

A Markovian jump system approach for the estimation and adaptive diagnosis of decreased power generation in wind farms

 ISSN 1751-8644
 doi: 0000000000
 www.ietdl.org

 Ester Sales-Setién^{1*} Ignacio Peñarrocha-Alós¹
¹ Department of Industrial Engineering and Design, Universitat Jaume I, Castelló, Spain

* E-mail: esales@uji.es

Abstract: In this work, we propose a Markovian jump model of the power generation system of a wind turbine and we present a closed-loop model-based observer to estimate the faults related to energy losses. The observer is designed through an \mathcal{H}_∞ -based optimization problem that optimally fixes the trade-off between the observer fault sensitivity and robustness. The fault estimates are then used in data-based decision mechanisms for achieving fault detection and isolation. The performance of the strategy is then ameliorated in a wind farm level scheme that uses a bank of the aforementioned observers and decision mechanisms. Finally, the proposed approach is tested using a well-known benchmark in the context of wind farm fault diagnosis.

1 Introduction

Wind energy is considered as a powerful source of sustainable energy. However, wind turbines (WTs) are expensive systems and their maintainability and reliability must be high. Fouling of the rotor blades with ice or insects, as well as erosion, is an important root cause of faults and it is expected to increase in significance as more WTs are situated in locations with higher wind speeds (WSs) [1]. As stated in [2], the major problem related to debris build-up and erosion is the reduction of the overall WT aerodynamic efficiency leading to unpredicted reductions in power production. Besides possible human safety risks, debris build-up may also cause mass and aerodynamic imbalances, damaging all WT components.

As discussed in [3], the diagnosis of debris build-up can be done using direct or indirect measurements. Direct methods are based on the detection of some change of physical properties such as mass, electrical or thermal properties. Hence, they often require extra equipment, which augment the installation and maintenance costs and the weight and space requirements. For its part, indirect methods are mainly based on detecting the reduction in power production. These methods do not require extra hardware because they use WT control measurements. They require, however, the WTs to be in operation. Their main disadvantage is that energy losses may be also the consequence of other phenomena which are, nonetheless, generally easily distinguishable from debris build-up. For instance, consider the occurrence of an electric fault in the WT generator. In this case, the deviation of the generator torque from its reference, which can be deduced from the generator torque measurement, provides an immediate and accurate diagnosis of the electric fault leading to its straightforward isolation from debris build-up or erosion [4].

The indirect fault diagnosis (FD) of decreased power generation due to debris build-up or erosion is one of the objectives of the realistic and widely accepted wind farm (WF) benchmark presented in [5]. The benchmark includes the WF power generation model which is based on the well-known power curve of a WT. Basically, two kinds of residuals are used in the bibliography [6–10] to diagnose power reductions in the benchmark. At a WT level, temporal residuals represent the inconsistencies between the power generation model output and the generated power measurement of a single WT. At a WF level, spatial residuals represent the inconsistencies between the generated power measurements of different WTs. Borcehrsen et al. [6] utilize open-loop temporal residuals in dynamical cumulative sums. Alternatively, Duviella et al. [7] present a FD approach based on spatial residuals and Blesa et al. [8] compute both spatial and

open-loop temporal residuals via nonlinear parameter varying parity equations. For its part, Badihi et al. [9] merge both approaches. The authors use the generation model of the WTs to compute the power differences among the WF; then, they compare them to the measured power differences. They also present an approach which replaces the generation model by fuzzy inference mechanisms. Similarly, Simani et al. [10] present fuzzy and neural networks techniques.

The power generation model of a WT is affected by various non-negligible disturbances: mean WS estimation errors, turbulences, vibrations and measurement noises [5]. Hence, the WT level open-loop temporal residuals in [6, 8] may be significantly disturbed. The spatial residuals in [7–9] are not affected by mean WS estimation errors because they compare WTs working under the same wind conditions. For instance, the methods in [9] require that all the WTs in the WF operate under the same wind conditions. However, in most cases, the WTs operate in different conditions and WS estimation errors still affect the spatial residuals. In all, it is of interest to develop closed-loop strategies with a better performance w.r.t. disturbances at a WT level and at a WF level for groups of WTs affected not only by identical but also by similar wind conditions. Specially, if the FD objectives require not only the information about the appearance and the location of a fault (fault detection and isolation or FDI) but also about its size and shape (fault estimation or FE). FE is of paramount importance for both real-time decisions and active fault tolerant control (AFTC) [11] such as power demand redistribution among the WF.

The closed-loop power generation system of a WT is a parametric loop because the WT operates in different WS zones and under different WT operating modes [12]. Moreover, the stochastic characteristic of the wind brings further difficulty to the design of closed-loop FE strategies. In the literature, many authors conclude that the WS behaviour can be explained as a Markovian process [13, 14]. This behaviour has been exploited in recent WT control schemes as the ones in [15–17]; however, up to date, there has been no work taking advantage of the Markovian behaviour of the wind to design FE strategies.

Motivated by the above background, in this work, we develop a Markovian jump system approach for closed-loop FE of decreased power generation in a WF. First, we develop a WT level approach; then, we extend it to the WF level for groups of nearby WTs working under similar wind conditions. FE techniques must be simultaneously robust against uncertainties and noises [18] and sensitive to faults through the accomplishment of certain trade-off between these properties [19]. In this paper, we use the \mathcal{H}_∞ performance of the proposed model-based observer to characterize these properties and

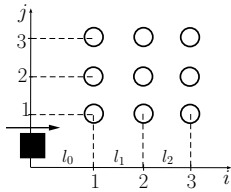


Fig. 1: Layout of the WF benchmark ($l_0 = 1150$ m and $l_1 = l_2 = 1138.44$ m). \rightarrow : Wind direction, \circ : Wind turbine, \blacksquare : Wind mast.

to set up an optimal observer design strategy based on linear matrix inequalities (LMIs). Providing a systematic performance-based optimal approach for tuning the FE observer is an advantage when compared to other algorithm designs, where some user expertise is necessary. For instance, numerical extensive simulations and trial and error procedures are necessary to tune the algorithms in [10].

To achieve FDI, we feed the fault estimates provided by the proposed observer into decision mechanisms based on thresholds. Model-based thresholding usually leads to too conservative results which are characterized by poor fault detectability and isolability [20], notably in the cases of Markovian jump systems [21]. Therefore, we utilize a data-driven FDI approach based on adaptive thresholds for evaluating the FE output provided by the model-based observer. The proposed adaptive mechanism enhances a tight FDI adjustment in the different WS zones and WT operating modes when compared to the constant thresholds in [6, 9, 10].

1.1 Challenges

The proposed Markovian jump system approach for FE and adaptive FDI entails the following challenges.

1. Obtaining a linear parameter varying state-space model of the power generation system of a WT, which is based on nonlinear power curves and consists of two operation modes. To do so, we define a parameter vector containing the WS acting on a WT and the difference between the demanded and the generated power. In this work, the power curves in the WF benchmark [5] are utilized. In reality, these curves can be obtained through different methods such as random forest, method of bins, k-nearest neighbours or support vector regression [22].
2. Modelling the stochastic behaviour of the parameters as Markovian processes. Obtaining a suitable partition of the parameter set taking its discontinuities into account. Computing the transition probabilities between subsets.
3. Building an augmented model-based observer for FE of decreased power generation in a WT. Designing the observer through a multiobjective optimization problem that guarantees certain optimal trade-off between the robustness against disturbances/uncertainties and the sensitivity to faults. To do so, we use the \mathcal{H}_∞ performance of the observer, which we formulate using linear matrix inequalities (LMIs) for Markovian processes and convex polytopic sets.
4. Building an adaptive threshold-based decision mechanism for FDI of decreased power generation in a WT. Proposing an algorithm for computing a tight adaptive threshold piecewise on the WT operation mode. Tuning the algorithm with fault-free datasets using data-driven techniques.
5. Extending the WT level model-based FE and data-driven FDI approach to the WF level. To do so, we group the WTs operating under similar wind conditions and we use a bank of the previous observers and decision mechanisms that automatically merges the information provided by the temporal and spatial relations in the WF according to the degree of shared uncertainty among the WTs.

1.2 Structure and notation

The remainder of this paper is organized as follows. Section 2 gives the problem formulation and Section 3 presents a Markovian jump modeling of the power generation system of a WT. In Section 4, we develop a wind turbine level model-based FE strategy and, in Section 5, the FE output is utilized in a data-based FDI algorithm.

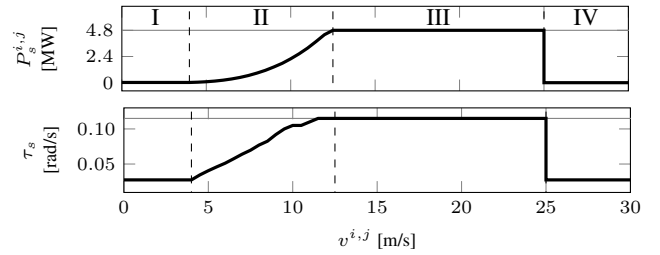


Fig. 2: Power curve $P_s^{i,j}(v^{i,j})$ and transfer coefficient curve $\tau_s(v^{i,j})$. Values borrowed from [5].

Then, in Section 6, these strategies are extended to the wind farm level. Simulation results are reported in Section 8 followed by some concluding remarks in Section 9.

Let a be some vector and A and B be some matrix. The size of a is denoted as n_a and $a[i]$ denotes the i -th element in a . $A \preceq 0$ means that A is negative semidefinite and similar applies to \succeq . The direct sum is represented as \oplus and the Kronecker product is represented as \otimes . I_n is the identity matrix of size $n \times n$, $\mathbf{1}_{n \times m}$ is a matrix of ones of size $n \times m$ and $0_{n \times m}$ is the zero matrix of size $n \times m$ (all of these matrices being of appropriate dimensions when the subindex is omitted). Expected value and absolute value are denoted by $\mathbb{E}\{\cdot\}$ and $|\cdot|$. Let x_k be a vector of stochastic signals at a sample k . We write $\|x_k\|_\infty \triangleq \max_i |x_k[i]|$ for the max norm of vector x_k . We write $\|x\|_\infty \triangleq \max_k \max_i |x_k[i]|$ for the l_∞ norm of signal x , $\|x\|_2^2 \triangleq \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K x_k^T x_k$ for its l_2 norm and $\|x\|_{\text{RMS}}^2 = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K x_k^T x_k$ for its RMS norm.

2 Problem Statement

2.1 Wind Farm Benchmark Description

Consider the WF benchmark [5] with 9 WTs of 4.8 MW and the layout shown in Fig. 1. We name the WTs according to the existing wind direction (see the example in Fig. 1) and we consider that the wind is perpendicular to the rows of WTs (which are numbered as $i = 1, \dots, Y$ and Y is the number of rows for the considered wind direction) and parallel to the columns of WTs (which are numbered as $j = 1, \dots, Z$ and Z is the number of columns for the considered wind direction). We denote the distance between two rows i and $i + 1$ as l_i . There is a wind mast that measures the WS at row $i = 0$ (i.e., $v^0 = \hat{v}^0 + \tilde{v}^0$, with v^0 being the WS at the wind mast, \hat{v}^0 being its measurement and \tilde{v}^0 being the corresponding sensor noise). The WS acting on the WT (i, j) , denoted as $v^{i,j}$, can be modelled as $v^{i,j} = v^i + \tilde{v}_t^{i,j}$, where v^i is the mean WS acting on the WTs in the i -th row and $\tilde{v}_t^{i,j}$ is a zero-mean turbulence component of known variance ($\sigma_t^2 = 0.2 \text{ m}^2/\text{s}^2$).

The static available power in the WT (i, j) , denoted as $P_s^{i,j}$, represents the theoretical maximal generated power in the WT and it depends on $v^{i,j}$ as shown in Fig. 2. This power curve is extensively used in monitoring of wind farms (e.g., [23]) and it consists of four zones delimited by three different WSs: the cut-in ($v_{\text{cut-in}} = 4$ m/s), the rated ($v_{\text{rated}} = 12.5$ m/s) and the cut-out ($v_{\text{cut-out}} = 25$ m/s) WS [12]:

- **Zone I** with $v^{i,j} < v_{\text{cut-in}}$. In this zone, the aerodynamic torque is not sufficient to overcome the WT inertia and $P_s^{i,j} = 0$.
- **Zone II** or partial load region with $v_{\text{cut-in}} \leq v^{i,j} < v_{\text{rated}}$. Maximum power point tracking (MPPT) techniques are performed in this region. For MPPT, the pitch angle is held at zero degrees and the generator moment is adjusted to keep the power coefficient at a maximum value. Hence, $P_s^{i,j}$ becomes a nonlinear function of $v^{i,j}$.
- **Zone III** or full load region with $v_{\text{rated}} \leq v^{i,j} < v_{\text{cut-out}}$. In this zone, the pitch angle is controlled to keep the static available power not higher than the WT nominal power (i.e., $P_s^{i,j} = P_{\text{nom}}$).
- **Zone IV** with $v^{i,j} \geq v_{\text{cut-out}}$. The turbine is pitched out to stop the rotation due to security reasons and $P_s^{i,j} = 0$.

