

A COMBINED QUALITATIVE/QUANTITATIVE APPROACH FOR FAULT ISOLATION IN CONTINUOUS DYNAMIC SYSTEMS

Eric-J. Manders ^{*,1} Sriram Narasimhan ^{*,2} Gautam Biswas ^{*}
Pieter J. Mosterman ^{**}

^{*} Department of Electrical Engineering and Computer Science
Vanderbilt University, P.O. Box 1679 Sta. B, Nashville, TN 37235

Email: {manders,nsriram,biswas}@vuse.vanderbilt.edu

^{**} Institute of Robotics and Mechatronics DLR Oberpfaffenhofen, P.O.

Box 1116, D-82230 Wessling, Germany

Email: Pieter.J.Mosterman@dlr.de

Abstract: The TRANSCEND system for fault detection and isolation in continuous systems uses qualitative reasoning methods to analyze transients caused by abrupt faults. Qualitative transient analysis avoids some of the computational difficulties associated with numerical schemes, but they lack discriminating power. This paper presents the formal basis for qualitative transient analysis, and then establishes the limits of the discriminatory power of this methodology. An integrated scheme that starts with qualitative fault isolation to narrow down possible fault hypotheses, and then uses a focused quantitative parameter estimation scheme to identify the true fault is developed. This approach provides a number of advantages over purely quantitative FDI schemes. *Copyright © 2000 IFAC*

Keywords: Fault detection and isolation, monitoring, qualitative analysis, transient analysis, parameter estimation.

1. INTRODUCTION

Model-based approaches for fault detection and isolation (FDI) in continuous dynamic systems employ relations imposed by the system configuration and functionality to compute *residuals*, that capture the discrepancies between nominal and observed behavior. Residual computation and analysis is non-trivial for complex systems, primarily because of stiffness, convergence, and intractability problems in dealing with the system's non-linear dynamics. To mitigate this, an FDI framework has been developed that derives residuals as *qualitative fault signatures*, and analyzes these residuals with a fault isolation observer mechanism based on a unique *progressive monitoring* scheme (Mosterman and Biswas 1999). This paper

demonstrates a combined qualitative and quantitative fault isolation process where the qualitative fault isolation scheme is followed by a directed quantitative parameter estimation step to resolve ambiguities among the fault candidates that cannot be distinguished by the qualitative signatures alone.

Fig. 1 illustrates the architecture of TRANSCEND, our qualitative model-based approach to diagnosis (Mosterman and Biswas 1999). Variables u , x , and y , are the input, state, and output vectors of the process under diagnostic scrutiny. A standard gain matrix observer scheme (Brammer and Siffling 1989) tracks the residual, $r = y - \hat{y}$ (\hat{y} is the predicted system behavior) to correct for small deviations in the estimated state vector \hat{x} . A unique aspect of the qualitative approach is the *symbol generation* unit, that uses robust methods to compute symbolic values of the magnitude deviation and slope of signal transients, r_s .

¹ Partially supported by a grant from Agilent Laboratories

² Partially supported by a grant from Xerox PARC

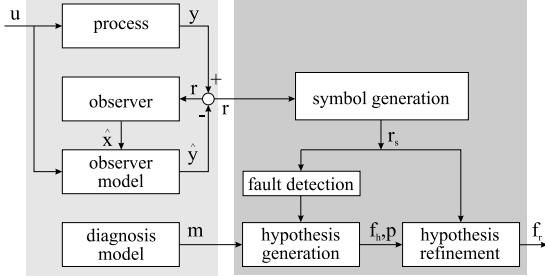


Fig. 1. TRANSCEND architecture.

Fault detection triggers a fault isolation scheme that consists of *hypothesis generation* and *hypothesis refinement*. Hypothesis generation uses the diagnosis model, m (implemented as a temporal causal graph (TCG) (Mosterman and Biswas 1999)), and the symbolic residuals, r_s , to generate a set of hypothesized fault candidates, f_h , and to predict behavior, p , for each fault candidate. During hypothesis refinement spurious candidates are eliminated from the set using progressive monitoring to match new observations against the predictions and derive the refined fault set, f_r .

