

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ  
СІКОРСЬКОГО»

Факультет електроніки

Кафедра акустичних та мультимедійних електронних систем

«На правах рукопису»  
УДК 681.8

«До захисту допущено»

Завідувач кафедри



С.А. Найда

(ініціали, прізвище)

« 06 » січня 2024 р.

**Магістерська дисертація**

зі спеціальності 171 Електроніка

(код і назва)

на тему: Акустичні основи якісного звукозапису

Виконав: студент II курсу, групи ДГ-21мп

(шифр групи)

Дойна Олексій Ярославович

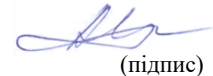
(прізвище, ім'я, по батькові)



(підпис)

Керівник: професор, д.т.н., професор Продеус А.М.

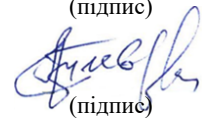
(посада, науковий ступінь, вчене звання, прізвище та ініціали)



(підпис)

Рецензент: к.т.н., доцент кафедри електронної інженерії Шуляк О.П.

(посада, науковий ступінь, вчене звання, прізвище та ініціали)



(підпис)

Засвідчую, що у цій дипломній роботі  
немає запозичень з праць інших авторів  
без відповідних посилань.

Студент:



(підпис)

Київ – 2024

**Національний технічний університет України**  
**«Київський політехнічний інститут імені Ігоря Сікорського»**

**Факультет електроніки**

**Кафедра акустичних та мультимедійних електронних систем**

Рівень вищої освіти – другий (магістерський)

Спеціальність – 171 «Електроніка»

Освітньо-професійна програма «Акустичні електронні системи та технології обробки акустичної інформації»

ЗАТВЕРДЖУЮ

Завідувач кафедри

\_\_\_\_\_ Сергій НАЙДА

« 01 » \_\_\_\_\_ 09 \_\_\_\_\_ 20 23р.

**ЗАВДАННЯ**

**на магістерську дисертацію студенту**

**Дойна Олексій Ярославович**

1. Тема дисертації «Акустичні основи якісного звукозапису», науковий керівник дисертації Продеус Аркадій Миколайович д. т. н., затверджені наказом по університету від « \_\_\_ » \_\_\_\_\_ 2023р. № \_\_\_\_\_

Музичні записи в концертній залі бажано робити із використанням пари мікрофонів штучної голови (прямий метод). Альтернативний (непрямий) метод двоканального запису полягає у використанні лише одного мікрофона для запису, із подальшою згорткою отриманого сигналу із бінауральною імпульсною характеристикою (БІХ) приміщення. Метою роботи є дослідження якості двоканальних записів музики, отриманих непрямим методом. Якість записів, отриманих непрямим методом, має оцінюватися суб'єктивним та об'єктивним (інструментальним) методами.

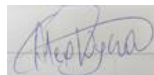
9. Дата видачі завдання 01.09.2023

---

## Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Аналіз літературних джерел	04.09–10.09	Виконано
2	Планування досліджень	11.09–17.09	Виконано
3	Підготовка музичного матеріалу	18.09–24.09	Виконано
4	Підготовка програмно-апаратного забезпечення	25.09–01.10	Виконано
5	Підготовка еталонних 2-канальних записів музики та 1-канальних дослідницьких записів	02.10–08.10	Виконано
6	Запис БІХ приміщення та створення за їх допомогою 2-канальних дослідницьких записів	09.10–15.10	Виконано
7	Суб'єктивна та об'єктивна оцінка якості отриманих 2-канальних записів. Опис отриманих результатів	16.10–22.10	Виконано
8	Оформлення 1-го та 2-го розділів магістерської дисертації, включно із переліком літературних джерел	23.10–26.10	Виконано

Студент



Олексій Дойна

Науковий керівник:



професор, д.т.н., професор Продеус А.М

## РЕФЕРАТ

*Дойна О. Я. Акустичні основи якісного звукозапису: магістерська дис. 171 Електроніка / Дойна Олексій Ярославович. – Київ, 2024. – 77 с.*

Ключові слова: бінауральне аудіо, імпульсні характеристики приміщення, згортка сигналу, суб'єктивне та об'єктивне оцінювання.

**Актуальність роботи.** Сучасний аудіофільський світ постійно розвивається — розширюються можливості сприйняття, відтворення та запису музики та аудіо-композицій людиною. Бінауральне аудіо здатне надати унікальний досвід прослуховування звуку, відтвореного в просторовій реальності, схожій на ту, що ми спостерігаємо в повсякденному житті. Бінауральне аудіо базується на концепції відтворення звуку так, як його сприймає слухова система людини в реальному часі в певному оточенні.

**Мета і завдання дослідження.** Метою даної роботи є дослідження можливостей створення бінаурального аудіо за допомогою одноканального дослідницького запису та бінауральних імпульсних характеристик (БІХ) приміщення.

Для досягнення поставленої мети необхідно вирішити такі завдання:

- Підготувати музичний матеріал
- Підготувати програмне забезпечення
- Підготувати еталонні 1-канальні дослідницькі записи
- Записати БІХ приміщення та створити 2-канальні дослідницькі записи
- Виконати суб'єктивну та об'єктивну оцінку якості отриманих 2-канальних записів
- Описати отримані результати

**Об'єкт дослідження.** Процес обробки одноканального запису звуку для створення синтезованого стереозапису.

**Предмет дослідження.** Порівняння якості синтезованого та реального стереозаписів із метою визначення ефективності розглянутого методу синтезу.

**Методи дослідження.** Експериментальні дослідження якості стереозаписів, отриманих шляхом згортки одноканального запису із записом бінауральної імпульсної характеристики приміщення.

## ABSTRACT

*Doina O. Y. Acoustic bases of high-quality sound recording: master's thesis. 171 Electronics / Oleksii Doyna - Kyiv, 2024. 77 p.*

Keywords: binaural audio, room impulse response, signal convolution, subjective and objective evaluation.

**Relevance of the work.** The modern audiophile world is constantly evolving - the possibilities of human perception, reproduction and recording of music and audio compositions are expanding. Binaural audio is able to provide a unique experience of listening to sound reproduced in a spatial reality similar to what we observe in everyday life. Binaural audio is based on the concept of reproducing sound as it is perceived by our ears in real time in a certain environment.

**The purpose and objectives of the study.** The purpose of this work is to investigate the possibilities of creating binaural audio using a single-channel research recording and a BIC room.

To achieve this goal, we need to solve the following tasks:

- Prepare musical material
- Prepare the software
- Prepare reference 1-channel test recordings
- Record the BIC room and create 2-channel research recordings
- Perform a subjective and objective assessment of the quality of the received 2-channel recordings
- Describe the results obtained

**Object of research.** The difference between real and artificial binaural recordings.

The subject of the study. Modern methods and tools for creating binaural recordings.

**Research methods.** Analysis of specialized literature on the subject of recording and audio processing, analysis and comparison of recording methods and creation of such recordings.

## Зміст

Вступ .....	10
Розділ 1: Аналіз літературних джерел та планування досліджень .....	12
1.1 Сучасний стан .....	12
1.3 Планування досліджень .....	18
1.4 Висновки до Розділу 1 .....	20
Розділ 2: Підготовка програмно-апаратного забезпечення та запис музичного матеріалу .....	21
2.1 Підготовка музичного матеріалу .....	21
2.2 Підготовка програмно-апаратного забезпечення.....	22
2.2.1 Апаратне забезпечення.....	22
2.2.2 Програмне забезпечення .....	30
2.3 Підготовка еталонних 2-канальних записів музики та 1-канальних дослідницьких записів.....	31
2.5 Висновки до 2 розділу .....	33
Розділ 3: Оцінювання якості 2-канальних дослідницьких записів .....	34
3.1 Суб'єктивне оцінювання якості 2-канальних дослідницьких записів за шкалою DMOS .....	35
3.2 Об'єктивне оцінювання якості 2-канальних дослідницьких записів .....	38
3.3 Отримані результати .....	40
3.4 Висновки до 3 розділу .....	54
Додатки.....	60
Додаток 1: Програма zgortka.m.....	60



Додаток 2: Програма ekspertiza3_m.m .....	62
Додаток 3: Програма filterlow.m.....	64
Додаток 4: Програма ocinka_ser.m .....	65
Додаток 5: Програма ocinka_vidpovidnosti.m.....	66
Додаток 6: Програма obj_quality.m .....	68
Додаток 7: Програма srr_func_correct2.m .....	72
Додаток 8: Програма lsd_func.m.....	73
Додаток 9: Програма enframe.m .....	74
Додаток 10: Програма l_rfft.m .....	777

## Вступ

Факсимільне або правдиве відносно оригіналу відтворення звукових подій викликає великий інтерес в акустиці та суміжних галузях і досліджується вже багато років. Одним із досягнень внаслідок цих дій є бінауральна технологія. Багато хто вважає бінауральну технологію дуже сучасною та новою, а дехто навіть вважає, що вона пов'язана суто з дослідженнями якості звуку і була винайдена виключно для них. Однак бінауральна технологія, особливо технологія запису, існує вже досить давно, і, насправді, перші кроки були зроблені приблизно в 1880 році.

Сучасний аудіофільський світ постійно розвивається — розширюються можливості сприйняття, відтворення та запису музики та аудіо-композицій людиною. Бінауральне аудіо здатне надати унікальний досвід прослуховування звуку, відтвореного в просторовій реальності, схожій на ту, що ми спостерігаємо в повсякденному житті. Бінауральне аудіо базується на концепції відтворення звуку так, як його сприймають наші вуха в реальному часі в певному оточенні.

Як це працює? Основні сигнали є "бінауральними" за своєю природою, що впливає з фізичного розділення між нашими двома вухами. Коли джерело звуку зміщується від центру, виникає ледь помітна затримка в часі між сигналом, що досягає ближнього вуха, і дальнього. Ця "міжауральна різниця в часі" ("Interaural Time Difference" або ITD) створює залежну від довжини хвилі різницю фаз, яку мозок миттєво інтерпретує як сигнал про напрямок руху. Чим правіше знаходиться джерело, тим більша затримка між звуком, що досягає правого вуха, і лівим. Різницю фаз стає важче виявити на високих частотах, тому цей сигнал найбільш ефективний на частотах нижче 1-1,5 кГц.

Але на додаток до цієї простої різниці в часі, фізична маса вашої голови поглинає звукову енергію, коли хвилі проходять від найближчого до джерела вуха до далекого. Зручно, що цей ефект затінення - "міжвушна різниця рівнів" ("Interaural Level Difference" або ILD) - діє в основному для високих частот, які поглинаються більше, ніж низькі частоти, які мають тенденцію до дифракції.

Отже, ця технологія використовує два мікрофони, розташовані аналогічно та пропорційно до положення вух людини, щоб створити стереозвучання максимально точно відображаюче будь-яке оточення. Але це не єдиний варіант розташування мікрофонів так само, як і не завжди використовується їх однакова кількість. Наприклад, музичні записи в концертній залі бажано проводити із використанням пари мікрофонів штучної голови (прямий метод). Альтернативний же (непрямий) метод двоканального запису полягає у використанні лише одного мікрофона для запису, із подальшою згортокою отриманого сигналу із бінауральною імпульсною характеристикою (БІХ) приміщення.

Метою даної роботи є дослідження якості двоканальних записів музики, отриманих непрямим методом. Якість записів, отриманих непрямим методом, буде оцінюватися суб'єктивним та об'єктивним (інструментальним) методами.

## **Розділ 1: Аналіз літературних джерел та планування досліджень**

### **1.1 Сучасний стан**

Актуальний стан теми бінаурального аудіо можна назвати достатньо захопливим, оскільки в наш час, на відміну від років відкриття та перших досліджень, технології та ресурси в області бінаурального аудіо стали більш ефективними та доступними. Бінауральне аудіо - це метод запису та відтворення звуку, який створює просторовий звуковий ефект, подібний до того, що сприймає людське вухо. Це дозволяє слухачам відчувати звук так, як він сприймається в реальному часі, що робить його дуже популярним у використанні для записів концертів. Запис концертів на штучні голови - це один із типів застосування бінаурального аудіо.

Традиційно записи здійснювалися двома методами: моно і стерео. При монозаписі використовується один мікрофон для запису звуку, а при стереозаписі - два мікрофони, віддалені один від одного. Бінауральний запис просуває стереозапис на крок вперед, розміщуючи два мікрофони у вушних порожнинах по обидва боки штучної голови. Оскільки вона відтворює щільність і форму людської голови, ці мікрофони вловлюють і обробляють звук точно так, як його чують людські вуха, зберігаючи міжвушні сигнали. Ефект найкраще відчувається в навушниках, з чітким розрізненням лівої та правої перспективи. Це проста концепція, але коли її реалізовано за допомогою високоякісних мікрофонів і відтворено через такі ж високоякісні навушники, ефект до незбагненності реалістичний: мозок лишається обманутим і змушеним повірити, що він чує звуки немов з перших вуст.

Однак для досягнення реалістичного звучання не тільки сама техніка важлива, але й належний аналіз приміщення, де проводиться запис. Останнє

десятиліття свідчить про збільшення інтересу до бінауральної імпульсної характеристики (БІХ) приміщення.

Бінауральна імпульсна характеристика приміщення - це важливий аспект при створенні реалістичного бінаурального аудіо. Ця характеристика вказує, як звук розсіюється в приміщенні і впливає на звуковий спектр під час запису. Сучасні програмні засоби дозволяють вимірювати БІХ приміщення для точного відтворення просторового звуку. Зазвичай, для цього використовують спеціальні апаратні та програмні засоби для аналізу відгуків приміщення на різних частотах. Це допомагає забезпечити високу якість звуку при відтворенні. Загалом, бінауральне аудіо і вивчення БІХ приміщення є цікавою областю досліджень та застосувань, особливо в сучасному світі, де вимоги до якості звуку щоразу зростають. Ці технології відкривають нові можливості для віртуальної реальності, аудіовізуального мистецтва та багатьох інших галузей.

Тема даного дослідження відрізняється від інших кількома важливими аспектами. У даному дослідженні буде розглянуто якість двоканальних записів музики, отриманих непрямым методом. У роботі використовуватиметься лише один мікрофон для запису, а потім проводитиметься згортка отриманого сигналу із бінауральною імпульсною характеристикою (БІХ) приміщення. Цей метод відмінний від традиційного прямого методу, де використовується пара мікрофонів, прикріплених до штучної голови.

Основною метою даної роботи є дослідження та порівняння якості звуку, отриманого за допомогою використання цього альтернативного методу запису зі стандартним прямим методом. У відмінну від більшості попередніх досліджень, буде оцінено якість записів не лише з точки зору об'єктивних параметрів, але і за допомогою суб'єктивних тестів, де реакція слухачів відіграє ключову роль у визначенні якості аудіо. Такий підхід дозволяє ширше

та прискіпливіше розглянути переваги та недоліки альтернативного методу запису, і визначити ситуації, в яких він може бути найбільш корисним.

Ця тема вивчалась у наступних джерелах:

У книзі "Spatial Audio" авторства Francis Rumsey та Tim McCormick [8] - ретельно розглянуто різні аспекти просторового аудіо, включаючи багато аспектів бінаурального аудіо. Нижче наведено детальний огляд результатів, представлених в цій книзі.

Основні поняття просторового звуку: в книзі розглянуто основні поняття, пов'язані з просторовим аудіо, включаючи стерео, мультіканальний та бінауральний звук. Автори розкривають, як звук поширюється у просторі і як цей процес впливає на сприйняття звуку;

Бінауральне аудіо та його принципи: в книзі пояснюється, як працює бінауральне аудіо, включаючи принципи, за якими вухоподібні мікрофони розташовані на голові, а також яким чином відбувається запис та відтворення звуку для досягнення просторового ефекту.

Методи запису бінаурального аудіо: в книзі досліджуються різні методи запису бінаурального аудіо, включаючи технічні аспекти та технології, що використовуються для створення бінауральних записів.

Сприйняття звуку та акустичні аспекти: в книзі розглянуто, як сприймається звук людськими вухами та як акустичні аспекти впливають на сприйняття звуку.

Застосування в різних сферах: автори виносять окремою темою для обговорення різні сфери діяльності, де бінауральне аудіо знаходить своє застосування, включаючи віртуальну реальність, аудіоінженерію, музику, аудіовізуальні медіа та багато інших.

Практичні приклади та технічні вирішення: в книзі надано приклади реальних сценаріїв та технічних вирішень для досягнення просторового звуку

через бінауральний метод. Автори докладно досліджують багато аспектів бінаурального аудіо та його ролі в створенні просторового звуку.

У книзі "The Technology of Binaural Listening" авторства William M. Hartmann [12] - ретельно досліджено різні аспекти бінаурального аудіо, розкриваючи технічні та психоакустичні деталі, а саме:

Природа бінаурального слухання: автором розглянуто фізіологічні аспекти бінаурального слухання та те, як людські вуха сприймають звук в просторі. Він вивчає, як відбувається аналіз різниці в часі та інтенсивності між звуком, який досягає лівого та правого вуха.

Принципи роботи бінауральних систем: автор докладно розглядає технічні аспекти бінауральних систем, включаючи принципи функціонування вухоподібних мікрофонів, які використовуються для запису бінаурального звуку.

Бінауральний запис та відтворення: у книзі розкриваються методи запису та відтворення бінаурального аудіо, включаючи технічні аспекти, які важливі для досягнення просторового звуку.

Взаємодія звуку та акустичні аспекти: автор вивчає, як звук взаємодіє з акустичними характеристиками простору та як це впливає на сприйняття звуку.

Психоакустичні аспекти бінаурального слухання: досліджується, як люди сприймають бінауральний звук, включаючи важливі аспекти психоакустики, такі як сприйняття напрямку та відстані звуку.

Застосування бінаурального аудіо: автор розглядає різні сфери та застосування бінаурального аудіо, включаючи віртуальну реальність, аудіоінженерію, музику, аудіовізуальні медіа та багато інших.

У книзі "Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio" авторства Agnieszka Roginska та Paul Geluso [6] - розкриваються:

Принципи мультимедійного аудіо: автори розпочинають з основ мультимедійного аудіо та розглядають, як звук створюється та сприймається в просторі.

Бінауральний та багатоканальний звук: автори детально досліджують обидва аспекти звуку, розкриваючи, як вони взаємодіють та відрізняються один від одного.

Технології запису та відтворення: у книзі розглянуто технічні аспекти запису та відтворення бінаурального та багатоканального звуку, включаючи використання вухоподібних мікрофонів та акустичні системи.

Психоакустичні аспекти: автори досліджують сприйняття звуку людьми, включаючи психоакустичні аспекти, такі як відмінності у сприйнятті напрямку та відстані.

Застосування в мистецтві та науці: в книзі розглянуто використання бінаурального та багатоканального аудіо в різних сферах, включаючи музику, аудіовізуальні медіа, віртуальну реальність та наукові дослідження.

Практичні поради та приклади: автори надають конкретні приклади та практичні поради щодо створення і відтворення бінаурального та багатоканального аудіо.

У книзі Art of sound reproduction авторства Watkinson J. [15] - ретельно досліджується багато аспектів бінаурального аудіо, включаючи технічні, психоакустичні та практичні аспекти, а саме:

Основи бінаурального аудіо: Книга починається з основ бінаурального аудіо, де розглянуто, як звук відтворюється у вухах слухача та як цей процес впливає на сприйняття звуку.



Технічні аспекти бінаурального запису: автори ґрунтовно вивчають технічні аспекти бінаурального запису, включаючи використання вухоподібних мікрофонів та обробку аудіоданих для досягнення просторового звуку.

