IOP Publishing Quantum Sci. Technol. 10 (2025) 045042 https://doi.org/10.1088/2058-9565/ae03e6

Quantum Science and Technology

Check for updates

OPEN ACCESS

RECEIVED
17 April 2025

REVISED
7 August 2025

ACCEPTED FOR PUBLICATION
5 September 2025

PUBLISHED
24 September 2025

Original Content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

PAPER

# Ensemble-based quantum token protocol benchmarked on IBM quantum processors

**Lucas Tsunaki[1], Bernd Bauerhenne[2], Malwin Xibraku[2], Martin E Garcia[2], Kilian Singer[2,*] and Boris Naydenov[1,3,*]**

[1] Department Spins in Energy Conversion and Quantum Information Science (ASPIN), Helmholtz-Zentrum Berlin für Materialien und Energie GmbH, Hahn-Meitner-Platz 1, 14109 Berlin, Germany
[2] Institute of Physics and Center for Interdisciplinary Nanostructure Science and Technology (CINSaT), University of Kassel, Heinrich-Plett-Strasse 40, 34132 Kassel, Germany
[3] Department of Physics, Freie Universität Berlin, 14195 Berlin, Germany
[*] Authors to whom any correspondence should be addressed.

**E-mail:** ks@uni-kassel.de and boris.naydenov@helmholtz-berlin.de

**Keywords:** quantum tokens, color centers, superconducting qubits, quantum cryptography, quantum memories

## Abstract

Quantum tokens envision to store unclonable quantum states in a physical device, with the goal of being used for personal authentication protocols, as required by banks. Still, the experimental realization of such devices faces many technical challenges, which can be partially mitigated using ensembles instead of single qubits. In this work, we thus propose an ensemble-based quantum token protocol, describing it through a simple yet general model based on a quantum mechanical observable. The protocol is benchmarked on five IBM quantum processors and a general hacker attack scenario is analyzed, in which the attacker attempts to read the bank token and forge a fake one, based on the information gained from this measurement. We experimentally demonstrate that the probability that the bank erroneously accepts a forged coin composed of multiple tokens can reach values below $10^{-22}$, while the probability that the bank accepts its own coin is above 0.999. The overall security of the protocol is therefore demonstrated within a hardware-agnostic framework, confirming the practical viability of the protocol in arbitrary quantum systems and thus paving the way for future applications with different ensembles of qubits, such as color center defects in solids.

## 1. Introduction

Quantum no-cloning theorem [1] is the basis for secure exchange of information in quantum communications and cryptography [2]. Another interesting application is the creation of mobile non-copyable storage units with a natural expiration date, i.e. a quantum token [3, 4]. The proposed quantum token protocol is schematically presented in figure 1. First, an issuing agent (e.g. a bank) generates many tokens, each in a different quantum state with angles $\theta_b^i$ and $\phi_b^i$ on the Bloch sphere. All quantum tokens compose together a physical device, which we denominate as a quantum coin. The device is then physically handed to the user, who stores it until authentication is required. To perform the authentication, the bank measures all tokens in the coin and compares them with the angles as they are originally prepared in each token. If a minimum number of tokens in the coin are measured in the correct state above a predetermined security threshold, the coin is successfully authenticated. Otherwise, a cloning attempt is heralded. In such a way, the quantum information is stored, transported and later authenticated rather than simply transmitted, as in quantum communication applications [5, 6]. If successful, such a technology would represent a robust system for the highest security requirements, such as bank cards or personal IDs.

Promising candidates for quantum systems with long coherence, room temperature operation, low power consumption, potential for miniaturization, ease of state manipulation and readout are color centers in transparent solid state substrates [7], such as the nitrogen-vacancy (NV) center in diamond [8]. When considering these systems, the use of ensembles as quantum tokens rather than single qubits could make this

© 2025 The Author(s). Published by IOP Publishing Ltd

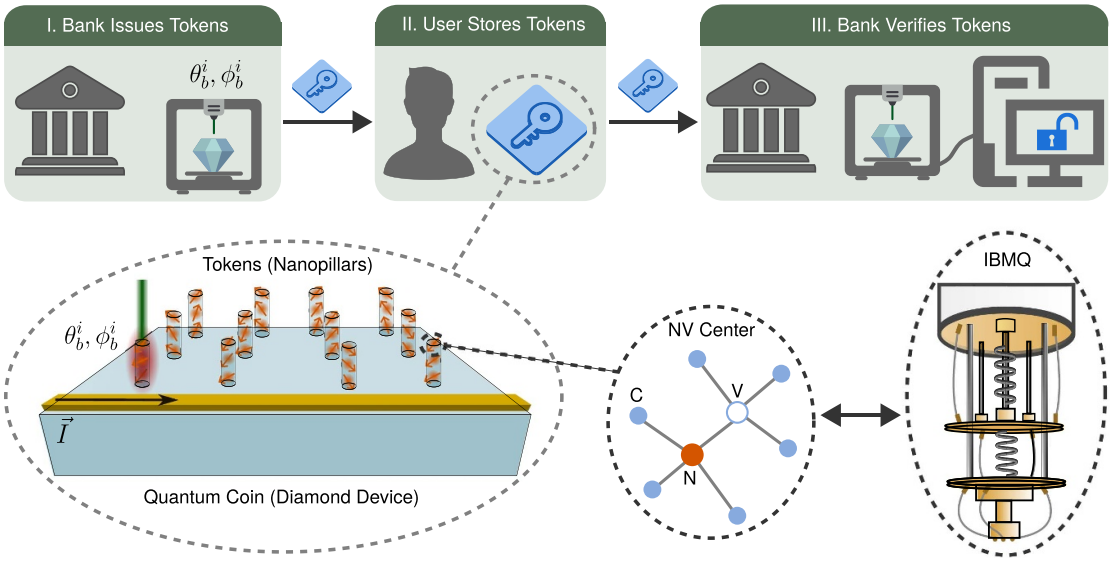


**Figure 1.** Proposed ensemble-based quantum token protocol and the diamond quantum coin device. An issuing agent (e.g. a bank) generates many quantum tokens with angles $\theta_b^i$ and $\phi_b^i$ in a quantum coin device. The device is then physically handed to the user, who stores it until authentication by the bank is required, which consists in measuring the tokens in the coin and comparing them with the previously generated angles. The coin is accepted if a minimum number of tokens in the coin above a certain security threshold is measured in the right state. The physical implementation of the tokens is envisioned to be achieved with diamond nanopillars (for optimal photon collection efficiency [15, 16]) containing nitrogen-vacancy center ensembles. Whereby adopting tokens composed of ensembles instead of single qubits, the application can be technology less demanding. In a combination of optical methods [13] and resonant electromagnetic pulses [14], the electronic spins of the NVs could be used as auxiliary qubits for measuring [22] and controlling [18] the nitrogen nuclear spins, used as memory qubits. Nonetheless, such physical implementation of a diamond quantum coin faces many technical challenges regarding material fabrication [23]. With that, a benchmark of the protocol within a hardware-agnostic framework is vital for future quantum token applications, which is accomplished with IBM quantum platforms. IBM quantum processor illustration adapted from 'Quantum Computer' (349208) by Aleksandar Radovanovic via Open Clipart.

application technologically less demanding [9]. Whereby entering a redundant quantum parallelism regime with many qubits performing the same protocol, the partial decoherence of one or more qubits of the ensemble will not result in a complete failure of the quantum token protocol. However, the conventional single qubit quantum token protocol [3] does not allow to use ensembles. As the quantum no-cloning theorem is not trivially applicable for ensembles, which already consist of identical copies of the individual qubits. In this work, we then demonstrate that an ensemble-based quantum token protocol is still safe and reproducible.

Additional challenges appear when dealing with ensembles rather than single qubits, which can be mitigated or corrected with the appropriate technique. For instance, many optimal quantum control methods, such as composite pulses [10, 11], are being developed to allow for optimal ensemble manipulation, taking inhomogeneous broadening due to slightly different qubit energies and control field amplitude into account. Another important factor is the ensemble separability, that is, the ability of an attacker to manipulate and measure individual qubits of the ensemble. In this case, by making measurements on different axes, an attacker can more efficiently read and create forged tokens, as discussed in detail in our companion paper [12]. For NVs and other color centers in contrast, the ensemble is typically already not separable, due to the diffraction limited optical polarization and readout [13] and to non-local microwave excitation [14] techniques common to them. Nonetheless, our method is still applicable if measurements on sub-ensembles are performed.

A physical candidate for a quantum coin device can be imagined as series of diamond nanopillars containing small ensembles of NVs (see figure 1). Each nanopillar then represents a quantum token, providing improved photon collection efficiency as compared to bulk diamond [15, 16]. The states of the tokens are optically initialized [13] by the bank and prepared in the electronic spin of the NVs into the states $\theta_b^i$ and $\phi_b^i$ with resonant microwave pulses [14], transmitted through an antenna micro-fabricated on the diamond surface [17]. The states are then transferred to nitrogen nuclear spins of the NVs or to nearby $^{13}$C with longer coherence times through a SWAP gate [18], which could potentially be implemented by dynamical decoupling of the electron spin [19]. The latter can be extended during storage, by also applying dynamical decoupling sequences to the nuclear spins for error mitigation [20]. In addition, the NV charge

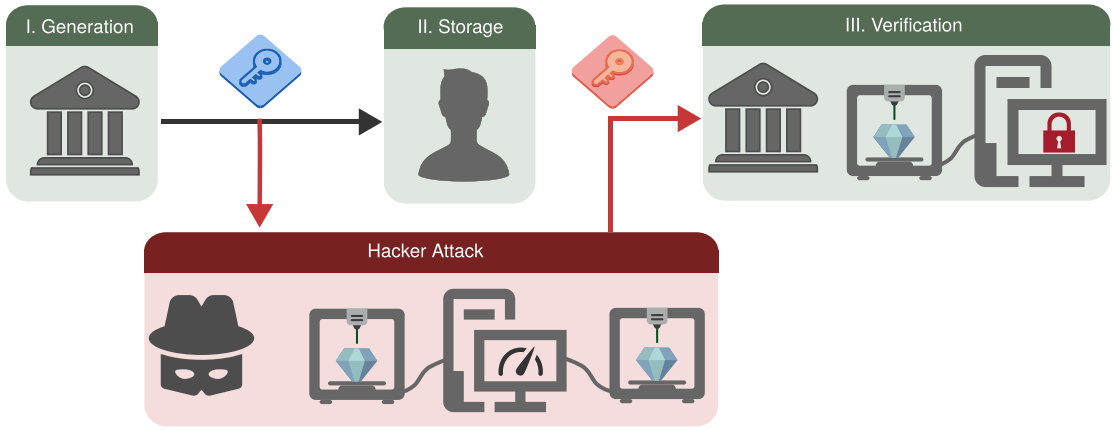


**Figure 2.** General hacker attack scenario considered in this work. Firstly, the attacker attempts to gain information about the quantum coin prepared by the bank by measuring its tokens. With this information, they forge fake tokens, which they use to attempt to impersonate the user. The bank needs to be able to discern the fake tokens from the ones which they originally prepared, where the disparity between the efficiency in which the bank generates its own tokens and the efficiency in which the attacker forges fake ones defines the security of the quantum token protocol.

can be changed during storage, leading to a spinless electronic state which further extends the nuclear coherence [21]. Finally, in order to verify the angles $\theta_{\mathrm{b}}^{i}$ and $\phi_{\mathrm{b}}^{i}$ of each token in the coin, the bank performs an optical single-shot readout of the nuclear spins using the electron spins as an ancilla qubit [22].

