Fault diagnosis based on dissipativity property

Author:
Lei, Q; Wang, R; Bao, J

Publication details:
Computers and Chemical Engineering
v. 108
pp. 360 - 371
0098-1354 (ISSN); 1873-4375 (ISSN)

Publication Date:
2018-01-04

Publisher DOI:
https://doi.org/10.1016/j.compchemeng.2017.09.024

Fault Diagnosis Based on Dissipativity Property

Qingyang Lei, Ruigang Wang, Jie Bao*

School of Chemical Engineering, The University of New South Wales,

UNSW, Sydney, NSW, 2052

Abstract

In this paper, a novel fault diagnosis scheme for linear process systems using dissipativity theory is developed. Dissipativity (supply rate) of a process is an input/output property, which may not be valid when a fault occurs. For a given process, dissipativity is not a unique property, with different dissipative supply rates reflecting different aspects of its dynamics. In this approach, the dissipativity of a process is "shaped" such that it is fault-sensitive (i.e., no longer valid when faults occur) and fault-selective (i.e., no longer valid when one particular fault occurs). By adopting the storage functions and supply rates in the Quadratic Difference Form (QdF), the dissipativity conditions are represented as quadratic functions of the input/output trajectories of the process, which captures much more detailed dynamical features compared to conventional dissipativity (e.g., QSR-type supply rates). These dissipativity properties are determined offline by solving an optimization problem with linear matrix inequality constraints. The online diagnosis algorithm involves checking of inequalities on input/output trajectories, which is much simpler compared to the diagnosis approaches based on observers or parameter estimation. The proposed approach is illustrated using a case study of fault diagnosis of a heat exchanger.

Keywords: Fault diagnosis; dissipativity; quadratic difference form

1 Introduction

Modern industrial processes are becoming very complex. The increasing dependence of complex processes on automatic control systems can make the plants susceptible to faults such as sensor/actuator failures. Therefore, fault detection (i.e., to identify if there is a fault) and diagnosis (i.e., to determine what fault occurs) are becoming an important issue in process control practice. Model-based fault detection approaches, including observer-based, parity equation-based and parameter estimation-based methods, utilize the mathematical models of the processes (referred to surveys [1, 2]). The general procedure of observer-based methods usually involves

*Corresponding author. e-mail: j.bao@unsw.edu.au.

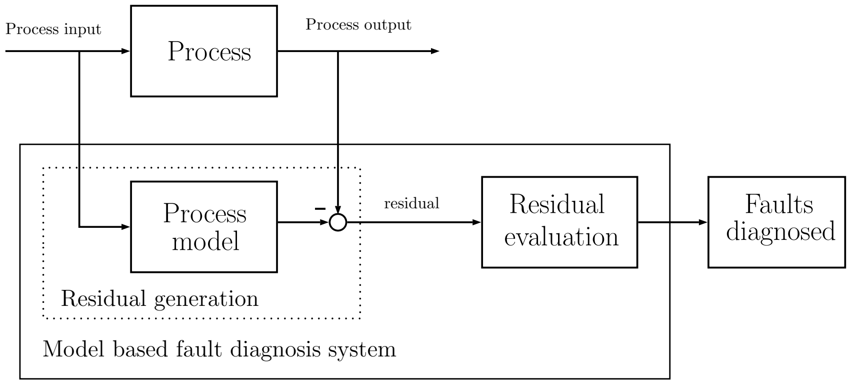


Figure 1: Model-based fault diagnosis scheme (adapted from [9])

two steps, residual generation and decision making [3], as depicted in Figure 1. The residuals are shaped such that they are sensitive to abnormal conditions. An example of observer-based method is the fault detection filter, proposed in [4] and [5]. Parity equation based methods (e.g., [6]) generate parity vector (residuals), that is used to check the consistency between process model and process outputs [7]. While they are simpler than observer-based approaches, parity equation based methods can be less effective in detecting faults and are limited to faults that do not include gross parameter drifts [1]. Another fault detection method is based on parameter estimation, which is formed on the basis of system identification techniques [8]. The basic idea is to identify the actual process parameters online, and compare them with the parameters of the fault-free process model.

Many of the above fault detection methods have been extended for fault diagnosis [9]. For observer-based methods, a bank of observers, one for each fault or a group of faults, are required for fault diagnosis. One intuitive idea is to make a residual sensitive to the fault that is concerned and robust to all other faults (i.e., structure residual fault isolation [10]). Alternatively, the residual can be shaped to be robust to all but one fault and also robust against uncertainties (i.e., generalized residual fault isolation [3]). Generally, a fault diagnosis method needs to generate several representative symptoms. For example, in [11], fault-symptom tables have been used, and systematic treatment of fault-symptom trees is based on approximative reasoning with if-then-rules by fuzzy logic. However, the implementation of above observer based approach can be complex, especially for large scale chemical processes [1]. Fault diagnosis methods based on parameter estimation methods are suitable for the diagnosis of multiplicative faults (with process parameter changes), but they require dynamic process input excitation which is often infeasible in online monitoring [11].

In this paper, a fault diagnosis approach is developed based on dissipativity theory. Dissipativity theory, introduced by Willems in [12], has become an important tool for system analysis

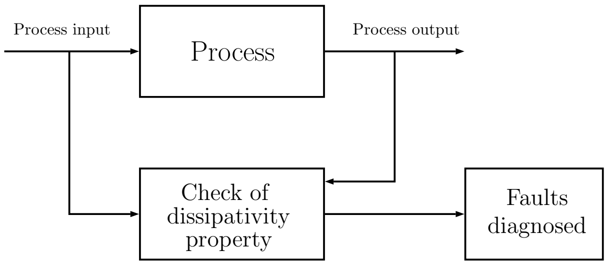


Figure 2: Dissipativity-based fault diagnosis scheme

and control design (e.g., [13, 14, 15]). Dissipativity (represented by a supply rate) is an input/output property of a system [12], representing the features of process dynamics, such as the gain and phase conditions and their combinations [16]. When a fault (e.g., a multiplicative fault, which is modeled by parameter changes [9]) occurs, it can be identified by checking the change of dissipativity property, as depicted in Figure 2. The dissipativity property of a process is not unique. For the same process, different aspects of the process dynamics can be captured by different supply rates. In this paper, the dissipativity properties of a process are shaped to be sensitive to different faults (fault-selective). The dissipativity shaping problem is formulated in linear matrix inequality (LMI) constraints, which can be easily solved offline using any semi-definite programming tools. Furthermore, a robust dissipativity condition is also developed, which is incorporated in the proposed fault diagnosis approach to reduce the rates of false alarm caused by uncertainties. Compared to existing approaches, the proposed approach is simpler to implement, as it does not require observers or parameter estimation.

Passivity condition (a special case of dissipativity) was used for fault detection and diagnosis for passive electronic circuits, as shown in [17]. However, the passivity condition is very coarse and may not capture sufficient dynamic details of process input output relationship, leading to limited capacity in fault detection and diagnosis. Another issue in existing passivity based approach is that it needs the full state information as the storage function is defined on state variables, which are usually unavailable in practice. To overcome the above problems, dissipativity in the Quadratic Difference Forms (QdF) (the discrete-time version of the dissipativity in Quadratic Differential Forms developed by Willems and Trentelman [18]) is adopted in this work, where both the storage functions and supply rates are defined as functions of input/output trajectories (as in [19, 20, 21]). This eliminates the need for state estimation for fault diagnosis. Furthermore, as a more general form of dissipativity, the QdF supply rates and storage functions can capture much more details of the dynamic features (e.g. the gain, phase or their combination *at different frequencies*) of the process, comparing to traditional QSR dissipativity [22, 23], leading to much more effective fault diagnosis.

This paper is organized as follows. The framework for analyzing faults using dissipativity theory is introduced in Section 2. The developments of the novel dissipativity-based fault diagnosis method is presented in Sections 3. The robust dissipativity condition for the proposed fault diagnosis approach is developed in Section 4. The proposed approach is illustrated on a heat exchanger case study in Section 5, followed by the discussion and conclusion in Section 6.

2 Fault Analysis using Dissipativity Theory

In this section, some important concepts of dissipativity theory and quadratic difference form (QdF) are introduced, followed by the framework of dissipativity based fault analysis, which is different from classical fault detection and diagnosis methods based on analytical redundancy.

2.1 Introduction to Dissipativity Theory

The dissipativity theory was first introduced by Willems in [12], as a framework for analyzing dynamical systems. While inspired by a class of systems which dissipate energy, the concept of dissipative systems is developed for general systems where the energy can be abstract and not necessarily physical [12, 15].

Consider a linear time-invariant process defined by the following state space equations

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k + Du_k \end{aligned} \tag{1}$$

where $x \in \mathcal{X} \subset \mathbb{R}^n$ are process variables, k is the time step, $u \in \mathcal{U} \subset \mathbb{R}^p$ is the input vector and $y \in \mathcal{Y} \subset \mathbb{R}^q$ is the output vector.

Definition 1 ([12])**.** *Consider the system described by* (1)*. Define a function $s(u_k, y_k)$ on input and output variables, called the supply rate. The system is said to be dissipative with respect to the supply rate $s(u_k, y_k)$ if there exists a positive semi-definite function $V(x_k)$ defined on the states, called the storage function, such that the following dissipativity inequality is satisfied*

$$V(x_{k+1}) - V(x_k) \leq s(u_k, y_k). \tag{2}$$

for all $x_k \in \mathcal{X}$, $u_k \in \mathcal{U}$ and k.

