Quantum generative adversarial networks

Pierre-Luc Dallaire-Demers

TDLS
Toronto| June 18, 2018

XANADU
Think With Light ™

Agenda

1) About Xanadu

2) Quantum Computing

3) QuGANs

About Xanadu

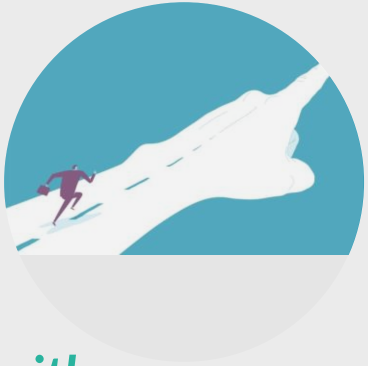
About Us

We're a **photonic-based** quantum computing company with expertise in:

Quantum computing hardware & software

Quantum & classical machine learning

High-performance computing & benchmarking



Currently developing and testing new algorithms, along with quantum photonic processing chips

About Us

Based in Toronto

VC Backed

25^{+} PhDs

Full-Stack

World-Class Team


Christian Weedbrook (CEO), PhD


Zachary Vernon, PhD


Kamil Bradler, PhD


Nicolas Quesada, PhD


Timjan Kalajdzievski, Current PhD


Thomas Bromley, PhD


Daiqin Su, PhD


Nathan Killoran, PhD


Andy Feng, BBA


Maria Schuld, PhD


Juan Miguel Arrazola, PhD


Kang Tan, PhD


Joshua Izaac, PhD


Brajesh Gupt, PhD


Pierre-Luc Dallaire-Demers, PhD


Krishnakumar Sabapathy, PhD


Razieh Annabestani, PhD


Haoyu Qi, PhD


Reihaneh Shahrokhshahi, PhD


Blair Morrison, PhD


Dylan Mahler, PhD


Casey Myers, PhD


Matteo Menotti, PhD


Ranier Sandoval, MBA


Mariam Naseem, MBA


Seth Lloyd, PhD


Dirk Englund, PhD

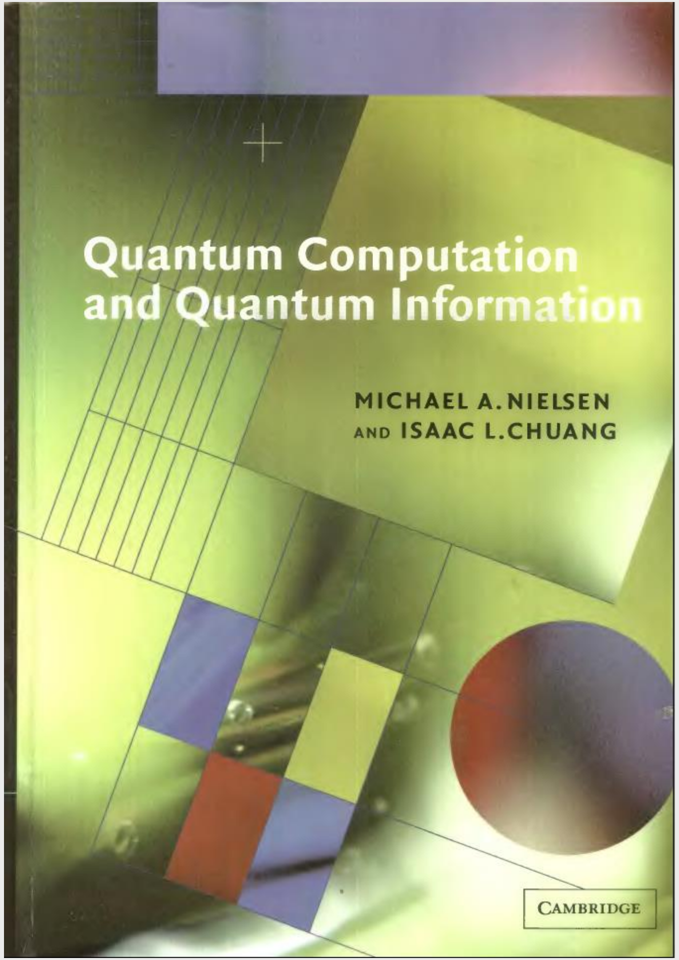

John Sipe, PhD


Marco Liscidini, PhD


Matthew Collins, PhD

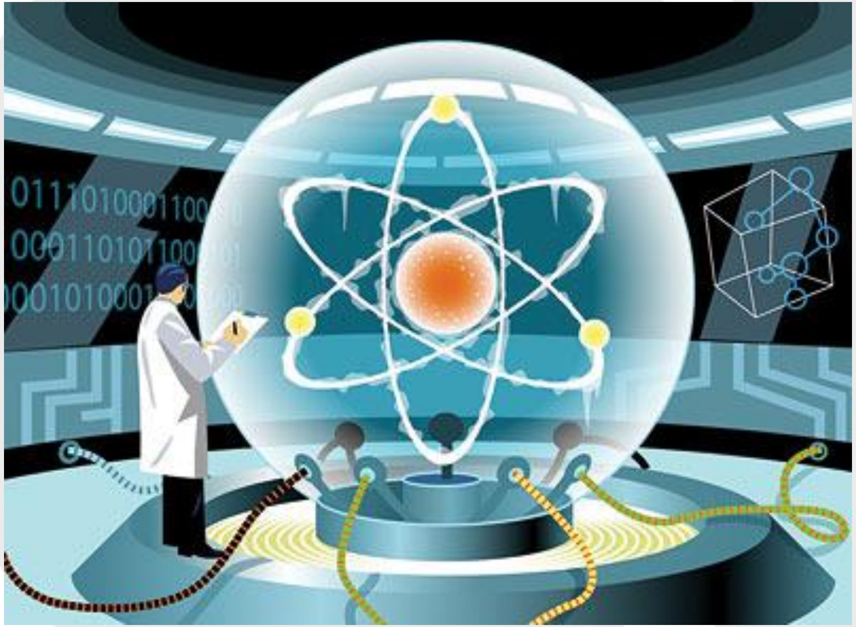
Quantum Computing

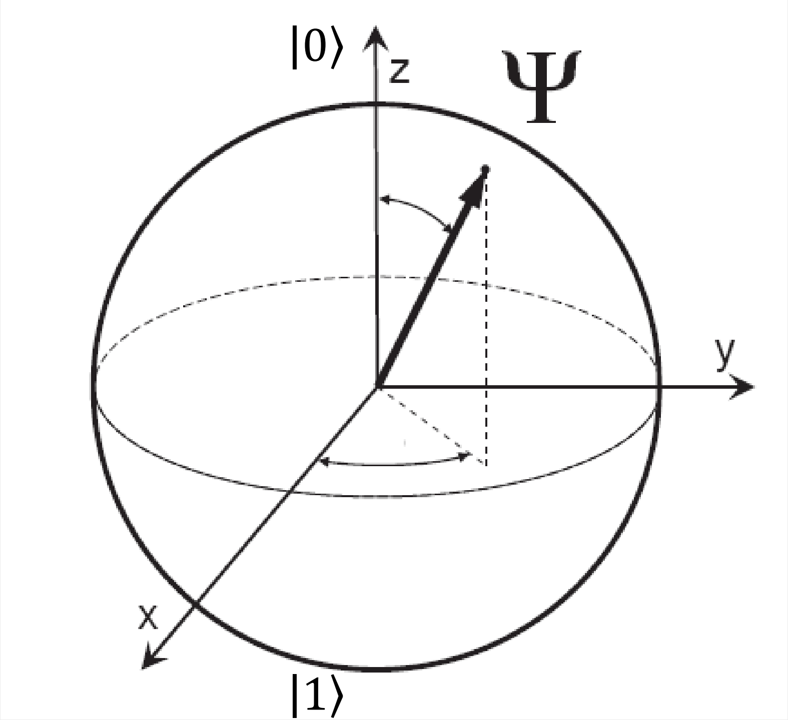
Best textbook



Nielsen, M.A. and Chuang, I., 2002.
Quantum computation and quantum information.

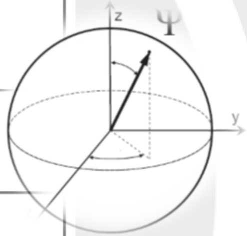
Qubits

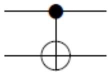
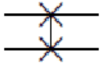
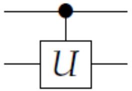


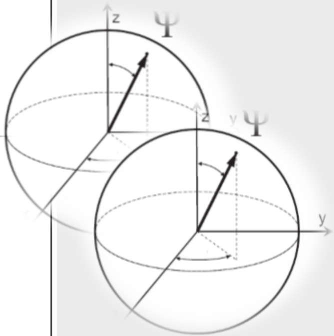


Quantum gates

Gate	Matrix representation
$\boxed{\mathrm{I}}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
$\boxed{\mathrm{X}}$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
$\boxed{\mathrm{Y}}$	$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$
$\boxed{\mathrm{Z}}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
$\boxed{\mathrm{H}}$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$
$\boxed{\mathrm{S}}$	$\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$
$\boxed{\mathrm{T}}$	$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}$
$\boxed{\mathrm{W}}$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}$

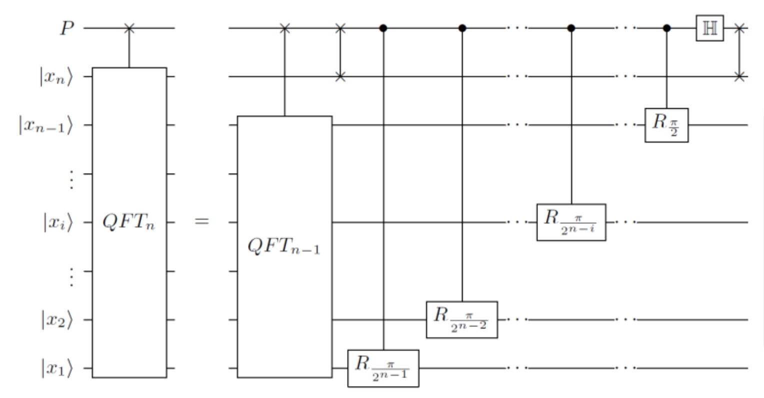


Name	Gate	Matrix representation
CNOT		$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$
SWAP		$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$
c-U		$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & \gamma & \delta \end{pmatrix}$



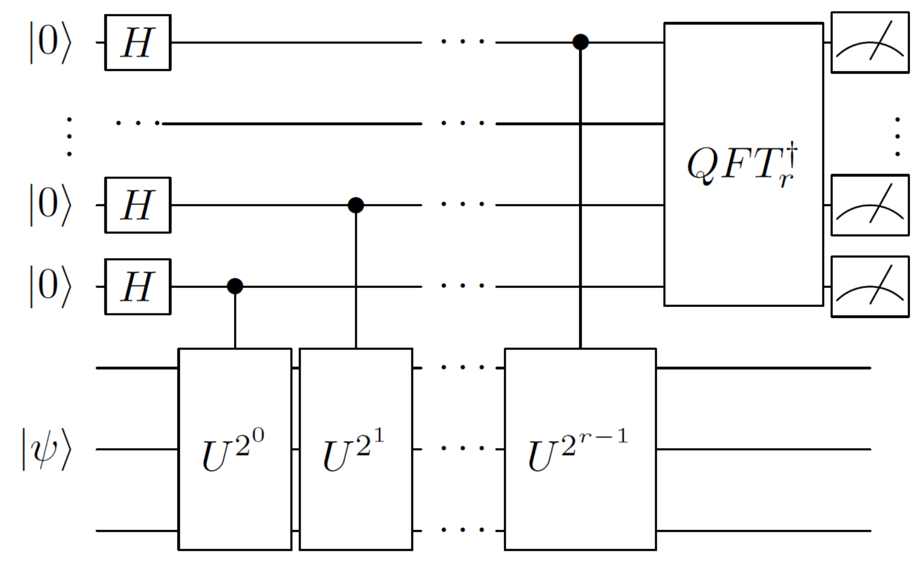
Quantum algorithms

Quantum Fourier transform



$$QFT\,|j\rangle \equiv \frac{1}{\sqrt{\mathcal{N}}}\sum_{k=0}^{\mathcal{N}-1} e^{\frac{2\pi i j k}{\mathcal{N}}}\,|k\rangle$$

Phase estimation



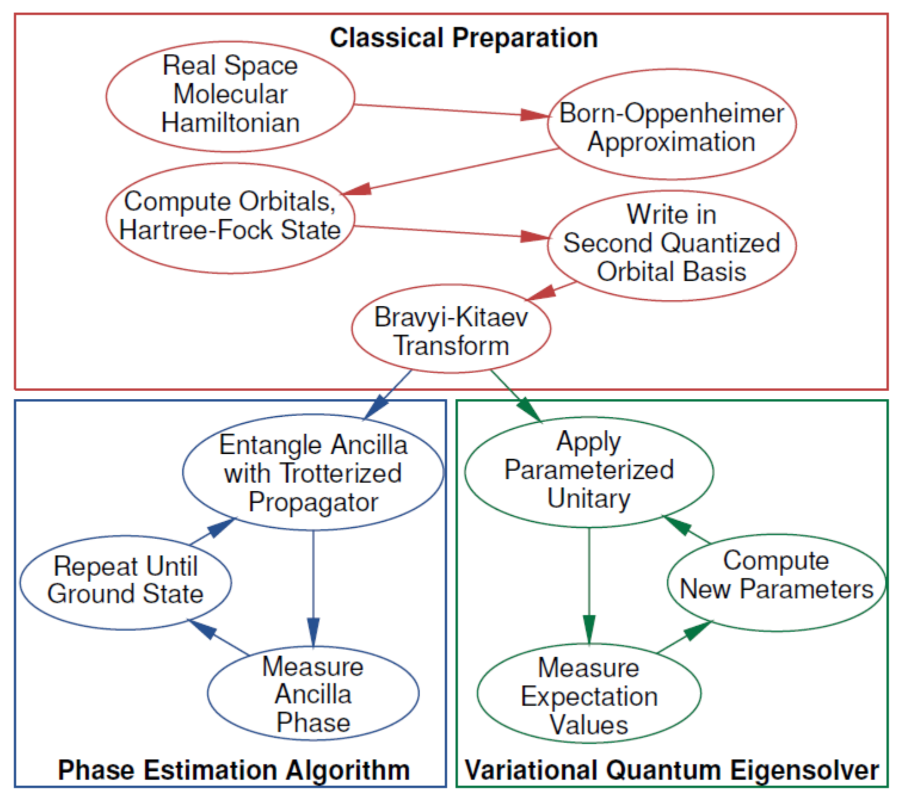
Quantum chemistry with variational circuits

$$H = \sum_{pq} t_{pq} a_p^\dagger a_q + \sum_{pqrs} v_{pqrs} a_p^\dagger a_q^\dagger a_s a_r$$

$$E_0 = \mathrm{tr}\left(H\rho_0\right) \qquad \rho_0 = \left|\Psi_0\right\rangle\left\langle\Psi_0\right|$$