Психоакустичні аспекти: у книзі досліджується, як люди сприймають бінауральний звук, включаючи важливі аспекти психоакустики, такі як сприйняття напрямку та відстані.

Акустичні характеристики: авторами розглядається вплив акустичних особливостей приміщення на бінауральний звук та як вони впливають на його якість та сприйняття.

Застосування бінаурального аудіо: в книзі обговорюються різні сфери для застосування бінаурального аудіо, включаючи віртуальну реальність, аудіоінженерію, музику, аудіовізуальні медіа та багато інших.

Практичні приклади та технічні рішення: автори надають конкретні приклади та практичні поради для створення і відтворення бінаурального аудіо.

У книзі *Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality* авторства Vorländer M. [14] - ретельно досліджено багато аспектів бінаурального аудіо та сприйняття звуку в реальних та віртуальних середовищах, а саме:

Психофізіологія слуху: автор розглядає основи психофізіології слуху та те, як люди сприймають звук в реальних та віртуальних середовищах.

Бінауральне слухання: автор ретельно досліджує, як працює бінауральне слухання, включаючи важливі аспекти відмінностей у часі та інтенсивності звуку, які досягають лівого та правого вуха.

Акустичні особливості простору: у книзі обговорюється вплив акустичних особливостей реальних та віртуальних середовищ на сприйняття звуку та створення просторового враження.

Звукові поля в реальних та віртуальних середовищах: автор досліджує створення звукових полів у різних середовищах та техніки їх симуляції віртуальних просторів.

Застосування в аудіовізуальних медіа та віртуальній реальності: у книзі вивчається використання бінаурального аудіо в аудіовізуальних медіа, віртуальній реальності, аудіоінженерії та інших сферах.

### **1.3 Планування досліджень**

Планування досліджень є ключовою складовою будь-якого наукового дослідження. Цей пункт буде зосереджений на важливих аспектах, пов'язаних з плануванням та організацією нашого дослідження як з урахуванням теоретичних аспектів, так і практичних завдань. Мета дослідження полягає в ретельному аналізі якості двоканальних записів музики, отриманих непрямим методом, і порівнянні їх із стандартними прямими методами.

Вище наведений план включає в себе наступні етапи:

#### **1. Підготовка музичного матеріалу**

Для дослідження обрано 5 творів класичної та 5 творів популярної музики. Кожен запис має тривалість 15-20 секунд. Музичний матеріал перетворюється на 1-канальні записи для відтворення через один гучномовець.

#### **2. Підготовка програмно-апаратного забезпечення**

Пункт включає в себе апаратне та програмне обладнання, необхідне для відтворення та запису звуку. Буде визначено обладнання для випромінення та записування сигналів, включаючи мікрофони, звукову карту та ноутбук. На

стороні програмного забезпечення, для аналізу та обробки сигналів, використовується Matlab та звуковий редактор, такий як Audacity.

### 3. Підготовка еталонних та дослідницьких записів

Включає запис еталонних 2-канальних музичних сигналів із використанням 2-х мікрофонів та штучної голови (ШГ), а також запис 1-канальних дослідницьких записів (без ШГ, 1 мікрофон).

### 4. Запис БІХ приміщення та створення за їх допомогою 2-канальних дослідницьких записів

Буде здійснено запис бінауральної імпульсної характеристики приміщення з використанням 2-х мікрофонів та ШГ. Запис буде використовуватися для обробки 1-канальних записів музики та створення 2-канальних дослідницьких записів.

### 5. Суб'єктивне оцінювання якості 2-канальних дослідницьких записів

Слухачі проведуть суб'єктивне оцінювання, порівнюючи еталонні та дослідницькі 2-канальні записи музики за 5-бальною шкалою.

### 6. Об'єктивне оцінювання якості 2-канальних дослідницьких записів

Будуть застосовані спеціальні комп'ютерні програми для порівняння еталонних та дослідницьких записів.

### 7. Опис отриманих результатів

Підсумковий аналіз результатів представляє собою побудову графіків, коментарів та висновків з отриманих суб'єктивних та об'єктивних оцінок якості записів.

## 1.4 Висновки до Розділу 1

В результаті аналізу сучасного стану досліджень у галузі бінаурального аудіо та імпульсних характеристик приміщення можна визначити важливість обраної теми і необхідності додаткових досліджень. Сучасні технології та акустичні методи дозволяють створювати звукові простори, максимально наближені до реального звукового середовища.

З аналізу літературних джерел видно, що багато авторів докладно вивчають технічні, психоакустичні та практичні аспекти бінаурального аудіо та імпульсних характеристик приміщення.

Планування досліджень включає докладну підготовку музичного матеріалу, вибір програмно-апаратного забезпечення, підготовку еталонних та дослідницьких записів, запис бінауральних імпульсних характеристик приміщення та проведення суб'єктивного та об'єктивного оцінювання якості записів. Цей план дозволить отримати об'єктивні дані, щодо якості двоканальних записів музики, порівняти їх із стандартними прямими методами та визначити переваги та недоліки використання бінаурального аудіо та імпульсних характеристик приміщення.

У підсумку, план досліджень ретельно підготований та охоплює всі необхідні кроки для досягнення поставлених цілей. Він дозволить систематично дослідити важливі аспекти бінаурального аудіо та імпульсних характеристик приміщення і отримати об'єктивні результати для подальшого аналізу та висновків.

## Розділ 2: Підготовка програмно-апаратного забезпечення та запис музичного матеріалу

### 2.1 Підготовка музичного матеріалу

Музичний матеріал, що використовується при проведенні досліджень – це 5 творів класичної музики та 5 творів популярної музики. Тривалість відрізка кожного запису – 15-20 секунд. Вихідний музичний матеріал має наступні параметри: 44100 Hz; 24 bit; 991kb/s; 2 channels.

Були обрані наступні твори:

1. “Get Lucky” – Daft Punk;
2. “Neon” – John Mayer;
3. “Grenade” – Bruno Mars;
4. “Stay” – Rihanna;
5. “Clair de Lune” – Claude Debussy;
6. “Piano Sonata No. 14” (Op.27, No. 2) – Ludwig van Beethoven;
7. “Nocturne in E Flat Major” (Op. 9 No. 2) – Frédéric Chopin;
8. “Winter” (The Four Seasons) – Antonio Vivaldi;
9. “Hungarian Rhapsody No. 2” – Franz Liszt;
10. “Right now” – Rihanna.

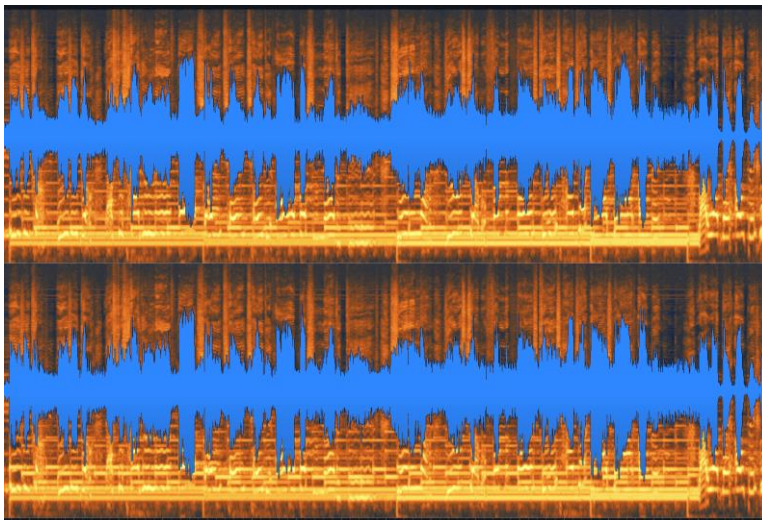


Рис. 1 – Спектрограма вихідного музичного матеріалу

## 2.2 Підготовка програмно-апаратного забезпечення

### 2.2.1 Апаратне забезпечення

Для випромінення сигналів потрібно:

- Акустично підготовлене приміщення із студійними гучномовцями;

Для запису 2-канального сигналу потрібно:

- 2 мікрофони;
- штучну голову
- зовнішню звукову карту;
- ноутбук.

Гучномовець, через який буде відтворений запис, це RCF Aura 5 із наступними технічними характеристиками [18]:

#### **Acoustical specifications:**

Frequency Response: 50 Hz - 20000 Hz

Max SPL @ 1m: 110 dB

Horizontal coverage angle: 110°

Vertical coverage angle: 70°

#### **Transducers:**

Dome Tweeter: 1.0" neo, 1.0" v.c

Woofers: 5.0", 1.0" v.c

Processor section

Crossover Frequencies: 2000Hz

Limiter: Soft Limiter

#### **Power section:**

Total Power: 200 W Peak, 100 W RMS

High frequencies: 50 W Peak, 25 W RMS

Low frequencies: 150 W Peak, 75 W RMS

Connections: XLR, Jack, RCA



Рис. 2 — Гучномовець RCF Аура Pro5 [18]

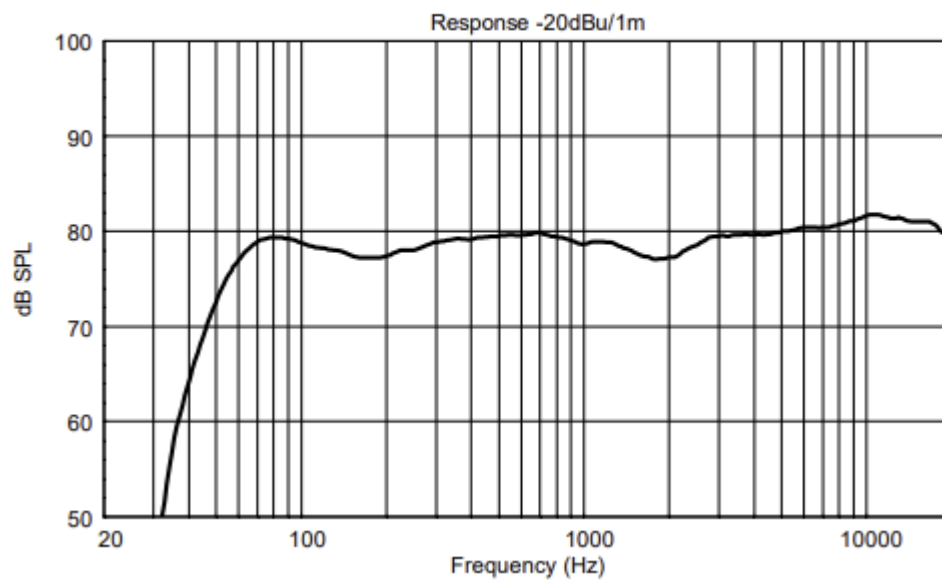


Рис. 3 — АЧХ гучномовця RCF Аура Pro5 [18]

Для запису використаємо конденсаторний мікрофон sE8 Electronics із наступними технічними характеристиками [19]:

Capsule: Hand-crafted 1/2" True Condenser

Polar patterns: Cardioid

Frequency range: 20 Hz - 20 kHz

Sensitivity: 25 mV / Pa (-32 dBV)

Max SPL: 139 / 149 / 159 dB (0 / 10 / 20 dB pad) (0.5% THD @ 1kHz)

Equivalent noise level: 13 dB

Dynamic range: 126 / 136 / 146 dB (0 / 10 / 20 dB attenuation pad)

Signal to noise ratio: 81 dB

Low-cut filter: 80 / 160 Hz, 6 dB / Oct, switchable

Attenuation pad: 10 / 20 dB, switchable

Current consumption: 2.7 mA

**Power:**

Electrical impedance: 110 Ohms

Recommended load impedance: >1k Ohms

Connectivity: 3-pin male XLR connector

Powering: 48V Phantom Power (P48 according to IEC 61938)



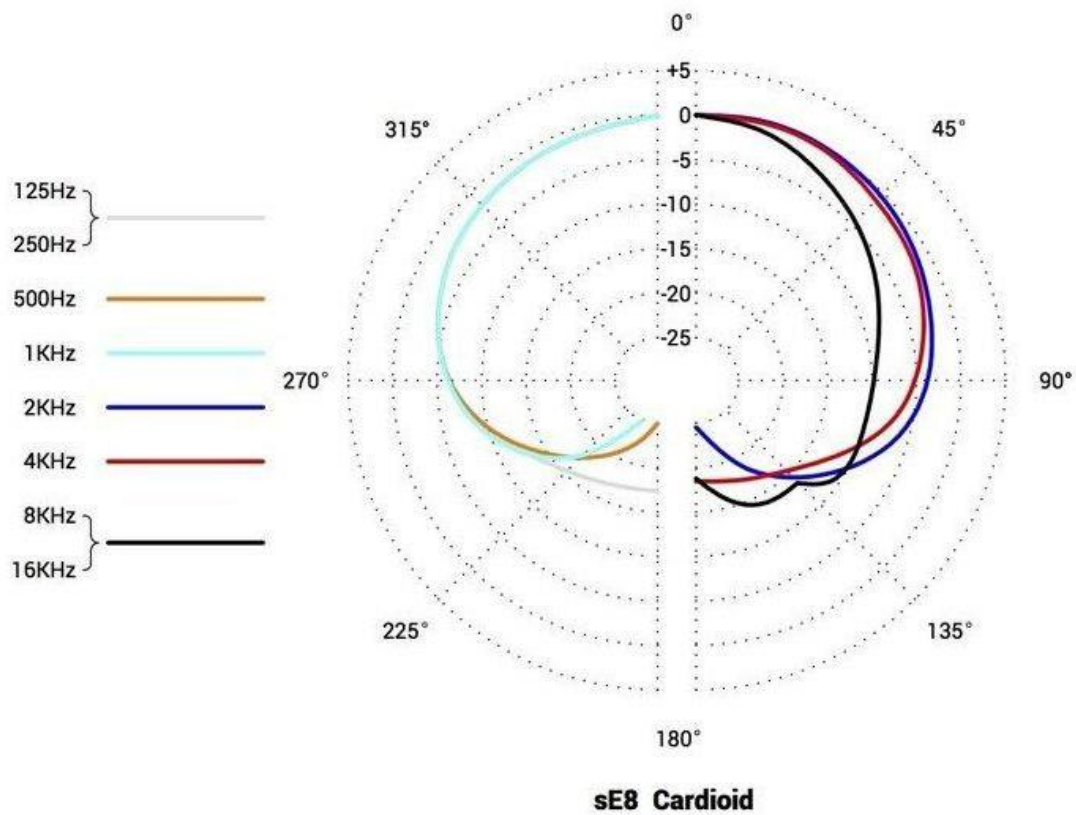


Рис. 4 — Діаграмма спрямованості мікрофона sE8 Electronics [19]



Рис. 5 — Мікрофон sE8 Electronics [19]

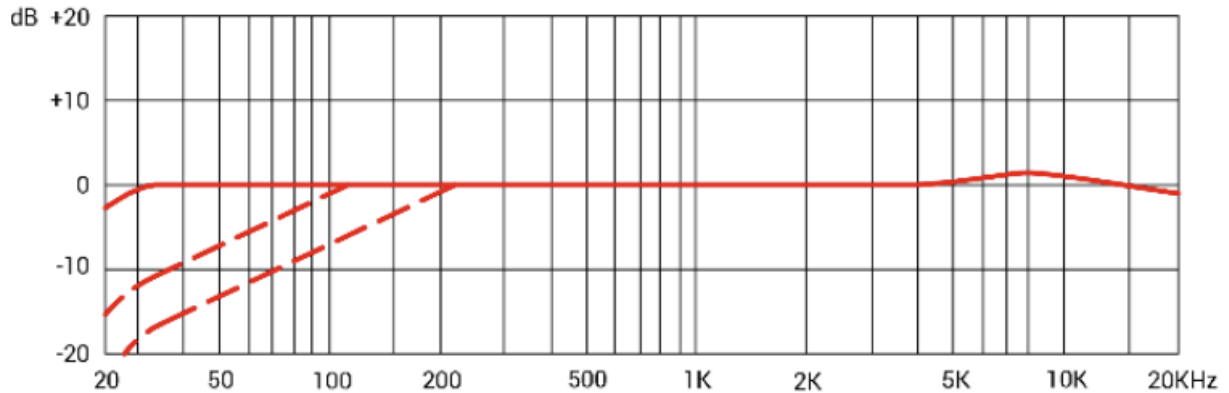


Рис. 6 — АЧХ мікрофона sE8 Electronics [19]

Аудіоінтерфейс Focusrite Clarett 4Pre має наступні технічні характеристики [20]:

**Configuration:**

Inputs 18: analogue (8), S/PDIF (2), ADAT (8)

Outputs 8: analogue (4), S/PDIF (2), HP (2)

Mixer Fully assignable 26-in/10-out mixer

Custom mixes 10 mono

Maximum custom mix inputs 18 mono

Supported sample rates 44.1 kHz, 48 kHz, 88.2 kHz, 96 kHz, 176.4 kHz & 192 kHz

**Microphone Inputs:**

Frequency Response 20 Hz - 35 kHz +/-0.5 dB at minimum gain

Dynamic Range 118 dB

THD+N 0.0009%

Noise EIN: <-128 dB (A-weighted)

Maximum input level +18 dBu at minimum gain

Gain Range 57 dB

Line & Monitor Outputs

Dynamic Range 118 dB

THD+N 0.0006%

Maximum Output Level (0 dBFS) +18 dBu, balanced line/TRS outputs



Рис. 7 — Фронтальна панель аудіо-інтерфейсу Focusrite Clarett 4Pre USB [20]



Рис. 8 — Задня панель аудіо-інтерфейсу Focusrite Clarett 4Pre USB [20]

Модель штучної голови, яка була використана це Free Space Pro II Binaural Microphone із наступними технічними характеристиками [21]:

### **Microphone Specifications:**

Directional Pattern: Omnidirectional

Cartridge Type: Pre-polarized condenser

Frequency Range: 20 Hz - 20 kHz, 3 dB soft boost at 8 - 20 kHz

Sensitivity, nominal,  $\pm 3$  dB at 1 kHz: 20 mV/Pa; -34 dB re. 1 V/Pa

S/N ratio (A-weighted), re. 1 kHz at 1 Pa (94 dB SPL): 71 dB(A)

Total harmonic distortion (THD) (Bare mic without ears): 126 dB SPL RMS, 129 dB SPL peak

Dynamic range: Typ. 106 dB

Max. SPL, peak before clipping: 134 dB

Output impedance: 30 - 40  $\Omega$

Power supply (for full performance): Min. 5 V - max. 50 V. 48 V phantom power  $\pm 4$  V for full performance

Microphone diameter: 5.4 mm (0.21 in)

Temperature range:  $-40^{\circ}\text{C}$  to  $45^{\circ}\text{C}$  ( $-40^{\circ}\text{F}$  to  $113^{\circ}\text{F}$ )

Relative humidity (RH): Up to 90%

Battery: 9V Alkaline



Рис. 9 — Фронтальна панель штучної голови Free Space Pro II Binaural Microphone [21]



Рис. 10 — Задня панель штучної голови Free Space Pro II Binaural Microphone  
[21]

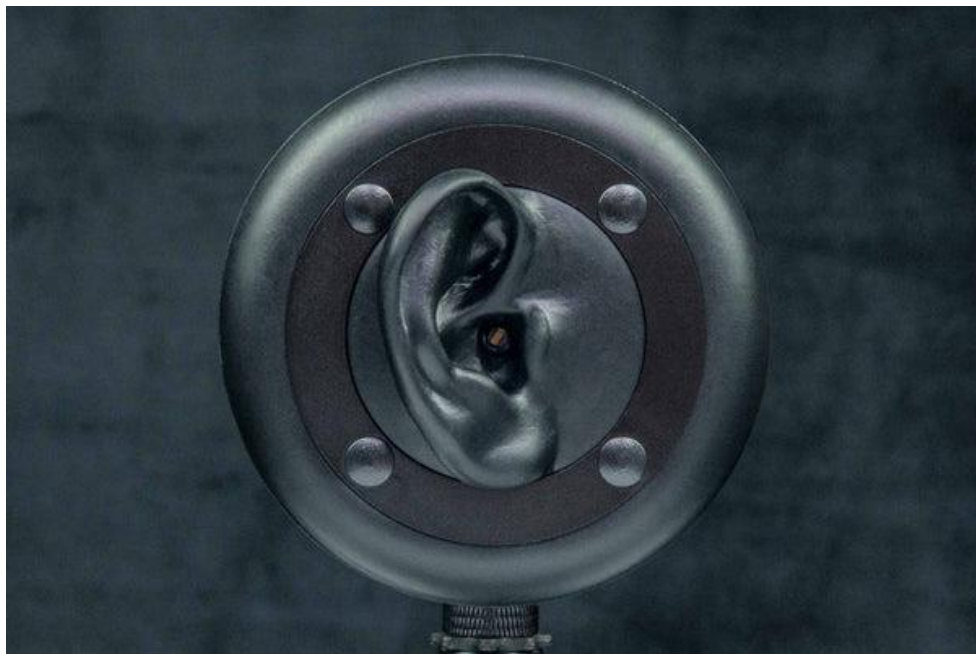


Рис. 11 — Модель людського вуха штучної голови Free Space Pro II Binaural  
Microphone [21]

### 2.2.2 Програмне забезпечення

Визначимося із видом звукового редактора для запису та прослуховування сигналів, а також для первинної підготовки музичного матеріалу. Для виконання цієї задачі скористаємось DAW Presonus Studio One v6.2.0, а також аудіо редактором від компанії iZotope RX 11 Pro Audio Editor.

