niveau en sortie des casques est calibré en amont avec un sonomètre, de sorte que le niveau sonore total (bruit de fond + son de VA) soit en moyenne à 58dB(A). Ce niveau a été choisi de sorte à mettre en oeuvre une écoute confortable durant toute la durée de l'expérience, tout en permettant d'entendre au mieux les sons de VA.

### Protocole

**Interface** L'interface de test est similaire à celle utilisée pendant les expériences précédente (voir figure 5.4). Un bouton situé sur la gauche, permet d'écouter le son de VA. Un curseur horizontal permet d'évaluer l'intrusivité du son écouté, sur une échelle allant de « Non Intrusif » à « Très Intrusif ». Un bouton, sur lequel est écrit « Je n'entends pas le son » permet au sujet d'indiquer qu'il-elle n'entend pas de son de VA si le cas se présente (cas où le son serait totalement masqué par le bruit de fond). Aucun score d'intrusivité

n'est attribué aux sons qui n'ont pas été entendus. Une barre horizontale, située en haut à droite, indique au sujet la progression de l'expérience. Enfin, un bouton situé sous le curseur horizontal permet de charger la page suivante une fois les sons évalués.

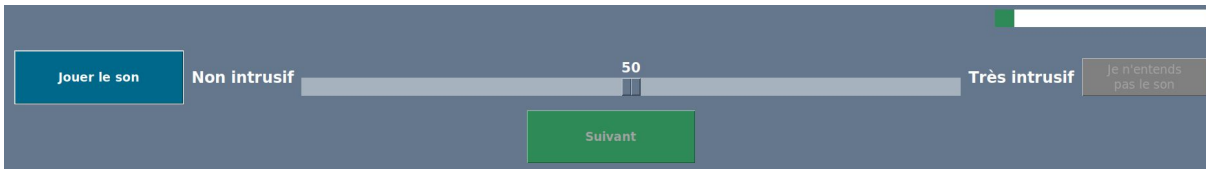


FIGURE 5.4 – Interface de test pour l'expérience 5. Le bouton « Jouer le son » permet d'écouter le son de VA à évaluer. Le curseur horizontal situé au centre permet d'évaluer l'intrusivité du son sur une échelle continue structurée allant de « Non intrusif » à « Très intrusif ». Cette évaluation est convertie en un score allant de 0 à 100. Un bouton « Je n'entends pas le son » permet au sujet d'indiquer qu'il ne lui est pas possible de noter l'intrusivité du son, car celui est trop faible par rapport au fond sonore. Un bouton « Suivant » permet de passer à l'évaluation du son suivant. Ce bouton n'est actif que si le curseur a été déplacé au moins une fois ou si le sujet a cliqué sur le bouton « Je n'entends pas le son ». Enfin, une barre de progression permet de savoir approximativement combien de sons il reste à évaluer.

**Procédure expérimentale** Deux pages de tutoriel sont données à lire au début de l'expérience et permettent au sujet de se familiariser avec l'interface et le contexte de l'expérience. Par ailleurs, deux exemples de sons sont donnés pour illustrer le concept d'intrusivité : un son peu intrusif et un son très intrusif<sup>5</sup>. Ces exemples nous ont été fournis par le designer sonore Andrea Cera. Le fond sonore commence dès la fin du tutoriel, et joue sans interruption jusqu'à la fin de l'expérience. Chaque sujet évalue 20 sons. L'expérience dure moins de 10 minutes pour chaque sujet.

## Analyse

Nous analysons les évaluations au travers d'un test de Fisher (LEBART, MORINEAU et PIRON 1995) dans un premier temps (carrés de type III). Celui-ci nous permet d'identifier les variables de conception pour lesquelles nous observons un effet significatif sur l'intrusivité du son généré. Un test post-hoc est effectué sur les variables pour lesquelles le test de Fisher montre un effet significatif. Ce test post-hoc (test de Tukey) nous permet alors d'identifier, pour chaque variable, les paires de modalités pour lesquelles les sons correspondants ont une intrusivité significativement différente.

---

5. Les deux exemples de sons peuvent être écoutés à l'adresse <https://souaille.github.io/conception/#ch5>

Suite aux tests statistiques, nous éliminons les modalités pour lesquelles il existe au moins une autre modalité donnant des sons significativement moins intrusifs. Cette approche diffère de celle utilisée dans le chapitre 4, où le même nombre de modalités était éliminé pour toutes les variables de conception. Nous évitons ainsi d'éliminer arbitrairement des modalités, bien que cela puisse amener à éliminer un moins grand nombre de modalités par variable.

### 5.2.3 Résultats

À l'issue de l'expérience, nous obtenons un score d'intrusivité compris entre 0 et 100 pour 480 configurations différentes du synthétiseur, chacune caractérisée par 19 variables de conception. 24 sujets ont chacun évalué 20 sons parmi les 480. Le test de Fisher nous indique que sept variables de conception parmi 19 ont un effet significatif ( $p < 0,05$ ) sur l'intrusivité :

- L'enveloppe de fréquence  $E_f$ ,
- La fréquence cible du premier oscillateur  $f_{syn}$ ,
- L'écart de fréquence autour de la fréquence de contrôle des oscillateurs  $r_{\Delta, syn}$ ,
- Le taux de modulation d'amplitude du bruit filtré  $h_{bf}$ .
- Le gain du signal de synthèse additive  $G$ ,
- L'enveloppe d'amplitude du signal de synthèse additive  $E_{a, syn}$ ,
- La sélection de la source sonore  $S$ .

Nous procédons donc à un test de Tukey pour chacune de ces variables, ce qui nous permet d'identifier les modalités à éliminer. Les résultats sont illustrés par la figure 5.5. Concernant l'enveloppe de fréquence  $E_f$ , le test post-hoc n'a permis d'éliminer aucune modalité, bien que le test de Fisher ait montré un effet significatif de cette variable. Pour la fréquence cible du premier oscillateur  $f_{syn}$ , les modalités éliminées sont celles correspondant aux fréquences les plus élevées (modalités 3 à 5). Concernant la modulation de fréquence de la section *SYNTHÈSE ADDITIVE*, la première modalité de  $r_{\Delta, syn}$ , correspondant à une fréquence faiblement modulée, est éliminée. Pour la modulation d'amplitude du bruit filtré, seule la modalité de  $h_{bf}$  correspondant à une absence de modulation est gardée (la modalité 1). Les modalités éliminées pour le gain  $G$  sont celles pour lesquelles le signal  $y_{syn}(t)$  est augmenté de 6dB (modalité 4) ou 12dB (modalité 5), par rapport au signal  $y_{bf}(t)$ . Une seule modalité de l'enveloppe d'amplitude du signal de synthèse additive  $E_{a, syn}$  est éliminée (modalité 4), correspondant à une enveloppe avec une courte attaque, un long entretien et une courte extinction (voir figure B.16). Enfin, pour

la variable  $S$  qui correspond au choix des sources présentes dans le signal total, seule la première modalité est éliminée, pour laquelle il n’y a pas de bruit filtré dans le signal.

### 5.2.4 Discussion

Contrairement à la méthode suivie dans le chapitre 4, nous avons choisi ici de ne pas éliminer un même nombre de modalités pour chaque variable. Une analyse statistique nous a permis d’identifier des variables de conceptions pour lesquelles certaines modalités donnent des sons significativement plus intrusifs que d’autres modalités. Ainsi, nous avons pu éliminer des modalités pour six variables sur les 19 variables de conception. Pour les variables  $f_{syn}$ ,  $G$  et  $S$ , les modalités restantes tendent à faire ressembler le timbre du son généré à celui du fond sonore. En effet, ce dernier est semblable à du bruit, dont la majorité de l’énergie est située dans les basses fréquences (voir figure 5.3). Or, les modalités restantes des variables  $G$  et  $S$  produisent des sons qui contiennent forcément du bruit filtré, et pour lesquels la composante de synthèse additive, si elle est présente, est à un niveau (en dB(A)) moins élevé que la composante bruitée. Les trois modalités de la variable  $f_{syn}$  correspondant aux fréquences les plus élevées (modalités 3 à 5) sont de même éliminées. Cette observation rejoint le point soulevé par CERA et MISDARIIS 2021, pour qui l’émergence du son étudié par rapport au fond sonore contribue à son intrusivité. Dans MISDARIIS, CERA et RODRIGUEZ 2019 et CERA et MISDARIIS 2021, les auteurs considèrent par ailleurs que des sons de haute fréquence, mais avec un niveau sonore bas, pourraient être peu intrusifs. Les résultats de cette expérience ne permettent malheureusement pas de confirmer cette hypothèse. Cela aurait peut-être pu être exploré en considérant le niveau sonore comme une variable de conception supplémentaire.

L’enveloppe éliminée pour la partie de synthèse additive (modalité 4 de la variable  $E_{a,syn}$ ) est celle pour laquelle le niveau sonore reste à sa valeur maximale pour la plus longue durée (il s’agit de la durée de la phase d’entretien). Perceptivement, cette modalité correspond donc aux sons pour lesquels la durée de la partie harmonique est la plus longue. Ce résultat semble aller dans le sens des préconisations faites dans MISDARIIS, CERA et RODRIGUEZ 2019 et CERA et MISDARIIS 2021, plutôt que dans celui des auteurs FAGERLÖNN, LARSSON et MACULEWICZ 2020, qui privilégient des sons longs et ressemblant à des sons de moteur (donc harmoniques, plutôt que bruités).

Les résultats pour les deux variables restantes  $r_{\Delta,syn}$  et  $h_{bf}$  sont plus complexes à interpréter. En effet, pour la variable  $r_{\Delta,syn}$ , qui contrôle l’intensité de la modulation de fréquence pour la section *SYNTHÈSE ADDITIVE*, les modalités conservées correspondent

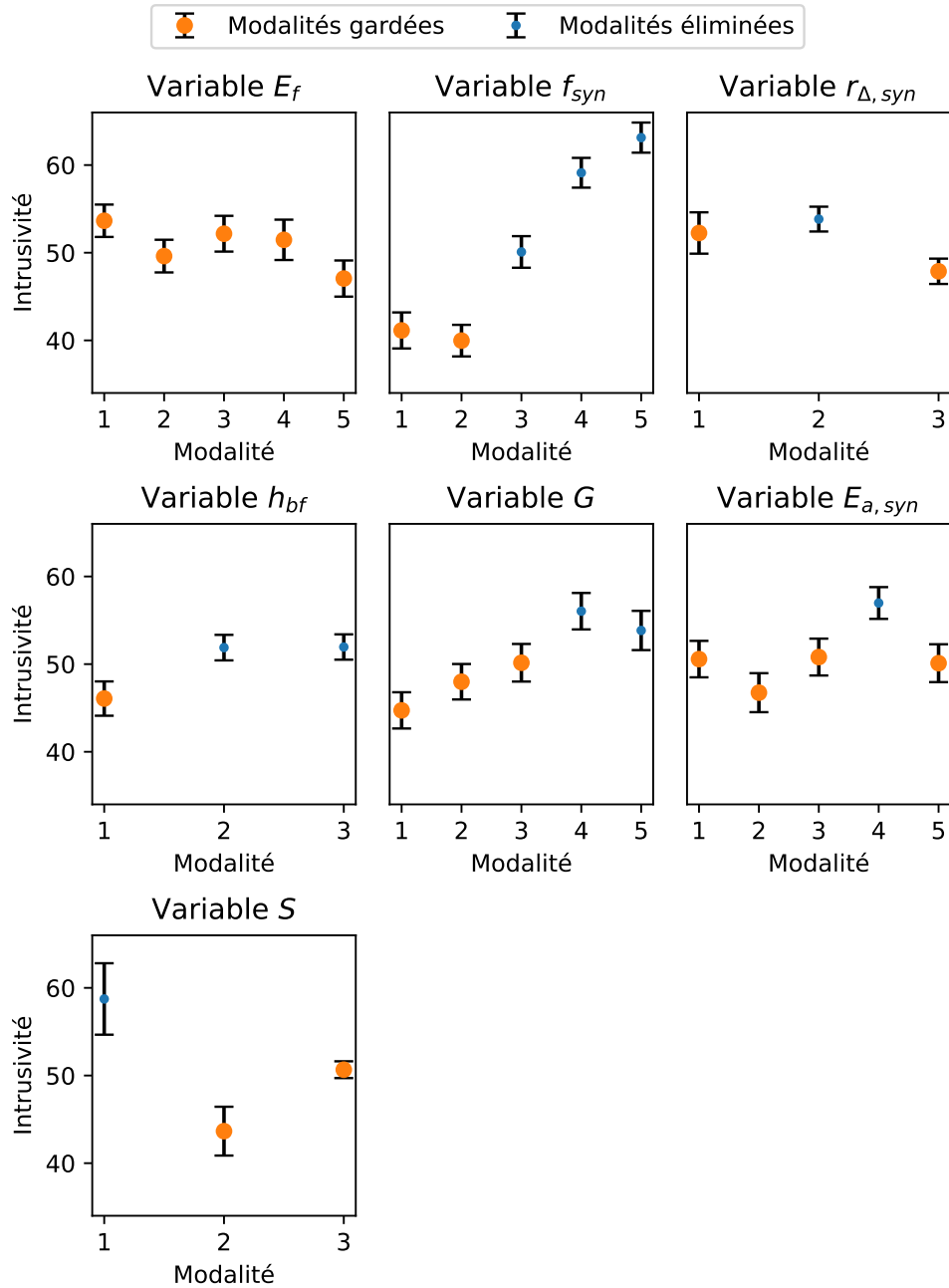


FIGURE 5.5 – Sélection de modalités. Les valeurs moyennes d'intrusivité sont affichées pour chacune des modalités des variables de conception pour lesquelles un effet significatif a été observé. Les modalités éliminées sont celles pour lesquelles le test post-hoc nous indique qu'il existe au moins une autre modalité pour laquelle l'intrusivité est significativement ( $p < 0,05$ ) plus faible, en moyenne. Les barres d'erreur correspondent à l'erreur type de la moyenne d'intrusivité pour chaque modalité.



à une absence de modulation (modalité 1) et à la modulation la plus forte (modalité 3). Pour la variable  $h_{bf}$ , qui contrôle le taux de modulation d’amplitude de la section *BRUIT FILTRÉ*, la modalité conservée (modalité 1) correspond à une absence de modulation. La modulation de fréquence et la modulation d’amplitude jouent toutes les deux sur la rugosité d’un son (ZWICKER et FASTL 2013). CERA et MISDARIIS 2021 mentionnent que la rugosité est susceptible de jouer sur l’intrusivité d’un son (une forte rugosité correspondant à une plus forte intrusivité). Les résultats pour la variable  $h_{bf}$  semblent confirmer cela. Cependant, la rugosité due à la modulation en fréquence d’un son dépend à la fois de l’indice de modulation et de la fréquence de modulation. Il est donc difficile de dire si les modalités de  $r_{\Delta, syn}$  qui sont éliminées correspondent en général, ou non, à des sons plus rugueux.

Enfin, nous pouvons remarquer que certaines des stratégies adoptées pendant la phase de conception du synthétiseur se rapprochent des méthodes de réduction d’espace de conception mentionnées dans SHAN et WANG 2010. De manière empirique, nous avons en effet été amené·e·s à combiner certaines variables (*mapping*) et à choisir des variables de conception (*screening*) et des modalités (*space reduction*) qui semblent avoir un effet sur le son produit. Ces choix ont été faits de manière informelle en se basant sur l’expertise et l’intuition des auteur·rice·s, à l’écoute des sons générés par le synthétiseur.

### 5.3 Expérience 6 : validation de la méthode de réduction d’espace

L’expérience précédente (expérience 5) nous a permis d’éliminer plusieurs modalités pour six des 19 variables de conception du problème. Nous avons pour cela utilisé une approche basée sur des tests statistiques, afin de limiter le risque d’élimination des « mauvaises » modalités. Cependant, comme cela a été soulevé dans le chapitre 4, notre méthode se base sur l’approximation de la réponse à minimiser par un modèle linéaire sans interaction. Cette approche ne nous garantit donc pas d’avoir préservé les sons les moins intrusifs du plan d’expériences complet. De plus, notre approche dépend de l’accord entre les sujets sur la notion d’intrusivité. C’est pourquoi une dernière expérience (expérience 6) est nécessaire, afin de confirmer ou non l’intérêt de la méthode proposée, comme étape préalable à une expérience d’optimisation perceptive interactive.

### 5.3.1 Objectif

Cette expérience a pour but de savoir si la réduction de l'espace de conception effectuée à l'issue de l'expérience 5 permet ou non de trouver des sons moins intrusifs pendant un processus d'optimisation perceptive interactive. Dans ce but, nous voulons comparer l'intrusivité des sons trouvés avec un Algorithme Générique Interactif (IGA), dans deux cas de figure, comme illustré sur la figure 5.6 :

1. Exploration de l'espace de conception complet (**EC**, voir chapitre 4) .
2. Exploration de l'espace de conception réduit grâce à des données perceptives (Approche Sujet (**AS**, voir chapitre 4).