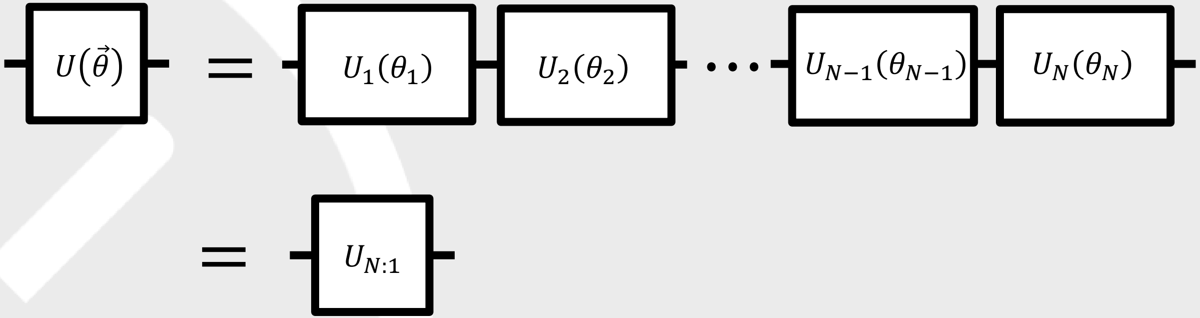
Trotterization is expensive!

$$e^{-iH\Delta t} \simeq \left(\prod_{i=1}^{K} e^{-\frac{iH_i\Delta t}{n_T}}\right)^{n_T} + \sum_{i<j}\frac{[H_i, H_j]\,(\Delta t)^2}{2n_T} + \ldots$$



Peruzzo *et al.* (2014)

Parametrized quantum circuits



$$U(\vec{\theta}) \equiv U_N(\theta_N)U_{N-1}(\theta_{N-1}) \dots U_2(\theta_2)U_1(\theta_1) = U_{N:1}$$

A parametrized gate: $U_j(\theta_j) = e^{-\frac{i}{2}\theta_j h_j}$

Self-adjoint generator: $h_j = h_j^\dagger$

Unitary operations

$$U_j^{-1}(\theta_j) = U_j^\dagger(\theta_j) = U_j(-\theta_j) = e^{+\frac{i}{2}\theta_j h_j}$$

$$U_j(\theta_j)U_j^\dagger(\theta_j) = I$$

The quantum subroutine

$$H = \sum_{pq} t_{pq} a_p^\dagger a_q + \sum_{pqrs} v_{pqrs} a_p^\dagger a_q^\dagger a_s a_r$$

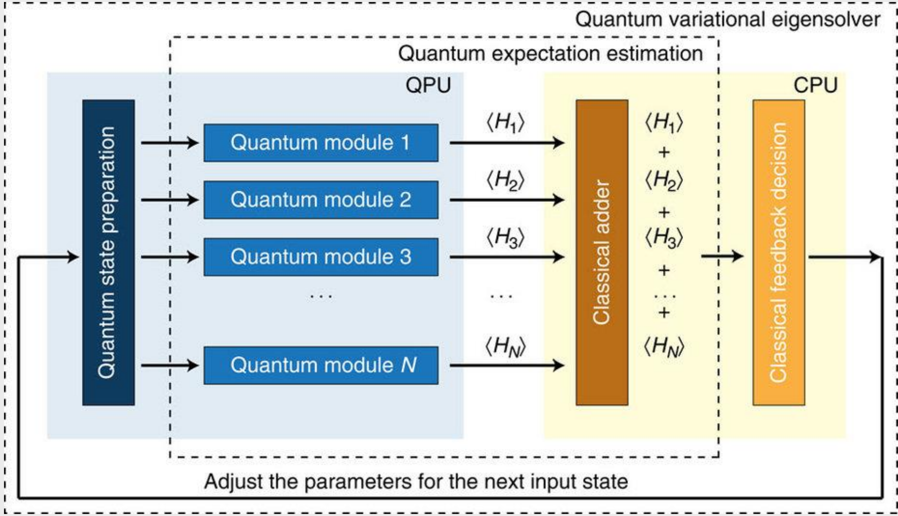
Hartree-Fock reference state

$$|\Phi_0\rangle = \prod_k a_k^\dagger |\mathrm{vac}\rangle$$

Variational unitary coupled cluster

$$\min_\Theta E(\Theta) = \langle \Psi(\Theta)| H |\Psi(\Theta)\rangle$$

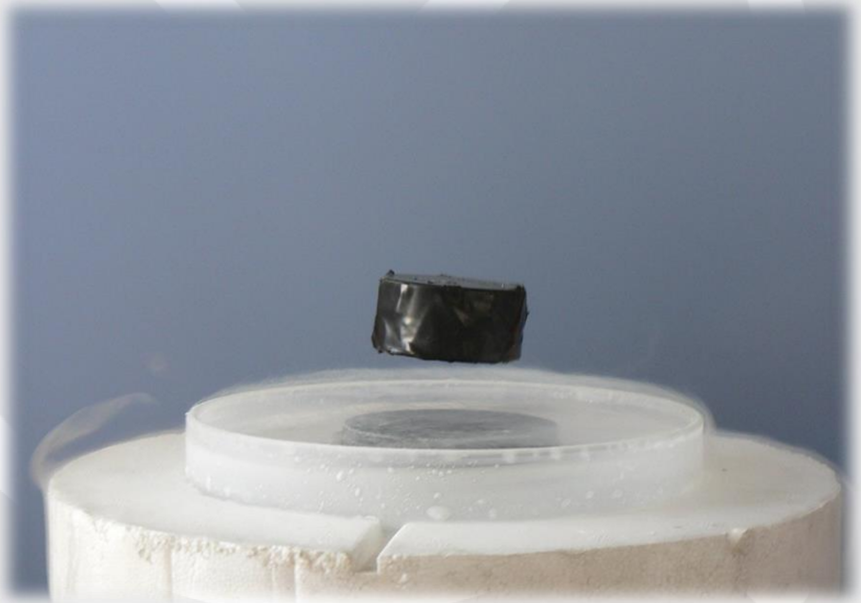
$$|\Psi(\Theta)\rangle = e^{i\left(\mathcal{T}(\Theta)+\mathcal{T}^\dagger(\Theta)\right)} |\Phi_0\rangle$$

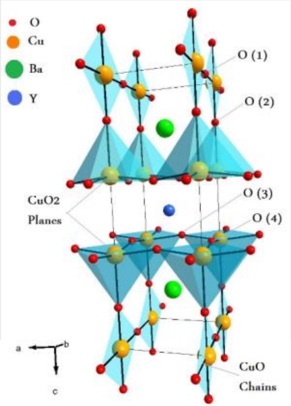


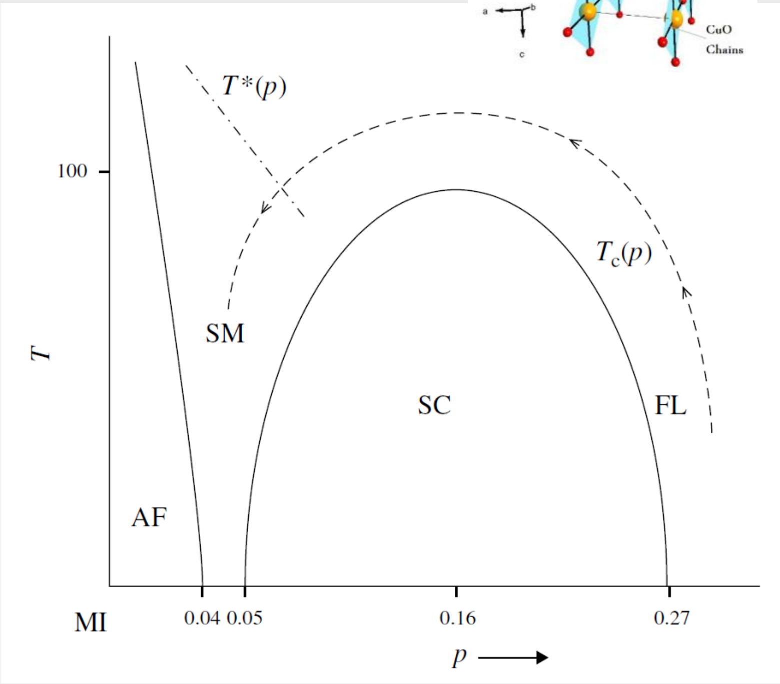
Peruzzo *et al.* (2014)

Applications - Superconductivity

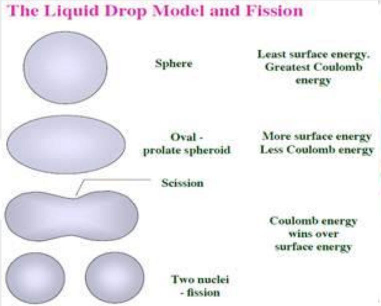
$$(\mathrm{CuO_2})_n\,\mathrm{A}_{n-1}\mathrm{X}$$

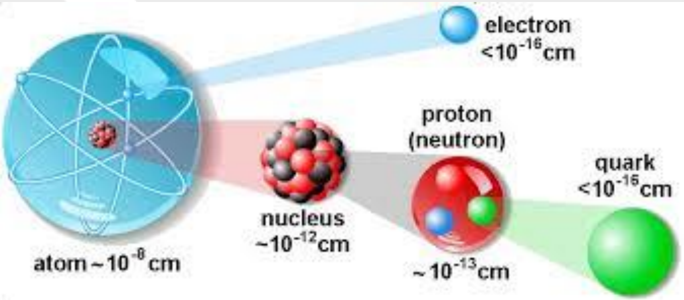


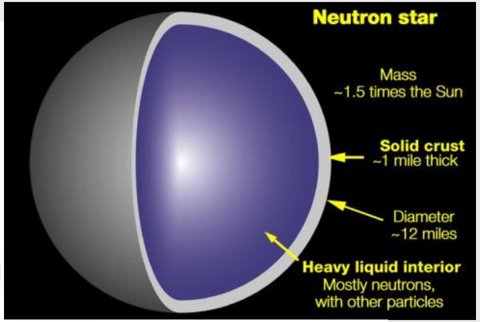


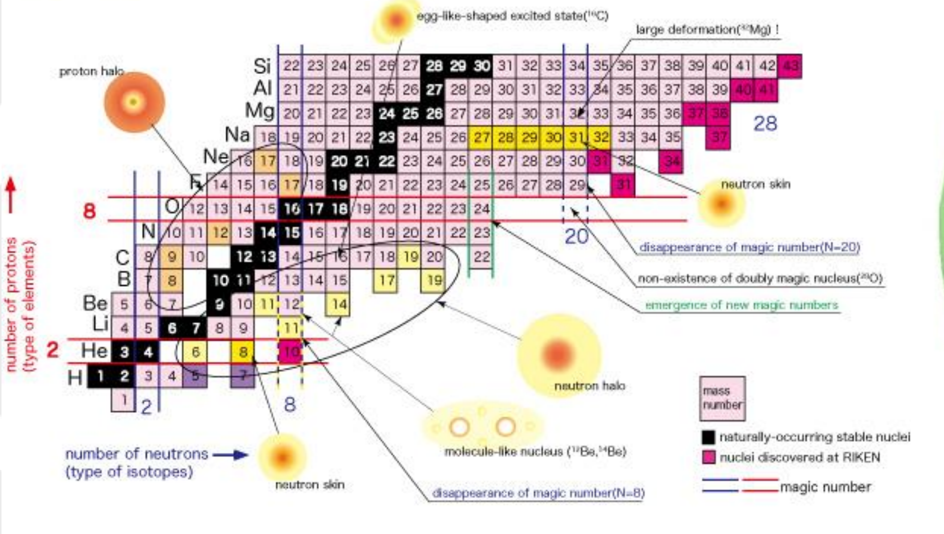


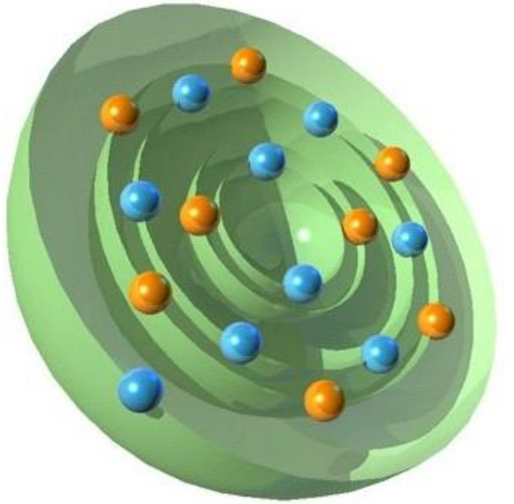
Applications – Nuclear physics











QuGANs

Quantum generative adversarial learning

Quantum generative adversarial learning

Seth Lloyd[1] and Christian Weedbrook[2]
[1] *Massachusetts Institute of Technology, Department of Mechanical Engineering, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA*
[2] *Xanadu, 372 Richmond Street W, Toronto, Ontario M5V 1X6, Canada*
(Dated: April 25, 2018)

Generative adversarial networks (GANs) represent a powerful tool for classical machine learning: a generator tries to create statistics for data that mimics those of a true data set, while a discriminator tries to discriminate between the true and fake data. The learning process for generator and discriminator can be thought of as an adversarial game, and under reasonable assumptions, the game converges to the point where the generator generates the same statistics as the true data and the discriminator is unable to discriminate between the true and the generated data. This paper introduces the notion of quantum generative adversarial networks (QuGANs), where the data consists either of quantum states, or of classical data, and the generator and discriminator are equipped with quantum information processors. We show that the unique fixed point of the quantum adversarial game also occurs when the generator produces the same statistics as the data. Since quantum systems are intrinsically probabilistic the proof of the quantum case is different from – and simpler than – the classical case. We show that when the data consists of samples of measurements made on high-dimensional spaces, quantum adversarial networks may exhibit an exponential advantage over classical adversarial networks.

