Bayesian Inference with Nonlinear Generative Models

Random Matrices and Random Landscapes

Ali Bereyhi[†], Bruno Loureiro[⋆], Florent Krzakala[⋆], Ralf R. Müller[†] and Hermann Schulz-Baldes[†]
[†]Friedrich-Alexander Universität Erlangen-Nürnberg and [⋆]École polytechnique fédérale de Lausanne
July 2022

Table of Contents

Start Point

Conclusions

Table of Contents

Table of Contents

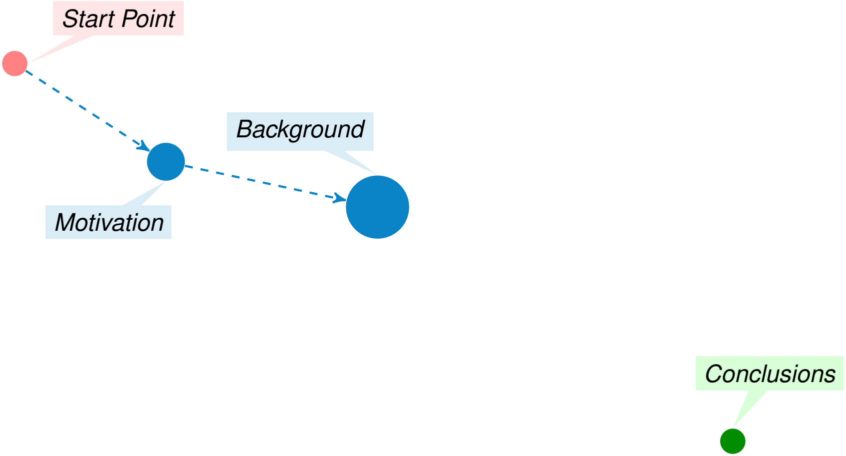
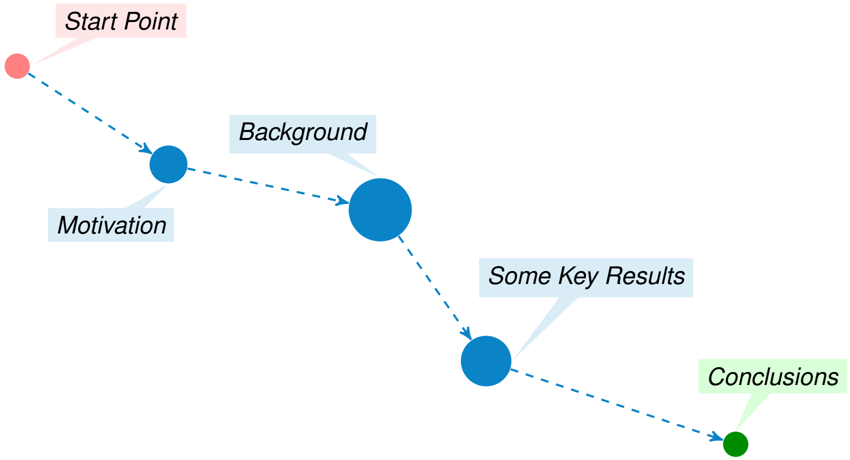


Table of Contents



Understating the Objective

What is the Core Problem?

$\boldsymbol{s} \in \mathbb{R}^D$ is mapped by $\mathcal{V} : \mathbb{R}^D \mapsto \mathbb{R}^N$

$$\boldsymbol{x} = \mathcal{V}\left(\boldsymbol{s}\right)$$

The mapping is then observed through a noisy channel

$$\boldsymbol{y} = \boldsymbol{x} + \mathbf{w}$$

What is the Core Problem?

$\boldsymbol{s} \in \mathbb{R}^D$ is mapped by $\mathcal{V} : \mathbb{R}^D \mapsto \mathbb{R}^N$

$$\boldsymbol{x} = \mathcal{V}(\boldsymbol{s})$$

The mapping is then observed through a noisy channel

$$\boldsymbol{y} = \boldsymbol{x} + \mathbf{w} \dashrightarrow \text{i.i.d. Gaussian vector}$$

What is the Core Problem?

$\boldsymbol{s} \in \mathbb{R}^D$ is mapped by $\mathcal{V} : \mathbb{R}^D \mapsto \mathbb{R}^N$

$$\boldsymbol{x} = \mathcal{V}(\boldsymbol{s})$$

The mapping is then observed through a noisy channel

$$\boldsymbol{y} = \boldsymbol{x} + \mathbf{w}$$ i.i.d. Gaussian vector

Ultimate Goal

Recover $\boldsymbol{s}$ from the noisy observations

Mapping via a Gaussian Field

Why should it be a new problem?

Mapping via a Gaussian Field

Why should it be a new problem? $\mathcal{V}(\cdot)$ is a Gaussian field

Mapping via a Gaussian Field

Why should it be a new problem? $\mathcal{V}(\cdot)$ is a Gaussian field

$$\boldsymbol{x} = \mathcal{V}(\boldsymbol{s}) = \begin{bmatrix} \mathcal{V}_1(\boldsymbol{s}) \\ \vdots \\ \mathcal{V}_N(\boldsymbol{s}) \end{bmatrix}$$

Gaussian

- The entries of $\boldsymbol{x}$ are independent zero-mean Gaussians conditioned to $\boldsymbol{s}$
- The covariance read

$$\mathbb{E}\left\{\mathcal{V}_n(\boldsymbol{s}_1)\,\mathcal{V}_n(\boldsymbol{s}_2)\right\} = \Phi\left(\frac{\langle \boldsymbol{s}_1; \boldsymbol{s}_2 \rangle}{K}\right)$$

Mapping via a Gaussian Field

Why should it be a new problem? $\mathcal{V}(\cdot)$ is a Gaussian field

$$\boldsymbol{x} = \mathcal{V}(\boldsymbol{s}) = \begin{bmatrix} \mathcal{V}_1(\boldsymbol{s}) \\ \vdots \\ \mathcal{V}_N(\boldsymbol{s}) \end{bmatrix}$$

Gaussian

- The entries of $\boldsymbol{x}$ are independent zero-mean Gaussians conditioned to $\boldsymbol{s}$
- The covariance read

$$\mathbb{E}\left\{\mathcal{V}_n(\boldsymbol{s}_1)\,\mathcal{V}_n(\boldsymbol{s}_2)\right\} = \Phi\left(\frac{\langle \boldsymbol{s}_1; \boldsymbol{s}_2 \rangle}{K}\right)$$

inner product

Mapping via a Gaussian Field

We are used to the *linear* model

Linear model is a Gaussian field of order one

$$\mathcal{V}_n\left(\boldsymbol{s}\right) = \langle \mathbf{a}_n ; \boldsymbol{s} \rangle$$

zero-mean Gaussian with variance $1/K$

whose covariance function is $\Phi\left(x\right) = x$

Mapping via a Gaussian Field

We are used to the *linear* model literature of information theory

Linear model is a Gaussian field of order one

$$\mathcal{V}_n(\boldsymbol{s}) = \langle \mathbf{a}_n ; \boldsymbol{s} \rangle$$ zero-mean Gaussian with variance $1/K$

whose covariance function is $\Phi(x) = x$

Mapping via a Gaussian Field

We are used to the *linear* model

Linear model is a Gaussian field of order one

$$\mathcal{V}_n(\boldsymbol{s}) = \langle \mathbf{a}_n ; \boldsymbol{s} \rangle$$

zero-mean Gaussian with variance $1/K$

whose covariance function is $\Phi(x) = x$

Mapping via a Gaussian Field

We are used to the *linear* model

Linear model is a Gaussian field of order one

$$\mathcal{V}_n(\boldsymbol{s}) = \langle \mathbf{a}_n; \boldsymbol{s} \rangle$$ zero-mean Gaussian with variance $1/K$

whose covariance function is $\Phi(x) = x$

But one can in general think of higher-order fields

Example: A purely quadratic field

$$\mathcal{V}_n(\boldsymbol{s}) = \langle \boldsymbol{s}; \mathbf{J}_n \boldsymbol{s} \rangle$$ zero-mean Gaussian with variance $1/K^2$

whose covariance function is $\Phi(x) = x^2$

Mapping via a Gaussian Field

Why should anyone care about a nonlinear model?

Mapping via a Gaussian Field

Why should anyone care about a nonlinear model?

