Transmission Design for Energy-Efficient Vehicular Networks with Multiple Delay-limited Applications

Danyan Lan, Chao Wang, Ping Wang, Fuqiang Liu, and Geyong Min

Abstract

Vehicular networking is potentially an effective solution to the problems in today's transportation system. However, realizing efficient communications among vehicles with satisfactory reliability and latency requirements is challenging, especially when diverse applications are taken into consideration. In this paper, we investigate a cross-layer energy-efficient transmission design for a class of vehicular communication networks, in which two pairs of vehicle-to-vehicle links non-orthogonally share the available spectrum. Each link desires to deliver two types of messages that can support different delay-limited applications. The periodically-generated heartbeat messages should be transmitted subject to a reliability requirement, and the randomly-appeared sensing messages should be delivered with finite latency. We propose a power control strategy to achieve high energy efficiency, while ensuring the expected quality-of-service requirements, based on both channel state information in the physical layer and queue state information in the media access control layer. Simulation results show that our proposed method notably outperforms conventional methods.

Index Terms

Cross-layer design, Energy efficiency, Lyapunov optimization, Vehicular networks

I. INTRODUCTION

Road safety, traffic efficiency and energy consumption have become the major concerns in modern road transportation systems. Establishing an intelligent transportation system (ITS) using information and communication technologies (ICT) has been widely accepted as the key solution to these issues. Equipping vehicles with advanced sensing and computing devices enables human drivers and artificial intelligence (AI) driving engines to attain a better understanding of the complex driving environment. Using wireless communication to connect vehicles and roadside infrastructure further enhances the sensing range and accuracy of individual vehicles.

For instance, a typical vehicular networking use case is illustrated in Fig. 1. Under the concept of cooperative adaptive cruise control (CACC), the vehicles V_{S_1} and V_{D_1} form a platoon. The leading vehicle V_{S_1} can be driven by human or AI, and V_{D_1} intends to closely follow V_{S_1}. If V_{S_1} periodically sends its status (e.g., speed and location) to V_{D_1}, the following vehicle can attain accurate knowledge regarding V_{S_1} to direct its driving operations. When V_{S_1} detects certain objects of interest in front of it, it can also share such information with V_{D_1} to enhance the latter's sensing range. The coordinated sensing and maneuvering actions of the platoon allow the vehicles to safely drive with high speed and small inter-vehicle distance. Hence ITS has been deemed to be a core application scenario of 5G technologies [1]. With the support of strong communication capability, vehicles can even share sensing, computing, storage, and communication resources to realize the concept of vehicular cloud networks and fulfill advanced tasks far beyond individual vehicle's capability [2].

However, enabling high-performance communication in vehicular networks is challenging, due to the complex signal propagation environment. Different applications pose diverse quality-of-service (QoS) requirements on delay, accuracy and throughput etc. Efficient and reliable transmission designs are needed, especially when multiple users coexist and the available resources (e.g., bandwidth and battery) are limited.

D. Lan, C. Wang, P. Wang, and F. Liu are with the Department of Information and Communication Engineering, Tongji University, Shanghai, China, C. Wang is also with the Department of Computer Science, University of Exeter, Exeter, UK. G. Min is with the Department of Computer Science, University of Exeter, Exeter, UK. Emails: {danyan_lan, pwang, liufuqiang}@tongji.edu.cn, {c.wang4, g.min}@exeter.ac.uk

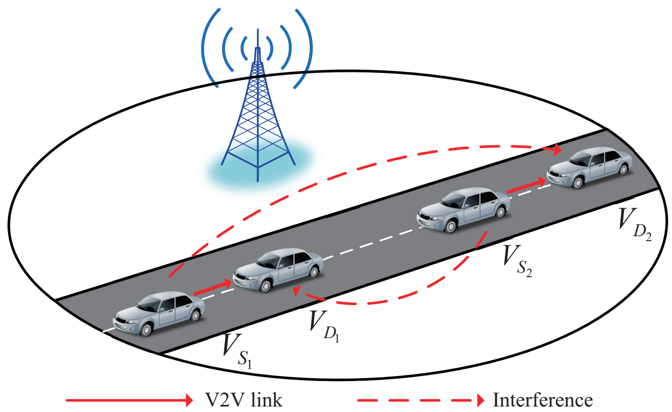

Fig. 1. System model: a vehicular communication network.

Wireless transmission design through power control, rate adaption, and channel assignment have been studied extensively [3]. Most conventional solutions are based on information theory and consider only the channel state information (CSI) in the physical (PHY) layer (e.g., [4]). Resources are allocated according to transmission opportunity, irrespective of urgency [5]. Hence transmissions are unaware of the delay performance and the resulting latency may be unlimited. To handle this issue, several recent works [6]–[9] have taken the queue state information (QSI) in the media access control (MAC) layer also into consideration. This allows a balance between transmission delay and other metrics such as power consumption and achievable rate. However, when dealing with multiple users, many works apply spectrally inefficient orthogonal channel sharing in order to avoid interference. More importantly, only one type of messages are considered. The solutions may not be sufficient in future vehicular networks. [10] investigates an interference channel where each user has multiple types of messages. The minimum power consumption is found to guarantee the desired QoS. But small power consumption does not necessarily mean efficient usage. Energy efficiency, which quantifies the amount of information successfully delivered by one Joule of energy, is a more important optimization objective in systems with rich information transmission demand but limited power budget.