The following (Q,S,R) type supply rate is commonly used:

$$s(u, y) = y^\top Qy + 2y^\top Su + u^\top Ru. \tag{3}$$

As aforementioned, full state measurements are usually unavailable in online monitoring and process control practice [1]. Therefore the traditional dissipativity condition given in (2), with storage function defined on state variables, cannot be directly used for fault detection and diagnosis. To overcome this difficulty, in this work, the behavior systems approach developed by Willems [24] is adopted. For continuous time systems, Willems and Trentelman introduced storage functions and supply rates in the "Quadratic Differential Forms" (QDF) which are functions of the input and output and their derivatives [18]. This was later extended to "Quadratic Difference Forms" (QdF) for discrete time systems by Kojima and Takaba [19, 20], as follows:

Definition 2 (adopted from [19])**.** *Consider the system described by the model in* (1)*. Define the extended input and output as*

$$\begin{aligned} \hat{u}_k &= \begin{pmatrix} u_k^\top & u_{k+1}^\top & \dots & u_{k+n_u}^\top \end{pmatrix}^\top \\ \hat{y}_k &= \begin{pmatrix} y_k^\top & y_{k+1}^\top & \dots & y_{k+n_y}^\top \end{pmatrix}^\top . \end{aligned} \tag{4}$$

for some finite n_u, n_y. Also define the storage function Q_Ψ and a supply rate Q_Φ in the following Quadratic Difference Form (QdF) ([25]), as functions of the extended input and output:

$$Q_\Psi(u_k, y_k) = \begin{pmatrix} \hat{y}_k \\ \hat{u}_k \end{pmatrix}^\top \tilde{\psi} \begin{pmatrix} \hat{y}_k \\ \hat{u}_k \end{pmatrix}, \quad Q_\Phi(u_k, y_k) = \begin{pmatrix} \hat{y}_k \\ \hat{u}_k \end{pmatrix}^\top \tilde{\phi} \begin{pmatrix} \hat{y}_k \\ \hat{u}_k \end{pmatrix}. \tag{5}$$

where $\tilde{\psi}$ and $\tilde{\phi}$ are the constant coefficient matrices respectively. The system is said to be dissipative with respect to supply rate Q_Φ if there exists a semi-positive definite storage functions $Q_\Psi(u_k, y_k)$ such that the following inequality is satisfied,

$$Q_\Psi(u_{k+1}, y_{k+1}) - Q_\Psi(u_k, y_k) \leq Q_\Phi(u_k, y_k) \tag{6}$$

for all $u_k \in \mathcal{U}$ and k.

The QdF in (5) can be rewritten as

$$Q_\Phi(u_k, y_k) = \sum_{i=0}^{N} \sum_{j=0}^{N} \begin{pmatrix} y_{k+i} \\ u_{k+i} \end{pmatrix}^T \tilde{\phi}_{ij} \begin{pmatrix} y_{k+j} \\ u_{k+j} \end{pmatrix} \tag{7}$$

where N =max(n_u, n_y) is called the degree of supply rate, which is the maximum number of forward steps in the supply rate. Here $\tilde{\phi}_{ij}$ is a sub-matrix of the coefficient matrix $\tilde{\phi}$ as shown below:

$$\tilde{\phi} = \begin{pmatrix} \tilde{\phi}_{00} & \cdots & \cdots & \cdots & \tilde{\phi}_{0,N} \\ \vdots & \ddots & & & \vdots \\ \vdots & & \tilde{\phi}_{i,j} & & \vdots \\ \vdots & & & \ddots & \vdots \\ \tilde{\phi}_{N,0} & \cdots & \cdots & \cdots & \tilde{\phi}_{N,N} \end{pmatrix}. \tag{8}$$

The above QdF is said to be induced by the symmetric two-variable polynomial matrix $\phi(\zeta, \eta)$ defined as

$$\phi(\zeta, \eta) = \sum_{i=0}^{N} \sum_{j=0}^{N} \zeta^i \tilde{\phi}_{ij} \eta^j \tag{9}$$

where the indeterminates ζ and η represent a forward step in time on the left and right of input/output signals respectively, *i.e.*, $y_k^T \zeta = y_{k+1}^T$ and $\eta y_k = y_{k+1}$. The two-variable polynomial matrix, $\phi(\zeta, \eta)$, was introduced in [18] as a general and compact representation of QdFs. For example, a dynamic supply rate

$$\begin{aligned} Q_\Phi(u_k, y_k) &= y_k^2 + y_{k+1}^2 + y_k u_{k+1} - 2u_{k+2}^2 \\ &= \begin{pmatrix} y_k \\ y_{k+1} \\ y_{k+2} \\ u_k \\ u_{k+1} \\ u_{k+2} \end{pmatrix}^\top \begin{pmatrix} 1 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} y_k \\ y_{k+1} \\ y_{k+2} \\ u_k \\ u_{k+1} \\ u_{k+2} \end{pmatrix} \end{aligned} \tag{10}$$

can also be represented by

$$Q_\Phi(u_k, y_k) = \begin{pmatrix} y_k \\ u_k \end{pmatrix}^\top \begin{pmatrix} 1 + \zeta\eta & 1/2\eta \\ 1/2\zeta & -2\zeta^2\eta^2 \end{pmatrix} \begin{pmatrix} y_k \\ u_k \end{pmatrix} \tag{11}$$

Therefore, the supply rate in (10) is said to be induced by the following polynomial matrix:

$$\phi(\zeta, \eta) = \begin{pmatrix} 1 + \zeta\eta & 1/2\eta \\ 1/2\zeta & -2\zeta^2\eta^2 \end{pmatrix}. \tag{12}$$

The advantage of adopting above notation of two-variable polynomial matrix is that the dynamic features of processes (e.g., frequency weighted gain and phase conditions) can be directly represented by dynamic supply rates in the QdF form. As shown in (17) that the two-variable polynomial matrix $\phi(\zeta, \eta)$ that represents a frequency weighted gain condition can be determined directly from the weighting function.

The QdF supply rate Q_Φ can be viewed as an extension of the commonly used QSR type supply rate in (3) by including future steps of inputs and outputs (trajectories) ([19, 20]):

$$Q_\Phi(u_k, y_k) = \begin{pmatrix} \hat{y}_k \\ \hat{u}_k \end{pmatrix}^\top \begin{pmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^\top & \tilde{R} \end{pmatrix} \begin{pmatrix} \hat{y}_k \\ \hat{u}_k \end{pmatrix} \tag{13}$$

QdF storage functions are defined similarly as functions of input and output trajectories rather than based on state variables.

The dissipation rate Q_Δ is defined as follows:

$$Q_\Delta(u_k, y_k) = Q_\Phi(u_k, y_k) - [Q_\Psi(u_{k+1}, y_{k+1}) - Q_\Psi(u_k, y_k)], \tag{14}$$

where a positive dissipation rate indicating the dissipativity condition is satisfied. With suitable choices of $Q_\Phi(u_k, y_k)$ and $Q_\Psi(u_k, y_k)$, when a fault occurs, the system will no longer be dissipative with respect to above supply rate $Q_\Phi(u_k, y_k)$ and storage function $Q_\Psi(u_k, y_k)$ (i.e., the dissipation rate becomes negative).

2.2 Dissipativity based Fault Analysis

In the proposed approach, the changes of process dynamic features caused by faults are captured by the changes of the dissipativity (storage functions and supply rates). For example, a multiplicative fault modeled by changes in the parameters of the model [9], leads to changes in the gain and phase at different frequencies, which can be in turn captured by changes in dissipativity. A process can have different dissipativity properties, capturing different aspects of its dynamic features. For example, if a process is dissipative with respect to a supply rate of

$$s(u, y) = -y^\top y + \gamma^2 u^\top u, \tag{15}$$

then the $\mathcal{L}_\infty$ norm of this process is bounded by γ. If a process is dissipative with respect to the supply rate

$$s(u, y) = y^\top u, \tag{16}$$

then this process is passive and phase bounded between $[-\pi/2, \pi/2]$. The QSR-dissipativity in (3) implies the conditions on both the gain and phase. It should be noted that the dissipativity concept is very general, not limited to physically dissipative systems. Any process (even unstable) can be dissipative with respect to certain storage functions and supply rates. Intuitively the changes in the gain and phase caused by a fault will invalidate a dissipativity condition. To illustrate this, we consider a number of simple examples:

Example 1. *Consider a static nominal process $y = u$. It has many dissipativity properties, among which we have (1) a supply rate of $s(u, y) = -y^\top y + u^\top u$ with a void storage function (in this case the dissipation inequality is $-y^\top y + u^\top u > 0$), representing the upper bound of the process gain (2) a supply rate of $s(u, y) = y^\top y - u^\top u$ with a void storage function (in this case the dissipation inequality is $y^\top y - u^\top u > 0$), representing the lower bound of the process gain. It can be easily verified that: if a fault that causes the process gain to increase occurs (e.g., the process becomes $y = 1.1u$), the first dissipation inequality will be invalid; if a fault that* *causes the process gain to decrease (e.g., the process becomes $y = 0.9u$), the second dissipation inequality will be invalid.*

For dynamic systems, the above simple dissipativity conditions are insufficient to capture the changes in the process dynamics when a fault occurs. The concept of QdF dissipativity is adopted. QdF supply rate allows for more information on system dynamics to be captured, including information related to gain and phase at different frequencies [18]. For example, for a single-input single-output linear system $G(z)$, a dissipativity condition with QSR-type of supply rate shown in (15) (with scalar $Q = -1$, $S = 0$ and $R = \gamma^2$) implies the $\mathcal{L}_\infty$ norm (an upper bound of $\mathcal{L}_2$ gain) of the system. In contrast, a corresponding QdF supply rate of

$$\phi_{\bar{g}}(\zeta, \eta) = \begin{pmatrix} -N^T(\zeta)N(\eta) & 0 \\ 0 & \gamma^2 d(\zeta)d(\eta)I \end{pmatrix}. \tag{17}$$

implies a frequency weighted $\mathcal{L}_\infty$ norm bound of $\|WG\|_\infty \leq \gamma$, where $W(z) = N(z)/d(z)$ (see [26] for details). If the weighting is chosen as $W(z) = G(z)^{-1}$ together with $\gamma = 1$, the dissipativity inequality with the supply rate given in (17) will be invalid if a fault causes the process gain to increase at *any* frequency. This is very important for fault detection and diagnosis applications as the QSR type supply rates can be too coarse for this purpose.

