Control over Gaussian Channels With and Without Source–Channel Separation

Elias Riedel Gårding

Royal Institute of Technology (KTH), Stockholm, Sweden

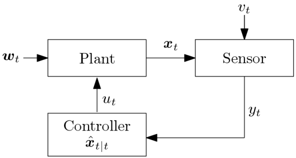
Supervisors: Victoria Kostina and Anatoly Khina
Electrical Engineering, Caltech

August 24, 2017

Control theory in a nutshell

Stabilize an unstable system (“plant”) using measurements in a feedback loop.

Traditional control



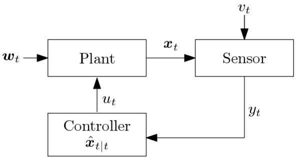
- Optimal strategies known

Elias Riedel Gårding

Control theory in a nutshell

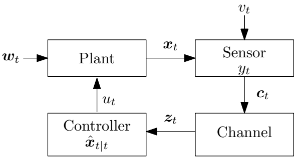
Stabilize an unstable system ("plant") using measurements in a feedback loop.

Traditional control



- Optimal strategies known
- Sensor and controller co-located

Networked control



- Channel limits communication
- Requires information theory

Networked control

Applications

Real-time wireless systems

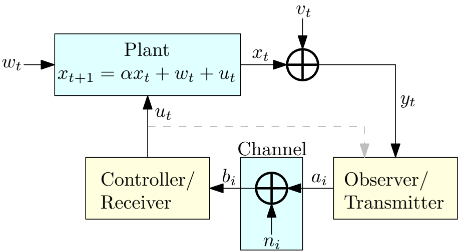
- Autonomous vehicles
- Remote surgery
- ...

http://www.bbc.com/future/story/20140516-i-operate-on-people-400km-away

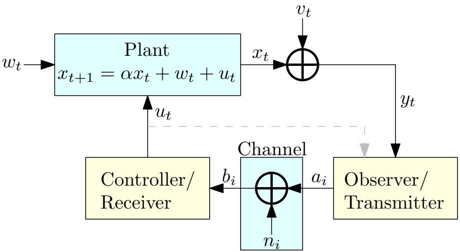
http://latam.pcmag.com/drones/1774/review/yuneec-typhoon-h-pro

Networked control

Problem setup



- Unstable if $|\alpha| > 1$
- Goal: Minimize $\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[x_t^2]$



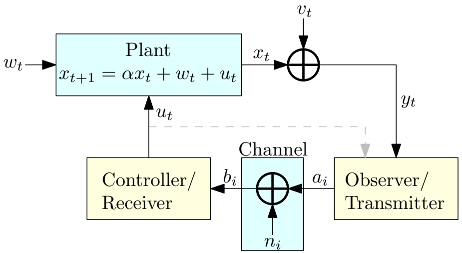
- Unstable if $|\alpha| > 1$
- Goal: Minimize $\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[x_t^2]$

Channel model

AWGN (**A**dditive **W**hite **G**aussian **N**oise)

Power constraint: $\mathbb{E}[a_i^2] \leq 1$

Networked control
Problem setup



- Unstable if $|\alpha| > 1$
- Goal: Minimize $\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[x_t^2\right]$

Channel model

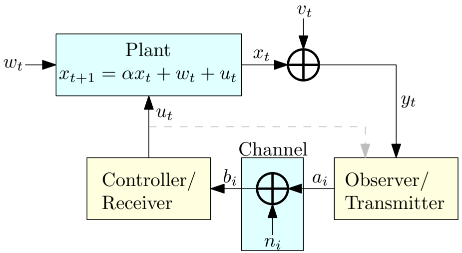
AWGN (**A**dditive **W**hite **G**aussian **N**oise)

Power constraint: $\mathbb{E}[a_i^2] \leq 1$

Signaling rate versus sampling rate

We may use the channel K times per time step.

We focus on $K = 1$ and $K = 2$.



- Unstable if $|\alpha| > 1$
- Goal: Minimize $\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[x_t^2]$

Channel model

AWGN (**A**dditive **W**hite **G**aussian **N**oise)

Power constraint: $\mathbb{E}[a_i^2] \leq 1$

Signaling rate versus sampling rate

We may use the channel K times per time step.

We focus on $K = 1$ and $K = 2$.

How to encode/decode?

Solution approaches

JSCC versus separation

Source–channel separation

Split encoding and decoding into smaller subproblems.

