Interference-Constrained Scheduling of a Cognitive Multihop Underwater Acoustic Network

Chen Peng, Urbashi Mitra
Ming Hsieh Department of Electrical and Computer Engineering
University of Southern California
Los Angeles, USA
{cpeng732, ubli}@usc.edu

***Abstract*—In this paper, optimal scheduling for a cognitive multihop underwater acoustic network with a primary user interference constraint is investigated. The network consists of primary and secondary users, with multihop transmission adopted for both of them. Key characteristics of underwater acoustic channels including large propagation delay, distance-and-frequency dependent attenuation, half-duplex modem constraint and inter-hop interference are taken into account in the design and analysis. The network scheduling goal is to maximize the end-to-end throughput of the whole system, while limiting the throughput loss of primary users. The framework of Partially Observable Markov Decision Process is applied to formulate the optimization problem and an optimal dynamic programming algorithm is derived. However, for this problem, the optimal dynamic programming solution is computationally intractable. Key properties are shown for the objective function, enabling the design of an approximate scheme which offers near optimal performance with significant complexity reduction. Numerical results show the proposed scheme significantly increases system throughput and while maintaining the primary throughput loss constraint.**

***Index Terms*—Cognitive underwater acoustic network, multihop transmission, partially observable Markov decision processes (POMDP), dynamic programming, approximation scheme**

I. Introduction

Underwater communications and networks are currently an active research area due to a wide range of applications including environment monitoring, oceanography data collection, underwater target tracking, *etc.* [1]. As electromagnetic and optical signals have limited transmission range, acoustic signals are employed for underwater communication [2]. With respect to terrestrial communications, underwater acoustic communications (UACs) possess distinguishing characteristics, such as long propagation delay, high path attenuation and limited spectrum bandwidth, *etc.* [3]. Therefore, to provide wider area coverage with high efficiency, many UAC systems utilize multihop transmission [4]–[6], where a long distance is divided into multiple hops.

Given the limited available bandwidth for UAC, efficient spectral utilization is desired. To this end, a cognitive radio approach has been proposed to be used in underwater acoustic networks (UANs) [1], [7], [8], which includes dynamic and opportunistic spectrum access.

Key aspects of our proposed framework have been previously considered. Multihop cooperative transmission and network scheduling for UAC systems has attracted intense research interest [5], [6], [9]–[14], where optimal parameter selection and scheduling are investigated for a single class of users. Spectrum allocation and utilization for cognitive UANs has been investigated for the single hop case in [7], [8] and a software-defined acoustic modem prototype is proposed in [15].

Herein, we propose to consider both cognitive scheduling and multihop communication systems. This joint problem formulation offers challenges with an increased decision dimension and the need for interference management. However, the multihop feature offers additional temporal/spatial re-use opportunities. In our system, we have a multihop network of primary users (PUs) that coexists with a multihop network of secondary users (SUs). The main contributions of this work are summarized as

- We model the complicated time-evolving behavior of cognitive multihop network in realistic underwater acoustic channel as a discrete time Markov chain, which captures the distinguishing underwater channel characteristics of long propagation delay and high path attenuation.
- We formulate the optimal scheduling problem with a primary user interference constraint as a constrained Partially Observable Markov Decision Process (POMDP) and derive an optimal dynamic programming (DP) algorithm for solving the problem.
- We prove properties of the reward-to-go functions, which enables the derivation of an efficient approximation algorithm.
- We evaluate performance of the proposed approximation scheme through simulation results and illustrate that it is adaptive to unique characteristics of underwater acoustic channel.

This paper is organized as follows. In Section II, the system model is described. The optimal scheduling problem with performance constraint is formulated in Section III. We propose an approximation scheme in Section IV and evaluate its performance via numerical results in Section V. Finally, we conclude the paper in Section VI.

II. System Model

In this section, we describe the underwater channel model, state space representation of our system and transition proba-

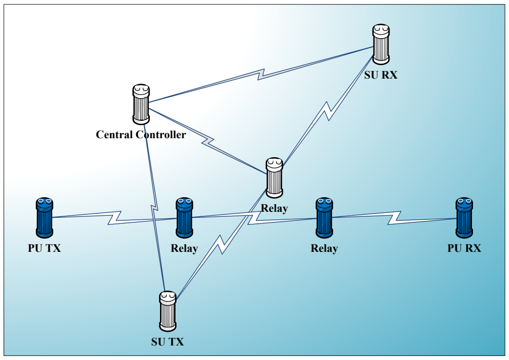

Fig. 1. A cognitive multihop underwater acoustic network of $N_P = 4$ PUs, $N_S = 3$ SUs and a central controller.

bilities for our cognitive multihop network.

A. Underwater Acoustic Channel Model

The attenuation experienced by an underwater acoustic signal transmitted over a distance d meters at carrier frequency f (in kHz), is given by (in dB):

$$10\log_{10}(A(d,f)/A_0) = \kappa 10\log_{10} d + \frac{d}{10^3} 10\log_{10} a(f), \quad (1)$$

where A_0 is a normalizing constant, κ denotes the spreading factor and $a(f)$ is the absorption coefficient [4]. The absorption coefficient can be expressed in dB/km using Thorp's empirical formulas for f in kiloHertz as [16]

$$10\log_{10} a(f) = 0.11\frac{f^2}{1+f^2} + 44\frac{f^2}{4100+f^2} + 2.75\times 10^{-4} f^2 + 0.003. \quad (2)$$

For simplicity and without loss of generality, fading and multipath effect are not taken into account. The following framework and analysis can be extended to the channel model with fading and multipath effect.

