Problems We Can Solve With a Helper

Haim H. Permuter
Stanford
haim1@stanford.edu

Yossef Steinberg
Technion
ysteinbe@ee.technion.ac.il

Tsachy Weissman
Stanford/Technion
weissman@ee.technion.ac.il

Outline

Outline

- Problem Formulation

Outline

- Problem Formulation
- Rate-Distortion with common rate-limited helper

Outline

- Problem Formulation
- Rate-Distortion with common rate-limited helper
    - Previous work

Outline

- Problem Formulation
- Rate-Distortion with common rate-limited helper
  - Previous work
  - The rate-distortion region, and important observations

Outline

- Problem Formulation
- Rate-Distortion with common rate-limited helper
    - Previous work
    - The rate-distortion region, and important observations
- Independent helper rates

Outline

- Problem Formulation
- Rate-Distortion with common rate-limited helper
  - Previous work
  - The rate-distortion region, and important observations
- Independent helper rates
- Helper with side information at the decoder

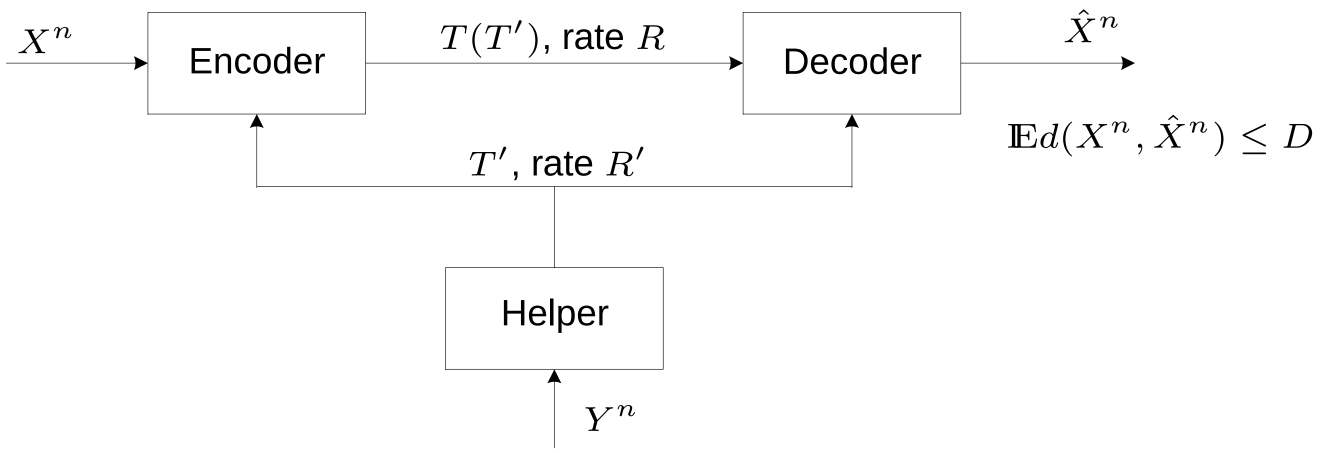
Outline

- Problem Formulation
- Rate-Distortion with common rate-limited helper
  - Previous work
  - The rate-distortion region, and important observations
- Independent helper rates
- Helper with side information at the decoder
- Cascade rate-distortion with a helper

Outline

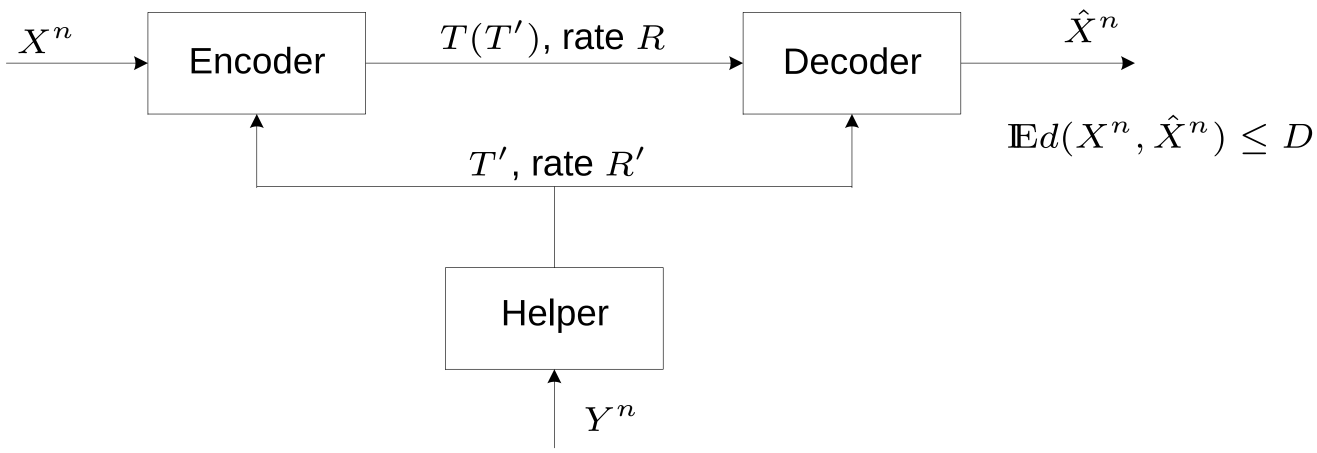
- Problem Formulation
- Rate-Distortion with common rate-limited helper
  - Previous work
  - The rate-distortion region, and important observations
- Independent helper rates
- Helper with side information at the decoder
- Cascade rate-distortion with a helper
- Examples

Problem Formulation



Goal: Compression of X^n, with rate-limited SI at both ends:

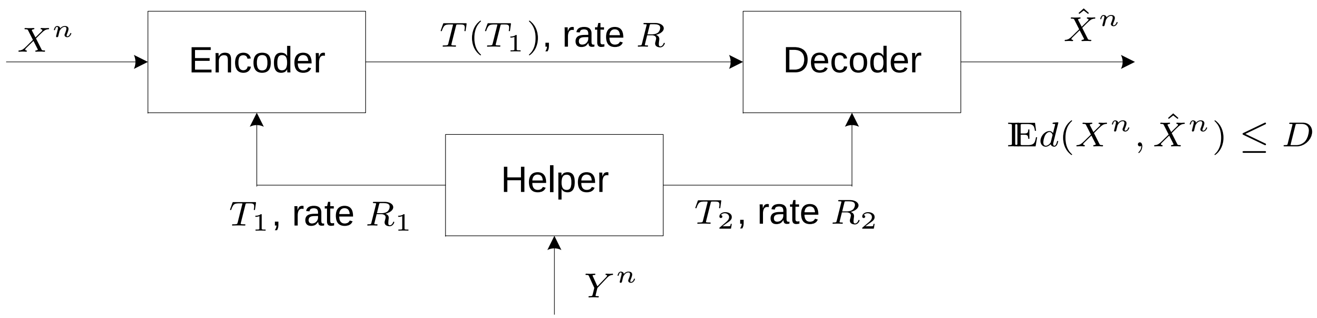
Problem Formulation



Goal: Compression of X^n, with rate-limited SI at both ends:

A. Common

Problem Formulation

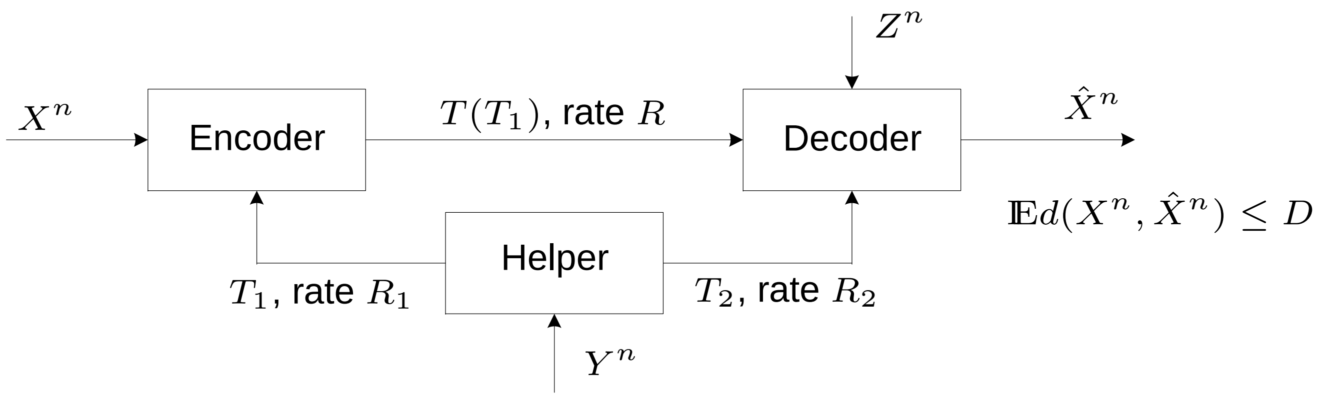


Goal: Compression of X^n, with rate-limited SI at both ends:

A. Common

B. Independent

Problem Formulation



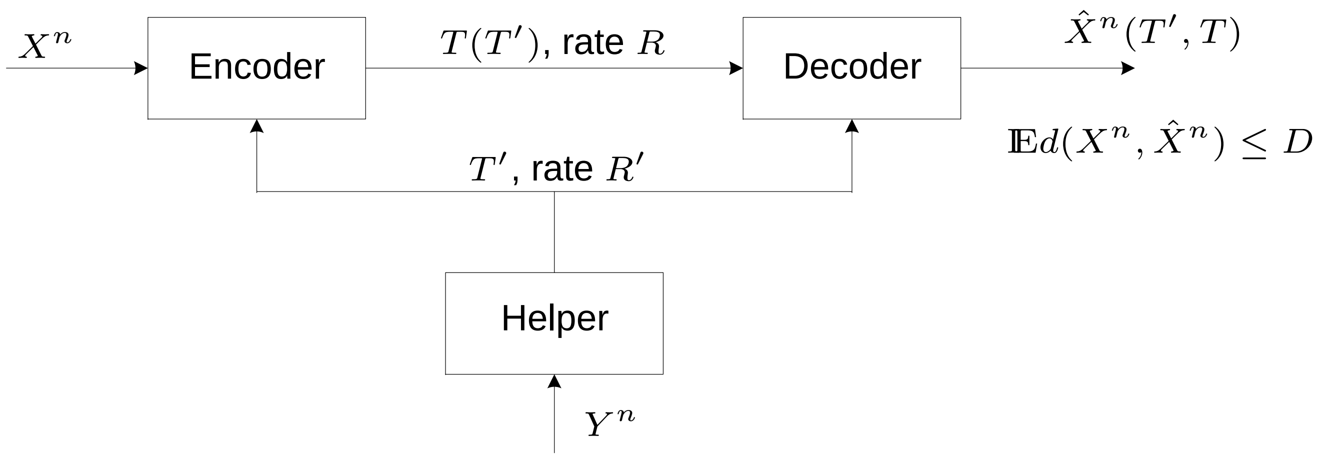
Goal: Compression of X^n, with rate-limited SI at both ends:

A. Common

B. Independent

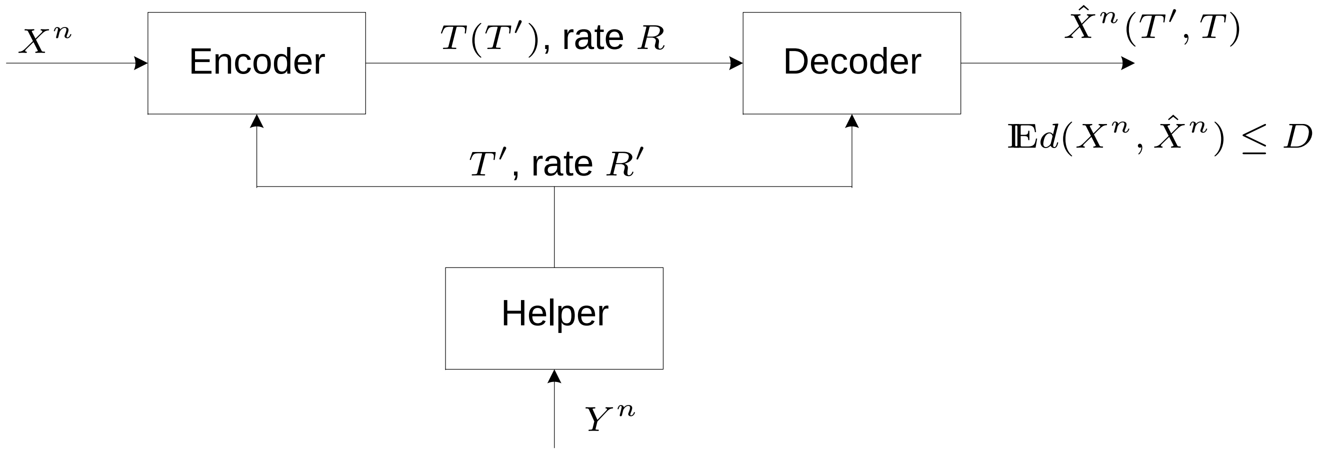
C. With decoder SI

Problem Formulation



A. Compression of X^n, with common rate-limited SI at both ends.

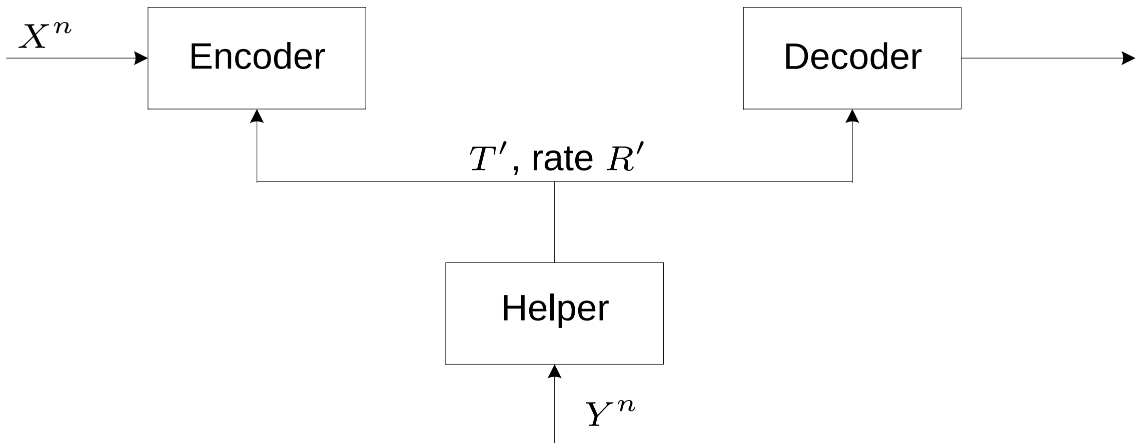
Problem Formulation



A. Compression of X^n, with common rate-limited SI at both ends.

The communication protocol:

Problem Formulation

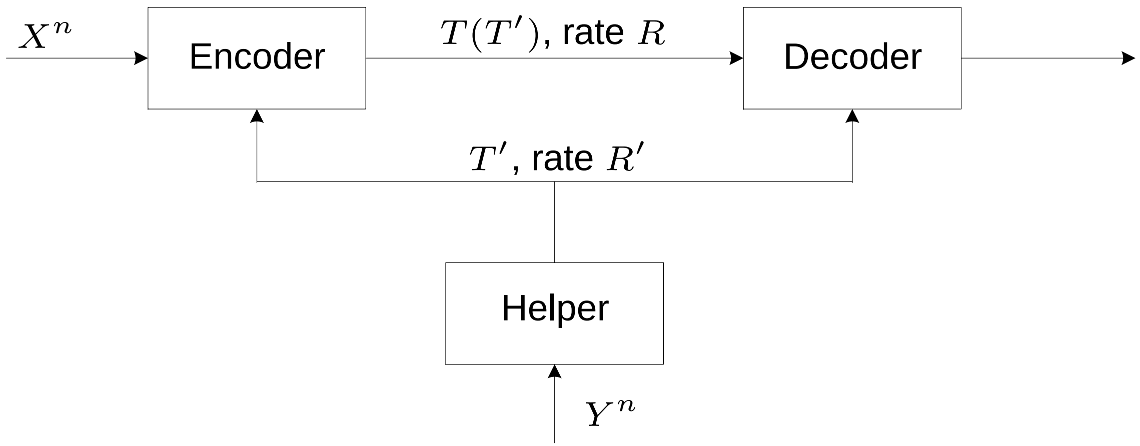


A. Compression of X^n, with common rate-limited SI at both ends.

The communication protocol:

1. A helper compresses side information Y^n to rate R', and sends it to the encoder and decoder (same message T' to both).

Problem Formulation

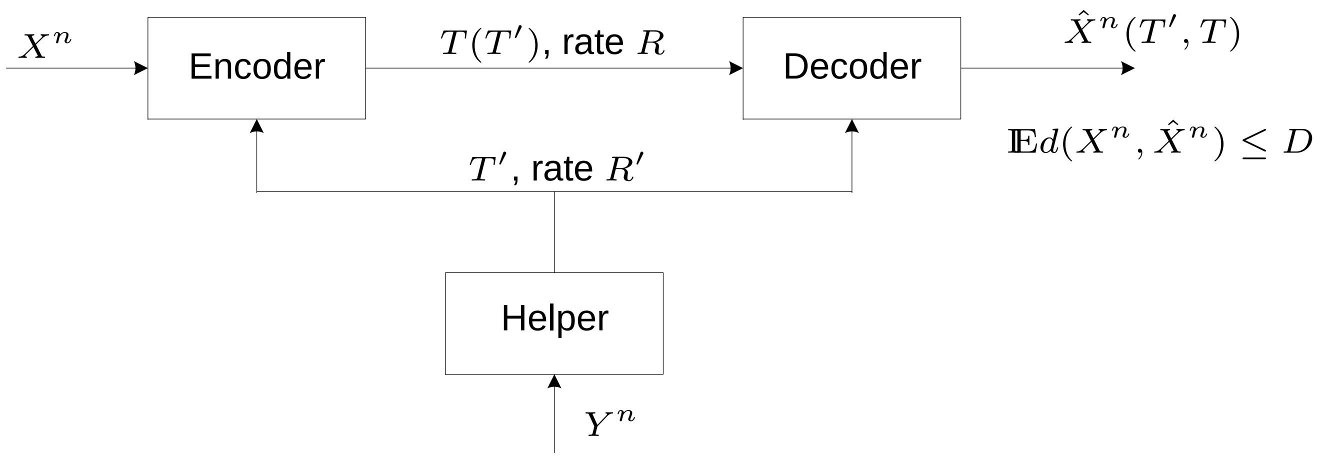


A. Compression of X^n, with common rate-limited SI at both ends.

The communication protocol:

1. A helper compresses side information Y^n to rate R', and sends it to the encoder and decoder (same message T' to both).
2. Given T', the source encoder sends a message T of rate R to the destination.