In this paper, we intend to address this issue. We consider a two-user non-orthogonal vehicular communication network in which messages with different QoS requirements are delivered within each vehicle-to-vehicle (V2V) link: A heartbeat message should be transmitted immediately with high reliability, and a sensing message should be sent with finite delay. Making use of both CSI and QSI, a delay-aware cross-layer transmission design that aims to maximize energy efficiency is investigated. Since the optimization problem has a non-convex fractional objective and time-average constraints, we apply a series of simplifications, based on the Lyapunov theory, to transform it into a solvable problem. Simulation results show that our method notably outperforms conventional approaches.

The remainder of the paper is organized as follows. Section II presents the considered system model and problem formulation. The original problem is transformed to a solvable problem and then solved via power control in Sections III and IV, respectively. Simulation results are discussed in Section V. Finally, Section VI concludes the paper.

II. System Model

We consider a vehicular communication network in which two pairs of independent V2V links share the same spectrum band under the coordination of the serving base station, as shown in Fig. 1. We denote the source and destination vehicles of the ith ($i \in \{1, 2\}$) link by V_{S_i} and V_{D_i}, respectively. The messages delivered in each link intend to support two different types of delay-limited safety applications. The first provides V_{D_i} with the real-time status of V_{S_i}. Such information is carried by heartbeat messages generated periodically with fixed data rate. The second type of application is sensor data sharing from V_{S_i} for extending the environment perception capability of V_{D_i}. These messages arrive at V_{S_i} randomly.

Message generation and transmission are conducted in unit time slots. We term the heartbeat messages *type-1 messages*. The constant data rate is r_i bits/slot. These messages contain real-time information and should be transmitted immediately to avoid being outdated. To guarantee V_{D_i} to attain sufficient knowledge of the status of V_{S_i}, a transmission reliability requirement is placed. Use $\phi_i[t]=1$ to denote that at time slot t, V_{S_i} successfully sends the type-1 message to V_{D_i}. Otherwise, $\phi_i[t]=0$. It is desired that the time average of $\phi_i[t]$, i.e., $\overline{\phi}_i=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\phi_i[t]$ is greater than a certain minimum value $\phi_{\min,i}$, e.g., 0.9, for $i\in\{1,2\}$.

The sensing messages are termed *type-2 messages*. For the ith V2V link at time slot t the arrival data rate (in bits) is $a_i[t]$, which is in general modeled by a Poisson-distributed random variable with parameter λ_i. Different from type-1 messages, which must be delivered immediately, the type-2 messages can tolerate some delay. This means they can be temporarily placed in the source queues (denoted by $\mathcal{Q}_1$ and $\mathcal{Q}_2$) to wait for proper transmission opportunities, as long as the queuing delay is bounded to be finite. Let $Q_i[t]$ denote the queue length at V_{S_i} and $b_i[t]$ denote the rate that the type-2 messages leave the queue, at time slot t. The queuing dynamics of $\mathcal{Q}_i$ is

$$Q_i[t+1]=\max\{Q_i[t]-b_i[t],0\}+a_i[t],\ i\in\{1,2\}. \tag{1}$$

A finite queuing delay implies queues being stable [11].

The V2V communications are conducted in a Nakagami-m block-fading environment. The fading coefficient between V_{S_j} and V_{D_i} $(i,j\in\{1,2\})$ at time slot t is denoted by $h_{ij}[t]$, where $|h_{ij}[t]|\sim$ Nakagami(m_{ij},Ω_{ij}) with parameters m_{ij} and Ω_{ij}. $h_{ij}[t]$ remains fixed in each time slot, and changes independently across different slots. Each source V_{S_i} encodes its messages using unit-power capacity-achieving Gaussian random codes. The received signal at V_{D_i} is

$$y_i[t]=h_{ii}[t]\sqrt{p_i[t]}x_i[t]+h_{ij}[t]\sqrt{p_j[t]}x_j[t]+n_i[t], \tag{2}$$

where $x_i[t]$ and $x_j[t]$ denote the desired and interference signals from V_{S_i} and V_{S_j} respectively, and $n_i[t]$ denotes additive white Gaussian noise (AWGN) with power N_0. The transmit power of V_{S_i}, denoted by $p_i[t]$, is constrained by a maximum power limit $p_{\max,i}$. The achievable transmission data rate (in bits/slot) between V_{S_i} and V_{D_i} at time slot t is

$$R_i[t]=\log_2\left(1+\frac{|h_{ii}[t]|^2p_i[t]}{|h_{ij}[t]|^2p_j[t]+N_0}\right). \tag{3}$$

In each time slot, a source has two options to encode its messages (termed *encoding actions*). It can divide its data rate into two parts, dedicated to the two messages respectively: the type-1 message is transmitted with rate r_i and type-2 message, $b_i[t]=R_i[t]-r_i$. We have $\phi_i[t]=1$, which is possible only when $R_i[t]\geq r_i$. In addition, V_{S_i} can choose to send only the type-2 message, as long as the type-1 messages' reliability condition is still satisfied. This means $\phi_i[t]=0$ and $b_i[t]=R_i[t]$. We use $\sigma_i=1$ and $\sigma_i=2$ to denote these two encoding actions at V_{S_i} respectively. All source encoding actions are included in set $\mathcal{A}=\{A=(\sigma_1,\sigma_2)|\sigma_1,\sigma_2\in\{1,2\}\}$.

