# Short Packet Communication with NOMA

**Varun Varindani**

Indian Institute of Science, Bangalore

April 6, 2019

# Introduction

- IoT to connect massive machines and devices in the future communication networks
- Short-packet with finite blocklength codes is considered to reduce the transmission latency
- Decoding error probability is not negligible since the blocklength is small

# NOMA

- NOMA schemes rely on a key assumption that the users' channel conditions are very different.
- Within the NOMA user pair, one user is assumed to be deployed close to the base station, and the other is far away from the base station
- NOMA offers great potential to achieve low-latency by serving multiple users simultaneously

# System Model

- Consider that two users, $u_i$, $i \in 1, 2$ are paired to perform NOMA
- Let $|h_1|^2 \geq |h_2|^2$
- $\alpha_i$, the power allocation coefficient satisfying $\alpha_1 \leq \alpha_2$ and, $\alpha_1 + \alpha_2 = 1$
- The received signals at $u_i$ is given by

$$y_i = h_i \left( \sqrt{\alpha_1 P} s_1 + \sqrt{\alpha_2 P} s_2 \right) + n_i \tag{1}$$

# Cont.

- At $u_2$, it decodes $s_2$ directly by treating $s_1$ as interference.
- Then the received SINR of decoding $s_2$ at $u_2$ is given by

$$\gamma_{22} = \alpha_2 |h_2|^2 / \left( \alpha_1 |h_2|^2 + 1/\rho \right) \quad (2)$$

where $\rho = P/\sigma^2$ is the transmit signal-to-noise ratio (SNR).

# Cont.

- Given the blocklength $m$, the number of data bits $N_i$ to $u_i$, the Shannon capacity $C(x) = log_2(1+x)$ and the channel dispersion $V(x)$
- The instantaneous BLER of decoding $s_2$ at $u_2$, denoted by $\epsilon_2$,

$$\epsilon_2 = \epsilon_{22} \approx Q\left(\frac{C(\gamma_{22}) - \frac{N_2}{m}}{\sqrt{V(\gamma_{22}/m)}}\right) = \varphi(\gamma_{22}, N_2, m) \qquad (3)$$

- The received SINR of decoding $s_2$ at $u_1$ is given by

$$\gamma_{12} = \alpha_2|h_1|^2/\left(\alpha_1|h_1|^2 + 1/\rho\right) \qquad (4)$$

$$\epsilon_{12} \approx \varphi(\gamma_{12}, N_2, m) \qquad (5)$$

# Cont.

- If $s_2$ can be successfully decoded and removed, $u_1$ then can decode $s_1$ in an interference-free manner with

$$\gamma_{11} = \alpha_1 |h_1|^2 \rho \tag{6}$$

$$\epsilon_1 \approx \varphi(\gamma_{11}, N_1, m) \tag{7}$$

- the instantaneous overall BLER at $u_1$ can be approximated as

$$\epsilon_1 = \epsilon_{12} + (1 - \epsilon_{12})\epsilon_{11} \approx \epsilon_{12} + \epsilon_{11} \tag{8}$$

# Problem Statement

- To reduce the blocklength,

$$\mathbf{P}: m^* = \min_{\alpha_1,\alpha_2} M(\alpha_1, \alpha_2) \quad (9)$$

$$\text{s.t. } \bar{\epsilon}_i \le \bar{\epsilon}_i^{th}, i \in \{1, 2\} \quad (10)$$

$$0 < \alpha_i < 1, \alpha_1 + \alpha_2 = 1 \quad (11)$$

- Noting that $\alpha_1 = 1 - \alpha_2$, $\bar{\epsilon}_i$ is a decreasing function of $m$, we simplify above problem as

$$\mathbf{P}: m^* = \min_{\alpha_1} M(\alpha_1) \quad (12)$$

$$\text{s.t. } \bar{\epsilon}_i = \bar{\epsilon}_i^{th}, i \in \{1, 2\} \quad (13)$$

$$0 < \alpha_1 < 1 \quad (14)$$

# Problem Solution

- Closed-form expression of the average BLER $\bar{\epsilon}_{ij}$,

$$\bar{\epsilon}_{ij} \approx \int_0^{\infty} Q\left(\frac{C(\gamma_{22}) - \frac{N_2}{m}}{\sqrt{V(\gamma_{22}/m)}}\right) f_{\gamma_{ij}}(x)dx, \quad i,j = \{1,2\} \quad (15)$$

- By using the first order Riemann integral approximation,

$$\bar{\epsilon}_2 \approx 2\beta_{2,m}/\big(\lambda\rho(\alpha_2 - \alpha_1\beta_{2,m})\big) \quad (16)$$

where $\beta_{j,m} = 2^{\frac{N_j}{m}} - 1$

- The minimum blocklength $m_2$ for $u_2$,

$$m_2 \approx N_2/log_2\left(\left(2+\lambda\rho\bar{\epsilon}_2^{th}\right)/\left(2+\alpha_1\lambda\rho\bar{\epsilon}_2^{th}\right)\right) \quad (17)$$

- Similarly, $\bar{\epsilon}_1$ can be approximated as

$$\bar{\epsilon}_1 \approx \big(\beta_{1,m}/(\lambda\rho\alpha_1)\big)^2 + (\bar{\epsilon}_2^{th})^2 \quad (18)$$

