Power Control of an Energy Harvesting Sensor for Remote State Estimation

Yuzhe Li, Fan Zhang, Daniel E. Quevedo, Vincent Lau, Subhrakanti Dey, and Ling Shi

Abstract—**We investigate sensor transmission power control for remote state estimation. Instead of using a conventional sensor, a sensor equipped with an energy harvester which can obtain energy from the external environment is utilized. We formulate power control of the energy harvesting sensor into an infinite time-horizon Markov decision process (MDP). To deal with the computation complexity associated with this multi-dimensional MDP, a continuous-time approach and perturbation analysis are used and a closed-form approximate value function is derived. Based on the approximation, we obtain a closed-form optimal power control solution which has a threshold-based structure. A numerical example is provided to evaluate the estimation performance of the optimal solution compared with other power scheduling schemes.**

Index Terms—**Energy harvesting, Markov decision process (MDP), power control, remote state estimation.**

I. INTRODUCTION

THE past decade has witnessed the rapid growth of wireless sensor networks (WSN) in fundamental research and practical applications. Due to advantages such as low cost, ease of installation, and self-power, wireless sensors play an increasingly important role in many applications, compared with traditional wired sensors. In a WSN, the sensors are typically equipped with batteries and expected to work for a long time [1]. In some applications, the number of sensors can be quite large (e.g., environment monitoring) or sensors may be located in dangerous environments [2] (e.g., chemical industry), which makes the replacement of batteries difficult or even impossible. Thus, finding a good power control schedule for the sensors is imperative.

To deal with energy constraints of wireless sensors, two possible solution categories are often adopted. One way is

to focus on the outlet of the energy, i.e., to develop more efficient sensor power control methods to improve the use of the batteries [3]–[6]. Existing results have demonstrated significant improvement of sensor lifetime and system performance under energy constraints using appropriate power control policies. The problem is, however, still not completely solved as the battery will eventually run out. At the same time, the optimization of sensor lifetime where energy is limited always leads to sacrifices such as estimation quality or system stability [7].

To overcome this limitation, an alternative way is to put emphasis on the inlet of the energy, i.e., using sensors equipped with an energy harvester to replace conventional battery-powered sensors. The energy harvester can obtain energy from the external environment (e.g., body heat, solar energy, piezoelectric energy, wind energy) and convert them into electrical energy which can be stored and used by the sensor [2]. In contrast to battery-powered devices, for sensors using this technology, the energy (but not the energy-rate) is typically unlimited as the harvester can generate power most of the time. But unlike the battery-powered sensor, which has known energy levels for future use, a sensor with energy harvester will be subject to uncertain future energy levels, as it is affected by the randomness of the external environment. For example, as stated in [8], both technical and theoretical challenges are posed by the inherent uncertainty and variability in wind energy when integrated into the electricity power grid. Due to the randomness of the amount of future harvested energy, power control and battery management require trading current transmission success probabilities for safeguarding against possible future energy shortage. Therefore, new challenges arise in the design and analysis of the communication strategy of the energy harvesting sensor which are summarized as follows.

1) **Randomness of Energy Arrivals**: For energy harvesting sensors, the information of the future energy constraints is not exactly available for the sensor before the harvesting. Therefore, a new approach is needed to handle the constraints related to the energy arrival randomness.
2) **Practical Energy Model**: In practice, the energy storage amount of the sensor is limited, which will introduce a more complex energy constraint in the stochastic optimization problem. Furthermore, non-i.i.d. and continuous energy arrivals (e.g., Markovian arrivals of solar energy described in [9]) are more difficult to deal with.
3) **Computation Complexity of High-dimensional MDP**: It is difficult to overcome the computation complexity and implicit solution form of the MDP problem in remote estimation with energy harvesting sensor.

To tackle with these difficulties, in the present work, we consider remote estimation with a wireless sensor having energy harvesting capability. Our present manuscript extends existing results to a more general model (non-i.i.d. energy arrival) with closed-form solution based on the continuous-time approximation method and the perturbation analysis to reduce the computation complexity. We derive an approximate optimal solution, which has a closed-form expression. We will show that, under the MDP framework, we can improve the performance based on the continuous-time approximation method in [10] and [11] and the perturbation analysis in [12]. The main contributions of this work are summarized as follows.

1) **Standard MDP Framework**: To handle the challenge about the randomness of energy arrivals, we prove that an associated power control design problem can be formulated into a standard MDP framework with infinite time-horizon and we provide the optimal solution (Theorem 3.3).

2) **Finite Energy Storage Amount Limitation and Non-i.i.d. Continuous Energy Arrivals**: To model the battery more practically, we consider limited energy storage amount of the sensor in this work. Furthermore, departing from [13] which considered discrete and i.i.d. energy arrivals, our model and solution also apply to non-i.i.d. and continuous energy arrivals (e.g., Markovian arrivals of solar energy as described in [9]).

3) **Closed-form Approximate Optimal Solution**: To deal with the computation complexity of the high-dimensional MDP problem, we obtain an approximate optimal solution under the MDP framework based on the continuous-time approximation method in [10] and [11] and the perturbation analysis in [12], which reduces the computation complexity significantly and has a closed-form expression (Theorem 4.4). The performance of the approximate optimal solution is also improved compared with sub-optimal solutions.

The remainder of the paper is organized as follows. Section II presents the system model and states the main problem of interest. Section III presents the optimal power schedule under the MDP framework. Section IV provides the approximate optimal solution based on the continuous-time approximation method and the perturbation analysis. Simulations and numerical examples are given in Section V. Section VI draws conclusions.

Notations: $\mathbb{Z}$ denotes the set of integers and $\mathbb{N}$ the positive integers. $\mathbb{R}$ is the set of real numbers. $\mathbb{R}^n$ is the n-dimensional Euclidean space. $\mathbb{S}^n_+$ (and $\mathbb{S}^n_{++}$) is the set of n by n positive semi-definite matrices (and positive definite matrices). When $X \in \mathbb{S}^n_+$ (and $\mathbb{S}^n_{++}$), we write $X \geqslant 0$ (and $X > 0$). $X \geqslant Y$ if $X - Y \in \mathbb{S}^n_+$. $\mathrm{Tr}(\cdot)$ is the trace of a matrix. The superscript $'$ stands for transposition. For functions f, f_1, f_2 with appropriate domains, $f_1 \circ f_2(x)$ stands for the function composition $f_1(f_2(x))$, and $f^n(x) \triangleq f(f^{n-1}(x))$, where $n \in \mathbb{N}$ and with $f^0(x) \triangleq x$. δ_{ij} is the Dirac delta function, i.e., δ_{ij} equals to 1 when $i = j$ and 0 otherwise. The notation $\mathbb{P}[\cdot]$ refers to probability and $\mathbb{E}[\cdot]$ to expectation.
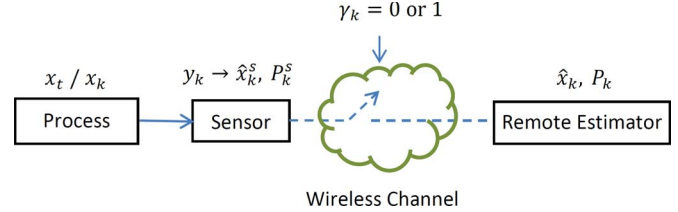

Fig. 1. System architecture.

II. STATE ESTIMATION WITH AN ENERGY HARVESTER

We consider the problem of remotely estimating the state of the following continuous-time linear time-invariant (LTI) system:

$$\dot{x}_t = \widetilde{A}x_t + \widetilde{w}_t \qquad (1)$$

where $x_t \in \mathbb{R}^{n_x}$ is the system state vector at time t, $\widetilde{w}_t \in \mathbb{R}^{n_x}$ is zero-mean i.i.d. Gaussian with $\mathbb{E}[\widetilde{w}_{t_1} \widetilde{w}'_{t_2}] = \delta_{t_1 t_2} \widetilde{Q}$ $(\widetilde{Q} \geqslant 0)$. The initial state x_0 is a zero-mean Gaussian random vector with covariance $\Pi_0 \geqslant 0$ and is uncorrelated with $\widetilde{w}_t$.

A sensor measures the state x_t discretely with sampling time τ [14]. The state and measurement dynamics of the sampled discrete-time LTI system can be expressed as follows [15]:

$$x_{k+1} = Ax_k + w_k \qquad (2)$$

$$y_k = Cx_k + v_k \qquad (3)$$

where $A = \exp(\widetilde{A}\tau)$, $w_k = \int_0^\tau \exp(\widetilde{A}s)\widetilde{w}[(k+1)\tau - s]ds$, $y_k \in \mathbb{R}^{n_y}$ is the measurement taken by the sensor, $v_k \in \mathbb{R}^{n_y}$ is zero-mean i.i.d. Gaussian with $\mathbb{E}[v_k(v_j)'] \triangleq \delta_{kj}R$ $(R > 0)$, $\mathbb{E}[w_k(v_j)'] = 0 \ \forall j, k \in \mathbb{N}$. The pair (A, C) is assumed to be observable and $(A, Q^{1/2})$ to be controllable.

Remark 2.1: Note that as $\widetilde{w}_t$ is zero-mean i.i.d. Gaussian, it is straightforward that $w_k = \int_0^\tau \exp(\widetilde{A}s)\widetilde{w}[(k+1)\tau - s]ds$ is also zero-mean i.i.d. Gaussian with $\mathbb{E}[w_k(w_j)'] = \delta_{kj}Q$, where $Q = \int_0^\tau \exp(\widetilde{A}s)\widetilde{Q}\exp(\widetilde{A}s)ds$. For system (1) and (2), the Kalman filter is the minimum mean-square error (MMSE) estimator for the state x_k. □

A. Sensor Local State Estimate

As depicted in Fig. 1, the sensor is embedded with an on-board processor, the so called "smart sensor" [16]. At each time k, the sensor first locally runs a regular Kalman filter to produce the MMSE estimate of the state x_k based on all the measurements it collects up to time k and then transmits the local estimate to the remote estimator.

Denote $\hat{x}^s_k$ and P^s_k as the sensor's local MMSE state estimate and the corresponding estimation error covariance, respectively, i.e.,

$$\hat{x}^s_k = \mathbb{E}[x_k | y_1, y_2, \ldots, y_k] \qquad (4)$$

$$P^s_k = \mathbb{E}\left[(x_k - \hat{x}^s_k)(x_k - \hat{x}^s_k)' | y_1, y_2, \ldots, y_k\right]. \qquad (5)$$

Thus the above quantities can be calculated recursively using standard Kalman filter update equations ([17]), where the recursion starts from $\hat{x}^s_0 = 0$ and $P^s_0 = \Pi_0 \geqslant 0$.

To facilitate our subsequent discussions, define the Lyapunov and Riccati operators $h, \tilde{g} : \mathbb{S}^n_+ \to \mathbb{S}^n_+$

$$h(X) \triangleq AXA' + Q \tag{6}$$

$$\tilde{g}(X) \triangleq X - XC'[CXC' + R]^{-1}CX. \tag{7}$$

Since the estimation error covariance of the sensor, P^s_k, converges to a steady-state value exponentially fast (see [17]), without loss of generality, we assume that the Kalman filter at the sensor side has entered the steady state, i.e.,

$$P^s_k = \overline{P}, \ k \geqslant 1 \tag{8}$$

where $\overline{P}$ is the steady-state error covariance given by the unique positive semi-definite solution of $\tilde{g} \circ h(X) = X$. In the steady-state of Kalman filter, the Kalman gain will become a constant K which can be computed offline and programmed into the sensor. The sensor only needs to compute the state estimate based on: $\hat{x}^s_k = A\hat{x}^s_{k-1} + K(y_k - CA\hat{x}^s_{k-1})$. In addition, the sensor does not need to compute the state estimation error covariance P_k. Hence the required computational effort is not overly burdensome for a state-of-the-art sensor node.