Problem Formulation

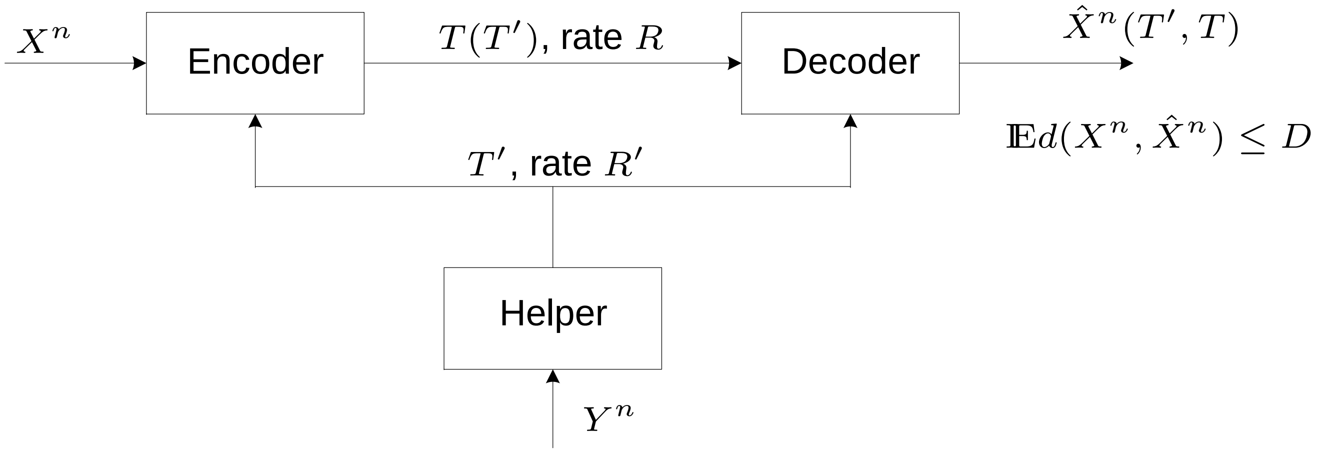


A. Compression of X^n, with common rate-limited SI at both ends.

The communication protocol:

1. A helper compresses side information Y^n to rate R', and sends it to the encoder and decoder (same message T' to both).
2. Given T', the source encoder sends a message T of rate R to the destination.
3. The destination constructs an estimate $\hat{X}^n$, based on T and T'.

Problem Formulation *(cont'd)*



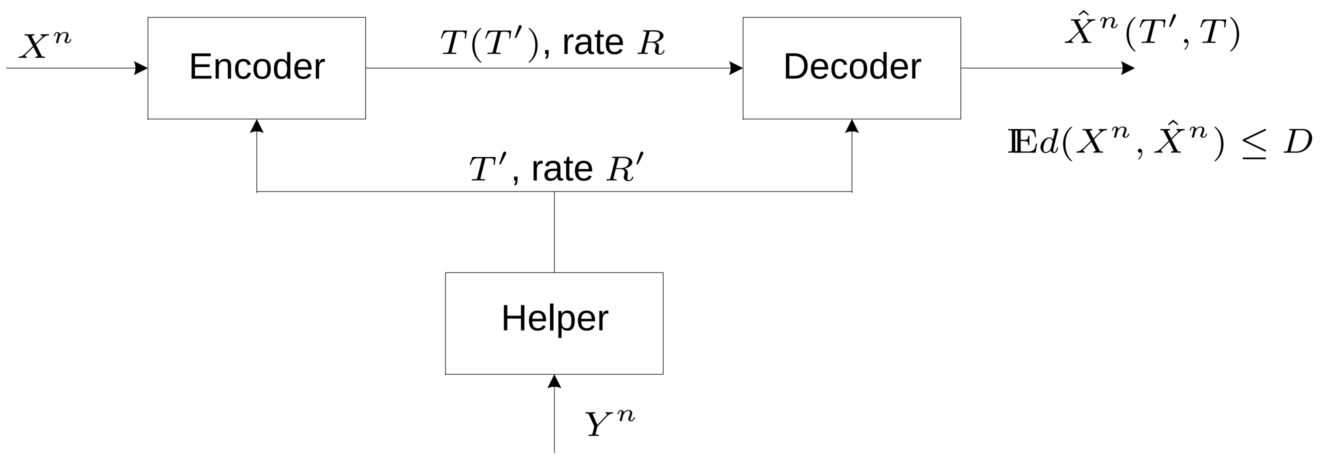
Model:

$(X, Y) \sim P_{X,Y}$, memoryless.

We are interested in the region of all achievable rate pairs and distortion

$$(R, R', D).$$

Problem Formulation *(cont'd)*



Model:

$(X, Y) \sim P_{X,Y}$, memoryless.

We are interested in the region of all achievable rate pairs and distortion

$$(R, R', D).$$

$\mathcal{R}(D)$ – the collection of all rate pairs (R, R') such that the triplet (R, R', D) is achievable.

Previous Work

- Lossless source coding with rate-limited SI at the decoder(s)

  Wyner 75.

  Ahlswede and Körner, 75

- Lossy source coding with SI at the decoder

  Wyner and Ziv, 76.

- Rate-distortion for correlated sources with partially separated encoders

  Kaspi, Ph.D. dissertation, 79.

  Kaspi and Berger, IEEE IT 1982.

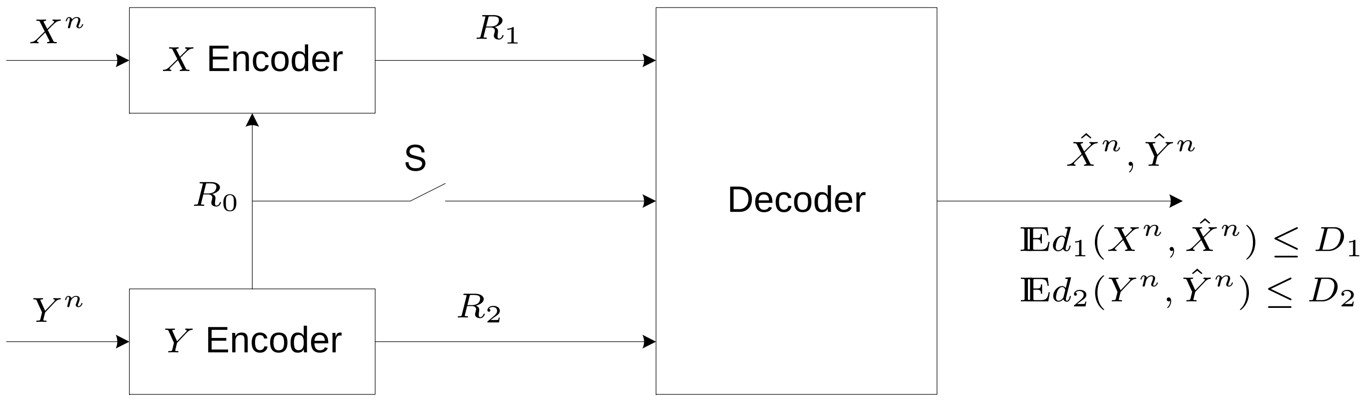
  Subsumes all previous models.

Previous Work *(cont'd)*

- Cooperative source coding with encoder breakdown,
  Dinkar Vasudevan and Etienne Perron, ISIT 2007, Nice.

Previous Work *(cont'd)*

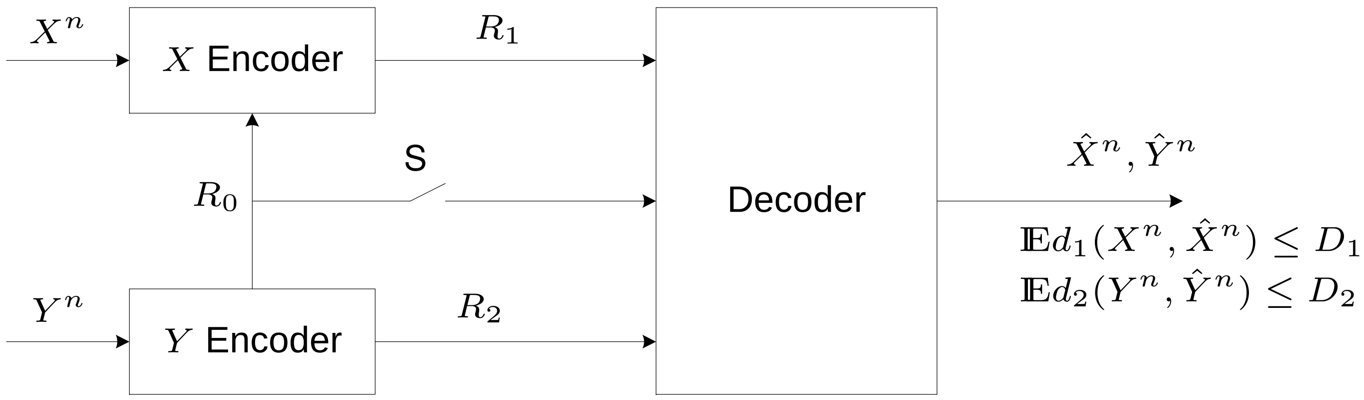
The Kaspi and Berger model:



Derived achievability results for switch closed/open.

Previous Work *(cont'd)*

The Kaspi and Berger model:

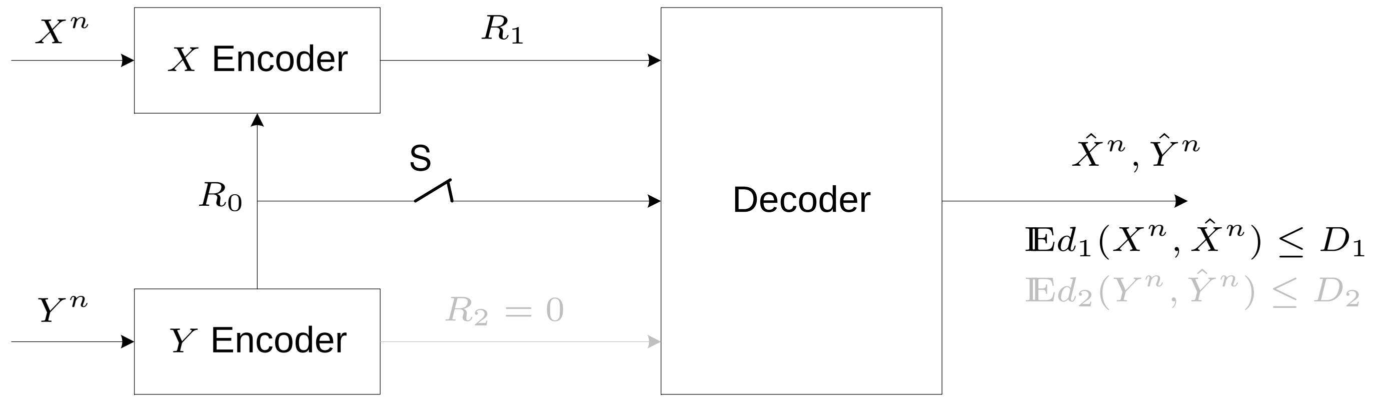


Derived achievability results for switch closed/open.

When S is closed, $R_2 = 0$, and D_2 free, this model reduces to our helper problem.

Previous Work *(cont'd)*

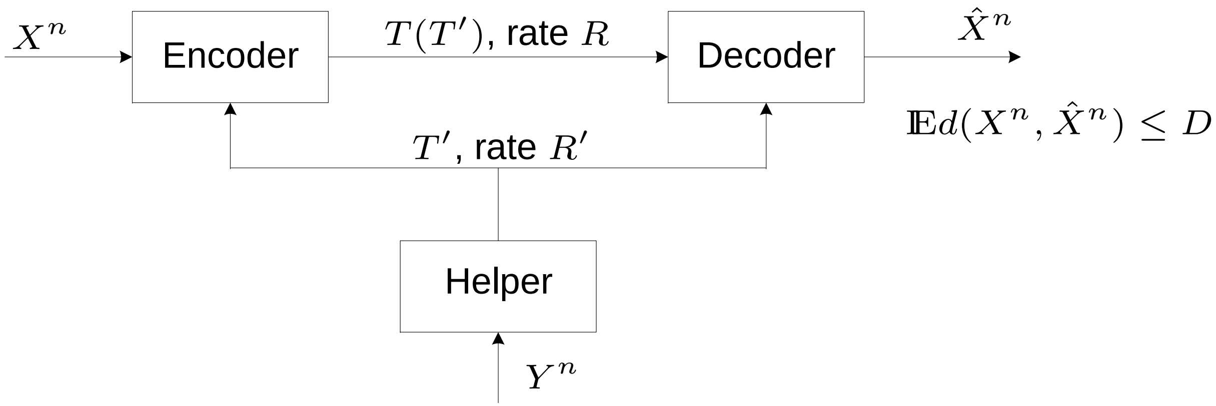
The Kaspi and Berger model:



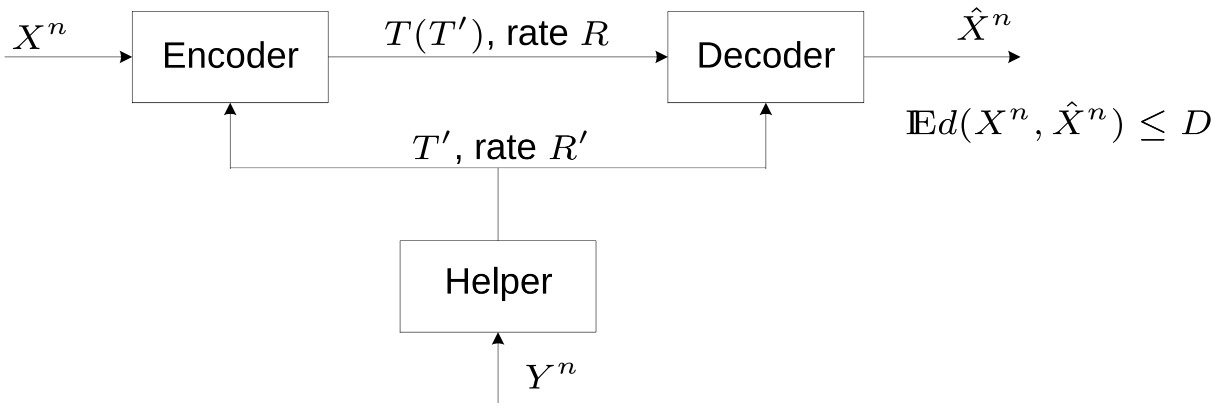
Derived achievability results for switch closed/open.

When S is closed, $R_2 = 0$, and D_2 free, this model reduces to our helper problem.

The rate-distortion region with common messages



The rate-distortion region with common messages



Theorem 1 [Vasudevan & Perron, 2007, PSW 2008]
$(R, R') \in \mathcal{R}(D)$ if and only if

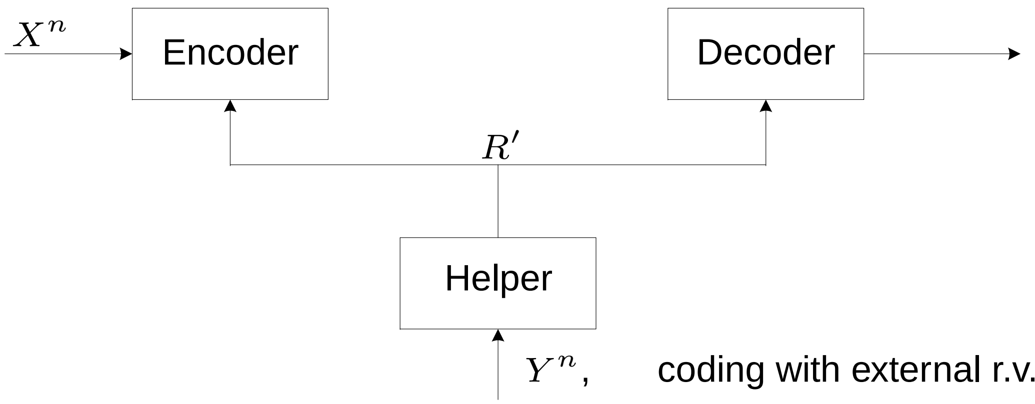
$$\begin{aligned} R' &\geq I(U;Y) \\ R &\geq I(X;\hat{X}|U) \\ D &\geq \mathbb{E}d(X,\hat{X}) \end{aligned}$$

for some joint distribution of the form

$$P_{X,Y} \cdot P_{U|Y} \cdot P_{\hat{X}|U,X}.$$

The R-D region with common messages *(cont'd)*

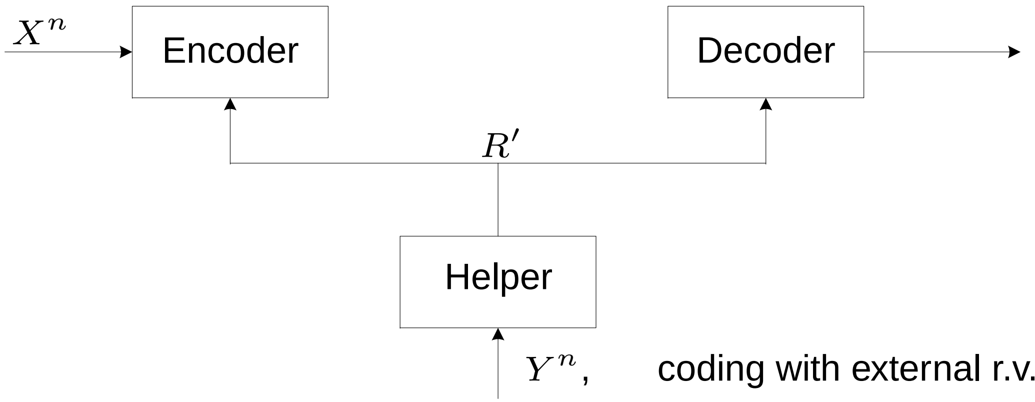
Direct Part



coding with external r.v. U

The R-D region with common messages (cont'd)

Direct Part

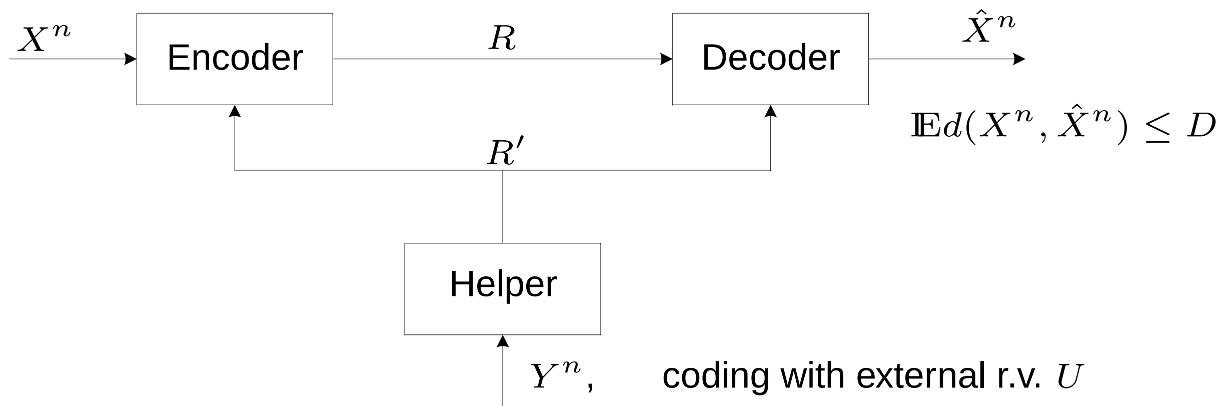


coding with external r.v. U

$$R' \geq I(U;Y) \qquad U - Y - (X, \hat{X})$$

The R-D region with common messages (cont'd)

Direct Part



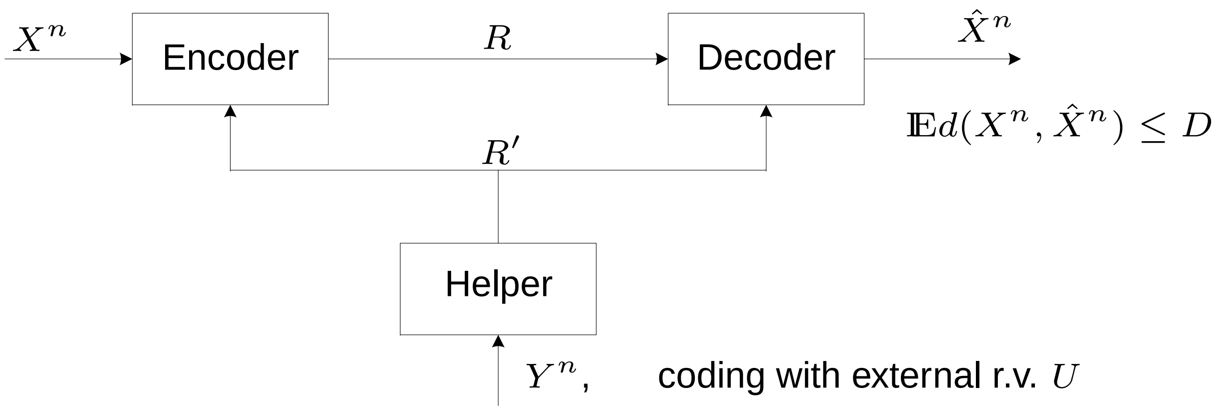
coding with external r.v. U

$$R' \geq I(U;Y) \qquad U - Y - (X, \hat{X})$$

Now both sides can condition on U

The R-D region with common messages (cont'd)

