# Hierarchical Modeling of Systems with Similar Components

# Milad Memarzadeh

Doctoral Candidate, Dept. of Civil and Env. Eng., Carnegie Mellon University, Pittsburgh, USA

# Matteo Pozzi

Assistant Professor, Dept. of Civil and Env., Carnegie Mellon University, Pittsburgh, USA

# J. Zico Kolter

Assistant Professor, Dept. of Computer Science, Carnegie Mellon University, Pittsburgh, USA

ABSTRACT: Identifying optimal management policies for systems made up by similar components is a challenging task, due to dependence in the components' behavior. In this setting, observations collected on one component are also relevant for learning the behavior of others. Probabilistic graphical models allow for consistent inference using all available data, taking dependence among components into account, while optimizing system operation. In this paper we propose a framework for management of systems made by similar components based on hierarchical Bayesian modeling, called Multiple Uncertain Partially Observable Markov Decision Processes (MU-POMDP), that overcomes some limitations of a previously proposed approaches. We describe a detailed numerical algorithm to learn the system parameters within this framework and we investigate its performance with an example of management of a wind farm (i.e., the system) made up by turbines of the same type (i.e., the components).

# 1. INTRODUCTION

Optimal policies for operation and maintenance of systems can be identified under the assumption of independent components, or of components governed by the same model. However, it is challenging to represent systems made by different but similar components. In this case, component parameters should be treated as a set of interdependent random variables, and inference performed so that observations collected on any component is also relevant, to a certain degree, in the management of others. Many infrastructure systems can be modeled using this framework.

Markov Decision Process (MDP) (Bertsekas 1996, Sutton and Barto 1998) and Partially Observable MDP (POMDP) (Smallwood and Sondik 1973, Sondik 1978) are classic frameworks for sequential operation. Specifically, POMDP assumes that evolution of the component state follows the Markov property, but allows for partial and/or imperfect observation of that state.

One limitation of these frameworks is that the stochastic modeling of the state evolution (i.e. the "transition probabilities") and, in case of POMDP, of the observations (i.e. the "emission probabilities") is not affected by uncertainty nor it is updated by data processing. However, in many applications, these models are uncertain, and different for each individual component. Ross et al. (2011) have proposed the Bayes-Adaptive POMDP (BA-POMDP) framework that allows to treat models as random variables and to update their distribution as data are analyzed during the management process. However, exact algorithms in this framework require high computational complexity and become easily intractable for systems with a high number components and states. Approximate methods for optimal management under model uncertainty have been proposed by Jaulmes et al. (2005a,b)recently, ourselves and, by

(Memarzadeh et al. 2013,2014a). The latter method, called PLUS (Planning and Learning in Uncertain dynamic Systems), is structured in two phases: learning and planning. In the learning phase, it makes use of Markov Chain Monte Carlo (MCMC) Gibbs sampling (Carter and Kohn 1994). The planning phase is based on an approximation which neglects the exploratory value of future learning of the model, to make the algorithm efficient and tractable. Details about PLUS can be found in Memarzadeh et al. (2014a). When applied to the management of a set of components, the learning approach of PLUS can be implemented in two alternative ways: independent models can be tuned for each component, or a single model can be updated by processing data from all components. While these two implementations are suitable for some applications, they are not for systems made by components that we assume to be similar, but not identical. Following a preliminary publication on systems with similar components (Memarzadeh et al. 2014b), in this paper we generalize PLUS to a framework we called Multiple Uncertain POMDP (MU-POMDP) that is based on hierarchical Bayesian modeling. This framework allows assuming specific levels of similarities among components, and consistently processing observations collected at system level.

The motivation for this research is to improve operation and maintenance of wind farms. MDP and POMDP have been applied to management in this field (Byon et al. 2010, Byon and Ding 2010, Nielsen and Sorensen 2012). Turbines can be understood as components, and their deterioration modeled as a stochastic depending on the actions taken. process, Epistemic uncertainties on deterioration of different components can be assumed as dependent, as turbines are instances of the same structural model. On the other hand, differences on component construction and position lead to slight or significant dissimilarities among deterioration processes. Similar considerations can be made on the monitoring systems providing observations about the condition state.

For these reasons, MU-POMDP can be a suitable framework for this application.

The following parts of this paper present the MU-POMDP framework (2), and compare its performance with that of PLUS on a numerical example (3) before drawing conclusions (4).

