Energy Management and Power Allocation for Underwater Acoustic Sensor Network

Lianyou Jing, *Student Member, IEEE*, Chengbing He, *Member, IEEE*, Jianguo Huang, *Senior Member, IEEE*, and Zhi Ding, *Fellow, IEEE*

***Abstract*— This paper investigates energy allocation in underwater acoustic nodes powered by energy harvesting. Our goal is to maximize the expected total amount of delivered data over a finite time slots. Two scenarios are considered for different knowledge levels of channel state information (CSI). In one scenario, the transmitter receives CSI at the end of each data transmission epoch and we consider the energy allocation problem for sensing and transmission. In the second scenario, the transmitter receives delayed CSI after multiple time slots, while only the energy allocation for transmission is considered. The underwater acoustic channel is modeled as a finite-state Markov chain to characterize its time varying nature. We employ a stochastic dynamic programming (DP) approach to derive the optimal allocation policy for both scenarios. To reduce the inherent computation complexity of DP approach, we also present a suboptimal algorithm by analyzing the structure of DP solution. Our results show the DP approach achieves substantial performance improvement that is preserved substantially by the suboptimal algorithm to provide a good performance-complexity tradeoff.**

***Index Terms*— Dynamic programming, energy allocation, energy harvesting, underwater acoustic sensor network.**

I. Introduction

UNDERWATER acoustic sensor networks and their acoustic data communications have found broad applications in areas such as environmental monitoring, offshore exploration, and disaster warning. Often, these networks are required to provide delay sensitive data collection and remain active for a prolonged period of deployment. Most underwater sensors cannot rely on traditional electrical power source because of environmental or practical constraints. As a result, limited lifetime is one key issue that confronts underwater sensing and communication networks more than their land-based counterpart. In general, underwater sensors can be powered by battery or through energy harvesting. When powered by battery, sensor and transmitter lifetime is severely limited. For rechargeable batteries, external battery charging can be very costly or even impractical. Hence, energy harvesting offers a potential alternative energy source for sensor battery recharging.

Energy harvesting demonstrates strong potential for powering underwater sensors by acquiring energy from the environment [1]–[5], such as solar, wind, tidal waves, among others. In [1], the authors proposed an environmental monitoring framework based on a wireless sensor networks powered by adaptive solar-energy harvesting mechanisms and tandem batteries. The study in [2] considered an electronic circuit for harvesting energy trickling from benthic sources and the long-term performance in powering sensors and devices in a littoral tidal basin. In [3], the author analyzed the applicability of solar cells as a power source for medusa-inspired biomimetic vehicles. In [4], the author proposed an underwater energy harvesting system based on plucked-driven piezoelectric. Experimental results show that the proposed energy harvester achieves a maximum power density of $350\mu W/cm^3$.

Harvested energy can be stored in powering batteries to prolong sensor network lifetime. However, energy harvesting based networks also present distinct challenges. Generally speaking, harvested energy may be sporadic and unpredictable. Thus, judicious use and management of available energy for optimized sensor data collection and transmission becomes an important issue to ensure successful utility of underwater networks. This work addresses the energy management problem for underwater sensors in acoustic communication networks.

For terrestrial networks, a number of research papers have studied optimization of transmission policy with an energy-harvesting transmitter [6], [7]. In [8]–[10], the authors studied the throughput maximization problem over a finite number of transmission blocks for energy harvesting systems. In [11] and [12], the authors investigated optimal power allocation to minimize the average outage probability in block-fading channels under energy harvesting constraint. The study in [13] and [14] considered the power allocation of point-to-point (P2P) communication system with joint energy harvesting and grid power supply. In [15]–[17], the authors studied the power allocation problem in multi-senor remote estimation systems. Additionally, [18] and [19] considered the effect of batteries suffering from charging and discharging inefficiency as energy storage devices.

However, the aforementioned works generally tend to stress the energy consumption during the transmission phase but focusing less on the data collection phase of sensors, i.e., the

Manuscript received June 16, 2017; accepted July 29, 2017. Date of publication August 9, 2017; date of current version September 8, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61471298, Grant 61531015, and Grant 61401499, and in part by the Fundamental Research Funds for the Central Universities under Grant 3102016AXXX03. The associate editor coordinating the review of this paper and approving it for publication was Dr. Amitava Chatterjee. *(Corresponding author: Lianyou Jing.)*

L. Jing, C. He, and J. Huang are with the School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: jingly369@mail.nwpu.edu.cn; hcb@nwpu.edu.cn; jghuang@nwpu.edu.cn).

Z. Ding is with the Department of Electrical and Computer Engineering, University of California at Davis, Davis, CA 95616 USA (e-mail: zding@ucdavis.edu).

Digital Object Identifier 10.1109/JSEN.2017.2737229

energy expended by sensing devices during data collection. Underwater sensors often need to collect information of interests through transducing, sampling, quantizing, compressing and storing. Our study takes into consideration the dual tasks of sensing and wireless communications in an underwater environment.

In this paper, we consider sensors equipped with underwater acoustic transmitters powered by energy harvesting. As we know, underwater acoustic communication channels are notoriously difficult link media [20]. Underwater acoustic channels usually exhibit time varying characteristics that depend on environmental and seasonal conditions in both short and long time scales. Reception failure in underwater systems is common, which necessitates error control mechanisms such as automatic repeat request (ARQ) protocol to retransmit dropped packets for reliability enhancement [21], [22]. Under ARQ, a transmitter sends multiple packets to its receiver. For each successful or failed packet, the destination sends back an acknowledgement (ACK) or an non-acknowledgement (NACK) back to the transmitter, respectively. Given an NACK, the transmitter would retransmit the failed packet up to a maximum retrial limit. In our system model, we shall consider such an ARQ mechanism and also assume that channel state information (CSI) is fed back to the transmitter. We model the time-varying underwater acoustic channel as a finite state Markov chain, which suitably represents the channel characteristics [23], [24]. We also model the rate of energy harvesting as a first-order Markov process. Given a finite energy capacity (battery) and a data buffer at the sensor, the transmitter must allocate its energy to either sense and collect data, or transmit data stored in the buffer. Our overall objective is to optimize energy allocation policies so as to maximize the expected number of successful packets over a finite-time horizon subject to both the buffer overflow and the energy capacity constraints.

There exist several related works that consider the dual sensor tasks. The authors of [25] studied the problem of energy allocation for data acquisition and transmission in general wireless sensor networks. In [26], the authors extended the consideration to include a single source with finite battery and data buffer. In [27], the authors studied the energy allocation problem between sensing and transmission to maximize the total amount of data transmitted. In [28], the objective is to minimize the mean squared-error distortion between the Gaussian source samples and their reconstructions over a flat fading channel. However, the existing works typically do not consider the effect of ACK/NACK feedback and as well as the impact of delayed CSI.

Our novel contributions in this paper are as follows. We optimize the allocation of limited energy capacity to maximize the number of expected successful packets within a finite time interval without data buffer overflow by modeling both energy harvesting and time-varying acoustic channel as Markov processes. We incorporate the practical impact of ACK/NACK feedback along with CSI availability. We considered cases of both delayed and non-delayed CSI feedbacks. Without feedback delay, the transmitter receives the feedback messages (CSI and ACK/NACK) at the end of each time slot and therefore knows the immediately previous CSI. We develop an optimized energy allocation policy to allocate energy to sense and transmit based on dynamic programming (DP) approach. In second scenario involving feedback delay, the received CSI contains less accurate information but is more practical due to round-trip delay. For this scenario, we allow infinite data backlog and focuses on the allocation of the transmission energy. To reduce the inherent high complexity of DP approach, we also propose a suboptimal online algorithm to tackle the performance-complexity tradeoff.

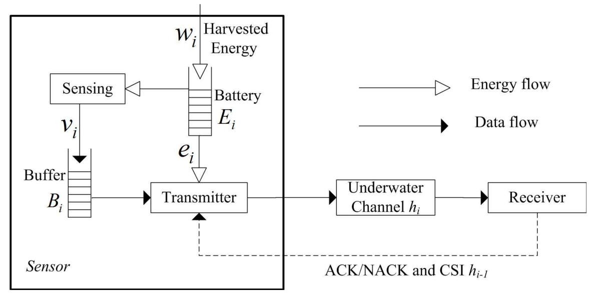


Fig. 1. The communication system model with energy harvesting.

