Diagnosis of Physical Systems With Hybrid Models Using Parametrized Causality

Pieter J. Mosterman*

Institute of Robotics and Mechatronics, DLR Oberpfaffenhofen
P.O. Box 1116, D-82230 Wessling, Germany
Pieter.J.Mosterman@dlr.de
http://www.op.dlr.de/~pjm

Abstract. Efficient algorithms exist for fault detection and isolation of physical systems based on functional redundancy. In a qualitative approach, this redundancy can be captured by a temporal causal graph (TCG), a directed graph that may include temporal information. However, in a detailed continuous model, time constants may be present that are beyond the bandwidth of the data acquisition system, which leads to incorrect fault isolation because of a difference in observed and modeled behavior. To solve this, the modeled time constants can be taken to be infinitely small, which results in a model with mixed continuous/discrete, *hybrid* behavior that is difficult to analyze because the causality of the directed graph may change. In this paper, to avoid the combinatorial explosion when using a bank of TCGs in parallel, causal paths are parametrized by the state of local switches. The result is a hybrid model that produces parametrized predictions that can be efficiently matched against observed behavior.

1 Introduction

To reduce cost, improve performance, and to manage the complexity of large engineered systems, functional redundancy can be employed in fault detection and isolation (FDI). In this approach, a system model links measured variables by their functional relations, facilitating the computation of redundant values for selected system variables. In general, the system model can be of a continuous or discrete nature. In case of a continuous model, often parameter and state estimation techniques based on a state space model of the system are used for FDI [1, 4]. In case of a discrete event approach, models that capture failure modes and transition sequences are applied [5, 15, 16]. Both these methods have proven themselves successful in their respective applications.

Previous work [8, 9] has focused on qualitative parameter estimation of continuous system models. These models are represented by a temporal causal graph (TCG) that is automatically derived from a bond graph model of a physical system [8, 11]. This work revealed the importance to design the model in harmony

* Pieter J. Mosterman is supported by a grant from the DFG Schwerpunktprogramm KONDISK.

with the data acquisition system, i.e., behavior that is beyond the bandwidth of the data acquisition system should not be included in the model as it leads to incorrect fault isolation [2].

Removing large and small parameters from the system model causes the following model characteristics that complicate the FDI task:

- Algebraic loops may emerge. Because of the passive behavior of physical processes, these algebraic loops have negative gain, and, therefore, any qualitative $\pm$ deviation is reversed when propagated around the loop. This, in turn, leads to many unknown values of system variables in a qualitative sense.
- In case of abrupt faults that cause mode changes, higher index systems may arise with algebraic constraints between time derivative behavior. These systems may exhibit impulsive behavior.
- The direction of the computational causality in the model may change. When abrupt faults cause component parameter changes to values that are taken to be infinitely large or small, they are effectively removed from the model, which changes the model configuration, and, in effect, the model becomes of a switched continuous, hybrid, nature.

Other work [3, 12], addresses the first two issues whereas this paper focuses on the hybrid diagnosis problem.

In order to deal with the change of causality, the TCG can be derived for each possible system configuration or *mode*. However, in case of many locally acting switches, the combinatorial explosion quickly leads to an intractable problem. These problems can be mitigated to some extent by dynamically generating the TCG of each possible system mode in response to a failure. This may still result in a problem with large computational complexity which can be further reduced by measuring system variables that indicate specifically which local switches may have occurred [13] and predictions for each of the variables that determine different causal assignments are required to be made and analyzed. Once a set of possible TCGs is available, Gaussian decision techniques have been applied to compute the most likely mode of continuous behavior [7].

Recent attention to hybrid diagnosis [7, 14] concentrates on efficiently processing a set of TCGs. This paper describes how a hybrid model can be made amenable to the diagnosis algorithms that were developed in previous work [8, 9] by systematically generating one parametrized TCG. In this graph, the directed links are enabled by conditionals that correspond to the mode in which these links are present. The result is a set of predictions that are parametrized by the state of the local switches and the diagnosis problem then becomes one of constraint satisfaction [17]. The solution to this constraint satisfaction problem contains the possible parameter changes (i.e., the faults) and the effect on the system mode that this is required to have.

2 Preliminaries

This section reviews the qualitative FDI approach developed in previous work [8, 9]. Instead of a temporal causal graph, though, the model representation format and processing will be in qualitative matrix algebra, which is easier to represent and to extend with the required notions.

