Approximating the model of a water distribution network as a Markov decision process

Misra, Rahul; Wisniewski, Rafal; Kallesøe, Carsten S.

Published in:
IFAC-PapersOnLine

DOI (link to publication from Publisher):
10.1016/j.ifacol.2022.09.107

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 55-20 (2022) 271–276

Approximating the model of a water distribution network as a Markov decision process ⋆

Rahul Misra * **Rafał Wisniewski** * **Carsten S. Kallesøe** *,**

* *Automation and Control section, Department of electronic systems, Aalborg University, Aalborg, Denmark (e-mail: {rmi, raf, csk}@es.aau.dk).*

** *Grundfos A/S, Poul Due Jensens Vej, Bjerringbro, Denmark (e-mail: ckallesoe@grundfos.com).*

Abstract: In this paper, our objective is to convert the model of a water distribution network described via Stochastic differential equations (SDE) into a Markov decision process (MDP). The motivation behind this work is that while MDP's represents the underlying dynamics for dynamic programming and reinforcement learning, the actual underlying model is best described via differential equations, and therefore, we would like to convert the SDE to MDP. We have applied Kushner's Markov chain approximation (MCA) method and verified it using a novel modified Monte Carlo method which can be considered as an alternative to the well-known Kushner's MCA. Both the methods approximate the value function and simulation studies show that the obtained value functions from both the methods converge to almost the same value.

Keywords: Modeling for control optimization, Markov decision process, Stochastic differential equations, Dynamic programming, Markov chain approximation, Monte Carlo methods.

1. INTRODUCTION

Efficient pressure management in a Water Distribution Network (WDN) is a complex control problem since it entails an inherently multi-input, multi-output system with control objectives of ensuring supply with minimal variance in pressure at demand side while ensuring energy efficiency of supply pumps. If the WDN is used for drinking water distribution than we would also need to control water quality (see Polycarpou et al. (2002) and the references therein). The WDN considered in this work consists of suppliers and consumers which are connected together by a piping network. Such a WDN can be modeled using graph theory which represents the topology of the network connecting individual components as stated in Tahavori et al. (2012). The uncertainty in consumption pattern is the reason for the stochastic nature of a water distribution network and therefore, we have used a stochastic differential equations (SDE) where the diffusion matrix takes into account the uncertainty due to demand side consumption.

For solving stochastic control problems, Dynamic Programming (DP) is one of the fundamental mathematical tools and forms the basis of modern reinforcement learning (RL) algorithms such as Q-learning, SARSA, policy gradient, actor-critic etc. (see Sutton and Barto (2018) and Bertsekas (2005) for an introduction to dynamic programming and reinforcement learning algorithms). A common feature of all these algorithms is that they assume that the underlying system is a Markov decision process (MDP) or its generalization such as a Stochastic game in Filar and Vrieze (2012). All the aforementioned settings have a finite number of states which makes it possible to represent the value of each state in a tabular representation. Bellman's optimality equation can than be used to find the optimal states and the corresponding sequence of control actions constitute the optimal control policy. However in a practical setting such as a water distribution network, the dynamics are modeled using differential equations derived from physical laws such as mass or energy conservation. The state space of such equations is infinite, thereby making traditional DP and RL methods not applicable in such settings. To overcome this limitation, function approximation based techniques such as least squares or a neural networks are used for approximating the value function or the control strategy or both simultaneously. Function approximation based RL algorithms have some inherent challenges such as non-convergence to the target in the case of off-policy learning with bootstrapping, overfitting and instability in the case of neural networks (see chapters 9, 11 and references therein from Sutton and Barto (2018)). The method of Markov chain approximation was developed in Kushner and Dupuis (2001) as a numerical method for solving stochastic control problems and can be considered as an alternative to standard functional approximation techniques. This idea has previously been considered in Munos and Bourgine (1997) as *finite-differences RL*. However, there is no existing literature on practical applications of *finite-differences RL* and this

⋆ This work is supported by Poul Due Jensen Foundation (Grundfos Foundation) under the project SWIfT (Smart water infrastructure).

10.1016/j.ifacol.2022.09.107

forms a motivation for this work. The key contributions of this work are summarized as follows:

- Markov chain approximation is applied on a novel SDE model of a WDN so as to obtain an MDP.
- A novel Monte Carlo based method is developed as an alternative to Markov chain approximation.
- To the best of authors knowledge, this is the first attempt at modeling a WDN as an SDE.

The rest of the paper is organized as follows. In the next Section, we present a model of WDN described via SDE's. In Section 3, we approximate the value function for a fixed policy and derive transition probabilities using the drift and diffusion terms of the model presented in Section 2. In Section 4, we derive the transition probabilities using a modified Monte Carlo method. In Section 5 we compare the results obtained from Section 3 and 4. Lastly, we present our conclusions and future work.

