ADVANCED AUDIO CODING

How AAC works, its improvements over MP3, and its applications in streaming platforms and digital broadcasting

Pedro Carneiro, Inês Águia, João Oliveira

BACKGROUND

Advanced Audio Coding was introduced in the **late 1990s**, companies like *Fraunhofer, Dolby, Sony and AT&T* joined forces to **overcome the growing mp3 limitations, and as part of the MPEG-2 and later MPEG-4 standards.**
Audio formats began to **shift from analog** (vinyl, cassettes tapes) **to digital** (CD and digital files).

Consumers demanded **portable audio formats**, with the rise of **portable audio players**.

MPEG-X -> Moving Picture Experts Group, to address different needs in audio and video compression

PROBLEMS

GROWING DEMAND FOR BETTER QUALITY

As portable devices and streaming platforms evolved, users demanded higher audio quality without sacrificing storage or streaming performance.

Consumer Expectations

POOR ERROR RESILIENCE

MP3 lacked robust error correction mechanisms, which made it more susceptible to issues like audio dropouts during streaming over unstable networks.

Streaming Interruptions

LARGER FILE SIZES

Although MP3 significantly reduced file sizes compared to uncompressed formats, it was still inefficient compared to emerging technologies, especially for mobile devices with limited storage.

Storage constraints

LACK OF FLEXIBILITY

MP3 was primarily designed for stereo audio, making it less suitable for multi-channel formats like surround sound, which were becoming increasingly popular in home theaters and digital broadcasting.

Limited Channel Support

LIMITED FREQUENCY RANGE

MP3 could only handle frequencies up to 16 kHz well (compared to human hearing up to ~20 kHz), which led to a noticeable loss of detail, especially in high-quality headphones or sound systems.

Reduced high-frequency audio

INEFFICIENCY AT LOW BIT RATES

Sound quality was compromised below 128kbps, more noticeable in complex audios such as orchestral music or high-pitched vocals making sound distorted or “muddy”. With the rise of the internet streaming , MP3’s inefficiency at low bit rates became problematic for users with limited bandwidth.

Sound quality degradation

SOLUTION

MAIN FEATURES

Like MP3, AAC is a lossy format, meaning it reduces file size by discarding inaudible data.
However, it uses more advanced techniques to achieve this.

AAC

PRE-PROCESSING AND ANALYSIS

PSYCHOACOUSTIC MODELLING

ADVANCED TOOLS FOR EFFICIENCY

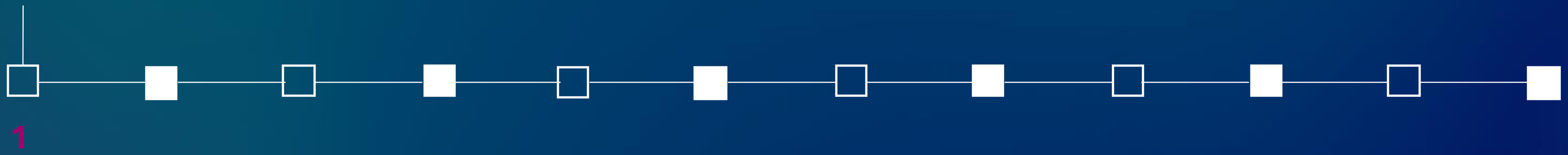
TIME-TO-FREQUENCY TRANSFORMATION

QUANTIZATION AND CODING

MULTIPLEXING AND PACKAGING

HOW DOES AAC WORK?

Pre-Processing
and
Analysis



Pre-Processing

Segmentation
into frames

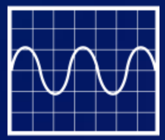
Psychoacoustic Analysis

Amplitude
Variation

Wave form
shape

Transient
Events

Periodicity



ADSR

HOW DOES AAC WORK?

MDCT – Modified Discrete Cosine Transform

Block Switching:

- **Long blocks(1024 samples)** for stationary sounds
- **Short blocks (128 samples)** for transient signals

HOW DOES AAC WORK?

Pre-Processing and Analysis

Psychoacoustic Modelling

Psychoacoustic Modelling

1

2

3

This model determines which audio can be safely discarded or reduced in precision, improving compression without a perceptible change in sound quality.

