# Estimates of loss function concentration in noisy parametrized quantum circuits

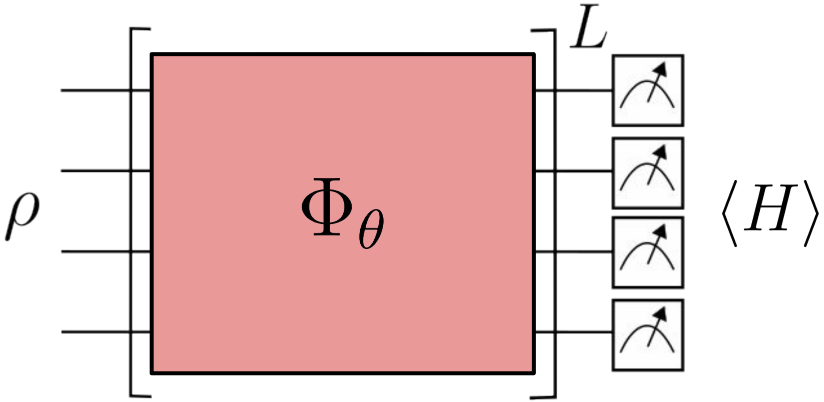
(arXiv:2410.01893)

**Giulio Crognaletti**, Michele Grossi and Angelo Bassi

**Quantum Computing@INFN, October 29th-31st 2024**

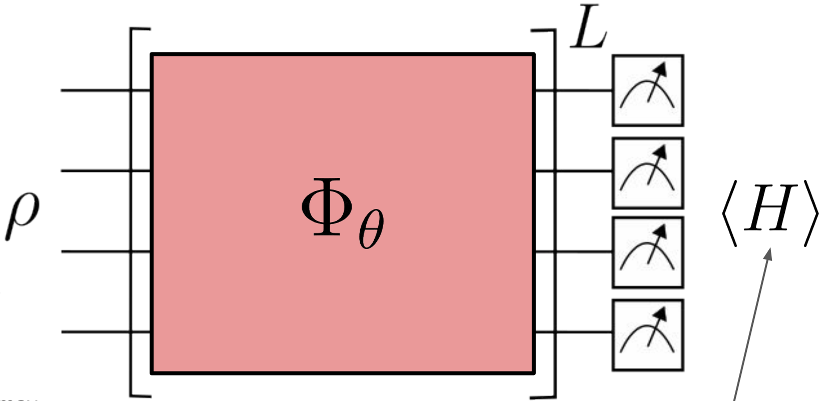
# Variational Quantum Algorithms in a nutshell

Variational Quantum Algorithms are based on a **hybrid** quantum classical optimization scheme.



# Variational Quantum Algorithms in a nutshell

Variational Quantum Algorithms are based on a **hybrid** quantum classical optimization scheme.




Initial state, may be a *reference state* or encode *quantum data*

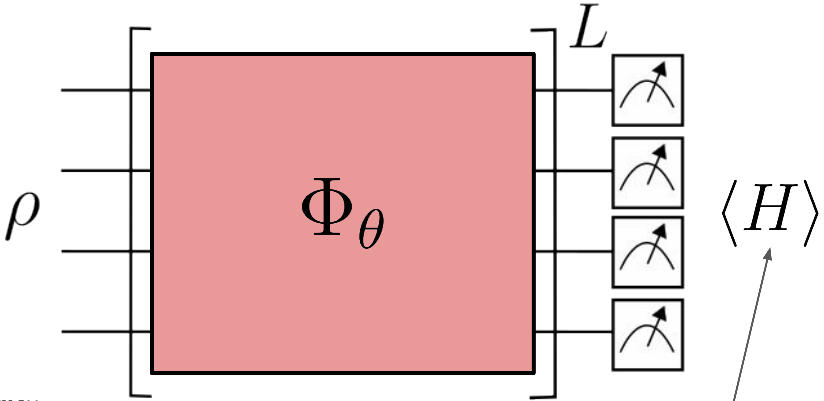
Parameterized quantum channel

An observable, generally the Hamiltonian for applications to condensed matter physics.

(generally referred to as loss operator in quantum ML)

# Variational Quantum Algorithms in a nutshell

Variational Quantum Algorithms are based on a **hybrid** quantum classical optimization scheme.



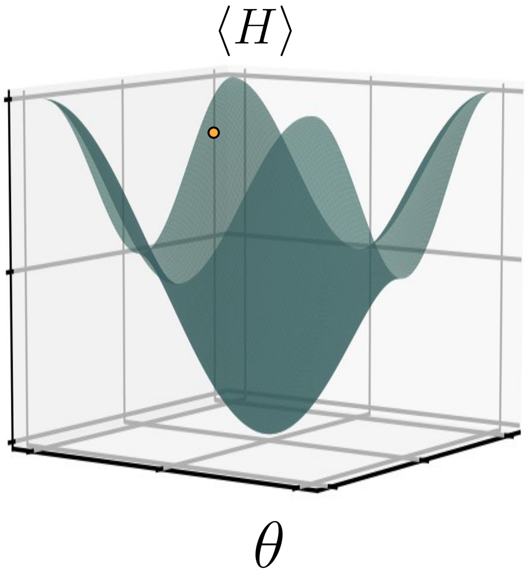

Initial state, may be a *reference state* or encode *quantum data*

Parameterized quantum channel

An observable, generally the Hamiltonian for applications to condensed matter physics.

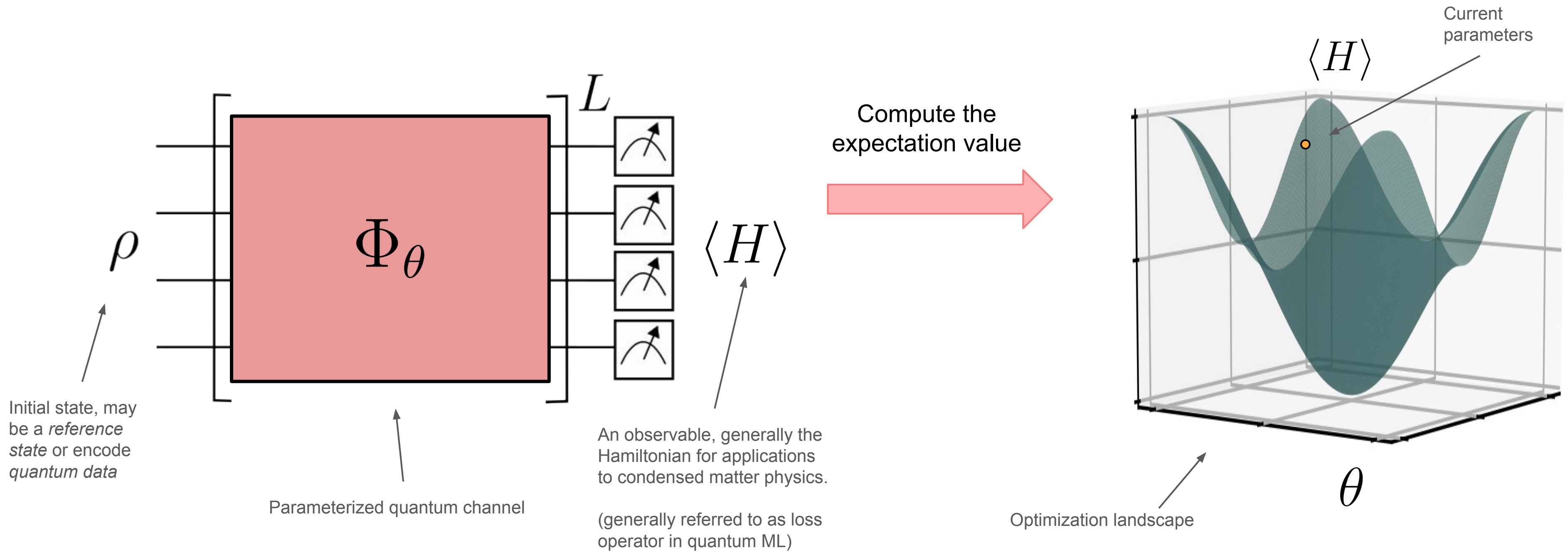
(generally referred to as loss operator in quantum ML)

Compute the expectation value



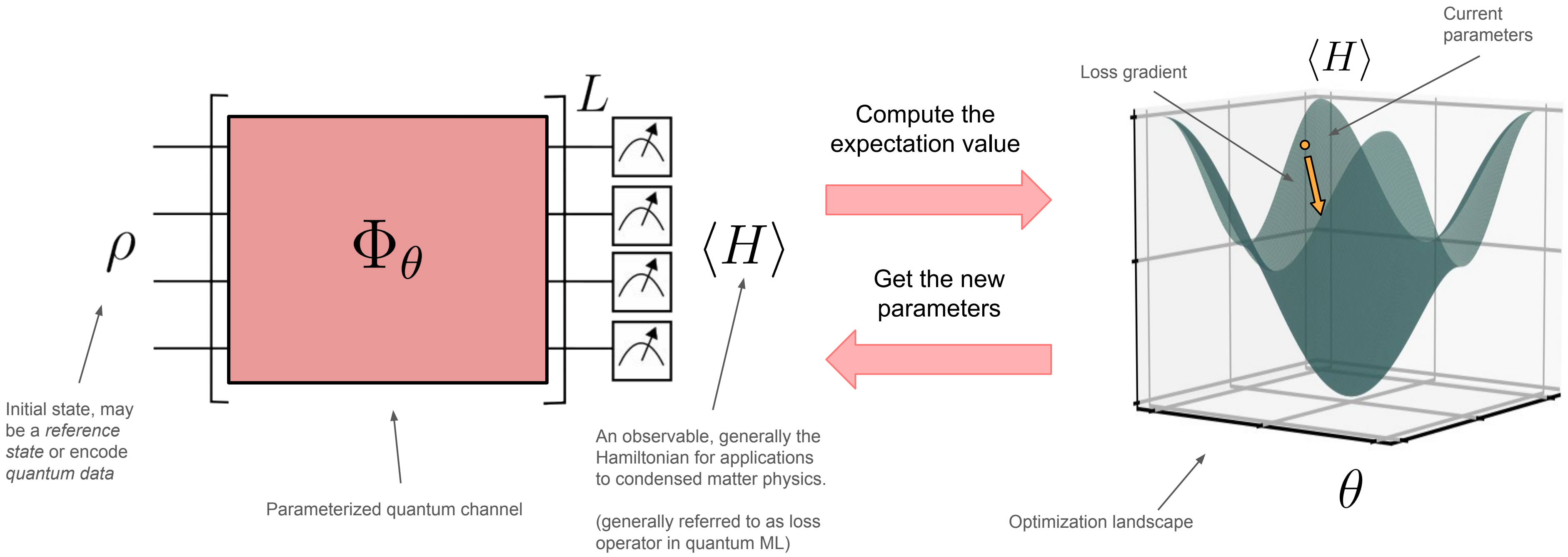
# Variational Quantum Algorithms in a nutshell

Variational Quantum Algorithms are based on a **hybrid** quantum classical optimization scheme.



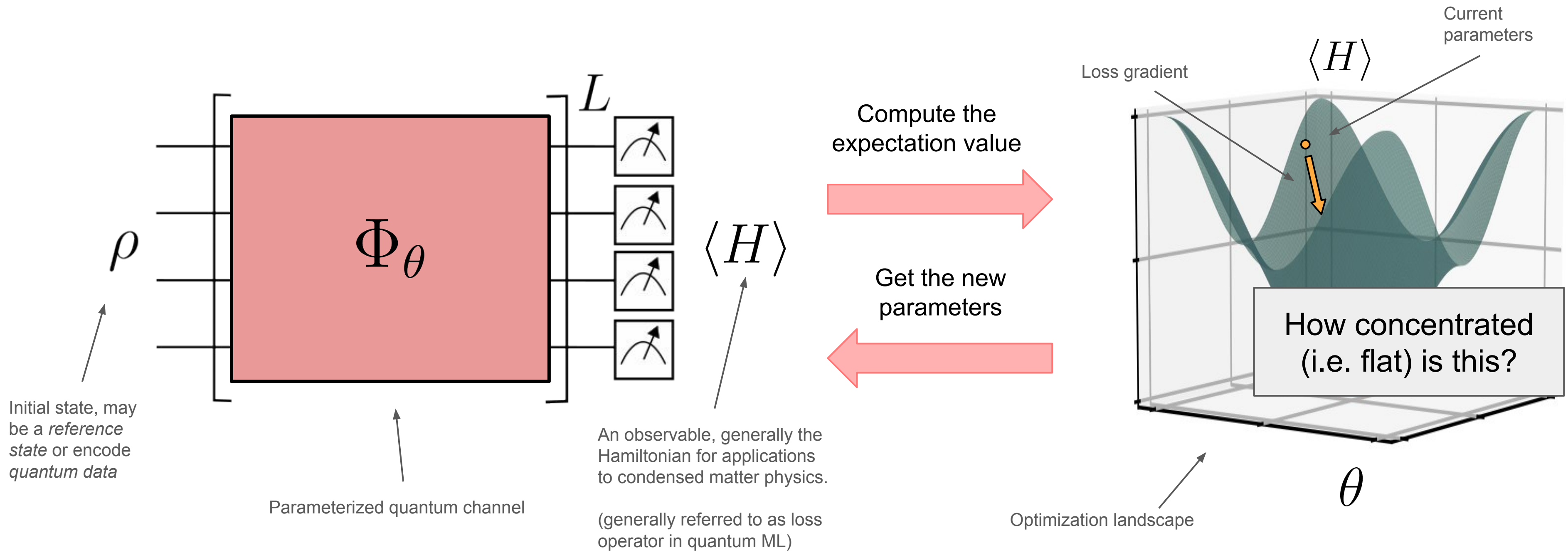
# Variational Quantum Algorithms in a nutshell

Variational Quantum Algorithms are based on a **hybrid** quantum classical optimization scheme.



