Here comes the $SU(N)$: multivariate quantum gates and gradients

By Roeland Wiersema

QTML - November 2023

VECTOR INSTITUTE

XANADU

UNIVERSITY OF WATERLOO

Here comes the $SU(N)$: multivariate quantum gates and gradients

By Roeland Wiersema

QTML - November 2023

VECTOR INSTITUTE

XANADU

UNIVERSITY OF WATERLOO

Variational Quantum Computing

Choosing an Ansatz for a variational quantum algorithm is hard

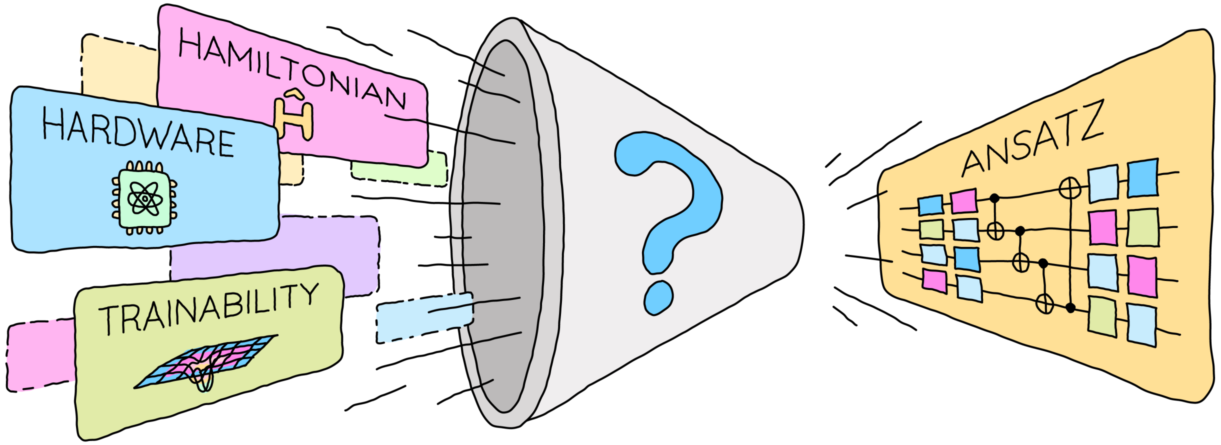


Image credit: Xanadu

Variational Quantum Computing

If you cannot find structure in the problem, one can pick a generic ansatz of this form:

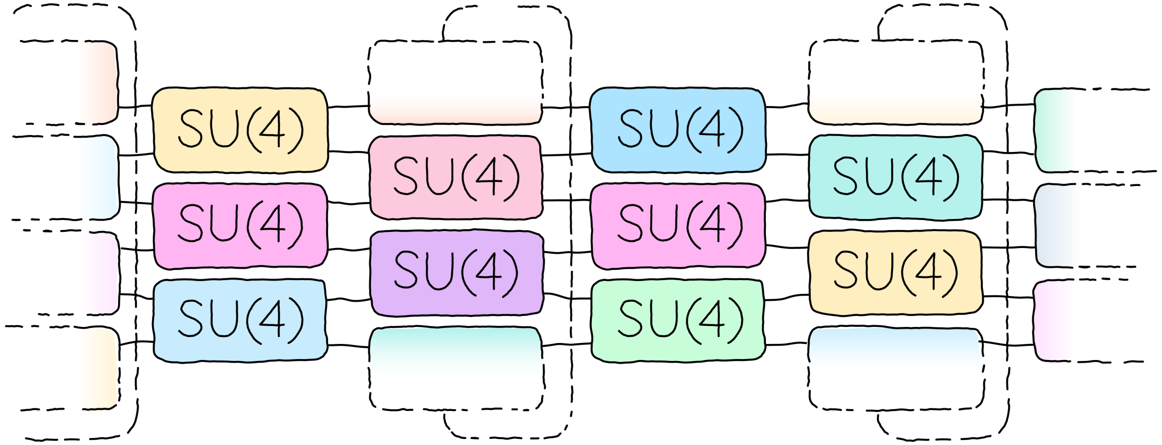
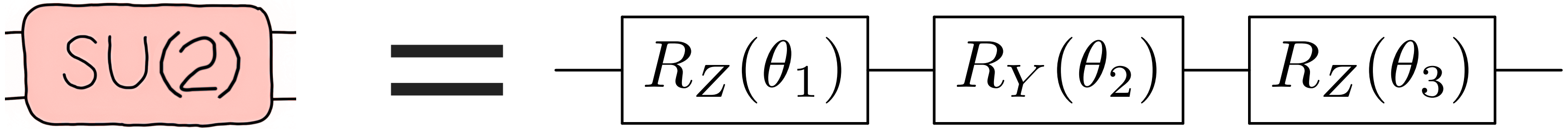


