Information Rates for Phase Noise Channels

Luca Barletta
Politecnico di Milano
luca.barletta@polimi.it

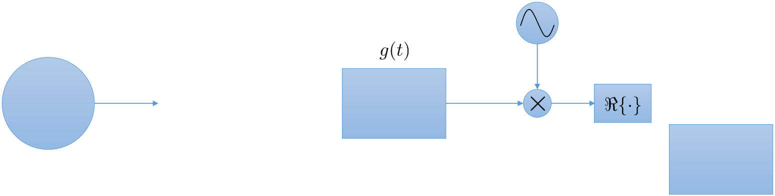
ESIT 2019
Antibes, July 15th - 19th, 2019

Outline

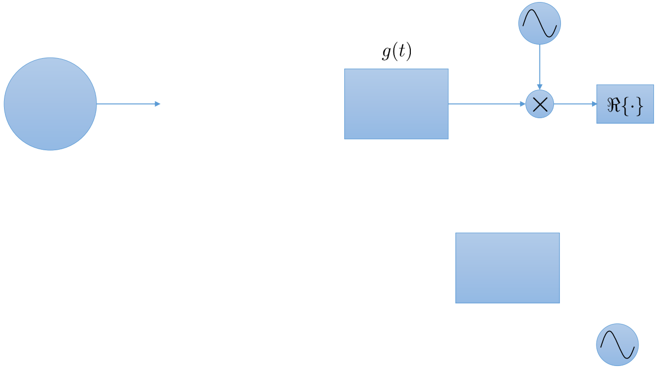
Outline

1. Motivation
2. From continuous to discrete time
3.

A Classic Communication Scheme



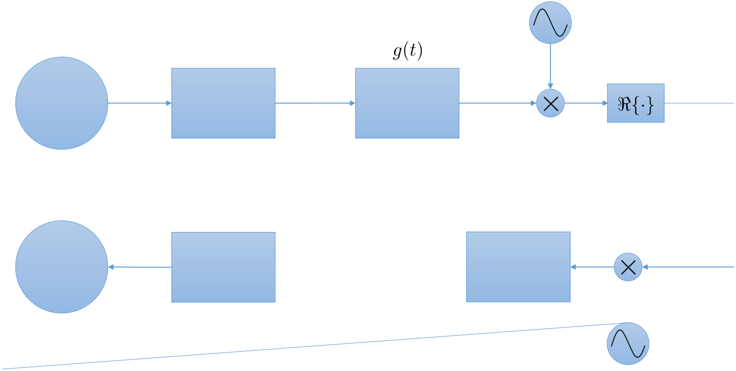
The AWGN Channel



$$Y_k = X_k + W_k$$

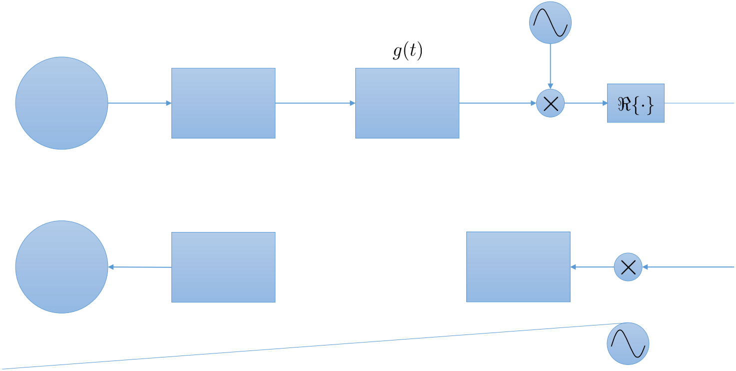
The Actual

The Actual AWGN Channel



Phase noise processes tx and rx

The Actual AWGN Channel



Phase noise processes $_{\text{tx}}$ and $_{\text{rx}}$
Sampling at kT_{symb} is no longer optimal!