Lloyd, S. and Weedbrook, C., 2018. Quantum generative adversarial learning. arXiv:1804.09139.

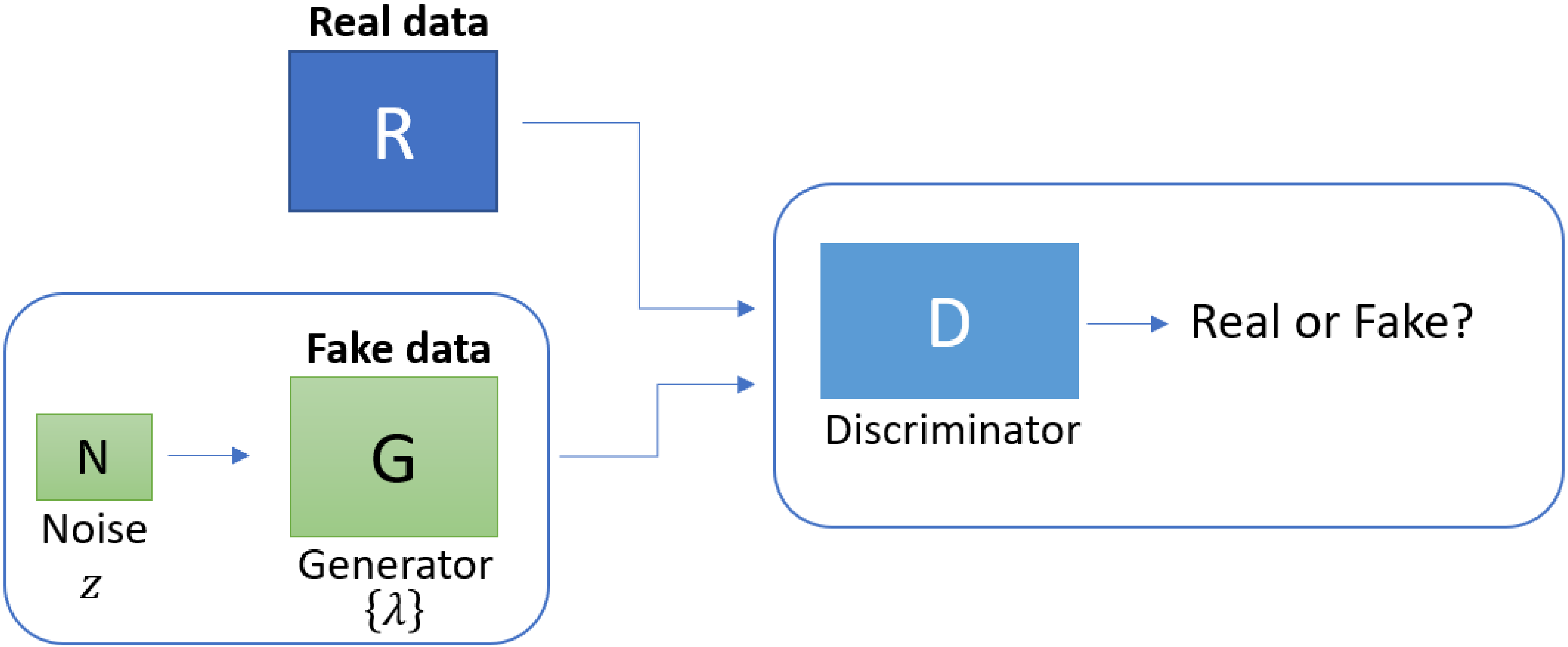
Quantum generative adversarial networks

Pierre-Luc Dallaire-Demers* and Nathan Killoran
Xanadu, 372 Richmond Street W, Toronto, Ontario M5V 1X6, Canada
(Dated: May 2, 2018)

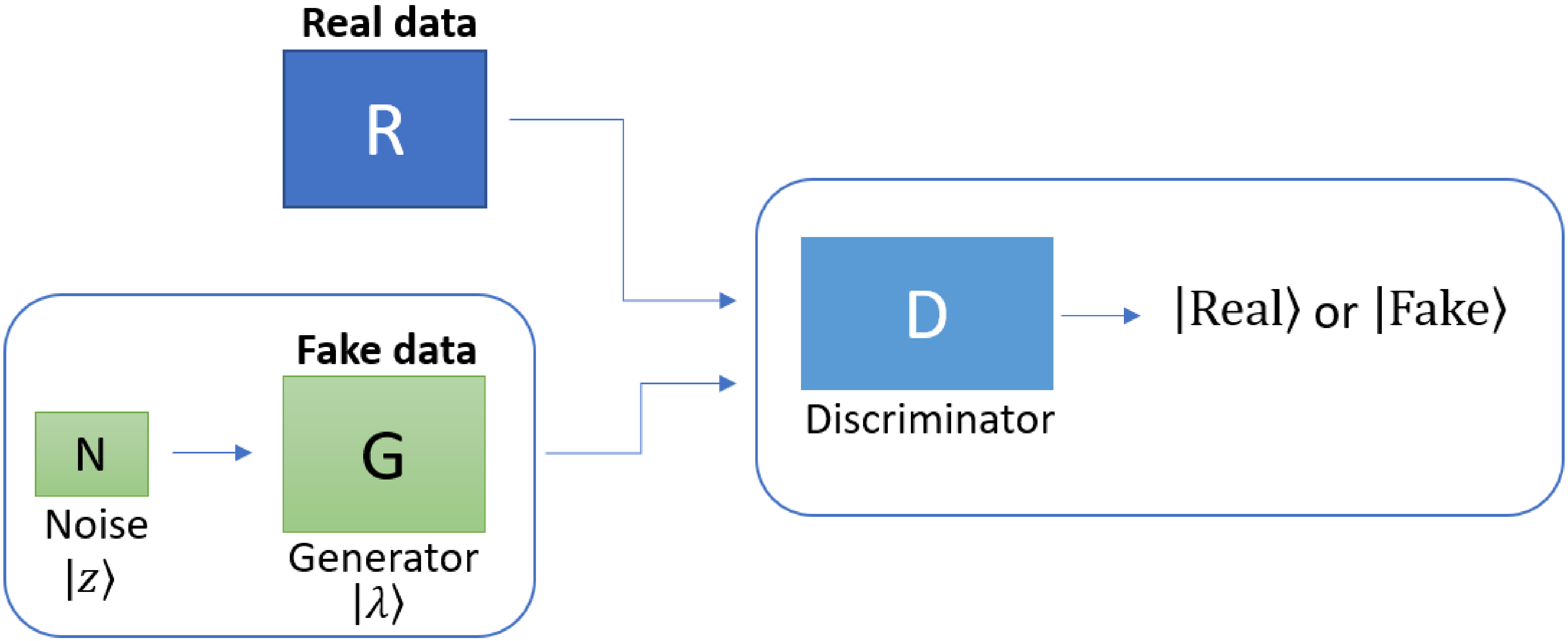
Quantum machine learning is expected to be one of the first potential general-purpose applications of near-term quantum devices. A major recent breakthrough in classical machine learning is the notion of generative adversarial training, where the gradients of a discriminator model are used to train a separate generative model. In this work and a companion paper, we extend adversarial training to the quantum domain and show how to construct generative adversarial networks using quantum circuits. Furthermore, we also show how to compute gradients – a key element in generative adversarial network training – using another quantum circuit. We give an example of a simple practical circuit ansatz to parametrize quantum machine learning models and perform a simple numerical experiment to demonstrate that quantum generative adversarial networks can be trained successfully.

Dallaire-Demers, P.L. and Killoran, N., 2018. Quantum generative adversarial networks. arXiv:1804.08641.

Classical GANs

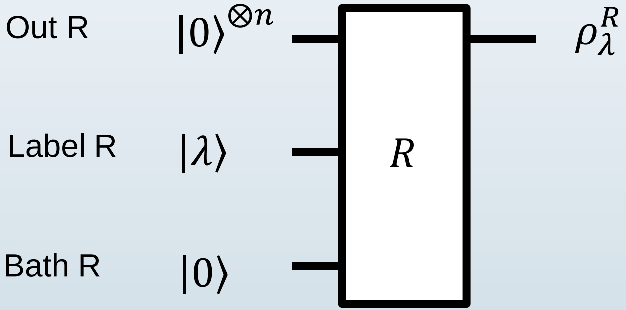


Quantum GANs

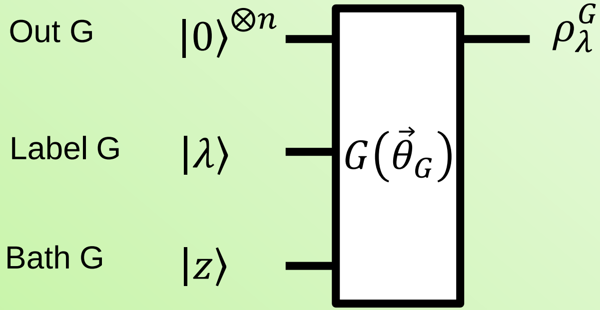


$$\rho = \{(p_1, |\psi_1\rangle); (p_2, |\psi_2\rangle); \ldots ; (p_d, |\psi_d\rangle)\}$$

Quantum sources of data



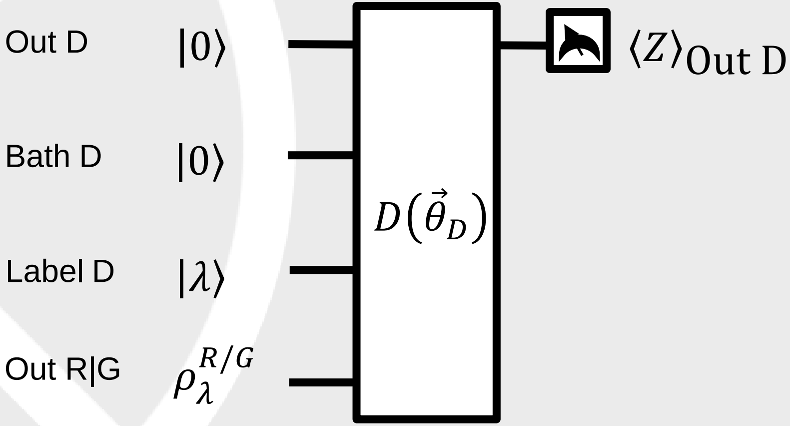
$$R(|\lambda\rangle) = \rho_\lambda^R$$



$$G\left(\vec{\theta}_G, |\lambda, z\rangle\right) = \rho_\lambda^G\left(\vec{\theta}_G, z\right)$$

Quantum discriminator

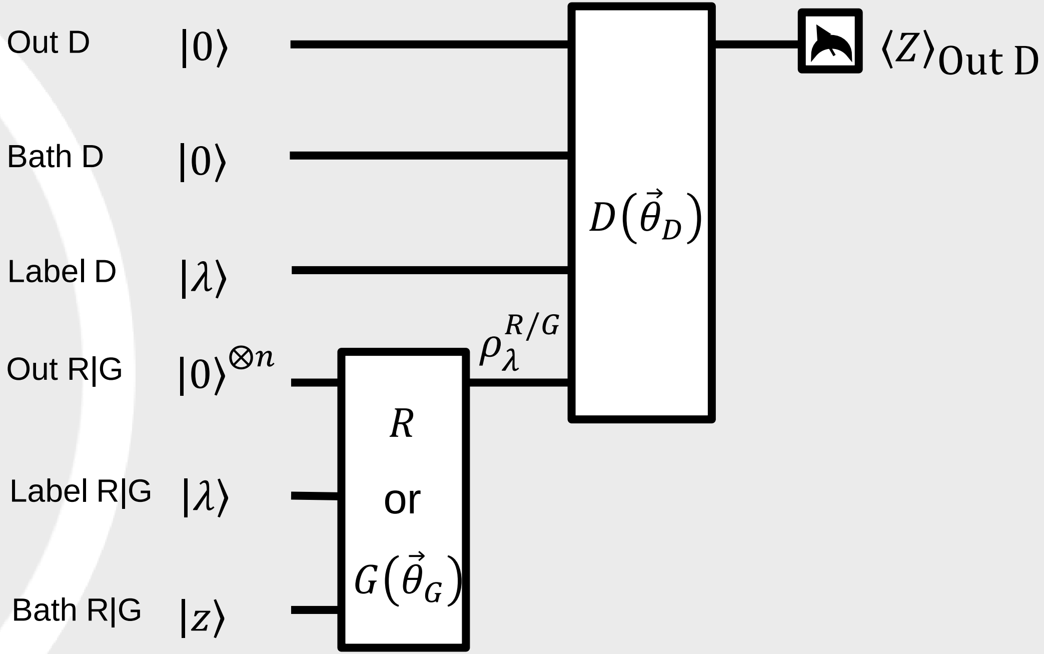
$$Z \equiv |\text{real}\rangle\langle\text{real}| - |\text{fake}\rangle\langle\text{fake}|$$



