Kilian Kiekenap, Andrea Ortiz, and Anja Klein, ”Energy-Optimal Short Packet Transmission for Time-Critical Control” in *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall) 2021*, September 2021.

# Energy-Optimal Short Packet Transmission for Time-Critical Control

Kilian Kiekenap, Andrea Ortiz and Anja Klein
Communications Engineering Lab, TU Darmstadt, Landgraf-Georg-Str. 4, 64283 Darmstadt, Germany
Email: {k.kiekenap, a.ortiz, a.klein}@nt.tu-darmstadt.de

***Abstract*—In this paper, the transmission energy for reliable communications with short packets and low latency requirements, e.g. for control applications, is minimized. Since the dynamics of the agents determine the allowed latencies for receiving control inputs, the requirements on latency and allowable packet error rate are individual, depending on the machine type. We consider a centralized environment with a single controller transmitting control commands wireless to multiple agents with given latency requirements. Also, the channel conditions are individual for each agent. Therefore, the optimal time-frequency resource allocation is derived for continuous time-frequency resource allocation. Since the resource allocation in OFDM systems like 5G is discrete, an algorithm to select the allocation from a resource grid with different resolutions is proposed and shown to achieve solutions with less than 0.5 dB increase in energy consumption compared to the continuous results. With numerical evaluation, the benefit of a channel-state- and deadline-aware solution is shown for a resource grid based on the 5G frame structure. On average, the gain of the proposed algorithm to an allocation only balancing the number of resources for each agent, as far as the deadlines allow, is about 50% energy saving.**

***Index Terms*—Communication and Control, Cyber-Physical Systems, URLLC, 5G**

## I. Introduction

Recent developments in industrial automation introduce wireless communication to production facilities for e.g. real-time monitoring or process control [1]. Currently, most Internet of Things (IoT) devices serve as home appliances, building control or environmental sensors [2]. In IoT for industrial manufacturing, called Industrial Internet of Things (IIoT), devices are part of industrial production processes and, therefore, directly embedded into control loops, which impose different demands on wireless communication systems compared to previous applications like voice, web browsing and video streaming. In control, the data amounts are small, in the order of tenths or a few hundred bytes, while the constraints on latency and packet error rate are even tighter than in other fields. Packet error rates as low as $10^{-9}$ and latencies of 0.25 to 5 ms are required [3]. The new 5G mobile radio standard is the first to define requirement profiles for these use cases [4].

We consider a model with a single central controller, which observes multiple agents and generates control commands for them, which are then transmitted via a wireless link. To maintain the requirements of the control process, each command has to be delivered successfully with a maximum latency and packet error rate. To achieve energy-optimal communication while maintaining the error rate and latency requirements, the allocation of time-frequency resources has to be adapted to the individual wireless channel conditions and transmission deadlines resulting from the control perspective. The capacity of a communication channel for infinite time-frequency resources according to Shannon is determined by the Signal-to-noise ratio (SNR). This is a valid assumption for transmissions of large amounts of data over infinite time-frequency resources, but since IIoT is especially about short packets and low latencies, this estimation is way too optimistic. To have a more realistic estimation for Ultra-Reliable Low Latency Communication (URLLC), Polyanskiy et al. derived a short packet formula, which gives a tighter bound on the achievable data rate given the available time-frequency resources, the SNR at the receiver and the allowable packet error rate in their seminal paper [5].

Since the allowable packet error rate and number of time-frequency resources are limited, the SNR at the receiver, which is determined by the transmit power and the channel gain, has to be tuned to meet the requirements. To increase the SNR, the transmit power has to be increased, which results in an increase of total energy consumption.

We will now give a brief overview of related work. In [6], a system with a single controller and agent was investigated. The main focus was to find the minimum SNR to maintain stability. The controller sends control commands over a wireless link to the agent. The effects of quantization to discrete commands and packet loss due to the short packet effect were considered. Further control requirements leading to latency restrictions were not considered. In [7], a power spectral density minimization was investigated. A limited number of time-frequency resources was split between an initial transmit and a potential retransmit, potentially saving energy, if no retransmission is needed. A maximum probability of a buffer underrun at the receiver had to be reached, thus including latency requirements. In this scenario, all available time-frequency resources are used for the single agent, while the multi-agent scenario is not considered. A multi-agent bandwidth minimization problem was studied in [8]. The channel was assumed to be frequency selective, but only known to the receiver. The transmit power per resource element was fixed. [9] maximized the energy efficiency in a scenario with multiple sensors transmitting data to agents. Sensors and agents were assumed to be in different mobile radio cells, the transmissions were done in URLLC style. The deadlines for these transmissions were

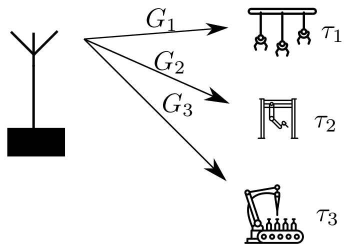


Fig. 1: System model

assumed identical, which is unrealistic for an industrial plant with agents belonging to different classes of machines.

