A Bayesian Approach for Active Fault Isolation with an Application to Leakage Localization in Water Distribution Networks

van Lagen, G.; Abraham, E.; Mohajerin Esfahani, P.

DOI
10.1109/TCST.2022.3201334

Publication date
2022

Document Version
Final published version

Published in
IEEE Transactions on Control Systems Technology

A Bayesian Approach for Active Fault Isolation With an Application to Leakage Localization in Water Distribution Networks

Gert van Lagen, Edo Abraham, and Peyman Mohajerin Esfahani

***Abstract*— This article proposes an active fault isolation method for application to water distribution networks (WDNs) to localize leaks. The method relies on the classification of observed outputs to a discrete set of hypothetical faults. Due to parametric uncertainties, the outputs are random vectors that follow unknown probability distribution functions (PDFs). The output PDFs corresponding to the considered faults are approximated using smooth kernel density estimation (SKDE). They are used to calculate the posterior probability of each hypothesis, given the observed outputs, by applying Bayes' rule. The difficulty to classify observed outputs to the right fault comes from the overlap between output PDFs. An active algorithm is introduced that proactively minimizes the joint overlap between the output PDFs by designing optimal control inputs. Due to physical limitations on control inputs and depending on the intensity of uncertainties, full separation, and hence fault isolation, cannot be guaranteed for a single observed output. Therefore, subsequent observations are used in an iterative framework, where the posterior probabilities of the previous time step serve as the prior probabilities for the next time step. The method is applied to locate leaks in a benchmark WDN for different levels of uncertainty in customer water demand and leakage magnitude. Improvements in the performance are observed in comparison to the best considered passive method from literature.**

***Index Terms*— Bayesian classification, leak localization, stochastic active fault diagnosis (AFD), water distribution networks (WDNs).**

I. Introduction

ONE of the main challenges for water utilities is the diagnosis and control of leakage from aging water distribution networks (WDNs). The early detection and management of leaks, in addition to reducing the cost in non-revenue water and conserving energy, is critical to mitigate deterioration of pipes and the surrounding infrastructure. Water loss due to leakage varies between 5% and above 50% of the supplied volume, respectively, for well-managed and older poorly maintained networks [1]. Leakage reduction beyond the economically optimal level of about 15% [2] is further motivated by stringent regulations and imminent risks. One risk is a poorer water quality due to temporarily negative pressures that allow intrusion of pollutants into the network, potentially jeopardizing public health [3]. A further threat is that very small leaks can gradually grow in size, eventually into pipe bursts, which can render (a part of) the network inoperable and result in damage to other infrastructure and also economical losses due to flooded areas. Leaks are also known to cause destructive and dangerous sink holes due to underground soil erosion [4]. Finally, reducing the annual global loss of 32 billion m^3 of potable water [5] will help alleviate the water stress caused by the mutually reinforcing global issues of rapid urbanization and increasing water scarcity.

Leakage analysis includes the fault diagnosis tasks of detection, isolation, identification, and estimation. These techniques, respectively, involve the determination of whether or not a leak is present, if so, to estimate its location, type, and magnitude [6]. This article focuses on leak localization techniques.

Different methods to locate leaks in WDNs have been proposed in literature. The conventional deterministic techniques include random and regular sounding surveys using listening sticks and acoustic loggers, and step-testing of district metered areas (DMAs) through gradual valve closures [7]. DMAs are subsystems that can be analytically isolated through segregation of WDN by means of (dynamic) boundary valves and metering the flows at remaining open connections [8], [9]. More advanced deterministic methods such as leak noise correlators, pig-mounted acoustic sensing, and gas-injection techniques [10] are the most precise at locating leaks. However all these techniques come with expensive equipment cost and are man-hour intensive, and so are not scalable. In addition, the suppression of leakage sound signatures by reduced pressures in active pressure management has also made these methods of limited application [7], [10].

To make those deterministic methods scalable, recent approaches use model-based analysis to reduce the search space. These methods use near real-time telemetry data from pressure sensors and flow meters distributed over the network and rely on a calibrated predictive hydraulic model. Based on the observed residuals that reflect how pressure measurements from the leaky reality deviate from the predicted pressure values in the absence of leakage, their aim

Manuscript received 6 November 2021; revised 8 April 2022; accepted 12 August 2022. The work of Peyman Mohajerin Esfahani was supported by the European Research Council (ERC) Grant TRUST-949796. Recommended by Associate Editor G. Orosz. *(Corresponding author: Edo Abraham.)*

Gert van Lagen is with Yunex Traffic, 2712 PN Zoetermeer, The Netherlands (e-mail: gertvlagen@gmail.com).

Edo Abraham is with the Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, 2628 CN Delft, The Netherlands (e-mail: e.abraham@tudelft.nl).

Peyman Mohajerin Esfahani is with the Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: p.mohajerinesfahani@tudelft.nl).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCST.2022.3201334.

Digital Object Identifier 10.1109/TCST.2022.3201334

is to designate a leak location from a discretized set of possibilities through comparison to offline generated residual signatures corresponding to the possible locations [11]. Recent developments toward this problem apply machine learning techniques such as k-nearest neighbors, neuro-fuzzy [12], Bayesian classifiers [13], and Fisher discriminant analysis [14] and Dempster–Shafer [15] to classify observed residuals to one of the possible leak locations, which have shown best results when applied over multiple time steps. Recently, there are also some progress to leverage these tools to go beyond the problem of detection to answer more complex questions such as estimating the time and intensity of the leakage [16], [17]. Similarly, Steffelbauer *et al.* [18] make use of the time-series analysis for detecting the start time static and growing leaks, and then model-based passive approaches for leak localization.

The difficulty to classify observed residuals to one of the possible leak locations comes from the overlap between the corresponding residual sets due to uncertainties, such that leak isolation cannot be achieved under all possible observed residuals. All these techniques in literature rely on nominal input–output data from the network, i.e., control input strategies are not adapted to improve leak localization, which we refer to as the passive fault diagnosis (PFD) methods. As the joint overlap in residual space increases with growing uncertainties, these PFD methods are of limited application. Therefore, in this work we present a novel active fault diagnosis (AFD) algorithm for faster and more reliable leak isolation. Where pressure control inputs are usually regulated at a minimum level using pressure regulating valves (PRVs) [9], we show that they can also be optimized to enhance active leakage isolation [19]. We will make use of output observations directly, rather than the common residual observations used in literature that subtract two random "output" variables, because composed random variables gain a higher joint spread and hence, unnecessarily, stochastically deteriorate the observed samples. We also introduce new control design strategies for pressure inputs that minimize the probability of misisolation, i.e., the overlap between output sets corresponding to considered leak locations. The output sets are described by probability distributions, which are estimated by means of smooth kernel density estimation (SKDE) [20] through extrapolation of output realizations from Monte Carlo simulations. Due to physical limitations and regulatory constraints, the pressure control inputs to the network are bounded. Hence, it is plausible that output sets cannot be fully separated, such that isolation is not guaranteed. However, by iteratively applying Bayes' rule over consecutive time steps, leak isolation can be improved in terms of reliability and speed compared with the PFD methods in [12], [13], and [14]. At nighttime, when user demand is low [21], the proposed AFD algorithm is applied to a benchmark network for different degrees of uncertainty. Its performance is compared with that of the best considered PFD method proposed in [13] with the slight adaptation of using the output space instead of the residual space.

This article is structured as follows. In Section II, the active fault isolation problem is stated and a motivating example for leak localization in a WDN is given. In Section III-A, an AFD algorithm is proposed that solves the stated problem and is directly applicable to locating leaks in a WDN, which is further elaborated in Section IV. In Section V, the simulation experiments are presented with a benchmark WDN as a case study, in which the AFD algorithm is tested for different leak scenarios and compared against a PFD method. Finally, conclusions are drawn and recommendations for future work are given in Section VI.

II. Problem Statement and Motivating Example

In this section, a class of models describing the steady-state of a nonlinear system subject to possible faults are introduced. Consider the set of algebraic equations

$$\text{Model}: \begin{cases} F(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{d}, K) = 0 \\ \boldsymbol{y} = C(\boldsymbol{x}) \end{cases} \tag{1}$$

where the function F models the steady-state of the system, the signal $\boldsymbol{x}$ denotes the state of the system, $\boldsymbol{u}$ is the control input, $\boldsymbol{d}$ denotes the natural disturbances that the system may encounter, and $\boldsymbol{y}$ is the available measurement signal. We highlight that the bold signals $\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{d}$, and $\boldsymbol{y}$ are time-varying and take vector values from $\mathbb{R}^{n_x}, \mathbb{R}^{n_u}, \mathbb{R}^{n_d}$, and $\mathbb{R}^{n_y}$, respectively. The parameter $K = [K_1, \ldots, K_{n_K}] \in \mathbb{R}^{n_K}$ is a constant vector representing n_K different possible faults, i.e., when the ith component of K is nonzero (i.e., $K_i \neq 0$), then the system is in the ith faulty mode.

The set of algebraic equations (1) essentially describes the input–output mapping between the variables $(\boldsymbol{d}, \boldsymbol{u}; K)$ and $\boldsymbol{y}$. In this view, the output can be explicitly described by $\boldsymbol{y}(\boldsymbol{d}, \boldsymbol{u}; K)$. For brevity, and with slight abuse of notation, we may use the shorthand notation

$$\boldsymbol{y}^{[i]} := \boldsymbol{y}(\boldsymbol{d}, \boldsymbol{u}; K_i), \qquad i \in \{1, \ldots, n_K\}$$

where $\boldsymbol{y}^{[i]}$ denotes the output of the system (1) in the presence of fault i, namely, when the only component of the vector K that is not zero is K_i. In this study, we treat $(\boldsymbol{d}; K_i)$ as random variables whose behavior follows a certain distribution from which we have access to historical data or sample realizations. We also reserve the symbol "$\widehat{\ }$" for sample realizations of the random variables, e.g., given realizations $(\widehat{\boldsymbol{d}}, \widehat{K}_i)$, we denote an output realization $\widehat{\boldsymbol{y}}^{[i]} = \boldsymbol{y}(\widehat{\boldsymbol{d}}, \boldsymbol{u}; \widehat{K}_i)$. The aim of this study is to address the following objective.

