2018 European Control Conference (ECC)
June 12-15, 2018. Limassol, Cyprus

Interval State Estimation of Hydraulics in Water Distribution Networks

S.G. Vrachimis, S. Timotheou, D.G. Eliades and M.M. Polycarpou

Abstract— **Water distribution networks are critical infrastructure systems, which are required to reliably deliver water to consumers. Hydraulic state estimation of flows and pressures is key for the efficient and dependable operation of water systems. However, in practice, hydraulic state estimation is a challenging task due to the scarcity of sensors compared to system states and the presence of several modeling uncertainties. It is a common practice for historical data to be used in the place of missing measurements, however the lack of a statistical characterization for the error of these measurements results in having no knowledge of the state estimate error magnitude. This paper presents a methodology for generating hydraulic state bounding estimates by modeling measurement and parametric uncertainties as intervals. The resulting nonlinear interval hydraulic equations are re-formulated using bounding linearization, a technique that restricts the nonlinearities within a convex set, and solved using linear optimization. An iterative procedure improves the bounding linearization, converging to the tightest possible bounds. Simulation results demonstrate that the proposed methodology produces tight state bounds and can be successfully used to detect water leakages in the network.**

I. Introduction

A Water Distribution Network (WDN) is responsible for delivering water to consumers from water sources. It comprises of pipes, storage tanks, reservoirs, water consumption points and hydraulic control elements such as pumps and valves. The water industry increasingly requires the estimation of WDN hydraulic state variables, such as water flows and pressures, in order to operate water systems efficiently, provide better customer service and assess the system behavior in order to detect and isolate water leaks or other emergency events. State estimation is enabled by gathering sensor measurements of flows and pressures at certain locations of the network through a Supervisory Control and Data Acquisition (SCADA) computer system. Then, using a mathematical model of the network, the state at all locations is estimated.

However, this is a challenging task due to the complexity and large area covered by water networks. Water outflow due to consumer demands is difficult to be measured accurately, as this would require a smart water meter at every residence. Thus, sensor measurements are scarce and the state estimation problem is under-determined. A common practice in WDN is to skeletonize the network by treating a group of consumers as a single demand point. It is then possible to use *pseudo-measurements*, which are demand estimates determined from population densities and historical data, to obtain an observable system configuration for state estimation [1]. Furthermore, modelling uncertainty is a cause of serious estimation errors. Pipe parameters, such as pipe roughness coefficients, are rarely known accurately. Even with an observable sensor configuration, model calibration is required a priori or during state estimation for the procedure to produce feasible solutions [2], [3].

The assumption of a known statistical characterization of sensor measurement error can lead to a serious miscalculation of the state estimation error in WDN. This is due to the use of pseudo-measurements, which do not have a statistical characterization and in the best case their estimation error is defined by an upper and lower bound. The case is similar for pipe parameters (e.g. length, diameter, roughness coefficient), for which the most accurate description that can be given for the error of the parameter value is an upper and lower bound.

An alternative approach uses bounds for the representation of measurement and model parameter uncertainty. In contrast to point state estimation methods, the use of bounded uncertainty can provide upper and lower bounds on the state variables. This method is referred to as *interval state estimation*. In many applications, such as leakage detection and contamination detection, the derivation of a range of possible values for the state of the WDN provides useful information for event and fault detection methodologies. Hydraulic state bounds can be used to generate bounds on chlorine concentration in the water network or other chemicals in the water, by taking into consideration the uncertainty on decay rate [4].

The use of bounds for the representation of measurement uncertainty and their incorporation into the state estimation cost function for WDN was introduced in [5]. This idea was developed in [6] as the *set-bounded* state estimation problem. The process of calculating uncertainty bounds for state estimates caused by inaccuracies of input data is referred to as Confidence Limit Analysis and it was solved using different approaches, including Neural Networks [7], the Error Maximization method [8], the Ellipsoid method and Linear Programming [9]. These methods have the disadvantage of using a linear approximation of the water network model, which does not guarantee that the calculated bounds contain all possible solutions based on the uncertainty. Monte-Carlo simulations use the nonlinear model to obtain uncertainty bounds by randomly generating and evaluating a large num-

This work was supported by the European Union Horizon 2020 programme under grant agreement no. 739551 (KIOS CoE).

S.G. Vrachimis, S. Timotheou, D.G. Eliades, M.M. Polycarpou are with the KIOS Research and Innovation Center of Excellence, Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus {vrachimis.stelios, timotheou.stelios, eldemet, mpolycar}@ucy.ac.cy

978-3-9524-2699-9 ©2018 EUCA

ber of parameter sets or realizations [10]. This approach is computationally intensive and even with a large number of simulations some extreme cases may not be covered.

