Acoustic Beamforming with Collaborating Hearing Aids

Olivier Roy and Martin Vetterli

Motivations

- Hearing aids are acoustic sensing devices (see Figure 1) that aim at compensating various hearing impairments:
  - Spectral shaping by means of frequency attenuation/amplification.
  - Beamforming by combining coherently signals acquired at multiple microphones.
- Most state-of-the-art systems involve two devices working independently of one another:
  - Limited beamforming capability.
  - Poor rejection of interfering signals.
- What is the need for collaboration between two hearing aids?
  - Uses the spatial extent offered by the head to provide better beamforming capability.
  - Improves spatial noise reduction and increases speech intelligibility in noisy environments.

Fig. 1: Different types of hearing aids.

Problem Setup

- The auditory scene is composed of a desired source S, an interferer I and some ambient noise N [see Figure 2(a)].

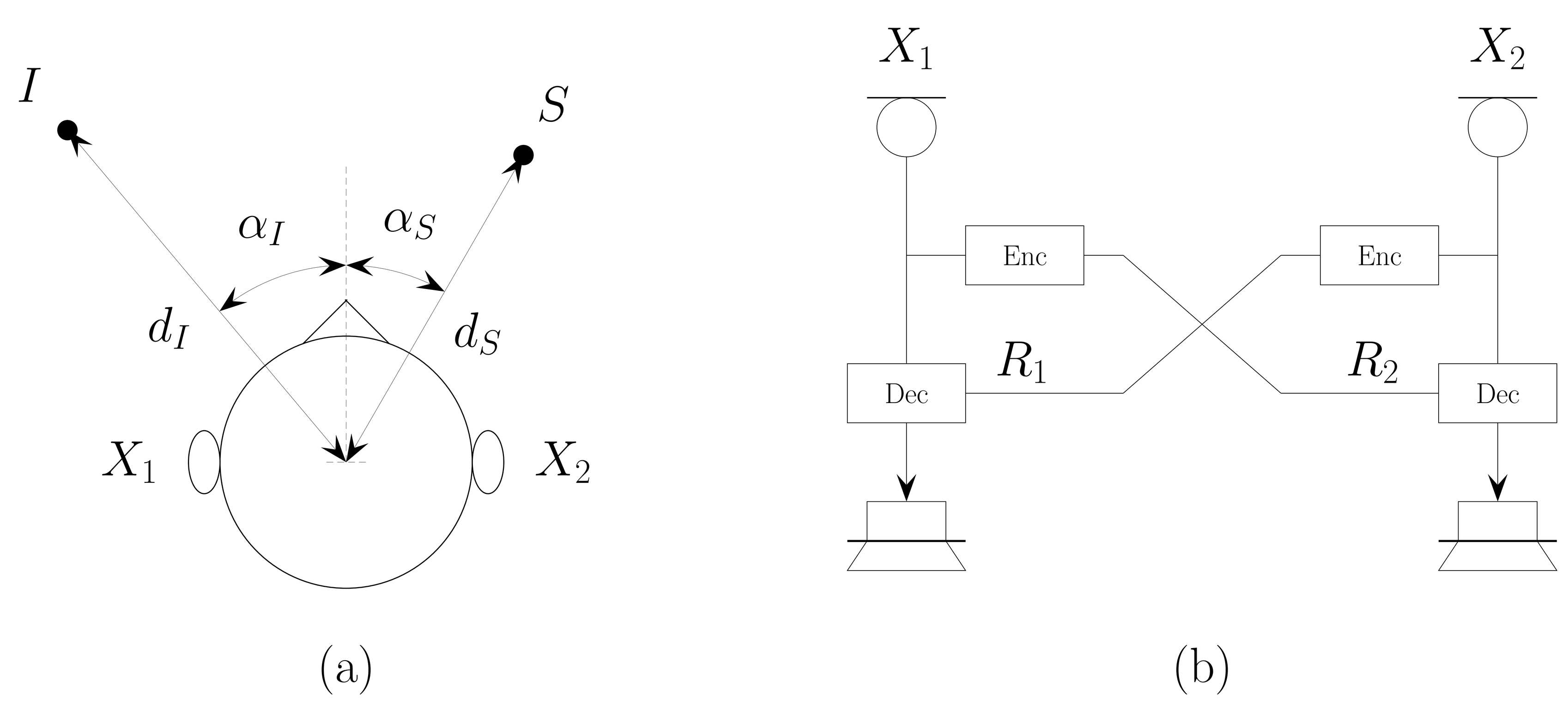

Fig. 2: Our hearing aids setup. (a) Typical head-related configuration. (b) Collaboration using a wireless communication link.

