Multiplexing URLLC Traffic Within eMBB Services in 5G NR: Fair Scheduling

Hao Yin, Lyutianyang Zhang, Sumit Roy

Department of Electrical and Computer Engineering, University of Washington, Seattle, USA

February, 2021

Published in IEEE Transactions on Communications

System Model

- One base station, downlink transmission, OFDMA, eMBB and URLLC users.
- Saturated eMBB traffic [2]: Each eMBB user has infinite amount of data to be served.

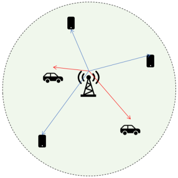


Figure: System model

Parameters

- Base frequency (use only when discuss channel model and do simulation, otherwise assume 0 for simplicity)

$$\mathfrak{f}_0 = 28 \quad [GHz] \tag{1}$$
$$= 28 \cdot 10^9 \quad [Hz] \tag{2}$$

- Number of time slots

$$\mathfrak{n} = 200 \tag{3}$$

- Number of time minislots per time slot

$$\mathfrak{m} = 7$$

- Position of base station

$$\mathfrak{x} = 500 \qquad [m] \tag{4}$$

$$\mathfrak{y} = 500 \qquad [m] \tag{5}$$

$$\mathfrak{z} = 100 \qquad [m] \tag{6}$$

- Number of eMBB users

$$\mathfrak{u} = 100 \tag{7}$$

- Begin position of eMBB users

$$X_{u,0} \sim Uniform(0, 1000) \qquad [m], \quad \forall u \tag{8}$$
$$Y_{u,0} \sim Uniform(0, 1000) \qquad [m], \quad \forall u \tag{9}$$
$$\mathfrak{z}_{u,0} = 0 \qquad [m], \quad \forall u \tag{10}$$

- End position of eMBB users

$$X_{u,\mathfrak{n}-1} \sim Uniform(0, 1000) \qquad [m], \quad \forall u \tag{11}$$
$$Y_{u,\mathfrak{n}-1} \sim Uniform(0, 1000) \qquad [m], \quad \forall u \tag{12}$$
$$\mathfrak{z}_{u,\mathfrak{n}-1} = 0 \qquad [m], \quad \forall u \tag{13}$$

- Number of URLLC users

$$\mathfrak{v} = 10 \tag{14}$$

- Begin position of URLLC users

$$X_{v,0,0} \sim Uniform(0, 1000) \qquad [m], \quad \forall v \tag{15}$$

$$Y_{v,0,0} \sim Uniform(0, 1000) \qquad [m], \quad \forall v \tag{16}$$

$$\mathfrak{z}_{v,0,0} = 0 \qquad [m], \quad \forall v \tag{17}$$

- End position of URLLC users

$$X_{v,\mathfrak{n}-1,\mathfrak{m}-1} \sim Uniform(0, 1000) \qquad [m], \quad \forall v \tag{18}$$

$$Y_{v,\mathfrak{n}-1,\mathfrak{m}-1} \sim Uniform(0, 1000) \qquad [m], \quad \forall v \tag{19}$$

$$\mathfrak{z}_{v,\mathfrak{n}-1,\mathfrak{m}-1} = 0 \qquad [m], \quad \forall v \tag{20}$$

- URLLC demand

$$R_{v,n,m} \sim Uniform(0, 4000)\left[\frac{bits}{minislot}\right], \quad \forall v, \forall n, \forall m \tag{21}$$

- URLLC demand peaks 500B per minislot, equivalently 14MBs.

- Channel bandwidth

$$\begin{aligned} \mathfrak{w}^{cn} &= 40 && [MHz] \\ &= 4 \cdot 10^7 && [Hz] \end{aligned}$$

- Numerology ($\mu \in \{0, 1\}$ is for Sub-6GHz)

$$\mu = 2$$

- Number of subchannels

$$\mathfrak{l} = 51$$

- Transmission power

$$\mathfrak{p}^{tx} = 100 \qquad [W] \tag{22}$$

$$= 20 \qquad [dB] \tag{23}$$

- Thermal noise density

$$\mathfrak{o} = -174 \qquad \left[\frac{dBm}{Hz}\right] \tag{24}$$

$$= -204 \qquad \left[\frac{dB}{Hz}\right] \tag{25}$$

Bandwidth

- Subcarrier spacing (SCS)

$$\zeta = 2^{\mu} \cdot 15 \quad [kHz] \tag{26}$$
$$= 60 \quad [kHz] \tag{27}$$

- Subchannel bandwidth

$$\mathfrak{w}^{sc} = 12 \cdot \zeta \quad [kHz] \tag{28}$$
$$= 2^{\mu} \cdot 180 \quad [kHz] \tag{29}$$
$$= 720 \quad [kHz] \tag{30}$$
$$= 7.2 \cdot 10^{5} \quad [Hz] \tag{31}$$

