The Stochastic Complexity of Spin Models

Are Pairwise Models Really Simple?

APS March Meeting 2019

Alberto Beretta, Claudia Battistin,
Clélia de Mulatier, Iacopo Mastromatteo, Matteo Marsili

Context

- Model complexity? Why is this interesting?
- What about spin models?

Complexity emerges from the problem of **Model Selection**

- **Finite data with random errors:** find the model that best captures the patterns hidden within the data…

- Ideally, we would like the model to be:

  — **not too simple**: to be able to fit well the data;

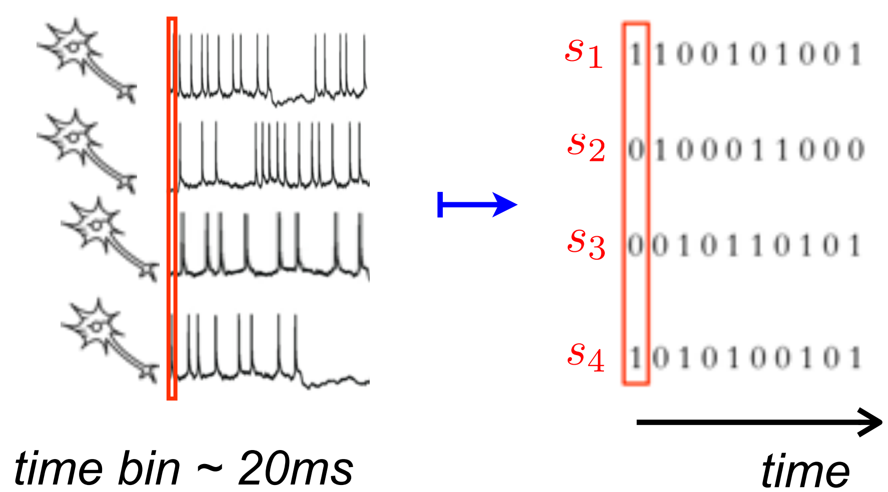
  — **not too complex**: to capture the main patterns of the data and not noise.

⟶ We would prefer a simple model, unless the data calls for a more complex one.

Spin Models are used for binary data

Spin Models are probabilistic models commonly used to analyze **binary datasets**.

— Neuronal data:



— Animal behavioral data:

[Talk] E. Amstrong

— Voting data

— Financial data (ex. Stock market);

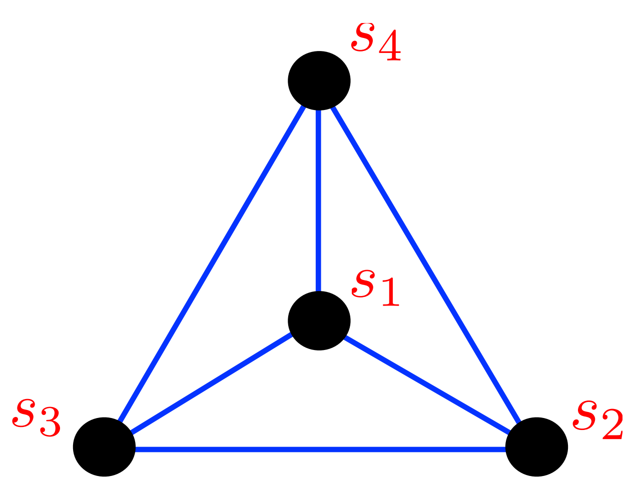
— Medical imaging data (for disease diagnoses)

— etc.

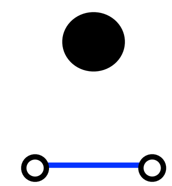
Ising Models are Spin Models

Fields and Pairwise Interactions

Ex. $n = 4$



Fully connected pairwise model



K=10 interactions, parametrized by:

$$\vec{g} = \{h_1, h_2, h_3, h_4, J_{13}, J_{23}, \cdots\}$$

Spins:

$$s_i \in \{-1, +1\}$$

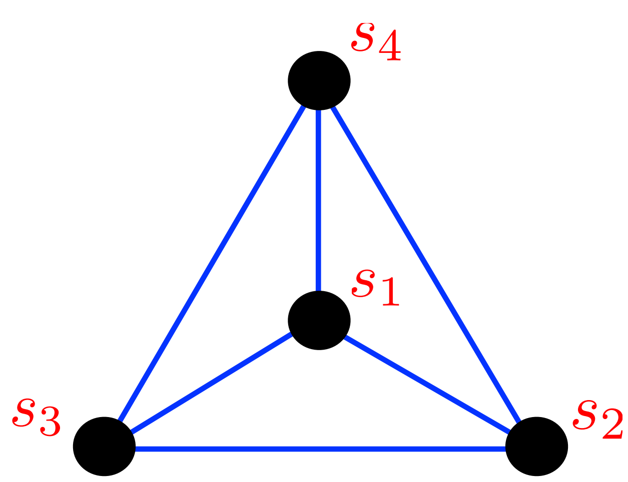
Probability that the n spins are in the **configuration** $\vec{s}$:

$$P(\vec{s}\,|\,\vec{g}, \mathcal{M}) = \frac{1}{Z_{\mathcal{M}}(\vec{g})} \exp\left(\sum_{i \in \mathcal{M}} h_i s_i + \sum_{pair(i,j) \in \mathcal{M}} J_{ij} s_i s_j \right)$$

