Utilizing Matrix Completion for Simulation and Optimization of Water Distribution Networks

Alaa Jamal

University of Illinois at Urbana-Champaign

Mashor Housh (✉ mhoush@univ.haifa.ac.il)

University of Haifa

Research Article

Keywords: Water Distribution Networks, Matrix Completion, Simulation, Optimization

Posted Date: June 7th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-469861/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Utilizing Matrix Completion for Simulation and Optimization of Water Distribution Networks

Alaa Jamal¹, Mashor Housh^{2*}

¹ Dept. of Civil and Environmental Engineering, Univ. of Illinois at Urbana Champaign, Champaign, IL, United States. Email: ajamal@illinois.edu

² Dept. of Natural Resources and Environmental Management, Univ. of Haifa, Haifa 3498838, Israel

* Correspondence: mhoush@univ.haifa.ac.il;

Keywords

Water Distribution Networks; Matrix Completion; Simulation; Optimization.

Acknowledgments

This study was funded by the Israeli Water Authority.

Abstract

Simulation of Water Distribution Networks (WDNs) constitutes a key element for the planning and management of water supply systems. This simulation involves estimating the flows and pressures by solving a linear set of mass conservation equations and a nonlinear set of energy conservation equations. The literature presents different formulations of heads-flows equations to derive the flows and heads in WDN. These formulations differ in terms of dimensionality, computational cost, and solution accuracy. Whereas this problem has been the subject of active research in the past, in the last decades a state of stagnation was reached and no new formulations were introduced. In this study, we propose a novel formulation that utilizes a matrix completion technique to construct a reduced-size nonlinear system of equations that guarantees both mass and energy conservation. Unlike former formulations that rely on the

topology of the network, in the proposed method we employ a matrix completion technique in which arbitrary entries are added to the equation system to facilitate its solution. The advantages of the proposed method are demonstrated in simulation and optimization settings. In the former, the method demonstrates improved scalability and accuracy as compared with other widely known formulations. In the latter, the new formulation leads to smaller optimization problems, which are otherwise intractable when the classical formulation is used. Our results reopen an old debate on the best formulation for WDN simulation and optimization tasks and show that the matrix completion technique is a viable solution option for the problem.

1 Introduction

Water Distribution Networks (WDNs) are large scale systems responsible for supplying drinking water from sources to consumers. These systems are composed of several links and nodes, such as pipes, pumps, valves, and reservoirs. The design, operation, and monitoring of WDNs rely heavily on WDN mathematical modeling to simulate the flows and pressures. For example, Savic and Walters (1997), Kapelan et al. (2005), Perelman et al. (2013), and many others used hydraulic simulation engines (e.g. EPANET, Rossman 2000) to derive optimal design and operation decisions for WDNs. Simulation engines are also useful for leakage estimation and detection (Martínez-Solano et al. 2017), system security monitoring (Housh and Ohar 2017), system maintenance and rehabilitation (Fuchs-Hanusch et al. 2008), and system resiliency and reliability assessment (Klise et al. 2017).

In general, the flows and pressures in a WDN are calculated by solving a linear set of mass conservation equations and a non-linear set of energy conservation equations that should be solved simultaneously, thus making the problem nonlinear (Giustolisi et al. 2012; Moosavian and Roodsari 2014).

Different formulations exist that can be utilized to derive the flows and heads in the WDN. These formulations differ in terms of dimensionality, computational cost, and solution accuracy. There is a rich body of literature on the development of formulation and solution methods for WDN modeling. Next, we provide a concise summary for key formulations and techniques in this domain.

One of the most widely used formulations is the Heads and Flows (H-Q) formulation, proposed by Todini and Pilati (1988) as the Global Gradient Algorithm (GGA). In this formulation, the nodal heads and flows in the links are considered unknowns within a simultaneous nonlinear system of equations that includes both the mass and energy conservation equations. The formulation showed excellent convergence and was adopted in the widely used EPANET hydraulic simulator (Rossman 2000). Later, Basha and Kassab (1996) proposed solving the equations with the H-Q formulation based on the theory of disturbances, which yields a fast convergence rate to final solutions. However, Moosavian (2017) argued that the accuracy of the method was not sufficient.

Overall, the H-Q formulation suffers from high computational cost due to the problem's dimensionality. In several papers, the H formulation was proposed, in which only the nodal heads are considered unknowns within a nonlinear system of equations that guarantees both mass and energy conservation. Martin and Peters (1963) proposed formulating all the WDN equations in terms of nodal heads and updating the heads iteratively using the Newton–Raphson method. Liu (1969) solved the H formulation using this method, while decomposing the Jacobian matrix into diagonal and nondiagonal matrices. This decomposition is utilized by observing that the nondiagonal matrix is negligible as compared to the diagonal one. However, Moosavian (2017) argued that this method suffers from divergence in the case of an inappropriate initial guess and in the case of large-scale networks.

An additional widely studied formulation is the Q formulation, in which the flows are the only unknowns within a nonlinear system of equations that guarantees both mass and energy conservation. In this formulation, the energy conservation equations are written around the loops in the network, where in each loop the headloss should be equal to zero. This formulation was shown to be numerically superior to the H formulation (Nielsen 1989). Wood and Charles (1972) proposed the linear theory method for solving the Q formulation. In this method, the nonlinear headloss is linearized in each iteration, using the zero and the current flow as anchors. A disadvantage of this method is the low convergence rate due to the oscillations in the iterative process. Similarly, Wood and Funk (1993) used the extended Taylor series to solve a linearized Q formulation iteratively. This method is currently used in the commercial software KYPIPE (Wood 1980).

Building on the Q formulation, several authors suggested partitioning the network graph into a spanning tree and a corresponding co-tree (i.e., non-spanning tree links), where the size of the chord is equal to the number of loops in the network. Such partitioning will allow the spanning tree flows to be written as a function of chord flows, and thus the chord flows are the only unknowns within a nonlinear system of equations that guarantees both the mass and the energy conservation. We call this approach the Qchord formulation. It is noteworthy that the number of chord flows (i.e., number of loops) is always smaller than the number of links, meaning that this system of equations will have fewer unknowns (Nielsen 1989; Rahal 1995; Elhay et al. 2014; Qiu et al. 2020). An additional technique for reducing the number of unknowns is to formulate the system of equations in terms of corrections of flow around each loop (these correction variables are denoted by ΔQ). We call this approach the Null-Space formulation, since the loops of the network can be represented by the null space of the network's incidence matrix (Epp and Fowler 1970; Benzi et al. 2005; Housh 2011; Abraham and Stoianov 2016). In this formulation, observing that the mass conservation is written as an undetermined system

of linear equations, the complete general solution to the underdetermined system can be characterized by adding a particular solution (i.e., an arbitrary flow distribution that satisfies the mass conservation) to a linear combination of the null space vectors. As such, mass conservation can be guaranteed by representing the flows as an affine map of ΔQ . The energy conservation equations are written around the loops in the network, and in each loop the headloss should be equal to zero. Thus, we obtain a nonlinear system of equations where both the number of equations and the number of unknowns equals the number of loops in the network. This reduced system of equations can be solved using the Newton–Raphson method (Abraham and Stoianov 2016).

