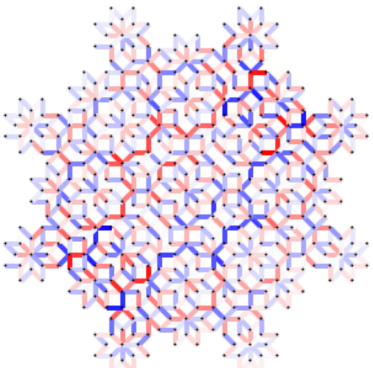
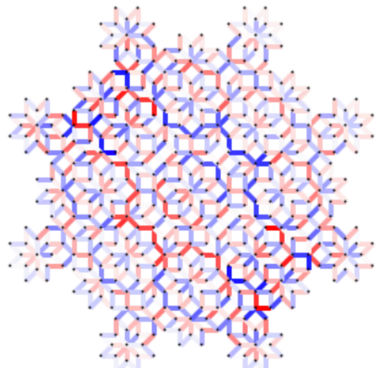
Operator content of quantum and classical lattice models from lossy compression theory

Maciej Koch-Janusz





University of Zurich UZH

MARIE CURIE ACTIONS

THE UNIVERSITY OF CHICAGO

HAIQU

Zohar Ringel | Sebastian Huber | Doruk Efe Gokmen | Matthew Schmitt | Michel Fruchart | Vincenzo Vitelli






האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

ETH zürich

THE UNIVERSITY OF CHICAGO

- **MKJ** and Zohar Ringel, *Nature Physics* ***14****, 578-582 (2018)*
- P. Lenngenhager, D.E. Gokmen, Z. Ringel, S.D. Huber and **MKJ**, *Phys. Rev. X* ***10****, 011037 (2020)*
- A. Gordon, A. Banerjee, **MKJ** and Z. Ringel, *Phys. Rev. Lett.* ***126****, 240601 (2021)*
- D.E. Gokmen, Z. Ringel, S.D. Huber and **MKJ**, *Phys. Rev. Lett.* ***127****, 240603 (2021)*
- D.E. Gokmen, Z. Ringel, S.D. Huber and **MKJ**, *Phys. Rev. E* ***104****, 064106 (2021)*

- D.E. Gokmen, S. Biswas, S.D. Huber, Z. Ringel, F. Flicker and **MKJ**, *arXiv:2301:11934 (2023)* *“ML for physical sciences” @NeurIPS 2023*
- L. Oppenheim, **MKJ**, S. Gazit, Z. Ringel *arXiv:2311.17994, "ML for physical sciences" @NeurIPS 2023*
- M. Schmitt, **MKJ**, M. Fruchart, D. Seara, V. Vitelli *arXiv:2312.06608, “ML for physical sciences” @NeurIPS 2023*

https://github.com/RSMI-NE/RSMI-NE

Outline

- Intro and motivation: RG, compression, and ML

- Information Bottleneck lossy compression
- Equivalence of “relevance” in IB and in RG

- The RSMI coarse-graining
- Example 1: operators and symmetries with RSMI
- Example 2: RG for quasi-periodic tilings
- Example 3: Dynamical systems

Effective descriptions

Effective descriptions

- We seek effective descriptions of **complex** systems in terms of **few variables**

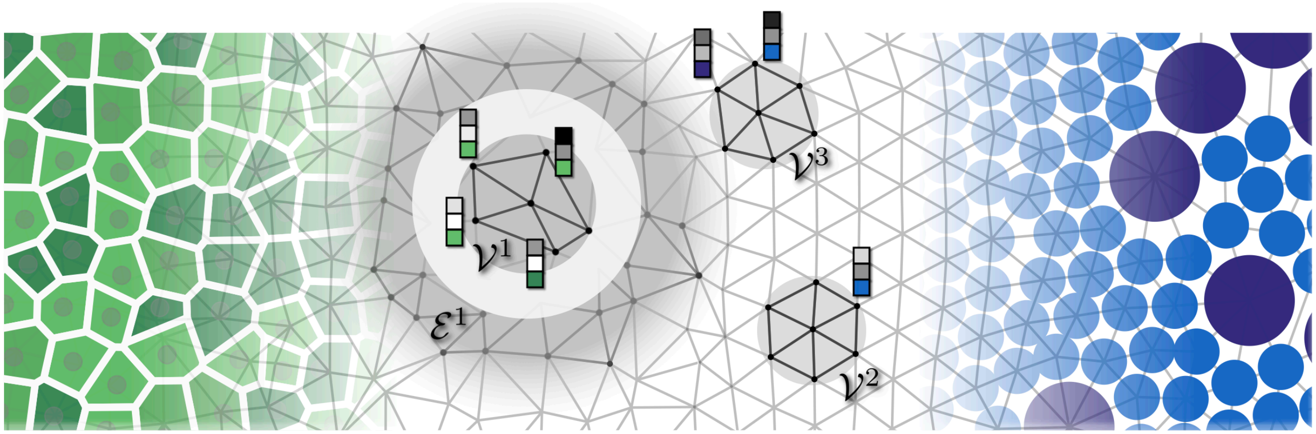
Effective descriptions

- We seek effective descriptions of **complex** systems in terms of **few variables**

- **Coarse-graining** connects microscopic and continuum descriptions

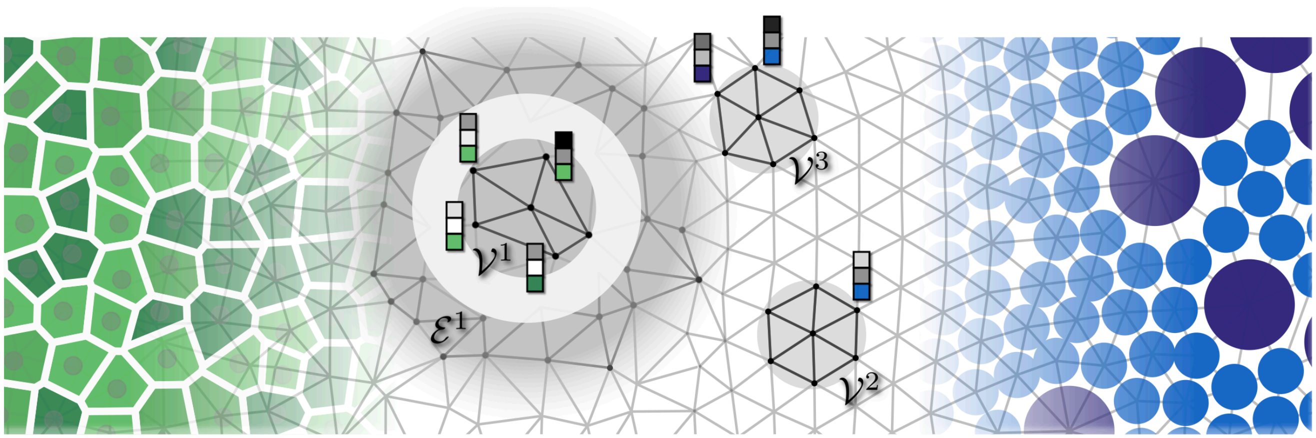
Effective descriptions

- We seek effective descriptions of **complex** systems in terms of **few variables**
- **Coarse-graining** connects microscopic and continuum descriptions

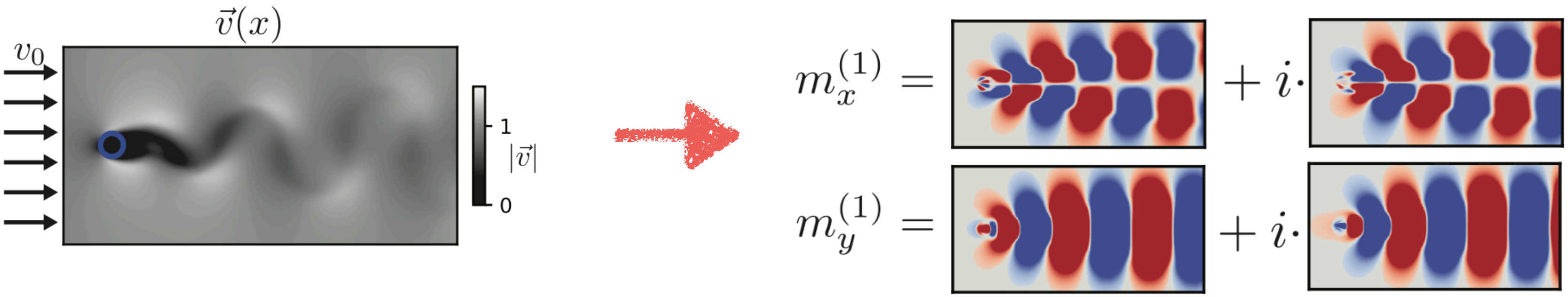


Effective descriptions

- We seek effective descriptions of **complex** systems in terms of **few variables**
- **Coarse-graining** connects microscopic and continuum descriptions



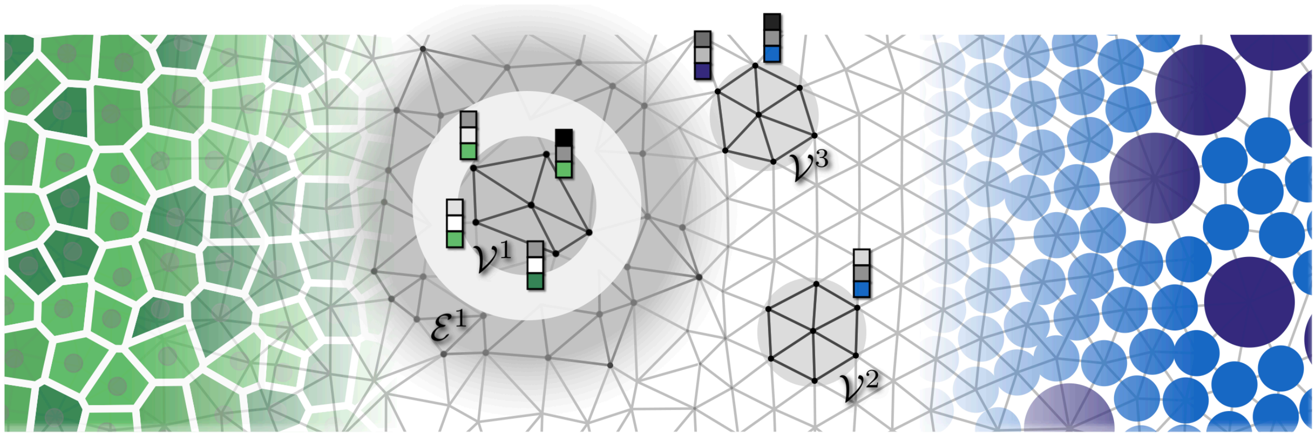
- Dynamical systems' model reductions



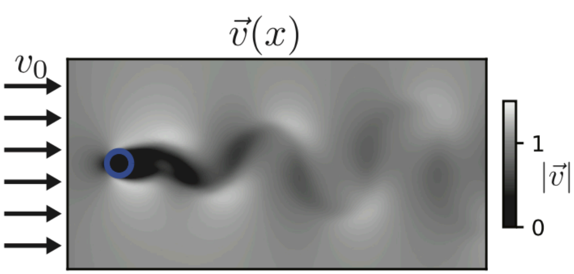
Effective descriptions

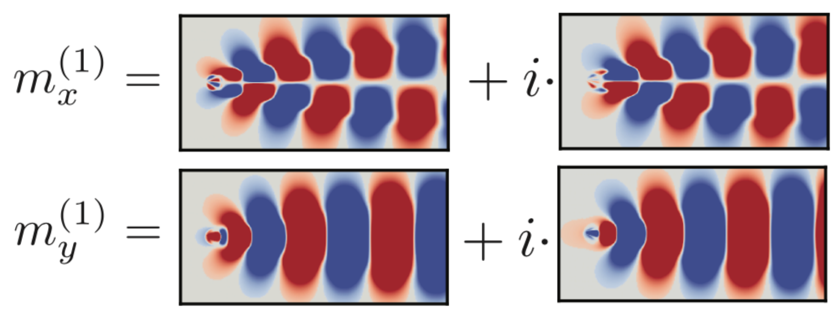
- We seek effective descriptions of **complex** systems in terms of **few variables**

- **Coarse-graining** connects microscopic and continuum descriptions



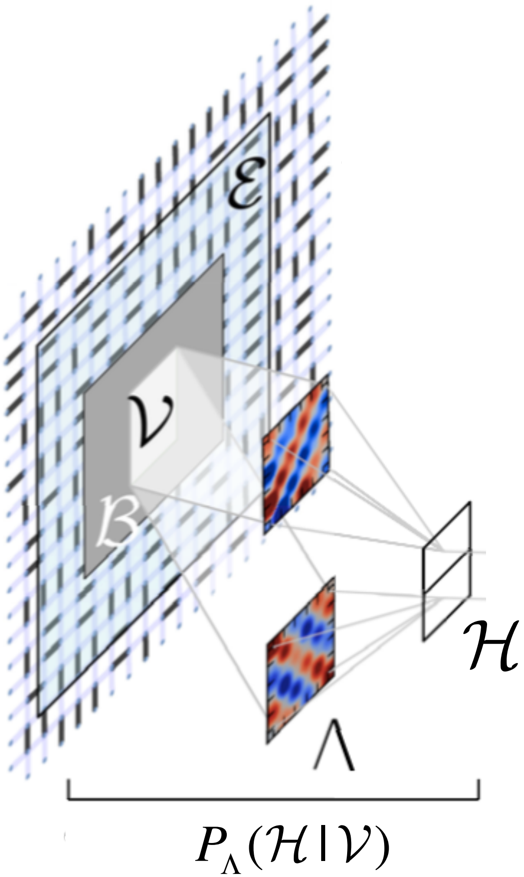
- Dynamical systems' model reductions



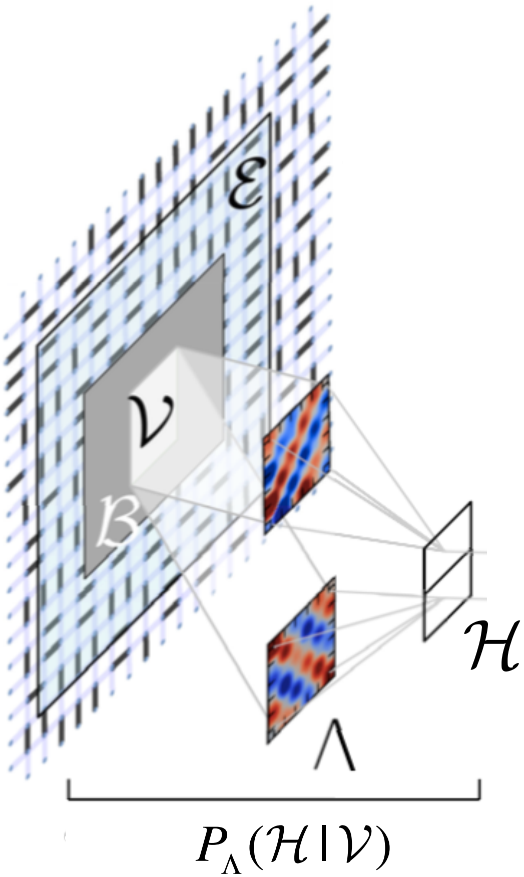


- Can one **systematically** derive an effective theory?

Motivating example: real-space RG

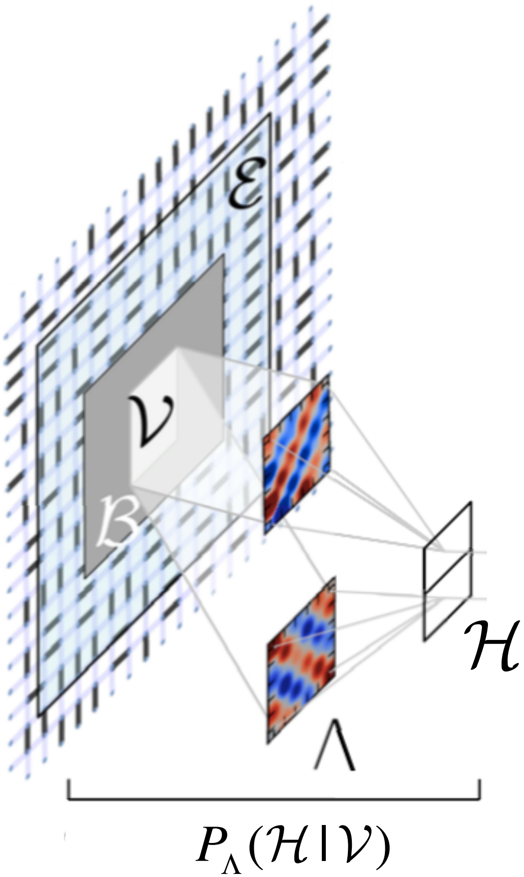


Motivating example: real-space RG



$$e^{\mathcal{K}'(\{\mathcal{H}_i\}_i)} \leftarrow P(\{H_i\}_i) = \sum_i e^{\mathcal{K}(\{\mathcal{V}_i\}_i)} P(\mathcal{H}_i|\mathcal{V}_i)$$

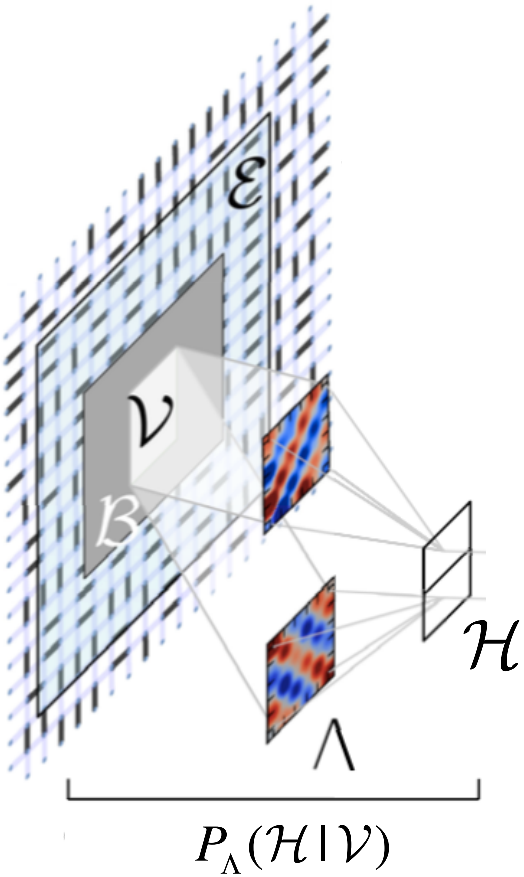
Motivating example: real-space RG



$$e^{\mathcal{K}'(\{\mathcal{H}_i\}_i)} \leftarrow P(\{H_i\}_i) = \sum_i e^{\mathcal{K}(\{\mathcal{V}_i\}_i)} P(\mathcal{H}_i|\mathcal{V}_i)$$

Task: find $P_\Lambda(\mathcal{H}|\mathcal{V})$ such that $\mathcal{H}$ tracks the most relevant degrees of freedom within region $\mathcal{V}$

Motivating example: real-space RG

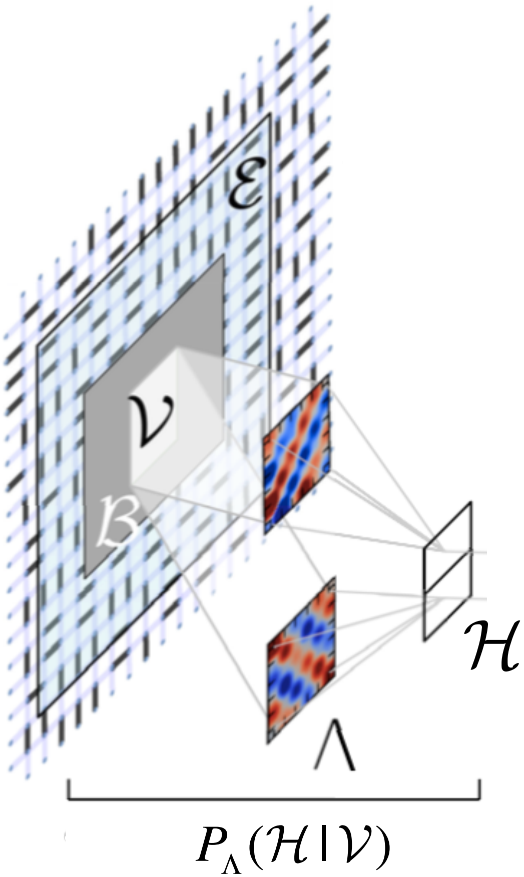


$$e^{\mathcal{K}'(\{\mathcal{H}_i\}_i)} \leftarrow P(\{H_i\}_i) = \sum_i e^{\mathcal{K}(\{\mathcal{V}_i\}_i)} P(\mathcal{H}_i|\mathcal{V}_i)$$

Task: find $P_\Lambda(\mathcal{H}|\mathcal{V})$ such that $\mathcal{H}$ tracks the most relevant degrees of freedom within region $\mathcal{V}$

Method: find max$[I_\Lambda(\mathcal{H}:\mathcal{E})]$ over params Λ

Motivating example: real-space RG



$$e^{\mathcal{K}'(\{\mathcal{H}_i\}_i)} \leftarrow P(\{H_i\}_i) = \sum_i e^{\mathcal{K}(\{\mathcal{V}_i\}_i)} P(\mathcal{H}_i|\mathcal{V}_i)$$

Task: find $P_\Lambda(\mathcal{H}|\mathcal{V})$ such that $\mathcal{H}$ tracks the most relevant degrees of freedom within region $\mathcal{V}$

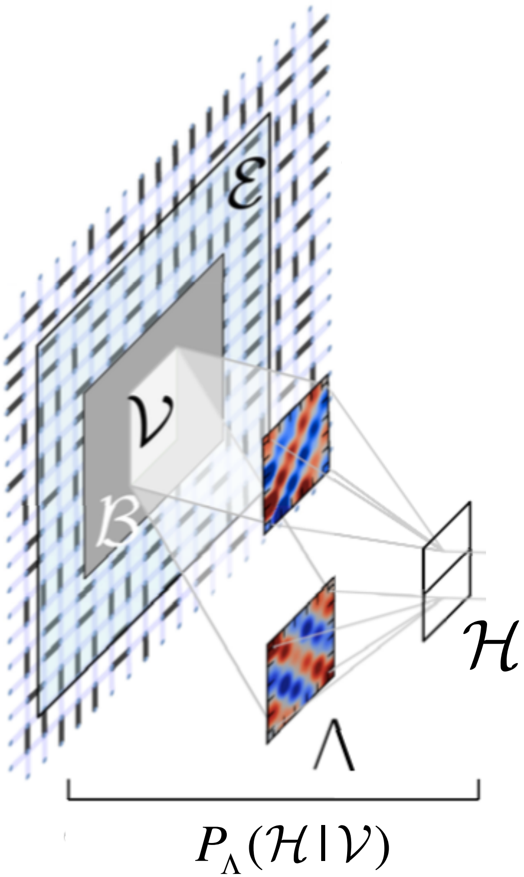
Method: find max[$I_\Lambda(\mathcal{H}:\mathcal{E})$] over params Λ

The Real-Space Mutual Information (RSMI)

Nature Physics ***14****, 578-582 (2018)*

Phys. Rev. X ***10****, 011037 (2020)*

Motivating example: real-space RG



$$e^{\mathcal{K}'(\{\mathcal{H}_i\}_i)} \leftarrow P(\{H_i\}_i) = \sum_i e^{\mathcal{K}(\{\mathcal{V}_i\}_i)} P(\mathcal{H}_i|\mathcal{V}_i)$$

Task: find $P_\Lambda(\mathcal{H}|\mathcal{V})$ such that $\mathcal{H}$ tracks the most relevant degrees of freedom within region $\mathcal{V}$

Method: find max$[I_\Lambda(\mathcal{H}:\mathcal{E})]$ over params Λ

The Real-Space Mutual Information (RSMI)

Nature Physics ***14****, 578-582 (2018)*

Phys. Rev. X ***10****, 011037 (2020)*

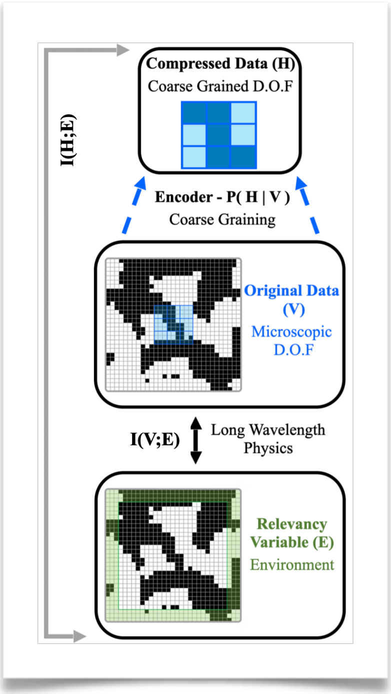
Insight: the optimal $P_\Lambda(\mathcal{H}|\mathcal{V})$ give access to the RG-relevant operators

Phys.Rev.Lett.126, 240601 (2021)

The three ingredients

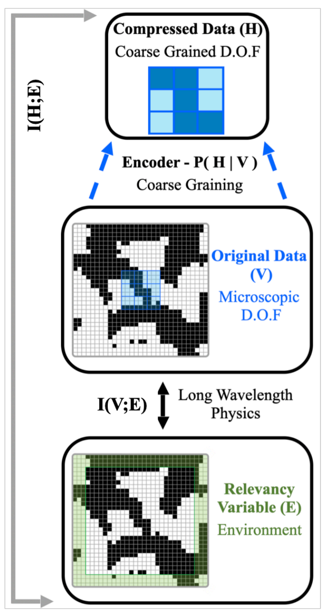
The three ingredients

- The physical principle:
  lossy compression
  maximising **I(H:E)**

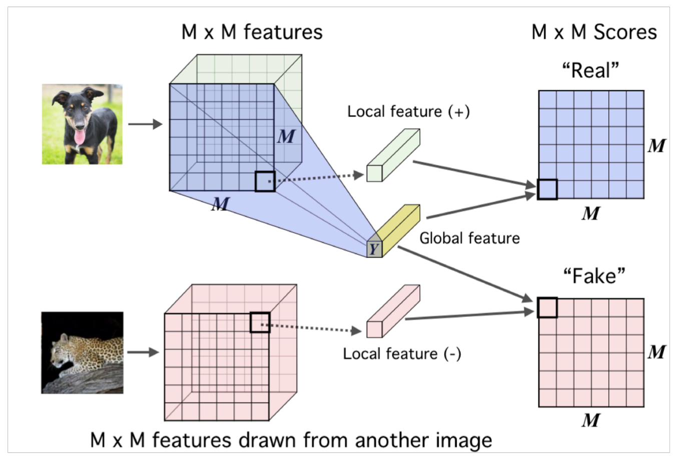


The three ingredients

- The physical principle: **lossy compression** maximising **I(H:E)**



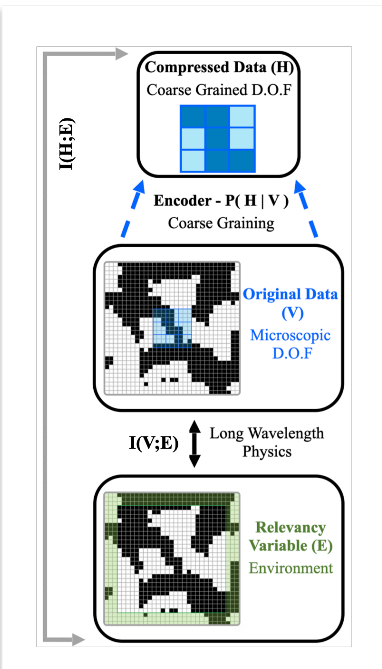
- The estimator of mutual information

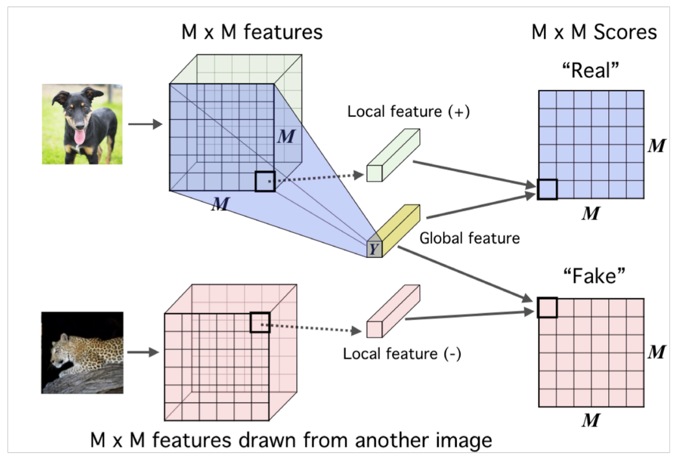


$$D_{KL}(\mathbb{P} \,||\, \mathbb{Q}) = \sup_{T:\Omega \to \mathbb{R}} \mathbb{E}_{\mathbb{P}}[T] - \log(\mathbb{E}_{\mathbb{Q}}[e^T])$$

