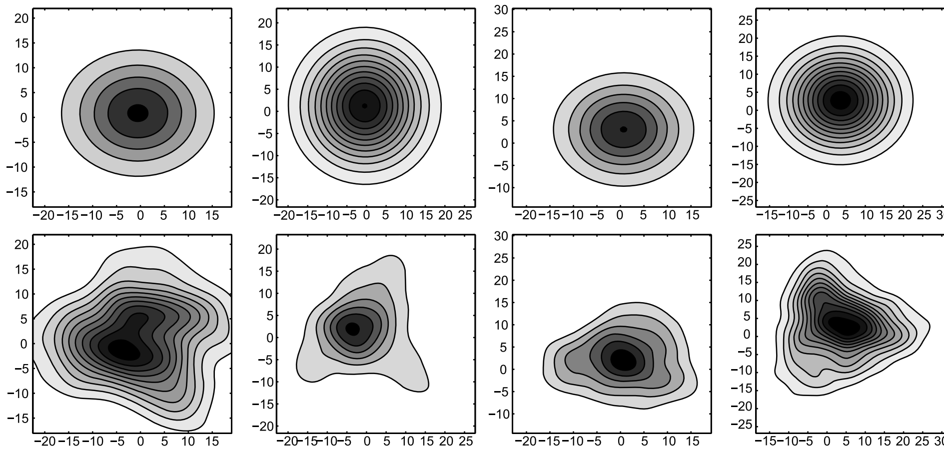

FIG. 3. *Contour plots of the best fit Gaussian (top) and kernel density estimate (bottom) for the top two principal components of the audio features associated with each of the four speakers present in the AMI_*20041210-1052 *meeting. Without capturing the non-Gaussianity of the speaker-specific emissions, the speakers are challenging to identify.*

that achieving state-of-the-art performance within our framework also relies on allowing for non-Gaussian emissions.

3. Dirichlet processes. A Dirichlet process (DP) is a distribution on probability measures on a measurable space Θ. This stochastic process is uniquely defined by a base measure H on Θ and a concentration parameter γ; we denote it by $\mathrm{DP}(\gamma, H)$. Consider a random probability measure $G_0 \sim \mathrm{DP}(\gamma, H)$. The DP is formally defined by the property that, for any finite partition $\{A_1, \dots, A_K\}$ of Θ,

$$(G_0(A_1), \dots, G_0(A_K)) | \gamma, H \sim \mathrm{Dir}(\gamma H(A_1), \dots, \gamma H(A_K)). \tag{3.1}$$

That is, the measure of a random probability distribution $G_0 \sim \mathrm{DP}(\gamma, H)$ on every finite partition of Θ follows a finite-dimensional Dirichlet *distribution* [Ferguson (1973)]. A more constructive definition of the DP was given by Sethuraman (1994). Consider a probability mass function (p.m.f.) $\{\beta_k\}_{k=1}^{\infty}$ on a countably infinite set, where the discrete probabilities are defined as follows:

$$\begin{aligned} v_k | \gamma &\sim \mathrm{Beta}(1, \gamma), \qquad k = 1, 2, \dots, \\ \beta_k &= v_k \prod_{\ell=1}^{k-1} (1 - v_\ell), \qquad k = 1, 2, \dots. \end{aligned} \tag{3.2}$$

In effect, we have divided a unit-length stick into lengths given by the weights β_k: the kth weight is a random proportion v_k of the remaining stick after the previous $(k-1)$ weights have been defined. This *stick-breaking construction* is generally