Enabling Pulse-level Programming, Compilation, and Execution in XACC

Thien Nguyen  
Computer Science and Mathematics  
Oak Ridge National Laboratory  
nguyentm@ornl.gov

Alexander McCaskey  
Computer Science and Mathematics  
Oak Ridge National Laboratory  
mccaskeyaj@ornl.gov

ABSTRACT
Noisy gate-model quantum processing units (QPUs) are currently available from vendors over the cloud, and digital quantum programming approaches exist to run low-depth circuits on physical hardware. These digital representations are ultimately lowered to pulse-level instructions by vendor quantum control systems to affect unitary evolution representative of the submitted digital circuit. Vendors are beginning to open this pulse-level control system to the public via specified interfaces. Robust programming methodologies, software frameworks, and backend simulation technologies for this analog model of quantum computation will prove critical to advancing pulse-level control research and development. Prototypical use cases for this include error mitigation, optimal pulse control, and physics-inspired pulse construction. Here we present an extension to the XACC quantum-classical software framework that enables pulse-level programming for superconducting, gate-model quantum computers, and a novel, general, and extensible pulse-level simulation backend for XACC that scales on classical compute clusters via MPI. Our work enables custom backend Hamiltonian definitions and gate-level compilation to available pulses with a focus on performance and scalability. We end with a demonstration of this capability, and show how to use XACC for pertinent pulse-level programming tasks.

1 INTRODUCTION
The emergence of noisy intermediate scale quantum (NISQ) [1] processing units and their accessibility to programmers via cloud services in recent years has led to major advances in software frameworks and programming models for quantum-classical computing. For gate-model quantum computing the majority of the programming framework approaches have focused on digital assembly submission to dedicated quantum job servers. This has led to a number of demonstrations of small-scale experiments in machine learning, quantum chemistry, nuclear physics, and quantum field theory [2–7]. However, in the NISQ-era, much stands to be gained from thinking about programming quantum hardware closer to the physical model of the system, specifically, programming at an analog, or pulse-instruction, level. Early results in this regard have demonstrated the utility of pulse-level control in areas such as error mitigation and optimal quantum control [8, 9]. Moreover, vendors are starting to propose and implement specifications for pulse-level job submission, opening up this novel, analog programming model to the community at large. In the near-term, programming models and software framework implementations will prove critical in enabling the benefits of low-level analog quantum control and programming. Moreover, robust simulation technologies are required to promote pulse-level research productivity due to the scarce and remote characteristic of available pulse-level physical backends.

Programming mechanisms for pulse-level quantum computation is very much at an early research and development phase. IBM has recently put forward the OpenPulse specification, with a corresponding early implementation in the Qiskit-Terra framework [10]. This implementation provides a high-level Pythonic approach for pulse-level program composition and specifically targets IBM physical backends. Pulse-level simulation tools exist but are available in an ad-hoc manner - most users leverage frameworks such as Qutip for quick prototyping and scripting [11]. These approaches lack integration with physical hardware backends and internal circuit representations for the typical quantum programming, compilation, and execution workflow. There is a need for a system-level programming, compilation, and execution workflow that enables scalable simulation on available distributed-memory compute architectures.

Here we detail our solution to the current lack of programming tools and simulation technologies for pulse-level
programming of gate-model quantum architectures. We leverage the system-level XACC heterogeneous quantum-classical software framework to provide a programming mechanism and associated compilation tooling for pulse-level quantum control systems. Initially we target the OpenPulse specification, but believe that the hardware-agnostic characteristic of XACC will enable us to extend this model for future quantum control specifications (such as Rigetti Quilt). We put forward a state-of-the-art pulse-level simulation backend for XACC that builds off the QuaC (Quantum in C) open quantum system dynamics solver. QuaC builds off of the PETSc matrix algebra library, and therefore our work immediately scales on distributed memory architectures via the Message Passing Interface (MPI). We detail programming, compilation, and execution layers of this overall pulse-level programming model, and provide detailed demonstrations exhibiting the use of this API. Our goal with this work is to provide a unique pulse-level programming capability with a robust simulation backend to enable research productivity for tasks in error mitigation, physics-inspired pulse learning, and quantum optimal control.

The paper is organized as follows. In Section 2, we review some key features and components of the XACC framework. Next, in 3, we briefly summarize the OpenPulse specification which we adopt in this work. Section 4 describes the overall software architecture of our pulse-level extension and how it fits in with the framework. Usage demonstrations and conclusions are given in Section 5 and Section 6, respectively.

2 XACC FRAMEWORK

Recent progress in the development of programmable quantum computing architectures has resulted in a number of programming models and associated software stack implementations enabling digital gate-level programming of these nascent architectures. Most efforts have been put forward by hardware vendors and typically only target the backend architecture the vendor provides. Most of these efforts are provided via high-level Pythonic circuit construction data structures, with a few specifically extending existing classical languages for heterogeneous quantum-classical computing. This landscape has necessitated the development of unified, system-level architectures for robust integration efforts across quantum architectures, languages, and compiler techniques [12].