Ising Models are Spin Models

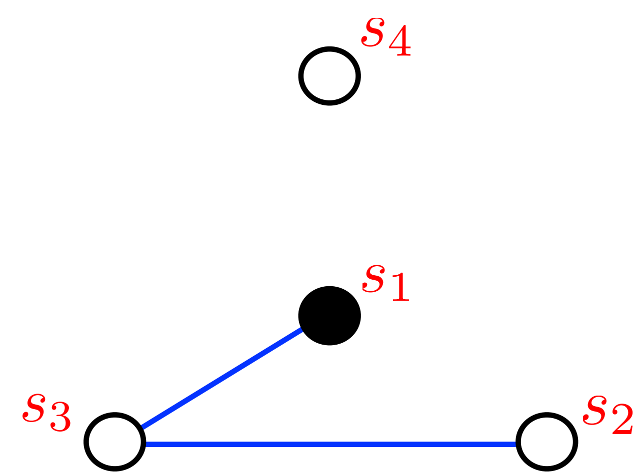
Fields and **Pairwise Interactions**

Ex. $n = 4$



Fully connected pairwise model

$$\vec{g} = \{h_1, h_2, h_3, h_4, J_{13}, J_{23}, \cdots\}$$



Pairwise model with **K=3 interactions**

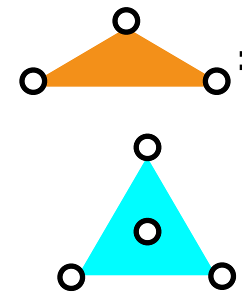
$$\vec{g} = \{h_1, J_{13}, J_{23}\}$$

Model $\mathcal{M}$ **= skeleton**

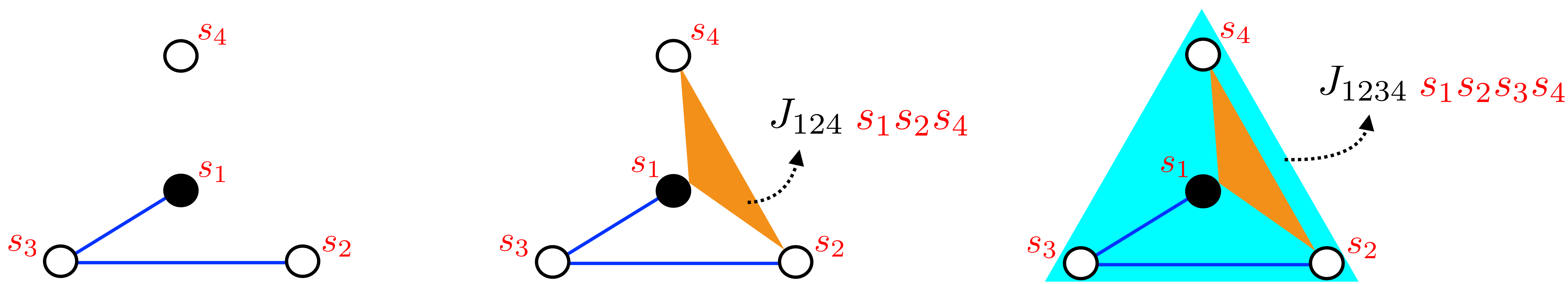
Parameters $\vec{g}$

General Spin Models

Interactions of Any Order



Ex. $n = 4$



Probability that the n spins are in the **configuration** $\vec{s}$:

$$P(\vec{s}\,|\,\vec{g},\mathcal{M}) = \frac{1}{Z_{\mathcal{M}}(\vec{g})}\ \exp\left(\sum_{k\in\mathcal{M}} g_k\ \phi_k(\vec{s})\right)$$

parameters (g_k) — *Spin operator* ($\phi_k(\vec{s})$)

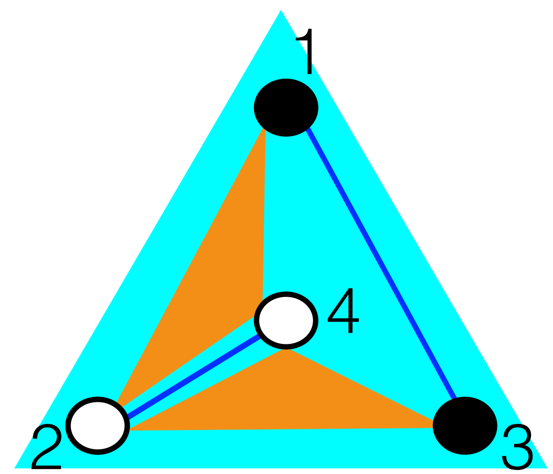
→ **Rigorous**

Which Model is the Simplest?

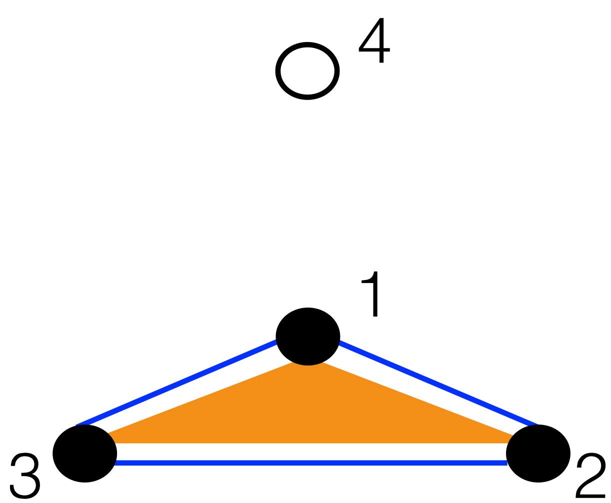
[Model = skeleton]

7 parameters

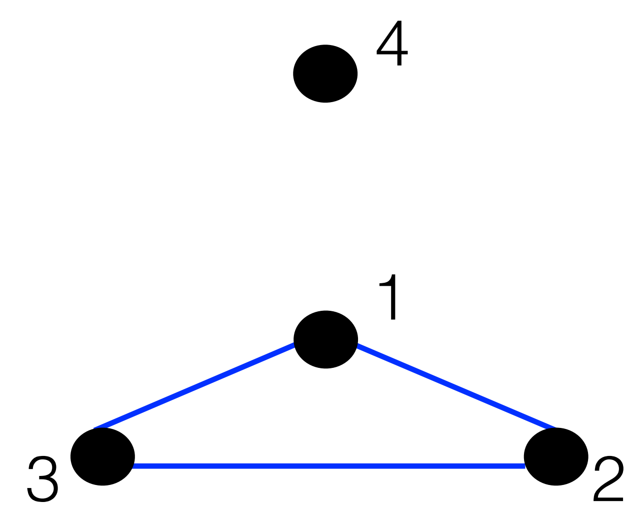
a)



b)

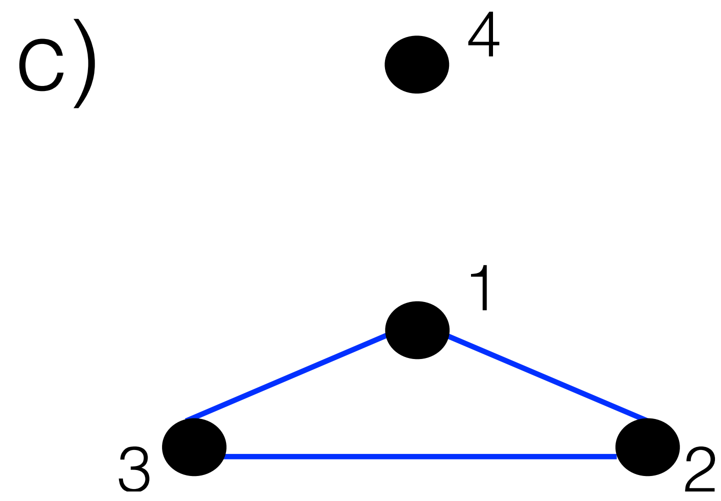


c)



How is simplicity/complexity **related to** the model **architecture?**

Are Pairwise Models Simpler?



