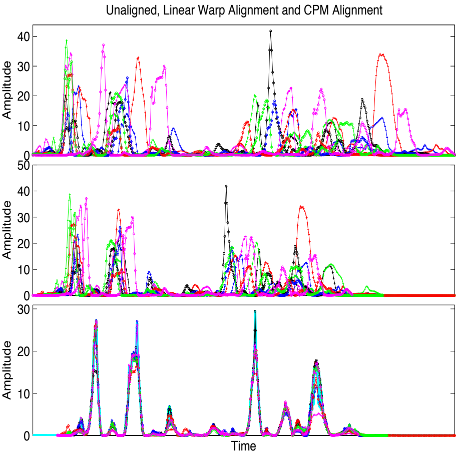
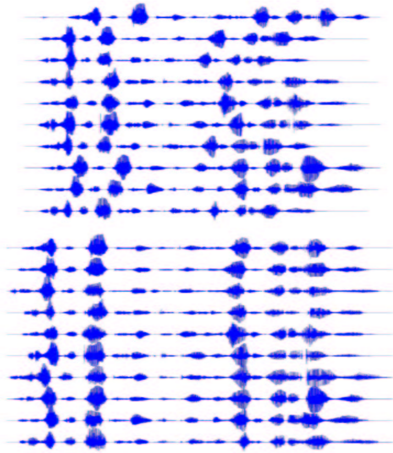
observed data. After training, the latent trace learned by the model represents a higher resolution 'fusion' of the experimental replicates. Figure 1 illustrate the model in action.



a)



b)

Figure 1: a) Top: ten replicated speech energy signals as described in Section 4), Middle: same signals, aligned using a linear warp with an offset, Bottom: aligned with CPM (the learned latent trace is also shown in cyan). b) Speech waveforms corresponding to energy signals in a), Top: unaligned originals, Bottom: aligned using CPM.

2 Defining the Continuous Profile Model (CPM)

The CPM is generative model for a set of K time series, $\vec{x}^k = (x_1^k, x_2^k, ..., x_{N^k}^k)$. The temporal sampling rate within each $\vec{x}^k$ need not be uniform, nor must it be the same across the different $\vec{x}^k$. Constraints on the variability of the sampling rate are discussed at the end of this section. For notational convenience, we henceforth assume $N^k = N$ for all k, but this is not a requirement of the model.

The CPM is set up as follows: We assume that there is a latent trace, $\vec{z} = (z_1, z_2, ..., z_M)$, a canonical representation of the set of noisy input replicate time series. Any given observed time series in the set is modeled as a non-uniformly subsampled version of the latent trace to which local scale transformations have been applied. Ideally, M would be infinite, or at least very large relative to N so that any experimental data could be mapped precisely to the correct underlying trace point. Aside from the computational impracticalities this would pose, great care to avoid overfitting would have to be taken. Thus in practice, we have used $M = (2 + \epsilon)N$ (double the resolution, plus some slack on each end) in our experiments and found this to be sufficient with $\epsilon < 0.2$. Because the resolution of the latent trace is higher than that of the observed time series, experimental time can be made effectively to speed up or slow down by advancing along the latent trace in larger or smaller jumps.

The subsampling and local scaling used during the generation of each observed time series are determined by a sequence of hidden state variables. Let the state sequence for observation k be $\vec{\pi}^k$. Each state in the state sequence maps to a time state/scale state pair: $\pi_i^k \rightarrow \{\tau_i^k, \phi_i^k\}$. Time states belong to the integer set $(1..M)$; scale states belong to an ordered set $(\phi_1..\phi_Q)$. (In our experiments we have used Q=7, evenly spaced scales in logarithmic space). States, π_i^k, and observation values, x_i^k, are related by the emission probability distribution: $A_{\pi_i^k}(x_i^k|\vec{z}) \equiv p(x_i^k|\pi_i^k, \vec{z}, \sigma, u^k) \equiv \mathcal{N}(x_i^k; z_{\tau_i^k}\phi_i^k u^k, \sigma)$, where σ