$$D\left(\vec{\theta}_D, |\lambda\rangle, \rho_\lambda^{R/G}\right)$$

The cost function

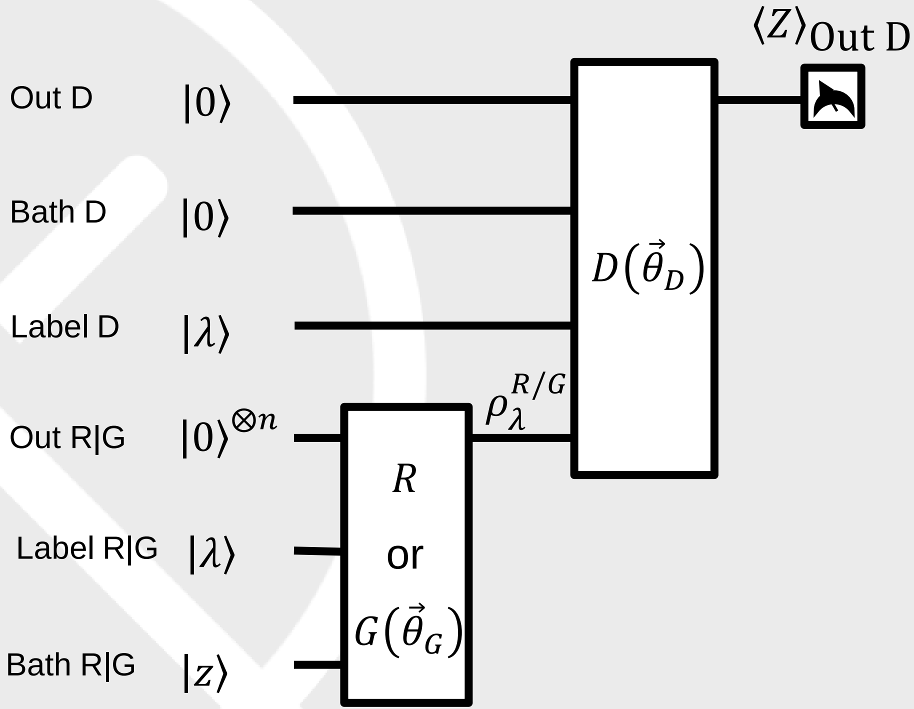
$$\min_{\vec{\theta}_G} \max_{\vec{\theta}_D} V(\vec{\theta}_D, \vec{\theta}_G)$$



$$V(\vec{\theta}_D, \vec{\theta}_G) = \frac{1}{\Lambda}\sum_{\lambda=1}^{\Lambda} \Pr\Big(\big(D\left(\vec{\theta}_D, |\lambda\rangle, R(|\lambda\rangle)\right) = |\text{real}\rangle\big) \cap \big(D\left(\vec{\theta}_D, |\lambda\rangle, G(\vec{\theta}_G, |\lambda, z\rangle)\right) = |\text{fake}\rangle\big)\Big)$$

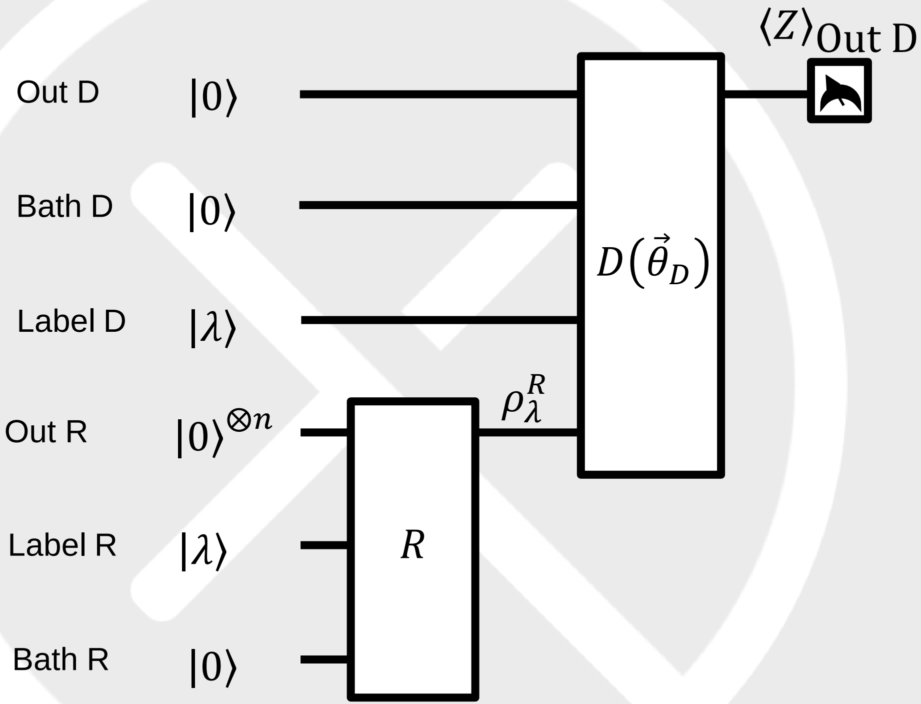
Can we formulate this in the language of quantum mechanics?

Notation – Initial state



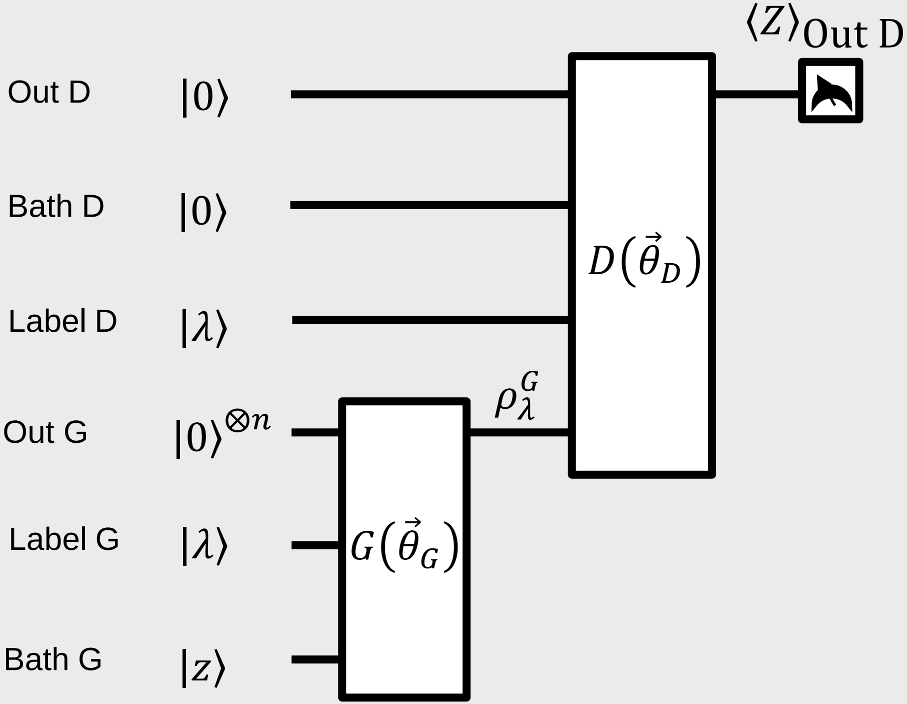
$$\rho_\lambda^0(z) = (|0\rangle\langle 0|)^{\otimes d+1} \otimes |\lambda\rangle\langle\lambda| \otimes (|0\rangle\langle 0|)^{\otimes n} \otimes |\lambda\rangle\langle\lambda| \otimes |z\rangle\langle z|$$

Notation – Source states



$$U_R = I^{\otimes(1+d+s)} \otimes R$$

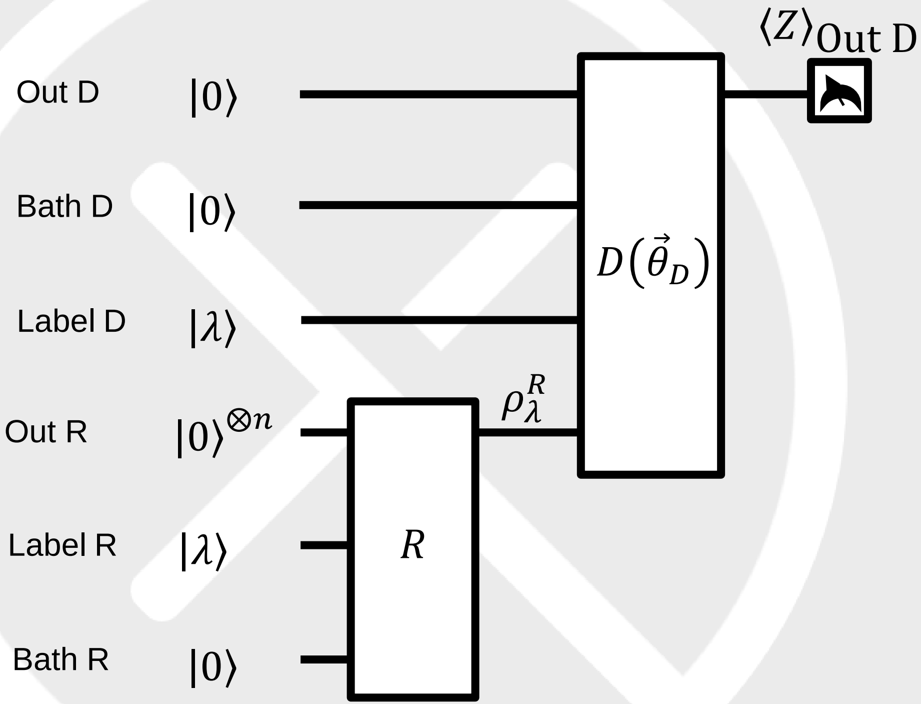
$$\rho_\lambda^R = U_R \rho_\lambda^0(0) U_R^\dagger$$

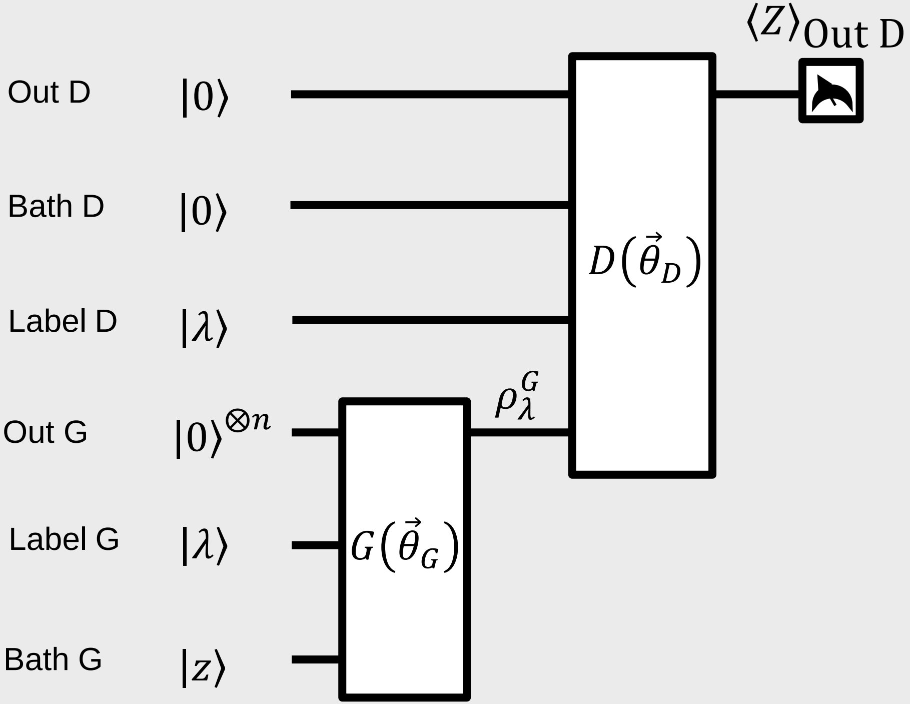


$$U_G(\vec{\theta}_G) = I^{\otimes(1+d+s)} \otimes G(\vec{\theta}_G)$$

$$\rho_\lambda^G(\vec{\theta}_G, z) = U_G(\vec{\theta}_G)\rho_\lambda^0(z)U_G^\dagger(\vec{\theta}_G)$$

Notation – Discriminator state

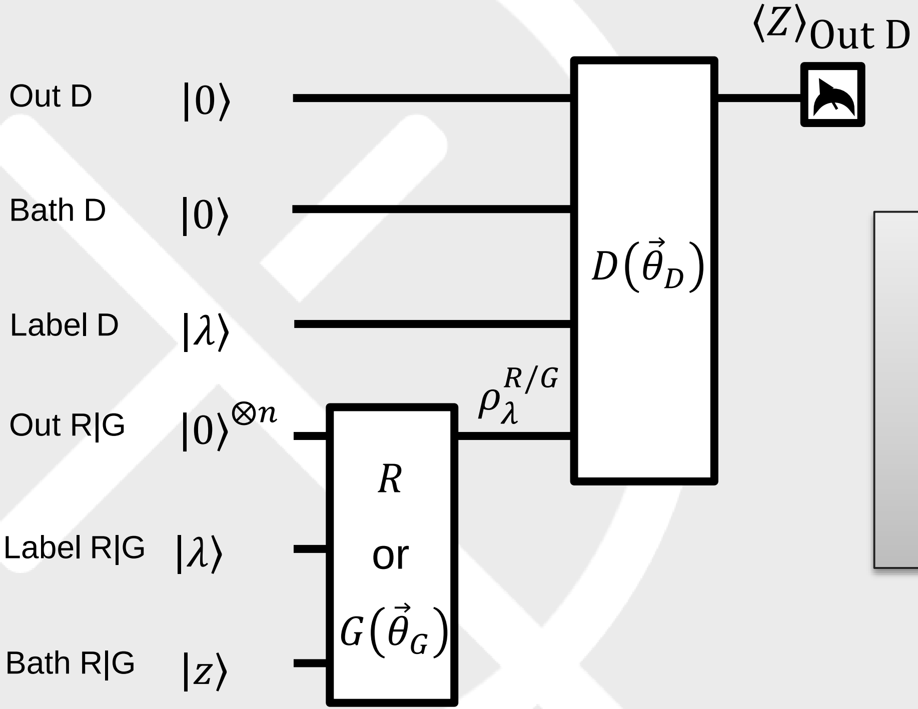




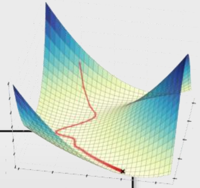
$$U_D(\vec{\theta}_D) = D(\vec{\theta}_D) \otimes I^{\otimes m}$$

$$\rho_\lambda^{DR}(\vec{\theta}_D) = U_D(\vec{\theta}_D)\rho_\lambda^R U_D^\dagger(\vec{\theta}_D) \qquad \rho_\lambda^{DG}(\vec{\theta}_D, \vec{\theta}_G, z) = U_D(\vec{\theta}_D)\rho_\lambda^G(\vec{\theta}_G, z)U_D^\dagger(\vec{\theta}_D)$$