Direct Part



coding with external r.v. U

$$R' \geq I(U;Y) \qquad U - Y - (X, \hat{X})$$

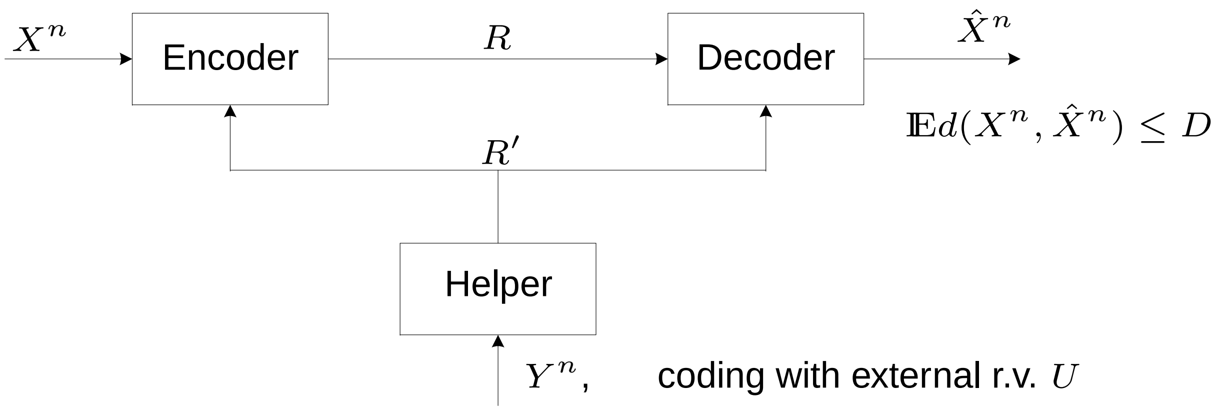
Now both sides can condition on U

$$R \geq I(X;\hat{X}|U) \qquad \hat{X} - (U,X) - Y$$

$$D \geq \mathbb{E}d(X,\hat{X})$$

The R-D region with common messages (cont'd)

Direct Part



coding with external r.v. U

$$R' \geq I(U;Y) \qquad U - Y - (X, \hat{X})$$

Now both sides can condition on U

$$R \geq I(X;\hat{X}|U) \qquad \hat{X} - (U,X) - Y$$
$$D \geq \mathbb{E}d(X,\hat{X})$$

$$P_{X,Y} \cdot P_{U|Y} \cdot P_{\hat{X}|U,X}.$$

The R-D region with common messages (cont'd)

Converse

In most converse proofs, the main issue is to get the Markov conditions.

Do we really need them?

The R-D region with common messages *(cont'd)*

Converse

In most converse proofs, the main issue is to get the Markov conditions.

Do we really need them?

$$
\begin{aligned}
R' &\geq I(U;Y) & & U - Y - (X,\hat{X}) \\
R &\geq I(X;\hat{X}|U) & & \hat{X} - (U,X) - Y \\
D &\geq \mathbb{E}d(X,\hat{X}) & & (*)
\end{aligned}
$$

The R-D region with common messages (cont'd)

Converse

In most converse proofs, the main issue is to get the Markov conditions.

Do we really need them?

$$\begin{aligned} R' &\geq I(U;Y) && U - Y - (X,\hat{X}) \\ R &\geq I(X;\hat{X}|U) && \hat{X} - (U,X) - Y \\ D &\geq \mathbb{E}d(X,\hat{X}) && (*) \end{aligned}$$

- $\hat{X}^n(j),\ j = 1, 2 \ldots$ are generated conditioned on U^n.

  U^n is known at both sides.

The R-D region with common messages (cont'd)

Converse

In most converse proofs, the main issue is to get the Markov conditions.

Do we really need them?

$$
\begin{aligned}
R' &\geq I(U;Y) && U - Y - (X,\hat{X}) \\
R &\geq I(X;\hat{X}|U) && \hat{X} - (U,X) - Y \\
D &\geq \mathbb{E}d(X,\hat{X}) && (*)
\end{aligned}
$$

- $\hat{X}^n(j),\ j = 1,2\ldots$ are generated conditioned on U^n.

  U^n is known at both sides.

- The specific codeword $\hat{X}^n(i)$ is chosen typical with X^n.

The R-D region with common messages (cont'd)

Converse

In most converse proofs, the main issue is to get the Markov conditions.

Do we really need them?

$$
\begin{aligned}
R' &\geq I(U;Y) && U - Y - (X,\hat{X}) \\
R &\geq I(X;\hat{X}|U) && \hat{X} - (U,X) - Y \\
D &\geq \mathbb{E}d(X,\hat{X}) && (*)
\end{aligned}
$$

- $\hat{X}^n(j),\ j = 1, 2 \ldots$ are generated conditioned on U^n.

  U^n is known at both sides.

- The specific codeword $\hat{X}^n(i)$ is chosen typical with X^n.

- If $(\hat{X}^n, U^n, X^n)$ are jointly typical, the distortion $\leq D$, due to $(*)$.

The R-D region with common messages (cont'd)

Converse

In most converse proofs, the main issue is to get the Markov conditions.

Do we really need them?

$$\begin{aligned} R' &\geq I(U;Y) & & U - Y - (X, \hat{X}) \\ R &\geq I(X;\hat{X}|U) & & \hat{X} - (U, X) - Y \\ D &\geq \mathbb{E}d(X,\hat{X}) & & (*) \end{aligned}$$

- $\hat{X}^n(j),\ j = 1, 2 \ldots$ are generated conditioned on U^n.

  U^n is known at both sides.
- The specific codeword $\hat{X}^n(i)$ is chosen typical with X^n.
- If $(\hat{X}^n, U^n, X^n)$ are jointly typical, the distortion $\leq D$, due to $(*)$.

$$\Longrightarrow \qquad P_{\hat{X},U,X,Y} \text{ need not be kept. Only } P_{\hat{X},U,X}.$$

The R-D region with common messages *(cont'd)*

Another way of saying the same thing: Let

$$\tilde{P}_{\hat{X},U,X,Y} = P_{X,Y} \cdot P_{U|Y} \cdot P_{\hat{X}|U,X,Y}$$
$$P_{\hat{X},U,X,Y} = P_{X,Y} \cdot P_{U|Y} \cdot P_{\hat{X}|U,X}$$

with $P_{\hat{X}|U,X}$ the conditional marginal induced by $\tilde{P}_{\hat{X},U,X,Y}$.

The R-D region with common messages (cont'd)

Another way of saying the same thing: Let

$$\tilde{P}_{\hat{X},U,X,Y} = P_{X,Y} \cdot P_{U|Y} \cdot P_{\hat{X}|U,X,Y}$$
$$P_{\hat{X},U,X,Y} = P_{X,Y} \cdot P_{U|Y} \cdot P_{\hat{X}|U,X}$$

with $P_{\hat{X}|U,X}$ the conditional marginal induced by $\tilde{P}_{\hat{X},U,X,Y}$.

The region

$$R' \geq I(U;Y)$$
$$R \geq I(X;\hat{X}|U)$$
$$D \geq \mathbb{E}d(X,\hat{X})$$

is the same under P and $\tilde{P}$.

The R-D region with common messages (cont'd)

Another way of saying the same thing: Let

$$\tilde{P}_{\hat{X},U,X,Y} = P_{X,Y} \cdot P_{U|Y} \cdot P_{\hat{X}|U,X,Y}$$
$$P_{\hat{X},U,X,Y} = P_{X,Y} \cdot P_{U|Y} \cdot P_{\hat{X}|U,X}$$

with $P_{\hat{X}|U,X}$ the conditional marginal induced by $\tilde{P}_{\hat{X},U,X,Y}$.

The region

$$R' \geq I(U;Y)$$
$$R \geq I(X;\hat{X}|U)$$
$$D \geq \mathbb{E}d(X,\hat{X})$$

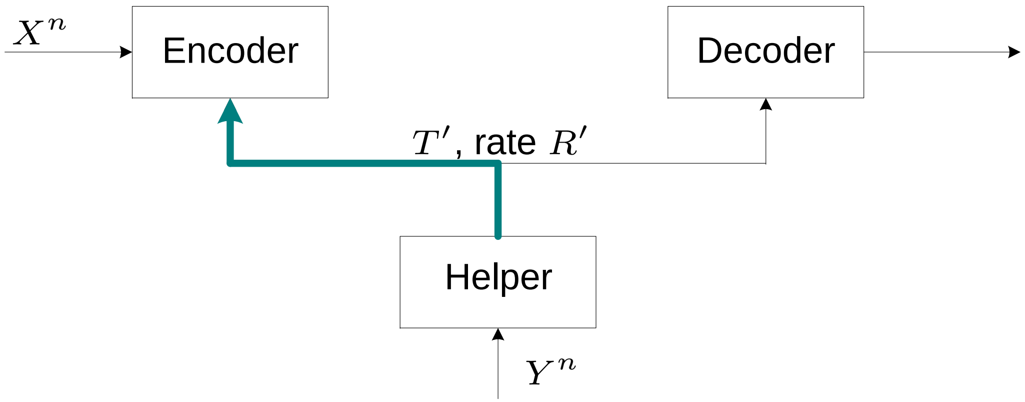
is the same under P and $\tilde{P}$.

Thus, it is enough to prove:

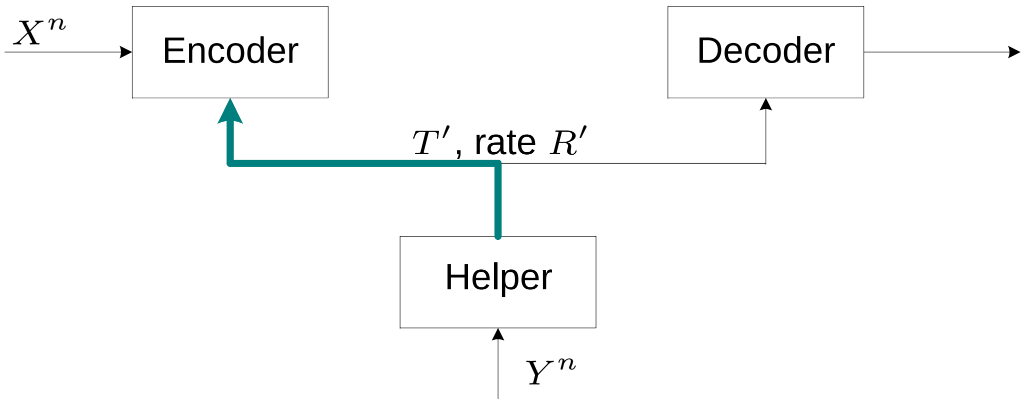
Achievability for any $P_{X,Y} \cdot P_{U|Y} \cdot P_{\hat{X}|U,X}$

Upper bound for $P_{X,Y} \cdot P_{U|Y} \cdot P_{\hat{X}|U,X,Y}$

A Second Look at the Helper

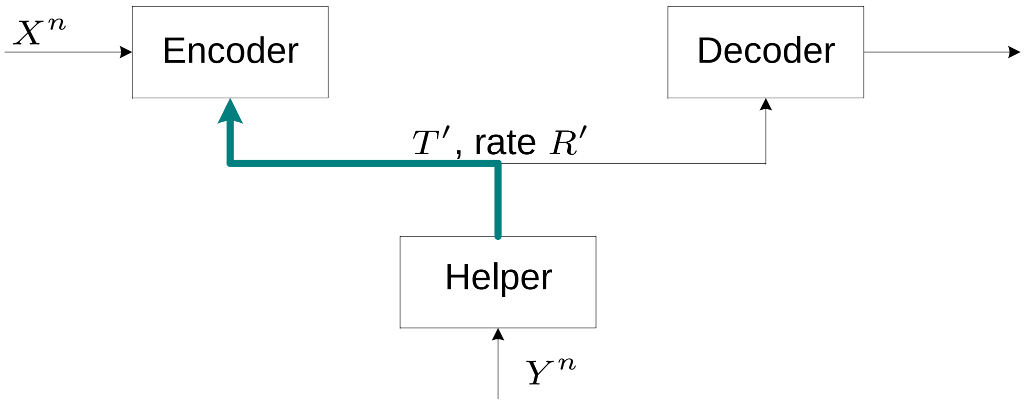


A Second Look at the Helper



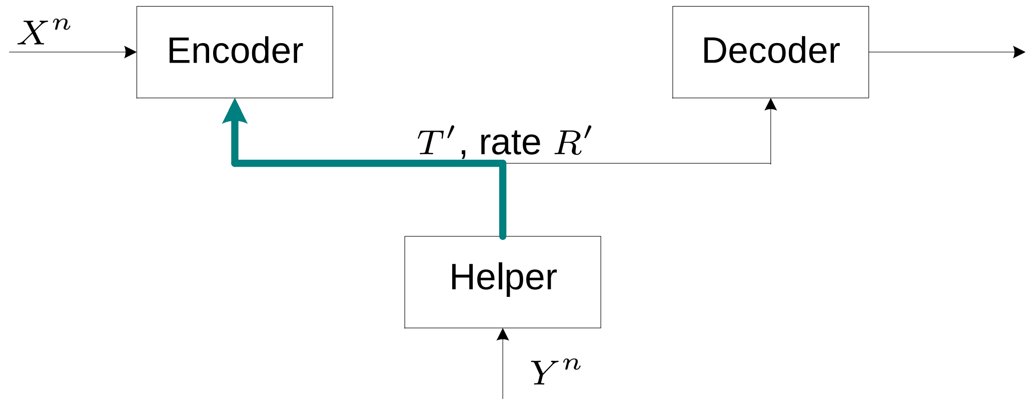
- The same stream T' is sent to the encoder and decoder

A Second Look at the Helper



- The same stream T' is sent to the encoder and decoder
- The encoder has X^n at hand, which can serve as SI, to further improve its knowledge about Y^n

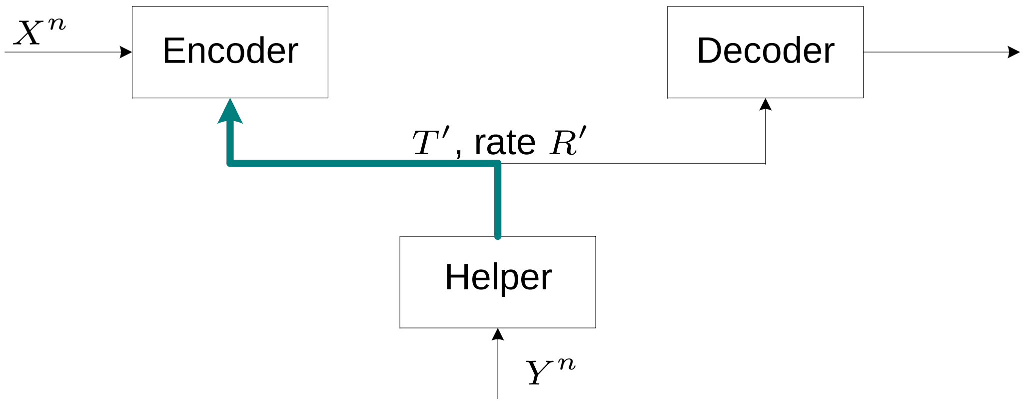
A Second Look at the Helper



- The same stream T' is sent to the encoder and decoder
- The encoder has X^n at hand, which can serve as SI, to further improve its knowledge about Y^n

$\Rightarrow$ “Rate distortion when side information may be absent," Heegard and Berger, *IEEE IT*, Nov. 1985.

A Second Look at the Helper

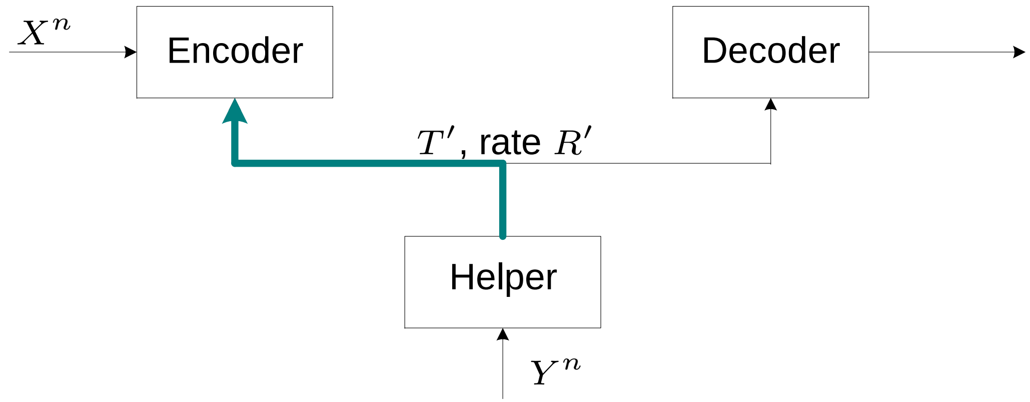


- The same stream T' is sent to the encoder and decoder
- The encoder has X^n at hand, which can serve as SI, to further improve its knowledge about Y^n

$\Rightarrow$ "Rate distortion when side information may be absent," Heegard and Berger, *IEEE IT*, Nov. 1985.

$$R' \geq I(U;Y) + I(V;Y|U,X), \qquad (U,V) - Y - X.$$

A Second Look at the Helper



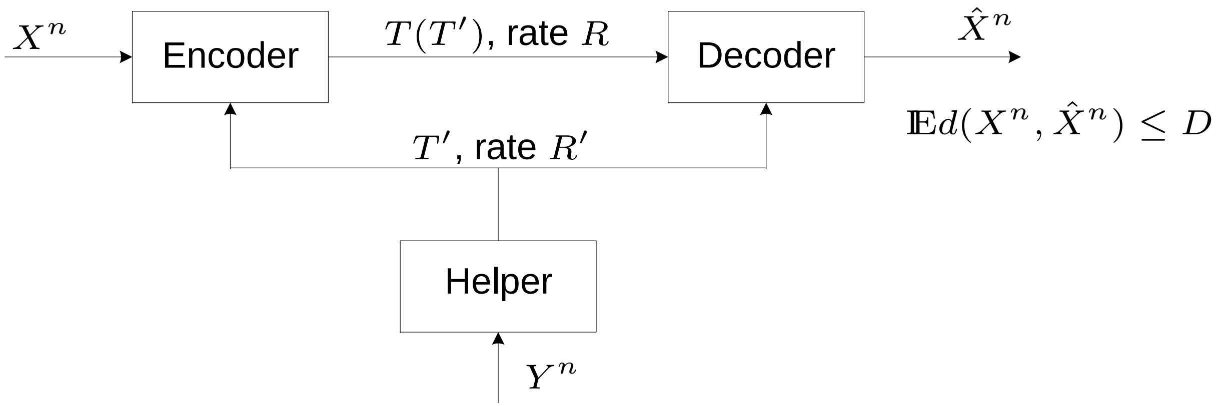
- The same stream T' is sent to the encoder and decoder
- The encoder has X^n at hand, which can serve as SI, to further improve its knowledge about Y^n

$\Rightarrow$ “Rate distortion when side information may be absent," Heegard and Berger, *IEEE IT*, Nov. 1985.

$$R' \geq I(U;Y) + I(V;Y|U,X), \qquad (U,V) - Y - X.$$

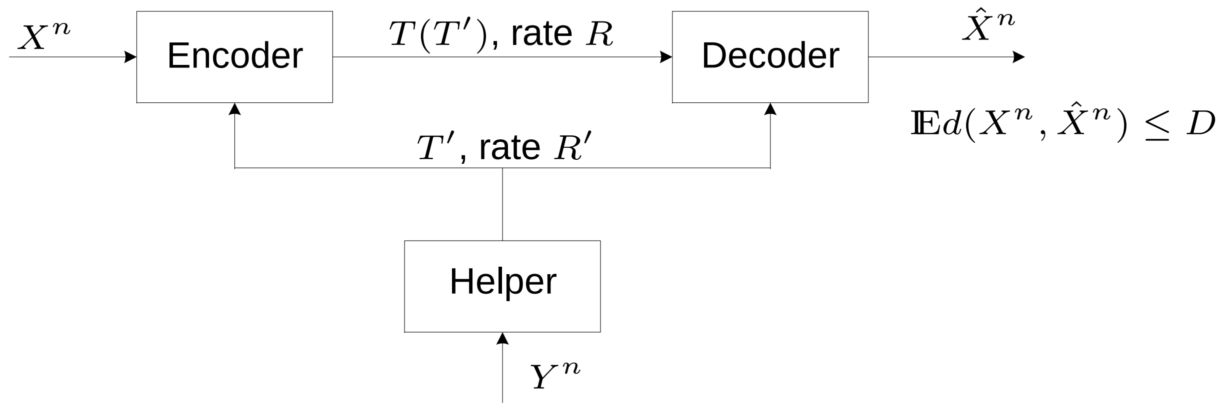
But in our result, we have only $R' \geq I(U;Y)$.

