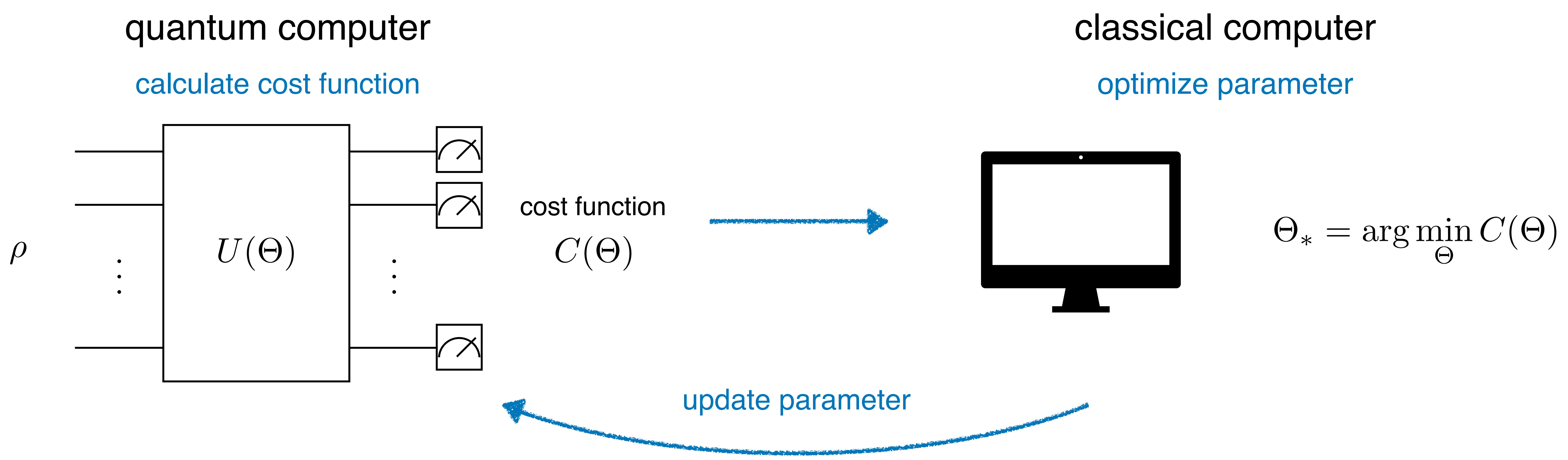
Gradient of cost function for general ansatz in variational quantum algorithm

Lento Nagano (ICEPP, The University of Tokyo)

Based on a work with R. Okubo (to appear in arXiv)

