Exact chiral symmetry with quantum signal processing

Hersh Singh
Fermi National Accelerator Laboratory

Frontiers of Lattice Fermions, YITP, Kyoto University • July 01, 2026

with H. Lamm, A. Roggero, L. Spagnoli • 2607.xxxxx

Where this is going

- Lattice QCD is moving toward a new frontier, **quantum simulation**: real-time dynamics, finite density, things Monte Carlo cannot reach.
- An old problem reappears in this new setting: how do we put **chiral fermions** on the lattice, now as a **Hamiltonian** we evolve on a quantum computer?
- Two classic answers give **exact** lattice chiral symmetry: **overlap** and **domain-wall** fermions.

> The question of this talk:
>
> *what is the best quantum algorithm for each, and*
> *what does comparing them tell us about the physics?*

Plan

1. **Exact chiral symmetry on the lattice**
   Ginsparg–Wilson, overlap, domain-wall
2. **A quantum-computing toolkit**
   block encoding & quantum signal processing
3. **Simulating chiral fermions**
   the construction and its cost

Exact chiral symmetry on the lattice

Two faces of the chiral fermion problem

The "easy" problem

A **global** chiral symmetry, with a 't Hooft anomaly. Think **QCD**.

- Perfectly fine as a global symmetry (it just cannot be gauged).
- Physical: the $\pi^0 \to \gamma\gamma$ rate.
- **Solved on the lattice**: overlap / domain-wall realize it *exactly*.

The "hard" problem

A **gauged** chiral symmetry: a chiral gauge theory. Think **electroweak**.

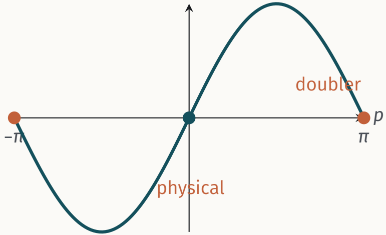
- Anomaly cancels, then the chiral symmetry is gauged.
- **No known** nonperturbative lattice construction.
- We don't yet know how to even *define* it.

This talk lives in the **easy** problem: bring exact lattice chiral symmetry to a **quantum computer**.

The problem, in one slide (you've seen this all week)

- A massless Dirac fermion has a **chiral symmetry** $\psi \to e^{i\theta\gamma_5}\psi$, generated by $\{\not{D}, \gamma_5\} = 0$.
- Naively discretize $\not{D}$: the propagator $\sim \sum_\mu \gamma^\mu \sin p_\mu$ has **extra zeros** at the corners of the Brillouin zone.

  $\Rightarrow$ **fermion doubling**: $1 \to 2^d$ species.
- **Nielsen–Ninomiya** [Nielsen, Ninomiya 1981]: you cannot have all of *locality* + *no doublers* + *exact* $\{D, \gamma_5\} = 0$ at once.
  - **Wilson's fix** [Wilson 1977]: add $\sum_\mu (1 - \cos p_\mu)$ to lift the doublers, but this **explicitly breaks** chiral symmetry, restored only as $a \to 0$.



$D(p) \sim \sin p$ vanishes at $p = 0$ *and* $p = \pi$.

Is there a better way?

There are infinitely many ways to discretize a fermion.
Wilson's is just one, and it sacrifices chiral symmetry completely.

So: **can we keep an *exact* chiral symmetry on the lattice, with no doublers, at finite lattice spacing?**

Remarkably, yes. The key is to ask for a *smarter* version of $\{D, \gamma_5\} = 0$.

▶ *see Shamir's and Sen's talks for the full doubling / Nielsen–Ninomiya story*

The Ginsparg–Wilson relation (1982)

- Nielsen–Ninomiya forbids *exact* $\{D, \gamma_5\} = 0$. Ginsparg & Wilson asked: what is the **mildest** way to relax it?
- Allow a right-hand side that vanishes in the continuum, but is local:

The Ginsparg–Wilson relation

$$\{D, \gamma_5\} = a\, D\, \gamma_5\, D$$

- The breaking is $O(a)$ and **local**

Lay dormant for ~ 15 years: nobody had a D satisfying it.

Then (late 90s): **overlap** and **domain-wall** fermions both realize it.