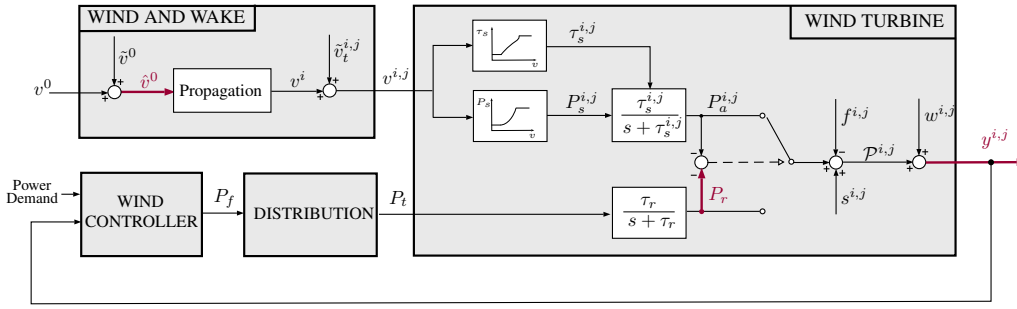


Fig. 3: Structure of the WF benchmark. The signals which are available for FE and FDI are depicted in red.

According to [5], changes in the generated power will be instantaneous; thus, the static available power in the WT (i, j) , denoted as $P_s^{i,j}$, is filtered by a first-order transfer function. This transfer function is characterized by the wind-dependent transfer coefficient τ_s in Fig.2. In all, the dynamic available power in a WT, named after $P_a^{i,j}$, is modelled as

$$\dot{P}_a^{i,j} = \tau_s(v^{i,j}) (P_s^{i,j}(v^{i,j}) - P_a^{i,j}). \quad (1)$$

The WF controller feeds each turbine with the WT static power reference, denoted as P_t . The WT also behaves as a low-pass filter to changes in the power reference; thus, the dynamic power reference of a turbine, P_r , fulfils

$$\dot{P}_r = \tau_r (P_t - P_r), \quad (2)$$

where τ_r is a known transfer function coefficient ($\tau_r = 1.2$ rad/s).

Depending on the values of the dynamic available power $P_a^{i,j}$ and the dynamic power reference P_r , each WT operates in one of the following two main working modes:

- If $P_a^{i,j} < P_r$, the control objective is to maximise the amount of power harvested from the wind and, thus, the generated power equals the dynamic available power in the WT.
- If $P_a^{i,j} \geq P_r$, the control objective is to maintain the generated electrical power equal to the reference.

This behaviour can be modelled as

$$\mathcal{P}^{i,j} = \begin{cases} P_a^{i,j} - f^{i,j} + s^{i,j} & \text{if } P_a^{i,j} < P_r \\ P_r - f^{i,j} + s^{i,j} & \text{otherwise} \end{cases}, \quad (3)$$

where $\mathcal{P}^{i,j}$ is the power generated by the turbine (i, j) and $s^{i,j}$ is a disturbance modelling the drive train oscillations that influence the electrical power generation (i.e., $s^{i,j} = \gamma_p \sin(\sigma_p 2\pi t)$ with $\gamma_p = 1000$ W and $\sigma_p = 10$ Hz). The additive element $f^{i,j}$ is the fault representing a decreased power generation and it is caused by changes in the aerodynamics of the WT due to phenomena such as debris build-up or blade erosion.

The power $\mathcal{P}^{i,j}$ is measured by a sensor whose noise, $w^{i,j}$, can be realistically described by a zero-mean Gaussian noise with variance $\sigma_w^2 = 2.5 \cdot 10^5$ W². Denoting the generated power measurement as $y^{i,j}$, we have

$$y^{i,j} = \mathcal{P}^{i,j} + w^{i,j}. \quad (4)$$

2.2 FE and FDI Signals

The objective of this work is to detect, isolate and estimate the faults representing a decreased power generation in the WTs of the WF. The generated power measurements $y^{i,j}$ ($i = 1, \dots, Y$, $j = 1, \dots, Z$), the dynamic power reference P_r and the WS measurement at the wind mast \hat{v}^0 , are available for FE and FDI (see Fig.3). We assume that the WS at the WTs is not measured and, thus, $v^{i,j}$ must be estimated. As detailed in Appendix 12.1, the measurements \hat{v}^0 can be used to estimate the mean WS v^i through some

propagation strategy. Denoting the propagated mean WS as \hat{v}^i , it yields

$$v^{i,j} = \hat{v}^i + \tilde{v}_p^i + \tilde{v}_t^{i,j}, \quad (5)$$

where \tilde{v}_p^i is the propagation error (i.e., $\tilde{v}_p^i = v^i - \hat{v}^i$), which derives from both the propagation model mismatch and the use of the noisy measurement \hat{v}^0 in the propagation scheme. Note that the total WS estimation error verifies $\tilde{v}^{i,j} = \tilde{v}_p^i + \tilde{v}_t^{i,j}$, where the propagation error \tilde{v}_p^i is common for all the turbines in the i -th row while the turbulence $\tilde{v}_t^{i,j}$ is different for each turbine (i, j) .

It is well known that most continuous-time control systems are implemented digitally [24]. Thus, we present a discrete-time FE and FDI algorithm. The following assumption on the fault $f^{i,j}$ is made.

Assumption 1. The variation $\delta_k^{i,j}$ of the fault $f^{i,j}$ (i.e., $\delta_k^{i,j} = f_{k+1}^{i,j} - f_k^{i,j}$) belongs to $l_2[0, \infty)$.

Remark 1. Assumption 1 is common in FE and it considers faults whose variations are slow with respect to the dynamics of the system and it can cover the typical faults in engineering systems such as abrupt faults and incipient faults [24]. In any case, the strategies developed in this work can be easily extended to faults with more complex dynamics (cf. [25]).

3 Markovian Jump State-Space Modeling

3.1 Parameter Varying Discrete State-Space Modeling

In this section, we develop the discrete state-space model of the power generation system described in Section 2. Let us first define

$$a(v^{i,j}) = e^{-\tau_s(v^{i,j})T_s}, \quad b(v^{i,j}) = 1 - a(v^{i,j}), \quad (6)$$

with T_s being the sampling time that we fix to 1 s. Introducing the state $x^{i,j} = P_a^{i,j}$ and the inputs $u(v^{i,j}) = P_s^{i,j}(v^{i,j})$ and $r = P_r$; it yields,

$$x_{k+1}^{i,j} = a(v_k^{i,j}) x_k^{i,j} + b(v_k^{i,j}) u(v_k^{i,j}), \quad (7a)$$

$$y_k^{i,j} = c(\Delta_k^{i,j}) x_k^{i,j} + d(\Delta_k^{i,j}) r_k - f_k^{i,j} + s_k^{i,j} + w_k^{i,j}, \quad (7b)$$

with $\Delta^{i,j}$ being the power difference defined as

$$\Delta^{i,j} = P_a^{i,j} - P_r = x^{i,j} - r, \quad (8)$$

$$c(\Delta^{i,j}) = \begin{cases} 1 & \text{if } \Delta^{i,j} < 0 \\ 0 & \text{otherwise} \end{cases} \quad d(\Delta^{i,j}) = \begin{cases} 0 & \text{if } \Delta^{i,j} < 0 \\ 1 & \text{otherwise} \end{cases}. \quad (9)$$

The faults verifying Assumption 1 can be modeled through

$$f_{k+1}^{i,j} = f_k^{i,j} + \delta_k^{i,j}, \quad (10)$$

where $\delta^{i,j}$ is the fault variation. Defining the parameter vector $\theta^{i,j} = [v^{i,j} \quad \Delta^{i,j}]^T$ and the extended state vector $z^{i,j} =$

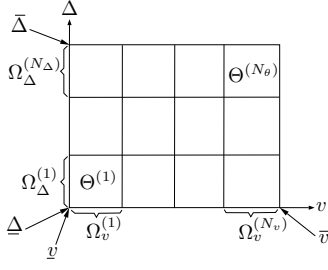


Fig. 4: Partition of the two-dimensional parameter set.

$[x^{i,j} \quad f^{i,j}]^T$, we augment the system (7) into

$$z_{k+1}^{i,j} = \underbrace{\begin{bmatrix} a(v_k^{i,j}) & 0 \\ 0 & 1 \end{bmatrix}}_{A(\theta_k^{i,j})} z_k^{i,j} + \underbrace{\begin{bmatrix} b(v_k^{i,j}) \\ 0 \end{bmatrix}}_{B(\theta_k^{i,j})} \underbrace{u(v_k^{i,j})}_{u(\theta_k^{i,j})} + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_E \delta_k^{i,j}, \quad (11a)$$

$$y_k^{i,j} = \underbrace{\begin{bmatrix} c(\Delta_k^{i,j}) & -1 \end{bmatrix}}_{C(\theta_k^{i,j})} z_k^{i,j} + \underbrace{d(\Delta_k^{i,j})}_{D(\theta_k^{i,j})} r_k + s_k^{i,j} + w_k^{i,j}, \quad (11b)$$

$$f_k^{i,j} = \underbrace{\begin{bmatrix} 0 & 1 \end{bmatrix}}_F z_k^{i,j}. \quad (11c)$$

For ease of readability, let us omit hereafter the dependence of the variables on the number of turbine (i, j) (e.g. z_k stands for $z_k^{i,j}$). The extended system model (11) results in

$$z_{k+1} = A(\theta_k) z_k + B(\theta_k) u(\theta_k) + E \delta_k, \quad (12a)$$

$$y_k = C(\theta_k) z_k + D(\theta_k) r_k + s_k + w_k, \quad (12b)$$

$$f_k = F z_k. \quad (12c)$$

The system dynamics (12) depends on the system matrices $A(\theta_k)$, $B(\theta_k)$, $C(\theta_k)$ and $D(\theta_k)$. According to (6) and (11), the behaviour of $A(\theta_k)$ and $B(\theta_k)$ switches between the estimated WS zones as shown in Fig.2. According to (9) and (11), $C(\theta_k)$ and $D(\theta_k)$ are switching matrices whose value depends on the WT working mode determined by the sign of the parameter Δ_k .

3.2 Markovian Jump Discrete State-space Modeling

The parameters v and Δ can be considered to be bounded by the sets

$$v \in \Omega_v, \quad \Omega_v := \{v < v < \bar{v}\}, \quad (13a)$$

$$\Delta \in \Omega_\Delta, \quad \Omega_\Delta := \{\underline{\Delta} < \Delta < \bar{\Delta}\}, \quad (13b)$$

where v , \bar{v} , $\underline{\Delta}$ and $\bar{\Delta}$ are the minimum and maximum possible values of these parameters (which we fix to $v = 0$ m/s, $\bar{v} = 30$ m/s, $\underline{\Delta} = -4.8$ MW and $\bar{\Delta} = 4.8$ MW). The parameter vector $\theta = \begin{bmatrix} v \\ \Delta \end{bmatrix}$ lies then in $\Theta = \Omega_v \times \Omega_\Delta$. Let us partition the parameter set Θ into N subsets $\Theta^{(p)}$ (i.e., $\Theta = \{\Theta^{(p)}\}_{p \in \{1, \dots, N\}}$) by dividing Ω_v into N_v intervals $\Omega_v^{(p_v)}$ (i.e., $\Omega_v = \{\Omega_v^{(p_v)}\}_{p_v \in \{1, \dots, N_v\}}$) and Ω_Δ into N_Δ intervals $\Omega_\Delta^{(p_\Delta)}$ (i.e., $\Omega_\Delta = \{\Omega_\Delta^{(p_\Delta)}\}_{p_\Delta \in \{1, \dots, N_\Delta\}}$), i.e.,

$$\Theta = \{(\Omega_v^{(1)}, \Omega_\Delta^{(1)}), \dots, (\Omega_v^{(1)}, \Omega_\Delta^{(p_\Delta)}), \dots, (\Omega_v^{(p_v)}, \Omega_\Delta^{(1)}), \dots, (\Omega_v^{(p_v)}, \Omega_\Delta^{(p_\Delta)})\} \quad (14)$$

with $N = N_v \cdot N_\Delta$ (see Fig.4). The partition is one such that each interval $\Omega_v^{(p_v)}$ lies in a single WS zone and each interval $\Omega_\Delta^{(p_\Delta)}$ lies in a single WT working mode (see Table 1).