- Standard approach in information theory: Bits!
- Optimal for long messages (but control needs short ones)
- No reason it should work well for control

Joint source–channel coding (JSCC)

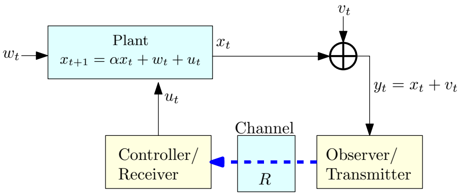
Design the encoder and decoder holistically.

- Much simpler and less computationally intensive
- Hypothesis: Gives much better control performance

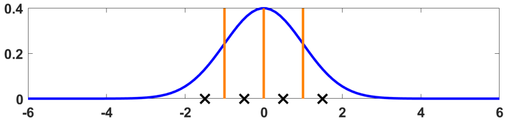
My task: Implement, simulate and compare

Solution approaches

Separation: source coding (quantization)

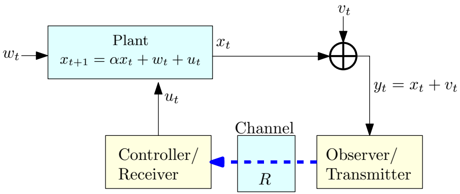


- Separation $\implies$ noiseless digital channel
- Encode the measurement as a fixed number of bits
- Quantization errors blow up
- Fixed quantizer won't work
- Optimal strategy known

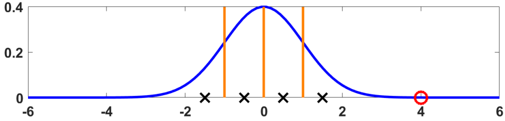


Solution approaches

Separation: source coding (quantization)



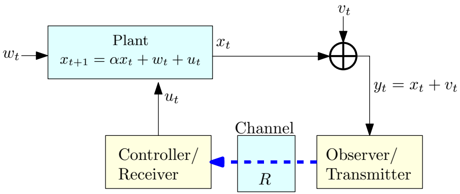
- Separation $\implies$ noiseless digital channel
- Encode the measurement as a fixed number of bits
- Quantization errors blow up
- Fixed quantizer won't work
- Optimal strategy known



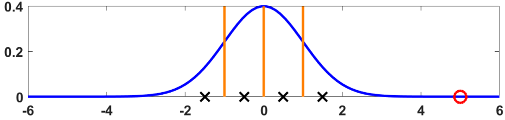
Caltech

Solution approaches

Separation: source coding (quantization)

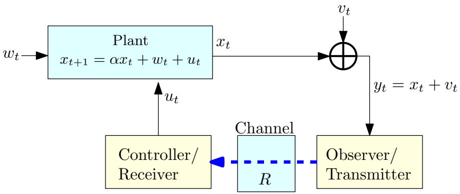


- Separation $\implies$ noiseless digital channel
- Encode the measurement as a fixed number of bits
- Quantization errors blow up
- Fixed quantizer won't work
- Optimal strategy known

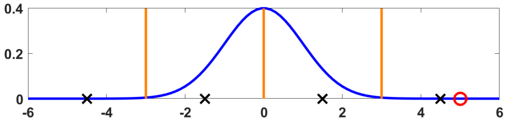


Solution approaches

Separation: source coding (quantization)



- Separation $\implies$ noiseless digital channel
- Encode the measurement as a fixed number of bits
- Quantization errors blow up
- Fixed quantizer won't work
- Optimal strategy known



Solution approaches

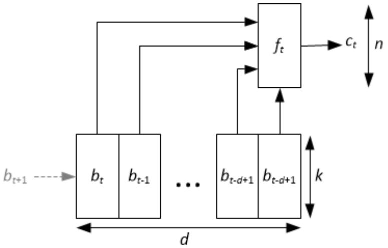
Separation: channel coding (tree codes)

Recall $x_{t+1} = \alpha x_t + w_t + u_t$

Anytime reliability

- Errors get magnified as α^t ($|\alpha| > 1$)
- Error probabilites must shrink as α^{-t}

- Each sent bit must depend on the entire input history
- Optimal decoding computationally infeasible
- Sequential decoding



Solution approaches

JSCC and spirals

For $K = 1$: **Optimal** encoding map $\mathbb{R} \to \mathbb{R}$ known: Transmit as is!

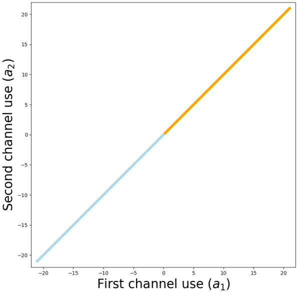
For $K = 1$: **Optimal** encoding map $\mathbb{R} \to \mathbb{R}$ known: Transmit as is!
Try the same for $K = 2$?