The rest of this paper is organized as follows. Section II presents the system model for energy harvesting. We formulate the optimization problem in section III. We propose the optimization solution based on the principle of dynamic programming in section IV. Section V address the performance-complexity tradeoff by presenting a suboptimal algorithm to reduce the computational complexity of DP solution. Section VI provides numerical results for the proposed solutions and section VII concludes the paper.

II. System Model

We consider a sensor-to-collector connection that involves a point-to-point (P2P) underwater acoustic communication link as shown in Fig. 1. The transmitter has a rechargeable battery with capacity E_{max} and a data buffer with size B_{max} data packets. The transmit node sensor queries an information source/transducer and stores the information data in its data buffer as packets. It also sends the data packets to a destination node over a time-varying underwater acoustic link. We do allow simultaneous sensing and transmission if needed.

We let the channel access be time-slotted with N total slots per scheduling period, indexed by $\{1, \ldots, N\}$. The slot time is denoted by T_{slot}. The sending node harvests energy from environment to store in a rechargeable fuel cell (battery). The harvested energy in i-th slot is denoted by w_i and assumed to be available for future time slots. We let energy be replenished at the end of each slot and assume that the transmitter is aware of the value of w_{i-1} at the beginning of slot i but unaware of the value of w_i. In our formulation, we divide the possible harvested energy into Q levels denoted by $w_i \in \{\delta_1, \ldots, \delta_Q\}$. We model the energy harvesting process as a first-order stationary Markov process with known state transition probabilities (estimated from practical tests).

Similarly, we model the acoustic channel state by a finite-state Markov chain [23], [24]. Assuming slowly time-varying and quasi-stationary channels, the channel state is stationary during each transmission slot and is allowed to change in the subsequent slot according to the Markov model. Let h_i denote the channel gain in time slot i which is quantized into K levels, $h_i \in \{\gamma_1, \ldots, \gamma_K\}$. In the system model, the state transition probabilities $P_h(h_i|h_{i-1})$ are given.

We utilize the ARQ error control at the data link-layer. Each data packet consists ℓ payload information bits and only one packet transmission may be attempted in one slot. If the destination fails to decode the transmitted packet successfully, it will send a NACK message to the transmitter. The transmitter, upon receiving a NACK, shall retransmit the failed packet in the next slot until it receives an ACK or is out of energy.

In time slot i, we let e_i denote the energy used for transmission. It is clear that the packet error probability (PEP) P_e depends on the forward error correction (FEC) technique, the packet length, the bit error rate, and the FEC redundancy. In our problem, the packet length and FEC are fixed. Then, the PEP is a monotonically decreasing function of e_i. For example, for low rate modulation such as BPSK and convolution code with codeword (packet) length ℓ in bits, the packet error probability can be modeled as [29]

$$P_e(h_i, e_i) = 1 - \left(1 - \sum_{d=d_{free}}^{\ell} A_d Q\left(\sqrt{\frac{2dh_ie_i}{T_{slot}N_f f_B}}\right)\right)^{\ell} \quad (1)$$

where A_d is the weight spectral coefficient of the convolution code whereas d_{free} is the free distance of the convolutional code. $Q(\cdot)$ is the tail Gaussian integration function, N_f is the noise power spectral density and f_B is the signal bandwidth.

For sensing in time slot i, we let υ_i denote the amount of data packet generated by sensor and let s_i be the amount of energy used for sensing. In general, υ_i is a monotonically nondecreasing and concave function in s_i. In this work, we assume a linear relationship, i.e. $\upsilon_i = \lambda s_i$.

III. Problem Formulation

We now describe the problem and our design objective. Our energy allocation problem aims to maximize the expected total transmission packets within a finite time duration. We consider two different scenarios in which either the feedback CSI is updated at the end of the present slot or with a finite delay.

A. Scenario 1: Timely CSI Feedback Without Delay

Fig. 2 shows update process of energy battery and data buffer. Let E_i denote the stored energy in the battery and B_i be the number of packets stored in its data buffer at beginning of i-th slot for $i = 1, 2, \cdots, N$. At slot i, the transmitter energy consumption equals to e_i to for transmission and s_i for sensing, respectively. Meanwhile, harvested energy ω_i is stored in the fuel cell. Accordingly, we can update the battery level E_i as follows

$$E_{i+1} = \min\{E_i - (e_i + s_i) + \omega_i, E_{max}\}, \quad (2)$$

in which E_{max} is the capacity of the battery. Equation (2) also indicates that certain amount of energy may need to be consumed to avoid battery overflow. Since we assume that energy is replenished at the end of each slot, we have another causality constraint about the battery charging

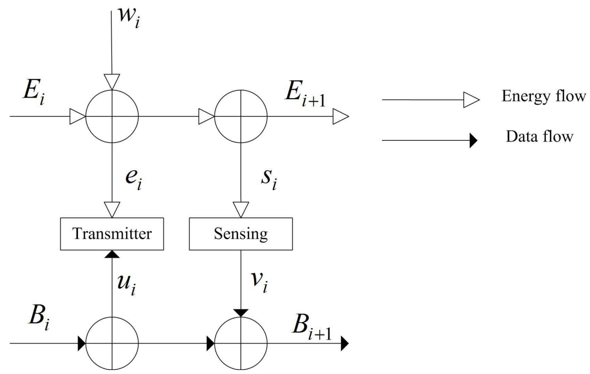


Fig. 2. The update process of energy battery and data buffer.

$$e_i + s_i \leq E_i. \quad (3)$$

That means the total amount of transmission energy and sensing energy should not exceed the energy stored in the battery.

Similarly, the amount of data packet in the buffer B_i is updated as follows

$$B_{i+1} = \min\{[B_i - u_i]^+ + \upsilon_i, B_{max}\} \quad (4)$$

where $[x]^+ = \max\{x, 0\}$, and u_i is the transmitted packet in i-th slot. in this paper, we consider a single packet scheme. Thus,

$$u_i = \begin{cases} 0, & \text{with probability } P_e(h_i, e_i) \\ 1, & \text{with probability } P_c(h_i, e_i) \end{cases} \quad (5)$$

where $P_c(h_i, e_i) = 1 - P_e(h_i, e_i)$ is the packet correct probability. Thus, the amount of data delivered in i-th time slot equal

$$\mathbb{E}[u_i] = 1 - P_e(h_i, e_i) = P_c(h_i, e_i) \quad (6)$$

and the objective function, i.e., the expected sum data packets within the scheduling period of N slots, are given by

$$J_1 = \max_{e_i, s_i} \sum_{i=1}^{N} \mathbb{E}[u_i]. \quad (7)$$

Consequently, energy allocation can be optimized by solving the following constrained optimization problem

$$\begin{aligned} &\max_{e_i, s_i} \sum_{i=1}^{N} \mathbb{E}[u_i] \\ &\text{s.t. } (2), (3), (4), \text{ and} \\ &\quad e_i \geq 0, s_i \geq 0 \end{aligned} \quad (8)$$

Under the no-delay CSI feedback environment, we consider how to allocate the available energy for sensing and transmission in each slot to maximize the expected throughout across the N time slots. At time slot i, the source node need to provide a trade-off between the s_i and e_i. If the energy allocated for sensing s_i is too large, the sensor may successfully acquire enough information packets in the buffer. However, the remaining energy that can be used for transmission will be small, likely resulting in a larger packet error probability and potential energy waste because of higher likelihood of packet failure and the need to retransmit the failed packet. On the other hand, if the e_i is too large, the packet error probability can be substantially reduced. However, the number of stored packets in the buffer may be low and insufficient to benefit from the lower packet error rate (PER) for transmission.

B. Scenario 2: Delayed Feedback

Without feedback delay, the transmitter always receives the feedback message at the start of each time slot. Hence, the transmitter knows the past slot CSI. However, in many practical systems, CSI feedback delay may be unavoidable because of roundtrip transmission delay or block based acknowledgement mode in which multiple ACK/NACK messages are sent periodically to reduce feedback overhead. For this reason, we now consider such a more practical case that the transmitter receives CSI and ACK/NACK after some time delay.

Delayed CSI feedback is very likely for underwater acoustic communications. The speed of sound travels at approximately 1500 m/s. A transmitter would prefer to send multiple packets in a row without waiting for individual feedbacks so as to avoid unnecessary network waste. This means that the feedback CSI becomes less accurate for current slot.