Consider the one-tank hydraulic system in Fig. 1. The functional relation for flow, f_R, through the outflow pipe is given by $f_R = \frac{p_R}{R}$, where p_R is the pressure drop across the pipe and R is the pipe resistance to flow. The pressure p_R depends on the pressure at the bottom of the tank, p_C, according to $p_R = p_C$ (i.e., the ambient pressure is assumed to be 0). The rate of change in the pressure, $\dot{p}_C$, at the bottom of the tank is given by $\dot{p}_C = \frac{1}{C} f_C$, where $f_C = f_{in} - f_R$ and f_{in} is the flow into the tank and C is the tank capacity.

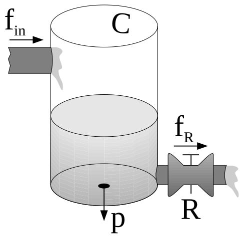

Fig. 1. A tank with in- and outflow.

To derive qualitative predictions, the system is written as a directed graph that captures the causal (directed) relations between system variables. For the one-tank system, the preferred (integral) causality model description is

$$\begin{bmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{bmatrix} \begin{bmatrix} p_C \\ f_C \\ f_R \\ p_R \end{bmatrix} = \begin{bmatrix} 0 & \lambda^{-1}C^{-1} & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & R^{-1} \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_C \\ f_C \\ f_R \\ p_R \end{bmatrix} + \begin{bmatrix} 0 \\ f_{in} \\ 0 \\ 0 \end{bmatrix} \tag{1}$$

where λ represents the time differentiation operator and λ^{-1} indicates integration over time. The corresponding temporal causal graph (TCG) is given in Fig. 2.

The TCG can be represented by a weighted adjacency matrix where the columns are cause and rows are the effect variables and the entries capture the parameters on the graph edges. This is called the temporal causal matrix (TCM), that is

$$\begin{bmatrix} 1 & \lambda^{-1}C^{-1} & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & R^{-1} \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_C \\ f_C \\ f_R \\ p_R \end{bmatrix} \tag{2}$$

for the TCG in Fig. 2.

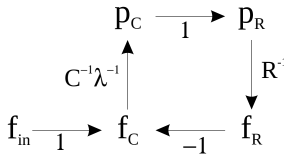

Fig. 2. TCG of the one-tank system.

Our diagnosis engine TRANSCEND [6] relies on qualitative information to achieve diagnosis. In this framework, only the three values $-$, 0, $+$ are used to indicate values that are too low, normal, and too high, with respect to some nominal value, respectively. For example, a value of a model variable that is measured to be above its nominal value is marked $+$. In case the outflow of the tank system in Fig. 1 is too high, this is represented by f_R^+.

Note that in a qualitative representation, the parameters R and C correspond to direct relations between variables, and, therefore, they can be replaced by value 1. This results in a qualitative system where 1 and -1 represent direct and reverse relations, respectively.

To find parameter deviations, in previous work a backpropagation algorithm is used. In qualitative matrix algebra this is equivalent to repeated multiplication of the initial deviation with the transpose TCM. Here, for f_R^+ this results in the sequence of vectors

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ R^{-1} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ C^{-1} \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ ? \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ ? \\ ? \end{bmatrix}, \begin{bmatrix} ? \\ 1 \\ ? \\ ? \end{bmatrix}, \begin{bmatrix} ? \\ ? \\ ? \\ ? \end{bmatrix}. \tag{3}$$

The parameters R^{-1} and C^{-1} are fault hypotheses and replaced by 1 after they are generated because R and C are positive parameters, and, therefore, in a qualitative framework they represent direct relations. Also, qualitatively $1 - 1$ is unknown, "?". Once all variables are unknown, no further parameter deviations can be hypothesized (the remaining candidates that are not generated in Eq. (3) are $-R^{-1}$ and $-C^{-1}$). The resulting set of possible faults is, therefore, R^{-1} or C^{-1} too high, i.e., $\{R^-, C^-\}$ (the remaining candidates are $\{R^+, C^+\}$). Physically, these fault candidates correspond to, e.g., leakage in the outflow pipe (R^-) or an object that has fallen into the tank (C^-).