The noise in an acoustic channel mainly comes from four sources: turbulence, waves, shipping and thermal noise [4] and is modeled as Gaussian. The overall power spectral density of the non-white, ambient noise in dB is 1μPa^2/Hz (*i.e.*, the power per unit bandwidth associated with the reference sound pressure level of 1μPa) can be approximated as [17]

$$10\log_{10} N(f) = \eta_0 - 18\log_{10} f. \quad (3)$$

B. State Space Representation

We consider a CM-UAN consisting of multihop chains of N_P PUs and N_S SUs respectively, as shown in Fig. 1. For multihop networks of both PUs and SUs, transmitter at one end of the multihop chain wants to send packets to receiver at the other end, with the help of relay nodes in between. There is also a central controller that takes observation data from each SU and determines behavior of SU network. Our goal is to maximize the end-to-end throughput of the whole system while restricting the throughput loss of PUs to certain level.

In our system, PUs adopt a fixed, decode-and-forward protocol without re-transmission. To take into account the half-duplex constraint of the acoustic modems and avoid the hidden terminal problem, each PU will stay idle for two time slots after transmission of a packet. We define the system state as the current action performed by each PU and the current buffer state of each SU. The channel quality information is not included in the system state, but reflected in the state transition probabilities introduced in the next part. The underlying state vector in time slot k for PUs is represented as

$$\boldsymbol{s}_k \triangleq (s_k^1, s_k^2, \ldots, s_k^{N_P-1})^T \in \boldsymbol{S} \equiv \{0,1\}^{N_P-1}, \quad (4)$$

where s_k^i is the state of the ith PU, which can be idle ($s_k^i = 0$) or transmitting ($s_k^i = 1$). The buffer state vector of SUs in time slot k is represented as

$$\boldsymbol{b}_k \triangleq (b_k^1, b_k^2, \ldots, b_k^{N_S-1})^T \in \boldsymbol{B} \equiv \{0,1\}^{N_S-1}, \quad (5)$$

where b_k^i is the state of the ith SU, which can be empty ($b_k^i = 0$) or not empty ($b_k^i = 1$).

Over L time slots, the SUs opportunistically access the spectrum to transmit packets, based on the decision of the central controller in each time slot, represented as

$$\boldsymbol{\delta}_k \triangleq (\delta_k^1, \delta_k^2, \ldots, \delta_k^{N_S-1})^T \in \boldsymbol{\Delta} \equiv \{0,1\}^{N_S-1}, \quad (6)$$

where δ_k^i is the decision variable for the ith SU, which can be sensing ($\delta_k^i = 0$) or transmitting ($\delta_k^i = 1$). $\boldsymbol{\delta}_k$ is dependent on $\boldsymbol{b}_k$ because a SU can only transmit when there is at least one packet in its buffer. Different from terrestrial cognitive radio networks, in our model, SUs can not conduct both sensing and transmission in one slot, because of the long propagation delay characteristic of underwater acoustic channel.

C. Transition Probability

The system state transition depends on two factors, new packet arrival and transmission result. The probability that a new packet arrives at the PU transmitter is denoted as $\alpha \in [0,1]$. The SU transmitter is assumed to be backlogged, *i.e.*, it always has packets to transmit. For each relay node of PUs, if a new packet is received successfully, it transmits the packet to the next node with probability 1. Transmission of the ith PU hop in time slot k is successful with probability $p_P^i(\boldsymbol{s}_k, \boldsymbol{\delta}_k) \in [0,1]$, which is approximated by

$$p_P^i(\boldsymbol{s}_k, \boldsymbol{\delta}_k) = 1 - (1 - Q(\sqrt{2\gamma_b(\boldsymbol{s}_k, \boldsymbol{\delta}_k)}))^N, \quad (7)$$

where N is the packet size in bits and γ_b is the per bit signal to interference and noise ratio (SINR), which is expressed as

$$\gamma_b(\boldsymbol{s}_k, \boldsymbol{\delta}_k) = \frac{|H(d,f)|^2}{I(\boldsymbol{s}_k, \boldsymbol{\delta}_k, f) + N(f)}, \quad (8)$$

where $I(\boldsymbol{s}_k, \boldsymbol{\delta}_k, f)$ is the inference power caused by other PUs and SUs and the function $Q(\cdot)$ is the tail distribution of the standard normal distribution function. The probability of successful transmission for SUs, denoted as p_S^i, is defined similarly.

Therefore we can model the PU state in the kth time slot $\boldsymbol{s}_k$ as a Markov chain, with transition matrix ($\boldsymbol{P}(\boldsymbol{\delta}_k)$) depending on SU decision vector $\boldsymbol{\delta}_k$.

