Noise-Resilient Quantum Machine Learning for Stability Assessment of Power Systems

Yifan Zhou, Member, IEEE, and Peng Zhang, Senior Member, IEEE

Abstract—Transient stability assessment (TSA) is a cornerstone for resilient operations of today’s interconnected power grids. This paper is a confluence of quantum computing, data science and machine learning to potentially address the power system TSA issue. We devise a quantum TSA (QTSA) method to enable efficient data-driven transient stability prediction for bulk power systems, which is the first attempt to tackle the TSA issue with quantum computing. Our contributions are three-fold: 1) A high expressibility, low-depth quantum circuit (HELD) is designed for accurate and noise-resilient TSA; 2) A quantum natural gradient descent algorithm is developed for efficient HELD training and 3) A systematical analysis on QTSA’s performance under various quantum factors is performed. QTSA underpins a foundation of quantum machine learning-enabled power grid stability analytics. It renders the intractable TSA straightforward and effortless in the Hilbert space, and therefore provides stability information for power system operations. Extensive experiments on quantum simulators and real quantum computers verify the accuracy, noise-resilience, scalability and universality of QTSA.

Index Terms—Quantum machine learning, quantum neural network, power system stability, transient stability assessment.

I. INTRODUCTION

Today’s power systems are undergoing an Enlightenment, where the confluence of big data, quantum computing and machine learning altogether is to drive a regime shift in the analysis and operation of our critical power infrastructures. Big data is the force behind the revolution: massive new types of intelligent electronic sensors such as synchronized phasor measurement units (PMUs), advanced metering infrastructure (AMI) meters and remote terminal units (RTUs) [14] are continuously generating gigantic volumes of data which allow for the development of data-driven power system analytics. Most recently, the successes in exploiting the potential of quantum supremacy [15], [16] shed lights on a “quantum leap” of computing capabilities. The power of quantum computing is derived from its ability to prepare and maintain complex superpositions of quantum states across many quantum degrees of freedom. While classically the number of required physical resources \( N \) grows exponentially with the system complexity \( n \), \( N \) grows linearly with \( n \) in a quantum
computer, resulting in exponential improvement over classical computing [17]. Furthermore, highly entangled states, very difficult to represent on classical computers, are easily represented on a quantum computer [18]. Therefore, power system problems, if formulated properly through programmable quantum circuits, can be executed efficiently on a quantum computer.

Inheriting the quantum speedup in tensor manipulation [19], the swift growth in quantum machine learning (QML) techniques [20], [21] ignites new hopes of developing scalable and efficient data-driven power system analytics. QML is promising for data processing and model training in high-dimensional space [22], [23]. Ideally, unique quantum operators such as superposition and entanglement, which can not be efficiently represented by classical operators, enable a superior representation of complicated data relationships [24], [25]. Nevertheless, the existing noisy quantum devices are still restrictive, hindering the implementing of QML if deep quantum circuits are needed.

This paper is the first attempt to unlock the potential of QML for power system TSA. A high expressibility, low-depth quantum circuit-based transient stability assessment (QTSAs) method is devised to enable effective data-driven transient stability prediction. The contributions of this work are threefold:

- We devise a high expressibility, low-depth quantum circuit (HELD) for QTSAs that is feasible to pursue on near-term devices, considering the noisy-intermediate-scale quantum (NISQ) era [26], and also general enough to be directly expandable to the noise-free quantum computer. Through quantitative evaluations, we show that HELD exhibits the best expressibility in exploring the Hilbert space and in performing the TSA classification.
- We establish a quantum natural gradient descent algorithm to enable an efficient training of the HELD circuit along the steepest direction in the output space.
- We design systematical studies to comprehensively evaluate the TSA results under quantum environments regarding the accuracy, fidelity, and noise-resilience. By extensive experiments, we demonstrate the consistent high performance of QTSAs for real-scale power systems on both quantum simulators and real quantum computers.

The remainder of the paper is organized as follows. Section II devises the QTSAs methodology, including the overall framework, the HELD circuit design, and the training algorithm. Section III offers extensive case studies on both quantum simulators and real IBM quantum computers. Section IV concludes the paper.

II. QUANTUM CLASSIFIER FOR POWER SYSTEM TSA

A. Power System Transient Stability Assessment (TSA)

Power system transients are generically modelled as a set of nonlinear differential algebraic equations:

\[
\begin{align*}
\dot{X} &= F_D(X,Y) \\
0 &= F_A(X,Y)
\end{align*}
\]

where \(X\) and \(Y\) separately denote the differential variables and algebraic variables; (1a) formulates the nonlinear dynamics of power devices, such as generators (e.g., synchronous machines, distributed energy resources), controllers (e.g., governors, excitors, inverters), power loads, etc; (1b) formulates the instantaneous power flow of the entire power grid.

TSA appraises a power system’s capability of resisting large disturbance [5]. Denote \(Z = (X, Y)\), and \(\phi(t, Z)\) as the orbit of (1a) starting from \(Z\). An asymptotically stable equilibrium point (SEP) \(Z_s\) of (1) satisfies that:\n
(a) \(Z_s\) is Lyapunov stable; \(b\) there exists an open neighborhood \(O\) of \(Z_s\) such that \(\forall Z \in O\) converges to \(Z_s\) when \(t\) approaches infinity [10]. The stability region of \(Z_s\) encloses all the states that can be attracted by \(Z_s\) within an infinite time:

\[
\mathcal{A}(Z_s) = \{Z \in \mathbb{R}^n : \lim_{t \to \infty} \phi(t, Z) = Z_s\} \tag{2}
\]

Stability region theory states that system stability after a large disturbance is determined by whether the post-disturbance state is within the stability region of an SEP [10]. Therefore, to formulate the data-driven TSA, the idea is to establish a direct mapping between the post-disturbance power system states and the stability results [6], [13].

\[
\mathcal{A}(Z_s) = \{Z \in \mathbb{R}^n : \lim_{t \to \infty} \phi(t, Z) = Z_s\} \tag{2}
\]

B. Preliminaries of Quantum Computing

Machine learning tasks frequently involve high-dimensional computation. Quantum computing exhibits promising efficacy in manipulating high-dimensional vectors by tensor product of Hilbert spaces [19], which motivates us to exploit a quantum-assisted data-driven TSA.

In quantum computing, a quantum state is represented by a vector in a Hilbert space. A qubit is the fundamental element of a quantum system, which is described as \(|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle\). Here, \(|0\rangle = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]\) and \(|1\rangle = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]\) respectively denote the two-dimensional basis states; \(\alpha_0\) and \(\alpha_1\) are complex numbers; \(|\alpha_0|^2\) and \(|\alpha_1|^2\) respectively denote the probability on \(|0\rangle\) and \(|1\rangle\) satisfying \(|\alpha_0|^2 + |\alpha_1|^2 = 1\). With \(n\) qubits, an arbitrary quantum state can be modelled as \(|\psi\rangle = \sum_{i=1}^{2^n} \alpha_i |i\rangle\), where \(|i\rangle\) represents the computational basis and \(\sum_{i=1}^{2^n} |\alpha_i|^2 = 1\). The superposition and entanglement enable creating a complete \(2^n\)-dimensional vector space and performing the corresponding manipulations with only \(n\) qubits [18].