Next, predictions of future system behavior are generated for each of the possible parameter deviations, R^- and C^-. From the TCM, their initial deviations are found to be

$$R^- \rightarrow \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, C^- \rightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{4}$$

To achieve a suffiently high order prediction for the measured variable, f_R, the initial deviation is repeatedly multiplied with the TCM. Here, a second order prediction requires eight such multiplications and for R^- this yields

$$\begin{bmatrix} p_C \\ f_C \\ f_R \\ p_R \end{bmatrix} = \begin{bmatrix} 1 & \lambda^{-1} & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}^8 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\lambda^{-1} + \lambda^{-2} \\ -1 + \lambda^{-1} \\ 1 - \lambda^{-1} + \lambda^{-2} \\ -\lambda^{-1} + \lambda^{-2} \end{bmatrix}. \tag{5}$$

The TCM raised to the power 8 can be computed off-line to be

$$\begin{bmatrix} 1 & \lambda^{-1} & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}^8 = \begin{bmatrix} 1 - \lambda^{-1} + \lambda^{-2} & \lambda^{-1} - \lambda^{-2} & -\lambda^{-1} + \lambda^{-2} & -\lambda^{-1} + \lambda^{-2} \\ -1 + \lambda^{-1} & 1 - \lambda^{-1} + \lambda^{-2} & -1 + \lambda^{-1} & -1 + \lambda^{-1} \\ 1 - \lambda^{-1} & \lambda^{-1} - \lambda^{-2} & 1 - \lambda^{-1} + \lambda^{-2} & 1 - \lambda^{-1} \\ 1 - \lambda^{-1} & \lambda^{-1} - \lambda^{-2} & -\lambda^{-1} + \lambda^{-2} & 1 - \lambda^{-1} + \lambda^{-2} \end{bmatrix} \tag{6}$$

and can be used for efficiently generating predictions for other fault candidates.

The polynomials in λ are equal to the qualitative signatures generated in previous work [8, 9]. For this example, the signature for the measured variable is f_R^{+-+}, where the superscripts indicate the qualitative values of the time derivative behavior with increasing order from left to right, i.e., there is a positive discontinuous change with negative slope that increases. For the pressure at the bottom of the tank, the prediction is p_C^{0-+}, i.e., no discontinuous change in pressure occurs and the pressure is decreasing.

This method works well if the system of equations that describes continuous behavior is fixed. However, in case discrete switches cause changes in the continuous model, signatures for each mode have to be generated. This quickly becomes intractable, and, therefore, for these system models a parametrized formulation is advantageous.

3 Hybrid Models for FDI

For the qualitative FDI approach to be effective, it is imperative that the modeled time constants are observable, i.e., within the bandwidth of the data acquisition system. If a parameter that models an abrupt fault changes to a very large or small value, it may correspond to a time constant that cannot be observed, and, therefore, this behavior needs to be abstracted from the model. This causes the model to be of a switched continuous, hybrid nature.

In general, modeled discontinuities result in causal changes. Therefore, the TCM may take several different forms and so do the corresponding predictions of future behavior, depending on whether a mode change occurs. Consider for example a valve that controls the outflow in Fig. 1 in a binary manner, i.e., either there is an outflow determined by the Bernoulli resistance ($\alpha_1 = 1$) or there is no outflow ($\alpha_1 = 0$). When the switch is modeled as a discontinuous change, the corresponding model includes a change in causality when the control valve switches its state. If it is open, the pressure p_C determines the outflow f_R and if it is closed, $f_R = 0$, which determines the pressure drop across the pipe to be $p_R = f_R R = 0$. To handle the change in TCM, the causal relations can be parametrized to make them dependent on the mode of operation.

To this end, first the system is described in a noncausal form by using implicit equations. An implicit model of the one tank consists of the following equations

$$0 = C\dot{p}_C - f_C \tag{7}$$

$$0 = f_C - f_{in} + f_R \tag{8}$$

$$0 = Rf_R - p_R \tag{9}$$

$$0 = \alpha_1(p_R - p_C) + (1 - \alpha_1)f_R \tag{10}$$

From Eq. (10), in case the control valve is open, $\alpha_1 = 1$, and $p_R = p_C$, when the control valve is closed, $\alpha_1 = 0$, and $f_R = 0$.