D. Observation Model and Sufficient Statistics

Scheduling of the SUs is based on channel occupancy measurements collected by the SUs at sensing mode. Denote the set of all sensing mode SUs in time slot k as M_k. We assume the channel sensing is conducted in centralized fashion such that each SU collects noisy measurements independently and report to the central controller. The observation model for the ith SU in time slot k is expressed as

$$y_{k,m} = \boldsymbol{c}_{k,m}^T \boldsymbol{s}_k + \boldsymbol{d}_{k,m}^T \boldsymbol{\delta}_k + n_{k,m}, \quad \forall m \in M_k, \tag{9}$$

where $n_{k,m} \sim \mathcal{N}(0, \sigma_N^2)$ is i.i.d. Gaussian noise over time and across SUs. $\boldsymbol{c}_{k,m}^T$ and $\boldsymbol{d}_{k,m}^T$ are measurement vectors for PUs and SUs respectively, containing channel attenuation information. Since observations are noisy the central controller can only make decision based on partially observed state information. So the optimal scheduling of a CM-UAN is a POMDP problem.

The information vector in time slot k, denoted as $\boldsymbol{I}_k$, can be expressed as

$$\boldsymbol{I}_k = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{k-1}, \boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_{k-1}), \tag{10}$$

which is of expanding dimension. However, for control purposes, we can use the probability distribution of state $\boldsymbol{s}_k$, denoted by $\boldsymbol{\omega}_k$, as a sufficient statistic [18]. $\boldsymbol{\omega}_k$ is also known as belief state and is defined as

$$\boldsymbol{\omega}_k = [\omega_k^1, \ldots, \omega_k^n]^T, \text{ where } \omega_k^j = \mathbb{P}(\boldsymbol{s}_k = j|\boldsymbol{I}_k), j \in \boldsymbol{S}_k. \tag{11}$$

Given observation $\boldsymbol{y}_k$, the central controller can update its belief state via Bayes' Rule as follows

$$\omega_{k,i} \triangleq \mathbb{P}(\boldsymbol{s}_k = i|\boldsymbol{y}_k) \tag{12}$$

$$= \frac{\left[\prod_{m \in M_k} \mathbb{P}(y_{k,m}|\boldsymbol{s}_k = i)\right] \left[\boldsymbol{\omega}_{k-1}^T \boldsymbol{P}(\boldsymbol{\delta}_{k-1})_i\right]}{\sum_{j \in \boldsymbol{S}_k} \left[\prod_{m \in M_k} \mathbb{P}(y_{k,m}|\boldsymbol{s}_k = j)\right] \left[\boldsymbol{\omega}_{k-1}^T \boldsymbol{P}(\boldsymbol{\delta}_{k-1})_j\right]}, \tag{13}$$

where $\boldsymbol{P}(\boldsymbol{\delta}_{k-1})_i$ represents the ith column of the transition matrix $\boldsymbol{P}(\boldsymbol{\delta}_{k-1})$ and $\mathbb{P}(y_{k,m}|\boldsymbol{s}_k = i)$ is the probability distribution function of a Gaussian distribution evaluated at $y_{k,m}$. Further, we denote the belief state update rule as $\Phi(\cdot)$, then $\boldsymbol{\omega}_k = \Phi(\boldsymbol{\omega}_{k-1}, \boldsymbol{\delta}_{k-1}, \boldsymbol{y}_k)$.

Notice that because of the large propagation delay, the central controller can only obtain $\boldsymbol{\omega}_k$ at the end of time slot k. Therefore the central controller has to make decision for time slot k only based on the belief $\boldsymbol{\omega}_{k-1}$ calculated from previous observation. The uncertainty existing in the time-evolving system comes from the following sources: random packet arrival, random transmission failure, imperfect and delayed observation.

III. The Optimal Scheduling Problem

In this section, we formulate the optimal interference-constrained scheduling problem. Our goal is to maximize the expected end-to-end throughput of the whole system over a finite horizon of L time slots. If we define, M_P and M_S as the packet sizes of PU network and SU network, respectively and the functions, $g_P(\cdot)$ and $g_S(\cdot)$, to be the throughput of PUs and SUs; then we can define the system throughput as follows

$$g(\boldsymbol{s}_k, \boldsymbol{\delta}_k) = g_P(\boldsymbol{s}_k, \boldsymbol{\delta}_k) + g_S(\boldsymbol{s}_k, \boldsymbol{\delta}_k) \tag{14}$$

$$\stackrel{(a)}{=} M_P \cdot s_k^{N_P-1} \cdot p_P^{N_P-1}(\boldsymbol{s}_k, \boldsymbol{\delta}_k) + M_S \cdot \delta_k^{N_S-1} \cdot p_S^{N_S-1}(\boldsymbol{s}_k, \boldsymbol{\delta}_k). \tag{15}$$

