See also [20] for a signal processing application of Dirichlet processes, specifically nonparametric modeling of excitations to a switching dynamical process. In this paper we make use of a variant of the HDP-HMM—the *sticky HDP-HMM* of [21]—to obtain improved control over the number of modes inferred; such control is crucial for the problems we examine. Our Bayesian nonparametric approach for learning switching dynamical processes extends the sticky HDP-HMM formulation to infer an unknown number of persistent dynamical modes and thereby capture a wider range of temporal dependencies. We then explore a method for learning which components of the underlying state vector contribute to the dynamics of each mode by employing *automatic relevance determination* (ARD) [22]–[24]. The resulting model allows for learning realizations of SLDS that switch between an unknown number of dynamical modes with possibly varying state dimensions, or switching VAR processes with varying autoregressive orders.

A. Previous System Identification Techniques

Paoletti *et al.* [25] provide a survey of recent approaches to identification of switching dynamical models. The most general formulation of the problem involves learning: 1) the number of dynamical modes, 2) the model order, and 3) the associated dynamic parameters. For noiseless switching VAR processes, Vidal *et al.* [26] present an exact algebraic approach, though relying on fixing a maximal mode space cardinality and autoregressive order. Psaradakis and Spagnolog [27] alternatively consider a penalized likelihood approach to identification of stochastic switching VAR processes.

For SLDS, identification is significantly more challenging and methods typically rely on simplifying assumptions such as deterministic dynamics or knowledge of the mode space. Huang *et al.* [28] present an approach that assumes deterministic dynamics and embeds the input/output data in a higher-dimensional space and finds the switching times by segmenting the data into distinct subspaces [29]. Kotsalis *et al.* [30] develop a balanced truncation algorithm for SLDS assuming the mode switches are independent and identically distributed (i.i.d.) within a fixed, finite set; the authors also present a method for model-order reduction of HMMs (see also [31]). In [32], a realization theory is presented for *generalized jump-Markov linear systems* in which the dynamic matrix depends both on the previous mode and current mode. Finally, when the number of dynamical modes is assumed known, Ghahramani and Hinton [33] present a variational approach to segmenting the data from a *mixture of experts* SLDS into the linear dynamical regimes and learning the associated dynamic parameters. For questions on observability and identifiability of SLDS in the absence of noise, see [34].

In the Bayesian approach that we adopt, we coherently incorporate noisy dynamics and uncertainty in the mode space cardinality. Our choice of prior penalizes more complicated models, both in terms of the number of modes and the state dimension describing each mode, allowing us to distinguish between the set of equivalent models described in [34]. Thus, instead of placing hard constraints on the model, we simply increase the posterior probability of simpler explanations of the data. As opposed to a penalized likelihood approach using *Akaike's information criterion* (AIC) [35] or the *Bayesian information criterion* (BIC) [36], our approach provides a model complexity penalty in a purely Bayesian manner.

In Section II, we provide background on the switching linear dynamical systems we consider herein and previous Bayesian nonparametric methods of learning HMMs. Our Bayesian nonparametric switching linear dynamical systems are described in Section III. We proceed by analyzing a conjugate prior on the dynamic parameters and a sparsity-inducing prior that allows for variable-order switching processes. The section concludes by outlining a Gibbs sampler for the proposed models. In Section IV we present results on synthetic and real datasets and in Section V we analyze a set of alternative formulations that are commonly found in the maneuvering target tracking and econometrics literature.

II. Background

A. Switching Linear Dynamic Systems

A state-space (SS) model consists of an underlying state, $\boldsymbol{x}_t \in \mathbb{R}^n$, with dynamics observed via $\boldsymbol{y}_t \in \mathbb{R}^d$. A linear time-invariant (LTI) SS model is given by

$$\boldsymbol{x}_t = A\boldsymbol{x}_{t-1} + \boldsymbol{e}_t \quad \boldsymbol{y}_t = C\boldsymbol{x}_t + \boldsymbol{w}_t \tag{1}$$

where $\boldsymbol{e}_t$ and $\boldsymbol{w}_t$ are independent Gaussian noise processes with covariances Σ and R, respectively.

An order r VAR process, denoted by $\mathrm{VAR}(r)$, with observations $\boldsymbol{y}_t \in \mathbb{R}^d$, can be defined as

$$\boldsymbol{y}_t = \sum_{i=1}^{r} A_i \boldsymbol{y}_{t-i} + \boldsymbol{e}_t \quad \boldsymbol{e}_t \sim \mathcal{N}(0, \Sigma). \tag{2}$$

Every $\mathrm{VAR}(r)$ process can be described in SS form, though the converse is not true for finite r [37].

The dynamical phenomena we examine in this paper exhibit behaviors better modeled as switches between a set of linear dynamical models. We define a *switching linear dynamical system* (SLDS) by

$$\begin{aligned} z_t \mid z_{t-1} &\sim \pi_{z_{t-1}} \\ \boldsymbol{x}_t &= A^{(z_t)}\boldsymbol{x}_{t-1} + \boldsymbol{e}_t^{(z_t)} \\ \boldsymbol{y}_t &= C\boldsymbol{x}_t + \boldsymbol{w}_t. \end{aligned} \tag{3}$$

The first-order Markov process z_t with transition distributions $\{\pi_j\}$ indexes the mode-specific LDS at time t, which is driven by Gaussian noise $\boldsymbol{e}_t^{(z_t)} \sim \mathcal{N}(0, \Sigma^{(z_t)})$. One can view the SLDS as an extension of the classical hidden Markov model (HMM) [11], which has the same mode evolution, but conditionally *independent* observations:

$$\begin{aligned} z_t \mid z_{t-1} &\sim \pi_{z_{t-1}} \\ y_t \mid z_t &\sim F(\theta_{z_t}) \end{aligned} \tag{4}$$

for an indexed family of distributions $F(\cdot)$ where θ_i are the *emission parameters* for mode i.