A quantum circuit is a computational routine on quantum states [21]. As illustrated in Fig. 1(a), a quantum circuit consist of an ordered sequence of quantum gates \(|\psi\rangle = A(p)|\psi_0\rangle\) followed by the measurements. When parameterized quantum gates are involved, the circuit is named as a variational quantum circuit (VQC). As illustrated in Fig. 1(b), a VQC implements parameterized quantum operations \(|\psi_P\rangle = A(p)|\psi_0\rangle\) according to a set of free parameters \(p\), and therefore the output state \(|\psi_P\rangle\) can be controlled by the VQC parameters. By optimizing \(p\), VQCs...
can be trained to perform various tasks (e.g., linear equation solver [27], eigensolver [28]).

Quantum machine learning (QML) takes advantages of VQCs to train a quantized mapping between inputs and outputs. Benefiting from the aforementioned quantum speedup, QML provides an exponential improvement on computational complexity in regard of distance evaluation and inner product compared with classical computing [19].

Our key innovation is a design of the high expressibility, low-depth circuit (HELD), as presented in Fig. 2. The high expressibility attribute addresses the dimensionality and non-linearity obstacles in TSA, enabling an efficacious stability prediction for real-scale power systems. The low-depth attribute addresses the nonnegligible noise and source limitations in the noisy-intermediate-scale quantum era, enabling an effective execution of QTSA on near-term quantum devices. Fig. 2 depicts the variational HELD circuit, whose kernel ingredients include (see the arrangement of quantum gates in Fig. 2(a) and the corresponding mathematical formulations in Section II-D):

1) Non-Gaussian feature encoding: $|\psi_E\rangle = U_E(p_E, Z, 0)|0\rangle$ which adopts parameterized, activation-enhanced quantum gates to enable a flexible, nonlinear and dimension-free encoding of power system stability features $Z$;

2) Properly-arranged Gaussian quantum gates: $|\psi_V\rangle = U_V(p_V)|\psi_E\rangle$ to efficiently represent the solution space;

3) Feature re-encoding: $|\psi\rangle = U_R(p_R, Z)|\psi_V\rangle$ for enhanced expressibility of nonlinear behaviors in power systems;

4) Repetitive layered structure: $|\psi\rangle = U(p, Z)|0\rangle \equiv U^{(1)} \cdots U^{(2)} U^{(1)}|0\rangle$ to realize a more expressive and entangled VQC, where $U^{(l)} = U_E^{(l)} U_V^{(l)} U_R^{(l)}$ denotes the unitary operation at $l$th layer and $p$ assembles the circuit parameters at each layer. The HELD circuit takes $|0\rangle$ as the input quantum state, whose state preparation is trivial.

Then, as presented in Fig. 2(b), QTSA employs a hybrid quantum-classical framework for HELD training. The parameterized HELD circuit is executed on a quantum computer as the feedforward functionality, and parameter optimization is executed on a classical computer as the backpropagation functionality. The two subroutines interact to train the VQC’s parameters. Here, we introduce a generalized quantum natural gradient to enable a more efficient training of QTSA: $p - \Delta p (\nabla L_{y, F_p}^{-1} \nabla y_p) \rightarrow p$, where $L$ is the nonconvex cross-entropy loss representing the correctness of quantum-embedded power system stability results; $y$ denotes the output of the quantum circuit; $\nabla y_p$ is the quantum gradient function; $\nabla L_p$ is the gradient of $L$ w.r.t. $y$; $F_p$ is the Fisher information matrix of the non-Gaussian quantum circuit; $\Delta p$ is the updating rule for circuit parameters (see details in Subsection II-E).

### C. Overall Framework of QTSA

The keystone of QTSA, different from classical machine learning techniques, is that the transient stability features in a Euclidean space $Z \in E$ are embedded into quantum states in a Hilbert space $|\psi\rangle \in \mathcal{H}$ through a VQC, which will explicitly separate the stable and unstable samples.

Our key innovation is a design of the high expressibility, low-depth circuit (HELD), as presented in Fig. 2. The high expressibility attribute addresses the dimensionality and non-linearity obstacles in TSA, enabling an efficacious stability prediction for real-scale power systems. The low-depth attribute addresses the nonnegligible noise and source limitations in the noisy-intermediate-scale quantum era, enabling an effective execution of QTSA on near-term quantum devices. Fig. 2 depicts the variational HELD circuit, whose kernel ingredients include (see the arrangement of quantum gates in Fig. 2(a) and the corresponding mathematical formulations in Section II-D):

1) Non-Gaussian feature encoding: $|\psi_E\rangle = U_E(p_E, Z, 0)|0\rangle$ which adopts parameterized, activation-enhanced quantum gates to enable a flexible, nonlinear and dimension-free encoding of power system stability features $Z$;

2) Properly-arranged Gaussian quantum gates: $|\psi_V\rangle = U_V(p_V)|\psi_E\rangle$ to efficiently represent the solution space;

3) Feature re-encoding: $|\psi\rangle = U_R(p_R, Z)|\psi_V\rangle$ for enhanced expressibility of nonlinear behaviors in power systems;

4) Repetitive layered structure: $|\psi\rangle = U(p, Z)|0\rangle \equiv U^{(1)} \cdots U^{(2)} U^{(1)}|0\rangle$ to realize a more expressive and entangled VQC, where $U^{(l)} = U_E^{(l)} U_V^{(l)} U_R^{(l)}$ denotes the unitary operation at $l$th layer and $p$ assembles the circuit parameters at each layer. The HELD circuit takes $|0\rangle$ as the input quantum state, whose state preparation is trivial.

Then, as presented in Fig. 2(b), QTSA employs a hybrid quantum-classical framework for HELD training. The parameterized HELD circuit is executed on a quantum computer as the feedforward functionality, and parameter optimization is executed on a classical computer as the backpropagation functionality. The two subroutines interact to train the VQC’s parameters. Here, we introduce a generalized quantum natural gradient to enable a more efficient training of QTSA: $p - \Delta p (\nabla L_{y, F_p}^{-1} \nabla y_p) \rightarrow p$, where $L$ is the nonconvex cross-entropy loss representing the correctness of quantum-embedded power system stability results; $y$ denotes the output of the quantum circuit; $\nabla y_p$ is the quantum gradient function; $\nabla L_p$ is the gradient of $L$ w.r.t. $y$; $F_p$ is the Fisher information matrix of the non-Gaussian quantum circuit; $\Delta p$ is the updating rule for circuit parameters (see details in Subsection II-E).

### D. Design of HELD

Our goal is to design a powerful high expressibility, low-depth quantum circuit that processes high-dimensional inputs, embeds nonlinear power features and thus can realize stability prediction for large-scale power systems. However, the existing noisy quantum devices only allow us to execute low-depth quantum circuits. To solve this dilemma, a high expressibility, low-depth quantum circuit is established as visualized in Fig. 2. Under a multilayer structure $U = U^{(1)} \cdots U^{(2)} U^{(1)}$, each layer of HELD consists of three blocks: a non-Gaussian feature encoding block $U_E^{(l)}$, a Gaussian variational block $U_V^{(l)}$ and a feature re-encoding block $U_R^{(l)}$ (see Fig. 2(a)):

$$U^{(l)}(p^{(l)}, Z) = U_R^{(l)}(p_R^{(l)}, Z) U_V^{(l)}(p_V^{(l)}) U_E^{(l)}(p_E^{(l)}, Z)$$

Here, $p^{(l)}$ denotes the parameters at layer $l$, i.e., an assembly of $p_E^{(l)}, p_V^{(l)}, p_R^{(l)}$ in each block.