In this paper, we propose a novel formulation that utilizes the matrix completion technique to construct a reduced-size nonlinear system of equations that guarantees both mass and energy conservation. Unlike former formulations that rely on the topology of the network (e.g., loops) and/or its original physical unknowns, in the proposed formulation we arbitrarily add "fictional" entries and auxiliary unknowns to facilitate direct mathematical manipulations.

More specifically, we employ a matrix completion technique, in which arbitrary entries are added to the system of equations in order to obtain invertible squared matrices that facilitate the elimination of both the H and Q variables in the problem. To the best of our knowledge, there is no published study exists that considered such a procedure to solve the hydraulic simulation of WDNs, although it holds advantages compared to other methods discussed in the literature, as is shown in the following. The idea of using the completion technique is inspired by Yang (2011) who used similar technique to solve bilinear system of equations.

Using benchmark networks, the proposed method is compared to (1) the H - Q formulation, (2) the Q chord formulation, and (3) the Null-Space formulation. These formulations were chosen because the first is widely used in the literature while in the latter two formulation, the number

of number of unknowns are reduced, similar to the proposed matrix completion formulation. Nonetheless, unlike the former methods, the suggested method does not rely on the topological properties (e.g., loops, spanning trees) of the network. We also demonstrated the manner in which the new formulation can facilitate full elimination of both Q-H variables when used within an optimization model. This elimination reduces the size of the decision variables, which results in an efficient solution of the optimization model.

2 State-of-the-art formulations

In the WDN hydraulic simulation problem the objective is to find the flows in the links (Q) and the heads at the nodes (H) such that the mass and the energy conservation are satisfied in the entire network.

2.1 Mass conservation

Mass conservation at each node can be formulated as:

$$\sum_{l \in IN_k} Q_l - \sum_{l \in OUT_k} Q_l = q_k \quad \forall k = 1 \dots nNodes \quad (1)$$

where Q_l is the flow in link l ; q_k is the demand at node k ; IN_k and OUT_k are sets of ingoing and outgoing links that are linked to node k , respectively; $nNodes$ is the number of the junction nodes in the network (i.e., not including source nodes).

2.2 Energy conservation

Energy conservation in each link can be formulated as:

$$H_{l_e} - H_{l_s} = \Delta H_l \quad \forall l = 1 \dots nLinks \quad (2)$$

where H_{l_e} and H_{l_s} are the heads at the end and start nodes of link l , respectively; $nLinks$ is the number of the links; ΔH_l is the head change of link l . In case of pipe, ΔH_l is the headloss (note the minus in Equation 3) which could be defined using the Hazen-William formula:

$$\Delta H_l = -R_l Q_l |Q_l|^{0.852} \quad (3)$$

where R_l is the Hazen-William resistance of pipe l . In cases where the link represents a pump, the head gain is defined as the pump characteristic curve that dictates the head gain as a function of the flow in the pump. For example, when the pump characteristic curve is given as quadratic function (Lansey and Mays 1989), ΔH_l can be defined as:

$$\Delta H_l = aQ_l^2 + bQ_l + c \quad (4)$$

where a , b and c are constants which are determined according to the specifications of the pump.

For nodes with fixed heads, Equation (5) should be satisfied:

$$H_k = H_{0,k} \quad \forall k \in FH \quad (5)$$

where FH is a set of nodes with predetermined fixed head (e.g. reservoir); $H_{0,k}$ is the given value of the fixed head at node k .

2.3 Matrix formulation

Equations 1-5 can be formulated in a matrix form. Equation 1, which dictates the mass conservation at the nodes, can be formulated as an underdetermined linear system of equations (underdetermined = more unknowns than equations) as shown in Equation (6).

$$\mathbf{A}^T \cdot \mathbf{Q} = \mathbf{q} \quad (6)$$

where $\mathbf{A} \in \mathbb{R}^{nLinks \times nNodes}$ is the oriented incidence matrix of the network in which each entry (l, k) equals 1 if link l is ingoing node k , -1 if the link l is outgoing node k , or 0 if there is no link between nodes l and k ; the vector $\mathbf{Q} \in \mathbb{R}^{nLinks \times 1}$ collects the unknown flows in all links; the vector $\mathbf{q} \in \mathbb{R}^{nNodes \times 1}$ collects the demands at all nodes; $(\cdot)^T$ is the transpose operation. Since $nLinks > nNodes$ in looped networks, the linear equation system above is undetermined, namely, the null space of $\mathbf{A}^T$ is not empty. If we define the matrix $\mathbf{N}$ as the orthonormal basis for the null space, then $\mathbf{N} \in \mathbb{R}^{nLinks \times nLoops}$ where $nLoops = nLinks - nNodes$ equals to the number of loops and quasi-loops (i.e. paths between constant head nodes, Qiu et al. 2020) in the network.

Equations 3-5 can be also presented in matrix form as shown in Equation (7).

$$\mathbf{A} \cdot \mathbf{H} = \mathbf{H}_0 + \Delta \mathbf{H}(\mathbf{Q}) \quad (7)$$

where the vector $\mathbf{H} \in \mathbb{R}^{nNodes \times 1}$ collects the heads at all nodes, $\Delta \mathbf{H}(\mathbf{Q}) \in \mathbb{R}^{nLinks \times 1}$ is a vector of functions representing the head loss/gain in the links. Note that element l in the vector $\Delta \mathbf{H}(\mathbf{Q})$, is one-dimensional function of the form $\Delta H_l(\mathbf{Q}_l)$; $\mathbf{H}_0 \in \mathbb{R}^{nLinks \times 1}$ is a vector which collects the values of the fixed heads in the network. Since the vector of functions $\Delta \mathbf{H}$ holds nonlinear terms, the system of equations in Equation (6) is nonlinear.