Table 1 Partition of the parameter sets Ω_v and Ω_Δ

p_v	Wind speed boundaries	Wind zone
1	$v \leq v < v_1$	Zone I
2	$v_1 \leq v < v_2$	Zone I
\vdots	\vdots	\vdots
n_1	$v_{n_1-1} \leq v < v_{cut-in}$	Zone I
$n_1 + 1$	$v_{cut-in} \leq v < v_{n_1+1}$	Zone II
\vdots	\vdots	\vdots
n_2	$v_{n_2-1} \leq v < v_{rated}$	Zone II
$n_2 + 1$	$v_{rated} \leq v < v_{n_2+1}$	Zone III
\vdots	\vdots	\vdots
n_3	$v_{n_3-1} \leq v < v_{cut-out}$	Zone III
$n_3 + 1$	$v_{cut-out} \leq v < v_{n_3+1}$	Zone IV
\vdots	\vdots	\vdots
N_v	$v_{N_v-1} \leq v < \bar{v}$	Zone IV

p_Δ	Power difference boundaries	Working mode
1	$\Delta \leq \Delta < \Delta_1$	$P_a < P_r$
2	$\Delta_1 \leq \Delta < \Delta_2$	$P_a < P_r$
\vdots	\vdots	\vdots
n'_1	$\Delta_{n'_1-1} \leq \Delta < 0$	$P_a < P_r$
$n'_1 + 1$	$0 \leq \Delta < \Delta_{n'_1+1}$	$P_a \geq P_r$
\vdots	\vdots	\vdots
N_Δ	$\Delta_{N_\Delta-1} \leq \Delta < \bar{\Delta}$	$P_a \geq P_r$

For such partition, the system state-space matrices $C(\theta_k)$ and $D(\theta_k)$ associated with $\theta_k \in \Theta^{(p)}$ are constant and they can be expressed as

$$C(\theta_k) = C_p, \quad D(\theta_k) = D_p. \quad (15)$$

The system state-space matrices $A(\theta_k)$ and $B(\theta_k)$ associated with $\theta_k \in \Theta^{(p)}$ are constant if the interval $\Omega_v^{(p_v)}$ defining $\Theta^{(p)}$ (i.e., $\Theta^{(p)} = (\Omega_v^{(p_v)}, \Omega_\Delta^{(p_\Delta)})$) lies in the WS zones I, III or IV (i.e., $\Omega_v^{(p_v)} \in [v, v_{cut-in}) \cup [v_{rated}, \bar{v})$) because the discrete transfer coefficient $a(v)$ is constant in these zones. If the interval $\Omega_v^{(p_v)}$ lies in the zone II (i.e., $\Omega_v^{(p_v)} \in [v_{cut-in}, v_{rated})$), these matrices are not constant because the discrete transfer coefficient $a(v)$ is described by the function in Fig.2. In all, we express the matrices $A(\theta_k)$ and $B(\theta_k)$ associated with $\theta_k \in \Theta^{(p)}$ as

$$A(\theta_k) = \sum_{m=1}^M \alpha_p^m(\theta_k) A_p^m, \quad B(\theta_k) = \sum_{m=1}^M \alpha_p^m(\theta_k) B_p^m, \quad (16)$$

with $\sum_{m=1}^M \alpha_p^m(\theta_k) = 1$ and $0 \leq \alpha_p^m(\theta_k) \leq 1$ for $m = 1, \dots, M$. The matrix A_p^m denotes the m -th vertex of the convex polytope in which $A(\theta_k)$ lies whenever $\theta_k \in \Theta^{(p)}$. Note that in the WS zones I, III and IV, we have $A_p^m = A_p$ and $\alpha_p^m = 1/M$ for $l = 1, \dots, M$. Similar applies to $B(\theta_k)$.

Define the membership signal ξ^v whose value at a sample k equals the number p_v of the interval $\Omega_v^{(p_v)}$ in which the WS v lies at the sample k ; similar applies to ξ^Δ w.r.t. Δ and to ξ w.r.t. θ :

$$\xi_k^v = p_v \quad \text{if} \quad v_k \in \Omega_v^{(p_v)}, \quad (17a)$$

$$\xi_k^\Delta = p_\Delta \quad \text{if} \quad \Delta_k \in \Omega_\Delta^{(p_\Delta)}, \quad (17b)$$

$$\xi_k = p \quad \text{if} \quad \theta_k \in \Theta^{(p)}. \quad (17c)$$

Research has shown that the stochastic behaviour of the mean WS can be represented as a Markovian process [13, 14]. This behaviour has been exploited in recent WT control strategies as the ones in [15–17]. Motivated by this research, we assume that $\{\xi_k^v\}$ is a discrete homogenous Markov chain taking values in the finite set

$\{1, \dots, N_v\}$. Since the mean WS affects the WT working mode, we also assume that $\{\xi_k^\Delta\}$ is a discrete homogenous Markov chain taking values in the finite set $\{1, \dots, N_\Delta\}$. Both assumptions are considered in the following assumption over $\{\xi_k\}$.

Assumption 2. *The membership $\{\xi_k\}$ is governed by a discrete homogenous Markov chain whose states are in the finite set $S = \{1, \dots, N\}$ with the transition probability matrix*

$$\Pi = \begin{bmatrix} \pi_{11} & \dots & \pi_{1N} \\ \vdots & \ddots & \vdots \\ \pi_{N1} & \dots & \pi_{NN} \end{bmatrix} \in \mathbb{R}^{N \times N} \quad (18)$$

and π_{pq} being the transition probability defined as

$$\pi_{pq} = \mathbf{Pr}\{\xi_{k+1} = q | \xi_k = p\}, \quad (19)$$

where $\sum_{q=1}^N \pi_{pq} = 1$ for any $p \in S$.

Remark 2. ([26]) *Provided (14), we have that $\Pi = \Pi^v \otimes \Pi^\Delta$, with $\Pi^v = [\pi_{p_v q_v}^v]_{p_v, q_v \in \{1, \dots, N_v\}}$ and $\Pi^\Delta = [\pi_{p_\Delta q_\Delta}^\Delta]_{p_\Delta, q_\Delta \in \{1, \dots, N_\Delta\}}$; $\pi_{p_v q_v}^v$ and $\pi_{p_\Delta q_\Delta}^\Delta$ being defined in analogy to (19) w.r.t. ξ^v and ξ^Δ .*

Remark 3. ([27]) *The probabilities π_{pq} can be obtained through numerical simulations by sampling the data in the separate subsets $\Theta^{(p)}$ and computing $\pi_{pq} = \frac{\text{observed transitions from state } p \text{ to } q}{\text{observed occurrences of state } p}$. Similar applies to $\pi_{p_v q_v}$ and $\pi_{p_\Delta q_\Delta}$.*

4 FE at a Wind Turbine Level

4.1 FE Architecture

To achieve FE, we build the following model-based observer for the extended system (12):

$$\hat{z}_{k+1} = A(\hat{\theta}_k) \hat{z}_k + B(\hat{\theta}_k) u(\hat{\theta}_k) + \quad (20a)$$

$$L(\hat{\theta}_k) (y_k - C(\hat{\theta}_k) \hat{z}_k - D(\hat{\theta}_k) r_k),$$

$$\hat{f}_k = F \hat{z}_k + K(\hat{\theta}_k) (y_k - C(\hat{\theta}_k) \hat{z}_k - D(\hat{\theta}_k) r_k), \quad (20b)$$

where $\hat{\theta} = [\hat{v} \ \hat{\Delta}]^T$ is the estimated parameter vector in which \hat{v} is the open-loop propagated WS (see (5)) and $\hat{\Delta}$ is the closed-loop estimated power difference computed as $\hat{\Delta} = P^r - \hat{x}$ (with \hat{x} being the first state of \hat{z}). The estimated parameter vector $\hat{\theta}$ does also lie in the set Θ , partitioned into the subsets $\{\Theta^{(p)}\}_{p \in \{1, \dots, N\}}$ as explained in Section 3.2.

The matrices $L(\hat{\theta}_k) \in \mathbb{R}^{2 \times 1}$ and $K(\hat{\theta}_k) \in \mathbb{R}^1$ are the parameter-dependent design gain matrices of appropriate dimensions. Provided the switching behaviour of the system state-space matrices in (15)-(16), we fix $L(\hat{\theta}_k) \in \mathbb{R}^{2 \times 1}$ and $K(\hat{\theta}_k) \in \mathbb{R}^1$ as

$$L(\hat{\theta}_k) = L_p \quad \text{if } \hat{\theta}_k \in \Theta^{(p)}, \quad (21a)$$

$$K(\hat{\theta}_k) = K_p \quad \text{if } \hat{\theta}_k \in \Theta^{(p)}, \quad (21b)$$

where L_p and K_p are constant matrices associated to $\hat{\theta}_k \in \Theta^{(p)}$ and must be designed. Given (17), these matrices can be also expressed as $L(\hat{\theta}_k) = L_p$ if $\hat{\xi}_k = p$ and $K(\hat{\theta}_k) = K_p$ if $\hat{\xi}_k = p$.

Remark 4. *Note that certain degree of conservatism is introduced when using switched observer gain matrices in the form of (21) for the regions lying in the WS zone II, where the state-space matrices $A(\hat{\theta}_k)$ and $B(\hat{\theta}_k)$ are not constant (see (16)). Even though switched polytopic gain matrices could be used instead, the small variations experienced by these matrices (see Fig.2) justify the choice in (21) that reduces the observer complexity.*

Define the extended state estimation error as $\tilde{z}_k = z_k - \hat{z}_k$ and the FE error as $\tilde{f}_k = f_k - \hat{f}_k$. It follows that

$$\tilde{z}_{k+1} = (A(\hat{\theta}_k) - L(\hat{\theta}_k) C(\hat{\theta}_k)) \tilde{z}_k + E \delta_k + \bar{R} g_k - \quad (22a)$$

$$L(\hat{\theta}_k) (R h_k + s_k + w_k),$$

$$\tilde{f}_k = (F - K(\hat{\theta}_k) C(\hat{\theta}_k)) \tilde{z}_k - K(\hat{\theta}_k) (R h_k + s_k + w_k), \quad (22b)$$

with $\bar{R} = [1 \ 0]^T$. The error depends then on the fault variation δ_k , the oscillation s_k , the noise w_k and the disturbances $g_k = g_k(\hat{\theta}_k, \theta_k)$ and $h_k = h_k(\hat{\theta}_k, \theta_k)$, which stem from using $\hat{\theta}$ instead of θ in the estimation algorithm. As (22) refers to the turbine (i, j) , g_k and h_k refer in fact to $g_k^{i,j}$ and $h_k^{i,j}$ defined as

$$g_k^{i,j} = (a(v_k^{i,j}) - a(\hat{v}_k^{i,j})) x_k^{i,j} + (\bar{u}(v_k^{i,j}) - \bar{u}(\hat{v}_k^{i,j})), \quad (23a)$$

$$h_k^{i,j} = (c(\Delta_k^{i,j}) - c(\hat{\Delta}_k^{i,j})) x_k^{i,j} + (d(\Delta_k^{i,j}) - d(\hat{\Delta}_k^{i,j})) r_k, \quad (23b)$$

with $\bar{u}(v_k^{i,j}) = b(v_k^{i,j}) u(v_k^{i,j})$. Specifically, $g_k^{i,j}$ stands for the disturbance derived from the open-loop wind propagation strategy and $h_k^{i,j}$ is the disturbance derived from the closed-loop dynamic available power estimation strategy. Note that $h_k^{i,j}$ is in fact equal to 0 (if the WT operating mode is correctly estimated) or to the difference $\Delta^{i,j}$ (otherwise).

Provided (5), $g_k^{i,j}$, which is caused by the total WS estimation error, can be expressed as

$$g_k^{i,j} = e_k^{i,j} + \varepsilon_k^{i,j}, \quad (24)$$

where $e_k^{i,j}$ is caused by the wind propagation error and $\varepsilon_k^{i,j}$ is caused by the wind turbulence, i.e.,

$$e_k^{i,j} = (a(v_k^{i,j}) - a(\hat{v}_k^{i,j})) x_k^{i,j} + (\bar{u}(v_k^{i,j}) - \bar{u}(\hat{v}_k^{i,j})), \quad (25a)$$

$$\varepsilon_k^{i,j} = (a(v_k^{i,j}) - a(\hat{v}_k^{i,j})) x_k^{i,j} + (\bar{u}(v_k^{i,j}) - \bar{u}(\hat{v}_k^{i,j})). \quad (25b)$$

In (25), the first summands are negligible w.r.t. to the second summands. Hence, we omit the dependence of the disturbance e^i on the column j . The following assumption on e^i , $\varepsilon^{i,j}$ and $h^{i,j}$ is made.

Assumption 3. *The disturbances e^i , $\varepsilon^{i,j}$ and $h^{i,j}$ can be considered to be bounded as*

$$\|e^i\|_\infty \leq \bar{e}_p, \quad \|\varepsilon^{i,j}\|_\infty \leq \bar{\varepsilon}_p, \quad \|h^{i,j}\|_\infty \leq \bar{h}_p, \quad \text{if } \theta \in \Theta^{(p)}. \quad (26)$$

Details on the computation of these bounds are included in Appendix 12.2. Note that it would have been possible to consider a row-dependent bound of the disturbance caused by the propagation error (i.e., $|e^i| \leq \bar{e}_p^i$) because this error generally increases with the distance to wind mast. For sake of simplicity, we introduce some conservatism by neglecting this dependence. Hence, \bar{e}_p verifies $\bar{e}_p = \max_i \bar{e}_p^i$.