We aim to find an energy-efficient transmission strategy, such that the V2V links can choose their encoding actions and transmission powers based on the knowledge of both CSI and QSI, to satisfy the desired message delivery requirements with maximized *energy efficiency* $\bar{\eta}$. Specifically, $\bar{\eta}$ is defined as the ratio of average data rate to average power consumption:

$$\bar{\eta}\triangleq\frac{\overline{R}}{\overline{P}}=\frac{\sum_{i=1}^{2}\alpha_i\overline{R}_i}{\sum_{i=1}^{2}\beta_i\overline{p}_i}, \tag{4}$$

where the individual time average transmission rate $\overline{R}_i=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}R_i[t]$, individual time average power usage $\overline{p}_i=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}p_i[t]$, and the weighting coefficients $0\leq\alpha_1\leq1$, $\alpha_2=1-\alpha_1$, $0\leq\beta_1\leq1$, and $\beta_2=1-\beta_1$ are determined to represent relative importance.

Therefore, we aim to solve the optimization problem:

$$\begin{aligned} \text{maximize}: \quad & \overline{\eta} \\ \text{s.t.: } & \text{C1}: \overline{\phi}_i \geq \phi_{\text{min},i}, i \in \{1,2\}, \\ & \text{C2}: \mathcal{Q}_i \text{ is stable}, i \in \{1,2\}, \\ & \text{C3}: 0 \leq p_i[t] \leq p_{\text{max},i}, i \in \{1,2\}. \end{aligned} \tag{5}$$

The constraint C1 represents the high-reliability requirement of the type-1 messages, C2 represents the finite-delay requirement of the type-2 messages, and C3 represents the sources' transmit power limits. Note that an additional condition is that the sources choose their encoding actions from $\mathcal{A}$.

In fact, the above problem may not have any feasible solution, if $p_{\text{max},1}$ or $p_{\text{max},2}$ is small. Setting $\alpha_1 = \alpha_2 = 0$ leads to a feasibility check problem. In this paper, to make the considered transmission design meaningful, we assume that $p_{\text{max},1}$ and $p_{\text{max},2}$ are sufficiently large such that feasible solution exists. However, even in this case, finding the solution of problem (5) is involved because: i) the objective function is a non-linear fractional function, and ii) both the objective function and constraints contain time-average operations. In what follows, we will use a series of steps to transform the original problem (5) into an optimization problem that can be solved in each individual time slot.

III. Optimization problem transformation

A. *Transformation of objective function*

We start from the objective function. Use $\boldsymbol{p} = [p_1[1], p_2[1], p_1[2], p_2[2], \cdots]$ to denote a power allocation vector. Let $\boldsymbol{p}^*$ denote the solution that achieves the optimal value $\overline{\eta}_{opt}$. Following [12], we have

$$\overline{\eta}_{opt} = \frac{\overline{R}(\boldsymbol{p}^*)}{\overline{P}(\boldsymbol{p}^*)} \geq \frac{\overline{R}(\boldsymbol{p})}{\overline{P}(\boldsymbol{p})}. \tag{6}$$

Rearranging this inequality leads to

$$\overline{R}(\boldsymbol{p}^*) - \overline{\eta}_{opt}\overline{P}(\boldsymbol{p}^*) = 0 \geq \overline{R}(\boldsymbol{p}) - \overline{\eta}_{opt}\overline{P}(\boldsymbol{p}).$$

Conversely, if the optimization objective is given by

$$\overline{R}(\boldsymbol{p}) - \overline{\eta}_{opt}\overline{P}(\boldsymbol{p}), \tag{7}$$

within the feasibility region, the inequality $0 = \overline{R}(\boldsymbol{p}^+) - \overline{\eta}_{opt}\overline{P}(\boldsymbol{p}^+) \geq \overline{R}(\boldsymbol{p}) - \overline{\eta}_{opt}\overline{P}(\boldsymbol{p})$ is satisfied, where $\boldsymbol{p}^+$ is the optimal solution. Then we obtain

$$\overline{\eta}_{opt} = \frac{\overline{R}(\boldsymbol{p}^+)}{\overline{P}(\boldsymbol{p}^+)} \geq \frac{\overline{R}(\boldsymbol{p})}{\overline{P}(\boldsymbol{p})}.$$

$\boldsymbol{p}^+$ is also the optimal solution of the problem (6). Following [12], the problem (5) can be transformed to an equivalent optimization problem:

$$\begin{aligned} \text{maximize}: \ & \overline{R}(\boldsymbol{p}) - \overline{\eta}_{\text{opt}}\overline{P}(\boldsymbol{p}) \\ & \text{s.t.: C1, C2, C3.} \end{aligned} \tag{8}$$

Further, we define a time function $\eta[t]$, with $\eta[1] = 0$ and

$$\eta[t] = \frac{\sum_{\tau=1}^{t-1} \left(\alpha_1 R_1[\tau] + \alpha_2 R_2[\tau]\right)}{\sum_{\tau=1}^{t-1} \left(\beta_1 p_1[\tau] + \beta_2 p_2[\tau]\right)}. \tag{9}$$

Since $\eta[t]$ depends only on the transmission rates and power consumption in the past time instants, it is a constant at time slot t. The optimization problem (8) is thus transformed to:

$$\begin{aligned}\text{maximize}:\ &\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{2}(\alpha_i R_i[t]-\beta_i\eta[t]p_i[t]) \\ \text{s.t.: C1, C2, C3.}\end{aligned} \tag{10}$$

At each time slot, the objective function is no longer a non-linear fractional function. In the next subsection, we will focus on handling the time average operations.