The experimental realization of such a diamond-based quantum coin is still an ongoing research. More specifically, it is currently technologically limited by fabrication techniques of such a device [23] and the aforementioned control and single shot readout methods for NV ensembles. Therefore, it is crucial to benchmark the ensemble token protocol on a hardware-agnostic high-level representation quantum processor, such as the IBM quantum platforms (IBMQ). This approach allows us to not only identify the quality parameters of the quantum token hardware, but also to quantify the efficiency of the proposed protocol, demonstrating its potential hardware interoperability. And lastly, by comparing the efficiency of the bank in preparing and authenticating its own tokens with the efficiency in which a hacker can clone the quantum token, we obtain a measure of the protocol's safety.

In this study, we thus consider one general possible attack scenario, as represented in figure 2. First, the hacker attempts to gain information about the quantum coin prepared by the bank by performing projective measurements to the tokens. Using this information, the attacker forges a fake token and passes it to the bank, seeking to impersonate the user. During authentication, the bank then needs to be able to discern the forged coin from the one it originally created. Although this scenario covers a general framework for many quantum token hardware candidates, there could potentially exist other more efficient attack methods [24, 25], not covered in this study.

For benchmarking the proposed protocol and quantifying its security against such hacker attacks, we used five different IBMQ superconducting processors [26–29]. Due to the increased memory costs in waveform generation for reproducing an ensemble of qubits with IBMQ, we invoke the ergodic principle, in which the ensemble average is substituted by a time-average of the same qubit prepared and measured multiple times. More experimental details are given in appendix B, while the proofs of mathematical equations are all presented in appendix A. Due to the increased number of mathematical variables in this work, table 2 provides a glossary of their definitions. All codes for experimental control and result analysis are open-source and provided at [30], which also contains a graphical user interface (GUI) where the results presented in this work can be calculated and visualized for the parameters of an arbitrary quantum token hardware.

This work is structured as follows. In section 2, we start by presenting a model for the measurement uncertainty which a general quantum token hardware is subject to, being then benchmarked in the IBMQ. The highlight of this approach is that it enables us to determine the main quality parameters that will describe most of the tokens' behavior. Building on this model, in section 3, the bank generation and authentication protocol is defined and benchmarked, which defines how well can the bank prepare and read its own tokens. Finally, in section 4, we execute the attack method and conduct a detailed analysis of the resulting protocol's security.

## 2. Uncertainty analysis

A primary obstacle in modeling a quantum processor as the IBMQ is the lack of complete access to its specifications, due to proprietary constraints. A crucial parameter that can be leveraged to extract valuable information about a physical system is its inherent uncertainty or noise, characterized by its statistical properties [31]. In addition to that, a noise based framework is common to other quantum system, respecting the protocol's universality. As will be shown, this uncertainty is never zero even without any experimental errors, due to the quantum uncertainty principle

For the description of the ensemble system we use a simple Hamiltonian-independent model that describes a general quantum state on the Bloch-sphere by the two angles $\theta$ and $\phi$, additionally to expressions for expectation values and variance of the ensemble measurements. Rigorously speaking, the state of the ensemble is in fact in a statistical mixture, described by a density matrix instead of a pure ket state. However, to simplify the theoretical treatment, we assume pure states for the quantum token ensemble—an approximation which holds notably well with the IBMQ, as will become clear in the following sections. In most physical systems, the angles $\theta$ and $\phi$ are not directly measured. Instead, the bank has an observable $\hat{N}$ given by

$$\hat{N} = \begin{bmatrix} N_0 & 0 \\ 0 & N_1 \end{bmatrix},$$

which can be used to measure the populations in the $|0\rangle$ and $|1\rangle$ states, where $N_0, N_1 \in \mathbb{R}$. This observable could be related to some elaborate quantity like the change of magnetic flux or charge in a superconducting circuit, as with IBMQ. But for a more concrete physical example, we take this observable related to a photon count, as is the case with NVs. For a general Bloch state with angles $\theta$ and $\phi$, the expectation value of $\hat{N}$ is

$$\langle \hat{N} \rangle = N_0 \cos^2\left(\frac{\theta}{2}\right) + N_1 \sin^2\left(\frac{\theta}{2}\right), \tag{1}$$

not depending on $\phi$. Ideally, the bank would have one of the eigenvalues as 0. In reality however, if we assume $|1\rangle$ to be the bright state such that $N_1 > N_0$, $N_0$ is related to the background counts and $N_1$ to the collection efficiency of the measurement. Opposite to the IBMQ, in NVs we have $N_1 < N_0$, which does not affect the end results of the quantum token. As it will become clear in sections 3 and 4, the quality of a quantum token can be characterized in terms of the normalized contrast between $N_1$ and $N_0$, $c \equiv (N_1 - N_0)/(N_0 + N_1)$. The universality of the protocol extends to systems which support a measurement scheme that allows for the measurement of an ensemble average $\hat{N}$, a set of universal rotations $\hat{R}(\theta, \phi)$ and can be approximately described by pure states.

A general assumption can be made that the total uncertainty $\sigma_N$ for a measurement of $\hat{N}$ is composed of three factors [32]. Firstly, any quantum system is subject to a quantum projection noise, or Heisenberg uncertainty, given by the operator variance $\sigma_q^2 = \langle \hat{N}^2 \rangle - \langle \hat{N} \rangle^2$. This value is highest when the states are in a maximum superposition with $\theta = \pi/2$ and zero when they are in the eigenbasis states, with $\theta = 0$ or $\theta = \pi$. Secondly, given the quantized nature of the emitted photons, following a Poisson statistics, the total uncertainty also has a contribution from shot noise as $\sigma_s^2 = \langle \hat{N} \rangle$. Contrarily to the quantum projection noise, the shot noise is maximum at $\theta = \pi$. Finally, we also consider a generic Gaussian experimental error given by a constant $\sigma_{\text{exp}}$, adding an offset to the uncertainty. This error can be defined to incorporate pulses error in the state preparation and other experimental noise sources. Assuming that these three components are statistically independent, they can be added in quadrature [33], resulting in

$$\begin{aligned} \sigma_N^2 &= \sigma_q^2 + \sigma_n^2 + \sigma_{\text{exp}}^2 \\ &= N_0 \cos^2\left(\frac{\theta}{2}\right)(1 + N_0) + N_1 \sin^2\left(\frac{\theta}{2}\right)(1 + N_1) \\ &\quad - \left[N_0 \cos^2\left(\frac{\theta}{2}\right) + N_1 \sin^2\left(\frac{\theta}{2}\right)\right]^2 + \sigma_{\text{exp}}^2. \end{aligned} \tag{2}$$

Depending on the values of $N_0$, $N_1$ and $\sigma_{\text{exp}}$, the total noise will be dominated by either one of the three components. Solutions of $\sigma_N$ for different combinations of values are shown in figure 8 in appendix A. We also provide a GUI application [30], which can be used to visualize $\sigma_N$ for arbitrary parameters input by the user.

To test this noise model, a Rabi measurement [14, 34] for $\langle \hat{N} \rangle$ was performed on the IBMQ as a function of $\theta$ for $\phi = 0$, while $\sigma_N$ is taken from the standard deviation of 100 shots. To better compare the data between the different quantum processors, both values are normalized by $N_0 + N_1$. The high-level representation of

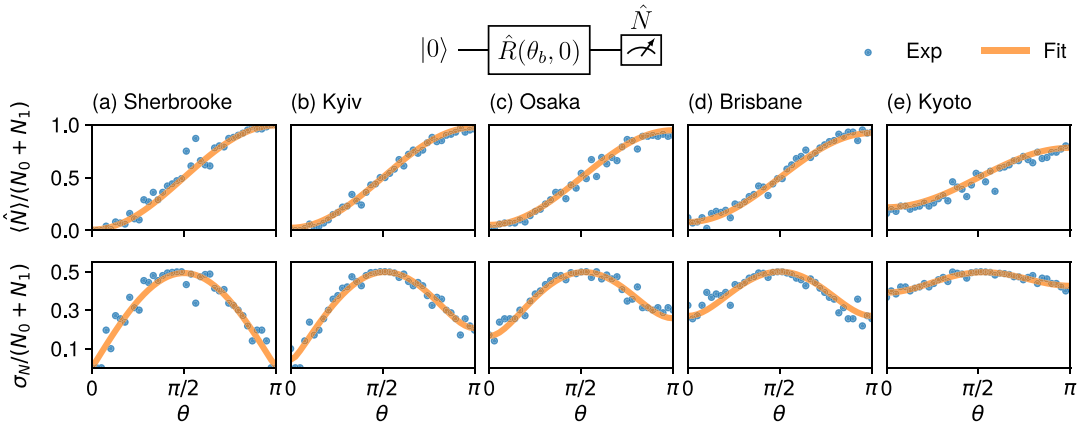


**Figure 3.** Normalized expectation values $\langle\hat{N}\rangle/(N_0+N_1)$ and uncertainty $\sigma_N/(N_0+N_1)$ as a function of $\theta$ measured with IBM quantum platforms (a) Sherbrooke, (b) Kyiv, (c) Osaka, (d) Brisbane and (e) Kyoto. $\langle\hat{N}\rangle$ values follow a Rabi oscillation from the state $|0\rangle$ to $|1\rangle$, while the uncertainties are maximum at $\theta=\pi/2$ due to quantum projection noise. Kyiv and Osaka are strongly influenced by shot noise, given that the uncertainty is higher at $\theta=\pi$ than at $\theta=0$. On the other hand, Brisbane and Kyoto are dominated by experimental noise, resulting in a constant offset to the the uncertainty. Last, Sherbrooke presents the highest contrast, with almost no contribution from shot and experimental noise. The fitted values of normalized contrast $c$ and $\sigma_{\text{exp}}$ are given in table 1.

