A generative approach to qualitative trend analysis for batch process fault diagnosis

Kris Villez*,†, Raghunathan Rengaswamy‡

*Purdue University, School of Chemical Engineering, West Lafayette, IN, USA

†Eawag, Department of Process Engineering, Dübendorf, Switzerland

kris.villez@eawag.ch

‡IIT Madras, Department of Chemical Engineering, Chennai, TN, India

raghur@iitm.ac.in

Abstract— Most of the existing methods for qualitative trend analysis are based on discriminative models. A disadvantage of such models is that many heuristic rules or local search methods are needed. Recently, an effort has been made to develop a globally optimal method for qualitative trend analysis. This method is based on a generative (rather than discriminative) model and has shown to lead to excellent performance. However, this method comes at an extreme computational demand which renders the methods unlikely for on-line application. In this work, an alternative method, while still generative in nature, is proposed which is shown to deliver the same performance while reducing the computational demand considerably.

I. INTRODUCTION

Qualitative trend analysis has been proposed most frequently to solve problems in the context of fault detection and identification (FDI) [5], [8], [10], [12]–[17]. The underlying philosophy is that many processes and their fault scenarios are ill-understood in the sense that the responses of key variables under different fault scenarios can only be described in a coarse-grained qualitative manner. If one has established a set of fault scenarios and their corresponding qualitative descriptions, the key challenge becomes to convert on-line data to matching qualitative representations, i.e., qualitative trend analysis. Existing methods for qualitative trend analysis are largely based on discriminative models, meaning that time series data is processed in a step-wise or iterative procedure to obtain a coarse-grained representation of the same. Such methods can be based on piece-wise polynomial function fitting [4], [10], wavelet analysis [1] or neural network modelling [12]. With each of these methods, the qualitative representation is conditional to the data in an explicit manner through the applied algorithm.

In most cases, a tabular look-up strategy is used to find a match between reported fault scenarios and the qualitative data representations. Unfortunately, such a search strategy can be unreliable because of noise in the data series, leading to incorrect qualitative representations. Also, typical tabular look-up strategies use an all-or-nothing approach in matching qualitative descriptions of fault scenarios and qualitative data representations. Some authors have attempted at accomodating this through fuzzy matching or quantification of similarities in shape [2], [10] though one may argue that these have stepped away from the original coarse-grained representation concept through making use of a quantitative representation.

In contrast to the majority of existing methods, a recently developed method avoids the use of tabular look-up search strategies [14]. In this method, a generative model in which the likelihood of the time series data is conditional to a proposed qualitative representation is proposed. The method fits spline functions -a special class of piece-wise polynomial functions- to the data series. The likelihood of a particular qualitative representation can then be evaluated by means of nonlinear constraints on the spline coefficients. Finding the qualitative representation which maximizes this likelihood cannot be solved analytically. Put otherwise, the desired, optimal qualitative representation is only implicitly related to a given time series through this non-linear optimization problem. In [14], [15], this problem is solved by means of a branch-and-bound search algorithm. Both lower and upper bounds for this problem can be found through solving a Second Order Cone Program (SOCP). Unfortunately, this becomes a lengthy procedure in most practical cases.

In this work the original nonlinearly constrained model is replaced with an alternative model based on a Hidden Markov Model (HMM). While this alternative model is only an approximation to the original model, it is still generative in nature and avoids a tabular look-up strategy. Results indicate that this approach is viable and delivers similar results as the original method while substantially reducing the computational load.

II. MATERIALS AND METHODS

A. General notes on qualitative trend analysis

To understand the remainder of the text, the following concepts and definitions are necessary. In general, qualitative trend analysis is concerned with the segmentation of time series into so called episodes. Such episodes are characterized by means of a start time, an end time, and a unique set of signs for the first and second derivative. This means that between the considered start and end time, the signs of the derivatives are considered to remain the same.