B. Wireless Communication Model

After obtaining the local state estimate $\hat{x}^s_k$, the sensor transmits it to the remote estimator over an Additive White Gaussian Noise (AWGN) channel using Quadrature Amplitude Modulation (QAM).[1] Denote ω_k as the transmission power for sending the QAM symbol at time slot k with duration τ.

Based on the analysis in [18], the approximate relationship between the symbol error rate (SER) and ω_k is given by

$$\text{SER} \approx \exp\left(-\beta\frac{\omega_k\tau}{N_0W}\right) \tag{9}$$

where β is a constant which depends on R, N_0 is the AWGN noise power spectral density, and W is the channel bandwidth [14]. The communication channel is assumed to be time-invariant, i.e., β, N_0, W, are constants during the entire communication session.[2] In practice, the remote estimator can detect symbol errors, e.g., via cyclic redundancy check (CRC). Thus taking account of the SER in the transmission of QAM symbols, a binary random process $\{\gamma_k\}$, $k \in \mathbb{N}$ (which indeed follows a Bernoulli distribution) can be used to characterize the equivalent communication channel for $\hat{x}^s_k$ between the sensor and the remote estimator, where:

$$\gamma_k = \begin{cases} 1, & \text{if } \hat{x}^s_k \text{ arrives error-free at time } k \\ 0, & \text{otherwise (regarded as dropout).} \end{cases} \tag{10}$$
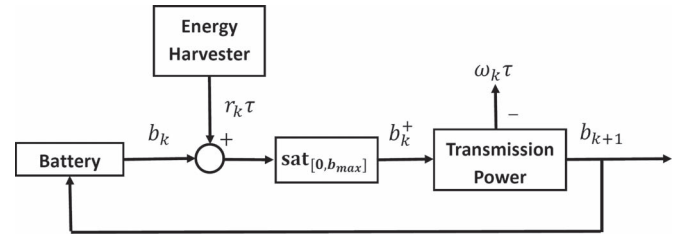

Fig. 2. Battery dynamics with harvested energy $r_k\tau$ and transmission power $\omega_k\tau$.

From (9), we have

$$\mathbb{P}[\gamma_k = 0|\omega_k] = \exp\left(-\beta\frac{\omega_k\tau}{N_0W}\right) \triangleq \lambda^{\omega_k\tau} \tag{11}$$

where $\lambda = \exp(-(\beta/N_0W)) \in (0, 1)$.

We assume that the remote estimator will send reliable ACKs to the sensor to indicate whether it has received the data packet successfully or not at time k.

C. Energy Harvester

The technology of energy harvesting enables the sensor to obtain energy from the external environment or other types of energy sources (e.g., body heat, solar energy, piezoelectric energy, wind energy) and converting them into electrical energy which can be stored.

We shall denote the energy harvesting rate (input power of the energy harvester) at the beginning of time slot k as r_k (Watt), hence, the amount of harvested energy within the time slot is $r_k\tau$ (Watt $\cdot$ s, i.e., Joule). We assume r_k is a random variable which follows a general distribution $\mathbb{P}[r_k]$ depending on the external environment (we will show later that we do not have a constraint on the exact distribution of r_k, thus the obtained results are quite general). In addition, the harvesting process is assumed to be independent of the physical process (1) and the measurements.

Denote the remaining energy level in the sensor's battery at the beginning of time step k as b_k. We further consider a finite energy storage constraint and denote the maximal energy amount that can be stored under the sampling time τ as $b_{\max} = \omega_{\max}\tau$ where $\omega_{\max}$ is the maximum transmission power per slot. Without loss of generality, we assume that the initial battery level b_1 is 0.

At the end of each time slot, after harvesting the energy $r_k\tau$, the battery level becomes

$$b^+_k \triangleq \min\{b_k + r_k\tau, b_{\max}\}. \tag{12}$$

Then the sensor needs to decide the transmission power ω_k used at time k to send the local state estimates to the remote estimator. After this procedure, the process moves to next time slot $k + 1$ and the battery level at the beginning of $k + 1$ is (as depicted in Fig. 2)

$$b_{k+1} = b^+_k - \omega_k\tau. \tag{13}$$

The energy consumption at each time slot cannot exceed the current available energy in the energy storage. To be more specific, the power consumption $\omega_k\tau$ should satisfy the following energy availability constraint $0 \leqslant \omega_k\tau \leqslant \min\{b_k + r_k\tau, b_{\max}\}$.

---

[1]Due to its high bandwidth efficiency, QAM is a common modulation scheme widely used in IEEE 802.11g/n as well as 3G and LTE systems. Specifically, $\hat{x}^s_k$ is quantized into R bits and mapped to one of the 2^R available QAM symbols.

[2]For time-varying channels, one can also formulate the problem in a similar way. This is left for future work.

For mathematical convenience, we consider a *normalized energy dynamics* (normalized per time slot)

$$\widetilde{b_{k+1}} = \widetilde{b_k^+} - \omega_k \tag{14}$$

where $\widetilde{b_{k+1}} = b_{k+1}/\tau$, $\widetilde{b_k^+} = b_k^+/\tau = \min\{\widetilde{b_{k+1}} + r_k, \omega_{\max}\}$. As such, the instantaneous power constraint becomes $0 \leqslant \omega_k \leqslant \min\{\widetilde{b_k} + r_k, \omega_{\max}\}$, where the range of the transmission power ω_k is determined by $\widetilde{b_k^+}$ and independent of τ.

As mentioned before, different power levels lead to different dropout rates, and thereby affect the estimation performance. While keeping the battery partly charged serves to "prepare for the future," one should also avoid wasting harvested energy due to battery capacity limitations. How to deal with the trade-off between the reservation for future and limited battery capacity is critical when designing power control protocols. This motivates the issue of energy management to be studied in Section III.

D. Remote State Estimation

Denote $\hat{x}_k$ and P_k as the remote estimator's own MMSE state estimate and the corresponding error covariance based on all the sensor data packets received up to time step k. The works [18] and [19] show that they can be calculated via the following procedure: once the sensor's local estimate arrives, the estimator synchronizes $\hat{x}_k$ with that of the sensor, i.e., $\hat{x}_k^s$; otherwise, the remote estimator just predicts $\hat{x}_k$ based on its previous estimate using the system model (2). From (10), the remote state estimate $\hat{x}_k$ thus obeys the recursion

$$\hat{x}_k = \begin{cases} \hat{x}_k^s, & \text{if } \gamma_k = 1 \\ A\hat{x}_{k-1}, & \text{if } \gamma_k = 0. \end{cases} \tag{15}$$

The corresponding state estimation error covariance P_k satisfies

$$P_k = \begin{cases} \overline{P}, & \text{if } \gamma_k = 1 \\ h(P_{k-1}), & \text{if } \gamma_k = 0. \end{cases} \tag{16}$$

III. Optimal Transmission Power Schedule

In this section, we define the transmission power control policy and formulate a sensor transmission power control problem. We will further give the optimality conditions for solving the associated stochastic control problem.

A. MDP Formulation

As in [20], we shall investigate energy management by formulating Problem 3.1 as an MDP problem, which requires the following definitions.

- *State*

  The ACKs from the remote estimator will enable the sensor to obtain P_{k-1}. Accordingly we define the state for the power management problem at the beginning of time step k as $\Phi_k = (b_k^+, P_{k-1})$, which consists of the battery

level at the beginning of time step k [see (12)] and the state estimate error covariance of the previous time step, P_{k-1}. Note that here we choose P_{k-1} because P_k is still unknown at the beginning of time step k. The initial state is denoted as Φ_0.

  From the recursion of P_k in (16), it is easy to see that at any time step $k_2 \geqslant k_1$, P_{k_2} can be written as $P_{k_2} = h^{k_2-k_1}(\overline{P})$, where k_1 is the latest time when the remote estimator successfully received sensor data. Since P_k only takes values in the infinitely countable set $\{\overline{P}, h(\overline{P}), h^2(\overline{P}), \ldots\}$, the state space $\mathbb{S}$ for Φ_k can be expressed as $\mathbb{S} = \{(b_k^+, P_{k-1})\}$, where $b_k^+ \in [0, b_{\max}]$, $P_{k-1} \in \{\overline{P}, h(\overline{P}), \ldots\}$.

- *Action*

  Let $\mathcal{F}(k) = \sigma(\Phi_i : i = 0, 1, 2 \ldots, k)$ be the minimum σ-algebra containing the set $\Phi_i : i = 0, 1, 2, \ldots, k$. At the beginning of the k-th slot, the sensor determines the power control action ω_k, which is $\mathcal{F}(k)$-adapted at time slot k. The actions set $\mathbb{A}$ can be expressed as $\mathbb{A} = [0, \widetilde{b_k^+}]$. Note that the range of the transmission power ω_k is determined by $\widetilde{b_k^+}$: the normalized energy level in the battery defined in (14) that is independent of τ.

- *Transition Probability*

  Given the control policy to be obtained, the random process $\{\Phi_k\}$ becomes a controlled Markov chain with a product space $\mathbb{S}$ of a continuous real value set and an infinitely countable set. As the recursion of the state estimation error covariance in (16) does not depend on the time index k, Φ is a time-homogeneous process. In such case, the controlled Markov chain is governed by the transition kernel $T(S, x, a)$ for $S \subset \mathbb{S}$, $x \in \mathbb{S}$ and $a \in \mathbb{A}$. Denote $\sigma(\mathbb{S})$ as a Borel σ-algebra on $\mathbb{S}$, then the transition kernel ([21]) $T : \sigma(\mathbb{S}) \times \mathbb{S} \times \mathbb{A} \rightarrow [0, 1]$ defines $\{\Phi_k\}$ as $\mathbb{P}(\Phi_{k+1} \in S | \Phi_k, \omega_k) = T(S, \Phi_k, \omega_k)$.

  We shall denote $T(y, x, a)$ as the density of the transition kernel $T(S, x, a) = \int_S T(y, x, a)dy$. The derivation of the closed-form expression for the density of the transition kernel is provided in Appendix A.

B. Problem of Interest

The objective of the remote estimation is to obtain an accurate state estimate $\hat{x}_k$. To be more specific, the sensor seeks to minimize the trace of the average expected state estimation error covariance at the remote estimator

$$J(\theta) = \limsup_{T \to \infty} \frac{1}{T} \sum_{k=1}^{T} \text{Tr}\left\{\mathbb{E}[P_k]\right\} \tag{17}$$

where $\theta = \{\omega_1, \omega_2, \ldots\}$ is the transmission power used at each time step and $\text{Tr}\{\mathbb{E}[P_k]\}$ is the per-stage cost. Note that, here we consider an infinite time-horizon which is a good approach for long-term running applications.

Due to the energy and battery constraints, we are interested in finding the optimal transmission power policy $\theta^\star$ for the sensor that solves the following constrained optimization problem:

Problem 3.1:

$$\min_{\theta} \quad J(\theta)$$

$$\text{s.t.} \quad 0 \leqslant \omega_k \leqslant \min\{\widetilde{b}_k + r_k, \omega_{\max}\}, \quad \forall k \in \mathbb{N}^+$$

where $\theta = \{\omega_1, \omega_2, \ldots\}$. □

As described in (17), the cost function (objective function) is the trace of average expected state estimate error covariance. Denote $v_k(\Phi_k, \omega_k)$ as the single stage cost function for time step k, $\text{Tr}\{\mathbb{E}[P_k]\}$. This means that as a result of choosing action ω_k when the remote estimator is in state Φ_k at time step k, the remote estimator receive a cost $v_k(\Phi_k, \omega_k)$.