**Matlab** – для **обробки сигналів**:

- формування БІХ потрібної тривалості та потрібного рівня;
- згортка 1-канального запису із БІХ (фільтрація нерекурсивним фільтром);
- проведення обчислень для отримання об'єктивних оцінок якості сигналів, отриманих непрямим методом;
- статистична обробка отриманих результатів (обчислення середнього та дисперсії отриманих суб'єктивних та об'єктивних оцінок якості) та побудова графіків, що представляють ці результати.

**програма Matlab** для отримання об'єктивних оцінок якості сигналів, отриманих непрямим методом;

**програма Matlab** для порівняння результатів суб'єктивного та об'єктивного оцінювання якості записів, отриманих непрямим методом.

## **2.3 Підготовка еталонних 2-канальних записів музики та 1-канальних дослідницьких записів**

У ході виконання дослідження було вирішено наступне завдання: проектування та реалізація запису еталонних бінауральних 2-канальних музичних сигналів у концертному залі, де для отримання натурального звучання використовувалася штучна голова (ШГ). Паралельно, також проводився запис 1-канальних музичних сигналів у безеховій студії, без використання ШГ, за допомогою одного мікрофону. Для забезпечення ефективного функціонування штучної голови та досягнення бінаурального акустичного ефекту, ШГ була розташована в концертному залі, де враховувалася специфіка звукового оточення, включаючи розташування акустичних систем та їх характеристики, акустичні властивості приміщення і параметри мікрофонів. Перед початком запису бінаурального аудіо у концертному залі, ми маємо оптимально налаштовані акустичні характеристики для надання найвищої якості звукового відтворення. Тривалість реверберації складає 1,8 секунди, що допомагає створити глибокий та емоційно насичений звуковий рельєф. Висока дифузність забезпечує рівномірне розповсюдження звуку по всьому залу, уникнення гарячих точок та створення комфортних умов для слухачів під час запису. Діаграма напрямленості оптимізована для створення однорідного звукового поля та враховує вплив різних частот на сприйняття. Звукопоглинаючі матеріали в залі допомагають зменшити відбиття звуку від стін і стелі, що сприяє покращенню якості звуку та зменшенню ехо. Спектральна рівномірність забезпечує рівномірну інтенсивність відтворення різних частот, незалежно від музичного жанру. Звукорозподіл може бути точно керований для різних зон і секторів, що дозволяє налаштовувати звукове поле під конкретні потреби виконавців та

слухачів. Також, зал оснащений системою звукозаглушення для контролю рівня шуму та створення комфортних умов під час запису бінаурального аудіо.

У результаті виконання запису було створено аудіозапис, який відтворює бінауральну передачу звуку у реальному концертному залі та аудіозапис, який представляє собою 1-канальний сигнал, записаний у безеховій студії. Цей підхід дозволяє отримати порівнювальні дані для аналізу якості та характеристик бінаурального аудіо в порівнянні з одноканальним аудіо, що відкриває можливість дослідження впливу використання штучної голови на сприйняття звуку та звукову імітацію.

#### **2.4 Запис БІХ приміщення та створення за їх допомогою 2-канальних дослідницьких записів**

У рамках проведеного експерименту було виконано запис бінауральної імпульсної характеристики приміщення з використанням штучної голови (ШГ). Підкреслимо, що цей запис здійснювався в тій самій точці приміщення, яка використовувалася для бінаурального аудіозапису. Далі, з отриманих даних, застосовуючи програмне середовище Matlab, було обчислено дослідницький 2-канальний запис. Для цього проводилася фільтрація 1-канального запису музики за допомогою бінауральної імпульсної характеристики приміщення. Цей процес мав на меті створити оптимальні умови для подальшого аналізу та порівняння якості звуку при відтворенні музики в реальному приміщенні з використанням бінаурального методу.



## 2.5 Висновки до 2 розділу

У другому розділі дослідження наведено результати підготовки програмно-апаратного забезпечення та запису музичного матеріалу для наступних досліджень.

Музичний матеріал, який використовувався при дослідженнях, складався з 5 творів класичної музики та 5 творів популярної музики. Кожен відрізок музичного матеріалу мав тривалість від 15 до 20 секунд і мав наступні технічні параметри: частота дискретизації 44100 Гц, 24 біти, бітрейт 991 кБ/с та два аудіоканали. Також було вибрано конкретні музичні твори для подальшого аналізу та порівняння.

Для відображення характеристик вихідного музичного матеріалу була побудована спектрограма. Щодо програмно-апаратного забезпечення, для випромінення сигналів використовувалося акустично підготовлене приміщення зі студійними гучномовцями. Для запису двоканального сигналу використовувалися два мікрофони та штучна голова. Крім того, використовувалися зовнішня звукова карта та ноутбук.

Для обробки сигналів та формування бінауральних імпульсних характеристик приміщення використовувалася програма Matlab. Зокрема, виконувалися такі операції, як формування БІХ необхідної тривалості та рівня, згортка одноканального запису з БІХ (фільтрація нерекурсивним фільтром), проведення обчислень для отримання об'єктивних оцінок якості сигналів, отриманих непрямим методом, а також статистична обробка отриманих результатів та побудова графіків для їх подальшого аналізу.

### Розділ 3: Оцінювання якості 2-канальних дослідницьких записів

Нижче описано процес створення двоканальних дослідницьких записів, які були виконані за допомогою бінауральної імпульсної характеристики. Створення файлів “Result (\*).wav” (програма zgortka.m).

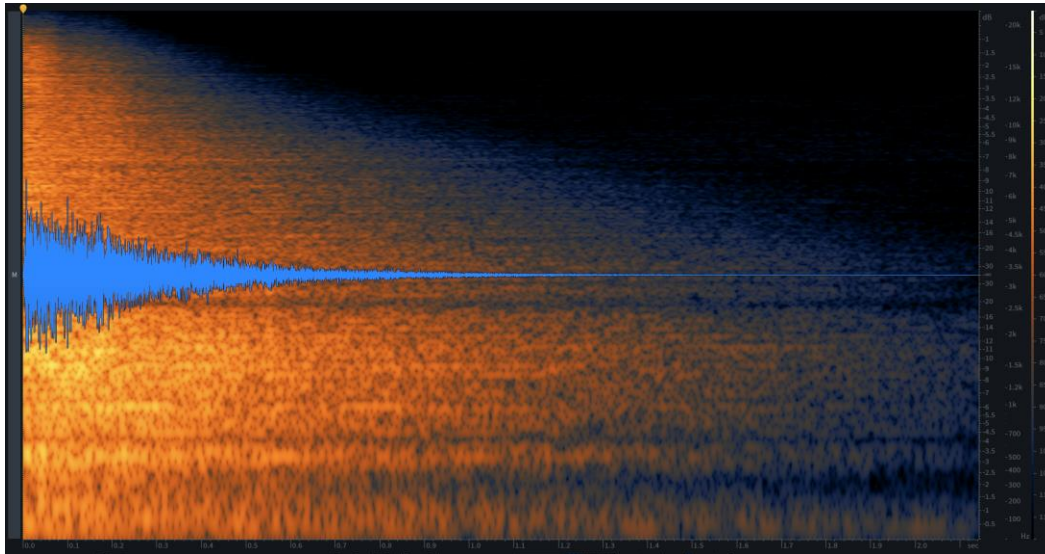


Рис. 12 — Спектрограма бінауральної імпульсної характеристики

Для цього за допомогою програмного забезпечення MatLab була створена програма згортки, яка сгенерувала 10 дослідницьких записів за допомогою БІХ, лістинг коду наведено у Додатку 1.

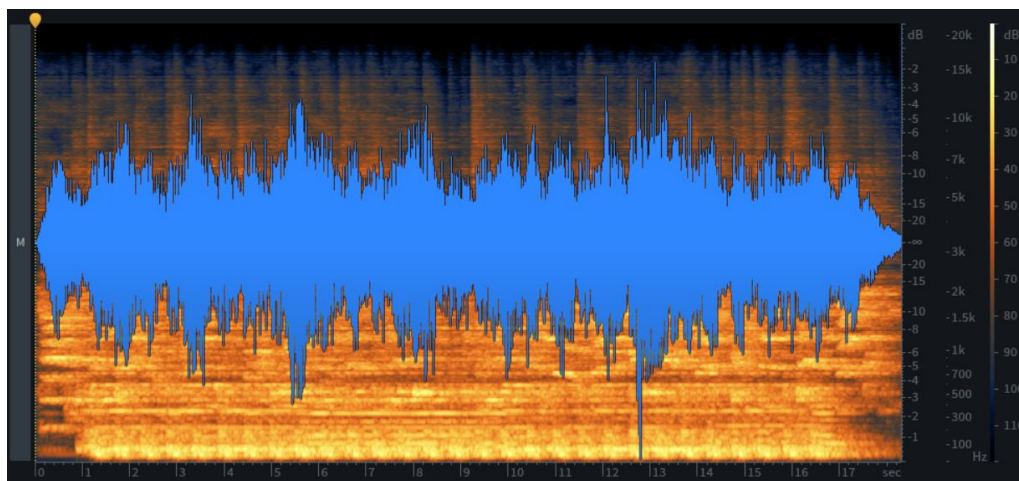


Рис. 13 — Спектрограма дослідницького запису

### 3.1 Суб'єктивне оцінювання якості 2-канальних дослідницьких записів за шкалою DMOS

Виконується суб'єктивне, кількома слухачами, на слух, порівняння еталонних 2-канальних музичних сигналів із дослідницькими 2-канальними записами.

Шкала оцінювання 5-бальна, дозволяється ставити дробові оцінки (наприклад, 3,5 або 4,7).

Таблиця 1

Шкала якості DMOS (Degradation Mean Opinion Score)

Бали	Що означають
5	Записи не відрізняються
4	Відчувається невелика різниця
3	Різниця помітна
2	Різниця суттєва
1	Різниця дуже суттєва

Припустимими є дробові значення оцінки (наприклад, 3.7 - три цілих та сім десятих). Результати оцінювання, що мають вигляд функціональної залежності DMOS(df), виводяться на графік та записуються на диск у вигляді файлу-структури (такий файл має розширення mat). Ім'я файлу при цьому складається із двох частин, що розділені знаком підкреслювання: перша частина є іменем звукового файлу, друга є прізвищем виконавця (наприклад: Rec1\_Doyna). Одержані графіки також записуються на диск (для подальшого використання у звіті) у вигляді графічного файлу із розширенням bmp із таким же ім'ям. Файли mat зберігаються для подальшої обробки (порівняння із результатами об'єктивного оцінювання).



Рис. 14 — Узагальнена схема досліджень, що проводяться

**Суб'єктивно оцінимо, які файли реальні, а які – ні (програма `osinka_vidpovidnosti.m`), лістинг коду у Додатку 5 [17].**

**Приклад діалогу при роботі програми:**

*Натисніть будь-яку клавішу для початку*

*Скільки секунд файлів хочете програвати? (секунд): 15*

*Програється музичний твір №1*

*Програється перший файл*

*Програється другий файл*

*Введіть номер аудіозапису, який ви вважаєте реальним (1 чи 2): 2*

*Програється музичний твір №2*

*Програється перший файл*

*Програється другий файл*

*Введіть номер аудіозапису, який ви вважаєте реальним (1 чи 2): 1*

.....

Таблиця 2

Відсоток вірного розпізнавання сигналів:

Особа 1	Особа 2	Особа 3	Особа 4	Особа 5
10%	30%	20%	40%	30%

Середній відсоток вірного розпізнавання сигналів складає всього 26%, що вказує нам на те, що досліджувана група осіб майже не може відрізнити реальний бінауральний запис від дослідницького, створеного штучно.

Вище наведені результати підкріплюють попередньо зроблені висновки стосовно того, наскільки легко надурити людський мозок і змусити його думати, що ми відчуваємо, чуємо чи бачимо щось насправді, навіть коли все – лише симуляція. Хед-сети віртуальної реальності та кінотеатри є найпростішими прикладами та доказами того, як сильно на нас впливають інновації та розвиток розуміння і вивчення аудіо-візуальних технологій. Водночас бінауральне аудіо відкриває купу можливостей для майбутнього використання даної технології, можливо колись науковці удосконалять, наприклад, слухові апарати для людей зі слуховими вадами, використовуючи саме бінауральні запис та відтворення звуку в реальному часі.

### 3.2 Об'єктивне оцінювання якості 2-канальних дослідницьких записів

Із застосуванням спеціальних комп'ютерних програм порівнюються еталонний та дослідницький 2-канальні записи музики. Якість сигналів визначається шляхом обчислення показників якості SSNR та LSD – при цьому, як і при суб'єктивному оцінюванні, відбувається порівняння сигналу  $y(t)$  на виході НЧ фільтра із еталонним сигналом  $x(t)$  на вході фільтра (рис. 14). Результати оцінювання, що мають вигляд функціональних залежностей SSNR(df) та LSD(df), виводяться на графіки та записуються на диск у вигляді файлів-структур (такі файли мають розширення mat). Ім'я файлу при цьому складається із двох частин, що розділені знаком підкреслювання: перша частина є назвою показника якості, друга є прізвищем виконавця (наприклад: SSNR\_Doyna). Одержані графіки вручну записуються на диск (для подальшого використання у звіті) у вигляді графічних файлів із розширенням bmp. Файли mat зберігаються для подальшої обробки (порівняння із результатами суб'єктивного оцінювання).

Дане оцінювання є 2-м етапом досліджень. Метою цього етапу є обчислення простих об'єктивних показників якості сигналів - сегментного відношення сигнал-шум (SSNR) та логарифмічного спотворення спектру (LSD) - та співставлення їх із результатами суб'єктивного оцінювання якості сигналів. Інша мета – співставлення результатів об'єктивного та суб'єктивного оцінювання. Сегментне відношення сигнал-шум SSNR (Segmental Signal to Noise Ratio) обчислюється за виразом:

$$SSNR = \frac{1}{M} \sum_{m=1}^M 10 \lg \left[ \frac{\sum_{n=R(m-1)+1}^{Rm} x^2(n, m)}{\sum_{n=R(m-1)+1}^{Rm} [x(n, m) - y(n, m)]^2} \right],$$

де  $x(n,m)$  и  $y(n,m)$ -  $n$ -е вибірки  $m$ -го фрейму чистого сигналу  $x(n)$  та спотвореного сигналу  $y(n)$ , відповідно;  $M$  - кількість фреймів;  $R$  - кількість вибірок у фреймі. При обчисленні показника SSNR враховуються тільки ті сегменти, для яких відношення сигналшум обмежено інтервалом  $[-10, +35]$  дБ. Логарифмічне спектральне спотворення LSD (Logarithmic Spectral Distortion) обчислюється за виразом:

$$LSD = \frac{2}{RM} \sum_{m=1}^M \sum_{j=1}^J |G\{X(j,m)\} - G\{Y(j,m)\}|,$$

$$G\{X(j,m)\} = \max\{20 \lg(|X(j,m)|), \delta\}, \quad \delta = \max_{l,k} \{20 \lg(|X(j,m)|)\} - 50,$$

де  $X(j,m)$  и  $Y(j,m)$  - дискретні перетворення Фур'є  $m$ -го фрейму сигналів  $x(n)$  и  $y(n)$ , відповідно,  $j$  - номер частотної вибірки,  $J$  - кількість частотних вибірок. Проводимо об'єктивне оцінювання якості сигналів (управляюча програма `obj_quality.m` та підпрограми функції `srr_func_correct2.m`, `lsd_func.m`, `enframe.m`, `rfft.m`): лістинг коду у Додатках 6-10 [17].