Problem 1 (Active Fault Isolation): Consider the system (1) under a single fault i^*, i.e., $K_j = 0$ if and only if $j \neq i^*$. Given the measurement signal $\boldsymbol{y}$ and statistical information of the natural disturbance $\boldsymbol{d}$, synthesize a sequence of input signal $\boldsymbol{u}$ and a diagnosis rule to maximize the probability of identifying the fault type i^*.

The relevance of Problem 1 is endorsed by the following motivating example.

Example 1 (Leak Localization in WDN): Water enters a WDN at n_u inlets and is supplied to consumers abstracted by n_K nodes that are connected to the inlets through a network of n_p pipes. The steady-state of a WDN can be described by a model of the form in (1), where the state $\boldsymbol{x} := \begin{bmatrix} \boldsymbol{q} & \boldsymbol{h} \end{bmatrix}^T \in \mathbb{R}^{n_x}$ consists of the flows $\boldsymbol{q} \in \mathbb{R}^{n_p}$ through the pipes (in m^3/s) and the hydraulic heads $\boldsymbol{h} \in \mathbb{R}^{n_K}$ at the nodes (in $\mathrm{mH_2O}$). The control inputs $\boldsymbol{u}$ are the inlet pressures of the network,

which are regulated using PRVs. The nodal consumer demands act like natural disturbances $\boldsymbol{d}$ on the network and need to be predicted using statistical information. The output of the network consists of the measured part of the system's state, where usually $n_y \ll n_x$. Consider the WDN under the presence of a single leak i^* at one of the n_K nodes, then, active isolation of this fault parametrized by K involves the synthesis of a sequence of control inputs $\boldsymbol{u}$ to maximize the probability of identifying the fault type i^* based on a sequence of measurements $\boldsymbol{y}$.

III. Proposed Methodology

In this section, we provide an AFD methodology built on a Bayesian perspective, an approach in which Bayes' theorem is used to update the probability for a hypothesis as more information is revealed to us. In the context of active fault isolation, roughly speaking, the hypothesis is our current belief about the probability of occurring for each fault (i.e., $K_i \neq 0$) and the information is the output measurement $\widehat{\boldsymbol{y}}^* = \boldsymbol{y}(\widehat{\boldsymbol{d}}, \boldsymbol{u}; \widehat{K}_{i^*})$ from the actual system, which is supposed to be generated by an unknown fault mode i^*. The proposed active fault isolation in this study comprises two main steps: 1) update our belief upon receiving an output measurement $\widehat{\boldsymbol{y}}^*$ and 2) introduce an appropriate input signal $\boldsymbol{u}$.

A. Bayesian Update of Hypotheses Probabilities

Recall that in the setting of this study we believe that the system is faulty and that one of the modes $i \in \{1, \dots, n_K\}$ occurs. Looking at the problem from a Bayesian perspective, it is then natural to define the hypothesis set $\mathbb{H} = \{1, \dots, n_K\}$ along with a probability distribution $\mathbb{P}$ representing our current (prior) belief about hypothesis candidates. Formally speaking

$$\mathbb{P}(i) := \text{Prob(fault mode: } i), \qquad i \in \mathbb{H}. \tag{2}$$

Recall also that given an input signal $\boldsymbol{u}$, the output of the system under fault mode i, denoted by $\boldsymbol{y}^{[i]} = \boldsymbol{y}(\boldsymbol{d}, \boldsymbol{u}; K_i)$, is a random variable whose distribution is induced by the distributions of the variables $(\boldsymbol{d}; K_i)$ through the algebraic equations (1). With this in mind, we denote the (conditional) distribution of the output measurements by

$$\boldsymbol{y}^{[i]} = \boldsymbol{y}(\boldsymbol{d}, \boldsymbol{u}; K_i) \sim \mathbb{P}(\mathrm{d}y|i, \boldsymbol{u}) = p(y|i, \boldsymbol{u})\mathrm{d}y \tag{3}$$

where $p(y|i, \boldsymbol{u})$ represents the probability density function; throughout this study, we assume such a density function exists. Given the definitions in (2) and (3), the marginal density distribution of the output measurement is

$$p(y|\boldsymbol{u}) = \sum_{j=1}^{n_K} p(y|j, \boldsymbol{u})\mathbb{P}(j). \tag{4}$$

Upon receiving a realization of the output $\widehat{\boldsymbol{y}}^*$ under the input signal $\boldsymbol{u}$, one can update the prior belief concerning the hypothesis candidates in (2) by means of Bayes' theorem through the relation

$$\mathbb{P}\big(i|\widehat{\boldsymbol{y}}^*, \boldsymbol{u}\big) = \frac{p(\widehat{\boldsymbol{y}}^*|i, \boldsymbol{u})}{p(\widehat{\boldsymbol{y}}^*|\boldsymbol{u})}\mathbb{P}(i) = \frac{p(\widehat{\boldsymbol{y}}^*|i, \boldsymbol{u})\mathbb{P}(i)}{\sum_{j=1}^{n_K} p(\widehat{\boldsymbol{y}}^*|j, \boldsymbol{u})\mathbb{P}(j)} \tag{5}$$

where the second equality follows from (4). The conditional distribution $\mathbb{P}(i|\widehat{\boldsymbol{y}}^*, \boldsymbol{u})$ in (5) is also known as the posterior distribution.

1) Approximation Techniques: Given the prior distribution (2), the key ingredient is the density function $p(\widehat{\boldsymbol{y}}^*|i, \boldsymbol{u})$ evaluated at the measurement $\widehat{\boldsymbol{y}}^*$; this quantity is also known as likelihood in the statistics literature [22, Ch. 4.4]. As pointed out earlier, this density function is essentially determined by the distributions of $(\boldsymbol{d}; K_i)$ through the system equations (1). In general, the analytical description of this density is not available and one has to resort to approximation techniques for numerical purposes. For instance, for each hypothesis $i \in \mathbb{H}$, given an input signal $\boldsymbol{u}$, and M realizations $(\widehat{\boldsymbol{d}}_m, \widehat{K}_{i,m})$, $m \in \{1, \dots, M\}$, one can simulate the system (1) and compute M output realizations $\widehat{\boldsymbol{y}}_m^{[i]}(\boldsymbol{u})$, $m \in \{1, \dots, M\}$; note that these realizations depend on the choice of $\boldsymbol{u}$. A single realization can be obtained by fixing $\boldsymbol{u}$ and i, generating a realization of $\boldsymbol{d}$ and solving (1) for $\boldsymbol{y}$. Now, since the required number of realizations for all the considered hypotheses scales proportional to $M \times n_K$, it becomes computationally very demanding to take a large M. Therefore, having a moderate number of output samples $\{\widehat{\boldsymbol{y}}_m^{[i]}(\boldsymbol{u})\}_{m \leq M}$, we use a kernel function $\kappa : \mathbb{R}^{n_y} \times \mathbb{R}^{n_y} \to \mathbb{R}_+$ to arrive at a smooth approximation of the conditional probability density distribution of $\boldsymbol{y}$ given $(i, \boldsymbol{u})$

$$p(y|i, \boldsymbol{u}) \approx \frac{1}{M}\sum_{m=1}^{M} \kappa\big(y, \widehat{\boldsymbol{y}}_m^{[i]}(\boldsymbol{u})\big). \tag{6}$$

Considering the approximation scheme (6), the approximation of the posterior distribution (5) then reduces to

$$\mathbb{P}\big(i|\widehat{\boldsymbol{y}}^*, \boldsymbol{u}\big) \approx \frac{\sum_{m=1}^{M} \kappa\big(\widehat{\boldsymbol{y}}^*, \widehat{\boldsymbol{y}}_m^{[i]}(\boldsymbol{u})\big)\mathbb{P}(i)}{\sum_{j=1}^{n_K}\sum_{m=1}^{M} \kappa\big(\widehat{\boldsymbol{y}}^*, \widehat{\boldsymbol{y}}_m^{[j]}(\boldsymbol{u})\big)\mathbb{P}(j)}.$$

In the Section V, we will provide a specific example of such kernels. Note that when the prior probability $\mathbb{P}(i) = 0$, then the posterior update (5) remains $\mathbb{P}(i|\widehat{\boldsymbol{y}}^*, \boldsymbol{u}) = 0$ irrespective of the observation $\widehat{\boldsymbol{y}}^*$. In this light, another approximation idea to practically improve the efficiency of the Bayesian update rule (5) is to introduce a threshold, say ζ, and set the posterior probability $\mathbb{P}(i|\widehat{\boldsymbol{y}}^*, \boldsymbol{u})$ to zero if $\mathbb{P}(i|\widehat{\boldsymbol{y}}^*, \boldsymbol{u}) \leq \zeta$. In this way, hypotheses with a close to zero belief are set to zero so that they can be neglected in the next iteration, speeding up the "knockout race." This approximation can be mathematically described as

$$\mathbb{P}\big(i|\widehat{\boldsymbol{y}}^*, \boldsymbol{u}\big) \leftarrow \begin{cases} \mathbb{P}\big(i|\widehat{\boldsymbol{y}}^*, \boldsymbol{u}\big) / \sum_{j \in I_\zeta} \mathbb{P}\big(j|\widehat{\boldsymbol{y}}^*, \boldsymbol{u}\big), & i \in I_\zeta \\ 0, & i \notin I_\zeta \end{cases} \tag{7}$$

where $I_\zeta := \{j \in \mathbb{H} : \mathbb{P}(j|\widehat{\boldsymbol{y}}^*, \boldsymbol{u}) > \zeta\}$. We note that kernelization is not the only way to construct an expression for unavailable distributions. Another strong candidate would be the approximate Bayesian computation method, of which more information can be found in [23].

