

Contents

1	Introduction	4
I	Literature, theory and background material	7
2	Literature review	8
2.1	Stochastic model predictive control	8
2.2	Switching model predictive control	12
3	Background theory	16
3.1	Probability theory	16
3.1.1	Discrete random variables	17
3.1.2	Continuous random variables	20
3.2	Graph theory	24
3.3	Probabilistic graphical models	25
3.3.1	Bayesian networks	26
3.3.2	Dynamic Bayesian networks	28
3.4	Control	30
3.4.1	Linear quadratic regulator control	30
3.4.2	Reference tracking	34
3.4.3	Linear quadratic Gaussian control	34
3.4.4	Model predictive control	36
3.5	Matrix identities	37
4	Hidden Markov models	38
4.1	Markov models	38
4.2	Hidden Markov models	39
4.2.1	Filtering	40
4.2.2	Smoothing	41
4.2.3	Viterbi decoding	42
4.2.4	Prediction	43
4.3	Burglar localisation problem	45
5	CSTR model	48

5.1	Qualitative analysis	49
5.2	Nonlinear model	51
5.3	Linearised models	53
II	Single model systems	57
6	Inference using linear models	58
6.1	Kalman filter	59
6.2	Kalman prediction	61
6.3	Smoothing and Viterbi decoding	63
6.4	Filtering the CSTR	64
7	Inference using nonlinear models	69
7.1	Sequential Monte Carlo methods	70
7.2	Particle filter	73
7.3	Particle prediction	75
7.4	Smoothing and Viterbi decoding	75
7.5	Filtering the CSTR	76
8	Stochastic linear control	81
8.1	Unconstrained stochastic control	82
8.2	Constrained stochastic control	85
8.3	Reference tracking	92
8.4	Linear system	92
8.5	Nonlinear system	102
8.6	Conclusion	115
III	Multiple model systems	117
9	Inference using linear hybrid models	118
9.1	Exact filtering	119
9.2	Rao-Blackwellised particle filter	120
9.3	Rao-Blackwellised particle prediction	121
9.4	Smoothing and Viterbi decoding	122
9.5	Filtering the CSTR	122
10	Stochastic switching linear control using linear hybrid models	131
10.1	Unconstrained switching control	134
10.1.1	Most likely model approach	134
10.1.2	Model averaging approach	141
10.2	Conclusion	141
11	Inference using nonlinear hybrid models	143

11.1	Exact filtering	144
11.2	Switching particle filter	144
11.3	Switching particle prediction	145
11.4	Smoothing and Viterbi decoding	146
11.5	Filtering the CSTR	146
12	Stochastic switching linear control using nonlinear hybrid models	152
12.1	Unconstrained switching control	153
12.2	Constrained switching control	157
12.3	Conclusion	164
13	Future work and conclusion	166
13.1	Parameter optimisation	166
13.2	Generalised graphical models	166
13.3	Filtering techniques	167
13.4	Conclusion	167

Part I

Literature, theory and background material

Chapter 2

Literature review

This dissertation primarily deals with stochastic model predictive control but applied within the context of probabilistic graphical models. Section 2.1 briefly covers some recent developments in stochastic model predictive control literature. Section 2.2 briefly covers control schemes, primarily based on model predictive control, where the model control is based upon is automatically adjusted based on plant measurements or manual control laws.

2.1 Stochastic model predictive control

Linear unconstrained stochastic control subject to white additive Gaussian noise is well studied in literature. The solution of the linear quadratic gaussian (LQG) controller,

$$\min_{\mathbf{u}} V(x_0, \mathbf{u}) = \mathbb{E} \left[\frac{1}{2} \sum_{k=0}^{N-1} (x_k^T Q x_k + u_k^T R u_k) + \frac{1}{2} x_N^T P_f x_N \right] \quad (2.1)$$

subject to $x_{t+1} = Ax_t + Bu_t + w_t$ with $w_t \sim \mathcal{N}(0, W)$,

is one of the most fundamental results in stochastic optimal control theory [13]. The reader should note that for stochastic processes it is customary to denote the current time step x_t by x_0 within the control optimisation problem. Also, x_t is a latent variable observable only through $y_t \sim \mathcal{N}(Cx_t, V)$. Boldface is used to denote a vector of vectors over time e.g. $\mathbf{u} = (u_0, u_1, \dots)$.