- ✔ The evidence hidden between the lines of Yan Fyodorov's work says that

Nonlinear models have secrecy potentials

Journal of Statistical Physics (2019) 175:789–818
https://doi.org/10.1007/s10955-018-02217-9

CrossMark

A Spin Glass Model for Reconstructing Nonlinearly Encrypted Signals Corrupted by Noise

Yan V. Fyodorov[1]

Received: 12 August 2018 / Accepted: 11 December 2018 / Published online: 12 January 2019
© The Author(s) 2019

Mapping via a Gaussian Field

Why should anyone care about a nonlinear model?

✔ The evidence hidden between the lines of Yan Fyodorov's work says that

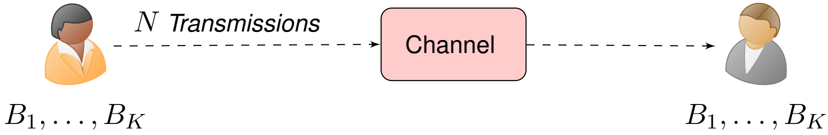
Nonlinear models have secrecy potentials

To understand this motivation, we need to take a quick look on

- Secure transmission over the wiretap channel
- Fyodorov's key observation

The Wiretap Channel

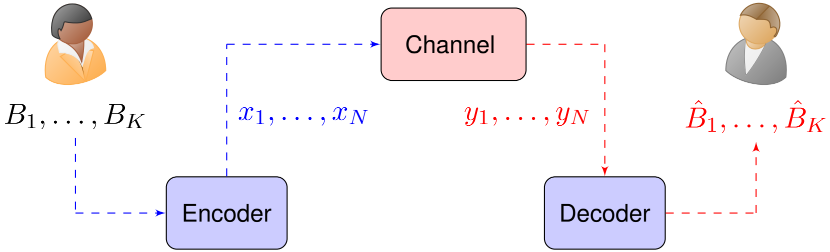
Review: Channel Coding



Transmission rate is

$$R = \frac{K}{N}$$

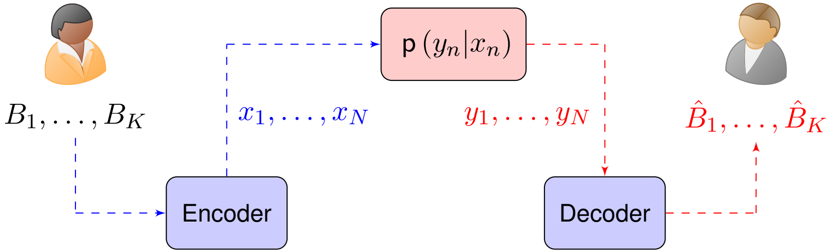
Review: Channel Coding



We desire to have reliability, i.e., with fixed rate $R = K/N$

$$\Pr\left\{(\hat{B}_1, \ldots, \hat{B}_K) \neq (B_1, \ldots, B_K)\right\} \to 0 \qquad \text{when } N \to \infty$$

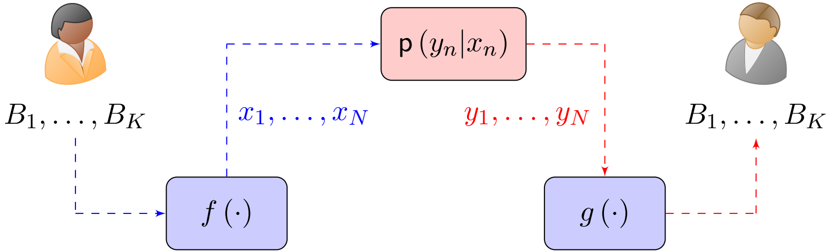
Review: Channel Coding



How we model the channel? By a conditional distribution

$$y_n \sim \mathsf{p}\left(y_n|x_n\right)$$

Review: Shannon's Answer (1948)

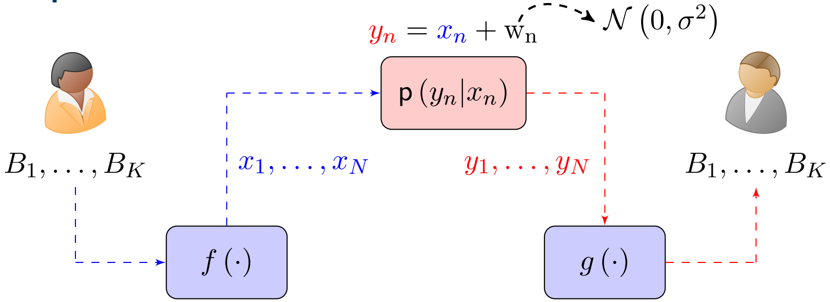


Channel Coding Theorem

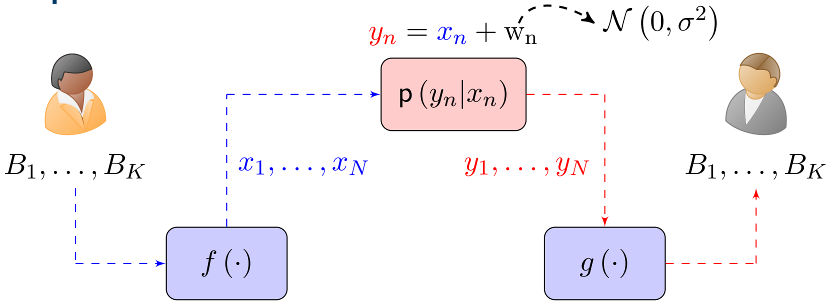
The maximum transmission rate for reliable communication is

$$C = \max_{p(x)} I(X;Y) = \max_{p(x)} \left[H(X) - H(X|Y) \right]$$

Example: Gaussian Channel



Example: Gaussian Channel

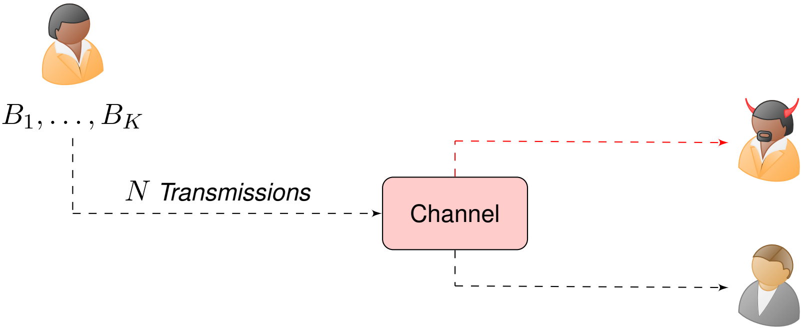


Channel Coding Theorem

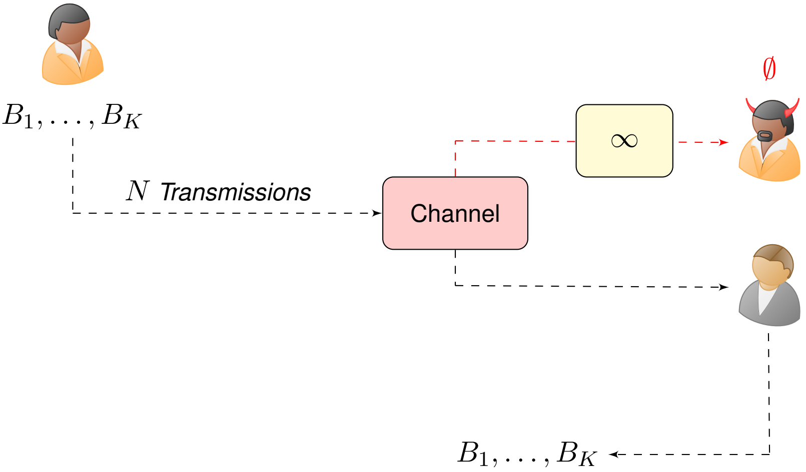
The capacity of the Gaussian channel is given by Gaussian input

