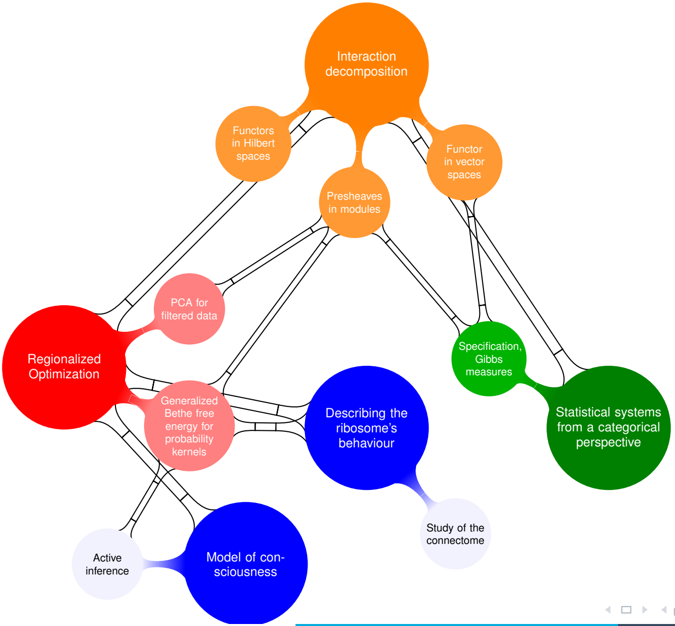
Intersection property, interaction decomposition, regionalized optimization and applications.

Gregoire Sergeant-Perthuis

IMJ-PRG
Université de Paris

PhD Defence, May 28 2021

Overview


Interaction decomposition
Functors in Hilbert spaces
Presheaves in modules
Functor in vector spaces
PCA for filtered data
Regionalized Optimization
Specification, Gibbs measures
Generalized Bethe free energy for probability kernels
Describing the ribosome's behaviour
Statistical systems from a categorical perspective
Study of the connectome
Active inference
Model of consciousness

Motivation

- Critical Brain hypothesis
  - The brain is near criticality, assembly of neurons exhibit criticality [MB11]
- Description of phase diagram, criticality without invariance by translation, without lattice or an a priori notion of space
- Statistical system with a categorical flavour:
  - Controlling the complexity of statistical systems

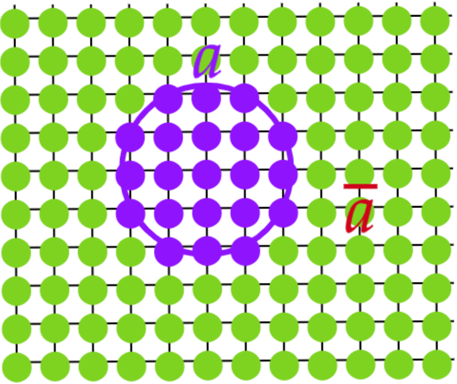
Standard approach to statistical systems and thermodynamic limit: Setting

- I index of (finite) random variables: $(X_i \in E_i, i \in I)$
- Global state space $\Omega = \prod_{i \in I} E_i$ denoted as E with $\sigma-$algebra $\mathscr{E}$,
- $\mathbb{P}(E)$ space of measures
- $\mathscr{H} = C_0(E)$ is the space of observables

- $a \subseteq I$ finite subset of I, $\mathscr{P}_f(I)$ the set of finite subsets
- $X_a \in E_a = \prod_{i \in a} E_i$ or $(E, \mathscr{E}_a)$ when seen in $(E, \mathscr{E})$
- for $b \subseteq a$, $i^a_b : E_a \to E_b$ in $\mathbf{Mes}(E_a, E_b)$
- $\mathrm{U}(a) \subseteq \mathscr{H}$ space of observable that depend only on the X_a
- $i^b_a : \mathrm{U}(b) \hookrightarrow \mathrm{U}(a)$

Standard approach to statistical systems and thermodynamic limit: Prescribed conditional probabilities

- Statistical system: collection of border conditions



- Probability kernel $p : E_{\overline{a}} \to E$,
- $p \in \mathbf{Kern}(E_a, E)$
- $\forall \omega_a \in E_a$, $p_{\omega_a} \in \mathbb{P}(E)$
- For $A \in \mathscr{E}$, $p(A|\omega_a)$ "$\cong$" $\mathbb{E}[A|\mathscr{E}_a]$