Let m denote the time slots delay of feedback message. The transmitter should determine whether or not to retransmit the current packet when it receives a ACK/NACK messages delayed by a fixed m time slots. Before receiving the message, the packet will not be dropped. In this way, the data buffer is updated as follows

$$B_{i+1} = \min\{[B_i - u_{i-m}]^+ + \lambda s_i, B_{max}\}, \tag{9}$$

which

$$u_{i-m} = \begin{cases} 0, & \text{with probability } P_e(h_{i-m}, e_{i-m}) \\ 1, & \text{with probability } P_c(h_{i-m}, e_{i-m}) \end{cases} \tag{10}$$

It is clear that the buffer state is no longer a simple Markov Chain. The longer the delay m, the more complex the buffer dynamics. To avoid over-complication, we consider a more restricted scenario of full data buffer by *assuming transmitters to have unlimited payload data to transmit.* In other words, for delayed feedback, we neglect the effect of energy in sensing. Fig. 3 shows the communication system model for scenario 2.

Then, the battery level is updated as follows

$$E_{i+1} = \min\{E_i - e_i + w_i, E_{max}\} \tag{11}$$

And the causality constraint on the battery is given by

$$e_i \leq E_i \tag{12}$$

the objective function is also modified into

$$J_1^m = \max_{e_i} \sum_{i=1}^{N} \mathbb{E}[u_i] = \max_{e_i} \cdot \sum_{i=1}^{N} [1 - P_e(h_i, e_i)] \tag{13}$$

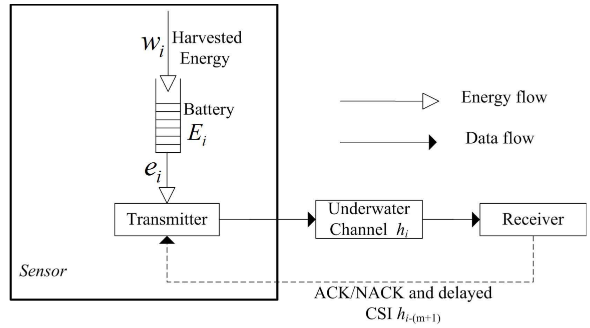


Fig. 3. The communication system model with energy harvesting for scenario 2.

Consequently, the optimal energy allocation can be obtained by solving the following optimization problem

$$\begin{aligned} &\max_{e_i} \sum_{i=1}^{N} \mathbb{E}[u_i] \\ &\text{s.t. } (11), (12), \quad e_i \geq 0 \end{aligned} \tag{14}$$

C. Comments and Remarks

In our models, both the buffer evolution and battery evolution are Markovian. Generally speaking, the optimization problems in (8) and (14) are hard to solve independently because the constraints for i-th slot depend on the $(i-1)$-th slot. Since packet error is a stochastic process, these two problems can be viewed as one of sequential decision making under channel state and energy harvesting rate uncertainties. For such problem, we can use a finite horizon Dynamic Programming (DP) [30] approach to achieve energy optimality, to be discussed next.

IV. Optimal Solution

In this section, we employ a finite horizon stochastic DP approach to solve the energy allocation problems of (8) and (14). DP approach divides the sequential optimization problem into a series of single stage problems. According to the Bellman's Optimality Principle [30], DP approach is a well known optimum mathematical method for solving the sequential decision making problems. There are a number of existing works employing DP approach to solve similar problems [24], [31]–[33].

A. Scenario 1: Timely Feedback Without Delay

In this problem, we consider the case which the CSI is updated together with ACK/NACK message at the end of each transmission slot. Namely, the transmitter knows the value of h_{i-1} and the ACK/NACK message of last slot at the beginning of i slot. The transmitter will decide whether or not to retransmit the packet depending on the ACK/NACK feedback.

Let $\mathbf{c}_i = (h_{i-1}, w_{i-1}, E_i, B_i)$ denote the state of the system in time slot i which includes the channel state h_{i-1}, the incoming harvested energy w_i, the stored energy in the battery E_i, and the buffer length B_i. We assume the initial state $\mathbf{c}_1 = (h_0, w_0, E_1, B_1)$ to be known. Starting from

Eqs. (2) and (4), the state transition probability is given by

$$\begin{aligned} &P(\mathbf{c}_{i+1}|\mathbf{c}_i, s_i, e_i) \\ &= P(h_i, w_i, E_{i+1}, B_{i+1}|h_{i-1}, w_{i-1}, E_i, B_i, s_i, e_i) \\ &= P(h_i, B_{i+1}|h_{i-1}, B_i, s_i, e_i)P(w_i, E_{i+1}|w_{i-1}, E_i, s_i, e_i) \\ &= P(B_{i+1}|h_i, B_i, s_i, e_i)P(h_i|h_{i-1}) \\ &\quad \times P(E_i|w_{i-1}, E_i, s_i, e_i)P(w_i|w_{i-1}) \\ &= \begin{cases} P(h_i|h_{i-1})P(w_i|w_{i-1}), & \text{if (2) and (4) are satisfied} \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{15}$$

Given the initial state $\mathbf{c}_1$, the maximum total expected number of packets $J_1(\mathbf{c}_1)$ can be obtained by recursively computing $J_N(\mathbf{c}_N)$, $J_{N-1}(\mathbf{c}_{N-1})$, $\cdots$, $J_2(\mathbf{c}_2)$ based on Bellman equation [30]. For the last time slot N within the scheduling period, the transmitter will use all remaining energy for transmission, i.e., $e_N = E_N$, to maximize the transmitted packet. Thus, we have

$$\begin{aligned} J_N(h_{N-1}, w_{N-1}, E_N, B_N) &= \max_{e_N, s_N} \mathbb{E}_{h_N}[u_N|h_{N-1}] \\ &= \sum_{k=1}^{K} P_c(\gamma_k, E_N)P(h_N = \gamma_k|h_{N-1}) \end{aligned} \tag{16}$$

At time slot i, we have

$$\begin{aligned} &J_i(h_{i-1}, w_{i-1}, E_i, B_i) \\ &= \max_{e_i, s_i} \Big\{\mathbb{E}_{h_i}[u_i|h_{i-1}] \\ &\quad + \mathbb{E}_{h_i, w_i}\left[J_{i+1}(h_i, w_i, E_{i+1}, B_{i+1})|h_{i-1}, w_{i-1}\right]\Big\} \qquad (17a) \\ &= \max_{e_i, s_i} \Bigg\{\sum_{k=1}^{K} P_c(\gamma_k, e_i)P(h_i = \gamma_k|h_{i-1}) \\ &\quad + \mathbb{E}_{h_i, w_i}\left[J_{i+1}(h_i, w_i, E_{i+1}, B_{i+1}^{(1)})|h_{i-1}, w_{i-1}\right] P_e(h_i, e_i) \\ &\quad + \mathbb{E}_{h_i, w_i}\left[J_{i+1}(h_i, w_i, E_{i+1}, B_{i+1}^{(2)})|h_{i-1}, w_{i-1}\right] P_c(h_i, e_i)\Bigg\} \qquad (17b) \\ &= \max_{e_i, s_i} \Bigg\{\sum_{k=1}^{K} P_c(\gamma_k, e_i)P(h_i = \gamma_k|h_{i-1}) \\ &\quad + \sum_{q=1}^{Q}\sum_{k=1}^{K} J_{i+1}(\gamma_k, \delta_q, E_{i+1}, B_{i+1}^{(1)})P_e(\gamma_k, e_i) \\ &\quad \cdot P(h_i = \gamma_k|h_{i-1})P(\delta_q|w_{i-1}) \\ &\quad + \sum_{q=1}^{Q}\sum_{k=1}^{K} J_{i+1}(\gamma_k, \delta_q, E_{i+1}, B_{i+1}^{(2)})P_c(\gamma_k, e_i) \\ &\quad \cdot P(h_i = \gamma_k|h_{i-1})P(\delta_q|w_{i-1})\Bigg\} \qquad (17c) \end{aligned}$$

in which

$$B_{i+1}^{(1)} = \min\{B_i + \lambda s_i, B_{max}\} \tag{17d}$$

$$B_{i+1}^{(2)} = \min\{[B_i - 1]^+ + \lambda s_i, B_{max}\} \tag{17e}$$

Note that $B_{i+1}^{(1)}$ and $B_{i+1}^{(2)}$ represent the numbers of packets in the data buffer at the beginning of slot $i+1$, with the probability of $P_e(h_i, e_i)$ and $P_c(h_i, e_i)$, respectively. The first term in (17a) represents the expected reward of current slot. The second term in (17a) is the expected future reward accumulated from slot $i+1$ to slot N. The DP approach includes two steps. In step 1, the sensor computes the Eq. (17a) to find the optimal $(e_i^*(\mathbf{c}_i),\ s_i^*(\mathbf{c}_i))$ for each system state of each time slot and records it as a lookup table. In step 2, the sensor sets the system state $\mathbf{c}_i = (h_{i-1}, w_{i-1}, E_i, B_i)$ and chooses the optimal $(e_i^*(\mathbf{c}_i), s_i^*(\mathbf{c}_i))$ based on the lookup table. The details of proposed optimal DP policy are summarized in Algorithm 1.

