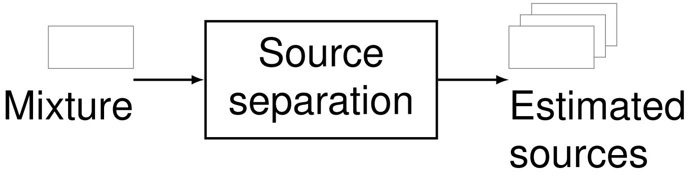
Generalized Constraints for NMF with Application to Informed Source Separation

Lehrstuhl und Institut für Nachrichtentechnik

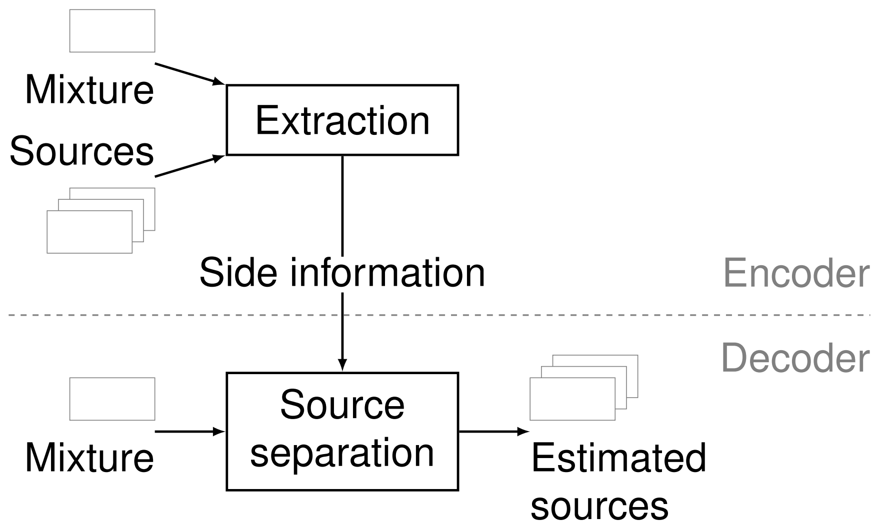
RWTH AACHEN UNIVERSITY

Content

Introduction



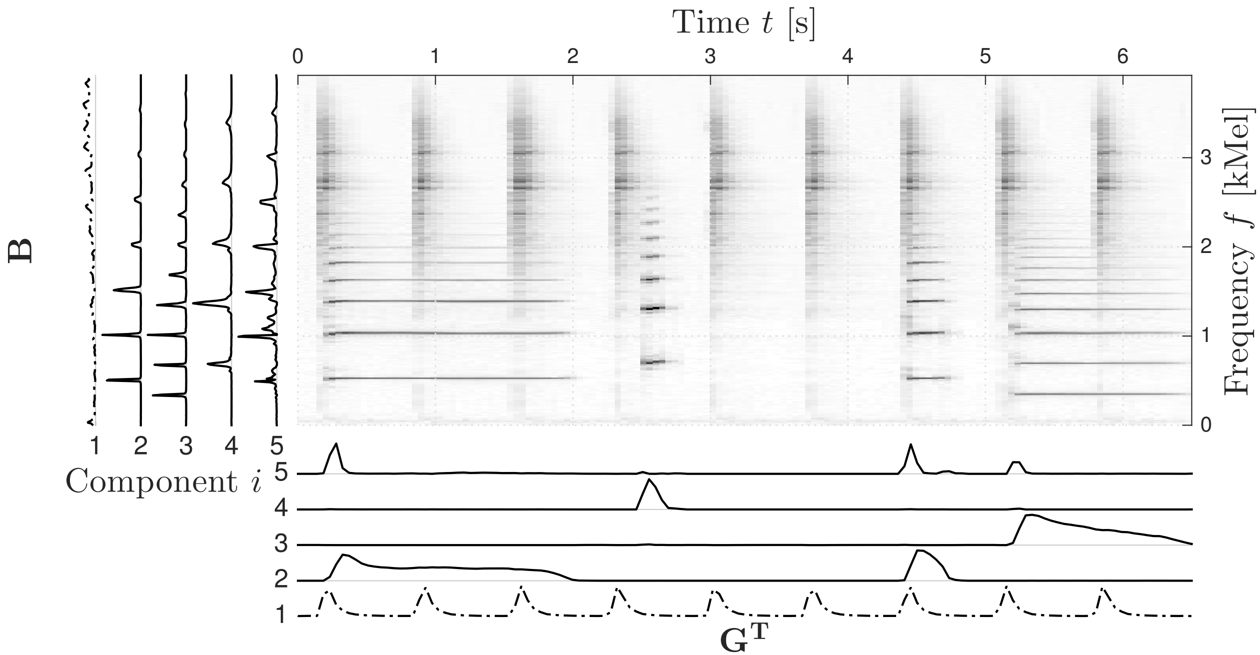
Introduction



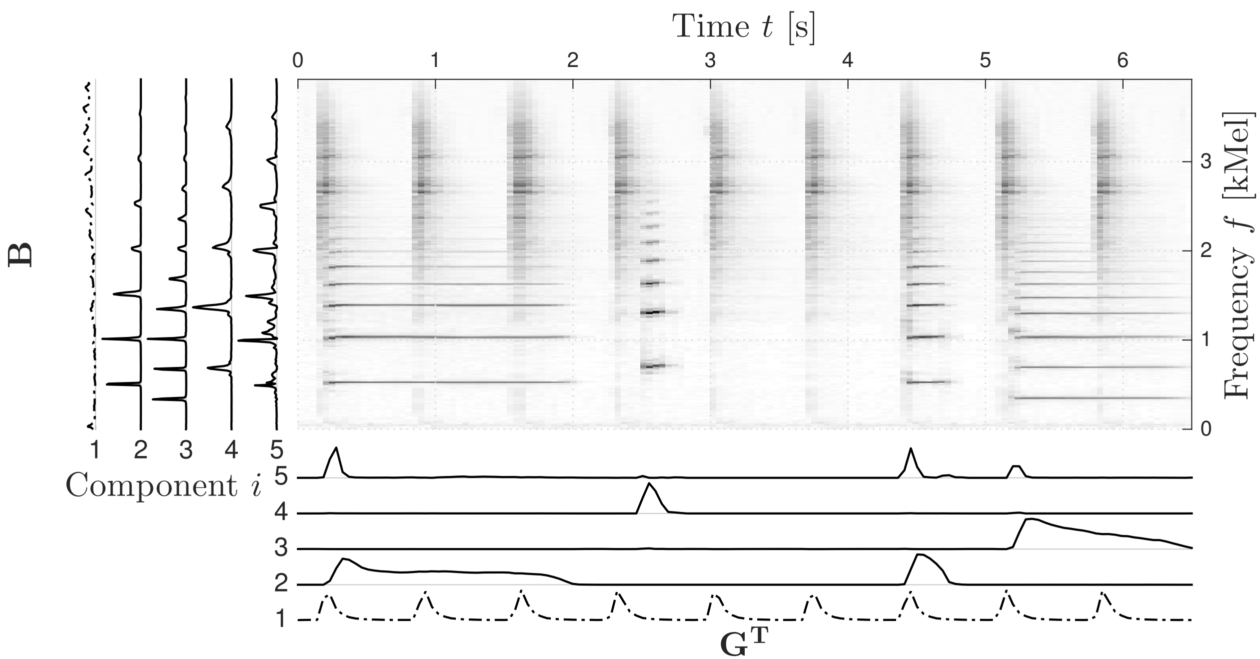
- Informed source separation (ISS) uses source separation for audio object coding
