The “Hallucination Bound” for the BSC

Anant Sahai and Stark Draper

Wireless Foundations
Department of Electrical Engineering and Computer Sciences
University of California at Berkeley

ECE Department
University of Wisconsin at Madison

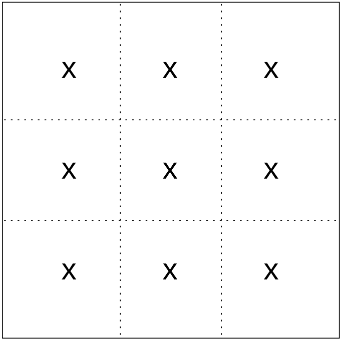
Major Support from NSF CISE

July 8th 2008: ISIT Tu-AM-3.3

Outline

1. Motivation and review
2. Block-coding converse: minimum error probability
3. Streaming converse: bit error probability

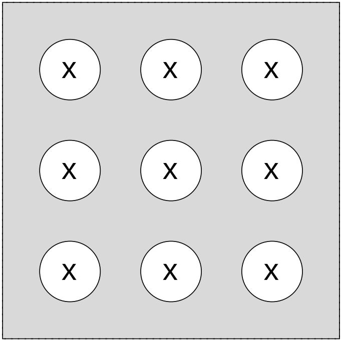
Review: Fixed blocks



- Hard decision regions cover space

- Feedback is pointless at high rates. (Dobrushin and Haroutunian)

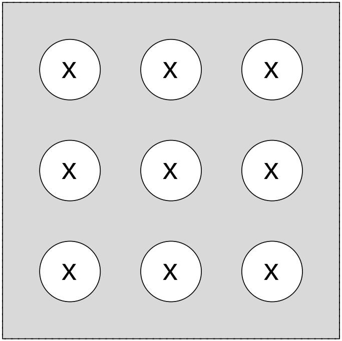
Review: Fixed blocks: Forney-68



- Decision regions catch the typical sets only

- Refuse to decide when ambiguous

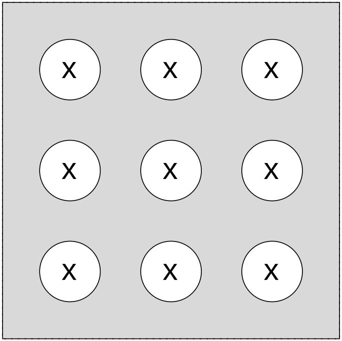
Review: Fixed blocks: Forney-68



- Decision regions catch the typical sets only
- 1 bit feedback can request retransmissions
- Can interpret as expected block-length

- Refuse to decide when ambiguous

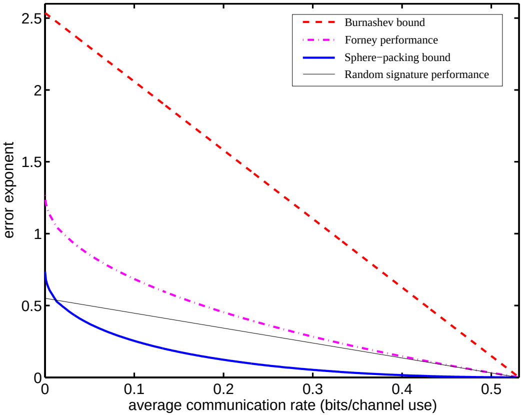
Review: Fixed blocks: Forney-68



- Refuse to decide when ambiguous

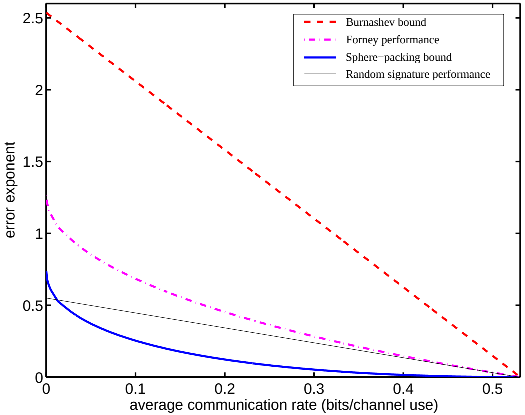
- Decision regions catch the typical sets only
- 1 bit feedback can request retransmissions
- Can interpret as expected block-length
- No converse except at zero rate