Using stochastic dynamic programming it is possible to show that the solution of (2.1) is merely the solution of the corresponding fully observed deterministic system, called the linear quadratic regulator (LQR), given the mean of the current state estimate x_0 . A significant drawback of the LQR controller, and by extension the LQG controller, is that it is inherently linear and unconstrained.

Conventional deterministic model predictive control (MPC) is very well studied in literature

[48] and can be seen as the constrained deterministic generalisation of the LQR controller,

$$\begin{aligned} \min_{\mathbf{u}} V(x_0, \mathbf{u}) &= \frac{1}{2} \sum_{k=0}^{N-1} (x_k^T Q x_k + u_k^T R u_k) + \frac{1}{2} x_N^T P_f x_N \\ \text{subject to } x_{t+1} &= A x_t + B u_t \\ \text{and } d^T x_t + e &\geq 0 \quad \forall t = 1, 2, \dots, N, \end{aligned} \tag{2.2}$$

for one constraint with prediction and control horizon length N . The multiple constraint generalisation is straightforward.

A further generalisation of deterministic MPC is stochastic MPC whereby either the variables, the constraints or both have stochastic elements. In current literature the trend is to convert all the stochastic elements of the control problem into deterministic ones. This usually makes the problem more tractable from an analytic and computational point of view.

This conversion is usually achieved via two distinct approaches. In the first approach, which is also the one we employ, the probability distributions are assumed to be Gaussian and the systems linear. This allows one to greatly simplify the problem at the cost of those relatively strong assumptions. The second approach is to use a particle/sampling approach. Here the probability distributions are approximated by particles/samples and no assumptions are made of form of the distributions. It is also not necessary to assume linear dynamics. The major practical drawback of this approach is that it can quickly become computationally intractable for large problems.

This is the approach taken by [8]. They attempt to solve the stochastic MPC problem,

$$\begin{aligned} \min_{\mathbf{u}} \mathbb{E} \left[\frac{1}{2} \sum_{k=0}^{N-1} (x_k^T Q x_k + u_k^T R u_k) + \frac{1}{2} x_N^T P_f x_N \right] \\ \text{subject to } x_{t+1} &= A x_t + B u_t + w_t \\ \text{and } \mathbb{E}[d^T x_t + e] &\geq 0 \quad \forall t = 1, \dots, N \\ \text{and } \Pr(d^T x_t + e \geq 0) &\geq p \quad \forall t = 1, \dots, N, \end{aligned} \tag{2.3}$$

by approximating the current and predicted distributions with particles. Note that w_t is some stochastic variable with known parametrisation.

In their approach a particle filter is used as the state estimator; the current state distribution is approximated by particles of equal weight. An integer variable is introduced for each particle at each predicted time step. The chance constraint is then enforced by requiring that at least a certain number of particles satisfy the constraint. It is not clearly stated but this approach is only valid for particles after resampling.

By using the particle approach to model distributions it is possible to convert the stochastic optimisation problem into a deterministic one. In the case where linear dynamics are used this becomes a mixed integer linear or quadratic programming problem depending on the objective function. The chance constraint becomes an integer constraint. Their algorithm is

appealing because it is not necessary to assume Gaussian distributions or linearity. However, it is possible that the algorithm can become computationally intractable due to the integer constraints which are used to approximate the chance constraints. Since it is necessary to include an integer variable for each particle at each time step in the prediction horizon the number of variables can become large. For large problems with long prediction horizons this can be problematic.

The approach taken by [37] is related to the sampling approach. They convert the stochastic chance constrained optimisation problem into a deterministic nonlinear optimisation problem. They then use a simulation approach to ensure that the chance constraints are satisfied. Their approach is numerically intensive due to the sampling and gradient estimation techniques. The approach taken by [4] uses a randomized optimisation algorithm in concert with the empirical mean of the variables. When the states approach a constraint a penalty method is used to heavily penalise the system to steer it away from the constraint. This causes the system to conservatively satisfy the constraints.