For brevity, we omit again the dependence on the number of row i and on the number column j . Taking (24) and (26) into account, the summands $\bar{R} g_k$ and h_k in (22) can be expressed as

$$\bar{R} g_k = \mathcal{G}(\hat{\theta}_k) \lambda_k + G(\hat{\theta}_k) \mu_k, \quad h_k = H(\hat{\theta}_k) \varphi_k, \quad (27)$$

with $\lambda_k \in [-1, 1]$, $\mu_k \in [-1, 1]$ and $\varphi_k \in [-1, 1]$ satisfying

$$\lambda_k = e_k / \bar{e}_p, \quad \mu_k = \varepsilon_k / \bar{\varepsilon}_p, \quad \varphi_k = h_k / \bar{h}_p, \quad (28)$$

if $\theta_k \in \Theta^{(p)}$ and where $\mathcal{G}(\hat{\theta}_k)$, $G(\hat{\theta}_k)$ and $H(\hat{\theta}_k)$ are defined as

$$\mathcal{G}(\hat{\theta}_k) = \mathcal{G}_p \quad \text{if } \hat{\theta}_k \in \Theta^{(p)}, \quad \mathcal{G}_p = \bar{R} \bar{e}_p, \quad (29a)$$

$$G(\hat{\theta}_k) = G_p \quad \text{if } \hat{\theta}_k \in \Theta^{(p)}, \quad G_p = \bar{R} \bar{\varepsilon}_p, \quad (29b)$$

$$H(\hat{\theta}_k) = H_p \quad \text{if } \hat{\theta}_k \in \Theta^{(p)}, \quad H_p = \bar{h}_p. \quad (29c)$$

In all, we rewrite (22) as

$$\tilde{z}_{k+1} = \mathcal{A}(\hat{\theta}_k) \tilde{z}_k + \mathcal{B}(\hat{\theta}_k) \mathcal{W}_k + E \delta_k, \quad (30a)$$

$$\tilde{f}_k = \mathcal{C}(\hat{\theta}_k) \tilde{z}_k + \mathcal{D}(\hat{\theta}_k) \mathcal{W}_k, \quad (30b)$$

with

$$\mathcal{A}(\hat{\theta}_k) = A(\hat{\theta}_k) - L(\hat{\theta}_k) C(\hat{\theta}_k),$$

$$C(\hat{\theta}_k) = F - K(\hat{\theta}_k) C(\hat{\theta}_k),$$

$$\mathcal{B}(\hat{\theta}_k) = [\mathcal{G}(\hat{\theta}_k) \quad G(\hat{\theta}_k) \quad -L(\hat{\theta}_k) H(\hat{\theta}_k) \quad -L(\hat{\theta}_k) \quad -L(\hat{\theta}_k)],$$

$$\mathcal{D}(\hat{\theta}_k) = [0 \quad 0 \quad -K(\hat{\theta}_k) H(\hat{\theta}_k) \quad -K(\hat{\theta}_k) \quad -K(\hat{\theta}_k)],$$

$$\mathcal{W}_k = [\lambda_k \quad \mu_k \quad \varphi_k \quad s_k \quad w_k]^T.$$

Note that $\mathcal{A}_p(\hat{\theta}_k)$, \mathcal{B}_p , \mathcal{C}_p and \mathcal{D}_p are defined using the corresponding matrices associated with $\theta_k \in \Theta^{(p)}$ (e.g., $\mathcal{A}_p(\hat{\theta}_k) = \sum_{m=1}^M \alpha_p^m(\hat{\theta}_k) (A_p^m - L_p C_p)$ and $\mathcal{C}_p = F - K_p C_p$).

4.2 FE Design

The FE error caused by δ describes the fault sensitivity of the observer and the error caused by \mathcal{W} describes the robustness of the observer against disturbances and noises. We characterize them through the \mathcal{H}_∞ performance of the observer w.r.t. δ and \mathcal{W} , respectively. To that end, we use the formulation based on LMIs in the following theorem. For the sake of generality, we denote the size of the inputs in (30) as n_δ , n_λ , n_μ , n_φ , n_s and n_w verifying $n_\delta = n_\lambda = n_\mu = n_\varphi = n_s = n_w = 1$ in this case of FE at a WT level.

Theorem 1. Consider the observer (20) applied to the system (12). If there exist positive scalars γ_λ , γ_μ , γ_φ , γ_s , γ_w and γ_δ ; full matrices of appropriate dimensions X_p , K_p and V_p ; and symmetric matrices of appropriate dimensions Q_p , for $p, q = 1, \dots, N$ and $m = 1, \dots, M$ fulfilling the LMIs

$$\Xi_p^m = \begin{bmatrix} \Omega_p & \bar{\Lambda}_p^m T & \bar{\Phi}_p^T & 0 \\ \bar{\Lambda}_p^m & Q_p & 0 & C_p^T \\ \bar{\Phi}_p & 0 & \Gamma & \Upsilon_p^T \\ 0 & C_p & \Upsilon_p & I \end{bmatrix} \succ 0, \quad (31)$$

with

$$\Omega_p = \bigoplus_{q=1}^N V_p + V_p^T - Q_q,$$

$$\bar{\Lambda}_p^m = [\sqrt{\pi_{p1}} \Lambda_p^m \quad \dots \quad \sqrt{\pi_{pN}} \Lambda_p^m], \quad \Lambda_p^m = V_p A_p^m - X_p C_p,$$

$$\bar{\Phi}_p = [\sqrt{\pi_{p1}} \Phi_p \quad \dots \quad \sqrt{\pi_{pN}} \Phi_p],$$

$$\Phi_p = [V_p G_p \quad V_p G_p \quad -X_p H_p \quad -X_p \quad -X_p \quad V_p E],$$

$$\Upsilon_p = [0 \quad 0 \quad -K_p H_p \quad -K_p \quad -K_p \quad 0],$$

$$C_p = F - K_p C_p,$$

$$\Gamma = \gamma_\lambda I_{n_\lambda} \oplus \gamma_\mu I_{n_\mu} \oplus \gamma_\varphi I_{n_\varphi} \oplus \gamma_s I_{n_s} \oplus \gamma_w I_{n_w} \oplus \gamma_\delta I_{n_\delta},$$

then defining $L_p = V_p^{-1} X_p$, the following holds for all $\hat{\theta}_k \in \Theta$.

1. In the absence of disturbances, noises and faults (i.e., $\mathcal{W}_k = 0$ and $\delta_k = 0$), the extended state estimation error (30) converges asymptotically to zero in average.

2. Under null initial conditions, the expected value of the FE error is bounded as

$$\mathbb{E}\{\|\tilde{f}\|_{\text{RMS}}^2\} < \gamma_\lambda n_\lambda \|\lambda\|_\infty^2 + \gamma_\mu n_\mu \|\mu\|_\infty^2 + \gamma_\varphi n_\varphi \|\varphi\|_\infty^2 + \gamma_s \|s\|_{\text{RMS}}^2 + \gamma_w \|w\|_{\text{RMS}}^2 + \gamma_\delta n_\delta \|\delta\|_\infty^2, \quad (32)$$

with $\mathbb{E}\{\|\tilde{f}\|_{\text{RMS}}^2\} = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K \mathbb{E}\{\tilde{f}_k^T \tilde{f}_k | \hat{\theta}_{k-1} \in \Theta^{(p)}\}$.

Proof: If (31) holds for $p, q = 1, \dots, N$, we have that $Q_q \succ 0$ and that V_p is a nonsingular matrix because $V_p + V_p^T - Q_q \succ 0$. Hence, we can state that $(Q_q - V_p) Q_q^{-1} (Q_q - V_p)^T \succeq 0$ which implies that

$$V_p + V_p^T - Q_q \preceq V_p Q_q^{-1} V_p^T.$$

Thus, the matrix inequalities which result from replacing $V_p + V_p^T - Q_q$ by $V_p Q_q^{-1} V_p^T$ in (31) are also positive definite. Let us substitute X_p by $V_p L_p$ and apply a congruence transformation by $(\bigoplus_{q=1}^N V_p^{-1}) \oplus I$ to the result. Multiplying the matrix inequality by $\alpha_p^m(\hat{\theta}_k)$, summing for $m = 1, \dots, M$, taking Schur's complements, pre-multiplying the result by $[\tilde{z}_k^T \quad \mathcal{W}_k^T \quad \delta_k^T]$ and post-multiplying by its transpose lead to

$$\begin{aligned} & \sum_{q=1}^N \pi_{pq} (\mathcal{A}_p(\hat{\theta}_k) \tilde{z}_k + \mathcal{B}_p \mathcal{W}_k + E \delta_k)^T Q_q (\mathcal{A}_p(\hat{\theta}_k) \tilde{z}_k + \\ & \mathcal{B}_p \mathcal{W}_k + E \delta_k) - \tilde{z}_k^T Q_p \tilde{z}_k + \tilde{f}_k^T \tilde{f}_k - \gamma_\lambda \lambda_k^T \lambda_k - \gamma_\mu \mu_k^T \mu_k - \\ & \gamma_\varphi \varphi_k^T \varphi_k - \gamma_s s_k^T s_k - \gamma_w w_k^T w_k - \gamma_\delta \delta_k^T \delta_k < 0, \end{aligned} \quad (33)$$

for all $p = 1, \dots, N$ and where we have taken into account that $\sum_{m=1}^M \alpha_p^m(\hat{\theta}_k) = 1$.

Now, let us define the Lyapunov function $V_k = \tilde{z}_k^T Q(\hat{\theta}_k) \tilde{z}_k$ equal to $V_k = \tilde{z}_k^T Q_p \tilde{z}_k$ for $\hat{\theta}_k \in \Theta^{(p)}$ and $p = 1, \dots, N$.

1. In the absence of disturbances noises and faults (i.e., $\mathcal{W}_k = 0$ and $\delta_k = 0$), expression (33) leads to

$$\tilde{z}_k^T \left(\sum_{q=1}^N \pi_{pq} \mathcal{A}_p(\hat{\theta}_k) Q_q \mathcal{A}_p(\hat{\theta}_k) \right) \tilde{z}_k - \tilde{z}_k^T Q_p \tilde{z}_k < 0,$$

for all $p = 1, \dots, N$, which assures $\mathbb{E}\{V_{k+1} | \hat{\theta}_k \in \Theta^{(p)}\} - V_k < 0$ guaranteeing that the extended state estimation error (30) converges asymptotically to zero in average for all $\hat{\theta}_k \in \Theta$.

2. For brevity, let us denote $\mathbb{E}\{V_{k+1} | \hat{\theta}_k \in \Theta^{(p)}\}$ as $\mathbb{E}\{V_{k+1} | \hat{\theta}_k\}$ and let us not include in the next that the inequalities are fulfilled for all $\hat{\theta}_k \in \Theta$. Taking conditional expectation given $\hat{\theta}_{k-1}$ over expression (33) leads to

$$\begin{aligned} & \mathbb{E}\{V_{k+1} | \hat{\theta}_k\} - \mathbb{E}\{V_k | \hat{\theta}_{k-1}\} + \mathbb{E}\{\tilde{f}_k^T \tilde{f}_k | \hat{\theta}_{k-1}\} - \\ & \gamma_\lambda n_\lambda \|\lambda\|_\infty^2 - \gamma_\mu n_\mu \|\mu\|_\infty^2 - \gamma_\varphi n_\varphi \|\varphi\|_\infty^2 - \\ & \gamma_s \|s\|_{\text{RMS}}^2 - \gamma_w \|w\|_{\text{RMS}}^2 - \gamma_\delta n_\delta \|\delta\|_\infty^2 < 0, \end{aligned} \quad (34)$$

because $\mathbb{E}\{V_{k+1} | \hat{\theta}_k, \hat{\theta}_{k-1}\} = \mathbb{E}\{V_{k+1} | \hat{\theta}_k\}$ and the exogenous signals are independent of $\hat{\theta}_{k-1}$. In (34), we have also taken into account that λ is a deterministic signal and

$$\mathbb{E}\{\lambda_k^T \lambda_k\} = \lambda_k^T \lambda_k \leq n_\lambda \|\lambda\|_\infty^2$$

(similar applies to μ , φ and δ), and that $\mathbb{E}\{s_k^T s_k\} = \|s\|_{\text{RMS}}^2$ because s_k is zero-mean (similar applies to w_k).

Under null initial conditions ($V_0 = 0$), adding the aforementioned expression from $k = 0$ to $k = K - 1$ leads to

$$\begin{aligned} & \sum_{k=0}^{K-1} \mathbb{E}\{\tilde{f}_k^T \tilde{f}_k | \hat{\theta}_{k-1}\} < \sum_{k=0}^{K-1} (\gamma_\lambda n_\lambda \|\lambda\|_\infty^2 + \gamma_\mu n_\mu \|\mu\|_\infty^2 + \\ & \gamma_\varphi n_\varphi \|\varphi\|_\infty^2 + \gamma_s \|s\|_{\text{RMS}}^2 + \gamma_w \|w\|_{\text{RMS}}^2 + \gamma_\delta n_\delta \|\delta\|_\infty^2), \end{aligned}$$

because $\sum_{k=0}^{K-1} (\mathbb{E}\{V_{k+1} | \hat{\theta}_k\} - \mathbb{E}\{V_k | \hat{\theta}_{k-1}\}) = \mathbb{E}\{V_{K+1} | \hat{\theta}_K\}$ and $\mathbb{E}\{V_{K+1} | \hat{\theta}_K\} > 0$. Dividing this expression by K and taking the limit when K tends to infinity leads to (32). \square

There exists thus a trade-off between ameliorating the fault tracking ability of the observer and its robustness [18]. From the practical viewpoint, it is of considerable interest to achieve certain tracking ability and to minimize the effect of the disturbances and noises on the estimations [19]. Thus, we propose to design the observer gain matrices through the following optimization problem:

$$\begin{aligned} & \text{minimize} \quad \gamma_\lambda n_\lambda \|\lambda\|_\infty^2 + \gamma_\mu n_\mu \|\mu\|_\infty^2 + \gamma_\varphi n_\varphi \|\varphi\|_\infty^2 + \\ & \quad \gamma_s \|s\|_{\text{RMS}}^2 + \gamma_w \|w\|_{\text{RMS}}^2 \\ & \text{subject to} \quad \gamma_\delta \leq \bar{\gamma}_\delta, \\ & \quad \Xi_p^m > 0, \quad \forall p, m \end{aligned} \quad (35)$$

along the variables $\gamma_\lambda, \gamma_\mu, \gamma_\varphi, \gamma_s, \gamma_w, \gamma_\delta, X_p, K_p, V_p, Q_p$ and P_p ($p = 1, \dots, N$) and with $\bar{\gamma}_\delta$ being the required \mathcal{H}_∞ performance describing the fault tracking ability of the observer.