- Guard bandwidth

$$\mathfrak{w}^{gd} = \frac{\mathfrak{w}^{cn} - \mathfrak{l} \cdot \mathfrak{w}^{sc}}{2} \quad [Hz] \tag{32}$$

$$= 16.4 \cdot 10^{5} \quad [Hz] \tag{33}$$

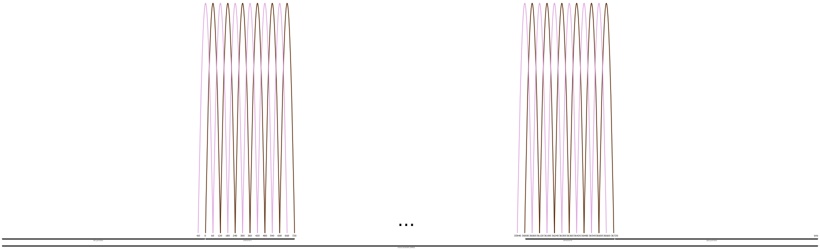


Figure: A Discrete Fourier Transform (DFT) coefficients example

Duration

- Time slot duration

$$
\begin{aligned}
\mathfrak{t}^{sl} &= \frac{1}{2^{\mu}} && [ms] \quad (34)\\
&= 0.25 && [ms] \quad (35)\\
&= 2.5 \cdot 10^{-4} && [s] \quad (36)
\end{aligned}
$$

- Time minislot duration

$$
\begin{aligned}
\mathfrak{t}^{ms} &= \frac{\mathfrak{t}^{sl}}{\mathfrak{m}} && [s] \quad (37)\\
&= \frac{1}{28} \cdot 10^{-3} && [s] \quad (38)
\end{aligned}
$$

Figure: Duration

OFDMA

- Chunk duration

$$t^{ck} = \frac{t^{sl}}{14} \quad [s] \tag{39}$$

$$= \frac{125}{7} \cdot 10^{-6} \quad [s] \tag{40}$$

- Symbol duration (enforced by DFT)

$$t^{sb} = \frac{1}{\zeta} \quad [ms] \tag{41}$$

$$= \frac{50}{3} \cdot 10^{-6} \quad [s] \tag{42}$$

- Cyclic prefix duration

$$\mathrm{t}^{cp} = \mathrm{t}^{ck} - \mathrm{t}^{sb} \qquad [s] \qquad (43)$$

$$= \frac{25}{21} \cdot 10^{-6} \qquad [s] \qquad (44)$$

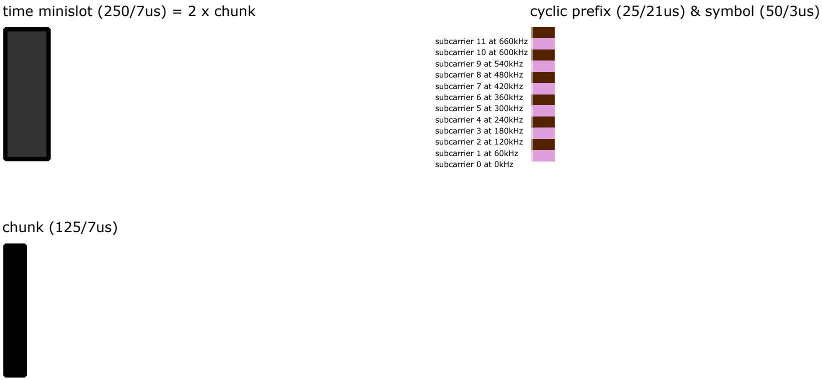


Figure: OFDMA

OFDM Analytical Concept

- Assume no cyclic prefix for simplicity

$$\mathfrak{t}^{ck} = \mathfrak{t}^{sb} = \frac{1}{\zeta} \tag{45}$$

Encoding

- Baseband modulation – binary phase shift keying (BPSK)
  - One symbol contains one bit
  - Non-quadrature communication i.e. no complex numbers

$$\theta(b) = \begin{cases} 0 & b = 0 \\ \pi & b = 1 \end{cases} [radians] \tag{46}$$

- Pulse

$$q(y,e,b,t) = \begin{cases} 10\cos(2\pi \frac{1}{\mathfrak{t}^{sb}} et + \theta(b)) & \mathfrak{t}^{sb} y \leq t < \mathfrak{t}^{sb}(y+1) \\ 0 & \text{otherwise} \end{cases} [V] \quad (47)$$

- where $y \in \mathbb{Z}_+, e \in \mathbb{Z}_+, b \in \{0,1\}, t \in \mathbb{R}$

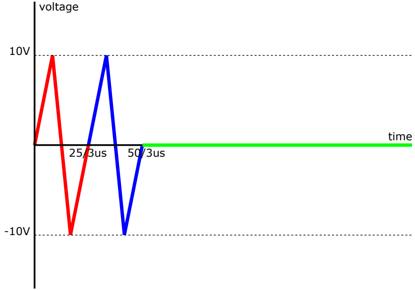


Figure: Pulse $q(0,2,0,t)$

- Do note that

$$\mathfrak{p}^{tx} = 10^2 \, [W] \tag{48}$$

- Passband modulation – frequency shifting
- The discussion followed considers $y_0 = 0, 1, \ldots, 2799$

- Subcarrier signal

$$\mathfrak{q}_{e,y_0}(t) = q(y_0, e, \mathfrak{b}_{e,y_0}, t) \tag{49}$$

$$= \begin{cases} 10\cos(2\pi\mathfrak{f}_e t + \theta(\mathfrak{b}_{e,y_0})) & \mathfrak{t}^{sb} y_0 \le t < \mathfrak{t}^{sb}(y_0+1) \\ 0 & \text{otherwise} \end{cases} \tag{50}$$

$$= \begin{cases} 10\cos(2\pi\mathfrak{f}_e t) & \mathfrak{t}^{sb} y_0 \le t < \mathfrak{t}^{sb}(y_0+1), \mathfrak{b}_{e,y_0} = 0 \\ 10\cos(2\pi\mathfrak{f}_e t + \pi) & \mathfrak{t}^{sb} y_0 \le t < \mathfrak{t}^{sb}(y_0+1), \mathfrak{b}_{e,y_0} = 1 \\ 0 & \text{otherwise} \end{cases} \tag{51}$$

$$\mathfrak{q}_{e,y_0}(t) = \begin{cases} 10\cos(2\pi \mathfrak{f}_e t) & \mathfrak{t}^{sb} y_0 \leq t < \mathfrak{t}^{sb}\left(y_0+1\right), \mathfrak{b}_{e,y_0} = 0 \\ -10\cos(2\pi \mathfrak{f}_e t) & \mathfrak{t}^{sb} y_0 \leq t < \mathfrak{t}^{sb}\left(y_0+1\right), \mathfrak{b}_{e,y_0} = 1 \; [V], \forall e \\ 0 & \text{otherwise} \end{cases} \tag{52}$$