Recently, the XACC system-level software framework has been developed to address these requirements. XACC provides an extensible, holistic framework for quantum-classical computing in C++ in a way that is agnostic to the backend quantum hardware type as well as front-end language transpilation and compilation [13]. At its core, XACC provides a unified and simplified quantum programming

Figure 1: The XACC framework provides a service-oriented architecture. Services are interfaces that expose core functionality for the typical quantum-classical programming, compilation, and execution workflow. Here we demonstrate the hierarchy of these services with concrete service implementations.
model that can cross hardware boundaries while not compromising efficiency in terms of quantum hardware utilization or developers productivity. Moreover, the system-level language approach ensures performance of classical processing and support code which is essential in the NISQ-era where hybrid algorithms are one of most viable approaches that can provide substantial quantum advantage.

As a comprehensive framework, XACC (see Fig. 1) adheres to the newly proposed QCOR language extension specification [14], which provides guidance on the programming and memory model, support data structures, and the execution model. XACC provides a reference implementation that is open-source, fully-functional, actively being maintained and developed [15]. The base implementation supports all major quantum assembly dialects, e.g. Open QASM, Quil, and Staq [16–18], as well as its own XASM dialect. XACC has back-end connections, e.g. via RESTful APIs, to hardware platforms provided by IBM, Rigetti, IonQ, and Dwave, thereby enabling a write-once-deploy/test-everywhere capability to programmers and researchers. Another key focus area is the integration of quantum simulation software to use as drop-in replacements for real hardware back-end during development. XACC support simulation backends from the IBM Terra/AER package [19], the Quantum++ library [20], the ITensor library [21], and the high-performance ExaTN tensor library [22]. This level of modularity and extensibility, and the overall strong focus on performance, significantly enhances developer productivity.

XACC provides an extensive set of libraries that are constantly updated. This includes algorithm implementations, from simple static algorithms, e.g. quantum Fourier transform, reversible classical computing, to novel quantum variational ansatz circuits, such as UCCSD, ASWAP, or quantum circuit structure learning. Observables, such as Pauli or spin operators, are a native construct within the framework, and enable XACC to automatically inject gates and measurement operations in order to compute pertinent expectation values. Moreover, XACC supports common hybrid algorithms like the variational quantum eigensolver and data-driven circuit learning routines, as well as characterization tools such as quantum process tomography. Again, these libraries are meant to provide developers with drop-in, backend-compatible, and fully-validated solutions without needing them to implement or maintain the code.

Lastly, XACC provides user-friendly Python language bindings via Pybind11 [23]. This is not a specification-driven feature but is equally important since C++, despite being highly performant, requires substantial familiarity with professional software development procedures, e.g., building and linking, that may represent a barrier to domain computational scientists. Thanks to the language bindings, users can use XACC services from their Python shell and, more importantly, can provide implementations to framework constructs, e.g. Accelerators, Compilers, Algorithms, etc., natively from Python while adhering to the common interfaces that have been exposed via the language bindings.

### 3 OPENPULSE SPECIFICATION

The OpenPulse specification [10] was proposed by IBM as an open and implementation agnostic standard for pulse-level control. The specification defines JSON schemas for system Hamiltonians, drive and control channel configurations, pulses, and pulse-sequence definitions. Since this is the first pulse-level standard put forward in the industry, our XACC implementation adheres to the OpenPulse specification.

#### 3.1 Hamiltonian

The specification defines a machine-readable encoding scheme for the system Hamiltonian as summarized in Table 1. Specifically, the Hamiltonian field \( h, \text{str} \) in the OpenPulse JSON object is a list of strings describing various terms in the Hamiltonian. These terms can be time-independent (controlled evolution) or time-dependent (driven by an input channel). The latter case is denoted by the characters \(|D|\) followed by a channel name (either a drive channel, \(d_i\), or a control channel, \(u_i\)).

| Type                        | Definition                                                                 | Examples                                                                 |
|-----------------------------|---------------------------------------------------------------------------|--------------------------------------------------------------------------|
| Operators                   | - Pauli operators \((I, X, Y, Z)\)                                       | \(2 \pi v_0 Z_0\)                                                        |
|                             | - Ladder operators \((SM, SP)\)                                          | \(2 \pi v_0 SP_0 SM_0\)                                                  |
|                             | - Number operators \((0, N)\)                                            | \(0.5 \delta \alpha (1 - O_0) + O_0\)                                   |
| Drive and Control Channel   | Special suffix \((|D|\) or \(|U|\) to denote the Hamiltonian term is modulated by either a drive or control channel) | \(g X_0 |D_0\)                                                          |
| Summation                   | A running sum of operator expression indexed by an integer               | \_SUM\[(i, 0, 9, -0.5*h\{i\}+Z\{i\})\]                                  |
|                             |                                                                           | \_SUM\[(i, 0, g\{i\}+X\{i\})\]|\(|D\{i\}\)|      |
| Variables                   | A map from symbolic variables to values                                   | \"vars\": \{"v\": 5.0, \"g\": 0.1\}                                 |
| Qubit dimensions            | A map from qubit indices to dimensions                                   | \"qub\": \{"\0\": 2, \"\1\": 2\}                                    |