Suppose that $v_k(\Phi_k, \omega_k | \Phi_{k+1})$ is the cost given Φ_{k+1}, then $v_k(\Phi_k, \omega_k)$ can be expressed as the expected value of $v_k(\Phi_k, \omega_k, \Phi_{k+1})$, which depends on the state of the remote estimator at time step k and at the next time $k+1$

$$v_k(\Phi_k, \omega_k) = \int_{\Phi_{k+1} \in \mathbb{S}} T(\Phi_{k+1}, \Phi_k, \omega_k) v_k(\Phi_k, \omega_k | \Phi_{k+1})$$

$$= \lambda^{\omega_k \tau} \text{Tr}\{h(P_{k-1})\} + [1 - \lambda^{\omega_k \tau}] \text{Tr}\{\overline{P}\}. \quad (18)$$

Without loss of generality, we assume that the costs can be calculated by the sensor prior to selecting a particular action. Define Θ as the policy for the sensor, which is a map from $\mathbb{S}$ to $\mathbb{A}$ such that the transmission power is given by $\omega_k = \Theta(\Phi_k)$.

Also denote the expected total cost under a policy Θ up to time-horizon T when the initial state of the system is Φ_0 as $V_T^{\Theta}(\Phi_0) \triangleq \mathbb{E}_{\Phi_0}^{\Theta}[\sum_{k=1}^{T} v_k(\Phi_k, \omega_k)]$.

The sensor needs to design the transmission power strategy Θ such that the estimation error is bounded. To be more specific, we have the following definition of admissible transmission power strategies.

Definition 3.2: A transmission power strategy for the sensor Θ is admissible if the estimation process is stable in the sense that the expected state estimation error covariance is bounded for every possible initial state, i.e., $\lim_{k \to \infty} \mathbb{E}_{\Phi_0}^{\Theta}[\text{Tr}\{P_k\}] < \infty$, $\forall \Phi_0$. □

Based on the results in [9], if the packet loss probability is relatively high compared to a relatively low energy harvesting rate, the admissible allocation strategy which stabilizes the estimation error in the asymptotic limit may not exist. In our work, we only consider the case when there exists an admissible policy as per Definition 3.2. More results regarding to the stability of the power allocation can be found in [9] and [22].

In the following discussion, we will aim to derive the corresponding policy assuming its existence. Once we obtain the policy, we will verify its admissibility of the such a policy, see Theorem 4.7.

Given the above, the performance metric is chosen as the average cost of a policy Θ given the initial value Φ_0, namely

$$J^{\Theta}(\Phi_0) \triangleq \lim_{T \to \infty} \frac{1}{T} V_T^{\Theta}(\Phi_0). \quad (19)$$

Note that (17) is the cost under a specific transmission power realization while (19) is the cost under a policy, which specifies the exact power used for different process state.

Therefore Problem 3.1 can be stated as finding the optimal admissible policy $\Theta^{\star}$ to minimize (19), i.e.,

$$J^{\star}(\Phi_0) = \min_{\Theta} J^{\Theta}(\Phi_0) \quad (20)$$

and $\Theta^{\star} = \arg\min_{\Theta} J^{\Theta}(\Phi_0)$.

C. Optimality Equation

Based on the theory of MDP [21], [23]–[25], the optimal policy $\Theta^{\star}$ is stationary and independent of the initial value (see [24, Prop. 7.4.1]). Thus the optimality conditions of (20) are given by the Bellman equation below. Note that to facilitate the discussion and subsequent calculations, without loss of generality, we slightly change the expressions of $J(\theta)$ and $v_k(\Phi_k, \omega_k)$ by multiplying them with τ (which will not change the solution of the following optimization problem).

Theorem 3.3: If there exists a pair $(J^{\star}, H^{\star}(\Phi_k))$ satisfying the following optimality equation (Bellman equation):

$$J^{\star}\tau + H^{\star}(\Phi_k) = \min_{\omega_k \in \mathbb{A}} \{v_k(\Phi_k, \omega_k)\tau + \mathbb{E}[H^{\star}(\Phi_{k+1}) | \Phi_k, \omega_k]\}$$

$$(21)$$

and for all admissible strategies Θ, $H^{\star}(\Phi_k)$ satisfies the transversality condition

$$\lim_{k \to \infty} \frac{1}{k} \mathbb{E}^{\Theta}[H^{\star}(\Phi_k) | \Phi_0] = 0 \quad (22)$$

then the value of this infinite-time horizon minimization problem is given by $J^{\star}$ and the power control policy $\omega_k^{\star} = \Theta^{\star}(\Phi_k)$, which achieves the minimum of the RHS in (21) for a given state Φ_k, is the optimal solution. Furthermore, $H^{\star}(\Phi_k)$ is the relative value function. □

The existence of a solution and independence with respect to initial conditions can be guaranteed under conditions provided in [9]. Note that (21) is not easy to solve (see also [9], [13]). As the state set is countably infinite (P_k takes values from $h^i(\overline{P})$, $i = 0, 1, 2, \ldots$) and the action set is continuous, one may solve it by value iteration after a fine discretization of the state and action space. This results in an exponential complexity that is beyond the capability of a state-of-the-art sensor node. The value function does not have a closed-form solution, which lacks valuable design insight. In addition, though the solution can be solved in principle (see [24]), if the system parameters keep changing during run-time, the sensor may not have the ability to compute the optimal policy online. This motivates us to consider how to obtain an approximate solution to this MDP problem, which can be expressed in a closed-form.

IV. APPROXIMATE OPTIMAL SOLUTION WITH CLOSED-FORM EXPRESSION

To overcome the computation complexity of the MDP problem stated in Section III, in this section, we provide a closed-form approximate solution to the optimality equation in (21).

A. Continuous-Time Approximation

Using the continuous-time approximation method in [10] and [11] and the perturbation analysis in [12], we can obtain a closed-form approximation of the original optimality equation (21) and the performance gap between the approximate solution and the optimal solution. We first calculate the approximation of right-hand side of (21).

Remark 4.1: For convenience, in the following discussion, we only investigate the case that r_k is i.i.d. distributed. We will show later that when r_k is non-i.i.d. distributed (e.g., Markovian), the main results and derivations will not change except that the state needs to be suitably augmented to include r_k. □

Theorem 4.2: Suppose that there exists a pair (J, H) satisfying:

1) the two-variable partial differential equation (PDE)

$$
\begin{aligned}
J = \min_{\omega_k} \Big\{ &-\frac{\beta}{N_0 W} \left[H\left(b_k^+, P_{k-1}\right) - H\left(b_k^+, \overline{P}\right) \right] \\
&- \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right) \Big\} \omega_k \\
&+ \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right) \mathbb{E}[r_{k+1}] \\
&+ \mathrm{Tr}\Big\{ \nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)\left(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q}\right) \Big\} \\
&+ \mathrm{Tr}\{P_{k-1}\}
\end{aligned}
\tag{23}
$$

where $\nabla_{b_k^+}$ and $\nabla_{P_{k-1}}$ are the partial derivatives with respect to the first and the second variable of $H(b_k^+, P_{k-1})$. The partial derivative of a scalar function with respect to a matrix $(\nabla_{P_{k-1}})$ is defined in the Appendix, see (28);

2) the value function $H(b_k^+, P_{k-1})$ is piecewise affine with respect to P_{k-1} and the coefficients of its partial derivatives of different orders with respect to b_k^+ are uniformly bounded.

Then, for any state Φ_k, we have $H^\star(\Phi_k) - H(\Phi_k) = \mathcal{O}(\tau)$.

Proof: See Appendix B. ∎

Theorem 4.2 implies that when the above two conditions are satisfied, $H(\Phi_k)$ can serve as an approximation to the optimal relative value function $H^\star(\Phi_k)$ with approximation error $\mathcal{O}(\tau)$, i.e., when $\tau \to 0$, $H^\star(\Phi_k) \to H(\Phi_k)$.

Furthermore, denote $\Theta^\star$ and $\bar{\Theta}^\star$ as the optimal policies for the Bellman (21) and the approximate Bellman (23), respectively. The corresponding objective values under each policy are represented by $J^\star$ and $\bar{J}^\star$. We have the following result about the performance gap.

Theorem 4.3: Given the optimal admissible policy $\bar{\Theta}^\star$ for the approximate Bellman (23), the performance gap between $J^\star$ and $\bar{J}^\star$ is given by $J^\star - \bar{J}^\star = \mathcal{O}(\tau)$.

Proof: See Appendix C. ∎

Therefore, based on Theorem 4.2 and 4.3, we can solve the PDE equation in (23) instead of solving the optimality equation in (21) to obtain an approximate relative function and optimal policy with performance gap bounded by $\mathcal{O}(\tau)$. This reduces the computation complexity significantly.

In the following subsection, we aim to solve the PDE in (23) to obtain the closed-form solution based on well-established PDE theories.

B. Closed-Form Approximate Value Function

In this subsection, we solve the multi-dimensional PDE (23) posed in Theorem 4.2 and obtain the approximate asymptotic relative value function $H(\Phi_k)$ and the approximate optimal power schedule which are given by the following theorem.

Theorem 4.4 [Asymptotic Solution of the PDE in (23)]: The approximate asymptotic solution of the relative value function to the PDE in (23) and satisfying the conditions in Theorem 4.2 is given by

$$
\begin{aligned}
&H\left(b_k^+, P_{k-1}\right) \\
&= \begin{cases}
H_1\left(b_k^+, P_{k-1}\right), \\
\quad \text{if } \Xi_1\left(b_k^+, P_{k-1}\right) \leqslant 0 \,\&\&\, \Xi_2\left(b_k^+, P_{k-1}\right) \leqslant 0 \\
H_2\left(b_k^+, P_{k-1}\right), \\
\quad \text{if } \Xi_1\left(b_k^+, P_{k-1}\right) > 0 \,\&\&\, \Xi_2\left(b_k^+, P_{k-1}\right) > 0 \\
\frac{1}{2}\left[H_1\left(b_k^+, P_{k-1}\right) + H_2\left(b_k^+, P_{k-1}\right)\right], \\
\quad \text{otherwise}
\end{cases}
\end{aligned}
$$

where

$$
H_i\left(b_k^+, P_{k-1}\right) = \mathrm{Tr}\left\{F_i\left(b_k^+\right) P_{k-1}\right\} + G_i(b_k^+), \quad i = 1, 2
$$

$F_1(b_k^+), F_2(b_k^+) \in \mathbb{R}^{n \times n}$ and $G_1(b_k^+), G_2(b_k^+) \in \mathbb{R}$ are continuous functions of b_k^+ given by (48), (49), (51) and (52), respectively, in Appendix D, and $\Xi_1(b_k^+, P_{k-1})$ and $\Xi_2(b_k^+, P_{k-1})$ are operators regarding to $H_1(b_k^+, P_{k-1})$ and $H_2(b_k^+, P_{k-1})$ defined in (53) in Appendix D.

The approximate optimal transmission power schedule $\bar{\Theta}^\star$ associated to H has an event-based threshold structure of the following form:

$$
\omega_k^\star = \begin{cases}
\max\{\omega_k \in \mathbb{A}_k\} = \widetilde{b_k^+}, & \text{if } \Xi\left(b_k^+, P_{k-1}\right) \leqslant 0 \\
0, & \text{otherwise.}
\end{cases}
\tag{24}
$$

where $\Xi(b_k^+, P_{k-1})$ is an operator regarding to $H(b_k^+, P_{k-1})$ and defined as

$$
\begin{aligned}
\Xi\left(b_k^+, P_{k-1}\right) \triangleq &-\frac{\beta}{N_0 W}\left[H\left(b_k^+, P_{k-1}\right) - H\left(b_k^+, \overline{P}\right)\right] \\
&- \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right).
\end{aligned}
\tag{25}
$$

Proof: See Appendix D. ∎

Remark 4.5: Note that in the proof to theorem 4.2 in Appendix B and its following discussion, the approximation in (29) is based on the assumption that ω_k is independent of τ, thus can be extracted from $\mathcal{O}(\tau^2)$. This is because the transmission

power ω_k is determined by $\widetilde{b}_k^+$: the normalized energy level in the battery defined in (14) that is independent of τ.