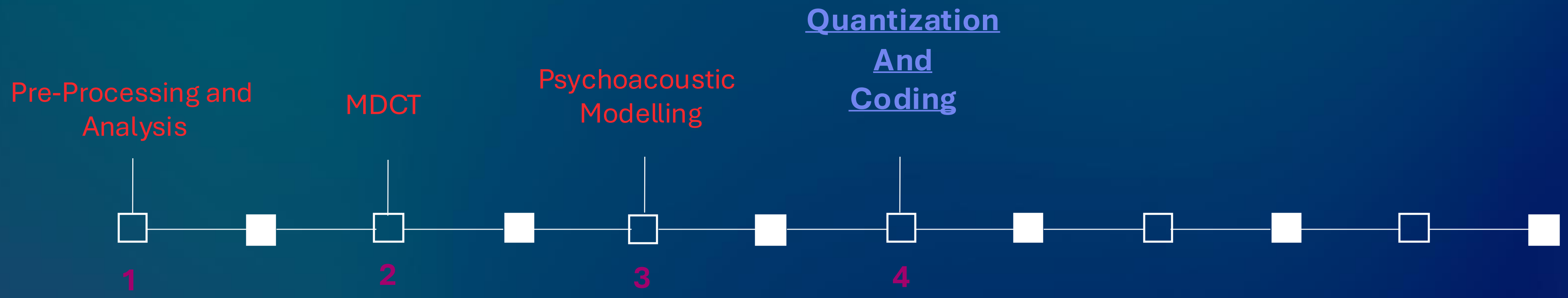
Spectrum Analysis

Auditory Masking Phenomena

Human Hearing Sensitivity/ Threshold

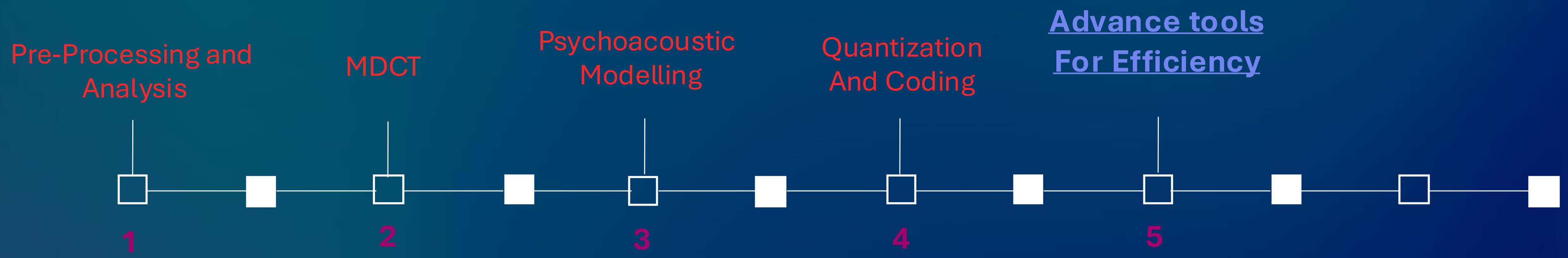
Quantization based on Perceptual Relevance

HOW DOES AAC WORK?



1. Quantize frequency to reduce bit rate, based on Psychoacoustic analysis
2. **Prediction Techniques**:
   1. **TNS** (Temporal Noise Shaping) – shape quantization noise over time (better handling transient signals)
   2. **Prediction Models**: temporal and spacial redundancy in the audio are predicted and encoded (reducing data size preserving audio fidelity)
3. **Entropy Coding (Huffman)** – the quantized values are further compressed using Huffman coding (shorted codes for more frequent values)

HOW DOES AAC WORK?



1. **SBR (Spectral Band Replication)** - used on HE-AAC (high-efficiency AAC)

   - SBR is a bandwidth extension technique that improves compression efficiency (particularly at low bit rates). Reconstructs high-frequency components based on a limited set of data.

2. **PS (Parametric Stereo):**

   - Tool used for encoding stereo audio information more efficiently by using parametric data. Specially used at low bit rate scenarios where traditional stereo encoding would consume excessive bandwidth.

HOW DOES AAC WORK?

1. Pre-Processing and Analysis
2. MDCT
3. Psychoacoustic Modelling
4. Quantization And Coding
5. Advance tools For Efficiency
6. Multiplexing And Packaging

1. **Bitstream Multiplexing**:

   - The encoded data is assembled into a final bitstream format compatible with AAC standards.

2. **Error Resilience** :

   - RVLC (Reversible Variable-Length Coding) : in case of an error, RVLC helps in recovering and decoding parts of the data stream even after error occurs
   - Data Partitioning : if error occurs , less critical data may be discarded preserving the more crucial information
   - Synchronization and Header information : headers and synchronization words are added to frames to mark the start of data packets, helping the decoder to detect and recover from errors more effectively.)
   - Error Concealment :
     - Noise substitution : inserting small amount of controlled noise to mask lost data
     - Interpolation : Using known data before and after the missing data to estimate and fill in the gap
   - EP (error protection layers) : include error protection layers that add extra parity or error checking bits to the stream (for real time requirements)

AAC VS MP3

	AAC	MP3
COMPRESSION EFFICIENCY	✓	✗
FREQUENCY RANGE	✓	✗
CHANNEL SUPPORT	✓	✗
PSYCHOACOUSTIC MODEL	✓	✗
ERROR RESILIENCE	✓	✗
FILE SIZE	✓	✗
COMPATIBILITY	✗	✓
BIT RATE FLEXIBILITY	✓	✗

APPLICATIONS IN STREAMING AND BROADCASTING

STREAMING PLATFORMS

- high efficiency at low bit rates allowing better streaming quality while conserving bandwidth.
- multi-channel support and better compression, high quality in video content while minimizing data usage.