The following details the mathematical representations of each block of the circuit designed in Fig. 2(a).

First, the feature encoding block (i.e., the first block in Fig. 2(a)) embeds the power system stability features into an $n$-qubit quantum state. This block applies single-qubit X rotations (i.e., RX gates) at each qubit. The representation of an RX gate is $R_X(\theta) = e^{-i \sigma_x \theta / 2} = [\cos(\theta / 2) \ -i \sin(\theta / 2) \ \cos(\theta / 2)]$, where $\theta$ is the rotating angle and $\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is the Pauli-X operator.
Different rotating angles of the RX gate lead to different quantum states [21]. Therefore, by embedding the power system features into the quantum gate angles, the output states can be controlled. Consequently, the feature encoding block is realized by a unitary transform as:

\[ U_E^{(l)}(p_E^{(l)}, Z) = \bigotimes_{k=1}^{n} \exp(-i\theta_0^{E,k}\sigma_{x,k}/2) \]  

where \((l)\) denotes the l-th layer; \(p_E^{(l)}\) denotes the parameters at the l-th layer; \(\otimes\) denotes the tensor product; \(\sigma_{x,k}\) denotes the Pauli-X operator applied on the k-th qubit; \(\theta_0^{E} = \sigma(p_E^{(l)} \cdot [Z; 1])\) denotes the rotating angles of the RX gates which adopt a single-layer perception of the input power system features to extract the expressibility on limited quantum resources; \(\sigma\) is the activation function, e.g. tanh, sigmoid, ReLU.

Then, the variational block is introduced to introduce entanglements between qubits, which is essential to enhance a quantum circuit’s ability to express the whole Hilbert space [29]. This block is composed of two parts: a ring structure of the RZZ gates (i.e., the CNOT-RZ-CNOT gates) on each two adjacent qubits [24] and parallel single-qubit Y rotations (i.e.,RY gates) on each qubit. The representations of an RZZ gate and an RY gate are respectively \(R_{\text{RZZ}}(\theta) = e^{-i\theta/2\sigma_z \otimes \sigma_z}\) and \(R_{\text{RY}}(\theta) = e^{-i\theta/2 = [\sin(\theta/2) - \sin(\theta/2)]\}, \) where \(\sigma_y = [0 \ 1 \ 1 \ 0]\) and \(\sigma_z = [1 \ 0 \ 0 \ 1]\) respectively denote the Pauli-Y and Pauli-Z operators. Therefore, the mathematical representation of the variational block is as follows:

\[ U_V^{(l)}(p_V^{(l)}) = \left( \prod_{k=1}^{n} I^{(k-1)} \right) \prod_{k=1}^{n} \exp(-i\theta_0^{V,k}\sigma_{z,k} \otimes \sigma_{z,k+1}/2) \]

where \(\theta_0^{V} = \theta_{V} + \sigma_{y,k}\) and \(\sigma_{z,k}\) separately denote the Pauli-Y and Pauli-Z operators on the k-th qubit. Specifically, for the 2-qubit HELD, no ring structure is required. The trainable rotations and CNOTs together generate highly entangled quantum states to capture non-trivial correlations in quantum states by a low-depth circuit, and therefore enable an efficient expression of the solution space [29].

Subsequently, feature encoding is repeated as a data-reencoding block [30] (i.e., the third block in Fig. 2(a)). The power system features are again embedded into the rotating angles of the RX gates to further enhance the expressibility for quantum stability assessment. In analogy to (4), the feature re-encoding block is formulated as:

\[ U_R^{(l)}(p_R^{(l)}, Z) = \bigotimes_{k=1}^{n} \exp(-i\theta_0^{R,k}\sigma_{x,k}/2) \]

where \(\theta_0^{R} = \sigma(p_R^{(l)} \cdot [Z; 1]).\)

An arbitrary qubit of the HELD circuit can be chosen as the output qubit. A single qubit describes a two-dimensional quantum state. A mapping between stable (unstable) and higher (lower) possibility on \([1]\) therefore enables the binary classification functionality of TSA. Without loss of generality, in this paper, the measurement on the first qubit, denoted as \(y_i\), is selected to generate the TSA prediction result. On a quantum simulator, the expectation value of \(y = [\langle 0 | U(p, Z) | M U(p, Z) | 0 \rangle]\) of the observable \(\mathcal{M}\) is computed, where \(p\) assembles circuit parameters (i.e., \(p^{(l)}\)) at each layer. On a real quantum machine, the multi-shot implementation is performed to estimate \(y\).

### E. QTSA Training Through Quantum Natural Gradient

In this subsection, we establish a generalized quantum natural gradient (QNG) algorithm for efficient QTSA training. QNG is a promising algorithm for quantum optimization. It outperforms the commonly-used quantum gradient descent because of the capability of searching the steepest descent direction in the output space [31], [32]. However, existing QNG algorithms mostly resolve quadratic objective minimization of Gaussian quantum circuits [33], [34], which do not adapt to the non-Gaussian HELD training. Therefore, the following develops a generalized QNG which is able to deal with arbitrary quantum circuits and nonconvex loss functions.

Consider a training set \(\{\{\hat{Z}_i, \hat{y}_i\}\}_{i=1}^{n}\), where \(\hat{Z}_i\) and \(\hat{y}_i\) respectively denote the TSA feature vector and the corresponding stability label of the \(i\)-th sample. The parameterized HELD circuit is optimized on the training set with a cross-entropy loss function:

\[ \mathcal{L}(p) = \sum_{i=1}^{n} l(y_i(p, \hat{Z}_i)) = \sum_{i=1}^{n} -\log(y_i(p, \hat{Z}_i)) \]

Here, \(y_i = [\langle 0 | U(p, \hat{Z}_i) | M U(p, \hat{Z}_i) | 0 \rangle]\) is the QTSA’s prediction of the \(i\)-th sample; \(\mathcal{M}\) is the corresponding observable, i.e., \(\mathcal{M}_i = [\langle 0 | 0 \rangle]\) for \(\hat{y}_i = 0\) (i.e., unstable) and \(\mathcal{M}_i = [1 | 1 \rangle\) for \(\hat{y}_i = 1\) (i.e., stable).

Denoting \(\theta\) as an ensemble of gate rotation angles (i.e., \(\theta_E, \theta_V, \theta_R\)), a second-order Taylor expansion of \(y_i\) w.r.t \(\theta\) yields:

\[ dy_i \approx \frac{\partial y_i}{\partial \theta} d\theta + \frac{1}{2} \frac{\partial^2 y_i}{\partial \theta^2} d\theta^T d\theta \]

Here, \(\frac{\partial y_i}{\partial \theta}\) is the quantum gradient calculated by the parameter-shift rule [32]; \(f_i(\theta)\) is the quantum Fisher information matrix regarding a Gaussian quantum circuit, i.e., \(f_{i,j} = \text{Re}((\frac{\partial^2 y_i}{\partial \theta^2}, \frac{\partial \psi_i}{\partial \theta^T}) - (\frac{\partial \psi_i}{\partial \theta^T}, \frac{\partial \psi_i}{\partial \theta^T})\) and \(\psi_i = U(p, \hat{Z}_i)\) based on the Fubini-Study metric [33].