### 2. THE MU-POMDP SCHEME

The MU-POMDP scheme targets the management of a set of components, each defined by a POMDP whose transition and emission probabilities are modeled as dependent random variables defined by hyper-parameters.

Figure 1 shows the graphical model of the MU-POMDP, where circles define random variables, squares decision variables, diamonds utility variables, dots fixed parameters, and arrows define dependence among variables. In that graph, *S* indicates the component state, *A* the maintenance action, *Z* the available observation and *R* the monetary reward (or loss, if value is negative). Subscript "k, t" refers the variable to component *k* at time *t*. Variables *Z* are shaded to indicate that they can be directly observed.

We consider a *K*-component system. For each k = 1, ..., K, model parameter  $\mathbf{T}_k$  indicates the transition probability of component *k*, and model parameter  $\mathbf{O}_k$  indicates the corresponding emission probability.



Figure 1. MU-POMDP scheme for a two-component system represented as a probabilistic graphical model.

Additional lavers model prior the distributions: hyper-parameters are marked as  $\alpha_T$ ,  $\beta_T$ ,  $\alpha_0$  and  $\beta_0$  in Figure 1: the first two define the dependence in the transitions, while the latter ones that of emissions. While model parameters are different for each component, hyper-parameters are common to the entire system. Parameter  $\mathbf{\eta}_T$  and  $\mathbf{\eta}_O$ ,  $\lambda_T$ , and  $\lambda_O$  define the distribution of those hyper-parameters. As apparent in Figure 1, the model parameters (e.g.  $\mathbf{T}_1$  and  $\mathbf{T}_2$ ) of different components are not marginally independent, because of the common hyper-parameter parents. Consequently, observations on any component, by affecting the knowledge of the hyper-parameters, affect in turn model parameters of other components as well. The overall management task is to select actions to minimize the expected sum of the discounted losses (Bertsekas 1996). As in the PLUS approach, MU-POMDP is composed by two phases: learning and planning. The planning phase is the same as PLUS (Memarzadeh et al. 2014a), as the samples produced by the learning phase are fully compatible with PLUS planning, without the need for any adjustment.

The learning phase represents the posterior distribution of all variables, conditional to all observations Z and actions A observed up to present time. In principle, once each distribution in the probabilistic graphical model is analytically defined, posterior distributions could be computed theoretically. However, exact learning is generally not feasible for the graph presented in Figure 1, and approximate methods need to be adopted.

# 2.1. MCMC updating scheme

Extending the approach used in PLUS, learning in MU-POMDP is based on Gibbs sampling, which is an effective implementation of MCMC. Figure 2 reports a scheme of the inference process. In the figure, the upper bar indicates a collection of variables, from the beginning of the management process up to a specific time. For example,  $\bar{S}_{k,t}$  indicates the state trajectory  $\{S_{k,1}, \dots, S_{k,t}\}$  for component *k*. The superscript (*j*) refers to the *j*-th samples generated by the

MCMC algorithm. At component level, the sampling of states and model parameters is identical to that adopted by PLUS. At system level. the hyper-parameters are generated conditional to the sampled model parameters for all components. This task can be accomplished, if needed, by using the Metropolis-Hastings (M-H) approach (MacKay 2003). In summary, Figure 2 can be read as a recipe for generating samples from the joint posterior distribution: after initialization, state trajectories and model parameters are sampled for each component, then hyper-parameters are sampled at system level, and these steps are alternated. The Burn-in phase is discarded and the remaining part of the random walk is kept.



Figure 2. The proposed MCMC sampling approach

# 2.2. Probabilistic models and inference

The graphical model in Figure 1 requires a specific assignment of marginal and conditional distributions for every random variable. A feasible set of distributions for MU-POMDP, inspired by Kemp et al. (2007), is defined as follows:

