Analysis of Uplink Scheduling for Haptic Communications

Maliheh Mahlouji, Toktam Mahmoodi

Centre for Telecommunications Research, Department of Informatics
King's College London, London WC2B 4BG, UK

Abstract—**While new mechanisms and configurations of the 5G radio are offering step forward in delivery of ultra-reliable low latency communication services in general, and haptic communications in particular, they could inversely impact the remainder of traffic services. In this paper, we investigate the uplink access procedure, how different advances in this procedure enhance delivery of haptic communication, and how it affects the remainder of traffic services in the network. We model this impact as the remainder of service, using stochastic network calculus. Our results show how best the tradeoff between faster or more resource efficient uplink access can be made depending on the rate of haptic data, which is directly relevant to the application domain of haptic communication.**

Index Terms—**bilateral teleoperation, 5GNR, URLLC, UL scheduling, stochastic network calculus (SNC).**

I. INTRODUCTION

Enabling ultra-reliable low latency communication (URLLC) is one of the main challenges of the next generation mobile networks, a.k.a. 5G. The standardization effort of New Radio (5GNR) has, thus, proposed series of enhancements to reduce the radio transmission delay and step forward in enabling URLLC for critical applications [1], [2]. However, configurations offering lower latency and higher reliability for the URLLC service, will inversely impact the remainder of traffic services (we refer to this as leftover traffic).

Haptic communication, particularly when used for critical applications, is one the finest examples of the URLLC services, sending haptic data (such as position/velocity, force/torque, texture, etc) from master to slave and receiving the haptic data at the master side. This round-trip communication should be performed within low and stable latency in order for stability of the control scheme to be maintained. There is a rich literature of teleoperation, focusing on balancing the control stability with transparency in order to compensate for communication delay [3], where control stability is maintained in the expense of lowering the quality of haptic communication. Therefore, communication latency will have the major impact on transparency of the haptic communication and hence will be a deciding factor in whether haptic communication could be used for critical and precise teleoperation [4].

In this paper, we focus on the Uplink (UL) scheduling procedure and the delay introduced through UL scheduling to the communication, as one of the main configuration of 5GNR for URLLC [1], [2]. To this end, we study four different UL scheduling: 1) the dynamic scheduling with scheduling request (SR) procedure for every transmission, as occurs in the 4G-LTE, 2) the semi-persistent scheduling, 3) an adaptation of semi-persistent scheduling that adapts the persistency depending the traffic, namely soft resource reservation, and finally 4) the fast UL, that is considered as an alternative for the 5G. We look into these four scheduling mechanisms, and how employing them for URLLC services will impact serving the remainder of traffic in the system. While network slicing could potentially provide isolation between these services [5], the way resources are split between different slices is a complex decision for which detailed studies, such as the study in this paper, are essential.

This paper is organized as follows: in section II, the four UL scheduling mechanisms under study are described in details, and the delay they incur to the communication is formulated. In section III, we present a model to compute radio access delay for the leftover traffic as a result of deploying any of the above mentioned UL scheduling techniques. In section V, we present the numerical results and finally, in section VI, concluding remarks are summarized.

II. UPLINK SCHEDULING MECHANISMS

In this section, we will review four potential UL scheduling mechanisms, based on existing scheduling in the LTE, but also potential proposals for ultra-low latency applications. These four scheduling mechanisms are chosen based on the literature ([2], [6], [7], [8], and [9]) in a way that they cover a diverse range of UL scheduling techniques in terms of their offered delay and communication resource consumption. The main goal of this section is to summarize delay incurred by each of these UL scheduling techniques.

A. Dynamic Scheduling

In Dynamic Scheduling (DS), user (UE) needs to send Scheduling Request (SR) to eNodeB (eNB) for each packet transmission. This is the UL scheduling technique used in LTE, and it provides significant benefits in terms of spectral efficiency, while also introducing an additional delay due to the involved procedure [6].
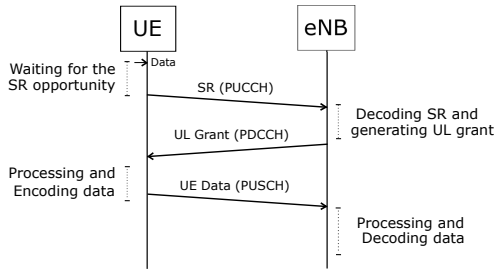
Fig. 1. Schematic of Scheduling Request (SR) procedure [7]

In order to send SR, the UE must wait till the next available SR-valid Physical Uplink Control Channel (PUCCH). In this paper, T_{SR} denotes the period of SR opportunity. Then, eNB will decode the SR and send an UL grant to the UE. After receiving the UL grant, UE decodes the grant and waits for the allocated grant to start transmission over Physical Uplink Shared Channel (PUSCH) [7].