**Table 1:** JSON schema of OpenPulse backend Hamiltonian specification
XACC framework implements provide a foundation upon the extensible, layered and future-proof interfaces that the quantum kernels as we will describe in detail next.

4.1 Pulse IR
In the XACC framework, the Instruction and CompositeInstruction intermediate representation (IR) interfaces are the building block of in-memory data structures capturing the semantics of a user-specified quantum program. In a usual setting, an Instruction represents a quantum kernel instruction, i.e. a quantum gate. To handle analog-like pulse instructions, we define a Pulse sub-class of Instruction (Fig. 3 (a)) that encapsulates extra metadata associated with analog control such as data samples, drive channel and pulse timing. We extend Instruction with typical getter and setter methods for this metadata, which by default (for digital gate instructions) has no effect. We implement these methods on Pulse to provide this metadata upon request. Moreover, we extend getInstruction with a isAnalog() method to easily indicate if a set of Instructions are digital or analog. This allows us to construct flexible combinations of instructions (analog or digital) when building up a CompositeInstruction, i.e. a quantum circuit/kernel.

In pulse mode, analog IR objects can be constructed via direct calls to the IR factory methods (Fig. 4(a) & 4(b)) or using XACC pulse library loader (Fig. 4(c).) The former offers maximal flexibility whereby arbitrary pulse instructions can be constructed from sample arrays. The latter provides a means for users to utilize any OpenPulse-conformed pulse library to build up their pulse program in XACC. Pre-defined and calibrated pulses are loaded into the instruction registry of the XACC IR provider for later use by referring to the pulse instruction name. In XACC, pulse composites that implement gates are prefixed with a ‘pulse::’ phrase. It is worth noting that, in pulse execution mode, XACC will automatically lower digital gate instructions to pulse composites as described in Section 4.2.

Besides being able to construct quantum kernels from pulse-level IR objects, the framework has extended its compilation infrastructure to handle automatic gate-to-pulse IR lowering. Hence, existing quantum circuits can be translated into pulse sequences and simulated dynamically within the framework itself. The existing XACC CompositeInstruction API allows us to add, remove, or replace component Instruction nodes, which is crucial for the pulse-aware lowering of quantum kernels as we will describe in detail next.

Figure 2: JSON description of a simple pulse and a quantum gate consisting of multiple pulses in a sequence (cmd-def).

Also, there is a set of pre-defined operators, such as Pauli operators (x, y, or z) or ladder operators (e.g. SM denotes the spin σ− or the photon destroy (a) operator), that can be used to construct those Hamiltonian terms. For brevity, the standard also allows running summation terms with integer indices to group terms that have the same structure.

All variables in the Hamiltonian expressions must have corresponding assignments in the vars map to resolve the Hamiltonian dynamics. Additionally, qubit subsystems can be modeled as higher-dimensional systems (e.g. qutrits or qudits) by specifying their dimensions in the qub field.

3.2 Pulses and Commands
Pulses (waveforms of drive signals) are defined as sequences of sample points (complex numbers). They are specified as a JSON object with a name and a samples field (see Fig. 2). A standard OpenPulse back-end configuration often contains a calibrated pulse library and a map from quantum gates to sequences of pulses. Each mapping entry is called a command definition, or cmd-def. A sample cmd-def JSON entry is also shown in Fig. 2.

As we can see, at the pulse level, a digital quantum gate is decomposed into a sequence of pulses (referred to by name) and frame changes (changing the local oscillator’s phase) on specific channels at specific time. The pulse sequence for each gate is defined independently referencing the gate start time as time zero. A back-end pulse library can implement an arbitrary set of universal gates and hence it is up to the high-level gate compiler to perform any necessary transformations and mappings.

4 SOFTWARE ARCHITECTURE
The extensible, layered and future-proof interfaces that the XACC framework implements provide a foundation upon which our pulse-level extension is built. We supplement the front-end intermediate representation with a pulse instruction extension and implement a middle-end pipeline to enable gate-to-pulse translations. The processed pulse-level program can then be executed on the QuaC simulator/accelerator backend which performs a fully dynamical simulation of the quantum system driven by those pulse instructions. Of course, with backend support for quantum hardware from IBM, XACC can take this pulse-level IR and execute it on physical architectures as well.
Enabling Pulse-level Programming, Compilation, and Execution in XACC

**Instruction**

- name: String
- bits: vector<int>
- parameters: vector<InstructionParameter>

**Gate**

- gateName: String
- isAnalog(): return False;

**Pulse**

- t0: Integer
- duration: Integer
- channel: String
- samples: vector<Complex>

- isAnalog(): return True;
- setChannel(String channel)
- setSamples(vector<Complex> samples)
- setStart(Integer start)

(a) C++ class polymorphism is used to describe both quantum gate and pulse instructions.