In this paper, a scenario with a single central controller and multiple agents is considered. The controller can sense the states of the agents, generates control commands according to the states and transmits them to the agents via a wireless link. The control commands are assumed to be short data packets of up to a few hundred bits in size. In contrast to the state of the art of [9], we consider the different dynamics of the various types of machines by means of the definition of agent-specific deadlines. Additionally, the maximum allowable packet error rate is constrained to a low constant value to account for the safety requirements of industrial production plants. The available bandwidth for transmission is limited. Under these constraints, we find the optimal time-frequency resource allocation to the agents minimizing the required energy for transmission.

For this purpose, we first formulate a problem with a continuous amount of resources for each agent in Sec. III. The given agent-specific maximum latencies lead to deadlines, when the transmission has to be finished the latest. The channel conditions are also given, as well as the common maximum packet error rate. This problem is shown to be convex. Then, we propose a gradient-based algorithm in Sec. IV to allocate the time-frequency resources in an Orthogonal Frequency Division Multiplex (OFDM) scheme in a quantized fashion. For comparison, an allocation balancing the number of resources for each agent, as far as the deadlines allow, is calculated. The three approaches are compared in Sec. V and the gradient based allocation is shown to be close to the continuous lower bound. Moreover, the balancing allocation of resources to all agents is shown to perform worse than the gradient-based algorithm.

The remainder of the paper is organized as follows: First, the system model is introduced. Then, the energy minimization problem is stated and its convexity is shown in Sec. III. In Sec. IV, the resource allocation algorithms are presented. Last, numerical results are shown to illustrate the performance of our proposed algorithm.

## II. System Model

The system consists of a single central controller and $M$ agents randomly distributed around the controller, as shown in Fig. 1. The central controller senses the control system states of all agents and generates control commands accordingly, which are then transmitted to the agents. The control system is assumed to be discrete-time with a time slot duration $T$. For each time slot a new control command is generated for every agent, the commands are all available at the beginning of the time slot. The time elapsed since the beginning of the time slot is denoted by $t, 0 \leq t \leq T$, $t = 0$ indicates where transmission starts. The performance of the control system is determined by the latency of the control commands, so each agent $m$ has an individual deadline $\tau_m, 0 \leq \tau_m \leq T$ for the successful reception of its command after the beginning of the time slot. The value of $\tau_m$ depends on the dynamics of agent $m$, where higher dynamics generally lead to shorter deadlines. Allocating resources to agent $m$ after its deadline $\tau_m$ has passed would not contribute to a timely reception, so we assume no resources after the deadline are allocated. Additionally, the probability of a lost control command must not exceed $p_c$ to keep the agents in a safe operation region.

Throughout this paper, a continuous quantity $x$ will be denoted by $x'(t)$, while its piecewise continuous counterpart will be denoted by $x_t$. The total bandwidth available for transmission is denoted by $B$. The time-variant bandwidth assigned to agent $m$ at time $t$ is $b'_m(t) \leq B$. $b'_m(t)$ is assumed to fulfill the uncertainty principle, i.e. it does not change arbitrarily fast. Moreover, the sum of all assignments must not exceed the total bandwidth, i.e.

$$\sum_{m=1}^{M} b'_m(t) \leq B \text{ for } 0 \leq t \leq T. \tag{1}$$

The commands for each agent, consisting of $N$ bits, are transmitted over a wireless channel, which is perfectly known at the central controller and the receiving agents. The agents are assumed to be stationary. The channel between the controller and every agent is modelled as line-of-sight (LOS). Thus, the channel is assumed to stay constant over $T$ and $B$. The power gain of the channel from the central controller to agent $m$ is denoted by the scalar channel gain $G_m$.