Algorithm 1 Energy Allocation Algorithm Based on Dynamic Programming for Scenario 1

1: Calculate $J_N(h_{N-1}, w_{N-1}, E_N, B_N)$, $\forall h_{N-1}$, $\forall w_{N-1}$, $\forall E_N$, $\forall B_N$ based on (16) and save in the table.
2: $i = N - 1$.
3: $\mathbf{c}_i = (h_{i-1}, w_{i-1}, E_i, B_i)$
4: **while** $i > 0$
5: Calculate $J_i(h_{i-1}, w_{i-1}, E_i, B_i)$, $\forall h_{i-1}, \forall w_{i-1}, \forall E_i, \forall B_i$ based on (17a)
6: Find the optimal $(e_i^*(\mathbf{c}_i),\ s_i^*(\mathbf{c}_i))$ to maximize $J_i(h_{i-1}, w_{i-1}, E_i, B_i)$
7: $i = i - 1$;
8: **end while**
9: **for** $i = 1$; $i \leq N$; $i = i + 1$ **do**
10: Find the optimal $(e_i^*(\mathbf{c}_i), s_i^*(\mathbf{c}_i))$ based on line 5.
11: Update the battery and data buffer based on (2) and (4)
12: **end for**

The DP approach can achieve optimal solution for this problem; however, its drawback is the high computational complexity. The DP method needs to first constructs a look-up table containing the optimal solution for each system state as shown in Algorithm 1. Once we obtain the table, we can easily find the optimal solution by table lookups. The computational complexity of DP method comes mainly from building the look-up table. A DP problem could be decomposed into multiple sub-problems and the computational complexity of DP problem has linear increase with the number of sub-problems (time slots in our problem). Due to the table contains all of system state at each time slot, the computational complexity of each subproblem is related to the number of system state. In order to evaluate the computational complexity, we set a foot-size ρ to search the optimal energy allocation. Then, the total number of state is $(\frac{E_{max}}{\rho} \times B_{max} \times K \times Q)$. For each time slot, we compute Eq. (17a) for each state. It results in long execution time as E_{max} getting large and ρ getting small. It will not be a big problem when we build a look-up table before transmission. However, in some case we may need to update some statistical information, such as channel transition probability, energy harvesting rate, which introduces a high real-time computation complexity. Then, we also propose a low complexity algorithm in next section.

B. Scenario 2: Delayed CSI and ARQ Feedback

In the second scenario that considers roundtrip delays and bandwidth efficiency, the transmitter receives CSI and

ARQ feedback after m slots. As a result, the transmitter only knows $h_{i-(m+1)}$ at the beginning of i-th slot. In this case, the total number of states grows to the order $O((\frac{E_{max}}{\rho})^{2m+1} \times B_{max} \times K \times Q)$.

To simplify the problem which involves large number of states, we assume the buffer to have an infinite data backlog. As a result, the number of system states is reduced into $\mathbf{c}_i = (w_{i-1}, h_{i-(m+1)}, E_i)$. Therefore, the channel state transition probability is

$$P^{(m+1)}(h_i|h_{i-(m+1)}) = (P_h)^{m+1} \tag{18}$$

When m increases, the feedback CSI becomes less and less informative. For sufficiently large m, the transition probability converges to the steady-state probability vector. Therefore, an upper bound to m denoted by M can be considered. This means that if $m > M$, then the feedback CSI is practically unusable for power control.

We again consider the DP approach. For the last time slot, the expected number of packet delivery is computed as

$$J_N^m(h_{N-(m+1)}, w_{N-1}, E_N) = \sum_{k=1}^{K} P_c(\gamma_k, E_N) P(h_N = \gamma_k | h_{N-(m+1)}) \tag{19}$$

For the i-th ($i > m$) time slot, the transmitter will receive the feedback channel state in each slot. Then,

$$\begin{aligned} J_i^m&(h_{i-(m+1)}, w_{i-1}, E_i) \\ &= \max_{e_i} \Big\{ \mathbb{E}_{h_i}[u_i | h_{i-(m+1)}] \\ &\qquad + \mathbb{E}_{h_i, w_i}\left[J_{i+1}(w_i, h_{i-m}, E_{i+1}) | h_{i-(m+1)}, w_{i-1} \right] \Big\} \\ &= \max_{e_i} \Bigg\{ \sum_{k=1}^{K} P_c(\gamma_k, e_i) P(h_i = \gamma_k | h_{i-(m+1)}) \\ &\qquad + \sum_{q=1}^{Q} \sum_{k=1}^{K} J_{i+1}(\gamma_k, \delta_q, E_{i+1}) \\ &\qquad\qquad \times P(h_{i-m} = \gamma_k | h_{i-(m+1)}) \cdot P(\delta_q | w_{i-1}) \Bigg\} \end{aligned} \tag{20}$$

in which E_{i+1} is updated according to (11) .

When $i \leq m$, the transmitter has not received the feedback message yet and has no information about the channel. We would use the steady-state probability π_k as the probability which the channel at state k.

$$\begin{aligned} J_i^m&(h_{i-(m+1)}, w_{i-1}, E_i) \\ &= \max_{e_i} \left\{ \mathbb{E}_{h_i}[u_i] + \mathbb{E}_{h_i, w_i}\left[J_{i+1}(w_i, h_{i-m}, E_i) | w_i \right] \right\} \\ &= \max_{e_i} \Bigg\{ \sum_{k=1}^{K} P_c(\gamma_k, e_i) \pi_k \\ &\qquad + \sum_{q=1}^{Q} \sum_{k=1}^{K} J_{i+1}(\gamma_k, \delta_q, E_{i+1}) \pi_k P(\delta_q | w_{i-1}) \Bigg\} \end{aligned} \tag{21}$$

When $m > M$, that means the feedback CSI is useless and the channel can be considered as independent. In this case, it will be similar with (21). The details of proposed optimal DP policy are summarized in Algorithm 2.

Algorithm 2 Energy Allocation Algorithm Based on Dynamic Programming for Scenario 2

1: Calculate $J_N^m(h_{N-1}, w_{N-(m+1)}, E_N)$, $\forall h_{N-1}$, $\forall w_{N-1}$, $\forall E_N$,based on (19) and save in the table.
2: $i = N - 1$.
3: $\mathbf{c}_i = (h_{i-1}, w_{i-(m+1)}, E_i)$
4: **while** $i > 0$
5: **if** $i > m$
6: Calculate $J_i^m(\mathbf{c}_i)$, $\forall h_{i-1}$, $\forall w_{i-1}$, $\forall E_i$ based on (20)
7: **else**
8: Calculate $J_i^m(\mathbf{c}_i)$, $\forall h_{i-1}$, $\forall w_{i-1}$, $\forall E_i$ based on (21)
9: **end**
10: Find the optimal $e_i^*(\mathbf{c}_i)$ to maximize $J_i^m(\mathbf{c}_i)$
11: $i = i - 1$;
12: **end while**
13: **for** $i = 1$; $i \leq N$; $i = i + 1$ **do**
14: Find the optimal $e_i^*(\mathbf{c}_i)$ based on line 9.
15: Update the battery based on (11).
16: **end for**

V. Suboptimal Myopic Solution for Performance-Complexity Tradeoff

It is apparent that the computational complexity of the optimal DP solution grows exponentially with the battery and buffer capacity and the number of states in the channel. In essence, sophisticated transmitter node may be well equipped with large energy and data capacity that render our DP solutions too complex to implement. In order to address this practical issue, we develop a myopic scheme for the energy allocation to reduce complexity without significant performance loss.

Recall that in the optimal DP method, there are two parts in (17a). The first part is the current reward, and the second part is the expected reward from all future slots. It introduces high computational complexity which becomes a major problem in practical implementation. Instead of compute the expected reward from all future slots, we can consider a myopic solution by focusing only on the expected reward of the next P slots. Since we only compute optimal solution for P slots, this myopic goal can substantially reduce the complexity when N is large.