In [36] the stochastic variables are assumed to be Gaussian and the stochastic optimisation problem is transformed into a nonlinear optimisation problem. Using the Gaussian assumption they are able to ensure feasibility and constraint satisfaction albeit conservatively. In [40] the feasibility of stochastically constrained predictive control is considered. Feasibility becomes a problem when predicting under uncertainty. Since the current state estimate is not precisely known and the evolution of the system is stochastic, the certainty of the predicted states often decreases with the prediction horizon. Ensuring constraint satisfaction can become problematic in such situations because of the large predicted uncertainty in the future. In [40] an algorithm enforcing joint chance constraints and recursive feasibility is discussed using a risk allocation approach. While [51] mainly concerns stochastic parameters in the optimisation problem it is shown that chance constraints can, in theory, be rewritten as deterministic constraints if the probability distributions are known and affine constraints are used.

In [54] and [55] an ellipsoidal approximation technique is used to ensure constraint satisfaction for

$$\begin{aligned} & \min_{\mathbf{u}} f(\mathbf{x}) \\ & \text{subject to } x_{t+1} = Ax_t + Bu_t \\ & \text{and } \Pr(d^T x_t + e \geq 0) \geq p \quad \forall t = 1, \dots, N. \end{aligned} \tag{2.4}$$

The authors use the expected value of the stochastic variables in the objective function and system dynamics. They also only consider deterministic linear objective functions. The randomness introduced by the stochastic variables is only addressed in the chance constraint.

If one assumes that each x_t is Gaussian with sufficient statistics (μ_t, Σ_t) and dimension n , then the chance constraint can be satisfied by ensuring that the area of each ellipse $(x - \mu_t)^T \Sigma_t^{-1} (x - \mu_t) = k^2$ for $t = 1, \dots, N$ is contained in the feasible region. We have that k^2 is

chosen by solving the integral equation of the Chi Squared distribution

$$\frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \int_0^{k^2} \mathcal{X}^{\frac{n}{2}-1} e^{-\frac{\mathcal{X}}{2}} d\mathcal{X} = p. \quad (2.5)$$

Using the ellipsoidal approximations the stochastic optimisation problem can be transformed into a second order conic optimisation problem. The authors ensure that each ellipse is contained in the feasible region by ensuring there exists sufficient “back-off” (the distances z_1, z_2, z_3 and z_4) between the predicted ellipses and the constraints as shown in Figure 2.1.

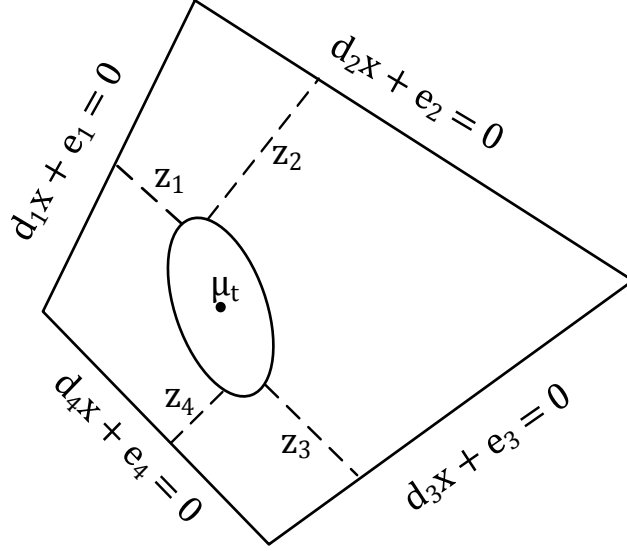


Figure 2.1: Ellipsoidal approximation to ensure chance constraint satisfaction as used by [54] and [55].