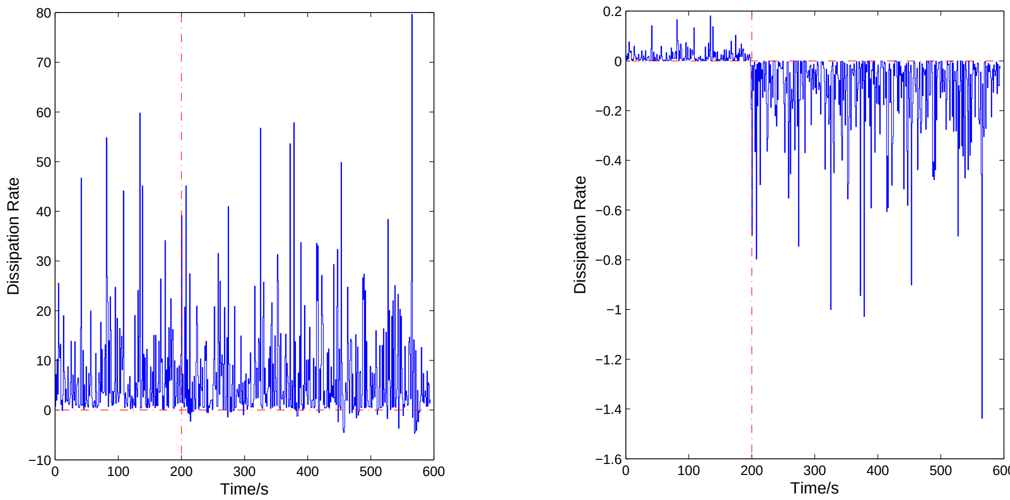
Example 2. *Consider a discrete process $G(z) = \frac{0.1+\Delta f_n}{z-0.9+\Delta f_d}$. For the fault-free case, $\Delta f_n = \Delta f_d = 0$ (with a steady-state gain of 1). A fault is modeled with $\Delta f_n = 0.1, \Delta f_d = 0$, i.e., $\tilde{G}_f(z) = \frac{0.2}{z-0.9}$ (with a steady-state gain of 2).*

The simplest form of supply rate that can identify changes in process gain is a QR supply rate (which is a traditional QSR supply rate with $S = 0$). In this simple example, one can choose the value of $Q = -1, R = 1.2^2$ such that the supply rate is as follows:

$$s(u, y) = -y^2 + 1.44u^2, \tag{18}$$

and the above supply rate implies an $\mathcal{L}_\infty$ norm ≤ 1.2 (gain increase to above 1.2 will be detected). The simulation result is shown in Figure 3a, with a fault occurring after 200s. As can be seen, some of the dissipation rates are negative which indicates the process with fault is no longer dissipative with respect to the original supply rate in (18). *However, the change of dissipation rate does not capture the fault effectively due to the fact that the QSR dissipativity only reflect the $\mathcal{L}_\infty$ norm of the system, which is too conservative for the fault detection and diagnosis purpose.*

To capture more process dynamic information, a QdF supply rate of the form in (17) *is*



(a) QSR Dissipativity

(b) QdF Dissipativity

Figure 3: Comparison of QSR and QdF dissipativity results of gain change

chosen with $W(z) = \frac{n(z)}{d(z)} = \frac{z-0.9}{0.1}$*. The polynomial matrix is then given below:*

$$\phi(\zeta,\eta) = \begin{pmatrix} -(\zeta-0.9)(\eta-0.9) & 0 \\ 0 & \gamma^2 0.1^2 \end{pmatrix}, \tag{19}$$

where $\gamma = 1.2$*, which induces a supply rate of*

$$Q_\Phi(u_k, y_k) = -0.81y_k^2 + 1.8y_k y_{k+1} - y_{k+1}^2 - 0.0144u_k^2. \tag{20}$$

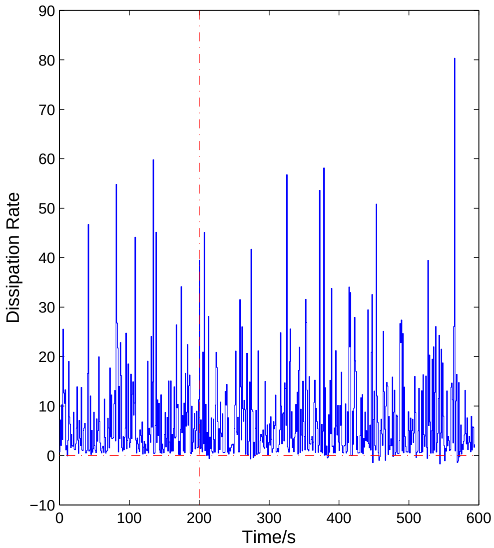
The coefficient matrix of above supply rate is:

$$\tilde{\phi} = \begin{pmatrix} -0.81 & 0.9 & 0 & 0 \\ 0.9 & -1 & 0 & 0 \\ 0 & 0 & 0.0144 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{21}$$

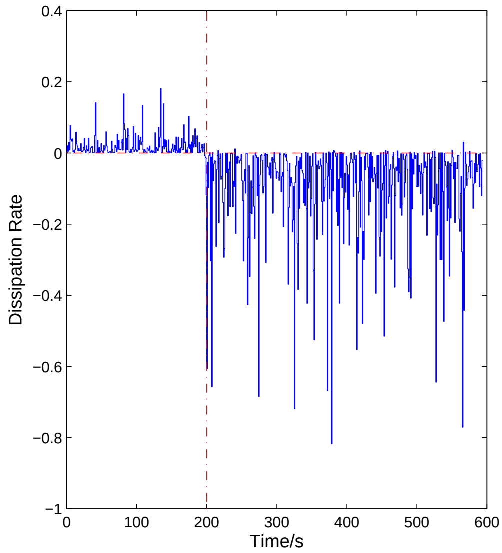
The result using QdF dissipativity is shown in Figure 3b. After the fault occurs, the dissipation rate becomes negative which indicates that process with fault is no longer dissipative with respect to the dissipativity property of original process. This shows that the QdF dissipativity can capture sufficient details of dynamic feature of the process for fault detection and diagnosis.

The next example shows that the change of a time constant of a process caused by a fault can also be reflected by the change in QdF dissipativity.

Example 3. *Consider the same discrete process in Example 2:* $G(z) = \frac{0.1+\Delta f_n}{z-0.9+\Delta f_d}$*, with another fault* $\Delta f_n = 0.08, \Delta f_d = 0.09$*, i.e.,* $\tilde{G}(z) = \frac{0.18}{z-0.81}$.



(a) QSR Dissipativity



(b) QdF Dissipativity

Figure 4: Comparison of QSR and QdF dissipativity results of time constant change

The same QSR and QdF supply rates in Example 2 are used in this example. The simulation results of QSR supply rate are shown in Figure 4a and the results of QdF dissipativity are shown in Figure 4b. It can be seen that fault detection based on the QSR dissipativity is ineffective in this example. This is because the QSR supply rate only implies an $\mathcal{L}_\infty$ norm bound which is still satisfied by the process with new fault. However, the QdF dissipativity, which implies a frequency weighted $\mathcal{L}_\infty$ norm bound, is no longer valid when the fault occurs, as the process with fault $\tilde{G}(z)$ has a gain increase in the high frequency range. The difference in gain between process with and without fault is shown in the Bode diagrams in Figure 5.

Theoretically, any multiplicative fault can be reflected by the change of a or some dissipativity property/properties. In general, the physical meanings of such dissipativity properties may not be explicit. To illustrate the point that the dissipativity property can be fault sensitive, we can choose the following dissipativity conditions representing the upper and lower bounds of the gain and phase of the nominal process at all frequencies, one of which will be invalid when any multiplicative fault occurs. Given a nominal process $G(z)$, choosing $W(z) = N(z)/d(z) = G(z)^{-1}$ we have $W(z)G(z) = I$, which should have gain to be 1 and a phase of 0 at all frequencies. A multiplicative fault (which causes changes in the input/output behavior through changes in process parameter) will cause the above one or both of the above conditions to change at certain frequencies, which means that one of the following conditions will be violated:

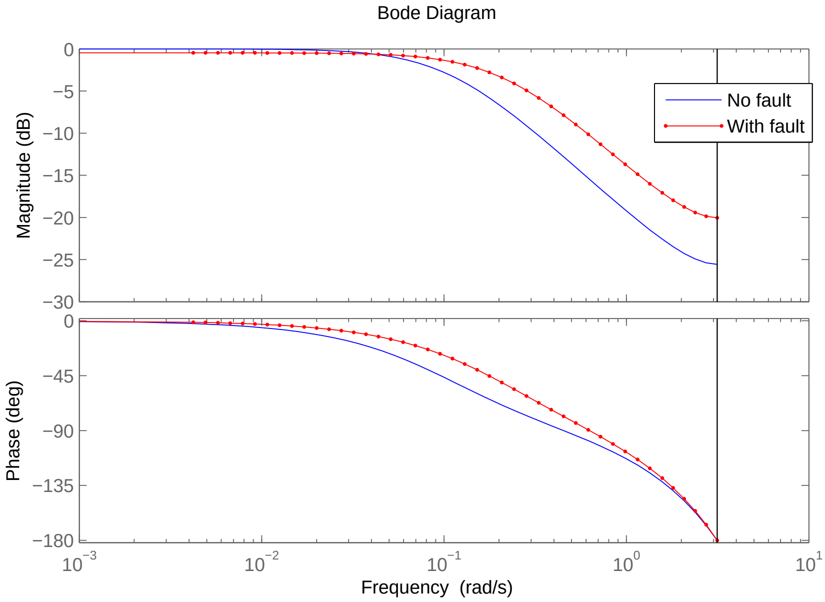


Figure 5: Bode plot of process with and without fault

- Upper bound of the gain: $\|WG\|_\infty \leq \gamma$, where $\gamma = 1 + \epsilon$, ϵ is a small number, which is implied by the following dissipativity supply rate:

$$\phi_{\bar{g}}(\zeta, \eta) = \begin{pmatrix} -N^T(\zeta)N(\eta) & 0 \\ 0 & \gamma^2 d(\zeta)d(\eta)I \end{pmatrix} \tag{22}$$

- Lower bound of the gain: $\min_\omega \sigma_{\min}(W(\omega)G(\omega)) \geq \gamma$, where $\gamma = 1-\epsilon$, ϵ is a small number, which is implied by the following dissipativity supply rate:

$$\phi_{\underline{\mathrm{g}}}(\zeta, \eta) = \begin{pmatrix} N^T(\zeta)N(\eta) & 0 \\ 0 & -\gamma^2 d(\zeta)d(\eta)I \end{pmatrix} \tag{23}$$

- Bound on phase lead: $(1 - z^{-1})W(z)G(z)$ is passive (i.e., a differentiator with $W(z)G(z)$ is passive), which is implied by the following dissipativity supply rate:

$$\phi_{\bar{p}}(\zeta, \eta) = \begin{pmatrix} 0 & 1/2(\zeta - 1)N^T(\zeta)d(\eta) \\ 1/2d(\zeta)(\eta - 1)N(\eta) & \bar{R} \end{pmatrix} \tag{24}$$

- Bound on phase lag: $\frac{1}{1-z^{-1}}W(z)G(z)$ is passive (i.e., an integrator with $W(z)G(z)$ is passive), which is implied by the following dissipativity supply rate:

$$\phi_{\underline{\mathrm{p}}}(\zeta, \eta) = \begin{pmatrix} 0 & 1/2N^T(\zeta)(\eta - 1)d(\eta) \\ 1/2(\zeta - 1)d(\zeta)N(\eta) & \underline{\mathrm{R}} \end{pmatrix} \tag{25}$$

Note that $W(z)$ can be non-causal and/or unstable [22, 23, 27]. Generally, the supply rates should be chosen so that they are sensitive to the changes of process models caused by faults. More general form of QdF (full matrix of $\phi(\zeta, \eta)$) supply rate can capture phase and gain information simultaneously. The detailed systematic algorithm of determination of supply rates is discussed in next section.