B. Transformation of time average operations

To remove the time-average operation in the optimization constraint C1, we follow [11] and first transform it into an equivalent queue stability constraint. Specifically, define virtual reliability queues $\mathcal{Y}_1$ and $\mathcal{Y}_2$. Use $Y_i[t]$ to represent the queue length of $\mathcal{Y}_i$ at time slot t. Set $Y_i[1]=0$ and

$$Y_i[t+1]=\max\{Y_i[t]-\phi_i[t],0\}+\phi_{\min,i},\ \ i\in\{1,2\}. \tag{11}$$

The condition that the queues $\mathcal{Y}_i$ are stable implies $\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\phi_i[t]\geq\phi_{\min,i}$, i.e. constraint C1. Thus in (10), C1 and C2 can be combined to form a new constraint

$$\text{C4}:\mathcal{Q}_i \text{ and } \mathcal{Y}_i \text{ are stable}, i\in\{1,2\}.$$

Next, we apply the Lyapunov optimization theory [11] to further simplify the optimization problem. Let $\mathbf{\Omega}[t]=\{Q_1[t],Q_2[t],Y_1[t],Y_2[t]\}$ denote the set of current queue backlogs at time slot t, and define the *quadratic Lyapunov function* $L(\mathbf{\Omega}[t])$ as a function of $\mathbf{\Omega}[t]$:

$$L(\mathbf{\Omega}[t])=\sum_{i=1}^{2}\left(u_iQ_i[t]^2+v_iY_i[t]^2\right), \tag{12}$$

in which u_i and v_i are constant weighting parameters. Further, define the *one-slot conditional Lyapunov drift function* as the expected (taken over transmission designs for dealing with random channel variations) change of the Lyapunov function over one time slot, given the knowledge of current QSI:

$$\mathbf{\Delta}\left(\mathbf{\Omega}[t]\right)=\mathbb{E}\left\{L(\mathbf{\Omega}[t+1])-L(\mathbf{\Omega}[t])|\mathbf{\Omega}[t]\right\}. \tag{13}$$

A negative value of $\mathbf{\Delta}(\mathbf{\Omega}[t])$ indicates that the sum queue length tends to reduce at the next time slot.

The optimization problem together with queue stability constraint C4 can be transformed into a *drift-plus-penalty* expression [11]. In the considered problem, the maximization of (10) subjects to C3 and C4 can be transformed to

$$\begin{aligned}\text{maximize}:&V\mathbb{E}\left\{\sum_{i=1}^{2}(\alpha_iR_i[t]-\beta_i\eta[t]p_i[t])|\mathbf{\Omega}[t]\right\}-\mathbf{\Delta}(\mathbf{\Omega}[t]) \\ \text{s.t.: C3.}\end{aligned} \tag{14}$$

The positive parameter V can be tuned to reveal the desired optimization tradeoff between achievable energy efficiency and queue lengths (both actual source queues $\mathcal{Q}_i$ and virtual queues $\mathcal{Y}_i$). Setting a large value of V means that we tend to put more emphasis on increasing system energy efficiency, with the cost of small transmission reliability of the type-1 messages (but still greater than the threshold $\phi_{\min,i}$) and large queuing delay of the type-2 messages (but still bounded).

Since $\mathbf{\Delta}(\mathbf{\Omega}[t])$ cannot be expressed explicitly, we use its upper bound to replace $\mathbf{\Delta}(\mathbf{\Omega}[t])$ in the optimization problem. Substituting the queuing dynamics of $\mathcal{Q}_i$ and $\mathcal{Y}_i$ into (13) and considering the inequality $(\max\{Q-b,0\}+c)^2 \leq Q^2+b^2+c^2+2Qc-2Qb$ for $Q\geq 0, b\geq 0, c\geq 0$, we can show that $\mathbf{\Delta}(\mathbf{\Omega}[t])$ can be upper-bounded as

$$\begin{aligned}\mathbf{\Delta}(\mathbf{\Omega}[t]) \leq & \mathbb{E}\left\{\sum_{i=1}^{2} u_i(b_i^2[t]+a_i^2[t]+2Q_i[t](a_i[t]-b_i[t]))\Big|\mathbf{\Omega}[t]\right\} \\ & +\mathbb{E}\left\{\sum_{i=1}^{2} v_i(\phi_i^2[t]+\phi_{\text{min},i}^2+2Y_i[t](\phi_{\text{min},i}-\phi_i[t]))\Big|\mathbf{\Omega}[t]\right\} \\ \leq & B-2\sum_{i=1}^{2} u_iQ_i[t](\mathbb{E}\{b_i[t]|\mathbf{\Omega}[t]\}-\lambda_i)-2\sum_{i=1}^{2} v_iY_i[t](\mathbb{E}\{\phi_i[t]|\mathbf{\Omega}[t]\}-\phi_{\text{min},i}),\end{aligned} \tag{15}$$

where $B=\sum_{i=1}^{2}(u_i\log_2(2^{1/ln2}+\mathbb{E}\{|h_{ii}[t]|^2\}p_{\text{max,i}}/N_0)+u_ia_{\text{max},i}^2+v_i\phi_{\text{min},i}^2)+v_i$ is a positive constant, and $a_{\text{max},i}$ is the maximum value of $a_i[t]$. Now we can use the last upper bound in (15) to replace $\mathbf{\Delta}(\mathbf{\Omega}[t])$ in (14). Notice that the parameters B, λ_i, $\phi_{\text{min},i}$ ($i\in\{1,2\}$) are constants. They are irrelevant to our power control design and can hence be directly discarded.