A Second Look at the Helper (cont'd)



$$\begin{aligned} R' &\geq I(U;Y) \\ R &\geq I(X;\hat{X}|U) \\ D &\geq \mathbb{E}d(X,\hat{X}) \end{aligned}$$

A Second Look at the Helper (cont'd)

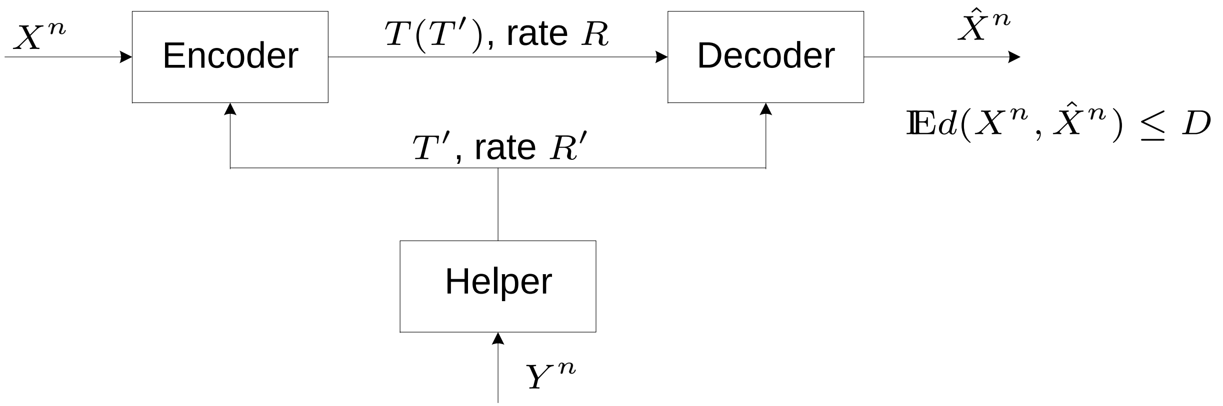


$$
\begin{aligned}
R' &\geq I(U;Y) \\
R &\geq I(X;\hat{X}|U) \\
D &\geq \mathbb{E}d(X,\hat{X})
\end{aligned}
$$

Conclusion:

When sending common message to the encoder and decoder, the source encoder does not use its SI X^n to improve his knowledge on Y^n beyond what is known at the decoder.

A Second Look at the Helper (cont'd)



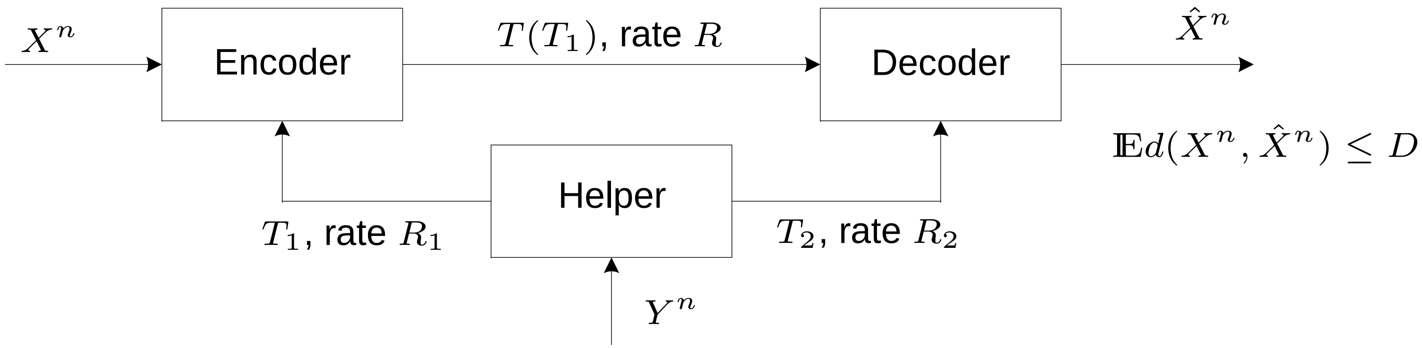
$$\begin{aligned} R' &\geq I(U;Y) \\ R &\geq I(X;\hat{X}|U) \\ D &\geq \mathbb{E}d(X,\hat{X}) \end{aligned}$$

Conclusion:

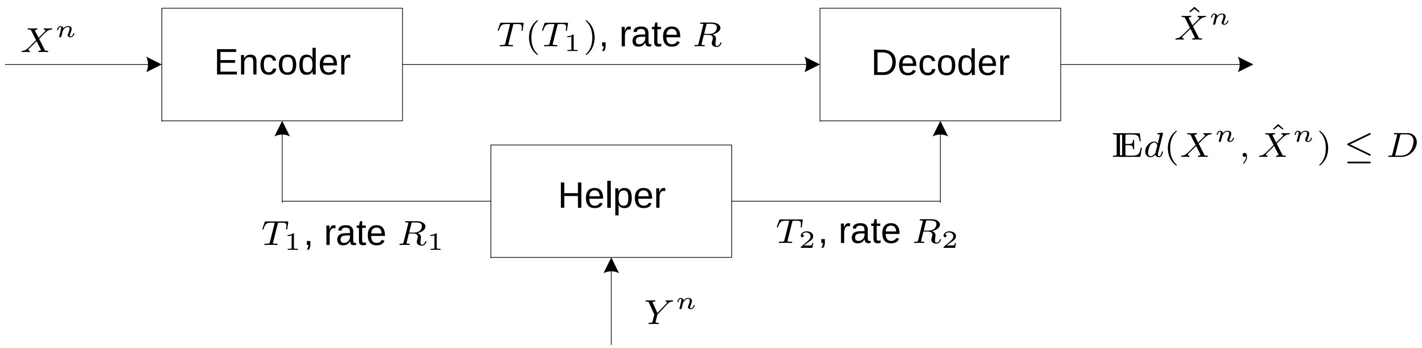
When sending common message to the encoder and decoder, the source encoder does not use its SI X^n to improve his knowledge on Y^n beyond what is known at the decoder.

Will he use extra rate?

Independent rates

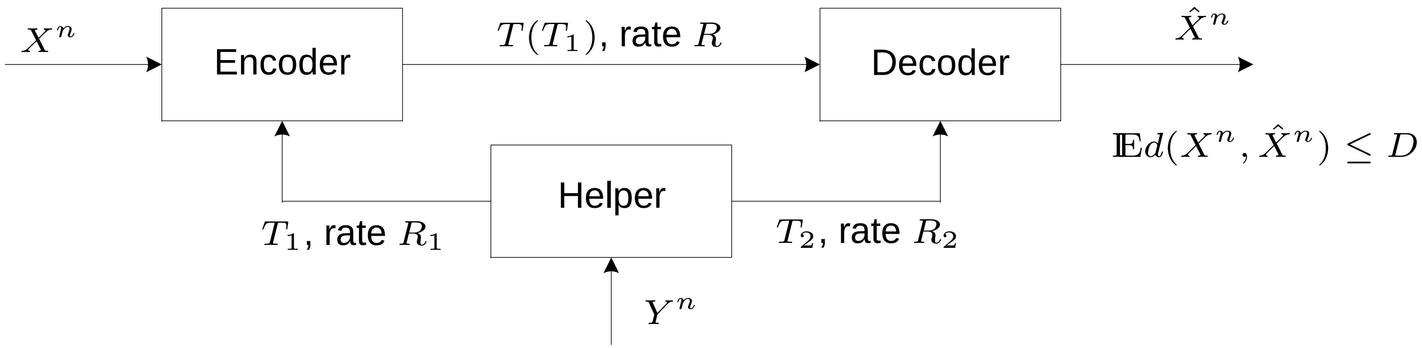


Independent rates



Theorem 2 Let $R_1 > R_2$. For any code with independent helper rates R_1, R_2, there exists a code with common helper rate R_2, with essentially the same performance (the same R and D).

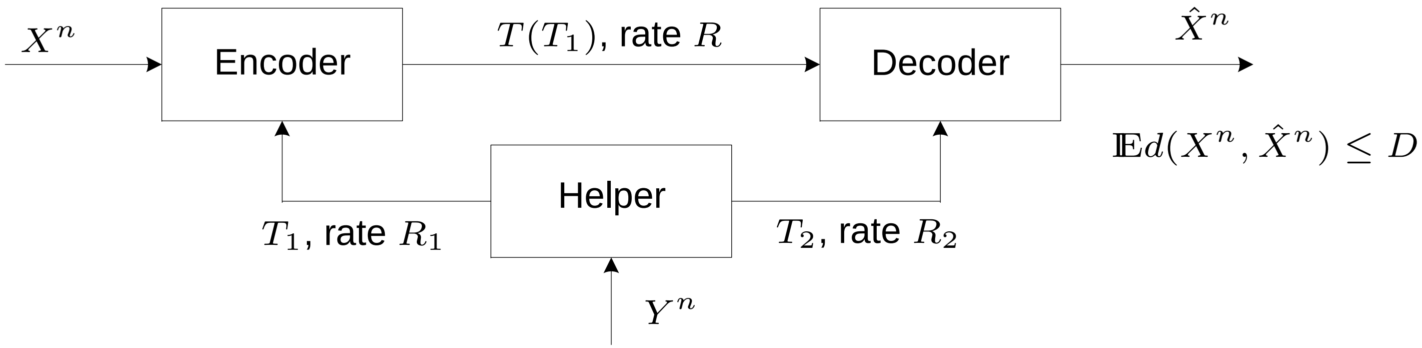
Independent rates



Theorem 2 Let $R_1 > R_2$. For any code with independent helper rates R_1, R_2, there exists a code with common helper rate R_2, with essentially the same performance (the same R and D).

Proof – Operational arguments.

Independent rates



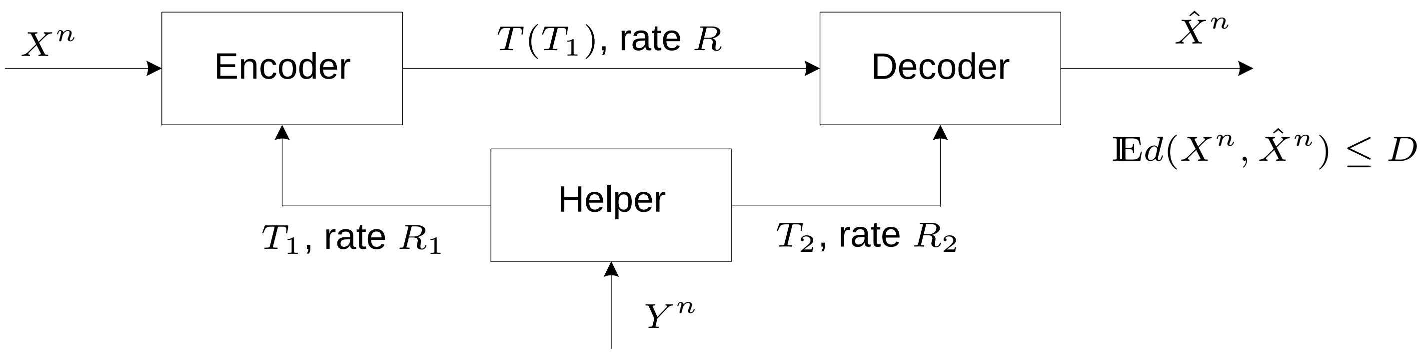
Theorem 2 Let $R_1 > R_2$. For any code with independent helper rates R_1, R_2, there exists a code with common helper rate R_2, with essentially the same performance (the same R and D).

Proof – Operational arguments.

Extra helper rate to the encoder is not used. Moreover, even when independent messages are allowed, sending a common helper message yields optimal performance. ($R_1 \geq R_2$).

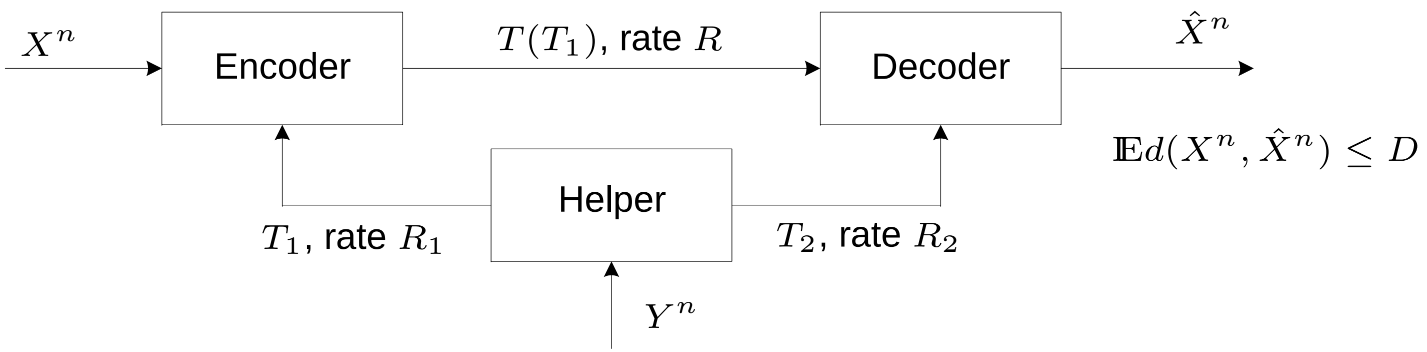
Independent rates (cont'd)

What about reducing R_1 below R_2?



Independent rates (cont'd)

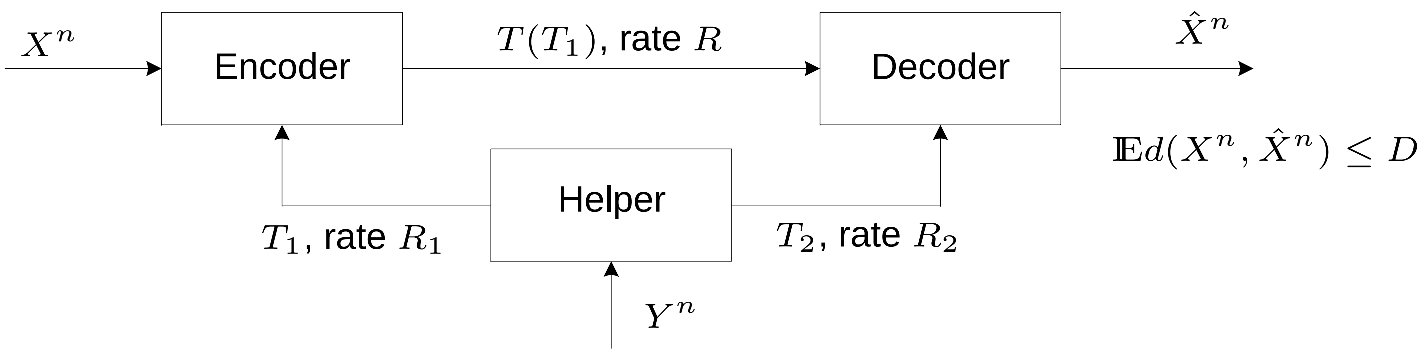
What about reducing R_1 below R_2?



Main idea: the SI X^n is not used to improve the encoder's knowledge on Y^n. But it can be used to reduce the rate R_1.

Independent rates (cont'd)

What about reducing R_1 below R_2?

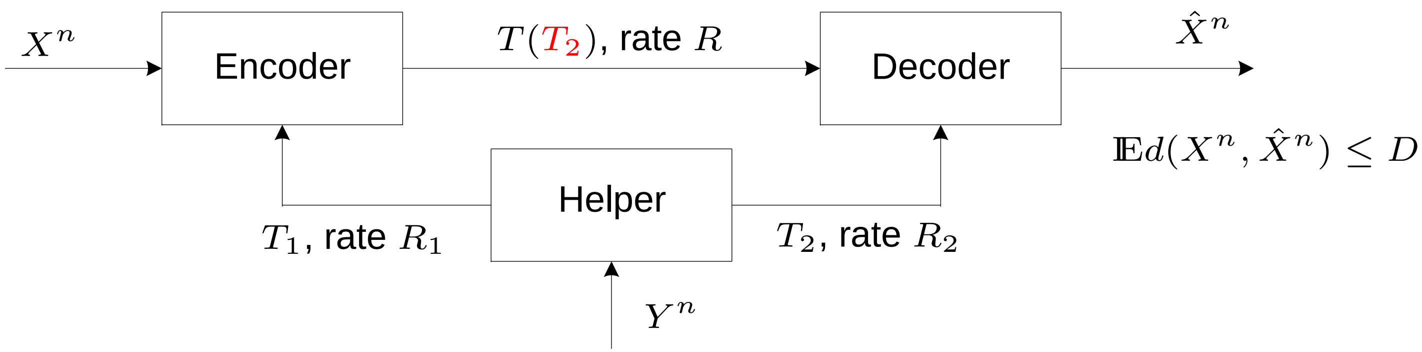


Main idea: the SI X^n is not used to improve the encoder's knowledge on Y^n. But it can be used to reduce the rate R_1.

$$\begin{aligned} R_2 &\geq I(U;Y) \\ R &\geq I(X;\hat{X}|U) \\ D &\geq \mathbb{E}d(X,\hat{X}) \end{aligned}$$

Independent rates *(cont'd)*

What about reducing R_1 below R_2?