As discussed previously, the hydraulic simulation problem could be achieved using different formulations as summarized in Table 1:

(a) The H-Q formulation directly solves Equations (6) and (7) as simultaneous nonlinear system of equations with $nLinks + nNodes$ unknowns/Equations. Thus, both $\mathbf{Q}$ and $\mathbf{H}$ are obtained as a direct outcome of the nonlinear equations' solver.

(b) The H-formulation extracts $\mathbf{Q}$ from Equation (7) as a function of nodal heads, $\mathbf{Q} = \Delta \mathbf{H}^{-1}(\mathbf{A} \cdot \mathbf{H} - \mathbf{H}_0)$, which is then substituted in Equation (6) to result in a nonlinear system of equations with $nNodes$ unknowns/Equations. The heads, $\mathbf{H}$, are obtained as a direct outcome of the nonlinear equations' solver while the flows could be obtained by substituting in $\mathbf{Q} = \Delta \mathbf{H}^{-1}(\mathbf{A} \cdot \mathbf{H} - \mathbf{H}_0)$. Noteworthy that, in Equation (7), each element l in the vector of functions $\Delta \mathbf{H}(\mathbf{Q})$, is a one-dimensional function of the form $\Delta \mathbf{H}_l(\mathbf{Q}_l)$, thus the functions inversion $\Delta \mathbf{H}^{-1}(\cdot)$ could be derived analytically (Martin and Peters 1963; Boulos et al. 2006).

(c) In the Q-formulation, Equation (6) is solved together with Equation (8) as simultaneous nonlinear system of equations with $nLinks$ unknowns/Equations. The flows, $\mathbf{Q}$, are obtained as a direct outcome of the nonlinear equations' solver while the heads, $\mathbf{H}$, could be obtained from Equation (7) which becomes linear after determining the flows.

$$\mathbf{N} \cdot (\mathbf{H}_0 + \Delta \mathbf{H}(\mathbf{Q})) = \mathbf{0} \quad (8)$$

(d) In the Qchord formulation, Equation (6) is re-written as:

$$\begin{aligned} \mathbf{A}^T \cdot \mathbf{Q} &= \mathbf{q} \\ \mathbf{A}_{ST} \cdot \mathbf{Q}_{ST} &= \mathbf{q} - \mathbf{A}_{ch} \cdot \mathbf{Q}_{ch} \\ \mathbf{Q}_{ST} &= \mathbf{A}_{ST}^{-1} \cdot \mathbf{q} - \mathbf{A}_{ch} \cdot \mathbf{Q}_{ch} \\ \mathbf{Q} &= \mathbf{P}_{ST} \cdot \mathbf{Q}_{ST} + \mathbf{P}_{ch} \cdot \mathbf{Q}_{ch} \\ \mathbf{Q} \mathbf{Q}_{ch} &= \mathbf{P}_{ST} \cdot \mathbf{A}_{ST}^{-1} \cdot \mathbf{q} - \mathbf{A}_{ch} \cdot \mathbf{Q}_{ch} + \mathbf{P}_{ch} \cdot \mathbf{Q}_{ch} \end{aligned} \quad (9)$$

where $\mathbf{A}_{ST}$ is sub-matrix of $\mathbf{A}^T$ which collects columns corresponding to a spanning tree; $\mathbf{A}_{ch}$ is sub-matrix of $\mathbf{A}^T$ which collects columns corresponding to co-tree; $\mathbf{P}_{st}$ and $\mathbf{P}_{ch}$ are permutation matrices to extract the spanning tree and the chord flows, respectively; $\mathbf{Q}_{ST}$ and $\mathbf{Q}_{ch}$ are flows in spanning tree and chord links, respectively.

After replacing $\mathbf{Q}$ in Equation (8) with $\mathbf{Q}(\mathbf{Q}_{ch})$ as obtained from Equation (9), it could be solved as nonlinear system of equations with $nLoops$ unknowns/Equations. The chord flows, $\mathbf{Q}_{ch}$, are obtained as a direct outcome of the nonlinear equations solver, the flows are obtained by substitution, from the function $\mathbf{Q}(\mathbf{Q}_{ch})$, while the heads, $\mathbf{H}$, could be obtained from Equation (7) which becomes linear after determining the flows.

(e) In the Null-Space formulation, Equation (6) is re-written as:

$$\begin{aligned} \mathbf{A}^T \cdot \mathbf{Q} &= \mathbf{q} \\ \mathbf{Q} \Delta \mathbf{Q} &= \mathbf{Q}_p + \mathbf{N} \cdot \Delta \mathbf{Q} \end{aligned} \tag{10}$$

where $\mathbf{Q}_p$ is particular solution to the undetermined system (i.e., arbitrary flow distribution that satisfies Equation 6); $\Delta \mathbf{Q}$ is corrections to flow around each loop.

After replacing $\mathbf{Q}$ in Equation (8) with $\mathbf{Q}(\Delta \mathbf{Q})$ as obtained from Equation (10), it could be solved as nonlinear system of equations with $nLoops$ unknowns/Equations. The flow corrections, $\Delta \mathbf{Q}$, are obtained as a direct outcome of the nonlinear equations' solver, the flows are obtained by substitution, from the function $\mathbf{Q}(\Delta \mathbf{Q})$, while the heads, $\mathbf{H}$, could be obtained from Equation (7) which becomes linear after determining the flows.

Table 1 Summary of formulations

Formulation	Unknowns/Equations	Solver Output	Reference
H-Q	nLinks+nNodes	Heads and Flows	Todini and Pilati (1988)

H	nNodes	Heads	Martin and Peters (1963)
Q	nLinks	Flows	Wood and Charles (1972)
Qchord	nLoops	Flows in chord tree	Nielsen (1989)
Null-Space	nLoops	Flow corrections in loops	Epp and Fowler (1970)
Matrix Completion	nLoops	Auxiliary variables	

3 Matrix Completion Formulation

The proposed method aims at obtaining the flows $\mathbf{Q}$ and heads $\mathbf{H}$ through an auxiliary problem using matrix completion technique. The matrix $\mathbf{A}^T$ is missing $nLoops$ independent rows to become squared full-rank matrix (i.e., invertible matrix). Let $\mathbf{A}_c^T \in \mathbb{R}^{nLoops \times nLinks}$ be an arbitrary completion matrix of $\mathbf{A}^T$, that collects $nLoops$ rows which are linearly independent of the rows in $\mathbf{A}^T$, then we can define the a squared invertible matrix $\mathbf{A}^T \in \mathbb{R}^{nLinks \times nLinks}$ as:

$$\mathbf{A}^T = \begin{bmatrix} \mathbf{A}^T \\ \mathbf{A}_c^T \end{bmatrix} \quad (11)$$

