Massive Access and Many-User Information Theory

Dongning Guo

Joint work with Lei Zhang, Jun Luo, Xu Chen, and Tsung-Yi Chen

Department of Electrical Engineering and Computer Science
Northwestern University

@ Conference on Applied Mathematics, The University of Hong Kong
August 25, 2016

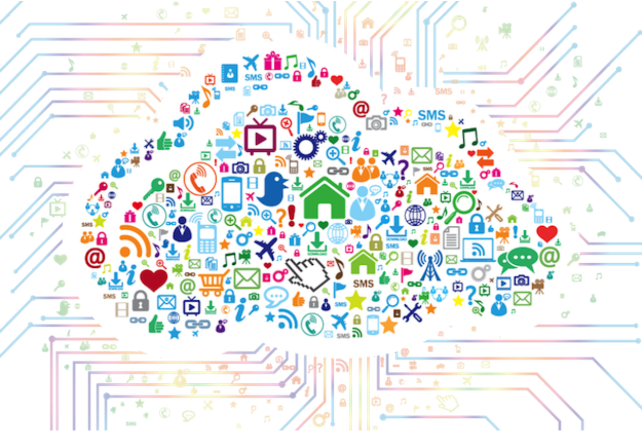
Motivation

Motivation

Motivation

Motivation

Internet of Things



Fundamental questions

1. ℓ devices in a cell; k of them are active.
2. ℓ and k are large numbers.
3. Massive grant free access in the uplink.
   Who transmitted?
   What are their messages?
4. Selective addressing in the downlink.
   Who should listen?
   What is the message for each receiver?

Fundamental questions

1. ℓ devices in a cell; k of them are active.
2. ℓ and k are large numbers.
3. Massive grant free access in the uplink.
   Who transmitted?
   What are their messages?
4. Selective addressing in the downlink.
   Who should listen?
   What is the message for each receiver?

Fundamental questions

1. ℓ devices in a cell; k of them are active.
2. ℓ and k are large numbers.
3. Massive grant free access in the uplink.
   Who transmitted?
   What are their messages?
4. Selective addressing in the downlink.
   Who should listen?
   What is the message for each receiver?

Fundamental questions

1. ℓ devices in a cell; k of them are active.
2. ℓ and k are large numbers.
3. Massive grant free access in the uplink.
   Who transmitted?
   What are their messages?
4. Selective addressing in the downlink.
   Who should listen?
   What is the message for each receiver?

Need a Many-User Information Theory

- Classical single-user Information Theory:
  1 user, coding blocklength $n \to \infty$.
- Multiuser Information Theory:
  k users (fixed, usually small), $n \to \infty$.
- Large-system analysis:
  $n \to \infty$ first, then $k \to \infty$.
- However, $k > n$ in many systems.
  E.g., large sensor networks, Internet of things.
- $n \to \infty$ for fixed k may be inaccurate and provide little insight.
- We propose a Many-User Information Theory:
  $k, n \to \infty$ simultaneously. For example:
  $k = \alpha n \to \infty$;
  $k = n^\alpha \to \infty$ (e.g., $k = 1,000,000$ devices, frame length $n = 1,000$).

Need a Many-User Information Theory

- Classical single-user Information Theory:
  1 user, coding blocklength $n \to \infty$.
- Multiuser Information Theory:
  k users (fixed, usually small), $n \to \infty$.
- Large-system analysis:
  $n \to \infty$ first, then $k \to \infty$.
- However, $k > n$ in many systems.
  E.g., large sensor networks, Internet of things.
- $n \to \infty$ for fixed k may be inaccurate and provide little insight.
- We propose a Many-User Information Theory:
  $k, n \to \infty$ simultaneously. For example:
  $k = \alpha n \to \infty$;
  $k = n^\alpha \to \infty$ (e.g., $k = 1,000,000$ devices, frame length $n = 1,000$).

Need a Many-User Information Theory

- Classical single-user Information Theory:
  1 user, coding blocklength $n \to \infty$.
- Multiuser Information Theory:
  k users (fixed, usually small), $n \to \infty$.
- Large-system analysis:
  $n \to \infty$ first, then $k \to \infty$.
- However, $k > n$ in many systems.
  E.g., large sensor networks, Internet of things.
- $n \to \infty$ for fixed k may be inaccurate and provide little insight.
- We propose a Many-User Information Theory:
  $k, n \to \infty$ simultaneously. For example:
  $k = \alpha n \to \infty$;
  $k = n^\alpha \to \infty$ (e.g., $k = 1,000,000$ devices, frame length $n = 1,000$).

Need a Many-User Information Theory

- Classical single-user Information Theory:
  1 user, coding blocklength $n \to \infty$.
- Multiuser Information Theory:
  k users (fixed, usually small), $n \to \infty$.
- Large-system analysis:
  $n \to \infty$ first, then $k \to \infty$.
- However, $k > n$ in many systems.
  E.g., large sensor networks, Internet of things.
- $n \to \infty$ for fixed k may be inaccurate and provide little insight.
- We propose a Many-User Information Theory:
  $k, n \to \infty$ simultaneously. For example:
  $k = \alpha n \to \infty$;
  $k = n^\alpha \to \infty$ (e.g., $k = 1,000,000$ devices, frame length $n = 1,000$).

Need a Many-User Information Theory

- Classical single-user Information Theory:
  1 user, coding blocklength $n \to \infty$.
- Multiuser Information Theory:
  k users (fixed, usually small), $n \to \infty$.
- Large-system analysis:
  $n \to \infty$ first, then $k \to \infty$.
- However, $k > n$ in many systems.
  E.g., large sensor networks, Internet of things.
- $n \to \infty$ for fixed k may be inaccurate and provide little insight.
- We propose a Many-User Information Theory:
  $k, n \to \infty$ simultaneously. For example:
  $k = \alpha n \to \infty$;
  $k = n^\alpha \to \infty$ (e.g., $k = 1,000,000$ devices, frame length $n = 1,000$).

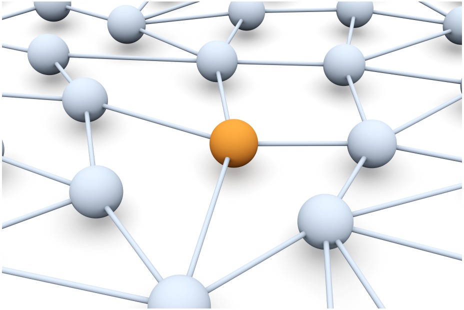
Need a Many-User Information Theory

- Classical single-user Information Theory:
  1 user, coding blocklength $n \to \infty$.
- Multiuser Information Theory:
  k users (fixed, usually small), $n \to \infty$.
- Large-system analysis:
  $n \to \infty$ first, then $k \to \infty$.
- However, $k > n$ in many systems.
  E.g., large sensor networks, Internet of things.
- $n \to \infty$ for fixed k may be inaccurate and provide little insight.
- We propose a Many-User Information Theory:
  $k, n \to \infty$ simultaneously. For example:
  $k = \alpha n \to \infty$;
  $k = n^\alpha \to \infty$ (e.g., $k = 1,000,000$ devices, frame length $n = 1,000$).

Outline

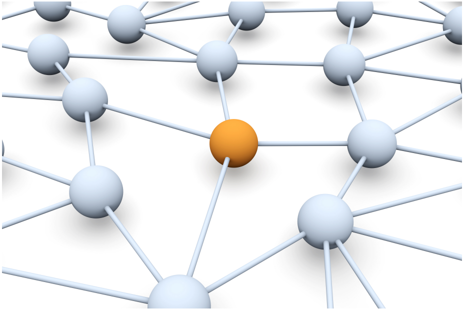
- (Almost practical) device identification
- Classical information theory
- Many-access channel
- Many-broadcast channel

Neighbor discovery/device identification



- To acquire the network interface addresses (NIAs) of all neighbors.
- Prior art: random access. Each node sends its NIA repeatedly with random delay.

Neighbor discovery/device identification



- To acquire the network interface addresses (NIAs) of all neighbors.
- Prior art: random access. Each node sends its NIA repeatedly with random delay.

Discovery is fundamentally sparse signal recovery

- b-bit address, $l = 2^b$ valid NIAs total.
- Node i sends signal $\boldsymbol{s}_i$.
- Multiaccess channel with path loss and fading:

$$\begin{aligned}\boldsymbol{Y} &= \sum_{i \in \text{neighborhood}} \boldsymbol{s}_i U_i + \boldsymbol{W} \\ &= \sum_{i=1}^{l} \boldsymbol{s}_i X_i + \boldsymbol{W} \\ &= \underline{\boldsymbol{S}}\boldsymbol{X} + \boldsymbol{W}\end{aligned}$$

- Given $\boldsymbol{Y}$, what is $\boldsymbol{X}$?
- $X_i \approxeq 0$ for all but a few neighbors.

Discovery is fundamentally sparse signal recovery

- b-bit address, $l = 2^b$ valid NIAs total.
- Node i sends signal $\boldsymbol{s}_i$.
- Multiaccess channel with path loss and fading:

$$\begin{aligned}\boldsymbol{Y} &= \sum_{i \in \text{neighborhood}} \boldsymbol{s}_i U_i + \boldsymbol{W} \\ &= \sum_{i=1}^{l} \boldsymbol{s}_i X_i + \boldsymbol{W} \\ &= \underline{\boldsymbol{S}}\boldsymbol{X} + \boldsymbol{W}\end{aligned}$$

- Given $\boldsymbol{Y}$, what is $\boldsymbol{X}$?
- $X_i \approxeq 0$ for all but a few neighbors.

Discovery is fundamentally sparse signal recovery

- b-bit address, $l = 2^b$ valid NIAs total.
- Node i sends signal $\boldsymbol{s}_i$.
- Multiaccess channel with path loss and fading:

$$\begin{aligned} \boldsymbol{Y} &= \sum_{i \in \text{neighborhood}} \boldsymbol{s}_i U_i + \boldsymbol{W} \\ &= \sum_{i=1}^{l} \boldsymbol{s}_i X_i + \boldsymbol{W} \\ &= \underline{\boldsymbol{S}} \boldsymbol{X} + \boldsymbol{W} \end{aligned}$$

- Given $\boldsymbol{Y}$, what is $\boldsymbol{X}$?
- $X_i \approxeq 0$ for all but a few neighbors.

Discovery is fundamentally sparse signal recovery

- b-bit address, $l = 2^b$ valid NIAs total.
- Node i sends signal $\boldsymbol{s}_i$.
- Multiaccess channel with path loss and fading:

$$\begin{aligned}\boldsymbol{Y} &= \sum_{i \in \text{neighborhood}} \boldsymbol{s}_i U_i + \boldsymbol{W} \\ &= \sum_{i=1}^{l} \boldsymbol{s}_i X_i + \boldsymbol{W} \\ &= \underline{\boldsymbol{S}} \boldsymbol{X} + \boldsymbol{W}\end{aligned}$$

- Given $\boldsymbol{Y}$, what is $\boldsymbol{X}$?
- $X_i \approxeq 0$ for all but a few neighbors.

Discovery is fundamentally sparse signal recovery

- b-bit address, $l = 2^b$ valid NIAs total.
- Node i sends signal $\boldsymbol{s}_i$.
- Multiaccess channel with path loss and fading:

$$
\begin{aligned}
\boldsymbol{Y} &= \sum_{i \in \text{neighborhood}} \boldsymbol{s}_i U_i + \boldsymbol{W} \\
&= \sum_{i=1}^{l} \boldsymbol{s}_i X_i + \boldsymbol{W} \\
&= \underline{\boldsymbol{S}} \boldsymbol{X} + \boldsymbol{W}
\end{aligned}
$$

- Given $\boldsymbol{Y}$, what is $\boldsymbol{X}$?
- $X_i \approxeq 0$ for all but a few neighbors.

Network-wide discovery in one frame interval

- Each node transmits a single frame of signature.
- Synchronized transmissions (can be relaxed).
- One key challenge is decoding complexity (need to scale to 2^{20}–2^{48} NIAs).
- Second-order Reed-Muller codes + chirp decoding algorithm [Calderbank, Gilbert & Strauss '06], [Howard, Calderbank & Searle '08].
- In an ad hoc network with half-duplex transceivers, use symbol erasures to achieve rapid on-off division duplex (RODD).

Network-wide discovery in one frame interval

- Each node transmits a single frame of signature.
- Synchronized transmissions (can be relaxed).
- One key challenge is decoding complexity (need to scale to 2^{20}–2^{48} NIAs).
- Second-order Reed-Muller codes + chirp decoding algorithm [Calderbank, Gilbert & Strauss '06], [Howard, Calderbank & Searle '08].
- In an ad hoc network with half-duplex transceivers, use symbol erasures to achieve rapid on-off division duplex (RODD).

Network-wide discovery in one frame interval

- Each node transmits a single frame of signature.
- Synchronized transmissions (can be relaxed).
- One key challenge is decoding complexity (need to scale to 2^{20}–2^{48} NIAs).
- Second-order Reed-Muller codes + chirp decoding algorithm [Calderbank, Gilbert & Strauss '06], [Howard, Calderbank & Searle '08].
- In an ad hoc network with half-duplex transceivers, use symbol erasures to achieve rapid on-off division duplex (RODD).

Network-wide discovery in one frame interval

- Each node transmits a single frame of signature.
- Synchronized transmissions (can be relaxed).
- One key challenge is decoding complexity (need to scale to 2^{20}–2^{48} NIAs).
- Second-order Reed-Muller codes + chirp decoding algorithm [Calderbank, Gilbert & Strauss '06], [Howard, Calderbank & Searle '08].
- In an ad hoc network with half-duplex transceivers, use symbol erasures to achieve rapid on-off division duplex (RODD).

Network-wide discovery in one frame interval

- Each node transmits a single frame of signature.
- Synchronized transmissions (can be relaxed).
- One key challenge is decoding complexity (need to scale to 2^{20}–2^{48} NIAs).
- Second-order Reed-Muller codes + chirp decoding algorithm [Calderbank, Gilbert & Strauss '06], [Howard, Calderbank & Searle '08].
- In an ad hoc network with half-duplex transceivers, use symbol erasures to achieve rapid on-off division duplex (RODD).