This paper establishes the basis for qualitative signatures and progressive monitoring in terms of the Taylor series expansion of the fault transient signal. Based on this, the limitations of the qualitative fault isolation method are derived, and a focused parameter estimation scheme is introduced to enable further refinement of the fault hypotheses. The reduced fault set obtained by applying the qualitative observers focuses and significantly reduces the computational complexity in the parameter estimation task. The method is illustrated with simulation experiments on a three-tank fluid system.

2. QUALITATIVE DIAGNOSIS FROM TRANSIENTS

Model parameters in the TCG correspond directly to system components, and a fault is a parameter value that deviates from its nominal value. This paper focuses on the class of *abrupt faults*.

Definition 1. (Abrupt fault). An instantaneous and persistent change in a parameter value.

The notion of instantaneous change is a modeling abstraction (e.g., see (Mosterman and Biswas 1998)), where the fault is defined as an instantaneous parameter change in the model. In a physical system this change is never truly instantaneous but the abstraction eliminates the steep nonlinearities and stiffness that occurs when simulating the behavior of such systems. An abrupt fault results in transient behaviors in system variables and the fault isolation task relies heavily on the characterization of the transients in the measurement data (Mosterman and Biswas 1999).

2.1 Transient detection and analysis

Transient detection implies a decision on whether the residual r is deviating significantly from 0. The need for sophisticated detection techniques is strongly dependent of the signal-to-noise ratio of the residual signal. In simulation studies, a detection scheme based on an instantaneous signal value has been used by applying simple threshold crossing detectors, for experiments where a small amount of noise is added to the simulation data. For signals where the noise presents a larger problem, more sophisticated methods have been studied that employ statistical signal processing techniques and can be designed to obtain desired sensitivity and specificity (Manders *et al.* 1999).

The FDI analysis in TRANSCEND assumes that discontinuous changes in variable values can only occur at the point of failure, thus system behavior is continuously differentiable before and after the occurrence of a fault. Therefore, the transient response in a measurement after the time point of failure, t_0 , can be approximated by the Taylor series expansion. If $y(t_0)$ is the value of the residual signal just after the occurrence of the fault, the k^{th} order Taylor series expansion for $y(t)$, $t \geq t_0$ is defined as:

$$y(t) = y(t_0) + y'(t_0) \frac{(t - t_0)}{1!} + y''(t_0) \frac{(t - t_0)^2}{2!} + \dots + y^{(k)}(t_0) \frac{(t - t_0)^k}{k!} + R_k(t),$$

where $R_k(t) = y^{(k+1)}(t') \frac{(t - t_0)^{k+1}}{(k+1)!}$ is a remainder term and $t' - t_0 < t$. For most well-behaved functions the series converges, i.e., $R_k(t) \rightarrow 0$ as $k \rightarrow \infty$ (Kreyszig 1972). In particular, if $|y^{(k+1)}|$ is bounded, the Taylor series is a good approximation of the true signal $y(t)$ when t is close to t_0 .

Consider the transient signal and its first through fourth order Taylor series approximation shown in Fig. 2. Conforming to the definition, as t increases from t_0 , approximations differ increasingly from the actual signal, but higher order approximations follow the signal for a longer time interval. The analysis of transient dynamics by interpreting the signal as a Taylor series approximation is the basis for describing the fault transient signal as a *fault signature*.

Definition 1. (Fault signature). The fault signature f is the set of $k + 1$ feature values consisting of the

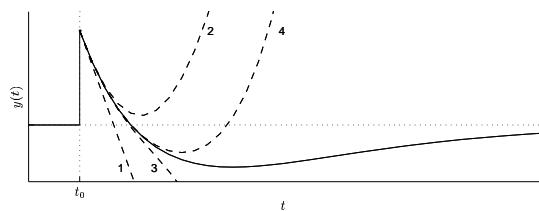


Fig. 2. Transient signal for a 2^{nd} order system (solid line) and 1^{st} through 4^{th} order Taylor series expansion (dashed lines) at $t = t_0^+$.

magnitude and the 1st through k^{th} order derivative values computed at t_0 from the residual y : $f = \{y(t_0), y'(t_0), y''(t_0), \dots, y^{(k)}(t_0)\}$

The notion of discontinuous change warrants additional explanation. Similar to the concept of an abrupt change in a model parameter, a discontinuous change is an abstraction in the observation of a physical system. What is considered as a discontinuous change is directly related to the sampling rate of the discrete time sampled signal.