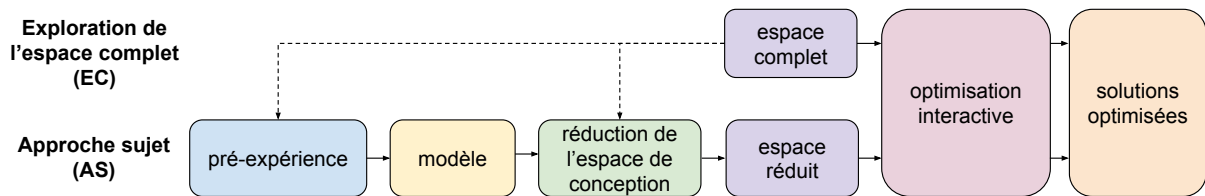


FIGURE 5.6 – Illustration du protocole de l'expérience 6.

### 5.3.2 Matériels et Méthodes

#### Stimuli

Les sons de VA et le fond sonore sont les mêmes que pour l'expérience 5 (voir section 5.2.2).

#### Plan d'expériences

Tous les sujets participent aux deux conditions expérimentales (**EC** et **AS**).

#### Participant·e·s

18 nouveaux sujets participent à cette expérience. 14 d'entre eux sont des élèves de l'École Centrale de Nantes. Les quatre personnes restantes font partie du personnel du Laboratoire des Sciences du Numérique de Nantes (LS2N).

## Apparatus

Le matériel et la procédure de calibration sont identiques à ceux de l'expérience 5 (voir section 5.2.2).

## Protocole

**Implémentation de l'IGA** L'IGA utilisé est le même que celui utilisé dans le chapitre 4, mais il est paramétré différemment. Il s'agit d'un IGA mono-objectif, avec un taux d'élitisme de  $p_{el} = 0,2$ , une sélection par tournoi binaire, un croisement uniforme et un taux de mutation  $p_{mu} = 0,2$ . Le nombre d'individus par génération est  $N = 10$ . Comme pour les expériences 1 et 4, ces valeurs ont été choisies empiriquement, en testant plusieurs fois la procédure d'optimisation. Les valeurs retenues sont celles qui semblaient assurer une évolution rapide des sons proposés vers des sons peu intrusifs, dans le nombre de générations choisi. Nous souhaitons pouvoir comparer les performances des méthodes **EC** et **AS**, pour un coût expérimental équivalent. Au cours de la pré-expérience que nous avons effectuée dans le cadre de la méthode **AS** (expérience 5), chaque sujet a fourni 20 évaluations. C'est pourquoi l'IGA dure 20 évaluations (soit deux générations) de moins pour l'approche **AS** que pour l'approche **EC**. Le nombre de générations est donc de sept pour l'exploration de l'espace complet (**EC**) et cinq pour l'approche sujet (**AS**). Les sons non entendus (bouton « Je n'entends pas le son », voir figure 5.4) ne sont pas pris en compte par l'IGA pour la sélection des élites, ni pour les tournois binaires de sélection des chromosomes parents.

**Interface** L'interface de test est identique à celle utilisée pour l'expérience 5 (voir figure 5.4, section 5.2.2). Cependant, le nombre de sons par page est cette fois-ci de 10 (correspondant au nombre de sons par génération de l'IGA).

**Procédure expérimentale** Les mêmes deux pages de tutoriels qu'à l'expérience 5.2 sont présentées en début d'expérience, afin de se familiariser avec l'interface, la tâche d'évaluation et la notion d'intrusivité. Chaque sujet explore simultanément l'espace de conception complet et l'espace réduit, sans être informé qu'il s'agit d'une expérience d'optimisation. Pour cela, les générations correspondant à chaque méthode sont présentées alternativement. La figure 5.7 illustre ce processus. Chaque sujet évalue en tout  $10 \times (7 + 5) = 120$  sons au cours de l'expérience 6. Celle-ci dure 16 minutes en moyenne.

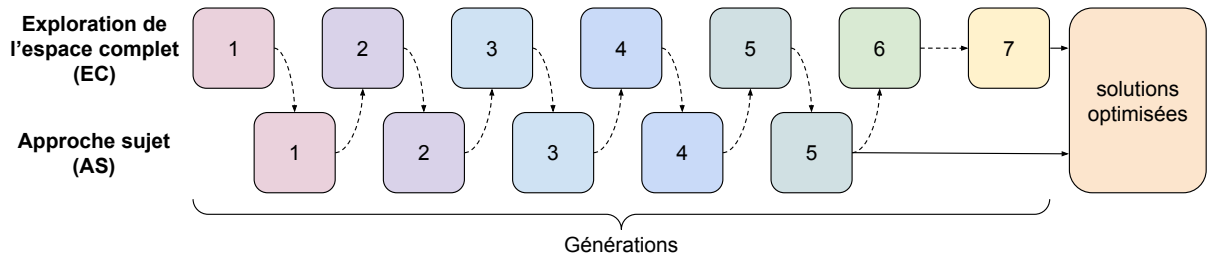


FIGURE 5.7 – Illustration de l’alternance des IGAs pendant l’expérience 6. La procédure commence par l’évaluation de la première génération de sons de la méthode **EC**. Ensuite, le sujet évalue la première génération de la méthode **AS**. Puis il-elle évalue la seconde génération de la méthode **EC**, puis la seconde génération de la méthode **AS**, et ainsi de suite.

## Analyse

Deux métriques sont calculées pour chaque sujet et chaque génération :

- L’intrusivité moyenne de tous les sons de la génération.
- L’intrusivité du meilleur son (l’intrusivité la plus faible).

Ces métriques sont comparées pour les cas **EC** et **AS** via un test statistique. Dans un premier temps, un test de Shapiro-Wilk nous permet de tester la normalité des données. En fonction des résultats de ce premier test, un test de Student apparié unilatéral ou un test des rangs signés de Wilcoxon unilatéral est utilisé pour comparer les valeurs moyennes des métriques calculées, en fonction de la méthode (**EC** ou **AS**). Par ailleurs, les distributions des modalités des meilleurs sons trouvés pour chaque méthode sont comparées pour chaque variable de conception. Le lien entre la méthode d’optimisation et la distribution des modalités est testé via un test du  $\chi^2$  d’indépendance. Pour chaque méthode d’optimisation, un test du  $\chi^2$  d’adéquation permet enfin, pour chaque variable, de tester si la distribution des modalités est significativement différente d’une distribution uniforme.

### 5.3.3 Résultats

#### Intrusivité moyenne

Dans un premier temps, nous nous intéressons à l’intrusivité moyenne des sons évalués à chaque génération. Pour cela, nous calculons, pour chaque sujet, l’intrusivité moyenne à chaque génération, puis nous calculons la moyenne de cette valeur sur l’ensemble des sujets. Les résultats sont présentés sur la figure 5.8. Nous y observons qu’en moyenne, les

sons trouvés avec la méthode sujet (**AS**) sont moins intrusifs que ceux trouvés en explorant l'espace complet (**EC**). Afin de savoir si cette différence est significative, nous testons tout d'abord la normalité des distributions desquelles les données de la figure sont issues. Un test de Shapiro-Wilk nous indique que l'hypothèse de normalité doit être rejetée au seuil de  $\alpha = 5\%$  pour chacune des générations de la méthode **AS**. Nous comparons donc les méthodes à chaque génération via un test des rangs signés de Wilcoxon unilatéral, car chaque sujet a évalué les deux conditions expérimentales, ce qui veut dire que nos données sont appariées. Les résultats des tests de Wilcoxon sont également indiqués sur la figure 5.8. Nous y observons que la méthode **AS** permet, en moyenne, d'explorer des sons significativement ( $p < 0,05$ ) moins intrusifs que la méthode **EC**, et ce pour l'ensemble des générations considérées. Nous observons par ailleurs une diminution de l'intrusivité moyenne pour des deux approches, au cours des générations.

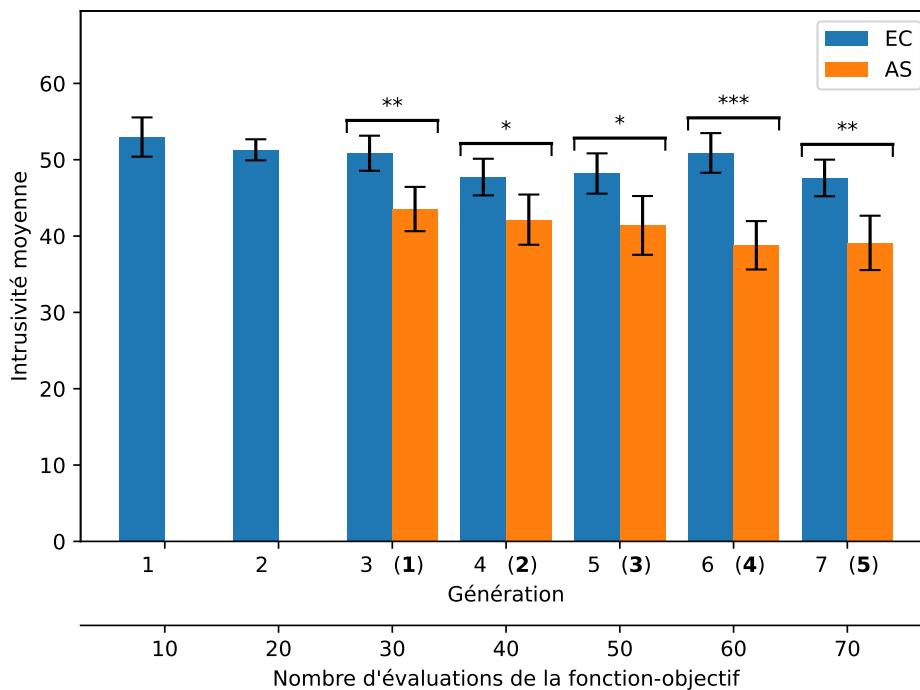


FIGURE 5.8 – Intrusivité moyenne des sons pendant l'expérience 6. Les valeurs correspondant à l'approche sujet (AS) commencent à la troisième génération de celles correspondant à l'exploration de l'espace complet (EC), pour prendre en compte les évaluations nécessaires à la pré-expérience dans le premier cas. Les nombres en gras entre parenthèses le long de l'axe des abscisses correspondent aux générations de l'approche AS. Les résultats des tests statistiques sur la différence entre les deux méthodes sont indiqués au dessus de chaque paire de données (\*\* :  $p < 0,01$ , \* :  $p < 0,05$ ). Les barres d'erreur correspondent à l'erreur-type de la moyenne.

### Intrusivité des meilleurs sons

Nous calculons ensuite l'intrusivité moyenne du meilleur son obtenu à chaque génération par chaque sujet, pour chacune des deux approches testées. Les valeurs correspondantes sont présentées dans la figure 5.9. De même que pour l'intrusivité moyenne, les différences entre les deux méthodes sont testées via un test des rangs signés de Wilcoxon unilatéral. Nous observons que l'intrusivité des meilleurs sons est en moyenne plus basse pour l'approche AS. Cependant, la différence entre les deux méthodes n'est significative ( $p < 0,05$ ) que pour la troisième génération de l'approche EC (la première génération de l'approche AS).

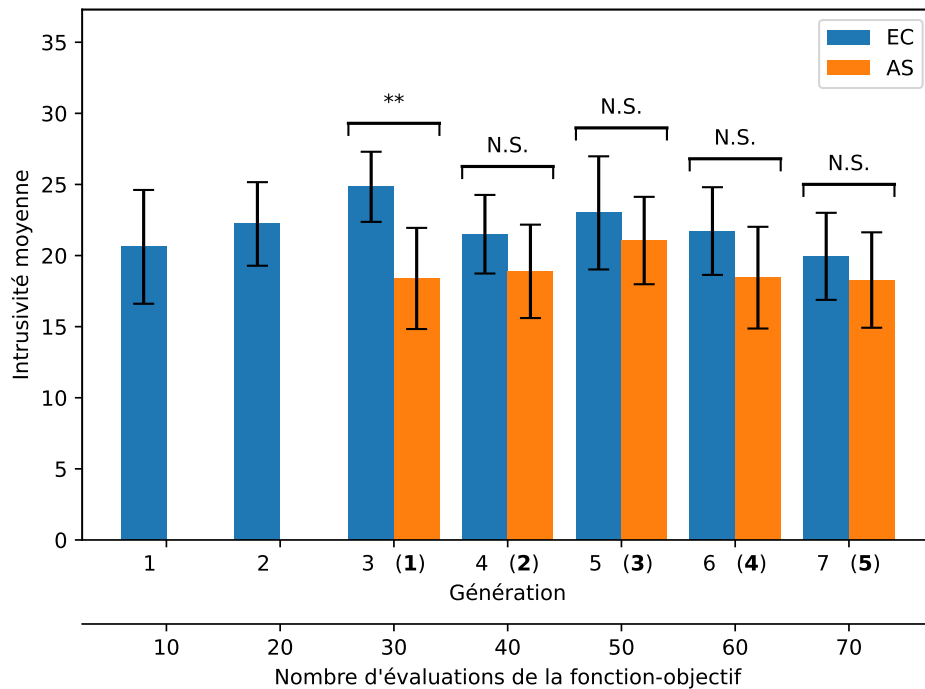


FIGURE 5.9 – Intrusivité moyenne des meilleurs sons pendant l'expérience 6. Les valeurs correspondant à l'approche sujet (**AS**) commencent à la troisième génération de celles correspondant à l'exploration de l'espace complet (**EC**), pour prendre en compte les évaluations nécessaires à la pré-expérience dans le premier cas. Les nombres en gras entre parenthèses le long de l'axe des abscisses correspondent aux générations de l'approche **AS**. Les résultats des tests statistiques sur la différence entre les deux méthodes sont indiqués au dessus de chaque paire de données (\*\*\*) :  $p < 0,001$ , \*\* :  $p < 0,01$ , \* :  $p < 0,05$ ). Les barres d'erreur correspondent à l'erreur-type de la moyenne.

## Analyse des meilleurs sons

Les figures 5.10 et 5.11 montrent les spectrogrammes des sons les moins intrusifs trouvés par chaque sujet, pour l'exploration de l'espace complet (**EC**) et l'approche sujet (**AS**), respectivement <sup>6</sup>. Ces sons sont ceux ayant obtenu la plus basse évaluation d'intrusivité à la dernière génération. Les meilleurs sons trouvés pour les deux méthodes semblent généralement être essentiellement composés de bruit filtré, avec peu d'éléments harmoniques. Pour l'exploration de l'espace complet (**EC**), nous pouvons cependant observer plusieurs sons (pour les sujets 4, 12 et 14 notamment) pour lesquels le bruit filtré est modulé en amplitude (stries verticales dans le spectrogramme). Or, à partir des résultats de l'expérience 5, seule une seule modalité de la variable  $h_{bf}$  a été gardée, correspondant à une absence de modulation d'amplitude pour le bruit filtré. Les sons avec un bruit filtré modulé en amplitude n'ont donc pas été explorés avec l'approche sujet (**AS**). De même, le meilleur son obtenu par le sujet 18 avec l'approche **EC** ne contient aucun bruit filtré, alors que cela est impossible avec l'approche **AS**.

La figure 5.12 présente les distributions des modalités des meilleurs sons trouvés avec l'IGA, pour les deux approches. Pour chaque variable de conception, un test du  $\chi^2$  d'indépendance est effectué, afin de savoir si la distribution des modalités pour les meilleurs sons est indépendante de l'approche utilisée pour l'optimisation (**EC** ou **AS**). Le test montre un lien entre la distribution des modalités et l'approche utilisée pour les variables  $h_{bf}$  ( $p < 0,001$ ),  $G$  ( $p < 0,05$ ) et  $S$  ( $p < 0,05$ ). Ces trois variables font partie de celles pour lesquelles nous avons éliminé des modalités suite à l'expérience 5. Concernant le taux de modulation d'amplitude du bruit filtré  $h_{bf}$ , les modalités des meilleurs sons trouvés avec l'approche **EC** semblent réparties de manière plutôt uniforme, alors que la procédure de réduction nous a amené à ne garder que la première modalité. Concernant le gain  $G$ , la répartition semble uniforme parmi les trois modalités gardées pour l'approche **AS** (correspondant aux gains les plus faibles, voir table B.7). Elle est également uniforme sur les cinq modalités explorées avec l'approche **EC**. Enfin, pour la variable  $S$ , qui gère la sélection des sources sonores (voir table B.7), la distribution semble uniforme sur les deux modalités gardées pour l'approche **AS**, tandis que pour l'approche **EC** la seconde modalité (correspondant à une absence de synthèse additive) est bien plus représentée.