B. Input Synthesis

This section focuses on synthesizing a feasible input signal $\boldsymbol{u}$ at each time instant to generate an "optimal"

measurement $\widehat{\boldsymbol{y}}^*$ for the Bayesian update described in (5). As a first step toward this goal, we need to formally define such an optimal process. Intuitively speaking, a key feature to isolate the fault modes is to separate the conditional distribution $\mathbb{P}(\mathrm{d}y|i,\boldsymbol{u}) = p(y|i,\boldsymbol{u})\mathrm{d}y$ from one another for all $i \in \mathbb{H}$. Note that in an ideal setting where these distributions have zero overlap, then in the posterior distribution update in (5), the quantity $p(\widehat{\boldsymbol{y}}^*|i,\boldsymbol{u}) = 0$ for all $i \neq i^*$. This means that Bayes' rule (5) immediately converges to the optimal distribution fully supported on the true fault mode i^*. In this view, we first choose a distance function $D(\mathbb{P}_1,\mathbb{P}_2)$ that essentially captures the overlaps of two distributions $\mathbb{P}_1$ and $\mathbb{P}_2$, i.e., $\mathcal{D}(\mathbb{P}_1,\mathbb{P}_2) \geq 0$ and is zero if and only if $\mathbb{P}_1 \equiv \mathbb{P}_2$. Given this distance function, we then introduce the objective function

$$J(\boldsymbol{u}) := \sum_{i,j=1}^{n_K} \mathbb{P}(i)\mathcal{D}(\mathbb{P}(\mathrm{d}y|i,\boldsymbol{u}),\mathbb{P}(\mathrm{d}y|j,\boldsymbol{u}))\mathbb{P}(j). \tag{8}$$

Our goal is to maximize the objective function (8), and thus to reduce the similarity between the marginal distributions $\mathbb{P}(\mathrm{d}y|i,\boldsymbol{u})$ for $i \in \mathbb{H}$, over the admissible set of inputs $\boldsymbol{u} \in \mathbb{U}$. The cumulative overlap between marginal distributions is weighted with the belief about their corresponding hypothesis candidates, such that the algorithm at any time focuses on separating the hypotheses with highest belief. From a computational perspective however, the function J in (8) may not be convex and it is not computationally feasible to solve $\max_{\boldsymbol{u}\in\mathbb{U}} J(\boldsymbol{u})$ per se. Therefore, we propose the projected gradient ascent update rule where at each iteration t, we only require to compute the gradient $(\partial J/\partial u)(u_t)$ at a given u_t. More formally, we propose

$$u_{t+1} = \Pi_{\mathbb{U}}\left[u_t + \eta\frac{\partial J}{\partial u}(u_t)\right] \qquad t \in \mathbb{N} \tag{9}$$

where $\Pi_{\mathbb{U}}$ is the projected operation on the set $\mathbb{U}$, and the constant η is a prespecified stepsize. The key ingredient to implement the input update (9) is the computation of $(\partial J/\partial u)$, the gradient of the cost function. This quantity indeed entails the behavior of the algebraic equations (1). We note that the choice of the distant function $\mathcal{D}(\mathbb{P}_1,\mathbb{P}_2)$ is a degree of freedom as long as the basic properties of a metric on the space of distributions are fulfilled. In anticipation of the application in the next section and for numerical purposes, we consider the Hellinger distance defined as follows:

$$\mathcal{D}(\mathbb{P}_1,\mathbb{P}_2) = 1 - \int \sqrt{p_1(y)p_2(y)}dy. \tag{10}$$

The Hellinger distance (10) is qualified as a metric, as opposed to the common KL-divergence measure. This metric is also perceived as the stochastic analog of the Euclidean distance. The metric can therefore be implemented intuitively and unambiguously, because the three basic axioms (identity of indiscernibles, symmetry, and the triangle inequality) hold. Specifically, the symmetry axiom is important in this application, because the degree of overlap does not change with perspective between two overlapping spatial objects, i.e., $\mathcal{D}(\mathbb{P}_1,\mathbb{P}_2) = \mathcal{D}(\mathbb{P}_2,\mathbb{P}_1)$.

Algorithm 1 Bayesian Based Active Fault Detection

Input: u_0, $\mathbb{P}_0$, η, ζ, $p_{\max}$, t_{max}
Output: $\boldsymbol{u}_t$, $\mathbb{P}_t$
Ensure: $t = 1$, $\mathbb{P}(i) = \mathbb{P}_0(i), \forall i \in \mathbb{H}$

```
while max_{i∈H} P(i) ≤ p_max and t ≤ t_max do
    Compute (∂J/∂u)(u_{t-1}) using (11)
    Update control u_t using (9)
    Measure real output ŷ_t^*
    for i ∈ H do
        Construct conditional distributions using (6)
        Sequentially update posterior distribution P(i|ŷ_t^*, u_t) using (5) and (7)
    end for
    t ← t + 1
end while
```

Proposition 1 (Cost Function Gradient): Let the cost function $J(\cdot)$ be defined as in (8) where the distance function is (10). Given $\boldsymbol{u} \in \mathbb{U}$ and realizations of random variables $\{\widehat{\boldsymbol{d}}_m, \widehat{K}_{i,m}\}$, $i \in I_\zeta$, let the set $\{\widehat{\boldsymbol{x}}(\boldsymbol{u})_m^{[i]}, \widehat{\boldsymbol{y}}(\boldsymbol{u})_m^{[i]} : i \in I_\zeta,\ 1 \leqslant m \leqslant M\}$ be "M" measurements of the model (1). Suppose κ to be a kernel function and $\mathbb{P}(\boldsymbol{y}|\boldsymbol{u})$ is approximated by (6). Then

$$\frac{\partial J}{\partial \boldsymbol{u}} = \frac{1}{2}\sum_{i,j}\mathbb{P}(i)\left(\int \gamma_{i,j}(\boldsymbol{y},\boldsymbol{u}) + \gamma_{j,i}(\boldsymbol{y},\boldsymbol{u})d\boldsymbol{y}\right)\mathbb{P}(j) \tag{11}$$

where the function $\gamma_{i,j}$ is defined as

$$\gamma_{i,j} := \frac{\sum_{m=1}^M \kappa\big(\boldsymbol{y},\widehat{\boldsymbol{y}}_m^{[i]}(\boldsymbol{u})\big)}{\sum_{m=1}^M \kappa\big(\boldsymbol{y},\widehat{\boldsymbol{y}}_m^{[j]}(\boldsymbol{u})\big)}\left(\frac{1}{M}\sum_{m=1}^M \frac{\partial \kappa}{\partial \widehat{\boldsymbol{y}}}\big(\boldsymbol{y},\widehat{\boldsymbol{y}}_m^{[j]}(\boldsymbol{u})\big)\frac{\partial \widehat{\boldsymbol{y}}^{[j]}}{\partial \boldsymbol{u}}\right)$$

$$\frac{\partial \widehat{\boldsymbol{y}}^{[j]}}{\partial \boldsymbol{u}} = \frac{\partial C}{\partial \boldsymbol{x}}\left(\frac{\partial F}{\partial \boldsymbol{x}}\right)^{-1}\frac{\partial F}{\partial \boldsymbol{u}}\big(\widehat{\boldsymbol{x}}_m^{[j]}(\boldsymbol{u}),\widehat{\boldsymbol{d}}_m,\boldsymbol{u},\widehat{K}_{j,m}\big).$$

Proof: The proof comprises the following steps.

1) Distance Function Gradient: Suppose $p_1(\boldsymbol{y}|\boldsymbol{u})$ and $p_2(\boldsymbol{y}|\boldsymbol{u})$ are two input-dependent density functions. Given distance function (10), we have by the chain rule

$$\frac{\partial}{\partial \boldsymbol{u}}\mathcal{D}(\mathbb{P}_1(\cdot),\mathbb{P}_2(\cdot)) = -\frac{1}{2}\int\left(\sqrt{\frac{p_2(\cdot)}{p_1(\cdot)}}\frac{\partial p_1(\cdot)}{\partial \boldsymbol{u}} + \sqrt{\frac{p_1(\cdot)}{p_2(\cdot)}}\frac{\partial p_2(\cdot)}{\partial \boldsymbol{u}}\right)d\boldsymbol{y}.$$

2) Output Perturbations: Suppose $\boldsymbol{y}(\boldsymbol{u})$ is the solution to (1). Then, given $(\boldsymbol{d}, K)$, we have

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{u}} = -\frac{\partial C(\boldsymbol{x})}{\partial \boldsymbol{x}}\left(\frac{\partial F}{\partial \boldsymbol{x}}\right)^{-1}\frac{\partial F}{\partial \boldsymbol{u}}.$$

Now the proof follows from observations 1) and 2), the description of the probability distribution (6) and (10). ■

We close this section with Algorithm 1 summarizing the proposed Bayesian approach comprising two pivotal steps of the input synthesis (9) and the posterior update rule (5).

IV. Leakage Localization in a WDN

In this section, the proposed methodology is further specified for the application to active leak localization in a WDN as described in Example 1. To this end, we first show how the mathematical model of WDNs falls into this category.