(b) C++ class structure represents different type of Hamiltonian terms.

Figure 3: Unified Modeling Language (UML) diagrams representing (a) IR Instruction and Pulse extension, and (b) Hamiltonian term polymorphism.

### 4.2 Pulse Program Compilation

An important feature that we want to enable for XACC is the ability to transpile or compile gate or assembly level descriptions of quantum circuits into analog pulses that have been arranged to individual channels at specific time slots. This capability previously is only available on non-public hardware vendor backend software stacks. Since we want to implement an end-to-end pulse-capable stack for XACC that spans from front-end to back-end, including a pulse emulator backend that we will introduce in Section 4.3, we also implement a generic digital quantum gate to pulse compilation workflow.

Starting with a CompositeInstruction tree structure, e.g. by compiling a quantum assembly program, XACC provides a flexible InstructionVisitor interface to traverse this Instruction tree and apply custom operations for each concrete Instruction (gate) type. Following this visitor pattern, we have implemented a two-pass gate-to-pulse lowering and scheduling procedure as described in Algorithm 1 & 2.

In the first pass, all gates are transpiled into pulses or pulse sequences via Instruction registry look-up. If a gate has a direct pulse command definition available, e.g. X or CNOT gates, it will be substituted by the associated pulse sequence. However, many quantum gates do not have direct pulse definitions, e.g. due to extended over-complete instruction sets. In that case, we rely on the universal instruction set that the registered pulse library implemented to decompose the gate into its universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses. After converting the original gate IR into universal gate equivalent and then lower those decomposed gates into pulses.
auto provider = xacc::getIRProvider("quantum");
auto compositeInst = provider->createComposite("test_pulse");
auto pulseWalshGaussian = make_shared<Pulse>("WalshGaussian", "d0");

// Pulse reference: arxiv:1809.03452 (sec 7.1)

std::vector<std::vector<double>> samples = {
  {0.0, 0.0}, {0.013434, 0.0}, {0.058597, 0.0}, ...
};

pulseWalshGaussian->setSamples(samples);

// Add the pulse instruction to the CompositeInstruction
compositeInst->addInstruction(pulseWalshGaussian);

(a) Define a pulse via C++ IR manipulation.

(b) Define a pulse via Python IR manipulation.

(c) Pulse IR construction by loading pulse library JSON file. All pulses and pulse sequences (command definitions) are loaded to the IR service registry. These pulse IR’s can then be retrieved from the IR provider.

Figure 4: XACC pulse-level IR construction via (a) C++ API, (b) Python API, and (c) IR provider.

quantum circuit. Therefore we must implement a compiler pass that schedules pulses globally. This high-level pulse scheduling is described in Algorithm 2. To achieve this, we have extended XACC with a new Scheduler service interface which we implement specifically for global pulse scheduling - the PulseScheduler.

This implementation basically block-shifts pulse-level CompositeInstructions, i.e. those pulse sequences that implement quantum gates, by adjusting the start time of component pulses. This maintains the relative timing of pulses within a pulse sequence as well as the atomicity of quantum gates once they are converted into pulses.

At the end of this procedure, we have a globally scheduled deck of pulse instructions, e.g. the one listed in Fig. 5, dictating analog signals on channels driving the system Hamiltonian. This pulse program can hence be submitted to a pulse-capable hardware backend or to the QuaC simulator backend that we have provided as part of the XACC suite.

4.3 QuaC Pulse Accelerator

The XACC framework defines an Accelerator service that provides a unified interface for all framework services that need access to a quantum computer, either via a remote REST connection or local simulator. Therefore, we also implement
Figure 5: Example of a pulse schedule. The time field of each pulse is shifted by the pulse scheduler to construct a global pulse schedule after assembling individual pulse sequences.

Algorithm 3: Channel Controller Tick algorithm: updating its internal state during time-stepping.

```
Data: Scheduled Pulse Program
Data: Current Time
Result: Update Channel Active Pulse Sequence
Result: Update Channel Accumulated FC Phase
foreach Active Pulse on Channel do
  if still active then
    Skip;
  else
    Remove current Pulse on Channel;
    Find next Pulse scheduled;
    Update current Pulse on Channel;
  end
end
foreach Channel do
  foreach Frame Change instructions on channel do
    if Activated @ current time then
      Update accumulated Phase on Channel;
      Remove processed Frame Change instruction;
    end
  end
end
```
far along its data sample list we are, (2) accounting for all frame changes (local oscillator phase manipulation) that have requested on any given channels, and (3) performing LO-mixing to compute the absolute time signals. The way in which our pulse controller handles these tasks is summarized in Algorithm 3.