Equation (6) can be written using $\mathbf{A}^T$ as:

$$\mathbf{A}^T \cdot \mathbf{Q} = \begin{bmatrix} \mathbf{q} \\ \mathbf{z} \end{bmatrix} \quad (12)$$

where $\mathbf{z} \in \mathbb{R}^{nLoops \times 1}$ is a vector of auxiliary variables. Notice that, unlike a "classical" linear equation system which only includes variables in the left hand side (e.g. Equation 6), Equation (12) is a linear system with extra variables on the right hand side. These auxiliary variables

have an important role in the completion technique; without these variables adding more equations would change the solution space of the original problem (Equation 6). But with these variables in the right hand side, the original solution space will be maintained. To see that the solution space is maintained, note that any solution $(\mathbf{Q}^*, \mathbf{z}^*)$ to Equation (12), must satisfy the first $nNodes$ equations, which is exactly the linear system in Equation (6). On the other hand, if $\mathbf{Q}^*$ is a solution for Equation (6), then we can set $\mathbf{z}^* = \mathbf{A}_c \cdot \mathbf{Q}^*$ such that $(\mathbf{Q}^*, \mathbf{z}^*)$ is a solution for Equation (12).

Recalling that $\mathbf{A}^T$ is invertible matrix, we can extract $\mathbf{Q}$ as a function of $\mathbf{z}$, as follows:

$$\mathbf{Q}(\mathbf{z}) = \mathbf{A}^{-T} \cdot \begin{bmatrix} \mathbf{q} \\ \mathbf{z} \end{bmatrix} \quad (13)$$

Note that there are many ways to complete the matrix $\mathbf{A}^T$, since there is far more than one choice of the matrix $\mathbf{A}_c^T$ that satisfies the linear independence requirement. In fact, randomly sampling the entries of $\mathbf{A}_c^T$ is likely to produce a valid completion matrix that satisfies the linear independence requirement. In this study, each entry of $\mathbf{A}_c^T$ is sampled arbitrarily from uniform distribution between 0 and 1.

Next, we replace $\mathbf{Q}$ in Equation (7) with $\mathbf{Q}(\mathbf{z})$ as obtained from Equation (13). This will yield overdetermined equations system (more equation than unknowns):

$$\mathbf{A} \cdot \mathbf{H} = \mathbf{H}_0 + \Delta \mathbf{H}(\mathbf{Q}(\mathbf{z})) \quad (14)$$

Equation (14) could be formulated using the squared invertible matrix, $\mathbf{A}$, as follows:

$$\mathbf{A} \cdot \begin{bmatrix} \mathbf{H} \\ \mathbf{w} \end{bmatrix} = \mathbf{H}_0 + \Delta \mathbf{H}(\mathbf{Q}(\mathbf{z})) \quad (15)$$

where $\mathbf{w} \in \mathbb{R}^{nLoops \times 1}$ is vector of auxiliary variables. This auxiliary vector is essential for matching the dimensions after using $\mathbf{A}$ instead of $\mathbf{A}$. From Equation (15), $\mathbf{H}$ and $\mathbf{w}$ could be calculated as a function of $\mathbf{z}$:

$$\begin{bmatrix} \mathbf{H} \\ \mathbf{w} \end{bmatrix} = \mathbf{A}^{-1} \cdot (\mathbf{H}_0 + \Delta\mathbf{H}(\mathbf{Q}(\mathbf{z}))) \quad (16)$$

Adding the $\mathbf{w}$ variables, however, will change the original energy conservation in Equation (7), unless we require that:

$$\mathbf{w} = \mathbf{0} \quad (17)$$

That is, we need to have:

$$\mathbf{w} = \mathbf{P}_w \cdot \begin{bmatrix} \mathbf{H} \\ \mathbf{w} \end{bmatrix} = \mathbf{P}_w \cdot \mathbf{A}^{-1} \cdot (\mathbf{H}_0 + \Delta\mathbf{H}(\mathbf{Q}(\mathbf{z}))) = \mathbf{0} \quad (18)$$

where $\mathbf{P}_w = [\mathbf{0}^{nNodes \times nNodes} \mid \mathbf{I}^{nLoops \times nLoops}]$ is a permutation matrix to extract $\mathbf{w}$ from $\begin{bmatrix} \mathbf{H} \\ \mathbf{w} \end{bmatrix}$; $\mathbf{0}$ is a matrix of zeros; $\mathbf{I}$ is the identity matrix.

To conclude, in the proposed completion method, we solve the following nonlinear system of equations with $nLoops$ unknowns/Equations, in which the unknowns are the $\mathbf{z}$ auxiliary variables:

$$\mathbf{P}_w \cdot \mathbf{A}^{-1} \cdot (\mathbf{H}_0 + \Delta\mathbf{H}(\mathbf{Q}(\mathbf{z}))) = \mathbf{0} \quad (19)$$

In Equation (19), vector $\mathbf{z}$ is obtained as a direct outcome of the nonlinear equations solver, the flows are obtained by substituting $\mathbf{z}$ in Equation (13), while the heads are obtained by substituting $\mathbf{z}$ in Equation (20).

$$\mathbf{H} = \mathbf{P}_H \cdot \begin{bmatrix} \mathbf{H} \\ \mathbf{w} \end{bmatrix} = \mathbf{P}_H \cdot \mathbf{A}^{-1} \cdot (\mathbf{H}_0 + \Delta \mathbf{H}(\mathbf{Q}(\mathbf{z}))) \quad (20)$$

where $\mathbf{P}_H = [\mathbf{I}^{nNodes \times nNodes} \mid \mathbf{0}^{nLoops \times nLoops}]$ is a permutation matrix to extract $\mathbf{H}$ from $\begin{bmatrix} \mathbf{H} \\ \mathbf{w} \end{bmatrix}$.

Noteworthy that the inversion of the matrix $\mathbf{A}$ should be performed one time for a specific network, thus the same $\mathbf{A}^{-1}$ is valid in case we run the network with different loading conditions as done in extended period analysis.