**Table 1.** Quality parameters of the token hardware. The values of normalized contrast $c$ and experimental uncertainty $\sigma_{\text{exp}}$ are obtained from the fit of $\langle\hat{N}\rangle$ and $\sigma_N$ (section 2), with the first representing how easy it is to determine the state of the system and the second is defined to incorporate all additional experimental errors. The mean experimental bank self-acceptance $\bar{n}_{\text{b}}$ represents the fraction of qubits in a token which the bank prepares and measures in the correct state (section 3). Analogously, $\bar{n}_{\text{f}}$ is the fraction of qubits which the bank accepts from a forged token (section 4). Finally, $p_{\text{f}}$ is the probability of acceptance of one forged token if we set the acceptance threshold $n_{\text{T}}$ such that the bank acceptance probability is $p_{\text{b}}>0.999$. Overall, Sherbrooke and Kyiv have the best quality parameters, such that a minor improvement of $c$ in comparison to Brisbane leads to a significant increase in the token security. Forgery measurements could not be concluded using Osaka and Kyoto due to the systems' retirement in August 2024.

| IBMQ | $c$ | $\sigma_{\text{exp}}/_{N_0+N_1}$ | $\bar{n}_{\text{b}}$ | $\bar{n}_{\text{f}}$ | $p_{\text{f}}$ |
|---|---|---|---|---|---|
| Sherbrooke | 0.986 | $10^{-5}$ | 0.990 | 0.685 | 0.065 |
| Kyiv | 0.950 | 0.026 | 0.992 | 0.682 | 0.059 |
| Osaka | 0.896 | 0.158 | 0.970 | — | — |
| Brisbane | 0.843 | 0.270 | 0.879 | 0.611 | 0.285 |
| Kyoto | 0.563 | 0.377 | 0.792 | — | — |

the quantum circuit of the experiment and the results are shown in figure 3. By fitting $\langle\hat{N}\rangle$ with equation (1) and $\sigma_N$ with equation (2), we obtain the values of the normalized contrast $c$ and $\sigma_{\text{exp}}$ (table 1).

All five quantum processing units show a maximum of $\sigma_N$ at $\theta=\pi/2$, resulting from a strong contribution from quantum projection noise to the total uncertainty. In Kyiv and Kyoto, a large contribution from shot noise is also observed, given by a higher uncertainty at $\theta=\pi$ over $\theta=0$. While in Brisbane and Kyoto the experimental error is the dominant term, adding a constant offset to the uncertainty. Finally, Sherbrooke has the highest contrast and is completely dominated by quantum projection noise, with negligible contribution from shot noise and experimental error. Which can be explained in terms of the hardware's slightly lower median readout and gate errors, apart from longer coherence times.

This prominent quantum projection noise can be used as an extra security measure for the bank. As an increased noise would indicate that the token is being measured in the wrong axis, which in turn would herald a cloning attempt by an attacker. However, in cases with inefficient read out—as typically is the case with NV color centers—shot noise can dominate quantum projection noise. In those cases, our protocol is still applicable, but more quantum tokens per coin are needed [12].

## 3. The bank protocol

The next important quantity to be benchmarked is the fraction of qubits which the bank prepares and reads successfully, without interference of an attacker. We assume a protocol as depicted in figure 4. Starting from the $|0\rangle$ state, the bank prepares a state with a rotation $\hat{R}(\theta_{\text{b}},\phi_{\text{b}})$. Subsequently, for authentication, the bank unrotates the token with the inverse operation $\hat{R}^{-1}(\theta_{\text{b}},\phi_{\text{b}})$ and makes a measurement on $\hat{N}$, yielding[4]

[4] Or $n_{\text{b}}=\langle\hat{N}\rangle/(N_0+N_1)$, if we had chosen $|1\rangle$ as the initial reference state or if $N_0>N_1$.

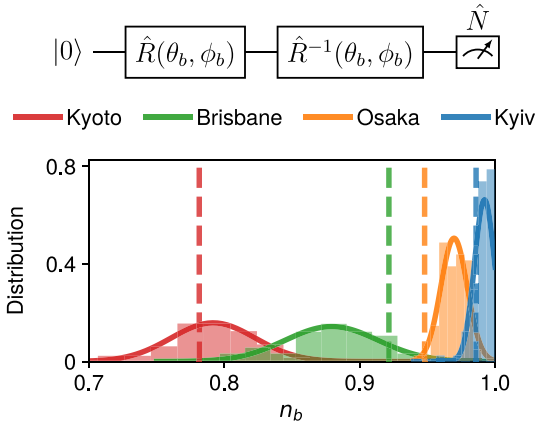


**Figure 4.** Self acceptance fraction $n_\mathrm{b}$ distributions of IBM quantum platforms Kyiv, Osaka, Brisbane and Kyoto measured for several combinations of linearly distributed angles $\theta_\mathrm{b}$ and $\phi_\mathrm{b}$. $n_\mathrm{b}$ is defined as the number of qubits in a token which are measured in the state $|0\rangle$ after a bank rotation of $\hat{R}(\theta_\mathrm{b}, \phi_\mathrm{b})$, followed by an inverse rotation and measurement of $\hat{N}$. The distributions are fitted with Gaussian functions, determining the acceptance range of each hardware, with a mean $\bar{n}_\mathrm{b}$ roughly around $N_1/(N_0 + N_1)$ (dashed lines). As the contrast of the hardware increases, the mean values of $n_\mathrm{b}$ get higher and the linewidth of the distributions get narrower. Sherbrooke results are omitted, as they overlap with Kyiv results.

$n_\mathrm{b} = 1 - \langle\hat{N}\rangle/(N_0 + N_1)$ qubits of the token at the initial state $|0\rangle$, where both the counts and the uncertainty are smallest.

This acceptance fraction is highly dependent on the specific pulse errors of the quantum hardware. Still, some general conclusions can be drawn just from the eigenvalues of the observable $\hat{N}$. If no rotation errors were made and there was no decoherence, the two rotations would cancel and the final state would be $|\Psi_\mathrm{f}\rangle = \hat{R}^{-1}(\theta_\mathrm{b}, \phi_\mathrm{b})\hat{R}(\theta_\mathrm{b}, \phi_\mathrm{b})|0\rangle = |0\rangle$. Although this is not the experimental reality, it gives a rough estimate on the average self-acceptance fraction of

$$\bar{n}_\mathrm{b} \approx 1 - \frac{\langle 0|\hat{N}|0\rangle}{N_0 + N_1} = \frac{N_1}{N_0 + N_1}. \tag{3}$$

Clearly, the higher the contrast between the dark counts $N_0$ and the bright counts $N_1$, the easier it is for the bank to distinguish the state. This also sets a limit for the self-acceptance fraction. Therefore, the bank should choose a threshold for the photon count $n_\mathrm{T}$ such that only tokens which show $n_\mathrm{b} > n_\mathrm{T}$ are accepted in the coin.

The acceptance fraction $n_\mathrm{b}$ was benchmarked in the five IBMQ for different combination of linearly distributed angles $\phi_\mathrm{b}$ and $\theta_\mathrm{b}$. In figure 4, the $n_\mathrm{b}$ distributions of each hardware are shown and fitted with Gaussian functions, defining the acceptance range of each token. For clearer visualization, the results from Sherbrooke are omitted in the figure, as they completely overlap with Kyiv, but can be found online [30]. The distributions present good agreement with their values of $N_1/(N_0 + N_1)$ fitted from $\langle\hat{N}\rangle$ and $\sigma_N$ (figure 2 and table 1). As the normalized contrast $c = (N_1 - N_0)/(N_0 + N_1)$ of the hardware increases, not only the mean values $\bar{n}_\mathrm{b}$ get higher, but the linewidth of the distributions get smaller. With Kyiv and Sherbrooke showing the highest self-acceptance fractions, while Kyoto has the worst, in accordance to the results from section 2. The $n_\mathrm{b}$ values as a function of $\theta_\mathrm{b}$ and $\phi_\mathrm{b}$ are shown in figure 10, where no angle dependence is observed. In hardware scenarios with non-optimized long pulses and short coherence times, we would expect smaller acceptance fractions for larger angles due to increased errors in longer pulses. Which is not the case in these IBMQ, given the high fidelity of their qubit operations.

## 4. An attack scenario

The last aspect which defines the quality of a quantum token hardware is how much better the bank can prepare and accepts its own tokens compared to tokens forged by an attacker. We imagine one general attack scenario where the attacker makes a measurement of the bank token, then uses the obtained information to forge a fake one and pass it to the bank. The attack scenario is schematically represented in figure 1.

As discussed in the previous section, the bank prepares the token with a rotation $\hat{R}(\theta_\mathrm{b}, \phi_\mathrm{b})$. In succession, the attacker unrotates with some chosen angles $\theta_\mathrm{a}$ and $\phi_\mathrm{a}$, leading to a final state as $|\Psi_\mathrm{f}\rangle = \hat{R}^{-1}(\theta_\mathrm{a}, \phi_\mathrm{a})\hat{R}(\theta_\mathrm{b}, \phi_\mathrm{b})|0\rangle$. Lastly, the attacker measures the observable $\hat{N}_\mathrm{a}$, which in practice is not necessarily the same observable the bank would use $\hat{N}_\mathrm{a} \neq \hat{N}$, as they do not have the same experimental setup to measure the token. However, for the sake of simplicity we consider the case where the faker can make their