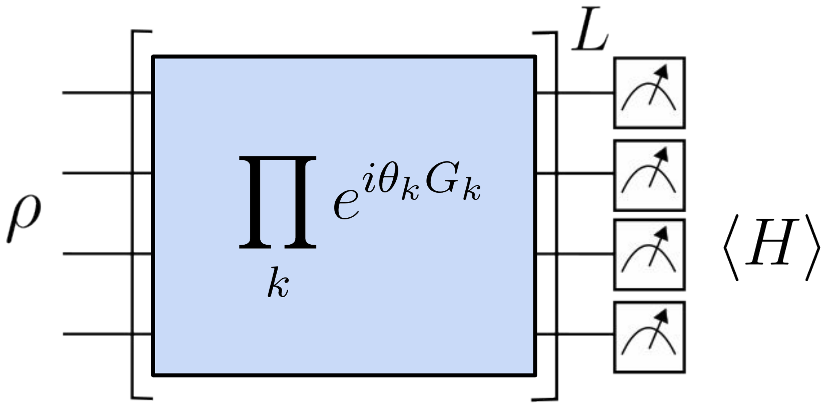
# Variational Quantum Algorithms in a nutshell

Variational Quantum Algorithms are based on a **hybrid** quantum classical optimization scheme.



# The emergence of Barren Plateaus (BP)

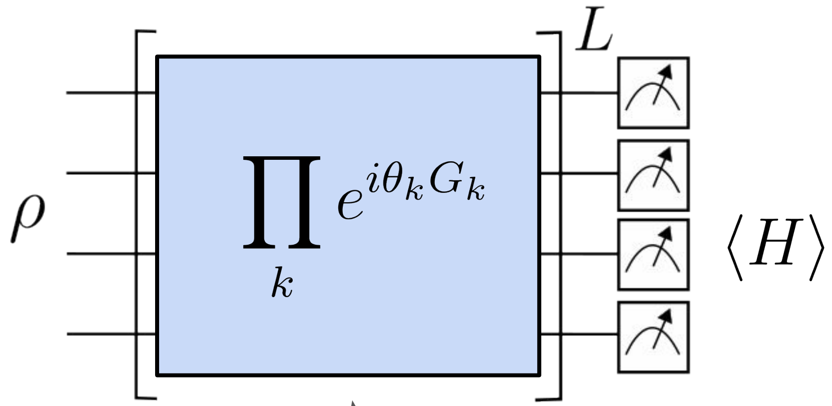
The conditions for concentration of the optimization landscape have been recently formalized for **deep**, **unitary** circuits, and hinge on the dimension of the **Dynamical Lie Algebra** (DLA) of the circuit.




E. Fontana et al., Nat. Comm. 2024, M. Ragone et al., Nat. Comm. 2024

# The emergence of Barren Plateaus (BP)

The conditions for concentration of the optimization landscape have been recently formalized for **deep**, **unitary** circuits, and hinge on the dimension of the **Dynamical Lie Algebra** (DLA) of the circuit.




Parameterized unitary circuit, with DLA

$$\mathfrak{g} = \langle \{ iG_k \} \rangle_{\mathrm{Lie}}$$

Structure of the DLA

$$\mathfrak{g} = \mathfrak{g}_1 \bigoplus \mathfrak{g}_2 \ldots \bigoplus \mathfrak{g}_J$$

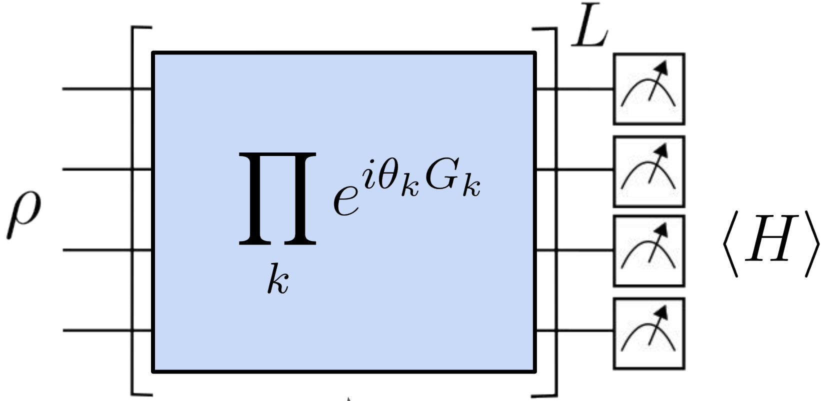
Variance of <H> w.r.t. the parameters

$$\mathbb{V}^{\infty}_{\rho,H} = \sum_j \frac{\mathcal{P}_j(\rho)\mathcal{P}_j(H)}{d_j}$$

E. Fontana et al., Nat. Comm. 2024, M. Ragone et al., Nat. Comm. 2024

# The emergence of Barren Plateaus (BP)

The conditions for concentration of the optimization landscape have been recently formalized for **deep**, **unitary** circuits, and hinge on the dimension of the **Dynamical Lie Algebra** (DLA) of the circuit.




Parameterized unitary circuit, with DLA

$$\mathfrak{g} = \langle \{ iG_k \} \rangle_{\text{Lie}}$$

Structure of the DLA

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \dots \oplus \mathfrak{g}_J$$

Variance of <H> w.r.t. the parameters

$$\mathbb{V}^{\infty}_{\rho,H} = \sum_j \frac{\mathcal{P}_j(\rho)\mathcal{P}_j(H)}{d_j}$$

Measure of purity in the respective sub-algera,

$$\mathcal{P}_j(A) = \sum_k \text{Tr}[B_k A]^2$$

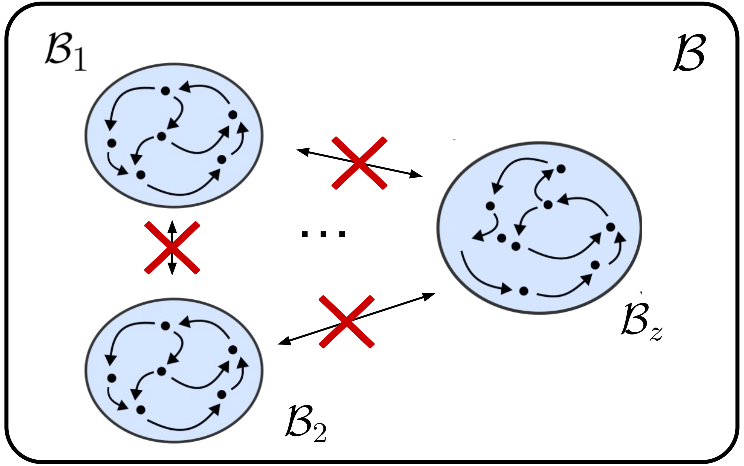
where $\{B_k\}$ is a basis of $g_j$

Dimension of the j-th component of the DLA

E. Fontana et al., Nat. Comm. 2024, M. Ragone et al., Nat. Comm. 2024

# The emergence of Barren Plateaus (BP)

One can visualize what's going on here using a **graph**. The space of bounded operators is represented here by a set of points. The action of Φ can only **connect some**, defining **invariant subspaces**.




Structure of the DLA

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \ldots \oplus \mathfrak{g}_J$$

Variance of <H> w.r.t. the parameters

$$\mathbb{V}^{\infty}_{\rho,H} = \sum_j \frac{\mathcal{P}_j(\rho)\mathcal{P}_j(H)}{d_j}$$

By definition, these subspaces **do not communicate**, and hence give an **independent** contribution.

# The emergence of Noise-Induced Barren Plateaus (NIBP)

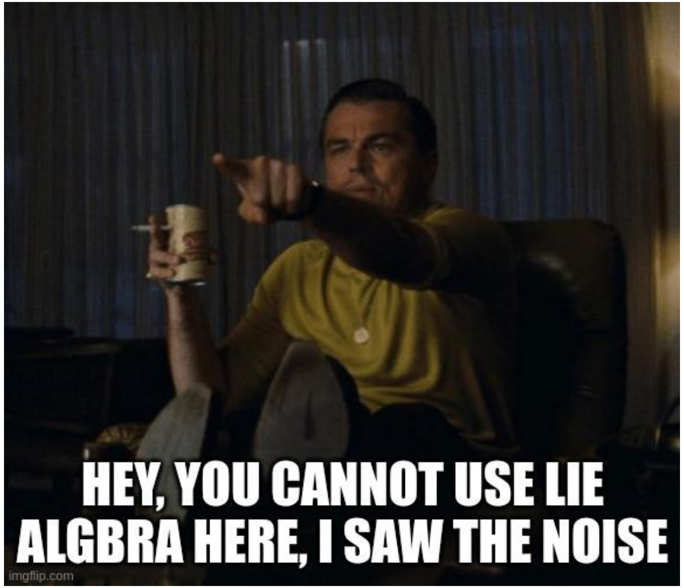
Any **substantial differences** in the presence of noise?

# The emergence of Noise-Induced Barren Plateaus (NIBP)

Any **substantial differences** in the presence of noise?

The previous analysis hinges on the **group structure** that **unitary** transformations have.

But **noise maps lack this structure**… which means no Lie-algebraic analyses here!

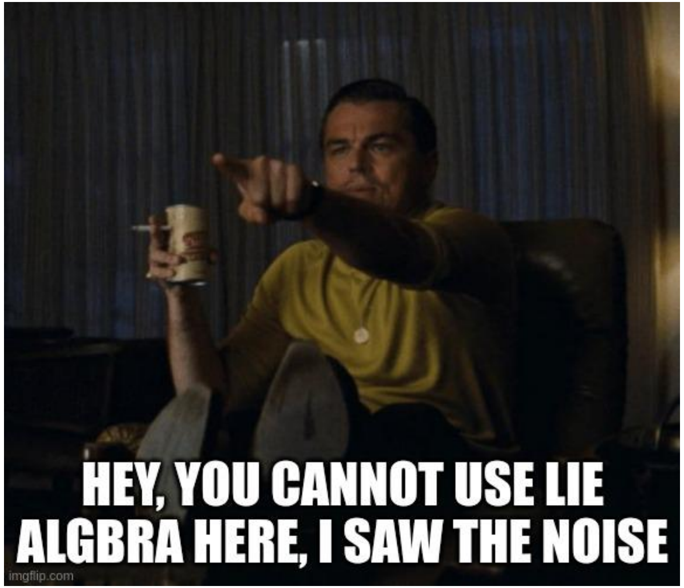


# The emergence of Noise-Induced Barren Plateaus (NIBP)

Any **substantial differences** in the presence of noise?

The previous analysis hinges on the **group structure** that **unitary** transformations have.

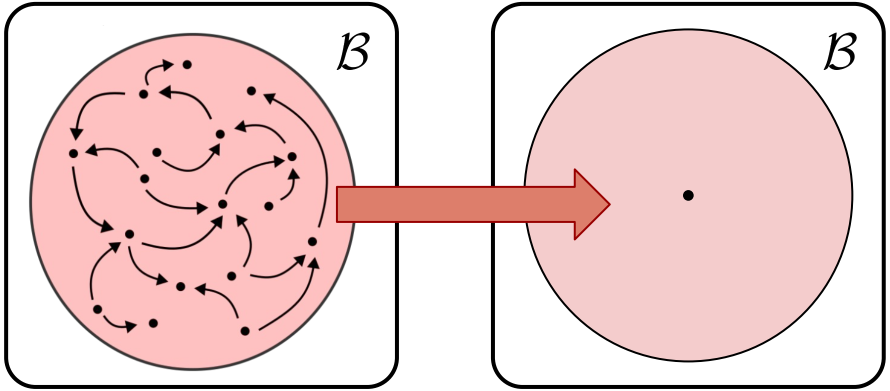
But **noise maps lack this structure**… which means no Lie-algebraic analyses here!




In this scenario, a much more dangerous effect is that of **decoherence**, which implies an **irreversible loss** of quantum **information** to the environment

# The emergence of Noise-Induced Barren Plateaus (NIBP)

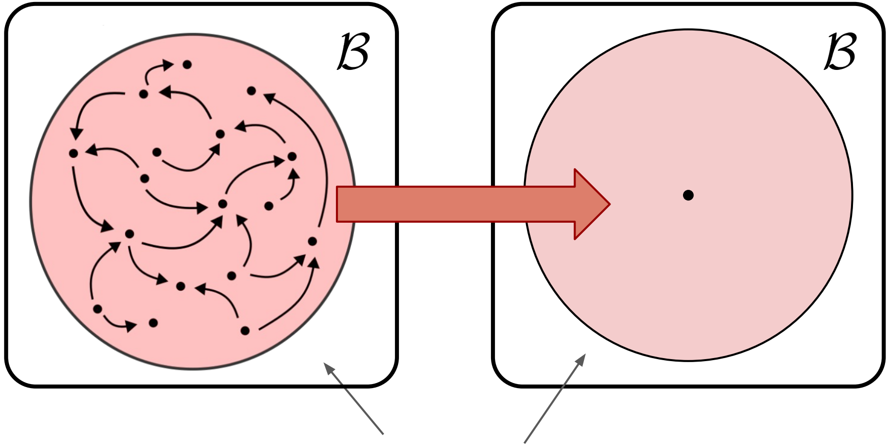
Indeed, the presence of invariant subspaces is now **irrelevant**. Basically, we are saying that loss of information is so dominant, that the variance decays **regardless of the symmetries** of the ansatz.




S. Wang et al., Nat. Comm. 2021

# The emergence of Noise-Induced Barren Plateaus (NIBP)

Indeed, the presence of invariant subspaces is now **irrelevant**. Basically, we are saying that loss of information is so dominant, that the variance decays **regardless of the symmetries** of the ansatz.

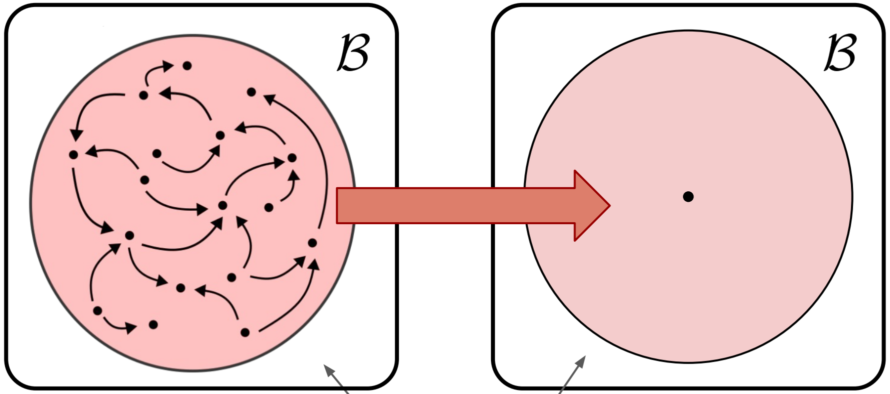



As the quantum information is lost, everything becomes **less and less distinguishable**. In the end, all points collapse into the same one, and all parameters inevitably produce the same output, i.e. the **variance vanishes**.