A. Model of a WDN

For the case of a WDN, the hypothesis candidates $\{1, \ldots, n_K\}$ correspond to possible leak locations at one of the n_K nodes. For a WDN at steady-state, the equations describing the state $\boldsymbol{x} = \begin{bmatrix}\boldsymbol{q} & \boldsymbol{h}\end{bmatrix}^T$ of the network under leak mode i can be represented by substitution of

$$F^{[i]}\big(\boldsymbol{x}^{[i]}, \boldsymbol{u}, \boldsymbol{d}, K^{[i]}\big) : \begin{cases} E^{[i]}\big(\boldsymbol{x}^{[i]}, \boldsymbol{u}, \boldsymbol{d}, \boldsymbol{g}^{[i]}\big) = 0 \\ \boldsymbol{g}^{[i]} = L\big(\boldsymbol{x}^{[i]}; K^{[i]}\big). \end{cases}$$

In (1), we have

$$\text{Model}^{[i]} : \begin{cases} E^{[i]}\big(\boldsymbol{x}^{[i]}, \boldsymbol{u}, \boldsymbol{d}, \boldsymbol{g}^{[i]}\big) = 0 \\ \boldsymbol{g}^{[i]} = L\big(\boldsymbol{x}^{[i]}; K^{[i]}\big) \\ \boldsymbol{y}^{[i]} = C\boldsymbol{x}^{[i]} \end{cases} \tag{12}$$

where following [24], the term $E^{[i]}(\cdot)$ takes the form:

$$\begin{aligned} E^{[i]}(\cdot) = &\begin{bmatrix} A_{11}\big(\boldsymbol{q}^{[i]}\big) & A_{12} \\ A_{12}^T & 0 \end{bmatrix}\begin{bmatrix}\boldsymbol{q}^{[i]} \\ \boldsymbol{h}^{[i]}\end{bmatrix} \\ &+ \begin{bmatrix} A_{10}\boldsymbol{u} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -\boldsymbol{d} - \boldsymbol{g}^{[i]} \end{bmatrix} = 0 \end{aligned} \tag{13}$$

with leak magnitude $\boldsymbol{g}^{[i]}$ in m^3/s being a function of the pressure $\boldsymbol{p}_i$ (part of the state vector $\boldsymbol{x}^{[i]}$) in mH_2O at the leak's location [25]. This yields

$$\boldsymbol{g}^{[i]} = L\big(\boldsymbol{x}^{[i]}; K^{[i]}\big) = K^{[i]}\boldsymbol{p}_i^{\alpha}, \quad \boldsymbol{p}_i = \rho_w g_c(\boldsymbol{h}_i - z_i). \tag{14}$$

The diagonal matrix $A_{11} \in \mathbb{R}^{n_p \times n_p}$ consists of the elements

$$A_{11}(j, j) = R_j \left|\boldsymbol{q}_j^{[i]}\right|^{\tau - 1}, \quad j = 1, \ldots, n_p$$

with R_j being the resistance coefficient of pipe j; see [26] for further information. The matrices $A_{12} \in \mathbb{R}^{n_p \times n_n}$ and $A_{10} \in \mathbb{R}^{n_p \times n_u}$ are the incidence matrices that denote the connectivity between the n_n unknown head nodes and the n_u nodes with regulated pressure heads $\boldsymbol{u} \in \mathcal{U}$, respectively. Using the Hazen–Williams (HW) energy loss model to describe the friction of pipes to flow, we have $\tau = 1.852$, $R_j = 10.670 L_j/(C_j^{1.852} D_j^{4.871})$ where L_j, C_j, and D_j denote the length in m, the unitless roughness coefficient, and the diameter in m of pipe j, respectively, [26]. The parameters describing the leak magnitude $(\rho_w, g_c, z_i, K_i, \alpha)$ denote the density of water in kg/m^3, gravity in m/s^2, elevation head z_i in mH_2O, and the discharge constant $K_i > 0$ and $0 < \alpha < 1$ are the parameters dependent on the leak size and pipe material. Finally, $\boldsymbol{y}^{[i]}$ denotes the measured part of state $\boldsymbol{x}^{[i]}$ with $C \in \mathbb{R}^{n_y \times n_x}$. With slight abuse of notation, we use $C\boldsymbol{x}$ instead of $C(\boldsymbol{x})$, because in this specific application $\boldsymbol{y}$ is a linear combination of $\boldsymbol{x}$.

B. Specifications

Thanks to the WDN's model built above, the proposed methodology can now be applied for active leak localization in a real WDN as described by (12). What complicates the application is that in a real-world setting the amount of leakage $\boldsymbol{g}$ is not known exactly. To mimic this situation, we therefore assume that only an estimate of $\boldsymbol{g}$ is available. For the M realizations needed to approximate the propagation of parametric uncertainties $(\boldsymbol{d}, \boldsymbol{g})$ into output distribution functions $\mathbb{P}(dy|i, \boldsymbol{u})$, $i \in \mathbb{H}$, it is assumed that the nodal demands are realizations $\hat{\boldsymbol{d}}_m$ from a Gaussian distribution $\hat{\mathbb{P}}_{\boldsymbol{d}}$. The availability of $\hat{\boldsymbol{g}}$ is mimicked by drawing a realization from the uniform distribution $\boldsymbol{g} \sim \mathcal{U}_{\boldsymbol{g}}(\boldsymbol{g}^-, \boldsymbol{g}^+)$ centered at the actual leak magnitude $\boldsymbol{g}$, i.e., $(\boldsymbol{g}^- + \boldsymbol{g}^+/2) = \boldsymbol{g}$. Finally, it is reasonably assumed that the parameters describing network characteristics such as pipe diameter D_j and roughness coefficient C_j are calibrated using historical data and are time-invariant within fault detection time scales.

C. Input Synthesis

Recall that synthesizing a feasible input signal $\boldsymbol{u}$ using the gradient ascent update rule as proposed in (9) iteratively maximizes the cost function $J(\boldsymbol{u})$ in (8). The key ingredient is the computation of $\nabla_{\boldsymbol{u}} J(u_t)$ which entails the behavior of WDN as described by (12). From Proposition 1, it follows that the only information needed to be able to compute (11) is to know how to compute $\nabla_{\boldsymbol{u}} \boldsymbol{y}$ and how to determine η.

Proposition 2 (WDN Cost Function Gradient): Suppose that $\boldsymbol{y}^{[i]}(\boldsymbol{u})$ is a realization of (12) under fault mode i. Then, the sensitivity of a realization with respect to the input u is a matrix $\mathbb{R}^{n_u \times n_y}$ defined as

$$\begin{aligned} \nabla_{\boldsymbol{u}} \boldsymbol{y}^{[i]} :&= \begin{bmatrix} \frac{\partial \boldsymbol{y}_1^{[i]}}{\partial \boldsymbol{u}_1} & \cdots & \frac{\partial \boldsymbol{y}_1^{[i]}}{\partial \boldsymbol{u}_{n_u}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \boldsymbol{y}_{n_y}^{[i]}}{\partial \boldsymbol{u}_1} & \cdots & \frac{\partial \boldsymbol{y}_{n_y}^{[i]}}{\partial \boldsymbol{u}_{n_u}} \end{bmatrix} = \left(\nabla_{\boldsymbol{g}} \boldsymbol{y}^{[i]} \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{h}_i}\right) \otimes \nabla_{\boldsymbol{u}} \boldsymbol{h}_i \\ &= \bar{K} C S_{\boldsymbol{g}}^{[i]} \otimes S_{\boldsymbol{u}}^{[i]}\big[n_p + i\big] \end{aligned}$$

where $\bar{K} = (\partial \boldsymbol{g}/\partial \boldsymbol{h}_i) > 0$, $\S_{\boldsymbol{g}}^{[i]} \in \mathbb{R}^{n_x}$, and $\S_{\boldsymbol{u}}^{[i]} \in \mathbb{R}^{n_x \times n_u}$ are the sensitivities of the state $\boldsymbol{x}$ with respect to leak magnitude $\boldsymbol{g}$ and the inputs $\boldsymbol{u}$, respectively, i.e.,

$$\S_{\boldsymbol{g}}^{[i]} := \left[\frac{\partial \boldsymbol{q}}{\partial \boldsymbol{g}} \ \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{g}}\right]^{[i],T}, \quad \S_{\boldsymbol{u}}^{[i]} := \big[\nabla_{\boldsymbol{u}} \boldsymbol{q} \ \nabla_{\boldsymbol{u}} \boldsymbol{h}\big]^{[i],T}. \tag{15}$$

Remark 1 (Measurement Input Sensitivity): Note that due to the Kronecker product, the term $\nabla_{\boldsymbol{u}} \boldsymbol{h}_i$ does not affect the direction of $\nabla_{\boldsymbol{u}} \hat{\boldsymbol{y}}_j^{[i]}$ but only its magnitude, which corresponds to the intuition that changing inputs at different locations has a different impact on h_i.

Proof: Application of the chain rule to a single element of $\nabla_{\boldsymbol{u}} \boldsymbol{y}^{[i]}$ yields: $(\partial \boldsymbol{y}_j^{[i]}/\partial \boldsymbol{u}_v) = (\partial \boldsymbol{y}_j^{[i]}/\partial \boldsymbol{g})(\partial \boldsymbol{g}/\partial \boldsymbol{h}_i)(\partial \boldsymbol{h}_i/\partial \boldsymbol{u}_v)$, which comprises of the following parts.