Remark 4.6: Note that the condition in (24) also depends on the parameter $C_2 = \mathbb{E}[r_{k+1}]$ as in (46), i.e., the expected harvested energy in the next time slot. When r_k is i.i.d distributed, C_2 can be regarded as a given parameter as in the proof. However, we do not need to impose restrictions on the exact distribution or other properties of r_{k+1}. The solution applies to all scenarios of stationary arrival sequences of harvested energy, including i.i.d. distribution or (hidden) Markov chain model. As stated in Remark 4.1, we just need to augment the hidden state and the corresponding state transition probability with r_k, then replace the distribution of r_{k+1} with conditional distribution, e.g., replace $\mathbb{E}[r_{k+1}]$ with $\mathbb{E}[r_{k+1}|r_k]$ or $\mathbb{E}[r_{k+1}|e_k]$, where e_k is the environment state at time k as in [20]. We will illustrate the details in the simulation part below. □

After obtaining the closed-form approximate value function H and the associated approximate optimal transmission power schedule, we now need to verify its admissibility of the corresponding transmission power schedule. The following theorem summarizes the results.

Theorem 4.7: Suppose the following conditions are satisfied:

1) $\mathbb{E}[F_i(r_k\tau)] > 1$, $i = 1, 2$, where $F_i(r_k\tau)$ is given by (48) and (51);
2) $\mathbb{P}[\gamma_k = 0|\omega_k = r_k] = \mathbb{E}[\lambda^{r_k\tau}] < (\rho/\|A\|^2)$, for some $\rho \in [0, 1)$ and $\|A\|$ is the largest eigenvalue of A.

Then under the approximate optimal transmission power schedule $\bar{\Theta}^\star$ associated to H in (24), the process is stable, i.e., the the corresponding transmission power schedule is admissible as defined in Definition 3.2.

Proof: See Appendix E. ∎

Remark 4.8: Intuitively, the two conditions in Theorem 4.7 specify the requirements on the average arrival rate of the harvested energy and the packet loss probability. Specifically, condition 2) is similar to the critical value of the packet dropout rate in Kalman filtering with intermittent observations [26]. □

In the next section, we will specify a distribution of r_k to show the implementation and performance of the approximate solution.

V. EXAMPLE AND SIMULATION

In this section, we provide a numerical example to evaluate the estimation performance of the approximate optimal solution and compare the solution with other power schedules.

We consider the Markov chain model described in [20], where there are two states of the external environment: G denotes the good condition (e.g., windy, sunny, etc.) and B denotes the bad condition which may alternate at every time step. At time k, the environment condition state is denoted as e_k and the transition of the two condition states between two time steps follows a Markov chain model (Fig. 3):

The transition probabilities are described by $\mathbb{P}(e_{k+1} = G|e_k = G) = \hat{p}_{00} = 0.7$, $\mathbb{P}(e_{k+1} = B|e_k = G) = \hat{p}_{01} = 0.3$, $\mathbb{P}(e_{k+1} = G|e_k = B) = \hat{p}_{10} = 0.2$, $\mathbb{P}(e_{k+1} = B|e_k = B) = \hat{p}_{11} = 0.8$. Due to the battery capacity limitation, when $r_k\tau >$
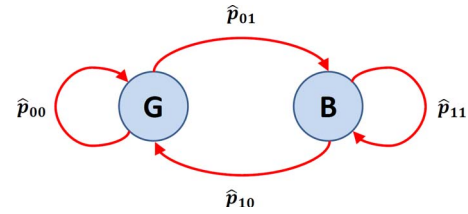

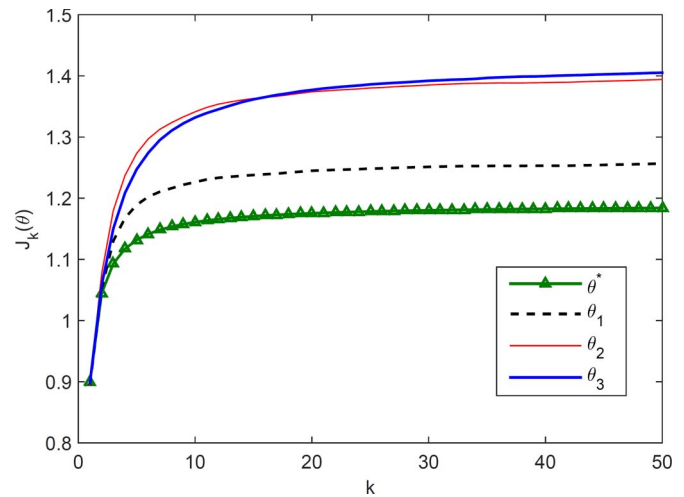Fig. 3. Markov chain model of environment condition.

Fig. 4. Estimation performance comparison of $\theta^\star$, θ_1, θ_2 and θ_3.

$b_{\max}$, we can regard it as $r_k\tau = b_{\max}$ and add up all the corresponding probabilities as $\mathbb{P}[r_k\tau = b_{\max}]$. Under different environment conditions, $r_k\tau$ follows different distributions: $r_k\tau|_{e_k=B} \sim \text{Uniform}(0, 2)$ and $r_k\tau|_{e_k=G} \sim \text{Uniform}(0, 3)$.

Define $J_k(\theta) = (1/k)\sum_{i=1}^{k} \text{Tr}(\mathbb{E}[P_i])$ as the empirical approximation (via 100 000 Monte Carlo simulations) of $J(\theta)$ [see (17)] at every time instant k. As stated in Remark 4.1, here we need to add r_k to the state and corresponding state transition probability, and replace the distribution of r_{k+1} with conditional distribution, e.g., replace $\mathbb{E}[r_{k+1}]$ with $\mathbb{E}[r_{k+1}|e_k]$.

We consider four sensor power schedules for comparison: $\theta^\star$ is the optimal solution to the original Bellman equation (21) (note that as the state in our work is continuous, we can only calculate the optimal solution numerically by discretization with incremental size of 0.05), θ_1 is our proposed approximate optimal solution, θ_2 is the sub-optimal one proposed in [20]

$$\omega_k\tau = \begin{cases} \min\{b_k + r_k\tau, R_0\}, & \text{if } e_k = G \\ \min\{b_k + r_k\tau, R_1\}, & \text{if } e_k = B \end{cases}$$

where the parameters to be designed $R_0 = 1$, and $R_1 = 2$ as in [20], and θ_3 is the "greedy" method, i.e., $\omega_k = r_k$, which refers to using all the harvested energy $r_k\tau$ to send the data packet at each time step. A scalar system is investigated with parameters: $A = 1.1$, $C = 0.7$, $R = Q = 0.8$, $\lambda = 0.7$ and $b_{\max} = 3$. The simulation results are shown in Fig. 4. Similar to the results in [20], the "greedy" method has a better performance only in the first several time steps, which is because the "greedy" method used all the harvested energy instead of reserving some

for the future. Our proposed approximate optimal method θ_1 has a better estimation performance than its counterparts and the gap between θ_1 and $\theta^\star$ is close.

VI. CONCLUSION

We investigated remote state estimation with an energy harvesting sensor. We formulated the power control of the harvesting sensor into an infinite time-horizon MDP problem. The continuous-time approach and perturbation analysis was used and a closed-form approximate value function was derived to deal with the high-dimensional MDP computation complexity. Based on the approximation, we also obtained a closed-form optimal power solution with a threshold-based structure. Numerical examples illustrated the implementation of the optimal solution and demonstrated the performance improvement of the estimation performance compared with other power schedules.

APPENDIX

A. Closed-Form Expression for the Transition Kernel Density

Assume that at time k, the state is $\Phi_k = (b_k^+, P_{k-1})$. Though P_k can take value from a countably infinite set, once P_{k-1} is given, based on the recursion in (16), there are only two possible states for P_k: $h(P_{k-1})$ and $\overline{P}$, associated to $\gamma_k = 0$ and $\gamma_k = 1$, respectively.

After the sensor chooses the transmission power ω_k and sends the data packet carrying $\hat{x}_k^s$, we can calculate the probability densities for different values of P_k.

Suppose $\Phi_{k+1} = (b_{k+1}^+, P_k)$. Clearly, when $P_k \neq \overline{P}$ and $h(P_{k-1})$, we have $T((b_{k+1}^+, P_k), (b_k^+, P_{k-1}), \omega_k) = 0$.

For the case P_k equals $\overline{P}$ or $h(P_{k-1})$, since $b_{k+1}^+ = \min\{b_k^+ - \omega_k\tau + r_{k+1}\tau, b_{\max}\}$, based on the battery level recursion in (13), when $r_{k+1}\tau \leqslant b_{\max} - (b_k^+\omega_k\tau)$ and $b_{k+1}^+ \leqslant b_{\max}$, i.e., the power will not exceed the maximum power level of the battery at the beginning of next time slot, we have $T((b_{k+1}^+, h(P_{k-1})), (b_k^+, P_{k-1}), \omega_k) = \lambda^{\omega_k\tau}f[r_{k+1} = b_{k+1}^+ - (b_k^+ - \omega_k\tau)]$, and $T((b_{k+1}^+, \overline{P}), (b_k^+, P_{k-1}), \omega_k) = (1 - \lambda^{\omega_k\tau})f[r_{k+1} = b_{k+1}^+ - (b_k^+ - \omega_k\tau)]$, where $f[r_{k+1} = X]$ is the probability density of r_{k+1} evaluated at X. Similarly, when $r_{k+1}\tau > b_{\max} - (b_k^+ + \omega_k\tau)$ and $b_{k+1}^+ = b_{\max}$, (i.e., the power will exceed the maximum level of the battery at the beginning of next time slot), we can directly calculate the probabilities of these state transitions: $\mathbb{P}((b_{k+1}^+, h(P_{k-1}))|(b_k^+, P_{k-1}), \omega_k) = \int_{b_{\max} - (b_k^+ + \omega_k\tau)}^{b_{\max}} \lambda^{\omega_k\tau}f[r_{k+1} = (X/\tau)]dX$, and $\mathbb{P}((b_{k+1}^+, \overline{P})|(b_k^+, P_{k-1}), \omega_k) = \int_{b_{\max} - (b_k^+ + \omega_k\tau)}^{b_{\max}}[1 - \lambda^{\omega_k\tau}]f[r_{k+1} = (X/\tau)]dX$.

B. Proof to Theorem 4.2

Proof: We first derive the expressions of terms in the Bellman equation in (21) using Taylor expansion and perturbation analysis.