Second order Reed-Muller signature generation

- Signature length $n = 2^m$.
- $P_{m\times m}$ is a binary symmetric matrix, $x, t \in \mathbb{Z}_2^m$:

$$\varphi_{P,t}(x) = \left(\sqrt{-1}\right)^{x^\top P x + 2t^\top x}.$$

- Codebook size up to $2^{m(m+3)/2}$:
  - $m = 5$, $n = 2^5 = 32$, l up to 2^{20} codewords.
  - $m = 10$, $n = 2^{10} = 1{,}024$, l up to 2^{65};
  - $m = 12$, $n = 2^{12} = 4{,}096$, l up to 2^{90}.
- Introduce about 50% erasures in case of virtual full duplex.

1	1	j	-j	j	-j	-1	-1	1	1	-j	j	j	-j	1	1	1	-1	j	j	j	j	-1	1	-1	1	j	j	-j	-j	-1	1
1	1	j	-j	j	-j	1	1	1	-1	-j	-j	-j	-j	1	-1	j	-j	1	1	-1	-1	-j	j	j	j	-1	1	1	-1	-j	-j
1	1	j	j	j	-j	-1	1	j	j	1	1	-1	1	j	-j	j	-j	-1	1	1	1	j	j	1	-1	-j	j	-j	-j	-1	-1
1	-1	j	-j	j	j	1	1	j	j	1	1	1	-1	j	-j	1	1	-j	-j	-j	j	1	-1	j	-j	-1	1	-1	-1	j	j
1	-j	1	j	j	1	-j	1	1	j	-1	j	j	-1	j	1	j	-1	j	1	-1	-j	1	-j	j	1	-j	1	-1	j	-1	-j
1	-1	1	1	j	-j	-j	-j	1	1	-1	1	-j	-j	-j	j	j	j	-j	j	-1	-1	-1	1	-j	j	-j	-j	-1	1	1	1
1	j	-1	-j	j	-1	j	-1	j	1	j	1	-1	j	1	-j	1	-j	-1	j	-j	-1	-j	-1	j	-1	j	-1	1	j	-1	-j
1	1	-1	-1	j	j	j	j	j	-j	j	-j	1	-1	-1	1	1	-1	1	-1	-j	j	j	-j	-j	-j	j	j	1	1	1	1

Second order Reed-Muller signature generation

- Signature length $n = 2^m$.
- $P_{m\times m}$ is a binary symmetric matrix, $x, t \in \mathbb{Z}_2^m$:

$$\varphi_{P,t}(x) = \left(\sqrt{-1}\right)^{x^\mathsf{T} P x + 2t^\mathsf{T} x}.$$

- Codebook size up to $2^{m(m+3)/2}$:
  - $m = 5$, $n = 2^5 = 32$, l up to 2^{20} codewords.
  - $m = 10$, $n = 2^{10} = 1{,}024$, l up to 2^{65};
  - $m = 12$, $n = 2^{12} = 4{,}096$, l up to 2^{90}.
- Introduce about 50% erasures in case of virtual full duplex.

1	1	j	-j	j	-j	-1	-1	1	1	-j	j	j	-j	1	1	1	-1	j	j	j	j	-1	1	-1	1	j	j	-j	-j	-1	1
1	1	j	-j	j	-j	1	1	1	-1	-j	-j	-j	-j	1	-1	j	-j	1	1	-1	-1	-j	j	j	j	-1	1	1	-1	-j	-j
1	1	j	j	j	-j	-1	1	j	j	1	1	-1	1	j	-j	j	-j	-1	1	1	1	j	j	1	-1	-j	j	-j	-j	-1	-1
1	-1	j	-j	j	j	1	1	j	j	1	1	1	-1	j	-j	1	1	-j	-j	-j	j	1	-1	j	-j	-1	1	-1	-1	j	j
1	-j	1	j	j	1	-j	1	1	j	-1	j	j	-1	j	1	j	-1	j	1	-1	-j	1	-j	j	1	-j	1	-1	j	-1	-j
1	-1	1	1	j	-j	-j	-j	1	1	-1	1	-j	-j	-j	j	j	j	-j	j	-1	-1	-1	1	-j	j	-j	-j	-1	1	1	1
1	j	-1	-j	j	-1	j	-1	j	1	j	1	-1	j	1	-j	1	-j	-1	j	-j	-1	-j	-1	j	-1	j	-1	1	j	-1	-j
1	1	-1	-1	j	j	j	j	j	-j	j	-j	1	-1	-1	1	1	-1	1	-1	-j	j	j	-j	-j	-j	j	j	1	1	1	1

Second order Reed-Muller signature generation

- Signature length $n = 2^m$.
- $P_{m\times m}$ is a binary symmetric matrix, $x, t \in \mathbb{Z}_2^m$:

$$\varphi_{P,t}(x) = \left(\sqrt{-1}\right)^{x^\mathsf{T} P x + 2t^\mathsf{T} x}.$$

- Codebook size up to $2^{m(m+3)/2}$:
  - $m = 5$, $n = 2^5 = 32$, l up to 2^{20} codewords.
  - $m = 10$, $n = 2^{10} = 1{,}024$, l up to 2^{65};
  - $m = 12$, $n = 2^{12} = 4{,}096$, l up to 2^{90}.
- Introduce about 50% erasures in case of virtual full duplex.

1	1	j	$-j$	j	$-j$	-1	-1	1	1	$-j$	j	j	$-j$	1	1	1	-1	j	j	j	j	-1	1	-1	1	j	j	$-j$	$-j$	-1	1
1	1	j	$-j$	j	$-j$	1	1	1	-1	$-j$	$-j$	$-j$	$-j$	1	-1	j	$-j$	1	1	-1	-1	$-j$	j	j	j	-1	1	1	-1	$-j$	$-j$
1	1	j	j	j	$-j$	-1	1	j	j	1	1	-1	1	j	$-j$	j	$-j$	-1	1	1	1	j	j	1	-1	$-j$	j	$-j$	$-j$	-1	-1
1	-1	j	$-j$	j	j	1	1	j	j	1	1	1	-1	j	$-j$	1	1	$-j$	$-j$	$-j$	j	1	-1	j	$-j$	-1	1	-1	-1	j	j
1	$-j$	1	j	j	1	$-j$	1	1	j	-1	j	j	-1	j	1	j	-1	j	1	-1	$-j$	1	$-j$	j	1	$-j$	1	-1	j	-1	$-j$
1	-1	1	1	j	$-j$	$-j$	$-j$	1	1	-1	1	$-j$	$-j$	$-j$	j	j	j	$-j$	j	-1	-1	-1	1	$-j$	j	$-j$	$-j$	-1	1	1	1
1	j	-1	$-j$	j	-1	j	-1	j	1	j	1	-1	j	1	$-j$	1	$-j$	-1	j	$-j$	-1	$-j$	-1	j	-1	j	-1	1	j	-1	$-j$
1	1	-1	-1	j	j	j	j	j	$-j$	j	$-j$	1	-1	-1	1	1	-1	1	-1	$-j$	j	j	$-j$	$-j$	$-j$	j	j	1	1	1	1

Second order Reed-Muller signature generation

- Signature length $n = 2^m$.
- $P_{m\times m}$ is a binary symmetric matrix, $x, t \in \mathbb{Z}_2^m$:

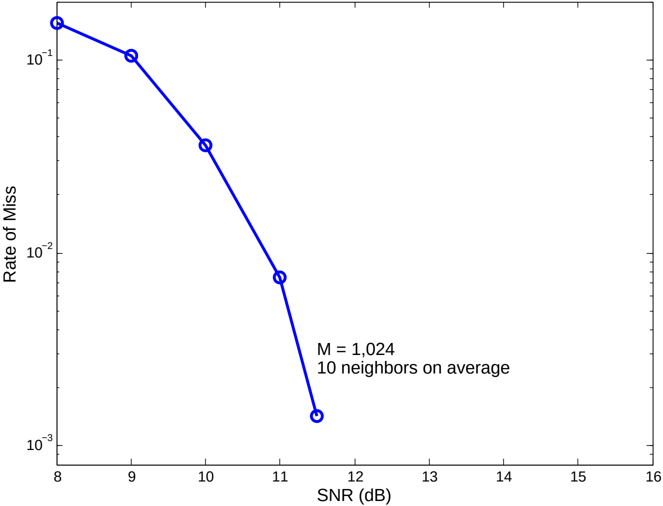
$$\varphi_{P,t}(x) = \left(\sqrt{-1}\right)^{x^\mathsf{T} P x + 2t^\mathsf{T} x}.$$

- Codebook size up to $2^{m(m+3)/2}$:
  - $m = 5$, $n = 2^5 = 32$, l up to 2^{20} codewords.
  - $m = 10$, $n = 2^{10} = 1{,}024$, l up to 2^{65};
  - $m = 12$, $n = 2^{12} = 4{,}096$, l up to 2^{90}.
- Introduce about 50% erasures in case of virtual full duplex.

0	0	j	0	j	-j	0	0	0	0	-j	j	0	-j	1	1	0	0	j	0	j	j	0	0	0	0	j	j	0	-j	-1	1
0	0	0	0	j	-j	0	1	0	-1	0	0	-j	0	0	0	0	0	0	0	-1	-1	0	j	0	j	0	0	1	0	0	0
1	1	0	0	j	0	-1	0	0	0	0	1	-1	1	j	-j	j	-j	0	0	1	0	j	0	0	0	0	j	-j	-j	-1	-1
0	0	0	-j	0	0	0	1	j	j	0	1	1	0	0	0	0	0	0	-j	0	0	0	-1	j	-j	0	1	-1	0	0	0
0	0	0	0	0	1	0	0	1	j	0	0	j	0	0	0	0	0	0	0	0	-j	0	0	j	1	0	0	-1	0	0	0
1	-1	1	1	j	-j	-j	-j	1	1	-1	0	0	0	0	0	j	j	-j	j	-1	-1	-1	1	-j	j	-j	0	0	0	0	0
1	0	0	0	0	-1	j	0	0	0	0	0	-1	j	1	-j	1	0	0	0	0	-1	-j	0	0	0	0	0	1	j	-1	-j
0	1	0	-1	j	0	j	0	j	-j	0	-j	0	-1	-1	1	0	-1	0	-1	-j	0	j	0	-j	-j	0	j	0	1	1	1

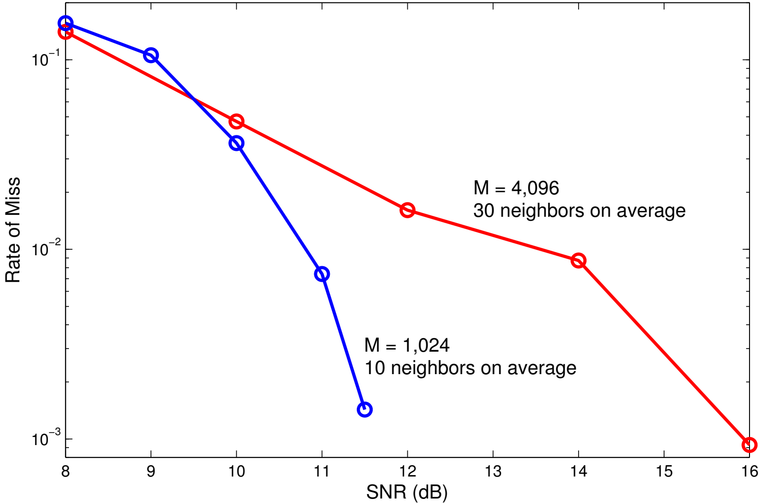
Error rate vs. SNR

2^{20} nodes, path loss exponent $= 3$, Rayleigh fading



Error rate vs. SNR

2^{20} nodes, path loss exponent $= 3$, Rayleigh fading



Comparison with random access

- $l = 2^{20}$ nodes, on average 10 neighbors, SNR $= 11.5$ dB

  Target $P_e = 0.002$

	Random access	RODD
# of frames	194	1
# of symbols	$\geq$ 194×20=3,880	1,024

- In addition, significant reduction of per-frame overhead.
- More results in:
  L. Zhang and D. Guo, "Virtual full duplex wireless broadcasting via compressed sensing," IEEE/ACM Trans. Networking, 2014.
  L. Zhang, J. Luo, and D. Guo, "Neighbor discovery for wireless networks via compressed sensing," Performance Evaluation, 2013.
  X. Chen and D. Guo, "Robust sublinear complexity Walsh-Hadamard transform with arbitrary sparse support," ISIT, 2015.

Comparison with random access

- $l = 2^{20}$ nodes, on average 10 neighbors, SNR $= 11.5$ dB
  Target $P_e = 0.002$

	Random access	RODD
# of frames	194	1
# of symbols	$\geq 194{\times}20{=}3{,}880$	1,024

- In addition, significant reduction of per-frame overhead.
- More results in:
  L. Zhang and D. Guo, “Virtual full duplex wireless broadcasting via compressed sensing,” IEEE/ACM Trans. Networking, 2014.
  L. Zhang, J. Luo, and D. Guo, “Neighbor discovery for wireless networks via compressed sensing,” Performance Evaluation, 2013.
  X. Chen and D. Guo, “Robust sublinear complexity Walsh-Hadamard transform with arbitrary sparse support,” ISIT, 2015.

Comparison with random access

- $l = 2^{20}$ nodes, on average 10 neighbors, SNR $= 11.5$ dB
  Target $P_e = 0.002$

	Random access	RODD
# of frames	194	1
# of symbols	$\geq$ 194×20=3,880	1,024

- In addition, significant reduction of per-frame overhead.
- More results in:
  L. Zhang and D. Guo, "Virtual full duplex wireless broadcasting via compressed sensing," IEEE/ACM Trans. Networking, 2014.
  L. Zhang, J. Luo, and D. Guo, "Neighbor discovery for wireless networks via compressed sensing," Performance Evaluation, 2013.
  X. Chen and D. Guo, "Robust sublinear complexity Walsh-Hadamard transform with arbitrary sparse support," ISIT, 2015.