The quantum cost function



Gradient update rules:

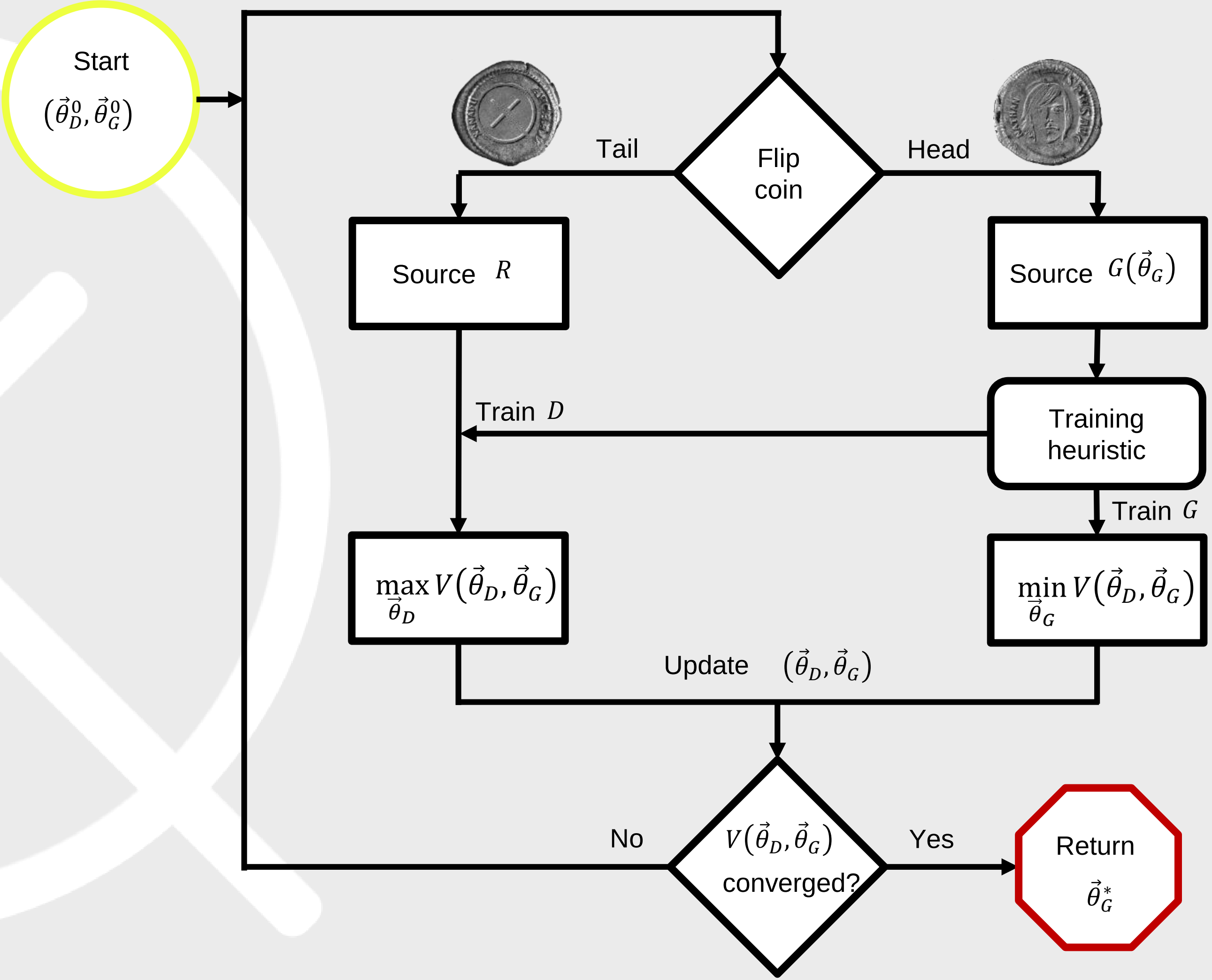


$$\vec{\theta}_D^{k+1} = \vec{\theta}_D^{k} + \chi_D^k \nabla_{\vec{\theta}_D} V\left(\vec{\theta}_D^k, \vec{\theta}_G^k\right)$$

$$\vec{\theta}_G^{k+1} = \vec{\theta}_G^{k} - \chi_G^k \nabla_{\vec{\theta}_G} V\left(\vec{\theta}_D^k, \vec{\theta}_G^k\right)$$

$$V(\vec{\theta}_D, \vec{\theta}_G) = \frac{1}{2} + \frac{1}{4\Lambda} \sum_{\lambda=1}^{\Lambda} \operatorname{tr}\left(\left(\rho_\lambda^{DR}(\vec{\theta}_D\ ) - \rho_\lambda^{DG}(\vec{\theta}_D, \vec{\theta}_G, z)\right) Z\right)$$

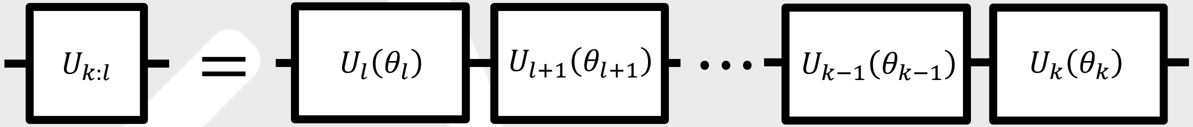
Training



Start
$(\vec{\theta}_D^0, \vec{\theta}_G^0)$
Flip coin
Tail
Head
Source R
Source $G(\vec{\theta}_G)$
Training heuristic
Train D
Train G
$\max_{\vec{\theta}_D} V(\vec{\theta}_D, \vec{\theta}_G)$
$\min_{\vec{\theta}_G} V(\vec{\theta}_D, \vec{\theta}_G)$
Update $(\vec{\theta}_D, \vec{\theta}_G)$
$V(\vec{\theta}_D, \vec{\theta}_G)$ converged?
No
Yes
Return
$\vec{\theta}_G^*$

Quantum Gradients

Chain rule for quantum gates



$$U_{k:l} \equiv U_k(\theta_k)U_{k-1}(\theta_{k-1}) \dots U_{l+1}(\theta_{l+1})U_l(\theta_l)$$

One gate:

$$U_j(\theta_j) = e^{-\frac{i}{2}\theta_j h_j}$$

Derivative of one gate (*Schrödinger equation*):

$$\frac{\partial}{\partial \theta_j} U_j(\theta_j) = -\frac{i}{2} h_j U_j(\theta_j)$$

Derivative for a sequence of gates:

$$\frac{\partial}{\partial \theta_j} U(\vec{\theta}) = -\frac{i}{2} U_{N:j+1} h_j U_{j:1}$$

Time-reversed notation

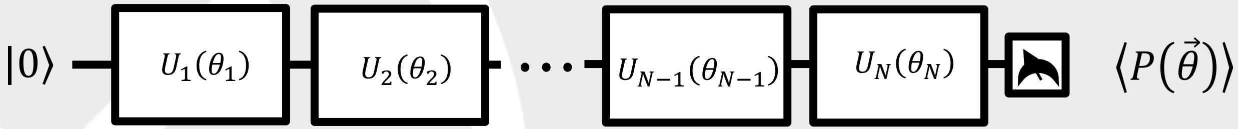
$$U_{l:k}^\dagger \equiv U_l^\dagger(\theta_l)U_{l+1}^\dagger(\theta_{l+1}) \dots U_{k-1}^\dagger(\theta_{k-1})U_k^\dagger(\theta_k)$$

$$U_j^\dagger(\theta_j) = e^{+\frac{i}{2}\theta_j h_j} \qquad h_j = h_j^\dagger$$

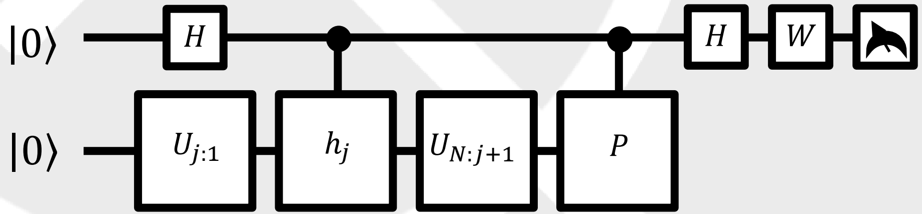
$$\frac{\partial}{\partial \theta_j} U^\dagger(\vec{\theta}) = +\frac{i}{2} U_{1:j}^\dagger h_j U_{j+1:N}^\dagger$$

A quantum circuit for gradients

$$\langle P(\vec{\theta})\rangle = \mathrm{tr}\left(\rho_0 U^\dagger(\vec{\theta}) P U(\vec{\theta})\right)$$



$$\frac{\partial}{\partial\theta_j}\langle P(\vec{\theta})\rangle = -\frac{i}{2}\mathrm{tr}\left(\rho_0 U^\dagger_{1:j}\left[U^\dagger_{j+1:N} P U_{N:j+1}, h_j\right] U_{j:1}\right)$$



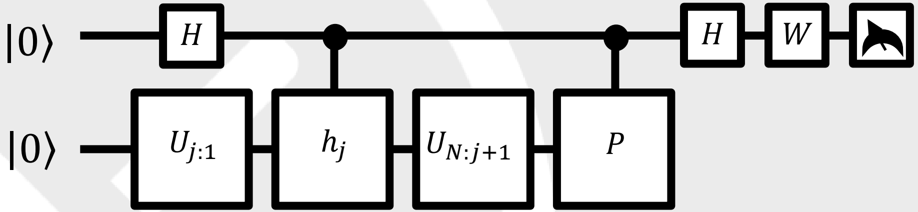
$$\langle Z\rangle_{\mathrm{Grad}} = \frac{\partial}{\partial\theta_j}\langle P(\vec{\theta})\rangle$$

$$\langle Z\rangle_{\mathrm{Grad}} = \Pr\left(|x_{\mathrm{Grad}}\rangle = |0\rangle\right) - \Pr\left(|x_{\mathrm{Grad}}\rangle = |1\rangle\right)$$

Self-adjoint & unitary

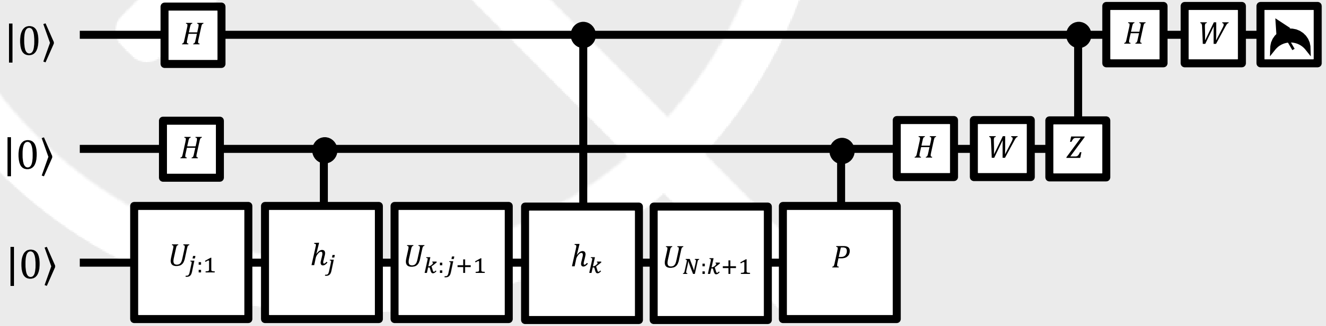
$$h_j = h_j^\dagger = h_j^{-1}$$

Quantum Hessian



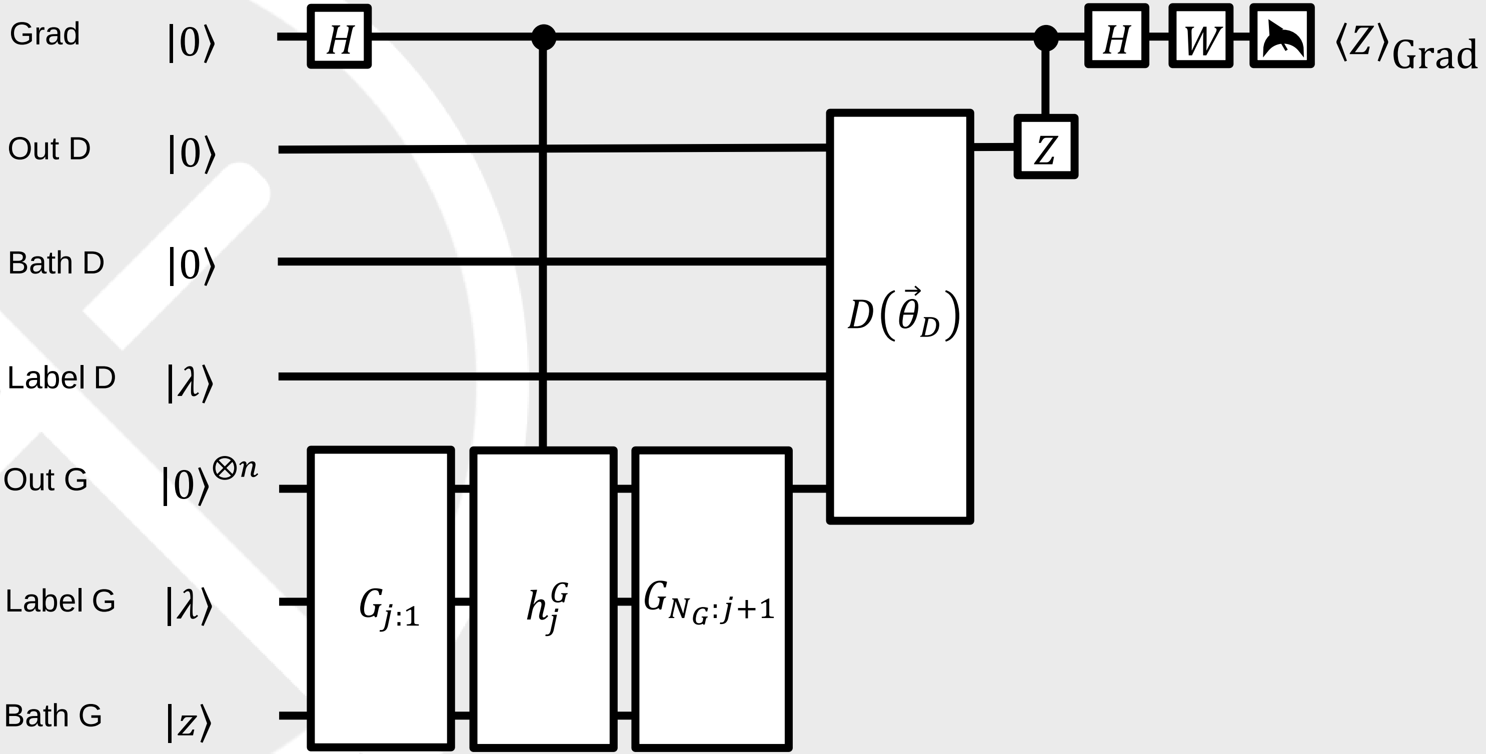
$$\langle Z\rangle_{\text{Grad}} = \frac{\partial}{\partial\theta_j}\langle P(\vec{\theta})\rangle$$