The main contribution of this work is the design of a methodology to produce tight hydraulic state bounds, by considering parameter and measurement uncertainties in the network. In contrast to other state bounding techniques, the proposed methodology calculates the bounds on state estimates using the nonlinear form of the network equations. This is achieved using bounding linearization, a technique which restricts the nonlinearities within a convex set, thus converting the hydraulic equations in a form where the minimum and maximum of each state can be found using linear optimization. Then, an iterative procedure is used to minimize the distance between state upper and lower bounds, by improving the bounding linearization at each step and converging to the tightest possible bounds.

This paper is organized as follows: Section II formulates the problem of hydraulic state estimation in WDN where the uncertainty on parameters and measurements is represented by intervals. Section III presents a methodology to solve this problem based on the Iterative Hydraulic Interval State Estimation (IHISE) algorithm. In Section IV this methodology is applied on benchmark water networks and its performance is assessed. Finally, we discuss the application of this method on issues concerning WDN, with an example of detecting water leakages.

II. Problem formulation

The topology of a WDN is described by a directed graph denoted as $\mathcal{G} = (\mathcal{N}, \mathcal{L})$. Let $\mathcal{N} = \{1, \cdots, n_n\}$ be the set of all nodes, where $|\mathcal{N}| = n_n$ is the total number of nodes. These represent junctions of pipes, consumer water demand locations, reservoirs and tanks. The unknown quantity associated with nodes is the *hydraulic head*, which is indicated by h_i, and is a specific measure of water pressure above a geodetic datum. Some nodes, such as reservoirs, have a known constant hydraulic head. Water tanks that have level sensors installed, can also be considered as nodes with known head. We define the set of nodes with unknown head $\mathcal{N}_u = \{1, \cdots, n_u\}$, where $|\mathcal{N}_u| = n_u$ is the number of nodes with unknown head. The set of nodes with known head is defined as $\mathcal{N}_h = \{n_u + 1, \cdots, n_n\}$, where $|\mathcal{N}_h| = n_h$ is the number of nodes with known head and $\mathcal{N} = \mathcal{N}_u \cup \mathcal{N}_h$. Each node with unknown head is associated with a water consumer demand at the node location, denoted by $q_{ext,i}$.

Let $\mathcal{L} = \{1, \cdots, n_l\}$ be the set of links, where $|\mathcal{L}| = n_l$ is the total number of links. These represent network pipes, water pumps and pipe valves, with the last two being the main hydraulic control elements in a water network. In this work we will consider networks that only have 'open/close' valves, which can be represented by changing the graph's topology. We define the set of links that represent pipes as $\mathcal{L}_p = \{1, \cdots, n_p\}$, where $|\mathcal{L}_p| = n_p$ is the total number of pipes. We also define the set of links that represent pumps as $\mathcal{L}_{pu} = \{n_p + 1, \cdots, n_l\}$, where $|\mathcal{L}_{pu}| = n_{pu}$ is the total number of pumps. The unknown quantity associated with links are the *water flows*, indicated by q_i.

A. Formulation of hydraulic equations

In this work we assume measurements are available at every discrete time step k for all nodal demand outflows and for all tank and reservoir levels. This sensor configuration guarantees the topological observability of the network [11]. The vector of measured external water demand outflow is indicated by $\bar{q}_{ext}(k) \in \mathbb{R}^{n_u}$. The known head vector, which results from tank and reservoir level measurements, is indicated by $\bar{h}_{ext}(k) \in \mathbb{R}^{n_l}$. The unknown state of the network are the water flows in pipes $\bar{q}(k) \in \mathbb{R}^{n_l}$ and the unknown hydraulic heads at nodes $\bar{h}(k) \in \mathbb{R}^{n_u}$. The state is calculated using a hydraulic model of a WDN, which is a set of equations derived from the laws of *1)* conservation of energy and *2)* conservation of mass in the network.

1) Conservation of energy equations: Energy in WDN is associated with the head and it is dissipated due to friction when water flows through a pipe, resulting in head-loss between two connected nodes. Head-loss depends on the water flow through the pipe but also on pipe parameters. Each pipe $i \in \mathcal{L}_p$ is characterized by pipe length l_i, pipe diameter d_i and pipe roughness coefficient c_i. All these quantities are used in the empirical Hazen-Williams (H-W) formula [12] to calculate head-loss. The effect of pipe parameters in this formula is aggregated in the H-W resistance coefficient r_i of each pipe, which is a function $f_{HW}^r : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}^+$ of pipe parameters, defined as:

$$r_i = f_{HW}^r\left(c_i, d_i, l_i\right). \tag{1}$$

The head-loss across pipe $i \in \mathcal{L}_p$ is then calculated using the H-W formula as follows:

$$f_i(q_i) = r_i \left|q_i\right|^{(\nu-1)} q_i \ , \tag{2}$$

where ν is a constant exponent associated with the H-W formula and q_i is the water flow in pipe i.

