Optimal robust fault detection of discrete-time LPV systems with measurement error-affected scheduling variables combining ZKF and pQP

Junbo Tan, Sorin Olaru, Feng Xu, Xueqian Wang, Bin Liang

Optimal Robust Fault Detection of Discrete-time LPV Systems with Measurement Error-affected Scheduling Variables Combining ZKF and pQP

Junbo Tan[1,2,3], Sorin Olaru[2], Feng Xu[3*], Xueqian Wang[3*] and Bin Liang[1]

[1]*Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, 100084 Beijing, P.R.China.*
[2]*Laboratory of Signals and Systems, Univ. Paris-Sud-CentraleSupelec-CNRS, Université Paris Saclay.*
[3]*Center for Artificial Intelligence and Robotics, Tsinghua Shenzhen International Graduate School, Tsinghua University, 518055 Shenzhen, P.R.China.*

SUMMARY

Optimal robust *state estimation* (SE) and *fault detection* (FD) methods of discrete-time *linear parameter varying* (LPV) systems with measurement error-affected scheduling variables are proposed under the boundedness assumption of system uncertainties. By using the weighted Frobenius norm of the generator matrix of SE zonotope to characterize the set size, the optimal observer gain can be computed by a *Zonotopic Kalman Filter* (ZKF) procedure for the purpose of observation. Meanwhile, by minimizing the influence of system uncertainties while maximizing that of faults on SE to enhance the sensitivity of FD, the optimal FD criterion is characterized based on an on-line fractional programming problem, which can be equivalently transformed into a *parametric quadratic programming* (pQP) problem. The pQP problem can be efficiently solved by searching the root of its nonlinear characteristic equation using secant method. In general, as long as sensors with sufficiently high precision are equipped to measure the scheduling variables, the bounds of measurement errors of scheduling variables can be less conservative than those direct bounds of scheduling variables, which can reduce the conservatism of FD or SE in this way. At the end of this paper, a case study based on a practical circuit model is used to illustrate the effectiveness of the proposed method.

Received . . .

KEY WORDS: Optimal Robust Faut Detection, State Estimation, LPV systems, Inexact Scheduling Variables.

1. INTRODUCTION

As industrial systems become more and more complex, there always exist faults occurrence during the operation of system. State estimation (SE) and fault detection (FD) are often key steps for the aim of advanced monitory and control requirements in many engineering applications [1][2]. Fault diagnosis techniques have attracted a great number of attentions since they play an important role in real-time detecting, isolating and estimating (identifying) the occurred faults in sensors, actuators or system plant itself [3] [4]. The model-based SE and FD rely on establishing an effective mathematical model for real physical systems by using differential and difference equations.

*Correspondence to: Feng Xu and Xueqian Wang
Email: tjb16@mails.tsinghua.edu.cn and {xu.feng, wang.xq}@sz.tsinghua.edu.cn

In recent years, there is a large volume of published studies dealing with LPV modeling methods since LPV systems can be considered as a representation of nonlinear systems but own a model structurally-similar with linear systems at each equilibrium point [5]. Some techniques for linear systems can be further used to handle sub-classes of nonlinear systems by means of LPV systems as a bridge collecting linear and nonlinear systems [6][7][8][9]. It is worthy to mention that most of works on SE and FD for LPV systems only consider perfectly known scheduling variables [10][11][12][13], while a more general situation is that the scheduling variables are often affected by the measurements noises and limited to the accuracy of equipped sensors [14].

In general, model-based SE and FD can be divided into two categories according to the different approaches dealing with system uncertainties. In particular, the stochastic approaches characterize system uncertainties using random variables to implement SE and FD [15][16], whose classical representation is Kalman Filtering technology. The deterministic approaches consider that system uncertainties are unknown and bounded to generate bounding sets for completion of SE and FD, such as interval observers, invariant-set methods, set-membership estimation and so on [2][17][18][19][20]. Since these set-theoretic methods consider the sets of system uncertainties and all possible situations of system states, the robustness of SE and FD can be guaranteed. [21] proposed a robust FD method for LPV systems using interval observer and zonotopes. The recent study in [22] combined the standard Kalman Filtering with zonotopic geometry to realize the optimal SE for linear time varying (LTV) systems based on the so-called ZKF procedure with a proof of robust convergence, which results in an explicit bridge between the zonotopic set-membership and the stochastic paradigms for Kalman Filtering by introducing the notion of covariation. Based on [22], [20] further considered the ZKF optimizing FD rather than SE for LTV systems by using generalized eigenvalues/eigenvectors technique, in which the main optimization problem is approximately transformed into maximizing the general Rayleigh quotient able to be solved by using generalized eigenvalues/eigenvectors. [23] used zonotopic unknown input observer (UIO) to implement robust fault detection and isolation for discrete-time LTV descriptor systems. R3-1

For a real physical system, system disturbances and measurement noises have an important effect on the results of SE and FD. The main idea behind most of robust SE and FD methods is to attenuate the impact of uncertainties as much as possible without providing any information of the estimation accuracy and detection effectiveness. FD filters should be designed such that the generated residual signal is sensitive to the faults while insensitive to system uncertainties[24]. Based on the work in [22], we further consider the optimal FD by maximizing the effect of faults on the size of the SE zonotope while minimizing that of system uncertainties to dramatically improve the performance of FD at the price of slightly sacrificing the accuracy of SE. Note that, compared with [20] using the generalized eigenvalues/eigenvectors to maximize the generalized Rayleigh quotient to obtain the approximately optimal gain of FD observer, our proposed method can obtain a strict global optimal solution for the gain design of FD observer. Meanwhile, the proposed method mainly deals with the optimal robust FD for LPV systems with measurement-error affected scheduling variables representing a wider practical application process compared to the LTV systems used in [20] and the ordinary LPV systems in [21]. Certainly, the computational burden of our proposed method is relatively higher than the generalized eigenvalues/eigenvectors proposed in [20]. However, the optimization problem of the proposed method to be solved at each step is just a convex quadratic programming problem, which has been able to be solved by many efficient t oolboxes, s uch as YALMIP [25], CVX [26] and so on. In this sense, the computation burden of our proposed method is also acceptable and tolerable. R3-1

For the clarity of this paper, the main contributions of this paper are summarized as follows:

- An optimal observer gain for observation purpose is computed by using the ZKF procedure for discrete-time LPV systems with measurement error-affected scheduling variables.
- Establish a fractional programming to maximize the effect of faults while minimize that of system uncertainties on SE to enhance the sensitivity of FD.
- Transform the fractional programming to the pQP problem to obtain optimal observer gain for FD purpose, which can be efficiently solved by searching a root for its characteristic equation using secant method.

The remainder of this paper is organized as follows. Section 2 mainly introduces a general model of discrete-time LPV systems affected by additive actuator and sensor faults and the corresponding design of FD observer considering measurement error-affected scheduling variables. The system behaviors and stability are analyzed and the ZKF-based optimal SE is proposed in Section 3. The optimal robust FD based on the pQP approach is further illustrated in Section 4. Simulation results are provided to show the effectiveness of the proposed method in Section 5. Finally, some conclusions are drawn in Section 6.

2. PLANT MODEL

This section introduces a general model of discrete-time LPV systems affected by additive actuator and sensor faults and the corresponding FD observer for this kind of LPV systems.

2.1. System Plant

The LPV plant under the effect of actuator and sensor faults is modeled as

$$x_{k+1} = A(\theta_k)x_k + B(\theta_k)u_k + G(\theta_k)f_k + E(\theta_k)w_k, \tag{1a}$$
$$y_k = C(\theta_k)x_k + D(\theta_k)u_k + H(\theta_k)s_k + P(\theta_k)v_k, \tag{1b}$$

where $A(\theta_k) \in \mathbb{R}^{n_x \times n_x}$, $B(\theta_k) \in \mathbb{R}^{n_x \times n_u}$, $G(\theta_k) \in \mathbb{R}^{n_x \times n_f}$, $E(\theta_k) \in \mathbb{R}^{n_x \times n_w}$, $C(\theta_k) \in \mathbb{R}^{n_y \times n_x}$, $D(\theta_k) \in \mathbb{R}^{n_y \times n_u}$, $H(\theta_k) \in \mathbb{R}^{n_y \times n_s}$ and $P(\theta_k) \in \mathbb{R}^{n_y \times n_v}$ are parametric matrices dependent on a scheduling vector $\theta_k \in \mathbb{R}^{n_\theta}$. k denotes the k-th discrete time instant. $x_k \in \mathbb{R}^{n_x}$ and $y_k \in \mathbb{R}^{n_y}$ denote the system state and output vectors, respectively. $u_k \in \mathbb{R}^{n_u}$, $w_k \in \mathbb{R}^{n_w}$ and $v_k \in \mathbb{R}^{n_v}$ denote the known inputs, the unknown inputs and the measurement noises, respectively. $f_k \in \mathbb{R}^{n_f}$ and $s_k \in \mathbb{R}^{n_s}$ represent the additive actuator and sensor faults, respectively. It is assumed that the unknown inputs w_k (including process disturbances, modeling errors, etc.) and the measurement noises v_k are bounded by zonotopes $\mathbf{W} = w^c \oplus M_w\mathbb{B}^{n_w} = \langle w_c, M_w \rangle$ and $\mathbf{V} = v^c \oplus M_v\mathbb{B}^{n_v} = \langle v^c, M_v \rangle$, respectively. The additive actuator faults f_k and sensor faults s_k are contained in zonotopes $\mathbf{F} = f^c \oplus M_f\mathbb{B}^{n_f} = \langle f_c, M_f \rangle$ and $\mathbf{S} = s^c \oplus M_s\mathbb{B}^{n_s} = \langle s^c, M_s \rangle$, respectively.

Assumption 2.1
The original system (1) is bounded-input, bounded-output (BIBO) stable. That is, given bounded input vector u_k, the system output y_k is also bounded.

It is assumed that the n_θ-dimensional scheduling vector θ_k is bounded by a polytopic hypercube $\mathbf{\Theta}$, i.e., $\theta_k \in \mathbf{\Theta}$, and θ_k^i is the i-th component of vector $\theta_k = [\theta_k^1, ..., \theta_k^i, ..., \theta_k^{n_\theta}]^T$. Therefore, a linear affine function $\Xi(\theta_k)$ of θ_k is also bounded by a polytopic set and can be written as the sum of vertex matrices of this set:

$$\Xi(\theta_k) = \sum_{i=1}^{\mathcal{N}} \lambda_i(\theta_k)\Xi_i, \tag{2}$$

where Ξ_i is the i-th vertex matrix of the set $\Xi(\mathbf{\Theta})$, $\mathcal{N} = 2^{n_\theta}$ is the number of vertex matrices and the weighting coefficients satisfy $\sum_{i=1}^{\mathcal{N}} \lambda_i(\theta_k) = 1, 0 \leq \lambda_i(\theta_k) \leq 1$. Here Ξ can represent matrices A, B, G, E, C, D, H and P in (1). In some existing works on SE and FD of LPV systems, it is often considered that the scheduling vector θ_k is perfectly measured. However, this is an unrealistic assumption because there always exists a measurement error between the actual value θ_k and the measured value $\hat{\theta}_k$, which is denoted as

$$\tilde{\theta}_k = \theta_k - \hat{\theta}_k, \tag{3}$$

where $\tilde{\theta}_k$ denotes the measurement error vector of sensors equipped in the real system to measure the scheduling variables. In general, the measurement errors of sensors are derived from their measurement precision and it is considered that the measurement precision of sensors used to obtain the bounds of measurement errors can be read from their performance specifications, which are R2-1

denoted as $\tilde{\theta}_k \in \tilde{\mathbf{\Theta}}$.

In general, the size of measurement error set $\tilde{\mathbf{\Theta}}$ of scheduling vector θ_k is much smaller than that of the polytopic hypercube $\mathbf{\Theta}$, based on which we are able to obtain less conservative results of SE or FD.