Definition 2. (Discontinuity). A change in a measured variable that exhibits a transient response faster than the sampling rate of the signals.

Fault isolation is then based on the comparison of the fault signatures with subsequent measurements made on the system. Performing this analysis quantitatively is an intractable problem. When a fault occurs, the exact magnitude of parameter value changes is unknown, so derivative values in the fault signature have to be computed from subsequent measurements. For complex nonlinear systems, this is a very difficult problem to solve, either by closed form analytic techniques or by numeric iteration. To address this a *qualitative constraint analysis* scheme, discussed in the next section, was developed for the fault isolation task.

In the qualitative framework, individual measurements are labeled as normal (0), above normal (+) and below normal (−). Similarly, derivatives take on values, increasing (+), steady (0), and decreasing (−). The fault signature in the qualitative framework then is the sequence of +, 0, or − – magnitude and k derivative values computed at the point of failure, t_0 . This fault signature is the basis for qualitative transient analysis using the progressive monitoring scheme.

Lemma 1. (Qualitative transient analysis). Transient dynamics are captured by evaluation of the direction of *abrupt change* at the point of failure (if it occurs), and the *signs* of the derivatives of the signal after the onset of a fault.

In this paper, all transient signals are considered idealized signals. This means that data is sampled at appropriate sampling rates and that noise does not play a role in determining the component values of a signature. Other work (Manders *et al.* 2000) describes statistical signal analysis algorithms for extracting transient features from noisy signals.

2.2 Fault Isolation

Fault isolation using residuals is traditionally achieved by designing multiple fault observers with a one-to-one correspondence between the individual fault

hypotheses and the observers (Patton and Chen 1997). In our work, we define each observer in terms of the *fault signatures*.

Fault detection triggers the fault isolation mechanism. The hypothesis generation algorithms, implemented as a two step process, fault hypothesis generation followed by fault signature generation for each hypothesis, is described in detail in (Mosterman and Biswas 1999). An observer, defined in terms of a set of fault signatures, one for each measurement, is designed for each fault hypothesis.

The fault signature is of order k when it includes derivative values of up to order k . The minimal practical fault signature consists of magnitude and first order derivative, the slope, of the signal. The choice of k is directly related to the concept of diagnosability (Mosterman and Biswas 1999). Ideally, the fault signature order and the set of measured variables are selected such that that all possible faults that can be hypothesized by the model can also be uniquely determined.

Comparing the fault signature with the feature vector obtained from the evolving transient data is the basis of a *progressive monitoring* scheme for tracking signal transients (Mosterman and Biswas 1999).

Lemma 2. (Progressive monitoring). Qualitative magnitude and slope of a fault transient are matched against a qualitative fault signature by starting in a sequence from a discontinuous magnitude and first order change to a succession of higher order derivatives.

Comparing the i^{th} and $(i+1)^{th}$ terms in the Taylor series, one can establish $|y^{(i)}(t_0)| > |y^{(i+1)}(t_0)| \frac{(t-t_0)}{(i+1)}$ for some period of time t . As t increases, at some point in time the inequality reverses. From that point in time, the higher order derivative dominates the lower one.

Lemma 2 provides the basis for progressive monitoring of signal dynamics using higher order derivatives. Starting from the point of failure, t_0 , the signal magnitude in response to the fault $y(t_0)$ determines the signal value. Immediately after that the first derivative of the signal dominates the dynamic behavior because small values of $(t - t_0)$ dominate higher powers $(t - t_0)^i$ in the Taylor series. As t increases, higher order derivatives in succession increasingly contribute to the dynamics of the signal.