Equality (a) holds because the throughput is achieved only when the transmission to the end receiver is successful. Denote the expected accumulated throughput for the system over a finite horizon as J, then optimization problem can be formulated as follows

$$\text{Maximize:} \quad J(\boldsymbol{\delta}) = \mathbb{E}\left\{\sum_{k=1}^{L} g(\boldsymbol{s}_k, \boldsymbol{\delta}_k)\right\} \tag{16}$$

$$\text{Subject to:} \quad \mathbb{E}\left\{\sum_{k=1}^{L} g_P(\boldsymbol{s}_k, \boldsymbol{\delta}_k^*)\right\} \geq \gamma \mathbb{E}\left\{\sum_{k=1}^{L} g_P(\boldsymbol{s}_k, \boldsymbol{0})\right\}, \tag{17}$$

where $\gamma \in (0, 1)$ is the PU throughput degradation coefficient. The constraint states that the expected throughput of PUs given by the optimal decision sequence ($\boldsymbol{\delta}^*$) should be larger than or equal to the γ weighted throughput of PUs in the case of no SU interference. Therefore the interference of SUs can be restricted by selecting specific value of γ.

The optimization problem is a constrained POMDP problem and we make use of the belief state $\boldsymbol{\omega}_k$ to rewrite the objective function as

$$J(\boldsymbol{\delta}) = \mathbb{E}\left\{\sum_{k=1}^{L} \boldsymbol{\omega}_k^T \cdot \boldsymbol{g}(\boldsymbol{\delta}_k)\right\}, \tag{18}$$

where $\boldsymbol{g}(\boldsymbol{\delta}_k)$ is the vector of system throughput at each PU state, defined as

$$\boldsymbol{g}(\boldsymbol{\delta}_k) = M_P \cdot s_k^{N_P-1} \boldsymbol{p}_P^{N_P-1}(\boldsymbol{\delta}_k) + M_S \cdot \delta_k^{N_S-1} \boldsymbol{p}_S^{N_S-1}(\boldsymbol{\delta}_k), \tag{19}$$

where $\boldsymbol{p}_P^{N_P-1}(\boldsymbol{\delta}_k)$ and $\boldsymbol{p}_S^{N_S-1}(\boldsymbol{\delta}_k)$ are vectors of the successful transmit probabilities consisting of $p_P^{N_P-1}(\boldsymbol{s}_k, \boldsymbol{\delta}_k)$ and $p_S^{N_S-1}(\boldsymbol{s}_k, \boldsymbol{\delta}_k)$ respectively, $\forall \boldsymbol{s}_k \in \boldsymbol{S}$.

The constrained optimization problem can be solved by dynamic programming. To take the constraint into account, we formulate the reward-to-go function $\hat{J}_k$ as a function of both the belief state $\boldsymbol{\omega}_k$ and the achieved PU throughput η_k. For $k = L-1, \ldots, 0$, the reward-to-go function $\hat{J}_k(\boldsymbol{\omega}_k, \eta_k)$ is related to $\hat{J}_{k+1}(\boldsymbol{\omega}_{k+1}, \eta_{k+1})$ through the recursion

$$\hat{J}_k(\boldsymbol{\omega}_k, \eta_k) = \max_{\boldsymbol{\delta}_k \in \boldsymbol{\Delta}} \Big\{\boldsymbol{\omega}_k^T \boldsymbol{g}(\boldsymbol{\delta}_k) + \mathop{\mathbb{E}}_{\boldsymbol{y}_{k+1}} \left[\hat{J}_{k+1}(\Phi(\boldsymbol{\omega}_k, \boldsymbol{\delta}_k, \boldsymbol{y}_{k+1}), \eta_k - \boldsymbol{\omega}_k^T \boldsymbol{g}_{\boldsymbol{P}}(\boldsymbol{\delta}_k))\right]\Big\}, \tag{20}$$

where $\Phi(\cdot)$ is the belief state update rule. The reward-to-go function for $k = L$ is given by

$$\hat{J}_L(\boldsymbol{\omega}_L, \eta_L(\boldsymbol{\delta}_L)) = \max_{\boldsymbol{\delta}_L \in \boldsymbol{\Delta}} \left\{\boldsymbol{\omega}_L^T \boldsymbol{g}(\boldsymbol{\delta}_L)\right\}. \tag{21}$$

However, solving the dynamic programming problem is intractable because of the uncountably infinite belief space, an exponentially large control space and the uncountably infinite observation space. Specifically, for m possible system states, d decision choices, and a discretization of the belief probability with n_1 levels, the number of belief states is $\mathcal{O}(n_1^m)$, where $m = \mathcal{O}(2^{N_P+N_S})$. Further, if we discretize the observation with n_2 levels, and the constraint value with n_3 levels, the computational complexity for determining the optimal decision sequence is $\mathcal{O}(d \cdot n_1^m n_2 n_3 \cdot L)$.

IV. Cognitive Time Slot Scheduling

In the previous section, we argued that optimal scheduling using dynamic programming is computationally intractable. In this section, we present a low-cost approximation scheme that can be implemented and run in real-time. We denote this method as the *Cognitive Time Slot Scheduling* scheme. We begin by reviewing the factors that contribute to computational complexity: an average performance constraint and delayed and imperfect observations. By exploring the unique characteristics of the system model and properties of the objective function, we will develop the approximation scheme.