CERN openlab Technical Workshop (online)
Mar. 23, 2022

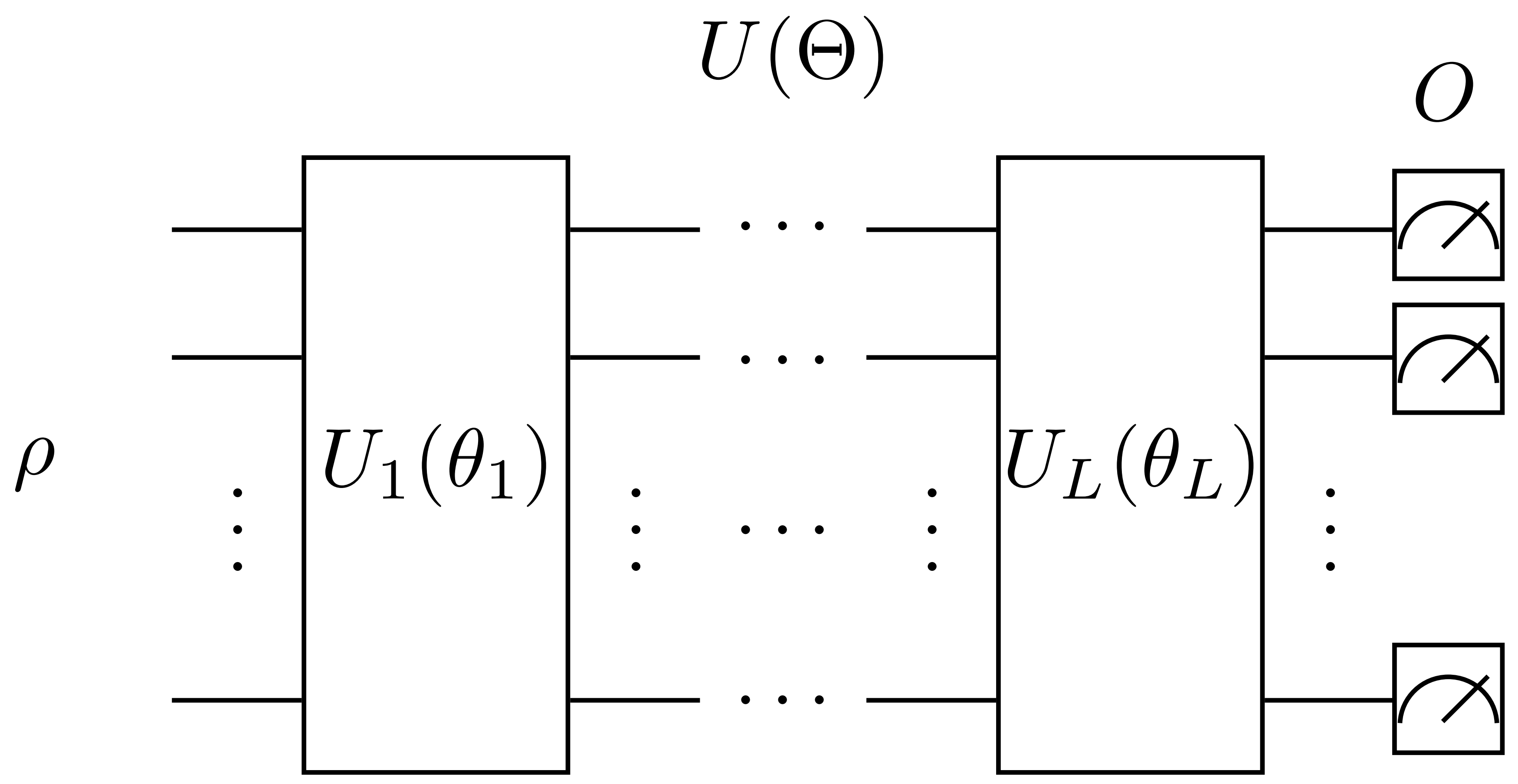
Variational quantum algorithm



- consider to be useful in NISQ era
- application:
  - variational quantum eigensolver (VQE): ground state energy in quantum many-body system
  - optimization problem (quantum alternative optimization algorithm; QAOA)
  - quantum machine learning

Ansatz for variational quantum algorithm

- cost function: $C(\Theta) = \text{Tr}\left[OU(\Theta)\rho U^\dagger(\Theta)\right]$
  - ρ: initial state
  - $U(\Theta)$: ansatz (parametrized unitary)
  - O: observable



- layered structure ansatz $U = \prod_{l=1}^{L} U_l(\theta_l)$
  - hardware efficient: $U_l = \prod_{k=1}^{K} e^{-i\theta_{l,k}\sigma_{l,k}} W_{l,k}$ ($\sigma_{l,k}$: Pauli matrix, $W_{k,l}$: unprametrized gate)
  - variational Hamiltonian ansatz: $U_l = \prod_{k=1}^{K} e^{-i\theta_{l,k}H_{l,k}}$ ($H_{l,k}$: traceless Hermitian operator)
- We don't assume detailed structure for each U_l