A combination of a specific sign for the first and second derivative is referred to as a primitive and is usually represented by a character. If one considers the available signs to be (strictly) negative, zero, or (strictly) positive, then

Characters (signs)	Sign of 1st derivative		
	-	0	+
Sign of 2nd derivative +	A (-,+)	–	B (+,+)
0	E (-,0)	F (0,0)	G (+,0)
-	D (-,-)	–	C (+,-)

TABLE I

OVERVIEW OF PRIMITIVES AS FUNCTION OF THE FIRST AND SECOND DERIVATIVE.

one can describe all primitives with 7 characters. Table I lists these primitives along with the derivative signs. While the characters can be assigned arbitrarily, they are assigned here so that (1) A corresponds to acceleration (positive second derivative), (2) D to a deceleration (negative second derivative), and (3) ABCD is a feasible qualitative sequence for a function with smooth first and second derivatives.

A complete segmentation into multiple contiguous episodes is referred to as a qualitative representation (QR). A sequence of primitives, thus a qualitative representation without time information, is referred to as a qualitative sequence (QS).

B. Algorithm for qualitative trend analysis

The proposed method for qualitative trend analysis consists of two steps. The two steps are explained in the following paragraphs.

1) Step 1: Probabilities for primitives via kernel regression: In the first step, one evaluates how likely it is whether the first and second derivative are positive, resp., negative. To this end, one first obtains a local polynomial fit to the data. This means that the vector of polynomial coefficients are estimated to a univariate data series consisting of pairs of independent variables, t_j, which usually is time, and measurements of the dependent variable, y_j ($j = 1 \ldots m$). To obtain a local fit, the data are given decreasing weights with increased distance from the considered sample of interest, i. Concretely, one thus optimizes the coefficients, β_i, for a sample of interest, i, according to the following objective function:

$$\min_{\beta_i} \sum_{j=1}^{m} K(t_i, t_j) \cdot (y_j - \mathbf{x}_j^T \cdot \beta_i)^2 \qquad (1)$$

with:

$$\mathbf{x}_j = \begin{bmatrix} (t_j - t_i)^0 & (t_j - t_i)^1 & \cdots & (t_j - t_i)^p \end{bmatrix}^T$$

The degree of the polynomial is p. For the purpose of qualitative trend analysis, choices for the polynomial are restricted to quadratic or higher-degree polynomials (degree 2 or higher, order 3 or higher). p is set to 2 in this study (quadratic polynomial). The weights are expressed by a kernel function, $K(x_i, x_j)$ in which case the local fitting is also known as kernel regression [6]. A wide variety of kernels exist. In our work, we use the so called tri-cube kernel:

$$K_{i,j} = K(t_i, t_j) = \left(1 - d_{i,j}^3\right)^3 \mid d \leq 1 \qquad (2)$$
$$d_{i,j} = \|t_i - t_j\| / \lambda$$

This kernel is parameterized by a meta-parameter, λ, which is known as the kernel width. Any data sample, j, with a distance to the considered sample i larger than this kernel width has zero weight and thus has no influence on the parameter estimates. For sufficiently small values for the kernel width on thereby obtains the local regression property of the method. With the above local polynomial model, one estimates a vector of polnomial coefficients for each data sample ($i = 1 \ldots m$). This can be solved analytically as follows:

$$\beta_i = H_i \cdot \mathbf{y} \qquad (3)$$

with:

$$H_i = (X_i' \cdot W_i \cdot X_i)^{-1} \cdot X_i' \cdot W_i$$
$$X_i = \begin{bmatrix} (\mathbf{t} - t_i)^0 & (\mathbf{t} - t_i)^1 & \cdots & (\mathbf{t} - t_i)^p \end{bmatrix}$$
$$W_i = \begin{bmatrix} K_{i,1} & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & & K_{i,j} & & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & K_{i,m} \end{bmatrix}$$

In the above equations, X_i is the $m \times (p+1)$ polynomial basis matrix and W_i is an $m \times m$ diagonal matrix with the weights for each data sample j for local regression in sample i. These equations demonstrate that the polynomial fitting problem corresponds to a linear projection scheme. As a result, one can also obtain point-wise covariances for the regression coefficients, assuming independent and identically distributed measurement errors following a Gaussian distribution, $N(0, \sigma_y)$,:

$$\Sigma_{\beta_i} = H_i \cdot \Sigma_y \cdot H_i^T \qquad (4)$$
$$\text{with: } \Sigma_y = \sigma_y \cdot \mathbf{I}_n$$