S. Wang et al., Nat. Comm. 2021

# The emergence of Noise-Induced Barren Plateaus (NIBP)

Indeed, the presence of invariant subspaces is now **irrelevant**. Basically, we are saying that loss of information is so dominant, that the variance decays **regardless of the symmetries** of the ansatz.




As the quantum information is lost, everything becomes **less and less distinguishable**. In the end, all points collapse into the same one, and all parameters inevitably produce the same output, i.e. the **variance vanishes**.

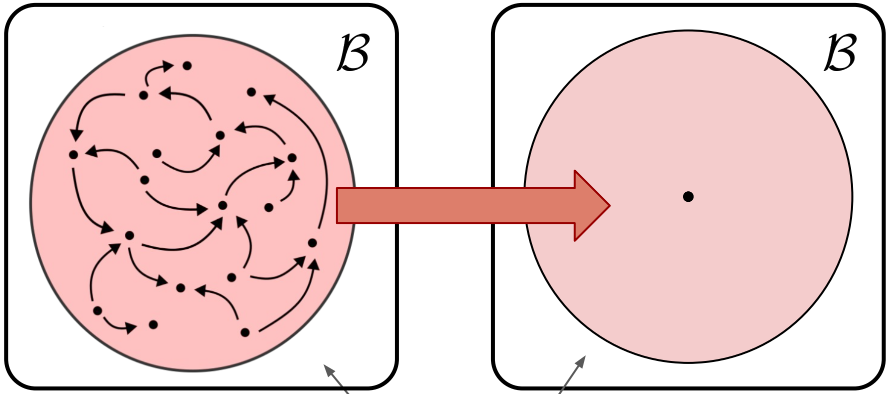
Variance of <H> w.r.t. the parameters

$$\mathbb{V}^{\infty}_{\rho,H} = 0$$

S. Wang et al., Nat. Comm. 2021

# The emergence of Noise-Induced Barren Plateaus (NIBP)

Indeed, the presence of invariant subspaces is now **irrelevant**. Basically, we are saying that loss of information is so dominant, that the variance decays **regardless of the symmetries** of the ansatz.




Variance of <H> w.r.t. the parameters

$$\mathbb{V}^{\infty}_{\rho,H} = 0$$

NIBP phenomenon, expressed in terms of the variance in the deep circuit limit

As the quantum information is lost, everything becomes **less and less distinguishable**. In the end, all points collapse into the same one, and all parameters inevitably produce the same output, i.e. the **variance vanishes**.

S. Wang et al., Nat. Comm. 2021

*Except the trivial one

# The ways of the Lord are not over…

Generally, noisy evolution is described by **quantum channels**, i.e. CPTP maps $\Phi(\rho)$. While there exists a wide variety of channels, we are interested in one major distinction:

| **Strictly contractive maps** | **Contractive maps** |
| --- | --- |
| Such maps always strictly **reduce** the norm of any operator they act on, in **all** subspaces* | Such maps **reduce** the norm of any operator they act on, but they also **preserve the norm** in some subspaces |
| $\|\Phi(A)\| < \|A\| \, \forall A$ | $\|\Phi(A)\| \leq \|A\|$ |
| e.g. depolarizing noise, p>0 | example: unitary transformations, but not only… |

# The ways of the Lord are not over…

Generally, noisy evolution is described by **quantum channels**, i.e. CPTP maps Φ(ρ). While there exists a wide variety of channels, we are interested in one major distinction:

**Strictly contractive maps**

Such maps always strictly **reduce** the norm of any operator they act on, in **all** subspaces*

$$\|\Phi(A)\| < \|A\| \; \forall A$$

e.g. depolarizing noise, p>0

**Contractive maps**

Such maps **reduce** the norm of any operator they act on, but they also **preserve the norm** in some subspaces

$$\|\Phi(A)\| \leq \|A\|$$

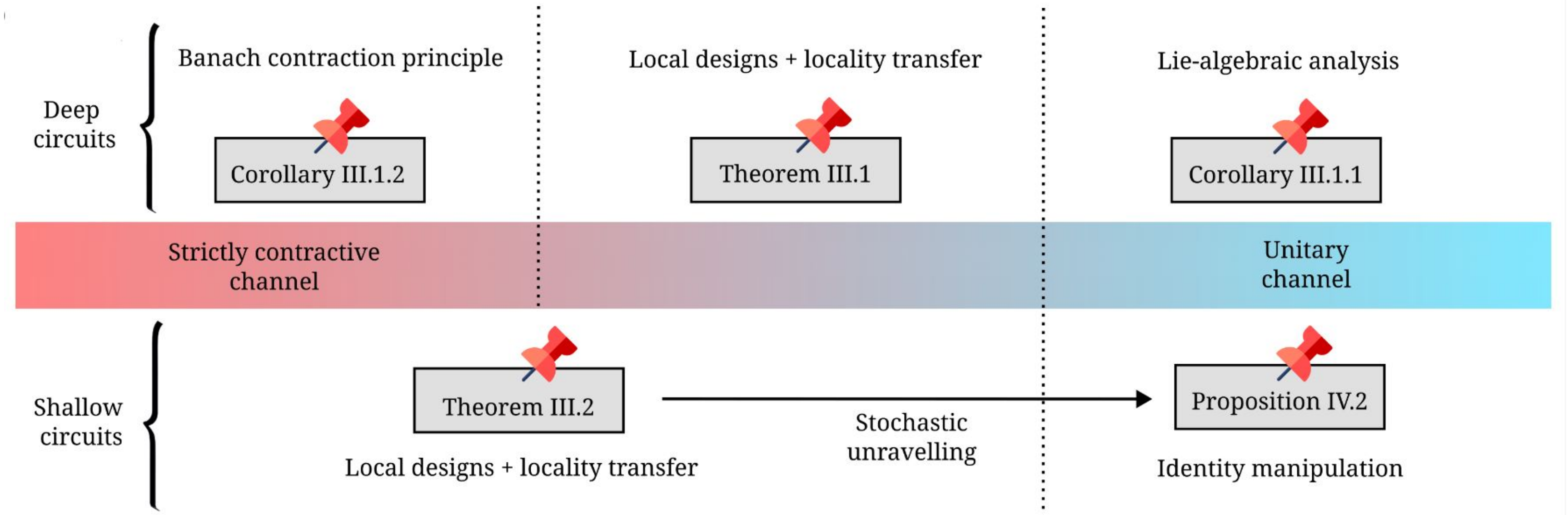
example: unitary transformations, but not only…

✗ Suffer from NIBP, we already know them well unfortunately…

? We know about unitary circuits, and maybe something for non-unital noise, what about the general case?

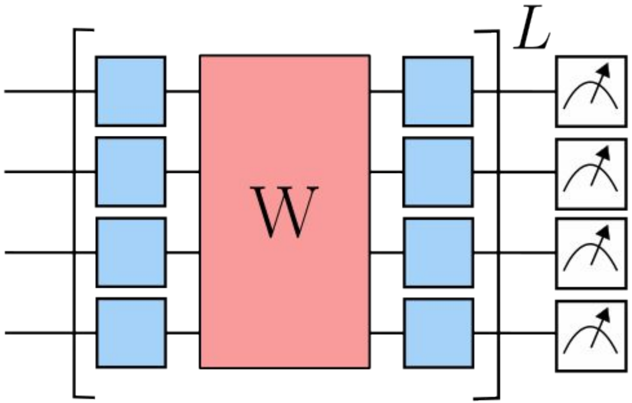
# …and they can give rise to different phenomena!

Indeed, BP and NIBP are just two **extreme cases** of a **richer** spectrum!



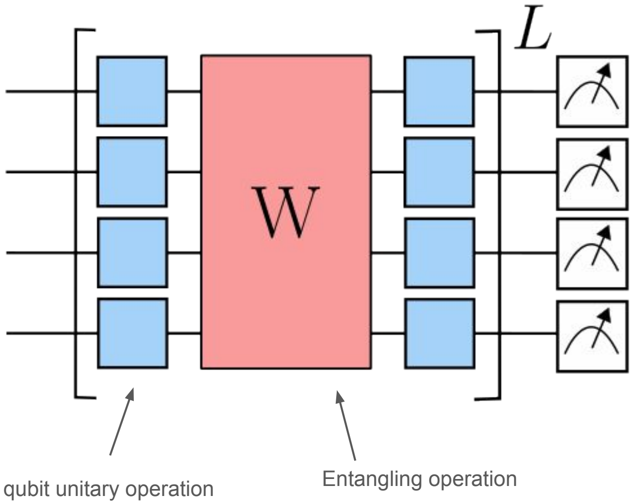
# How we plan to study it?

To investigate the main mechanisms at play in this general scenario, we devise a simple yet general model, comprised of **local unitary designs** and **general quantum channels**



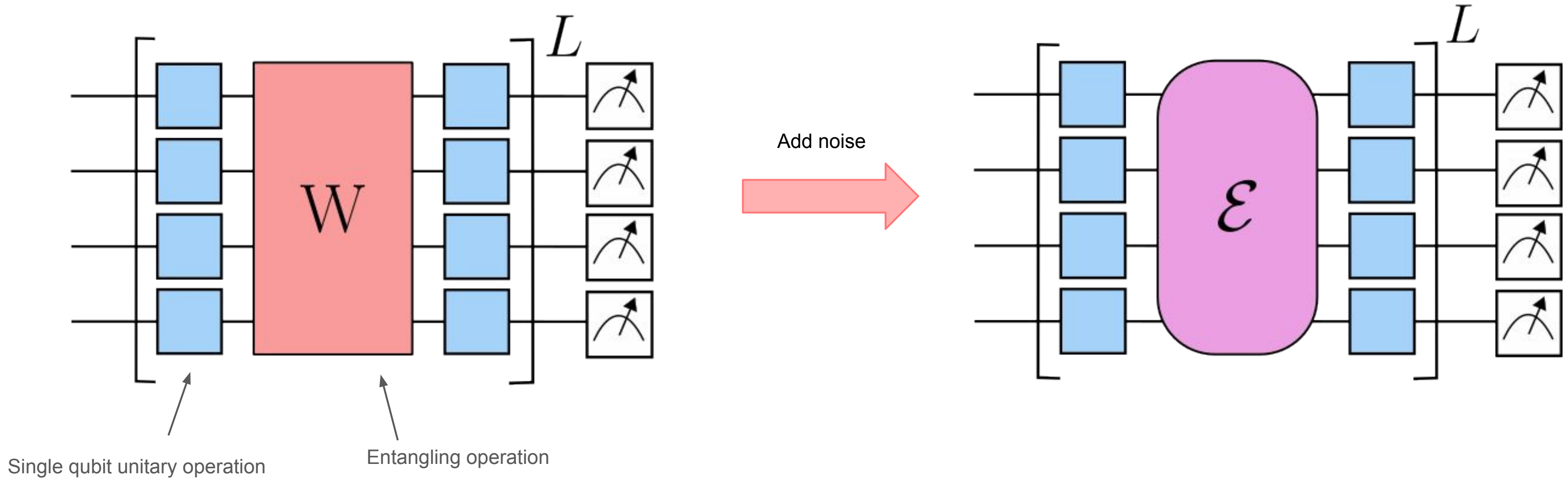
# How we plan to study it?

To investigate the main mechanisms at play in this general scenario, we devise a simple yet general model, comprised of **local unitary designs** and **general quantum channels**



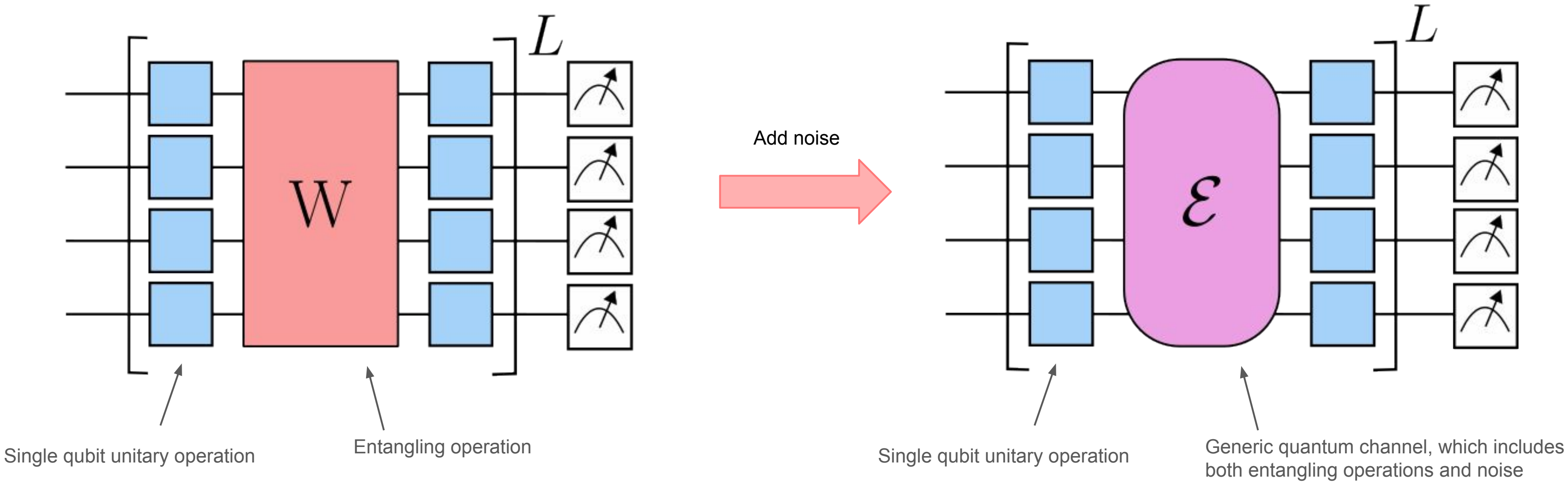
# How we plan to study it?