Is simplicity/complexity **related to** the **order of the interactions?**

Complexity of Spin Models

- **Define** Model **Complexity?**
- Complexity of Spin Models? **Thought Experiment…**

Model Complexity

[J. Rissanen] *Fisher Information and Stochastic Complexity* (1996)

$$\mathrm{COMP}(\mathcal{M}) = \boxed{\frac{K}{2}\log\frac{N}{2\pi}} + c_{\mathcal{M}} + O\left(\frac{1}{N}\right)$$

Due to Number of Parameters *K*

Due to Geometry

$$c_{\mathcal{M}} = \log\left[\int \sqrt{\det I(\boldsymbol{g})}\ \mathrm{d}^K\boldsymbol{g}\right]$$

- $c_{\mathcal{M}}$ more complex models are more flexible, they can fit well broad type of data patterns.

[I. J. Myung, V. Balasubramanian, M. A. Pitt] *Counting probability distributions: Differential geometry and model selection*

- Difficult to compute

Thought Experiment

Bob's dataset:

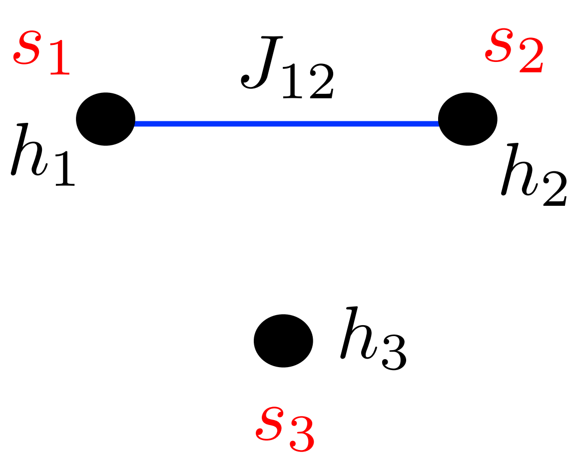
$$\hat{s} = \{\vec{s}^{\,(i)}\}$$

Thought Experiment

Bob's dataset:

$$\hat{s} = \{\vec{s}^{(i)}\}$$

"magic" MS ↓ rank all $P(\mathcal{M} \mid \hat{s})$

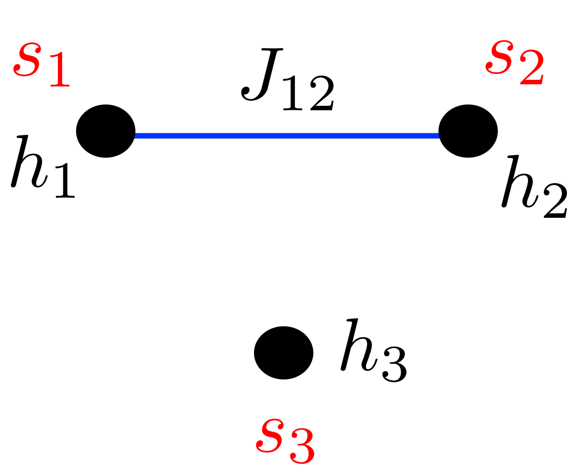


Thought Experiment

Bob's dataset:

$$\hat{s} = \{\vec{s}^{(i)}\}$$

"magic" MS



$$\xrightarrow{\mathcal{T}}$$

$$\begin{cases} \sigma_1 = s_1 s_2 s_3 \\ \sigma_2 = s_2 \\ \sigma_3 = s_3 \end{cases}$$

Alice's dataset:

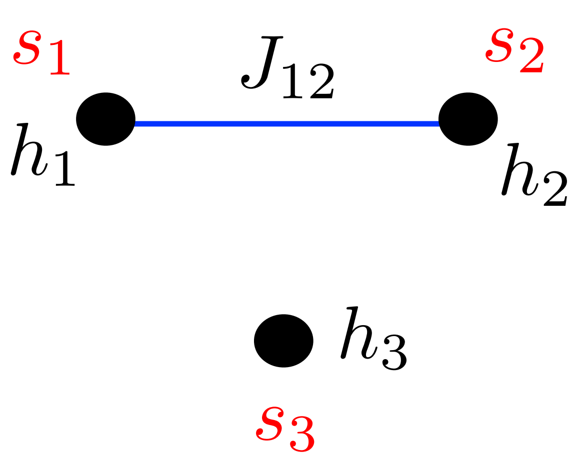
$$\hat{\sigma} = \{\vec{\sigma}^{(i)}\}$$

Thought Experiment

Bob's dataset:

$$\hat{s} = \{\vec{s}^{\,(i)}\}$$

"magic" MS



$$\xrightarrow{\mathcal{T}}$$

$$\begin{cases} \sigma_1 = s_1 s_2 s_3 \\ \sigma_2 = s_2 \\ \sigma_3 = s_3 \end{cases}$$

Alice's dataset:

$$\hat{\sigma} = \{\vec{\sigma}^{\,(i)}\}$$

"magic" MS

???