Standard approach to statistical systems and thermodynamic limit: Proper kernel

Standard Definition: Proper Kernel [Geo88]

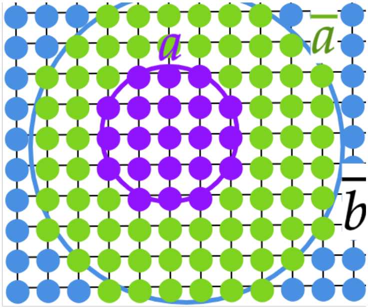
Let $\mathscr{E} \subseteq \mathscr{E}_1$ be two σ-algebras of a set E, a kernel $p \in \textbf{Kern}\,((E,\mathscr{E}_1),(E,\mathscr{E}))$ is proper if and only if, for any $A \in \mathscr{E}$, any $B \in \mathscr{E}_1$ and any $\omega \in E$,

$$p(A \cap B|\omega) = p(A|\omega)1[\omega \in B]$$

For f a $\mathscr{E}_1-$measurable function,

$$p(f|\omega) = \int f(x)p(dx|\omega) = f(\omega)$$

Standard approach to statistical systems and thermodynamic limit: tower rule



- Tower rule: for $A \in \mathscr{E}$,

$$\mathbb{E}\left[\mathbb{E}[A|\mathscr{F}_{\overline{a}}|\mathscr{F}_{\overline{b}}]\right] = \mathbb{E}[A|\mathscr{F}_{\overline{b}}]$$

- For kernels, for $\omega \in E$,

$$p_{\overline{b}}p_{\overline{a}}(A|\omega) = \int p(A|x_{\omega_{\overline{a}}})p(dx|\omega_{\overline{b}})$$

$$p_{\overline{b}}\left(p_{\overline{a}}(A|.)|\omega_{\overline{b}}\right) = p_{\overline{b}}(A|\omega_{\overline{b}})$$

Standard approach to statistical systems and thermodynamic limit: Specification

Standard Definition: Specification [Geo88]

A specification with parameter set I and state spaces $(E, \mathscr{E})$ is a collection $(\gamma_a, a \in \mathscr{P}_f(I))$ of proper kernels such that for any $a \in \mathscr{P}_f(I)$, $\gamma_a \in \textbf{Kern}\,((E, \mathscr{E}_{\overline{a}}), (E, \mathscr{E}))$ and which satisfies that for any $a \subseteq b$, i.e $\overline{b} \subseteq \overline{a}$, any $A \in \mathscr{E}$ and $\omega \in E$,

$$\gamma_b \gamma_a(A|\omega) = \gamma_b(A|\omega)$$

Standard approach to statistical systems and thermodynamic limit: Gibbs measures

Definition: Gibbs measures

Let γ be a specification with state space E, the set of probability measures,

$$\mathscr{G}(\gamma) = \{\mu \in \mathbb{P}(E) \quad : \quad \mathbb{E}_\mu(A|\mathscr{E}_{\overline{a}}) = \gamma_a(A|.)\ \mu \text{ a.s.}\}$$

is called the set of Gibbs measures of γ.

Categories of measurable spaces and probability kernels

- **Mes**: Objects are measurable space, Morphisms are measurable applications
- **Kern**: Objects are measurable spaces, Morphisms are probability kernel
- **Mes** is a subcategory of **Kern**

Remark: One element measurable space

Let $*$ be the one element measurable space, and E any measurable space,

$$\mathbf{Kern}(*, E) \cong \mathbb{P}(E)$$

From classical description to categorical description

Specification *Presheaf*

$$\begin{tikzcd}
& * \arrow[d] \arrow[ddl, bend right] \arrow[ddr, bend left] & \\
& E \arrow[dl, dotted, shift left] \arrow[dr, dotted, shift left] \arrow[dd, dotted, shift left] & \\
E_c \arrow[ur, shift left] & & E_a \arrow[ul, shift left] \\
& E_b \arrow[uu, shift left] &
\end{tikzcd}
\quad \overset{\textit{Proposition } 8.3.1}{\rightsquigarrow} \quad
\begin{tikzcd}
& * \arrow[d] \arrow[ddl, bend right] \arrow[ddr, bend left] & \\
& E \arrow[dl, dotted, shift left] \arrow[dr, dotted, shift left] \arrow[dd, dotted, shift left] & \\
E_c \arrow[ur, shift left] \arrow[rr] \arrow[dr] & & E_a \arrow[ul, shift left] \\
& E_b \arrow[uu, shift left] \arrow[ur] &
\end{tikzcd}$$