4.3.3 QuaC Simulator. The numerical back-end of our implementation, so-called Quantum in C · quaC [24], as shown in Fig. 6 is a quantum dynamical solver based on the Portable, Extensible Toolkit for Scientific Computation (PETSc) library [25, 26]. Built upon PETSc’s scalable (parallel) data-structures and runtimes, the quaC back-end in XACC supports multi-node parallelism via the industry-standard Message Passing Interface (MPI), which promises extended scalability across classical computing infrastructure.

As a standalone package, QuaC provides a flexible open quantum systems simulator capable of solving generic Lindblad master equations (1) as well as applying gate-type operators.

\[
\frac{d\rho}{dt} = -i[H, \rho] + \sum_k \gamma_k (A_k\rho A_k^\dagger - \frac{1}{2} A_k^\dagger A_k\rho - \frac{1}{2}\rho A_k^\dagger A_k) = \mathcal{L}(\rho)
\]

(1)

The vectorized density matrix (\(\rho\)) and the superoperator (\(\mathcal{L}\)) representing the right-hand side of the master equation are modeled by the abstract PETSc vector (\(\text{vec}\)) and matrix (\(\text{mat}\)) objects whose implementation details, such as memory allocation and storage across multiple MPI processes, are handled implicitly by PETSc for maximal portability. Specifically, the superoperator matrix is stored in the compressed sparse row (CSR) format across multiple (parallel) processes using the PETSc-builtin MPI AIJ type.

Qubit sub-system operators, e.g. Pauli operators, which can be used in the OpenPulse Hamiltonian, are implicitly modeled internally as functions that retrieve non-zero matrix elements and indices to take advantage of their sparsity. These functions are used to project operators’ local elements into the overall RHS superoperator matrix when we process Hamiltonian terms, i.e. executing the apply() method (Fig. 3(b)) on each term. All channel-driven terms in the Hamiltonian are modeled as a time-dependent function to compute the Jacobian which is a PETSc callback at each time step.

It is worth noting that PETSc offers a wide variety of time integrators, such as forward and backward methods, symplectic methods, backward differentiation formula (BDF) method, etc. By default, QuaC uses \(\text{TSRK3BS}\) integrator type (third-order Runge-Kutta scheme of Bogacki-Shampine with 2nd order embedded method) when solving the master equation.

As depicted in Fig. 6, all these low-level numerical computation details are encapsulated within the \texttt{xacc::Accelerator} interface thanks to the high-level system model, pulse channel controller, and accelerator buffer data structures. The first and second objects translate the input Hamiltonian, backend configurations (e.g. LO frequencies), and the pulse program into QuaC and PETSc data structures and functions required for solving the pulse-driven system dynamics. The third object (a core abstraction from XACC modeling a register of qubits and exposing pertinent backend-execution results) captures all necessary outputs from the simulation such as the probability distribution which can be used to generate a measurement shot count distribution.

4.3.4 Pulse Generator Utility. In an effort to improve the usability and ease of integration, we also put forward a set of utility functions that generate pulse data series for common pulse shapes, such as square or Gaussian pulses. Given the analytical parameters of the envelope function, e.g. the width (variance) parameter of a Gaussian function, the sampling time (\(dt\)), and the pulse duration, these utilities will generate the corresponding array of data samples. In the most generic form, users can also supply a string-based functional expression which will then be parsed and evaluated to compute the time-domain samples.

These pulse generation utilities are also made available to the Python language binding for Pythonic programming. Python has a rich collection of signal processing libraries, e.g. \texttt{scipy.signal}, which can be used independently or in conjunction with these XACC-provided facilities to construct input pulses.

5 DEMONSTRATION

In this section, we seek to demonstrate the usage of XACC pulse-level application programming interfaces to accomplish tasks ranging from simple physics experiments to complex, cross-service applications involving integration of new and existing XACC components.

