Title: Machine Learning Physics: From Quantum Mechanics to Holographic Geometry

Speakers: Yi-Zhuang You

Collection: Machine Learning for Quantum Design

Date: July 11, 2019 - 11:30 AM

URL: http://pirsa.org/19070014

Abstract: Inspired by the "third wave" of artificial intelligence (AI), machine learning has found rapid applications in various topics of physics research. Perhaps one of the most ambitious goals of machine learning physics is to develop novel approaches that ultimately allows AI to discover new concepts and governing equations of physics from experimental observations. In this talk, I will present our progress in applying machine learning technique to reveal the quantum wave function of Bose-Einstein condensate (BEC) and the holographic geometry of conformal field theories. In the first part, we apply machine translation to learn the mapping between potential and density profiles of BEC and show how the concept of quantum wave function can emerge in the latent space of the translator and how the Schrodinger equation is formulated as a recurrent neural network. In the second part, we design a generative model to learn the field theory configuration of the XY model and show how the machine can identify the holographic bulk degrees of freedom and use them to probe the emergent holographic geometry.

Machine Learning Physics
From Quantum Mechanics to Holographic Duality

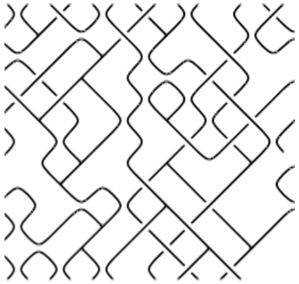
Yi-Zhuang You
University of California, San Diego

Machine Learning for Quantum Design
Perimeter Institute, July 2019

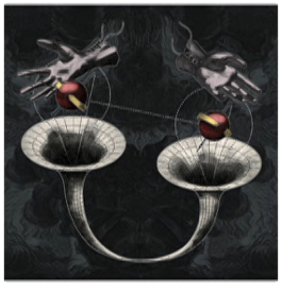
Machine Learning Physics

- Emergent phenomenon — a central theme of condensed matter physics.

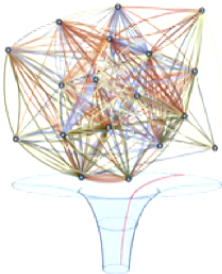
Weyl semimetal
(emergent particle)



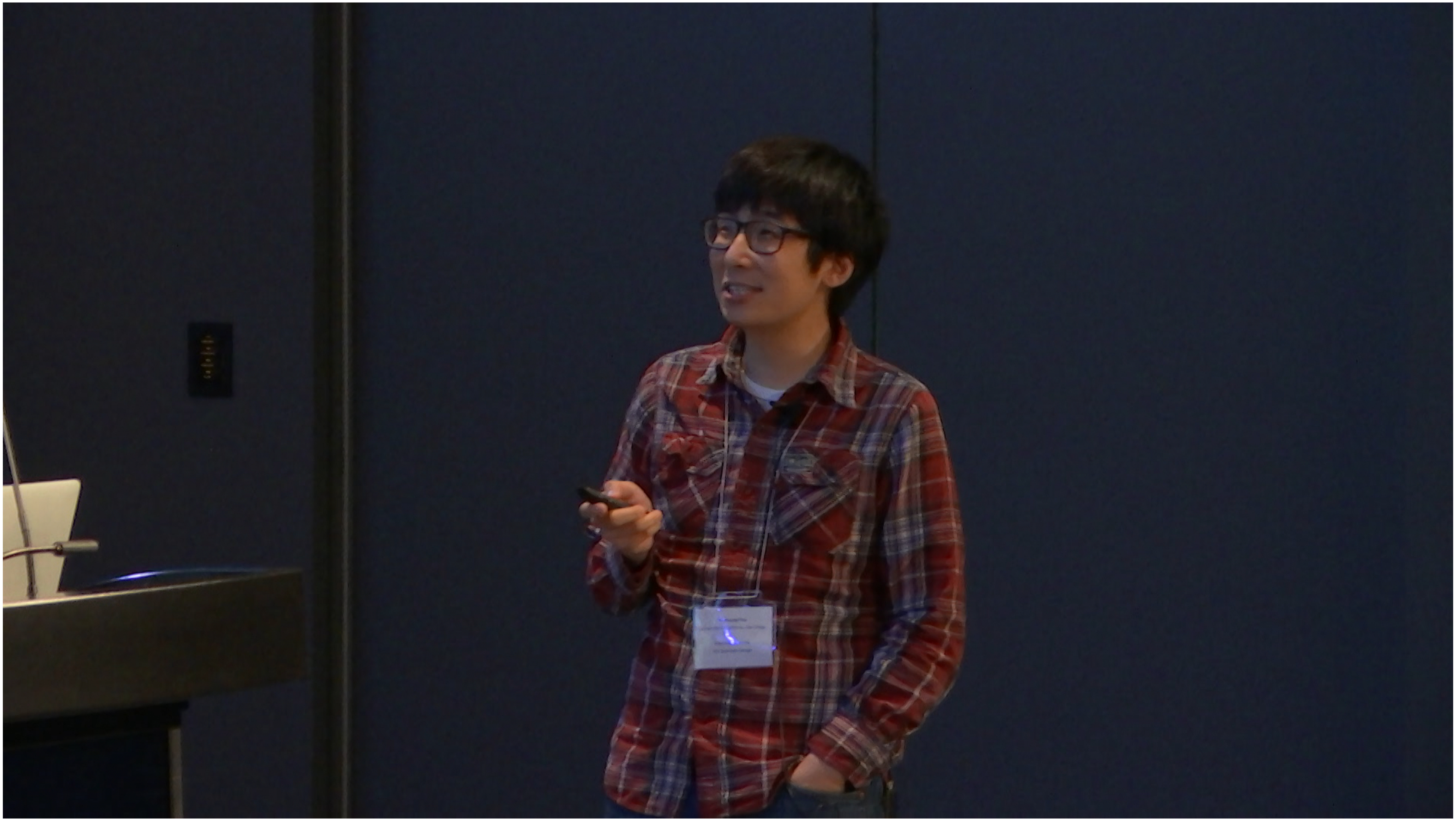
String net condensation
(emergent force)



ER = EPR
(emergent spacetime)

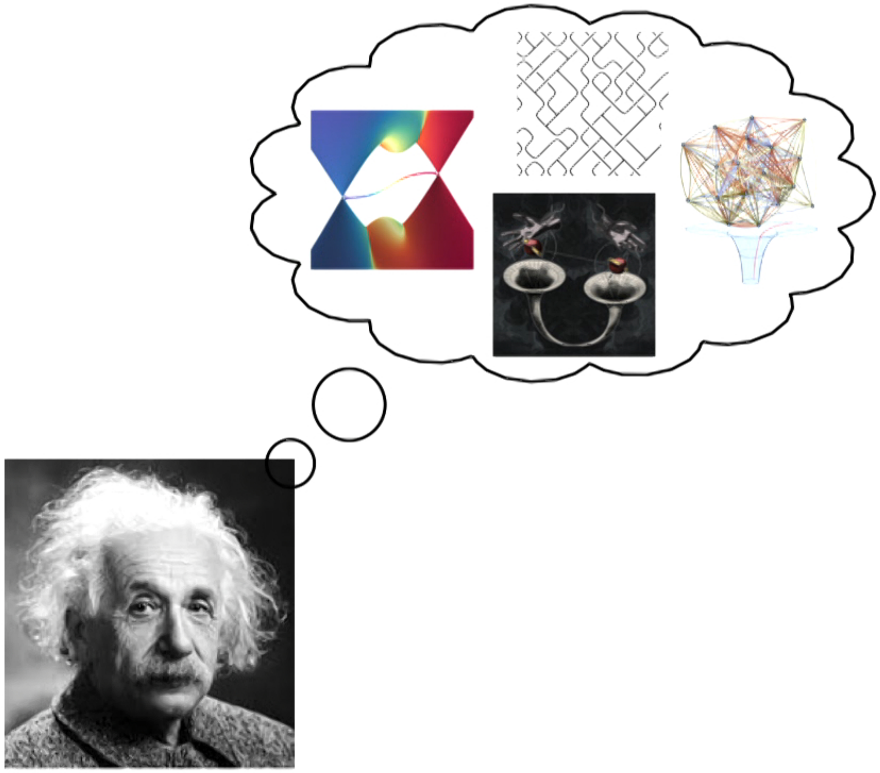


SYK model
(emergent gravity)