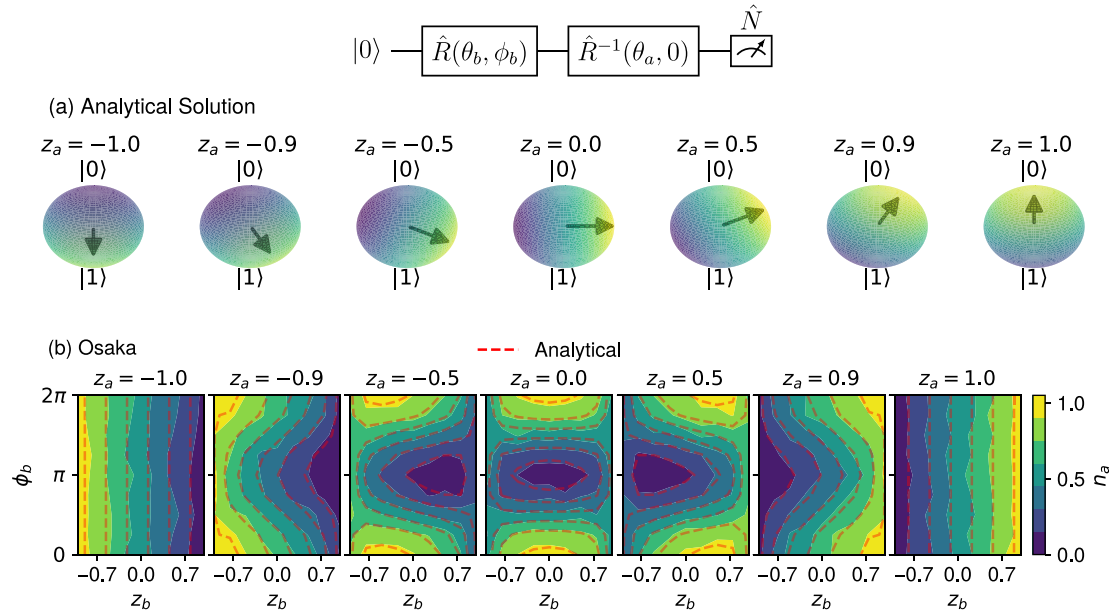


**Figure 5.** (a) Analytical solution and (b) experimental data from IBM quantum platform Osaka for the fraction of qubits in a token which are measured in the state $|0\rangle$ by an attacker $n_a$ as a function of the bank preparation angles $z_b = \cos\theta_b$, $\phi_b$ and the attacker measurement angle $z_a = \cos\theta_a$ for $\phi_a = 0$. The analytical solution for (a) is taken from equation (4) with $c = 1$, while the red-dashed lines in (b) show the solution for $c$ from table 1. The strong agreement of the analytical model with the experimental data evidences that the hardware can be well described only in terms of the operator $\hat{N}$ and its uncertainty. Clearly, the closer the measurement angles are to the bank angles, the higher $n_a$ will be. By making a measurement attempt, the attacker is able to conclude that the bank is in a line of the Bloch sphere solution of the parametric equation resulting from equation (4) with fixed $n_a$. The attacker then uses this knowledge to forge a fake token and pass it to the bank.

measurements as accurate as the bank, with $\hat{N}_a = \hat{N}$, which is the case in the IBMQ. Thus, analogously to equation (3), the fraction of qubits $n_a$ in the token which are measured in the state $|0\rangle$ by the attacker is

$$2n_a = 1 + c\left[\cos\theta_a \cos\theta_b + \sin\theta_a \sin\theta_b \cos\left(\phi_b - \phi_a\right)\right]. \tag{4}$$

Integrating and averaging over the bank angles $\theta_b$ and $\phi_b$ (appendix A), we obtain $\bar{n}_a = 1/2$, which does not depend on $c$. In this way, $n_a$ is centered around 1/2 and $c$ corresponds to a normalized contrast, as expected. The analytical solution of $n_a$ for $c = 1$ is plotted over the Bloch sphere in figure 5(a), where the arrows represent the measurement vectors of the attacker with $z_a = \cos\theta_a$. As observed, the closer the attacker measurement angles $\theta_a$ and $\phi_a$ are to the bank token $\theta_b$ and $\phi_b$, the higher is the value of $n_a$ measured by the attacker. The two pole measurements at $z_a = -1$ and $+1$ are not symmetric, due to the increased shot noise at $|1\rangle$. The GUI application [30] also provides visualization of $n_a$ solutions for arbitrary user defined parameters.

This attacker measurement scheme was performed on the five IBMQs as a function of the angles $z_b = \cos(\theta_b)$, $\phi_b$ and $z_a = \cos(\theta_a)$ for a fixed $\phi_a$. The resulting angle dependence of $n_a$ is shown for IBMQ Osaka with $\phi_a = 0$ in figure 5(b), with the Bloch sphere being projected into a plane for better visualization. The results for the other four quantum processors are shown in figure 11, with measurement angles $\phi_a = 0$ and $\phi_a = \pi/2$. The data is also compared with the analytical expression from equation (4) for the value of $c$ fitted from the uncertainty measurement (table 1). In all cases, the measurements show a strong agreement with the analytical formula, further indicating that the quality of a quantum token can be fully characterized in terms of the $\hat{N}$ operator eigenvalues.

Based on the measured value of $n_a$, the attacker can deduce the possible angles of the bank token $\theta_b$ and $\phi_b$ by solving equation (4). This results in a solution line over the Bloch sphere, as visualized in figure 5. However, due to the ambiguity of the solution, the attacker cannot unequivocally determine which is the valid point in the line with just one measurement. In this study, we consider a case where the attacker uses the information obtained from the measurement and randomly chooses one of the solutions to generate a forged token with angles $\theta_f$ and $\phi_f$, which is already a much more robust attack than randomly guessing a point in the whole Bloch sphere. More efficient methods where the attacker can subdivide the ensemble into parts and measure each of them in different axes are discussed in detail in our companion paper [12], even though this is not a realistic scenario with most color center systems (section 1).

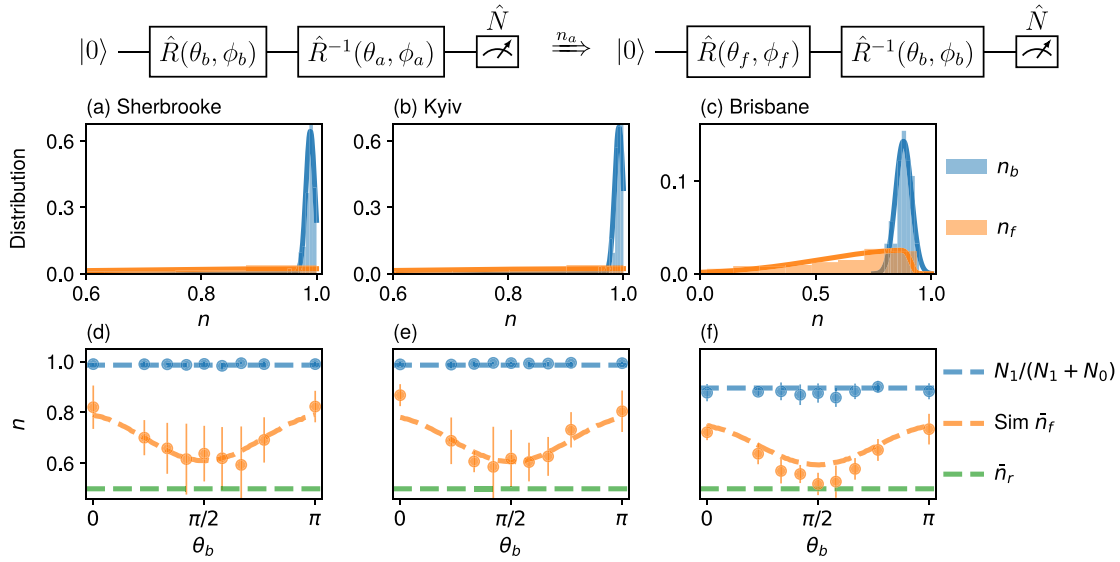


**Figure 6.** Forged tokens distributions $n_f$ compared to the bank distributions $n_b$ measured in IBM quantum platforms (a) Sherbrooke, (b) Kyiv and (c) Brisbane. First, the attacker makes a measurement as described in figure 5. Using the measured values of $n_a$, they try to forge a fake token and pass it to the bank with angles $\theta_f$ and $\phi_f$, which the bank will measure resulting in $n_f$ qubits in the $|0\rangle$ state. The $n_f$ distributions follow a phenomenological skew normal distribution, with a large standard deviation and negative scale. Most of its values are above the threshold for randomly generated tokens $\bar{n}_r = 1/2$ (green dashed line), indicating that the forger procedure is robust, but still worse than the verification procedure of the bank. In Sherbrooke and Kyiv, the forged tokens distributions are completely overshadowed by the narrow bank distribution. (d), (e) and (f) The angle dependence of $n_f$ with $\theta_b$ indicates that the bank tokens are more easily cloneable at the poles rather than the equator, due to a smaller state distribution there. The $n_f$ distributions are compared with their simulated values (figure 7(a)), while the $n_b$ distributions are compared with their expected values of $N_1/(N_1+N_0)$.

If the attacker chooses the measurement angle as $\theta_a = 0$ or $\pi$, then $\sin\theta_a = 0$ and equation (4) can be simply inverted leading to a solution for the angle of the forged token as

$$\theta_f = \arccos\left(\frac{2n_a - 1}{\cos\theta_a c}\right).$$

In this case, $\theta_f$ is completely determined, while $\phi_f$ is undetermined and the attacker must choose an arbitrary value. This can be observed for $z_a = \pm 1$ measurements shown in figure 5(b), where $n_a$ is completely independent of $\phi_b$. It may be the case, however, that $\theta_f$ has no real solution due to experimental errors and thus the attacker simply guesses the angles. If instead, the attacker does not measure the token at the poles, one can solve equation (4) for $\phi_b$, which gives two solutions

$$\phi_f = \phi_a \pm \arccos\left[\frac{(2n_a - 1)/c - \cos\theta_a\cos\theta_f}{\sin\theta_a\sin\theta_f}\right].$$

In order to obtain a valid solution for $\phi_f$, the argument of the inverse cosine function must be located in the interval $[-1, 1]$. This provides a condition for valid $\theta_f$ values as shown in equation (A1). After randomly choosing a $\theta_f$ value inside the interval, the attacker takes one of the two solutions of $\phi_f$. This procedure for forging a fake token is illustrated in figure 9.

The proposed token forgery protocol was implemented for different combinations of attack angles ($\theta_a$, $\phi_a$) and linearly distributed bank angles ($\theta_b$, $\phi_b$), based on the $n_a$ values measured in figure 11. This results in a fraction of $n_f$ qubits in the token which are measured in the $|0\rangle$ state by the bank, where $n_f$ is equivalent to equation (4) substituting the attacker measurement angles by the forged angles $a \to f$. In this way, if an attacker randomly generates fake tokens, they will get $\bar{n}_r = 1/2$ fraction of qubits accepted by the bank, independent of $c$. The experimental distributions of $n_f$ are shown in figure 6 for (a) Sherbrooke, (b) Kyiv and (c) Brisbane[5]. Unlike the bank self-acceptance fraction $n_b$, the forger distributions do not follow a Gaussian function. Instead, we phenomenologically model them by a skew normal distribution [35] with a large standard deviation and negative shape. It is evident that the forger is able to more efficiently generate forged tokens than the average for randomly guessing $\bar{n}_r = 1/2$, thus demonstrating the robustness of the forger

[5] The measurements on Osaka and Kyoto could not be completed due to their retirement in August 2024.