Par ailleurs, pour chaque approche et chaque variable de conception, un test du  $\chi^2$  d'adéquation est effectué pour savoir si la distribution des modalités est différente d'une distribution uniforme. Pour l'approche **AS**, nous prenons en compte que certaines mo-

---

6. Les sons correspondants peuvent être écoutés à l'adresse <https://souaille.github.io/conception/#sec53>

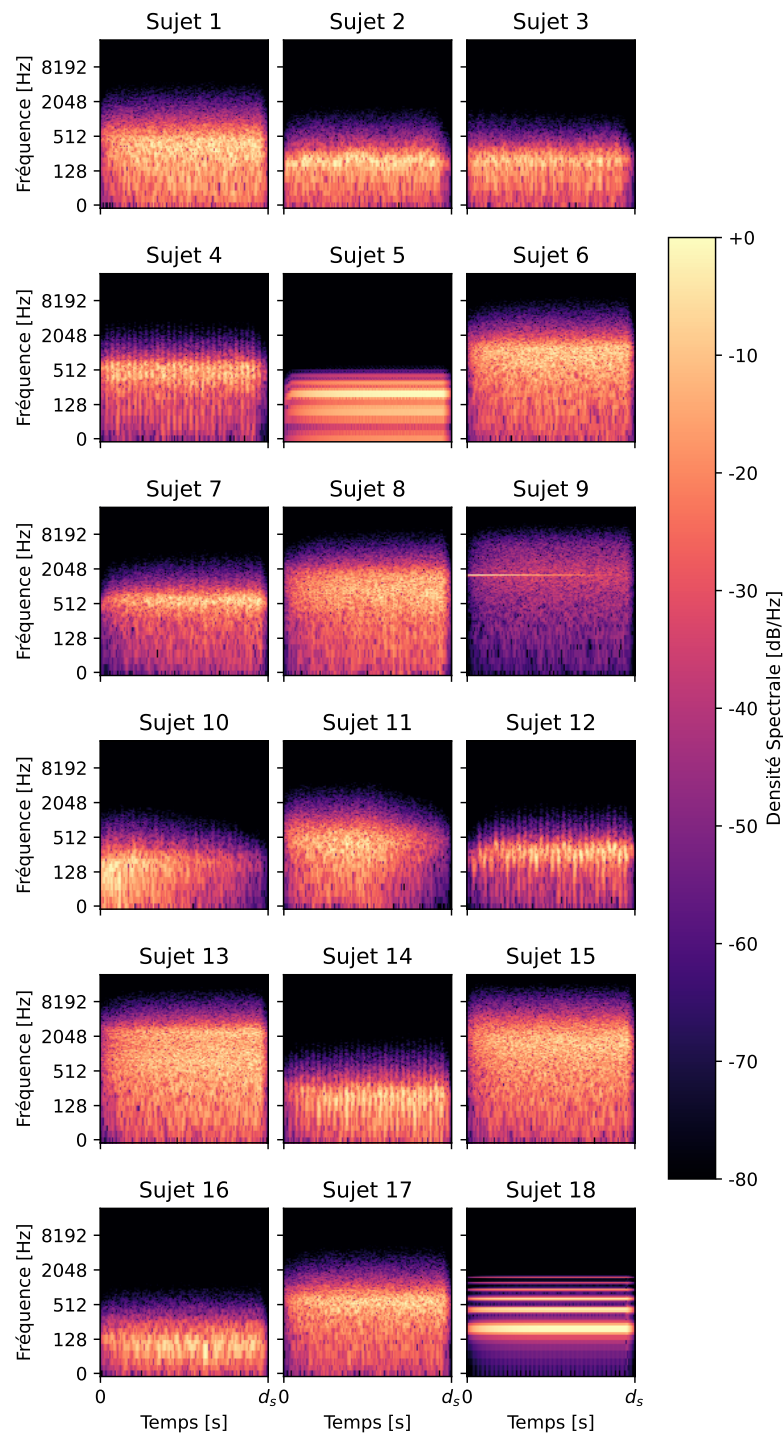


FIGURE 5.10 – Spectrogrammes des sons les moins intrusifs trouvés en explorant l'espace complet (EC). Il s'agit, pour chaque sujet, du son de la dernière génération dont l'intrusivité est la plus faible.



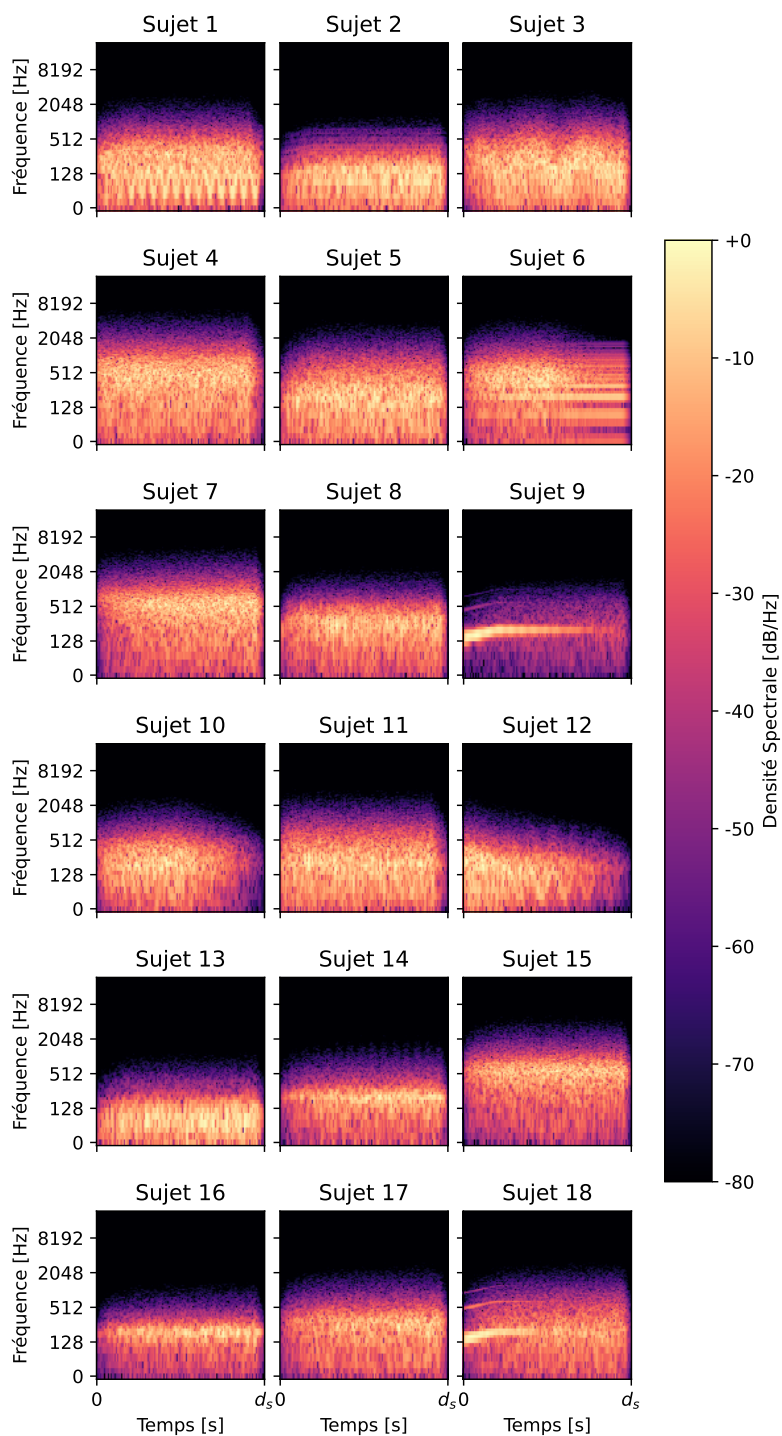


FIGURE 5.11 – Spectrogrammes des sons les moins intrusifs trouvés pour l’approche sujet (AS). Il s’agit, pour chaque sujet, du son de la dernière génération dont l’intrusivité est la plus faible.

dalités ont été éliminées. Pour l'exploration de l'espace complet (**EC**), les modalités des variables  $F$  ( $p < 0,05$ ),  $h_{syn}$  ( $p < 0,05$ ) et  $S$  ( $p < 0,001$ ) ont une distribution significativement différente d'une distribution uniforme. Pour la variable  $S$ , cela vient conforter l'observation effectuée précédemment, à savoir que les meilleurs sons trouvés par l'approche **EC** sont, pour la plupart, uniquement composés de bruit filtré (modalité 2 de la variable  $S$ ). Concernant la variable  $F$ , qui contrôle la forme d'onde générée par synthèse additive, nous observons que la modalité 4 est sur-représentée. Celle-ci correspond à une onde sinusoïdale, où seul le premier des six oscillateurs a une amplitude non nulle (voir table B.3 en annexe). Enfin, pour la variable  $h_{syn}$ , il semble, au regard de la figure 5.12, que les meilleurs sons obtenus avec l'approche **EC** tendent à avoir une légère modulation d'amplitude pour la section *SYNTHÈSE ADDITIVE* (modalité 2, voir table B.5 en annexe). Le test du  $\chi^2$  d'adéquation sur les modalités des meilleurs sons trouvés avec l'approche sujet (**AS**) nous indique que seule la variable  $p_{inh}$  ( $p < 0,01$ ) possède une distribution différente d'une distribution uniforme. En effet, bien que nous n'ayons éliminé aucune modalité pour cette variable, aucun des meilleurs sons obtenus à la fin du processus d'optimisation ne présente la seconde modalité (correspondant à une faible inharmonicité, voir table B.3 en annexe).

### 5.3.4 Discussion

Dans un premier temps, nous remarquons que la méthode de réduction de l'espace de conception a effectivement permis de réduire l'intrusivité moyenne des sons explorés, comme le montre la figure 5.8. Grâce à l'expérience 5, nous avons donc pu extraire des informations pertinentes pour notre tâche, avec un coût expérimental relativement faible, celle-ci durant moins de 10 minutes pour chaque sujet. De même, la méthode de sélection et d'élimination de modalités proposée permet une réduction de l'espace de conception qui reste adaptée à un nouveau panel de sujets, les panels des expériences 5 et 6 étant différents l'un de l'autre. Deuxièmement, les résultats observés concernant les meilleurs sons évalués sont similaires à ceux obtenus lors des simulations effectuées pendant l'expérience 4. En effet, réduire la taille de l'espace de conception avec la méthode proposée ne permet d'obtenir de meilleurs sons qu'au début du processus d'optimisation. À partir de la génération 4 pour l'exploration de l'espace complet (**EC**) (génération 2 pour l'approche sujet (**AS**)), les meilleures évaluations fournies par les sujets ne sont pas significativement différentes entre les deux méthodes (**EC** et **AS**). La méthode **EC** ne permet cependant pas de trouver de meilleurs sons que la méthode **AS** dans le nombre de générations imparti.

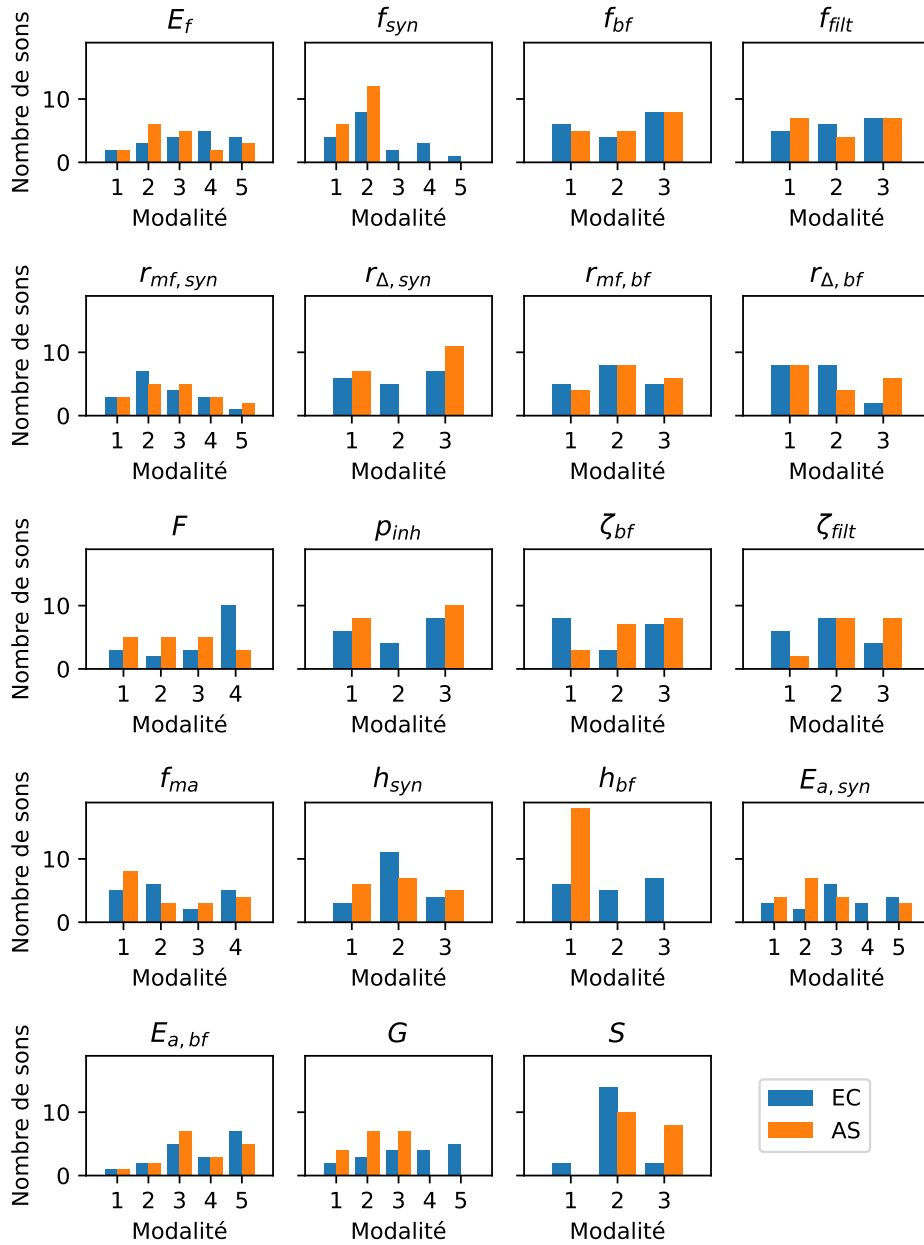


FIGURE 5.12 – Distributions des modalités des meilleurs sons pour chaque variable de conception, pour l'exploration de l'espace complet (EC) et l'approche sujet (AS).

Il semble donc pertinent d'approcher un problème d'optimisation perceptive interactive via un protocole expérimental hybride, où l'algorithme d'optimisation est seulement utilisé dans un deuxième temps, après une pré-expérience plus classique visant à dégrossir le problème. Cette piste est notamment mentionnée par PETIOT, DENJEAN et al. 2020, qui proposent cependant d'utiliser la pré-expérience non pas pour éliminer des modalités (*Space reduction*, cf. chapitre 4), mais pour éliminer les variables peu influentes (*Screening*). Cette stratégie peut donc être adaptée en fonction du type de problème et de la méthode d'optimisation choisie. En effet, comme nous l'avons mentionné dans le chapitre 4, il ne nous a pas semblé que le *screening* soit l'approche la mieux adaptée pour une optimisation avec un IGA, mais il est vraisemblable qu'elle puisse faciliter le processus d'optimisation pour d'autres algorithmes.

Dans un second temps, l'analyse des distributions des modalités pour les meilleurs sons obtenus avec les méthodes **EC** et **AS** montre que les meilleurs sont trouvés en explorant l'espace complet ne suivent pas nécessairement les recommandations faites suite à la pré-expérience (expérience 5). C'est notamment le cas pour les variables  $h_{bf}$  et  $G$ . Nous remarquons cependant une cohérence vis-à-vis de la présence du bruit filtré dans les sons les moins intrusifs. Dans le cas de l'approche **AS**, cette propriété est conditionnée par la réduction de l'espace effectuée en amont. Pour l'approche **EC**, cette propriété se retrouve en particulier dans la distribution des modalités de la variable  $S$ . En effet, 16 sons sur 18 contiennent du bruit filtré et seuls quatre sur 18 comportent de la synthèse additive. Comme nous l'avons mentionné dans la discussion de l'expérience 5, cette propriété est cohérente avec ce qui est mentionné par CERA et MISDARIIS 2021 concernant le rôle que joue l'émergence dans le percept d'intrusivité. Cette expérience d'optimisation semble, comme l'expérience 5, montrer que les sons peu intrusifs tendent à ressembler au fond sonore. Une analyse des interactions entre les variables de conception permettrait certainement d'expliquer une partie des différences observées entre les meilleurs sons obtenus avec les deux méthodes, et notamment les raisons pour lesquelles certaines modalités éliminées après l'expérience 5 ne sont pas sous-représentées dans les meilleurs sons trouvés avec la méthode **EC**. En effet, il est possible que certains paramétrages du synthétiseur produisent des sons similaires, ou du moins des sons qui ont une intrusivité semblable.

## 5.4 Conclusion du cinquième chapitre

Dans ce chapitre nous avons mis en oeuvre une méthode de réduction d'espace de conception développée dans le chapitre 4, au travers d'expériences perceptives réelles. Nous nous sommes intéressé·e·s au design de sons peu intrusifs pour les véhicules autonomes. En terme de design sonore, nous nous sommes basé·e·s sur l'utilisation d'un synthétiseur inspiré de la synthèse modulaire, pour la grande diversité de sons que celle-ci peut produire. Une première expérience (expérience 5) nous a permis d'identifier les modalités des variables de conception peu susceptibles de produire des sons faiblement intrusifs et de les éliminer de l'espace de conception. Lors d'une seconde expérience (expérience 6) des sujets ont utilisé un Algorithme Génétique Interactif (IGA) pour explorer simultanément l'espace de conception initial (approche **EC**) et l'espace de conception réduit (approche **AS**). Cette expérience nous a permis d'évaluer la qualité de la réduction d'espace de conception effectuée, en comparant l'intrusivité des sons trouvés dans les deux conditions expérimentales. Nous avons pu ainsi répondre positivement à la quatrième question de recherche :

### Quatrième question de recherche

La méthode de réduction d'espace de conception proposée dans le cadre d'une expérience simulée est-elle efficace dans le cadre d'une expérience réelle ?