The transmission is performed interference free by using frequency division multiple access on the available bandwidth $B$ and time $T$ for each agent. The integral of $b'_m(t)$ with respect to $t$ corresponds to the time-frequency resources of agent $m$, denoted by $n'_m$

$$n'_m = \int_0^T b'_m(t) dt. \tag{2}$$

The Power-Spectral-Density (PSD) of the transmit power for agent $m$ is denoted by $q_m$. It is assumed to stay constant for the whole transmission. The total energy $E_m$ spent for the transmission to agent $m$ is then given by

$$E_m = q_m \int_0^T b'_m(t) dt. \tag{3}$$

To account for the deadlines $\tau_m$ in (3), the assigned bandwidth for agent $m$, $b'_m(t)$, must be set to zero for $t > \tau_m$. The receiver noise is assumed to be Additive white Gaussian noise (AWGN), whose power $\sigma_m^2(t)$ depends only on the noise PSD

$N_0$ and the bandwidth $b'_m(t) \leq B$ assigned to agent $m$ at $t$, i.e. $\sigma_m^2(t) = N_0 b'_m(t)$. The SNR at agent $m$ is then

$$\gamma_m = \frac{q_m b'_m(t)}{\sigma_m^2} = \frac{q_m b'_m(t)}{N_0 b'_m(t)} = \frac{q_m}{N_0}. \tag{4}$$

Since the commands are short and transmit time and bandwidth is limited, the well-known Shannon capacity formula

$$C_m = \log_2(1 + \gamma_m) \tag{5}$$

for error-free transmission is too optimistic to determine the minimum SNR and has to be extended for short packets. Therefore, also the channel dispersion for agent $m$,

$$V_m = \gamma_m \frac{2 + \gamma_m}{(1 + \gamma_m)^2} \log_2^2(e), \tag{6}$$

has to be considered. In [9] the approximation

$$V_m \approx \log_2^2(e) \tag{7}$$

is given, which is valid for $\gamma_m \geq 5$dB. For the strict demands on $p_c$, the short packet sizes $N$ and limited resources $n_m$, generally $\gamma_m \geq 5$dB is required. For an AWGN channel, the normal approximation from [5] gives the short packet formula for a packet error rate $p_c$, given a certain packet size $N$, the number of time-frequency resources $n_m$, the channel dispersion $V_m$ and the SNR at the receiver $\gamma_m$. The packet error probability $p_c, m$ for agent $m$ can then be approximated by

$$p_{c,m} \approx Q\left(\frac{n_m C_m - N + \frac{\log_2 n_m}{2}}{\sqrt{n_m V_m}}\right), \tag{8}$$

where $Q(\cdot)$ is the Gaussian Q-function.

The minimal Shannon capacity corrected for short packets $C_{\text{corr},m}$ and therefore $E_m$ required to fulfill the latency and error rate requirements $p_{c,m}$ for each agent $m$ can be calculated using a reformulated version of (8):

$$C_{\text{corr},m} \approx \frac{1}{n'_m}\left(\sqrt{n'_m V_m} Q^{-1}(p_c) + N - \frac{\log_2(n'_m)}{2}\right) \tag{9}$$

$$E_m(n'_m, G_m, N, p_c) = (2^{C_{\text{corr},m}} - 1)\frac{N_0 n'_m}{G_m} \tag{10}$$

In the next section, the minimum total energy for continuous $n_m$ will be derived. In practice, however, a continuous allocation of time-frequency resources is not possible. Therefore, we follow the approach used in mobile radio standards like 5G New Radio (NR) to implement the OFDM scheme, dividing the time-frequency plane into a grid of rectangles, called resource elements. The total available bandwidth $B$ is split into $N_{\text{sc}}$ subcarriers. The subcarrier bandwidth is $b_{\text{sc}}$, such that $B = b_{\text{sc}} N_{\text{sc}}$. The number of OFDM symbols per time slot is $N_{\text{sym}}$, such that $T = N_{\text{sym}} t_{\text{sym}}$. Considering the available bandwidth, we define a time-frequency resource element as $t_{\text{sym}} b_{\text{sc}}$. Each resource element is identified by its time index $t, t = 1, \ldots, N_{\text{sym}}$ and subcarrier index $s, s = 1, \ldots, N_{\text{sc}}$. The number of resource elements for agent $m$ is denoted by $n_m$. The complete distribution of resource elements is collected in the vector $\boldsymbol{n} = [n_1, \ldots, n_M]^{\mathrm{T}}$. Due to the nature of OFDM, the area of one resource element is always $t_{\text{sym}} b_{\text{sc}} = 1\text{s} \cdot \text{Hz}$. In Sec. IV, two algorithms to derive distributions of resource elements $\boldsymbol{n}$ are proposed.