2. MODELING OF WDN USING SDE

A general model of the WDN with $\mathcal{N}$ controllers is defined by state equations of the following type

$$dx = \left(\mathrm{f}(x) + \sum_{k=1}^{\mathcal{N}} \mathrm{g}^k(x)u^k\right)dt + \sigma(x)d\mathrm{w}, \qquad (1)$$

where $x \in X \subseteq \mathbb{R}^n$ is a stochastic process which represents the free flows in the WDN, $u \in U \subseteq \mathbb{R}^p$ represents the pressure control action due to pumps, $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^p$ represent the drift, $\sigma : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^l$ represents the diffusion with $a(\cdot) = \sigma(\cdot)\sigma(\cdot)^T$ being the corresponding diffusion matrix, w represents the standard Wiener process on $\mathbb{R}^l$. For simplicity, we shall consider the WDN network shown in fig. 1 with 2 controllers (2 pumping stations) although, the methods developed in this paper are extendable to complex controlled systems with more control inputs. The aim of the pumping stations 1

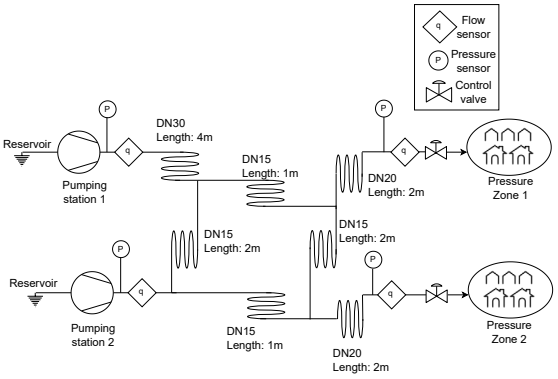

Fig. 1. Process and Instrumentation diagram

and 2 in fig. 1 is to minimize the pressure variation at pressure zones 1 and 2. The WDN is modeled by a directed graph $\Gamma = \{N, E\}$ in fig. 2 where $N = \{n_1, \cdots, n_{11}\}$ represents the nodes or vertices and $E = \{e_1, \cdots, e_{11}\}$ represents the edges where components of WDN such as pipes, pumps and valves are connected. The pumps in fig. 1 are represented as edges e_{10} and e_{11} in fig. 2 and the demands are represented by edges e_8 and e_9. We begin by defining the incidence matrix H which encodes the interconnections between different components of the WDN as shown in the fig. 1.

$$H_{i,j} = \begin{cases} -1, & \text{if the } j^{th} \text{ edge is entering } i^{th} \text{ node,} \\ 0, & \text{if the } j^{th} \text{ edge is not connected to the } i^{th} \text{ node,} \\ 1, & \text{if the } j^{th} \text{ edge is leaving } i^{th} \text{ node,} \end{cases} \qquad (2)$$

and the cycle matrix B which encodes the information about the edges which belong to cycles (or loops) and their orientation.

$$B_{i,j} = \begin{cases} -1, & \text{if the } j^{th} \text{ edge belongs to the } i^{th} \text{ cycle and their directions disagree,} \\ 0, & \text{if the } j^{th} \text{ edge does not belong to the , } i^{th} \text{ cycle,} \\ 1, & \text{if the } j^{th} \text{ edge belongs to the } i^{th} \text{ cycle and their directions agree.} \end{cases} \qquad (3)$$

The model is defined with following assumptions.

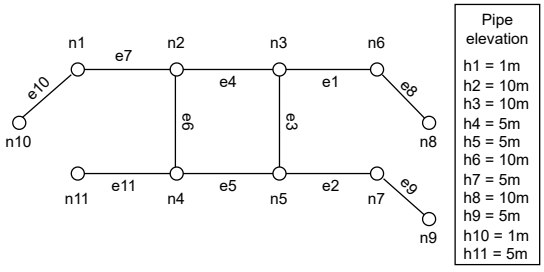

Fig. 2. Graph of the WDN in fig. 1 with pipe elevation $\bar{z}$.

Assumption 2.1. The graph Γ is a connected graph.

Assumption 2.2. We require existence and uniqueness of solutions of (1). Let Ω represent the sample space, $\mathcal{F}$ represent the event space and P represent the probability function. Strong existence holds if for a given probability space $(\Omega, \mathcal{F}, P)$, a filtration $\mathcal{F}_t$, an $\mathcal{F}_t$-Wiener process w and an $\mathcal{F}_0$- measurable initial condition $x(0)$, there exists an $\mathcal{F}_t$-adapted process $x(t)$ satisfying (1) for all $t \geq 0$. Furthermore, uniqueness holds if for any two sample paths $x_1(t), x_2(t)$, $P\{x_1(0) = x_2(0)\} = 1 \implies P\{x_1(t) = x_2(t) \forall t \geq 0\} = 1$.