4 Application

4.1 Simulation

The proposed matrix completion technique for simulating WDNs is tested over an illustrative example and five benchmark networks. The solution obtained from the proposed methods is compared to: (a) H-Q formulation; (b) Qchord formulation and (c) Null-Space formulation. All the computations were carried out in Matlab R2020a with a processor of Intel® Core™ i7-10510 CPU @ 1.80GHz 2.30 GHz and 16.0 GB RAM. All the formulations are solved using the *fsolve* solver for system of nonlinear equations which is shipped with the optimization toolbox of Matlab. The results of the illustrative example are detailed in Appendix A.

We performed hydraulic simulation for five networks which are available from EPANET-MATLAB Toolkit (Eliades et al. 2016). A summary of the network characteristics is given in Table 2. The initial guess of each run is sampled from uniform distribution which ranges in practical bounds on the variables (e.g. maximum flow cannot exceed total demand).

Table 2 Benchmark Networks

Network	Nodes	Links
KL	935	1274

Balerna	443	454
FossPoly1	36	58
Hanoi	31	34
Jilin	27	34

Fig. 1 shows the average Maximum Absolute Error (MAE) of the flows obtained from 20 different initial guesses. The bars show the average MAE, compared to the solution from EPANET, while the error intervals show the minimum and maximum MAE for each network and for each method. In comparison to H-Q method, satisfactory results (below $0.07 \text{ m}^3\text{hr}^{-1}$ in all runs) observed in the Completion method in all the networks. Examining the average MAE, Null-Space and Qchord methods obtained the largest MAE in Balerna, KL, and Hanoi networks. Despite the inferior estimation of Completion method in FossPoly1 network, the estimation is yet acceptable. Examining the minimum and maximum MAE shows that the H-Q method obtained a very narrow interval between minimum and maximum values in all networks. This indicates that the H-Q method is the most robust for changing in the initial guess. Noteworthy that the minimum MAE of Completion method is the same as the minimum MAE in the H-Q method in all the networks. On the other hand, Null-Space and Qchord could not converge to a good solution in any of the runs in Balerna network.

Fig. 1 Average, minimum and maximum MAE of the flows for benchmark networks

While the H-Q method may be superior in estimation accuracy and stability, it might be inferior in terms of computational efficiency because of the need to update both the flow and head variables. To examine the computation time, Fig. 2 shows the average, minimum and maximum CPU time required for each network.

The results show that the computation time for the H-Q formulation is the highest among all method in all network. For example, noting the logarithm scale in the y-axis, the Completion formulation is 7 times faster than the H-Q formulation in the KL network. The Completion and Null-Space formulations produced the fastest results in average. However, comparing the Null-Space and Completion formulations in terms of accuracy (Fig. 1), tips the scale toward the Completion formulation since it produced smaller average MAE in large networks such as KL and Balerna. For example, in Balerna network both the Null-Space and Completion solved the problem with about 2 seconds while the H-Q method required an average solution time of 200 seconds. However, the MAE in the Completion formulation is 0.03 CMH compared to 1000 CMH in the Null-Space formulation.

Fig. 2 Average, minimum and maximum CPU time for benchmark networks

4.2 Optimization

The results show that the Completion method holds advantages compared to other methods for solving the hydraulic simulation problem. In fact, the benefit for this method, due to its feature of size reduction, is further highlighted when solving optimization problems that involve hydraulic simulation in the constraints. The proposed formulation can facilitate full elimination of both Q-H variables, thus reducing the number of the decision variables significantly in the optimization model.

Many optimization problems in the water distribution systems literature involve hydraulic simulation; the most typical are the optimal design problem (Alperovits and Shamir, 1977), optimal operation problem (Zessler and Shamir, 1989). In the former, the decision maker seeks optimal sizing of components (e.g. pipes, pump stations, tanks, etc.), while in the latter problem, the decision maker seeks optimal scheduling of pumps to minimize operation cost. Many studies used simulation-optimization framework for problems that involve hydraulic

simulations. In this framework, meta-heuristic technique, such as genetic algorithm (Savic and Walters 1997) is coupled with available hydraulic simulation tools (e.g. EPANET) to perform objective function and constraints evaluation. As such, following this approach does not require explicit formulation of the hydraulic simulation equations that involves Q-H variables. Nonetheless, with advancement of new analytical optimization solvers, derivative-based optimization methods are regained research interest for their computational efficiency and their independence of parameter tuning processes (Mala-Jetmarova et al. 2017; Qiu et al. 2020). Unlike the simulation-optimization framework, when using analytical optimization solvers, the formulation of the optimization problem should explicitly include the hydraulic simulation within the problem's constraints. The straightforward approach for doing so is to include H-Q variables as decision variables in the optimization problem along with the mass and energy balance Equations (6)-(7) as sets of equality constraints. For example, under these settings, the optimal design problem for a gravitational WDS could be formulated as in Equation (21)

$$\begin{aligned}
& \min_{\mathbf{r}, \mathbf{Q}^s, \mathbf{H}^s} f(\mathbf{r}) \\
& \text{s.t.} \\
& \mathbf{A}^T \cdot \mathbf{Q}^s = \mathbf{q}^s \quad \forall s = 1 \dots N_s \\
& \mathbf{A} \cdot \mathbf{H}^s = \mathbf{H}_0 + \mathbf{r} \circ \mathbf{L} \circ \mathbf{Q}^s \circ |\mathbf{Q}^s|^{0.852} \quad \forall s = 1 \dots N_s \\
& \mathbf{H}_{\min} \leq \mathbf{H}^s \leq \mathbf{H}_{\max} \quad \forall s = 1 \dots N_s \\
& \mathbf{r}_{\min} \leq \mathbf{r} \leq \mathbf{r}_{\max}
\end{aligned} \tag{21}$$

where $\mathbf{r} \in \mathbb{R}^{nLinks \times 1}$ is the per-unit-length resistance; $\mathbf{Q}^s, \mathbf{H}^s$ vectors of flows and heads for demand load condition s ; N_s number of demand load conditions; $\mathbf{L} \in \mathbb{R}^{nLinks \times 1}$ vector of pipes length; $\mathbf{H}_{\min}, \mathbf{H}_{\max}$ are vectors for minimum and maximum nodal heads; $|\mathbf{Q}^s|$ is the element wise absolute value of vector $\mathbf{Q}^s$; $\circ$ is the element wise Hadamard product; $f(\mathbf{r})$ is the cost function. Note that there is one-to-one function that relates $\mathbf{r}$ to the diameters $\mathbf{D}$, i.e. ($\mathbf{r} = g(\mathbf{D})$) as shown in Equation 22). As such, by seeking optimal per-unit-length resistance we determine the

diameters of the pipes. Still, solving the optimization problem in terms of $\mathbf{r}$ instead of the diameters, as suggested in Equation (21), is advantageous since the formulation in terms of $\mathbf{r}$ involves less nonlinear terms. The last constraint limits the resistance within minimum ($\mathbf{r}_{\min}$) and maximum ($\mathbf{r}_{\max}$) which represent maximum and minimum allowed diameters, respectively.

$$\mathbf{r} = g(\mathbf{D}) = 1.526 \cdot 10^7 \cdot C^{-1.852} \cdot D^{-4.871} \quad (22)$$

where C is the roughness coefficient and D is diameter in centimeters.