## III. Problem Formulation

### A. General formulation

The overall goal is to minimize the total energy $E = \sum_{m=1}^{M} E_m(n'_m, G_m, N, p_c)$ used for the command transmission. Since $B$ as well as the available time, due to the deadlines, is limited, $n'_m$ is also limited. The energy minimization problem for the continuous resource case is then

$$\min_{n'_1, \ldots, n'_M} \sum_{m=1}^{M} E_m(n'_m, G_m, N, p_c), \tag{11a}$$

s.t.

$$\sum_{m=1}^{M} b'_m(t) \leq B \text{ for } 0 \leq t \leq T, \tag{11b}$$

$$b'_m(t) = 0 \text{ for } \tau_m < t \leq T, m = 1, \ldots, M, \tag{11c}$$

where (11b) enforces the bandwidth limitation and (11c) effectively restricts the transmission to $0 \leq t \leq \tau_m$.

### B. Convex reformulation

Problem (11a)-(11c) is hard to tackle, because the solution space is non-convex. We will now implement constraints on $n_m$ and $b'_m(t)$ to get a convex subset of the original solution space, still containing the optimal solution.

First, we restrict on $n_m$ to make (11a) convex. In [9], the partial convexity of (10) in $n_m$ up to an inflection point $n_{m,\text{thr}}$, i.e. for $n_m \leq n_{m,\text{thr}}$, is shown. Furthermore, the number of resources $n_m$ achieving the global minimum of (10), $n_{m,\text{min}}$, is shown to be $0 \leq n_{m,\text{min}} \leq n_{m,\text{thr}}$. Thus, $n_m \leq n_{m,\text{thr}}$ will turn (10) and (11a) into convex functions in $n_m$.

Since (11a) is only based on $n_m$ and not on $b'_m(t)$ directly, $b'_m(t)$ can be restricted to be piecewise constant, without further restrictions on $n_m$. The values of the constant pieces are then collected in a vector $\boldsymbol{b}_m$.

As a consequence, (2) becomes a sum of rectangular areas. The width of the rectangles is selected as the distance between two consecutive deadlines. With the auxiliary variable $\tau_0 = 0$, we have the convex problem

$$\min_{n_1, \ldots, n_M} \sum_{m=1}^{M} E_m(n'_m, G_m, N, p_c) \tag{12a}$$

$$n_m := \sum_{k=1}^{M} (\tau_k - \tau_{k-1}) b_{m,k} \tag{12b}$$

s.t.

$$\sum_{m=1}^{M} b_{m,k} \leq B \text{ for } k = 1, \ldots, M \tag{12c}$$

$$b_{m,k} = 0 \text{ for } m = 1, \ldots, M, k = 1, \ldots, m \tag{12d}$$

$$n_m \leq n_{m,\text{thr}} \text{ for } m = 1, \ldots, M \tag{12e}$$

## IV. Resource Scheduling Algorithms

### A. Gradient-Based Resource Scheduling Algorithm

The continuous allocation of time-frequency resources is not possible in an OFDM scheme, which splits the time-frequency plane into a grid of small rectangles. Therefore, (12a)-(12e) can only be used as a lower bound on $E$. To find a solution for the discrete-time and discrete-bandwidth problem, two scheduling algorithms are developed. The first algorithm is based on the fact that (12a) is convex in $n_m$ for all $m$ up to $n_{m,\text{thr}}$. All resource elements are iteratively allocated to the agents. In each iteration, the resource elements previously allocated $\boldsymbol{n} = [n_1, \dots, n_M]^{\mathrm{T}}$ determine the possible reduction of $E_m$ for each agent $m$, if an additional resource element is allocated to it. Therefore, the gradient of $E$, $\frac{\partial E}{\partial \boldsymbol{n}}(\boldsymbol{n})$ is used as the decision criterion to select the agent for the resource element in the current iteration. The resource elements can be distributed to the agents according to (12d) and (12e). However, not all resource elements are beneficial to all agents because of (12d). For resources at $t$, only agents with $\tau_m \geq t$ can benefit. The larger $t$, the more deadlines $\tau_m$ have passed, hence less agents will benefit from these resources. Before allocating the resource elements, the level of competition, i.e. how many agents can actually benefit form a certain resource element, must be calculated for each resource element.