We shall now derive the model of the WDN in a deterministic sense for simplicity. Thereafter, the model is extended to include stochastic noise on the states as consumption of water is uncertain. Let $\mathcal{T}$ be a arbitrary spanning tree of the graph Γ and let $\mathcal{C}$ be the chords of the same. $H_{\mathcal{T}}$ and $H_{\mathcal{C}}$ represent the incidence matrix of the spanning tree $\mathcal{T}$ and corresponding chords $\mathcal{C}$ respectively. For a WDN with graph Γ, H can be partitioned as $H = [\, H_{\mathcal{C}} \;\; H_{\mathcal{T}} \,]$. We now need to choose a reference node. The reference node is chosen to be one of the supply nodes (n_1 or n_4) since it is only at the nodes with non-zero demands that water can flow in or out of the network. In this setup, we have chosen the reference node as n_1 and defined $\bar{H}_{\mathcal{T}}$ as the reduced incidence matrix of $\mathcal{T}$ (obtained by removing the row corresponding to the reference node) and $\bar{H}_{\mathcal{C}}$ as the corresponding chord matrix. We then define the free flow vectors $q_{\mathcal{C}}$ (flow in one of the chords for example e_6) and $d \in \mathbb{R}^{n-1}$ (external flows e_8, e_9 and e_{11} excluding the flow e_{10} in the reference node),

$$q = B^T q_{\mathcal{C}} + B_d^T F d, \qquad (4)$$

where $B = \begin{bmatrix} I & -\bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} \end{bmatrix}$ is the standard cycle matrix ($\bar{H}_{\mathcal{T}}$ is always invertible as given in Deo (2017)), $B_d = \begin{bmatrix} 0 & \bar{H}_{\mathcal{T}}^{-T} \end{bmatrix}$ represents the cycle matrix which includes free flows and F is a matrix of 1's and 0's which selects the consumption nodes (n_6 and n_7) in Γ (fig. 2). The individual components of a water distribution network are characterized by the pressure drop across them (analogous to the voltage drop across resistors, inductors and capacitors in an electrical circuit). For a pipe element r in the network having length L (m), cross-section area A (m^2) with water density ρ (kg/m^3), the pressure drop (bar) across its ends i, j is given by the following equation.

$$p_i - p_j = J\dot{q_r} + Rf(q_r) - \rho g \Delta z_r, \tag{5}$$

where $J = \frac{L\rho}{A}$, $\dot{(\cdot)}$ notation represents the usual time derivative $\frac{d(\cdot)}{dt}$, q_r is the flow (m^3/hr) in pipe r, g is the gravitational constant, R is a diagonal matrix which represents the friction factor of the pipes, $f_i(q) = |q_i|\, q_i$ and $\Delta z_r = z_i - z_j$ is the difference in geodesic level of the pipe. The term $f_i(q)$ represents the nonlinear pipe resistance to flow q and it preserves the direction of flow (see Jensen and Wisniewski (2011)). The parameters of friction factor R can be calculated as per Chapter 2 in Swamee and Sharma (2008). Equation (5) is a consequence of Newton's second law of motion and is derived in equation 2.4 in Pétursson Geir Bjarni (2015). The pump is modeled as a centrifugal pump and can be represented as a positive pressure difference u^k across the k^{th} pump as shown in the following equation,

$$p_i - p_j = u^k. \tag{6}$$

In the sequel, u without superscript indicates a vector consisting of all $\mathcal{N}$ pump controls. We can now define the state vector for (1) consisting of free flows as $x = [q_{\mathcal{C}} \quad d]^T$. The total pressure drop Δp across all the components of WDN is obtained by combining (5) and (6) as follows.

$$\Delta p = J\dot{q} + Rf(q) - \rho g \Delta z + u. \tag{7}$$

Recall Kirchhoff's mesh law which is expressed as $B\Delta p = 0$ which is combined with (7) as follows.

$$0 = BJ\dot{q} + BRf(q) - B\Delta z + B\rho g G u, \tag{8}$$

where q is a linear function of q_c and d as per (4) and G is a matrix of 1's and 0's which selects the edges in Γ (fig. 2) on which the controllers are connected. The rest of the derivation will consider pressure at the consumer side which is obtained by

$$\bar{p} = \bar{H}_T^{-T} \Delta p_T = \bar{H}_T^{-T} (J_T \dot{q}_T + R_T f_T(q_T) - \rho g \Delta z_T + G_T u). \tag{9}$$