The number of decision variables in the nonlinear optimization problem given in Equation (21) equals to $nLinks + nLinks + nNodes \cdot N_s$ ($nLinks$ of diameter variables, and $nLinks$ of flow variables and $nNodes$ of head variables for each demand loading condition). Using the Completion approach one can write an equivalent optimization problem (Equation 23), which only has $nLinks + nLoops \cdot N_s$ decision variables ($nLinks$ of diameter variables, and $nLoops$ of z variables for each demand loading condition). For increasing number of demand loads, the size of the optimization problem in Equation (23) is significantly smaller than the one in Equation (21). This reduction in optimization problem size leads to efficient solution, as will be demonstrated next. Noteworthy that the last line in Equation (23) is not an equality constraint, but it is a definition which should be substituted in the first two constraints. Hence, the optimization problem in Equation (23) involves only $\mathbf{r}$ and $\mathbf{z}^s$ for each loading condition, s , as decision variables.

$$\begin{aligned}
& \min_{\mathbf{r}, \mathbf{z}^s} f(\mathbf{r}) \\
& \text{s.t.} \\
& \mathbf{P}_w \cdot \mathbf{A}^{-1} \cdot (\mathbf{H}_0 + \Delta \mathbf{H}^s) = \mathbf{0} \quad \forall s = 1 \dots N_s \\
& \mathbf{H}_{\min} \leq \mathbf{P}_H \cdot \mathbf{A}^{-1} \cdot (\mathbf{H}_0 + \Delta \mathbf{H}^s) \leq \mathbf{H}_{\max} \quad \forall s = 1 \dots N_s \\
& \mathbf{r}_{\min} \leq \mathbf{r} \leq \mathbf{r}_{\max}
\end{aligned} \tag{23}$$

where

$$\Delta \mathbf{H}^s \equiv \mathbf{r} \circ \mathbf{L} \circ \left(\mathbf{A}^{-T} \cdot \left[\frac{\mathbf{q}^s}{\mathbf{z}^s} \right] \right) \circ \left| \mathbf{A}^{-T} \cdot \left[\frac{\mathbf{q}^s}{\mathbf{z}^s} \right] \right|^{0.852} \quad \forall s = 1 \dots N_s$$

To demonstrate the advantage of using the proposed Completion method within an optimization problem we compared the optimization problem in Equation (21) to the reduced size problem in Equation (23) on gravitational Hanoi network introduced by Fujiwara and Khang (1990). The network consists of 34 pipes and 31 demand nodes supplied by a single reservoir at a constant head of 100 (m). The minimum head is 30 (m) and the maximum head is 100 (m) at all nodes. The allowed diameter of all pipes ranges between 12-40 inches. The Hazen-Williams coefficient of all pipes is 130. The lengths of all pipes are listed in Fujiwara and Khang (1990). To create different demand loads, the demand was sampled from joint normal distribution assuming mean demand equal to the demand in Fujiwara and Khang (1990). The standard deviation is defined as 10% of the mean demand for each node. Moreover, to consider different demand patterns and correlation, the system nodes were partitioned to two demand zones: zone 1 at nodes 1:15, zone 2 at nodes 16:32. Intrazone correlation is set to 0.8 reflecting that consumers in the same zone follow similar demand pattern, whilst interzone correlation is set to -0.6 to reflect different patterns in the two zones.

Using the above configuration, Hanoi optimal design problem is solved for increasing number of demand loads ranging from 1 demand load up to 30 demand loads. This process creates different optimization problems with different sizes. The optimization problems are solved using the fmincon solver (SQP algorithm), which is shipped with the optimization toolbox of

Matlab. Since the optimization problem is non-convex, we used multistart strategy, which solves the optimization problems 50 times each. Fig. 3 shows the box and whisker plot of the optimal cost statistics for both formulations and for different problem sizes (i.e., increasing N_s). The optimal cost is analyzed for runs that terminated successfully out of 50 runs from different initial guess. The box shows the mean, median, the first and third quartile while the whiskers shows the minimum and maximum cost. The figure shows that the mean optimal cost achieved by the reduced problem (i.e., r-z formulation of the completion formulation) in Equation (23) is generally better than the r- H-Q formulation in Equation (21). However, the true exam is the best optimal cost, again the results show that the r-z formulation obtained superior results, except when solving the problem with one loading condition in which the best cost of the r-z formulation is slightly outperformed. Fig. 4 shows the number of runs which succeeded to declare successful termination. The results show that the r-H-Q formulation failed to reach declare successful termination in more than half of the runs when the number demand loads is greater than 10. For example, when $N_s = 30$ the r-Q-H formulation converged to a solution only in 7 out 50 runs while the r-z formulation converged to a solution in 39 out of 50 runs. This could be explained by the significant difference in the size of the optimization problems. For the r-z formulation, when $N_s = 30$, the optimization problem involves $34 + 3 \cdot 30 = 124$ decision variables whilst the r-H-Q formulation involves $34 + (32 + 34) \cdot 30 = 2014$ decision variables. This reduction in the problem size is also beneficial in the mean CPU time required by the optimization solver. Fig. 5 show the mean CPU time of the successful runs, while Fig. 6 shows the mean CPU time of the failed runs. The results show that the reduced r-z formulation can obtain the optimal solution faster than the r-H-Q formulation. While the difference in the CPU time is negligible for small size optimization problems (i.e., $N_s = 1,5$), the CPU time saving can reach factor 15 for large optimization problems with $N_s = 30$.