Therefore, the algorithm consists of two phases. First, the level of competition for each resource element is determined. Second, the resource elements are allocated to the agents, starting with the resource elements with the lowest level of competition. If multiple agents can use a resource element, the agent who achieves a greater energy reduction with this additional resource element gets it. In the first phase, the level of competition is stored in matrix $\boldsymbol{C} \in \mathbb{N}^{N_{\text{sc}} \times N_{\text{sym}}}$. The element $c_{s,t}$ of matrix $\boldsymbol{C}$ contains the number of agents, which can use the resource element at subcarrier $s$ and time instant $t$. To calculate $\boldsymbol{C}$, first, the three dimensional array $\boldsymbol{R} \in \{0;1\}^{N_{\text{sc}} \times N_{\text{sym}} \times M}$ is generated. The element $r_{s,t,m}$ is set to 1, if agent $m$ can use the resource element at subcarrier $s$ and time instant $t$, and to 0 otherwise. Finally, the array $\boldsymbol{R}$ is summed up along the third dimension to get $\boldsymbol{C}$, i.e. $c_{s,t} = \sum_{m=1}^{M} r_{s,t,m}$.

In the second phase, the resource elements are allocated to the agents in increasing level of competition, starting with elements with $c_{s,t} = 1$ up to $c_{s,t} = M$. Elements with $c_{s,t} = 0$ are neglected, because no agent benefits from them. Now, all resource elements with the current level of competition are determined and their subcarrier and time indices $s$ and $t$ are stored in the vectors $\boldsymbol{s}$ and $\boldsymbol{t}$, respectively. In each iteration, one resource element identified by corresponding $s$ and $t$ from $\boldsymbol{s}$ and $\boldsymbol{t}$ is considered. First, the agents competing for this element are stored in the vector $\boldsymbol{m}$. Then, the current total number $n_m$ of elements allocated to agent $m$ is calculated. The agent $m$ from $\boldsymbol{m}$ with the smallest derivative $g_m = \frac{\partial E_m}{\partial n_m}(n_m)$ is assigned the resource element, because this results in the greatest reduction of $E$. If $g_m \geq 0$, $n_{m,\text{min}}$ is achieved, agent $m$ has no benefit from any more resources. The allocation is stored in the three dimensional array $\boldsymbol{A} \in \{0;1\}^{N_{\text{sc}} \times N_{\text{sym}} \times M}$.

**Algorithm 1** Scheduling algorithm

**Input:** $\tau_1, \dots, \tau_M$
**Output:** $n_1, \dots, n_M$

PHASE 1: Calculate levels of competition
1: **for** $t = 1 : N_{\text{sym}}$ **do**
2:   **for** $s = 1 : N_{\text{sc}}$ **do**
3:     **for** $m = 1 : M$ **do**
4:       **if** $t \leq \tau_m$ **then** ▷ check, if resource element at $s$ and $t$ is before deadline of agent $m$
5:         $r_{s,t,m} = 1$
6:       **else**
7:         $r_{s,t,m} = 0$
8:       **end if**
9:     **end for**
10:   **end for**
11: **end for**
12: $c_{s,t} = \sum_{m=1}^{M} r_{s,t,m}$ ▷ calculate levels of competition

PHASE 2: Allocate resource elements to agents
13: **for** $l = 1 : M$ **do**
14:   $(\boldsymbol{s}, \boldsymbol{t}) = \text{findindex}(c_{s,t} == l)$ ▷ find all resource elements for current level of competition
15:   **for** $(s,t) \in (\boldsymbol{s}, \boldsymbol{t})$ **do**
16:     $n_m = \sum_{m=1}^{M} \sum_{s=1}^{N_{sc}} a_{s,t,m}$ ▷ calculate current numbers of resource elements
17:     $\boldsymbol{m} = \text{findindex}(r_{s,t} == 1)$ ▷ find all agents competing for this element
18:     $g_m = \frac{\partial E_m}{\partial n_m}(n_m)$ ▷ calculate derivative for $m \in \boldsymbol{m}$
19:     $o = \text{sort}(\boldsymbol{g})$
20:     **for** $m \in o$ **do**
21:       **if** $m \in \boldsymbol{m}$ **then** ▷ check if element usable by agent $m$
22:         **if** $g_m < 0$ **then** ▷ check if not yet larger than $n_{m,\text{min}}$
23:           $a_{s,t,m} = 1$
24:           break for
25:         **else**
26:           $a_{s,t,m} = 0$
27:         **end if**
28:       **end if**
29:     **end for**
30:   **end for**
31: **end for**
32: $n_{m,\text{opt}} = \sum_{m=1}^{M} \sum_{s=1}^{N_{sc}} a_{s,t,m}$ ▷ calculate final numbers of resource elements

Finally, all allocations from $\boldsymbol{A}$ are combined to get the total numbers of resources $n_m$. If the resource elements can only be assigned in groups like the physical resource blocks in 5G NR [10], the $N_{\text{sc}}$ for the algorithm has to be reduced accordingly. For the calculation of $g_m$ and the final counting to get $n_m$, the size of one resource element group has to be adapted. A summary of the code is presented in Alg. 1.