3 Fault Diagnosis using Dissipativity Properties

The dissipativity based fault diagnosis approach is developed in this section. The proposed fault diagnosis approach is comprised of two stages: (1) Process dissipativity shaping for fault diagnosis (offline), where the dissipativity property is shaped to be sensitive to a particular fault; (2) Online fault diagnosis, where multiple dissipativity conditions associated to different faults are evaluated. This is a tractable method as most of the computational burdens are shifted to the offline design, allowing for simple method for online fault diagnosis. As shown in Figure 2, in the online stage, model of the process, observer and residual generator are not required.

In this section, a fault diagnosis approach is developed to determine if a particular fault (with prior knowledge) has occurred. In proposed dissipativity based framework, multiplicative faults are considered, whose models are in the following form [9]:

$$\begin{aligned} x_{k+1} &= (A_0 + \Delta A_q)x_k + (B_0 + \Delta B_q)u_k \\ y_k &= (C_0 + \Delta C_q)x_k + (D_0 + \Delta D_q)u_k \end{aligned} \tag{26}$$

where A_0, B_0, C_0, D_0 represent the nominal model, different faults (fault 1 to fault m) are described by ΔA_q, ΔB_q, ΔC_q, ΔD_q ($q = 1 \ldots m$). Denote $A_q = A_0 + \Delta A_q$, and B_q, C_q, D_q are similarly defined. Subsequently, process model with fault q, are described by A_q, B_q, C_q, D_q, referred to as Σ_q. The proposed fault diagnosis approach is developed based on process models with and without faults (without disturbance and noise). The effects of disturbances and noises are addressed by using a robust threshold developed in Section 4.

3.1 Process Dissipativity Shaping for Fault Diagnosis

As we discussed earlier, a given process can have many dissipativity properties (the dissipativity property is not unique), each of which describes different aspects of its process dynamics. The basic idea of dissipativity based fault diagnosis is as follows: to identify certain fault, one dissipativity property of the normal process is shaped such that it is *not* valid when a fault (to be identified) occurs but is valid when all other faults occur. That is, the dissipativity property is shaped to be fault-selective, which is presented in Proposition 1. The shaping is formulated as an LMI problem, the numerical solution of which provides the coefficient matrices of the QdF supply rate and storage function. The optimization problems with convex objective and LMI constrains can be efficiently solved using any semi-definite programming tools [28].

Proposition 1. *Consider a process which is modeled by Σ_0 when it is normal and is described by Σ_q when one fault q occurs. Each $\Sigma_q, (q = 0, \ldots, m)$ is described as below:*

$$\begin{aligned} x_{k+1} &= A_q x_k + B_q u_k \\ y_k &= C_q x_k + D_q u_k. \end{aligned} \tag{27}$$

Define a fault index $j, j = 1, \ldots, m$ to denote the fault needs to be identified. The j-th dissipativity of the system Σ_0 consists of the storage function Q_Ψ^j and supply rate Q_Φ^j with corresponding coefficient matrices $\tilde{\psi}^j$ and $\tilde{\phi}^j$ partitioned as $\tilde{\psi}^j = \begin{pmatrix} \tilde{X}^j & \tilde{Y}^j \\ \tilde{Y}^{j\top} & \tilde{Z}^j \end{pmatrix}$ and $\tilde{\phi}^j = \begin{pmatrix} \tilde{Q}^j & \tilde{S}^j \\ \tilde{S}^{j\top} & \tilde{R}^j \end{pmatrix}$, respectively. The j-th dissipativity property of the system, which is satisfied by $\Sigma_q, (q = 0, \ldots, m, q \neq j)$ but not satisfied by Σ_j, can be determined by solving the following LMI feasibility problem with decision variables $\tilde{Q}^j, \tilde{S}^j, \tilde{R}^j, \tilde{X}^j, \tilde{Y}^j, \tilde{Z}^j$:

$$\mathbb{T}_q^j > 0,\ 0 \leq q \leq m,\ q \neq j \tag{28}$$

$$\mathbb{T}_q^j \not> 0, q = j \tag{29}$$

$$\tilde{\psi}^j > 0 \tag{30}$$

where (29) *implies that $\mathbb{T}_q^j$ is indefinite and $\mathbb{T}_q^j = \begin{pmatrix} T_{q,11}^j & T_{q,12}^j \\ T_{q,12}^{j\top} & T_{q,22}^j \end{pmatrix}$, $q = 0, \ldots, m$ with*

$$\begin{aligned} T_{q,11}^j =& \hat{C}_q^\top (\tilde{Q}^j - \hat{X}^j) \hat{C}_q \\ T_{q,12}^j =& \hat{C}_q^\top (\tilde{Q}^j - \hat{X}^j) \hat{D}_q + \hat{C}_q^\top (\tilde{S}^j - \hat{Y}^j) \\ T_{q,22}^j =& \hat{D}_q^\top (\tilde{Q}^j - \hat{X}^j) \hat{D}_q + \hat{D}_q^\top (\tilde{S}^j - \hat{Y}^j) + (\tilde{S}^j - \hat{Y}^j)^\top \hat{D}_q + (\tilde{R}^j - \hat{Z}^j) \end{aligned} \tag{31}$$

$$\hat{C}_q = \begin{pmatrix} C_q \\ C_q A_q \\ \vdots \\ C_q A_q^N \end{pmatrix}, \hat{D}_q = \begin{pmatrix} D_q & 0 & \cdots & 0 & 0 \\ C_q B_q & D_q & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_q A_q^{N-1} B_q & C_q A_q^{N-2} B_q & \cdots & C_q B_q & D_q \end{pmatrix} \tag{32}$$

and $\hat{X}^j = \begin{pmatrix} 0 & 0 \\ 0 & \tilde{X}^j \end{pmatrix} - \begin{pmatrix} \tilde{X}^j & 0 \\ 0 & 0 \end{pmatrix}$, $\hat{Y}^j, \hat{Z}^j$ are similarly defined.

Proof. Inequality (30) implies that the storage function Q_Ψ is positive definite. For each Σ_q, taking forward steps for the state space equation in (27), we have

$$\begin{pmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+N} \end{pmatrix} = \begin{pmatrix} C_q \\ C_q A_q \\ \vdots \\ C_q A_q^N \end{pmatrix} x_k + \begin{pmatrix} D_q & 0 & \cdots & 0 & 0 \\ C_q B_q & D_q & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_q A_q^{N-1} B_q & C_q A_q^{N-2} B_q & \cdots & C_q B_q & D_q \end{pmatrix} \begin{pmatrix} u_k \\ u_{k+1} \\ \vdots \\ u_{k+N} \end{pmatrix}, \tag{33}$$

which can be represented in a compact form:

$$\hat{y}_k = \hat{C}_q x_k + \hat{D}_q \hat{u}_k. \tag{34}$$

The QdF dissipativity inequality (2) can be rewritten as:

$$\begin{pmatrix} \hat{y}_k \\ \hat{u}_k \end{pmatrix}^\top \begin{pmatrix} \hat{X} & \hat{Y} \\ \hat{Y}^\top & \hat{Z} \end{pmatrix} \begin{pmatrix} \hat{y}_k \\ \hat{u}_k \end{pmatrix} \leq \begin{pmatrix} \hat{y}_k \\ \hat{u}_k \end{pmatrix}^\top \begin{pmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^\top & \tilde{R} \end{pmatrix} \begin{pmatrix} \hat{y}_k \\ \hat{u}_k \end{pmatrix}. \tag{35}$$

From (34) and (35), we can derive

$$\begin{pmatrix} x_k \\ \hat{u}_k \end{pmatrix}^\top \begin{pmatrix} T_{q,11} & T_{q,12} \\ T_{q,12}^\top & T_{q,22} \end{pmatrix} \begin{pmatrix} x_k \\ \hat{u}_k \end{pmatrix} > 0. \tag{36}$$

Since x_k and $\hat{u}_k$ are independent, system Σ_q is dissipative if $\mathbb{T}_q > 0$ and it does not satisfy the above dissipativity condition if $\mathbb{T}_q \not> 0$. □

The fault index j is introduced for the convenience of distinguishing between different dissipativity properties of the (same) normal process model, as the dissipativity properties are not unique. For example, denote $j = 1$ when shaping dissipativity property for diagnosis of fault 1, which leads to $\mathbb{T}_q^1 > 0$ $(q = 0, 2 \dots m)$ and $\mathbb{T}_1^1 \not> 0$.

A sufficient condition for the indefinite LMI condition in (29) which can be numerically solved is

$$\mathrm{tr}(\mathbb{T}_q) < 0. \tag{37}$$

The LMI conditions in (28)-(30) are solved simultaneously to determine the dissipativity property of the nominal process which is not valid when certain fault occurs. The dissipation inequality of which can be checked numerically by calculating the dissipation rate in real time using the plant data for fault diagnosis (as detailed in Section 3.2). Based on Proposition 1, the dissipativity shaping approach for fault diagnosis of fault j is stated as follows.

Problem 1 (Dissipativity Shaping for Fault Diagnosis)**.** *Consider a process which is modeled by Σ_0 when it is normal and is described by Σ_q when fault q occurs (Σ_q includes nominal model* *and multiplicative fault ΔA_q, ΔB_q, ΔC_q, ΔD_q). The dissipativity properties shaped for fault diagnosis of Fault j can be determined by solving the following problem:*

$$\begin{aligned}
\min_{\substack{\tilde{Q}^j,\tilde{S}^j,\tilde{R}^j,\\ \tilde{X}^j,\tilde{Y}^j,\tilde{Z}^j,\alpha^j}} \quad & \alpha^j \\
\text{s.t.} \quad & \tilde{\psi}^j > 0, \\
& \mathbb{T}_q^j > 0,\ q = 0,\ldots,m,\ q \neq j, \\
& \mathrm{tr}(\mathbb{T}_j^j) < 0, \\
& \mathbb{T}_j^j - \alpha^j I < 0.
\end{aligned} \tag{38}$$

The idea behind above LMIs can be understood as follows. To *isolate* the j-th fault, the dissipativity property of the nominal process is determined such that: (1) the dissipation inequality is satisfied by the process models with all faults *except* the j-th fault, with positive definite $\mathbb{T}_q^j, q = 0,\ldots,m,\ q \neq j$ (i.e., the "common" dissipativity property of the nominal process and the models with all faults *except* the j-th fault); (2) the dissipation inequality is not satisfied by the process model with the j-th fault, with non-positive definite $\mathbb{T}_q^j, q = j$ (and the minimized maximum eigenvalue). A total of j sets of dissipativity properties can be determined using above problem formulation. Therefore, when fault j occurs, the process with fault j will no longer be dissipative with respect to the j-th dissipativity property. Consequently, the j-th dissipativity inequality is not satisfied, which will be indicated by negative dissipation rate Q_Δ^j.