The matrices J, G and R can also be partitioned as $J = [\, J_{\mathcal{C}} \;\; J_{\mathcal{T}} \,]$, $G = [\, G_{\mathcal{C}} \;\; G_{\mathcal{T}} \,]$ and $R = [\, R_{\mathcal{C}} \;\; R_{\mathcal{T}} \,]$ respectively. The pressure at all nodes open to atmosphere $F^T \bar{p}$ can be set to 0, where F is a matrix of 1's and 0's which selects the nodes open to the atmospheric pressure. This means that

$$\begin{aligned} 0 = F^T \bar{p} = F^T \bar{H}_{\mathcal{T}}^{-T} J_T \dot{q}_T + F^T \bar{H}_{\mathcal{T}}^{-T} R_{\mathcal{T}} f_{\mathcal{T}}(q_{\mathcal{T}}) \\ - \rho g F^T \bar{H}_{\mathcal{T}}^{-T} \Delta z_{\mathcal{T}} + F^T \bar{H}_{\mathcal{T}}^{-T} G_{\mathcal{T}} u, \end{aligned} \tag{10}$$

which together with definition of $B = \begin{bmatrix} I & -\bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} \end{bmatrix}$ and $B_d = \begin{bmatrix} 0 & \bar{H}_{\mathcal{T}}^{-T} \end{bmatrix}$ gives us

$$\begin{aligned} \underbrace{\begin{bmatrix} BJB^T & BJB_d^T F \\ F^T B_d J B^T & F^T B_d J B_d^T F \end{bmatrix}}_{J_{ex}} \begin{bmatrix} \dot{q}_C \\ \dot{d} \end{bmatrix} = - \underbrace{\begin{bmatrix} I & -\bar{H}_C^T \bar{H}_T^{-T} \\ 0 & F^T \bar{H}_T^{-T} \end{bmatrix}}_{A_{ex}} \\ \left(\begin{bmatrix} R_C & 0 \\ 0 & R_T \end{bmatrix} f(q) - \rho g \begin{bmatrix} G_C \\ G_T \end{bmatrix} u \right) + \rho g \underbrace{\begin{bmatrix} 0 \\ F^T \end{bmatrix}}_{F_{ex}} (\bar{z} - z_0) \end{aligned} \tag{11}$$

The pump and pipe dynamics can be combined together as drift terms f and g in (1) as follows.

$$\begin{aligned} \mathrm{f}(x(t)) + \sum_{k=1}^{\mathcal{N}} \mathrm{g}^k(x(t)) u^k = -{J_{ex}}^{-1} (A_{ex}(Rf(q) \\ - \rho g \sum_{k=1}^{\mathcal{N}} G^k u^k) + \rho g F_{ex} (\bar{z} - z_0)), \end{aligned} \tag{12}$$

where G^k is the column of matrix G corresponding to player k, F_{ex} is a matrix of 1's and 0's which selects the edges in the graph Γ (fig. 2) on which the demands are connected (in this case the pressure zones), $\bar{z}$ is the physical height of each node as shown in fig. 2, z_0 is the height of reference node. Note that, the matrix J_{ex} and A_{ex} combine the incidence matrix H and cycle matrices B and B_d with component models. The stochastic nature of consumption at nodes n_8 and n_9 in fig. 2 is captured as a Wiener process.

Assumption 2.3. The diagonal terms of diffusion matrix are dominant over off-diagonal terms as follows

$$a_{ii}(x) - \sum_{j:j \neq i} |a_{ij}(x)| \geq 0. \tag{13}$$

This assumption is valid for WDN as in practice only the free flows at end-user edges (e_8 and e_9 in fig. 2) are correlated with each other. The correlation between free-flows in the other edges are accounted for in the drift term (12). Therefore, the off-diagonal terms are negligible relative to diagonal terms and can be ignored. The pressure measurement at a selected node in the WDN is derived from equations (5) and (6) by replacing p_i, p_j with $\bar{p}, p_0$ as follows.

$$\begin{aligned} C(\bar{p} - p_0) = C \bar{H}_{\mathcal{T}}^{-T} J_{\mathcal{T}} \dot{q}_{\mathcal{T}} + C \bar{H}_{\mathcal{T}}^{-T} R_{\mathcal{T}} (q_{\mathcal{T}}) \\ - \rho g C (\bar{z} - z_0) - \rho g C \bar{H}_{\mathcal{T}}^{-T} G u, \end{aligned} \tag{14}$$

where C is a matrix of 1's and 0's which selects the measured quantities in Γ (fig. 2), $q_{\mathcal{T}}$ is the flow in the spanning tree $\mathcal{T}$ of Γ in fig. 2 (it includes both the chord flow q_c and the consumer flow d), $R_{\mathcal{T}}$ is the associated friction factor matrix and p_0 is the reference pressure. Now, let Π_T project the component flows q onto the tree flow $q_{\mathcal{T}}$, such that $\dot{q}_{\mathcal{T}} = \Pi_{\mathcal{T}} \dot{q}$. This means that

$$\begin{aligned} \bar{p} - p_0 = \bar{H}_{\mathcal{T}}^{-T} J_{\mathcal{T}} \Pi_{\mathcal{T}} \left(B^T \dot{q}_{\mathcal{C}} + B_d^T F \dot{d} \right) + \\ \bar{H}_{\mathcal{T}}^{-T} R_{\mathcal{T}} \left(\Pi_{\mathcal{T}} \left(B^T q_{\mathcal{C}} + B_d^T F d \right) \right) \\ - \rho g (\bar{z} - z_0) - \rho g \bar{H}_{\mathcal{T}}^{-T} G u. \end{aligned} \tag{15}$$