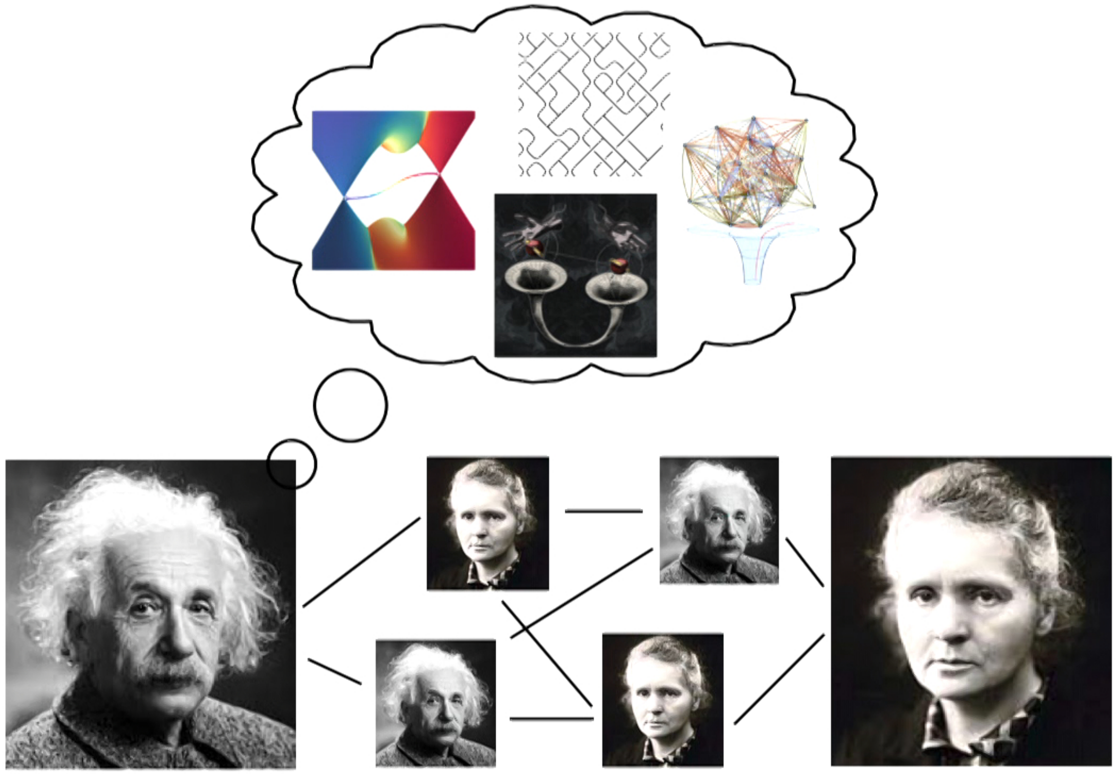
Machine Learning Physics

- Aren't all these physics theories themselves also emergent phenomena?



Machine Learning Physics

- Aren't all these physics theories themselves also emergent phenomena?

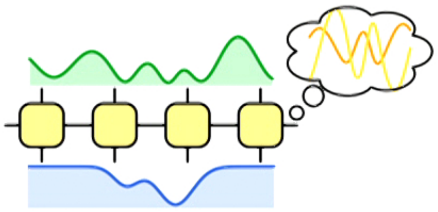


Machine Learning Physics

- Goal: investigate whether artificial neural networks can be used to discover physical concepts and laws from observation data.

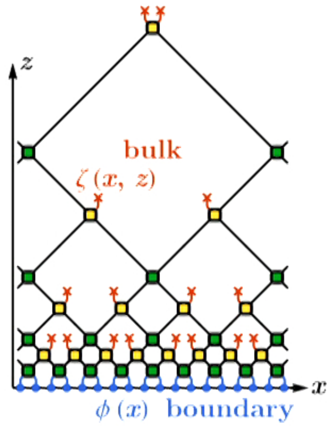
- Examples
  - Machine learning quantum mechanics — recurrent autoencoder

    C Wang, H Zhai, Y-Z You. arXiv: 1901.11103



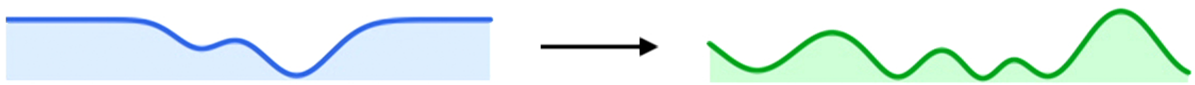
  - Machine learning renormalization group and holographic duality — flow-based deep generative model

    H Hu, S-H Li, L Wang, Y-Z You. arXiv: 1903.00804



Potential and Density Data

- Suppose quantum mechanics has not been formulated so far
- Yet, amazingly, we know how to perform cold atom experiments of Bose-Einstein condensate (BEC)



Potential profile
(optical speckles,
optical tweezers …)

BEC Density profile
(in-situ measurement)

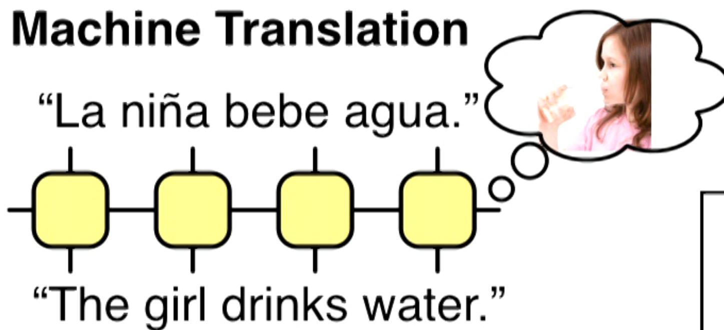
Billy et. al., Nature (2008),
Henderson et.al., NJP (2009) …

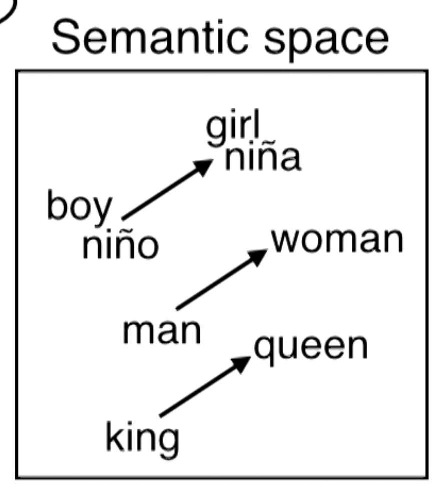
Issac Tamblyn's talk

- Questions
  - Can quantum mechanics (QM) be discovered as the most natural theory to explain the experiment?
  - Will the machine develop alternative form of QM?