- The collection $(i_{\overline{a}}, a \in \mathscr{P}(I))$ encoded the functor of observables

Categorical approach: new definition of Specification

Definition 8.3.6: Specification

Let $\mathscr{A}$ be a poset, a specification is a couple (G, F) of presheaf-functor where $G : \mathscr{A}^{op} \to \mathbf{Mes}$ and $F : \mathscr{A} \to \mathbf{Kern}$ such that for any $a, b \in \mathscr{A}$ with $b \leq a$,

$$G_b^a F_a^b = \mathrm{id}$$

Categorical approach: new definition of Gibbs measures of a specification

Definition 8.3.7: Gibbs measures for specifications

Let $\gamma = (G, F)$ be a specification over $\mathscr{A}$, we shall call the Gibbs measures of γ the sections of F,

$$\mathscr{G}(\gamma) = \{P_a \in \mathbb{P}(F(a)), a \in \mathscr{A} | \quad \forall b \leq a, F_a^b P_b = P_a\}$$

Dictionary of concepts in the categorical setting

Statistical system

- Specification
- Gibbs measure
- Independent random variables

Statistical system in **Kern**

- Couple of presheaf/ functor
- Limit of the functor
- ?

L^∞ presheaf

- $L^\infty(E)$: space of bounded measurable functions
- L^∞ a presheaf **Mes** $\to$ **Vect**
  - For $f \in L^\infty(E)$, $h : E_1 \to E$,

$$L^\infty(h)(f) = f \circ h$$

- L^∞ a presheaf **Kern** $\to$ **Vect**
  - For $f \in L^\infty(E)$, $\omega_1 \in E_1$ and $p : E_1 \to E$,

$$L^\infty(p)(f, \omega_1) = \int f(\omega) p(d\omega|\omega_1)$$

Categorical approach: Decomposable specifications

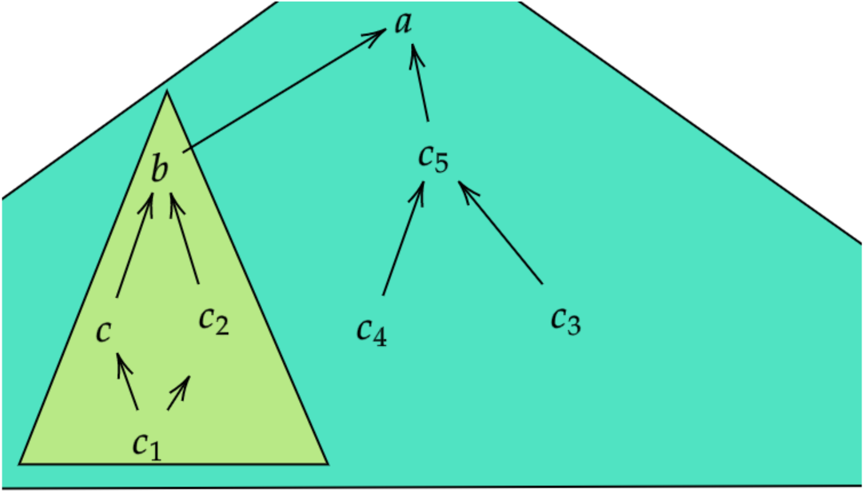
Definition 8.4.1: Decomposability

A specification γ is decomposable if $L^\infty\gamma$ is decomposable, i.e. there is a collection of vector spaces $(S_a, a \in \mathscr{A})$ such that for any $a, b \in \mathscr{A}$ such that $b \leq a$,

$$L^\infty G(a) = L^\infty F(a) \cong \bigoplus_{c \leq a} S_c$$

$$L^\infty G_a^b \cong i_a^b : \bigoplus_{c \leq b} S_c \to \bigoplus_{c \leq a} S_c, \quad L^\infty F_b^a \cong p_b^a : \bigoplus_{c \leq a} S_c \to \bigoplus_{c \leq b} S_c$$

Categorical approach: direct sum of constant functors and presheaves



Results for decomposable specifications

- Decomposable specification imply acyclic functors (Chapter 7) and acyclic presheaves (Chapter 8)
    - Extend Kellerer's result on the marginal problem and cohomological interpretation.
- Characterization of Gibbs measures of decomposable specifications (Chapter 8)
- Independent random variables are a particular case of Decomposable specifications

Why decomposability? When decomposability?