1) Expressions of $h(X)$, $H(\Phi_{k+1})$ and $v_k(\Phi_k, \omega_k)$ Using Taylor Expansion: Recall that: $A = \exp(\widetilde{A}\tau)$, and $Q = \int_0^\tau \exp(\widetilde{A}s)\widetilde{Q}\exp(\widetilde{A}s)ds$, based on the definition of Lyapunov operator $h(X)$ in (6) and the Taylor expansion of the exponen-

tial function, we have the equivalent expression of Lyapunov operator $h(X)$ with respect to τ

$$
\begin{aligned}
h(X) &= AXA' + Q \\
&= \exp(\widetilde{A}\tau)X\exp(\widetilde{A}\tau)' + \int_0^\tau \exp(\widetilde{A}s)\widetilde{Q}\exp(\widetilde{A}s)ds \\
&= \left[I + \widetilde{A}\tau + \mathcal{O}(\tau^2)\right]X\left[I + \widetilde{A}\tau + \mathcal{O}(\tau^2)\right]' \\
&\quad + \int_0^\tau (I + \widetilde{A}s)\widetilde{Q}(I + \widetilde{A}s)ds \\
&= X + \widetilde{A}\tau X + X\widetilde{A}'\tau + \int_0^\tau (\widetilde{Q} + \widetilde{A}s\widetilde{Q} + \widetilde{Q}\widetilde{A}'s)ds + X\mathcal{O}(\tau^2) \\
&= X + (\widetilde{A}X + X\widetilde{A}' + \widetilde{Q})\tau + X\mathcal{O}(\tau^2). \quad (26)
\end{aligned}
$$

Remark A.1: Note that the above expression can be regarded as an equivalent expression using Taylor expansion respect to τ given other variables. The notation $\mathcal{O}(\tau^2)$ represents the high-order terms of τ with system parameters (e.g., $\widetilde{A}$, $\widetilde{Q}$). □

Based on the perturbation analysis in [11] and [12], when τ is small, we have $r_{k+1}\tau - \omega_k\tau \leqslant b_{\max} - b_k^+$ and $b_k^+ - \omega_k\tau + r_{k+1}\tau \leqslant b_{\max}$, i.e., the power will not exceed the maximum power level of the battery at the beginning of next time slot. Therefore, the equation $b_{k+1}^+ = \min\{b_k^+ - \omega_k\tau + r_{k+1}\tau, b_{\max}\}$ is equivalent to

$$
b_{k+1}^+ = b_k^+ + (r_{k+1} - \omega_k)\tau. \quad (27)
$$

In the following discussion, for convenience, we will use Ψ to denote a general smooth function to represent all the residual cross terms consisting of Φ_k, $H(\Phi_k)$ and ω_k (we will show later that the exact forms of such terms are not needed and Ψ can also represent the combination of multiple Ψ terms). Combining (26) and (27), we obtain the Taylor expansion of $H(b_{k+1}^+, h(P_{k-1}))$ around the point (b_k^+, P_{k-1}) as

$$
\begin{aligned}
&H\left(b_{k+1}^+, h(P_{k-1})\right) \\
&= H\left(b_k^+, P_{k-1}\right) + \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right)\left(b_{k+1}^+ - b_k^+\right) \\
&\quad + \text{Tr}\left\{\nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)\left(h(P_{k-1}) - P_{k-1}\right)\right\} \\
&\quad + \Psi\left(\Phi_k, H(\Phi_k), \omega_k\right)\mathcal{O}(\tau^2) \\
&= H\left(b_k^+, P_{k-1}\right) + \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right)\left(r_{k+1} - \omega_k\right)\tau \\
&\quad + \text{Tr}\left\{\nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)\left(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q}\right)\tau\right\} \\
&\quad + \Psi\left(\Phi_k, H(\Phi_k), \omega_k\right)\mathcal{O}(\tau^2) \\
&H\left(b_{k+1}^+, \overline{P}\right) \\
&= H\left(b_k^+, \overline{P}\right) + \nabla_{b_k^+} H\left(b_k^+, \overline{P}\right)\left(b_{k+1}^+ - b_k^+\right) \\
&\quad + \Psi\left(\Phi_k, H(\Phi_k), \omega_k\right)\mathcal{O}(\tau^2) \\
&= H\left(b_k^+, \overline{P}\right) + \nabla_{b_k^+} H\left(b_k^+, \overline{P}\right)\left(r_{k+1} - \omega_k\right)\tau \\
&\quad + \Psi\left(\Phi_k, H(\Phi_k), \omega_k\right)\mathcal{O}(\tau^2)
\end{aligned}
$$

where $\nabla_{b_k^+}$ and $\nabla_{P_{k-1}}$ are the partial derivatives with respect to the first and the second variable of $H(b_k^+, P_{k-1})$. Note that

$\nabla_{P_{k-1}}$ is the derivative of a scalar function $H(b_k^+, P_{k-1})$ with respect to a matrix variable P_{k-1}, which is defined as

$$\nabla_{P_{k-1}} = \left[\frac{\partial H\left(b_k^+, P_{k-1}\right)}{\partial P_{k-1_{ij}}} \right]_{n_x \times n_x} \tag{28}$$

where $P_{k-1_{ij}}$ denote the elements of the matrix P_{k-1}.

Remark A.2: Based on the previous Taylor expansions, supposing that $H(\Phi_k)$ satisfies the second condition in Theorem 4.2, since the value function $H(b_k^+, P_{k-1})$ is piecewise affine with respect to P_{k-1}, we can observe that $\Psi(\Phi_k, H(\Phi_k), \omega_k)$ consists of linear combination of terms of H or the coefficients of the derivatives of H multiplied by functions of P_{k-1} with order one, i.e., in the form of $\Psi(\Phi_k, H(\Phi_k), \omega_k) = \sum_{i=1}^{\infty} \mathrm{Tr}\{\eta_i \nabla_{b_k^+}^i H(\Phi_k)\tau^i\}$, where certain ω_k, b_k^+ and other constant or bounded system parameters are omitted in the coefficients for simplicity. In addition, as the partial derivatives of different orders with respect to b_k^+ are bounded, the omitted coefficients are also bounded. □

Note that

$$\mathbb{P}[\gamma_k = 0 | \omega_k]$$
$$= \exp\left(-\beta \frac{\omega_k \tau}{N_0 W}\right)$$
$$= 1 - \beta \frac{\omega_k \tau}{N_0 W} + \Psi(\omega_k)\mathcal{O}(\tau^2) \tag{29}$$

and $\mathbb{P}[\gamma_k = 1 | \omega_k] = 1 - \exp(-\beta(\omega_k \tau / N_0 W)) = \beta(\omega_k \tau / N_0 W) + \Psi(\omega_k)\mathcal{O}(\tau^2)$, where the approximations are based on the fact that ω_k is independent of τ, thus can be extracted from the $\mathcal{O}(\tau^2)$. This is because the transmission power ω_k is determined by $\widetilde{b_k^+}$: the normalized energy level in the battery defined in (14) that is independent of τ. Conditioning on whether the data packet arrived successfully or not, we have

$$H\left(b_{k+1}^+, h(P_{k-1})\right) \mathbb{P}[\gamma_k = 0]$$
$$= H\left(b_{k+1}^+, h(P_{k-1})\right)\left(1 - \beta \frac{\omega_k \tau}{N_0 W} + \Psi(\omega_k)\mathcal{O}(\tau^2)\right)$$
$$= H\left(b_k^+, P_{k-1}\right) + \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right)(r_{k+1} - \omega_k)\tau$$
$$+ \mathrm{Tr}\left\{\nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q})\tau\right\}$$
$$- \beta \frac{\omega_k \tau}{N_0 W} H\left(b_k^+, P_{k-1}\right) + \Psi\left(\Phi_k, H(\Phi_k), \omega_k\right)\mathcal{O}(\tau^2)$$

and $H(b_{k+1}^+, \overline{P})\mathbb{P}[\gamma_k = 1] = \beta(\omega_k \tau / N_0 W)H(b_k^+, \overline{P}) + \Psi(\Phi_k, H(\Phi_k), \omega_k)\mathcal{O}(\tau^2)$. Therefore, the approximation of right-hand side of (21) can be written as

$$\mathbb{E}\left[H(\Phi_{k+1}) | \Phi_k, \omega_k, r_{k+1}\right]$$
$$= \mathbb{E}\left[H\left(b_{k+1}^+, h(P_{k-1})\right)\mathbb{P}[\gamma_k = 0]\right.$$
$$\left. + H\left(b_{k+1}^+, \overline{P}\right)\mathbb{P}[\gamma_k = 1] | b_k^+, P_{k-1}, \omega_k, r_{k+1}\right]$$
$$= H\left(b_k^+, P_{k-1}\right) + \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right)(r_{k+1} - \omega_k)\tau$$
$$+ \mathrm{Tr}\left\{\nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q})\tau\right\}$$
$$- \beta \frac{\omega_k \tau}{N_0 W}\left[H\left(b_k^+, P_{k-1}\right) - H\left(b_k^+, \overline{P}\right)\right]$$
$$+ \mathbb{E}\left[\Psi\left(\Phi_k, H(\Phi_k), \omega_k\right)\right]\mathcal{O}(\tau^2). \tag{30}$$

Based on (30) and taking expectation over r_{k+1}, we have

$$\mathbb{E}\left[H(\Phi_{k+1}) | \Phi_k, \omega_k\right]$$
$$= \int_{X=0}^{\infty} \mathbb{E}\left[H(\Phi_{k+1}) | \Phi_k, \omega_k, r_{k+1}\right] \mathbb{P}\left[r_{k+1} = \frac{X}{\tau}\right] dX$$
$$= H\left(b_k^+, P_{k-1}\right) + \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right)\{\mathbb{E}[r_{k+1}] - \omega_k\}\tau$$
$$+ \mathrm{Tr}\left\{\nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q})\tau\right\}$$
$$- \beta \frac{\omega_k \tau}{N_0 W}\left[H\left(b_k^+, P_{k-1}\right) - H\left(b_k^+, \overline{P}\right)\right]$$
$$+ \mathbb{E}\left[\Psi\left(\Phi_k, H(\Phi_k), \omega_k\right)\right]\mathcal{O}(\tau^2). \tag{31}$$

Similarly, the approximation of $v_k(\Phi_k, \omega_k)$ in (18) can be written as

$$v_k(\Phi_k, \omega_k)$$
$$= \lambda^{\omega_k \tau}\mathrm{Tr}\left\{h(P_{k-1})\right\} + [1 - \lambda^{\omega_k \tau}]\mathrm{Tr}\{\overline{P}\}$$
$$= \exp\left(-\beta \frac{\omega_k \tau}{N_0 W}\right)\mathrm{Tr}\left\{h(P_{k-1})\right\}$$
$$+ \left[1 - \exp\left(-\beta \frac{\omega_k \tau}{N_0 W}\right)\right]\mathrm{Tr}\{\overline{P}\}$$
$$= \mathrm{Tr}\left\{P_{k-1} + \left(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q}\right)\tau\right\}$$
$$- \beta \frac{\omega_k \tau}{N_0 W}\mathrm{Tr}\{P_{k-1} - \overline{P}\} + P_{k-1}\mathcal{O}(\tau^2). \tag{32}$$

After obtaining the approximate expression for terms in the Bellman (21), we are now ready to investigate the relationship between the original Bellman (21) and the approximate Bellman (23) in Theorem 4.2.

2) Relationship Between Bellman Equation (21) and (23): To facilitate our subsequent discussion, based on the Bellman (21), we denote

$$\Upsilon^\star(J, H, \Phi_k, \omega_k)$$
$$\triangleq \frac{1}{\tau}\left\{v_k(\Phi_k, \omega_k)\tau + \mathbb{E}\left[H(\Phi_{k+1}) | \Phi_k, \omega_k\right] - H(\Phi_k)\right\} - J \tag{33}$$

We can substitute $\mathbb{E}[H(\Phi_{k+1}) | \Phi_k, \omega_k]$ and $v_k(\Phi_k, \omega_k)$ in (33) with (31) and (32) to obtain

$$\Upsilon^\star(J, H, \Phi_k, \omega_k)$$
$$= \mathrm{Tr}\{P_{k-1}\} - \beta \frac{\omega_k}{N_0 W}\left[H\left(b_k^+, P_{k-1}\right) - H\left(b_k^+, \overline{P}\right)\right]$$
$$+ \mathrm{Tr}\left\{\nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q})\right\}$$
$$+ \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right)(\mathbb{E}[r_{k+1}] - \omega_k) - J$$
$$+ \mathbb{E}\left[\Psi\left(\Phi_k, H(\Phi_k), \omega_k\right)\right]\mathcal{O}(\tau)$$
$$= \left\{-\frac{\beta}{N_0 W}\left[H\left(b_k^+, P_{k-1}\right) - H\left(b_k^+, \overline{P}\right)\right]\right.$$
$$\left. - \nabla_{b_k^+} H(b_k^+, P_{k-1})\right\}\omega_k + \nabla_m H\left(b_k^+, P_{k-1}\right)\mathbb{E}[r_{k+1}]$$
$$+ \mathrm{Tr}\left\{\nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q})\right\}$$
$$+ \mathrm{Tr}\{P_{k-1}\} - J + \mathbb{E}\left[\Psi\left(\Phi_k, H(\Phi_k), \omega_k\right)\right]\mathcal{O}(\tau). \tag{34}$$

Similar to $\Upsilon^\star(J, H, \Phi_k, \omega_k)$, based on the approximate Bellman (23), we denote

$$
\begin{aligned}
&\Upsilon(J, H, \Phi_k, \omega_k) \\
&\triangleq \Bigg\{ -\frac{\beta}{N_0 W} \big[H(b_k^+, P_{k-1}) - H(b_k^+, \overline{P}) \big] - \nabla_{b_k^+} H(b_k^+, P_{k-1}) \Bigg\} \omega_k \\
&\quad + \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right) \mathbb{E}[r_{k+1}] \\
&\quad + \mathrm{Tr}\left\{ \nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right) \left(\widetilde{A} P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q} \right) \right\} \\
&\quad + \mathrm{Tr}\{P_{k-1}\} - J.
\end{aligned}
\tag{35}
$$

Combing (33) and (35), we can rewrite (34) as

$$
\Upsilon^\star(J, H, \Phi_k, \omega_k) = \Upsilon(J, H, \Phi_k, \omega_k) + \mathbb{E}[\Psi(\Phi_k, H(\Phi_k), \omega_k)]\mathcal{O}(\tau).
\tag{36}
$$

Based on the above discussion, we have the following result.