Finally, we replace the problem of solving optimization with expectations by opportunistically maximizing the objective function at each individual time slot [11]. To simplify presentation, in the remainder of the paper, we omit the time index t. The original problem (5) becomes the following form:

$$\begin{aligned}&\text{maximize}:\ V\sum_{i=1}^{2}(\alpha_iR_i-\beta_i\eta p_i)+2\sum_{i=1}^{2}(u_iQ_ib_i+v_iY_i\phi_i) \\ &\text{s.t.:}\ 0\leq p_i[t]\leq p_{\text{max},i}, i\in\{1,2\}.\end{aligned} \tag{16}$$

So far, we have transformed (5) into a problem which can be solved in each individual time slot. However, finding the solution of (16) is still challenging since different encoding actions in $\mathcal{A}$ lead to different forms of the problem, each of which has a particular non-convex objective function and feasibility region. To address this issue, in the next section we individually discuss each encoding action. It can be shown that for each action $A\in\mathcal{A}$, the corresponding optimization problem can be approximated by a solvable concave optimization problem. The overall solution can thus be found by finding the best one among all encoding actions.

IV. Power control design at each time slot

As discussed in Section II, there are four encoding actions in $\mathcal{A}$. For each individual action, the objective function is non-convex due to inter-user interference caused by non-orthogonal transmission. Let $\boldsymbol{p}=[p_1,p_2]$ denote the transmit power vector and $\mathcal{P}_A$ denote the feasibility region when encoding action $A\in\mathcal{A}$ is adopted. When the problem is feasible, applying the *difference of two convex functions programming* (D.C. programming) [13], we can rewrite the problem (16) as

$$\begin{aligned}&\text{maximize}:\ \gamma_A\triangleq f_A(\boldsymbol{p})-g_A(\boldsymbol{p}) \\ &\text{s.t.:}\ \boldsymbol{p}\in\mathcal{P}_A,\end{aligned} \tag{17}$$

where γ_A represents the objective function, and $f_A(\boldsymbol{p})$ and $g_A(\boldsymbol{p})$ are functions determined by the chosen action A.

According to Taylor's expansion, $g_A(\boldsymbol{p}^{(k)})+\nabla g_A(\boldsymbol{p}^{(k)})^T(\boldsymbol{p}-\boldsymbol{p}^{(k)})$ is the first-order expansion of $g_A(\boldsymbol{p})$ at point $\boldsymbol{p}^{(k)}$, where $\boldsymbol{p}^{(k)}$ denotes the optimal solution at the $(k-1)$th iteration. Considering its concave characteristics [14], it is always larger than $g_A(\boldsymbol{p})$. Maximizing $-g_A(\boldsymbol{p})$ is equivalent to minimizing the upper bound of $g_A(\boldsymbol{p})$. As long as $f_A(\boldsymbol{p})$, $g_A(\boldsymbol{p})$, $\boldsymbol{p}\in\mathcal{P}_A$ can be shown to be concave, the optimization problem (17) can be approximated by the following concave optimization problem, and then solved according to the method in [13]:

$$\begin{aligned} \text{maximize}:\ & f_A(\boldsymbol{p}) - g_A(\boldsymbol{p}^{(k)}) - \nabla g_A(\boldsymbol{p}^{(k)})^T(\boldsymbol{p} - \boldsymbol{p}^{(k)}) \\ \text{s.t.:}\ & \boldsymbol{p} \in \mathcal{P}_A, \end{aligned} \tag{18}$$

Certainly, for an encoding action A, it is possible that due to the random nature of fading and limited power budgets, the feasibility region $\mathcal{P}_A$ is an empty set. The problem (16) for this action does not have feasible solution (but by considering all four actions, the problem is always feasible). If this is the case, we can directly set $p_1 = p_2 = 0$ and set the objective function γ_A to be a small negative value. In what follows, we will present $f_A(\boldsymbol{p})$, $g_A(\boldsymbol{p})$, $\boldsymbol{p} \in \mathcal{P}_A$ for each action in $\mathcal{A}$.