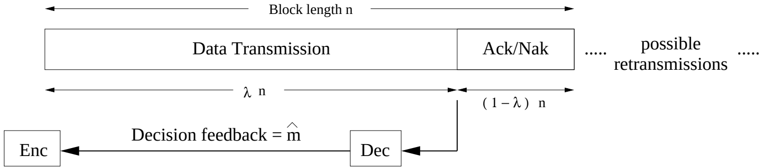
Review: Fixed blocks, Soft deadlines: Burnashev-76



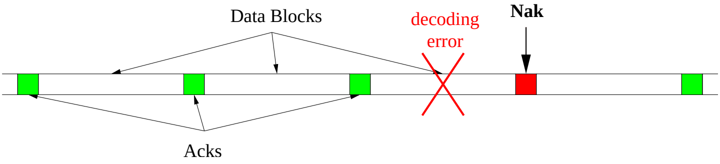
- Showed $C_1(1 - \frac{R}{C})$ was a bound where $C_1 = \max_{i,j} D(p_i||p_j)$
- Considered expected stopping time and used Martingale arguments.

Review: Fixed blocks, Soft deadlines: Burnashev-76



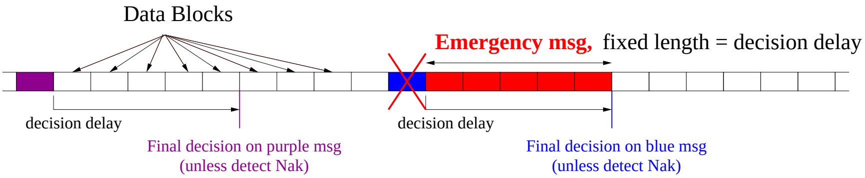


Streaming: an opportunity presents itself



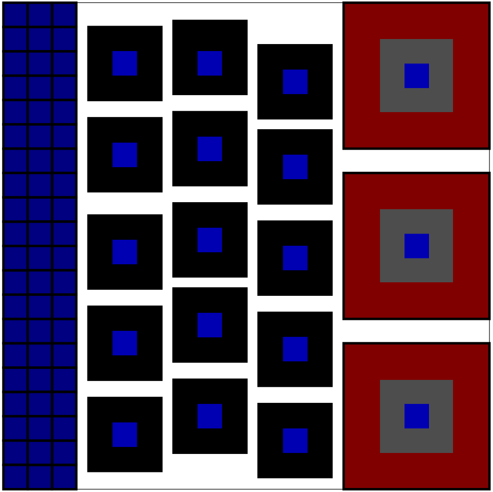
- What if we only sent NAKs when needed?

Sliding blocks with collective punishment only (Kudryashov-79)



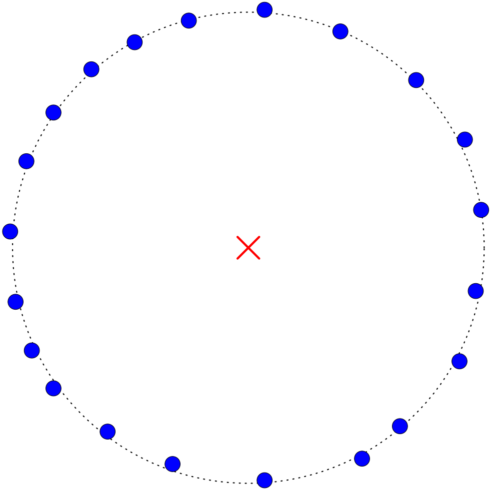
- Make packet length n much smaller than soft deadline Δ.
- A NAK collectively denies the past $\frac{\Delta}{n} - 1$ packets
- Error only if $\frac{\Delta}{n} - 1$ NAKs are all missed

Reason for hope: Csiszar's result '80

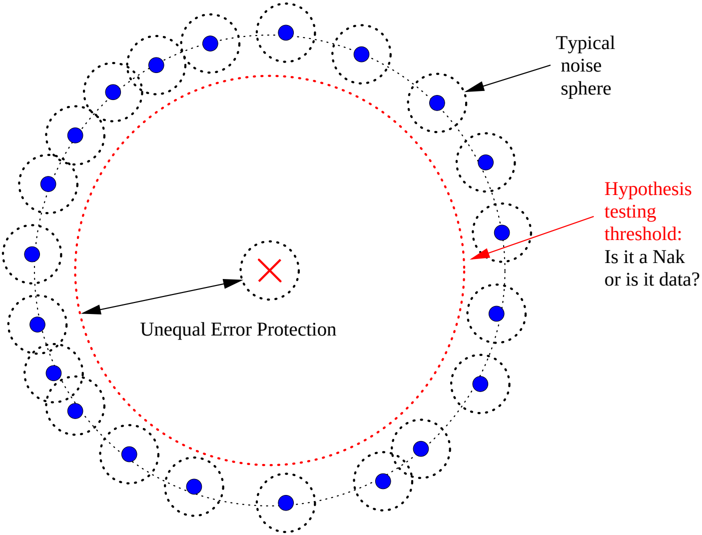


Can pack-in control messages at lower rates and give each subset their own random-coding bound!