Remark 5. The use of the signals (28) in the design process allows considering the differences in the bounds (26) among the different subsets $\Theta^{(p)}$. Effectively, we now have that the value of these bounds is introduced through the matrices \mathcal{G}_p, G_p and H_p and $\|\lambda\|_\infty^2 = \|\mu\|_\infty^2 = \|\varphi\|_\infty^2 = 1$. If the signals e, ϵ and h were used instead, the maximum value of the bounds (26) (e.g., $\max_p \bar{e}_p$) would be the only information introduced in the design. Note also that if the bounds (26) are not completely accurate, $\|\lambda\|_\infty^2, \|\mu\|_\infty^2$ and $\|\varphi\|_\infty^2$ can be considered as tuning parameters regarding the bounding accuracy. Similar applies to $\|s\|_{\text{RMS}}^2$ and $\|w\|_{\text{RMS}}^2$ that can be easily derived from the known parameters γ_p and σ_w but whose values can be increased with the uncertainty over these parameters.

Taking Remark 4 and Remark 5 into account, we deduce that the conservatism is reduced when a lot of gridding intervals are utilized for the partition of the set Θ . Hence, the density N of the grid $\{\Theta^{(p)}\}_{p \in \{1, \dots, N\}}$ is to be determined from a trade-off between having a few gridding intervals that ensure reduced computational burden but introduce conservatism or having a lot of gridding intervals causing heavy computational times but reduced performance conservatism [28]. As specified in Appendix 12.2, we choose $N_v = 6$ and $N_\Delta = 4$. This gridding is a posteriori validated through the numerical simulations in Section 8.

5 FDI at a Wind Turbine Level

We set the following decision for FDI:

$$\begin{cases} \text{if } |\hat{f}_k| \geq J_k & \text{Fault} \\ \text{otherwise} & \text{No fault} \end{cases}, \quad (36)$$

where J is an adaptive threshold covering the uncertainties affecting the fault estimate in fault-free conditions. As recalled in [20], the model-based setting of thresholds usually leads to too conservative thresholds which result in poor fault diagnosability. It is the state of the art in real applications to optimally set thresholds on the basis of tests in the real application environment. In this context, we propose to compute the threshold through a multivariate linear model over the

estimated parameters $\hat{\theta}$ and to obtain the coefficients of this model using a set of fault-free training data as detailed hereafter.

Remark 6. Norm-based constant thresholding can be performed using the RMS-norm bound of the FE error in (32) (with $\gamma_\delta = 0$). Assuming zero-mean disturbances, one can apply the Chebyshev's inequality and obtain a stochastic threshold for certain confidence level. However, this approach ignores the real statistical distribution of the error and leads to too loose bounds [29].

Provided the switching behaviour of the system, we use different coefficients for each WT working mode. For this purpose, we define $\hat{\eta}^M$ ($M = 1, 2, 3$) as a signal indicating whether the WT is estimated to be operating in mode M :

$$\hat{\eta}_k^1 = \begin{cases} 1 & \text{if } \hat{\xi}_k^\Delta \in [1, n'_1 - n'] \\ 0 & \text{otherwise} \end{cases}, \quad (37a)$$

$$\hat{\eta}_k^2 = \begin{cases} 1 & \text{if } \hat{\xi}_k^\Delta \in (n'_1 - n', n'_1 - n'] \\ 0 & \text{otherwise} \end{cases}, \quad (37b)$$

$$\hat{\eta}_k^3 = \begin{cases} 1 & \text{if } \hat{\xi}_k^\Delta \in (n'_1 - n', N_\Delta] \\ 0 & \text{otherwise} \end{cases}, \quad (37c)$$

with $\hat{\xi}_k^\Delta$ being the estimated membership function of power difference (i.e., $\hat{\xi}_k^\Delta = p_\Delta \mid \hat{\Delta}_k \in \Omega_\Delta^{(p_\Delta)}$) and n' being a design parameter that defines an uncertain intermediate mode. Effectively, mode $M = 1$ refers to the cases where $\hat{\Delta}_k < \Delta_{n'_1 - n'} \leq 0$ and the generated power equals the dynamic available power in the WT. Mode $M = 3$ refers to the cases where $0 \leq \Delta_{n'_1 - n'} \leq \hat{\Delta}_k$ and the generated power equals the dynamic power reference (see Table 1). For its part, mode $M = 2$ is the transition mode defined by $n' > 0$ that takes the working mode estimation error into account. In all, the proposed multivariate linear model for computing the adaptive threshold is

$$J_k = X_k \beta, \quad (38)$$

with X_k being the variable vector defined as

$$X_k = [\hat{\eta}_k^1 \quad \hat{\eta}_k^2 \quad \hat{\eta}_k^3 \mid \hat{\eta}_k^1 \hat{v}_k \quad \hat{\eta}_k^2 \hat{v}_k \quad \hat{\eta}_k^3 \hat{v}_k \mid \hat{\eta}_k^1 \hat{\Delta}_k \quad \hat{\eta}_k^2 \hat{\Delta}_k \quad \hat{\eta}_k^3 \hat{\Delta}_k]$$

and $\beta \in \mathbb{R}^{9 \times 1}$ being the coefficient vector. Thus, in Mode M , the threshold verifies $J_k = \beta[M] + \hat{v}_k \beta[M+3] + \hat{\Delta}_k \beta[M+6]$.

Remark 7. Less conservative thresholds would be achievable by taking account of all the switchings of $\hat{\theta}_k$ among the subsets of the parameter vector. To do so, the alternative variable vector $\bar{X}_k = [\hat{\eta}_k^1 \dots \hat{\eta}_k^N] \otimes [1 \quad \hat{\theta}_k^T]$ (with $\bar{\beta} \in \mathbb{R}^{3N \times 1}$) would be used instead. Here, $\hat{\eta}_k^p$ is equal to 1 if $\hat{\xi}_k = p$ and it is equal to 0 otherwise. Similarly, it would be possible to take account of the process dynamic behaviour by using the alternative variable vector $\bar{X}'_k = [X_k \dots X_{k-O}]$ (with $\bar{\beta}' \in \mathbb{R}^{3O \times 1}$ and O being the number of past considered samples). However, the simplified form in (38) reduces the complexity of the threshold computation while it has provided satisfactory results in the simulations in Section 8.

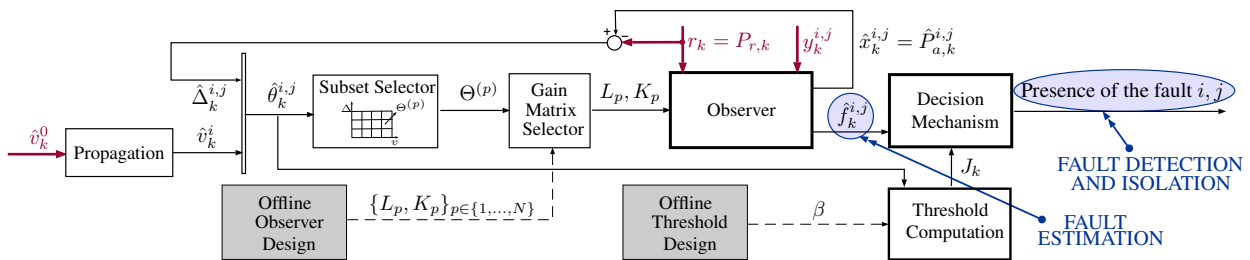


Fig. 5: FE and FDI strategy at a wind turbine level.

To obtain β , we collect the fault-free training data of N_T samples, we run the estimator (20) and we construct the matrices

$$\mathcal{X} = \begin{bmatrix} X_1^T & \dots & X_{N_T}^T \end{bmatrix}^T, \quad \hat{\mathcal{F}} = \begin{bmatrix} |\hat{f}_1| & \dots & |\hat{f}_{N_T}| \end{bmatrix}^T. \quad (39)$$

Remark 8. To obtain a fault-free training dataset two approaches are possible. If a realistic simulator of the WF is available, the designer can excite it with all possible power demands and WS profiles and collect the output data. Alternatively, the designer can collect historical data covering all possible power demands and WS profiles. For newly-built systems, one can use either the corresponding simulator or the historical data of its initial operation by assuming that no faults affect the system during this initial period.

Then, we solve the following optimization problem

$$\underset{\beta, \epsilon > 0}{\text{minimize}} \quad \sigma (\hat{\mathcal{F}} - \mathcal{X} \beta)^T (\hat{\mathcal{F}} - \mathcal{X} \beta) + (1 - \sigma) \epsilon N_T^2 \quad (40)$$

$$\text{subject to} \quad |\hat{f}_k| < \mathcal{X}_k \beta < |\hat{f}_k| + \epsilon, \quad k = 1, \dots, N_T$$

with $\sigma \in [0, 1]$ being a weighting factor. The optimization problem (40) guarantees that the collected fault estimates do not exceed the threshold and it minimizes a weighted combination of the quadratic accumulated difference and the maximum difference (denoted as ϵ) between the threshold and the simulated data. Note that the multiplicand N_T^2 in (40) is used to regularize the minimized terms. The proposed FE and FDI strategies are summarized in Fig.5.

Remark 9. The sampling False Alarm Rate (FAR) obtained from (40) is equal to 0. If the resulting threshold is too conservative, one can fix $\text{FAR} = \frac{\phi}{T}$ by performing ϕ iterations in which the optimization problem (40) is solved and the most conservative data (i.e., $|\hat{f}_k|$ and X_k such that $\bar{k} = \text{argmin}_{k \in [1, N]} |\hat{f}_k| - X_k \beta$) is eliminated from \mathcal{F} and \mathcal{X} for the next iteration.

Remark 10. For brevity, we have omitted the dependence of the variables on the number of row i and on the number of column j . Hence, let us remark that the inequality in (36) stands in fact for $|\hat{f}_k^{i,j}| \geq J_k$. The threshold J does not depend on the number of WT (i, j) because the disturbances and noises affecting the estimation error $\tilde{f}^{i,j}$ are equally bounded for all the WTs in the WF.

6 FE and FDI at a Wind Farm Level

The performance of the observer (20) and the diagnoser (36) may be compromised if the bounds of the disturbances in Assumption 3 are too large. A solution to mitigate this effect is to totally decouple the FE error from these signals instead of attenuating their effect through the robust design in (35). However, a necessary condition for achieving disturbance decoupling is that the faults and the disturbances have a totally decoupled effect on the measurements and this is not possible from a WT level perspective because there is just one measurement which is simultaneously affected by the faults and the disturbances. At the WF level, we can consider altogether the WTs in the same row because the disturbance caused by the wind propagation error (i.e., the disturbance e^i) is common to all these WTs. As detailed hereafter, when all the WTs in a row are prone to the power fault, e^i does not verify disturbance decoupling conditions either. However, in this case, it is possible to build a bank of observers and diagnosers that allow to achieve the decoupling from e^i at the cost of some fault simultaneity restrictions (cf. [30]).

Remark 11. We have here considered that e^i is caused by the wind propagation error and that $e^{i,j}$ takes account of the uncertainties caused by the individual turbulence. If for any wind direction or for any other WF layouts there were no sufficient WTs per row in order to build the bank of observers, we would consider groups i of close WTs and divide the uncertainty caused by the wind propagation error into a common disturbance e^i and a non-common disturbance which would be considered together with the turbulence by $e^{i,j}$.