- Active listening and remixing of music (e.g. karaoke) requires audio objects

Content

Nonnegative matrix factorization



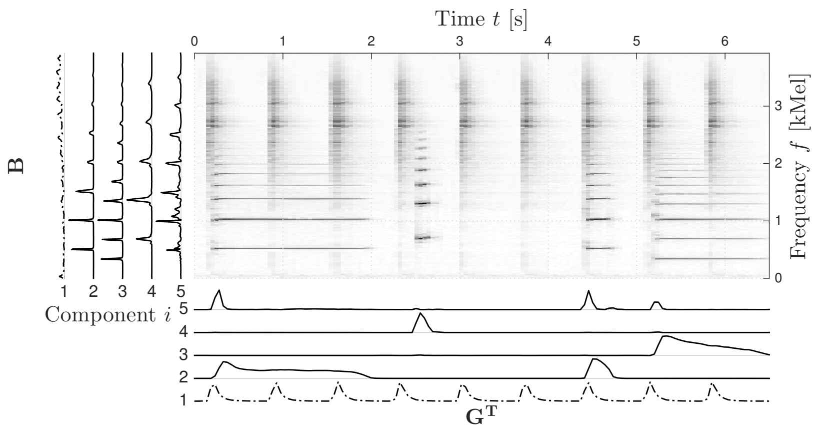
$$\mathbf{X} = |\underline{\mathbf{X}}| \approx \mathbf{B}\mathbf{G}^\mathbf{T}, \quad \mathbf{X} \in \mathbb{R}_+^{F\times T}, \mathbf{B} \in \mathbb{R}_+^{F\times I}, \mathbf{G} \in \mathbb{R}_+^{T\times I}$$



