Annual Reviews in Control 47 (2019) 35–50

Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Review

Input design for active fault diagnosis

Tor Aksel N. Heirung, Ali Mesbah*

Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA

A R T I C L E I N F O

Article history:
Received 13 January 2019
Revised 5 March 2019
Accepted 8 March 2019
Available online 27 March 2019

Keywords:
Active fault diagnosis
Optimal input design
Model discrimination
Bayesian decision theory
System uncertainty

A B S T R A C T

Reliably diagnosing faults and malfunctions has become increasingly challenging in modern technical systems because of their growing complexity as well as increasingly stringent requirements on safety, availability, and high-performance operation. Traditional methods for fault detection and diagnosis rely on nominal input–output data, which can contain insufficient information to support reliable conclusions. Recent years have witnessed a growing interest in active fault diagnosis, which addresses this issue by injecting input signals specifically designed to reveal the fault status of the system. This paper provides an overview of state-of-the-art methods for input design for active fault diagnosis and discusses the primary considerations in the formulation and solution of the input-design problem. We also discuss the primary challenges and suggest avenues for future research in this rapidly evolving field.

Contents

* Corresponding author.
E-mail address: mesbah@berkeley.edu (A. Mesbah).

https://doi.org/10.1016/j.arcontrol.2019.03.002
1367-5788/

1. Introduction

Faults and malfunctions can happen in any modern technical system, with potentially detrimental effects on safety, performance, reliability, environmental footprint, and economics. In 2013, every Boeing 787 Dreamliner was grounded indefinitely after battery failures had occurred in two planes, with enormous consequences for the finances and reputations of the affected airlines, the manufacturer, and its suppliers (Williard, He, Hendricks, & Pecht, 2013). The 2005 series of explosions and fires at the BP refinery in Texas City, in part caused by an overflowing isomerization column, resulted in 15 fatalities and 180 injuries (Khan & Amyotte, 2007; Manca & Brambilla, 2012). Before losing control of El Al Flight 1862 in the 1992 accident, in which both engines on the starboard side detached because of material fatigue, the pilots were able to keep the plane in the air for almost fifteen minutes. Had the fault been diagnosed during this time, followed by appropriate action, the disaster could have been averted (Alwi, Edwards, Stroosma, & Mulder, 2008; Maciejowski & Jones, 2003). Reliable and timely diagnosis of faults is not only critical to safety, reliability, availability, and maintainability of a system, it is also essential in ensuring a system's ability to function as designed (Isermann, 2006). However, the growing complexity and strict performance requirements of modern technical systems have made reliable fault diagnosis increasingly challenging.

Fault diagnosis is generally a multi-step process, commonly including fault *detection, isolation, identification*, and *estimation*. Informally, these terms in turn refer to: determining whether or not the system is fault free; if not, which part of the system is faulty; the type of fault that has occurred in that part; and the magnitude of the fault (e.g., Blanke, Kinnaert, Lunze, & Staroswiecki, 2006). This paper deals with fault diagnosis in its entirety, rather than treating these activities individually. In particular, we focus on the problem of enhancing fault diagnosis through the design of system inputs, which is known as *active fault diagnosis*, or AFD. The remainder of this section gives an overview of some common types of faults, contrasts the active and passive approaches to fault diagnosis, discusses advantages of the active approach, highlights some connections to related branches of the control literature, and states the objective of the paper.

For clarity, we conform to the common practice of distinguishing between faults and failures, since these two terms are sometimes conflated in the literature. While a fault may cause a reduction in a system's ability to perform the tasks for which it is designed, a failure is generally understood as an event that renders the system inoperable. The two terms can thus be defined as follows (after Blanke et al., 2006; Isermann, 2006; Varga, 2017).

Definition 1 (Fault). A *fault* in a dynamic system is an anomalous variation in a characteristic system property that causes an unacceptable deviation from the specified limits of normal operation.

Definition 2 (Failure)**.** A *failure* is generally an irrecoverable event that renders the system incapable of operating such that it fulfills its purpose.

Hence, a failure is more critical than a fault, and a fault may lead to a failure unless diagnosed and managed appropriately.

Much of the literature makes a distinction between faults that arise from structural and gradual changes in the system. *Structural* changes are discrete events, such as actuators that are stuck in some position, the complete loss of a sensor, or a system component that breaks entirely. Faults arising from structural changes are often *abrupt*. Conversely, faults that stem form *gradual* changes can increase in severity or magnitude over time; examples include actuators that become slower to respond because of wear, sensor biases, and system components that suffer from issues like leaks or changing material characteristics. *Incipient* faults are in their earliest stages, primarily of the gradual type. Finally, a structural fault about to happen is *impending*.

1.1. Active versus passive approaches to fault diagnosis

The growing complexity of modern technical systems has made faults possibly more frequent and harder to diagnose. Generally, an important consideration in the design of technical systems is the potential occurrence of faults and failures to ensure some level of inherent robustness to such anomalies through the system design. For example, sensor and actuator redundancy can enable graceful degradation of system performance in the event of certain faults. Nonetheless, systematically accounting for all potential faults in the system design stage is impractical or impossible. This has motivated the use of fault diagnostics during operation, which are typically developed once the system is designed. However, a complex design, as well as feedback control and system uncertainties, can significantly limit the ability to diagnose faults (Sampath, Lafortune, & Teneketzis, 1998). Therefore, there has been a growing interest in the development of methods for faster and more reliable fault diagnosis during operation (Campbell & Nikoukhah, 2004; Zhang, 1989).

Fault diagnosis approaches are commonly classified as *active* or *passive*. The latter approach, also known as non-invasive, generally relies on comparing recorded input–output data to some reference data, which can be historical or generated through simulation. Importantly, the system is not perturbed to investigate its fault status. Comprehensive survey papers (e.g., Venkatasubramanian, Rengaswamy, Yin, & Kavuri, 2003, Venkatasubramanian, Rengaswamy, & Kavuri, 2003; Venkatasubramanian, Rengaswamy, Kavuri, & Yin, 2003) and a growing number of textbooks, such as Chen and Patton (1999), Chiang, Russell, and Braatz (2001), Blanke et al. (2006), Isermann (2006), Gonzalez, Qi, and Huang (2016), and Varga (2017), discuss passive methods in detail. Algorithms for passive fault diagnosis are broadly classified as data or model based (Venkatasubramanian, Rengaswamy, Yin, et al., 2003; Venkatasubramanian, Rengaswamy, & Kavuri, 2003; Venkatasubramanian, Rengaswamy, Kavuri, & Yin, 2003). Model-based methods generally require a model for every fault. These models are often based on first principles, but can also be identified from data (e.g., see Ljung, 1999). In contrast, data-based methods rely less on domain knowledge about the system, with a stronger focus on analysis of large historical data sets to characterize fault-free and different types of faulty operation.

A shortcoming of passive approaches arises from the potential lack of diagnostically relevant information in the input–output data generated while the system is operated under the assumption that no fault has occurred. We refer to this as *normal* operation, and say the system then generates *nominal* input–output data. Note that normal operation does not imply that no fault has occurred, and is thus distinct from *fault-free operation*.[1] That is, passive approaches do not account for the fact that nominal operating data may not be sufficiently informative for reliable fault diagnosis. Common reasons of this lack of diagnostically relevant information include system uncertainties and the presence of feedback controllers. Incomplete knowledge, or uncertainty, about the system and its state can result from inadequate measurements (including issues such as low signal-to-noise ratio, which lowers the information content in the measurements) and system disturbances that may not be readily distinguishable from faults through analysis of nominal operating data. Similarly, feedback controllers, the purpose of which

[1] Some authors use *nominal* as an antonym for faulty in the context of operation and models, a convention we do not follow here.

is partly to compensate for system variations, can mask the effects of faults. Thus, faults may remain undiagnosed for an extended period of time when nominal operating data is used for fault diagnosis. This situation can be particularly critical in the case of incipient faults that lead to failures unless diagnosed early.

Active fault diagnosis has primarily emerged to address the shortcomings of the passive approach of using nominal input–output data. Algorithms for AFD are generally model based; that is, a set of models is used to predict the behavior of different faults. The central idea in AFD is to design input signals that, when applied to the system, increase the amount of diagnostically relevant information in the input–output data. The designed input signals can be applied to the system when some performance-monitoring metric indicates abnormal operation (Qin, 1998; Severson, Chaiwatanodom, & Braatz, 2016; Zagrobelny, Ji, & Rawlings, 2013), or as part of a diagnostic routine, possibly periodically, to verify whether or not a fault has occurred. Such an input signal, commonly referred to as an *auxiliary* or *test* signal, is generally designed to ensure maximal or full separation between the model predictions corresponding to the different modes of operation. The fault hypotheses can then be discarded if their respective model predictions are in sufficient disagreement with new measurements. Note that the active approach to fault diagnosis is closely related to optimal experiment design for model discrimination (e.g., Atkinson & Cox, 1974; Mélykúti, August, Papachristodoulou, & El-Samad, 2010), where the design objective is to determine which model, among a set of candidates, best predicts the data. Similar to classic methods for model discrimination, the early AFD approaches do not rely on numerical optimization (Kerestecioğlu & Zarrop, 1991; Zhang, 1989). This is in contrast to more recent ones, which often pose the input-design problem as a dynamic optimization problem, as discussed in this paper.

1.2. Advantages of active fault diagnosis

The main motivation for AFD is to diagnose faults faster and more reliably relative to passive approaches. An active approach enables proactive investigation of the system's fault status while systematically accounting for uncertainty and operational constraints. Accounting for uncertainty is particularly important when dealing with faults that are small in magnitude or develop slowly. The effects of these faults may be indistinguishable from the variation arising from system uncertainty and disturbances; as a result they can become challenging to diagnose based on nominal input–output data. Injecting a test signal into the system can enable diagnosing small-magnitude faults before their severity increases. Another important feature of AFD is the ability to account for system constraints while injecting the test signal. Physical limits on capacity and actuation are common reasons for specifying constraints. Other reasons include ensuring that the test signal does not compromise operational safety or leads to unacceptable reduction in performance while investigating the fault status.

Active fault diagnosis fits into a larger context of automatic control systems that are tolerant to system uncertainties and undesirable events. As recognized by Åström (1991), "Fault diagnosis is an essential ingredient property of an intelligent system." In the context of fault detection, diagnosis, and tolerance, a brief comment on the use of the terms *passive* and *active* is in order. As described, active fault diagnosis involves manipulating the system input to actively investigate the fault status. The meaning of the word *active* in AFD is different from its meaning in the context of fault-tolerant control, or FTC. An active fault-tolerant controller responds to diagnosed faults through reconfiguration of the control strategy (Zhang & Jiang, 2008). That is, *active* refers to the controller's response to a diagnosed fault, as opposed to the diagnosis approach.[2]

1.3. Objective of the paper

Active fault diagnosis has matured significantly over the last two decades, with a diverse set of AFD approaches for a wide range of problems reported in the literature. The objective of this paper is to introduce researchers unfamiliar with the topic to the various directions in the literature, provide a systematic overview of available input-design methods for AFD, discuss important challenges and considerations involved in the formulation and solution of input-design problems, and present some opportunities and possible directions for future research.

This paper is organized as follows. Section 2 first provides some background on active fault diagnosis along with a discussion on relevant concepts. An overview of the first decade of the literature follows, along with a brief discussion of how the modern developments relate to these early results. Section 3 provides a comprehensive overview of the main contributions in the literature, centered around what we consider the three primary approaches to formulating input-design problems in AFD, along with some technical detail. We then discuss extension and variations on these three directions, followed by a discussion on results for nonlinear systems and Markov jump systems. We discuss some aspects of implementation of AFD methods in Section 4, with a focus on open- and closed-loop approaches as well as how AFD and control can be performed simultaneously. Section 5 concludes the paper with a brief discussion and some ideas for future research.

2. Input design for AFD

2.1. Background

A fundamental question in fault diagnosis is whether or not it is possible to diagnose a fault. This gives rise to the property *diagnosability*, the use of which varies across the literature, and there is no agreed-upon definition of diagnosability in the context of active fault diagnosis. Sampath, Sengupta, Lafortune, Sinnamohideen, and Teneketzis (1995), Sampath et al. (1998), Paoli and Lafortune (2005), and Bazille, Fabre, and Genest (2017) use definitions in the context of stochastic automata and finite state machines, and Dunia and Qin (1998) posit a set of conditions based on subspace identification; see also Saberi, Stoorvogel, Sannuti, and Niemann (2000). There are also variations in the literature in using diagnosability to refer to specific faults or to the system as a whole.

Here, we use the term as follows when referring to the diagnosability of a fault and a system. For a *fault* to be diagnosable, it needs to be possible to generate input–output data that is sufficiently informative for ascertaining, with a desired confidence in finite time and without violating any specified constraints, whether the fault has occurred. Note the relevance of constraints in this context: considerations such as safety limits may prohibit generating the required data. A typical example of generating data that reveals the occurrence of a fault is brake testing in a car. When driving at a constant velocity on a straight road there is no indication whether a car's brakes have failed. Tapping the brakes generates a response, which in turn generates data that can be analyzed to diagnose a brake fault. If all considered faults are diagnosable, we say the *system* is diagnosable.