Derivative of derivative



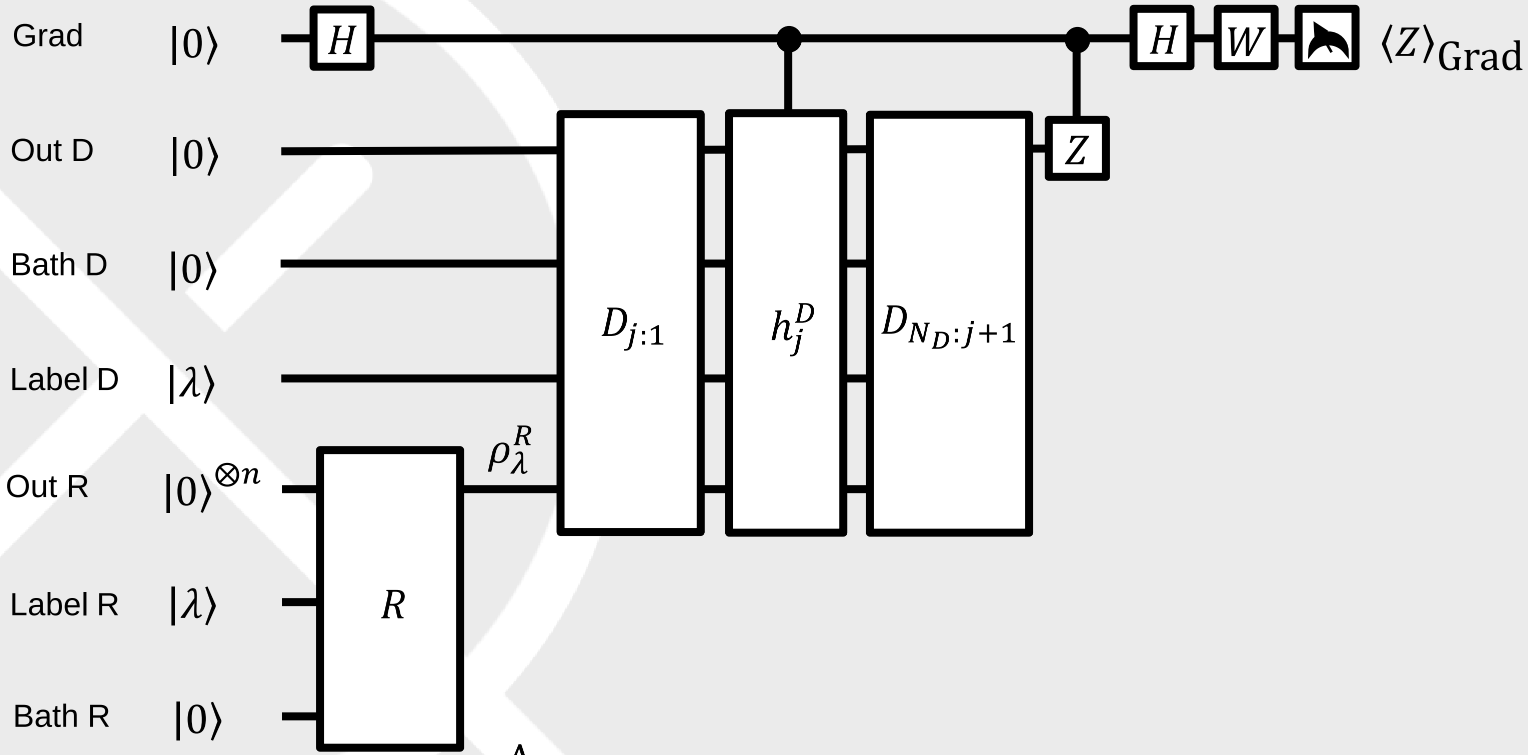
$$\langle Z\rangle_{\text{Hess}} = \frac{\partial}{\partial\theta_k}\frac{\partial}{\partial\theta_j}\langle P(\vec{\theta})\rangle$$

Gradient generator



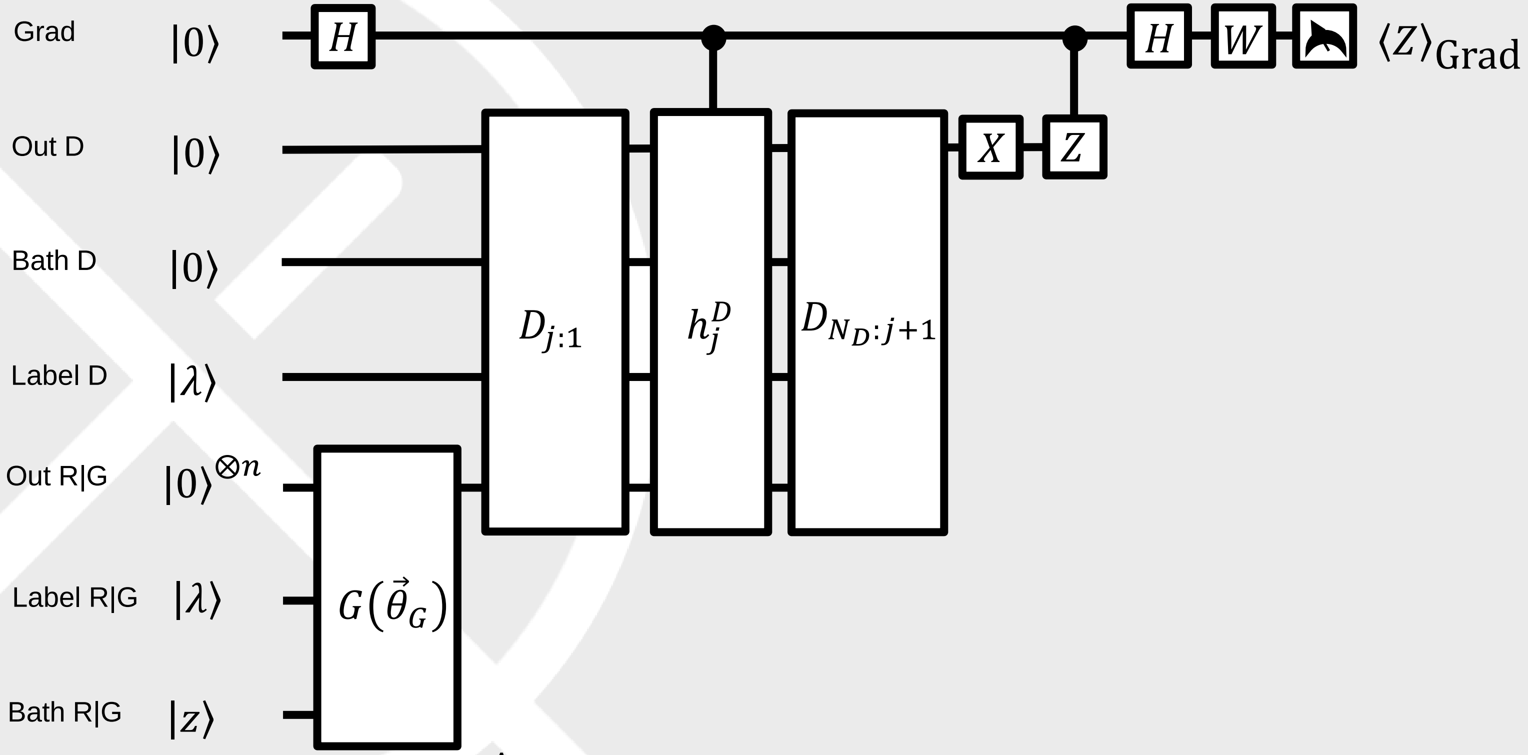
$$\frac{\partial}{\partial\theta_{Gj}}V(\vec{\theta}_D,\vec{\theta}_G)=\frac{i}{8\Lambda}\sum_{\lambda=1}^{\Lambda}\mathrm{tr}\left(\rho_\lambda^0(z)U_{G,1:j}^\dagger\left[U_{G,j+1:N_G}^\dagger U_D^\dagger(\vec{\theta}_D)ZU_D(\vec{\theta}_D)U_{G,N_G:j+1},h_j^G\right]U_{G,j:1}\right)$$

Gradient discriminator (R)



$$\frac{\partial}{\partial\theta_{Dj}}V(\vec{\theta}_D,\vec{\theta}_G)=-\frac{i}{8\Lambda}\sum_{\lambda=1}^{\Lambda}\mathrm{tr}\left(\rho_\lambda^R U_{D,1:j}^\dagger\left[U_{D,j+1:N_D}^\dagger Z U_{D,N_D:j+1},h_j^D\right]U_{D,j:1}\right)$$

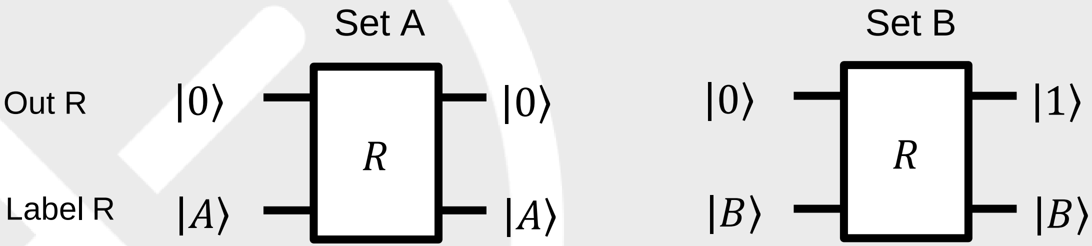
Gradient discriminator (G)

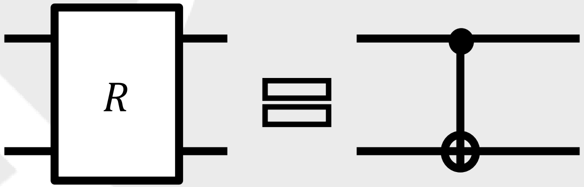


$$\frac{\partial}{\partial \theta_{Dj}} V(\vec{\theta}_D, \vec{\theta}_G) = \frac{i}{8\Lambda} \sum_{\lambda=1}^{\Lambda} \mathrm{tr}\left(\rho_\lambda^G(\vec{\theta}_G, z) U_{D,1:j}^\dagger \left[U_{D,j+1:N_D}^\dagger Z U_{D,N_D:j+1}, h_j^D\right] U_{D,j:1}\right)$$