Thought Experiment

Bob's dataset:

$$\hat{s} = \{\vec{s}^{\,(i)}\}$$

$$\xrightarrow{\;\mathcal{T}\;}$$

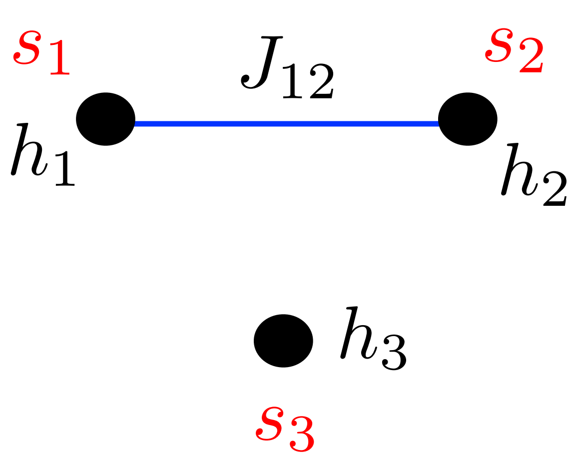
$$\begin{cases} \sigma_1 = s_1 s_2 s_3 \\ \sigma_2 = s_2 \\ \sigma_3 = s_3 \end{cases}$$

Alice's dataset:

$$\hat{\sigma} = \{\vec{\sigma}^{\,(i)}\}$$

"magic" MS ↓

↓ "magic" MS



$$\overset{\mathcal{T}}{\Longrightarrow}$$

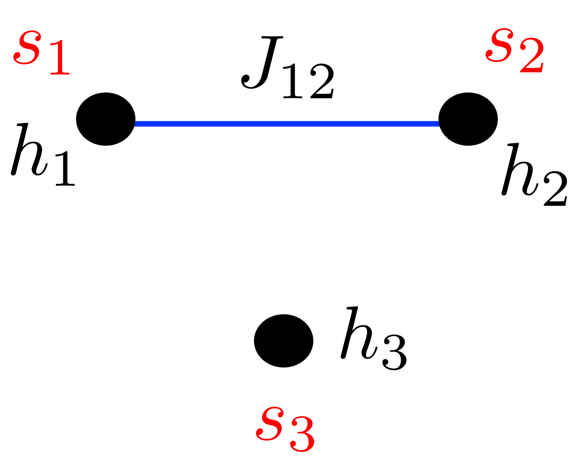
???

Thought Experiment

Bob's dataset:

$$\hat{s} = \{\vec{s}^{(i)}\}$$

"magic" MS



$$\xrightarrow{\mathcal{T}}$$

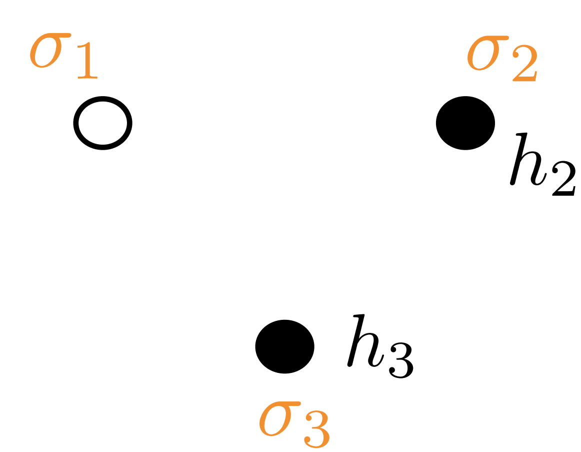
$$\begin{cases} \sigma_1 = s_1 s_2 s_3 \\ \sigma_2 = s_2 \\ \sigma_3 = s_3 \end{cases}$$

$$\overset{\mathcal{T}}{\Longrightarrow}$$

Alice's dataset:

$$\hat{\sigma} = \{\vec{\sigma}^{(i)}\}$$

"magic" MS



As:

$$\sigma_2 = s_2$$

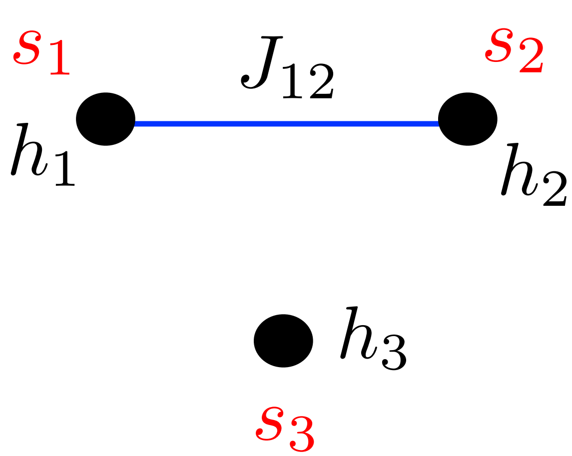
$$\sigma_3 = s_3$$

Thought Experiment

Bob's dataset:

$$\hat{s} = \{\vec{s}^{\,(i)}\}$$

"magic" MS



$$\xrightarrow{\mathcal{T}}$$

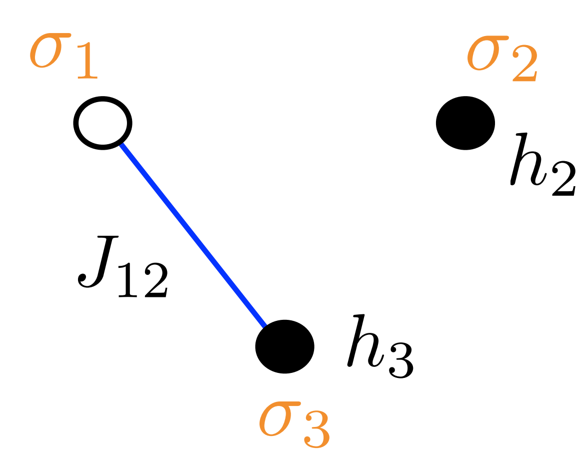
$$\begin{cases} \sigma_1 = s_1 s_2 s_3 \\ \sigma_2 = s_2 \\ \sigma_3 = s_3 \end{cases}$$

$$\overset{\mathcal{T}}{\Longrightarrow}$$

Alice's dataset:

$$\hat{\sigma} = \{\vec{\sigma}^{\,(i)}\}$$

"magic" MS



As: $s_1 s_2 = \sigma_1 \sigma_3$

Thought Experiment

Bob's dataset:

$$\hat{s} = \{\vec{s}^{\,(i)}\}$$

"magic" MS

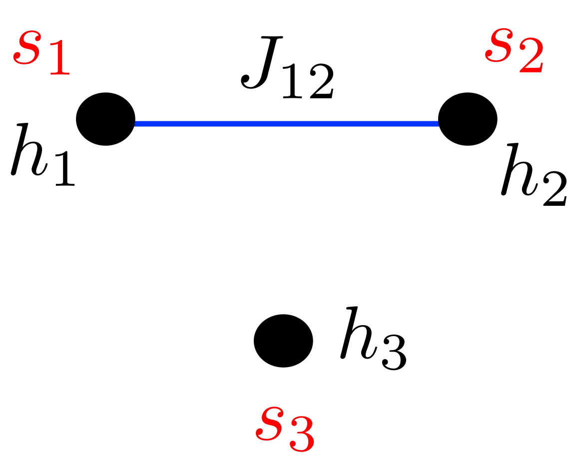
$$\xrightarrow{\mathcal{T}}$$

$$\begin{cases} \sigma_1 = s_1 s_2 s_3 \\ \sigma_2 = s_2 \\ \sigma_3 = s_3 \end{cases}$$