$$C = \frac{1}{2} \log \left(1 + \frac{1}{\sigma^2} \right)$$

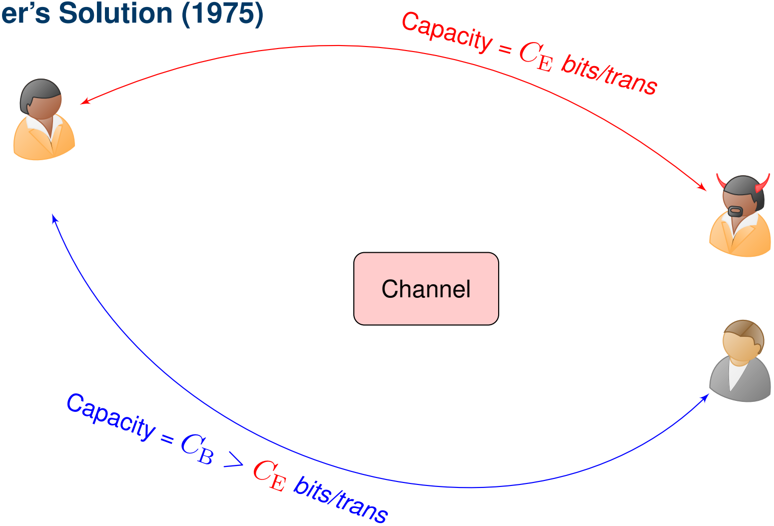
Secure Channel Coding



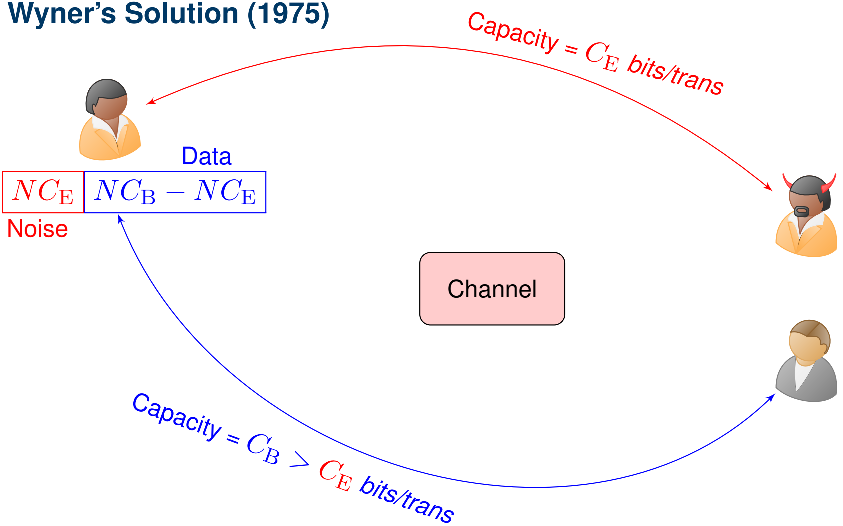
Secure Channel Coding



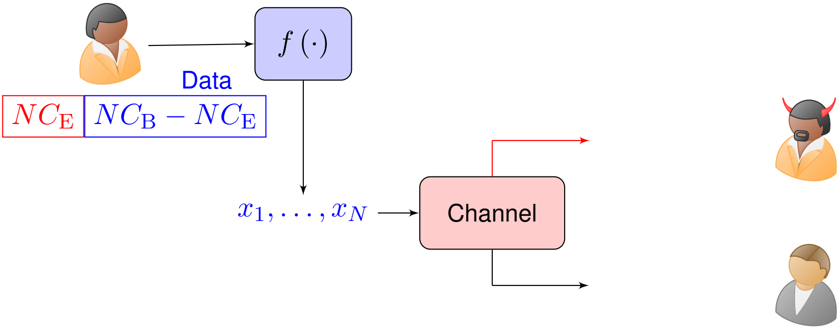
Wyner's Solution (1975)



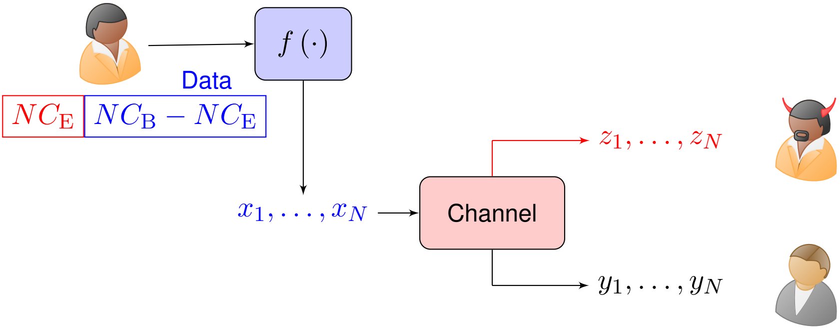
Wyner's Solution (1975)



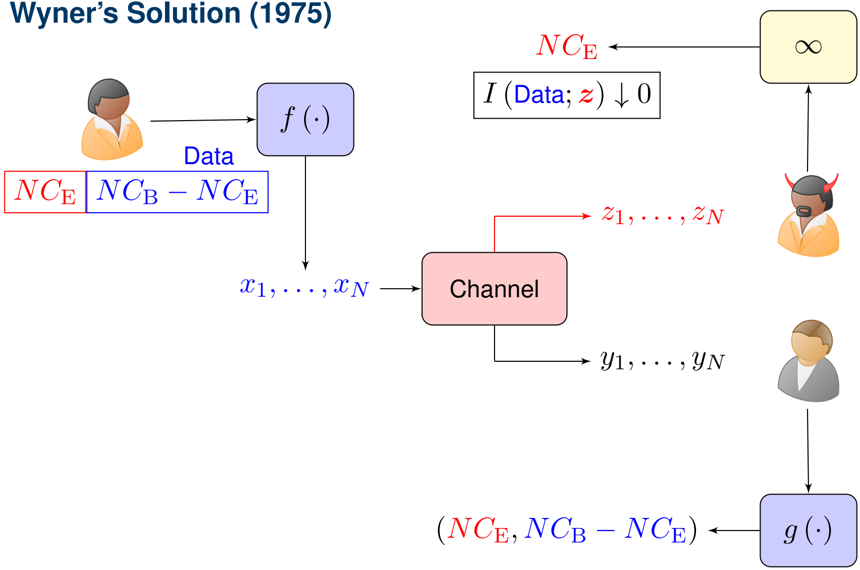
Wyner's Solution (1975)



Wyner's Solution (1975)



Wyner's Solution (1975)



Wyner's Solution (1975)

The *maximum secure communication rate* in the wiretap channel is

$$C_{\text{Secure}} = \max_{p(x)} \left[I\left(X;Y\right) - I\left(X;Z\right) \right]^{+}$$

Wyner's Solution (1975)

The *maximum secure communication rate* in the wiretap channel is

$$C_{\text{Secure}} = \max_{p(x)} \left[I\left(X;Y\right) - I\left(X;Z\right) \right]^{+}$$

Proof

The proof is given via random binning

Wyner's Solution (1975)

The *maximum secure communication rate* in the wiretap channel is

$$C_{\mathrm{Secure}} = \max_{p(x)} \left[I\left(X;Y\right) - I\left(X;Z\right) \right]^{+}$$

Proof

The proof is given via random binning

Gaussian Wiretap Channel (Leung and Hellman 1978)

The capacity of a Gaussian wiretap channel is achieved by a Gaussian input

$$C_{\mathrm{Secure}} = \frac{1}{2} \left[\log\left(1 + \frac{1}{\sigma_{\mathrm{B}}^2}\right) - \log\left(1 + \frac{1}{\sigma_{\mathrm{E}}^2}\right) \right]^{+}$$

Wyner's Solution (1975)

The *maximum secure communication rate* in the wiretap channel is

$$C_{\text{Secure}} = \max_{p(x)} \left[I\left(X;Y\right) - I\left(X;Z\right) \right]^{+}$$

Proof

The proof is given via random binning

Gaussian Wiretap Channel (Leung and Hellman 1978)

The capacity of a Gaussian wiretap channel is achieved by a Gaussian input

$$C_{\text{Secure}} = \frac{1}{2}\left[\log\left(1+\frac{1}{\sigma_{\mathrm{B}}^2}\right) - \log\left(1+\frac{1}{\sigma_{\mathrm{E}}^2}\right) \right]^{+}$$