In dynamic transient analysis, a current normal measurement or slope value cannot be used to eliminate a fault candidate, because, there is no guarantee that this measurement will not deviate at some later time. The exception to this is the case when a discontinuous change can be inferred from the signature, because any discontinuous change in the measurements should manifest itself at the point of failure. Therefore the ability to reliably detect discontinuous changes in the measurement data enhances fault isolation.

Qualitative transient analysis based on the progressive monitoring scheme established in Lemma 2 becomes the basis for tracking system behavior and eliminating inconsistent fault hypotheses till the true fault is isolated.

As an example, the methodology is applied to a three-tank fluid system in a simulation. The tank system and corresponding TCG model are shown in Fig. 3. The tank capacities, C_1 , C_2 , and C_3 , and pipe resistances, R_{12} , R_{23} , and R_b , constitute the set of model parameters of the physical components and thus the possible fault candidates. The e_i and f_i vertices shown in the TCG correspond to tank pressure and pipe flow rate variables, respectively. Circled vertices in the TCG indicate the measured variables. In this example the pressure in the third tank (e_{10}) and the flow rate in the pipe connecting tank 1 and tank 2 (f_3) are measured.

Consider the fault situation where the capacity of tank 1 decreases abruptly (indicated with C_1^-), as might happen when an object falls into the tank. The resulting transients in the measurements are shown in Fig. 4(a). The initial deviation that triggers the isolation scheme is an abrupt increase in f_3 , indicated by f_3^+ . Fault detection and hypothesis generation based on this deviation produces the set of fault candidates with predicted fault signatures for both measured variables, as shown in Fig. 4(b). This figure lists the complete qualitative fault isolation sequence. Step 0 is defined as the initial fault isolation step. Abrupt change detection allows the elimination of three of the six initial fault hypotheses in the first step. At step 1, e_{10} and f_3 are reported as $(0, 0)$ and $(+, -)$, respectively. All three fault hypotheses are still consistent with the observations. At step 2, e_{10} crosses the threshold and is reported to be $(+, +)$. Application of progressive monitoring based on Lemma 2 results in fault candidates C_1^- and R_{12}^- still remaining consistent with the observations but C_2^+ does not. Further refinement of the fault hypotheses is not possible with the given measurements.

2.3 Analysis of the Qualitative Fault Isolation Scheme

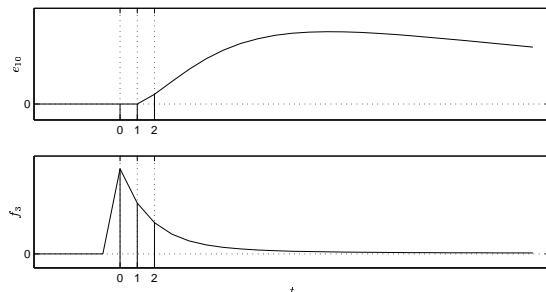
The Taylor series expansion example above illustrates the two primary assumptions of this scheme:

- (1) the signal sampling rate is fast enough to pick up all significant qualitative changes in signal magnitudes and slopes, and
- (2) abrupt changes in signal magnitudes can be reliably detected.

A fault signature of order k results on 2^k distinct signatures. Combining this with the above assumptions and Lemma 2, one may come to the conclusion that for a system with K possible faults, complete diagnosability can be achieved with $\lceil \log_2 K \rceil$ measurements. However, careful analysis of the progressive monitoring framework reveals that this is not the case. As a

first step, the occurrence of an abrupt fault can have one of three distinct effects on a measured signal: (i) an observed positive discontinuity in the signal, (ii) an observed negative discontinuity in the signal, and (iii) no discontinuity observed in the measured signal. The example for the three-tank system discussed above shows that reliable discontinuity detection plays a major role in pruning the fault hypothesis. Following the abrupt change, the following observation patterns can occur: (a) $(+, +)$, (b) $(+, -)$, (c) $(-, -)$, and (d) $(-, +)$, implying that there are at least four distinct fault signatures that can be recognized after an abrupt change.