Ginsparg, Wilson 1982

A genuine lattice chiral symmetry (Lüscher)

- The GW relation looks like a cheat, but it hides an **exact** symmetry.
- Define a **modified chirality** operator

$$\hat{\gamma}_5 = \gamma_5(\mathbb{1} - aD), \qquad \hat{\gamma}_5 \xrightarrow{a\to 0} \gamma_5.$$

- Then GW is *equivalent* to the statement that the action is invariant under

$$\delta\psi = \hat{\gamma}_5\,\psi, \qquad \delta\bar{\psi} = \bar{\psi}\,\gamma_5.$$

The lattice theory has an **exact** chiral symmetry at finite a, just generated by $\hat{\gamma}_5$ instead of γ_5.
No fine-tuning, correct anomaly, no doublers.

Lüscher 1998

Solution 1: the overlap operator

- Neuberger wrote down an explicit GW solution from the Wilson operator H_{w}:

Overlap Dirac operator

$$D_{\mathrm{ov}} = \mathbb{1} + \gamma^5\, \varepsilon(H_{\mathrm{w}}), \qquad \varepsilon(X) = \frac{X}{\sqrt{X^\dagger X}}, \qquad H_{\mathrm{w}} = \gamma^5 D_W$$

- The **matrix sign function** $\varepsilon(H_{\mathrm{w}})$ is the whole story: $\varepsilon(H_{\mathrm{w}})^2 = \mathbb{1}$ is exactly what makes GW hold.
- **Price:** $\varepsilon(H_{\mathrm{w}})$ is a *non-local* operator; it couples every site to every other site.

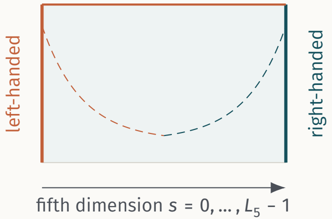
On the lattice we never have $\varepsilon(H_{\mathrm{w}})$ exactly: we *approximate* it. Hold that thought.

▸ *see Kikukawa's, Shamir's, Fukaya's talks for more on overlap fermions*

Neuberger 1998

Solution 2: domain-wall fermions

- Kaplan's idea: add an **extra dimension** of size L_5.
- A $(d+1)$-dim Wilson fermion with a mass *defect* traps **chiral zero-modes on the two boundaries**:
  - left wall → left-handed mode
  - right wall → right-handed mode
- As $L_5 \to \infty$ the walls decouple ⇒ **exact** chiral symmetry, locally in d dimensions.
- Locality is bought back at the cost of **extra qubits/sites**.



Boundary modes, exponentially localized.

Kaplan 1992, 1993

The two are secretly the same operator

- Integrate out the bulk of a domain-wall stack: the **boundary** theory is exactly an overlap fermion.
- At *finite* L_5, the domain-wall operator approximates the sign function:

$$\varepsilon(H_{\mathrm{w}}) \longrightarrow \tanh(L_5 H_{\mathrm{w}}),$$

  which becomes exact as $L_5 \to \infty$.
- So the **fifth dimension** L_5 is nothing but a **tool to approximate** $\varepsilon(H_{\mathrm{w}})$.

Overlap = exact sign function, non-local, few sites.
Domain-wall = approximate sign function via L_5, local, many sites.

Narayanan, Neuberger 1993 • Kikukawa, Noguchi 1999

For a quantum computer: go Hamiltonian

- Quantum computers evolve states in real time: $|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle$.
- So we want these fermions as a **Hamiltonian**, not a Euclidean action.
- Both formulations port over. The single-particle Wilson Hamiltonian

$$h_{\mathrm{w}} = \sum_i i\,\Gamma^i \delta_i + \Gamma^0\left(m - \tfrac{r}{2}\Delta\right)$$

  is **local** and **gapped** (smallest |eigenvalue| = m).
- **Domain-wall**: just the Wilson Hamiltonian in $d+1$ dimensions, h_{dw}, an honest local lattice Hamiltonian.

Domain-wall is local and obvious. But what *is* the overlap Hamiltonian?

The overlap Hamiltonian (Creutz–Horvath–Neuberger 2002)

- A natural single-particle ansatz, built directly from the Wilson h_{w}:

Overlap Hamiltonian

$$h_{\mathrm{ov}} = \Gamma^0 D_{\mathrm{ov}} = \Gamma^0 + \varepsilon(h_{\mathrm{w}}), \qquad \varepsilon(h_{\mathrm{w}}) = h_{\mathrm{w}} / \sqrt{h_{\mathrm{w}}^\dagger h_{\mathrm{w}}}.$$

- Choosing $0 < m < 2r$ gives a **single massless Dirac fermion**, no doublers.
- It carries an **exact lattice chiral symmetry**, generated by a **modified chirality**:

$$Q_5 = \psi^\dagger \hat{\gamma}_5 \psi, \qquad \hat{\gamma}_5 = \tfrac{1}{2}\gamma_5\big(1 + \Gamma^0 \varepsilon(h_{\mathrm{w}})\big), \qquad [\hat{\gamma}_5, h_{\mathrm{ov}}] = 0.$$

- $\hat{\gamma}_5 \to \gamma_5$ in the continuum; at finite a it is the GW-deformed chirality.