DIGITAL BROADCASTING

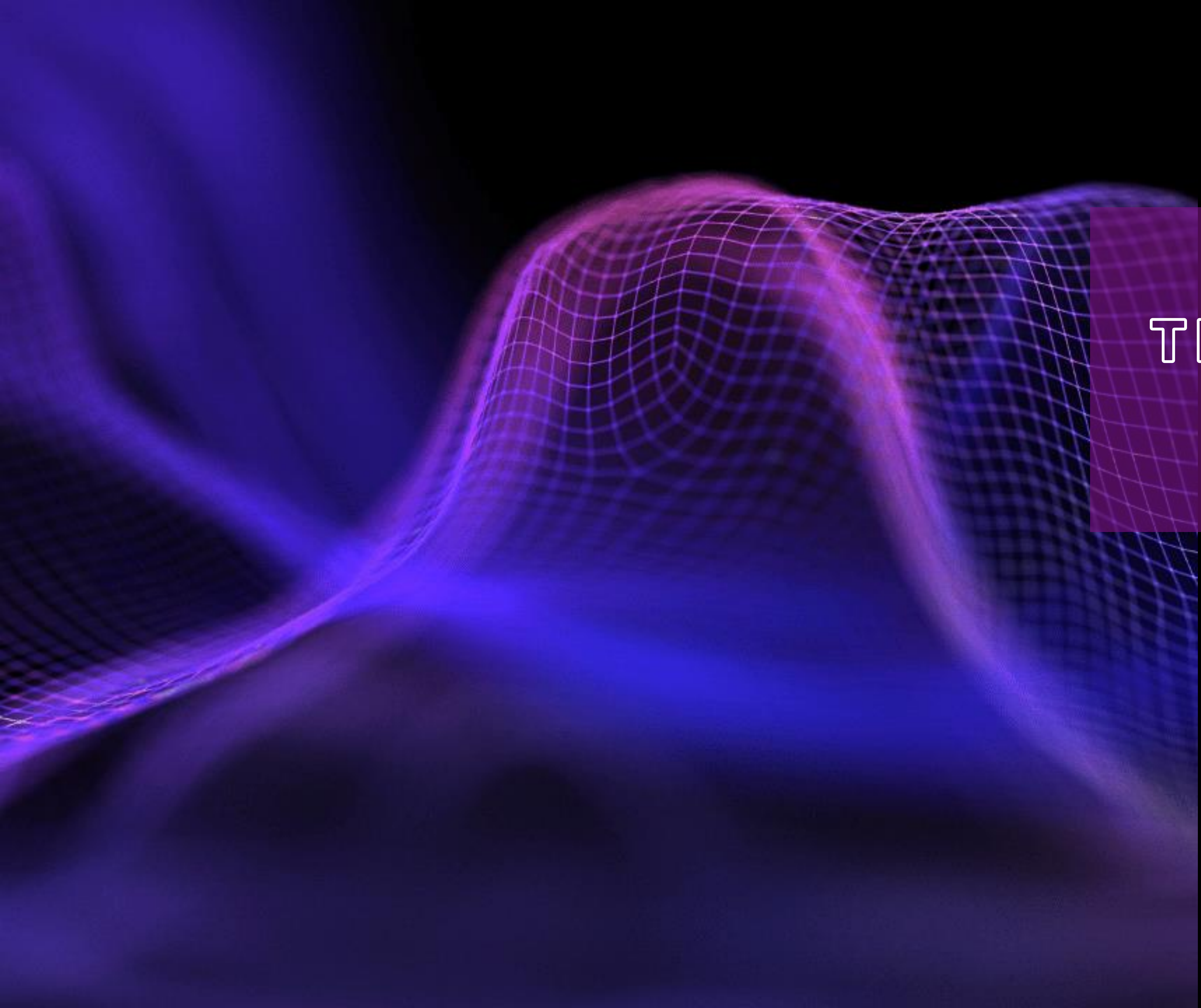
- **Audio (DAB+)** : improved compression, error resilience and ability to handle multichannel configurations for transmitting high-quality audio over limited bandwidth
- **Television (DVB, etc …)** : enables the efficient delivery of high-fidelity (low quality loss) audio in digital TV broadcasting supporting both stereo and multichannel audio.

DBV –> Digital Video Broadcasting
DAB+ -> Digital Audio broadcasting Plus

DRAWBACKS AND CONCLUSION

- AAC has an important role in modern audio technology, **offering a balance between efficiency, quality and versatility**
- High quality sound at **lower bit rates**, great for streaming and broadcasting
- **Lossy** nature leads to **reduction in audio quality when compared to lossless formats like FLAC**, making it less suitable for archival purposes where maximum fidelity is required
- May face **compatibility** issues with **older devices**
- **More complex and resource intensive encoding**, what can lead in **longer processing times** when converting large audio libraries
- Advanced variants like **xHE-AAC emerge**, allowing AAC to keep **growing** and solidifying its status as **a robust and efficient audio format for the digital age**

xHe-AAC -> Extend High Efficiency Advanced Audio Coding



THANK YOU!