Succinct Fermion Data Structures

Joseph Carolan[1] Luke Schaeffer[2]

[1]University of Maryland College Park

[2]University of Waterloo

Data Structures

- Consider storing <*thing*> in 1 out of K configurations
- We will want to query/update <*thing*>
- The choice of representation impacts efficiency
- For example, consider storing bitstrings $\{0,1\}^n$
- We can store it as a **bitstring**:

0	0	1	0	1	0	0	0	1	0	0	0

  ($\rightarrow$) easy to flip a bit!
- Or as a sorted **list of pointers**:

2	4	8

  ($\rightarrow$) easy to find the t-th one!

Succinct Data Structures

- Consider storing <*thing*> in 1 out of K configurations
- Needs at least $\mathcal{I} := \lceil \log K \rceil$ bits
- Usually trade-off: space $\leftrightarrow$ time

Definition

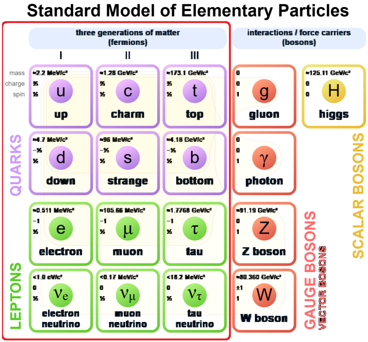
Succinct: A representation using $\mathcal{I} + o(\mathcal{I})$ many bits
($\rightarrow$) Fraction of unnecessary information goes to 0

- Succinct data structures have found many (classical) applications in big data

Fermions

What is a **fermion**?

- Building block of matter
- Any half-odd integer spin particle
- Show up in physics, chemistry, etc.

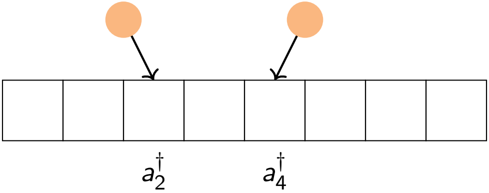


Fermion Data Structures

- Quantum computers are made up of qubits
- To simulate physics/chemistry, we want to encode fermions
- Idea: find qubit operators that obey fermion rules
- We will see that this is a data structures problem
- In particular:
  (1) Represent **bitstrings** $\{0, 1\}^M$ on a quantum computer
  (2) Support efficient **sign-rank** and **bit-flip** queries

Fermions

What we will need:



- Fermion := a thing you put in a mode
- $a_j^\dagger$ puts a fermion in the j-th mode
- Finite number of modes ($M = 8$)
- Finite number of fermions ($F = 2$)
- At most one fermion per mode
- Fermions anticommute

Fermions

Definition

Majorana operators (more convenient basis):

$$\gamma_j := a_j^\dagger + a_j \qquad \bar{\gamma}_j := i(a_j^\dagger - a_j)$$

- Fock states, denoted $|\mathbf{b}\rangle_f$ correspond to bitstrings $\mathbf{b} \in \{0,1\}^M$
- The state $\left|0^M\right\rangle_f$ has no fermions ("vacuum state")

Definition

Fock states are acted on by majorana operators as:

$$\gamma_j \left|\mathbf{b}_1 ... \mathbf{b}_j ... \mathbf{b}_M\right\rangle_f = \left(\prod_{n=1}^{j-1} (-1)^{\mathbf{b}_n}\right) \left|\mathbf{b}_1 ... (\neg \mathbf{b}_j) ... \mathbf{b}_M\right\rangle_f ,$$

$$\bar{\gamma}_j \left|\mathbf{b}_1 ... \mathbf{b}_j ... \mathbf{b}_M\right\rangle_f = i \cdot \left(\prod_{n=1}^{j} (-1)^{\mathbf{b}_n}\right) \left|\mathbf{b}_1 ... (\neg \mathbf{b}_j) ... \mathbf{b}_M\right\rangle_f .$$