Image credit: Xanadu

General SU(N) parameterization

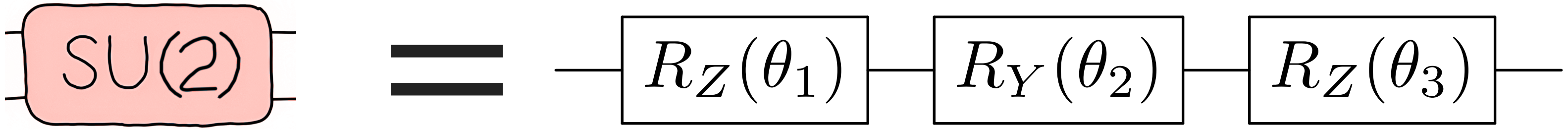


$$R_z(\theta) = \exp\{iZ\theta\}$$

Quantum gate

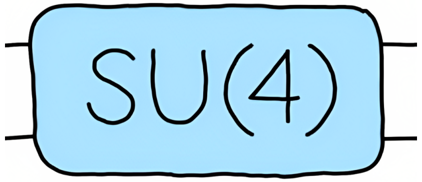
Vatan, F., & Williams, C. (2004). *Physical Review A*, ***69***(3), 032315.

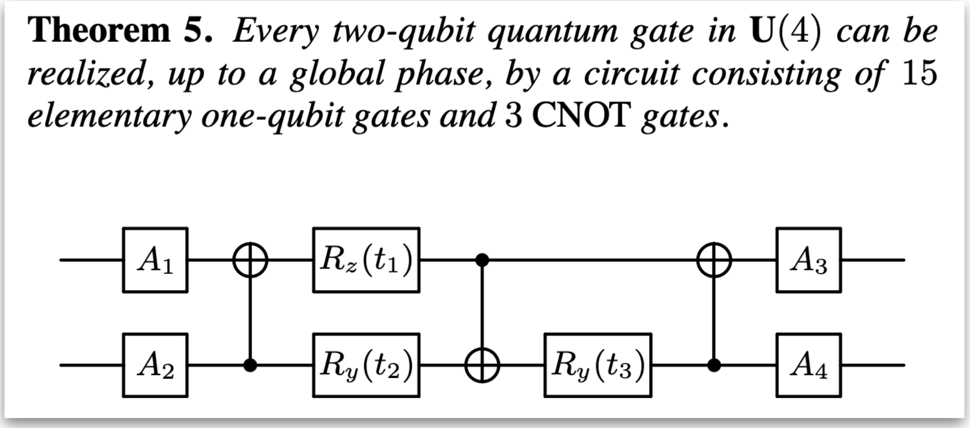
General SU(N) parameterization



$$R_z(\theta) = \exp\{iZ\theta\}$$

Quantum gate





Theorem 5. *Every two-qubit quantum gate in* $\mathbf{U}(4)$ *can be realized, up to a global phase, by a circuit consisting of* 15 *elementary one-qubit gates and* 3 CNOT *gates.*

Are there better parameterizations out there?

Vatan, F., & Williams, C. (2004). *Physical Review A*, ***69***(3), 032315.

General SU(N) parameterization

- Parameterize an element of the Lie algebra $\mathfrak{su}(N)$

$$A(\boldsymbol{\theta}) = \sum_{m}^{4^N-1} \theta_m G_m$$

$$U(\boldsymbol{\theta}) = e^{A(\boldsymbol{\theta})}$$

- With $\theta_m \in \mathbb{R}$ and G_m a Pauli string multiplied by the imaginary unit i.

Example 1

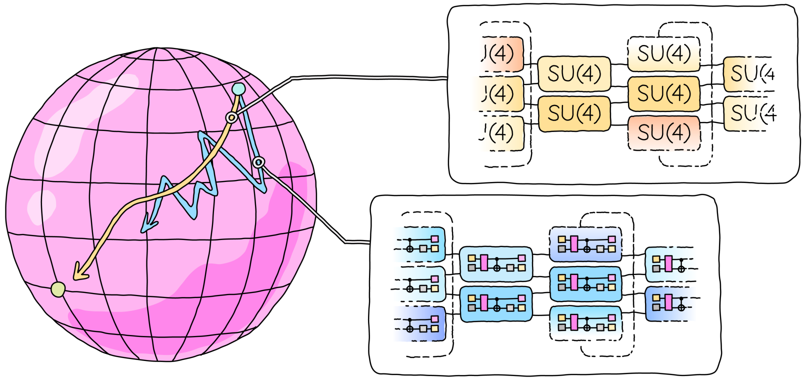
$$R_x(\theta) = e^{i\theta X}$$

Example 2

$$A(\theta_1, \theta_2) = e^{i(\theta_1 X + \theta_2 Y)}$$

Image credit: Xanadu

Gradients of SU(N) gates



$$\min_{\boldsymbol{\theta}} C(\boldsymbol{\theta}) = \min_{\boldsymbol{\theta}} \mathrm{Tr}\{U(\boldsymbol{\theta})\rho U^{\dagger}(\boldsymbol{\theta})H\}$$

