

6) Gibbs Sampling Approximation

We sample from $p(u_n | \mathbf{s}_{j \neq n}, \mathbf{u}_{j \neq n}, \mathbf{x}, \theta)$ and from $p(s_n | \mathbf{u}, s_{j \neq n}, \mathbf{x}, \theta)$:

$$s_n \sim \begin{cases} p_{mR}(s; \nu_n, \Psi_n^{-1}) & \text{if } u_n = 1 \\ \delta_0(s) & \text{if } u_n = 0. \end{cases}$$

7) Sampling from $p_{mR}(s; \nu_n, \Psi_n^{-1})$

$$\frac{1}{Z_{mR}} \left((s - \mu) e^{-\frac{(s-\mu)^2}{2\sigma_{mR}}} + \mu e^{-\frac{(s-\mu)^2}{2\sigma_{mR}}} \right)$$

With probability $p(s > \mu)$ we can sample from a mixture of a shifted Rayleigh distribution and a Gaussian. For $s < \mu$ we have a Gaussian upper bound on the distribution and use rejection sampling.

10) Conclusions

- We developed a linear generative signal model to describe polyphonic piano music.
- The introduced modified Rayleigh distribution is a conjugate prior for the Gaussian mean and closely models the note amplitudes.
- A Gibbs sampler can be implemented to sample from the posterior distribution of the note activation.
- Only based on the assumption of sparsity and the non-negative prior, the algorithm extracts note-like features.

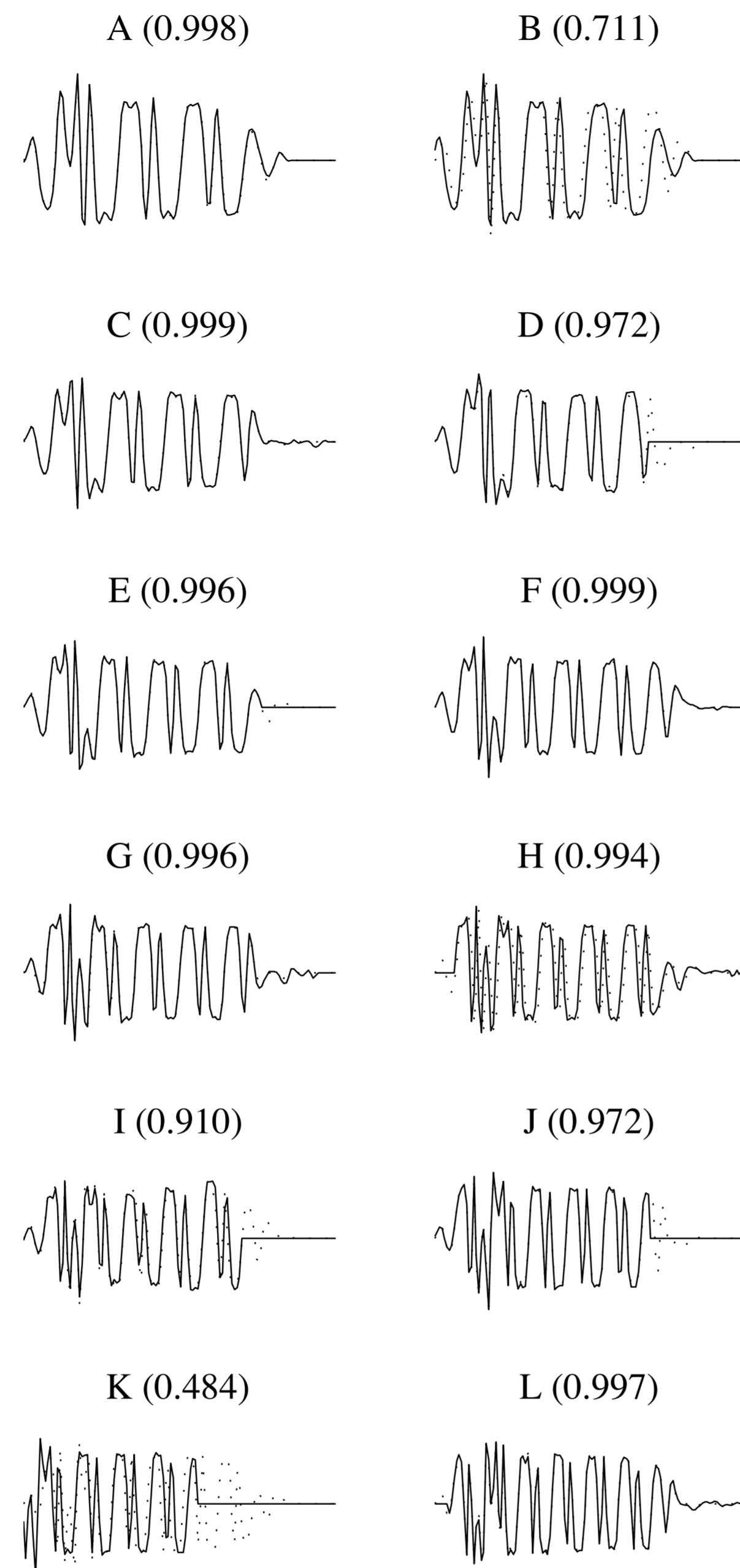


Fig. 2. Learned toy problem features.