While this breakthrough is important – we build on it in our approach – the authors do not state that they are in fact using a form of the Mahalanobis distance to enforce their chance constraints. The approach of using confidence ellipsoids is further refined in [9]. The ellipsoidal approximation technique is also further investigated in [7]; they show that it is possible to reformulate joint chance constraints using univariate Gaussian distributions.

Although [59] and [60] primarily deal with univariate problems they show that if the underlying system is linear and Gaussian, it is possible to manipulate the constrained stochastic problem shown in (2.3) into a deterministic problem. Their analysis allows the stochastic objective function to be transformed into its deterministic equivalent using the properties of Gaussian integrals. This development is quite important because it allows one to directly evaluate the stochastic objective function. The constraints are handled by directly evaluating the Gaussian integral corresponding to the chance constraint in the univariate case. The authors allude to the fact that this becomes computationally intractable in higher dimensions and suggest that the approach in [54] be used. The authors also suggest a way to handle the situation where covariance matrix grows without bound in unstable systems. This is related to the feasibility problems discussed earlier.

In this dissertation the benefits gained by designing controllers within the framework of

probabilistic graphical models is illustrated. This work results in the following conclusions:

1. Under the assumption of normality and linearity it is possible to convert the stochastic objective function of (2.1) into its deterministic equivalent. The analysis is closely related to the work of [59] and [60] but we show that these results are immediately obvious from within the framework of probabilistic graphical models. Thus it is possible to solve the LQG problem without resorting to stochastic dynamic programming.
2. We generalise our analysis to stochastic MPC and show that by using the statistically important metric, the Mahalanobis distance, we arrive at a technique for enforcing chance constraints which is very closely related to the approach by [54] and [55]. Under the assumption of linearity and normality we show that the constraint satisfaction is ensured. Owing to the use of the Mahalanobis distance metric we provide some theoretical support for the use of the “ellipsoidal approximation” technique if the underlying system is nonlinear or not exactly Gaussian.
3. Combining the previous results we show that it is possible to write the joint chance constrained stochastic MPC problem as a deterministic MPC problem. Additionally we show that the joint chance constraints can be written in a linear format. The entire optimisation problem can then be written in the standard form for quadratic programming optimisation. Standard deterministic MPC solution techniques can then be used to solve the stochastic problem.
4. We compare the effect different inference techniques have on the quality of the MPC. Specifically, if the underlying system is nonlinear it can be beneficial to use nonlinear particle filtering techniques as opposed to inherently linear techniques like the Kalman filter.

Lastly, measurement and system noise is ubiquitous in real life systems. Therefore most modern model predictive control systems use an inference (filtering) technique to estimate the current system state. By using a filter the control system can be seen as using a probabilistic graphical model; therefore we are not actually introducing anything exotic but rather highlighting a connection between two rich fields.

2.2 Switching model predictive control

In model predictive control the model of the plant is used to predict the future behaviour of the plant given some inputs which are optimised according to some performance criterion. Classically this model is linear and time invariant. An example of such a model is

$$\begin{aligned}x_{t+1} &= Ax_t + Bu_t + w_t \\ y_t &= Cx_t + v_t.\end{aligned}\tag{2.6}$$

Since the model includes system and measurement noise a state estimator would typically be used to infer the state, x_t , from some observation y_t . The mean of the current state, $\mathbb{E}[x_t]$ together with the deterministic state model $x_{t+1} = Ax_t + Bu_t$ would then be used for prediction [48]. This assumption allows for the use of advanced constrained optimisation algorithms, typically quadratic programming algorithms. From a practical perspective robust and fast optimisation is crucial because it allows control inputs to be calculated on-line [39].

Unfortunately modelling errors or omissions often cause poor controller performance within the context of MPC. This is often observed as steady state offset i.e. the controller takes no more action and the system is not at the set point. In certain cases it is possible to account for plant/model mismatch or asymptotically constant disturbances by incorporating a latent disturbance model within the MPC framework. This is classically called zero offset regulation [48] and is achieved by augmenting the system model

$$\begin{aligned}x_{t+1} &= Ax_t + Bu_t + B_d d_t + w_t \\d_{t+1} &= d_t \\y_t &= Cx_t + C_d d_t + v_t.\end{aligned}\tag{2.7}$$

By using a state estimator the latent integrating disturbance d_t can be estimated. The model used for prediction is then augmented to incorporate this disturbance $x_{t+1} = Ax_t + Bu_t + \mathbb{E}[d_t]$ [34]. The problem with this approach is that it assumes that the model (A, B) is at least somewhat representative of the underlying dynamics.

