Can this algorithm lead to more optimal hardware design?

Yes! Here is one example. Imagine that you are designing a two-qubit quantum gate. You discover some tricks which can make the gate faster. But, doing so increases the degree of error incurred by this gate. How should this gate-speed vs. error trade-off be balanced? Our work gives a route to quantifying gate error in terms of algorithm runtime. By understanding how both gate error and gate speed contribute to algorithm runtime, we can choose the optimal balance between them. Beyond gate design, we hope these insights will be useful in developing effective error correction schemes, comparing different qubit modalities, and in other scenarios where error rates and operating speeds are relevant.

Could you say a few sentences about your error model? Coherent? Incoherent? Leakage out of computational subspace?

The error model considered in our work is rough. The jargon description is that we model error on circuit blocks with a global depolarizing channel. Much like the Bohr model of the atom, we view this as a starting point and source of motivation for future development in this direction. Important work is being done on so-called "noise-tailoring" techniques, which make error-modeling more tractable. Our findings highlight the value of work on noise characterization and tailoring.

You've shown the optimal number of Grover iterations depends on the number of qubits and gate fidelity. Is there a further dependence of the type/complexity of observable being estimated on the optimal number of Grover iterations?

The short answer is yes. Consider a binary observable that, when represented as a unitary, requires a deep circuit. The deeper this circuit is, the more error will be incurred. The optimal number of Grover iterates will be fewer, so the runtime-to-target-accuracy will be increased. Our general model, in terms of the circuit layer decay rate, includes such cases. But we also consider the specific case of a Pauli string observable, which can be implemented as a layer of single-qubit gates. For this particular case, we evaluate the runtime in terms of the parameters of qubit number, gate fidelity, etc.