2.2. Design of FD Observer

We consider designing a Luenberger-structure FD observer for the discrete-time LPV system (1) as

$$\hat{x}_{k+1} = A(\hat{\theta}_k)\hat{x}_k + B(\hat{\theta}_k)u_k + L(\hat{\theta}_k)(y_k - \hat{y}_k) \tag{4a}$$

$$\hat{y}_k = C(\hat{\theta}_k)\hat{x}_k + D(\hat{\theta}_k)u_k, \tag{4b}$$

where $\hat{x}_k \in \mathbb{R}^{n_x}$ and $\hat{y}_k \in \mathbb{R}^{n_y}$ are the estimated state vector and output vector of (1), respectively. $\hat{\theta}_k = \left[\hat{\theta}_k^1, \hat{\theta}_k^2, ..., \hat{\theta}_k^{n_\theta}\right]^T$ denotes an actual measurement of θ_k with

$$\hat{\theta}_k \in \mathbf{\Theta},$$

where $\hat{\theta}_k$ denotes the measurement values of scheduling vector θ_k and $\hat{\theta}_k^i$ denotes the i-th component of $\hat{\theta}_k$. Note that, by projecting $\hat{\theta}_k$ into the bounding set $\mathbf{\Theta}$, we can obtain the measurement $\hat{\theta}_k$ confined inside $\mathbf{\Theta}$ (see [27] for more detailed explanations on this point). Therefore, the following polytopic decomposition

$$\Xi(\hat{\theta}_k) = \sum_{i=1}^{\mathcal{N}} \lambda_i(\hat{\theta}_k)\Xi_i$$

can be obtained. Here, it is considered that $\Xi(\theta_k)$ (or $\Xi(\hat{\theta}_k)$) is the affine function of θ_k (or $\hat{\theta}_k$). That is

$$\Xi(\theta_k) = \Xi^0 + \sum_{i=1}^{n_\theta} \Xi^i \theta_k^i, \Xi(\hat{\theta}_k) = \Xi^0 + \sum_{i=1}^{n_\theta} \Xi^i \hat{\theta}_k^i \tag{5}$$

where Ξ^0, Ξ^1, $\cdots$, and Ξ^{n_θ} are constant matrices.

By combining (5) and (3), we have the following results:

$$\Xi(\theta_k) = \Xi(\hat{\theta}_k) + \Xi(\tilde{\theta}_k). \tag{6}$$

where $\Xi(\tilde{\theta}_k) = \sum_{i=1}^{n_\theta} \Xi^i \tilde{\theta}_k^i$.

Remark 2.1

It must be emphasized that the notations Ξ_i in (2) and Ξ^i in (5) have different meanings. The former denotes the i-th vertex matrix of the hypercube set $\Xi(\mathbf{\Theta})$, while the latter denotes the coefficient matrix of the i-th component of θ_k (or $\hat{\theta}_k$) of linear form of $\Xi(\theta_k)$ (or $\Xi(\hat{\theta}_k)$).

For the convenience of illustration, we simplify the symbol $\Xi(\theta_k)$ ($\Xi(\hat{\theta}_k)$ or $\Xi(\tilde{\theta}_k)$) as Ξ_k^θ ($\Xi^{\hat{\theta}}{}_k$ or $\Xi_k^{\tilde{\theta}}$). For a clear explanation of the proposed methods, both the affine and polytopic forms of LPV system are used in the mathematical derivations of this paper. R2-2

3. SYSTEM BEHAVIORS

This section mainly derives the state-estimation-error dynamics and optimal SE set in the healthy situation by using a ZKF procedure.

3.1. Analysis of System Behaviors

With (1) and (4), the state-estimation error is defined as

$$e_k = x_k - \hat{x}_k, \tag{7}$$

whose dynamics can be further derived as

$$\begin{aligned} e_{k+1} =&(A_k^\theta - L_k^{\hat{\theta}} C_k^\theta)e_k + (A_k^{\tilde{\theta}} - L_k^{\hat{\theta}} C_k^{\tilde{\theta}})\hat{x}_k + (B_k^{\tilde{\theta}} - L_k^{\hat{\theta}} D_k^{\tilde{\theta}})u_k \\ &+ G_k^\theta f_k + E_k^\theta w_k - L_k^{\hat{\theta}} H_k^\theta s_k - L_k^{\hat{\theta}} P_k^\theta v_k. \end{aligned} \tag{8}$$

Moreover, the output estimation error (i.e., the residual signal) is defined as

$$r_k = y_k - \hat{y}_k = C_k^\theta e_k + C_k^{\tilde{\theta}} \hat{x}_k + D_k^{\tilde{\theta}} u_k + H_k^\theta s_k + P_k^\theta v_k. \tag{9}$$

We consider the dynamics (8) and (9) in the healthy situation without the effect of the fault signals f_k and s_k:

$$\bar{e}_{k+1} =(A_k^\theta - L_k^{\hat{\theta}} C_k^\theta)\bar{e}_k + (A_k^{\tilde{\theta}} - L_k^{\hat{\theta}} C_k^{\tilde{\theta}})\hat{x}_k + (B_k^{\tilde{\theta}} - L_k^{\hat{\theta}} D_k^{\tilde{\theta}})u_k + E_k^\theta w_k - L_k^{\hat{\theta}} P_k^\theta v_k, \tag{10a}$$

$$\bar{r}_k =C_k^\theta \bar{e}_k + C_k^{\tilde{\theta}} \hat{x}_k + D_k^{\tilde{\theta}} u_k + P_k^\theta v_k. \tag{10b}$$

The set version of the dynamics (10) can be derived as

$$\begin{aligned} \bar{\mathbf{E}}_{k+1} =&(A_k^{\hat{\theta}} + A^{\tilde{\mathbf{\Theta}}} - L_k^{\hat{\theta}} C_k^{\hat{\theta}} - L_k^{\hat{\theta}} C^{\tilde{\mathbf{\Theta}}})\bar{\mathbf{E}}_k \oplus (A^{\tilde{\mathbf{\Theta}}} - L_k^{\hat{\theta}} C^{\tilde{\mathbf{\Theta}}})\hat{x}_k \oplus (B^{\tilde{\mathbf{\Theta}}} - L_k^{\hat{\theta}} D^{\tilde{\mathbf{\Theta}}})u_k \\ &\oplus (E_k^{\hat{\theta}} + E^{\tilde{\mathbf{\Theta}}})\mathbf{W} \oplus (-L_k^{\hat{\theta}} P_k^{\hat{\theta}} - L_k^{\hat{\theta}} P^{\tilde{\mathbf{\Theta}}})\mathbf{V}, \end{aligned} \tag{11a}$$

$$\bar{\mathbf{R}}_k =(C_k^{\hat{\theta}} + C^{\tilde{\mathbf{\Theta}}})\bar{\mathbf{E}}_k \oplus C^{\tilde{\mathbf{\Theta}}}\hat{x}_k \oplus D^{\tilde{\mathbf{\Theta}}}u_k \oplus (P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}})\mathbf{V}. \tag{11b}$$

where $A^{\tilde{\mathbf{\Theta}}}$, $B^{\tilde{\mathbf{\Theta}}}$, $C^{\tilde{\mathbf{\Theta}}}$, $D^{\tilde{\mathbf{\Theta}}}$, $E^{\tilde{\mathbf{\Theta}}}$ and $P^{\tilde{\mathbf{\Theta}}}$ are the interval matrices containing $A_k^{\tilde{\theta}}$, $B_k^{\tilde{\theta}}$, $C_k^{\tilde{\theta}}$, $D_k^{\tilde{\theta}}$, $E_k^{\tilde{\theta}}$ and $P_k^{\tilde{\theta}}$, respectively. More detailed introduction about interval analysis and operation can be referred to [28]. Furthermore, the center-generator matrix form of (11) can be obtained as

$$\begin{aligned} M_{k+1}^{\bar{e}} = \Big[&(A_k^{\hat{\theta}} - L_k^{\hat{\theta}} C_k^{\hat{\theta}})\bar{M}_k^{\bar{e}} \quad \mathbf{seg}(\diamond(A^{\tilde{\mathbf{\Theta}}}\bar{M}_k^{\bar{e}})) \quad \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}}\bar{e}_k^c) \quad -L_k^{\hat{\theta}}\mathbf{seg}(\diamond(C^{\tilde{\mathbf{\Theta}}}\bar{M}_k^{\bar{e}})) \\ &-L_k^{\hat{\theta}}\mathbf{rad}(C^{\tilde{\Theta}}\bar{e}_k^c) \quad -L_k^{\hat{\theta}}\mathbf{rad}(C^{\tilde{\mathbf{\Theta}}}\hat{x}_k) \quad -L_k^{\hat{\theta}}\mathbf{rad}(D^{\tilde{\mathbf{\Theta}}}u_k) \quad \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}}\hat{x}_k) \\ &\mathbf{rad}(B^{\tilde{\mathbf{\Theta}}}u_k) \quad -L_k^{\hat{\theta}}\mathbf{seg}(\diamond((P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}})M_v)) \quad -L_k^{\hat{\theta}}\mathbf{rad}((P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}})v^c) \\ &\mathbf{seg}(\diamond((E_k^{\hat{\theta}} + E^{\tilde{\mathbf{\Theta}}})M_w)) \quad \mathbf{rad}((E_k^{\hat{\theta}} + E^{\tilde{\mathbf{\Theta}}})w^c) \Big], \end{aligned} \tag{12a}$$

$$\begin{aligned} \bar{e}_{k+1}^c = &(A_k^{\hat{\theta}} + \mathbf{mid}(A^{\tilde{\mathbf{\Theta}}}) - L_k^{\hat{\theta}} C_k^{\hat{\theta}} - L_k^{\hat{\theta}}\mathbf{mid}(C^{\tilde{\mathbf{\Theta}}}))\bar{e}_k^c + (\mathbf{mid}(A^{\tilde{\mathbf{\Theta}}}) - L_k^{\hat{\theta}}\mathbf{mid}(C^{\tilde{\mathbf{\Theta}}}))\hat{x}_k \\ &+ (\mathbf{mid}(B^{\tilde{\mathbf{\Theta}}}) - L_k^{\hat{\theta}}\mathbf{mid}(D^{\tilde{\mathbf{\Theta}}}))u_k - L_k^{\hat{\theta}}(P_k^{\hat{\theta}} + \mathbf{mid}(P^{\tilde{\mathbf{\Theta}}}))v^c + (E_k^{\hat{\theta}} + \mathbf{mid}(E^{\tilde{\mathbf{\Theta}}}))w^c, \end{aligned} \tag{12b}$$

$$\begin{aligned} M_k^{\bar{r}} = \Big[&\mathbf{seg}(\diamond((C_k^{\hat{\theta}} + C^{\tilde{\mathbf{\Theta}}})\bar{M}_k^{\bar{e}})) \quad \mathbf{rad}((C_k^{\hat{\theta}} + C^{\tilde{\mathbf{\Theta}}})\bar{e}_k^c) \quad \mathbf{rad}(C^{\tilde{\mathbf{\Theta}}}\hat{x}_k) \quad \mathbf{rad}(D^{\tilde{\mathbf{\Theta}}}u_k) \\ &\mathbf{seg}(\diamond((P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}})M_v)) \quad \mathbf{rad}((P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}})v^c) \Big], \end{aligned} \tag{12c}$$

$$\bar{r}_k^c = (C_k^{\hat{\theta}} + \mathbf{mid}(C^{\tilde{\mathbf{\Theta}}}))\bar{e}_k^c + \mathbf{mid}(C^{\tilde{\mathbf{\Theta}}})\hat{x}_k + \mathbf{mid}(D^{\tilde{\mathbf{\Theta}}})u_k + (P_k^{\hat{\theta}} + \mathbf{mid}(P^{\tilde{\mathbf{\Theta}}}))v^c \tag{12d}$$

where $M_k^{\bar{e}}$, $\bar{e}_k^c$, $M_k^{\bar{r}}$ and $\bar{r}_k^c$ are the generator matrices and center of $\bar{\mathbf{E}}_k$ and $\bar{\mathbf{R}}_k$, respectively. $\bar{M}_k^{\bar{e}}$ is the order-reduction matrix of $M_k^{\bar{e}}$, i.e., $\bar{M}_k^{\bar{e}} = \downarrow_{\lambda,W}(M_k^{\bar{e}})$ (Zonotope-order reduction is a contribution from [22]. Please refer Appendix B for more details.) R2-3

The FD criterion lies in real-time checking whether

$$\bar{r}_k \in \bar{\mathbf{R}}_k \tag{13}$$

holds or not. (13) is equivalent to the following relationship:

$$\mathbf{0} \in \{-\bar{r}_k\} \oplus \bar{\mathbf{R}}_k = \mathbf{R}_k = \langle c_{\mathbf{R}_k}, M_{\mathbf{R}_k} \rangle. \tag{14}$$

If there is a violation of (14), i.e., $\mathbf{0} \notin \mathbf{R}_k$, it is considered that faults have occurred in system (1). Otherwise, we still consider that the system (1) operates in the healthy situation.