The point-wise conditional distributions for the linear and quadratic polynomial coefficient are then:

$$\beta_{1,i} \sim N(\beta_i(2), \Sigma_{\beta_i}(2,2)) = N(\mu_{1,i}, \sigma_{1,i}) \qquad (5)$$
$$\beta_{2,i} \sim N(\beta_i(3), \Sigma_{\beta_i}(3,3)) = N(\mu_{2,i}, \sigma_{2,i})$$

Then, the probability that the first (second) derivative is positive, P1 (P2), corresponds to the integration of the normal distribution as follows:

$$P_{1,i} = \int_{u=0}^{+\infty} \frac{1}{\sigma_{1,i}\sqrt{2\pi}} e^{-\frac{(u-\mu_{1,i})^2}{2 \cdot \sigma_{1,i}^2}} \, du \qquad (6)$$
$$P_{2,i} = \int_{u=0}^{+\infty} \frac{1}{\sigma_{2,i}\sqrt{2\pi}} e^{-\frac{(u-\mu_{2,i})^2}{2 \cdot \sigma_{2,i}^2}} \, du$$

Similarly, the likelihood for negative derivatives are computed as $1 - P1$ and $1 - P2$. Finally, to compute the likelihood of a particular qualitative state or primitive, $P_i(s)$, one multiplies the corresponding probablities. Note that only primitives A, B, C, D are considered in this study, leading

to the computation of four probabilities, $P_i(prim), prim \in \{A, B, C, D\}$, at each sample, i, as follows:

$$
\begin{aligned}
P_i(A) &= (1 - P_{1,i}) \cdot P_{2,i} \quad\quad (7) \\
P_i(B) &= P_{1,i} \cdot P_{2,i} \\
P_i(C) &= P_{1,i} \cdot (1 - P_{2,i}) \\
P_i(D) &= (1 - P_{1,i}) \cdot (1 - P_{2,i})
\end{aligned}
$$

Note that the above approach to compute the probability of a particular primitive at a given sample time is approximate because the correlation between the estimates of $\beta_{1,i}$ and $\beta_{2,i}$ is not accounted for.

2) Step 2: Optimal state sequence estimation: In Step 1 of the algorithm, probabilities are assigned to all considered primitves at each data sample. To obtain a most likely sequence of such primitives, the Viterbi algorithm is used. To this end, a qualitative state of the monitored process is assumed to evolve from data sample to data sample according to a stochastic first-order discrete Markov process. This model is thus associated with a given QS and the number of Markov states matches the length of the QS:

$$
\pi_{QS,k}(s) = \sum_a T_{QS}(s, q) \cdot \pi_{QS,k-1}(q) \quad\quad (8)
$$

where $\pi_{QS,k}(s)$ and $\pi_{QS,k-1}(q)$ represent the probablities of the state s, resp. q, at time k, resp. $k - 1$. T_{QS} is a matrix of transition probabilities. The subscript QS indicates that one has a different Markov chain for different QSs. The transition probabilities are set up so that all states are non-ergodic (there is no path back to any given state once a transition is made from this same state), that there is a single source and a single sink state, and that a single, linear path is formed from the source state to sink state. The states are indexed incrementally from the source to sink state (1 to n). Practically this means that the elements of the diagonal and the subdiagonal of the transition matrix are non-zero while all other elements are zero:

$$
T_{QS} = \begin{bmatrix}
\tau_{1,1} & 0 & 0 & \cdots & 0 \\
\tau_{2,1} & \tau_{2,2} & 0 & \cdots & 0 \\
0 & \tau_{3,2} & \tau_{3,3} & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & \tau_{n,n}
\end{bmatrix} \quad (9)
$$