Inspiration from Machine Translation

- Motivation: developments in machine translation
  - Sequence-to-sequence mapping (RNN, LSTM …)





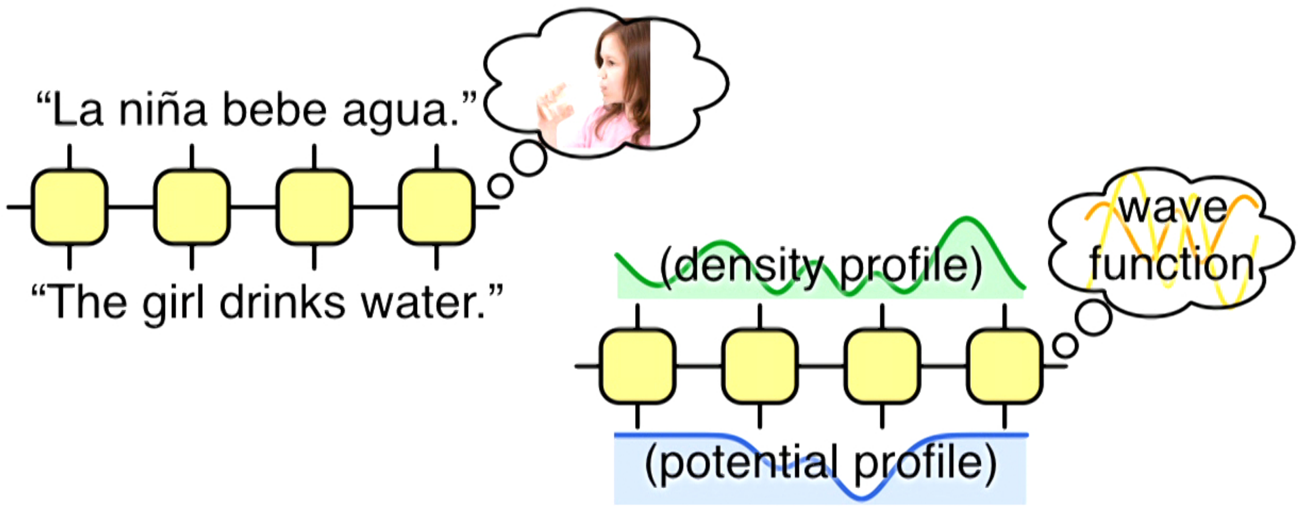
  - Representation learning (word2vec …)

king – man + woman ≈ queen

T Mikolov, SW Yih, G Zweig. NAACL-HLT-2013
also Paul's talk

Inspiration from Machine Translation

- Motivation: developments in machine translation
  - Train the neural network model to perform a task
  - Discover concepts and relations in representation space

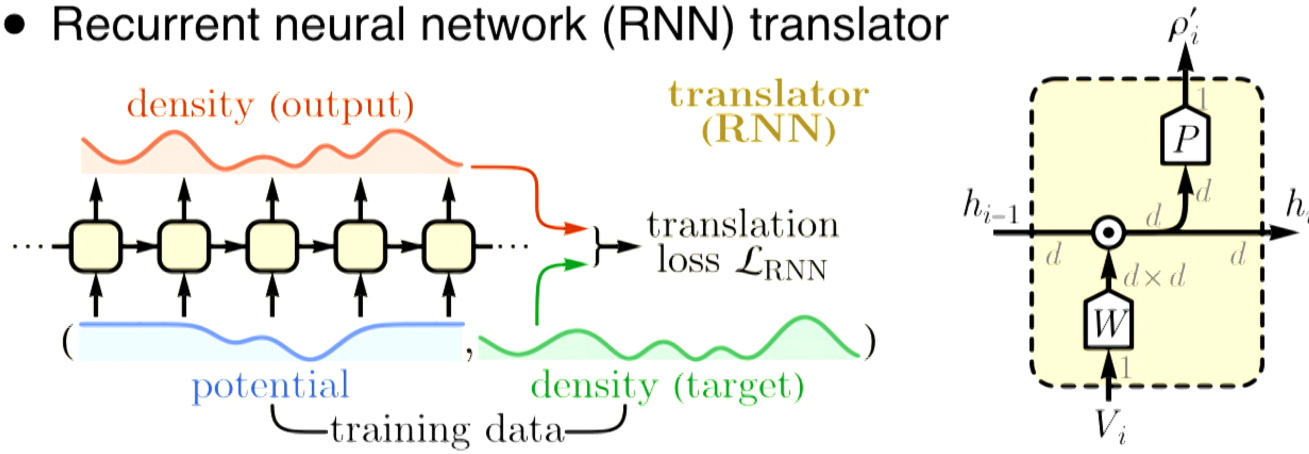


Similar setup but different task: S Pilati, P Pieri, Scientific Reports (2019)

- **Task**: potential-to-density mapping
- **Latent variables**: wave function?

Potential-to-Density Translator

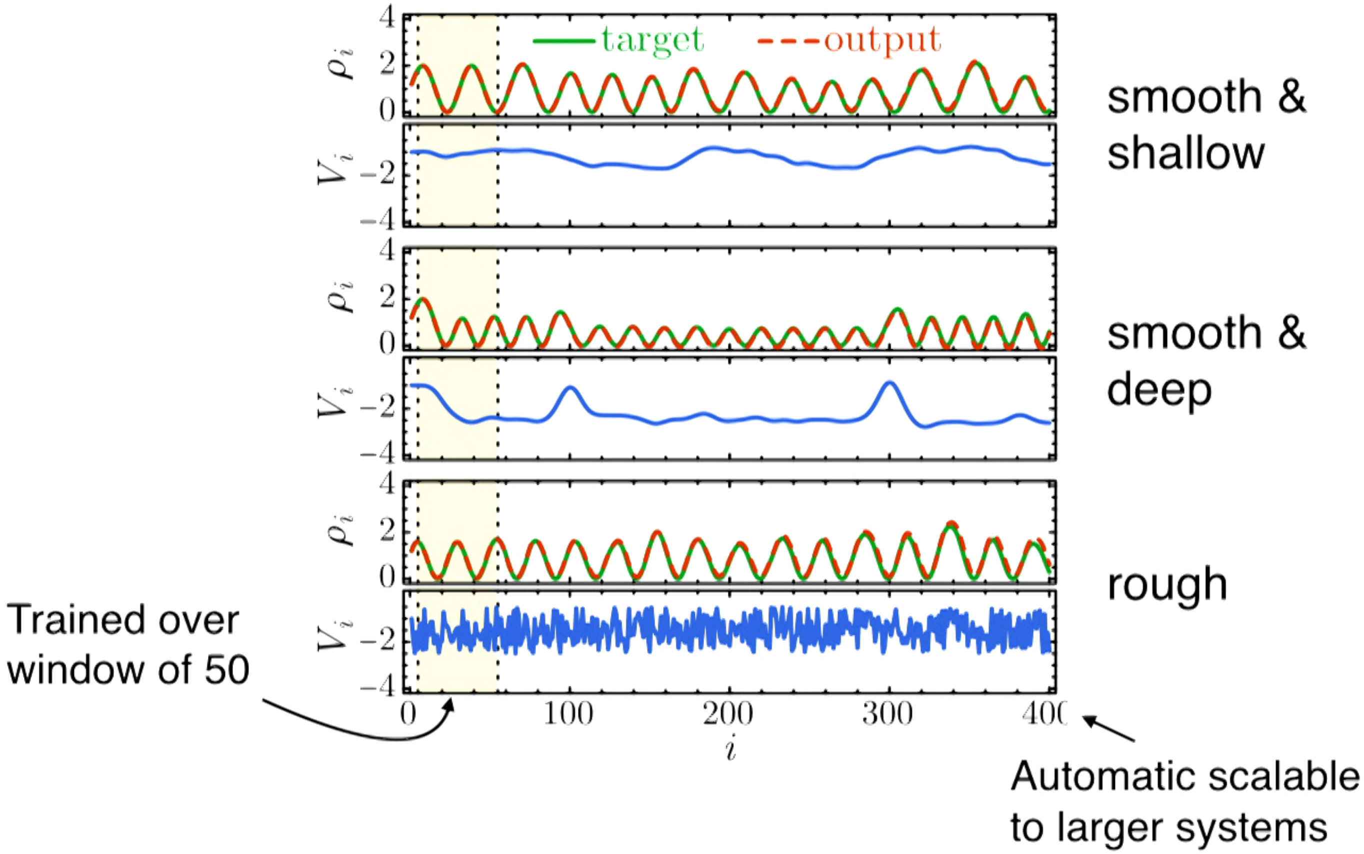
- Recurrent neural network (RNN) translator



- Discretize the 1D space, collect training data by simulation
- Input: potential sequence V_i
- Update: hidden state $h_i = W(V_i) \cdot h_{i-1}$
- Output: density sequence $\rho'_i = P(h_i)$
- Minimize translation loss $\mathcal{L}_{\text{RNN}} = \sum_{i \in \text{window}} (\rho'_i - \rho_i)^2$