To investigate the main mechanisms at play in this general scenario, we devise a simple yet general model, comprised of **local unitary designs** and **general quantum channels**



# How we plan to study it?

To investigate the main mechanisms at play in this general scenario, we devise a simple yet general model, comprised of **local unitary designs** and **general quantum channels**



# The general formula

Using **non-negative matrix theory,** we are able to **exactly** compute the **variance** in the deep circuit limit. In particular, if we focus for simplicity on **single qubit noise**, we get:

$$\mathbb{V}^{\infty}_{\rho,H} = \sum_z \frac{(\ell_\rho)_z(\ell_H)_z}{d_z} + \frac{(\ell_\rho)_z(A\ell_H)_z}{d_z}$$

# The general formula

Using **non-negative matrix theory,** we are able to **exactly** compute the **variance** in the deep circuit limit. In particular, if we focus for simplicity on **single qubit noise**, we get:

**Variance** of loss function in the limit of **large number of layers**

$$\mathbb{V}^{\infty}_{\rho,H} = \sum_z \frac{(\ell_\rho)_z(\ell_H)_z}{d_z} + \frac{(\ell_\rho)_z(A\ell_H)_z}{d_z}$$

# The general formula

Using **non-negative matrix theory,** we are able to **exactly** compute the **variance** in the deep circuit limit.
In particular, if we focus for simplicity on **single qubit noise**, we get:

**Variance** of loss function in the limit of l**arge number of layers**

$$\mathbb{V}_{\rho,H}^{\infty} = \sum_{z} \frac{(\ell_\rho)_z (\ell_H)_z}{d_z} + \frac{(\ell_\rho)_z (A\ell_H)_z}{d_z}$$

Sum over **invariant subspaces of both** the single qubit blocks and the intermediate channel

# The general formula

Using **non-negative matrix theory,** we are able to **exactly** compute the **variance** in the deep circuit limit.
In particular, if we focus for simplicity on **single qubit noise**, we get:

**Variance** of loss function in the limit of **large number of layers**

**Locality** (= purity) of the initial state and observable

$$\mathbb{V}^{\infty}_{\rho,H} = \sum_z \frac{(\ell_\rho)_z(\ell_H)_z}{d_z} + \frac{(\ell_\rho)_z(A\ell_H)_z}{d_z}$$

Sum over **invariant subspaces of both** the single qubit blocks and the intermediate channel

**Dimension** of such subspaces

# The general formula

Using **non-negative matrix theory,** we are able to **exactly** compute the **variance** in the deep circuit limit. In particular, if we focus for simplicity on **single qubit noise**, we get:

**Variance** of loss function in the limit of **large number of layers**

**Locality** (= purity) of the initial state and observable

**Absorption matrix**, represents the absorption to the invariant subspaces

$$\mathbb{V}^{\infty}_{\rho,H} = \sum_z \frac{(\ell_\rho)_z(\ell_H)_z}{d_z} + \frac{(\ell_\rho)_z(A\ell_H)_z}{d_z}$$

Sum over **invariant subspaces of both** the single qubit blocks and the intermediate channel

**Dimension** of such subspaces

# The general formula

Using **non-negative matrix theory,** we are able to **exactly** compute the **variance** in the deep circuit limit. In particular, if we focus for simplicity on **single qubit noise**, we get:

$$\mathbb{V}^{\infty}_{\rho,H} = \sum_z \frac{(\ell_\rho)_z(\ell_H)_z}{d_z} + \frac{(\ell_\rho)_z(A\ell_H)_z}{d_z}$$

**Variance** of loss function in the limit of **large number of layers**

Sum over **invariant subspaces of both** the single qubit blocks and the intermediate channel

**Locality** (= purity) of the initial state and observable

**Absorption matrix**, represents the absorption to the invariant subspaces

**Dimension** of such subspaces

**Unitary contribution**

**Noise contribution**

**Holds rigorously for aperiodic circuits. For periodic circuits only holds for the Cesàro sum of all depths L.

# The general formula

Using **non-negative matrix theory,** we are able to **exactly** compute the **variance** in the deep circuit limit. In particular, if we focus for simplicity on **single qubit noise**, we get:

$$\mathbb{V}^{\infty}_{\rho,H} = \sum_z \frac{(\ell_\rho)_z (\ell_H)_z}{d_z} + \frac{(\ell_\rho)_z (A\ell_H)_z}{d_z}$$

**Variance** of loss function in the limit of **large number of layers**

Sum over **invariant subspaces of both** the single qubit blocks and the intermediate channel

**Locality** (= purity) of the initial state and observable

**Absorption matrix**, represents the absorption to the invariant subspaces

**Dimension** of such subspaces

**Unitary contribution**

**Noise contribution**

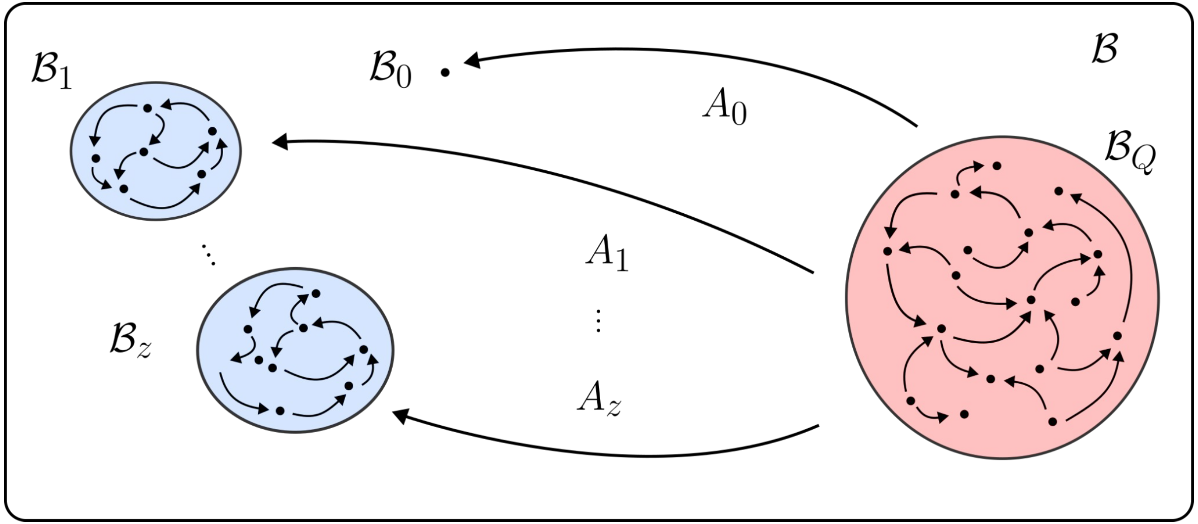
More specifically, we have **Theorem III.1 (Deep circuits):**

*The variance converges** exponentially fast in the number of layers to the above value, i.e.*

$$\left|\mathbb{V}^{L}_{\rho,H} - \mathbb{V}^{\infty}_{\rho,H}\right| \in O\left(e^{-\beta L}\|H\|_2^2\right)$$

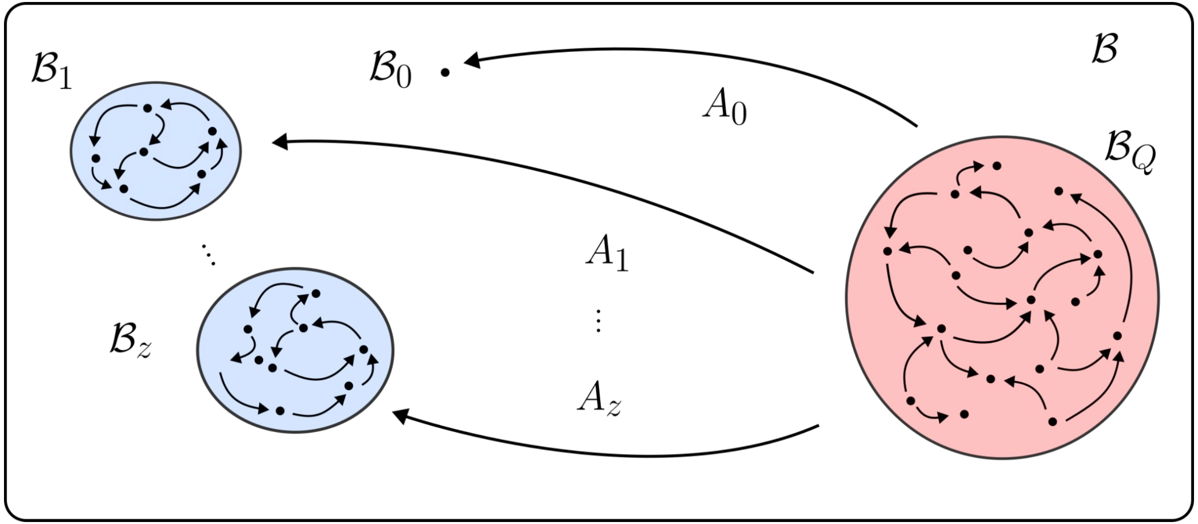
# An intuitive picture

We can visualize what’s going on here again by a **graph**, summarizing the **interactions** in the **space of bounded operators**



# An intuitive picture

We can visualize what's going on here again by a **graph**, summarizing the **interactions** in the **space of bounded operators**

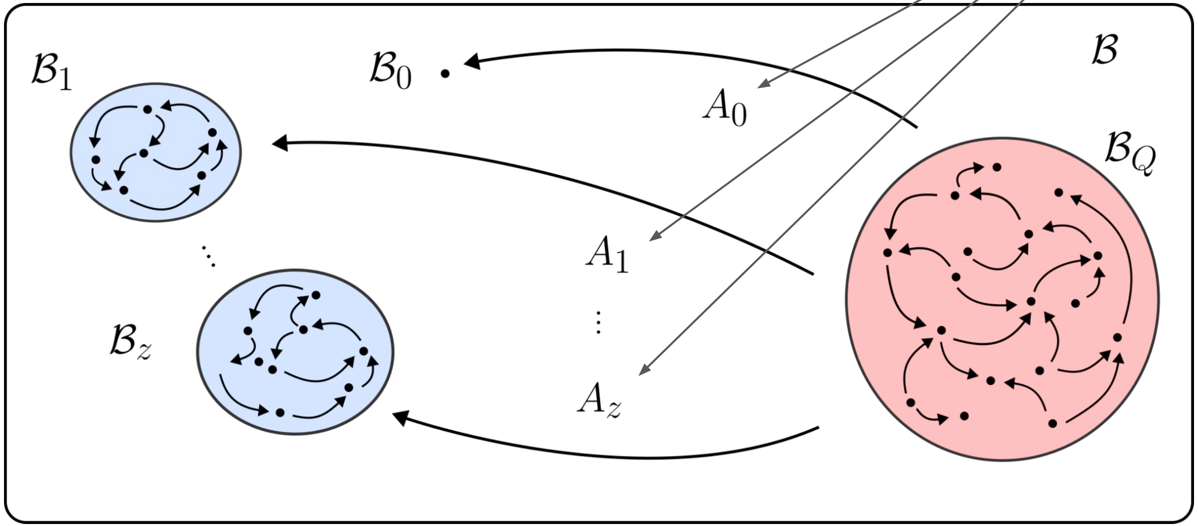



**Strictly contractive** subspaces may be all connected or have symmetries, but it doesn't matter, as their contribution **vanish anyway.**

# An intuitive picture

We can visualize what's going on here again by a **graph**, summarizing the **interactions** in the **space of bounded operators**

**Strictly contractive** subspaces can in general **extend outside**, i.e. they are not invariant. But **once they reach** an invariant subspace, **they are bound to stay there**! These are what makes **contributions from noise** possible!



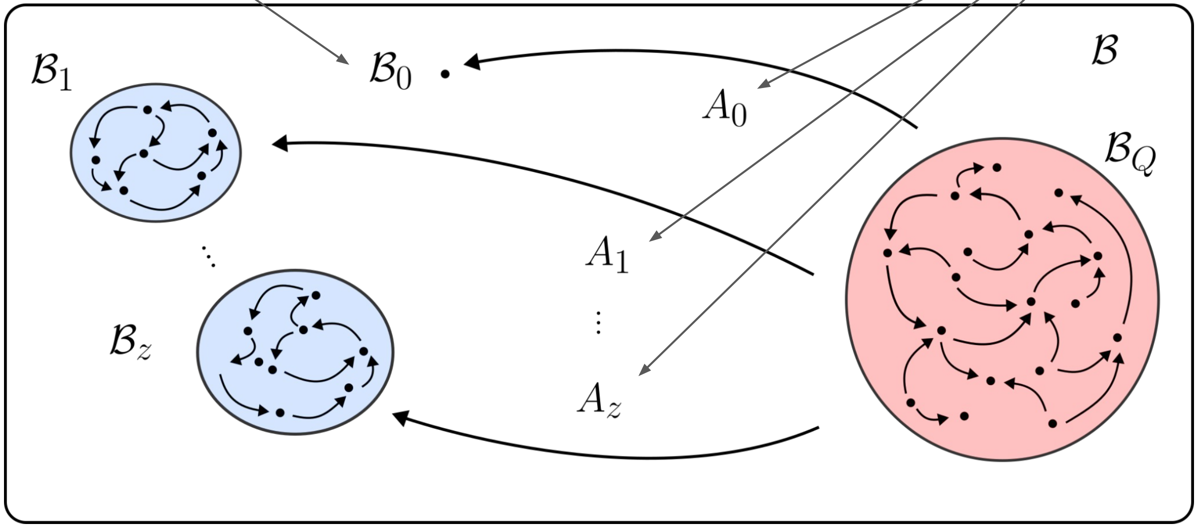

**Strictly contractive** subspaces may be all connected or have symmetries, but it doesn't matter, as their contribution **vanish anyway.**

# An intuitive picture

We can visualize what's going on here again by a **graph**, summarizing the **interactions** in the **space of bounded operators**

**The trivial subspace** normally **doesn't contribute** to the variance **by itself,** but here it has a **great importance,** since it can **collect** contributions coming from **strictly contractive subspaces!**

**Strictly contractive** subspaces can in general **extend outside**, i.e. they are not invariant. But **once they reach** an invariant subspace, **they are bound to stay there**! These are what makes **contributions from noise** possible!