Fermions

Definitions

For bitstring $\mathbf{b} \in \{0,1\}^n$, index $j \in [n]$, define:

$$\text{sgn-rank}(\mathbf{b}, j) := \prod_{n=1}^{j} (-1)^{\mathbf{b}_n} \qquad (\text{“sign rank”})$$

$$\text{bit-flip}(\mathbf{b}, j) := \mathbf{b}_1 \dots (\neg \mathbf{b}_j) \dots \mathbf{b}_n \qquad (\text{“bit flip”})$$

With these in hand, can rewrite:

Definition

$$\gamma_j \left|\mathbf{b}\right\rangle_f = \text{sgn-rank}(\mathbf{b}, j-1) \left|\text{bit-flip}(\mathbf{b}, j)\right\rangle_f$$

$$\bar{\gamma}_j \left|\mathbf{b}\right\rangle_f = i \cdot \text{sgn-rank}(\mathbf{b}, j) \left|\text{bit-flip}(\mathbf{b}, j)\right\rangle_f$$

Example Fermion Data Structures

- The complexity of majoranas is intimately tied to the complexity of running most simulation algorithms
- The natural representation maps $|\mathbf{b}\rangle_f$ to $|\mathbf{b}\rangle$
- Called "Jordan Wigner" [1]
- The sgn-rank becomes a prefix list of Pauli Z's
- The bit-flip becomes a single Pauli X
- **Problem:** requires M qubits, even when F is small
- **Problem:** requires $\Omega(M)$ gates for most operations

Succinct Fermion Data Structures

- Early quantum computers will be very small
- Therefore, a recent line of work aims to improve space efficiency of fermion encodings
- Many works exploit that physical systems often conserve the total number of particles
- This result: how to represent such fermions **succinctly** and **efficiently** on a quantum computer
- Note: time efficiency refers to quantum **circuit** complexity

Sparse Systems

Suppose there are F (conserved) fermions in M modes, satisfying

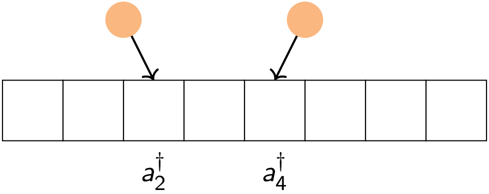
$$F = o(M)$$

- The minimum space usage is $\mathcal{I} := \lceil \log \binom{M}{F} \rceil$
- Prior works:

Encoding	Space	Time
Optimal Degree[6]	$\Omega(\mathcal{I}^2 \log^2 M)$	$\Omega(\mathcal{I}^2 \log^3 M)$
Qubit Tapering/Segment[3, 4]	$M - o(M)$	$\Omega(F^2)$
Configuration Interaction[2]	$\Omega(F \log M)$	*Incomparable*
Permutation Basis[5]	$\mathcal{I}$	$\Omega(M^2 2^{\mathcal{I}})$
Qubit Efficient[7]	$\mathcal{I}$	$2^{\Omega(\mathcal{I})}$

- Note that prior succinct structures require exponential time

Sorted List Encoding



- A Fock state (above) is essentially a bitstring
- We can store it as a bitstring:

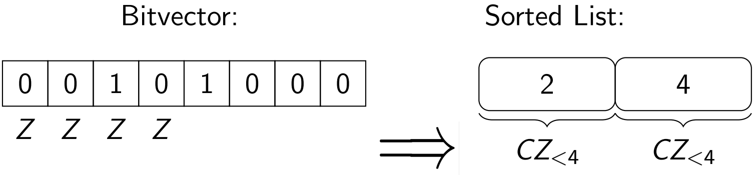
0	0	1	0	1	0	0	0

- Or as a sorted list of pointers:

2	4

Performing Sign Rank

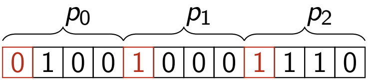
- Recall that **sign-rank** is a prefix string of Z's on the bitstring
- On a sorted list, this can be done with register comparisons



- Performing **bit-flip** is an insertion/deletion
- Somewhat more intricate, but can be done in linear time
- Unfortunately, **neither** representation is succinct over $F = o(M)$!