Main idea: the SI X^n is not used to improve the encoder's knowledge on Y^n. But it can be used to reduce the rate R_1.

$$\begin{aligned} R_2 &\ge I(U;Y) \\ R &\ge I(X;\hat{X}|U) \\ D &\ge \mathbb{E}d(X,\hat{X}) \end{aligned}$$

We can reduce R_1 slightly below R_2, and the encoder still be able to decode T_2:

$$R_1 \ge I(U;Y) - I(U;X)$$

Example

Doubly symmetric binary source with Hamming distortion measure

$$X = Y \oplus Z, \quad Y \sim \text{Bernoulli}(\frac{1}{2}), \quad Z \sim \text{Bernoulli}(p_z)$$

The region with common helper:

$$R \geq H_b(H_b^{-1}([1 - R']_+) * p_z) - H_b(D).$$

Example

Doubly symmetric binary source with Hamming distortion measure

$$X = Y \oplus Z, \quad Y \sim \text{Bernoulli}(\frac{1}{2}), \quad Z \sim \text{Bernoulli}(p_z)$$

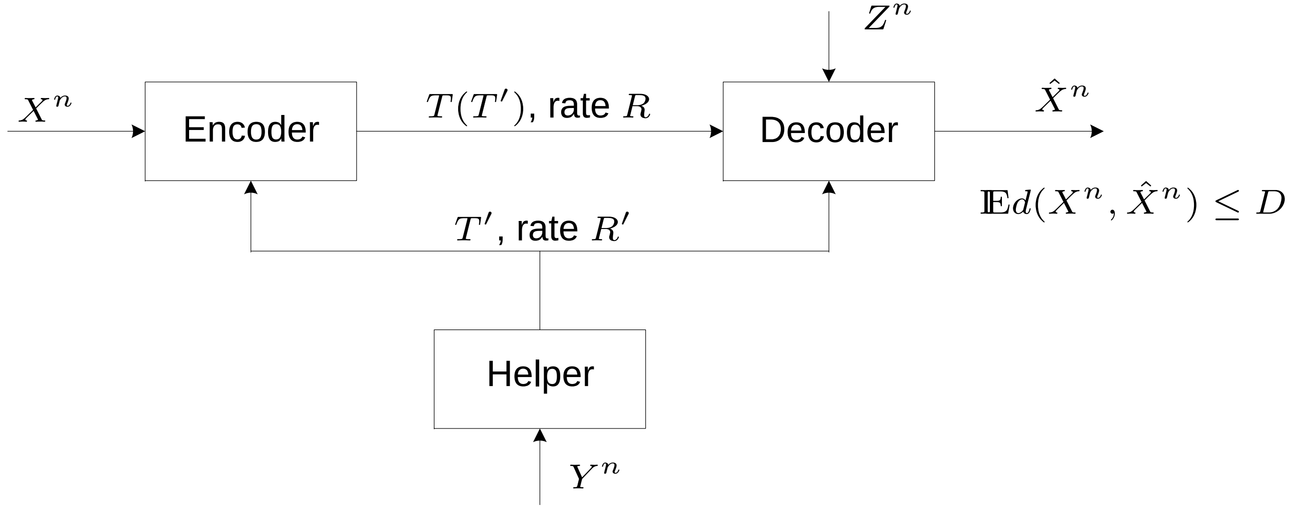
The region with common helper:

$$R \geq H_b(H_b^{-1}([1 - R']_+) * p_z) - H_b(D).$$

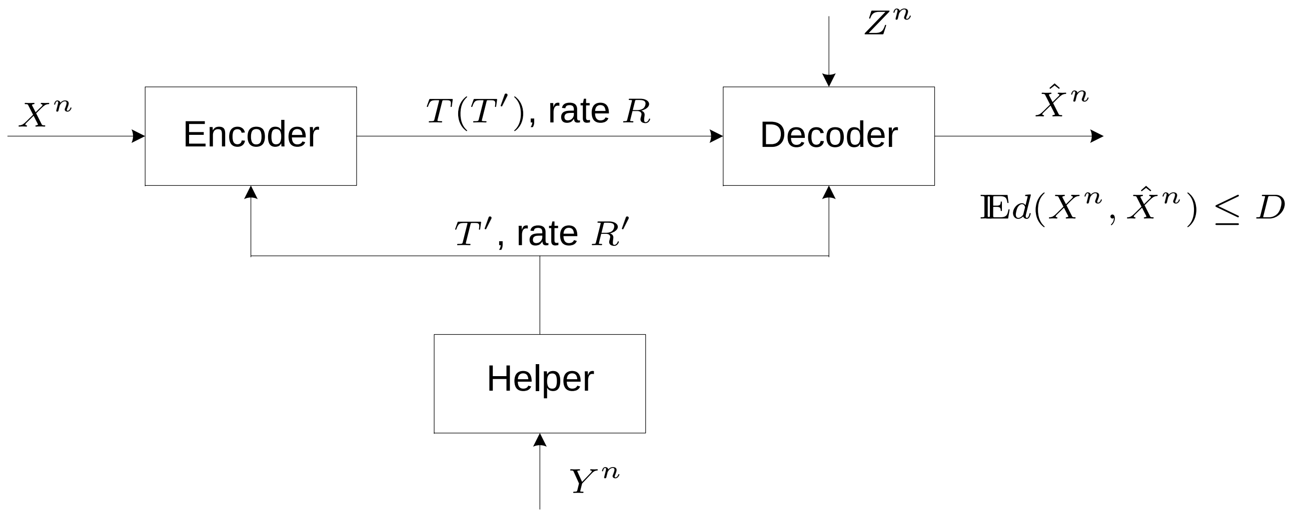
The region with two rates, $R_2 \geq R_1$

$$\begin{aligned} R &\geq H_b(H_b^{-1}([1 - R_2]_+) * p_z) - H_b(D) \\ R_1 &\geq H_b(H_b^{-1}([1 - R_2]_+) * p_z) - [1 - R_2]_+. \end{aligned}$$

Common Helper and Decoder SI

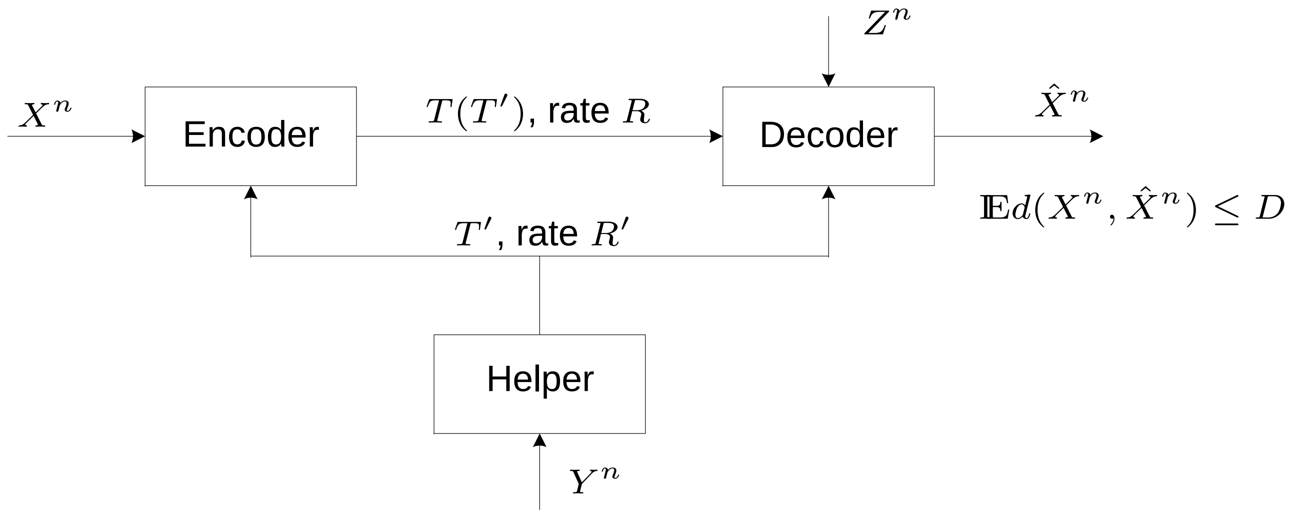


Common Helper and Decoder SI



- We assume the Markov chain $Y - X - Z$

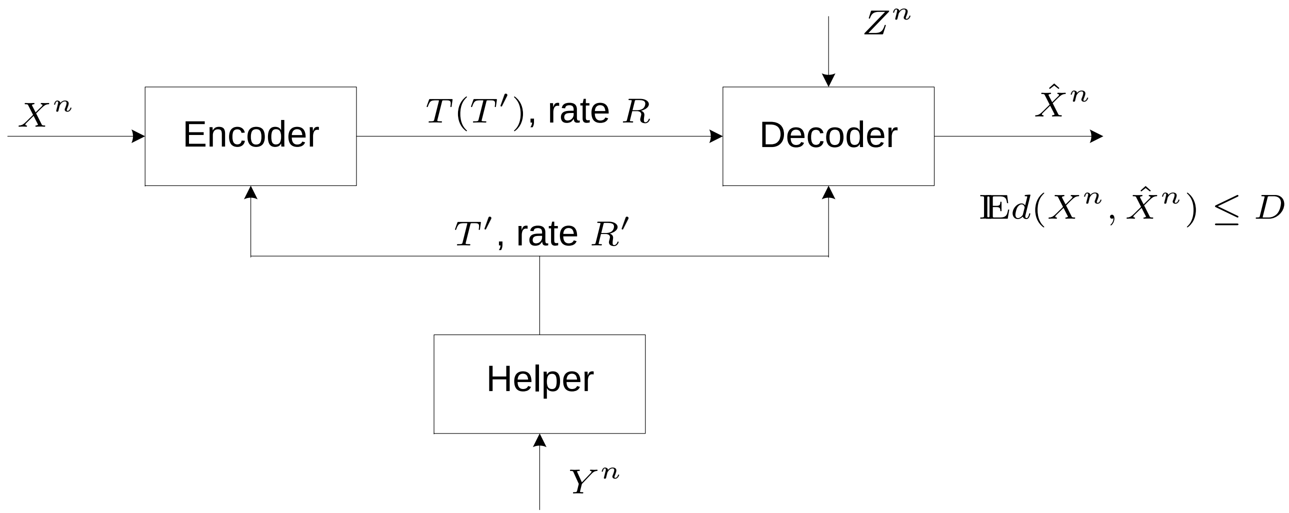
Common Helper and Decoder SI



- We assume the Markov chain $Y - X - Z$
- The definitions and communication protocol are as before, with

$$\hat{X}^n = \hat{X}^n(T', T, Z^n).$$

Common Helper and Decoder SI



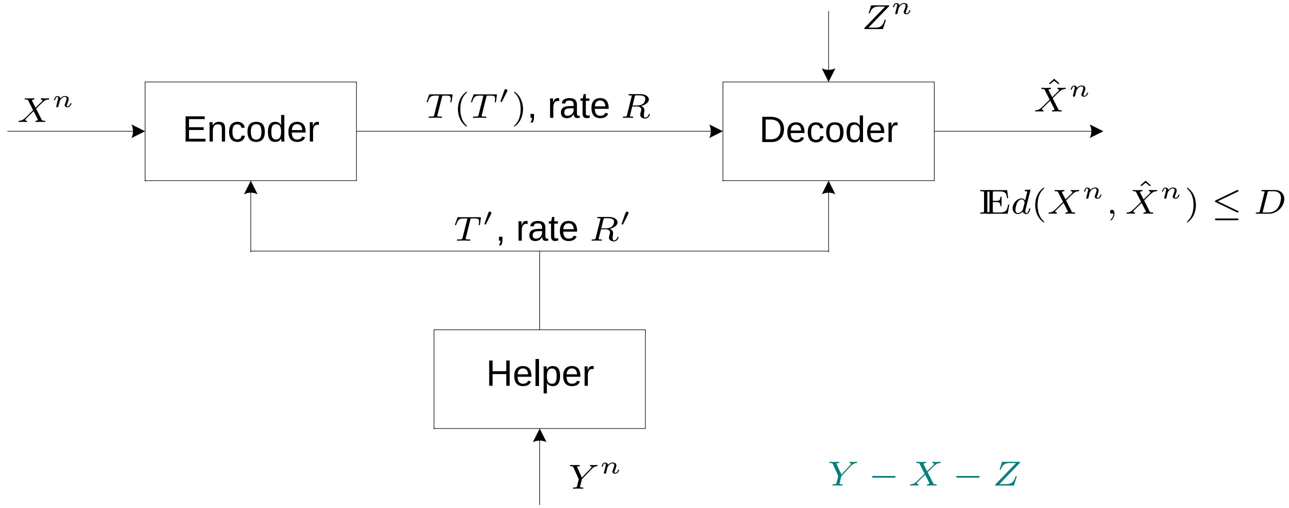
- We assume the Markov chain $Y - X - Z$
- The definitions and communication protocol are as before, with

$$\hat{X}^n = \hat{X}^n(T', T, Z^n).$$

- The rate distortion region is denoted $\mathcal{R}(D|Z)$.

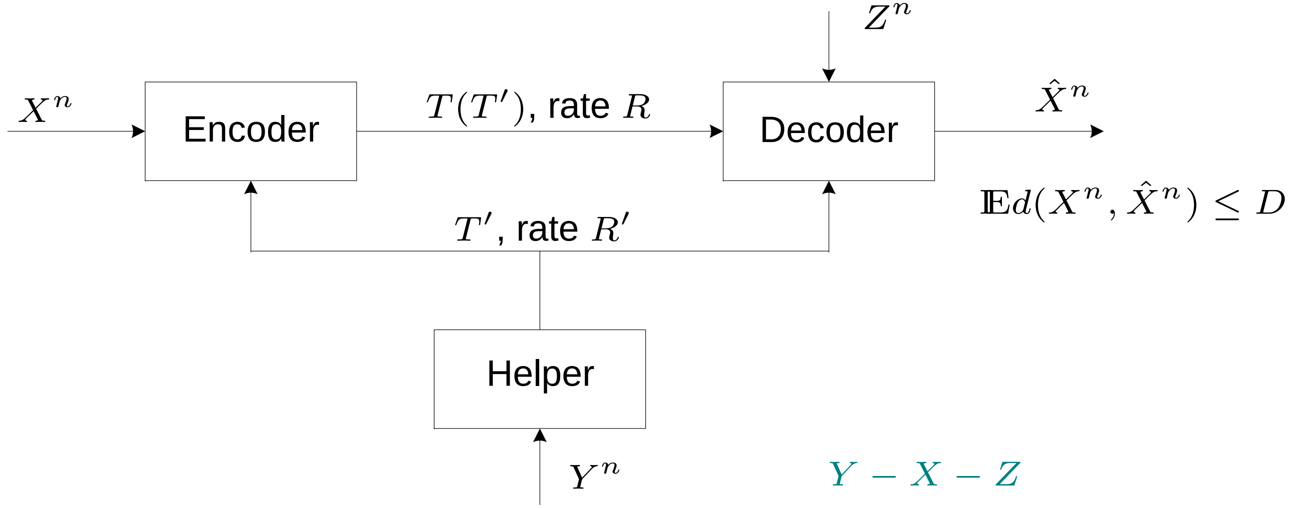
Main Result

Main Result



$Y - X - Z$

Main Result



$Y - X - Z$

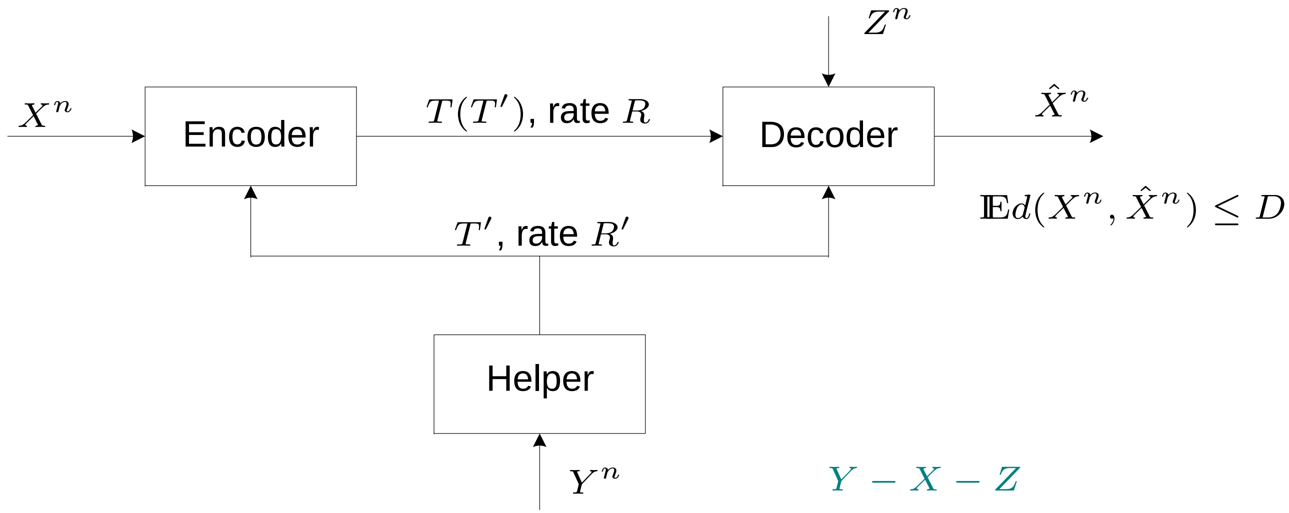
Theorem 3 $(R, R') \in \mathcal{R}(D|Z)$ if and only if

$$\begin{aligned} R' &\geq I(U;Y|Z) \\ R &\geq I(V;X|U,Z) \\ D &\geq \mathbb{E}d(X, \hat{X}(U,V,Z)) \end{aligned}$$

for some joint distribution of the form

$$P_{X,Y} \cdot P_{Z|X} \cdot P_{U|Y} \cdot P_{V|U,X}.$$

Main Result



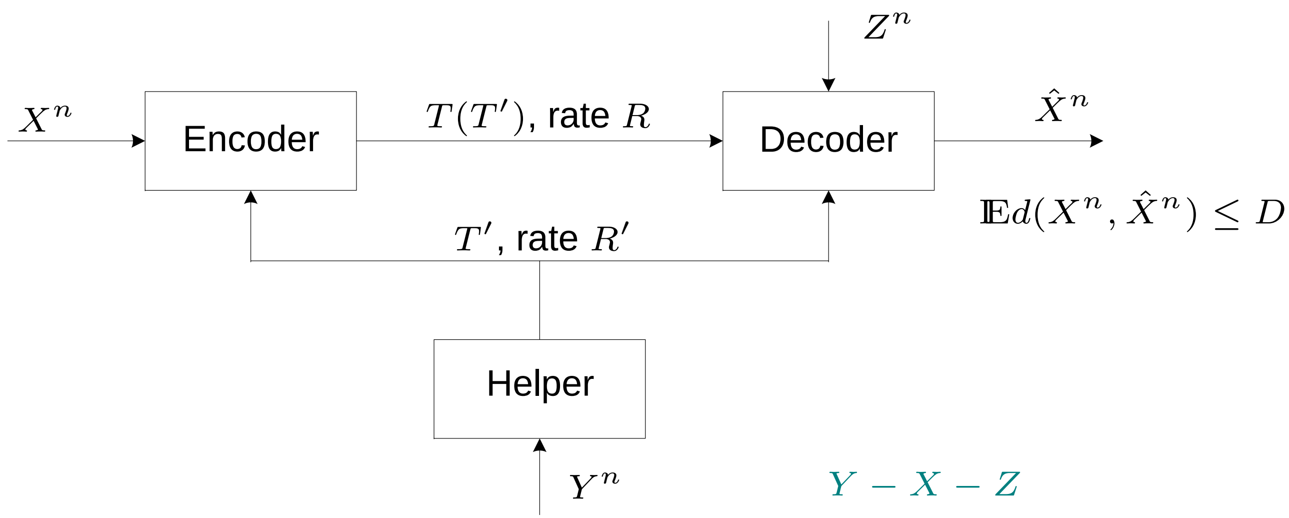
Theorem 3 $(R, R') \in \mathcal{R}(D|Z)$ if and only if

$$\begin{aligned} R' &\geq I(U;Y|Z) & \quad U - Y - (X,V,Z) \\ R &\geq I(V;X|U,Z) & \quad V - (U,X) - (Y,Z) \\ D &\geq \mathbb{E}d(X, \hat{X}(U,V,Z)) \end{aligned}$$

for some joint distribution of the form

$$P_{X,Y} \cdot P_{Z|X} \cdot P_{U|Y} \cdot P_{V|U,X}.$$

Main Result



Theorem 3 $(R, R') \in \mathcal{R}(D|Z)$ if and only if

$$\begin{aligned} R' &\geq I(U;Y|Z) && U - Y - (X,V,Z) \\ R &\geq I(V;X|U,Z) && V - (U,X,Y) - Z \\ D &\geq \mathbb{E}d(X, \hat{X}(U,V,Z)) \end{aligned}$$

for some joint distribution of the form

$$P_{X,Y} \cdot P_{Z|X} \cdot P_{U|Y} \cdot P_{V|U,X,Y}.$$

Markov Invariancy

Let

$$\begin{aligned} \tilde{P}_{U,V,X,Y,Z} &= P_{X,Y} \cdot P_{Z|X} \cdot P_{U|Y} \cdot P_{V|U,X,Y} \\ P_{V,U,X,Y,Z} &= P_{X,Y} \cdot P_{Z|X} \cdot P_{U|Y} \cdot P_{V|U,X} \end{aligned}$$

with $P_{V|U,X}$ the conditional marginal induced by $\tilde{P}_{U,V,X,Y,Z}$.