The TCM for this system of equations contains the relations between each of the variables. For example, Eq. (7) embodies a temporal relation between p_C and f_C and Eq. (10) a direct relation between p_C and p_R that is only active when $\alpha_1 \neq 0$. The TCM then becomes

$$\begin{bmatrix} 1 & \lambda^{-1}C^{-1} & 0 & \alpha_1 \\ \lambda C & 1 & -1 & 0 \\ 0 & -1 & 1 & R^{-1} \\ \alpha_1 & 0 & R & 1 \end{bmatrix} \begin{bmatrix} p_C \\ f_C \\ f_R \\ p_R \end{bmatrix} \tag{11}$$

and causal links from p_C to p_R and from p_R to p_C are only active when the system is in mode α_1. A special case arises for $\alpha_1 = 0$ which implies $f_R = 0$. This effect is not present in the TCM because it is not a relation between variables. However, it contains essential diagnostic information about system behavior that can be included by an input vector

$$\begin{bmatrix} 0 \\ 0 \\ -(1-\alpha_1) \\ 0 \end{bmatrix} \tag{12}$$

where the $-$ sign is because the flow, f_R, is positive during normal operation, and, therefore, its deviation is $-$ when the valve closes (possibly inadvertently).

Diagnosis now proceeds to predict future behavior, y_f, for each hypothesized fault, f, and both possible configurations ($\alpha_1 = 0$ and $\alpha_1 = 1$). To this end, the TCM, A, raised to a sufficiently high power, n, operates on the sum of the input vector, u, and each of the initial deviations, d_f, generated from the hypothesized faults,

$$y_f = A^n(d_f + u) \tag{13}$$

These predictions are then compared against actual observations to prune the fault hypotheses and find the correct fault.

Note that, to facilitate a qualitative algebra, the $(1 - \alpha)$ construct with $\alpha \in \{0, 1\}$ cannot be used to (de)activate relations because in a qualitative sense $(1-\alpha)$ is unknown instead of 0. Therefore, $\neg\alpha$ is used to indicate a quantitative evaluation of $(1 - \alpha)$ so that $\neg\alpha$ produces a value $\{0, 1\}$.

For the initial deviation that corresponds to R^- in Eq. (4) and the input vector in Eq. (12), after multiplying with the TCM five times, the prediction becomes

$$\begin{bmatrix} \alpha_1 - \lambda^{-1} \\ \alpha_1\lambda - 1 + \alpha_1\lambda^{-1} \\ -\alpha_1\lambda + 1 - \alpha_1\lambda^{-1} \\ -\alpha_1\lambda + 1 - \alpha_1\lambda^{-1} \end{bmatrix} (1 - \neg\alpha_1) \tag{14}$$

Compared with the prediction derived from the explicit system in Section 2 this shows impulsive behavior because of the positive powers of λ and other spurious behavior because *all* possible relations are present in the TCM. In other words, for a given causal assignment all other relations are present as well even though these may not be consistent with the given causal assignment.

To demonstrate that such an extensive set of relations quickly leads to contradiction, consider an implicit relation $0 = x_1 + x_2 + x_3$ with TCM

$$\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{15}$$

Because in a qualitative sense $1-1$ is unknown, this leads to unknown predictions as soon as the TCM is raised to a power > 1 (e.g., $x_1^+ \rightarrow x_2^- \rightarrow x_3^+ \rightarrow x_1^-$, and x_1 is unknown). This problem can be circumvented by committing to one causal assignment only. In matrix form, this is achieved by using binary selection variables, $k_i \in \{0, 1\}$,

$$\begin{bmatrix} 1 & -k_1 k_2 & -k_1 k_2 \\ -k_1 \neg k_2 & 1 & -k_1 \neg k_2 \\ -k_2 \neg k_1 & -k_2 \neg k_1 & 1 \end{bmatrix} \tag{16}$$

and the matrix is invariant under multiplication.

In summary, to design an approach for diagnosis based on hybrid models, the TCM is derived from an implicit model formulation that includes mode selection parameters, α_i, to switch between equations. The possible causal assignments of ternary and higher relations are then made mutually exclusive by introducing selection parameters, k_i. If possible, the parameters α_i can be related to k_i and the TCM contains only mode selection parameters, α_i, and, therefore, produces fault hypotheses and predictions that are parametrized by α_i only.