### B. Deadline-Aware Balancing Scheduling Algorithm

For comparison, a simpler resource element balancing algorithm is developed. The gradient of $E$ is not considered, but rather the number of resource elements already allocated to agent $m$ is the decision criterion for the current iteration. The resource element is allocated to the agent with the least number of resources, i.e. with the lowest $n_m$ whose deadline has not yet passed and, thus, can use the resource element of the current iteration. This is done by replacing line 18 in Alg. 1 by $g_m = n_m$. As a consequence, an equal allocation, as far as the deadlines allow, is achieved. Moreover, the channel gain $G_m$ is not considered.

## V. Numerical Results

The numerical results are generated for parameters based on the 5G NR standard. In particular, we consider the frame

| | |
|---|---|
| Carrier frequency $f_c$ | 6 GHz |
| Noise power spectral density $N_0$ | $-174\,\frac{\text{dBm}}{\text{Hz}}$ |
| Number of subcarriers $N_{\text{sc}}$ | 4 |
| Subcarrier bandwidth $B_{sc}$ | 15 kHz |
| Symbol duration $t_{\text{sym}}$ | 66.666 $\mu$s |
| Deadlines $\tau_m$ | $\tau_m \sim \mathcal{U}(14t_{\text{sym}}; 70t_{\text{sym}})$ |
| Maximum packet error probability $p_c$ | $10^{-9}$ |
| Number of OFDM symbols $N_{\text{sym}}$ | 70 |

TABLE I: Simulation parameters

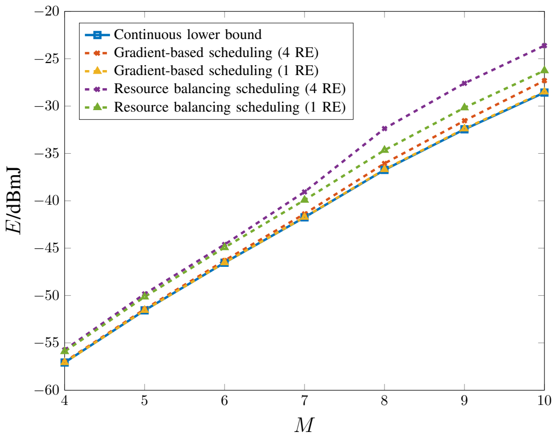


Fig. 2: Energy for different number $M$ of agents

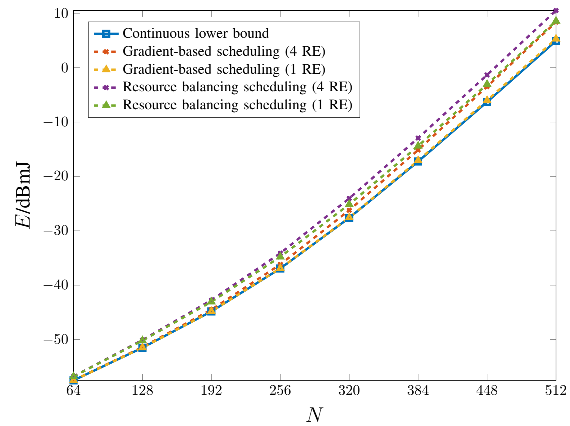


Fig. 3: Energy for different packet sizes $N$

structure [10] and the possibility to make shorter time allocations instead of assigning a whole frame to an agent, so-called minislots [11]. The carrier frequency $f_c$ is chosen to be 6 GHz, corresponding to unlicensed band n96 of 5G NR. The channel is assumed to be pure LOS, so $G_m$ only depends on the distance of agent $m$ to the central controller, but not on the subcarrier frequency or the time $t$. Random deadlines $\tau_1, \ldots, \tau_M$ are used and the results are derived from a Monte-Carlo simulation. Each Monte-Carlo run uses a new set of deadlines, drawn from the uniform distribution $\mathcal{U}(14t_{\text{sym}}; 70t_{\text{sym}})$ for each $\tau_m$ to ensure $\tau_m \leq T$ on the one hand and make the problem feasible on the other hand. Parameters common to all simulations are given in Table I. The allocations generated by the gradient-based scheduling algorithm described in Sec. IV-A and the balancing scheduling algorithm described in Sec. IV-B are compared to the continuous lower bound derived in Sec. III-B. Both, the gradient-based scheduling algorithm and the balancing scheduling algorithm, are used to either assign a single resource element or four resource elements from a single OFDM symbol per iteration.