An **exact** chiral symmetry on the lattice Hamiltonian, at the price of the nonlocal sign function $\varepsilon(h_{\mathrm{w}})$.

Why the overlap Hamiltonian is subtle

- The overlap operator was born **Euclidean**: a kernel in an *action* $\bar{\psi} D_{\mathrm{ov}} \psi$. A **Hamiltonian** is a different beast: a Hermitian generator of real-time evolution at fixed time. **No canonical recipe** connects the two.
- Domain-wall is a local $(d+1)$-dim theory: its transfer matrix gives a Hamiltonian for free. For overlap, what the analog even *is* remains an **open, active question**. [Clancy 2024 • HS 2025 • Chattejee, Pace, Shao 2024 • Gioia, Thorngren 2025, • Misumi 2025]
- One proposal, $h_{\mathrm{ov}} = \Gamma^0 + \varepsilon(h_{\mathrm{w}})$, trades the extra dimension for **all-to-all** interactions: no ultralocality [Creutz, Horvath, Neuberger 2002]

  ▸ *see Kikukawa's talk: how the CHN overlap and Hamiltonian domain-wall fermions relate*
- Its chiral charge is a **dynamical** $\hat{\gamma}_5 = \frac{1}{2}\gamma_5(1 + \Gamma^0 \varepsilon(h_{\mathrm{w}}))$: the symmetry generator *depends on the Hamiltonian*, unlike the fixed continuum γ_5.

Compact vs. noncompact

Euclidean

$$\psi \to e^{i\theta\hat{\gamma}_5}\psi, \quad \bar{\psi} \to \bar{\psi}\, e^{i\theta\gamma_5}$$

with $\hat{\gamma}_5 = \gamma_5(1 - aD)$, $\hat{\gamma}_5^2 = 1$, eigenvalues ± 1.

θ is a 2π-periodic angle $\Rightarrow$ a **compact** $U(1)$.

Anomaly is in the measure:
$\mathrm{Tr}\,\hat{\gamma}_5 = \text{index} \in \mathbb{Z}$.

Hamiltonian

$$Q_5 = \psi^\dagger\, \hat{\gamma}_5\, \psi, \quad \hat{\gamma}_5 = \tfrac{1}{2}\gamma_5(1 + \Gamma^0 \varepsilon(h_{\mathrm{w}}))$$

Now $\psi^\dagger$ is **conjugate** to ψ, and $\hat{\gamma}_5^2 \neq 1$: its spectrum is *not* ± 1.

So $e^{i\theta Q_5}$ **never closes up** $\Rightarrow$ a **noncompact** $\mathbb{R}$.

The compact Euclidean chiral $U(1)$ becomes a **noncompact** $\mathbb{R}$ in the Hamiltonian

▸ *see talks by Shao, Gioia, Thorngren, Zakharov, Ueda for other exact chiral symmetries in Hamiltonian formulation*

A quantum-computing toolkit

What a quantum computer actually does for us

- State of n qubits = a vector in $\mathbb{C}^{2^n}$, an **exponentially large** Hilbert space.
- The machine applies **unitary** gates. Our target is real-time evolution

$$|\psi(t)\rangle = e^{-iHt}\,|\psi(0)\rangle.$$

- For lattice fermions, H is a $2^n \times 2^n$ matrix we never write down: we only know how to *act* with it.

The core challenge

H is **Hermitian, not unitary**. A quantum computer can only apply *unitaries*.

How do we apply a function of H, like e^{-iHt}, or a sign function $\varepsilon(H)$?

First instinct: just Trotterize?

- The standard near-term recipe: split $H = \sum_k H_k$ into easy pieces and

$$e^{-iHt} \approx \left(\prod_k e^{-iH_k t/n}\right)^n.$$

- **Great when H is local**: shallow, local circuits, few ancillas. Deservedly popular.