Comparison with random access

- $l = 2^{20}$ nodes, on average 10 neighbors, SNR $= 11.5$ dB
  Target $P_e = 0.002$

	Random access	RODD
# of frames	194	1
# of symbols	$\geq$ 194×20=3,880	1,024

- In addition, significant reduction of per-frame overhead.
- More results in:
  L. Zhang and D. Guo, "Virtual full duplex wireless broadcasting via compressed sensing," IEEE/ACM Trans. Networking, 2014.
  L. Zhang, J. Luo, and D. Guo, "Neighbor discovery for wireless networks via compressed sensing," Performance Evaluation, 2013.
  X. Chen and D. Guo, "Robust sublinear complexity Walsh-Hadamard transform with arbitrary sparse support," ISIT, 2015.

Comparison with random access

- $l = 2^{20}$ nodes, on average 10 neighbors, SNR $= 11.5$ dB
  Target $P_e = 0.002$

	Random access	RODD
# of frames	194	1
# of symbols	$\geq$ 194×20=3,880	1,024

- In addition, significant reduction of per-frame overhead.
- More results in:
  L. Zhang and D. Guo, "Virtual full duplex wireless broadcasting via compressed sensing," IEEE/ACM Trans. Networking, 2014.
  L. Zhang, J. Luo, and D. Guo, "Neighbor discovery for wireless networks via compressed sensing," Performance Evaluation, 2013.
  X. Chen and D. Guo, "Robust sublinear complexity Walsh-Hadamard transform with arbitrary sparse support," ISIT, 2015.

What are the fundamental limits?

Outline

- (Almost practical) neighbor discovery
- Classical information theory: a digression
- Many-access channel
- Many-broadcast channel

Classical (single-user) information theory

- [Shannon–MacMillan–Breiman '48, '60] for discrete stationary ergodic sequence,

$$-\frac{1}{n}\log p_{X_1,\dots,X_n}(X_1,\dots,X_n) \xrightarrow{\text{a.s.}} \mathcal{H}$$

- Typical set

$$T_\epsilon^{(n)} = \left\{(x_1,\cdots,x_n) : \left| -\frac{1}{n}\log p_{X_1,\cdots,X_n}(x_1,\dots,x_n) - \mathcal{H}\right| \le \epsilon\right\}$$

- Asymptotic equipartition property:
  - Almost all sequences that occur are typical, $\lim_{n\to\infty} P\left\{T_\epsilon^{(n)}\right\} = 1$;
  - There are about $2^{n\mathcal{H}}$ of them, $|T_\epsilon^{(n)}| \approx 2^{n\mathcal{H}}$.
- Hence Shannon's lossless source coding theorem.
- A joint AEP is the workhorse for Shannon's channel coding theorem and rate distortion theorem.

Classical (single-user) information theory

- [Shannon–MacMillan–Breiman '48, '60] for discrete stationary ergodic sequence,

$$-\frac{1}{n}\log p_{X_1,\dots,X_n}(X_1,\dots,X_n) \overset{\text{a.s.}}{\to} \mathcal{H}$$

- Typical set

$$T_\epsilon^{(n)} = \left\{(x_1,\cdots,x_n) : \left| -\frac{1}{n}\log p_{X_1,\cdots,X_n}(x_1,\dots,x_n) - \mathcal{H}\right| \le \epsilon\right\}$$

- Asymptotic equipartition property:
  - Almost all sequences that occur are typical, $\lim_{n\to\infty} P\left\{T_\epsilon^{(n)}\right\} = 1$;
  - There are about $2^{n\mathcal{H}}$ of them, $|T_\epsilon^{(n)}| \approx 2^{n\mathcal{H}}$.
- Hence Shannon's lossless source coding theorem.
- A joint AEP is the workhorse for Shannon's channel coding theorem and rate distortion theorem.

Classical (single-user) information theory

- [Shannon–MacMillan–Breiman '48, '60] for discrete stationary ergodic sequence,

$$-\frac{1}{n}\log p_{X_1,\ldots,X_n}(X_1,\ldots,X_n) \overset{\text{a.s.}}{\to} \mathcal{H}$$

- Typical set

$$T_\epsilon^{(n)} = \left\{(x_1,\cdots,x_n) : \left| -\frac{1}{n}\log p_{X_1,\cdots,X_n}(x_1,\ldots,x_n) - \mathcal{H}\right| \le \epsilon\right\}$$

- Asymptotic equipartition property:
  - Almost all sequences that occur are typical, $\lim_{n\to\infty} P\left\{T_\epsilon^{(n)}\right\} = 1$;
  - There are about $2^{n\mathcal{H}}$ of them, $|T_\epsilon^{(n)}| \approx 2^{n\mathcal{H}}$.
- Hence Shannon's lossless source coding theorem.
- A joint AEP is the workhorse for Shannon's channel coding theorem and rate distortion theorem.

Classical (single-user) information theory

- [Shannon–MacMillan–Breiman '48, '60] for discrete stationary ergodic sequence,

$$-\frac{1}{n}\log p_{X_1,\dots,X_n}(X_1,\dots,X_n) \stackrel{\text{a.s.}}{\to} \mathcal{H}$$

- Typical set

$$T_\epsilon^{(n)} = \left\{(x_1,\cdots,x_n) : \left| -\frac{1}{n}\log p_{X_1,\cdots,X_n}(x_1,\dots,x_n) - \mathcal{H}\right| \le \epsilon\right\}$$

- Asymptotic equipartition property:
  - Almost all sequences that occur are typical, $\lim_{n\to\infty} P\left\{T_\epsilon^{(n)}\right\} = 1$;
  - There are about $2^{n\mathcal{H}}$ of them, $|T_\epsilon^{(n)}| \approx 2^{n\mathcal{H}}$.
- Hence Shannon's lossless source coding theorem.
- A joint AEP is the workhorse for Shannon's channel coding theorem and rate distortion theorem.

Classical (single-user) information theory

- [Shannon–MacMillan–Breiman '48, '60] for discrete stationary ergodic sequence,

$$-\frac{1}{n}\log p_{X_1,\dots,X_n}(X_1,\dots,X_n) \overset{\text{a.s.}}{\to} \mathcal{H}$$

- Typical set

$$T_\epsilon^{(n)} = \left\{(x_1,\cdots,x_n) : \left| -\frac{1}{n}\log p_{X_1,\cdots,X_n}(x_1,\dots,x_n) - \mathcal{H}\right| \le \epsilon\right\}$$

- Asymptotic equipartition property:
  - Almost all sequences that occur are typical, $\lim_{n\to\infty} P\left\{T_\epsilon^{(n)}\right\} = 1$;
  - There are about $2^{n\mathcal{H}}$ of them, $|T_\epsilon^{(n)}| \approx 2^{n\mathcal{H}}$.
- Hence Shannon's lossless source coding theorem.
- A joint AEP is the workhorse for Shannon's channel coding theorem and rate distortion theorem.

Classical (single-user) information theory

- [Shannon–MacMillan–Breiman '48, '60] for discrete stationary ergodic sequence,

$$-\frac{1}{n}\log p_{X_1,\dots,X_n}(X_1,\dots,X_n) \overset{\text{a.s.}}{\to} \mathcal{H}$$

- Typical set

$$T_\epsilon^{(n)} = \left\{(x_1,\cdots,x_n) : \left| -\frac{1}{n}\log p_{X_1,\cdots,X_n}(x_1,\dots,x_n) - \mathcal{H}\right| \le \epsilon\right\}$$

- Asymptotic equipartition property:
  - Almost all sequences that occur are typical, $\lim_{n\to\infty} P\left\{T_\epsilon^{(n)}\right\} = 1$;
  - There are about $2^{n\mathcal{H}}$ of them, $|T_\epsilon^{(n)}| \approx 2^{n\mathcal{H}}$.
- Hence Shannon's lossless source coding theorem.
- A joint AEP is the workhorse for Shannon's channel coding theorem and rate distortion theorem.

Classical (single-user) information theory

- [Shannon–MacMillan–Breiman '48, '60] for discrete stationary ergodic sequence,

$$-\frac{1}{n}\log p_{X_1,\dots,X_n}(X_1,\dots,X_n) \stackrel{\text{a.s.}}{\to} \mathcal{H}$$

- Typical set

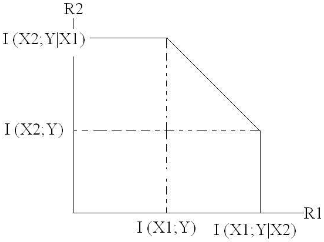
$$T_\epsilon^{(n)} = \left\{(x_1,\cdots,x_n) : \left| -\frac{1}{n}\log p_{X_1,\cdots,X_n}(x_1,\dots,x_n) - \mathcal{H}\right| \le \epsilon\right\}$$

- Asymptotic equipartition property:
  - Almost all sequences that occur are typical, $\lim_{n\to\infty} P\left\{T_\epsilon^{(n)}\right\} = 1$;
  - There are about $2^{n\mathcal{H}}$ of them, $|T_\epsilon^{(n)}| \approx 2^{n\mathcal{H}}$.
- Hence Shannon's lossless source coding theorem.
- A joint AEP is the workhorse for Shannon's channel coding theorem and rate distortion theorem.

Classical multiaccess channel

$$\begin{array}{l} w_1 \longmapsto (X_{11}, \ldots, X_{1n}) \searrow \\ \qquad\qquad\qquad\qquad\qquad\quad \boxed{P_{Y|X_1,X_2}} \longrightarrow (Y_1, \ldots, Y_n) \longmapsto (\hat{w}_1, \hat{w}_2) \\ w_2 \longmapsto (X_{21}, \ldots, X_{2n}) \nearrow \end{array}$$

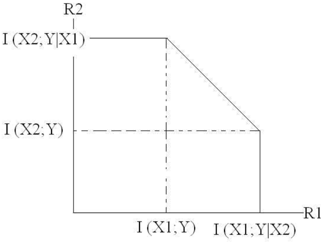
The capacity region is due to Ahlswede (1971) and Liao (1972). It can be achieved by random coding, joint typicality decoding, and time sharing.



Classical multiaccess channel

$$
\begin{array}{l}
w_1 \longmapsto (X_{11}, \dots, X_{1n}) \searrow \\
\qquad\qquad\qquad\qquad\qquad\quad \boxed{P_{Y|X_1,X_2}} \longrightarrow (Y_1, \dots, Y_n) \longmapsto (\hat{w}_1, \hat{w}_2) \\
w_2 \longmapsto (X_{21}, \dots, X_{2n}) \nearrow
\end{array}
$$

The capacity region is due to Ahlswede (1971) and Liao (1972). It can be achieved by random coding, joint typicality decoding, and time sharing.



Jointly typical set

$$
\begin{aligned}
T_\epsilon^{(n)} = \Big\{ (\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}) : & \left| \frac{1}{n} \log \frac{1}{p_{\boldsymbol{X}_1}(\boldsymbol{x}_1)} - H(X_1) \right| < \epsilon \\
& \left| \frac{1}{n} \log \frac{1}{p_{\boldsymbol{X}_2}(\boldsymbol{x}_2)} - H(X_2) \right| < \epsilon \\
& \left| \frac{1}{n} \log \frac{1}{p_{\boldsymbol{Y}}(\boldsymbol{y})} - H(Y) \right| < \epsilon \\
& \left| \frac{1}{n} \log \frac{1}{p_{\boldsymbol{X}_1\boldsymbol{Y}}(\boldsymbol{x}_1, \boldsymbol{y})} - H(X_1, Y) \right| < \epsilon \\
& \left| \frac{1}{n} \log \frac{1}{p_{\boldsymbol{X}_2\boldsymbol{Y}}(\boldsymbol{x}_2, \boldsymbol{y})} - H(X_2, Y) \right| < \epsilon \\
& \left| \frac{1}{n} \log \frac{1}{p_{\boldsymbol{X}_1\boldsymbol{X}_2}(\boldsymbol{x}_1, \boldsymbol{x}_2)} - H(X_1, X_2) \right| < \epsilon \\
& \left| \frac{1}{n} \log \frac{1}{p_{\boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{Y}}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y})} - H(X_1, X_2, Y) \right| < \epsilon \Big\}
\end{aligned}
$$

The "empirical entropy" converges to the entropy for all subsets of (X_1, X_2, Y).