[2] Note that a *passive* fault-tolerant controller is designed for robustness to faults; that is, the controller is capable of maintaining acceptable performance over a range of faults without reacting to the faults by modifying the control law (Zhang & Jiang, 2008).

When a system is diagnosable, fast and reliable fault diagnosis relies on two key factors: the information content of the input–output data and a *diagnostic algorithm* that analyzes the data to diagnose faults. If the data contains little or no information relevant to fault diagnosis, the diagnostic algorithm, no matter how well designed, may be unable to diagnose faults in a timely and reliable manner. Conversely, information-rich data may not be useful if the diagnostic algorithm is not adequate. Hence, the problem of active fault diagnosis consists of two components: (1) designing an input signal that, when applied to the system, can generate sufficiently informative input–output data for fault diagnosis, and (2) analyzing the data using a suitable diagnostic algorithm to infer the occurrence of faults. This is in contrast to passive fault diagnosis that generally relies on nominal input–output data, taking no active steps toward increasing the information content of data. Note that in general, these two components are not separate and their efficacies can be significantly interdependent. That is, a test signal that is optimal for one diagnostic algorithm may be unsuitable for another, and vice versa. Thus, active fault diagnosis can be generally defined as follows.

Definition 3 (Active fault diagnosis)**.** *Active fault diagnosis* consists of designing an appropriate test signal, injecting it into the system, and using the resulting input–output data for diagnosis of faults.

It follows from this definition that input design is the main aspect differentiating active and passive approaches to fault diagnosis, while the diagnostic algorithms used for passive and active approaches generally do not differ. Diagnostic algorithms have been thoroughly surveyed (e.g., see Blanke et al., 2006; Chiang et al., 2001; Isermann, 2006; Venkatasubramanian, Rengaswamy, Yin, et al., 2003; Venkatasubramanian, Rengaswamy, & Kavuri, 2003; Venkatasubramanian, Rengaswamy, Kavuri, & Yin, 2003) and discussing these is beyond the scope of this paper. Rather, we focus on the *input-design problem*.

Problem (Input design for fault diagnosis)**.** Given a model of the system, a set of fault models, and a diagnostic algorithm, design an input signal that results in data sufficiently informative for diagnosing a fault with some specified or maximized confidence, within a specified or minimal time, and without violating system constraints.

In the following, *diagnosis experiment* refers to the time period over which the input signal is applied to the system. Generally, input design for AFD has multiple objectives: (1) ensure diagnosis with an acceptable (typically prescribed) confidence, (2) ensure the diagnosis experiment is not excessively long, (3) prevent excessive operational disruption during the diagnosis experiment, and (4) ensure that the system remains within the specified operational constraints during the diagnosis experiment. Specifying an upper bound on, or minimizing the energy of, the test signal in the input-design problem is one way of ensuring its application does not result in excessive disruption. That is, the input signal can be designed to be minimally intrusive to operation during the diagnosis experiment. This is particularly important when there are risks associated with the experiment, or when there is a direct monetary cost associated with perturbing the system with an input signal designed to improve the diagnosis.

Modern input design for AFD generally involves formulating a dynamic optimization problem that consists of an objective function, one model for fault-free operation and one model for each fault, and commonly input and state constraints. The extent to which the system dynamics are nonlinear, the number of fault models, and how system uncertainties are accounted for all contribute to the complexity of an input-design problem. In the simplest case of two models, one for fault-free operation and one fault model, the problem often simplifies significantly; see, e.g., Campbell, Horton, and Nikoukhah (2002) and Blackmore and Williams (2005, 2006). The majority of research on active fault diagnosis considers linear systems. The two most common model types are data-driven input–output models affected by Gaussian white noise, such as ARMAX, and state-space models that are generally based on first principles. In recent years, active fault diagnosis for nonlinear systems, systems governed by differential-algebraic equations, and jump Markov processes are receiving increasing attention.

System uncertainty is generally described using (deterministic) bounded sets or probability distributions, although as discussed below there are input-design methods that rely on hybrid descriptions. Modeling uncertainty with bounded sets in an input-design problem enables prediction of the system outputs as bounded sets. When there is no overlap between any of these predicted output sets from the models, a system measurement can be reconciled with only one model, thus guaranteeing fault diagnosis. This approach to AFD is known as *set based*. Conversely, when describing uncertainties with probability distributions, the input-design objective generally involves separating the resulting predicted output distributions. If these distributions have infinite support, they necessarily overlap to some extent. In this case, it is not possible to guarantee diagnosis. The goal of input design is then typically either to minimize the probability of misdiagnosis or to ensure it is below a specified value. This is referred to as a *probabilistic* or *stochastic* approach to AFD. Note that when all distributions have finite supports, probabilistic AFD approaches enable guaranteed fault diagnosis through full separation of the predicted output probability distributions, similar to set-based approaches.

The remainder of this section gives an overview of the early developments in input design for active fault diagnosis. We outline the field's evolution from the early input-design methods based on linear data-driven models before transitioning to the following section, which focuses on modern approaches.

2.2. Early results on AFD for linear systems

The first AFD results for linear systems are presented in the paper by Zhang and Zarrop (1988) and in the monographs by Zhang (1989) and Kerestecioğlu (1993). These early contributions consider input design for improved fault diagnosis using two ARMAX models, one for fault-free operation and one for operation under a fault. These formulations rely on the assumption that the disturbances and measurement noise are Gaussian. Zhang and Zarrop (1988) introduce a measure for the statistical distance between the predicted output distributions and show that increasing the generated information, as measured by this distance, leads to faster fault diagnosis. The test signal is designed by maximizing this measure of diagnostic information subject to magnitude and bias constraints on the input. Notably, Zhang and Zarrop (1988) present an algorithm for multi-stage input design with implementation on a moving horizon. Zhang (1989) explores this framework more thoroughly, including both open- and closed-loop implementation of test signals designed to minimize diagnosis delay, which is the time between the onset of a fault and the diagnosis. Additionally, Zhang (1989) proposes a design criterion based on the Kullback–Leibler divergence (Kullback & Leibler, 1951), a measure of the similarity of two probability distributions. A large Kullback–Leibler divergence between two predicted output distributions indicates the test signal causes the model outputs to diverge, so that there is little overlap between the distributions. For fault diagnosis, this means there is a low probability of misdiagnosis. Kerestecioğlu and Zarrop (1991) compare offline and online input design for two ARMAX model hypotheses. The design objectives here are to decrease the expected diagnosis time and to ensure that the rate of false alarms is below a given threshold. They formulate an online input-design problem as finding a static feedback gain that achieves these goals. The paper shows there exists a trade-off between reducing the diagnosis delay and the rate of false alarms, and that the test signals designed offline have a simple structure: a sine wave with one single frequency. Later, Kerestecioğlu and Zarrop (1994) extend their approach to multiple fault models and constraints on the output variance, the input energy, and the rate of false alarms. In this more general case, the optimal signals generated offline are sine waves that contain a finite number of frequencies. Output-variance constraints are also considered by Uosaki, Tanaka, and Sugiyama (1984) in the related field of model discrimination. Instead of using ARMAX models, Uosaki, Takata, and Hatanaka (1993) pose the input-design problem in terms of increasing the Kullback–Leibler divergence for output distributions predicted using one ARX model for fault-free and one for faulty operation. While not improving the diagnosis rates or worsening the rates of false alarms, the designed test signals result in faster diagnosis on average. The authors later investigate a variant of this problem in the frequency domain (Hatanaka & Uosaki, 1994), as well as for the case of multiple models (Hatanaka & Uosaki, 1996; 1999a; 1999b) by maximizing the minimum Kullback–Leibler divergence for any pair of models.

While Bayesian statistics were used in passive fault diagnosis before the onset of research on AFD (e.g., see Chow & Willsky, 1984), most of the above approaches rely on statistical tests such as the cumulative-sum, or CUSUM, test or sequential probability-ratio test (SPRT). Kerestecioğlu and Zarrop (1989) introduce the Bayesian framework to active fault diagnosis and formulate input design as a sequential decision-making problem. They demonstrate the potential of the Bayesian formulation through analysis based on dynamic programming (Bellman, 1957).

All of the methods for input design discussed so far assume that each model hypothesis is an exact representation of the respective fault dynamics, with the stochastic process disturbances the only source of uncertainty. That is, the model parameters and the initial conditions are assumed known in each hypothesis. Kerestecioğlu and Çetin (1997a) develop an AFD approach for ARMAX models that allows the fault model to have parameters that change in a known direction but by an unknown magnitude with respect to the model for fault-free operation (in which the parameters are known and fixed). Separately, Kerestecioğlu and Çetin (1997b, 2004) consider the case where the direction of the parameter change in the fault model is also unknown. A main finding is that the designed test signals have simple structures. A single-frequency sinusoidal input is sufficient when the change direction is known. When the direction is unknown, the number of required frequencies is determined by the number of poles and zeros in the transfer function. Furthermore, the standard trade-off between fast diagnosis and the rate of false alarms exists in the case of a known change direction. However, when the direction of parameter change is unknown, the test signal can enable both faster diagnosis and a lower rate of false alarms. Note that in this approach, neither a probability distribution nor a closed set is assigned to the unknown parameters, offering a flexible framework for input design with a fault model that has a high level of parametric uncertainty.

The literature from this first decade of research on active fault diagnosis discussed above largely focuses on deriving closed-form expressions for the optimal input signals. Typical structures of these test signals include bang-bang (Zhang & Zarrop, 1988) and a sum of sinusoids (e.g., Kerestecioğlu & Çetin, 1997b). While these early AFD approaches rely on simple computational methods to determine optimal test signals for scalar systems, the underlying assumptions are somewhat restrictive and only a limited class of input-design problems lends itself to analytical solution. As discussed below, more general models and problem formulations enable posing a range of complex problems, with optimal inputs primarily designed through numerically solving dynamic optimization problems.

2.3. Toward modern AFD methods

A new paradigm, utilizing more modern tools and theory, started in the late 1990s, with Nikoukhah (1998) introducing a general formulation of the input-design problem based on state-space models and set-based uncertainty formulations. In this two-model approach, fault-free operation and one fault are modeled using two linear time-varying state-space models, with process disturbances and measurement noise described by deterministic, bounded sets. The formulation uses convex polytopes to formulate bounds on the input signal and on the uncertain disturbances and measurement noise. This deterministic problem formulation enables guaranteed fault diagnosis; that is, the designed test signal guarantees that the bounded sets of predicted model outputs do not intersect.

The use of state-space models has been prominent in the literature on active fault diagnosis since Nikoukhah's 1998 paper. Most of the subsequent development can be classified as one of three separate directions. One of these directions is the set-based approach described above, which has been extended in a series of papers through the introduction of a specific type of polytopes, *zonotopes*, to the input-design problem. A second direction considers disturbances and measurement noise that are sequences of random variables. As a consequence, this is a probabilistic framework and involves increasing the probability of correct diagnosis, rather than providing a guarantee. The third direction relies on bounding the energy of the uncertainty that enters the system and does permit guaranteed diagnosis. In the following section, we discuss these three directions in more detail. With that basis, we then discuss the different variations and extensions proposed in the literature.

3. Optimization-based formulations and solution methods

We here present three primary approaches to modeling the uncertainty within the framework of state-space models, along with corresponding formulations of the input-design problem. These three formulations span a broad range of problems in active fault diagnosis, and form a basis for the main directions in current research.

After presenting a fairly general state-space model, of which the three problem formulations use special cases, we introduce the problems in order of complexity. Arguably, the probabilistic framework introduced by Blackmore and Williams (2006) is the conceptually simplest, relies on the least involved mathematical formulation, and is computationally the cheapest. The set-based problem formulation introduced by Nikoukhah (1998) using polytopes, and later extended through the use of zonotopes in a series of papers, starting with Scott, Findeisen, Braatz, and Raimondo (2013), is mathematically more involved. The AFD problems introduced by Blackmore and Williams (2006) and Scott et al. (2013) both involve numerically solving discrete-time dynamic optimization problems that share many features with those commonly posed in model predictive control (see, e.g., Rawlings, Mayne, & Diehl, 2017). The final approach we discuss here, most comprehensively presented by Campbell and Nikoukhah (2004), is distinct from the two others in several ways, as demonstrated below. After presenting the basic forms of each of these problem types, we discuss variations, special cases, and extensions to each.

Table 1
Classification of some primary modern input-design methods. The contributions are categorized as considering linear or nonlinear systems, using probabilistic, set-based, or hybrid uncertainty descriptions, the number of many models $n_m = n_f + 1$ the method is developed for. A * indicates a method that allows state constraints.