**Strictly contractive** subspaces may be all connected or have symmetries, but it doesn't matter, as their contribution **vanish anyway.**

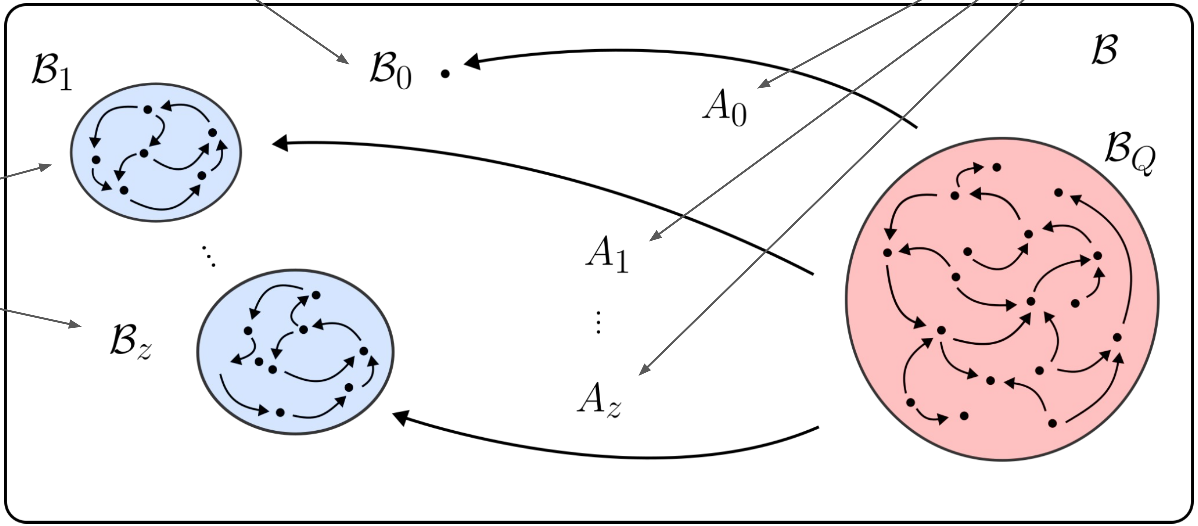
# An intuitive picture

We can visualize what's going on here again by a **graph**, summarizing the **interactions** in the **space of bounded operators**

**The trivial subspace** normally **doesn't contribute** to the variance **by itself,** but here it has a **great importance,** since it can **collect** contributions coming from **strictly contractive subspaces!**

**Strictly contractive** subspaces can in general **extend outside**, i.e. they are not invariant. But **once they reach** an invariant subspace, **they are bound to stay there**! These are what makes **contributions from noise** possible!

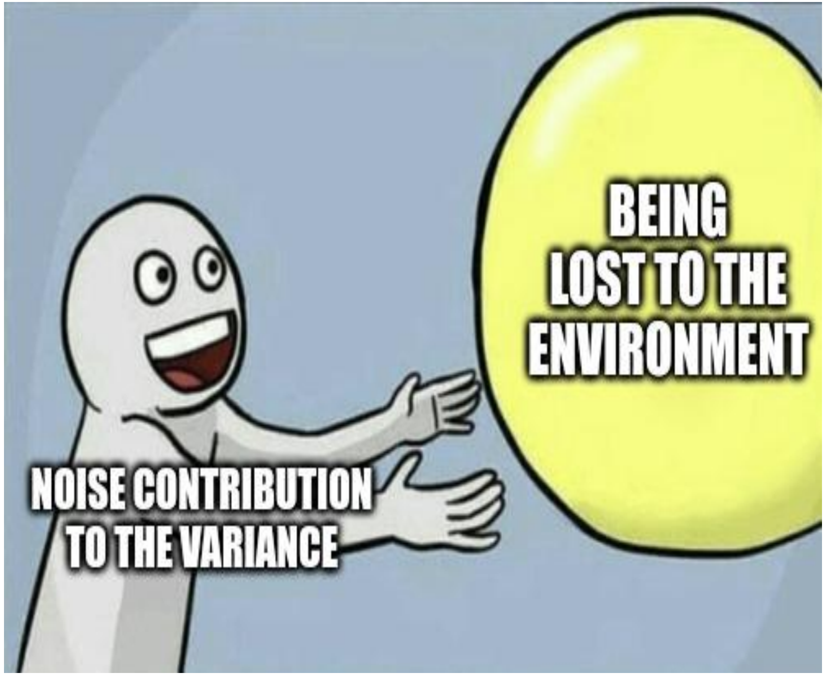
**Norm-preserving subspaces,** exactly as in the unitary case, cannot communicate with the outside. Their contribution is **proportional to their dimension.**

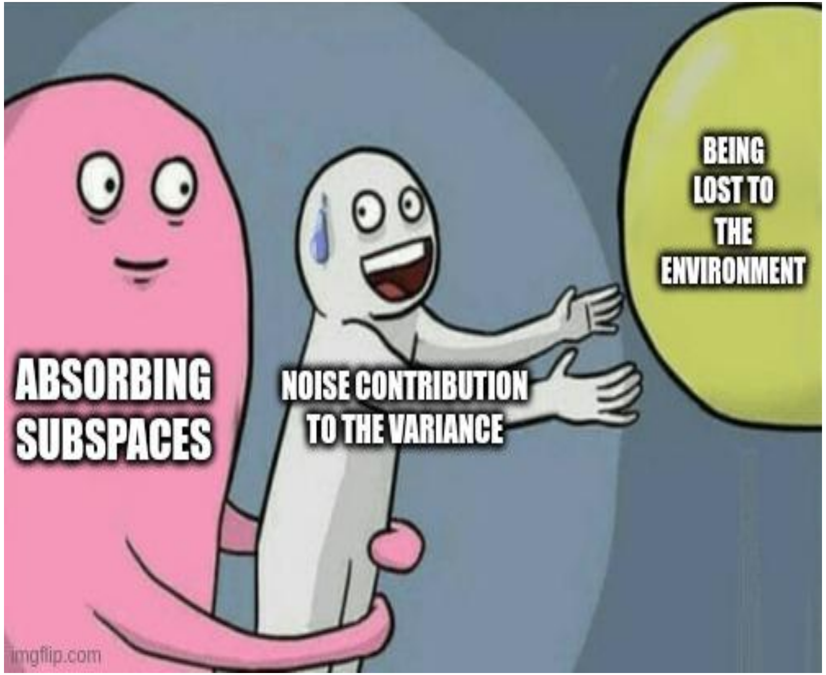



**Strictly contractive** subspaces may be all connected or have symmetries, but it doesn't matter, as their contribution **vanish anyway.**

# TL;DR

There is a meme template for everything these days :)



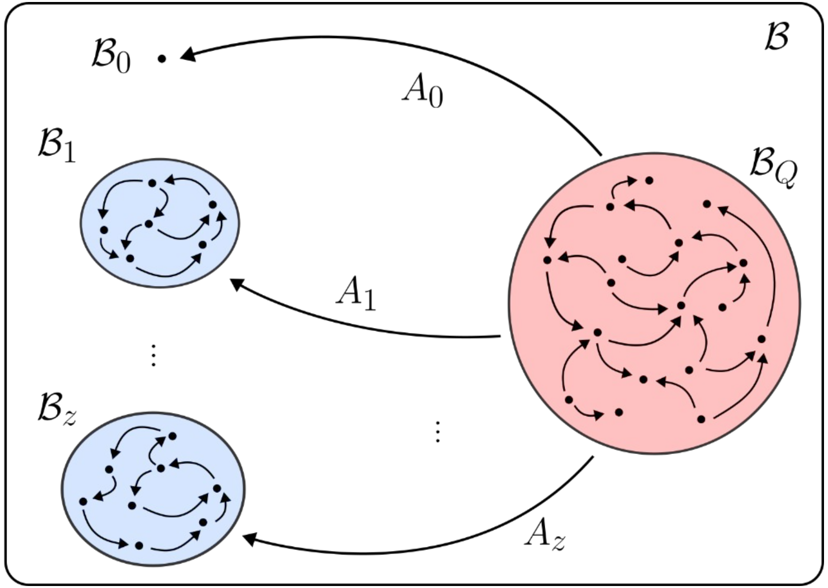
# Examples: Unitary circuits

The absorption term A vanishes in this case because unitary dynamics is **reversible**.

**Corollary III.1.1 (Deep, unitary circuits).**

*If the channel is unitary, then the absorption terms vanish.*

$$\mathbb{V}^{\infty}_{\rho,H} = \sum_z \frac{(\ell_\rho)_z(\ell_H)_z}{d_z}$$



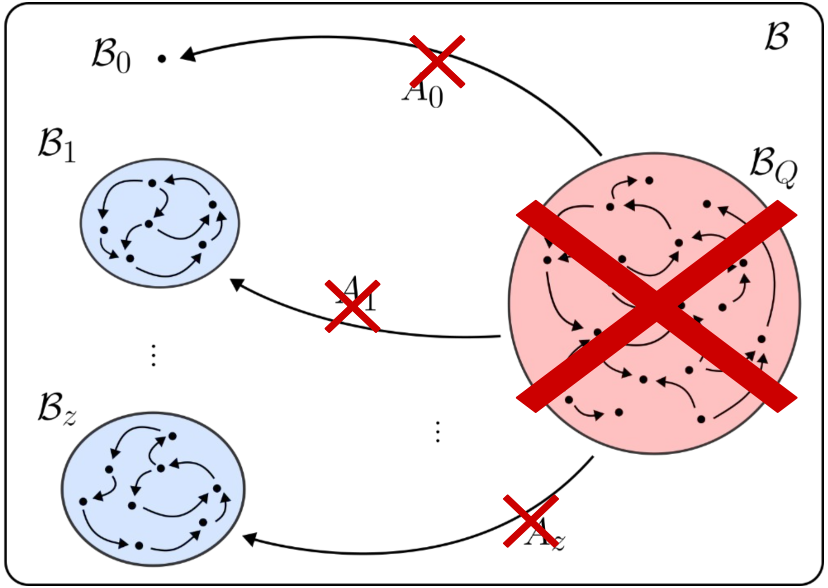
# Examples: Unitary circuits

The absorption term A vanishes in this case because unitary dynamics is **reversible**.

**Corollary III.1.1 (Deep, unitary circuits).**

*If the channel is unitary, then the absorption terms vanish.*

$$\mathbb{V}^{\infty}_{\rho,H} = \sum_{z} \frac{(\ell_\rho)_z(\ell_H)_z}{d_z}$$



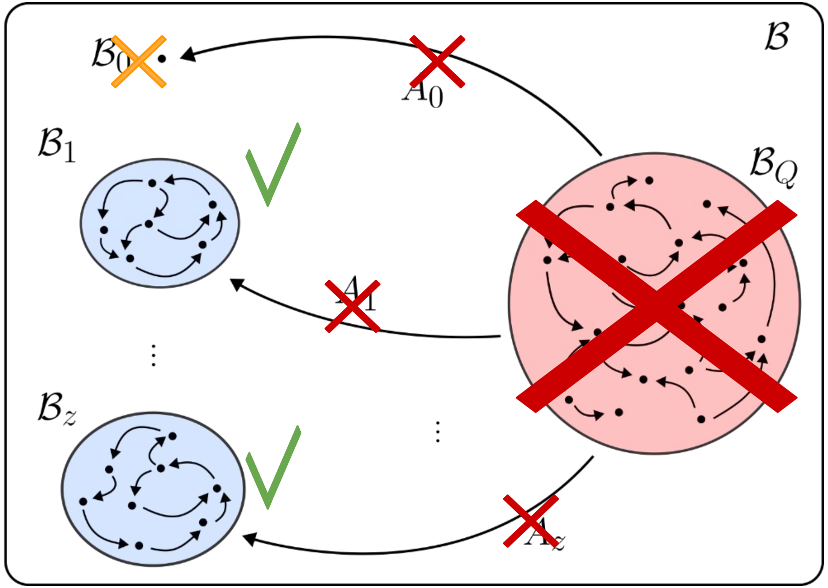
# Examples: Unitary circuits

The absorption term A vanishes in this case because unitary dynamics is **reversible**.

**Corollary III.1.1 (Deep, unitary circuits).**

*If the channel is unitary, then the absorption terms vanish.*

$$\mathbb{V}^{\infty}_{\rho,H} = \sum_z \frac{(\ell_\rho)_z (\ell_H)_z}{d_z}$$



# Examples: Contractive circuits

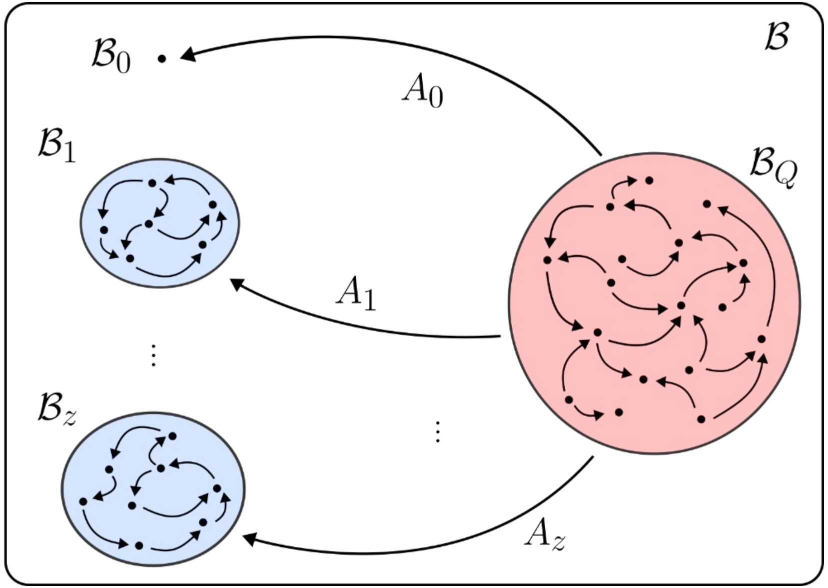
Conversely, since each invariant subspace norm-preserving to ‘*survive*’ the deep circuit limit, in the strictly contractive case **only the absorption terms** remain.