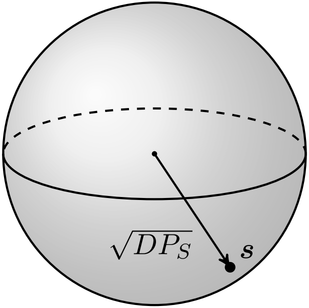
noise variance to Bob

noise variance to Eve

Earlier Result by Fyodorov

Nonlinear Encryption by Gaussian Fields

Fyodorov uses the nonlinear model to encrypt data on a *hypersphere*

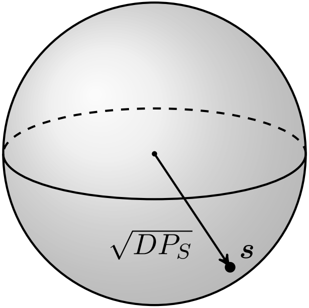


- $\boldsymbol{s}$ is uniform on a D-dimensional hypersphere

$$\|\boldsymbol{s}\|^2 = DP_S$$

Nonlinear Encryption by Gaussian Fields

Fyodorov uses the nonlinear model to encrypt data on a *hypersphere*



- $\boldsymbol{s}$ is uniform on a D-dimensional hypersphere

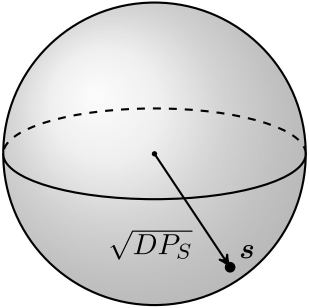
$$\|\boldsymbol{s}\|^2 = DP_S$$

- Data is encrypted by a nonlinear field

$$\boldsymbol{x} = \mathcal{V}(\boldsymbol{s})$$

Nonlinear Encryption by Gaussian Fields

Fyodorov uses the nonlinear model to encrypt data on a *hypersphere*



- $\boldsymbol{s}$ is uniform on a D-dimensional hypersphere

$$\|\boldsymbol{s}\|^2 = DP_S$$

- Data is encrypted by a nonlinear field

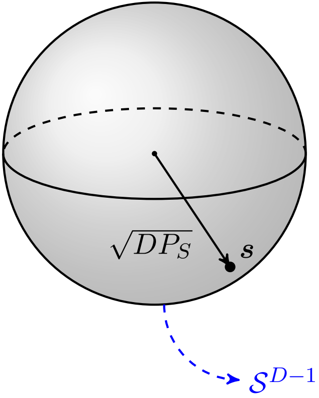
$$\boldsymbol{x} = \mathcal{V}(\boldsymbol{s})$$

- $\boldsymbol{x}$ is observed through a Gaussian channel

$$\boldsymbol{y} = \boldsymbol{x} + \mathbf{w}$$

Nonlinear Encryption by Gaussian Fields

Fyodorov uses the nonlinear model to encrypt data on a *hypersphere*



- $\boldsymbol{s}$ is uniform on a D-dimensional hypersphere

$$\|\boldsymbol{s}\|^2 = DP_S$$

- Data is encrypted by a nonlinear field

$$\boldsymbol{x} = \mathcal{V}(\boldsymbol{s})$$

- $\boldsymbol{x}$ is observed through a Gaussian channel

$$\boldsymbol{y} = \boldsymbol{x} + \mathbf{w}$$

$$\mathcal{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}_N\right)$$

Nonlinear Encryption by Gaussian Fields

To recover $\boldsymbol{s}$ from $\boldsymbol{y}$, Fyodorov uses

- from the regression viewpoint "*the method of least-squares*"
- in the Bayesian framework "*the maximum-a-posteriori estimator*"

and finds

$$\hat{\boldsymbol{s}} = \underset{\mathbf{u} \in \mathcal{S}^{D-1}}{\operatorname{argmin}} \left\| \boldsymbol{y} - \mathcal{V}\left(\mathbf{u}\right) \right\|^2$$

Nonlinear Encryption by Gaussian Fields

To recover $\boldsymbol{s}$ from $\boldsymbol{y}$, Fyodorov uses

- from the regression viewpoint "*the method of least-squares*"
- in the Bayesian framework "*the maximum-a-posteriori estimator*"

and finds

$$\hat{\boldsymbol{s}} = \underset{\mathbf{u} \in \mathcal{S}^{D-1}}{\operatorname{argmin}} \left\| \boldsymbol{y} - \mathcal{V}(\mathbf{u}) \right\|^2$$

Main Result: He determines the asymptotic overlap

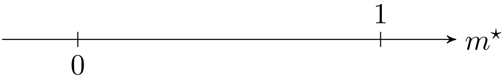
$$m^{\star} = \lim_{D,N \uparrow \infty} \frac{\mathbb{E}\left\{ \langle \boldsymbol{s}; \hat{\boldsymbol{s}} \rangle \right\}}{D P_S}$$

with D/N kept fixed via the replica method considering the full-RSB ansatz

Nonlinear Encryption by Gaussian Fields

What does the overlap mean in the wiretap setting?

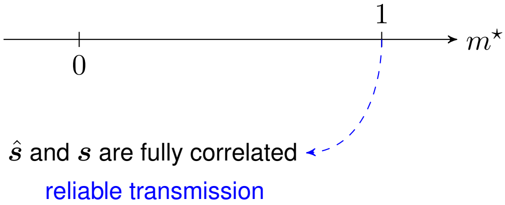
It characterizes reliability and security



Nonlinear Encryption by Gaussian Fields

What does the overlap mean in the wiretap setting?

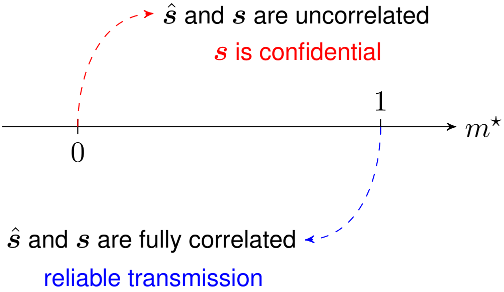
It characterizes reliability and security



Nonlinear Encryption by Gaussian Fields

What does the overlap mean in the wiretap setting?

It characterizes reliability and security



Key Observations By Fyodorov

Fyodorov reports the following key findings in his paper

- For any Gaussian field containing a linear term

  the overlap never touches $m^\star = 0$

Key Observations By Fyodorov

Fyodorov reports the following key findings in his paper

- For any Gaussian field containing a linear term

  the overlap never touches $m^\star = 0$

- With a *strictly* nonlinear Gaussian field however

  the overlap exhibits a second-order phase transition!

Key Observations By Fyodorov

Fyodorov reports the following key findings in his paper

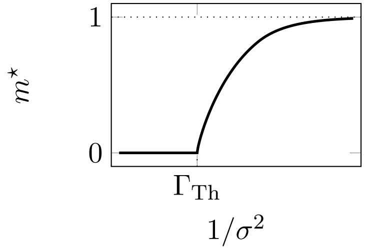
- For any Gaussian field containing a linear term

  the overlap never touches $m^\star = 0$

- With a *strictly* nonlinear Gaussian field however

  the overlap exhibits a second-order phase transition!

Example: Purely Quadratic Field



Key Observations By Fyodorov

Fyodorov suggests that this can be used to provide perfect secrecy

> The existence of a sharp NSR threshold $\hat{\gamma}_c$ in the pure quadratic encryption case may have useful consequences for security of transmitting the encrypted signal. Indeed, it is a quite common assumption that an eavesdropper may get access to the transmitted signal by a channel with inferior quality, characterized by higher level of noise. This may then result in impossibility for eavesdroppers to reconstruct the quadratically encoded signal even if the encoding algorithm is perfectly known to them.

Key Observations By Fyodorov

Fyodorov suggests that this can be used to provide perfect secrecy

> The existence of a sharp NSR threshold $\hat{\gamma}_c$ in the pure quadratic encryption case may have useful consequences for security of transmitting the encrypted signal. Indeed, it is a quite common assumption that an eavesdropper may get access to the transmitted signal by a channel with inferior quality, characterized by higher level of noise. This may then result in impossibility for eavesdroppers to reconstruct the quadratically encoded signal even if the encoding algorithm is perfectly known to them.

Well! This is the wiretap channel of Wyner!