### 3.3 Отримані результати

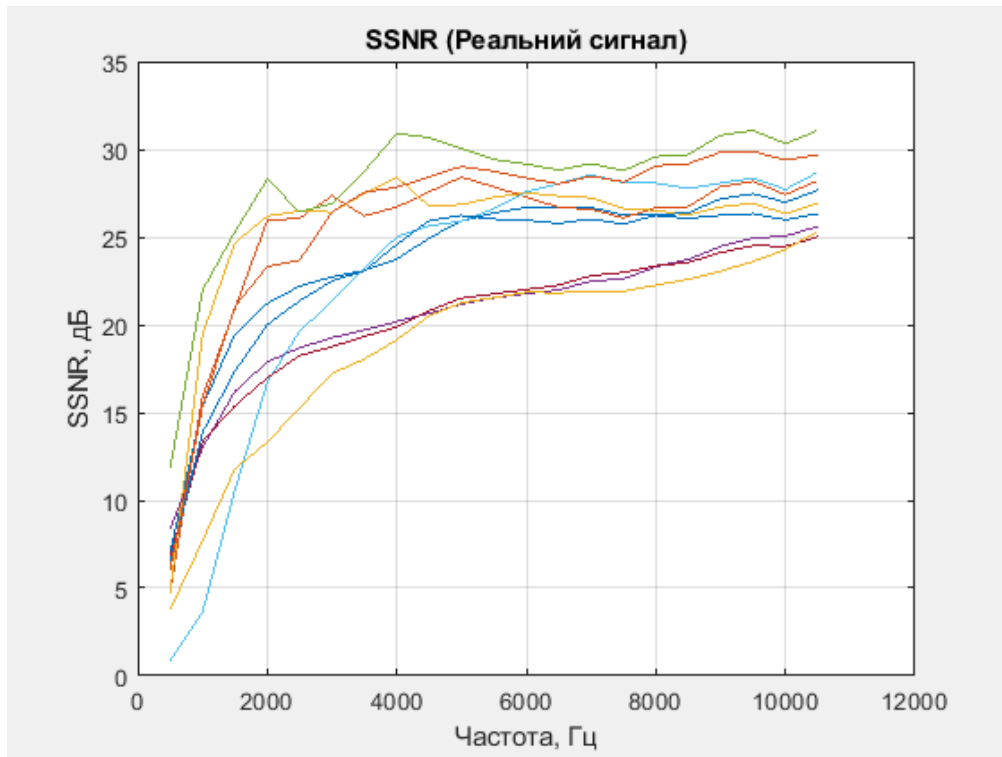


Рис. 19 — Графік SSNR запису на штучну голову

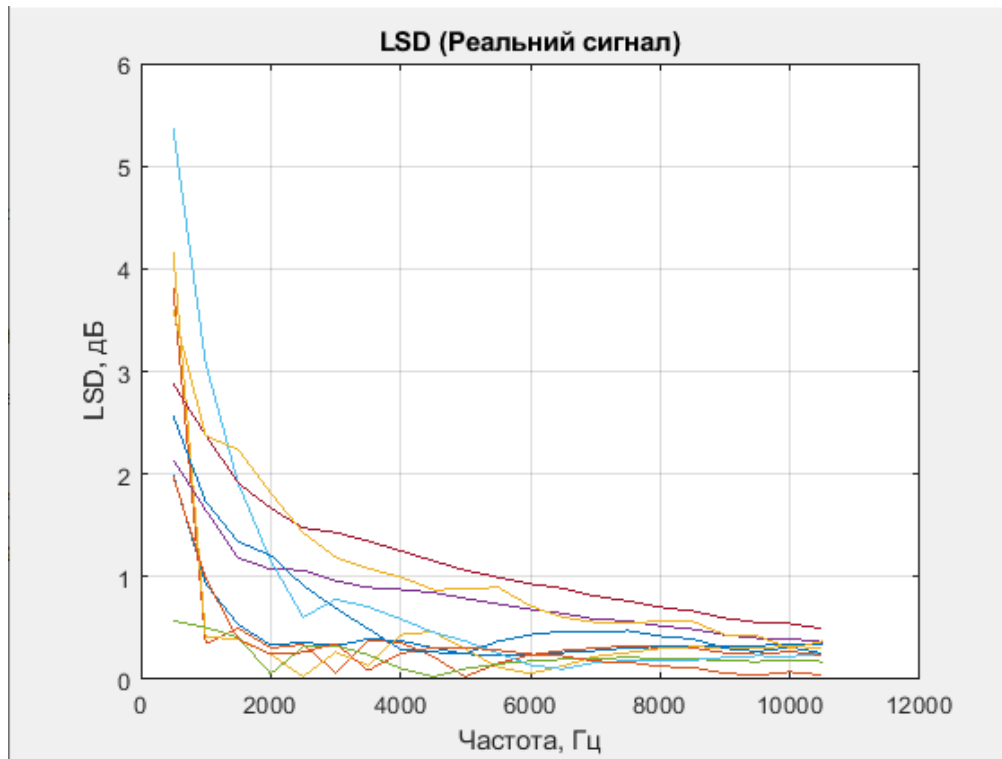


Рис. 20 — Графік LSD запису на штучну голову



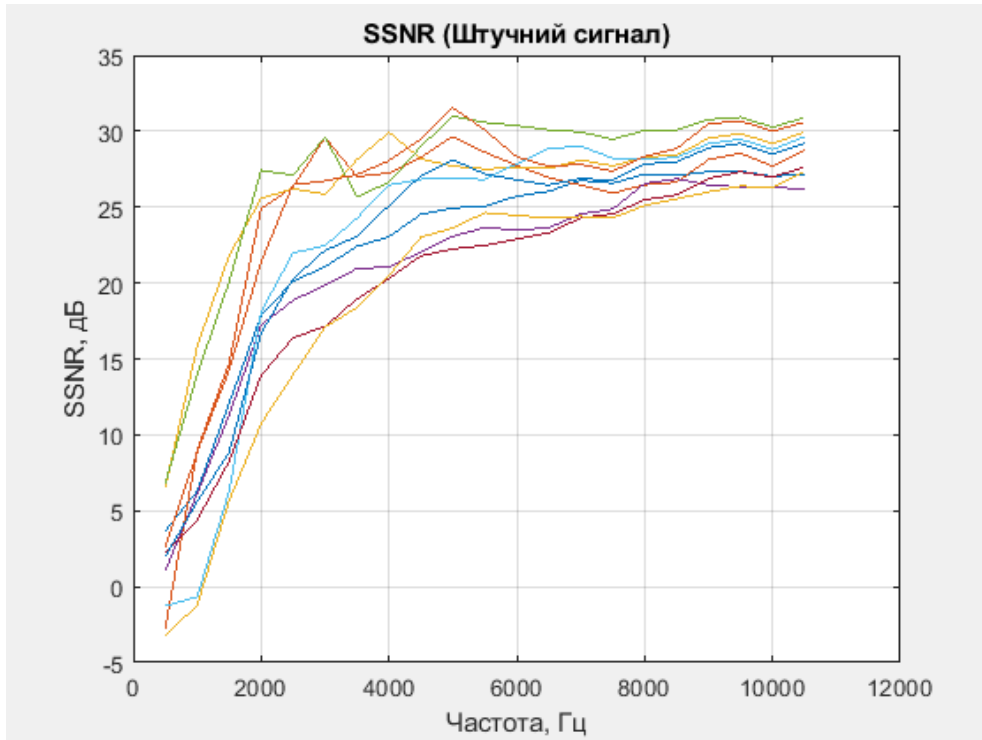


Рис. 21 — Графік SSNR дослідницького запису

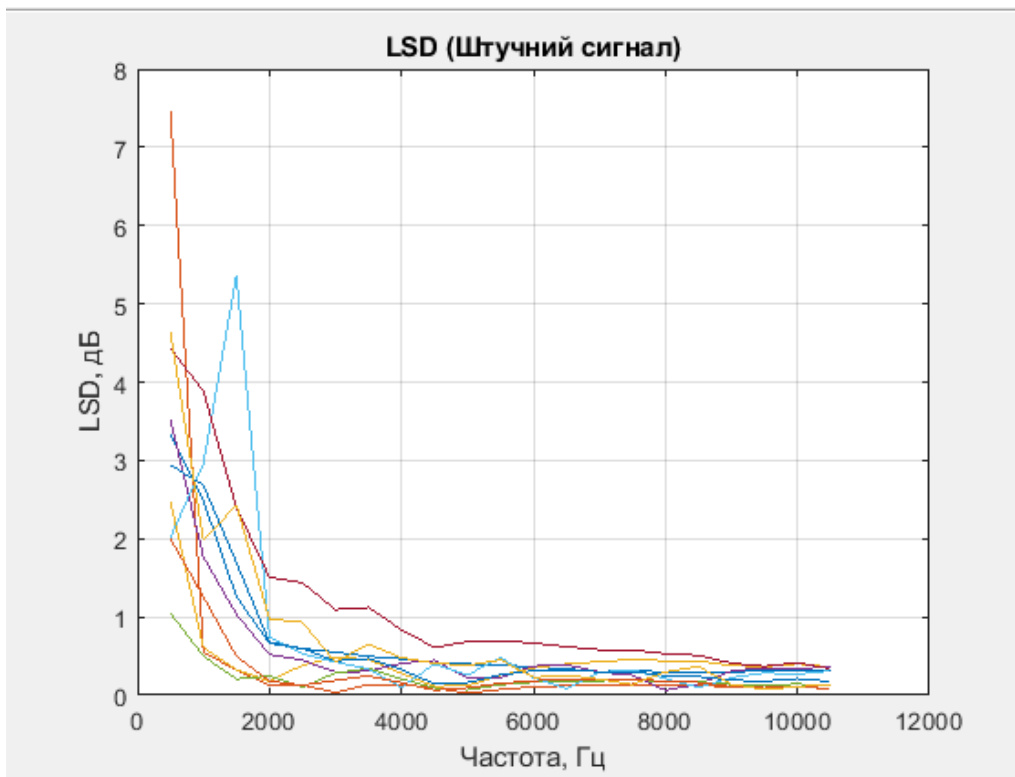


Рис. 22 — Графік LSD дослідницького запису

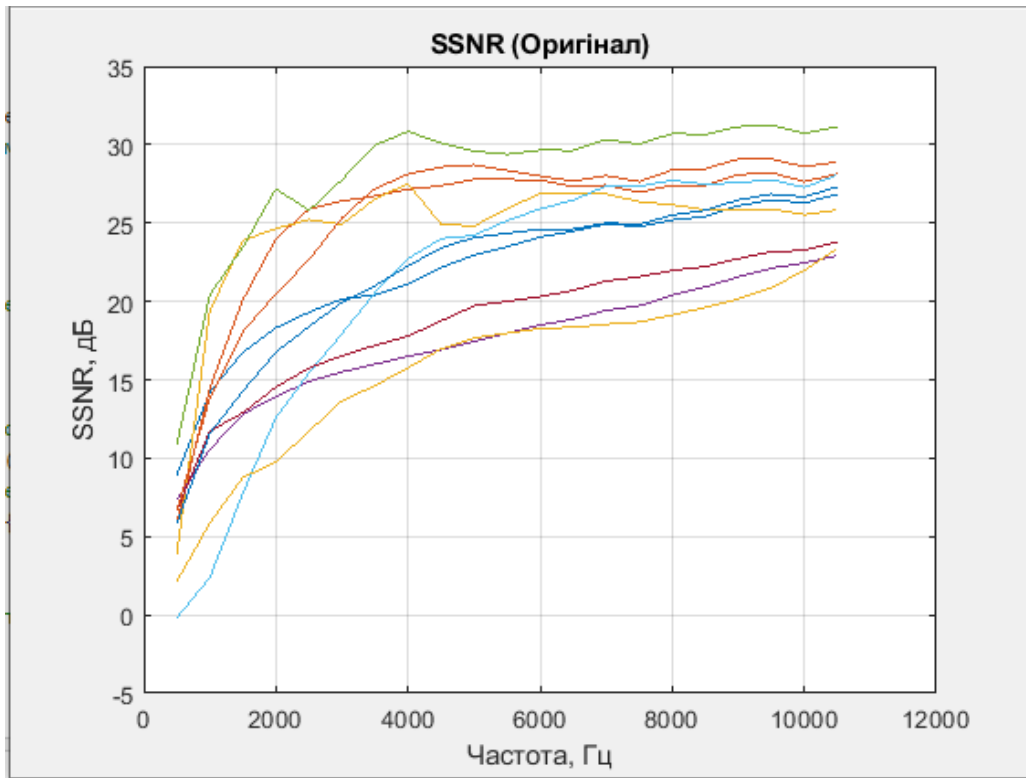


Рис. 23 — Графік SSNR оригінальної платівки

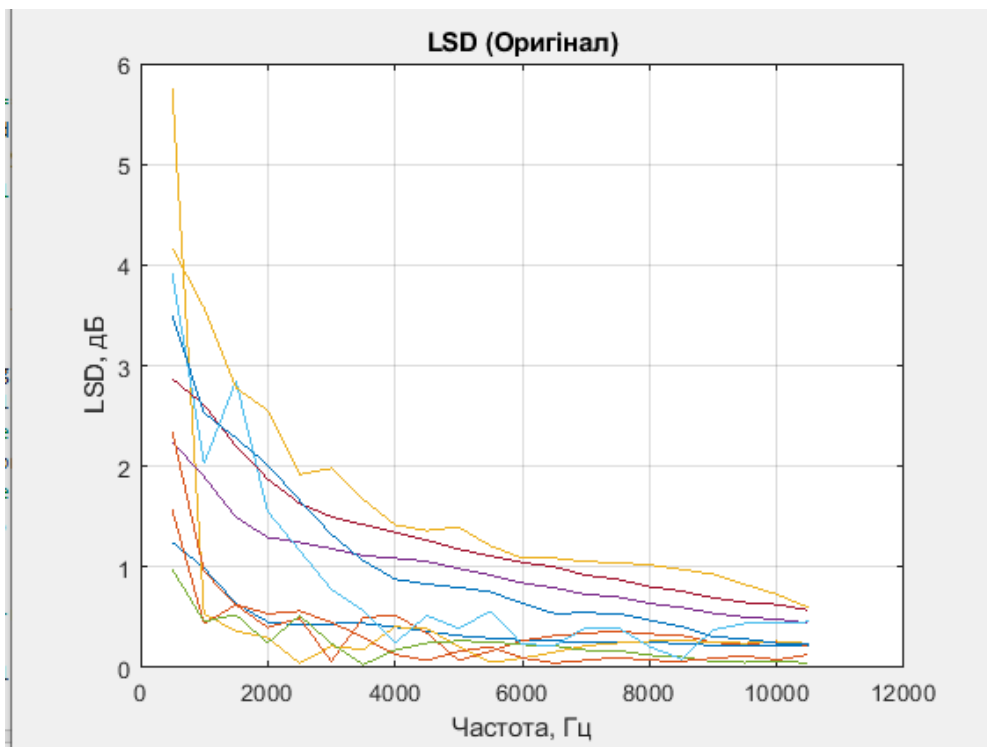


Рис. 24 — 1 Графік LSD оригінальної платівки

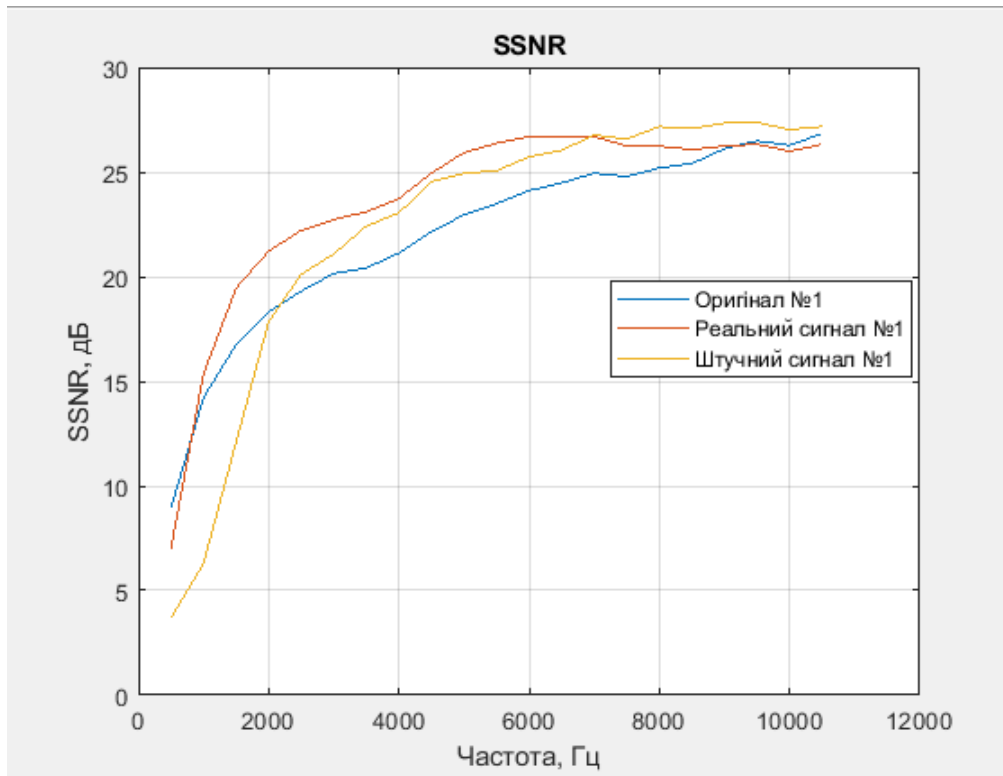


Рис. 25 — Графік порівняння SSNR для першого твору

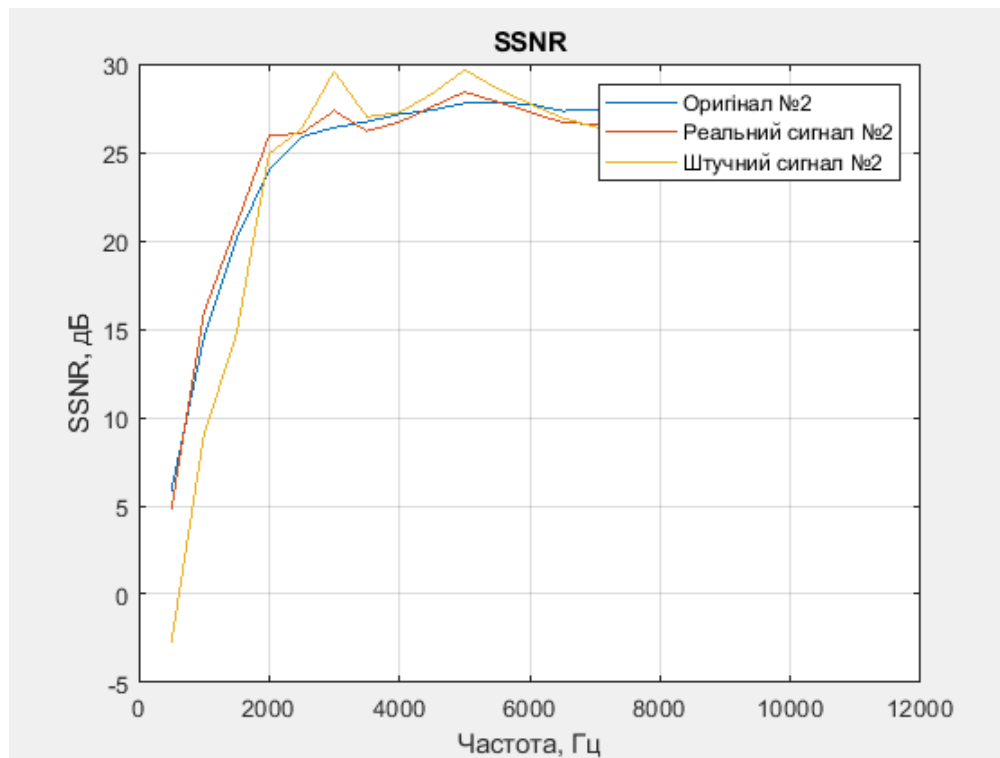


Рис. 26 — Графік порівняння SSNR для другого твору

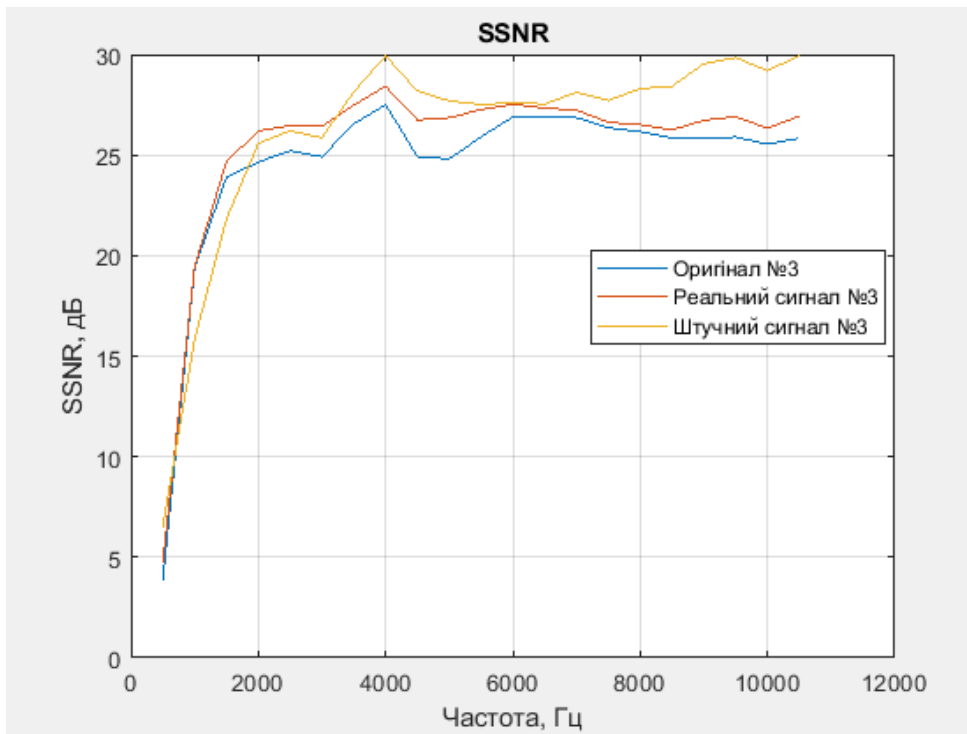


Рис. 27 — Графік порівняння SSNR для третього твору

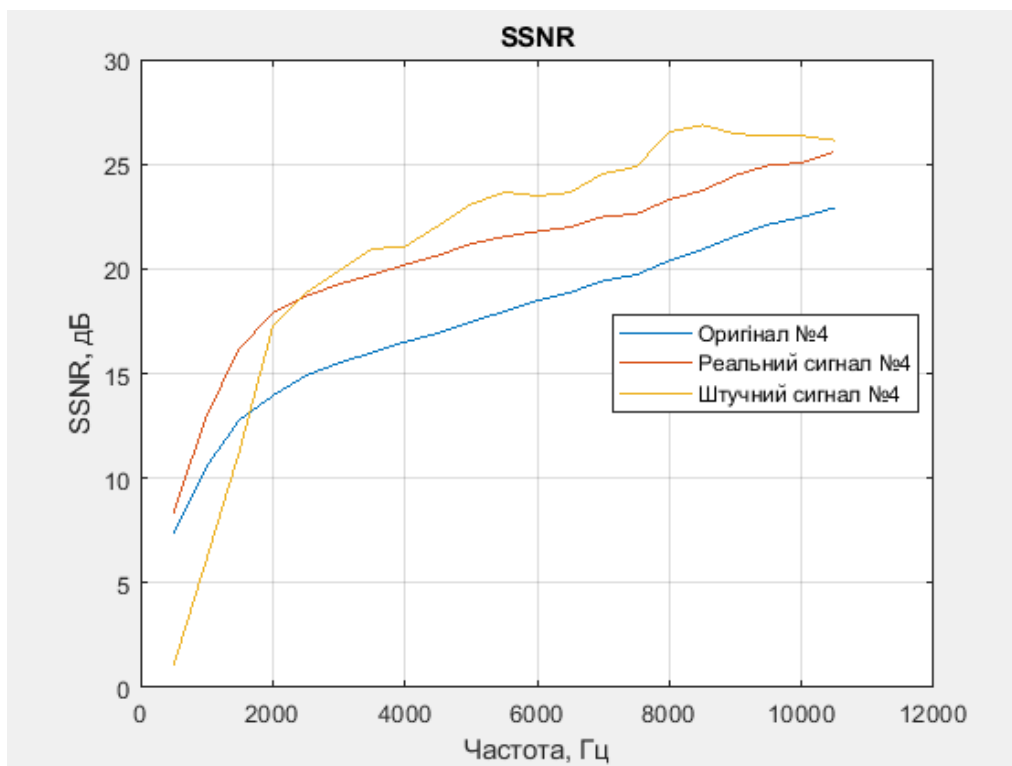


Рис. 28 — Графік порівняння SSNR для четвертого твору

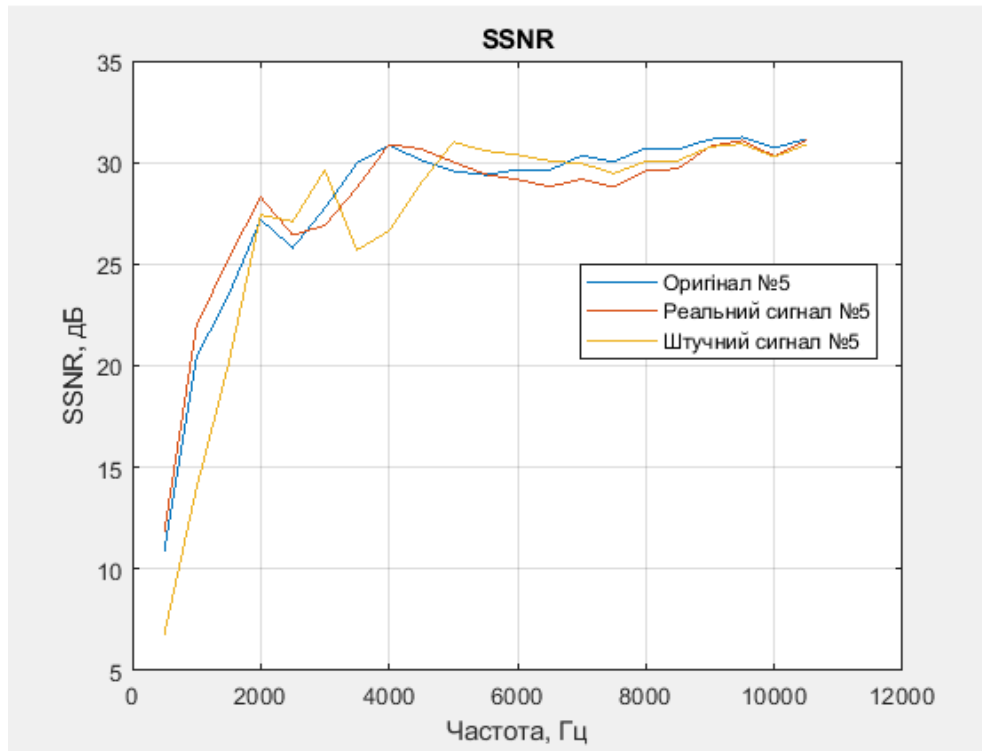


Рис. 29 — Графік порівняння SSNR для п'ятого твору

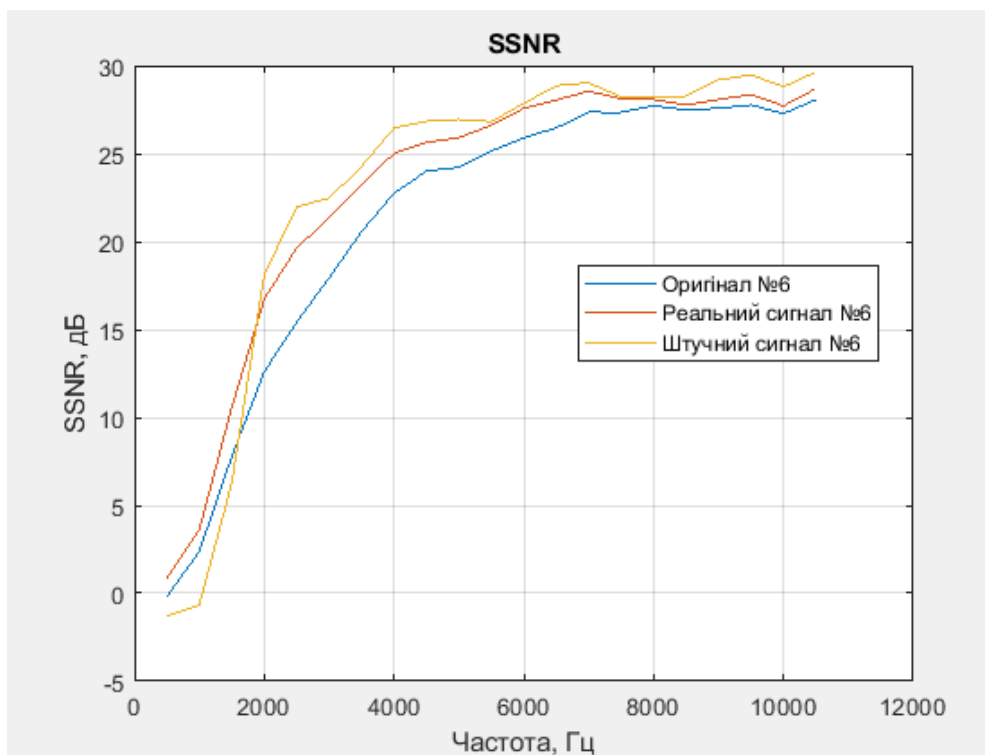


Рис. 30 — Графік порівняння SSNR для шостого твору

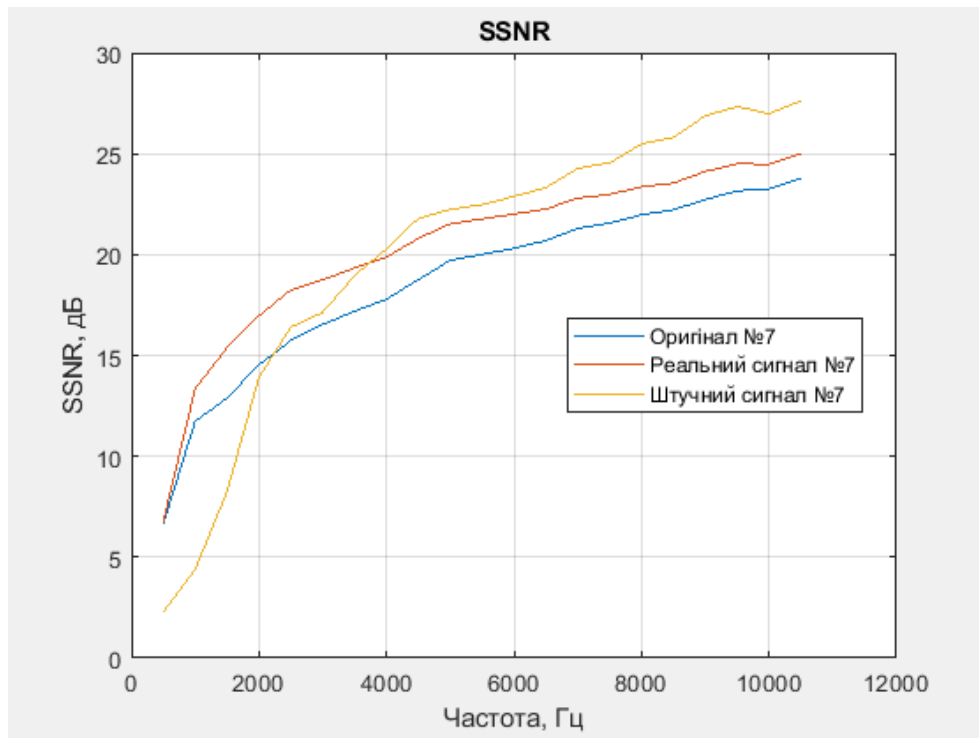


Рис. 31 — Графік порівняння SSNR для сьомого твору

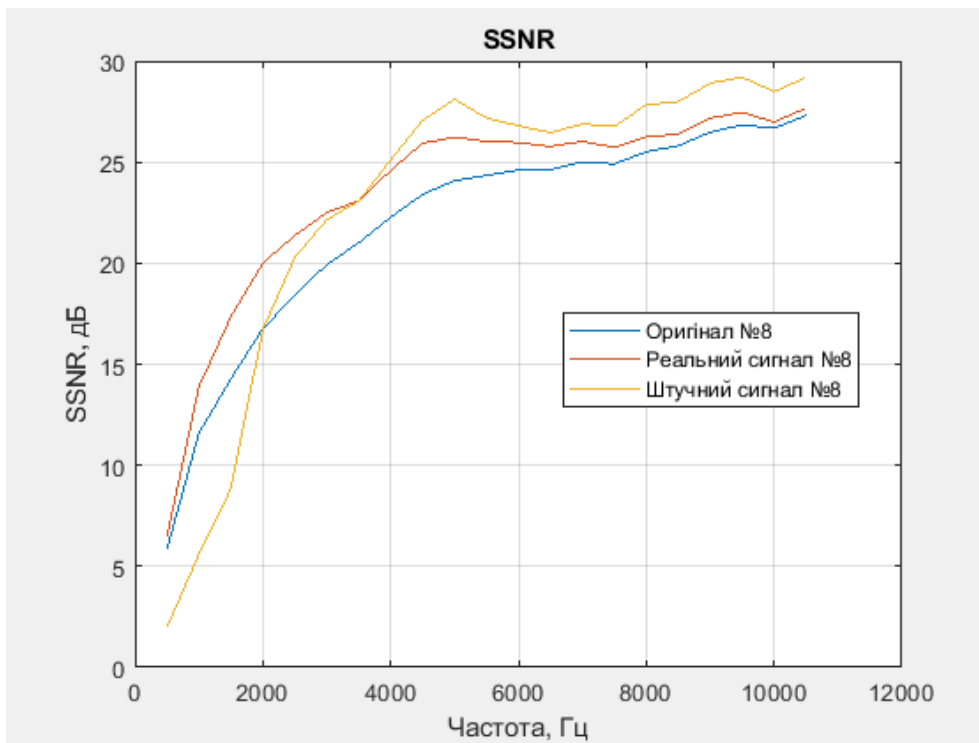


Рис. 32 — Графік порівняння SSNR для восьмого твору



Рис. 33 — Графік порівняння SSNR для дев'ятого твору

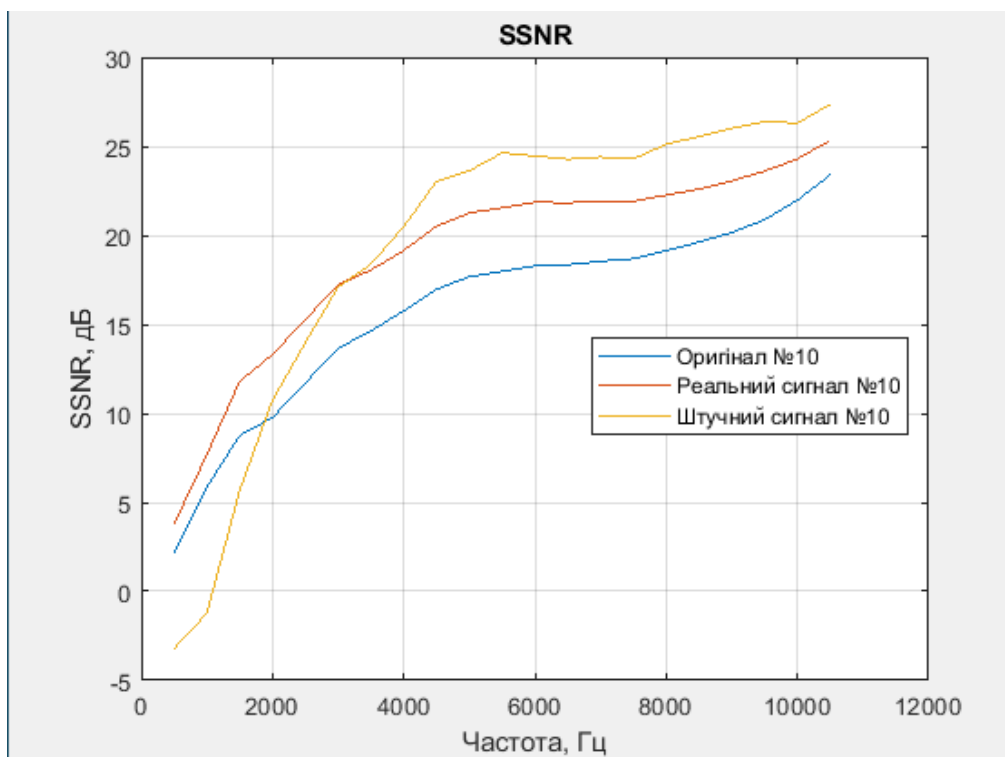


Рис. 34 — Графік порівняння SSNR для десятого твору

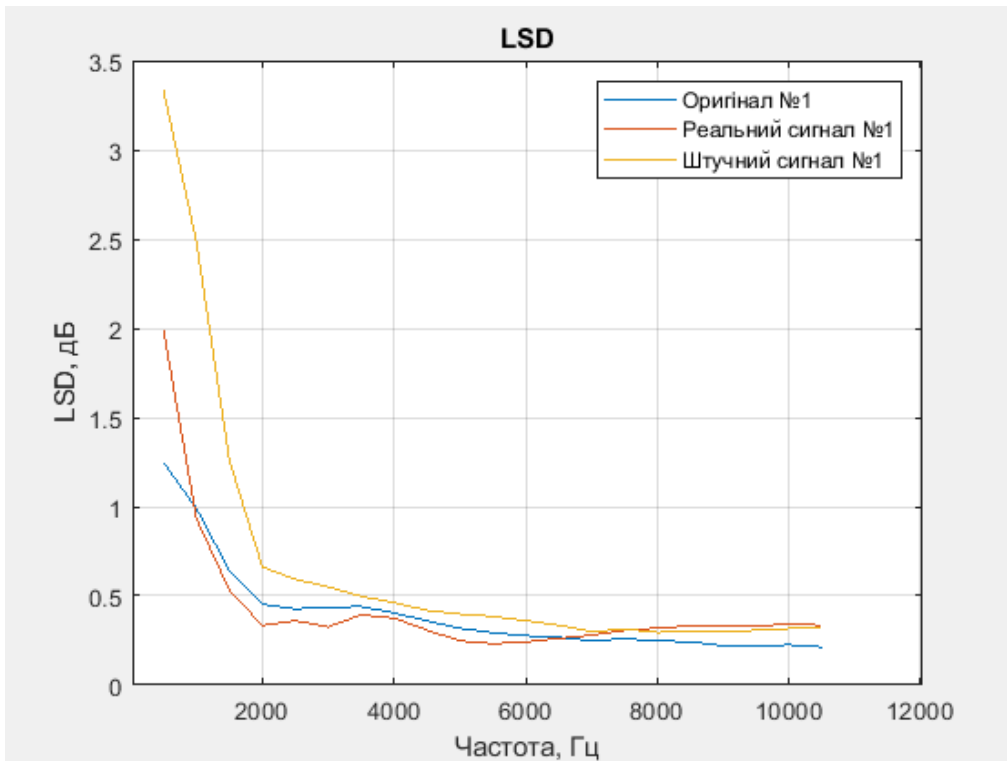


Рис. 35 — Графік порівняння LSD для першого твору

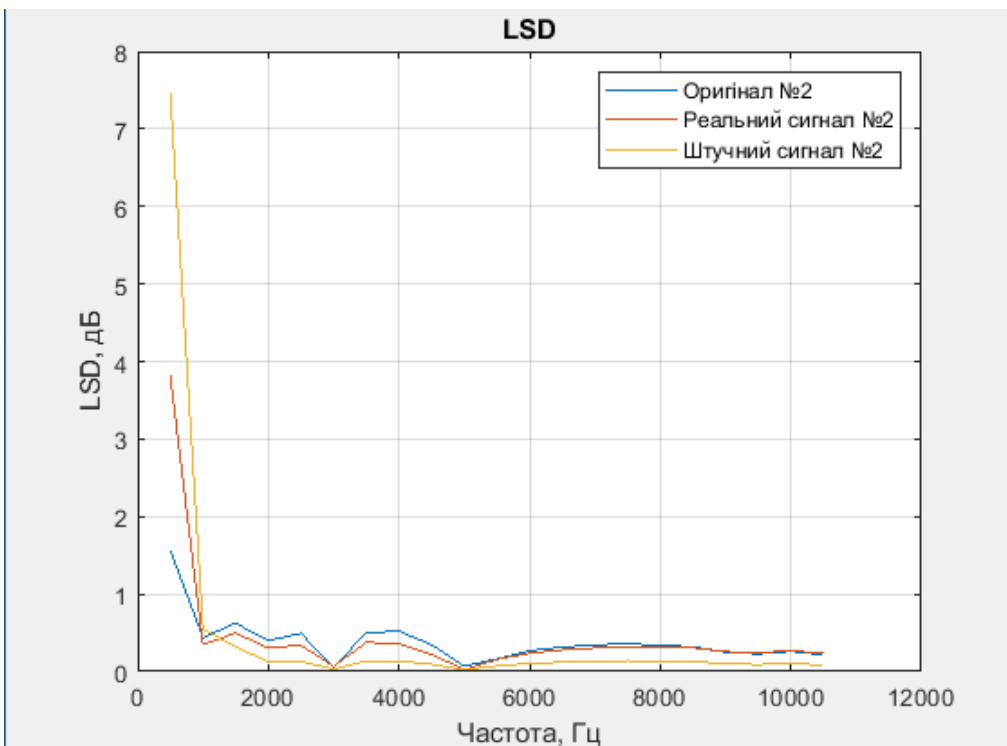


Рис. 36 — Графік порівняння LSD для другого твору



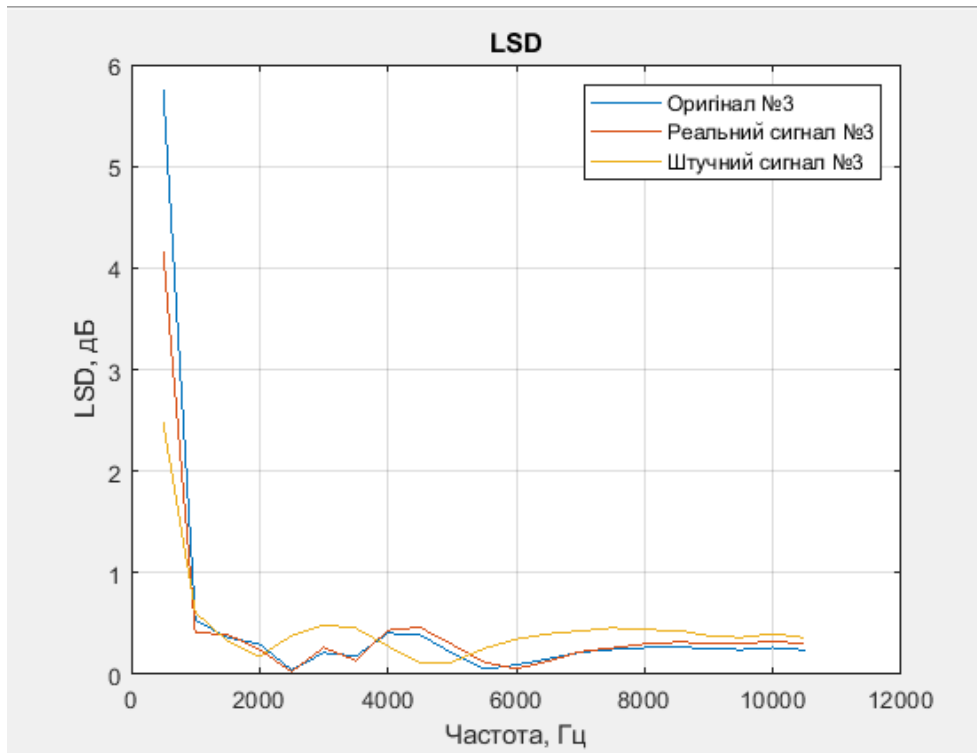


Рис. 37 — Графік порівняння LSD для третього твору

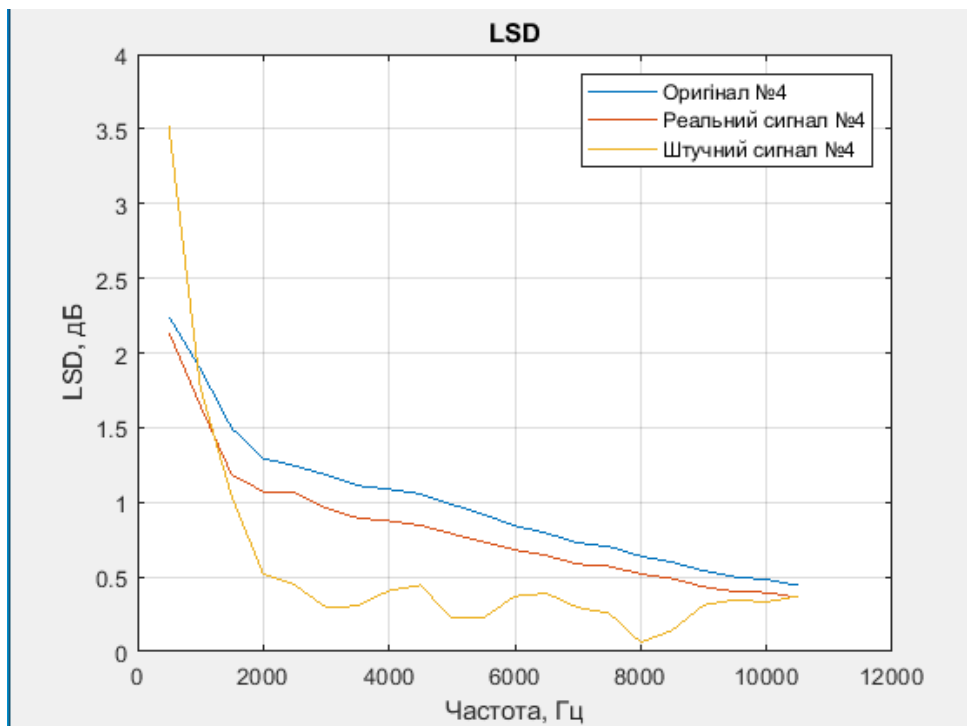


Рис. 38 — Графік порівняння LSD для четвертого твору

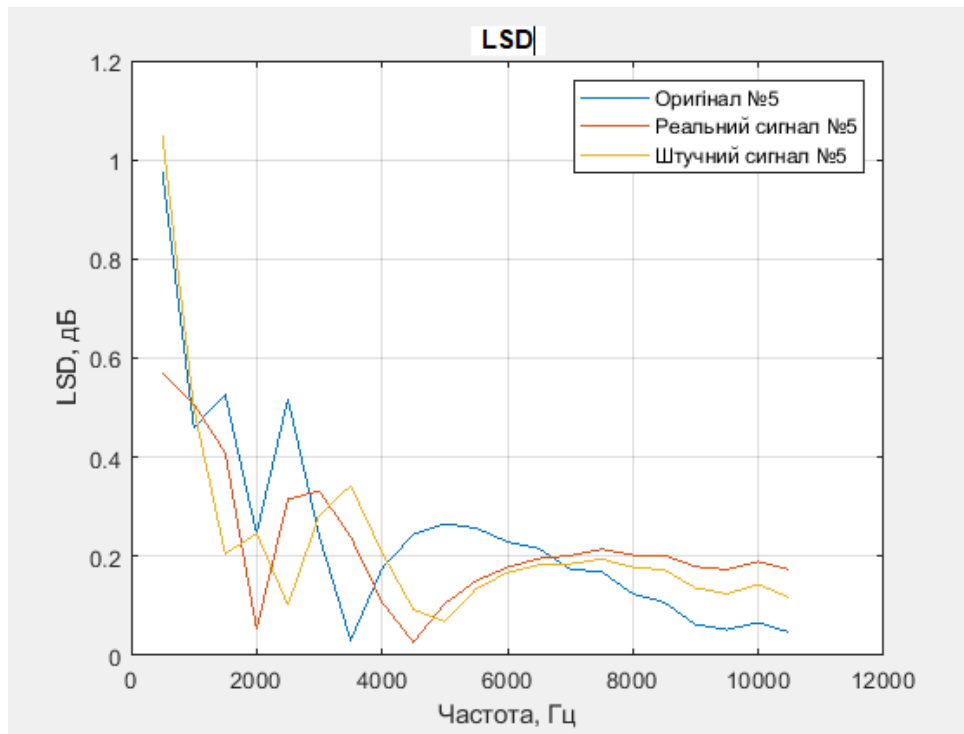


Рис. 39 — Графік порівняння LSD для п'ятого твору

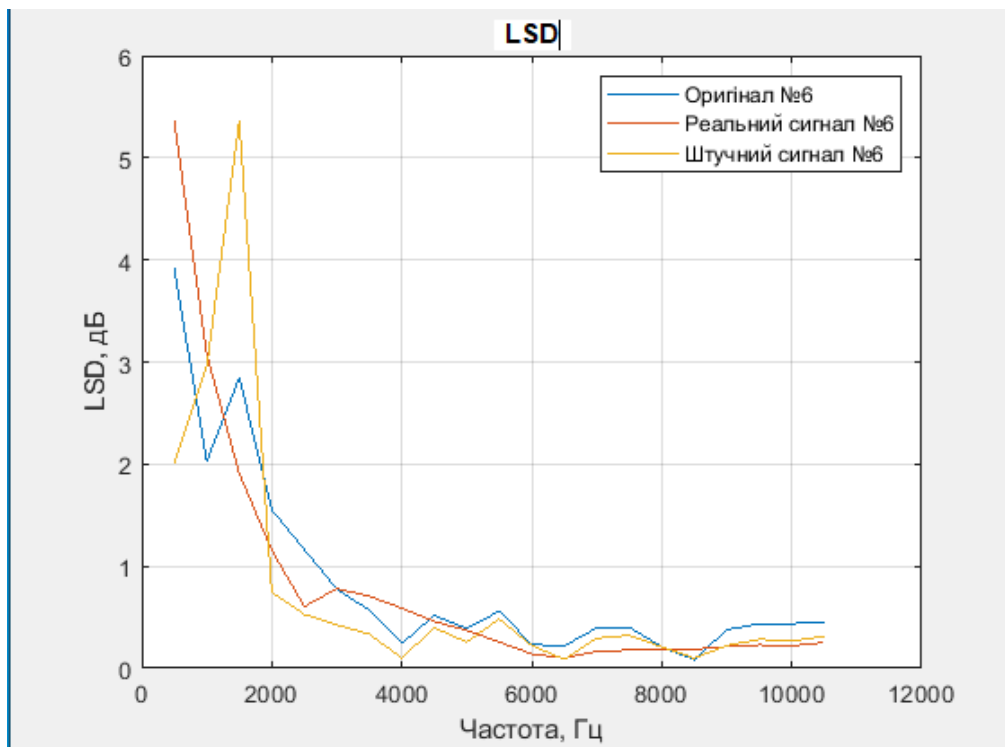


Рис. 40 — Графік порівняння LSD для шостого твору

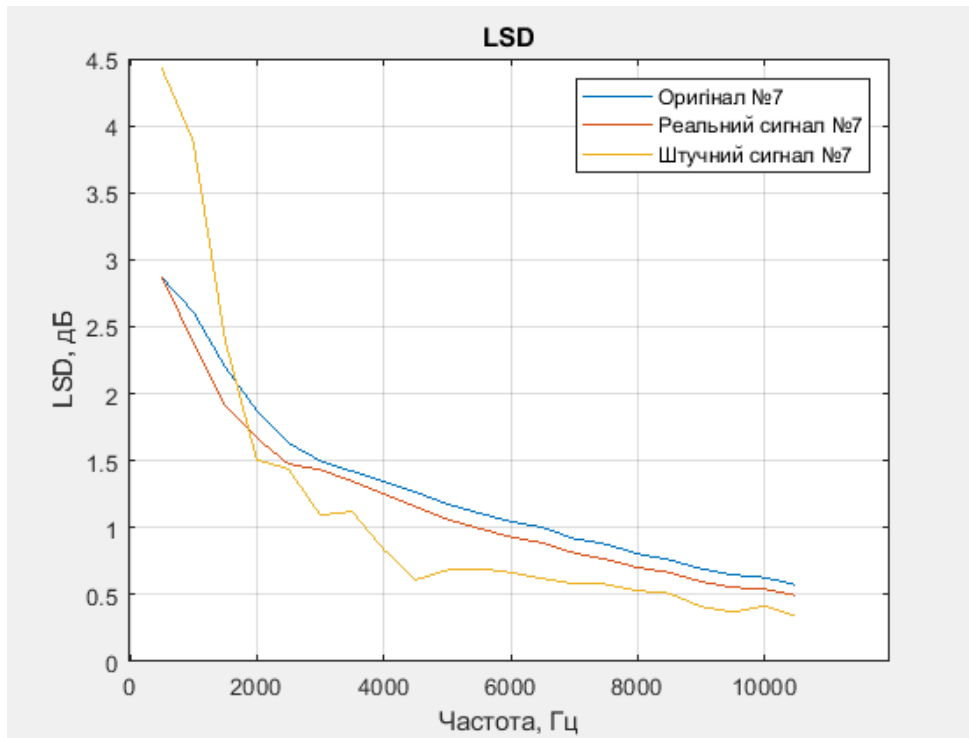


Рис. 41 — Графік порівняння LSD для сьомого твору

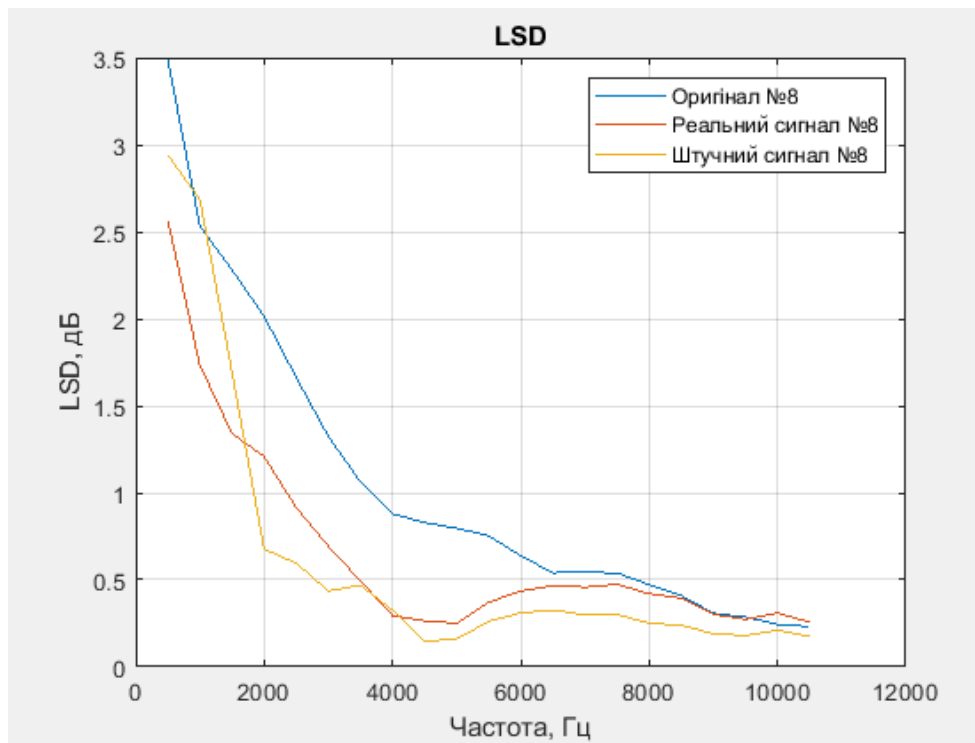


Рис. 42 — Графік порівняння LSD для восьмого твору

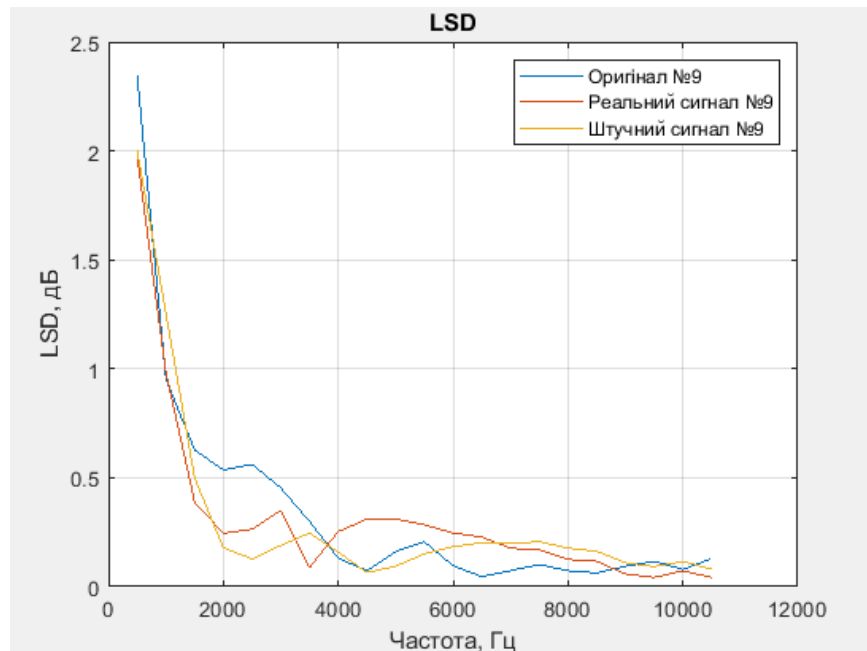


Рис. 43 — Графік порівняння LSD для дев'ятого твору

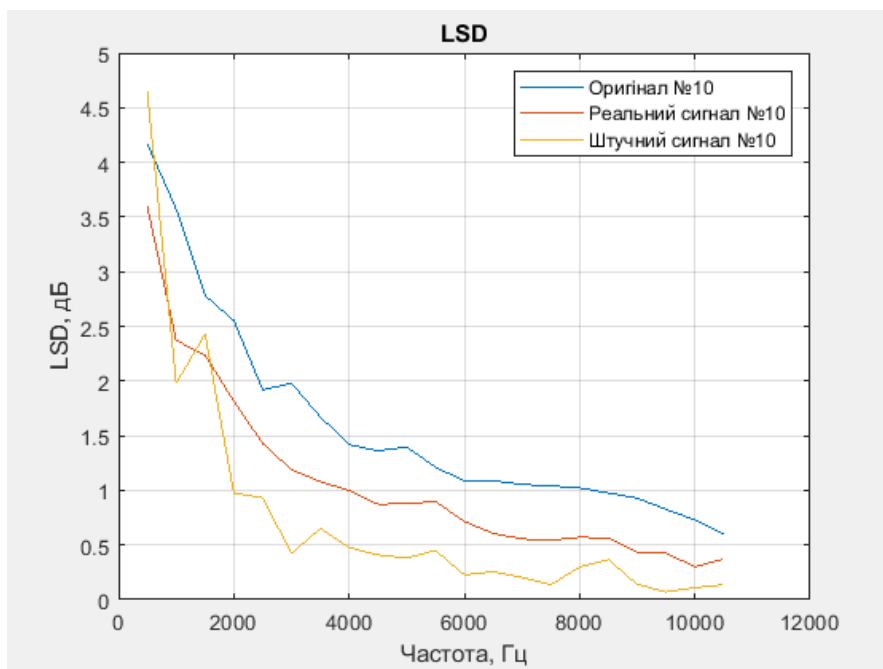


Рис. 44 — Графік порівняння LSD для десятого твору

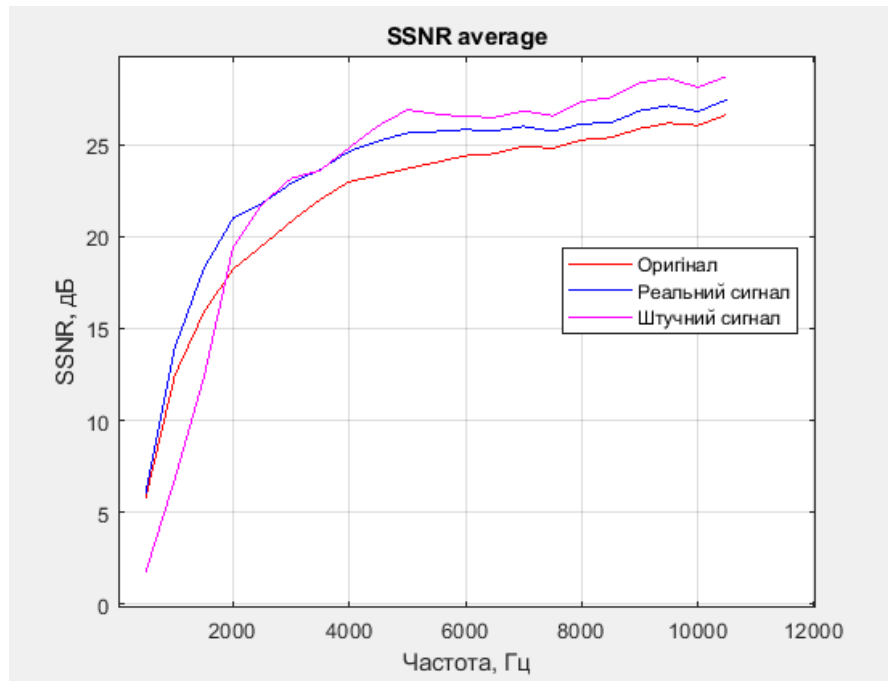


Рис. 45 — Графік усередненого SSNR

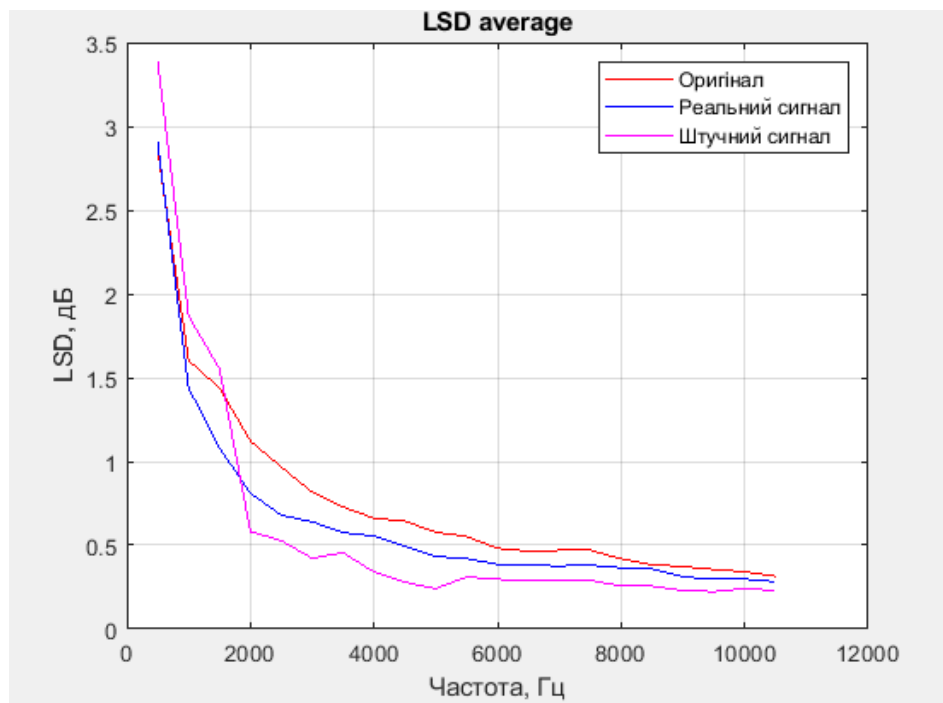


Рис. 46 — Графік усередненого LSD

### 3.4 Висновки до 3 розділу

У даному розділі наведено результати оцінювання якості 2-канальних дослідницьких записів з використанням бінауральної імпульсної характеристики та аналізу суб'єктивних та об'єктивних показників.

Створені записи, що порівнювалися із еталонними 2-канальними музичними сигналами, піддавалися суб'єктивному оцінюванню п'ятьма слухачами. Отримані функціональні залежності DMOS(df) були усереднені та представлені на графіках, в результаті чого виявлено, що суб'єктивне оцінювання вказує на суттєву складність розрізнення реальних та штучно створених бінауральних записів. Усереднені результати були відображені графічно та збережені у вигляді файлів-структур.

Додатково, використовуючи об'єктивні показники якості, такі як SSNR та LSD, було проведено об'єктивне оцінювання якості сигналів, результати якого також візуалізовані та порівняні з суб'єктивними оцінками. Отримані результати можуть вказувати на важливі особливості сприйняття бінауральних аудіозаписів людським слухом.

Загальний висновок дослідження вказує на важливість подальших робіт у розумінні та поліпшенні якості бінаурального аудіо, і отримані результати можуть слугувати важливим внеском у це поле знань.