- The minimum blocklength $m_1$ for $u_1$,

$$m_1 \approx N_1/log_2\left(1+\alpha_1\lambda\rho\sqrt{\bar{\epsilon}_1^{th}-(\bar{\epsilon}_2^{th}/2)^2}\right) \quad (19)$$

- **Lemma 1:** The optimal power allocation coefficient $\alpha_1^*$ can be solved under the condition that $m_1 = m_2$.
- The optimal power allocation coefficient $\alpha_1^*$ can be approximated as

$$\alpha_1^* \approx \left(\left(\left(2+\lambda\rho\bar{\epsilon}_2^{th}\right)/2\right)^{\frac{N_1}{N_2}}-1\right)/\left(\lambda\rho\sqrt{\bar{\epsilon}_1^{th}}\right) \qquad (20)$$

- The minimum common blocklength to guarantee the target average reliability of users,

$$m^* \approx N_2/log_2\left(\left(2+\lambda\rho\bar{\epsilon}_2^{th}\right)/\left(2+\alpha_1^*\lambda\rho\bar{\epsilon}_2^{th}\right)\right) \qquad (21)$$

# Comparison with OMA

- In conventional OMA (e.g., TDMA) systems, the minimum sum blocklength $\hat{m}^*$ is the summation of $\hat{m}_i$ for $i \in \{1, 2\}$

$$\hat{m}_1 \approx N_1/log_2\left(1 + \lambda\rho\sqrt{\bar{\epsilon}_1^{th}}\right) \tag{22}$$

$$\hat{m}_2 \approx N_2/log_2\left(\left(2 + \lambda\rho\bar{\epsilon}_2^{th}\right)/2\right) \tag{23}$$

- The blocklength reduction of NOMA compared to OMA can be characterized as

$$\triangle_m = \hat{m}^* - m^* \approx N_1/log_2\left(1 + \lambda\rho\sqrt{\bar{\epsilon}_1^{th}}\right) = \hat{m}_1 \tag{24}$$

# Remarks

- Given the reliability constraints, OMA needs $\hat{m}_1 + \hat{m}_2$ channel uses to serve $u_1$ and $u_2$. However, in NOMA, by properly optimizing the power allocation among the served users, only $\hat{m}_2$ channel uses are needed to serve $u_1$ and $u_2$ simultaneously.

# Another Problem

- Effective throughput at $u_i$ is defined by

$$\bar{T}_i = \frac{N_i}{N} R_i (1 - \bar{\epsilon}_i) \tag{25}$$

- Transmission rate at $u_i$ can be approximated by

$$R_i \approx log_2(1 + \gamma_i) - \sqrt{\frac{V_i}{N_i}} \frac{Q^{-1}(\epsilon_i)}{ln2} \tag{26}$$

where $V_i$ is the channel dispersion

- For a given transmission rate $R_i$, the decoding error probability at user $i$ is approximated by

$$\epsilon_i \approx Q(f(\gamma_i, N_i, R_i)) \tag{27}$$

where $f(\gamma_i, N_i, R_i) = \ln 2 \sqrt{\frac{N_i}{1-(1+\gamma_i)^{-2}}}(\log_2(1+\gamma_i) - R_i)$

- Optimization problem

$$\max_{\triangle} \bar{T}_1 \tag{28}$$

$$\bar{T}_2 \geq T_0 \tag{29}$$

$$P_1 N_1 + P_2 N_2 \leq PN \tag{30}$$

$$N_1 = N_2 = N, \text{ for NOMA} \tag{31}$$

$$N_1 + N_2 = N, \text{ for OMA} \tag{32}$$

- Optimization problem for NOMA

$$\max_{\triangle} \bar{T}_1 \tag{33}$$

$$P_1 + P_2 \leq P \tag{34}$$

$$\bar{T}_2 \geq T_0 \tag{35}$$

- The effective throughput at $u_1$ is given by

$$\bar{T}_1 = R_1(1 - \epsilon_1 + \epsilon_1 \epsilon_2^1 - \epsilon_2^1 \epsilon_1^{'}) \tag{36}$$

- The effective throughput at $u_2$ is given by

$$\bar{T}_1 = R_2(1 - \epsilon_2) \tag{37}$$

# Cont.

- **Lemma 1:**The equality in the power constraint, i.e., $P_1 + P_2 = P$, is always guaranteed in order to maximize $T_1$ subject to $T_2 \geq T_0$
- **Lemma 2:** The equality in the effective throughput constraint is always guaranteed, i.e., $T_2 = T_0$, in order to maximize $T_1$ subject to $T_2 \geq T_0$
- Optimal design:
  **Step 1:** Determine $R_2$ for a feasible $P_2$
- Determine the value of $R_2$ that maximizes $T_1$ for given feasible $P_1$ and $P_2$

# Cont.

- **Proposition 2:** The solution of $R_2$ to $T(R_2) = T_0$ can be obtained by the fixed-point iteration

$$R_2 = F(R_2) = \frac{T_0}{1 - Q(f(\gamma_2, R_2))} \tag{38}$$

  **Step 2:** Determine $R_1$ for given $P_1$, $P_2$, and $R_2$.

- The value of $R_1$ that maximizes the effective throughput $T_1$ is determined

  **Step 3:** Determine a strict lower bound on $P_2$

  **Step 4:** Determine the optimal power allocation