Proof of multiaccess channel capacity

- Two users transmit $\boldsymbol{X}_1(w_1)$ and $\boldsymbol{X}_2(w_2)$ from random codebooks.
- Receiver puts out the first $(\hat{w}_1, \hat{w}_2)$ satisfying $(\boldsymbol{X}_1(\hat{w}_1), \boldsymbol{X}_2(\hat{w}_2), \boldsymbol{Y}) \in T_\epsilon^{(n)}$.
- Let $E_{w_1 w_2} = \{(\boldsymbol{X}_1(w_1), \boldsymbol{X}_2(w_2), \boldsymbol{Y}) \in T_\epsilon^{(n)}\}$. Then

$$
\begin{aligned}
P_e \le\ & \underbrace{P(E_{11}^c)}_{\to 0 \text{ by AEP}} \\
& + \underbrace{\sum_{w_1=1, w_2\neq 1} P(E_{1w_2})}_{\Rightarrow \text{bound on } R_2} + \underbrace{\sum_{w_1\neq 1, w_2=1} P(E_{w_1 1})}_{\Rightarrow \text{bound on } R_1} + \underbrace{\sum_{w_1\neq 1, w_2\neq 1} P(E_{w_1 w_2})}_{\Rightarrow \text{bound on } R_1+R_2} \\
\to\ & 0
\end{aligned}
$$

Proof of multiaccess channel capacity

- Two users transmit $\boldsymbol{X}_1(w_1)$ and $\boldsymbol{X}_2(w_2)$ from random codebooks.
- Receiver puts out the first $(\hat{w}_1, \hat{w}_2)$ satisfying $(\boldsymbol{X}_1(\hat{w}_1), \boldsymbol{X}_2(\hat{w}_2), \boldsymbol{Y}) \in T_\epsilon^{(n)}$.
- Let $E_{w_1 w_2} = \{(\boldsymbol{X}_1(w_1), \boldsymbol{X}_2(w_2), \boldsymbol{Y}) \in T_\epsilon^{(n)}\}$. Then

$$P_e \leq \underbrace{P(E_{11}^c)}_{\to 0 \text{ by AEP}} + \underbrace{\sum_{w_1=1, w_2\neq 1} P(E_{1w_2})}_{\Rightarrow \text{bound on } R_2} + \underbrace{\sum_{w_1\neq 1, w_2=1} P(E_{w_1 1})}_{\Rightarrow \text{bound on } R_1} + \underbrace{\sum_{w_1\neq 1, w_2\neq 1} P(E_{w_1 w_2})}_{\Rightarrow \text{bound on } R_1+R_2}$$

$$\to 0$$

Proof of multiaccess channel capacity

- Two users transmit $\boldsymbol{X}_1(w_1)$ and $\boldsymbol{X}_2(w_2)$ from random codebooks.
- Receiver puts out the first $(\hat{w}_1, \hat{w}_2)$ satisfying $(\boldsymbol{X}_1(\hat{w}_1), \boldsymbol{X}_2(\hat{w}_2), \boldsymbol{Y}) \in T_\epsilon^{(n)}$.
- Let $E_{w_1 w_2} = \{(\boldsymbol{X}_1(w_1), \boldsymbol{X}_2(w_2), \boldsymbol{Y}) \in T_\epsilon^{(n)}\}$. Then

$$P_e \leq \underbrace{P(E_{11}^c)}_{\to 0 \text{ by AEP}} + \underbrace{\sum_{w_1=1, w_2\neq 1} P(E_{1w_2})}_{\Rightarrow \text{bound on } R_2} + \underbrace{\sum_{w_1\neq 1, w_2=1} P(E_{w_1 1})}_{\Rightarrow \text{bound on } R_1} + \underbrace{\sum_{w_1\neq 1, w_2\neq 1} P(E_{w_1 w_2})}_{\Rightarrow \text{bound on } R_1+R_2}$$

$$\to 0$$

The large-system limit is ill-suited for many-user

- Example: The sum rate of the k-user Gaussian multiaccess channel:

$$C_{\text{sum}} = \frac{1}{2}\log(1+k\gamma) \to \infty$$
$$\frac{1}{k}C_{\text{sum}} = \frac{1}{2k}\log(1+k\gamma) \to 0$$

- *"When the total number of senders is very large, so that there is a lot of interference, we can still send a total amount of information that is arbitrary large even though the rate per individual sender goes to 0."* —Cover & Thomas, Elements of Information Theory.

- Rate or capacity in bits per channel use is ill-suited for many-user systems.

The large-system limit is ill-suited for many-user

- Example: The sum rate of the k-user Gaussian multiaccess channel:

$$C_{\text{sum}} = \frac{1}{2}\log(1+k\gamma) \to \infty$$
$$\frac{1}{k}C_{\text{sum}} = \frac{1}{2k}\log(1+k\gamma) \to 0$$

- *"When the total number of senders is very large, so that there is a lot of interference, we can still send a total amount of information that is arbitrary large even though the rate per individual sender goes to 0."* —Cover & Thomas, Elements of Information Theory.

- Rate or capacity in bits per channel use is ill-suited for many-user systems.

The large-system limit is ill-suited for many-user

- Example: The sum rate of the k-user Gaussian multiaccess channel:

$$C_{\text{sum}} = \frac{1}{2}\log(1+k\gamma) \to \infty$$
$$\frac{1}{k}C_{\text{sum}} = \frac{1}{2k}\log(1+k\gamma) \to 0$$

- *"When the total number of senders is very large, so that there is a lot of interference, we can still send a total amount of information that is arbitrary large even though the rate per individual sender goes to 0."* —Cover & Thomas, Elements of Information Theory.

- Rate or capacity in bits per channel use is ill-suited for many-user systems.

Outline

- (Almost practical) neighbor discovery
- Classical information theory
- Many-access channel
- Many-broadcast channel

Gaussian many-access channel (MnAC)

$$\boldsymbol{Y} = \sum_{j=1}^{\ell} \boldsymbol{S}_j(w_j) + \boldsymbol{Z}$$

- Many (ℓ) transmitters, each active w.p. $\alpha \in (0, 1]$ in a block.
- Average number of active users:

$$k = \alpha\ell.$$

- Message of user j: w_j, corresponding codeword $\boldsymbol{S}_j(w_j)$.
- User j is silent if $w_j = 0$, $\boldsymbol{S}_j(0) = \boldsymbol{0}$.

Gaussian many-access channel (MnAC)

$$\boldsymbol{Y} = \sum_{j=1}^{\ell} \boldsymbol{S}_j(w_j) + \boldsymbol{Z}$$

- Many (ℓ) transmitters, each active w.p. $\alpha \in (0, 1]$ in a block.
- Average number of active users:

$$k = \alpha\ell.$$

- Message of user j: w_j, corresponding codeword $\boldsymbol{S}_j(w_j)$.
- User j is silent if $w_j = 0$, $\boldsymbol{S}_j(0) = \boldsymbol{0}$.

Gaussian many-access channel (MnAC)

$$\boldsymbol{Y} = \sum_{j=1}^{\ell} \boldsymbol{S}_j(w_j) + \boldsymbol{Z}$$

- Many (ℓ) transmitters, each active w.p. $\alpha \in (0, 1]$ in a block.
- Average number of active users:

$$k = \alpha\ell.$$

- Message of user j: w_j, corresponding codeword $\boldsymbol{S}_j(w_j)$.
- User j is silent if $w_j = 0$, $\boldsymbol{S}_j(0) = \boldsymbol{0}$.

Gaussian many-access channel (MnAC)

$$\boldsymbol{Y} = \sum_{j=1}^{\ell} \boldsymbol{S}_j(w_j) + \boldsymbol{Z}$$

- Many (ℓ) transmitters, each active w.p. $\alpha \in (0, 1]$ in a block.
- Average number of active users:

$$k = \alpha\ell.$$

- Message of user j: w_j, corresponding codeword $\boldsymbol{S}_j(w_j)$.
- User j is silent if $w_j = 0$, $\boldsymbol{S}_j(0) = \mathbf{0}$.

New challenges

- Time sharing is not good in general. If $n = 1,000$, $k = 2,000$, an average user has half a channel use!
- Classical joint typicality does not apply as $n, \ell \to \infty$:

$$\begin{bmatrix} X^n_{1,1} & X^n_{1,2} & \dots & X^n_{1,n} \\ X^n_{2,1} & X^n_{2,2} & \dots & X^n_{2,n} \\ \vdots & \vdots & & \vdots \\ X^n_{\ell,1} & X^n_{\ell,2} & \dots & X^n_{\ell,n} \end{bmatrix}$$

- The union bound fails as the number of error events 2^k grows exponentially with n.

New challenges

- Time sharing is not good in general. If $n = 1,000$, $k = 2,000$, an average user has half a channel use!
- Classical joint typicality does not apply as $n, \ell \to \infty$:

$$\begin{bmatrix} X^n_{1,1} & X^n_{1,2} & \dots & X^n_{1,n} \\ X^n_{2,1} & X^n_{2,2} & \dots & X^n_{2,n} \\ \vdots & \vdots & & \vdots \\ X^n_{\ell,1} & X^n_{\ell,2} & \dots & X^n_{\ell,n} \end{bmatrix}$$

- The union bound fails as the number of error events 2^k grows exponentially with n.

New challenges

- Time sharing is not good in general. If $n = 1,000$, $k = 2,000$, an average user has half a channel use!
- Classical joint typicality does not apply as $n, \ell \to \infty$:

$$\begin{bmatrix} X_{1,1}^n & X_{1,2}^n & \dots & X_{1,n}^n \\ X_{2,1}^n & X_{2,2}^n & \dots & X_{2,n}^n \\ \vdots & \vdots & & \vdots \\ X_{\ell,1}^n & X_{\ell,2}^n & \dots & X_{\ell,n}^n \end{bmatrix}$$

- The union bound fails as the number of error events 2^k grows exponentially with n.

Identification code for the memoryless Gaussian MnAC

1. Let ℓ denote the total number of users.
2. If user j is inactive ($W_j = 0$), it transmits $\mathbf{0}$;
3. If user j is active ($W_j = 1$), it transmits $\boldsymbol{s}_j$,

$$\frac{1}{n_\ell}\sum_{i=1}^{n_\ell} s_{ji}^2 \le \gamma.$$

4. On average, k_ℓ users are active with i.i.d. activities, $P\{W_j = 1\} = \frac{k_\ell}{\ell}$.
5. Average identification error probability:

$$p_\ell = P\left\{\mathcal{D}(\boldsymbol{Y}) \neq (W_1, \dots, W_\ell)\right\}.$$

Identification code for the memoryless Gaussian MnAC

1. Let ℓ denote the total number of users.
2. If user j is inactive ($W_j = 0$), it transmits $\mathbf{0}$;
3. If user j is active ($W_j = 1$), it transmits s_j,

$$\frac{1}{n_\ell} \sum_{i=1}^{n_\ell} s_{ji}^2 \leq \gamma.$$

4. On average, k_ℓ users are active with i.i.d. activities, $P\{W_j = 1\} = \frac{k_\ell}{\ell}$.
5. Average identification error probability:

$$p_\ell = P\left\{\mathcal{D}(\boldsymbol{Y}) \neq (W_1, \ldots, W_\ell)\right\}.$$

Identification code for the memoryless Gaussian MnAC

1. Let ℓ denote the total number of users.
2. If user j is inactive ($W_j = 0$), it transmits $\mathbf{0}$;
3. If user j is active ($W_j = 1$), it transmits $\boldsymbol{s}_j$,

$$\frac{1}{n_\ell} \sum_{i=1}^{n_\ell} s_{ji}^2 \leq \gamma.$$

4. On average, k_ℓ users are active with i.i.d. activities, $P\{W_j = 1\} = \frac{k_\ell}{\ell}$.
5. Average identification error probability:

$$p_\ell = P\left\{\mathcal{D}(\boldsymbol{Y}) \neq (W_1, \ldots, W_\ell)\right\}.$$

Identification code for the memoryless Gaussian MnAC

1. Let ℓ denote the total number of users.
2. If user j is inactive $(W_j = 0)$, it transmits $\mathbf{0}$;
3. If user j is active $(W_j = 1)$, it transmits $\boldsymbol{s}_j$,

$$\frac{1}{n_\ell} \sum_{i=1}^{n_\ell} s_{ji}^2 \leq \gamma.$$

4. On average, k_ℓ users are active with i.i.d. activities, $P\{W_j = 1\} = \frac{k_\ell}{\ell}$.
5. Average identification error probability:

$$p_\ell = P\left\{\mathcal{D}(\boldsymbol{Y}) \neq (W_1, \dots, W_\ell)\right\}.$$

Identification code for the memoryless Gaussian MnAC

1. Let ℓ denote the total number of users.
2. If user j is inactive $(W_j = 0)$, it transmits $\mathbf{0}$;
3. If user j is active $(W_j = 1)$, it transmits $\boldsymbol{s}_j$,

$$\frac{1}{n_\ell}\sum_{i=1}^{n_\ell} s_{ji}^2 \leq \gamma.$$

4. On average, k_ℓ users are active with i.i.d. activities, $P\{W_j = 1\} = \frac{k_\ell}{\ell}$.
5. Average identification error probability:

$$p_\ell = P\left\{\mathcal{D}(\boldsymbol{Y}) \neq (W_1, \ldots, W_\ell)\right\}.$$

Main result 1: identification cost

Theorem

Let

$$n_\ell = \frac{\ell \, H_2(k_\ell/\ell)}{\frac{1}{2}\log(1 + k_\ell \gamma)}.$$

For every $\epsilon \in (0,1)$, as $\ell \to \infty$, arbitrarily reliable identification ($p_\ell \to 0$) is achievable with $(1+\epsilon)n_\ell$ channel uses; whereas $p_\ell \to 0$ is not achievable with $(1-\epsilon)n_\ell$ channel uses. Here

$$H_2(\alpha) = \alpha \log \frac{1}{\alpha} + (1-\alpha) \log \frac{1}{1-\alpha}.$$

- Intuition: The activity uncertainty of ℓ users is $\ell H_2(k_\ell/\ell)$.
- Achievable code: random Gaussian sequences as signatures.

Main result 1: identification cost

Theorem

Let

$$n_\ell = \frac{\ell\, H_2(k_\ell/\ell)}{\frac{1}{2}\log(1+k_\ell\gamma)}.$$

For every $\epsilon \in (0,1)$, as $\ell \to \infty$, arbitrarily reliable identification ($p_\ell \to 0$) is achievable with $(1+\epsilon)n_\ell$ channel uses; whereas $p_\ell \to 0$ is not achievable with $(1-\epsilon)n_\ell$ channel uses. Here

$$H_2(\alpha) = \alpha \log \frac{1}{\alpha} + (1-\alpha)\log\frac{1}{1-\alpha}.$$

- Intuition: The activity uncertainty of ℓ users is $\ell H_2(k_\ell/\ell)$.
- Achievable code: random Gaussian sequences as signatures.

Main result 1: identification cost

Theorem

Let

$$n_\ell = \frac{\ell\, H_2(k_\ell/\ell)}{\frac{1}{2}\log(1+k_\ell\gamma)}.$$

For every $\epsilon \in (0,1)$*, as* $\ell \to \infty$*, arbitrarily reliable identification (*$p_\ell \to 0$*) is achievable with* $(1+\epsilon)n_\ell$ *channel uses; whereas* $p_\ell \to 0$ *is not achievable with* $(1-\epsilon)n_\ell$ *channel uses. Here*

$$H_2(\alpha) = \alpha \log\frac{1}{\alpha} + (1-\alpha)\log\frac{1}{1-\alpha}.$$

- Intuition: The activity uncertainty of ℓ users is $\ell H_2(k_\ell/\ell)$.
- Achievable code: random Gaussian sequences as signatures.