In Section 2.2, we illustrated the change of dissipativity properties caused by faults (Examples 1-3) using special types of supply rates which are associated with process gain (the static, $\mathcal{L}_\infty$norm and frequency dependent gain conditions). The approach presented in Problem 1 gives a general dissipativity property with no limit on the type of supply rates (e.g., captures the process dynamic features of both gain and phase), which may lead to more effective fault diagnosis.

Successful shaping in the design stage leads to each dissipativity property being "fault-selective" (only sensitive to certain fault). A numerical solution to Problem 1 provides the coefficient matrices of the j-th dissipativity property of the process, which are used for fault diagnosis by checking the validity of the dissipativity inequality. Some faults may cause similar changes of process dynamic features, therefore hard to classify under the general problem formulation in Problem 1. If above LMI problem does not have a feasible solution, an alternative way is to adopt the "divide-and-conquer" approach. That is, break all possible faults into several subgroups, then formulate the LMIs similar to Problem 1 and forms a hierarchical diagnosis system.

The proposed approach diagnoses faults based on the change of process input-output relationships captured by the dissipativity properties. The storage function and supply rate are optimized to be most sensitive to the faults to be diagnosed by solving the LMI problem in Problem 1. The detectability of the fault using the proposed approach depends on the feasibility of this LMI problem. To illustrate this point, consider the following process nominal model:

$$x_{k+1} = \begin{pmatrix} 1.4 & -0.48 \\ 1 & 0 \end{pmatrix} x_k + \begin{pmatrix} 0.5 \\ 0 \end{pmatrix} u_k,\ y_k = \begin{pmatrix} 0 & 0.4 \end{pmatrix} x_k \tag{39}$$

with a faulty model:

$$x_{k+1} = \begin{pmatrix} 1.4 & -0.48 \\ 1 & 0 \end{pmatrix} x_k + \begin{pmatrix} 0.25 \\ 0 \end{pmatrix} u_k,\ y_k = \begin{pmatrix} 0 & 0.8 \end{pmatrix} x_k. \tag{40}$$

In this case, while the state-space models of the process in normal and faulty conditions are different, the input-output relationships are unchanged. As such, Problem 1 is infeasible.

3.2 Online Dissipativity Inequality Evaluation for Fault Diagnosis

Online fault diagnosis is performed by checking different dissipativity inequalities at each sample instance k. While in (4) the extended variables are defined with future inputs and outputs u_{k+N} and y_{k+N}, the history of process input and output is used in process monitoring. Therefore, if the QdF order is N, the extended input and output at sample instance $k-N$ are used for fault diagnosis at sample instance k, denoted as $\check{u}$ and $\check{y}$ respectively:

$$\begin{aligned} \check{u}_k &= \begin{pmatrix} u_{k-N}^\top & u_{k+1-N}^\top & \dots & u_k^\top \end{pmatrix}^\top \\ \check{y}_k &= \begin{pmatrix} y_{k-N}^\top & y_{k+1-N}^\top & \dots & y_k^\top \end{pmatrix}^\top . \end{aligned} \tag{41}$$

At every sampling step k, the dissipation rate for diagnosis of each fault j can be represented as

$$Q_\Delta^j(u_k, y_k) = \begin{pmatrix} \check{y}_k \\ \check{u}_k \end{pmatrix}^\top \begin{pmatrix} \tilde{Q}^j - \hat{X}^j & \tilde{S}^j - \hat{Y}^j \\ \tilde{S}^{j\top} - \hat{Y}^{j\top} & \tilde{R}^j - \hat{Z}^j \end{pmatrix} \begin{pmatrix} \check{y}_k \\ \check{u}_k \end{pmatrix}. \tag{42}$$

Multiple dissipation rates $Q_\Delta^j, j = 1, \dots m$ are calculated simultaneously. If the calculated dissipation rate Q_Δ^j is negative we can draw the conclusion that this system is not dissipative with respect to the j-th dissipativity property and infer the j-th fault has occurred. As the dissipativity properties are shaped to be sensitive to individual faults, multiple dissipativity conditions need to be verified to perform fault diagnosis, as shown below:

Procedure 1 (Online fault diagnosis)**.** *Online fault diagnosis are performed with the following steps:*

1. Collect inputs and outputs at sampling step k into extended form $\breve{y}_k, \breve{u}_k$;

2. Calculate multiple dissipation rate $Q^j_\Delta(u_k, y_k),\ j = 1, \ldots, m$ for sampling step k;

3. If the dissipation rate $Q^j_\Delta(u_k, y_k)$ is smaller than threshold r, the associate fault j is identified;

4. Move to the next sampling step $k = k + 1$, and go to Step 1.

Online evaluation of dissipation rate can be easily carried out with equation (42), with a minor computational effort. The choice of dissipativity threshold r is an important issue as it will affect the accuracy of fault diagnosis. Theoretically the threshold can be $r = 0$ if there is no disturbance or noise. The false alarm rate of the proposed diagnosis approach with Gaussian noises and disturbances is analysed in Section 4. A dynamic threshold is also developed to reduce false alarms.

4 Dissipativity Threshold for Fault Diagnosis

Noises and disturbances are inevitable in practice and may lead to false alarms. As such, it is important to analyze the effects of noises and disturbances on the dissipativity of process models so that a threshold of dissipation rate r can be employed to improve the robustness of the proposed fault diagnosis approach. In this section, only Gaussian-type of process disturbances and measure noises are considered. For future work, it is possible to extend it to disturbance signals over a certain frequency range by choosing dissipativity property which is less sensitive to disturbances.

In general, the threshold can be time-varying, in the form of $r(k)$ where k is the time step such that

$$\Omega(k) = \begin{cases} 1, & Q_\Delta(u_k, y_k) < r(k) \\ 0, & \text{otherwise} \end{cases} \tag{43}$$

where $\Omega(k) = 1$ indicates that fault is diagnosed at time step k. The dissipation rate $Q_\Delta(u_k, y_k)$ is defined by (14). For model (1) which is not perturbed by noises, theoretically the threshold function can be chosen as $r(k) = 0$. Here, we consider the nominal model Σ_0 perturbed by Gaussian disturbances/noises as follows:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + \overline{w}_k \\ y_k &= Cx_k + Du_k + \overline{v}_k \end{aligned} \tag{44}$$

where $\overline{w}_k, \overline{v}_k$ are disturbances and noises, respectively. Assume that the covariance matrices of $\overline{w}_k, \overline{v}_k$ are known as W, V respectively. Taking following linear transform

$$\overline{w}_k = Ew_k, \quad \overline{v}_k = Fv_k \tag{45}$$

where $W = E^TE, V = F^TF$ and $w_k(i)$ and $v_k(j)$ are jointly independent random variables with standard normal distribution $\mathcal{N}(0,1)$ for $k \geq 1, 1 \leq i \leq s, 1 \leq j \leq q$, (44) can be rewritten as follows:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Ew_k \\ y_k &= Cx_k + Du_k + Fv_k. \end{aligned} \tag{46}$$

The dissipation rate for model with uncertainty can be represented as follows:

$$Q_\Delta(\hat{u}_k, \tilde{y}_k) = \begin{pmatrix} \tilde{y}_k \\ \hat{u}_k \end{pmatrix}^T \begin{pmatrix} Q_\delta & S_\delta \\ S_\delta^T & R_\delta \end{pmatrix} \begin{pmatrix} \tilde{y}_k \\ \hat{u}_k \end{pmatrix} \tag{47}$$

where $Q_\delta = \tilde{Q} - \hat{X}, S_\delta = \tilde{S} - \hat{Y}, R_\delta = \tilde{R} - \hat{Z}$, and $\tilde{y}_k$ is the extended output trajectory of the noise-perturbed model (46) with initial state x_k and extended input trajectory $\hat{u}_k$. The notation $\hat{y}_k$ denotes the extended output trajectory of the nominal model (1) with initial state x_k and input trajectory $\hat{u}_k$. Trajectory $\tilde{y}_k$ is actually the trajectory $\hat{y}_k$ perturbed by noises w, v, which has the following representation:

$$\tilde{y}_k = \hat{y}_k + G\hat{z}_k \tag{48}$$

where $\hat{z}_k = \begin{pmatrix} \hat{w}_k^T & \hat{v}_k^T \end{pmatrix}^T$ and $G = \begin{pmatrix} \hat{E} & \hat{F} \end{pmatrix}$ with

$$\hat{E} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ CE & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{N-1}E & CA^{N-2}E & \cdots & 0 \end{pmatrix}, \quad \hat{F} = \begin{pmatrix} F & 0 & \cdots & 0 \\ 0 & F & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F \end{pmatrix}. \tag{49}$$

The dissipation rate $Q_\Delta(\hat{u}_k, \tilde{y}_k)$ can be reformulated as follows:

$$\begin{aligned} Q_\Delta(\hat{u}_k, \tilde{y}_k) &= \begin{pmatrix} \hat{y}_k + G\hat{z}_k \\ \hat{u}_k \end{pmatrix}^T \begin{pmatrix} Q_\delta & S_\delta \\ S_\delta^T & R_\delta \end{pmatrix} \begin{pmatrix} \hat{y}_k + G\hat{z}_k \\ \hat{u}_k \end{pmatrix} \\ &= Q_\Delta(\hat{u}_k, \hat{y}_k) + 2[\hat{u}_k^T S_\delta + \hat{y}_k^T Q_\delta] G\hat{z}_k + \hat{z}_k^T G^T Q_\delta G\hat{z}_k \\ &= Q_\Delta(\hat{u}_k, \hat{y}_k) + 2[\hat{u}_k^T S_\delta + \tilde{y}_k^T Q_\delta] G\hat{z}_k - \hat{z}_k^T G^T Q_\delta G\hat{z}_k \\ &= Q_\Delta(\hat{u}_k, \hat{y}_k) + H_k\hat{z}_k - \hat{z}_k^T J\hat{z}_k \end{aligned} \tag{50}$$

where

$$H_k = 2[\hat{u}_k^T S_\delta + \tilde{y}_k^T Q_\delta]G, \ J = G^T Q_\delta G. \tag{51}$$

The term $Q_\Delta(\hat{u}_k, \tilde{y}_k)$ is the dissipation rate for nominal model and $J \leq 0$. The term $H_k \hat{z}_k$ is a random variable with $\mathcal{N}(0, \|H_k\|_2)$-distribution where $\|\cdot\|_2$ denotes the Euclidean norm of a vector. The term $-\hat{z}_k^T J \hat{z}_k$ is a random variable with $\chi^2(-\operatorname{tr}(J))$-distribution. Theoretically, the probability of false alarms for a given threshold can be determined based on the following proposition:

Proposition 2. *Consider a nominal model* (27) *with a dissipation rate given by* (47)*. Assume the following threshold function:*

$$r(k) = -\lambda \|H_k\|_2 - \mu \operatorname{tr}(J) \tag{52}$$

where constants $\lambda > 0$, and $\mu > 0$ are tuning parameters. The extended input and output trajectories $\hat{u}_k$ and $\tilde{y}_k$ are generated by noise-perturbed model (46)*. Then the probability of $Q_\Delta(\hat{u}_k, \tilde{y}_k) < r(k)$ (false alarms) is upper bounded by*

$$\mathbb{P}\{Q_\Delta(\hat{u}_k, \tilde{y}_k) < r(k)\} \leq \int_0^\infty \int_{-\infty}^{r(k)-x} \frac{x^{a-1} e^{-(x+y^2/\sigma_k^2)/2}}{2^a \sqrt{2\pi}\Gamma(a)\sigma_k} \, dxdy \tag{53}$$

where $a = -\operatorname{tr}(J)/2$ and $\sigma_k = \|H_k\|_2$.