**Corollary III.1.2 (Deep, contractive circuits).**

*If the channel has no unitary subspaces, then only the absorption term survives.*

$$\mathbb{V}^{\infty}_{\rho,H} = \frac{(A\ell_H)_0}{d}$$

*In particular, if the channel is unital, then* $\mathbb{V}^{\infty}_{\rho,H} = 0$.



# Examples: Contractive circuits

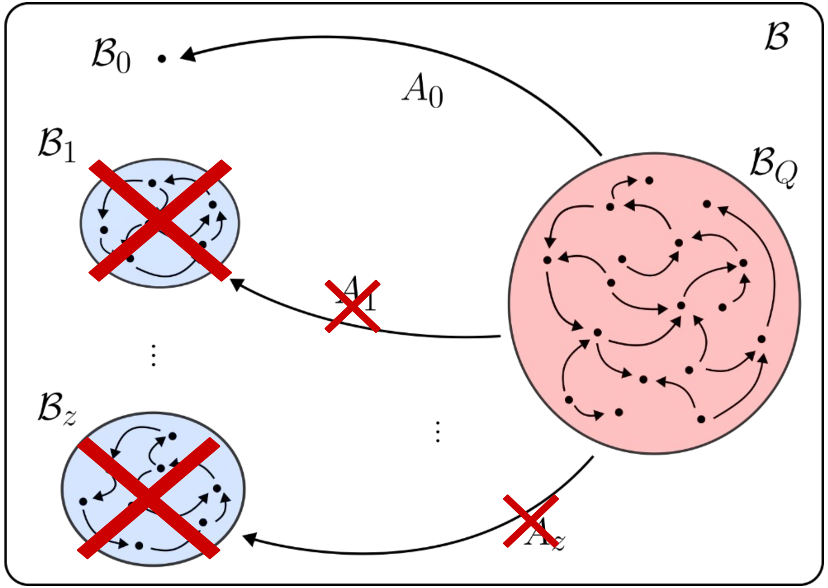
Conversely, since each invariant subspace norm-preserving to '*survive*' the deep circuit limit, in the strictly contractive case **only the absorption terms** remain.

**Corollary III.1.2 (Deep, contractive circuits).**

*If the channel has no unitary subspaces, then only the absorption term survives.*

$$\mathbb{V}^{\infty}_{\rho,H} = \frac{(A\ell_H)_0}{d}$$

*In particular, if the channel is unital, then* $\mathbb{V}^{\infty}_{\rho,H} = 0\,.$



# Examples: Contractive circuits

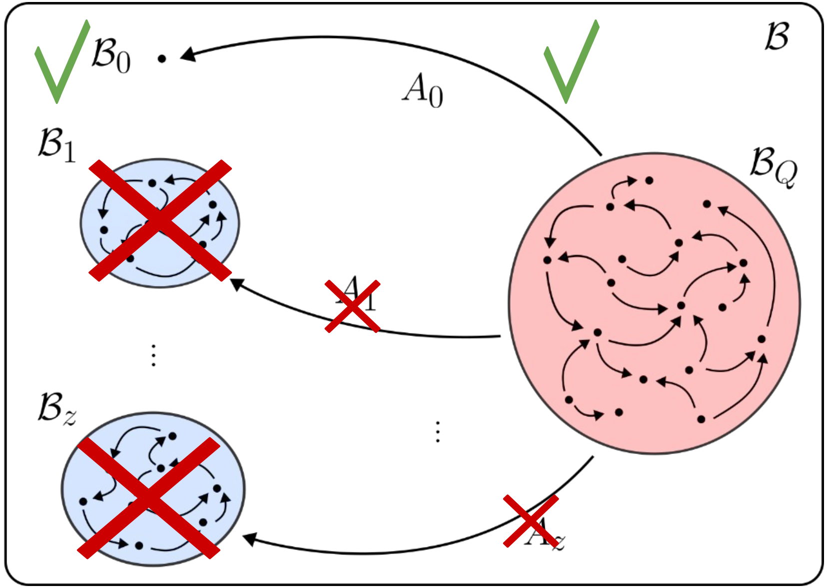
Conversely, since each invariant subspace norm-preserving to '*survive*' the deep circuit limit, in the strictly contractive case **only the absorption terms** remain.

**Corollary III.1.2 (Deep, contractive circuits).**

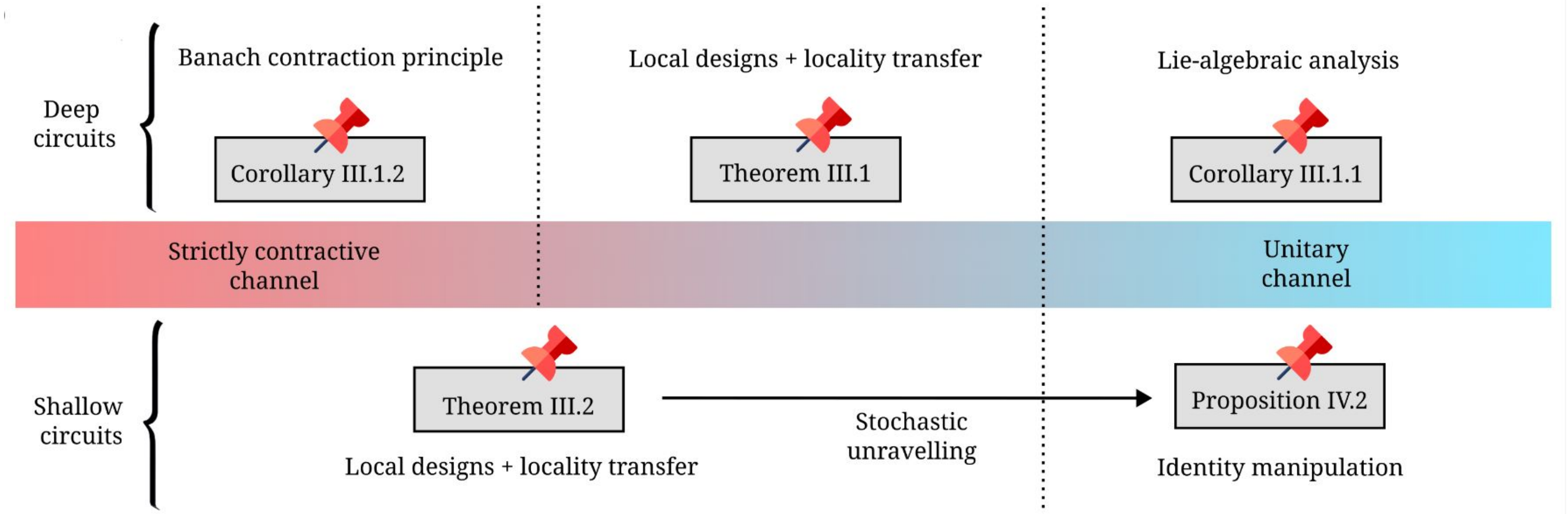
*If the channel has no unitary subspaces, then only the absorption term survives.*

$$\mathbb{V}^{\infty}_{\rho,H} = \frac{(A\ell_H)_0}{d}$$

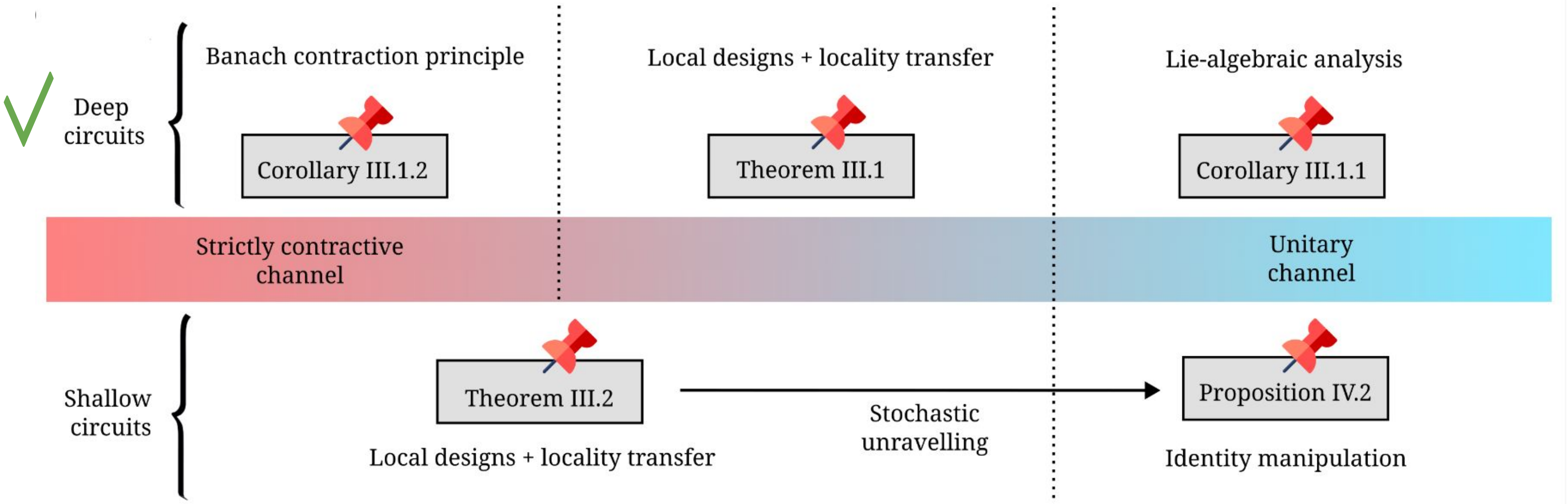
*In particular, if the channel is unital, then* $\mathbb{V}^{\infty}_{\rho,H} = 0$.



# Summary

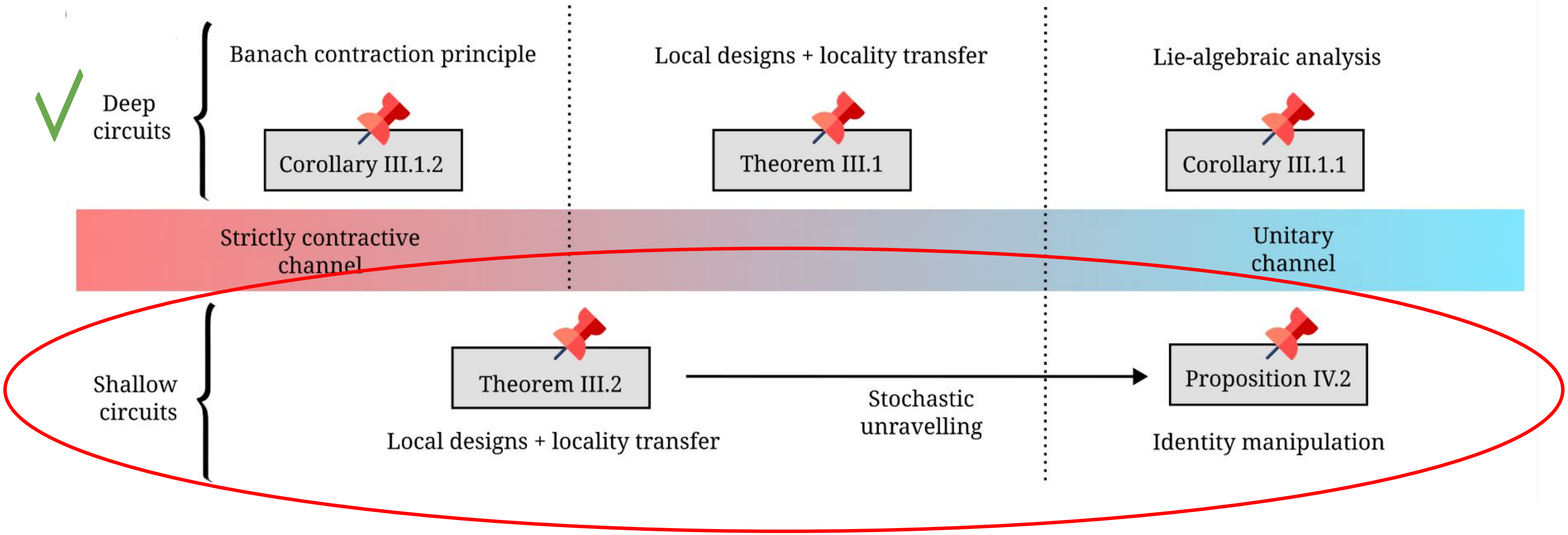

Deep
circuits
Banach contraction principle
Corollary III.1.2
Local designs + locality transfer
Theorem III.1
Lie-algebraic analysis
Corollary III.1.1
Strictly contractive
channel
Unitary
channel
Shallow
circuits
Theorem III.2
Local designs + locality transfer
Stochastic
unravelling
Proposition IV.2
Identity manipulation

# Summary



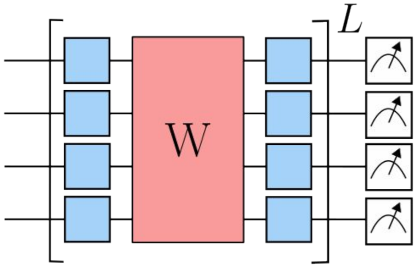
Deep circuits
Banach contraction principle
Corollary III.1.2
Local designs + locality transfer
Theorem III.1
Lie-algebraic analysis
Corollary III.1.1
Strictly contractive channel
Unitary channel
Shallow circuits
Theorem III.2
Local designs + locality transfer
Stochastic unravelling
Proposition IV.2
Identity manipulation

# Summary


Deep
circuits
Banach contraction principle
Corollary III.1.2
Local designs + locality transfer
Theorem III.1
Lie-algebraic analysis
Corollary III.1.1
Strictly contractive
channel
Unitary
channel
Shallow
circuits
Theorem III.2
Local designs + locality transfer
Stochastic
unravelling
Proposition IV.2
Identity manipulation

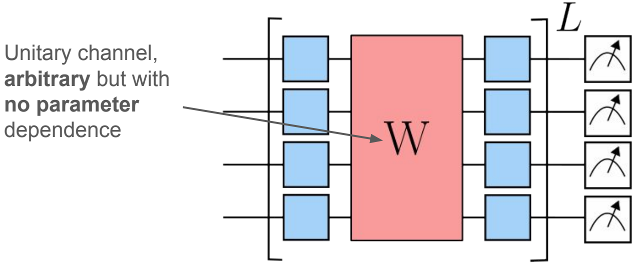
# Quantum Residual Networks

A simple way to escape the results shown here is to resort to **shallow circuit**. Even for deep circuits, we can get away from the main theorem if the circuit is **effectively shallow**.