$$\begin{aligned} \alpha_{T} &\sim \text{Exponential}(\lambda_{T}), & \alpha_{O} \sim \text{Exponential}(\lambda_{O}) \\ \boldsymbol{\beta}_{T} &\sim \text{Dirichlet}(\boldsymbol{\eta}_{T}), & \boldsymbol{\beta}_{O} \sim \text{Dirichlet}(\boldsymbol{\eta}_{O}) \\ \boldsymbol{T}_{k} &\sim \text{Dirichlet}(\alpha_{T}\boldsymbol{\beta}_{T}), & \boldsymbol{O}_{k} \sim \text{Dirichlet}(\alpha_{O}\boldsymbol{\beta}_{O}) \\ \boldsymbol{T}_{k} \perp \boldsymbol{T}_{l} \mid \alpha_{T}, \boldsymbol{\beta}_{T}, & \boldsymbol{O}_{k} \perp \boldsymbol{O}_{l} \mid \alpha_{O}, \boldsymbol{\beta}_{O} \\ \boldsymbol{S}_{k,t} \mid \boldsymbol{S}_{k,t-1}, \boldsymbol{A}_{k,t} \sim \text{Multinomial}(\boldsymbol{T}_{k}) \\ \boldsymbol{Z}_{k,t} \mid \boldsymbol{S}_{k,t}, \boldsymbol{A}_{k,t} \sim \text{Multinomial}(\boldsymbol{O}_{k}) \end{aligned}$$
(1)

As in the PLUS approach, state trajectories are sampled using forward filtering backward

sampling (FFBS) (Fruhwirth-Schnatter 2006). The Dirichlet distribution on model parameters is appropriate in this context, because it is conjugate prior for the multinomial distribution, and this facilitates the implementation of the Gibbs approach. Because of this, samples of model parameters can be easily generated, as in the PLUS algorithm. It is worth clarifying the role of hyper-parameters  $\alpha$  and  $\beta$  in the definition of the prior distribution of model parameters. Matrix  $\beta$  defines the expected value of the corresponding model parameters, while scalar variable  $\alpha$  affects the uncertainty of model parameters: a high value of  $\alpha$  induces a low variance.

input: $\lambda_T, \mathbf{\eta}_T, \mathbf{T}, \sigma_{\alpha}, c_{\beta}, J$
initialize $\alpha_T^{(1)}, \boldsymbol{\beta}_T^{(1)}$
for $j = 1,, J$ do
sample $r \sim Normal(0,1)$
$lpha_T' = \exp\left[\log\left(lpha_T^{(j)} ight) + r  imes \sigma_lpha ight]$
sample $oldsymbol{eta}_T'$ ~Dirichlet $\left(c_oldsymbol{eta}_T^{(j)} ight)$
$p-ratio = P(\mathbf{T}, \alpha_T', \boldsymbol{\beta}_T', \lambda_T, \boldsymbol{\eta}_T) / P\left(\mathbf{T}, \alpha_T^{(j)}, \boldsymbol{\beta}_T^{(j)}, \lambda_T, \boldsymbol{\eta}_T\right)$
$q\text{-ratio} = \frac{\text{Dirichlet}(\boldsymbol{\beta}_T': c_{\boldsymbol{\beta}} \boldsymbol{\beta}_T^{(j)})}{\text{Dirichlet}(\boldsymbol{\beta}_T^{(j)}: c_{\boldsymbol{\beta}} \boldsymbol{\beta}_T')} \times \frac{\alpha_T'}{\alpha_T^{(j)}}$
$accept = p-ratio \times q-ratio$
sample $r' \sim \text{Uniform}(0,1)$
if accept $> r'$
$lpha_T^{(j+1)}=lpha_T'$ , $oldsymbol{eta}_T^{(j+1)}=oldsymbol{eta}_T'$
else
$lpha_T^{(j+1)} = lpha_T^{(j)}, \ oldsymbol{eta}_T^{(j+1)} = oldsymbol{eta}_T^{(j)}$
end
end
output: hyper-parameters $\alpha_T^{(J+1)}$ , $\beta_T^{(J+1)}$

*Figure 3. M-H algorithm for sampling hyperparameters on transition probabilities.* 