System	Uncertainty	n_m	References
Linear	Probabilistic	2	Kim et al. (2013); Blackmore and Williams (2005)*
		≥2	Paulson et al. (2018); Blackmore and Williams (2006)*
	Set-based	2	Nikoukhah (1998); Andjelkovic and Campbell (2011)*
		≥2	Tabatabaeipour (2015); Scott et al. (2014)*
	Hybrid	2	–
		≥2	Scott et al. (2013)*; Marseglia et al. (2014)*
Nonlinear	Probabilistic	2	–
		≥2	Mesbah et al. (2014)*; Paulson et al. (2017)*; Martin-Casas and Mesbah (2018)*
	Set-based	2	Campbell et al. (2006); Andjelkovic et al. (2008)
		≥2	Campbell et al. (2002); Paulson et al. (2014)

We then provide an overview of AFD for nonlinear systems, followed by Markov jump systems.

Table 1 classifies some of the primary contributions discussed in this section. The table organizes the referenced papers according to system type (linear or nonlinear), uncertainty description (probabilistic, set-based, or hybrid), and the number of models considered.

The three approaches we discuss here all consider variations on a system with two or more linear discrete-time models of the form

$$x(k+1) = Ax(k) + Bu(k) + B_w w(k) + r \tag{1a}$$

$$y(k) = Cx(k) + Du(k) + D_v v(k) + s \tag{1b}$$

with time index k, state $x(k)$, input $u(k)$, disturbance $w(k)$, output $y(k)$, and measurement noise $v(k)$, all vectors with dimensions n_x, n_u, n_w, n_y, and n_v. Generally, all matrices, as well as the vectors r and s if included, implicitly depend on k in the sense that faults are modeled as changes in these quantities. Using $i = 0, 1, 2, \ldots, n_f$ as a model index, with n_f the number of fault models, we sometimes index the quantities in the formulation (1) (e.g., A_i) when explicitly distinguishing the models adds clarity. A model M_i is then uniquely defined by $(A_i, B_i, B_{w,i}, r_i, C_i, D_i, D_{v,i}, s_i)$. The fault-free model is indexed with 0 whereas $i \geq 1$ represents a fault model. Note that parts of the literature use *nominal* model instead of *fault free*; in this paper we only use *fault free* to distinguish from *faulty* to describe models in this context.

3.1. Probabilistic AFD

The probabilistic AFD problem for multiple fault hypotheses introduced by Blackmore and Williams (2006) considers a set of models of the form (1) with r and s both zero for all models and B_w and D_v identity matrices of appropriate dimensions. That is, the faults are modeled through changes in the matrices (A_i, B_i, C_i, D_i) as opposed to through the additive signals r and s.

The uncertain quantities in this formulation are the disturbances w, the measurement error v, and initial state $x(0)$, all specified with normal distributions. Specifically, $w(k) \sim \mathcal{N}(0, Q)$ and $v(k) \sim \mathcal{N}(0, R)$ are both sequences of independent and identically distributed random variables with Q and R known; the mean and covariance of $x(0)$ are also known, and $x(0)$, $w(k)$, and $v(j)$ are mutually independent for all $k, j \in \{0, 1, 2, \ldots\}$. Since the model parameters are assumed known, the predicted system states and outputs are Gaussian random variables, which makes it trivial to calculate their expected values. When the probability distributions have infinite support, as is the case for the normal distribution, it is not possible to define guaranteed diagnosis through eliminating any overlap between the predicted probability distributions. A natural problem is then designing an input sequence $\mathbf{u}$ that generates an output sequence $\mathbf{y}$, such that when analyzed the data minimizes

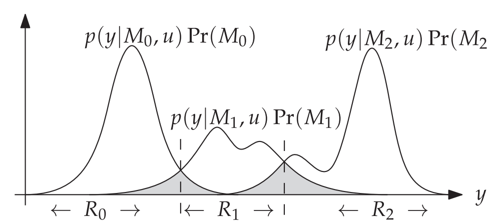

Fig. 1. Illustration of three regions used for hypothesis selection under the decision rule (2). The shaded areas indicate the probability of misdiagnosis Pr(error).

the probability of selecting the wrong model, or misdiagnosing the fault in other words. This problem can be formulated as outlined in the following.

Central to minimizing the probability of misdiagnosis is the Bayesian decision rule for model hypothesis selection, which minimizes the risk of misclassifying observations. That is, this rule minimizes the risk of selecting an incorrect fault model or hypothesis, given the a set of input–output data (Hellman & Raviv, 1970). Given the two sequences of inputs and outputs $\mathbf{u}$ and $\mathbf{y}$, the Bayesian decision rule can be expressed as

$$\text{select } M_{i^*} \text{ such that } i^* = \arg\max_i \Pr(M_i \,|\, \mathbf{y}, \mathbf{u}), \tag{2}$$

where $\Pr(M_i \,|\, \mathbf{y}, \mathbf{u}) = p(\mathbf{y} \,|\, M_i, \mathbf{u}) \Pr(M_i) / p(\mathbf{y} \,|\, \mathbf{u})$ by Bayes' theorem.[3] Here, the initial knowledge about the system's fault status is specified through the prior probabilities $\Pr(M_i)$. Under the decision rule (2), we can define a region R_i such that (2) selects hypothesis M_i when the system observations $\mathbf{y}$ fall into this region. Fig. 1 illustrates three such regions for densities of arbitrary form. We define these regions as

$$R_i : \left\{ \mathbf{y} : p(\mathbf{y} \,|\, M_i, \mathbf{u}) \Pr(M_i) > p(\mathbf{y} \,|\, M_j, \mathbf{u}) \Pr(M_j) \;\; \forall j \neq i \right\}.$$

With this definition, the decision rule (2) selects M_j when $\mathbf{y} \in R_j$, which is a selection error when M_i is the correct hypothesis. The probability of this specific misdiagnosis is $\Pr(\mathbf{y} \in R_j, M_i \,|\, \mathbf{u})$. The sum of this quantity over all pairs (i, j) is the probability of any misdiagnosis, $\Pr(\text{error})$, which is more generally known as the probability of hypothesis-selection error or the Bayes risk. Using Bayes' theorem, we have

$$\Pr(\text{error}) = \sum_{i=0}^{n_f} \sum_{j=i+1}^{n_f} \int_{R_j} p(\mathbf{y} \,|\, M_i, \mathbf{u}) \Pr(M_i) d\mathbf{y}. \tag{3}$$

The above discussion makes that it clear that reducing the size of the regions R_i lowers the risk of misdiagnosis. A primary challenge in this framework is evaluating $\Pr(\text{error})$. The multivariate

[3] The term $p(\mathbf{y} \,|\, \mathbf{u})$, known as the *evidence*, is the same for all models and we therefore omit it below in the expressions that compare $\Pr(M_i \,|\, \mathbf{y}, \mathbf{u})$ for different models M_i.

integrals involved in determining Pr(error) can be expensive to evaluate numerically, and the regions R_i may be difficult or impossible to determine, in particular for higher-dimensional space or arbitrary probability distributions. Various upper bounds on Pr(error) are therefore proposed to minimize the probability of misdiagnosis. Blackmore and Williams (2006) use the bound derived by Matusita (1971), formulated in terms of the Bhattacharyya coefficient (Bhattacharyya, 1943; Kailath, 1967), which is one measure of affinity between two distributions. This affinity measure can be difficult to evaluate for arbitrary distributions, but when the models are linear with additive Gaussian disturbances and noise, the bound can be expressed explicitly as a nonlinear function of the control-input sequence $\mathbf{u}$.

Minimizing the selection-error bound derived by Matusita (1971), evaluated over a finite horizon N of future inputs and outputs, subject to the state-space model candidates, results in a nonconvex nonlinear program, or NLP. The NLP has $N(n_\mathrm{f}+1)(n_x+n_u)$ variables in a straight-forward full-space implementation. There is a range of methods that can efficiently solve an NLP to local optimality, but finding the global minimum requires more computationally intensive algorithms (Tawarmalani & Sahinidis, 2002). Since the models are linear and the Gaussian disturbances and noise are the only model uncertainties, hard constraints the control inputs and expected values of the state can be included without increasing the problem complexity. In the case of only two models, the expression for the bound simplifies further and can be minimized through solving a concave quadratic program, or QP (Blackmore & Williams, 2005). However, like the multiple-fault case, this two-model problem requires sophisticated algorithms to determine the global optimizer, rather than a local solution.

The only source of nonlinearity in this NLP is the bound on Pr(error), which depends on the predicted mean state values and output covariances. While the error bound is a function of large covariance matrices that capture the output cross-correlation in time, these matrices are not functions of the decision variables. They can therefore be computed prior to solving the input-design problem.

3.2. A set-based formulation

Scott et al. (2013, 2014) develop a formulation that uses sets to model bounded disturbances and measurement noise. The primary difference between this framework and the one introduced by Nikoukhah (1998) is the use of zonotopes rather than standard convex polytopes to specify uncertainty bounds and the ability to consider more than one fault model. Furthermore, rather than framing the diagnosis problem directly in terms of models, Scott, Findeisen, Braatz, and Raimondo (2014) diagnose fault *scenarios*, defined as a *sequence* of models, with one model for each time index k on the horizon of interest. The frameworks of Nikoukhah (1998) and Scott et al. (2014) both involve determining a *separating input sequence*, which generates outputs that are guaranteed to be consistent with at most one fault model or scenario. In other words, the reachable output sets that correspond to the fault scenarios must be disjoint. This idea is illustrated using zonotopes for a three-model problem in Fig. 2, in which the output reachable sets are not fully separated.

Scott et al. (2014) use r and s in (1) to model additive faults, such as actuator and sensor biases. Unlike Blackmore and Williams (2006), who specify uncertainty in the disturbances $w(k)$, noise $v(k)$, and initial state $x(0)$ with normal distributions, Scott et al. assume that $x(0) \in X_0$ and $w(k) \in W$, $v(k) \in V$ for all k on the finite horizon of interest N, where X_0, W, and V are zonotopes. In contrast to general convex polytopes, which can be defined as the intersection of a set of half-spaces, zonotopes are centrally symmetric convex polytopes that can be described as Minkowski sums of line segments. In the generator representation used by Scott et al. (2014), an n-dimensional zonotope Z is defined by its center $c \in \mathbb{R}^n$ and a set of generators $g_1, g_2, \ldots, g_{n_g}$ as

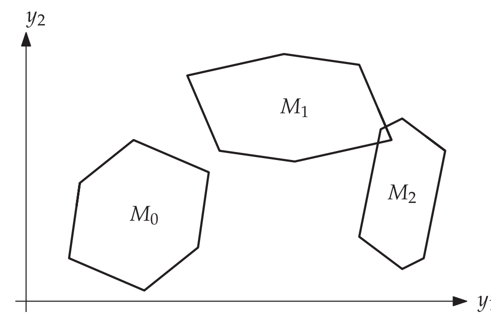

Fig. 2. Illustration of three zonotopes that correspond to output reachable sets for three models at some future time. The zonotopes for the fault models M_1 and M_2 have some overlap but the zonotope representing M_0 is fully separated.

$$Z = \{G\xi + c : \xi \in \mathbb{R}^{n_g},\ \|\xi\|_\infty \leq 1\}, \tag{4}$$

where $G = [g_1\ g_2\ \cdots\ g_{n_g}]$. The paper motivates the use zonotopes to specify uncertainty from a computational point of view. In the approach developed by Nikoukhah (1998), in which X_0, W, and V are specified as convex polytopes, characterizing the set of separating inputs involves determining the complement of a convex polytope, and this polytope is determined through projection. Scott et al. (2014) argue that this operation, an Nn_u-dimensional projection of a polytope of dimension $N(n_y+n_u)$, is computationally intractable for polytopes of dimension higher than about 10. By contrast, zonotopes enable efficiently and reliably computing the reachable sets and the set of separating inputs.

The input-design problem posed here is minimizing a quadratic cost function subject to a given set of fault scenarios formulated in terms of the linear state-space models of the form (1) and a set of constraints. This latter set includes constraints that ensure the input sequence is separating, linear input constraints, and constraints that specify the system states remain robustly within a given polytope for all considered scenarios. Scott et al. (2014) show that each of the separation constraints, one for every pair of fault scenarios, can be formulated as a linear program (LP). This permits formulating the input-design problem as a bilevel program, which can be reformulated into a single-level program by replacing the inner linear programs with their necessary and sufficient optimality conditions. Since the inner complementarity constraints for the Lagrange multipliers are nonconvex, they can be included as constraints in the outer program by introducing binary variables to the problem. These integer-valued variables render the design problem a mixed-integer quadratic program, or MIQP.