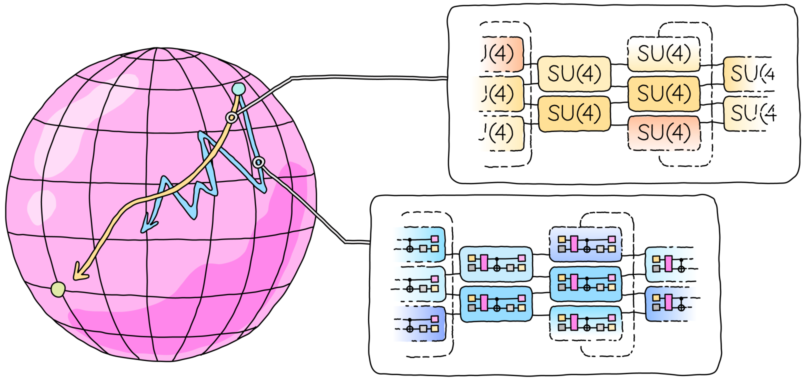
VQE Optimization

Schuld, M., Bergholm, V., Gogolin, C., Izaac, J., & Killoran, N. (2019). *Physical Review A*, 99(3), 032331.
Mitarai, K., & Fujii, K. (2021). *New Journal of Physics*, 23(2), 023021.
Kyriienko, O., & Elfving, V. E. (2021). *Physical Review A*, 104(5), 052417.
Wierichs, D., Izaac, J., Wang, C., & Lin, C. Y. Y. (2022). *Quantum,* 6, 677.
Izmaylov, A. F., Lang, R. A., & Yen, T. C. (2021). *Physical Review A*, *104*(6), 062443.

Gradients of SU(N) gates

Parameter-shift rule (P has two eigenvalues):

$$V(t) = e^{itP} \rightarrow \partial_t C(t) = \frac{1}{2}\left(C(t+\delta) - C(t-\delta)\right)$$



$$\min_{\boldsymbol{\theta}} C(\boldsymbol{\theta}) = \min_{\boldsymbol{\theta}} \mathrm{Tr}\{U(\boldsymbol{\theta})\rho U^{\dagger}(\boldsymbol{\theta})H\}$$

VQE Optimization

Schuld, M., Bergholm, V., Gogolin, C., Izaac, J., & Killoran, N. (2019). *Physical Review A*, 99(3), 032331.
Mitarai, K., & Fujii, K. (2021). *New Journal of Physics*, 23(2), 023021.
Kyriienko, O., & Elfving, V. E. (2021). *Physical Review A*, 104(5), 052417.
Wierichs, D., Izaac, J., Wang, C., & Lin, C. Y. Y. (2022). *Quantum,* 6, 677.
Izmaylov, A. F., Lang, R. A., & Yen, T. C. (2021). *Physical Review A*, *104*(6), 062443.

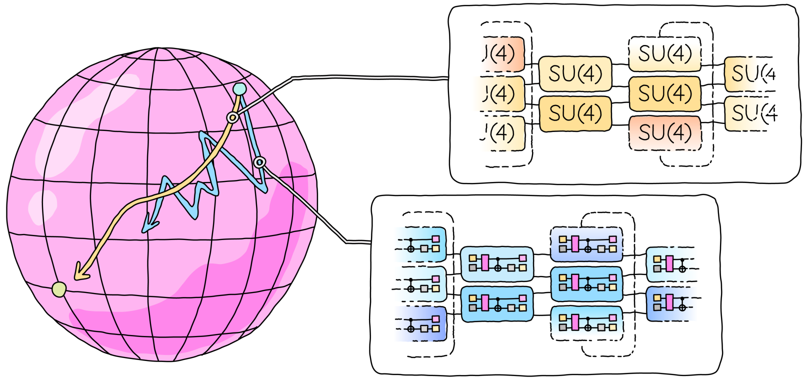
Gradients of SU(N) gates

Parameter-shift rule (P has two eigenvalues):

$$V(t) = e^{itP} \to \partial_t C(t) = \frac{1}{2}\left(C(t+\delta) - C(t-\delta)\right)$$