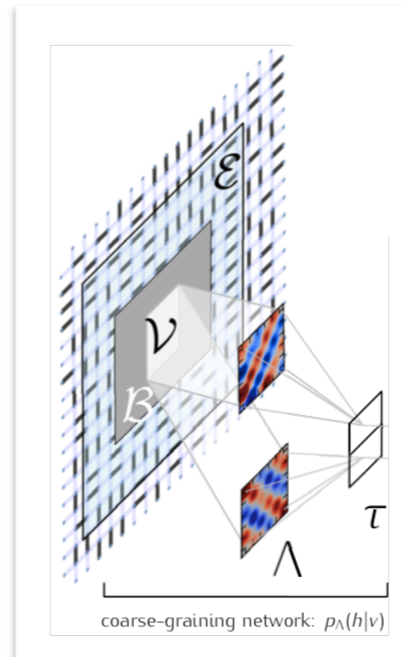
The three ingredients

- The physical principle: **lossy compression** maximising **I(H:E)**
- The estimator of mutual information
- The coarse-graining ansatz family **P(H I V)**





$$D_{KL}(\mathbb{P} \,||\, \mathbb{Q}) = \sup_{T:\Omega\to\mathbb{R}} \mathbb{E}_{\mathbb{P}}[T] - \log(\mathbb{E}_{\mathbb{Q}}[e^T])$$



https://github.com/RSMI-NE/RSMI-NE

Lossy compression

Lossy compression

We have a complicated signal ***V***

dim $d = N_{pixels}$

Lossy compression

We have a complicated signal ***V***

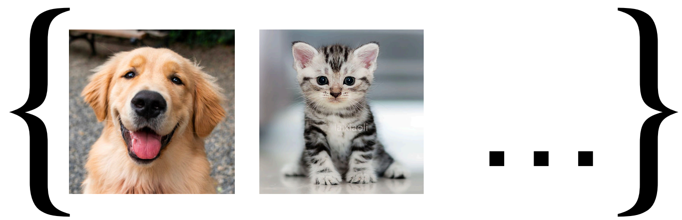
dim $d = N_{pixels}$

We'd like to **compress** it to a variable ***H,*** using a mapping **p(h I v)**

dim $d = 4 \ll N_{px}$

Lossy compression

We have a complicated signal ***V***

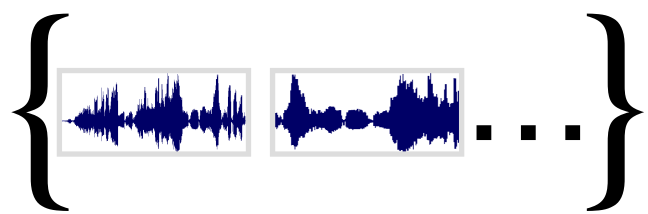


dim $d = N_{pixels}$

We'd like to **compress** it to a variable ***H,*** using a mapping **p(h l v)**

dim $d = 4 \ll N_{px}$

So that ***H*** retains **relevant** information for the down-stream task, **implicitly defined** by correlations with ***Y***



The information bottleneck (IB) compression

The information bottleneck (IB) compression

- **Relevance** defined implicitly, by correlations with a signal variable

The information bottleneck (IB) compression

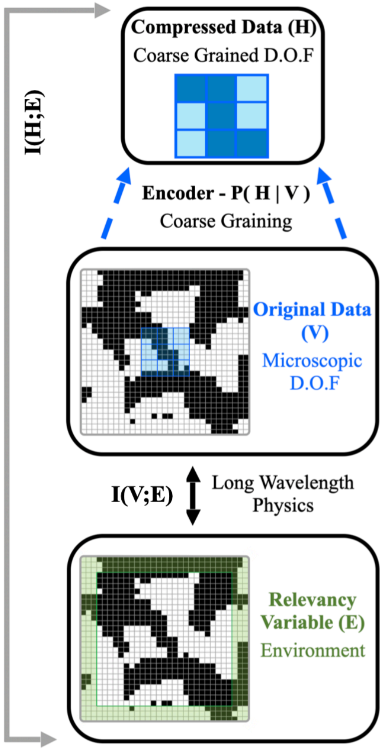
- **Relevance** defined implicitly, by correlations with a signal variable

Tishby, Pereira, Bialek (1999)
Tishby Slonim NeurIPS (2000)

The information bottleneck (IB) compression

- **Relevance** defined implicitly, by correlations with a signal variable

Tishby, Pereira, Bialek (1999)
Tishby Slonim **NeurIPS** (2000)

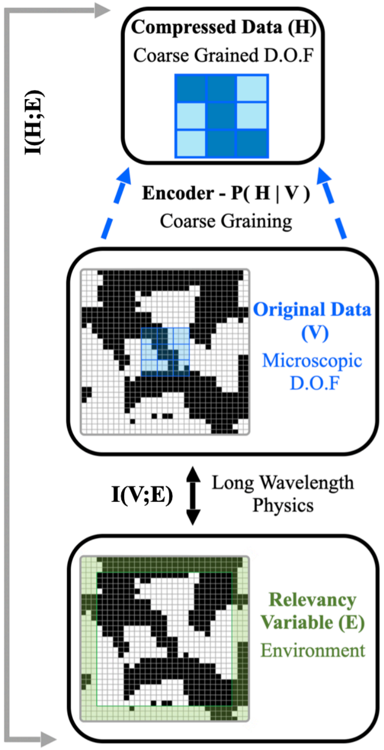


The information bottleneck (IB) compression

- **Relevance** defined implicitly, by correlations with a signal variable
- Optimal compression of relevant information is a variational problem

Tishby, Pereira, Bialek (1999)
Tishby Slonim **NeurIPS** (2000)

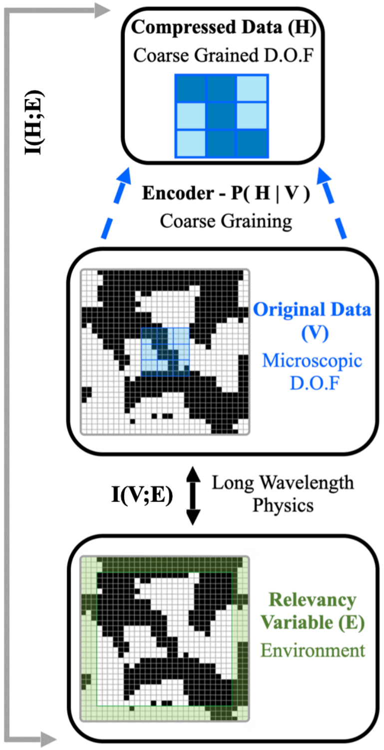
$$\min_{P(H|V)} \mathcal{L}_{IB}[P(H|V)] \equiv \min_{P(H|V)} I(V;H) - \beta I(H;E)$$



The information bottleneck (IB) compression

- **Relevance** defined implicitly, by correlations with a signal variable
- Optimal compression of relevant information is a variational problem

Tishby, Pereira, Bialek (1999)
Tishby Slonim **NeurIPS** (2000)



$$\min_{P(H|V)} \mathcal{L}_{IB}[P(H|V)] \equiv \min_{P(H|V)} I(V;H) - \beta I(H;E)$$

- IB equations:

$$P(h|v) = \frac{P(h)}{Z} \exp\left(-\beta \sum_e P(e|v) \log\left[\frac{P(e|v)}{P(e|h)}\right]\right)$$

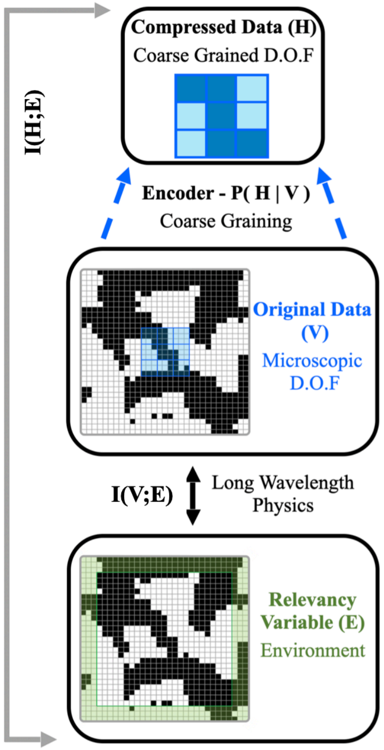
$$P(h) = \sum_v P(h|v)P(v)$$

$$P(e|h) = \sum_v P(e|v)P(v|h)$$

The information bottleneck (IB) compression

- **Relevance** defined implicitly, by correlations with a signal variable
- Optimal compression of relevant information is a variational problem

Tishby, Pereira, Bialek (1999)
Tishby Slonim **NeurIPS** (2000)



$$\min_{P(H|V)} \mathcal{L}_{IB}[P(H|V)] \equiv \min_{P(H|V)} I(V;H) - \beta I(H;E)$$

- IB equations:

$$P(h|v) = \frac{P(h)}{Z} \exp\left(-\beta \sum_e P(e|v) \log\left[\frac{P(e|v)}{P(e|h)}\right]\right)$$

$$P(h) = \sum_v P(h|v) P(v)$$

$$P(e|h) = \sum_v P(e|v) P(v|h)$$

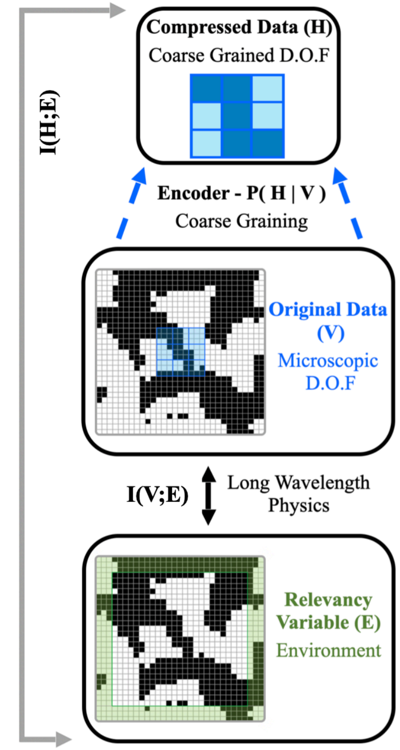
- Optimal IB encoder goes through a sequence of permutation **symmetry breaking** transitions

Gedeon et al. **Entropy (2012)**, 14(3) 456-479

The information bottleneck (IB) compression

- **Relevance** defined implicitly, by correlations with a signal variable
- Optimal compression of relevant information is a variational problem

Tishby, Pereira, Bialek (1999)
Tishby Slonim **NeurIPS** (2000)



$$\min_{P(H|V)} \mathcal{L}_{IB}[P(H|V)] \equiv \min_{P(H|V)} I(V;H) - \beta I(H;E)$$

- IB equations:

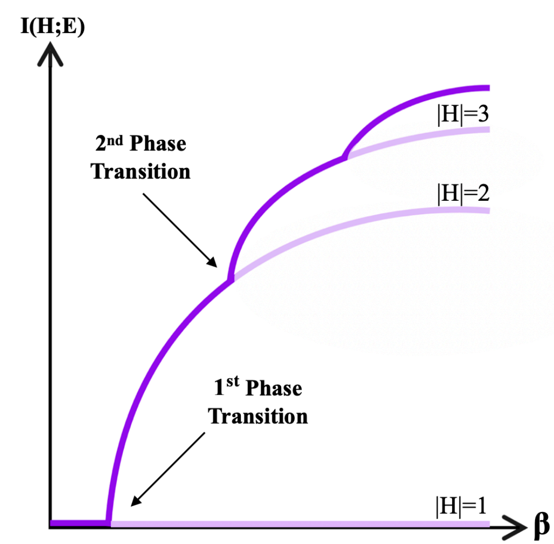
$$P(h|v) = \frac{P(h)}{Z} \exp\left(-\beta \sum_e P(e|v) \log\left[\frac{P(e|v)}{P(e|h)}\right]\right)$$

$$P(h) = \sum_v P(h|v)P(v)$$

$$P(e|h) = \sum_v P(e|v)P(v|h)$$

- Optimal IB encoder goes through a sequence of permutation **symmetry breaking** transitions

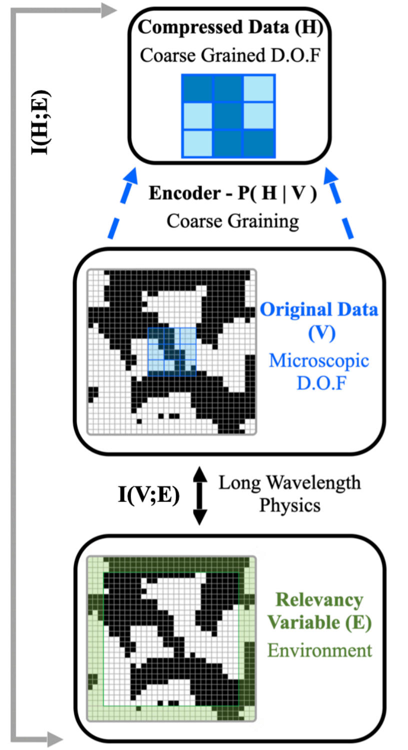
Gedeon et al. **Entropy (2012)**, 14(3) 456-479



The information bottleneck (IB) compression

- **Relevance** defined implicitly, by correlations with a signal variable
- Optimal compression of relevant information is a variational problem

Tishby, Pereira, Bialek (1999)
Tishby Slonim **NeurIPS** (2000)



$$\min_{P(H|V)} \mathcal{L}_{IB}[P(H|V)] \equiv \min_{P(H|V)} I(V;H) - \beta I(H;E)$$

- IB equations:

$$P(h|v) = \frac{P(h)}{Z} \exp\left(-\beta \sum_e P(e|v) \log\left[\frac{P(e|v)}{P(e|h)}\right]\right)$$

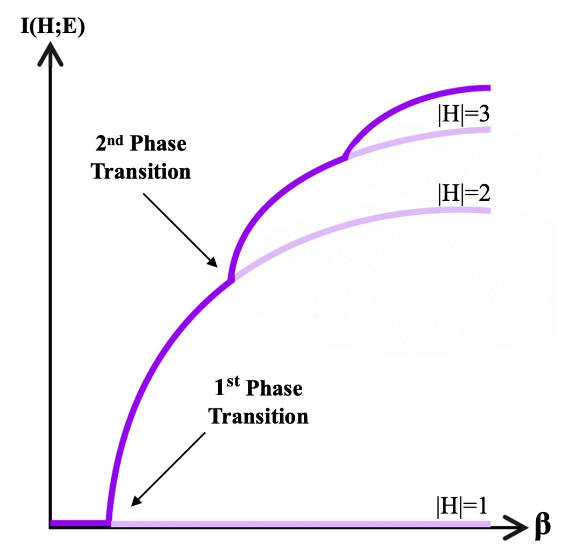
$$P(h) = \sum_v P(h|v)P(v)$$

$$P(e|h) = \sum_v P(e|v)P(v|h)$$

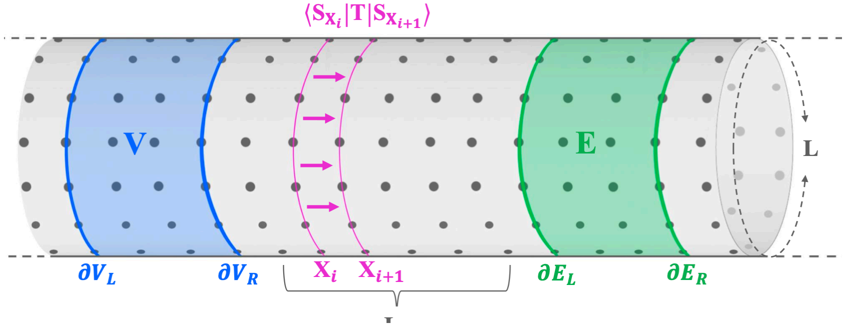
- Optimal IB encoder goes through a sequence of permutation **symmetry breaking** transitions

Gedeon et al. **Entropy (2012)**, 14(3) 456-479

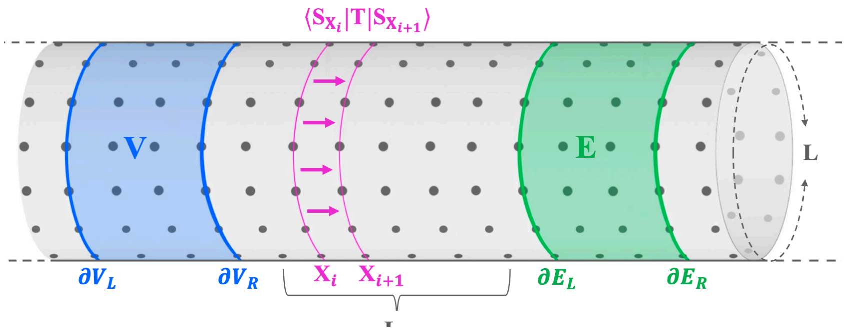
- RSMI arises in the infinite β limit, and finite alphabet



Equivalence of the IB and RG relevance

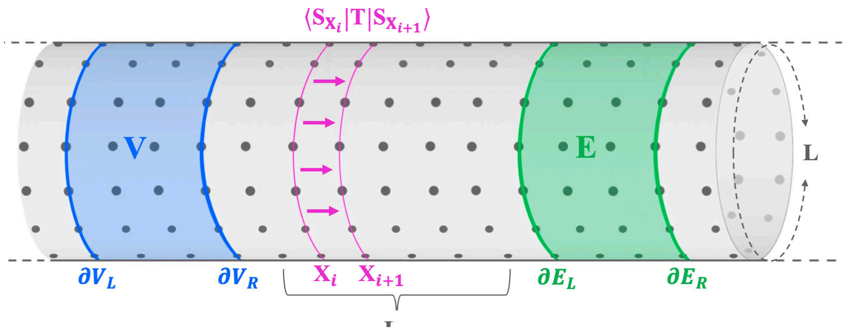


Equivalence of the IB and RG relevance



- We want to solve the IB eqs. for the optimal encoder $P(h|v)$ at a fixed β

Equivalence of the IB and RG relevance

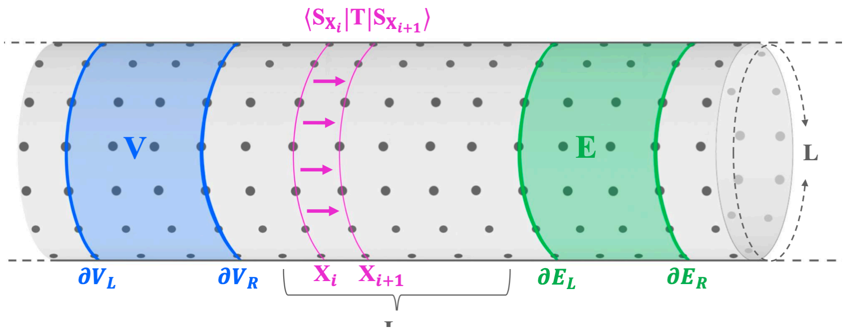


- Distributions in IB equations can be written using **transfer matrices** $Z = \mathrm{Tr}\left[\mathbb{T}^{L_\infty}\right]$
- Eigenvectors/eigenvalues of transfer matrices have direct relation to CFT operator content

  Cardy **J. Phys. A: Math. Gen. 17,** L385 (1984)
  Bloete et al. **Phys. Rev.Lett.** 56, 742 (1986)

- We want to solve the IB eqs. for the optimal encoder $P(h|v)$ at a fixed β

Equivalence of the IB and RG relevance



- Distributions in IB equations can be written using **transfer matrices** $Z = \mathrm{Tr}\left[\mathbb{T}^{L_\infty}\right]$
- Eigenvectors/eigenvalues of transfer matrices have direct relation to CFT operator content

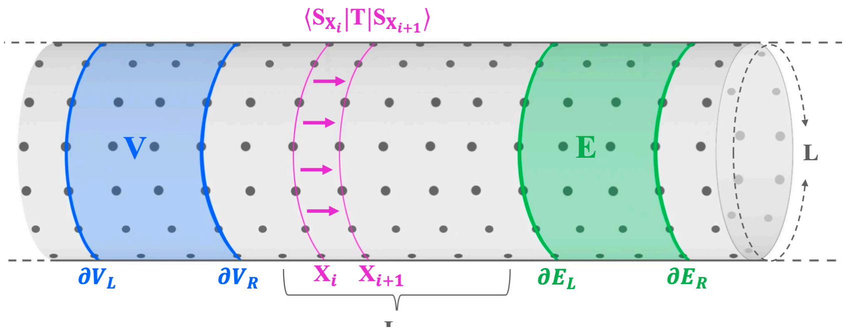
  Cardy **J. Phys. A: Math. Gen. 17,** L385 (1984)
  Bloete et al. **Phys. Rev.Lett.** 56, 742 (1986)

- We want to solve the IB eqs. for the optimal encoder $P(h|v)$ at a fixed β

- All quantities of interest are functions of matrix elements of RG-relevant operators:

$$P(v|e) = N^{-1}P(v)\left[1 + \frac{\langle\partial V_R|1\rangle\langle 1|\partial E_L\rangle}{\langle\partial V_R|0\rangle\langle 0|\partial E_L\rangle}\left(\frac{\lambda_1}{\lambda_0}\right)^{L_B}\right]$$
$$= N^{-1}P(v)\left[1 + \epsilon r_e r_v\right]$$

Equivalence of the IB and RG relevance



- Distributions in IB equations can be written using **transfer matrices** $Z = \mathrm{Tr}\left[\mathbb{T}^{L_\infty}\right]$
- Eigenvectors/eigenvalues of transfer matrices have direct relation to CFT operator content

  Cardy **J. Phys. A: Math. Gen. 17,** L385 (1984)
  Bloete et al. **Phys. Rev.Lett.** 56, 742 (1986)

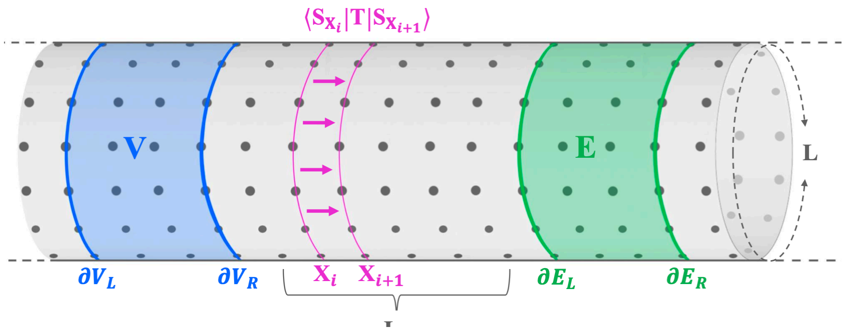
- We want to solve the IB eqs. for the optimal encoder $P(h|v)$ at a fixed β

- All quantities of interest are functions of matrix elements of RG-relevant operators:

$$P(v|e) = N^{-1}P(v)\left[1 + \frac{\langle\partial V_R|1\rangle\langle 1|\partial E_L\rangle}{\langle\partial V_R|0\rangle\langle 0|\partial E_L\rangle}\left(\frac{\lambda_1}{\lambda_0}\right)^{L_B}\right]$$
$$= N^{-1}P(v)\left[1 + \epsilon r_e r_v\right]$$

$$r_v = \frac{\langle 1|\partial V_R\rangle}{\langle 0|\partial V_R\rangle} = \frac{\langle 0|\phi_{\Delta_1}|\partial V_R\rangle}{\langle 0|\partial V_R\rangle}$$
$$r_e = \frac{\langle\partial E_L|1\rangle}{\langle\partial E_L|0\rangle} = \frac{\langle\partial E_L|\phi_{\Delta_1}|0\rangle}{\langle\partial E_L|0\rangle}$$

Equivalence of the IB and RG relevance



- Distributions in IB equations can be written using **transfer matrices** $Z = \mathrm{Tr}\left[\mathbb{T}^{L_\infty}\right]$
- Eigenvectors/eigenvalues of transfer matrices have direct relation to CFT operator content

  Cardy **J. Phys. A: Math. Gen. 17,** L385 (1984)
  Bloete et al. **Phys. Rev.Lett.** 56, 742 (1986)

- We want to solve the IB eqs. for the optimal encoder $P(h|v)$ at a fixed β

- All quantities of interest are functions of matrix elements of RG-relevant operators:

$$P(v|e) = N^{-1}P(v)\left[1 + \frac{\langle\partial V_R|1\rangle\langle 1|\partial E_L\rangle}{\langle\partial V_R|0\rangle\langle 0|\partial E_L\rangle}\left(\frac{\lambda_1}{\lambda_0}\right)^{L_B}\right]$$
$$= N^{-1}P(v)\left[1 + \epsilon r_e r_v\right]$$

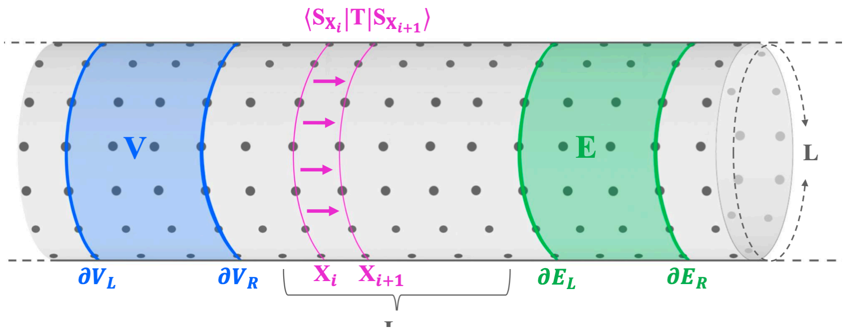
$$r_v = \frac{\langle 1|\partial V_R\rangle}{\langle 0|\partial V_R\rangle} = \frac{\langle 0|\phi_{\Delta_1}|\partial V_R\rangle}{\langle 0|\partial V_R\rangle}$$
$$r_e = \frac{\langle\partial E_L|1\rangle}{\langle\partial E_L|0\rangle} = \frac{\langle\partial E_L|\phi_{\Delta_1}|0\rangle}{\langle\partial E_L|0\rangle}$$

- The **IB-optimal** encoder only depends on **RG-relevant** data:

$$P(h|v) = P(h|r_v) \propto P(h)e^{\beta\epsilon^2 r_v \langle r_v\rangle_h}$$

Gordon, Banerjee, **MKJ,** Ringel, *PRL **126**, 240601 (2021)*

Equivalence of the IB and RG relevance



- Distributions in IB equations can be written using **transfer matrices** $Z = \mathrm{Tr}\left[\mathbb{T}^{L_\infty}\right]$
- Eigenvectors/eigenvalues of transfer matrices have direct relation to CFT operator content

  Cardy **J. Phys. A: Math. Gen. 17,** L385 (1984)
  Bloete et al. **Phys. Rev.Lett.** 56, 742 (1986)

- We want to solve the IB eqs. for the optimal encoder $P(h|v)$ at a fixed β

- All quantities of interest are functions of matrix elements of RG-relevant operators:

$$P(v|e) = N^{-1}P(v)\left[1 + \frac{\langle\partial V_R|1\rangle\langle 1|\partial E_L\rangle}{\langle\partial V_R|0\rangle\langle 0|\partial E_L\rangle}\left(\frac{\lambda_1}{\lambda_0}\right)^{L_B}\right]$$
$$= N^{-1}P(v)\left[1 + \epsilon r_e r_v\right]$$

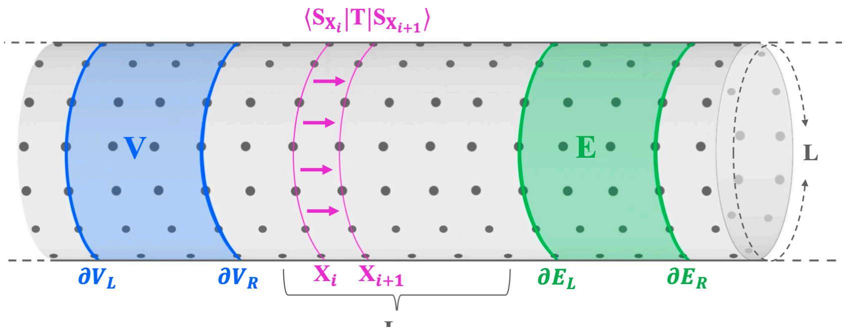
$$r_v = \frac{\langle 1|\partial V_R\rangle}{\langle 0|\partial V_R\rangle} = \frac{\langle 0|\phi_{\Delta_1}|\partial V_R\rangle}{\langle 0|\partial V_R\rangle}$$
$$r_e = \frac{\langle\partial E_L|1\rangle}{\langle\partial E_L|0\rangle} = \frac{\langle\partial E_L|\phi_{\Delta_1}|0\rangle}{\langle\partial E_L|0\rangle}$$

- The **IB-optimal** encoder only depends on **RG-relevant** data:

$$P(h|v) = P(h|r_v) \propto P(h)e^{\beta\epsilon^2 r_v\langle r_v\rangle_h}$$

Gordon, Banerjee, **MKJ**, Ringel, *PRL **126**, 240601 (2021)*

Equivalence of the IB and RG relevance



- Distributions in IB equations can be written using **transfer matrices** $Z = \mathrm{Tr}\left[\mathbb{T}^{L_\infty}\right]$
- Eigenvectors/eigenvalues of transfer matrices have direct relation to CFT operator content

  Cardy **J. Phys. A: Math. Gen. 17,** L385 (1984)
  Bloete et al. **Phys. Rev.Lett.** 56, 742 (1986)

- We want to solve the IB eqs. for the optimal encoder $P(h|v)$ at a fixed β

- All quantities of interest are functions of matrix elements of RG-relevant operators:

$$P(v|e) = N^{-1}P(v)\left[1 + \frac{\langle\partial V_R|1\rangle\langle 1|\partial E_L\rangle}{\langle\partial V_R|0\rangle\langle 0|\partial E_L\rangle}\left(\frac{\lambda_1}{\lambda_0}\right)^{L_B}\right]$$
$$= N^{-1}P(v)\left[1 + \epsilon r_e r_v\right]$$

$$r_v = \frac{\langle 1|\partial V_R\rangle}{\langle 0|\partial V_R\rangle} = \frac{\langle 0|\phi_{\Delta_1}|\partial V_R\rangle}{\langle 0|\partial V_R\rangle}$$
$$r_e = \frac{\langle \partial E_L|1\rangle}{\langle \partial E_L|0\rangle} = \frac{\langle \partial E_L|\phi_{\Delta_1}|0\rangle}{\langle \partial E_L|0\rangle}$$

- The **IB-optimal** encoder only depends on **RG-relevant** data:

$$P(h|v) = P(h|r_v) \propto P(h)e^{\beta\epsilon^2 r_v \langle r_v\rangle_h}$$

Gordon, Banerjee, **MKJ,** Ringel, *PRL **126**, 240601 (2021)*

- Expand around first transition $\beta = \beta_{c,1} + t$:

Perturbative IB

- Expand around first transition $\beta = \beta_{c,1} + t$:

$$P(h|r_v) = \frac{1}{|H|} + t b_{r_v}(h)$$

$$\sum_h b_{r_v}(h) = 0,$$

- Expand around first transition $\beta = \beta_{c,1} + t$:

$$P(h|r_v) = \frac{1}{|H|} + t b_{r_v}(h)$$

$$\sum_h b_{r_v}(h) = 0,$$

Differentiation from trivial encoder

Perturbative IB

- Expand around first transition $\beta = \beta_{c,1} + t$:

$$P(h|r_v) = \frac{1}{|H|} + t b_{r_v}(h)$$

Differentiation from trivial encoder

$$\sum_h b_{r_v}(h) = 0,$$

- Trivial solution exists always, but a nontrivial one appears when: $\beta_{c,1}^{-1} = \epsilon^2 + o(\epsilon^2)$