The energy equations for all the network links, considering (2), can be written as follows:

$$f_i(q_i) + \sum_{j \in \mathcal{N}_u} \left(B_{ij}\ h_j\right) = h_{ext,i}, \quad i \in \mathcal{L}, \tag{3}$$

where h_j is the unknown head of node $j \in \mathcal{N}_u$; $B \in \mathbb{R}^{n_l \times n_u}$ is the incidence flow matrix, indicating the connectivity of nodes with links; element $B_{ij} = +1(-1)$ if the direction of link i enters (leaves from) node j, otherwise $B_{ij} = 0$ (nodes with known head are excluded from this matrix); $h_{ext,i}$ is the sum of known (measured) heads that appear in each equation $i \in \mathcal{L}$. In vector notation, the known head vector is given by $\bar{h}_{ext} \in \mathbb{R}^{n_l}$.

2) Conservation of mass equations: The conservation of mass law for a node $i \in \mathcal{N}_u$ can be summarized as follows: the sum of branch water flows from pipes incident at a node must be equal to the node's external water demand. The conservation of mass equations, considering all the nodes of the network, can be written using the incidence flow matrix as follows:

$$\sum_{j\in\mathcal{L}} \left(B_{ij}^{\top}\, q_j\right) = q_{ext,i}, \quad i \in \mathcal{N}_u \tag{4}$$

where $\bar{q}_{ext,i}$ is the external water demand at node $i \in \mathcal{N}_u$. In vector notation, the external water demands for all nodes are given by $\bar{q}_{ext} \in \mathbb{R}^{n_u}$. Equations (3) and (4) define the network state, which are the water flows in pipes and hydraulic heads at nodes and is indicated by $\bar{x} = [\bar{q}^{\top}\ \bar{h}^{\top}]^{\top}$.

B. Measurement and parameter uncertainty

Sensor measurement and model parameter uncertainty is modeled in this work using intervals. For notational convenience, we adopt the convention of denoting intervals in bold. Let $\bar{\boldsymbol{v}} = [\bar{v}^l, \bar{v}^u]$ be a closed interval vector, where $\bar{v}^l$ is the lower bound vector and $\bar{v}^u$ is the upper bound vector. This then represents the set of real number vectors for which component-wise the following holds: $\bar{\boldsymbol{v}} = \{\bar{v} \in \mathbb{R}^n : \bar{v}^l \leq \bar{v} \leq \bar{v}^u\}$, where n is the size of the vector. Note that any calculations performed in equations containing intervals must be done using interval arithmetic [13].

The uncertain water demands and the reservoir/tank levels are given by the interval vectors:

$$\bar{\mathbf{q}}_{ext} = \left[\bar{q}_{ext}^{\,l},\ \bar{q}_{ext}^{\,u}\right] \tag{5}$$

$$\bar{\mathbf{h}}_{ext} = \left[\bar{h}_{ext}^{\,l},\ \bar{h}_{ext}^{\,u}\right]. \tag{6}$$

We also consider the uncertainty on pipe parameters, which are the roughness coefficients c_i, diameter d_i and length l_i. For a certain pipe i, those will be aggregated in an uncertain H-W coefficient $\mathbf{r}_i$, such that:

$$\bar{\mathbf{r}} = \left[\bar{r}^{\,l},\ \bar{r}^{\,u}\right]. \tag{7}$$

An interval H-W coefficient will transform the pipe headloss function given by (2), into an interval function given by:

$$\boldsymbol{f}_i\left(q_i\right) = \mathbf{r}_i\ |q_i|^{\nu-1} q_i, \quad i \in \mathcal{L}_p \tag{8}$$

C. Problem definition

The problem of solving the hydraulic equations of a WDN when these contain uncertainty in the form of intervals, is reduced to finding the set of all point solutions for the state $\bar{x} = [\bar{q}^{\top}\ \bar{h}^{\top}]^{\top}$, that satisfy the following Nonlinear Interval System of Equations (NISE):

$$\boldsymbol{f}_i\left(q_i\right) + \sum_{j\in\mathcal{N}_u} \left(B_{ij}\ h_j\right) = \boldsymbol{h}_{ext,i}, \qquad i \in \mathcal{L} \tag{9a}$$

$$\sum_{j\in\mathcal{L}} \left(B_{ij}^{\top}\ q_j\right) = \mathbf{q}_{ext,i}, \qquad i \in \mathcal{N}_u \tag{9b}$$

As this set of solutions may have a rather complex form, the following 'interval solutions' are most often considered [14]: 1) Interval Hull (IH) solution is the smallest interval vector $\bar{\mathbf{x}}$ containing all solutions. 2) Outer Interval (OI) solution is any interval vector enclosing the interval hull solution. 3) INner Interval (INI) solution is any interval vector that is a subset of the interval hull solution.