Enfin, sur la base d'une analyse des sons trouvés lors des deux expériences nous avons pu, en lien avec la littérature, contribuer à dégager des propriétés sonores contribuant à une faible intrusivité.

**Points clés à retenir**

- Un protocole expérimental hybride, où un algorithme d'optimisation est utilisé après une pré-expérience plus classique visant à dégrossir le problème, semble être une piste intéressante pour approcher un problème d'optimisation perceptive interactive.
- La méthode générale proposée est applicable à d'autres problèmes d'optimisation que celui traité dans ce chapitre. Il est tout à fait possible de l'adapter en considérant notamment d'autres stratégies de réduction d'espace, comme le *screening*, et d'autres algorithmes d'optimisation.
- Nos résultats semblent confirmer que l'intrusivité d'un son dépend en grande partie de son émergence vis-à-vis de l'environnement sonore ambiant.



# CONCLUSION GÉNÉRALE ET PERSPECTIVES

---

## 5.5 Conclusion

Les travaux présentés dans ce manuscrit de thèse proposent des outils pour l'intégration de l'expérience utilisateur dans un processus de design sonore. Celui-ci vise à étudier la perception humaine via des expériences d'optimisation interactive au cours desquelles des utilisateur·rice·s évaluent des sons selon des attributs perceptifs d'intérêt. La méthode proposée comporte trois grandes parties : la définition des variables de conception et de leurs niveaux, les expériences d'optimisation interactive et enfin l'analyse des résultats obtenus individuellement par les participant·e·s. Cette approche est motivée par la possibilité d'explorer des espaces de conception plus grands qu'avec une approche basée sur l'étude de corpus de sons, plus classiquement utilisée. Elle requiert cependant le choix d'un modèle de synthèse sonore paramétrique adapté au problème, qui définit l'espace sonore qui pourra être exploré.

Dans ce manuscrit, nous avons d'abord cherché à définir un protocole expérimental d'optimisation interactive approprié à la conception de sons. Nous avons choisi d'utiliser les Algorithmes Génétiques Interactifs (IGAs) qui ont déjà fait leurs preuves pour l'optimisation de produits via l'intégration d'évaluations perceptives. Deux premières expériences (Expériences 1 et 2) ont permis de confirmer l'intérêt des IGAs, en application à l'optimisation de sons d'alerte de Véhicule Silencieux (VS) pour un problème bi-objectifs. Les deux fonctions-objectif à minimiser étaient le temps de réaction et le désagrément perçu. Un modèle de synthèse sonore spécifique à ce problème a été utilisé, inspiré de précédentes études sur le sujet. Les performances des sons obtenus individuellement par les participant·e·s avec l'IGA étaient meilleures que celles de sons conçus empiriquement et de sons choisis aléatoirement. Les IGAs sont simples à implémenter et ne requièrent pas de compétence particulière de la part des utilisateur·rice·s. Nous avons choisi de mesurer un temps de réaction via la pression sur une touche d'un clavier d'ordinateur, ainsi que de recueillir des jugements subjectifs via une échelle sémantique continue. Cependant, il



---

est possible d'adapter la méthode de d'évaluation perceptive au problème considéré.

Nous avons ensuite abordé la troisième partie du protocole d'optimisation interactive proposé, qui concerne l'analyse des résultats obtenus individuellement par un groupe de participant·e·s. Là où il existe de nombreux outils à disposition dans le cadre d'un protocole plus classique pour l'étude de la perception, seules quelques études proposent des méthodes dans le cas de données obtenues avec des algorithmes d'optimisation interactive. Les méthodes trouvées dans la littérature se basent généralement sur une étude probabiliste des meilleurs designs obtenus par les sujets, ou bien sur la construction de modèles simples à partir des évaluations de ces derniers. Cependant, elles font souvent l'hypothèse d'une indépendance entre les variables de conception, ne permettant pas de produire des recommandations dans le cas du problème bi-objectifs considéré. La méthode d'analyse proposée dans ce manuscrit est basée sur une approche probabiliste, mais tente en amont de prendre en compte les dépendances entre les variables de conception. La connaissance de ces dépendances permet le calcul d'un score pour chaque son, à partir des probabilités marginales calculées pour chaque groupe de variables indépendant. Nous avons appliqué la méthode proposée à l'analyse des résultats obtenus au cours de la première expérience portant sur le design de son d'alerte de VS. Le score calculé permet de classer les sons de l'espace de conception selon leur probabilité d'avoir de bonnes performances. Dans notre cas, nous avons considéré qu'un son a de bonnes performances s'il fait partie des sons appartenant au front de Pareto obtenu par l'un des sujets à la fin du processus d'optimisation. Cependant, ce critère est flexible et peut être adapté à d'autres types de problèmes d'optimisation interactive (mono ou multi-objectifs). La troisième expérience nous a montré que les sons qui ont obtenu le meilleur score selon la méthode proposée sont des propositions de design pertinentes par rapport au problème d'optimisation selon un nouveau panel de sujets. En particulier, les sons proposés présentent de meilleures performances (au sens de Pareto) que des sons sélectionnés aléatoirement dans l'espace de conception. De plus, certains des sons proposés présentent des performances équivalentes (au sens de Pareto) à celles de sons choisis empiriquement. La méthode proposée a donc permis de généraliser les résultats individuels de sujets ayant participé à une expérience d'optimisation interactive. Cette étape est indispensable pour pouvoir informer un processus de design sonore.

Nous avons finalement traité de la première étape du protocole proposé, à savoir le choix des variables de conception et de leurs niveaux. Ce choix est important pour s'assurer de converger dans un nombre raisonnable d'itérations, en raison de la fatigue de

---

l'utilisateur-riche. Or, cette question n'est pas vraiment abordée dans la littérature. Bien que la taille de l'espace de conception soit parfois adaptée au nombre d'itérations prévu, les raisons du choix des variables et de leurs modalités sont rarement expliquées. Cependant, pour un problème de design sonore portant sur une problématique nouvelle ou mal connue, le-a scientifique ou le-a designer-euse sonore n'ont pas nécessairement l'expertise pour faire un choix optimal. Nous avons considéré le choix de l'espace de conception comme un problème de réduction consistant à sélectionner, parmi un espace relativement grand qui inclut toutes les options que l'on souhaite explorer, les variables et modalités les plus pertinentes pour le problème. En nous inspirant de méthodes provenant d'autres disciplines scientifiques et d'ingénierie, nous avons proposé une approche basée sur la réduction du nombre de niveaux par variable de conception. Cette approche se base sur la construction d'un approximateur de la fonction-objectif à minimiser, ensuite utilisé pour identifier les modalités peu susceptibles de donner des solutions intéressantes. Nous avons étudié cette approche via des expériences d'optimisation interactive simulées (Expérience 4), portant sur le même problème de design sonore de son de VS que précédemment. Le problème bi-objectifs de départ a cependant été séparé en deux problèmes mono-objectif. Deux types d'approximateur de la fonction-objectif ont été testés : des descripteurs acoustiques (Approche Descripteur) et des modèles construits à partir de données perceptives (Approche Sujet). Pour cette quatrième expérience, nous avons choisi de réduire arbitrairement toutes les variables du même nombre de niveaux. Les résultats des simulations ont montré l'intérêt de l'approche proposée, qui permet d'améliorer les solutions trouvées au début du processus d'optimisation. Cependant, les simulations effectuées n'ont pas intégré certains facteurs, comme les interactions entre les variables de conception et le bruit dans dans les évaluations des sons par les sujets.

Dans le dernier chapitre de ce manuscrit, nous avons souhaité appliquer la méthode de réduction d'espace de conception proposée à un problème d'optimisation interactive réel. Nous avons abordé le problème du design d'un son d'intérieur de Véhicule Autonome (VA) peu intrusif, visant à communiquer aux passager-ère-s des informations sur les actions du véhicule en cours ou à venir. Nous avons commencé par la création d'un synthétiseur original et expressif, permettant de produire une grande diversité de sons. Celui-ci est basé sur une architecture modulaire et fait appel à plusieurs méthodes de synthèse (additive et soustractive notamment). Nous avons traité ce problème via l'approche dite sujet, c'est-à-dire en procédant dans un premier temps à une pré-expérience (Expérience 5), visant à récolter les données nécessaires à la réduction de l'espace de conception. De la même ma-

---

nière que pour les simulations, nous avons réduit l'espace de conception en éliminant les modalités donnant, en moyenne, les sons les plus intrusifs. Cependant, nous n'éliminons plus le même nombre de modalités pour toutes les variables, mais basons cette fois-ci le choix des modalités à éliminer sur les résultats de tests statistiques. Ainsi, la réduction opérée est moins grande, mais nous mitigeons le risque d'éliminer des modalités correspondant à une faible intrusivité en moyenne. Dans un second temps, nous testons l'effet de la réduction d'espace sur un processus d'optimisation interactive, via une dernière expérience (Expérience 6). Celle-ci montre, comme pour les simulations effectuées au chapitre précédent, une amélioration des solutions trouvées au début du processus d'optimisation. Ces résultats montrent l'intérêt d'effectuer une réduction informée de l'espace de conception avant une expérience d'optimisation interactive. Par ailleurs, le principe général de la méthode proposée peut s'adapter à n'importe quel problème d'optimisation interactive, parce que la pré-expérience fait appel au même matériel et à la même méthode d'évaluation que l'expérience d'optimisation et nécessite peu de préparation supplémentaire, si ce n'est la définition d'un plan d'expérience approprié.

Les travaux effectués pendant cette thèse ont donc permis d'explorer en détail les grandes étapes d'un protocole de conception interactive en design sonore, et de proposer des outils pour chacune d'entre elles, au travers de deux exemples d'application concrets. Nous avons pu mettre en évidence l'intérêt d'un tel protocole, mais également certaines de ses limites.

## 5.6 Perspectives

Plusieurs perspectives de recherche peuvent ainsi être proposées :

- Comme pour beaucoup de métaheuristiques inspirées de la nature, il n'existe pas de méthode universelle pour régler des paramètres des Algorithmes Génétiques (FAUSTO et al. 2020), et par extension pour les Algorithmes Génétiques Interactifs. Ce réglage est donc la plupart du temps effectué empiriquement, à partir de l'expertise personnelle de l'expérimentateur·rice. De plus, l'impact de ce réglage sur les phénomènes d'exploration et d'exploitation est particulièrement difficile à mesurer, alors que l'équilibre entre ces derniers conditionne les performances de l'algorithme. En plus de cela, la structure même des IGAs (choix et agencement des opérateurs génétiques, type de retour demandé à l'utilisateur·rice) est souvent spécifique au problème traité. Ces aspects rendent le processus d'optimisation

---

peu interprétable en détail, et les algorithmes difficiles à comparer entre eux. Il serait enrichissant d'explorer l'utilisation d'algorithmes plus simples et non paramétriques, afin de savoir si des alternatives aux Algorithmes Évolutionnistes (IEC) peuvent fonctionner et aussi de mieux comprendre les mécanismes en jeu dans l'optimisation interactive de sons,

- Les sons recommandés suite à l'analyse des résultats individuels pour le design de son d'alerte de VS sont plutôt situés vers les extrêmes du front de Pareto des sons évalués pendant l'expérience 3. Investiguer la cause de ce résultat permettrait de savoir si celui-ci est dû à un potentiel désaccord entre les sujets, à l'IGA utilisé (ou au réglage de ses paramètres) ou bien à la méthode d'analyse proposée. Dans le cas où l'un des deux derniers points serait en cause, des pistes d'amélioration pourraient être proposées, pour pouvoir faire des recommandations plus variées en termes de compromis entre les deux fonctions-objectifs,
- Notre méthode de réduction de l'espace de conception est basée sur l'élimination de modalités. Il est probable que d'autres approches, faisant par exemple appel à une analyse de sensibilité (SALTELLI et al. 2008), soient mieux adaptées pour d'autres algorithmes d'optimisation. C'est une raison de plus pour laquelle il serait intéressant d'explorer l'utilisation d'autres algorithmes,
- Davantage d'études sont nécessaires pour savoir si les résultats obtenus pendant l'expérience 6, moins marqués que ceux observés pendant les simulations (Expérience 5), sont dûs à la méthode de réduction d'espace proposée, ou bien par exemple à la tâche demandée aux sujets (intrusivité difficile à évaluer via le protocole proposé ou difficile de distinguer l'intrusivité de certains sons). En ce sens, il pourrait être éclairant de tester la méthode avec d'autres problèmes de design sonore,
- Une approche souvent proposée pour améliorer les performances des IGAs consiste à utiliser un modèle des évaluations du sujet pour faciliter le processus d'optimisation (PEI et TAKAGI 2018). Ce modèle peut être construit en avance ou pendant l'optimisation, et mis à jour au cours des générations. Le modèle construit suite à la pré-expérience dans la méthode réduction d'espace proposée pourrait également être utilisé dans ce but,
- D'autres auteur·rice·s tel·le·s que MATA et al. (2018) proposent de prendre en compte l'ensemble des évaluations d'une cohorte de sujets directement pendant le processus d'optimisation. Dans ce cas, tous les sujets participent à l'expérience

---

simultanément et leurs évaluations sont traitées par un unique IGA. Ceci présente l'avantage de récolter des informations qui sont directement représentatives d'un groupe de sujets. Une telle approche pourrait être intéressante dans le cas de la co-conception de sons,

- Il serait pertinent d'accorder une plus grande attention à la validité écologique de l'expérience d'optimisation interactive. En effet, le processus d'optimisation requiert de pouvoir générer et évaluer un grand nombre de propositions de design de manière successive et rapidement, ce pourquoi les travaux présentés dans cette thèse se sont limités à des méthodes de présentation et d'évaluation simples. Une présentation des sons dans un contexte plus proche du contexte d'utilisation finale pourrait permettre une meilleure intégration de l'expérience utilisateur dans le processus de conception. Cette démarche pourrait être facilitée par l'utilisation d'algorithmes plus performants, nécessitant moins d'évaluations de la part de l'utilisateur·rice,
- Plus généralement concernant les deux précédents points, nous avons fait le choix d'utiliser des outils d'optimisation interactive pour étudier des attributs perceptifs, via un mode d'interaction relativement limité. Il serait intéressant d'explorer l'aspect interactif des outils proposés, par exemple via d'autres méthodes d'évaluation des sons, ou encore dans une démarche de co-conception.

En conclusion, les travaux effectués au cours de cette thèse ont permis d'explorer un protocole expérimental d'aide au design sonore basé des outils d'optimisation interactive. Les méthodes utilisées, encore peu appliquées à l'étude la perception sonore, ont donné des résultats prometteurs pour les deux problème applicatifs considérés. Plusieurs pistes de recherche ont été proposées, afin d'améliorer les performances des outils proposés et d'intégrer au mieux l'expérience utilisateur dans le processus de conception.