A. *Encoding action* $A = (1, 1)$

In this case, both sources intend to deliver their type-1 messages. Recall that the type-1 message for V_{D_i} has fixed rate r_i. Hence to guarantee successful transmission, we need to have $r_1 \le R_1$ and $r_2 \le R_2$. The remaining part of transmission rate $b_i = R_i - r_i$ is left for the type-2 message. After some mathematical manipulations, we can show that in (17) the functions $f_A(\boldsymbol{p})$ and $g_A(\boldsymbol{p})$ are both concave:

$$\begin{aligned} f_A(\boldsymbol{p}) = &(V\alpha_1 + 2u_1Q_1)\log_2(N_0 + |h_{11}|^2 p_1 + |h_{12}|^2 p_2) \\ &+ (V\alpha_2 + 2u_2Q_2)\log_2(N_0 + |h_{21}|^2 p_1 + |h_{22}|^2 p_2) \\ &- \sum_{i=1}^{2}\left[\beta_i V\eta p_i + 2u_iQ_ir_i - 2v_iY_i\right], \end{aligned}$$

$$\begin{aligned} g_A(\boldsymbol{p}) = &(V\alpha_1 + 2u_1Q_1)\log_2(N_0 + |h_{12}|^2 p_2) \\ &+ (V\alpha_2 + 2u_2Q_2)\log_2(N_0 + |h_{21}|^2 p_1). \end{aligned}$$

The conditions $r_1 \le R_1$ and $r_2 \le R_2$ lead to

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \succeq \begin{bmatrix} 0 & \frac{|h_{12}|^2(2^{r_1}-1)}{|h_{11}|^2} \\ \frac{|h_{21}|^2(2^{r_2}-1)}{|h_{22}|^2} & 0 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} + \begin{bmatrix} \frac{(2^{r_1}-1)N_0}{|h_{11}|^2} \\ \frac{(2^{r_2}-1)N_0}{|h_{22}|^2} \end{bmatrix}. \tag{19}$$

To ensure a non-empty feasibility region, the following condition must also satisfy (i.e., the maximum power budget must be greater than the minimum power needed in (19))

$$p_{\max,i} \ge \frac{N_0\left((2^{r_i}-1)|h_{ji}|^2/|h_{ii}|^2 + k\right)}{|h_{ji}|^2(1-k)},\ \ i \ne j, \tag{20}$$

where $k = (2^{r_1}-1)(2^{r_2}-1)\frac{|h_{12}|^2|h_{21}|^2}{|h_{11}|^2|h_{22}|^2}$ [4]. Therefore, the feasibility region $\mathcal{P}_A$ is the union of constraints (19) and (20), which are both convex. If $\mathcal{P}_A$ is not empty, (17) can be solved using (18). Otherwise, as mentioned before, if A leads to an infeasible power control problem, we set $p_1 = p_2 = 0$.

B. *Encoding action* $A = (1, 2)$

Under this encoding action, only V_{S_1} intends to transmit its type-1 message, which requires $R_1 \ge r_1$. The type-2 messages can be transmitted in the two V2V links with rates $b_1 = R_1 - r_1$ and $b_2 = R_2$, respectively. Setting $\phi_1 = 1$ and $\phi_2 = 0$ in (16), we can attain the following concave functions:

$$\begin{aligned} f_A(\boldsymbol{p}) = &(V\alpha_1 + 2u_1Q_1)\log_2(N_0 + |h_{11}|^2 p_1 + |h_{12}|^2 p_2) \\ &+ (V\alpha_2 + 2u_2Q_2)\log_2(N_0 + |h_{21}|^2 p_1 + |h_{22}|^2 p_2) \\ &- 2u_1Q_1r_1 + 2v_1Y_1 - \sum_{i=1}^{2}\beta_i V\eta p_i, \end{aligned}$$

$$g_A(\boldsymbol{p}) = (V\alpha_1 + 2u_1Q_1)\log_2(N_0 + |h_{12}|^2 p_2) \\ + (V\alpha_2 + 2u_2Q_2)\log_2(N_0 + |h_{21}|^2 p_1).$$

The power limit of V_{S_1} must be sufficiently large to ensure the existence of feasible solution. To this end, we have

$$\frac{(2^{r_1}-1)|h_{12}|^2}{|h_{11}|^2} p_2 + \frac{(2^{r_1}-1)N_0}{|h_{11}|^2} \leq p_1 \leq p_{\max,1}. \tag{21}$$

The conditions $0 \leq p_2 \leq p_{\max,2}$ and (21) form the feasibility region $\mathcal{P}_A$. Both conditions are concave. The problem (17) can be solved using (18) if $\mathcal{P}_A$ is not empty.

C. *Encoding action* $A = (2, 1)$

Now only V_{S_2} desires to send its type-1 message. The analysis is similar to the above case. We can straightforwardly attain $f_A(\boldsymbol{p})$, $g_A(\boldsymbol{p})$, and $\mathcal{P}_A$ by swapping the indexes 1 and 2 in the results provided in the above subsection.

D. *Encoding action* $A = (2, 2)$

In this case, both sources send only type-2 messages with $b_1 = R_1$ and $b_2 = R_2$. Setting $\phi_1 = \phi_2 = 0$ in (16) leads to

$$f_A(\boldsymbol{p}) = (V\alpha_1 + 2u_1Q_1)\log_2(N_0 + |h_{11}|^2 p_1 + |h_{12}|^2 p_2) \\ + (V\alpha_2 + 2u_2Q_2)\log_2(N_0 + |h_{21}|^2 p_1 + |h_{22}|^2 p_2) - \sum_{i=1}^{2} \beta_i V \eta p_i,$$

$$g_A(\boldsymbol{p}) = (V\alpha_1 + 2u_1Q_1)\log_2(N_0 + |h_{12}|^2 p_2) \\ + (V\alpha_2 + 2u_2Q_2)\log_2(N_0 + |h_{21}|^2 p_1).$$

The problem is always feasible. The condition C3 defines $\mathcal{P}_A$.