Most approaches in the literature deal with the problem of solving Linear Interval Systems of Equations (LISE), due to the complexity that arises when using interval arithmetic [15], [13]. When using interval arithmetic to solve LISE, the solution interval is inherently an overestimation, thus many usually iterative methods have been proposed to approximate the IH solution, such as the Gauss Elimination method, LU decomposition method and the iterative Jacobi method [16].

Finding the interval hull solution is an NP-hard problem so alternative approaches deal with the problem of finding an OI solution to LISE, while the INI solution is used as a measure of the overestimation of the solution [17], [18]. The problem of finding the IH solution to LISE can also be solved using optimization techniques [19]. Some approaches have been proposed for solving NISE, such as [20] which uses affine arithmetic to represent the equations and interval linearization to deal with the nonlinearities. In [21], a similar problem of solving nonlinear interval number programming problems was investigated. Other applied methods for interval analysis are given in [22].

III. Iterative Hydraulic Interval State Estimation

In this work we propose an iterative method for approximating the IH solution of the NISE in (9), named *Iterative Hydraulic Interval State Estimation* (IHISE). The main steps of the IHISE algorithm are:

Step 1: Find initial bounds on the state variables using physical constraints of the system.
Step 2: Use bounding linearization to bound the nonlinearities in a convex set.
Step 3: Formulate a Linear Program using the resulting linear inequalities.
Step 4: Solve the Linear Program and produce new bounds on the system state.
Step 5: Iteratively approximate the IH solution of (9) by repeating steps 2 to 4 until convergence of bounds.

Next, the five steps of IHISE are described in detail.

A. Step 1: Initial bounds on state vector

The first step of the IHISE algorithm is to impose initial bounds on the state vector $\bar{x} = [\bar{q}^{\top}\ \bar{h}^{\top}]^{\top}$ which should be an OI solution of (9). The initial bounds on the unknown head vector $\bar{h}$ are given by the interval vector $\bar{\mathbf{h}}^{(0)} = \left[\bar{h}^{l\,(0)},\ \bar{h}^{u\,(0)}\right]$ and are chosen by considering physical properties of the network. The minimum head vector $\bar{h}^{l\,(0)}$ is set equal to the elevation of each node and the maximum head vector $\bar{h}^{u\,(0)}$ is set equal to the sum of reservoir and pump heads, which is physically the maximum head that any node in the network can have.

The special structure (9a), in which each equation contains only one flow state q_i, allows us to use the initial bounds on heads $\mathbf{h}^{(0)}$ to find the initial bounds on the flows. We rewrite (9a) as follows:

$$\boldsymbol{f}_i\left(q_i\right) = -\sum_{j\in\mathcal{N}_u} \left(B_{ij}\ \mathbf{h}_j^{(0)}\right) + \mathbf{h}_{ext,i} = \mathbf{y}_i, \tag{10}$$

where $\mathbf{y}_i = \left[y_i^l, y_i^u\right]$ is a known interval derived from the known terms in (10) using interval analysis. The function $\boldsymbol{f}_i\left(q_i\right)$, when $i \in \mathcal{L}_p$, is given by (8). This function is *inclusion isotonic* [13] meaning that if $\mathbf{q}_i^1 \subseteq \mathbf{q}_i^2$ then $\boldsymbol{f}(q_i^1) \subseteq \boldsymbol{f}(q_i^2)$. This property allows us to use the bounds on this function to find the bounds on the unknown pipe flows, by rearranging (10) with respect to $q_i, i \in \mathcal{L}_p$.The initial bounds on the flow state vector are denoted by $\bar{\mathbf{q}}^{(0)} = \left[\bar{q}^{\,l\,(0)}, \bar{q}^{\,u\,(0)}\right]$.

B. Step 2: Bounding linearization of interval nonlinear terms

This step aims at enclosing in a convex set the interval nonlinear terms $\boldsymbol{f}_i(q_i)$ in (9a) given an interval for $q_i \in \left[q_i^l, q_i^u\right]$, using bounding linearization [20]. The interval nonlinear function $\boldsymbol{f}_i(q_i)$ can be enclosed between two lines, a lower line $f_{L,i}^l(q_i) = \lambda_i^l q_i + \beta_i^l$ and an upper line $f_{L,i}^u(q_i) = \lambda_i^u q_i + \beta_i^u$ such that:

$$f_{L,i}^l(q_i) \leq \boldsymbol{f}_i(q_i) \leq f_{L,i}^u(q_i), \ \forall q_i \in \left[q_i^l, q_i^u\right], \ i \in \mathcal{L} \quad (11)$$