$$\mathbf{X} = |\underline{\mathbf{X}}| \approx \mathbf{B}\mathbf{G}^\mathbf{T}, \quad \mathbf{X} \in \mathbb{R}_+^{F\times T}, \mathbf{B} \in \mathbb{R}_+^{F\times I}, \mathbf{G} \in \mathbb{R}_+^{T\times I}$$

$$\min \quad d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T})$$

Constrained NMF



$$\min \quad d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathrm{T}) + \alpha c(\mathbf{G}) + \gamma c(\mathbf{B})$$

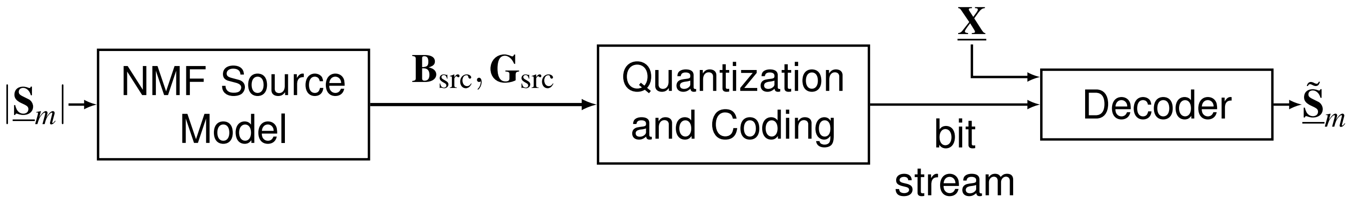
Temporal continuity and temporal sparseness [Virtanen 2007]

$$c_\mathrm{tc}(\mathbf{G}) = \sum_{i=1}^{I} \frac{1}{\sigma_i^2} \sum_{t=2}^{T} [g_{t,i} - g_{t-1,i}]^2 = \sum_i c_\mathrm{tc}(\mathbf{g}_i)$$

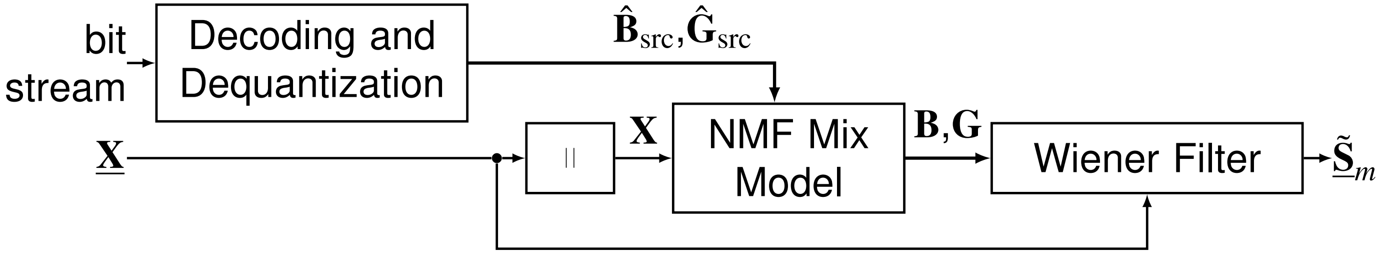
$$c_\mathrm{s}(\mathbf{G}) = \sum_{i=1}^{I} \sum_{t=1}^{T} \frac{g_{t,i}}{\sigma_i} = \sum_i c_\mathrm{s}(\mathbf{g}_i) \qquad \text{with } \sigma_i = \sqrt{1/T \sum_t g_{t,i}^2}$$