But for exactly-chiral fermions it falls short

- **Precision is costly:** the error is *polynomial* in $1/\epsilon$ (first order $\sim t^2/\epsilon$), versus the optimal $t + \log\frac{1}{\epsilon}$.
- **Nothing to split:** the overlap $\varepsilon(h_{\mathrm{w}})$ is *all-to-all*, not a sum of a few easy terms.
- **It breaks the symmetry:** each Trotter step adds chiral-symmetry violation with no single controlled knob.

Why quantum signal processing?

We want one method that fixes all three Trotter problems at once. That method is **QSP**.

- **Optimal in precision and time:** cost ~ $t + \log \frac{1}{\epsilon}$, **exponentially** better accuracy than Trotter, linear in t, matching the lower bound.
- **Functions of H, directly:** it realizes *any* polynomial $p(H)$, including the sign function $\varepsilon(h_{\mathrm{w}})$. The nonlocality becomes a **feature**, not an obstacle.
- **One knob for the symmetry:** the polynomial degree M sets the GW violation ϵ_e: a single, controlled dial.

One framework gives both the time evolution e^{-iHt} and the overlap's $\varepsilon(h_{\mathrm{w}})$, optimally and with controlled chiral symmetry. Next: how it works.

Idea 1: hide H inside a unitary ("block encoding")

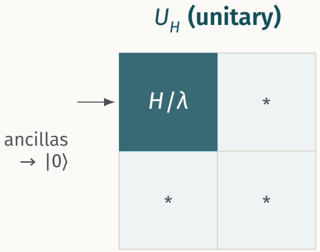
- Embed the (rescaled) H as the **top-left corner** of a bigger unitary U_H, using a few **ancilla** qubits:

$$U_H = \begin{pmatrix} H/\lambda & * \\ * & * \end{pmatrix}$$

- Concretely: prepare ancillas in $|0\rangle$, apply U_H, measure ancillas in $|0\rangle$:

$$\langle 0^a | \, U_H \, | 0^a \rangle \;=\; H/\lambda.$$

- For our *local* Wilson h_{w} this is cheap: a sum of a few unitaries (shifts × γ-matrices), cost $O(Q)$.



Low, Chuang 2019 • Gilyén et al. 2019

Idea 2: quantum signal processing (QSP)

- Given a block encoding of H, **interleave** it with simple ancilla rotations (phases $\phi_1, \dots, \phi_M$):

$$U_H \quad \to \quad R(\phi_M)\, U_H\, R(\phi_{M-1}) \cdots U_H\, R(\phi_1).$$

- Magic: this new circuit block-encodes a **polynomial** $p(H)$ of the original:

$$\langle 0|\,(\text{circuit})\,|0\rangle \;=\; p(H), \qquad \deg p = M.$$

- Choose the phases $\{\phi_k\} \Rightarrow$ choose *any* polynomial you like (with $|p| \leq 1$).

QSP turns a block encoding of H into a **"polynomial-of-H machine"**:
pick a degree-M polynomial p, pay M uses of U_H, and get $p(H)$.

Low, Chuang 2019 • Gilyén et al. 2019

Two polynomials we want

Time evolution

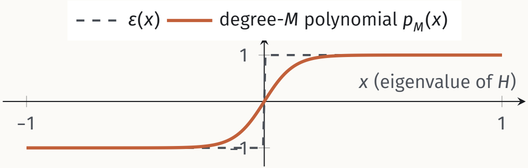
$$e^{-iHt} \approx \sum_{k=0}^{M} c_k\, T_k(H), \quad \text{degree } M = O(\lambda t + \log \tfrac{1}{\epsilon}).$$

Asymptotically optimal simulation.

The sign function

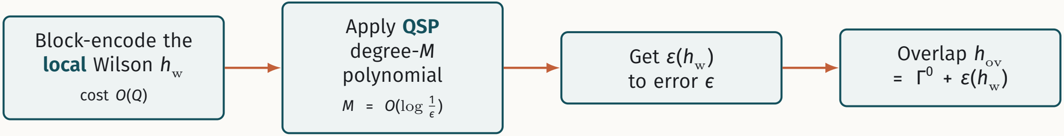
$\varepsilon(H) \approx p_M(H)$, degree $M = O(\kappa^{-1} \log \frac{1}{\epsilon})$ (κ = gap).

Exactly what overlap fermions need.



Simulating chiral fermions

The recipe for overlap fermions



- The hard, non-local object (the sign function) is **never built by hand**. QSP manufactures it from the cheap local operator.
- Then standard QSP time-evolution runs $e^{-iH_{\mathrm{ov}}t}$.
- Gauge fields slot in cleanly: replace shifts by *covariant* shifts $A_i \to A_i\, U_{(x,i)}$. Asymptotic scaling unchanged.

