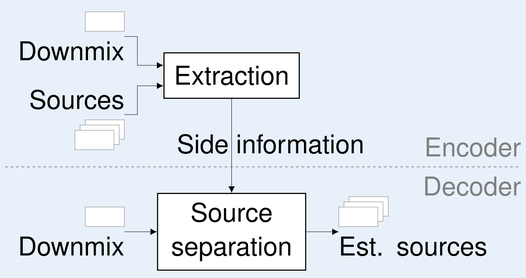
ADAPTIVE CODING OF NON-NEGATIVE FACTORIZATION PARAMETERS WITH APPLICATION TO INFORMED SOURCE SEPARATION

Max Bläser, Christian Rohlfing, Yingbo Gao, Mathias Wien

Institut für Nachrichtentechnik, RWTH Aachen University

Informed source separation

- ISS aims at generating a multichannel signal based on a downmix
- Numerous applications, such as adaptive rendering on loudspeakers arrays, karaoke or active listening

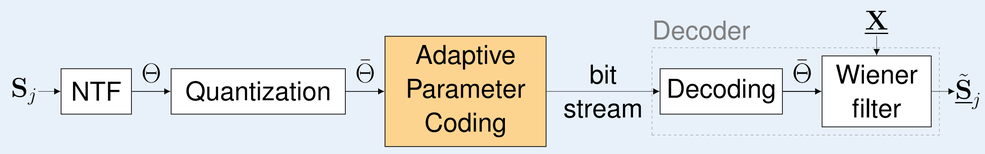


Parametric audio upmixing

Parametric audio upmixing with Non-negative Tensor Factorization (NTF) consists of two steps **[1]**:

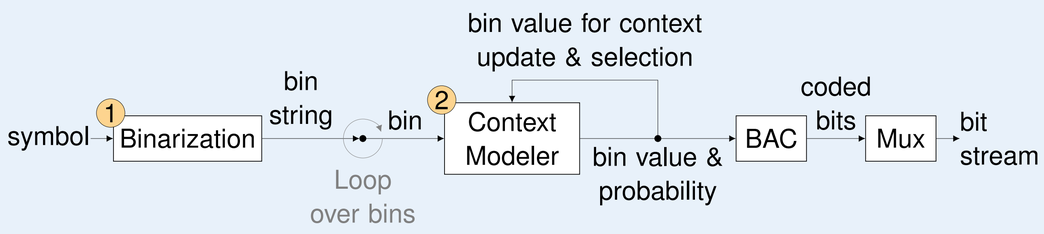
- **Encoder**: Sources and downmix perfectly known
  - NTF on source spectrogram $\mathbf{S}_j$ which computes parameters $\Theta = \{\mathbf{W}, \mathbf{H}, \mathbf{Q}\}$
  - Scalar quantization and subsequent *coding* of parameters with GZIP
- **Decoder**: Only mix $\underline{\mathbf{X}}$ available Sources estimated with Wiener-filtering given $\underline{\mathbf{X}}$ and quantized parameters $\bar{\Theta}$ yielding $\underline{\tilde{\mathbf{S}}}_j$

Our contribution: Code NTF parameters adaptively to exploit *conditional entropy*



Context-based adaptive binary arithmetic coding

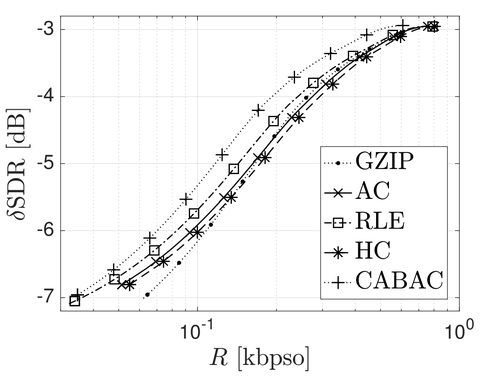
- CABAC **[2]** used in latest video coding standards, e.g. HEVC
- Exploit dependencies within the signal's statistics ⇒ *Conditional entropy*



1. Non-binary symbols binarized yielding bin-strings
2. Context modeling for tracking bin probabilities $p\left(\mathrm{bin} \mid \mathrm{ctx}\right)$ for Binary Arithmetic Coding (BAC)
   - Context model ctx maps to state of information available at en- and decoder
   - Context model design ⇒ performance towards conditional entropy

Experiments

Comparison with reference methods



- GZIP (·-·)
- Arithmetic Coding (AC, -×-)
- Run-length encoding (RLE, -□-)
- Huffman coding (HC, -✳-)
- Proposed method (CABAC, -+-)