A schematic of SR procedure in LTE is depicted in Figure 1. This procedure repeats for every UL transmission. Therefore, assuming successful signal transmission and reception in all SR procedure, the maximum UL radio access latency for each packet using dynamic scheduling can be expressed as d_{DS}, according to Equation (1).

$$
\begin{aligned}
d_{DS} \leq & T_{SR} + T_{tx(UE \to eNB)} + T_{proc(eNB)} + T_{tx(eNB \to UE)} + \\
& T_{proc(UE)} + T_{tx(UE \to eNB)} + T_{proc(eNB)} \\
= & T_{SR} + TTI + TTI + TTI + \\
& TTI + TTI + TTI = T_{SR} + 6TTI,
\end{aligned}
\tag{1}
$$

where T_{SR} is maximum waiting time for SR opportunity on PUCCH, $T_{tx(UE \to eNB)}$ and $T_{tx(eNB \to UE)}$ denote packet transmission time from UE to eNB and from eNB to UE, respectively. $T_{proc(UE)}$ and $T_{proc(eNB)}$ denote processing time for decoding and encoding packets in UE and eNB, respectively. Based on the 3GPP documentation [7], each of the elements contributing to the total delay, i.e d_{DS}, can be estimated as a function of Transmission Time Interval (TTI) and SR period T_{SR}, as seen in the evolution of Equation (1). The processing delays is estimated as three TTI in Release 14 [7], while for 5GNR it is estimated as one TTI [2]. Since our aim is to study the use of dynamic scheduling for haptic communications, and since we are referring to haptic radio delay, we consider the one TTI processing delay.

B. Semi-Persistent Scheduling

In latency critical applications, such as haptic communication that is the focus of this paper, repeating scheduling request procedure for every packets, as in dynamic scheduling, might not guarantee the application requirements. Also, depending on the pattern of data traffic, dynamic scheduling might introduce sig-

nificant additional overhead. One available solution for reducing scheduling latency is static allocation of UL grant with a given periodicity to the traffic. This mechanism, referred to in the literature as Semi-Persistent Scheduling (SPS), is an effective way to reduce scheduling latency of deterministic traffic like VoIP, but it will have inverse effect on the spectral efficiency in the case of bursty traffic due to un-used allocated resource in non-burst periods. Also, SPS might fail to transmit all packets during burst periods [8].

In SPS, after initialization, i.e. after the first UL grant, the radio access latency for the rest of UL transmissions can be expressed by d_{SPS}, according to Equation (2).

$$
\begin{aligned}
d_{SPS} \leq & T_{pg} + T_{RRB(UE)} + T_{proc(UE)} \\
& + T_{tx(UE \to eNB)} + T_{proc(eNB)} \\
= & T_{pg} + TTI + TTI + TTI + TTI = T_{pg} + 4TTI,
\end{aligned}
\tag{2}
$$

where T_{pg} is the period of the pre-allocated UL grant on Physical Downlink Control Channel (PDCCH), and $T_{RRB(UE)}$ denotes the time required by the UE to read the received Resource Block (RB) from eNB on PDCCH. Other notations are the same as in Equation (1). Similar to calculation of d_{DS}, the estimated values of each delay component are extracted from [7].

C. Soft-Resource Reservation

To improve the spectral efficiency of SPS for bursty traffic, Soft-Resource Reservation (SRR) [9] has been proposed. In this scheduling mechanism, the UL grant is reserved until the UE has data to transmit, otherwise, the UL grant is released, and for further transmission the UE needs to send another SR. Therefore, UL radio access latency during burst period follows Equation (2) and during non-burst period follows Equation (1).

D. Fast UL access

An enhanced version of SPS in order to accommodate the needs of URLLC communication, is proposed as the fast UL access (FA) [7]. In the fast uplink access, the UE has opportunity to transmit UL data every TTI after the data is ready to transmit. This scheduling mechanism makes UL radio access delay the same as DL radio access delay (this scheduling mechanism is also further discussed in [2]).