Proof. The probability density function of random variables $x = -\hat{z}_k^T J \hat{z}_k$ and $y = H_k \hat{z}^k$ are given by $f(x) = \frac{x^{a-1} e^{-x/2}}{2^a \Gamma(a)}$ and $g(y) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{y^2}{2\sigma_k^2}}$, respectively. Since $Q_\Delta(\hat{u}_k, \tilde{y}_k) \geq 0$, then the upper bound of a false alarm can be estimated as:

$$\begin{aligned} \mathbb{P}\{Q_\Delta(\hat{u}_k, \tilde{y}_k) < r(k)\} &\leq \mathbb{P}\{H_k \hat{z}_k - \hat{z}_k^T J \hat{z}_k < r(k)\} \\ &= \int_0^\infty \int_{-\infty}^{r(k)-x} \frac{x^{a-1} e^{-(x+y^2/\sigma_k^2)/2}}{2^a \sqrt{2\pi}\Gamma(a)\sigma_k} \, dxdy. \end{aligned} \tag{54}$$

□

In theory, the parameters of λ and μ in the threshold function (52) can be determined by iteratively evaluating (53) such that a required (low) probability of false alarms is achieved. By assuming $H_k \hat{z}_k$ and $-\hat{z}_k^T J \hat{z}_k$ are independent random variables, λ and μ can be chosen according to the empirical rules for normal distribution and p-value table for χ^2-distribution (as in [29]), which will give a rough upper bound of the false alarm events. A larger λ and a smaller μ imply stronger confidence level, but may also lead to more "missed alarm" events (false negative errors).

5 Illustrative Example

To illustrate the proposed approach, a simple case study is presented in this section. A heat exchanger [11] as shown in Figure 6 is studied. In this process, steam mass flow $\dot{m}_s$ is measured

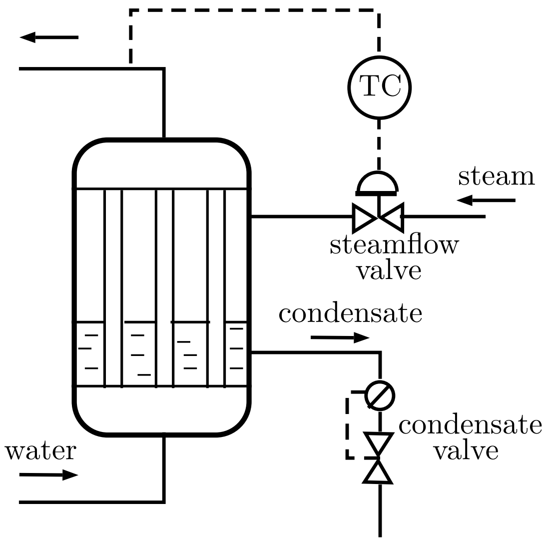


Figure 6: A Heat Exchanger Example[11]

as input variable, and the outlet fluid temperature ϑ_{fo} is measured as output variable, the inlet fluid temperature is denoted as ϑ_{fi}. The operating point is:

$$\dot{m}_f = 3000\ kg/h;\ \dot{m}_s = 50\ kg/h;\ \vartheta_{fi} = 60°C;\ \vartheta_{fo} \approx 70°C \tag{55}$$

The system is linearized around the operating point. The approximate transfer function is

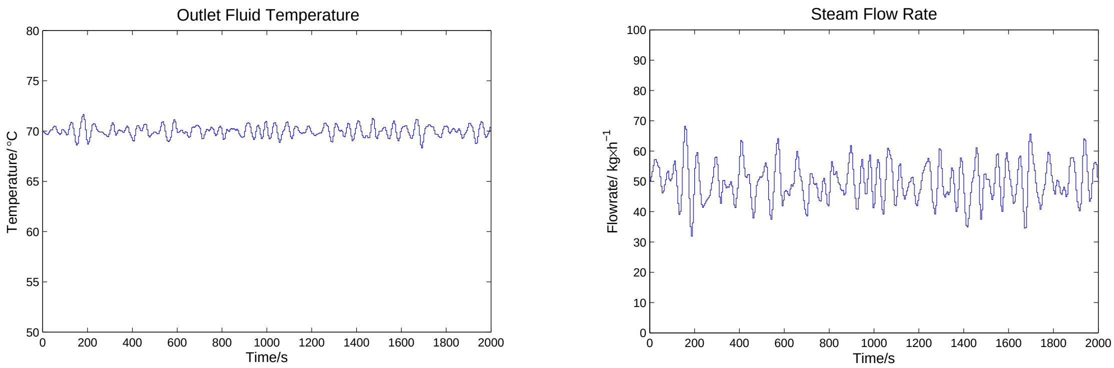
$$\tilde{G}(s) = \frac{\Delta\vartheta_{fo}(s)}{\Delta\dot{m}_s(s)} = \frac{K_s}{(1+T_{1s}s)(1+T_{2s}s)} \tag{56}$$

with

$$K_s = \frac{r}{\dot{m}_f c_f},\ T_{1s} = \frac{1}{v_f}(1+\frac{A_w\rho_w c_w}{A_f\rho_f c_f}),\ T_{2s} = \frac{A_w\rho_w c_w}{\alpha_{wf}U_f}\frac{1}{1+\frac{A_w\rho_w c_w}{A_f\rho_f c_f}}, \tag{57}$$

where the parameters and subscripts are listed as follows:

Parameters		Subscripts	
A	cross-sectional area	f	fluid
c	specific heat capacity	s	steam
$m, \dot{m}$	mass, mass flowrate	w	wall
r	evaporation heat	i	inlet
U	periphery of one tube	o	outlet
v	velocity in the tube		
α	heat transfer coefficient		
ϑ	temperature		
ρ	density		



(a) Outlet Temperature

(b) Steam Flow Rate

Figure 7: Normal Operating Condition

There are disturbances in inlet water flow rate $\dot{m}_f$, a controller is implemented to regulate the outlet temperature. One fault in the heat exchanger is the presence of air (inert gas) in steam space, later referred to as Fault 1. And another fault, closed condensate valve, are later referred to as Fault 2. In the simulation studies, it is assume that the process is operating normally in the beginning and then one of the faults occurs after 600 seconds. Under normal operating condition, the outlet water temperature and steam flow rate are shown in Figures 7a and 7b, respectively.

Take steam flow $\dot{m}_s$ as input and outlet water temperature ϑ_{fo} as output, the transfer function is converted to a discrete state space model denoted as Σ_0 is shown below:

$$\begin{aligned} x_{k+1} &= A_0 x_k + B_0 u_k \\ y_k &= C_0 x_k \end{aligned} \tag{58}$$

where

$$A_0 = \begin{pmatrix} 1.1676 & -0.6675 \\ 0.5 & 0 \end{pmatrix}, \ B_0 = \begin{pmatrix} 0.25 \\ 0 \end{pmatrix}, \ C_0 = \begin{pmatrix} 0.067 & 0.093 \end{pmatrix} \tag{59}$$

The state space models of process with two faults are as follows:

Fault 1 (inert gas in steam space), described by Σ_1:

$$\begin{aligned} x_{k+1} &= A_1 x_k + B_1 u_k \\ y_k &= C_1 x_k \end{aligned} \tag{60}$$

where

$$A_1 = A_0 + \Delta A_1, \ B_1 = B_0 + \Delta B_1, \ C_1 = C_0 + \Delta C_1 \tag{61}$$

and $\Delta A_1, \Delta B_1, \Delta C_1$ represent the multiplicative fault in the process. For simplicity reason, parameters of process with Fault 1 are presented directly, $\Delta A_1, \Delta B_1, \Delta C_1$ are omitted (e.g., in this case $\Delta C_1 = (0.0482 \;\; 0.0524)$).

$$A_1 = \begin{pmatrix} 1.0045 & -0.5045 \\ 0.5 & 0 \end{pmatrix},\ B_1 = \begin{pmatrix} 0.25 \\ 0 \end{pmatrix},\ C_1 = \begin{pmatrix} 0.1152 & 0.1454 \end{pmatrix} \tag{62}$$

Fault 2 (closed condensate valve), described by Σ_2 (defined with A_2, B_2, C_2), are presented similar to Fault 1, with

$$A_2 = \begin{pmatrix} 0.9669 & -0.3756 \\ 0.5 & 0 \end{pmatrix},\ B_2 = \begin{pmatrix} 0.25 \\ 0 \end{pmatrix},\ C_2 = \begin{pmatrix} 0.1066 & 0.1224 \end{pmatrix} \tag{63}$$

Online monitoring of the process is performed by calculating dissipation rates at each sampling instance k using the dissipativity property Q^1_Φ, Q^1_Ψ and Q^2_Φ, Q^2_Ψ. In this example, the QdF order is chosen to be 3. The parameters of the dynamic thresholds are chosen as $\lambda = 1.2$, $\mu = 1$. From Proposition 2, the probability of false alarms is roughly upper bounded by 9.4%. However, from simulation results (as shown in Figure 8 and 9), the actual probability of false alarms is much lower.