First Try: *Let Fyodorov and Wyner Meet*

Encryption on Top of Sphere Coding

We have two Gaussian channels:

- One to the good guy *Bob* with noise variance σ_{B}^2
- One to the bad guy *Eve* with noise variance σ_{E}^2

Encryption on Top of Sphere Coding

We have two Gaussian channels:

- One to the good guy *Bob* with noise variance σ_{B}^2
- One to the bad guy *Eve* with noise variance σ_{E}^2

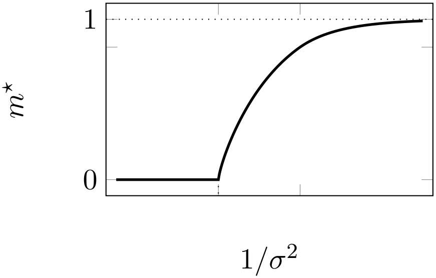
We spend some power on noise, say ξ, and remaining $1-\xi$ on the signal

Encryption on Top of Sphere Coding

We have two Gaussian channels:

- One to the good guy *Bob* with noise variance σ_{B}^2
- One to the bad guy *Eve* with noise variance σ_{E}^2

We spend some power on noise, say ξ, and remaining $1-\xi$ on the signal

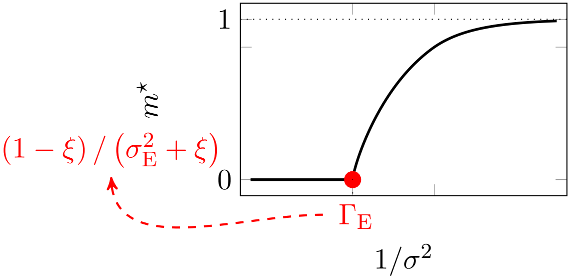


Encryption on Top of Sphere Coding

We have two Gaussian channels:

- One to the good guy *Bob* with noise variance σ_{B}^2
- One to the bad guy *Eve* with noise variance σ_{E}^2

We spend some power on noise, say ξ, and remaining $1-\xi$ on the signal

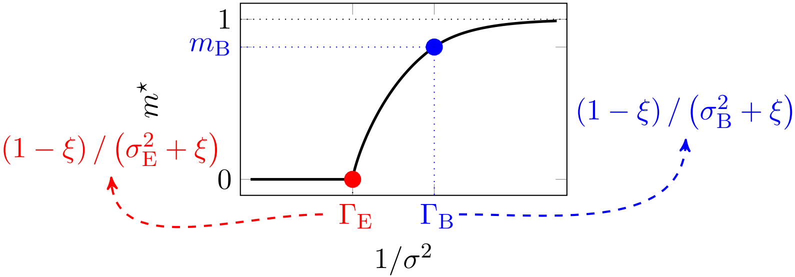


Encryption on Top of Sphere Coding

We have two Gaussian channels:

- One to the good guy *Bob* with noise variance σ_{B}^2
- One to the bad guy *Eve* with noise variance σ_{E}^2

We spend some power on noise, say ξ, and remaining $1-\xi$ on the signal

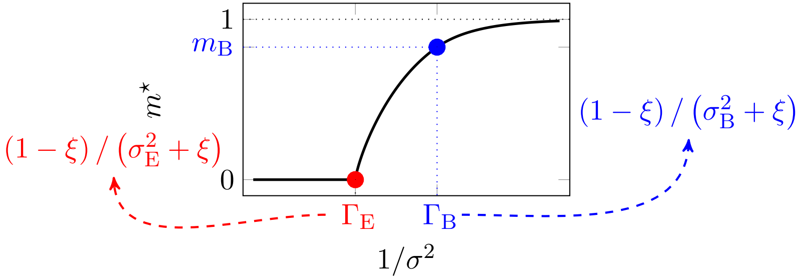


Encryption on Top of Sphere Coding

We have two Gaussian channels:

- One to the good guy *Bob* with noise variance σ_{B}^2
- One to the bad guy *Eve* with noise variance σ_{E}^2

We spend some power on noise, say ξ, and remaining $1-\xi$ on the signal

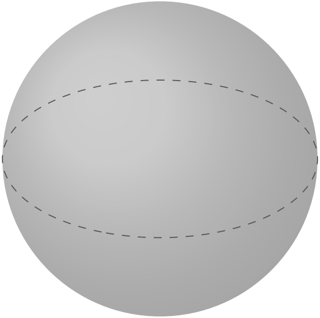


Then, we set the minimum distance $d_{\mathrm{B}} = 1 - m_{\mathrm{B}}$

Encryption on Top of Sphere Coding

We now go on the hypersphere and put 2^K points, such that

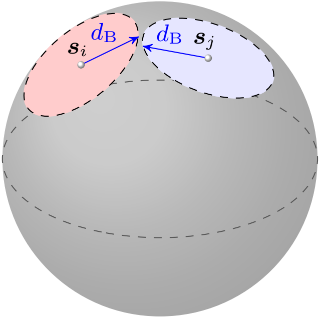
$$\min_{i \neq j} \frac{\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2}{DP_S} \geq 2d_\mathrm{B}$$



Encryption on Top of Sphere Coding

We now go on the hypersphere and put 2^K points, such that

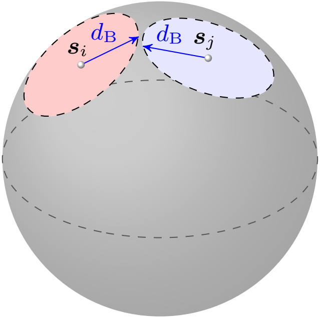
$$\min_{i \neq j} \frac{\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2}{DP_S} \geq 2 d_{\mathrm{B}}$$



Encryption on Top of Sphere Coding

We now go on the hypersphere and put 2^K points, such that

$$\min_{i \neq j} \frac{\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2}{DP_S} \geq 2d_{\mathrm{B}}$$



We have perfect secrecy if we encrypt $\boldsymbol{s}_i$ by

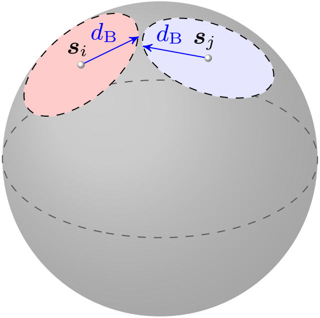
the purely quadratic field

The *secure transmission rate* is then $R = K/N$

Encryption on Top of Sphere Coding

We now go on the hypersphere and put 2^K points, such that

$$\min_{i\neq j} \frac{\|\boldsymbol{s}_i - \boldsymbol{s}_j\|^2}{DP_S} \geq 2d_{\mathrm{B}}$$



We have perfect secrecy if we encrypt $\boldsymbol{s}_i$ by

the purely quadratic field

The *secure transmission rate* is then $R = K/N$

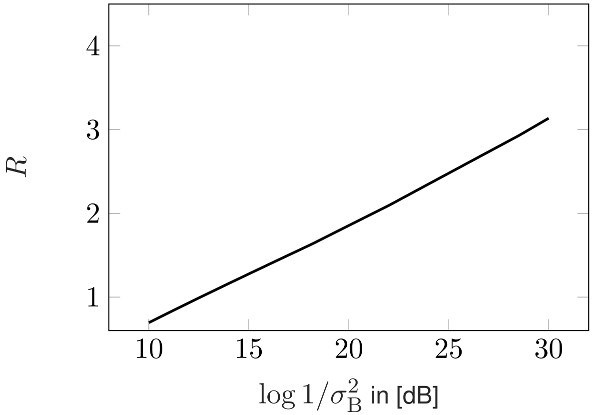
Best thing we can do is to

put as much points as possible

Finding maximum K is sphere covering

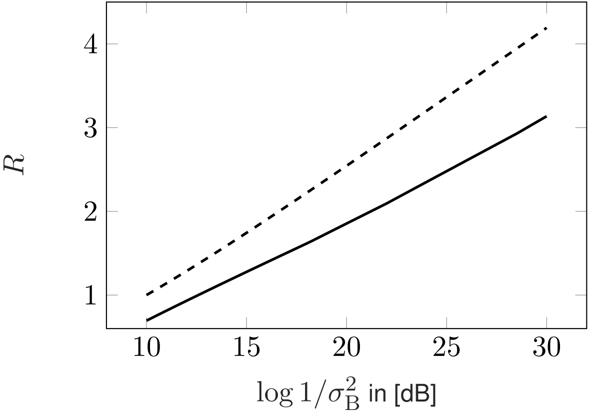
Let's use an optimistic bound

Encryption on Top of Sphere Coding



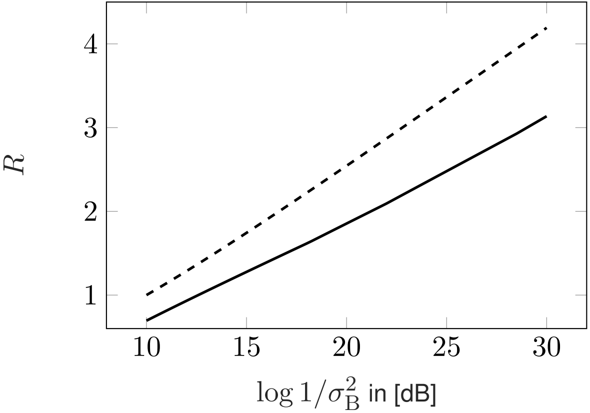
Encryption on Top of Sphere Coding

It didn't really end up well!