Each qualitative state as defined for the above Markov chain is associated with a single primitive. Note that the opposite is not necessarily true. This makes it possible to form a qualitative sequence with multiple states corresponding to the same primitive (e.g., BCBC). Importantly, the likelihood of a qualitative state conditional to the data, $P_i(s)$, is equivalent to the likelihood of the primitive associated with this state, $P_i(prim)$. Using this equivalance, one finds the maximum likelihood sequence of qualitative states by means of Viterbi algorithm. This recursive algorithm is well-known for HMM based sequence estimation [11] and is implemented as follows for qualitative analysis:

$$
\begin{aligned}
L_{QS,0}(s) &= \pi_{QS,0}(s) \\
L_{QS,i}(s) &= P_i(s). \max_{q \in S} (T_{QS}(s, q) \cdot L_{QS,i-1}(q)) \quad (10)
\end{aligned}
$$

where $L_{i,QS}(s)$ represents the likelihood of the most probable state sequence with s as state at sample i. $T_{QS}(s, q)$ are the transition probabilities as above and $P(s)_i$ are the state probabilities as computed in Eq. 8 for the corresponding primitives. The values for $\pi_{QS,0}(s)$ represent the initial probabilities for each state. These are set so that the first state in the linear chain has probability of one and all others probability of zero. The maximization in Eq. 10 determines the predecessor state, q, at each time instant. For $i = m$ and $s = n$, thus at the end of the data series and for the last state in the Markov chain, one obtains the maxmimum likelihood value, $L_{QS,m}(n)$ for the considered sequence. Note that this likelihood is only an approximation of the true likelihood because (1) correlation between $\beta_{1,i}$ and $\beta_{2,i}$ is not accounted for and (2) cross-correlation of the same estimates across the time index are not accounted for. While the correlation between the the polynomial coefficients could be accounted for by approximate integration of the multivariate normal distribution, one cannot account for cross-correlation easily since the Viterbi algorithm requires additivity (of the log-likelihood terms) to work properly.

C. Fault diagnosis based on qualitative analysis

To use the above method for fault diagnosis, the same strategy as in [15] is used. Namely, one associates a QS and prior likelihood with each fault scenario (S). The considered QSs and associated HMM models are defined by a process experts who interpret plotted time series on a regular basis as part of their process monitoring task set. In benchmark cases, like this one, one can inspect and analyze the noise-free data for the true qualitative sequence. Note that qualitative simulation [7] or Signed Directed Graphs [9] can be used to determine likely sequences starting from a detailed process model. As discussed above, an HMM is set up for each QS, and thus for each fault scenario. Based on the maximum likelihood as obtained for each HMM with the same data, one can then compute the a posteriori likelihood of each scenario (S) conditional to the data as follows:

$$
\begin{aligned}
L_{post}(S) &= L_{cond}(S) \cdot L_{prior}(S) \quad\quad (11) \\
&= L_{QS,m}(n) \cdot L_{prior}(S)
\end{aligned}
$$

Thus, the a posteriori likelihood of a scenario, $L_{post}(S)$, is computed as the product of the prior likelihood of the scenario ($L_{prior}(S)$) and the conditional likelihood ($L_{cond}(S)$). The prior likelihood is given by the model user. The conditional likelihood of the scenario is approximated as the maximum likelihood of the associated QS and HMM as obtained through the Viterbi algorithm. Therefore, one executes the Viterbi algorithm once for each fault scenario, each time with a different Markov chain corresponding to the respective qualitative sequence. Fault diagnosis then ends with selecting the scenario with a maximum posterior likelihood.

D. Benchmark model simulations

The data set used in this study has been used in previous studies to benchmark existing and new techniques for qualitative trend analysis [14], [17]. The data set represents
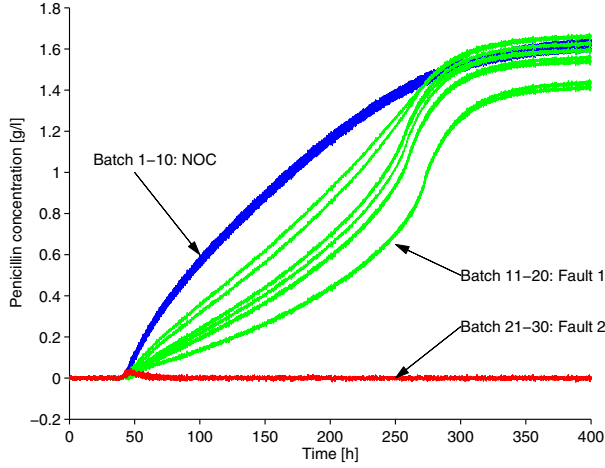
Fig. 1. Penicillin concentration measurement profiles for 30 batches. Taken from [15]