Lemma A.3: Assume that the pair $(J^\star, H^\star)$ satisfies the Bellman (21) and the transversality condition in (22), and that the pair (J, H) satisfies the approximate Bellman equation in (23) and the transversality condition in (22). Then for any state Φ_k we have

1) $\min_{\omega_k} \Upsilon^\star(J^\star, H^\star, \Phi_k, \omega_k) = 0$,
2) $\min_{\omega_k} \Upsilon(J, H, \Phi_k, \omega_k) = 0$,
3) $\min_{\omega_k} \Upsilon^\star(J, H, \Phi_k, \omega_k) = \mathcal{O}(\tau)$.

Proof: Based on (21) and (23), the first two conclusions are straightforward to verify. For the third conclusion, using the relationship in (36), we have

$$
\begin{aligned}
&\min_{\omega_k} \Upsilon^\star(J, H, \Phi_k, \omega_k) \\
&= \min_{\omega_k} \left\{ \Upsilon(J, H, \Phi_k, \omega_k) + \mathbb{E}\left[\Psi(\Phi_k, H(\Phi_k), \omega_k)\right]\mathcal{O}(\tau) \right\} \\
&\geqslant \min_{\omega_k} \Upsilon(J, H, \Phi_k, \omega_k) + \min_{\omega_k}\mathbb{E}\left[\Psi(\Phi_k, H(\Phi_k), \omega_k)\right]\mathcal{O}(\tau) \\
&= 0 + \min_{\omega_k}\mathbb{E}\left[\Psi(\Phi_k, H(\Phi_k), \omega_k)\right]\mathcal{O}(\tau) \\
&= \min_{\omega_k}\mathbb{E}\left[\Psi(\Phi_k, H(\Phi_k), \omega_k)\right]\mathcal{O}(\tau).
\end{aligned}
$$

On the other hand

$$
\begin{aligned}
&\min_{\omega_k} \Upsilon^\star(J, H, \Phi_k, \omega_k) \\
&= \min_{\omega_k} \left\{ \Upsilon(J, H, \Phi_k, \omega_k) + \mathbb{E}\left[\Psi(\Phi_k, H(\Phi_k), \omega_k)\right]\mathcal{O}(\tau) \right\} \\
&\leqslant \min_{\omega_k} \Upsilon(J, H, \Phi_k, \omega_k) + \mathbb{E}\left[\Psi\left(\Phi_k, H(\Phi_k), \omega_k^-\right)\right]\mathcal{O}(\tau) \\
&= \mathbb{E}\left[\Psi\left(\Phi_k, H(\Phi_k), \omega_k^-\right)\right]\mathcal{O}(\tau)
\end{aligned}
$$

where $\omega_k^- = \arg\min_{\omega_k} \Upsilon(J, H, \Phi_k, \omega_k)$.

Since $H(\Phi_k)$ is assumed to satisfy the second condition in Theorem 4.2, the finite summations of power functions with respect to P_{k-1} are bounded as stated in Remark A.2. Then we have

$$
\begin{aligned}
\mathbb{E}\left[\Psi(\Phi_k, H, \omega_k)\right] &= \mathbb{E}\left[\sum_{i=1}^{\infty} \mathrm{Tr}\left\{ \eta_i \nabla_{b_k^+}^i H(\Phi_k)\tau^i \right\} \right] \\
&\leqslant \mathbb{E}\left[\sum_{i=1}^{\infty} \mathrm{Tr}\left\{ \mathcal{M}\tau^i P_{k-1} \right\} \right] \\
&= \mathbb{E}\left[\frac{\mathcal{M}\tau}{1-\tau} \mathrm{Tr}\{P_{k-1}\} \right]
\end{aligned}
$$

where $\mathcal{M}$ is the uniform bound of the coefficients of the derivatives of H with resect to b_k^+ and other corresponding coefficients. Under the admissible strategy defined in Definition 3.2, $\mathbb{E}[\mathrm{Tr}\{P_{k-1}\}]$ is bounded. Consequently, we have $\Upsilon^\star(J, H, \Phi_k) = \mathcal{O}(\tau), \forall \Phi_k.$ ∎

3) Difference Between $H^\star(\Phi_k)$ and $H(\Phi_k)$: Based on Lemma A.3, we can prove the main result of Theorem 4.2 concerning the difference between $H^\star(\Phi_k)$ and $H(\Phi_k)$. Suppose that the pairs $(J^\star, H^\star)$ and (J, H) satisfy the conditions in Lemma A.3, respectively, then we have $\Upsilon^\star(J, H, \Phi_k) = \mathcal{O}(\tau), \forall \Phi_k.$ When $\tau \to 0$, $\Upsilon^\star(J, H, \Phi_k) \to 0, \forall \Phi_k$, i.e., the pair (J, H) satisfies the conditions in the Bellman (21). As the pair $(J^\star, H^\star)$ is the unique solution to the Bellman (21), we have $H^\star(\Phi_k) = H(\Phi_k)$ when $\tau \to 0$, i.e., $H^\star(\Phi_k) - H(\Phi_k) = \mathcal{O}(\tau).$ ∎

C. Proof to Theorem 4.3

To facilitate the following discussion, we denote:

$$
\Upsilon(H, \Phi_k, \omega_k) \triangleq \Upsilon(J, H, \Phi_k, \omega_k) + J
\tag{37}
$$

$$
\begin{aligned}
\Upsilon^\star(H, \Phi_k, \omega_k) &\triangleq \Upsilon^\star(J, H, \Phi_k, \omega_k) + J \\
&= \Upsilon(H, \Phi_k, \omega_k) + \mathcal{O}(\tau).
\end{aligned}
\tag{38}
$$

Based on Lemma A.3, it is straightforward that

$$
\min_{\omega_k} \Upsilon(H, \Phi_k, \omega_k) = \Upsilon(H, \Phi_k, \bar{\omega}^\star)
\tag{39}
$$

and $J^\star = \min_{\omega_k} \Upsilon^\star(H, \Phi_k, \omega_k) = \Upsilon^\star(H^\star, \Phi_k, \omega^\star) = \Upsilon(H^\star, \Phi_k, \omega_k^\star) + \mathcal{O}(\tau)$, where $\bar{\omega}^\star$ and $\omega^\star$ are the optimal power under policy $\bar{\Theta}^\star$ and $\Theta^\star$, respectively.

On the other hand, given the optimal admissible policy $\bar{\Theta}^\star$ for the approximate Bellman (23), we have

$$
\begin{aligned}
\mathbb{E}^{\bar{\Theta}^\star}\left[\mathbb{E}\left[H^\star(\Phi_{k+1})|\Phi_k, \omega_k\right] \right] &= \mathbb{E}^{\bar{\Theta}^\star}\left[H^\star(\Phi_{k+1})|\Phi_k\right] \\
&= \mathbb{E}^{\bar{\Theta}^\star}\left[H^\star(\Phi_{k+1})\right] \\
&= \mathbb{E}^{\bar{\Theta}^\star}\left[H^\star(\Phi_k)\right]
\end{aligned}
\tag{40}
$$

where the different time index in Φ only represents different states of the process. Based on (19) and the above discussion, the corresponding objective value $\bar{J}^\star$ under $\bar{\Theta}^\star$ can be calculated as follows:

$$
\begin{aligned}
\bar{J}^\star &= \lim_{T\to\infty} \frac{1}{T} \mathbb{E}^{\bar{\Theta}^\star}\left[\sum_{k=1}^{T} v_k(\Phi_k, \omega_k) \right] \\
&= \mathbb{E}^{\bar{\Theta}^\star}\left[v_k(\Phi_k, \bar{\omega}^\star) \right] \\
&= \mathbb{E}^{\bar{\Theta}^\star}\left[v_k(\Phi_k, \bar{\omega}^\star) + \frac{1}{\tau}\left(\mathbb{E}\left[H^\star(\Phi_{k+1})|\Phi_k, \bar{\omega}^\star\right] - H^\star(\Phi_k) \right) \right] \\
&= \mathbb{E}^{\bar{\Theta}^\star}\left[\Upsilon^\star(H, \Phi_k, \bar{\omega}^\star) \right] \\
&= \mathbb{E}^{\bar{\Theta}^\star}\left[\Upsilon(H^\star, \Phi_k, \bar{\omega}^\star) + \mathcal{O}(\tau) \right] \\
&\leqslant \mathbb{E}^{\bar{\Theta}^\star}\left[\Upsilon(H, \Phi_k, \omega^\star) + \mathcal{O}(\tau) \right].
\end{aligned}
$$

Therefore, due to the Lipschitz continuity of operator $\Upsilon(H, \Phi_k, \omega_k)$ [24], we have

$$
\begin{aligned}
J^\star - \bar{J}^\star &= \mathbb{E}^{\bar\Theta^\star}\left[\Upsilon(H, \Phi_k, \omega^\star) - \Upsilon(H, \Phi_k, \omega^\star) + \mathcal{O}(\tau)\right] \\
&\leqslant \alpha \|H^\star - H\| + \mathcal{O}(\tau) = \mathcal{O}(\tau)
\end{aligned}
$$

where α is some constant and $\|\cdot\|$ is any norm.

D. Proof to Theorem 4.4

1) Approximate Optimal Transmission Power Solution:
Since we have $\Xi(b_k^+, P_{k-1}) = -(\beta/N_0 W)[H(b_k^+, P_{k-1}) - H(b_k^+, \overline{P})] - \nabla_{b_k^+} H(b_k^+, P_{k-1})$, we can rewrite J in (23) in a simpler form

$$
\begin{aligned}
J = &\min_{\omega_k}\left\{\Xi\left(b_k^+, P_{k-1}\right)\right\}\omega_k \\
&+ \operatorname{Tr}\left\{\nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)\left(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q}\right)\right\} \\
&+ \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right)\mathbb{E}[r_{k+1}] + \operatorname{Tr}\{P_{k-1}\}. \quad (41)
\end{aligned}
$$

Lemma A.4: The relative function $H(b_k^+, P_{k-1})$ has following properties:
1) $H(b_k^+, P_{k-1}) > H(b_k^+, \overline{P}), \forall P_{k-1} \geqslant \overline{P}$,
2) $\nabla_{b_k^+} H(b_k^+, P_{k-1}) \leqslant 0$.

Proof: As $H(b_k^+, P_{k-1})$ is the relative value function representing the optimal expected average state estimation error covariance value obtained starting from the state (b_k^+, P_{k-1}). As more energy and lower initial error covariance are always beneficial (proved in [18]), it is straightforward that $H(b_k^+, P_{k-1})$ is monotonic increasing with P_{k-1} and decreasing with b_k^+, which finishes the proof. ∎

Based on Lemma A.4, we have $-(\beta/N_0 W)[H(b_k^+, P_{k-1}) - H(b_k^+, \overline{P})] \leqslant 0$ and $-\nabla_{b_k^+} H(b_k^+, P_{k-1}) \geqslant 0$. Therefore, $\Xi(b_k^+, P_{k-1})$ is not necessarily positive or negative for all b_k^+ and P_{k-1}.