A. Apply Local instead of Global Constraints

We take the first step to revise the constraint over the whole horizon, which greatly complicates the dynamic programming solution by expanding dimension of the solution space. Therefore, we approximate the original problem by posing stricter local constraints such that in each time slot k, the degradation level calculated from slot k to the last slot L must always be lower than the degradation coefficient, which is expressed as

$$\mathbb{E}\left\{\sum_{k=i}^{L} g_P(\boldsymbol{s}_k, \boldsymbol{\delta}_k^*(\boldsymbol{b}_k))\right\} \geq \gamma \cdot \mathbb{E}\left\{\sum_{k=i}^{L} g_P(\boldsymbol{s}_k, \mathbf{0})\right\}, \quad \forall i \in \{1, 2, \ldots, L\}. \tag{22}$$

The original constraint (17) only requires the inequality holds for $i = 1$ while the stricter constraint (22) requires it holds for $\forall i \in \{1, 2, \ldots, L\}$.

Proposition 1. *The feasible set of (22) is not empty and whenever a sequence of decisions satisfies (22), it also satisfies (17).*

Proof. We prove this proposition by induction. First, for $i = L$, there exists decisions that satisfy (22) because $\boldsymbol{\delta}_L^* = \mathbf{0}$ is always a valid choice. Next, let us assume that for $i = m+1$, a valid sequence of decisions satisfying $\mathbb{E}\left\{\sum_{k=m+1}^{L} g_P(\boldsymbol{s}_k, \boldsymbol{\delta}_k^*)\right\} \geq \gamma \cdot \mathbb{E}\left\{\sum_{k=m+1}^{L} g_P(\boldsymbol{s}_k, \mathbf{0})\right\}$ is found. Then for $i = m$, we want to select $\boldsymbol{\delta}_m^*$ such that the following expression is positive,

$$\mathbb{E}\left\{\sum_{k=m}^{L} g_P(\boldsymbol{s}_k, \boldsymbol{\delta}_k^*)\right\} - \gamma \cdot \mathbb{E}\left\{\sum_{k=m}^{L} g_P(\boldsymbol{s}_k, \mathbf{0})\right\} \tag{23}$$

$$= \mathbb{E}\left\{g_P(\boldsymbol{s}_m, \boldsymbol{\delta}_m^*) + \sum_{k=m+1}^{L} g_P(\boldsymbol{s}_k, \boldsymbol{\delta}_k^*)\right\} - \gamma \cdot \mathbb{E}\left\{g_P(\boldsymbol{s}_m, \mathbf{0}) + \sum_{k=m+1}^{L} g_P(\boldsymbol{s}_k, \mathbf{0})\right\} \tag{24}$$

$$\geq \mathbb{E}\left\{g_P(\boldsymbol{s}_m, \boldsymbol{\delta}_m^*)\right\} - \gamma \cdot \mathbb{E}\left\{g_P(\boldsymbol{s}_m, \mathbf{0})\right\}. \tag{25}$$

Such $\boldsymbol{\delta}_m^*(\boldsymbol{b}_m)$ always exists because at least $\boldsymbol{\delta}_L^* = \mathbf{0}$ satisfies (25). Thus by induction, a sequence of decisions satisfying (22) can be found for $\forall i \in \{1, 2, \ldots, L\}$. When $i = 1$, (22) is the same as the original constraint, which is satisfied by any sequence of decisions that satisfies (22). ☐

This approximation provides a lower bound of $\hat{J}_k(\boldsymbol{\omega}_k, \eta_k)$, denoted as $\underline{\hat{J}}_k(\boldsymbol{\omega}_k)$, which has a simpler form

$$\underline{\hat{J}}_k(\boldsymbol{\omega}_k) = \max_{\boldsymbol{\delta}_k \in \boldsymbol{\Delta}} \left\{\boldsymbol{\omega}_k^T \boldsymbol{g}(\boldsymbol{\delta}_k) + \mathbb{E}\left[\hat{J}_{k+1}(\Phi(\boldsymbol{\omega}_k, \boldsymbol{\delta}_k, \boldsymbol{y}_{k+1}))\right]\right\}. \tag{26}$$

B. An Upper Bound of the Reward Function

After reducing the complexity caused by the global constraint, the optimization problem remains computationally expensive to solve, because of the uncountably infinite belief space. To resolve this issue, we explore several properties of the reward-to-go function.

Lemma 2. *The function $\underline{\hat{J}}_k(\boldsymbol{\omega}_k)$ is positively homogeneous of degree* 1.

Proof. We prove this lemma by induction. First, we notice that $\underline{\hat{J}}_L$ is positively homogeneous of degree 1 since

$$\underline{\hat{J}}_L(\mu \cdot \boldsymbol{\omega}_L) = \max_{\boldsymbol{\delta}_L \in \boldsymbol{\Delta}} \left[\mu \cdot \boldsymbol{\omega}_L^T \cdot \boldsymbol{g}(\boldsymbol{\delta}_L)\right] = \mu \cdot \underline{\hat{J}}_L(\boldsymbol{\omega}_L). \tag{27}$$