The goal is to define the line parameters such as to minimize the area of the resulting convex set. Since $\boldsymbol{f}_i\left(q_i\right)$ is inclusion isotonic, the procedure needs to deal only with the two following ”boundary” functions:

$$f_i^l(q_i) = r_i^l |q_i|^{\nu-1} q_i \quad (12a)$$
$$f_i^u(q_i) = r_i^u |q_i|^{\nu-1} q_i. \quad (12b)$$

First the procedure to define the lower line $f_{L,i}^l(q_i)$ is described. The slope λ_i^l is calculated using the lower two points of $\boldsymbol{f}_i\left(q_i\right)$ at the boundaries q_i^l and q_i^u:

$$\lambda_i^l = \frac{\min\left(f_i^l\left(q_i^u\right), f_i^u\left(q_i^u\right)\right) - \min\left(f_i^l\left(q_i^l\right), f_i^u\left(q_i^l\right)\right)}{q_i^u - q_i^l}. \quad (13)$$

To minimize the area of the convex set but also make sure that (11) is satisfied, the offset β_i^l is defined by solving analytically the following optimization problem, using the Karush-Kuhn-Tucker optimality conditions:

$$\begin{array}{ll} \max & \beta_i^l \\ s.t. & \begin{cases} f_{L,i}^l(q_i) \leqslant f_i^l(q_i) \\ f_{L,i}^l(q_i) \leqslant f_i^u(q_i) \\ q_i^l \leqslant q_i \leqslant q_i^u \end{cases} \end{array} \quad (14)$$

The procedure for finding the upper line $f_{L,i}^u(q_i)$ parameters is similar, thus is omitted. The convex set defined by the lines in the range $q_i \in \left[q_i^l, q_i^u\right]$, will be referred to as the *approximated uncertainty area*.

C. Step 3: Formulation of Linear Program

The bounding linearization of Section III-B, yields linear inequality constraints for the interval functions $\boldsymbol{f}_i(q_i)$. These inequalities can replace $\boldsymbol{f}_i(q_i)$ in (9) with new variables ζ_i yielding:

$$h_{ext,i}^l \leq \zeta_i + \sum_{j \in \mathcal{N}_u} \left(B_{ij}\ h_j\right) \leq h_{ext,i}^u, \qquad i \in \mathcal{L} \quad (15)$$
$$q_{ext,i}^l \leq \sum_{j \in \mathcal{L}} \left(B_{ij}^\top\ q_j\right) \leq q_{ext,i}^u, \qquad i \in \mathcal{N}_u \quad (16)$$
$$\lambda_i^l q_i + \beta_i^l \leq \zeta_i \leq \lambda_i^u q_i + \beta_i^u, \qquad i \in \mathcal{L}, \quad (17)$$
$$q_i^l \leq q_i \leq q_i^u \qquad i \in \mathcal{L}, \quad (18)$$
$$h_i^l \leq h_i \leq h_i^u \qquad i \in \mathcal{N}_u. \quad (19)$$

Using the constraints (15)-(19), a Linear Programming (LP) problem can be formulated that minimizes or maximizes a specific variable z_i, where $\bar{z} = [\bar{q}^\top\ \bar{h}^\top\ \bar{\zeta}^\top]^\top$:

LPmin:

$$\left\{ \begin{array}{ll} \min\limits_{\bar{z}} & z_i \\ s.t. & (15)-(19) \end{array} \right\}$$

LPmax:

$$\left\{ \begin{array}{ll} \max\limits_{\bar{z}} & z_i \\ s.t. & (15)-(19) \end{array} \right\}$$

D. Step 4: Solution of the linear interval system of equations

The objective of the optimization problem formulated in Section III-C is to find the lower and upper bounds on the state vector $\bar{x}$ that satisfy the inequalities (15)-(19). To achieve this, a total of $2(n_l + n_u)$ LPs are solved, where each problem derives either the lower or upper bound of an individual state variable, indicated by $z_i^*, \quad i \in \{1, \cdots, n_l + n_u\}$. This procedure is described in Algorithm 1.

Algorithm 1 Solution of LISE using LP

begin
1: **for** $i = 1$ **to** $n_l + n_u$ **do**
2: Minimize z_i by solving problem **LPmin**
3: $x_i^l = z_i^*$
4: Maximize z_i by solving problem **LPmax**
5: $x_i^u = z_i^*$
6: **end for**
7: $\bar{\mathbf{x}} = \left[\bar{x}^{\,l},\ \bar{x}^{\,u}\right]$
return $\bar{\mathbf{x}}$