NMF-based ISS

Encoder



Decoder



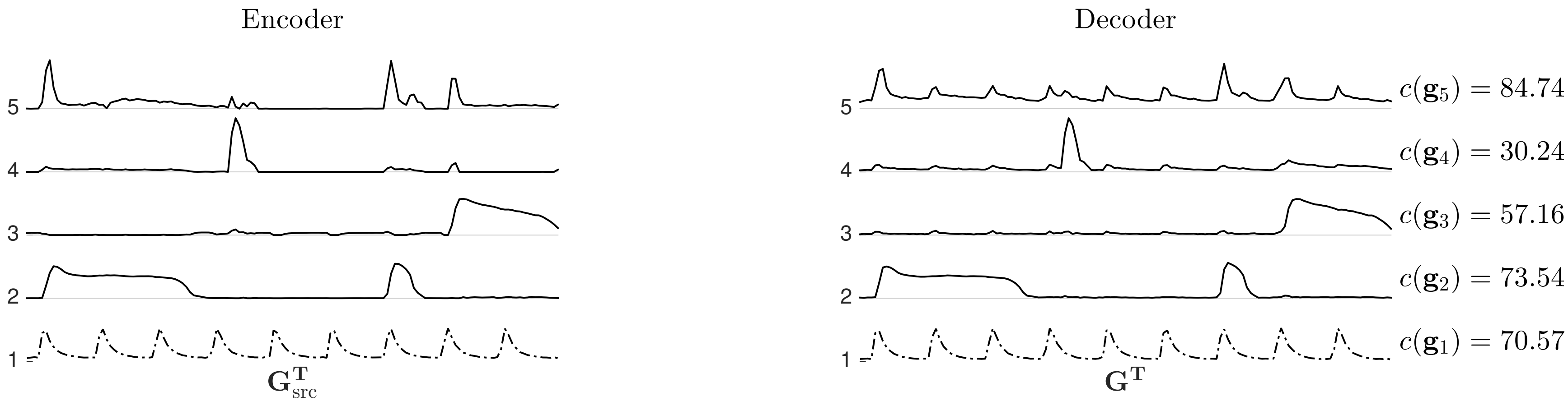
[Rohlfing, Becker, Wien 2016]

Content

Generalized Constraints for NMF

Cost function

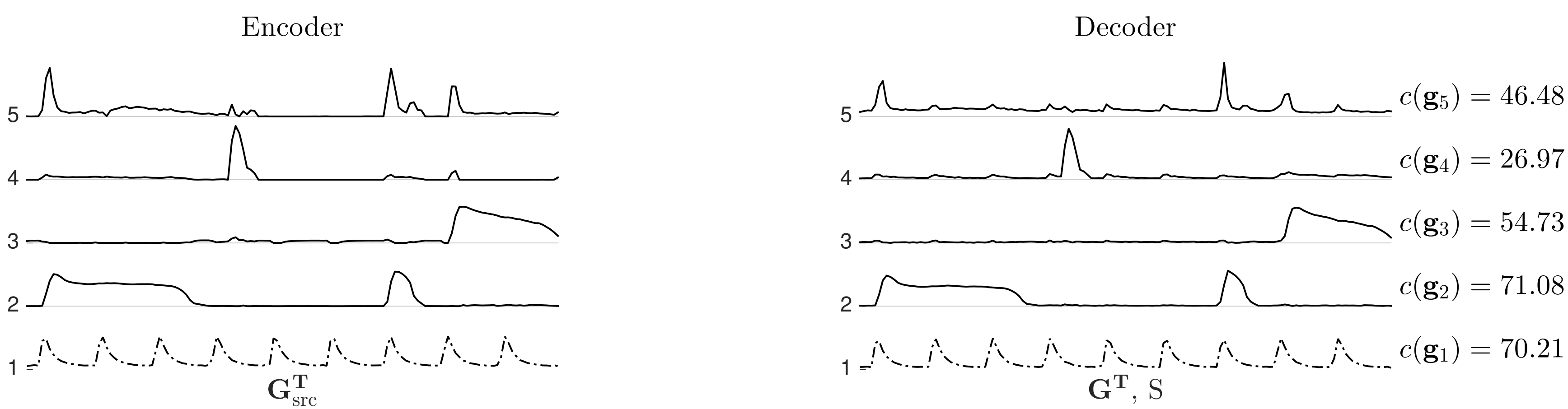
$$\min \quad d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^{\mathbf{T}})$$