- (Superpositioned) channel signal

$$\mathfrak{q}_{y_0}(t) = \sum_{e=0}^{611} \mathfrak{q}_{e,y_0}(t)$$
$$= \begin{cases} \sum_{e=0}^{611} 10\cos(2\pi\mathfrak{f}_e t + \theta(\mathfrak{b}_{e,y_0})) & \mathfrak{t}^{sb} y_0 \leq t < \mathfrak{t}^{sb}(y_0+1) \\ 0 & \text{otherwise} \end{cases}$$
$$[V]$$

Decoding

- Observation

$$
\begin{aligned}
&\mathfrak{q}_{y_0}(t)\\
=&10\cos(2\pi 0^{[kHz]}t+\theta(\mathfrak{b}_{0,y_0}))+\cdots+\\
&10\cos(2\pi 660^{[kHz]}t+\theta(\mathfrak{b}_{11,y_0}))\\
+&\dots\\
+&10\cos(2\pi 36000^{[kHz]}t+\theta(\mathfrak{b}_{600,y_0}))+\cdots+\\
&10\cos(2\pi 36660^{[kHz]}t+\theta(\mathfrak{b}_{611,y_0}))\\
=&\sum_{e=0}^{611}10\cos(2\pi\zeta e t+\theta(\mathfrak{b}_{e,y_0}))\,[V]\,,\quad \forall \mathfrak{t}^{sb}y_0\le t<\mathfrak{t}^{sb}(y_0+1)
\end{aligned}
$$

- Orthogonal basis

$$\forall i\colon \forall j\colon i \neq j \rightarrow \langle \cos(2\pi\zeta it), \cos(2\pi\zeta jt)\rangle = 0 \tag{53}$$

- Proof

$$\int_{\mathfrak{t}^{sb}y}^{\mathfrak{t}^{sb}y+\mathfrak{t}^{sb}} \cos(2\pi\zeta it)\cos(2\pi\zeta jt)dt \tag{54}$$

$$=\int_{0}^{\mathfrak{t}^{sb}} \cos(2\pi\zeta it)\cos(2\pi\zeta jt)dt \tag{55}$$

$$=\int_{0}^{\mathfrak{t}^{sb}} \frac{\cos(2\pi\zeta(i-j)t)+\cos(2\pi\zeta(i+j)t)}{2}dt \tag{56}$$

$$=\frac{1}{2}\left(\frac{\sin(2\pi\zeta(i-j)\mathfrak{t}^{sb})}{2\pi\zeta(i-j)}+\frac{\sin(2\pi\zeta(i+j)\mathfrak{t}^{sb})}{2\pi\zeta(i+j)}\right) \tag{57}$$

$$=0, \quad \forall i \neq j, \forall y \tag{58}$$

- By (53) and (52), $b_{0,y_0}, b_{1,y_0}, \ldots, b_{611,y_0}$ can be decoded from the following coefficients

$$
\begin{aligned}
&\mathfrak{q}_{y_0}(t) \\
=&\sum_{e=0}^{611} 10\cos(2\pi\zeta e t + \theta(\mathfrak{b}_{e,y_0})) \\
=&c_{0,y_0}\cos(2\pi 0^{[kHz]} t) + \cdots + c_{11,y_0}\cos(2\pi 660^{[kHz]} t) \\
+&\ldots \\
+&c_{600,y_0}\cos(2\pi 36000^{[kHz]} t) + \cdots + c_{611,y_0}\cos(2\pi 36660^{[kHz]} t), \\
&\forall \mathfrak{t}^{sb} y_0 \le t < \mathfrak{t}^{sb}(y_0 + 1)
\end{aligned}
$$

- Fourier's idea

$$c_{e,y_0} = \frac{\langle \mathfrak{q}_{y_0}(t), \cos(2\pi\zeta et)\rangle}{\langle \cos(2\pi\zeta et), \cos(2\pi\zeta et)\rangle} \tag{59}$$

$$= 2\zeta\langle \mathfrak{q}_{y_0}(t), \cos(2\pi\zeta et)\rangle, \quad \forall e \tag{60}$$

- Decode

$$b_{e,y_0} = \begin{cases} 0 & c_{e,y_0} \geq 0 \\ 1 & c_{e,y_0} < 0 \end{cases}, \quad \forall e \tag{61}$$

System Framework

- 3 eMBB users $\{0, 1, 2\}$.
- 2 URLLC users $\{0, 1\}$.