We first build a $P + 1$ slots (including current slot) look-up table as described in the full DP solution. Namely, we find out an optimal solution for $P + 1$ slots and the reward denoted by $\hat{J}_p(\mathbf{c}_p)$. The compute method is similar with (16) and (17a) but using $(P + 1)$ instead of N. Hence, for the i-th ($i \leq N - P$) slot, the optimal energy allocation to maximize the expected reward of next P slots can be directly obtained by maximizing $\hat{J}_1(\mathbf{c}_i)$. However, when $i > N - P$, the optimal energy allocation to maximize the expected reward of next $N - i$ slots can be get from to maximize $\hat{J}_{i-(N-P)+1}(\mathbf{c}_i)$.

More specific details of suboptimal policy are summarized in algorithm 3.

If we choose $P = 1$, the total expected number of packets is replaced by

$$\hat{J}_i(\mathbf{c}_i) = \max_{e_i,s_i} \left\{ \mathbb{E}_{h_i}[u_i|h_{i-1}] + \mathbb{E}_{h_i,w_i}\left[u_{i+1}|h_{i-1}, w_{i-1}\right] \right\} \tag{22}$$

$$= \max_{e_i,s_i} \left\{ \sum_{k=1}^{K} P_c(\gamma_k, e_i) P(h_i = \gamma_k|h_{i-1}) + \sum_{k=1}^{K}\sum_{q=1}^{Q} P_c(\gamma_k, e_{i+1}) \cdot P(h_{i+1} = \gamma_k|h_{i-1}) P(w_i = \delta_q|w_{i-1}) \right\} \tag{23}$$

where

$$e_{i+1} = \begin{cases} \min\{E_i - (e_i + s_i) + \delta_q, E_{max}\}, & \text{if } B_{i+1} \geq 1 \\ 0, & \text{otherwise} \end{cases} \tag{24}$$

Thus, if $P = 1$, then for the next slot, the low complexity solution reduces to using all the energy to transmit the packet if the data buffer is not empty. In fact, this scheme is a myopic view of the future reward (by only considering the next slot). In fact, it is a greedy strategy of the energy consumption when considering a shortest horizon of $P = 1$. This suboptimal approach determines the current slot energy allocation by developing a greedy approach for the next slot. Even though the next slot does not consume all the energy in the battery, the residue energy can be determined in subsequent steps.

Algorithm 3 Energy Allocation Based on Suboptimal Algorithm

Calculate $\hat{J}_{P+1}(h_P, w_P, E_{P+1}, B_{P+1}), \forall h_P, \forall w_P, \forall E_{P+1}, \forall B_{P+1}$ based on (16) and save in the table.
$i = P$.
$\mathbf{c}_i = (h_{i-1}, w_{i-1}, E_i, B_i)$
while $i > 0$
  Calculate $\hat{J}_i(\mathbf{c}_i), \forall h_{i-1}, \forall w_{i-1}, \forall E_i, \forall B_i$ based on (17a)
  Find the optimal $(e_i^*(\mathbf{c}_i), s_i^*(\mathbf{c}_i))$ to maximize $\hat{J}_i(\mathbf{c}_i)$
  $i = i - 1$;
end while
for $i = 1$; $i \leq N - P$; $i = i + 1$ **do**
  Find the optimal $(e_i^*(\mathbf{c}_i), s_i^*(\mathbf{c}_i))$ to maximize $\hat{J}_1(\mathbf{c}_i)$
  Update the battery and data buffer based on (2) and (4)
end for
for $i = N - P + 1$; $i \leq N$; $i = i + 1$ **do**
  Find the optimal $(e_i^*(\mathbf{c}_i), s_i^*(\mathbf{c}_i))$ to maximize $\hat{J}_{i-(N-P)+1}(\mathbf{c}_i)$
  Update the battery and data buffer based on (2) and (4)
end for

VI. Simulation Result

A. Underwater Acoustic Channel Model

The underwater acoustic channel and terrestrial radio channel have different parameters. In this section, we bravely introduce a simple underwater acoustic channel model [34], [35]. In this model, we only consider about the effect of the transmission loss and the ambient noise. The received signal-to-noise ratio (SNR) is given by

$$\text{SNR}(r, f) = \frac{P_s}{A(r, f) N_f f_B} \tag{25}$$

where P_s is the power of transmitted signal, r is the transmission distance, f is the signal frequency, $A(r, f)$ is the transmission loss and $N(f)$ is the power spectral density (PSD) of ambient noise, f_B is the bandwidth. A distinguishing property of acoustic channels is that the transmission loss for acoustic waves will increase with both propagation distance r and frequency f [34], [36]. Because of the restrict of paper length, we do not give the calculation formula of $A(r, f)$ and N_f. More details about $A(r, f)$ and N_f can be found in [34]. The average channel gain $\bar{h}$ can be approximately calculated by the transmission loss.

$$\bar{h} = \sqrt{\frac{1}{A(d, f)}} \tag{26}$$

The channel model proposed in [34] is time-invariant, and in our problem, we consider about time-variant channel. Then, we make a modification on the channel model. In underwater acoustic communications, the communication link often exhibits an ON/OFF behavior with coherence time lasting several seconds. Therefore, the channel is modeled as two state Markov Chain model (i.e., the Gilbert-Elliot model) [23], [24]: "Good (1)" and "Bad (0)", and the transition matrix is denoted by [24]

$$P_h = \begin{bmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{bmatrix} \tag{27}$$

where $P_{i,j}$ represents the transition probability from channel state $i = 0, 1$ to channel state $j = 0, 1$. The steady-state probabilities are given by $\pi_0 = P_{10}/(P_{10} + P_{01}) = 1/3$ whereas $\pi_1 = 1 - \pi_0 = 2/3$. We use the steady-state probabilities to replace the probabilities of channel state at i-th slot. And we assume that the channel gain in "Bad (0)" state is a half of the channel gain in "Good (1)" state, which means $R_h = \gamma_1/\gamma_2 = 2$. Then, we can obtain the channel gains in the "Bad (0)" and "Good (1)" states given the average channel gain $\bar{h}$ and the steady-state probabilities π_0.

In (25), P_s is the power of transmitted signal. As we know, the transmitter sensor converts electrical power into acoustic power in the water applied by underwater acoustic transducer. The total efficiency of the transmitter sensor η_e is the product of electro-acoustic efficiency and power amplifier efficiency. Then, the power of transmitted signal is given by

$$P_s = \eta_e \frac{e_i}{T_{slot}} \tag{28}$$

In this paper, the total electro-acoustic conversion efficiency is $\eta_e = 19\%$ [37].

In order to simplify the calculation, we assume that the energy harvesting rate take values from two state set $\{\delta_1, \delta_2\}$, and the transition matrix is given by

$$P_\delta = \begin{bmatrix} 0.75 & 0.25 \\ 0.5 & 0.5 \end{bmatrix} \tag{29}$$

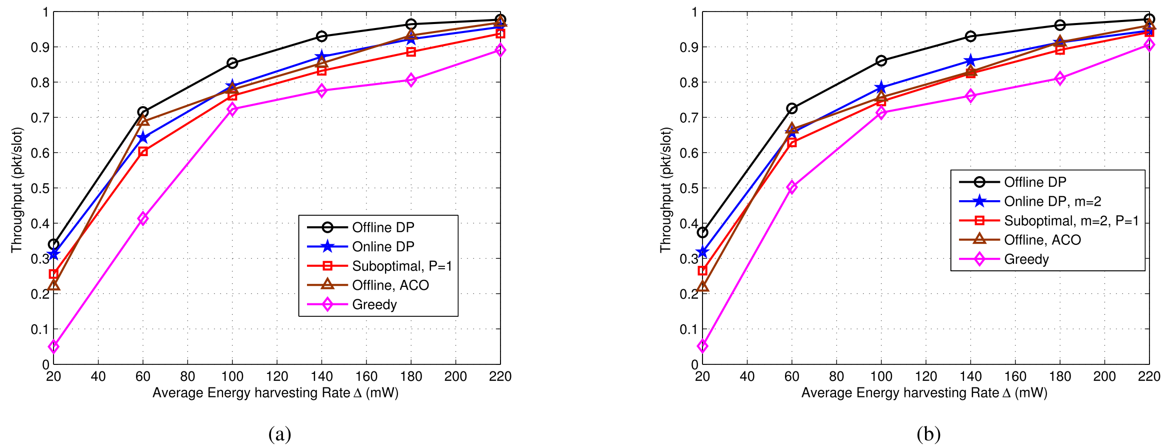


Fig. 4. The average throughput of different policies for different average energy harvesting rate. (a) Scenario 1. (b) Scenario 2.