Now assume this property holds for $2 \leq k \leq L$, then $\underline{\hat{J}}_{k-1}$ is positively homogeneous of degree 1 because

$$\underline{\hat{J}}_{k-1}(\mu \boldsymbol{\omega}_{k-1}) = \max_{\boldsymbol{\delta}_{k-1} \in \boldsymbol{\Delta}} \left[\mu \boldsymbol{\omega}_{k-1}^T \boldsymbol{g}(\boldsymbol{\delta}_{k-1}) + \underline{\hat{J}}_k(\Phi(\mu \boldsymbol{\omega}_{k-1}))\right] \stackrel{(a)}{=} \mu \max_{\boldsymbol{\delta}_{k-1} \in \boldsymbol{\Delta}} \left[\boldsymbol{\omega}_{k-1}^T \boldsymbol{g}(\boldsymbol{\delta}_{k-1}) + \underline{\hat{J}}_k(\Phi(\boldsymbol{\omega}_{k-1}))\right]. \tag{28}$$

Equality (a) holds because the belief update function Φ is linear, which concludes on the desired statement. ☐

Lemma 3. *The function $\underline{\hat{J}}_k(\boldsymbol{\omega}_k)$ is a convex function.*

Proof. For $k = L$, the throughput given by each possible decision choice, $\boldsymbol{\omega}_L^T \cdot \boldsymbol{g}(\boldsymbol{\delta}_L)$ is a linear function of $\boldsymbol{\omega}_L$, and the point-wise maximum of linear functions, $\underline{\hat{J}}_L(\boldsymbol{\omega}_L)$, is a convex function. Next, assume $\underline{\hat{J}}_m(\boldsymbol{\omega}_m)$ $(2 \leq m \leq L-1)$ is a convex function. The form of $\underline{\hat{J}}_{m-1}(\boldsymbol{\omega}_{m-1})$ is as follows

$$\underline{\hat{J}}_{m-1}(\boldsymbol{\omega}_{m-1}) = \max_{\boldsymbol{\delta}_{m-1}} \left[\boldsymbol{\omega}_{m-1}^T \boldsymbol{g}(\boldsymbol{\delta}_{m-1}) + \underline{\hat{J}}_m(\Phi(\boldsymbol{\omega}_{m-1}))\right]. \tag{29}$$

For arbitrary decision $\boldsymbol{\delta}_{m-1}$, $\boldsymbol{\omega}_{m-1}^T \cdot \boldsymbol{g}(\boldsymbol{\delta}_{m-1}) + \underline{\hat{J}}_m(\Phi(\boldsymbol{\omega}_{m-1}))$ is the summation of a linear function and a convex function, which is a convex function. Finally, $\underline{\hat{J}}_{m-1}(\boldsymbol{\omega}_{m-1})$ is the point-wise maximum of convex functions, which is a convex function as well. By induction, the reward-to-go function $\underline{\hat{J}}_k(\omega_k), \forall k \in \{1, \ldots, L\}$, is a convex function. □

These two properties of the reward-to-go function enables us to obtain an upper bound of $\underline{\hat{J}}_k(\boldsymbol{\omega}_k)$. If we decompose the belief state as

$$\boldsymbol{\omega}_k = c_1 \boldsymbol{v}_1 + \cdots + c_n \boldsymbol{v}_n, \tag{30}$$

where $\boldsymbol{v}_i$s come from an arbitrary set of belief state vectors and $\sum_{i=1}^n c_i = 1$. Then $\underline{\hat{J}}_k(\boldsymbol{\omega}_k)$ satisfies the following inequality

$$\underline{\hat{J}}_k(\boldsymbol{\omega}_k) \leq \overline{J_k}(\boldsymbol{\omega}_k) = c_1 \underline{\hat{J}}_k(\boldsymbol{v}_1) + \cdots + c_n \underline{\hat{J}}_k(\boldsymbol{v}_n), \tag{31}$$

where $\overline{J_k}(\boldsymbol{\omega}_k)$ represents an upper bound of $\underline{\hat{J}}_k(\boldsymbol{\omega}_k)$. Therefore we can approximate the optimal reward-to-go function by the optimal value of its upper bound. The remaining goal is to optimize each term of the sum on the right-hand-side of (31).

C. Upper Bound of the Future Reward

The previous approximation enables us to obtain an approximate value for the reward-to-go function $\hat{J}_k(\boldsymbol{\omega}_k)$ by a linear combination of $\hat{J}_k(\boldsymbol{v}_i)$. However, it remains difficult to compute $\hat{J}_k(\boldsymbol{v}_i)$. We observe that the reward-to-go function (26) consists of two parts: an immediate reward and a future reward. The computational complexity lies in the future reward, because of the uncountably infinite observation space. To resolve this issue, we approximate the future reward by its upper bound, denoted as $\tilde{J}_k(\boldsymbol{\omega}_k)$, which is obtained by assuming no error in future observations. It is obvious that decision making under observation error gives lower rewards. With noiseless observation, delayed observation only results in a finite number of possible belief vectors. In this way we have finite states in the belief space and we can solve the problem by value iteration method, with the following recursion:

$$\tilde{J}_k(\boldsymbol{\omega}_k) = \max_{\boldsymbol{\delta}_k \in \boldsymbol{\Delta}} \Bigg[\boldsymbol{\omega}_k^T \boldsymbol{g}(\boldsymbol{\delta}_k) + \sum_{\boldsymbol{\omega}_{k+1}} P(\boldsymbol{\omega}_{k+1} | \boldsymbol{\omega}_k) \tilde{J}_{k+1}(\boldsymbol{\omega}_{k+1}) \Bigg], \tag{32}$$

which enables us to obtain upper bounds for future rewards iteratively.