Based on (4)–(6), substituting \(d\theta = \frac{\partial y_i}{\partial \theta} d\theta\) into (8) yields:

\[ dy_i \approx \frac{\partial y_i}{\partial p} dp + \frac{1}{2} dp^T \left( \frac{\partial^2 y_i}{\partial \theta^2} \frac{\partial \psi_i}{\partial \theta^T} \right) dp \]

\[ F_i(p) = \frac{\partial \psi_i^T}{\partial \theta} \frac{\partial \psi_i}{\partial \theta} \]

establishes the quantum Fisher matrix for a generic non-Gaussian quantum circuit. \(F_i\) represents an approximation of Hessian matrix of \(y_i\) at \(p\) and therefore enables searching the local minimal of \(y_i\) around \(p\) with second-order terms:

\[ g^y_p = \arg \min_{\Delta p} \left[ \frac{\partial y_i}{\partial p} \Delta p + \frac{1}{2} \Delta p^T F_i(p) \Delta p \right] = -F_i^+(p) \frac{\partial y_i}{\partial p} \]

(10)
where $g^p_k$ denotes the natural gradient of $y$ w.r.t $p$; $F^+_i$ denotes the pseudo-inverse of $F_i$.

Correspondingly, QNG is established for (7), where $L$-optimization is iteratively solved at the steepest descent direction in the distribution space:

$$
g^L_{p,k} = - \sum_{i \in B} \frac{\partial l}{\partial y_i} F^+_i (p_k) \frac{\partial y_i}{\partial p_k}$$

$$p_{k+1} = p_k - \frac{\eta}{\sqrt{\nu_k + \xi}} \hat{m}_k$$

Here, $B$ denotes the batch set; $k$ denotes the $k$-th iteration; $\eta$ denotes the learning rate; $\nu_k$ and $\xi$ jointly realize an adaptive learning rate [35], where $\hat{m}_k = m_k/(1 - \beta_k)$, $\nu_k = v_k/(1 - \beta_k)$, $m_k = \beta_1 m_{k-1} + (1 - \beta_1) g^L_{p,k}$, and $\nu_k = \beta_2 v_{k-1} + (1 - \beta_2) (g^L_{p,k})^2$; $\xi$ is a small positive constant to ensure the numerical stability; $\beta_1$ and $\beta_2$ are hyperparameters governing the moment updates.

Next, we show that the non-Gaussian quantum Fisher matrix $F_l$ can be effectively approximated in a block-diagonal form. For concision, the subscript $i$ is omitted in the following derivation. Based on (4)–(6), we have:

$$\frac{\partial \theta}{\partial p} = \begin{bmatrix} \nabla \theta^{(1)}_p & 0 & \cdots & 0 \\ 0 & \nabla \theta^{(2)}_p & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \nabla \theta^{(L)}_p \end{bmatrix}$$

where:

$$\nabla \theta^{(1)}_p = \begin{bmatrix} \nabla \sigma[Z;1]^T \\ 0 \\ I \\ 0 \end{bmatrix} \begin{bmatrix} \theta^{(1)}_E \\ \theta^{(2)}_I \\ \theta^{(L)}_R \end{bmatrix}$$

According to [33], the Gaussian Fisher matrix $f$ also follows a block-diagonal form. Consequently, the non-Gaussian Fisher matrix is constructed as:

$$F(p) = \begin{bmatrix} F^{(1)}(p) & 0 & \cdots & 0 \\ 0 & F^{(2)}(p) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F^{(L)}(p) \end{bmatrix}$$

Here, $F^{(l)}(p) = (\nabla \theta^{(l)}_p)^T f^{(l)} \nabla \theta^{(l)}_p$, $f^{(l)}(\theta)$ denotes the Gaussian Fisher matrix block of the $l$-th layer, where $f^{(l)}_{jk} = \langle \psi^{(l)} | K_j K_k | \psi^{(l)} \rangle - \langle \psi^{(l)} | K_j | \psi^{(l)} \rangle \langle \psi^{(l)} | K_k | \psi^{(l)} \rangle$ [33]. Matrix multiplication and pseudo-inverse of a block-diagonalized $F(p)$ would be highly efficient.

Aforementioned derivations establish the generalized QNG algorithm, which adopts Riemannian metrics and enables efficient training of the HEld circuit with arbitrary gates and targets. Therefore, the optimization of HEld will be performed along the steepest descent direction in the output space, whereas the conventional quantum gradient descent is only capable of updating parameters following the steepest descent direction in the parameter space with Euclidean metrics.

### III. Numerical Experiments

Extensive experiments of QTSA are conducted on typical power systems ranging from benchmark grids to a real-scale U.S. power grids. QTSA is coded with Qiskit [36] and Pennylane [37], and is implemented on both a noise-free quantum simulator (ibmq_qasm_simulator, a 32-qubit simulator) and real IBM quantum devices (a 20-qubit machine ibmq_boeblingen, and a 7-qubit machine ibm_lagos). Power system transient simulations with random fault locations and random fault clearing time are conducted by Power System Toolbox (PST) [38]. Those physical model-based simulations will generate the features for QTSA (such as power outputs, rotor angles and rotor speeds of generators, power flows through transmission lines, and bus voltage angles and amplitudes) as well as providing the ground truth of the system stability.

The following sections investigate the capability of QTSA in assessing power system stability on both noise-free simulators and near-term noisy quantum computers, as well as the performance of different quantum circuit designs for QTSA.

#### A. Validity of QTSA on Noise-Free Quantum Simulator

Four typical power systems are studied:

- **Single-machine infinite-bus (SMIB) system**, one of the most widely-used test systems in power system research. Since the stability region of SMIB (i.e., a 2-dimensional region) can be analytically computed, SMIB always serves as an indispensable benchmark for testing the performance of a TSA method.

- **Two-area system**, another most widely-used benchmark system exhibiting both local and inter-area oscillation modes. The system models and parameters are constructed from a real North American power system [5].

- **Northeast Power Coordinating Council (NPCC) test system**, a real Northeastern US power system which was involved in the Northeast blackout of 2003.

- **IEEE 300-bus system**, a large power system which contains 69 generators, 304 transmission lines, and 195 loads. Generators are formulated using the classical synchronous machine model.

Appendix A presents the test system parameters and the settings of the training and testing sets. The overall QTSA procedure is exemplified via the SMIB system. We then verify the efficacy of QTSA on all the test systems.

1. **Training Preparation:** Various dynamics samples are generated on the SMIB system (see Fig. 3(a)). Time-domain simulations of SMIB under random disturbances are performed to obtain stability features (i.e., rotor angle and rotor speed in this case). Fig. 3(b) shows both the stable cases with dampened oscillations and the unstable cases where power generators run out-of-step. The initial states of the post-disturbance system
therefore construct the feature set for QTSA training, as shown by the samples in Fig. 3(c) (green for stable samples and red for unstable samples). Our hope is that QTSA can directly distinguish the stable and unstable samples in the Hilbert space.