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Carrier frequency f	10 kHz
Bandwidth f_B	5 kHz
Transmission distance r	2 km
Channel gain $[\gamma_1, \gamma_2]$	$[2.28 \times 10^{-4}, 4.56 \times 10^{-4}]$
Time slot T_{slot}	0.5 s
Packet size ℓ	2500 bit
Code rate	1/2
λ	200 pkt/J
N	20
R_δ	2
E_{max}	1 J
B_{max}	10 pkt
Δ	40 mW
E_1	0.1 J
B_1	2 pkt
ρ	0.01 J

where $P_{\delta_{ij}}$ represents the probability of the energy harvesting state going from state δ_i to state δ_j. The steady-state probabilities, respectively, are then given by $[P_{\delta_1}, P_{\delta_2}] = [2/3, 1/3]$. And the same as channel state, we calculate the value δ_1 and δ_2 according the average energy harvesting rate Δ, steady-state probabilities and the ratio $R_\Delta = \delta_2/\delta_1$.

Unless otherwise specified, the simulation parameters of system are presented in Table I. The noise power spectral density N_f and the transmission loss $A(r, f)$ are calculated as in [34].

For comparison, we consider *Greedy* policy as a baseline policy for online decision. For the first scenario which considers the sensing energy, the sensor will acquire 1 packet of payload data when the number of packets in the data buffer is below 2. Otherwise, the sensor will not expend energy to sense. The energy allocation policy is given by

$$e_i = \begin{cases} E_i, & \text{if } B_i > 1 \\ E_i - 1/\lambda, & \text{otherwise} \end{cases} \tag{30}$$

$$s_i = \begin{cases} 0, & \text{if } B_i > 1 \\ 1/\lambda, & \text{otherwise} \end{cases} \tag{31}$$

Recall that the second scenario (of delayed feedback) does not consider the sensing energy. Hence, the transmitter always uses all stored energy whenever available, i.e.,

$$e_i = E_i \tag{32}$$

For the suboptimal algorithm, we choose $P = 1$ unless otherwise specified.

In the case discussed above, the transmitter only knows the past channel state and energy harvesting rate, which is called *online* case. Here, we also test the *offline* case in which the transmitter has full knowledge about the CSI and energy harvesting. Thus, the *offline* case can be as the performance upper bound. For DP approach, it is similar to algorithm 1 and algorithm 2 but the state is $\mathbf{c}_i = (E_i, B_i)$.

In order to more accurately evaluate the performance of the proposed DP method, we also test the ant colony optimization (ACO) algorithm [38] to solve this problem in *offline* case. In the ACO algorithm, heuristic function is proportional to the packet correct probability. The number of ants equals to the number of feasible solutions. The colony evaporation rate is 0.5. The relative influence of pheromone level is 1 and the the relative influence of heuristic function is 5. The number of iterations is 200. For all methods, the number of Monte Carlo simulation is 500. For more details on the ACO, we refer our readers to references [38].

B. Effect of Average Energy Harvesting Rate

Fig. 4(a) and Fig. 4(b) show the impact of the average energy harvesting rate Δ on the throughput for scenario 1 and scenario 2, respectively. All policies exhibit improved performance as Δ increases. It is intuitive that Δ growth gives the transmitter more energy to use at each slot, hence the performance improvement in general.

From these results, one can see that the *offline DP* has the best performance (as upper bound) since it knows all the information about the energy harvesting and channel state. Hence, the performance shall improve because the transmitter has more flexibility to allocate the energy for various time slots. The *offline ACO* has the similar performance as the *online DP*. This also shows that the proposed DP method is

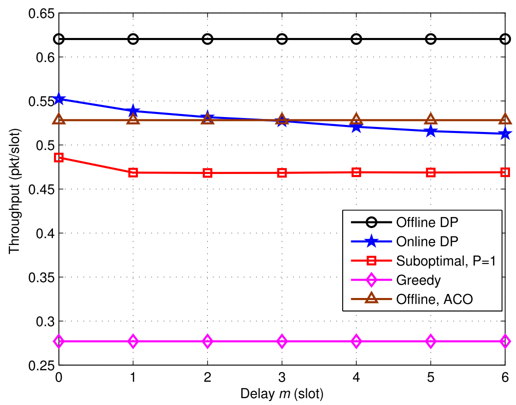


Fig. 5. The average throughput of different policies for different feedback delay for scenario 2.

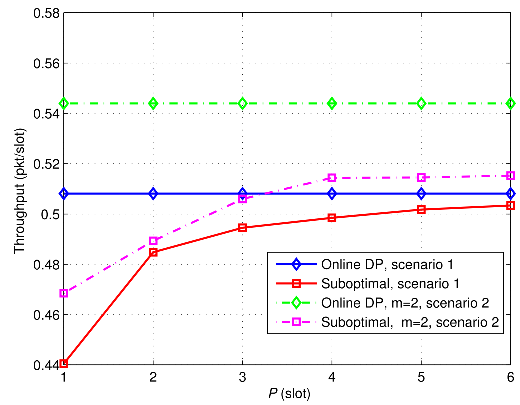


Fig. 6. The impact of P for suboptimal algorithm.

superior to the ACO method. On the other hand, the *online DP* generates the best possible online allocation policy to maximize the sum of current slot reward and the expected reward of future slots. However, *online DP* does suffer from high computational complexity. The suboptimal algorithm take a myopic view of future slot which it only involves the next P slots. Therefore, the performance of throughput degrades when compared with that of *online DP*. But the performance of the two methods is very close.

C. Impact of Feedback Delay

Fig. 5. shows the performance of different policies versus the delay m for scenario 2. It is clear that the performance of *offline* case is unchanged with the delay m since it always knows the actual channel information. The *offline DP* still has the best performance and this genie aided result provides a performance upper bound. When m increases, the feedback channel information becomes less accurate. Hence, the performance of *online DP* continually degrades. It is noted that the performance of *online DP* method still exhibits the similar performance as the *offline ACO* method. Meanwhile, the performance of suboptimal algorithm degrades when $m = 1$ increases. When m becomes greater than 1, the performance of suboptimal holds steady, which means that, for $m > 1$, the delayed feedback provides little value for the suboptimal algorithm.

D. Suboptimal Solutions and Performance-Complexity Tradeoff

In this subsection, we consider about the performance-complexity of the suboptimal algorithm. Fig. 6 show the results of the suboptimal algorithm with different choices of P. Clearly, the throughput of suboptimal algorithm improves for larger P and higher complexity. For scenario 1, when $P = 6$, the performance of suboptimal algorithm is already very close to the exact online DP approach. On the other hand, when CSI feedback has larger delays (e.g., $m = 2$), the performance gap between the DP and suboptimal algorithm becomes larger. Clearly, the suboptimal algorithm is more sensitive to the CSI feedback delay.

E. Channel Condition Effect

We further compare the performance over different channel conditions. First, we define a parameter η to represent throughput gain. For example, if $T_\alpha(\pi_0)$ and $T_\beta(\pi_0)$ is the throughput by policy α and policy β, respectively, then, $\eta = \frac{T_\alpha(\pi_0) - T_\beta(\pi_0)}{T_\beta(\pi_0)}$.

Fig.7(a) and Fig.7(b) present the performance of different policies in terms of throughput gain for different values of steady-state probability π_0 relative to *Greedy* policy. From the results, we can see that the throughput gain improves with increasing steady-state probability π_0. That means when the channel conditions are more likely to be poor, the improvement of our proposed optimization policy is more pronounced. This is quite intuitive since under more favorable channel conditions, *Greedy* policy also can already obtain a good performance.