The computational complexity of an MIQP is primarily determined by the number of integer variables. In the proposed formulation, there are $2Qn_g$ binary variables, where Q is the number of scenarios and n_g the number of zonotope generators used to represent the separation constraint for each scenario pair, assuming the same number of generators are used in each. The authors suggest two approaches to reduce the number of binary variables. The first of these involves a systematic method for eliminating scenario pairs (reducing Q) as well as a discussion of how over-approximating the zonotopes significantly reduces the number of binary variables at the cost of a potential small increase in the optimal objective-function value. The other approach relies on observer-based diagnosis, which through a more conservative definition of a separating input reduces the required number of generators n_g by a factor of N. This more conservative requirement imposes only that the output of the set-based observers are separated at time N, which reduces the dimension of the sets that define separation from $(N+1)n_y$ to n_y. Removing N from the complexity consideration enables applying this method to problems that require a long horizon for guaranteed separation.

The MIQP formulation developed by Scott et al. (2014) allows some flexibility in the choice of objective. While the authors minimize the two-norm of the input sequence in the paper, they note that the separation condition can be used within a predictive control formulation while retaining the MIQP structure of the dynamic optimization problem.

3.3. AFD with energy-bounded uncertainty

The discrete-time two-model input-design problem posed by Campbell and Nikoukhah (2004) uses a state-space formulation that differs slightly from (1) in that the disturbances and noise are collected in one vector $\nu^\top(k) = [w^\top(k),\ v^\top(k)]$. Accordingly, the matrices $M_\nu = [B_w,\ 0]$ and $N_\nu = [0,\ D_v]$ are used instead of B_w and D_v, with the 0s representing zero-block matrices of appropriate dimensions, such that $M_\nu \nu(k)$ replaces $B_w w(k)$ and $N_\nu \nu(k)$ replaces $D_v v(k)$.[4] The vectors r and s do not appear in this formulation. Campbell and Nikoukhah demonstrate how parametric model uncertainty can be included through augmenting $\nu(k)$ with a signal that depends on the specified uncertainty. They also consider the slightly more general problem of using part of the input signal for control and the other part for active diagnosis, replacing $Bu(k)$ in the state-space model (1) with $Bv(k) + \bar{B}u(k)$ where $v(k)$ is the test signal. In this subsection, we do not consider simultaneous AFD and control and therefore disregard $\bar{B}u(k)$. To keep the notation consistent with Campbell and Nikoukhah (2004), we here use $v(k)$ to denote the test signal, as opposed to measurement noise.

Broadly, the framework involves bounding the uncertainty that arises from the unknown initial condition $x_i(0)$ and the additive signal $\nu_i(0)$ for both models $i = 0,\ 1$. The uncertainty in $x_i(0)$ is specified with respect to some known $\bar{x}_i(0)$ and the additive signal has bounded energy, rather than being bounded at every point in time. Together this forms the uncertainty-measure bound

$$S_i(\nu, k) = (x_i(0) - \bar{x}_i(0))^\top P_i^{-1}(0)(x_i(0) - \bar{x}_i(0)) + \sum_{j=0}^{k} \nu_i^\top(j) J_i \nu_i(j) < 1,\ \ \forall k \in [0,\ N-1] \tag{5}$$

with $P_i(0)$ a symmetric positive definite matrix used to specify the initial-state uncertainty and J_i a diagonal matrix with $+1$ and -1 entries that specify which uncertain components are bounded. When there is additive uncertainty only, J_i is the identity and $k = N-1$. This type of energy bound is motivated by cases when the uncertainty is primarily in the power density of the exogenous signals (Nikoukhah, Campbell, & Delebecque, 2000).

The uncertainty bound (5) is used to determine whether or not some given input–output data $\{v(j), y(j)\}_{j=0}^{N}$ can be reconciled with one or both of the models. That is, if the given data results in $S_i(v,\ k) < 1$ for both $i = 0$ and 1 it is inconclusive whether a fault has occurred. Conversely, if $S_i(v,\ k) < 1$ holds for only one model, the occurrence of a fault is unambiguous. Since the framework assumes that one of the two models describes the behavior of the system at any given time k, it is never the case that $S_i(v,\ k) \geq 1$ for $i = 0$ and 1 simultaneously. The input-design problem is thus to synthesize a signal v for which $S_i(v,\ k) < 1$ holds for either $i = 0$ or 1 but not both simultaneously. An input signal that accomplishes this separates the models and is called *proper*. Accordingly, v is proper if no set of noise, input, output, and state trajectories exists that satisfies the model equations (1) and the energy bound (5) for both $i = 0$ and 1 simultaneously. That is, the input signal v is proper if and only if there exists at least one k between 0 and $N-1$ for which $S_i(v,\ k) \geq 1$ for either $i = 0$ or 1. To formulate a condition on v in terms of the uncertainty measure $S_i(v,\ k)$, it is sufficient to check the largest value $\max(S_0(v,\ k), S_1(v,\ k))$. If there exists a set of feasible realizations of the signals $(\nu_0, \nu_1, u, y, x_0, x_1)$ that results in $\max(S_0(v,\ k), S_1(v,\ k)) < 1$, we can conclude that v is *not* proper. Conversely, if there exists a set of these signals that causes the largest of the uncertainty measures $S_i(v,\ k)$ to have a minimum value that is 1 or greater for some k, v is proper. This condition for v being proper can be formulated as

$$\sigma(v, k) \geq 1 \quad \text{for some } k \tag{6}$$

with

$$\sigma(v, k) = \inf_{\substack{\nu_0, \nu_1, u, y, \\ x_0, x_1}} \max(S_0(v, k),\ S_1(v, k)). \tag{7}$$

To develop an algorithm for synthesizing a proper test signal, the authors rely on rewriting the maximum of two numbers, $\max(c_1, c_2)$, as $\max_{0 \leq \beta \leq 1}(\beta c_1 + (1-\beta)c_2)$. They then show that in the right-hand side of (7), "inf max" is equivalent to "max inf." By defining the function

$$\phi_\beta(v, k) = \inf_{\substack{\nu_0, \nu_1, u, y, \\ x_0, x_1}} \beta S_0(v, k) + (1-\beta) S_1(v, k), \tag{8}$$

$\sigma(v,\ k)$ can be written as

$$\sigma(v, k) = \max_{0 \leq \beta \leq 1} \phi_\beta(v, k). \tag{9}$$

The resulting optimization problem, the solution to which is the minimum-energy proper test signal, is

$$\min_{v} \sum_{k=0}^{N} v_k^\top V_k v_k \quad \text{subject to} \quad \max_{\substack{0 \leq \beta \leq 1 \\ 0 \leq k \leq N}} \phi_\beta(v, k) \geq 1, \tag{10}$$

where V_k is a positive definite matrix. Note that this basic formulation of this approach does not include constraints on the system input, state, or output. Several of the extensions discussed below do, however, permit constraints.

Campbell and Nikoukhah (2004) propose a solution approach that involves constructing $\phi_\beta(v, k) = v_k^\top \Omega_{k,\beta} v_k$, with $\Omega_{k,\beta}$ a matrix constructed recursively from the matrices in the models (1) and uncertainty measure (5). The proposed algorithm then requires the solution of two eigenvalue problems. First, solve

$$\lambda = \max_{k,\beta} \text{"largest eigenvalue of } \Omega_{k,\beta} V_k^{-1}\text{"} \tag{11}$$

and denote the maximizer by $(k^*,\ \beta^*)$, with k^* the optimal length of the test signal. With λ and this maximizer, determine the optimal test signal v_{k^*} by solving the eigenvalue problem

$$(\lambda v_{k^*} - \Omega_{k^*,\beta^*}) v_{k^*} = 0. \tag{12}$$

The solution algorithm outlined here is the most straightforward one among the many Campbell and Nikoukhah (2004) propose for a range of variations on this problem. They also thoroughly discuss the continuous-time version of the problem.

3.4. Linear systems: Variants, extensions, and special cases

3.4.1. Probabilistic methods

Kim, Raimondo, and Braatz (2013) study a two-model AFD problem similar to the one considered by Blackmore and Williams (2005). Instead of minimizing the probability of misdiagnosis, however, Kim et al. use the Kullback–Leibler divergence to quantify the distance between the predicted output distributions. Maximizing the Kullback–Leibler divergence leads to a nonconvex optimization problem that can become computationally prohibitive. Thus, as a tractable alternative, the input-design problem

[4] Note that the matrix M_ν has no relation to the model M_i.

is formulated in terms of maximizing the geometric distance between the output distributions. The latter problem has a concave objective function and a convex constraint set, and its solution can be determined by solving a series of semidefinite programming (SDP) problems.

Paulson, Heirung, Braatz, and Mesbah (2018) consider an input-design problem for multiple linear state-space models with Gaussian disturbances and noise, similar to Blackmore and Williams (2006). However, the resulting test signal maximizes the pairwise sum of the Bhattacharyya distances and the formulation incorporates input constraints only. The constraint set forms a polytopic convex feasible area, which implies that the global optimum is at one of its vertices. The solution can therefore be found efficiently through full enumeration of a moderate number of vertices. A major advantage of this formulation is that the number of vertices is independent of the number of fault models n_f and the dimension of the system state n_x, and depends only on the horizon length N and input dimension n_u. Including state constraints in this approach increases its complexity but the feasible area remains a convex polytope.

3.4.2. Hybrid probabilistic-deterministic formulations

Scott, Marseglia, Magni, Braatz, and Raimondo (2013) argue that, in practice, probabilistic formulations for input design often lead to acceptable confidence in fault diagnosis with test signals that are less aggressive than those designed by deterministic approaches. They propose a hybrid framework for input design based on disturbances and measurement noise that have known finite-support uniform probability distributions, with the supports described by zonotopes. This approach allows defining a two-fold diagnosis objective: a high probability of diagnosis at the end of some specified short time horizon in addition to guaranteed diagnosis after the full horizon N. Later, Marseglia, Scott, Magni, Braatz, and Raimondo (2014) relax the assumption of uniform probability. By applying a scenario approach to solve the input-design problem, they develop a method that allows arbitrary probability distributions on finite supports and derive a guaranteed lower bound on the probability of diagnosis that depends on the number of samples.

Hatanaka and Uosaki (2000) also present a hybrid approach to input design, which extends their earlier results to the case of bounded parameter uncertainties combined with Gaussian white noise, resulting in a mixed stochastic-deterministic formulation. Unlike the approaches discussed in the previous paragraph (Marseglia et al., 2014; Scott et al., 2013), Hatanaka and Uosaki consider ARX models, rather than state-space models.

3.4.3. Set-based methods

Despite their computationally attractive features, standard zonotopes have several shortcomings that limit their applicability. Two important shortcomings are that zonotopes are not closed under intersection and that they are symmetric. Hence, they cannot accurately represent strongly asymmetric sets, such as those that result from the intersection of two centrally symmetric sets. Scott, Raimondo, Marseglia, and Braatz (2016) address these issues by introducing *constrained* zonotopes. Constrained zonotopes are not necessarily centrally symmetric, and are thus more flexible than their standard counterpart. As a result, they enable the computation of tight enclosures at moderate computational cost and offer a simple mechanism for trading off accuracy with computational efficiency. Scott et al. (2016) show that their application in fault diagnosis can lead to faster diagnosis. To further reduce the computational complexity of the approach presented in Scott et al. (2014), Scott et al. (2016) also introduce a zonotope order-reduction technique. Yang and Scott (2018) compare this technique to others available in the literature. Despite the potential of these order-reduction techniques in lowering the computational burden, the cost of solving an MIQP can be prohibitive for online application. Marseglia and Raimondo (2017) propose using multi-parametric programming (Dua, Bozinis, & Pistikopoulos, 2002) to move most of the computational cost offline. While multi-parametric programming also involves expensive computations, in particular for a large number of fault models, the authors address this challenge by considering only two models at the time. They show through simulations that this simplification result in AFD performance that is comparable to considering all models simultaneously.

Tabatabaeipour (2015) introduces a different different set-based approach, relying on both zonotopes and conventional polytopes. Zonotopes are here used for computational efficiency in the set operations that need not be exact, whereas regular polytopes are used to avoid approximations in the part of the algorithm that falsifies model hypotheses from data.

3.4.4. Formulations with energy bounds

In the AFD approaches developed by Campbell, Nikoukhah, and coworkers, the input-design problem is typically posed as a dynamic optimization problem (Campbell & Nikoukhah, 2004). In contrast to the discrete-time formulation discussed above, the majority of results in the bounded-energy framework apply to continuous-time models. While there is some variation between the different versions of the problem they consider, their formulations generally take the form of a bi-level optimization problem. The fundamental idea is that the inner problem ensures the test signal is proper, which guarantees separation of the predicted output sets, while the outer problem minimizes the energy of the proper input. In the continuous-time version, this results in a two-point boundary-value problem as a necessary condition, and the authors discuss the development of a specialized solution algorithm and its software implementation. The special case of two continuous-time models and finite diagnosis time involves the solution of Riccati equations (Nikoukhah, Campbell, Horton, & Delebecque, 2002). Campbell et al. (2002) explore other optimization-based formulations for the case of an arbitrary number of models, including Riccati theory, the calculus of variations, directly minimizing the energy of the proper input signal, and Euler–Lagrange theory. These continuous-time input-design problems are formulated for implementation in a commercially available solver for dynamic optimization problems.