Generalized parameter-shift rules (H Hermitian, R unique evs.):

$$W(t) = e^{itH} \to \partial_t C(t) = \sum_{i=1}^{R} \Delta_i \left(C(t+\delta_i) - C(t-\delta_i)\right)$$



$$\min_{\boldsymbol{\theta}} C(\boldsymbol{\theta}) = \min_{\boldsymbol{\theta}} \mathrm{Tr}\{U(\boldsymbol{\theta})\rho U^\dagger(\boldsymbol{\theta})H\}$$

VQE Optimization

Schuld, M., Bergholm, V., Gogolin, C., Izaac, J., & Killoran, N. (2019). *Physical Review A*, 99(3), 032331.
Mitarai, K., & Fujii, K. (2021). *New Journal of Physics*, 23(2), 023021.
Kyriienko, O., & Elfving, V. E. (2021). *Physical Review A*, 104(5), 052417.
Wierichs, D., Izaac, J., Wang, C., & Lin, C. Y. Y. (2022). *Quantum,* 6, 677.
Izmaylov, A. F., Lang, R. A., & Yen, T. C. (2021). *Physical Review A*, *104*(6), 062443.

Gradients of SU(N) gates

Parameter-shift rule (P has two eigenvalues):

$$V(t) = e^{itP} \rightarrow \partial_t C(t) = \frac{1}{2}\big(C(t+\delta) - C(t-\delta)\big)$$

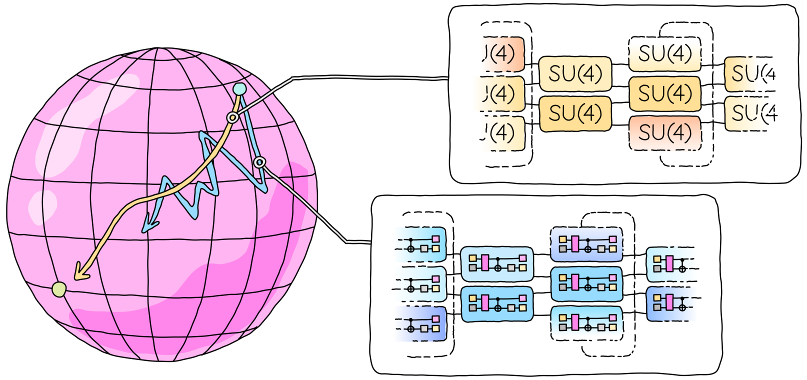
Generalized parameter-shift rules (H Hermitian, R unique evs.):

$$W(t) = e^{itH} \rightarrow \partial_t C(t) = \sum_{i=1}^{R} \Delta_i \big(C(t+\delta_i) - C(t-\delta_i)\big)$$

Gates generated by non-commuting operators:

$$U(\theta_1, \theta_2) = e^{i(\theta_1 X + \theta_2 Y)} \rightarrow \partial_{\theta_1} C(\theta_1, \theta_2) = ???$$

Our work



$$\min_{\boldsymbol{\theta}} C(\boldsymbol{\theta}) = \min_{\boldsymbol{\theta}} \mathrm{Tr}\{U(\boldsymbol{\theta})\rho U^\dagger(\boldsymbol{\theta})H\}$$

VQE Optimization

Schuld, M., Bergholm, V., Gogolin, C., Izaac, J., & Killoran, N. (2019). *Physical Review A*, 99(3), 032331.
Mitarai, K., & Fujii, K. (2021). *New Journal of Physics*, 23(2), 023021.
Kyriienko, O., & Elfving, V. E. (2021). *Physical Review A*, 104(5), 052417.
Wierichs, D., Izaac, J., Wang, C., & Lin, C. Y. Y. (2022). *Quantum,* 6, 677.
Izmaylov, A. F., Lang, R. A., & Yen, T. C. (2021). *Physical Review A*, *104*(6), 062443.

Gradients of SU(N) gates

- Analytically: an infinite series

$$\partial_{\theta_1} U(\theta_1, \theta_2) = U(\theta_1, \theta_2) \sum_{p=0}^{\infty} \frac{(-1)^p}{(p+1)!} (\mathrm{ad}_{\theta_1 X + \theta_2 Y})^p X = \underbrace{U(\theta_1, \theta_2)}_{\mathrm{SU}(N)} \underbrace{\Omega_{\theta_1}(\theta_1, \theta_2)}_{\mathfrak{su}(N)}$$

Gradients of SU(N) gates

- Analytically: an infinite series

$$\partial_{\theta_1} U(\theta_1, \theta_2) = U(\theta_1, \theta_2) \sum_{p=0}^{\infty} \frac{(-1)^p}{(p+1)!} (\mathrm{ad}_{\theta_1 X + \theta_2 Y})^p X = \underbrace{U(\theta_1, \theta_2)}_{\mathrm{SU}(N)} \underbrace{\Omega_{\theta_1}(\theta_1, \theta_2)}_{\mathfrak{su}(N)}$$