Succinct Encoding

- Consider $M = 16$, $F = 3$
- Sorted list, in binary:

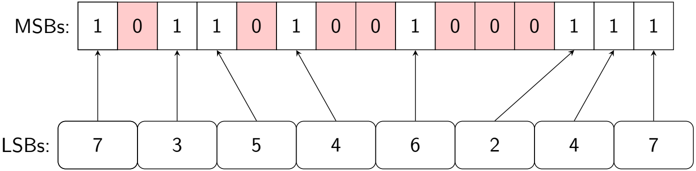


- The red qubits are non-decreasing
- Possibilities: $000, 001, 011, 111$, so 2 qubits suffice
- There is redundancy in top $\approx \log F$ qubits of each pointer
- Store them via stars and bars instead
- This cuts down $O(\log F)$ MSBs to $O(1)$ bits
- This is a **succinct** representation[1]

[1] When $F = o(M)$

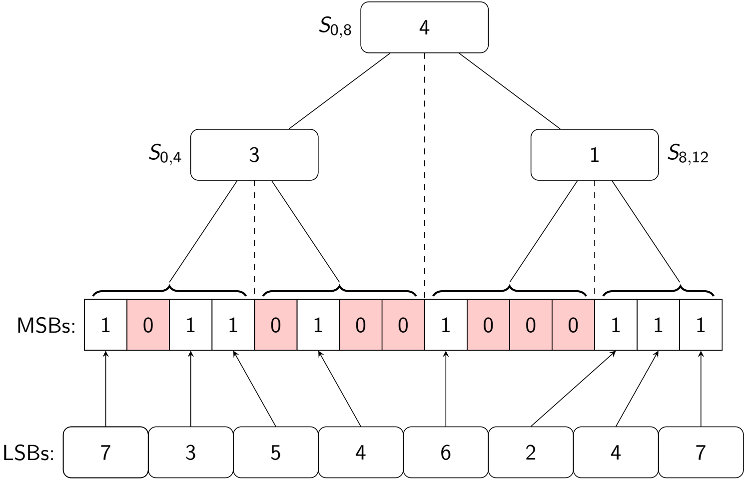
Succinct Representation

- The MSBs and LSBs can be paired up from the ordering
- **Problem:** The MSBs cannot be efficiently accessed



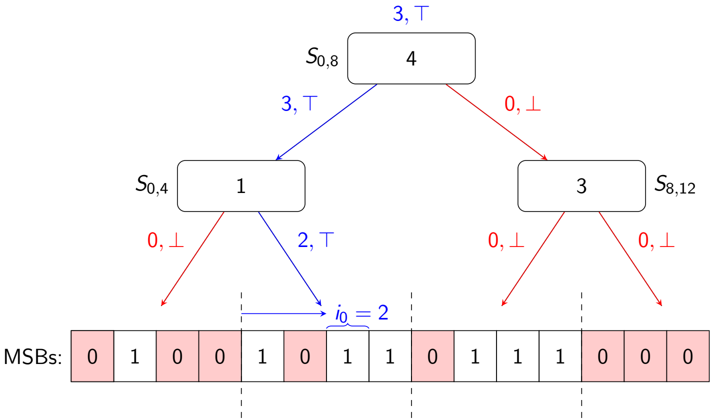
Succinct Tree Structure

- We remedy this with a certain **sum tree**
- Each node stores the number of ones in its left subtree