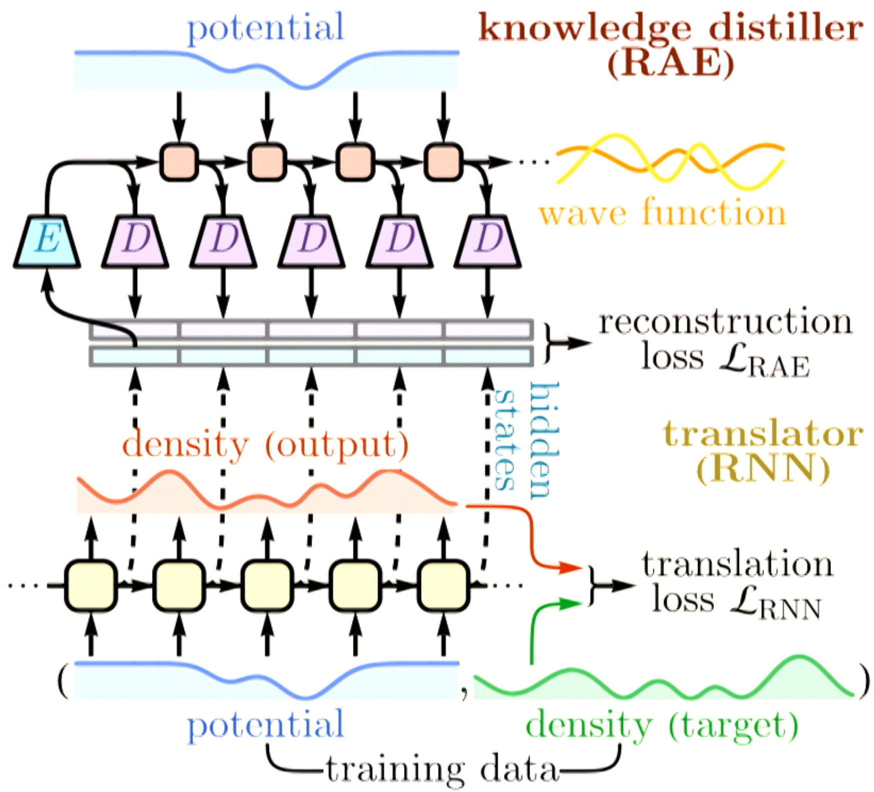
Performance of the Translator

- Performance of the RNN translator



Introspective Learning

- Introspective Learning

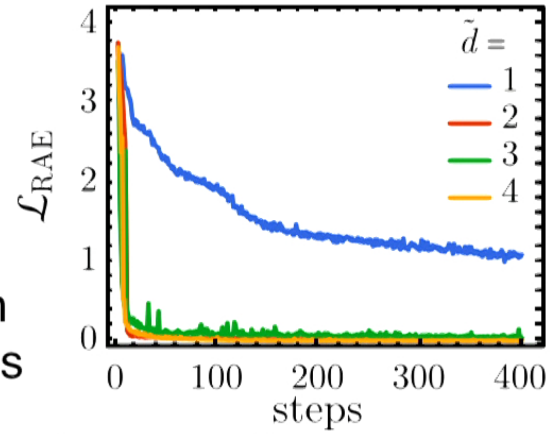


High-level machine only interface with the neural activation of the low-level machine

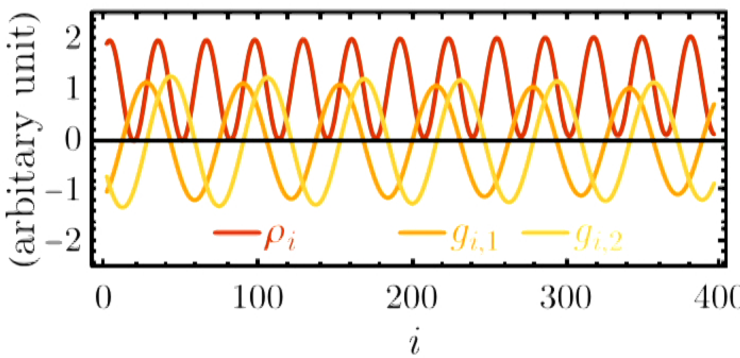
Low-level machine deal with training / experimental data

Emergent Quantum Mechanics

- Imposing information bottleneck
  - Squeezing the latent space dim
  - Monitor the reconstruction loss of the knowledge distiller
  - Abrupt increase of loss only when latent dim < 2 ⇒ two real variables



- Quantum wave function and its 1st order derivative



Update rules

$$\begin{bmatrix} g_{i+1,1} \\ g_{i+1,2} \end{bmatrix} = \begin{bmatrix} 1 & a \\ aV_i & 1 \end{bmatrix} \begin{bmatrix} g_{i,1} \\ g_{i,2} \end{bmatrix}$$

matching Schrödinger Eq.

$$\partial_x^2 \psi(x) = V(x)\psi(x)$$

Alternative Forms of Quantum Mechanics

- If we relax the information bottle neck → alternative forms of quantum machines can also emerge, e.g.

$$\partial_x \begin{bmatrix} \rho(x) \\ \rho'(x) \\ \rho''(x) \end{bmatrix} = \begin{bmatrix} 0 & 2 & 0 \\ V(x) & 0 & 1 \\ 0 & 2V(x) & 0 \end{bmatrix} \begin{bmatrix} \rho(x) \\ \rho'(x) \\ \rho''(x) \end{bmatrix}$$

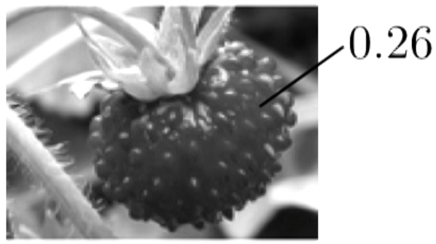
  - Hidden variables: density $\rho(x) = |\psi(x)|^2$ and derivatives
  - But requires at least three real variables

- Wave function + Schrödinger equation formulation of QM is indeed the most parsimonious theory that have emerged in our neural network.

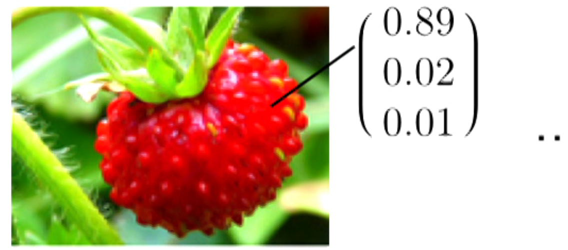
Machine Learning Renormalization Group

Quantum Field Theory as Image Dataset

- A field: a mapping from spacetime to some target manifold

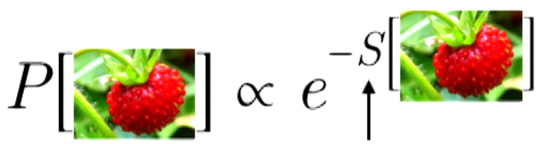


Scalar fields



Vector fields

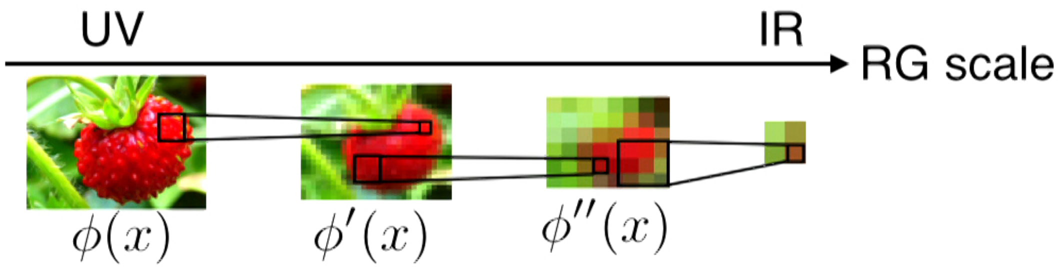
- A quantum field theory (QFT): a model that assigns an action (= negative log likelihood) to every field configuration.



- Can we build a generative model to represent a QFT?