Therefore, with fast UL access, latency critical application user can send UL packet immediately after packet generation on the next available TTI without getting permission from eNB. The maximum UL radio access latency, d_{FA}, could, therefore, be calculated as in Equation (3).

$$
\begin{aligned}
d_{FA} \leq & TTI + T_{proc(UE)} + T_{tx(UE \to eNB)} + T_{proc(eNB)} \\
= & TTI + TTI + TTI + TTI = 4TTI.
\end{aligned}
\tag{3}
$$

Delay components of Equation (3) is similar to Equation (2), except that T_{pg} for fast UL access is equal
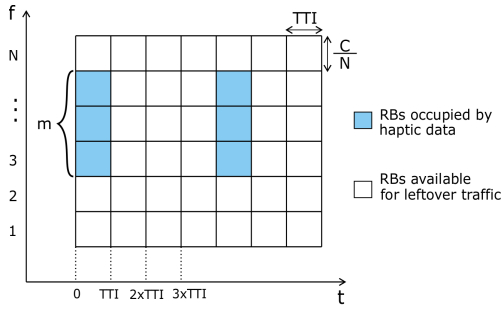
Fig. 2. System Model

to the TTI, since latency critical user can potentially transmit in each TTI without getting permission. Also, in fast UL access, there is no need to read the PDCCH before UL transmission and thus, $T_{RRB(UE)}$ equals zero.

III. DELAY ANALYSIS FOR LEFTOVER TRAFFIC

A. System Model

We consider a constant rate multi-channel server consisting of N frequency channels with total transmission rate of C. Transmission rate of each channel is considered as C/N as depicted in Figure 2. All channels are synchronized in time and each square represents one Resource Block (RB) of time × frequency. The time unit for allocating RBs is one Transmission Time Interval (TTI). The coloured blocks are occupied by haptic data and the blank blocks will remain for use of leftover traffic. Since haptic data requires ultra-low latency, we assume that for transmission of each packet, only one block is occupied in time and it will spread in frequency to fulfil the resource block it needs. Denoting the amount of time × frequency resource needed for haptic data by ρ, the number of frequency blocks occupied by haptic data, denoted by m, is computed by the following formula:

$$m = \lceil \frac{N \times \rho}{C \times TTI} \rceil \ . \tag{4}$$

After occupation of the coloured blocks by haptic data (i.e. m blocks for each packet transmission opportunity), we will use Stochastic Network Calculus (SNC) to compute the service remained for the leftover traffic and the resulting stochastic delay bound.

B. Methodology

The Stochastic Network Calculus (SNC) [10] is a mathematical method being used for analyzing delay and backlog in a wide class of stochastic traffic and service models. In this paper, we will apply SNC to analyze the effect of using different UL scheduling mechanisms, for haptic communication, on the delay of the remainder of traffic, i.e. leftover service delay. Here, we will provide some definitions, from [11], in the SNC that will be used in the next subsection for modeling and analysis.

Definition 1. *Stochastic Arrival Curve (SAC):* The virtual backlog centric (v.b.c) SAC α bounded by function f is defined for a flow if for all $x \geq 0$ and all $0 \leq \tau \leq t$ there holds:

$$P\{ \sup_{0 \leq \tau \leq t} [A(\tau,t) - \alpha(t-\tau)] \geq x\} \leq f(x) \ . \tag{5}$$

where $A(\tau,t)$ is the cumulative arrival traffic during the time $(\tau,t]$.

Definition 2. *Stochastic Service Curve (SSC):* The weak SSC β bounded by function g is defined for a server if for all $x \geq 0$ and $t \geq 0$ there holds:

$$P\{A \otimes \beta(t) - A^\star(t) > x\} \leq g(x)$$
$$A \otimes \beta(t) = \inf_{0 \leq \tau \leq t} \{A(\tau) + \beta(t-\tau)\} \ , \tag{6}$$

where $A^\star(t)$ is cumulative departure traffic during the time $(0,t]$ and $\otimes$ is the min-plus convolution.

Theorem 1. *Stochastic Delay Bound :* The stochastic delay $D(t)$ of the arriving flow in a queuing system with SAC and SSC as defined in the Definitions 1 and 2, respectively, is upper bounded as:

$$P\{D(t) > h(\alpha+x,\beta)\} \leq f \otimes g(x) \ , \tag{7}$$

where

$$h(\alpha+x,\beta) = \sup_{0 \leq \tau \leq t} \{\inf\{s \geq 0 : \alpha(\tau)+x \leq \beta(\tau+s)\}\} \tag{8}$$

$h(\alpha+x,\beta)$ is the maximum horizontal distance between β and $\alpha+x$.