Perturbative IB

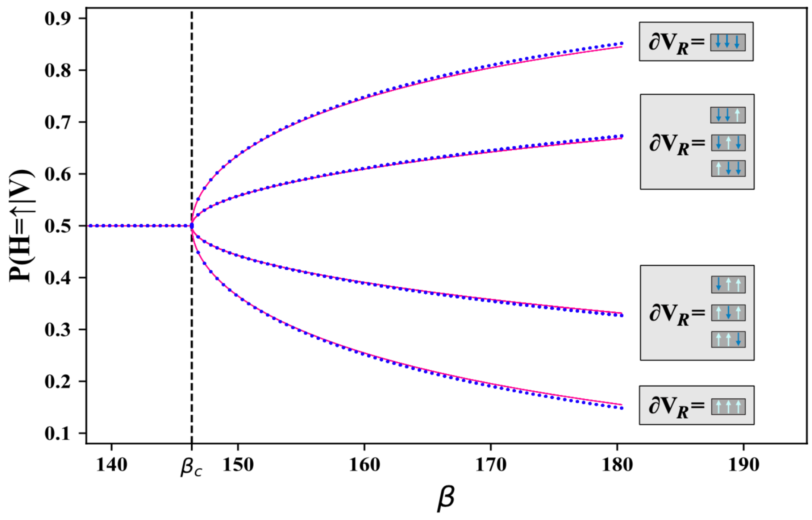
- Expand around first transition $\beta = \beta_{c,1} + t$:

$$P(h|r_v) = \frac{1}{|H|} + t b_{r_v}(h)$$

$$\sum_h b_{r_v}(h) = 0,$$

Differentiation from trivial encoder

- Trivial solution exists always, but a nontrivial one appears when: $\beta_{c,1}^{-1} = \epsilon^2 + o(\epsilon^2)$



Estimating MI with neural networks

Estimating MI with neural networks

$$I(X:Y) := H(X) - H(X|Y)$$

Estimating MI with neural networks

$$I(X:Y) := H(X) - H(X|Y)$$

$$= \mathbb{E}_{p(x,y)}\left[\log \frac{p(x,y)}{p(x)p(y)}\right]$$

Estimating MI with neural networks

- Common methods are computationally demanding and/or not differentiable

$$I(X:Y) := H(X) - H(X|Y)$$
$$= \mathbb{E}_{p(x,y)}\left[\log \frac{p(x,y)}{p(x)p(y)}\right]$$

Estimating MI with neural networks

- Common methods are computationally demanding and/or not differentiable

$$I(X:Y) := H(X) - H(X|Y)$$

$$= \mathbb{E}_{p(x,y)}\left[\log \frac{p(x,y)}{p(x)p(y)}\right]$$

$$=: D_{\mathrm{KL}}\left[p(x,y)||p(x)p(y)\right]$$

Estimating MI with neural networks

- Common methods are computationally demanding and/or not differentiable

$$I(X:Y) := H(X) - H(X|Y)$$
$$= \mathbb{E}_{p(x,y)}\left[\log \frac{p(x,y)}{p(x)p(y)}\right]$$
$$=: D_{\mathrm{KL}}\left[p(x,y)||p(x)p(y)\right]$$

- Estimation by optimising **variational bounds:**

Estimating MI with neural networks

- Common methods are computationally demanding and/or not differentiable

$$I(X:Y) := H(X) - H(X|Y)$$
$$= \mathbb{E}_{p(x,y)}\left[\log \frac{p(x,y)}{p(x)p(y)}\right]$$
$$=: D_{\mathrm{KL}}\left[p(x,y)||p(x)p(y)\right]$$

- Estimation by optimising **variational bounds:**

$$D_{KL}(\mathbb{P} \,||\, \mathbb{Q}) = \sup_{T:\Omega\to\mathbb{R}} \mathbb{E}_{\mathbb{P}}[T] - \log(\mathbb{E}_{\mathbb{Q}}[e^T])$$

Estimating MI with neural networks

- Common methods are computationally demanding and/or not differentiable

$$I(X:Y) := H(X) - H(X|Y)$$
$$= \mathbb{E}_{p(x,y)}\left[\log \frac{p(x,y)}{p(x)p(y)}\right]$$
$$=: D_{\mathrm{KL}}\left[p(x,y)||p(x)p(y)\right]$$

- Estimation by optimising **variational bounds:**

$$D_{KL}(\mathbb{P} \,||\, \mathbb{Q}) = \sup_{T:\Omega\to\mathbb{R}} \mathbb{E}_{\mathbb{P}}[T] - \log(\mathbb{E}_{\mathbb{Q}}[e^T])$$

Make T a neural network! [*MINE*, Belghazi et al. (2018)]

Estimating MI with neural networks

- Common methods are computationally demanding and/or not differentiable

$$I(X:Y) := H(X) - H(X|Y)$$
$$= \mathbb{E}_{p(x,y)}\left[\log \frac{p(x,y)}{p(x)p(y)}\right]$$
$$=: D_{\mathrm{KL}}\left[p(x,y)||p(x)p(y)\right]$$

- Estimation by optimising **variational bounds:**

$$D_{KL}(\mathbb{P} \,||\, \mathbb{Q}) = \sup_{T:\Omega\to\mathbb{R}} \mathbb{E}_{\mathbb{P}}[T] - \log(\mathbb{E}_{\mathbb{Q}}[e^T])$$

Make T a neural network! [*MINE*, Belghazi et al. (2018)]

- MINE can be improved in many respects (e.g. variance)

InfoNCE, van den Oord et al. (2018)

Poole et al. ICMLR (2019) “On variational bounds of mutual information”

Estimating MI with neural networks

- Common methods are computationally demanding and/or not differentiable

$$I(X:Y) := H(X) - H(X|Y)$$
$$= \mathbb{E}_{p(x,y)}\left[\log \frac{p(x,y)}{p(x)p(y)}\right]$$
$$=: D_{\mathrm{KL}}\left[p(x,y)||p(x)p(y)\right]$$

- Estimation by optimising **variational bounds:**

$$D_{KL}(\mathbb{P} \,||\, \mathbb{Q}) = \sup_{T:\Omega\to\mathbb{R}} \mathbb{E}_{\mathbb{P}}[T] - \log(\mathbb{E}_{\mathbb{Q}}[e^T])$$

Make T a neural network! [*MINE*, Belghazi et al. (2018)]

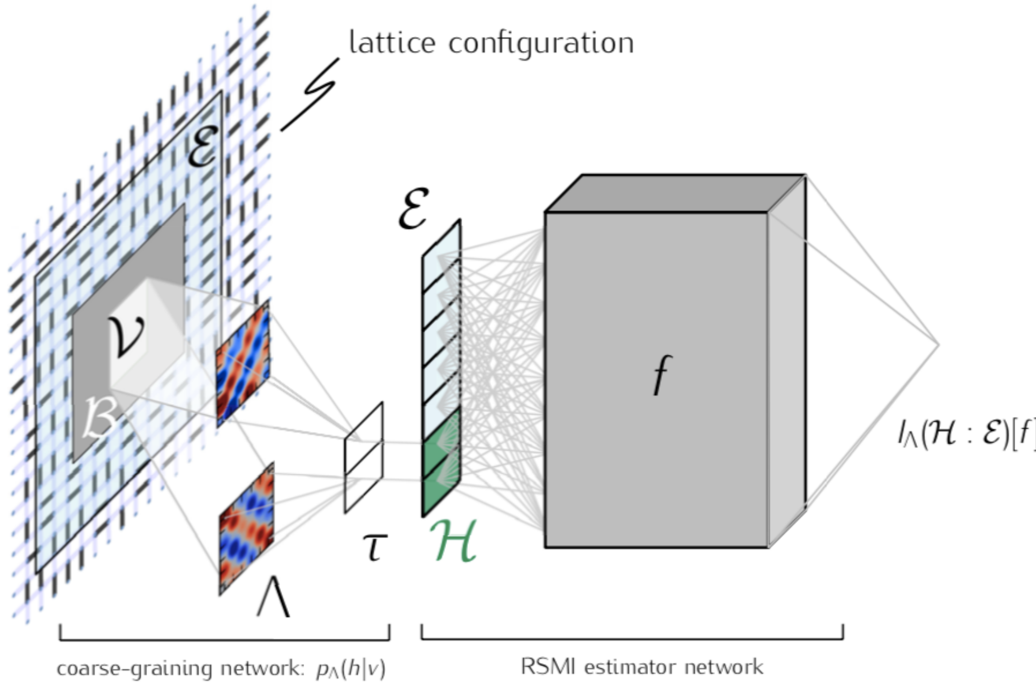
- MINE can be improved in many respects (e.g. variance)

InfoNCE, van den Oord et al. (2018)

Poole et al. ICMLR (2019) "On variational bounds of mutual information"

- We get a **parametric, differentiable**, and **tight** lower bound on MI

The RSMI-NE network

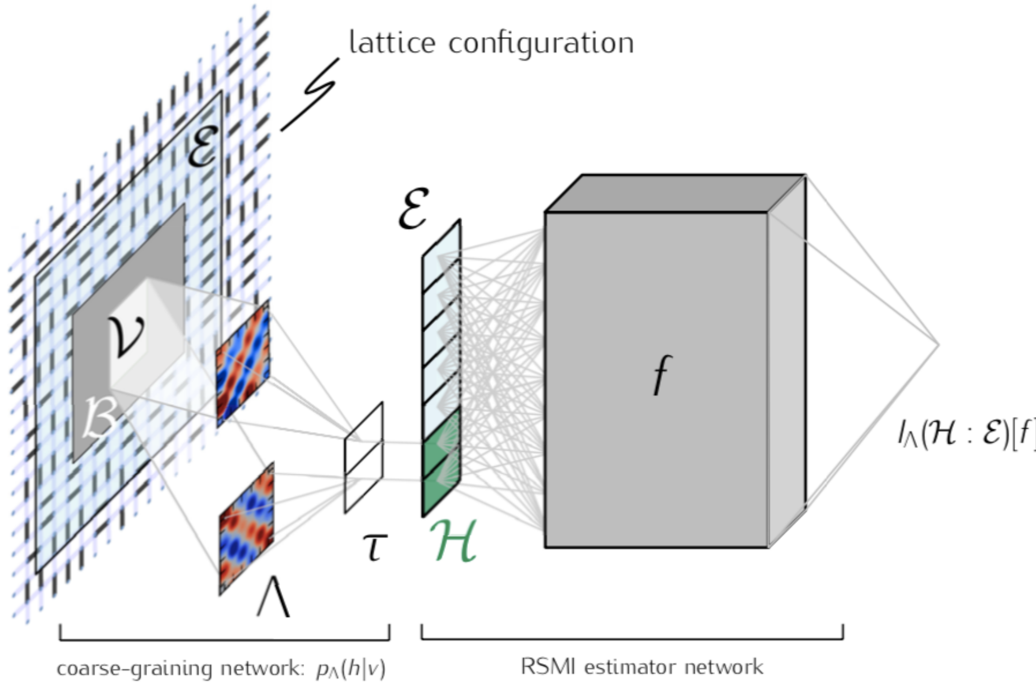


https://github.com/RSMI-NE/RSMI-NE

Phys.Rev. E, 104, 064106 (2021)
Phys.Rev.Lett 127, 240603 (2021)

The RSMI-NE network

$$p_\Lambda(h|v) : v \mapsto h = \tau \circ (\Lambda \cdot v)$$



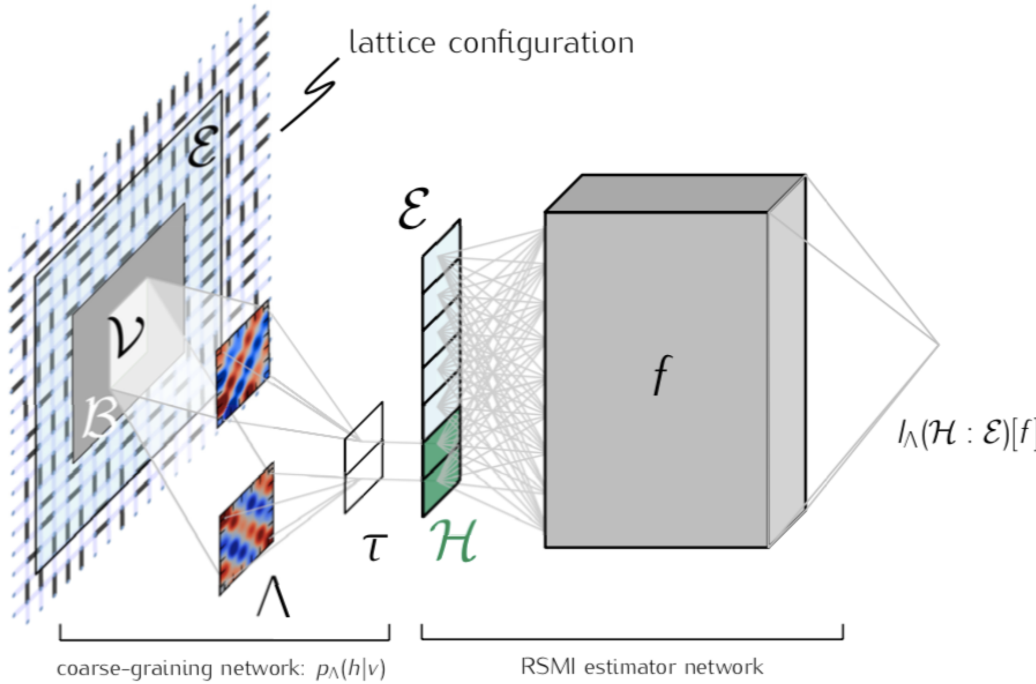
https://github.com/RSMI-NE/RSMI-NE

Phys.Rev. E, 104, 064106 (2021)
Phys.Rev.Lett 127, 240603 (2021)

The RSMI-NE network

$$p_\Lambda(h|v) : v \mapsto h = \tau \circ (\Lambda \cdot v)$$

- The coarse-graining network(s)

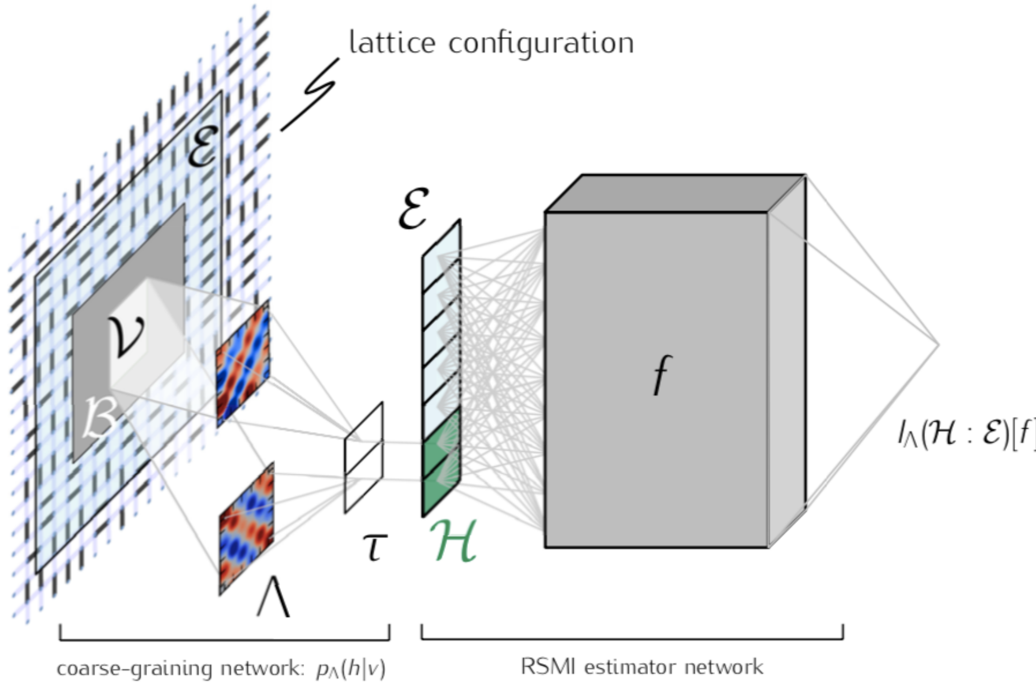


Phys.Rev. E, 104, 064106 (2021)
Phys.Rev.Lett 127, 240603 (2021)

https://github.com/RSMI-NE/RSMI-NE

The RSMI-NE network

$$p_\Lambda(h|v) : v \mapsto h = \tau \circ (\Lambda \cdot v)$$



- The coarse-graining network(s)

- **Differentiable** discretization

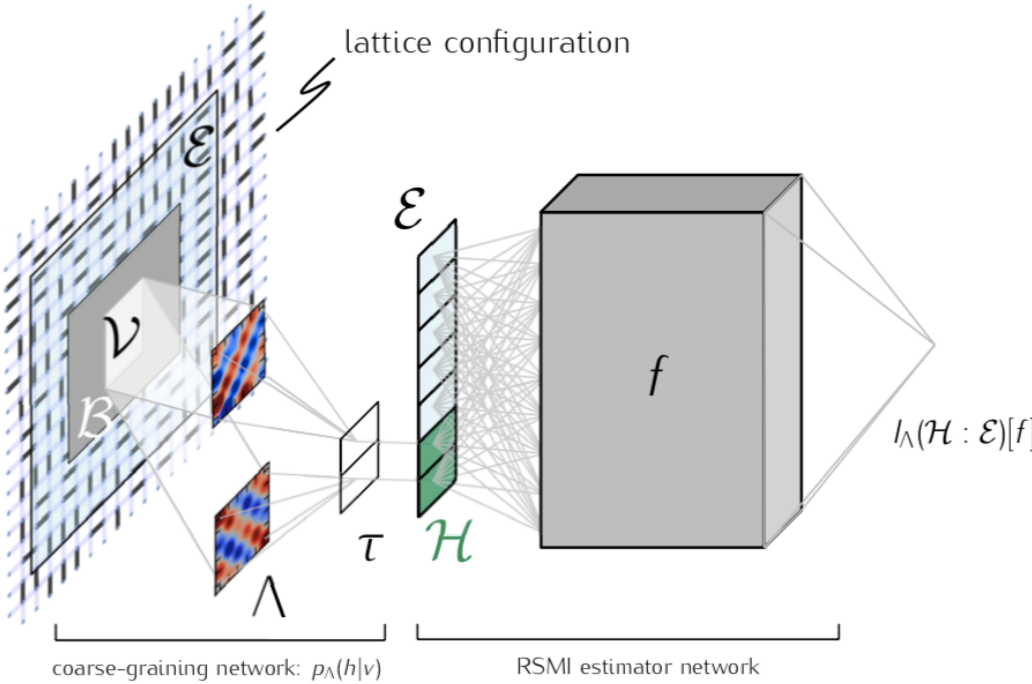
  Bengio, Leonard, Courville **arXiv**:1308:3432
  Jang, Gu, Poole **ICLR (2017**)
  Maddison, Mnih, Teh **ICLR (2017)**

Phys.Rev. E, 104, 064106 (2021)
Phys.Rev.Lett 127, 240603 (2021)

https://github.com/RSMI-NE/RSMI-NE

The RSMI-NE network

$$p_\Lambda(h|v) : v \mapsto h = \tau \circ (\Lambda \cdot v)$$



- The coarse-graining network(s)

- **Differentiable** discretization

  Bengio, Leonard, Courville **arXiv**:1308:3432
  Jang, Gu, Poole **ICLR (2017)**
  Maddison, Mnih, Teh **ICLR (2017)**

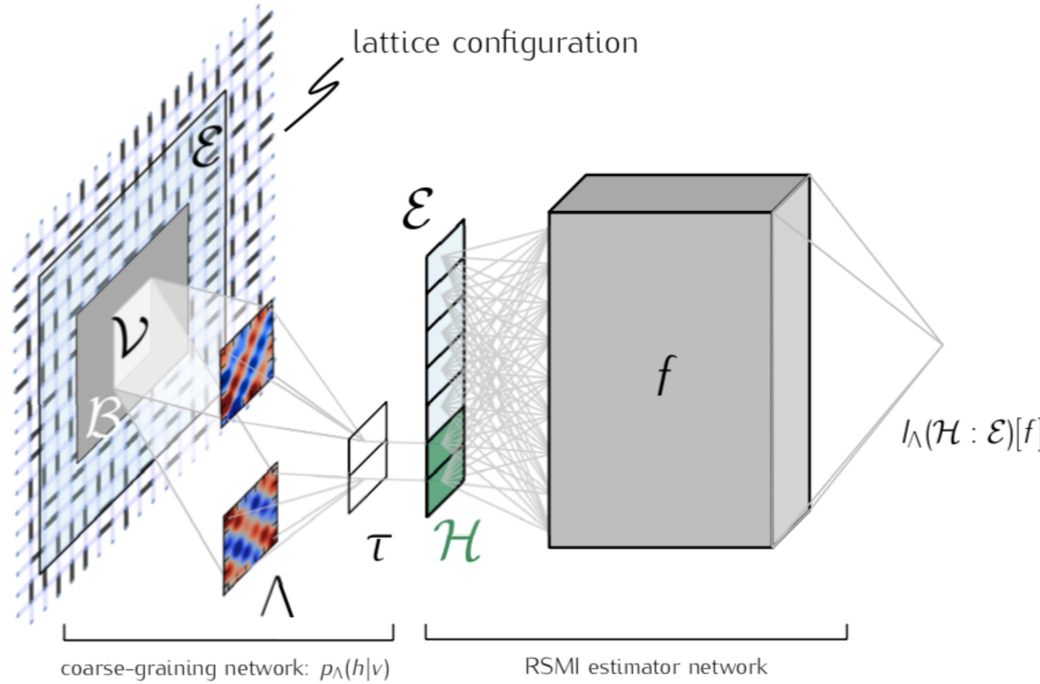
- The RSMI estimator and the coarse-grainer are stacked

Phys.Rev. E, 104, 064106 (2021)
Phys.Rev.Lett 127, 240603 (2021)

https://github.com/RSMI-NE/RSMI-NE

The RSMI-NE network

$$p_\Lambda(h|v) : v \mapsto h = \tau \circ (\Lambda \cdot v)$$



- The coarse-graining network(s)
- **Differentiable** discretization

  Bengio, Leonard, Courville **arXiv**:1308:3432
  Jang, Gu, Poole **ICLR (2017)**
  Maddison, Mnih, Teh **ICLR (2017)**

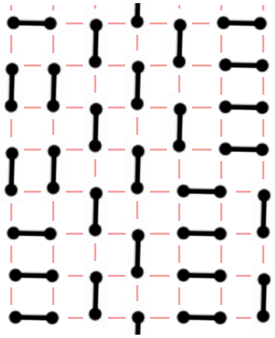
- The RSMI estimator and the coarse-grainer are stacked
- Co-trained with SGD as a single network **(differentiable, upper bounded!)**

Phys.Rev. E, 104, 064106 (2021)
Phys.Rev.Lett 127, 240603 (2021)

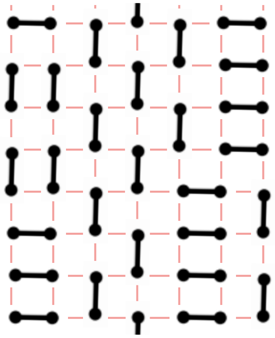
https://github.com/RSMI-NE/RSMI-NE

Example: the interacting dimer model

Example: the interacting dimer model



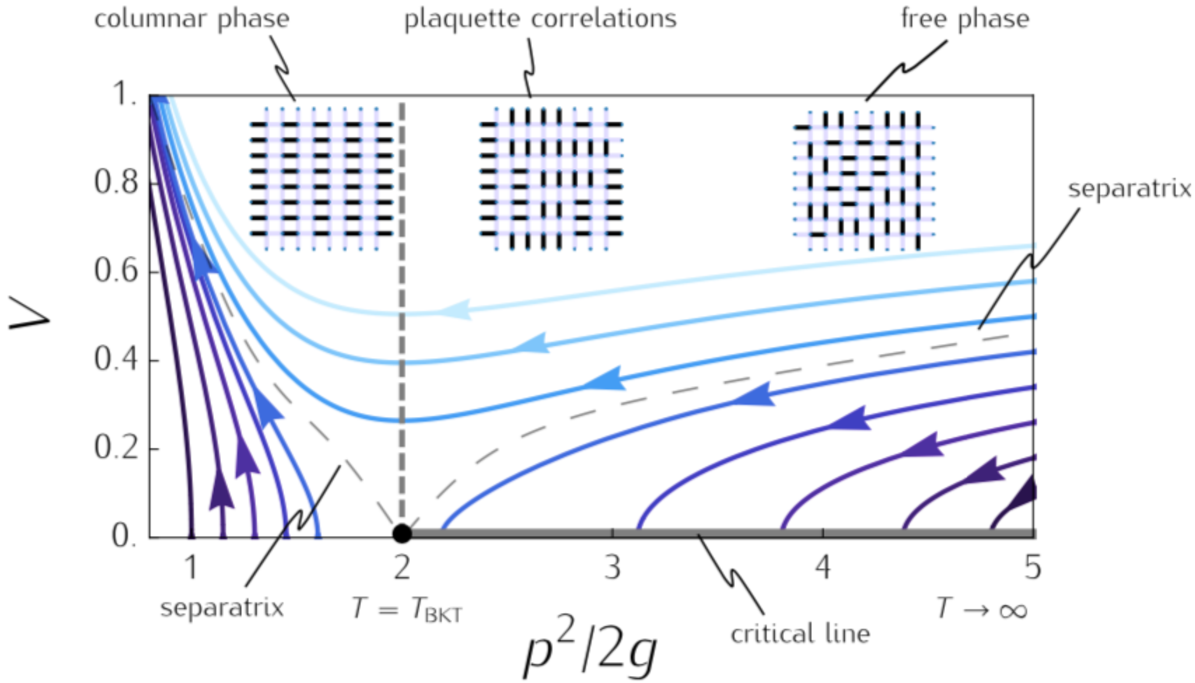
Example: the interacting dimer model

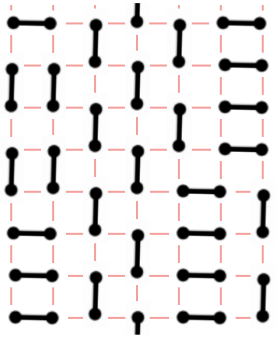


$$Z = \sum_c \exp(-E_c/T),$$

$$E_c = v[(N^c(\text{=}) + N^c(\text{II}))].$$

Example: the interacting dimer model





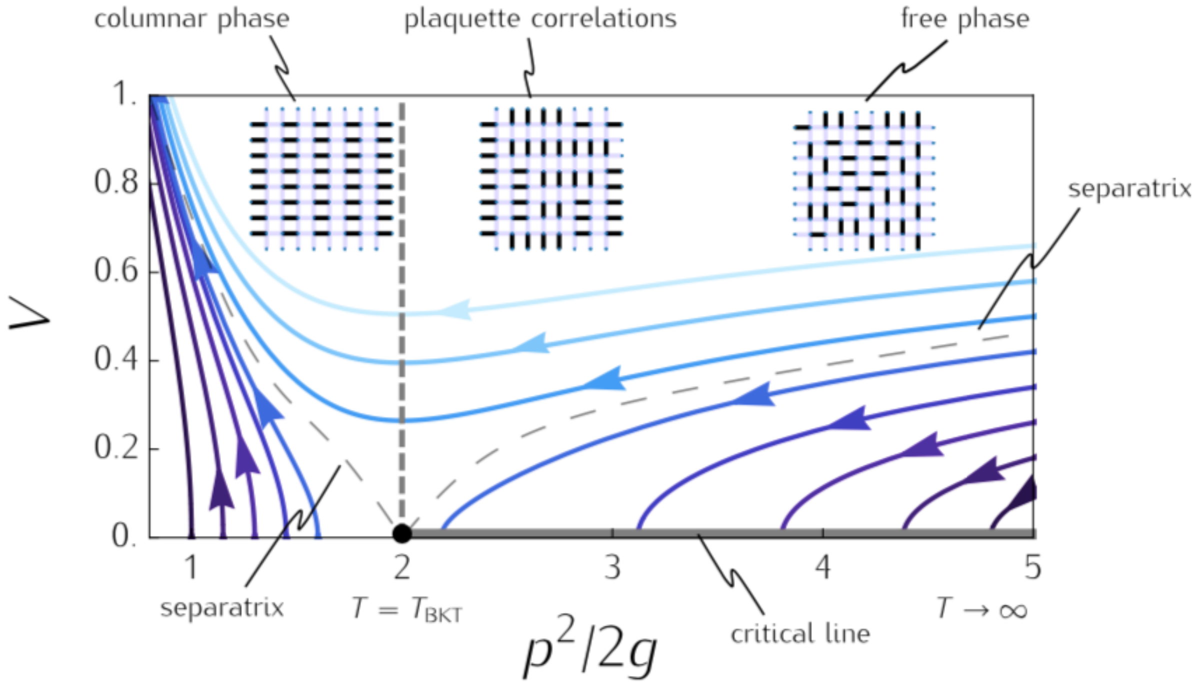
$$S[\varphi(\mathbf{r})] = \int \mathrm{d}^2\mathbf{r} \left[\frac{g(T)}{2} |\nabla \varphi(\mathbf{r})|^2 + V \cos\left(4\varphi(\mathbf{r})\right) \right]$$

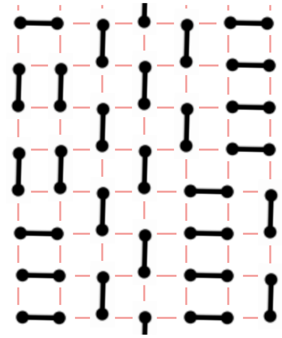
$$Z = \sum_c \exp(-E_c/T),$$

$$E_c = v[(N^c(=) + N^c(\|))].$$

Alet et al. **PRE 74**, 041124 (2006)