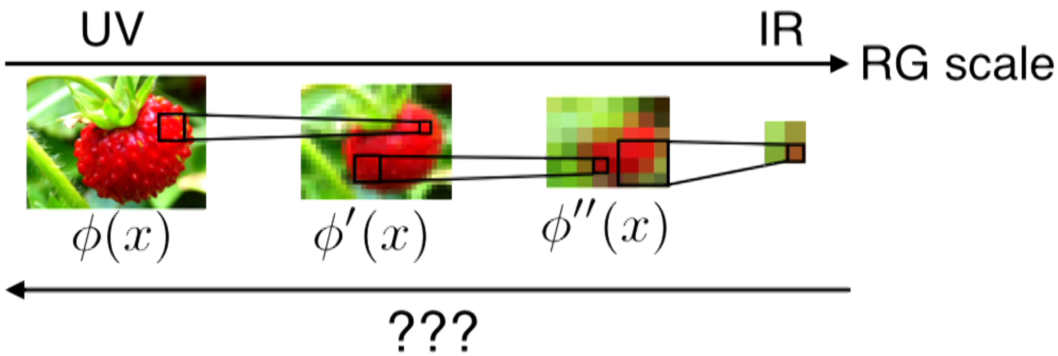
Renormalization Group as Generative Model

- Renormalization "group" (RG): progressively coarse-graining the field (like a convolutional neural network)



Renormalization Group as Generative Model

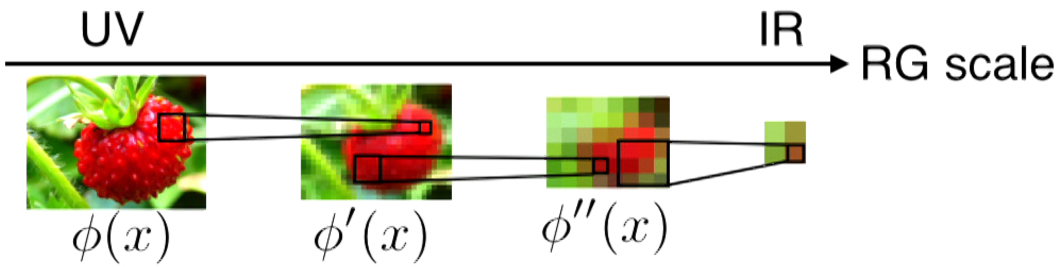
- Renormalization "group" (RG): progressively coarse-graining the field (like a convolutional neural network)



Traditional RG is not invertible…

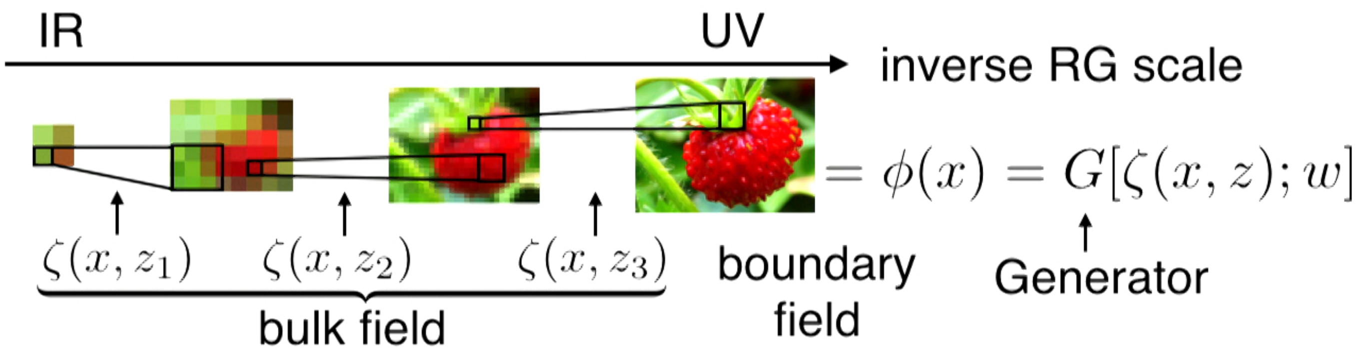
Renormalization Group as Generative Model

- Renormalization "group" (RG): progressively coarse-graining the field, like a convolutional neural network



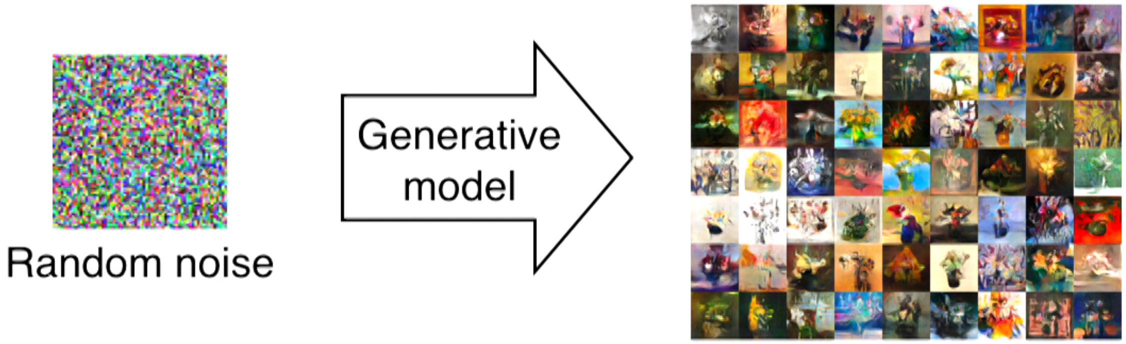
- Inverse RG: a hierarchical generative model

Cédric Bény, NJP (2013)



Generative Models

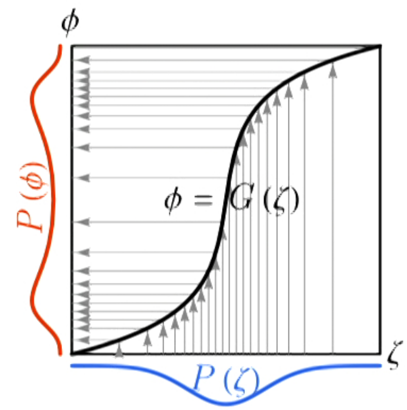
- Flow-based generative model: generate images from noise (latent variables) by an invertible non-linear transformation



- Generative model deforms the probability distribution, sample ζ to generate ϕ

$$\phi = G(\zeta)$$

$$P(\phi) = P(\zeta)\left(\frac{\partial G(\zeta)}{\partial \zeta}\right)^{-1}$$



Generative Models

- What are the advantages of flow-based models compared to energy-based models (e.g. Boltzmann machines)?
  - Differentiable log likelihood allows gradient to propagate through probability to train the model.

$$\mathcal{L} = \mathsf{KL}(P_{\mathrm{dat}}||P_{\mathrm{mdl}}) \qquad P_{\mathrm{mdl}}(\phi) = P_{\mathrm{prior}}(\zeta)\left(\frac{\partial G(\zeta)}{\partial \zeta}\right)^{-1}$$

  - Direct sampling allows efficient sample generation

$$\phi = G(\zeta)$$

  - Bijectivity allows inference of latent encoding

$$\zeta = G^{-1}(\phi)$$

- Generative models with tractable likelihood
  - Flow-based: Zhang, E, Wang (2018)
  - Autoregressive: Wu, Wang, Zhang, PRL(2019), Sharir et.al. (2019)
  - Tensor networks: Han et.al. PRX(2018)

Neural Network Renormalization Group

- Generative model deforms noise to QFT

Li, Wang, PRL (2018)
Hu, Li, Wang, You (2019)