5.1 Arbitrary pulse definition

As a first example, we describe how arbitrary pulse waveforms can be described via direct construction and manipulation of pulse-level IR instances (see Fig. 7). In particular, the setSamples method of the \texttt{Pulse} class (see also Fig. 3(a)) can be used to attach a list of sample data points to a pulse IR. The \texttt{xacc::getAccelerator} API function enables one to retrieve an \texttt{Accelerator} back-end by name from the framework service registry. The \texttt{CompositeInstruction}, i.e. quantum program, which contains pulse-level instructions can be simulated/executed on the back-end by invoking the \texttt{execute} method and passing references to the program and the \texttt{AccelerationBuffer} (see Fig. 6) to which the results are persisted.
auto quaC = xacc::getAccelerator("QuaC");
// Create a simple pulse program using IR
auto provider = xacc::getIRProvider("quantum");
auto compositeInst = provider->createComposite("test.pulse");
// Walsh-Gaussian pulse
auto pulseWalshGaussian =
  make_shared<Pulse>("WalshGaussian", "00");
// Pulse reference: https://arxiv.org/abs/1809.03452 (sec 7.1)
std::vector<std::vector<double>> samples = {
  {0.0, 0.0}, {0.013434, 0.0}, {0.058597, 0.0}, ...
};
pulseWalshGaussian->setSamples(samples);
// Add the pulse and execute
compositeInst->addInstruction(pulseWalshGaussian);
auto quaC = xacc::getAccelerator("QuaC", "system-model", systemModel),
// Run simulation
quaC->execute(qubitReg, compositeInst);

Figure 7: Define a pulse via IR (Intermediate Representation) manipulation.

5.2 Finding resonance frequencies

In this example, we demonstrate the way the system Hamiltonian is specified (Fig. 8) and how to configure various simulation parameters to experiment on- and off-resonance driving of qubits.

Specifically, we will apply an X gate (in xasm dialect, Fig. 9), which will be translated to a corresponding pulse, on each qubit. Since the X-gate pulse will actually be mixed with the LO signal before being applied to the qubit, we can vary the LO frequency in order to find the qubit resonance frequency. In the below simulation, we sweep the frequency of the local oscillators of drive channel 0 (line 20, Fig. 9), and drive channel 1 (line 45, Fig. 9), where the X-gate pulses are applied and collect the probability that the corresponding qubit is measured in the |1⟩ state (line 42, Fig. 9.) The results are shown in Fig. 10 which, as expected, verify the resonance characteristics of the back-end Hamiltonian.

This example also demonstrates the end-to-end gate-to-pulse transformation capabilities of the XACC framework. In particular, rather than needing users to manually specify the pulses, which they can certainly do as in 5.1, digital gates are automatically lowered into pulses and dynamically simulated once the quaC pulse back-end is requested. In other words, other quantum assembly language based
circuits can also be simulated at the pulse level targeting a specific hardware model via the QuaC simulator back-end.

5.3 Qutrit simulation

The XACC pulse-level infrastructure that we have implemented is hardware-agnostic in that it supports not only two-level systems (qubits) but also arbitrary-dimension systems such as qutrits and qudits. In addition, the ability to model higher-dimensional systems allows us to accurately simulate qubit systems in their extended Hilbert space beyond the computational qubit subspace.

For example, we can model superconducting transmon qubits as three-level systems (eq. (2)) in order to compute the amount of leakage into the undesirable (third) basis state.

\[
H = \Omega(t)(\hat{a} + \hat{a}^\dagger) + \omega_{01}\hat{n} + \frac{\alpha}{2}(\hat{n} - 1)\hat{n}
\]  

(2)

Theoretically, the unwanted (leakage) transition (\(|1\rangle \rightarrow |2\rangle\)) is separated from the principle transition (\(|0\rangle \rightarrow |1\rangle\)) by the anharmonicity (\(\alpha\)), i.e. \(\omega_{12} = \omega_{01} + \alpha\). Consequently, noise or Fourier-components of the driving signal at that anharmonicity frequency will result in leakage. We demonstrate this leakage by driving a transmon qubit which is described by the OpenPulse Hamiltonian in Fig. 11 with a square π-pulse plus an additional fictitious noise term of the form, \(\text{Amplitude} \times \cos(\alpha t)\). By varying this Amplitude parameter, we can examine the leakage while driving the system from \(|0\rangle\) to \(|1\rangle\). Fig. 12 directly demonstrates this effect, and we see that increasing noise levels lead to expected leakage into the \(|2\rangle\) state.

5.4 Pulse optimization

XACC is expanding on the QCOR Optimizer concept [14] for pulse-level, optimal quantum control. We leave a detailed description of this capability for a future manuscript, but want to demonstrate its utility here for pulse-level programming and compilation. Fig. 13 demonstrates how one may use standard optimal control techniques (e.g. GOAT, GRAPE, CRAB, etc.) to create a pulse-level program that affects targeted unitary evolution under a certain set of input parameters. This snippet shows how one might leverage this to optimize an arbitrary truncated Fourier series in order to implement an \(X\) gate. This is useful when, e.g., we want to constrain the bandwidth of the drive signal.
Enabling Pulse-level Programming, Compilation, and Execution in XACC

```cpp
1 xacc::HeterogeneousMap configs { ...
2 std::make_pair("method", "GOAT"), ...
3 std::make_pair("optimizer", "ml-pack"), ...
4 std::make_pair("dimension", dimension), ...
5 // Target unitary is an X gate on qubit 0
6 std::make_pair("target-U", "X0"), ...
7 // Control parameter (used in the control function)
8 std::make_pair("control-params", { ...
9 0.100447*cos(7*0.1*t) + 0.120409*sin(7*0.1*t) +
10 0.215306*cos(6*0.1*t) + 0.118562*sin(6*0.1*t) +
11 0.177682*cos(5*0.1*t) + 0.126134*sin(5*0.1*t) +
12 0.100447*cos(7*0.1*t) + 0.120409*sin(7*0.1*t) +
13 0.103292*cos(8*0.1*t) + 0.108712*sin(8*0.1*t)
14 }); ...
15 // Fourier series (first 8 terms only)
16 std::make_pair("control-funcs", { ...
17 } ...
18 // Initial params
19 std::make_pair("initial-parameters", initParams), ...
20 std::make_pair("max-time", tMax) ...
21 }; ...
22 auto optimizer = xacc::getOptimizer("quantum-control", configs); ...
23 auto result = optimizer->optimize(); ...
24 auto optimal_a_b_vec = result.second; ...
```