Generalized Constraints for NMF

Cost function

$$\min \quad d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T}) + \alpha \underbrace{\sum_{i=1}^{I} c(\mathbf{g}_i)}_{=c(\mathbf{G})}$$

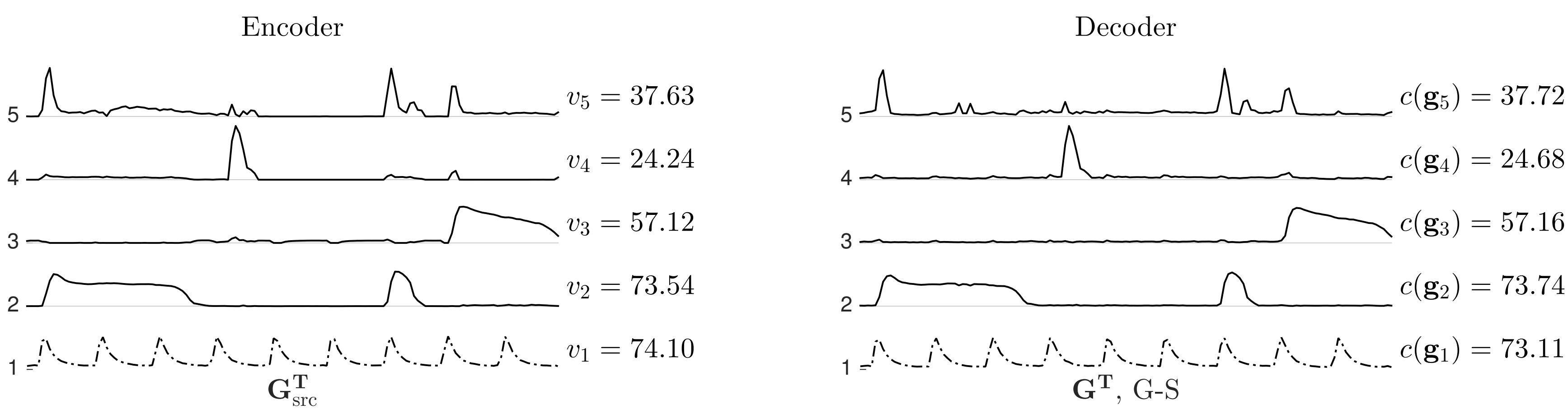


Generalized Constraints for NMF

Cost function

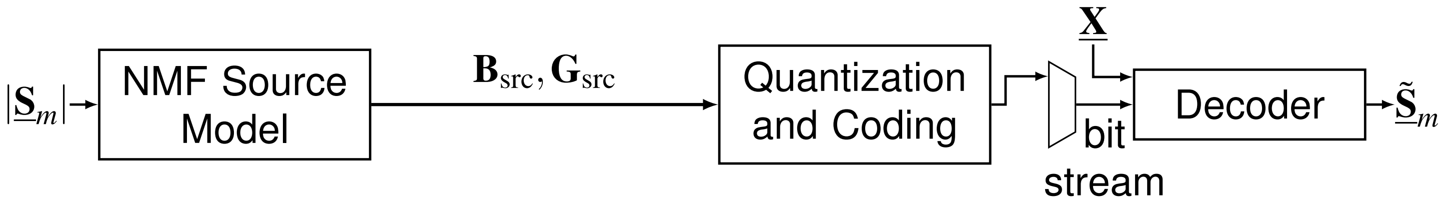
$$\min \quad d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T}) + \alpha \underbrace{\sum_{i=1}^{I} |c(\mathbf{g}_i) - \hat{v}_i|^p}_{=\bar{c}_p(\mathbf{G})}$$