Markov Invariancy

Let

$$\tilde{P}_{U,V,X,Y,Z} = P_{X,Y} \cdot P_{Z|X} \cdot P_{U|Y} \cdot P_{V|U,X,Y}$$
$$P_{V,U,X,Y,Z} = P_{X,Y} \cdot P_{Z|X} \cdot P_{U|Y} \cdot P_{V|U,X}$$

with $P_{V|U,X}$ the conditional marginal induced by $\tilde{P}_{U,V,X,Y,Z}$.

The region

$$R' \geq I(U;Y|Z)$$
$$R \geq I(V;X|U,Z)$$
$$D \geq \mathbb{E}d(X,\hat{X}(U,V,Z))$$

is the same under P and $\tilde{P}$.

Markov Invariancy

Let

$$\tilde{P}_{U,V,X,Y,Z} = P_{X,Y} \cdot P_{Z|X} \cdot P_{U|Y} \cdot P_{V|U,X,Y}$$
$$P_{V,U,X,Y,Z} = P_{X,Y} \cdot P_{Z|X} \cdot P_{U|Y} \cdot P_{V|U,X}$$

with $P_{V|U,X}$ the conditional marginal induced by $\tilde{P}_{U,V,X,Y,Z}$.

The region

$$R' \geq I(U;Y|Z)$$
$$R \geq I(V;X|U,Z)$$
$$D \geq \mathbb{E}d(X,\hat{X}(U,V,Z))$$

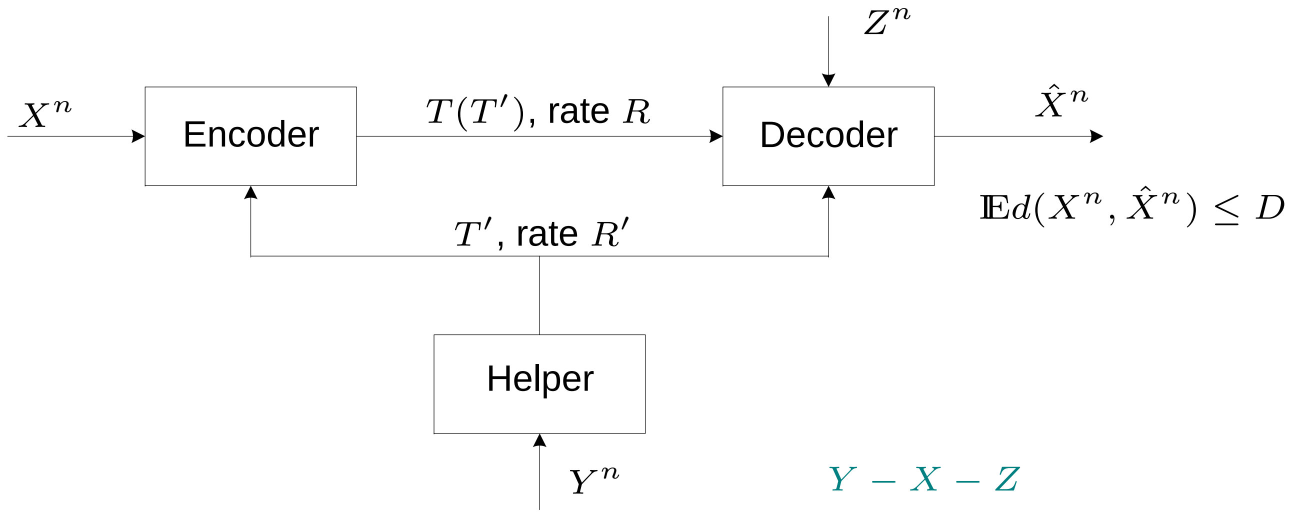
is the same under P and $\tilde{P}$.

Thus, it is enough to prove:

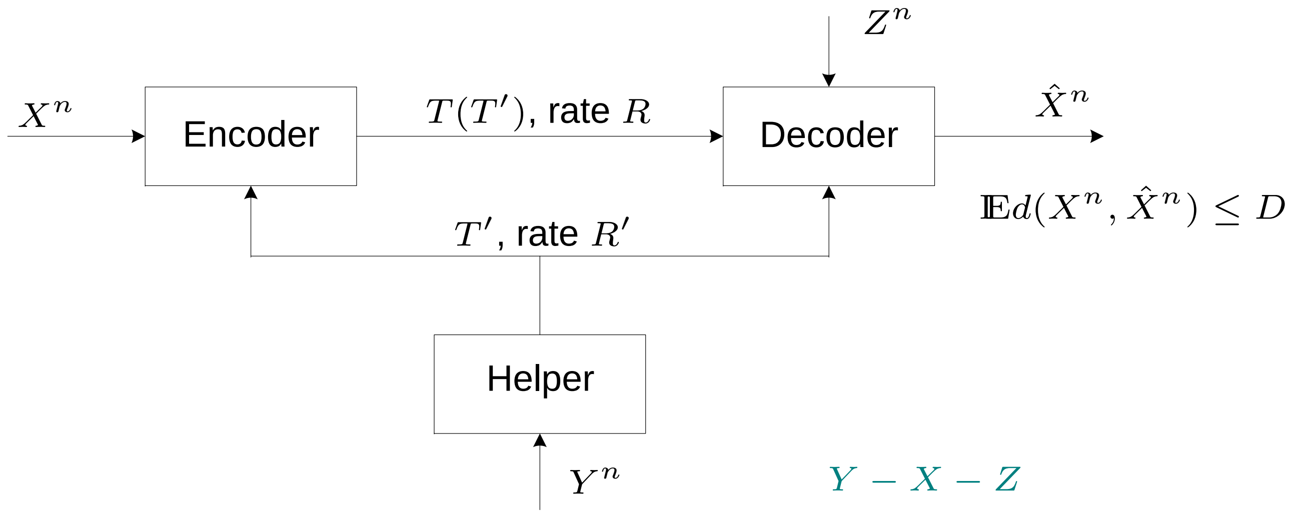
Achievability for any $P_{X,Y} \cdot P_{Z|X} \cdot P_{U|Y} \cdot P_{V|U,X}$

Upper bound for $P_{X,Y} \cdot P_{Z|X} \cdot P_{U|Y} \cdot P_{V|U,X,Y}$

SI Utilization

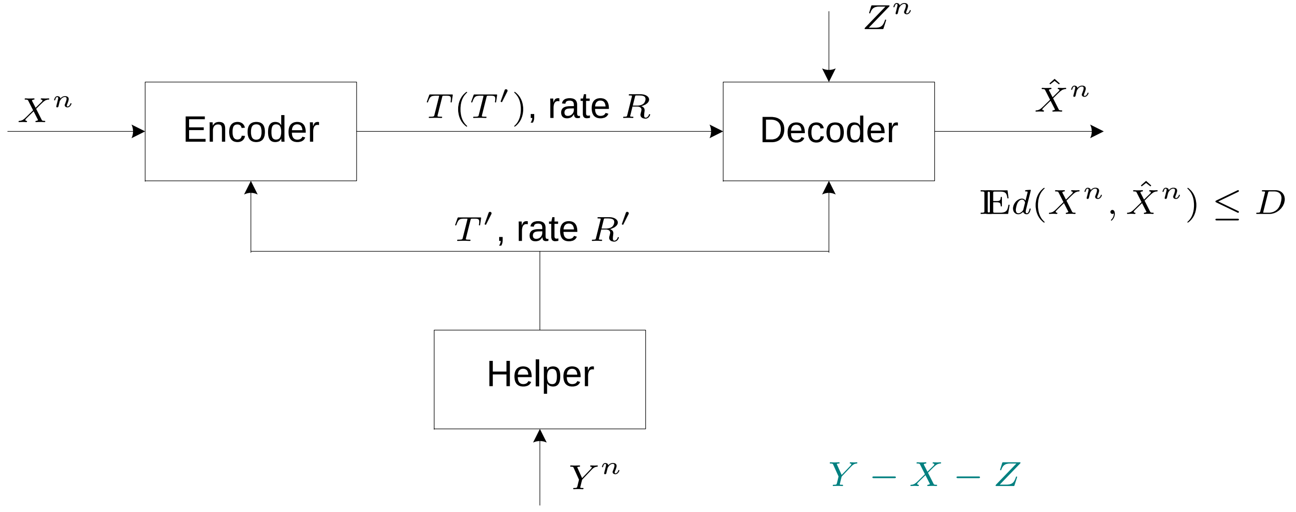


SI Utilization



Due to $Y - X - Z$, X is a better SI than Z, in coding for Y (helper).

SI Utilization



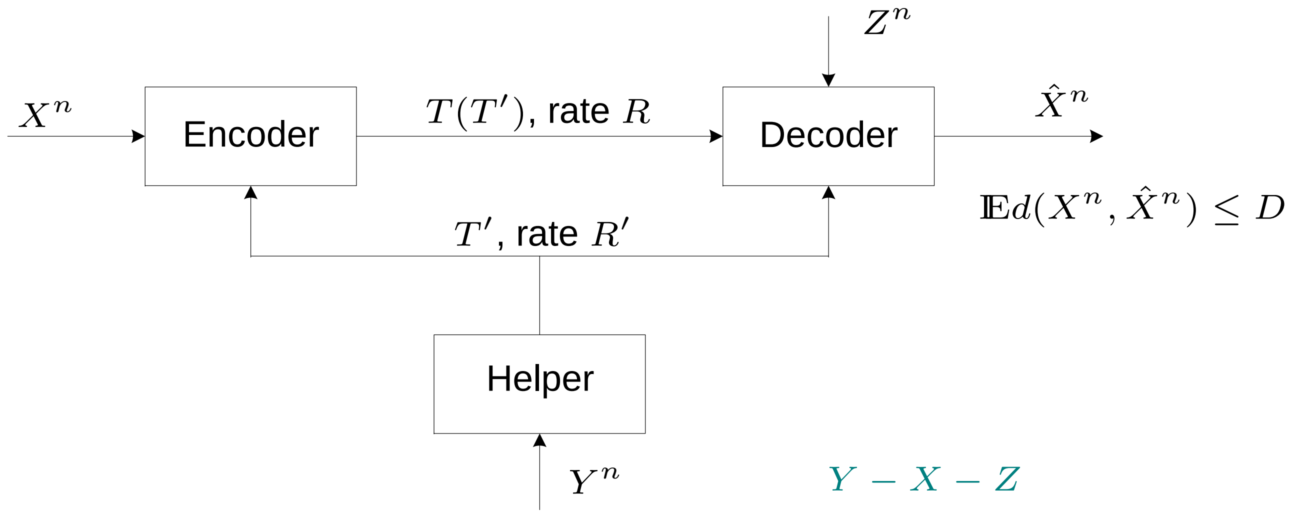
Due to $Y - X - Z$, X is a better SI than Z, in coding for Y (helper).

$\Rightarrow$ Would expect

$$R' \ge I(U;Y|Z) + I(W;Y|U,X), \qquad (U,W) - Y - X.$$

“Rate distortion when side information may be absent," Heegard & Berger, 1985.

SI Utilization



Due to $Y - X - Z$, X is a better SI than Z, in coding for Y (helper).

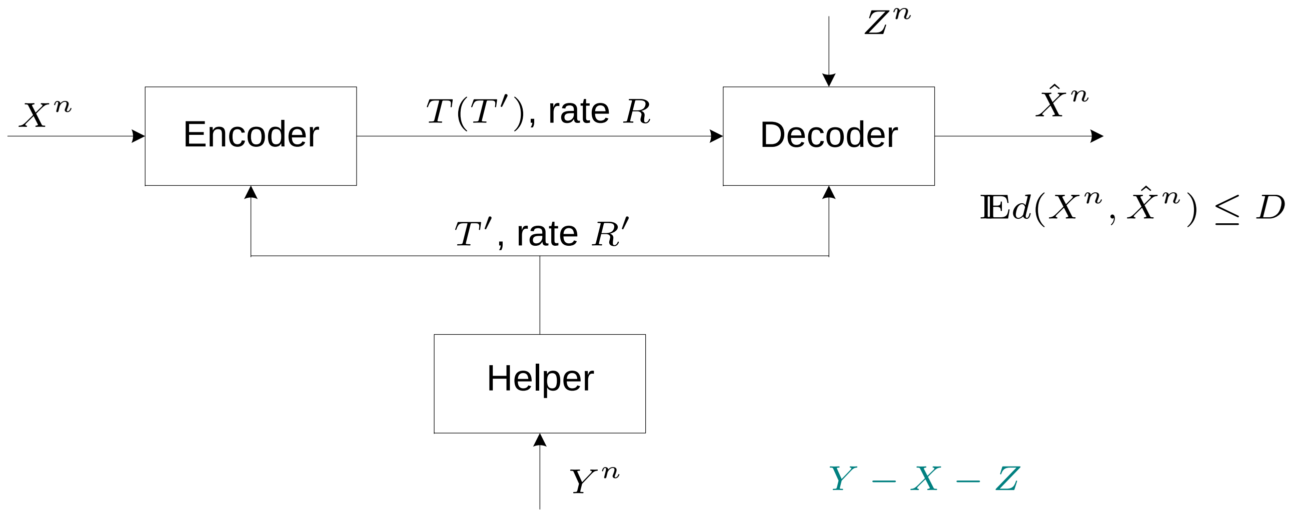
$\Rightarrow$ Would expect

$$R' \geq I(U;Y|Z) + I(W;Y|U,X), \qquad (U,W) - Y - X.$$

"Rate distortion when side information may be absent," Heegard & Berger, 1985.

Yet, our result says: $R' \geq I(U;Y|Z)$

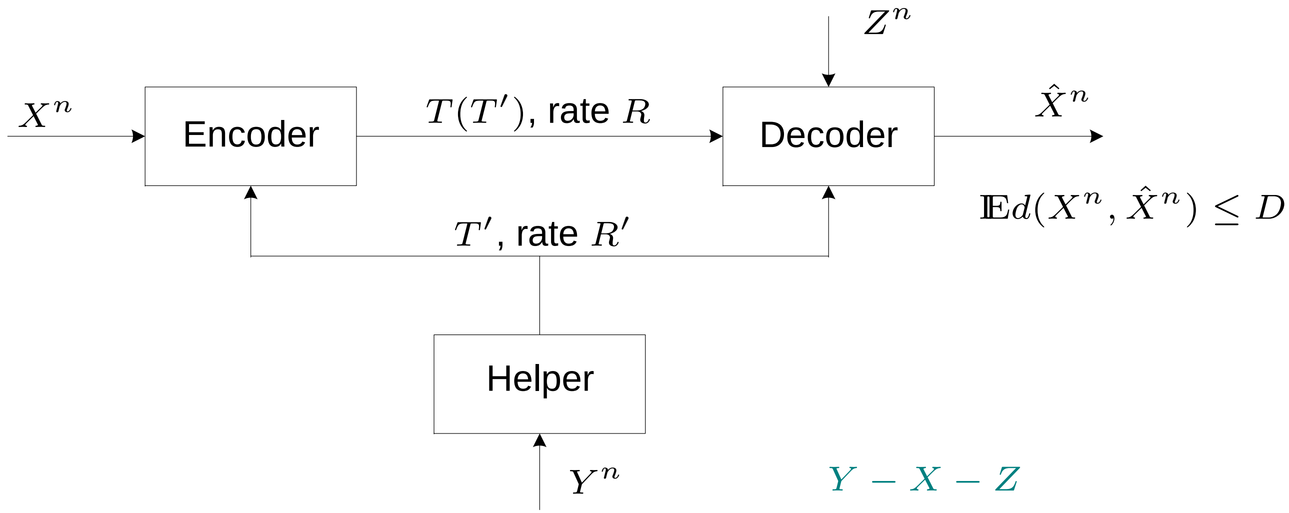
SI Utilization



$Y - X - Z$

$$\begin{aligned} R' &\geq I(U;Y|Z) \\ R &\geq I(V;X|U,Z) \\ D &\geq \mathbb{E}d(X, \hat{X}(U,V,Z)) \end{aligned}$$

SI Utilization

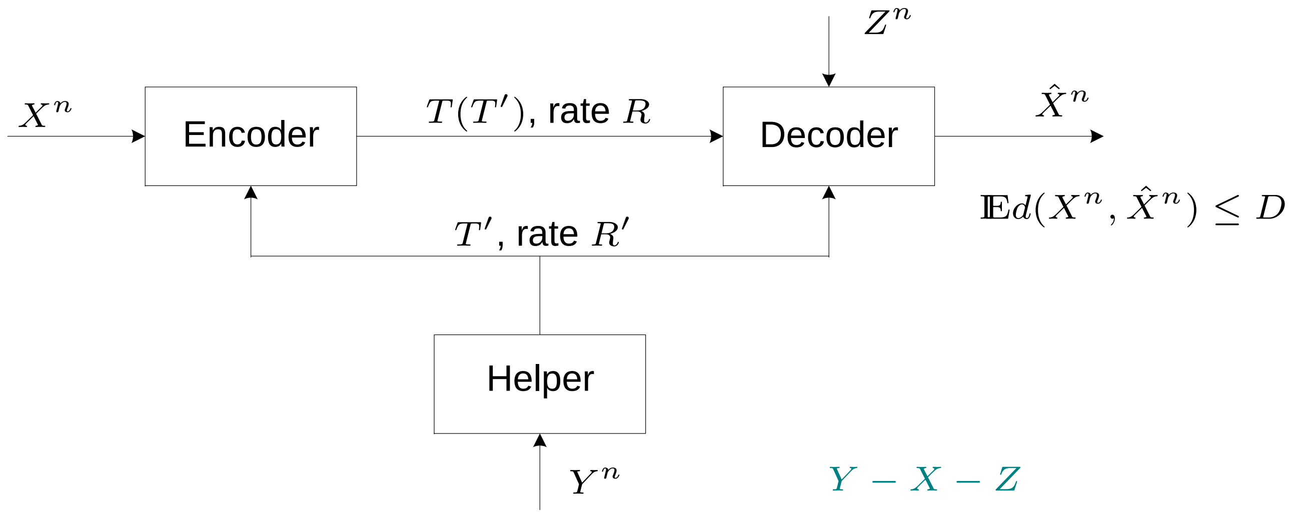


$$
\begin{aligned}
R' &\geq I(U;Y|Z) \\
R &\geq I(V;X|U,Z) \\
D &\geq \mathbb{E}d(X, \hat{X}(U,V,Z))
\end{aligned}
$$

Conclusion:

When sending a common message to the encoder and decoder, the source encoder does not use the extra SI X^n.