With approximated future rewards, we can optimize each individual term of the sum on the right-hand-side of (31) and the optimal decision for arbitrary belief state $\boldsymbol{\omega}_k$ can be constructed by a stochastic time-sharing approach.

V. Numerical Results

In this section, we numerically evaluate the performance of the CTSS method presented in Section IV. We consider a CM-UAN with $N_P = 5$ PUs and $N_S = 5$ SUs. The end-to-end distances for both the PU and SU multihop networks are set as 10 km. The carrier frequency is $f = 20$ kHz with 5 kHz bandwidth. The total transmission power for the PUs is 5 W and 4 W for the SUs. Taking both propagation delay and transmission delay into account, the time slot length is set to 2 s. The performance is evaluated via simulations over a finite horizon of $L = 1000$ slots, and each Monte Carlo run is repeated 100 times.

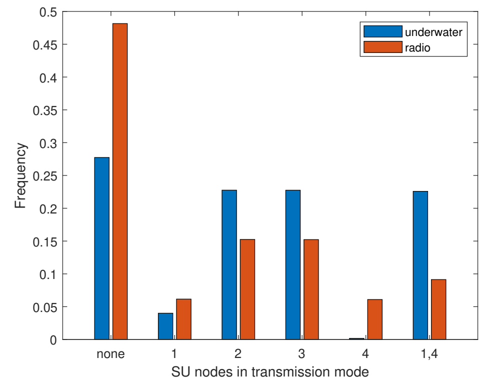

Fig. 2. Frequency of SU decisions.

Our proposed CTSS directly exploits the strong attenuation as a function of range in UACs. To underscore this feature, we simulate a terrestrial radio network by modifying the high propagation delay and using free space attenuation [19]. A simplified algorithm based CTSS is applied to the terrestrial radio network. Fig.2 plots decision frequency for the UAN and radio networks where the x-axis describes which node is transmitting. There are two noteworthy observations. First, SUs in radio network stay idle for a larger fraction of time. Second, SUs in the radio network are less motivated to let two nodes (1 and 4) transmit at the same time, because they will interfere with each other in a way that does not occur in the underwater network. The high path loss challenges scheduling complexity by necessitating a multihop structure, but it provides spatial reuse opportunities as well.

For comparison purposes, we consider a fixed threshold time division multiplexing (TDM) algorithm. In order to take the constraint into account, the total L horizon is divided into two segments, where SUs are only allowed to transmit in the second segment. The length of the each segment is determined by achieving the PU constraint. Based on channel measurements, the SUs make a decision to transmit if the channel occupancy probability is lower than a fixed threshold. Here we select the fixed threshold to be 0.5 such that the optimal decision maximizes the probability of successful transmission. In addition to the static TDM scheme, we also compute the PU throughput in the absence of any SUs.

Fig.3 plots both PU throughput and total (PU+SU) throughput as a function of the packet arrival rate α. Here, we fix the PU throughput degradation coefficient γ to be 0.85 and provide the curve depicting this constraint (PU constraint). We notice that both CTSS and the static TDM scheme satisfy the PU throughput constraint while CTSS is able to get closer to the actual constraint. In terms of total throughput, the proposed

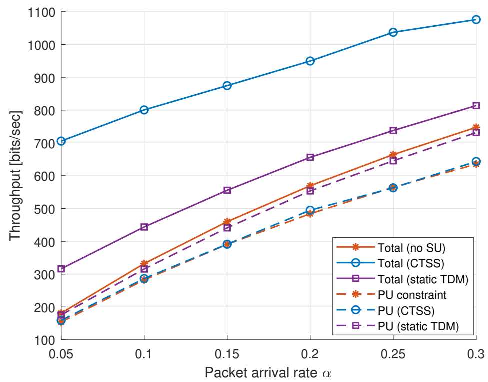

Fig. 3. Throughput versus packet arrival rate α.

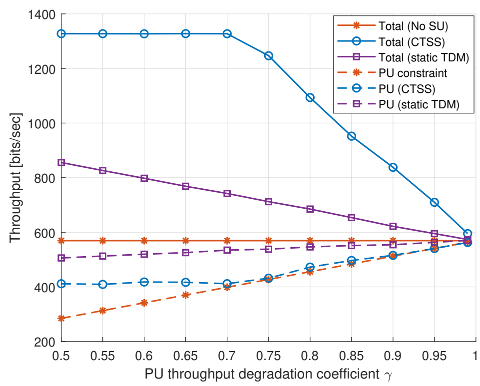

Fig. 4. Throughput versus degradation coefficient γ.

scheme provides 114% improvement on average, compared with the case with no SU transmission, while the static TDM scheme provides 27% improvement on average.