C. Modeling and Analysis

In this part, we model the upper bound delay of leftover traffic, which refers to the remainder of traffic, after scheduling the haptic communication traffic. We formulate this problem similar to the case of cognitive radio with primary users, being haptic communication flows, and secondary users, being the remainder of traffic [12]. Since in SNC, a service curve is defined as a lower bound service provided to the traffic, in order to compute the leftover service curve, we should consider the worst case time period $(\tau,t]$ $0 \leq \tau \leq t$. The leftover service would be greater than its worst case during $(\tau,t]$:

$$S_{lo} \geq \frac{(N-m)C}{N}(t-\tau) + \frac{mC}{N}(t-\tau-RTTI-2TTI)^+, \tag{9}$$

where S_{lo} denotes the guaranteed leftover service, $a^+ = max\{a,0\}$, and R is the number of RBs occupied by haptic traffic during the worst case time period $(\tau,t]$. The first sentence in the right hand side is the service available for leftover traffic in $(N-m)$ sub-channels that weren't used by haptic traffic, thus all period of $(\tau,t]$ can be used by the leftover traffic. The second sentence represents the leftover service in m sub-channels that are occupied by haptic traffic, thus, the R time periods have been subtracted from $(\tau,t]$ time period. Also, two

TTIs are excluded since in the worst case, τ might be just after TTI start and t might be just before TTI end.

The right hand side of Equation (9) is defined as leftover Stochastic Service Curve (SSC) β_{lo} which is a function of time period $(\tau, t]$:

$$\beta_{lo}(t-\tau) \triangleq \frac{(N-m)C}{N}(t-\tau) + \frac{mC}{N}(t-\tau-RTTI-2TTI). \tag{10}$$

Note that we remove the function a^+ for simplicity and since $a^+ \geq a$ which still satisfies Inequality (9). We consider Stochastic Arrival Curve (SAC) of leftover traffic as compound Poisson Levy process which is the summation of all packet lengths arrived till time t, with packet lengths of i.i.d. random variables with distribution of $exp(\sigma)$ and number of packets with distribution of independent Poisson process with parameter λ. The corresponding SAC would be [12]:

$$\alpha_{lo}(t) = \frac{\lambda t}{\theta}(e^{\theta\sigma}-1) \quad , \quad f(x) = e^{-\theta x}, \tag{11}$$

where $\theta \geq 0$ (free parameter). Let $d_0 = h(\alpha_{lo}(t) + x, \beta_{lo})$, according to the definition in Equation (8), we can compute d_0 by the equation $\beta_{lo}(d_0) = x$. We assume the communication channel will serve all packets of leftover traffic (the communication channel is error free). Hence, $g(x)$ in Equation (7) equals zero and $f \otimes g(x) = f(x)$. By substituting into Equation (7), we can get:

$$P\{d_{lo} > d_0\} \leq e^{-\theta\beta_{lo}(d_0)}, \tag{12}$$

where d_{lo} denotes leftover traffic delay. One can compute the upper bound of d_{lo} with outage probability ε (i.e. d_0) when the right hand side of the Equation (12) is equal to $1-\varepsilon$. Note that to have a stable queue, the arrival rate must be less than or equal to the service rate, thus, $\lambda(e^{\theta\sigma}-1)/\theta$ must be less than the coefficient of $(t-\tau)$ in $\beta_{lo}(t-\tau)$. Also note that we can get the tightest upper bound with the maximum θ that satisfies the stability condition.

IV. EFFECT OF HAPTIC UL SCHEDULING ON LEFTOVER TRAFFIC

To model the effect of UL scheduling mechanism used for haptic communications, as a case of URLLC, on the leftover service and delay, firstly, we need to consider a model for haptic traffic packet arrivals. Based on experiments in [13], we consider periodic bursty traffic for haptic data as depicted in Figure 3. Bursts arrive with period of T_p, their duration is T_b, and the inter arrival time during the burst period is T_{ib}. Also, during non-burst period, haptic data arrives every T_{nb}. Note that in the worst case time period, there would be an excess burst other than the bursts inside the periodic traffic like the one shown at the end of T_p in Figure 3.

In our analysis, we assume that haptic data is a part of the global closed control loop (such as position, force
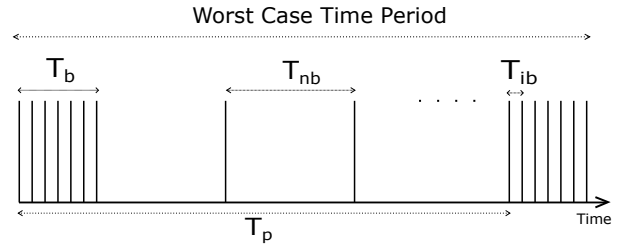

Fig. 3. Model of haptic traffic packet arrivals. In this figure, T_p is the duration of one traffic period, T_b is duration of each burst, T_{nb} is packet inter arrival time in non-burst periods, and T_{ib} is packet inter arrival time during burst periods.

and velocity in teleoperation system) and thus, when there is an opportunity to start transmission to eNB, only the latest version of haptic data is meaningful to be transmitted (as the most up-to-date position, force, etc.) and the previous un-transmitted data will be discarded.