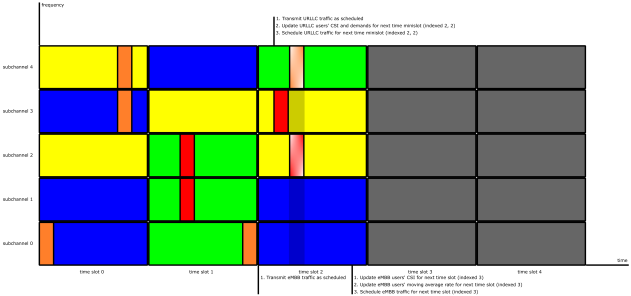


Figure: System framework

Channel Model

- Path loss
- Shadowing
- Thermal noise
- Co-channel interference
- Channel fading
- Inter-symbol interference (ISI)

Path Loss and Shadowing

- Subchannel frequency

$$\mathfrak{f}_l = \mathfrak{f}_0 + l\mathfrak{w}^{sc} \, [Hz] \, , \quad \forall l \tag{62}$$

- 3D distance

$$\mathfrak{d}_{u,n} = \sqrt{(X_{u,n} - \mathfrak{x})^2 + (Y_{u,n} - \mathfrak{y})^2 + (\mathfrak{z}_{u,n} - \mathfrak{z})^2} \quad [m] \, , \quad \forall u, \forall n$$

$$\mathfrak{d}_{v,n,m} = \sqrt{(X_{v,n,m} - \mathfrak{x})^2 + (Y_{v,n,m} - \mathfrak{y})^2 + (\mathfrak{z}_{v,n,m} - \mathfrak{z})^2} \quad [m] \, , \quad \forall v, \forall n, \forall m$$

- Path loss and shadowing – close-in (CI) free space reference distance model

$$h(f, d) = 20 \log_{10} \frac{4\pi f}{3 \cdot 10^8} + 10\epsilon^{pl} \log_{10} d + \sigma^{sd} \tag{63}$$

- Path loss exponent and shadowing standard deviation for urban macro-cellular (UMa) line-of-sight (LoS) over frequency and 3D distance ranging from 2 to 73.5GHz and 58 to 930m are derived via minimum mean square error (MMSE) fit by [3]

$$\epsilon^{pl} = 2 \tag{64}$$

$$\sigma^{sd} = 4.6\,[dB] \tag{65}$$

- Path loss and shadowing

$$\mathfrak{h}_{u,n,l} = h(\mathfrak{f}_l, \mathfrak{d}_{u,n}) \qquad [dB]\,, \quad \forall u, \forall n, \forall l \tag{66}$$

$$\mathfrak{h}_{v,n,m,l} = h(\mathfrak{f}_l, \mathfrak{d}_{v,n,m}) \qquad [dB]\,, \quad \forall v, \forall n, \forall m, \forall l \tag{67}$$

- Received power – path loss and shadowing model

$$\mathfrak{p}_{u,n,l} = \mathfrak{p}^{tx} - \mathfrak{h}_{u,n,l} \qquad [dB]\,, \quad \forall u, \forall n, \forall l \tag{68}$$

$$\mathfrak{p}_{v,n,m,l} = \mathfrak{p}^{tx} - \mathfrak{h}_{v,n,m,l} \qquad [dB]\,, \quad \forall v, \forall n, \forall m, \forall l \tag{69}$$

Thermal Noise and Co-channel Interference

- Subchannel thermal noise

$$\mathfrak{o}^{sc} = 10^{\frac{\mathfrak{o}}{10}} \mathfrak{w}^{sc} \quad [W] \tag{70}$$

$$= 10^{\frac{-204}{10}} \cdot 7.2 \cdot 10^5 \quad [W] \tag{71}$$

$$\approx -145 \quad [dB] \tag{72}$$

- Co-channel interference does not exist in this system model.

- Shannon-Hartley theorem

$$\mathfrak{r}_{u,n,l} = \mathfrak{w}^{sc} \log_2 \left(1 + 10^{\frac{\mathfrak{p}_{u,n,l} - \mathfrak{o}^{sc}}{10}}\right) \mathfrak{t}^{sl} \left[\frac{bits}{slot}\right],$$

$$\forall u, \forall n, \forall l$$

$$\mathfrak{r}_{v,n,m,l} = \mathfrak{w}^{sc} \log_2 \left(1 + 10^{\frac{\mathfrak{p}_{v,n,m,l} - \mathfrak{o}^{sc}}{10}}\right) \mathfrak{t}^{ms} \left[\frac{bits}{minislot}\right],$$

$$\forall v, \forall n, \forall m, \forall l$$

Channel Fading and ISI

- Channel fading is resolved by OFDM's cyclic prefix.
- ISI is resolved by OFDM's multi-carrier and cylic prefix.
  - Multi-carrier has relatively long symbol duration compared to single-carrier transmitting same number of symbols over same frequency bandwidth during same time duration.
  - Cyclic prefix also serves as inter-symbol guard duration.
- Note on resolving ISI using equalization:
  - Channel gain is first measured as received signal of a unit impulse transmission, and then used with received signal to decode transmitted signal (I do not understand how this is done).
  - OFDM performs equalization in frequency domain (I do not understand how this is done).