- Numerically: matrix derivative via automatic differentiation

$$\longrightarrow \partial_{\theta_i} \exp A(\theta) \longrightarrow U^\dagger(\theta) \partial_{\theta_i} \exp A(\theta) = \Omega_{\theta_i}(\theta)$$

Gradients of SU(N) gates

- The gradient calculation then becomes:

$$\frac{\partial}{\partial \theta_l} C(\boldsymbol{\theta}) = \mathrm{Tr}\left\{ \left(\frac{\partial}{\partial \theta_l} U(\boldsymbol{\theta}) \right) \rho U^\dagger(\boldsymbol{\theta}) H \right\} + \mathrm{h.c.}$$

Gradients of SU(N) gates

- The gradient calculation then becomes:

$$\frac{\partial}{\partial \theta_l} C(\boldsymbol{\theta}) = \mathrm{Tr}\left\{ \left(\frac{\partial}{\partial \theta_l} U(\boldsymbol{\theta}) \right) \rho U^\dagger(\boldsymbol{\theta}) H \right\} + \mathrm{h.c.}$$

$$\frac{\partial}{\partial \theta_l} C(\boldsymbol{\theta}) = \mathrm{Tr}\left\{ U(\boldsymbol{\theta}) \Omega_l(\boldsymbol{\theta}) \rho U^\dagger(\boldsymbol{\theta}) H \right\} + \mathrm{h.c.}$$

$$\partial_{\theta_1} U(\theta_1, \theta_2) = \underbrace{U(\theta_1, \theta_2)}_{\mathrm{SU}(N)} \underbrace{\Omega_{\theta_1}(\theta_1, \theta_2)}_{\mathfrak{su}(N)}$$

Calculate matrix derivative $\partial_{\theta_l} U(\boldsymbol{\theta})$ and obtain $\Omega_l(\boldsymbol{\theta})$ by multiplying $\partial_{\theta_l} U(\boldsymbol{\theta})$ on the **left** with $U^\dagger(\boldsymbol{\theta})$

Gradients of SU(N) gates

- The gradient calculation then becomes:

$$\partial_{\theta_1} U(\theta_1, \theta_2) = \underbrace{U(\theta_1, \theta_2)}_{\mathrm{SU}(N)} \underbrace{\Omega_{\theta_1}(\theta_1, \theta_2)}_{\mathfrak{su}(N)}$$

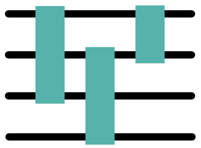
$$\frac{\partial}{\partial \theta_l} C(\boldsymbol{\theta}) = \mathrm{Tr}\left\{ \left(\frac{\partial}{\partial \theta_l} U(\boldsymbol{\theta}) \right) \rho U^\dagger(\boldsymbol{\theta}) H \right\} + \mathrm{h.c.}$$

Calculate matrix derivative $\partial_{\theta_l} U(\boldsymbol{\theta})$ and obtain $\Omega_l(\boldsymbol{\theta})$ by multiplying $\partial_{\theta_l} U(\boldsymbol{\theta})$ on the **left** with $U^\dagger(\boldsymbol{\theta})$

$$\frac{\partial}{\partial \theta_l} C(\boldsymbol{\theta}) = \mathrm{Tr}\left\{ U(\boldsymbol{\theta}) \Omega_l(\boldsymbol{\theta}) \rho U^\dagger(\boldsymbol{\theta}) H \right\} + \mathrm{h.c.}$$

Add the gate generated by $\Omega_l(\boldsymbol{\theta})$ to the circuit and calculate the **total derivative** of the new gate with a standard parameter-shift rule on quantum hardware

$$\frac{\partial}{\partial \theta_l} C(\boldsymbol{\theta}) = \frac{d}{dt} \mathrm{Tr}\left\{ U(\boldsymbol{\theta}) e^{t\Omega_l(\boldsymbol{\theta})} \rho U^\dagger(\boldsymbol{\theta}) H \right\} + \mathrm{h.c.} \Big|_{t=0}$$



Gradients of SU(N) gates

Mari, A., Bromley, T. R., & Killoran, N. (2021). *Physical Review A*, *103*(1), 012405.
Banchi, L., & Crooks, G. E. (2021). *Quantum*, *5*, 386.

Gradients of SU(N) gates

Finite difference

$$\partial_{\mathrm{FD},\theta_j} C(\boldsymbol{\theta}) = \frac{1}{\delta}\left[C\left(\boldsymbol{\theta} + \frac{\delta}{2}\boldsymbol{e}_j\right) - C\left(\boldsymbol{\theta} - \frac{\delta}{2}\boldsymbol{e}_j\right)\right]$$

- Pro: simple to implement
- Con: biased estimator for the gradient

Mari, A., Bromley, T. R., & Killoran, N. (2021). *Physical Review A*, *103*(1), 012405.
Banchi, L., & Crooks, G. E. (2021). *Quantum*, *5*, 386.