Linear models (A, B) are often derived by linearising nonlinear models about a single operating point. These models are traditionally used in MPC design; a drawback with this approach is that linear models are fundamentally only accurate in a small region around the point of linearisation. The further the system moves away from the linearisation point the worse the accuracy of the model becomes because the model is no longer a good approximation of the underlying system dynamics. A possible solution to this problem is nonlinear MPC. In nonlinear MPC a full nonlinear model is used for controller design; it is hoped that this more complicated model is accurate over the entire problem domain (or a larger part thereof than the corresponding linear model). Unfortunately this approach is often computationally intensive, especially for large problems, because it inherently requires nonlinear, usually non-convex, optimisation. This is a subject of much current research [18].

Another approach is to approximate a potentially complex nonlinear system by a set of linear functions which are valid in certain ranges. An early attempt at this idea [5] integrated logical rules for switching between different system dynamics and constraints. Using that approach a set of logical rules were integrated into the optimisation problem to yield, in the setting of standard MPC, a mixed integer quadratic program with linear constraints. This approach is called mixed logical dynamical modelling (MLD) in literature; the system dynamics are then specified by

$$x_{t+1} = \sum_{i=1}^I \delta_i A_i x_t + \sum_{i=1}^I \delta_i B_i u_t\tag{2.8}$$

where $i = 1, 2, \dots, I$ are the indices of the models approximating the underlying problem.

The binary variable $\delta_i \in [0, 1]$ selects which linear model to use at each step in the prediction horizon within the framework of MPC. Constraints on δ_i allow the optimisation algorithm to switch between models based on its position in the state space. The drawbacks with this approach is that the “IF-THEN-ELSE” rules need to be fully specified before hand and the mixed integer optimisation problem can become computationally intractable.

The work by [21] and [50] elaborate on this approach. Both papers deal with the control of continuously stirred tank reactors (CSTRs) throughout different operating regimes. CSTRs are a good case study because they often have multiple steady states, constraints and can be quite nonlinear. The approach of [21] and [50] is to linearise the underlying nonlinear CSTR model about different operating points and use those models to approximate the true nonlinear dynamics. Computational difficulties are reported because the complexity of the mixed integer quadratic problem scales exponentially with the number of variables.

The approach by [46] is similar except that they use ellipsoidal regions to develop a multiple model MPC. Again the approach by [33] is similar except that they attempt to reduce the complexity of the mixed integer quadratic problem by providing guidelines on selecting the number of linear models used. More linear models leads to better control but the optimisation problem becomes correspondingly more difficult. Finally, [45] also investigates hybrid systems but uses Bayes’ theorem to assign weights to the different models as opposed to only having one model active in each region. In their approach δ_i is a continuous variable in the range $[0, 1]$ and $\sum_{i=1}^I \delta_i = 1$. While their approach is also computationally intensive, a nonlinear optimisation problem needs to be solved, an effort is made to take advantage of the problem structure to attenuate this problem.

Loosely related to the idea of model switching is model based fault detection. In [28] it is found that almost 70% of the reviewed papers dealing with fault detection use observer or parameter estimation methods. The basic idea behind this approach is to estimate the system outputs using an observer and to compare this to some model of the system. The difference, often called the residual, is then used to estimate the probability of a fault in the process [23]. The approach followed in [57] combines elements of model switching and state observations. They use a filter to construct a residual generator which is used to evaluate whether or not the system has a fault. The filter can switch between the different linear system models based on the current regime of the system.