As mentioned before, the free flows are stochastic due to uncertain consumption and this uncertainty has been modeled as a Brownian motion. As a consequence of this, the pressure measurement is represented via Ito integral (reviewed in Kushner and Dupuis (2001)). We begin by defining $y = \bar{p} - p_0$ and $T_r = \begin{bmatrix} B^T & B_d^T F \end{bmatrix}$ then,

$$y = \int_s^t \underbrace{(\bar{H}_{\mathcal{T}}^{-T} J_{\mathcal{T}} \Pi_{\mathcal{T}} T)}_{\mathrm{h}} dx + \bar{H}_{\mathcal{T}}^{-T} R_{\mathcal{T}} \Pi_{\mathcal{T}} T_r x(s) - \rho g (\bar{z} - z_0) - \rho g \bar{H}_{\mathcal{T}}^{-T} G u(s). \quad (16)$$

We assume that h is a right-continuous, adapted and a locally bounded process. We approximate the Ito integral in (16) using a mesh m with grid size $\mathcal{Y}_m \to 0$ as,

$$\int_s^t \mathrm{h} dx = \lim_{m\to\infty} \sum_{[t_{i-1},t]\in\mathcal{Y}_m} \mathrm{h}_{t_{i-1}} (x_{t_i} - x_{t_{i-1}}). \quad (17)$$

We shall now define the control objectives of both the players as the following stage cost functions for a player k,

$$c^k(y, x, u) = y^T W_1 y + W_2 \left| (G^k)^T A_{ex}^T x u^k \right|, \quad (18)$$

where $|\cdot|$ represents the standard 1-norm, W_1 and W_2 are normalized weights. The first term in (18) represents the pressure variance at consumer nodes and the control objective for the pumps is to minimize the same. The second term in (18) represents the energy consumption (kW) for a pumping station k.

3. MARKOV CHAIN APPROXIMATION

We consider Markov chain approximation (MCA) which is a numerical method for solving stochastic control problem. MCA was introduced in Kushner and Dupuis (2001) and it works as follows. Firstly, we consider the reachable states under a given control policy and thereafter, we calculate the transition probabilities and the interpolation time by applying finite differences method on the Hamilton-Jacobi-Bellman equation (see fig. 3 for an illustration of MCA). We define a state grid for the WDN based on normalizing

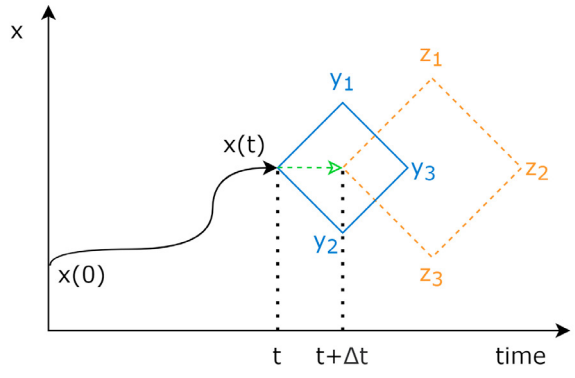

Fig. 3. Illustration of MCA. Here we see a sample path of (1) starting at $x(0)$ and reaching $x(t)$ at time t, y_1, y_2 and y_3 represent the reachable states for $x(t)$ in time interval Δt. Any realized state $x(t+\Delta t)$ (shown by the green arrow) can be found as a convex combination of extremities of blue polygon. Once the next state $x(t+\Delta t)$ is reached the process repeats as shown by orange polygon with reachable states z_1, z_2 and z_3.

the amount of water flowing in the network. Consider the k^{th} edge of the WDN denoted by e_k. Suppose that the total amount of water which can flow through e_k is Q. Than by considering how much percentage of Q is flowing through e_k, we can assign a number between 1 to 100 indicating the flow through e_k. This allows us to construct a grid where all free flows are normalized. Consider $\mathcal{B} \subset \mathbb{R}^n$ as a compact set on state space with boundary absorption. The discounted stochastic control problem with $\tau = \min\{t : x(t) \notin \mathcal{B}\}$ representing the minimum time taken to exit $\mathcal{B}$ can now be defined as follows.

$$\min_u \mathop{\mathbb{E}}_{x,u(\cdot)} \left(\int_s^\tau \gamma^t c\left(x(t), u(t)\right) dt + \gamma^\tau c_\tau\left(x_\tau\right) \right) \mid x(s) \quad (19a)$$

$$\text{s.t. } dx(t) = (\mathrm{f}(x(t)) + \mathrm{g}(x(t))u(t))dt + \sigma d\mathrm{w}(t), \quad (19b)$$

$$u(t) \in U, \quad x(t) \in \mathcal{B}, \quad (19c)$$

where $x(s)$ is the given state at some starting time s, γ is the discount factor and U is the admissible controls. Let the stochastic integral in the objective function of (19) be denoted by

$$Q(x, u(\cdot)) = \mathop{\mathbb{E}}_{x,u(\cdot)} \left(\int_s^\tau \gamma^t c\left(x(t), u(t)\right) dt + \gamma^\tau c_\tau\left(x_\tau\right) \right). \quad (20)$$