In Fig. 2 the required energy $E$ for different numbers $M$ of agents is shown. The available resources are kept constant as in Table I. The agents are spaced equidistant on a straight line starting at the central controller. The agent $m = 1$ is at a distance $d_{\text{min}} = 5$ m from the central controller, $m = M$ at $d_{\text{max}} = 100$ m. The packet size is $N = 256$ bits, which is in the center of the range for $N$, where (8) is valid [5]. The more agents are in the scenario, the less resources per agent are available, therefore the required energy $E$ increases. In a highly constrained scenario, i.e. $n_m \ll n_{m,\text{min}}$, changes in $n_m$ have greater influence on $E_m$, because (10) is strictly convex in $n_m$ for $n_m < n_{m,\text{min}}$. The continuous lower bound gives the minimum $E$, if there were no quantization effects of $n_m$ on $E$. The influence of the coarse grid with the allocation in blocks of four resource elements on the performance becomes apparent especially for $M \geq 8$. For $M \geq 8$, the gradient based scheduler is about 0.1 dB worse than the optimum, in single resource element case and about 1.25 dB in the 4-resource element case. Meanwhile, the balancing algorithm needs 5 dB and 2.3 dB more than the lower bound, respectively.

In Fig. 3, the influence of different packet sizes $N$ on the required energy $E$ is investigated. The number of agents is $M = 7$, the agents are spaced equidistant from $d_{\text{min}} = 5$m to $d_{\text{max}} = 100$m. Since the number of resources and agents is fixed, the scenario becomes more constrained when the packet size increases. This is because the more bits are transmitted, the larger is the number of bits per resource element. The effect on $E$ is similar to the previous result, due the curvature of (10) for small $n_m$ in constrained scenarios. The gradient-based scheduling algorithm for a single resource element gets results about 0.3 dB worse than the optimum derived by solving (11a) even for high $N$. The balancing benchmark scheduler always needs about 3dB more energy than the gradient-based algorithm, even with the fine grid of only one resource element, because it does not consider the different gradients of $E_m$ caused by the different channel gains $G_m$ and packet sizes $N$.

In Fig. 4, the influence of different channel gains $G_m$ on the required energy $E$ is shown. The $M = 7$ agents are placed again equidistantly, agent $m = 1$ is at $d_{\text{min}} = 5$ m, but agent $m = M$ is varied from $d_{\text{max}} = 10$ m to $d_{\text{max}} = 160$ m. The agents in between are placed accordingly to keep the equidistant positioning. The higher $d_{\text{max}}$, the greater is the distance between neighbouring agents and, thus, their difference in $G_m$. The optimal resource allocation has to account for this difference. Since the balancing scheduler only takes the number of resources into account, the energy requirement is up to 3 dB higher than for the gradient based scheduler. This is

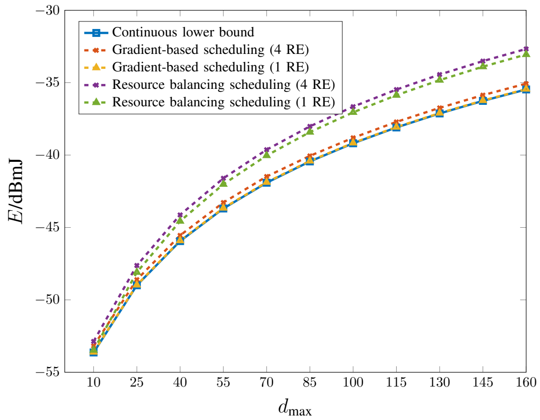


Fig. 4: Energy for varying channel conditions $G_m$

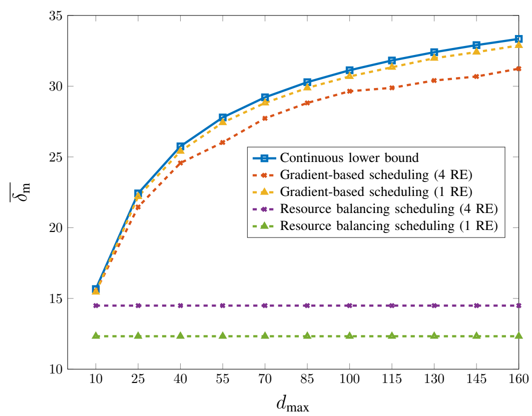


Fig. 5: Average difference of minimum and maximum $n_m$, $\delta_m$

an interesting result especially for scenarios, where non-line-of-sight propagation leads to largely different channel gains.