E. *Control action selection*

We have presented the impact of every individual encoding action on the power control design at each time slot. For action A, the power vector that leads to the optimal value of objective function γ_A can be found (a small negative value of γ_A for infeasible problem). Then one can compare the objective functions and select the encoding action and power vector that have the highest γ_A. In other words, at each time slot t, power control in the considered vehicular communication network is as follows:

$$\boldsymbol{p}[t] = \arg\max_{A \in \mathcal{A}} \gamma_A[t]. \tag{22}$$

In the next section, we will use simulation results to demonstrate the performance of our transmission design.

V. Performance evaluation

We take the application scenario described in Section I as an example. Each pair of source and destination vehicles form a platoon. The leading vehicle V_{S_i} periodically shares its speed and location information (type-1 messages), as well as sensing results of certain objects of interest (type-2 messages) with the following vehicle V_{D_i}. The inter-vehicle distance within each platoon is roughly 10 m. The platoons are separated by about 80 m, which means that the interfering distance between V_{S_1} and V_{D_2} is 100 m and that between V_{S_2} and V_{D_1} is 80 m.

In our simulations, the Nakagami-m fading parameters m_{ij} are set to 2 for all $i, j \in \{1, 2\}$, and Ω_{ij} is determined by considering V2V path loss $PL_{ij} = 103.4 + 24.2 \log_{10}(d_{ij})$ [15], where d_{ij} denotes distance. The source power limits are $p_{\max,i} = -85$ dBm, and the noise power spectral density is -174 dBm/Hz with transmission bandwidth 100 kHz. In this section, we consider a symmetric system, such that both V2V links have the same performance requirements and transmission parameters. The type-1 messages' data rates are assumed to be $r_1 = r_2 = 0.5$. Their transmission reliability requirements are $\phi_{\text{min},1} = \phi_{\text{min},2} = 90\%$. The type-2 messages are generated with average rate $\lambda_1 = \lambda_2 = 2$. The weights in energy efficiency evaluation are set to $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0.5$, $u_1 = u_2 = 1.1$, and $v_1 = v_2 = 3$. The system model can be straightforwardly generalized to be asymmetric.

We compare our transmission design with three benchmark approaches. The first performs design using only CSI (we term it *CSI-based* design). Specifically, it adopts the power control strategy proposed in [4] and tries to use the minimum power to guarantee the fixed transmission rates of $R_1[t] = R_2[t] = 2.5 + \theta$ bits/slot. The powers consumed for supporting 0.5 bits and $2 + \theta$ bits are dedicated respectively to the type-1 and type-2 messages, where the parameter θ is chosen to counterbalance the impact of channel outage and to maintain the queue lengths to be similar to our method (if $\theta = 0$, the lengths of queues $\mathcal{Q}_i$ will increase with time).

The remaining two benchmark methods take into account both CSI and QSI. The second method is termed *power minimization-based (PM-based)* design. It follows the approach presented in [10] and tends to use the minimum power to guarantee the performance of the two types of messages. The final method is termed *orthogonal transmission-based (OT-based)* design, where the two V2V links orthogonally share the available channel. The maximum transmission data rate between V_{S_i} and V_{D_i} is $R_i[t] = \frac{1}{2}\log_2(1 + \frac{2p_i[t]|h_{ii}[t]|^2}{N_0})$ bits/slot. We can follow exactly the same procedure presented earlier to formulate and transform the optimization problem to maximize energy efficiency (omitted due to page limit). Inter-user interference is avoided, with the cost of inefficient channel usage (because of the pre-log scaling factor $\frac{1}{2}$).

Fig. 2 shows the achievable energy efficiency comparison of our method (termed *EE-based* design) with the above three approaches, when they all attain the similar performance of the two types of messages, as shown in Fig. 3. Clearly, our method achieves the highest energy efficiency, because it properly utilizes the knowledge of both CSI and QSI, adopts efficient non-orthogonal channel sharing, and uses energy efficiency to direct its transmission and power control design. Due to the lack of QSI (i.e., unaware of delay), the CSI-based design tries to guarantee the QoS requirements of the two types of messages using a naive fashion. To ensure both types of messages to be transmitted successfully, it demands a very high data rate at every time slot, even when the channel is weak, the reliability of type-1 message is already high, and/or the queue length for the type-2 message is already small. This can cause high probability of channel outage when the power budget is relatively small. Even if $p_{\max,i}$ are sufficiently large, the method also leads to unnecessarily large reliability of the type-1 messages (see Fig. 3) and low energy efficiency. The PM-based design intends to minimize the power consumption of the whole system. But a low power usage does not always mean a high energy efficiency. Finally, the OT-based design suffers from the inefficient orthogonal transmission, because in the considered system the mutual interference between the two V2V links is not very large and orthogonally activating the sources is not necessary. The tradeoff between energy efficiency and message transmission performance in our proposed solution can be further flexibly balanced by adjusting the parameters V, α_i and β_i in (4), and u_i and v_i in (12), according to application demands.