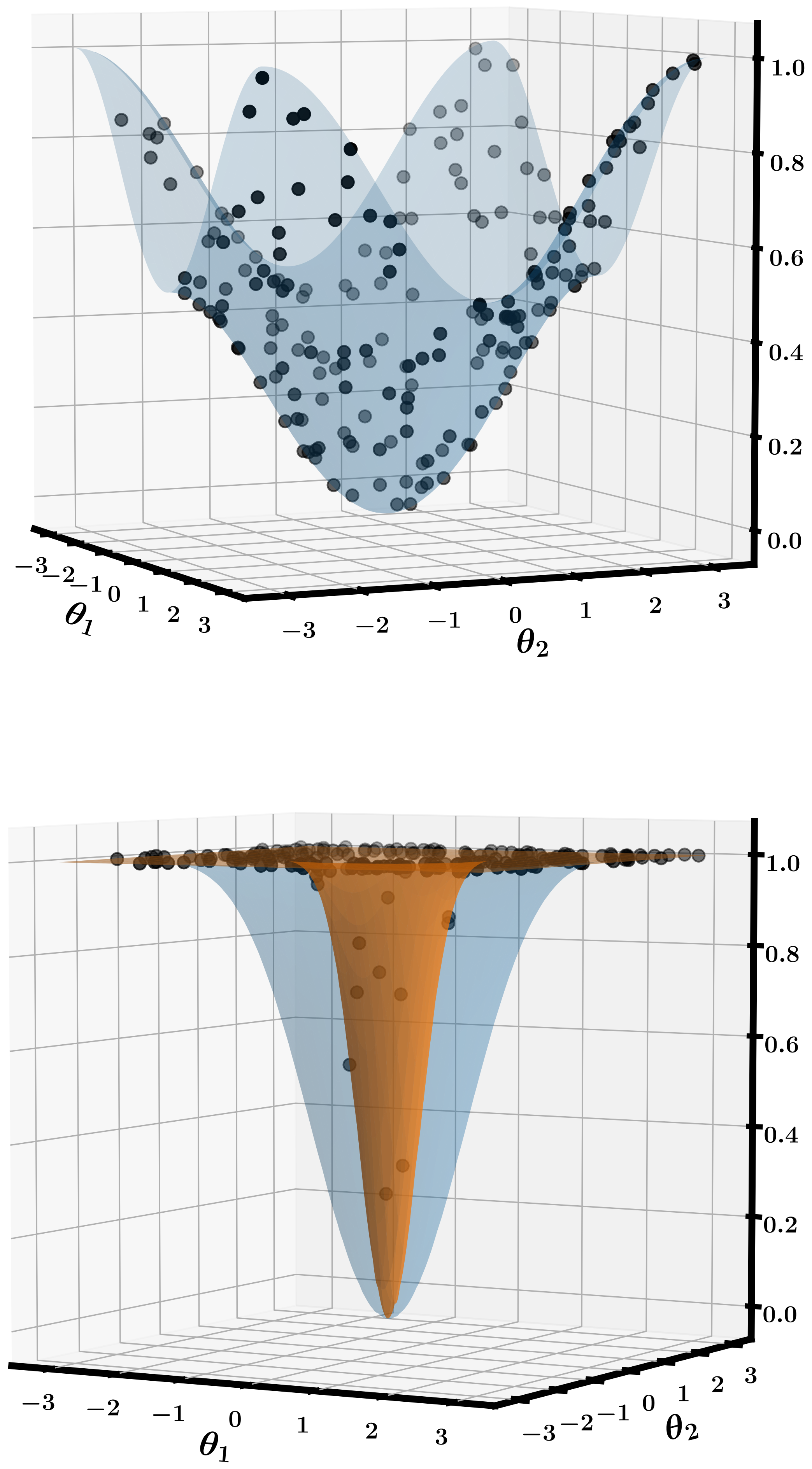
Barren plateau

- take random initial parameters
- gradient of cost function: $\partial C := \partial C / \partial \theta_{k,l}$
- definition of barren plateau: [McClean et.al. '18]

$$\forall c, P(|\partial C| \geq c) \in \mathcal{O}(2^{-n}) \quad (n\text{: \# of qubits})$$

- Chebyshev's inequality: $P(|\partial C| \geq c) \leq \mathbb{E}\left[(\partial C)^2\right]/c^2$