Offline URLLC Puncturing

- The system maximizes eMBB traffic's total average rate and fairness (73a).
- For each time slot, the system allocates at most one eMBB user to each subchannel (73b).
- For each time slot, the system either schedules or un-schedules a subchannel to each eMBB user (73c).
- For each time minislot, the system allocates at most one URLLC user to each subchannel. Also, it schedules l^{th} subchannel from u^{th} eMBB user to a URLLC user only if it schedules the subchannel to the eMBB user (73d).
- For each time minislot, the system serves URLLC demands without delay (73e).
- For each time minislot, the system employs URLLC puncturing instead of superposition (73f).

$$
\begin{aligned}
\underset{\boldsymbol{\alpha},\boldsymbol{\beta}}{\text{maximize}} \quad & \sum_{u} \ln \bar{r}_u && && \text{(73a)}\\
\text{subject to} \quad & \sum_{u} \alpha_{u,n,l} \leq 1, && \forall n, \forall l, && \text{(73b)}\\
& \alpha_{u,n,l} \in \{0,1\}, && \forall u, \forall n, \forall l, && \text{(73c)}\\
& \sum_{v} \beta_{v,u,n,m,l} \leq \alpha_{u,n,l}, && \forall u, \forall n, \forall m, \forall l, && \text{(73d)}\\
& r_{v,n,m} \geq R_{v,n,m}, && \forall v, \forall n, \forall m, && \text{(73e)}\\
& \beta_{v,u,n,m,l} \in \{0,1\}, && \forall v, \forall u, \forall n, \forall m, \forall l && \text{(73f)}
\end{aligned}
$$

where

$$\bar{r}_u = \frac{1}{\mathfrak{n}} \sum_{n,m,l} \left(\alpha_{u,n,l} - \sum_{v} \beta_{v,u,n,m,l} \right) \frac{\mathfrak{r}_{u,n,l}}{\mathfrak{m}} \qquad \left[\frac{bits}{slot} \right], \quad (74)$$

$$\forall u \quad (75)$$

$$r_{v,n,m} = \sum_{u,l} \beta_{v,u,n,m,l} \mathfrak{r}^{sc}_{v,n,m} \qquad \left[\frac{bits}{minislot} \right], \quad (76)$$

$$\forall v, \forall n, \forall m \quad (77)$$

$$\mathfrak{r}^{sc}_{v,n,m} = \min_{l} \left\{ \mathfrak{r}_{v,n,m,l} \right\} \qquad \left[\frac{bits}{minislot} \right] \quad (78)$$

$$\forall v, \forall n, \forall m \quad (79)$$

Offline Multiplexing URLLC Puncturing – eMBB

- Schedule eMBB traffic using only eMBB channel state information (CSI)

$$\begin{aligned} \underset{\boldsymbol{\alpha}}{\text{maximize}} \quad & \sum_u \ln \bar{r}_u && && \text{(80a)} \\ \text{subject to} \quad & \sum_u \alpha_{u,n,l} \leq 1, && \forall n, \forall l, && \text{(80b)} \\ & \alpha_{u,n,l} \in \{0,1\}\,, && \forall u, \forall n, \forall l && \text{(80c)} \end{aligned}$$

- By first relaxing the binary constraint (80c) and then employing subgradient method, we should be able to prove by [2] and by total unimodularity that proportional fairness (PF) scheduling algorithm gives an asymptotically optimal solution.

Online Multiplexing URLLC Puncturing – eMBB

- Relax binary constraint

$$\underset{\boldsymbol{\alpha}'}{\text{maximize}} \quad \sum_u \ln \bar{r}'_u \tag{81a}$$

$$\text{subject to} \quad \sum_u \alpha'_{u,n,l} \le 1, \quad \forall n, \forall l, \tag{81b}$$

$$\alpha'_{u,n,l} \ge 0, \quad \forall u, \forall n, \forall l \tag{81c}$$

- Note that the constraint $\alpha'_{u,n,l} \le 1, \quad \forall u, \forall n, \forall l$ is implied for this problem.
- The optimal value of this problem is always greater than the optimal value of problem (80) (proof needed).

- Since the feasible average rate set is convex (proof needed), [2] shows that the following scheduling policy is asymptotically optimal: For $n_0 = 0, 1, \ldots, \mathfrak{n} - 1$

$$\boldsymbol{\alpha}'^{*}_{n_0} \in \arg\max_{\boldsymbol{\alpha}'_{n_0}} \left\{ \nabla_g(\tilde{\mathbf{r}}'_{n_0})^{\mathsf{T}} \mathbf{r}'_{n_0} \,\middle|\, (81b), (81c) \right\} \tag{82}$$

- where

$$g : \mathbb{R}_+^{\mathfrak{u}} \longrightarrow \mathbb{R} \tag{83}$$

$$\mathbf{r} \longmapsto \sum_u \ln r_u \tag{84}$$

- For $n_0 = 0, 1, \ldots, \mathfrak{n} - 1$

$$\begin{aligned} \underset{\boldsymbol{\alpha}'_{n_0}}{\text{maximize}} \quad & \sum_u \frac{1}{\tilde{r}'_{u,n_0}} r'_{u,n_0} && \text{(85a)} \\ \text{subject to} \quad & \sum_u \alpha'_{u,n_0,l} \leq 1, \quad \forall l, && \text{(85b)} \\ & \alpha'_{u,n_0,l} \geq 0, \quad \forall u, \forall l && \text{(85c)} \end{aligned}$$

- where

$$\tilde{r}'_{u,n_0} = \begin{cases} 1 & n_0 = 0 \\ (1-\epsilon)\,\tilde{r}'_{u,n_0-1} + \epsilon r'_{u,n_0-1} & \text{otherwise} \end{cases}, \left[\frac{bits}{slot}\right] \quad \forall u \tag{86}$$

$$\epsilon = 0.1 \tag{87}$$

- Since linear program (85) is totally unimodular (proof needed), it has binary solution(s).
- Combining the above statements, binary solution(s) of problem (85) gives asymtotically maximum objective value for problem (80) (proof needed).
- Also, do notice that the PF scheduling algorithm optimizes problem (85)[1].