4 A Case Study

To make the implicit approach suitable for diagnosis, it must deal with additional causal paths and the possible conflicts. Consider the two tank system in Fig. 3 with externally controlled outflow valves on the left and right and a pressure controlled valve between the left and right tank. An implicit quantitative model of this system could look like

$$\begin{aligned}
0 &= -f_{in} + f_{C_1} + f_{R_{b1}} + f_{R_{12}} \\
0 &= \alpha_1(-p_{C_1} + p_{R_{12}} + p_{C_2}) + (1-\alpha_1) f_{R_{12}} \\
0 &= \alpha_2(p_{C_1} - p_{R_{b1}}) + (1-\alpha_2) f_{R_{b1}} \\
0 &= \alpha_3(p_{C_2} - p_{R_{b2}}) + (1-\alpha_3) f_{R_{b2}} \\
0 &= f_{C_2} + f_{R_{b2}} - f_{R_{12}} \\
0 &= C_1 \dot{p}_{C_1} - f_{C_1} \\
0 &= C_2 \dot{p}_{C_2} - f_{C_2} \\
0 &= p_{R_{b1}} - R_{b1} f_{R_{b1}} \\
0 &= p_{R_{b2}} - R_{b2} f_{R_{b2}} \\
0 &= p_{R_{12}} - R_{12} f_{R_{12}}
\end{aligned} \tag{17}$$

where α_i are mode selection parameters and α_1, α_2, and α_3 correspond to the state of the middle, left, and right valves in Fig. 3, respectively, where $\alpha_i = 0$ implies the valve is closed and $\alpha_i = 1$ that the valve is open.

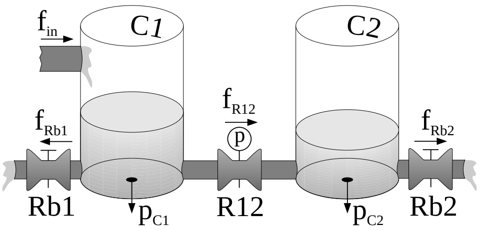

Fig. 3. Two tanks with outflow valves and a pressure controlled connecting valve.

This model contains a number of ternary relations (input variables are not considered as fault candidates) and when a deviation is propagated, multiple possible paths are taken. To prevent this, the paths can be parametrized as demonstrated in Section 3 (the binary relations are mutually consistent),

$$\begin{bmatrix} 1 & -k_1 k_2 & -k_1 k_2 & 0 & 0 & \boxed{\lambda C_1} & 0 & 0 & 0 & 0 \\ \boxed{-k_1 \neg k_2} & 1 & \boxed{-k_1 \neg k_2} & 0 & 0 & 0 & 0 & 0 & R_{b1}^{-1} & 0 \\ \boxed{-k_2 \neg k_1} & \boxed{-k_2 \neg k_1} & 1 & \boxed{k_3 k_4} & \boxed{k_3 k_4} & 0 & R_{12}^{-1} & 0 & 0 & 0 \\ 0 & 0 & k_3 \neg k_4 & 1 & -k_3 \neg k_4 & 0 & 0 & \boxed{\lambda C_2} & 0 & 0 \\ 0 & 0 & \boxed{k_4 \neg k_3} & \boxed{-k_4 \neg k_3} & 1 & 0 & 0 & 0 & 0 & \boxed{R_{b2}^{-1}} \\ \lambda^{-1} C_1^{-1} & 0 & 0 & 0 & 0 & 1 & \boxed{\alpha_1 k_5 k_6} & \boxed{\alpha_1 k_5 k_6} & \boxed{\alpha_2} & 0 \\ 0 & 0 & R_{12} & 0 & 0 & \alpha_1 k_6 \neg k_5 & 1 & -\alpha_1 k_6 \neg k_5 & 0 & 0 \\ 0 & 0 & 0 & \lambda^{-1} C_2^{-1} & 0 & \boxed{\alpha_1 k_5 \neg k_6} & \boxed{-\alpha_1 k_5 \neg k_6} & 1 & 0 & \boxed{\alpha_3} \\ 0 & R_{b1} & 0 & 0 & 0 & \alpha_2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & R_{b2} & 0 & 0 & \alpha_3 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{C_1} \\ f_{R_{b1}} \\ f_{R_{12}} \\ f_{C_2} \\ f_{R_{b2}} \\ p_{C_1} \\ p_{R_{12}} \\ p_{C_2} \\ p_{R_{b1}} \\ p_{R_{b2}} \end{bmatrix} \tag{18}$$