## Висновки

В даній дисертаційній роботі представлено наступні результати досліджень можливості створення якісних двоканальних (стерео) записів музики на основі одноканального запису та оцінок бінауральних імпульсних характеристик приміщення.

- Визначено актуальність теми використання бінауральних імпульсних характеристик приміщення для створення двоканальних записів музики та вказано на потенційну можливість даної технології створювати записи звуку, якість яких є максимально наближеною до якості звуку в реальному приміщенні.
- Виконано аналітичний огляд літературних джерел, де розглядаються технічні, психоакустичні та практичні аспекти досліджуваної теми, що дозволило детально розробити план експериментальних досліджень. Основними складовими розробленого плану є підготовка музичного матеріалу, формування програмно-апаратного забезпечення, запис бінауральних імпульсних характеристик, проведення суб'єктивного та об'єктивного оцінювання якості створених записів.
- Виконано формування музичного матеріалу для досліджень, визначено параметри сигналів для подальшої обробки, визначено параметри акустично підготовленого приміщення та перелік необхідних пристроїв для проведення експериментальних досліджень.
- Отримано оцінки бінауральних імпульсних характеристик приміщення та розроблено алгоритми та комп'ютерні програми їх використання для створення та оцінювання якості звукових записів.

- Отримано оцінки якості створених 2-канальних записів шляхом оцінювання суб'єктивної міри якості DMOS та обчислення об'єктивних мір якості SSNR та LSD.
- Виконано порівняння усереднених результатів суб'єктивного оцінювання з результатами об'єктивного оцінювання, за результатами якого вказано на складність розрізнення реальних та штучно створених бінауральних записів, що свідчить про високу якість синтезованих 2-канальних записів.



## Література

1. Abouchacra K. S. Binaural and spatial hearing in real and virtual environments. *Ear and hearing*. 1998. Vol. 19, no. 3. P. 250–252. URL: <https://doi.org/10.1097/00003446-199806000-00009> (дата звернення: 02.10.2023).
2. Audio programming book / M. V. Mathews et al. MIT Press, 2010. 920 p.
3. Auditory processing of speechlike sounds. *Acoustic phonetics*. 2000. URL: <https://doi.org/10.7551/mitpress/1072.003.0005> (дата звернення: 12.10.2023).
4. Green D. M. Binaural phenomena. *An introduction to hearing*. London, 2021. P. 200–234. URL: <https://doi.org/10.4324/9781003167143-8> (дата звернення: 05.11.2023).
5. Immersive audio signal processing / ed. by S. Bharitkar, C. Kyriakakis. New York, NY : Springer New York, 2006. URL: <https://doi.org/10.1007/0-387-28503-2> (дата звернення: 17.11.2023).
6. Immersive sound / ed. by A. Roginska, P. Geluso. New York ; London : Routledge, 2017. : Routledge, 2017. URL: <https://doi.org/10.4324/9781315707525> (дата звернення: 17.11.2023).
7. Johnson K. Acoustic and auditory phonetics. Wiley & Sons, Incorporated, John, 2011. 256 p.
8. Jones D. Spatial audio. *Nature*. 2002. Т. 417, № 6888. С. 503. URL: <https://doi.org/10.1038/417503a> (дата звернення: 19.11.2023).
9. Roginska A. Binaural audio through headphones. *Immersive sound*. New York ; London : Routledge, 2017., 2017. P. 88–123. URL: <https://doi.org/10.4324/9781315707525-5> (дата звернення: 19.11.2023).

10. Roginska A., Geluso P. Immersive sound: the art and science of binaural and multi-channel audio. Taylor & Francis Group, 2017. 364 p.
11. Surround sound recording techniques. *Spatial audio*. 2012. С. 199–244. URL: <https://doi.org/10.4324/9780080498195-12> (дата звернення: 19.11.2023).
12. Technology of binaural listening. Springer-Verlag Berlin and Heidelberg GmbH &, 2013. (дата звернення: 19.11.2023).
13. The art of immersion: how the digital generation is remaking Hollywood, Madison Avenue, and the way we tell stories. *Choice reviews online*. 2011. Vol. 49, no. 03. P. 49–1292–49–1292. URL: <https://doi.org/10.5860/choice.49-1292> (дата звернення: 19.11.2023).