Main result 1: identification cost

Theorem

Let

$$n_\ell = \frac{\ell\, H_2(k_\ell/\ell)}{\frac{1}{2}\log(1+k_\ell\gamma)}.$$

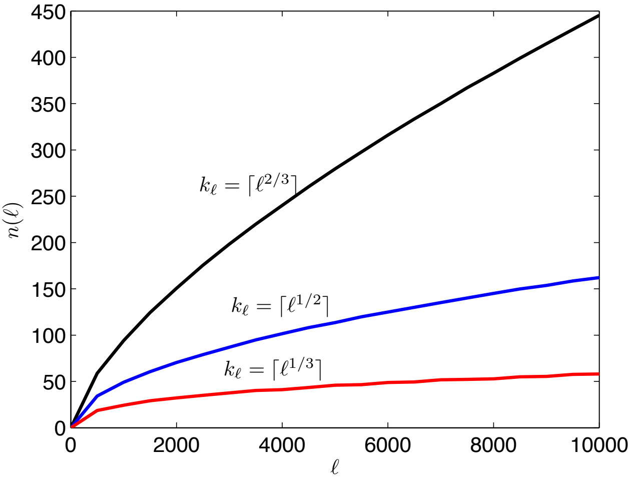
For every $\epsilon \in (0,1)$, as $\ell \to \infty$, arbitrarily reliable identification ($p_\ell \to 0$) is achievable with $(1+\epsilon)n_\ell$ channel uses; whereas $p_\ell \to 0$ is not achievable with $(1-\epsilon)n_\ell$ channel uses. Here

$$H_2(\alpha) = \alpha \log \frac{1}{\alpha} + (1-\alpha)\log\frac{1}{1-\alpha}.$$

- Intuition: The activity uncertainty of ℓ users is $\ell H_2(k_\ell/\ell)$.
- Achievable code: random Gaussian sequences as signatures.

Identification cost vs. user number

$\gamma = 10$ dB



Identification and channel code for the Gaussian MnAC

An (M, n) symmetric code:

1. Encoders $\mathcal{E}_k : \{0, \dots, M\} \to \mathcal{S}_k^n$ yields codewords $\boldsymbol{s}_k(0), \cdots, \boldsymbol{s}_k(M)$.

$$\frac{1}{n}\sum_{i=1}^{n} s_{ki}^2(w) \leq \gamma, \quad \forall w \in \{1, \cdots, M\}.$$

$\boldsymbol{s}_k(0) = \mathbf{0}$.

2. Decoder $\mathcal{D} : \mathcal{Y}^n \to \{0, \dots, M\}^{\ell_n}$.

I.i.d. messages $\{W_k\}$,

$$P\{W_k = w\} = \begin{cases} 1 - \alpha, & w = 0, \\ \frac{\alpha}{M}, & w \in \{1, \cdots, M\}. \end{cases}$$

$$P_e^{(n)} = P\left\{\mathcal{D}(\boldsymbol{Y}) \neq (W_1, \dots, W_{\ell_n})\right\}.$$

Identification and channel code for the Gaussian MnAC

An (M,n) symmetric code:

1. Encoders $\mathcal{E}_k : \{0,\dots,M\} \to \mathcal{S}_k^n$ yields codewords $\boldsymbol{s}_k(0),\cdots,\boldsymbol{s}_k(M)$.

$$\frac{1}{n}\sum_{i=1}^{n} s_{ki}^2(w) \le \gamma, \quad \forall w \in \{1,\cdots,M\}.$$

   $\boldsymbol{s}_k(0) = \mathbf{0}$.

2. Decoder $\mathcal{D} : \mathcal{Y}^n \to \{0,\dots,M\}^{\ell_n}$.

I.i.d. messages $\{W_k\}$,

$$P\{W_k = w\} = \begin{cases} 1-\alpha, & w = 0, \\ \frac{\alpha}{M}, & w \in \{1,\cdots,M\}. \end{cases}$$

$$P_e^{(n)} = P\left\{\mathcal{D}(\boldsymbol{Y}) \neq (W_1,\dots,W_{\ell_n})\right\}.$$

Identification and channel code for the Gaussian MnAC

An (M, n) symmetric code:

1. Encoders $\mathcal{E}_k : \{0, \dots, M\} \to \mathcal{S}_k^n$ yields codewords $\boldsymbol{s}_k(0), \cdots, \boldsymbol{s}_k(M)$.

$$\frac{1}{n}\sum_{i=1}^{n} s_{ki}^2(w) \leq \gamma, \quad \forall w \in \{1, \cdots, M\}.$$

   $\boldsymbol{s}_k(0) = \mathbf{0}$.

2. Decoder $\mathcal{D} : \mathcal{Y}^n \to \{0, \dots, M\}^{\ell_n}$.

I.i.d. messages $\{W_k\}$,

$$P\{W_k = w\} = \begin{cases} 1-\alpha, & w = 0, \\ \frac{\alpha}{M}, & w \in \{1, \cdots, M\}. \end{cases}$$

$$P_e^{(n)} = P\left\{\mathcal{D}(\boldsymbol{Y}) \neq (W_1, \dots, W_{\ell_n})\right\}.$$

Capacity

- $(v(n))_{n=1}^{\infty}$ with $v(n) > 1$ is a sequence of asymptotically achievable message lengths for the MnAC if there exists a sequence of $(\lceil \exp(v(n)) \rceil, n)$ codes such that $P_e^{(n)}$ vanishes as $n \to \infty$.
- $\mathcal{B} = (B(n))_{n=1}^{\infty}$ is said to be *a* symmetric capacity of the MnAC channel if, for every $\epsilon \in (0,1)$, $((1-\epsilon)B(n))$ is asymptotically achievable but $((1+\epsilon)B(n))$ is not.
- If $(B(n))$ is a capacity, then $(B(n) + o(B(n)))$ is also a capacity.
- In classical IT, $B(n) = nC$;
  In Many-User IT, $B(n)$ may be nonlinear.

Capacity

- $(v(n))_{n=1}^{\infty}$ with $v(n) > 1$ is a sequence of asymptotically achievable message lengths for the MnAC if there exists a sequence of $(\lceil \exp(v(n)) \rceil, n)$ codes such that $P_e^{(n)}$ vanishes as $n \to \infty$.
- $\mathcal{B} = (B(n))_{n=1}^{\infty}$ is said to be *a* symmetric capacity of the MnAC channel if, for every $\epsilon \in (0, 1)$, $((1 - \epsilon)B(n))$ is asymptotically achievable but $((1 + \epsilon)B(n))$ is not.
- If $(B(n))$ is a capacity, then $(B(n) + o(B(n)))$ is also a capacity.
- In classical IT, $B(n) = nC$;
  In Many-User IT, $B(n)$ may be nonlinear.

Capacity

- $(v(n))_{n=1}^{\infty}$ with $v(n) > 1$ is a sequence of asymptotically achievable message lengths for the MnAC if there exists a sequence of $(\lceil \exp(v(n)) \rceil, n)$ codes such that $P_e^{(n)}$ vanishes as $n \to \infty$.
- $\mathcal{B} = (B(n))_{n=1}^{\infty}$ is said to be *a* symmetric capacity of the MnAC channel if, for every $\epsilon \in (0, 1)$, $((1 - \epsilon)B(n))$ is asymptotically achievable but $((1 + \epsilon)B(n))$ is not.
- If $(B(n))$ is a capacity, then $(B(n) + o(B(n)))$ is also a capacity.
- In classical IT, $B(n) = nC$;
  In Many-User IT, $B(n)$ may be nonlinear.

Capacity

- $(v(n))_{n=1}^{\infty}$ with $v(n) > 1$ is a sequence of asymptotically achievable message lengths for the MnAC if there exists a sequence of $(\lceil \exp(v(n)) \rceil, n)$ codes such that $P_e^{(n)}$ vanishes as $n \to \infty$.
- $\mathcal{B} = (B(n))_{n=1}^{\infty}$ is said to be *a* symmetric capacity of the MnAC channel if, for every $\epsilon \in (0,1)$, $((1-\epsilon)B(n))$ is asymptotically achievable but $((1+\epsilon)B(n))$ is not.
- If $(B(n))$ is a capacity, then $(B(n) + o(B(n)))$ is also a capacity.
- In classical IT, $B(n) = nC$;
  In Many-User IT, $B(n)$ may be nonlinear.

Main result 2: capacity of Gaussian MnAC

Theorem (Symmetric capacity)

Suppose $k_n \to \infty$, $k_n = O(n)$, *and* $\ell_n = o(e^{\delta k_n})$, $\forall \delta > 0$.

$$B(n) = \left(\frac{n}{2k_n} \log(1 + k_n \gamma) - \frac{H_2(\alpha)}{\alpha} \right)^+$$
$$= \left(B_1(n) - \frac{H_2(\alpha)}{\alpha} \right)^+ \quad \textit{nats}$$

- The capacity is $B_1(n)$ if $\alpha = 1$ or the set of active users is known.
- The penalty $H_2(\alpha)/\alpha$ is the total amount of activity uncertainty divided by the number of active users.
- If $H_2(\alpha)/\alpha > B_1(n)$, an average user cannot send 1 bit reliably.
- Large-system analysis ($k \to \infty$ after $n \to \infty$) obliterates the identification cost.

Main result 2: capacity of Gaussian MnAC

Theorem (Symmetric capacity)

Suppose $k_n \to \infty$, $k_n = O(n)$, *and* $\ell_n = o(e^{\delta k_n})$, $\forall \delta > 0$.

$$B(n) = \left(\frac{n}{2k_n} \log(1 + k_n \gamma) - \frac{H_2(\alpha)}{\alpha} \right)^+$$
$$= \left(B_1(n) - \frac{H_2(\alpha)}{\alpha} \right)^+ \quad \textit{nats}$$

- The capacity is $B_1(n)$ if $\alpha = 1$ or the set of active users is known.
- The penalty $H_2(\alpha)/\alpha$ is the total amount of activity uncertainty divided by the number of active users.
- If $H_2(\alpha)/\alpha > B_1(n)$, an average user cannot send 1 bit reliably.
- Large-system analysis ($k \to \infty$ after $n \to \infty$) obliterates the identification cost.

Main result 2: capacity of Gaussian MnAC

Theorem (Symmetric capacity)

Suppose $k_n \to \infty$, $k_n = O(n)$, *and* $\ell_n = o(e^{\delta k_n})$, $\forall \delta > 0$.

$$B(n) = \left(\frac{n}{2k_n} \log(1 + k_n \gamma) - \frac{H_2(\alpha)}{\alpha} \right)^+$$
$$= \left(B_1(n) - \frac{H_2(\alpha)}{\alpha} \right)^+ \quad \textit{nats}$$

- The capacity is $B_1(n)$ if $\alpha = 1$ or the set of active users is known.
- The penalty $H_2(\alpha)/\alpha$ is the total amount of activity uncertainty divided by the number of active users.
- If $H_2(\alpha)/\alpha > B_1(n)$, an average user cannot send 1 bit reliably.
- Large-system analysis ($k \to \infty$ after $n \to \infty$) obliterates the identification cost.

Main result 2: capacity of Gaussian MnAC

Theorem (Symmetric capacity)

Suppose $k_n \to \infty$, $k_n = O(n)$, *and* $\ell_n = o(e^{\delta k_n})$, $\forall \delta > 0$.

$$B(n) = \left(\frac{n}{2k_n} \log(1 + k_n \gamma) - \frac{H_2(\alpha)}{\alpha} \right)^+$$
$$= \left(B_1(n) - \frac{H_2(\alpha)}{\alpha} \right)^+ \quad \textit{nats}$$

- The capacity is $B_1(n)$ if $\alpha = 1$ or the set of active users is known.
- The penalty $H_2(\alpha)/\alpha$ is the total amount of activity uncertainty divided by the number of active users.
- If $H_2(\alpha)/\alpha > B_1(n)$, an average user cannot send 1 bit reliably.
- Large-system analysis ($k \to \infty$ after $n \to \infty$) obliterates the identification cost.

Main result 2: capacity of Gaussian MnAC

Theorem (Symmetric capacity)

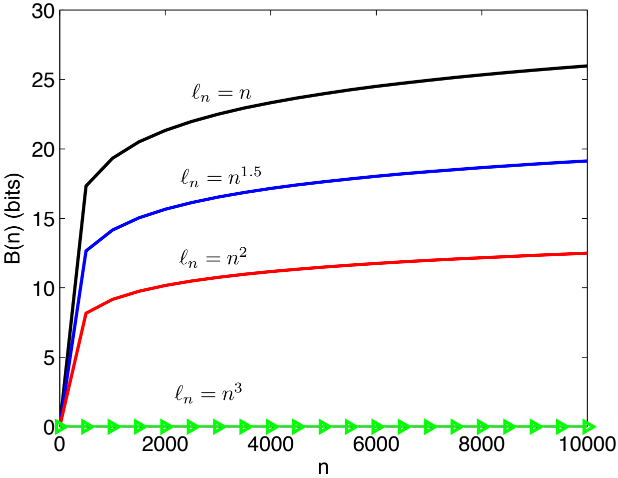
Suppose $k_n \to \infty$, $k_n = O(n)$, *and* $\ell_n = o(e^{\delta k_n})$, $\forall \delta > 0$.

$$B(n) = \left(\frac{n}{2k_n} \log(1 + k_n \gamma) - \frac{H_2(\alpha)}{\alpha} \right)^+$$
$$= \left(B_1(n) - \frac{H_2(\alpha)}{\alpha} \right)^+ \quad \textit{nats}$$

- The capacity is $B_1(n)$ if $\alpha = 1$ or the set of active users is known.
- The penalty $H_2(\alpha)/\alpha$ is the total amount of activity uncertainty divided by the number of active users.
- If $H_2(\alpha)/\alpha > B_1(n)$, an average user cannot send 1 bit reliably.
- Large-system analysis ($k \to \infty$ after $n \to \infty$) obliterates the identification cost.