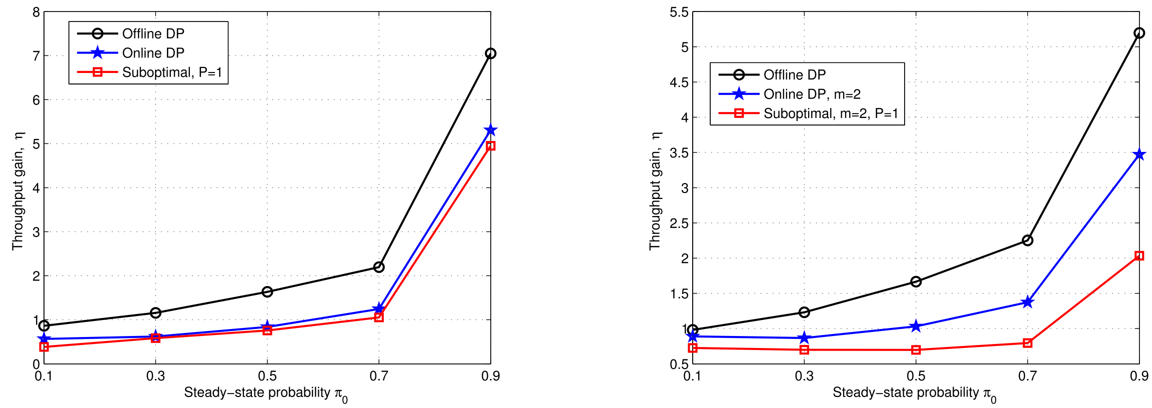
F. Sensitivity to CSI Accuracy

Thus far, we assumed the feedback state information to be accurate and error-free. In practical underwater acoustic systems, channels cannot be estimated accurate. To examine the sensitivity of our method to the accuracy of channel estimation. Fig.8(a) and Fig.8(b) present the average throughput of different policies for different levels of channel estimation error probability. Note that the *offline DP* has full knowledge of CSI whereas the *Greedy* policy does not need to estimate the channel. Hence their throughput is insensitive to the accuracy of CSI.

For scenario 1, when channel estimation error probability increases, the performance degrades correspondingly for the *online DP* and suboptimal algorithms and the *online DP* algorithm is more sensitive to error probability. When the error probability is larger than 0.9, the *online DP* is even worse than Greedy; For scenario 2, the performance of suboptimal policy at $m = 2$ is almost insensitive. That is because when $m \geq 1$, the CSI feedback no longer affects the suboptimal policy, as shown earlier from Fig. 5.

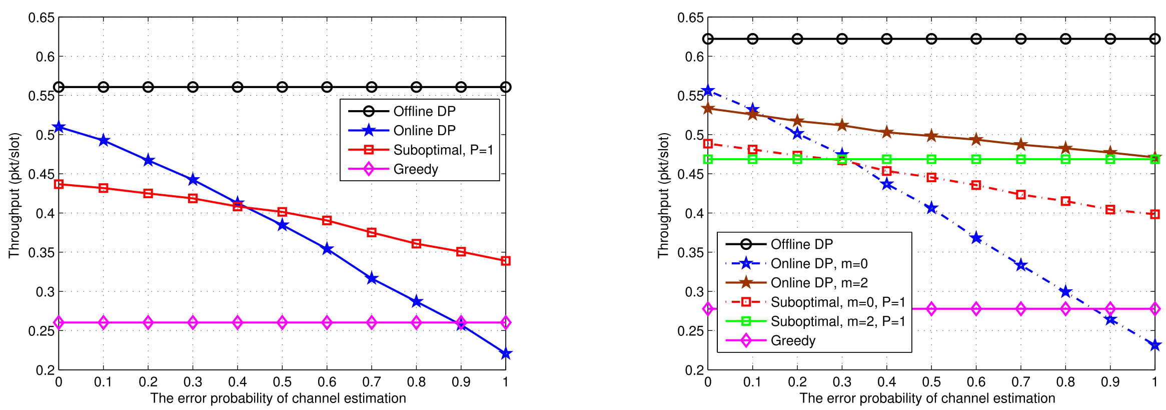
G. Buffer Backlog and Battery Capacity Effect

In this test we consider the effect of buffer backlog size. For our second scenario, we let data buffer have infinite data backlog. As a result, we do not consider sensing energy.



(a) (b)

Fig. 7. The average throughput of different policies under different channel steady-state probability. (a) Scenario 1. (b) Scenario 2.



(a) (b)

Fig. 8. The average throughput of different policies for different channel estimation error probability. (a) Scenario 1. (b) Scenario 2.

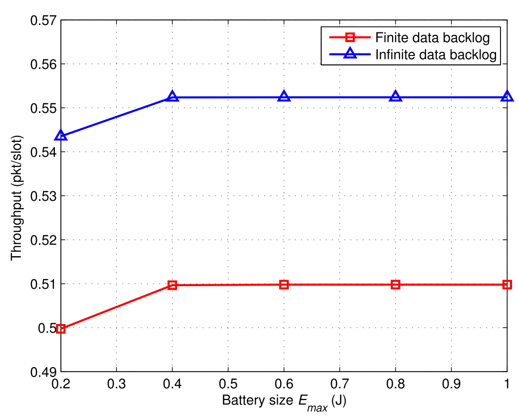


Fig. 9. The average throughput difference of finite and infinite data backlog.

Fig.9 shows that the performance difference of proposed *online DP* method due to the infinite data backlog assumption with different battery capacity E_{max}.

As E_{max} increases, the average throughout also grows. It is intuitive that the transmitter can store more harvested energy for further slots if it has a larger battery capacity. Hence, the performance shall improve because the transmitter has more flexibility to allocate the energy for various time slots. Nevertheless, the throughout growth tends to saturate when E_{max} reaches beyond a certain value. The reason is because for sufficiently large energy capacity, all harvested energy can be stored and optimally allocated. Any additional increase of capacity would no longer contribute to the sensing, transmission, or the success of packet delivery.

It is obvious that the infinite data backlog has a better throughput performance because it no longer worries about energy for sensing and has more energy for transmission. The performance gap between finite data and infinite backlog is 0.0426(pkt/slot) when $E_{max} = 1$ J.

VII. Conclusion

This work investigates the energy allocation between data acquisition and transmission for networked underwater sensors to maximize the expected data packet delivery under energy harvesting constraints. We consider two different scenarios: one involving receiver feedback delays and one without. In first cases, the transmitter must take into consideration the dual need for data sensing and data transmission. We tackle both problem scenarios by formulating a finite horizon optimization and deriving a Dynamic Programming approach. To simplify the complexity of the second scenario involving feedback delays, we assume the transmitter to have infinite data backlog to focus on the need of data transmission. We also propose a suboptimal policy based on a myopic view of future slots to further reduce the high computational complexity. Simulation results demonstrate the performance gain of the proposed policies as well as the impact of various practical factors such as battery size, average energy harvesting rate, channel steady-state distribution, and CSI accuracy.

REFERENCES

[1] C. Alippi, R. Camplani, C. Galperti, and M. Roveri, “A robust, adaptive, solar-powered WSN framework for aquatic environmental monitoring,” *IEEE Sensors J.*, vol. 11, no. 1, pp. 45–55, Jan. 2011.

[2] P. R. Bandyopadhyay, D. P. Thivierge, F. M. McNeilly, and A. Fredette, “An electronic circuit for trickle charge harvesting from littoral microbial fuel cells,” *IEEE J. Ocean. Eng.*, vol. 38, no. 1, pp. 32–42, Jan. 2013.

[3] K. B. Joshi, J. H. Costello, and S. Priya, “Estimation of solar energy harvested for autonomous jellyfish vehicles (AJVs),” *IEEE J. Ocean. Eng.*, vol. 36, no. 4, pp. 539–551, Oct. 2011.

[4] D. Toma, J. del Rio, M. Carbonell-Ventura, and J. Masalles, “Underwater energy harvesting system based on plucked-driven piezoelectrics,” in *Proc. IEEE OCEANS*, Genoa, Italy, May 2015, pp. 1–5.

[5] A. Bereketli and S. Bilgen, “Remotely powered underwater acoustic sensor networks,” *IEEE Sensors J.*, vol. 12, no. 12, pp. 3467–3472, Dec. 2012.

[6] S. He, J. Chen, F. Jiang, D. K. Y. Yau, G. Xing, and Y. Sun, “Energy provisioning in wireless rechargeable sensor networks,” *IEEE Trans. Mobile Comput.*, vol. 12, no. 10, pp. 1931–1942, Oct. 2013.

[7] A. Mehrabi and K. Kim, “Maximizing data collection throughput on a path in energy harvesting sensor networks using a mobile sink,” *IEEE Trans. Mobile Comput.*, vol. 15, no. 3, pp. 690–704, Mar. 2016.

[8] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, “Transmission with energy harvesting nodes in fading wireless channels: Optimal policies,” *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1732–1743, Sep. 2011.

[9] J. Yang and S. Ulukus, “Optimal packet scheduling in an energy harvesting communication system,” *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 220–230, Jan. 2012.

[10] C. K. Ho and R. Zhang, “Optimal energy allocation for wireless communications with energy harvesting constraints,” *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4808–4818, Sep. 2012.