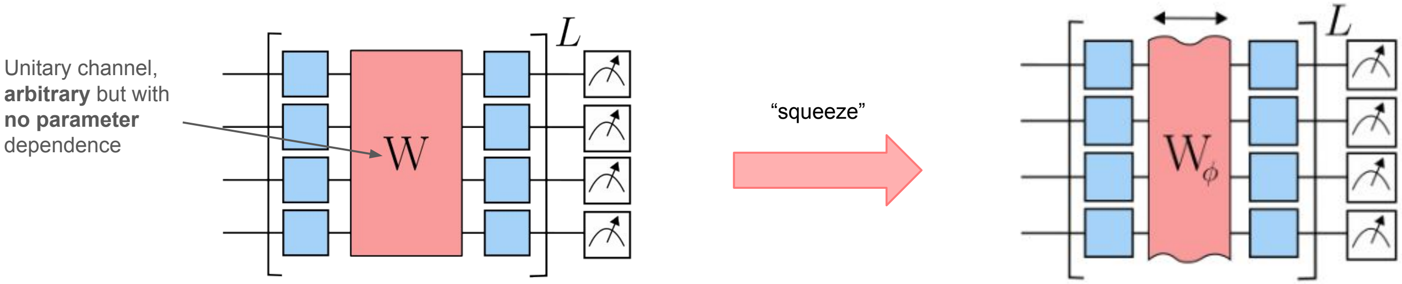
# Quantum Residual Networks

A simple way to escape the results shown here is to resort to **shallow circuit**. Even for deep circuits, we can get away from the main theorem if the circuit is **effectively shallow**.



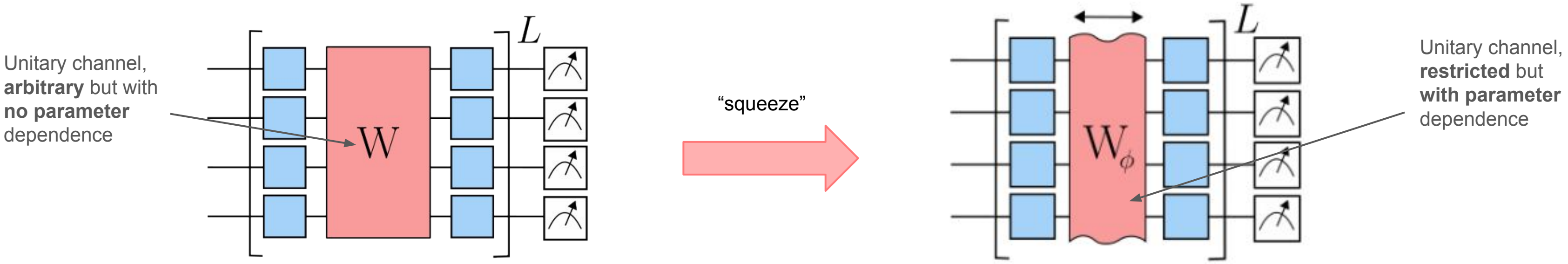
# Quantum Residual Networks

A simple way to escape the results shown here is to resort to **shallow circuit**. Even for deep circuits, we can get away from the main theorem if the circuit is **effectively shallow**.



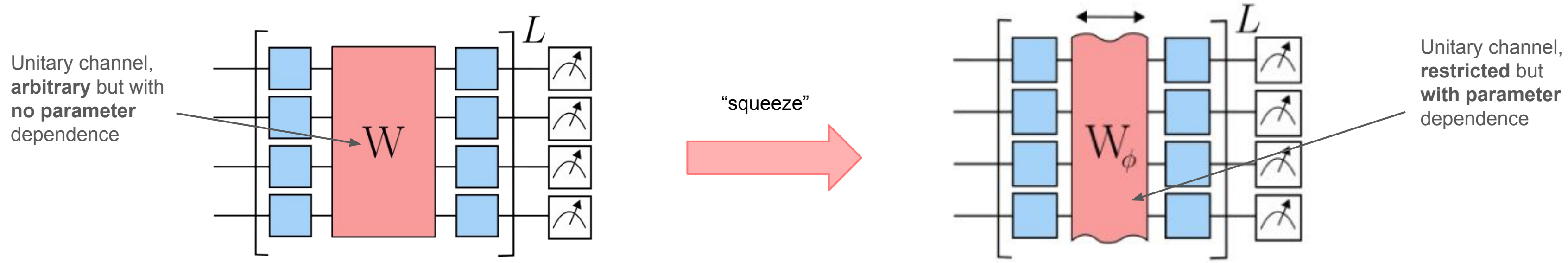
# Quantum Residual Networks

A simple way to escape the results shown here is to resort to **shallow circuit**. Even for deep circuits, we can get away from the main theorem if the circuit is **effectively shallow**.



# Quantum Residual Networks

A simple way to escape the results shown here is to resort to **shallow circuit**. Even for deep circuits, we can get away from the main theorem if the circuit is **effectively shallow**.




**Proposition IV.1 (QUResNet).**
*If the channel is unitary, and parameterized, with parameters having variance scaling inversely proportional to L, then*

$$\mathbb{V}^{L}_{\rho,H} \geq F(n)(\ell_\rho, \ell_H)$$

*where F(n) decays polynomially with n.*

# An intuitive picture

This result still hinges, although indirectly, on the deep circuit limit result, as it works by enforcing a **slower speed of convergence** to the process, acting on the constant **β**

$$\left.\begin{aligned}\left|\mathbb{V}_{\rho,H}^{L} - \mathbb{V}_{\rho,H}^{\infty}\right| \in O\left(e^{-\beta L}\|H\|_2^2\right) \\ \beta \in O\left(1/L\right)\end{aligned}\right\}$$

# An intuitive picture

This result still hinges, although indirectly, on the deep circuit limit result, as it works by enforcing a **slower speed of convergence** to the process, acting on the constant **β**

Theorem III.1 in the aperiodic regime

$$\left|\mathbb{V}^{L}_{\rho,H} - \mathbb{V}^{\infty}_{\rho,H}\right| \in O\left(e^{-\beta L}\|H\|_2^2\right)$$

$$\beta \in O\left(1/L\right)$$

Upper bound on β, explicitly requiring it to go to zero at least as fast as the number of layers increase.

→ No formal bound to ensure convergence after **any number** of layers

# An intuitive picture

This result still hinges, although indirectly, on the deep circuit limit result, as it works by enforcing a **slower speed of convergence** to the process, acting on the constant **β**

Theorem III.1 in the aperiodic regime

$$\left|\mathbb{V}^{L}_{\rho,H} - \mathbb{V}^{\infty}_{\rho,H}\right| \in O\left(e^{-\beta L}\|H\|_2^2\right)$$

$$\beta \in O\left(1/L\right)$$

Upper bound on β, explicitly requiring it to go to zero at least as fast as the number of layers increase.

→ No formal bound to ensure convergence after **any number** of layers

This regime is achieved also in the **noisy regime**, when the noise is **sufficiently weak.**

# Stochastic unravelling of channels and QResNets

For any given noise map, we are guaranteed to be able to describe it as a **stochastic unravelling**, i.e. as the **expectation value** of some simpler operators. For certain maps, such operators can even be **unitary**!

$$\mathcal{E}(\rho) = \mathbb{E}_\phi \left\{ W_\phi \rho W_\phi^\dagger \right\} \qquad \{W_\phi\} \subset U(d)$$

# Stochastic unravelling of channels and QResNets

For any given noise map, we are guaranteed to be able to describe it as a **stochastic unravelling**, i.e. as the **expectation value** of some simpler operators. For certain maps, such operators can even be **unitary**!

$$\mathcal{E}(\rho) = \mathbb{E}_\phi \left\{ W_\phi \rho W_\phi^\dagger \right\} \qquad \{W_\phi\} \subset U(d)$$

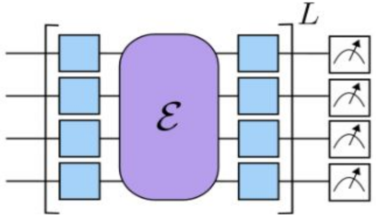
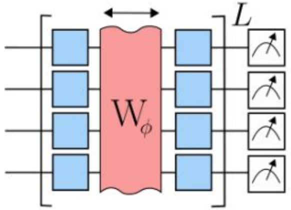
**Proposition IV.2 (informal)**
The variance of a channel is always smaller than the variance of the corresponding stochastic unravelling, with equality holding if and only if the channel is unitary.

# Stochastic unravelling of channels and QResNets

For any given noise map, we are guaranteed to be able to describe it as a **stochastic unravelling**, i.e. as the **expectation value** of some simpler operators. For certain maps, such operators can even be **unitary**!

$$\mathcal{E}(\rho) = \mathbb{E}_\phi \left\{ W_\phi \rho W_\phi^\dagger \right\} \qquad \{W_\phi\} \subset U(d)$$

**Proposition IV.2 (informal)**
The variance of a channel is always smaller than the variance of the corresponding stochastic unravelling, with equality holding if and only if the channel is unitary.

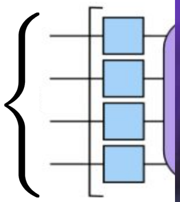
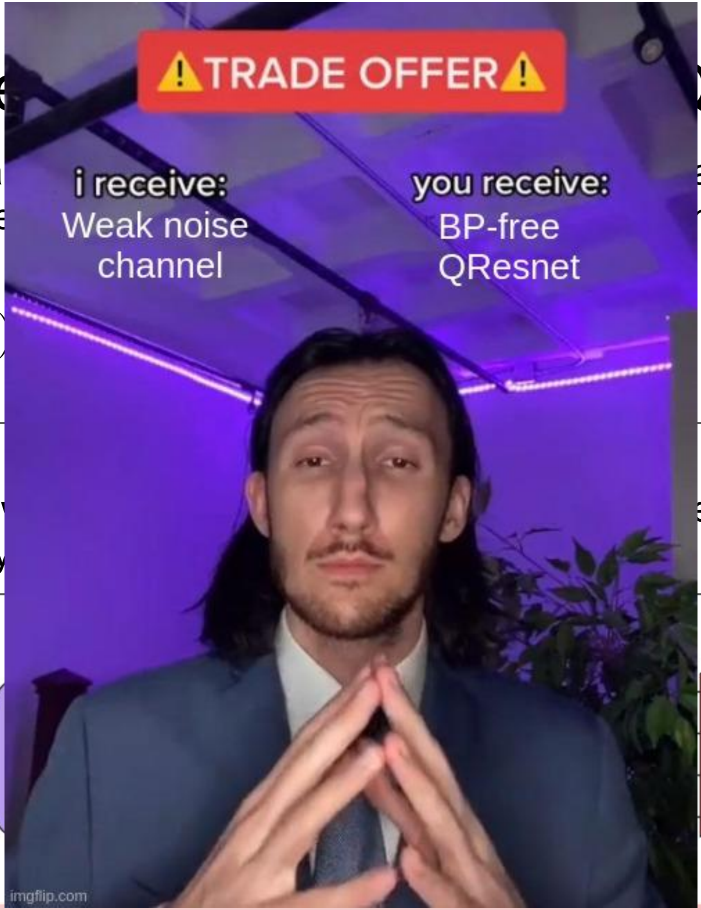
$\mathrm{Var}_\theta \Bigg\{$  $\Bigg\} \leq \mathrm{Var}_{\theta,\phi} \Bigg\{$  $\Bigg\}$

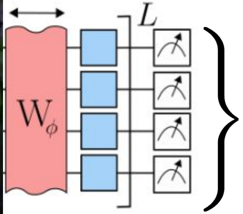
# Stochastic unrave… …QResNets

For any given noise map, we a… …a **stochastic unravelling**, i.e. as the **expectation value** of some… …n operators can even be **unitary**!