E. Step 5: Iterative solution of the nonlinear interval system of equations

In Algorithm 1 the linearized version of the original problem in (9) is solved. This is an outer interval solution to the nonlinear problem, which guarantees to include the interval hull solution. To find the smallest possible interval $\bar{\mathbf{x}} = \left[\bar{x}^{\,l},\ \bar{x}^{\,u}\right]$ that satisfies (9), an iterative method is used. At each iteration m, the range $\bar{\mathbf{x}}_{bnd}$ in which the optimization algorithm searches for an optimal solution becomes smaller and is redefined as follows:

$$\bar{\mathbf{x}}_{bnd}^{(m+1)} = \bar{\mathbf{x}}_{bnd}^{(m)} \cap \bar{\mathbf{x}}^{\,(m+1)}, \quad (20)$$

where $\bar{\mathbf{x}}^{(m+1)}$ are the bounds calculated for the state vector $\bar{x}$ at iteration m. The iterations stop when the bounds on the state vector remain relatively unchanged, i.e. the change is smaller than a small number ϵ. The number ϵ defines the largest allowable absolute error for the calculated bounds of

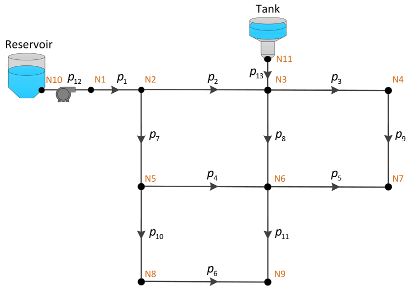

Fig. 1. The benchmark network 'Net1', on which the IHISE algorithm is demonstrated.

each state, e.g. for flow states $\epsilon = 0.01\,\mathrm{m}^3/\mathrm{h}$. The relative change in bounds is checked as follows:

$$e^{(m)} \triangleq \left| \left(\bar{x}^{u\,(m)} - \bar{x}^{l\,(m)}\right) - \left(\bar{x}^{u\,(m-1)} - \bar{x}^{l\,(m-1)}\right) \right|_1 . \tag{21}$$

The procedure for the solution of the NISE is described in Algorithm 2. The resulting bounds are an OI solution of these equations that approximate closely the IH solution.

Algorithm 2 Iterative solution of NISE

begin

Define initial bounds $\bar{\mathbf{h}}^{(0)}$
Calculate initial bounds $\bar{\mathbf{q}}^{(0)}$ as described in Step 1.
$\bar{\mathbf{x}}_{bnd}^{(0)} = \left[\bar{\mathbf{q}}^{(0)\top} \ \bar{\mathbf{h}}^{(0)\top}\right]^\top$
$m = 0$
while $e^{(m)} > \epsilon$ **do**
  Bounding linearization of (9) for $\bar{x} \in \bar{\mathbf{x}}_{bnd}^{(m)}$
  Formulation of problems **LPmin** and **LPmax**
  Find $\bar{\mathbf{x}}^{(m+1)}$ using **Algorithm 1**
  $\bar{\mathbf{x}}_{bnd}^{(m+1)} = \bar{\mathbf{x}}_{bnd}^{(m)} \cap \bar{\mathbf{x}}^{(m+1)}$
  $m = m + 1;$
end while

return $\bar{\mathbf{x}}^{(m)}$

IV. Case Study

A. Illustrative example

The benchmark network 'Net1', which is shown in Fig. 1 and is provided by the EPANET WDN modeling software [23], was used to highlight the efficiency of the IHISE algorithm. The network consists of 11 nodes, of which one is a reservoir and one is a water tank. The head at the water tank is assumed measurable with a level sensor. The connecting links between nodes consist of 12 pipes and a water pump. Nominal water demand patterns, which resemble the pattern observed in real WDN, were assigned at each demand node.

The IHISE algorithm was used on this network to generate bounds on water flows in pipes and hydraulic heads at nodes for a period of 24 hours. The considered measurement uncertainty was $\pm 2\%$ on water demands at nodes. This is the typical error given by manufacturers of water flow meters, while any pseudo-measurements at demand nodes are assumed here to have the same uncertainty as flow meters. Modeling uncertainty was considered and it corresponds to $\pm 2\%$ on pipe Hazen-Williams coefficients.

Additionally, the same bounds were generated using Monte-Carlo Simulations (MCS) of the network where the uncertain parameters were varied randomly at each simulation. The number of simulations $(30,000)$ was sufficient to obtain the best possible bounds, and we will assume that the MCS bounds is the IH solution. For the given network, on a personal computer with Intel Core i5-2400 CPU at 3.10GHz, the IHISE algorithm requires 0.68 seconds to solve each hydraulic step, while for the defined number of MCS the simulation time was 56.8 seconds.

Simulation results are illustrated in Fig. 2, where the bounds on water flows for selected pipes are plotted, and Fig. 3 where the bounds on hydraulic heads for selected nodes are plotted. The IHISE bounds are compared with bounds generated using MCS and it is shown that the best possible bounds are approximated closely and the area they define are a subset of the area defined by the IHISE bounds.