# PRÉCISIONS SUR LA MÉTHODE D'ÉLIMINATION DE NIVEAUX

---

## A.1 Démonstration du calcul de $G_i(m)$

Nous considérons un paramétrage comme étant une variable aléatoire  $D = (V_1, \dots, V_I)$ , qui suit une distribution discrète uniforme sur tout l'espace de conception  $E = \prod_{i=1}^I \llbracket 1; M_i \rrbracket$ . Nous appelons  $G_i(m)$  l'espérance conditionnelle de  $g(D)$  sachant  $V_i$ , c'est-à-dire l'espérance de  $g(D)$  lorsque  $V_i = m$ .

$$G_i(m) = \mathbb{E}[g(V_1, \dots, V_I) \mid V_i = m] \quad (\text{A.1})$$

Celle-ci peut être calculée comme :

$$G_i(m) = \frac{1}{\prod_{j \neq i} M_j} \sum_{v_1} \cdots \sum_{v_{i-1}} \sum_{v_{i+1}} \cdots \sum_{v_I} g(v_1, \dots, v_{i-1}, m, v_{i+1}, \dots, v_I) \quad (\text{A.2})$$

En effet, la fonction de masse de  $D$  vaut :

$$f_D(d) = \frac{1}{\prod_{i=1}^I M_i} \quad (\text{A.3})$$

Soit  $Y = g(D)$ . Alors  $G_i(m)$  peut s'exprimer comme :

$$\begin{aligned} G_i(m) &= \sum_y y \mathbb{P}[g(V_1, \dots, V_I) = y \mid V_i = m] \\ &= \sum_y y \frac{\mathbb{P}[(g(V_1, \dots, V_I) = y) \cap (V_i = m)]}{\mathbb{P}[V_i = m]} \end{aligned} \quad (\text{A.4})$$

---

La distribution de  $D$  est uniforme, donc  $\mathbb{P}[V_i = m] = \frac{1}{M_i}$ . Par conséquent :

$$G_i(m) = M_i \sum_y y \mathbb{P}[(g(V_1, \dots, V_I) = y) \cap (V_i = m)] \quad (\text{A.5})$$

La probabilité à droite de l'équation A.5 peut s'exprimer à partir de la fonction de masse  $f_D(d)$  :

$$\begin{aligned} G_i(m) &= M_i \sum_y y \sum_{dt.q.(g(d)=y) \cap (V_i=m)} f_D(d) \\ &= M_i \sum_{dt.q.V_i=m} g(d) f_D(d) \end{aligned} \quad (\text{A.6})$$

Enfin, grâce l'équation A.3 nous avons :

$$\begin{aligned} G_i(m) &= \frac{M_i}{\prod_{j=1}^I M_j} \sum_{dt.q.V_i=m} g(d) \\ &= \frac{1}{\prod_{j \neq i} M_j} \sum_{dt.q.V_i=m} g(d) \\ &= \frac{1}{\prod_{j \neq i} M_j} \sum_{v_1} \cdots \sum_{v_{i-1}} \sum_{v_{i+1}} \cdots \sum_{v_I} g(v_1, \dots, v_{i-1}, m, v_{i+1}, \dots, v_I) \end{aligned} \quad (\text{A.7})$$

## A.2 Démonstration de $\mathbb{E}[g(D')] < \mathbb{E}[g(D)]$

Pour chaque variable  $V_i$ , nous réduisons le nombre de modalités à  $k_i$ , avec  $1 \leq k_i < M_i$ , en écartant les  $M_i - k_i$  modalités pour lesquelles la valeur de  $G_i(m)$  est la plus grande. Soit  $D'$  la variable aléatoire correspondant à l'espace de conception réduit  $E' = \prod_{i=1}^I \llbracket 1; k_i \rrbracket$ . Par convention, numérotons  $m$  tel que les valeurs de  $G_i(m)$  soient par ordre croissant :

$$G_i(n) < G_i(m), \forall n < m \quad (\text{A.8})$$

En éliminant les niveaux  $V_i$  qui maximisent  $G_i(m)$ , nous réduisons l'espérance de  $g$  :

$$\mathbb{E}[g(D')] < \mathbb{E}[g(D)] \quad (\text{A.9})$$

---

En effet, l'espérance de  $Y = g(D)$  vaut :

$$\begin{aligned}
\mathbb{E}[g(D)] &= \mathbb{E}[\mathbb{E}[g(D) \mid V_i]] \\
&= \sum_{m=1}^{M_i} \mathbb{P}[V_i = m] \mathbb{E}[g(D) \mid V_i = m] \\
&= \frac{1}{M_i} \sum_{m=1}^{M_i} \mathbb{E}[g(D) \mid V_i = m] \\
&= \frac{1}{M_i} \sum_{m=1}^{M_i} G_i(m)
\end{aligned} \tag{A.10}$$

De même, l'espérance de  $Y' = g(D')$  vaut :

$$\mathbb{E}[g(D')] = \frac{1}{k_i} \sum_{m=1}^{k_i} G_i(m) \tag{A.11}$$

Enfin, l'équation A.8 nous donne :

$$\begin{aligned}
&\sum_{n=1}^{k_i} \frac{G_i(n)}{k_i} < \sum_{m=k_i+1}^{M_i} \frac{G_i(m)}{M_i - k_i} \\
\Leftrightarrow \frac{M_i - k_i}{M_i} \sum_{n=1}^{k_i} \frac{G_i(n)}{k_i} &< \sum_{m=k_i+1}^{M_i} \frac{G_i(m)}{M_i} \\
\Leftrightarrow \sum_{n=1}^{k_i} \frac{G_i(n)}{k_i} &< \sum_{m=1}^{M_i} \frac{G_i(m)}{M_i} \\
\Leftrightarrow \mathbb{E}[g(D')] &< \mathbb{E}[g(D)]
\end{aligned} \tag{A.12}$$



# DÉTAIL DES SECTIONS DU SYNTHÉTISEUR DE SONS DE VA

---

Nous nous inspirons du logiciel Puredata pour la représentation graphique du synthétiseur utilisée dans ce manuscrit, également basée sur des modules interconnectés. Cependant, afin de faciliter la compréhension du fonctionnement du synthétiseur, nous utilisons une schématisation de plus haut niveau. Les modules et les données présentés dans les figures ne sont donc pas exactement les mêmes que dans les *patches* Puredata utilisés en réalité. De plus, nous faisons un choix différent dans les types de données représentées : nous faisons la distinction entre des paramètres à valeur constante et des signaux dépendant du temps (représentés en gras). Le schéma générique d'un module est présenté sur la figure B.1.

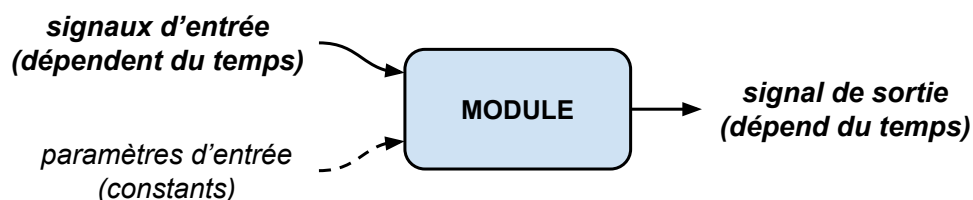


FIGURE B.1 – Module générique. Les flèches en traits pleins représentent des signaux dépendant du temps, tandis que la flèche en trait pointillés représente un paramètre à valeur constante. Les noms des signaux sont indiqués en gras.

## B.1 Détail des sections

Les figures B.2 à B.5 détaillent le fonctionnement des quatre sections du synthétiseur. Chaque section est composée de plusieurs modules interconnectés, dont les paramètres sont détaillés ci-après.

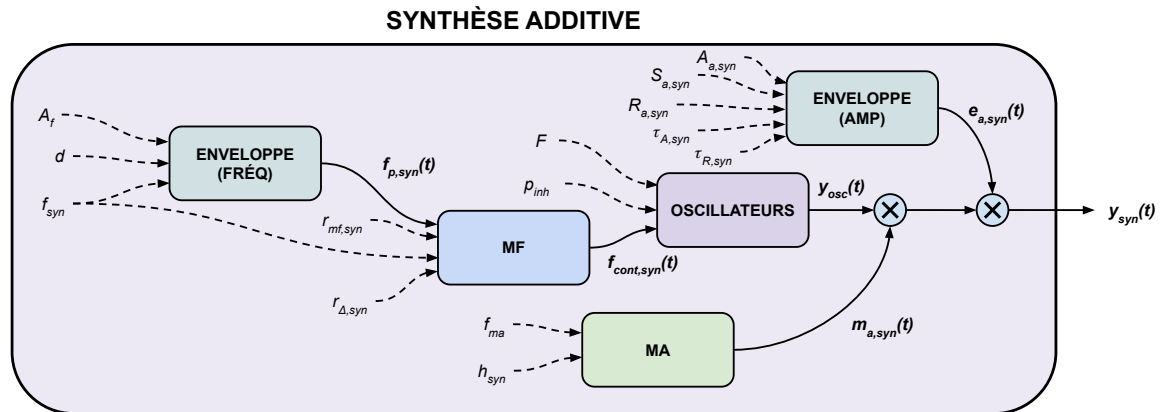


FIGURE B.2 – Section *SYNTHÈSE ADDITIVE*, composée de cinq modules. Un module *OSCILLATEURS* génère une somme d'oscillateurs sinusoidaux. Les fréquences instantanées des oscillateurs sont contrôlées par un module *ENVELOPPE (FRÉQ)* et un modulateur de fréquence *MF*. Enfin, le signal produit par le module *OSCILLATEURS* est affecté par un modulateur d'amplitude *MA* et un module *ENVELOPPE (AMP)*. Les variables présentées sur la figure sont détaillées dans la suite de cette annexe.

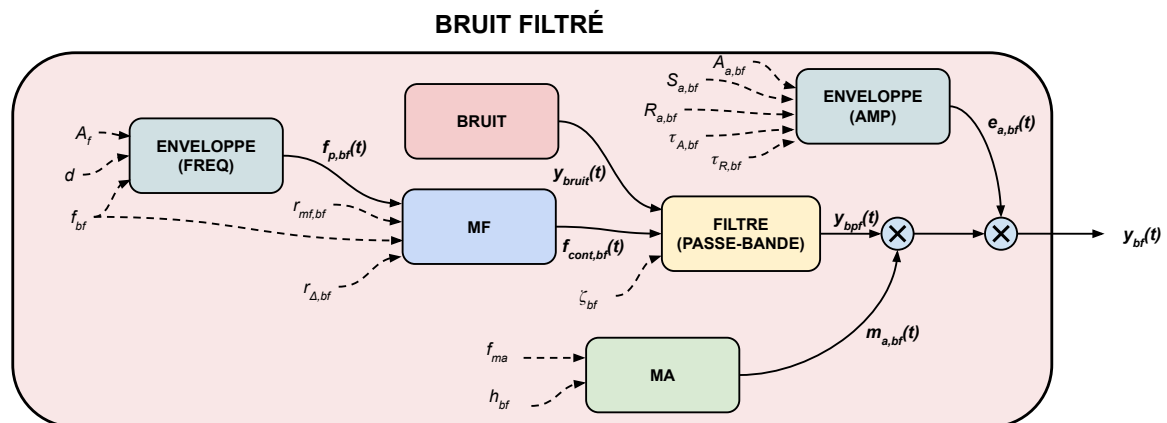


FIGURE B.3 – Section *BRUIT FILTRÉ*. Un module *BRUIT* génère un bruit rose, filtré par le module *FILTRE (PASSE-BANDE)*. La fréquence centrale du filtre est contrôlée par un module *ENVELOPPE (FRÉQ)* et un modulateur de fréquence *MF*. Enfin, le signal de bruit filtré est affecté par un modulateur d'amplitude *MA* et un module *ENVELOPPE (AMP)*. Les variables présentées sur la figure sont détaillées dans la suite de cette annexe.

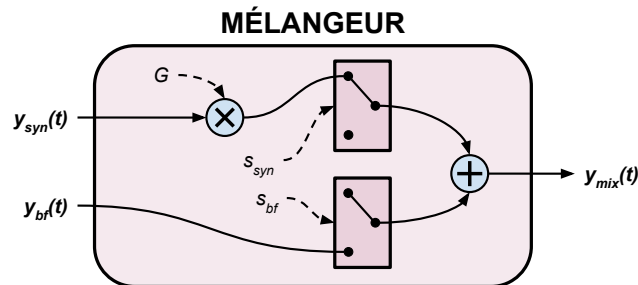


FIGURE B.4 – Section *MÊLANGEUR*. Un gain  $G$  est appliqué au signal  $y_{syn}(t)$  provenant de la section *SYNTHÈSE ADDITIVE*. Deux interrupteurs permettent d’inclure ou non les signaux  $y_{syn}(t)$  et  $y_{bf}(t)$ , respectivement issus des sections *SYNTHÈSE ADDITIVE* et *BRUIT FILTRÉ*. Les variables présentées sur la figure sont détaillées dans la suite de cette annexe.

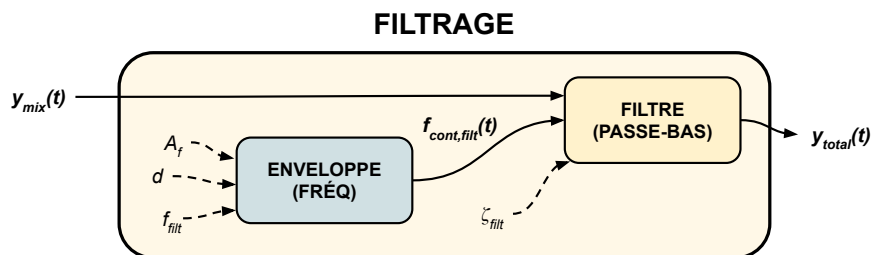


FIGURE B.5 – Section *FILTRAGE*. Le signal  $y_{mix}(t)$ , résultant du mélange des deux modules de synthèse additive et de bruit filtré, passe par un filtre passe-bas, dont la fréquence de coupure est modulée par une enveloppe. Les variables présentées sur la figure sont détaillées dans la suite de cette annexe.

---

## B.2 Détails des modules

Certains modules sont présents dans plusieurs sections. Afin d'expliquer leur fonctionnement de manière générique, des variables avec l'indice  $*$  sont utilisées dans la suite de ce manuscrit. L'indice  $*$  signifie indifféremment *syn*, *bf* ou *filt*, respectivement en référence aux sections *SYNTHÈSE ADDITIVE*, *BRUIT FILTRÉ* et *FILTRAGE*. Par exemple, la variable  $y_*$  désigne indifféremment les variables  $y_{syn}$ ,  $y_{bf}$  et  $y_{filt}$ , respectivement présentes dans les sections *SYNTHÈSE ADDITIVE*, *BRUIT FILTRÉ* et *FILTRAGE*.

Ces trois sections contiennent un même module appelé *ENVELOPPE (FRÉQ)*. Ce dernier, détaillé en figure B.6, génère une fréquence instantanée  $f_{p,*}(t)$  à partir de trois paramètres d'enveloppe que nous appelons  $A_f$  (pour *Attack* ou *Attaque*), une fréquence cible  $f_*$  et un écart  $d$ .  $A_f$  correspond au temps que prend la fréquence de sortie  $f_{p,*}(t)$  pour atteindre  $f_*$ , en partant de la fréquence  $f_* + d$  (voir figure B.7).  $d$  correspond à l'écart entre la fréquence de départ et la fréquence cible, en demi-tons. Les valeurs prises par ces paramètres sont présentées dans le tableau B.1. Le paramètre d'attaque  $A_f$  est défini comme un pourcentage de la durée totale du son généré  $d_s$ . Cette dernière est fixée à  $d_s = 1,5$ s pour tous les sons. Les paramètres  $A_f$  et  $d$  sont regroupés sous une seule variable de conception  $E_f = (A_f, d)$ , pouvant prendre cinq modalités :

1. Attaque courte, fréquence ascendante.
2. Attaque longue, fréquence ascendante.
3. Fréquence constante.
4. Attaque longue, fréquence descendante.
5. Attaque courte, fréquence descendante.

Seule la variable  $f_*$  peut prendre une valeur spécifique à la section dans laquelle le module *ENVELOPPE (FRÉQ)* est placé. Les valeurs de  $f_{bf}$  et  $f_{filt}$  sont définies comme des multiples entiers de  $f_{syn}$ . L'enveloppe de fréquence est donc identique pour les trois sections *SYNTHÈSE ADDITIVE*, *BRUIT FILTRÉ* et *FILTRAGE*, à un facteur entier près.

Le module suivant, présent dans les sections *SYNTHÈSE ADDITIVE* et *BRUIT FILTRÉ*, est le module *MF*. Ce dernier, dont le détail est présenté dans la figure B.8, prend la fréquence porteuse  $f_{p,*}(t)$  et la module avec une onde sinusoïdale. La fréquence instantanée en sortie vaut :

$$f_{cont,*}(t) = f_{p,*}(t) + r_{\Delta,*} \cdot f_* \cdot \sin \left( 2\pi \int_0^t r_{mf,*} \cdot f_{p,*}(\tau) d\tau \right) \quad (\text{B.1})$$

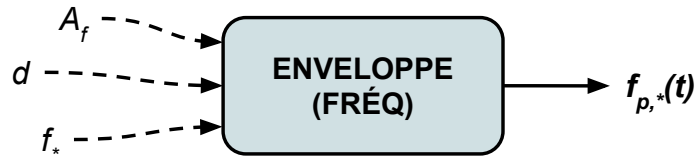


FIGURE B.6 – Schéma du module *ENVELOPPE (FRÉQ)*.

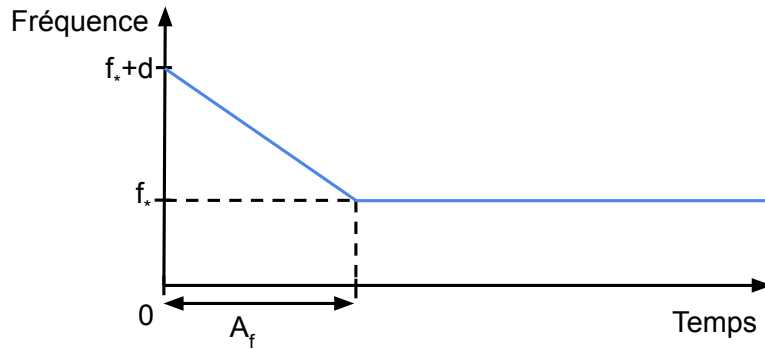


FIGURE B.7 – Fréquence  $f_{p,syn}(t)$  au cours du temps.  $A_f$  représente le temps que met le signal à atteindre sa valeur cible  $f_{syn}$  et  $d$  l'écart de départ par rapport à cette valeur cible, en demi-tons.

TABLE B.1 – Variables de conception pour les modules *ENVELOPPE (FRÉQ)*.

Variable	Unité	Modalité				
		1	2	3	4	5
$E_f = \begin{bmatrix} A_f \\ d \end{bmatrix}$	$\%(d_s)$	5	20	—	20	5
	demi-tons	-5	-5	0	+5	+5
$f_{syn}$	Hz	100	200	400	800	1600
$f_{bf}$	Hz	$f_{syn}$	$2 \cdot f_{syn}$	$3 \cdot f_{syn}$	—	—
$f_{filt}$	Hz	$f_{syn}$	$2 \cdot f_{syn}$	$3 \cdot f_{syn}$	—	—

TABLE B.2 – Variables de conception pour les modules *MF*.