8) Fig 2. Toy Problem

- A toy problem generated using the statistics of a real piano performance was analysed.
- Left. Original features (dotted lines) and learned features (solid lines).
- 10 out of 12 features have been learned while features B and K have not.
- The correlation between each original feature and closest learned feature are given in brackets.
- The model parameters converged to $\lambda_\epsilon = 172$, $\lambda_p = 1.05$, $\lambda_u = 13.15$ and $\mu = 0.45$.

9) Fig 3. Piano Notes

- A polyphonic piano recording of Ludwig van Beethoven's Bagatelle No. 1 Opus 33 was analysed.
- Of the 100 features, 47 converged. These are shown in the time-domain and the spectral-domain (right).
- The harmonic structure of the features is visible and the features are found to resemble individual notes.
- The found parameters were: $\lambda_\epsilon = 58$, $\lambda_p = 1.1$, $\lambda_u = 8.3$ and $\mu = -0.039$.

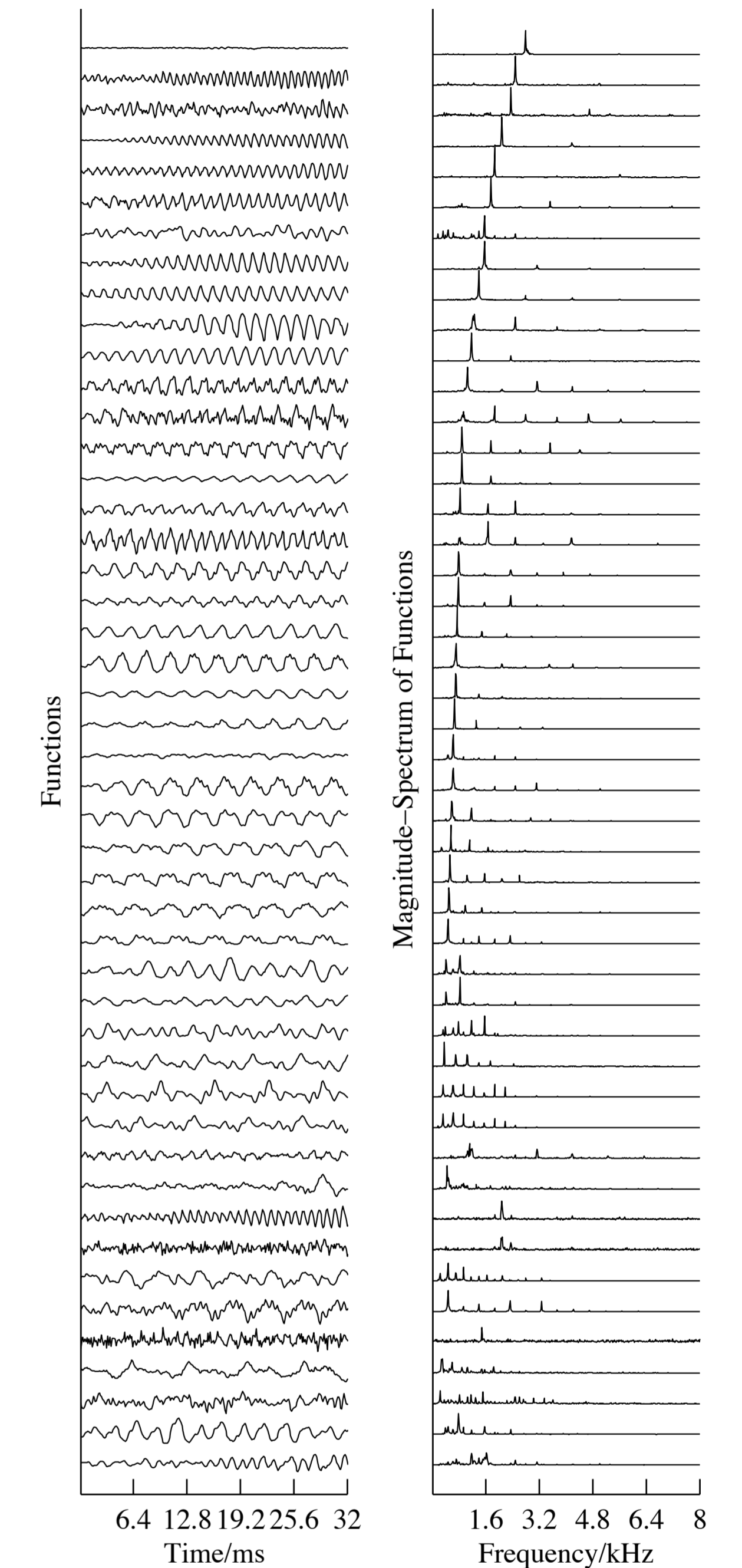


Fig. 3. Learned piano notes.

11) References

- [1] D. Wipf and B. Rao, "Sparse Bayesian learning for basis selection," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2153–2164, 2004.
- [2] D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 21, pp. 788–791, Oct 1999.
- [3] T. Blumensath and M. Davies, "Unsupervised learning of sparse and shift-invariant decompositions of polyphonic music," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2004.
- [4] P. Sallee and B. A. Olshausen, "Learning sparse multiscale image representations," in *Advances in Neural Information Processing Systems*. MIT Press, 2003, pp. 1327–1334.

centre for digital music

Department of Electronic Engineering

 Queen Mary
University of London

Mile End Road, London, E1 4NS, UK