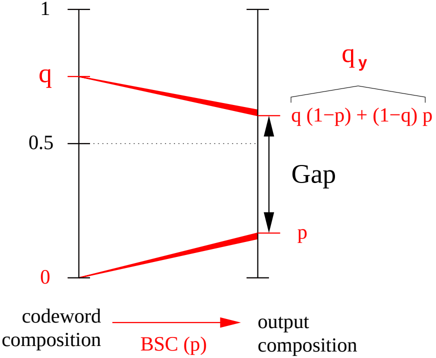
Even simpler: need only one special message



Even simpler: need only one special message

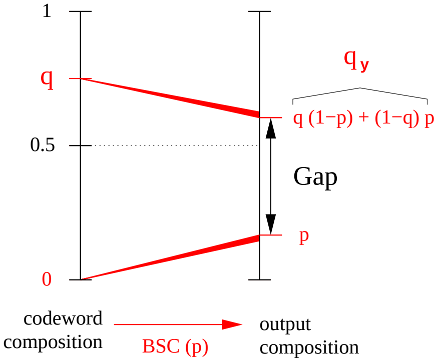


Specialize to BSC case



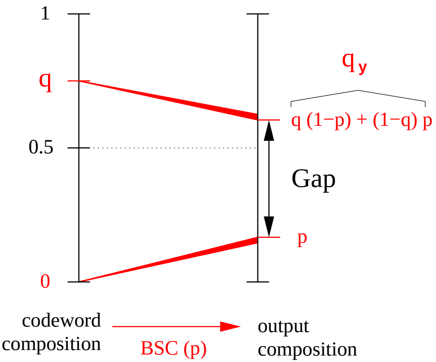
- Use all zero for NAK

Specialize to BSC case



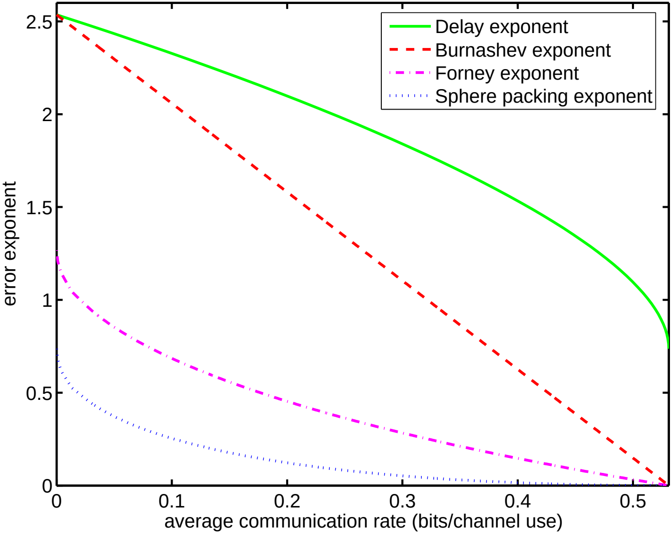
- Use all zero for NAK
- Use composition q code for data:
  $R < H(q_y) - H(p)$

Specialize to BSC case



- Use all zero for NAK
- Use composition q code for data:
  $R < H(q_y) - H(p)$
- Probability of missed NAK is $2^{-nD(q_y||p)}$

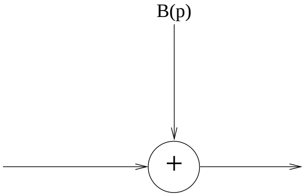
Resulting exponents



Outline

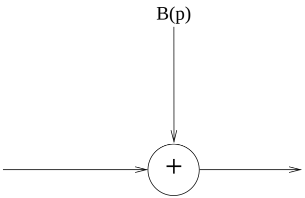
1. Motivation and review
2. Block-coding converse: minimum error probability
3. Streaming converse: bit error probability

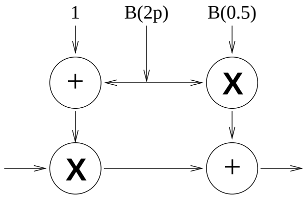
Towards the "Hallucination Bound"



- No converse for Forney, unlike Burnashev and Sphere-packing.
- What about for the minimum probability of error with feedback?