Encryption on Top of Sphere Coding

It didn't really end up well! But David MacKay would have told us so



Second Try: Introducing Fyodorov to Wyner

Encoding via Nonlinear Gaussian Fields

We represent the bits via $\boldsymbol{s} \in \{\pm 1\}^K$ and directly pass them through the field

$$\boldsymbol{x} = \mathcal{V}\left(\boldsymbol{s}\right)$$

At the output of the Gaussian channel, we have

$$\boldsymbol{y} = \boldsymbol{x} + \mathbf{w}$$

Encoding via Nonlinear Gaussian Fields

We represent the bits via $\boldsymbol{s} \in \{\pm 1\}^K$ and directly pass them through the field

$$\boldsymbol{x} = \mathcal{V}\left(\boldsymbol{s}\right)$$

At the output of the Gaussian channel, we have

$$\boldsymbol{y} = \boldsymbol{x} + \mathbf{w}$$

We then recover $\boldsymbol{s}$ via the Bayesian optimal algorithm

$$\hat{\boldsymbol{s}} = \mathbb{E}\left\{\boldsymbol{s}|\boldsymbol{y}\right\} = \frac{\mathbb{E}_{\boldsymbol{s}}\left\{\boldsymbol{s}\exp\left\{-\frac{\|\boldsymbol{y}-\mathcal{V}\left(\boldsymbol{s}\right)\|^2}{2\sigma^2}\right\}\right\}}{\mathbb{E}_{\boldsymbol{s}}\left\{\exp\left\{-\frac{\|\boldsymbol{y}-\mathcal{V}\left(\boldsymbol{s}\right)\|^2}{2\sigma^2}\right\}\right\}}$$

Asymptotics via the Replica Method

We could now do some replica calculations

- Define the variational problem

finding the mutual information $\equiv$ *finding a free energy*

Asymptotics via the Replica Method

We could now do some replica calculations

- Define the variational problem

  finding the mutual information $\equiv$ *finding a free energy*

- Using the replica method to find the free energy

  We focus on the Bayesian optimal case, and hence the RS solution

Asymptotics via the Replica Method

We could now do some replica calculations

- Define the variational problem

finding the mutual information $\equiv$ *finding a free energy*

- Using the replica method to find the free energy

We focus on the Bayesian optimal case, and hence the RS solution

RS Solution: Free energy (*for you*) or mutual information (*for me*) reads

$$\mathcal{L}_m = \frac{1}{2} \log \left(1 + \xi_m \right) + \mathcal{Q}_m$$

The overlap $m^\star$

$$m^\star = \operatorname*{argmin}_{m \in [0,1]} \; \mathcal{L}_m$$

Asymptotics via the Replica Method

We could now do some replica calculations

- Define the variational problem

finding the mutual information $\equiv$ *finding a free energy*

- Using the replica method to find the free energy

We focus on the Bayesian optimal case, and hence the RS solution

RS Solution: Free energy (*for you*) or mutual information (*for me*) reads

$$\mathcal{L}_m = \frac{1}{2} \log\left(1 + \xi_m\right) + \mathcal{Q}_m$$

in terms of $\Phi\left(\cdot\right), \dots$

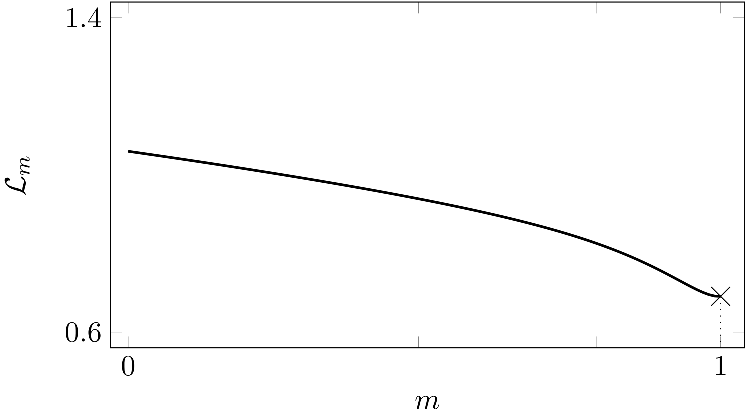
The overlap $m^\star$

$$m^\star = \underset{m \in [0,1]}{\operatorname{argmin}} \ \mathcal{L}_m$$

RS Solution: Linear Field

We start with a conventional linear field $\Phi\left(x\right) = x$

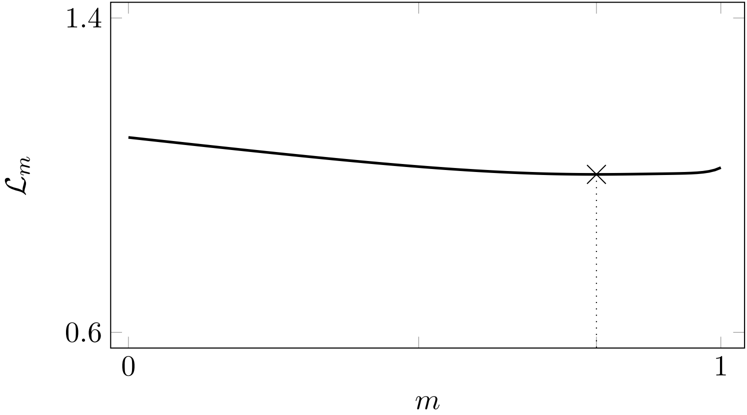
start with small $R = K/N$ and gradually increase it



RS Solution: Linear Field

We start with a conventional linear field $\Phi(x) = x$

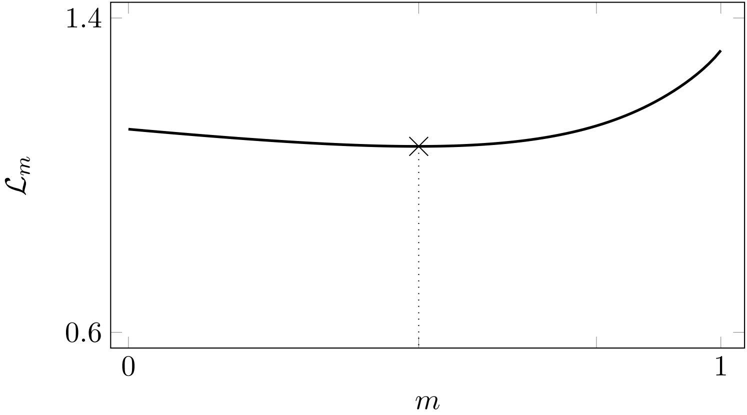
start with small $R = K/N$ *and gradually increase it*