2) QTSA Training: A HELD circuit with 2 qubits, 6 layers (denoted by HELD(2,6)) is constructed for SMIB. Fig. 3(d) to (f) show how an untrained HELD evolves into a well-trained one. We use Bloch sphere and quantum states overlap to evaluate the QTSA efficacy in the Hilbert space. A point \((\theta, \phi)\) on the Bloch sphere\(^1\) represents a single-qubit state \(\cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle\). The data overlap between each pair of samples represents the fidelity between the corresponding quantum states.

Starting with a randomly-initialized HELD, Fig. 3(d) shows that the stable and unstable samples are mixed on the Bloch sphere and quantum states corresponding to stable and unstable samples are highly overlapped, meaning the untrained circuit cannot be used for TSA. Then, Fig. 3(e) shows the evolution of loss function during the QTSA training, where the classification accuracy exceeds 99% for both stable and unstable samples at the final stage. The Hilbert space observations in Fig. 3(f) demonstrate the efficacy of the trained HELD. It can be seen that the well-trained quantum circuit successfully embeds the stable (unstable) samples to the lower-(upper-) half Bloch sphere. As can be seen from the data overlap, the mutual fidelity between samples from the same class is almost 1, while that from different classes is close to 0. Both observations verify a successful QTSA being able to clearly separate the opposite classes through the VQC.

3) Stability Insights by QTSA: Fig. 3(g) shows that QTSA is able to accurately discover power system stability regions \(A(Z_s)\) (see (2)). As illustrated by the simulation results, there is a faithful match between the stable/unstable regions learned by QTSA (i.e., the green and red regions) and those obtained from model-based analytical solutions (i.e., the black line). Stability

---

\(^1\)The Bloch sphere geometrically represents a single-qubit quantum states on a unit three-dimensional sphere. A state \(|\psi\rangle = \alpha |0\rangle + \beta |1\rangle\) can always be rewritten as \(|\psi\rangle = e^{i\gamma} (\cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle)\), where \(\gamma, \theta, \phi\) are real numbers. Therefore, \(\theta\) and \(\phi\), as the spherical co-ordinates, jointly define a point on the Bloch sphere for \(|\psi\rangle\).

---

Fig. 3. Demonstration of QTSA procedure on the SMIB test system. Subplots (a)–(c) demonstrate training preparation. Subplots (d)–(f) demonstrate QTSA training. Subplots (g)–(i) demonstrate the stability insights obtained by QTSA learning. (a): Topology of the SMIB system. (b): Time-domain simulations of SMIB under random disturbances. (c): Samples in the classical feature space (green for stable samples and red for unstable samples). (d): Samples embedded in the Hilbert space by a randomly-initialized HELD. (e): Loss function evolution during QTSA training process and the final confusion matrix. (f): Samples embedded in the Hilbert space by the well-trained HELD, where the stable and unstable samples are distinctly separated. (g): Stable and unstable regions discovered by the well-trained HELD and its comparison with the physical model-based analytic stability boundary (i.e., black line). (h): Contour plot of probability on measuring \(|1\rangle\) from HELD. Either a high probability on \(|1\rangle\) or a high probability on \(|0\rangle\) (i.e., low probability on \(|1\rangle\)) reflects a high fidelity of transient stability prediction. (i): Stability boundaries discovered by QTSA under different fidelity levels.
regions discovered by QTSA allow power system operators to identify the system stability under disturbances, i.e., the SMIB system will remain stable if its post-contingency state locates in the green zone, whereas it must collapse if the post-contingency state locates in the red zone.

Another interesting observation is that QTSA offers not only the stability classification results but also the fidelity of stability or instability. Fidelity measures the similarity of quantum states \( |\psi_1\rangle \) and \( |\psi_2\rangle \). The fidelities between \( |\psi_1\rangle \) and stability (i.e., \( |1\rangle \)) and instability (i.e., \( |0\rangle \)) are respectively \( F_{1} = |\langle \psi_{1} | 1 \rangle|^{2} = |\alpha_{1}|^{2} \) and \( F_{0} = |\langle \psi_{1} | 0 \rangle|^{2} = |\alpha_{0}|^{2} \), which can be directly measured from the multi-shot executions of the quantum circuit.

As shown in Fig. 3(h), for a post-contingency state falling in the stable zone, the farther it deviates from the stability boundary, the system stability (\( |1\rangle \)) would be assured with a higher probability. Similarly, the system would collapse (\( |0\rangle \)) with a higher probability if its state lies in the deeper red zone. The operating points with probabilities near 0.5 indicate marginal cases, which are observed to be located near the stability boundary. The probabilities associated with system stability provide valuable new information for system operators and decision makers to make more suitable planning, operation and remedial action schemes. Either the risk-averse or risk-takers will then selectively use such information to optimize the social welfare and improve electricity resilience.

Further, QTSA is able to calculate stability regions at different risk levels. Fig. 3(i) exhibits the stability boundaries corresponding to different fidelity thresholds. Being less risk-tolerant leads to shrunken stability regions because higher fidelity levels are adopted.

4) Verification on Practical Power Systems: We further exhibit the versatility and efficacy of QTSA in the stability assessment of large power systems. Fig. 4 presents the QTSA results for the two-area system, the NPCC system and the IEEE 300-bus system. For each system, as shown in Fig. 4(b), (f) and (j), their high-dimensional features are irregularly distributed in the classical space and their stable/unstable samples are highly mixed. Once a trained HELD embeds the classical features into quantum states, it successfully arranges the stable and unstable samples onto the upper- and lower- halves of a Bloch sphere, rendering the challenging power system stability identification straightforward and effortless in the Hilbert space.

Table I quantitatively evaluates the accuracy of QTSA. For the small- and medium-scale power system, QTSA achieves high

\[ F(\psi_{1}, \psi_{2}) = |\langle \psi_{1} | \psi_{2} \rangle|^{2} = (\sum \alpha_{1i} \alpha_{2i})^{2}. \]
Increasing the number of qubits also leads to improved accuracy on both the training set (> 99%) and test set (> 98%). Even for the large-scale systems such as the NPCC grid and IEEE 300-bus grid, QTSA exhibits satisfactory performance of 98% accuracy on the training set and 95% accuracy on the testing set. QTSA is further compared with two data-driven TSA methods based on classical machine learning, i.e., a support vector machine (SVM) with the radial basis function kernel, and a deep neural network (DNN) with a three-layer perceptron architecture implemented in Scikit-learn [39]. Both the F1-scores and accuracy validate QTSA’s competency and high performance in comparison with classical data-driven TSA approaches.

\[ \text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \]
\[ \text{Precision} = \frac{TP}{TP + FP} \]
\[ \text{Recall} = \frac{TP}{TP + FN} \]
\[ \text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \]

1 Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN}, F1-Score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.