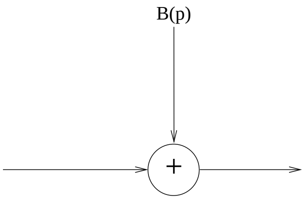
Towards the "Hallucination Bound"

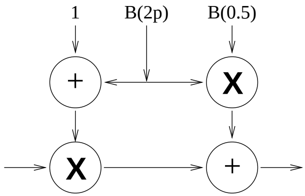




- No converse for Forney, unlike Burnashev and Sphere-packing.
- What about for the minimum probability of error with feedback?

Towards the "Hallucination Bound"

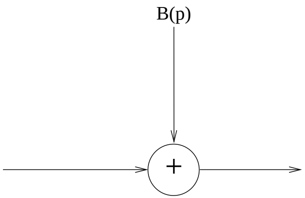


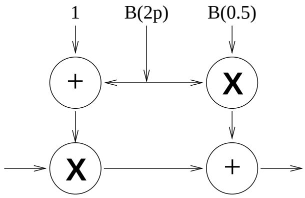


- No converse for Forney, unlike Burnashev and Sphere-packing.
- What about for the minimum probability of error with feedback?

- Trivial bound: let channel disconnect $(2p)^n$

Towards the "Hallucination Bound"

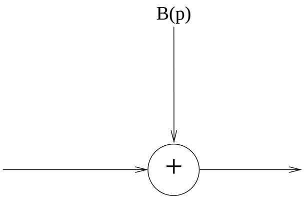


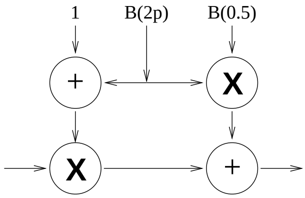


- No converse for Forney, unlike Burnashev and Sphere-packing.
- What about for the minimum probability of error with feedback?

- Trivial bound: let channel disconnect $(2p)^n$
- What is the chance we land on something $\frac{1}{2^n}$

Towards the “Hallucination Bound”

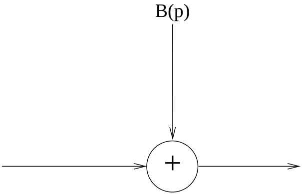


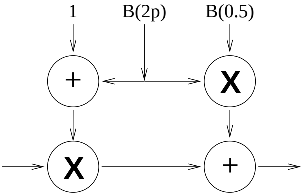


- No converse for Forney, unlike Burnashev and Sphere-packing.
- What about for the minimum probability of error with feedback?

- Trivial bound: let channel disconnect $(2p)^n$
- What is the chance we land on something $\frac{1}{2^n}$
- How many decode to normal codewords? $> 2^{nR}$

Towards the "Hallucination Bound"

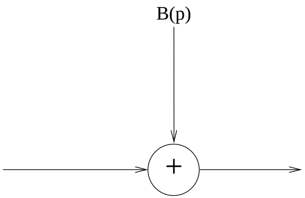


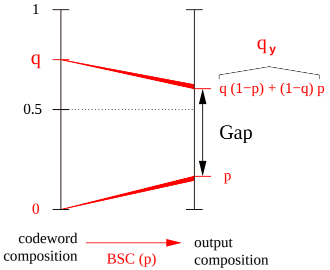


- No converse for Forney, unlike Burnashev and Sphere-packing.
- What about for the minimum probability of error with feedback?

- Trivial bound: let channel disconnect $(2p)^n$
- What is the chance we land on something $\frac{1}{2^n}$
- How many decode to normal codewords? $> 2^{nR}$
- Exponent at most: $\log_2 \frac{1}{p} - R$

Towards the "Hallucination Bound"

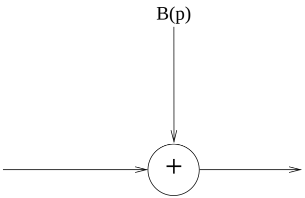


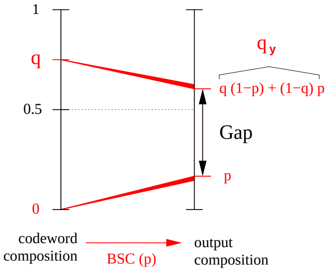


- No converse for Forney, unlike Burnashev and Sphere-packing.
- What about for the minimum probability of error with feedback?