To do so, let us first define the auxiliary vectors

$$\mathbf{f}^l(i) = [f^{i,1} \dots f^{i,l-1} \quad f^{i,l+1} \dots f^{i,Z}]^T \in \mathbb{R}^{Z-1}, \quad (41a)$$

$$\boldsymbol{\delta}^l(i) = [\delta^{i,1} \dots \delta^{i,l-1} \quad \delta^{i,l+1} \dots \delta^{i,Z}]^T \in \mathbb{R}^{Z-1}, \quad (41b)$$

where Z denoted the number columns for the considered wind direction. The vector $\mathbf{f}^l(i)$ considers all the faults of the WTs in the i -th row but the fault of the WT in the l -th column $f^{i,l}$. Similar applies to $\boldsymbol{\delta}^l(i)$ w.r.t. the fault variations in (10). The model of the power generation systems of the i -th row of WTs including the dynamics of $\mathbf{f}^l(i)$ can be written as shown in (42) where $\mathbf{z}^l(i) \in \mathbb{R}^{2Z-1}$ is the extended state vector, $\mathbf{y}(i) \in \mathbb{R}^Z$ is the measurement vector and $\mathbf{w}(i) \in \mathbb{R}^Z$ and $\mathbf{s}(i) \in \mathbb{R}^Z$ are the disturbance vectors. For its part, $\boldsymbol{\theta}(i)$ stands for the parameter vector defined as

$$\boldsymbol{\theta}(i) = [\theta^{i,1} \dots \theta^{i,Z}]^T \in \mathbb{R}^{2Z}. \quad (43)$$

Assumption 4. The parameter vectors $\theta^{i,j}$ of all the turbines in the i -th row belong to the same subset $\Theta^{(p)}$ at each sample k (i.e., $\theta_k^{i,j} \in \Theta^{(p)}$ for $j = 1, \dots, Z$). Hence, $\boldsymbol{\theta}(i)$ lies in the ordered parameter set Θ of N subsets $\Theta^{(p)}$ (i.e., $\Theta = \{\Theta^{(1)}, \dots, \Theta^{(N)}\}$) defined as $\Theta^{(p)} = (\Theta^{(p)}, \dots, \Theta^{(p)})$.

Remark 12. Assumption 4 seeks reducing the computational burden of the design and implementation of a switched observer based on the model (42). However, in reality, it may in fact happen that not all the parameter vectors $\theta^{i,j}$ of the i -th row belong to the same subset $\Theta^{(p)}$ at certain sample k . In such cases, we neglect these differences and we consider that $\boldsymbol{\theta}_k(i)$ belongs to the closest subset $\Theta^{(p)}$. In any case, if more precise results were required, Assumption 4 would be omitted and the parameter set would be alternatively defined as $\Theta' = \{\Omega_v \times \Omega_\Delta \times \dots \times \Omega_v \times \Omega_\Delta\}$. The partition of Ω_Δ and Ω_v as detailed in Section 3.2 would then lead to N^Z subsets $\Theta^{(p)}$.

Omitting hereafter the dependence on the number of row i (e.g., \mathbf{z}_k^l stands for $\mathbf{z}_k^l(i)$), it yields

$$\mathbf{z}_{k+1}^l = \mathbf{A}(\boldsymbol{\theta}_k) \mathbf{z}_k^l + \mathbf{B}(\boldsymbol{\theta}_k) \mathbf{u}(\boldsymbol{\theta}_k) + \mathbf{E} \boldsymbol{\delta}_k^l, \quad (44a)$$

$$\mathbf{y}_k = \mathbf{C}^l(\boldsymbol{\theta}_k) \mathbf{z}_k^l + \mathbf{D}(\boldsymbol{\theta}_k) r_k + \mathbf{U}^l f_k^{i,l} + \mathbf{s}_k + \mathbf{w}_k, \quad (44b)$$

$$\mathbf{f}_k^l = \mathbf{F} \mathbf{z}_k^l, \quad (44c)$$

with $\mathbf{F} = \begin{bmatrix} 0 & \mathbf{I}_{Z-1} \end{bmatrix}$.

6.1 FE and FDI Architecture

In analogy to (20), the following model-based observer is built for the extended system (44):

$$\hat{\mathbf{z}}_{k+1}^l = \mathbf{A}(\hat{\boldsymbol{\theta}}_k) \hat{\mathbf{z}}_k^l + \mathbf{B}(\hat{\boldsymbol{\theta}}_k) \mathbf{u}(\hat{\boldsymbol{\theta}}_k) + \quad (45a)$$

$$\mathbf{L}^l(\hat{\boldsymbol{\theta}}_k) (\mathbf{y}_k - \mathbf{C}^l(\hat{\boldsymbol{\theta}}_k) \hat{\mathbf{z}}_k^l - \mathbf{D}(\hat{\boldsymbol{\theta}}_k) r_k),$$

$$\hat{\mathbf{f}}_k^l = \mathbf{F} \hat{\mathbf{z}}_k^l + \mathbf{K}^l(\hat{\boldsymbol{\theta}}_k) (\mathbf{y}_k - \mathbf{C}^l(\hat{\boldsymbol{\theta}}_k) \hat{\mathbf{z}}_k^l - \mathbf{D}(\hat{\boldsymbol{\theta}}_k) r_k). \quad (45b)$$

where $\mathbf{L}^l(\hat{\boldsymbol{\theta}}) \in \mathbb{R}^{(2Z-1) \times Z}$ and $\mathbf{K}^l(\hat{\boldsymbol{\theta}}) \in \mathbb{R}^{(Z-1) \times Z}$ are the gain matrices of the observer that we fix to

$$\mathbf{L}^l(\hat{\boldsymbol{\theta}}) = \mathbf{L}_p^l \quad \text{if } \hat{\boldsymbol{\theta}} \in \Theta^{(p)}, \quad (46a)$$

$$\mathbf{K}^l(\hat{\boldsymbol{\theta}}) = \mathbf{K}_p^l \quad \text{if } \hat{\boldsymbol{\theta}} \in \Theta^{(p)}. \quad (46b)$$

Define $\tilde{\mathbf{z}}_k^l = \mathbf{z}_k^l - \hat{\mathbf{z}}_k^l$ and $\tilde{\mathbf{f}}_k^l = \mathbf{f}_k^l - \hat{\mathbf{f}}_k^l$. It follows that

$$\tilde{\mathbf{z}}_{k+1}^l = \mathbf{A}^l(\hat{\boldsymbol{\theta}}_k) \tilde{\mathbf{z}}_k^l + \mathbf{B}^l(\hat{\boldsymbol{\theta}}_k) \mathcal{W}_k + \mathbf{E} \boldsymbol{\delta}_k^l - \mathbf{L}^l(\hat{\boldsymbol{\theta}}_k) \mathbf{U}^l f_k^{i,l}, \quad (47a)$$

$$\tilde{\mathbf{f}}_k^l = \mathbf{C}^l(\hat{\boldsymbol{\theta}}_k) \tilde{\mathbf{z}}_k^l + \mathbf{D}^l(\hat{\boldsymbol{\theta}}_k) \mathcal{W}_k - \mathbf{K}^l(\hat{\boldsymbol{\theta}}_k) \mathbf{U}^l f_k^{i,l}, \quad (47b)$$

$$z_{k+1}^l(i) = \underbrace{\begin{bmatrix} a(v_k^{i,1}) & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & a(v_k^{i,Z}) & 0 \\ 0 & 0 & 0 & I_{Z-1} \end{bmatrix}}_{A(\theta_k(i))} \underbrace{\begin{bmatrix} x_k^{i,1} \\ \vdots \\ x_k^{i,Z} \\ \mathbf{f}^l(i) \end{bmatrix}}_{z_k^l(i)} + \underbrace{\begin{bmatrix} b(v_k^{i,1}) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & b(v_k^{i,Z}) \\ 0 & 0 & 0 \end{bmatrix}}_{B(\theta_k(i))} \underbrace{\begin{bmatrix} u(v_k^{i,1}) \\ \vdots \\ u(v_k^{i,Z}) \end{bmatrix}}_{\mathbf{u}(\theta_k(i))} + \underbrace{\begin{bmatrix} 0 \\ I_{Z-1} \end{bmatrix}}_E \delta_k^l(i), \quad (42a)$$

$$\underbrace{\begin{bmatrix} y_k^{i,1} \\ \vdots \\ y_k^{i,Z} \end{bmatrix}}_{\mathbf{y}_k(i)} = \underbrace{\begin{bmatrix} c(\Delta_k^{i,1}) & 0 & 0 & -I_{l-1} & 0 \\ 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & c(\Delta_k^{i,Z}) & 0 & -I_{Z-l} \end{bmatrix}}_{C^l(\theta_k(i))} z_k^l(i) + \underbrace{\begin{bmatrix} d(\Delta_k^{i,1}) \\ \vdots \\ d(\Delta_k^{i,Z}) \end{bmatrix}}_{D(\theta_k(i))} r_k + \underbrace{\begin{bmatrix} 0_{(l-1) \times 1} \\ 1 \\ 0_{(Z-l) \times 1} \end{bmatrix}}_{U^l} J_k^{i,l} + \underbrace{\begin{bmatrix} s_k^{i,1} \\ \vdots \\ s_k^{i,Z} \end{bmatrix}}_{\mathbf{s}_k(i)} + \underbrace{\begin{bmatrix} w_k^{i,1} \\ \vdots \\ w_k^{i,Z} \end{bmatrix}}_{\mathbf{w}_k(i)}, \quad (42b)$$

with

$$\begin{aligned} \mathcal{A}^l(\hat{\theta}_k) &= A(\hat{\theta}_k) - L^l(\hat{\theta}_k) C^l(\hat{\theta}_k), \\ \mathcal{C}^l(\hat{\theta}_k) &= F - K^l(\hat{\theta}_k) C^l(\hat{\theta}_k), \\ \mathcal{B}^l(\hat{\theta}_k) &= [\mathcal{G}(\hat{\theta}_k) \quad G(\hat{\theta}_k) \quad -L^l(\hat{\theta}_k)H(\hat{\theta}_k) \quad -L^l(\hat{\theta}_k) \quad -L^l(\hat{\theta}_k)], \\ \mathcal{D}^l(\hat{\theta}_k) &= [0 \quad 0 \quad -K^l(\hat{\theta}_k)H(\hat{\theta}_k) \quad -K^l(\hat{\theta}_k) \quad -K^l(\hat{\theta}_k)], \\ \mathcal{W}_k &= [\lambda_k \quad \mu_k \quad \varphi_k \quad s_k \quad w_k]^T. \end{aligned}$$

We have that λ_k , μ_k and φ_k stand in fact for $\lambda_k(i)$, $\mu_k(i)$ and $\varphi_k(i)$ which are defined in analogy to (28) as

$$\lambda_k(i) = e_k^{i,1}/\bar{\varepsilon}_p, \quad \mu_k(i) = \begin{bmatrix} \varepsilon_k^{i,1}/\bar{\varepsilon}_p \\ \vdots \\ \varepsilon_k^{i,Z}/\bar{\varepsilon}_p \end{bmatrix}, \quad \varphi_k(i) = \begin{bmatrix} h_k^{i,1}/\bar{h}_p \\ \vdots \\ h_k^{i,Z}/\bar{h}_p \end{bmatrix},$$

if $\theta_k \in \Theta^{(p)}$. For its part, $\mathcal{G}(\hat{\theta}_k)$, $G(\hat{\theta}_k)$ and $H(\hat{\theta}_k)$ satisfy

$$\mathcal{G}(\hat{\theta}_k) = \mathcal{G}_p \quad \text{if } \hat{\theta}_k \in \Theta^{(p)}, \quad \mathcal{G}_p = \bar{R}_0 \bar{\varepsilon}_p, \quad (48a)$$

$$G(\hat{\theta}_k) = G_p \quad \text{if } \hat{\theta}_k \in \Theta^{(p)}, \quad G_p = \bar{R} \bar{\varepsilon}_p, \quad (48b)$$

$$H(\hat{\theta}_k) = H_p \quad \text{if } \hat{\theta}_k \in \Theta^{(p)}, \quad H_p = R \bar{h}_p, \quad (48c)$$

with $\bar{R}_0 = \begin{bmatrix} I_{Z \times 1} \\ 0_{(Z-1) \times 1} \end{bmatrix}$, $\bar{R} = \begin{bmatrix} I_Z \\ 0_{(Z-1) \times 1} \end{bmatrix}$ and $R = I_Z$. As previously specified, $\lambda_k(i)$ is a single disturbance modelling the wind propagation error and it affects all the turbines simultaneously.

In all, the signals affecting the l -th observer are the fault generators δ^l (with $n_\delta = Z - 1$), the disturbances in \mathcal{W} (with $n_\lambda = 1$ and $n_\mu = n_\varphi = n_s = n_w = Z$) and the fault $f^{i,l}$ which can be now considered as a new ‘‘disturbance’’. Following the approaches in [30], one can verify that it is now possible to decouple \mathbf{f}^l from λ if the new ‘‘disturbance’’ $f^{i,l}$ is ignored. Note that if $f^{i,l}$ was not ignored and its dynamics was considered together with the dynamics of \mathbf{f}^l , λ would not verify disturbance decoupling conditions because it would not be possible to distinguish this disturbance from the occurrence of simultaneous faults in all the turbines in the row.

Hence, we propose to ignore the presence of the fault $f^{i,l}$ and to design the observer (45) through the optimization problem

$$\begin{aligned} \text{minimize} \quad & \gamma_\mu n_\mu \|\mu\|_\infty^2 + \gamma_\varphi n_\varphi \|\varphi\|_\infty^2 + \gamma_s \|\mathbf{s}\|_{\text{RMS}}^2 + \\ & \gamma_w \|\mathbf{w}\|_{\text{RMS}}^2 \\ \text{subject to} \quad & \gamma_\lambda = 0, \quad \gamma_\delta \leq \bar{\gamma}_\delta, \\ & \Xi_p^m \succ 0, \quad \forall p, m \end{aligned} \quad (49)$$

with Ξ_p^m built in analogy to Ξ_p^m with the matrices in (47) replacing the matrices in (30). The new constraint $\gamma_\lambda = 0$ ensures now the decoupling from λ .

