we simply illustrate that our improved inference procedure robustly explores the posterior, enabling this large-scale analysis and producing promising results.

10. Discussion. We have presented a Bayesian nonparametric framework for discovering dynamical behaviors common to multiple time series. Our formulation reposes on the beta process, which provides a prior distribution on overlapping subsets of binary features. This prior allows both for commonality and series-specific variability in the use of dynamic behaviors. We additionally developed an exact sampling algorithm for the BP-AR-HMM model, as well as novel split-merge moves and data-driven birth moves which efficiently explore the unbounded feature space. The utility of our BP-AR-HMM was demonstrated on the task of segmenting a large set of MoCap sequences. Although we focused on switching VAR processes, our approach (and sampling algorithms) could also be applied to other Markov switching processes, such as switching linear dynamical systems.

The idea proposed herein of a feature-based approach to relating multiple time series is not limited to nonparametric modeling. One could just as easily employ these ideas within a parametric model that prespecifies the number of possible dynamic behaviors. We emphasize, however, that conditioned on the infinite feature vectors of our BP-AR-HMM, which are guaranteed to be sparse, our model reduces to a collection of Markov switching processes on a *finite* state space. The beta process simply allows for flexibility in the overall number of globally shared behaviors, and computationally we do not rely on any truncations of this infinite model.

One area of future work is further improving the split-merge proposals. Despite the clear benefits of these proposals, we found sometimes that one "true" state would be split among several recovered features. The root of the splitting issue is twofold. One is the issue of mixing, which the annealing partially addresses, however, the fundamental issue of maintaining the reversibility of split-merge moves limits the acceptance rates due to the combinatorial number of configurations. The second is due to modeling issues. Our model assumes that the dynamic behavior parameters (i.e., VAR parameters) are identical between time series and do not change over time. This assumption can be problematic in grouping related dynamic behaviors and might be addressed via hierarchical models of behaviors or by ideas similar to those of the *dependent Dirchlet process* [Griffin and Steel (2006), MacEachern (1999)] that allows for time-varying parameters.

Overall, the MoCap results appeared to be fairly robust to examples of only slightly dissimilar behaviors, such as squatting to different levels or twisting at different rates. However, in cases such as the running motion where only portions of the body moved in the same way while others did not, the behaviors can be split (e.g., third jogging example in Figure 5). This observation could motivate *local partition processes* [Dunson (2009, 2010)] rather than *global partition processes*. That is, our current model assumes that the grouping of observations into behavior categories occurs along all components of the observation vector rather than just