Fig. 2. Top: Time series for a batch belonging to the second scenario (Fault 1). Bottom: Normalized 1st and 2nd derivative with point-wise 3-σ confidence intervals.

three groups of ten batch fermentations according to the benchmark model presented in [3]. Each group corresponds to one of three scenarios. The scenarios are respectively (1) Normal operation conditions (NOC), (2) Fault 1: Reduced saturation constant, and (3) Fault 2: Reduced feed rate. The analyzed time series are the noisy measurements of Penicillin concentration and consists of 5000 samples spaced uniformly over the span of 400 hours, which is the batch length for all batches. Note that this fixed batch length is not a restriction of the method as the scenario likelihoods can easily be computed for different values of m.

III. RESULTS

A. Data description

In Figure 1 the noisy concentration measurements are plotted as a function of sampling index. It can be seen that the qualitative description of the profiles is quite distinct for each group. This suggests that a qualitative trend analysis may suffice to properly identify the different batch conditions without relying on a complex mathematical model. In particular, the identified sequences are BC, BCBC, and BCDA respectively.

B. Model setup

For the first step of the algorithm, one needs to select (1) an order for the polynomial, (2) a kernel function (Tri-Cube kernel) and (3) a kernel width, λ. The order and kernel were already selected to be two (quadratic polynomial) and the tri-cube kernel. The results shown below are obtained with a kernel width, λ, of 256 which was set through trial and error. In order to execute the second step of the proposed algorithm for qualitative trend analysis, three different Markov chains are set up. These correspond to the sequences BC, BCBC, and BCDA respectively. The Markov transition matrices are
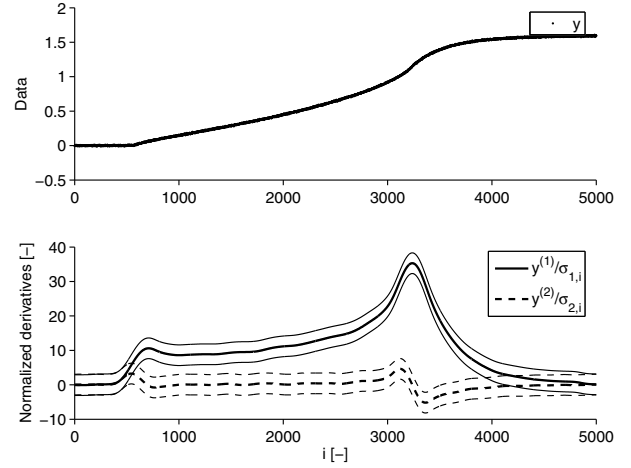
set up as follows:

$$T_{BC} = \begin{bmatrix} 1-\gamma & 0 \\ \gamma & 1 \end{bmatrix} \quad (12)$$

$$T_{BCBC} = T_{BCDA} \quad (13)$$

$$= \begin{bmatrix} 1-\gamma & 0 & 0 & 0 \\ \gamma & 1-\gamma & 0 & 0 \\ 0 & \gamma & 1-\gamma & 0 \\ 0 & 0 & \gamma & 1-\gamma \end{bmatrix}$$

Note that even though the transition matrices for the BCBC and BCDA sequences are the same, the states are not. Indeed, for the BCBC sequence state 3 and 4 correspond to B and C primitives while for the BCDA sequence the same states correspond to D and A primitives. The value for γ is 0.01 in all reported results. It is noted that the values for λ and γ were tuned by trial and error to obtain the best possible results so far.

C. Demonstration

The proposed method for qualitative trend analysis is demonstrated with one batch where fault scenario 2 is present. The sampled data is shown in the top panel of Figure 2. Visual inspection suggests that the corresponding BCBC shape is indeed most plausible for this time series. In the first step of the algorithm, one obtains a local quadratic polynomial fit based on kernel regression. Based on this fit, one can evaluate the first and second derivatives as well as their point-wise standard deviation. For vizualization, these derivatives are normalized with these standard deviations in the bottom panel of Figure 2. The first derivative is positive over the majority of the time series length. The second derivative fluctuates with peaking positive, negative, positive, and negative values (in this order).