The optimal value function can than be obtained by minimizing $Q(x, u(\cdot))$ with respect to $u(\cdot)$ for each state as follows

$$V(x) = \inf_{u(\cdot)} Q(x, u(\cdot)). \quad (21)$$

Let $\xi(s) \in \mathbb{R}^n$ denote the state $x(s)$ at some time s, than the next reachable discrete state in the set $\mathcal{B}$ is denoted by $\xi(s+1) = x(s) + e_i h$, where h is an approximation parameter which indicates the coarseness of the state-grid and $e_i \in \mathbb{R}^n$ is the basis vector in i^{th} direction of the state space. The grid is shown in fig. 4. Note that, we are

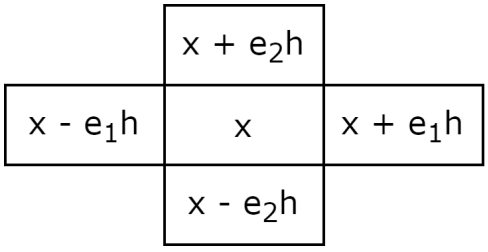

Fig. 4. Illustration of state space as a grid. Any state whose numerical value is within the boundary of a particular grid box is considered as a part of the same.

considering only state transitions of the type $x \pm e_i h$ and not of the type $x \pm e_i h \pm e_j h$ since the interpolation time is small (depending on the drift and diffusion as we shall see later) and the diffusion matrix is diagonal. We shall now apply the well-known Ito's lemma (which can be thought of as chain rule for SDE's) in order to obtain evolution of $Q(x, u(\cdot))$ with time. We begin by defining the differential operator $\mathcal{L}$ as follows. Let $a_{ij}(x)$, $i, j = 1, \cdots, n$ be an element of the diffusion matrix $a(x)$, than we can define $\mathcal{L}$ for (1) as follows.

$$\mathcal{L} = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} (\mathrm{f}_i(x) + \mathrm{g}_i(x)u) \frac{\partial}{\partial x_i} \quad (22)$$

Applying Ito's differential operator on (19) gives us the stochastic analogue of the well-known Hamilton-Jacobi-Bellman equation as follows.

$$\begin{aligned}&\mathcal{L}V(x(t)) + \gamma^t c(x(t), u(t)) = 0,\\ &V(x(\tau)) = c(x(\tau)), V(x(s)) = c(x(s)),\end{aligned} \tag{23}$$

where $V(x(\tau)) = c(x(\tau))$ and $V(x(s)) = c(x(s))$ represents the boundary condition. Expanding the differential operator in (23) gives us,

$$\frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2 V(x)}{\partial x_i \partial x_j} + \sum_{i=1}^{n}(\mathrm{f}_i(x) + \mathrm{g}_i(x)u)\frac{\partial V(x)}{\partial x_i} + \gamma^t c(x,u) = 0, \tag{24}$$

where we have written $x(t), u(t)$ as x, u for notational simplicity. In the sequel, we shall refer to the drift term $\mathrm{f}_i(x) + \mathrm{g}_i(x)u$ as $b_i(x,u)$ for the sake of notational simplicity. We now apply finite-differences with space approximation parameter h on (24) as follows.

$$\begin{aligned}&\frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x)\frac{V(x+e_i h) + V(x-e_i h) - 2V(x)}{h^2}\\ &\quad + \sum_{i=1}^{n}|b_i(x,u)|^{+}\frac{V(x+e_i h) - V(x)}{h}\\ &- \sum_{i=1}^{n}|b_i(x,u)|^{-}\frac{V(x) - V(x-e_i h)}{h} + \gamma^t c(x,u) = 0,\end{aligned} \tag{25}$$

where $|b_i(x,u)|^{+} = \max(b_i(x,u), 0)$ and $|b_i(x,u)|^{-} = \max(-b_i(x,u), 0)$. This is the standard "upwind" scheme in numerical analysis of hyperbolic partial differential equations (see LeVeque (2007)). The intuition behind this scheme is that the approximation of V should be in the same direction as the drift of (1). It can also be verified that $|b_i(x,u)|^{+} + |b_i(x,u)|^{-} = |b_i(x,u)|$. Since the diffusion matrix is diagonal as per assumption 2.3, we can rearrange the terms in (25) so as to obtain transition probabilities $p(x, x \pm e_i h \mid u)$ and interpolation time Δt as follows.