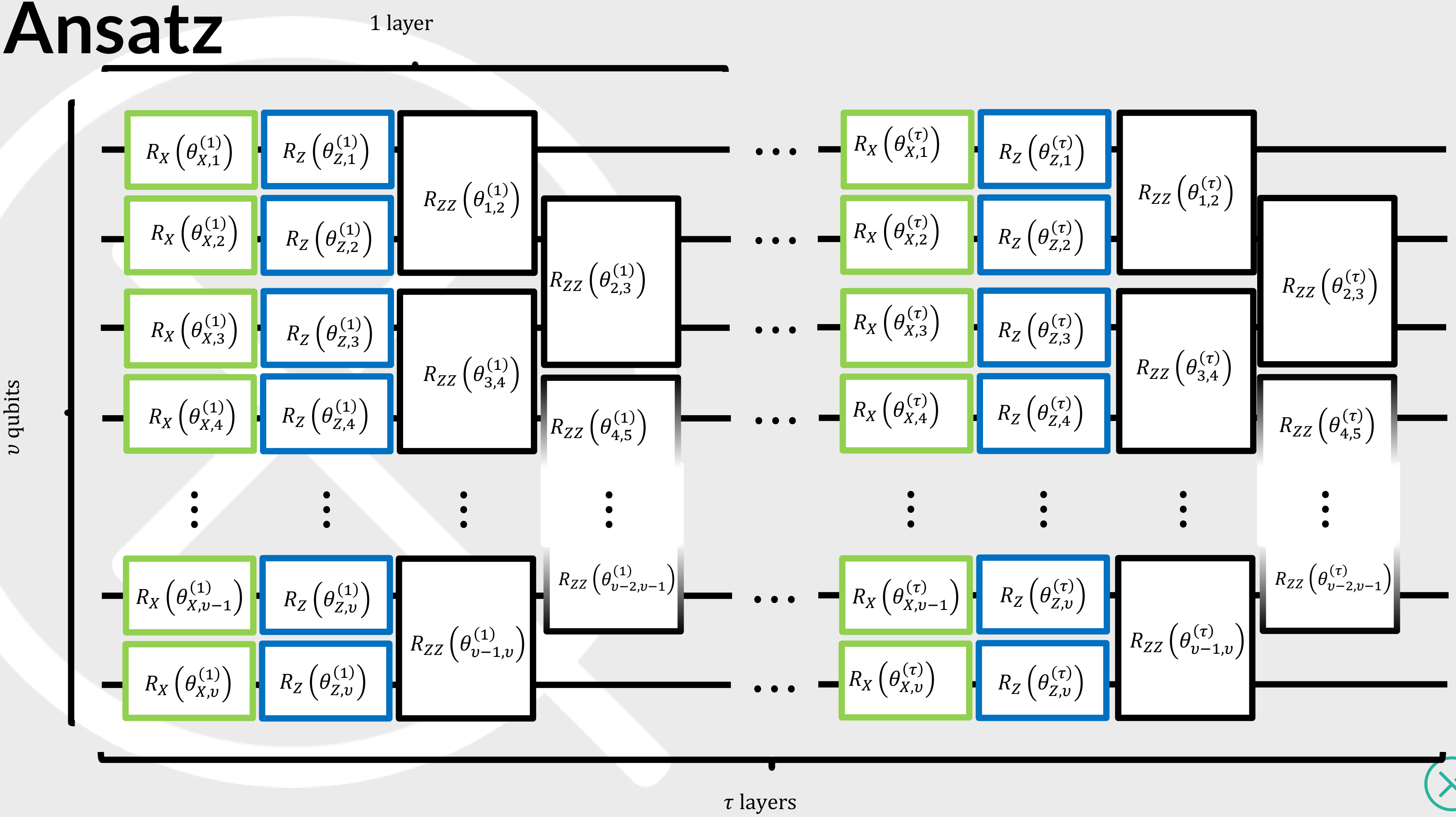
Numerics

Simplest example

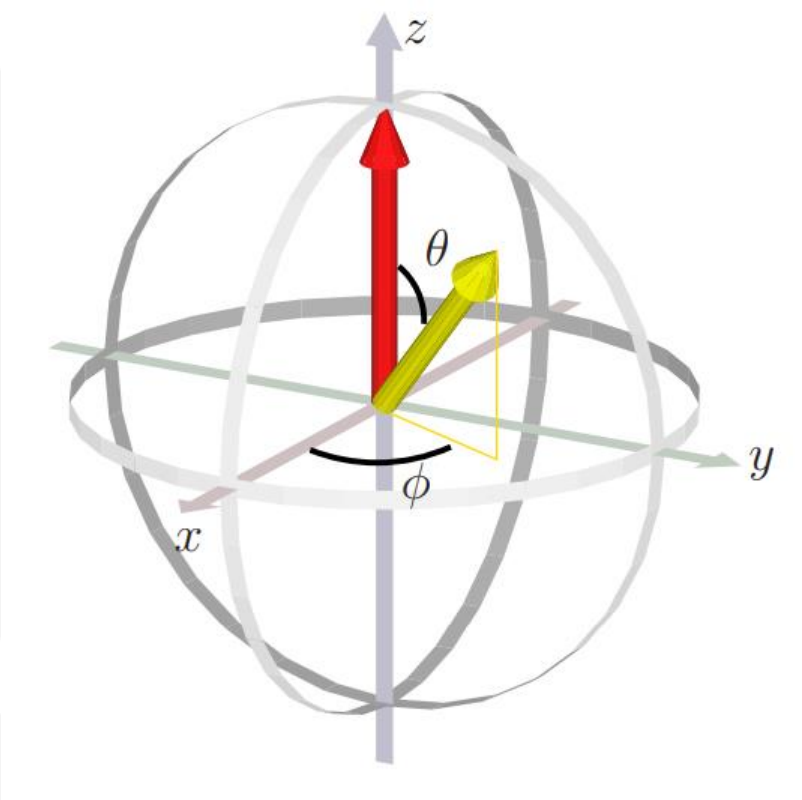


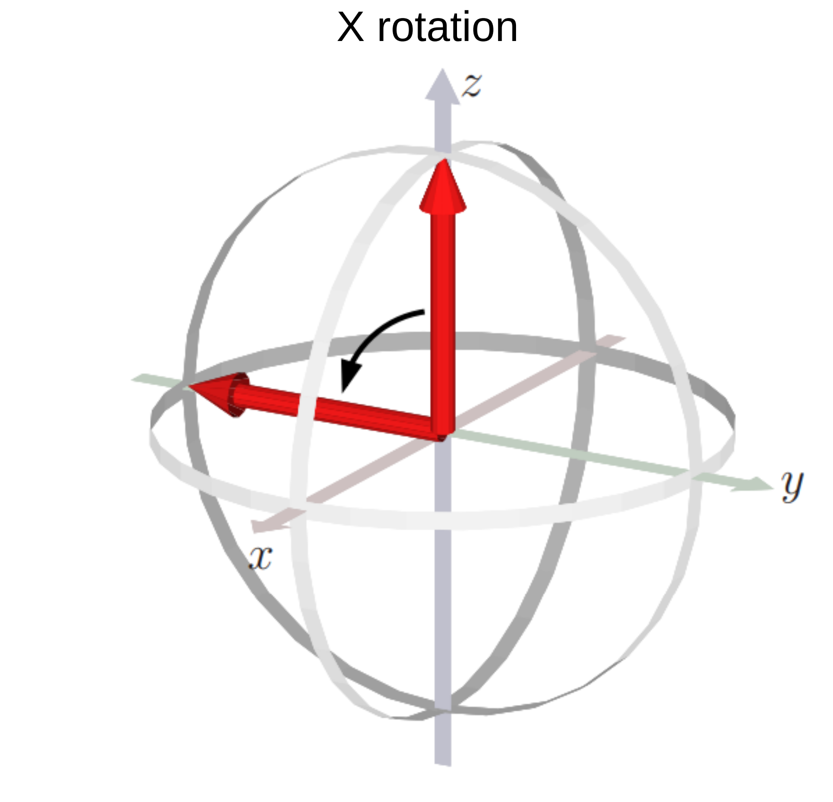


Ansatz


1 layer
υ qubits
τ layers
$R_X\left(\theta_{X,1}^{(1)}\right)$
$R_Z\left(\theta_{Z,1}^{(1)}\right)$
$R_{ZZ}\left(\theta_{1,2}^{(1)}\right)$
$R_X\left(\theta_{X,2}^{(1)}\right)$
$R_Z\left(\theta_{Z,2}^{(1)}\right)$
$R_{ZZ}\left(\theta_{2,3}^{(1)}\right)$
$R_X\left(\theta_{X,3}^{(1)}\right)$
$R_Z\left(\theta_{Z,3}^{(1)}\right)$
$R_{ZZ}\left(\theta_{3,4}^{(1)}\right)$
$R_X\left(\theta_{X,4}^{(1)}\right)$
$R_Z\left(\theta_{Z,4}^{(1)}\right)$
$R_{ZZ}\left(\theta_{4,5}^{(1)}\right)$
$R_{ZZ}\left(\theta_{\upsilon-2,\upsilon-1}^{(1)}\right)$
$R_X\left(\theta_{X,\upsilon-1}^{(1)}\right)$
$R_Z\left(\theta_{Z,\upsilon}^{(1)}\right)$
$R_{ZZ}\left(\theta_{\upsilon-1,\upsilon}^{(1)}\right)$
$R_X\left(\theta_{X,\upsilon}^{(1)}\right)$
$R_Z\left(\theta_{Z,\upsilon}^{(1)}\right)$
$R_X\left(\theta_{X,1}^{(\tau)}\right)$
$R_Z\left(\theta_{Z,1}^{(\tau)}\right)$
$R_{ZZ}\left(\theta_{1,2}^{(\tau)}\right)$
$R_X\left(\theta_{X,2}^{(\tau)}\right)$
$R_Z\left(\theta_{Z,2}^{(\tau)}\right)$
$R_{ZZ}\left(\theta_{2,3}^{(\tau)}\right)$
$R_X\left(\theta_{X,3}^{(\tau)}\right)$
$R_Z\left(\theta_{Z,3}^{(\tau)}\right)$
$R_{ZZ}\left(\theta_{3,4}^{(\tau)}\right)$
$R_X\left(\theta_{X,4}^{(\tau)}\right)$
$R_Z\left(\theta_{Z,4}^{(\tau)}\right)$
$R_{ZZ}\left(\theta_{4,5}^{(\tau)}\right)$
$R_{ZZ}\left(\theta_{\upsilon-2,\upsilon-1}^{(\tau)}\right)$
$R_X\left(\theta_{X,\upsilon-1}^{(\tau)}\right)$
$R_Z\left(\theta_{Z,\upsilon}^{(\tau)}\right)$
$R_{ZZ}\left(\theta_{\upsilon-1,\upsilon}^{(\tau)}\right)$
$R_X\left(\theta_{X,\upsilon}^{(\tau)}\right)$
$R_Z\left(\theta_{Z,\upsilon}^{(\tau)}\right)$

Single-qubit rotations

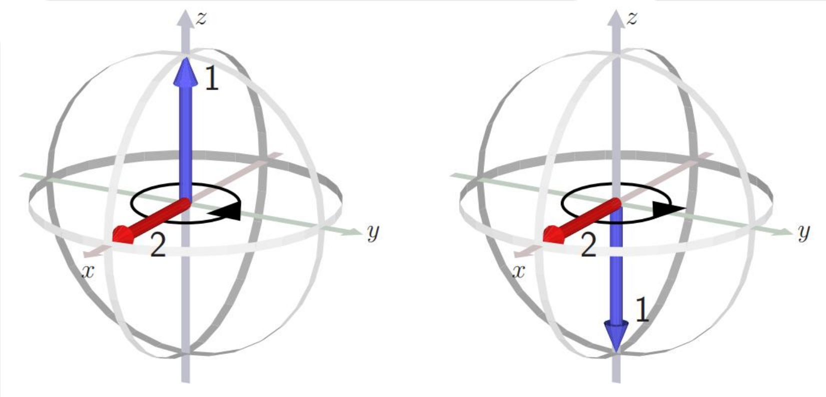




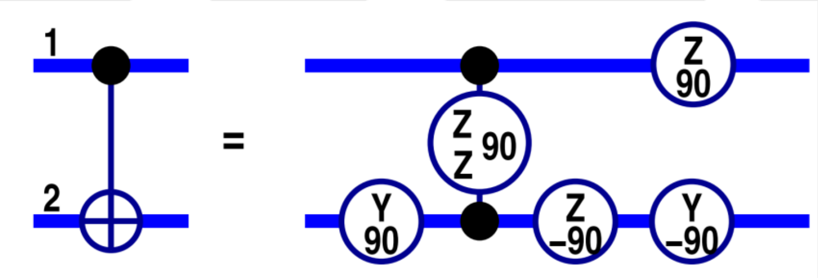
Laflamme, R., et al. "Introduction to NMR quantum information processing." arXiv: quant-ph/0207172 (2002).