From one particle to the many-body Hamiltonian

- Everything so far is the **single-particle** h, a $Q \times Q$ matrix. The trick: **promote** it to an operator on only

$$q = \lceil \log_2 Q \rceil = O(\log Q) \text{ qubits.}$$

  All the QSP and sign-function work lives in this **exponentially compressed** register.

- But we evolve the **many-body** Hamiltonian on the full Fock space:

$$H = \psi^\dagger h \, \psi \qquad \text{on } Q \text{ qubits (Jordan–Wigner).}$$

- A standard construction wraps the single-particle block encoding with $O(Q)$ extra gates and one call to h:

$$O(Q) \text{ gates}, \quad O(\log Q) \text{ ancillas}$$

> The sign function is built in a $\log Q$-qubit space; second-quantization adds the $O(Q)$ factor, the source of the linear system-size scaling.

What does it cost?

- The Wilson Hamiltonian is **gapped**: smallest |eigenvalue| = m, independent of system size. So κ is $O(1)$ and

$$M = O\left(\log \tfrac{1}{\epsilon}\right) \quad \text{(only \textit{logarithmic} in the accuracy).}$$

- Each QSP step is one use of the $O(Q)$ block encoding, so applying $\varepsilon(h_{\mathrm{w}})$ costs

$$O\left(Q \, \log \tfrac{1}{\epsilon}\right).$$

- Full real-time evolution of the overlap fermion:

$$O\Big(\underbrace{Q \log \tfrac{1}{\epsilon_e}}_{\text{sign fn}} \cdot \Big(Qt + \log \tfrac{1}{\epsilon_t}\Big)\Big).$$

Q = number of fermionic modes; ϵ_e = sign-function error, ϵ_t = time error.

sign-function scaling: Low, Chuang 2017 • Gilyén et al. 2019

The chiral symmetry is broken only *softly*

- Truncating to degree M means $p_M(h_{\mathrm{w}})$ is *not exactly* a sign function:

$$\|p_M(h_{\mathrm{w}})^2 - \mathbb{1}\| \;\leq\; 2\epsilon_e.$$

- One short computation (anticommutators of γ's) turns this into a bound on the **Ginsparg–Wilson violation**:

$$\left\| \{D, \gamma_5\} - D\gamma_5 D \right\| \leq 2\epsilon_e \qquad \Longrightarrow \qquad \left\| [\,\hat{\gamma}_5,\, h\,] \right\| \leq 2\epsilon_e.$$

- The exact chiral symmetry is recovered **exponentially fast**: ϵ_e is what we dial with the polynomial degree M.

Chiral symmetry violation is **controlled directly** by
the QSP error ϵ_e, and costs only $\log(1/\epsilon_e)$ to shrink.

And domain-wall fermions?

- Domain-wall stays a **local** Hamiltonian in $d+1$ dimensions: no sign function to approximate.
- Locality is a real asset on a quantum computer
- But you pay in **memory**: every one of the Q modes is copied L_5 times.
- The residual GW violation is set by the wall separation: $\sim e^{-cL_5}$.

Two knobs (M for overlap, L_5 for domain-wall), both controlling the same chiral-symmetry violation. Coincidence?

local simulation: Haah et al. 2018 • Rhodes et al. 2024

The punchline

The scoreboard

	Overlap	**Wilson**	**Domain-wall**
Qubits (memory)	Q	Q	$Q\,L_5$
Gates: block-encode h	$Q \log \frac{1}{\epsilon_e}$	Q	$Q\,L_5$
Gates: time evolution	$Q \log \frac{1}{\epsilon_e} (Qt + \log \frac{1}{\epsilon_t})$	$Qt \ \log \frac{Qt}{\epsilon_t}$	$QL_5 t \ \log \frac{QL_5 t}{\epsilon_t}$
Chiral (GW) violation	ϵ_e	$O(1)$	e^{-cL_5}

Wilson & domain-wall scalings: Rhodes et al. 2024

- **Wilson**: cheapest, but chiral symmetry broken at $O(1)$.
- **Overlap**: fewest qubits, deeper circuits (the $\log \frac{1}{\epsilon_e}$ overhead).
- **Domain-wall**: shallow local circuits, but a factor L_5 in memory.

Stare at the last row.