Motivation for decomposable specifications: Factor spaces and Factorization spaces

$E = \prod_{i \in I} E_i$. For $a \subseteq I$, $E_a = \prod_{i \in a} E_i$ and $\pi_a : E \to E_a$

Definition: Factor spaces

The $a-$factor space denotes $\mathrm{U}(a)$ is the set of functions, f, of $\mathbb{R}^E$ that factor through π_a, i.e there is $\tilde{f} \in \mathbb{R}^{E_a}$, $f = \tilde{f}\pi_a$

Particular case of an Interaction decomposition : Decomposition into interaction subspaces

Theorem: Decomposition into interaction subspaces [Spe79][Lau96]

There is a collection of vector subspaces of $\mathbb{R}^E$, $(S_a, a \subseteq \mathscr{P}(I))$, such that, for any $a \subseteq \mathscr{P}(I)$,

$$U(a) = \bigoplus_{b \subseteq a} S_b$$

and any two S_a, S_b, with $a \neq b$, are orthogonal to one another.

How to desbribe dependencies?

- A graphical model is a way to express the interactions between random variables from the connectivity properties of a graph

$$X \text{ —— } Z \text{ —— } Y$$

Gibbs random field

- Hamiltonian
- $H(x, y, z) = \Phi_{1,3}(x, z) + \Phi_{2,3}(y, z)$

Markov random field

- Factorization space
- $P_H = f(x, z)g(y, z)$
- $X \perp\!\!\!\perp Y | Z$

$$P(x, y, z) = \frac{e^{-\beta H(x,y,z)}}{\sum_{(x,y,z) \in X \times Y \times Z} e^{-\beta H(x,y,z)}}$$

$\mathscr{B}$ potential space, $\mathscr{B}$ factorization space

$\mathscr{B}$-potential

$$\mathrm{U}(\mathscr{B}) = \sum_{a \in \mathscr{B}} \mathrm{U}(a)$$

$\mathscr{B}$ factorization space,

$$\mathscr{F}_{\mathscr{B}} = exp(\mathrm{U}(\mathscr{B}))$$

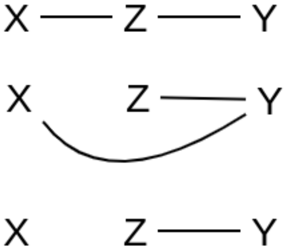
- We can restric our attention to lower sets.
- For $\mathscr{B} \subseteq \mathscr{P}(I)$, $\hat{\mathscr{B}} = \{b \in \mathscr{P}(I) : \ \exists a \in \mathscr{B}, b \subseteq a\}$
- If $\hat{\mathscr{B}} = \mathscr{B}$, $\mathscr{B}$ is a lower set of $\mathscr{P}(I)$.
- Sets of lower set: $\mathscr{U}(\mathscr{P}(I))$.

Graphoid intersection property

Proposition: Graphoid intersection property

Let X, Y, Z be three random variables that take values in a finite set and for which the probability density $P_{X,Y,Z}$ is strictly positive, then,

$$X \perp\!\!\!\perp Y|Z \text{ and } X \perp\!\!\!\perp Z|Y \implies X \perp\!\!\!\perp (Y, Z)$$



- $X \to 1$, $Y \to 2$, $Z \to 3$
- $\mathscr{A} = \{\{1,3\},\{2,3\}\}$
- $X \perp\!\!\!\perp Y|Z \iff P_{X,Y,Z} \in \mathscr{F}_{\mathscr{A}}$
- $\mathscr{B} = \{\{1,2\},\{2,3\}\}$
- $X \perp\!\!\!\perp Z|Y \iff P_{X,Y,Z} \in \mathscr{F}_{\mathscr{B}}$
- $\mathscr{F}_{\hat{\mathscr{A}}} \cap \mathscr{F}_{\hat{\mathscr{B}}} \subseteq \mathscr{F}_{\hat{\mathscr{A}} \cap \hat{\mathscr{B}}}$