Fig. 3 Box and whisker plot of the optimal cost statistics for different demand loads

Fig. 4 The number of runs with successful termination for different demand loads

Fig. 5 The mean CPU time of the successful runs for different demand loads

Fig. 6 The mean CPU time of the failed runs for different demand loads

5 Conclusion

In this paper, we proposed a novel formulation of the mass and energy equation for simulating WDNs. The proposed method utilizes the matrix completion technique to construct a reduced-size nonlinear system of equations that guarantees both mass and energy conservation. Unlike in former formulations that rely on the topology of the network, in the proposed method we employ a matrix completion technique in which arbitrary entries are added to the system of equations to obtain invertible squared matrices that facilitate the elimination of both the H and Q variables in the problem. The advantage of the proposed method in terms of accuracy and run time in comparison to other formulations is demonstrated through five benchmark networks. As compared to the full H - Q formulation (the method used in EPANET), the proposed method is superior in terms of run time while maintaining acceptable accuracy. To further examine the advantages of the method, we developed an optimization model that utilizes the proposed formulation for modeling the hydraulic variables in the system. By virtue of the elimination of the Q - H variables, a compact optimization problem is obtained, reducing significantly the size of the decision variable. Therefore, better optimal solution quality is obtained in terms of the objective function, run time, and rate of successful optimization runs. Additional salient features of the proposed method include that (1) it is possible to write infinitely many equivalent auxiliary problems that guarantee mass and energy conservation for the same network and (2) the construction of the auxiliary problem can be efficiently updated in the case where the network is run with different loading conditions, since many of the algebraic operations (e.g., matrices inversion) are independent of the loading conditions. The former feature suggests that different completion strategies may result in different formulations for the same problem. Here, we used a completion strategy of randomly sampled entries in the completion matrix. Nonetheless, future research should consider different completion strategies that rely on the network characteristics to enhance its performance further.

Appendix A

Illustrative Example

A synthetic network was generated as shown in Fig. 7. It consists of 7 pipes, 4 nodes and one fixed head reservoir. The lengths of the pipes are 1000 meters each, the diameters are 200 millimeters, and the Hazen-Williams coefficient is 100 in all pipes. The head at the reservoir is fixed at 10 meters and the elevation of all network's nodes is 0 meters. The nodal demand is 20 CMH at all nodes. Fig. 8 shows the matrices definitions required by the completion technique as detailed in Equations (11-20).

Fig. 7 Illustrative Example

To analyze the robustness of the methods, we performed 100 runs where each run starts from different initial guess. The initial guess of each run was sampled from uniform distribution between -100 and 100 ($20 \times 5 = 100$ CMH). The average and the standard deviation (STD) of

the MAE of the flows as obtained from the 100 runs are detailed in Table 3. The standard deviation is presented as an indication for the sensitivity of each formulation to the initial guess. The results in Table 3 show that the H-Q method outperformed other methods in terms of accuracy of estimation of flows as it has lower average MAE. Comparing the other methods shows that the completion formulation outperforms other methods in average MAE and it is more robust since the standard deviation of the MAE is lower than the Qchord and Null-Space formulations.

$$\begin{aligned}
 \tilde{\mathbf{A}}^T &= \begin{bmatrix} \mathbf{A}^T \\ \mathbf{A}_c \end{bmatrix} = \begin{bmatrix} -1 & 1 & 1 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ \hline 0.91 & 0.59 & 0.68 & 0.45 & 0.14 & 0.95 & 0.80 \\ 0.63 & 0.39 & 0.68 & 0.69 & 0.36 & 0.08 & 0.74 \\ 0.99 & 0.33 & 0.37 & 0.35 & 0.79 & 0.90 & 0.28 \end{bmatrix} \\
 \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{P}_r} \cdot \underbrace{\begin{bmatrix} -0.23 & 0.18 & -0.01 & -0.42 & 0.13 & -0.08 & 0.53 \\ -0.45 & -0.79 & -1.12 & -1.02 & 0.56 & 0.30 & -0.37 \\ 1.00 & 1.00 & 1.00 & 1.00 & 0.00 & 0.00 & 0.00 \\ -0.35 & -0.75 & -0.20 & -0.67 & -0.44 & 0.79 & 0.31 \\ 0.10 & 0.04 & -0.08 & 0.35 & -1.01 & 0.49 & 0.67 \\ 0.12 & -0.07 & 0.19 & 0.25 & 0.57 & -0.88 & 0.23 \\ -0.22 & 0.03 & -0.11 & 0.40 & 0.43 & 0.38 & -0.90 \end{bmatrix}}_{\tilde{\mathbf{A}}^{-1}} \cdot \underbrace{\begin{bmatrix} 0 \\ 0 \\ 10 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{H}_r} + \underbrace{\begin{bmatrix} R_1 \cdot Q_1(z) \cdot |Q_1(z)|^{0.852} \\ R_2 \cdot Q_2(z) \cdot |Q_2(z)|^{0.852} \\ R_3 \cdot Q_3(z) \cdot |Q_3(z)|^{0.852} \\ R_4 \cdot Q_4(z) \cdot |Q_4(z)|^{0.852} \\ R_5 \cdot Q_5(z) \cdot |Q_5(z)|^{0.852} \\ R_6 \cdot Q_6(z) \cdot |Q_6(z)|^{0.852} \\ R_7 \cdot Q_7(z) \cdot |Q_7(z)|^{0.852} \end{bmatrix}}_{\Delta \mathbf{H}(\mathbf{Q}(z))} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow \text{SOLVE for } \mathbf{z}^* \\
 \underbrace{\begin{bmatrix} 20 \\ 20 \\ 20 \\ 20 \\ z_1^* \\ z_2^* \\ z_3^* \end{bmatrix}}_{\mathbf{Q}(\mathbf{z}) = \tilde{\mathbf{A}}^{-T} \cdot} \cdot \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}}_{\mathbf{P}_r} \cdot \underbrace{\begin{bmatrix} \mathbf{H} \\ \mathbf{w} \end{bmatrix}}_{\mathbf{H}}
 \end{aligned}$$

Fig. 8 The matrices generated in the illustrative example

Table 3 MAE results in the illustrative example.

	MAE ($m^3 hr^{-1}$)	
Method	Average	STD
H-Q	0.003×10^{-3}	0.007×10^{-3}
Qchord	0.960×10^{-3}	1.456×10^{-3}
Null-Space	1.565×10^{-3}	2.300×10^{-3}
Completion	0.030×10^{-3}	0.139×10^{-3}

Declarations

Funding

This study was funded by the Israeli Water Authority

Conflict of interest/Competing interests

Authors have no conflict of interest to declare that are relevant to the content of this article.

Availability of data and material

Data are retrieved from <https://github.com/ShenWang9202/GP4WFP1/tree/master/networks>

Code availability

Matlab code is available from the author upon request.