Based on value of the function $\Xi(b_k^+, P_{k-1})$, we can obtain an event-based threshold structure optimal solution to (34) in the following form:

$$
\omega_k^\star = \begin{cases} \max\{\omega_k \in \mathbb{A}_k\} = \widetilde{b_k^+}, & \text{if } \Xi\left(b_k^+, P_{k-1}\right) \leqslant 0 \\ 0, & \text{otherwise.} \end{cases} \quad (42)
$$

To be more specific, the approximate optimal transmission power of Problem 3.1 is given by (24).

Next we need to obtain the closed-form expression for the relative value function H by solving the PDE in (23). Substituting the ω_k in (23) with $\omega_k^\star$ in (42), when $\Xi(b_k^+, P_{k-1}) \leqslant 0$, we can get

$$
\begin{aligned}
J = &-\frac{\beta b_k^+}{N_0 W \tau}\left[H\left(b_k^+, P_{k-1}\right) - H\left(b_k^+, \overline{P}\right)\right] \\
&- \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right)\frac{b_k^+}{\tau} + \nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right)\mathbb{E}[r_{k+1}] \\
&+ \operatorname{Tr}\left\{\nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)\left(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q}\right)\right\} \\
&+ \operatorname{Tr}\{P_{k-1}\}. \quad (43)
\end{aligned}
$$

Otherwise, we have

$$
\begin{aligned}
J = &\nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right)\mathbb{E}[r_{k+1}] \\
&+ \operatorname{Tr}\left\{\nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)\left(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q}\right)\right\} \\
&+ \operatorname{Tr}\{P_{k-1}\}. \quad (44)
\end{aligned}
$$

2) Solution to (43):
To solve (43), we first assume that the relative value function $H(b_k^+, P_{k-1})$ has a solution in the form of

$$
H_1(m, L) = \operatorname{Tr}\{F_1(m)L\} + G_1(m) \quad (45)
$$

where $F_1(m) \in \mathbb{R}^{n \times n}$ and $G_1(m) \in \mathbb{R}$ are continuous functions.

Substituting the general solution into (43), we can obtain that

$$
\begin{aligned}
J = \operatorname{Tr}\Bigg\{&\left[C_1 m F_1(m) - \frac{1}{\tau}m\dot{F}_1(m) + C_2\dot{F}_1(m)\right. \\
&\left.+ F_1(m)\widetilde{A} + \widetilde{A}'F_1(m) + I\right]L\Bigg\} \\
&+ \operatorname{Tr}\left\{-C_1 m F_1(m)\overline{P} + F_1(m)\widetilde{Q}\right\} \\
&- \frac{1}{\tau}m\dot{G}_1(m) + C_2\dot{G}_1(m) \quad (46)
\end{aligned}
$$

where $C_1 \triangleq -(\beta/N_0 W)$, $C_2 \triangleq \mathbb{E}[r_{k+1}]$ and $\dot{G}_1(m)$, $\dot{F}_1(m)$ are the first-order derivatives of corresponding functions.

First, we solve the part with L and require the following equation holding for any m and L: $C_1 m F_1(m) - (1/\tau)m\dot{F}_1(m) + C_2\dot{F}_1(m) + F_1(m)\widetilde{A} + \widetilde{A}'F_1(m) + I = 0$, i.e., $C_1 m f_1(m)_{ij} - (1/\tau)m\dot{f}_1(m)_{ij} + C_2\dot{f}_1(m)_{ij} + \sum_{r=1}^{n} a_{rj} f_1(m)_{ir} + \sum_{r=1}^{n} a_{ri} f_1(m)_{rj} + \delta_{ij} = 0$, $i, j = 1, 2, \ldots, n$, where $f_1(m)_{ij}$, $\dot{f}_1(m)_{ij}$ and a_{ij} are the elements of corresponding matrices.

By using the method of dominant balance (MDB) [12], we can obtain the asymptotic solution of $f_1(m)_{ij}$

$$
f_1(m)_{ij} = \left(C_2 - \frac{1}{\tau}m\right)^{C_1 C_2 \tau^2 + a_{ij}\delta_{ij}\tau} \exp(C_1 m\tau) \quad (47)
$$

thus we have

$$
F_1(m) = \{f_1(m)_{ij}\} \quad (48)
$$

where $f_1(m)_{ij}$ is given in (47).

For the second part without L in (46), similarly we require that: $\operatorname{Tr}\{-C_1 m F_1(m)\overline{P} + F_1(m)\widetilde{Q}\} - (1/\tau)m\dot{G}_1(m) + C_2\dot{G}_1(m) = J$. We can write the above PDE into elements: $-C_1\operatorname{Tr}\{\overline{P}\}m f_1(m)_{ij} - (1/\tau)mn\delta_{ij}\dot{G}_1(m) + C_2 n\delta_{ij}\dot{G}_1(m) + \sum_{r=1}^{n} q_{rj} f_1(m)_{ir} = n\delta_{ij}J$, where $\widetilde{Q} = \{q_{ij}\}$. To deal with the sum of functions in the above PDE, we can separate it into two parts: $-C_1\operatorname{Tr}\{\overline{P}\}m f_1(m)_{ij} - (1/\tau)mn\delta_{ij}\dot{G}_1(m) + C_2 n\delta_{ij}\dot{G}_1(m) = 0$ and $-(1/\tau)m\dot{G}_1(m) + C_2\dot{G}_1(m) + q_{rj}f_1(m)_{ir} = n\delta_{ij}J$, where $i, j, r = 1, 2, \ldots, n$ and the solution of $G_1(m)$ is the sum of the solutions to the above two PDEs. Using MDB we can obtain the solutions for each PDE: $G_{11}(m) = (1/n)C_1 C_2\operatorname{Tr}\{\overline{P}\}\log(C_2\tau - m) - C_1(m/n)$ and

$G_{12}(m)_{ij} = \int (C_1 b_{ji}^+ m / (C_2 \tau - m)) f_1(m)_{ij} dm$. Therefore, we have

$$G_1(m) = G_{11}(m) + \sum_{i=1}^{n} \sum_{j=1}^{n} \left[G_{12}(m)_{ij} \right]. \tag{49}$$

Note that we do not need solve the indefinite integrals above. As we will show later, we only need a closed-form expression of the derivative of $G(m)$. Given the above approximate solution of (43) is given by (48) and (49).

3) Solution to (44): To solve (44), we can also assume the relative value function $H(b_k^+, P_{k-1})$ has a solution is in the form of

$$H_2(m, L) = \text{Tr}\left\{F_2(m)L\right\} + G_2(m) \tag{50}$$

where $F_2(m) \in \mathbb{R}^{n \times n}$ and $G_2(m) \in \mathbb{R}$ are continuous functions. Following the same procedure as for the case considered above, we first require that $C_2 \dot{F}_2(m) + F_2(m)\widetilde{A} + \widetilde{A}' F_2(m) + \widetilde{Q} = 0$. Then using the MDB method, we have the asymptotic solution of $\widetilde{f}_2(m)_{ij}$

$$f_2(m)_{ij} = \begin{cases} C_1 \exp\left(-\dfrac{a_{ij} m}{C_2}\right) - \dfrac{q_{ij}}{a_{ij}}, & \text{if } i \neq j \\ C_1 - \dfrac{q_{ij} m}{C_2}, & \text{if } i = j. \end{cases}$$

Therefore, $F_2(m)$ is given by

$$F_2(m) = \left\{ \widetilde{f}_2(m)_{ij} \right\}. \tag{51}$$

For the second part of (44) without L, similarly, we require $\text{Tr}\{C_2 \dot{G}_1(m) \widetilde{F}_1(m) \widetilde{Q}\} = J$, and obtain

$$G_2(m) = G_{21}(m) + \sum_{i=1}^{n} \sum_{j=1}^{n} \left[G_{22}(m)_{ij} \right] \tag{52}$$

where $G_{21}(m) = (m/C_2)$, $G_{22}(m)_{ij} = \int -(q_{ji}/C_2) f_2(m)_{ij} dm$. Clearly, the solutions obtained for H satisfies the two conditions in Theorem 4.2.

4) Combination of Solutions to (43) and (44): The approximate solution of the relative value function $H(b_k^+, P_{k-1})$ to (23) is given by the combination of solutions to (43) and (44). The threshold condition (42) for choosing the transmission power implies that $H(b_k^+, P_{k-1})$ has a piecewise structure to fit these two solutions together

$$H(m, L) = \begin{cases} H_1(m, L), \\ \quad \text{if } \Xi_1(m, L) \leqslant 0 \ \&\& \ \Xi_2(m, L) \leqslant 0 \\ H_2(m, L), \\ \quad \text{if } \Xi_1(m, L) > 0 \&\& \Xi_2(m, L) > 0 \\ \frac{1}{2}\left[H_1(m, L) + H_2(m, L) \right], \\ \quad \text{otherwise} \end{cases} \tag{53}$$

where Ξ_1 and Ξ_2 are operators defined as in (25) except replacing H with H_1 in (45) and H_2 in (50), respectively.

E. Proof to Theorem 4.7

To investigate the stability of P_k under the proposed approximate optimal transmission power schedule associated to H in (53), first we define a Lyapunov function as $L(b_k^+, P_{k-1}) \stackrel{\Delta}{=}$

$H(b_k^+, P_{k-1})$, the conditional Lyapunov drift as $dL(P_{k-1}) = \mathbb{E}^{\bar{\Theta}^\star}[L(b_{k+1}^+, P_k) - L(b_k^+, P_{k-1})|P_{k-1}]$, and the conditional state estimation error covariance drift as $d(P_{k-1}) = \mathbb{E}^{\bar{\Theta}^\star}[\text{Tr}\{P_k - P_{k-1}\}|P_{k-1}]$.

Now we want to investigate the relation between $dL(P_{k-1})$ and $d(P_{k-1})$. On one hand, when P_{k-1} becomes large, given P_{k-1}, we have

$$\begin{aligned} &dL(P_{k-1}) \\ &= \mathbb{E}^{\bar{\Theta}^\star} \left[L\left(b_{k+1}^+, P_k\right) - L\left(b_k^+, P_{k-1}\right) | P_{k-1} \right] \\ &= \mathbb{E}^{\bar{\Theta}^\star} \left[H\left(b_{k+1}^+, P_k\right) - H\left(b_k^+, P_{k-1}\right) | P_{k-1} \right] \\ &= \mathbb{E}^{b_k^+} \left[\mathbb{E}^{\bar{\Theta}^\star} \left[H\left(b_{k+1}^+, P_k\right) - H\left(b_k^+, P_{k-1}\right) | P_{k-1}, b_k^+ \right] \right] \\ &= \mathbb{E}^{b_k^+} \left[\mathbb{E}^{\bar{\Theta}^\star} \left[\text{Tr}\left\{ \nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)\left(P_k - P_{k-1}\right) \right\} \right.\right. \\ &\qquad\qquad + \left.\left. \Psi\left(\Phi_k, H(\Phi_k), \omega_k\right) \mathcal{O}(\tau) | P_{k-1}, b_k^+ \right] \right] \\ &\geqslant \mathbb{E}^{b_k^+} \left[\mathbb{E}^{\bar{\Theta}^\star} \left[\text{Tr}\{ \nabla_{P_{k-1}} H(b_k^+, P_{k-1})(P_k - P_{k-1})\} | P_{k-1}, b_k^+ \right] \right] \\ &\geqslant \mathbb{E}^{b_k^+} \left[\nabla_{P_{k-1}} H(b_k^+, P_{k-1}) | b_k^+ \right] \mathbb{E}^{\bar{\Theta}^\star}[\text{Tr}\{(P_k - P_{k-1})\} | P_{k-1}] \\ &\geqslant \mathbb{E}^{\bar{\Theta}^\star} \left[\text{Tr}\left\{ P_k - P_{k-1} \right\} | P_{k-1} \right] = (P_{k-1}) \end{aligned}$$

where the first inequality holds for certain positive constant C and is due to the properties of the relative value function H (increasing with P_k and decreasing with b_k^+), the second inequality is based on the inequalities for the trace of matrix product in [27] and the third inequality holds when $\nabla_{P_{k-1}} H(b_k^+, P_{k-1}) = F(b_k^+)$ satisfies the condition 1) of this theorem.