Case (iii) above implies the presence of integrating energy storage elements in the direct path from the component parameter to the measurement signal in the TCG. Depending on the number of such integrating elements in the path, a corresponding number of derivative terms (1^{st} , 2^{nd} , \dots) may be 0 in the fault signature. Applying progressive monitoring, the initial direction of change for this signal will be the first non-zero term in the fault signature. Given the assumption that the transient is appropriately sampled, the observed signal will necessarily exhibit a $(+, +)$ or a $(-, -)$ pattern. Therefore, initially, there are only two distinct fault signatures associated with a measured signal that does not undergo a discontinuous change in response to an abrupt fault. The limitations of the strictly qualitative analysis can now be summarized in the following lemma.



(a) Transient data for measurements e_{10} (top) and f_3 , with TRANSCEND diagnosis steps indicated.

Step 0	Step 1	Step 2
actual $e_{10}:$ 0 · $f_3:$ + -	actual $e_{10}:$ 0 · $f_3:$ + -	actual $e_{10}:$ + · $f_3:$ + -
$R_{12}^+ e_{10}:$ 0 0 + - $f_3:$ + - + -	$R_{12}^- e_{10}:$ 0 0 + - $f_3:$ + - + -	$R_{12}^- e_{10}:$ 0 0 + - $f_3:$ + - + -
$C_2^+ e_{10}:$ 0 - + - $f_3:$ + - + -	$C_2^+ e_{10}:$ 0 - + - $f_3:$ + - + -	$C_1^- e_{10}:$ 0 0 + - $f_3:$ + - + -
$R_{23}^- e_{10}:$ 0 + - + $f_3:$ 0 + - +		
$C_3^+ e_{10}:$ - + - + $f_3:$ 0 + - +		
$R_b^- e_{10}:$ 0 - + - $f_3:$ 0 0 + -		
$C_1^- e_{10}:$ 0 0 + - $f_3:$ + - + -		

(b) Diagnosis results.

Fig. 4. Fault detection and isolation for a fault C_1^- .

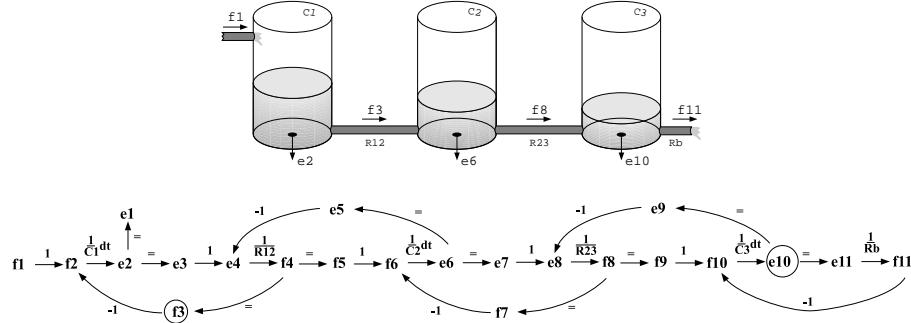


Fig. 3. Three-tank fluid system and its TCG model representation.

Lemma 3. (Discriminatory power). In a purely qualitative framework, only the following characteristics of a signal can be used to discriminate among faults:

- (1) if there is an abrupt change, the direction of abrupt change plus the direction of change immediately following the abrupt change. This implies that there are four distinct fault signatures.
- (2) if there is no abrupt change, the first direction of change in the signal. Therefore, this case has two unique fault signatures.

For the case where the signal does not undergo an abrupt change, higher order derivatives beyond the first non zero derivative have no discriminatory power. Consider two faults with second order signatures $<0, +, +>$ and $<0, +, ->$, respectively for a particular measurement. In both cases, the signal shows no discontinuous change at the point of failure, and subsequently matches a $<+, +, \cdot>$ signature, where \cdot can be a $+$ or $-$. Subsequently, even if the signal slope is measured to be $-$, the $<+, +, +>$ cannot be eliminated, because a higher order derivative effect that is not captured in the second order signature could be $-$. This problem can only be overcome by modeling more quantitative information about the signal time constants, and thus the system parameters.