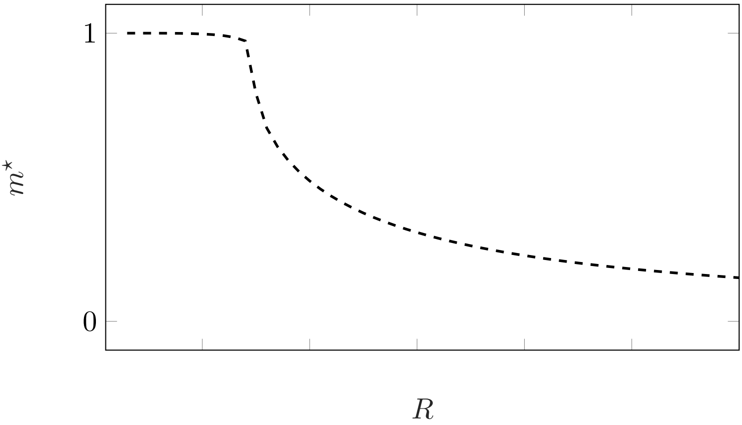
RS Solution: Linear Field

We start with a conventional linear field $\Phi\left(x\right) = x$

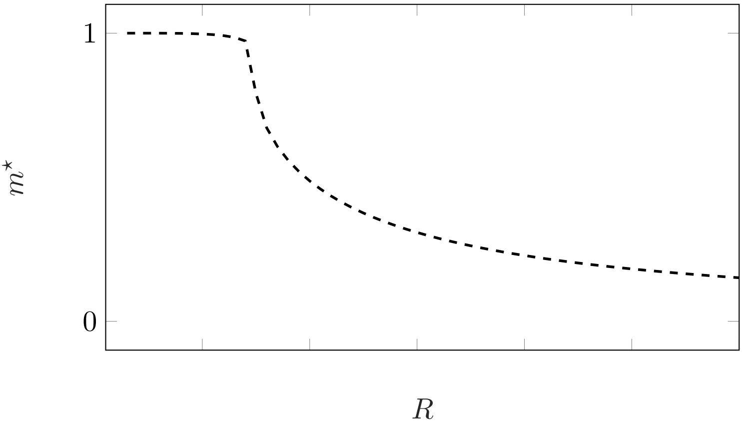
start with small $R = K/N$ and gradually increase it



RS Solution: Linear Field



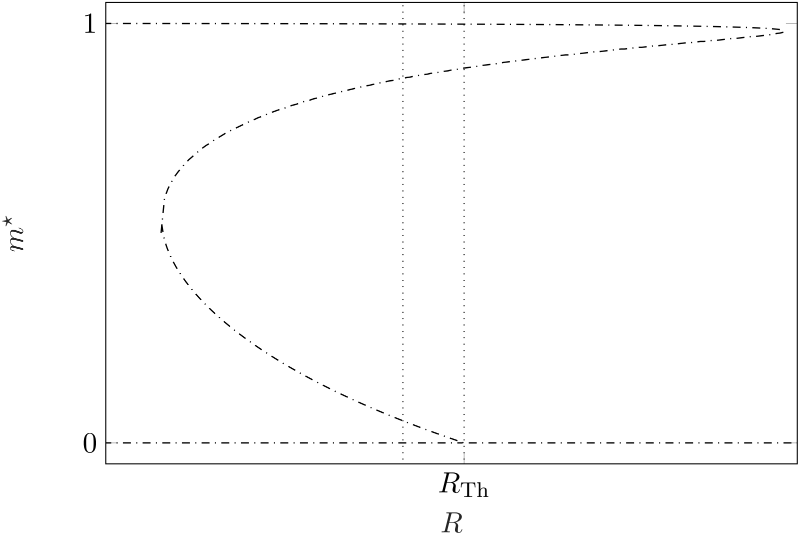
RS Solution: Linear Field



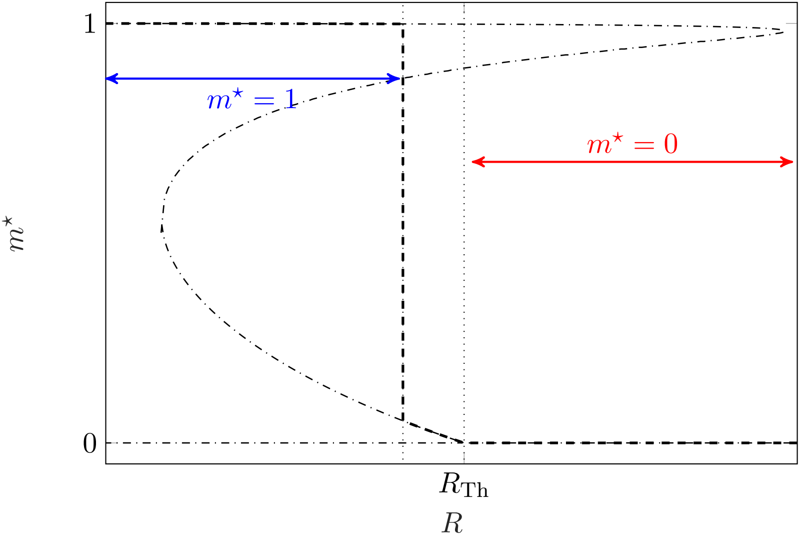
We can easily show that

$$\boxed{\textit{overlap never touches } m^\star = 0}$$

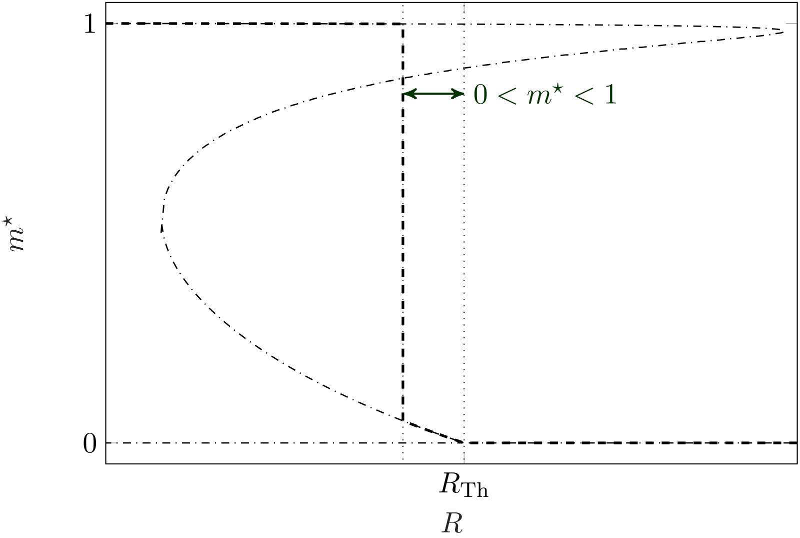
RS Solution: Purely Quadratic Field



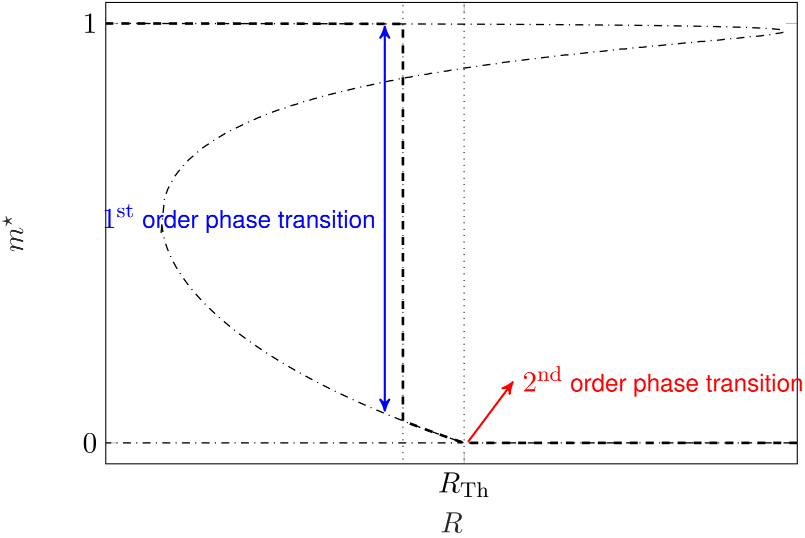
RS Solution: Purely Quadratic Field



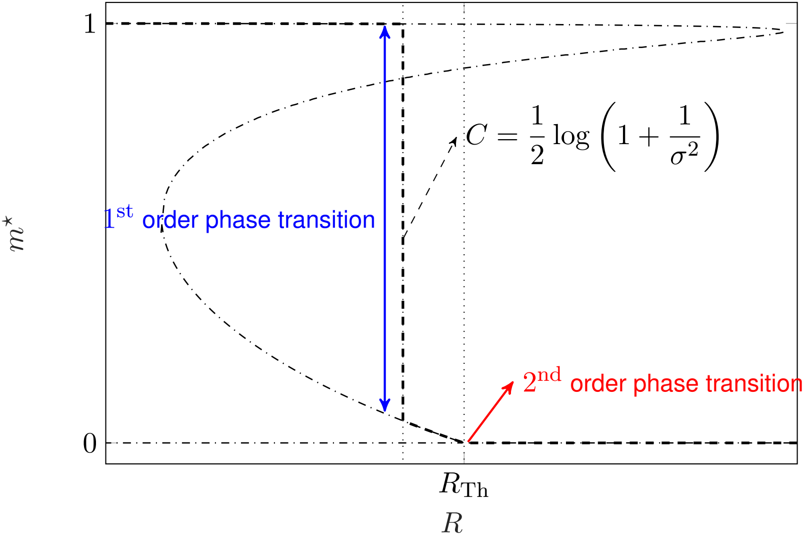
RS Solution: Purely Quadratic Field



RS Solution: Purely Quadratic Field



RS Solution: Purely Quadratic Field

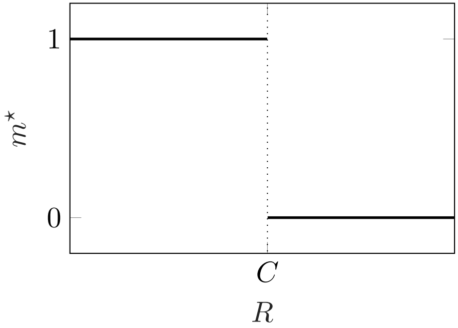


RS Solution: Higher Order Fields

What if we go for higher orders?