Related to this class of model based fault detection algorithms is the switching model filter. The corresponding probabilistic graphical model is shown in Figure 2.2. Classically these types of models are used to infer the current state estimate (called filtering) given a set of models. Intuitively, the model which best describes the current observation (y_t) is assigned the most weight in the state estimate. This allows for the modelling of nonlinear and multi-modal distributions by combining models based on the inferred switching variable (s) [43]. These models may be used for both filtering and event detection (see [56] for an example)

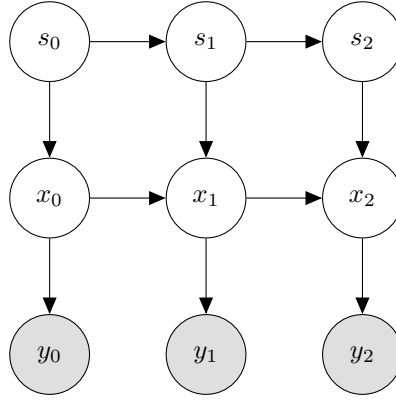


Figure 2.2: Switching filter graphical model.

which, in the setting of control, can be interpreted as fault detection.

In this dissertation we also illustrate the potential control benefits gained by designing predictive controllers using the switching model probabilistic graphical model of Figure 2.2. We investigate and show the following:

1. Using switching model filters it is possible to accurately estimate the current state as well as estimate the probability of each model supporting the current observation. The state estimate can then be used within the controller framework alluded to in Section 2.1. Using this approach it is also possible to change the model control is based upon. The linear model which best describes the current observations is used for control purposes. In this way we gain the benefit of model switching but do not incur the significant computational cost when it is incorporated into the optimisation problem as is the trend in current literature. We call this approach the switching controller algorithm.
2. We investigate using the Rao-Blackwellised particle filter as the switching model filter. In this context the resultant most likely linear model is used to move the underlying system to different regions in state space. It was found that the approach caused controller instability because the current most likely model is often not accurate enough to steer the system to the target.
3. We also investigate using the switching particle filter as the switching model filter. In this context the filter was used to identify when a process fault occurred and, based on this event, switch the model control is based upon. It was found that the algorithm successfully stabilised and regulated the nonlinear underlying system. Furthermore, since the controllers of Section 2.1 were used the control scheme was computationally efficient.

Even more so than Section 2.1, the approach used here relies on the application of probabilistic graphical models to controller design.

Part II

Single model systems

Part III

Multiple model systems

Bibliography

- [1] Y. Bar-Shalom, X.R. Li, and T. Kirubarajan. *Estimation with applications to tracking and navigation*. John Wiley and Sons, 2001.
- [2] D. Barber. Expectation correction for smoothed inference in switching linear dynamical systems. *Journal of Machine Learning*, 7:2515–2540, 2006.
- [3] D. Barber. *Bayesian Reasoning and Machine Learning*. Cambridge University Press, 2012.
- [4] I. Batina, A.A. Stoorvogel, and S. Weiland. Optimal control of linear, stochastic systems with state and input constraints. In *Proceedings of the 41st IEEE Conference on Decision and Control*, 2002.
- [5] A. Bemporad and M. Morari. Control of systems integrating logic, dynamics, and constraints. *Automatics*, 35:407–427, 1999.
- [6] C.M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [7] L. Blackmore, Hui Li, and B. Williams. A probabilistic approach to optimal robust path planning with obstacles. In *American Control Conference*, June 2006.
- [8] L. Blackmore, O. Masahiro, A. Bektassov, and B.C. Williams. A probabilistic particle-control approximation of chance-constrained stochastic predictive control. *IEEE Transactions on Robotics*, 26, 2010.
- [9] M. Cannon, B. Kouvaritakis, and X. Wu. Probabilistic constrained mpc for multiplicative and additive stochastic uncertainty. *IEEE Transactions on Automatic Control*, 54(7), 2009.
- [10] A.L. Cervantes, O.E. Agamennoni, and J.L. Figueroa. A nonlinear model predictive control system based on weiner piecewise linear models. *Journal of Process Control*, 13:655–666, 2003.
- [11] R. Chen and J.S. Liu. Mixture kalman filters. *Journal of Royal Statistical Society*, 62(3):493–508, 2000.