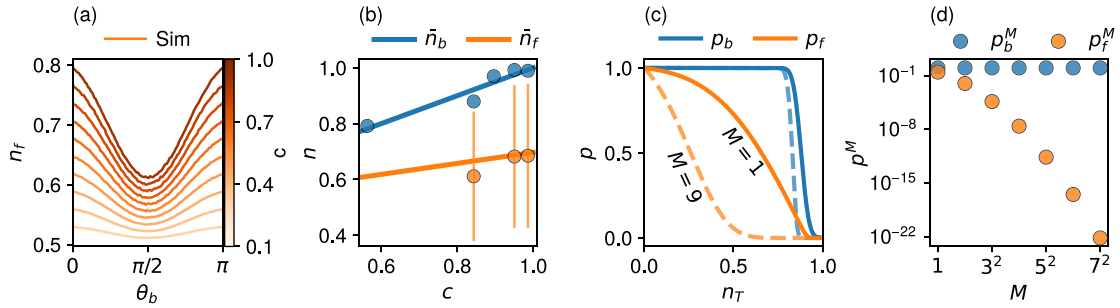


**Figure 7.** (a) Simulated values of $n_f$ for different normalized contrasts $c$. The mean values of $n_f$ increase with $c$, indicating that the token is more easily cloneable. (b) Simulated forger acceptance fractions $\bar{n}_f$ and bank self-acceptance $\bar{n}_b = N_1/(N_1 + N_0)$ as a function of the normalized contrast $c$. Experimental results for the five IBM quantum platforms are also shown, with good agreement with the model. Error bars for $n_b$ values are smaller than the points, while $n_f$ present large standard deviations due to the nature of the forger protocol. As $c$ increases, both values increase linearly, but faster for the bank. Meaning that the protocol becomes safer. (c) Acceptance probability for the forger $p_f$ and the bank $p_b$ calculated from the fitted distributions as a function of the acceptance threshold $n_T$ for 1 and 9 tokens in Brisbane. With a single token, $p_f$ is non-negligible for $n_T$ around 0.8, but with the increase in the number of tokens the chance of acceptance of a fake token approaches zero. (d) Acceptance probabilities $p_f^M$ and $p_b^M$ of IBMQ Brisbane as a function of the number of tokens in the coin $M$. The acceptance threshold is chosen such that $p_b^M > 0.999$. This shows that $p_f^M$ can be made negligible with just a few tokens.

protocol. Still, the distribution from the forged tokens are below the bank self-acceptance distributions. Where in Sherbrooke and Kyiv, the behavior of $n_f$ is completely dominated by the narrow $n_b$ distributions.

By examining the values of $n_f$ as a function of the bank polar angle $\theta_b$ (figures 6(d)–(f)), we observe that the attacker has a larger chance of success if the bank tokens are closer to the poles rather than at the equator of the Bloch sphere. This is a direct result from the probability distribution of the angle $\theta$ in spherical coordinates system $f_\theta(\theta) = \sin(\theta)/2$, being larger at $\theta = \pi/2$. Meaning that there is a higher density of states on the equator than at the poles for the attacker to guess from. Thus, the bank should exploit this result and prepare more tokens at the equator, differently from these measurements. The experimental values of $n_f$ are also in good agreement with simulation, considering the contrast values $c$ obtained from the uncertainties fit (section 2). Small discrepancies between the simulated and experimental values can be attributed to contrast parameter estimation errors.

Additional simulations for different contrast values $c$ are shown in figure 7(a), indicating an increase of $n_f$ with $c$. In figure 7(b), the simulated mean values of $\bar{n}_f$ are compared with the experimental data from IBMQ, showing a linear dependence with $c$. The $n_f$ distributions have intrinsically large standard deviations, due to the nature of the forger protocol. In addition, the experimental values of $n_b$ (section 3) are compared with its theoretical values from equation (3). The results show that the tokens become more easily clonable as the contrast of the hardware increases, given that the attacker can better rely on their measurement of the bank token. On the other hand, $\bar{n}_b \approx N_1/(N_1 + N_0)$ is also linearly dependent on $c$, but with a larger linear coefficient. Which in the end, makes the token safer.

The danger of the bank token being successfully cloned is related to the intersection of the $n_b$ and $n_f$ distributions in figures 6(a)–(c). Therefore, the bank should carefully choose the acceptance threshold $n_T$ such that most of its own tokens are accepted, while forged tokens are rejected. We define the probability of acceptance $p_b$ of the bank tokens and $p_f$ of the forged tokens as the integral of the fitted distributions above the acceptance threshold $n_T$. These acceptance probabilities can be visualized and calculated for arbitrary user defined parameters in our GUI application as well [30].

The bank has the freedom to adjust the acceptance threshold $n_T$ such that its own acceptance probability $p_b$ is set to desired values, while the forger acceptance probability is minimum. Figure 7(c) shows both acceptance probabilities as a function of $n_T$ in IBMQ Brisbane, the hardware with the lowest contrast. If the bank chooses a value of $p_b > 0.999$, this results in $p_f$ values for the three IBMQ which are shown in table 1. It can be seen that a small improvement of less than 15% in contrast from Brisbane to Kyiv leads to a reduction of $p_f$ by a factor of almost 5. This significant increase in security with just a small improvement in the token quality is caused by the highly non-linear behavior of the acceptance probability, being proportional to the overlap area between the two $n_f$ and $n_b$ distributions. Therefore, as the difference between the two distributions increase linearly with $c$ (figure 7(b)) and the linewidth of $n_b$ gets smaller, the resulting token security rapidly increases.

Another important quantity for the security of the protocol is the number of tokens in the coin $M$. Likewise, it should also be optimized by the bank to increase security, as the acceptance probability of the entire coin becomes $p^M$. In figure 7(c), we observe that the acceptance probability of a forged token is non-negligible for $M = 1$. However, if the number of tokens in the coin is increased to $M = 9$, $p_{\mathrm{f}}^9$ is greatly reduced, while the probability of acceptance for a bank $p_{\mathrm{b}}^9$ coin is less affected. To further demonstrate this, figure 7(d) shows the acceptance probability $p_{\mathrm{f}}^M$ of IBMQ Brisbane for different values of $M$, where $n_{\mathrm{T}}$ is chosen such that the acceptance probability of the bank tokens are always above $p_{\mathrm{b}}^M > 0.999$. $M$ values are taken imagining a square lattice of tokens in the coin device. This shows that with a reduced number of tokens smaller than $M = 7^2 = 49$, it is already possible to obtain acceptance probabilities of forged tokens below $10^{-22}$, while keeping high acceptances for the bank verification. Therefore, simply by increasing the number of tokens in a coin, the security of the protocol can be enhanced to desired levels, even with hardware of poor performance.

## 5. Conclusion and outlook

Overall, this study proposes and realizes an ensemble-based quantum token protocol by benchmarking it on the IBM superconducting processors. The experimental noise characterization of the quantum hardware permits to calculate the eigenvalues and uncertainties of the observable $\hat{N}$. Where the large demonstrated quantum projection noise can be used as an extra protection layer, heralding a cloning effort if the token is measured in the wrong axis. In turn, the contrast between the dark and bright states $c$ almost entirely defines the quality of the quantum coin, as observed in the bank self-acceptance and forger measurements. The model's independence from the exact Hamiltonian of the system and its other intrinsic properties demonstrates the hardware-agnosticism of the protocol. Furthermore, the high agreement between theory and the experiments shows that the model is complete, but still general to a large class of quantum systems that support a measurement observable operator $\hat{N}$ and a set of universal rotations $\hat{R}(\theta, \phi)$.

Our protocol also shows great potential due to the ongoing quantum hardware evolution, as a minor improvement in the hardware quality represented by the contrast $c$ leads to significant improvements in security. But even for hardware with low performance, it was shown that the security can be increased to arbitrarily high standards simply by increasing the number of tokens in the coin, within realistic values. Additional security can be obtained by preparing more states in the equator than at the poles of the Bloch sphere, due to the increased proficiency for an attacker to forge tokens prepared at the poles. We further provide an open-source tool with GUI [30], where parameters for an arbitrary hardware configuration can be specified and the resulting token security following our protocol can be estimated.

While the IBMQ are arguably the most advanced quantum processing units available for public use, these superconducting architectures [27–29] are currently not optimal candidates for such a quantum coin device. They operate at low temperatures, making them hardly compact and mobile. Our results thus pave the way for the application of the protocol with other systems, such as NV centers, even in situations when single shot readout protocols are not feasible. Still, certain technological steps need to be overcome in the fabrication of such quantum coins [23]. Also the optimization of control techniques and error mitigation [36] with ensembles should be further improved to extend the coherence times to application realistic values of more than a few seconds.

The results presented here demonstrate the security of the ensemble-based quantum token protocol under one possible attack scenario, where the hacker performs projective measurements to the bank tokens and attempts to forge a fake coin. Although this attack scenario can be considered a general framework for quantum systems, the demonstrated safety does not imply that the proposed protocol is resilient to all other imaginable attack scenarios. This motivates further research into the field, in order that this promising proposal can meet high standards of securities for personal authentication technologies. An additional hacking scenario, not considered in section 4, would be for the attacker to introduce an external entangled qubit to the system and perform a quantum state transfer with the coin tokens, i.e. a SWAP gate [37]. In this way, quantum no-cloning theorem is not violated, as the state is not cloned but stolen, or in quantum mechanical terms, it is swapped. A concrete example of this is if the attacker uses a scanning tip with an NV [38] which can be positioned close to the coin tokens. Such that a hyperfine interaction between the scanning NV and token NVs is established [39], generating the means for a SWAP operation. Nonetheless, this technique presents many technical challenges, with a possible fidelity much lower than the bank self-acceptance probability. Additionally, this would herald the cloning once the bank token is read, given by a larger quantum projection noise.

Further questions regarding the security of quantum token devices are outlined in this work. Namely, the token ensemble is assumed inseparable, which is trivially achieved in color centers, but may not be the case in other candidate systems. Otherwise, if the attacker could make measurements in more than one axis, they would be able to more precisely forge tokens [12]. This does not invalidate the token protocol, but would merely require more tokens in the coin device. Another point which can degrade the protocol's security comes from the fact that our model was developed taking into account pure quantum token states, without considering the intrinsic statistical mixed nature of the ensemble token states. Although this approximation is largely valid for the five IBMQ analyzed here, given the high agreement between the theory and experiments, this can potentially not be the case for other ensemble systems as well. In these cases, the density matrix formalism would need to be used to describe the tokens, with extra variables for the Bloch vector length and uncertainties of the angles.

## Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/lucas-tsunaki/quantum-token.

## Acknowledgments

This work was supported by the German Federal Ministry of Education and Research (BMBF) under the project '*DIamant-basiert QuantenTOKen*' (DIQTOK—nº 16KISQ034).

## Author contributions

All authors participated in the discussion, writing and revision of the manuscript. L T conceptualized, performed and analyzed the experiments. L T, B B, M X and K S theoretically developed the protocol. M E G, K S and B N conceptualized the quantum token device, acquired funding and supervised the work.

## Appendix A. Mathematical proofs

**Proof of equation (1).** A quantum state with arbitrary angles $\theta$ and $\phi$ is described by

$$|\Psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle.$$

The expectation value of $\langle\hat{N}\rangle = \langle\Psi|\hat{N}|\Psi\rangle$ is written as

$$\langle\hat{N}\rangle = \begin{bmatrix}\cos\left(\frac{\theta}{2}\right)\\ e^{i\phi}\sin\left(\frac{\theta}{2}\right)\end{bmatrix}^{T}\begin{bmatrix}N_0 & 0\\ 0 & N_1\end{bmatrix}\begin{bmatrix}\cos\left(\frac{\theta}{2}\right)\\ e^{i\phi}\sin\left(\frac{\theta}{2}\right)\end{bmatrix}.$$

Which results in

$$\langle\hat{N}\rangle = N_0\cos^2\left(\frac{\theta}{2}\right) + N_1\sin^2\left(\frac{\theta}{2}\right).$$

□

**Proof of equation (2).** Regarding the observable uncertainty $\sigma_N$, we first need to calculate the $\langle\hat{N}^2\rangle$ term as

$$\langle\hat{N}^2\rangle = \begin{bmatrix}\cos\left(\frac{\theta}{2}\right)\\ e^{i\phi}\sin\left(\frac{\theta}{2}\right)\end{bmatrix}^{T}\begin{bmatrix}N_0^2 & 0\\ 0 & N_1^2\end{bmatrix}\begin{bmatrix}\cos\left(\frac{\theta}{2}\right)\\ e^{i\phi}\sin\left(\frac{\theta}{2}\right)\end{bmatrix}.$$

Which results in

$$\langle\hat{N}^2\rangle = N_0^2\cos^2\left(\frac{\theta}{2}\right) + N_1^2\sin^2\left(\frac{\theta}{2}\right).$$

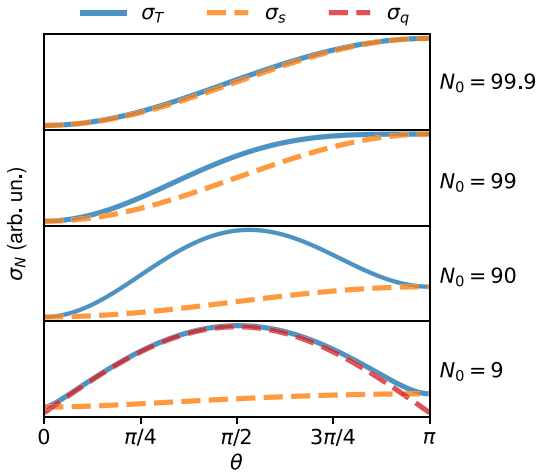


**Figure 8.** $\sigma_N$ for different values of $N_0$ with fixed $N_1 = 100$ and $\sigma_{\text{exp}} = 0$. At large $N_0$, the noise is completely dominated by shot noise, which is larger at $\theta = \pi$. On the other hand, as the difference between $N_1$ and $N_0$ increases, the quantum uncertainty becomes the dominating term, being higher at $\theta = \pi/2$.

Thus, adding all terms together we get

$$\begin{aligned}\sigma_N^2 &= \sigma_q^2 + \sigma_n^2 + \sigma_{\text{exp}}^2 \\ &= \langle \hat{N}^2 \rangle - \langle \hat{N} \rangle^2 + \langle \hat{N} \rangle + \sigma_{\text{exp}}^2 \\ &= N_0^2 \cos^2\left(\frac{\theta}{2}\right) + N_1^2 \sin^2\left(\frac{\theta}{2}\right) \\ &\quad - \left[N_0 \cos^2\left(\frac{\theta}{2}\right) + N_1 \sin^2\left(\frac{\theta}{2}\right)\right]^2 \\ &\quad + N_0 \cos^2\left(\frac{\theta}{2}\right) + N_1 \sin^2\left(\frac{\theta}{2}\right) + \sigma_{\text{exp}}.\end{aligned}$$

Rearranging them results in

$$\begin{aligned}\sigma_N^2 = N_0 \cos^2\left(\frac{\theta}{2}\right)(1 + N_0) + N_1 \sin^2\left(\frac{\theta}{2}\right)(1 + N_1) \\ - \left[N_0 \cos^2\left(\frac{\theta}{2}\right) + n_1 \sin^2\left(\frac{\theta}{2}\right)\right]^2 + \sigma_{\text{exp}}^2.\end{aligned}$$

Solutions of $\sigma_N$ for different combinations of $N_0$ and $N_1$ are shown in figure 8 □

**Proof of equation (4).** Now considering the measurement performed by an attacker, we first calculate the final state after a rotation of the bank with $\theta_b$ and $\phi_b$, followed by an attacker rotation by $\theta_a$ and $\phi_a$. This gives

$$\begin{aligned}|\Psi_f\rangle &= \hat{R}^{-1}(\theta_a, \phi_a)\,\hat{R}(\theta_b, \phi_b)\,|0\rangle \\ &= \begin{bmatrix} \cos\left(\frac{\theta_a}{2}\right) & i\sin\left(\frac{\theta_a}{2}\right) e^{-i\phi_a} \\ i\sin\left(\frac{\theta_a}{2}\right) e^{i\phi_a} & \cos\left(\frac{\theta_a}{2}\right) \end{bmatrix} \\ &\quad \times \begin{bmatrix} \cos\left(\frac{\theta_b}{2}\right) & -i\sin\left(\frac{\theta_b}{2}\right) e^{-i\phi_b} \\ -i\sin\left(\frac{\theta_b}{2}\right) e^{i\phi_b} & \cos\left(\frac{\theta_b}{2}\right) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \cos\left(\frac{\theta_a}{2}\right) & i\sin\left(\frac{\theta_a}{2}\right) e^{-i\phi_a} \\ i\sin\left(\frac{\theta_a}{2}\right) e^{i\phi_a} & \cos\left(\frac{\theta_a}{2}\right) \end{bmatrix} \begin{bmatrix} \cos\left(\frac{\theta_b}{2}\right) \\ -i\sin\left(\frac{\theta_b}{2}\right) e^{i\phi_b} \end{bmatrix} \\ &= \begin{bmatrix} \cos\left(\frac{\theta_a}{2}\right)\cos\left(\frac{\theta_b}{2}\right) + \sin\left(\frac{\theta_a}{2}\right)\sin\left(\frac{\theta_b}{2}\right) e^{i(\phi_b - \phi_a)} \\ i\sin\left(\frac{\theta_a}{2}\right)\cos\left(\frac{\theta_b}{2}\right) e^{i\phi_a} - i\sin\left(\frac{\theta_b}{2}\right)\cos\left(\frac{\theta_a}{2}\right) e^{i\phi_b} \end{bmatrix},\end{aligned}$$

where we use the rotation operator as defined in [37]. The expectation value of $\hat{N}$ for this state is

$$N_0 \left| \cos\left(\frac{\theta_a}{2}\right)\cos\left(\frac{\theta_b}{2}\right) + \sin\left(\frac{\theta_a}{2}\right)\sin\left(\frac{\theta_b}{2}\right) e^{i(\phi_b-\phi_a)} \right|^2 + N_1 \left| \sin\left(\frac{\theta_a}{2}\right)\cos\left(\frac{\theta_b}{2}\right) e^{i\phi_a} - \sin\left(\frac{\theta_b}{2}\right)\cos\left(\frac{\theta_a}{2}\right) e^{i\phi_b} \right|^2 .$$

Using basic trigonometric relations for $\cos(2x)$ and $\sin(2x)$, the term with $N_0$ gives

$$\cos^2\left(\frac{\theta_a}{2}\right)\cos^2\left(\frac{\theta_b}{2}\right) + \sin^2\left(\frac{\theta_a}{2}\right)\sin^2\left(\frac{\theta_b}{2}\right) + \cos\left(\frac{\theta_a}{2}\right)\cos\left(\frac{\theta_b}{2}\right)\sin\left(\frac{\theta_a}{2}\right)\sin\left(\frac{\theta_b}{2}\right) 2\cos(\phi_b-\phi_a)$$
$$= \frac{1}{2}\left[1 + \cos\theta_a\cos\theta_b + \sin\theta_a\sin\theta_b\cos(\phi_b-\phi_a)\right],$$

while the term with $N_1$ yields

$$\sin^2\left(\frac{\theta_a}{2}\right)\cos^2\left(\frac{\theta_b}{2}\right) + \sin^2\left(\frac{\theta_b}{2}\right)\cos^2\left(\frac{\theta_a}{2}\right) - \sin\left(\frac{\theta_a}{2}\right)\cos\left(\frac{\theta_b}{2}\right)\sin\left(\frac{\theta_b}{2}\right)\cos\left(\frac{\theta_a}{2}\right) 2\cos(\phi_b-\phi_a)$$
$$= \frac{1}{2}\left[1 - \cos\theta_a\cos\theta_b - \sin\theta_a\sin\theta_b\cos(\phi_b-\phi_a)\right].$$

Adding both terms together we get

$$\frac{N_0+N_1}{2} + \frac{N_0-N_1}{2}\left[\cos\theta_a\cos\theta_b + \sin\theta_a\sin\theta_b\cos(\phi_b-\phi_a)\right].$$

Finally, the fraction of qubits measured in the $|0\rangle$ state by the attacker $n_a$ will be

$$1 - n_a = \frac{\langle\Psi_f|N|\Psi_f\rangle}{N_0+N_1},$$

which leads to

$$2n_a = 1 + \frac{N_1-N_0}{N_0+N_1}\left[\cos\theta_a\cos\theta_b + \sin\theta_a\sin\theta_b\cos(\phi_b-\phi_a)\right].$$