here: $\quad v_i = c(\mathbf{g}_{\mathrm{src},i})$

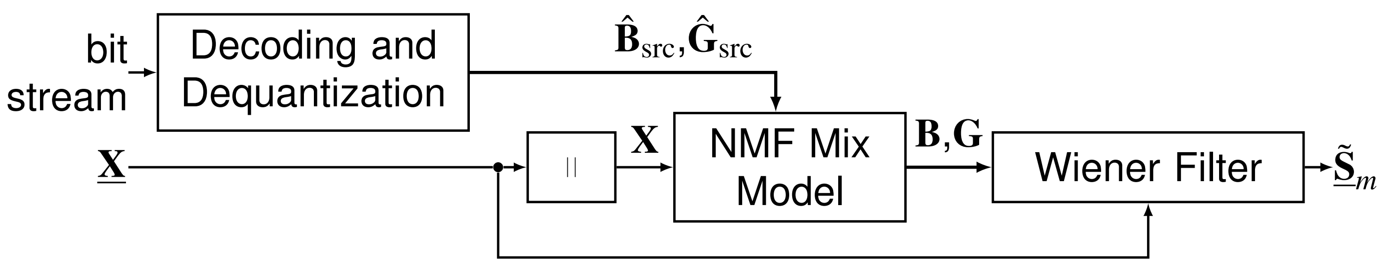


NMF-based ISS – revised

Encoder

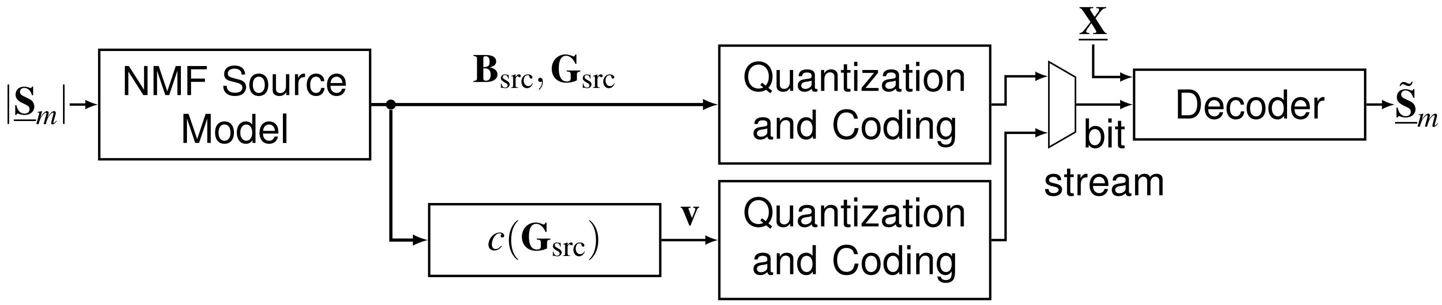


Decoder

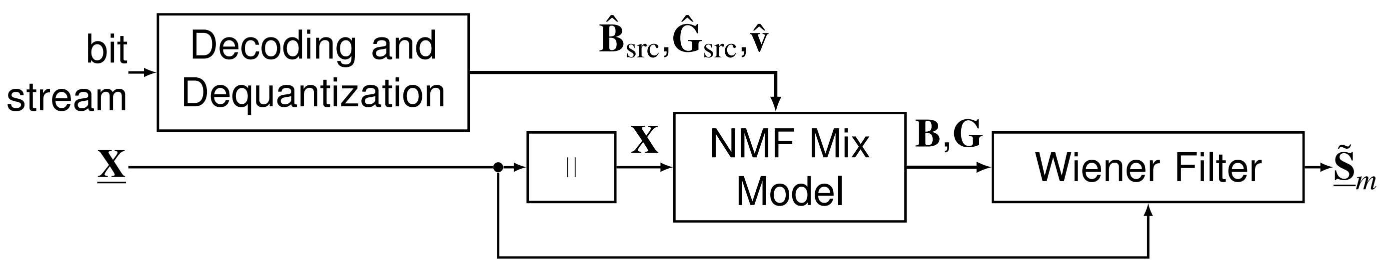


NMF-based ISS – revised

Encoder



Decoder



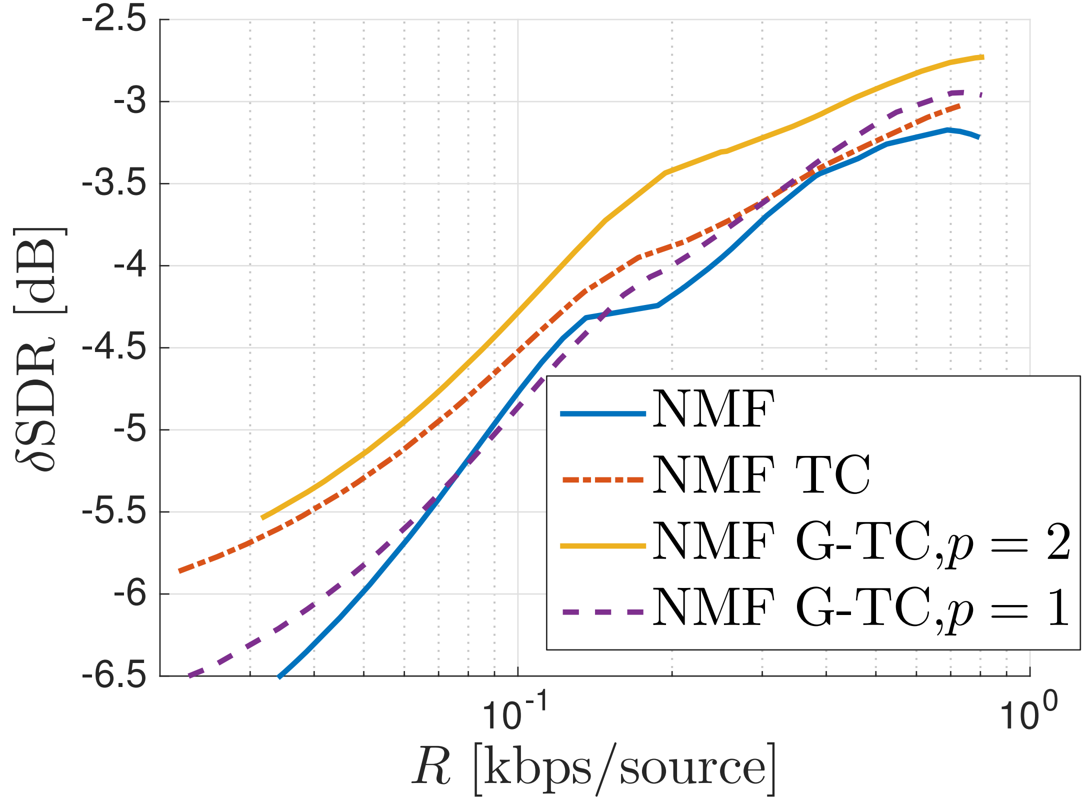
Content

Setup

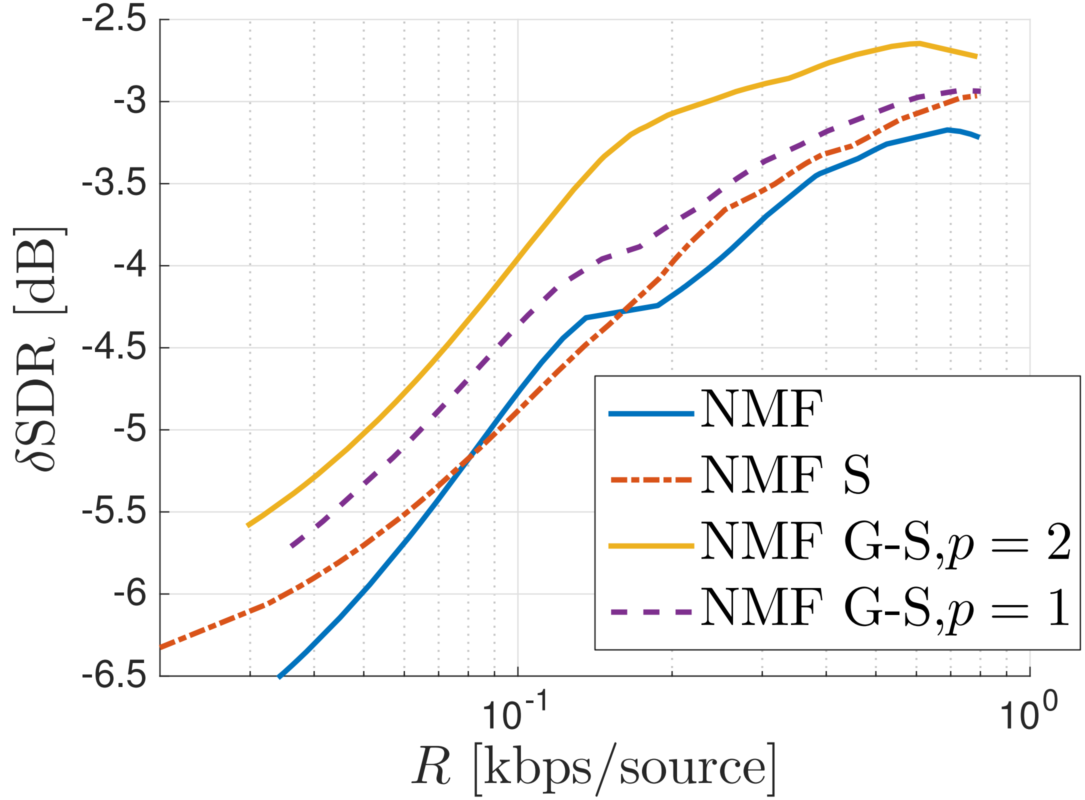
Dataset and metrics

- 5 monaural mixtures from QUASI database (3 to 6 sources), $20\,\mathrm{s}$ long
- Quality measure: signal-to-distortion ratio (SDR), in relation to oracle SDR (optimal TF-masking): δSDR. Parameter bit rate R in kbps / source.

Results Generalized Temporal Continuity



Results Generalized Temporal Sparseness



Content

Conclusion and Outlook

Conclusion

- Proposed NMF with generalized constraints
- Generalization due to convergence towards certain target values
- Application to informed source separation

Outlook

- Use generalized constraints for blind source separation
- Transmit target values per NMF iteration

Thank you for your attention

Any questions?

Update rules