Nikoukhah et al. (2000) address the design of a minimum-energy proper test signal for finite time horizons with discrete-time models, as presented in Section 3.3. This paper also investigates asymptotic behavior and shows that the optimal test signal converges to pure sinusoids as the diagnosis horizon goes to infinity. Nikoukhah, Campbell, and Delebecque (2001) extend this finite-time input-design framework to include parametric uncertainty in the state-space models, represented by bounded perturbations to the model matrices. The framework is also extended by Campbell et al. (2002) to handle more than two models and certain nonlinearities, but without uncertain model parameters. A different extension, developed by Nikoukhah and Campbell (2003), generalizes the input-design problem to design objectives other than minimizing the energy of the auxiliary signal. The proposed quadratic formulation of the cost function enables more detailed specification of desired system behavior during testing. Campbell and Nikoukhah (2004) comprehensively discuss all of these problem variations.

The input-design problem for incipient faults is investigated by Nikoukhah and Campbell (2006b) through representing the fault by drift in the A, B, C, and D matrices in the model (1). Subsequently, Nikoukhah and Campbell (2008) extend this framework to account for multiplicative model uncertainty. The more general case of simultaneous incipient faults is studied by Fair and Campbell (2009b) for two faults and by Fair and Campbell (2009a) for more than two faults; see also Nikoukhah, Campbell, and Drake (2010).

Nikoukhah and Campbell (2006a) and Ashari, Nikoukhah, and Campbell (2011) consider the problem of incorporating information about the initial system state through bounding the total uncertainty from the initial state and the disturbances and noise. This approach enables accounting for known input signals to the system (for example control signals) and handling a broader class of faults, including those modeled by a bias. The uncertainty on the initial condition and the additive signals are treated separately in the formulation developed by Campbell and Scott (2016), with both box-type and ball-type bounds on the initial condition. Andjelkovic and Campbell (2011) consider the two-model problem with bounded multiplicative parametric model uncertainty and constraints. This framework can handle constraints on the input as well as on the system state. The approaches to input design by Nikoukhah, Campbell, and coworkers discussed above all result in a continuously-varying test signal. Implementing such a signal is in many cases not practical. Choe, Campbell, and Nikoukhah (2009) address this issue through the design of test signals that are piecewise constant, where the length of each constant input can be different from the measurement interval.

Blanchini et al. (2017) propose a computationally efficient method for guaranteed diagnosis in linear continuous-time systems with multiple fault models. The method accounts for disturbances and measurement noise with known bounds, and synthesizes a test signal that is either constant or sinusoidal. Their approach allows offline computation of the optimal inputs through solving convex optimization problems.

The approach developed by Andjelkovic and Campbell (2011) is extended by Scott and Campbell (2014) to handle systems governed by a linear differential-algebraic equation (DEA) set, potentially of high index. The introduction of algebraic state variables poses several challenges, including that an optimal test signal may not exist. The authors show how in this case a modification to their proposed algorithm enables the synthesis of a near-optimal signal.

3.5. AFD for nonlinear systems

Active fault diagnosis for nonlinear systems has received increasing attention in the last decade. However, the literature on input design for nonlinear systems is still limited. Potential reasons include the computational challenges associated with uncertainty propagation through nonlinear models as well as the increased complexity of the associated optimization problem. Campbell et al. (2002) are among the first to address AFD for nonlinear systems. Their proposed algorithm is applicable to a specific class of nonlinear systems: small nonlinearities in the state, meaning the norm of the nonlinear function is bounded by a small number for all time in the entire state space; and control inputs that enter the state equation through a nonlinear function. This method assumes no knowledge of the initial state, allows more than two models, and uses minimum energy of the auxiliary signal as the design criterion. Campbell, Drake, Andjelkovic, Sweetingham, and Choe (2006) extend the input-design framework of Campbell and Nikoukhah (2004) to nonlinear systems using linearization, and propose an algorithm for evaluating the performance of the designed test signal. Andjelkovic, Sweetingham, and Campbell (2008) use separate approaches to deal with small and large nonlinearities. For smaller nonlinearities, they show that linearization leads to acceptable input design. The paper analyzes theoretically the consequences of linearization and establishes conditions under which the input signal resulting from linearization meets the specified design goals. With larger and more general nonlinearities, the authors directly formulate the input-design problem in terms of the nonlinear models and solve the optimization problem with a direct-transcription approach.

Paulson, Raimondo, Findeisen, Braatz, and Streif (2014) investigate input design for guaranteed fault diagnosis in nonlinear systems. They consider nonlinearities that are either polynomial or rational and assume that model parameters, noise, and disturbances are all unknown but bounded. The authors derive a bilevel optimization problem, in which the inner problem is convex for a given input signal and ensures the system output can be reconciled with exactly one model; the outer problem minimizes the norm of the input signal. The separating input is the solution to a convex relaxation of this problem. Mesbah, Streif, Findeisen, and Braatz (2014) propose a method for input design for general nonlinear systems in which the model parameters and initial conditions are unknown but have known probability distributions. The proposed method does not consider disturbances and measurement noise. The paper formulates a computationally tractable input-design problem through approximating the probabilistic model uncertainties by truncated polynomial chaos expansions (Wiener, 1938; Xiu & Karniadakis, 2002). Polynomial chaos allows handling non-Gaussian distributions, including ones with finite support. The formulation proposed by Mesbah et al. (2014) also incorporates constraints on the nominal state trajectories and hard constraints on the inputs. The objective for input design is maximizing the sum of the Hellinger distances between every pair of model output predictions. In the NLP solved to determine the optimal input, the coefficients in the polynomial chaos expansions are determined through repeated simulation of the models. Paulson, Martin-Casas, and Mesbah (2017) propose an input-design method that in addition to probabilistic model uncertainties in parameters and initial conditions also permits stochastic disturbances and noise. The paper proposes two methods for propagating the probabilistic uncertainties through the nonlinear models, one involving linearization along predicted model trajectories and another using the unscented transform. Both methods represent the output predictions through mean and covariance approximations, which are used to evaluate an approximate bound on the probability of misdiagnosis, adapted from Blackmore and Williams (2006). Constraints on the state are formulated as joint chance constraints, evaluated through a moment-based approximation. This results in an input-design problem formulated as an NLP that can be solved with standard techniques and software. Martin-Casas and Mesbah (2018) also consider nonlinear systems with parametric uncertainty and stochastic disturbances and noise. They develop a sample-based distance measure, similar to the k-nearest neighbors algorithm, to separate the predicted output distributions, which are propagated using generalized polynomial chaos.

3.6. Markov jump systems and AFD

In the input-design methods discussed above, transition probabilities between modes of operation are not specified. This type of information, such as the probability of a given fault occurring when the system is fault free, or the probability of one specific fault occurring right after another, may not be available in practice. When it is available, however, it may significantly benefit active fault diagnosis. Alternatively, the probabilities can be used to assign priority to the faults considered most critical. A system with known probability transitions between different modes of behavior, each with separate models, is called a *Markov jump system*. A key challenge in input design for this type of systems is that the number of possible fault-mode sequences grows exponentially with time, rendering the problem intractable without a strategy to limit the number of sequences considered.

Table 2
Some primary input-design methods classified as either open- and closed-loop implementation and according to the type of optimization problem solved.

Implementation	Optimization problem	References
Open loop	LP	Nikoukhah (1998); Scola, Nikoukhah, and Delebecque (2003)
	QP	Blackmore and Williams (2005)
	MIQP	Scott et al. (2014)
	SDP	Kim, Shen, Nagy, and Braatz (2013)
	NLP	Blackmore and Williams (2006)
	Other	Blanchini et al. (2017); Campbell and Nikoukhah (2004)
Closed loop	MIQP	Marseglia et al. (2017); Raimondo et al. (2013)
	NLP	Martin-Casas and Mesbah (2018); Paulson et al. (2017)
	Other	Paulson et al. (2018)

Šimandl, Punčochář, and Královec (2005) develop an input-design framework for Markov jump linear systems with Gaussian noise and disturbances. This formulation allows unknown model parameters and initial state but assumes these quantities have known probability distributions. They formulate the problem using dynamic programming and propose a computationally tractable approach for obtaining an approximate solution. Blackmore, Rajamanoharan, and Williams (2008) propose an input-design method for Markov jump linear systems with Gaussian distributions for the initial state, disturbances, and noise, but with no parametric uncertainty. They extend their approach for multiple models (Blackmore & Williams, 2006) to the case of specified transition probabilities. To limit the number of sequences considered in designing the input, the proposed algorithm includes a feature for pruning the tree of possible mode switches. The formulations proposed by Šimandl et al. (2005) and Blackmore and Williams (2006) consider finite time horizons. For the case of infinite horizons, Punčochář, Škach, and Šimandl (2015b) solve the input-design problem using approximate dynamic programming, while Škach, Punčochář, and Lewis (2016) propose the use of temporal-difference learning.

Punčochář, Král, and Šimandl (2009) consider Markov jump nonlinear systems with a single input and a single output, whereas Punčochář and Šimandl (2014) and Punčochář, Škach, and Šimandl (2015a) tackle the more general case of multiple inputs and outputs. None of these approaches make assumptions on the system uncertainties except for the probability distributions being known. Punčochář et al. (2009) propose using neural nets to model the fault modes. In their more recent work, Punčochář and coworkers consider both the case of perfect and imperfect state information, and use approximate dynamic programming to solve the input-design problem (Punčochář & Šimandl, 2014; Punčochář et al., 2015a). Škach and Punčochář (2017) address the special case of Gaussian disturbances and measurement noise and use reinforcement learning as a solution approach, a technique closely related to the ones used by Punčochář and Šimandl (2014) and Punčochář et al. (2015a). Škach, Punčochář, and Straka (2017) and Škach, Straka, and Punčochář (2017) extend this framework by not restricting the stochastic sequences to a specific type of distribution.

4. Implementation of input design for AFD

This section provides a discussion on two implementation considerations in active fault diagnosis. Of these, the most important is whether to design the inputs in an open- or closed-loop fashion, which is related to the issue of online versus offline input design. Table 2 categorizes some of the primary methods according to whether they implement the input signal in closed or open loop and notes the type of optimization problem solved to determine the signal. Our discussion then focuses on AFD integrated with feedback control. The considerations involved in the question of whether to combine control and diagnosis objectives are largely outside the scope of this paper. We therefore keep this discussion brief, despite the considerable body of literature addressing this topic.

4.1. Open- and closed-loop input design

Input design for active fault diagnosis can be implemented as an open-loop design problem, or as a closed-loop problem based on updating the test signal applied to the system whenever new information becomes available. Much of the literature on open-loop AFD considers offline design of an input signal that has no online dependence on the measurements. Nikoukhah et al. (2000) argue this is the preferable approach to AFD since a test signal that depends on the most recent measurement establishes a feedback that modifies the dynamics of the system, potentially causing instability. Note that online input design does not mean there is a closed loop, in which every input depends on the most recent measurement or state estimate. For example, in the input-design method developed by Blackmore and Williams (2006), the test signal is designed based on the most recent state estimate, but not redesigned when subsequent measurements become available. That is, when a fault is detected or suspected, the test signal is designed online based on the state estimate at that time. The signal is then applied to the system in open loop for the duration of the planned diagnosis experiment. This online design approach, based on the current state information, necessitates the signal is computed fast enough so that the state information is not outdated by the time the signal is applied. Clearly, an advantage of offline design is that the requirements on computational complexity of the input-design problem are less strict. Thus, an offline approach allows more complex problem formulations, including the use of more rigorous models or uncertainty descriptions. However, both online and closed-loop input design can offer advantages such as lower risk of constraint violation and faster diagnosis (e.g. Raimondo, Braatz, & Scott, 2013). This is often referred to as a less conservative approach, since using the most recent available information in designing the input enables mitigating some uncertainty.

In closed-loop approaches to input design, the current input is at any given time dependent on the most recent state estimate. This does not necessarily imply that an input-design problem is solved online at a regular interval, since the mapping from state estimate to input can be computed offline; e.g., see Punčochář et al. (2015b). The majority of the input-design methods discussed in the previous sections do not consider closed-loop implementation of the design procedure. However, the attention toward this approach is increasing, with *moving-horizon* strategies the most common. While closed-loop input design has been investigated since some of the earliest work on active fault diagnosis (Zhang & Zarrop, 1988), the following contributions are more representative of the modern approach to moving-horizon input design.