Example: the interacting dimer model





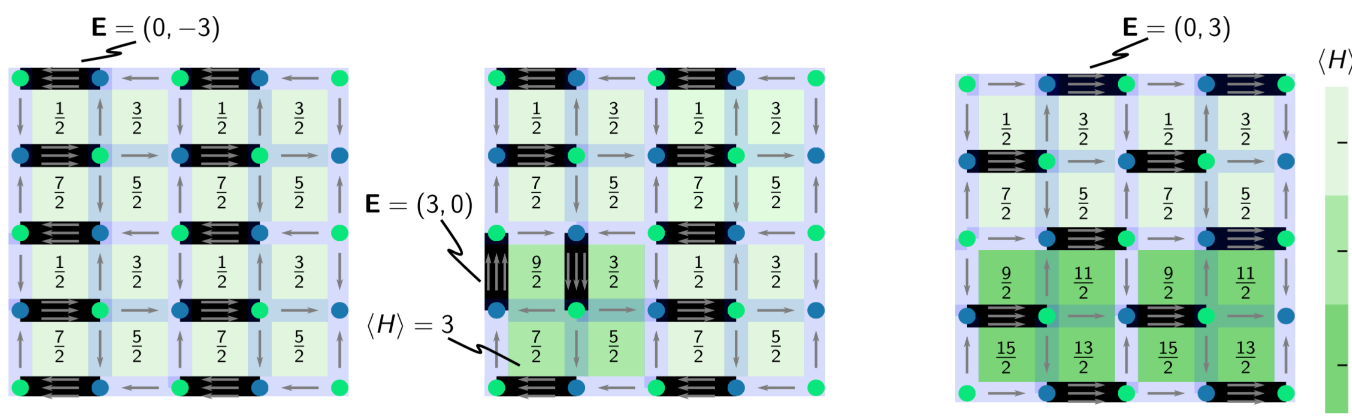
$$Z = \sum_c \exp(-E_c/T),$$

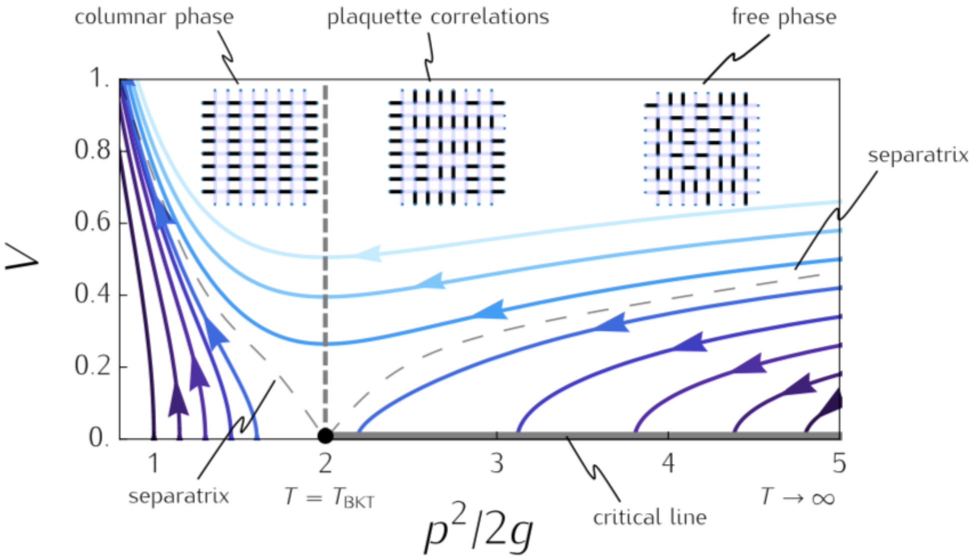
$$E_c = v[(N^c(\mathbf{=}) + N^c(\mathbf{II}))].$$

$$S[\varphi(\mathbf{r})] = \int \mathrm{d}^2\mathbf{r} \left[\frac{g(T)}{2} |\nabla\varphi(\mathbf{r})|^2 + V \cos\left(4\varphi(\mathbf{r})\right) \right]$$

Alet et al. **PRE 74**, 041124 (2006)

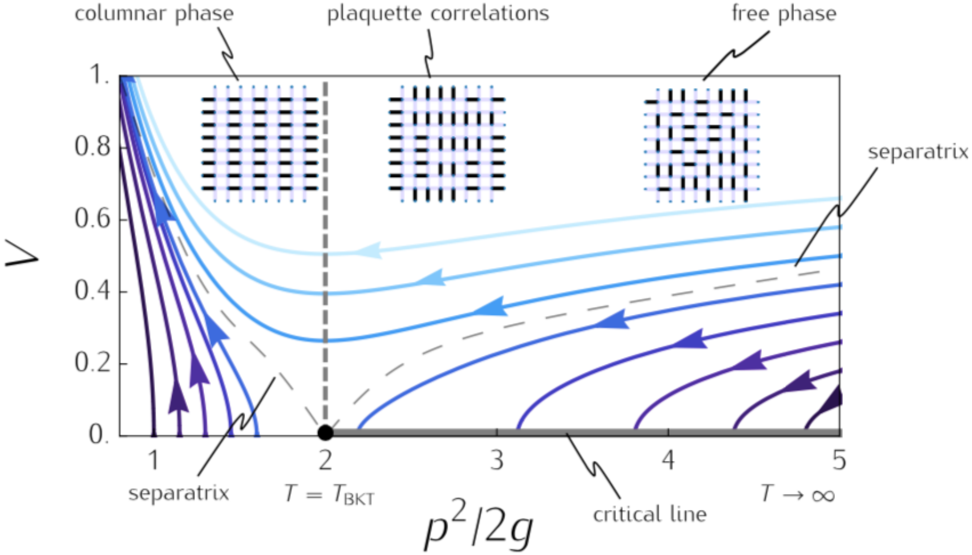
RG of dimer model:
mapping to height field



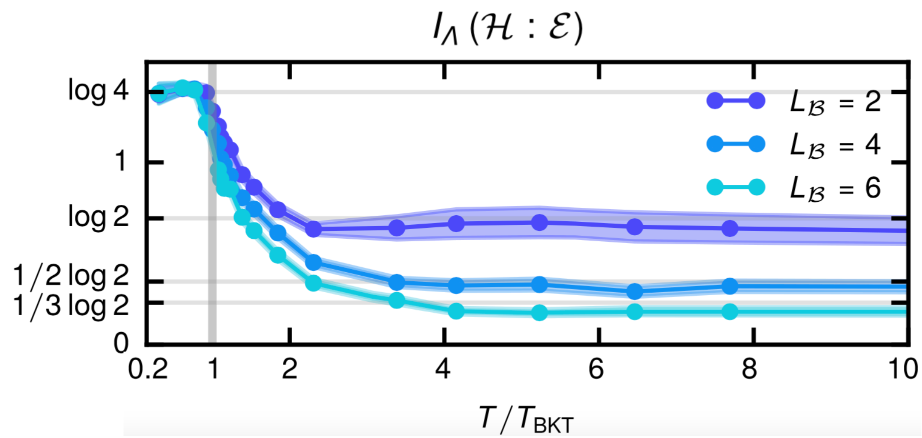


https://github.com/RSMI-NE/RSMI-NE

Phys.Rev.Lett 127, 240603 (2021)

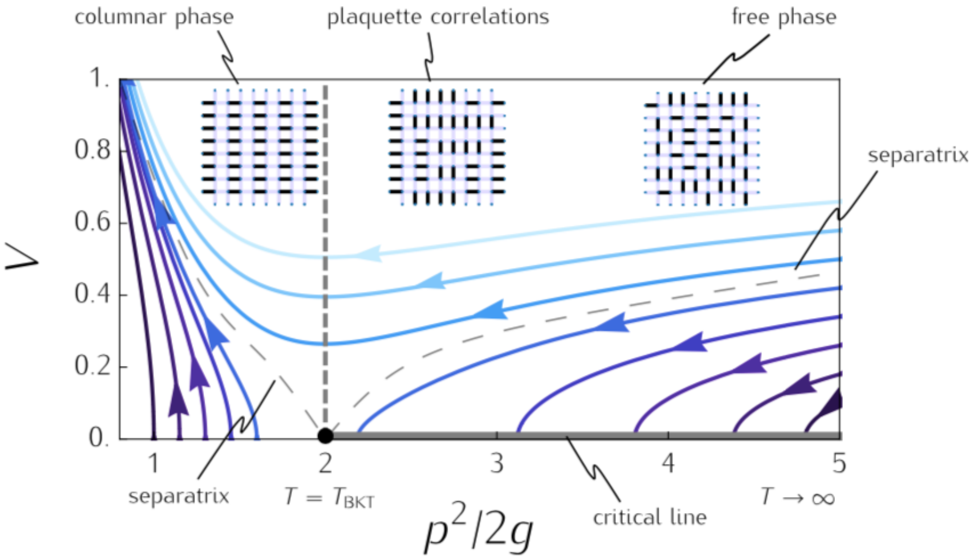


- Total RSMI with the *optimal* filter

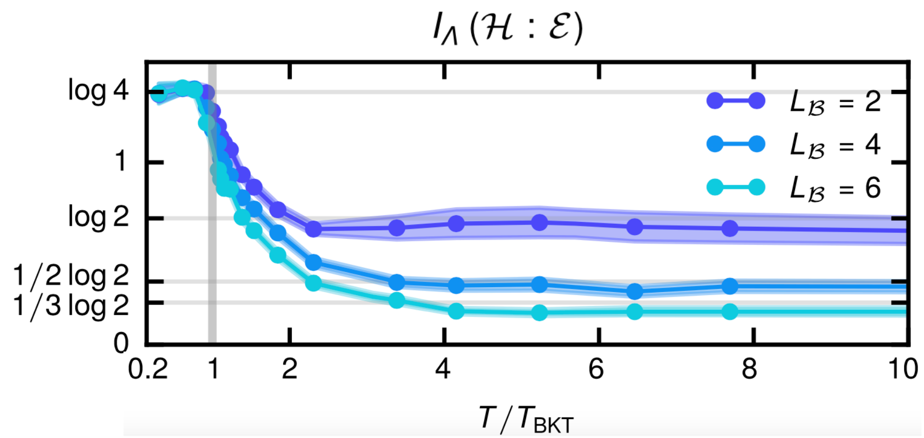


https://github.com/RSMI-NE/RSMI-NE

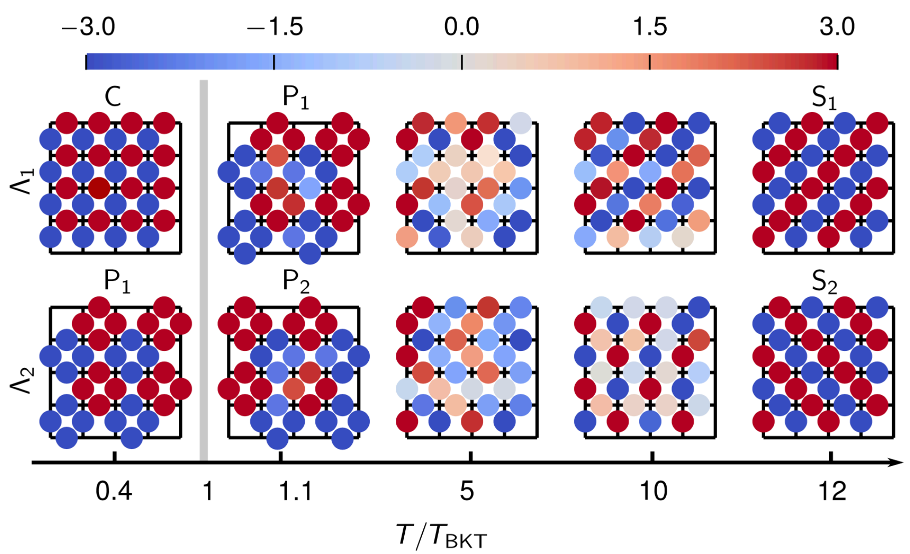
Phys.Rev.Lett 127, 240603 (2021)



- Total RSMI with the *optimal* filter



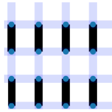
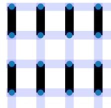
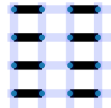
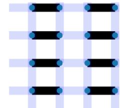
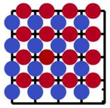
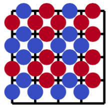
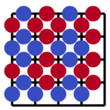
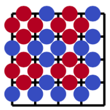
- The optimal filters ***depend*** on T



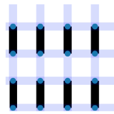
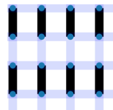
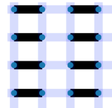
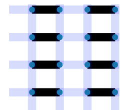
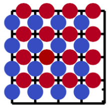
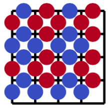
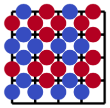
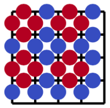
https://github.com/RSMI-NE/RSMI-NE

Phys.Rev.Lett 127, 240603 (2021)

- Pairs of C/P filters **label broken symmetry** states

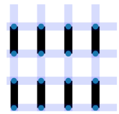
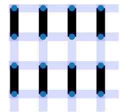
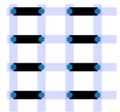
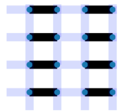
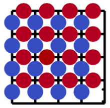
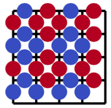
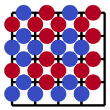
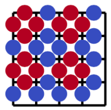
$C(\mathbf{r})$				
Λ_{C}  Λ_{P1} 	$(-1,-1)$	$(-1,+1)$	$(+1,-1)$	$(+1,+1)$
Λ_{P1}  Λ_{P2} 	$(-1,-1)$	$(+1,+1)$	$(-1,+1)$	$(+1,-1)$

- Pairs of C/P filters **label broken symmetry** states

$C(\mathbf{r})$				
Λ_{C}  Λ_{P1} 	$(-1,-1)$	$(-1,+1)$	$(+1,-1)$	$(+1,+1)$
Λ_{P1}  Λ_{P2} 	$(-1,-1)$	$(+1,+1)$	$(-1,+1)$	$(+1,-1)$

- Filters **define** order parameters:

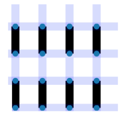
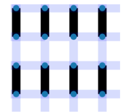
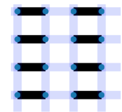
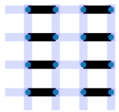
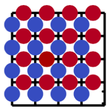
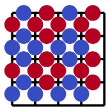
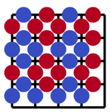
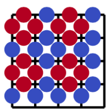
- Pairs of C/P filters **label broken symmetry states**

$C(\mathbf{r})$				
Λ_{C}  Λ_{P1} 	$(-1,-1)$	$(-1,+1)$	$(+1,-1)$	$(+1,+1)$
Λ_{P1}  Λ_{P2} 	$(-1,-1)$	$(+1,+1)$	$(-1,+1)$	$(+1,-1)$

- Filters **define** order parameters:

$$D_i := \mathbb{E}\left[\frac{1}{N_{\mathcal{V}}}\sum_k \tau \circ (\Lambda_i \cdot \mathcal{V}_k)\right]$$

- Pairs of C/P filters **label broken symmetry** states

$C(\mathbf{r})$				
Λ_C  Λ_{P1} 	$(-1,-1)$	$(-1,+1)$	$(+1,-1)$	$(+1,+1)$
Λ_{P1}  Λ_{P2} 	$(-1,-1)$	$(+1,+1)$	$(-1,+1)$	$(+1,-1)$

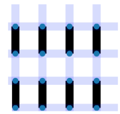
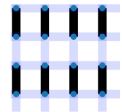
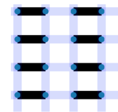
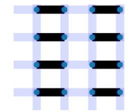
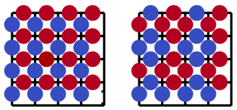
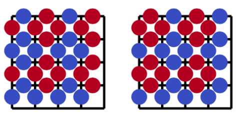
- Filters **define** order parameters:

$$D_i := \mathbb{E}\left[\frac{1}{N_{\mathcal{V}}}\sum_k \tau \circ (\Lambda_i \cdot \mathcal{V}_k)\right]$$

$$\mathrm{DSB} := \mathbb{E}\left[\sum_k \tau \circ (\Lambda_{\mathrm{C}} \cdot \mathcal{V}_k)\right]$$

Alet et al. **PRE 74**, 041124 (2006)

- Pairs of C/P filters **label broken symmetry states**

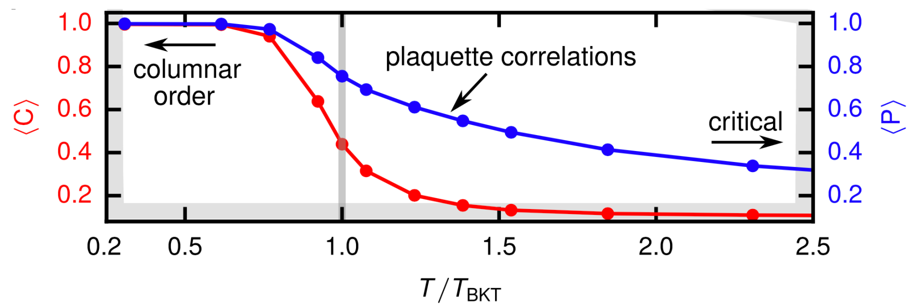
$C(\mathbf{r})$				
Λ_C Λ_{P1} 	$(-1,-1)$	$(-1,+1)$	$(+1,-1)$	$(+1,+1)$
Λ_{P1} Λ_{P2} 	$(-1,-1)$	$(+1,+1)$	$(-1,+1)$	$(+1,-1)$

- Filters **define** order parameters:

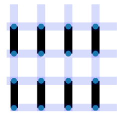
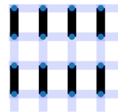
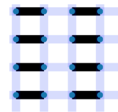
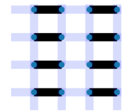
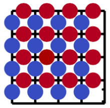
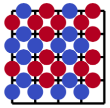
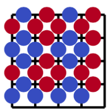
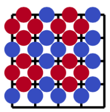
$$D_i := \mathbb{E}\left[\frac{1}{N_{\mathcal{V}}}\sum_k \tau \circ (\Lambda_i \cdot \mathcal{V}_k)\right]$$

$$\mathrm{DSB} := \mathbb{E}\left[\sum_k \tau \circ (\Lambda_C \cdot \mathcal{V}_k)\right]$$

Alet et al. **PRE 74**, 041124 (2006)



- Pairs of C/P filters **label broken symmetry states**

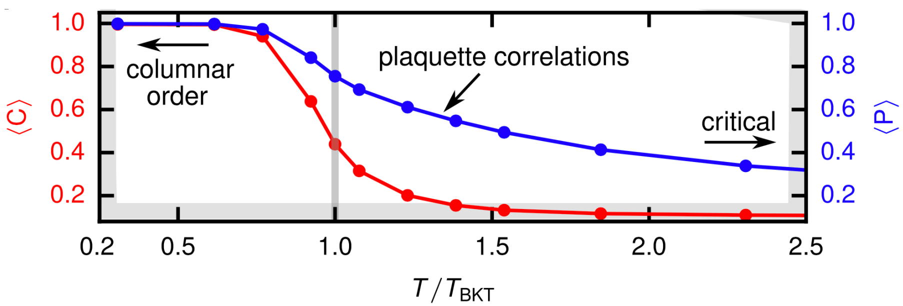
$C(\mathbf{r})$				
Λ_C  Λ_{P1} 	$(-1,-1)$	$(-1,+1)$	$(+1,-1)$	$(+1,+1)$
Λ_{P1}  Λ_{P2} 	$(-1,-1)$	$(+1,+1)$	$(-1,+1)$	$(+1,-1)$

- Filters **define** order parameters:

$$D_i := \mathbb{E}\left[\frac{1}{N_{\mathcal{V}}}\sum_k \tau \circ (\Lambda_i \cdot \mathcal{V}_k)\right]$$

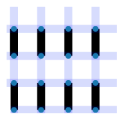
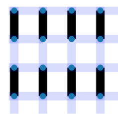
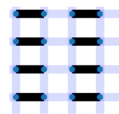
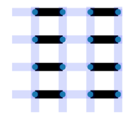
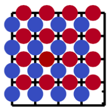
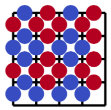
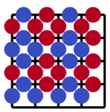
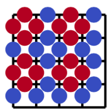
$$\mathrm{DSB} := \mathbb{E}\left[\sum_k \tau \circ (\Lambda_\mathrm{C} \cdot \mathcal{V}_k)\right]$$

Alet et al. **PRE 74**, 041124 (2006)



- Filters **are** relevant operators:

- Pairs of C/P filters **label broken symmetry states**

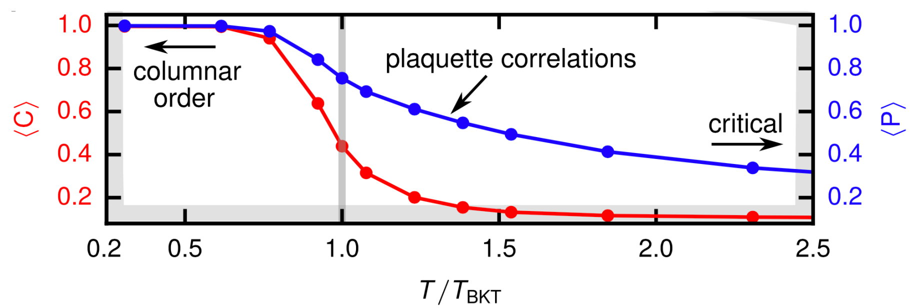
$C(\mathbf{r})$				
Λ_C  Λ_{P1} 	$(-1,-1)$	$(-1,+1)$	$(+1,-1)$	$(+1,+1)$
Λ_{P1}  Λ_{P2} 	$(-1,-1)$	$(+1,+1)$	$(-1,+1)$	$(+1,-1)$

- Filters **define** order parameters:

$$D_i := \mathbb{E}\left[\frac{1}{N_{\mathcal{V}}}\sum_k \tau \circ (\Lambda_i \cdot \mathcal{V}_k)\right]$$

$$\mathrm{DSB} := \mathbb{E}\left[\sum_k \tau \circ (\Lambda_C \cdot \mathcal{V}_k)\right]$$

Alet et al. **PRE 74**, 041124 (2006)

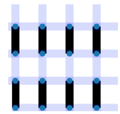
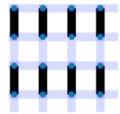
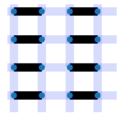
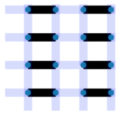
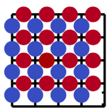
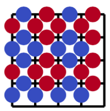
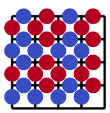
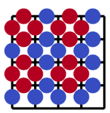


- Filters **are** relevant operators:

$$\mathcal{O}_n(\varphi) = (\cos(n\varphi), \sin(n\varphi))$$

Papanikolaou et al. **PRB 76**, 134514 (2007)

- Pairs of C/P filters **label broken symmetry** states

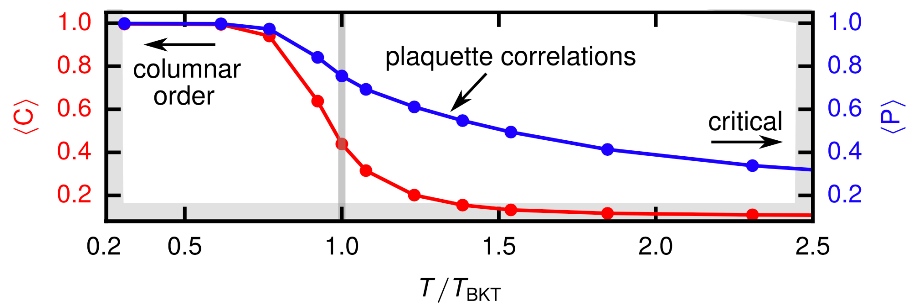
$C(\mathbf{r})$				
Λ_C  Λ_{P1} 	$(-1,-1)$	$(-1,+1)$	$(+1,-1)$	$(+1,+1)$
Λ_{P1}  Λ_{P2} 	$(-1,-1)$	$(+1,+1)$	$(-1,+1)$	$(+1,-1)$
$\varphi(\mathbf{r})$	$\frac{\pi}{2}$	$\frac{3\pi}{2}$	π	0
$\mathcal{O}_1(\varphi) = (\cos\varphi, \sin\varphi)$	$(0,1)$	$(0,-1)$	$(-1,0)$	$(+1,0)$
$\mathcal{O}_2(\varphi) = \cos(2\varphi)$	-1	-1	$+1$	$+1$

- Filters **define** order parameters:

$$D_i := \mathbb{E}\left[\frac{1}{N_{\mathcal{V}}}\sum_k \tau \circ (\Lambda_i \cdot \mathcal{V}_k)\right]$$

$$\mathrm{DSB} := \mathbb{E}\left[\sum_k \tau \circ (\Lambda_{\mathrm{C}} \cdot \mathcal{V}_k)\right]$$

Alet et al. **PRE 74**, 041124 (2006)

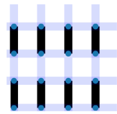
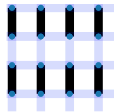
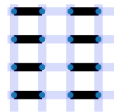
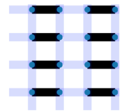
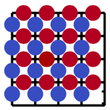
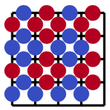
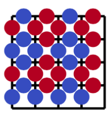
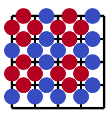


- Filters **are** relevant operators:

$$\mathcal{O}_n(\varphi) = (\cos(n\varphi), \sin(n\varphi))$$

Papanikolaou et al. **PRB 76**, 134514 (2007)

- Pairs of C/P filters **label broken symmetry states**

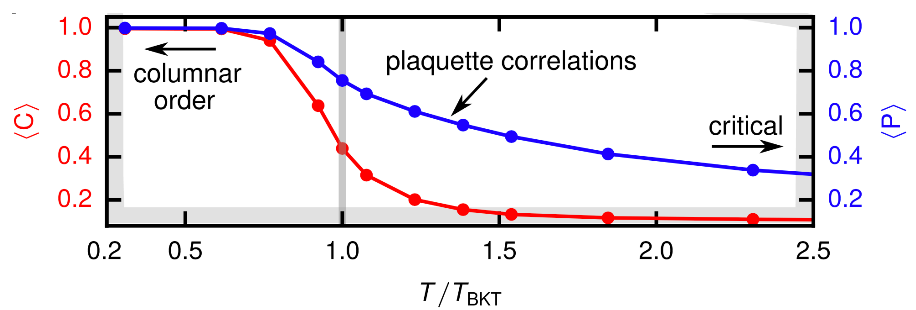
$C(\mathbf{r})$				
Λ_C  Λ_{P1} 	$(-1,-1)$	$(-1,+1)$	$(+1,-1)$	$(+1,+1)$
Λ_{P1}  Λ_{P2} 	$(-1,-1)$	$(+1,+1)$	$(-1,+1)$	$(+1,-1)$
$\varphi(\mathbf{r})$	$\frac{\pi}{2}$	$\frac{3\pi}{2}$	π	0
$\mathcal{O}_1(\varphi) = (\cos\varphi, \sin\varphi)$	$(0,1)$	$(0,-1)$	$(-1,0)$	$(+1,0)$
$\mathcal{O}_2(\varphi) = \cos(2\varphi)$	-1	-1	$+1$	$+1$

- Filters **define** order parameters:

$$D_i := \mathbb{E}\left[\frac{1}{N_\mathcal{V}}\sum_k \tau \circ (\Lambda_i \cdot \mathcal{V}_k)\right]$$

$$\mathrm{DSB} := \mathbb{E}\left[\sum_k \tau \circ (\Lambda_\mathrm{C} \cdot \mathcal{V}_k)\right]$$

Alet et al. **PRE 74**, 041124 (2006)



- Filters **are** relevant operators:

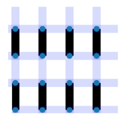
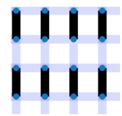
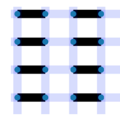
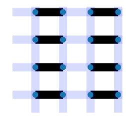
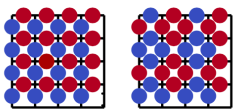
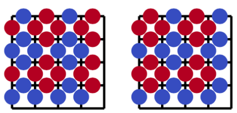
$$\mathcal{O}_n(\varphi) = (\cos(n\varphi), \sin(n\varphi))$$

Papanikolaou et al. **PRB 76**, 134514 (2007)

$$(\Lambda_\mathrm{P1}, \Lambda_\mathrm{P1}) \circ \varphi = (\cos(\varphi + \pi/4), \sin(\varphi + \pi/4))$$

$$\Lambda_\mathrm{C} \circ \varphi = \cos(2\varphi)$$

- Pairs of C/P filters **label broken symmetry states**

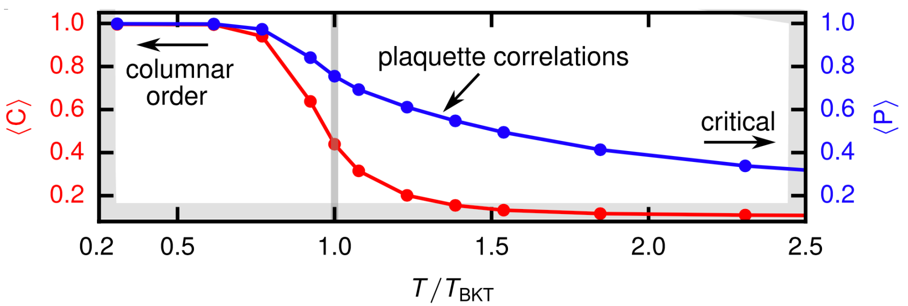
$C(\mathbf{r})$				
Λ_{C} Λ_{P1} 	$(-1,-1)$	$(-1,+1)$	$(+1,-1)$	$(+1,+1)$
Λ_{P1} Λ_{P2} 	$(-1,-1)$	$(+1,+1)$	$(-1,+1)$	$(+1,-1)$
$\varphi(\mathbf{r})$	$\frac{\pi}{2}$	$\frac{3\pi}{2}$	π	0
$\mathcal{O}_1(\varphi) = (\cos\varphi, \sin\varphi)$	$(0,1)$	$(0,-1)$	$(-1,0)$	$(+1,0)$
$\mathcal{O}_2(\varphi) = \cos(2\varphi)$	-1	-1	$+1$	$+1$

- Filters **define** order parameters:

$$D_i := \mathbb{E}\left[\frac{1}{N_{\mathcal{V}}}\sum_k \tau \circ (\Lambda_i \cdot \mathcal{V}_k)\right]$$

$$\mathrm{DSB} := \mathbb{E}\left[\sum_k \tau \circ (\Lambda_{\mathrm{C}} \cdot \mathcal{V}_k)\right]$$

Alet et al. **PRE 74**, 041124 (2006)