Samples of hyper-parameters  $\alpha_T$  and  $\beta_T$  can be generated following the recipe reported in Figure 3, based on the M-H algorithm (the corresponding procedure for  $\alpha_0$  and  $\beta_0$  being identical, with obvious changes in the input variables). Input variables are the parameters defining the prior distribution ( $\lambda_T$  and  $\eta_T$ ), the transition probabilities for all components (**T**) which, following Gibbs, are obtained by sampling. Inputs  $\sigma_{\alpha}$  and  $c_{\beta}$  control the step size of the M-H proposal distribution in the direction of  $\alpha_T$  and  $\beta_T$  respectively, while *J* is the length of the Markov Chain. In Figure 3, Dirichlet(**x**; **y**) is the value assumed by the Dirichlet distribution with parameters **y** at **x**. *P* indicates the un-normalized joint distribution of hyper-parameters and model parameters that, following Eq. (1), reads:

$$P(\mathbf{T}, \alpha_T, \boldsymbol{\beta}_T, \lambda_T, \boldsymbol{\eta}_T) = \lambda_T \exp(-\lambda_T \alpha_T) \\ \times \text{Dirichlet}(\boldsymbol{\beta}_T; \boldsymbol{\eta}_T) \\ \times \prod_{k=1}^{K} \text{Dirichlet}(\mathbf{T}_k; \alpha_T, \boldsymbol{\beta}_T)$$
(2)

At any state during the management process, this procedure provides samples of model parameters, hyper-parameters and state, to be used in the planning phase.

# 3. NUMERICAL VALIDATION

# 3.1. PLUS as alternative processing approach

We compare MU-POMDP performance with PLUS on a numerical example similar to that presented in Memarzadeh et al. (2014a). Figure 4 shows the graphical models for PLUS algorithm used for comparison, which models transition and emission for all components as identical.

Actually, it should be noted that, for practical applications of PLUS, the additional layer of hyper-parameters is unnecessary and the prior of model parameters can be defined directly. However, in order to achieve a fair comparison between MU-POMDP and PLUS, we make use of the same arrangement of layers and values for  $\lambda$ ,  $\eta$ , and for conditional distributions of hyper-parameters and model parameters, to get the same marginal distribution for model parameters across the two approaches.

#### 3.2. Parameters for numerical investigation

For the purpose of validation, we consider a wind farm made up by 5 turbines of the same type placed in similar environmental conditions. The state condition of each turbine is discretized into three possible states where s = 1 refers to an intact structure, s = 2 to a damaged one, and s = 3 to the failure of the turbine; the agent

receives observations from a set of four possible observations where z = 1 suggests that the turbine is undamaged, z = 2 and z = 3 indicate two symptoms of damage, and z = 4 indicates the failure of the turbine; three actions are available: Do-Nothing (DN), Repair (RE), and Visual Inspection (VI). When the agent chooses DN, the condition state of the turbine degrades owing to fatigue and aging, potentially causing a structural failure and a relevant economical loss. In turn, the agent can perform a costly intervention (i.e., RE) to avoid failure and improve the condition state of the turbines. VI better measures the condition state of the turbine (that evolves according to the degradation model, as for DN). Each time step is assumed to be six months, and the agent takes one action per turbine at each time step.



Figure 4. Probabilistic graphical models of PLUS

Parameters of priors for hyper-parameters are fixed as follow:

$$\lambda_T = \lambda_0 = \frac{1}{1000}$$
$$\boldsymbol{\eta}_{T,\text{DN,VI}} = \kappa \times \begin{bmatrix} 0.57 & 0.28 & 0.15 \\ 0 & 0.67 & 0.33 \\ 0 & 0 & 1 \end{bmatrix}$$
$$\boldsymbol{\eta}_{T,\text{RE}} = \kappa \times \begin{bmatrix} 0.67 & 0.33 & 0 \\ 0.67 & 0.33 & 0 \\ 0.67 & 0.33 & 0 \end{bmatrix}$$

$$\boldsymbol{\eta}_{o,\mathrm{DN,RE}} = \kappa \times \begin{bmatrix} 0.57 & 0.28 & 0.15 & 0\\ 0.15 & 0.57 & 0.28 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}$$
$$\boldsymbol{\eta}_{o,\mathrm{VI}} = \kappa \times \begin{bmatrix} 0.67 & 0.33 & 0 & 0\\ 0.33 & 0 & 0.67 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where subscripts report the action symbol,  $\kappa$ controls the skewness of the prior and has been fixed to 50, so that the corresponding coefficient of variation of the samples is about 0.26. Parameter  $\lambda$  controls the correlation among the model parameters across components: as  $\lambda$ decreases, the correlation increases, and it is about 75% given the values reported above. Entries in square brackets define the expected value of transition and emission probabilities: for example, the expected value of the probability that the undamaged turbine becomes damaged under DN is 28%. The costs for repair, visual inspections and down-time due to failure are assumed to be US \$10,000, \$500, and \$50,000, respectively. The discount factor is assumed to be  $\gamma = 0.95$ . The initial belief state for all turbines is defined as  $b_0 = \begin{bmatrix} 0.8 & 0.2 & 0 \end{bmatrix}$ , which means that the agent believes that, at the beginning of the process, the turbines are in the intact state with 80% probability and in damaged state with 20% probability.