Posterior (boundary) ↓ Prior (bulk) ↓ Generator ↓

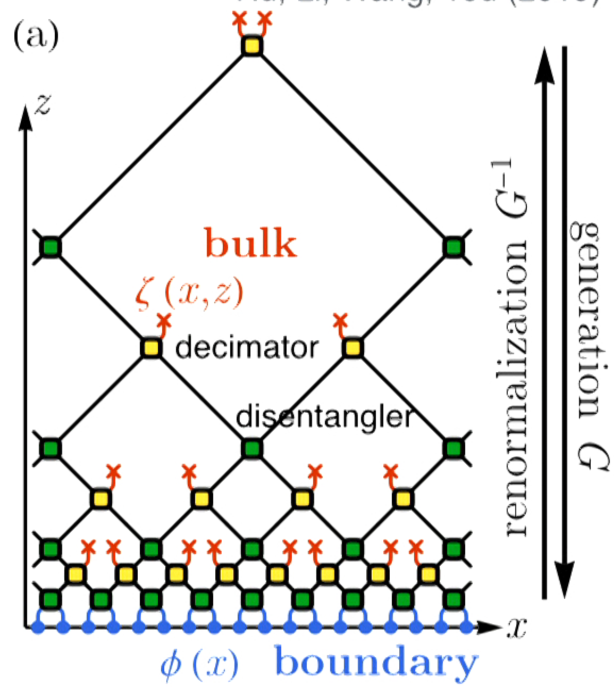
$$P_\partial[\phi] = P[\zeta] \det \left(\frac{\delta G[\zeta]}{\delta \zeta} \right)^{-1}$$

Model

↕ Minimize KL divergence

Target (QFT)

$$P[\phi] = e^{-S[\phi]}/Z$$

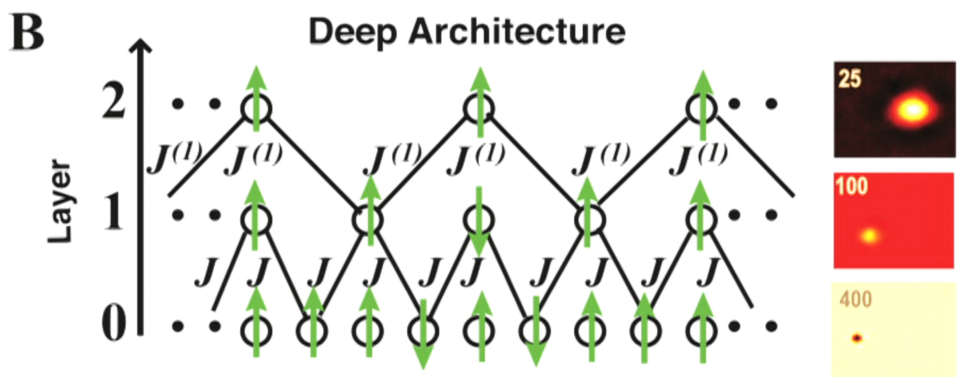


MERA network - Vidal (2006)

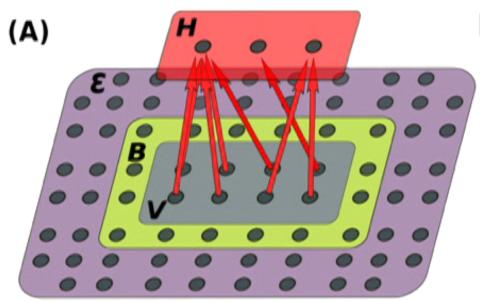
- How to choose the prior?
  Our choice: independent Gaussian $P[\zeta] \propto e^{-\zeta^2}$ (Why?)

Information Theoretic Goal of RG

- Renormalization Group = Deep Learning? Mehta, Schwab (2014)



- Maximal Real-Space Mutual Information (maxRMI) principle



Koch-Janusz, Ringel (2017)

$$\max I(H, \mathcal{E})$$

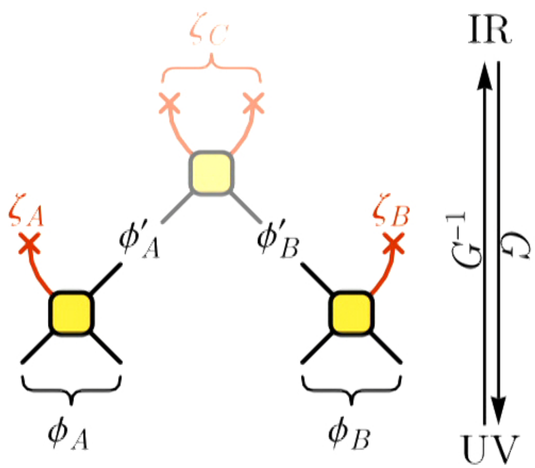
Coarse-grained freedom

Environment


IETER

Information Theoretic Goal of RG

- Minimal Bulk Mutual Information (minBMI) principle



- maxRMI: $\max I(\phi'_A : \phi_B)$
- minBMI: $\min I(\zeta_A : \zeta_B)$

Two objectives are related

$$I(\phi'_A : \phi_B) + I(\zeta_A : \zeta_B)$$
$$= I(\phi_A, \phi_B) = \text{const.}$$

Hu, Li, Wang, You (2109)

- The objectives are two-folded
  - Generate the QFT on the boundary

$$\min \mathrm{KL}(P_\partial[\phi] || e^{-S[\phi]})$$

  - Disentangle the QFT in the bulk $P[\zeta] \propto e^{-||\zeta||^2}$

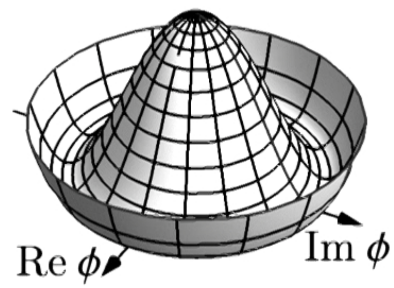
Complex ϕ^4 Model in 2D

- Lattice field theory on square lattice

$$S[\phi] = -t \sum_{\langle ij \rangle} \phi_i^* \phi_j + \sum_i (\mu |\phi_i|^2 + \lambda |\phi_i|^4)$$

  - Effectively 2D XY model $\phi_i = \sqrt{\rho} e^{\mathrm{i}\theta_i}$

$$S[\theta] = -\frac{1}{T} \sum_{\langle ij \rangle} \cos(\theta_i - \theta_j)$$

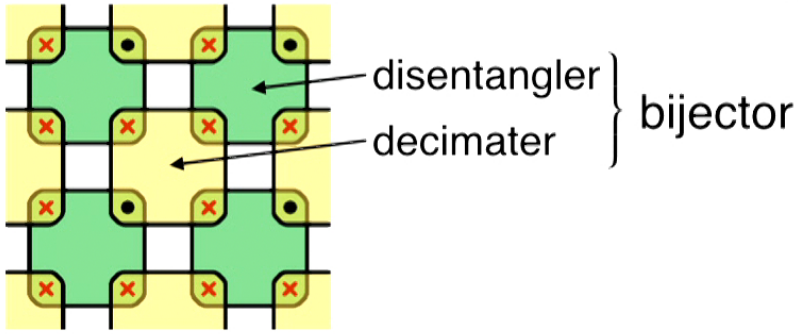


$\langle \phi_i^* \phi_j \rangle \sim r_{ij}^{\alpha}$ $\qquad$ $\langle \phi_i^* \phi_j \rangle \sim e^{-r_{ij}/\xi}$

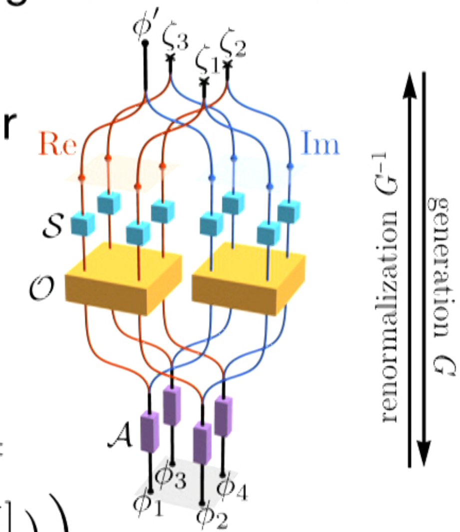
Luttinger liquid (CFT) — T_{KT} — Disordered → T (effectively tuned by t)