For fault diagnosis, each of the dissipativity property Q^j_Φ and Q^j_Ψ are shaped such that they are associated to different Fault j. As shown in a following example:

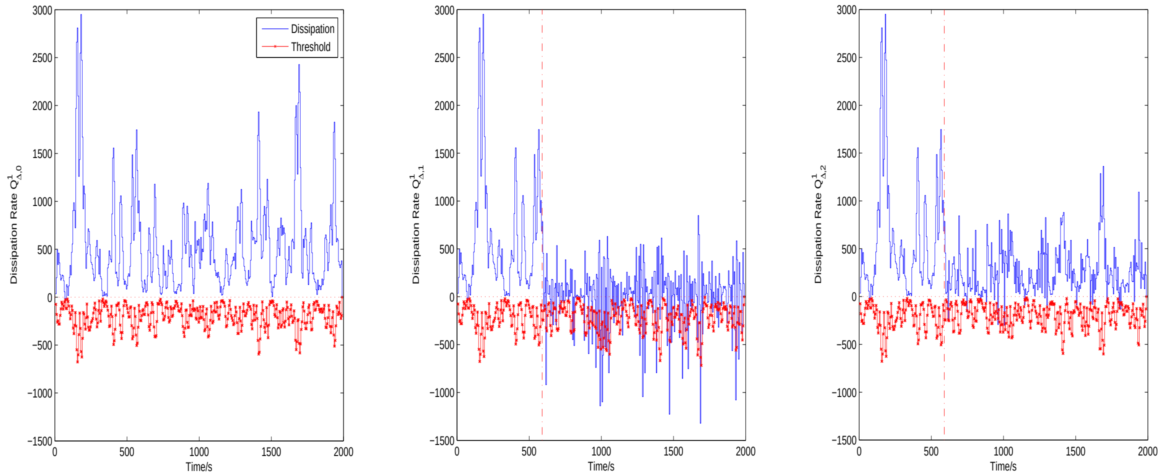
Example 4 (Fault diagnosis for Fault 1)**.** *Solve the following problem:*

$$\begin{aligned} \min_{\substack{\tilde{Q}^1, \tilde{S}^1, \tilde{R}^1, \\ \tilde{X}^1, \tilde{Y}^1, \tilde{Z}^1, \alpha^1}} \quad & \alpha^1 \\ \text{s.t.} \quad & \tilde{\psi}^1 > 0,\ \mathbb{T}^1_0 > 0,\ \mathbb{T}^1_2 > 0 \\ & \mathrm{tr}(\mathbb{T}^1_1) < 0,\ \mathbb{T}^1_1 - \alpha^1 I < 0. \end{aligned} \tag{64}$$

The coefficient matrices used for diagnosis of Fault 1, $\tilde{\psi}^1 = \begin{pmatrix} \tilde{X}^1 & \tilde{Y}^1 \\ \tilde{Y}^{1\top} & \tilde{Z}^1 \end{pmatrix}$, $\tilde{\phi}^1 = \begin{pmatrix} \tilde{Q}^1 & \tilde{S}^1 \\ \tilde{S}^{1\top} & \tilde{R}^1 \end{pmatrix}$ are determined from solving above LMI problem and provided in the Appendix. Matrices $\mathbb{T}^1_q, q = 0, 1, 2$ are defined using A_q, B_q, C_q as in Proposition 1. The coefficient matrices for diagnosis of Fault 2, $\tilde{\phi}^2$ and $\tilde{\psi}^2$ are also provided in the Appendix.

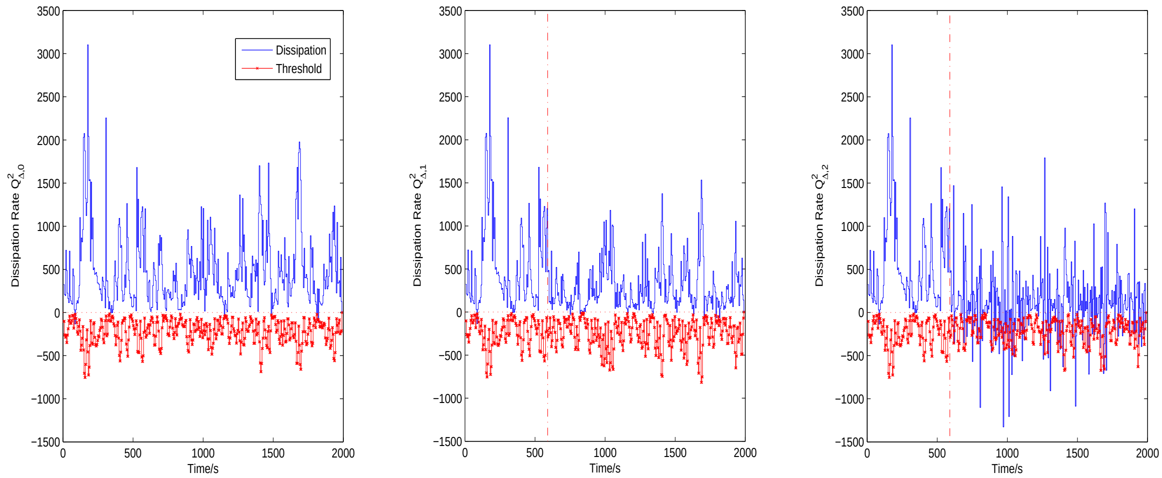
For fault diagnosis, two dissipation rates Q^1_Δ and Q^2_Δ are calculated simultaneously. Dissipation rates are calculated using input output trajectories and the coefficient matrices of the dissipativity property $\tilde{\phi}^1, \tilde{\psi}^1$ and $\tilde{\phi}^2, \tilde{\psi}^2$. Fault diagnosis results are shown in Figures 8 and 9. As shown in Figure 8b, only the dissipation rate $Q^1_{\Delta,1}$ (Q^1_Δ calculated based on the input/output trajectories from process with Fault 1) violates the threshold, while $Q^1_{\Delta,0}$ in Figure 8a (dissipation rate of the same dissipativity property but using input output of normal process Σ_0) and $Q^1_{\Delta,2}$ in Figure 8c are within the threshold. As such, Fault 1 is diagnosed.

Likewise, Fault 2 is diagnosed with dissipativity property shaped as following example:



(a) Nominal Process Σ_0 (b) Process with Fault 1 Σ_1 (c) Process with Fault 2 Σ_2

Figure 8: Diagnosis of Fault 1



(a) Nominal Process Σ_0 (b) Process with Fault 1 Σ_1 (c) Process with Fault 2 Σ_2

Figure 9: Diagnosis of Fault 2

Example 5 (Fault diagnosis for Fault 2)**.** *Solve the following problem:*

$$\begin{aligned} \min_{\substack{\tilde{Q}^2,\tilde{S}^2,\tilde{R}^2,\\ \tilde{X}^2,\tilde{Y}^2,\tilde{Z}^2,\alpha^2}} \quad & \alpha^2 \\ \text{s.t.} \quad & \tilde{\psi}^2 > 0,\ \mathbb{T}_0^2 > 0,\ \mathbb{T}_1^2 > 0 \\ & \operatorname{tr}(\mathbb{T}_2^2) < 0,\ \mathbb{T}_2^2 - \alpha^2 I < 0 \end{aligned} \tag{65}$$

Different fault indices (superscript) in Example 4 and Example 5 are adopted to indicate different dissipativity properties of the nominal process. For diagnosis of Fault 2, dissipation rate are calculated as (42) using $\tilde{\phi}^2$ and $\tilde{\psi}^2$ at each sample instance. As shown in Figure 9, only the dissipation rate $Q^2_{\Delta,2}$ in Figure 9c (Q^2_Δ calculated with the input/output trajectories from process with Fault 2) violate the threshold, while $Q^2_{\Delta,0}$ (Figure 9a) and $Q^2_{\Delta,1}$ (Figure 9b) are within the threshold.

6 Discussion and Conclusion

A fault diagnosis scheme based on dissipativity theory has been developed. For a given nominal process model, its dissipativity property is shaped to be sensitive to different faults (fault-selective). Fault diagnosis is performed by evaluating dissipativity inequalities derived from various dissipativity properties. A fault is diagnosed when its associated dissipation inequality is violated, leading to a simple fault diagnosis scheme, without observers/estimators (as depicted in Figure 2). The proposed dissipativity-based approach is suitable for handling multiplicative faults. While the process models are used to determine the dissipativity property of the process, they are not directly used in online fault diagnosis. As such the proposed approach can be classified as a model-based approach but not in the traditional way. The effects of noise and disturbance are handled by a dynamic threshold which can be tuned to achieve different confidence levels. It is also worth pointing out that the proposed approach can be applied to general (linear) processes as the dissipativity concept in this work is a process input-output property of any systems, not limited to physically dissipative systems.

The proposed dissipativity based method diagnoses fault based on changes in process input/output dynamic relationships, including any changes in process models such as parameters, coefficients or even system order (e.g., when a fault is demonstrated by a side reaction). As such the proposed approach is particularly suitable for diagnosing multiplicative faults.

The proposed approach in its current form does not optimize the fault diagnosis algorithm such that it is insensitive to unknown disturbances. While the proposed approach can deal with unknown disturbances using a threshold, rigorous results (dynamic thresholds) are limited to Gaussian noises and disturbances.

Future work will include the extension of the proposed approach to deal with bounded unknown disturbances. The disturbance model will be considered. The dissipativity of process model will be optimized to be sensitive to faults to be diagnosed but less sensitive to the disturbances.

Other possible research includes the extension of the dissipativity based fault diagnosis to plantwide processes with interactions between process units. The dissipativity property is very useful in interaction analysis for plantwide systems and the dissipativity theory is an ideal tool for analysis and control of large-scale systems, for example, in [23, 30, 31, 32]. It is also possible to incorporate the proposed fault diagnosis method with a dissipativity based fault-tolerant control (FTC) approach. An integrated fault diagnosis and fault tolerant control approach can provide efficient response to enhance fault recovery [33]. The FTCs can be designed such that the dissipativity of the closed-loop system is valid when faults occur to ensure the stability and performance, as shown promising from our recent work [34, 35].

Acknowledgements

This work is supported by Australian Research Council Discovery Project DP160101810. The first author wishes to acknowledge the financial support of UNSW University International Postgraduate Award. The second author wishes to acknowledge the financial support of the CSC (China Scholarship Council) Scholarship and UNSW Tuition Fee Scholarship.

References

[1] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, "A review of process fault detection and diagnosis: Part I Quantitative model-based methods," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 293–311, 2003.

[2] P. Frank, S. Ding, and T. Marcu, "Model-based fault diagnosis in technical processes," *Transactions of the Institute of Measurement and Control*, vol. 22, no. 1, pp. 57–101, 2000.

[3] P. M. Frank, "Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results," *Automatica*, vol. 26, no. 3, pp. 459–474, 1990.

[4] R. V. Beard, "Failure accomodation in linear systems through self-reorganization." Ph.D. dissertation, Massachusetts Institute of Technology, 1971.