3. PARAMETER ESTIMATION

Qualitative methods for diagnosis are robust in that they apply in uncertain environments, and avoid the computational difficulties associated with the stiffness and convergence problems of numerical schemes. However, for the reasons discussed in Sec. 2.3 the inability to incorporate time constant information adversely affects their ability to discriminate faults that show no qualitative differences, or differ only in higher order transient effects. Fig. 4 shows that the qualitative fault isolation scheme is unable to distinguish between fault hypotheses, C_1^- and R_{12}^- for the three-tank system.

To achieve higher resolution, a quantitative analysis approach, illustrated in Fig. 5 is introduced into the fault isolation scheme. The idea is to express the transient behavior as a function of the hypothesized fault

parameters derived from f_r , and estimate the value of these parameters from the available measurements.

The bond graph modeling approach (Rosenberg and Karnopp 1983) is the starting point for deriving both the state equation based observer model and the TCG based diagnosis model. This paradigm provides a direct mapping from physical component parameters to the standard state equation form:

$$\dot{x} = Ax + Bu \quad (1)$$

$$y = Cx + Du. \quad (2)$$

The quantitative parameter estimation scheme is implemented by expressing the coefficients of the matrices, A , B , C , and D , in terms of the single parameter corresponding to the hypothesized fault, and using the nominal (known) values for all other component parameter values. Like the qualitative fault observers, a separate parameter estimator is initiated for each fault hypothesis in f_r . Given the observation vector $y(t)$ ($t \geq t_0$), a standard least squares estimation method is applied to derive the fault parameter values. Fault parameters for which the error term, i.e., the difference between the predicted (\hat{y}) and observed measurements (y), do not converge to zero are eliminated. The decision test for convergence to zero is implemented as a statistical hypothesis testing scheme. Predicted measurement values are computed from the estimated fault parameter values. The parameter estimation is executed only for the coefficients of the state matrices in which the fault parameters appear. As a result, the form of the nonlinear functions used for the estimation task are simplified, which in turn reduces the complexity of the least squares estimation, and numerical convergence is easier to achieve.

The state vector x for the three-tank system is defined by the pressure variables in the three tanks, i.e., $x = [e_2 \ e_6 \ e_{10}]^T$. The input vector, $u = [f_1]$ and the output vector y in the example are the measured variables, $[e_{10} \ f_3]^T$. The symbolic form of the matrices A , B ,

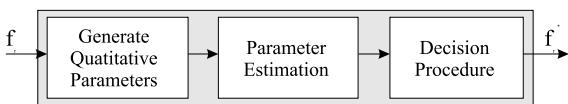


Fig. 5. Extending fault isolation with quantitative parameter estimation methods.

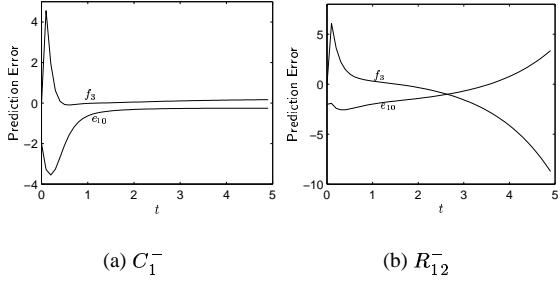


Fig. 6. Parameter estimation for remaining candidates after qualitative fault isolation

C and D , derived from the bond graph model as a function of the component parameters are:

$$A = \begin{bmatrix} -\frac{1}{R_{12}C_1} & \frac{1}{R_{12}C_1} & 0 \\ \frac{1}{R_{12}C_2} & \frac{1}{C_2} \left(\frac{1}{R_{23}} - \frac{1}{R_{12}} \right) & -\frac{1}{R_{23}C_2} \\ 0 & \frac{1}{R_{23}C_3} & -\frac{1}{C_3} \left(\frac{1}{R_{23}} + \frac{1}{R_b} \right) \end{bmatrix},$$

$$B = \begin{bmatrix} \frac{1}{C_1} \\ 0 \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{R_{12}} & -\frac{1}{R_{12}} & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Consider fault hypothesis, C_1^- . Inspection of the A matrix shows that this parameter appears only in two of the matrix coefficients, a_{00} and a_{01} . This simplifies the parameter estimation task for this fault hypothesis, and system identification techniques have to be applied to derive the new values for a_{00} and a_{01} .