$$\begin{aligned}V(x) = &\underbrace{\frac{a_{ii}(x)/2 + h\,|b_i(x,u)|^{+}}{\sum_{i=1}^{n}(a_{ii}(x) + h\,|b_i(x,u)|)}}_{p(x, x+e_i h|u)} V(x+e_i h)\\ &+ \underbrace{\frac{a_{ii}(x)/2 + h\,|b_i(x,u)|^{-}}{\sum_{i=1}^{n}(a_{ii}(x) + h\,|b_i(x,u)|)}}_{p(x, x-e_i h|u)} V(x-e_i h)\\ &+ \underbrace{\frac{h^2}{\sum_{i=1}^{n}(a_{ii}(x) + h\,|b_i(x,u)|)}}_{\Delta t}\gamma^t c(x,u).\end{aligned} \tag{26}$$

The probability of state remaining unchanged is given by

$$p(x, x \mid u) = 1 - (p(x, x+e_i h \mid u) + p(x, x-e_i h \mid u)). \tag{27}$$

Thus, we obtain an MDP which is defined by the tuple $(\mathcal{X}, \mathcal{U}, P, \gamma, \mathcal{C})$ where $\mathcal{X} = \{x, x \pm e_i h \mid i = 1, \cdots, n\}$, $\mathcal{U}$ are discreatized control actions, P is the transition probabilities for the next state, $\gamma \in (0,1)$ is the discount factor and $\mathcal{C} = \gamma^t c(x,u)\Delta t$ is the stage cost.

4. VERIFICATION USING MODIFIED MONTE CARLO METHOD

The convergence proof for MCA is given in Kushner and Dupuis (2001) and is based on probabilistic arguments whose intuition is explained in fig. 3. We would like to verify in practice how good the approximation (26) is and we do that using a modified Monte Carlo method stated in Algorithm 1. Algorithm 1 is based on the sample paths of (1) and it calculates transition probabilities based on number of times a grid square $\Delta^h x(s)$ (with the entire state space being represented by the grid $\Delta^h x$) corresponding to state $x(s)$ is visited. This is in contrast to MCA which calculates transition probabilities based solely on drift and diffusion terms of (1) as done in (26). The

Algorithm 1 Monte Carlo method for constructing P

Input: State $\xi(s) = x(s)$, Sate space grid $\Delta^h x$, Maximum time T for a sample path, No. of sample paths to be evaluated M
for all sample paths $m \leq M$ **do**
  while $s < T$ **do**
    Find grid square $\Delta^h x(s)$ in $\Delta^h x$ which corresponds to $x(s)$
    Increment counter $MC(\Delta^h x(s))$ corresponding to $\Delta^h x(s)$ by 1
    Obtain realization of $x(s+1)$ using modified Euler-Maruyama method with interpolation time Δt
    $s = s + \Delta t$
  end while
  for all $i \leq n$, find $\xi_i(s+1)$ based on drift **do**
    if $b_i(x,u) \geq 0$ **then**
      $\xi_i(s+1) = \xi(s) + 1$
    else
      $\xi_i(s+1) = \xi(s) - 1$
    end if
  end for
  $p(\xi_i(s), \xi_i(s+1)) = \frac{MC_i(\Delta^h x(s+1))}{\sum_{i=1}^{n} MC_i(\Delta^h x)}\ \forall i < n$
  $p(\xi(s), \xi(s)) = 1 - \left(\sum_{i=1}^{n}\frac{MC_i(\Delta^h x(s+1))}{\sum_{i=1}^{n} MC_i(\Delta^h x)}\right)$
end for

sample paths of (1) are realized using Euler-Maruyama method (see Kloeden and Platen (1992)) modified with interpolation time Δt obtained via (26). The interpolation time Δt is used instead of standard equally spaced time intervals so as to retain similar time intervals as MCA for comparison. Furthermore Algorithm 1, finds next state component $\xi_i(s+1)$ on the grid $\Delta^h x$ based on the sign of the corresponding i^{th} component of the drift vector $b_i(x,u)$. This is done for ensuring that both Algorithm 1 and MCA are considering the same future state. Since Algorithm 1 is a Monte Carlo based method, it requires evolution of a complete sample path (from $x(0)$ to $x(T)$, where $[0,T]$ is the time interval for numerical evaluation of (1)) prior to constructing P. This implies that we cannot compare P obtained from both the methods directly since they are evaluated at different times. This motivates us to consider the approximate value function constructed by both the methods as a means of comparison. The approximate value function from modified Monte Carlo method is

obtained by substituting the probabilities obtained from Algorithm 1 into (26) as coefficients of $V(x + e_i h)$ and $V(x - e_i h)$.