For this model, the causality of some of the binary relations is fixed for each possible mode and incorporating this *a priori* knowledge leads to a more constrained model. For example, the relation $0 = \alpha_2(p_{C_1} - p_{R_{b1}})$ leads to two entries in the TCM, one for $p_{C_1} \overset{\alpha_2}{\rightarrow} p_{R_{b1}}$ and one for $p_{R_{b1}} \overset{\alpha_2}{\rightarrow} p_{C_1}$. Analysis reveals that the latter causal relation is never used for any configuration of valve states, and, therefore, the corresponding entry in the TCM can be removed. The matrix entries in Eq. (18) that vanish because of pre-processing are marked by a bounding box.

The causality of the ternary relations can be analyzed exhaustively because it only involves a limited number of local constraints. Causal analysis of the system of equations shows that although the causality of the ternary equations may change, the changed causality corresponds to the vanishing (deactivating) of an edge. For example, the causality of $0 = \alpha_1(-p_{C_1} + p_{R_{12}} + p_{C_2})$ changes when α_1 changes its value. But, for the state $\neg\alpha_1$, the equation is not active anymore. Therefore, this need not be explicitly modeled, and the relation between the α_i

and k_i degrades to the fixed values $k_1 = 1$, $k_2 = 1$, $k_3 = 1$, $k_4 = 0$, $k_5 = 0$, and $k_6 = 1$.

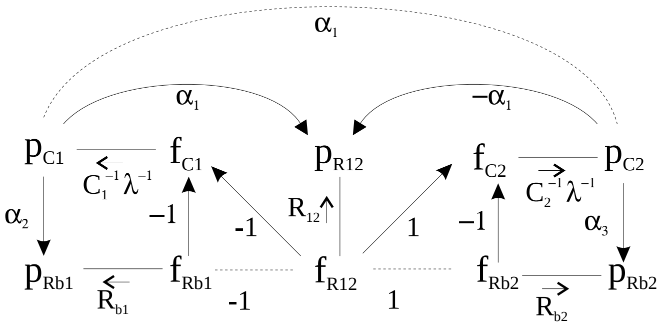

Fig. 4. The temporal causal graph of the two-tank system.

In Fig. 4 the temporal graph of the TCM is shown to clarify the relations between system variables. The dashed edges are those that are present in the original implicit formulation because of ternary relations but that are removed based on a mode dependent causal analysis. The undirected edges are implicit binary relations and can be decomposed into two edges with opposite direction (corresponding to the two entries in the TCM) to be compatible with the temporal causal graph format used in previous work [8, 9]. Note that in many cases, graph propagation is more efficient than matrix multiplication, especially in case of sparse matrices.

After replacing the parameters with their qualitative equivalent, the resulting TCM is given by

$$\begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \boxed{1} \\ \lambda^{-1} & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & 0 & \alpha_1 & 1 & -\alpha_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda^{-1} & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 & 0 & \alpha_2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} & 0 & 0 & \alpha_3 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{C_1} \\ f_{R_{b1}} \\ f_{R_{12}} \\ f_{C_2} \\ f_{R_{b2}} \\ p_{C_1} \\ p_{R_{12}} \\ p_{C_2} \\ p_{R_{b1}} \\ p_{R_{b2}} \end{bmatrix} \tag{19}$$

where the boxed entries are those that correspond to bidirectional, non-causal, edges (in this particular case, these could still be made mode-dependent, where the entries above the diagonal become α_i and below become $\neg\alpha_i$).

The predictions of the TCM are parametrized by the active mode. This leads to more efficient diagnosis compared to the use of a bank of TCMs, which, in this case of three switches, would consist of eight TCMs that need to be processed separately. For example, in case of a measurement $f^{+}_{R_{b2}}$, R^{-}_{b2} is one of the fault hypotheses that results in the prediction

$$\begin{bmatrix} -\alpha_1\lambda^{-1} + \alpha_1\lambda^{-2} \\ -\alpha_1\alpha_2\lambda^{-2} \\ \alpha_1\lambda^{-1} - \alpha_1\lambda^{-2} \\ -1 + \alpha_1\lambda^{-1} + \alpha_3\lambda^{-1} - \alpha_1\lambda^{-2} \\ 1 - \alpha_3\lambda^{-1} + \alpha_3\lambda^{-2} \\ -\alpha_1\lambda^{-2} \\ \alpha_1\lambda^{-1} - \alpha_1\lambda^{-2} \\ -\lambda^{-1} + \alpha_3\lambda^{-2} \\ -\alpha_1\alpha_2\lambda^{-2} \\ 1 - \alpha_3\lambda^{-1} + \alpha_3\lambda^{-2} \end{bmatrix}. \tag{20}$$