#### 3.3. Scheme for numerical validation

To investigate the performance of MU-POMDP and compare it with PLUS, we simulate the response of a system characterized by model  $\Theta^* = \{\Theta_1^*, \Theta_2^*, ..., \Theta_k^*\}$ , where  $\Theta_k^* = \{\mathbf{T}_k^*, \mathbf{O}_k^*\}$ defines transition and emission probabilities for component *k*. Comparison is performed in terms of effectiveness of learning and planning.

For learning, we evaluate the accuracy in inferring condition state (i.e. the accuracy of beliefs). At time step *t*, the probability distribution of states for the entire farm is defined as  $P(\mathbf{s}_{t+1}|\mathbf{\Theta}^*, \overline{\mathbf{Z}}_t, \overline{\mathbf{A}}_t)$ , where  $\mathbf{s}_t = \{s_{1,t}, s_{2,t}, \dots, s_{K,t}\}$ ,  $\overline{\mathbf{Z}}_t = \{\overline{Z}_{1,t}, \overline{Z}_{2,t}, \dots, \overline{Z}_{K,t}\}$ ,  $\overline{\mathbf{A}}_t = \{\overline{A}_{1,t}, \overline{A}_{2,t}, \dots, \overline{A}_{K,t}\}$ . Corresponding distributions not knowing the actual models are  $P(\mathbf{s}_{t+1}|\overline{\mathbf{Z}}_t, \overline{\mathbf{A}}_t, \mathcal{F})$ , where framework  $\mathcal{F}$  indicate

MU-POMDP or PLUS. This latter distribution can be approximated via Monte Carlo:

$$P(\mathbf{s}_{t+1}|\overline{\mathbf{Z}}_t,\overline{\mathbf{A}}_t,\mathcal{F}) \cong \sum_{j=1}^J P(\mathbf{s}_{t+1}|\overline{\mathbf{Z}}_t,\overline{\mathbf{A}}_t,\mathbf{\Theta}^{(j)})$$
(3)

where samples  $\{\mathbf{\Theta}^{(j)}\}_{j=1}^{J} \sim p(\mathbf{\Theta}|\mathbf{\bar{Z}}_{t}, \mathbf{\bar{A}}_{t}, \mathcal{F})$  are generated according to the posterior distribution within framework  $\mathcal{F}$ , using the scheme outlined in Figure 2.

Error in the inference can be measured by the Kullback-Leibler (KL) divergence (Cover and Thomas 2006), that is a measure of information lost when one distribution is used to approximate another:

$$\varepsilon(\mathbf{O}^*, \mathbf{\bar{Z}}_t, \mathbf{\bar{A}}_t, \mathcal{F}) = \\ \operatorname{KL}[P(\mathbf{s}_{t+1} \mid \mathbf{O}^*, \mathbf{\bar{Z}}_t, \mathbf{\bar{A}}_t), P(\mathbf{s}_{t+1} \mid \mathbf{\bar{Z}}_t, \mathbf{\bar{A}}_t, \mathcal{F})]$$
(4)

Function  $\varepsilon(\mathbf{\Theta}^*, \overline{\mathbf{Z}}_t, \overline{\mathbf{A}}_t, \mathcal{F})$  depends on the specific framework (MU-POMDP or PLUS) and realization of model, actions, observations. Despite expected value can be taken (as done in Memarzadeh et al. 2014b), in this paper we validate the effectiveness of MU-POMDP vs PLUS on a specific realization. To do so, we have sampled farm model  $\Theta^*$  from the MU-POMDP priors outlined above, and actions  $\overline{\mathbf{A}}_t$ and observations  $\overline{\mathbf{Z}}_t$  consequently. For t = 60and 5 turbines, we estimate via sampling that  $\varepsilon(\mathbf{\Theta}^*, \overline{\mathbf{Z}}_t, \overline{\mathbf{A}}_t, \mathcal{F} = \text{PLUS}) = 5.47\%$ and  $\varepsilon(\mathbf{\Theta}^*, \overline{\mathbf{Z}}_t, \overline{\mathbf{A}}_t, \mathcal{F} = \mathrm{MU}_\mathrm{POMDP}) = 2.93\%$ . As expected, since data are simulated from the MU-POMDP framework, this is more effective in identifying the correct beliefs.