14. Vorländer M. Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality. Springer London, Limited, 2007. (дата звернення: 21.11.2023).
15. Watkinson J. Art of sound reproduction. Taylor & Francis Group, 2012. (дата звернення: 21.11.2023).
16. К.С. Замша, Б.В. Лозинский, Ю.А. Митяй, Е.С. Степановская, А.Н. Продеус. Объективное и субъективное оценивание качества речевых сигналов с ограниченной полосой частот. - *Electronics and Communications*, Vol. 21, № 1(90), 2016 (дата звернення: 21.11.2023).
17. Печена, М. В. Автоматизована система оцінювання якості спотворених акустичних сигналів : магістерська дисертація : 171 Електроніка / Печена Марина Василівна. - Київ, 2021. - 94 с. Доступ: [https://ela.kpi.ua/bitstream/123456789/49769/1/Pechena\\_mahistr.pdf](https://ela.kpi.ua/bitstream/123456789/49769/1/Pechena_mahistr.pdf)
18. Студійні монітори RCF AYRA PRO5. Доступ: <https://ua.prodj.com.ua/studio-monitors/rcf-ayra-pro5.html>

- 19.Інструментальні мікрофони SE ELECTRONICS SE8. Доступ:  
<https://ua.prodj.com.ua/instrumental-microphones/se-electronics-s-e-8.html>
- 20.Звукові карти Focusrite Clarett 4Pre USB. Доступ:  
<https://ua.prodj.com.ua/sound-cards/focusrite-clarett-4-pre-usb.html>
- 21.Free Space Pro II Binaural Microphone. Доступ:  
<https://3diosound.com/products/free-space-pro-binaural-microphone>

## Додатки

### Додаток 1: Програма zgortka.m

```
% === x_h_convolve.m ===
% Це програма згортки сигналу x із БІХ h приміщення.
% Сигнал x може бути 1-канальним або 2-канальним
% БІХ h є двоканальним сигналом
%
% === План обчислень: ===
% 1) зчитуємо h, нормуємо його та фіксуємо Fs (частота дискретиз.)
% 2) вирізаємо із h відрізок тривалістю T=0.5 сек (починаючи із "голови")
% 3) зчитуємо сигнал x, аналізуємо кількість каналів та fs (частота дискретиз.)
% 4) передискретизуємо сигнал із fs в Fs, якщо це потрібно
% 5) якщо x є 1-канальним, робимо його 2-канальним
% 6) обчислюємо у як згортку x та h
% 7) нормуємо у та записуємо результат на диск під потрібним ім'ям

clc; clear

T = 2;      % тривалість "голови" БІХ в секундах
% 1) зчитуємо h, нормуємо його та фіксуємо Fs (частота дискретиз.)
filename_h = input('Вкажіть ім'я файлу запису БІХ (без лапок): ', 's');
[h,Fs] = audioread(filename_h);
mxh = max(abs(h));
[minmxh,minx] = min(mxh);
h = h/minmxh;

% 2) вирізаємо із h відрізок тривалістю T=2 сек (починаючи із "голови")
NT = round(T*Fs);      % тривалість "голови" БІХ у вибірках (семплах)
hpor = 0.5*minmxh;     % поріг, перевищення якого дає початок "голови" h
n0 = find(diff(sign(abs(h(:,minx))-hpor)),1); % відлік початку "голови" БІХ
h1 = h(n0:n0+NT,:);   % відрізок БІХ тривалістю T сек

stop_check = 1;

while (stop_check > 0)
% 3) зчитуємо сигнал x, аналізуємо кількість каналів та fs (частота дискретиз.)
filename_x = input('Вкажіть ім'я файлу запису сигналу (без лапок): ', 's');
[x,fs] = audioread(filename_x);
numchan = size(x,2);   % кількість каналів в сигналі

% 4) передискретизуємо сигнал із fs в Fs, якщо це потрібно
if fs ~= Fs
    x1 = resample(x,Fs,fs); % передискретизація в частоту дискретизації БІХ
    if numchan == 1
        y1 = filter(h1(:,1),1,x1(:));
        y = [[y1],[y1]];
    else
        x2 = x1;
        y1 = filter(h1(:,1),1,x2(:,1));
        y2 = filter(h1(:,2),1,x2(:,2));
        y = [[y1],[y2]];
    end
else

```

```

if numchan == 1
    y1 = filter(h1(:),1,x(:));
    y = [[y1],[y1]];
else
    x2 = x;
    y1 = filter(h1(:,1),1,x2(:,1));
    y2 = filter(h1(:,2),1,x2(:,2));
    y = [[y1],[y2]];
end
end

%% 5) якщо x є 1-канальним, робимо його 2-канальним
% if numchan == 1
%     x2 = [x1 x1];
% else
%     x2 = x1;
% end

% 6) обчислюємо y як згортку x та h
% y1 = filter(h1(:,1),1,x2(:,1));
% y2 = filter(h1(:,2),1,x2(:,2));
% y = [y1 y2];

% 7) нормуємо y та записуємо результат на диск під потрібним ім'ям