The obtained estimates and standard deviations are used to compute the probabilities of the qualitative states, $P(s)_i$. Figure 3 shows these probabilities as a function of sampling

time. As one might expect, one observes alternating periods in which the highest probabilities correspond to B, C, B, and C primitives.

In Step 2 of the algorithm, the individual point-wise probabilities are combined to obtain the maximum likelihood for a given qualitative sequence as well as the corresponding transition times. The resulting qualitative representations for each of the considered sequences (BC, BCBC, BCDA) are vizualized in Figure 4. In each of the considered sequences, one identifies a BC transition at the same time, which is at sample 3237. This corresponds to the inflection point in the data series which is the easiest to identify based on visual inspection. In the case of the BCBC sequence, two additional inflection points are found at samples 695 and 992. This is also likely based on visual inspection. In the case of the BCDA sequence, a maximum and inflection point are identified at sample 4998 and 4999 respectively. This suggests that the BCDA sequence is rather unlikely given that these are the very last sample times where one could possibly locate these transitions.

The corresponding log-likelihoods, $L_m(n)$, are shown in Figure 5. The BCBC sequence results in the largest likelihood. Since the prior probabilities for the sequences were all set equal, the fault diagnosis task ends with the assignment of this batch to scenario 2, which is correct.

D. Fault diagnosis performance

The same exercise was repeated for all 30 batches. Figure 6 shows the likelihoods for each sequence as a function of batch index. One can see that for batch 1 to 10 (scenario 1), the BC sequence is found most likely. Similarly, for batch 11 to 20 (scenario 2) one finds the BCBC sequence most likely and for batch 21 to 30 one finds the BCDA sequence most likely. As a consequence, this means that a fault diagnosis strategy based on maximum likelihood leads to excellent performance.
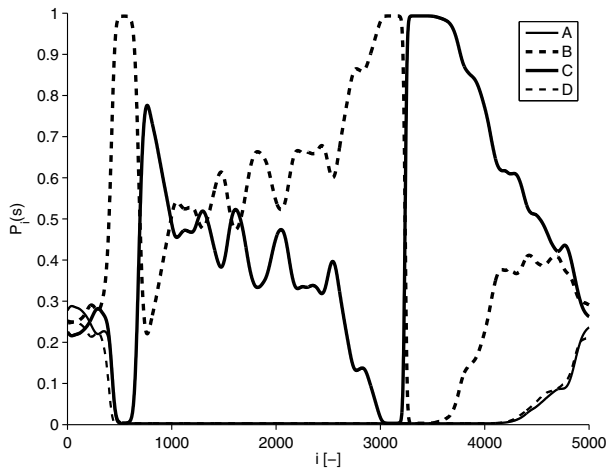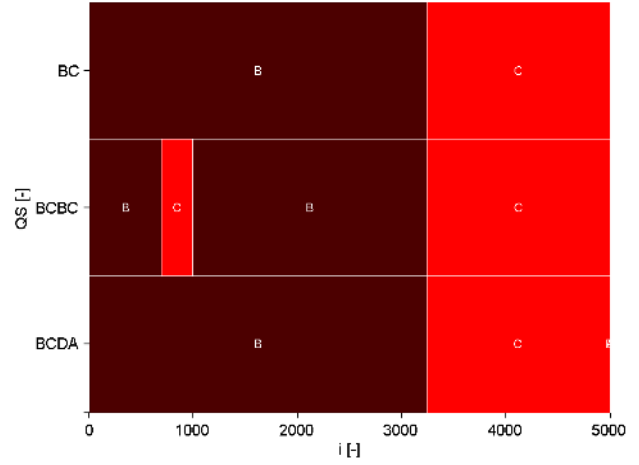

Fig. 4. Qualitative representations (QRs) resulting from applying the qualitative trend analysis algorithm using three different HMMs corresponding to BC, BCBC, and BCDA sequences.

The above fault diagnosis performance is the same as obtained with the deterministic global optimization approach as taken in [15]. Importantly, this can be attributed in part to the fact that two parameters were tuned to obtain these results. Indeed, one does not necessarily obtain the same performance with different values for λ and γ. In contrast, the original method in [14], [15] does not require such tuning.