- Filters **are** relevant operators:

$$\mathcal{O}_n(\varphi) = (\cos(n\varphi), \sin(n\varphi))$$

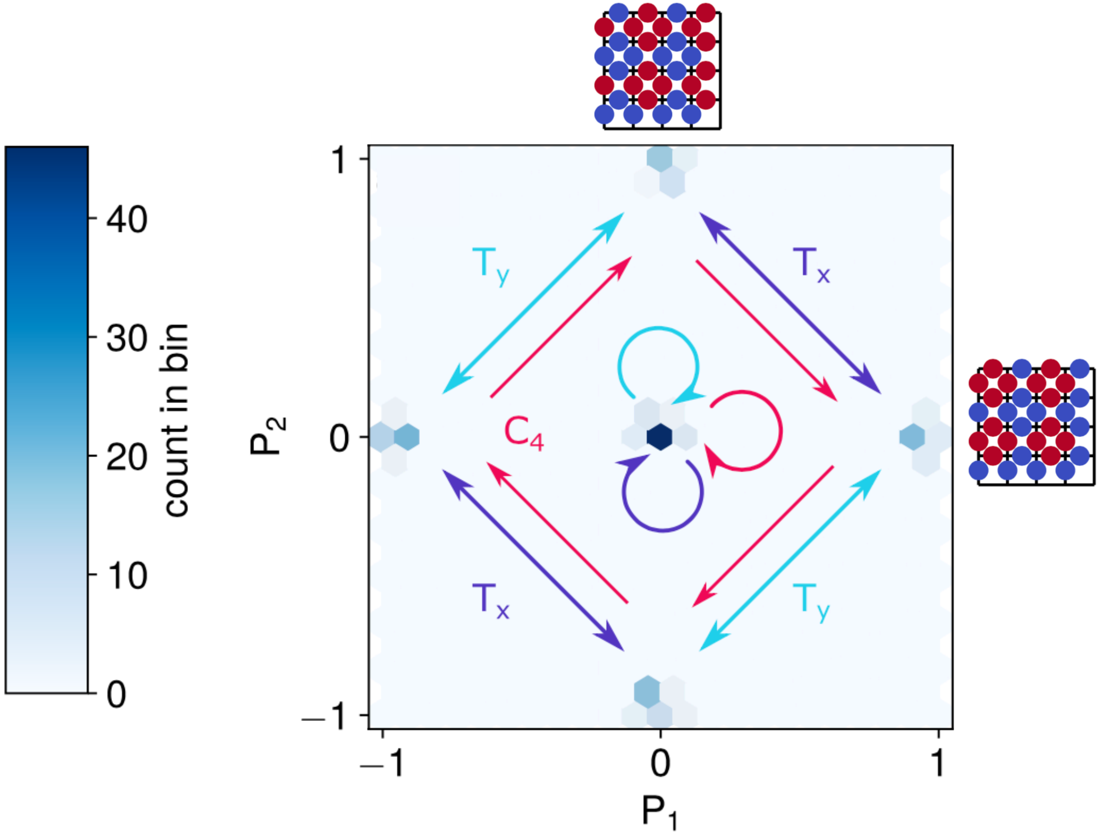
Papanikolaou et al. **PRB 76**, 134514 (2007)

$$(\Lambda_{\mathrm{P1}}, \Lambda_{\mathrm{P1}}) \circ \varphi = (\cos(\varphi + \pi/4), \sin(\varphi + \pi/4))$$

$$\Lambda_{\mathrm{C}} \circ \varphi = \cos(2\varphi)$$

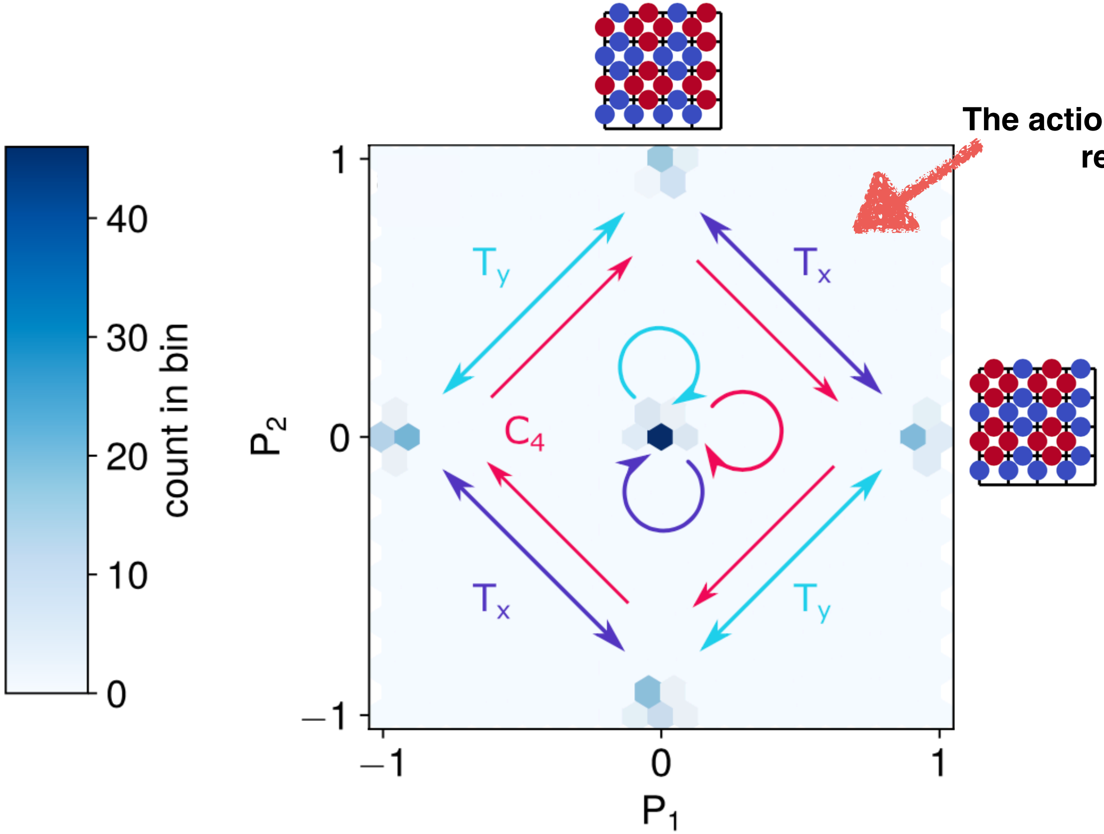
(Also: staggered filters are gradients of the height field)

Symmetries in the RSMI ensemble



PRE, 104, 064106 (2021)
PRL 127, 240603 (2021)

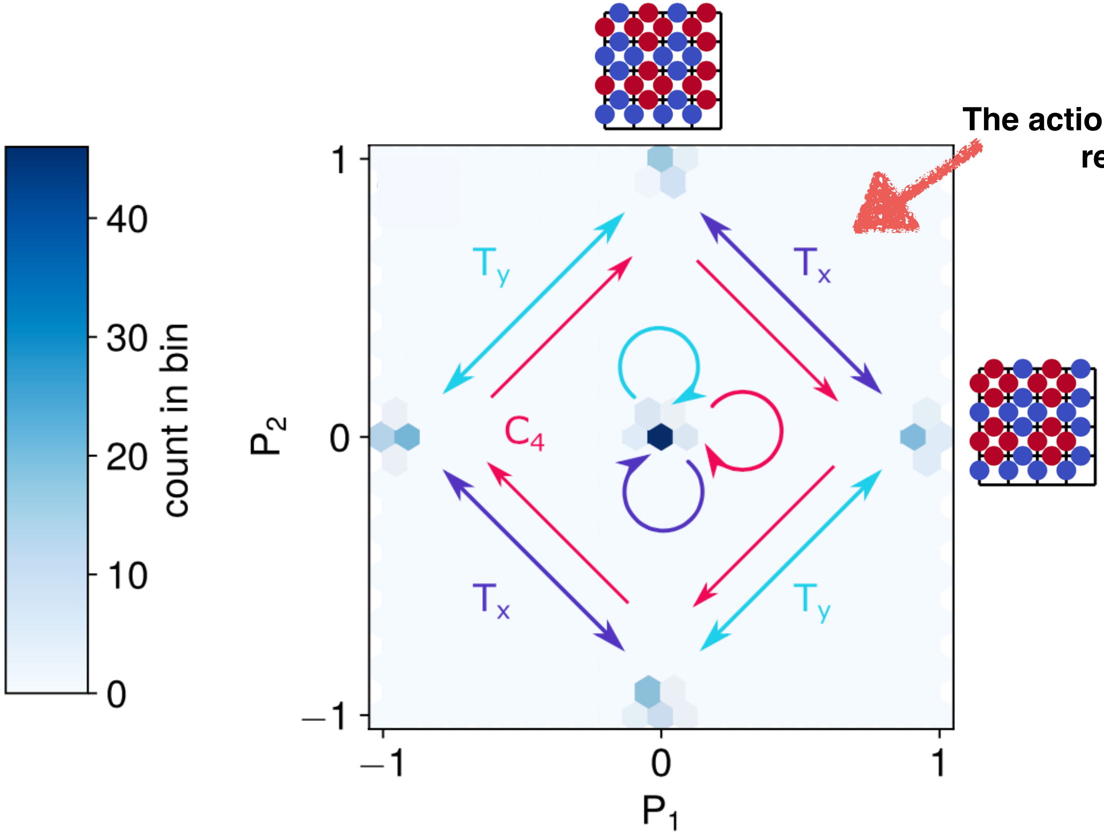
Symmetries in the RSMI ensemble



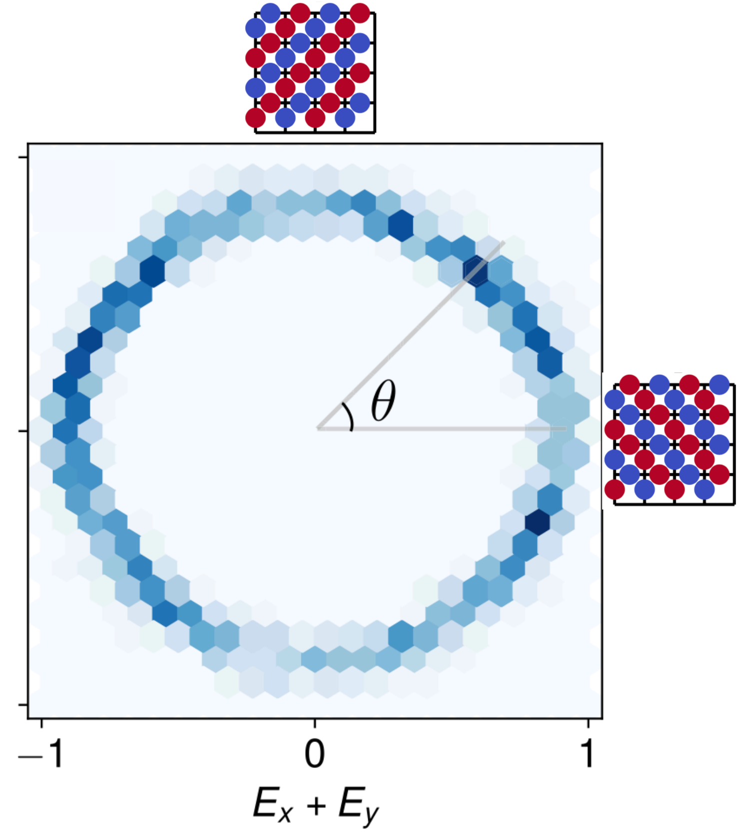
The action of symmetries of the physical state are represented on the space of filters

PRE, 104, 064106 (2021)
PRL 127, 240603 (2021)

Symmetries in the RSMI ensemble

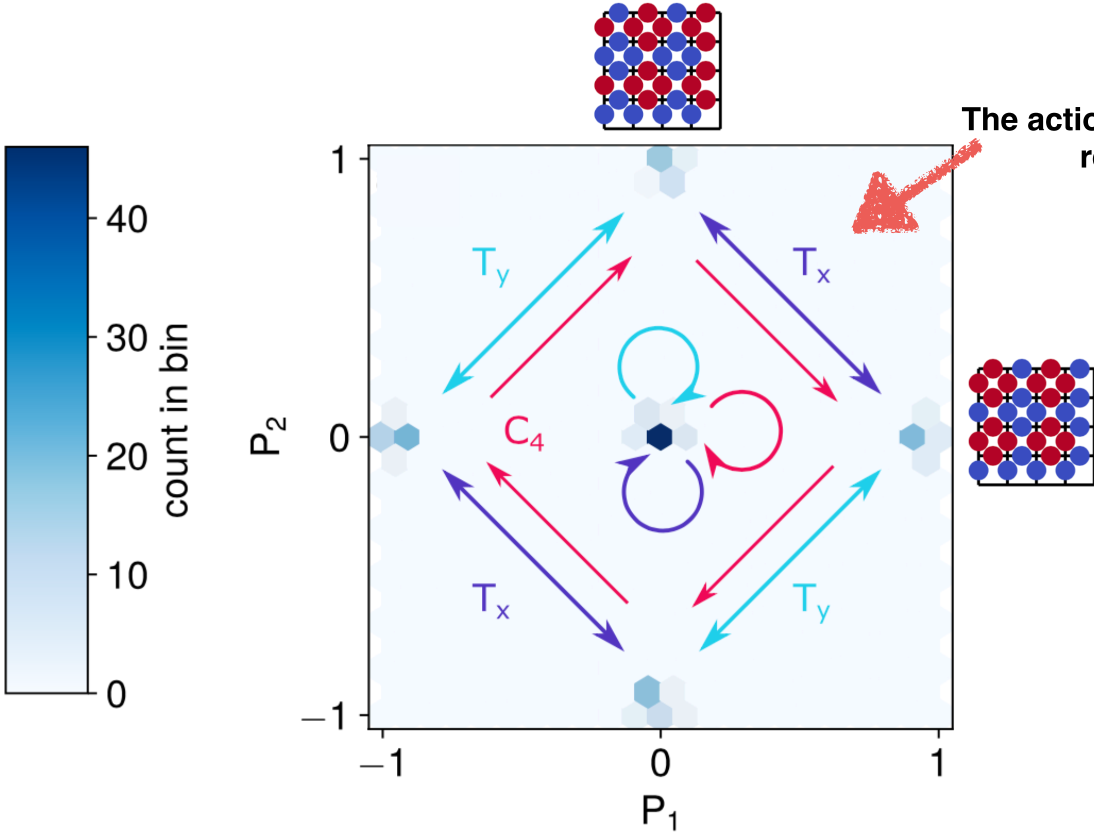


The action of symmetries of the physical state are represented on the space of filters

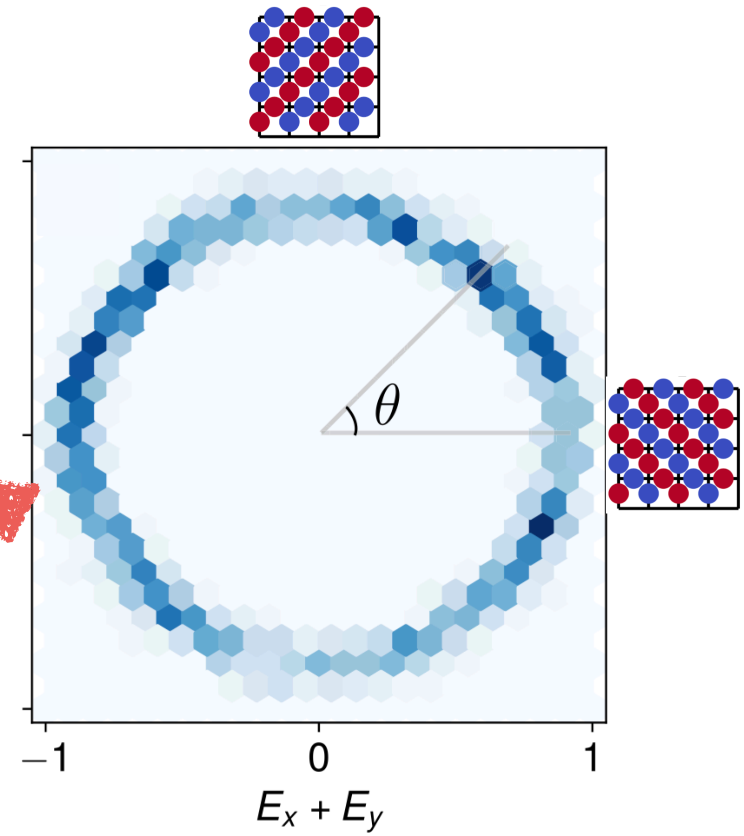


PRE, 104, 064106 (2021)
PRL 127, 240603 (2021)

Symmetries in the RSMI ensemble



The action of symmetries of the physical state are represented on the space of filters



Emergent U(1):

$$R(\vartheta)(\Lambda_{S_1}, \Lambda_{S_2}) \sim [\cos(\vartheta)E_x - \sin(\vartheta)E_y, \sin(\vartheta)E_x + \cos(\vartheta)E_y]$$

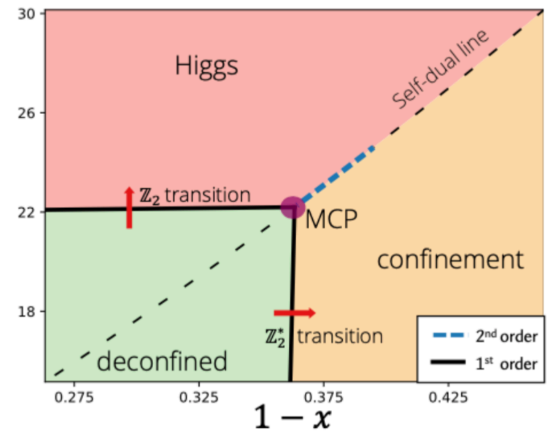
PRE, 104, 064106 (2021)
PRL 127, 240603 (2021)

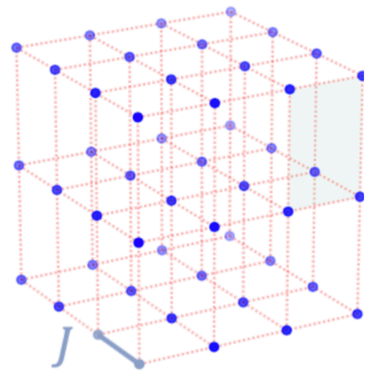
Lattice gauge theories

L. Oppenheim, **MKJ**, S. Gazit, Z. Ringel *arXiv:2311.17994*

Lattice gauge theories

$$\mathcal{S}_{\text{SD-IHG}} = K \sum_{\square} \prod_{\langle i,j \rangle \in \square} \sigma_{ij} + J \sum_{\langle i,j \rangle} \tau_i \sigma_{ij} \tau_j$$

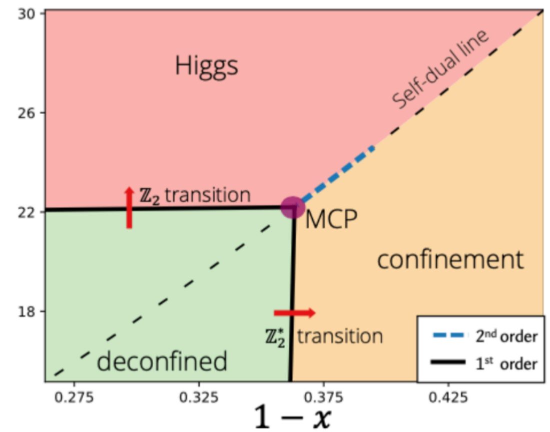


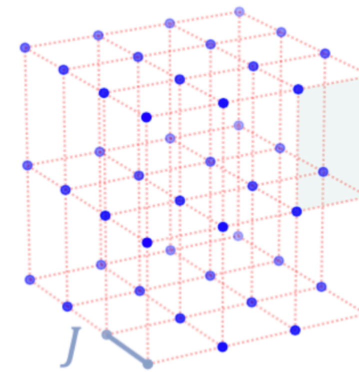


L. Oppenheim, **MKJ**, S. Gazit, Z. Ringel *arXiv:2311.17994*

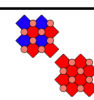
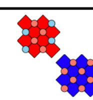
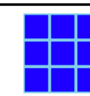
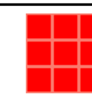
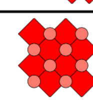
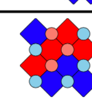
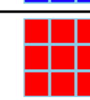
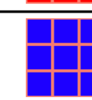
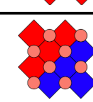
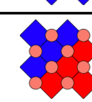
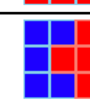
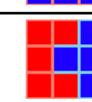
Lattice gauge theories

$$\mathcal{S}_{\text{SD-IHG}} = K \sum_{\square} \prod_{\langle i,j \rangle \in \square} \sigma_{ij} + J \sum_{\langle i,j \rangle} \tau_i \sigma_{ij} \tau_j$$





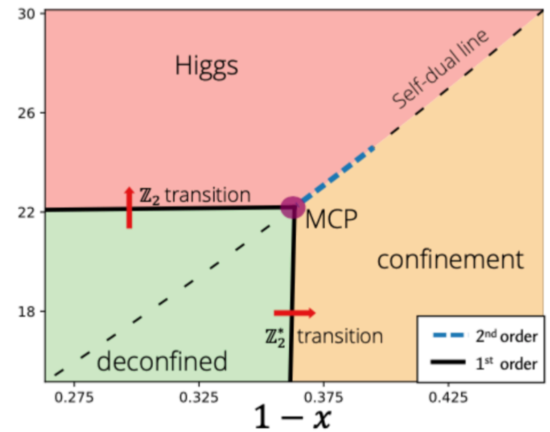
- We can identify subleading operators (or the absence of the expected ones)

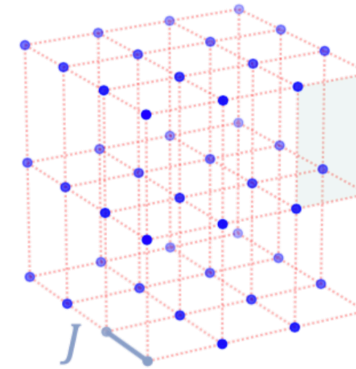
AT-TFI				SD-IHG			
RSMI-NE Scaling Dimension {Expected[46]}	Analytic Operator {Deg.}	Neural Operator Projection: Maximum	Neural Operator Projection: Minimum	RSMI-NE Scaling Dimension {Expected[17]}	Analytic Operator {Deg.}	Neural Operator Projection: Maximum	Neural Operator Projection: Minimum
1.24(1) 1.22(1) {1.23629}	$\langle\sigma\rangle^2 - \langle\tau\rangle^2$ $\langle\sigma\rangle\langle\tau\rangle$ {2}			1.24(1) {1.222}	$\langle A\rangle$ {1}		
1.49(2) {1.51136}	$\langle\sigma\rangle^2 + \langle\tau\rangle^2$ {1}			1.54(2) {1.502}	$\langle S\rangle$ {1}		
2.02(3) {2.0}	$\langle\sigma\rangle\langle\partial\tau\rangle - \langle\tau\rangle\langle\partial\sigma\rangle$ {3}			2.20(6) {2.222}	$\langle\partial A\rangle$ {3}		

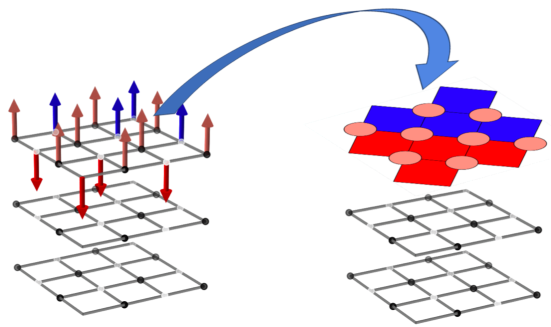
L. Oppenheim, **MKJ**, S. Gazit, Z. Ringel *arXiv:2311.17994*

Lattice gauge theories

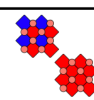
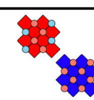
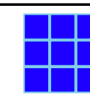
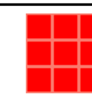
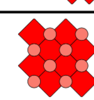
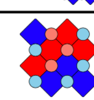
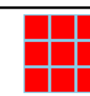
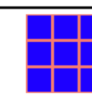
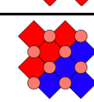
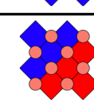
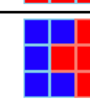
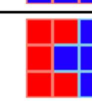
$$\mathcal{S}_{\text{SD-IHG}} = K \sum_{\square} \prod_{\langle i,j \rangle \in \square} \sigma_{ij} + J \sum_{\langle i,j \rangle} \tau_i \sigma_{ij} \tau_j$$







- We can identify subleading operators (or the absence of the expected ones)

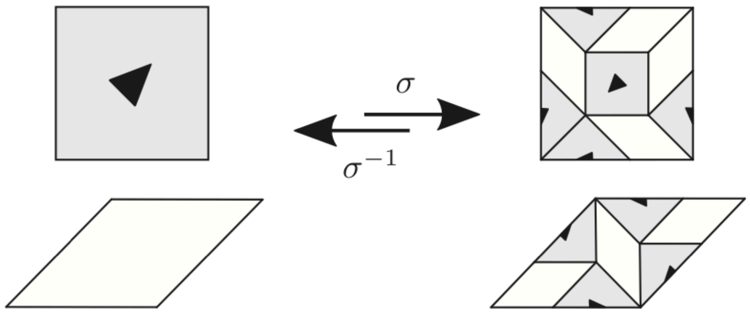
AT-TFI				SD-IHG			
RSMI-NE Scaling Dimension {Expected[46]}	Analytic Operator {Deg.}	Neural Operator Projection: Maximum	Neural Operator Projection: Minimum	RSMI-NE Scaling Dimension {Expected[17]}	Analytic Operator {Deg.}	Neural Operator Projection: Maximum	Neural Operator Projection: Minimum
1.24(1) 1.22(1) {1.23629}	$\langle\sigma\rangle^2 - \langle\tau\rangle^2$ $\langle\sigma\rangle\langle\tau\rangle$ {2}			1.24(1) {1.222}	$\langle A \rangle$ {1}		
1.49(2) {1.51136}	$\langle\sigma\rangle^2 + \langle\tau\rangle^2$ {1}			1.54(2) {1.502}	$\langle S \rangle$ {1}		
2.02(3) {2.0}	$\langle\sigma\rangle\langle\partial\tau\rangle - \langle\tau\rangle\langle\partial\sigma\rangle$ {3}			2.20(6) {2.222}	$\langle \partial A \rangle$ {3}		

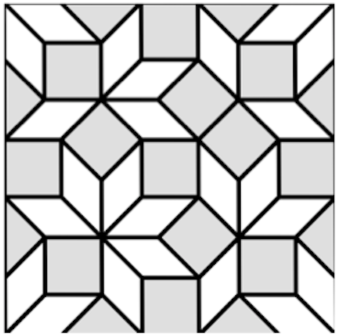
L. Oppenheim, **MKJ**, S. Gazit, Z. Ringel *arXiv:2311.17994*

RG for quasi-periodic systems

RG for quasi-periodic systems

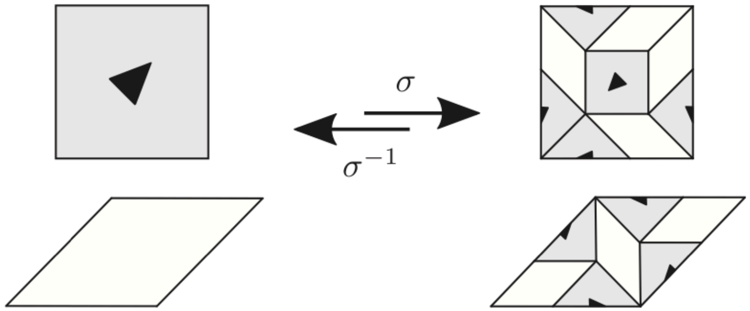
- The quasiperiodic Amman Beenker (AB) tiling is generated hierarchically:

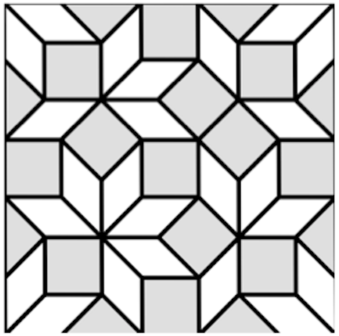


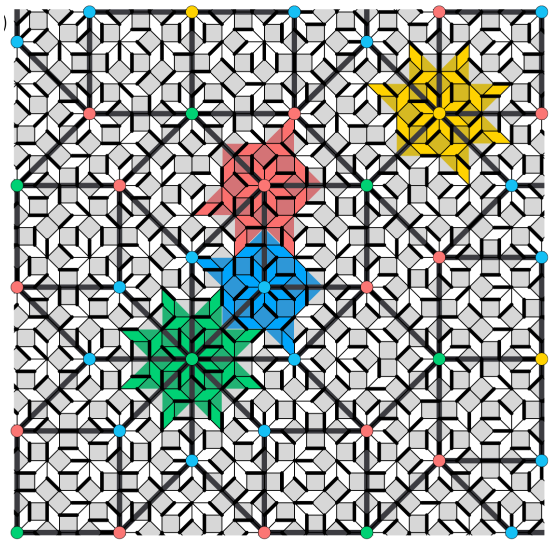


RG for quasi-periodic systems

- The quasiperiodic Amman Beenker (AB) tiling is generated hierarchically:

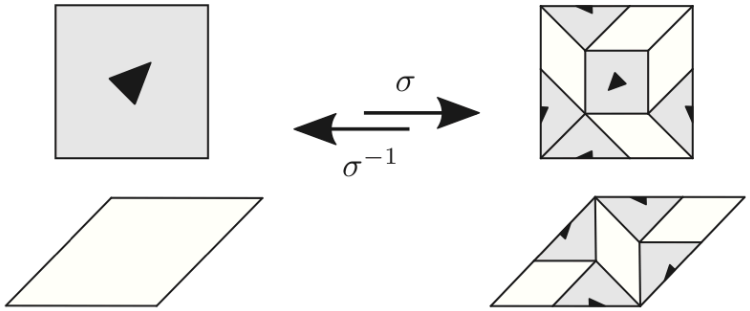


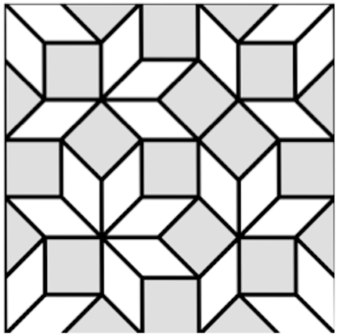




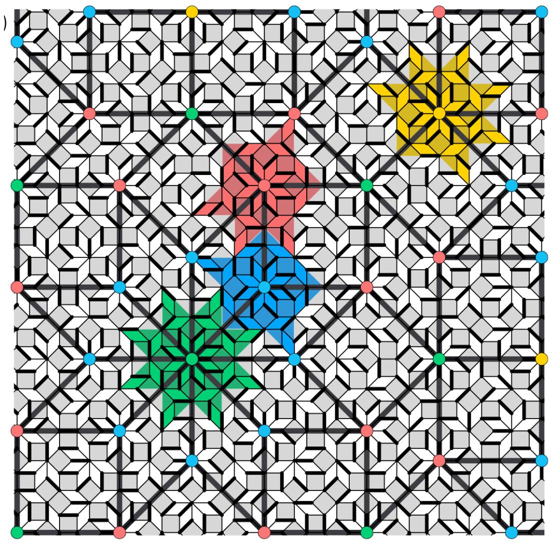
RG for quasi-periodic systems

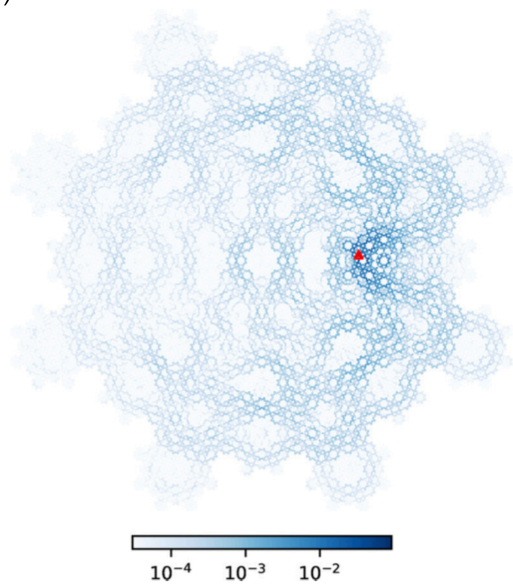
- The quasiperiodic Amman Beenker (AB) tiling is generated hierarchically:





- AB tiling admits perfect dimer covering, and shows evidence of power-law dimer correlations: **Phys. Rev. B** 106, 094202 (2020)

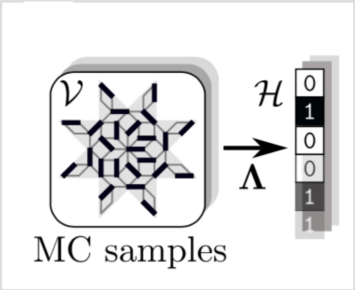




RG for quasi-periodic systems

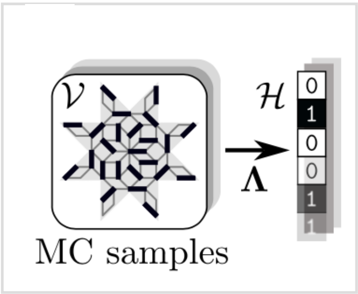
D.E. Gokmen, S. Biswas, S.D. Huber, Z. Ringel, F. Flicker and **MKJ**, *arXiv:2301:11934 (2023)*

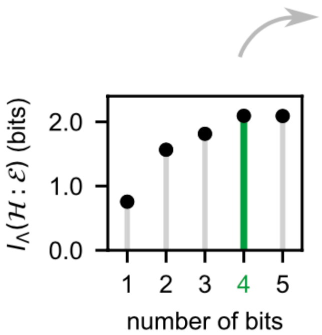
RG for quasi-periodic systems

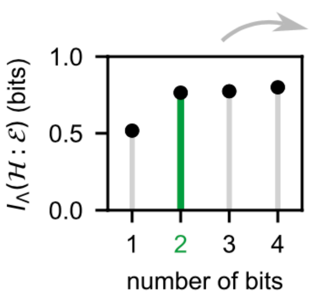


D.E. Gokmen, S. Biswas, S.D. Huber, Z. Ringel, F. Flicker and **MKJ**, *arXiv:2301:11934 (2023)*

RG for quasi-periodic systems

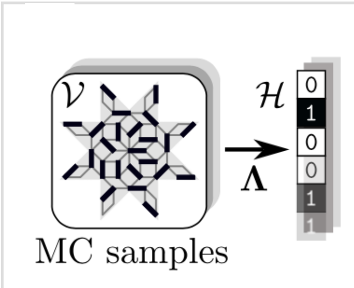


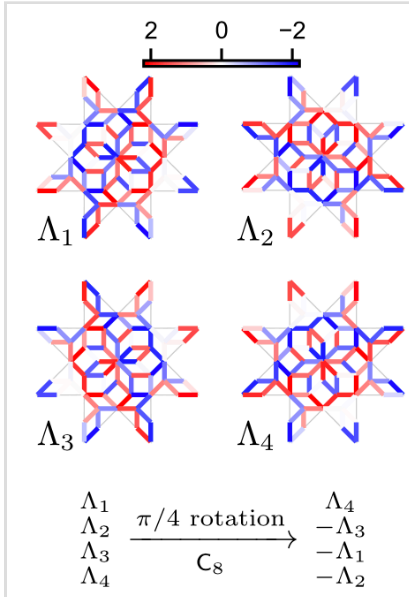


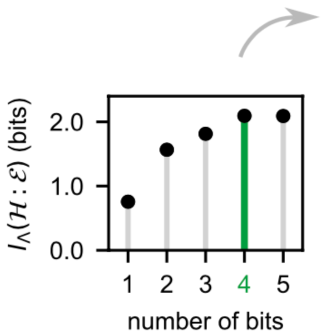


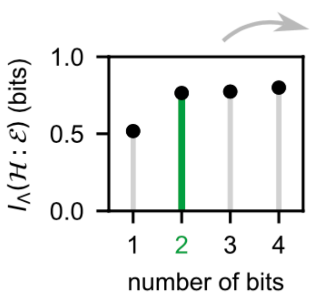
D.E. Gokmen, S. Biswas, S.D. Huber, Z. Ringel, F. Flicker and **MKJ**, *arXiv:2301:11934 (2023)*

RG for quasi-periodic systems



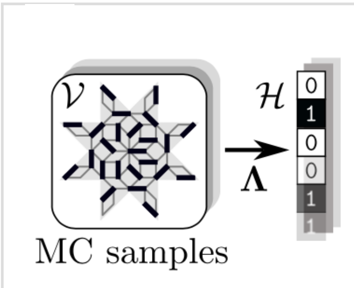


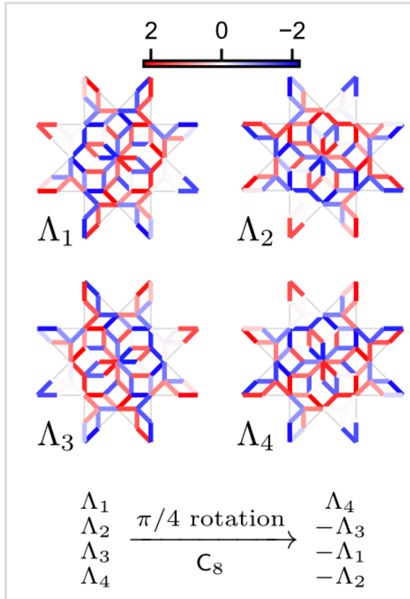


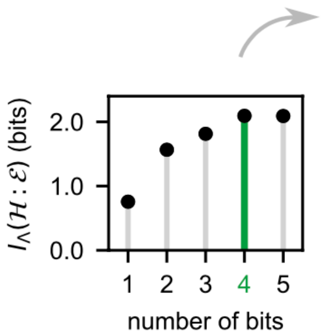


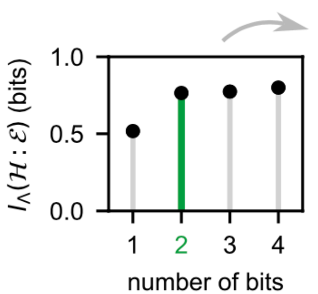
D.E. Gokmen, S. Biswas, S.D. Huber, Z. Ringel, F. Flicker and **MKJ**, *arXiv:2301:11934 (2023)*

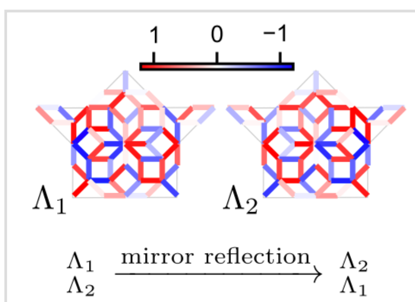
RG for quasi-periodic systems





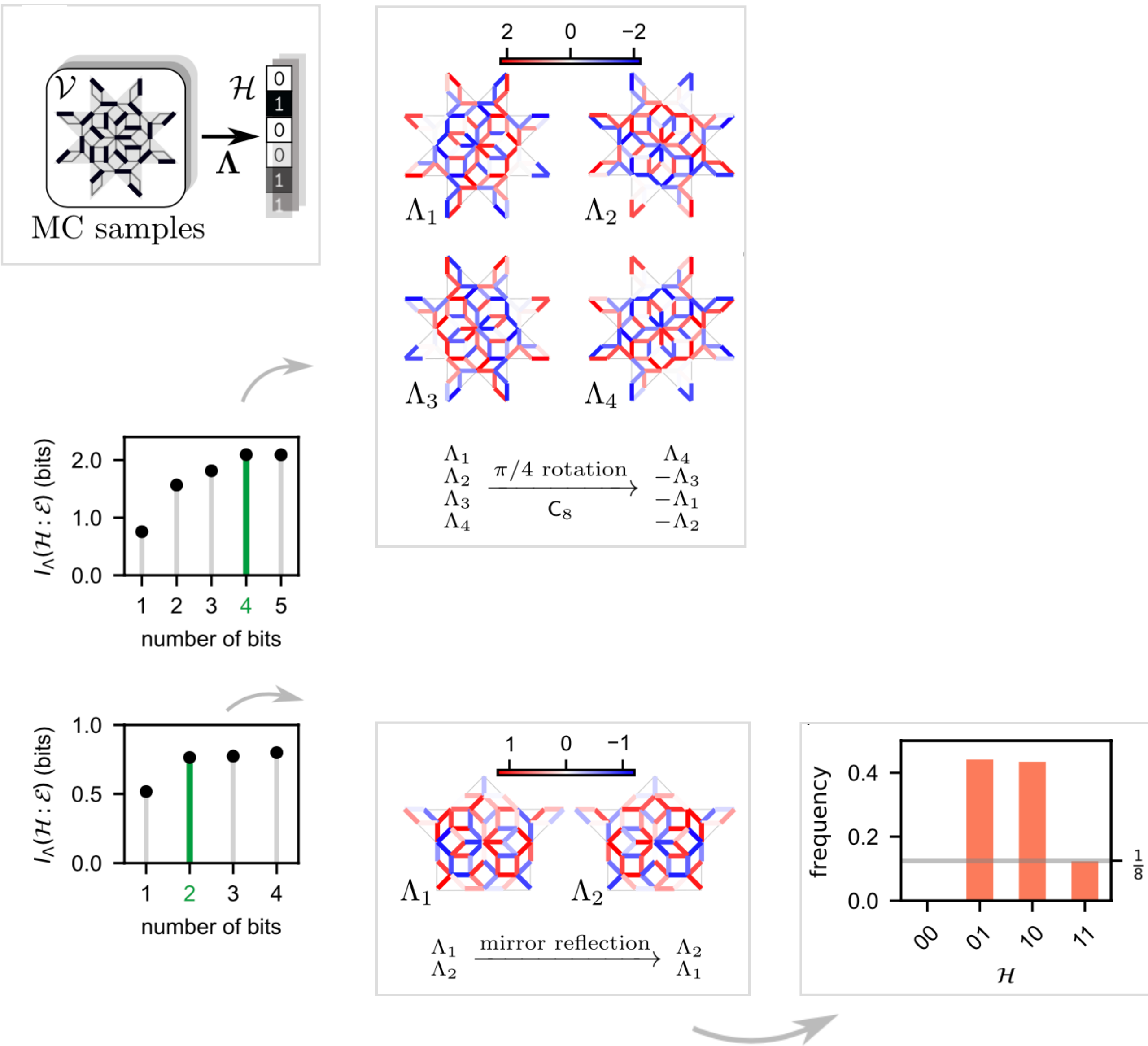






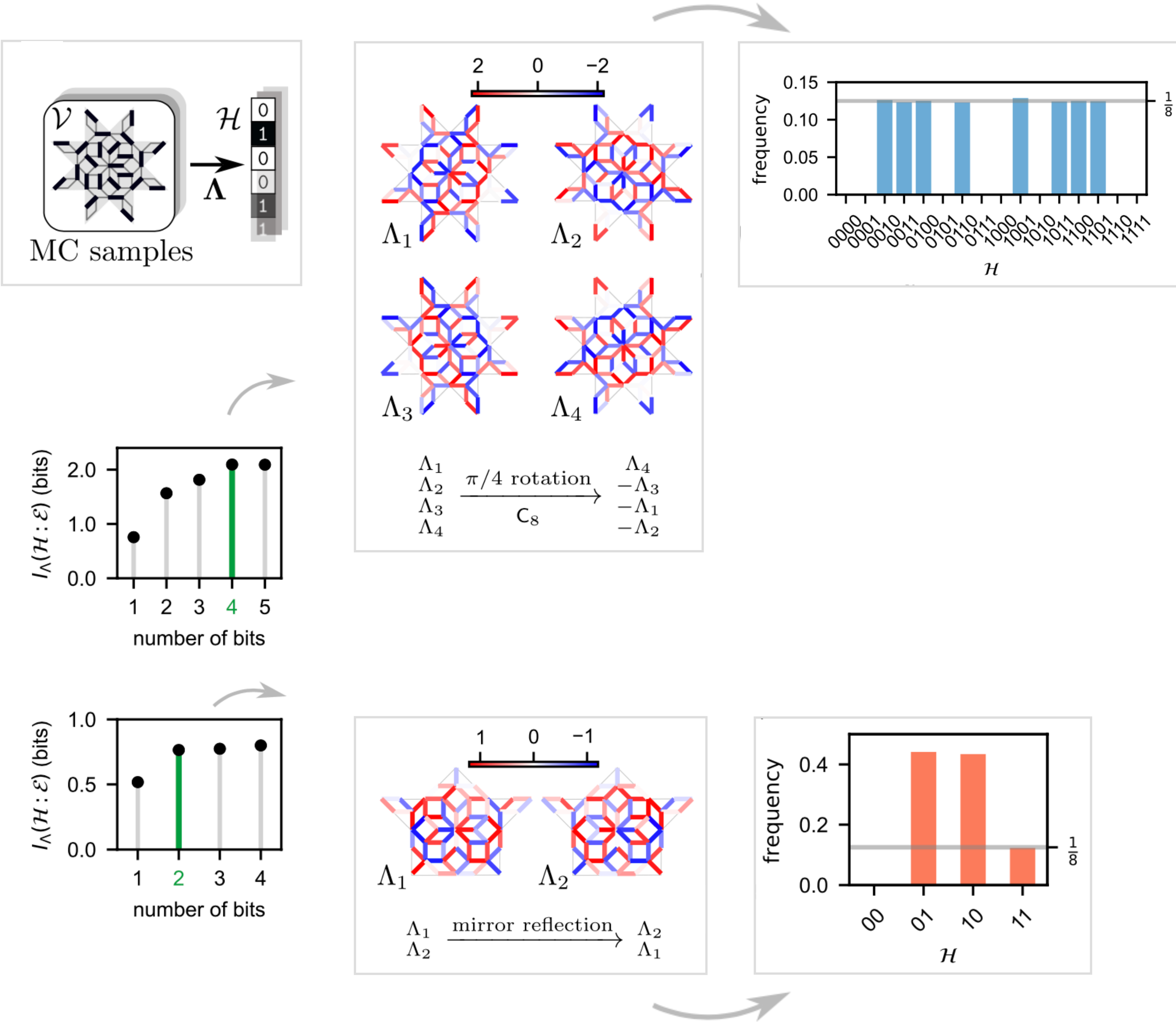
D.E. Gokmen, S. Biswas, S.D. Huber, Z. Ringel, F. Flicker and **MKJ**, *arXiv:2301:11934 (2023)*

RG for quasi-periodic systems



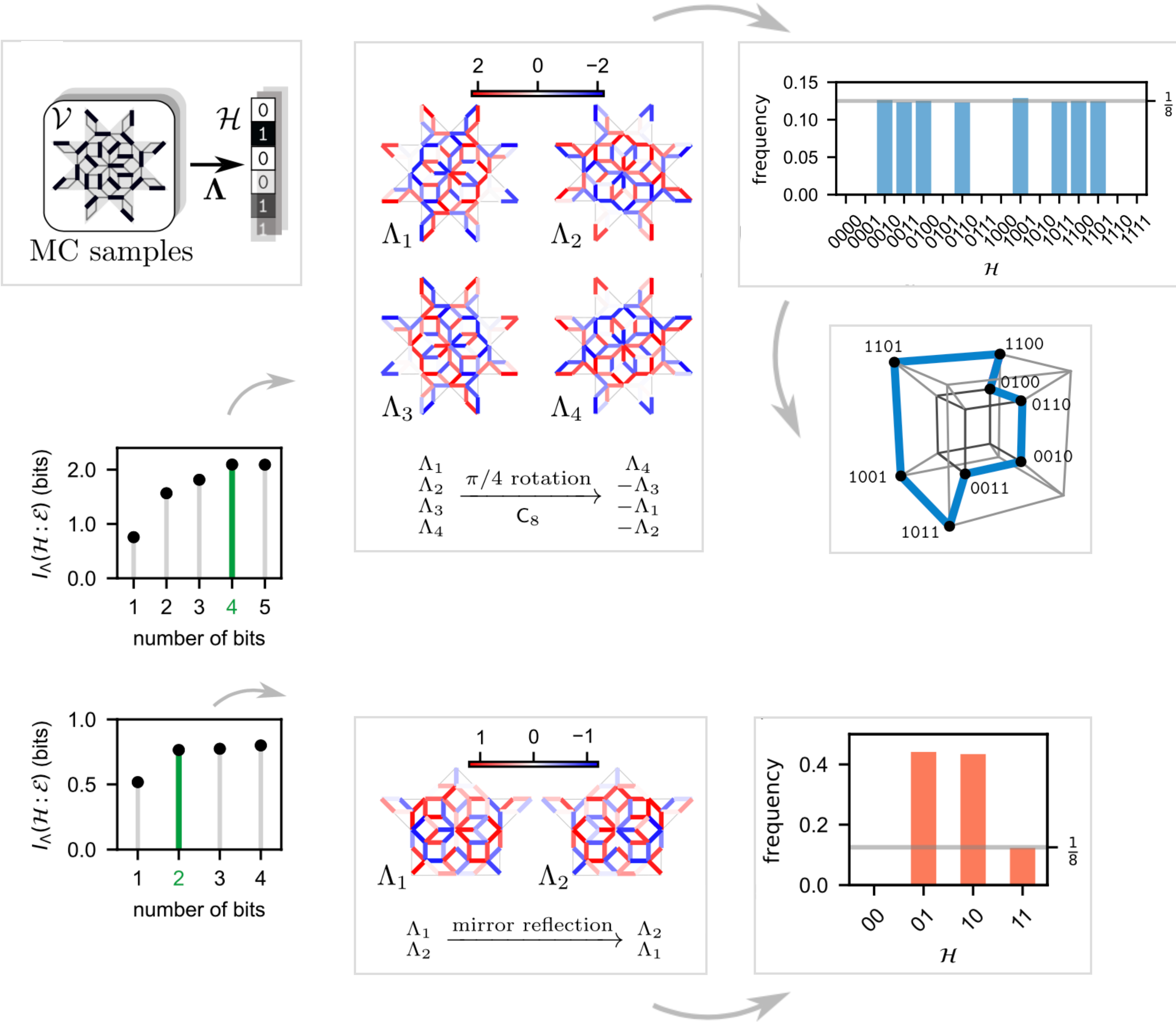
D.E. Gokmen, S. Biswas, S.D. Huber, Z. Ringel, F. Flicker and **MKJ**, *arXiv:2301:11934 (2023)*

RG for quasi-periodic systems



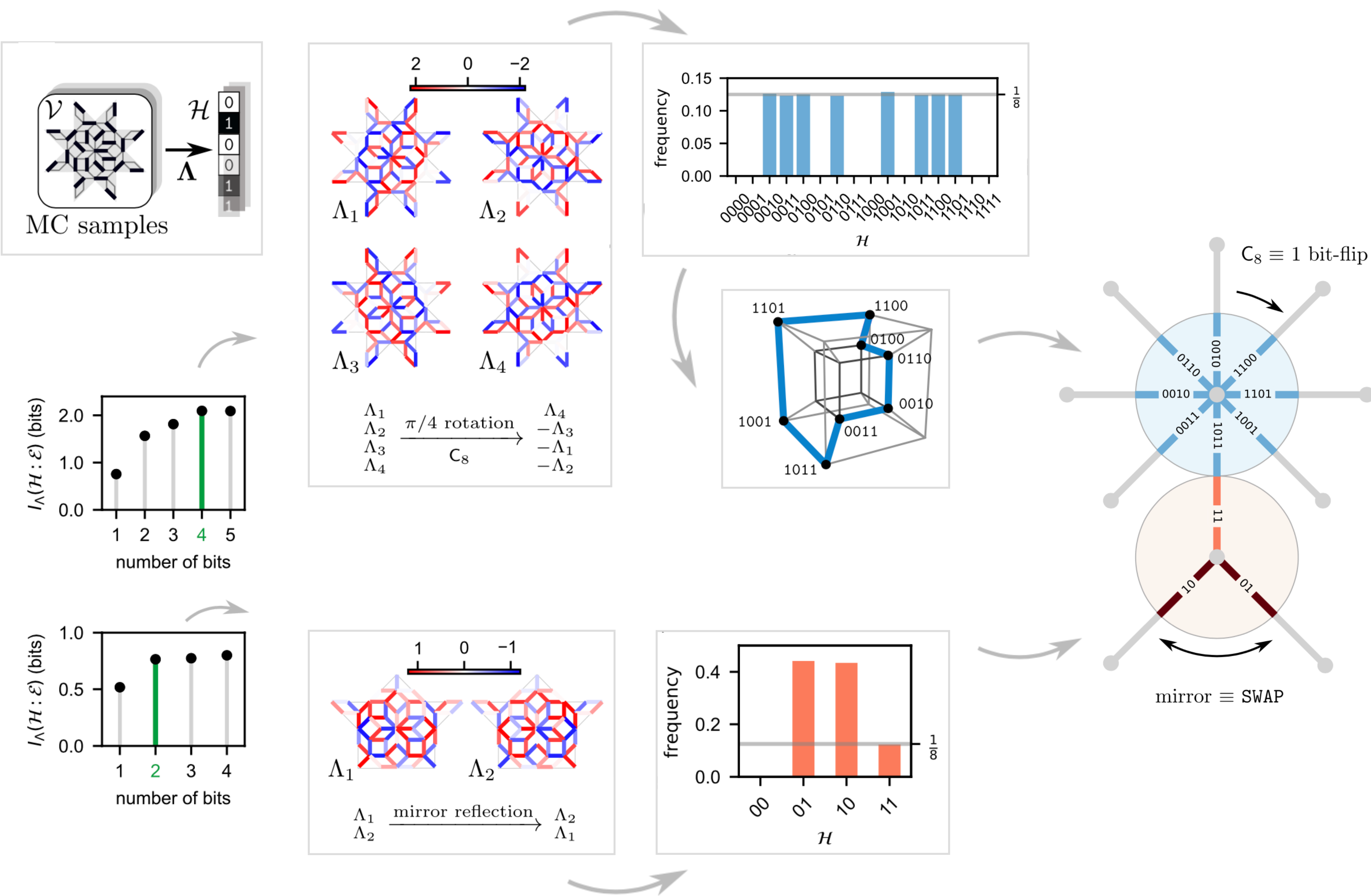
D.E. Gokmen, S. Biswas, S.D. Huber, Z. Ringel, F. Flicker and **MKJ**, *arXiv:2301:11934 (2023)*

RG for quasi-periodic systems



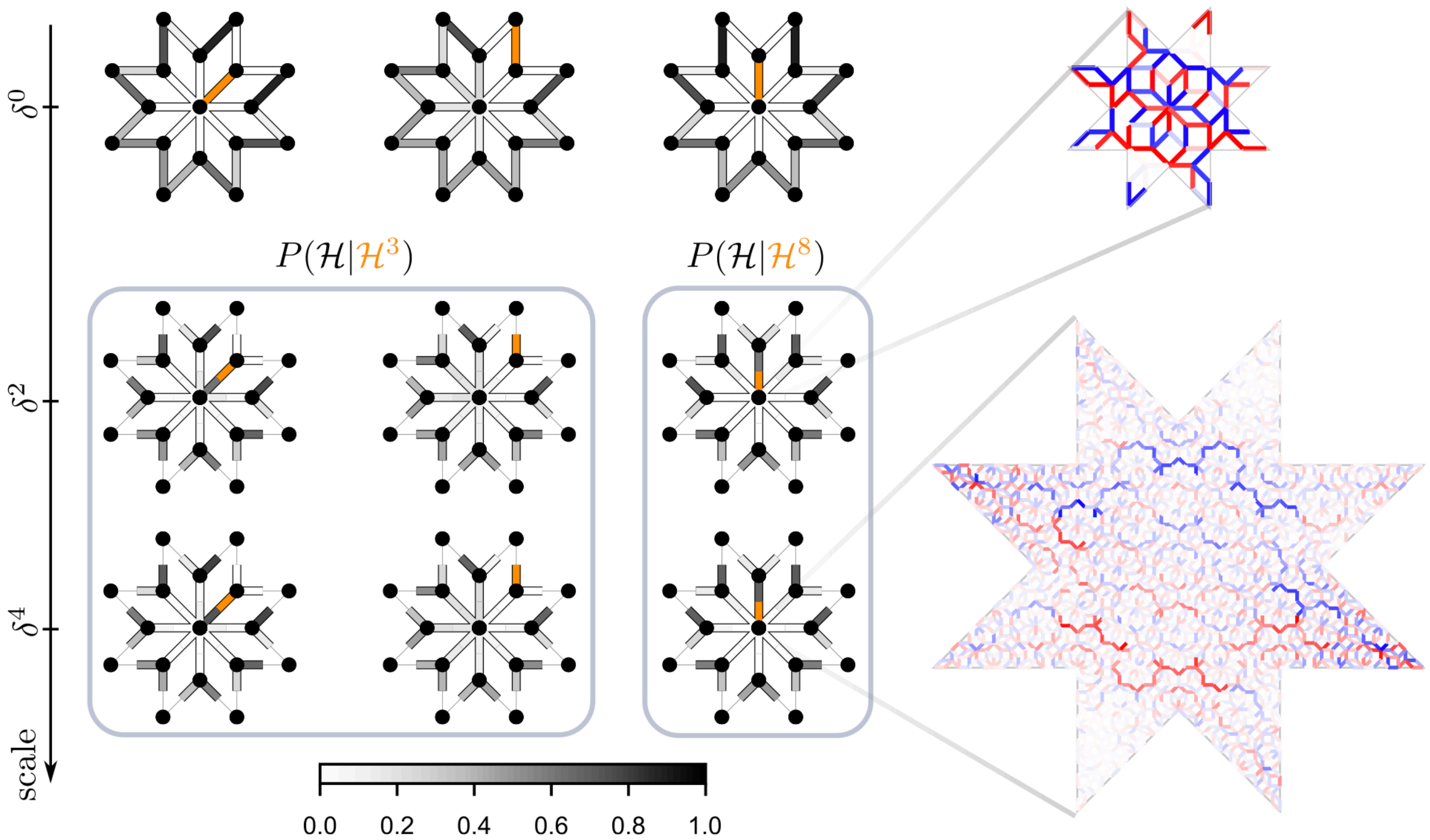
D.E. Gokmen, S. Biswas, S.D. Huber, Z. Ringel, F. Flicker and **MKJ**, *arXiv:2301:11934 (2023)*

RG for quasi-periodic systems



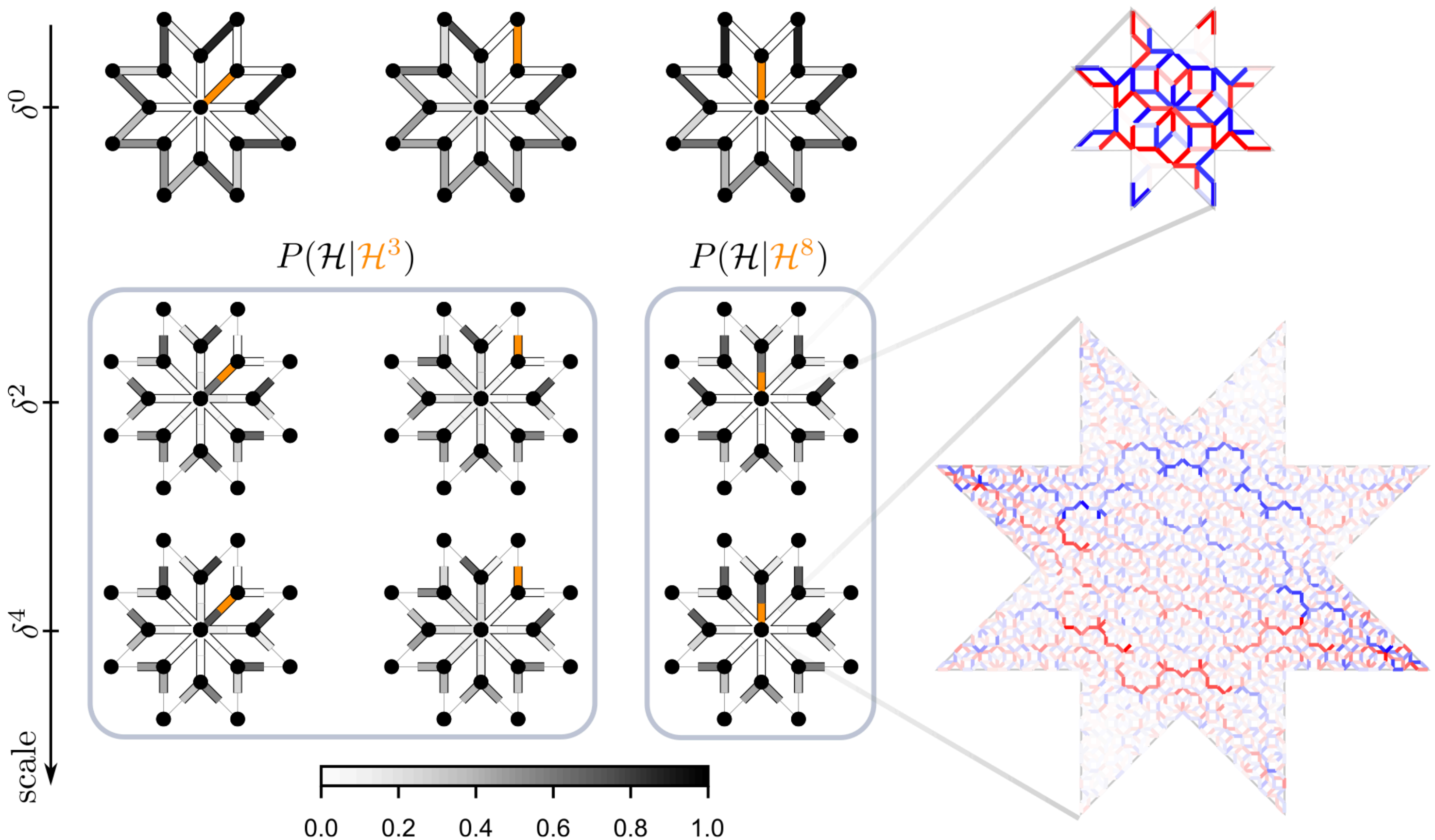
D.E. Gokmen, S. Biswas, S.D. Huber, Z. Ringel, F. Flicker and **MKJ**, *arXiv:2301:11934 (2023)*