In the simulation experiments, the normal values for all parameters are set to 1. Two sets of parameterized matrices, A , B , C , and D , are constructed for the two fault hypotheses, C_1^- , and R_{12}^- . The transient response for measurements e_{10} and f_3 (illustrated in Fig. 4(a)), are used to compute the numeric values of C_1 and R_{12} . The derived values are then used to predict the values for the same measurements, and the prediction error, $e = \hat{y} - y$, is plotted in Fig. 6. For predictions with parameter, C_1 , the error e converges toward 0. However, for prediction with parameter R_{12} , the error e diverges for both measurements, indicating that C_1^- is the true fault. Currently, the decision procedure is implemented as a statistical hypothesis testing algorithm that checks for the convergence of e to 0 at a predetermined confidence level.

4. DISCUSSION AND CONCLUSIONS

This paper presents a systematic analysis of an approach to FDI that combines qualitative and quantitative analysis for robust fault isolation. The Taylor series expansion of transient signals provides the basis for the construction of qualitative fault signatures and the progressive monitoring scheme for tracking fault

transients. The limits of the discriminatory ability of the qualitative scheme could then be established based on a formal analysis. To improve the isolation task, a focused parameter estimation method is developed that works in conjunction with the qualitative scheme to enable isolation of the true fault candidate. This methodology allows the state equations of the system to be parameterized in terms of the hypothesized fault parameters, and in the process creates a simpler formulation for quantitative analysis. This mitigates a number of computational problems that arise with traditional numeric schemes. Simulation experiments conducted on a three-tank fluid system demonstrate the effectiveness of the methodology.

In future work, the quantitative analysis will be applied to more complex systems, such as the automobile engine test bed used in previous experiments (Manders et al. 2000). The parameter estimation problem for nonlinear dynamic systems must also be addressed, and the challenge there will be to derive simplified parameterized input-output representations (cf. (Zhang et al. 1998)) from the state equations for the parameter estimation task.

5. REFERENCES

Brammer, K. and G. Siffling (1989). Kalman-Bucy Filters. Artec House, Norwood MA.

Kreyszig, E. (1972). Advanced Engineering Mathematics, Third Ed. John Wiley, New York.

Manders, E.-J., G. Biswas, P.J. Mosterman, L.A. Barford and R.J. Barnett (2000). Signal interpretation for monitoring and diagnosis, a cooling system testbed. *IEEE Trans. on Instrumentation and Measurement*. To Appear.

Manders, E.-J., P.J. Mosterman and G. Biswas (1999). Signal to symbol transformation techniques for robust diagnosis in TRANSCEND. In: Tenth International Workshop on Principles of Diagnosis. Loch Awe, Scotland. pp. 155–165.

Mosterman, P. J. and G. Biswas (1998). A theory of discontinuities in physical system models. *Journal of the Franklin Institute: Engineering and Applied Mathematics* **335B**(3), 401–439.

Mosterman, P.J. and G. Biswas (1999). Diagnosis of continuous valued systems in transient operating regions. *IEEE Trans. on Systems, Man and Cybernetics* **29**(6), 554–565.

Patton, R.J. and J. Chen (1997). Observer-based fault detection and isolation: Robustness and applications.. *Control Engineering Practice* **5**(5), 671–682.

Rosenberg, R. C. and D. Karnopp (1983). Introduction to Physical System Dynamics. McGraw-Hill Publishing Company. New York, New York.

Zhang, Q., M. Basseville and A. Benveniste (1998). Fault detection and isolation in nonlinear dynamic systems: A combined input-output and local approach. *Automatica* **38**(11), 1421–1429.