3.2. Stability analysis

According to (10a), it can be found that the dynamics (10a) of the SE error $\bar{e}_k$ involves in the state estimation $\hat{x}_k$. Thus, in order to analyze the stability of the dynamics (10a), we integrate (4a) and (10a) to construct the following augmented dynamics:

$$\zeta_{k+1} = \Phi_0 \zeta_k + \Phi_1 \mu_k + \Phi_2 w_k + \Phi_3 v_k, \tag{15}$$

where

$$\Phi_0 = \begin{bmatrix} A_k^{\hat{\theta}} - L_k^{\hat{\theta}} C_k^{\hat{\theta}} & 0 \\ A_k^{\tilde{\theta}} - L_k^{\hat{\theta}} C_k^{\tilde{\theta}} & A_k^{\theta} - L_k^{\hat{\theta}} C_k^{\theta} \end{bmatrix}, \Phi_1 = \begin{bmatrix} L_k^{\hat{\theta}} & B_k^{\hat{\theta}} - L_k^{\hat{\theta}} D_k^{\hat{\theta}} \\ 0 & B_k^{\tilde{\theta}} - L_k^{\hat{\theta}} D_k^{\tilde{\theta}} \end{bmatrix},$$
$$\Phi_2 = \begin{bmatrix} 0 \\ E_k^{\theta} \end{bmatrix}, \Phi_3 = \begin{bmatrix} 0 \\ -L_k^{\hat{\theta}} P_k^{\theta} \end{bmatrix}, \zeta_k = \begin{bmatrix} \hat{x}_k \\ \bar{e}_k \end{bmatrix}, \mu_k = \begin{bmatrix} y_k \\ u_k \end{bmatrix}.$$

As we can see, the input μ_k of the augmented dynamics (15) is composed of the known system input u_k and the known system output y_k. Under Assumption 2.1, the augmented input μ_k is bounded. Meanwhile, considering that the uncertain factors w_k and v_k are also bounded, i.e., $w_k \in \mathbf{W}$ and $v_k \in \mathbf{V}$, they will not affect the BIBO stability of the augmented dynamics (15). Thus, we directly consider the stability of the following nominal system:

$$\bar{\zeta}_{k+1} = \Phi_0 \bar{\zeta}_k, \tag{16}$$

Here, stability results in Lemmas 3.1 and 3.2 taken from [29] and [30] are first recalled for the stability analysis of the dynamics (16).

Lemma 3.1
Considering the discrete LPV system:

$$x_{k+1} = \begin{bmatrix} \mathcal{A}_{11,k}^{\hat{\theta}} & \mathbf{0} \\ \mathcal{A}_{21,k}^{\hat{\theta}} & \mathcal{A}_{22,k}^{\hat{\theta}} \end{bmatrix} x_k \tag{17}$$

and assume that $\mathcal{A}_{11,k}^{\hat{\theta}}$ and $\mathcal{A}_{22,k}^{\hat{\theta}}$ are poly-quadratically stable[†], respectively, then the LPV system (17) is poly-quadratically stable.

Lemma 3.2
The dynamics $x_{k+1} = \mathcal{A}_k^{\hat{\theta}} x_k$ is poly-quadratically stable if and only if there exists symmetric definite matrices S_i, S_j and matrices G_i such that

$$\begin{bmatrix} G_i + G_i^T - S_i & * \\ \mathcal{A}_i G_i & S_j \end{bmatrix} \succ 0, \forall\, i,\ j = 1, 2,\ , N, \tag{18}$$

where the symbol $*$ denotes the transpose of $\mathcal{A}_i G_i$. In this case, the time-varying parameter dependent Lyapunov function for the stability is given as $V(x_k, \lambda_{i,k}^{\hat{\theta}}) = x_k^T P_k^{\hat{\theta}} x_k$ with $P_k^{\hat{\theta}} = \sum_{i=1}^{N} \lambda_{i,k}^{\hat{\theta}} S_i^{-1}$, $\sum_{i=1}^{N} \lambda_{i,k}^{\hat{\theta}} = 1$ and $0 \leq \lambda_{i,k}^{\hat{\theta}} \leq 1$.

[†]The expression on the stability of a matrix $\mathcal{A}_k^{\hat{\theta}}$ means the stability of the system $x_{k+1} = \mathcal{A}_k^{\hat{\theta}} x_k$.

Theorem 3.1
The dynamics (16) is poly-quadratically stable if there exist symmetric definite matrices S_i, S_j and matrices G_i such that

$$\begin{bmatrix} G_i + G_i^T - S_i & * \\ (A_i - L_t C_i)G_i & S_j \end{bmatrix} \succ 0 \tag{19}$$

$\forall\, i,\, j,\, t = 1,\, 2, \cdots,\, \mathcal{N}$ with the indices defined as $A_k^\theta = \sum_{i=1}^{\mathcal{N}} \lambda_{i,k}^\theta A_i$, $A_k^{\hat{\theta}} = \sum_{i=1}^{\mathcal{N}} \lambda_{i,k}^{\hat{\theta}} A_i$, $C_k^\theta = \sum_{i=1}^{\mathcal{N}} \lambda_{i,k}^\theta C_i$, $C_k^{\hat{\theta}} = \sum_{i=1}^{\mathcal{N}} \lambda_{i,k}^{\hat{\theta}} C_i$, and $L_k^{\hat{\theta}} = \sum_{t=1}^{\mathcal{N}} \lambda_{t,k}^{\hat{\theta}} L_t$, where the symbol $*$ denotes the transpose of $(A_i - L_t C_i)G_i$.

Proof
From Lemma 3.1, the stability condition of Φ_0 can be formulated into the simultaneous stability problems of $A_k^{\hat{\theta}} - L_k^{\hat{\theta}} C_k^{\hat{\theta}}$ and $A_k^\theta - L_k^{\hat{\theta}} C_k^\theta$. For the stability conditions of $A_k^{\hat{\theta}} - L_k^{\hat{\theta}} C_k^{\hat{\theta}}$, by using the results in Lemma 3.2, we can first derive the following stability conditions:

$$\begin{bmatrix} G_i + G_i^T - S_i & * \\ \Lambda_i G_i & S_j \end{bmatrix} \succ 0$$

with $\Lambda_i = A_i - L_k^{\hat{\theta}} C_i$ by decomposing $A_k^{\hat{\theta}} = \sum_{i=1}^{\mathcal{N}} \lambda_{i,k}^{\hat{\theta}} A_i$ and $C_k^{\hat{\theta}} = \sum_{i=1}^{\mathcal{N}} \lambda_{i,k}^{\hat{\theta}} C_i$ for $i, j = 1, 2, ..., \mathcal{N}$, where S_i and S_j are symmetric definite matrices with proper dimensions, and G_i are matrices with proper dimensions. Furthermore, with the polytopic decomposition of $L_k^{\hat{\theta}} = \sum_{t=1}^{\mathcal{N}} \lambda_{t,k}^{\hat{\theta}} L_t$, we can have $\Lambda_i = \sum_{t=1}^{\mathcal{N}} \lambda_{t,k}^{\hat{\theta}} (A_i - L_t C_i)$. Thus, we are able to further have

$$\begin{bmatrix} G_i + G_i^T - S_i & * \\ \Lambda_i G_i & S_j \end{bmatrix} = \sum_{t=1}^{\mathcal{N}} \lambda_{t,k}^{\hat{\theta}} \begin{bmatrix} G_i + G_i^T - S_i & * \\ (A_i - L_t C_i)G_i & S_j \end{bmatrix}$$

which implies that as long as (19) holds for $\forall t = 1, 2, ..., \mathcal{N}$, the matrix $\begin{bmatrix} G_i + G_i^T - S_i & * \\ \Lambda_i G_i & S_j \end{bmatrix}$ is also positive definite. Similarly, we can also derive the same stability conditions for $A_k^\theta - L_k^{\hat{\theta}} C_k^\theta$. Therefore, according to Lemmas 3.1 and 3.2, the dynamics (16) is quadratically stable. □

3.3. Optimal observer gain for robust SE

Regarding the design of the gain for SE purposes, the main goal is to reduce the effect of system uncertainties on SE. In the healthy situation, the estimated state set $\mathbf{X}_k$ at time instant k can be computed by $\mathbf{X}_k = \hat{x}_k \oplus \bar{\mathbf{E}}_k$. It can be found that the precision of SE is determined by the size of the set $\bar{\mathbf{E}}_k$. That means, if we can compute the gain matrix $L_k^{\hat{\theta}}$ to minimize the size of $\bar{\mathbf{E}}_k$, we can obtain the optimal SE. It is known that the shape and size of a zonotope are completely determined by its generator matrix. Because the weighted Frobenius norm of a matrix sufficiently considers its effects, we use the weighted Frobenius norm of the generator matrix of a zonotope to describe its size here (the details on the weighted Frobenius norm based size of zonotope is referred to [31] and [22]). As presented in (11a) and (12a), the size of zonotope $\bar{\mathbf{E}}_{k+1}$ can be directly assessed by the $M_{k+1}^{\bar{e}}$. Moreover, for brevity, we use the square of the weighted Frobenius norm of the generator matrix $M_{k+1}^{\bar{e}}$ to describe the size of $\bar{\mathbf{E}}_{k+1}$:

$$\| \langle \bar{e}_{k+1}^c, M_{k+1}^{\bar{e}} \rangle \|_{F,W}^2 = \| M_{k+1}^{\bar{e}} \|_{F,W}^2 = \mathbf{tr}(W M_{k+1}^{\bar{e}} M_{k+1}^{\bar{e}^T}).$$

Theorem 3.2
(Optimal Gain for Robust SE) Considering $\bar{e}_k \in \bar{\mathbf{E}}_k = \langle \bar{e}_k^c, \bar{M}_k^{\bar{e}} \rangle$ at time instant k, the optimal observer gain L_k^* minimizing the weighted Frobenius-norm based size of bounding zonotope $\bar{\mathbf{E}}_{k+1}$ at time instant $k+1$ is computed as

$$L_k^* = A_k^{\hat{\theta}} \mathcal{K}^*, \tag{20}$$

with

$$\begin{aligned}
\mathcal{K}^* =& \mathcal{L}\mathcal{S}^{-1}, \mathcal{L} = \mathcal{Q}_M C_k^{\hat{\theta}^T}, \mathcal{S} = C_k^{\hat{\theta}} \mathcal{Q}_M C_k^{\hat{\theta}^T} + \mathcal{Q}_{\tilde{\mathbf{\Theta}}} + \mathcal{Q}_v, \mathcal{Q}_M = \bar{M}_k^{\bar{e}} \bar{M}_k^{\bar{e}^T}, \\
\mathcal{Q}_v =& \mathbf{seg}(\diamond((P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}}) M_v)) \mathbf{seg}(\diamond((P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}}) M_v))^T, \\
\mathcal{Q}_{\tilde{\mathbf{\Theta}}} =& \mathbf{seg}(\diamond(C^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{\bar{e}})) \mathbf{seg}(\diamond(C^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{\bar{e}}))^T + \mathbf{rad}(C^{\tilde{\mathbf{\Theta}}} \bar{e}_k^c) \mathbf{rad}(C^{\tilde{\mathbf{\Theta}}} \bar{e}_k^c)^T + \mathbf{rad}(C^{\tilde{\mathbf{\Theta}}} \hat{x}_k) \mathbf{rad}(C^{\tilde{\mathbf{\Theta}}} \hat{x}_k)^T \\
& + \mathbf{rad}(D^{\tilde{\mathbf{\Theta}}} u_k) \mathbf{rad}(D^{\tilde{\mathbf{\Theta}}} u_k)^T + \mathbf{rad}((P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}}) v^c) \mathbf{rad}((P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}}) v^c)^T.
\end{aligned}$$

Proof
According to (12a), we have

$$\begin{aligned}
\|M_{k+1}^{\bar{e}}\|_{F,W}^2 =& \mathbf{tr}(W M_{k+1}^{\bar{e}} M_{k+1}^{\bar{e}^T}) \\
=& \mathbf{tr} \Big(W(A_k^{\hat{\theta}} - L_k^{\hat{\theta}} C_k^{\hat{\theta}}) \mathcal{Q}_M (A_k^{\hat{\theta}} - L_k^{\hat{\theta}} C_k^{\hat{\theta}})^T + W L_k^{\hat{\theta}} (\mathcal{Q}_v + \mathcal{Q}_{\tilde{\Theta}}) L_k^{\hat{\theta}^T} \\
& + W \mathbf{seg}(\diamond(A^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{\bar{e}})) \mathbf{seg}(\diamond(A^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{\bar{e}}))^T + W \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}} \bar{e}_k^c) \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}} \bar{e}_k^c)^T \\
& + W \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}} \hat{x}_k) \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}} \hat{x}_k)^T + W \mathbf{rad}(B^{\tilde{\mathbf{\Theta}}} u_k) \mathbf{rad}(B^{\tilde{\mathbf{\Theta}}} u_k)^T \\
& + W \mathbf{seg}(\diamond((E_k^{\hat{\theta}} + E^{\tilde{\mathbf{\Theta}}}) M_w)) \mathbf{seg}(\diamond((E_k^{\hat{\theta}} + E^{\tilde{\mathbf{\Theta}}}) M_w))^T \\
& + W \mathbf{rad}((E_k^{\hat{\theta}} + E^{\tilde{\mathbf{\Theta}}}) w^c) \mathbf{rad}((E_k^{\hat{\theta}} + E^{\tilde{\mathbf{\Theta}}}) w^c)^T \Big).
\end{aligned}$$