Figure 13: Using XACC pulse optimization service (Gradient optimization of analytic controls - GOAT) to optimize a waveform given by a Fourier series (lines 13-21) to implement a target unitary (line 6) which can be given as a text input as in this example or as a matrix.

The optimal parameters that were found by the XACC pulse optimization plugin can then be verified on the QuaC backend as shown in the Python script in Fig. 14. The drive signal (analytical form), line 1-10, which is the result from the pulse optimizer, can be converted to a pulse sequence using the pulse generation utility (line 13) that we have described in 4.3.4. It can be then verified by the simulator back-end to confirm that the optimized pulse indeed implemented an X gate. Fig. 15 shows the time-domain data resulting from executing the pulse instruction in Fig. 14.

As we can see, the probability of $|1\rangle$ state and the expectation value of the $Z$ operator can confirm that this pulse indeed implement an $X$ gate.

Within the XACC framework, there are many modular and well-defined services that users can leverage for pulse-level experiments. For instance, the framework provides various standardized optimization and machine-learning services which can handle pulse-level optimal control problems. In a similar vein, by adhering to common framework interfaces, pulse-level services, especially the pulse-capable back-end, that we implement in this work, are interoperable with all other framework components that are listed in Fig. 1.

5.5 MPI Demonstration

Finally, as explained in Section 4.3.3, the QuaC Accelerator simulation backend is built upon the standard MPI runtime environment (through PETSc), and can therefore be easily parallelized to suit the underlying classical compute architecture. Here, we seek to demonstrate that capability by checking the simulation runtime under various MPI parallelization configurations, specifically the number of MPI processes.
The results of these scaling experiments are shown in Fig. 16 where we have normalized the runtime of the single-process configuration as one. In this test, we vary the system size in terms of the number of qubits and perform a pulse simulation whereby all qubits are driven by an arbitrary pulse.

![Figure 16: Performance scaling vs. number of MPI processes. The bigger the system size (in terms of the number of qubits, \(n_q\)) is, the closer to ideal (linear) performance scaling we can achieve. When simulating small systems, e.g. less than 5 qubits, it’s best to use the single processing mode. Hardware descriptions: desktop workstation, Core i7-8700, 6 (12) cores (threads).](image)

There are two scaling regimes of interest, (1) the perfect, linear scaling denoted by the green dash line in Fig. 16 and (2) the zero-scaling horizontal line at 1.0 (the red dash line in Fig. 16.) Anything below the horizontal line (2) is not suitable for MPI multi-processing, i.e. the inter-process communication and synchronization overheads exceed any parallel-processing gain.

In this regard, we can see that systems of small sizes (e.g. 4-5 qubits) are best to be simulated using a single MPI process. On the other hand, large systems can take advantage of MPI parallelism with a considerable amount of speed-up. The larger the system size is, the closer we can get to the ideal scaling regime.

6 DISCUSSION AND FUTURE WORK

We have extended the XACC framework to implement a unified end-to-end programming model consisting of multiple layers of abstraction from high-level constructs of algorithms and observables via gate-level representations to pulse-level decomposition. Thanks to this extension, entry-level users can experiment with dynamical simulation or pulse-level execution of their quantum applications without needing to change any of their code. On the other hand, by defining the common pulse-related API’s for the framework, professional users can access those API’s to provide custom, high-performance implementations, e.g. a custom pulse library and gate to pulse mapping, a specialized pulse scheduler, or a performant pulse-level simulator, and contribute those implementations as services to the framework. This work provides a much needed capability in the nascent quantum programming, compilation, and execution research landscape. We anticipate this work to enable new efforts with regards to quantum optimal-control, error mitigation and noise-suppression, and the development of novel multi-qubit pulse sequences for physical system simulation.