Phase Noise Channels

Phase noise arises due to

- Imperfections in the oscillator circuits at the transceivers

Phase Noise Channels

Phase noise arises due to

- Imperfections in the oscillator circuits at the transceivers
  - Even for high-quality oscillators: if the continuous-time waveform is processed by long lters at the receiver (e.g., long symbol time, OFDM systems), the phase uncertainty accumulates!

Phase Noise Channels

Phase noise arises due to

Phase Noise Channels

Questions:

- Models for phase noise channels

Phase Noise Channels

Questions:

-

Phase Noise Channels

Questions:

- Models for phase noise channels
- Impact of phase noise on channel capacity
- Signal and code design

Outline

Representation of continuous-time waveforms

- Assume input waveform $\{X(t)\}_{t=0}^{T}$ is square integrable:

$$X(t) \in L^2[0;T] \quad \rightarrow \quad \int_0^T |X(t)|^2 dt < \infty$$

Representation of continuous-time waveforms

- Assume input waveform $\{X(t)\}_t^T$

Mutual information for random waveforms

- The average mutual information between $\{X(t)\}_{t=0}^{T}$ and $\{Y(t)\}_{t=0}^{T}$ is [Gallager, 1968]

$$I\left(\{X(t)\}_{t=0}^{T}; \{Y(t)\}_{t=0}^{T}\right) = \lim_{n \to \infty} I(X_1 \ldots X_n; Y_1 \ldots Y_n)$$

(if it exists)

Mutual information for random waveforms

- The average mutual information between $\{X(t)\}_{t=0}^{T}$ and $\{Y(t)\}_{t=0}^{T}$ is [Gallager, 1968]

$$I\left(\{X(t)\}_{t=0}^{T};\{Y(t)\}_{t=0}^{T}\right)=\lim_{n\to\infty} I(X_1\ldots X_n;Y_1\ldots Y_n)$$

(if it exists)

Example: mutual information for a phase noise channel

- Input-output relation:

$$Y(t) = X(t)e^{j\quad(t)} + W(t)$$

Example: mutual information for a phase noise channel

- Input-output relation:

$$Y(t) = X(t)e^{j\quad(t)} + W(t)$$

- Choose an incomplete orthonormal basis as

$$_m(t) = \frac{1}{\sqrt{\quad}}\mathrm{rect}\left(\frac{t \quad m \quad + \quad /2}{\quad}\right); \qquad m = 1,\ldots,n; \qquad = T/n$$

Example: mutual information for a phase noise channel

- Discretized model:

$$Y_m = X_m \int_{(m\ \ 1)}^{m} \frac{e^{j\ \ (t)}}{\ } dt + W_m$$

Example: mutual information for a phase noise channel

-

Example: mutual information for a phase noise channel

- Discretized model:

$$Y_m = X_m \int_{(m \quad 1)}^{m} \frac{e^{j \quad (t)}}{\qquad} dt + W_m$$

- Both amplitude fading and phase noise!
- Commonly used discrete-time phase noise channel model:

$$Y_m = X_m e^{j \quad m} + W_m; \qquad {}_m = \quad ((m \quad 1) \quad)$$

Example: mutual information for a phase noise channel

- Discretized model:

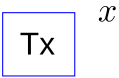
$$Y_m = X_m \int_{(m \quad 1)}^{m} \frac{e^{j \quad (t)}}{} dt + W_m$$

- Both amplitude fading and phase noise!
- Commonly used discrete-time phase noise channel model:

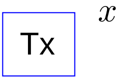
$$Y_m = X_m e^{j \quad m} + W_m; \qquad {}_m = \quad ((m \quad 1) \quad)$$

- How di erent are the two models in terms of capacity?

Receivers with nite time resolution

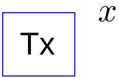


Receivers with nite time resolution



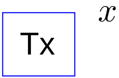
- How small should be?