It can be found that $\|M_{k+1}^{\bar{e}}\|_{F,W}^2$ is convex with respect to $L_k^{\hat{\theta}}$. Moreover, $L_k^{\hat{\theta}}$ is a free parametric matrix and there are no constraints on $L_k^{\hat{\theta}}$. Therefore, we can directly compute the optimal value L_k^* to obtain the minimal value of $\|M_{k+1}^{\bar{e}}\|_{F,W}^2$ by solving the partial differential equation

$$\frac{\partial \|M_{k+1}^{\bar{e}}\|_{F,W}^2}{\partial L_k^{\hat{\theta}}} = 0. \tag{21}$$

We solve (21) and obtain $A_k^{\hat{\theta}} \mathcal{Q}_M C_k^{\hat{\theta}^T} = L_k^{\hat{\theta}} (C_k^{\hat{\theta}} \mathcal{Q}_M C_k^{\hat{\theta}^T} + \mathcal{Q}_{\tilde{\mathbf{\Theta}}} + \mathcal{Q}_v)$. By setting $\mathcal{K}^*$, $\mathcal{L}$ and $\mathcal{S}$ as in Theorem 3.2, since $\mathcal{S} = C_k^{\hat{\theta}} \mathcal{Q}_M C_k^{\hat{\theta}^T} + \mathcal{Q}_{\tilde{\mathbf{\Theta}}} + \mathcal{Q}_v$ is positive definite, an explicit solution can be obtained in Theorem 3.2. □

Remark 3.1
Solving the optimal parametric matrix L_k^* in Theorem 3.2 is actually equivalent to a Kalman filter procedure searching for the optimal Kalman gain matrix but with bounded disturbances and noises instead of stochastic disturbances and noises. Since here the zonotopic framework and Kalman filter are combined to solve the problem of optimal parametric matrix, we denote this optimization procedure as zonotopic Kalman filter (ZKF). More information on this point can be referred to [22], which shows that there is a strong analogy between the Kalman filter and ZKF where the usually Gaussian probability density functions are replaced by zonotopic sets.

4. OPTIMAL OBSERVER GAIN FOR ROBUST FD

In this section, we aim to design the optimal observer gain for the purposes of robust FD. Such a gain is computed to maximize the effect of faults while minimize the effect of disturbances to enhance the sensitivity of FD.

4.1. Structure of dynamics

Let us first consider the state-estimation-error dynamics (8) in the faulty situation, which can split into two independent dynamical systems:

$$\bar{e}_{k+1} = (A_k^{\theta} - L_k^{\hat{\theta}} C_k^{\theta}) \bar{e}_k + (A_k^{\tilde{\theta}} - L_k^{\hat{\theta}} C_k^{\tilde{\theta}}) \hat{x}_k + (B_k^{\tilde{\theta}} - L_k^{\hat{\theta}} D_k^{\tilde{\theta}}) u_k + E_k^{\theta} w_k - L_k^{\hat{\theta}} P_k^{\theta} v_k, \tag{22a}$$

$$e_{k+1}^{fs} = (A_k^{\theta} - L_k^{\hat{\theta}} C_k^{\theta}) e_k^{fs} + G_k^{\theta} f_k - L_k^{\hat{\theta}} H_k^{\theta} s_k, \tag{22b}$$

with $e_k = \bar{e}_k + e_k^{fs}$. It can be found that (22) allows the seperation of the effects of the disturbances (i.e., w_k and v_k) and the faults (i.e., f_k and s_k). Furthermore, (22a) is considered for the computation of the observer gain for SE purposes, which has been clearly illustrated in Section 3. Next, tuning the observer gain for FD purposes is done by combining (22a) and (22b). As a result, the advantage of the proposed FD scheme lies in the fact that not only the effect of the system disturbances is considered, but also the *relative* influence of the disturbances and the faults are characterized to formulate an optimization criterion to satisfy the optimal FD goal.

For the robustness of FD, we consider the set version of (22b)

$$\mathbf{E}_{k+1}^{fs} = (A_k^{\hat{\theta}} + A^{\tilde{\mathbf{\Theta}}} - L_k^{\hat{\theta}} C_k^{\hat{\theta}} - L_k^{\hat{\theta}} C^{\tilde{\mathbf{\Theta}}}) \mathbf{E}_k^{fs} \oplus (G_k^{\hat{\theta}} + G^{\tilde{\mathbf{\Theta}}}) \mathbf{F} \oplus (-L_k^{\hat{\theta}})(H_k^{\hat{\theta}} + H^{\tilde{\mathbf{\Theta}}}) \mathbf{S}. \tag{23}$$

Furthermore, the center-generator matrix form of (23) can be obtained as

$$\begin{aligned} M_{k+1}^{fs} = \Big[& (A_k^{\hat{\theta}} - L_k^{\hat{\theta}} C_k^{\hat{\theta}}) \bar{M}_k^{fs} \quad \mathbf{seg}(\diamond(A^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{fs})) \quad \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}} e_k^{fs,c}) \quad -L_k^{\hat{\theta}} \mathbf{seg}(\diamond(C^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{fs})) \\ & -L_k^{\hat{\theta}} \mathbf{rad}(C^{\tilde{\mathbf{\Theta}}} e_k^{fs,c}) \quad \mathbf{seg}(\diamond((G_k^{\hat{\theta}} + G^{\tilde{\mathbf{\Theta}}}) M_f)) \quad \mathbf{rad}((G_k^{\hat{\theta}} + G^{\tilde{\mathbf{\Theta}}}) f^c) \\ & -L_k^{\hat{\theta}} \mathbf{seg}(\diamond((H_k^{\hat{\theta}} + H^{\tilde{\mathbf{\Theta}}}) M_s)) \quad -L_k^{\hat{\theta}} \mathbf{rad}((H_k^{\hat{\theta}} + H^{\tilde{\mathbf{\Theta}}}) s^c) \Big], \end{aligned} \tag{24a}$$

$$\begin{aligned} e_{k+1}^{fs,c} = & (A_k^{\hat{\theta}} + \mathbf{mid}(A^{\tilde{\mathbf{\Theta}}}) - L_k^{\hat{\theta}} C_k^{\hat{\theta}} - L_k^{\hat{\theta}} \mathbf{mid}(C^{\tilde{\theta}})) e_k^{fs,c} + (G_k^{\hat{\theta}} + \mathbf{mid}(G^{\tilde{\mathbf{\Theta}}})) f^c \\ & - L_k^{\hat{\theta}} (H_k^{\hat{\theta}} + \mathbf{mid}(H^{\tilde{\mathbf{\Theta}}})) s^c, \end{aligned} \tag{24b}$$

where $M_k^{fs,c}$ and $e_k^{fs,c}$ are the generator matrix and center of $\mathbf{E}_k^{fs}$. $\bar{M}_k^{fs}$ is the order-reduction matrix of M_k^{fs}, i.e., $\bar{M}_k^{fs} = \downarrow_{\lambda,W} (M_k^{fs})$.

4.2. Optimal observer gain for FD purposes

Now, we concentrate on the design of optimal observer gain to improve the sensitivity of robust FD with respect to the system disturbances. We consider the following optimization criterion minimizing the influences of disturbances while maximizing that of faults:

$$J = \frac{\|\bar{\mathbf{E}}_{k+1}\|_{F,W}^2}{\|\mathbf{E}_{k+1}^{fs}\|_{F,W}^2}. \tag{25}$$

For the convenience of handling the optimization criterion (25), Theorem 4.1 gives an equivalent form of the FD optimization criterion (25).

Theorem 4.1
Considering $\xi = \mathbf{vec}(L_k^{\hat{\theta}})$, the FD optimization criterion (25) can be parameterized with ξ as

$$J(\xi) = \frac{\xi^T (\mathcal{Z}_1 \otimes W) \xi + \mathbf{vec}(\mathcal{Z}_2)^T \xi + \mathbf{tr}(\mathcal{Z}_3)}{\xi^T (\mathcal{Z}_4 \otimes W) \xi + \mathbf{vec}(\mathcal{Z}_5)^T \xi + \mathbf{tr}(\mathcal{Z}_6)}, \tag{26}$$

with

$$
\begin{aligned}
\mathcal{Z}_1 =& C_k^{\hat{\theta}} \bar{M}_k^{\bar{e}} \bar{M}_k^{\bar{e}^T} C_k^{\hat{\theta}^T} + \mathbf{seg}(\diamond(C^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{\bar{e}})) \mathbf{seg}(\diamond(C^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{\bar{e}}))^T + \mathbf{rad}(C^{\tilde{\mathbf{\Theta}}} \bar{e}_k^c) \mathbf{rad}(C^{\tilde{\mathbf{\Theta}}} \bar{e}_k^c)^T \\
&+ \mathbf{rad}(C^{\tilde{\mathbf{\Theta}}} \hat{x}_k) \mathbf{rad}(C^{\tilde{\mathbf{\Theta}}} \hat{x}_k)^T + \mathbf{rad}((P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}}) v^c) \mathbf{rad}((P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}}) v^c)^T \\
&+ \mathbf{rad}(D^{\tilde{\mathbf{\Theta}}} u_k) \mathbf{rad}(D^{\tilde{\mathbf{\Theta}}} u_k)^T + \mathbf{seg}(\diamond((P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}}) M_v)) \mathbf{seg}(\diamond((P_k^{\hat{\theta}} + P^{\tilde{\mathbf{\Theta}}}) M_v))^T, \\
\mathcal{Z}_2 =& -2W A_k^{\hat{\theta}} \bar{M}_k^{\bar{e}} \bar{M}_k^{\bar{e}^T} C_k^{\hat{\theta}^T}, \\
\mathcal{Z}_3 =& W A_k^{\hat{\theta}} \bar{M}_k^{\bar{e}} \bar{M}_k^{\bar{e}^T} A_k^{\hat{\theta}^T} + W \mathbf{seg}(\diamond(A^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{\bar{e}})) \mathbf{seg}(\diamond(A^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{\bar{e}}))^T + W \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}} \bar{e}_k^c) \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}} \bar{e}_k^c)^T \\
&+ W \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}} \hat{x}_k) \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}} \hat{x}_k)^T + W \mathbf{seg}(\diamond((E_k^{\hat{\theta}} + E^{\tilde{\mathbf{\Theta}}}) M_w)) \mathbf{seg}(\diamond((E_k^{\hat{\theta}} + E^{\tilde{\mathbf{\Theta}}}) M_w))^T \\
&+ W \mathbf{rad}(B^{\tilde{\mathbf{\Theta}}} u_k) \mathbf{rad}(B^{\tilde{\mathbf{\Theta}}} u_k)^T + W \mathbf{rad}((E_k^{\hat{\theta}} + E^{\tilde{\mathbf{\Theta}}}) w^c) \mathbf{rad}((E_k^{\hat{\theta}} + E^{\tilde{\mathbf{\Theta}}}) w^c)^T, \\
\mathcal{Z}_4 =& C_k^{\hat{\theta}} \bar{M}_k^{fs} \bar{M}_k^{fs^T} C_k^{\hat{\theta}^T} + \mathbf{seg}(\diamond(C^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{fs})) \mathbf{seg}(\diamond(C^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{fs}))^T + \mathbf{rad}(C^{\tilde{\mathbf{\Theta}}} e_k^{fs,c}) \mathbf{rad}(C^{\tilde{\mathbf{\Theta}}} e_k^{fs,c})^T \\
&+ \mathbf{seg}(\diamond((H_k^{\hat{\theta}} + H^{\tilde{\mathbf{\Theta}}}) M_s)) \mathbf{seg}(\diamond((H_k^{\hat{\theta}} + H^{\tilde{\mathbf{\Theta}}}) M_s))^T \\
&+ \mathbf{rad}(\diamond((H_k^{\hat{\theta}} + H^{\tilde{\mathbf{\Theta}}}) s^c)) \mathbf{rad}(\diamond((H_k^{\hat{\theta}} + H^{\tilde{\mathbf{\Theta}}}) s^c))^T, \\
\mathcal{Z}_5 =& -2W A_k^{\hat{\theta}} \bar{M}_k^{fs} \bar{M}_k^{fs^T} C_k^{\hat{\theta}^T}, \\
\mathcal{Z}_6 =& W A_k^{\hat{\theta}} \bar{M}_k^{fs} \bar{M}_k^{fs^T} A_k^{\hat{\theta}^T} + W \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}} e_k^{fs,c}) \mathbf{rad}(A^{\tilde{\mathbf{\Theta}}} e_k^{fs,c})^T \\
&+ W \mathbf{seg}(\diamond(A^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{fs})) \mathbf{seg}(\diamond(A^{\tilde{\mathbf{\Theta}}} \bar{M}_k^{fs}))^T + W \mathbf{seg}(\diamond(G_k^{\hat{\theta}} + G^{\tilde{\mathbf{\Theta}}}) M_f) \mathbf{seg}(\diamond(G_k^{\hat{\theta}} + G^{\tilde{\mathbf{\Theta}}}) M_f)^T \\
&+ W \mathbf{rad}((G_k^{\hat{\theta}} + G^{\tilde{\mathbf{\Theta}}}) f^c) \mathbf{rad}((G_k^{\hat{\theta}} + G^{\tilde{\mathbf{\Theta}}}) f^c)^T.
\end{aligned}
$$