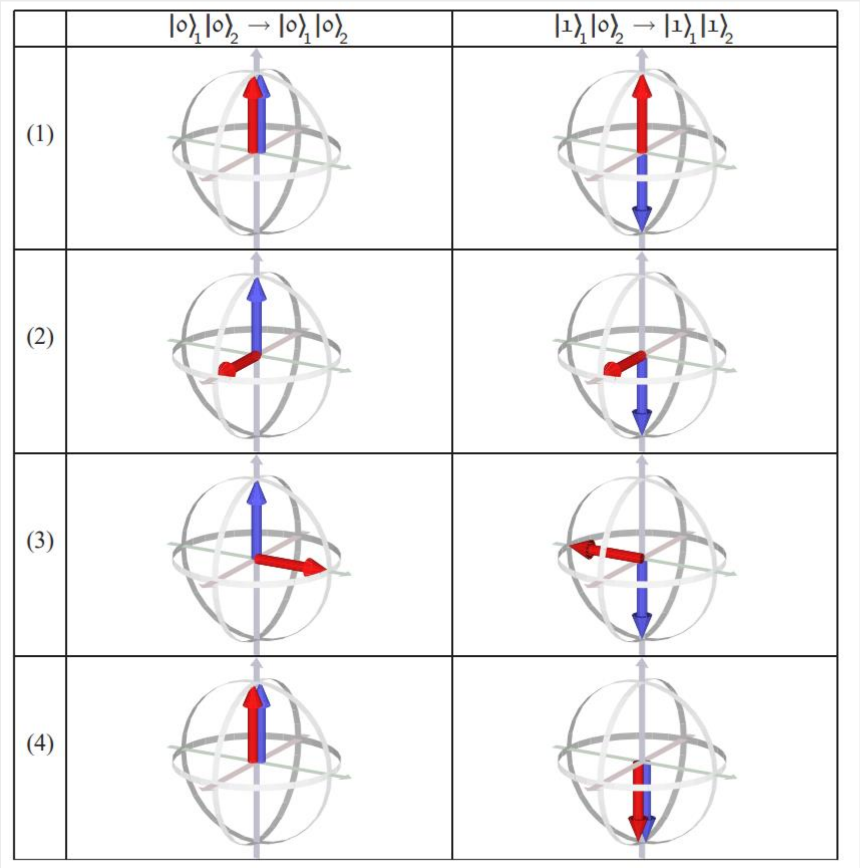
Two-qubit rotations

ZZ Interaction



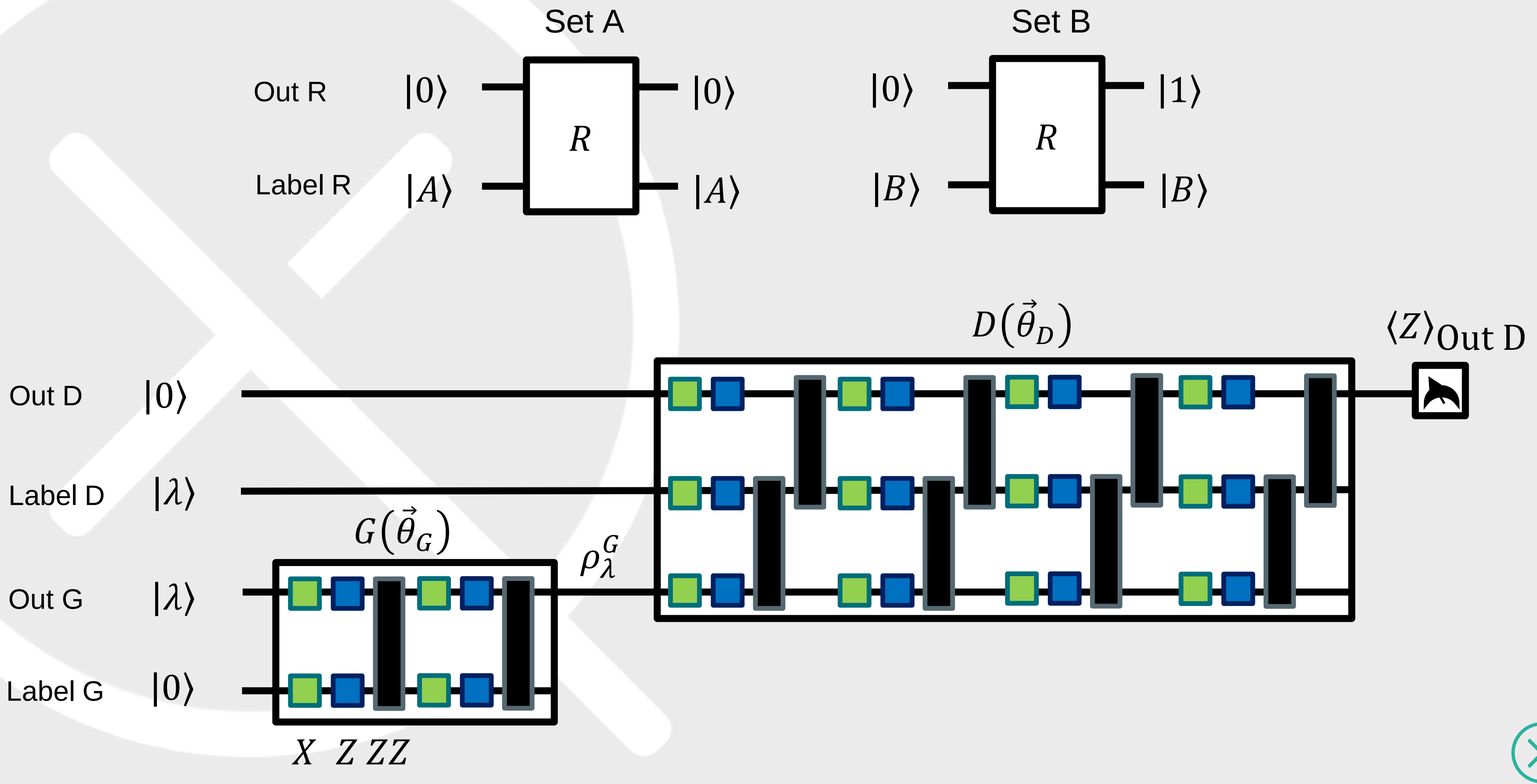
CNOT



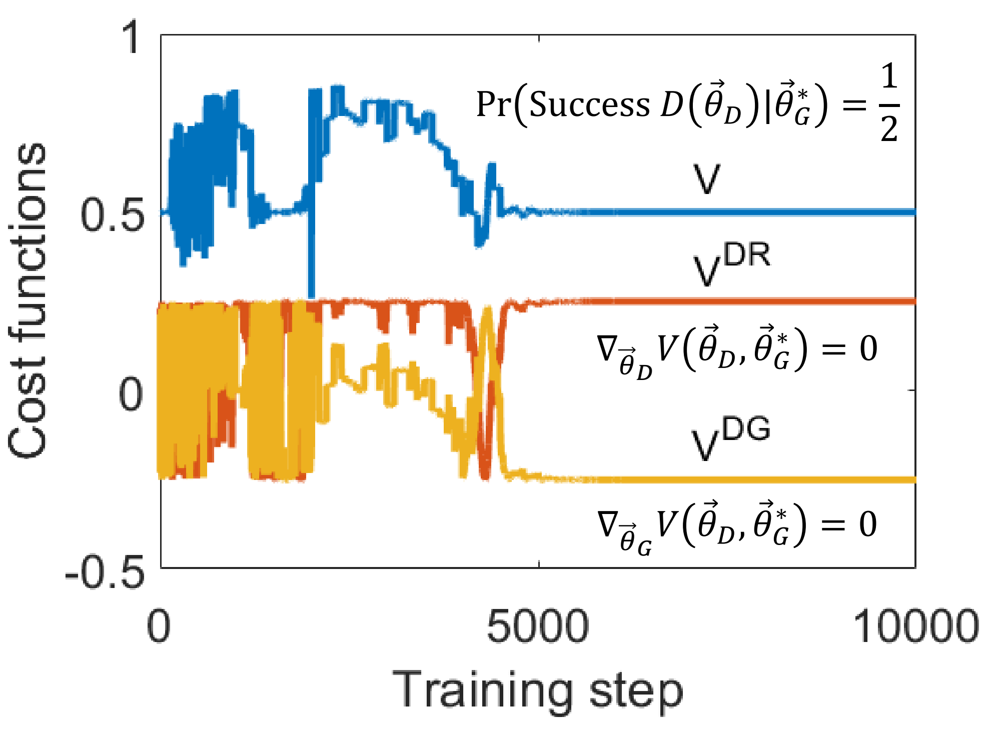


Laflamme, R., et al. "Introduction to NMR quantum information processing." arXiv: quant-ph/0207172 (2002).

Solving the simple example



Numerical training



$$V^{DR}(\vec{\theta}_D) = \frac{1}{4\Lambda} \sum_{\lambda=1}^{\Lambda} \mathrm{tr}\left(\rho_\lambda^{DR}(\vec{\theta}_D) Z\right)$$

$$V^{DG}(\vec{\theta}_D, \vec{\theta}_G) = -\frac{1}{4\Lambda} \sum_{\lambda=1}^{\Lambda} \mathrm{tr}\left(\rho_\lambda^{DG}(\vec{\theta}_D, \vec{\theta}_G, z) Z\right)$$

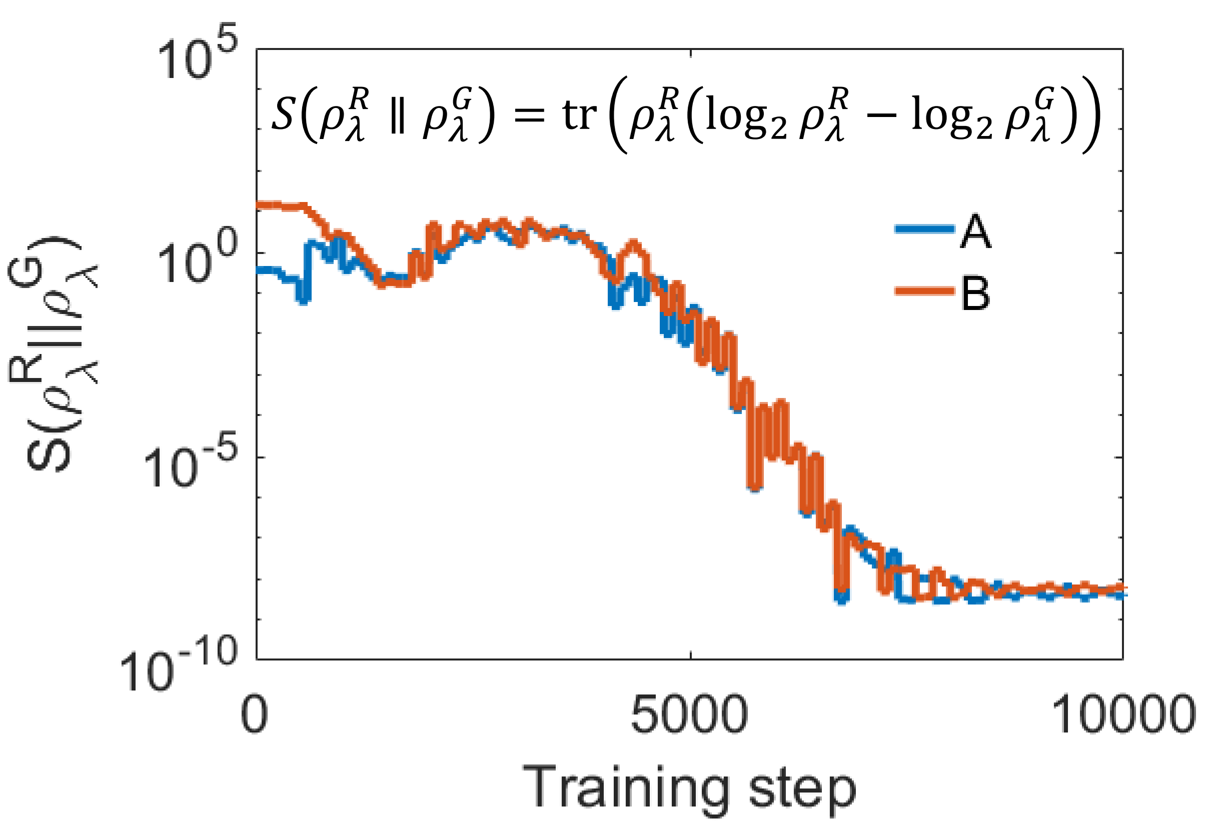
$$V(\vec{\theta}_D, \vec{\theta}_G) = \frac{1}{2} + V^{DR}(\vec{\theta}_D) + V^{DG}(\vec{\theta}_D, \vec{\theta}_G)$$

Performance analysis

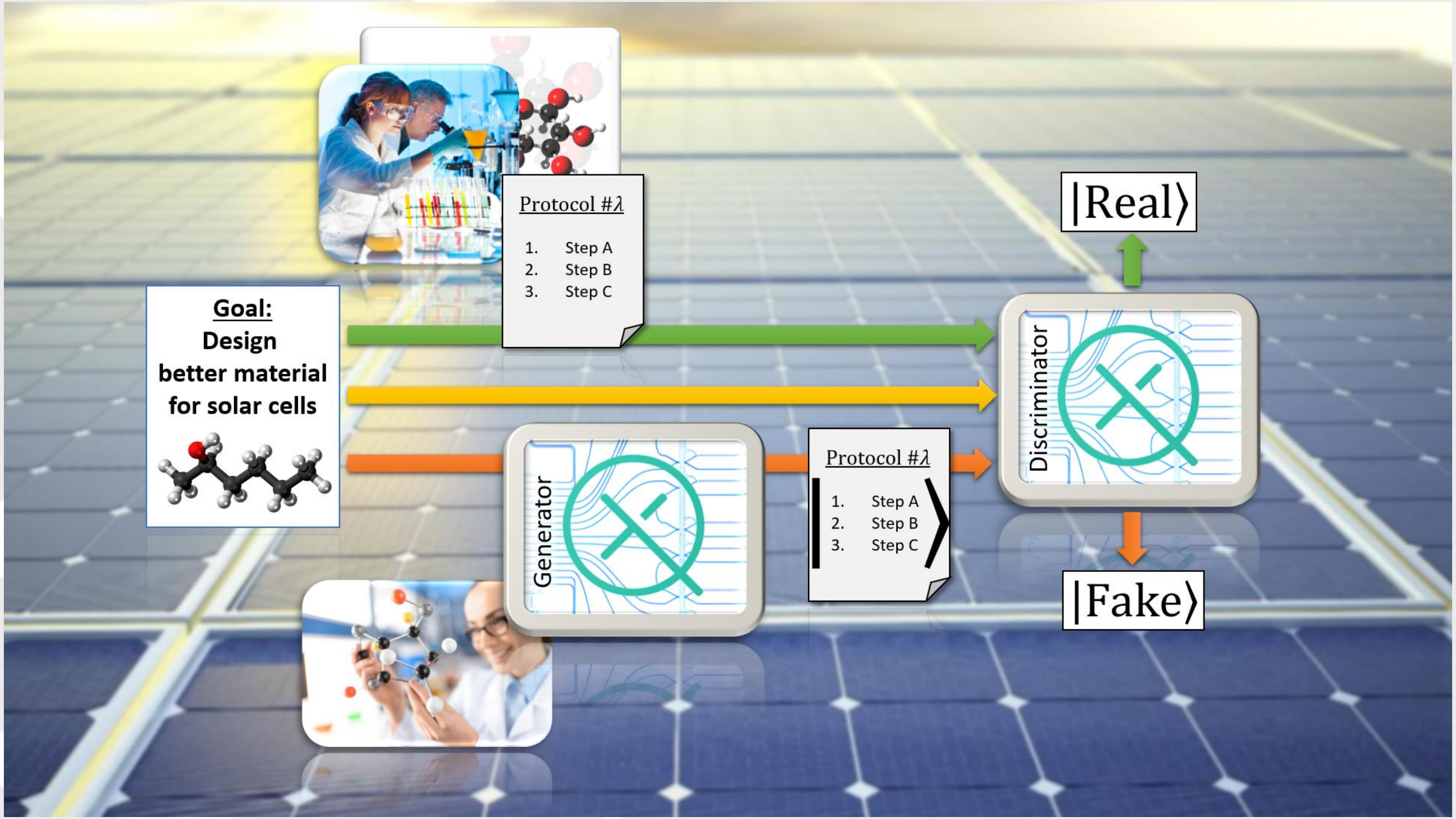
$$\frac{1}{2} C(\vec{\theta}_G) \leq \Pr\left(\text{Success } D(\vec{\theta}_D)|\vec{\theta}_G\right) \leq 1 - \frac{1}{2} C(\vec{\theta}_G)$$

$$r_{\min} \leq C(\vec{\theta}_G) \equiv \mathrm{tr}\left(\rho^R \rho^G(\vec{\theta}_G)\right) \leq \mathrm{tr}\left((\rho^R)^2\right)$$

Quantum cross-entropy



Outlook



Protocol #λ
1. Step A
2. Step B
3. Step C
Goal:
Design
better material
for solar cells
Generator
Protocol #λ
1. Step A
2. Step B
3. Step C
Discriminator
|Real⟩
|Fake⟩


Thank You
XANADU
Think With Light TM