Capacity (message length) vs. blocklength

$\gamma = 10$ dB, $k_n = n/4$



Proof using an equivalent model

$$\boldsymbol{Y} = \underline{\boldsymbol{S}}\boldsymbol{X} + \boldsymbol{Z}$$

- Concatenated codebook

$$\underline{\boldsymbol{S}} = [\boldsymbol{s}_1(1), \cdots, \boldsymbol{s}_1(M), \ldots, \boldsymbol{s}_{\ell_n}(1), \cdots, \boldsymbol{s}_{\ell_n}(M)]_{n \times (M\ell_n)}$$

- $\boldsymbol{X} \in \{0,1\}^{M\ell_n}$ selects the codewords:

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \vdots \\ \boldsymbol{X}_{\ell_n} \end{bmatrix}$$

$\boldsymbol{X}_k = \boldsymbol{0}$ or $\boldsymbol{e}_j = [0 \ldots 0\ 1\ 0 \ldots 0]^\top$

Proof using an equivalent model

$$\boldsymbol{Y} = \underline{\boldsymbol{S}}\boldsymbol{X} + \boldsymbol{Z}$$

- Concatenated codebook

$$\underline{\boldsymbol{S}} = [\boldsymbol{s}_1(1), \cdots, \boldsymbol{s}_1(M), \ldots, \boldsymbol{s}_{\ell_n}(1), \cdots, \boldsymbol{s}_{\ell_n}(M)]_{n \times (M\ell_n)}$$

- $\boldsymbol{X} \in \{0,1\}^{M\ell_n}$ selects the codewords:

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \vdots \\ \boldsymbol{X}_{\ell_n} \end{bmatrix}$$

$\boldsymbol{X}_k = \boldsymbol{0}$ or $\boldsymbol{e}_j = [0 \ldots 0\ 1\ 0 \ldots 0]^\top$

Proof using an equivalent model

$$\boldsymbol{Y} = \underline{\boldsymbol{S}}\boldsymbol{X} + \boldsymbol{Z}$$

- Concatenated codebook

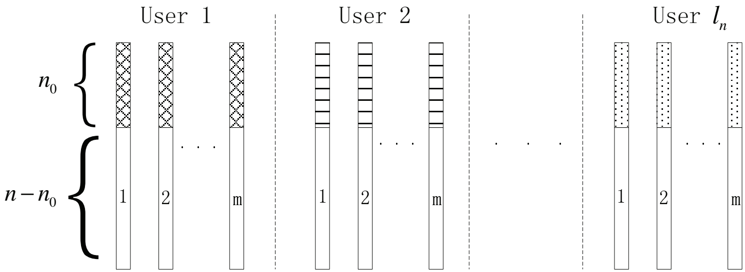
$$\underline{\boldsymbol{S}} = [\boldsymbol{s}_1(1), \cdots, \boldsymbol{s}_1(M), \ldots, \boldsymbol{s}_{\ell_n}(1), \cdots, \boldsymbol{s}_{\ell_n}(M)]_{n\times(M\ell_n)}$$

- $\boldsymbol{X} \in \{0,1\}^{M\ell_n}$ selects the codewords:

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \vdots \\ \boldsymbol{X}_{\ell_n} \end{bmatrix}$$

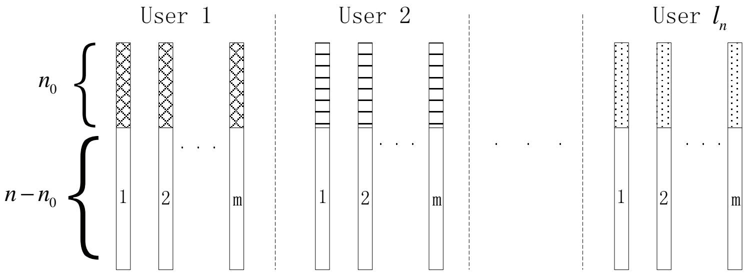
$\boldsymbol{X}_k = \mathbf{0}$ or $\boldsymbol{e}_j = [0\ldots0\ 1\ 0\ldots0]^\top$

Proof of achievability: The codebooks



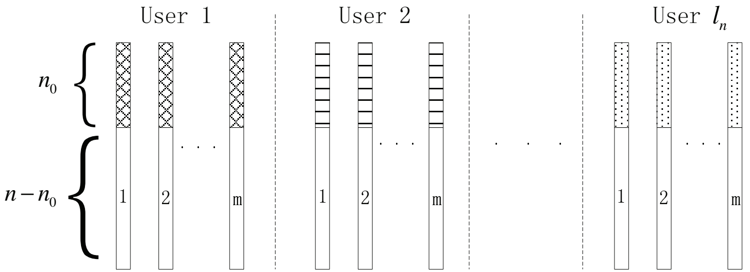
- The first n_0 symbols form a user-specific signature. The remaining symbols carry data.
- The capacity is achieved by separate identification and decoding.
- X. Chen and D. Guo, "Gaussian many-access channels with random user activities," ISIT 2014.

Proof of achievability: The codebooks



- The first n_0 symbols form a user-specific signature. The remaining symbols carry data.
- The capacity is achieved by separate identification and decoding.
- X. Chen and D. Guo, "Gaussian many-access channels with random user activities," ISIT 2014.

Proof of achievability: The codebooks



- The first n_0 symbols form a user-specific signature. The remaining symbols carry data.
- The capacity is achieved by separate identification and decoding.
- X. Chen and D. Guo, "Gaussian many-access channels with random user activities," ISIT 2014.

Achievability: Separate identification and decoding

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}^a_{n_0\times 1} \\ \boldsymbol{Y}^b_{(n-n_0)\times 1} \end{bmatrix} = \begin{bmatrix} \underline{\boldsymbol{S}}^a \boldsymbol{X}^a + \boldsymbol{Z}^a \\ \underline{\boldsymbol{S}}^b \boldsymbol{X} + \boldsymbol{Z}^b \end{bmatrix}$$

1. Identification:

$$\begin{aligned} \text{minimize} \quad & \|\boldsymbol{Y}^a - \underline{\boldsymbol{S}}^a \boldsymbol{x}^a\|_2 \\ \text{subject to} \quad & \boldsymbol{x}^a \in \{0,1\}^{\ell_n} \\ & \sum_{i=1}^{l_n} x_i^a \le (1 + 2k_n^{-\frac{1}{3}})k_n \end{aligned}$$

2. ML joint message decoding based on result of identification.

$$\begin{aligned} \text{minimize} \quad & \|\boldsymbol{Y}^b - \underline{\boldsymbol{S}}^b \boldsymbol{x}\|_2 \\ \text{subject to} \quad & \boldsymbol{x}_k = \boldsymbol{e}_{w_k}, \ \forall k = 1, \dots, k_n. \end{aligned}$$

Achievability: Separate identification and decoding

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}^a_{n_0 \times 1} \\ \boldsymbol{Y}^b_{(n-n_0)\times 1} \end{bmatrix} = \begin{bmatrix} \underline{\boldsymbol{S}}^a \boldsymbol{X}^a + \boldsymbol{Z}^a \\ \underline{\boldsymbol{S}}^b \boldsymbol{X} + \boldsymbol{Z}^b \end{bmatrix}$$

1. Identification:

$$\begin{aligned} \text{minimize} \quad & \|\boldsymbol{Y}^a - \underline{\boldsymbol{S}}^a \boldsymbol{x}^a\|_2 \\ \text{subject to} \quad & \boldsymbol{x}^a \in \{0,1\}^{\ell_n} \\ & \sum_{i=1}^{l_n} x_i^a \le (1 + 2k_n^{-\frac{1}{3}})k_n \end{aligned}$$

2. ML joint message decoding based on result of identification.

$$\begin{aligned} \text{minimize} \quad & \|\boldsymbol{Y}^b - \underline{\boldsymbol{S}}^b \boldsymbol{x}\|_2 \\ \text{subject to} \quad & \boldsymbol{x}_k = \boldsymbol{e}_{w_k}, \;\; \forall k = 1, \dots, k_n. \end{aligned}$$

Achievability: Separate identification and decoding

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}^a_{n_0 \times 1} \\ \boldsymbol{Y}^b_{(n-n_0)\times 1} \end{bmatrix} = \begin{bmatrix} \underline{\boldsymbol{S}}^a \boldsymbol{X}^a + \boldsymbol{Z}^a \\ \underline{\boldsymbol{S}}^b \boldsymbol{X} + \boldsymbol{Z}^b \end{bmatrix}$$

1. Identification:

$$\begin{aligned} \text{minimize} \quad & \|\boldsymbol{Y}^a - \underline{\boldsymbol{S}}^a \boldsymbol{x}^a\|_2 \\ \text{subject to} \quad & \boldsymbol{x}^a \in \{0,1\}^{\ell_n} \\ & \sum_{i=1}^{l_n} x_i^a \le (1 + 2k_n^{-\frac{1}{3}}) k_n \end{aligned}$$

2. ML joint message decoding based on result of identification.

$$\begin{aligned} \text{minimize} \quad & \|\boldsymbol{Y}^b - \underline{\boldsymbol{S}}^b \boldsymbol{x}\|_2 \\ \text{subject to} \quad & \boldsymbol{x}_k = \boldsymbol{e}_{w_k}, \ \ \forall k = 1, \dots, k_n. \end{aligned}$$

Identification error probability

- Simple union bound fails for the exponential number of error events.
- The identification error event is decomposed to polynomial number of events (t_1 misses and t_2 false alarms):

$$\mathcal{E} = \bigcup_{t_1,t_2} \mathcal{E}_{t_1,t_2}$$

- Using techniques from Gallager '68 to upper bound the error

$$\mathsf{P}\left\{\mathcal{E}_{t_1,t_2} | \boldsymbol{X}^a = \boldsymbol{x}^a\right\} \leq e^{-nh_n(t_1,t_2)}$$

where for large enough n, $h_n(t_1,t_2) \geq c_0(\epsilon) > 0$, $\forall (t_1,t_2)$.

- Correct identification with high probability with n_0 channel uses.

Identification error probability

- Simple union bound fails for the exponential number of error events.
- The identification error event is decomposed to polynomial number of events (t_1 misses and t_2 false alarms):

$$\mathcal{E} = \bigcup_{t_1, t_2} \mathcal{E}_{t_1, t_2}$$

- Using techniques from Gallager '68 to upper bound the error

$$\mathsf{P}\left\{\mathcal{E}_{t_1,t_2} | \boldsymbol{X}^a = \boldsymbol{x}^a\right\} \le e^{-n h_n(t_1, t_2)}$$

where for large enough n, $h_n(t_1, t_2) \ge c_0(\epsilon) > 0$, $\forall (t_1, t_2)$.
- Correct identification with high probability with n_0 channel uses.

Identification error probability

- Simple union bound fails for the exponential number of error events.
- The identification error event is decomposed to polynomial number of events (t_1 misses and t_2 false alarms):

$$\mathcal{E} = \bigcup_{t_1,t_2} \mathcal{E}_{t_1,t_2}$$

- Using techniques from Gallager '68 to upper bound the error

$$\mathsf{P}\left\{\mathcal{E}_{t_1,t_2} | \boldsymbol{X}^a = \boldsymbol{x}^a\right\} \leq e^{-nh_n(t_1,t_2)}$$

where for large enough n, $h_n(t_1,t_2) \geq c_0(\epsilon) > 0$, $\forall(t_1,t_2)$.

- Correct identification with high probability with n_0 channel uses.

Identification error probability

- Simple union bound fails for the exponential number of error events.
- The identification error event is decomposed to polynomial number of events (t_1 misses and t_2 false alarms):

$$\mathcal{E} = \bigcup_{t_1,t_2} \mathcal{E}_{t_1,t_2}$$

- Using techniques from Gallager '68 to upper bound the error

$$\mathsf{P}\left\{\mathcal{E}_{t_1,t_2} | \boldsymbol{X}^a = \boldsymbol{x}^a\right\} \leq e^{-nh_n(t_1,t_2)}$$

  where for large enough n, $h_n(t_1, t_2) \geq c_0(\epsilon) > 0$, $\forall (t_1, t_2)$.
- Correct identification with high probability with n_0 channel uses.

ML decoding error probability

- Characterized similarly by lower bounding the error exponents.
- k_n active users, decomposed to k_n events according to # of users in error:

$$\mathcal{E} = \bigcup_{k=1}^{k_n} E_k$$

- Error exponent

$$P\{E_k\} \leq e^{-nf(k,\rho)}, \quad \forall \rho \in [0,1]$$

$$f(k,\rho) = E_0\left(\frac{k}{k_n}, \rho\right) - \rho\frac{k}{n}\log M - \frac{k_n}{n}H_2\left(\frac{k}{k_n}\right).$$

- Let $\log M = B(n) - \epsilon n/k_n$. For large enough n, $\exists d(\epsilon) > 0$, s.t.

$$\min_k \max_{\rho\in[0,1]} f(k,\rho) \geq d(\epsilon).$$

- The decoding error probability vanishes. Hence the achievability.

ML decoding error probability

- Characterized similarly by lower bounding the error exponents.
- k_n active users, decomposed to k_n events according to # of users in error:

$$\mathcal{E} = \bigcup_{k=1}^{k_n} E_k$$

- Error exponent

$$P\{E_k\} \leq e^{-nf(k,\rho)}, \quad \forall \rho \in [0,1]$$
$$f(k,\rho) = E_0\left(\frac{k}{k_n}, \rho\right) - \rho \frac{k}{n} \log M - \frac{k_n}{n} H_2\left(\frac{k}{k_n}\right).$$

- Let $\log M = B(n) - \epsilon n / k_n$. For large enough n, $\exists d(\epsilon) > 0$, s.t.

$$\min_k \max_{\rho \in [0,1]} f(k,\rho) \geq d(\epsilon).$$

- The decoding error probability vanishes. Hence the achievability.