1)

$$\frac{\partial \boldsymbol{y}_j^{[i]}}{\partial \boldsymbol{g}} = C S_{\boldsymbol{g}}^{[i]}.$$

2) As we assume an underlying leak model of the form in (14), this unidentifiable derivative can be elaborated as

$$\bar{K} = \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{h}_i} = \rho_w g_c \alpha K_i (\rho_w g_c (h_i - z_i))^{\alpha - 1} > 0$$

3)

$$\frac{\partial \boldsymbol{h}_i}{\partial \boldsymbol{u}_v} = S_{\boldsymbol{u}}^{[i]}[n_p + i; v].$$

Now the proof follows from 1), 2) and 3). ■

D. Sensitivities

It is often claimed that the sensitivity matrix of the state with respect to leak magnitude $\boldsymbol{g}$ "is extremely difficult to calculate analytically" [11] because the nonlinear hydraulic equations in (13) are implicit. In this article, the sensitivities are computed analytically using the implicit function theorem, which reduces the computational complexity and makes the proposed approach tractable. In the following proposition, we address when the sensitivities required in (15) can be computed efficiently.

Proposition 3 (Analytical Description of $\S_{\boldsymbol{g}}^{[i]}$, $\S_{\boldsymbol{u}}^{[i]}$): Let $\boldsymbol{x}^*$ be a solution to algebraic equation (13) where the Jacobian $\partial F^{[i]}/\partial \boldsymbol{x}$ at $\boldsymbol{x}^*$ is invertible, i.e., the equation is nondegenerate. Then, the sensitivity matrices $\S_{\boldsymbol{g}}^{[i]}$ and $\S_{\boldsymbol{u}}^{[i]}$ in (15) can be calculated via

$$\begin{bmatrix} \S_{\boldsymbol{g}}^{[i]} & \S_{\boldsymbol{u}}^{[i]} \end{bmatrix} = \begin{bmatrix} \mathrm{diag}(\gamma_i) A_{11}(\hat{\boldsymbol{q}}^{[i]}) & A_{12} \\ A_{12}^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} B_{\boldsymbol{g}}^{[i]} & B_{\boldsymbol{u}}^{[i]} \end{bmatrix}.$$

Proof: Since the function $F^{[i]}(\cdot)$ is continuously differentiable around such a solution $\boldsymbol{x}^*$ [24, Appendix. 1], by the implicit function theorem [27, Th. A.2], we have

$$\frac{\partial F^{[i]}(\cdot)}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{g}} = -\frac{\partial F^{[i]}(\cdot)}{\partial \boldsymbol{g}}$$

$$\underbrace{\begin{bmatrix} \mathbb{N}_A A_{11}(\hat{\boldsymbol{q}}^{[i]}) & A_{12} \\ A_{12}^T & 0 \end{bmatrix}}_{A^{[i]}} \underbrace{\begin{bmatrix} \frac{\partial \boldsymbol{q}}{\partial \boldsymbol{g}} \\ \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{g}} \end{bmatrix}^{[i]}}_{\S_{\boldsymbol{g}}^{[i]}} = \underbrace{\begin{bmatrix} 0 \\ I^{[i]} \end{bmatrix}}_{B_{\boldsymbol{g}}^{[i]}}$$

where $\mathbb{N}_A = \mathrm{diag}(\gamma_i)$, $i = 1, \ldots, n_p$ and the only nonzero element of vector $I^{[i]} \in \mathbb{R}^{n_n}$ is $I^{[i]}(i) = 1$. Therefore, sensitivity $\S_{\boldsymbol{g}}^{[i]}$ can be computed by solving the n_x linear equations, i.e., $\S_{\boldsymbol{g}}^{[i]} = A^{[i]} \backslash B_{\boldsymbol{g}}^{[i]}$.
Likewise, at the same steady-state solution $\boldsymbol{x}^*$, we can write

$$\frac{\partial F^{[i]}(\cdot)}{\partial \boldsymbol{x}} \nabla_{\boldsymbol{u}} x = -\nabla_{\boldsymbol{u}} F^{[i]}(\cdot)$$
$$A^{[i]} \S_{\boldsymbol{u}}^{[i]} = B_{\boldsymbol{u}}^{[i]}$$

where $B_{\boldsymbol{u}}^{[i]} = \begin{bmatrix} A_{10} \\ 0 \end{bmatrix}$, and we obtain $\S_{\boldsymbol{u}}^{[i]} = A^{[i]} \backslash B_{\boldsymbol{u}}^{[i]}$. ■

Remark 2 (Computational complexity): The complexity of the proposed algorithm is determined by solving $n_n \times n_x \times (M+1+n_u)$ linear algebraic equations where n_n is the number of unknown nodes, n_x is the number of the states including all the flows and hydraulic heads, n_u is the number of inlets, and M is the number of realizations.

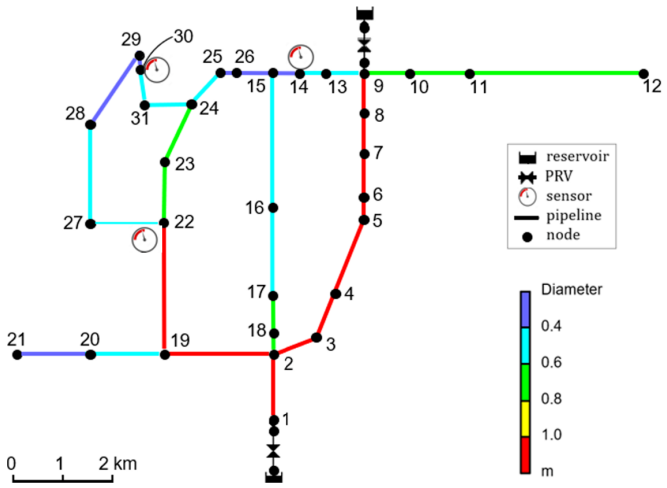

Fig. 1. Benchmark Hanoi trunk network.

V. Case Study

In this section, the proposed AFD method is numerically evaluated by comparing to a PFD method under different levels of demand uncertainty. It is assumed that no prior information is available about where approximately the leak is located, such that the initial belief is a uniform distribution over the nodes.

A. Hanoi Network

To assess Algorithm 1, its performance is tested on the benchmark Hanoi network [28] and compared against the state-of-the-art PFD method introduced in [13] with the slight difference that the output space is directly used here. This is implemented using Algorithm 1 without the active control rules, i.e., lines 2 and 3 are skipped. Fig. 1 shows a schematic representation of this network, which is fed by two reservoirs. This trunk model consists of 31 nodes connected by 33 pipes. At nodes 14, 22, and 30, pressure loggers are installed. The pressures at nodes 1 and 9 are controlled by means of PRVs, which will be referred to as inputs $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, respectively. To demonstrate the algorithm's ability to handle multiple inputs,the default Hanoi network from [28] is extended with an extra reservoir and PRV at node 9. The WDN contains two trees (i.e., acyclic, connected subgraphs of the network): 9-10-11-12 and 19-20-21. Leaks at these nodes of equal magnitude affect the pressure distribution across the looped part of the network identically and are therefore not isolable. Therefore, as is done in model reduction for WDNs [29], the nodes in these trees are grouped into corresponding sets and are represented by the root node of the tree nodes 9 and 19, respectively. This prevents the algorithm from getting stuck in an attempt to isolate nonisolable leaks.

The size of each time step during fault diagnosis is in the order of minutes, such that steady-state can be considered and dynamic subsecond processes can be neglected.

B. Simulation Setup

Since the real Hanoi network is not available to validate the proposed AFD algorithm, we resort to simulating the

TABLE I
LEAK MODEL CHARACTERISTICS

K	α	$\rho_w[kg/m^3]$	$g_c[m/s^2]$
0.0005	0.75	0.5	9.81

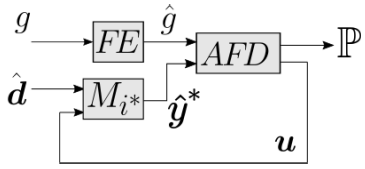

Fig. 2. Schematic representation of the simulation setup to mimic a real implementation of the AFD algorithm on the Hanoi WDN for a single time step.

hydraulics described via (1) for numerical purposes. To this end, a single leak scenario is investigated where the node i^* is the location of the leak whose level of the leakage is described by (14). In this setting, the leak parameters used in the simulation experiments are listed in Table I.

The obtained model M_{i^*} is used as if it were the real Hanoi network. Fig. 2 shows a schematic representation of how inputs and outputs of the "real" network are imitated and how these interact with the AFD method. The consumer demands $\hat{\boldsymbol{d}}_t$ projected at the nodes of the "real" (M_{i^*}) network are determined by

$$\boldsymbol{d}_{k,j} = \max\big(\boldsymbol{\xi}_d, 0\big)\mu_t b_j \quad \forall j \in \mathcal{I}_n \tag{16}$$

where b_j is the base demand of the jth node, μ_t is the demand multiplier at time step t, and $\boldsymbol{\xi}_d$ is randomly drawn from normal distribution $\boldsymbol{\xi}_{\boldsymbol{d}} \sim \mathcal{N}(1, \sigma_{\boldsymbol{d}})$ [30].