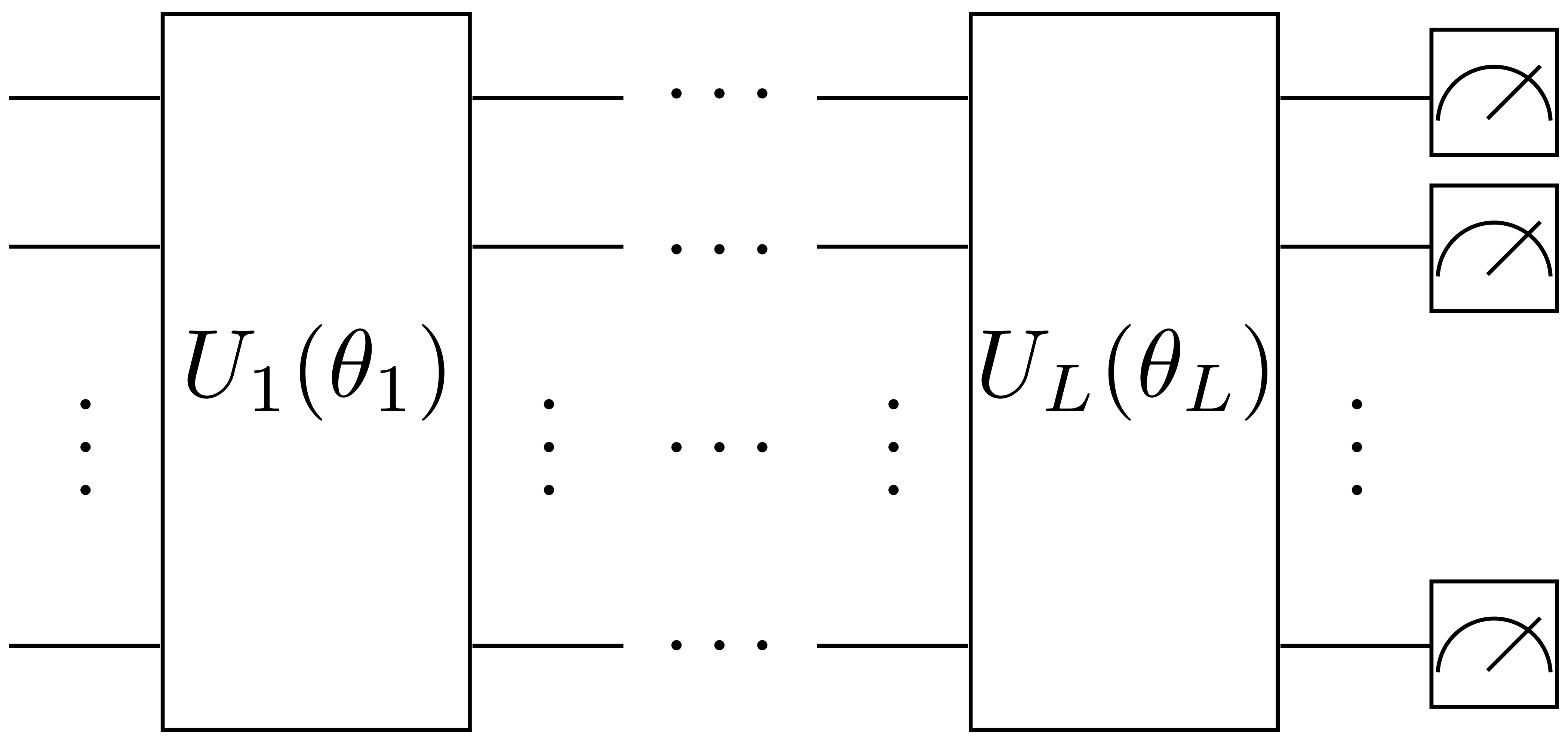
  $\rightarrow$ sufficient condition for BP: $\mathbb{E}\left[(\partial C)^2\right] \in \mathcal{O}(2^{-n})$
- known results (for ansatz with certain structure):
  - maximally expressive ansatz $\Rightarrow$ BP [McClean et.al. '18 , Cerezo et.al. '20]
  - more expressive ansatz $\Rightarrow$ flatter landscape [Holmes et.al. '21]
- We show that these results hold independent of detailed structure for U_l



[figure from Cerezo et.al. '20]

Ansatz expressibility

- deep hardware efficient ansatz
  → maximally expressive ansatz (unitary 2-design) [e.g., McClean et.al. '18]
- $\mathscr{A}(U)$:= extent to which given ansatz U can span
- "expressibility" $\epsilon(U) := \mathscr{A}_{\text{2-design}} - \mathscr{A}(U)$
  - expressive ansatz ⇒small ϵ
  - $\epsilon = 0 \Rightarrow$ 2-design
- We will evaluate second moment $\mathbb{E}\left[(\partial C)^2\right]$
  - with assuming 2-design
  - without assuming 2-design
    →upper bound on $\mathbb{E}\left[(\partial C)^2\right]$