Weaker Intersection property: Extension for factorization spaces

Proposition [SP]: Weaker Intersection property (for factor spaces)

For any collection of lower sets of $\mathscr{P}(I)$, $(\mathscr{B}_j, j \in J)$,

$$\bigcap_{j \in J} \mathrm{U}(\mathscr{B}_j) = \mathrm{U}(\bigcap_{j \in J} \mathscr{B}_j)$$

- Reducing the proof of the Hammersley-Clifford Theorem to a property of graphs. (Chapter 2 or [SP19a])

Interaction decomposition: Functors from a poset to **Vect**

- **Gr** V the poset of vector subspaces of a vector space V
- $[\mathscr{A}, \mathbf{Gr}\ V]$ the collection of increasing functions from $\mathscr{A}$ to **Gr** V

Definition 3.3.1: Decomposable collection of vector subspaces

$\mathrm{U} \in [\mathscr{A}, \mathbf{Gr}\ V]$ is decomposable if and only if there is a collection of vector subspaces $(S_a \subseteq V, a \in \mathscr{A})$ such that for any $a \in \mathscr{A}$

$$\mathrm{U}(a) \cong \bigoplus_{b \leq a} S_b$$

and for $b \in \mathscr{A}$ with $b \leq a$, U_a^b is isomorphic to the inclusion $\bigoplus_{c \leq b} S_c \to \bigoplus_{c \leq a} S_c$. We will call $(S_a, a \in \mathscr{A})$ a (interaction) decomposition of U.

Interaction decomposition : Functors from a poset to **Vect**, equivalence theorem

Definition 3.3.3: Intersection property

Let $\mathscr{A}$ be any poset, an increasing function $\mathrm{U} \in [\mathscr{A}, \mathbf{Gr}\ V]$ is said to verify the intersection property (I) if and only if,

$$\forall \mathscr{B}, \mathscr{C} \in \mathscr{U}(\mathscr{A}), \quad \sum_{b \in \mathscr{B}} \mathrm{U}(b) \cap \sum_{c \in \mathscr{C}} \mathrm{U}(c) \subseteq \sum_{a \in \mathscr{B} \cap \mathscr{C}} \mathrm{U}(a) \qquad \text{(I)}$$

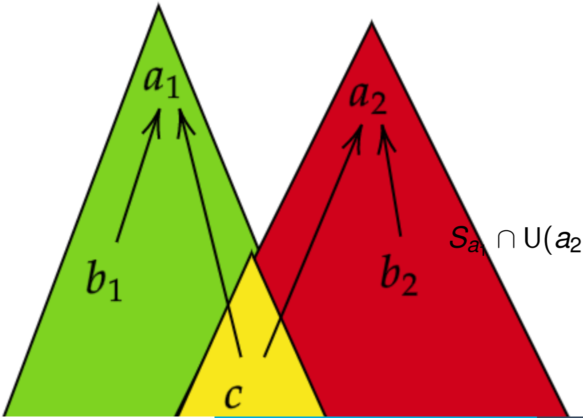
Theorem [SP]: Equivalence theorem

If $\mathscr{A}$ is a well-founded poset, $\mathrm{U} \in [\mathscr{A}, \mathbf{Gr}\ V]$ is decomposable if and only if U verifies (I).

- Extension to functors from $\mathscr{A}$ to **Vect** [SP19b], Chapter 3 Proposition 3.2.1

Intersection property: a key element for building interaction decompositions

$$\forall a \in \mathscr{A},\ \mathrm{U}(a) = S_a \oplus \sum_{b \lneq a} \mathrm{U}(b)$$



$$S_{a_1} \cap \mathrm{U}(a_2) \subseteq \sum_{b \lneq a_1} \mathrm{U}(b)$$

More around the interaction decomposition

- Interaction decomposition and intersection property for presheaves from a poset to the category of modules (Equivalence Theorem 4.4.1 Chapter 4, [SP20])
- Interaction decomposition and intersection property for functors from a poset to the the category of Hilbert spaces, with morphisms isomorphisms
  Equivalence Theorem 5.4.1 Chapter 5 $\implies$ Generalization of Chaos decomposition

Interaction decomposition for presheaves: Definition of decomposability

- When a collection of vector subspace is decomposable there are several decompositions possible.
- Additional data of collection of projectors to distinguish them.