The output at each time step is generated by solving the nonlinear hydraulic equation of M_{i^*} and obtaining its output $\tilde{\boldsymbol{y}}_t$. The fault estimation (FE) block in Fig. 2 models the magnitude estimation of leak $\boldsymbol{g}$ and predicts its value according to $\hat{\boldsymbol{g}}_t = \boldsymbol{g}_t + \boldsymbol{\xi}_{\boldsymbol{g}}$ where $\boldsymbol{\xi}_{\boldsymbol{g}} \sim \mathcal{N}(0, \sigma_{\boldsymbol{g}})$. The AFD block takes as input the estimated leak magnitude $\hat{\boldsymbol{g}}_t$ and the output observation $\tilde{\boldsymbol{y}}_t$, and based on these updates the input $\boldsymbol{u}_t$ and likelihood vector $\mathbb{P}_t$. Within the AFD block, the leak magnitude probability distribution function (PDF) $\mathbb{P}_{\boldsymbol{g}}$ and output PDF $\mathbb{P}_{\boldsymbol{y}}$ are estimated at each time step t and used for execution of algorithm 1. The leak magnitude PDF is assumed to be uniformly distributed around the estimated leak magnitude, i.e., $\boldsymbol{g} \sim \mathcal{U}(\hat{\boldsymbol{g}} - 3\sigma_{\boldsymbol{g}}, \hat{\boldsymbol{g}} + 3\sigma_{\boldsymbol{g}})$. The output PDF is estimated as described in Section IV. For this setup, a Gaussian kernel was used with bandwidth determination based on Scott's rule [31]. The control inputs have the upper bound $\boldsymbol{u}_{\max}$ and are lower bounded by the required minimal service level pressure head of 15 m at the critical point, node 30 in this case being the node in the lowest pressure area of the network [9]. In any case, the pressure at this critical point needs to be maintained above its critical level. Therefore, depending on the scenario, the inlet pressure heads $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ have a different lower bound per steady-state. This is the reason that the inputs in Fig. 3(b) (passive method) are not straight lines. Furthermore, a constraint $\Delta_{\boldsymbol{u},\max}$ is imposed on the input change rate due to PRV's characteristics. The leak magnitude is bounded by about $[0.02, 0.05]$ m^3/s $\Leftrightarrow$ $[1.0, 2.5\%]$ of the mean total demand between 0 A.M. and 5 A.M. All the necessary simulation

TABLE II
SIMULATION CONSTANTS

M	ζ	$\boldsymbol{u}_{max}[mH_2O]$	$p_{\max}$
80	$5/10^4$	100	0.95
$\Delta_{\boldsymbol{u},max}[mH_2O]$	η	$\sigma_{\boldsymbol{g}}[m^3/s]$	t_{max}
5	50	0.003	60

constants are specified in Table II. Since the amount of computing power required to perform the simulations depends on M, its value has been minimized. Lower values for M make that the PDF estimates are filled with gaps, such that the actual output space is not covered. Likewise, the stepsize η is a design parameter and has been optimized in such a fashion that the gradient in (9) controls $\Delta_{\boldsymbol{u}} = \eta(\partial J/\partial u)$. Taking a too large η makes that always $\Delta_{\boldsymbol{u}} = \Delta_{\boldsymbol{u},\max}$, such that the gradient is in fact out of play. On the other hand when η goes to zero, the active algorithm becomes passive. The simulation experiments were performed using the WNTR Python package [32] and its built-in hydraulic solver.

C. Experiments and Scenarios

The following numerical experiments are set up and performed five times for scenarios with different nodal demand realizations.

1) i^* is varied over all 26 considered leak locations, i.e., all the classes of nodes where trees are aggregated into the corresponding root nodes. The AFD algorithm is directly compared with its PFD counterpart. The two algorithms have identical initial conditions and are activated during nighttime between 0 A.M. and 5 A.M. with a time step of 5 min. The algorithms try to isolate the leak location within this time frame.
2) Step 1) is repeated for two different values of the demand distribution variance $\sigma_{\boldsymbol{d}}$, namely, $\sigma_{\boldsymbol{d}} \in [0.01, 0.10]$.

To measure the performance of the AFD and PFD methods, accuracy and average distance are used as metrics. Accuracy is measured by the percentage of leaks that are classified into correct leak location within the time frame of 0 A.M. and 5 A.M. The average distance is the mean distance in kilometers between the "as classified" and the actual leak location i^*, calculated using Dijkstra's algorithm [33].

Remark 3: The experiments simulate fault diagnosis at nighttime, because the ratio between leakage and total inflow at the inlets is the largest in these hours [34]. Nodal consumer demands are more predictable as well, i.e., have a lower variance, such that leaks to a lesser extend are getting obscured by the increased uncertainty imposed by a higher variance [21].

D. Results

Fig. 3(a) and (b) shows the input and likelihood trajectories resulting from a single AFD and PFD diagnosis with a leak at node 1, i.e., $i^* = 1$, and demand distribution variance $\sigma_{\boldsymbol{d}} = 0.10$. Fig. 4(a) and (b) shows similar trajectories for a leak at node 6, i.e., $i^* = 6$, under equal conditions. The spatial distribution of the nodes is shown in Fig. 1. All the scenarios

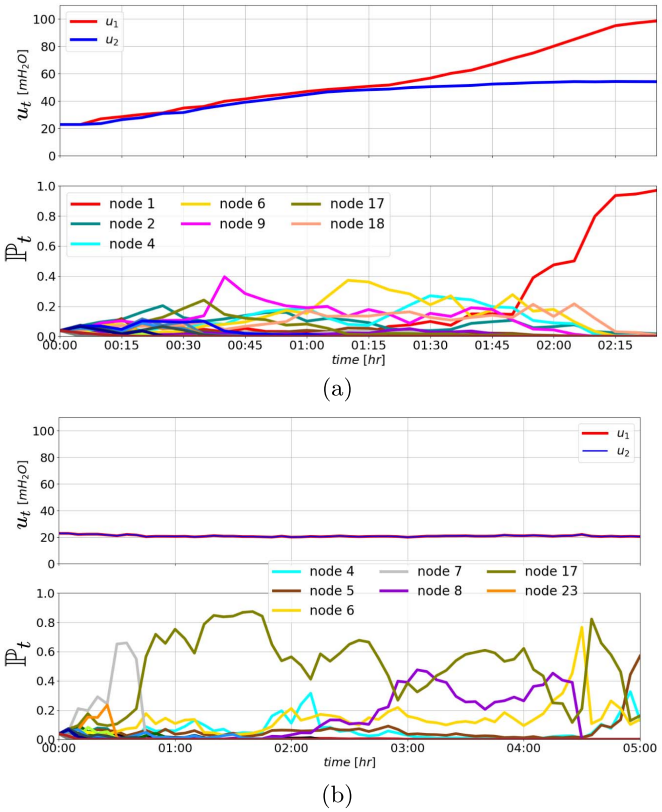

Fig. 3. Input and likelihood trajectories for a single experiment with initially equal scenarios resulting from (a) AFD and (b) PFD diagnosis ($i^* = 1$).

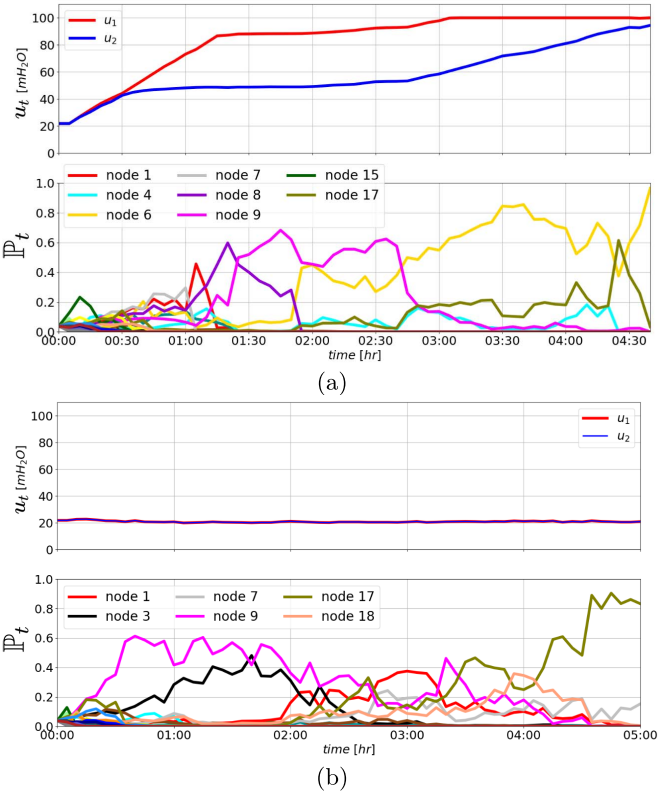

Fig. 4. Input and likelihood trajectories for a single experiment with initially equal scenarios resulting from (a) AFD and (b) PFD diagnosis ($i^* = 6$).

are equally initiated regarding demands and leak magnitude estimation.

In both the cases ($i^* = 1$, $i^* = 6$) the AFD trajectories can be roughly divided into two periods: 1) leak area selection and 2) isolation of the most likely leak node or location, described below.

1) The pressures in the network are low and many hypotheses are initially posed. The algorithm aims to fan-out all the corresponding densely packed output PDFs by increasing the control inputs, sometimes at the maximum allowed rate. In this way, it aims to stepwise maximize the objective function $J(\cdot)$ in (8).
2) In this period, many hypotheses have been rejected based on a sequence of output measurements, and the algorithm focuses on separating the output PDFs of the remaining hypotheses. It is observed that the inputs diverge due to their different effect on separating the remaining output PDFs, i.e., $(\partial J/\partial \boldsymbol{u}_1)$ and $(\partial J/\partial \boldsymbol{u}_2)$ diverge from each other. As the likelihood vector $\mathbb{P}_t$ changes over time (line 7 in Algorithm 1), the varying dominant hypotheses determine the stagnation or increase of the different inputs. This clearly demonstrates the adaptive behavior of the AFD algorithm, which determines the input directions based on the "current belief." When the hypotheses of two spatially closely related nodes have both a high likelihood, it is observed that the inputs tend to grow faster due to the relative high overlap between their output PDFs. It can be observed that as a rule of thumb, when output PDFs of the most likely hypotheses at a certain time step have little overlap, the inputs stagnate. In contrast, when the dominant hypotheses have a high overlap of output PDFs in the full output space, the input vector takes a step in the direction in which maximum separation of the output PDFs corresponding to the dominant hypotheses is expected, that is, maximizing J with respect to $\boldsymbol{u}$.