y = y/max(max(abs(y)))*0.99;
filename_y = input('Вкажіть ім"я файлу згортки (без лапок): ', 's');
audiowrite(filename_y,y,Fs);

stop_check = input('Продовжувати обробку сигналів? (1 - Так, 0 - Ні) ');
fprintf('\n');
end

```

## Додаток 2 [17]: Програма ekspertiza3\_m.m

```
% === ekspertiza3_m.m =====
% управляющая программа для субъективной оценки качества звучания речевого сигнала
% при различных значениях полосы частот, занимаемой сигналом
% разработана 13.09.2015, разработчик Митяй Ю.
% модернизирована 20.09.2015 и 1.09.2021, Продеус А.

% Шкала качества MOS при прослушивании речевого сигнала
%   Качество речи   Оценка
%   Отличное       5
%   Хорошее        4
%   Удовлетворительное 3
%   Посредственное 2
%   Плохое         1

clc           %очистка командного окна
clear        %очистка памяти данных программы , т.н. workspace
fs=44100;    %частота дискретизации, Гц

v=input('Введите имя эталонного wav-файла (без кавычек, без расширения wav): ', 's');
%ввод имени файла
[y, Fd]=audioread([v '.wav']);           %чтение эталонного звукового файла

N=input('Укажите время прослушивания (не более 15 сек): '); %ввод времени
прослушивания
y=y(1:Fd*N);           %вырезание первых Fd*N отсчетов сигнала

if Fd~=fs             % передискретизация загруженного файла
    y=resample(y, fs, Fd);
end

y = y/(max(abs(y)))*0.99; % нормирование по максимуму
playery = audioplayer(y,fs);

df=5e2:5e2:10.5e3;    %задание вектора значений полосы частот сигнала
z=randperm(length(df)); %создание вектора целых чисел от 1 до "размер вектора
df" и перемешивание их в случайном порядке
%disp('Эталонный файл'); %вывод на экран информации
pause on             %включение команды "пауза"
%pause (0.5)
%wavplay(y,fs);     %проигрывание эталонного файла

c = zeros(1,length(z)); % выделение области
ocenka = zeros(1,length(z)); % выделение области

tic                 % включение таймера
clc                 %очистка командного окна

for k=1:length(z)  %перебор значений полосы, НЧ фильтрация и прослушивание

    % === НЧ фильтрация ===
    df1 = df(z(k));
    x = filterlow(y,df1,fs);
    % === конец НЧ фильтрации ===
```

```

x = x/(max(abs(x)))*0.99; % нормирование по максимуму

% запрос на прослушивание эталонного файла
quest = input('Слушать эталонный файл? "1" - Да, "0" - Нет: ');
if quest == 1
    disp('Эталонный файл'); %вывод на экран информации
    pause (0.5)
    play(playery); %проигрывание эталонного файла
end

pause(length(y)/fs)
disp('Искаженный файл №'); %отображение номера опыта
disp(k);
pause(0.6); %пауза 0,6 сек
playerx = audioplayer(x,fs);
play(playerx); %воспроизведение фильтрованного сигнала
g=input('Введите оценку от 1 до 5 (можно дробную): '); %считывание оценки с
командного окна
c(k)=g; %запись оценки в вектор с номером, равным номеру цикла
clear x; %удаление фильтрованного сигнала
clc %очистка командного окна
end %окончание перебора значений полосы

disp('Экспертиза завершена');

% упорядочивание рез-тов оценивания в порядке возрастания значений полосы
for m=1:length(z)
    l=find(z==m);
    ocenka(m)=c(l); % упорядоченные результаты записаны в массив ocenka
end

% сохранение результатов экспертизы
b=input('Введите латинцей вашу фамилию(без кавычек): ','s');
b1 = [v '_' b]; % формирование имени файла с результатами
save(b1,'df','ocenka'); % запись результатов в файл

%построение графика результатов
plot (df,ocenka);grid on;
axis ([400 11e3 0.8 5.2]); % задание границ осей
xlabel ('Полоса частот, Гц');ylabel('MOS');title([v ' - ' b]); % надписи на графике
toc

```

### Додаток 3 [17]: Програма filterlow.m

```
function x = filterlow(y,df,fs)

Fpass = df;
Fstop = Fpass*1.15;
Dpass = 0.057501127785;
Dstop = 0.00031622776602;
dens = 20;
[N, Fo, Ao, W] = firpmord([Fpass, Fstop]/(fs/2), [1 0], [Dpass, Dstop]);
b = firpm(N, Fo, Ao, W, {dens});
Hd = dfilt.dffir(b);
x=filter(Hd, y);
```



