U-DREAM: Unsupervised Dereverberation guided by a Reverberation Model

Additional proofs

Louis Bahrman, Mathieu Fontaine, Gaël Richard

1 Training-less variant: Closed-form solution

Eq. (29), provides the formulation of the training-less variant of the dereverberation procedure. The proposed Monte-Carlo sampling solving scheme allows to deal with any reverberation model, in the case no closed-form solution of Eq. (29) exists. This is for instance the case with Polack's model and half-Gaussian noise used on synthetic data.

Yet, in the case where the reverberation model is based on Polack's model with Gaussian noise, detailed in Eq. (17), a solution can be found by considering a time-domain convolution operation. We detail this closed-form solution hereunder. Eq. (29) is:

$$\hat{\boldsymbol{S}} = \underset{\boldsymbol{S}}{\operatorname{argmin}}\, \mathbb{E}_{p(h|\mathcal{A}_{w_A}(\boldsymbol{Y}))} \left[\|\boldsymbol{Y} - \mathcal{C}(\boldsymbol{S}, h)\|_F^2 \right] \tag{29}$$

The time-frequency convolution operator $\mathcal{C}$ is equivalent for the time-domain linear convolution operator $\star$. Hence, Eq. (29) can be equivalently written in time-domain, as:

$$\hat{s} = \underset{s}{\operatorname{argmin}}\, \mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))} \left[\|y - s \star h\|_2^2 \right] \tag{P1}$$

This problem is a convex problem, as the expectation is a linear operation, for all h, $\left[\|y - s \star h\|_2^2 \right]$ is convex in s, (square norm of a linear operator). From Eq. (17) we can write $s \star h$ as $s + s \star h_r$, where h_r is a Gaussian vector. Note that Eq. (17) defines the covariance matrix of h_r to be diagonal. In this case:

$$\begin{aligned}
\hat{s} &= \underset{s}{\operatorname{argmin}}\, \mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))} \left[\|y - s - (s \star h_r)\|_2^2 \right] && \text{(P2)}\\
&= \underset{s}{\operatorname{argmin}}\, \mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))} \left[\|y - s\|_2^2 - 2(y-s)^T (s \star h_r) + \|s \star h_r\|_2^2 \right] && \text{(P3)}\\
&= \underset{s}{\operatorname{argmin}}\, \mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))} \left[\|y - s\|_2^2 \right] - 2(y-s)^T \left(s \star \mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))} \left[h_r \right] \right) + \mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))} \left[\|s \star h_r\|_2^2 \right], && \text{(P4)}
\end{aligned}$$

by linearity of the expectation. Neither the measured signal y, nor s depend on h_r (which is the sampled RIR and not the ground-truth RIR), and h_r is centered as per Eq. (17), the following holds:

$$\hat{s} = \underset{s}{\operatorname{argmin}}\, \|y - s\|_2^2 + \mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))} \left[\|s \star h_r\|_2^2 \right]. \tag{P5}$$

Let $\mathcal{L}$ the loss of this optimization problem, and H the Toeplitz convolution matrix associated with the convolution by h_r. We have:

$$\frac{\partial \mathcal{L}}{\partial s} = 2(s-y) + \mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))}\left[2H^T H s\right] \tag{P6}$$

$$= 2\left(I + \mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))}\left[H^T H\right]\right) s - 2y, \tag{P7}$$

by linearity of the expectation and the derivative of the square norm being integrable. $\mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))}\left[H^T H\right]$ is a diagonal matrix. Indeed, at row i and column j,

$$\mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))}\left[H^T H\right][i,j] = \sum_k \mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))}\left[h_r(k-i)h_r(k-j)\right], \tag{P8}$$

which is nonzero if and only if (iff) $i = j$, as h_r is a Gaussian vector with a diagonal covariance matrix. Then, denoting the diagonal matrix $D = \mathbb{E}_{p(h|\mathcal{A}_{w_A}(y))}\left[H^T H\right]$, in Eq. (P7), we have:

$$\frac{\partial L}{\partial s} = 0 \tag{P9}$$

$$\text{iff } (D+I)s = y. \tag{P10}$$

This means that in the Gaussian case, solving Eq. (P1) corresponds to scaling each sample of the reverberant signal y independently.

While this solution provides an intuition of the solution of Eq. (29) in a very simple case (Polack's model with a white noise distribution), it is not valid in the time-frequency domain with cross-band convolution. No closed-form solution could be found with a half-Normal distribution used in our proposed adaptation of Polack's model to simulate synthetic RIRs.