Definition 4.2.4: Decomposable collection of projectors

Let $\mathrm{U} \in [\mathscr{A}, \mathbf{Gr}\, V]$ be decomposable, let $(\pi_a, a \in \mathscr{A})$ be a collection of projectors onto the $\mathrm{U}(a)$; this family is decomposable if and only if there is a decomposition of U, $(S_a, a \in \mathscr{A})$, such that for any $b \leq a$,

$$\pi|_{\mathrm{U}(a)}^{\mathrm{U}(b)} \cong p_b^a$$

where p_b^a is the projection of $\bigoplus_{c \leq a} S_c$ onto $\bigoplus_{c \leq b} S_c$.

Interaction decomposition for presheaves: a particular case when considering Meet semi-lattices

Definition: Meet semi-lattice

Let $\mathscr{A}$ be a poset, $a, b \in \mathscr{A}$. $\mathscr{A}$ has a meet for (a, b) when there is d such that,

$$\forall c \in \mathscr{A}, \quad c \leq a \quad \& \quad c \leq b \implies c \leq d$$

d is unique and we shall note it $a \cap b$.
We will call meet semi-lattice any poset that has all meets for any couple.

Interaction decomposition for presheaves: Equivalence theorem

Definition 4.3.6: Intersection property for collection of projectors

Let $\mathscr{A}$ be a finite meet semi-lattice, and let $(\pi_a, a \in \mathscr{A})$ be a collection of projectors. This collection satisfies the intersection property when,

$$\pi_a \pi_b = \pi_{a \cap b} \tag{I}$$

Theorem [SP]: Equivalence theorem

A collection of projectors is decomposable if and only if it satisfies the intersection property

Particular case of decomposable specification

- $\mathbb{P} \in \mathbb{P}(E)$ defines a collection of projectors $\mathbb{E}[\cdot|\, \mathrm{U}(a)], a \subseteq I$

Corollary 4.3.2: Interaction Decompositions for factor spaces

Let I be a finite set, $(E_i, i \in I)$ a collection of finite sets, and $\mathbb{P}$ a probability measure on E, $(\mathbb{E}_a[\cdot|\mathscr{F}_a], a \in \mathscr{P}(I))$ is decomposable if and only if $\mathbb{P}$ is a product measure, i.e if there is $(p_i \in \mathbb{P}(E_i), i \in I)$ such that $\mathbb{P} = \bigotimes_{i \in I} p_i$.

- Independent statistical systems are a particular case of decomposable specifications

Dictionary of concepts in the categorical setting

Statistical system

- Specification
- Gibbs measure
- Independent random variables

Statistical system in **Kern**

- Couple of presheaf/ functor
- Limit of the functor
- Decomposable specifications

Bayesian point of view

Bayesian point of view:

- Θ space of parameters, Ω space of the observations
- A kernel $p : \Theta \to \Omega$: for $\theta \in \Theta$, $p_\theta \in \mathbb{P}(\Omega)$
- A prior: $P_0 \in \mathbb{P}(\Theta)$
- Update of beliefs: posterior after making an observation ω,

$$P(\theta|\omega) = \frac{P_0(\theta)p_\theta(\omega)}{\sum_{\theta\in\Theta} P_0(\theta)p_\theta(\omega)}$$

Variational inference

- Approximate the posterior
- Natural notion of length between two distributions $Q, P \in \mathbb{P}(\theta)$: Relative Entropy or Kullback-Leibler divergence

$$S[Q|P] = \sum_{\theta \in \Theta} Q(\theta) \ln \frac{Q(\theta)}{P(\theta)}$$

- In Statistical Mechanics: free energy with respect to a Hamiltonian H and at temperature T,

$$F(Q) = \mathbb{E}_Q[H] - TS(Q)$$

- Find Q in a family of probability distributions that minimizes $S[Q|P]$

Region-based free energy approximation, a motivation for Regionalized optimization

Definition: Region-based free energy approximation [YFW05] or Generalized Bethe free energy

Let I be a finite set and let $E = \prod_{i \in I} E_i$ be a product of finite sets and $\mathscr{A}$ a subposet of $\mathscr{P}(I)$. Yedidia, Freeman, Weiss consider for collections $Q = (Q_a \in \mathbb{P}(E_a), a \in \mathscr{A})$ of measures compatible by marginalization, a free energy built from the entropy of each probability measure Q_a,

$$F_{Bethe}(Q) = \sum_{a \in \mathscr{A}} c(a) \left(\mathbb{E}_{Q_a}[H_a] - S(Q_a) \right) \tag{0.1}$$

with $(H_a \in \mathrm{U}(a), a \in \mathscr{A})$ a collection of Hamiltonians.