Šimandl et al. (2005) address the problem of closed-loop input design for stochastic linear systems using dynamic programming. Since the resulting Bellman equation is intractable to solve directly, they propose an approximate rolling-horizon solution method. Raimondo et al. (2013) propose two moving-horizon approaches to AFD for linear systems with set-based uncertainty descriptions. The first is based on solving the open-loop input-design problem developed by Scott et al. (2013) at every sampling time and applying the first element of the resulting input sequence. They show that this approach leads to significant advantages relative to applying the entire initial open-loop input sequence. The advantages include faster diagnosis and smaller input norm while ensuring guaranteed diagnosis within the specified time. However, the computational cost associated with online solution of the input-design problem is significant, which may render this closed-loop AFD approach impractical. This motivates the development of their other moving-horizon approach, which uses an explicit control law to shift most of the computational burden offline. While this approach also achieves faster diagnosis and smaller input norm relative to applying the precomputed open-loop input sequence, the improvements are not as significant as in the first approach. Both of the closed-loop AFD algorithms proposed by Raimondo et al. (2013) are further developed using constrained zonotopes by Raimondo, Marseglia, Braatz, and Scott (2016). Paulson et al. (2018) develops a computationally efficient solution method for moving-horizon input design in stochastic linear systems that enables updating the test signal at a high frequency, which leads to a low rate of misdiagnosis. For stochastic nonlinear systems with parametric uncertainty, both Paulson et al. (2017) and Martin-Casas and Mesbah (2018) propose closed-loop input-design methods. Paulson et al. (2017) compare their closed-loop approach to implementing the first input sequence in open loop and find that redesigning the test signal on a moving horizon results in a lower rate of misdiagnosis.

4.2. AFD and closed-loop control

There are two primary ways in which active fault diagnosis and closed-loop control interact. In the first, the test signal designed for AFD is applied to a system that is under closed-loop control. That is, the feedback control law is given a priori and must be accounted for in the input design. In the other, the system input is designed with a dual purpose: to control the system and to improve diagnosis. The overall design goal is optimizing some overall performance metric, and this strategy is thus similar to the dual control paradigm (Feldbaum, 1961). The methods in this latter class are distinct from the others discussed in this paper in that their design objective is not only improving fault diagnosis.

Ashari, Nikoukhah, and Campbell (2009a, 2009b, 2009c, 2012a, 2012b) investigate input design for systems under static linear feedback control. The test signal is designed offline and added to the control signal so that the sum is injected into the system. The main findings are that for finite-horizon input design, feedback control cannot reduce the worst-case input energy required for guaranteed diagnosis. However, feedback control can improve performance during the testing period as measured by the quadratic control cost (Ashari, Nikoukhah, & Campbell, 2012b). Furthermore, (Ashari, Nikoukhah, & Campbell, 2012a) find that for infinite-horizon input design, the optimal test signal is always sinusoidal.

Marseglia, Raimondo, Magni, and Mesbah (2017) also consider active fault diagnosis for a system under feedback control, adding the test signal to the control input. In their approach, however, an open-loop test signal is calculated at every sampling time on a moving horizon, with the first element of this input sequence being added to the control signal. The open-loop sequence is designed to separate the output set from the model that currently best matches the system measurements from those of the other models. Their analysis shows that the presence of the feedback controller can increase the diagnosis time, but when compared to the closed-loop AFD approach of Raimondo et al. (2016), input design under feedback control can reduce the cost of the online computations and result in faster diagnosis with an input that has smaller norm.

The other approach to combined AFD and closed-loop control, which is analogous to dual control, does not involve input design with the feedback law given a priori. Rather, the aim is designing inputs that strike an appropriate balance between fault diagnosis and control objectives while ensuring fast and reliable diagnosis. An early approach to this problem involves designing a controller that improves fault isolation by making the effect of any fault clearer from the input–output data. Nett, Jacobsen, and Miller (1988) and Jacobsen and Nett (1991) propose using the *four-parameter controller* (Nett, 1986) for this purpose. Their approach involves integrated design that considers control and diagnostic performance together. The analysis demonstrates the interaction between control and diagnosis and identifies benefits of considering both simultaneously. Niemann and Poulsen (2005) propose a similar approach for active diagnosis of parametric faults where the feedback controller is defined in terms of four parameters, designed using Youla–Jabr–Bongiorno–Kucera (YJBK) parameterization. A series of papers extends this framework to cases such as more than one fault model (Niemann, Poulsen, & Bækgaard, 2007) and multiple-input multiple-output (MIMO) systems (Niemann & Poulsen, 2014). See Ding (2009) for a comprehensive review of these and related methods.

Another approach to integrating AFD and control is to add diagnosis constraints to a control formulation. One such method is developed by Raimondo, Marseglia, Braatz, and Scott (2013), in which a robust model predictive control (MPC) formulation is augmented with constraints that guarantee separation of all model output sets at the end of the prediction horizon. Xu, Olaru, Puig, Ocampo-Martinez, and Niculesco (2014) propose a similar robust MPC framework, also relying on integrating active fault diagnosis through modifying the constraints. This approach, however, can handle sensor faults only and the constraints are modified when a fault is detected. Heirung and Mesbah (2017) and Heirung, Santos, and Mesbah (2019) develop MPC frameworks for stochastic nonlinear systems with model-structure uncertainty and demonstrate how their approaches can actively improve fault diagnosis. The latter control algorithm minimizes a multi-objective cost function that includes an approximate bound on the probability of selecting the wrong model (similar to Paulson et al., 2017) in addition to standard control objectives.

Šimandl and Punčochář (2009) pose the problem of integrated control and AFD as a stochastic optimal control problem. They analyze the problem using dynamic programming and pose three special cases that arise by varying a single parameter in the formulation: optimal control without AFD, optimal AFD without control, and a combination. Šimandl, Široký, and Punčochář (2011) investigate this approach further. A simple solution strategy that avoids dynamic programming and instead relies on open-loop optimization is proposed by Široký, Šimandl, Axehill, and Punčochář (2011). Here the integrated problem of control and active diagnosis is posed in three different ways: minimizing the input energy subject to an upper bound on the probability of misdiagnosis, minimizing the probability of misdiagnosis subject to an upper bound on the input energy, and minimizing a weighted sum of the input energy and the probability of misdiagnosis. These stochastic methods, among others, are summarized and compared by Škach and Punčochář (2015); see also Punčochář and Škach (2018), which includes a thorough discussion of the four-parameter approach we discuss above.

5. Discussion and opportunities for future research

Active fault diagnosis is fundamentally a problem of resolving uncertainty through model-based input design. In addition to the uncertainty that arises from the system's unknown fault status, each fault model can contain uncertain parameters and initial conditions and the system can be affected by exogenous uncertainties such as measurement noise and disturbances. The choices made in modeling faults and uncertainty dictate the formulation complexity of an input-design problem. Formulations with bounded non-probabilistic uncertainty generally guarantee diagnosis by ensuring separation of the output sets predicted from the fault models. Other important considerations in these formulations include minimizing the input energy or the length of the diagnosis experiment. Conversely, using probability distributions to describe uncertainty can result in some overlap between the predicted output sets. Instead of requiring full separation of these sets, a natural objective in probabilistic input design is minimizing either the overlap between the output distributions or the probability of misdiagnosis. Input-design problems for AFD can also be formulated as minimum-time problems, in which the goal typically is to minimize the time required for achieving full output separation or a specified probability of misdiagnosis, subject to input and state constraints, such as a bound on the available input energy. Alternatively, a weighted sum of two or more of these goals can form a preferable objective for AFD. In this type of multi-objective input-design problems, the choice of weights for the different objective terms can be guided by Pareto-analysis. In general, deciding how to formulate an input-design problem to achieve the desired diagnostic performance is a challenging task. For example, in probabilistic input-design problems it is not necessarily clear whether minimizing the misdiagnosis probability or the input energy, subject to a bound on the other, would result in a more effective test signal. There are currently no comprehensive studies available that investigate the consequences of such formulation choices, which warrants further research.

Input-design methods for AFD rely on separate models for each possible fault. When there is limited knowledge on the dynamic effects of a fault, developing an adequate fault model can pose a significant challenge. As faults are rare occurrences, it is unlikely that input–output data that can accurately describe faults is available, rendering data-driven model development an impractical option. Furthermore, the formulation of each fault model, as well as the total number of models, can have a significant impact on the computational complexity of the input-design problem. Some important modeling considerations include the degree of fidelity of the models, such as whether or not to include nonlinear effects with minor impact on the system behavior, whether and how to account for system uncertainties that may be of lesser importance for diagnosis, and the number of potential faults accounted for in the input-design problem. Arguably, another key choice is whether to model uncertainties using bounded sets or probability distributions (or some combination), in particular since exact bounds or probability distributions are rarely known in practice. Since models are generally imperfect representations of a system, especially in the case of faults exhibiting behavior that is difficult to predict, the value of using accurate uncertainty descriptions remains elusive. That is, an accurate uncertainty description offers questionable value if propagated through an inaccurate model. In other words, a test signal that guarantees diagnosis in theory does not offer the same guarantee when applied to the actual system if the models are imperfect and the uncertainty descriptions are inaccurate. Similarly, a theoretic upper bound on the probability of misdiagnosis may not relate to the frequency of diagnostic errors observed in practice. If the diagnosis performance is deemed inadequate using a particular method, the input-design problem can be reformulated by reconsidering the diagnosis objectives and constraints, as discussed above, or switching between set-based and probabilistic uncertainty descriptions.

Isolated and integrated design: With stricter requirements and growing complexity of modern technical systems, ensuring safe, reliable, and high-performance operation has become increasingly challenging. In particular, this exacerbates the nuanced interactions between the design of test signals and the controller (Šimandl & Punčochář, 2009), the test signal and the fault diagnoser, the fault detector and fault diagnoser (Kerestecioğlu & Zarrop, 1989; 1991), and the engineered system itself (Sampath et al., 1998). As a result, it is prohibitively challenging to consider all these interactions simultaneously in an integrated design of all these components. Common practice is therefore isolated design, often the system, the controller, the fault detector and diagnoser, and the test signal, in that order. Further research is necessary to gain a deeper understanding of these interactions and the consequences of various simplifying assumptions made in developing tractable approaches to design. In particular, controllers with an integrated capability to initiate AFD after the detection of a fault, while ensuring continued operation, have the potential to expand the system's capabilities for self maintenance in a manner that is less disruptive than injecting a pure test signal.

Computational complexity: The majority of approaches to input design for active fault diagnosis involves solving challenging nonconvex dynamic optimization problems. Much of the computational work can be done offline, either by computing a mapping from state estimate to input through some approximate method (e.g., Punčochář et al., 2015b) or by having the input be independent of the state estimate (Nikoukhah et al., 2000). As discussed in this paper, however, the literature largely focuses on the design of input signals with given uncertain initial conditions. Furthermore, the implementation of this design methodology on a receding horizon is receiving increased attention (e.g., Paulson et al., 2018; Paulson et al., 2017; Raimondo et al., 2016). Widespread adoption of closed-loop AFD, either through a receding-horizon implementation or through determining an approximate mapping offline are partly held back by the computational challenges involved in these approaches. Thus, investigations into problem formulations and solution algorithms that significantly lower the computational complexity can greatly contribute to the application of these methods to a broader array of problems. This is of particular importance for real-time applications and in large-scale systems.

Uncertainty propagation: Rigorous propagation of uncertainty is a significant source of computational complexity in many of the approaches discussed in this paper. This issue arises in both linear and nonlinear systems, and in both probabilistic and set-based problem formulations. Except for the linear-Gaussian case (Blackmore & Williams, 2006), predicting the output distributions or sets involves trading off computational cost against accuracy and precision. Significant progress in this area will help enable applications where computational cost is a limiting factor. This is of particular relevance for nonlinear systems, where further work in this area is also necessary to provide the diagnosis probabilities and guarantees available in the linear case. Note that some methods, such as the one proposed by Blanchini et al. (2017), avoid explicit uncertainty propagation, shift the computational burden offline, and can guarantee diagnosis for moderately large linear systems. This approach thus represents a potentially promising avenue for future research. The method developed by Blackmore and Williams (2006) is capable of locally optimal input design for large linear systems, owing to the low computational cost of propagating Gaussian probability densities through linear models. However, the extent to which globally optimal test signals, which can be prohibitively expensive to compute, are necessary for acceptable performance on large systems, is an open question. The issue of whether the minimized upper bound on the probability of misdiagnosis is sufficiently tight to be a suitable objective also warrants further research.

Distributed AFD: An interesting direction for lowering the computational requirements is distributing the AFD effort over connected subsystems. A distributed approach[5] involves designing a set of modules, each specific to a subsystem, that communicate with each other to accomplish a common task. This is in contrast to designing one central algorithm for the entire system. The promising results on decentralized AFD (that is, with no communication between subsystem-specific modules) demonstrated by Raimondo, Boem, Gallo, and Parisini (2016) motivate further research on this type of design approach, potentially opening up new areas of application through posing simpler problems and easier integration with existing control systems.