The most obvious advantage of the newly proposed method lies in the computational speed. Figure 7 shows the time need for each qualitative sequence as a function of batch index on a Pentium IV (3 GHz, 1 GB RAM) computer. One can see that this time is lower than 10 seconds for every sequence and batch and remains relatively constant across the set of batches. This stands in clear contrast with the method in [14] where computational demand for the same task was

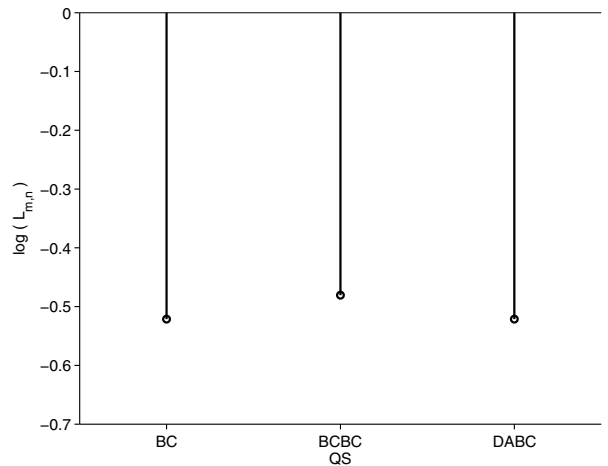Fig. 3. Probability of each primitive based on the normal distribution of the 1st and 2nd derivative.

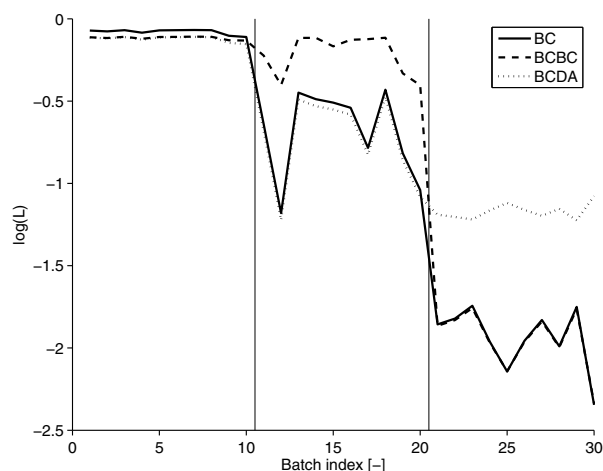Fig. 5. Maximum a posteriori (MAP) likelihoods for each of the sequences (BC, BCBC, BCDA)

Fig. 6. A posteriori likelihoods for each sequence as function of batch index.

reported to be as high as 20 hours and highly dependent on the actual batch scenario.

IV. CONCLUSIONS

In this contribution, a novel method for qualitative trend analysis is proposed. The method is based on a two-step algorithm in which one first assigns probabilities for different qualitative states and then obtains a maximum likelihood sequence of qualitative states based on a Markov chain representation and by using the Viterbi algorithm. Following this, the likelihood value for different sequences can be used to determine the most likely sequence and associated batch process condition. Results indicate that, pending proper tuning, one can obtain excellent fault diagnosis performance as was obtained in previous studies based on a global optimization approach. However, with the proposed method
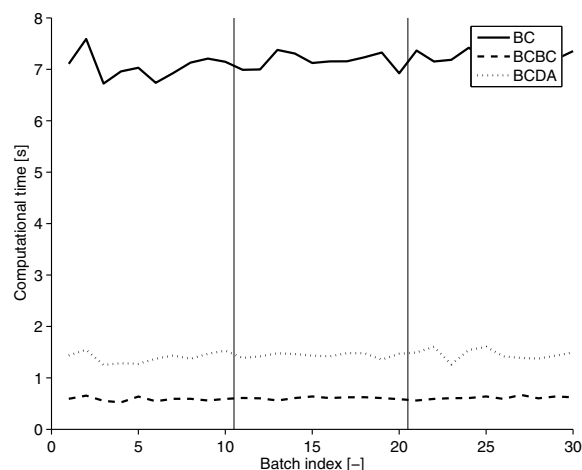

Fig. 7. Computational time needed on a Pentium IV (3 GHz, 1 GB RAM) computer for each sequence as function of batch index.

computational time is reduced to the point where fast, on-line implementation is a realistic perspective. Further research will be oriented at determining (1) the sensitivity of the method to the choice of kernel function, parameters of the kernel function and the Markov model and measurement noise levels as well as (2) obtaining an on-line version of the proposed algorithm. Importantly, the HMMs are considered a given and the considered primitives exclude those with zero derivatives. This may not alway be realistic and for this reason future research will also focus on (1) the identification of the structure and parameters of such HMM models and (2) use of a more expressive set of primitives.