Training Scheme

- Architecture: flow-based hierarchical generative model



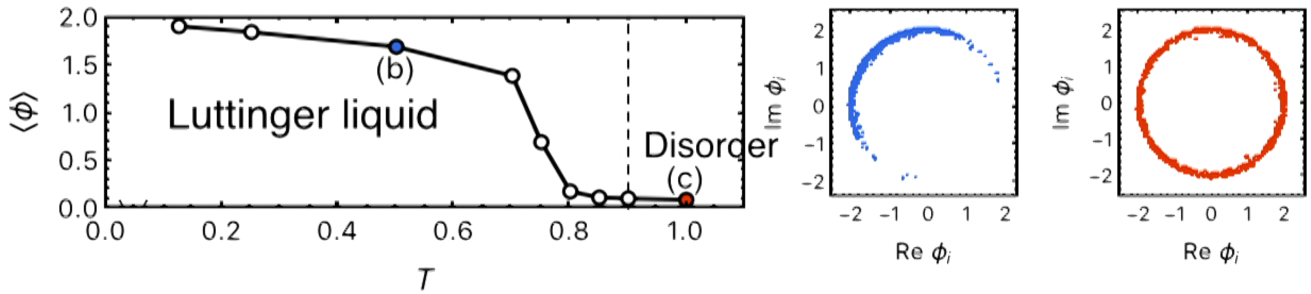
(top view of one layer)



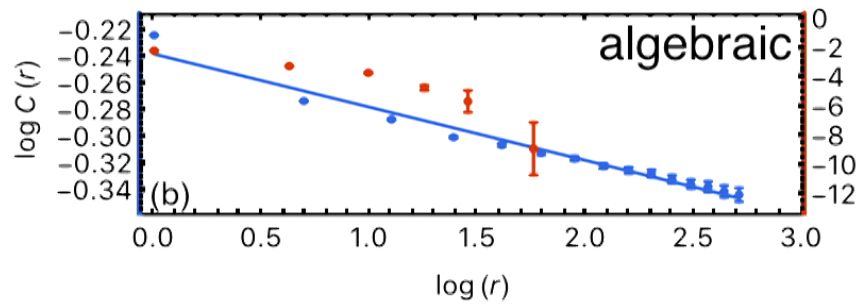
- Objective: $\min \mathsf{KL}(P_\partial[\phi]||e^{-S[\phi]}) =$
$\mathbb{E}_{[\zeta]}\left(S[G[\zeta]] - ||\zeta||^2 - \ln\det\left(\frac{\delta G[\zeta]}{\delta\zeta}\right)\right)$
  - Sample ζ from bulk, push to the boundary $\phi = G[\zeta]$
  - Forward: evaluate loss function
  - Backward: propagate gradient to train bijectors

Performance of the Generative Model

- Let us first make sure that the machine learns the correct physics from the given action.
  - Phase diagram (32x32 finite size lattice)



  - Correlation function



Machine Learning Holography

- Training a generative model establishes a holographic duality

$$\min \mathsf{KL}(P[\zeta]\det(\delta_\zeta G[\zeta])^{-1} || e^{-S[\phi]})$$

CFT (boundary) | **AdS** (bulk)

$$Z = \mathrm{Tr}_{[\phi]}\, e^{-S[\phi]} \longleftrightarrow Z = \mathrm{Tr}_{[\zeta]}\, P[\zeta]\det(\delta_\zeta G[\zeta])^{-1}$$

Field theory in flat space
- massless field $\phi(x)$

(Classical) gravity + matter
- massive matter $\zeta(x, z)$
- on background $G[\cdot; w]$

Features in dataset
- image $\phi(x)$

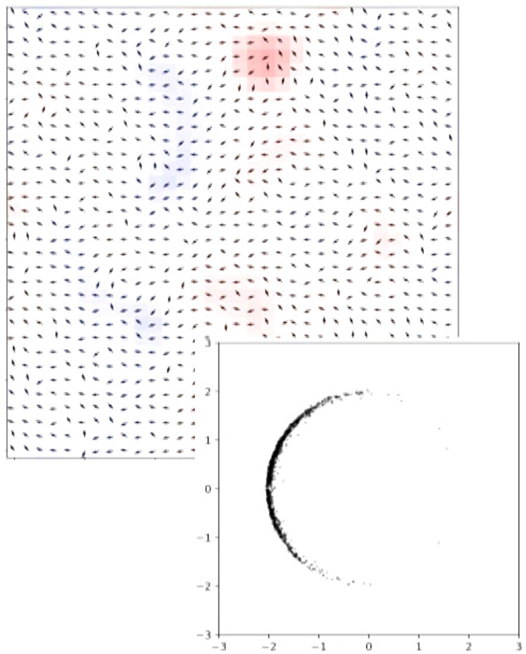
Deep generative model
- latent representation $\zeta(x, z)$
- neural network $G[\cdot; w]$

Efficient Sampling from the Bulk

- Sampling: holographic mapping from bulk to boundary
  - Massive field in the bulk → Critical field on the boundary
  - Local update in the bulk → Global update on the boundary

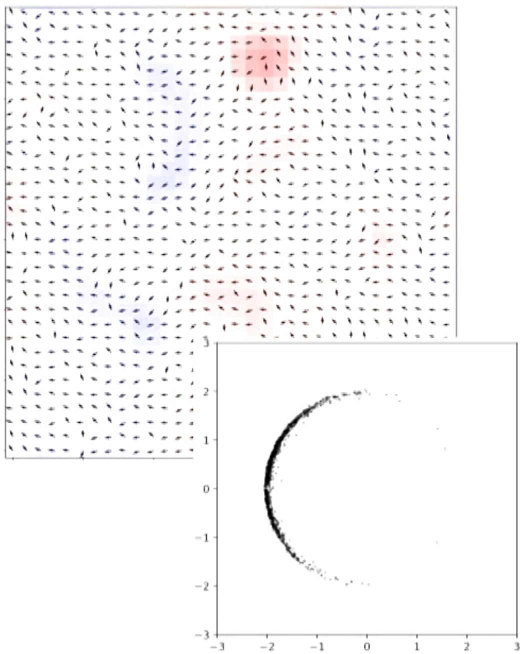
Efficient Sampling from the Bulk

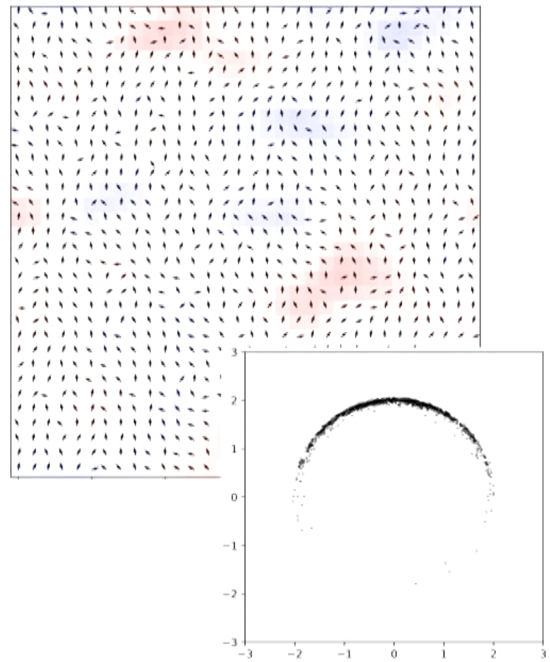
- Sampling: holographic mapping from bulk to boundary
  - Massive field in the bulk → Critical field on the boundary
  - Local update in the bulk → Global update on the boundary