Figure 5 reports examples of the inference process, plotting samples for one entry in the transition matrix under DN, for component 1 and 2. The red star shows the value used for simulating the data, while the green points shows the samples generated from the prior distribution under MU-POMDP (a), the posterior at t = 60 (b), and the posterior for PLUS (c). It is worth noting that variables are dependent under MU-POMDP and identical under PLUS (and that the generating model assumes similar but not identical values).



Figure 5. Examples of samples of model parameter (green dots) and exact value (red star) for MU-POMDP (a-b) and PLUS (c).

In light of this, it is worth describing in details how different agents consider the collected observations, for the sake of inferring

the model parameters. According to the MU-POMDP formulation, observations can be partitioned in two subsets. As shown in Figure 1, observations  $\mathbf{Z}_k$ , collected on component k, are particularly useful to infer model parameters  $\boldsymbol{\theta}_k = \{\mathbf{T}_k, \mathbf{O}_k\}, \text{ and we can call them "direct}$ measures". On the other hand, observations  $\mathbf{Z}_{l\neq k}$ , collected on all components except the k-th one, are also useful for inferring  $\theta_k$ , but only via the hyper-parameters  $\alpha$  and  $\beta$ , and we call them "indirect measures". In the limit for K and tgoing to infinite, the set of indirect measures is equivalent to a perfect observation of the hyperparameters. This, however, would not allow getting a perfect prediction of  $\mathbf{\theta}_k$ . On the other hand, for t going to infinite and if all actions are sufficiently explored, the set of direct measures correspond to observing  $\mathbf{\theta}_k$  directly. Figure 4 shows that PLUS does not apply the distinction between direct and indirect measures: it puts all measures on the same level, for the sake of inferring  $\boldsymbol{\theta}$ .

We can quantify the effect of learning by measuring the KL at different times: Table 1, reports the KL for transition and emission probabilities, computed with a formula similar to Eq.4. It can be noted how KL decays with time, and how that of MU-POMDP is less than that of PLUS.

1 LOS jor ir unsilion una emission models			
	t = 0	t = 60	
MU-POMDP: transition	0.0215	0.0348	
PLUS: transition	0.0158	0.0161	
MU-POMDP: emission	0.0039	0.0120	
PLUS: emission	0.0037	0.0053	

Table 1. Comparison of KL of MU-POMDP andPLUS for transition and emission models

#### 3.4. Simulation of the planning phase

In the last numerical campaign, we investigate the economic impact of adopting MU-POMDP, showing how the more accurate learning algorithm, which accounts for discrepancies in the component models, allows for a more effective planning phase.

Figure 6 shows the cost of operation and maintenance for the farm (i.e. the negative reward) as a function of the time step for (i) an agent with perfect knowledge about the actual model parameters (black line), (ii) an agent following MU-POMDP (red line), and (iii) an agent following PLUS (blue line). Estimates are based 16 independent simulations in the time domain. Agent (i) represents a lower limit for the cost, as she has no uncertainty on the model parameters. For the specific example, the benefit of adopting MU-POMDP instead of PLUS can be quantified in about 2K\$ per turbine per time steps, that is a value similar to the gap between a perfect knowledge on the model and MU-POMDP. Of course, these values depend on the specific application, and are not constant in time. Coefficient of variation for results is around 0.2.



*Figure 6. Average cost of O&M per time and turbine for three agents.* 

#### 4. CONCLUSIONS

We have proposed the MU-POMDP framework using hierarchical Bayesian modeling approach. It extends our previously proposed PLUS algorithm by allowing selecting a level of similarity between components of the system. The computational complexity of the MU-POMDP framework is higher than PLUS as it requires an extra layer of hyper-parameters to model the dependence among components and it with number of components. is linear Specifically, MU-POMDP makes use of the M-H scheme, and this asks for numerical calibration,

e.g. in the selection of appropriate proposal distributions and burn-in phases. The practical efficiency of this framework will depend on its detailed numerical implementation. However, we have shown on a simple application that it has the potential for significant improvement with respect to other formulations. Furthermore, applications to systems with high costs for operation and maintenance, as wind farms, easily justify the adoption of accurate and computationally complex frameworks.