\begin{table}[h]
\centering
\caption{Comparison of accuracy: QTSA vs. TSA based on classical machine learning methods}
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
 & & & & & & \\
 & Training Set & Test Set & & & & \\
 & QML & SVM & DNN & QML & SVM & DNN \\
\hline
SMIB System & & & & & & \\
Accuracy & 0.9906 & 0.9525 & 0.9913 & 0.9810 & 0.9510 & 0.9865 \\
Precision & 0.9867 & 0.9672 & 0.9867 & 0.9726 & 0.9754 & 0.9855 \\
Recall & 0.9911 & 0.9185 & 0.9926 & 0.9820 & 0.9050 & 0.9820 \\
P1-Score & 0.9889 & 0.9422 & 0.9897 & 0.9773 & 0.9389 & 0.9837 \\
\hline
Two-Area System & & & & & & \\
Accuracy & 0.9975 & 0.9913 & 0.9924 & 0.9660 & 0.9900 & 0.9850 \\
Precision & 0.9972 & 0.9917 & 0.9917 & 0.9795 & 0.9866 & 0.9779 \\
Recall & 0.9991 & 0.9981 & 1.0000 & 1.0000 & 0.9985 & 1.0000 \\
P1-Score & 0.9981 & 0.9949 & 0.9958 & 0.9896 & 0.9925 & 0.9888 \\
\hline
NPCC System & & & & & & \\
Accuracy & 0.9890 & 0.9550 & 0.9765 & 0.9530 & 0.9470 & 0.9560 \\
Precision & 0.9931 & 0.9442 & 0.9822 & 0.9560 & 0.9455 & 0.9582 \\
Recall & 0.9846 & 0.9854 & 0.9798 & 0.9722 & 0.9757 & 0.9757 \\
P1-Score & 0.9938 & 0.9644 & 0.9810 & 0.9640 & 0.9604 & 0.9669 \\
\hline
IEEE 300-bus System & & & & & & \\
Accuracy & 0.9890 & 0.9385 & 0.9645 & 0.963 & 0.9300 & 0.9460 \\
Precision & 0.9768 & 0.9349 & 0.9573 & 0.9674 & 0.937 & 0.9397 \\
Recall & 0.9931 & 0.9741 & 0.9901 & 0.9775 & 0.9595 & 0.9820 \\
P1-Score & 0.9849 & 0.9541 & 0.9734 & 0.9724 & 0.9481 & 0.9604 \\
\hline
\end{tabular}
\end{table}

- Increasing the number of qubits also leads to improved QTSA performance, whereas the expressibility saturation is again observed. A noteworthy observation is that QTSA manifests impressive expressibility even with a single qubit, where the classification accuracy still reaches 97.38%.
- Although more layers and qubits definitely boost the expressibility of HELD, an overscale quantum circuit is not recommended. Such a circuit would demand prohibitively expensive quantum resources due to the saturation phenomenon. It also aggravates the training burden of the variational quantum circuit accompanied with the increased parameters. Finally, a large-scale quantum circuit is prone to the noisy quantum environment, leading to a sharply declined TSA performance on real quantum devices (which will be further discussed in the next subsection).

2) Impact of Circuit Layer Structure: Fig. 6 compares the performance of HELD with those of three typical variational quantum circuit (VQC) structures, i.e., an Instantaneous Quantum Polynomial (IQP) circuit [24], the Quantum Approximate Optimization Algorithm (QAOA) inspired circuit [25], and a data reuploading circuit [30]. Our HELD circuit exhibits the best expressibility benefiting from its non-Gaussian feature encoding and entanglement layers.

Further, we quantitatively analyze the expressibility of HELD. Two expressibility indices are introduced: \( E_{Haar} \) describes a VQC’s ability to express the Hilbert space [29], and \( E_{TSA} \) describes a VQC’s ability to express the TSA classification.

\[ E_{Haar} = D_{KL}(P_{VQC}(f, p) || P_{Haar}(f)) \] (15)
\[ E_{TSA} = D_{KL}(P_{VQC}(f_y, p^*) || P_{TSA}(f_y)) \] (16)

Both \( E_{Haar} \) and \( E_{TSA} \) are defined based on the Kullback-Leibler (KL) divergence \( D_{KL}(P || Q) = \sum_i P(i) \ln \frac{P(i)}{Q(i)} \), which is widely used in machine learning to quantify the difference between an estimated distribution \( P \) and the ideal distribution \( Q \). A smaller KL-divergence indicates a better expressibility. Specifically, \( E_{Haar} \) denotes the KL divergence between \( P_{VQC}(f, p) \) (i.e., the distribution of fidelity \( f \) estimated from a VQC parameterized by \( p \)) and the Haar distribution \( P_{Haar}(f) \) (i.e., the probability distribution under the Haar random states) [29]. \( E_{TSA} \) denotes the KL divergence between \( P_{VQC}(f_y, p^*) \) (i.e., the distribution of the stability fidelity \( f_y = |\langle 1 | U(p^*) | 0 \rangle |^2 \) estimated by the optimized VQC) and

\[ \text{Fig. 5. QTSA performances with different HELD circuit scales. The architecture of HELD is depicted in Fig. 2. An improved separation between the stable and unstable classes is indicated by improved accuracy and decreased } \text{TSA} \text{ efficiency}. \]
Fig. 6. Quantum TSA performances with different quantum circuit designs. Our HELD circuit employs 2 qubits and 6 layers. The IQP and QAOA circuits both employ 3 qubits and 10 layers. Three indices are employed to evaluate the transient stability prediction capability of quantum circuits: classification accuracy, $F_1$-Score and $\text{Tr}(\sigma_0\sigma_1)$. An improved separation between the stable and unstable classes is indicated by improved accuracy, increased $F_1$-Score and decreased $\text{Tr}(\sigma_0\sigma_1)$.

Fig. 7. Expressibility of different VQCs for exploring the Hilbert space.

Fig. 8. Expressibility of different VQCs for TSA classifications.

C. Validity of QTSA on Real, Noisy Quantum Environment

Today’s quantum devices are limited by gate errors and qubit connectivity, which hinders the implementation of deep quantum circuits. As an ultimate test of the practicality of QTSA, we now implement it on a real quantum computer and systematically verify it under noisy quantum computing environments.

1) Verification on the Real IBM Quantum Computers:

In this subsection, we run QTSA on real IBM Quantum devices. QTSA of the SMIB system, 2-area system and NPCC system is tested on ibmq_boelingen which is a 20-qubit superconducting quantum computer (QC). QTSA of the IEEE 300-bus system is tested on a 7-qubit QC ibm_lagos. The number of shots is set as 1024. Fig. 9 presents the configuration of ibmq_boelingen as well as its error map. The average error rates of the single-qubit Hadamard gate and the double-qubit CNOT gate are respectively 0.038% and 1.308%.

Table II presents QTSA’s accuracy obtained from real IBM quantum computers, where the stability identification results are again obtained by HELD(2,6). Compared with those results on the noise-free IBM Quantum simulator ibmq_qasm_simulator, a surprising finding is that a noisy real quantum computing environment has little effect on the performance of QTSA and the Hilbert space effectively. Nevertheless, simulation results in Fig. 8 show that those VQCs exhibit essentially different expressibility for TSA classifications. HELD achieves the best performance regarding $E_{TSA}$. In contrast, other types of VQCs lead to considerable KL divergence between the true stability distribution and the estimated stability distribution, which conforms with the fact that they mix plenty of stable and unstable samples. This observation shows that HELD not only explores the Hilbert space effectively but also well expresses the stable and unstable samples.
the overall classification is still of high quality. Even in the most challenging case of the IEEE 300-bus system, only a 0.5% decrease in the accuracy is reported, and the QTSA results remain compelling as compared with those from classical machine learning algorithms (see Table I). This experiment exhibits the effectiveness of QTSA on the near-term quantum devices and verifies the inherent resilience of QTSA against noisy quantum computing environments.