Reading the table: $M \leftrightarrow L_5$

- Domain-wall GW violation ~ e^{-cL_5}. Overlap GW violation ~ ϵ_e.
- Set them equal:

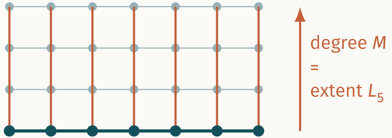
$$\boxed{\epsilon_e \sim e^{-cL_5} \qquad \Longleftrightarrow \qquad L_5 \sim \log\frac{1}{\epsilon_e} \sim M}$$

- The QSP **polynomial degree** M plays *exactly* the role of the domain-wall **fifth-dimension extent** L_5:
  - both control the chiral-symmetry violation,
  - both add the *same* logarithmic cost,
  - both interpolate between local and fully non-local.

When QSP approximates $\varepsilon(h_w)$ with a degree-M polynomial,
it is **secretly reconstructing the extra dimension** of domain-wall fermions.

Why this is the *same* extra dimension

- h_w is **nearest-neighbor** (sparsity $s \sim d$).
- A degree-M polynomial $p_M(h_w)$ reaches at most M sites away: sparsity $\sim d\,M$.
- So increasing M literally **spreads the operator's support** by M lattice spacings.
- Domain-wall: hopping L_5 steps along the fifth dimension spreads the boundary operator by L_5 spacings. **Same picture.**



$M = 1$: local Wilson. $M \to \infty$: fully non-local overlap.

Overlap *traded the extra dimension for all-to-all interactions*;
QSP's all-to-all polynomial **is** that dimension, grown back.

This is the quantum version of an old story

- Classical lattice QCD has wrestled with $\varepsilon(H_{\mathrm{w}})$ for decades: [review: Blum, Shamir 2026]
  - rational (Zolotarev) approximations for overlap,
  - the Möbius / truncated domain-wall formulations.
- They all hit the *same* scaling $M = O(\kappa^{-1} \log \frac{1}{\epsilon_e})$, with the approximation order playing the role of L_5.
- **QSP provides the unifying quantum analog**: the polynomial degree is the single knob, the quantum analog of L_5

The "overlap = boundary of domain-wall" equivalence re-emerges as a statement about *quantum algorithms*.

Summary

- We built **QSP-based quantum algorithms** for overlap and domain-wall fermions, both keeping an **exact** (softly broken) lattice chiral symmetry.
- For overlap, the non-local sign function is **manufactured by QSP** from the cheap local Wilson operator, at cost $O(Q \log \frac{1}{\epsilon_e})$, with GW violation $O(\epsilon_e)$.
- The central insight:

The QSP polynomial degree $M \sim \log \frac{1}{\epsilon_e}$ **is** the domain-wall fifth dimension L_5.

- A clean trade-off you can choose by hardware:

Overlap: few qubits (Q), deeper non-local circuits.
Domain-wall: shallow local circuits, $L_5\times$ the memory.

Outlook

- Hardware-related
  - **Realistic gauge theories:** explicit resource estimates for $SU(3)$; the block encoding already accommodates non-Abelian links.
  - **Near-term hardware:** benchmarking the overlap/domain-wall trade-off on real devices.
- Conceptual
  - Is there a better formulation of Hamlitonian overlap for quantum devices?
  - Chiral gauge theories will likely suffer from a sign problem — need to develop quantum simulation techniques in parallel
  - Quantum algorithms may shed new light on the physics

Complementary routes to lattice chirality at this workshop:

▸ *see Aoki's & Shao's talks (modified Villain, exact anomalies)* • *Xu's, Zakharov's, Ueda's talks (symmetric mass generation)*

Thank you!

Backup: the modified chirality operator

- With the approximate sign $E_M \approx \varepsilon(h_{\mathrm{w}})$, define $\hat{V} = \gamma^0 E_M$ and

$$\hat{\gamma}_5 = \frac{1}{2}\gamma_5\,(\mathbb{1} - \hat{V}).$$

- Exact chirality needs $\hat{V}$ unitary; truncation gives $\|\hat{V}^\dagger\hat{V} - \mathbb{1}\| \le 2\epsilon_e$.
- The commutator collapses to

$$[\hat{\gamma}_5, h] = -\gamma^0(\mathbb{1} - \hat{V}\hat{V}^\dagger)\gamma_5 \quad\Rightarrow\quad \|[\hat{\gamma}_5, h]\| \le 2\epsilon_e.$$

- Since E_M already comes out of the QSP circuit, implementing $\hat{\gamma}_5$ costs the *same* as one application of the overlap operator.