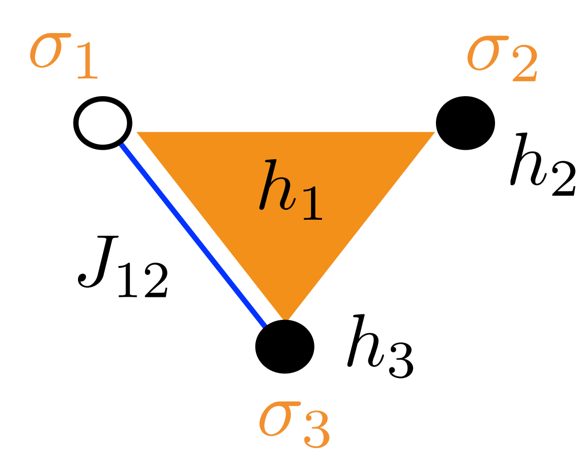
Alice's dataset:

$$\hat{\sigma} = \{\vec{\sigma}^{\,(i)}\}$$

"magic" MS



$$\overset{\mathcal{T}}{\Longrightarrow}$$



As: $s_1 = \sigma_1 \sigma_2 \sigma_3$

Thought Experiment

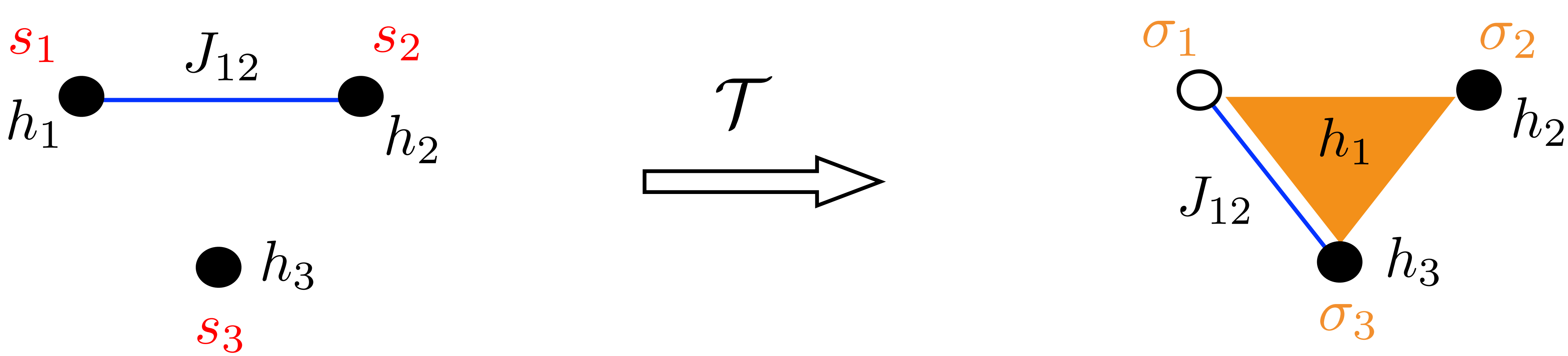
Bob's dataset: $\hat{s} = \{\vec{s}^{\,(i)}\}$

$\xrightarrow{\mathcal{T}}$

Alice's dataset: $\hat{\sigma} = \{\vec{\sigma}^{\,(i)}\}$

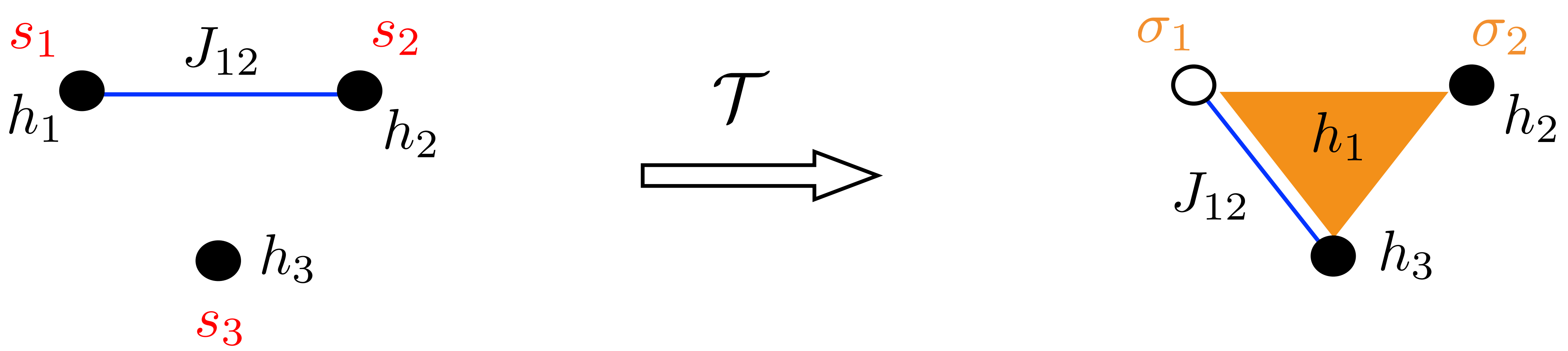
$$\begin{cases} \sigma_1 = s_1 s_2 s_3 \\ \sigma_2 = s_2 \\ \sigma_3 = s_3 \end{cases}$$

"magic" MS ↓ "magic" MS



⟶ These 2 models must be **As Complex**!!