SI Utilization

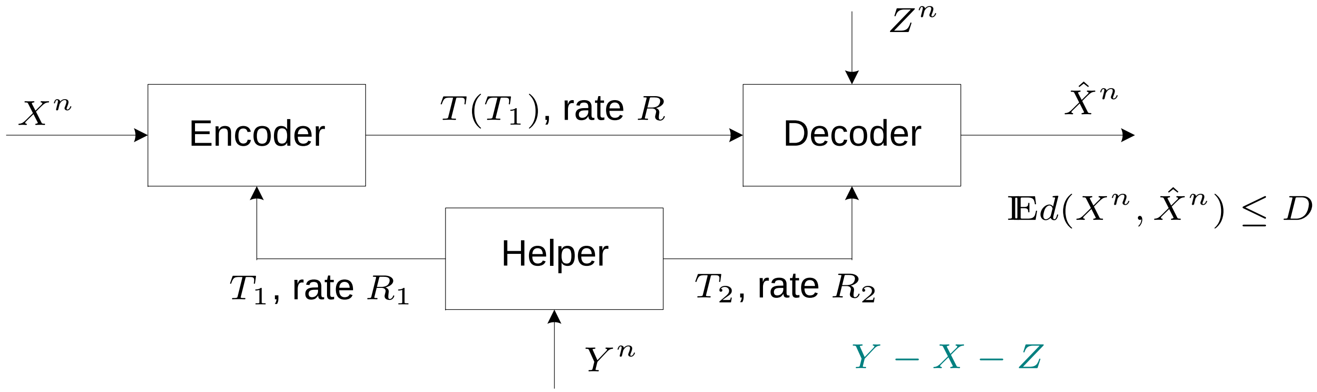


$$
\begin{aligned}
R' &\geq I(U;Y|Z) \\
R &\geq I(V;X|U,Z) \\
D &\geq \mathbb{E}d(X,\hat{X}(U,V,Z))
\end{aligned}
$$

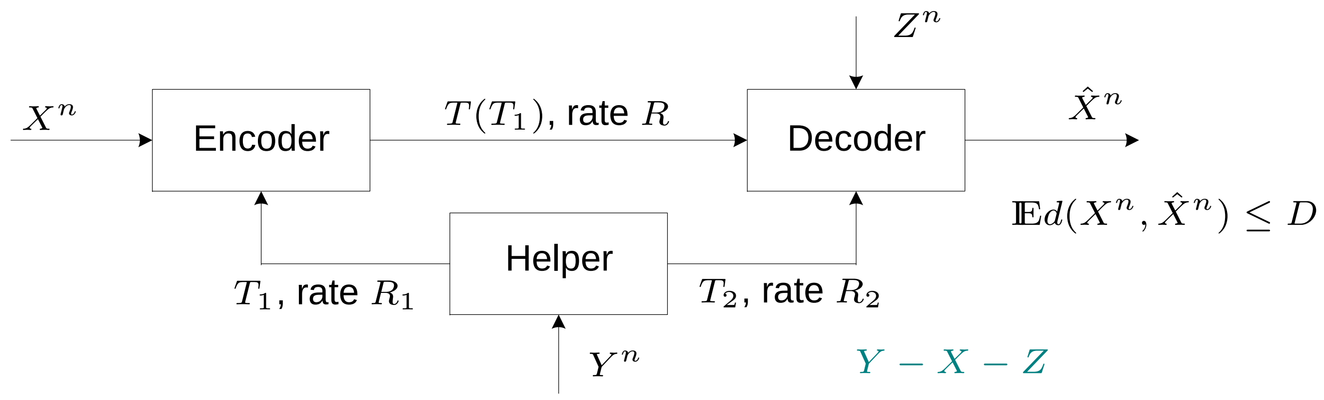
Conclusion:
When sending a common message to the encoder and decoder, the source encoder does not use the extra SI X^n.

Neither would he use extra rate.

Independent Rates With SI - Main Result

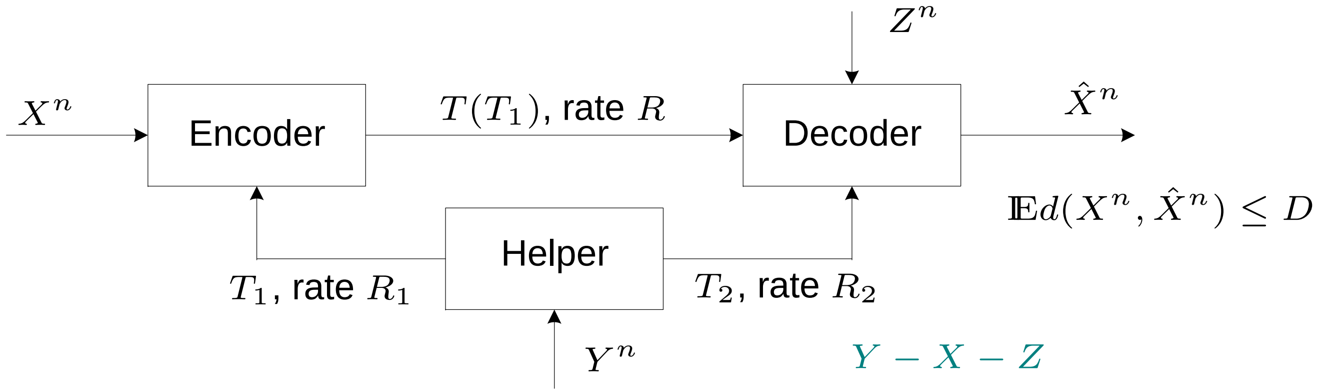


Independent Rates With SI - Main Result



Theorem 4 Let $R_1 > R_2$, and let the side information Z satisfy $Y - X - Z$. For any code with independent helper rates R_1, R_2, there exists a code with common helper rate R_2, with essentially the same performance (the same R and D).

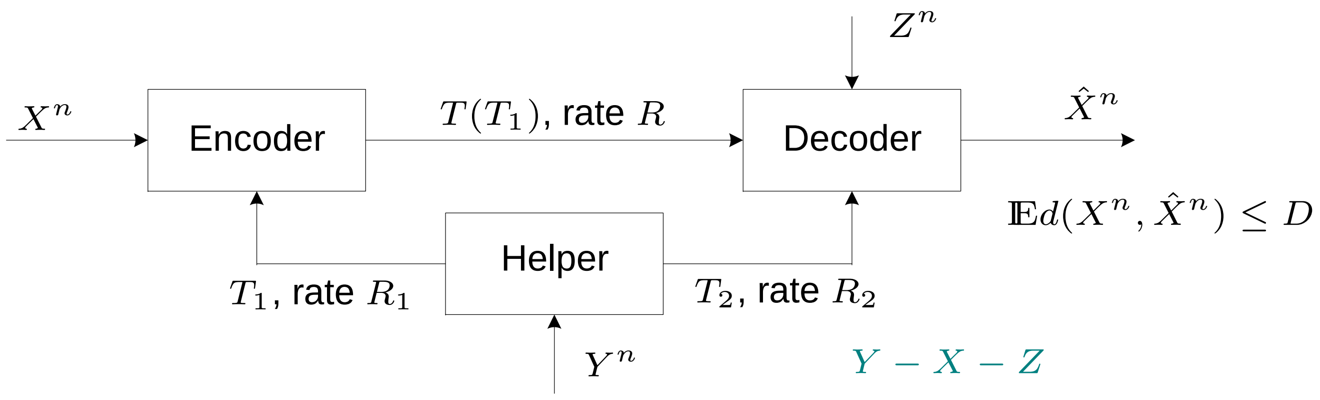
Independent Rates With SI - Main Result



Theorem 4 Let $R_1 > R_2$, and let the side information Z satisfy $Y - X - Z$. For any code with independent helper rates R_1, R_2, there exists a code with common helper rate R_2, with essentially the same performance (the same R and D).

Proof – Operational arguments.

Independent Rates With SI - Main Result



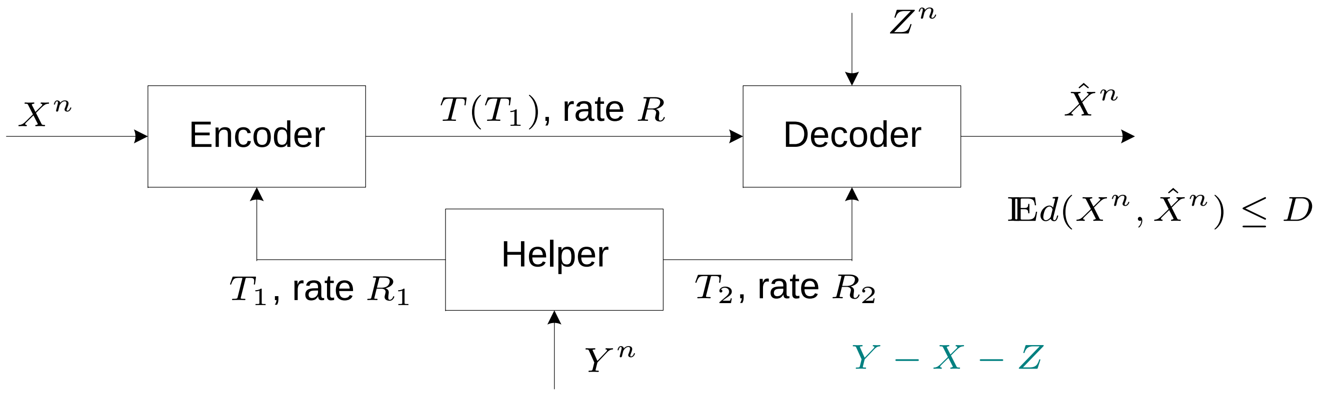
Theorem 4 Let $R_1 > R_2$, and let the side information Z satisfy $Y - X - Z$. For any code with independent helper rates R_1, R_2, there exists a code with common helper rate R_2, with essentially the same performance (the same R and D).

Proof – Operational arguments.

Extra helper rate to the encoder is not used. Moreover, even when independent messages are allowed, sending a common helper message yields optimal performance. $(R_1 \geq R_2)$.

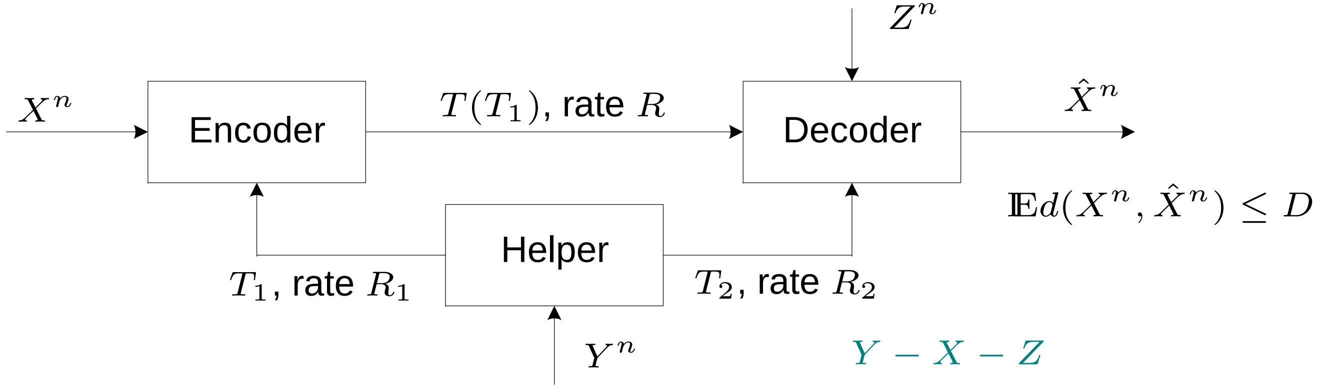
Independent Rates With SI (cont'd)

What about reducing R_1 below R_2?



Independent Rates With SI *(cont'd)*

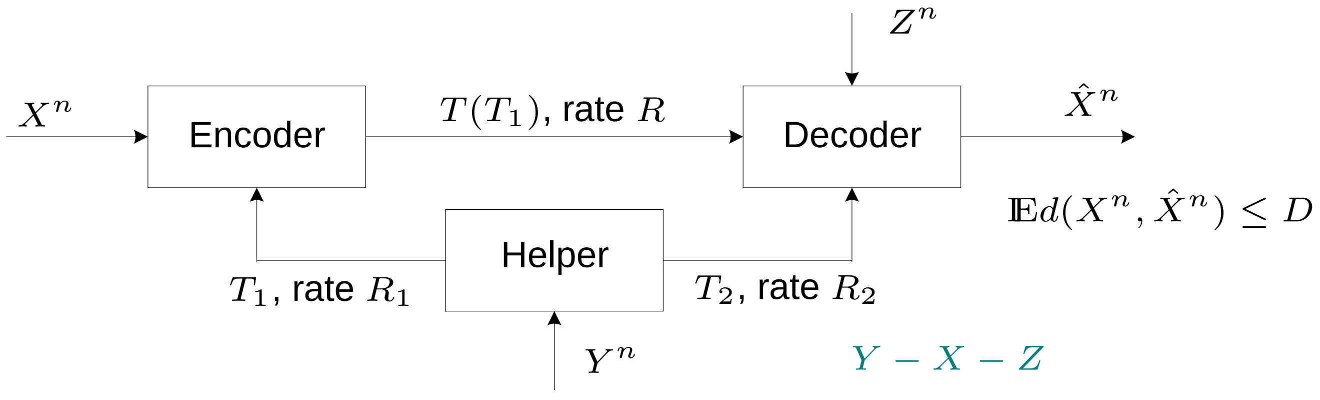
What about reducing R_1 below R_2?



$$
\begin{aligned}
R' &\geq I(U;Y|Z) = I(U;Y) - I(U;Z) \\
R &\geq I(V;X|U,Z) \\
D &\geq \mathbb{E}d(X, \hat{X}(U,V,Z))
\end{aligned}
$$

Independent Rates With SI *(cont'd)*

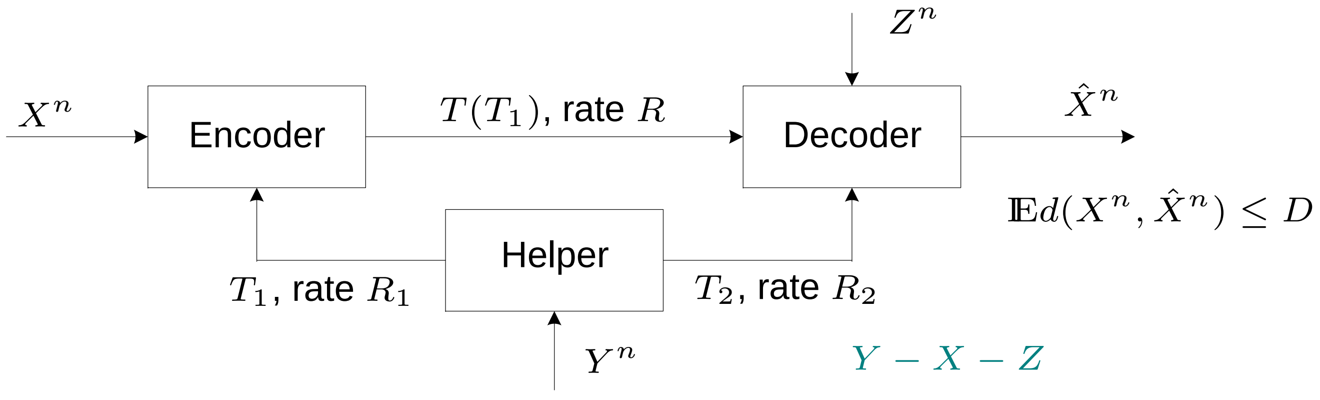
What about reducing R_1 below R_2?



$$
\begin{aligned}
R_2 &\geq I(U;Y|Z) = I(U;Y) - I(U;Z) \\
R &\geq I(V;X|U,Z) \\
D &\geq \mathbb{E}d(X, \hat{X}(U,V,Z))
\end{aligned}
$$

Independent Rates With SI *(cont'd)*

What about reducing R_1 below R_2?

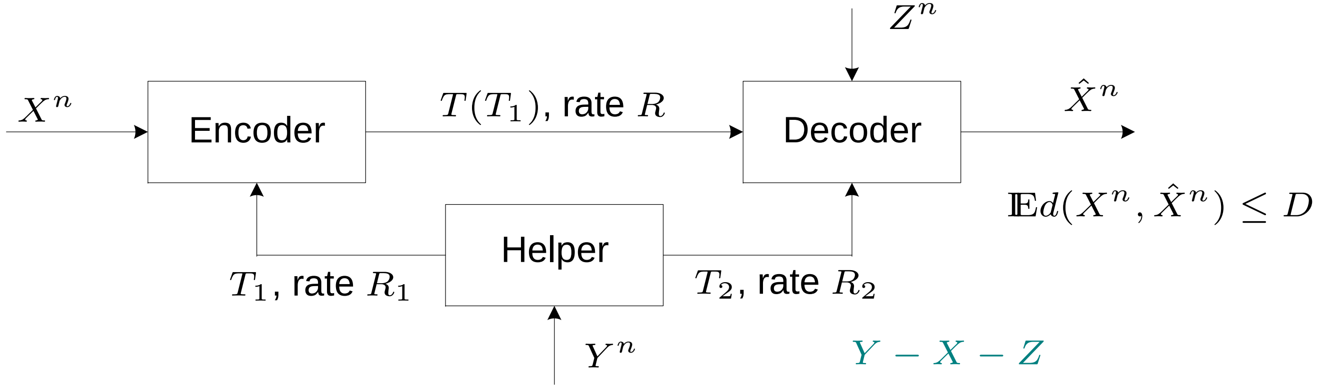


$$
\begin{aligned}
R_2 &\ge I(U;Y|Z) = I(U;Y) - I(U;Z) \\
R &\ge I(V;X|U,Z) \\
D &\ge \mathbb{E}d(X, \hat{X}(U,V,Z))
\end{aligned}
$$

Main idea – use binning w.r.t X to reduce the *rate* to the encoder (but not to improve the quality of $\hat{Y}^n$)

Independent Rates With SI *(cont'd)*

What about reducing R_1 below R_2?



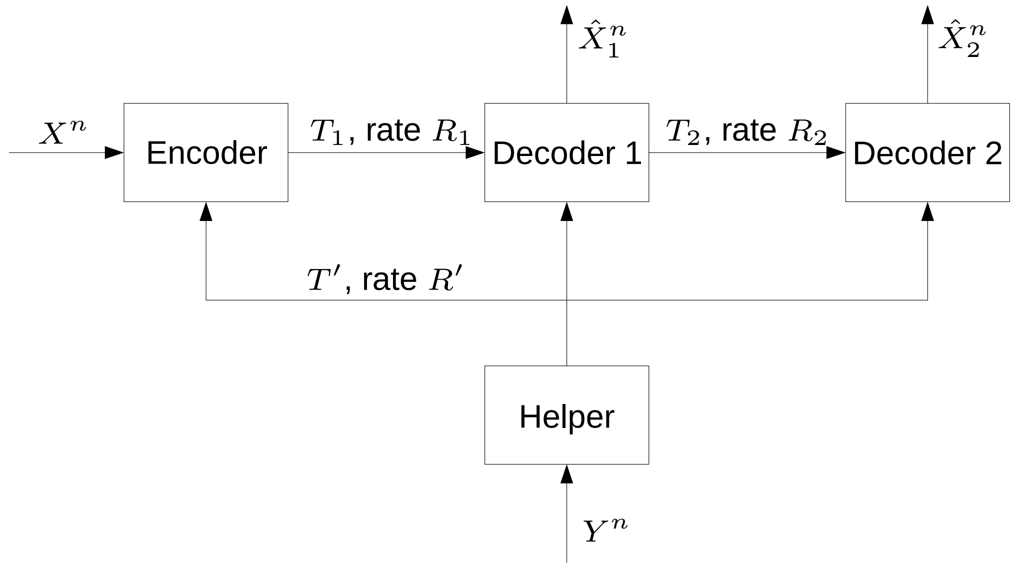
$$
\begin{aligned}
R_2 &\geq I(U;Y|Z) = I(U;Y) - I(U;Z) \\
R &\geq I(V;X|U,Z) \\
D &\geq \mathbb{E}d(X, \hat{X}(U,V,Z))
\end{aligned}
$$

Main idea – use binning w.r.t X to reduce the *rate* to the encoder (but not to improve the quality of $\hat{Y}^n$)

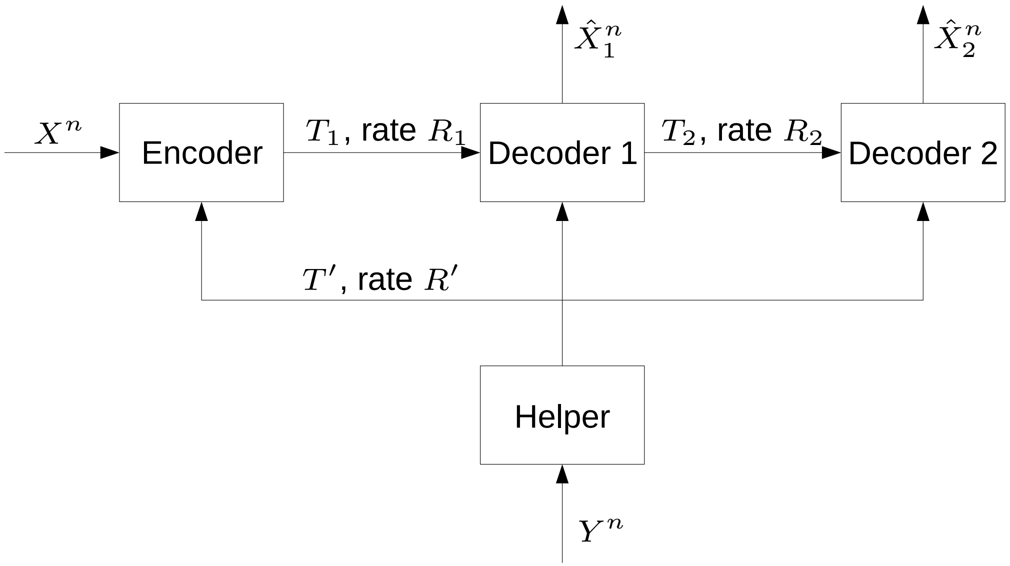
$$
\begin{aligned}
R_1 &\geq I(U;Y|Z) - I(U;X|Z) \\
&= I(U;Y) - I(U;X)
\end{aligned}
$$

Due to $U - Y - X - Z$, R_1 is strictly less than R_2.