Variable	Unité	Modalité				
		1	2	3	4	5
$r_{mf, syn}$	%	1	5	10	50	80
$r_{\Delta, syn}$	%	0	10	50	—	—
$r_{mf, bf}$	%	1	5	10	—	—
$r_{\Delta, bf}$	%	0	10	30	—	—

L'écart à la fréquence de la porteuse est proportionnel à  $f_*$  et vaut  $r_{\Delta,*} \cdot f_*$ . La fréquence instantanée du signal modulant est proportionnelle à  $f_{p,*}(t)$  et vaut  $r_{mf,*} \cdot f_{p,*}(t)$ . Celle-ci suit donc l'enveloppe de fréquence générée par le module *ENVELOPPE (FRÉQ)*. Les valeurs prises par les variables  $r_{\Delta,*}$  et  $r_{mf,*}$  sont détaillées dans le tableau B.2. Plus l'indice de la modalité de  $r_{mf,*}$  est élevé, plus la fréquence de modulation est élevée. De même, plus l'indice de la modalité de  $r_{\Delta,*}$  est élevé, plus l'écart de fréquence est élevé. Lorsque  $r_{\Delta,*} = 0$  (modalité 1), il n'y a pas de modulation de fréquence.

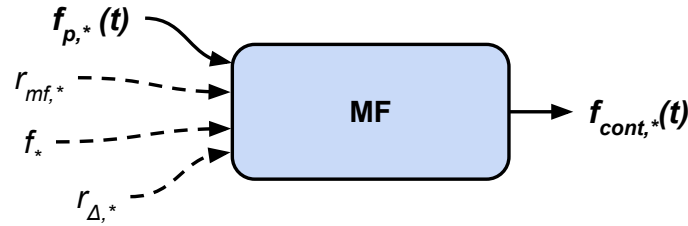


FIGURE B.8 – Schéma du module *MF*.  $f_{p,*}(t)$  est la fréquence instantanée de la porteuse. La fréquence de modulation est proportionnelle à la fréquence de la porteuse, et est obtenue en multipliant celle-ci par  $r_{mf,*}$ . L'écart en fréquence est le produit  $r_{\Delta,*} \cdot f_*$ .  $f_{cont,*}(t)$  est la fréquence instantanée en sortie du module *MF*.

Dans la section *SYNTHÈSE ADDITIVE*, le son est généré par le module *OSCILLATEURS*. Celui-ci (voir figure B.9) est composé de six oscillateurs sinusoïdaux, dont les fréquences sont des multiples d'une fréquence de contrôle  $f_{cont, syn}(t)$ , qui peut varier au cours du temps. Un paramètre,  $p_{inh}$ , permet de rendre le rapport entre les oscillateurs plus ou moins harmonique. La fréquence instantanée  $f_{osc,i}(t)$  de l'oscillateur  $i$  est donnée par :

$$f_{osc,i}(t) = i \cdot f_{cont, syn}(t) \cdot p_{inh}^{i-1} \quad (\text{B.2})$$

Nous appelons  $F = (a_{osc,1}, \dots, a_{osc,6})$  le paramètre à six dimensions correspondant aux

amplitudes des oscillateurs. Le signal généré par l'oscillateur  $i$  vaut :

$$y_{osc,i}(t) = a_{osc,i} \cdot \sin\left(2\pi \int_0^t f_{osc,i}(\tau) d\tau\right) \quad (\text{B.3})$$

Le signal de sortie  $y_{osc}(t)$  du module *OSCILLATEURS* est la somme des  $y_{osc,i}(t)$  :

$$y_{osc}(t) = \sum_{i=1}^6 y_{osc,i}(t) \quad (\text{B.4})$$

Les valeurs des variables contrôlant le module *OSCILLATEURS* sont détaillées dans le tableau B.3. La variable  $F$ , qui regroupe les amplitudes des six oscillateurs, possède quatre modalités correspondant aux amplitudes des six premières harmoniques de formes d'ondes couramment utilisées en synthèse sonore :

1. Onde carrée.
2. Onde triangulaire.
3. Onde en dents de scie.
4. Onde sinusoïdale.

Les amplitudes des oscillateurs sont normalisées de sorte que l'amplitude de l'onde résultante soit égale à 1. Le pourcentage d'inharmonicité  $p_{inh}$  peut prendre trois modalités. Plus l'indice de la modalité de  $p_{inh}$  est élevé, plus le son généré est inharmonique. Pour la modalité 1, les fréquences des oscillateurs sont réparties de manière harmonique.

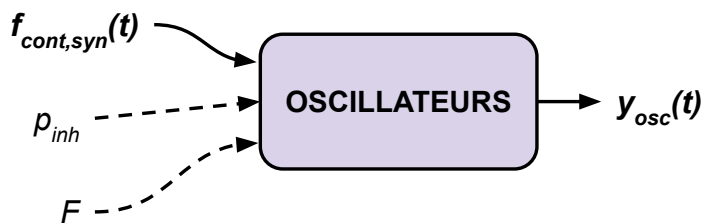


FIGURE B.9 – Schéma du module *OSCILLATEURS*. La fréquence de contrôle  $f_{cont,syn}(t)$  et le pourcentage d'inharmonicité  $p_{inh}$  gèrent les fréquences des oscillateurs, tandis que le paramètre  $F$  définit leurs amplitudes.

Pour la section *BRUIT FILTRÉ*, le son est généré par un module *BRUIT*. Ce dernier (voir figure B.10) ne prend aucun paramètre ni signal en entrée et génère en sortie un

TABLE B.3 – Variables de conception pour le module *OSCILLATEURS*.

Variable	Unité	Modalité				
		1	2	3	4	
$F =$	$a_{osc,1}$	—	1,2732	0,8106	-0,6366	1
	$a_{osc,2}$	—	0	0	0,3183	0
	$a_{osc,3}$	—	0,4244	-0,0901	-0,2122	0
	$a_{osc,4}$	—	0	0	0,1562	0
	$a_{osc,5}$	—	0,2546	0,0324	-0,1273	0
	$a_{osc,6}$	—	0	0	0,1061	0
$p_{inh}$	%	0	1	5	—	

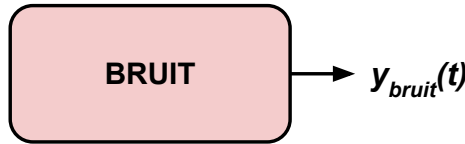


FIGURE B.10 – Schéma du module *BRUIT*.  $y_{bruit}(t)$  est un bruit rose.

bruit rose  $y_{bruit}(t)$ , c'est-à-dire un bruit dont la densité spectrale par bande d'octave est constante.

Le signal  $y_{bruit}(t)$  passe ensuite par un module *FILTRE (PASSE-BANDE)*. Un module similaire, appelé *FILTRE (PASSE-BAS)*, est présent dans la section *FILTRAGE* et affecte la totalité du son généré par le synthétiseur. Ces deux modules (voir figure B.11 et B.12) sont des implémentations de filtres à variable d'état (*State Variable Filters*; DUTILLEUX 1989; WISHNICK 2014), choisis car la fréquence centrale (ou de coupure) et la résonance (c'est-à-dire la valeur du facteur de qualité  $Q$ , ou inversement du facteur d'atténuation  $\zeta$ ) peuvent être contrôlées séparément et être dépendantes du temps. Les coefficients des versions passe-bas, passe-bande et passe-haut d'un même filtre sont calculés simultanément, ce qui permet de passer facilement de l'un à l'autre. Les signaux  $f_{cont,bf}(t)$  et  $f_{cont,flt}(t)$  contrôlent respectivement la fréquence centrale du filtre passe-bande de la section *BRUIT FILTRÉ* et la fréquence de coupure du filtre passe-bas de la section *FILTRAGE*. Les variables  $\zeta_*$  désignent les facteurs d'atténuation des deux filtres, dont les valeurs possibles sont indiquées dans le tableau B.4. Une valeur élevée de  $\zeta_*$  correspond à une faible résonance, et inversement.

Les signaux  $y_{osc}(t)$  et  $y_{bruit}(t)$  sont tous les deux multipliés par un signal modulant  $m_{a,*}(t)$ . Dans les deux cas, le signal  $m_{a,*}(t)$  est généré par un module *MA* (voir figure



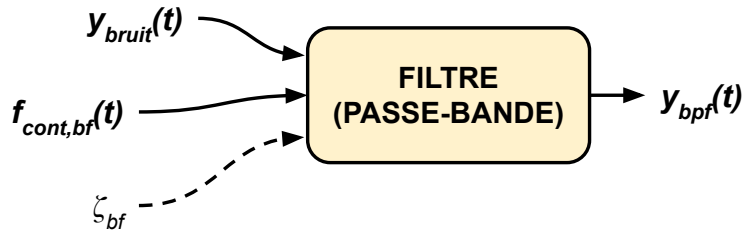


FIGURE B.11 – Schéma du module *FILTRE (PASSE-BANDE)*. Le signal d'entrée  $y_{bruit}(t)$  passe par un filtre passe-bande, de fréquence centrale  $f_{cont,bf}(t)$  et dont le facteur d'atténuation est  $\zeta_{bf}$ .

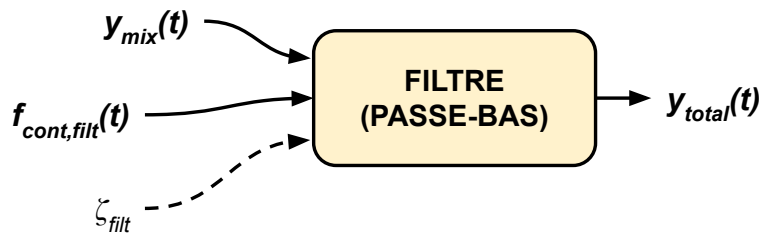


FIGURE B.12 – Schéma du module *FILTRE (PASSE-BAS)*. Le signal d'entrée  $y_{mix}(t)$  passe par un filtre passe-bas, de fréquence de coupure  $f_{cont,filt}(t)$  et dont le facteur d'atténuation est  $\zeta_{filt}$ .

TABLE B.4 – Variables de conception pour les modules *FILTRE (PASSE-BANDE)* et *FILTRE (PASSE-BAS)*, appartenant respectivement aux sections *BRUIT FILTRÉ* et *FILTRAGE*.

Variable	Unité	Modalité		
		1	2	3
$\zeta_{bf}$	—	0,1	0,3	0,4
$\zeta_{filt}$	—	0,2	0,3	0,5

TABLE B.5 – Variables de conception pour les modules *MA*.

Variable	Unité	Modalité			
		1	2	3	4
$f_{ma}$	Hz	10	20	50	100
$h_{syn}$	%	0	20	80	—
$h_{bf}$	%	0	20	80	—

B.13), avec l'équation suivante :

$$m_{a,*}(t) = (1 + h_*) \cdot \sin(2\pi f_{mat}) \quad (\text{B.5})$$

$h_*$  désigne le taux de modulation, et  $f_{ma}$  la fréquence de modulation. Cette dernière est identique pour les deux sections *SYNTHÈSE ADDITIVE* et *BRUIT FILTRÉ*, afin de préserver la cohérence global du son produit. Les valeurs que peuvent prendre les variables  $h_*$  et  $f_{ma}$  sont détaillées dans le tableau B.5. Plus la valeur de  $h_*$  est élevée et plus la modulation est importante. Lorsque  $h_* = 0$  (modalité 1), il n'y a pas de modulation d'amplitude.

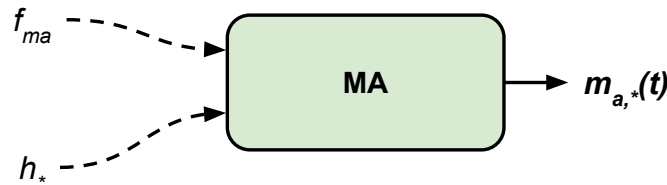


FIGURE B.13 – Schéma du module *MA*.  $f_{ma}$  est la fréquence du signal modulant et  $h_*$  est l'indice de modulation.

Ensuite, le dernier module présent dans les sections *SYNTHÈSE ADDITIVE* et *BRUIT FILTRÉ* est un module générateur d'enveloppe appelé *ENV (AMP)*. Ce dernier (voir figure B.14) est à chaque fois appliqué en fin de section et permet de définir la durée du signal généré. Il est contrôlé par cinq paramètres que sont  $A_{a,*}$  (pour *Attack* ou *Attaque*),  $S_{a,*}$  (pour *Sustain* ou *Entretien*),  $R_{a,*}$  (pour *Release* ou *Extinction*), la constante de temps d'attaque  $\tau_{A,*}$  et la constante de temps d'extinction  $\tau_{R,*}$ . La figure B.15 montre deux exemples de courbes d'enveloppe générées par le module. Les courbes des phases d'attaque et d'extinction sont respectivement similaires à celles de la charge et de la décharge d'un condensateur. Les équations correspondantes ont été modifiées pour que

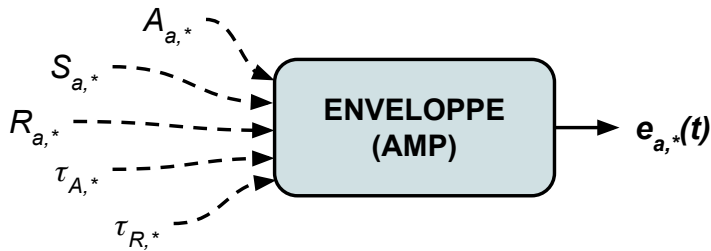


FIGURE B.14 – Schéma du module *ENVELOPPE (AMP)*. Les paramètres  $A_{a,*}$ ,  $S_{a,*}$  et  $R_{a,*}$  correspondent aux durées des trois phases de l'enveloppe (respectivement attaque, entretien et extinction). Les paramètres  $\tau_{A,*}$  et  $\tau_{R,*}$  sont des constantes de temps qui affectent la forme des courbes d'attaque et d'extinction.

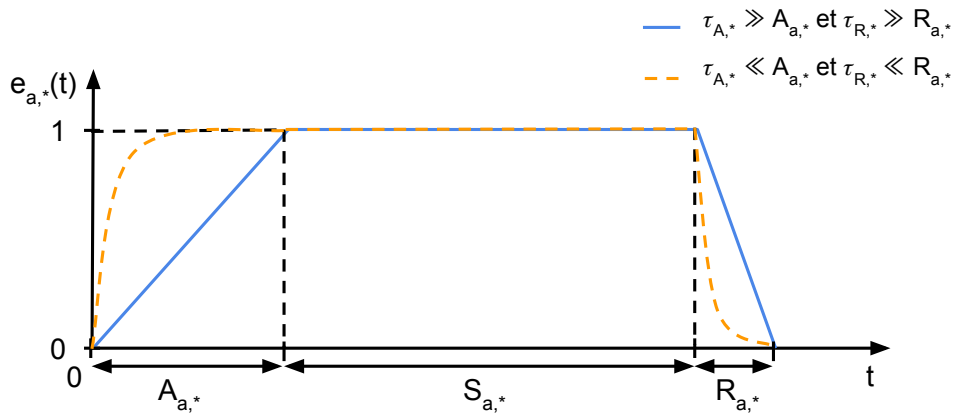


FIGURE B.15 – Enveloppe  $e_{a,*}(t)$  au cours du temps. La courbe en trait plein montre la forme de l'attaque et de l'extinction lorsque les constantes de temps correspondantes sont très grandes. À l'inverse, la courbe en pointillés correspond à des constantes de temps très faibles.

TABLE B.6 – Variables de conception pour le module *ENVELOPPE (AMP)*.

Variable	Unité	Modalité				
		1	2	3	4	5
$E_{a, syn} = \begin{bmatrix} A_{a, syn} \\ R_{a, syn} \end{bmatrix}$	$\%(d_s)$	5	50	95	10	50
	$\%(d_s)$	95	50	5	10	5
$E_{a, bf} = \begin{bmatrix} A_{a, bf} \\ R_{a, bf} \end{bmatrix}$	$\%(d_s)$	5	50	95	10	50
	$\%(d_s)$	95	50	5	10	5

l'enveloppe  $e_{a,*}(t)$  atteint une valeur de 1 lorsque  $t = A_{a,*}$  et une valeur de 0 lorsque  $t = A_{a,*} + S_{a,*} + R_{a,*}$ . Les constantes de temps  $\tau_{A,*}$  et  $\tau_{R,*}$  permettent de modifier la forme des courbes d'attaque et d'extinction. De faibles valeurs correspondent à une attaque ou une extinction rapide, alors que de grandes valeurs correspondent à des courbes quasi linéaires. Pendant l'attaque, l'équation de l'enveloppe  $e_{a,*}(t)$  est :

$$e_{a,*}(t) = \frac{1 - e^{-\frac{t}{\tau_{A,*}}}}{1 - e^{-\frac{A_{a,*}}{\tau_{A,*}}}} \quad (\text{B.6})$$

Son équation pendant la phase d'extinction est :

$$e_{a,*}(t) = \frac{e^{-\frac{t-(A_{a,*}+S_{a,*})}{\tau_{R,*}}} - e^{-\frac{R_{a,*}}{\tau_{R,*}}}}{1 - e^{-\frac{R_{a,*}}{\tau_{R,*}}}} \quad (\text{B.7})$$

Les valeurs que peuvent prendre les paramètres du module *ENVELOPPE (AMP)* sont détaillées dans le tableau B.6. Chaque paire de paramètres  $A_{a,*}$  et  $R_{a,*}$  est regroupée sous une seule variable de conception  $E_{a,*}$ . Chaque variable  $E_{a,*}$  possède cinq modalités, correspondant à différents profils d'enveloppe, illustrés sur la figure B.16. Les modalités de  $E_{a, syn}$  et  $E_{a, bf}$  correspondent aux mêmes profils d'enveloppe. Le paramètre  $S_{a,*}$  n'est pas inclus dans  $E_{a,*}$ , car sa valeur est déduite des valeurs des paramètres  $A_{a,*}$  et  $R_{a,*}$ , ainsi que de la durée du signal  $d_s$ . Les valeurs de  $\tau_{A,*}$  et  $\tau_{R,*}$  sont calculées à partir des valeurs de  $A_{a,*}$  et  $R_{a,*}$ , avec  $\tau_{A,*} = A_{a,*}/3$  et  $\tau_{R,*} = R_{a,*}/3$ .