[11] C. Huang, R. Zhang, and S. Cui, “Optimal power allocation for outage probability minimization in fading channels with energy harvesting constraints,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 1074–1087, Feb. 2014.

[12] S. Wei, W. Guan, and K. J. R. Liu, “Power scheduling for energy harvesting wireless communications with battery capacity constraint,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4640–4653, Aug. 2015.

[13] J. Gong, S. Zhou, and Z. Niu, “Optimal power allocation for energy harvesting and power grid coexisting wireless communication systems,” *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 3040–3049, Jul. 2013.

[14] I. Ahmed, A. Ikhlef, D. W. K. Ng, and R. Schober, “Power allocation for an energy harvesting transmitter with hybrid energy sources,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6255–6267, Dec. 2013.

[15] M. Nourian, S. Dey, and A. Ahlén, “Distortion minimization in multi-sensor estimation with energy harvesting,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 524–539, Mar. 2015.

[16] S. Knorn, S. Dey, A. Ahlén, and D. E. Quevedo, “Distortion minimization in multi-sensor estimation using energy harvesting and energy sharing,” *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2848–2863, Jun. 2015.

[17] Y. Zhao, B. Chen, and R. Zhang, “Optimal power management for remote estimation with an energy harvesting sensor,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6471–6480, Nov. 2015.

[18] K. Tutuncuoglu, A. Yener, and S. Ulukus, “Optimum policies for an energy harvesting transmitter under energy storage losses,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 467–481, Mar. 2015.

[19] B. Devillers and D. Gunduz, “A general framework for the optimization of energy harvesting communication systems with battery imperfections,” *J. Commun. Netw.*, vol. 14, no. 2, pp. 130–139, Apr. 2012.

[20] M. Stojanovic and J. Preisig, “Underwater acoustic communication channels: Propagation models and statistical characterization,” *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 84–89, Jan. 2009.

[21] D. Pompili, T. Melodia, and I. F. Akyildiz, “Distributed routing algorithms for underwater acoustic sensor networks,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2934–2944, Sep. 2010.

[22] J. Yu, H. Chen, L. Xie, and J.-H. Cui, “Performance analysis of hybrid ARQ schemes in underwater acoustic networks,” in *Proc. Oceans*, Sep. 2014, pp. 1–6.

[23] B. Tomasi and J. C. Preisig, “Energy efficient transmission scheduling for non-stationary underwater acoustic channels,” in *Proc. ACM WUWNet*, Kaohsiung, Taiwan, 2013, pp. 27:1–27:2.

[24] B. Tomasi and J. C. Preisig, “Energy-efficient transmission strategies for delay constrained traffic with limited feedback,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1369–1379, Mar. 2015.

[25] P. Castiglione, O. Simeone, E. Erkip, and T. Zemen, “Energy management policies for energy-neutral source-channel coding,” *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2668–2678, Sep. 2012.

[26] P. Castiglione and G. Matz, “Energy-neutral source-channel coding with battery and memory size constraints,” *IEEE Trans. Commun.*, vol. 62, no. 4, pp. 1373–1381, Apr. 2014.

[27] S. Mao, M. H. Cheung, and V. W. S. Wong, “Joint energy allocation for sensing and transmission in rechargeable wireless sensor networks,” *IEEE Trans. Veh. Technol.*, vol. 63, no. 6, pp. 2862–2875, Jul. 2014.

[28] O. Orhan, D. Gündüz, and E. Erkip, “Source-channel coding under energy, delay, and buffer constraints,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3836–3849, Jul. 2015.

[29] A. Yadav, M. Goonewardena, W. Ajib, and H. Elbiaze, “Novel retransmission scheme for energy harvesting transmitter and receiver,” in *Proc. IEEE ICC*, London, U.K., Jun. 2015, pp. 3198–3203.

[30] D. P. Bertsekas, *Dynamic Programming and Optimal Control (2 Vol Set)*. Belmont, MA, USA: Athena Scientific, 2000.

[31] A. Fu, E. Modiano, and J. N. Tsitsiklis, “Optimal transmission scheduling over a fading channel with energy and deadline constraints,” *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 630–641, Mar. 2006.

[32] J. Lee and N. Jindal, “Energy-efficient scheduling of delay constrained traffic over fading channels,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1866–1875, Apr. 2009.

[33] R. Srivastava and C. E. Koksal, “Energy optimal transmission scheduling in wireless sensor networks,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1550–1560, May 2010.

[34] M. Stojanovic, “On the relationship between capacity and distance in an underwater acoustic communication channel,” in *Proc. 1st ACM Int. Workshop Underwater Netw. (Wuwnet)*, Los Angeles, CA, USA, Sep. 2006, pp. 41–47.

[35] H. Wang, S. Wang, E. Zhang, and J. Zou, “A network coding based hybrid ARQ protocol for underwater acoustic sensor networks,” *Sensors*, vol. 16, no. 9, p. 1444, 2016.

[36] A. Valera, P. W. Q. Lee, H.-P. Tan, H. Liang, and W. K. G. Seah, “Implementation and evaluation of multihop ARQ for reliable communications in underwater acoustic networks,” in *Proc. Oceans*, May 2009, pp. 1–6.

[37] Z. Yang, L. Kaizhuo, Z. Qunfei, H. Dengfeng, and F. Xuchao, “An embedded signal transmitter for UAC based on H-bridge class-D power amplifier,” in *Proc. IEEE Int. Conf. Signal Process., Commun. Comput. (ICSPCC)*, Aug. 2016, pp. 1–5.

[38] M. Dorigo, V. Maniezzo, and A. Colorni, “Ant system: Optimization by a colony of cooperating agents,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 26, no. 1, pp. 29–41, Feb. 1996.

Lianyou Jing received the B. S. degree in the electronic and communications engineering from the School of Marine Science and Technology, Northwestern Polytechnical University, Xi’an, China, in 2010, where he is currently pursuing the Ph.D. degree. His current research focuses on underwater acoustic communications.

Chengbing He received the B.S., M.S., and Ph.D. degrees from the School of Marine Science and Technology, Northwestern Polytechnical University (NPU), Xi’an, China, in 2003, 2006, and 2009, respectively, all in electronic and communications engineering. He has been a Faculty Member with NPU since 2009, where he is currently an Associate Professor. His current research focuses on underwater acoustic communications.

Jianguo Huang received the B.S. degree from Northwestern Polytechnical University (NPU), Xi'an, China, in 1967. He has been a Faculty Member with NPU since 1967 and a Professor since 1987. He is a Doctoral Advisor in information and communication engineering. He is also the Former Dean of the College of Marine Engineering. He is teaching and conducting research at NPU in signal and information processing, electronic engineering, wireless communication, and underwater acoustic communication with NPU. His general research interests include modern signal processing, array signal processing, wireless, and underwater acoustic communication theory and application. He is a member of review committee for National Science and Technology Awards, a Senior Member of the IEEE Signal Processing Society, the Communication Society, and the Ocean Engineering Society, the Chair of the IEEE Xi'an Section, Fellow of Chinese Society of Acoustics, and the Administrative Director of the Signal Processing Society of Shaanxi Province.

Zhi Ding received the Ph.D. degree in electrical engineering from Cornell University, in 1990. From 1990 to 2000, he was a Faculty Member with Auburn University and later with the University of Iowa. He has held visiting positions at Australian National University, the Hong Kong University of Science and Technology, the NASA Lewis Research Center, and the USAF Wright Laboratory. He is currently a Professor of Engineering and Entrepreneurship with the University of California at Davis, Davis, CA, USA. He has active collaboration with researchers from several countries including Australia, China, Japan, Canada, Taiwan, South Korea, Singapore, and Hong Kong. He is serving on the technical programs of several workshops and conferences. He has co-authored *Modern Digital and Analog Communication Systems*, (Oxford University Press, 2009, 4th Edition). He is an IEEE Distinguished Lecturer (Circuits and Systems Society from 2004 to 2006 and the Communications Society from 2008 to 2009). He served as an IEEE Transactions on Wireless Communications Steering Committee Member from 2007 to 2009 and the Chair from 2009 to 2010. He received the 2012 IEEE Wireless Communication Recognition Award from the IEEE Communications Society. He was a member of the Technical Committee on Statistical Signal and Array Processing and a member of the Technical Committee on Signal Processing for Communications from 1994 to 2003. He was the Technical Program Chair of the 2006 IEEE Globecom. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1994 to 1997 and from 2001 to 2004, and an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS from 2002 to 2005.