With the aforementioned assumptions, we try to compute the upper bound delay of leftover traffic for each of the UL scheduling mechanisms described in section II, using the model detailed in Section III. In fact, we only need to find $\beta_{lo}(t-\tau)$ in Equation (10) and then use it to find d_0 in Inequality (12).

A. Dynamic Scheduling

If haptic packets inter arrival time during burst periods satisfies Inequality (13), then no haptic data packet would be discarded by the UE. In other words, this inequality ensures that the UE can start transmission of the arrived packet before the arrival of any new packet.

$$T_{ib} > T_{SR} + T_{tx(UE \to eNB)} + T_{proc(eNB)} + T_{tx(eNB \to UE)} , \tag{13}$$

In Inequality (13), the right hand side is the first four sentences of Equation (1) representing the DS scheduling delay before encoding packets by the UE. Therefore, the SSC in this case would be,

$$\begin{aligned} \beta_{lo-DS}(t-\tau) &\triangleq \frac{(N-m)C}{N}(t-\tau)+ \\ &\frac{mC}{N}(t-\tau-n_p R_p TTI - R_b TTI - 2TTI) , \\ n_p &= \lfloor \frac{t-\tau}{T_p} \rfloor \quad , \quad R_b = \lfloor \frac{T_b}{T_{ib}} \rfloor \\ R_{nb} &= \lfloor \frac{T_p - T_b}{T_{nb}} \rfloor \quad , \quad R_p = R_b + R_{nb} , \end{aligned} \tag{14}$$

where n_p is the number of whole traffic periods within $t-\tau$. R_b, R_{nb}, and R_p are the number of transmitted haptic packets during each burst period T_b, each non-burst period $(T_p - T_b)$, and each traffic period T_p, respectively. As we discussed before, each haptic packet consumes one TTI in time domain but it spreads in frequency domain to achieve lower latency.

If T_{ib} doesn't satisfy the Inequality (13), but is still greater than twice of the right hand side of the Inequality (13), half of the haptic packets during burst

period would noti have the chance to be transmitted and therefore R_b in Equation (14) would be $R_b = \lfloor \frac{T_b}{2T_{ib}} \rfloor$. For lower T_{ib}, R_b changes accordingly in a same way.

B. Semi-Persistent Scheduling

In the SPS, a periodic grant is allocated for the haptic UE regardless of whether it has any packet to transmit on this or not. Assuming the period of the allocated grant is T_{pg}, the SSC would be,

$$\beta_{lo-SPS}(t-\tau) \triangleq \frac{(N-m)C}{N}(t-\tau)+$$
$$\frac{mC}{N}(t-\tau-n_p R_{pg}TTI - R_{bg}TTI - 2TTI) \quad (15)$$
$$R_{pg} = \lfloor \frac{T_p}{T_{pg}} \rfloor \quad , \quad R_{bg} = \lfloor \frac{T_b}{T_{pg}} \rfloor \; ,$$

where R_{pg} and R_{bg} are the number of allocated grants during each haptic traffic period and each haptic burst periods, respectively. In SPS, however, some allocated resources might be unused during non-burst periods and also some haptic packets might be discarded during burst periods.

C. Soft-Resource Reservation

In the SRR, during burst periods the UE keeps the periodic allocated grant (which has the period of T_{pg}) and transmit on the reserved RBs, but during non-burst periods, it will send SR and waits for the permission to transmit. Therefore, from the leftover RBs point of view, it is the same as SPS during bursts and the same as DS outside the bursts. Therefore, its SSC would be:

$$\beta_{lo-SRR}(t-\tau) \triangleq \frac{(N-m)C}{N}(t-\tau)+$$
$$\frac{mC}{N}(t-\tau-n_p(R_{bg}+R_{nb})TTI - R_{bg}TTI - 2TTI) \; ,$$
$$(16)$$

where n_p, R_{bg}, and R_{nb} are the same as previously defined in Equation (14) and (15).

D. Fast UL access

In fast UL access, if haptic packet inter arrival time is greater than TTI, no packet would be discarded and every packet would have the chance to be transmitted. In that case, SSC would be the same as DS in Equation (14). However, haptic packets would experience much less delay since there is no need for them to send SR and wait for the UL grant before starting transmission. If T_{ib} <TTI, some packets might be discarded and R_b would differ accordingly as discussed in section IV-A.