RG for quasi-periodic systems



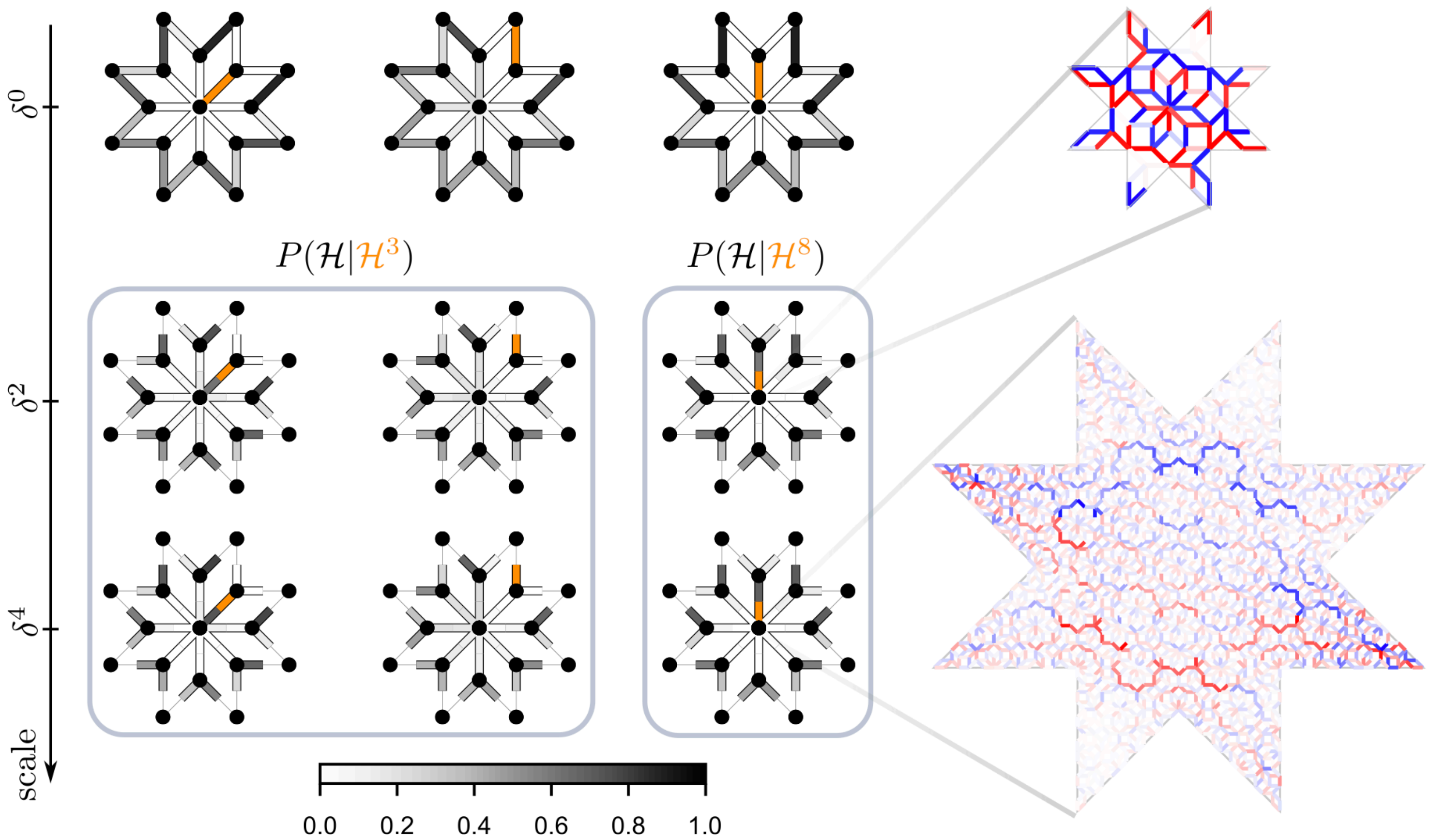
RG for quasi-periodic systems

- The compression map reveals effective **super-dimers** on a larger scale:



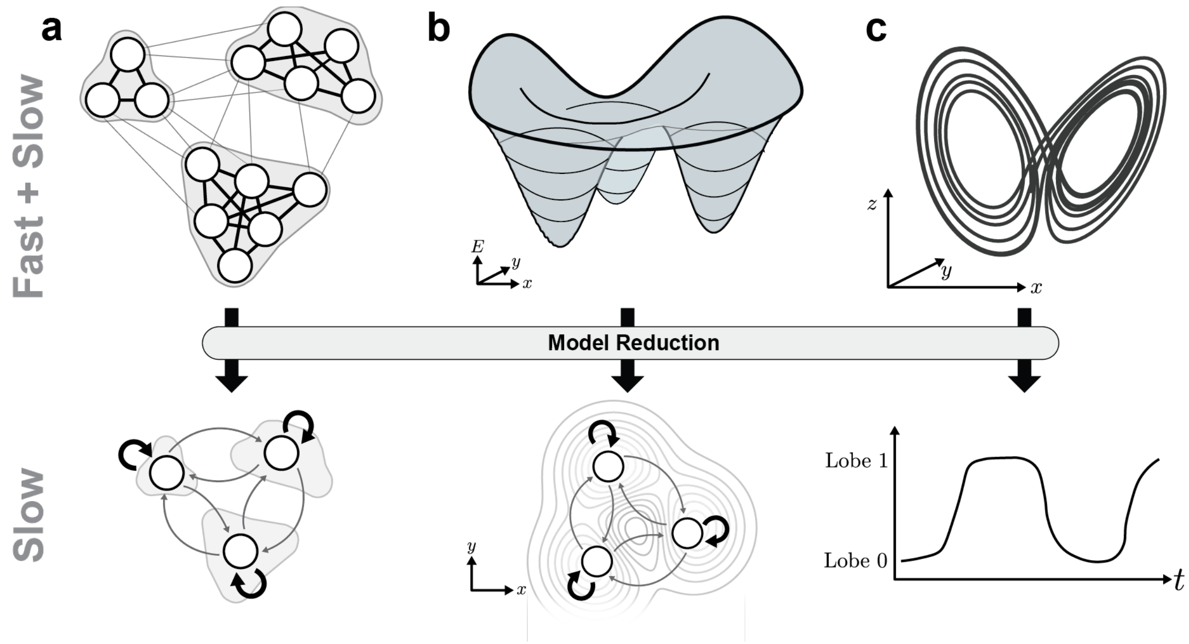
RG for quasi-periodic systems

- The compression map reveals effective **super-dimers** on a larger scale:



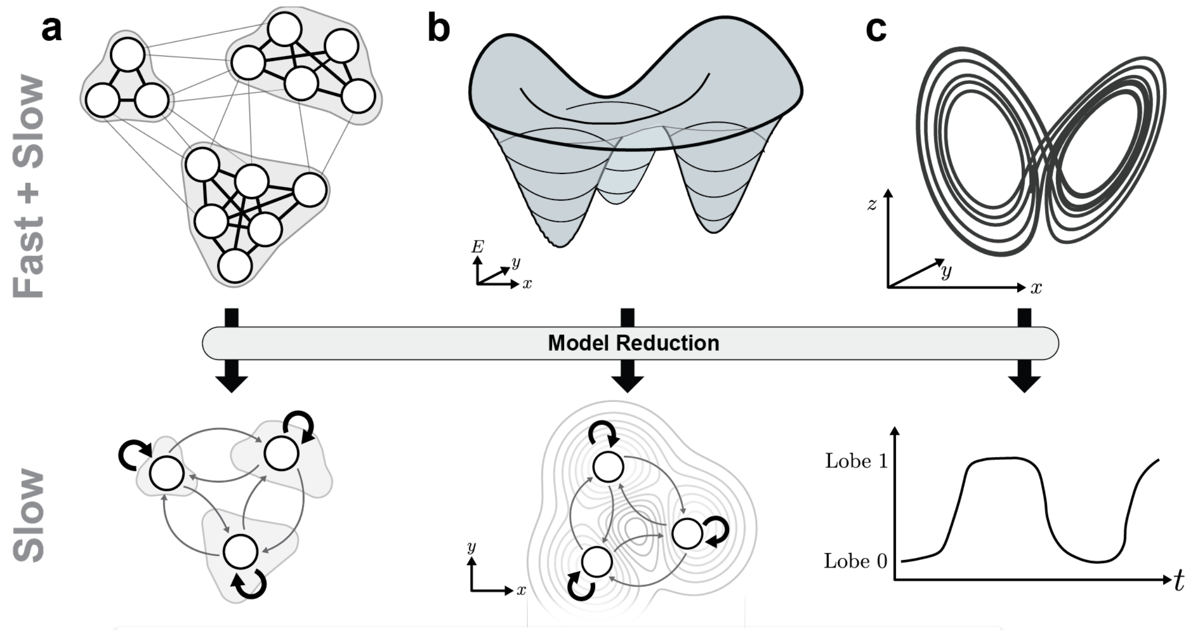
- The same compression maps persist across **multiple scales**

Dynamical systems



$$p^*_\beta(h|x) = \frac{1}{\mathcal{N}(x)}\, p^*_\beta(h)\, \exp\left[\beta \sum_n \mathrm{e}^{\lambda_n \Delta t} \phi_n(x) f_n(h)\right]$$

Dynamical systems

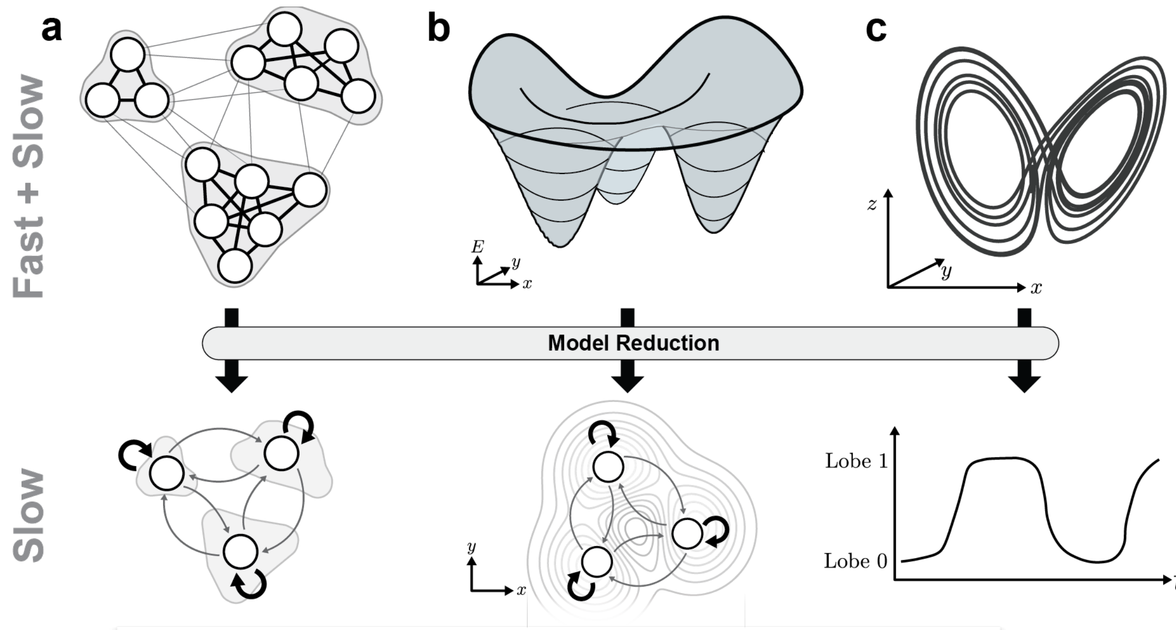


$$\mathcal{L}_{\mathrm{IB}}[p_{H_t|X_t}] = I(X_t, H_t) - \beta I(X_{t+\Delta t}, H_t)$$

- Compress to preserve information about the **future state** of the system.

$$p^*_\beta(h|x) = \frac{1}{\mathcal{N}(x)}\, p^*_\beta(h)\, \exp\left[\beta \sum_n \mathrm{e}^{\lambda_n \Delta t} \phi_n(x) f_n(h)\right]$$

Dynamical systems



$$\mathcal{L}_{\text{IB}}[p_{H_t|X_t}] = I(X_t, H_t) - \beta I(X_{t+\Delta t}, H_t)$$

- Compress to preserve information about the **future state** of the system.

- The **IB-optimal** encoder is determined by the eigenmodes of the transfer operator

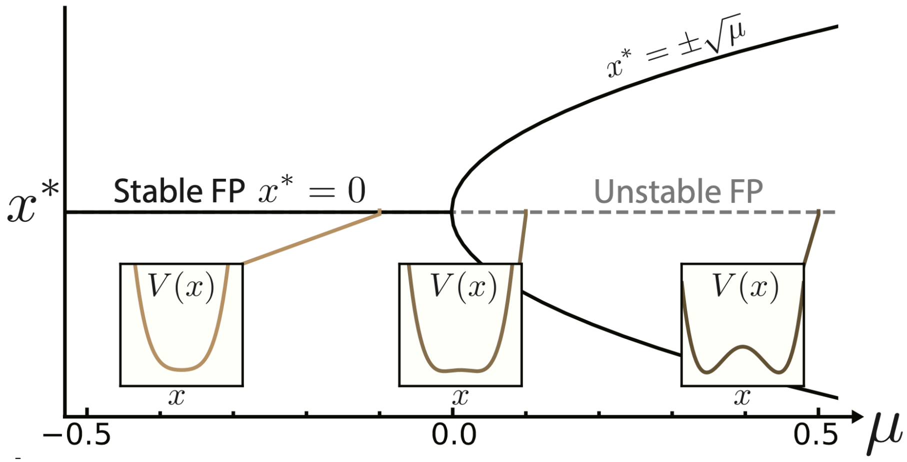
$$p^*_\beta(h|x) = \frac{1}{\mathcal{N}(x)}\, p^*_\beta(h)\, \exp\left[\beta \sum_n \mathrm{e}^{\lambda_n \Delta t} \phi_n(x) f_n(h)\right]$$

Brownian particle in a potential

$$\dot{x}_t = -\partial_x V(x_t) + \sigma \eta_t.$$

$$V(x) = \frac{1}{4}(\mu - x^2)^2$$

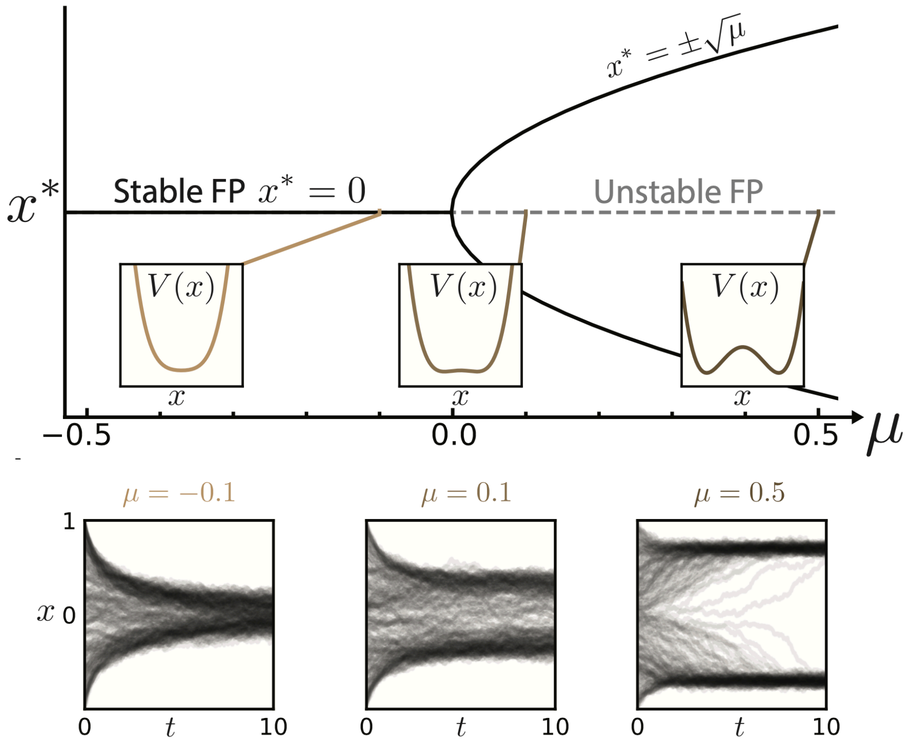
Brownian particle in a potential



$$\dot{x}_t = -\partial_x V(x_t) + \sigma \eta_t.$$

$$V(x) = \frac{1}{4}(\mu - x^2)^2$$

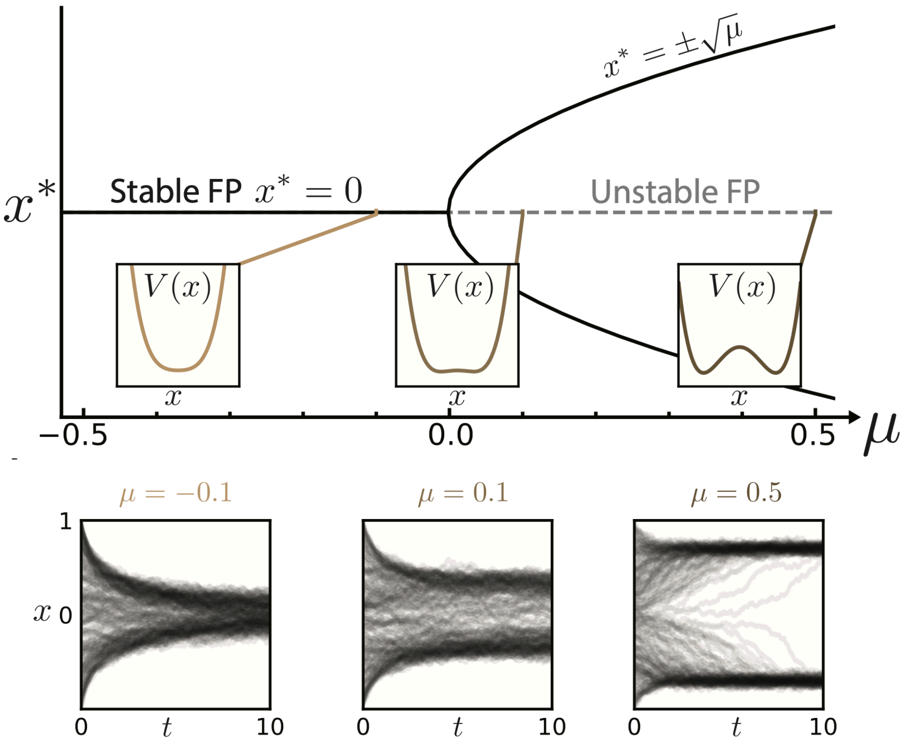
Brownian particle in a potential



$$\dot{x}_t = -\partial_x V(x_t) + \sigma \eta_t.$$

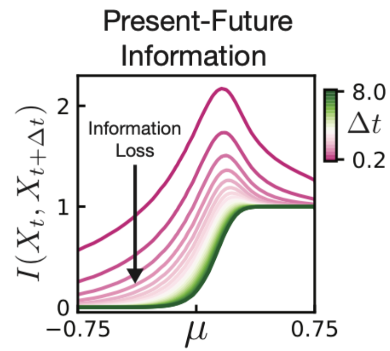
$$V(x) = \frac{1}{4}(\mu - x^2)^2$$

Brownian particle in a potential

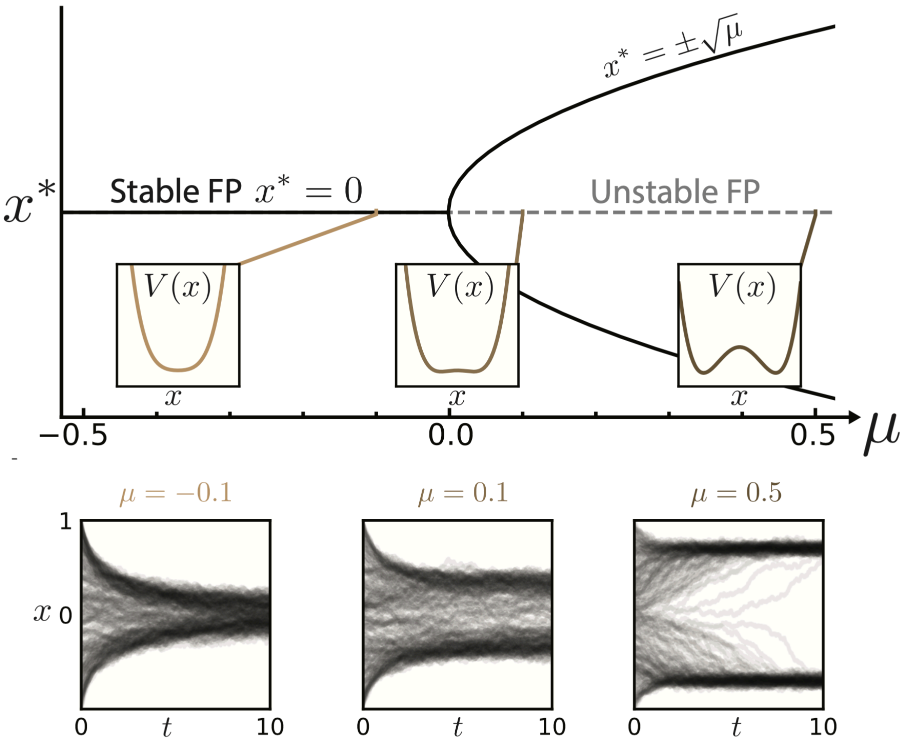


$$\dot{x}_t = -\partial_x V(x_t) + \sigma \eta_t.$$

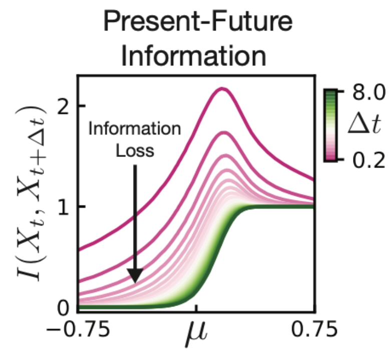
$$V(x) = \frac{1}{4}(\mu - x^2)^2$$

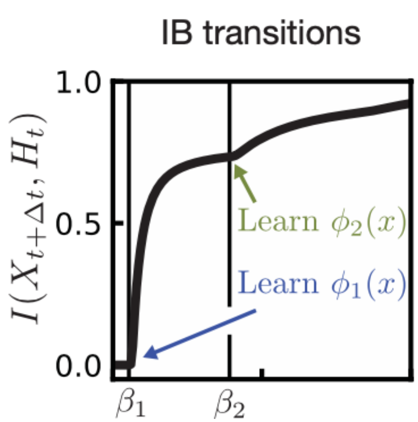


Brownian particle in a potential



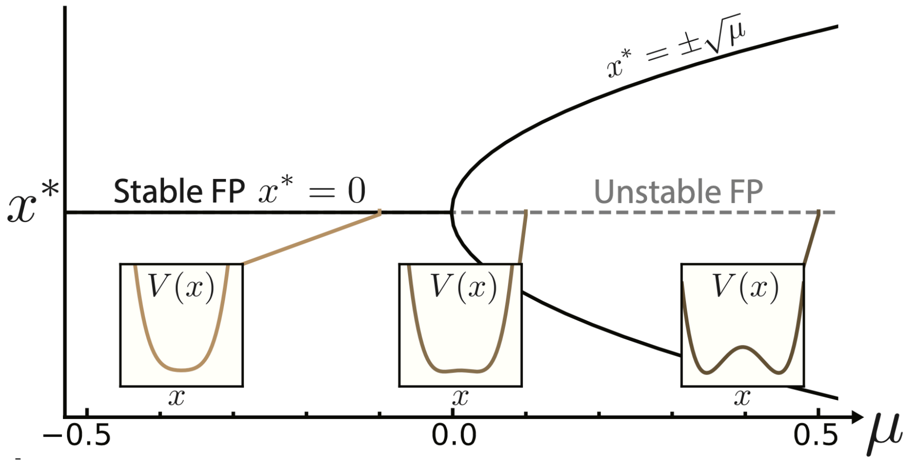
$$\dot{x}_t = -\partial_x V(x_t) + \sigma\eta_t.$$
$$V(x) = \frac{1}{4}(\mu - x^2)^2$$

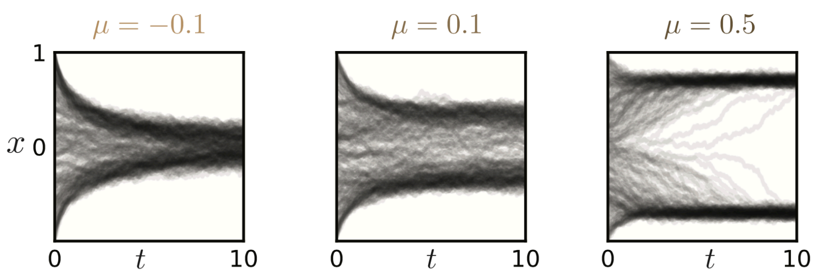




- IB learns the transfer operator eigenmodes

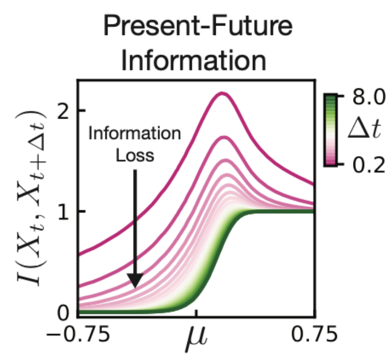
Brownian particle in a potential

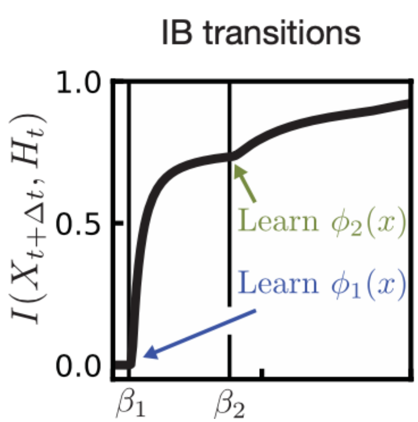




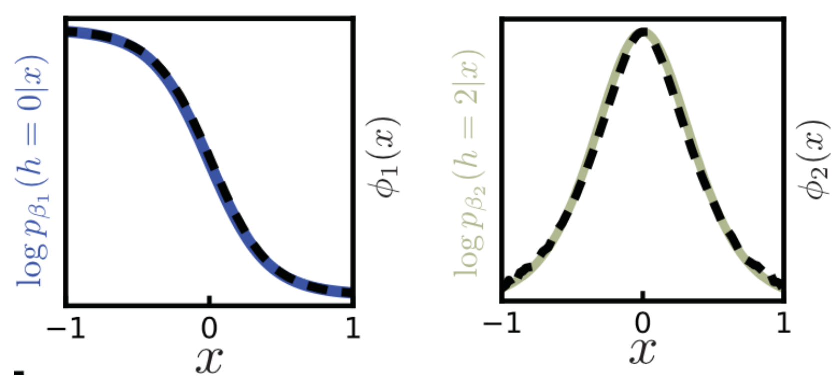
$$\dot{x}_t = -\partial_x V(x_t) + \sigma\eta_t.$$

$$V(x) = \frac{1}{4}(\mu - x^2)^2$$

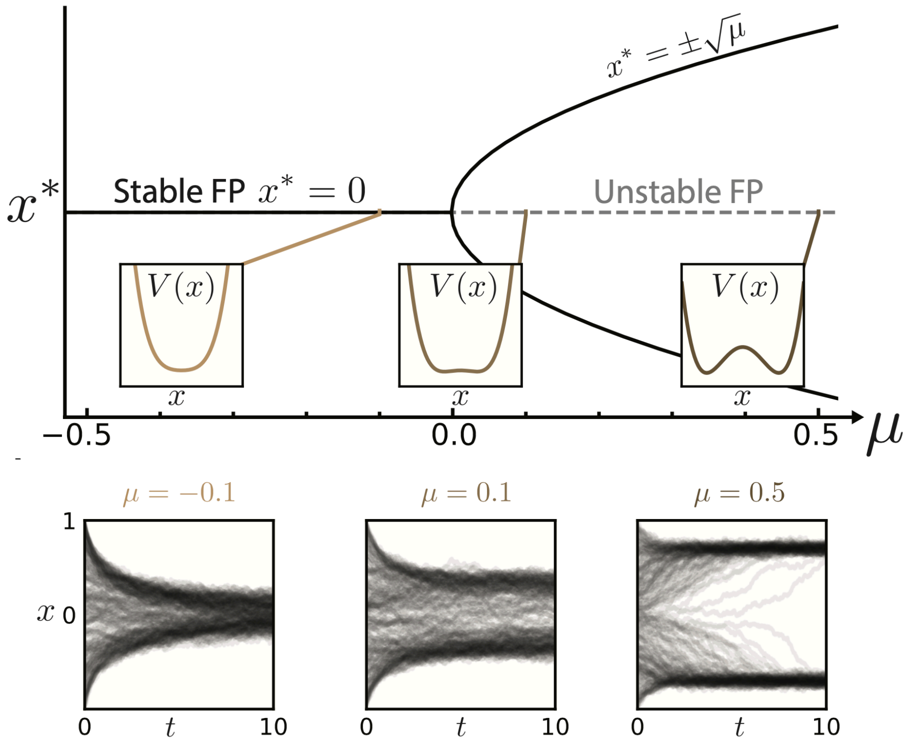




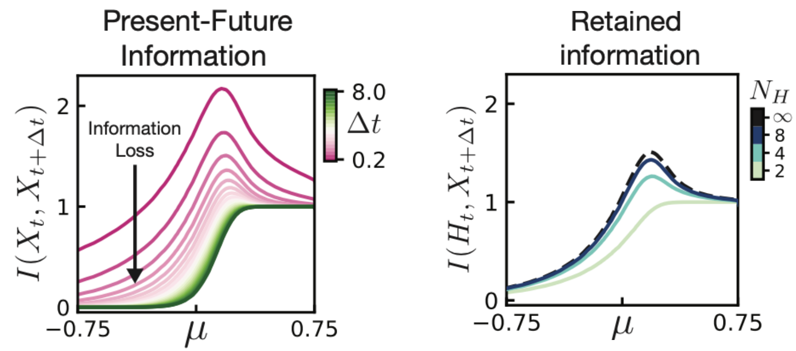
- IB learns the transfer operator eigenmodes