□

**Proof of $\bar{n}_a = 1/2$.** To get the average value of $\bar{n}_a$, we integrate equation (4) over the whole Bloch sphere

$$2\bar{n}_a = \int_0^{2\pi} d\phi_b \int_0^{\pi} \frac{\sin\theta_b}{2} d\theta_b \left\{1 + c\left[\cos\theta_a\cos\theta_b + \sin\theta_a\sin\theta_b\cos(\phi_b-\phi_a)\right]\right\},$$

The first term simply gives 1, while the two other result in

$$2\bar{n}_a = 1 + \frac{c}{2}\cos\theta_a\int_0^{\pi}\sin\theta_b\cos\theta_b d\theta_b + \frac{c}{2}\sin\theta_a\int_0^{\pi}\sin^2\theta_b d\theta_b\int_0^{2\pi}\cos(\phi_b-\phi_a)d\phi_b.$$

It is straight forward to see that both integrals in the second and third term result in 0. Thus, the final average value of $n_a$ is simply

$$\bar{n}_a = \frac{1}{2}.$$

□

**Interval of valid solutions for $\theta_f$.** In order for the parametric relation from equation (4) to have real solution for $\phi_f$ we need that the argument inside the inverse cosine function is in the interval $[-1,1]$. Defining $\alpha \equiv (2n_a-1)/c$, this results in

$$(\alpha - \cos\theta_a\cos\theta_f)^2 < (\sin\theta_a\sin\theta_f)^2$$
$$\alpha^2 - 2\alpha\cos\theta_a\cos\theta_f + \cos^2\theta_a\cos^2\theta_f - \sin^2\theta_a\sin^2\theta_f < 0.$$

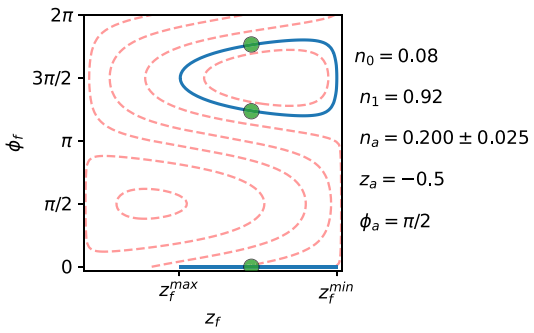


**Figure 9.** Solution for forger angles $\theta_f$ and $\phi_f$ with fixed $n_a$ inside the interval $0.200 \pm 0.025$, depicting how the attacker can forge a fake token based on their measurement of the bank token. First, the attacker calculates the interval solution for $z_f$ and then randomly takes a value inside the interval. Last, they solve equation (4) for $\phi_f$ and choose one of the two possible solutions.

Again, using the $\cos(2x)$ relation we have

$$\alpha^2 - 2\alpha\cos\theta_a\cos\theta_f + \frac{\cos 2\theta_a}{2} + \frac{\cos^2\theta_f - \sin^2\theta_f}{2} < 0.$$

Now we can eliminate the dependency on $-\sin^2\theta_f$ substituting it by $\cos^2\theta_f - 1$ and get to a quadratic equation for $z_f = \cos\theta_f$ as

$$z_f^2 + z_f(-2\alpha\cos\theta_a) + \left(\alpha^2 + \frac{\cos 2\theta_a}{2} - \frac{1}{2}\right) > 0.$$

Using a simple quadratic formula solution this results in interval of solution for $z_f$ of

$$\begin{aligned} \max\left\{\alpha\cos\theta_a - \sqrt{\Delta}, -1\right\} &\leqslant z_f \\ &\leqslant \min\left\{\alpha\cos\theta_a + \sqrt{\Delta}, 1\right\}, \end{aligned} \tag{A1}$$

with

$$\Delta = \alpha^2\cos^2\theta_a - \alpha^2 - \frac{\cos 2\theta_a}{2} + \frac{1}{2}.$$

The solution interval for $z_f$ is depicted in figure 9 □

## Appendix B. Experimental methods

To fully benchmark the protocol and validate our model, measurements were performed in five distinct IBMQ of the Eagle family: Sherbrooke, Kyiv, Brisbane, Kyoto (retired) and Osaka (retired). All of them have similar superconducting architecture [27–29], with 127 physical qubits and the same native operations. Still, their exact specifications are not publicly available. Comparatively, Sherbrooke presented the longest mean coherence time $T_2$, while Brisbane had the smallest gate error and Kyiv the smallest readout error, based on IBM calibration reports. Kyoto, on the other hand, had the worse parameters of the five. Small changes in these performance parameters were observed during the experiments conducted from April 2024 to December 2024, but without much implication to the results here presented. Due to high memory costs for waveform generation and increased gate and readout errors, an ergodic approximation was used for the qubit ensemble, where a single qubit was averaged over time instead of using multiple qubits at once. The use of such an ergodic approximation is further validated by the assumption of the inseparability of the ensemble.

Experiments were run in the hardware through Qiskit software development tool [26] in Python, version 1.1.1 and older. All Qiskit codes used for measurements and data analysis are open-source, provided at the author's GitHub repository [30]. The high-level representation of the quantum circuits were defined with the Quantum Circuit class. Which were then transpiled to the hardware's native gates without gate approximation, using the qubit with longest $T_2$ and shortest gate operation, while considering the specific timing constrains of each quantum hardware. Finally, the protocols were run with Qiskit Runtime, while the corresponding observable $\hat{N}$ expectation values are taken from the job counts in the $|0\rangle$ and $|1\rangle$ states.

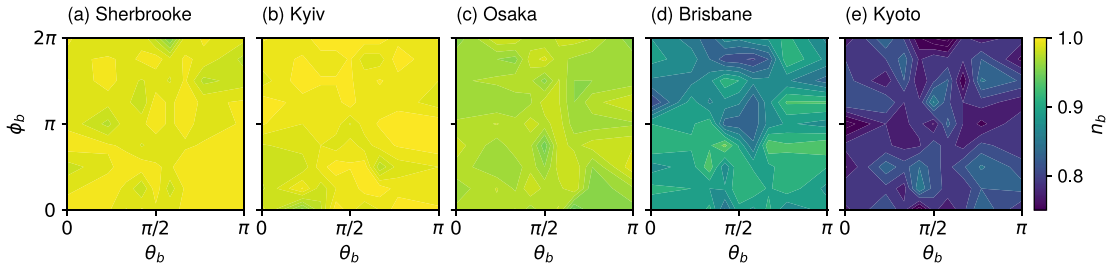


**Figure 10.** Angle dependence of the self-acceptance fraction $n_\mathrm{b}$ for IBM quantum platforms (a) Sherbrooke, (b) Kyiv, (c) Osaka, (d) Brisbane and (e) Kyoto. In hardware with unoptimized control and short coherence times, smaller $n_\mathrm{b}$ would be expected for larger angles due to prolonged pulses. Which is not the case here for the IBMQ.

**Table 2.** Glossary of the main variables in the text.

| Variable | Physical meaning |
|---|---|
| $\theta$ | Polar angle on the Bloch sphere |
| $\phi$ | Azimuthal angle on the Bloch sphere |
| $z$ | $\cos\theta$ |
| $\theta_\mathrm{b}, \phi_\mathrm{b}$ | Angles which the bank prepares and measures the token |
| $\theta_\mathrm{a}, \phi_\mathrm{a}$ | Angles used by the attacker to measure the bank token |
| $\theta_\mathrm{f}, \phi_\mathrm{f}$ | Angles forged by the attacker |
| $\hat{N}$ | System's observable (Example: photon count) |
| $N_0$ | $\hat{N}$ eigenvalue for $\|0\rangle$ (Example: dark counts) |
| $N_1$ | $\hat{N}$ eigenvalue for $\|1\rangle$ (Example: bright counts) |
| $c$ | Normalized contrast between $N_1$ and $N_0$ |
| $\sigma_N$ | Total uncertainty of $\hat{N}$ |
| $\sigma_q$ | Heisenberg uncertainty |
| $\sigma_\mathrm{s}$ | Shot noise |
| $\sigma_\mathrm{exp}$ | Experimental noise |
| $\hat{R}(\theta, \phi)$ | Rotation operator with angles $\theta$ and $\phi$ |
| $n$ | Fraction of qubits measured in the $\|0\rangle$ state |
| $n_\mathrm{b}$ | $n$ if the bank prepares and measures without an attacker |
| $n_\mathrm{a}$ | $n$ measured by the attacker with angles $\theta_\mathrm{a}, \phi_\mathrm{a}$ for a bank token |
| $n_\mathrm{f}$ | $n$ measured by the bank for a forged token |
| $n_\mathrm{r}$ | $n$ measured by the bank for randomly forged tokens |
| $n_\mathrm{T}$ | Minimum $n$ threshold for the token to be accepted by the bank |
| $M$ | Number of tokens in the device |
| $p_\mathrm{b}$ | Self-acceptance probability of the bank tokens |
| $p_\mathrm{f}$ | Acceptance probability of forged tokens |
| $\alpha$ | $(2n_\mathrm{a} - 1)/c$ |