- The signal observed at hearing aid k ($k = 1, 2$) can be written as

$$\begin{aligned} X_k(t) &= S_k(t) + I_k(t) + N_k(t) \\ &= h_k(t) * S(t) + g_k(t) * I(t) + N_k(t)\,. \end{aligned}$$

- S, I and N_k are independent jointly Gaussian stationary random processes with mean zero and bandlimited power spectral density (PSD) Φ_S, Φ_I and Φ_{N_k}, respectively.
- The filter h_k (resp. g_k) denotes the *head-related impulse response* (HRIR) from the source (resp. the interferer) to hearing aid k. Their Fourier transform is referred to as *head-related transfer function* (HRTF).
- The two hearing aids are allowed to collaborate using a wireless communication link [see Figure 2(b)]:
  - The problem is symmetric. We look at it from the perspective of hearing aid 1.
  - Hearing aid 2 relays its acquired signal to hearing aid 1.
  - We want to reconstruct S_1 with minimum mean-squared error (MMSE).
- We define the *gain-rate* function

$$G(R) = \frac{D(0)}{D(R)}\,.$$

- Goal: to characterize the optimal gain-rate tradeoff provided by this collaboration.

Distributed Source Coding

- Our setup is identified as a source coding problem with side information at the decoder (indirect, noisy or remote Wyner-Ziv [1]).
- The optimal rate-distortion tradeoff for stationary random sources can be computed as [2]:

$$R(\theta) = \frac{1}{4\pi}\int_{-\infty}^{\infty} \max\left\{0, \log_2 \frac{\Phi_e(\Omega)}{\theta}\right\} d\Omega \quad \text{[b/s]}$$

$$D(\theta) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_{S_1|X_1,X_2}(\Omega)\, d\Omega + \frac{1}{2\pi}\int_{-\infty}^{\infty} \min\{\theta, \Phi_e(\Omega)\}\, d\Omega \quad \text{[MSE/s]}$$

where $\Phi_e = \Phi_{S_1|X_1} - \Phi_{S_1|X_1,X_2}$ and $\theta \in (0, \operatorname{ess\,sup}_\Omega \Phi_e(\Omega)]$. $\Phi_{X|Y}$ denotes the PSD of the error process $X - \mathrm{E}[X|Y]$.

- If the side information X_1 is disregarded in the encoding process, we obtain the following (suboptimal) rate-distortion tradeoff:

$$\tilde{R}(\theta) = \frac{1}{4\pi}\int_{-\infty}^{\infty} \max\left\{0, \log_2 \frac{\tilde{\Phi}_e(\Omega)}{\theta}\right\} d\Omega \quad \text{[b/s]}$$

$$\begin{aligned}\tilde{D}(\theta) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_{S_1|X_2}(\Omega)\, d\Omega + \frac{1}{2\pi}\int_{-\infty}^{\infty} \min\{\theta, \tilde{\Phi}_e(\Omega)\}\, d\Omega \\ &\quad - \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_{S_1}(\Omega) - \Phi_{S_1|V}(\Omega)\, d\Omega \quad \text{[MSE/s]}\end{aligned}$$

where $\tilde{\Phi}_e = \Phi_{S_1} - \Phi_{S_1|X_2}$, $V = X_1 - \mathrm{E}[X_1|U]$ and $\theta \in (0, \operatorname{ess\,sup}_\Omega \tilde{\Phi}_e(\Omega)]$. The process U corresponds to the signal received at the decoder and can be described by an optimal forward test channel for the remote source coding problem.

Rate-Constrained Beamforming Gain

- We model the head as a sphere for which an analytical expression of the HRTFs can be computed [3].
- An example of gain-rate function in the absence of interferer is given in Figure 3(a). We observe the rate-loss incurred by neglecting the presence of side information.
- The normalized distortion obtained for a particular rate and frequency as a function of the interferer's position is plotted in Figure 3(b).