## Додаток 4 [17]: Програма oscinka\_ser.m

```
clc;
clear all;
close all;

status_continue = 1;
i = 1;

df_1_array = [], [];
ocenska_1_array = [], [];

while (status_continue > 0)
file_name = input('Введіть назву файла оцінок штучного запису ', 's');
temp_FILE = load(file_name);
df_1_array(:,i) = temp_FILE.df;
ocenska_1_array(:,i) = temp_FILE.ocenska;

status_continue = input('Продовжувати додавання даних з файлів оцінок штучного запису? (1 - Так, 0 - Ні) ');
fprintf('\n');
i = i + 1;
end

status_continue = 1;
i = 1;

df_2_array = [], [];
ocenska_2_array = [], [];

while (status_continue > 0)
file_name = input('Введіть назву файла оцінок реального запису ', 's');
temp_FILE = load(file_name);
df_2_array(:,i) = temp_FILE.df;
ocenska_2_array(:,i) = temp_FILE.ocenska;

status_continue = input('Продовжувати додавання даних з файлів оцінок реального запису? (1 - Так, 0 - Ні) ');
fprintf('\n');
i = i + 1;
end

ocenska_1_mean = mean(ocenska_1_array,2);
df_1_mean = mean(df_1_array,2);
ocenska_2_mean = mean(ocenska_2_array,2);
df_2_mean = mean(df_2_array,2);
figure;
hold on
plot (df_1_mean,ocenska_1_mean);
plot (df_2_mean,ocenska_2_mean);
axis ([400 11e3 0.8 5.2]); % задание границ осей
xlabel ('Полоса частот, Гц');ylabel('MOS');title(['Результати усереднення суб`ективного оцінювання']); % надписи на графике
legend('Штучний запис', 'Реальний запис');
grid on;
hold off
```

## Додаток 5 [17]: Програма ocinka\_vidpovidnosti.m

```
clc;
clear all;
close all;

o = input('Натисніть будь-яку клавішу для початку');
delete_sound = input('Скільки секунд файлів хочете програвати? (секунд) ');
disp(sprintf('\n'));
ocenka = [[],[],[[]];

files_number = 1; %% КІЛЬКІСТЬ ФАЙЛІВ!!!
for i = 1:files_number
    temp_1 = 0;
    check = 0;
    disp(sprintf('Програється музичний твір №%d',i));
    [y, Fs] = audioread(sprintf('Original (%d).wav',i));

    player = audioplayer(y, Fs);
    y = y(1:delete_sound*Fs);
    play(player);
    pause(length(y)/Fs);

    while(temp_1 < 2)
        a = rand(1);
        if((a>0.5 && check~=1) || check==2)
            [y, Fs] = audioread(sprintf('Result (%d).wav',i));
            if(check==2)
                real_test_temp = 1;
            else
                real_test_temp = 2;
            end
            check = 1;
            %disp('not real');
        else
            [y, Fs] = audioread(sprintf('Etalon (%d).wav',i));
            check = 2;
            %disp('real');
        end

        if(temp_1==0)
            disp('Програється перша версія файл');
        else
            disp('Програється другий файл');
        end

        player = audioplayer(y, Fs);
        y = y(1:delete_sound*Fs);
        play(player);
        pause(length(y)/Fs);
        stop(player);

        temp_1 = temp_1 + 1;
    end
end
```

```
ocenka(i,1) = input('Введіть номер аудіозапису, який ви вважаєте реальним (1 чи  
2) ');  
ocenka(i,2) = real_test_temp;  
if(ocenka(i,1) == real_test_temp)  
    ocenka(i,3) = 1;  
else  
    ocenka(i,3) = 0;  
end  
disp(sprintf('\n'));  
end  
disp('Оцінені номери програних файлів, що вважаються реальними:');  
disp(ocenka(:,1));  
disp('Реальні номери програних файлів, що вважаються реальними:');  
disp(ocenka(:,2));  
  
disp('Відсоток вірного розпізнавання сигналів:');  
d = size(ocenka);  
disp(sum(ocenka(3))/d(1)*100);
```

## Додаток 6 [17]: Програма obj\_quality.m

```
% === obj_quality.m ===
% Програма забезпечує обчислення показників SSNR і LSD,
% порівнюючи вихідний речовий сигнал і результат його НЧ фільтрації.
%
% Характеристика набору НЧ фільтрів Ремеза: полоси пропускання від 500 Гц
% до 10500 Гц з кроком 500 Гц; ширина перехідної зони 15% від полоси
% пропускання; нерівномірність 1 дБ в полосі пропускання;
% подавлення сигналу мінус 80 дБ в області заперення

% тривалість кожного з 8 сигналів 5 секунд - при цьому загальне час обчислень
% показників SSNR і LSD становить приблизно 5 хвилин; це час обчислень
% автоматично відображається в кінці на моніторі

clc          %очистка командного вікна
clear       %очистка пам'яті даних програми , т.н. workspace

tic
% === Initial parameters ===
Fs = 44100; % Sampling Frequency
Fpass_begin = 500; % first passband frequency
Fpass_end = 10500; % last passband frequency
Fpass_step = 500; % passband frequency step
nsteps = (Fpass_end - Fpass_begin)/Fpass_step + 1; % filters quantity
nsteps1 = (7500 - Fpass_begin)/Fpass_step + 1; % filters quantity

ssnr = zeros(nsteps,8); % виділення області
lsd = zeros(nsteps,8); % виділення області

% === filter parameters ===
Dpass = 0.057501127785; % Passband Ripple (1 dB)
Dstop = 0.00031622776602; % Stopband Attenuation (-70 dB)
dens = 20; % Density Factor

files_number = 1; %% КІЛЬКІСТЬ ФАЙЛІВ!!!
for m = 1:files_number*3 % кількість дикторів

if(m<(files_number+1))
    filename = sprintf('Etalon (%d).wav',m);
elseif(m<(2*files_number+1))
    filename = sprintf('Result (%d).wav',m-files_number);
else
    filename = sprintf('Original (%d).wav',m-2*files_number);
end

% === input signal ===
[x,Fread] = audioread(filename);
x = x(1:(Fread*5)); %%Обмежуємо лише 5-ма секундами
% x1 = resample(x,2,1); % передискр. с 22050 в 44100 Гц
% x = x1;
% clear x1;

% === Calculation of filter coeffs ===
```

```

for i = 1:nsteps
    Fpass = Fpass_begin + Fpass_step*(i-1); % Passband Frequency
    Fstop = Fpass*1.15; % Stopband Frequency
% Calculate the filter order
[N, Fo, Ao, W] = firpmord([Fpass, Fstop]/(Fs/2), [1 0], [Dpass, Dstop]);
% Calculate the filter coefficients
a = firpm(N, Fo, Ao, W, {dens});
len_a = length(a); % length of a

% анализ на четность-нечетность
if mod(len_a,2)==0
    len_a1 = len_a + 1;
    arg_a = 1:len_a;
    arg_a1 = 0.5:len_a+0.5;
    a1 = interp1(arg_a,a,arg_a1,'spline');
    a = a1;
    len_a = length(a); % length of a
end

%len_a05 = round(0.5*len_a); % half of len_a
len_a05 = floor(0.5*len_a); % half of len_a

% filtering
y = fftfilt(a,x);

% циклический сдвиг результата фильтрации на половину длины ИХ
y = circshift(y,-len_a05);

% откусывание последних отсчетов после циклического сдвига
y = y(1:end-len_a05);
x = x(1:end-len_a05);

% === нормирование сигналов по мощности ===
stdx = std(x);
x = x/stdx;
stdy = std(y);
y = y/stdy;

% quality indexes
ssnr(i,m) = srr_func_correct2(x,y,Fs); % вычисление показателя SRR
lsd(i,m) = lsd_func(x,y,Fs); % вычисление показателя LSD
end
end

% сохранение результатов объективного оценивания в файле
b=input('Введите латиницей вашу фамилию(в кавычках): ');
b1 = ['SSNR_' b]; % формирование имени файла с результатами
b2 = ['LSD_' b]; % формирование имени файла с результатами
save(b1,'ssnr'); % запись оценки SSNR в файл
save(b2,'lsd'); % запись оценки LSD в файл

% === Graphics ===
f = Fpass_begin:Fpass_step:Fpass_end;

```

```

% Графики для РЕАЛЬНИХ СИГНАЛІВ
figure;
for k = 1:files_number
plot(f,ssnr(:,k)); % segmental SNR
xlabel('Частота, Гц'); ylabel('SSNR, дБ'); hold on;
title(['SSNR (Реальний сигнал) ' b]); grid on;
end
figure;
for k = 1:files_number
plot(f,lsd(:,k)); % LSD
xlabel('Частота, Гц'); ylabel('LSD, дБ'); hold on;
title(['LSD (Реальний сигнал) ' b]); grid on;
end

% Графики для ШТУЧНИХ СИГНАЛІВ
figure;
for k = (files_number+1):2*files_number
plot(f,ssnr(:,k)); % segmental SNR
xlabel('Частота, Гц'); ylabel('SSNR, дБ'); hold on;
title(['SSNR (Штучний сигнал) ' b]); grid on;
end
figure;
for k = (files_number+1):2*files_number
plot(f,lsd(:,k)); % LSD
xlabel('Частота, Гц'); ylabel('LSD, дБ'); hold on;
title(['LSD (Штучний сигнал) ' b]); grid on;
end

% Графики для ОРИГІНАЛІВ
figure;
for k = (2*files_number+1):3*files_number
plot(f,ssnr(:,k)); % segmental SNR
xlabel('Частота, Гц'); ylabel('SSNR, дБ'); hold on;
title(['SSNR (Оригінал) ' b]); grid on;
end
figure;
for k = (2*files_number+1):3*files_number
plot(f,lsd(:,k)); % LSD
xlabel('Частота, Гц'); ylabel('LSD, дБ'); hold on;
title(['LSD (Оригінал) ' b]); grid on;
end

% Графіки порівняння сигналів
for k = 1:files_number
figure;
plot(f,ssnr(:,k+2*files_number)); % segmental SNR
hold on;
plot(f,ssnr(:,k));
plot(f,ssnr(:,k+files_number)); % segmental SNR
legend(sprintf('Оригінал №%d',k),sprintf('Реальний сигнал №%d',k),sprintf('Штучний
сигнал №%d',k));
xlabel('Частота, Гц'); ylabel('SSNR, дБ');
title(['SSNR ' b]); hold off; grid on;
end

```

```

for k = 1:files_number
figure;
plot(f,lsd(:,k+2*files_number)); % segmental SNR
hold on;
plot(f,lsd(:,k));
plot(f,lsd(:,k+files_number)); % segmental SNR
legend(sprintf('Оригінал №%d',k),sprintf('Реальний сигнал №%d',k),sprintf('Штучний
сигнал №%d',k));
xlabel('Частота, Гц'); ylabel('LSD, дБ');
title(['LSD ' b]); hold off; grid on;
end

```

%Графіки порівняння середніх значень SSNR та LSD

```

figure;
ssnrmean_1 = mean(ssnr(:,1:files_number),2); % усреднение SSNR
ssnrmean_2 = mean(ssnr(:,(files_number+1):2*files_number),2); %
усреднение SSNR
ssnrmean_3 = mean(ssnr(:,(2*files_number+1):3*files_number),2); %
усреднение SSNR
plot(f,ssnrmean_3,'r');
hold on;
plot(f,ssnrmean_1,'b');
plot(f,ssnrmean_2,'m');
xlabel('Частота, Гц'); ylabel('SSNR, дБ');
legend('Оригінал','Реальний сигнал','Штучний сигнал');
title(['SSNR average ' b]); hold off; grid on;

```

```

figure;
lsdmean_1 = mean(lsd(:,1:files_number),2); % усреднение LSD
lsdmean_2 = mean(lsd(:,(files_number+1):2*files_number),2); %
усреднение LSD
lsdmean_3 = mean(lsd(:,(2*files_number+1):3*files_number),2); %
усреднение LSD
plot(f,lsdmean_3,'r');
hold on;
plot(f,lsdmean_1,'b');
plot(f,lsdmean_2,'m');
xlabel('Частота, Гц'); ylabel('LSD, дБ');
legend('Оригінал','Реальний сигнал','Штучний сигнал');
title(['LSD average ' b]); hold off; grid on;

```

toc





## Додаток 8 [17]: Програма lsd\_func.m

```
function lsd = lsd_func(x,y,fs)

%           LSD = 2/(K*L)* sum_l&k|LX(1,k)-LY(1,k)|
%
% LX(1,k) = max_l_k {20*log10[X(1,k)],delta};
% delta = max_l_k {20*log10[X(1,k)]} - 50;
%
% =====

no=2;           % overlap factor = (fft length)/(frame increment)
ti=16e-3;      % desired frame increment (16 ms)
ni=round(ti*fs); % frame increment in samples
nf=ni*no;      % fft length
w=sqrt(hamming(nf+1))'; w(end)=[]; % for now always use sqrt hamming window
w=w/sqrt(sum(w(1:ni:nf).^2)); % normalize to give overall gain of 1
rfm = 'r';

[x1,tt]=enframe(x,w,ni,rfm);
x1f=rfft(x1,nf,2);
X=abs(x1f); % amplitude spectrum of input speech
[L,K2]=size(X);% L - number of frames; K2 - number of spectrum samples
Xlog = 20*log10(X);
deltax = max(max(Xlog))-50;
for l = 1:L
    for k = 1:K2
        LX(1,k)=max([Xlog(1,k) deltax]);
    end
end

[y1,tt]=enframe(y,w,ni,rfm);
y1f=rfft(y1,nf,2);
Y=abs(y1f); % amplitude spectrum of input speech
%[L,K2]=size(Y);% nr - number of frames; nf2 - number of spectrum samples
Ylog = 20*log10(Y);
deltay = max(max(Ylog))-50;
for l = 1:L
    for k = 1:K2
        LY(1,k)=max([Ylog(1,k) deltay]);
    end
end

lsd = mean(mean(abs(LX-LY)));
```

## Додаток 9 [17]: Програма `enframe.m`

```
function [f,t,w]=enframe(x,win,inc,m)
%ENFRAME split signal up into (overlapping) frames: one per row. [F,T]=(X,WIN,INC)
%
% Usage: (1) f=enframe(x,n)      % split into frames of length n
%
%         (2) f=enframe(x,hamming(n,'periodic'),n/4)    % use a 75% overlapped
Hamming window of length n
%
%         (3) frequency domain frame-based processing:
%
%             S=...;                % input signal
%             OV=2;                  % overlap factor of 2 (4 is also
often used)
%             INC=20;                % set frame increment in samples
%             NW=INC*OV;              % DFT window length
%             W=sqrt(hamming(NW,'periodic')); % omit sqrt if OV=4
%             W=W/sqrt(sum(W(1:INC:NW).^2)); % normalize window
%             F=rfft(enframe(S,W,INC),NW,2); % do STFT: one row per time frame,
+ve frequencies only
%             ... process frames ...
%             X=overlapadd(irfft(F,NW,2),W,INC); % reconstitute the time waveform
(omit "X=" to plot waveform)
%
% Inputs:  x    input signal
%         win  window or window length in samples
%         inc  frame increment in samples
%         m    mode input:
%             'z' zero pad to fill up final frame
%             'r' reflect last few samples for final frame
%             'A' calculate window times as the centre of mass
%             'E' calculate window times as the centre of energy
%
% Outputs: f    enframed data - one frame per row
%         t    fractional time in samples at the centre of each frame
%         w    window function used
%
% By default, the number of frames will be rounded down to the nearest
% integer and the last few samples of x() will be ignored unless its length
% is no more than a multiple of inc. If the 'z' or 'r' options are given,
% the number of frame will instead be rounded up and no samples will be ignored.
%
% Example of frame-based processing:
%         INC=20                                % set frame increment in
samples
%         NW=INC*2                               % oversample by a factor
of 2 (4 is also often used)
%         S=cos((0:NW*7)*6*pi/NW);                % example input signal
%         W=sqrt(hamming(NW,'periodic')); % sqrt hamming window of period NW
%         F=enframe(S,W,INC);                    % split into frames
%         ... process frames ...
%         X=overlapadd(F,W,INC);                  % reconstitute the time waveform
(omit "X=" to plot waveform)
```

```

% Bugs/Suggestions:
% (1) Possible additional mode options:
%     'u'  modify window for first and last few frames to ensure WOLA
%     'a'  normalize window to give a mean of unity after overlaps
%     'e'  normalize window to give an energy of unity after overlaps
%     'wm' use Hamming window
%     'wn' use Hanning window
%     'x'  include all frames that include any of the x samples

%     Copyright (C) Mike Brookes 1997-2012
%     Version: $Id: enframe.m 2470 2012-11-02 15:27:24Z dmb $
%
% VOICEBOX is a MATLAB toolbox for speech processing.
% Home page: http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This program is free software; you can redistribute it and/or modify
% it under the terms of the GNU General Public License as published by
% the Free Software Foundation; either version 2 of the License, or
% (at your option) any later version.
%
% This program is distributed in the hope that it will be useful,
% but WITHOUT ANY WARRANTY; without even the implied warranty of
% MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
% GNU General Public License for more details.
%
% You can obtain a copy of the GNU General Public License from
% http://www.gnu.org/copyleft/gpl.html or by writing to
% Free Software Foundation, Inc., 675 Mass Ave, Cambridge, MA 02139, USA.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

nx=length(x(:));
if nargin<2 || isempty(win)
    win=nx;
end
if nargin<4 || isempty(m)
    m='';
end
nwin=length(win);
if nwin == 1
    lw = win;
    w = ones(1,lw);
else
    lw = nwin;
    w = win(:)';
end
if (nargin < 3) || isempty(inc)
    inc = lw;
end
nli=nx-lw+inc;
nf = fix((nli)/inc);
na=nli-inc*nf;
f=zeros(nf,lw);
indf= inc*(0:(nf-1)).';
inds = (1:lw);

```

```

f(:) = x(indf(:,ones(1,lw))+inds(ones(nf,1),:));
if nargin>3 && (any(m=='z') || any(m=='r')) && na>0
    if any(m=='r')
        ix=1+mod(nx-na:nx-na+lw-1,2*nx);
        f(nf+1,:)=x(ix+(ix>nx).*(2*nx+1-2*ix));
    else
        f(nf+1,1:na)=x(1+nx-na:nx);
    end
    nf=size(f,1);
end
if (nwin > 1) % if we have a non-unity window
    f = f .* w(ones(nf,1),:);
end
if nargin>1
    if any(m=='E')
        t0=sum((1:lw).*w.^2)/sum(w.^2);
    elseif any(m=='E')
        t0=sum((1:lw).*w)/sum(w);
    else
        t0=(1+lw)/2;
    end
    t=t0+inc*(0:(nf-1)).';
end

```

## Додаток 10 [17]: Програма l rfft.m

```
function y=rfft(x,n,d)
%RFFT Calculate the DFT of real data Y=(X,N,D)
% Data is truncated/padded to length N if specified.
% N even: (N+2)/2 points are returned with
%         the first and last being real
% N odd: (N+1)/2 points are returned with the
%         first being real
% In all cases fix(1+N/2) points are returned
% D is the dimension along which to do the DFT

% Copyright (C) Mike Brookes 1998
% Version: $Id: rfft.m 713 2011-10-16 14:45:43Z dmb $
%
% VOICEBOX is a MATLAB toolbox for speech processing.
% Home page: http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This program is free software; you can redistribute it and/or modify
% it under the terms of the GNU General Public License as published by
% the Free Software Foundation; either version 2 of the License, or
% (at your option) any later version.
%
% This program is distributed in the hope that it will be useful,
% but WITHOUT ANY WARRANTY; without even the implied warranty of
% MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
% GNU General Public License for more details.
%
% You can obtain a copy of the GNU General Public License from
% http://www.gnu.org/copyleft/gpl.html or by writing to
% Free Software Foundation, Inc., 675 Mass Ave, Cambridge, MA 02139, USA.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

s=size(x);
if prod(s)==1
    y=x;
else
    if nargin <3 || isempty(d)
        d=find(s>1,1);
        if nargin<2
            n=s(d);
        end
    end
    if isempty(n)
s
    end
    y=fft(x,n,d);
    y=reshape(y,prod(s(1:d-1)),n,prod(s(d+1:end)));
    s(d)=1+fix(n/2);
    y(:,s(d)+1:end,:)=[];
    y=reshape(y,s);
end
```