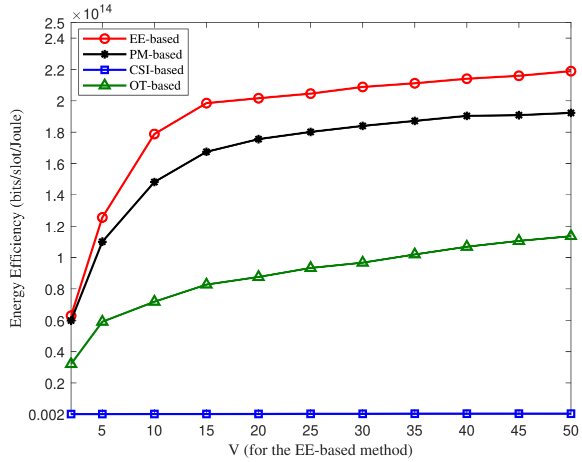

Fig. 2. Achievable energy efficiency versus V.

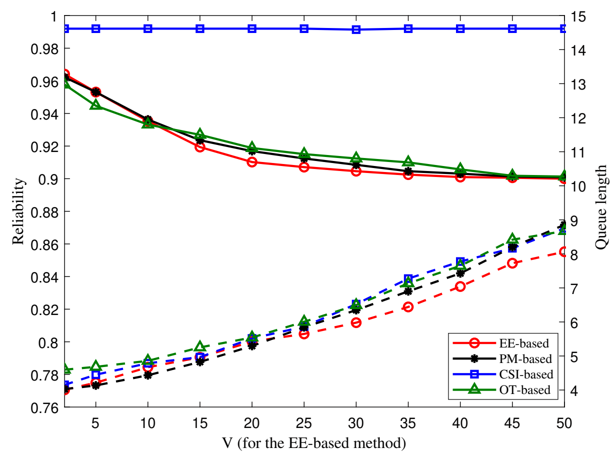

Fig. 3. Average reliability of type-1 messages (solid lines) and average queue lengths of type-2 messages (dashed lines).

VI. Conclusion

We have proposed a transmission strategy that allocates powers in a class of vehicular communication networks to achieve high energy efficiency. The networks contain concurrent V2V links, each of which desires to deliver messages to support different types of delay-limited applications. Both CSI in the PHY layer and QSI in the MAC layer are used to direct the transmission design so that transmission opportunity and urgency can be properly balanced. The advantages of the proposed method over serveral conventional solutions have been verified by simulations. The impact of imperfect CSI and QSI, and distributed decision making, through e.g., reinforcement learning, are deemed as meaningful future works.

Acknowledgment

This work was funded in part by the National Natural Science Foundation of China (61771343). This is also a part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 752979. This work reflects only the authors' view and the EU Commission is not responsible for any use that may be made of the information it contains. C. Wang is the correspondence author.

REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jan. 2014.

[2] E. Lee, E.-K. Lee, M. Gerla, and S. Y. Oh, "Vehicular cloud networking: architecture and design principles," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 148–155, Feb. 2014.

[3] S. Stańczak, M. Wiczanowski, and H. Boche, *Fundamentals of Resource Allocation in Wireless Networks: Theory and Algorithms*. Springer Science & Business Media, 2009.

[4] H. Farhadi, C. Wang, and M. Skoglund, "Fixed-rate transmission over fading interference channels using point-to-point Gaussian codes," *IEEE Trans. Commun.*, vol. 63, no. 10, pp. 3633–3644, Oct. 2015.

[5] Y. Cui, V. K. N. Lau, R. Wang, H. Huang, and S. Zhang, "A survey on delay-aware resource control for wireless systems — large deviation theory, stochastic Lyapunov drift and distributed stochastic learning," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1677–1701, Mar. 2012.

[6] S. Huang, B. Liang, and J. Li, "Distributed interference and delay aware design for D2D communication in large wireless networks with adaptive interference estimation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3924–3939, Jun. 2017.

[7] X. Chen, W. Chen, J. Lee, and N. B. Shroff, "Delay-optimal buffer-aware scheduling with adaptive transmission," *IEEE Trans. Commun.*, vol. 65, no. 7, pp. 2917–2930, Jul. 2017.

[8] M. Sheng, Y. Li, X. Wang, J. Li, and Y. Shi, "Energy efficiency and delay tradeoff in device-to-device communications underlaying cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 1, pp. 92–106, Jan. 2016.

[9] B. Yang, Y. Shen, Q. Han, C. Chen, X. Guan, and W. Zhang, "Energy-efficient resource allocation for time-varying ofdma relay systems with hybrid energy supplies," *IEEE Syst. J.*, vol. 12, pp. 702–713, Mar. 2018.

[10] M. Li, P. Wang, C. Wang, F. Liu, and Y. Ji, "Cross-layer transmission design in a class of delay-aware two-user interference networks," *Int. J. Syst. Control Inf. Process.*, vol. 2, no. 3, pp. 191–218, Jun. 2018.

[11] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queuing Systems*. Morgan & Claypool, 2010.

[12] W. Dinkelbach, "On nonlinear fractional programming," *Manag. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967.

[13] H. H. Kha, H. D. Tuan, and H. H. Nguyen, "Fast global optimal power allocation in wireless networks by local d.c. programming," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 510–515, Feb. 2012.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ. Press, 2004.

[15] *Further advancements for E-UTRA physical layer aspects (Release 9)*, 3GPP TR 36.814 Std., 2009.