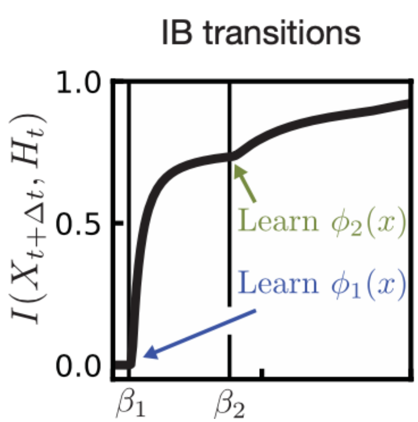


Brownian particle in a potential

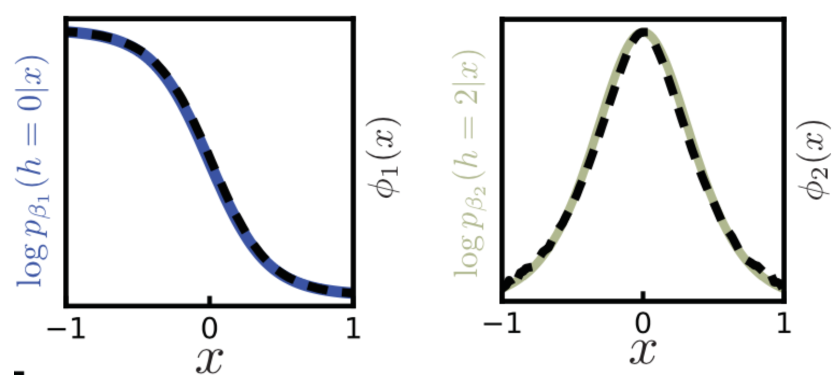


$$\dot{x}_t = -\partial_x V(x_t) + \sigma \eta_t.$$
$$V(x) = \frac{1}{4}(\mu - x^2)^2$$

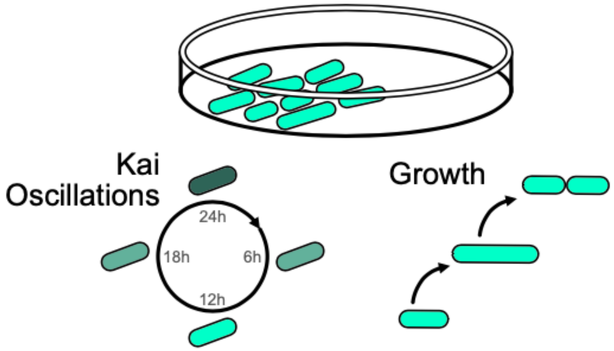




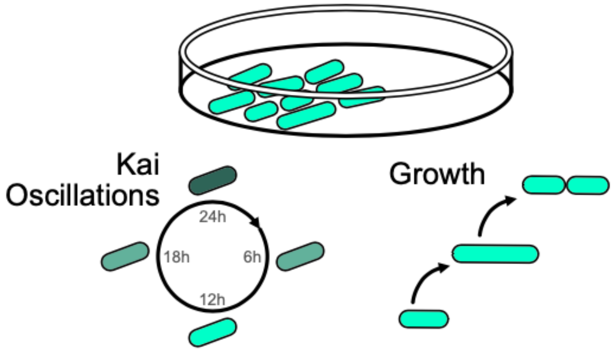
- IB learns the transfer operator eigenmodes

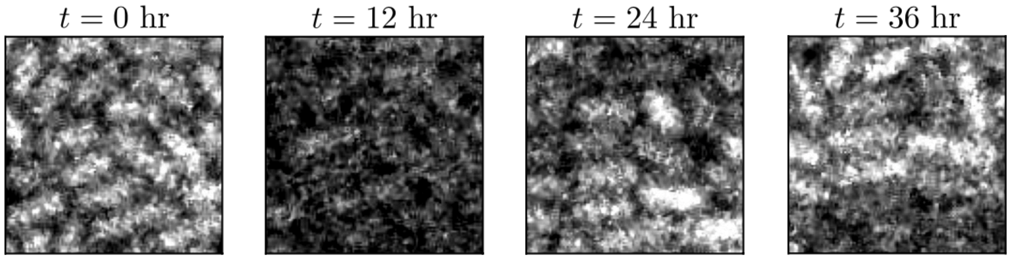


Cyanobacteria experiments

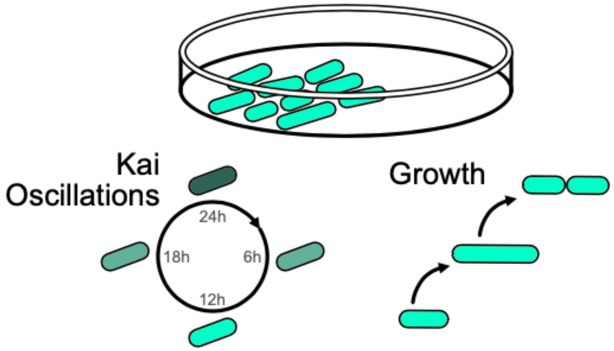


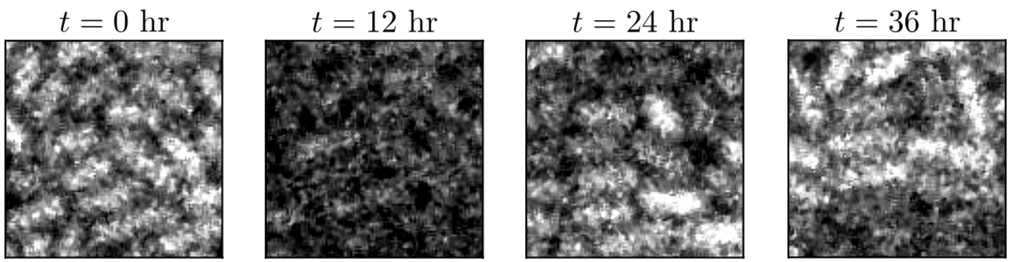
Cyanobacteria experiments

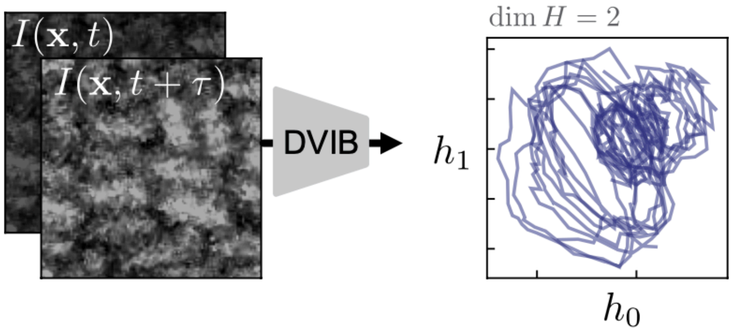




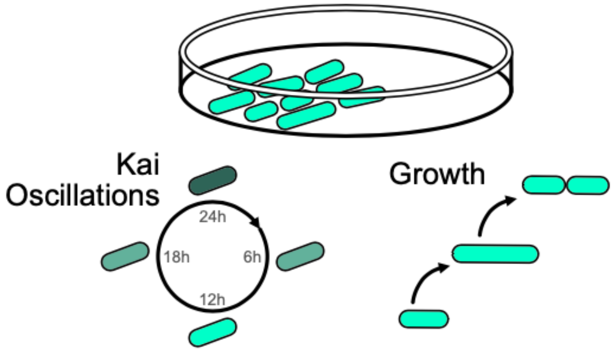
Cyanobacteria experiments

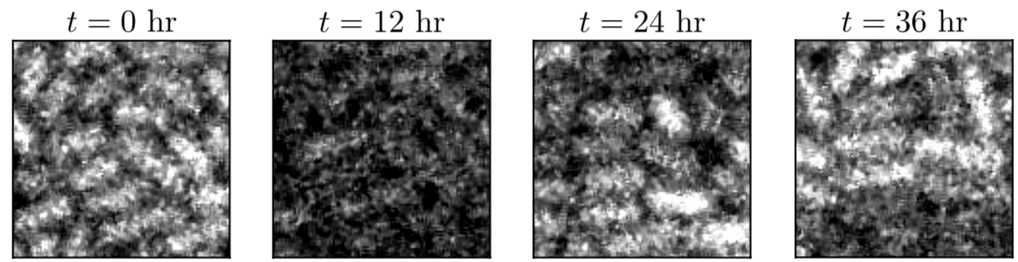




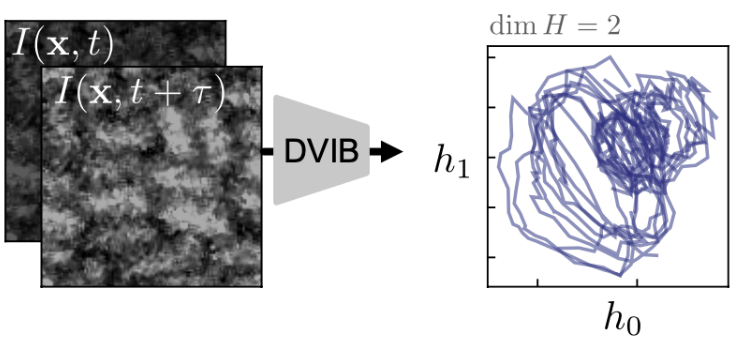


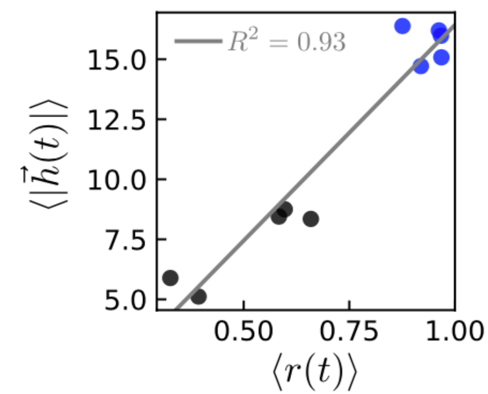
Cyanobacteria experiments



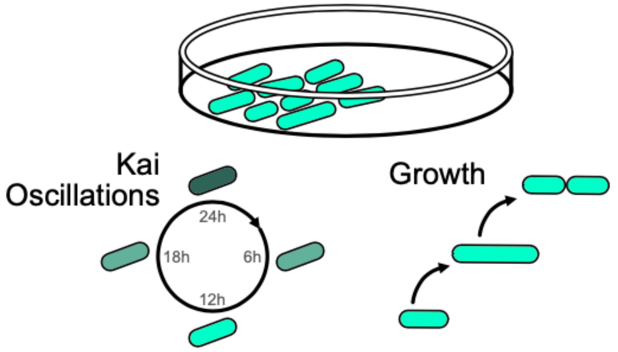


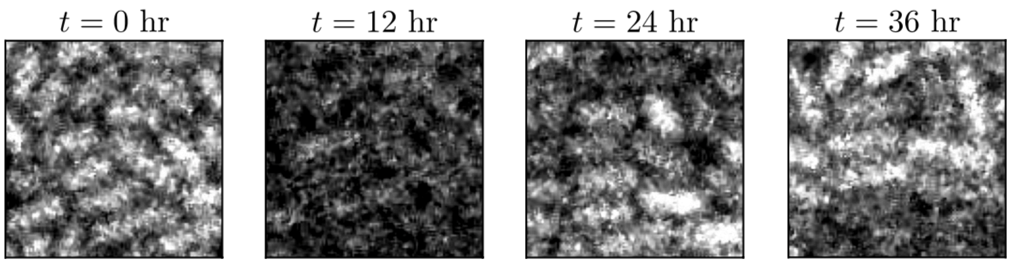
- Dynamics in latent space reveals populations differing by **synchronisation**



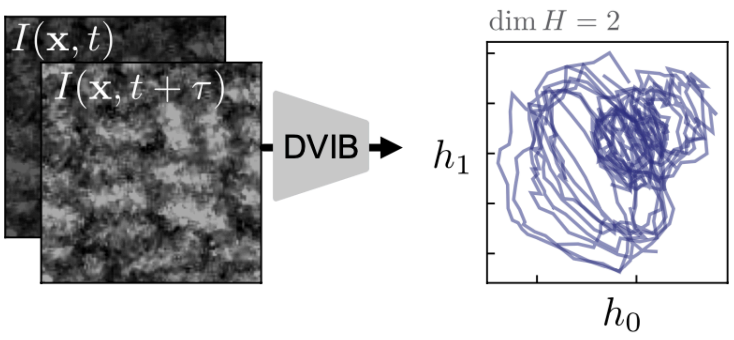


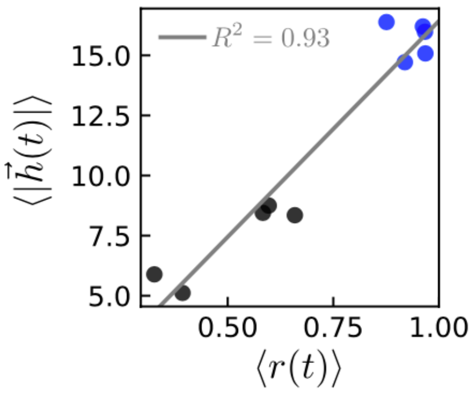
Cyanobacteria experiments

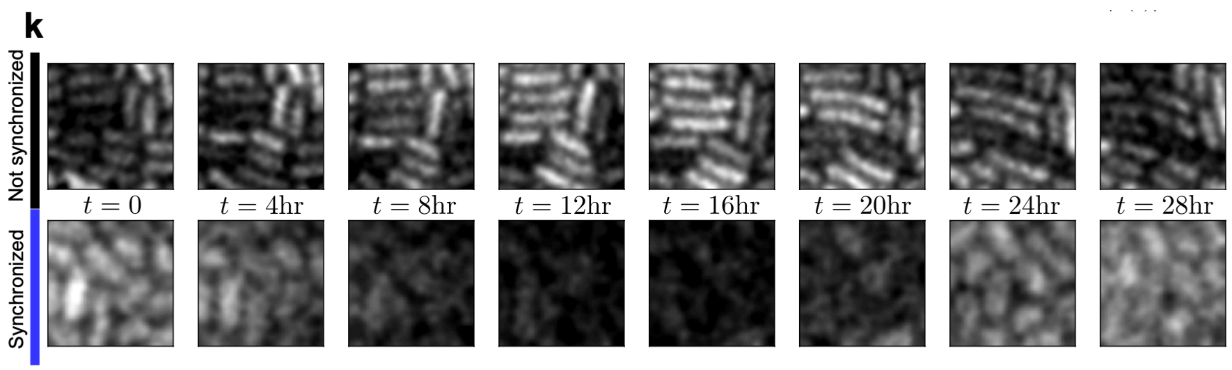




- Dynamics in latent space reveals populations differing by **synchronisation**







Outlook

- Applications to 3D stat.mech. models
- Automating discovering the algebraic properties satisfied by the operators
- Dynamical graphs
- Application to experimental data: soft-matter,

Operator content of theories on different lattices, equilibrium or not, can be extracted using compression theory tools from raw data alone.

$$\Lambda_{\mathrm{S1}} \cdot \mathcal{V}(\mathbf{r}) = (-1)^{x+y} N_{|}(\mathbf{r}),$$
$$\Lambda_{\mathrm{S2}} \cdot \mathcal{V}(\mathbf{r}) = (-1)^{x+y+1} N_{-}(\mathbf{r})$$

$$N_{|}(\mathbf{r}) = \frac{1}{4} + \frac{(-1)^{x+y+1}}{2\pi}\partial_x\varphi(\mathbf{r}) + (-1)^y \sin\varphi(\mathbf{r})$$
$$N_{-}(\mathbf{r}) = \frac{1}{4} + \frac{(-1)^{x+y}}{2\pi}\partial_y\varphi(\mathbf{r}) + (-1)^x \cos\varphi(\mathbf{r})$$

Papanikolaou et al. PRB 76, 134514 (2007)

$$\mathcal{H}_1 \sim \sum_{\mathbf{r}\in\mathcal{V}} \Lambda_{\mathrm{S1}} \cdot \mathcal{V}(\mathbf{r}) = \sum_{\mathbf{r}\in\mathcal{V}} \left[\frac{(-1)^{x+y}}{4} + \frac{\partial_x\varphi(\mathbf{r})}{2\pi} + (-1)^x \sin\varphi(\mathbf{r})\right] = \sum_{\mathbf{r}\in\mathcal{V}} \left[\frac{\partial_x\varphi(\mathbf{r})}{2\pi} + (-1)^x \sin\varphi(\mathbf{r})\right]$$
$$\mathcal{H}_2 \sim \sum_{\mathbf{r}\in\mathcal{V}} \Lambda_{\mathrm{S2}} \cdot \mathcal{V}(\mathbf{r}) = \sum_{\mathbf{r}\in\mathcal{V}} \left[\frac{(-1)^{x+y}}{4} + \frac{\partial_y\varphi(\mathbf{r})}{2\pi} + (-1)^y \cos\varphi(\mathbf{r})\right] = \sum_{\mathbf{r}\in\mathcal{V}} \left[\frac{\partial_y\varphi(\mathbf{r})}{2\pi} + (-1)^y \cos\varphi(\mathbf{r})\right]$$

$$\mathcal{H} \propto \tau \circ \nabla \langle\varphi(\mathbf{r})\rangle_{\mathbf{r}\in\mathcal{V}}$$

- The staggered filters:

$$\Lambda_{\mathrm{S1}} \cdot \mathcal{V}(\mathbf{r}) = (-1)^{x+y} N_{|}(\mathbf{r}),$$
$$\Lambda_{\mathrm{S2}} \cdot \mathcal{V}(\mathbf{r}) = (-1)^{x+y+1} N_{-}(\mathbf{r})$$

$$N_{|}(\mathbf{r}) = \frac{1}{4} + \frac{(-1)^{x+y+1}}{2\pi} \partial_x \varphi(\mathbf{r}) + (-1)^y \sin \varphi(\mathbf{r})$$
$$N_{-}(\mathbf{r}) = \frac{1}{4} + \frac{(-1)^{x+y}}{2\pi} \partial_y \varphi(\mathbf{r}) + (-1)^x \cos \varphi(\mathbf{r})$$

Papanikolaou et al. PRB 76, 134514 (2007)

$$\mathcal{H}_1 \sim \sum_{\mathbf{r}\in\mathcal{V}} \Lambda_{\mathrm{S1}} \cdot \mathcal{V}(\mathbf{r}) = \sum_{\mathbf{r}\in\mathcal{V}} \left[\frac{(-1)^{x+y}}{4} + \frac{\partial_x \varphi(\mathbf{r})}{2\pi} + (-1)^x \sin \varphi(\mathbf{r}) \right] = \sum_{\mathbf{r}\in\mathcal{V}} \left[\frac{\partial_x \varphi(\mathbf{r})}{2\pi} + (-1)^x \sin \varphi(\mathbf{r}) \right]$$
$$\mathcal{H}_2 \sim \sum_{\mathbf{r}\in\mathcal{V}} \Lambda_{\mathrm{S2}} \cdot \mathcal{V}(\mathbf{r}) = \sum_{\mathbf{r}\in\mathcal{V}} \left[\frac{(-1)^{x+y}}{4} + \frac{\partial_y \varphi(\mathbf{r})}{2\pi} + (-1)^y \cos \varphi(\mathbf{r}) \right] = \sum_{\mathbf{r}\in\mathcal{V}} \left[\frac{\partial_y \varphi(\mathbf{r})}{2\pi} + (-1)^y \cos \varphi(\mathbf{r}) \right]$$

$$\mathcal{H} \propto \tau \circ \nabla \langle \varphi(\mathbf{r}) \rangle_{\mathbf{r}\in\mathcal{V}}$$

- The staggered filters:

$$\Lambda_{S1} \cdot \mathcal{V}(\mathbf{r}) = (-1)^{x+y} N_{|}(\mathbf{r}),$$
$$\Lambda_{S2} \cdot \mathcal{V}(\mathbf{r}) = (-1)^{x+y+1} N_{-}(\mathbf{r})$$

$$N_{|}(\mathbf{r}) = \frac{1}{4} + \frac{(-1)^{x+y+1}}{2\pi}\partial_x \varphi(\mathbf{r}) + (-1)^y \sin\varphi(\mathbf{r})$$
$$N_{-}(\mathbf{r}) = \frac{1}{4} + \frac{(-1)^{x+y}}{2\pi}\partial_y \varphi(\mathbf{r}) + (-1)^x \cos\varphi(\mathbf{r})$$

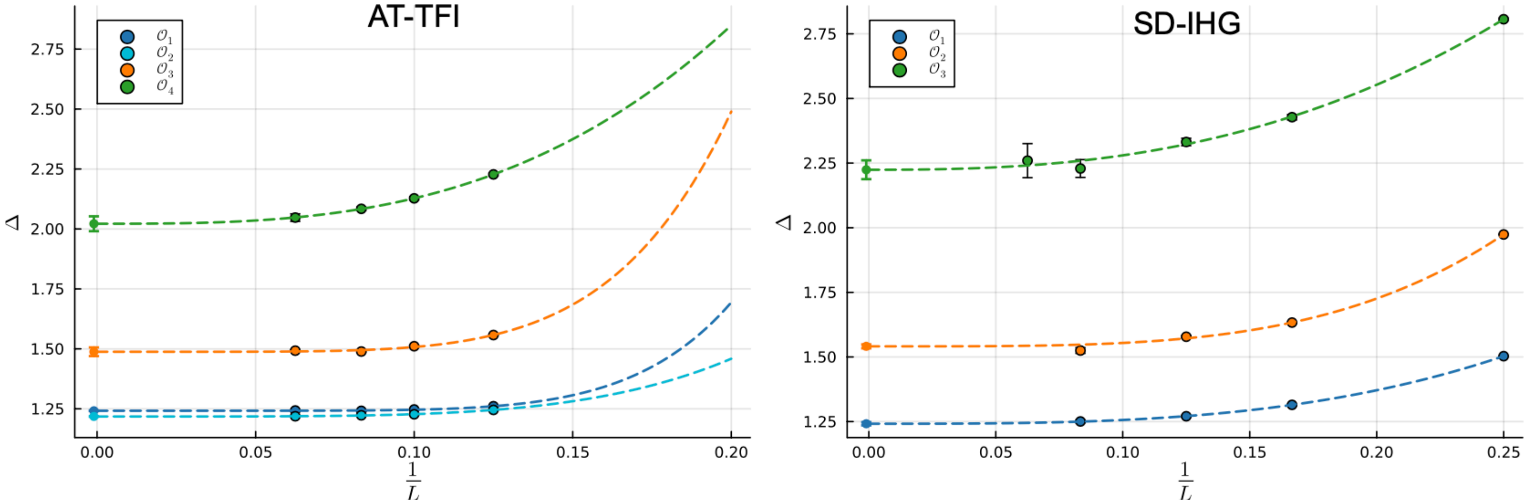
Papanikolaou et al. **PRB 76**, 134514 (2007)

$$\mathcal{H}_1 \sim \sum_{\mathbf{r}\in\mathcal{V}} \Lambda_{S1} \cdot \mathcal{V}(\mathbf{r}) = \sum_{\mathbf{r}\in\mathcal{V}} \left[\frac{(-1)^{x+y}}{4} + \frac{\partial_x \varphi(\mathbf{r})}{2\pi} + (-1)^x \sin\varphi(\mathbf{r})\right] = \sum_{\mathbf{r}\in\mathcal{V}} \left[\frac{\partial_x \varphi(\mathbf{r})}{2\pi} + (-1)^x \sin\varphi(\mathbf{r})\right]$$
$$\mathcal{H}_2 \sim \sum_{\mathbf{r}\in\mathcal{V}} \Lambda_{S2} \cdot \mathcal{V}(\mathbf{r}) = \sum_{\mathbf{r}\in\mathcal{V}} \left[\frac{(-1)^{x+y}}{4} + \frac{\partial_y \varphi(\mathbf{r})}{2\pi} + (-1)^y \cos\varphi(\mathbf{r})\right] = \sum_{\mathbf{r}\in\mathcal{V}} \left[\frac{\partial_y \varphi(\mathbf{r})}{2\pi} + (-1)^y \cos\varphi(\mathbf{r})\right]$$

- Expanding sin/cos and averaging we obtain:

$$\mathcal{H} \propto \tau \circ \nabla \langle \varphi(\mathbf{r}) \rangle_{\mathbf{r}\in\mathcal{V}}$$

Critical exponents for the Self-dual Ising-Higgs Gauge theory and the Ashkin-Teller model



Filters as scaling operators

https://github.com/RSMI-NE/RSMI-NE

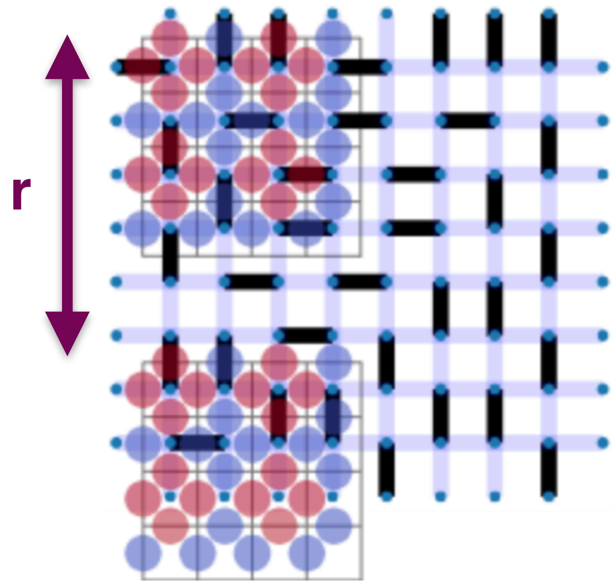
Phys.Rev.Lett 127, 240603 (2021)

Filters as scaling operators

- They can used as operators in correlation functions

Filters as scaling operators

- They can used as operators in correlation functions

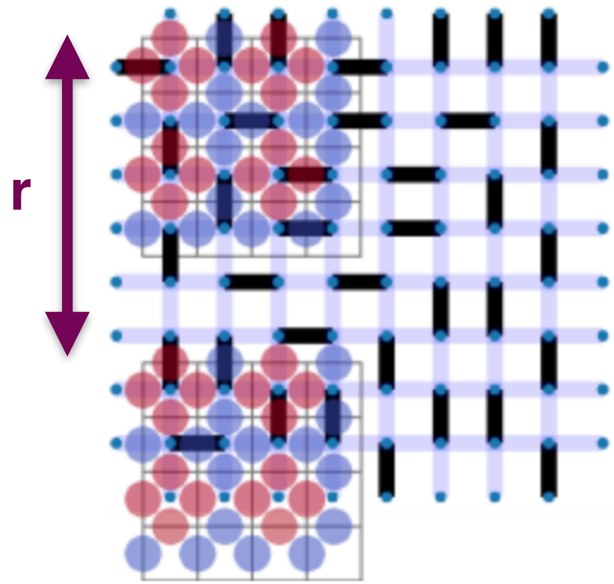


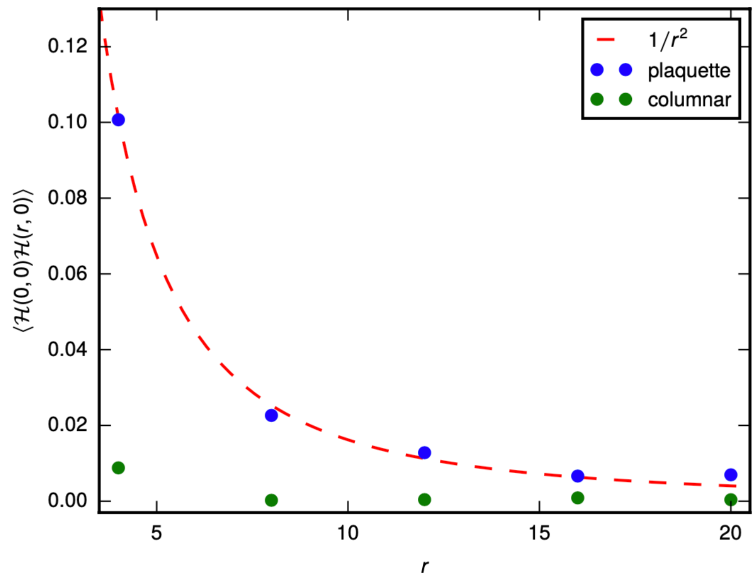
https://github.com/RSMI-NE/RSMI-NE

Phys.Rev.Lett 127, 240603 (2021)

Filters as scaling operators

- They can used as operators in correlation functions





https://github.com/RSMI-NE/RSMI-NE

Phys.Rev.Lett 127, 240603 (2021)