Cascade Rate-Distortion With a Helper



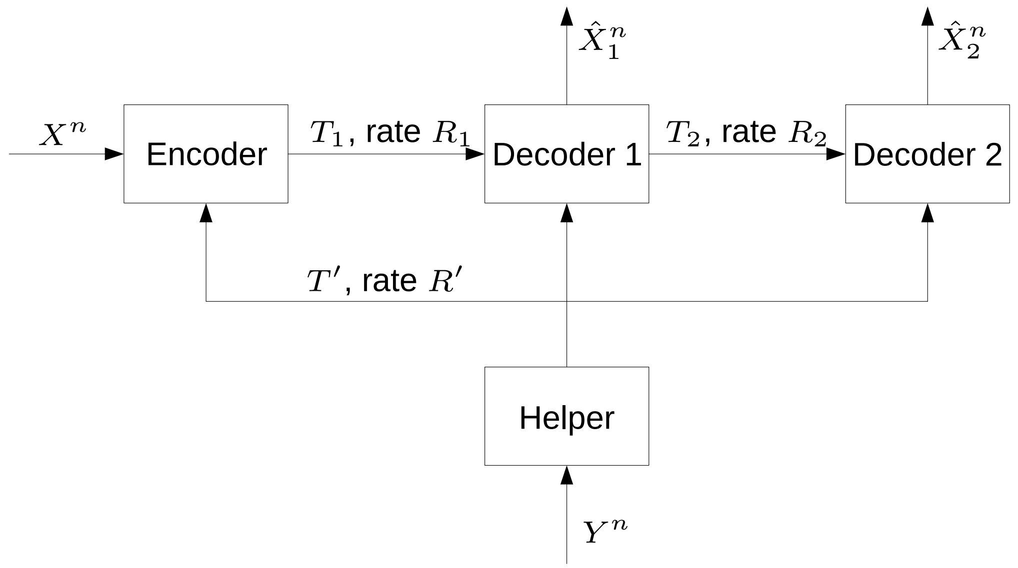
Cascade Rate-Distortion With a Helper



- The definitions and communication protocol are as before, with

$$\hat{X}_i^n = \hat{X}_i^n(T', T_i), \qquad i = 1, 2,$$

Cascade Rate-Distortion With a Helper



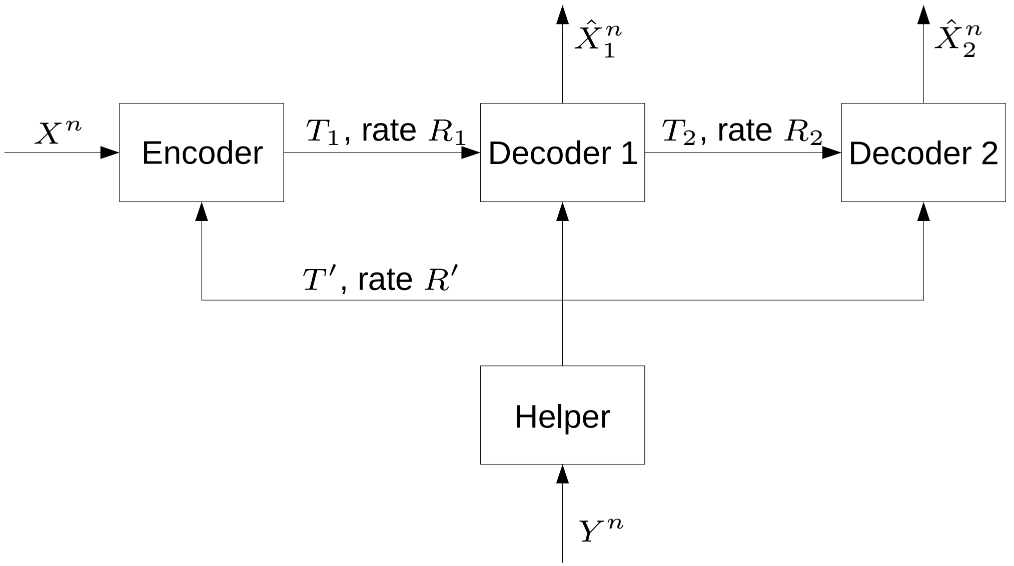
- The definitions and communication protocol are as before, with

$$\hat{X}_i^n = \hat{X}_i^n(T', T_i), \qquad i = 1, 2,$$

and possibly different distortion measures

$$\mathbb{E} d_i(X^n, \hat{X}_i^n) \leq D_i, \quad i = 1, 2.$$

Cascade Rate-Distortion With a Helper



- The definitions and communication protocol are as before, with

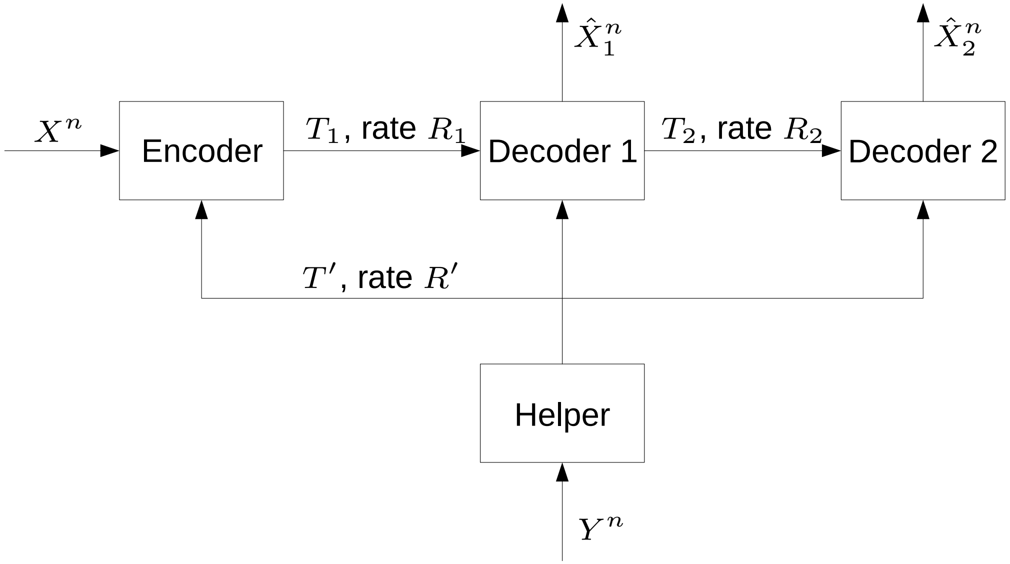
$$\hat{X}_i^n = \hat{X}_i^n(T', T_i), \qquad i = 1, 2,$$

and possibly different distortion measures

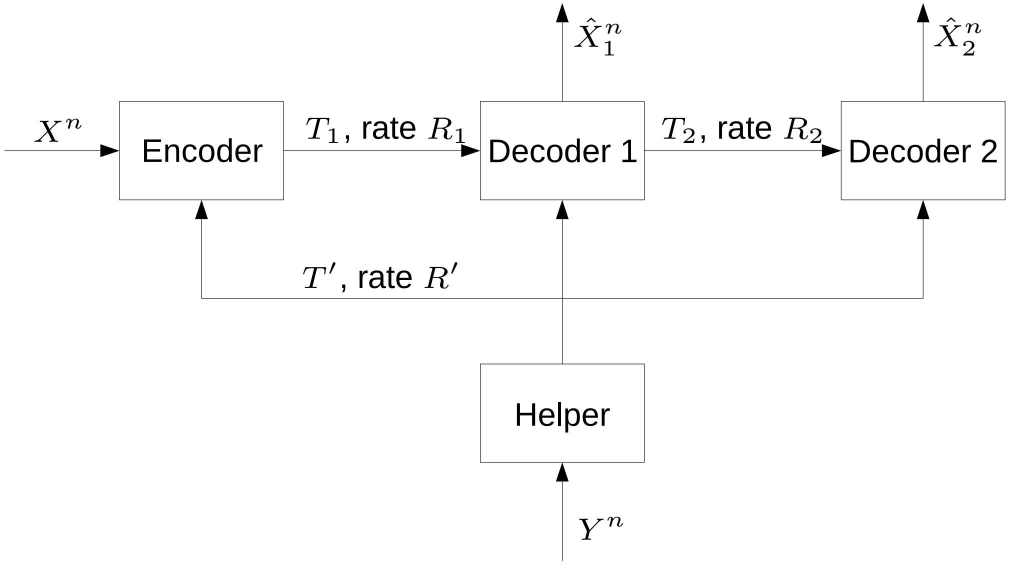
$$\mathbb{E}d_i(X^n, \hat{X}_i^n) \leq D_i, \quad i = 1, 2.$$

- the rate distortion region is denoted $\mathcal{R}_c(D_1, D_2)$.

Main Result


$\hat{X}_1^n$
$\hat{X}_2^n$
X^n
Encoder
T_1, rate R_1
Decoder 1
T_2, rate R_2
Decoder 2
T', rate R'
Helper
Y^n

Main Result



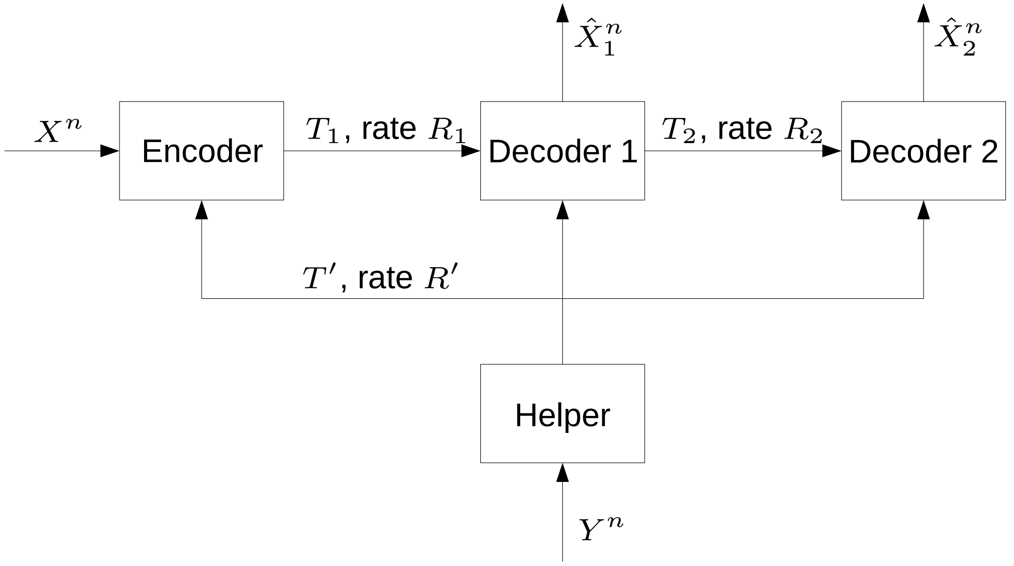
Theorem 5 $(R', R_1, R_2) \in \mathcal{R}_c(D_1, D_2)$ if and only if

$$\begin{aligned} R' &\geq I(Y;V) \\ R_1 &\geq I(X;\hat{X}_1,\hat{X}_2|V) \\ R_2 &\geq I(X;\hat{X}_2|V) \\ D_i &\geq \mathbb{E}d(X,\hat{X}_i), \quad i=1,2 \end{aligned}$$

for some joint distribution of the form

$$P_{X,Y} \cdot P_{V|Y} \cdot P_{U|Y} \cdot P_{\hat{X}_1,\hat{X}_2|V,X}.$$

Main Result



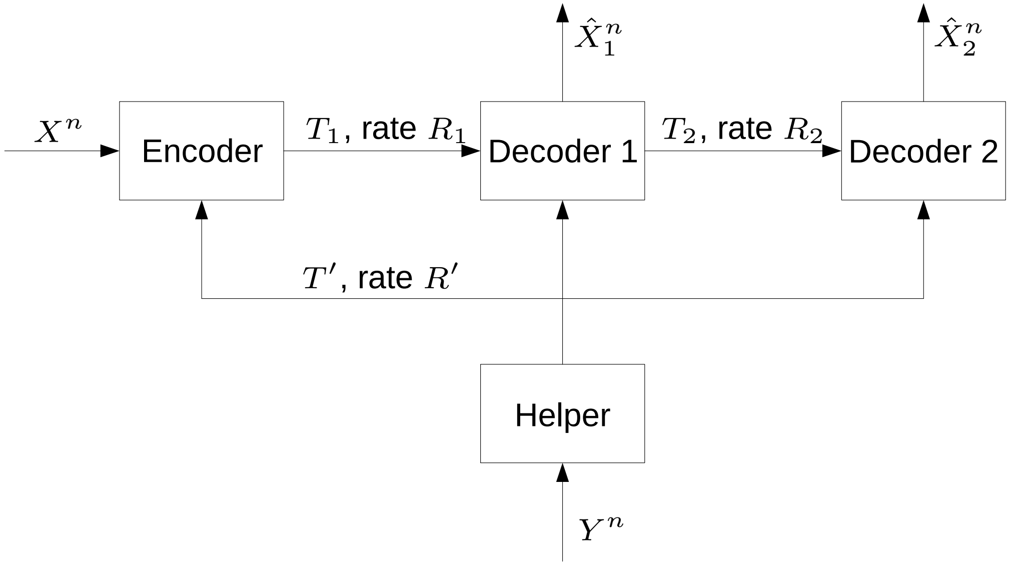
Theorem 5 $(R', R_1, R_2) \in \mathcal{R}_c(D_1, D_2)$ if and only if

$$\begin{aligned} R' &\geq I(Y;V) & V - Y - X \\ R_1 &\geq I(X;\hat{X}_1,\hat{X}_2|V) & Y - (V,X) - (\hat{X}_1,\hat{X}_2) \\ R_2 &\geq I(X;\hat{X}_2|V) & \\ D_i &\geq \mathbb{E}d(X,\hat{X}_i), \quad i=1,2 & \end{aligned}$$

for some joint distribution of the form

$$P_{X,Y} \cdot P_{V|Y} \cdot P_{U|Y} \cdot P_{\hat{X}_1,\hat{X}_2|V,X}.$$

Main Result



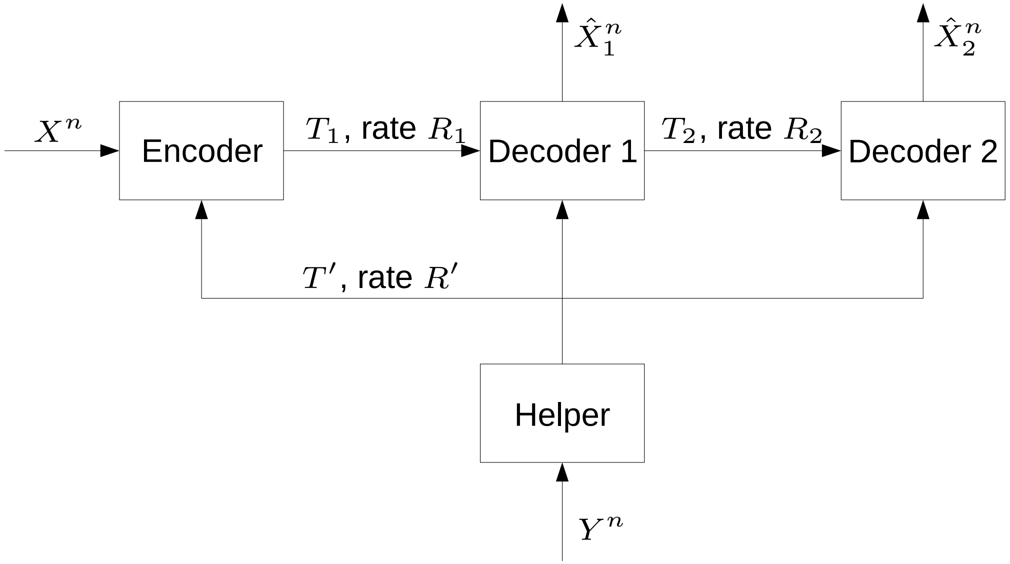
Theorem 5 $(R', R_1, R_2) \in \mathcal{R}_c(D_1, D_2)$ if and only if

$$\begin{aligned} R' &\geq I(Y;V) & & V - Y - X \\ R_1 &\geq I(X;\hat{X}_1,\hat{X}_2|V) & & Y - (V,X) - (\hat{X}_1,\hat{X}_2) \\ R_2 &\geq I(X;\hat{X}_2|V) & & \\ D_i &\geq \mathbb{E}d(X,\hat{X}_i), \quad i=1,2 & & \end{aligned}$$

for some joint distribution of the form

$$P_{X,Y} \cdot P_{V|Y} \cdot P_{U|Y} \cdot P_{\hat{X}_1,\hat{X}_2|V,X,Y}.$$

Main Result



Theorem 5 $(R', R_1, R_2) \in \mathcal{R}_c(D_1, D_2)$ if and only if

$$\begin{aligned} R' &\geq I(Y;V) \qquad\qquad V - Y - X \\ R_1 &\geq I(X;\hat{X}_1,\hat{X}_2|V) \\ R_2 &\geq I(X;\hat{X}_2|V) \\ D_i &\geq \mathbb{E}d(X,\hat{X}_i), \quad i=1,2 \end{aligned}$$

for some joint distribution of the form

$$P_{X,Y} \cdot P_{V|Y} \cdot P_{U|Y} \cdot P_{\hat{X}_1,\hat{X}_2|V,X,Y}.$$

Summary

- Solved the problem of rate-distortion with helper and SI at the decoder, for the case of common helper message.
- Implies a solution to the case of independent helper messages, with $R_1 \geq R_2$.
- Can reduce slightly the helper rate to the encoder (R_1), and get the same performance as with $R_1 \geq R_2$ (but this depends on external random variables).
- Solved the model of cascade rate-distortion with a helper. Here helper must send common messages; cannot play with rates.
- Have also a solution for the two way problem with a helper.

Thank You!