[1] Example

Online Multiplexing URLLC Puncturing – URLLC

- For $m_0 = 0, 1, \ldots, \mathfrak{m} - 1$ in $n_0 = 0, 1, \ldots, \mathfrak{n} - 1$

$$
\begin{aligned}
\underset{\boldsymbol{\beta}_{n_0,m_0}}{\text{maximize}} \quad & \sum_u \ln \tilde{\phi}_{u,n_0,m_0} && && \text{(88a)}\\
\text{subject to} \quad & \sum_v \beta_{v,n_0,m_0,l} \leq 1, && \forall l, && \text{(88b)}\\
& \sum_l \beta_{v,n_0,m_0,l} = D_{v,n_0,m_0}, && \forall v, && \text{(88c)}\\
& \beta_{v,n_0,m_0,l} \in \{0,1\}, && \forall v, \forall l && \text{(88d)}
\end{aligned}
$$

- where

$$
\tilde{\phi}_{u,n_0,m_0} = (1-\epsilon)\,\tilde{\psi}_{u,n_0-1} + \epsilon\phi_{u,n_0,m_0} \left[\frac{bits}{minislot}\right], \quad \forall u \quad (89)
$$

- Punctured eMBB moving average minislot rate

$$\tilde{\psi}_{u,n_0-1} = \begin{cases} 0 & n_0 = 0 \\ \frac{1}{\mathfrak{m}} & n_0 = 1 \\ (1-\epsilon)\,\tilde{\psi}_{u,n_0-2} + \epsilon \frac{r_{u,n_0-2}}{\mathfrak{m}} & \text{otherwise} \end{cases} \tag{90}$$

$$\left[\frac{bits}{minislot}\right], \quad \forall u \tag{91}$$

- Modified eMBB 'peak' minislot rate

$$\varphi_{u,n_0,m_0} = \begin{cases} \frac{r'^{*}_{u,n_0}}{\mathfrak{m}} & m_0 = 0 \\ \phi_{u,n_0,m_0-1} & \text{otherwise} \end{cases} \tag{92}$$

$$\left[\frac{bits}{minislot}\right], \quad \forall u \tag{93}$$

- Modified eMBB minislot rate

$$\phi_{u,n_0,m_0} = \left(1 - \frac{\sum_l \left(\alpha_{u,n_0,l} \wedge \sum_v \beta_{v,n_0,m_0,l}\right)}{\sum_l \alpha_{u,n_0,l}}\right) \varphi_{u,n_0,m_0} \tag{94}$$

$$\left[\frac{bits}{minislot}\right], \quad \forall u \tag{95}$$

$$\tag{96}$$

- Linearize modified eMBB minislot rate in β_{n_0,m_0} [2]

$$\phi_{u,n_0,m_0} = \left(1 - \frac{\langle B_{n_0,m_0}(:\,,col(A_{u,n_0})), 1\rangle}{\sum_l \alpha_{u,n_0,l}}\right) \varphi_{u,n_0,m_0} \tag{97}$$

$$\left[\frac{bits}{minislot}\right], \quad \forall u \tag{98}$$

[2] Example

Convex

- Equivalent program 1

$$\underset{\boldsymbol{\beta}_{n_0,m_0}}{\text{maximize}} \quad \sum_u \ln \tilde{\phi}_{u,n_0,m_0} \tag{99a}$$

$$\text{subject to} \quad \sum_v \beta_{v,n_0,m_0,l} \leq 1, \quad \forall l, \tag{99b}$$

$$\sum_l \beta_{v,n_0,m_0,l} = D_{v,n_0,m_0}, \quad \forall v, \tag{99c}$$

$$\beta_{v,n_0,m_0,l}\left(1 - \beta_{v,n_0,m_0,l}\right) = 0, \quad \forall v, \forall l \tag{99d}$$

Convex (Continued)

- Equivalent program 2

$$\underset{\boldsymbol{\beta}_{n_0,m_0}}{\text{maximize}} \quad \sum_u \ln \tilde{\phi}_{u,n_0,m_0} \tag{100a}$$

$$\text{subject to} \quad \sum_v \beta_{v,n_0,m_0,l} \leq 1, \quad \forall l, \tag{100b}$$

$$\sum_l \beta_{v,n_0,m_0,l} = D_{v,n_0,m_0}, \quad \forall v, \tag{100c}$$

$$\beta_{v,n_0,m_0,l} \left(1 - \beta_{v,n_0,m_0,l}\right) \geq 0, \quad \forall v, \forall l, \tag{100d}$$

$$\sum_{v,l} \beta_{v,n_0,m_0,l} \left(1 - \beta_{v,n_0,m_0,l}\right) = 0 \tag{100e}$$

Convex (Continued)

- Equivalent program 3

$$\underset{\boldsymbol{\beta}_{n_0,m_0}}{\text{maximize}} \quad \sum_{u} \ln \tilde{\phi}_{u,n_0,m_0} \tag{101a}$$

$$\text{subject to} \quad \sum_{v} \beta_{v,n_0,m_0,l} \leq 1, \quad \forall l, \tag{101b}$$

$$\sum_{l} \beta_{v,n_0,m_0,l} = D_{v,n_0,m_0}, \quad \forall v, \tag{101c}$$

$$\beta_{v,n_0,m_0,l} \geq 0, \quad \forall v, \forall l, \tag{101d}$$

$$\sum_{v,l} \beta_{v,n_0,m_0,l} \left(1 - \beta_{v,n_0,m_0,l}\right) = 0 \tag{101e}$$