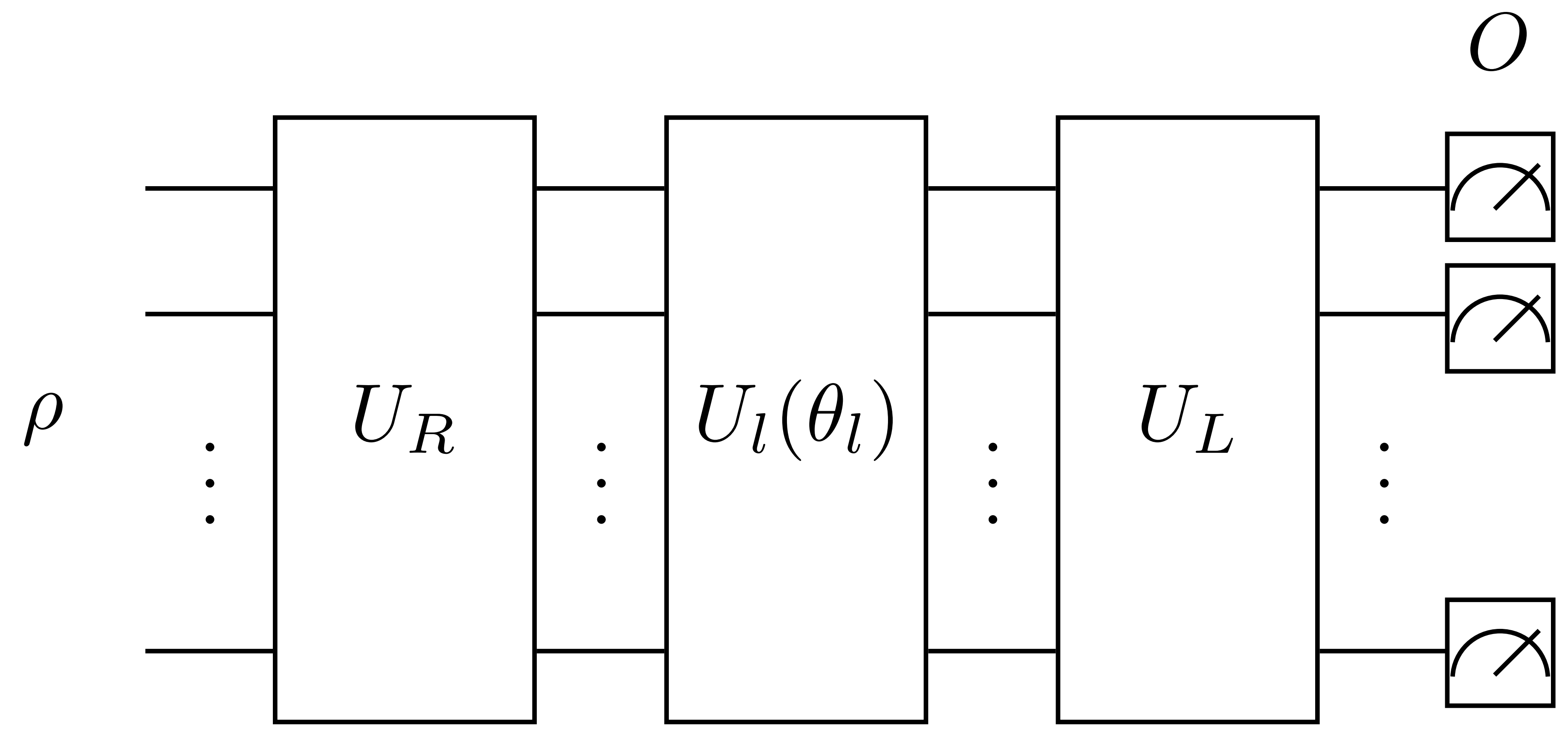


BP for highly expressive ansatz

- consider derivative w.r.t. $\theta_{l,k}$: $\partial C := \partial C / \partial \theta_{l,k}$ with $\theta_l = (\theta_{l1}, \cdots, \theta_{lk}, \cdots \theta_{lK})$
- divide ansatz as $U(\Theta) = U_L U_l(\theta_l) U_R$
- dimension of Hilbert space: $d = 2^n$
- fully explicit expression for second moment assuming U_L, U_R form unitary 2-design

$$\mathbb{E}\left[(\partial C)^2\right] = \frac{2d}{(d^2-1)^2}\left(\mathrm{Tr}\left[\rho^2\right] - \frac{\mathrm{Tr}\left[\rho\right]}{d}\right)\left(\mathrm{Tr}\left[O^2\right] - \frac{\mathrm{Tr}\left[O\right]}{d}\right) F(U_l)$$

- formula to see BP or not BP
- example: symmetry preserving ansatz
  - BP/non-BP depending on a sector in which ρ lives