REFERENCES

[1] B. R. Bakshi and G. Stephanopoulos. Representation of process trends - Part III. multiscale extraction of trends from process data. *Comput. Chem. Eng.*, 18:267–302, 1994.
[2] B. Balasko, S. nemeth, and J. Abonyi. Qualitative analysis of segmented time-series by sequence alignment. In *HUCI2006*, 2006. Appeared on CD-ROM.
[3] G. Birol, C. Undey, and A. Çinar. A modular simulation package for fed-batch fermentation: penicillin production. *Comput. Chem. Eng.*, 26:1553–1565, 2002.
[4] S. Charbonnier, C. Garcia-Beltan, C. Cadet, and S. Gentil. Trends extraction and analysis for complex system monitoring and decision support. *Eng. Appl. Artif. Intell.*, 18:21–36, 2005.
[5] J. Colomer, J. Melendez, and F. Gamero. Qualitative representation of process trends for situation assessment based on cases. In *IFAC World Congress 2002, Barcelona, Spain.*, 2002.
[6] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning. Data Mining, Inference, and Prediction.* Springer, NY, USA, 2001.
[7] B. J. Kuipers. *Qualitative Reasoning: Modeling and simulation with incomplete knowledge.* MIT Press, Cambridge, MA, USA, 1994.
[8] M. R. Maurya, P. K. Paritosh, R. Rengaswamy, and V. Venkatasubramanian. A framework for on-line trend extraction and fault diagnosis. *Engineering Applications of Artificial Intelligence*, 23:950–960, 2010.
[9] M. R. Maurya, R. Rengaswamy, and V. Venkatasubramanian. Application of signed digraphs-based analysis for fault diagnosis of chemical process flowsheets. *Eng. Appl. Artif. Intell.*, 17:501–518, 2004.
[10] M. R. Maurya, R. Rengaswamy, and V. Venkatasubramanian. Fault diagnosis by qualitative trend analysis of the principal components. *Comput. Chem. Eng.*, 83:1122–1132, 2005.
[11] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.
[12] R. Rengaswamy and V. Venkatasubramanian. A syntactic pattern-recognition approach for process monitoring and fault diagnosis. *Eng. Appl. Artif. Intell.*, 8:35–51, 1995.
[13] M. Rubio, Colomer J., M. L. Ruiz, Colprim J., and J. Mélendez. Qualitative trends for situations assessment in SBR wastewater treatment process. In *Proceedings of the 4th ECAI Workshop on Binding Environmental Sciences and Artificial Intelligence (BESAI), August 2004, Valencia, Spain.*, 2004.
[14] K. Villez, R. Rengaswamy, and V. Venkatasubramanian. Generalized shape constrained spline fitting for qualitative analysis of trends. Submitted:In review, 2012.
[15] K. Villez, L. Rieger, B. Keser, and V. Venkatasubramanian. Probabilistic qualitative analysis for fault detection and identification of an on-line phosphate analyzer. *International Journal of Advances in Engineering Sciences and Applied Mathematics*, 4:6777, 2012.
[16] K. Villez, C. Rosén, F. Anctil, C. Duchesne, and P. A. Vanrolleghem. Qualitative representation of trends: an alternative approach to process diagnosis and control. *Wat. Sci. Technol.*, 57(10):1525–1532, 2008.
[17] K. Villez, C. Rosén, F. Anctil, C. Duchesne, and P. A. Vanrolleghem. Qualitative representation of trends (QRT): Extended method for identification of consecutive inflection points. *Comp. Chem. Eng.*, 48:187–199, 2012.