- [12] J.J. Dabrowski and J.P. de Villiers. A method for classification and context based behavioural modelling of dynamical systems applied to maritime piracy. *Expert Systems with Applications*, 2014.
- [13] B.N. Datta. *Numerical Methods for Linear Control Systems - Design and Analysis*. Elsevier, 2004.
- [14] F. Daum and J. Huang. Particle flow for nonlinear filters. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 5920–5923, May 2011.
- [15] M. Davidian. Applied longitudinal data analysis. North Carolina State University, 2005.
- [16] J.P. de Villiers, S.J. Godsill, and S.S. Singh. Particle predictive control. *Journal of Statistical Planning and Inference*, 141:1753–1763, 2001.
- [17] N. Deo. *Graph Theory with Applications to Engineering and Computer Science*. Prentice-Hall, 1974.
- [18] M. Diehl, H.J. Ferreau, and N. Haverbeke. Efficient numerical methods for nonlinear mpc and moving horizon estimation. *Control and Information Sciences*, 384:391–417, 2009.
- [19] A. Doucet and A.M. Johansen. A tutorial on particle filtering and smoothing: fifteen years later. Technical report, The Institute of Statistical Mathematics, 2008.
- [20] A.D. Doucet, N.J. Gordon, and V. Krishnamurthy. Particle filters for state estimation of jump markov linear systems. *IEEE Transactions on Signal Processing*, 49(3):613–624, March 2001.
- [21] J. Du, C. Song, and P. Li. Modeling and control of a continuous stirred tank reactor based on a mixed logical dynamical model. *Chinese Journal of Chemical Engineering*, 15(4):533–538, 2007.
- [22] The Economist. In praise of bayes. Article in Magazine, September 2000.
- [23] C. Edwards, S.K. Spurgeon, and R.J. Patton. Sliding mode observers for fault detection and isolation. *Automatica*, 36:541–553, 200.
- [24] H.C. Edwards and D.E. Penny. *Elementary Differential Equations*. Pearson, 6th edition edition, 2009.
- [25] W. Forst and D. Hoffmann. *Optimisation - Theory and Practice*. Springer, 2010.
- [26] O.R. Gonzalez and A.G. Kelkar. *Electrical Engineering Handbook*. Academic Press, 2005.
- [27] N.J. Gordon, D.J. Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proceedings-F*, 140(2):107–113, 1993.

- [28] R. Isermann and P. Balle. Trends in the application of model based fault detection and diagnosis of technical processes. *Control Engineering Practice*, 5(5):709–719, 1997.
- [29] K. Ito and K. Xiong. Gaussian filters for nonlinear filtering problems. *IEEE Transactions on Automatic Control*, 45(5):910–928, 2000.
- [30] R. J. Jang and C.T. Sun. *Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence*. Prentice-Hall, 1996.
- [31] D. Koller and N. Friedman. *Probabilistic Graphical Models*. MIT Press, 2009.
- [32] K. B. Korb and A. E. Nicholson. *Bayesian Artificial Intelligence*. Series in Computer Science and Data Analysis. Chapman & Hall, first edition edition, 2004.
- [33] M. Kvasnica, M. Herceg, L. Cirka, and M. Fikar. Model predictive control of a cstr: a hybrid modeling approach. *Chemical Papers*, 64(3):301–309, 2010.
- [34] J.H. Lee, M. Morari, and C.E. Garcia. *Model Predictive Control*. Prentice Hall, 2004.
- [35] U.N. Lerner. *Hybrid Bayesian Networks for Reasoning about Complex Systems*. PhD thesis, Stanford Univesity, 2002.
- [36] P. Li, M. Wendt, H. Arellano-Garcia, and G. Wozny. Optimal operation of distrillation processes under uncertain inflows accumulated in a feed tank. *American Institute of Chemical Engineers*, 2002.
- [37] P. Li, M. Wendt, and G. Wozny. A probabilistically constrained model predictive controller. *Automatica*, 38:1171–1176, 2002.
- [38] W.L. Luyben. *Process Modeling, Simulation and Control for Chemical Engineers*. McGraw-Hill, 2nd edition edition, 1990.
- [39] J.M. Maciejowski. *Predictive Control with constraints*. Prentice-Hall, 2002.
- [40] O. Masahiro. Joint chance-constrained model predictive control with probabilistic resolvability. *American Control Conference*, 2012.
- [41] P. Mhaskar, N.H. El-Farra, and P.D. Christofides. Stabilization of nonlinear systems with state and control constraints using lyapunov-based predictive control. *Systems and Control Letters*, 55:650–659, 2006.
- [42] K.P. Murphy. Switching kalman filters. Technical report, Compaq Cambridge Research Lab, 1998.
- [43] K.P. Murphy. *Dynamic Bayesian Networks: Representation, Inference and Learning*. PhD thesis, University of California, Berkeley, 2002.
- [44] K.P. Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012.