Now we consider a simple "greedy" policy as $\Theta_g : \omega_k = \widetilde{b_k^+} = r_k$. Clearly, based on the results in [9], when condition 2) of this theorem holds, the greedy policy is admissible in the sense that P_k is bounded.

On the other hand, the Lyapunov drift when P_k becomes large is given by

$$\begin{aligned} &dL(P_{k-1}) \\ &= \mathbb{E}^{\bar{\Theta}^\star} \left[H\left(b_{k+1}^+, P_k\right) - H\left(b_k^+, P_{k-1}\right) | P_{k-1} \right] \\ &= \mathbb{E}^{\bar{\Theta}^\star} \left[\nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right) \left\{ \mathbb{E}[r_{k+1}] - \omega_k \right\} \tau \right. \\ &\qquad + \text{Tr}\left\{ \nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q})\tau \right\} \\ &\qquad - \beta \frac{\omega_k \tau}{N_0 W} \left[H\left(b_k^+, P_{k-1}\right) - H\left(b_k^+, \overline{P}\right) \right] \\ &\qquad + \left. \mathbb{E}\left[\Psi\left(\Phi_k, H(\Phi_k), \omega_k\right) \right] \mathcal{O}(\tau^2) | P_{k-1} \right] \\ &\leqslant \mathbb{E}^{\Theta_g} \left[\nabla_{b_k^+} H\left(b_k^+, P_{k-1}\right) \left\{ \mathbb{E}[r_{k+1}] - \omega_k \right\} \tau \right. \\ &\qquad + \text{Tr}\left\{ \nabla_{P_{k-1}} H\left(b_k^+, P_{k-1}\right)(\widetilde{A}P_{k-1} + P_{k-1}\widetilde{A}' + \widetilde{Q})\tau \right\} \\ &\qquad - \beta \frac{\omega_k \tau}{N_0 W} \left[H\left(b_k^+, P_{k-1}\right) - H\left(b_k^+, \overline{P}\right) \right] \\ &\qquad + \left. \mathbb{E}\left[\Psi\left(\Phi_k, H(\Phi_k), \omega_k\right) \right] \mathcal{O}(\tau^2) | P_{k-1} \right] \\ &\leqslant J\tau - \mathbb{E}^{\Theta_g} \left[v_k(\Phi_k, \omega_k)\tau \right] \\ &= J\tau - \mathbb{E}^{\Theta_g} \left[\lambda^{\omega_k \tau} \text{Tr}\left\{ h(P_{k-1}) \right\} - [1 - \lambda^{\omega_k \tau}] \text{Tr}\{\overline{P}\} \right] < 0 \end{aligned}$$

where the first and the second inequality follow from the definition of $\Upsilon^\star(J, H, \Phi_k, \omega_k)$ in (33) and (34), $v_k(\Phi_k, \omega_k)$ in (32) and the result in Lemma A.3 that $\bar{\Theta}^\star$ minimizes $\Upsilon^\star(J, H, \Phi_k, \omega_k)$ rather than Θ_g (but Θ_g is a stablizing policy making the RHS of the first inequality bounded). The third inequality holds when P_{k-1} becomes large.

Combining the above two inequalities between $dL(P_{k-1})$ and $d(P_{k-1})$ yields

$$d(P_{k-1}) \leqslant dL(P_{k-1}) < 0 \qquad (54)$$

Based on the Lyapunov theory [28], the negative state drift in (54) ensures the stability of $\{P_k\}$, i.e., the admissibility of our proposed policy $\bar{\Theta}^{\star}$.

REFERENCES

[1] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Comput. Netw.*, vol. 52, no. 12, pp. 2292–2330, 2008.

[2] 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.

[3] A. A. Aziz, Y. A. Sekercioglu, P. Fitzpatrick, and M. Ivanovich, "A survey on distributed topology control techniques for extending the lifetime of battery powered wireless sensor networks," *IEEE Commun. Surveys Tutorials*, vol. 15, no. 1, pp. 121–144, 2013.

[4] N. A. Pantazis and D. D. Vergados, "A survey on power control issues in wireless sensor networks," *IEEE Commun. Tutorials*, vol. 9, no. 4, pp. 86–107, 2007.

[5] Y. Zhang, S. He, and J. Chen, "Data gathering optimization by dynamic sensing and routing in rechargeable sensor networks," *IEEE/ACM Trans. Netw.*, to be published, doi: 10.1109/TNET.2015.2425146.

[6] P. Cheng, Y. Qi, K. Xin, J. Chen, and L. Xie, "Energy-efficient data forwarding for state estimation in multi-hop wireless sensor networks," *IEEE Trans. Autom. Control*, vol. 61, no. 5, pp. 1322–1327, May 2016.

[7] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: Survey and implications," *IEEE Commun. Surveys Tutorials*, vol. 13, no. 3, pp. 443–461, 2011.

[8] E. Y. Bitar, R. Rajagopal, P. P. Khargonekar, K. Poolla, and P. Varaiya, "Bringing wind energy to market," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1225–1235, Aug. 2012.

[9] M. Nourian, A. Leong, and S. Dey, "Optimal energy allocation for Kalman filtering over packet dropping links with imperfect acknowledgments and energy harvesting constraints," *IEEE Trans. Autom. Control*, vol. 59, no. 8, pp. 2128–2143, Aug. 2014.

[10] W. Chen *et al.*, "Approximate dynamic programming using fluid and diffusion approximations with applications to power management," in *Proc. 48th IEEE Conf. Decision Control*, 2009, pp. 3575–3580.

[11] V. K. Lau, F. Zhang, and Y. Cui, "Low complexity delay-constrained beamforming for multi-user MIMO systems with imperfect CSIT," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4090–4099, Aug. 2013.

[12] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory*. New York, NY, USA: Springer, 1999, vol. 1.

[13] A. Nayyar, T. Başar, D. Teneketzis, and V. V. Veeravalli, "Optimal strategies for communication and remote estimation with an energy harvesting sensor," *IEEE Trans. Autom. Control*, vol. 58, no. 9, pp. 2246–2260, Sep. 2013.

[14] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. New York, NY, USA: McGraw-Hill, 2007.

[15] M. Micheli and M. I. Jordan, "Random sampling of a continuous-time stochastic dynamical system," in *Proc. 15th Int. Symp. Math. Theory Netw. Syst.*, 2002, Citeseer.

[16] P. Hovareshti, V. Gupta, and J. S. Baras, "Sensor scheduling using smart sensors," in *Proc. 46th IEEE Conf. Decision Control*, 2007, pp. 494–499.

[17] B. D. O. Anderson and J. B. Moore, "Detectability and stabilizability of time-varying discrete-time linear systems," *SIAM J. Control Optim.*, vol. 19, no. 1, pp. 20–32, 1981.

[18] Y. Li, D. E. Quevedo, V. Lau, and L. Shi, "Optimal periodic transmission power schedules for remote estimation of ARMA processes," *IEEE Trans. Signal Process.*, vol. 61, no. 24, pp. 6164–6174, Dec. 2013.

[19] L. Shi, M. Epstein, and R. M. Murray, "Kalman filtering over a packet-dropping network: A probabilistic perspective," *IEEE Trans. Autom. Control*, vol. 55, no. 3, pp. 594–604, Mar. 2010.

[20] Y. Li, D. E. Quevedo, V. Lau, S. Dey, and L. Shi, "Transmission power scheduling for energy harvesting sensor in remote state estimation," presented at the 19th IFAC World Congr. Int. Fed. Autom. Control, Cape Town, South Africa, 2014.

[21] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[22] D. E. Quevedo, A. Ahlén, A. S. Leong, and S. Dey, "On Kalman filtering over fading wireless channels with controlled transmission powers," *Automatica*, vol. 48, no. 7, pp. 1306–1316, 2012.

[23] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY, USA: Wiley, 2009, vol. 414.

[24] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, WA, USA: Athena, 1995.

[25] O. Hernández-Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes*. New York, NY, USA: Springer, 1996.

[26] B. Sinopoli *et al.*, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[27] Y. Fang *et al.*, "Inequalities for the trace of matrix product," *IEEE Trans. Autom. Control*, vol. 39, no. 12, pp. 2489–2490, Dec. 1994.

[28] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic power allocation and routing for time-varying wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 89–103, Jan. 2005.

Yuzhe Li received the B.S. degree in mechanics from Peking University, Beijing, China, in 2011, and the Ph.D. degree in electronic and computer engineering from the Hong Kong University of Science and Technology, Hong Kong, in 2015. Between June 2013 and August 2013, he was a visiting scholar in the University of Newcastle, Newcastle. Australia. He is currently a Postdoctoral Fellow at the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada.

His current research interests include cyber-physical system security, sensor power control and networked state estimation.

Fan Zhang revived the B.Eng. (First Class Hons) from Chu Kochen Honors College, Zhejiang University, Zhejiang, China, in 2010, and the Ph.D. degree from Hong Kong University of Science and Technology, Hong Kong, in 2015.

He is now a Senior Researcher in the Future Network Theory Lab of Huawei Technologies. His research interests include cross-layer QoE-aware resource allocation for 5G wireless systems, low complexity dynamic programming and control, and networked control systems.

Daniel E. Quevedo received Ing. Civ. Electrónico and the M.Sc. degrees from Universidad Técnica Federico Santa María, Chile, in 2000, and the Ph.D. degree from the University of Newcastle, Newcastle, Australia, in 2005.

He holds the Chair in Automatic Control at Paderborn University, Paderborn, Germany. He is Editor of the *International Journal of Robust and Nonlinear Control*. His research mainly focuses on networked estimation and control and on predictive control of power converters.

Prof. Quevedo serves as Chair of the IEEE CSS Technical Committee on Networks and Communication Systems. In 2009, he was awarded a five-year fellowship from the Australian Research Council.

Vincent Lau received the B.Eng. (Distinction first Hons) from the University of Hong Kong, Hong Kong, in 1992, and the Ph.D. degree from the Cambridge University, Cambridge, U.K., in 1997.

He joined Bell Labs from 1997–2004 and the Department of Electrical and Computer Engineering, Hong Kong University of Science and Technology (HKUST), Hong Kong, in 2004. He is currently a Chair Professor and the Founding Director of Huawei-HKUST Joint Innovation Lab at HKUST. His current research focus includes robust cross layer optimization for wireless systems, massive MIMO, compressed sensing, networked control systems, as well as PHY caching for wireless networks.

Subhrakanti Dey received the B.Tech. and M.Tech. degrees from Indian Institute of Technology, Kharagpur, India, in 1991 and 1993, respectively, and the Ph.D. degree from Australian National University, Canberra, Australia, in 1996.

He is currently a Professor with the Department of Engineering Sciences, Uppsala University, Sweden. His current research interests include networked control systems, wireless communications and networks, signal processing for sensor networks, and stochastic and adaptive signal processing and control. He served as Associate Editor for *Elsevier Systems and Control Letters* during 2003–2013.

Prof. Dey currently serves on the Editorial Board of IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS. He was also an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING during 2007–2010 and the IEEE TRANSACTIONS ON AUTOMATIC CONTROL during 2004–2007.

Ling Shi received the B.S. degree in electrical and electronic engineering from Hong Kong University of Science and Technology, Hong Kong, in 2002 and the Ph.D. degree in control and dynamical systems from California Institute of Technology, Pasadena, CA, USA, in 2008.

He is currently an Associate Professor at the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology. His research interests include networked control systems, wireless sensor networks, event-based state estimation and sensor scheduling, and smart energy systems. He has been serving as a subject editor for *International Journal of Robust and Nonlinear Control* since 2015.

Dr. Shi served as an Associate Editor for a special issue on secure control of cyber physical systems in the IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS in 2015.