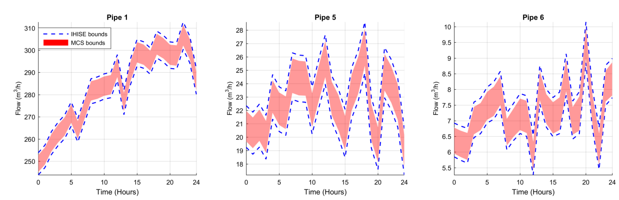

Fig. 2. Comparison of selected pipes water flow bounds, generated by Monte-Carlo Simulations (red area) and the IHISE algorithm (blue dashed lines).

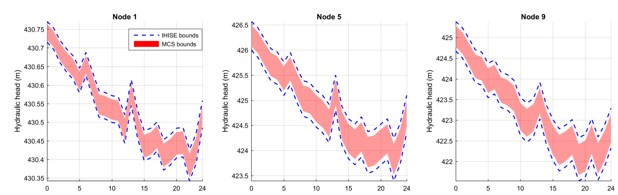

Fig. 3. Comparison of selected nodes hydraulic head bounds, generated by Monte-Carlo Simulations (red area) and the IHISE algorithm (blue dashed lines).

B. Leakage detection

Next we demonstrate how the IHISE algorithm can be utilized in a WDN for detecting an abnormal event, which could possibly be a leakage in the network. In this scenario, the network shown in Fig. 1 is used. Sensors measure the water demands at all nodes and the water levels in the tank and reservoir. This is a typical setup for a water transport network, which is the network of main pipes that distribute water within a city. The uncertainties considered are the same as in the previous example in Section IV-A.

Additionally, an extra flow sensor is installed on pipe 1 to measure the flow through the pump, as it is often the case in WDN. The flow measurements of the sensor on pipe 1, indicated by $q_1^s(k)$, should at all time steps k be a subset of the IHISE bounds, i.e.: $q_1^s(k) \in [q_1^l(k), q_1^u(k)]$,

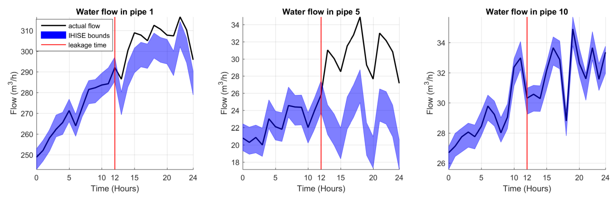

Fig. 4. Example of leakage detection using the IHISE bounds: The measurements from a sensor in pipe 1 (black line), violate the IHISE bounds (blue area), thus the leakage in pipe 5 starting at the $12th$ hour is detected, but not localized.

where $[q_1^l(k), q_1^u(k)]$ is the interval flow state calculated by the IHISE algorithm. If at any time step k the sensor flow reading $q_1^s(k) \notin [q_1^l(k), q_1^u(k)]$, it would indicate to the network operator that an abnormal event took place in the network, since measurement and modeling uncertainties were considered in the derivation of the IHISE bounds.

As an example of this, an abrupt leakage event, i.e. a pipe burst, was simulated in pipe 5, at hour 12 with a magnitude of 10 m^3/h, which is the typical rate of water lost during a large pipe burst. In Fig. 4 it is shown that there is a violation of the flow bounds generated by the IHISE algorithm for the illustrated pipes. Using the sensor on pipe 1, the leakage event is detected in the following time step since $q_1^s(k) \notin [q_1^l(k), q_1^u(k)]$. It should be noted that the placement of the extra sensor is critical because, as illustrated in Fig. 4, if the extra sensor was installed on pipe 10 the bounds would not be violated, i.e. $q_{10}^s(k) \in [q_{10}^l(k), q_{10}^u(k)]$ even in the presence of the leak.

V. Conclusions

In this work the problem of estimating bounds on WDN hydraulic states is addressed. A new methodology is proposed that generates interval state estimates. The proposed *Iterative Hydraulic Interval State Estimation* (IHISE) algorithm generates bounds on hydraulic states of the network, by taking into account the water demand uncertainty and modeling uncertainty in the form of uncertain pipe parameters. The uncertainties are modeled as intervals. The results show that the proposed methodology is able to generate tight bounds on hydraulic states and can be used to detect leakages in the network.

There are several issues that can be investigated for future work. One direction is to use the generated bounds to apply fault-detection methods that detect and localize leakages in the network. Additionally, the bounds on hydraulic states of the network can be used to generate bounds on water quality states, since the dynamics of hydraulic and quality states of a water network are interconnected. Finally, the IHISE algorithm can be generalized to a broader class of large-scale interconnected systems with clearly defined applicability restrictions.