Comparing the trajectories of $\mathbb{P}_t$ between the PFD and AFD in these two experiments shows that the PFD method has a poorer performance in terms of speed, accuracy, and decisiveness. The different hypotheses struggle for precedence, but their corresponding output PDFs clearly have too much overlap which makes that the PFD algorithm has no further region selection. Of course this does not mean that the AFD algorithm is superior in every experiment. Both the algorithms are heavily dependent on the "separating quality" of observed outputs at each time step. To show that the AFD algorithm has an overall better performance than its PFD counterpart, in Fig. 5 the accuracy and average distance time-lapse trajectories over all 130 scenarios (26 experiments repeated with five demand realizations) are plotted for the two different levels of demand uncertainty (different values of σ_d).

The upper plot shows how the accuracy evolves over diagnosis time for different values of demand distribution variance σ_d. Likewise, the lower plot shows the development over time for the average "distance to actual leak" performance metric in kilometers. Compared with the PFD algorithm, the accuracy of the AFD algorithm is higher by 12% and 9% for the demand

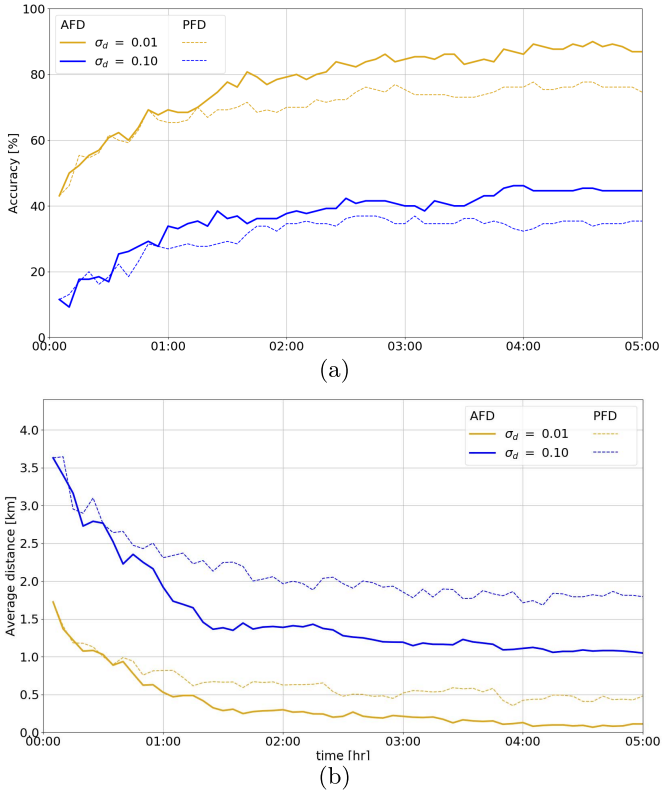

Fig. 5. Performance of the AFD and PFD methods over a 5-h diagnosis period. The performance measures are averaged over all the experiments, where leaks are placed at different nodes in each experiment. Two scenarios with different levels of demand uncertainty are considered. (a) Accuracy in finding exact leak nodes (%) and (b) average distance of nodes classified as leak location to actual leak nodes (km).

variance values of 0.01 and 0.1, respectively. Similarly, the average distance is lower for AFD by, respectively, 0.37 and 0.75 km. Finally, the mean diagnosis time for AFD is 117 and 164 min, whereas the PFD takes 45 and 50 min longer, respectively.

E. Robustness of the Proposed AFD Algorithm

Algorithm 1 depends on a number of parameter choices as inputs, whose values are presented in Table II. Some of these are hyperparameters as their value controls the Bayesian learning process of the algorithm. Here, we investigate the sensitivity of the AFD's performance to some of these parameter choices.

1) Hyperparameter M: The parameter M denotes the number of realizations sampled to construct probability density functions corresponding to the different hypotheses as in (6). As such it affects how well we estimate the output distributions at each iteration, and therefore the overall performance of the algorithm. We performed closed-loop simulations of our AFD algorithm for hyperparameter M; for each M in $\{10, 20, 40, 60, 80, 100, 120\}$, the algorithm's performance was tested using different performance metrics. Table III shows performance for the diagnosis with respect to M, while the other parameters in Table II were fixed. As an example, a leak at node 3 is simulated and the algorithm is tested with 15 closed-loop experiments to assess its average performance and variances in performance; 15 experiments were found, *a posteriori*, to be sufficient to show convergence in these performance metrics shown in Table III.

TABLE III

SENSITIVITY ANALYSIS FOR PARAMETER M. AS AN EXAMPLE, A LEAK AT NODE 3 IS SIMULATED, WITH 15 CLOSED-LOOP EXPERIMENTS TO ASSESS THE ALGORITHM'S AVERAGE PERFORMANCE AND VARIANCES IN PERFORMANCE; 15 EXPERIMENTS WERE FOUND TO BE SUFFICIENT, *a Posteriori*, TO GET CONVERGENCE FOR THESE PERFORMANCE METRICS

M	10	20	40	60	80	100	120
Accuracy [%]	27	13	7	11	31	15	25
Average distance [km]	2.2	1.7	1.4	1.0	1.1	1.2	1.1
$SD_{distance}$ [km]	2.0	1.3	0.9	0.7	0.9	0.6	0.6

From this example analysis in Table III, we can observe the following regarding the sensitivity of performance to M.

1) The average distance between the node classified as leaky and the actual leaky node i^* drops with increasing M. After $M \geqslant 60$, it stagnates at ~ 1 km distance for this network example. The variance $\text{Var}_{\text{distance}}$ of this metric also saturates beyond $M = 80$.
2) When M is increased, the algorithms takes more iterations to reject hypotheses, because the overlap of output PDF's increases with M. The result is that the algorithm converges slower with increasing M.
3) When M is small (eg. $M = 10$), the algorithm converges very fast and its results can therefore be qualified as quick guesses with outliers far outside the neighborhood of i^*.
4) M largely affects the computation speed because the WDN needs to be simulated M times, at each iteration of the AFD algorithm and for each hypothesis with nonzero "belief." However, it is also important to emphasize that these network simulations between hypothesis and control updates are fully parallelizable (for loop in lines 5–8 of Algorithm 1); network simulations can be done in parallel across the different hypotheses with nonzero "belief" and across the M samples made for each hypothesis.

From similar simulations over many leaks, we can conclude that the proposed AFD algorithm was found to be robust to this hyperparameter values since a sufficiently large M value could be found to ascertain good performance in terms of distance to actual leak and with low variance of this performance. For the network example in this article, and as also depicted in Table III, $M = 80$ (the value selected for the experiments in Table II) gives a good trade-off between computational burden and performance.

2) Other (Hyper)Parameters: As shown in Fig. 5, the uncertainty of nodal demands σ_d has a big impact on the performance of the algorithm as it accounts for the uncertainty within the system under normal operations, even without a leak. Unlike for M, σ_d has little impact on the algorithm convergence rate but rather does affect its ability to find the accurate leak location or proximity to it. As diurnal demand uncertainty becomes larger, the impact of a leak on the measured output variable $\widehat{\boldsymbol{y}}^*$ falls within normal operations, and therefore its identification becomes less accurate and less precise in distance to actual leak.

Other parameters in Table II did not affect the algorithm performance or were controlled design or system parameters. These are as follows.

1) ζ: This hyperparameter can always be set sufficiently close to zero and could be controlled to have no influence on the algorithm's performance. It has an influence on the computational speed, since hypotheses that have a belief lower than zeta get set to zero, after which the belief vector is renormalized. This is because at each iteration, the algorithm has to construct the PDF's of all unrejected hypotheses. However, this burden is fully parallelizable and so can be mitigated with parallel computational resources.
2) $\mathbf{u}_{\max}$: this parameter comes from regulatory constraints on system pressure and only sets a maximum on the input vector, such that the pressure in a physical WDN does not exceed its maximum allowable pressure.
3) $p_{\max}$: this parameter is used as a stopping criterion; whenever the belief of a single hypothesis exceeds $p_{\max}$, set to 0.95 here; the algorithm terminates and qualifies that node to be the leaky one.
4) $\Delta_{\mathbf{u},\max}$: this parameter limits the stepsize in u in a single control time step. This often comes from pressure control valve operational constraints.
5) η: this hyperparameter determines how active the algorithm is. When chosen close to zero, the active algorithm converges to its passive counterpart. So it does not directly make the algorithm robust but rather controls it to be more or less active.
6) σ_g: this is a design parameter for the experiments.
7) $t_{\max}$: this is a design parameter for the experiments.

VI. Conclusion and Future Directions

A tractable active fault isolation method is proposed for a class of nonlinear models subject to faults and applied to locate leaks in a WDN with uncertain user demands and unknown leak magnitude. The method relies on the classification of output observations to a discrete set of hypotheses. The uncertainties are captured by output PDFs which are used to iteratively update the posterior probability of each hypothesis in a Bayesian framework. The AFD algorithm proactively minimizes the joint overlap between output PDFs by designing optimal control inputs. A new numerically scalable approach for synthesizing such control inputs on the fly is derived. The performance is tested for two levels of demand uncertainty and compared with the PFD counterpart method. Improvements of the performance metrics accuracy and average distance as well as diagnosis speed are observed. It can be concluded that the AFD method is more reliable and faster compared with its state-of-the-art PFD counterpart. It is further shown that the AFD algorithm updates the inputs in an economical way, i.e., the inputs are only adjusted when this is in favor of the objective. The robustness of the AFD algorithm was also tested, showing that the hyperparameter values could be selected appropriately to guarantee good performance. The number of output realizations of the system, sampled to estimate the output probability density functions corresponding to the different hypotheses, was shown as the main hyperparameter that affects the performance and computational burden. We note that the number of system simulations required at each iteration, which grows linearly with number of realizations sampled and number of hypothesis not yet rejected, is fully parallelizable and may not be burdensome even for large number of realizations sampled. Future follow-up studies are encouraged to study optimal sensor and input placement to facilitate AFD.