Fig. 10(a) further investigates the rationale behind the resilience of QTSA. Accordingly, there are four types of results: green dots (samples predicted as stable on both the simulator \textit{i.e.}, \textit{ibmq_qasm_simulator} and the real QC \textit{ibmq_boelingen}); red dots (samples predicted as unstable on both devices); yellow dots (samples predicted as stable on the simulator but unstable on the real QC); blue dots (samples predicted as unstable on the simulator but stable on the real QC). It can be seen that samples of inconsistent prediction from \textit{ibmq_boelingen} and \textit{ibmq_qasm_simulator} (\textit{i.e.}, the blue and yellow dots) are mostly distributed around the stability boundary. Further, Fig. 10(a) exemplifies the quantum state probability measured from HELD by two typical samples. Sample (A) illustrates an identical prediction on both devices, which represents the most cases (green and red dots), whereas sample (B) illustrates an opposite prediction, which is observed mainly around the stability boundary. The measurement probability distribution of samples (A) and (B) shows that the measurement probabilities from \textit{ibmq_boelingen} and \textit{ibmq_qasm_simulator} do not differ much. This indicates HELD(2,6) is only slightly perturbed by the noisy quantum device. Therefore, only for the low-fidelity areas where the quantum circuit output already has a similar probability on both \{0\} and \{1\}, a slight perturbation from the noisy environment would possibly produce an opposite classification result, as exemplified by sample (B). According to our previous study in Fig. 3(h), those samples locate in a narrow area around the stability boundary, which are actually of low stability margin, and thus their stability by their very nature is uncertain. Whereas, for the high-fidelity areas as exemplified by sample (A), the quantum circuit output in the noisy device does not change the final QTSA prediction. Therefore, QTSA generates reliable results for most samples even on noisy QCs.

Even though HELD(2,6) maintains high accuracy, a larger scale HELD circuit may fail to produce similar results. Fig. 10(b) illustrates the performance of HELD of different scales on noisy QCs, where HELD(2,6) as the default case, HELD(2,10) with 10 layers and HELD(3,6) with an additional qubit are compared.

- On the first qubit (\textit{i.e.}, the output qubit), we measure its probability on the basis quantum state \{1\} based on 1024 shots. Denote the probability on \{1\} as \(|\alpha_1|^2\). By gathering \(|\alpha_1|^2\) for each sample \(i (\forall i = 1, \ldots, n)\), the probability density function of \(|\alpha_1|^2\) can be estimated. The x-axis and y-axis of Fig. 10(b) correspondingly refer to \(|\alpha_1|^2\) and the PDF/CDF of \(|\alpha_1|^2\). To best classify the stable and unstable samples, \(|\alpha_1|^2\) is expected to be concentrated around 0 and 1 (\textit{i.e.}, high fidelity for instability/stability).

- On the noise-free \textit{ibmq_qasm_simulator}, all circuits achieve high accuracy over 99% with a notable feature that quantum measurements mainly accumulate around 0 and 1. This indicates that most samples are predicted at high confidences, which is coincident with Fig. 3(h), \textit{i.e.}, only the narrow areas around the stability boundary are of low fidelity.

- For HELD(2,6) running on the noisy \textit{ibmq_boelingen}, the peaks of the probability distribution slightly shift towards the center compared with that from the simulator, indicating a slightly decreased prediction fidelity.

- For HELD(2,10), although its accuracy on the noisy QC remains comparable to that of HELD(2,6), the shift of probability distribution peaks is more obvious, reflecting a further decreased fidelity on the noisy prediction.

- For HELD(3,6) involving more qubits, more CNOT gates and deeper circuit depth, the accuracy on the real QC deteriorates down to 52.02%, making the evaluation on the noise-free simulator meaningless. The probability distribution histogram shows most of the measurements center around 0.5, indicating that output states from QTSA are nearly random due to the noisy environment and quantum decoherence.

- The degraded performance of HELD(3,6) is induced by the increased number of CNOT gates. On the one hand,

| Table II |
| --- |
| **COMPARISON OF QTSA ACCURACY: NOISY REAL QUANTUM DEVICES VS. NOISE-FREE SIMULATOR ibmq_qasm_simulator** |
| Training Set | Test Set |
| --- | --- |
| **SMIB System** | | |
| Accuracy | 0.9819 | 0.9906 |
| Precision | 0.9708 | 0.9867 |
| Recall | 0.9867 | 0.9911 |
| F1 Score | 0.9787 | 0.9889 |
| **Two-Area System** | | |
| Accuracy | 0.9907 | 0.9975 |
| Precision | 0.9882 | 0.9972 |
| Recall | 0.9980 | 0.9991 |
| F1 Score | 0.9931 | 0.9981 |
| **NPCC System** | | |
| Accuracy | 0.9770 | 0.9900 |
| Precision | 0.9727 | 0.9831 |
| Recall | 0.9806 | 0.9846 |
| F1 Score | 0.9767 | 0.9838 |
| **IEEE 300-bus System** | | |
| Accuracy | 0.9780 | 0.9800 |
| Precision | 0.9746 | 0.9768 |
| Recall | 0.9924 | 0.9931 |
| F1 Score | 0.9834 | 0.9849 |

[Fig. 9. Configuration and error map of the ibmq_boelingen 20-qubit quantum computer. Figure courtesy of IBM TJ Watson Research Center [40].]
Fig. 10. Implementation of QTSA on the noisy IBM Quantum computer (ibmq_boelingen). (a): Stability predictions obtained on the simulator ibmq_qasm_simulator and the real noisy QC ibmq_boelingen. Two typical samples exemplify the quantum state probability measured from HELD. Sample (A) illustrates an identical prediction on both devices, which represents the most cases (green and red dots), whereas sample (B) illustrates an opposite prediction, which is observed mainly around the stability boundary. (b): Probability distribution function (PDF) and cumulative distribution function (CDF) of the probability of measuring $|1\rangle$ from different-scaled HELD circuits.

![Stability region discovery by ibmq_boelingen](image)

![Distribution of probability on [1] with different HELDs](image)

**Fig. 11.** Transpiled HELD circuit on the real QC ibm_lagos.

3The IBM quantum computer ibm_lagos only has 5 basis gates, i.e., CX (the controlled-X gate), ID (the identity gate), RZ (the single-qubit rotation about the Z axis), X (the Pauli-X gate) and SX (the sqrt-X gate). Therefore, when executed on ibm_lagos, the RX and RY gates designed in Fig. 2 are transpiled to basis gates (i.e., RZ gates, CX gates and SX gates) in Fig. 11.

Additionally, Table III presents the runtime of the quantum-based TSA on the real QC ibm_lagos. The runtime of each job in IBM Qiskit mainly comprises the time consumption for transpilation, validation, and execution.

2) Impact of Various Noise Levels: Since different quantum devices may have different noise features, it is necessary to systematically examine the QTSA’s performance under various noisy environments. We adjust the noise levels through a noise module provisioned by IBM and the noise settings are modified from the ibmq_boelingen’s real noise data. As Fig. 12 shows, excessive gate error and decoherence time certainly lead to the collapse of the quantum circuit. Nevertheless, high TSA accuracy can still be obtained with gate errors smaller than 2% or the thermal relaxation time larger than $10\mu s$, which can be easily satisfied by today’s quantum hardware. The QTSA’s excellent performance is therefore universal.