In the absence of the ignored fault $f^{i,l}$, we can thus set the following decisions ($j = 1, \dots, l - 1, l + 1, \dots, Z$):

$$\begin{cases} \text{if } |\hat{\mathbf{f}}_k^l[j]| \geq J_k & \text{Fault } \mathbf{f}^l[j] \\ \text{otherwise} & \text{No fault } \mathbf{f}^l[j] \end{cases}. \quad (50)$$

The adaptive threshold J does not depend on the number of turbine (i, j) nor on the number of ignored fault l (see Remark 10) and it is designed following the strategy in Section 5. The decoupling from λ enhances FDI because J in (50) is now smaller than J in (36).

However, if the fault $f^{i,l}$ is present in the system, the decision mechanisms (50) are no longer reliable. Hence, we build a bank of $l = 1, \dots, Z$ observers, each of them taking account of the fault \mathbf{f}^l and ignoring the fault $f^{i,l'}$. An observer l and the corresponding decision mechanisms are reliable if the absence of the fault $f^{i,l}$ is diagnosed by the decision mechanisms of at least another reliable observer l' ($l' \neq l$). In turn, the reliability of l' implies that the decision mechanisms of the l -th observer diagnose the absence of the fault $f^{i,l'}$. In all, the proposed bank of observer and decision mechanisms enables FE and FDI whenever 2 of the faults in a row i are not present in the system (i.e., there are no more than $Z - 2$ simultaneous faults in a row). FE and FDI are then achieved with any of the reliable observers and decision mechanisms of the bank.

7 Benefits and practical considerations of the proposed approach

Compared to the relevant existing literature, the benefits of the proposed approach are the following.

- As detailed in Section 2.2, the proposed approach utilizes a reduced number of measurements: the power reference, the generated power and the WS at the wind mast. This reduces the information needs w.r.t. other techniques such as the one in [6] (requiring the rotor speed measurement) or in [10] (requiring the collective pitch angle measurement). It does not either require the information about the presumed fault size as it is in [6].
- The residual-based techniques in [6–8] are focused on FDI tasks. Contrariwise, the proposed approach is focused on both FE (Section 4) and FDI (Section 5), being more suitable for the development of AFTC strategies.
- The proposed closed-loop approach is more actively robust against disturbances and uncertainties than the open-loop methods in [6–9].
- The Markovian jump system approach is a non-conservative procedure to handle the nonlinearities in the power generation model. It facilitates the adaptation to the different levels of uncertainties and disturbances along the parameter set, leading to a less restrictive compromise between fault sensitivity and robustness.
- The proposed observer design in Section 4.2 guarantees certain fault sensitivity with optimal disturbance rejection. This performance is not guaranteed in [6–9], where more ad-hoc and user knowledge-based tuning procedures are utilized to set this trade-off.

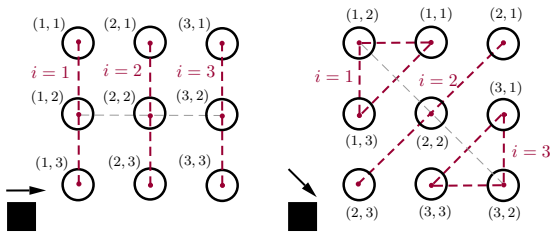


Fig. 6: Rows/groups of WTs in the WF benchmark depending on the wind direction (\rightarrow). \circ : Wind turbine, \blacksquare : Wind mast.

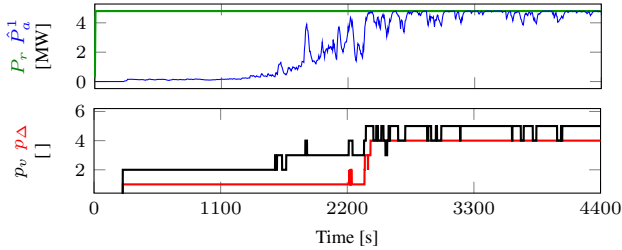


Fig. 7: Case A. Description and subsets. Turbine ($i = 1, j = 1$).

- In contrast to the constant thresholds in [6, 9, 10], the adaptive FDI mechanisms presented in Section 5 allow the adjustment of the thresholds to the different WS zones and WT operating modes. The proposed data-driven design of the FDI thresholds provides tight bounds for the fault estimates obtained via Markovian observers. This allows a more rapid detection and isolation of small faults w.r.t. model-based adaptive thresholds as the ones used in [8], which moreover require precise bounds of the uncertainties.

- The WF level scheme proposed in Section 6 is based on a systematic multi-input multi-output (MIMO) observer that automatically merges the information acquired from temporal and spatial inconsistencies depending on the level of shared uncertainties among the group of considered WTs. Hence, it is more easily extensible to different WF layouts and wind direction than the residuals in [7–10], which assume identical wind conditions among groups of WTs.

For its part, practitioners should pay attention to the following considerations when applying the proposed FE and FDI approach.

- The FE observer (20) is based on the model in Section 3. Although this observer is robust against uncertainties (see Section 4.2), a minimum fitting of the model parameters (e.g., τ_r) is necessary.
- Practitioners must take into account the variations of the uncertainties along the parameter set in order to suitably grid it. See Appendix 12.2 including an appropriate gridding for the WF benchmark.
- In practice, the transition probabilities among the subsets of the parameter set can be computed as detailed in Remark 3.
- The design of the FE observer (20) is based on the convex optimization problem (35), which can be addressed using semidefinite programming solvers (e.g., YALMIP [31]).
- The data-driven design of the FDI decision mechanism (36) requires a fault-free training dataset. See Remark 8 for practical details on the obtention of this dataset.
- The WT level approach is independent of the wind direction (WD); however, the WF level approach depends on the WD. Hence, different designs must be performed for different WDs. We propose to group these WDs (e.g., $[0-45]^\circ$ and $(45-90)^\circ$) and perform a design for each of these groups (considering the worst-case WD). The number of considered WD groups depends on the performance sought by the user. As the number of groups increases, the performance ameliorates at the cost of more offline design computations.

Remark 13. *The proposed FE and FDI strategy consists on a diagnostic tool which processes the available control measurements to diagnose power faults (Fig. 5). This diagnostic tool does not affect the system. Remedial actions based on the fault information provided by the tool will be considered in future works.*

Table 2 FE and FDI in the fault-free WT. Turbine ($i = 1, j = 1$).

Case	Case 1	Case 2	Case 3
P_r, P_a [MW]			
WT Level [MW]			
[pu]			
WF Level [MW]			
[pu]			

8 Simulation Results

In this section, the proposed strategies are numerically validated using the WF benchmark simulator [5]. The benchmark is used as a reference in the literature because it contains a realistic number of nonlinearities, uncertainties and disturbances. The simulator includes a WS profile of 4400 s that covers the whole WS range with challenging WS variations (detailed in Fig.10 of Appendix 12.1) and it considers the wind directions in Fig. 6.

The proposed WT level strategy is independent of the WF layout and of the wind direction. The WF level approach requires to group the nearby turbines and to bound the shared uncertainties caused by the wind (see Remark 11). In the first wind direction, we simply group the turbines by rows. In the second wind direction, we group the WTs as shown in Fig. 6. Due to space constraints, in the following, we just include the simulation results for the first wind direction. Similar WF level results are obtained for the other wind direction.

First, we create a fault-free training dataset by considering 10 different dynamic power reference cases (e.g., Case A in Fig.7). Note that it is not possible to consider different WS scenarios because only one WS profile is available in the benchmark; however, this WS profile is complex and considers a wide range of complex variations. The transitions experienced by the training dataset (e.g., Case A in Fig.7) are utilized for computing the transition probability matrix of the Markovian process. These probabilities are computed with the open-loop estimated power difference because the closed-loop estimates are not yet available. Then, we perform the design of the WT level observer (20) through the optimization problem (35) and of the WF level observer (45) through the optimization problem (49). For both designs, we fix the required \mathcal{H}_∞ performance describing the observer fault sensitivity to $\bar{\gamma}_\delta = 1$. The values included in the designs are specified in Table 3 and Table 4 of Appendix 12.2. The fault-free training dataset is subsequently utilized to obtain the adaptive threshold algorithms of the data-based decision mechanisms. For each training case, we run the designed estimator ((20) for the WT level and (45) for the WF level). Then, we construct the matrices (39) with all the estimated data and we solve the optimization problem. All the designs are set up in YALMIP. For simplification, we omit the numerical results.

Table 2 shows the simulation results for three validation cases (Cases 1-3 with 10^3 validation samples). The estimates are in black

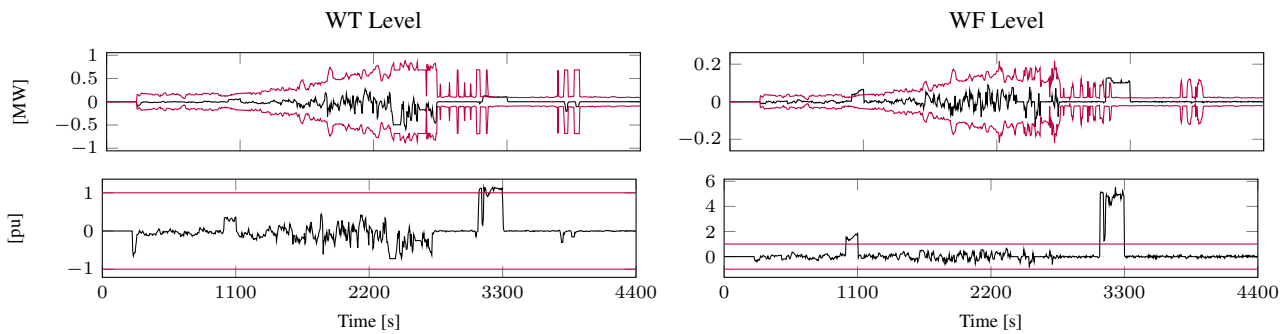


Fig. 8: FE and FDI in the WT ($i = 1, j = 1$). Case 1 with 40% power degradation during **T1** and 3% power degradation during **T2**.

and the thresholds are in red. For ease of comprehension, we include both the figures comparing the real estimates in MW with the real thresholds and the figures comparing normalized FE variables with unitary thresholds. First, we deduce that more accurate fault estimates are obtained when the WT is able to generate the amount of requested power. Effectively, in this WT working mode, the uncertainties derived from the WS estimation error disappear. Second, we verify that the WF level approach enhances fault diagnosability w.r.t. the WT level approach: the minimum diagnosable faults are ~ 0.2 MW at the WF level and ~ 0.8 MW at the WT level. Let us remark that the norm-based constant thresholds in Remark 6 are $\sim 10^2$ MW (WF level) and $\sim 10^3$ MW (WT level), being not applicable in practice without any further assumption on the statistical distributions of the disturbances.

Now, let us consider Case 1 and suppose that the turbine ($i = 1, j = 1$) is affected by a 40% power degradation during the period **T1** (defined as [1000, 1100] s) and by a 3% power degradation during the period **T2** (defined as [3200, 3300] s). The WT level and WF level FE and FDI results are depicted in Fig.8. The first fault is only diagnosed at the WF level and the second fault is diagnosed at both levels (although being more clearly distinguishable at the WF level). Effectively, relative power degradations are more difficult to diagnose during **T1** than **T2** because the amount of generated power is smaller and the level of uncertainties is bigger.

Finally, we consider the whole WF and the Case 2 in Table 2. We assume that the turbine ($i = 1, j = 3$) is affected by a relative power degradation of 50% during **T1** and that the turbine ($i = 2, j = 1$) is affected by a relative power degradation of 20% during **T2**. Fig.9 shows the WF level simulation results for all the WTs in the WF. These $4 \cdot 10^4$ samples prove that the proposed approach allows FE and FDI in the WF. Moreover, the proposed approach allows simultaneous faults in the farm: if the WT level approach is utilized, no simultaneity restrictions apply; if the WF level approach is utilized, the only restriction for achieving FE is that there should be two simultaneous fault-free turbines per row (i.e., only a faulty turbine per row in this case).

9 Conclusion

This paper proposed a novel closed-loop scheme for the estimation of decreased power generation in WTs due to blade erosion and debris build-up. The model-based strategy is based on the Markovian jump system model of the power generation system of a WT and it thus covers all the WS zones and WT operating modes. The FE output is also utilized for FDI in tighten data-based adaptive decision mechanisms. The approach is first applied to a single WT and then generalized from a WF perspective. The WF approach has the advantage of being easily applicable to different wind directions and layouts by simply adjusting the level of shared uncertainty by the considered groups of nearby turbines. The proposed strategies are tested in a realistic numerical simulator in the context of FD in WFs. Future work includes their application to a small-scale physical simulator. Also, the extension of the proposed approach to other WT systems and the use of the fault estimates in AFTC strategies highlight as immediate future work.