Proof
Considering the FD optimization criterion (25), we have

$$
J = \frac{\|\bar{\mathbf{E}}_{k+1}\|_{F,W}^2}{\|\mathbf{E}_{k+1}^{fs}\|_{F,W}^2} = \frac{\mathbf{tr}(W M_{k+1}^{\bar{e}} M_{k+1}^{\bar{e}^T})}{\mathbf{tr}(W M_{k+1}^{fs} M_{k+1}^{fs^T})}. \tag{27}
$$

By setting $\mathcal{Z}_1$, $\mathcal{Z}_2$, $\mathcal{Z}_3$, $\mathcal{Z}_4$, $\mathcal{Z}_5$ and $\mathcal{Z}_6$ as in Theorem 4.1, based on (12a) and (24a), (27) is further computed as

$$
J = \frac{\mathbf{tr}(W L_k^{\hat{\theta}} \mathcal{Z}_1 L_k^{\hat{\theta}^T}) + \mathbf{tr}(\mathcal{Z}_2^T L_k^{\hat{\theta}}) + \mathbf{tr}(\mathcal{Z}_3)}{\mathbf{tr}(W L_k^{\hat{\theta}} \mathcal{Z}_4 L_k^{\hat{\theta}^T}) + \mathbf{tr}(\mathcal{Z}_5^T L_k^{\hat{\theta}}) + \mathbf{tr}(\mathcal{Z}_6)}. \tag{28}
$$

Then, according to the relationship between the trace function $\mathbf{tr}(\cdot)$ and the vectorization $\mathbf{vec}(\cdot)$ of matrices, we can obtain (26). □

Next, we consider the minimization of $J(\xi)$ in (26). For the convenience of analysis, we assume that ξ is bounded, i.e. $\|\xi\|_\infty \leq \bar{\xi}$, where $\bar{\xi}$ is a constant positive scalar. Note that, we introduce the boundedness assumption $\|\xi\|_\infty \leq \bar{\xi}$ just for the convenience of analysis. Since we can let $\bar{\xi}$ be enough large considering a real physical system, in this case, it is in fact equivalent to an unconstrained optimization problem. Therefore, we formulate the following fractional programming problem:

$$
\mathbf{min} \quad \frac{J_1(\xi)}{J_2(\xi)} \qquad \mathbf{s.t.} \quad \|\xi\|_\infty \leq \bar{\xi}, \tag{29}
$$

where $J_1(\xi) = \xi^T(\mathcal{Z}_1 \otimes W)\xi + \mathbf{vec}(\mathcal{Z}_2)^T\xi + \mathbf{tr}(\mathcal{Z}_3)$ and $J_2(\xi) = \xi^T(\mathcal{Z}_4 \otimes W)\xi + \mathbf{vec}(\mathcal{Z}_5)^T\xi + \mathbf{tr}(\mathcal{Z}_6)$. Since $J(\xi)$ has the rational fraction form and is not convex function with respect to ξ, the fractional optimization problem (29) could not be solved directly within a convex optimization framework. Therefore, we need to change the form of the fractional programming problem (29). We are interested in the relationship between the original fractional programming problem (29) and the

following parametric programming:

$$\mathbf{min} \quad J_1(\xi) - \beta J_2(\xi) \qquad \mathbf{s.t.} \quad \|\xi\|_\infty \leq \bar{\xi}, \tag{30}$$

where $\beta \in \mathbb{R}$ is a scalar parameter. As we know, $J_1(\xi) = \mathbf{tr}(WM_{k+1}^{\bar{e}}M_{k+1}^{\bar{e}^T}) > 0$ and $J_2(\xi) = \mathbf{tr}(WM_{k+1}^{fs}M_{k+1}^{fs^T}) > 0\,, \forall \xi \in \mathcal{X} = \{\xi \in \mathbb{R} | \, \|\xi\|_\infty \leq \bar{\xi}\}$. Thus, both optimization problems (29) and (30) have global optimal since $J_1(\xi)$ and $J_2(\xi)$ are continuous on the compact set $\mathcal{X}$.

Proposition 4.1
The following two statements are equivalent:

$$(\mathbf{I}) \quad \min_{\xi \in \mathcal{X}} \quad \frac{J_1(\xi)}{J_2(\xi)} = \beta, \qquad (\mathbf{II}) \quad \min_{\xi \in \mathcal{X}} \{J_1(\xi) - \beta J_2(\xi)\} = 0.$$

Proof
We first present $(\mathbf{I}) \Rightarrow (\mathbf{II})$. Let ξ_0 be a solution of the problem (29). Then we have

$$\beta = \frac{J_1(\xi_0)}{J_2(\xi_0)} \leq \frac{J_1(\xi)}{J_2(\xi)}, \quad \forall \xi \in \mathcal{X},$$

which implies

$$\begin{aligned} &J_1(\xi) - \beta J_2(\xi) \geq 0, \quad \forall \xi \in \mathcal{X}, \\ &J_1(\xi_0) - \beta J_2(\xi_0) = 0. \end{aligned}$$

Thus, we have $\min_{\xi \in \mathcal{X}} \{J_1(\xi) - \beta J_2(\xi)\} = 0$ and the minimum is attained at ξ_0.

Next we show $(\mathbf{II}) \Rightarrow (\mathbf{I})$. Let ξ_0 be a solution of (30). Then we have

$$0 = J_1(\xi_0) - \beta J_2(\xi_0) \leq J_1(\xi) - \beta J_2(\xi), \quad \forall \xi \in \mathcal{X}.$$

Dividing the above inequality by $J_2(\xi) > 0$, we obtain

$$\frac{J_1(\xi)}{J_2(\xi)} \geq \beta, \ \forall \xi \in \mathcal{X}, \ \text{and} \quad \frac{J_1(\xi_0)}{J_2(\xi_0)} = \beta.$$

Thus, we have $\min_{\xi \in \mathcal{X}} \frac{J_1(\xi)}{J_2(\xi)} = \beta$ and the minimum is attained at ξ_0. □

According to Proposition 4.1, it can be found that, if we can find a scalar parameter β such that the optimal value of the problem (30) is 0, then the optimal solution of (30) is also the optimal solution of (29). Therefore, we will consider solving the parametric programming problem (30) instead of the original fractional programming problem (29). We first define a function $S(\beta) : \mathbb{R} \to \mathbb{R}$ as

$$S(\beta) = \min_{\xi \in \mathcal{X}} \{J_1(\xi) - \beta J_2(\xi)\}. \tag{31}$$

Our purpose is to search a value β^* such that the characteristic equation $S(\beta^*) = 0$. In this case, β^* is the minimal value of the problem (29) based on Proposition 4.1. In order to solve the characteristic equation $S(\beta) = 0$, we should first capture the properties of the function $S(\beta)$. The following theorem illustrates some properties of the function $S(\beta)$.

Theorem 4.2
The function $S(\beta)$ is concave, continuous, strictly decreasing over $\mathbb{R}$. The characteristic equation $S(\beta) = 0$ has a unique solution.

Proof
Let us first consider the *hypograph* of $S(\beta)$:

$$\mathbf{hypo}(S) = \{(\beta, t) \in \mathbb{R}^2 | t \leq S(\beta)\}. \tag{32}$$

Combining the definition (31) of $S(\beta)$ and (32), we obtain

$$\mathbf{hypo}(S) = \left\{(\beta,t)\middle|\min_{\xi\in\mathcal{X}}\{J_1(\xi)-\beta J_2(\xi)\}\geq t\right\} = \left\{(\beta,t)\middle|J_1(\xi_i)-\beta J_2(\xi_i)\geq t, \forall \xi_i\in\mathcal{X}\right\},$$

which means that $\mathbf{hypo}(S)$ is the intersection of a family of halfspaces $\left\{(\beta,t)\middle|\beta J_2(\xi_i)+t\leq J_1(\xi_i)\right\}$, $\xi_i\in\mathcal{X}$. Thus, $\mathbf{hypo}(S)$ is a convex set and $S(\beta)$ is a concave function over $\mathbb{R}$. Since $S(\beta)$ is a concave mapping from $\mathbb{R}$ to $\mathbb{R}$, we can easily obtain the continuity of $S(\beta)$ over $\mathbb{R}$.

Furthermore, we choose $\beta_1<\beta_2$ and let $\xi_{\beta_1}\in\mathbf{argmin}_{\xi\in\mathcal{X}}\{J_1(\xi)-\beta_1 J_2(\xi)\}$. Then, we have

$$\begin{aligned} S(\beta_1) &= \min_{\xi\in\mathcal{X}}\{J_1(\xi)-\beta_1 J_2(\xi)\} = J_1(\xi_{\beta_1})-\beta_1 J_2(\xi_{\beta_1}) \\ &> J_1(\xi_{\beta_1})-\beta_2 J_2(\xi_{\beta_1}) \geq \min_{\xi\in\mathcal{X}}\{J_1(\xi)-\beta_2 J_2(\xi)\} = S(\beta_2). \end{aligned}$$

Therefore, the function $S(\beta)$ is strictly decreasing over $\mathbb{R}$.

Since $J_1(\xi)$ and $J_2(\xi)$ are positive, continuous and $\mathcal{X}$ is compact, $J_1(\xi)$ and $J_2(\xi)$ are bounded. That means, there exist positive real scalars ϕ_1, ψ_1, ϕ_2 and ψ_2 such that $0<\phi_1\leq J_1(\xi)\leq\psi_1$ and $0<\phi_2\leq J_2(\xi)\leq\psi_2$. Thus, for any $\beta\in\mathbb{R}$, we have

$$S(\beta) = \min_{\xi\in\mathcal{X}}\{J_1(\xi)-\beta J_2(\xi)\} = J_1(\xi_\beta)-\beta J_2(\xi_\beta)\leq\psi_1-\beta\phi_2 \tag{33a}$$

$$S(\beta) = \min_{\xi\in\mathcal{X}}\{J_1(\xi)-\beta J_2(\xi)\} = J_1(\xi_\beta)-\beta J_2(\xi_\beta)\geq\phi_1-\beta\psi_2 \tag{33b}$$

where $\xi_\beta\in\mathbf{argmin}_{\xi\in\mathcal{X}}\{J_1(\xi)-\beta J_2(\xi)\}$. Since $0<\phi_2\leq\psi_2$, combing (33a) and (33b), we can obtain $\lim_{\beta\to+\infty}S(\beta)=-\infty$ and $\lim_{\beta\to-\infty}S(\beta)=+\infty$. Since $S(\beta)$ is strictly decreasing, based on the Zero Point Theorem [32] [‡], we can conclude that the characteristic equation $S(\beta)=0$ has a unique solution. □

In the following, we concentrate on solving the characteristic equation $S(\beta)=0$. According to the definition $S(\beta)=\mathbf{min}_{\xi\in\mathcal{X}}\{\bar{J}(\xi)\}$ with $\bar{J}(\xi)=J_1(\xi)-\beta J_2(\xi)=\xi^T(\mathcal{Z}_1\otimes W-\beta\mathcal{Z}_4\otimes W)\xi+(\mathbf{vec}(\mathcal{Z}_2)^T-\beta\mathbf{vec}(\mathcal{Z}_5)^T)\xi+\mathbf{tr}(\mathcal{Z}_3)-\beta\mathbf{tr}(\mathcal{Z}_6)$ is a quadratic function of ξ. Therefore, if $\bar{J}(\xi)$ is a non-convex quadratic function regarding ξ, one always has $\mathbf{min}_{\xi\in\mathcal{X}}\{\bar{J}(\xi)\}<0$ for the enough large set $\mathcal{X}$. That means, it is a necessary condition for the characteristic equation $S(\beta)=0$ that $\bar{J}(\xi)$ is a convex quadratic function regarding ξ. We can obtain the upper bound $\tilde{\beta}$ of β^* by solving the following LMI-based optimization:

$$\tilde{\beta} = \mathbf{max}\quad \beta \qquad \mathbf{s.t.}\quad \mathcal{Z}_1\otimes W-\beta\mathcal{Z}_4\otimes W\succeq 0. \tag{34}$$

Meanwhile, considering $S(0)=\mathbf{min}_{\xi\in\mathcal{X}}\{J_1(\xi)\}>0$ and combining the properties of $S(\beta)$ in Theorem 4.2, we can roughly draw the profile of the function $S(\beta)$ over $\mathbb{R}$ as shown in Figure 1.