3) Remarks: (a) With HELD, quantum computers with limited qubit resources can still handle high-dimensional TSA problems and achieve satisfactory accuracy. The reason is because HELD employs TSA features as rotation parameters of the controllable quantum gates (see the parameterized rotation angles $\theta$ in (4), (5) and (6)) rather than encoding them into

**TABLE III**

| HELD Scale | Transpilation | Validation | Execution | Sum |
|------------|---------------|------------|-----------|-----|
| (2.2)      | 0.919         | 1.173      | 4.931     | 7.023 |
| (2.4)      | 0.750         | 1.221      | 4.892     | 6.863 |
| (2.6)      | 0.842         | 1.136      | 5.022     | 7.000 |
| (2.8)      | 0.766         | 0.999      | 4.988     | 6.753 |
| (3.2)      | 0.900         | 1.579      | 5.296     | 7.775 |
| (3.4)      | 0.625         | 1.177      | 5.078     | 6.880 |
| (3.6)      | 0.832         | 0.845      | 5.179     | 6.856 |
| (3.8)      | 0.924         | 1.080      | 5.174     | 7.042 |
quantum states. The feature dimension that QTSA can handle is independent of the number of qubits. (b) Due to quantum noises, it is more challenging to implement a quantum circuit on near-term quantum computers than on quantum simulators or noise-free quantum computers in a distant future. This necessitates the design of a low-depth quantum circuit. (c) The unique potentials of QTSA over classical ML-based TSA are as follows. First, the execution of the quantum circuit refers to computations in the Hilbert space, as presented in (4)–(6), which inherits the exponential quantum speedup in tensor products. Second, QTSA reports comparable accuracy against classical DNN while saving more than 70% parameters, which conforms to the observations in existing studies that QML can be more expressive [41].

IV. CONCLUSION

This paper unlocks the potential of quantum machine learning in power system data-driven analytics. The key innovation is to develop a HELD circuit-centric transient stability assessment to enable an efficient transient stability analysis for resilient and secured decision-making of real-world power systems. The new quantum transient stability analysis algorithm achieves outstanding performances both on a simulator and a real IBM quantum devices (ibmq_boeblingen, a 20-qubit machine, and ibm_lagos, a 7-qubit machine). Therefore, it has the penitential to resolve the TSA challenge in this noisy-intermediate-scale quantum (NISQ) era.

APPENDIX A

TEST SYSTEM PARAMETERS AND TRAINING SETTINGS

This appendix provides the test system parameters and dataset settings.

SMIB system: The dynamical model of the SMIB system is detailed as follows [5]:

\[
\begin{align*}
\dot{\delta} &= \omega_0 (\omega - 1) \\
\dot{\omega} &= \frac{1}{T_j} \left( P_m - P_e - D (\omega - 1) \right)
\end{align*}
\]

(17)

Here, \(\delta\) denotes the power angle (unit: rad) of the generator; \(\omega\) denotes the rotor speed (unit: p.u.) of the generator; \(T_j\) and \(D\) respectively denote the inertia constant and damping coefficient; \(P_m\) is the mechanical power; \(P_e = E'U_0 \sin \delta/x_0\) is the electrical power where \(E'\), \(U_0\) and \(x_0\) respectively denote the generator voltage, the infinite bus voltage and the equivalent impedance. Parameters of the SMIB system are: \(\omega_0 = 120\pi\) rad/s, \(T_j = 6\) s, \(D = 5\) p.u., \(P_m = 1.2\) p.u., \(E' = 1.05\) p.u., \(U_0 = 1\) p.u., \(x_0 = 0.5\) p.u.

Two-area system and NPCC system: All the parameters of the systems are from the Power System Toolbox (PST) [38]. The voltage-behind-transient-reactance model is adopted for all the synchronous generators.

IEEE 300-bus system: Parameters of the network and the operating point are from Matpower [42]. Dynamical models and parameters of the generators are from PST [38]. The classical generator model is adopted with a typical parameter setting (e.g., the d-axis synchronous reactance is 0.081 p.u. and the inertia constant is 30.3s).

Dataset settings: For the SMIB system, we perform time-domain simulations with randomly-generated initial states. For other test systems, we perform time-domain simulations under three-phase short circuit faults. The initial operating points are identical to those from PST/Matpower. The fault location is randomly selected at different transmission lines. The fault clearing time is randomly set between 0.1s-0.35s for the 2-area system, 0.1s-0.55s for the NPCC system, and 0.1s-0.4s for the IEEE 300-bus system.

The obtained data are randomly divided into the training and testing sets.

ACKNOWLEDGMENT

The authors would like to acknowledge the Brookhaven National Laboratory operated IBM-Q Hub. We thank Prof. Tzu-Chieh Wei (C.N. Yang Institute for Theoretical Physics and Department of Physics and Astronomy, Stony Brook University) for the valuable suggestions on the manuscript.

REFERENCES

[1] M. Singh, “‘California and Texas are warnings’: Blackouts show US deeply unprepared for the climate crisis,” Guardian, Feb. 19, 2021.
[2] California ISO, California Public Utilities Commission, and California Energy Commission, “Root cause analysis: Mid-August 2020 extreme heat wave,” Tech. Rep., Jan. 13, 2021.
[3] Department of Energy, “Transforming the nation’s electricity system: The second installment of the quadrennial energy review,” Tech. Rep., Washington DC, USA, 2017.
[4] B. A. Carreras, D. E. Newman, and I. Dobson, “North american blackout time series statistics and implications for blackout risk,” IEEE Trans. Power Syst., vol. 31, no. 6, pp. 4406–4414, Nov. 2016.
[5] P. Kundur, Power System Stability and Control. New York, NY, USA: McGraw-Hill, 1994.

Fig. 12. Impact of noisy quantum environment on QTSA performance. (a): Left subplot shows QTSA’s success probabilities of all samples (sorted from the sample of the lowest success probability to that of the highest) over a range of gate errors. Right subplot shows QTSA’s overall accuracy on the test set over a range of gate errors. (b): QTSA’s success probabilities of all sample as well as the overall accuracy over a range of thermal relaxation time \(T_1\) (i.e., the time it takes for a qubit to decay from the excited state to the ground state).
Yifan Zhou (Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2014 and 2019, respectively. She is currently a Postdoctoral Researcher with Stony Brook University, Stony Brook, NY, USA. Her research interests include microgrid stability and control, formal methods and reachability analysis, and quantum computing.

Peng Zhang (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of British Columbia, Vancouver, BC, Canada, in 2009. He is currently a Full Professor of electrical and computer engineering, and a SUNY Empire Innovation Professor with Stony Brook University, Stony Brook, NY, USA. He has a joint appointment with Brookhaven National Laboratory, as a Staff Scientist. Previously, he was a Centennial Associate Professor and a Francis L. Castleman Associate Professor with Stony Brook University, Stony Brook, NY, USA. He was a System Planning Engineer with BC Hydro and Power Authority, Canada, during 2006–2010. His research interests include AI-enabled smart grids, quantum-enabled power grids, networked microgrids, power system stability and control, cybersecurity, and formal methods and reachability analysis.

Dr. Zhang is an individual member of CIGRE. He is the Editor of the IEEE TRANSACTIONS ON POWER SYSTEMS, IEEE POWER AND ENERGY SOCIETY LETTERS, and IEEE JOURNAL OF OCEANIC ENGINEERING.