Thought Experiment



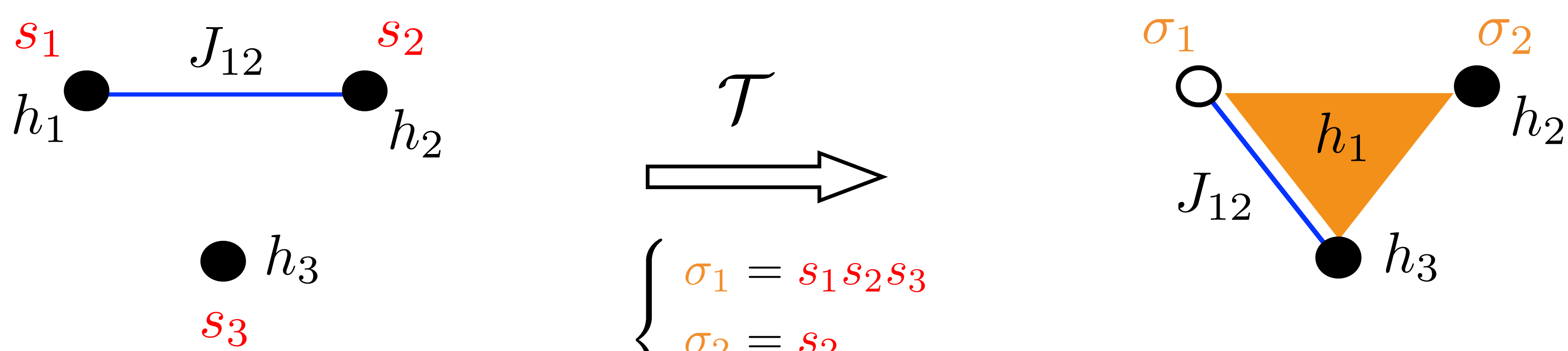
Same complexity!

<u>First Conclusions:</u>

- Difficult to guess from the look of the models if one is more complex.
- In particular:
  - — **pairwise models** are **not necessarily simpler**
  - — **complexity** is **not defined by the order** of the interactions

Some Results and Perspectives…

"*Gauge Transformations*"



$$\begin{cases} \sigma_1 = s_1 s_2 s_3 \\ \sigma_2 = s_2 \\ \sigma_3 = s_3 \end{cases}$$

$\mathcal{T}$ is a **change of basis**

It **preserves:**

— The **number of interactions** in the model;

— The **intrinsic architecture** of the model (***loop* structure**).

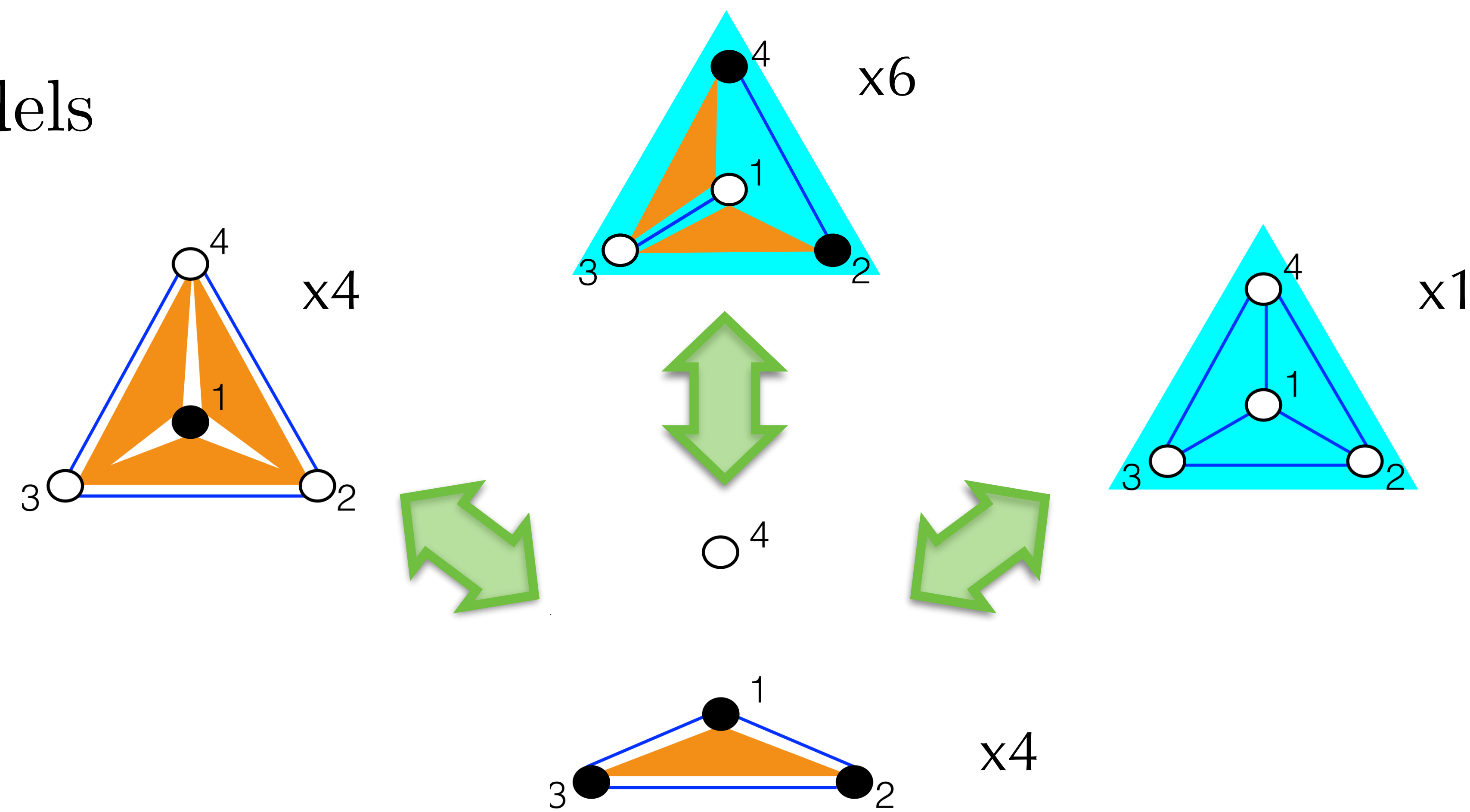
"Complexity Classes"