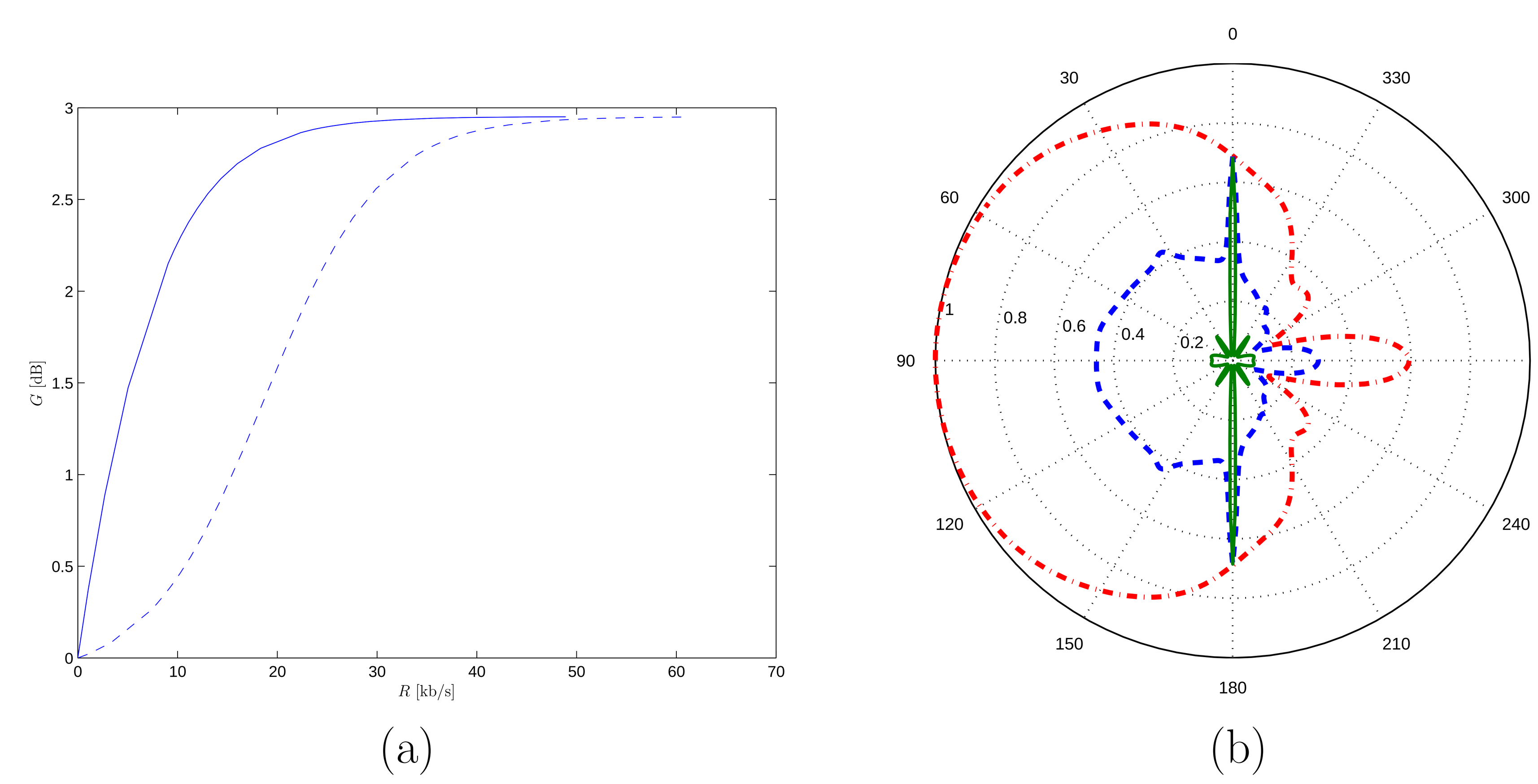

Fig. 3: Collaborative beamforming. (a) Gain-rate function with (plain) and without (dashed) Wyner-Ziv coding. (b) Normalized distortion for $f = \Omega/2\pi = 3000[Hz]$ and $R = 0$ (dash-dotted), $R = 0.1$ (dashed) and $R = 1$ (solid) [b/s/Hz].

References

[1] H. Yamamoto and K. Itoh, "Source coding theory for multiterminal communication systems with a remote source," *Trans. IECE Japan*, vol. E63, pp. 700–706, October 1980.

[2] O. Roy and M. Vetterli, "Rate-constrained beamforming for collaborating hearing aids," *accepted to IEEE International Symposium on Information Theory*, 2006.

[3] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," *Journal of the Acoustical Society of America*, vol. 5, pp. 3048–3058, November 1998.

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Audiovisual Communications Laboratory (LCAV), École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne

email: {olivier.roy,martin.vetterli}@epfl.ch web: http://lcavwww.epfl.ch/~{oroy,vetterli}