5. SIMULATION RESULTS

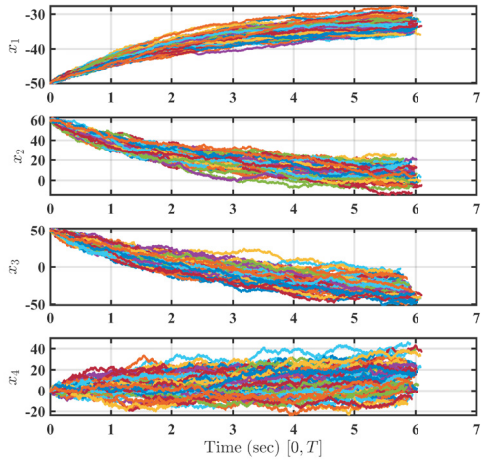

Fig. 5. Sample paths for initial $x = [-50\ 60\ 50\ 0]^T$

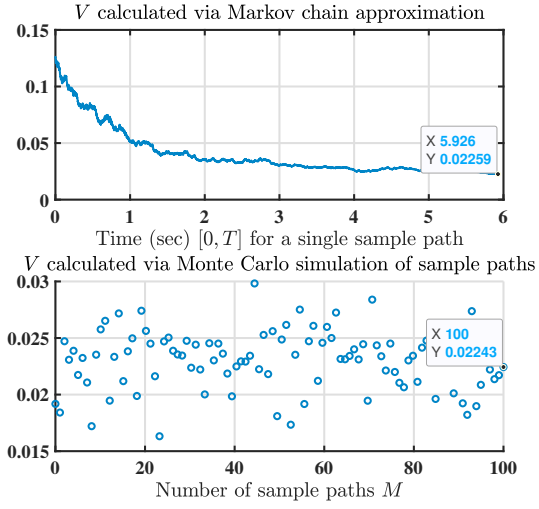

Fig. 6. The obtained value functions converge to almost same value.

In fig. 5, we see 100 sample paths of (1) obtained using modified Euler-Maruyama method. These sample paths form the basis for calculating value function using Algorithm 1. In fig. 6, we see convergence of value function obtained from MCA in upper subplot and convergence of value function obtained from Algorithm 1 in the lower subplot. Furthermore, we have also simulated both the methods for different initial conditions and the results are summarized in the table 1.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented a model of WDN and converted the same to an MDP using MCA. We have also proposed Algorithm 1 as an alternative for MCA.

Initial condition	V obtained from MCA	V obtained from Algorithm 1
$x(0) = [-60\ 60\ 50\ 50]^T$	0.0170	0.0168
$x(0) = [-50\ 50\ 30\ 60]^T$	0.0157	0.0158
$x(0) = [-50\ 60\ 50\ 10]^T$	0.0194	0.0198

Table 1. Value functions obtained for different initial conditions

Simulation results show in fig. 6 show that the approximated value functions converge and are reasonably close. In future, we would like to apply MCA for WDN's with dynamic games as discussed in Kushner (2002).

ACKNOWLEDGEMENTS

Financial support from the Poul Due Jensen Foundation (Grundfos Foundation) for this research is gratefully acknowledged.

REFERENCES

Bertsekas, D.P. (2005). *Dynamic programming and optimal control*, volume 1, 2. Athena Scientific.

Deo, N. (2017). *Graph theory with applications to engineering and computer science*. Courier Dover Publications.

Filar, J. and Vrieze, K. (2012). *Competitive Markov decision processes*. Springer Science & Business Media.

Jensen, T.N. and Wisniewski, R. (2011). Global practical stabilisation of large-scale hydraulic networks. *IET control theory & applications*, 5(11), 1335–1342.

Kloeden, P.E. and Platen, E. (1992). Stochastic differential equations. In *Numerical Solution of Stochastic Differential Equations*, 103–160. Springer.

Kushner, H.J. (2002). Numerical approximations for stochastic differential games. *SIAM journal on control and optimization*, 41(2), 457–486.

Kushner, H.J. and Dupuis, P.G. (2001). *Numerical methods for stochastic control problems in continuous time*, volume 24. Springer Science & Business Media.

LeVeque, R.J. (2007). *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems*. SIAM.

Munos, R. and Bourgine, P. (1997). Reinforcement learning for continuous stochastic control problems. In *NIPS*, 1029–1035.

Polycarpou, M.M., Uber, J.G., Wang, Z., Shang, F., and Brdys, M. (2002). Feedback control of water quality. *IEEE Control Systems Magazine*, 22(3), 68–87.

Pétursson Geir Bjarni, M.M. (2015). *Energy Optimisation of Water Distribution Networks: Using a distributed pumping solution for optimal pressure control*. Department of Control Engineering, Aalborg University.

Sutton, R.S. and Barto, A.G. (2018). *Reinforcement learning: An introduction*. MIT press.

Swamee, P.K. and Sharma, A.K. (2008). *Design of water supply pipe networks*. John Wiley & Sons.

Tahavori, M., Kallesøe, C.S., Leth, J., and Wisniewski, R. (2012). Modeling of water supply systems: Circuit theoretic approach. In *2012 7th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 1561–1566. IEEE.