classes of equivalent models

Ex. 1 class with 15 models

4 spins

7 interactions

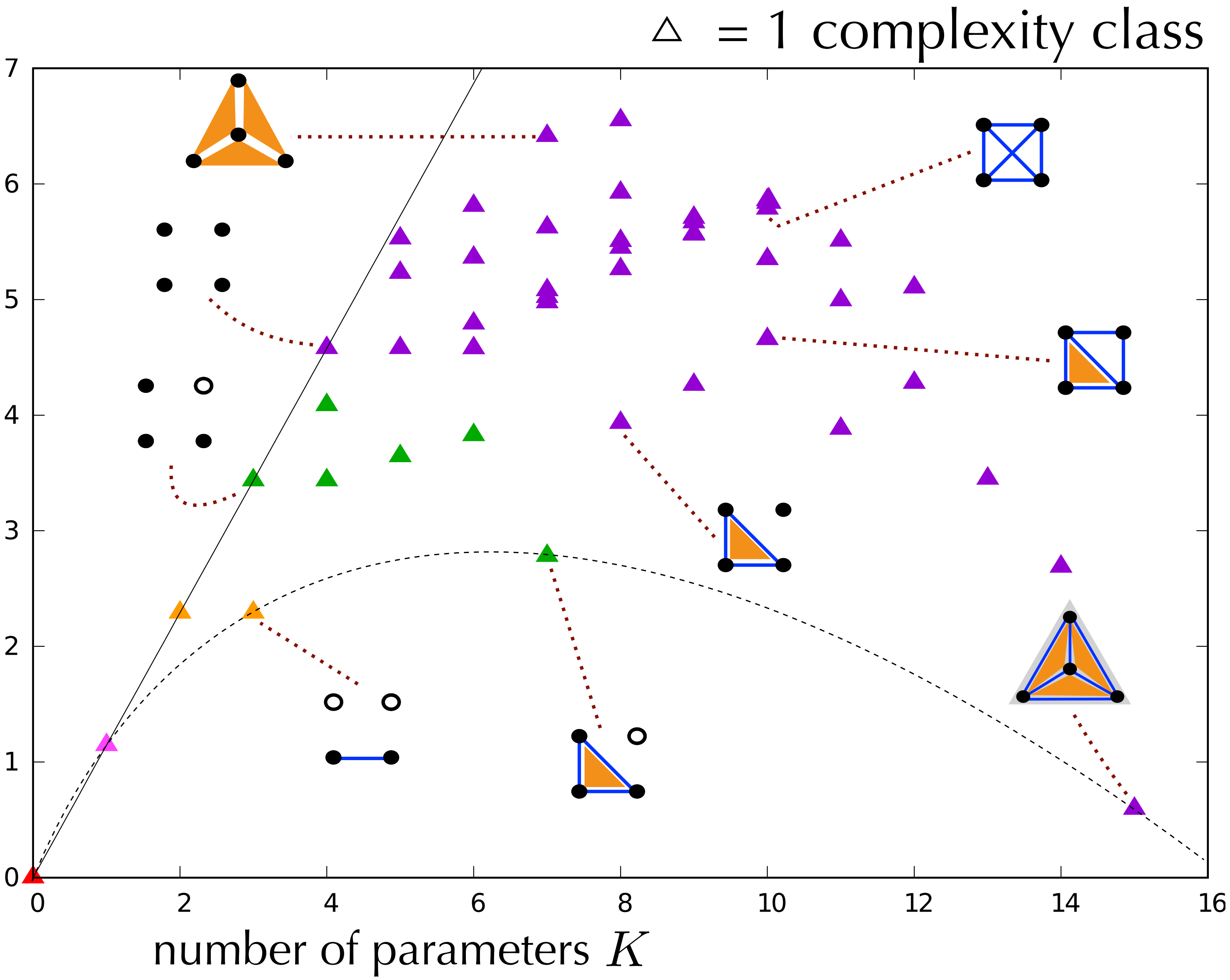


"Same" Model seen in different bases

All the **Same Complexity**

Ex. Complexity for $n = 4$

Ex. $n = 4$ spins



32 768 models, **only 46 classes**

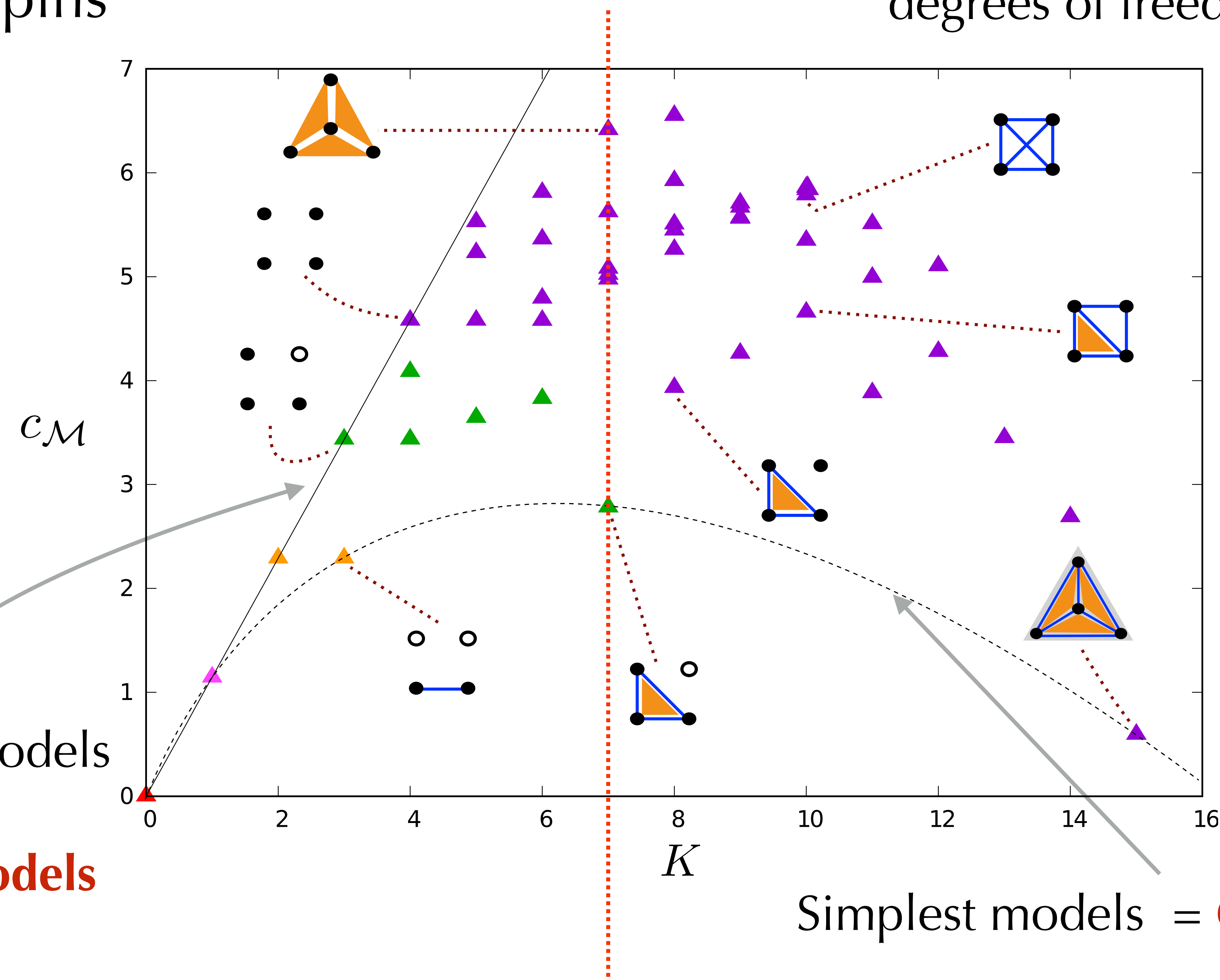
At fixed K?