Fig.4 plots throughput as a function of PU throughput degradation coefficient γ where the packet arrival rate α is fixed to be 0.2. We observe that when γ is low, CTSS does not suppress PU throughput as much as possible. This is expected since, with lower γ, after some point, PU transmission becomes more rewarding. In terms of total throughput, CTSS provides 92% improvement on average, while the static TDM scheme provides 25% improvement on average. Additionally, notice that the system throughput of the proposed scheme decreases when γ increases to large values. When $\gamma = 1$, CTSS does not improve throughput as no interference is allowed.

Lastly, with fixed end-to-end distance, we investigate the relationship between throughput and the number of relays by simulation, which indicates there is a clear optimal number of relays that maximizes total throughput. Analysis of the optimal relay number will be conducted in our future work.

VI. Conclusions

In this paper, we addressed the scheduling problem of a CM-UAN with an interference constraint. We first propose a model for system behavior which captures the unique characteristics of UAC via POMDP. The optimal dynamic programming solution is investigated and shown to be computationally expensive. By proving key properties of the objective function, an approximate scheme, CTSS, is proposed. This scheme uniquely exploits the strong attenuation in UACs to enable increased sharing of the temporal resource. The performance of our scheme is evaluated via numerical results, which shows that the CTSS scheme is well matched to the underwater channel and provides significant throughput gain with limited loss to the primary user.

References

[1] Y. Luo, L. Pu, M. Zuba, Z. Peng, and J.-H. Cui, "Challenges and opportunities of underwater cognitive acoustic networks," *IEEE Transactions on Emerging Topics in Computing*, vol. 2, no. 2, pp. 198–211, 2014.

[2] E. M. Sozer, M. Stojanovic, and J. G. Proakis, "Underwater acoustic networks," *IEEE journal of oceanic engineering*, vol. 25, no. 1, pp. 72–83, 2000.

[3] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE communications magazine*, vol. 47, no. 1, pp. 84–89, 2009.

[4] M. Stojanovic, "On the relationship between capacity and distance in an underwater acoustic communication channel," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 11, no. 4, pp. 34–43, 2007.

[5] W. Zhang, M. Stojanovic, and U. Mitra, "Analysis of a linear multihop underwater acoustic network," *IEEE Journal of Oceanic Engineering*, vol. 35, no. 4, pp. 961–970, 2010.

[6] C. Carbonelli, S.-H. Chen, and U. Mitra, "Error propagation analysis for underwater cooperative multi-hop communications," *Ad Hoc Networks*, vol. 7, no. 4, pp. 759–769, 2009.

[7] N. Baldo, P. Casari, and M. Zorzi, "Cognitive spectrum access for underwater acoustic communications," in *ICC Workshops-2008 IEEE International Conference on Communications Workshops*. IEEE, 2008, pp. 518–523.

[8] B. Mishachandar and S. Vairamuthu, "An underwater cognitive acoustic network strategy for efficient spectrum utilization," *Applied Acoustics*, vol. 175, p. 107861, 2021.

[9] W. Chen and U. Mitra, "Packet scheduling for multihopped underwater acoustic communication networks," in *OCEANS 2007*. IEEE, 2007, pp. 1–6.

[10] W. Chen, M. J. Neely, and U. Mitra, "Delay-constrained energy-efficient scheduling over a multihop link," in *2007 IEEE International Symposium on Information Theory*. IEEE, 2007, pp. 1406–1410.

[11] S. Basagni, C. Petrioli, R. Petroccia, and M. Stojanovic, "Choosing the packet size in multi-hop underwater networks," in *OCEANS'10 IEEE SYDNEY*. IEEE, 2010, pp. 1–9.

[12] R. Cao, L. Yang, and F. Qu, "On the capacity and system design of relay-aided underwater acoustic communications," in *2010 IEEE Wireless Communication and Networking Conference*. IEEE, 2010, pp. 1–6.

[13] M. Vajapeyam, S. Vedantam, U. Mitra, J. C. Preisig, and M. Stojanovic, "Distributed space–time cooperative schemes for underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 33, no. 4, pp. 489–501, 2008.

[14] Z. Han, Y. L. Sun, and H. Shi, "Cooperative transmission for underwater acoustic communications," in *2008 IEEE International Conference on Communications*. IEEE, 2008, pp. 2028–2032.

[15] E. Demirors, G. Sklivanitis, T. Melodia, S. N. Batalama, and D. A. Pados, "Software-defined underwater acoustic networks: toward a high-rate real-time reconfigurable modem," *IEEE Communications Magazine*, vol. 53, no. 11, pp. 64–71, 2015.

[16] L. M. Brekhovskikh, Y. P. Lysanov, and R. T. Beyer, "Fundamentals of ocean acoustics," 1991.

[17] J. M. Jornet, M. Stojanovic, and M. Zorzi, "On joint frequency and power allocation in a cross-layer protocol for underwater acoustic networks," *IEEE Journal of Oceanic engineering*, vol. 35, no. 4, pp. 936–947, 2010.

[18] D. P. Bertsekas *et al.*, *Dynamic programming and optimal control: Vol. 1*. Athena scientific Belmont, 2000.

[19] J. C. Whitaker, *The electronics handbook*. Crc Press, 2018.