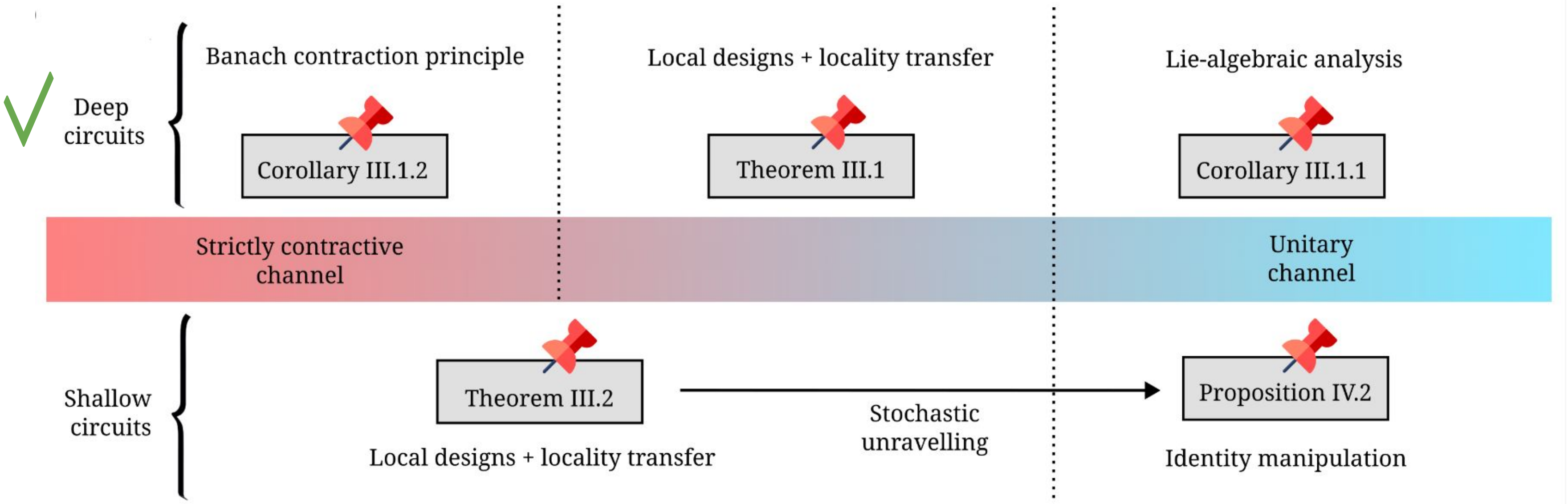
$$\mathcal{E}(\rho) \ldots \subset U(d)$$

**Proposition IV.2 (informal)**
The variance of a channel is al… …esponding stochastic unravelling, with equality holding if and only…

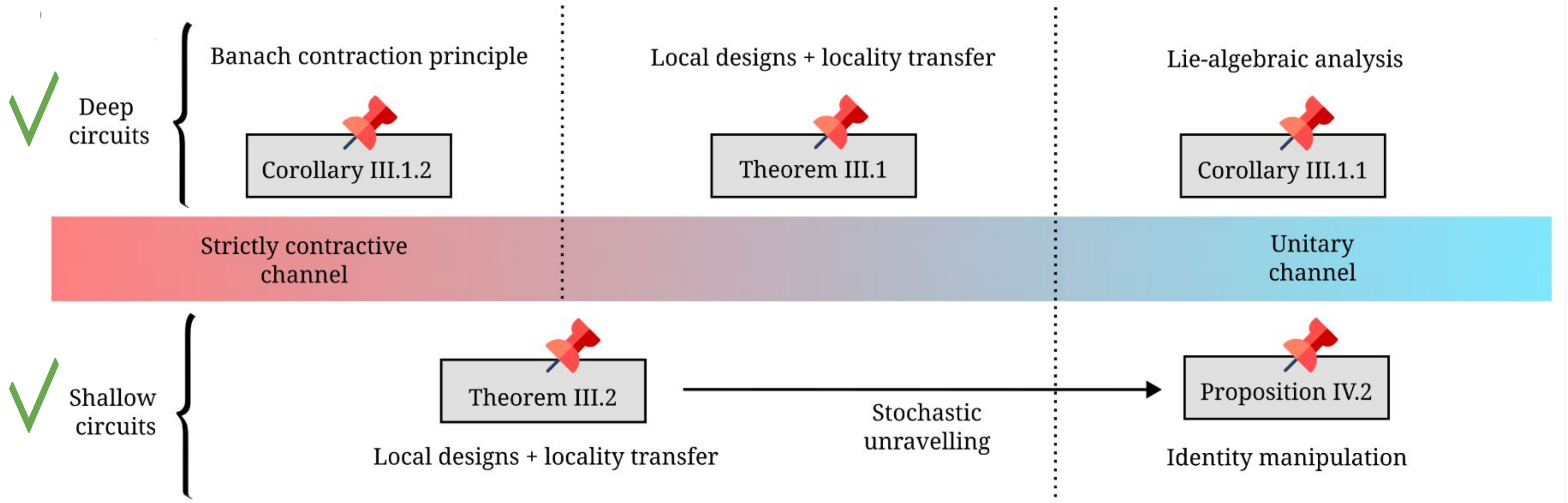
$\mathrm{Var}_\theta$  



# Summary


Deep circuits
Banach contraction principle
Corollary III.1.2
Local designs + locality transfer
Theorem III.1
Lie-algebraic analysis
Corollary III.1.1
Strictly contractive channel
Unitary channel
Shallow circuits
Theorem III.2
Local designs + locality transfer
Stochastic unravelling
Proposition IV.2
Identity manipulation

# Summary


Deep
circuits
Banach contraction principle
Corollary III.1.2
Local designs + locality transfer
Theorem III.1
Lie-algebraic analysis
Corollary III.1.1
Strictly contractive
channel
Unitary
channel
Shallow
circuits
Theorem III.2
Local designs + locality transfer
Stochastic
unravelling
Proposition IV.2
Identity manipulation

# Conclusions

Takeaway messages:

- **Concentration** for noisy circuits is more **intricate** than one can naively expect.
- Noise can **not only induce** barren plateaus, but **also prevent it** if engineered correctly.
- **Weak noise** maps are closely related to **small angle-like initialization** strategies.

# Thanks for your attention!

## Any question?

giulio.crognaletti@phd.units.it

# Extra: the theoretical framework

Generic quantum channels **do not have a group structure**, therefore the algebraic arguments do not apply. However, we can still divide the space of bounded operator based on the **unitary part of the circuit**.

Space of **bounded operators** on H

Space of **bounded, traceless operators** acting non-trivially on **k** qubits

$$\mathcal{B} = \bigoplus_{\kappa \in \{0,1\}^M} \mathcal{B}_\kappa$$

Dimension of $B_k$

Orthonormal basis of $B_k$

$$(\ell_A)_\kappa = \sum_{j=1}^{d_\kappa} \mathrm{Tr}[B_j A]^2$$

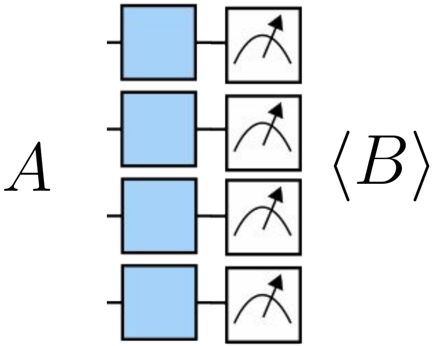
Locality vector

General operator in B

Since the unitary part is divided into qubits, these subspaces are linked to how "local" an operator is, and we can define a measure analogous to the previously mentioned **purity** in this context as well.

# Extra: locality and locality transfer matrix

The action of the unitary and local part of the circuit is **easily computed** using the tools from Weingarten calculus in the Lie-algebraic picture, and yields a result analogous to the previously mentioned.




Diagonal matrix

Locality vectors of the corresponding operators

$$(a, b) = a^t D^{-1} b = \sum_{\kappa} \frac{a_\kappa b_\kappa}{d_\kappa}$$

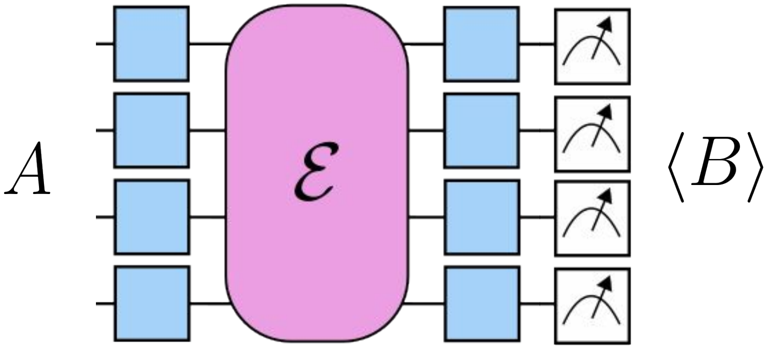
Convenient scalar product in this context

Dimension of the operator subspace k

This can be interpreted a **diagonal matrix** applied to the locality vectors.

# Extra: locality and locality transfer matrix

This work extends this idea, showing how the action of a **generic quantum channel** is described instead by a more **generic matrix**.




Generic, non-negative matrix

$$(a, b) = a^t D^{-1} T b = \sum_{\kappa} \frac{a_\kappa (Tb)_\kappa}{d_\kappa}$$

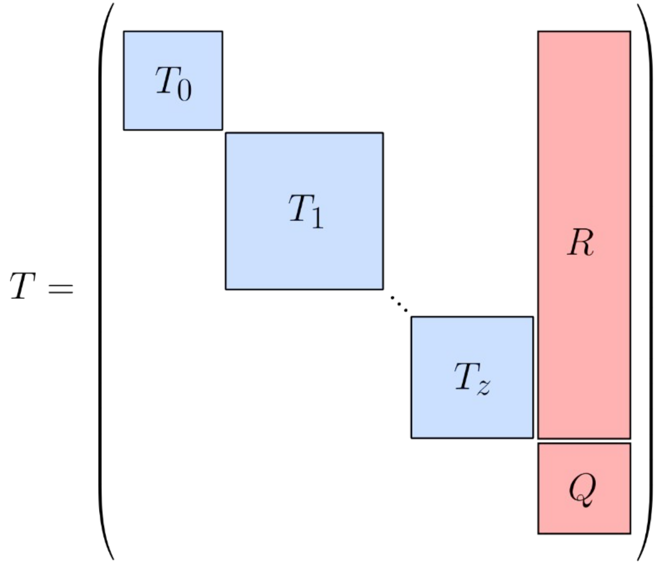
Formal definition of T

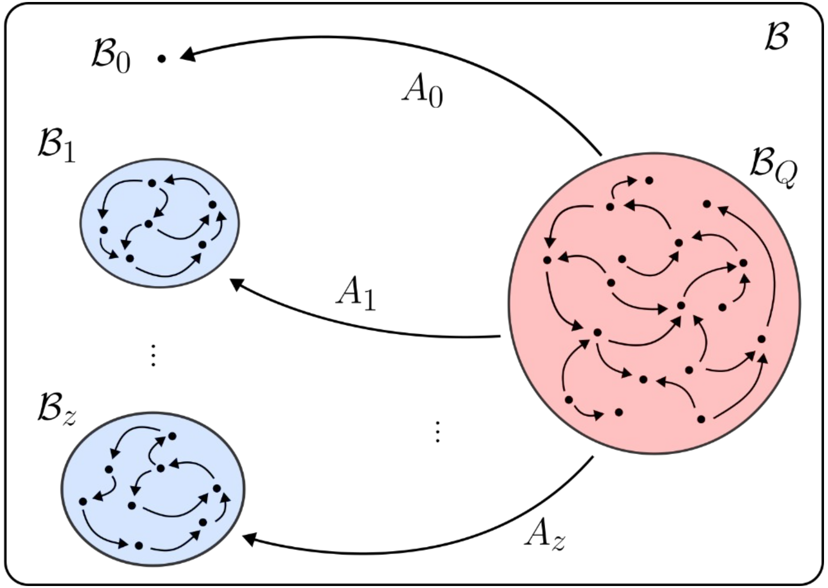
$$T_{\kappa,\lambda} = \frac{1}{d_\kappa} \sum_{j=1}^{d_\kappa} (\ell_{\mathcal{E}^\dagger(B_j)})_\lambda$$

We can now analyze the **spectral properties** of T to deduce the main properties of the **deep circuit** limit.

# Extra: irreducible blocks and absorption coefficients

The matrix T is most fruitfully examined **graphically** as an adjacency matrix for some graph. This graph represents the subspaces of B put in **communication** by the intermediate channel.



# Examples: Non-unital noise and entanglement

Employing the last result one can obtain a variance with **different scaling** as a function of noise strength based on the **entangling power** of each layer.

Specific example for the intermediate channel

P is a projector. A good approximation for rapidly entangling an slowly entangling gate

$$\mathcal{E}(\rho) = (1-p)W\rho W^\dagger + p\tilde{\rho} \quad \Rightarrow \quad \mathbb{V}^{\infty}_{\rho,H} = \left(\frac{p}{2-p} - p^2\right)(\ell_\rho, P\ell_H) + p^2(\ell_\rho, \ell_H)$$

Slowly entangling limit:

P≈1

$$\mathbb{V}^{\infty}_{\rho,H} = \frac{p}{2-p}(\ell_{\tilde{\rho}}, \ell_H)$$

**Linear** dependence on noise strength

Rapidly entangling limit:

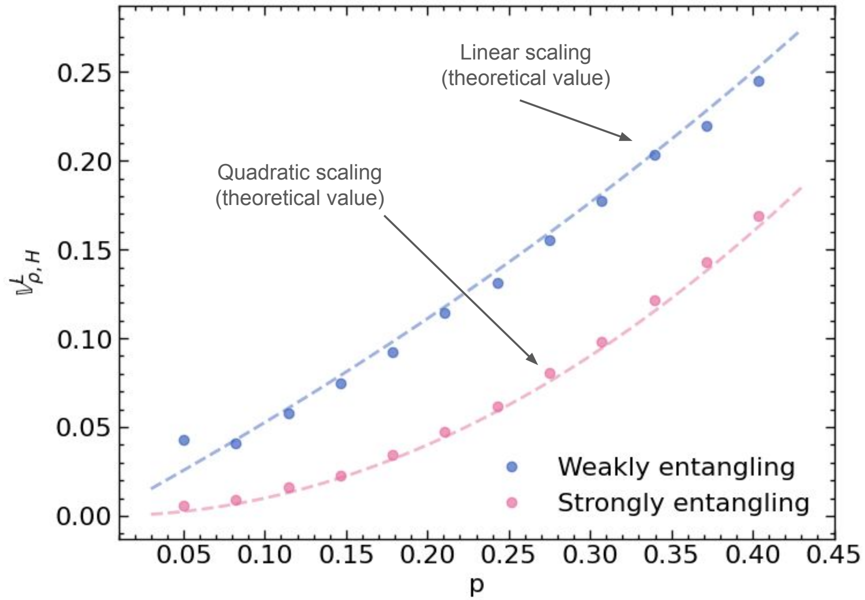
P≈$d_k$/d

$$\mathbb{V}^{\infty}_{\rho,H} = p^2(\ell_{\tilde{\rho}}, \ell_H)$$

**Quadratic** dependence on noise strength

# Examples: Non-unital noise and entanglement

Variance in the deep circuit limit as a function of noise strength




Speed of convergence to the deep circuit limit as a function of L