Convex (Continued)

- Equivalent program 4

$$\underset{\boldsymbol{\beta}_{n_0,m_0}}{\text{minimize}} \quad -\sum_{u} \ln \tilde{\phi}_{u,n_0,m_0} \tag{102a}$$

$$\text{subject to} \quad \sum_{v} \beta_{v,n_0,m_0,l} \leq 1, \quad \forall l, \tag{102b}$$

$$\sum_{l} \beta_{v,n_0,m_0,l} = D_{v,n_0,m_0}, \quad \forall v, \tag{102c}$$

$$\beta_{v,n_0,m_0,l} \geq 0, \quad \forall v, \forall l, \tag{102d}$$

$$\sum_{v,l} \beta_{v,n_0,m_0,l} \left(1 - \beta_{v,n_0,m_0,l}\right) = 0 \tag{102e}$$

Convex (Continued)

- Polyhedron

$$\mathscr{P} = \left\{ \boldsymbol{\beta}_{n_0,m_0} \middle| \begin{aligned} \sum_{v} \beta_{v,n_0,m_0,l} &\leq 1 \\ \sum_{l} \beta_{v,n_0,m_0,l} &= D_{v,n_0,m_0} \\ \beta_{v,n_0,m_0,l} &\geq 0 \end{aligned} \right\} \tag{103}$$

Convex (Continued)

- Equivalent program 5

$$\underset{\boldsymbol{\beta}_{n_0,m_0}}{\text{minimize}} \quad -\sum_{u} \ln \tilde{\phi}_{u,n_0,m_0} \tag{104a}$$

$$\text{subject to} \quad \boldsymbol{\beta}_{n_0,m_0} \in \mathscr{P}, \tag{104b}$$

$$\sum_{v,l} \beta_{v,n_0,m_0,l}\left(1 - \beta_{v,n_0,m_0,l}\right) = 0 \tag{104c}$$

Convex (Continued)

- Feasible solution

$$\boldsymbol{\beta}^{fs} \in \mathscr{P} \cap \left\{ \boldsymbol{\beta}_{n_0,m_0} \middle| \sum_{v,l} \beta_{v,n_0,m_0,l} \left(1 - \beta_{v,n_0,m_0,l}\right) = 0 \right\} \quad (105)$$

Convex (Continued)

- Penalized program

$$\begin{aligned} \underset{\boldsymbol{\beta}_{n_0,m_0}}{\text{minimize}} \quad & -\sum_{u} \ln \tilde{\phi}_{u,n_0,m_0} + \xi \sum_{v,l} \beta_{v,n_0,m_0,l} \left(1 - \beta_{v,n_0,m_0,l}\right) && \text{(106a)} \\ \text{subject to} \quad & \boldsymbol{\beta}_{n_0,m_0} \in \mathscr{P}, && \text{(106b)} \\ & \sum_{v,l} \beta_{v,n_0,m_0,l} \left(1 - \beta_{v,n_0,m_0,l}\right) \geq 0 && \text{(106c)} \end{aligned}$$

- where $\xi > \frac{\left(-\sum_u \ln \tilde{\phi}_{u,n_0,m_0}\right)^{\text{at } \boldsymbol{\beta}^{fs}} - \inf_{\boldsymbol{\beta}_{n_0,m_0}} \left\{-\sum_u \ln \tilde{\phi}_{u,n_0,m_0} \middle| \boldsymbol{\beta}_{n_0,m_0} \in \mathscr{P}, \sum_{v,l} \beta_{v,n_0,m_0,l}\left(1-\beta_{v,n_0,m_0,l}\right) \geq 0\right\}}{\inf_{\boldsymbol{\beta}_{n_0,m_0}} \left\{\sum_{v,l} \beta_{v,n_0,m_0,l}\left(1-\beta_{v,n_0,m_0,l}\right) \middle| \boldsymbol{\beta}_{n_0,m_0} \in vertices(\mathscr{P}), \sum_{v,l} \beta_{v,n_0,m_0,l}\left(1-\beta_{v,n_0,m_0,l}\right) > 0\right\}} \geq 0$
- In practice, ξ is set to a very large number.

Convex (Continued)

Theorem 1 [1]

If $-\sum_u \ln \tilde{\phi}_{u,n_0,m_0}$ and $\sum_{v,l} \beta_{v,n_0,m_0,l}(1-\beta_{v,n_0,m_0,l})$ are concave in $\boldsymbol{\beta}_{n_0,m_0}$, then problem (104) and (106) are equivalent.

Convex (Continued)

- 'Equivalent' program 6

$$
\begin{aligned}
\underset{\boldsymbol{\beta}_{n_0,m_0}}{\text{minimize}} \quad & -\sum_{u} \ln \tilde{\phi}_{u,n_0,m_0} + \xi \sum_{v,l} \beta_{v,n_0,m_0,l}\left(1 - \beta_{v,n_0,m_0,l}\right) \\
\text{subject to} \quad & \sum_{v} \beta_{v,n_0,m_0,l} \leq 1, \qquad \forall l, \\
& \sum_{l} \beta_{v,n_0,m_0,l} = D_{v,n_0,m_0}, \qquad \forall v, \\
& \beta_{v,n_0,m_0,l} \geq 0, \qquad \forall v, \forall l, \\
& \sum_{v,l} \beta_{v,n_0,m_0,l}\left(1 - \beta_{v,n_0,m_0,l}\right) \geq 0
\end{aligned}
$$