Efficient Sampling from the Bulk

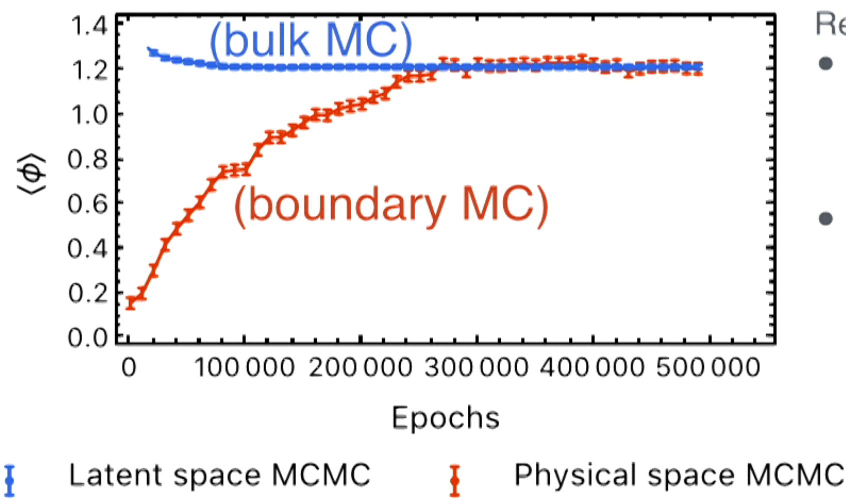
- Sampling: holographic mapping from bulk to boundary
  - Massive field in the bulk → Critical field on the boundary
  - Local update in the bulk → Global update on the boundary





Efficient Sampling from the Bulk

- Sampling: holographic mapping from bulk to boundary
  - Massive field in the bulk → Critical field on the boundary
  - Local update in the bulk → Global update on the boundary
- Order parameters converges faster using bulk MCMC.



Related topics:

- Self-learning MC
  Huang, Wang, PRB (2017)
  Liu, Qi, Meng, Fu, PRB(2017)
  …
- Super-resolution sampling
  Efthymiou, Beach, Melko (2019)

Probing Holographic Bulk Geometry

- Inference: holographic mapping from boundary to bulk
  - Push the boundary field distribution back into the bulk → Bulk Effective Theory

$$S_{\mathrm{eff}}[\zeta] = ||\zeta||^2 + \ln P_\partial[G[\zeta]] - \ln P[G[\zeta]]$$

model (P_∂) ↑ target (P) ↑

- Bulk field will have residual correlation
  - Pessimist: model is too weak, training is not perfect …
  - Optimist: important message about bulk geometry!

$$S_{\mathrm{eff}}[\zeta] = \int_{\mathcal{M}} g^{\mu\nu}\partial_\mu\zeta^*\partial_\nu\zeta + m^2|\zeta|^2 + u|\zeta|^4 + \cdots$$

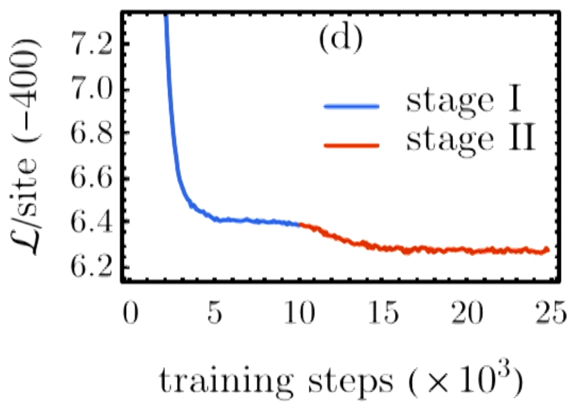
$$\langle\zeta_i\zeta_j^*\rangle \text{ or } I(\zeta_i:\zeta_j) \sim e^{-d_{ij}/\xi}$$

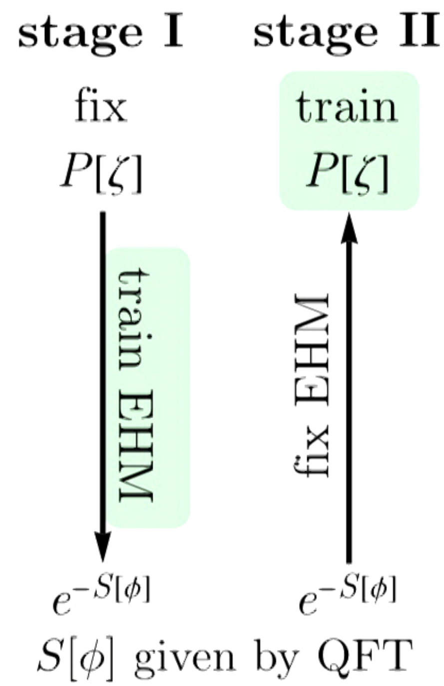
correlation mutual info.

Probing Holographic Bulk Geometry

- To fit the quadratic terms, we can model the prior by correlated Gaussian

$$P[\zeta] \sim e^{-\zeta^\dagger K \zeta}$$





- Geometry from information

$$d_{ij} = -\xi I(\zeta_i : \zeta_j) = \frac{\xi}{2} \ln\left(1 - \frac{\langle \zeta_i^* \zeta_j \rangle^2}{\langle \zeta_i^* \zeta_i \rangle \langle \zeta_j^* \zeta_j \rangle}\right)$$

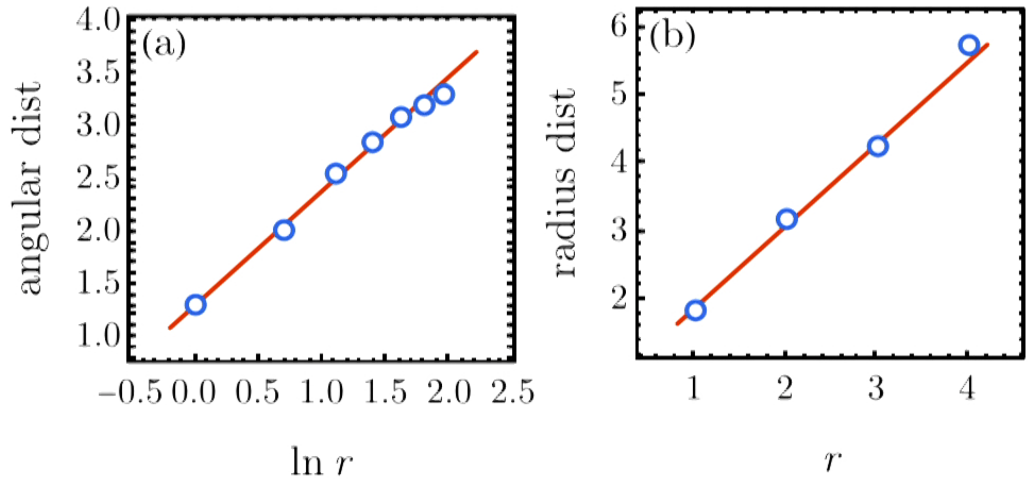
geodesic distance

X-L Qi (2013)

Probing Holographic Bulk Geometry

- Apply to Luttinger liquid CFT, measure the bulk distance

$$d(x,y,z|x+r,y,z) \sim \ln r \qquad d(x,y,z|x,y,z+r) \sim r$$



- Result matches hyperbolic geometry ~ AdS

$$\mathrm{d}s^2 = \frac{1}{z^2}(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2)$$

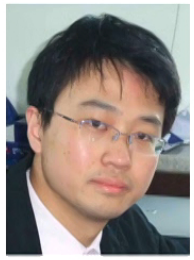
Summary

- We demonstrated several examples of machine learning physics. The common theme:
  - Train the machine on a task (but we don't use it!)
  - Open up the neural network for emergent physics

	Task	Emergent physics
ML Quantum Mechanics arXiv: 1901.11103	Potential-density mapping	Wave function + Schrödinger eq.
ML Holographic Mapping arXiv:1903.00804	Quantum field generation	RG scheme, bulk effective theory

Acknowledgment

- ML Quantum Mechanics

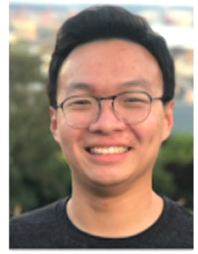



Ce Wang   Hui Zhai

(Tsinghua University)

C Wang, H Zhai, Y-Z You. arXiv: 1901.11103

- Neural-RG and Holography



Hong-Ye Hu   Shuo-Hui Li   Lei Wang

(UCSD)   (IOP, CAS)

H Hu, S-H Li, L Wang, Y-Z You. arXiv: 1903.00804