- Repetition: $\begin{cases} a_1 = s \\ a_2 = s \end{cases}$

Solution approaches

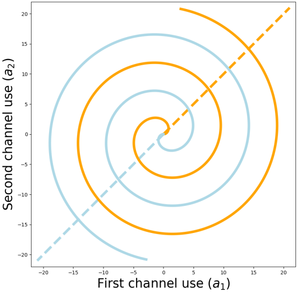
JSCC and spirals

For $K = 1$: **Optimal** encoding map $\mathbb{R} \to \mathbb{R}$ known: Transmit as is!
Try the same for $K = 2$?



- Repetition: $\begin{cases} a_1 = s \\ a_2 = s \end{cases}$
- Waste of space! Stretch?

For $K = 1$: **Optimal** encoding map $\mathbb{R} \to \mathbb{R}$ known: Transmit as is!
Try the same for $K = 2$?

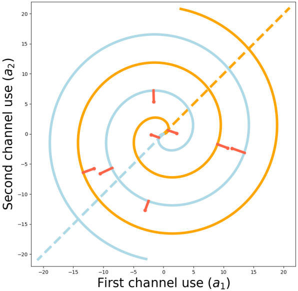


- Repetition: $\begin{cases} a_1 = s \\ a_2 = s \end{cases}$
- Waste of space! Stretch?
- Spiral: $\begin{cases} a_1 = \sqrt{2}s\cos(\omega|s|) \\ a_2 = \sqrt{2}s\sin(\omega|s|) \end{cases}$

Solution approaches

JSCC and spirals

For $K = 1$: **Optimal** encoding map $\mathbb{R} \to \mathbb{R}$ known: Transmit as is!
Try the same for $K = 2$?



- Repetition: $\begin{cases} a_1 = s \\ a_2 = s \end{cases}$
- Waste of space! Stretch?
- Spiral: $\begin{cases} a_1 = \sqrt{2}s\cos(\omega\lvert s\rvert) \\ a_2 = \sqrt{2}s\sin(\omega\lvert s\rvert) \end{cases}$
- Much better!
- Tightness–crossover tradeoff

Implementation

Python 3 with NumPy + SciPy, ∼1400 lines of code

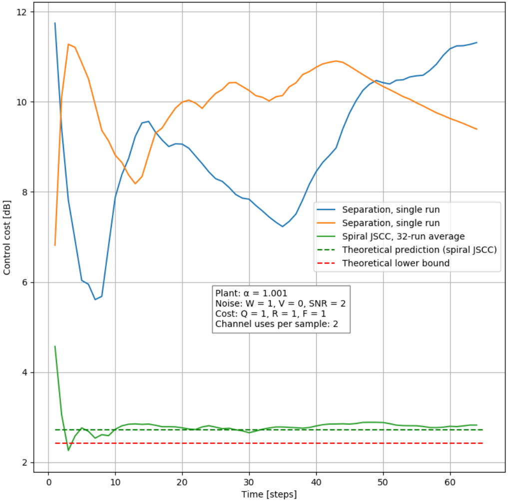
`https://github.com/eliasrg/SURF2017`

Side benefit: Implementations of

- Spiral encoder and maximum-likelihood (ML) decoder
- The Lloyd–Max quantization algorithm
- Tree/convolutional codes, especially the "stack" sequential decoding algorithm

Most important lesson learned: `git cherry-pick -n` and `git revert -n`

Results


Control cost [dB]
Time [steps]
12
10
8
6
4
2
0
10
20
30
40
50
60
Separation, single run
Separation, single run
Spiral JSCC, 32-run average
Theoretical prediction (spiral JSCC)
Theoretical lower bound
Plant: α = 1.001
Noise: W = 1, V = 0, SNR = 2
Cost: Q = 1, R = 1, F = 1
Channel uses per sample: 2

Further directions

Remaining work

- Adapt separation scheme for observation noise
- Send more than one bit per channel use using pulse amplitude modulation (PAM)
- Improve numerical stability

Further exploration

- Examine JSCC maps other than the spiral for $K = 2$
- Develop JSCC maps for $K = 3, 4, 5, 6, \ldots$

Thanks to

Victoria Kostina (Caltech)

Anatoly Khina (Caltech)

Caltech SFP and SURF

Henrik Sandberg (KTH)

Video source: Andreas Eder, Tobias Glück, TU Vienna ACIN CDS, `http://www.acin.tuwien.ac.at/` (`https://youtu.be/cyN-CRNrb3E`)