Convex (Continued)

- 'Equivalent' program 7

$$
\begin{aligned}
\underset{\boldsymbol{\beta}_{n_0,m_0}}{\text{minimize}} \quad & -\sum_{u} \ln \tilde{\phi}_{u,n_0,m_0} + \xi \sum_{v,l} \beta_{v,n_0,m_0,l}\left(1-\beta_{v,n_0,m_0,l}\right) && \text{(107a)}\\
\text{subject to} \quad & \sum_{v} \beta_{v,n_0,m_0,l} \le 1, \quad \forall l, && \text{(107b)}\\
& \sum_{l} \beta_{v,n_0,m_0,l} = D_{v,n_0,m_0}, \quad \forall v, && \text{(107c)}\\
& \beta_{v,n_0,m_0,l} \ge 0, \quad \forall v, \forall l && \text{(107d)}
\end{aligned}
$$

Convex (Continued)

- Define

$$g'(\boldsymbol{\beta}_{n_0,m_0}) = \sum_{v,l} \beta_{v,n_0,m_0,l}\left(1 - \beta_{v,n_0,m_0,l}\right) \tag{108}$$

- Since g' is a differentiable concave function

$$g'(\boldsymbol{\beta}_{n_0,m_0}) \leq g'(\boldsymbol{\beta}^{fs}) + \nabla_{g'}(\boldsymbol{\beta}^{fs})^{\mathsf{T}}\left(\boldsymbol{\beta}_{n_0,m_0} - \boldsymbol{\beta}^{fs}\right) \tag{109}$$

$$= \sum_{v,l} \beta^{fs}_{v,l}\left(1 - \beta^{fs}_{v,l}\right) + \sum_{v,l}\left(1 - 2\beta^{fs}_{v,l}\right)\left(\beta_{v,n_0,m_0,l} - \beta^{fs}_{v,l}\right) \tag{110}$$

$$= approx(\boldsymbol{\beta}_{n_0,m_0}) \tag{111}$$

Convex (Continued)

- Subroutine solving difference of convex (DC) programming

$$
\begin{aligned}
\underset{\boldsymbol{\beta}_{n_0,m_0}}{\text{minimize}} \quad & -\sum_u \ln \tilde{\phi}_{u,n_0,m_0} + \xi approx(\boldsymbol{\beta}_{n_0,m_0}) && \text{(112a)} \\
\text{subject to} \quad & \sum_v \beta_{v,n_0,m_0,l} \leq 1, \quad \forall l, && \text{(112b)} \\
& \sum_l \beta_{v,n_0,m_0,l} = D_{v,n_0,m_0}, \quad \forall v, && \text{(112c)} \\
& \beta_{v,n_0,m_0,l} \geq 0, \quad \forall v, \forall l && \text{(112d)}
\end{aligned}
$$

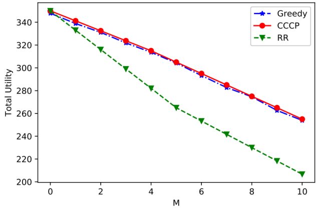
Greedy

- The idea is to puncture ‘wealthies’ eMBB users in the current time slot[3].

[3] Example

Issues

- Due to missing of constraint (88c), the obvious optimal solution of the convex method is to not allocate any subchannels to URLLC users.
- It is then questionable why greedy method can perform as good as convex one.



(a) Total utility vs URLLC users M per mini-slot.

Figure: Experiment result

Issues – Convex

- Since $-\sum_u \ln \tilde{\phi}_{u,n_0,m_0}$ is not concave in $\boldsymbol{\beta}_{n_0,m_0}$, theorem 1 is not applicable, and hence the penalized program's optimality is questionable[4].
- Even if we assume the penalized program's optimality, the subroutine solving it does not have optimality bounds (in single cell case, my algorithm is optimal).
- There is no theory backing up their objective function for URLLC's online multiplexing problem (my model possesses analytical justification).

[4] Example

Issues – Greedy

- In greedy algorithm, unlike PF, eMBB subchannel-wise rates are not considered when puncturing (94) (do note that my model does take this into account (74)).
- In greedy algorithm, moving average rate employs outdated information (fixing this is straightforward).
- Greedy algorithm does not have optimality bounds (in multicell case, my algorithm has a guaranteed approximation ratio).

Contributions

- Based on the existing eMBB scheduler, a model for joint scheduling of punctured eMBB and URLLC traffic as an optimization problem that maximizes the eventual aggregate utility of the eMBB users subject to latency constraints for the URLLC users.
- Model for the delay and reliability of URLLC traffic from a media access control (MAC) layer perspective.
- Two new resource allocation algorithms to align with practical implementation for downlink scheduling in the 5G system.

References

[1] Thi Hoai An Le, Dinh Tao Pham, and Huynh Van Ngai. "Exact Penalty and Error Bounds in DC Programming". In: *Journal of Global Optimization* 52.3 (2012), pp. 509–535. DOI: 10.1007/s10898-011-9765-3.

[2] Alexander L. Stolyar. "On the Asymptotic Optimality of the Gradient Scheduling Algorithm for Multiuser Throughput Allocation". In: *Operations Research* 53.1 (2005), pp. 12–25. DOI: 10.1287/opre.1040.0156.

[3] Shu Sun et al. "Propagation Path Loss Models for 5G Urban Micro- and Macro-Cellular Scenarios". In: *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*. 2016, pp. 1–6. DOI: 10.1109/VTCSpring.2016.7504435.