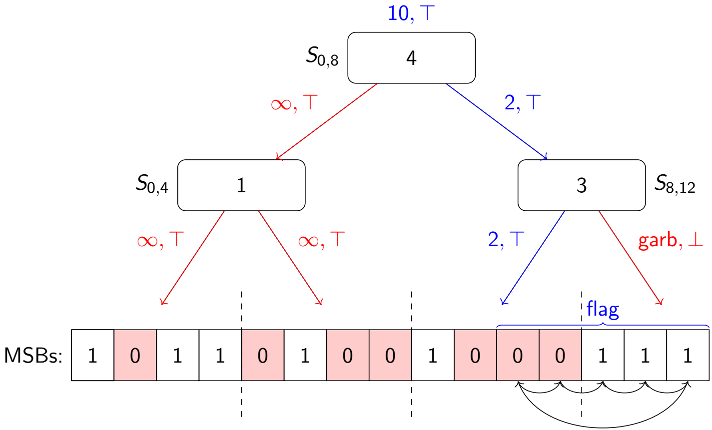
Performing Comparisons

- Comparisons involve reconstructing a given MSB efficiently
- To do this, we walk down our sum tree



Performing bit flips

- Bit flips correspond to inserting/deleting from MSBs
- Result: a list rotation



Main Results

Suppose there are F (conserved) fermions in M modes, satisfying

$$F = o(M)$$

Encoding	Space	Time
Optimal Degree[6]	$\Omega(\mathcal{I}^2 \log^2 M)$	$\Omega(\mathcal{I}^2 \log^3 M)$
Qubit Tapering/Segment[3, 4]	$M - o(M)$	$\Omega(F^2)$
Configuration Interaction	$\Omega(F \log M)$	*Incomparable*
Permutation Basis[5]	$\mathcal{I}$	$\Omega(M^2 2^{\mathcal{I}})$
Qubit Efficient[7]	$\mathcal{I}$	$2^{\Omega(\mathcal{I})}$
This work	$\mathcal{I} + o(\mathcal{I})$	$O(\mathcal{I})$

- We achieve succinctness with **linear** gate complexity
- This beats many prior encodings in time as well as space
- Bonus: our circuits are **log depth**

Prior Work on Dense Systems

Suppose there are F (conserved) fermions in M modes, satisfying

$$F = \Theta(M).$$

- "Constant filling" is often relevant in highly entangled systems
- However, few known ways to save meaningful space

Encoding	Space	Time
Qubit Tapering/Segment[3, 4]	$M - O(1)$	$\Omega(\mathcal{I}^2)$
Permutation Basis[5]	$\mathcal{I}$	$\Omega(\mathcal{I}^2 2^{\mathcal{I}})$
Qubit Efficient[7]	$\mathcal{I}$	$2^{\Omega(\mathcal{I})}$

- Prior succinct structures require exponential complexity
- Idea from last slides is not succinct here
- Even a **single bit** of redundancy per fermion is too much

Implicit Labels

- Starting point: enumerate $\binom{M}{F}$ strings in lexicographic order
- Store a pointer to positions in this list
- Clearly space optimal, $\mathcal{I}$ usage!
- But how to do gates **efficiently?**
- Prior work [5, 7] explores similar ideas, results in

$$\text{Gate Cost} = \exp(\mathcal{I})$$

An Observation

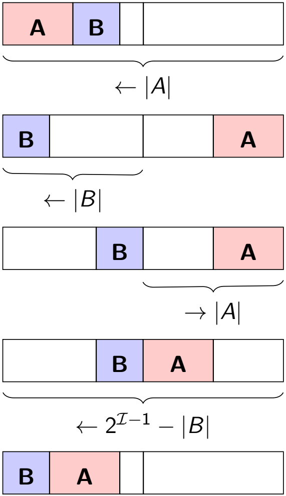
- Consider lexicographic order
- The MSB can be "accessed"
- Computed by a **fixed comparison** on label
- Can apply a phase, i.e. for sign-rank
- **Problem:** How to bitflip?
- **Problem:** LSB is hidden...