RS Solution: Higher Order Fields

What if we go for higher orders?



The moral of story is

- Strictly nonlinear fields show *all-or-nothing* behavior
- The phase transition exactly occurs at Shannon's capacity

RS Solution: Higher Order Fields

We now only need to

- set the eavesdropper in the "*nothing*" area
- set the legitimate receiver in the "*all*" area

RS Solution: Higher Order Fields

We now only need to

- set the eavesdropper in the "*nothing*" area
- set the legitimate receiver in the "*all*" area

We have done this and verified that

Wyner's bound is achieved

RS Solution: Higher Order Fields

We now only need to

- set the eavesdropper in the "*nothing*" area
- set the legitimate receiver in the "*all*" area

We have done this and verified that

Wyner's bound is achieved

You might now tell me that

Forget about Wyner, we have found a good channel code

RS Solution: Higher Order Fields

We now only need to

- set the eavesdropper in the "*nothing*" area
- set the legitimate receiver in the "*all*" area

We have done this and verified that

Wyner's bound is achieved

You might now tell me that

Forget about Wyner, we have found a good channel code

My answer would then be

Yes and No!

A Step Back and Conclusions

Earlier Work by Sourlas (1989)

After the initial draft, I found out that Sourlas published a paper in 1989

LETTERS TO NATURE

Spin-glass models as error-correcting codes

Nicolas Sourlas

Laboratoire de Physique Théorique de l'Ecole Normale Supérieure,
24 rue Lhomond, 75231 Paris Cédex 05, France

Earlier Work by Sourlas (1989)

After the initial draft, I found out that Sourlas published a paper in 1989

EUROPHYSICS LETTERS 20 January 1994

Europhys. Lett., **25** (3), pp. 159-164 (1994)

Spin Glasses, Error-Correcting Codes and Finite-Temperature Decoding.

N. SOURLAS (*)

Laboratoire de Physique Théorique de l'Ecole Normale Supérieure (**)
24 rue Lhomond, 75231 Paris Cedex 05, France

LETTERS TO NATURE

Spin-glass models as error-correcting codes

Nicolas Sourlas

Laboratoire de Physique Théorique de l'Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cédex 05, France

Earlier Work by Sourlas (1989)

After the initial draft, I found out that Sourlas published a paper in 1989

LETTERS TO NATURE

Spin-glass models as error-correcting codes

Nicolas Sourlas

Laboratoire de Physique Théorique de l'Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cédex 05, France

EUROPHYSICS LETTERS 20 January 1994

Europhys. Lett., **25** (3), pp. 159-164 (1994)

Spin Glasses, Error-Correcting Codes and Finite-Temperature Decoding.

N. SOURLAS (*)

Laboratoire de Physique Théorique de l'Ecole Normale Supérieure (**) *24 rue Lhomond, 75231 Paris Cedex 05, France*

EUROPHYSICS LETTERS 1 January 1999

Europhys. Lett., **45** (1), pp. 97-103 (1999)

Statistical mechanics of error-correcting codes

Y. KABASHIMA[1](*) and D. SAAD[2](**)

[1] *Department of Computational Intelligence and Systems Science Tokyo Institute of Technology - Yokohama 226, Japan*

[2] *The Neural Computing Research Group, Aston University Birmingham B4 7ET, UK*

Earlier Work by Sourlas (1989)

After the initial draft, I found out that Sourlas published a paper in 1989

LETTERS TO NATURE

Spin-glass models as error-correcting codes

Nicolas Sourlas

Laboratoire de Physique Théorique de l'Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cédex 05, France

EUROPHYSICS LETTERS 20 January 1994

Europhys. Lett., **25** (3), pp. 159-164 (1994)

Spin Glasses, Error-Correcting Codes and Finite-Temperature Decoding.

N. SOURLAS (*)

Laboratoire de Physique Théorique de l'Ecole Normale Supérieure (**) *24 rue Lhomond, 75231 Paris Cedex 05, France*

EUROPHYSICS LETTERS 1 January 1999

Europhys. Lett., **45** (1), pp. 97-103 (1999)

Statistical mechanics of error-correcting codes

Y. KABASHIMA[1](*) and D. SAAD[2](**)

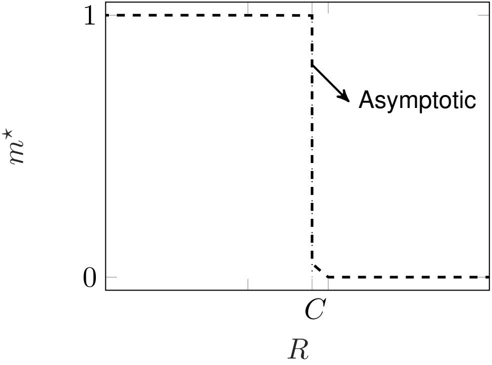
[1] *Department of Computational Intelligence and Systems Science Tokyo Institute of Technology - Yokohama 226, Japan*
[2] *The Neural Computing Research Group, Aston University Birmingham B4 7ET, UK*

It however did not find its way to information theory literature

as people considered it to be impractical

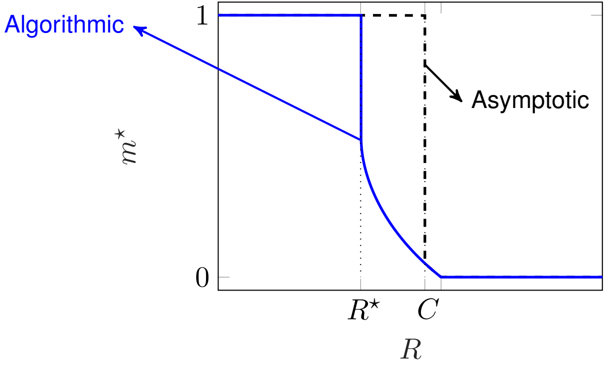
Comments on Practicability

To be honest, “people” were partially right!



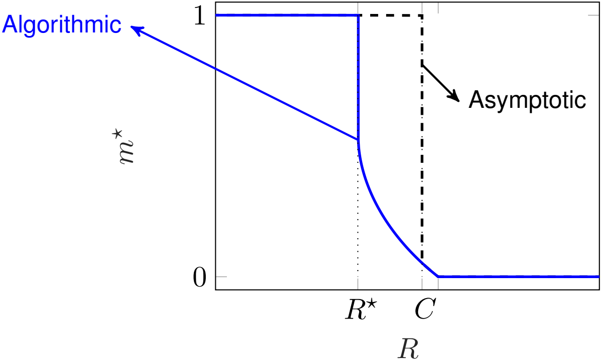
Comments on Practicability

To be honest, “people” were partially right!



Comments on Practicability

To be honest, “people” were partially right!



For higher-order fields, we are always algorithmically at $m = 0$

unless we replace the random field with a spatially-coupled one!

Conclusions

Nonlinear models show *all-or-nothing* property

- It has a direct application to secure coding
- It seems to provide secrecy to other learning problems

Conclusions

Nonlinear models show *all-or-nothing* property

- It has a direct application to secure coding
- It seems to provide secrecy to other learning problems

What am I looking for right now?

- Other applications for nonlinear generative models
- AMP-based implementation of Bayesian inference on nonlinear models

Time for Questions

Special thanks to

- Yan Fyodorov
- Lenka Zdeborová
- Nicolas Macris