According to Figure 1, we need to search a root $\beta^*\in[0,\tilde{\beta}]$ such that the characteristic equation $S(\beta^*)=0$. Since the definition of $S(\beta)$ involves the minimum operator, it is generally non-differentiable. We could not use the classical Newton method to search the root of the characteristic equation $S(\beta)=0$. In order to avoid the use of differential operation, we use the secant method to solve the characteristic equation $S(\beta)=0$. Although the order of convergence for the secant method is generally lower than that of Newton's method, the derivatives $S'(\beta)$ need not be evaluated, which is a definite computational advantage. For more details on the exact order of convergence for the secant method, the reviewer can refer to [33]. Since the convergence of secant method can be guaranteed, the terminal condition for searching the root of the equation $S(\beta)=0$ always holds. The whole algorithm summarized in Algorithm 1.

[‡] If f is a function which is continuous at every point of the interval $[a,b]$ and $f(a)<0$, $f(b)>0$ then $f(x)=0$ at some point $x\in(a,b)$.

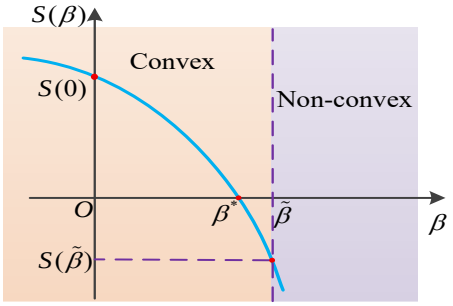

Figure 1: The profile of function $S(\beta)$ over $\mathbb{R}$.

Algorithm 1: Solve $S(\beta) = 0$ using secant method

1: **Initialization**:
(1) Give the approximation precision ϵ and set $i = 0$;
(2) Solve the optimization problem (34) to obtain the value of $\tilde{\beta}$;
(3) Choose $\beta_{c1} = \tilde{\beta}$ and $\beta_{c2} = 0$;
(4) Solve the optimization problem (31) to obtain optimal values $S(\beta_{c1})$ and $S(\beta_{c2})$;

2: **Update β_{next} using Secant Method**:
(3) Compute β_{next} using secant method:

$$\beta_{next} = \beta_{c1} - S(\beta_{c1}) \frac{\beta_{c1} - \beta_{c2}}{S(\beta_{c1}) - S(\beta_{c2})}.$$

(4) Solve the optimization problem (31) to obtain global optimum $\xi_{\beta_{next}}$ and its optimal values $S(\beta_{next})$;
(5) **if** $S(\beta_{next}) < 0$ **then** $\beta_{c1} \leftarrow \beta_{next}$ **else** $\beta_{c2} \leftarrow \beta_{next}$;

3: **Check Termination Condition**:
if $|S(\beta_{next})| \leq \epsilon$ **then**
$\beta^* = \beta_{next}$ and $\xi^* = \xi_{\beta_{next}}$. Terminate algorithm.
else Return Step 2.

After getting the value of ξ^* according to Algorithm 1, we can obtain $L_k^* = \mathbf{mat}(\xi^*)$ by using the matrixing operation. Obviously, by using this gain L_k^* for the FD purpose, we can also implement robust SE (not optimal in general). What is interesting is that even by considering the FD purpose with the pQP method, there are no significant differences from the perspective of SE compared with ZKF, while the performance of FD for the pQP method has an obvious improvement. Regarding this point, it will be illustrated and verified in the following simulation.

5. ILLUSTRATIVE EXAMPLE

In this section, we consider an electric circuit taken from [34] as a case study to illustrate the effectiveness of proposed optimal SE and FD method based on ZKF and pQP. The whole electric circuit chart is shown in Figure 2, where $R_i (i = 1, ..., 8)$ and L_1 and L_2 represent the resistors and inductors, respectively. Moreover, R_1 and R_6 are time-varying resistors and we consider them as the scheduling variables, i.e., $\theta(1) = R_1$ and $\theta(2) = R_6$. $\theta(1)$ ranges from 9Ω to 11Ω and $\theta(2)$ ranges from 25Ω to 27Ω. $e_1(t)$ and $e_2(t)$ are the DC voltage sources used as inputs and $e_1(t) = 1\text{v}$, $e_2(t) = -1\text{v}$.

The system states are the loop currents $i_1(t)$ and $i_2(t)$, and the measured outputs are the voltages of resistors R_1 and R_6, i.e., $y(t) = \begin{bmatrix} U_{R_1}(t) & U_{R_6}(t) \end{bmatrix}^T$.

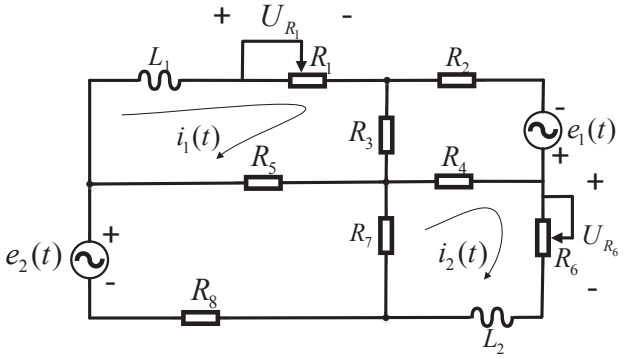

Figure 2: A circuit diagram.

According to the basic circuit theory and Kirchoff's laws, we can derive the following differential equations:

$$\begin{aligned}
L_1\frac{di_1(t)}{dt} &= (-\theta(1) - R_3 - R_5 + \frac{R_3^2}{R_{234}} + \frac{R_5^2}{R_{578}})i_1(t) \\
&\quad + (\frac{R_3R_4}{R_{234}} + \frac{R_5R_7}{R_{578}})i_2(t) + \frac{R_3}{R_{234}}e_1(t) + \frac{R_5}{R_{578}}e_2(t) \\
L_2\frac{i_2(t)}{dt} &= (-R_4 - R_7 - \theta(2) + \frac{R_4^2}{R_{234}} + \frac{R_7^2}{R_{578}})i_2(t) \\
&\quad + (\frac{R_3R_4}{R_{234}} + \frac{R_5R_7}{R_{578}})i_1(t) + \frac{R_4}{R_{234}}e_1(t) + \frac{R_7}{R_{578}}e_2(t), \\
y(t) &= \begin{bmatrix} U_{R_1} \\ U_{R_6} \end{bmatrix} = \begin{bmatrix} \theta(1) & 0 \\ 0 & \theta(2) \end{bmatrix} \begin{bmatrix} i_1(t) \\ i_2(t) \end{bmatrix},
\end{aligned}$$

where $R_{234} = R_2 + R_3 + R_4$, $R_{578} = R_5 + R_7 + R_8$, $R_2 = 17\Omega$, $R_3 = 3\Omega$, $R_4 = 5\Omega$, $R_5 = 2\Omega$, $R_7 = 8\Omega$, $R_8 = 10\Omega$, $L_1 = 0.3H$ and $L_2 = 0.65H$.

With a sampling time of $T_s = 0.01s$, we can transform the circuit model into the following discrete-time LPV model by using the first order forward Euler difference method:

$$\begin{aligned}
x_{k+1} &= A(\theta_k)x_k + B(\theta_k)u_k, \\
y_k &= C(\theta_k)x_k + D(\theta_k)u_k
\end{aligned}$$

with

$$\begin{aligned}
A(\theta_k) &= \begin{bmatrix} -0.0333\theta_k(1) + 0.8520 & 0.0467 \\ 0.0323 & -0.0154\theta_k(2) + 0.8646 \end{bmatrix}, \\
B(\theta_k) &= \begin{bmatrix} 0.0040 & 0.0033 \\ 0.0031 & 00062 \end{bmatrix}, C(\theta_k) = \begin{bmatrix} \theta_k(1) & 0 \\ 0 & \theta_k(2) \end{bmatrix}, D(\theta_k) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

Here we consider the measurement errors of scheduling vector θ as

$$\tilde{\theta}_k \in \tilde{\mathbf{\Theta}} = \begin{bmatrix} (-0.02, & 0.02) \\ (-0.02, & 0.02) \end{bmatrix}.$$

Furthermore, in order to verify the effectiveness of the proposed method, it is assumed that the unknown input ω_k and measurement noise v_k are bounded by

$$\mathbf{W} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} 0.03 & 0 \\ 0 & 0.03 \end{bmatrix} \mathbb{B}^2, \mathbf{V} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} 0.03 & 0 \\ 0 & 0.03 \end{bmatrix} \mathbb{B}^2,$$

whose distribution matrices are given as

$$E(\theta_k) = \begin{bmatrix} 0.4693 & 0.1496 \\ 0.1346 & 0.4748 \end{bmatrix}, P(\theta_k) = \begin{bmatrix} 0.8147 & 0.9134 \\ 0.9058 & 0.6324 \end{bmatrix},$$

respectively. Meanwhile, we assume that the additive actuator faults f_k and sensor faults s_k are respectively bounded by

$$\mathbf{F} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} \mathbb{B}^2, \mathbf{S} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} \mathbb{B}^2$$

and their distribution matrices are given by

$$G(\theta_k) = \begin{bmatrix} 0.4382 & 0.6513 \\ 0.6332 & 0.4894 \end{bmatrix}, H(\theta_k) = \begin{bmatrix} 0.8147 & 0.1270 \\ 0.9058 & 0.9134 \end{bmatrix}.$$

Furthermore, we set the following fault scenarios:

$$f_k = \begin{cases} [0,0]^T, & k \leq 20, \\ [0.35, 0.4]^T, & k \geq 21. \end{cases} \quad s_k = \begin{cases} [0,0]^T, & k \leq 20, \\ [0.45, 0.3]^T, & k \geq 21. \end{cases}$$

That means, the healthy scenario is considered from the beginning of the simulation to the time instant $k = 20$. Then, the actuator faults f_k and sensor faults s_k are simultaneously activated at $k = 21$ and they last until the end of simulation. Figure 3 shows the SE results when the actuator faults f_k and sensor faults s_k occur at time instant $k = 21$. Each subplot corresponds to one component of the system state x_k and the interval hull $\mathbf{X}_k$. For the convenience of illustration, we take the first subplot to present the SE results. The red dash line represents the real system state $x_k(1)$. The region between the blue lines denotes the interval hull $\mathbf{X}_k(1)$ by considering the ZKF method to compute the observer gain. The region between the green lines denotes the SE interval hull $\mathbf{X}_k(1)$ using the pQP method to compute the observer gain. It can be found that from time instant $k = 0$ to $k = 21$ when the system is only affected by the disturbances and noises, there are no big differences from the aspect of SE for both ZKF and pQP methods. Certainly, it is normal to obtain a little better SE result for the ZKF method since the ZKF optimization objective only considers the observation performance while the pQP optimization criterion further considers the FD purposes. After the time instant $k = 21$, the actuator faults f_k and the sensor faults s_k are simultaneously activated, whose effect on SE can be seen from Figure 3. We can find that the pQP has a larger inconsistency of the observation after the occurrence of faults than the ZKF, which implies that an increased sensitivity of FD can be obtained considering the pQP method to compute the observer gain.