References

[1] C. J. Hutton, Z. Kapelan, L. Vamvakeridou-Lyroudia, and D. A. Savic, "Dealing with uncertainty in water distribution system models: A framework for real-time modeling and data assimilation," *Water Resources Planning and Management*, vol. 140, pp. 169–183, 2014.

[2] D. A. Savic, Z. S. Kapelan, and P. M. Jonkergouw, "Quo vadis water distribution model calibration?" *Urban Water Journal*, vol. 6, pp. 3–22, 2009.

[3] D. Kang and K. Lansey, "Demand and roughness estimation in water distribution systems," *Water Resources Planning and Management*, vol. 137, pp. 20–30, 2011.

[4] S. G. Vrachimis, D. G. Eliades, and M. M. Polycarpou, "The backtracking uncertainty bounding algorithm for chlorine sensor fault detection," in *Proceedings of 13th Computing and Control for the Water Industry Conference, CCWI 2015*, vol. 119, no. 1, 2015, pp. 613–622.

[5] A. Bargiela and G. D. Hainsworth, "Pressure and flow uncertainty in water systems," *Water Resources Planning and Management*, vol. 115, no. 2, pp. 212–229, 1989.

[6] M. Brdys and K. Chen, *Joint state and parameter estimation of dynamic water supply systems with unknown but bounded uncertainty*. John Wiley & Sons, Inc., 1994.

[7] B. Gabrys and A. Bargiela, "Integrated neural based system for state estimation and confidence limit analysis in water networks," in *Proceedings of European Simulation Symposium ESS'96*, 1997, pp. 398–402.

[8] C. T. Arsene, D. Al-Dabass, and J. Hartley, "Confidence limit analysis of water distribution systems based on a least squares loop flows state estimation technique," in *Proceedings of 2011 UKSim 5th European Symposium on Computer Modeling and Simulation*. IEEE, 2011, pp. 94–101.

[9] A. Bargiela, W. Pedrycz, and M. Tanaka, "A study of uncertain state estimation," *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans.*, vol. 33, no. 3, pp. 288–301, 2003.

[10] D. G. Eliades, D. Stavrou, S. G. Vrachimis, C. G. Panayiotou, and M. M. Polycarpou, "Contamination event detection using multi-level thresholds," in *Proceedings of 13th Computing and Control for the Water Industry Conference, CCWI 2015*, vol. 119, no. 1, 2015, pp. 1429–1438.

[11] S. Díaz, R. Mínguez, and J. González, "Topological observability analysis in water distribution systems," *Journal of Water Resources Planning and Management*, vol. 143, no. 5, p. 06017001, 2017.

[12] P. F. Boulos, K. E. Lansey, and B. W. Karney, *Comprehensive water distribution systems analysis handbook for engineers and planners*. American Water Works Association, 2006.

[13] R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to interval analysis*. Society for Industrial and Applied Mathematics, 2009.

[14] L. V. Kolev, "Parameterized solution of linear interval parametric systems," *Applied Mathematics and Computation*, vol. 246, pp. 229–246, 2014.

[15] M. Daumas, D. Lester, and C. Munoz, "Verified real number calculations: A library for interval arithmetic," *IEEE Transactions on Computers*, vol. 58, no. 2, pp. 226–237, 2009.

[16] E. Zieniuk, M. Kapturczak, and A. Kużelewski, "Solving interval systems of equations obtained during the numerical solution of boundary value problems," *Computational and Applied Mathematics*, vol. 35, no. 2, pp. 629–638, 2015.

[17] A. Neumaier and A. Pownuk, "Linear systems with large uncertainties, with applications to truss structures," *Reliable Computing*, vol. 13, no. 2, pp. 149–172, 2007.

[18] L. V. Kolev, "A method for outer interval solution of parametrized systems of linear interval equations," *Reliable Computing*, vol. 10, pp. 227–239, 2004.

[19] G. H. Huang and M. F. Cao, "Analysis of solution methods for interval linear programming," *Journal of Environmental Informatics*, vol. 17, no. 2, pp. 54–64, 2011.

[20] L. V. Kolev, "An improved interval linearization for solving nonlinear problems," *Numerical Algorithms*, vol. 37, pp. 213–224, 2004.

[21] C. Jiang, X. Han, and G. P. Liu, "A sequential nonlinear interval number programming method for uncertain structures," *Computer Methods in Applied Mechanics and Engineering*, vol. 197, no. 49, pp. 4250–4265, 2008.

[22] L. Jaulin, M. Kieffer, O. Didrit, and É. Walter, *Applied Interval Analysis: With Examples in Parameter and State Estimation, Robust Control and Robotics*, 2001.

[23] L. A. Rossman, "Epanet 2: users manual," 2000.