Cost function

$$\min \quad d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T}) + c(\mathbf{B}) + c(\mathbf{G})$$

with
$$\nabla_\mathbf{B} d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T}) = \underbrace{\nabla_\mathbf{B}^+ d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T})}_{\geq 0} - \underbrace{\nabla_\mathbf{B}^- d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T})}_{\geq 0}$$

$$\nabla_\mathbf{G} d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T}) = \dots$$

$$\nabla_\mathbf{G} c(\mathbf{G}) = \dots$$

Update rules

$$b_{fi} \leftarrow b_{fi} \frac{\left[\nabla_\mathbf{B}^- d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T})\right]_{fi}}{\left[\nabla_\mathbf{B}^+ d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T})\right]_{fi}}$$

$$g_{ti} \leftarrow g_{ti} \frac{\left[\nabla_\mathbf{G}^- d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T})\right]_{ti} + \alpha \left[\nabla_\mathbf{G}^- c(\mathbf{G})\right]_{ti}}{\left[\nabla_\mathbf{G}^+ d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T})\right]_{ti} + \alpha \left[\nabla_\mathbf{G}^+ c(\mathbf{G})\right]_{ti}}$$

Update rules generalization

Cost function

$$\min \quad d_\beta(\mathbf{X}|\mathbf{B}\mathbf{G}^\mathbf{T}) + \alpha \underbrace{\sum_{i=1}^{I} |c(\mathbf{g}_i) - \hat{v}_i|^p}_{=\bar{c}_p(\mathbf{G})}$$

Update rules $p = 1$

$$\left[\nabla^+ \bar{c}_1(\mathbf{G})\right]_{ti} = \begin{cases} [\nabla^+ c(\mathbf{G})]_{ti} & \text{if } c(\mathbf{g}_i) > \hat{v}_i \\ [\nabla^- c(\mathbf{G})]_{ti} & \text{otherwise.} \end{cases}$$

$$\left[\nabla^- \bar{c}_1(\mathbf{G})\right]_{ti} = \begin{cases} [\nabla^- c(\mathbf{G})]_{ti} & \text{if } c(\mathbf{g}_i) > \hat{v}_i \\ [\nabla^+ c(\mathbf{G})]_{ti} & \text{otherwise.} \end{cases}$$

Update rules $p = 2$

$$\left[\nabla^+ \bar{c}_2(\mathbf{G})\right]_{ti} = 2\left\{c(\mathbf{g}_i)\left[\nabla^+ c(\mathbf{G})\right]_{ti} + \hat{v}_i\left[\nabla^- c(\mathbf{G})\right]_{ti}\right\}$$

$$\left[\nabla^- \bar{c}_2(\mathbf{G})\right]_{ti} = 2\left\{c(\mathbf{g}_i)\left[\nabla^- c(\mathbf{G})\right]_{ti} + \hat{v}_i\left[\nabla^+ c(\mathbf{G})\right]_{ti}\right\} .$$