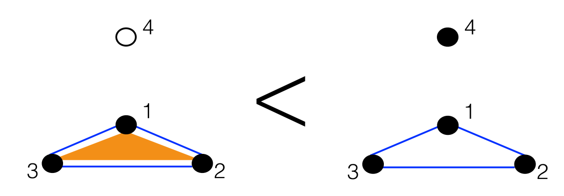
Ex. $n = 4$ spins

$c_{\mathcal{M}}$ increases with the number of degrees of freedom of the models

Degrees of freedom:

- 0
- 1
- 2
- 3





Most complex models = **Independent models**

Simplest models = **Complete models**

At fixed K

Conclusion

Complexity

does not dependent on the order of the interactions

depends on how interactions are arranged in the model

Simplest models?

[At fixed K]

= **the most constraints** between the interactions:

— less degrees of freedom;

— as compact as possible.

Simpler models: implement **more constraints**

As a result, they can account for **less variety of data types.**

Easier to falsify

Some Perspectives…

- **Model selection *within class*:** Compare on Max Log-Likelihood only

- **Change the basis of the data** to facilitate model selection:
  Is there a basis in which the best model would be pairwise?

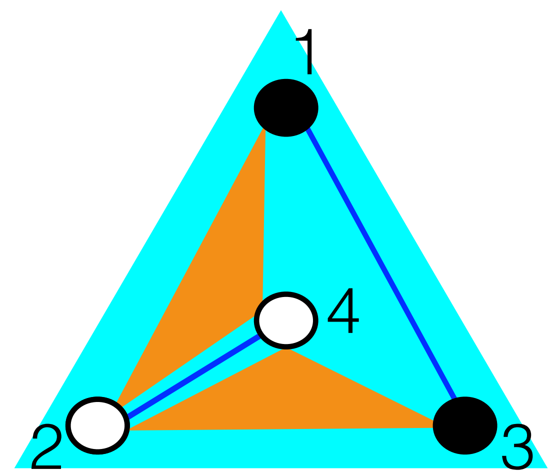
- Model selection among models of ***minimally complex classes***.

- Is the high complexity of pairwise models at the ***Origin of Pairwise Sufficiency?***

[Ref] L. Merchan, I. Nemenman
On the Sufficiency of Pairwise Interactions in Maximum Entropy Models of Networks

So… Which Model is the Simplest?

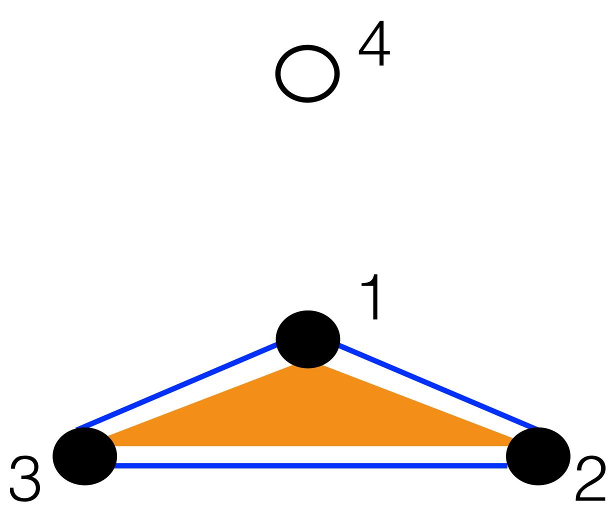
a)



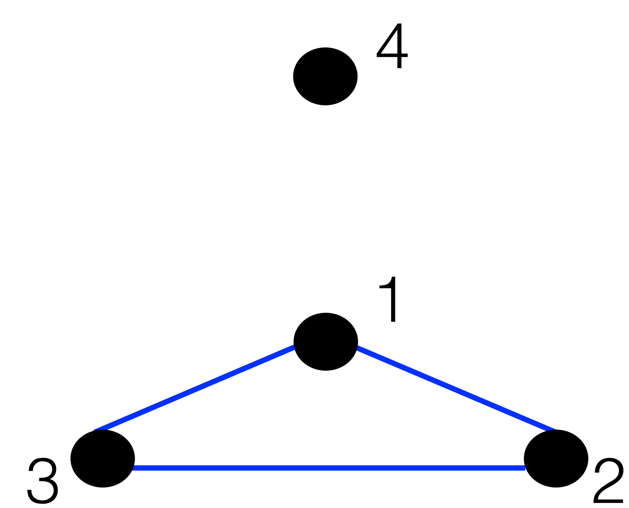
[Model = skeleton]

7 parameters

b)



c)



Questions?

Matteo Marsili

Alberto Beretta

Claudia Battistin

Iacopo Mastromatteo

The Abdus Salam International Centre for Theoretical Physics (ICTP)

The Stochastic Complexity of Spin Models: Are Pairwise Models Really Simple?

Entropy **2018**, *20*(10), 739