ML decoding error probability

- Characterized similarly by lower bounding the error exponents.
- k_n active users, decomposed to k_n events according to # of users in error:

$$\mathcal{E} = \bigcup_{k=1}^{k_n} E_k$$

- Error exponent

$$P\{E_k\} \le e^{-nf(k,\rho)}, \quad \forall \rho \in [0,1]$$
$$f(k,\rho) = E_0\left(\frac{k}{k_n},\rho\right) - \rho\frac{k}{n}\log M - \frac{k_n}{n}H_2\left(\frac{k}{k_n}\right).$$

- Let $\log M = B(n) - \epsilon n/k_n$. For large enough n, $\exists d(\epsilon) > 0$, s.t.

$$\min_k \max_{\rho\in[0,1]} f(k,\rho) \ge d(\epsilon).$$

- The decoding error probability vanishes. Hence the achievability.

ML decoding error probability

- Characterized similarly by lower bounding the error exponents.
- k_n active users, decomposed to k_n events according to # of users in error:

$$\mathcal{E} = \bigcup_{k=1}^{k_n} E_k$$

- Error exponent

$$P\{E_k\} \leq e^{-nf(k,\rho)}, \quad \forall \rho \in [0,1]$$
$$f(k,\rho) = E_0\left(\frac{k}{k_n}, \rho\right) - \rho\frac{k}{n}\log M - \frac{k_n}{n}H_2\left(\frac{k}{k_n}\right).$$

- Let $\log M = B(n) - \epsilon n/k_n$. For large enough n, $\exists d(\epsilon) > 0$, s.t.

$$\min_k \max_{\rho\in[0,1]} f(k,\rho) \geq d(\epsilon).$$

- The decoding error probability vanishes. Hence the achievability.

ML decoding error probability

- Characterized similarly by lower bounding the error exponents.
- k_n active users, decomposed to k_n events according to # of users in error:

$$\mathcal{E} = \bigcup_{k=1}^{k_n} E_k$$

- Error exponent

$$P\{E_k\} \leq e^{-nf(k,\rho)}, \quad \forall \rho \in [0,1]$$

$$f(k,\rho) = E_0\left(\frac{k}{k_n}, \rho\right) - \rho\frac{k}{n}\log M - \frac{k_n}{n}H_2\left(\frac{k}{k_n}\right).$$

- Let $\log M = B(n) - \epsilon n/k_n$. For large enough n, $\exists d(\epsilon) > 0$, s.t.

$$\min_k \max_{\rho\in[0,1]} f(k,\rho) \geq d(\epsilon).$$

- The decoding error probability vanishes. Hence the achievability.

Recap

Theorem (Symmetric capacity)

$$B(n) = \left(\frac{n}{2k_n} \log(1 + k_n\gamma) - \frac{H_2(\alpha_n)}{\alpha_n} \right)^+$$
$$= \left(B_1(n) - \frac{H_2(\alpha_n)}{\alpha_n} \right)^+ \quad \textit{nats.}$$

- Achieved by using random Gaussian codebooks with separate indentification and decoding.
- Large-system analysis ($k \to \infty$ after $n \to \infty$) would obliterate the identification cost.

Recap

Theorem (Symmetric capacity)

$$B(n) = \left(\frac{n}{2k_n} \log(1 + k_n \gamma) - \frac{H_2(\alpha_n)}{\alpha_n} \right)^+$$
$$= \left(B_1(n) - \frac{H_2(\alpha_n)}{\alpha_n} \right)^+ \quad \textit{nats.}$$

- Achieved by using random Gaussian codebooks with separate indentification and decoding.
- Large-system analysis ($k \to \infty$ after $n \to \infty$) would obliterate the identification cost.

Recap

Theorem (Symmetric capacity)

$$B(n) = \left(\frac{n}{2k_n} \log(1 + k_n\gamma) - \frac{H_2(\alpha_n)}{\alpha_n} \right)^+$$
$$= \left(B_1(n) - \frac{H_2(\alpha_n)}{\alpha_n} \right)^+ \quad \textit{nats.}$$

- Achieved by using random Gaussian codebooks with separate indentification and decoding.
- Large-system analysis ($k \to \infty$ after $n \to \infty$) would obliterate the identification cost.

Outline

- (Almost practical) neighbor discovery
- Classical information theory
- Many-access channel
- Many-broadcast channel

Classical degraded broadcast channels (BC)

- 2-user BC: $P_{Y_1 Y_2|X}$.
- Degraded if X–Y_1–Y_2 is Markov.
- The capacity region is

$$\bigcup_{P_{XU}:U-X-Y_1-Y_2} \left\{ (R_1, R_2) : \begin{array}{c} 0 \le R_2 \le I(U;Y_2) \\ 0 \le R_1 \le I(X;Y_1|U) \end{array} \right\}$$

- Generalizing to a k-user degraded BC:

$$R_j \le I(U_j; Y_j | U_{j+1}), \quad j = 1, \dots, k$$

where $(0 = U_{k+1})$–U_k–$\dots$–$(U_1 = X)$–Y_1–$\dots$–Y_k.

Classical degraded broadcast channels (BC)

- 2-user BC: $P_{Y_1Y_2|X}$.
- Degraded if X–Y_1–Y_2 is Markov.
- The capacity region is

$$\bigcup_{P_{XU}:U-X-Y_1-Y_2}\left\{(R_1,R_2): \begin{array}{c} 0\le R_2\le I(U;Y_2) \\ 0\le R_1\le I(X;Y_1|U) \end{array}\right\}$$

- Generalizing to a k-user degraded BC:

$$R_j \le I(U_j;Y_j|U_{j+1}),\quad j=1,\dots,k$$

where $(0=U_{k+1})$–U_k–$\dots$–$(U_1=X)$–Y_1–$\dots$–Y_k.

Classical degraded broadcast channels (BC)

- 2-user BC: $P_{Y_1Y_2|X}$.
- Degraded if X–Y_1–Y_2 is Markov.
- The capacity region is

$$\bigcup_{P_{XU}:U-X-Y_1-Y_2} \left\{ (R_1, R_2) : \begin{array}{c} 0 \le R_2 \le I(U;Y_2) \\ 0 \le R_1 \le I(X;Y_1|U) \end{array} \right\}$$

- Generalizing to a k-user degraded BC:

$$R_j \le I(U_j;Y_j|U_{j+1}), \quad j = 1, \dots, k$$

where $(0 = U_{k+1})$–U_k–$\dots$–$(U_1 = X)$–Y_1–$\dots$–Y_k.

Classical degraded broadcast channels (BC)

- 2-user BC: $P_{Y_1 Y_2|X}$.
- Degraded if X–Y_1–Y_2 is Markov.
- The capacity region is

$$\bigcup_{P_{XU}:U-X-Y_1-Y_2} \left\{ (R_1, R_2) : \begin{array}{c} 0 \le R_2 \le I(U;Y_2) \\ 0 \le R_1 \le I(X;Y_1|U) \end{array} \right\}$$

- Generalizing to a k-user degraded BC:

$$R_j \le I(U_j; Y_j | U_{j+1}), \quad j = 1, \ldots, k$$

where $(0 = U_{k+1})$–U_k–$\ldots$–$(U_1 = X)$–Y_1–$\ldots$–Y_k.

Gaussian degraded many-broadcast channel (MnBC)

- k_n channel outputs:

$$Y_j = X + \sigma_{n,j} Z_j, \quad j = 1, 2, \dots, k_n$$

  where Z_j i.i.d.$\sim \mathcal{N}(0,1)$ and $\sigma_{n,j} \le \sigma_{n,j+1}$.
- $k_n \to \infty$ monotonically. $k_n = O(n)$.
- Power constraint γ.
- How many bits can one send to each user reliably?
- Noise level as a triangular array $(\sigma_{n,j} : n = 1, 2, \dots; j = 1, \dots, k_n)$.

Gaussian degraded many-broadcast channel (MnBC)

- k_n channel outputs:

$$Y_j = X + \sigma_{n,j} Z_j, \quad j = 1, 2, \ldots, k_n$$

  where Z_j i.i.d.$\sim \mathcal{N}(0,1)$ and $\sigma_{n,j} \leq \sigma_{n,j+1}$.
- $k_n \to \infty$ monotonically. $k_n = O(n)$.
- Power constraint γ.
- How many bits can one send to each user reliably?
- Noise level as a triangular array $(\sigma_{n,j} : n = 1, 2, \ldots; j = 1, \ldots, k_n)$.

Gaussian degraded many-broadcast channel (MnBC)

- k_n channel outputs:

$$Y_j = X + \sigma_{n,j} Z_j, \quad j = 1, 2, \ldots, k_n$$

  where Z_j i.i.d.$\sim \mathcal{N}(0,1)$ and $\sigma_{n,j} \leq \sigma_{n,j+1}$.
- $k_n \to \infty$ monotonically. $k_n = O(n)$.
- Power constraint γ.
- How many bits can one send to each user reliably?
- Noise level as a triangular array $(\sigma_{n,j} : n = 1, 2, \ldots; j = 1, \ldots, k_n)$.

Gaussian degraded many-broadcast channel (MnBC)

- k_n channel outputs:

$$Y_j = X + \sigma_{n,j} Z_j, \quad j = 1, 2, \ldots, k_n$$

  where Z_j i.i.d.$\sim \mathcal{N}(0,1)$ and $\sigma_{n,j} \leq \sigma_{n,j+1}$.
- $k_n \to \infty$ monotonically. $k_n = O(n)$.
- Power constraint γ.
- How many bits can one send to each user reliably?
- Noise level as a triangular array $(\sigma_{n,j} : n = 1, 2, \ldots; j = 1, \ldots, k_n)$.

Gaussian degraded many-broadcast channel (MnBC)

- k_n channel outputs:

$$Y_j = X + \sigma_{n,j} Z_j, \quad j = 1, 2, \dots, k_n$$

  where Z_j i.i.d.$\sim \mathcal{N}(0,1)$ and $\sigma_{n,j} \le \sigma_{n,j+1}$.
- $k_n \to \infty$ monotonically. $k_n = O(n)$.
- Power constraint γ.
- How many bits can one send to each user reliably?
- Noise level as a triangular array $(\sigma_{n,j} : n = 1, 2, \dots; j = 1, \dots, k_n)$.

Definitions

- A triangular array

$$V = \begin{pmatrix} v_{1,1} & & & & \\ v_{2,1} & v_{2,2} & & & \\ v_{3,1} & v_{3,2} & v_{3,3} & & \\ & \vdots & & & \\ v_{n,1} & v_{n,2} & \cdots & \cdots & v_{n,k_n} \\ & \vdots & & \vdots & \end{pmatrix}$$

- A triangular array $V = (v_{n,j} : n = 1, 2, \dots; j = 1, \dots, k_n)$ describes an asymptotically achievable message length array for an MnBC if there exists a sequence of $\left(\left(2^{\lceil v_{n,j} \rceil} \right)_{j=1}^{k_n}, n \right)$ codes s.t. $\lim_{n\to\infty} P_e^{(n)} = 0$.
- The message length capacity of an MnBC is a collection of triangular arrays $\mathcal{B}$ such that $(1-\delta)\mathcal{B}$ is asymptotically achievable and $(1+\delta)\mathcal{B}$ is not asymptotically achievable.

Definitions

- A triangular array

$$V = \begin{pmatrix} v_{1,1} & & & & \\ v_{2,1} & v_{2,2} & & & \\ v_{3,1} & v_{3,2} & v_{3,3} & & \\ & \vdots & & & \\ v_{n,1} & v_{n,2} & \dots & \dots & v_{n,k_n} \\ & \vdots & & \vdots & \end{pmatrix}$$

- A triangular array $V = (v_{n,j} : n = 1, 2, \dots ; j = 1, \dots, k_n)$ describes an asymptotically achievable message length array for an MnBC if there exists a sequence of $\left(\left(2^{\lceil v_{n,j} \rceil} \right)_{j=1}^{k_n}, n \right)$ codes s.t. $\lim_{n \to \infty} P_e^{(n)} = 0$.
- The message length capacity of an MnBC is a collection of triangular arrays $\mathcal{B}$ such that $(1 - \delta)\mathcal{B}$ is asymptotically achievable and $(1 + \delta)\mathcal{B}$ is not asymptotically achievable.

Definitions

- A triangular array

$$V = \begin{pmatrix} v_{1,1} & & & & \\ v_{2,1} & v_{2,2} & & & \\ v_{3,1} & v_{3,2} & v_{3,3} & & \\ & \vdots & & & \\ v_{n,1} & v_{n,2} & \dots & \dots & v_{n,k_n} \\ & \vdots & & \vdots & \end{pmatrix}$$

- A triangular array $V = (v_{n,j} : n = 1, 2, \dots ; j = 1, \dots, k_n)$ describes an asymptotically achievable message length array for an MnBC if there exists a sequence of $\left(\left(2^{\lceil v_{n,j} \rceil}\right)_{j=1}^{k_n}, n\right)$ codes s.t. $\lim_{n\to\infty} P_e^{(n)} = 0$.
- The message length capacity of an MnBC is a collection of triangular arrays $\mathcal{B}$ such that $(1-\delta)\mathcal{B}$ is asymptotically achievable and $(1+\delta)\mathcal{B}$ is not asymptotically achievable.