V. NUMERICAL RESULTS

This section presents the numerical results of the model described earlier. We assume haptic communication services use one of the four discussed UL scheduling and we see the impact of this on the remaining service rate, used for the leftover traffic. It is important

parameter	value
leftover traffic	[14]
λ	4 packets/sec
σ	1500 bytes/sec
Communication Channel	[2], [7]
TTI	0.125, 0.25, 0.5, 1 ms
T_{pg}	10×TTI
T_{SR}	TTI
SNC parameters	[12]
ε	10^{-5}
RB	10^{-3}
N	10
Traffic timing parameters	[9]
T_{ib}	1-3 ms
T_{nb}	50 ms
T_b	200 ms
T_p	1000 ms

to note that the UL scheduling delay for the leftover traffic is not included in the leftover delay plots here, and the presented delay is just the impact of remainder of service.

These analyses are performed using the parameters in Table I. The haptic traffic model is based on the presented results in [13], this traffic pattern is transmitted data after haptic lossy encoder has been applied to the haptic data. Using the assumptions in Table I and Equations (1)-(3), maximum radio access delay resulted by UL scheduling for haptic packets, are: 7TTI, for DS, 4TTI, for FA, and 14TTI for SPS and SRR (during the burst period).

Main performance indicators studied in this paper are as follows. First, we study how using each of the scheduling would affect the packet drop rate of haptic communications. Dropping haptic packets here is based on the assumption that a haptic packet will only be transmitted if it can be scheduled upon its arrival and it will not be buffered (the queue length is zero). This assumption relies on the fact that haptic data refers to physical attributes of position, velocity, etc, and when an update value of these attributes become available, the old value is not valid anymore. In addition, and due to the fact that the haptic data is encoded data, and every single packet should be transmitted for delivering high quality haptic communication, we consider only zero-packet-drop regions as our operational region. Second performance metric we look into is the *remainder of service*, referring to what service rate remains for the leftover traffic. We also examine the upper bound delay that can be offered to the leftover traffic, using this remainder of service.

Figure 4(a) shows the haptic packet drop rate for different UL scheduling mechanisms when TTI = 0.5 ms and T_{pg} = 10×TTI. It can be seen that using SPS and SRR cannot offer zero packet drop rate in the observed range of T_{ib}, thus, SPS and SRR are not good candidates with the above mentioned timing (i.e. TTI and T_{pg}). In addition, for DS, as soon as arriving haptic data has higher frequency than one packet per 2 ms, packet loss
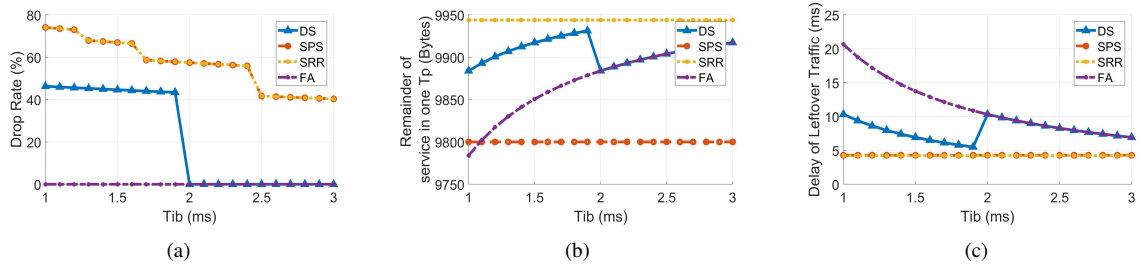
Fig. 4. (a) Haptic packet drop rate, (b) remainder of service, (c) leftover traffic upper bound delay, Vs. packet inter arrival time of haptic traffic during burst periods (T_{ib}), using different UL scheduling schemes with TTI = 0.5 ms.

is introduced. It can be concluded that for T_{ib} below 2 ms, only FA can be used as UL scheduling mechanism for haptic communication, and for T_{ib} between 2 and 3 ms, both FA and DS can be used.

Figure 4(b) shows the remainder of service (to be used by the leftover traffic) in one period of haptic traffic vs packet inter arrival time of haptic data during burst periods, i.e. T_{ib} at TTI = 0.5 ms (T_{pg} is still 5 ms). Observing from this figure, the remainder of service when SRR is used is higher than when SPS is used, since SRR releases the grant in non-burst periods. Another important observation is when SPS and SRR are used, the leftover service is independent of T_{ib}, since the pre-allocated periodic grant with the constant period of T_{pg} is occupied by the haptic traffic regardless of its data generation frequency.

Further observation from Figure 4(b) shows for T_{ib} larger than 2 ms, the remainder of service behaves very close when either of DS and FA are used; this is the region in which both DS and FA offer zero packet drop. Also note that by using DS or FA, the remainder of service increases with the increase in T_{ib}, since the larger T_{ib}, the fewer packets are generated by the haptic communication flow and thus, the more resource would remain for the leftover traffic. Given that the remainder of service does not show significant differences, deploying different UL scheduling techniques, we examine the upper bound delay offered by this service closely.