Receivers with nite time resolution



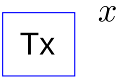
- How small should be?
  - It depends on the statistics of $f \quad (t) g$

Receivers with nite time resolution



- How small should be?
  - It depends on the statistics of $f \quad (t) g$
  - In general, as small as possible
- Oversampling helps! $\& \quad =) \quad I(X_1 \quad X_n ; Y_1 \quad Y$

Receivers with nite time resolution



- How small should be?
  - It depends on the statistics of $f \quad (t)g$
  - In general, as small as possible
- Oversampling helps! $\& \quad =) \; I(X_1 \quad X_n; Y_1 \quad Y_n) \%$

Oversampled channel model

$$Y_{mL+\ell} = X_m F_{mL+\ell} + W_m; \quad \ell = 1;\dots;L$$
$$m = 1;\dots;n$$

$$F_{mL+\ell} = \int_{(mL+\ell-1)\Delta}^{(mL+\ell)\Delta} \frac{e^{j\theta(t)}}{\Delta} dt; \qquad \Delta = \frac{T_{\text{symb}}}{L}$$

Wiener phase noise

- Widely used for modeling oscillators
- Also known as Brownian motion:

$$(t) = \quad (0) + \int_0^t Z(t')dt'$$

where Z is a white Gaussian process

Wiener phase noise

- Widely used for modeling oscillators
- Also known as Brownian motion:

$$(t) = \quad (0) + \int_0^t Z(t')dt'$$

where Z is a white Gaussian process

- Variance increases: $(t) \quad N(0; \;^2 t)$

Wiener phase noise

- Widely used for modeling oscillators
- Also known as Brownian motion:

$$(t) = (0) + \int_0^t Z(t')dt'$$

where Z is a white Gaussian process

- Variance increases: $(t) \quad N(0; {}^2 t)$
- Samples are not independent: $E[\ (t)\ (s)] = {}^2 \min\{t; s\}$

Wiener phase noise

- Widely used for modeling oscillators
- Also known as Brownian motion:

$$(t) = \quad (0) + \int_0^t Z(t')dt'$$

  where Z is a white Gaussian process
- Variance increases: $(t) \quad N(0; \; ^2 t)$
- Samples are not independent: $\mathrm{E}[\;(t)\;(s)] = \; ^2 \min\{t; s\}$
- Process with memory!
- Oversampling, i.e. $L > 1$, increases information rates

Wiener phase noise channel

De ne $_k =$ $((k$ $1)$ $)$ and N_k $N(0;1)$:

$$_k = \;_{k\;\;1} + \; p\overline{}\, N_k$$

Wiener phase noise channel

De ne $_k = \ ((k \quad 1) \quad)$ and $N_k \quad N(0;1)$:

$$ {}_k = \quad {}_{k \quad 1} + \sqrt{\ }\, N_k $$

$$ Y_{mL+\ell} = X_m e^{j \quad {}_{mL+\ell}} \int_{(mL+\ell \quad 1)}^{(mL+\ell)} \frac{e^{j(\quad (t) \quad {}_{mL+\ell})}}{\ } \mathrm{d}t + W_{mL+\ell} $$

Wiener phase noise channel

De ne $_k$ = ((k 1)) and N_k $N(0; 1)$:

$$_k = _{k \; 1} + p\text{—} N_k$$

$$Y_{mL+`} = X_m e^{j \; mL+`} \int_{(mL+` \; 1)}^{(mL+`)} \frac{e^{j(\; (t) \; mL+`)}}{} \mathrm{d}t + W_{mL+`}$$

Contour plot of the unnormalized fading pdf for = 6 and = 1. (Y. Wang *et al.*, TCOM 2006, vol. 54, no. 5)

Oversampled Discrete-time Wiener Phase Noise Channel

- Let us study the capacity of a simpler model

$$ {}_k = {}_{k\ \ 1} + p\text{—}N_k $$

$$ Y_{mL} $$

Oversampled Discrete-time Wiener Phase Noise Channel

- Let us study the capacity of a simpler model

$$\theta_k = \theta_{k-1} + \sqrt{\gamma} N_k$$

$$Y_{mL+\ell} = X_m e^{j\theta_{mL+\ell}} + W_{mL+\ell}$$

- Capacity under an average power constraint:

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T |X(t)|^2 \, dt \le P$$

Oversampled Discrete-time Wiener Phase Noise Channel

- Let us study the capacity of a simpler model

$$_k = \ _{k\ 1} + \sqrt{}\, N_k$$

$$Y_{mL+\ell} = X_m e^{j\ \ _{mL+\ell}} + W_{mL+\ell}$$

- Capacity under an average power constraint:

$$\mathbb{E}\ \frac{1}{T}\int_0^T |X(t)|^2\,\mathrm{d}t \quad P \quad \Longrightarrow \quad \mathbb{E}\ |X_m|^2 \quad P$$

How to compute capacity

Challenges:

- Channel with memory

How to compute capacity

Challenges:

- Channel with memory
- How to deal with oversampling?

How to compute capacity

Challenges:

- Channel with memory
- How to deal with oversampling?
- Unknown optimal input distribution

How to compute capacity

Challenges:

- Channel with memory

How to compute capacity

Challenges:

- Channel with memory
- How to deal with oversampling?
- Unknown optimal input distribution

We will see how to:

- Get rid of the memory
- Compute an upper bound to capacity

How to compute capacity

Challenges:

- Channel with memory
- How to deal with oversampling?
- Unknown optimal input distribution

We will see how to:

- Get rid of the memory
- Compute an upper bound to capacity
- We assume iid X_m's with $\underline{\angle X_m} \quad U[0;2\ )$

A capacity upper bound

Define $X_1^n = X_1 \ldots X_n$

$$\frac{1}{n} I\left(X_1^n ; \mathbf{Y}_1^n\right) = \frac{1}{n} \sum_{m=1}^{n} I\left(X_1^n ; \mathbf{Y}_m \mid \mathbf{Y}_1^{m-1}\right)$$

A capacity upper bound

De ne $X_1^n = X_1 \quad X_n$

$$\frac{1}{n} I\,(X_1^n\,;\mathbf{Y}_1^n) = \frac{1}{n}\sum_{m=1}^{n} I \quad X_1^n;\mathbf{Y}_m\, j\,\mathbf{Y}_1^{m \quad 1}$$

$$\text{(DPI)} \qquad \frac{1}{n}\sum_{m=1}^{n} I \quad X_1^n; \quad {}_{mL+1};\mathbf{Y}_m\, j\,\mathbf{Y}_1^{m \quad 1}$$

A capacity upper bound

De ne $X_1^n = X_1 \quad X_n$

$$\frac{1}{n} I\left(X_1^n; \mathbf{Y}_1^n\right) = \frac{1}{n} \sum_{m=1}^{n} I\left(X_1^n; \mathbf{Y}_m \,j\, \mathbf{Y}_1^{m-1}\right)$$

$$\text{(DPI)} \quad \frac{1}{n} \sum_{m=1}^{n} I\left(X_1^n; {}_{mL+1}; \mathbf{Y}_m \,j\, \mathbf{Y}_1^{m-1}\right)$$

$$= \frac{1}{n} \sum_{m=1}^{n} I\left(X_1^n; \mathbf{Y}_m \,j\, {}_{mL+1}; \mathbf{Y}_1^{m-1}\right) + I\left({}_{mL+1}; \mathbf{Y}_m \,j\, \mathbf{Y}_1^{m-1}\right)$$

A capacity upper bound

De ne $X_1^n = X_1 \quad X_n$

$$\frac{1}{n} I\left(X_1^n; \mathbf{Y}_1^n\right) = \frac{1}{n}\sum_{m=1}^{n} I \quad X_1^n; \mathbf{Y}_m \,|\, \mathbf{Y}_1^{m \quad 1}$$

$$\text{(DPI)} \quad \frac{1}{n}\sum_{m=1}^{n} I \quad X_1^n; \quad {}_{mL+1}; \mathbf{Y}_m \,|\, \mathbf{Y}_1^{m \quad 1}$$

$$= \frac{1}{n}\sum_{m=1}^{n} I \quad X_1^n; \mathbf{Y}_m \,|\, \quad {}_{mL+1}; \mathbf{Y}_1^{m \quad 1} \quad + I \quad {}_{mL+1}; \mathbf{Y}_m \,|\, \mathbf{Y}_1^{m \quad 1}$$

$$\text{(Markov)} = \frac{1}{n}\sum_{m=1}^{n} I\left(X_m; \mathbf{Y}_m \,|\, \quad {}_{mL+1}\right) + I \quad {}_{mL+1}; \mathbf{Y}_m \,|\, \mathbf{Y}_1^{m \quad 1}$$