Setup:

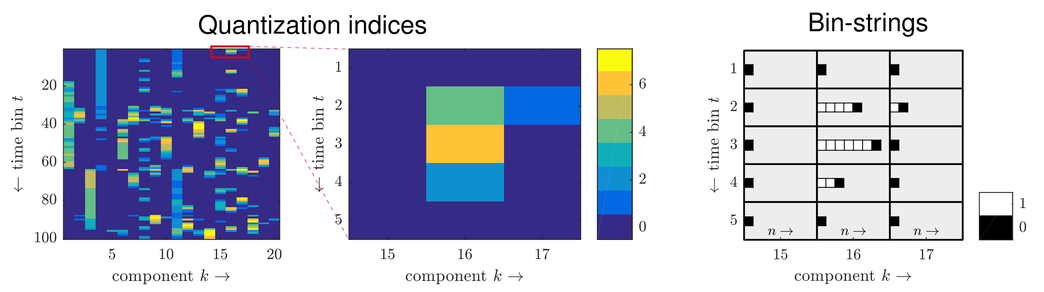
- NTF minimizes Kullback-Leibler divergence with $\{1, 2 \ldots 10\}$ components per source
- Quantization in log-domain with $\{2, 4, 8, 16\}$ levels
- Evaluated on DSD100 test set (**[3]**, 100 mixtures à 4 sources)

⇒ CABAC outperforms all other reference methods

⇒ RLE adequate low-complexity alternative to CABAC. Applicable since data has Markov property

Adaptive parameter coding

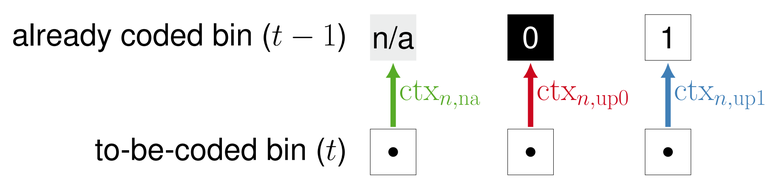
Binarization ①



Binarization of quantization indices $g_{t,k}$ with Truncated Unary:
Sequence of $g_{t,k}$ '1's terminated by a single '0'.

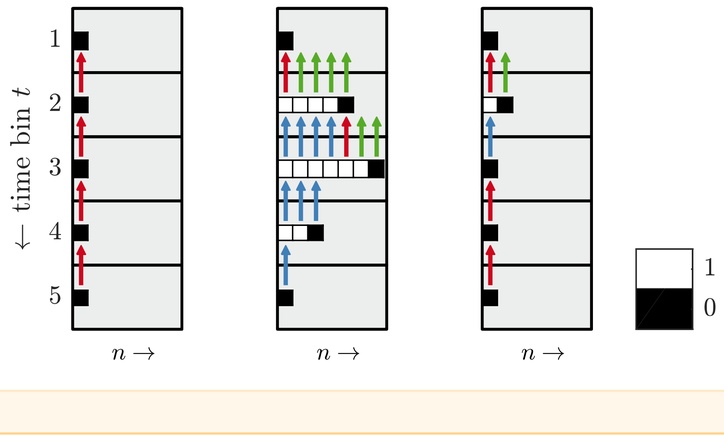
Context model design ②

- Exploit typical structure in NTF parameters (sparseness / continuity)
- Context models chosen on *bin-level* dependent on already coded bin within column k
- Bins at same bin position n considered



- $\mathrm{ctx}_{n,\mathrm{up0}}$ with $n = 1$ exploits sparseness
- Upper limit of modeled bins N_{LBP} ⇒ $3N_{\mathrm{LBP}} + 1$ context models in total

Exemplary context model selections



[1] A. Liutkus, J. Pinel, R. Badeau, L. Girin, and G. Richard, "Informed source separation through spectrogram coding and data embedding," *Signal Processing*, vol. 92, no. 8, pp. 1937 – 1949, 2012.

[2] D. Marpe, H. Schwarz, and T. Wiegand, "Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 620 – 636, 2003.

[3] SiSEC "MUS 2016" task. `http://sisec.inria.fr`.

Find MATLAB implementation of CABAC and code for the ISS method here:

`https://github.com/christianrohlfing/ISScabac`

rohlfing@ient.rwth-aachen.de

www.ient.rwth-aachen.de

Institut für Nachrichtentechnik, Melatener Str. 23, 52074 Aachen

ICASSP 2018