[5] H. L. Jones, "Failure detection in linear systems." Ph.D. dissertation, Massachusetts Institute of Technology, 1973.

[6] E. Chow and A. Willsky, "Analytical redundancy and the design of robust failure detection systems," *IEEE Transactions on Automatic control*, vol. 29, no. 7, pp. 603–614, 1984.

[7] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniquespart i: Fault diagnosis with model-based and signal-based approaches," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3757–3767, 2015.

[8] S. Simani, C. Fantuzzi, and R. J. Patton, *Model-based Fault Diagnosis in Dynamic Systems Using Identification Techniques.* Springer Science & Business Media, 2013.

[9] S. Ding, *Model-based fault diagnosis techniques: design schemes, algorithms, and tools.* Springer Science & Business Media, 2008.

[10] J. J. Gertler, "Survey of model-based failure detection and isolation in complex plants," *IEEE Control Systems Magazine*, vol. 8, no. 6, pp. 3–11, 1988.

[11] R. Isermann, *Fault-diagnosis applications: model-based condition monitoring: actuators, drives, machinery, plants, sensors, and fault-tolerant systems.* Springer Science & Business Media, 2011.

[12] J. C. Willems, "Dissipative dynamical systems: Part I General theory," *Archive for Rational Mechanics and Analysis*, vol. 45, no. 5, pp. 321–351, 1972.

[13] R. Sepulchre, M. Janković, and P. Kokotović, *Constructive Nonlinear Control.* London: Springer-Verlag, 1997.

[14] B. E. Ydstie and A. A. Alonso, "Process systems and passivity via the Clausius-Planck inequality," *Systems & Control Letters*, vol. 30, no. 5, pp. 253–264, 1997.

[15] J. Bao and P. L. Lee, *Process control: the passive systems approach.* Springer Science & Business Media, 2007.

[16] O. J. Rojas, J. Bao, and P. L. Lee, "On dissipativity, passivity and dynamic operability of nonlinear processes," *Journal of Process Control*, vol. 18, no. 5, pp. 515–526, 2008.

[17] W. Chen, S. Ding, A. Khan, and M. Abid, "Energy based fault detection for dissipative systems," in *Control and Fault-Tolerant Systems (SysTol), Conference on.* IEEE, 2010, Conference Proceedings, pp. 517–521.

[18] J. C. Willems and H. L. Trentelman, "On quadratic differential forms," *SIAM Journal on Control and Optimization*, vol. 36, no. 5, pp. 1703–1749, 1998.

[19] C. Kojima and K. Takaba, "A generalized Lyapunov stability theorem for discrete-time systems based on quadratic difference forms," in *Proc. Decision and Control 2005 and European Control Conference 2005 (CDC-ECC'05).* IEEE, 2005, pp. 2911–2916.

[20] ——, "An LMI condition for asymptotic stability of discrete-time system based on quadratic difference forms," in *IEEE Conference on Computer-Aided Control Systems Design*, 2006.

[21] O. Kaneko and T. Fujii, "When is a storage function a state function in discrete time?" *SIAM journal on control and optimization*, vol. 42, no. 4, pp. 1374–1394, 2003.

[22] I. Pendharkar and H. K. Pillai, "Systems with sector bound nonlinearities: A behavioral approach," *Systems & Control Letters*, vol. 57, no. 2, pp. 112–122, 2008.

[23] M. J. Tippett and J. Bao, "Control of plant-wide systems using dynamic supply rates," *Automatica*, vol. 50, no. 1, pp. 44–52, 2014.

[24] J. C. Willems, "The behavioral approach to open and interconnected systems," *Control Systems, IEEE*, vol. 27, no. 6, pp. 46–99, 2007.

[25] O. Kaneko and T. Fujii, "Discrete-time average positivity and spectral factorization in a behavioral framework," *Systems & control letters*, vol. 39, no. 1, pp. 31–44, 2000.

[26] M. J. Tippett and J. Bao, "Distributed model predictive control based on dissipativity," *AIChE Journal*, vol. 59, no. 3, pp. 787–804, 2013.

[27] M. N. Belur and H. L. Trentelman, "Synthesis of dissipative behaviors with dynamics in the weighting matrices," in *Proceedings of the 16th International Symposium on Mathematical Theory of Networks & Systems, Leuven, Belgium*, 2004.

[28] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory.* SIAM, 1994, vol. 15.

[29] P. Westfall and K. S. Henning, *Understanding advanced statistical methods.* CRC Press, 2013.

[30] O. Rojas, R. Setiawan, J. Bao, and P. Lee, "Dynamic operability analysis of nonlinear process networks based on dissipativity," *AIChE Journal*, vol. 55, pp. 963–982, 2009.

[31] P. J. Moylan and D. J. Hill, "Stability criteria for large-scale systems," *Automatic Control, IEEE Transactions on*, vol. 23, no. 2, pp. 143–149, 1978.

[32] D. Hioe, J. Bao, and B. E. Ydstie, "Dissipativity analysis for networks of process systems," *Computers & Chemical Engineering*, vol. 50, pp. 207–219, 2013.

[33] P. Mhaskar, J. Liu, and P. D. Christofides, *Fault-tolerant process control: methods and applications.* Springer Science & Business Media, 2012.

[34] J. Bao, W. Z. Zhang, and P. L. Lee, "Passivity-based decentralized failure-tolerant control," *Industrial & Engineering Chemistry Research*, vol. 41, no. 23, pp. 5702–5715, 2002.

[35] J. Bao, W. Z. Zhang, and P. Lee, "Decentralized fault-tolerant control system design for unstable processes," *Chemical Engineering Science*, vol. 58, no. 22, pp. 5045–5054, 2003.

Appendix A Dissipativity Properties in Illustrative Example

The supply rate used for fault diagnosis in Section 5 is provided below. The dissipativity properties for fault diagnosis of Fault 1 are derived from solving the LMIs as defined in Example 4. The supply rate coefficient matrices are in the following form

$$\tilde{Q}^1 = \begin{pmatrix} 2741.10 & -12340.1 & 18139.9 & -8284.64 \\ -12340.1 & 47762.5 & -51945.1 & 11610.8 \\ 18139.9 & -51945.1 & 13874.4 & 34892.7 \\ -8284.64 & 11610.8 & 34892.7 & -49888.5 \end{pmatrix}, \tag{66}$$

$$\tilde{S}^1 = \begin{pmatrix} -163.137 & -116.036 & 261.768 & 0.00006 \\ 657.734 & 722.292 & -579.226 & -0.00021 \\ -771.036 & -1404.06 & -368.657 & 0.00008 \\ 213.700 & 845.191 & 958.468 & -0.00003 \end{pmatrix}, \tag{67}$$

$$\tilde{R}^1 = \begin{pmatrix} 7.08691 & 9.66420 & -8.77291 & 0 \\ 9.66420 & 5.65613 & -21.9709 & 0.00001 \\ -8.77291 & -21.9709 & -13.6113 & 0.00001 \\ 0 & 0.00001 & 0.00001 & 6.28622 \end{pmatrix}. \tag{68}$$

The difference storage function coefficient matrices are

$$\hat{X}^1 = \begin{pmatrix} -4.16875 & -0.00016 & 0.00003 & 0 \\ -0.00016 & 0.00281 & -0.00011 & -0.00003 \\ 0.00003 & -0.00011 & -0.00363 & 0.00027 \\ 0 & -0.00003 & 0.00027 & 4.16957 \end{pmatrix}, \tag{69}$$

$$\hat{Y}^1 = \begin{pmatrix} -0.00007 & -0.00009 & 0 & 0 \\ 0.00006 & -0.00012 & 0.00017 & 0 \\ -0.00015 & -0.00002 & 0.00024 & -0.00009 \\ 0 & 0.00015 & -0.00004 & -0.00006 \end{pmatrix}, \tag{70}$$

$$\hat{Z}^1 = \begin{pmatrix} -4.16776 & 0.00001 & -0.00001 & 0 \\ 0.00001 & 0 & -0.00002 & 0.00001 \\ -0.00001 & -0.00002 & -0.00632 & 0.00001 \\ 0 & 0.00001 & 0.00001 & 4.17407 \end{pmatrix}. \tag{71}$$

The dissipativity properties for fault diagnosis of Fault 2 are derived from solving the LMIs defined in Example 5. The supply rate coefficient matrices are in the following form

$$\tilde{Q}^2 = \begin{pmatrix} -2992.09 & 10942.4 & -10450.9 & 1278.83 \\ 10942.4 & -28855.1 & 2310.44 & 24802.6 \\ -10450.9 & 2310.44 & 86625.7 & -99504.8 \\ 1278.83 & 24802.6 & -99504.8 & 88487.5 \end{pmatrix}, \tag{72}$$

$$\tilde{S}^2 = \begin{pmatrix} 157.842 & 168.525 & -113.632 & -0.00007 \\ -445.937 & -829.778 & -294.010 & 0.00057 \\ 137.835 & 1421.88 & 2106.36 & -0.00239 \\ 266.184 & -820.270 & -2100.47 & -0.00698 \end{pmatrix}, \tag{73}$$

$$\tilde{R}^2 = \begin{pmatrix} -8.21229 & -11.2498 & -2.03379 & 0 \\ -11.2498 & -0.59347 & 24.7836 & -0.00033 \\ -2.03379 & 24.7836 & 49.1902 & 0 \\ 0 & -0.00033 & 0 & 5.19460 \end{pmatrix}. \tag{74}$$

The difference storage function coefficient matrices are

$$\hat{X}^2 = \begin{pmatrix} -3.19201 & 0.00578 & 0.00486 & 0 \\ 0.00578 & -0.01275 & -0.01095 & -0.00486 \\ 0.00486 & -0.01095 & 0.00144 & 0.00517 \\ 0 & -0.00486 & 0.00517 & 3.20332 \end{pmatrix}, \tag{75}$$

$$\hat{Y}^2 = \begin{pmatrix} -0.00292 & -0.00204 & -0.00033 & 0 \\ -0.00209 & 0.00455 & 0.00432 & 0.00033 \\ -0.00161 & 0.00402 & 0.00523 & -0.00228 \\ 0 & 0.00161 & -0.00193 & -0.00685 \end{pmatrix}, \tag{76}$$

$$\hat{Z}^2 = \begin{pmatrix} -3.20022 & -0.00012 & 0.00034 & 0 \\ -0.00012 & -0.00347 & 0.00012 & -0.00034 \\ 0.00034 & 0.00012 & -0.22226 & 0 \\ 0 & -0.00034 & 0 & 3.42595 \end{pmatrix}. \tag{77}$$