We note that pulse-level programming is an emerging field which we believe will certainly experience rapid iteration and advancement. Hence, our immediate future goals are to keep the framework relevant to latest industry standards and trends. In particular, we have plans to:

1. Adopt other pulse-level programming models as they emerge, e.g., Rigetti’s \texttt{Qulit}, in a continuous integration fashion.
2. Enable translation (where applicable) between pulse-level program representations across physical hardware implementations.
3. Expand and improve our pulse-related service capabilities such as applying quantum optimal control to pulse synthesis and optimizing quantum programs at pulse level (e.g. pulse scheduling).
4. Develop standardized pulse-level emulator environments which include high-quality, fully-calibrated quantum hardware system models and pulse libraries for learning and testing purposes.

ACKNOWLEDGEMENTS

This work has been supported by the US Department of Energy (DOE) Office of Science Advanced Scientific Computing Research (ASCR) Quantum Computing Application Teams (QCAT), Quantum Algorithms Team (QAT), and Accelerated Research in Quantum Computing (ARQC). ORNL is managed by UT-Battelle, LLC, for the US Department of Energy under contract no. DE-AC05-00OR22725. The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan.

REFERENCES

[1] John Preskill. Quantum computing in the nisq era and beyond. Quantum, 2:79, 2018.
[2] Kathleen E. Hamilton, Eugene F. Dumitrescu, and Raphael C. Pooser. Generative model benchmarks for superconducting qubits. Phys. Rev. A, 99:062323, Jun 2019.
[3] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski. Cloud quantum
computing of an atomic nucleus. Phys. Rev. Lett., 120:210501, May 2018.

[4] Alexander J. McCaskey, Zachary P. Parks, Jacek Jakowski, Shirley V. Moore, Titas D. Morris, Travis S. Humble, and Raphael C. Pooser. Quantum chemistry as a benchmark for near-term quantum computers. npj Quantum Information, 5(1):99, 2019.

[5] N. Klco, E. F. Dumitrescu, A. J. McCaskey, T. D. Morris, R. C. Pooser, M. Sanz, E. Solano, P. Lougovski, and M. J. Savage. Quantum-classical computation of schwinger model dynamics using quantum computer. Phys. Rev. A, 98:032331, Sep 2018.

[6] P. J. J. O’Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffery, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis. Scalable quantum simulation of molecular energies. Phys. Rev. X, 6:031007, Jul 2016.

[7] Abhinav Kandala, Anton Zuchuatop, Kristan Temme, Maika Takita, Markus Brink, Jerry M Chow, and Jay M Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. Nature, 549:242–245, sep 2017.

[8] Tiffany M. Mintz, Alexander J McCaskey, Eugene F Dumitrescu, Shirley V. Moore, Titus D. Morris, Travis S. Humble, and Raphael C. Pooser. Validating quantum-classical programming models with tensor network simulations. PLOS ONE, 13(12):1-19, 12 2018.

[9] Extran. http://github.com/ornl-qci/exatn. Accessed: 2020-03-17.

[10] Otte Matthew and Finkel Hal. Qua â ˘AˇT quantum in c, 2020.

[11] Alexander McCaskey, Eugene Dumitrescu, Dmitry Lyakh, and Travis Humble. Validating quantum-classical programming models with tensor network simulations. npj Quantum Information, 5(1):99, 2019.

[12] Qiskit. An open-source framework for quantum computing, 2019.

[13] Harry Putra Rudy, Mingi Ryu, Ninad Sathaye, Chris Schnabel, Ed-die Schoute, Kanav Setia, Yunong Shi, Adenilton Silva, Yukio Siraiuchi, Seyon Sirajah, John A. Smolin, Mathias Soeke, Hitomi Takahashi, Ivan Taverneli, Charles Taylor, Pete Tayloure, Kenzo Training, Matthew, Treinisch, Wes Turner, Desiree Vogt-Lee, Christophe Vuillot, Jonathan A. Wildstrom, Jessica Wilson, Erick Winston, Christopher Wood, Stephen Wood, Stefan Wörner, Ismed Yunus Akhalwata, and Christa Zoufalt. Qiskit: An open-source framework for quantum computing, 2019.

[14] Vladi Gheorghiu. Quantum ++: A modern C++ quantum computing library. arXiv e-prints, page arXiv:1412.4704, December 2014.

[15] Soeken. https://github.com/pybind/pybind11. Accessed: 2020-03-17.

[16] Shrirang Abhyankar, Jed Brown, Emil M Constantinescu, Debojoti Ghosh, Barry F Smith, and Hong Zhang. Petsc/c: A modern scalable odesdae solver library. arXiv preprint arXiv:1806.01437, 2018.

[17] Satish Balay, Shrirang Abhyankar, Mark Adams, Jed Brown, Peter Brune, Kris Buschelman, Lisandro Dalcín, Alp Dener, Victor Eijkhout, W Gropp, et al. Petsc users manual. 2019.

IEEE Transactions on Computers: Special Issue on Quantum Computing, 54(9):1036–1058, 9 2005.