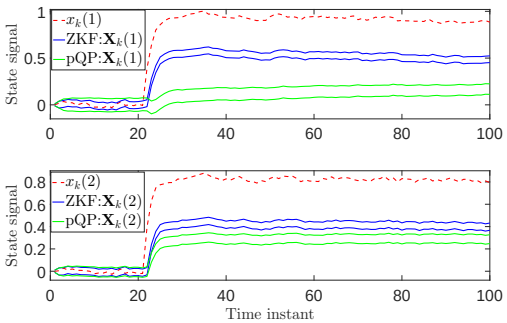

Figure 3: State estimation in case of actuator faults f_k and sensor faults s_k.

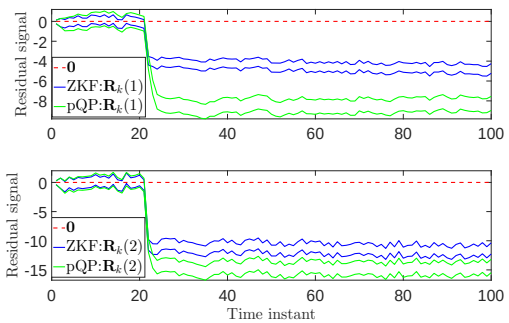

Figure 4: Fault detection in case of actuator faults f_k and sensor faults s_k.

For further illustration on the results of FD specifically, Figure 4 show the FD results for both ZKF and pQP methods. As it is shown in Figure 4, form time instant $k = 0$ to $k = 20$, the origin

$\mathbf{0}$ is contained in the residual set $\mathbf{R}_k$ generated from both ZKF and pQP methods. In this case, we consider that the system still operates in the healthy situation. After the occurrence of faults at time instant $k = 21$, the origin $\mathbf{0}$ is no longer contained in the residual set $\mathbf{R}_k$ for the ZKF and the pQP, which means that both two methods has detected the occurrence of faults. Further, the higher sensitivity of FD for the pQP approach in comparison with the ZKF method can be found since its generated residual set $\mathbf{R}_k$ moves further from the non-faulty origin $\mathbf{0}$.

Furthermore, for the completion of quantitative analysis, we use the ratio between the weighted Frobenius norm of the center $c_{\mathbf{R}_k}$ and the generator matrix $M_{\mathbf{R}_k}$ of the residual set $\mathbf{R}_k$ to quantitatively characterize the sensitivity of FD for both the ZKF and pQP methods. Figure 5 shows the computed results regarding $\frac{\|c_{\mathbf{R}_k}\|^2_{F,W}}{\|M_{\mathbf{R}_k}\|^2_{F,W}}$, where we can see that the ratio $\frac{\|c_{\mathbf{R}_k}\|^2_{F,W}}{\|M_{\mathbf{R}_k}\|^2_{F,W}}$ of the pQP method is much larger than that of the ZKF approach after the occurrence of faults. That means, there is an obvious enhancement regarding the sensitivity of FD for the pQP approach in comparison with the ZKF approach.

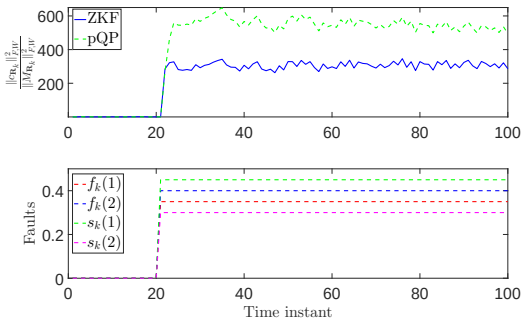

Figure 5: Ratio $\frac{\|c_{\mathbf{R}_k}\|^2_{F,W}}{\|M_{\mathbf{R}_k}\|^2_{F,W}}$ and fault profiles in case of actuator faults f_k and sensor faults s_k.

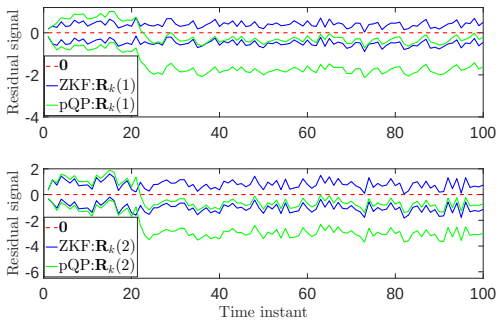

Figure 6: FD results in case of actuator faults f_k and sensor faults s_k with smaller magnitudes.

We consider another group of actuator and sensor faults with smaller magnitudes. The fault scenarios are set as follows:

$$f_k = \begin{cases} [0,0]^T, & k \le 20, \\ [0.05, 0.04]^T, & k \ge 21. \end{cases} \quad s_k = \begin{cases} [0,0]^T, & k \le 20, \\ [0.03, 0.06]^T, & k \ge 21. \end{cases}$$

The FD results for both ZKF and pQP methods are shown in Figure 6. It can be found that from time instant $k = 0$ to $k = 20$, the residual sets $\mathbf{R}_k$ for both ZKF and pQP contain the origin $\mathbf{0}$, which implies that there is no fault occurrence and the system operates in healthy situation. From time instant $k = 21$, the residual set $\mathbf{R}_k$ generated by the pQP method no longer contains the origin $\mathbf{0}$, which means that the pQP method has detected the occurrence of faults. However, since the residual set $\mathbf{R}_k$ generated by the ZKF method contains the origin $\mathbf{0}$ during the whole stage of system operation, the occurred faults could not be detected using the ZKF approach. R2-5

Furthermore, we make a comparison with the proposed generalized eigenvalues/eigenvectors (GVV) based FD method in [20] on the performance of FD results. The fault scenarios are given by

$$f_k = \begin{cases} [0,0]^T, & k \le 20, \\ [0.8, 1.3]^T, & k \ge 21. \end{cases} \quad s_k = \begin{cases} [0,0]^T, & k \le 20, \\ [-0.8, -2.3]^T, & k \ge 21. \end{cases}$$

The GVV method is used to compute the optimal ξ^* for the objective function $J(\xi)$ in (26) to construct the optimal gain matrix $L^*_k = \mathbf{mat}(\xi^*)$ at each step. Compared with the proposed pQP method, there is a relative advantage for the GVV method on the computation time. By using a laptop with 4G RAM, Core m3-6Y30 @ 0.90GHz CPU and 128GB SSD, the average time to compute the optimal gain matrix L^*_k at each step for the GVV method during whole simulation stage is 0.1582s. While for the proposed pQP method, the average time at each step is 1.5743s. Although the computation speed of the pQP method is a little slower than that of the GVV method, R2-6

the proposed pQP method has a better FD result. The FD results for both the GVV method and the pQP method are shown in Figure 7. It can be found that from the time instant $k = 0$ to $k = 20$, the residual sets $\mathbf{R}_k$ for both the GVV method and the pQP method contain the origin $\mathbf{0}$. This implies that the system is healthy. However, from the time instant $k \geq 21$, the residual sets $\mathbf{R}_k$ for both methods no longer contain the origin $\mathbf{0}$. Thus, we consider that both methods have detected the faults f_k and s_k, and the system becomes faulty. Note that, although both methods successfully implement FD, the proposed pPQ-based method has a higher FD sensitivity than the GVV method. We show the ratio $\frac{\|c_{\mathbf{R}_k}\|^2_{F,W}}{\|M_{\mathbf{R}_k}\|^2_{F,W}}$ for both methods in Figure 8. It can be found that after the fault occurrence, the ratio $\frac{\|c_{\mathbf{R}_k}\|^2_{F,W}}{\|M_{\mathbf{R}_k}\|^2_{F,W}}$ of the pQP method is larger than that of the GVV method, which implies the proposed method has a higher FD sensitivity. R2-5

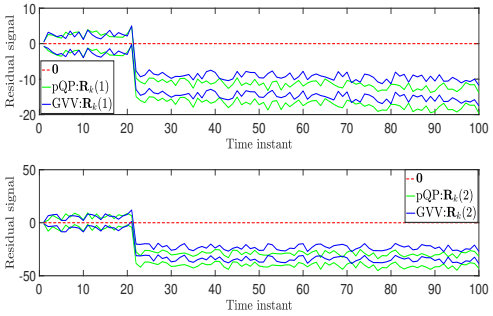

Figure 7: FD results for both the GVV method and the pQP method.

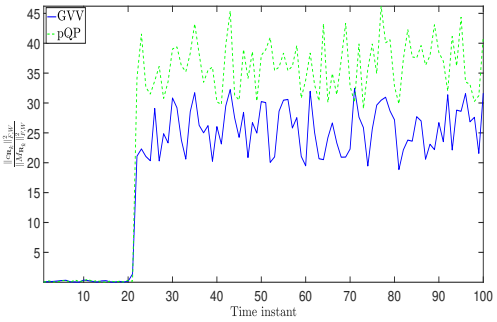

Figure 8: Ratio $\frac{\|c_{\mathbf{R}_k}\|^2_{F,W}}{\|M_{\mathbf{R}_k}\|^2_{F,W}}$ for both the GVV method and the pQP method.

6. CONCLUSIONS

This paper proposes an optimal SE and FD method for discrete-time LPV systems with measurement error-affected scheduling variables by combining ZKF and pQP under the condition that system uncertainties are bounded. Only considering the observation purpose, the optimal SE results can be obtained based on the ZKF procedure by using the weighted Frobenius norm to characterize the size of the estimated state set. Furthermore, by maximizing the effect of faults on the error of SE with respect to the system uncertainties, the optimal FD criterion can be formulated into an on-line fractional programming problem equivalent to the pQP problem which can be efficiently solved by searching the root of the characteristic equation using the secant method. The pQP approach has an obvious enhancement of sensitivity of FD in comparison with the ZKF approach. In the future, we will further consider improving the performance of fault isolation and estimation based on the idea in this paper.

ACKNOWLEDGEMENTS

This work was partially supported by the National Natural Science Foundation of China (No.U1813216 and No.61803221), the Science and Technology Planning Project of Guangdong Province (No. 2017B010116001), and the Basic Research Program of Shenzhen JCYJ20170817152701660 and JCYJ20170412171459177).