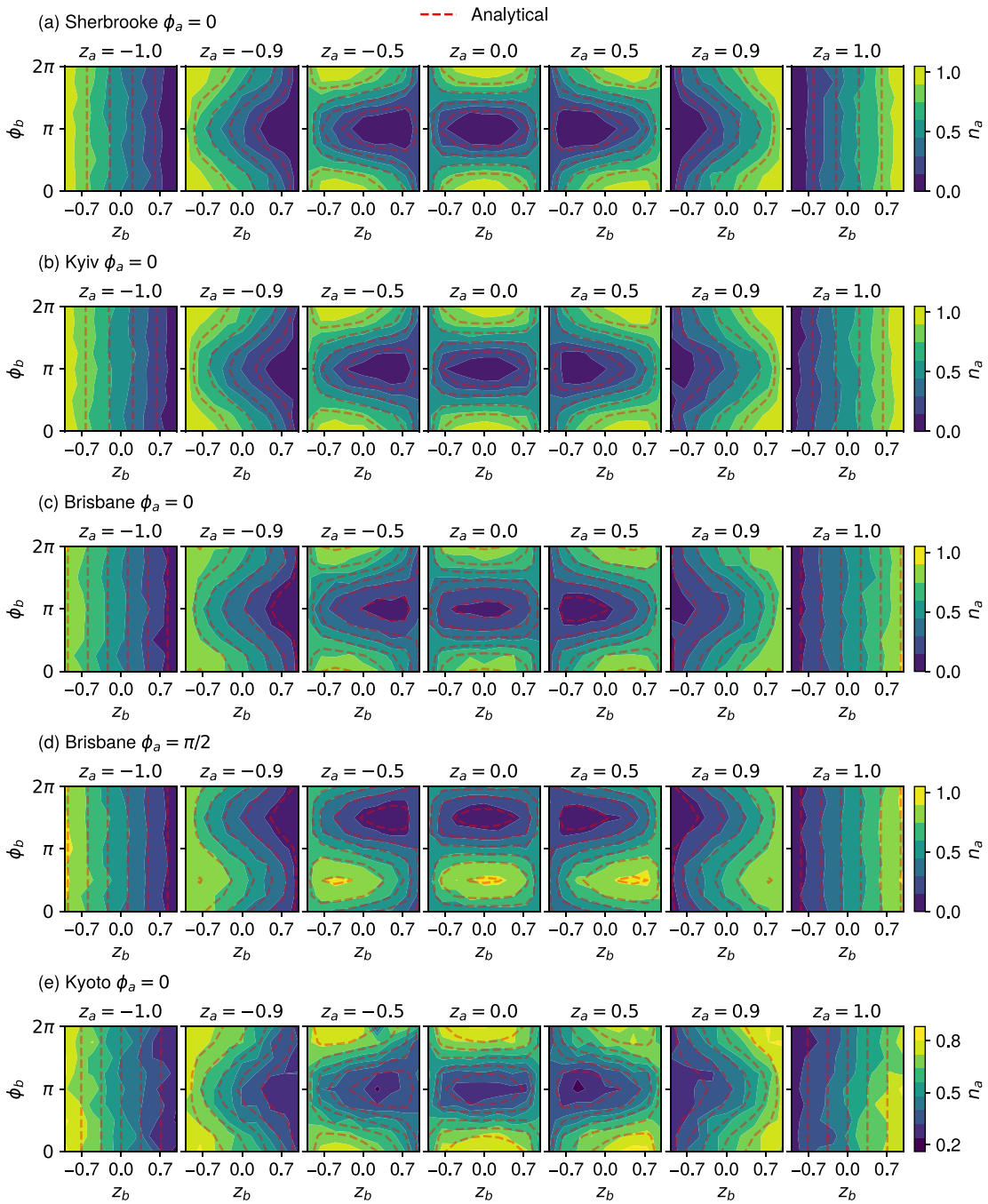


**Figure 11.** Angle dependence of $n_a$ measured with $\phi_a = 0$ in IBM quantum platforms (a) Sherbrooke, (b) Kyiv, (c) Brisbane and (e) Kyoto, in addition to (d) Brisbane measured with $\phi_a = \pi/2$. All of them are compared with the analytical solution from equation (4) of (red dashed lines). All hardware show strong agreement with the analytical formula, except Kyoto which has more noise due to its increased experimental uncertainty, as well as a smaller contrast (table 1). Furthermore, the attacker angle of $\phi_a = \pi/2$ in Brisbane simply shifts the angle dependence of $n_a$ compared to $\phi_a = 0$.

## ORCID iDs

Lucas Tsunaki 0009-0003-3534-6300
Bernd Bauerhenne 0000-0002-3397-2290
Malwin Xibraku 0009-0002-3183-6188
Martin E Garcia 0000-0003-2418-1902
Kilian Singer 0000-0001-9726-0367
Boris Naydenov 0000-0002-5215-3880

## References


[1] Park J L 1970 The concept of transition in quantum mechanics *Found. Phys.* **1** 23
[2] Portmann C and Renner R 2022 Security in quantum cryptography *Rev. Mod. Phys.* **94** 025008
[3] Pastawski F, Yao N Y, Jiang L, Lukin M D and Cirac J I 2012 Unforgeable noise-tolerant quantum tokens *PNAS* **109** 16079
[4] Singer K T T, Popov C and Naydenov B 2022 *Patent de* 10 2022 107 528 a1 2023.10.05: Verfahren zum erstellen eines quanten-datentokens
[5] Pompili M *et al* 2021 Realization of a multinode quantum network of remote solid-state qubits *Science* **372** 259
[6] Pfaff W *et al* 2014 Unconditional quantum teleportation between distant solid-state quantum bits *Science* **345** 532
[7] Tilley R J D 2014 Color centers *Encyclopedia of Color Science and Technology* ed R Luo (Springer) pp 1–9
[8] Pezzagna S and Meijer J 2021 Quantum computer based on color centers in diamond *Appl. Phys. Rev.* **8** 011308
[9] Unden T *et al* 2018 Coherent control of solid state nuclear spin nano-ensembles *npj Quantum Inf.* **4** 39
[10] Torosov B T and Vitanov N V 2022 Experimental demonstration of composite pulses on IBM's quantum computer *Phys. Rev. Appl.* **18** 034062
[11] Wang X, Bishop L S, Kestner J, Barnes E, Sun K and Das Sarma S 2012 Composite pulses for robust universal control of singlet–triplet qubits *Nat. Commun.* **3** 997
[12] Bauerhenne B, Tsunaki L, Thieme J, Naydenov B and Singer K 2025 Advanced attacks on qubit-ensemble based quantum coins *Quantum Sci. Technol.* (https://doi.org/10.1088/2058-9565/ae03e7)
[13] Gruber A, Dräbenstedt A, Tietz C, Fleury L, Wrachtrup J and von Borczyskowski C 1997 Scanning confocal optical microscopy and magnetic resonance on single defect centers *Science* **276** 2012
[14] Jelezko F, Gaebel T, Popa I, Gruber A and Wrachtrup J 2004 Observation of coherent oscillations in a single electron spin *Phys. Rev. Lett.* **92** 076401
[15] Volkova K *et al* 2022 Optical and spin properties of NV center ensembles in diamond nano-pillars *Nanomaterials* **12** 1516
[16] Momenzadeh S A, Stöhr R J, de Oliveira F F, Brunner A, Denisenko A, Yang S, Reinhard F and Wrachtrup J 2015 Nanoengineered diamond waveguide as a robust bright platform for nanomagnetometry using shallow nitrogen vacancy centers *Nano Lett.* **15** 165
[17] Toyli D M, Weis C D, Fuchs G D, Schenkel T and Awschalom D D 2010 Chip-scale nanofabrication of single spins and spin arrays in diamond *Nano Lett.* **10** 3168
[18] Xie T, Zhao Z, Xu S, Kong X, Yang Z, Wang M, Wang Y, Shi F and Du J 2023 99.92%-fidelity CNOT gates in solids by noise filtering *Phys. Rev. Lett.* **130** 030601
[19] Tsunaki L, Singh A, Volkova K, Trofimov S, Pregnolato T, Schröder T and Naydenov B 2025 Ambiguous resonances in multipulse quantum sensing with nitrogen-vacancy centers *Phys. Rev. A* **111** 022606
[20] Maurer P C *et al* 2012 Room-temperature quantum bit memory exceeding one second *Science* **336** 1283
[21] Pfender M *et al* 2017 *Nano Lett.* **17** 5931
[22] Neumann P, Beck J, Steiner M, Rempp F, Fedder H, Hemmer P R, Wrachtrup J and Jelezko F 2010 Single-shot readout of a single nuclear spin *Science* **329** 542
[23] Delgado M M, Tsunaki L, Michaelson S, Kuntumalla M K, Reithmaier J P, Hoffman A, Naydenov B and Popov C 2025 Impact of annealing and nanostructuring on properties of NV centers created by different techniques *Diam. Relat. Mater.* **154** 112126
[24] Bozzio M, Cavaillès A, Diamanti E, Kent A and Pitalúa-García D 2021 Multiphoton and side-channel attacks in mistrustful quantum cryptography *PRX Quantum* **2** 030338
[25] Kwiat P, Weinfurter H, Herzog T, Zeilinger A and Kasevich M A 1995 Interaction-free measurement *Phys. Rev. Lett.* **74** 4763
[26] Javadi-Abhari A *et al* 2024 Quantum computing with qiskit (arXiv:2405.08810)
[27] Kandala A, Wei K X, Srinivasan S, Magesan E, Carnevale S, Keefe G, Klaus D, Dial O and McKay D 2021 Demonstration of a high-fidelity CNOT gate for fixed-frequency transmons with engineered ZZ suppression *Phys. Rev. Lett.* **127** 130501
[28] Bravyi S, Cross A W, Gambetta J M, Maslov D, Rall P and Yoder T J 2024 High-threshold and low-overhead fault-tolerant quantum memory *Nature* **627** 778
[29] Glick J R, Gujarati T P, Córcoles A D, Kim Y, Kandala A, Gambetta J M and Temme K 2024 Covariant quantum kernels for data with group structure *Nat. Phys.* **20** 479
[30] Tsunaki L 2024 Quantum token (available at: https://github.com/lucas-tsunaki/quantum-token)
[31] Callen H B and Welton T A 1951 Irreversibility and generalized noise *Phys. Rev.* **83** 34
[32] Demkowicz-Dobrzański R, Jarzyna M and Kołodyński J 2015 *Chapter Four - Quantum Limits in Optical Interferometry* (*Progress in Optics*) vol 60, ed E Wolf (Elsevier) pp 345–435
[33] Itano W M, Bergquist J C, Bollinger J J, Gilligan J M, Heinzen D J, Moore F L, Raizen M G and Wineland D J 1993 Quantum projection noise: population fluctuations in two-level systems *Phys. Rev. A* **47** 3554
[34] Rabi I I 1937 Space quantization in a gyrating magnetic field *Phys. Rev.* **51** 652
[35] O'hagan A and Leonard T 1976 Bayes estimation subject to uncertainty about parameter constraints *Biometrika* **63** 201
[36] Waldherr G *et al* 2014 Quantum error correction in a solid-state hybrid spin register *Nature* **506** 204
[37] Oliveira I, Sarthour R Jr, Bonagamba T, Azevedo E and Freitas J C 2007 *NMR Quantum Information Processing* (Elsevier)
[38] Maletinsky P, Hong S, Grinolds M S, Hausmann B, Lukin M D, Walsworth R L, Loncar M and Yacoby A 2012 A robust scanning diamond sensor for nanoscale imaging with single nitrogen-vacancy centres *Nat. Nanotechnol.* **7** 320
[39] Dolde F, Jakobi I, Naydenov B, Zhao N, Pezzagna S, Trautmann C, Meijer J, Neumann P, Jelezko F and Wrachtrup J 2013 Room-temperature entanglement between single defect spins in diamond *Nat. Phys.* **9** 139