A capacity upper bound

De ne $X_1^n = X_1 \quad\quad X_n$

$$\frac{1}{n} I(X_1^n;\mathbf{Y}_1^n) = \frac{1}{n}\sum_{m=1}^{n} I \quad X_1^n;\mathbf{Y}_m \,|\, \mathbf{Y}_1^{m \quad 1}$$

$$\text{(DPI)} \quad \frac{1}{n}\sum_{m=1}^{n} I \quad X_1^n; \quad _{mL+1};\mathbf{Y}_m \,|\, \mathbf{Y}_1^{m \quad 1}$$

$$= \frac{1}{n}\sum_{m=1}^{n} I \quad X_1^n;\mathbf{Y}_m \,|\, \quad _{mL+1};\mathbf{Y}_1^{m \quad 1} \quad + I \quad\quad _{mL+1};\mathbf{Y}_m \,|\, \mathbf{Y}_1^{m \quad 1}$$

$$\text{(Markov)} = \frac{1}{n}\sum_{m=1}^{n} I(X_m;\mathbf{Y}_m \,|\, \quad _{mL+1}) + I \quad\quad _{mL+1};\mathbf{Y}_m \,|\, \mathbf{Y}_1^{m \quad 1}$$

$$\text{(Stationary)} = I(X_1;\mathbf{Y}_1 \,|\, \quad _{L+1}) + \frac{1}{n}\sum_{m=1}^{n} I \quad\quad _{mL+1};\mathbf{Y}_m \,|\, \mathbf{Y}_1^{m \quad 1}$$

A capacity upper bound

Assumption: iid X_m's with $\underline{\angle X_m} \quad U[0; 2\ )$

$$I \quad {}_{mL+1}; \mathbf{Y}_m j \, \mathbf{Y}_1^{m \quad 1}$$

$$\text{(DPI)} \quad I \quad {}_{mL+1}; \quad {}_{mL+1} \quad \underline{\angle X_m} \quad p\,\overline{}\, N_{mL+1} \quad \mathbf{Y}_1^{m \quad 1}$$

$$= 0$$

Polar decomposition

$$
\begin{aligned}
I(X_1; \mathbf{Y}_1 | \;_{L+1}) &= I(|X_1|; \angle X_1; \mathbf{Y}_1 | \;_{L+1}) \\
&= \underbrace{I(|X_1|; \mathbf{Y}_1 | \;_{L+1})}_{\text{amplitude mod.}} + \underbrace{I(\angle X_1; \mathbf{Y}_1 | \;_{L+1}; |X_1|)}_{\text{phase mod.}}
\end{aligned}
$$

Polar decomposition

$$
\begin{aligned}
I(X_1; \mathbf{Y}_1 | {}_{L+1}) &= I(|X_1|, \angle X_1; \mathbf{Y}_1 | {}_{L+1}) \\
&= \underbrace{I(|X_1|; \mathbf{Y}_1 | {}_{L+1})}_{\text{amplitude mod.}} + \underbrace{I(\angle X_1; \mathbf{Y}_1 | {}_{L+1}, |X_1|)}_{\text{phase mod.}}
\end{aligned}
$$

- Different bounding techniques are used for the two terms

Polar decomposition

$$I(X_1; \mathbf{Y}_1 | \;\; {}_{L+1}) = I(|X_1|; \angle X_1; \mathbf{Y}_1 | \;\; {}_{L+1})$$

$$= \underbrace{I(|X_1|; \mathbf{Y}_1 | \;\; {}_{L+1})}_{\text{amplitude mod.}} + \underbrace{I(\angle X_1; \mathbf{Y}_1 | \;\; {}_{L+1}; |X_1|)}_{\text{phase mod.}}$$