Defining diagnosability: A rigorous and readily applicable definition of diagnosability (cf. the discussion in Section 2.1) does not currently exist in the literature. Whether or not a system is diagnosable, or in some sense partially diagnosable, and under which conditions, are fundamental questions in AFD. A definition that is both precise and easily translated into a test of diagnosability has the potential to greatly advance the theoretical foundation for active fault diagnosis. Moreover, a diagnosability test can guide the initial system and control designs as well as inform the design of the test signal.

Practical applications: Active fault diagnosis has not seen widespread adoption in practical applications. The field would greatly benefit from research on real-world systems, with demonstrations of the viability and benefits of AFD. Such results will help identify new areas of application as well as clarify the most critical challenges and obstacles that require increased attention from the research community.

Conflict of interest

None.

Acknowledgments

This work is partially supported by NSF grant 1705706. We are also grateful to the three referees whose insightful comments and suggestions helped improve the paper.

References

Alwi, H., Edwards, C., Stroosma, O., & Mulder, J. A. (2008). Piloted sliding mode FTC simulator evaluation for the Elal Flight 1862 incident. *AIAA guidance, navigation and control conference and exhibit*. Honolulu, HI.

Andjelkovic, I., & Campbell, S. L. (2011). Direct optimization determination of auxiliary test signals for linear problems with model uncertainty. In *IEEE conference on decision and control* (pp. 909–914). Orlando, FL.

Andjelkovic, I., Sweetingham, K., & Campbell, S. L. (2008). Active fault detection in nonlinear systems using auxiliary signals. In *American control conference* (pp. 2142–2147). Seattle, WA.

Ashari, A. E., Nikoukhah, R., & Campbell, S. L. (2009a). Active robust fault detection of closed-loop systems: general cost case. In *IFAC symposium on fault detection, supervision and safety of technical processes* (pp. 585–590). Barcelona, Spain.

Ashari, A. E., Nikoukhah, R., & Campbell, S. L. (2009b). Asymptotic behavior and solution approximation of active robust fault detection for closed-loop systems. In *IEEE conference on decision and control* (pp. 1026–1031). Shanghai, P.R. China.

Ashari, A. E., Nikoukhah, R., & Campbell, S. L. (2009c). Feedback in active fault detection. In *IFAC symposium on system identification* (pp. 192–196). Saint-Malo, France.

Ashari, A. E., Nikoukhah, R., & Campbell, S. L. (2011). Auxiliary signal design for robust active fault detection of linear discrete-time systems. *Automatica, 47*(9), 1887–1895.

Ashari, A. E., Nikoukhah, R., & Campbell, S. L. (2012a). Active robust fault detection in closed-loop systems: Quadratic optimization approach. *IEEE Transactions on Automatic Control, 57*(10), 2532–2544.

Ashari, A. E., Nikoukhah, R., & Campbell, S. L. (2012b). Effects of feedback on active fault detection. *Automatica, 48*(5), 866–872.

Åström, K. J. (1991). Intelligent control. In *European control conference* (pp. 2328–2339). Grenoble, France.

Atkinson, A. C., & Cox, D. R. (1974). Planning experiments for discriminating between models. *Journal of the Royal Statistical Society. Series B (Methodological), 36*(3), 321–348.

Bazille, H., Fabre, E., & Genest, B. (2017). Diagnosability degree of stochastic discrete event systems. In *IEEE conference on decision and control* (pp. 5726–5731). Melbourne, Australia.

Bellman, R. (1957). *Dynamic programming*. Princeton University Press.

Bhattacharyya, A. (1943). On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society, 35*, 99–109.

Blackmore, L., Rajamanoharan, S., & Williams, B. C. (2008). Active estimation for jump Markov linear systems. *IEEE Transactions on Automatic Control, 53*(10), 2223–2236.

Blackmore, L., & Williams, B. C. (2005). Finite horizon control design for optimal model discrimination. In *IEEE conference on decision and control* (pp. 3795–3802). Seville, Spain.

Blackmore, L., & Williams, B. C. (2006). Finite horizon control design for optimal discrimination between several models. In *IEEE conference on decision and control* (pp. 1147–1152). San Diego, CA.

Blanchini, F., Casagrande, D., Giordano, G., Miani, S., Olaru, S., & Reppa, V. (2017). Active fault isolation: A duality-based approach via convex programming. *SIAM Journal on Control and Optimization, 55*(3), 1619–1640.

Blanke, M., Kinnaert, M., Lunze, J., & Staroswiecki, M. (2006). *Diagnosis and Fault–Tolerant control* (2nd). Springer.

Campbell, S. L., Drake, K. J., Andjelkovic, I., Sweetingham, K., & Choe, D. (2006). Model based failure detection using test signals from linearizations: A case study. In *IEEE conference on computer aided control systems design* (pp. 2659–2664). Munich, Germany.

Campbell, S. L., Horton, K. G., & Nikoukhah, R. (2002). Auxiliary signal design for rapid multi-model identification using optimization. *Automatica, 38*(8), 1313–1325.

Campbell, S. L., & Nikoukhah, R. (2004). *Auxiliary signal design for failure detection*. Princeton University Press.

Campbell, S. L., & Scott, J. R. (2016). Auxiliary signal design for failure detection with initial prior information. *Southeastcon*. Norfolk, VA.

Chen, J., & Patton, R. J. (1999). *Robust model-based fault diagnosis for dynamic systems*. Springer.

Chiang, L. H., Russell, E. L., & Braatz, R. D. (2001). *Fault detection and diagnosis in industrial systems*. Springer.

Choe, D., Campbell, S. L., & Nikoukhah, R. (2009). Optimal piecewise-constant signal design for active fault detection. *International Journal of Control, 82*(1), 130–146.

Chow, E. Y., & Willsky, A. S. (1984). Bayesian design of decision rules for failure detection. *IEEE Transactions on Aerospace and Electronic Systems, 20*(6), 761–774.

Ding, S. X. (2009). Integrated design of feedback controllers and fault detectors. *Annual Reviews in Control, 33*, 124–135.

Dua, V., Bozinis, N. A., & Pistikopoulos, E. N. (2002). A multiparametric programming approach for mixed integer and quadratic engineering problems. *Computers & Chemical Engineering, 26*, 715–733.

Dunia, R., & Qin, S. J. (1998). Subspace approach to multidimensional fault identification and reconstruction. *AIChE Journal, 44*(8), 1813–1831.

Fair, M., & Campbell, S. L. (2009a). Active incipient fault detection with more than two simultaneous faults. In *IEEE international conference on systems, man, and cybernetics* (pp. 3322–3327). San Antonio, TX.

Fair, M., & Campbell, S. L. (2009b). Active incipient fault detection with two simultaneous faults. In *IFAC symposium on fault detection, supervision and safety of technical processes* (pp. 573–578). Barcelona, Spain.

Feldbaum, A. A. (1961). Dual-control theory. I. *Automation and Remote Control, 21*(9), 874–880.

Gonzalez, R., Qi, F., & Huang, B. (2016). *Process control system fault diagnosis: A Bayesian approach*. Wiley.

Hatanaka, T., & Uosaki, K. (1994). Optimal auxiliary input for fault detection — frequency domain approach. In *IFAC symposium on system identification* (pp. 1069–1074). Copenhagen, Denmark.

Hatanaka, T., & Uosaki, K. (1996). Optimal auxiliary input for fault detection and fault diagnosis. In *IEEE international symposium on computer-aided control system design* (pp. 117–122). Dearborn, MI.

Hatanaka, T., & Uosaki, K. (1999a). Frequency domain approach to optimal auxiliary input design for fault diagnosis. In *European control conference* (pp. 1717–1722). Karlsruhe, Germany.

Hatanaka, T., & Uosaki, K. (1999b). Optimal auxiliary input design for fault diagnosis. In *IFAC world congress* (pp. 3862–3867). Beijing, China.

Hatanaka, T., & Uosaki, K. (2000). Optimal auxiliary input design for fault detection of systems with model uncertainty. In *IFAC symposium on system identification* (pp. 487–492). Santa Barbara, CA.

[5] See, e.g., Scattolini (2009) for an overview of distributed design in predictive control.

Heirung, T. A. N., & Mesbah, A. (2017). Stochastic nonlinear model predictive control with active model discrimination: a closed-loop fault diagnosis application. In *IFAC world congress* (pp. 16504–16509). Toulouse, France.

Heirung, T. A. N., Santos, T. L. M., & Mesbah, A. (2019). Model predictive control with active learning for stochastic systems with structural model uncertainty: Online model discrimination. *Computers & Chemical Engineering*. submitted.

Hellman, M. E., & Raviv, J. (1970). Probability of error, equivocation, and the Chernoff bound. *IEEE Transactions on Information Theory, 16*(4), 368–372.

Isermann, R. (2006). *Fault-diagnosis systems: An introduction from fault detection to fault tolerance*. Springer.

Jacobsen, C. A., & Nett, C. N. (1991). An integrated approach to controls and diagnostics using the four parameter controller. *IEEE Control Systems, 11*(6), 22–29.

Kailath, T. (1967). The divergence and Bhattacharyya distance measures in signal selection. *IEEE Transactions on Communication Technology, 15*(1), 52–60.

Kerestecioğlu, F. (1993). *Change detection and input design in dynamical systems*. Wiley.

Kerestecioğlu, F., & Çetin, I. (1997a). Auxiliary input design for detecting changes towards partially known hypotheses. In *IFAC fault detection, supervision and safety of technical processes* (pp. 1023–1028). Hull, UK.

Kerestecioğlu, F., & Çetin, I. (1997b). Auxiliary signal design for detecting changes towards unknown hypotheses. In *Ieee international symposium on intelligent control* (pp. 297–302). Istanbul, Turkey.

Kerestecioğlu, F., & Çetin, I. (2004). Optimal input design for the detection of changes towards unknown hypotheses. *International Journal of Systems Science, 35*(7), 435–444.

Kerestecioğlu, F., & Zarrop, M. B. (1989). Bayesian approach to optimal input design for failure detection and diagnosis. In *IFAC symposium on adaptive systems in control and signal processing* (pp. 525–529). Glasgow, UK.

Kerestecioğlu, F., & Zarrop, M. B. (1991). Optimal input design for change detection in dynamical systems. In *European control conference* (pp. 321–326). Grenoble, France.

Kerestecioğlu, F., & Zarrop, M. B. (1994). Input design for detection of abrupt changes in dynamical systems. *International Journal of Control, 59*(4), 1063–1084.

Khan, F. I., & Amyotte, P. R. (2007). Modeling of BP Texas City refinery incident. *Journal of Loss Prevention in the Process Industries, 20*(4–6), 387–395.

Kim, K.-K. K., Raimondo, D. M., & Braatz, R. D. (2013). Optimum input design for fault detection and diagnosis: model-based prediction and statistical distance measures. In *European control conference* (pp. 1940–1945). Zürich, Switzerland.

Kim, K.-K. K., Shen, D. E., Nagy, Z. K., & Braatz, R. D. (2013). Wiener's polynomial chaos for the analysis and control of nonlinear dynamical systems with probabilistic uncertainties. *IEEE Control Systems, 33*(5), 58–67.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *Annals of Statistics, 22*(1), 79–86.

Ljung, L. (1999). *System identification. Theory for the user* (2nd). Prentice-Hall.

Maciejowski, J. M., & Jones, C. N. (2003). MPC fault-tolerant flight control case study: Flight 1862. In *IFAC fault detection, supervision and safety of technical processes* (pp. 119–124).

Manca, D., & Brambilla, S. (2012). Dynamic simulation of the BP Texas City refinery accident. *Journal of Loss Prevention in the Process Industries, 25*(6), 950–957.

Marseglia, G. R., & Raimondo, D. M. (2017). Active fault diagnosis: A multi-parametric approach. *Automatica, 79*, 223–230.

Marseglia, G. R., Raimondo, D. M., Magni, L., & Mesbah, A. (2017). A probabilistic framework for reference design for guaranteed fault diagnosis under closed-loop control. In *IEEE conference on decision and control* (pp. 5739–5744). Melbourne, Australia.

Marseglia, G. R., Scott, J. K., Magni, L., Braatz, R. D., & Raimondo, D. M. (2014). A hybrid stochastic-deterministic approach for active fault diagnosis using scenario optimization. In *IFAC world congress* (pp. 1102–1107). Cape Town, South Africa.

Martin-Casas, M., & Mesbah, A. (2018). Active fault diagnosis for stochastic nonlinear systems: online probabilistic model discrimination. *IFAC international symposium on advanced control of chemical processes*. Shenyang, China.

Matusita, K. (1971). Some properties of affinity and applications. *Annals of the Institute of Statistical Mathematics, 23*, 137–155.

Mélykúti, B., August, E., Papachristodoulou, A., & El-Samad, H. (2010). Discriminating between rival biochemical network models: Three approaches to optimal experiment design. *BMC Systems Biology, 4*(38).