Enfin, la section *MÉLANGEUR* (voir figure B.4) ne comporte pas de module mais uniquement des opérations élémentaires. Le gain  $G$  est d'abord appliqué au signal  $y_{syn}(t)$ . Puis, deux interrupteurs permettent de choisir quels signaux d'entrée composeront la sortie  $y_{mix}(t)$ . Ces deux interrupteurs sont contrôlés par les variables  $s_{syn}$  et  $s_{bf}$ . Lorsque  $s_{syn} = 1$ ,  $G \cdot y_{syn}(t)$  est ajouté à  $y_{mix}(t)$ . Lorsque  $s_{syn} = 0$ ,  $y_{mix}(t)$  ne contient pas  $G \cdot y_{syn}(t)$ .

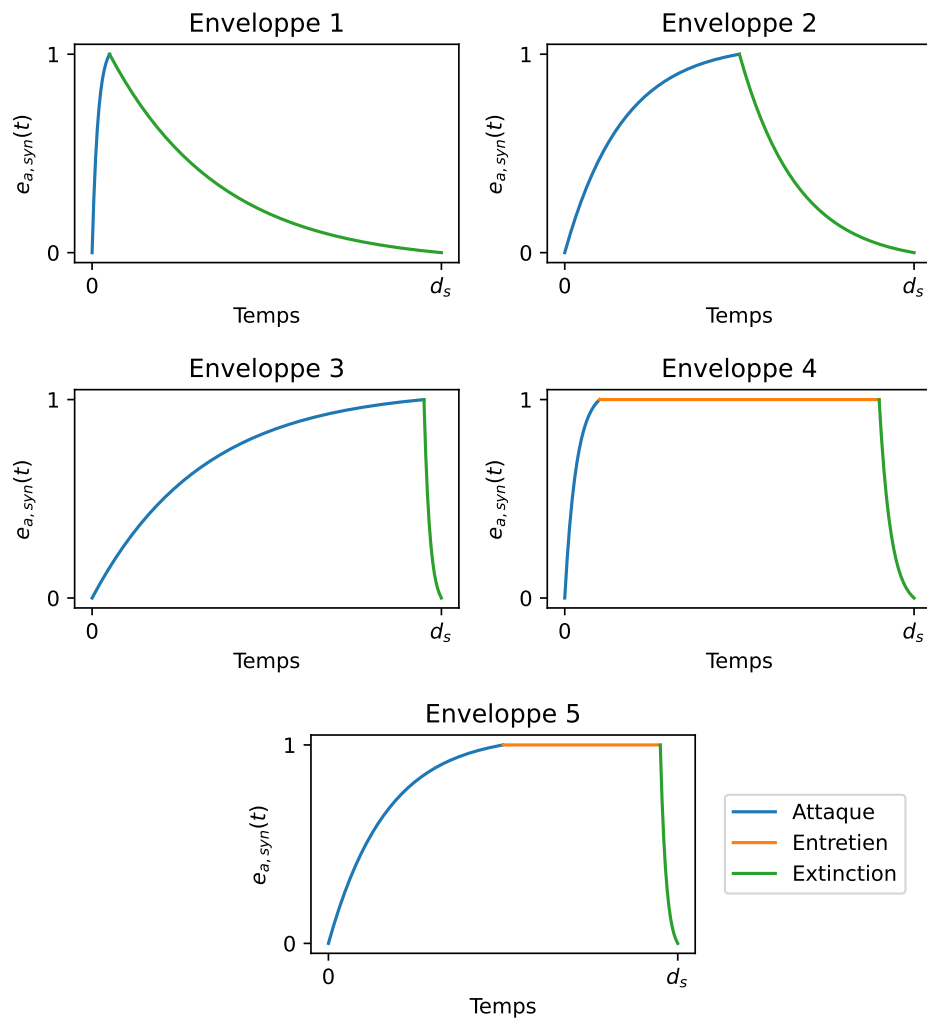


FIGURE B.16 – Profils d’enveloppe de la variable  $E_{a,*}$ . Les couleurs représentent les trois phase de l’enveloppe, attaque (A), entretien (S) et extinction (R).

TABLE B.7 – Variables de conception pour la section *MÉLANGEUR*.

Variable	Unité	Modalité				
		1	2	3	4	5
$G$	dB	-12	-6	0	+6	+12
$S =$	$(s_{syn},$	1	0	1	—	—
	$s_{bf})$	0	1	1	—	—

De même, la présence du signal  $y_{bf}(t)$  dans  $y_{mix}(t)$  est contrôlée par  $s_{bf}$ . Le signal de sortie  $y_{mix}(t)$  vaut :

$$y_{mix}(t) = G \cdot s_{syn} \cdot y_{syn}(t) + s_{bf} \cdot y_{bf}(t) \quad (\text{B.8})$$

Les variables  $s_{syn}$  et  $s_{bf}$  sont regroupées sous une variable  $S = (s_{syn}, s_{bf})$  qui peut prendre trois modalités :

1. Synthèse additive uniquement.
2. Bruit filtré uniquement.
3. Synthèse additive et bruit filtré.

Les valeurs des variables  $G$  et  $S$  sont décrites dans le tableau B.7.

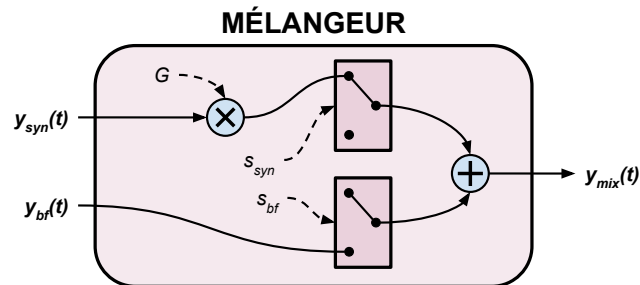


FIGURE B.17 – Schéma du module *MÉLANGEUR*.



# BIBLIOGRAPHIE

---

- ABDI, Hervé, Betty EDELMAN, Dominique VALENTIN et W Jay DOWLING (2009), *Experimental design and analysis for psychology*, Oxford University Press.
- ANTONY, Jiju (2014), *Design of experiments for engineers and scientists*, Elsevier.
- BANERJEE, Amit, Juan C QUIROZ et Sushil J LOUIS (2008), « A model of creative design using collaborative interactive genetic algorithms », in : *Design Computing and Cognition'08*, Springer, p. 397-416.
- BARRATT, Emma L et Nick J DAVIS (2015), « Autonomous Sensory Meridian Response (ASMR) : a flow-like mental state », in : *PeerJ* 3, e851.
- BEATTIE, David, Lynne BAILLIE et Martin HALVEY (2017), « Exploring how drivers perceive spatial earcons in automated vehicles », in : *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1.3, p. 1-24.
- BEATTIE, David, Lynne BAILLIE, Martin HALVEY et Rod MCCALL (2014), « What's around the corner ? Enhancing driver awareness in autonomous vehicles via in-vehicle spatial auditory displays », in : *Proceedings of the 8th nordic conference on human-computer interaction : fun, fast, foundational*, p. 189-198.
- BEATTIE, David, Baillie LYNNE, Martin HALVEY et Roderick MCCALL (2013), « Maintaining a sense of control in autonomous vehicles via auditory feedback », in.
- BERENGUT, David (2006), *Statistics for experimenters : design, innovation, and discovery*.
- BEZAT, Marie-Céline (2007), « Perception des bruits d'impact Application au bruit de fermeture de porte automobile », thèse de doct., Université de Provence-Aix-Marseille I.
- BRINTRUP, Alexandra Melike, Jeremy RAMSDEN, Hideyuki TAKAGI et Ashutosh TIWARI (2008), « Ergonomic chair design by fusing qualitative and quantitative criteria using interactive genetic algorithms », in : *IEEE Transactions on Evolutionary Computation* 12.3, p. 343-354.
- BRINTRUP, Alexandra Melike, Jeremy RAMSDEN et Ashutosh TIWARI (2007), « An interactive genetic algorithm-based framework for handling qualitative criteria in design optimization », in : *Computers in Industry* 58.3, p. 279-291.



- 
- CERA, Andrea et Nicolas MISDARIIS (2021), « Intrusiveness, Annoyance and Sound Design », in : *LINKs 5-6*, p. 184-189.
- CHEN, Kuan-Ting et Huei-Yen Winnie CHEN (2021), « Manipulating music to communicate automation reliability in conditionally automated driving : A driving simulator study », in : *International Journal of Human-Computer Studies* 145, p. 102518.
- ČREPINŠEK, Matej, Shih-Hsi LIU et Marjan MERNIK (2013), « Exploration and exploitation in evolutionary algorithms : A survey », in : *ACM computing surveys (CSUR)* 45.3, p. 1-33.
- DAHLSTEDT, Palle (2001), « A MutaSynth in parameter space : interactive composition through evolution », in : *Organised Sound* 6.2, p. 121-124.
- DANIEL, Peter et Reinhard WEBER (1997), « Psychoacoustical roughness : Implementation of an optimized model », in : *Acta Acustica united with Acustica* 83.1, p. 113-123.
- DEB, Kalyanmoy, Samir AGRAWAL, Amrit PRATAP et Tanaka MEYARIVAN (2000), « A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization : NSGA-II », in : *International conference on parallel problem solving from nature*, Springer, p. 849-858.
- DEFREVILLE, Boris (2005), « Caractérisation de la qualité sonore de l'environnement urbain : une approche physique et perceptive basée sur l'identification des sources sonores », thèse de doct., Université de Cergy Pontoise.
- DENJEAN, Sébastien (2015), « Sonification des véhicules électriques par illusions auditives : étude de l'intégration audiovisuelle de la perception du mouvement automobile en simulateur de conduite », thèse de doct., Aix-Marseille.
- DESOEUVRE, Gaylord, Florent RICHARD, Vincent ROUSSARIE et Marie Céline BEZAT (2008), « Hartis, a re-synthesis tool for vehicles sound design », in : *The Journal of the Acoustical Society of America* 123.5, p. 3247-3247.
- DU, Na, Jacob HASPIEL, Qiaoning ZHANG, Dawn TILBURY, Anuj K PRADHAN, X Jessie YANG et Lionel P ROBERT JR (2019), « Look who's talking now : Implications of AV's explanations on driver's trust, AV preference, anxiety and mental workload », in : *Transportation research part C : emerging technologies* 104, p. 428-442.
- DUTILLEUX, Pierre (1989), « Simple to operate digital time varying filters », in : *Audio Engineering Society Convention 86*, Audio Engineering Society.
- ELLERMEIER, Wolfgang, Markus MADER et Peter DANIEL (2004), « Scaling the unpleasantness of sounds according to the BTL model : Ratio-scale representation and psychoacoustical analysis », in : *Acta Acustica united with Acustica* 90.1, p. 101-107.

- 
- ESLING, Philippe, Naotake MASUDA, Adrien BARDET, Romeo DESPRES et Axel CHEMLA-ROMEU-SANTOS (2019), « Flow synthesizer : Universal audio synthesizer control with normalizing flows », in : *Applied Sciences* 10.1, p. 302.
- EU REGULATION (mai 2014), *Regulation (EU) No 540/2014 of the European Parliament and of the Council of 16 April 2014 on the Sound Level of Motor Vehicles and of Replacement Silencing Systems, and Amending Directive 2007/46/EC and Repealing Directive 70/157/EEC Text with EEA Relevance.*
- FAGERLÖNN, Johan, Pontus LARSSON et Justyna MACULEWICZ (2020), « The sound of trust : sonification of car intentions and perception in a context of autonomous drive », in : *International Journal of Human Factors and Ergonomics* 7.4, p. 343-358.
- FAUSTO, Fernando, Adolfo REYNA-ORTA, Erik CUEVAS, Ángel G ANDRADE et Marco PEREZ-CISNEROS (2020), « From ants to whales : metaheuristics for all tastes », in : *Artificial Intelligence Review* 53.1, p. 753-810.
- FUKUMOTO, Makoto et Koki HIJIKURO (2020), « Optimization of sound of autonomous sensory meridian response with interactive genetic algorithm », in : *2020 Joint 11th International Conference on Soft Computing and Intelligent Systems and 21st International Symposium on Advanced Intelligent Systems (SCIS-ISIS)*, IEEE, p. 1-6.
- FUKUMOTO, Makoto, Ryota YAMAMOTO et Shintaro OGAWA (2013), « The efficiency of interactive differential evolution in creation of sound contents : In comparison with interactive genetic algorithm », in : *International Journal of Software Innovation (IJSI)* 1.2, p. 16-27.
- GANG, Nick, Srinath SIBI, Romain MICHON, Brian MOK, Chris CHAFE et Wendy JU (2018), « Don't Be alarmed : Sonifying autonomous vehicle perception to increase situation awareness », in : *Proceedings of the 10th international conference on automotive user interfaces and interactive vehicular applications*, p. 237-246.
- GUSKI, Rainer, U FELSCHER-SUHR et R SCHUEMER (1999), « The concept of noise annoyance : How international experts see it », in : *Journal of sound and vibration* 223.4, p. 513-527.
- HUANG, Nicholas et Mounya ELHILALI (2017), « Auditory salience using natural soundscapes », in : *The Journal of the Acoustical Society of America* 141.3, p. 2163-2176.
- HWANG, Ching-Lai et Kwangsun YOON (1981), *Multiple Attribute Decision Making*, sous la dir. de M. BECKMANN et H. P. KÜNZI, t. 186, Lecture Notes in Economics and Mathematical Systems, Berlin, Heidelberg : Springer Berlin Heidelberg, ISBN : 978-3-540-10558-9 978-3-642-48318-9.

- 
- INTERNATIONAL, SAE (2021), *J3016\_202104 : Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, Warrendale.
- JOHNSON, Colin G (1999), « Exploring the sound-space of synthesis algorithms using interactive genetic algorithms », in : *Proceedings of the AISB'99 Symposium on Musical Creativity*, Society for the Study of Artificial Intelligence et Simulation of Behaviour, p. 20-27.
- KACZMAREK, Tomasz et Anna PREIS (2010), « Annoyance of time-varying road-traffic noise », in : *Archives of acoustics* 35.3.
- KERBER, Stefan et Hugo FASTL (2008), « Prediction of perceptibility of vehicle exterior noise in background noise », in : *Tagungsband Fortschritte der Akustik-DAGA# 2008, Dresden*, p. 623-624.
- LAGEAT, Thierry, Sandor CZELLAR et Gilles LAURENT (2003), « Engineering hedonic attributes to generate perceptions of luxury : Consumer perception of an everyday sound », in : *Marketing Letters* 14.2, p. 97-109.
- LANGEVELD, Lau, René van EGMOND, Reinier JANSEN et Elif ÖZCAN (2013), « Product sound design : Intentional and consequential sounds », in : *Advances in industrial design engineering* 47.3.
- LARSSON, Pontus, Justyna MACULEWICZ, Johan FAGERLÖNN et Max LACHMANN (2019), « Auditory displays for automated driving-challenges and opportunities », in.
- LARSSON, Pontus, Anders OPPERUD, Krister FREDRIKSSON et Daniel VÄSTFJÄLL (2009), « Emotional and behavioural response to auditory icons and earcons in driver-vehicle interfaces », in : *Proc. 21st International Technical Conference on Enhanced Safety of Vehicles, Germany*.
- LARSSON, Pontus et Daniel VÄSTFJÄLL (2013), « Emotional and behavioural responses to auditory interfaces in commercial vehicles », in : *International Journal of Vehicle Noise and Vibration* 9.1-2, p. 75-95.
- LAVANDIER, Catherine et Boris DEFRÉVILLE (2006), « The contribution of sound source characteristics in the assessment of urban soundscapes », in : *Acta acustica united with Acustica* 92.6, p. 912-921.
- LEBART, Ludovic, Alain MORINEAU et Marie PIRON (1995), *Statistique exploratoire multidimensionnelle*, t. 3, Dunod Paris.
- LEE, Sang Kwon, Seung Min LEE, Taejin SHIN et Manug HAN (2017), « Objective evaluation of the sound quality of the warning sound of electric vehicles with a consideration

- 
- of the masking effect : Annoyance and detectability », in : *International Journal of Automotive Technology* 18.4, p. 699-705.
- LEGRAND, Pierrick, Vincent PÉAN, Esther HARBOUN-COHEN, Jacques LEVY-VEHEL, Bruno FRACHET, Evelyne LUTTON, Pierre COLLET et al. (2007), « Interactive evolution for cochlear implants fitting », in : *Genetic Programming and Evolvable Machines* 8.4, p. 319-354.
- LEIBA, Raphaël (2017), « Conception d'un outil de diagnostic de la gêne sonore en milieu urbain », thèse de doct., Université Pierre et Marie Curie-Paris VI.
- LEJEUNE, Michel (2004), *Statistique : La théorie et ses applications*, Springer Science & Business Media.
- LOH, Wei-Liem (1996), « On Latin hypercube sampling », in : *The annals of statistics* 24.5, p. 2058-2080.
- MAFFIOLO, Valérie (1999), « De la caractérisation sémantique et acoustique de la qualité sonore de l'environnement urbain : structuration des représentations mentales et influence sur l'appréciation qualitative : application aux ambiances sonores de Paris », thèse de doct., Le Mans.
- MATA, Ivan, Georges FADEL, Anthony GARLAND et Winfried ZANKER (2018), « Affordance based interactive genetic algorithm (ABIGA) », in : *Design Science* 4.
- MATHEW, Thomas, Bimal K SINHA et André I KHURI (2011), *Statistical tests for mixed linear models*, John Wiley & Sons.
- MCDERMOTT, James, Michael O'NEILL et Niall JL GRIFFITH (2010), « Interactive EC control of synthesized timbre », in : *Evolutionary Computation* 18.2, p. 277-303.
- MEIGNAN, David, Sigrid KNUST, Jean-Marc FRAYRET, Gilles PESANT et Nicolas GAUD (2015), « A review and taxonomy of interactive optimization methods in operations research », in : *ACM Transactions on Interactive Intelligent Systems (TiIS)* 5.3, p. 1-43.
- MIKI, Mitsunori, Hiroko ORITA, Sanae H WAKE et Tomoyuki HIROYASU (2006), « Design of sign sounds using an interactive genetic algorithm », in : *2006 IEEE International Conference on Systems, Man and Cybernetics*, t. 4, IEEE, p. 3486-3490.
- MISDARIIS, Nicolas (2017), « Recherche-projet en design sonore : le cas emblématique du véhicule électrique », in : *Sciences du design* 5.1, p. 115-130.
- (2018), « Sciences du Design Sonore-Approche intégrée du design sonore au sein de la recherche en design », thèse de doct., Université de Technologie Compiègne (UTC).
- (2021), « Design Sonore & Transport-Le son de la mobilité », in.