The benefit of assigning resources $n_m$ based on the gradient of $E$, compared to the balancing scheduling for different channel gains $G_m$ is investigated in Fig. 5. The setup is the same as in Fig. 4. The difference between the largest and the smallest $n_m$ for all agents, averaged over all runs with the same $d_{\max}$, is shown. The continuous lower bound suggests a larger difference in the assigned $n_m$ is beneficial in terms of energy consumption by assigning more resources to agents with low $G_m$. For the gradient based scheduler, almost the same $\delta_m$ as the continuous lower bound is attained. Especially for larger $d_{\max}$, the 4-resource element case cannot achieve the results of the finer resource grid, since the adaptation is worse due to the coarser grid. The difference in the distances $r_m$ and, thus, the channel gains $G_m$ is not considered by the balancing scheduler, resulting in the costant average differences.

## VI. Conclusion

In this paper, the time-frequency resource allocation for a single central controller transmitting control commands to multiple agents, was optimized for minimum energy consumption. The agents needed to receive the control commands before an individual deadline. The resulting continuous minimization problem was shown to be convex. For application in mobile radio systems like 5G, the resource allocation has to be done based on fixed size resource elements, turning the problem into a mixed integer problem. An algorithm to find a scheduling of these resource elements based on the gradient of the required transmit energy was proposed and compared to a simple resource balancing algorithm only considering the deadlines. The gradient-based algorithm was shown to perform only about 0.3 dB worse in terms of required energy than the continuous optimum and showed improvements of more than 50% to the balancing algorithm, especially if the channel gains are very different.


## Acknowledgement

This work has been performed in the context of the DFG Collaborative Research Center (CRC) 1053 MAKI, the HMWK LOEWE Center emergenCITY and the BMBF project Open6GHub.


## References


[1] W. Khan, M. Rehman, H. Zangoti, M. Afzal, N. Armi, and K. Salah, "Industrial Internet of Things: Recent advances, enabling technologies and open challenges," *Computers & Electrical Engineering*, vol. 81, p. 106522, 2020.

[2] "Cisco Annual Internet Report (2018-2023) White Paper," Tech. Rep., March 2020. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

[3] P. Schulz, M. Matthe, H. Klessig, M. Simsek, G. Fettweis, J. Ansari, S. A. Ashraf, B. Almeroth, J. Voigt, I. Riedel, A. Puschmann, A. Mitschele-Thiel, M. Muller, T. Elste, and M. Windisch, "Latency Critical IoT Applications in 5G: Perspective on the Design of Radio Interface and Network Architecture," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 70–78, February 2017.

[4] 3GPP Technical Report 38.824, "Study on physical layer enhancements for NR ultra-reliable and low latency case (URLLC) (Release16)," ETSI, Tech. Rep., 2019, v16.0.0.

[5] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel Coding Rate in the Finite Blocklength Regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[6] W. Liu, G. Nair, Y. Li, D. Nesic, B. Vucetic, and H. V. Poor, "On the Latency, Rate, and Reliability Tradeoff in Wireless Networked Control Systems for IIoT," *IEEE Internet of Things Journal*, vol. 8, no. 2, pp. 723–733, 2021.

[7] F. E. da Silva, A. L. V. Iaremczuk, R. D. Souza, G. Brante, G. L. Moritz, and S. Hussain, "Hybrid ARQ in Wireless Packetized Predictive Control," *IEEE Sensors Letters*, vol. 4, no. 5, pp. 1–4, 2020.

[8] Y. Wu, D. Qiao, and H. Qian, "Efficient Bandwidth Allocation for URLLC in Frequency-Selective Fading Channels," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.

[9] C. Sun, C. She, C. Yang, T. Q. S. Quek, Y. Li, and B. Vucetic, "Optimizing Resource Allocation in the Short Blocklength Regime for Ultra-Reliable and Low-Latency Communications," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 402–415, 2019.

[10] 3GPP Technical Report 38.211, "Physical channels and modulation (Release16)," ETSI, Tech. Rep., 2021, v16.0.0.

[11] 3GPP Technical Report 38.912, "Study on New Radio (NR) access technology (Release16)," ETSI, Tech. Rep., 2020, v16.0.0.