Gradients of SU(N) gates

Finite difference

$$\partial_{\mathrm{FD},\theta_j} C(\boldsymbol{\theta}) = \frac{1}{\delta}\left[C\left(\boldsymbol{\theta} + \frac{\delta}{2}\boldsymbol{e}_j\right) - C\left(\boldsymbol{\theta} - \frac{\delta}{2}\boldsymbol{e}_j\right)\right]$$

- Pro: simple to implement
- Con: biased estimator for the gradient

Stochastic parameter-shift

$$\partial_{\mathrm{SPS},\theta_j} C(\boldsymbol{\theta}) = \int_0^1 ds\,\left(C_+(\boldsymbol{\theta}, s) - C_-(\boldsymbol{\theta}, s)\right)$$

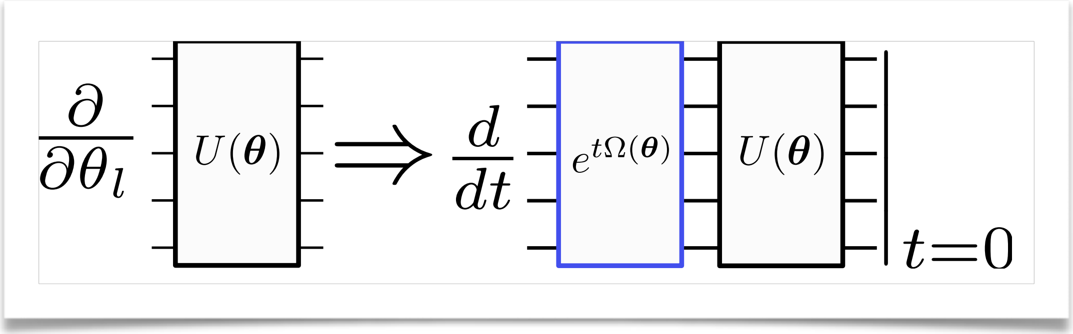
- Pro: unbiased estimator
- Con: requires Monte Carlo estimate of integral

Mari, A., Bromley, T. R., & Killoran, N. (2021). *Physical Review A*, *103*(1), 012405.
Banchi, L., & Crooks, G. E. (2021). *Quantum*, *5*, 386.

Gradients of SU(N) gates

$\mathrm{SU}(N)$ gate gradients

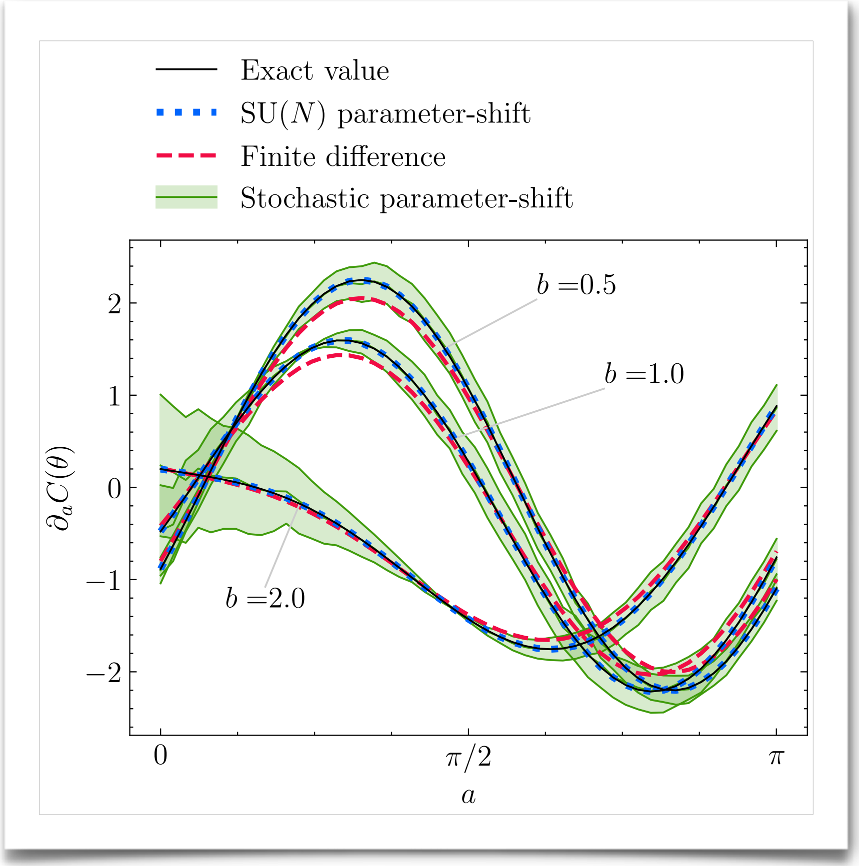
- Unbiased estimator
- Easy to implement
- Sample complexity is the same as the generalized parameter-shift rule



$$\frac{\partial}{\partial \theta_l} C(\boldsymbol{\theta}) = \frac{d}{dt} \mathrm{Tr}\left\{ U(\boldsymbol{\theta}) e^{t\Omega_l(\boldsymbol{\theta})} \rho U^\dagger(\boldsymbol{\theta}) H \right\} + \mathrm{h.c.} \Big|_{t=0}$$