Authors' contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Alaa Jamal and Mashor Housh. The first draft of

the manuscript was written by Alaa Jamal and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

References

- Abraham, E., & Stoianov, I. (2016). An efficient null space inexact Newton method for hydraulic simulation of water distribution networks. *arXiv preprint arXiv:1602.02630*.
- Basha, H. A., & Kassab, B. G. (1996). Analysis of water distribution systems using a perturbation method. *Applied mathematical modelling*, 20(4), 290-297.
- Benzi, M., Golub, G. H., & Liesen, J. (2005). Numerical solution of saddle point problems. *Acta numerica*, 14, 1-137.
- Boulos, P. F., Lansey, K. E., & Karney, B. W. (2006). Comprehensive water distribution systems analysis handbook for engineers and planners. American Water Works Association.
- Elhay, S., Simpson, A. R., Deuerlein, J., Alexander, B., & Schilders, W. H. (2014). Reformulated co-tree flows method competitive with the global gradient algorithm for solving water distribution system equations. *Journal of Water Resources Planning and Management*, 140(12), 04014040.
- Epp, R., & Fowler, A. G. (1970). Efficient code for steady-state flows in networks. *Journal of the hydraulics division*, 96(1), 43-56.

Fuchs-Hanusch, D., Gangl, G., Kornberger, B., Kölbl, J., Hofrichter, J., & Kainz, H. (2008). PIREM—Pipe rehabilitation management developing a decision support system for rehabilitation planning of water mains. *Water Practice and Technology*, 3(1).

Giustolisi, O., Laucelli, D., Berardi, L., & Savić, D. A. (2012). Computationally efficient modeling method for large water network analysis. *Journal of Hydraulic Engineering*, 138(4), 313-326.

Housh, M. (2011). Optimal Multi-Year Management of Regional Water Resources Systems under Uncertainty. Technion-Israel Institute of Technology, Faculty of Civil and Environmental Engineering.

Housh, M., & Ohar, Z. (2017). Integrating physically based simulators with event detection systems: Multi-site detection approach. *Water research*, 110, 180-191.

Kapelan, Z. S., Savic, D. A., & Walters, G. A. (2005). Multiobjective design of water distribution systems under uncertainty. *Water Resources Research*, 41(11).

Klise, K. A., Bynum, M., Moriarty, D., & Murray, R. (2017). A software framework for assessing the resilience of drinking water systems to disasters with an example earthquake case study. *Environmental modelling & software*, 95, 420-431.

Lansey, K. E., & Mays, L. W. (1989). Optimization model for water distribution system design. *Journal of Hydraulic Engineering*, 115(10), 1401-1418.

Liu, K. T. H. (1969, August). The numerical analysis of water supply networks by digital computers. In *Thirteenth congress of the international association for hydraulic research* (Vol. 1, pp. 36-43).

Martin, D. W., & Peters, G. (1963). The application of Newton's method to network analysis by digital computer. *Journal of the institute of Water Engineers*, 17(2), 115.

Martínez-Solano, F. J., Iglesias-Rey, P. L., & Arce, S. X. M. (2017). Simultaneous Calibration of Leakages, Demands and Losses from Measurements. Application to the Guayaquil Network (Ecuador). *Procedia Engineering*, 186, 397-404.

Moosavian, N. (2017). Multilinear method for hydraulic analysis of pipe networks. *Journal of Irrigation and Drainage Engineering*, 143(8), 04017020.

- Moosavian, N., & Roodsari, B. K. (2014). Soccer league competition algorithm: a novel meta-heuristic algorithm for optimal design of water distribution networks. *Swarm and Evolutionary Computation*, 17, 14-24.
- Nielsen, H. B. (1989). Methods for analyzing pipe networks. *Journal of Hydraulic Engineering*, 115(2), 139-157.
- Perelman, L., Housh, M., & Ostfeld, A. (2013). Robust optimization for water distribution systems least cost design. *Water Resources Research*, 49(10), 6795-6809.
- Qiu, M., Housh, M., & Ostfeld, A. (2020). A Two-Stage LP-NLP Methodology for the Least-Cost Design and Operation of Water Distribution Systems. *Water*, 12(5), 1364.
- Rahal, H. (1995). A co-tree flows formulation for steady state in water distribution networks. *Advances in Engineering Software*, 22(3), 169-178.
- Rossman, L. A. (2000). EPANET 2: users manual.
- Savic, D. A., & Walters, G. A. (1997). Genetic algorithms for least-cost design of water distribution networks. *Journal of water resources planning and management*, 123(2), 67-77.
- Yang, D.. Solution theory for systems of bilinear equations. B.Sc. Thesis; College of William and Mary; 2011. <https://scholarworks.wm.edu/cgi/viewcontent.cgi?article=1424&context=honorsthesis>
- Todini, E., & Pilati, S. (1988). A gradient algorithm for the analysis of pipe networks. In *Computer applications in water supply: vol. 1---systems analysis and simulation* (pp. 1-20).
- Wood, D. J. (1980). User's manual computer analysis of flow in pipe networks including extended period simulations. *Unpublished) Technical Report, Department of Civil Engineering, University of Kentucky, Lexington (KY)*.
- Wood, D. J., & Charles, C. O. (1972). Hydraulic network analysis using linear theory. *Journal of the Hydraulics division*, 98(7), 1157-1170.
- Wood, D. J., & Funk, J. E. (1993). Hydraulic analysis of water distribution systems. *Water Supply systems, state of the art and future trends*, E. Cabrera and F. Martinez, eds., Computational Mechanics Publications, Southampton, 41-85.

Figures

Figure 1

Average, minimum and maximum MAE of the flows for benchmark networks

Figure 2

Average, minimum and maximum CPU time for benchmark networks

Figure 3

Box and whisker plot of the optimal cost statistics for different demand loads

Figure 4

The number of runs with successful termination for different demand loads

Figure 5

The mean CPU time of the successful runs for different demand loads

Figure 6

The mean CPU time of the failed runs for different demand loads

Figure 7

Illustrative Example

$$\mathbf{A}^0 = \left[\frac{\mathbf{A}^T}{\mathbf{A}_c} \right]$$

$$\left[\begin{array}{cccc|ccc} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 4 & 4 & 4 & 2 & 4 & 4 & 3 \end{array} \right]$$

$$Q(z) = \overset{0}{A} \overset{0}{z}.$$

The matrices generated in the illustrative example