Capacity results

- Let auxiliary variables $U_1, \ldots, U_{k_n}$ be jointly Gaussian.
- Power allocation described by a triangular array $(\alpha_{n,j} : n = 1, 2, \ldots; j = 1, \ldots, k_n)$, with $\sum_{j=1}^{k_n} \alpha_{n,j} = 1$.
- User j's SNR is then $\alpha_{n,j}\gamma/\sigma_j^2$.
- An asymptotically achievable triangular array:

$$B_{n,j} = B_j(n) = \frac{n}{2} \log \left(1 + \frac{\alpha_{n,j}\gamma}{\sigma_{n,j}^2 + \sum_{i=1}^{j-1} \alpha_{n,i}\gamma} \right).$$

Capacity results

- Let auxiliary variables $U_1, \dots, U_{k_n}$ be jointly Gaussian.
- Power allocation described by a triangular array $(\alpha_{n,j} : n = 1, 2, \dots ; j = 1, \dots, k_n)$, with $\sum_{j=1}^{k_n} \alpha_{n,j} = 1$.
- User j's SNR is then $\alpha_{n,j}\gamma/\sigma_j^2$.
- An asymptotically achievable triangular array:

$$B_{n,j} = B_j(n) = \frac{n}{2} \log \left(1 + \frac{\alpha_{n,j}\gamma}{\sigma_{n,j}^2 + \sum_{i=1}^{j-1} \alpha_{n,i}\gamma} \right).$$

Capacity results

- Let auxiliary variables $U_1, \ldots, U_{k_n}$ be jointly Gaussian.
- Power allocation described by a triangular array $(\alpha_{n,j} : n = 1, 2, \ldots; j = 1, \ldots, k_n)$, with $\sum_{j=1}^{k_n} \alpha_{n,j} = 1$.
- User j's SNR is then $\alpha_{n,j}\gamma/\sigma_j^2$.
- An asymptotically achievable triangular array:

$$B_{n,j} = B_j(n) = \frac{n}{2} \log \left(1 + \frac{\alpha_{n,j}\gamma}{\sigma_{n,j}^2 + \sum_{i=1}^{j-1} \alpha_{n,i}\gamma} \right).$$

Capacity results

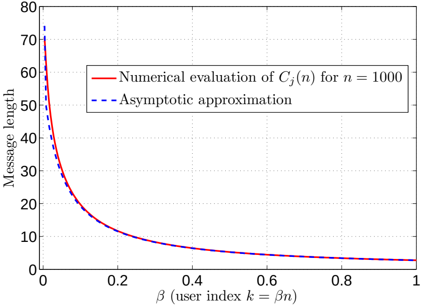
- Let auxiliary variables $U_1, \ldots, U_{k_n}$ be jointly Gaussian.
- Power allocation described by a triangular array $(\alpha_{n,j} : n = 1, 2, \ldots; \, j = 1, \ldots, k_n)$, with $\sum_{j=1}^{k_n} \alpha_{n,j} = 1$.
- User j's SNR is then $\alpha_{n,j}\gamma/\sigma_j^2$.
- An asymptotically achievable triangular array:

$$B_{n,j} = B_j(n) = \frac{n}{2} \log \left(1 + \frac{\alpha_{n,j}\gamma}{\sigma_{n,j}^2 + \sum_{i=1}^{j-1} \alpha_{n,i}\gamma} \right).$$

Gaussian MnBC: a numerical example

$n = 1000$, $\gamma = 20$, $k_n = 250$ (i.e., $c = 1/4$), and

$$\sigma_j^2 = \exp\left[-\frac{j}{(1+j)^2}\right].$$



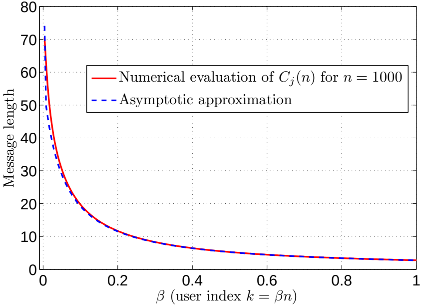
Using uniform power allocation:

- Can send ≥ 1 bits reliably to all users;
- Can send ≥ 10 bits reliably to about 20% of the users.

Gaussian MnBC: a numerical example

$n = 1000$, $\gamma = 20$, $k_n = 250$ (i.e., $c = 1/4$), and

$$\sigma_j^2 = \exp\left[-\frac{j}{(1+j)^2}\right].$$



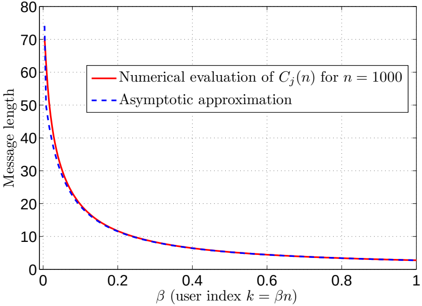
Using uniform power allocation:

- Can send ≥ 1 bits reliably to all users;
- Can send ≥ 10 bits reliably to about 20% of the users.

Gaussian MnBC: a numerical example

$n = 1000$, $\gamma = 20$, $k_n = 250$ (i.e., $c = 1/4$), and

$$\sigma_j^2 = \exp\left[-\frac{j}{(1+j)^2}\right].$$



Using uniform power allocation:

- Can send ≥ 1 bits reliably to all users;
- Can send ≥ 10 bits reliably to about 20% of the users.

Grouping

- What is the optimal strategy if a fraction q of the users can be dropped?
- It's optimal to drop the group of qk_n least capable users.
- In general, multiple groups with different rates—leading to a notion of "capacity region."

Grouping

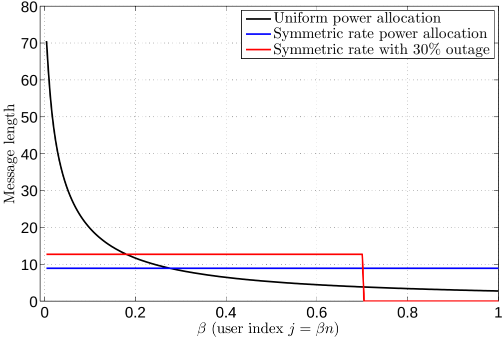
- What is the optimal strategy if a fraction q of the users can be dropped?
- It's optimal to drop the group of qk_n least capable users.
- In general, multiple groups with different rates—leading to a notion of "capacity region."

Grouping

- What is the optimal strategy if a fraction q of the users can be dropped?
- It's optimal to drop the group of qk_n least capable users.
- In general, multiple groups with different rates—leading to a notion of "capacity region."

Grouping

- What is the optimal strategy if a fraction q of the users can be dropped?
- It's optimal to drop the group of qk_n least capable users.
- In general, multiple groups with different rates—leading to a notion of "capacity region."



Many-Source coding?

$$\begin{bmatrix} X_{1,1}^n & X_{1,2}^n & \dots & X_{1,n}^n \\ X_{2,1}^n & X_{2,2}^n & \dots & X_{2,n}^n \\ \vdots & \vdots & & \vdots \\ X_{l_n,1}^n & X_{l_n,2}^n & \dots & X_{l_n,n}^n \end{bmatrix}$$

- No joint typicality in general.
- A certain joint typicality can be established if the sources form a Markov chain.

Many-Source coding?

$$\begin{bmatrix} X_{1,1}^n & X_{1,2}^n & \dots & X_{1,n}^n \\ X_{2,1}^n & X_{2,2}^n & \dots & X_{2,n}^n \\ \vdots & \vdots & & \vdots \\ X_{l_n,1}^n & X_{l_n,2}^n & \dots & X_{l_n,n}^n \end{bmatrix}$$

- No joint typicality in general.
- A certain joint typicality can be established if the sources form a Markov chain.

Conclusion

- Proposed a new many-user paradigm.
- Determined the minimum device identification cost.
- Determined the symmetric capacity of the Gaussian many-access channel with random user activities.
- Capacity results for the Gaussian degraded many-broadcast channel also develped (not shown here).
- large-system $\neq$ many-user.
- Ongoing: other many-user channel models, source coding, rate distortion.
- The goal is to develop a Many-User Information Theory for emerging many-user systems.

Conclusion

- Proposed a new many-user paradigm.
- Determined the minimum device identification cost.
- Determined the symmetric capacity of the Gaussian many-access channel with random user activities.
- Capacity results for the Gaussian degraded many-broadcast channel also develped (not shown here).
- large-system $\neq$ many-user.
- Ongoing: other many-user channel models, source coding, rate distortion.
- The goal is to develop a Many-User Information Theory for emerging many-user systems.

Conclusion

- Proposed a new many-user paradigm.
- Determined the minimum device identification cost.
- Determined the symmetric capacity of the Gaussian many-access channel with random user activities.
- Capacity results for the Gaussian degraded many-broadcast channel also develped (not shown here).
- large-system $\neq$ many-user.
- Ongoing: other many-user channel models, source coding, rate distortion.
- The goal is to develop a Many-User Information Theory for emerging many-user systems.

Conclusion

- Proposed a new many-user paradigm.
- Determined the minimum device identification cost.
- Determined the symmetric capacity of the Gaussian many-access channel with random user activities.
- Capacity results for the Gaussian degraded many-broadcast channel also develped (not shown here).
- large-system $\neq$ many-user.
- Ongoing: other many-user channel models, source coding, rate distortion.
- The goal is to develop a Many-User Information Theory for emerging many-user systems.

Conclusion

- Proposed a new many-user paradigm.
- Determined the minimum device identification cost.
- Determined the symmetric capacity of the Gaussian many-access channel with random user activities.
- Capacity results for the Gaussian degraded many-broadcast channel also develped (not shown here).
- large-system $\neq$ many-user.
- Ongoing: other many-user channel models, source coding, rate distortion.
- The goal is to develop a Many-User Information Theory for emerging many-user systems.

Conclusion

- Proposed a new many-user paradigm.
- Determined the minimum device identification cost.
- Determined the symmetric capacity of the Gaussian many-access channel with random user activities.
- Capacity results for the Gaussian degraded many-broadcast channel also develped (not shown here).
- large-system $\neq$ many-user.
- Ongoing: other many-user channel models, source coding, rate distortion.
- The goal is to develop a Many-User Information Theory for emerging many-user systems.

Conclusion

- Proposed a new many-user paradigm.
- Determined the minimum device identification cost.
- Determined the symmetric capacity of the Gaussian many-access channel with random user activities.
- Capacity results for the Gaussian degraded many-broadcast channel also develped (not shown here).
- large-system $\neq$ many-user.
- Ongoing: other many-user channel models, source coding, rate distortion.
- The goal is to develop a Many-User Information Theory for emerging many-user systems.

Remark: a related work

The only prior many-user model in the literature is the noiseless binary adder channel:

S.-C. Chang and E. Weldon, "Coding for t-user multiple-access channels," IEEE Trans. Inform. Theory, vol. 25, no. 6, pp. 684-691, 1979.

The number of users and blocklength taken to infinity simultaneously.

They studied uniquely decodable multiuser codes and the capacity.

Remark: large-system ≠ many-user

- Large-system analysis of CDMA and MIMO:
  Send to infinity the number of users and spreading factor simultaneously with fixed ratio; or the number of transmit and receive antennas with fixed ratio;
  Blocklength $n \to \infty$ before that.
  [Foschini & Gans '96, Telatar '99, Verdú & Shamai '99, Tanaka '02, Guo & Verdú '05, Huh, Tulino & Caire '12]
- Massive MIMO:
  First, $n \to \infty$.
  Then, send the number of antennas to infinity.
  [Rusek, Persson, Lau, Larsson, Marzetta, Edfors & Tufvesson '13, Hoydis, ten Brink & Debbah '13]
- The CEO problem [Berger & Zhang '96].
  $n \to \infty$ before the number of agents.
- Broadcast strategy for point-to-point slow-fading channels [Shamai '97].
  $n \to \infty$ before the number of layers.

Remark: large-system $\neq$ many-user

- Large-system analysis of CDMA and MIMO:
  Send to infinity the number of users and spreading factor simultaneously with fixed ratio; or the number of transmit and receive antennas with fixed ratio;
  Blocklength $n \to \infty$ before that.
  [Foschini & Gans '96, Telatar '99, Verdú & Shamai '99, Tanaka '02, Guo & Verdú '05, Huh, Tulino & Caire '12]
- Massive MIMO:
  First, $n \to \infty$.
  Then, send the number of antennas to infinity.
  [Rusek, Persson, Lau, Larsson, Marzetta, Edfors & Tufvesson '13, Hoydis, ten Brink & Debbah '13]
- The CEO problem [Berger & Zhang '96].
  $n \to \infty$ before the number of agents.
- Broadcast strategy for point-to-point slow-fading channels [Shamai '97].
  $n \to \infty$ before the number of layers.

Remark: large-system $\neq$ many-user

- Large-system analysis of CDMA and MIMO:
  Send to infinity the number of users and spreading factor simultaneously with fixed ratio; or the number of transmit and receive antennas with fixed ratio;
  Blocklength $n \to \infty$ before that.
  [Foschini & Gans '96, Telatar '99, Verdú & Shamai '99, Tanaka '02, Guo & Verdú '05, Huh, Tulino & Caire '12]
- Massive MIMO:
  First, $n \to \infty$.
  Then, send the number of antennas to infinity.
  [Rusek, Persson, Lau, Larsson, Marzetta, Edfors & Tufvesson '13, Hoydis, ten Brink & Debbah '13]
- The CEO problem [Berger & Zhang '96].
  $n \to \infty$ before the number of agents.
- Broadcast strategy for point-to-point slow-fading channels [Shamai '97].
  $n \to \infty$ before the number of layers.

Remark: large-system $\neq$ many-user

- Large-system analysis of CDMA and MIMO:
  Send to infinity the number of users and spreading factor simultaneously with fixed ratio; or the number of transmit and receive antennas with fixed ratio;
  Blocklength $n \to \infty$ before that.
  [Foschini & Gans '96, Telatar '99, Verdú & Shamai '99, Tanaka '02, Guo & Verdú '05, Huh, Tulino & Caire '12]
- Massive MIMO:
  First, $n \to \infty$.
  Then, send the number of antennas to infinity.
  [Rusek, Persson, Lau, Larsson, Marzetta, Edfors & Tufvesson '13, Hoydis, ten Brink & Debbah '13]
- The CEO problem [Berger & Zhang '96].
  $n \to \infty$ before the number of agents.
- Broadcast strategy for point-to-point slow-fading channels [Shamai '97].
  $n \to \infty$ before the number of layers.