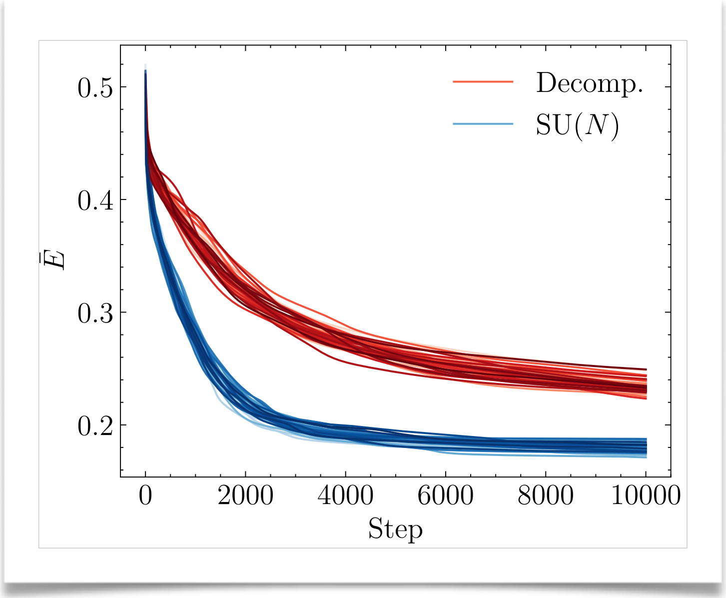
Toy example

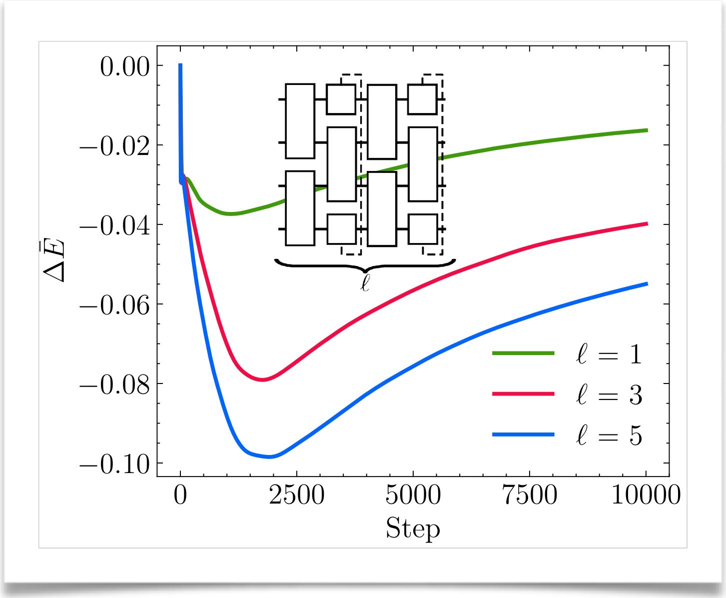
- We consider a random single qubit Hamiltonian H
- The generator of our gate is given by $A(\boldsymbol{\theta}) = iaX + ibY$
- Then, we calculate the partial derivative with respect to a



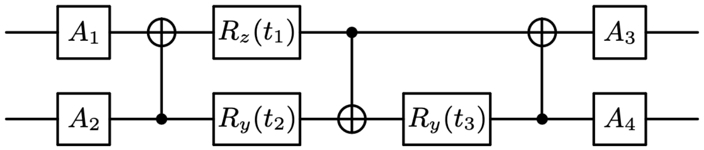
Numerical study

- We consider 100 random 10-qubit Hamiltonians with a bricklayer ansatz

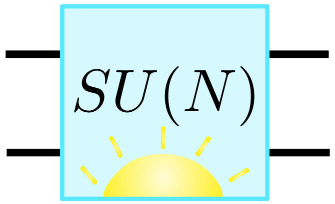




$\ell = 5$ circuit



V.S.