- Each normal message needs $2^{nH(p)}$ to decode to it.
- Must claim $2^{n(R+H(p))}$ volume.

Towards the "Hallucination Bound"

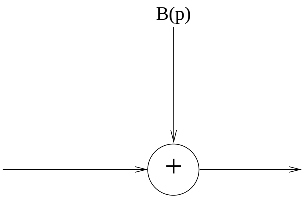


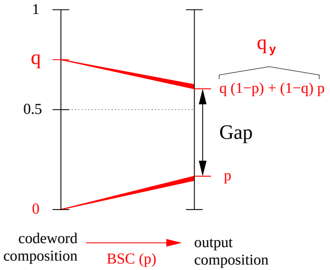


- No converse for Forney, unlike Burnashev and Sphere-packing.
- What about for the minimum probability of error with feedback?

- Each normal message needs $2^{nH(p)}$ to decode to it.
- Must claim $2^{n(R+H(p))}$ volume.
- Place special-message as far away as possible.

Towards the “Hallucination Bound”





- No converse for Forney, unlike Burnashev and Sphere-packing.
- What about for the minimum probability of error with feedback?

- Each normal message needs $2^{nH(p)}$ to decode to it.
- Must claim $2^{n(R+H(p))}$ volume.
- Place special-message as far away as possible.
- Matches achievability!

Generalizes to general symmetric channels

- Consider fixed block-length and moderate probability of correct decoding for “many” regular codewords.
- Berger’s source-coding game reveals that typical channel outputs can only reach $\mathcal{P}_y$ output types.

Generalizes to general symmetric channels

- Consider fixed block-length and moderate probability of correct decoding for "many" regular codewords.
- Berger's source-coding game reveals that typical channel outputs can only reach $\mathcal{P}_y$ output types.
- 2^{nR} regular messages could have their normal (non-erased) decoding regions packed into any of those types.

Generalizes to general symmetric channels

- Consider fixed block-length and moderate probability of correct decoding for “many” regular codewords.
- Berger’s source-coding game reveals that typical channel outputs can only reach $\mathcal{P}_y$ output types.
- 2^{nR} regular messages could have their normal (non-erased) decoding regions packed into any of those types.
- Typical footprint at least $\min_{p_x} 2^{nH(Y|X)}$ inside.

Generalizes to general symmetric channels

- Consider fixed block-length and moderate probability of correct decoding for "many" regular codewords.
- Berger's source-coding game reveals that typical channel outputs can only reach $\mathcal{P}_y$ output types.
- 2^{nR} regular messages could have their normal (non-erased) decoding regions packed into any of those types.
- Typical footprint at least $\min_{p_x} 2^{nH(Y|X)}$ inside.
- Lower bound: pick most distant $\widetilde{p}_y \in \mathcal{P}_y$.

$$\approx 2^{-n(D(p_y||\widetilde{p}_y)+H(p_y))} 2^{nR+nH(Y|X)}$$

Generalizes to general symmetric channels

- Consider fixed block-length and moderate probability of correct decoding for "many" regular codewords.
- Berger's source-coding game reveals that typical channel outputs can only reach $\mathcal{P}_y$ output types.
- 2^{nR} regular messages could have their normal (non-erased) decoding regions packed into any of those types.
- Typical footprint at least $\min_{p_x} 2^{nH(Y|X)}$ inside.
- Lower bound: pick most distant $\widetilde{p}_y \in \mathcal{P}_y$.

$$\approx 2^{-n(D(p_y||\widetilde{p}_y)+H(p_y))} 2^{nR+nH(Y|X)}$$

- Allow convex hull of

$$E_{hal}(R) = \max_{I(p_x,P)\geq R} \max_{\widetilde{p}_y \in \mathcal{P}_y} [D(P(p_x)||\widetilde{p}_y) + (I(p_x,P) - R)]$$

Generalizes to general symmetric channels

- Consider fixed block-length and moderate probability of correct decoding for "many" regular codewords.
- Berger's source-coding game reveals that typical channel outputs can only reach $\mathcal{P}_y$ output types.
- 2^{nR} regular messages could have their normal (non-erased) decoding regions packed into any of those types.
- Typical footprint at least $\min_{p_x} 2^{nH(Y|X)}$ inside.
- Lower bound: pick most distant $\widetilde{p}_y \in \mathcal{P}_y$.