$$\left.\begin{aligned}\mathcal{L}(00011) &= 0\\ \mathcal{L}(00101) &= 1\\ \mathcal{L}(00110) &= 2\\ \mathcal{L}(01001) &= 3\\ \mathcal{L}(01010) &= 4\\ \mathcal{L}(01100) &= 5\end{aligned}\right\}\ \textbf{MSB } 0$$

$$\left.\begin{aligned}\mathcal{L}(10001) &= 6\\ \mathcal{L}(10010) &= 7\\ \mathcal{L}(10100) &= 8\\ \mathcal{L}(11000) &= 9\end{aligned}\right\}\ \textbf{MSB } 1$$

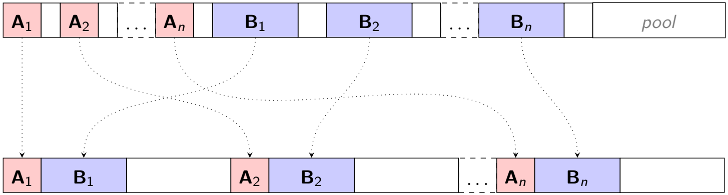
Performing Bit Flips

- We include a pool of unused labels, allowing some insertions/deletions
- We can swap contiguous labels using controlled modular arithmetic
- Can be broken into "rotations"
- Each rotation is a controlled modular addition



Reaching Other Bits

- We will walk through a sequence of M orderings
- Each ordering will “expose” a different bit
- Each re-ordering is a sequence of many transposition



- Key insight: **modular arithmetic** allows us to do many transpositions at once, for the same cost as one transposition
- Many more technical details swept under the rug...

Main Results

Suppose there are F (conserved) fermions in M modes, satisfying

$$F = \Theta(M).$$

- Comparison to prior space efficient encodings:

Encoding	Space	Time
Qubit Tapering/Segment[3, 4]	$M - O(1)$	$\Omega(\mathcal{I}^2)$
Permutation Basis[5]	$\mathcal{I}$	$\Omega(\mathcal{I}^2 2^{\mathcal{I}})$
Qubit Efficient[7]	$\mathcal{I}$	$2^{\Omega(\mathcal{I})}$
This work	$\mathcal{I} + O(1)$	$O(\mathcal{I}^3)$

- Our work is (even better than) succinct, uses $O(1)$ ancilla
- Achieves exponentially improved time complexity for majorana operators

Conclusion

- Representing fermions is storing bit-vectors
- It suffices to perform **sign rank** and **bit flip** efficiently
- We use data structures ideas to significantly improve on space and time of prior encodings
- Many open questions:
  (1) What about other physical symmetries?
  (2) Bring down the $O(\mathcal{I}^3)$ scaling in implicit structure?
  (3) Reduce the T-gate count?

Thank you!

[1] P. Jordan and E. Wigner. "On the Paulian prohibition of equivalence.". In: *Z. Physik* 47 (1928), pp. 631–651.

[2] Ryan Babbush et al. "Exponentially more precise quantum simulation of fermions in the configuration interaction representation". In: *Quantum Science and Technology* 3.1 (Dec. 2017), p. 015006.

[3] Sergey Bravyi et al. *Tapering off qubits to simulate fermionic Hamiltonians*. 2017.

[4] Mark Steudtner and Stephanie Wehner. "Fermion-to-qubit mappings with varying resource requirements for quantum simulation". In: *New Journal of Physics* 20.6 (June 2018), p. 063010.

[5] Brent Harrison et al. *Reducing the qubit requirement of Jordan-Wigner encodings of N-mode, K-fermion systems from N to* $\lceil \log_2 \binom{N}{K} \rceil$. Nov. 2022.

[6] William Kirby et al. "Second-Quantized Fermionic Operators with Polylogarithmic Qubit and Gate Complexity". In: *PRX Quantum* 3 (2 June 2022), p. 020351.

[7] Yu Shee et al. "Qubit-efficient encoding scheme for quantum simulations of electronic structure". In: *Phys. Rev. Res.* 4 (2 May 2022), p. 023154.