REFERENCES

1. M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki. *Diagnosis and Fault-Tolerant Control*. Springer-Verlag, Berlin, Germany, 2006.
2. F. Blanchini and S. Miani. *Set-theoretic Methods in Control*. Birkhäuser Boston, 2008.
3. Y.M. Zhang and J. Jiang. Bibliographical review on reconfigurable fault-tolerant control systems. *Annual Reviews in Control*, 32(2):229 – 252, 2008.
4. Y. Wang, V. Puig, and G. Cembrano. Non-linear economic model predictive control of water distribution networks. *Journal of Process Control*, 56:23–34, 2017.
5. J. Mohammadpou and C.W. Scherer. *Control of Linear Parameter Varying Systems with Applications*. Springer US, 2012.
6. J.S. Shamma and M. Athans. Gain scheduling: Potential hazards and possible remdies. *IEEE Control Systems*, 12(3):101–107, 1992.
7. J. Boker, Z. Szabo, and G. Stikkel. Failure detection for quasi LPV systems. In *Proceedings of the 41th IEEE Conference on Decision and Control*, Las Vegas, USA, December, 2002.
8. M. Sato. Gain-scheduled flight controller using bounded inexact scheduling parameters. *IEEE Transactions on Control Systems Technology*, 26(3):1074–1082, 2018.
9. S. Gómez-Peñate, G. Valencia-Palomo, F.R. López-Estrada, C.M. Atorga-Zaragoza, R.A. Osornio-Rios, and I. Santos-Ruiz. Sensor fault diagnosis based on a H1 sliding mode and unknown input observer for Takagi-Sugeno systems with uncertain premise variables. *Asian Journal of Control*, 21(1):339–353, 2019.
10. H. Hamdi, M. Rodrigues, C. Mechmeche, D. Theilliol, and N.B. Braiek. State estimation for polytopic LPV descriptor systems: application to fault diagnosis. In *the 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, Barcelona, Spain, July 2009.
11. G.I. Bara, J. Daafouz, J. Ragot, and F. Kratz. State estimation for affine LPV systems. In *Proceedings of the 39th IEEE Conference on Decision and Control*, Sydney, Australia, December 2000.
12. F. Blanchini and S. Miani. Stabilization of lpv systems: state feedback, state estimation and duality. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, Hawaii USA, December 2003.
13. A.H. Hassanabadi, M. Shafiee, and V. Puig. Robust fault detection of singular LPV systems with multiple time–varying delays. *International Journal of Applied Mathematics and Computer Science*, 26(1):45–61.
14. F. Xu, J.B. Tan, Y. Wang, X.Q. Wang, B. Liang, and B. Yuan. Robust fault detection and set-theoretic UIO for discrete-time LPV systems with state and output equations scheduled by inexact scheduling variables. *IEEE Transactions on Automatic Control*, 2019.
15. R.E. Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.
16. P.S. Maybeck. *Stochastic models, estimation, and control*. Academic press, 1982.
17. T. Alamo, J.M. Bravo, and E.F. Camacho. Guaranteed state estimation by zonotopes. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, Hawaii, USA, December 2003.
18. X.J. Li and G.H. Yang. Fault detection in finite frequency domain for takagi-sugeno fuzzy systems with sensor faults. *IEEE Transactions on Cybernetics*, 44(8):1446–1458, 2013.
19. G. Tao. Direct adaptive actuator failure compensation control: a tutorial. *Journal of Control and Decision*, 1(1):75–101, 2014.
20. M. Pourasghar, C. Combastel, V. Puig, and C. Ocampo-Martinez. FD-ZKF: A zonotopic kalman filter optimizing fault detection rather than state estimation. *Journal of Process Control*, (73):89–102, 2019.
21. F. Nejjari, V. Puig, S. Montes de Oca, and A. Sadeghzadeh. Robust fault detection for LPV systems using interval observers and zonotopes. In *Proceedings of 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, Beijing, China, December 16-18 2009.
22. C. Combastel. Zonotopes and kalman observers: Gain optimality under distinct uncertainty paradigms and robust convergence. *Automatica*, 55:265–273, 2015.
23. Y. Wang, V. Puig, F. Xu, and G. Cembrano. Robust fault detection and isolation based on zonotopic unknown input observers for discrete-time descriptor systems. *Journal of the Franklin Institute*, 356(10):5293 – 5314, 2019.
24. J. Chen and R.J. Patton. *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers, 1999.
25. J. Löfberg. Automatic robust convex programming. *Optimization methods and software*, 27(1):115–129, 2012.
26. M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008.
27. R. Nazari, M.M. Seron, and J.A. De Donà. On virtual actuators for lpv systems under errors in the measurement of the varying parameter. In *Proceedings of the 5th Australian Control Conference (AUCC)*, Gold Coast, Australia, November, 2015.
28. R.E. Moore, R.B. Kearfott, and M.J. Cloud. *Introduction to Interval Analysis*. Siam, 2009.
29. J. Daafouz and J. Bernussou. Parameter dependent lyapunov functions for discrete time systems with time varying parametric uncertainties. *Systems & Control Letters*, 43(5):355–399, 2001.
30. D. Rotondo, F. Nejjari, and V. Puig. A virtual actuator and sensor approach for fault tolerant control of LPV systems. *Journal of Process Control*, 24(3):203 – 222, 2014.
31. T. Alamo, J.M. Bravo, and E.F. Camacho. Guaranteed state estimation by zonotopes. *Automatica*, 41(6):1035–1043, 2005.
32. S.B Russ. A translation of bolzano's paper on the intermediate value theorem. *Historia Mathematica*, 7(2):156 – 185, 1980.
33. M. Raydan. Exact order of convergence of the secant method. *Journal of Optimization Theory and Applications*, 78(3):541–551, 1993.
34. M.Rodrigues, H.Hamdi, D.Theilliol, C. Mechmeche, and N. Benhadj Braiek. Fault diagnosis based on adaptive polytopic observer for LPV descriptor systems. In *8th IFAC Symposium on Fault Detection, SAFEPROCESS*, Mexico City, Mexico, August 29-31 2012.

35. X.D. Zhang. *Matrix Analysis and Application*. Tsinghua University Press, 2016.
36. W.Kuhn. Rigorously computed orbits of dynamical systems without the wrapping effect. *Computing*, 61(1), 1998.

Appendices

A. MATRIX CALCULUS

For a matrix $X \in \mathbb{R}^{n\times n}$, $\mathbf{tr}(X)$ denotes the trace of X, respectively. The compatible matrix full of 1 are denoted as $\mathbf{I}$, and I_n denotes the n-dimensional identity matrix. $\mathbf{diag}(x)$ denotes a diagonal matrix whose diagonal elements are composed of a vector $x \in \mathbb{R}^n$. Let $X = [x_{ij}]_{m\times n} \in \mathbb{R}^{m\times n}$ and $\mathbf{vec}(X)$ denotes the vectorization of matrix X, i.e., $\mathbf{vec}(X) = x = [x_{11}\cdots x_{m1}, \cdots, x_{1i}\cdots x_{mi}, \cdots, x_{1n}\cdots x_{mn}]^T \in \mathbb{R}^{mn}$. Inversely, the matrixing operation of a vector $x \in \mathbb{R}^{mn}$ is $\mathbf{mat}(x) = X = [x_{ij}]_{m\times n} \in \mathbb{R}^{m\times n}$. For matrices $X \in \mathbb{R}^{n\times n}$ and $Y \in \mathbb{R}^{m\times m}$, the Kronecker product of these two matrices is denoted by

$$X \otimes Y = \begin{bmatrix} x_{11}Y & \cdots & x_{1n}Y \\ \vdots & \ddots & \vdots \\ x_{n1}Y & \cdots & x_{nn}Y \end{bmatrix}_{mn\times mn}.$$

$X \succ 0$ denotes positive definiteness if the scalar $x^T Xx$ is positive for arbitrary non-zero column vector x of real numbers. Let X, A, B, C and D be matrices with appropriate dimensions. According to [35], we can obtain

$$\begin{aligned}
&\mathbf{tr}(A) = \mathbf{tr}(A^T), \mathbf{tr}(AB) = \mathbf{tr}(BA), (A\otimes B)^T = A^T \otimes B^T,\\
&\frac{\partial}{\partial X}\mathbf{tr}(AX^T B) = A^T B^T, \mathbf{tr}(A^T B) = (\mathbf{vec}(A))^T \mathbf{vec}(B),\\
&\frac{\partial}{\partial X}\mathbf{tr}(AXBX^T C) = BX^T CA + B^T X^T A^T C^T,\\
&\mathbf{tr}(ABCD) = (\mathbf{vec}(D^T))^T (C^T \otimes A)\mathbf{vec}(B).
\end{aligned}$$

For $X \in \mathbb{R}^{n\times m}$, $W \in \mathbb{R}^{n\times n}$ and $W = W^T \succ 0$, the weighted Frobenius norm of X is defined by $\|X\|_{F,W} = \sqrt{\mathbf{tr}(X^T WX)}$, and $\|X\|_F = \sqrt{\mathbf{tr}(X^T X)}$, obtained by setting $W = I_n$, is the non-weighted Frobenius norm. Moreover, $\|X\|^2_{F,W} = \sum_{i=1}^m \|x_i\|^2_W$, where $\|x_i\|_W = \sqrt{x_i^T W x_i}$ is a weighted vector norm in $\mathbb{R}^n$ and x_i is the i-th column of the matrix X.

B. INTERVALS AND ZONOTOPES

Based on [28], the closed interval denoted by $[a, b]$ is the set of real numbers given by $x = [a, b] = \{s \in \mathbb{R} | a \le s \le b\}$. Although various other types of intervals (open, half-open) appear throughout mathematics, our work in this paper only centers primarily on closed intervals. The radius of an interval x is defined and denoted by $\mathbf{rad}(x) = \frac{b-a}{2}$. The midpoint of an interval x is given by $\mathbf{mid}(x) = \frac{a+b}{2}$.

An interval matrix is defined as a matrix whose elements are interval variables, i.e.,

$$X = \begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m,n} \end{bmatrix},$$

where $x_{i,j}$ $(\forall i = 1, ..., m,\ j = 1, ..., n)$ is an interval variable. If the number of row or column of the interval matrix X is 1, X degenerates to an interval vector. The radius of an interval matrix X is

defined as

$$\mathbf{rad}(X) = \begin{bmatrix} \mathbf{rad}(x_{1,1}) & \cdots & \mathbf{rad}(x_{1,n}) \\ \vdots & \ddots & \vdots \\ \mathbf{rad}(x_{m,1}) & \cdots & \mathbf{rad}(x_{m,n}) \end{bmatrix},$$

and the midpoint of interval matrix X is given by

$$\mathbf{mid}(X) = \begin{bmatrix} \mathbf{mid}(x_{1,1}) & \cdots & \mathbf{mid}(x_{1,n}) \\ \vdots & \ddots & \vdots \\ \mathbf{mid}(x_{m,1}) & \cdots & \mathbf{mid}(x_{m,n}) \end{bmatrix}.$$

According to [31] and [36], the definition and properties of zonotopes are given as follows. Readers can refer [31] and [36] for more details on zonotopes.

Definition B.1
A zonotope Z is defined as $Z = g \oplus H\mathbb{B}^t$, where g and H are its center and generator matrix, respectively, $\mathbb{B}^t$ is an interval vector composed of t unitary intervals $[-1, 1]$. Here we simplify $Z = g \oplus H\mathbb{B}^t$ into $Z = \langle g, H \rangle$.

Property B.1
Given two zonotopes $Z_1 = \langle g_1, H_1 \rangle$ and $Z_2 = \langle g_2, H_2 \rangle$, $Z_1 \oplus Z_2 = \langle g_1 + g_2, [H_1 \ H_2] \rangle$.

Property B.2
Given a zonotope $Z = \langle g, H \rangle$ and a compatible matrix K, $KZ = \langle Kg, KH \rangle$.

Given a zonotope $Z = \langle g, H \rangle$, the interval hull of Z is defined as $\square(Z) = \langle g, \mathbf{diag}(|H|\mathbf{1}) \rangle$. For a zonotope $Z = \langle g, H \rangle$ with $H \in \mathbb{R}^{n \times b}$, according to [22], the weighted zonotope reduction operator $\downarrow_{\lambda,W}(H)$ denotes an order reduction of $Z = \langle g, H \rangle$ from the order b to λ, which first sorts the columns of H on decreasing weighted vector norm and then encloses the zonotope $\langle 0, H_< \rangle$ generated by the $b - \lambda + 1$ smaller columns into a box, i.e., R2-3

$$H = \begin{bmatrix} h_1 & \cdots & h_j & \cdots & h_b \end{bmatrix}, \|h_j\|_W^2 \geq \|h_{j+1}\|_W^2$$

and if $b \leq \lambda$, then $\downarrow_{\lambda,W}(H) = H$. Otherwise, we have $\downarrow_{\lambda,W}(H) = \begin{bmatrix} H_> & \mathbf{diag}(|H_<|\mathbf{1}) \end{bmatrix}$ with $H_> = \begin{bmatrix} h_1 & \cdots & h_{\lambda-1} \end{bmatrix}, H_< = \begin{bmatrix} h_\lambda & \cdots & h_b \end{bmatrix}$, and the weighted zonotope reduction operator $\downarrow_{\lambda,W}(H)$ satisfies the inclusion property $\langle g, H \rangle \subseteq \langle g, \downarrow_{\lambda,W}(H) \rangle$, where $\lambda \geq n$.

Property B.3
Given a family of zonotopes denoted by $Z = \langle g, \mathbf{H} \rangle$, where $\mathbf{H} \in \mathbb{R}^{n \times m}$ is an interval matrix, a zonotope inclusion $\diamond(Z)$ is defined by

$$\diamond(Z) = \langle g, \begin{bmatrix} \mathbf{mid}(\mathbf{H}) & H \end{bmatrix} \rangle,$$

where the matrix H is a diagonal matrix with

$$H_{ii} = \sum_{j=1}^{m} \mathbf{rad}(\mathbf{H}_{ij}), i = 1, 2, \cdots, n,$$

where $\mathbf{mid}(\cdot)$ and $\mathbf{rad}(\cdot)$ compute the midpoint and radius of interval matrices.

Property B.4
Given $Z_{k+1} = \mathbf{A}Z_k \oplus \mathbf{B}u_k$, where $\mathbf{A}$ and $\mathbf{B}$ are interval matrices and u_k is the input at time instant k, if Z_k is a zonotope with the center g_k and segment matrix H_k, Z_{k+1} can be bounded by a zonotope

$$Z_{k+1}^e = \langle g_{k+1}, H_{k+1} \rangle$$

with

$$g_{k+1} = \mathbf{mid}(\mathbf{A})g_k + \mathbf{mid}(\mathbf{B})u_k, H_{k+1} = \begin{bmatrix} J_1 & J_2 & J_3 \end{bmatrix},$$
$$J_1 = \mathbf{seg}(\diamond(\mathbf{A}H_k)), J_2 = \mathbf{rad}(\mathbf{A}g_k), J_3 = \mathbf{rad}(\mathbf{B}u_k),$$

where $\mathbf{seg}(\cdot)$ computes the segment matrix of a zonotope.