- 
- MISDARIIS, Nicolas et Andrea CERA (2013), « Sound signature of Quiet Vehicles : state of the art and experience feedbacks », in : *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, t. 247, 5, Institute of Noise Control Engineering, p. 3333-3342.
- MISDARIIS, Nicolas, Andrea CERA, Eugenie LEVALLOIS et Christophe LOCQUETEAU (2012), « Do electric cars have to make noise? An emblematic opportunity for designing sounds and soundscapes », in : *Acoustics 2012*.
- MISDARIIS, Nicolas, Andrea CERA et William RODRIGUEZ (2019), « Electric and Autonomous Vehicle : from Sound Quality to Innovative Sound Design », in : *ICA-International Congress on Acoustics*.
- MISDARIIS, Nicolas, Anais GRUSON et Patrick SUSINI (2013), « Detectability study of warning signals in urban background noises : a first step for designing the sound of electric vehicles », in : *Proceedings of Meetings on Acoustics ICA2013*, t. 19, 1, Acoustical Society of America, p. 040032.
- MOREL, Julien, Catherine MARQUIS-FAVRE et L-A GILLE (2016), « Noise annoyance assessment of various urban road vehicle pass-by noises in isolation and combined with industrial noise : A laboratory study », in : *Applied Acoustics* 101, p. 47-57.
- NOMURA, Kota et Makoto FUKUMOTO (2017), « Asynchronous distributed interactive genetic algorithm for creating music melody reflecting multiple users' feelings », in : *2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, IEEE, p. 645-650.
- NORMAN DONALD, A (2013), *The design of everyday things*, MIT Press.
- PAGÈS, Jérôme et Emmanuel PÉRINEL (2007), « Blocs incomplets équilibrés versus plans optimaux », in : *Journal de la Société Française de Statistique* 148.2, p. 99-112.
- PARIZET, Etienne, Erald GUYADER et Valery NOSULENKO (2008), « Analysis of car door closing sound quality », in : *Applied acoustics* 69.1, p. 12-22.
- PARIZET, Etienne, Ryan ROBART, Wolfgang ELLERMEIER, Karl JANSSENS, Fabio BIANCARDI, Manfred HAIDER, David QUINN et Jean-Christophe CHAMARD (2014), « Warning sounds for electric vehicles », in : *Fisita World Automotive Congress 2014*, p. 1.
- PEI, Yan et Hideyuki TAKAGI (2018), « Research progress survey on interactive evolutionary computation », in : *Journal of Ambient Intelligence and Humanized Computing*, p. 1-14.

- 
- PETIOT, Jean-François, Francisco CERVANTES CHAVEZ, Ludivine BOIVIN et al. (2015), « A comparison of conjoint analysis and interactive genetic algorithms for the study of product semantics », in : *DS 80-5 Proceedings of the 20th International Conference on Engineering Design (ICED 15) Vol 5 : Design Methods and Tools-Part 1, Milan, Italy, 27-30.07. 15*, p. 011-020.
- PETIOT, Jean-François, Sébastien DENJEAN, Sébastien VILLA et Emmanuelle DIAZ (2020), « Design of warning sounds using an Interactive Genetic Algorithm », in : *Forum Acusticum*, p. 107-114.
- PETIOT, Jean-François, Bjørn G KRISTENSEN et Anja M MAIER (2013), « How should an electric vehicle sound? User and expert perception », in : *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, t. 55928, American Society of Mechanical Engineers, V005T06A028.
- PETIOT, Jean-François, Killian LEGEAY et Mathieu LAGRANGE (2019), « Optimization of the sound of electric vehicles according to unpleasantness and detectability », in : *Proceedings of the Design Society : International Conference on Engineering Design*, t. 1, 1, Cambridge University Press, p. 3949-3958.
- PETIOT, Jean-François et Bernard YANNOU (2004), « Measuring consumer perceptions for a better comprehension, specification and assessment of product semantics », in : *International Journal of industrial ergonomics* 33.6, p. 507-525.
- POIRSON, Emilie, Jean-François PETIOT, Emmanuel ALIOUAT, Ludivine BOIVIN et David BLUMENTHAL (2010), « Study of the convergence of Interactive Genetic Algorithm in iterative user's tests : application to car dashboard design », in : *International conference on Integrated Design and Manufacturing in Mechanical Engineering IDMME 2010*, p. 001-010.
- POIRSON, Emilie, Jean-François PETIOT, Ludivine BOIVIN et David BLUMENTHAL (2013), « Eliciting user perceptions using assessment tests based on an interactive genetic algorithm », in : *Journal of Mechanical Design* 135.3, p. 031004.
- POVEDA-MARTINEZ, Pedro, Ramon PERAL-ORTS, Nuria CAMPILLO-DAVO, Josue NESCOLARDESELVA, Miguel LLORET-CLIMENT et Jaime RAMIS-SORIANO (2017), « Study of the effectiveness of electric vehicle warning sounds depending on the urban environment », in : *Applied Acoustics* 116, p. 317-328.
- PREIS, Anna (1987), « Intrusive sounds », in : *Applied Acoustics* 20.2, p. 101-127.
- PUCKETTE, Miller (2007), *The Theory and Techniques of Electronic Music*, World Scientific Publishing Company.

- 
- PULKKI, Ville (2001), « Localization of amplitude-panned virtual sources II : Two-and three-dimensional panning », in : *Journal of the Audio Engineering Society* 49.9, p. 753-767.
- REFAEILZADEH, Payam, Lei TANG et Huan LIU (2009), « Cross-validation. », in : *Encyclopedia of database systems* 5, p. 532-538.
- ROADS, Curtis (1996), *The computer music tutorial*, MIT press.
- (2004), *Microsound*, The MIT Press.
- ROBART, Ryan, Etienne PARIZET, Jean-Christophe CHAMARD, Karl JANSSENS, Fabio BIANCARDI, Joseph SCHLITTENLACHER, Wolfgang ELLERMEIER, Perceval PONDROM, James COCKRAM, Paul SPEED-ANDREWS et al. (2013), « eVADER : A Perceptual Approach to Finding Minimum Warning Sound Requirements for Quiet Cars. », in : *AIA-DAGA 2013 Conference on Acoustics*, p. 1.
- SALTELLI, Andrea, Marco RATTO, Terry ANDRES, Francesca CAMPOLONGO, Jessica CARIBONI, Debora GATELLI, Michaela SAISANA et Stefano TARANTOLA (2008), *Global sensitivity analysis : the primer*, John Wiley & Sons.
- SCHAFER, Raymond Murray (2010), *Le paysage sonore : le monde comme musique*, Éditions Wildproject.
- SEAGO, Allan (2013), « A new interaction strategy for musical timbre design », in : *Music and human-computer interaction*, Springer, p. 153-169.
- SHAN, Songqing et G Gary WANG (2010), « Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions », in : *Structural and multidisciplinary optimization* 41.2, p. 219-241.
- SOUAILLE, Tom, Nicolas MISDARIIS, Jean-François PETIOT et Mathieu LAGRANGE (2022), « Conception interactive en design sonore—Réduction d’un espace de design appliquée aux sons d’usage du véhicule autonome », in : *16ème Congrès Français d’Acoustique, CFA2022*.
- SOUAILLE, Tom, Jean-François PETIOT, Mathieu LAGRANGE et Nicolas MISDARIIS (2021), « Extracting Design Recommendations from Interactive Genetic Algorithm Experiments : Application to the Design of Sounds for Electric Vehicles », in : *Proceedings of the Design Society* 1, p. 1567-1576.
- SOUAILLE, Tom, Jean-François PETIOT, Nicolas MISDARIIS et Mathieu LAGRANGE (2022), « An interactive bi-objective optimisation process to guide the design of electric vehicle warning sounds », in : *Design Science* 8, e26.

- 
- STEINBACH, Lisa et M Ercan ALTINSOY (2019), « Prediction of annoyance evaluations of electric vehicle noise by using artificial neural networks », in : *Applied Acoustics* 145, p. 149-158.
- SUSINI, Patrick, Olivier HOUIX et Nicolas MISDARIIS (2014), « Sound design : an applied, experimental framework to study the perception of everyday sounds », in : *The New Soundtrack* 4.2, p. 103-121.
- SUSINI, Patrick, Guillaume LEMAITRE et Stephen MCADAMS (2012), « Psychological measurement for sound description and evaluation », in : *Measurements with persons : Theory, methods, and implementation areas* 227.
- SUSINI, Patrick, Stephen MCADAMS, Suzanne WINSBERG, Ivan PERRY, Sandrine VIEILLARD et Xavier RODET (2004), « Characterizing the sound quality of air-conditioning noise », in : *Applied acoustics* 65.8, p. 763-790.
- SYLCOTT, Brian, Jeremy J MICHALEK et Jonathan CAGAN (2015), « Exploring the role of interaction effects in visual conjoint analysis », in : *Journal of Mechanical Design* 137.9, p. 094503.
- TAKAGI, Hideyuki (2001), « Interactive evolutionary computation : Fusion of the capabilities of EC optimization and human evaluation », in : *Proceedings of the IEEE* 89.9, p. 1275-1296.
- (2012), « Interactive evolutionary computation for analyzing human awareness mechanisms », in : *Applied Computational Intelligence and Soft Computing* 2012.
- TAKAGI, Hideyuki et Miho OHSAKI (2007), « Interactive evolutionary computation-based hearing aid fitting », in : *IEEE Transactions on Evolutionary Computation* 11.3, p. 414-427.
- TSENG, Ian, Jonathan CAGAN et Kenneth KOTOVSKY (2012), « Concurrent optimization of computationally learned stylistic form and functional goals », in.
- UN REGULATION (jan. 2017), *Regulation No 138 of the Economic Commission for Europe of the United Nations (UNECE) — Uniform Provisions Concerning the Approval of Quiet Road Transport Vehicles with Regard to Their Reduced Audibility [2017/71]*.
- WAKEFIELD, Gregory H, Chris van den HONERT, Wendy PARKINSON et Sean LINEAWEAVER (2005), « Genetic algorithms for adaptive psychophysical procedures : recipient-directed design of speech-processor MAPs », in : *Ear and hearing* 26.4, 57S-72S.
- WILSON, Alex, Bruno M FAZENDA et al. (2016), « An evolutionary computation approach to intelligent music production informed by experimentally gathered domain



- 
- knowledge », in : *Proceedings of the 2nd AES Workshop on Intelligent Music Production*.
- WILSON, Alex et Bruno M FAZENDA (2019), « User-guided rendering of audio objects using an interactive genetic algorithm », in : *Journal of the Audio Engineering Society* 67.7/8, p. 522-530.
- WISHNICK, Aaron (2014), « Time-Varying Filters for Musical Applications. », in : *DAFx*, p. 69-76.
- XIN, Bin, Lu CHEN, Jie CHEN, Hisao ISHIBUCHI, Kaoru HIROTA et Bo LIU (2018), « Interactive multiobjective optimization : A review of the state-of-the-art », in : *IEEE Access* 6, p. 41256-41279.
- YEE-KING, Matthew John (2016), « The use of interactive genetic algorithms in sound design : a comparison study », in : *Computers in Entertainment : Special Issue on Musical Metacreation*.
- ZATTRA, Laura, Nicolas DONIN, Nicolas MISDARIIS, Frank PECQUET et David FIERRO (2021), « Sound Design as Viewed by Sound Designers : A Questionnaire About People, Practice and Definitions », in : *Doing Research in Sound Design*, Focal Press, p. 297-323.
- ZITZLER, Eckart (1999), *Evolutionary algorithms for multiobjective optimization : Methods and applications*, t. 63, Citeseer.
- ZITZLER, Eckart et Lothar THIELE (1999), « Multiobjective evolutionary algorithms : a comparative case study and the strength Pareto approach », in : *IEEE transactions on Evolutionary Computation* 3.4, p. 257-271.
- ZWICKER, Eberhard et Hugo FASTL (2013), *Psychoacoustics : Facts and models*, t. 22, Springer Science & Business Media.
- ZWICKER, Eberhard, Hugo FASTL, Ulrich WIDMANN, Kenji KURAKATA, Sonoko KUWANO et Seiichiro NAMBA (1991), « Program for calculating loudness according to DIN 45631 (ISO 532B) », in : *Journal of the Acoustical Society of Japan (E)* 12.1, p. 39-42.



---

**Titre :** Conception Interactive en Design Sonore

**Mots clés :** Design sonore, Évaluation perceptive, Algorithme génétique interactif, Expérience utilisateur, Véhicule autonome, Optimisation multi-objectifs

**Résumé :** La problématique de cette thèse porte sur la définition d'outils et de méthodes interactives pour le design sonore, permettant d'intégrer l'expérience utilisateur dans le processus de conception. Nous proposons un paradigme expérimental d'étude de la perception sonore, basé l'optimisation interactive de sons. En particulier, nous utilisons des Algorithmes Génétiques Interactifs (IGAs) pour résoudre deux problèmes de design sonore: le design de sons d'alerte de véhicules électriques ou hybrides facilement détectables et peu désagréables, ainsi que le design de sons d'intérieur peu intrusifs pour le véhicule autonome. Au travers de ces exemples, nous montrons l'intérêt de l'utilisation des IGAs pour le paradigme proposé.

Nous proposons également une méthode d'analyse et de généralisation des résultats obtenus individuellement lors d'une expérience d'optimisation interactive. Nous montrons expérimentalement que cette méthode permet de produire des recommandations de design sonore pertinentes pour répondre au problème d'optimisation. Enfin, nous proposons une méthode de réduction informée d'un espace de conception, préalable à une expérience d'optimisation interactive. Nous montrons expérimentalement que réduire l'espace de conception avec cette méthode permet d'améliorer la qualité des sons trouvés au cours du processus d'optimisation.

---

**Title :** Interactive Methods for Sound Design

**Keywords :** Sound design, Subjective evaluation, Interactive genetic algorithm, User experience, Autonomous vehicle, Multi-objective optimization

**Abstract :** This thesis deals with the definition of interactive tools and methods for sound design, that involve user experience in the design process. We propose an experimental paradigm for studying sound perception, based on the interactive optimization of sounds. In particular, we use Interactive Genetic Algorithms (IGAs) to solve two optimization problems: the design of the exterior sound of an electric or hybrid vehicle, that is easily detectable while not being too unpleasant, as well as the design of sounds for the interior of an autonomous vehicle. Through these examples, we show that IGAs are promising methods to be used in the proposed paradigm.

We also propose a method for analyzing and generalizing the individual results obtained during an interactive optimization experiment. Through experiments, we show that the proposed method allows for making sound design recommendations that are relevant for the optimization problem. Finally, we propose an informed method for reducing the design space prior to an interactive optimization experiment. We show, through experiments, that reducing the design space with the proposed method improves the quality of the sounds found during the optimization process.