Regionalized optimization: Optimization with constraints given by a presheaf

- Generalization for presheaves from a poset to the category of finite vector spaces
- Simple algorithm for finding their critical points when the presheaf is decomposable
- PCA for filtered data
- Free energy approximation for diagrams in **Kern**

Contributions

- Chapter 2: Weak Intersection property for factor spaces, reducing Hammersley-Clifford theorem to a property of graphs.
  Theorem 2.4.1, Corollary 2.5.2
- Chapter 3: Intersection property is equivalent to the existence of an interaction decomposition for injective functors from a well founded poset to the category of vector spaces
  Theorem 3.5.1
- Chapter 4: Equivalence theorem for presheaves in the category of modules and description of interaction decompositions for factor spaces
  Theorem 4.4.1, Corollary 4.3.2
- Chapter 5: Equivalence theorem for functors in the category of Hilbert spaces with morphisms isometries (generalization of the Chaos decomposition)
  Theorem 5.4.1

Contributions

- Chapter 6, with D. Bennequin, O. Peltre, J.P. Vigneaux : Extra-fine sheaves, their acyclicity, homological interpretation and extension of Kellerer's result for the marginal problem (injective functor case)
  Theorem 6.4.3, Theorem 6.5.3
- Chapter 7: acyclicity of decomposable presheaves
  Theorem 7.2.1
- Chapter 8: Reformulation of Gibbs measures for diagrams over a poset in the category of probability kernels, characterization for decomposable specification.
  Theorem 8.5.1

Contributions

- Chapter 9: Formulation of a global optimization problem from a collection of local ones; applications are a regionalized version of the PCA for data provided with a filtration, an extension of the free energy underlying the General Belief Propagation to diagrams over a poset in the category of probability kernels. When the presheaf is decomposable we provide a simple algorithm for finding the critical points.
  Theorem 9.2.1, Theorem 9.2.3, Proposition 9.3.3
- Chapter 10 with Y. Timsit and D. Bennequin: statistical properties of the graph of the ribosome, definition of a conditional statistical test.
  Theorem 10.6.1

Sincere Thanks

I would like to thank very much Daniel Bennequin for his supervision,

I would like to thank Olivier Peltre and Juan Pablo-Vigneaux for our brothership for these past and I hope future years.

I would like to thank Youri Timsit, David Rudrauf, Olivier Belli, Yvain Tisserand, Kaie Kubjas, Elias Tsigaridas for our stimulating collaborations.

Finally I would like to thank Frédéric Hélein for his constant support during these years and Christian Lorenzi for permitting this PhD to happen.

Thank you very much for your attention

Thank you very much for your attention!

References I

Hans-Otto Georgii, *Gibbs measures and phase transitions*, de Gruyter, 1988.

Steffen L. Lauritzen, *Graphical models*, Oxford Science Publications, 1996.

Thierry Mora and William Bialek, *Are biological systems poised at criticality?*, Journal of Statistical Physics volume (2011).

Grégoire Sergeant-Perthuis, *Bayesian/graphoid intersection property for factorisation models*, arXiv:1903.06026v1, 2019.

Grégoire Sergeant-Perthuis, *Intersection property and interaction decomposition*, arXiv:1904.09017v2, 2019.

Grégoire Sergeant-Perthuis, *Interaction decomposition for presheaves*, arXiv:2008.09029, 2020.

References II

Terry P. Speed, *A note on nearest-neighbour gibbs and markov probabilities*, Sankhyā: The Indian Journal of Statistics, Series A (1979).

Jonathan S Yedidia, William T Freeman, and Yair Weiss, *Constructing free-energy approximations and generalized belief propagation algorithms*, IEEE Transactions on information theory **51** (2005), no. 7, 2282–2312.