$$\approx 2^{-n(D(p_y||\widetilde{p}_y)+H(p_y))} 2^{nR+nH(Y|X)}$$

- Allow convex hull of

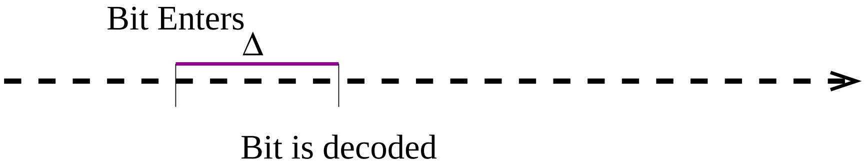
$$E_{hal}(R) = \max_{I(p_x,P)\geq R} \max_{\widetilde{p}_y \in \mathcal{P}_y} [D(P(p_x)||\widetilde{p}_y) + (I(p_x,P) - R)]$$

- On Friday Borade, *et al* will show a more general proof that holds with expected block-length.

Outline

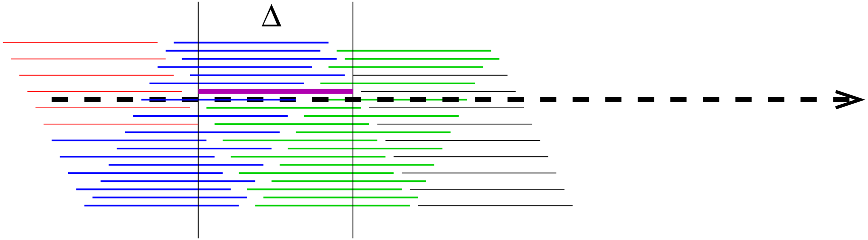
1. Motivation and review
2. Block-coding converse: minimum error probability
3. Streaming converse: bit error probability

The basic intuition



- If we hallucinate for Δ, we will miss our opportunity to decode.

The basic intuition



- If we hallucinate for Δ, we will miss our opportunity to decode.
- Challenge: overlaps with other bit-footprints.

The idea of the proof

- View a block of many $n\Delta$.

The idea of the proof

- View a block of many $n\Delta$.
- Count the volume of normal decoding regions in two different ways.

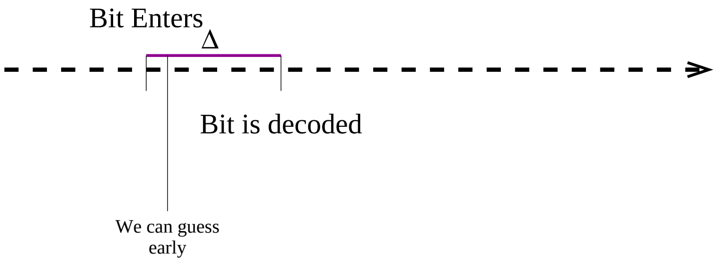
The idea of the proof

- View a block of many $n\Delta$.
- Count the volume of normal decoding regions in two different ways.
  - As a big block code based on correct decoding.

The idea of the proof

- View a block of many $n\Delta$.
- Count the volume of normal decoding regions in two different ways.
  - As a big block code based on correct decoding.
  - With a tree-structure based on the desired exponent.

The idea of the proof



- View a block of many $n\Delta$.
- Count the volume of normal decoding regions in two different ways.
  - As a big block code based on correct decoding.
  - With a tree-structure based on the desired exponent.
- Technical condition: impose sequentiality on the code: we can usually guess the answer even before the deadline runs out.

Conclusions

- The “Hallucination Bound” is the probability that that the decoder imagines that everything is normal despite your best efforts to tell it otherwise.
- This corresponds to the best probability of error for a special message in the fixed block-code setting.

Conclusions

- The “Hallucination Bound” is the probability that that the decoder imagines that everything is normal despite your best efforts to tell it otherwise.
- This corresponds to the best probability of error for a special message in the fixed block-code setting.
- Future work

Conclusions

- The “Hallucination Bound” is the probability that that the decoder imagines that everything is normal despite your best efforts to tell it otherwise.
- This corresponds to the best probability of error for a special message in the fixed block-code setting.
- Future work
  - Eliminate the technical condition for the streaming case.

Conclusions

- The "Hallucination Bound" is the probability that that the decoder imagines that everything is normal despite your best efforts to tell it otherwise.
- This corresponds to the best probability of error for a special message in the fixed block-code setting.
- Future work
  - Eliminate the technical condition for the streaming case.
  - Get a two-way "Hallucination bound" for the case of noisy feedback.