Mesbah, A., Streif, S., Findeisen, R., & Braatz, R. D. (2014). Active fault diagnosis for nonlinear systems with probabilistic uncertainties. In *IFAC world congress* (pp. 7079–7084). Cape Town, South Africa.

Nett, C. N. (1986). Algebraic aspects of linear control system stability. *IEEE Transactions on Automatic Control, 31*(10), 941–949.

Nett, C. N., Jacobsen, C. A., & Miller, A. T. (1988). An integrated approach to controls and diagnostics: the 4-parameter controller. In *American control conference* (pp. 824–835). Atlanta, GA.

Niemann, H. H., & Poulsen, N. K. (2005). Active fault diagnosis in closed-loop systems. In *IFAC world congress* (pp. 448–453). Prague, Czech Republic.

Niemann, H. H., & Poulsen, N. K. (2014). Active fault detection in MIMO systems. In *American control conference* (pp. 1975–1980). Portland, OR.

Niemann, H. H., Poulsen, N. K., & Bækgaard, M. A. B. (2007). A multi-model approach for system diagnosis. In *American control conference* (pp. 2539–2544). New York City, NY.

Nikoukhah, R. (1998). Guaranteed active failure detection and isolation for linear dynamical systems. *Automatica, 34*(11), 1345–1358.

Nikoukhah, R., & Campbell, S. L. (2003). Auxiliary signal design for robust active failure detection: the general cost case. In *IFAC fault detection, supervision and safety of technical processes* (pp. 251–256). Washington, DC.

Nikoukhah, R., & Campbell, S. L. (2006a). Auxiliary signal design for active failure detection in uncertain linear systems with a priori information. *Automatica, 42*(2), 219–228.

Nikoukhah, R., & Campbell, S. L. (2006b). Robust detection of incipient faults: an active approach. *Mediterranean conference on control and automation*. Ancona, Italy.

Nikoukhah, R., & Campbell, S. L. (2008). On the detection of small parameter variations in linear uncertain systems. *European Journal of Control, 14*(2), 158–171.

Nikoukhah, R., Campbell, S. L., & Delebecque, F. (2000). Detection signal design for failure detection: A robust approach. *International Journal of Adaptive Control and Signal Processing, 14*(7), 701–724.

Nikoukhah, R., Campbell, S. L., & Delebecque, F. (2001). Auxiliary signal design for failure detection in uncertain systems. *Mediterranean conference on control and automation*. Dubrovnik, Croatia.

Nikoukhah, R., Campbell, S. L., & Drake, K. J. (2010). An active approach for detection of incipient faults. *International Journal of Systems Science, 41*(2), 241–257.

Nikoukhah, R., Campbell, S. L., Horton, K. G., & Delebecque, F. (2002). Auxiliary signal design for robust multimodel identification. *IEEE Transactions on Automatic Control, 47*(1), 158–164.

Paoli, A., & Lafortune, S. (2005). Safe diagnosability for fault-tolerant supervision of discrete-event systems. *Automatica, 41*(8), 1335–1347.

Paulson, J. A., Heirung, T. A. N., Braatz, R. D., & Mesbah, A. (2018). Closed-loop active fault diagnosis for stochastic linear systems. In *American control conference* (pp. 735–741). Milwaukee, WI.

Paulson, J. A., Martin-Casas, M., & Mesbah, A. (2017). Input design for online fault diagnosis of nonlinear systems with stochastic uncertainty. *Industrial & Engineering Chemistry Research, 56*, 9593–9605.

Paulson, J. A., Raimondo, D. M., Findeisen, R., Braatz, R. D., & Streif, S. (2014). Guaranteed active fault diagnosis for uncertain nonlinear systems. In *European control conference* (pp. 926–931). Strasbourg, France.

Punčochář, I., Král, L., & Šimandl, M. (2009). Active fault detection for neural network based control of non-linear stochastic systems. In *IFAC fault detection, supervision and safety of technical processes* (pp. 125–130). Barcelona, Spain.

Punčochář, I., & Šimandl, M. (2014). On infinite horizon active fault diagnosis for a class of non-linear non-Gaussian systems. *International Journal of Applied Mathematics and Computer Science, 24*(4), 795–807.

Punčochář, I., & Škach, J. (2018). A survey of active fault diagnosis methods. In *IFAC symposium on fault detection, supervision and safety for technical processes* (pp. 1091–1098). Warsaw, Poland.

Punčochář, I., Škach, J., & Šimandl, M. (2015a). Adaptive generalized policy iteration in active fault detection and control. In *IFAC symposium on fault detection, supervision and safety of technical processes* (pp. 505–510). Paris, France.

Punčochář, I., Škach, J., & Šimandl, M. (2015b). Infinite time horizon active fault diagnosis based on approximate dynamic programming. In *IEEE conference on decision and control* (pp. 4456–4461). Osaka, Japan.

Qin, S. J. (1998). Control performance monitoring – A review and assessment. *Computers & Chemical Engineering, 23*(2), 173–186.

Raimondo, D. M., Boem, F., Gallo, A., & Parisini, T. (2016). A decentralized fault–tolerant control scheme based on active fault diagnosis. In *IEEE conference on decision and control* (pp. 2164–2169). Las Vegas, NV.

Raimondo, D. M., Braatz, R. D., & Scott, J. K. (2013). Active fault diagnosis using moving horizon input design. In *European control conference* (pp. 3131–3136). Zürich, Switzerland.

Raimondo, D. M., Marseglia, G. R., Braatz, R. D., & Scott, J. K. (2013). Fault-tolerant model predictive control with active fault isolation. In *Control and fault-tolerant systems* (pp. 444–449). Nice, France.

Raimondo, D. M., Marseglia, G. R., Braatz, R. D., & Scott, J. K. (2016). Closed-loop input design for guaranteed fault diagnosis using set-valued observers. *Automatica, 74*, 107–117.

Rawlings, J. B., Mayne, D. Q., & Diehl, M. (2017). *Model predictive control: Theory and design* (2nd). Nob Hill Publishing.

Saberi, A., Stoorvogel, A. A., Sannuti, P., & Niemann, H. H. (2000). Fundamental problems in fault detection and identification. *International Journal of Robust and Nonlinear Control, 10*(1), 1209–1236.

Sampath, M., Lafortune, S., & Teneketzis, D. (1998). Active diagnosis of discrete-event systems. *IEEE Transactions on Automatic Control, 43*(7), 908–929.

Sampath, M., Sengupta, R., Lafortune, S., Sinnamohideen, K., & Teneketzis, D. (1995). Diagnosability of discrete-Event systems. *IEEE Transactions on Automatic Control, 40*(9), 1555–1575.

Scattolini, R. (2009). Architectures for distributed and hierarchical model predictive control – A review. *Journal of Process Control, 19*(5), 723–731.

Scola, H. R., Nikoukhah, R., & Delebecque, F. (2003). Test signal design for failure detection: A linear programming approach. *International Journal of Applied Mathematics and Computer Science, 13*(4), 515–526.

Scott, J. K., Findeisen, R., Braatz, R. D., & Raimondo, D. M. (2013). Design of active inputs for set-based fault diagnosis. In *American control conference* (pp. 3567–3572). Washington, DC.

Scott, J. K., Findeisen, R., Braatz, R. D., & Raimondo, D. M. (2014). Input design for guaranteed fault diagnosis using zonotopes. *Automatica, 50*(6), 1580–1589.

Scott, J. K., Marseglia, G. R., Magni, L., Braatz, R. D., & Raimondo, D. M. (2013). A hybrid stochastic-deterministic input design method for active fault diagnosis. In *IEEE conference on decision and control* (pp. 5656–5661). Florence, Italy.

Scott, J. K., Raimondo, D. M., Marseglia, G. R., & Braatz, R. D. (2016). Constrained zonotopes: A new tool for set-based estimation and fault detection. *Automatica, 69*, 126–136.

Scott, J. R., & Campbell, S. L. (2014). Auxiliary signal design for failure detection in differential-algebraic equations. *Numerical Algebra, Control & Optimization, 4*(2), 151–179.

Severson, K., Chaiwatanodom, P., & Braatz, R. D. (2016). Perspectives on process monitoring of industrial systems. *Annual Reviews in Control, 42*, 190–220.

Šimandl, M., & Punčochář, I. (2009). Active fault detection and control: Unified formulation and optimal design. *Automatica, 45*(9), 2052–2059.

Šimandl, M., Punčochář, I., & Královec, J. (2005). Rolling horizon for active fault detection. In *IEEE conference on decision and control* (pp. 3789–3794). Seville, Spain.

Šimandl, M., Široký, J., & Punčochář, I. (2011). New special cases of general active change detection and control problem. In *IFAC world congress* (pp. 4260–4265). Milan, Italy.

Široký, J., Šimandl, M., Axehill, D., & Punčochář, I. (2011). An optimization approach to resolve the competing aims of active fault detection and control. In *IEEE conference on decision and control* (pp. 3712–3717). Orlando, FL.

Škach, J., & Punčochář, I. (2015). Active fault detection: A comparison of probabilistic methods. *European workshop on advanced control and diagnosis*. Pilsen, Czech Republic.

Škach, J., & Punčochář, I. (2017). Input design for fault detection using extended Kalman filter and reinforcement learning. In *IFAC world congress* (pp. 7302–7307). Toulouse, France.

Škach, J., Punčochář, I., & Lewis, F. L. (2016). Optimal active fault diagnosis by temporal-difference learning. In *IEEE conference on decision and control* (pp. 2146–2151). Las Vegas, NV.

Škach, J., Punčochář, I., & Straka, O. (2017). Active fault diagnosis for jump Markov nonlinear systems. In *IFAC world congress* (pp. 7308–7313). Toulouse, France.

Škach, J., Straka, O., & Punčochář, I. (2017). Efficient active fault diagnosis using adaptive particle filter. In *IEEE conference on decision and control* (pp. 5732–5738). Melbourne, Australia.

Tabatabaeipour, S. M. (2015). Active fault detection and isolation of discrete-time linear time-varying systems: A set-membership approach. *International Journal of Systems Science, 46*(11), 1917–1933.

Tawarmalani, M., & Sahinidis, N. V. (2002). Convexification and global optimization in continuous and mixed-Integer nonlinear programming. theory, algorithms, software, and applications. *Nonconvex Optimization and Its Applications*. Springer.

Uosaki, K., Takata, N., & Hatanaka, T. (1993). Optimal auxiliary input for on-line fault detection and fault diagnosis. In *IFAC world congress* (pp. 441–446). Sydney, Australia.

Uosaki, K., Tanaka, I., & Sugiyama, H. (1984). Optimal input design for autoregressive model discrimination with constrained output variance. *IEEE Transactions on Automatic Control, 29*(4), 348–350.

Varga, A. (2017). Solving fault diagnosis problems. linear synthesis techniques. *Studies in Systems, Decision and Control*: 84. Springer.

Venkatasubramanian, V., Rengaswamy, R., & Kavuri, S. N. (2003). A review of process fault detection and diagnosis. part II: Qualitative models and search strategies. *Computers & Chemical Engineering, 27*(3), 313–326.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N., & Yin, K. (2003). A review of process fault detection and diagnosis. part III: Process history based methods. *Computers & Chemical Engineering, 27*(3), 327–346.

Venkatasubramanian, V., Rengaswamy, R., Yin, K., & Kavuri, S. N. (2003). A review of process fault detection and diagnosis. part I: Quantitative model-based methods. *Computers & Chemical Engineering, 27*(3), 293–311.

Wiener, N. (1938). The homogeneous chaos. *American Journal of Mathematics, 60*(4), 897–936.

Williard, N., He, W., Hendricks, C., & Pecht, M. (2013). Lessons learned from the 787 dreamliner issue on lithium-ion battery reliability. *Energies, 6*(9), 4682–4695.

Xiu, D., & Karniadakis, G. E. (2002). The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing, 24*(2), 619–644.

Xu, F., Olaru, S., Puig, V., Ocampo-Martinez, C., & Niculesco, S.-I. (2014). Sensor-fault tolerance using robust MPC with set-based state estimation and active fault isolation. In *IEEE conference on decision and control* (pp. 4953–4958). Los Angeles, CA.

Yang, X., & Scott, J. K. (2018). A comparison of zonotope order reduction techniques. *Automatica, 95*, 378–384.

Zagrobelny, M., Ji, L., & Rawlings, J. B. (2013). Quis custodiet ipsos custodes? *Annual Reviews in Control, 37*(2), 260–270.

Zhang, X. J. (1989). *Auxiliary signal design in fault detection and diagnosis*. Springer.

Zhang, X. J., & Zarrop, M. B. (1988). Auxiliary signals for improving on-line fault detection. In *IEE international conference on control* (pp. 414–419). Oxford, UK.

Zhang, Y., & Jiang, J. (2008). Bibliographical review on reconfigurable fault-tolerant control systems. *Annual Reviews in Control, 32*(2), 229–252.