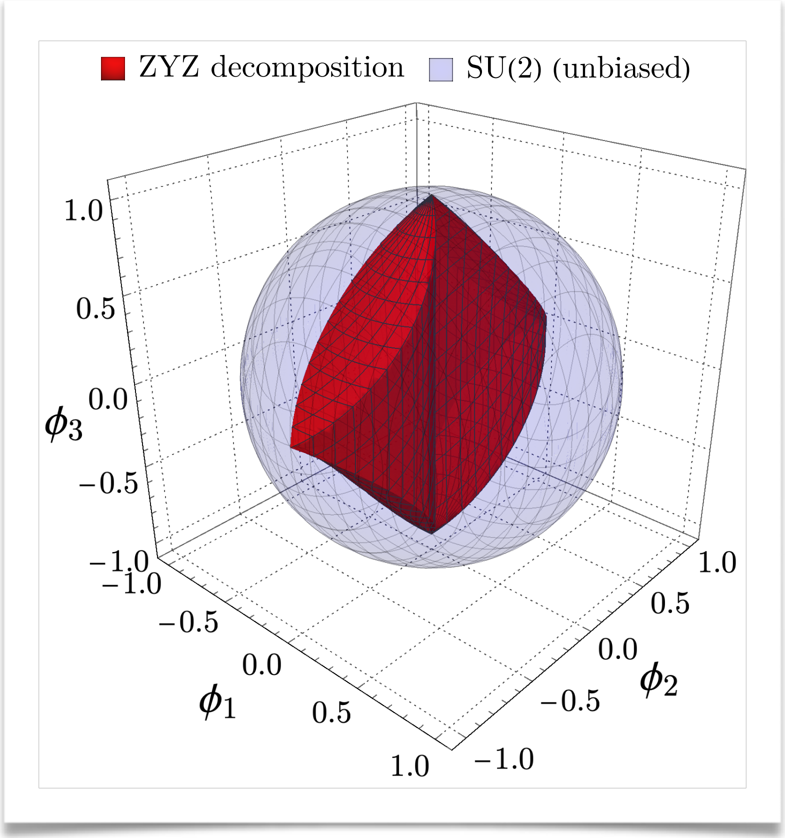
Uniform parameterization of the group

- We call a product of gates that parameterizes $\mathrm{SU}(N)$ a *decomposed* gate

$$V(\boldsymbol{\theta}) = \text{—}\boxed{R_Z(\theta_1)}\text{—}\boxed{R_Y(\theta_2)}\text{—}\boxed{R_Z(\theta_3)}\text{—}$$

- If we normalize $\boldsymbol{\theta}$, we can write this as in our parametrization as

$$V(\boldsymbol{\theta}) = \exp\left\{\phi_1(\boldsymbol{\theta})X + \phi_2(\boldsymbol{\theta})Y + \phi_3(\boldsymbol{\theta})Z\right\}$$

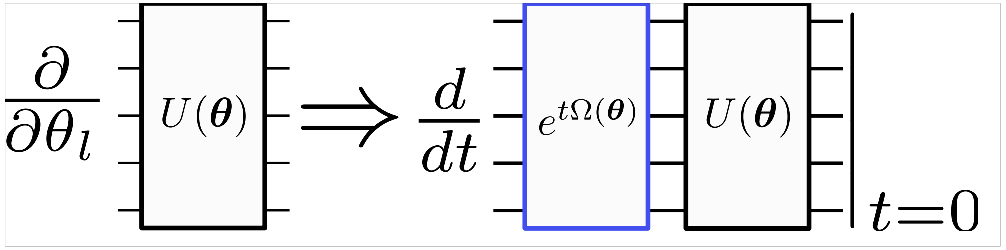


Resource estimation via the dynamical Lie algebra of the gate

- The dynamical Lie algebra is given by $\mathfrak{g} = \text{span}\{[\,..[G_k, [G_i, G_j]]]\}$
- The effective generator $\Omega(\theta) \in \mathfrak{g}$
- The number of circuit evaluations depends on the number of unique eigenvalues, which is related to the structure of $\mathfrak{g}$.

$$A(\boldsymbol{\theta}) = \sum_{m}^{4^N-1} \theta_m G_m$$

$$U(\boldsymbol{\theta}) = e^{A(\boldsymbol{\theta})}$$



Conclusion

- We proposed a new parameterization for general $\mathrm{SU}(N)$ gates and provide a method to obtain the gradients of this gate on hardware
- Leveraging auto-differentiation in different ways could motivate new variational algorithms
- Instead of focusing on expressivity of ansätze we should try to understand **how** we are parameterizing circuits

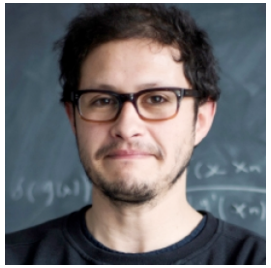
Paper

Link to demo

Dylan Lewis

David Wierichs

Juan Carrasquilla



Nathan Killoran