- [45] N. Nandola and S. Bhartiya. A multiple model approach for predictive control of non-linear hybrid systems. *Journal of Process Control*, 18(2):131–148, 2008.
- [46] L. Ozkan, M. V. Kothare, and C. Georgakis. Model predictive control of nonlinear systems using piecewise linear models. *Computers and Chemical Engineering*, 24:793–799, 2000.
- [47] T. Pan, S. Li, and W.J. Cai. Lazy learning based online identification and adaptive pid control: a case study for cstr process. *Industrial Engineering Chemical Research*, 46:472–480, 2007.
- [48] J.B. Rawlings and D.Q. Mayne. *Model Predictive Control*. Nob Hill Publishing, 2009.
- [49] B. Reiser. Confidence intervals for the mahalanobis distance. *Communications in Statistics: Simulation and Computation*, 30(1):37–45, 2001.
- [50] Y. Sakakura, M. Noda, H. Nishitani, Y. Yamashita, M. Yoshida, and S. Matsumoto. Application of a hybrid control approach to highly nonlinear chemical processes. *Computer Aided Chemical Engineering*, 21:1515–1520, 2006.
- [51] A.T. Schwarm and Nikolaou. Chance constrained model predictive control. Technical report, University of Houston and Texas A&M University, 1999.
- [52] C. Snyder, T. Bengtsson, P. Bickel, and J. Anderson. Obstacles to high-dimensional particle filtering. *Mathematical Advances in Data Assimilation*, 2008.
- [53] S.J. Streicher, S.E. Wilken, and C. Sandrock. Eigenvector analysis for the ranking of control loop importance. *Computer Aided Chemical Engineering*, 33:835–840, 2014.
- [54] D.H. van Hessem and O.H. Bosgra. Closed-loop stochastic dynamic process optimisation under input and state constraints. In *Proceedings of the American Control Conference*, 2002.
- [55] D.H. van Hessem, C.W. Scherer, and O.H. Bosgra. Lmi-based closed-loop economic optimisation of stochastic process operation under state and input constraints. In *Proceedings of the 40th IEEE Conference on Decision and Control*, 2001.
- [56] H. Veeraraghavan, P. Schrater, and N. Papanikolopoulos. Switching kalman filter based approach for tracking and event detection at traffic intersections. *Intelligent Control*, 2005.
- [57] D. Wang, W. Wang, and P. Shi. Robust fault detection for switched linear systems with state delays. *Systems, Man and Cybernetics*, 39(3):800–805, 2009.
- [58] R.S. Wills. Google’s pagerank: the math behind the search engine. Technical report, North Carolina State University, 2006.

- [59] J. Yan and R.R. Bitmead. Model predictive control and state estimation: a network example. In *15th Triennial World Conference of IFAC*, 2002.
- [60] J. Yan and R.R. Bitmead. Incorporating state estimation into model predictive control and its application to network traffic control. *Automatica*, 41:595–604, 2005.
- [61] M.B. Yazdi and M.R. Jahed-Motlagh. Stabilization of a cstr with two arbitrarily switching modes using model state feedback linearisation. *Chemical Engineering Journal*, 155(3):838–843, 2009.