#### ACKNOWLEDGEMENTS

This work is supported in part by the Pennsylvania Infrastructure Technology Alliance, a partnership of Carnegie Mellon, Lehigh University and the Commonwealth of Pennsylvania's Department of Community and Economic Development (DCED), via grant PITA YR16 31571.1.9.1042204. The authors also acknowledge Kevin Wigell and EverPower Wind Holdings for their collaboration.

### 5. REFERENCES

- Bertsekas, D.P. (1996). "Dynamic programming and optimal control" *Athena Scientific*, Belmont, MA.
- Byon, E., Ntaimo, L., and Ding, Y. (2010). "Optimal maintenance strategies for wind turbine system under stochastic weather conditions" *IEEE Transaction on Reliability*, 59(2), 393-404.
- Byon, E., and Ding, Y. (2010). "Season-dependent condition-based maintenance for a wind turbine using a partially observed Markov decision process" *IEEE transaction on Power Systems*, 25(4), 1823-1834.
- Carter, C.K., and Kohn, R. (1994). "On Gibbs sampling for state space models" *Biometrika*, 81(3), 541-553.
- Cover, T.M., and Thomas, J.A. (2006). "Elements of information theory" *John Wiley & Sons, Inc.*
- Fruhwirth-Schnatter, S. (2006). "Finite mixture and Markov switching models" *Springer*, New York.
- Jaulmes, R., Pineau, J., and Precup, D. (2005a). "Active learning in partially observable Markov decision processes" *European Conf. on Machine Learning*, Porto, Portugal, 601-608.
- Jaulmes, R., Pineau, J., and Precup, D. (2005b). "Learning in non-stationary partially observable Markov decision processes" *Eur. Conf. on*

Machine Learning Workshop on Reinf. Learning in Non-Stationary Environments, Porto, Portugal.

- Jensen, F.V., and Nielsen, T.D. (2007). "Bayesian networks and decision graphs" *Springer*.
- Kemp, C, Perfors, A., and Tenenbaum, J.B. (2007). "Learning overhypotheses with hierarchical Bayesian models" *Developmental Science*, 10(3), 307-331.
- MacKay, D.J.C. (2003). "Information theory, inference, and learning algorithms" *Cambridge University Press*, Cambridge, UK.
- Memarzadeh, M., Pozzi, M., and Kolter, J.Z. (2013). "Probabilistic learning and planning for optimal management of wind farms" *Proc.* 9<sup>th</sup> Int. Workshop on Structural Health Monitoring, Stanford, CA, 2720-2728.
- Memarzadeh, M., Pozzi, M., and Kolter, J.Z. (2014a). "Optimal planning and learning in uncertain environments for the management of wind farms" *ASCE J. of Computing in Civil Engineering*, DOI: 10.1061/(ASCE)CP. 1943-5487.0000390.
- Memarzadeh, M., Pozzi, M., and Kolter, J.Z. (2014b). "Managing systems made up by similar components: A probabilistic framework for the maintenance of wind farms" *Proc.* 6<sup>th</sup> World Conf. on Structural Control and Monitoring, Barcelona, Spain.
- Nielesen, J.S., and Sorensen, J.D. (2012). "Maintenance optimization for offshore wind turbines using POMDP" Proc. 16<sup>th</sup> Conf. of Int. Federation for Information Processing on Reliability and Optimization of Structural Systems, 175-182.
- Ross, S., Pineau, B., Chaib-draa, B., and Kreitmann, P. (2011). "A Bayesian approach for learning and planning in partially observable Markov decision process" *J. Machine Learning Research*, 12, 1729-1770.
- Sutton, R.S., and Barto, A.G. (1998). "Reinforcement learning: An introduction" *MIT Press*, Cambridge, MA.
- Smallwood, R.D., and Sondik, E.J. (1973). "The optimal control of partially observable Markov processes over a finite horizon" *Operations research*, 21(5), 1071-1088.
- Sondik, E.J. (1978). "The optimal control of partially observable Markov processes over the infinite horizon" *Operations Research*, 26(2), 282-304.