Observation from Figure 4(c) shows that the upper bound latency is affected more significantly. In other words, if the remainder of service is used for traffic with latency constraint, deploying any of the UL scheduling will have a major impact on its performance, while this is not the case for no delay constraint traffic. Figure 4(c) shows the upper bound delay of leftover traffic, computed based on Equation (12). As we can see, for T_{ib} between 2 and 3 ms, DS and FA perform similarly. In summary, results in Figures 4(a)-4(c) confirm the superiority of FA UL scheduling mechanism for haptic communication at TTI = 0.5 ms and also possibility of using DS for $T_{ib} \geq 2$ ms.

To see the impact of different TTI values, we also vary TTI from 0.125 ms to 1 ms, and plot upper bound delay for the leftover traffic and the packet drop rate of haptic communication flow for the four UL scheduling mechanisms (Figure 5). In general, lowering the TTI will lower the packet drop rate for haptic communication flow, while increasing the upper bound delay for the leftover traffic. For DS and FA, wherever the haptic drop rate is zero, the leftover delay is equal for all TTIs. The reason is that we assumed constant RB and m will change according to the Equation (4). Since m is inversely proportional to TTI, the effect of changing TTI is completely compensated (see Equation (14)). However, we can use FA for all TTIs in the whole range of T_{ib} since haptic packet drop rate is always zero, while we cannot use DS when TTI = 0.5 ms for $T_{ib} < 2$ ms and when TTI = 1 ms. Interestingly, we can use SPS and SRR when TTI = 0.125 ms for $T_{ib} \geq 1.3$ ms as well as when TTI = 0.25 ms for $T_{ib} \geq 2.5$ ms, but still the upper bound delay of the leftover traffic with those parameters are close to the upper bound delay of FA and DS. However, the maximum haptic packet delay would be 14TTI for SPS ans SRR, while it would be 7TTI for DS and 4TTI for FA.

Note that for the TTIs and T_{ib}s that SPS and SRR are usable, both are almost the same in terms of the upper bound delay for the leftover traffic, but for example observing the results for $T_{pg} = 4 \times$TTI at TTI = 0.125 ms shows that SRR results in 25% lower leftover delay than SPS (because of releasing the grant in non-burst periods).

VI. Concluding Remarks

We have reviewed different UL scheduling mechanisms that can be used for haptic communication. Those include Dynamic Scheduling (DS), Semi-Persistent Scheduling (SPS), Soft-Resource Reservation (SRR), and Fast UL Access (FA). For each mechanism, we have modeled remaining service and delay for the leftover traffic after consuming the resources by haptic packets. Using the model, we obtained the regions that each UL scheduling mechanism can be used for communication of haptic packets. we observed that FA can be used for all TTIs and all typical inter arrival time of haptic packets during the burst period (based on the traffic model in [14]) and also, this mechanism provides the lowest haptic packet radio access delay, while it has almost the same effect on the remainder of traffic (in terms of remainder service and upper bound of leftover delay) as other applicable UL scheduling mechanisms.
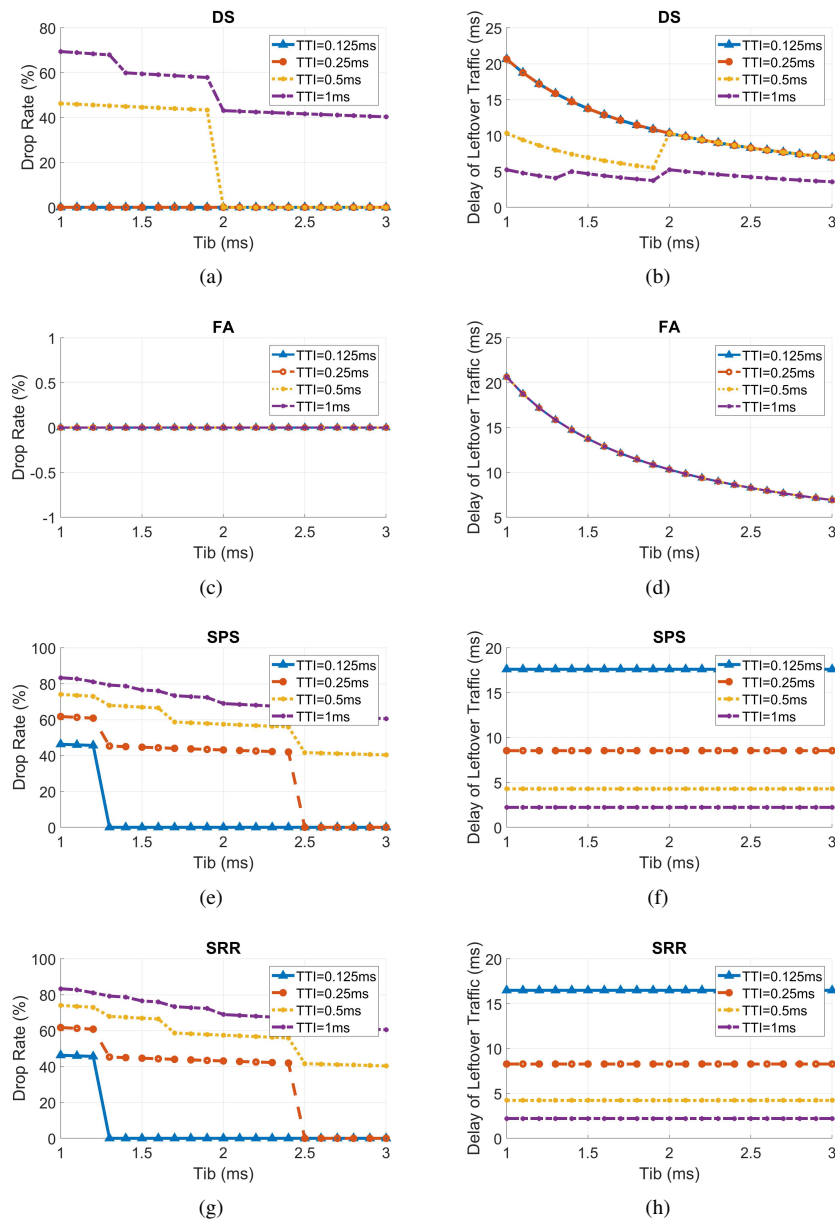
Fig. 5. Haptic packet drop rate (left) and leftover delay (right) Vs. packet inter arrival time of haptic traffic during burst periods (T_{ib}).