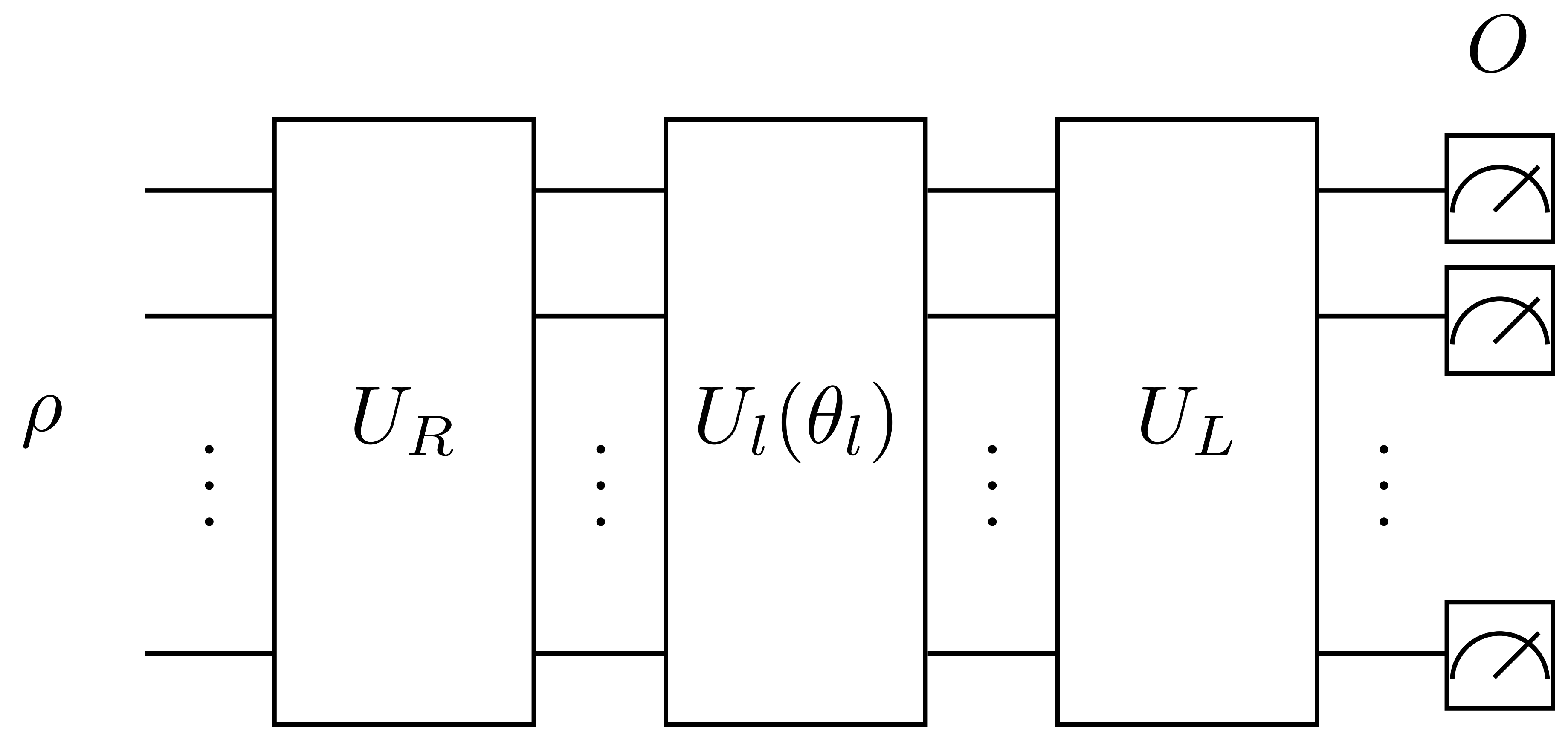
Upper bound on second moment

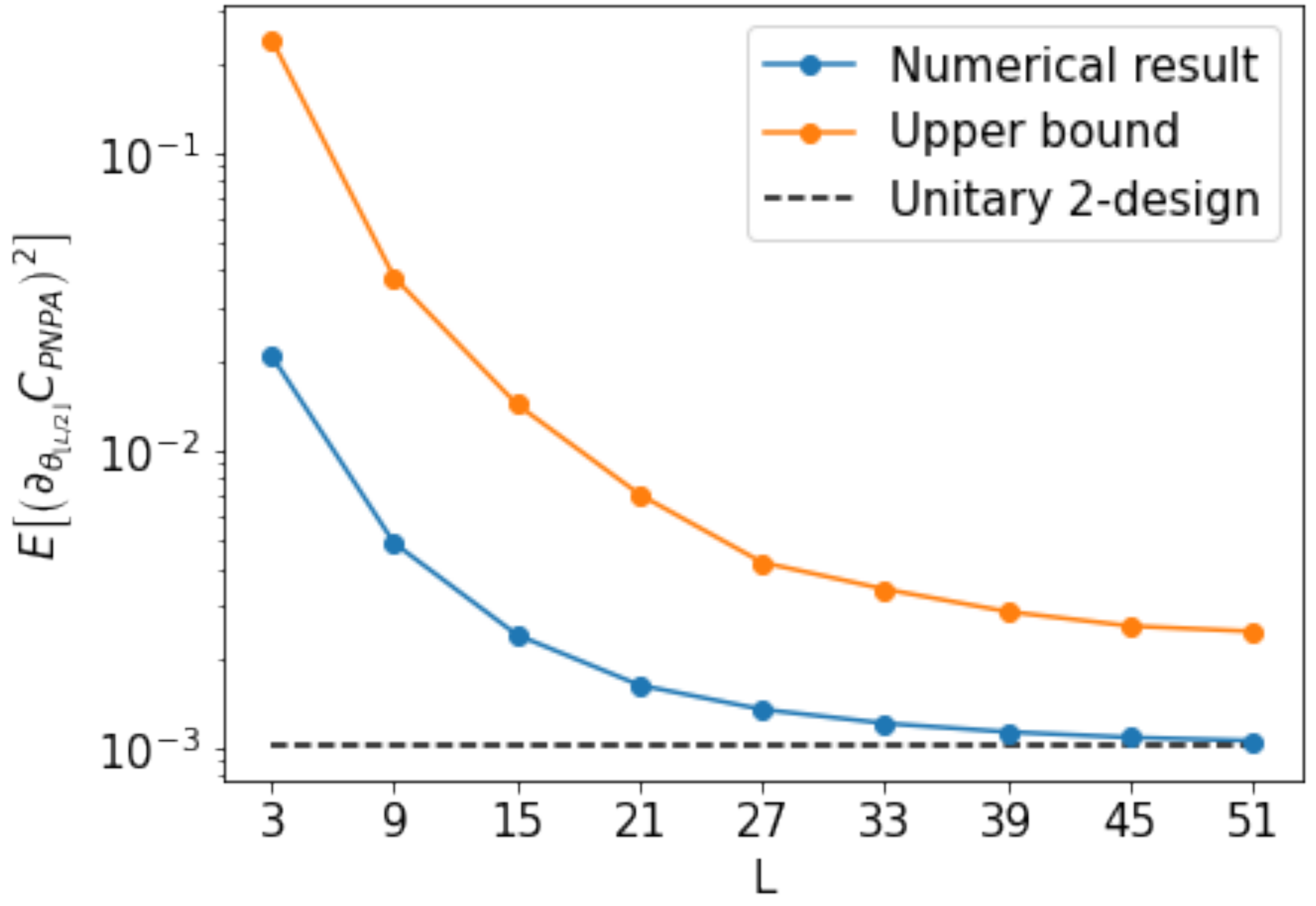
- evaluate second moment without 2-design assumption
- expressibility for U_L, U_R : ϵ_L, ϵ_R
- second moment is upper bounded by ϵ_L, ϵ_R :

$$\mathbb{E}\left[(\partial C)^2\right] \leq a_0 + a_1\epsilon_L + a_2\epsilon_R + a_3\epsilon_L\epsilon_R$$

  (a_i: positive constants which depend on d, ρ, U_l, O)

- expressive ansatz → small gradient
- numerical check
  for particle number preserving ansatz [Gard-Zhu, et. al. '19]
  - correct, but not so tight
  - observe "correlation" between right and left hand side





Summary

- Barren plateau can be diagnosed by evaluating second moment of gradient
- We evaluated second moment with/without 2-design assumption
  - to have explicit d, ρ, U_M, O dependence
  - to have upper bound on second moment
- Future directions:
  - application of our formula (e.g. circumvent BP)
  - tighter bound for second moment
  - analytically prove correlation?
  - effect of noise?