In addition, the input vectors for $\neg\alpha_1$, $\neg\alpha_2$ and $\neg\alpha_3$ are determined to be

$$\neg\alpha_1 \rightarrow \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \neg\alpha_2 \rightarrow \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \neg\alpha_3 \rightarrow \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{21}$$

and their effect is propagated as well. For $\neg\alpha_1$, this leads to the prediction for p_{C_1} to be $\neg\alpha_1\lambda^{-1} - \neg\alpha_1\alpha_2\lambda^{-2} - \neg\alpha_1\alpha_1\lambda^{-2}$, or $\neg\alpha_1\lambda^{-1} - \neg\alpha_1\alpha_2\lambda^{-2}$. The combined prediction for p_{C_1} becomes

$$\neg\alpha_1\lambda^{-1} - \neg\alpha_1\alpha_2\lambda^{-2} - \alpha_1\lambda^{-2} \tag{22}$$

The parametrized predictions can be matched against further measurements (e.g., $p_{C_1}^{0+}$, where the second order derivative is not measured). In case α_1, i.e., the pressure controlled connecting valve remains open, the prediction for p_{C1} is $-\lambda^{-2}$, a falling level of liquid in C_1 with second order behavior. This is inconsistent with the $p_{C_1}^{0+}$ observation and the fault $R_{b_2}^{-}[\alpha_1]$ is rejected as a possible explanation of the anomalous system behavior. If the new pressure in C_2 causes the connecting valve to close, the predicted behavior of p_{C_1} changes. This can be derived by evaluating the prediction with $\neg\alpha_1$, which yields $\lambda^{-1} - \alpha_2\lambda^{-2}$, i.e., the liquid level in C_1 rises. In case the left outflow valve remains open, α_2, the rate of increase decreases but if this outflow valve closes, the level continues to rise. It is easily verified that the predictions of both fault hypotheses ($R_{b2}^{-}[\neg\alpha_1\alpha_2]$ and $R_{b2}^{-}[\neg\alpha_1\neg\alpha_2]$) are consistent with the $p_{C_1}^{0+}$ measurement, and, therefore, possible causes of the observed anomalous behavior. Further measurements are needed to prune this set of candidates, as described in detail elsewhere [8, 9].

5 Conclusions

Algorithms and hybrid models for diagnosis of physical systems are required to deal with configuration changes between modes of operation but the combinatorial explosion prohibits a global enumeration approach. This papers shows that mode changes can be modeled by locally activating and deactivating relations between system variables. When relations are (de)activated, the causal effect between system variables may change. This is handled by including all possible relations between system variables. Because of the presence of relations not describing system behavior in a given mode, the model may foster conflicting relations, which is solved by introducing parameters to enforce mutual exclusion between different causal assignments on individual relations. Performing local analyses establishes the relation between these parameters and mode selection parameters. The resulting method generates conditional predictions that depend on the mode of the system which allows for efficient execution of the diagnosis algorithms.

The presented method allows for a declarative prediction of future system behavior. It has not taken yet taken into account imperative mode switching functionality (e.g., a switching constraint such as $p_1 > p_2$ causes $\alpha_2 = 1$). Including this may constrain possible mode changes, and, therefore, further prune the set of hypothesized candidates.

Note that the analysis of interacting local switches is automated in HyBrSim [10] based on analysis of *causal areas* in a bond graph. This forms the basis for future research into automatically performing the pre-processing of the relations between mode selection parameters and those that ensure mutual exclusion of different causal assignments. This should facilitate scaling the approach, because the complexity increases exponentially only with interacting switches within one causal area. So, e.g., for k causal areas with m switches, instead of 2^{km} modes, $k2^m$ modes have to be analyzed, and typically if a hybrid bond graph modeling approach is useful, the number of switches that interact directly, i.e., without dynamic behavior, is low.