REFERENCES

[1] "Technical Specification Group Radio Access Network NR Physical Channels and Modulation (Release 15) ," tech. rep., TS 38.211, 3GPP, 2018.

[2] J. Sachs, G. Wikstrom, T. Dudda, R. Baldemair, and K. Kitti-chokechai, "5G Radio Network Design for Ultra-Reliable Low-Latency Communication," *IEEE Network*, vol. 32, no. 2, pp. 24–31, 2018.

[3] P. Arcara and C. Melchiorri, "Control Schemes for Teleoperation With Time Delay: A Comparative Study," *Robotics and Autonomous Systems*, vol. 38, no. 1, pp. 49–64, 2002.

[4] K. Antonakoglou, X. Xu, E. Steinbach, T. Mahmoodi, and M. Dohler, "Towards Haptic Communications over the 5G Tactile Internet," *IEEE Communications Surveys Tutorials*, 2018.

[5] M. Jiang, M. Condoluci, and T. Mahmoodi, "Network Slicing in 5G: An Auction-Based Model," in *IEEE ICC, Paris, France*, 2017.

[6] A. Patra, V. Pauli, and L. Yu, "Packet Scheduling for Real-Time Communication Over LTE Systems," in *IFIP Wireless Days, Valencia, Spain*, 2013.

[7] "Study on Latency Reduction Techniques for LTE (Release 14)," tech. rep., TR 36.881, 3GPP, 2016.

[8] J. B. Seo and V. C. M. Leung, "Performance Modeling and Stability of Semi-Persistent Scheduling With Initial Random Access in LTE," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4446–4456, 2012.

[9] M. Condoluci, T. Mahmoodi, E. Steinbach, and M. Dohler, "Soft Resource Reservation for Low-Delayed Teleoperation Over Mobile Networks," *IEEE Access*, vol. 5, pp. 10445–10455, 2017.

[10] M. Fidler, "An End-to-End Probabilistic Network Calculus With Moment Generating Functions," in *IEEE/ACM IWQoS , New Haven, CT, USA*, 2006.

[11] Y. Jiang, "A Basic Stochastic Network Calculus," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 4, pp. 123–134, 2006.

[12] Y. Zhang, Y. Gao, B. Gong, X. Zhang, and D. Yang, "Transmission Energy Consumption Analysis of a Multi-Channel Cognitive Radio Network Applying Stochastic Network Calculus," in *PIMRC, Washington, DC, USA*, 2014.

[13] X. Xu, Q. Liu, and E. G. Steinbach, "Toward QoE-Driven Dynamic Control Scheme Switching for Time-Delayed Teleoperation Systems: A Dedicated Case Study," *CoRR*, vol. abs/1705.05613, 2017.

[14] "Traffic Model for Legacy GPRS MTC." GP 160060, 3GPP GERAN meeting 69, 2016.