References

[1] A. Gupta and K. D. Kulat, "A selective literature review on leak management techniques for water distribution system," *Water Resour. Manage.*, vol. 32, no. 10, pp. 3247–3269, Aug. 2018.

[2] OECD. (2016). *Water Governance in Cities*. [Online]. Available: https://www.oecd-ilibrary.org/content/publication/9789264251090-en

[3] T. Sakomoto, M. Lutaaya, and E. Abraham, "Managing water quality in intermittent supply systems: The case of Mukono town, Uganda," *Water*, vol. 12, no. 3, p. 806, Mar. 2020.

[4] T. T. T. Zan, H. B. Lim, K.-J. Wong, A. J. Whittle, and B.-S. Lee, "Event detection and localization in urban water distribution network," *IEEE Sensors J.*, vol. 14, no. 12, pp. 4134–4142, Dec. 2014.

[5] Q. Xu, R. Liu, Q. Chen, and R. Li, "Review on water leakage control in distribution networks and the associated environmental benefits," *J. Environ. Sci.*, vol. 26, no. 5, pp. 955–961, May 2014.

[6] M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki, *Diagnosis and Fault-Tolerant Control*, 3rd ed. Berlin, Germany: Springer-Verlag, 2015.

[7] Z. Wu, "Innovative optimization model for water distribution leakage detection," Bentley Syst., Watertown, NY, USA, Tech. Rep., Sep. 2008, pp. 1–8. Accessed: May 15, 2021. [Online]. Available: http://aquageo.es/wp-content/uploads/2012/10/INNOVATIVE-OPTIMIZATION-MODEL-FOR-WATER-DISTRIBUTION-LEAKAGE-DETECTION.pdf

[8] W. Li *et al.*, "Development of systems for detection, early warning, and control of pipeline leakage in drinking water distribution: A case study," *J. Environ. Sci.*, vol. 23, no. 11, pp. 1816–1822, Nov. 2011.

[9] R. Wright, E. Abraham, P. Parpas, and I. Stoianov, "Control of water distribution networks with dynamic DMA topology using strictly feasible sequential convex programming," *Water Resour. Res.*, vol. 51, no. 12, pp. 9925–9941, Dec. 2015.

[10] R. Puust, Z. Kapelan, D. A. Savic, and T. Koppel, "A review of methods for leakage management in pipe networks," *Urban Water J.*, vol. 7, no. 1, pp. 25–45, 2010.

[11] R. Perez *et al.*, "Leak localization in water networks: A model-based methodology using pressure sensors applied to a real network in Barcelona [applications of control]," *IEEE Control Syst.*, vol. 34, no. 4, pp. 24–36, Aug. 2014.

[12] D. Wachla, P. Przystalka, and W. Moczulski, "A method of leakage location in water distribution networks using artificial neuro-fuzzy system," *IFAC-PapersOnLine*, vol. 48, no. 21, pp. 1216–1223, 2015.

[13] A. Soldevila *et al.*, "Leak localization in water distribution networks using Bayesian classifiers," *J. Process Control*, vol. 55, pp. 1–9, Jul. 2017.

[14] G. Romero-Tapia, M. J. Fuente, and V. Puig, "Leak localization in water distribution networks using Fisher discriminant analysis," *IFAC-PapersOnLine*, vol. 51, no. 24, pp. 929–934, 2018.

[15] A. Soldevila, J. Blesa, T. N. Jensen, S. Tornil-Sin, R. M. Fernández-Cantí, and V. Puig, "Leak localization method for water-distribution networks using a data-driven model and Dempster–Shafer reasoning," *IEEE Trans. Control Syst. Technol.*, vol. 29, no. 3, pp. 937–948, May 2020.

[16] A. Soldevila, G. Boracchi, M. Roveri, S. Tornil-Sin, and V. Puig, "Leak detection and localization in water distribution networks by combining expert knowledge and data-driven models," *Neural Comput. Appl.*, vol. 34, no. 6, pp. 4759–4779, Mar. 2022.

[17] P. Irofti, L. Romero-Ben, F. Stoican, and V. Puig, "Data-driven leak localization in water distribution networks via dictionary learning and graph-based interpolation," 2021, *arXiv:2110.06372*.

[18] D. B. Steffelbauer, J. Deuerlein, D. Gilbert, E. Abraham, and O. Piller, "Pressure-leak duality for leak detection and localization in water distribution systems," *J. Water Resour. Planning Manage.*, vol. 148, no. 3, Mar. 2022, Art. no. 04021106.

[19] T. A. N. Heirung and A. Mesbah, "Input design for active fault diagnosis," *Annu. Rev. Control*, vol. 47, pp. 35–50, Mar. 2019.

[20] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. London, U.K.: Chapman & Hall, 1986.

[21] R. Gomes, A. Sá Marques, and J. Sousa, "Estimation of the benefits yielded by pressure management in water distribution systems," *Urban Water J.*, vol. 8, no. 2, pp. 65–77, Apr. 2011.
[22] J. Shao, *Mathematical Statistics*. Berlin, Germany: Springer, 2003.
[23] J.-M. Marin, P. Pudlo, C. P. Robert, and R. J. Ryder, "Approximate Bayesian computational methods," *Statist. Comput.*, vol. 22, no. 6, pp. 1167–1180, Nov. 2012.
[24] E. Abraham and I. Stoianov, "Sparse null space algorithms for hydraulic analysis of large-scale water supply networks," *J. Hydraulic Eng.*, vol. 142, no. 3, p. 99, Mar. 2016.
[25] J. Schwaller and J. E. van Zyl, "Modeling the pressure-leakage response of water distribution systems based on individual leak behavior," *J. Hydraulic Eng.*, vol. 141, no. 5, May 2015, Art. no. 4014089.
[26] E. Todini and S. Pilati, "A gradient algorithm for the analysis of pipe networks," in *Computer Applications in Water Supply* (System Analysis and Simulation), vol. 1. London, U.K.: Wiley, 1988, pp. 1–20.
[27] J. Nocedal and S. Wright, *Numerical Optimization*. Berlin, Germany: Springer, 2006.
[28] O. Fujiwara and D. B. Khang, "A two-phase decomposition method for optimal design of looped water distribution networks," *Water Resour. Res.*, vol. 26, no. 4, pp. 539–549, Apr. 1990.
[29] F. Pecci, E. Abraham, and I. Stoianov, "Model reduction and outer approximation for optimizing the placement of control valves in complex water networks," *J. Water Resour. Planning Manage.*, vol. 145, no. 5, May 2019, Art. no. 04019014.
[30] D. Hart *et al.*, "Quantifying hydraulic and water quality uncertainty to inform sampling of drinking water distribution systems," *J. Water Resour. Planning Manage.*, vol. 145, no. 1, Jan. 2019, Art. no. 04018084.
[31] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*. Hoboken, NJ, USA: Wiley, 2015.
[32] K. A. Klise, R. Murray, and T. Haxton, "An overview of the water network tool for resilience (WNTR)," in *Proc. WDSA/CCWI Joint Conf.*, Harvard, vol. 1, 2018.
[33] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numer. Math.*, vol. 1, no. 1, pp. 269–271, Oct. 1959.
[34] J. Thornton, R. Sturm, and G. Kunkel, *Water Loss Control Manual*. New York, NY, USA: McGraw-Hill, 2002.

Gert van Lagen received the B.Sc. degree in mechanical engineering and the M.Sc. degree in systems and control from the Delft University of Technology, Delft, The Netherlands, in 2020.

His technical interests are in control systems engineering, robotics, and complex problem solving in the broadest sense.

Edo Abraham received the M.Eng. degree in electrical and electronic engineering and the Ph.D. degree in control engineering (aeronautics) from the Imperial College London, London, U.K., in 2008 and 2013, respectively.

From 2014 to 2016, he was a Research Associate with the Environmental and Water Resources Engineering Group, Imperial College London. Since 2016, he has been an Assistant Professor in water and control with the Faculty of Civil Engineering and Geosciences, Delft University of Technology (TU Delft), Delft, The Netherlands. His research focuses are mainly in the application of mathematical optimization, control, and systems theory to advance water management and environmental engineering and their nexuses with energy and agricultural systems. His application areas span from control open canal water systems, water distribution networks, and irrigation to infrastructure planning in energy transition.

Dr. Abraham currently serves as an Associate Editor for the journal *Energy Systems* (Springer) and the *Journal of Hydroinformatics*.

Peyman Mohajerin Esfahani received the B.Sc. and M.Sc. degrees in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2005 and 2007, respectively, and the Ph.D. degree in control from ETH Zürich, Zürich, Switzerland, in 2014.

He joined the Delft University of Technology (TU Delft), Delft, The Netherlands, in October 2016, as an Assistant Professor, where he is currently an Associate Professor with the Delft Center for Systems and Control and the Co-Director of the Delft-AI Energy Laboratory. Prior to joining TU Delft, he held several research appointments at EPFL, Lausanne, Switzerland, ETH Zürich, and the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, from 2014 to 2016. His research interests include theoretical and practical aspects of decision-making problems in uncertain and dynamic environments, with applications to control and security of large-scale and distributed systems.

Dr. Mohajerin Esfahani was one of the three finalists for the Young Researcher Prize in Continuous Optimization awarded by the Mathematical Optimization Society in 2016, and a recipient of the 2016 George S. Axelby Outstanding Paper Award from the IEEE Control Systems Society. He received the ERC Starting Grant and the INFORMS Frederick W. Lanchester Prize in 2020. He was a recipient of the 2022 European Control Award. He currently serves as an Associate Editor for *Operations Research*, *Open Journal of Mathematical Optimization*, and IEEE TRANSACTIONS ON AUTOMATIC CONTROL.