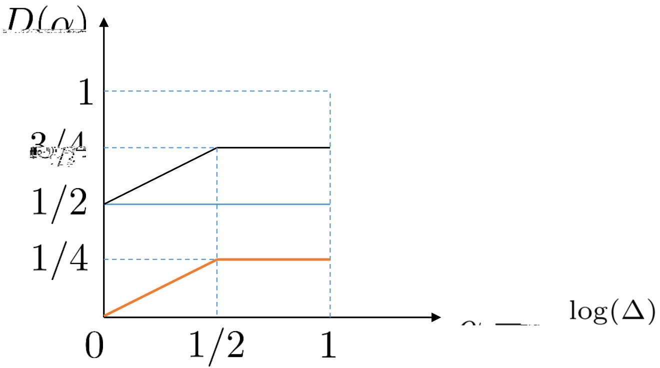
- Different bounding techniques are used for the two terms
  - Reveal all phase noise samples to the receiver: amplitude mod. on AWGN channel
  - Application of the I-MMSE formula to the phase mod. term

$$C \quad \frac{1}{2}\log\left(1+\frac{P}{2}\right) + \log(2\;) + \frac{1}{2}\log\left(P\left(\sqrt{1+\frac{4}{{}^2P\;{}^2}}\quad 1\right)\right)$$

Degrees of freedom

Define the degrees of freedom as

$$D(\;) = \lim_{P \to \infty} \frac{C(P;\quad = P\quad;\;)}{\log(P)}$$



Conclusions

- The capacity pre-log depends on the growth rate of receiver's time resolution

Conclusions

- The capacity pre-log depends on the growth rate of receiver's time resolution
- If $1/\sqrt{P}$, an asymptotic capacity pre-log of 0.75 can be achieved, but not surpassed

Conclusions

- The capacity pre-log depends on the growth rate of receiver's time resolution
- If $1/\sqrt{P}$, an asymptotic capacity pre-log of 0.75 can be achieved, but not surpassed
- Even with infinite time resolution $C_{\mathrm{PN}} < C_{\mathrm{AWGN}}$

Conclusions

- The capacity pre-log depends on the growth rate of receiver's time resolution
- If $\quad 1/\sqrt{P}$, an asymptotic capacity pre-log of 0.75 can be achieved, but not surpassed
- Even with infinite time resolution $C_{\text{PN}} < C_{\text{AWGN}}$
  - For any $\quad > 0$! Even for very high-quality oscillators!

Conclusions

- The capacity pre-log depends on the growth rate of receiver's time resolution
- If $1/\sqrt{P}$, an asymptotic capacity pre-log of 0.75 can be achieved, but not surpassed
- Even with infinite time resolution $C_{\mathrm{PN}} < C_{\mathrm{AWGN}}$
  - For any > 0! Even for very high-quality oscillators!
- Conjectures:

Conclusions

- The capacity pre-log depends on the growth rate of receiver's time resolution
- If $1/\sqrt{P}$, an asymptotic capacity pre-log of 0.75 can be achieved, but not surpassed
- Even with infinite time resolution $C_{\mathrm{PN}} < C_{\mathrm{AWGN}}$
  - For any > 0! Even for very high-quality oscillators!
- Conjectures:
  - Simplifying the model by discarding the amplitude fading is too much

Conclusions

- The capacity pre-log depends on the growth rate of receiver's time resolution
- If $1/\sqrt{P}$, an asymptotic capacity pre-log of 0.75 can be achieved, but not surpassed
- Even with in nite time resolution $C_{\mathrm{PN}} < C_{\mathrm{AWGN}}$
    - For any > 0! Even for very high-quality oscillators!
- Conjectures:
    - Simplifying the model by discarding the amplitude fading is too much
    - The fundamental tension between additive noise and phase noise limits the degrees of freedom

Thanks for your attention!

Derivation of the average power constraint

$$\lim_{T \to \infty} \frac{1}{T}$$