References

1. R.N. Clark, P.M. Frank, and R.J. Patton. Introduction. In Ron Patton, Paul Frank, and Robert Clark, editors, *Fault Diagnosis in Dynamic Systems: Theory and Applications*, chapter 1, pages 1–19. Prentice-Hall, UK, 1989.
2. Philippus J. Feenstra, Eric J. Manders, Pieter J. Mosterman, Gautam Biswas, and Joel Barnett. Modeling and instrumentation for fault detection and isolation of a cooling system. In *Proceedings of the IEEE Southeastern Conference 2000*, pages 365–372, Nashville, TN, April 2000.
3. Philippus J. Feenstra, Pieter J. Mosterman, Gautam Biswas, and Peter C. Breedveld. Bond graph modeling procedures for fault detection and isolation of complex flow processes. In *International Conference on Bond Graph Modeling and Simulation (ICBGM '01)*, pages 77–82, Phoenix, AZ, January 2001.
4. Paul Frank. Fault diagnosis: A survey and some new results. *Automatica: IFAC Journal*, 26(3):459–474, 1990.
5. Jan Lunze. Diagnosis of quantised systems by means of timed discrete-event representations. In Nancy Lynch and Bruce H. Krogh, editors, *Lecture Notes in Computer Science, Hybrid Systems: Computation and Control*, pages 258–271, Berlin, 2000. Springer-Verlag.
6. Eric Jan Manders, Pieter J. Mosterman, and Gautam Biswas. Signal to Symbol Transformation Techniques for Robust Diagnosis in Transcend. In *Tenth Inter-*

national Workshop on Principles of Diagnosis, pages 155–165, Lock Awe Hotel, Scotland, June 1999.

7. Sheila McIlraith, Gautam Biswas, Dan Clancy, and Vineet Gupta. Hybrid systems diagnosis. In Nancy Lynch and Bruce H. Krogh, editors, *Lecture Notes in Computer Science, Hybrid Systems: Computation and Control*, pages 282–295, Berlin, 2000. Springer-Verlag.
8. Pieter J. Mosterman. *Hybrid Dynamic Systems: A hybrid bond graph modeling paradigm and its application in diagnosis*. PhD dissertation, Vanderbilt University, 1997.
9. Pieter J. Mosterman and Gautam Biswas. Diagnosis of continuous valued systems in transient operating regions. *IEEE Transactions on Systems, Man, and Cybernetics*, 29(6):554–565, November 1999.
10. Pieter J. Mosterman and Gautam Biswas. A Java Implementation of an Environment for Hybrid Modeling and Simulation of Physical Systems. In *International Conference on Bond Graph Modeling (ICBGM '99)*, pages 157–162, San Francisco, January 1999.
11. Pieter J. Mosterman, Ravi Kapadia, and Gautam Biswas. Using Bond Graphs for Diagnosis of Dynamic Physical Systems. In *Sixth International Workshop on Principles of Diagnosis*, pages 81–85, Goslar, Germany, October 1995.
12. Pieter J. Mosterman, Eric J. Manders, and Gautam Biswas. Qualitative Dynamic Behavior of Physical System Models With Algebraic Loops. In *Eleventh International Workshop on Principles of Diagnosis*, pages 155–162, Morelia, Mexico, 2000.
13. Sriram Narasimhan, Gautam Biswas, Gabor Karsai, Tal Pasternak, and Feng Zhao. Building observers to address fault isolation and control problems in hybrid dynamic systems. In *Proceedings of the 2000 IEEE International Conference on Systems, Man, and Cybernetics*, pages 2393–2398, Nashville, TN, 2000.
14. Sriram Narasimhan, Feng Zhao, Gautam Biswas, and Elmer Hung. An integrated framework for combining global and local analyses in diagnosing hybrid systems. In *Eleventh International Workshop on Principles of Diagnosis*, pages 163–170, Morelia, Mexico, June 2000.
15. M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D.C. Teneketzis. Failure diagnosis using discrete-event models. *IEEE Transactions on Automatic Control*, 40(9):1555–1575, September 1995.
16. M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D.C. Teneketzis. Failure diagnosis using discrete-event models. *IEEE Transactions on Control Systems Technology*, 4(2):105–124, March 1996.
17. Marc Torrens, Rainer Weigel, and Boi Faltings. Java constraint library: bringing constraint technology on the internet using the java language. In *AAAI-97*, Providence, Rhode Island, 1997. Workshop on Constraints and Agents.