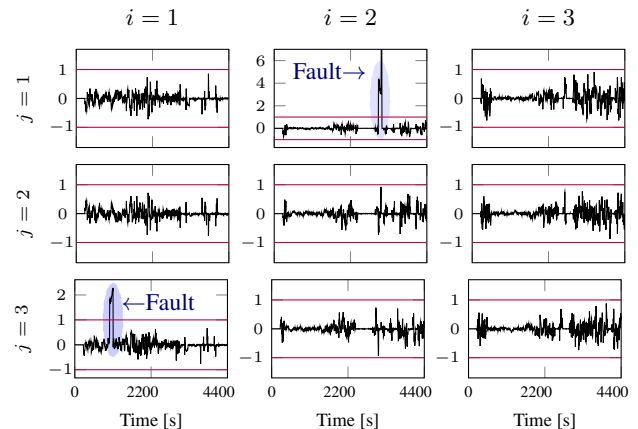


Fig. 9: WF level FE and FDI in the WF [pu]. Case 2 with 50% power degradation in the turbine ($i = 1, j = 3$) during **T1** and 20% power degradation in the turbine ($i = 2, j = 1$) during **T2**.

10 Acknowledgements

This work has been supported by the Spanish MECD (Grant FPU14/01592), the Spanish MINECO (Project TEC2015-69155-R) and the Universitat Jaume I (Project P11B2015-42).

11 References

- 1 R. C. McKenna, P. O. vd Leye, and W. Fichtner, "Key challenges and prospects for large wind turbines," *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 1212–1221, 2016.
- 2 N. Dalili, A. Edrissi, and R. Cariveau, "A review of surface engineering issues critical to wind turbine performance," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 2, pp. 428–438, 2009.
- 3 O. Parent and A. Ilinca, "Anti-icing and de-icing techniques for wind turbines: Critical review," *Cold Regions Science and Technology*, vol. 65, no. 1, pp. 88–96, 2011.
- 4 P. F. Odgaard, J. Stoustrup, and M. Kinnaert, "Fault-tolerant control of wind turbines: A benchmark model," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 4, pp. 1168–1182, 2013.
- 5 P. F. Odgaard and J. Stoustrup, "Fault tolerant wind farm control – A benchmark model," in *IEEE International Conference on Control Applications (CCA), 2013*, pp. 412–417, IEEE, 2013.
- 6 A. B. Borchers, J. A. Larsen, and J. Stoustrup, "Fault detection and load distribution for the wind farm challenge," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 4316–4321, 2014.
- 7 E. Duviella, L. Serir, and M. Sayed-Mouchaweh, "An evolving classification approach for fault diagnosis and prognosis of a wind farm," in *2nd Conference on Control and Fault-Tolerant Systems (SysTol), 2013*, pp. 377–382, IEEE, 2013.
- 8 J. Blesa, P. Jiménez, D. Rotondo, F. Nejari, and V. Puig, "An interval NLPV parity equations approach for fault detection and isolation of a wind farm," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3794–3805, 2015.
- 9 H. Badihi, Y. Zhang, and H. Hong, "Fault-tolerant cooperative control in an offshore wind farm using model-free and model-based fault detection and diagnosis approaches," *Applied Energy*, 2017.
- 10 S. Simani, P. Castaldi, and S. Farsoni, "Data-driven fault diagnosis of a wind farm benchmark model," *Energies*, vol. 10, no. 7, p. 866, 2017.
- 11 Y. Zhang and J. Jiang, "Bibliographical review on reconfigurable fault-tolerant control systems," *Annual Reviews in Control*, vol. 32, no. 2, pp. 229–252, 2008.

12 Z. Civelek, M. Lüy, E. Çam, and H. Mamur, "A new fuzzy logic proportional controller approach applied to individual pitch angle for wind turbine load mitigation," *Renewable Energy*, vol. 111, pp. 708–717, 2017.

13 G. Mavromatidis, K. Orehounig, and J. Carmeliet, "A review of uncertainty characterisation approaches for the optimal design of distributed energy systems," *Renewable and Sustainable Energy Reviews*, vol. 88, pp. 258–277, 2018.

14 A. Carpinone, M. Giorgio, R. Langella, and A. Testa, "Markov chain modeling for very-short-term wind power forecasting," *Electric Power Systems Research*, vol. 122, pp. 152–158, 2015.

15 Z. Lin, J. Liu, Q. Wu, and Y. Niu, "Mixed H_2/H_∞ pitch control of wind turbine with a markovian jump model," *International Journal of Control*, vol. 91, no. 1, pp. 156–169, 2018.

16 X.-k. Dai, Y. Song, M. Schüller, and D. Schramm, "Stability analysis of wind turbines combined with rechargeable batteries based on markov jump linear systems," in *Advanced Computational Methods in Energy, Power, Electric Vehicles, and Their Integration*, pp. 126–136, Springer, 2017.

17 Z. Lin, J. Liu, W. Zhang, and Y. Niu, "Regional pole placement of wind turbine generator system via a markovian approach," *IET Control Theory & Applications*, vol. 10, no. 15, pp. 1771–1781, 2016.

18 Z. Gao, X. Liu, and M. Z. Chen, "Unknown input observer-based robust fault estimation for systems corrupted by partially decoupled disturbances," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2537–2547, 2016.

19 P. Zhang and S. X. Ding, "An integrated trade-off design of observer based fault detection systems," *Automatica*, vol. 44, no. 7, pp. 1886–1894, 2008.

20 S. X. Ding, *Data-driven design of fault diagnosis and fault-tolerant control systems*. Springer Science & Business Media, 2014.

21 J. Saijai, S. X. Ding, A. Abdo, B. Shen, and W. Damlakhi, "Threshold computation for fault detection in linear discrete-time markov jump systems," *International Journal of Adaptive Control and Signal Processing*, vol. 28, no. 11, pp. 1106–1127, 2014.

22 E. Gonzalez, B. Stephen, D. Infield, and J. J. Melero, "Using high-frequency scada data for wind turbine performance monitoring: A sensitivity study," *Renewable Energy*, vol. 131, pp. 841–853, 2019.

23 A. Marvuglia and A. Messineo, "Monitoring of wind farms' power curves using machine learning techniques," *Applied Energy*, vol. 98, pp. 574–583, 2012.

24 K. Zhang, B. Jiang, V. Cocquemot, and H. Zhang, "A framework of robust fault estimation observer design for continuous-time/discrete-time systems," *Optimal Control Applications and Methods*, vol. 34, no. 4, pp. 442–457, 2013.

25 E. Sales-Setién and I. Peñarocha-Alós, "Trade-offs on fault estimation via proportional multiple-integral and multiple-resonant observers for discrete-time systems," *IET Control Theory & Applications*, vol. 13, no. 5, pp. 659–671, 2019.

26 R. Serfozo, *Basics of applied stochastic processes*. Springer Science & Business Media, 2009.

27 B. L. Nelson, *Stochastic modeling: analysis & simulation*. Courier Corporation, 2010.

28 K. Z. Østergaard, J. Stoustrup, and P. Brath, "Linear parameter varying control of wind turbines covering both partial load and full load conditions," *International Journal of Robust and Nonlinear Control*, vol. 19, no. 1, pp. 92–116, 2009.

29 I. Peñarocha, D. Dolz, and R. Sanchis, "Performance trade-offs for networked jump observer-based fault diagnosis," *IEEE Transactions on Signal Processing*, vol. 63, no. 10, pp. 2692–2703, 2015.

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

31 J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *IEEE International Symposium on Computer Aided Control Systems Design, 2004*, pp. 284–289, IEEE, 2004.

32 T. Knudsen, T. Bak, and M. Soltani, "Prediction models for wind speed at turbine locations in a wind farm," *Wind Energy*, vol. 14, no. 7, pp. 877–894, 2011.

12 Appendices

12.1 Wind Speed Estimation

From \hat{v}^0 , it is possible to estimate the mean WS acting on the WTs in the i -th row through different propagation models available in the literature, see [32] and the references therein. Provided the wind and wake model of the benchmark in [5], we compute the estimated mean WS using the following propagation model:

$$\hat{v}^i(t) = \frac{1}{K_i} \hat{v}^{i-1}(t - t_p^i(t)), \quad t_p^i(t) = \frac{l_i}{\hat{v}_f^0(t)/(K_i)^{i-1}},$$

where K_i is a wake factor and satisfies $K_i = 1$ for $i = 1$ and $K_i = 0.9$ for $i = 2, \dots, Y - 1$. The term $t_p^i(t)$ represents the propagation time between two consecutive rows and \hat{v}_f^0 results from filtering \hat{v}^0 using the low-pass filter $G_f(s) = 1/(s + 1)$.

The WF benchmark provides the real signals $v^{i,j}$ for numerical validations of the wind propagation schemes. Fig.10 shows the WS estimation results for the row $i = 1$ in Fig. 1. Recall that the estimates are common to all the turbines in a row i .

Table 3 Partition of the set Ω_v .

p_v	Boundaries	A_p^1	A_p^2	\bar{e}_p [m/s]	$\bar{\varepsilon}_p$ [m/s]
1	$0 \leq v < 4$	$0.973 \oplus -1$	$0.973 \oplus -1$	0.55	0.13
2	$4 \leq v < 7$	$0.973 \oplus +1$	$0.934 \oplus -1$	1.95	0.66
3	$7 \leq v < 11.25$	$0.934 \oplus -1$	$0.895 \oplus -1$	3.25	1.78
4	$11.25 \leq v < 12.5$	$0.895 \oplus -1$	$0.891 \oplus -1$	1.44	1.34
5	$12.5 \leq v < 25$	$0.891 \oplus +1$	$0.891 \oplus -1$	0.82	0.07
6	$25 \leq v < 30$	$0.973 \oplus -1$	$0.973 \oplus -1$	0.55	0.07

Table 4 Partition of the set Ω_Δ .

p_Δ	Boundaries	C_p	D_p	\bar{h}_p [MW]
1	$-4.8 \leq \Delta < -0.48$	[1 -1]	0	0
2	$-0.48 \leq \Delta < 0$	[1 -1]	0	0.48
3	$0 \leq \Delta < 0.48$	[0 -1]	1	0.48
4	$0.48 \leq \Delta < 4.8$	[0 -1]	1	0

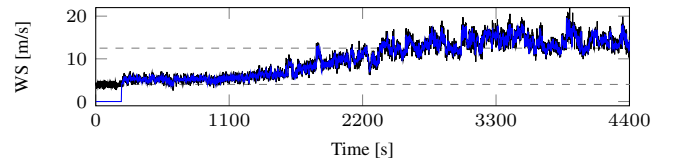


Fig. 10: Wind estimation for the WTs in the row $i = 1$ [m/s]. Black: Real ($j = 1$), Blue: Estimation. Dashed black: Wind zones.

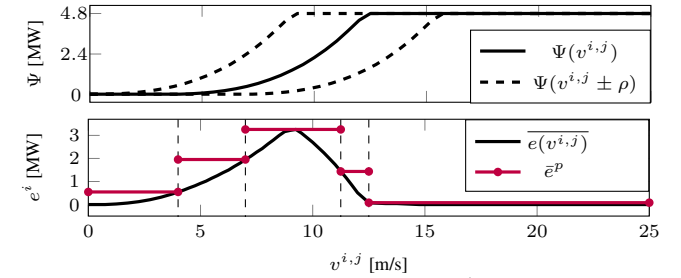


Fig. 11: Bounds \bar{e}^p of the disturbance e^i ($\rho = 4$ m/s).

12.2 Uncertainties and Partition of the Parameter Set

Let us first bound the WS estimation error. Since v^i is not available for numerical validations, we cannot separately compute the propagation error and the turbulence. Hence, we bound the turbulence error as $|\tilde{v}_t^{i,j}| \leq 3\sigma_t$ and we conservatively bound the wind propagation error as $|\tilde{v}_p^i| \leq \rho$ with ρ being the 99-th percentile of the total error $\tilde{v}^{i,j}$.

Then, we compute the bounds of the uncertainties e^i and $\varepsilon^{i,j}$, which are caused by these wind estimation errors. Define $\Psi(v^{i,j}) = a(v^{i,j})\bar{x} + \bar{u}(v^{i,j})$ with $\bar{x} = 4.8$ MW. The bounds of e^i and $\varepsilon^{i,j}$ for a certain WS $v^{i,j}$ verify

$$\overline{e^i(v^{i,j})} = \max \left\{ |\Psi(v^{i,j}) - \Psi(v^{i,j} \pm \rho)| \right\},$$

$$\overline{\varepsilon^{i,j}(v^{i,j})} = \max \left\{ \begin{array}{l} |\Psi(v^{i,j} + \rho) - \Psi(v^{i,j} + \rho \pm 3\sigma_t)| \\ |\Psi(v^{i,j} - \rho) - \Psi(v^{i,j} - \rho \pm 3\sigma_t)| \end{array} \right\}.$$

Fig.11 shows the results regarding e^i . Smaller bounds are obtained for $\varepsilon^{i,j}$ (not included here due to space constraints). Provided these results and given the form of $a(v^{i,j})$, we divide Ω_v into $N_v = 6$ subsets as shown in Table 3. The table includes the value of the matrices in (16) (B_p^m is deduced from A_p^m) and the bounds of the uncertainties related to the WS estimation in Fig.10.

Provided the definition of the uncertainties $h^{i,j}$ and given the binary functions $c(\Delta^{i,j})$ and $d(\Delta^{i,j})$, we divide the set Ω_Δ into $N_\Delta = 4$ subsets as detailed in Table 4. We have here considered that no WT operating mode estimation error appears above the 10% of the nominal power (i.e., 0.48 MW).